

JAMIT Annual Meeting 2021 第40回日本医用画像工学会大会

- 会期:2021年10月13日(水)~10月15日(金) ハンズオンセミナー:10月16日(土)
- 会 場:慶應義塾大学日吉キャンパス協生館/オンライン
- 大会長: 陣崎 雅弘(慶應義塾大学)



主	催	一般社団法人 日本医用画像工学会	
		(The Japanese Society of Medical I	Imaging Technology)
後	援	一般社団法人 日本画像医療システム	工業会
		(JIRA:Japan Medical Imaging and R	adiological Systems Industries Association)
		一般社団法人 保健医療福祉情報シス	マテム工業会
		(JAHIS: Japanese Association of Heal	Ithcare Information Systems Industry)
		一般社団法人 雷子情報技術産業協会	슈
		(JEITA: Japan Electronics and Informatic	n Technology Industries Association)
協	替	医用画像情報学会	医用画像認知研究会
		应用物理学会	可視化情報学会
		画像雷子学会	5.1%(1)(4)(4)(4) 着護理工学会
		三次元画像コンファレンス宝行委員会	情報処理学会
		雪子情報通信受 会	日本医学物理学会
		电子中型位于公司	口本区于初建于五日本区于杨建于五日本区于杨建于五日本区于杨建于五日本区于杨建于五日本区于杨建于五日本区于杨建于五日本区于杨建于五日本区于五日本区于五日本区于五日本区于五日本区于五日本区于五日本区于五日本区
		口本区于成初旅子云	口本应该有報子云 日本枝库学技術学会
			口午起日叔区子云
		日本脳神栓し子云	口平瓜射标拉帕子云

学会長ご挨拶

第40回日本医用画像工学会大会(JAMIT 2021)を、会期10/13(水)~10/15(金)にわたって慶應義塾大学日吉キャンパスの協 生館とオンラインのハイブリッド開催で、陣崎雅弘先生大会長のもとで開催させていただくことになりました。COVID-19や東京オリ ンピック開催がどうなるか全く予測できない状況のもと、早めに10月開催と現地・オンラインを併用したハイブリッド開催を決定して、 これまで陣崎先生を中心に精力的に準備を進めてきました。一年以上前から準備を進めて来られた陣崎先生他大会関係者の皆様方に厚 くお礼を申しあげます。

さて、JAMIT 大会は第35回大会(JAMIT2016)頃から AI ブームの上昇気流に乗り、更にはチュートリアル講演、ハンズオンセミ ナー、特別講演、シンポジウムに魅力的なものを準備して、毎年大成功を収めて来ました、参加登録者の総数が近年は毎年300人を越 えており JAMIT 大会への医用画像関係者の大きな期待が感じとれます。今回の大会のテーマは「新たな潮流を目指して」に設定され ており、AI に関する充実した企画を残しつつ近未来における医用画像工学の新しい展開が感じとれる盛り沢山の内容になっています。 企画を中心に簡単にその一端を紹介させていただきます。特別講演は、大変著名な革新的なバイオイメージングに関する講演を運化学 研究所の宮脇敦史先生に、従来のコンピュータとは原理が異なる量子コンピュータで何ができどう使うかに関する講演を慶應義塾大学 の伊藤公平先生にしていただけることが決まりました。また、AI に関しては、例年好評を博しており JAMIT 大会に定着したハンズオ ンセミナーに加え、AI 技術が完成に近づいた現状をふまえ次の時期を意識したシンポジウム「AI 教育を考える」、AI の臨床応用に関 する特別企画「画像 AI 最前線」を準備しています。JAMIT お家芸のイメージングに関しては、大会長の陣崎先生も先駆的に取り組 んでいるテーマの「立位のイメージング」、空間分解能の極限を目指す「画像で微細構造がどこまで見えるか」、「脳のイメージングの 最前線」を準備しています。一般演題は3月に演題募集を開始して、AI に関する発表を中心に約100 件の投稿があると期待しており、 JAMIT2018, JAMI2019 に引き続きポスターセッション (ティザー+ポスター)の形態で実施します。JAMIT 会員の方も非会員の方も、 積極的に演題の応募をお願いします。最後に、昨年度はオンライン開催のため実施できなかったランチョンセミナーも会場で実施する 予定です。

さて, JAMIT は念願であった任意団体から法人への移行を2020年12月18日に達成して, JAMIT2021が法人化後に開催される初 めての大会になります.法人化の際に定款や規程を大幅に見直しして,今まで中日に実施していた「総会」を「会員集会」という形で, 表彰を「田中栄一記念賞」に加え「功労賞」と「功績賞」の表彰を毎年行うことになっています. それ以外にも法人化により JAMIT 大会の様相が変わる部分もあると予想され,新しい法人 JAMIT の大会を是非覗いてみてください.

最後に, JAMIT2021 はハイブリッド開催ですが, 現地会場の慶應義塾大学日吉キャンパス協生館周辺は, 大会 HP の画像にあるように 11 月には銀杏(いちょう)並木が大変きれいな場所と聞いています. 10 月開催ですから銀杏が黄色く色づくにはまだ早い時期と思われますが, COVID-19 が 10 月までに収まり安全が回復して現地に多くの方が集まっての JAMIT 大会になることを心から祈っています.

日本医用画像工学会 会長 工藤 博幸(筑波大学)

大会長ご挨拶

第40回日本医用画像工学会(JAMIT)大会を2021年10月13~15日に慶應義塾大学日吉キャンパスの協生館にて開催させて頂くことになりました.大会のテーマは、「新たな潮流を目指して」に致しました.

HPに掲載しました写真は、日吉キャンパスの正面入り口から写したものです。秋の開催ということで、11月に見頃を迎える紅葉した銀杏並木に致しました。この坂道を上り切った奥に昨年建て替えたばかりの新「日吉記念館」が見えています。新たな潮流を目指して坂道を上っていったら、新しい世界が見えてきたというイメージでこの写真を選びました。

特別講演は先駆的に時代を切り拓くお仕事をされているお二人にお願い致しました。お一人目は、本学理工学部物理情報工学科教授の伊藤公平先生です。今年、大学の学長に相当する塾長に就任されました。量子コンピュータの第一人者であられ、「量子コンピュータをどう使いこなすか」というタイトルで、次世代解析法の可能性と課題をお話して頂きます。伊藤先生は、私にとっては医工連携のシンボル的存在でいらっしゃり、私も多岐にわたって連携させて頂いております。お二人目は、今年の3月に日本学士院賞を受賞された理化学研究所の宮脇敦史先生です。医用画像工学の目標は人体を切らずに更なる可視化を進めていくことだと思います。現在、3次元4次元像で組織レベルの可視化が行われているところですが、究極は人の細胞内動態の可視化だと思っています。宮脇先生は動物を用いて細胞内動態の可視化を探究しておられます。私とは大学時代の同期で、共にこのキャンパスで学び、語り合った仲です。

シンポジウムは、新たな切り口での話題を2つほど設けました。1つ目は、これまでもこの学会の主要テーマとして取り上げられて きた人工知能関連ですが、ソフト開発や手法の話ではなく、これからの時代を担う AIの人材育を取り上げました。「AI 教育を考える」 というテーマで、どのように AI 人材育成を行っていくのかを、昨年文科省の医療データ人材育成拠点形成事業に採択された、東北大 学と名古屋大学の先生方にもご登壇頂き, 議論したいと思います。2つ目は、「横断画像で微細構造がどこまで見えるか」というテーマで、 超音波、CT、MRI、PET などの様々なモダリティーを一同に並べて、最新の分解能でそれぞれどこまで見えるようになっているかと いうことを総合的に聞いて頂きたいと思います。特別講演の宮脇先生の、細胞内動態を可視化する技術の分解能との違いを感じて頂け ればと思います。

特別企画は、機能評価という視点からの話題を2つ提供したいと思います.まず、「重力下の人体を可視化する~立位のイメージン グ~」というタイトルで、私が推進している立位での機能評価の現状を取り上げました.これまで画像の多くは臥位で撮影され、器質 的疾患の評価に用いられて、生命寿命の延伸に役立ってきました.これからは、立位で人体機能を評価することにより、健康寿命の延 伸に役立つものになっていくと考えています.もう1つは、「脳のイメージング最前線」と題して、今多くの画像研究者の興味を強く 惹いていると思われる脳の機能,構造,連絡性,すなわち「コネクトーム」と、老廃物排泄系である「Glymphatic システム」の話題です. これらの異常は睡眠や痴呆との関連も指摘されており、立位イメージと同様、人体機能評価の観点から大きな潮流の1つになると思っ ています.

この学会は、多くの場合基礎工学系の先生が大会長を担っておられますが、医学系の大会長は2014年の慈恵医科大学放射線科教授(当時)の福田国彦先生以来7年振りになります.放射線科医の立場から、臨床に近い話題も多く聞いて頂けるように致しました.いずれのトピックも新たな潮流と思えるようなものを選定致しましたが、この大会を通して、医用画像工学の今後の方向性の一端を感じてもらうことができればと思います.

すっかり定着した,チュートリアル講演会,深層学習ハンズオンセミナーは例年通り開催させて頂きますので,奮って応募ください. コロナ禍でハイブリッド開催を予定しており,現地参加もWEB参加も可能ですので,多くの方にご参加頂けますことを願っております.

第40回日本医用画像工学会大会

大会長 陣崎 雅弘 (慶應義塾大学)



慶應義塾大学日吉キャンパス協生館 (神奈川県横浜市港北区日吉 4-1-1)

【交通アクセス】

東急東横線・東急目黒線・横浜市市営地下鉄グリーンライン 日吉駅下車(徒歩1分)
● 渋谷 - 日吉:約25分(急行約20分)
● 横浜 - 日吉:約20分(急行約15分)

- 新横浜 菊名 日吉: 20 分
- ※ 日吉駅に停車するのは「普通」および「急行」電車です。「特急」は停まりません。

協生館 2F



参加者へのご案内

- 1. 会期: 2021年10月13日(水)~10月15日(金) ハンズオンセミナー: 10月16日(土)
- 2. 会場:慶應義塾大学 日吉キャンパス 協生館 及び オンライン (Zoom ウェビナー使用) 〒223-8526 神奈川県横浜市港北区日吉 四丁目 1番 1号 TEL: 045-564-2500 (内線) 37115 (協生館運営センター) <u>http://www.kcc.keio.ac.jp/index.html</u> 講演会場: 2階 藤原洋記念ホール

3. 参加登録

参加登録ページよりオンラインにて参加登録を完了してください.現地参加される際は,参加費の決 済完了メールをプリントアウトして受付までご持参ください.

http://jamit2021.jamit.jp/registration.html

- 1) 登録期間:7月1日(木)~10月15日(金)16:00
- 2)参加費:正会員(賛助会員含む)13,000円 学生会員5,000円 非会員20,000円

4. 総 会

日時:10月14日(木) 13:00~14:00

5. 予稿集について

ダウンロード (PDF) 形式で作成し、ダウンロードパスワードを参加者にお知らせいたします.

6. 発表者資格

一般演題の筆頭発表者は、日本医用画像工学会の会員に限ります.非会員の方は必ず8月末日までに 入会のお申込みをいただき、9月20日(火)までに会費を納入してください.

7. 入会手続き及びお問い合わせ

日本医用画像工学会事務局(<u>http://www.jamit.jp/</u>)
〒104-0033 東京都中央区新川 1-5-19 6 階 (株) メイプロジェクト内
TEL: 03-6264-9071 FAX: 03-6264-8344 E-mail: jamit@may-pro.net

8. 利益相反(COI) について

筆頭発表者が,日本医用画像工学会における「利益相反の取扱いに関する規程」 (<u>http://www.jamit.jp/outline/agreement/rieki.html</u>)において開示の対象となる場合,この規定に 基づいて,発表者の申告した利益相反の状態が予稿集に開示されます.

9. 特別講演

特別講演1:10月14日(木) 10:50~11:40

量子コンピュータをどう使いこなすか

伊藤 公平 (慶應義塾大学)

座 長:工藤 博幸(筑波大学/日本医用画像工学会 会長)

特別講演2:10月15日(金) 11:00~11:50

革新的バイオイメージング技術

宮脇 敦史(理化学研究所)

座 長:陣崎 雅弘 (慶應義塾大学/第40回日本医用画像工学会大会 大会長)

10. 第 11 回 JAMIT チュートリアル講演会(教育委員会企画)

(コニカミノルタ科学技術振興財団 JAMIT ハンズオンセミナー 連携企画)

「GitHub とその周辺:大学と企業でのプロジェクト管理」

第40回日本医用画像工学会大会のプログラムの一環として「第11回日本医用画像工学会(JAMIT) チュートリアル講演会~GitHub とその周辺:大学と企業でのプロジェクト管理~」を第1日目の午 後に開催いたします.チュートリアルでは,医用画像工学に携わる若手研究者や最新の動向を得たい 第一線の研究者を対象として,現在の研究に役立つテーマを専門の研究者が講演します.

日 時:10月13日(水) 13:10~16:00

座 長:木戸 尚治(大阪大学)/滝沢 穂高(筑波大学)

講演1 未定

中田 典生 (東京慈恵会医科大学)

講演2 バイオインフォマティクス分野のアプリケーション開発における GitHub の活 用とその実際

小野 浩雅(情報・システム研究機構)

講演3 医療機器開発企業における共同研究のための環境と実際の運用について(事例 を交えて)

大村 和元 (GE ヘルスケア・ジャパン株式会社)

11. コニカミノルタ科学技術振興財団 JAMIT ハンズオンセミナー

深層学習ハンズオンセミナー~再スタート・深層学習~

(概要)深層学習の実行環境は、比較的簡単に構築できる.本年のハンズオンは PyTorch を利用する.参加者は.事前に提供されるセットアップ資料に基づいて、自分自身のコンピュータに PyTorch / Python による実行環境を構築し、その上で5つの課題について、医用画像を利用した基本的な課題を実行する.また、必要に応じて DICOM ファイルの読み込みなどの前処理の方法や、構築したモデルの保存/読み込みといった実践的な方法についても触れる.そして、課題に応じた評価方法についても考え方と処理方法を実践的に習得する.

(本年の特徴)本ハンズオンは、コニカミノルタ科学技術振興財団助成事業として、初心者や初学 者である学生/若手研究者の支援を目的として実施する.学会期間中5回(各回60分程度)のセ ミナーを実施する.5回の内容は以下の特徴がある.小さめのデータベースと取り組みやすい内容 を用意する.

- 第1回 入門編:環境構築から自然画像の分類
- 第2回 画像の分類と回帰:自分で構築するデータベースの利用
- 第3回 領域抽出と領域分割:物体検出とセマンティックセグメンテーション
- 第4回 異常検知: AutoEncoder による教師なし学習入門
- 第5回 画像生成:フェイク画像の作成入門

(開催形態・会場) ハンズオンセミナーは,大会サテライトイベントとしてオンライン開催で実施し ます.事前に公開するセットアップ資料と題材に基づいて,セミナー開催までにご自身の PC に環 境を構築してください.なお,オンライン参加のためには Zoom による接続が必要です.

(参加登録) ハンズオンの参加は、ハンズオン専用ウェブページから可能です(9月15日開始).

http://micv.sakura.ne.jp/jamitho/

なお,JAMIT2021 大会への事前登録が必要です.JAMIT2021 大会の参加登録がない方は,ハンズオン に参加できません.JAMIT2021 大会参加者は,ハンズオンに無料で参加できます.

(登録後のスケジュール) 10月 1日 セットアップ資料公開

10月 8日 コース資料の公開

10月16日 セミナー開催(オンライン開催)

- (各回の予定:60分)
 00:00~00:05 達成目標の確認
 00:05~00:10 プログラム実行のデモ
 00:10~00:20 自分のPCでの実行確認
 00:20~00:30 スクリプトの説明
 00:30~00:40 結果の確認/評価方法
 00:40~00:50 自分のPCでの実行結果確認
 00:50~01:00 問題点や限界の説明とまとめ
 原 武史(岐阜大学)
 李 鎔範(新潟大学)
 中田 典生(東京慈恵会医科大学)
 - 小田 昌宏(名古屋大学)
 - 滝沢 穂高(筑波大学)

(協力)本セミナーは、コニカミノルタ科学技術振興財団の支援を得て実施しています.また、岐 阜大学人工知能研究推進センター、東海国立大学機構医療健康データ統合研究教育拠点の協力で実 施します.

12. シンポジウム

シンポジウム1:10月14日(木) 15:50~17:20

AI 教育を考える

座 長:橋本 正弘 (慶應義塾大学) / 森 健策(名古屋大学) 講演1 デザイン思考にもとづく医療 AI 研究のための人材育成 植田 琢也(東北大学)

6

- 講演2 医療 AI 人材教育における現状と情報学研究者としての展望 森 健策(名古屋大学)
- 講演3 国際競争に打ち勝つ AI 人材を育成するために何が必要か? 鈴木 賢治(東京工業大学)
- 講演4 CG・画像処理分野での深層学習の実践的教育 金森 由博(筑波大学)
- シンポジウム2:10月15日(金) 13:00~14:30
 - 横断画像で微細構造がどこまで見えるか
 - 座 長:清水 昭伸(東京農工大学)/花岡 昇平(東京大学)
 - 講演1 光超音波イメージングによる表在リンパ管の観察梶田 大樹 (慶應義塾大学)
 - 講演2 超音波 SMI による微細血管の描出

自 二郎(川崎医科大学)

- 講演3 高精細 CT 岩澤 多恵(神奈川県立循環器呼吸器病センター)
- 講演4 7 テスラ MRI 画像による微細脳構造評価
 - 岡田 知久(京都大学)
- 講演 5 デジタル(半導体検出器)PET
 - 堀越 浩幸(群馬県立がんセンター)

13. 特別企画

特別企画1:10月14日(木) 14:10~15:40

重力下の人体を可視化する~立位のイメージング~

- 座 長:佐藤 嘉伸(奈良先端科学技術大学院大学)/陣崎 雅弘(慶應義塾大学)
- 講演1 X線動態撮影
 - 田中 利恵(金沢大学)
- 講演2 立位 CT の開発と応用~健康長寿の時代に向けて~
 - 山田 祥岳 (慶應義塾大学)
- 講演3 Gravity MRI による多姿勢イメージング
 - 宮地 利明(金沢大学)
- 講演4 挑戦する PET:世界初のヘルメット型 PET は座位ブームを起こすか? 山谷 泰賀(量子科学技術研究開発機構)
- 特別企画 2:10 月 15 日(金) 14:40~16:10

脳のイメージング最前線

- 座 長: 增谷 佳孝 (広島市立大学)
- 講演1 T2 解析による Glymphatic pathway の可視化押尾 晃一(順天堂大学)
- 講演2 拡散画像で見る間質液動態
 - 田岡 俊昭(名古屋大学)

講演3 非侵襲画像技術による霊長類脳コネクトーム解明研究

林 拓也(理化学研究所)

14. スポンサードセミナー

スポンサードセミナー1:10月14日(木) 11:50~12:20

Amazon SageMaker を使って医用画像の機械学習をはじめる方法

- 橋本 正弘 (慶応義塾大学)
- 今井 真宏 (アマゾンウェブサービスジャパン株式会社)

座 長:中田 典生 (東京慈恵会医科大学)

スポンサードセミナー2:10月14日(木) 12:20~12:50

NVIDIA における医療 AI の取り組み - 医用画像 AI を加速する Clara Imaging のご紹介-

阮 佩穎(エヌビディア合同会社)

座 長:中田 典生 (東京慈恵会医科大学)

スポンサードセミナー3:10月15日(金) 12:00~12:50

薬事承認された画像用 AI ソフト

- 座 長:木戸 尚治(大阪大学)
- 講演1 富士フイルムの AI 技術「REiLI」の現状-最新の取り組みのご紹介-久永 隆治(富士フイルム株式会社)
- 講演2 Siemens Healthineers の AI 技術を応用した画像診断支援サービスの特長と活 用事例

岩田 和浩 (シーメンスヘルスケア株式会社)

- 講演3 画像診断における AI の実装へ向けた課題と取り組み 大越 厚(GE ヘルスケア・ジャパン株式会社)
- 講演4 "Abierto Reading Support Solution"画像診断領域における AI 活用の取組み
 渡辺 和之(キヤノンメディカルシステムズ株式会社)

ご協力団体

本大会は下記企業及び団体にご協力いただきました. 心より感謝申し上げます.

- 後 援:一般社団法人 日本画像医療システム工業会
 - 一般社団法人 保健医療福祉情報システム工業会
 - 一般社団法人 電子情報技術産業協会
- 助 成:公益財団法人 コニカミノルタ科学技術振興財団
- 企業展示:キヤノンメディカルシステムズ株式会社
 コニカミノルタ株式会社
 GE ヘルスケア・ジャパン株式会社

株式会社ジーデップ・アドバンス 株式会社 NOBORI 富士フイルムメディカル株式会社 菱洋エレクトロ株式会社

広告掲載:株式会社インナービジョン
 株式会社オーム
 クレアボ・テクノロジーズ株式会社
 PSP株式会社
 富士通 Japan 株式会社

協

- 賛: 医用画像情報学会
 - 公益社団法人 応用物理学会
 - 一般社団法人 可視化情報学会
 - 一般社団法人 画像電子学会
 - 看護理工学会
 - 三次元画像コンファレンス実行委員会
 - 一般社団法人 情報処理学会
 - 一般社団法人 電子情報通信学会
 - 公益社団法人 日本医学物理学会
 - 公益社団法人 日本医学放射線学会
 - 一般社団法人 日本医療情報学会
 - 一般社団法人 日本核医学会
 - 特定非営利活動法人 日本核医学技術学会
 - 日本画像医学会
 - 一般社団法人 日本コンピュータ外科学会
 - 一般社団法人 日本磁気共鳴医学会
 - 一般社団法人 日本写真学会
 - 公益社団法人 日本生体医工学会
 - 公益社団法人 日本超音波医学会
 - 日本脳神経 CI 学会
 - 公益社団法人 日本放射線技術学会

田中栄一先生追悼の集い

日本医用画像工学会(JAMIT)の前身である「CT 物理技術研究会」、「医用画像工学研究会」の会 長をされ、放射線計測、核医学、画像再構成など多彩な分野で著しい業績を挙げられた田中栄一先 生が 2021 年 8 月 21 日(土)に逝去されました。JAMIT を創設した方と言っても過言ではありませ ん。ご冥福を心よりお祈り申し上げます。先生は、近年でも JAMIT 大会に積極的に参加され、JAMIT を大変気にかけ心配してくださっており、田中栄一記念賞の原資となるご寄付もいただいておりま した。田中栄一先生へ追悼の意を表すとともに先生の優れた業績を多くの方に知ってもらうため、 先生と一緒に研究や開発を行った方々に田中栄一先生の業績や思い出などをお話いただきます。具 体的な内容は以下の通りです。

日時: 2021年10月13日(水)18:15~19:45

- 会場: JAMIT2021 大会メインホール (慶応義塾大学 日吉キャンパス 協生館)
- 司会: 尾川浩一(法政大学)
- 1. 開式のことば
- 2. 田中栄一先生ご業績の概要紹介
- 3. 黙とう
- 田中栄一先生との想い出
 (登壇予定者:河田燕,飯沼武,掛川誠,山崎統四郎,村山秀雄,山谷泰賀,工藤博幸)
- ビデオメッセージ紹介
 (予定者:外山比南子,野原功全,山下貴司, David Townsend, Michel Defrise, Grant T. Gullberg)
- 6. 閉式のことば
- ※ オンライン参加の場合は Zoom の以下の接続先から接続願います。

https://us02web.zoom.us/j/89422465612?pwd=ZmpmcFBWWlYvRWhMVHc3N3BLbTBoUT09

ウェビナーID:894 2246 5612

パスコード:879327

- ※ 平服でご参加ください。
- ※本集いのみ参加される方は、JAMIT2021 大会の参加登録は不要です。

日本医用画像工学会(JAMIT)会長 工藤博幸 「田中栄一先生追悼の集い」開催事務局 山谷泰賀、村山秀雄

第1日目/10月13日(水)				
	メイン会場(藤原洋記念ホール)	展示会場		
13:00	13:00-13:10 開会式	13:00-17:00		
13:10	13:00-13:10 開会式 13:10-16:00 第11回JAMITチュートリアル講演会(教育委員会企画) コニカミノルタ科学技術振興財団, JAMITノンズオンセミナー連携企画 TL「GitHub とその周辺:大学と企業でのプロジェクト管理」 13:10-14:00 TL1「(未定)」 中田 典生 14:10-15:00 TL2「バイオインフォマティクス分野のアプリケーション開発におけるGitHubの活用とその実際」 小野 浩雅 15:10-16:00 TL3「医療機器開発企業における共同研究のための環境と実際の運用について(事例を交えて)」 大村 和元 座長:木戸 尚治/滝沢 穂高			
16:15	16:15-18:00 一般演題セッション1 P1「CADx, 画像分類」(P1-01~19)			
	座長:野村 行弘/目加田慶人			
18:15	18:15-19:45 田中栄一先生追悼の集い			

第2日目/10月14日(木)							
	メイン会場(藤原洋記念ホール)	展示会場					
9:00	9:00-10:40 一般演題セッション2 P2「CADe, 計測, Segmentation」(P2-01~21)	9:00-17:00					
	座長:清水 昭伸/管 幹生	企業展示					
10:50	10:50-11:40 特別講演1 SL1「量子コンピュータをどう使いこなすか」						
	伊藤 公平						
	座長:工藤 博幸						
11:50	11:50-12:20 スポンサードセミナー1 SS1「Amazon SageMakerを使って医用画像の機械学習をはじめる方法」						
	橋本 正弘/今井 真宏						
	座長:中田 典生						
12:20	12:20-12:50 スポンサードセミナー2 SS2「NVIDIAにおける医療AIの取り組み -医用画像AIを加速するClara Imagingのご紹介-」						
	阮 佩穎						
	座長:中田 典生						
13:00	13:00-14:00 総会						
14:10	14:10-15:40 特別企画1 SE1「重カ下の人体を可視化する~立位のイメージング~」						
	田中 利恵/山田 祥岳/宮地 利明/山谷 泰賀						
	座長:佐藤 嘉伸/陣崎 雅弘						
15:50	15:50-17:20 シンボジウム1 SY1「AI教育を考える」						
	植田 琢也/森 健策/鈴木 賢治/金森 由博						
	座長:橋本 正弘/森 健策						
17:30	17:30-18:40 一般演題セッション3 P3「医用画像:X線写真,核医学,超音波,その他」 (P3-01~14)						
	座長:小尾 高史/羽石 秀昭						

第3日目/10月15日(金)				
	メイン会場(藤原洋記念ホール)	展示会場		
9:00	9:00-10:50 一般演題セッション4 P4「医用画像:CT, MRI」(P4-01~22)	9:00-16:00		
	座長:伊藤 聡志/滝沢 穂高	企業展示		
11:00	11:00-11:50 特別講演2 SL2「革新的バイオイメージング技術」			
	宮脇 敦史			
	座長:陣崎 雅弘			
12:00	12:00-12:50 スポンサードセミナー3 SS3「薬事承認された画像用AIソフト」			
	久永 隆治/岩田 和浩/大越 厚/渡辺 和之			
	座長:木戸 尚治			
13:00	13:00-14:30 シンボジウム2 SY2「横断画像で微細構造がどこまで見えるか」			
	梶田 大樹/畠 二郎/岩澤 多恵/岡田 知久/堀越 浩幸			
	座長:清水 昭伸/花岡 昇平			
14:40	14:40–16:10 特別企画 2 SE2「脳のイメージング最前線」			
	押尾 晃一/田岡 俊昭/林 拓也			
	座長:増谷 佳孝			
16:20	16:20-17:05 一般演題セッション5 P5「イメージング,再構成,CAS,その他」(P5-01~10)			
	座長:北坂 孝幸/森 健策			
17:05	17:05-17:10 閉会式			

ハンズオンセミナー「再スタート・深層学習」/10月16日(土)					
	オンライン				
10:00	10:00-11:00 ハンズオンセミナー(1) HS1「入門編:環境構築から自然画像の分類」				
11:10	11:10-12:10 ハンズオンセミナー(2) HS2「画像の分類と回帰:自分で構築するデータベースの利用」				
13:00	13:00-14:00 ハンズオンセミナー(3) HS3「領域抽出と領域分割:物体検出とセマンティックセグメンテーション」				
14:10	14:10-15:10 ハンズオンセミナー(4) HS4「異常検知:AutoEncoderによる教師なし学習入門」				
15:20	15:20-16:20 ハンズオンセミナー(5) HS5「画像生成:フェイク画像の作成入門」				

特別講演

10月14日(木)

メイン会場

特別講演1

10:50~11:40 SL1「量子コンピュータをどう使いこなすか」

座長:工藤 博幸(筑波大学/JAMIT 学会長)

 SL1
 量子コンピュータをどう使いこなすか

 伊藤 公平
 慶應義塾長,慶應義塾大学理工学部 教授

10月15日(金)

メイン会場

特別講演2

11:00~11:50 SL2「革新的バイオイメージング技術」

座長: 陣崎 雅弘

(慶應義塾大学/JAMIT2021大会長)

 SL2
 革新的バイオイメージング技術

 宮脇 敦史
 理化学研究所 脳神経化学研究センター チームリーダー, 光量子工学研究センター チームリーダー

第11回 JAMIT チュートリアル講演会

10月13日(水)

メイン会場

第 11 回 JAMIT チュートリアル講演会

13:10~16:00 | TL [GitHub とその周辺:大学と企業でのプロジェクト管理]

座長:木戸尚治(大阪大学)

滝沢 穂高(筑波大学)

TL1

未定

中田 典生 東京慈恵会医科大学

- TL2 バイオインフォマティクス分野のアプリケーション開発における GitHub の活用とその実際 小野 浩雅 情報・システム研究機構
- TL3
 医療機器開発企業における共同研究のための環境と実際の運用について(事例を交えて)

 大村 和元
 GE ヘルスケア・ジャパン株式会社

シンポジウム

10月14日(木)

メイン会場

シンポジウム1

15:50 ~ 17:20

座長:橋本 正弘(慶應義塾大学)

森 健策(名古屋大学)

- SY1-1
 デザイン思考にもとづく医療 AI 研究のための人材育成 植田 琢也

 東北大学 医学系研究科 (文科省 AI 人材育成教育拠点)
- SY1-2
 医療 AI 人材教育における現状と情報学研究者としての展望

 森
 健策

 名古屋大学大学院情報学研究科(文科省)
 AI 人材育成教育拠点)

SY1「AI 教育を考える」

- SY1-3
 国際競争に打ち勝つ AI 人材を育成するために何が必要か?

 鈴木 賢治
 東京工業大学 科学技術創成研究院
- SY1-4
 CG・画像処理分野での深層学習の実践的教育

 金森<由博</td>
 筑波大学 システム情報系

10月15日(金)

メイン会場

シンポジウム2

13:00 ~ 14:30	SY 2「横断画像で微細構造がどこまで見えるか」							
	座長:清水	昭伸(東京農工大学)						
	花岡	昇平(東京大学)						

 SY2-1
 光超音波イメージングによる表在リンパ管の観察

 梶田<大樹</td>
 慶應義塾大学

 医学部
 形成外科

- SY2-2
 超音波 SMI による微細血管の描出 島

 山崎医科大学
 検査診断学

 内視鏡・超音波部門
- SY2-3
 高精細 CT

 岩澤 多恵
 神奈川県立循環器呼吸器病センター 放射線科
- SY2-4 7 テスラ MRI 画像による微細脳構造評価 岡田 知久 京都大学大学院医学研究科 脳機能総合研究センター
- SY2-5
 デジタル(半導体検出器)PET

 堀越
 浩幸

 群馬県立がんセンター
 放射線科

特別企画

10月14日(木)

メイン会場

特別企画1

14:10~15:40 SE1「重力下の人体を可視化する~立位のイメージング~」

座長:佐藤 嘉伸 (奈良先端科学技術大学院大学)

陣崎雅弘(慶應義塾大学)

- SE1-1
 X 線動態撮影

 田中 理恵
 金沢大学 医薬保健研究域
- SE1-2 立位 CT の開発と応用~健康長寿の時代に向けて~ 山田 祥岳 慶應義塾大学 医学部 放射線科学教室(診断)
- SE1-3
 Gravity MRI による多姿勢イメージング

 宮地 利明
 金沢大学 医薬保健研究域
- SE1-4
 挑戦する PET:世界初のヘルメット型 PET は座位ブームを起こすか?

 山谷 泰賀
 量子科学技術研究開発機構

メイン会場

特別企画2

14:40~16:10 SE2「脳のイメージング最前線」

座長: 増谷 佳孝 (広島市立大学)

 SE2-1
 T2 解析による Glymphatic pathway の可視化 押尾

 規定
 現

 順天堂大学
 医学部

 SE2-2
 拡散画像で見る間質液動態

 田岡 俊昭
 名古屋大学 大学院医学系研究科 革新的生体可視化技術開発産学協同研究講座

SE2-3 非侵襲画像技術による霊長類脳コネクトーム解明研究 林 拓也 理化学研究所 生命機能科学研究センター

スポンサードセミナー

10月14日(木)

メイン会場

スポンサードセミナー1

11:50 ~ 12:20

SS1 「Amazon SageMaker を使って

医用画像の機械学習をはじめる方法」

座長:中田 典生 (東京慈恵会医科大学)

SS1Amazon SageMaker を使って医用画像の機械学習をはじめる方法
橋本 正弘
慶應義塾大学医学部 放射線科 (診断) 特任助教
今井 真宏
Amazon Web Services Japan PublicSector シニアソリューションアーキテクト

スポンサードセミナー2

12:20~12:50 SS2 NVIDIA における医療 AI の取り組み

- 医用画像 AI を加速する Clara Imaging のご紹介 -」

座長:中田 典生 (東京慈恵会医科大学)

 SS2
 NVIDIA における医療 AI の取り組み - 医用画像 AI を加速する Clara Imaging のご紹介 -

 阮
 佩穎

エヌビディア合同会社 シニア ディープラーニング ソリューションアーキテクト

10月15日(金)

メイン会場

読影支

スポンサードセミナー3

12:00~12:50 SS3「薬事承認された画像用 AI ソフト」

座長:木戸 尚治 (大阪大学)

SS3-1	富士フイルムの AI 技術「REiLI」の現状 – 最新の取り組みのご紹介 – 久永 隆治 富士フイルム株式会社 メディカルシステム事業部・マネージャー
SS3-2	Siemens Healthineers の AI 技術を応用した画像診断支援サービスの特長と活用事例 岩田 和浩 シーメンスヘルスケア株式会社 デジタルヘルス& SYNGO 事業部 AI プロダクトマネージャー
SS3-3	画像診断における AI の実装へ向けた課題と取り組み 大越 厚 GE ヘルスケア・ジャパン株式会社 エジソン・ソリューション本部 部長
SS3-4	 "Abierto Reading Support Solution" 画像診断領域における AI 活用の取組み 渡辺 和之 キヤノンメディカルシステムズ株式会社 ヘルスケア IT 事業統括部 ヘルスケア IT 事業部 援ソリューション推進責任者

コニカミノルタ科学技術振興財団

JAMITハンズオンセミナー

(運営スタッフ)	原	武史	(岐阜大学)
	李	鎔範	(新潟大学)
	中田	典生	(東京慈恵会医科大学)
	小田	昌宏	(名古屋大学)
	滝沢	穂高	(筑波大学)

ハンズオンセミナー(1)

10:00~11:00 HS1「再スタート・深層学習」

HS1 入門編:環境構築から自然画像の分類

ハンズオンセミナー(2)

11:10~12:10 HS2「再スタート・深層学習」

HS2 画像の分類と回帰:自分で構築するデータベースの利用

ハンズオンセミナー(3)

13:00~14:00 HS3「再スタート・深層学習」

HS3 領域抽出と領域分割:物体検出とセマンティックセグメンテーション

ハンズオンセミナー(4)

14:10~15:10 HS4「再スタート・深層学習」

HS4 異常検知: AutoEncoder による教師なし学習入門

ハンズオンセミナー(5)

15:20~16:20 **HS5「再スタート・深層学習」**

HS5 画像生成:フェイク画像の作成入門

プログラム

10月13日(水)

メイン会場

開会式

13:00 ~ 13:10

第 11 回 JAMIT チュートリアル講演会

13:10~16:00 | TL [GitHub とその周辺:大学と企業でのプロジェクト管理]

座長:木戸 尚治(大阪大学)

滝沢 穂高(筑波大学)

- TL1
 未定

 中田 典生
 東京慈恵会医科大学
- TL2 バイオインフォマティクス分野のアプリケーション開発における GitHub の活用とその実際 小野 浩雅 情報・システム研究機構
- TL3
 医療機器開発企業における共同研究のための環境と実際の運用について(事例を交えて)

 大村 和元
 GE ヘルスケア・ジャパン株式会社

一般演題セッション1

16:15 ~ 18:00

| P1「CADx,画像分類」

座長:野村 行弘(千葉大学) 目加田慶人(中京大学)

- P1-01
 胸部 CT 像からの COVID-19 症例の自動分類手法

 小田 昌宏¹²,鄭 通²,林 雄一郎²,大竹 義人³⁴,橋本 正弘⁵,明石 敏昭⁶,森 健策²¹⁴

 ¹名古屋大学情報連携推進本部情報戦略室,²名古屋大学大学院情報学研究科,³奈良先端科学技術大学院大学先端科学技術研究科情報科学領域,⁴国立情報学研究所医療ビッグデータ研究センター,⁵慶

 應義塾大学医学部放射線科,⁶順天堂大学医学部放射線医学講座
- P1-02
 極度に不均衡な頭部 CT 画像データセットにおける Self-attention class balanced DenseNet_ LSTM 構造を用いたくも膜下出血 CT 画像分類

魯 仲陽¹, 小田 昌宏^{2,1}, 林 雄一郎¹, 胡 涛¹, 伊東 隼人¹, 渡谷 岳行³, 阿部 修³, 橋本 正弘⁴, 陣崎 雅弘⁴, 森 健策^{15,6}

¹名古屋大学大学院情報学研究科,²名古屋大学情報連携推進本部情報戦略室,³東京大学医学部附属 病院放射線科,⁴慶應義塾大学医学部放射線科学教室,⁵名古屋大学情報基盤センター,⁶国立情報学 研究所 医療ビッグデータ研究センター

- P1-03
 AI による乳がん腫瘤影の良悪性鑑別処理における画質の相違が CAD 性能に与える影響の分析

 工藤 颯馬¹,安倍 和弥¹,武尾 英哉¹,永井 優一²,縄野 繁³

 ¹神奈川工科大学,²国立がん研究センター東病院,³新松戸中央総合病院
- P1-04
 マルチチャンネル化と M-Net を用いた嚥下時 X 線動画からの頸椎椎間板の抽出と評価実験

 郡司絵莉華¹,目片幸二郎²,滝沢 穂高¹,工藤 博幸¹

 ¹ 筑波大学 情報理工学位プログラム,²四條畷学園大学 リハビリテーション学部 作業療法学専攻
- P1-05
 EfficientNet を用いたアンサンブル学習に基づく放射線皮膚炎グレード判定手法の開発

 和田 清隆¹²,渡邊 陸²,新野 将史²,野口 康介²,荻野 尚¹

 ¹メディポリス国際陽子線治療センター,²鹿児島大学大学院 理工学研究科
- P1-06
 畳み込み自己符号化器つき Deep Autoencoding Gaussian Mixture Model を用いた胸部 CT 画像

 における異常検知
 栗林 雅刀¹,間普 真吾¹,若本 亮佑¹,呉本 尭²,木戸 尚治³

 ¹山口大学,²日本工業大学,³大阪大学
- P1-07
 骨シンチグラムを用いた診断支援システムの再学習による性能変化

 島田 夏帆¹,斉藤 篤¹,大崎 洋充²,東山 滋明³,河邉 譲治⁴,中岡 竜介⁵,清水 昭伸¹

 ¹東京農工大学大学院工学研究院,²群馬県立県民健康科学大学大学院診療放射線学研究科,³大阪市 立大学医学部附属病院核医学科,⁴大阪市立大学医学部附属病院核医学科,⁵国立医薬品食品衛生研究 所医療機器部
- P1-08
 CycleGAN を用いた甲状腺組織画像のドメイン変換と識別システムへの応用

 一氏
 良仁¹,間普
 真吾¹,八田<</td>
 聡美²,稲井
 邦博²,木戸
 尚治³

 ¹山口大学,²福井大学,³大阪大学
- P1-09
 アテンション誘導を利用した胸部 CT 画像による COVID-19 の診断支援 高手山雄亮¹, 斉藤 篤¹, 橋本 正弘², 大竹 義人³, 明石 敏昭⁴, 清水 昭伸¹

 ¹東京農工大学大学院工学研究院,²慶應義塾大学医学部放射線科(診断),³奈良先端科学技術大学院 大学,⁴順天堂大学 放射線診断学講座
- P1-10
 口腔診断支援のための舌の色特徴に基づく機械学習を用いた疾患予測

 野口 桂冴¹,斎藤 一郎²,吉村裕一郎³,並木 隆雄⁴,渡辺 悠紀⁴,中口 俊哉⁵

 ¹千葉大学大学院融合理工学府基幹工学専攻医工学コース,² 鶴見大学歯学部 病理学講座,³富山大学

 附属病院放射線部,⁴千葉大学大学院医学研究院和漢診療学,⁵千葉大学フロンティア医工学センター
- P1-11
 腸閉塞・イレウスの病変箇所特定における診断支援システムの精度評価
 小田 紘久¹,林 雄一郎²,北坂 孝幸³,玉田 雄大⁴,滝本愛太朗¹,檜 顕成¹,内田 広夫¹,
 鈴木耕次郎⁵,小田 昌宏⁶,森 健策^{2.7.8}
 ¹名古屋大学大学院医学系研究科,²名古屋大学大学院情報学研究科,³愛知工業大学情報科学部,⁴名
 古屋大学医学部,⁵愛知医科大学放射線科,⁶名古屋大学情報連携統括本部,⁷名古屋大学情報基盤センター,⁸国立情報学研究所医療ビッグデータ研究センター
- P1-12
 手術標本及び生検標本の H&E 染色組織画像を用いた畳み込みニューラルネットによる肝細胞癌

 の識別
 高木 弦¹、山口 雅浩¹、阿部 時也²、橋口 明典²、坂元 亨宇²

 ¹東京工業大学工学院、²慶應義塾大学医学部病理学教室
- P1-13
 骨粗鬆症 CT 検診におけるコンピュータ診断支援システムの開発

 宮崎 優太¹, 鈴木 秀宣¹, 河田 佳樹¹, 仁木 登¹, 松元 祐司², 土田 敬明², 楠本 昌彦²,

 金子 昌弘³

 ¹徳島大学創成科学研究科理工学専攻光システムコース, ²国立ガン研究センター中央病院, ³保険会

 館クリニック

- P1-14
 機械学習による軟部腫瘍 MRI 画像の良悪性判別

 ハックハスナイン¹⁴, 弘實
 透³, 橋本 正弘¹, 東田 直樹¹, 有田 祐起¹, 庄野 寛太², 金内友里恵²,

 カーントラン¹, 植竹
 望⁴, 陣崎 雅弘¹

 ¹慶應義塾大学医学部整形外科, ²慶應義塾大学工学部, ³慶應義塾大学医学部整形外科, ⁴GE ヘル

 スケアジャパン(株)
- P1-15
 転移学習による手部骨髄浮腫の識別精度向上についての研究

 潘
 冬平¹,小田
 昌宏²¹,片山
 耕³,大久保学宣³,森
 健策¹⁴⁵

 ¹名古屋大学大学院情報学研究科,²名古屋大学情報連携推進本部情報戦略室,³片山整形外科リウマ

 チ科クリニック,⁴名古屋大学情報基盤センター,⁵国立情報学研究所医療ビッグデータ研究センター
- P1-16
 肺胞蛋白症 CT 画像における重症度分類器構築の検討

 山浦賢太郎¹²,小田 未来⁴,石井 晴之⁴,北村 信隆⁵⁶,田中 崇裕⁵⁶,中田 光⁵⁷,野中 尋史¹³

 ¹長岡技術科学大学,²技術支援センター,³情報・経営システム工学専攻,⁴杏林大学医学部,⁵新潟

 大学医歯学総合病院,⁶臨床研究推進センター,⁷高度医療開発センター
- P1-17
 歯科パノラマ X 線画像における分岐を含む SSD とヒューリスティクスによる歯列の認識手法

 森下 拓水¹、村松千左子²、清野 雄多³、周 向栄³、高橋 龍⁴、林 達郎⁴、西山 航⁵、

 原 武史³、有地 淑子⁶、有地榮一郎⁶、勝又 明敏⁵、藤田 広志³

 ¹岐阜大学大学院自然科学技術研究科知能理工学専攻、²滋賀大学データサイエンス学部、³岐阜大学

 工学部電気電子・情報工学科、⁴メディア株式会社、⁵朝日大学歯学部口腔病態医療学講座歯科放射線

 学分野、⁶愛知学院大学歯学部歯科放射線学講座
- P1-18
 3D-Attention Branch Networks を用いた脳 MRI 画像における低悪性度グリオーマの 1p/19q 共 欠失推定法

 田中
 大貴¹, 檜作
 彰良¹, 中山
 良平¹

1立命館大学大学院 理工学研究科

P1-19 体幹部定位放射線治療を受けた非小細胞肺癌患者における癌の進行予測のためのレディオミクス 解析

兄玉 拓巳¹,有村 秀孝²,二宮 健太¹,白川 友子³,吉武 忠正⁴,塩山 善之⁵
 ¹九州大学大学院医学系学府保健学専攻,²九州大学医学研究院保健学部門,³国立病院機構 九州がんセンター,⁴九州大学病院 放射線科,⁵九州国際重粒子線がん治療センター

10月13日(水)

メイン会場

田中栄ー先生追悼の集い

18:15 ~ 19:45

メイン会場

一般演題セッション2

9:00 ⁄	~ 10:40	P2 [CADe	, 計測,	Segmen	tation				
					座長:	清水	昭伸(東京農	 【工大学)	
						菅	幹生(千葉大	:学)	
P2-01	3D CNN 森田 ¹ 兵屆 民病	を用いた小児脳 (蓮 ¹ , 安藤 沙耶 ^{軍県立大学} 大学院 院	CT 画像か ² ,藤田 ジ 工学研究科	らの脳成長 大輔 ¹ ,新居 , ² 兵庫県立	年 齢推定 研 学 ¹ ,安 尼崎総合医	千究 藤久美- 療セング	子 ³ , 石藏 利 ター, ³ 神戸市	L一 ³ ,小橋 [「立医療センタ	¹ ー中央市
P2-02	自己符号 中原 ジ ¹ 山口	化器を用いた慢性 ^{祥喜 1} , 間普 真吾 1大学	b閉塞性肺 ¹ , 平野 着	疾患におけ _{網彦¹,村田}	る身体活動 順之 ¹ , 土	b性のう 居 恵	予測に関する 子 ¹ ,松永 和	研究 1人 ¹	
P2-03	機械学習 中津 ↓ ¹ 兵庫	・深層学習を用い _{東平 ¹, 盛田 健人 ^{重県立大学} 工学研}	、た手 X 親 ² ,藤田 究科, ² 三	豪画像におけ 大輔 ¹ ,小橋 重大学	るリウマ ラ 昌司 ¹	≁mTS	S 推定法		
P2-04	機械学習 篠原 ¹ 近台	によるコルポスニ 寿広 ¹ ,藤井 一輝 &大学生物理工学部	1ープ画像 ¹ ,村上 生命情報コ	における白 幸祐 ² , 中迫 二学科, ² 近畿	色上皮領域 昇 ¹ ,松 大学医学部	伐検出 料 村 謙日 産科婦ノ	青度向上のた ^{五² 人科学教室}	めの検討	
P2-05	覚醒下脳 出手法の 佐藤 ¹ 公ゴ	腫瘍摘出術におけ 提案 生馬 ¹ ,田村 学 なはこだて未来大学	ける手術工 ² ,山口 ギ システム	程同定のた 智子 ² ,藤野 情報科学部,	めの深層学 雄一 ¹ ,吉 ² 東京女子	^全 習を月 光喜太臣 医科大学	目いた顕微鏡 郎 ² , 村垣 善 学 先端生命医	画 像からの 特 ^注 曲 (本) (本) (本) (本) (本) (本) (本) (本)	寺徴量抽 賢 ²
P2-06	Transfor 岡崎 森 ¹ 名己 放射 医療	mer ベースのモラ 真治 ¹ ,林雄一郎 健索 ¹⁵⁶ 5屋大学大学院情報 線科, ⁴ 順天堂大学 ビッグデータ研究も	^ギ ルを用い ¹ ,小田 学研究科, 医学部放射 <i>zンター</i>	た肺領域に 昌宏 ²¹ ,橋オ ² 名古屋大与 时線医学講座	おける所 ^x 正弘 ³ , 学情報連携排 , ⁵ 名古屋ナ	し文から 陣崎 登 進本部 大学情報	らの疾患名推 雅弘 ³ ,明石 B情報戦略室, B基盤センター	出 敏昭 ⁴ ,青木 ³ 慶應義塾大: -, ⁶ 国立情報:	茂樹 ⁴ , 学医学部 学研究所
P2-07	取り下げ								
P2-08	3次元 C 山本 村津 ¹ 兵属	T 画像を用いた栲 侃利 ¹ ,藤田 大輔 裕嗣 ³ ,小橋 昌司 ⁵ 県立大学,工学部,	 擬ボーリ ¹、ラシェ ²姫路獨制 	ング調査に ドゥーララー 協大学, ³ 製銀	基 づく脆 弱 マン ¹ ,八オ t,広畑記念病	号性骨 盤 、 直美 「院	整骨折検出法 ² ,林 圭 [:]	による特徴量 吾 ³ , 圓尾 明	との検討 宏 ³ ,
P2-09	膵臓の超 竹内 清水 ¹ 岐 第一	音波内視鏡から作 祐慈 ¹ ,岩佐 悠平 雅仁 ³ ³ ² 大学大学院 自然和 内科, ⁴ 岐阜大学 コ	F成した2 ² ,岩下 斗学技術研 二学部 電気	次元ヒスト 拓司 ³ , 原 究科 知能理 電子・情報工	グラムを用 武史 ⁴ , J 工学専攻, ² 工学科	引いた 周 向 岐阜市	萃癌自動分類 句栄 ⁴ ,上村 民病院 消化器	晋也 ³ , 三田 器内科, ³ 岐阜	直樹 ³ , 大学病院

- P2-10
 多段の画素異常検知による FDG-PET/CT 上のがん骨転移候補検出
 山口 明乃¹,根本 充貴²,甲斐田勇人³⁴,木村 裕一¹,永岡 隆¹,山田 誉大⁴,花岡 宏平⁴, 北島 一宏⁵, 槌谷 達也⁶,石井 一成³⁴
 ¹近畿大学大学院 生物理工学研究科,²近畿大学 生物理工学部,³近畿大学 医学部 放射線医学教室 放 射線診断学部門,⁴近畿大学高度先端総合医療センター PET 分子イメージング部,⁵ 兵庫医科大学 放 射線医学講座,⁶ 兵庫医科大学病院 放射線技術部
- P2-11
 PSA density 算出のためのセグメンテーションによる前立腺体積計測法の検討 安倍和弥¹,武尾英哉¹,永井優一²,黒木嘉典³,縄野 繁⁴

 ¹神奈川工科大学,²国立がん研究センター東病院,³にいむら病院,⁴新松戸中央総合病院
- P2-12 Perforation Detection from Endoscopy Videos Using Model Training with Synthesised Images by GAN

蔣 凱¹, 伊東 隼人¹, 小田 昌宏²¹, 奥村 大志³, 森 悠一³, 三澤 将史³, 林 武雅³, 工藤 進英³, 森 健策¹⁴⁵

¹名古屋大学 大学院情報学研究科,²名古屋大学 情報連携推進本部情報戦略室,³昭和大学 横浜市北 部病院消化器センター,⁴名古屋大学 情報基盤センター,⁵国立情報学研究所 医療ビッグデータ研究 センター

 P2-13
 実体顕微鏡画像における血管構造セグメンテーション

 呉
 運恒¹,小田
 昌宏⁵¹,林
 雄一郎¹,武部
 貴則³,森
 健策¹²⁴

 ¹名古屋大学大学院情報学研究科、²名古屋大学情報基盤センター、³東京医科歯科大学統合研究機構、
 ⁴国立情報学研究所医療ビッグデータ研究センター、⁵名古屋大学情報連携推進本部

P2-14 深度情報を利用した FCN による腹腔鏡映像からの血管領域自動抽出の検討 榎本 圭吾¹,林 雄一郎¹,北坂 孝幸²,小田 昌宏³¹,伊藤 雅昭⁴,竹下 修由⁴,三澤 一成⁵, 森 健策^{1.67}

¹名古屋大学大学院情報学研究科,²愛知工業大学情報科学部,³名古屋大学情報連携推進本部情報戦略室,⁴国立がん研究センター東病院,⁵愛知県がんセンター,⁶名古屋大学情報基盤センター,⁷国立 情報学研究所医療ビッグデータ研究センター

- P2-15 Attention 機構を導入したグラフニューラルネットワークによる腹部動脈血管名自動対応付け 出口 智也¹,林 雄一郎¹,北坂 孝幸²,小田 昌宏³¹,三澤 一成⁴,森 健策^{1,5,6} ¹名古屋大学大学院情報学研究科,²愛知工業大学情報科学部,³名古屋大学情報連携推進本部情報戦 略室,⁴愛知県がんセンター消化器外科,⁵名古屋大学情報基盤センター,⁶国立情報学研究所医療ビッ グデータ研究センター
- P2-16
 Cascaded Convolutional Network を用いた 3D 腎臓腫瘍セマンティックセグメンテーション

 趙
 武楊¹, 申
 忱¹, 小田
 昌宏²¹, 林
 雄一郎¹, 東田
 直樹³, 橋本
 正弘³, 陣崎
 雅弘³,

 森
 健策^{1,45}
 ¹タエ屋土学
 土学院/長報学研究科
 ²タ土屋土学
 唐朝連進進進本部長報酬略容科
 ³ 唐爾美朝土学 医

¹名古屋大学 大学院情報学研究科,²名古屋大学 情報連携推進本部情報戦略室科,³慶應義塾大学 医 学部放射線科,⁴名古屋大学 情報基盤センター,⁵国立情報学研究所 医療ビッグデータ研究センター

- P2-17 Bayesian U-Net を用いた脊椎セグメンテーションの不確実性予測:大規模 CT データベースに おける年齢・性別統計モデリングに向けて
- 程 卓¹, 大竹 義人¹, スーフィーマーゼン¹, 上村 圭亮¹, 重松 秀樹², 池尻 正樹², 合田 憲人³, 佐藤 真一³, 橋本 正弘⁴, 明石 敏昭⁵, 佐藤 嘉伸¹ ¹奈良先端科学技術大学院大学,²奈良県立医科大学,³国立情報学研究所,⁴慶應義塾大学,⁵順天堂 大学
- P2-18
 弱教師あり学習を用いた MR 画像における血管奇形腫瘍の領域の自動抽出

 深谷 航生¹,原 武史²,野崎 太希³,周 向栄²,藤田 広志²

 ¹岐阜大学大学院自然科学技術研究科知能理工学専攻,²岐阜大学工学部電気電子・情報工学科,³ 聖 路加国際病院放射線科
- P2-19 取り下げ

P2-20 頭部 thick-slice CT 画像からの階層的脳内血腫領域抽出法の提案 岡 和範¹,藤田 大輔¹,野原 康伸²,井上 創造³,有村 公一⁴,飯原 弘二⁵,小橋 昌司¹ ¹兵庫県立大学 大学院 工学研究科,²熊本大学,³九州工業大学,⁴九州大学,⁵国立循環器病研究 センター

 P2-21
 CycleGAN と Bayesian U-Net による Cross-modality セグメンテーション: CT 画像学習データ を用いた MRI での自動筋体積計測

 李 贛萍¹, 大竹 義人¹, 中西 直樹¹, 谷口 匡史², 八木 優英², 市橋 則明², スーフィーマーゼン¹, 上村 圭亮¹³, 高尾 正樹³, 菅野 伸彦³, 佐藤 嘉伸¹

 ¹奈良先端科学技術大学院大学,²京都大学,³大阪大学

10月14日(木)

メイン会場

特別講演1

10:50~11:40 SL1「量子コンピュータをどう使いこなすか」

座長:工藤 博幸(筑波大学/JAMIT 学会長)

 SL1
 量子コンピュータをどう使いこなすか

 伊藤 公平
 慶應義塾長,慶應義塾大学理工学部 教授

スポンサードセミナー1

11:50 ~ 12:20 SS1 「Amazon SageMaker を使って

医用画像の機械学習をはじめる方法」

座長:中田 典生 (東京慈恵会医科大学)

SS1 Amazon SageMaker を使って医用画像の機械学習をはじめる方法

橋本 正弘¹, 今井 真宏²

1慶應義塾大学医学部 放射線科(診断)特任助教,

²Amazon Web Services Japan PublicSector シニアソリューションアーキテクト

メイン会場

スポンサードセミナー2

12:20~12:50 SS2 NVIDIA における医療 AI の取り組み

- 医用画像 AI を加速する Clara Imaging のご紹介 -」

座長:中田 典生 (東京慈恵会医科大学)

 SS2
 NVIDIA における医療 AI の取り組み - 医用画像 AI を加速する Clara Imaging のご紹介 - 阮 佩穎

エヌビディア合同会社 シニア ディープラーニング ソリューションアーキテクト

総会



特別企画1

14:10~15:40 SE1「重力下の人体を可視化する~立位のイメージング~」

座長: 佐藤 嘉伸(奈良先端科学技術大学院大学) 陣崎 雅弘(慶應義塾大学)

- SE1-1
 X 線動態撮影

 田中 理恵
 金沢大学 医薬保健研究域
- SE1-2 立位 CT の開発と応用~健康長寿の時代に向けて~ 山田 祥岳 慶應義塾大学 医学部 放射線科学教室(診断)
- SE1-3
 Gravity MRI による多姿勢イメージング 宮地 利明 金沢大学 医薬保健研究域
- SE1-4
 挑戦する PET:世界初のヘルメット型 PET は座位ブームを起こすか?

 山谷 泰賀
 量子科学技術研究開発機構

メイン会場

シンポジウム1

15:50 ~ 17:20 SY1「AI 教育を考える」

座長:橋本 正弘 (慶應義塾大学)

森 健策(名古屋大学)

- SY1-1
 デザイン思考にもとづく医療 AI 研究のための人材育成 植田 琢也 東北大学 医学系研究科(文科省 AI 人材育成教育拠点)
- SY1-2
 医療 AI 人材教育における現状と情報学研究者としての展望

 森
 健策

 名古屋大学大学院情報学研究科(文科省 AI 人材育成教育拠点)
- SY1-3
 国際競争に打ち勝つ AI 人材を育成するために何が必要か?

 鈴木 賢治
 東京工業大学 科学技術創成研究院
- SY1-4
 CG・画像処理分野での深層学習の実践的教育

 金森 由博
 筑波大学 システム情報系
- 一般演題セッション3

17:30~18:40 P3 [医用画像:X線写真,核医学,超音波,その他]

座長:小尾 高史(東京工業大学)

羽石 秀昭(千葉大学)

- P3-01
 ポータブル型トモシンセシスにおける X 線源と検出器の位置関係推定法 堀

 塩
 挙輔¹,小池
 貴久¹,只野
 喜一¹,橋本
 雄幸¹

 ¹ 杏林大学大学院保健学研究科
 ¹
- P3-02 深層学習とデジタルファントムを用いた骨陰影低減技術の開発 五島 風汰¹,田中 利恵²,小田 昌宏³,森 健策⁴,高田 宗尚⁵,田村 昌也⁵,松本 勲⁵ ¹金沢大学大学院 医薬保健学総合研究科,²金沢大学 医薬保健研究域附属 AI ホスピタル・マクロ シグナルダイナミクス研究開発センター,³名古屋大学情報連携推進本部情報戦略室,⁴名古屋大学大 学院情報学研究科,⁵金沢大学呼吸器外科
- P3-03 解剖学的情報を併用した教師なし深層学習による PET 画像のノイズ除去 大西 佑弥¹,橋本二三生¹,大手 希望¹,大庭 弘行¹,大田 良亮¹,吉川 悦次¹,尾内 康臣² ¹浜松ホトニクス株式会社 中央研究所,²浜松医科大学 光尖端医学教育研究センター 生体機能イメー ジング研究室
- P3-04
 ピンホール SPECT における散乱線の定量と除去法の比較 島田
 良¹、村田
 一心²、尾川
 浩一³

 ¹法政大学大学院理工学研究科、²国立天文台、³法政大学大学院理工学研究科
- P3-05
 デコンボリューション法による静止型マルチピンホール SPECT システムの空間分解能改善 大越 迪¹,村田 一心²,尾川 浩一³

 「法政大学大学院理工学研究科,²国立天文台,³法政大学理工学部

- P3-06
 Deep Image Prior による SPECT 画像の画質改善の試み

 矢部 暁登¹, 對馬 結太¹, 福原 健太¹, 松原 佳亮², 越野 一博³, 渡部 浩司⁴, 銭谷 勉¹

 ¹ 弘前大学大学院理工学研究科,²秋田県立循環器・脳脊髄センター研究所脳血管研究センター放射線

 医学研究部,³北海道情報大学経営情報学部システム情報学科,⁴ 東北大学サイクロトロン・ラジオア

 イソトープセンター
- P3-07
 近接マルチスプリッティングを用いた Row-Action 型 PET 画像再構成法の体系的な構成

 定方 数弥¹,金 喜正¹, 工藤 博幸¹

 ¹ 筑波大学大学院 理工情報生命学術院 システム情報工学研究群 情報理工学位プログラム
- P3-08 Higher SNR PET image prediction using a deep learning model with efficient channel attention YinTuo¹, 小尾 高史²

1 東京工業大学 情報通信系, 2 東京工業大学 科学技術創成研究院

- P3-09 深層学習により PET/CT 早期ダイナミック画像から遅延画像を合成する手法の初期検討 ブーサルチャタクリリトゥ¹, 立花 泰彦¹, 小畠 隆行¹, 西井 龍一¹, 東 達也¹ ¹量子医学医療部門放射線医学総合研究所 分子イメージング診断治療研究部
- P3-10 腹部超音波スクリーニング支援のための深層学習による撮影断面推定に関する初期検討 目加田慶人¹,道満 恵介¹,小川 眞広²,西田直生志³,工藤 正後³ ¹中京大学工学部,²日本大学病院消化器内科,³近畿大学医学部
- P3-11
 深層学習を用いた超音波画像における腎嚢胞の計測補助機能の開発

 金内友里恵¹,橋本 正弘²,東田 直樹²,HAQUEHasnine²³,陣崎 雅弘²,榊原 康文¹

 ¹應義塾大学理工学部,²慶應義塾大学医学部,³GE ヘルスケア・ジャパン
- P3-12 ピットパターン特徴量の解析に向けた超拡大内視鏡画像の再構成法に関する初期的検討 伊東 隼人¹,小田 昌宏¹²,森 悠一³⁴,三澤 将史³,工藤 進英³,森 健策¹⁵⁶ ¹名古屋大学 大学院情報学研究科,²名古屋大学 情報連携推進本部情報戦略室,³昭和大学横浜市北部 病院 消化器センター,⁴オスロー大学 臨床効果研究グループ,⁵名古屋大学 情報基盤センター,⁶国 立情報学研究所 医療ビッグデータ研究センター
- P3-13
 X 線暗視野法を用いた乳腺における乳管内腺腔の自動抽出および3次元解析手法の開発

 谷口 華奈¹,砂口 尚輝¹,市原 周²,西村理恵子²,岩越 朱里²,村上 善子²,島雄 大介³, 湯浅 哲也⁴,安藤 正海⁵

 ¹名古屋大学大学院 医学系研究科,²名古屋医療センター 病理診断科,³北海道科学大学 保健医療学部, ⁴山形大学大学院 理工学研究科,⁵高エネルギー加速器研究機構

P3-14 取り下げ

メイン会場

ー般演題セッション4

9:00~10:50 P4 医用画像:CT. MRI

座長:伊藤 聡志(宇都宮大学)

滝沢 穂高(筑波大学)

 P4-01
 圧縮センシング CT 画像再構成のための FBP 法組み込み型主双対反復法とコーンビーム CT への拡張

 工藤
 博幸¹, 王
 婷¹

1 筑波大学システム情報系情報工学域

P4-02 生体軟組織の屈折コントラスト CT に適用するための機械学習によるリングアーチファクト除去 法の開発

> 黄 卓然¹,砂口 尚輝¹,島雄 大介²,市原 周³,湯浅 哲也⁴,安藤 正海⁵ ¹名古屋大学大学院医学系研究科,²北海道科学大学診療放射線学科,³名古屋医療センター病理診断科, ⁴山形大学大学院理工学研究科,⁵高エネルギー加速器研究機構

 P4-03
 サイクル敵対的生成ネットワークによる胸部4次元コーンビームCTの画質改善

 日井
 桂介¹²,尾川
 浩一³,後藤
 政実¹,坂野
 康昌¹,京極
 伸介¹,代田
 浩之¹

 ¹ 順天堂大学保健医療学部診療放射線学科,² 順天堂大学医学部放射線治療学講座,³法政大学理工学
 部

- P4-04 マルチ近接スプリッティングを用いた Row-Action 型 TV 正則化 CT 画像再構成法の体系的な構成
 - 金 喜正¹, 定方 数弥¹, 工藤 博幸² ¹筑波大学大学院システム情報工学研究群情報理工学位プログラム,²筑波大学システム情報系情報工 学専攻教授
- P4-05 インテリア CT 厳密解法における安定な画像再構成が可能な先験情報領域配置の解明 伊藤 優太¹,藤井 克哉²,工藤 博幸² ¹筑波大学大学院システム情報工学研究群情報理工学位プログラム,²筑波大学システム情報系
- P4-06
 CT 画像再構成におけるインテリア問題とエクステリア問題の統一解法と安定性解明 藤井 克哉¹, 工藤 博幸¹

 ¹筑波大学システム情報系
- P4-07 異常データ検知トモグラフィ ABD-CT の概念提唱と事例紹介 工藤 博幸 筑波大学システム情報系情報工学域
- P4-08
 距離マップを利用した肺マイクロ CT 像からの肺胞抽出

 椎名
 健¹,小田 紘久²,鄭
 通¹,中村 彰太²,林 雄一郎¹,小田 昌宏³¹,森
 健策¹⁴⁵

 ¹名古屋大学 大学院情報学研究科,²名古屋大学 大学院医学系研究科,³名古屋大学 情報連携推進本部,
 ⁴名古屋大学 情報基盤センター,⁵国立情報学研究所 医療ビッグデータ研究センター
- P4-09 非ラグランジアン解法を用いた少数方向投影からの Micro-CT 画像再構成 熊切 俊夫¹,岡本 尚之²,工藤 博幸³,羽石 秀昭⁴ ¹千葉大学大学院融合理工学府数学情報科学専攻,²千葉大学大学院融合理工学府基幹工学専攻,³筑 波大学システム情報系,⁴千葉大学フロンティア医工学センター

 P4-10
 圧縮センシングを深層学習を組み合わせた低線量 CT 画像再構成法 CSDL-Net のヘリカルコーン

 ビーム CT への拡張
 川又健太郎¹、数尾
 昴¹、工藤
 博幸¹

¹筑波大学システム情報系

- P4-11 Pix2pix を用いた任意形状の肺がん CT 画像生成及び CT 画像様ノイズ再現の試み
 - 遠田 凉¹, 寺本 篤司¹, 辻本 正和², 外山 宏³, 近藤 征史³, 今泉 和良³, 齋藤 邦明¹, 藤田 広志⁴ ¹藤田医科大学大学院 保健学研究科, ²藤田医科大学病院, ³藤田医科大学 医学部, ⁴岐阜大学 工 学部
- P4-12 コーンビーム CT の散乱線除去のための機械学習ネットワーク 村田 一心¹,尾川 浩一²,臼井 桂介³ ¹国立天文台,²法政大学理工学部応用情報工学科,³順天堂大学保健医療学部診療放射線学科
- P4-13
 臨床 マイクロ CT 画像間位置合わせ済みデータセットを用いた臨床 CT 像の超解像

 鄭
 通¹,小田 紘久¹,林 雄一郎¹,中村 彰太²,森 雅樹³,高畠 博嗣⁴,名取 博⁵,

 小田 昌宏⁶¹,森
 健策^{1,78}

 ¹名古屋大学大学院情報学研究科,²名古屋大学大学院医学系研究科,³札幌厚生病院,⁴札幌南三条病院,

 ⁵恵和会西岡病院,⁶名古屋大学情報連携推進本部情報戦略室,⁷名古屋大学情報基盤センター,⁸国立 情報学研究所医療ビッグデータ研究センター
- P4-14
 敵対生成ネットワークによる COVID-19 の CT 画像の生成モデル

 川崎
 良¹, 斉藤
 第¹, 橋本
 正弘², 大竹
 義人³, 明石
 敏昭⁴, 清水
 昭伸¹

 ¹東京農工大学大学院工学研究院,²慶應義塾大学医学部放射線科,³奈良先端科学技術大学院大学,
 ⁴順天堂大学放射線診断学講座
- P4-15
 表現に合わせた生成モデルによる非造影から動脈造影 CT への変換

 胡
 涛¹、小田
 昌宏¹、林
 雄一郎¹、魯
 仲陽¹、明石
 敏昭²、青木
 茂樹²、森
 健策¹³⁴

 ¹名古屋大学大学院情報学研究科、²順天堂大学大学院医学研究科、³名古屋大学情報基盤センター、
 ⁴国立情報学研究所 医療ビッグデータ研究センター
- P4-16
 多重解像度画像のグループ型学習を利用した新たな MR 深層学習再構成の検討

 大内 翔平¹²,伊藤 聡志¹
 1¹宇都宮大学大学院 工学研究科 システム創成工学専攻,²日本学術振興会特別研究員 DC
- P4-17
 信号空間の深層学習を利用した位相拡散フーリエ変換に基づいた MR 超解像

 山登 一輝¹, 若槻 泰迪¹, 伊藤 聡志¹

 ¹宇都宮大学工学部
- P4-18 位相拡散フーリエ変換法を導入した MR 深層学習再構成における信号間引きパターンの検討 金澤 崇大¹,山登 一輝¹,伊藤 聡志¹ ¹宇都宮大学大学院 地域創生科学研究科 情報電気電子システム工学プログラム
- P4-19
 MR 画像からの 3 次元導電率分布の回帰による TMS のための誘導電場の回帰精度の改善牧

 豊大¹、横田 達也¹、平田 晃正²、本谷 秀堅¹

 ¹名古屋工業大学、²福島県立医科大学
- P4-20
 PROPELLER MRI における CS と CNN を利用した再構成法の検討

 松元 祐太¹,堀 拳輔¹,只野 喜一²,久原 重英²,遠藤 祐太²,橋本 雄幸²

 ¹ 杏林大学 大学院 保健学研究科保健学専攻 診療放射線学分野,² 杏林大学 保健学部 診療放射線技術

 学科
- P4-21
 AD-MIL を用いた頭部 MR 画像における回復期脳梗塞患者の帰結予測

 東
 優大¹, 寺本 篤司¹, 岡崎 英人², 武田湖太郎³, 前田 繁信⁴

 ¹藤田医科大学大学院保健学研究科,²藤田医科大学医学部連携リハビリテーション医学,³藤田医科 大学保健衛生学部リハビリテーション学科,⁴藤田医科大学七栗記念病院

 P4-22
 腹部 Cine-MRI における小腸蠕動運動のオプティカルフローによる可視化および腹膜炎の鑑別

 川原 稔暉¹,井上 明星²,岩本祐太郎¹,古川 顕³,BolorkhandBatsaikhan³,陳 延偉¹

 ¹立命館大学大学院 情報理工学研究科,²東近江総合医療センター,³東京都立大学

<u>10月15日(金)</u> メイン会場

特別講演2

11:00~11:50 SL2「革新的バイオイメージング技術」

座長:陣崎 雅弘

(慶應義塾大学/JAMIT2021 大会長)

SL2 革新的バイオイメージング技術 宮脇 敦史 理化学研究所 脳神経化学研究センター チームリーダー、光量子工学研究センター チームリーダー

スポンサードセミナー3

援ソリューション推進責任者

SS3「薬事承認された画像用 AI ソフト」 12:00 ~ 12:50 座長:木戸 尚治(大阪大学) SS3-1 富士フイルムの AI 技術「REiLI」の現状 - 最新の取り組みのご紹介 -久永 隆治 富士フイルム株式会社 メディカルシステム事業部・マネージャー SS3-2 Siemens Healthineers の AI 技術を応用した画像診断支援サービスの特長と活用事例 岩田 和浩 シーメンスヘルスケア株式会社 デジタルヘルス& SYNGO 事業部 AI プロダクトマネージャー SS3-3 画像診断における AI の実装へ向けた課題と取り組み 大越 厚 GE ヘルスケア・ジャパン株式会社 エジソン・ソリューション本部 部長 SS3-4 "Abierto Reading Support Solution" 画像診断領域における AI 活用の取組み 渡辺 和之 キヤノンメディカルシステムズ株式会社 ヘルスケア IT 事業統括部 ヘルスケア IT 事業部 読影支

メイン会場

シンポジウム2

13:00~14:30 SY 2 「横断画像で微細構造がどこまで見えるか」

座長:清水 昭伸(東京農工大学)

花岡 昇平 (東京大学)

- SY2-1
 光超音波イメージングによる表在リンパ管の観察

 梶田 大樹¹, 鈴木 悠史¹, 渡部 紫秀¹, 岡部 圭介¹, 佐久間 恒², 大滝真梨香¹, 今西 宣晶³,

 陣崎 雅弘⁴, 相磯 貞和³⁵, 貴志 和生¹

 ¹慶應義塾大学 医学部 形成外科, ²東京歯科大学 市川総合病院 形成外科, ³慶應義塾大学 医

 学部 解剖学教室, ⁴慶應義塾大学 医学部 放射線科学教室 (診断), ⁵株式会社 Luxonus
- SY2-2
 超音波 SMI による微細血管の描出 島

 山崎医科大学
 検査診断学
- SY2-3
 高精細 CT

 岩澤<多恵</td>
 神奈川県立循環器呼吸器病センター
- SY2-4
 7 テスラ MRI 画像による微細脳構造評価 岡田 知久 京都大学大学院医学研究科 脳機能総合研究センター
- SY2-5
 デジタル(半導体検出器)PET

 堀越
 浩幸

 群馬県立がんセンター
 放射線科

特別企画2

14:40~16:10 SE2「脳のイメージング最前線」

座長: 増谷 佳孝 (広島市立大学)

- SE2-1T2 解析による Glymphatic pathway の可視化
押尾抑尾晃一
順天堂大学順天堂大学医学部
放射線科
- SE2-2 拡散画像で見る間質液動態 田岡 俊昭 名古屋大学 大学院医学系研究科 革新的生体可視化技術開発産学協同研究講座
- SE2-3 非侵襲画像技術による霊長類脳コネクトーム解明研究 林 拓也 理化学研究所 生命機能科学研究センター

ー般演題セッション5

16:20~17:05 **P5「イメージング,再構成,CAS,その他」**

座長:北坂 孝幸 (愛知工業大学)

森 健策(名古屋大学)

- P5-01
 2次元フーリエ変換を用いたリングアーチファクト低減フィルターの特性評価

 米山
 明男¹, 馬場
 理香², 河本
 正秀¹

 ¹九州シンクロトロン光研究センター, ²(株)日立製作所研究開発グループ
- P5-02
 GP-GPUを用いた光子輸送シミュレーションの高速化 瀬在 翔太¹,村田 一心²,乳井 嘉之³,尾川 浩一⁴

 ¹法政大学大学院 理工学研究科,²国立天文台,³東京都立大学 健康福祉学部,⁴法政大学 理工学 部
- P5-03 深層学習を利用した信号の間引きパターンに頑健な圧縮センシング再構成の基礎検討 渋井 雅希¹, 山登 一輝¹, 伊藤 聡志¹ ¹宇都宮大学大学院 地域創生科学研究科 情報電気電子システムプログラム
- P5-04
 WGI コンプトン画像再構成条件の実験的検討

 田島英朗¹,吉田英治¹,仁科 匠¹²,菅 幹生²,脇坂秀克¹,高橋美和子¹,永津弘太郎¹, 辻 厚至¹,鎌田 圭³⁴,吉川 彰³⁴, KatiaParodi⁵,山谷泰賀¹

 ¹量子科学技術研究開発機構,²千葉大学,³東北大学,⁴株式会社 C&A,⁵ミュンヘン大学
- P5-05 膝半月板の 2.5 次元統計的形状モデル構築における主成分分析と非負値行列因子分解の比較 繩手 就矢¹, 増谷 佳孝¹ ¹広島市立大学大学院 情報科学研究科
- P5-06 腹腔鏡映像からの手術箇所の認識による手術ナビゲーションの検討 林 雄一郎¹, 三澤 一成², 森 健策^{1,3} ¹名古屋大学大学院情報学研究科,²愛知県がんセンター消化器外科,³国立情報学研究所医療ビッグデー タ研究センター
- P5-07
 VR Organ Puzzle: A Virtual Reality Application for Human Anatomy Education

 李
 斯頎¹, 林 雄一郎¹, 藤原 道隆², 小田 昌宏³¹, 森
 健策^{1,45}

 ¹名古屋大学大学院情報学研究科, ²名古屋大学医学部附属病院医療機器総合管理部. ³名古屋大学情報連携推進本部情報戦略室, ⁴名古屋大学情報基盤センター, ⁵国立情報学研究所医療ビッグデータ研究センター
- P5-08
 複数術野視点映像を用いた機械学習による外科手術工程の自動分類

 藤井 亮輔¹,梶田 大樹²,高詰 佳史³,青木 義満¹

 ¹慶應義塾大学理工学研究科 総合デザイン工学専攻 電気電子工学専修,²慶應義塾大学医学部 形成外科,³慶應義塾大学医学部 解剖学教室
- P5-09 腹腔鏡の視野外における術具誤接触防止システムの開発 岸 知哉¹,中口 俊哉²,長尾 俊哉³,駒村 賢三⁴,西澤 祐吏⁵ ¹千葉大学大学院融合理工学府基幹工学専攻医工学コース,²千葉大学フロンティア医工学センター, ³京セラ株式会社東京事業所,⁴株式会社ソフケン,⁵国立がん研究センター東病院
- P5-10 取り下げ

閉会式

17:05 ~ 17:10

10月16日(土)

オンライン

ハンズオンセミナー(1)

10:00~11:00 HS1「再スタート・深層学習」

HS1 入門編:環境構築から自然画像の分類

ハンズオンセミナー(2)

11:10~12:10 HS2「再スタート・深層学習」

HS2 画像の分類と回帰:自分で構築するデータベースの利用

ハンズオンセミナー(3)

13:00~14:00 HS3「再スタート・深層学習」

HS3 領域抽出と領域分割:物体検出とセマンティックセグメンテーション

ハンズオンセミナー(4)

14:10~15:10 HS4「再スタート・深層学習」

HS4 異常検知: AutoEncoder による教師なし学習入門

メイン会場

ハンズオンセミナー (5)

15:20~16:20 HS5「再スタート・深層学習」

HS5 画像生成:フェイク画像の作成入門

抄 録

(著者から提出された内容をそのまま掲載します)

10月13日(水)

メイン会場

第11回 JAMIT チュートリアル講演会

13:10~16:00 GitHub とその周辺:大学と企業でのプロジェクト管理

座長:木戸 尚治(大阪大学)滝沢 穂高(筑波大学)

TL1 未定

○中田 典生 東京慈恵会医科大学

TL2 バイオインフォマティクス分野のアプリケーション開発における GitHub の活用とその実際

○小野 浩雅

情報・システム研究機構

GitHubは、生命科学分野においてもバイオインフォマティクスを中心に利用が進んでいる.ここ数年で開発されている生命 科学分野の解析ツールやデータベースの多くは、GitHub上でソースコードが公開・共有されており、世界中の誰もが開発に参 加することができる、ソースコード以外にも文書や画像など、どんな種類のドキュメントでも、すべての段階の草案やアップデー ト履歴を保存し管理することができるプロジェクト管理のメリットは大きい、本講演では、演者の所属するセンターにおける いくつかのアプリケーション開発を例に、GitHubの活用法の実際を紹介する.

TL3 医療機器開発企業における共同研究のための環境と実際の運用について(事例を交えて)

○大村 和元

GE ヘルスケア・ジャパン株式会社

最近になって世の中での当初の熱は冷めてきているようにも思うが、第三次 AI ブームから AI を用いた医用画像解析に関す る研究は相変わらず多く、弊社においても AI を用いた製品の開発や AI に関係するお客様との共同研究がここ数年増加してい る. AI の研究、特に医用画像ペースでの研究を始めようとした場合、環境の構築とデータの収集が最初のハードルとなること が多いと思うが、環境構築方法の情報や海外まで目を広げればフリーで扱えるデータ情報の共有も行われており、さらにいく つかの企業においては開発環境としてのプラットフォームの構

築,提供も進められている. 医療機器開発企業が進める共同研究は多岐にわたるが、本講演では医用画像を用いた AI に関連 する共同研究にフォーカスをして、その進め方や環境について事例を挙げつつ概説する. またグローバルな環境で各地で進め られている研究と比較して日本で研究を進めるにあたっての難しさについても言及したい.

ー般演題セッション1

16:15 ~ 18:00 | **P1 [CA**

P1「CADx,画像分類」

座長:野村 行弘(千葉大学) 目加田慶人(中京大学)

P1-01 胸部 CT 像からの COVID-19 症例の自動分類手法

○小田 昌宏¹²,鄭 通²,林 雄一郎²,大竹 義人³⁴,橋本 正弘⁵,明石 敏昭⁶,森 健策²¹⁴
¹名古屋大学情報連携推進本部情報戦略室、²名古屋大学大学院情報学研究科、³奈良先端科学技術大学院大学先端科学 技術研究科情報科学領域、⁴国立情報学研究所医療ビッグデータ研究センター、⁵慶應義塾大学医学部放射線科、⁶順天 堂大学医学部放射線医学講座

本稿では、胸部 CT 像をもとに新型コロナウイルス感染症(COVID-19) 症例か否かを自動分類する手法を述べる.2019年 12月に発生した COVID-19は発生から1年以上経過した現在も多数の感染者を生じており、世界の人々の生活様式を変化させ 経済活動に悪影響を与えている.COVID-19を含むウイルス性肺炎の診断において胸部 CT 像が有用であり、医師の診断支援 のため、CT 像を用いた COVID-19 診断支援手法の確立が求められる。本稿では深層学習を用い、CT 像から COVID-19 の画像 所見典型度を自動推定する手法を提案する。まず COVID-19により生じる異常陰影に対応した肺野セグメンテーション手法を 用いて肺野の正常及び異常陰影領域をセグメンテーションする。セグメンテーション結果を用いて肺野ボリューム画像を作成 し、これを COVID-19 典型度推定 3D CNN を用いて自動分類する。3D CNN は局所及び大域的な画像特徴を抽出可能な構造と した。また、3D CNN の学習においては 3D 画像変形を用いたデータ拡張を利用し、多様な肺野形状に対応可能とした。提案 手法を COVID-19 症例の胸部 CT 像に適用し、COVID-19 の画像所見典型度を良好に推定可能であることを確認した。

P1-02 極度に不均衡な頭部 CT 画像データセットにおける Self-attention class balanced DenseNet_ LSTM 構造を用いたくも膜下出血 CT 画像分類

〇魯 仲陽¹, 小田 昌宏²¹, 林 雄一郎¹, 胡 涛¹, 伊東 隼人¹, 渡谷 岳行³, 阿部 修³, 橋本 正弘⁴, 陣崎 雅弘⁴, 森 健策^{1.5.6}

¹名古屋大学大学院情報学研究科.²名古屋大学情報連携推進本部情報戦略室.³東京大学医学部附属病院放射線科. ⁴慶應義塾大学医学部放射線科学教室.⁵名古屋大学情報基盤センター,⁶国立情報学研究所 医療ビッグデータ研究センター

Deep learning algorithms have obtained many remarkable achievements in the computer vision field, whereas their poor capacity toward the data-imbalanced challenge cannot make people satisfied. In this paper, we propose a novel class imbalanced network framework with self-attention mechanism, SACB-DenseNet_LSTM to handle the data-imbalanced issue in the mission in order to classify the subarachnoid haemorrhage CT scans from the extremely imbalanced brain CT dataset. In our proposed framework, we utilize a self-attention mechanism to build up a self-attention block to estimate the feature importance of individual feature extracted by the feature extractor (here we employ DenseNet-121 with class balanced loss as our feature extractor). Moreover, we integrate the feature importance vector with the input feature, and fed them to the LSTM sequentially to acquire the consecutive information from adjacent CT scans for further improving the classification performance. We performed experiments on an extremely imbalanced brain CT dataset, including 33 cases with subarachnoid haemorrhage and 2519 cases of normal cases. Our experiments showed that after inducting the self-attention mechanism, in contrast to the previous work, the classification performance improved from 0.726 F-value to 0.734 F-value.

P1-03 AI による乳がん腫瘤影の良悪性鑑別処理における画質の相違が CAD 性能に与える影響の分析

○工藤 颯馬¹, 安倍 和弥¹, 武尾 英哉¹, 永井 優一², 縄野 繁³

¹神奈川工科大学,²国立がん研究センター東病院,³新松戸中央総合病院

近年,医療の現場では体内撮影装置の高精細化により生成される医用画像の量が増大し,そしてそれを読影する医師の負担 も急増している.この問題を解決するために,診断の補助としてコンピュータ画像支援診断(CAD)の機能を向上させる必要 がある.

そのひとつとして、本研究では乳がん腫瘤を対象とし、AI 深層学習の1つである転移学習と事前に画像の回転,明度,コント ラストを調整した画像を用いて学習させる画像拡張を用いて乳がん検診における乳がん腫瘍良悪性鑑別処理 AI の開発を行い, 検診画像と診断画像の比較により画像診断の精度改善にどのような要素が必要か分析した.

開発した良悪性鑑別処理の精度は85%と医師の判別と同等の値になったが,診断画像の90%には及ばず,その理由として画 像比較によりノイズとコントラストの悪さが影響していると判明し,これらの改善を行うことが検診におけるより精度の高い CAD システムの構築に必要だと分析した.
P1-04 マルチチャンネル化と M-Net を用いた嚥下時 X 線動画からの頸椎椎間板の抽出と評価実験

○郡司絵莉華¹, 目片幸二郎², 滝沢 穂高¹, 工藤 博幸¹ ¹ 筑波大学 情報理工学位プログラム,²四條畷学園大学 リハビリテーション学部 作業療法学専攻

栄養低下や誤嚥を引き起こすことがある嚥下障害の詳細なメカニズムは未だに解明されていない. そこで本研究では頸部構 造物を動態解析し,嚥下障害との関係を調査することを目指し,嚥下時X線動画(videofluorography, VF)から頸椎椎間板の抽 出を行う. VFの各フレームに,モルフォロジー演算とRANSACによる放物線あてはめ処理などを適用し,頸部領域を抽出する. VFの各フレームは濃淡画像であり,45種類の線形,非線形の画像フィルタを適用することで同数の特徴画像を生成し,その 中から3種類を選択し,カラー画像のRGBチャンネルに入力することでマルチチャンネル化画像(MC 画像)を作成する. MC 画像をセグメンテーション用のネットワークである M-Net に適用し,画素値ベースのF値をエネルギ値として,焼きなまし法 に基づくフレームワークによって,特徴の組み合わせの最適化を行う.本報告では,3-fold 3-phase cross validation という評価 実験手法も提案し,実際の58 症例のVF に適用した実験結果を示す.

P1-05 EfficientNet を用いたアンサンブル学習に基づく放射線皮膚炎グレード判定手法の開発

○和田 清隆¹²,渡邊 睦²,新野 将史²,野口 康介²,获野 尚¹ ¹メディポリス国際陽子線治療センター,²鹿児島大学大学院 理工学研究科

放射線皮膚炎の評価は、臨床で用いる有害事象共通用語規準(Common Terminology Criteria for Adverse Events: CTCAE) を用いて軽度のグレード1から有害事象による死亡の5まで5段階に分類される.しかし、視覚的な評価に基づいているため 個人の経験や知識に左右される課題がある.そこで我々は、少ない症例に対して人工症例画像を生成し、深層学習を用いた放 射線皮膚炎のグレード判定システムを構築してきた.本稿では、近年提案された EfficientNet モデルを用いた複数モデル方式 に基づき、アンサンブル学習に基づく最終的な分類を行う手法を新たに開発し、放射線皮膚炎グレード判定の曖昧性解消、シ ステムの精度,汎化性能の向上を行った研究について述べる. EfficientNet B0 ~ B7を用いて,画像解像度とデータ拡張法(Data Augmentation, Rand Augmentation および人工症例画像生成)の条件を変えた学習モデルを複数作成し、グレード毎と全体の 正答率を比較することにより、有効性を評価した.

P1-06 畳み込み自己符号化器つき Deep Autoencoding Gaussian Mixture Model を用いた胸部 CT 画像 における異常検知

○栗林 雅刀¹,間普 真吾¹,若本 亮佑¹,呉本 尭²,木戸 尚治³ ¹山口大学,²日本工業大学,³大阪大学

近年,デジタル医用画像の撮影技術が発展し,医師一人当たりの読影を要する医用画像数は増加傾向にある.そこで,医師 の負担を軽減するために,深層学習を用いて医用画像を分類することで,医療診断の支援をする研究が進められている.深層 学習には、ラベル付けされた大量の画像が必要となるが,疾患のある画像は疾患のない画像に比べて数が圧倒的に少ないため, 学習に必要な異常データが十分に得られないという問題がある.本研究では,疾患のない画像のみを学習データに用いて,正 常領域を表現する分布をつくり,正常領域から外れたデータを異常とみなすシステムを構築する.画像の特徴抽出には畳み込 み自己符号化器,クラスタリングには混合正規分布モデルを用い,この2つを同時に学習する畳み込み自己符号化器つき Deep Autoencoding Gaussian Mixture Model という異常検知手法を胸部 CT 画像の異常陰影検知に応用する.

P1-07 骨シンチグラムを用いた診断支援システムの再学習による性能変化

○島田 夏帆¹, 斉藤 篤¹, 大崎 洋充², 東山 滋明³, 河邊 譲治⁴, 中岡 竜介⁵, 清水 昭伸¹ ¹東京農工大学大学院工学研究院,²群馬県立県民健康科学大学大学院診療放射線学研究科,³大阪市立大学医学部附属病院核医学科,⁴大阪市立大学医学部附属病院核医学科,⁵国立医薬品食品衛生研究所医療機器部

本報告では、骨シンチグラムを用いた診断支援システムの再学習による性能変化について報告する. 医療現場で実際に使用 されている高集積検出処理のネットワークを用いて、市販後学習による性能変化をシミュレーションする.まず、手持ちのデー タを市販前と市販後に仮想的に分割する.次に、市販前の学習データを用いてネットワークを学習する.さらに、市販後の学 習データを用いて再学習を行い、市販後のテストデータを用いて性能の変化を評価する.複数の施設のデータを用いて、市販 前と市販後のデータの分割法を変えて繰り返し評価を行うことで、性能変化をもたらした要因や再学習の有用性について議論 する.

P1-08 CvcleGANを用いた甲状腺組織画像のドメイン変換と識別システムへの応用

○一氏 良仁¹, 間普 真吾¹, 八田 聡美², 稲井 邦博², 木戸 尚治³ ¹山口大学, ²福井大学, ³大阪大学

近年,深層学習を用いた医用画像の識別について多くの研究がなされており,臨床応用へ向けたさらなる性能向上が期待さ れている.しかし,深層学習には大量の教師ラベル付きデータが必要であるものの,特に医療分野では大量の教師データを用 意することは難しく,一つの医療施設のみのデータでは不十分であることも多い.したがって,複数の施設から得たデータの 活用が必要である.複数施設のデータを使用する場合,施設ごとのデータの取得条件等の違いにより特徴が異なるため,デー タを一定の基準にそろえる正規化を行うことで,データの特徴の分布を一致させる必要がある.本研究では、2つの施設から取 得した甲状腺組織画像を対象に,甲状腺がんの識別を行うシステムを構築することを目的とし,甲状腺組織画像に CycleGAN によるドメイン変換を適用することで、2つのデータセットの正規化を行い,深層ニューラルネットワークの学習および識別 に利用した.識別性能を評価する実験の結果,ある施設Aのデータセットと、CycleGANによる正規化を行った施設Bのデー タセットを混合し訓練データとした場合に学習の効果が高まることを明らかにした.一方,施設Aのデータセットで学習した ネットワークで,正規化した施設Bのデータセットを識別するには改善が必要であることが明らかとなった.

P1-09 アテンション誘導を利用した胸部 CT 画像による COVID-19 の診断支援

○高手山雄亮¹, 斉藤 篤¹, 橋本 正弘², 大竹 義人³, 明石 敏昭⁴, 清水 昭伸¹

¹東京農工大学大学院工学研究院,²慶應義塾大学医学部放射線科(診断),³奈良先端科学技術大学院大学,⁴順天堂大 学 放射線診断学講座

本報告では、COVID-19の症例を含む胸部 CT 画像を対象にした疾病クラスの分類について述べる.単純な Convolutional Neural Network ベースの手法では、広範囲の特徴を抽出することが難しく、COVID-19の典型的な特徴である両側性の病変な どを見落としてしまう可能性がある.提案手法は、病変を含むスライスのみを選択するスライス選択ネットワークと、選択さ れたスライスからクラス分類を行う病変分類ネットワークで構成される.本報告では病変分類ネットワークについて述べるが、これは、ResNet18 をベースとして Attention 誘導機構を組み合わせることで、ネットワークの注目部位を COVID-19 の特徴的 な部位に誘導可能なモデルである.247 症例を用いて性能を評価した結果を示し、提案法の有効性について議論をする.

P1-10 口腔診断支援のための舌の色特徴に基づく機械学習を用いた疾患予測

○野口 桂冴¹, 斎藤 一郎², 吉村裕一郎³, 並木 隆雄⁴, 渡辺 悠紀⁴, 中口 俊哉⁵

¹千葉大学大学院融合理工学府基幹工学専攻医工学コース,²鶴見大学歯学部 病理学講座,³富山大学附属病院放射線部, ⁴千葉大学大学院医学研究院和漢診療学,⁵千葉大学フロンティア医工学センター

舌診は舌特徴を捉え, 簡便かつ非侵襲的に健康状態を把握する診断法であるが, 医師の主観や経験に依存するため, 定量化 が求められている.本研究ではドライマウスや口腔疾患を誘発するシェーグレン症候群(SS)について, 診断支援を目的に舌 画像に基づく解析を行った. SS 患者 19 名を含む, 計 60 名において舌画像を取得した. 舌形状規格化のため, 舌を5角形に近 似し同一の解析領域を指定した. 舌色を CIELAB 空間で取得し Support Vector Machine, Random Forest, Logistic Regression を訓練した. SVM が最も高い性能を示し, 交差検証における平均感度 76.5%, 特異度 61.6%となった. 本手法は現行の診断基 準の一つである唾液量検査と比較し同等の診断性能を示しており, 衛生面と簡便性において優位であるため, コロナ禍に適し た新たな診断法となることが示唆された.

P1-11 腸閉塞・イレウスの病変箇所特定における診断支援システムの精度評価

〇小田 紘久¹,林 雄一郎²,北坂 孝幸³,玉田 雄大⁴, 滝本愛太朗¹,檜 顕成¹,内田 広夫¹, 鈴木耕次郎⁵,小田 昌宏⁶,森 健策^{2,7,8}

¹名古屋大学大学院医学系研究科,²名古屋大学大学院情報学研究科,³愛知工業大学情報科学部,⁴名古屋大学医学部, ⁵愛知医科大学放射線科,⁶名古屋大学情報連携統括本部,⁷名古屋大学情報基盤センター,⁸国立情報学研究所医療ビッ グデータ研究センター

本稿では、腸閉塞・イレウス診断支援システムを用いて、それらの疾患の病変箇所を探索する場合における.腸管抽出結果の探索支援への有効性に関する精度評価を行う.腸閉塞ないしイレウスの画像診断において、医師は CT 像において腸管の走行を追跡し、病変箇所、すなわち絞扼や麻痺などのため内容物が流れにくくなっている病変箇所を探索する.腸閉塞・イレウス診断支援システムを用いることで、CT 像上である程度の太さをもった区間がそれぞれ抽出され、立体的に可視化される.このとき、区間のひとつの端点が病変箇所を表す可能性が高い.このため本稿では、病変箇所の探索における、腸管抽出結果の探索支援への有効性に関する精度評価を行う.ここで、抽出結果の端点が病変箇所に近接しており、また端点の数が少ないほど、医師は容易に病変箇所を探索することが可能である.そこで、抽出結果の端点と病変箇所をさす点との最短距離ならびに、症例ごとに生成される端点の数を計測する.実験では、CT 像 110 症例において腸管領域を抽出し、その結果を評価した.病変箇所は症例ごと手動でアノテーションした.多くの症例において,病変箇所の至近まで腸管領域が抽出できていることを確認した.

P1-12 手術標本及び生検標本の H&E 染色組織画像を用いた畳み込みニューラルネットによる肝細胞癌 の識別

○高木 弦¹,山口 雅浩¹,阿部 時也²,橋口 明典²,坂元 亭宇² 1 東京工業大学工学院, 2 慶應義塾大学医学部病理学教室

本研究では、H&E 染色病理標本のデジタルスライド画像に対して畳み込みニューラルネットワーク (CNN) による肝細胞癌 と非癌の識別を行った. CNN として VGG16 を採用し, 高解像度の病理画像を分割して学習 させ, その結果を統合する方法 について検討するとともに、学習率などについて改善を行い、先行研究に対する精度向上を図った。2048 × 2048 画素の手術 標本及び生検標本の画像それぞれ約 1,000 枚に対して 5-fold の交差検証を行った結果,先行研究において特徴量ベースの SVM で約 90% の精度であったデータに対して,正解率約 97% と極めて高い精度を得ることができた.このとき,生検標本と手術 標本の両者の画像を合わせて教師データとした際に、高精度でロバストな識別器が作成できることが示された、さらに、生検 標本と手術 標本の両方を学習させた際に, CNN が両者に共通する特徴を抽出しているのかについても考察を行った.

骨粗鬆症 CT 検診におけるコンピュータ診断支援システムの開発 P1-13

○宮崎 優太¹, 鈴木 秀宣¹, 河田 佳樹¹, 仁木 登¹, 松元 祐司², 土田 敬明², 楠本 昌彦², 金子 昌弘3

¹徳島大学創成科学研究科理工学専攻光システムコース、²国立ガン研究センター中央病院、³保険会館クリニック

骨粗鬆症とは、高齢化に伴い骨の強度が低下して骨折しやすくなる状態を指す、骨粗鬆症になり骨が脆くなると、つまづい て手や肘をつく、くしゃみをしたなどの僅かな衝撃で骨折してしまうことがある. 前腕部, 大腿部付近の骨が特に骨折しやす くなり、寝たきりの大きな原因になっている、骨粗鬆症の予防には早期発見・治療が最も重要となってくる、 本研究では CT 検診によって撮影された画像から骨粗鬆症を提案する診断支援システムの Web ベースでの開発を目指す

P1-14 機械学習による軟部腫瘍 MRI 画像の良悪性判別

○ハックハスナイン^{1,4}, 弘實 透³, 橋本 正弘¹, 東田 直樹¹, 有田 祐起¹, 庄野 寛太², 金内友里恵², カーントラン¹, 植竹 望4, 陣崎 雅弘1

¹ 慶應義塾大学医学部整形外科,² 慶應義塾大学工学部,³ 慶應義塾大学医学部整形外科,⁴ G E ヘルスケアジャパン(株)

Standardized protocol for how MR imaging should be used for patients with suspected soft tissue sarcoma (STS) disease is widely varied among institutions. Image-level detection of malignancy of STS is a challenging task for radiologists and orthopedic surgeons. In this research work, we investigate which MR imaging protocol or protocols contains most malignancy signature of STS.

Retrospective Gd-T1W(gadolinium-enhanced), T2W, FS-T1W, DWI, T1W MRI of 80 patients with histologically proven subtypes of STS were examined. We grouped UPS, myxofibrosarcoma, myxoid liposarcoma, leiomyosarcoma, synovial sarcoma, dedifferentiated liposarcoma subtypes as malignant and schwannoma, lipoma, ALT, myxoma, desmoid subtypes as benign. After manual segmentation and preprocessing texture-based radiomics features and deep CNN features from pretrained VGG19 model along with feature reduction methods and machine learning modeling ware used for malignancy detection performance comparison. Only combined imaging features of Gd-T1W and T2W were used due to its most sample size on the cohorts. 14 patient Images scans were used as test dataset to evaluate our claims.

Detection accuracy of using radiomics feature of Gd-T1W and T2W+Gd-T1W was 86% and 78% respectively. Whereas CNN features of same MR images was found to be 93%

In our investigation Gd-T1W scanned MR image contains most malignancy signature and CNN features of image can most accurately detect STS malignancy. This research can optimize the MR scan protocol for STS disease and the risk of oversight of malignant disease by radiologists and orthopedic surgeons.

P1-15 転移学習による手部骨髄浮腫の識別精度向上についての研究

冬平¹,小田 昌宏^{2,1},片山 健策 1,4,5 耕³, 大久保学宣³, 森 ○潘 ¹名古屋大学大学院情報学研究科,²名古屋大学情報連携推進本部情報戦略室,³片山整形外科リウマチ科クリニック, ⁴名古屋大学情報基盤センター,⁵国立情報学研究所医療ビッグデータ研究センター

Bone marrow edema (BME) is a main symptom of early rheumatoid arthritis (RA). By finding bone marrow edema, early diagnosis and treatment for RA will be possible. Although we can observe BME from Magnetic Resonance Imaging (MRI) scans, the cost of taking MRI images is quite high. To develop an economic solution, we aim to use X-ray images to recognize BME. While previous work verified the feasibility of using CNN to classify BME, we encountered the problem of the shortage of data. To overcome this problem and achieve a better performance, we propose a new method that uses transform learning from the RSNA bone age dataset. We use a pretrained weight from RSNA bone age regression task. From the result, the additional knowledge from the RSNA bone age dataset boost the performance of classification.

P1-16 肺胞蛋白症 CT 画像における重症度分類器構築の検討

○山浦賢太郎¹²,小田 未来⁴,石井 晴之⁴,北村 信隆⁵⁶,田中 崇裕⁵⁶,中田 光⁵⁷,野中 尋史¹³ ¹長岡技術科学大学,²技術支援センター,³情報・経営システム工学専攻,⁴杏林大学医学部,⁵新潟大学医歯学総合病 院,⁶臨床研究推進センター,⁷高度医療開発センター

肺胞蛋白症では肺胞腔内及び終末気管支内にサーファクタント様物質が異常に貯留することで労作時呼吸困難等の症状が現 れる.診断時の CT 画像の主要所見としては、すりガラス影及び小葉間隔壁肥厚等がある. 無症状でも本症では、CT の肺野異 常陰影は明瞭に観察されるため、他の臨床パラメーターよりも病態を鋭敏に反映しているのではないかと考えられている. そ こで本研究では、主として CT 値濃度に基づく肺胞蛋白症 CT 画像の重症度分類器の構築を検討した. 肺野を様々な症例及び 撮影条件で事前学習した U-net で立体抽出した後、3次元で立方格子に分け、立方体単位で CT 値の濃度統計量、ヘシアン固 有値を求め、これらを特徴量として k-means クラスタリングを行った. 管理区分重症度毎に比較検討し、重症化するにつれ CT 値濃度は非常に淡い濃度から徐々に濃くなり、すりガラス影は濃度でレベル分けする必要があることが示唆された.

P1-17 歯科パノラマ X 線画像における分岐を含む SSD とヒューリスティクスによる歯列の認識手法

○森下 拓水¹, 村松千左子², 清野 雄多³, 周 向栄³, 高橋 龍⁴, 林 達郎⁴, 西山 航⁵, 原 武史³, 有地 淑子⁶, 有地榮一郎⁶, 勝又 明敏⁵, 藤田 広志³
¹岐阜大学大学院自然科学技術研究科知能理工学専攻, ²滋賀大学データサイエンス学部, ³岐阜大学工学部電気電子・

「岐阜人学人学院自然科学技術研究科知能理上学専攻」。滋賀入学データサイエンス学部,「岐阜人学上学部電気電子・ 情報工学科,⁴メディア株式会社,⁵朝日大学歯学部口腔病態医療学講座歯科放射線学分野,⁶愛知学院大学歯学部歯科 放射線学講座

本研究の目的は、歯科パノラマX線画像を解析し、読影に必要な情報を自動抽出して歯科医師の診断補助に寄与することで ある.その初期段階として、950 症例を用いて画像内の歯牙の検出と歯種の分類を行った.Single Shot Multibox Detector (SSD) に分岐を与えたネットワークを提案し、歯種を区別しない1クラス検出と、16クラス(上顎と下顎で区別した中切歯、側切歯、 犬歯、第一小臼歯、第二小臼歯、第一大臼歯、第二大臼歯、第三大臼歯)検出を行った.また、これらの結果を用いて統合処 理や偽陽性削除などの後処理を行い、左右を区別した32クラスに分類した.結果、32 歯種について、検出率 98.0%、誤検出数 0.43 個/画像、分類率 96.7% となり、提案手法の有効性が示された.

P1-18 3D-Attention Branch Networks を用いた脳 MRI 画像における低悪性度グリオーマの 1p/19q 共 欠失推定法

○田中 大貴¹, 檜作 彰良¹, 中山 良平¹ ¹立命館大学大学院 理工学研究科

本研究では、腫瘍領域の強調を行う Attention 機構を導入した 3D-Attention Branch Networks を用いて、脳 MRI 画像上の 低悪性度グリオーマにおける 1p/19q 共欠失の有無を推定する手法を開発した.実験試料は、The Cancer Imaging Archive が 公開している 159 症例(1p/19q 共欠失有:102 症例, 無:57 症例)の T2 強調画像を用いた.提案手法は、まず、特徴抽出 機構より抽出した特徴マップに対し Attention 機構を適用し、腫瘍領域に注目した新たな特徴マップを生成した.そして、こ の特徴マップを用いた予測機構により、1p/19q 共欠失の有無を推定した.提案手法を実験試料に適用した結果、正答率 78.0% (124/159)、感度 79.4%(81/102)、特異度 75.4%(43/57)が得られ、提案手法が 1p/19q 共欠失の有無の推定に導入できる可能 性がある.

P1-19 体幹部定位放射線治療を受けた非小細胞肺癌患者における癌の進行予測のためのレディオミクス 解析

 ○兒玉 拓巳¹,有村 秀孝²,二宮 健太¹,白川 友子³,吉武 忠正⁴,塩山 善之⁵
 ¹九州大学大学院医学系学府保健学専攻,²九州大学医学研究院保健学部門,³国立病院機構 九州がんセンター,⁴九 州大学病院 放射線科,⁵九州国際重粒子線がん治療センター

[Purpose] The treatment options for patients with Stage I lung cancer are surgery and stereotactic body radiation therapy (SBRT), because their outcomes were almost the same according to a survey using a Japanese multi-institutional database (Onishi et al, 2011). Therefore, it is important to choose the more appropriate treatment for each patient based on pretreatment prediction of patients' prognoses. This study explored a radiomics prediction for progression using planning computed tomography (CT) images in non-small cell lung cancer (NSCLC) patients treated with SBRT. [Method] SBRT treatment planning CT images with dimensions of 512 × 512 pixels, a field of view of 400 mm, and a slice thickness of 3 mm were selected from 125 NSCLC patients. A total of 486 radiomics features were calculated from original and wavelet decomposition images using histogram and texture analysis. Patients were divided into 88 training cases and 37 test cases, and time-to-progressions (TTPs) for the patients were analyzed by a Kaplan-Meier (KM) method. Signatures were constructed with 8 features selected according to the absolute values of the regression coefficients in a Cox proportional hazard model. Those patients were stratified into high-risk and low-risk groups using the medians of radiomics scores of the signatures. Radiomics scores were derived from the feature values and their corresponding coefficients. The prognostic predictability was evaluated using statistically significant differences (p-values, log-rank test) between KM curves of high- and low-risk groups. [Result] The p-values between the KM curves between the two groups were 5.98×10^{-5} for training cases and 1.95× 10-2 for test cases. [Conclusion] This study demonstrated that the high- and low-risk of progression after SBRT in NSCLC patients can be predictable by using image features derived from pretreatment planning CT images.

メイン会場

一般演題セッション2

9:00~10:40 P2 [CADe, 計測, Segmentation]

座長:清水 昭伸(東京農工大学)菅 幹生(千葉大学)

P2-01 3D CNN を用いた小児脳 CT 画像からの脳成長年齢推定研究

○森田 蓮¹, 安藤 沙耶², 藤田 大輔¹, 新居 学¹, 安藤久美子³, 石藏 礼一³, 小橋 昌司¹ ¹兵庫県立大学 大学院 工学研究科, ²兵庫県立尼崎総合医療センター, ³神戸市立医療センター中央市民病院

未発達や未熟児などの脳の発達を評価するために, 脳画像が用いられている.しかし, 脳画像から小児の脳の発達度を定量 的に評価する方法はなく,現状では医師の経験に基づいて診断が行われている.そのため,読影可能な医師の不足, 定量性の 欠如が問題である.本研究では,小児の脳 CT 画像から脳の発達年齢を予測する手法を提案する.この手法では,CT 画像から 顕蓋領域を抽出し,姿勢と位置の補正を行う.また,3次元畳み込みニューラルネットワーク(3D CNN)を用いて CT 画像 から特徴を抽出し,全結合層で脳の発達年齢を予測する新しいネットワークモデルを提案し,予測に使用する.このモデルの 性能を,0歳から3歳までの脳神経医学的に異常のない小児60人を用いて評価した結果,予測年齢と患者の実年齢の間の二乗 平均平方根誤差は7.80(月)で,相関係数は0.801であった.

P2-02 自己符号化器を用いた慢性閉塞性肺疾患における身体活動性の予測に関する研究 ○中原 祥喜¹,間普 真吾¹,平野 綱彦¹,村田 順之¹,土居 恵子¹,松永 和人¹

¹山口大学

慢性閉塞性肺疾患(COPD)は、タバコ煙などの有害粒子を長期に吸入することで生じる肺の炎症である. COPD に罹患す ると、患者は身体活動性が低下し日常生活が制限される(身体活動性障害).そして身体活動性障害が COPD の悪化につなが るという循環に陥ってしまう.しかし、身体活動性障害を引き起こしている根本的な原因は、未だ明らかになっていない.ま た、検査にはコストがかかるという問題もある.そのため、毎回全ての検査を行うことはできず、簡易的な検査しか行うこと ができない.そこで本研究では、各患者の検査項目から、身体活動性を予測し身体活動性を低下させる要因を明らかにするこ とを目的とする.これにより、身体活動性低下の要因に集中して治療できるようになるため、コストを削減することができる. 具体的には、一般的な3層あるいは4層パーセプトロン(MLP)、および自己符号化器による事前学習を行う3層あるいは4層 MLPを構築し、身体活動性が低下しているクラスと低下していない正常のクラスの識別を行い、その識別率を評価するとと もに、識別に重要な特徴を明らかにする.

P2-03 機械学習・深層学習を用いた手 X 線画像におけるリウマチ mTSS 推定法

○中津 康平¹,盛田 健人²,藤田 大輔¹,小橋 昌司¹ ¹兵庫県立大学 工学研究科,²三重大学

現在、関節リウマチの国内患者数は60万人から100万人であると推定され、毎年数万人ずつ増加している、関節リウマチの 診断には、手足のX線画像から算出するmTSS(modified Total Shape Score)が標準的な診断法として用いられているが、診 断の問題点として、医師の目視による主観的診断であること、スコアの判断箇所が多いことが挙げられている。そのため、医 師の負担軽減や診断の迅速性・正確性向上のためにmTSスコアを自動で評価するコンピュータ支援診断(CAD)システムの 開発が期待されている.本研究では、手関節リウマチのmTSS予測法に関して、リッジ回帰(RR)を用いた手法と、畳み込みニュー ラルネットワーク(CNN)の3つのモデル(VGG16, DenseNet201, Xception)を用いた手法の比較を行う、4手法の比較の ために、90名のRA患者を対象に、手のX線画像を用いた実験を行った。実験結果は、erosionのmTSS予測はRR、JSNの mTSS 予測はVGG16で最も良い結果が得られた。

P2-04 機械学習によるコルポスコープ画像における白色上皮領域検出精度向上のための検討

○篠原 寿広¹,藤井 一輝¹,村上 幸祐²,中迫 昇¹,松村 謙臣² ¹近畿大学生物理工学部生命情報工学科.²近畿大学医学部産科婦人科学教室

子宮頸癌の診断において、コルポスコープとよばれる拡大鏡を用いて、目視による検査が行われている.本検査はコルポス コピーとよばれ、子宮頸癌の確定診断を行う組織診のための組織を採取する部位を決定する非常に重要な検査である.しかし ながら、コルポスコピーの精度は、検査者の知識と経験に大きく依存する.そこで、われわれは、機械学習を用いたコルポス コピー支援を検討している.これまでに初期検討として、子宮頸部の代表的な病変の一つである白色上皮を対象に、分割した コルポスコープ画像に対して、Convolutional Neural Networkを用いて、白色上皮の判別を行った.本稿では、判別精度向上 のため、画像の分割条件について詳細に検討した.また、分割画像を元に戻し、周辺の分割画像の判別結果を重み付き平均す ることにより、判別精度が向上することを確認した.

P2-05 覚醒下脳腫瘍摘出術における手術工程同定のための深層学習を用いた顕微鏡画像からの特徴量抽 出手法の提案

○佐藤 生馬¹,田村 学²,山口 智子²,藤野 雄一¹,吉光喜太郎²,村垣 善浩²,正宗 賢² ¹公立はこだて未来大学 システム情報科学部.²東京女子医科大学 先端生命医科学研究所

覚醒下脳腫瘍摘出手術では、最大限の腫瘍摘出と最小限の術後合併症リスクを実現するため、患者の脳構造や機能を把握し、 腫瘍摘出領域を決定することが重要である.しかし、脳腫瘍の種類・位置や脳機能位置、患者の年齢などによって適した手術 方針や工程は異なる.よって、執刀医は術中に患者の脳構造や機能を把握し、適切な腫瘍摘出領域を決定する必要がある.こ の執刀医の意思決定は、執刀医の知識や経験に依存しているため、手術スタッフの作業の最適化や若手医師の教育には手術工 程の可視化と解析が重要である.また、手術工程の可視化と解析の実現のためには、手術工程の自動同定が必要であると考え られる.本研究では、手術工程の可視化と解析に向けた手術工程同定のために、手術顕微鏡動画から深層学習を用いた特徴量 抽出手法の提案を行う.

P2-06 Transformer ベースのモデルを用いた肺領域における所見文からの疾患名抽出

○岡崎 真治¹, 林 雄一郎¹, 小田 昌宏²¹, 橋本 正弘³, 陣崎 雅弘³, 明石 敏昭⁴, 青木 茂樹⁴, 森 健索^{15.6}

¹名古屋大学大学院情報学研究科,²名古屋大学情報連携推進本部情報戦略室,³慶應義塾大学医学部放射線科,⁴順天 堂大学医学部放射線医学講座,⁵名古屋大学情報基盤センター,⁶国立情報学研究所医療ビッグデータ研究センター

本稿では、Transformer ベースのモデルを用いた肺領域における所見文からの疾患名抽出について述べる。画像処理において、 一般的にはモデルの性能向上のため多くの教師データが必要である。しかし、医用画像のアノテーションには膨大なコストが かかる問題がある。そこで、放射線科医が CT 画像を撮影した際に作成し、患者の状態を詳細に記載している所見文を利用する。 本研究ではラベル付きデータセット構築の初期検討として、Transformer ベースの手法を用い、所見文から肺領域における疾 患分類を行うことで疾患名が付与された画像データセット構築を行った。実験の結果、Transformer ベースの手法を用いるこ とで良好な結果を得られた。

P2-07 取り下げ

P2-08 3次元 CT 画像を用いた模擬ボーリング調査に基づく脆弱性骨盤骨折検出法による特徴量の検討 ○山本 侃利¹,藤田 大輔¹, ラシェドゥーララーマン¹,八木 直美²,林 圭吾³, 圓尾 明宏³.

1兵庫県立大学 工学部, 2姫路獨協大学, 3製鉄広畑記念病院

骨粗鬆症による高齢者脆弱性骨盤骨折は、外傷に因らず、自覚症状が顕著でなく、CT 画像上でその検出が容易ではないため、 発見後の治療が遅れ、転位が進行し、機能的予後回復が得られない場合がある。そこで、医師の診断能向上のため、CT 画像か ら骨盤脆弱性骨折を自動的に検出する医師の診断支援システムが求められている。従来手法では、単純X線画像やCT 画像に よる2次元画像解析に基づくため、3次元的に分布する微小な脆弱性骨折の検出が困難であった。そこで我々は、新しい手法 として、3次元 CT 画像を用いて、骨表から骨内部にかけて3次元的に骨折有無を探索するボーリング調査法を模した自動骨盤 骨折検出法 (BSFD 法; boring survey based fracture detection)を提案した。本研究では、BSFD 法における特徴量抽出法につ いて検討する。BSFD 法では、3次元 CT 画像から骨表同値面を求め、同値面上の各点に CT 値で構成される3次元特徴ベク トルを割り当て、学習済みの3次元畳み込みニューラルネットワーク (CNN) モデルにより、各点において骨折確率を求める。 各点でアノテーションされた骨折領域からの3次元 Chamfer 距離から求められた骨折確率を用いて、CNN を学習する。ここ で、3次元特徴ベクトルとして、骨内領域、骨内外領域、加えて領域範囲の拡大を検討して、検出性能を評価する。提案手法 を110人の被験者のデータで検証した結果、学習データでは AUC0.90、評価データでは AUC0.84 を確認した。

村津 裕嗣³, 小橋 昌司¹

P2-09 膵臓の超音波内視鏡から作成した2次元ヒストグラムを用いた膵癌自動分類

〇竹内 祐慈¹, 岩佐 悠平², 岩下 拓司³, 原 武史⁴, 周 向栄⁴, 上村 晋也³, 三田 直樹³, 清水 雅仁³

¹岐阜大学大学院 自然科学技術研究科 知能理工学専攻,²岐阜市民病院 消化器内科,³岐阜大学病院 第一内科,⁴岐阜 大学 工学部 電気電子・情報工学科

現在最も有用性が高い診断方法とされる超音波内視鏡下穿刺吸引生検では、出血や膵炎、穿刺ラインへの播種など、患者に 負担がかかることが知られている。本研究の目的は、撮影された超音波内視鏡から深層学習を用いて膵癌の診断を行い、患者 の負担の少ない診断法を確立することである。また膵癌の診断には、造影超音波内視鏡の時系列の造影パターンを横軸を時間軸、 縦軸を平均輝度値として表した time intensity curve(TIC)を用いた診断が診断能の改善に有用とする報告もある。そこで本手 法では、TIC に1 フレームごとの輝度値の頻度の情報を与えた 2 次元ヒストグラムを作成し、学習に用いることで分類精度の 向上を試みた。

P2-10 多段の画素異常検知による FDG-PET/CT 上のがん骨転移候補検出

〇山口 明乃¹, 根本 充貴², 甲斐田勇人³⁴, 木村 裕一¹, 永岡 隆¹, 山田 誉大⁴, 花岡 宏平⁴, 北島 一宏⁵, 槌谷 達也⁶, 石井 一成³⁴

¹近畿大学大学院 生物理工学研究科,²近畿大学 生物理工学部,³近畿大学 医学部 放射線医学教室 放射線診断学部門, ⁴近畿大学高度先端総合医療センター PET 分子イメージング部,⁵ 兵庫医科大学 放射線医学講座,⁶ 兵庫医科大学病院 放射線技術部

FDG-PET/CT 画像診断の支援技術として、マハラノビス距離を用いた異常検知と one-class support vector machine (OCSVM) による異常検知を骨画素に適用したがん骨転移候補の検出法を提案する.

CT 値の閾値処理等で自動抽出した骨領域内全画素に対し、2つの異常検知を行う.まず CT 値と SUV の 2 次元空間上で正 常骨画素からのマハラノビス距離を計測し、距離の閾値処理により骨転移が疑われる領域を粗抽出する.次に CT 値, SUV, CT 値と SUV の平均曲率とガウス曲率、マハラノビス距離の計7 特徴量を入力とした OCSVM による異常検知を行う.得られ た異常の極大点を骨転移候補とする.

FDG-PET/CT 画像 40 例を用いて学習した異常検知処理を 46 の骨転移病変を含む FDG-PET/CT 画像 20 例に適用した結果, 感度 80.4%(38/46), FP 数 314.6 点 / 例の性能が得られた.

P2-11 PSA density 算出のためのセグメンテーションによる前立腺体積計測法の検討

○安倍 和弥¹, 武尾 英哉¹, 永井 優一², 黒木 嘉典³, 縄野 繁⁴
¹神奈川工科大学, ²国立がん研究センター東病院, ³にいむら病院, ⁴新松戸中央総合病院

現在, 医療の様々な分野において画像支援診断(CAD: Computer Aided Diagnosis)の研究・開発が行われている. また, AI 技術を用いた臓器・病変検出においても積極的な検討が行われている.

前立腺がんは欧米人に羅漢率の高い病変であり、日本においても 2025 年には男性の羅漢率1位の病変となると予想されている。その発見には PSA 測定を用いた検査の精度が高く用いられているが、PSA 値は前立腺肥大症などでも高値が出ることがある。そこで、感度の向上を目的として PSA 値を前立腺の体積で除する PSA density(PSAD) を用いている。しかし、現在の PSAD 算出に用いる体積は楕円体法により楕円に見立てて縦径、横径、上下径から算出するやや正確性に欠けるものである。

本研究では、U-Net を用いた領域抽出手法を用いて前立腺の体積を計測、PSAD 算出に用いる手法について研究を行う.これにより従来法よりも正確な体積算出と PSAD 算出が行え、前立腺がんのスクリーニング精度向上が期待できる.

P2-12 Perforation Detection from Endoscopy Videos Using Model Training with Synthesised Images by GAN

> 〇蔣 凱¹, 伊東 隼人¹, 小田 昌宏²¹, 奥村 大志³, 森 悠一³, 三澤 将史³, 林 武雅³, 工藤 進英³, 森 健策¹⁴⁵

¹名古屋大学 大学院情報学研究科,²名古屋大学 情報連携推進本部情報戦略室,³昭和大学 横浜市北部病院消化器セン ター,⁴名古屋大学 情報基盤センター,⁵国立情報学研究所 医療ビッグデータ研究センター

Endoscopic submucosal dissection (ESD) is a minimally invasive treatment for early gastric cancer. In ESD, a physician directly removes the mucosa around the lesion under internal endoscopy by using the flush knife. However, the flush knife may accidentally pierce the colonic wall and generate a perforation on it. If doctors overlook a small perforation, the patient may need emergency open surgery since perforation can easily cause peritonitis. Therefore, there is a potential demand for a computer-aided diagnosis (CAD) system, which prevents the overlooking happened in the ESD. Researchers have used deep learning methods for perforation frame classification in endoscopy videos. However, the number of perforated frames is very rare. Therefore, it is very hard to use a supervised learning method for perforation frame classification in a limited-size dataset. We propose to use the Generative Adversarial Network (GAN) to generate virtual images to augmented the dataset. The generated image can be used in deep learning model training to improve the accuracy of classification. We used ModaNet and ResNet-50 as classification models and trained on the original dataset and augmented dataset, respectively. Experiments show models trained on the dataset augmented by the GAN have better classification accuracy.

P2-13 実体顕微鏡画像における血管構造セグメンテーション

 ○呉 運恒¹,小田 昌宏⁵¹,林 雄一郎¹,武部 貴則³,森 健策^{1,2,4}
 ¹名古屋大学大学院情報学研究科,²名古屋大学情報基盤センター,³東京医科歯科大学統合研究機構,⁴国立情報学研 究所医療ビッグデータ研究センター,⁵名古屋大学情報連携推進本部

In recent years, researchers have cultivated some three-dimensional tissue in vitro which is similar to human organs in biological functions called organoid. Organoid can be used in many fields such as drug development and regenerative medicine. However, it will take researchers a lot of time and energy to observe the transplantation process and transplantation results. With the development of artificial intelligence and image processing technology, it is necessary to use these technologies to automatically analyze stereomicroscope images that are taken during organoid transplantation and then assist researchers to improve their research efficiency. The vessels of the host are an important factor to analyze organoid transplantation results. Therefore, in this study, we preprocess the stereomicroscope image to remove the noise, and then improve the contrast of the image, so that the vascular structure which is difficult to be observed by the naked eye in the original image can also be well observed. Then we use the line detector to extract the vascular structure from stereomicroscope images. The experimental results show that the vascular structure is well segmented, even for low contrast and noisy image. This study enables researchers to better observe the vascular structure in low contrast stereomicroscope images and lays a foundation for the subsequent research for stereomicroscope image in organoid research.

P2-14 深度情報を利用した FCN による腹腔鏡映像からの血管領域自動抽出の検討

〇榎本 圭吾¹, 林 雄一郎¹, 北坂 孝幸², 小田 昌宏^{3,1}, 伊藤 雅昭⁴, 竹下 修由⁴, 三澤 一成⁵, 森 健策^{16,7}

¹名古屋大学大学院情報学研究科,²愛知工業大学情報科学部,³名古屋大学情報連携推進本部情報戦略室,⁴国立がん 研究センター東病院,⁵愛知県がんセンター,⁶名古屋大学情報基盤センター,⁷国立情報学研究所医療ビッグデータ研 究センター

本稿では、深度情報を利用した腹腔鏡映像からの血管領域自動抽出手法について述べる. 腹腔鏡下手術は低侵襲であるため 患者への負担は小さいが、手術難易度が高く術者にとっての負担は大きい. そのため、術者を支援するシステムの開発が求め られている. 特に,血管は手術を進めるうえで重要な体内構造であり,損傷により出血や合併症を引き起こす可能性があるため、 腹腔鏡映像から血管領域を自動抽出する研究が行われている. 従来の血管領域自動抽出手法では、FCN によって RGB 画像か ら色情報に基づき血管領域を抽出していた. 本稿では、腹腔鏡映像から深度情報を推定し、RGB-D 画像を用いた FCN による 血管領域抽出手法を提案する. 実験の結果、従来手法と比べて、鉗子とのオクルージョンや出血がみられるシーンの抽出精度 が向上した.

P2-15 Attention 機構を導入したグラフニューラルネットワークによる腹部動脈血管名自動対応付け

○出口 智也¹,林 雄一郎¹,北坂 孝幸²,小田 昌宏³¹,三澤 一成⁴,森 健策¹⁵⁶

¹名古屋大学大学院情報学研究科,²愛知工業大学情報科学部,³名古屋大学情報連携推進本部情報戦略室,⁴愛知県がんセンター消化器外科,⁵名古屋大学情報基盤センター,⁶国立情報学研究所医療ビッグデータ研究センター

本稿では、3次元腹部 CT 像から抽出した動脈領域に対し、Attention 機構を導入したグラフニューラルネットワークを用い て血管名自動対応付けの検討を行ったので報告する。

腹部動脈は個人差が大きく複雑な構造を持ち,種類も多様であるため,構造の把握は困難である.そこで,外科手術において医師が患者の血管構造を把握する際に医師の負担を軽減するため,血管領域に血管名を対応付けすることが提案されている.

これまで、血管構造をグラフ構造と捉え、血管の位置や太さ、長さなどを特徴量としてグラフ畳み込みネットワークを用い て腹部動脈領域に対して血管名自動対応付けを行う手法が提案されており、その有用性が示されている.

本稿では、Attention 機構を用いるグラフニューラルネットワークを利用した血管名自動対応付けの検討を行った. 100 症例の CT 像に対して血管名自動対応付けを行った結果、平均 F 値は 89.6% であった.

P2-16 Cascaded Convolutional Network を用いた 3D 腎臓腫瘍セマンティックセグメンテーション

〇趙 武楊¹, 申 tt^{1} , 小田 昌宏^{2,1}, 林 雄一郎¹, 東田 直樹³, 橋本 正弘³, 陣崎 雅弘³, 森 健策^{1,4,5}

¹名古屋大学 大学院情報学研究科,²名古屋大学 情報連携推進本部情報戦略室科,³慶應義塾大学 医学部放射線科, ⁴名古屋大学 情報基盤センター,⁵国立情報学研究所 医療ビッグデータ研究センター

Automated segmentation of kidney tumors from CT volumes is an important step for kidney tumor treatment, which can assist doctors to make surgery planning and diagnosis. In recent years, deep learning-based methods have been performed on medical image segmentation. However, for kidney tumor lesions, the irregularity of tumor shape and low contrast of lesion boundaries in CT images make kidney tumor segmentation especially challenging. In this paper, we described a two-stage pipeline for kidney tumor segmentation in 3D CT images using convolutional networks. In the first stage, we get the coarse location of the kidney using the lightweight network and then crop VOI (Volume-of-interest) from 3D CT images. Then the second stage, we precisely segment the kidney tumor from the cropped patch in the first stage. We used our own datasets for the experiment, and the results also show that our method for kidney tumor segmentation is effective.

P2-17 Bayesian U-Net を用いた脊椎セグメンテーションの不確実性予測:大規模 CT データベースに おける年齢・性別統計モデリングに向けて

〇程 卓¹, 大竹 義人¹, スーフィーマーゼン¹, 上村 圭亮¹, 重松 秀樹², 池尻 正樹², 合田 憲人³, 佐藤 真一³, 橋本 正弘⁴, 明石 敏昭⁵, 佐藤 嘉仲¹

1奈良先端科学技術大学院大学,2奈良県立医科大学,3国立情報学研究所,4慶應義塾大学,5順天堂大学

Quantifying segmentation uncertainty has become an important task due to the large diversity in anatomical structures, such as vertebrae. A previously proposed Bayesian U-Net demonstrated a correlation between the Monte Carlo (MC) dropout sampling-based uncertainty prediction and the segmentation accuracy in an application of the muscle segmentation. However, the effectiveness of this approach in vertebrae segmentation has not been validated. In this work, we integrate the MC dropout sampling in a framework that achieved high vertebrae segmentation accuracy and landmark detection rate in the MICCAI 2019 challenge. A 6-fold cross-validation study using the public CT dataset VerSe 2019 is conducted. The proposed method achieves an average Dice coefficient (DC) of 0.922 ± 0.012 , with a correlation coefficient between DC and corresponding uncertainty of -0.533 among the correctly identified vertebrae (landmark detection rate is 92.0%, n=793/862). Furthermore, we validate our approach on 10 CT volumes from a large-scale CT database collected independently from the training dataset, on which an average DC of 0.980 ± 0.007 and a correlation coefficient of -0.608 are achieved among the correctly identified vertebrae segmentation accuracy, which would be helpful for the age- and gender-dependent statistical modeling in a large-scale CT database in our future work.

P2-18 弱教師あり学習を用いた MR 画像における血管奇形腫瘍の領域の自動抽出

○深谷 航生¹, 原 武史², 野崎 太希³, 周 向栄², 藤田 広志²

¹ 岐阜大学大学院自然科学技術研究科知能理工学専攻,² 岐阜大学工学部電気電子·情報工学科,³ 聖路加国際病院放射線科

血管奇形腫瘍に関する診断や治療においては領域の正確な把握が重要である.現場では、目視による診察や、撮像された MR 画像から医師が主観で治療の効果判定を行っている.本研究の目的は、MRI を用いて撮像した画像において、深層学習を 利用した血管奇形腫瘍の領域の自動抽出手法を開発し、定量的評価につなげることである.本研究が対象とするクリッペルト レノネーウェーバー症候群は症例数が非常に少なく、腫瘍がびまん性を有するため、正解画像の作成が困難である.そこで、 Attension Map と前後処理を用いて、画像単位で付与されたラベルのみから血管奇形腫瘍領域の自動抽出を行った.

P2-19 取り下げ

P2-20 頭部 thick-slice CT 画像からの階層的脳内血腫領域抽出法の提案

○岡 和範¹,藤田 大輔¹,野原 康伸²,井上 創造³,有村 公一⁴,飯原 弘二⁵,小橋 昌司¹ ¹兵庫県立大学 大学院 工学研究科,²熊本大学,³九州工業大学,⁴九州大学,⁵国立循環器病研究センター

脳血管疾患は、日本の2019年の死因第4位であり、死亡者数は約10万人に上る.その要因の一つは脳内血腫(ICH; intracerebral hemorrhage)である. ICH は専門医でも読影が難しく,診断に時間を要するため、医師の負担軽減や診断の迅速性・ 正確性向上のために、脳CT 画像からの ICH 領域自動抽出法が必要とされている.従来研究として、CNN によるセマンティッ クセグメンテーションに基づく手法が研究されているが、実用化には不十分で、特に高吸収を示す脳室内やクモ膜下出血領域 の過抽出が多い.本研究では、高吸収領域の過抽出を削減する脳 CT 画像からの ICH 領域自動抽出法を提案する.提案法は、 ICH の解剖構造に注目し、高吸収領域の抽出、深層学習を用いたクラス分類による階層的手法を提案する.このモデルの性能を、 脳室内出血を有する被験者 33 名を用いて評価し、本手法の深層学習モデルが高吸収領域における過抽出領域削減に有効である ことを示した.

P2-21 CycleGAN と Bayesian U-Net による Cross-modality セグメンテーション: CT 画像学習データ を用いた MRI での自動筋体積計測

○李 贛萍¹, 大竹 義人¹, 中西 直樹¹, 谷口 匡史², 八木 優英², 市橋 則明², スーフィーマーゼン¹, 上村 圭亮¹³, 高尾 正樹³, 菅野 伸彦³, 佐藤 嘉伸¹
¹奈良先端科学技術大学院大学,²京都大学,³大阪大学

Generative methods based on neural networks have been proved efficient in cross-modality medical image translation and segmentation tasks. Meanwhile, the scarcity of medical datasets with high-quality annotations has become the bottleneck of the segmentation performance of neural networks. The complex anatomical structures make the annotation task extremely time-consuming. In this work, we investigate the feasibility of the image translation from a specific modality (e.g., MRI) to another (e.g., CT) with CycleGAN for the cross-modality segmentation, where annotations performed on one modality are used for training of the segmentation model on another modality. We first trained a CycleGAN model using training data set of 462 CTs and 136 MRIs, and a Bayesian U-net segmentation model using training data set of 20 manually annotated CTs. The CycleGAN translated the target MRI to CT-like images, which were automatically segmented by the Bayesian U-net. The experiments on quadriceps muscles using fully-manually segmented MRIs of 57 subjects showed a dice coefficient of 0.723 \pm 0.128 (mean \pm std).

特別講演1

10:50~11:40 SL1「量子コンピュータをどう使いこなすか」

座長:工藤 博幸(筑波大学/JAMIT 学会長)

SL1 量子コンピュータをどう使いこなすか

○伊藤 公平

慶應義塾長,慶應義塾大学理工学部 教授

最近,「量子コンピュータ」という言葉が様々なメディアで取り上げられています.その動作原理は単純ではありませんが, それゆえに自由度が高く,特定の課題を解く,計算することに関してはスパコンの性能を凌駕することが期待されています. そこで一つでも多くの「特定の課題」を見出し,自然科学や産業界の発展に役立てようとする研究活動が活性化しています. 本講演では,量子コンピュータ研究開発の最先端を紹介し,そのうえで,どのような課題に対して量子コンピュータが適して いるかを議論します.画像処理,最適化,機械学習といった具体的な問題に対する現状を紹介し,将来の医療応用の可能性を 考えていただきます.

スポンサードセミナー1

11:50~12:20 SS1 「Amazon SageMaker を使って

医用画像の機械学習をはじめる方法」

座長:中田 典生 (東京慈恵会医科大学)

SS1 Amazon SageMaker を使って医用画像の機械学習をはじめる方法

○橋本 正弘¹, 今井 真宏²

¹ 慶應義塾大学医学部 放射線科(診断)特任助教, ²Amazon Web Services Japan PublicSector シニアソリューショ ンアーキテクト

昨今,機械学習を利用した画像解析技術が劇的な進化を遂げており,様々な産業における活用が勢いを増しています. Amazon SageMaker は,機械学習を簡単にはじめたり,そのプロセスを効率化するための Amazon Web Services が提供する機械学習プラットフォームサービスです.

このセッションでは、Amazon SageMaker の概要から DICOM 形式の画像の利用や学習に加え、利用を開始する際のベスト プラクティスまで、既にサービスをご利用いただいている慶應義塾大学医学部 放射線科(診断)橋本先生をお迎えしてご説明 いたします。

スポンサードセミナー2

12:20~12:50 SS2 NVIDIA における医療 AI の取り組み

- 医用画像 AI を加速する Clara Imaging のご紹介 -」

座長:中田 典生 (東京慈恵会医科大学)

SS2 NVIDIA における医療 AI の取り組み - 医用画像 AI を加速する Clara Imaging のご紹介 -

○阮 佩穎

エヌビディア合同会社 シニア ディープラーニング ソリューションアーキテクト

NVIDIA は 10 年以上ヘルスケアや医療画像処理の機能向上に取り組んでおり,現在多くの医療機器ベンダーに採用され NVIDIA の GPU が活用されています,昨今は脳腫瘍や肺がんの検出する AI 機能を提供し,直近では COVID-19 のセグメ ンテーションなど多くの AI 機能を活用したソリューション等も提供しています.本セミナーこれらの機能を実装している「 NVIDIA Clara Imaging」を中心に NVIDIA の医療 AI の取り組みをご紹介いたします.

特別企画1

14:10 ~ 15:40

SE1「重力下の人体を可視化する~立位のイメージング~」

座長: 佐藤 嘉伸 (奈良先端科学技術大学院大学)

陣崎 雅弘 (慶應義塾大学)

【企画の意図】人間は日常の多くを立位・座位で過ごしていますが,立位・座位での人体の可視化は十分できていません.立 位では,器質的疾患のみならず多くの人体機能の動態評価も可能になり,機能的疾患の評価を通して"健康寿命"の延伸に貢 献することが期待できます.また,検査効率や占有スペースの視点からも臥位検査と比べて恩恵をもたらします.この企画を 通して,立位のイメージングという新しい潮流を感じて頂きたいと思います.

SE1-1 X 線動態撮影

金沢大学 医薬保健研究域

近年のフラットパネルX線検出器(FPD)の進化により,従来の単純X線撮影と同程度の被ばく線量での動態検査が実現した. 呼吸過程を撮影した胸部X線動画像には、横隔膜・胸郭・心壁の各種運動と、呼吸や肺血流による肺野内の濃度変化(=X線 透過性変化)などが投影されている.立ったままの一番自然な状態で、簡便かつ迅速に心肺機能を評価できる点に特徴がある. 胸部領域を中心に、X線動態撮影の現状と将来展望を解説する.

SE1-2 立位 CT の開発と応用~健康長寿の時代に向けて~

○山田 祥岳

慶應義塾大学 医学部 放射線科学教室 (診断)

立位や座位で全身を3次元像で撮影できる画像診断機器はこれまでなかった. 慶應義塾では,国内メーカーと共同で320列 全身撮影用立位 CT を開発し,当院に世界第1号機を導入した.器質的疾患は臥位でも評価できるが,身体機能の中には立位 でないと評価困難なものがある.立位における身体機能評価は,超高齢化で健康長寿が求められる時代において重要性が高い. また,立位 CT は完全非接触自動化が可能で,ウイズコロナの時代においては威力を発揮する.重力下での正常解剖や各種病 態の解明を進めていきたい.

SE1-3 Gravity MRI による多姿勢イメージング

○宮地 利明

金沢大学 医薬保健研究域

人間は地球上で生きる限り,必然的に重力の影響を受ける.重力が臓器や組織に及ぼす生理的な作用や形態の変化は、立位 や臥位などの体位によって異なる.そこで我々は、様々な姿勢で磁気共鳴画像(MRI)が得られる装置(Gravity MRI)を開発 した.そして体位による機能・形態の違いを Gravity MRI によって測定し、生体局所の重力代償能や環境恒常性を評価するプ ロジェクトを立ち上げた.ここでは中枢神経系を始め、内臓や骨軟部などにおける本プロジェクトの成果を紹介させて頂く.

SE1-4 挑戦する PET:世界初のヘルメット型 PET は座位ブームを起こすか?

○山谷 泰賀

量子科学技術研究開発機構

ヘルメット型. これが,理想的な頭部専用 PET 装置は何か?という問いへの私の答えである. 定量性を高め,検査時間を短 くし,被ばく量をさらに抑えるためには,放射線検出感度を十分に高めることが不可欠である. 頭部にフィットするような半 球状の検出器配置は,最も少ない検出器数で最も感度を高めることができる形状であろう. 7年間の研究開発を経て,世界初 となるヘルメット型 PET がいよいよ実用化されようとしている. PET としては珍しい座位型である点も特徴である. 患者にとっ ては検査への抵抗感が少なく,施設にとっては省スペースのメリットもあるだろう. 本講演では,座位 PET が秘める可能性に ついても議論したい.

[○]田中 理恵

シンポジウム1

15:50~17:20 SY1 [AI 教育を考える]

座長:橋本 正弘 (慶應義塾大学)森 健策(名古屋大学)

【企画の意図】AI研究は日々進歩しておりますが、AIを支える人材が国内で不足していることが指摘されて久しいです.この状況を踏まえて、昨年文科省が医療データ人材育成拠点形成事業を公募し、東北大学と名古屋大学が採択されています.AI人材育成はどこのアカデミアでも大きな課題であり、このシンポジウムでは4つの大学からAI教育への取り組みの現状をお話頂きます.教育のあり方について、情報共有の場になればと思います.

SY1-1 デザイン思考にもとづく医療 AI 研究のための人材育成

○植田 琢也

東北大学 医学系研究科 (文科省 AI人材育成教育拠点)

医療分野は AI を用いた技術開発が期待できる分野として期待をあつめる一方,データの構築・医療倫理や安全性の担保など,他の産業分野と一線を画した様々なロードブロックがあります.医療的・技術的な視点の双方向から医療課題をとらえ,実践的に AI 技術を応用できる人材の育成が求められています.

東北大学では、これまで培ってきたデザイン思考に基づく医療開発のノウハウを生かし、医療的価値と技術革新が交わる場所を探し、AI技術を医療現場での実課題解決に活用できる人材育成のとり組みを進めています。

SY1-2 医療 AI 人材教育における現状と情報学研究者としての展望

○森 健策

名古屋大学大学院情報学研究科(文科省 AI人材育成教育拠点)

本講演では、医療 AI 人材教育における現状と情報学研究者としての展望について述べる.人工知能技術、とりわけ、機械学 習は世の中の様々なところで利用されており、社会を大きく変革する一つの原動力にもなっている.このような技術の開発に 携わる研究者の育成も重要であるが、技術を正しく理解し、そして使いこなす人材の育成が急務となっている.これは、医療 分野においても例外ではない.本講演では、名古屋大学などにおいて行われている AI 人材育成プログラムなどを踏まえながら、 医用画像工学分野における人材について、情報学研究者の立場から述べたいと思う.

SY1-3 国際競争に打ち勝つ AI 人材を育成するために何が必要か?

○鈴木 賢治

東京工業大学 科学技術創成研究院

最近,深層学習が革新的な技術として注目され,様々な分野で盛んに応用されている.産業界,医学界を問わず AI 人材の需要は極めて高いが,AI 人材は圧倒的に不足しており,大学での AI 人材の育成が急務である.本講演では,東京工業大学での AI 教育と AI 人材の育成を紹介するとともに,米国の大学の AI 教育についても触れる.また,国プロによる AI 基盤技術の開発と医療 AI への応用を紹介し,AI 基礎に裏打ちされた AI 開発,システマティックな実装と解析,的確な評価ができる人材が必要であることを述べる.

SY1-4 CG・画像処理分野での深層学習の実践的教育

○金森 由博

筑波大学 システム情報系

CGや画像処理の研究分野では深層学習が基本技術として普及した.しかし大学学部の教育では座学で基礎知識に触れる程度であり,教育と研究のレベルの乖離が深刻に拡大している.そこで講演者は、学部3年次を対象に、当該分野における深層学習の応用事例を実践的に学ぶ授業を立ち上げた.各学生が個別に課題を設定し、Google Colaboratory上で実験してミーティングでフィードバック、というサイクルを反復することで、研究活動も併せて体験できる.

ー般演題セッション3

17:30 ~ 18:40

| P3「医用画像:X 線写真,核医学,超音波,その他」

座長:小尾 高史(東京工業大学)羽石 秀昭(千葉大学)

P3-01 ポータブル型トモシンセシスにおける X 線源と検出器の位置関係推定法

○堀 拳輔¹,小池 貴久¹,只野 喜一¹,橋本 雄幸¹ ¹杏林大学大学院保健学研究科

ポータブル X 線撮影は病室から撮影室への移動が困難な患者の X 線画像を得るために用いられているが、一般に 2 次元投影 画像が得られるのみである.ポータブル X 線撮影においてトモセンシセスが可能であれば、より多くの診断情報が得られる. しかし、撮影時においてトモセンシセス画像再構成に必要不可欠な検出器と X 線源の位置関係を一意に定めることは容易では ない.我々は専用のマーカーパネルを検出器に装着するだけで検出器と X 線源の位置関係を推定し、ポータブル型トモセンシ セスを可能とする方法を開発した.この方法は、金属マーカーの重心位置と投影画像上の金属マーカーの重心位置を用いるこ とで、検出器と X 線源の位置関係を幾何学的に推定計算するものである.数値シミュレーションにより提案手法の推定精度を 検討した結果、誤差は 1%以内であり、ポータブル X 線撮影でもトモシンセシスを実現できる可能性が示唆された.

P3-02 深層学習とデジタルファントムを用いた骨陰影低減技術の開発

○五島 風汰¹,田中 利恵²,小田 昌宏³,森 健策⁴,高田 宗尚⁵,田村 昌也⁵,松本 勲⁵
¹金沢大学大学院 医薬保健学総合研究科,²金沢大学 医薬保健研究域附属 AI ホスピタル・マクロシグナルダイナミクス研究開発センター,³名古屋大学情報連携推進本部情報戦略室,⁴名古屋大学大学院情報学研究科,⁵金沢大学呼吸器外科

呼吸過程を撮影した胸部X線動画像による肺癌の胸壁浸潤判定の実現可能性が報告されているが、骨陰影が定量的な動態解 析の妨げとなっている.本研究の目的は、様々な投影方向で取得された胸部X線動画像に適用可能な骨陰影低減技術を開発す ることである.実人体の4次元CT画像をモデル化したデジタルファントム20体から疑似胸部X線動画像(骨あり/なし)を 7200枚作成し、敵対的生成ネットワーク(pix2pix)を学習させた.デジタルファントムから作成した推論用データセット(骨 あり/なし)を用いて評価した結果、ピーク対雑音比(PSNR)35.0[dB]以上、構造的類似度指標(SSIM)0.950以上と、いずれ も高値を示した.また、臨床画像に適用したところ、時間軸方向の低減効果にばらつきがみられたものの、すべての投影方向 の画像の骨陰影を選択的に低減できていることを確認した.動画像の処理が可能となるよう、時間軸方向の情報を加味したネッ トワーク構造の改良が今後の課題である.

P3-03 解剖学的情報を併用した教師なし深層学習による PET 画像のノイズ除去

○大西 佑弥¹,橋本二三生¹,大手 希望¹,大庭 弘行¹,大田 良亮¹,吉川 悦次¹,尾内 康臣² ¹浜松ホトニクス株式会社 中央研究所、²浜松医科大学 光尖端医学教育研究センター 生体機能イメージング研究室

PET 画像のノイズ除去に CNN を応用する場合,一般的に高線量または長時間計測および低線量または短時間計測での画像 ペアを多数準備する必要がある。そこで我々は教師なし CNN を利用し,さらに MR 画像の解剖学的情報を併用することでノ イズ除去性能の向上を検討する。ネットワークは MR 画像から解剖学的特徴を抽出する CNN と PET 画像のノイズ除去を行う CNN が2種類のアテンションゲートによって結合されており,アテンションゲートを介した解剖学的情報の入力は MR 画像の 微細構造やセマンティック特徴をより明示的に利用することが可能である。18F-FDG,11C-raclopride,18F-florbetapir による 数値シミュレーションおよび実測データを用いて評価を行った結果,従来手法と比較し定量性を維持しつつノイズが抑制され たことから、PET 計測時間の短縮や RI 投与量の削減による被ばく線量の低減が期待される。

P3-04 ピンホール SPECT における散乱線の定量と除去法の比較

○島田 良¹,村田 一心²,尾川 浩一³ ¹法政大学大学院理工学研究科,²国立天文台,³法政大学大学院理工学研究科

本研究では SPECT でピンホールコリメータを装着して、データ収集を行うピンホール SPECT システムにおける散乱線の混入量を定量的に明らかにした.対象とする臓器は心筋であり、周囲にある体積の大きな臓器(主に肝臓)から受ける散乱線の影響を調べるため、シミュレーションではMCATファントムを用いた.また散乱線除去法として提案されている TEW 法と 散乱カーネルを用いたデコンボリューション法を用いた散乱線除去の精度を比較した.この結果、エネルギーウィンドウ内に 混入した散乱線成分は2割から3割であった.また散乱線の除去精度は両者でほぼ同等であった.

P3-05 デコンボリューション法による静止型マルチピンホール SPECT システムの空間分解能改善

○大越 迪¹,村田 一心²,尾川 浩一³

1法政大学大学院理工学研究科,2国立天文台,3法政大学理工学部

通常の SPECT システムでは、データを取得する際に検出器を回転させるため、データ収集に時間がかかってしまうという 課題がある。そこで近年、検出器を回転させない静止型ピンホール SPECT の研究が進められている。このシステムでは、複 数のピンホールを有するガンマカメラを複数台使用しているが、ピンホール径を大きくすると、画質の劣化が起こってしまう。 これを補正する方法として7 rays 法などが提案されてきたが、ピンホール径が大きい場合には近似精度が悪く、種々の補正と の絡みから十分な方法とは言えなかった。そこで本研究では、点拡がり関数を推定し、マルチピンホールコリメータで取得し た投影データに対して実空間でデコンボリューションを行い空間分解能の改善を試みた。再構成画像の画質を評価したところ、 従来の方法と比べて空間分解能の改善、PSNR 値の向上を確認することができた。

P3-06 Deep Image Prior による SPECT 画像の画質改善の試み

○矢部 暁登¹,對馬 結太¹,福原 健太¹,松原 佳亮²,越野 一博³,渡部 浩司⁴,銭谷 勉¹ ¹弘前大学大学院理工学研究科,²秋田県立循環器・脳脊髄センター研究所脳血管研究センター放射線医学研究部,³北 海道情報大学経営情報学部システム情報学科,⁴東北大学サイクロトロン・ラジオアイソトープセンター

SPECT(Single photon emission computed tomography)装置は目的臓器に集積した放射性薬剤から放出されるガンマ線を検 出し、体内の薬剤の分布を画像化することで、体の機能が可視化できる. 我々のグループでは高解像度 SPECT 装置を開発し ており、通常の SPECT 装置より高解像度の SPECT 画像を得ることができるが、検出効率が低く十分なカウントが得られな いため、画質が劣化してしまう. そこで、事前学習が不要で、ターゲット画像のみで画像を復元できる DIP(Deep Image Prior) の利用を検討した. 本研究では DIP を用いて SPECT 画像の画質改善を目的とする. 評価実験では数値ファントムを用いて統 計ノイズを付加した投影データを生成し、DIP 処理をした投影データと DIP 処理をしていない投影データを OS-EM 法で画像 再構成した. 作成された再構成画像を数値ファントムと比較評価した結果、DIP 処理によって PSNR は 32.26dB から 39.16dB, SSIM は 0.860 から 0.928 に向上した. また、実測データに DIP 処理した結果、再構成画像全体のノイズが低減され、ファントムの構造が鮮明になった.

P3-07 近接マルチスプリッティングを用いた Row-Action 型 PET 画像再構成法の体系的な構成

○定方 数弥¹, 金 喜正¹, 工藤 博幸¹

1筑波大学大学院 理工情報生命学術院 システム情報工学研究群 情報理工学位プログラム

PET 装置の画像再構成には、最尤推定(ML)に基づき構成された評価関数を反復法で最小化する方法が用いられている. 反復法としては、高速に収束するため投影データをサブセットに分割してサブセット毎に画像更新を行うOSEM 法が用いられ るが雑音特性が悪いリミットサイクルに収束する問題点があり、それを解決するためにステップサイズをゼロに近づけるゼロ ステップサイズ制御を導入した改良版の RAMLA 法や DRAMA 法も用いられる.本研究では、評価関数を部分評価関数の和 の形に分解して部分評価関数毎に逐次的に近接写像を計算する近接マルチスプリッティングと呼ばれる数学的枠組みに注目し て、高速に評価関数の最小値に収束する Row-Action 型の反復法を構成する体系的な理論を構築した.近接マルチスプリッティ ングとしては、Passty のスプリッティング、Dykstra 型スプリッティング、近接最小化トリック付き Dykstra 型スプリッティ ングの3種類を用いたが、いずれも PET 画像再構成に使用した先行研究がなく十分な新規性を有する.また、3つの枠組みと も原理は凸評価関数の一項一項をばらして部分評価関数の和の形に分解して各々の近接写像を逐次的に計算する点は共通であ るが、3手法でステップサイズ制御規則、評価関数に摂動項の有無、数学的な収束性に違いがある.3手法を整理して単一の 体系的な理論にまとめた結果、工学の視点から非常に分かりやすく見通しが良くなり画像再構成に使う道筋を拓いたといえる. 更に、PET 画像再構成においてリミットサイクルの回避はサブセット数の削減やゼロステップサイズ制御を用いるのが主流で あるが、2番目と3番目の手法は最大サブセット数の Row-Action 型かつステップサイズ制御を全く用いずにリミットサイク ルを回避できる理想的な手法になっている.シミュレーション実験と全身腫瘍検査の PET 実データを用いた画像再構成実験を 行った結果、提案手法が OSEM 法や DRAMA 法と比較して優位性があることを確認した.

P3-08 Higher SNR PET image prediction using a deep learning model with efficient channel attention

○ YinTuo¹,小尾 高史²
 ¹東京工業大学 情報通信系,²東京工業大学 科学技術創成研究院

In this study, a deep residual efficient channel attention (RCA) network is proposed for positron emission tomography (PET) image denoising. Since adjoining slices share similar brain structure while having different random noise, we take the concatenation of adjoining three slices as an input volume to learn distinguishing noise from consistent structure. Deep RCA network consists of a cascade of five residual groups. Each residual group contains three repetitions of RCA block. RCA module adaptively rescales channel-wise features by learning channel statistics. Short skip connection and long skip connection are employed to bypass abundant low-level features and reach a very deep network. 20 brain models are utilized to simulate PET emission data in the real case. The noisy activity maps reconstructed using maximum likelihood attenuation correction factors (MLACF) algorithm with 50 iterations are used as training data. Gaussian filter, nonlocal mean (NLM) algorithm, and deep RCA network with a single slice as input are selected as reference methods. Qualitative results shows that proposed method with adjoining three slices as input is capable of preserving fine details of tissues and restoring high-quality PET images. Quantitative analysis further confirms the superior performance of this approach over the reference methods with structural similarity index measure (SSIM) of 0.89, peak signal-to-noise ratio (PSNR) of 21.64 dB, and normalized root mean square error (NRMSE) of 3.32%.

P3-09 深層学習により PET/CT 早期ダイナミック画像から遅延画像を合成する手法の初期検討

○ブーサルチャタクリリトゥ¹, 立花 泰彦¹, 小畠 隆行¹, 西井 龍一¹, 東 達也¹
 ¹量子医学医療部門放射線医学総合研究所 分子イメージング診断治療研究部

Initial and delayed scans (dual-time-point scan: DualS) are widely used to improve the diagnostic efficacy of PET/CT for accurate delineation and staging. The acquisition process however is time consuming. To eliminate the need of these delayed scans, we propose a CNN-LSTM based approach for generating delayed phase image series images from the 5 min dynamic scan (DynS) images acquired right after the early phase scan at the rate of 10secs per frame.

DynS images were acquired in addition to DualS images. The CNN-LSTM network was trained to generate delay phase images (included in DualS) from the DynS images to predict the delayed phase image. Qualitative and quantitative analysis of the results obtained in this primary study shows significant accuracy in comparing the actual delayed scan and image generated by our model. Further validation for clinical use will be performed.

P3-10 腹部超音波スクリーニング支援のための深層学習による撮影断面推定に関する初期検討

〇目加田慶人¹, 道満 恵介¹, 小川 眞広², 西田直生志³, 工藤 正俊³

1中京大学工学部,2日本大学病院消化器内科,3近畿大学医学部

本稿では、腹部超音波スクリーニングにおいて撮影された画像に対して、それが日大方式の25 断面撮影法のどの断面に対応 しているかを推定する深層学習手法の初期的検討について述べる。25 断面撮影法は腹部スクリーニングにおける診断の網羅性 を保証するものであり、画像から撮影された断面が推定できることで腫瘍等が腹部のどの位置に存在しているのかを知ること ができる。25 の断面画像には比較的類似した見えの画像も存在しているため、画像特徴の類似した断面をまとめたクラスとし て扱う分類器に引き続き、まとめられたクラスと判定された画像をそれぞれの断面に分類する分類器を利用する2 段階の分類 アルゴリズムを開発した。25 断面を記録した 267 例の検査データセットを対象とした。25 クラス分類を単純に適用した場合の F 値 0.790 に対して、2 段階にすることでその F 値は 0.836 に向上した。

P3-11 深層学習を用いた超音波画像における腎嚢胞の計測補助機能の開発

○金内友里恵¹,橋本 正弘²,東田 直樹², HAQUEHasnine^{2,3},陣崎 雅弘²,榊原 康文¹
 ¹應義塾大学理工学部,²慶應義塾大学医学部,³GE ヘルスケア・ジャパン

超音波検査において, 臓器や腫瘍の計測のために2点1組の計測マーカーが付与される. 計測マーカーに基づいて対象所 見の位置やサイズが計測される. 本研究は腹部超音波検査における計測マーカーの付与をサポートすることを目的として, 深 層学習を用いた超音波画像における計測補助機能の開発を行なった. 対象は, 超音波検査の中でも計測頻度の高い腎嚢胞とし た. 慶應義塾大学病院で 2019 年 1 月から 2020 年 5 月に撮影された腹部超音波画像 170,538 枚のうち, 腎嚢胞が計測されている 2,664 枚を使用して, 画像を入力すると計測マーカーの位置を表す distance map を予測するようにモデルを訓練した. モデル は UNet++ を用いた. 出力された distance map から決定した計測マーカーの位置について, 座標誤差を用いて予測精度を評価 したところ, 平均 39pixel, 中央値 20pixel という結果になった.

P3-12 ピットパターン特徴量の解析に向けた超拡大内視鏡画像の再構成法に関する初期的検討

○伊東 隼人¹,小田 昌宏¹²,森 悠一³⁴,三澤 将史³,工藤 進英³,森 健策¹⁵⁶

¹名古屋大学 大学院情報学研究科,²名古屋大学 情報連携推進本部情報戦略室,³昭和大学横浜市北部病院 消化器セン ター,⁴オスロー大学 臨床効果研究グループ,⁵名古屋大学 情報基盤センター,⁶国立情報学研究所 医療ビッグデータ 研究センター

超拡大内視鏡画像から抽出したピットパターンを表す特徴量,これを入力として元の画像を再構成する深層学習手法を提案 する.拡大内視鏡画像の病理類型分類においてはピットパターンを表すハンドクラフト特徴量の有効性が示されており,さら には特徴量選択による分類の高精度化も報告されている.しかし,分類において画像中のどのようなパターンが特に重要であ るか,といった画像パターンと抽出・選択した特徴量の詳細な対応関係は未だ不明瞭である.特徴量を元の画像として再構成 する技術が求められているものの,ピットパターン特徴量の抽出は画像全体の統計情報であるために元画像と特徴量の間には 恒等写像が存在せず,特徴量からの画像再構成は不良設定問題である.本稿では画像集合と抽出したピットパターン特徴量の ペアの間でのみ成り立つ画像再構成法を検討し,特徴量選択の可視化に向けた検討結果を報告する.

P3-13 X線暗視野法を用いた乳腺における乳管内腺腔の自動抽出および3次元解析手法の開発

○谷口 華奈¹, 砂口 尚輝¹, 市原 周², 西村理恵子², 岩越 朱里², 村上 善子², 島雄 大介³, 湯浅 哲也⁴, 安藤 正海⁵

¹名古屋大学大学院 医学系研究科,²名古屋医療センター 病理診断科,³北海道科学大学 保健医療学部,⁴山形大学大 学院 理工学研究科,⁵高エネルギー加速器研究機構

乳腺疾患の病理診断において、2次元で観察する組織標本上は、非浸潤性乳管癌 (DCIS)の一部は特徴的な篩状構造を示すが、 良性の通常型乳管過形成 (UDH)も篩状様の構造を示すことがある。そのため、2次元的観察では DCIS との鑑別が困難なこと がある。しかし、これらの3次元構築は異なるため、腺腔の3次元構築の可視化は貴重な情報をもたらす。我々はこれまで X 線暗視野法による屈折コントラスト CT 像から腺腔を手動抽出し、3次元可視化・解析を行い、DCIS と UDH 各1 症例におい て3次元腺腔構造に差異があることを示した。しかし、乳管内には膨大な数の腺腔があるため、手動抽出では労力が大きく狭 い領域の観察に限られる。本研究では、短時間での腺腔抽出を実現するために機械学習を用いた自動抽出および解析手法を提 案する。約 10 症例の CT 像から自動腺腔抽出を試み、手動による抽出結果を正解として比較し、高い類似度が得られた。

P3-14 取り下げ

メイン会場

ー般演題セッション4

9:00 ~ 10:50 **P4「医用画像:CT, MRI」**

座長: 伊藤 聡志 (宇都宮大学)

滝沢 穂高(筑波大学)

P4-01 圧縮センシング CT 画像再構成のための FBP 法組み込み型主双対反復法とコーンビーム CT への拡張

○工藤 博幸¹, 王 婷¹

1筑波大学システム情報系情報工学域

スパースビュー CT や低線量 CT の画像再構成はトータルバリエーション (TV) と最小2乗誤差データ項の和を最小化する 圧縮センシングにより行われ,評価関数最小化には Chambolle-Pock アルゴリズムや FISTA などの反復法が用いられる.しか し,これらの反復法は収束が非常に遅く計算量が膨大であり,オーダードサブセット (OS) 法などの経験的な手法を組み合わ せ高速化が行われているのが現状である.しかし,OS 法は評価関数の厳密な最小値に収束せず並列化が困難である欠点がある. 筆者らは,既に新しい別の高速化の数学的枠組みとして,FBP 法の構造を主双対反復法に埋め込んで高速化を実現する FBP 法組み込み型主双対反復法を平行ビーム CT とファンビーム CT に対して提案している (Kudo et al., Proc.SPIE, 2016, Wang et al., Phys.Med.Biol., 2019).本手法は,厳密に評価関数の最小値に収束し FBP 法の構造が基本であるため並列化が容易で利便性 が高い.本研究では,FBP 法組み込み型主双対反復法の凸最適化に基づく新しい数学的導出を試み,鞍点問題を解く前処理付 き近接点法と見ることができることを示し,収束性を満たす FBP 法前処理フィルタの条件を導出し前論文では経験的で曖昧で あった収束性の問題をすっきり解明した.また,FBP 法組み込み型主双対反復法を 360 度スキャンとショートスキャンの円軌 道コーンビーム CT 画像再構成に拡張した.実験では、コーンビームのスパースビュー CT 及び低線量 CT の画像再構成のシミュ レーション実験を行い,本手法が有効に動作することを示した.更に,256 列検出器のキャノンコーンビーム CT 装置実機投影 データの画像再構成を行い,良好な結果を得た.

P4-02 生体軟組織の屈折コントラスト CT に適用するための機械学習によるリングアーチファクト除去 法の開発

○黄 卓然¹,砂口 尚輝¹,島雄 大介²,市原 周³,湯浅 哲也⁴,安藤 正海⁵

¹名古屋大学大学院医学系研究科,²北海道科学大学診療放射線学科,³名古屋医療センター病理診断科,⁴山形大学大 学院理工学研究科,⁵高エネルギー加速器研究機構

我々は屈折コントラスト CT (RCT) のリングアーチファクト (RA) を効果的に除去するために機械学習を用いた手法の開発を進めている.この手法は、RCT の投影像が微分像であることに着目し、機械学習によるネットワークの最適化の際に、その疎の性質を取り入れているところに特徴がある.以前、我々は512×512 画素のラット胎児の CT にこの手法を適用し優れた RA 除去効果を得た.一方で、本手法を実用するには、少なくとも 2000×2000 以上のサイズの画像に適用できる必要があるが、以前は開発環境の制約があり試すことができなかった.

本研究では,独自に開発した RCT の撮像を模擬したシミュレーターを用いて 2000 × 2000 のシミュレーションデータを生成し,本手法内の条件付き敵対的生成ネットワークで学習させる. そのネットワークを用いて,様々な生体軟組織を撮像した RCT に適用した結果,高い RA 除去効果が得られた.

P4-03 サイクル敵対的生成ネットワークによる胸部 4 次元コーンビーム CT の画質改善

○臼井 桂介¹², 尾川 浩一³, 後藤 政実¹, 坂野 康昌¹, 京極 伸介¹, 代田 浩之¹ ¹順天堂大学保健医療学部診療放射線学科,²順天堂大学医学部放射線治療学講座,³法政大学理工学部

放射線治療装置に搭載されたコーンビーム CT は、投影データを呼吸位相毎に分割し再構成することで4次元画像が取得で きる.一方で、この投影データの疎性のため深刻なアーチファクトが生じており、標的位置の正確な同定が困難である。そこで、 サイクル敵対的生成ネットワークによる胸部4次元コーンビーム CT の画質改善を行った。本研究では、胸部20例の4次元コー ンビーム CT およびヘリカル CT (各 1000枚)を学習データとしてサイクル敵対的生成ネットワークによる画質改善モデルを 作成した。次に学習データとは異なる4次元コーンビーム CT に対して学習済モデルによる画質改善を実装し、肺野、軟部組 織および骨部に関心領域を設定して画素値の平均誤差と標準偏差をヘリカル CT と比較した。その結果、平均画素値誤差およ び標準偏差はそれぞれ約80%および約50% 改善したことから、本学習モデルによる高い画質改善効果を実証することができた。

P4-04 マルチ近接スプリッティングを用いた Row-Action 型 TV 正則化 CT 画像再構成法の体系的な構成

○金 喜正¹, 定方 数弥¹, 工藤 博幸²

¹筑波大学大学院システム情報工学研究群情報理工学位プログラム,²筑波大学システム情報系情報工学専攻教授

圧縮センシングによる低線量 CT やスパースビュー CT の画像再構成はトータルバリエーション (TV) 正則化項と最小2乗 誤差のデータ項の和の形をした評価関数を反復法で最小化することにより行われる.反復法としては Chambolle-Pock アルゴリ ズムや FISTA 法がよく用いられるが同時反復型の構造であるため収束が非常に遅い.本研究では、高速に評価関数の最小値に 収束する Row-Action 型の構造を持つ反復法を、近接マルチスプリッティングと呼ばれる数学的枠組みを用いて構成する体系的 な方法論を構築した.アルゴリズム構築に利用する枠組みは、Passty のスプリッティング、Dykstra 型スプリッティング、近 接最小化トリック付き Dykstra 型スプリッティングの3つであり、いずれも CT 画像再構成に使われた先行研究がなく十分な 新規性がある.また、3つの枠組みとも原理は凸評価関数の一項一項をバラして部分評価関数の和の形に分解して各々の近接 写像を逐次近似的に計算する点は共通であるが、3手法でステップサイズの制御規制、評価関数に摂動項の有無、数学的な収 束性に違いがある.3手法を体系的に整理して、単一の体型的な理論にまとめた結果、工学の視点から非常に見通しが良くなり、 画像再構成に使う道筋を拓いたと言える.腹部 CT 実画像を用いて低線量 CT を想定したシミュレーション実験を行った結果、 3手法とも評価関数の最小値に上手く収束して、多くの研究でよく使われる経験的な高速化法である OS-SART 法や OS-SIRT 法に優れることが確認できた.

P4-05 インテリア CT 厳密解法における安定な画像再構成が可能な先験情報領域配置の解明

○伊藤 優太¹,藤井 克哉²,工藤 博幸²

¹筑波大学大学院システム情報工学研究群情報理工学位プログラム,²筑波大学システム情報系

インテリア CT 画像再構成の厳密解法では、X 線を照射して投影データを測定する関心領域(ROI)内部の任意の小領域 B の画像値が既知の先験情報があれば ROI 全体で解が一意に定まる一方,Bから離れた場所ほど逆変換の安定性が悪く再構成 誤差が大きくなることが知られる。ROi 内のなるべく大きな領域で安定に画像再構成可能な先験情報領域の配置がどのよう なものであるかは、当該テーマの未解決問題である。本研究では、先験情報領域 B が複数個(3 個以上)の小領域から構成 される場合のインテリア CT 画像再構成の安定性を先行研究(Defrise et al: Inverse Problems, 2006, Coudurier et al: Inverse Problems, 2008)の安定性解析手法を拡張して数学的に考察して、複数の先験情報領域で結ばれた多角形領域内では逆変換は安 定で安定に画像再構成できることを示した。

P4-06 CT 画像再構成におけるインテリア問題とエクステリア問題の統一解法と安定性解明

○藤井 克哉¹, 工藤 博幸¹ ¹筑波大学システム情報系

近年,被ばく量を軽減するための新たな CT 画像再構成法が精力的に研究されている.不完全投影データからの再構成でよ く知られた問題は投影データ欠損が互いに外側と内側に起きている場合の再構成問題であり,それぞれインテリア問題,エク ステリア問題と呼ばれる.安定性解析において前者は Defrise や Coudurier らの先行研究によって安定な問題となることが肯定 的に示されているが,後者は Cormack により不安定な問題となることが知られている.両者の相違は投影データの欠損が内側 か外側かのみではあるが用いる数学的枠組みは大きく異なるため,扱いが難しい.

そこで本研究では、エクステリア問題の安定性も含めた両問題の統一的な理論を提案する.そのための鍵は Cormack の第二 再生公式と先験情報領域の幾何的な配置であり、これらによりインテリア問題、混合インテリアエクステリア問題、更には未 解決であったエクステリアの安定性を統合的に保証することが可能になることを示した.またこれらのシミュレーション実験 を行い理論の妥当性を検証した.

P4-07 異常データ検知トモグラフィ ABD-CT の概念提唱と事例紹介

○工藤 博幸

筑波大学システム情報系情報工学域

トモグラフィ(CT)において、測定した投影データの一部に異常データが含まれることがよく発生する。例えば、医療用 X 線 CT において被写体内部に金属が含まれる場合に発生するビームハードニング、検出感度特性のむらで発生するリングアー ティファクト、暗視野 CT で固い結晶構造の被写体で発生する回折コントラストアーティファクト、などがこれに相当する. このような状況において、フィルタ補正逆投影(FBP)法で画像再構成を行うと異常データの影響が画像全体に広がり強いアー ティファクトが発生することがよく知られる。本発表では、異常データを画像再構成の過程で自動的に検知して除外して画像 再構成を行う、新しいトモグラフィの枠組み ABD-CT (Abnormal Data Detected CT)を提案する。ABD-CT では、画像再構 成の評価関数として異常データを除外する効果がある L0 ノルムをデータ項に用い、トータルバリエーション(TV)と L0 ノル ムの和 f(x)=beta*(x_TV+'Ax-b'_0^0 を最小化することにより画像再構成を行う、異常データが発生しやすい場所が事前に既知 であればその情報も加味して評価関数 f(x)は異常データになりやすい場所をまとめてグループ L0-LASSO により設計する。本 手法の有効性を示す実データの再構成例を紹介する。投影角度毎に X 線のエネルギーが異なる超極端なビームハードニングを 発生するミリ秒トモグラフィ、医療用コーンビーム CT と放射光 CT におけるリングアーティファクト、暗視野 CT で問題と なる回折コントラストアーティファクト、医療用 CT のメタルアーティファクトである。全てシミュレーションではなく実デー タであるが、ABD-CT の枠組みは強力でいずれも高精度の画像再構成に成功している。驚くべき画質改善を実現している。同 じ問題やデータに深層学習の適用も試みているが、正直に述べると現状では ABD-CT と同等の高画質の画像を再構成すること には成功しておらず深層学習の方が学習データ収集などはるかに手間がかかる。

P4-08 距離マップを利用した肺マイクロ CT 像からの肺胞抽出

 ○椎名 健¹,小田 紘久²,鄭 通¹,中村 彰太²,林 雄一郎¹,小田 昌宏³¹,森 健策^{1,4,5}
 ¹名古屋大学 大学院情報学研究科,²名古屋大学 大学院医学系研究科,³名古屋大学 情報連携推進本部,⁴名古屋大学 情報基盤センター,⁵国立情報学研究所 医療ビッグデータ研究センター

本稿では、肺のマイクロ CT 像における肺胞抽出手法を提案する.末梢肺には多様な微細構造が存在するが、肺胞と呼ばれ る袋状の構造が最も多く見られる.肺胞は、切除肺標本からマイクロ CT 像として画像化できるが、肺胞は複雑かつ三次元的 に分布しており、直感的な理解には自動抽出による可視化が重要である.しかしながら、肺マイクロ CT 像はノイズが多いほか、 隣接する肺胞同士で壁の存在しない箇所があることなどが原因で、単純なアルゴリズムでの抽出は困難である.提案手法では、 具体的には、1)二値化や小成分除去による肺胞壁の抽出、2)肺胞壁からの距離画像における極大点を肺胞の中心と仮定、3)過 検出された肺胞中心の除外、4)Watershedによる肺胞領域の抽出を行う.肺がん患者の切除肺標本のマイクロ CT 像を用いて 評価を行った.定量・定性評価により、良好な結果を得た.

P4-09 非ラグランジアン解法を用いた少数方向投影からの Micro-CT 画像再構成

○熊切 俊夫¹, 岡本 尚之², 工藤 博幸³, 羽石 秀昭⁴

¹千葉大学大学院融合理工学府数学情報科学専攻,²千葉大学大学院融合理工学府基幹工学専攻,³筑波大学システム情報系,⁴千葉大学フロンティア医工学センター

病理診断は微細構造の観察に有用であるが、2次元情報しか得られないといった問題点が存在する.そこで病理診断と同程 度の分解能で3次元構造が観察可能な Micro-CT を導入することで、網羅的な検査が可能になる.一方で、高画質な Micro-CT 像の取得には長時間撮影が必要となる. CT 撮影の高速化には少数方向からの撮影が有効であるが取得データ削減により、解析 的画像再構成手法(FBP 法)ではスポーク状のアーチファクトが発生してしまう.本研究では少数方向投影条件下における画 像再構成手法を Micro-CT に適用することで画質向上を目指す.本稿ではラグランジアンや双対変数を用いない非ラグランジ アン解法に基づいた手法を適用した.この手法は新型近接勾配法を Ordered-Subsets 反復法に拡張することで厳密解に高速で 収束することが可能であり、尚且つ実装も簡便であるといった特徴がある.実験の結果,FBP 法と比較しスポーク状のアーチファ クトを低減し,鮮明な物体構造が得られた.

P4-10 圧縮センシングを深層学習を組み合わせた低線量 CT 画像再構成法 CSDL-Net のヘリカルコーン ビーム CT への拡張

○川又健太郎¹,数尾 昴¹,工藤 博幸¹ ¹筑波大学システム情報系

低線量 CT の画像再構成に有効な手法として、圧縮センシング(CS)とディープラーニング(DL)が精力的に研究され実用 化が進んでいる.測定条件が悪い場合、CS の欠点は画像の滑らかな濃度変化やテクスチャーが消失して階段状アーティファク トが発生しやすいこと、DL の欠点は個人差が大きい学習データから学びにくい構造物(特に血管などの低コントラスト構造物) が消失しやすいこと、であることが知られている。各々の欠点を相補に補い合い単独で用いる場合より高画質の画像を得るた め、森ら(工藤,森,JAMIT2020)は、CS と DL を組み合わせて高画質の画像を生成する CSDL-Net を提案した。本研究では、 CSDL-Net を普及している CT 装置の構成方式であるヘリカルコーンビーム CT の 3 次元画像再構成に拡張する検討を行った. 拡張した 3D-CSDL-Net では、まず 3 次元画像を FBP 法で再構成した後に TV でデノイジングを行い、デノイジングなし画像 とデノイジングあり画像のペアをスライスごとに出力が 1 チャンネルの 2D 畳み込みニューラルネットワーク(CNN)に入力し、 最終的な再構成画像を求める。また、デノイジングしたい単一のスライスを CNN に入力する場合と、デノイジングしたいスラ イスとその前後のスライスを CNN に入力する場合を検討した。腹部 CT 撮影の実画像を用いたシミュレーション実験を行った 結果、3D の場合においても CSDL-Net が有効で各々を単独で用いる場合と比較してより高画質の画像を生成できることが明ら かになった.

P4-11 Pix2pix を用いた任意形状の肺がん CT 画像生成及び CT 画像様ノイズ再現の試み

○遠田 $凉^1$, 寺本 篤司¹, 辻本 正和², 外山 $宏^3$, 近藤 征史³, 今泉 和良³, 齋藤 邦明¹, 藤田 広志⁴

1藤田医科大学大学院 保健学研究科, 2藤田医科大学病院, 3藤田医科大学 医学部, 4岐阜大学 工学部

現在,人工知能による画像生成技術である敵対的生成ネットワーク (GAN)の応用が,医用画像に対しても進められている. 現状は,生成画像の多様性の問題から,応用は症例の豊富な疾患に留まっている.任意の病変形状の画像生成が可能であれば, 希少疾患に対する応用も実現し得ると考えられる.そこで本研究では,GANの派生型であるPix2pixを利用し,スケッチによ り病変形状を制御した画像生成を行うことを目的とした.また,GANの生成画像においては,本来あるべきノイズが消失する 課題もあるため,ノイズの再現も検討事項とした.自然画像の場合,画像を複数のパッチに分割して識別器に入力することで, 生成画像の品質向上が確認されているため,この手法の有無で比較を行った.初期的な検討として,症例が豊富かつ多様な病 変形状を呈す肺がんのCT 画像を対象とし,画像生成が可能であることを確認した.また,パッチ非分割時にCT 画像様ノイ ズが良好に再現された.

P4-12 コーンビーム CT の散乱線除去のための機械学習ネットワーク

○村田 一心¹, 尾川 浩一², 臼井 桂介³ ¹国立天文台,²法政大学理工学部応用情報工学科,³順天堂大学保健医療学部診療放射線学科

本研究では、コーンビーム CT 撮影で発生する散乱線を除去するために適した機械学習ネットワークを提案する.コーンビーム CT は散乱線の影響を強く受けるため、再構成画像の画質が劣化しやすい.そのため、散乱線を精度良く除去できるための ソフトウェアが必要とされている.先行研究で我々は、機械学習が散乱線除去に有効であることは示したものの、ネットワーク構成の最適化は考慮できていなかった.そこで本研究では、散乱線の特性とネットワークの性質との関係を考察した.散乱線の特性を考慮するため、通常の畳み込み層に加え、全結合層を導入した.さらに、通常の投影画像と垂直方向の画像を同時に入力することで、擬似的に3次元空間を考慮できるようにした.ネットワークの学習および評価には、モンテカルロ法で作成 した画像を用いた.その結果、提案ネットワークは従来より高精度で散乱線を除去できることがわかった.

P4-13 臨床 – マイクロ CT 画像間位置合わせ済みデータセットを用いた臨床 CT 像の超解像

○鄭 通¹,小田 紘久¹,林 雄一郎¹,中村 彰太²,森 雅樹³,高畠 博嗣⁴,名取 博⁵,

小田 昌宏 6.1, 森 健策 1.7.8

線診断学講座

¹名古屋大学大学院情報学研究科,²名古屋大学大学院医学系研究科,³札幌厚生病院,⁴札幌南三条病院,⁵恵和会西岡 病院,⁶名古屋大学情報連携推進本部情報戦略室,⁷名古屋大学情報基盤センター,⁸国立情報学研究所医療ビッグデー タ研究センター

This paper proposes a super-resolution (SR) method, for performing SR of clinical CT images utilizing newly-built micro CT – clinical CT images database. Precise non-invasive diagnosis of lung cancer typically utilizes clinical CT images for making surgical plans. Due to the resolution limitations of clinical CT (about 5003 μ m3 / voxel), it is difficult to obtain enough pathological information, such as the invasion condition at alveoli level. On the other hand, micro CT scanning allows the acquisition of volumes of lung specimens with much higher resolution (503 μ m3 / voxel or higher). Thus, if micro CT could be applied to guide SR of clinical CT images, it would be possible to acquire precise anatomical information from clinical CT images. Conventionally, paired low resolution (LR) and high-resolution (HR) images are required for training SR models. Thus, we build the paired clinical CT (LR) and micro CT (HR) dataset by conducting feature-point based registration of clinical CT images. The proposed SR method was trained on the paired clinical CT – micro CT dataset. Experimental results showed that SR results trained on paired clinical CT – micro CT dataset are qualitatively and quantitatively and quantitatively.

P4-14 敵対生成ネットワークによる COVID-19 の CT 画像の生成モデル

○川崎 良¹, 斉藤 篤¹, 橋本 正弘², 大竹 義人³, 明石 敏昭⁴, 清水 昭伸¹ ¹東京農工大学大学院工学研究院, ²慶應義塾大学医学部放射線科, ³奈良先端科学技術大学院大学, ⁴順天堂大学放射

本報告では、COVID-19の典型症例の胸部 CT 画像を対象に、敵対的生成ネットワーク(Generative Adversarial Networks; GAN)を用いた生成モデルの構築について述べる。COVID-19の典型症例の CT 画像は、病変が末梢優位性と両側性を持つと 報告されている。本研究では、この特徴を示す胸部 CT 画像の生成が可能なモデルを2種類の GAN を使って開発した。まず、 胸部 CT 画像は Axial 面のスライス画像のサイズが 512 × 512pixel と大きく、そのままでは学習が困難である。そこで高解像 度の画像を安定して正確に生成可能な Progressive growing を取り入れた PGGAN を用いた。また、2つのネットワークで構成 された Generator を組合せることで高品質な画像生成を実現した StyleGAN を用いた。それぞれのモデルによって生成された 画像に対して、Fréchet Inception Distance (FID) による定量評価や、血管、気管支、病変の特徴に関する視覚的な評価指標に よる定性評価を行い、二つのモデルの優劣について議論する。

P4-15 表現に合わせた生成モデルによる非造影から動脈造影 CT への変換

 ○胡 涛¹,小田 昌宏¹,林 雄一郎¹,魯 仲陽¹,明石 敏昭²,青木 茂樹²,森 健策¹³⁴
 ¹名古屋大学大学院情報学研究科,²順天堂大学大学院医学研究科,³名古屋大学情報基盤センター,⁴国立情報学研究 所 医療ビッグデータ研究センター

This paper proposes a deep representation-based network to improve the CT modality translation from non-contrast CT to artery contrast CT (ACT). ACT has been widely applied to the diagnosis of abdominal vascular diseases. For the imaging of ACT, contrast agent is introduced to the patients' body prior to the scan. To avoid the potential risk led by contrast agent, deep learning technique has been utilized to simulate ACT from non-contrast CT. In this paper, we propose to improve the translation performance via a representation-aligned network. Different from conventional generative adversarial networks (GANs), the proposed generator is trained to simultaneously fool the discriminator and approximate the real samples in the embedding space derived from target domain. The results show that the proposed network is able to generate ACT with more sematic information and higher accuracy.

P4-16 多重解像度画像のグループ型学習を利用した新たな MR 深層学習再構成の検討

○大内 翔平¹², 伊藤 聡志¹

¹宇都宮大学大学院工学研究科システム創成工学専攻,²日本学術振興会特別研究員DC

近年, CNN を用いた MR 画像の再構成法が数多く提案されているが,信号の間引きに由来する高周波帯域の復元と折り返し アーチファクトの除去には依然として課題が残っている.また,位相情報を持つ複素 MR 画像に対応した CNN 再構成法の提 案は少ない.本研究では、多重解像度解析手法である eFREBAS 変換と画像空間型の CNN 再構成法を併用した新たな深層学 習再構成法として eFREBAS-CNN を提案する.提案法では、eFREBAS 展開により、異なる周波数成分を持つ複数枚のサブ画 像を生成する.これらのサブ画像を被写体構造の保存度合に応じてグループ分割し、グループ毎に別個の CNN で学習と再構成 を行う.eFREBAS 変換は複素変換であるため、提案法は位相を含む MR 画像に適用可能な特長を有する.再構成シミュレーショ ンの結果、既存の位相考慮型の再構成法と比べて提案法では位相分布を持つ MR 画像の PSNR を改善でき、さらに画像の詳細 構造の復元性能を高めることができた.

P417 信号空間の深層学習を利用した位相拡散フーリエ変換に基づいた MR 超解像

○山登 一輝¹, 若槻 泰迪¹, 伊藤 聡志¹

1 宇都宮大学工学部

位相拡散フーリエ変換法 (PSFT) の信号は, 被写体が実関数であるという拘束条件を利用することで, 撮像後の数値処理によって, 撮像中に収集していない信号の外挿が可能であり, 高解像度化した再構成像を取得できる.

著者らの研究グループでは、これまで反復的手法により MR 信号を外挿し、再構成像の空間分解能改善を行ってきた. しかしながら、反復的方法では分解能の改善効果は画像の中央部分で小さく、画像の再生には 20 回程度反復する必要がある. 本稿では深層学習再構成の一種である Generic-ADMM-Net を利用し、PSFT 信号の信号外挿と空間分解能の改善を図った. 提案手法の有効性を確認する実験を行ったところ、信号収集量を 25% に制限した PSFT 信号を深層学習に入力し、制限なし 信号による再構成像を教師データとして学習した結果、折り返し雑音が低減し、再構成像の高分解能化を確認できた.

さらに、位相成分を含む画像に対して位相補正を導入することで、より位相成分を含む画像に対する分解能向上にも取り組 んだ。

P4-18 位相拡散フーリエ変換法を導入した MR 深層学習再構成における信号間引きパターンの検討

○金澤 崇大¹, 山登 一輝¹, 伊藤 聡志¹

1字都宮大学大学院 地域創生科学研究科 情報電気電子システム工学プログラム

MRI に圧縮センシングを応用する場合に非線形エンコーディングを行う位相拡散フーリエ変換映像法を使用すると再構成像の画質を改善できることが報告されている.昨年の大会では、交互方向乗数法(ADMM)をCNN化した ADMM-Net を使用し、信号収集比が33%の場合に汎用法であるフーリエ変換映像法に比べて高品質な画像を再生できることを示した.

本研究では、信号収集比を25%から40%と変え、また、信号間引きパターンを等間隔、ランダム、ポアソン分布等に変えた 場合の検討を加えた.その結果、信号収集比によって最適な位相拡散フーリエ変換映像法の位相拡散係数が変化すること、フー リエ変換映像法では実現不可能な等間隔間引きによっても画像再構成が可能であることが示された.

P4-19 MR 画像からの 3 次元導電率分布の回帰による TMS のための誘導電場の回帰精度の改善

○牧 豊大¹, 横田 達也¹, 平田 晃正², 本谷 秀堅¹

1名古屋工業大学,2福島県立医科大学

TMSとはうつ病の治療などを目的として導入された治療法であり、頭部にコイルを当て磁場を発生させることで、脳内の神 経細胞を非侵襲的に電気刺激する手法である.TMSを行う上で最も重要なのは最適なコイル位置の同定であり、コイルによる 脳への電気刺激は重要な指標である.脳内に誘発される電場の分布は、3次元導電率モデルから物理シミュレーションされる. MRI 画像を脳の部位ごとに領域分割し、その部位に対応する既知の導電率を割り当てることで導電率モデルは構築される.し かし、これらの推定は実時間で行うことができない、そこで本研究では、ニューラルネットワークを用いて誘導電場を実時間 推定する機構を提案する.MRI 画像から脳の部位ごとに領域分割する過程及び、導電率モデルから電場を推定する過程の両方 をニューラルネットワークで推定する.また領域分割せず電場を直接推定するよりも推定精度が高くなることを示すことがで きた.

P4-20 PROPELLER MRI における CS と CNN を利用した再構成法の検討

○松元 祐太¹, 堀 拳輔¹, 只野 喜一², 久原 重英², 遠藤 祐太², 橋本 雄幸²

1 杏林大学 大学院 保健学研究科保健学専攻 診療放射線学分野, 2 杏林大学 保健学部 診療放射線技術学科

PROPELLER は blade と呼ばれる帯状領域を様々な方向から取得し、k 空間に再配置する方法である. この方法では、blade 間での位相と回転を利用して被写体の動きを補正できる. しかし、多くの blade を収集すると撮像時間が長くなる欠点がある. 本研究では、PROPELLER で取得されたk 空間データを圧縮センシング (CS) で再構成し、畳み込みニューラルネットワーク (CNN) で処理することで画質向上を試みた. blade の幅と数を変化させたサンプリング率が異なるk 空間データにおいて、CS で再構成した場合、CS を利用せずに再構成した画像を CNN で処理した場合、CS と CNN を併用した場合の3パターンを比較 した. サンプリング率が 45%以下のとき、CS と CNN を併用した方法の評価値が最も良くなったことから、サンプリング率が 低い場合において CS と CNN を併用することで画質向上できる可能性が示唆された.

P4-21 AD-MIL を用いた頭部 MR 画像における回復期脳梗塞患者の帰結予測

○東 優大¹, 寺本 篤司¹, 岡崎 英人², 武田湖太郎³, 前田 繁信⁴

¹藤田医科大学大学院保健学研究科,²藤田医科大学医学部連携リハビリテーション医学,³藤田医科大学保健衛生学部 リハビリテーション学科,⁴藤田医科大学七栗記念病院

現在, 脳梗塞は要介護状態に至る疾患のなかで主要な疾患の1つである. 要介護状態の患者が日常動作の改善や在宅復帰を 目指すに際して, リハビリテーションが重要となる. その際に各々の患者に適したリハビリテーション計画を立案するには, 正確な予後予測が必要不可欠である. そのため本研究では, 回復期脳梗塞患者の MR 画像を CNN に入力し, 弱教師あり学習 手法である AD-MIL を用いて帰結良悪の予測を行った. まず, リハビリテーションを実施する前に撮影した MR 画像を入力し, CNN モデルの1つである VGG-16 の全結合層から 1024 個の特徴量を抽出した. そして, アテンション機構に基づいて, その 特徴量より重み付けされた特徴ベクトルを算出し, 症例ごとに帰結良好および不良の分類を行った. 結果は, 帰結良好の分類 精度が 86.7%, 帰結不良では 61.5% となり, 総合正解率が 75.0%であった. 以上の結果より, 本手法が回復期脳梗塞患者の帰結 予測に有効であることが示唆された.

P4-22 腹部 Cine-MRI における小腸蠕動運動のオプティカルフローによる可視化および腹膜炎の鑑別

○川原 稔暉¹, 井上 明星², 岩本祐太郎¹, 古川 顕³, BolorkhandBatsaikhan³, 陳 延偉¹
 ¹立命館大学大学院 情報理工学研究科,²東近江総合医療センター,³東京都立大学

小腸は人間の生命維持活動に必要不可欠な臓器であり、食物の運搬、吸収、消化機能を担う.これらの機能に異常をきたした場合、腸閉塞や腹膜炎など様々な疾患につながるリスクがあることから、こうした異常の早期発見のために Cine-MRI を用いて小腸の蠕動運動を評価する研究が行われている.先行研究では深層学習を用いて局所領域の小腸領域をセグメンテーションし、その面積変動を評価しているが、広範囲の小腸領域の動きを評価できていない.また、小腸の動きを数値的に評価している研究はあるものの、小腸の直接的な運動を可視化する研究は行われていない.本研究では Cine-MRI に対してオプティカルフローを計算することで、広範囲の腸の運動を評価し、これらの課題を解決するとともに腹膜炎の鑑別も行う.

The small bowel is an organ that is essential for human life support and is responsible for the transport, absorption, and digestion of food. Since abnormalities in these functions can lead to various diseases such as intestinal obstruction and peritonitis, research has been conducted to evaluate the peristalsis of the small intestine using Cine-MRI for early detection of such abnormalities. The existing method, which evaluates the intestine by area variation using segmentation with deep learning, crops a local region of the intestine and evaluates the intestine, but does not evaluate the global range of intestinal movements. In addition, although some studies have evaluated bowel movement numerically, no study has been conducted to visualize direct bowel movement. In this study, we evaluate global range of bowel movements by calculating optical flow on Cine-MRI to solve these problems and also to differentiate peritonitis.

特別講演2

| 11:00~11:50 | **SL2「革新的バイオイメージング技術」**

座長: 陣崎 雅弘 (慶應義塾大学 / JAMIT2021 大会長)

SL2 革新的バイオイメージング技術

○宮脇 敦史

理化学研究所 脳神経化学研究センター チームリーダー, 光量子工学研究センター チームリーダー

細胞内の生体分子の挙動を追跡しながら、大洋を泳ぐクジラの群を想起する.クジラの回遊を人工衛星で追うアルゴスシス テムのことだ.背鰭に電波発信器を装着したクジラを海に戻す時,なんとか自分の種の群に戻ってくれることをスタッフは願う. 今でこそ小さい発信器だが昔はこれが大きかった.厄介なものをぶら下げた奴と仲間から警戒される危険があった.クジラの 回遊が潮流や餌となる小魚の群とどう関わるのか,種の異なるクジラの群の間にどのような相互作用があるのか.捕鯨の時代 を超えて人間は海の同胞の真の姿を理解しようと試みてきた.

バイオイメージング技術において発信器の代りに活躍するのが蛍光や発光のプローブである. 我々は細胞の心を掴むための スパイ分子を開発している. 材料となるのは光を吸収・放出するタンパク質である. そうしたスパイ分子を活用して様々な生 命現象を可視化する技術の実際を紹介する. また, 光と相互作用するタンパク質が「光と生命との相互作用」を巡る研究から 生まれ, それらの生物学的存在意義に関する我々の理解を超えてますます有用になっていく過程を考察したい. 日本に棲息の クラゲ, サンゴ, ウナギ, ホタルに由来するタンパク質が近未来の医学に貢献する可能性を論じてみたい.

超ミクロ決死隊を結成し、微小管をジェットコースターのように滑走し、核移行 signal の旗を掲げて chromatin の jungle に潜り込む,,, そんな遊び心をもちたいと思う. 大切なのは科学の力を総動員することと想像力を逞しくすること. そして whale watching を楽しむような心のゆとりが serendipitous な発見を引き寄せるのだと信じている.

スポンサードセミナー3

12:00~12:50 SS3「薬事承認された画像用 AI ソフト」

座長:木戸 尚治(大阪大学)

SS3-1 富士フイルムの AI 技術「REiLI」の現状 – 最新の取り組みのご紹介 –

○久永 隆治

富士フイルム株式会社 メディカルシステム事業部・マネージャー

当社はこれまで様々な画像処理技術および画像認識技術を開発して、医用画像装置・情報システム搭載してきました。最近 の深層学習などの技術の発展に伴い、当社が培ってきた画像技術の上にこれらの新技術を組み合わせることによって、従来の 画像認識よりも高いレベルで診断を支援する AI 技術の開発を進めています。今回のセミナーでは当社の AI 技術「REiLI」の 取り組みに関して、これまで開発された様々な機能と最新の取り組みを中心に紹介致します。

SS3-2 Siemens Healthineers の AI 技術を応用した画像診断支援サービスの特長と活用事例

○岩田 和浩

シーメンスヘルスケア株式会社 デジタルヘルス& SYNGO 事業部 AI プロダクトマネージャー

Siemens Healthineers では、現在「医療のデジタル化の推進」を一つの大きな柱として掲げ、医療にかかわる人々のバリュー向上を実現するというミッションを理念に技術・製品・サービス開発を進めている. その中心的な役割として、現在 AI (人工 知能) 技術の応用を急速に進めている.

その中の一つの製品として、AI技術を応用した画像診断支援および放射線治療計画支援のプラットフォーム「AI-Rad Companion」の提供を行っており、2021年8月において、4つのシリーズを展開し、既に10の機能が使用可能である.

本講演では、それぞれのシリーズの特長および活用事例についてご紹介させていただくと同時に、開発背景や今後の展望 についても触れさせていただく.

SS3-3 画像診断における AI の実装へ向けた課題と取り組み

○大越 厚

GE ヘルスケア・ジャパン株式会社 エジソン・ソリューション本部 部長

画像診断領域における AI の活用は、検査時・撮影時における画質向上、被ばく線量の最適化、ワークフローの最適化などに 寄与するものと、読影時における診断効率や精度の向上に寄与するものに大別出来る.更に読影支援については、事前適用型 と即時適用型に分類される.これら各分類における AI アプリケーションの典型例を紹介した上で、それらを既存の業務フロー の中に実装する際の課題と、その課題の解決に向けた取り組みを紹介する.

SS3-4 "Abierto Reading Support Solution" 画像診断領域における AI 活用の取組み

○渡辺 和之

キヤノンメディカルシステムズ株式会社 ヘルスケア IT 事業統括部 ヘルスケア IT 事業部 読影支援ソリューション推進責任者

近年, AI による画像解析の開発が加速しているが, 出力結果は DICOM SC 形式が多く修正できず, インタラクティブな読 影ができないなど課題もある.

"Abierto Reading Support Solution"はこれらの課題を解決するため、"Findings Navigator"を開発,解析結果を読影フロー に適したレイアウトで表示,また医師による訂正が可能,さらに 3D 表示できる特長を持っている.これにより様々なアプリケー ションの結果を読影ワークフローに組み込み読影ワークフローの効率化が期待できる.本講演では "Findings Navigator"の他, 最新のアプリケーションも紹介する.

シンポジウム2

13:00 ~ 14:30

SY 2 「横断画像で微細構造がどこまで見えるか」

座長:清水 昭伸(東京農工大学)

花岡 昇平 (東京大学)

【企画の意図】この数年で様々な画像機器が著しく向上してきました.高分解能化に至る手法は検査法毎に様々ですが,高分 解能化により可視化できる世界が少しずつ変化してきています.特別講演2(宮脇先生)の蛍光法で可視化できる世界との間 には大きな隔たりがまだありますが,各検査の高分解能化の原理と可視化できる対象の変化を新たな潮流として感じて頂きた いと思います.

SY2-1 光超音波イメージングによる表在リンパ管の観察

○梶田 大樹¹, 鈴木 悠史¹, 渡部 紫秀¹, 岡部 圭介¹, 佐久間 恒², 大滝真梨香¹, 今西 宣晶³, 陣崎 雅弘⁴, 相磯 貞和³⁵, 貴志 和生¹

¹慶應義塾大学 医学部 形成外科,²東京歯科大学 市川総合病院 形成外科,³慶應義塾大学 医学部 解剖学教室, ⁴慶應義塾大学 医学部 放射線科学教室(診断),⁵株式会社 Luxonus

光超音波イメージングとは、光を照射された物体から発生した超音波(光音響信号)を検出し、非侵襲的に対象を画像化す る技術である.血液中のヘモグロビンから信号を得ることで、微細な血管の画像を得ることができる.慶應義塾大学では、ヒ ト生体のリンパ管内に色素を入れることで、表在のリンパ管を 0.2mm の解像度で 3 次元的に描出することに成功した.本演題 では、リンパ浮腫患者を対象とした臨床研究で得られた興味深い所見について紹介する.

SY2-2 超音波 SMI による微細血管の描出

○畠 二郎

川崎医科大学 検査診断学 内視鏡·超音波部門

Superb Microvascular Imaging (SMI) は従来のカラードプラ法の欠点である motion artifact を軽減した高フレームレート で高分解能,かつリアルタイムな血流の表示法である.SMIを用いて造影剤を使用することなく臓器や病変の微細血流をリア ルタイムに観察することでより詳細な病態の評価が可能となり,すでに多くの新たな知見が報告されている.本セッションで は腹部や表在など様々な領域の炎症や腫瘍における画像を供覧し,その臨床的有用性に関して報告する.

SY2-3 高精細 CT

○岩澤 多恵

神奈川県立循環器呼吸器病センター 放射線科

従来の CT はマトリックスサイズが 512x512,スライス厚が 0.5-0.6mm であったが,高精細 CT は 2048x2048,0.25mm 厚ま で解像度を上げられる.本発表では,我々が日常臨床で使用している高精細 CT の画像で肺の末梢構造がどこまで見えるか説 明する.高精細 CT では Reid の 2 次小葉を認識できるので,COVID-19 肺炎における肺容積減少の評価に役立つ.

SY2-4 7 テスラ MRI 画像による微細脳構造評価

○岡田 知久

京都大学大学院医学研究科 脳機能総合研究センター

7 テスラ MRI 装置は薬機法未承認ですが, EU・米国では既に診療用機器として承認されています. 静磁場強度比 7/3 の 2.3 倍を超える高い信号ノイズ比が得られることが判っており, 脳を中心に小構造の研究が現在, 大いに進められています. ドパ ミン神経が豊富な黒質 nigrosome-1 領域という小領域を対象としたパーキンソン病での評価や大脳皮質の層構造別の賦活検出 などを含めた微細構造評価研究の現状をご紹介いたします.

SY2-5 デジタル (半導体検出器) PET

○堀越 浩幸

群馬県立がんセンター 放射線科 診断部

従来のPET は存在診断には優れていたが、低解像度のため、がんの浸潤度評価は困難であった.しかし、半導体検出器搭載のPET では、高分解能の画像取得が可能になっている.さらに息止めによるPET 画像の収集も可能になりがん進展度評価、リンパ節の転移診断に寄与している。本講演では、半導体検出器搭載PET-CT の特徴、有用性について報告し、MRI などの他 モダリティとの fusion 画像作成についても紹介したい.

特別企画2

14:40 ~ 16:10 SE2

SE2「脳のイメージング最前線」

座長: 增谷 佳孝 (広島市立大学)

【企画の意図】現在,画像研究者の間での大きなトピックの1つが,脳の機能,構造,連絡性,すなわち「コネクトーム」と, 老廃物排泄系である「Glymphatic システム」ではないかと思います.これらの異常は睡眠や痴呆との関連も指摘されており, 立位のイメージングと共に人体機能評価の観点から大きな潮流の1つになると思っています.

SE2-1 T2 解析による Glymphatic pathway の可視化

○押尾 晃一

順天堂大学 医学部 放射線科

近年 glymphatic system という仮説が注目を集めている.これは本来リンパ系の存在しない脳において、リンパ系と同様に 脳内の老廃物を排泄する系で、アルツハイマー病等のメカニズムに深く関わっていることが示唆されている.これまで動物に よる研究がほとんどで、ヒトにおける glymphatic system の実態は明らかになっていない. 今回 T2 データを解析することで 老廃物の排泄経路を可視化することを試みた.

SE2-2 拡散画像で見る間質液動態

○田岡 俊昭

名古屋大学 大学院医学系研究科 革新的生体可視化技術開発産学協同研究講座

脳間質液動態の評価にはトレーサー研究が主に行われてきたが、我々は拡散画像による評価を試みている.血管周囲腔の走 行方向に限定した拡散能が間質液動態を反映すると仮定し、拡散テンソルデータから間質の水分子の動態を評価する手法とし て, diffusion tensor image analysis along the perivascular space (DTI-ALPS)という評価方法を提案した.DTI-ALPS法を用いて、 アルツハイマー病やパーキンソン病、正常圧水頭症など様々な疾患での間質液動態の異常が報告されている.

SE2-3 非侵襲画像技術による霊長類脳コネクトーム解明研究

○林 拓也

理化学研究所 生命機能科学研究センター

ニューロンは他のニューロンと共につながり,共に発火し,共に死に至ります.こうしたニューロンの共生的特性に基づい て人は記憶や認知などの高次脳機能を発揮し,脳の病にも至ります.昨今のMRI技術の進歩により人の大脳皮質分画,機能的 連絡性,皮質微細構造等が明らかになりつつあります.MRI技術を更に高度化しヒト以外の霊長類脳を調べ,人と霊長類脳の 種間共通のコネクトームや病の理解を進める研究プロジェクトを紹介します.

一般演題セッション5

P5「イメージング,再構成,CAS,その他」 $16:20 \sim 17:05$

座長:北坂 孝幸 (愛知工業大学)

健策(名古屋大学) 森

P5-01 2次元フーリエ変換を用いたリングアーチファクト低減フィルターの特性評価

○米山 明男¹. 馬場 理香². 河本 正秀¹

¹九州シンクロトロン光研究センター,²(株)目立製作所研究開発グループ

2次元フーリエ変換を用いたリングアーチファクト低減フィルターは、リングアーチファクトがサイノグラム上の縦ストラ イプとなって現れることを利用して、縦方向のハイパスフィルターと横方向のローパスフィルターの組合せにより、縦ストラ イプ、すなわちリングアーチファクトを低減するフィルターである、単純なフーリエ変換から構成されるため、パラメータの 設定が単純で、高速かつ汎用性が高いことが特徴である。これまでに放射光マイクロ CT で取得した種子、木材、及び化石等 に適用し、CT 値を保持したまま X 線画像検出器に起因したリングアーチファクトを大幅に低減することができた、一方で、 CT 値が大きく変動する領域において、円周方向に CT 値が滲む(尾を引く)ことも判明した.本発表では、本滲みの起源と特 性,及びその対策について検討した結果を紹介する.

P5-02 GP-GPUを用いた光子輸送シミュレーションの高速化

○瀬在 翔太¹,村田 一心²,乳井 嘉之³,尾川 浩一⁴ 1法政大学大学院 理工学研究科, 2国立天文台, 3東京都立大学 健康福祉学部, 4法政大学 理工学部

マンモグラフィにおける光子輸送シミュレーションではコヒーレント散乱を考慮する必要があるため、計算量が大きくなり、 EGS5 のような汎用ソフトでは膨大な処理時間を要する. そこで本研究では、GPUを用いた独自のシミュレーションコードを 開発し、計算速度の向上を試みた.コヒーレント散乱の計算に必要な散乱角の確率密度分布および分子形状因子は EGS5 を参 考にし、C/C++ でコーディングした、開発したコードの有効性を検証するため、簡易的な光子輸送シミュレーションを行い、 計算精度および計算時間を EGS5 による結果と比較した、その結果、自作コードでは EGS5 と同等の結果を 10 倍以上の速度で 計算できることを確認した.

P5-03 深層学習を利用した信号の間引きパターンに頑健な圧縮センシング再構成の基礎検討

○渋井 雅希1. 山登 一輝1. 伊藤 聡志1

1 宇都宮大学大学院 地域創生科学研究科 情報電気電子システムプログラム

圧縮センシング (Compressed Sensing: CS) を MRI に応用する CS-MRI は少数の信号から MR 画像を再構成できるため撮像 時間の短縮化が図れるが,画像再構成には多くの時間を要することや画質劣化などの問題がある.近年,ディープラーニング(DL) を用いた CS の再構成は高速, かつ高画質であることが報告されており, 大きな注目を集めている. 一方で, 学習に多くの時間 を費やすこと、また、信号の間引きパターンが変わると再学習が必要であるなどの課題がある、そこで、本研究では、信号 -画像間学習による深層学習再構成を利用し、信号の間引きパターンによらず良質な画像を再構成できる頑健な方法について検 討を行った. ADMM をネットワーク化した ADMM-CSnet を使用し、複数の信号間引きパターンをネットワークに学習させた ところ単一の間引きパターンを学習したネットワークより高い PSNR を得ることができたので報告する.

P5-04

WGI コンプトン画像再構成条件の実験的検討

匠^{1,2} 菅 ○田島 英朗¹, 吉田 英治¹, 仁科 幹生², 脇坂 秀克¹, 高橋美和子¹, 永津弘太郎¹, 彰^{3,4}, KatiaParodi⁵, 山谷 泰賀¹ 厚至¹,鎌田 圭^{3,4}, 吉川 辻 ¹量子科学技術研究開発機構,²千葉大学,³東北大学,⁴株式会社 C&A,⁵ミュンヘン大学

昨年の発表では、我々が開発を進めている WGI(Whole Gamma Imaging)の小動物実証機におけるコンプトンイメージングに よって、高精細なマウス画像が得られたことを報告した、高精細な画像が得られた要因の一つとして、WGI がフルリング型で あることが考えられるが、実際にどの程度寄与しているのか明らかにはなっていない、そこで、本研究では、画像再構成に利 用する測定データを限定し、フルリングの場合と比較することで、物体を囲むリングの範囲の変化が再構成画像へ与える影響 を検討した.円筒型ファントムの測定データを用い,再構成画像を視覚的に評価した結果,散乱検出器に囲まれる範囲が180° 未満の領域が物体にある場合,画像にアーチファクトが生じたが,物体全域が180°以上囲まれている場合には,フルリングと 同様にアーチファクトのない画像が得られた.よって、WGIコンプトン画像再構成の完全性条件として、対象視野全域を180° 以上囲む必要があることが示唆された.

P5-05 膝半月板の 2.5 次元統計的形状モデル構築における主成分分析と非負値行列因子分解の比較

○繩手 就矢¹, 增谷 佳孝¹

1広島市立大学大学院 情報科学研究科

膝半月板は外傷や加齢によって損傷し、厚さのみならずトポロジーの変化も含む多様な形態変化が観察される.3次元形状 の評価に基づく診断には MRI 検査が使用され、抽出形状の定量的な解析は診断に有用な情報をもたらすと期待される.我々は、 膝半月板形状の厚さのみの情報を抽出した2.5次元の統計的形状モデルの構築を研究している.同モデルの構築では、半月板形 状サンプルが2次元の濃淡画像で表現され、各画素は厚さの値を持つ.本研究では、その画像サンプル群の統計解析に主成分 分析と非負値行列因子分解を用い、基底パターンへの分解結果の視覚的比較、形状再現における残差などの定量比較を行った ので報告する.使用する基底数の影響についても検討し、比較結果に基づいて両手法の特性および得失を議論し、診断に有効 な情報をもたらす2.5次元の統計的形状モデルの構築を目指す.

P5-06 腹腔鏡映像からの手術箇所の認識による手術ナビゲーションの検討

○林 雄一郎¹, 三澤 一成², 森 健策^{1,3}

1名古屋大学大学院情報学研究科,2愛知県がんセンター消化器外科,3国立情報学研究所医療ビッグデータ研究センター

本稿では、腹腔鏡映像からの手術箇所の認識による腹腔鏡下手術ナビゲーションの検討について述べる.近年、手術中の術 野映像を解析して手術を支援する研究が行われている.本研究では、腹腔鏡下手術支援のために腹腔鏡から得られる腹腔鏡映 像から手術箇所を認識して手術ナビゲーションを行う手法を検討する.まず、腹腔鏡映像から畳み込みニューラルネットワー クを用いて手術を行っている解剖学的な領域を認識する.次に、認識結果を用いて手術箇所付近の解剖学的構造が確認できる 仮想腹腔鏡画像を手術支援画像として提示する.提案手法を腹腔鏡下胃切除術において使用し、手術箇所の認識と手術支援画 像の提示が可能であることを確認した.

P5-07 VR Organ Puzzle: A Virtual Reality Application for Human Anatomy Education

○李 斯頎¹,林 雄一郎¹,藤原 道隆²,小田 昌宏³¹,森 健策^{1,4,5}
¹名古屋大学大学院情報学研究科,²名古屋大学医学部附属病院医療機器総合管理部,³名古屋大学情報連携推進本部情報戦略室.⁴名古屋大学情報基盤センター.⁵国立情報学研究所医療ビッグデータ研究センター

Human body is a complicated system and consists of numerous organs. Knowing human anatomy is essential for elementary education and science popularization. Until now, the education and popularization are mainly implemented by textbooks and educational materials. A more immersive and intuitive way is necessary. Modern virtual reality (VR) technology based on head mounted display can provide an immersive 3D environment and an intuitive interactive method, which has already shown the potential in educational field, like medical and military training. In this paper, we developed a virtual reality application, VR Organ Puzzle, to help with the education of human anatomy. Some abdominal organs are modeled from the segmentation result of CT volumes and rendered in the VR environment. Users can grasp and rotate organ models by controllers in an intuitive way. By returning organ models to proper positions and orientations, users can learn correct human anatomy knowledge.

P5-08 複数術野視点映像を用いた機械学習による外科手術工程の自動分類

○藤井 亮輔¹, 梶田 大樹², 高詰 佳史³, 青木 義満¹

¹慶應義塾大学理工学研究科 総合デザイン工学専攻 電気電子工学専修,²慶應義塾大学医学部 形成外科,³慶應義 塾大学医学部 解剖学教室

近年,機械学習を活用した画像認識の発展に伴い数多くの研究が行われている中,外科手術シーンにおける術野映像を用い た研究は術野撮影でほとんどない.しかし今回,新たに制作されたマルチカメラ搭載型無影灯を用いてデータとして十分な視 認性を持つ外科手術映像を入手した.本研究ではこの映像データを機械学習に用いた外科手術における工程分類精度の検証を, 術野映像の入力を単一視点と複数視点の2パターンで比較し行なった.検証の結果,外科手術の工程分類における複数視点入 力の有効性を明らかにすることができた.この結果は,手術映像の管理や編集の手間の自動化の更なる高精度化をもたらし, 医療ミス発見の易化や術者の技能評価につながるデータ活用を促し,医療発展への貢献の可能性を示した.

P5-09 腹腔鏡の視野外における術具誤接触防止システムの開発

〇岸 知哉¹, 中口 俊哉², 長尾 俊哉³, 駒村 賢三⁴, 西澤 祐吏⁵

¹千葉大学大学院融合理工学府基幹工学専攻医工学コース,²千葉大学フロンティア医工学センター,³京セラ株式会社 東京事業所,⁴株式会社ソフケン,⁵国立がん研究センター東病院

腹腔鏡下手術は視野が腹腔鏡映像のみのため視野が狭い.また,特にロボット支援腹腔鏡下手術の場合は,腹腔鏡の視野外 において術具と近接組織との誤接触が発生し医療事故に繋がる.この腹腔鏡視野外での術具誤接触を防止するため、本研究で は術具先端と近接組織間の距離を計測し,危険と判断されれば術者に音声や画面表示などで警告を行うシステムを提案する. このシステムを実現するため先端に距離センサと小型カメラを装着した専用トロッカーを考案した.まず距離センサを用いて トロッカー・組織間の距離を計測し,小型カメラで撮影した画像よりトロッカーから突出する術具の長さを計測する.これら の距離の差分から術具先端と組織間の距離を算出する.また,警告対象となるのは腹腔鏡視野外であるため,術具が腹腔鏡の 視野外に位置しているか判定する手法を提案した.ドライボックス環境にて提案システムの評価実験を行った結果,有用性が 示唆された.

P5-10 取り下げ

以降の予稿は,著者から提出された原稿を査読なしで, そのまま掲載しています.

著者らが自らの内容を MEDICAL IMAGING TECHNOLOGY

(MIT) 誌へ投稿する場合,その原稿にこの PDF に類似した内容が 含まれていても,二重投稿とは判断しません.

発行日:2021年10月12日

発行者:陣崎 雅弘

製作所:トーヨー企画株式会社

胸部 CT 像からの COVID-19 症例の自動分類手法

小田 昌宏^{*1,2} 鄭 通^{*2} 林 雄一郎^{*2} 大竹 義人^{*3,4} 橋本 正弘^{*5} 明石 敏明^{*6} 青木 茂樹^{*6} 森 健策^{*2,1,4}

要旨

本稿では、胸部 CT 像をもとに新型コロナウイルス感染症(COVID-19)症例の可能性が高いか低いかを自動分類する手法を述べる. 2019 年 12 月に発生した COVID-19 は発生から 1 年以上経過した現在も多数の 感染者を生じており、世界の人々の生活様式を変化させ経済活動に悪影響を与えている. COVID-19 を含 むウイルス性肺炎の診断において胸部 CT 像が有用であり、医師の診断支援のため、CT 像を用いた COVID-19 診断支援手法の確立が求められる.本稿では深層学習を用い、COVID-19 の画像所見典型度に基 づいて CT 像を自動分類する手法を提案する.まず COVID-19 により生じる異常陰影に対応した肺野セグ メンテーション手法を用いて肺野の正常及び異常陰影領域をセグメンテーションする.セグメンテーショ ン結果を用いて肺野ボリューム画像を作成し、これを COVID-19 典型度推定 3D CNN を用いて自動分類す る.分類先クラスは COVID-19 症例の可能性が高いまたは低いの 2 クラスである.3D CNN は局所及び大 域的な画像特徴を抽出可能な構造とした.また、3D CNN の学習においては 3D 画像変形を用いたデータ拡 張を利用し、多様な肺野形状に対応可能とした.提案手法を COVID-19 感染例および非感染例の胸部 CT 像 818 例に適用し 84.8%の精度で自動分類が可能であることを確認した.

キーワード: COVID-19, 診断支援, CAD, 分類, 深層学習

1. はじめに

新型コロナウイルス感染症(COVID-19)が 急速に世界に広まり,現在も多くの感染者と 死亡者を生じている.2021年7月末時点の世

*1 名古屋大学情報連携推進本部情報戦 略室

〔〒464-8601 名古屋市千種区不老町〕 e-mail: moda@i.nagoya-u.ac.jp

- *2 名古屋大学大学院情報学研究科
- *3 奈良先端科学技術大学院大学先端科学技術研究科情報科学領域

*4 国立情報学研究所医療ビッグデータ研究センター

*5 慶應義塾大学医学部放射線科

*6 順天堂大学医学部放射線医学講座

投稿受付: 2021 年7月30日

界の COVID-19 累計感染者数は約1億9千万 人,死亡者数は約4百万人に達している[1]. 本邦の累計感染者数は約86万人,死亡者数は約1万5千人である[2].感染者への適切な治 療実施と新規感染発生防止のため,正確かつ 迅速な検査法が求められる.

胸部 X 線 CT 像(以下 CT 像) は短時間で 撮影可能であり, CT 像を用いた COVID-19 検査は 97%の感度を持つことが示されている [3]. 日本は人口あたりの CT 装置保有数が高 いため, COVID-19 検査で CT 像を活用でき ると考えられる. CT 像を用いた医師の診断 における見落とし軽減と判断の定量化のため, COVID-19 診断支援システムの実現が求めら れる.

本稿では, COVID-19 典型度に基づいて CT 像を自動分類する手法を提案する. COVID-19 典型度は医師の画像所見に基づいて決定した 分類であり、COVID-19 に典型的な画像所見 を示すかそうでないかの2種類である.本手 法では CT 像に対して肺野領域セグメンテー ションを適用し、その結果を利用して肺野領 域の絞り込みを行う.そして 3D CNN

(convolutional neural network)を用いて典型 度に準じた分類を行う.

2. 手法

2.1 肺野領域セグメンテーション

文献[4]の手法を用いて CT 像から肺野領域 セグメンテーションを行う.本手法は体軸方 向に連続した 3 枚の Axial 画像(2D 画像)を 入力とし, FCN (fully convolutional network) を用いてセグメンテーション結果を得る. Axial 画像単位のセグメンテーション結果を 再構成することで 3D の肺野領域セグメンテ ーション結果とする.

文献[4]の手法は肺野内を正常領域,異常陰 影領域の2種類へと分割して出力するが,こ の後の処理では正常領域,異常陰影領域を区 別せず扱うため,出力された2種類の領域の 論理和を肺野領域とする.

2.2 COVID-19 典型度に基づく分類

CT 像と肺野領域を入力とし, 3D CNN を用 いて COVID-19 に典型的な画像所見を示すか そうでないかの 2 種類へと自動分類する.

1) 肺野ボリュームの作成

CT像から,肺野領域を1画素以上含むAxial 画像を取り出す.取り出した Axial 画像上で 肺野領域以外の領域内の画素値は特定の値で 置き換える.これらの処理後の Axial 画像を3 次元的に再構成し,肺野のみを含む肺野ボリ ューム画像を作成する.肺野ボリューム画像 は 160x160x160 画素のサイズへとスケーリン グする.

2) Data augmentation

様々な肺野形状パターンを分類モデルに学 習 さ せ る た め , 形 状 に 関 す る data augmentation を行う.具体的には,学習用 CT 像から得られた肺野ボリューム画像に対して 3D の回転,平行移動,非剛体変形を適用し, 生成肺野ボリューム画像を作成する.

3) 分類モデル

3D CNN を分類モデルとして用いる.本モ デルは 3D 画像を入力, COVID-19 に典型的な 画像所見を示すかそうでないかの 2 クラスそ れぞれの尤度を出力とする.

本モデルは、2 つの畳み込み層と max pooling 層からなるユニットを複数持つ. 画像 空間における様々な範囲からの特徴量抽出を 行うことを目指し、一部の畳み込み層には dilated convolution block を使用した. Dilated convolution block の中では複数の dilation rate の dilated convolution[5]を行い、その結果を統 合する. 最終ユニットの後に global average pooling 層を持ち、2 クラスそれぞれの尤度を 出力する. 尤度の高いクラスが分類結果とな る.

4) 学習と分類

学習用 CT 像から作成した肺野ボリューム 画像と生成肺野ボリューム画像を用いて分類 モデルの学習を行う.学習の後,自動分類を 行う CT 像を用いて処理を行い,分類結果を 得る.

3. 実験及び考察

本邦の複数の医療機関で撮影された COVID-19 感染例及び非感染例の CT 像 818 例を用いて提案手法の評価を行った.このう ち 654 例を学習,164 例を分類精度評価に用 いた.Data augmentation では生成肺野ボリュ ーム画像を 1600 作成し学習に使用した.分類 モデル学習時のミニバッチ数は 2, epoch 数は 100 とした.分類の正解ラベルは医師の画像 所見に基づいて決定した.提案手法による 164 例の評価用 CT 像の分類精度は 84.8% とな った.

提案手法では良好な精度で CT 像を自動分 類することが可能であった.本成果は COVID-19 診断支援システムを実現する上で 重要であり,臨床での医師の判断誤り軽減な どに貢献するものと考えられる.

COVID-19 の影響を判断する上で肺野状態

の確認が重要となる. CT 像の撮影範囲は様々 であり,肺野だけでなく腹部等が広範囲に画 像に含まれている場合がある.そのため自動 分類の前処理として肺野領域セグメンテーシ ョン手法を用い,自動分類に用いる画像範囲 を適切に限定することが必須である.提案手 法では文献[4]のセグメンテーション手法を 用いることで,自動分類の処理対象を肺野内 の正常及び異常陰影領域のみに適切に限定す ることができた.この点は CT 像全体を用い て自動分類を行う場合と比較し,分類精度向 上に貢献していると考えられる.

自動分類モデルとして dilated convolution block を用いた 3D CNN を使用した.様々な 空間的広がりを持つ異常陰影から分類に貢献 する特徴を抽出する際,様々な dilation rate を使用した dilated convolution block が効果的 に働いていると思われる.

3. むすび

本稿では COVID-19 の診断支援を目的とし た CT 像自動分類手法を提案した.提案手法 では, CT 像から肺野領域をセグメンテーシ ョンした後, 3D CNN を用いて COVID-19 に 典型的な画像所見を示すかそうでないかの 2 クラスへと症例の自動分類を行う. 3D CNN では dilated convolution block を導入して肺野 内の広範囲の空間的特徴を利用した分類を実 現した. CT 像 818 例を用いた実験では,提 案手法は評価用 CT 像 164 例を 84.8%の精度 で自動分類することができた. 今後の課題と して,分類モデルの構造の変更,より広範囲 の CT 値を持つ異常陰影領域への対応などが 挙げられる.

謝辞

本研究の一部は, JSPS 科研費 17H00867,

17K20099, AMED 研究費「医療ビッグデータ 利活用を促進するクラウド基盤・AI 画像解析 に関する研究」JP19lk1010036, JP20lk1010036, AMED 研究費「画像診断ナショナルデータベ ース実現のための開発研究」JP20lk1010025, NICT 研究費「COVID-19 肺炎の CT 画像を AI 解析するためのプラットフォーム開発と 実証展開」222A03, JSPS 二国間交流事業, 公益財団法人 堀科学芸術振興財団,公益財団 法人 人工知能研究振興財団によった.

利益相反の有無

なし

文 献

- [1] Coronavirus Update, https://www.worldometers.info/coronavirus/ (2021 年7月26日アクセス)
- [2] 厚生労働省 国内の発生状況など、
 https://www.mhlw.go.jp/stf/covid-19/kokun
 ainohasseijoukyou.html
 (2021 年 7 月 26 日アクセス)
- [3] Ai T, Yang Z, Hou H, et al.: Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. Radiology 296(2): 2020
- [4] Oda M, Hayashi Y, Otake Y, te al.: Lung infection and normal region segmentation from CT volumes of COVID-19 cases. Proc. SPIE Medical Imaging 11597, 115972X-1-6, 2021
- [5] Yu F, Koltun V: Multi-scale Context Aggregation by Dilated Convolutions. ICLR, 2016

Automated Classification Method of COVID-19 Case

From Chest CT Volume

Masahiro ODA^{*1,2}, Tong ZHENG^{*2}, Yuichiro HAYASHI^{*2}, Yoshito OTAKE^{*3,4}, Masahiro HASHIMOTO^{*5}, Toshiaki AKASHI^{*6}, Shigeki AOKI^{*6}, Kensaku MORI^{*2,1,4}

- *1 Information Strategy Office, Information and Communications, Nagoya University
- *2 Graduate School of Informatics, Nagoya University
- *3 Division of Information Science, Graduate School of Science and Technology, Nara Institute of Science and Technology
- *4 Research Center for Medical Bigdata, National Institute of Informatics
- *5 Department of Radiology, Keio University School of Medicine
- *6 Department of Radiology, Juntendo University

We propose an automated method to classify a chest CT volume into classes that correspond to likelihood to be COVID-19 case. COVID-19 is spreading over the world, causing the large number of infected patients. In the diagnosis of viral pneumonia, including COVID-19, CT volume is quite helpful. A diagnosis assistance method for COVID-19 cases that utilizes CT volume is required. This paper proposes an automated classification method that classifies a chest CT volume into two classes, including high and low likelihoods of COVID-19 infection. We apply a lung region segmentation method to a CT volume that can segment regions, including lung normal and infected regions. We generate a lung volume using the CT volume and the segmentation result. A 3D CNN classifies the lung volume for COVID-19 classification. The 3D CNN has dilated convolution blocks that extract both local and global spatial features from a volume. We used 3D data augmentation method to 818 CT volumes, including COVID-19 cases and non-COVID-19 cases. The classification accuracy of the proposed method was 84.8%.

Key words: COVID-19, Computer-aided diagnosis, CAD, Classification, Deep learning

Self-attention Class Balanced DenseNet_LSTM framework for

Subarachnoid Hemorrhage CT image Classification

on Extremely Imbalanced Brain CT Dataset

Zhongyang LU^{*1}, Masahiro ODA^{*2,1}, Yuichiro HAYASHI^{*1},

Tao HU^{*1}, Hayato ITOH^{*1}, Takeyuki WATADANI^{*3}, Osamu ABE^{*3},

Masahiro HASHIMOTO^{*4}, Masahiro JINZAKI^{*4}, Kensaku MORI^{*1,5,6}

Abstract

In this paper, we propose a novel class imbalanced network framework with a self-attention mechanism, SACB-DenseNet_LSTM, to handle the data-imbalanced issue in the Subarachnoid hemorrhage classification mission in order to classify the subarachnoid hemorrhage CT scans from the extremely imbalanced brain CT dataset. We redesign the Class Balanced DenseNet_LSTM to introduce the self-attention mechanism for improving the classification performance. In our proposed framework, we utilize a self-attention mechanism to build up a self-attention block to estimate the feature importance of individual features extracted by the feature extractor (here, we employed DenseNet-121 with class balanced loss as our feature extractor). Moreover, we integrated the feature importance vector with the input feature and fed them to the LSTM sequentially to acquire the consecutive information from adjacent CT scans to further improve the classification performance. To evaluate our method's effectiveness, we conducted experiments on an extremely imbalanced brain CT dataset, including 33 cases with subarachnoid hemorrhage and 2519 cases of normal cases. Our experiments showed that after inducting the self-attention mechanism, in contrast to the previous work, the classification performance improved from 0.726 F-score to 0.734 F-score. This indicates our framework is more effective toward the imbalanced data issue in the Subarachnoid Hemorrhage classification.

Keywords : self-attention, imbalanced classification, DenseNet_LSTM, Subarachnoid Hemorrhage classification

*1 Graduate School of Informatics, Nagoya University, Aichi, Japan [Furo-cho, Chikusa-ku, Nagoya, 464-8601, Japan] e-mail: zlu@mori.m.is.nagoya-u.ac.jp

*2 Information Strategy Office, Information and Communications, Nagoya University, Aichi, Japan

*3 Department of Radiology, The University of Tokyo Hospital, Tokyo, Japan

*4 Department of Radiology, Keio University School of Medicine, Tokyo, Japan

*5 Information Technology Center, Nagoya University, Nagoya, Aichi, Japan

*6 Research Center for Medical Bigdata, National Institute of Informatics, Tokyo, Japan

1. Introduction

Subarachnoid Hemorrhage (SAH) is a fatal disease caused by acute hemorrhage in the subarachnoid space. Conventionally, the diagnosis of SAH is mainly dependent on medical image technology, Computerized Tomography (CT) or Magnetic Resonance Imaging (MRI). Meanwhile, CT is the first choice for SAH diagnosis [1]. Figure 1 demonstrates an example for SAH CT slice and no-SAH CT slice, respectively. Considering the tremendous workload of detecting SAH CT slice by slice, an AI-based automatic SAH diagnosis system is wished to be developed.

Several previous researches have gained many remarkable achievements, as the deep Learning based method, Chilamkuthy [2] used a base CNN to detect ICH CT scans and compared the model's performance to that of radiologists. Phong [3] utilized three types of convolutional neural networks: LeNet, GoogLeNet, and Inception-ResNet to diagnose brain hemorrhage. Another approach to address that Arbabshirani [4] trained a 3D CNN model on a huge ICH dataset, including 46,583 ICH cases (around 2 million images). Whereas, as the most situations in the real world, in contrast to no-SAH, SAH may not frequently happen in the clinical situation. Moreover, most previous works on imbalanced SAH CT image classification are conducted on the balanced dataset. Lu proposed a Class Balanced DenseNet_LSTM framework [5] to handle the imbalanced data challenge in SAH CT classification on an extremely imbalanced dataset.

In this study, we refer to Lu's work [5] and proposed a novel framework named SACB-DenseNet_LSTM, which introduces a self-attention mechanism to improve the classifier's performance. The key idea is to learn the feature importance vectors via Self-Attention Network (SAN) from the represented features extracted by the feature extractor. Using this framework, we show that the self-attention mechanism can improve the SAH CT classification even on an extremely imbalanced dataset.

2. Methodology

Overview. In this section, we demonstrate an overview of our proposed method. Inspired by the self-attention mechanism, we utilized an attention-based neural network architecture to estimate the feature importance of the input feature vector. We consider that feature importance vectors indicate the important feature that contributes to the prediction. We combine the feature importance vector with the input feature vector to inform a new vector in order to increase the classifier's performance.





Fig. 1 Example of no-SAH CT and SAH CT, where the gray arachnoid texture in the red bounding box of the right picture is SAH regions.

feature importance



Fig. 2 Illustration of the producing feature importance vectors. The green area demonstrates the structure of the Self-Attention Network. The key part of SAN is the attention layers, including multi-head attentions, where the element-wise product with input space is computed on each forward pass.

Feature importance estimation. Feature importance estimation refers to an approach to estimate the importance of each feature in the input feature space, and generally, we consider the crucial parts that contribute to the final decision of classifiers. The simplest example of feature importance estimation is dividing the features into similar and dissimilar ones. A change of perspective, it is similar to the feature selection. In this paper, we referred to the Blaz's work [6], and proposed a class balanced self-attention network to estimate the importance of the class imbalanced feature space.

Self-Attention Networks. Blaz firstly proposes Self-Attention Network (SAN) approach, and the idea is from the works about the attention mechanism [7][8]. The Self-Attention Network (SAN) architecture can be described as the following formula in mathematics,

$$\Omega(\mathbf{X}) = \frac{1}{k} \bigoplus (\mathbf{X} \otimes softmax(\mathbf{W}_{latt}^{k} \mathbf{X} + \boldsymbol{b}_{latt}^{k}))$$
(1)

where the attention layer Ω with multi-attention heads is the key part of the SAN (*k* represents the number of attention heads), and we only use this attention layer to analyze feature importance. Therefore, unlike the basic SAN, we do not count on SAN to make predictions. Hence we remove the last linear layer, and output the feature importance vectors directly. The illustration of this process is illustrated in Fig. 2. **X** represents the feature vectors set, which is produced by Convolutional Neural Network from input data set. We describe the size of **X** is $|I| \times |F|$ in Fig. 2. We employed *I* to represents the assembly of CT images, and |I| is the number of images in the dataset. *F* is the feature vector extracted by the feature extractor. |F| is the number of features in one feature vector. Additionally, the signal \bigoplus represents the Hadamard summation and \otimes is Hadamard product. W_{latt}^k and b_{latt}^k represent the weights and bias of attention layers respectively and the size of W_{latt}^k is a learnable $|F| \times |F|$ weight matrix which can be considered as the


Fig. 3 Overview of proposed SACB_DenseNet_LSTM. We integrate the extracted features and their feature importance vectors, and feed the new vectors to LSTM to compute the predicated label probability.

relations between features. The main point of SAN is discovering the relation of a given feature with itself. The network output is a set of vectors. Each element of the vector points out the importance of each feature to the prediction.

Sequential and attention information combination. Due to SAH regions showing spatial continuity characteristics in the cases, we consider a latent relationship between the feature importance vectors. Hence, we combine a feature vector and its feature importance vector and utilize LSTM to dig the latent relationship to improve the classifier's performance. Figure 3 illustrates this process.

3. Experiments

Datasets introduction and partition. Two dataset were used in our experiments. The first one, 19,530 cases brain CT volumes collected by RSNA, including 8,003 cases with ICH and 11,527 no hemorrhage cases, were applied as the source domain for pre-training. In the 8,003 ICH cases, every CT slice was labeled with six kinds of labels, including five kinds of subtypes of ICH covering subarachnoid hemorrhage (SAH), subdural hemorrhage (SDH), epidural hemorrhage (EDH), intraparenchymal hemorrhage (IPH), and intraventricular hemorrhage (IVH). As our main experiment dataset, the second one was collected by the National Institute of Informatics (NII), including 2,519 cases no hemorrhage brain CT and 33 cases brain CT labeled with SAH or not slice-by-slice. We used 1,632 cases (including 20 cases SAH data) as training dataset, 409 cases (including 6 SAH cases), and 511 cases (including 7 SAH cases) as validation and test data respectively. We classify CT images slice by slice in the work, so the dataset is extremely imbalanced (607 slices SAH CT and 108,737 no-SAH CT).

Data augmentation strategy. Half of the CT images in training data were randomly selected to be data augmented in every training period, including rotation (around ± 10 degree), in addition vertical flip and horizontal flip were also used in data augmentation.



Fig. 4 Feature importance distributions. The picture illustrates the feature importance distribution of feature importance vectors in the four prediction situations. The abscissa is the features in the feature vector, and the ordinate is the attention value (from 0 to 1). We noticed that the SAH and No-SAH have entirely different feature importance distributions, but the same predictions show almost same distribution. False classification samples seem like the hybrid of the TP and TN.

Experimental setting. We conduct all experiments by the Pytorch framework, and we choosen the DenseNet-121 as our backbone network. On the pre-training procedure, we trained the DenseNet-121 for 200 epochs to get available weights on the RSNA dataset. Then, we train the DenseNet-121 for 200 epochs on NII data as a feature extractor. Following Lu's work [5], we employ Class balanced loss [9] (CB-loss) as our objective function. Undergoing feature extraction processing, a piece of CT slice was represented by a 1024 dimensional feature vector, so the |F| is 1024 in this work. Then, we input these feature vectors into SAN, and train SAN for 200 epochs. Finally, we integrate the feature and their importance vectors. We input the new vectors to the LSTM frameworks and trained for 200 epochs.

4. Results and discussion

Results. Table 1 shows the quantitative evaluation results of our proposed method and previous methods. We compared our method with several previous works, and as Table 1 indicates, no matter Densenet-121 or CB_DenseNet_LSTM, when they introduced self-attention network, the performance have made superior improvement, and our proposed method SACB_DenseNet_LSTM has achieved the best performance, improving the classification performance from 0.726 F-score to 0.734 F-score. Additionally, compared with the DenseNet, the model's introduced SAN performance has improved from 0.638 F-score to 0.691 F-score. Meanwhile, the precision score has been increased sharply from 0.5777 to 0.704.

For more intuitive comprehension, we visualize the feature importance vectors to reveal features' contributions to the prediction. As Fig. 4 shows, we noticed that the SAH CT or no-SAH CT have an entirely different feature importance distribution, and the CT scans in the same predictions show almost the same feature importance distribution. For example, in the true positive samples, the 619th and 857th feature play a vital role in positive identification. Similarly, in the true negative samples, the 863th feature plays decisive role in the prediction. Additionally, the first, 669th and

Experiments	F-measure	Recall	Precision
DenseNet-121	0.638	0.721	0.577
DenseNet+SAN	0.691	0.679	0.704
CB_DenseNet_LSTM [5]	0.726	0.936	0.594
SACB_DenseNet_LSTM	0.734	0.927	0.608

Table 1 Quantitative evaluation results of experiments.

880th features have high response value, implying the high feature importance value of these feature is a character of negative samples.

Discussion. As Fig. 4 presents, no matter the geometry, the feature importance distributions of feature importance vectors in the same prediction are almost the same, i.e., the high response value of some features reveals the high probability of the samples belonging to certain predictions. We consider which contains some latent information and can help improve classification performance. Hence we combine the feature importance vectors and the feature vectors and analyze them via LSTM. Combined with Table 1, the results confirmed our hypothesis, the performance improved, which means the feature importance can contribute to classification. Whereas, we notice that even though the SAN can increase the precision score, the shortcoming of the degraded precision score after using LSTM still exists, which indicates the method can still be improved.

5. Conclusions

In our studies, motivated by the self-attention mechanism, we refined the original Self-Attention Network (SAN) and introduced CB_DenseNet_LSTM to propose a new SACB_DenseNet_LSTM framework. The proposed method improves the classification performance by estimating the importance of each element of the feature vector. Feature importance vector was obtained via SAN.

As the results showed, in contrast to the previous work, the classification performance improved from 0.726 F-score to 0.734 F-score. This indicates our framework is more effective toward the imbalanced data issue in the Subarachnoid Hemorrhage classification. Nevertheless, the relatively low precision score is still an unsolved problem. For fixing the problem, some methods will be considered in the future.

Competing interests

None.

Acknowledgement

Thank for the help and advices from the members of Mori laboratory and AMED NII team. A part of this research was supported by the AMED Grant Numbers JP19lk1010036, JP20lk1010036, and MEXT/JSPS 26108006, 17H00867.

References

- [1] Abraham M K, Chang W T, et al.: Subarachnoid hemorrhage. Emergency Medicine Clinics 34(4): 901-916, 2016
- [2] Chilamkurthy S, Ghosh R, Tanamala S, et al.: Deep learning algorithms for detection of critical findings in head CT scans: a

retrospective study. The Lancet 392: 2388-2396, 2018

- [3] Phong T D, Duong H N, Nguyen H T, et al.: Brain hemorrhage diagnosis by using deep learning. International Conference on Machine Learning and Soft Computing: 34-39, 2017
- [4] Arbabshirani M R, Fornwalt B K, Mongelluzzo G J, et al.: Advanced machine learning in action: identification of intracranial hemorrhage on computed tomography scans of the head with clinical workflow integration. NPJ digital medicine **1**(1): 1-7, 2018
- [5] Lu Z, Oda M, Hayashi Y, et al.: Extremely imbalanced subarachnoid hemorrhage detection based on DenseNet-LSTM network with class-balanced loss and transfer learning. Medical Imaging 2021: Computer-Aided Diagnosis. International Society for Optics and Photonics 11597: 115971Z, 2021.
- [6] Škrlj B, Džeroski S, Lavrač N, et al. Feature importance estimation with self-attention networks. arXiv preprint arXiv:2002.04464, 2020.
- [7] Chorowski J, Bahdanau D, Serdyuk D, et al.: Attention-based models for speech recognition. in Advances in Neural Information Processing Systems 1:577-585, 2015.
- [8] Vaswani A, Shazeer N, Parmar N, et al.: Attention is all you need, in Advances in Neural In- formation Processing Systems: 5998–6008, 2017.
- [9] Cui Y, Jia M, Lin TY, et al.: Class-balanced loss based on effective number of samples. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition: 9268-9277, 2019

AI による乳がん腫瘤影の良悪性鑑別処理における

画質の相違が CAD 性能に与える影響の分析

工藤 颯馬*1 安倍 和弥*2 武尾 英哉*2

永井 優一*3 縄野 繁*4

要旨

近年,医療の現場では体内撮影装置の高精細化により生成される医用画像の量が増大し,そしてそれを 読影する医師の負担も急増している.この問題を解決するために,診断の補助としてコンピュータ画像支援 診断(CAD)の機能を向上させる必要がある.そのひとつとして,本研究では乳がん腫瘤を対象とし,AI深 層学習の1つである転移学習と,事前に画像の回転・明度・コントラストを調整した画像を用いて学習さ せる画像拡張を用いて乳がん検診における乳がん腫瘍良悪性鑑別処理の開発を行った.また,検診画像と 診断画像の比較により画像診断の精度改善にどのような要素が必要か分析した.検診画像を用いた良悪性 鑑別処理の制度は85%と医師の正診率に近い値になったが,診断画像の90%には及ばなかった.その結果 が画像比較によりノイズとコントラストの悪さが原因であると判明し,これらの改善を行うことが検診に おけるより精度の高い鑑別CADシステムの構築に必要だと分析した.

キーワード:乳がん,検診,診断,CAD,良悪性鑑別

1. はじめに

1.1 背景

近年, 医療の現場では CT や MRI など撮影装 置の高精細化により生成される医用画像の量 が増大している.撮影される画像の枚数は患者 一人あたり数百枚にもおよび,それを読影する 医師の負担も急増している.そのため診断の補 助を行うコンピュータ画像支援診断(Computer Aided Diagnosis/Detection: CAD)の開発が非常 に重要視されている[1].

- *1 神奈川工科大学 大学院 工学研究
- 科 電気電子工学専攻
- 〔〒243-0292 厚木市下荻野 1030〕

e-mail: abex0930@ele.kanagawa-it.ac.jp

- *2 神奈川工科大学 工学部電気電子情報工学科
- *3 国立がん研究センター東病院
- *4 新松戸中央総合病院

乳がん良悪性鑑別の領域で CAD 研究の成果 は、例えば井上謙一の手法[2]のような、診断画 像において Alexnet を利用して 96%の正診率の 高性能の CAD システムの構築や、中川らの鋸 歯形状領域検出法[3]のような腫瘤辺縁形状認 識による鑑別結果の根拠を明確に医師に指示 することができるもの、福岡らの手法[4]では AI 学習のためのデータベースに日本人の画像を 用い ANN (Artificial Neural Network) に医師が 鑑別を行う際の診断基準を特徴量として定量 化する CAD システム開発が過去に行われてい る.

しかし、これらの研究で用いている画像は診 断マンモグラフィのものであり、本邦では、検 診マンモグラフィで乳がんの良悪性鑑別を行 う試みはいまだ少ない.乳がんの検診に不可欠 なマンモグラフィはX線撮影であり、被ばくの リスク、余命損失について飯沼武の研究[5]では 若年層ほど大きくなるという研究結果が存在 する.

1.2 マンモグラフィの区分について

乳腺診療における乳房画像検査として、マ ンモグラフィと乳房超音波検査は必須であ る.乳がん検診マンモグラフィ要精査者に対 しては、拡大スポット撮影などの特殊撮影の 追加や乳房超音波検査が行われる.

日本のマンモグラフィガイドラインのカテゴ リーは、乳がん検診のマネジメントに使用で きるが、精密検査機関や乳腺専門外来におけ る乳房画像検査に求められるマネジメントに 使用できない.

この問題を解決すべく日本乳癌学会は、乳が ん検診関連2団体(日本乳癌検診学会・日本 乳がん検診精度管理中央機構)の協力のも

と、「検診カテゴリーと診断カテゴリーに基づ く乳がん検診精検報告書作成マニュアル」を 作成した.

乳がん検診におけるマネジメントは要精査 か精査不要の判断である.日本では、従来の 悪性確信度を表した日本のマンモグラフィガ イドラインカテゴリーを「検診マンモグラフ ィカテゴリー」と改称し、検診カテゴリー3以 上を要精検とする.

診断カテゴリーは乳房画像検査診断カテゴ リーを意味しており、乳腺診療のマネジメント に直結するカテゴリーである。乳がん検診で実 施されるマンモグラフィを「検診マンモグラフ ィ」、精密検査機関や乳腺専門外来において実 施されるマンモグラフィを「診断マンモグラフ ィ」と称し、マンモグラフィ検査は「検診」と 「診断」を必ず区別することとする。[6]

検診マンモグラフィ段階で乳がんの良悪性 鑑別を実現することができれば、診断マンモグ ラフィ撮影で二方向撮影の良悪性鑑別を削減 し、すぐに生検に移ることが出来る.

本論文では、そのために乳がん検査画像の画

質の違いが CAD の性能に与える影響を分析し, 改善点を明らかにすることによって,検診マン モグラフィ画像を使用して AI による乳がんの 良悪性鑑別を可能にするために必要な画像処 理を導出することを目的とする.

2. 検診画像と診断画像の相違

2.1 診断用 X 線画像の画質評価

診断用 X 線画像の画質は,明るい部分・暗い 部分の明暗の差(明暗比)である画像コントラ スト,画像の微細部分の明瞭性を表わす物理的 尺度である 鮮鋭度,フィルム上の濃度ばらつき に起因する粒状のランダムなテクスチャであ る粒状性によって評価される.

この時画像コントラストは散乱線の影響を 受け低下し,粒状性はノイズによって低下する 特性がある.

画像処理において粒状性は鮮明度とコント ラストの 2 つに負の相関性がある.鮮明度とコ ントラストは正の相関関係である.

2.2 ノイズの影響

ノイズの量は粒状性に影響する.X線撮影に おいて支配的なノイズは量子ノイズ(フォトン ノイズ)であり,X線量子数とそのゆらぎによ るX線検出量の統計的変動から発生する.ノイ ズレベルが高い(粒状性が悪い)と,コントラ ストの低い信号はノイズに埋もれてしまうた めに識別することが困難である.(図1)[7].

ノイズを低減する方法として抜本的な改善 方法である線量増加は,被ばく線量の増加につ ながる.よって,感度の良いシステムを使用す る,ノイズ除去処理を行うといった方法がとら れる.しかし,粒状性はコントラストと負の相関 があり,単純にノイズ除去処理を行うと高周波 領域に影響するコントラストが悪くなり,画像 のぼけがひどくなる可能性がある.



図 1 ノイズの影響(文献[7]より引用)

2.3 散乱線の影響

X線菅から被写体に照射された X線の一部 は透過(通過)し,一部は被写体内に吸収される. ただ,被写体を透過できなかった X線の全てが 被験者に吸収されるわけではない.吸収されな かった X線のなかには,体内にある臓器や骨,ま たは被験者が寝ている台も含め,あらゆるモノ と衝突して,反射するように散乱するものがあ る[8].このあらゆるモノに衝突して,いろんな方 向に飛び散ってしまう X線のことを散乱線と いう.照射した X線(一次 X線)が人体内で相 互作用して二次的な X線(散乱 X線)が発生 し,X線画像に影響を及ぼす.かぶりやボケを発 生させ,画像のコントラストを低下させる要因 である(図 2).

散乱線を低減する方法として管電圧を低く する,被写体を圧迫する(図2上部),X線菅に備 え付けられている照射範囲を限定する「絞り」 (図2下部)がある.

X線が照射されない範囲は散乱 X線が発生す ることもなければ,被験者の被ばくにも影響し ない.X線菅から放出される X線の大きさと 形状を診断に必要な領域に制限することで,散 乱 X線の発生を抑え,画質の劣化を防ぐことが できる.



2.4 検診画像と診断画像の相違

乳がん検診で実施されるマンモグラフィを 「検診マンモグラフィ」、精密検査機関や乳腺 専門外来において実施されるマンモグラフィ を「診断マンモグラフィ」と称し、マンモグ ラフィ検査は「検診」と「診断」を必ず区別 することとする.これは、「検診マンモグラフ ィ」が、2方向もしくは1方向撮影のマンモ グラフィを実施し, 無症状者から乳癌が疑わ れる所見を発見して要精検か否かのマネジメ ントを決定する画像検査であるのに対し, 「診断マンモグラフィ」は、有症状の患者や 乳がん検診で要精検となった受診者に対し て, 2 方向撮影はもちろん, 拡大スポットな どの特殊撮影も実施したうえで、生検実施の 必要性の有無や経過観察といった マネジメントを決定するための診断乳房画像 検査の一つだからである[6].

この撮影方法の違いにより,診断画像は検 診画像と比べ,同一の患者で異なる様式のマ ンモグラフィ画像が複数枚撮影され,より正 確な良悪性鑑別が可能である.

3. 課題と目的

3.1 課題

CAD は第二の意見として,医師の見落としや 診断に要する時間の短縮により,医療事故の減 少や医師の負担の軽減が期待できる.しかし,そ れには前提として CAD の精度を高めて信用で きる手段として確立しなければいけないとい うのが課題である.

3.2 目的

本研究は、検診画像、診断画像の2つのデー タベースを用いて乳がん腫瘤の良悪性鑑別処 理を開発し、判別性能を確認する.その後、検 診と診断(二次精検)の検出性能を比較し、そ の違いは診断画像と検診画像の何が原因なの かを分析することを目的とする.

4. 実験手法

4.1 概要

今回の研究では乳がん腫瘤検診画像と診断画 像の2種類のデータベースで乳がん腫瘤の良悪 性鑑別処理を構築し,正診率を比較する.CAD システム構築はMATLABを用いて行った.

その後,検診画像と診断画像のコントラスト, 粒状性の比較を通じて CAD 性能の違いにどの ような原因があるか画像処理を用いて確認す る.

4.2 詳細な手順

AI 学習に使う画像は良性腫瘤,悪性腫瘤の画像の腫瘤周辺を切り取ったもの各250枚を用意して,以下の手法で実験を行う.N数度は5とする.

AI による乳がん腫瘤の良悪性鑑別処理シス テムは

- 基礎的な機械学習であるパーセプトロン・三層 NN 学習法
- 2. 深層学習である CNN を中間層 1~6 層
- 3. ある領域で学習させたモデルを別の領域

で適応させる転移学習の中の VGG16,VGG19,Alexnetの三種

の11種のネットワークを用いる.

さらに,以下のように画像拡張で検診画像と診 断画像のデータベースをさらに2つに分けて学 習を行った.

- 1. 画像をそのまま学習に使う画像拡張なしの 方法
- 2. 事前に画像の回転,明度,コントラストを調 整した画像を学習データに追加して学習さ せる(画像拡張あり)

検診画像と診断画像の比較手順について以下 に示す.

a)粒状性の測定

- 1. 腫瘤影部分5×5cmを切り出した画像(図 3 左)をフーリエ変換する.
- 2. 変換した画像(図 3 中央)の中央の信号 を0にすることで低周波を除去する.
- この画像(図3右)を逆フーリエ変換し て信号のヒストグラムを確認することで ノイズの量を確認する.

中央の円は直径が画像の縦横の半分である.

この手法の場合の粒状性の良さは上記ヒス トグラムの標準偏差(元画像の高周波領域のば らつき)が小さいほど画像のノイズの量が少な いと言える.



図 3 散乱線の影響

b)コントラスト(散乱線の量の推定)の測定 画像の濃度ヒストグラムを確認し標準偏差 の値の大小でコントラストを確認する.

コントラストは画像濃度ヒストグラムの標 準偏差(元画像の画素のばらつき)が大きい ほど良い値である.

5. 実験結果

5.1 検診画像を対象とした AI 学習の実験結果

検診画像を用いた腫瘤良悪性鑑別処理の実 験結果を以下の表1に示す.

5.2 診断画像を対象とした AI 学習の実験結果

診断画像を使用した実験結果を表2に示す.

最も高い正診率は画像拡張ありの VGG16 を 使用した学習ネットワークであり、専門医の正 診率と言われている 80~90%と同等の正診率 を出せる AI の開発がされた.

検診・診断各画像拡張の結果から1番精度が 高かったものをまとめたものを表3に記載する.

検診画像の最大精度は画像拡張ありの Alexnet で85%であり,診断画像の最大精度は画 像拡張ありのVGG16で90%となった.診断画像 の方が5%高い精度であった.

表1 検診画像を用いた乳がん腫瘤良悪性鑑 別処理の結果

	精	度(%)
	拡張なし	拡張あり
パーセプトロン	54	59
3層ニューラル ネットワーク	53	64
CNN (中間1層)	57	68
CNN(中間2層)	55	69
CNN(中間3層)	60	69
CNN(中間4層)	60	71
CNN(中間5層)	59	71
CNN(中間6層)	60	73
VGG16	63	80
VGG19	64	83
Alexnet	64	85

表 2 診断画像を用いた乳がん腫瘤良悪性鑑 別処理の結果

	精度	(%)
	拡張無し	拡張あり
パーセプトロン	69	60
3層ニューラル	75	68
ネットワーク		
CNN (中間1層)	71	70
CNN (中間2層)	66	65
CNN (中間3層)	73	70
CNN (中間4層)	73	78
CNN (中間5層)	77	82
CNN (中間6層)	68	80
VGG16	73	90
VGG19	69	80
Alexnet	69	83

表3 AI 学習による腫瘤良悪性鑑別処理の結 果の中で精度の高かったもの

検診の	最大精度	診断の:	最大精度
なし	拡張あり	なし	拡張あり
64%	85%	77%	90%
VGG19 Alexnet	Alexnet	CNN (中間5層)	VGG16

5.3 検診画像と診断画像の比較結果

5.3.1 粒状性の比較結果

4.2.a)の手法で検診診断共に 4 例で粒状性の 比較をした.図4に信号レベルとその頻度を記 録したヒストグラムを検診画像と診断画像で 2 例ずつ示す.標準偏差平均は検診画像が 2.7225, 診断画像が 0.5855 であった.

検診画像が診断画像の4倍近い標準偏差とか なり画像のノイズが大きいということがわか る.



図 5 コントラストの検査結果例

5.3.2 コントラスト(散乱線の量)の比較結果

4.2.b)の手法で検診診断共に4例でコントラ ストの比較をした.図5に濃度レベルとその頻 度のヒストグラムを検診画像と診断画像で2例 ずつ示す.標準偏差平均は検診画像が17.485,診 断画像が38.20075であった.

標準偏差の平均は診断画像が検診画像の2 倍近い値となり、検診画像は診断画像よりコントラストが低いことがわかる.すなわち、 検診画像は散乱線の量が多いと推測される.

6. 考察

検診画像と診断画像の AI による良悪性鑑別 処理の精度に関わる原因として,画像を解析し た結果,診断・検診画像のノイズ検査における標 準偏差平均(画像の粒状性の良さ)において検 診画像が診断画像の約 4 倍のノイズ量,コント ラストの良さが 1/2 であり,このノイズの大きさ とコントラストの悪さが学習精度に悪影響を 及ぼしている一因であると考えられる.

これらの差は検診と診断に使われる撮影機 器の違いであり,改善するには機器の進歩,画像 処理の必要がある.しかし,画像処理で粒状性の 改善のためにノイズ除去をするだけではコン トラストの低下による画像のボケが予想され る.

また,不正解画像は石灰化が写りこんでいる 物が多かった.この場合,乳がん石灰化の AI 学 習ネットワークを併用してすれば,より精度の 高い良悪性鑑別処理が行われる可能性がある.

7. まとめと課題

今回の研究では、検診画像と診断画像の相違 点による CAD システムへの影響を正診率とし て数値に示すため、中川らの鋸歯形状領域検出 法[3]のような腫瘤辺縁形状認識による鑑別結 果の根拠を明確に医師に指示することができ るもの、福岡らの手法[4]のような特徴量が可視 化される CAD システムの構築ではなく、井上 謙一の手法[2]のように転移学習による高い性 能を持つ CAD システムの構築を行った.

結果として検診画像において 85%,診断画像

において 90%の正診率と医師の読影に近い値 を導き出すこと,正診率の違いを検診画像と診 断画像の相違の解析を画像の空間周波数フー リエ変換による高調波除去を使った散乱性測 定,粒状性の測定で分析することに成功した.

しかし,診断画像との差を是正するためには, 画像処理や石灰化の AI 学習ネットワークとの 併用が必要であり今後の課題となる.

考察で述べたように,画像処理において粒 状性とコントラストは負の相関にあるため,画 像処理には粒状性とコントラストのバランス をとるようなものを用いる必要がある.

この問題を解決するために散乱線,粒状性の 同時改善が可能な画像処理を行い,それにより 生成された画像で学習を行うことが必要であ る.

その画像処理方法の一つとして富士フィル ムが開発した「Virtual Grid(バーチャルグリッ ド)」が挙げられる. Virtual Grid 技術は,被写体 内での X 線散乱により引き起こされる画像の コントラスト低下と粒状性低下の両方を改善 するため、「コントラスト改善処理」と「粒状改 善処理」から構成される[9]. 今回の研究で明ら かになった課題を解決するため、現在この技術 を参考とした画質改善処理を検討している.

利益相反の有無

利益相反 なし

文 献

- [1] 土井邦雄:医用画像とコンピュータ支援診断 ~現状と将来の可能性~. 映像 情報メディア学会誌,65(4):p.427-431, 2011
- [2] 井上謙一:畳み込みニューラルネット ワークを用いたマンモグラフィの自動 読影システムの構築:第6回「医療情 報学会・人工知能学会 AIM 合同研究 会」: 2018
- [3] 中川俊明,桜井寛之,原武史,他:乳房X
 線写真における腫瘤陰影の自動良悪性
 鑑別システムの開発:生体医工学:日

本エム・イー学会誌 43(3) : 2005 : p 437-446

- [4] 福岡大輔,原武史,福岡広志:乳房X線
 写真における腫瘤良悪性鑑別システムの開発:医用電子と生体工学 39-1:
 2001:p24-29
- [5] 飯沼武:2年間隔マンモグラフィ検診の利益リスク分析:14回日本乳癌検診学会:2004年
- [6] 検診カテゴリーと診断カテゴリーに基 づく乳がん検診精検報告書作成マニュ アル:日本乳癌学会編:金原出版株式会

社:2019

- [7] 桂川茂彦:医用画像情報学:南山堂: 2002:p101-103
- [8] 滝川厚:散乱 X 線が X 線画像に及ぼす
 影響:人間と科学 県立広島大学保健福
 祉学部誌 14 巻(1 号) 2014 p23-35
- [9] 川村隆浩,内藤慧,岡野佳代,他:新画 像処理「Virtual Grid(バーチャルグリッ ド)技術」の開発:X線検査の画質と作 業性の向上:富士フイルム R&D 統括 本部技術戦略部:富士フイルム研究報 告(60): 2015:p21-27

Analysis of the Effect of Image Quality Differences on CAD Performance in AI-Based Classification of Benign and Malignant of Breast Cancer Masses

Soma KUDO^{*1}, Kazuya ABE^{*2}, Hideya TAKEO^{*2}, Yuuichi NAGAI^{*3}, Shigeru NAWANO^{*4}

*1 Graduate School of Electronic & Electrical Engineering, Kanagawa Institute of Technology

*2 Kanagawa Institute of Technology

*3 National Cancer Center Hospital East

*4 Shinmatsudo Central General Hospital

Abstract

In recent years, the amount of medical images generated in the medical field has been increasing due to the high resolution of internal imaging equipment, and the burden on doctors who read these images has also been increasing rapidly. In order to solve this problem, it is necessary to improve the function of computer-aided diagnosis (CAD) as a diagnostic aid to solve this problem, we developed an AI system for discriminating between benign and malignant breast cancer tumors in breast cancer screening by using transfer learning, which is one of the AI deep learning methods, and image enhancement, which is trained by using images whose rotation, brightness, and contrast have been adjusted in advance. We analyzed what factors are necessary to improve the accuracy of diagnostic imaging. The system for classification of benign and malignant diseases using the medical checkup images was 85%, which was equivalent to the physician's discrimination, but it was not as good as 90% for the diagnostic images, and the results were found to be caused by noise and poor contrast through image comparison. We analyzed that these improvements are necessary for the construction of a more accurate CAD system in medical checkups.

Key words: Breast cancer, Screening, Diagnosis, CAD, Benign and Malignant classification

マルチチャンネル化と M-Net を用いた嚥下時 X 線動画から

の頸椎椎間板の抽出と評価実験

郡司 絵莉華*1 目片 幸二郎*2 滝沢 穂高*1 工藤 博幸*1

要旨

嚥下障害は栄養低下や誤嚥を引き起こすことがあり、患者の生活の質を低下させるが、その詳細なメカニ ズムは未だに解明されていない、本研究では、嚥下時 X 線動画(VF)にマルチチャンネル化(Multichannelization, MC)、特徴画像選択と M-Net を適用し、頸椎椎間板を抽出する手法を提案する、濃淡画像で ある X 線動画の各フレームにモルフォロジー演算と Random sample consensus による放物線あてはめ処理な どを適用し、頸部領域を抽出する. 抽出した頸部領域について線形フィルタや非線形フィルタを用いて特 徴画像を生成し、マルチチャンネル画像(MC 画像)を生成する. この MC 画像を M-Net に入力し、抽出精度 を画素値単位の F 値を用いて評価する. 焼きなまし法を用いて抽出精度が最大となる特徴画像の組み合わ せを求める. 本論文では、MC 画像による領域抽出を評価するための 3-fold 3-phase cross-validation という実 験方法を用いる. MC を用いなかった際の F 値は 0.730、用いた際の F 値は 0.800 となり、精度向上を確認 した.

キーワード:嚥下時X線動画,頸椎椎間板,マルチチャンネル化,M-Net,焼きなまし法

1. はじめに

嚥下障害は食べ物を飲み込むことが困難に なることによる栄養低下や,むせることにより 誤嚥性肺炎を引き起こすことがある[1][2].特に 誤嚥性肺炎は,日本において高齢化に伴い増加 する疾病の一つであり,2019年の厚生労働省の 調査によると,死因別死者数の第6位である[3]. 嚥下障害の原因としては,脳血管障害等の疾病 や頸部手術の副作用などであることが知られ ているが[4][5].そのメカニズムは未だに解明さ れていない.そこで,我々は頸椎構造物と嚥下 障害との関係を調査するために,頸椎椎間板に 着目し,頸椎椎間板の抽出を試みている.本報 告では嚥下時 X 線動画において,マルチチャ ンネル化(Multi-channelization, MC)[6][7]と特徴

*1 筑波大学

〔〒305-8573 つくば市天王台 1-1-1〕
e-mail: gunji.erika@pr.cs.tsukuba.ac.jp
*2 四条畷学園大学
投稿受付: 2021年06月18日

選択,及び M-Net を用いて頸椎椎間板の抽出を 行う.



2. 方法

1) 頸部領域の抽出

先行研究に基づき[8], モルフォロジー演算



図 2. マルチチャンネル化処理,特徴選択, M-Net による頸椎椎間板の抽出

と Random sample consensus (RANSAC)による 放物線あてはめ処理などを適用して首中心点 を特定する. 図1の(a)の緑色は首中心線を示 し,青色の放物線は頭部,首部,肩部の境界 を示す.これらの線の交点を上から順に点 p_{head} , $p_{shoulder}$ とし,この2点の中心点を首中 心点 p_{center} とする(同図(b)を参照). p_{center} を中 心に260 × 320ピクセルの関心領域(Region of interest, ROI)を頸部領域として抽出後(同図 (c)を参照),256 × 256ピクセルに正規化す る.

2) 頸椎椎間板のセグメンテーション

図2に提案手法のフローチャートを示す. VFの各フレームは8ビット濃淡画像である. この濃淡画像からM種類の特徴画像を作成 し,その中から3種類の特徴画像を選択し, それぞれをRGBチャンネルに入力したカラー 画像を生成する.これをマルチチャンネル画 像(MC画像)[6][7]と呼ぶ.MC画像を図3に示 すM-Net[9]に入力し,椎間板の候補領域を抽 出した予測画像を出力し,判別分析法で二値 化し,椎間板領域を抽出し,画素単位のF値 で精度を評価する.得られた精度をもとに, 焼きなまし法による特徴画像の組み合わせの 最適化を行う.本研究では F 値が最も良かっ た特徴画像を RGB チャンネルに入力した MC 画像を初期状態とし,エネルギー関数を1– F値,冷却スケジュールを $T' = T \times 0.9$ として最 適化を行った.



3. 実験

本研究では,嚥下時X線動画 58 症例からな るデータセットを実験に用いる.各症例から舌 骨最大挙上フレームの椎間板の位置を特定し, 椎間板を手動抽出し,正解データとする.本手 法では, M-Net の学習, 特徴画像の評価, M-Net による予測の評価について, それぞれ画像デー タが必要になる. そこで藤中らが提案した 3fold 3-phase cross validation [7]による評価を行う. この評価手法では、データセットを3つのサブ セットに分割し、学習、特徴評価、予測の3つ のプロセスにそれぞれ適用する. 3つのサブセ ットの順番を変えて3回繰り返す.なお,各サ ブセットで正常症例数と異常症例数の比率を 揃える. M-Net に用いるパラメータは、学習の エポック数 100, バッチサイズ を 1, 最適化手 法には Adam, Base learning rate を 0.001, 損失 関数にはバイナリクロスエントロピーである. なお,本研究は神戸赤十字病院の倫理委員会の 承認を受けて行われた.

4. 考察とまとめ

表1にMCを適用したときの平均F値と適用 しなかったときの平均F値を示す.MCを適用 したときの方が,有意に抽出精度が高かった (p<0.05).

$\overline{\alpha}$ I: 計価症未(MC: マルケケア ヤノ ホル f	表 1:	評価結果(MC:	マルチチャンネル化)
---	------	----------	-----------	---

MC	Precision	Recall	F 値
あり	0.837	0.790	0.800
なし	0.803	0.700	0.730

図4と図5はマルチチャンネル化を適用し たことによって精度が向上した症例と低下し た症例を示す.図中で青色の画素は未検出画素, 赤色の画素は過剰検出画素,緑色の画素は正し く抽出した画素を表す.

図4に示した症例では、マルチチャンネル化 を行わなかった場合(現画像をそのまま M-Net に入力した場合)に検出されなかった椎間板が、 マルチチャンネル化によって検出できるよう になり、抽出精度が向上していることが分かる. 一方、図5の症例では、マルチチャンネル化に より、検出されなくなる部分があるが、その割 合は小さいことが分かる.今回実験に用いたデ ータセットでは、マルチチャンネル化により、 精度が低下する場合があるものの,精度が向上 する場合に比べて小さく,全体としてはマルチ チャンネル化を用いることが有効であること が分かった.

今回はマルチチャンネル化におけるチャン ネル数を 3 としたが, 今後の課題として4チ ャンネル以上に増やして実験することがあげ られる.

MC なし	MC あり	患者 種類
		異常
F 値 = 0	F 値=0.693	
(I)II	lilli,	正常
F 値=0.241	F 値=0.862	

図4: 精度が向上した結果



図5:精度が低下した結果

謝辞

研究遂行にあたり貴重なご協力を賜った 京都大学大学院医学研究科人間健康科学系専 攻リハビリテーション科学コース 松林潤氏, 岡山大学整形外科 瀧川朋亨氏,神戸赤十字病 院整形外科 戸田一潔氏,神戸赤十字病院整形 外科 伊藤康夫氏に深謝いたします.

利益相反の有無

なし

文 献

[1] 日本気管食道科学会「嚥下障害」, http://www.kishoku.gr.jp/public/disease04.html

[2] World Gastroenterology Organisation,

"Dysphagia"

http://www.worldgastroenterology.org/guidelines/gl obal-guidelines/dysphagia

 [3] 厚生労働省,「令和元年(2019年)人口動態統計月報年計(概数)の概況」,

 約

 かの概況」,

 https://www.mhlw.go.jp/toukei/saikin/hw/jinkou/ge

ppo/nengai19/dl/gaikyouR1.pdf

[4] 平山三智子,西川光一「頚椎後方手術後に 脳神経麻痺による嚥下障害を

生じた 3 症例」日本臨床麻酔学会誌, Vol.32, No.7, pp.943-947, 2012.

[5] 林寿光ほか「頸椎手術後に悪化した嚥下障 害例に対する手術治療経験」

日本嚥下医学会, Vol.54, No.2, pp.152-156, 2008.

[6] Fujinaka Ayano, Takizawa Hotaka, Mekata Kojiro, Kudo Hiroyuki, "Segmentation of cervical intervertebral disks from videofluorography based on semantic segmentation, multi-channelization,

and feature selection: comparison of U-Net, FPN, LinkNet, and PSPNet" The 16th Asian Conference on Computer Aided Surgery, 2020.

[7] 藤中彩乃「深層学習とマルチチャンネル化 を用いた嚥下時 X 線動画からの頸椎椎間板領 域の形態解析に関する研究」筑波大学大学院シ ステム情報研究科コンピュータサイエンス専 攻博士論文, 2021

[8] 郡司 絵莉華,藤中 彩乃,目片 幸二郎, 滝 沢 穂高,工藤 博幸,「U-Net を用いた X 線動 画からの頸椎椎間板の抽出に関する基礎的検 討」,第39回日本医用画像工学会(JAMIT)大 会/2020-09-17--2020-09-17

[9] Kang Cheol Kim, Hyun Cheol Cho, Tae Jun Jang, Jong Mun Choi, and Jin Keun Seo "Automatic detection and segmentation of lumbar vertebra from X-ray images for compression fracture evaluation" arXivs-prints, Apr 2019.

Extraction of cervical intervertebral discs from videofluorography during swallowing using multi-channelization and M-Net and its

performance evaluation

Erika Gunji*1, Kojiro Mekata*2, Hotaka Takizawa*1, Hiroyuki Kudo*1 *1 University of Tsukuba *2 Shijonawate Gakuen University

Dysphagia may make it difficult for patients to swallow food and drink, and therefore it will lower their quality of life. The mechanism of dysphagia has not been elucidated yet. In this report, cervical intervertebral discs are extracted from videofluorography (VF) by use of our multi-channelization (MC) technique, our image feature selection technique, and M-Net. First, a neck region is extracted based on morphological operations and RANSAC-based parabola fitting. In the MC technique, M feature images are generated from one VF frame image by using linear and non-linear image filters. Three feature images are selected to generate one color image, which is called a MC image, and the MC image is input into M-Net. The combination of the three feature images is optimized by the simulated annealing technique. In this work, a performance evaluation method, called 3-fold 3-phase cross-validation, is used to evaluate the extraction accuracy obtained by the MC technique. The pixel-based F-measure of the proposed method is 0.800, whereas that of the conventional M-Net is 0.730.

Key words: Videofluorography, Cervical intervertebral disks, Multi-channelization, M-Net, Simulated annealing

EfficientNet を用いたアンサンブル学習に基づく

放射線皮膚炎グレード判定手法の開発

和田 清隆*1.2 渡邊 睦*2 新野 将史*2 野口 康介*2 荻野 尚*1

要旨

放射線皮膚炎の評価は、臨床で用いる有害事象共通用語規準(Common Terminology Criteria for Adverse Events: CTCAE)に基づいた軽度のグレード1から有害事象による死亡の5まで5段階に分類される.しかし、視覚的な評価に基づいているため個人の経験や知識に左右される課題がある。著者らは、学習に用いる症例数が少ない課題に対処するため、人工症例画像を生成し、深層学習を用いた放射線皮膚炎のグレード判定システムを構築してきた.本稿では、近年提案された EfficientNet モデルを用いたモデルを作成した.評価者のグレード判定に相違が生じた画像についてアンサンブル学習に基づく最終的な分類を行った手法について述べる. EfficientNet-B0~B7を用いて、画像解像度とデータ拡張法の条件を変えた学習モデルは、86.4%の正答率であった.また、グレード判定に相違が生じた画像について、EfficientNet モデルでアンサンブル学習を行った.ベイズの定理に基づく最大事後確率推定法(ベイズ推定)を用いた評価実験により、提案手法の有効性を確認した.

キーワード: 放射線皮膚炎, ディープラーニング, EfficientNet, アンサンブル学習, ベイズの定理

1. はじめに

放射線皮膚炎とは,放射線治療における放 射線が皮膚を通過することに起因する放射線 治療の有害事象である[1].放射線皮膚炎は, 放射線治療における一般的な有害事象であり, 放射線治療を受ける患者の 90%以上に発症 すると報告されている[2,3].多くの臨床では 医師,看護師が有害事象共通用語規準

(Common Terminology Criteria for Adverse Events: CTCAE) [4]に基づくグレードの判定 を行い,皮膚ケアを行っている.表1に示す ように,その程度は軽度のグレード1から有 害事象による死亡の5まで分類している.放射線皮膚炎のグレードを正しく評価し,症状 に応じた一貫性のあるケアが放射線皮膚炎の

*1 メディポリス国際陽子線治療センター 〔〒891-0304 鹿児島県指宿市東方 4423〕 e-mail: wada-kiyotaka@medipolis.org *2 鹿児島大学大学院理工学研究科

表1 有害事象共通用語規準 ver.4

Grade1	Grade2	Grade3	Grade4	Grade5
わずかな紅斑 や乾性落屑	 中等度和 年等の 年度の ただら ただ たど にした たど にした にした	皺や襞以外の 部位の湿性落 屑;軽度の外 傷や摩擦によ り出血する	生命を脅かす; 皮膚全層の壊 死や潰瘍;病 変部より自然 に出血する	死亡

低減と患者の Quality of life (QOL) 改善に繋 がる.しかし, CTCAE に基づく放射線皮膚 炎のグレード判定は, 評価者自身の基準の解 釈と評価経験に左右されやすく, 客観的評価 として難しい問題がある.これまで, 放射線 皮膚炎画像データベースを利用した DCNN に関する研究は, 我々の知る限り報告されて いない.そこで, 著者らは放射線皮膚炎のグ レード判定を統一した基準で行うために, デ ィープラーニング (Deep convolutional neural network: DCNN)を用いた放射線皮膚炎のグ レード判定を補助するシステム (Hybrid generation method Radiation dermatitis grading support system: Hyb-RDGS) を開発した[5]. Hyb-RDGS は、学習済の VGG16 モデル[6]の
 一部を再利用する Fine-tuning を行い、新たな
 学習モデルを作成したものである.

DCNN を用いた放射線皮膚炎のグレード判 定で課題となったのは、図1に示すように判 定基準が曖昧なため評価者の判定に相違が生 じる複数 (グレード) 判定画像があること, また、グレード高い症例が少ないことによる データの不均衡であった. Hyb-RDGS におい ては、これらの課題に対処するため、まず評 価者である放射線治療医、放射線治療に習熟 した看護師によってグレード分けを行った. 画質が不鮮明な画像や判定結果が異なる画像 を除外し, 整合性のとれた学習データを選別 した. グレードの高い症例数が少ないことに 対しては、ポアソン合成手法を用いた人工症 例画像を生成した. DCNN の不足する学習画 像を補う手法として,元の学習画像に画像処 理を加えてデータ量を増やす水増し処理

(data augmentation: DA)を行うとともに、人工症例画像と混合した(ハイブリッド生成法)
 学習データを作成した.この成果より著者らは、Hyb-RDGSのグレード判定の正答率が85.1%を達成したことを報告した[5].

本稿では,近年 Mingxing らによって提案された Efficientnet を用いて新たな判定手法を開発し,精度と学習効率を向上した結果について述べる[7]. さらに評価者のグレード判定に相違が生じた複数(グレード)判定について, EfficientNet を用いたアンサンブル学習に基づく判定手法を提案する.



図 1 DCNN を用いた放射線皮膚炎グレード判 定の課題

2. 提案手法

図2に提案手法の概要を示す.



- (1)はじめに放射線皮膚炎画像を正解画像と 複数(グレード)判定画像に分類した.
- (2) EfficientNet モデルのスケーリング B0~
 B7 と解像度の条件を変えたモデル構成
 を評価し、グレード判定精度の高い
 EfficientNet モデルを複数選定した.
- (3)データ拡張の条件を変えた複数のデータ セットを作成し、(2) で選定したモデル に用いた.モデル構成を評価し、グレー ド判定精度の高い EfficientNet モデルを 選定した.
- (4) (3) で選定したモデルを使用してアンサンブル学習を行い、複数(グレード)判定画像に対するグレード判定を行った.

1) EfficientNet モデルの構成

近年,高精度な画像認識を達成するため, 学習の過程で構造自体を自動的に探索する Neural Architecture Search (NAS) 🕆 compound scaling method (複合スケーリング法)を用い てスケーリングを徐々に大きくし, 効率的な 構造を探索する EfficientNet が登場した. 従来 の Convolutional neural network (CNN) のスケ ーリングは, CNN の主なパラメータである深 さや広さを任意に増大させるか、大きな解像 度を使用することであった.手動チューニン グを必要とし,最適な値とはいえない場合も あった. また, モデルのパラメータ増加に対 して精度の向上に限界があった. これに対し て EfficientNet は, ネットワークの構造を変え ずに深さと広さと解像度の比率を固定してス ケーリングアップしていく. そのため, 他の パラメータ数をあまり増やすことなく精度を あげることができる.パラメータと計算量が

他の DCNN よりも小さいため速度も向上し ている. EfficientNet は、比率を固定してスケ ーリングしているので, そのスケールする値 によって B0~B7 まである. しかし, スケー リングアップによって精度はあがるもののモ デルが大きくなるとその恩恵が受けにくいと されている[7]. モデルスケーリングについて Sheeela らは、道路状況カメラ画像の分類に EfficientNet-B0, B4 を使用したと報告してい る [8]. EfficientNet モデルが分類に有効であ ったとされているが、リソースの問題から他 のモデルのスケーリングは課題とされている. EfficientNet モデルのスケーリングの違いに ついて研究された報告は十分でない.近年で は、スマホなどの小型端末にも乗せられる高 性能 CNN として MobileNet が開発されている [9-11]. Srinivasu らは, 7 つの皮膚疾患分類に MobileNet V2 & Long Short Term Memory (LSTM) をベースとした moblie-size の DCNN を開発したと報告した. Srinivasu らの 提案モデルの目標解像度は 224×224 画素と されている[12]. 本稿では, Mobile-size の DCNN で報告されている解像度を参考に解像 度を設定した.はじめに解像度の条件を変え た学習用データを複数作成し, 放射線皮膚炎 画像のグレード判定を行う学習モデルの構成

2) グレード判定モデルの作成

について検討した.

著者らは、元の画像に変換を加えて DA を 行うとともに、新たな症例を人工的に生成す る人工症例画像(Poisson image editing: PID) も症例画像として利用した Hyb-RDGS を作成 している[5].しかし、DA は画像データにど のような操作が有効なのか試行に時間を要し てしまい、人工症例画像を生成する作業を加 えなければならない.そこで、本稿では、Ekin らによって自動的に DA を選択してくれる手 法として提案された Rand Augment (RA)を 用いたデータセットを新たに作成した.これ までデータ拡張の最適値を探す手法では、 Auto Augment などの手法があったが、最適な 拡張となるパラメータを探索するコストが非 常に大きいという欠点があった[13]. RA は, それを10の2乗オーダーで最適なデータ拡張 を見つけ出すことを実現している[14].

本稿では, RA, DA および人工症例画像で 作成したデータセットを用いて EfficientNet モデルを作成した. それぞれのモデルのグレ ード判定の正答率を評価し, アンサンブル学 習を行うモデルを探索した.

3) アンサンブル学習

Esteva らは、3 種類の皮膚腫瘍の良悪性を 識別する分類器を作成したと報告している [15].また,Fujisawa らは、14 種類の皮膚腫 瘍の良悪性を識別する分類器を作成したと報 告されている[16].しかし、放射炎皮膚炎の グレード判定は、これらの皮膚腫瘍診断が病 理検査を基にした確定診断であるのに対し、 評価基準が表1に示した CTCAE の記述だけ であり客観的な評価として難しい問題がある. そのため、評価者間のグレード判定の相違が 生じることがある.本稿では、この相違が生 じたグレード判定について EfficientNet を用 いた判定モデルを提案し、複数(グレード) 判定画像に対する評価方法としてアンサンブ ル学習を検討した.

アンサンブル学習は,複数の分類器の結果 を統合することで一つの分類器よりも予測精 度の向上及び汎化性能に有効な手法として用 いられてきた[17,18]. CNN を用いた画像認識 においてもアンサンブル学習は有効であるこ とが分かっている[19].また,統計学で提案 されたベイズ定理や所属確率を用いた予測法 も研究されている[20].

本稿では、アンサンブルの方法として、複数の EfficientNet モデルのグレード判定にベイズ定理に基づく最大事後確率推定法(ベイズ推定)を適用して予測する手法を提案する [21].

ベイズ推定

ベイズ推定は、ベイズの定理と組み合わせ

て確率的推論を行う統計的手法である. ベイ ズの定理は,尤度関数が理論的に設定できな い場合にも適用でき,実用性が高いため,臨 床医学の診断推論に活かせる定理として用い られている[22,23].

放射線皮膚炎画像を判定するモデルでは, 画像がある感度/特異度(尤度比)をもつグレ ードのパターン(x),各グレードのクラスの正 答率をクラス(y)とする. p(y|x)はグレードの パターン(x)を推定した後で,クラス(y)の判 定確率,p(y)はパターン(x)と推定する前のク ラス(y)の判定確率を表す.それぞれ事後確率, 事前確率とよぶ.

ここで事前確率と尤度比が分かれば事後確 率が求められる.事後確率p(y|x)が最大とな るクラスyにパターンxを分類すれば,認識誤 差が最適なパターンになる.ベイズの定理に よると

$$p(y|x) = \frac{p(x|y)p(y)}{p(x)} \propto p(x|y)p(y) \quad (1)$$

ここで, *p*(*y*|*x*)の(*y*)に関する最大化を *p*(*x*|*y*)*p*(*y*)の最大化に変換できる.ニューラ ルネットワークでは,事後最大化解は,交差 エントロピーの損失最小化(ベイズ推定量) で解くことができる.アンサンブル学習の各 モデルの事後確率を用いてベイズ推定量を算 出し,複数(グレード)判定画像に対する最 終グレードを判定する.複数(グレード)判 定画像に対するベイズ推定量は,2 つの確率 分布 *p*,*q*に対して(2)式で与えられる.

$$H(p,q) = -\sum_{x} p(x) \log q(x)$$
 (2)

本稿のグレード判定について, p を「真の 分布」, q を「予測の分布」とすると, EfficientNet モデルによる予測が正解に近い ほどpとqの誤差が小さくなると考えること ができる.図3に最終判定の導出過程を示す. ここでは, 3 つのモデルが, それぞれ正解で あると判定したグレードを出力する. ベイズ の定理より事後確率が求められ, グレード判 定の予測の分布 q となる. また, 評価者のグ レード判定が1または, 2 であった複数(グ レード)判定画像を真の確率分布 p とする. これより,(2)式を用いてベイズ推定量を算 出する.各モデルの真のグレード判定におけ るベイズ推定量が最も小さいモデルの判定を 最終グレード判定とする.



図3 最終判定の導出過程

3. 実験

1) 実験データ

実験では、当施設で陽子線治療を受けた患者の放射線皮膚炎画像を放射線腫瘍医2名と 放射線治療に習熟した看護師4名が、 CTCAEversion4をもとにグレード分けした上 でデータベースから抽出した.放射線皮膚炎 画像649枚(グレード1/2/3/4:318/95/205/31 枚)をEfficientNet モデルの作成に使用した.

次に放射線皮膚炎評価の複数(グレード) 判定画像 60 枚(グレード1または、2 と判定 された画像 30 枚、グレード2 または、3 と判 定された画像 30 枚)を抽出し、アンサンブル 学習のテストデータとして使用した.

2) データセット

EfficientNet モデルの構成の検証には,解像 度を256×256,384×384 画素および512×512 画素に変換させた画像を学習データとして使 用した. DCNN の学習は,バッチサイズを8, エポック数50回とした.学習データは,水増 し処理(Rand Augmentationのみ)を行った臨 床放射線皮膚炎画像650枚,テストデータ110 枚とした.解像度の条件を変えた学習データ とEfficientNet モデルのそれぞれ B0~B7 との 構成について精度を比較し、最適なモデル構 成を検討した.

次に選定したモデルのデータセットを検討 するため、水増し処理と人工症例画像を生成 し、データ拡張法の条件を変えたデータセッ トを4通り作成した.データセットの内訳を 表2に示す.人工症例画像は、症例数が少な いグレード4を対象とし、放射線皮膚炎画像 を正常皮膚画像にポアソン合成手法で作成し た[5].

表 2 データセット内訳

	データ拡張法	学習データ (枚)	テストデータ (枚)
А	Rand Augmentation	1401	110
В	Data Augmentation	1133	110
С	Rand Augmentation (PIE*)	1445 (195)	110
D	Rand Augmentation (Rand Augmentation+PIE)	1626 (183+195)	110
		*DIE· D/	visconn image editing

3) 評価方法

解像度の条件を変えた EfficientNet モデル 構成の検証は、それぞれの解像度について 3 分割交差検証を行い、test accuracy の高いモデ ルを選定した.

選定したモデルと表2に示したそれぞれの データセットについて、3分割交差検証を行った.それぞれのモデルの sensitivity (感度), precision (適合率), F1-measure (感度と適合 率の調和平均)を以下の算出式から求め、モ デルの性能を評価した.ここで、i は対象の

グレードである.
sensitivity_i =
$$\frac{TruePositive_i}{TruePositive_i + FalseNegative_i}$$
 (3)

$$precision_{i} = \frac{TruePositive_{i}}{TruePositive_{i} + FalsePositive_{i}}$$
(4)

F1 measure
$$_{i} = 2(\frac{sensitivity_{i} \times precision_{i}}{sensitivity_{i} + precision_{i}})$$
 (5)

複数(グレード)判定については,正答率 の高い3個のモデルでアンサンブル学習した 結果を用いた.それぞれのモデルで判定され た結果からベイズ推定により最終グレード判 定を行った.提案手法の最終グレード判定に ついては,評価者と一致したグレード判定さ れた割合とした.例えば,評価者の複数(グ レード) 判定画像がグレード1または,2の 場合にアンサンブル学習による最終グレード 判定が1または,2と判定した割合である.

4. 結果

1) モデルの構成比較

EfficientNet-B0~B7 と解像度の構成比較を 図 4 (a), B7 の学習曲線および B0~B7 の 学習時間を図 4 (b) に示す.スケーリングが 小さい B0~B3,および最も大きい B7 では, 解像度によってばらつきがみられた.B4~B7 では,256×256,512×512 画素において test accuracy の向上がみられた.



(a) 各モデルと解像度の構成における精度



(b) B7 の学習曲線と B0~B7 の学習時間 図 4 EfficientNet-B0~B7 と解像度の構成比較

 2) EffcientNet を用いたグレード判定モデルの 作成

EffcientNet モデルと解像度の構成比較から test accuracy の高い以下の3つのモデル構成 (test accuracy の平均) を選択した.

- ・EfficientNet-B1, 256 画素 (80.6%)
- ・EfficientNet-B4, 384 画素 (80.9%)
- EfficientNet-B7, 512 画素 (84.2%)

それぞれのモデルにおいて表2に示した A ~ D のデータセットを使用し, グレード判定 モデルを作成した. 表3に各モデルの全体の 正答率を示す.

表3 グレード判定モデルの正答率

データ	B1	B4	B7	
セット	256×256	384×384	512×512	
А	77.3±2.7%	75.2±6.8%	82.1±1.9%	
В	77.0±7.6%	86.4±4.2%	80.9±3.1%	
С	63.6±6.9%	59.4±2.1%	75.5±5.1%	
D	64.5±7.2%	68.2±6.9%	76.4±9.2%	

次にそれぞれのモデル性能について,対象 グレードの *sensitivity*, *precision*, *F-measure* を求めた. 表 4 にモデルの性能評価結果の平 均値を示す.

			B 256	×256			B4 384×384		B7 512×512				
Datasets		グレード1	グレード2	グレード3	グレード4	グレード1	グレード2	グレード3	グレード4	グレード1	グレード2	グレード3	グレード4
	sensitivity	0.888	0.674	0.753	0.783	0.637	0.831	0.711	0.852	0.744	0.878	0.800	0.883
Α	precision	0.669	0.795	0.779	0.979	0.784	0.607	0.831	0.897	0.882	0.725	0.867	0.855
	F-measure	0.763	0.729	0.766	0.870	0.703	0.701	0.766	0.874	0.807	0.794	0.832	0.869
	sensitivity	0.913	0.630	0.737	0.667	0.911	0.867	0.844	0.817	0.978	0.644	0.833	0.767
В	precision	0.613	0.795	0.795	0.955	0.882	0.830	0.826	0.961	0.677	0.879	0.862	0.979
	F-measure	0.734	0.703	0.765	0.785	0.896	0.848	0.835	0.883	0.800	0.744	0.847	0.860
	sensitivity	0.916	0.767	0.556	0.250	0.578	0.711	0.675	0.480	0.622	0.854	0.744	0.833
С	precision	0.585	0.793	0.667	0.484	0.559	0.736	0.767	0.400	0.862	0.717	0.838	0.641
	F-measure	0.714	0.780	0.606	0.330	0.568	0.723	0.718	0.436	0.723	0.779	0.788	0.725
	sensitivity	0.878	0.500	0.719	0.417	0.844	0.711	0.667	0.417	0.889	0.828	0.615	0.650
D	precision	0.585	0.726	0.653	0.735	0.724	0.674	0.625	0.735	0.800	0.700	0.757	0.780
	F-measure	0.702	0.592	0.684	0.532	0.779	0.692	0.645	0.532	0.842	0.759	0.679	0.709

表4 各モデルの性能評価結果

3) アンサンブル学習

各モデルの性能評価から, *F1-measure* の高 い以下の3個のモデルを選定した.

・EfficientNet-B1, 256 画素 (データセット A)

•EfficientNet-B4, 384 画素 (データセット B)

・EfficientNet-B7, 512 画素 (データセット A)

次にベイズ推定により各グレードの事後確 率を算出した.事後確率を基に真のグレード におけるベイズ推定量を算出し,その値が最 も小さいモデルの判定結果を最終判定とした. 表5に(a)事後確率と(b)各モデルのベイ ズ推定量を示す.また,評価者がグレード1 または,2と判定した複数(グレード)判定 画像に対する最終判定の結果を表6に示す. 最終判定が評価者の判定グレード1または,2

と一致した割合は96.7%,2または,3と一致 した割合は93.3%であった.

表5 事後確率とベイズ推定量

(a) 各モデルの事後確率

推定グレード (事後確率q(x))	B1 256 A	B4 384 B	B7 512 A
グレード1	0.9892	0.9942	0.9846
グレ ー ド2	0.9293	0.9712	0.9675
グレード3	0.9549	0.9740	0.9675
グレード4	0.9833	0.9871	0.9898

(b) 各モデルのベイズ推定量

真のグレード (真の確率p(x))	B1 256 A	B4 384 B	B7 512 A
グレード1 (1,0,0,0)	0.0047	0.0183	0.0106
グレード2 (0,1,0,0)	0.0318	0.0127	0.0143
グレード3 (0,0,1,0)	0.0200	0.0114	0.0144
グレード4 (0,0,0,1)	0.0073	0.0056	0.0044

表6複数(グレード)判定画像の最終グレー ド判定結果(グレード1または,2)

	アンサンブル学習		ベイズ推定量				
EfficienNet	B1	B4	B7	B1	B4	B7	
~	256	384 D	512	256	384 D	512	最終判定
アータセット	A 256	Б 294	A 512	A 256	Б 294	A 512	
	230	304	312	230	0.0102	0.0142	1
No.1	1	1	2	0.0047	0.0183	0.0143	1
No.2	1	I	1	0.0047	0.0183	0.0106	1
No.3	2	1	2	0.0318	0.0183	0.0143	1
No.4	2	2	2	0.0318	0.0127	0.0143	2
No.5	1	1	2	0.0047	0.0183	0.0143	1
No.6	2	1	1	0.0318	0.0183	0.0106	1
No.7	1	1	3	0.0047	0.0183	0.0144	1
No.8	1	1	1	0.0047	0.0183	0.0106	1
No.9	3	1	2	0.0200	0.0183	0.0143	2
No.10	2	1	2	0.0318	0.0183	0.0143	2
No.11	1	1	1	0.0047	0.0183	0.0106	1
No.12	1	1	3	0.0047	0.0183	0.0144	1
No.13	2	3	2	0.0318	0.0114	0.0143	3
No.14	1	2	3	0.0047	0.0127	0.0144	1
No.15	1	1	2	0.0047	0.0183	0.0143	1
No.16	2	1	2	0.0318	0.0183	0.0143	2
No.17	2	2	1	0.0318	0.0127	0.0106	1
No 18		2	2	0.0047	0.0127	0.0143	1
No 19	1	-	2	0.0047	0.0183	0.0143	1
No.19	2	2	2	0.0219	0.0105	0.0144	2
No.20	2	2	2	0.0318	0.0127	0.0144	2
No.21	2	2	2	0.0318	0.0127	0.0143	2
N0.22	1	2	2	0.0047	0.0127	0.0145	1
No.23	1	I	2	0.0047	0.0183	0.0143	1
No.24	1	2	2	0.0047	0.0127	0.0143	1
No.25	3	1	2	0.0200	0.0183	0.0143	2
No.26	1	2	1	0.0047	0.0127	0.0106	1
No.27	2	1	2	0.0318	0.0183	0.0143	2
No.28	1	1	2	0.0047	0.0183	0.0143	1
No.29	1	2	2	0.0047	0.0127	0.0143	1
No.30	2	2	2	0.0318	0.0127	0.0143	2
評価者のグレ	- ドギ	一定と一	・致した	割合(グレ	ノード1また	は,2)	96.7%

5. 考察

1) EfficientNet モデルの構成について

図4(a)に示したように,モデルのスケー リングが大きくなるに従い test accuracy が向 上した.しかし,各モデルの構成を比較する とモデルのスケーリングが低い B0~B3 では, 解像度によって test accuracy にばらつきがみ られた.特に B1は,画素数によって 65~81% となり test accuracy に 16%の差が生じた.こ れに対して B4~B7は,76~84%でありその 差は 8%であった. EfficientNet モデルは,ス ケーリングが低いほど解像度の影響を受けや すいと考えられた.一方,スケーリングが最 も高い B7 では,384 画素の精度が低下した. 本検証は,解像度を一定にした前評価でテス トデータにおける accuracy の低下,loss の上 昇が起きていないエポック数を選択している. EfficientNet の複合スケーリングは,解像度だ けでなくネットワークの深さ,幅を固定され た比率でスケールアップする[7].図4(b)の 学習曲線で accuracy の低下傾向が見られたこ とから,解像度のみ変化させた場合でも最も スケールアップされた B7 のモデルは過学習 を起こしている可能性も考えられる.

これまで DCNN では, 層を深く, 幅を広く, 解像度を上げること、どれか一つの要素を増 やすことで精度を向上させてきた. ネットワ ークの精度は重要な要素だが、 パラメータ数 が多くなりハードウェアのメモリに限界があ った. 前述したように, Srinivasu らの moblie-size の DCNN 研究報告の目標解像度 は 224×224 画素とされている[12]. 本稿の EfficientNet モデルの解像度は 256~512 画素 を設定した.表3、表4に示したように必ず しも高解像度の画像の精度が高いというわけ ではなく、解像度とスケーリングの構成評価 を行うことで、より汎化性能、精度の高いグ レード判定モデルが作成できると考えられる. また、解像度の違いは学習時間への影響は及 ぼさないが, モデルスケーリングが大きくな ると学習時間の延長に繋がる.

2) グレード判定モデルの作成について

表3に示したように本稿で作成した最も正 答率の高い EfficientNet のモデルは、データセ ットに384 画素のDAを用いた EfficienNet-B4 モデルであった. EfficientNet-B4 のグレード 判定モデルの正答率は86.4%であり、著者ら が開発した Hyb-RDGS の正答率85.1%に対し て1.3%上回る正答率を達成した.表7に Hyb-RDG と本稿で作成したグレード判定モ デルと性能比較を示す.

表7 先行研究と本稿で作成したモデルの性 能比較

	Hyb-RDGS			本稿で作成したモデル			゠デル	
モデル構成	VGG16モデルを Fine-tuningしたDCNN			EfficientNet-B4				
データセット	人工症例画像とDAの混合				D	A		
正答率	85.1%				86.	4%		
グレード	1	2	3	4	1	2	3	4
sensitivity	0.911	0.867	0.844	0.817	0.900	0.778	0.833	0.933
precision	0.882	0.830	0.826	0.961	0.794	0.898	0.882	0.847
F-measure	0.896	0.848	0.835	0.883	0.844	0.833	0.856	0.885

課題であった症例数の少ないグレード4の 性能は、Hyb-RDGS に対して EfficientNet モデ ルは precision が 1.13% (0.961/0.847) 低下し たものの sensitivity、F-measure はそれぞれ 1.14% (0.817/0.933)、1.03% (0.855/0.883) 向 上した. Hyb-RDGS は、放射線皮膚炎画像デ ータの不均衡に対して DA に PID による人工 症例画像を加えたデータセットを用いて性能 が向上した. しかし、EfficientNet モデルでは、 表 3 に示したように RA と DA を用いたデー タセットA,Bに対して人工症例を用いた C,

D では正答率が 14% (77.3-63.6) ~27% (86.4-59.4) 低下した.特にCにおけるグレ ード4については512から256 画素の解像度 低下に伴い, sensitivity, F-measure は58% (0.833-0.250),40% (0.725-0.330)の低下を 示した(表4).人工症例画像は,その特徴で ある炎症部を正常皮膚画像に埋め込み作成さ れる.そのため,本質的な炎症部のデータ拡 張とは異なり解像度の低いデータセットでは, 学習不足になっていた可能性が考えられる. これに対して,DA は各データに最大3種類 の画像処理を行い,RA は14 個の画像処理を 対象として,k 個の画像処理操作がそれぞれ 均等に行われる[10].炎症部を含む画像処理 がデータ拡張に有効に働いたと考えられる.

本稿で用いた EfficientNet は, MobileNet と ResNet のスケールアップと NAS によって作 られた新しいモデルのスケールアップを行っ た新しいモデルである. EfficientNet は B0~ B7 までのスケーリングアップしたモデルが 提案されており, ResNet や NAS などの精度 比較が報告されている[7]. Srinivasu らは MobileNet を用いた最小限のパフォーマンス 評価として皮膚疾患の分類精度は 85.3%であ ったと報告している.しかし,特徴抽出のラ ンダム性に欠けており一連の画像テストでは, 80%未満に低下したとされている[12]. EfficientNet-B0~B7における解像度,データ セットの構成など,学習データパターンによ る研究報告は少なく十分でない.本稿では, 解像度とデータ拡張の条件を変えた EfficientNetを用いたモデルを作成し,評価し た.EfficientNetモデルは解像度およびデータ セットに依存するものの,Mobile-sizeの低解 像度,DAのデータ拡張で高い性能を示した.

3) アンサンブル学習による最終グレード判定 放射線皮膚炎の複数(グレード)判定画像 に対して,3個の EfficientNet モデルでアンサ ンブル学習を行った.それぞれの判定結果に 対してベイズ推定を用いた最終グレード判定 を行う手法を提案した.

アンサンブル学習を行った EfficienNet-B1, B4 および B7 モデルのそれぞれの性能は,正 答率が 77.3%,86.4%および 82.1%であった. 本稿では,それぞれの EfficientNet モデルのグ レード毎の正答率を基に予測された事後確率 から,真のグレードであるとされた場合のベ イズ推定量を(2)式から求めた.(2)式は, 確率分布 p と q の近似性を表現している対数 関数である.q は,真のグレードと同じにな る確率を表現するので,近似すると誤差が小 さくなり,近似しなくなると誤差が大きくな る.つまり,それぞれの EfficientNet モデルが 判定したグレードについて,真の確率分布に 近いベイズ推定量を評価することで最も正解 に近いグレードを予測できる.

また、本稿でアンサンブル学習に用いた EfficientNet モデルは、データの不均衡を補う データセットを複数作成したものである.表 4 に示したように、それぞれグレード毎の sensitivity が異なる.提案したベイズ推定は、 モデルの sensitivity が異なれば、事後確率も 異なりその結果、ベイズ推定量も変わること になる. ベイズ推定による判定例をあげると モデル A, B, C の判定がグレード 1, 2, 2 と判定されたとする. 単純多数決では, グレ ード2と判定してしまう. ここでベイズ推定 から真のグレードに対する確率で評価する. B, C はグレード 2 に対するベイズ推定量は 小さいが, グレード1に対しては, グレード A の方がベイズ推定量の小さいモデルである とする. 真のグレードに最も近いモデル A の 判定 (グレード 1) 結果が最終判定となる.

本稿で提案したアンサンブル学習で複数 (グレード)判定に対する判定結果は,グレ ード1~4の分類を行うモデルにおいて,例え ば,評価者がグレード1または,2とした判 定(3または,4と判定しない)と一致した割 合は96.7%であった.また,EfficientNetモデ ルの作成については,解像度の条件を変えた 24パターンとデータ拡張の条件を変えた4パ ターンの構成を検討した.最終的にアンサン ブル学習に精度の高い3個のモデルを選択し た.少ないモデルであっても,ベイズ推定を 用いて推論することで不確実なグレード判定 の推定に有効であることを確認した.

3. まとめ

本稿では、アンサンブル学習に基づく EfficientNet モデルを用いた放射線皮膚炎の グレード判定システムを開発した.作成した EfficientNet モデルは、Mobile-size の低解像度、 DA によるデータ拡張でグレード判定の正答 率は 86.4%を達成した.

また、グレード判定の複数(グレード)判 定について、アンサンブル学習にベイズ推定 を用いる手法を提案した.アンサンブル学習 に用いる分類器の数が少ない場合にも有効で あり、特に評価者の判定が相違するグレード 1と2、2と3において、評価者の判定と高い 精度で一致する性能を持つことを確認した.

本稿で提案した EfficientNet モデルを用い たアンサンブル学習は,評価者のグレード判 定が相違した場合の手助けとなり,汎化性能, 効率化が期待できると考えられる.

謝辞

ご多忙の中,本研究のグレード評価にご協 力いただきましたメディポリス国際陽子線治 療センターの医師,看護師の方々に深謝いた します.

利益相反の有無

利益相反無し.

文 献

- [1] Trueman E: Managing radiotherapy induced skin reactions in the community. J community Nurs 27(4): 16-24, 2013.
- [2] Chan RJ, Mann J, Tripcony L, et al.: Natural oil-based emulsion containing allantoin versus aqueous cream for managing radiation-induced skin reactions in patients with cancer: a phase 3, double-blind, randomized, controlled trial. Int J Radiat Oncol Biol Phys 90: 756-64, 2014.
- Jonathan Leventhal, Melissa Rasar Young: Radiation Dermatitis: Recognition, Prevention, and Management. Radiother Oncol 31: (12): 885-887, 894-899, 2017.
- [4] National Cancer Institute CTEP: Common Terminology Criteria for Adverse Events version 4.0 2010; MedDRA v12.0 Code 10037767.
- [5] 和田清隆,渡邊睦,新地真大,他:ハ イブリッド生成法によるディープラ ーニングを用いた放射線皮膚炎グレ ード判定システムに関する研究.日 本放射線技術学会誌 77(8),2021. (in press)
- [6] K. Simonyan, A. Zisserman. Very deep convolutional networks for large-scale image recognition. The IEEE international Conference on Computer Vision (ICCV); arXiv: 1409. 1556v6 [cs.CV] 2015.

- [7] Tan, Mingxing, Quoc V. Le. Efficient-Net: Rethinking Model Scaling for Convolutional Neural Networks. arXiv preprint arXiv:1905.11946, 2019.
- [8] Sheela Ramanna, Cenker Sengoz, Scott Kehler, Dat Pham. Near real-time map building with labelling and classification of road conditions using convolutional neiral networks. arXiv: 2001. 09947v1 [cs.CV], 2020.
- [9] Howard, Andrew G., et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861,2017.
- [10] Sandler, Mark, et al. Mobilenetv2: Inverted residuals and linear bottlenecks.
 Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.
- [11] Howard, Andrew, et al. Searching for mobilenetv3. arXiv preprint arXiv:1905.02244 ,2019.
- [12] Srinivasu P.N, Sivasai J.G, Ijaz M.F, et al. Classification of skin disease using deep learning neural networks with mobilenet V2 anf LSTM. Sensors 21: 2852, 2021.
- [13] Ekin D. Cubuk, Barret Zoph, Dandelion Mane, et al. AutoAugmnet: Learning augmentation strategies from data. CVPR, 2019.
- [14] Ekin D. Cubuk, Barret Zoph, Jonathon Shlens, Quoc V. Le. RandAugment: Practical automated data augmentation with a reduced search space. arXiv: 1909. 13719v2 [cs.CV], 2019.
- [15] Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 542: 115-118, 2017.
- [16] Y Fujisawa, Y Otomo, Y Ogata, et al.

Deep-learning-based, computer-aided classifier developed with a small dataset of clinical images surpasses board-certified dermatologists in skin tumour diagnosis. British Journal of Dermatology 180: 373-381, 2019.

- Gao Huang, Yixuan Li, Geoff Pleiss, Zhuang Liu, et al.Snapshot ensembles: Train 1, get M for free. arXiv:1704.00109[cs.LG], 2017.
- [18] 上田 修功. アンサンブル学習-識別 器の識別性能向上法および情報統合 の数理-. 社団法人情報処理学会. CVIM: 145 (26), 2004.
- [19] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. arXiv: 1512.03385[cs.CV], 2015.
- [20] 高橋和子.クラス所属確率を利用した アンサンブル学習. The 24 th Annual Conference of the Iapanese Society for Artificial Intelligence, 2010.
- [21] 杉山将,鈴木大慈.機械学習のための 数学.情報処理,vol56(5),10-15,2015.
- [22] 手良向聡.希少がんに対する臨床試 験デザイン. The Japanese Journal of Pediatric Hematology/Oncology 56(5): 425-428, 2019,.
- [23] 佐々木春喜.診断推論と確率-ベット サイドでのベイズ定理-.日本プライ マリ・ケア連合学会誌 36(3): 191-197, 2013.

Development of Radiation Dermatitis grading method based on Ensemble learning using EfficientNet

Kiyotaka WADA*1, Mutsumi WATANABE*2, Masafumi SHINNO*2, Kousuke Noguchi*2, Takashi Ogino*1

*1 Medipolis Proton Therapy and Research Center

*2 Graduate School of Science and Engineering Kagoshima University.

Radiation dermatitis is classified into 5 grades from 1, representing mild disease to 5, death due to adverse events based on the Common Terminology Criteria for Adverse Events used in clinical practice. However, it is based on visual assessment; therefore, there are issues that depend on individual experience and knowledge. The authors have generated artificial case images and constructed a hybrid generation method for a radiation dermatitis grading support system using deep learning to deal with the limitation posed by the relatively small number of cases used for learning. In this paper, we created a model using the recently proposed EfficientNet model and describe a method for the final classification of ambiguous images based on ensemble learning. We describe a method for the final classification based on ensemble learning for images that are graded differently by evaluators. The learning model using Efficient-Net-B0 to B7 with different image resolution and data extension method conditions showed an overall accuracy of 86.4%. Moreover, ensemble learning was performed with the EfficientNet model to resolve the ambiguity of grade judgment. The effectiveness of the proposed method was confirmed via an evaluation experiment using the maximum a posteriori estimation method based on Bayesian' theorem (Bayesian estimation).

Key words: radiation dermatitis, deep learning, EfficientNet, ensemble learning, Bayes' theorem

著者紹介



和田 清隆 (わだ きよたか) メディポリス国際陽子線治療センター・鹿 児島大学大学院理工学研究科博士後期課 程在籍.放射線治療専門放射線技師.放射 線取扱主任者第1種.陽子線治療,医用画 像を中心とした研究に従事.

畳み込み自己符号化器つき Deep Autoencoding Gaussian Mixture Model を用いた胸部 CT 画像における異常検知

栗林 雅刀*1 間普 真吾*1 若本 亮佑*1

呉本 尭*2 木戸 尚治*3

要旨

近年、デジタル医用画像の撮影技術が発展し、医師一人当たりの読影を要する医用画像数は増加傾向にあ る.そこで、医師の負担を軽減するために、深層学習を用いて医用画像を分類することで、医療診断の支 援を行う研究が進められている.深層学習には、ラベル付けされた大量の画像が必要となるが、疾患のあ る画像は疾患のない画像に比べて数が圧倒的に少ないため、学習に必要な異常データが十分に得られない という問題がある.本研究では、疾患のない画像のみを学習データに用いて、正常領域を表現する分布を つくり、正常領域から外れたデータを異常とみなすシステムを構築する.画像の特徴抽出には畳み込み自 己符号化器、クラスタリングには混合正規分布モデルを用い、この2つを同時に学習する畳み込み自己符 号化器つき Deep Autoencoding Gaussian Mixture Model という異常検知手法を胸部 CT 画像の異常陰影検知に 応用する.

キーワード:深層学習,自己符号化器,混合正規分布モデル,異常検知,胸部 CT

1. はじめに

近年の医療分野では,X線装置やCT装置な どの撮影機器の高性能化に伴い,膨大な数の医 用画像データが得られるようになってきた.高 齢化による患者数の増加により,医師1人が診 る画像数も増加するため,医師にかかる負担が 増える問題がある.そのため,膨大な数の医用

連絡先:

間普真吾,山口大学大学院創成科学研究科 〔〒755-8611 宇部市常盤台 2-16-1〕

e-mail: a064vgu@yamaguchi-u.ac.jp

*1 山口大学大学院創成科学研究科

*2 日本工業大学先進工学部情報メディア 工学科

*3 大阪大学大学院医学系研究科人工知能 画像診断学共同研究講座

投稿受付: 2021 年 6 月 18 日

画像データを画像認識によって処理し、医師の
 診断を支援するコンピュータ支援診断
 (Computer-Aided Diagnosis, CAD)の研究が盛ん
 に行われている.

CADとは、放射線画像をはじめとする医用画 像に対して、コンピュータで定量的に解析され た結果を「第2の意見」として利用する「医師 による診断」である[1]. 深層学習を用いた大量 の医用画像データの解析は、データ自体に、疾 患の有無を示すラベルが付与されたものを使 用することが多いが、このラベルは、医師によ って付けられるため、大量のデータにラベルを 付与する作業は医師の負担になってしまう.こ れは教師あり学習と呼ばれるもので、これに対 し、データセットにラベルを付与する必要のな いものが教師なし学習[2]である.また、疾患の ある画像は,疾患のない画像に比べて数が非常 に少ないため,データの確保が難しいという問 題もある.

本問題の解決策として、本研究では、疾患の ない正常画像のみで画像分類システムを構築 することを目的とする.具体的には、畳み込み 自己符号化器(Convolutional Autoencoder, CAE)[3]で特徴を抽出し、混合正規分布モデル (Gaussian Mixture Model, GMM)[4]で異常検知を 行う Deep Autoencoding Gaussian Mixture Model (DAGMM)を構築する.

2. 方法

DAGMM(Deep Autoencoding Gaussian Mixture Model)は、特徴抽出とクラスタリングを同時に行うことで、画像内から異常検知に適した特徴を捉えることができる異常検知手法である[5].

異常検知で従来使われていた典型的な手法 の例として,CAE による特徴抽出とOne-class SVM[6]によるクラスタリングの組合せ(以下, 従来手法)がある.従来手法の問題点は,特徴抽 出(CAE)と正常領域生成(One-class SVM)を別々 に学習するため,異常検知に必要な特徴量が得 られにくい点にある.それに対し,特徴抽出と クラスタリングを同時に行うことができる DAGMM は,異常検知に適した特徴抽出を可能 にしており,医用画像の分類にも有用であると 考えた.

1) 関心領域(ROI)画像の作成

本研究では、山口大学医学部附属病院で撮影 された胸部 CT 画像を用い、びまん性肺疾患の 陰影識別を行った.胸部 CT 画像を32×32[pixel] の関心領域(Region of Interest, ROI)に分割し、こ れを DAGMM の入力とする.ここで、ROI 画像 の生成方法を図1に示す.各 CT 画像には、陰 影の箇所が示されたマスク画像が用意されて いる.このマスク画像は、3名の放射線科医の 指導を下に作成されている.マスク画像の左上 から32×32[pixel]の領域を走査して行き、領域 内に含まれるマスク部分が7割以上の場合のみ、 対応する胸部 CT 画像から同領域を抽出し実験 用の ROI とした.



図1ROI 画像の生成方法

2) DAGMM の構成

本研究で用いる DAGMM の構造を図 2 に示 す.入力した画像の特徴抽出は CAE を用いた 圧縮ネットワーク(Compression network), クラ スタリングは推定ネットワーク(Estimation network)と GMM により実現する.

圧縮ネットワークでは、入力xの次元を圧縮 し、特徴量 Z_c に符号化する.そして、この特徴 量 Z_c を復号化することにより、再構成画像x'を 得る.この圧縮ネットワークの自己符号化器で は、以下のような入力 $x \ge x'$ の誤差を求めて特徴 量 Z_r とする.

$$Z_r = (d_1, d_2) = \left(\frac{\|x - x'\|_2}{\|x\|_2}, \frac{x \cdot x'}{\|x\|_2 \|x'\|_2}\right) (1)$$

ここで、特徴量Z_cと再構成誤差Z_rを連結させて 新たな特徴量Zとし、これを推定ネットワーク への入力とする.

推定ネットワークでは、特徴量Zがどのクラ スに含まれるかの所属確率 $\hat{\gamma}$ を全結合ニューラ ルネットワークにより出力する.この所属確率 $\hat{\gamma}$ と特徴量Zを用いて,混合比 π_k ,平均行列 μ_k , 共分散行列 Σ_k の3つのパラメータを計算するこ とで,混合正規分布を生成する.3つのパラメ



図 2 DAGMM の構造

 $- \mathfrak{p}_{\pi_k}, \mu_k, \Sigma_k$ のそれぞれの計算式は以下の通りである.

$$\pi_k = \sum_{i=1}^N \frac{\hat{\gamma}_{ik}}{N} \tag{2}$$

$$\mu_{k} = \frac{\sum_{i=1}^{N} \hat{\gamma}_{ik} Z_{i}}{\sum_{i=1}^{N} \hat{\gamma}_{ik}}$$
(3)

$$\Sigma_{k} = \frac{\sum_{i=1}^{N} \hat{\gamma}_{ik} (Z_{i} - \mu_{k}) (Z_{i} - \mu_{k})^{T}}{\sum_{i=1}^{N} \hat{\gamma}_{ik}}$$
(4)

これらのパラメータにより GMM を生成したの ち,特徴量Zのエネルギーが計算される. エネル ギー関数は以下の式で表される.

E(Z)

$$= -\log\left(\sum_{k=1}^{K} \pi_{k} \frac{\exp\left(-\frac{1}{2}(Z-\mu_{k})^{T} \Sigma_{k}^{-1}(Z-\mu_{k})\right)}{\sqrt{|2\pi\Sigma_{k}|}}\right)$$
(5)

入力したデータが GMM の分布の中心に近づく ほど特徴量Zのエネルギーは小さくなり,分布 の中心から離れるほど特徴量Zのエネルギーは 大きくなる.

DAGMM は以下の目的関数を最小化するよう に学習を行う.

$$E(Z) = \frac{1}{N} \sum_{i=1}^{N} ||x_i - x_i'||_2^2 + \frac{\lambda_1}{N} \sum_{i=1}^{N} E(Z_i) + \frac{\lambda_2}{N} \sum_{k=1}^{K} \sum_{j=1}^{d} \frac{1}{\Sigma_{kjj}}$$
(6)

右辺第1項は、入力画像xと再構成画像x'のL2 ノルムであり、入力画像が自己符号化器によっ て再構成された画像に似ているほど値が小さ くなる.右辺第2項は、式(5)で表されたエネル ギーであり、特徴量Zが混合正規分布の中心に 近いほど値が小さくなる.右辺第3項は、共分 散行列の対角要素を0としないための正則化項 である.このとき、dは圧縮ネットワークで得ら れる特徴量Zの次元数である.また本研究では、 式(6)の λ_1 、 λ_2 の値を $\lambda_1 = 0.1$ 、 $\lambda_2 = 0.001$ とした. DAGMM は、式(6)の目的関数を用いることで、 自己符号化器を学習しながら、クラスタがまと まるような特徴抽出を学習することができる. 本研究では、学習に用いるデータは全て正常 データである.テスト時に入力された正常デー タは、自己符号化器で正しく復元できることに より、エネルギーが小さくなる.一方で異常デ ータは、自己符号化器で正しく復元できないこ とにより、エネルギーが大きくなる.異常検知 では、このエネルギーの大小によって、正常と 異常を分類する.なお、DAGMMのネットワー ク構造は次章で説明する.

3. 実験と結果

1) 使用するデータセット

図3にびまん性肺疾患の陰影6種類の画像例 を示す.これら6種類の陰影パターンには浸潤 影(con),粒状影(nod),肺気腫(emp),すりガラス 陰影(ggo),蜂巣肺(hcm),正常(nor)が含まれてい る.

表1に本実験で使用する陰影パターンごとの ROI 画像のデータ数を示す.本実験では,異常 検知アルゴリズムの学習と評価のため,異常陰 影の5種類[浸潤影,粒状影,肺気腫,すりガラ ス陰影,蜂巣肺]は区別せず,一まとめにして「異 常クラス」とし,「正常」との2クラス分類とし た.正常のうち,1,659枚を訓練データとして学 習し,訓練データに含まれない514枚の正常と, 計503枚の異常をテストデータとして使用した.



図3 びまん性肺疾患の陰影6種類

クラス	訓練データ	テストデータ
正常	1,659	514
浸潤影		100
粒状影		107
肺気腫		96
すりガラス陰影		98
蜂巣肺		102

2) 学習モデル

encodel

encode2

28 × 28 × 16 10 × 10 × 64

encode3

encode4

図5 圧縮ネットワークのモデル図

(DAGMM with CAE)

decode1

 $10 \times 10 \times 64$

decode2

 $28 \times 28 \times 16$

本実験で使用する, DAGMM[5]及び DAGMM with CAE の圧縮ネットワークのモデル図を図4, 5にそれぞれ示す. DAGMM は入力を1次元ベ クトルで受取る構造であり、元の 32×32[pixel] のデータを,長さ1024の1次元ベクトルに変 換した. DAGMM with CAE で用いた CAE の畳 み込みフィルタサイズは5×5[pixel]であり、ス トライド幅は1である. プーリングは,最大プ ーリング(max pooling)を使用し、プーリングの 領域は 2×2[pixel]とした.



Dropout率 誤差関数 勾配降下法の最適化アルゴリズム Adam(lr=0.0001, beta_1=0.9, beta_2=0.999) 学習回数 バッチサイズ

ここで,図4,5に関連するモデルのパラメータ を表 2,3 に示す.自己符号化器及び GMM の学 習は表 2,3 で示されたパラメータによって行 う.

次に, DAGMM 及び DAGMM with CAE の推 定ネットワークのモデル図をそれぞれ図 6,7に 示す. 全結合ニューラルネットワークによって 構成されており、入力は1次元ベクトルである. DAGMM では,長さ16のベクトル(特徴量)に, 式(1)で求めた2つの再構成誤差を連結させた, 長さ18の1次元ベクトルを入力した.DAGMM with CAE では、5×5×64 の特徴量を長さ 1600 の1次元ベクトルに平坦化し、2つの再構成誤 差を連結させた,長さ 1602 の1 次元ベクトル を入力した. ニューラルネットワークによって 4 つのクラスタに分類されるが、クラスタ数は 次節で説明する評価指標の AUC 値が最も高く なるよう設定した.

3) 評価指標

本実験での評価指標は, ROC 曲線下の面積で ある AUC(Area under the Curve)を用いた. ROC 曲線の算出のため,正常と異常の分類境界を決

表 2 DAGMM に関するパラメータ

	encode1	tanh	
活性化関数	encode2	linear	
	decode1	tanh	
	decode2	linear	
	dense l	tanh	
	dense2	softmax	
Dropout率	0.5		
誤差関数	元論文[5]に準拠		
勾配降下法の最適化アルゴリズム	Adam(lr=0.001, beta_1=0.9, beta_2=0.999		
学習回数	500		
バッチサイズ	1659		

表 3 DAGMM with CAE に関するパラメータ

活性化関数

encode 1

encode3

encode5

decode2

decode4

decode5

dense1

dense2

05

元論文[5]に準拠

500

1659

ReLU

ReLU

linear

なし

なし

linear ReLU

softmax



図7 推定ネットワークのモデル図 (DAGMM with CAE)

めるエネルギーの閾値を 11 個設定し変化させ た.各手法における AUC の 10 回平均を表 4 に 示す.表4より, CAE + One-class SVM の平均 AUC が 0.55 であるのに対して, DAGMM の平 均 AUC が 0.72 であることから,特徴抽出とク ラスタリングの同時学習を行ったことによる 識別精度の向上が見られた.また,全結合層の みの Autoencoder を使用した DAGMM と比較し, CAE を使用した DAGMM の平均 AUC が 0.74 であった.本研究で用いた ROI 画像に対しても 畳み込みニューラルネットワークを用いた特 徴抽出器の方が有用であることが分かる.

表4 谷	各手法における	らAUCの	10 回平均
------	---------	-------	--------

手法	平均AUC
CAE + One-class SVM	0.55
DAGMM	0.72
DAGMM with CAE	0.74



図 9 各陰影に対するエネルギー (DAGMM with CAE)

4) エネルギー分布に関する考察

テストデータにおける各陰影に対するエネ ルギーの値を, DAGMM および DAGMM with CAE について図 8,9 にそれぞれ示す.図 8,9 において,緑の点は,正常のテストデータのエ ネルギーを表しており,赤い点は,異常のテス トデータのエネルギーを表している. DAGMM を用いた異常検知では、5 種類の異常陰影のう ち,3種類[浸潤影(con),すりガラス陰影(ggo), 蜂巣肺(hcm)]において正常と異常の分離が見ら れたが、2種類[粒状影(nod), 肺気腫(emp)]にお いては、ほとんど分離されていないことが分か る. 一方, DAGMM with CAE を用いた異常検知 では, DAGMM で分離できなかった 2 種類[粒 状影(nod), 肺気腫(emp)]において, 一部分離が 見られた.しかし、画像に対する汎用性が一般 に高い畳み込み層を有する CAE でも,正常と 異常を分離できる特徴量を得るには、

限界があ ることが分かる.また,異常陰影パターンには, 局所領域画像のみでは判別が難しい症例もあ るため, 胸部 CT 画像全体も用いることで, 大 域領域も学習できるシステムを構築する必要 があると考えられる.

4. まとめ

本研究では、疾患画像の数が少なく、教師ラ ベルのついた画像の収集が困難であるという 条件の下,正常画像のみを用いた教師なし学習 手法である DAGMM を用いて,胸部 CT 画像の 異常検知を行った.

実験の結果より,特徴抽出とクラスタリング の学習を別々に行っていた従来手法と比較し, 同時学習を可能にした DAGMM の方が高い異 常検知性能を示した.しかし,改良を要する箇 所も見られた.特に,異常検知に適した特徴量 を得るという点で,畳み込み自己符号化器では 表現能力の限界があることが分かった.

今後の課題としては,前述の通り,自己符号 化器の改良を行うことが挙げられる.具体的に は,大量のデータで事前学習済みの特徴抽出器 によって ROI 画像の特徴量を得,それを DAGMM の入力として学習する異常検知シス テム[7]の構築を考える.また,本実験では局所 領域画像のみから陰影パターンを学習する必 要があった.しかし,胸部 CT 全体画像も context として用いる学習を行えば,陰影の周辺情報も 考慮した学習が可能となるため,識別が難しか った陰影パターンの精度も向上すると予想さ れる.

謝辞

本研究は JSPS 科研費 JP19K12120, 21H03840 の助成を受けたものです.

利益相反の有無

なし.

文 献

- [1] 藤田広志: 医用画像のためのコンピュー タ支援診断システムの開発の現状と将来. 日本写真学会誌, 2003, pp484-490
- [2] Ghahramani Z.: Unsupervised Learning, Advanced Lectures on Machine Learning, pp72-112, 2004
- [3] Chen M, Shi X, Zhang Y et al: Deep Features Learning for Medical Image Analysis with Convolutional Autoencoder Neural Network, IEEE Transactions on Big Data, 2017
- [4] Constantinopoulos C, Titsias MK, Likas A: Bayesian Feature and Model Selection for Gaussian Mixture Models, IEEE Transactions on Pattern Analysis and Machine Intelligence, pp1013-1018, 2006
- Zong B, Song Q, Min MR et al: Deep Autoencoding Gaussian Mixture Model for Unsupervised Anomaly Detection, International Conference on Learning Representations, 2018
- [6] Chen Y, Zhou X, Huang TS: One-class SVM for learning in image retrieval, IEEE International Conference on Image Processing, 2001
- [7] Tuluptceva N, Bakker B, Fedulova I et al: Anomaly Detection with Deep Perceptual Autoencoders, arXiv:2006.13265, 2020

Anomaly Detection for Chest CT Images Using Deep Autoencoding

Gaussian Mixture Model with Convolutional Autoencoder

Masato KURIBAYASHI*1, Shingo MABU*1, Ryosuke WAKAMOTO*1, Takashi KUREMOTO*2, Shoji KIDO*3

*1 Yamaguchi University

*2 Nippon Institute of Technology

*3 Osaka University

In recent years, with the development of digital medical imaging technology, the number of medical images that need to be read by each doctor has been increasing. In order to reduce the burden on doctors, research is being conducted to support medical diagnosis by classifying medical images using deep learning. Deep learning requires a large number of labeled images, but since the number of images with diseases is much smaller than those without diseases, it is difficult to obtain enough abnormal data for learning. In this study, we construct a system that creates a distribution that represents a normal region using only images without diseases as training data, and considers data that deviates from the normal region to be abnormal. We use a convolutional autoencoder for image feature extraction and a Gaussian mixture model for making normal clusters. These feature extraction and clustering models are simultaneously trained by the framework of Deep Autoencoding Gaussian Mixture Model. We apply this anomaly detection model to the detection of abnormal opacities in chest CT images.

Key words: Deep learning, Autoencoder, Gaussian Mixture Model, Anomaly detection, Chest CT
骨シンチグラムの陽性高集積検出システムの市販後学習に

よる性能変化

島田 夏帆^{*1}, 斉藤 篤^{*1}, 大﨑 洋充^{*2}, 東山 滋明^{*3}, 河邊 譲治^{*3},

中岡 竜介*4, 清水 昭伸*1

要旨

本報告では、骨シンチグラムの陽性高集積検出システムの、初期の学習データとは異なるデータ、例え ば市販後のデータを用いた再学習による性能変化について報告する.まず、手持ちのデータを市販前と市 販後に仮想的に分割する.次に、市販前の学習データを用いてネットワークを学習する.さらに、市販前 と市販後の学習データを用いて再学習を行い、市販後のテストデータを用いて性能の変化を評価する.複 数の施設のデータを用いて、市販前と市販後のデータの分割法を変えて繰り返し評価を行うことで、性能 変化をもたらした要因や再学習の有用性について議論する.

キーワード:骨シンチグラム、高集積検出、再学習, butterfly ネットワーク

1. はじめに

骨シンチグラム上の骨転移による異常集積の検 出は容易ではなく[1], コンピュータによる検出支 援が求められる.本研究室では骨格認識と陽性高 集積検出の2つの処理[2][3]からなるシステムを開 発している.このシステムの性能は,開発時と実運 用時のデータの画質などが異なる場合,期待され る性能が得られない可能性がある.従って,システ ムを実運用時の症例データを用いて再学習を行う ことで臨床現場での性能向上が期待できる.しか し,再学習後にどの程度性能が向上するかは,再学 習に用いるデータの性質に左右される[4].

本稿では、骨シンチグラムの高集積検出処理の ネットワーク[3]を用い、市販後の再学習による性 能の変化を調べ、性能へ影響を与える要因を明ら かにする.

2. 高集積検出システム

BtrflyNet[5]に Bottleneck[6]を追加した,

*1 東京農工大学大学院工学研究院

[〒184-8588 東京都小金井市中町 2-24-16] e-mail:

s215595t@st.go.tuat.ac.jp

*2 群馬県立県民健康科学大学大学院診療放射 線学研究科

*3 大阪市立大学医学部附属病院核医学科

*4 国立医薬品食品衛生研究所医療機器部



入力は同一症例の前後面骨シンチグラム,出力は 前後面それぞれの高集積検出結果(陽性,陰性,背 景)である.ネットワークの学習の損失関数には Class weighted softmax cross entropy(WSCE)を用いた (式(1)).

$$p_{ni} = softmax(y_{ni}) = \frac{e^{y_{ni}}}{\sum_{k=1}^{C} e^{y_{nk}}}$$
$$w_i = \frac{N - \sum_{n=1}^{N} t_{ni}}{N}$$
(1)
$$L_{WSCE} = -\frac{1}{N} \sum_{n=1}^{N} \sum_{i=1}^{C} w_i t_{ni} log(p_{ni})$$

ここで、Nは全画素数、nは画素のインデックス、 t_{ni} は正解ラベル、 p_{ni} は出力 y_{ni} の確率である.

3. 実験条件と実験結果

3.1 実験条件

6 施設から収集した骨シンチグラムのデータセットを,5施設の市販前学習用データセット,およ

び、1 施設の市販後学習(2/3)・評価用(1/3)データ

セットに分けた. ここで, 市販後の1 施設を注目施 設と呼ぶ. 次に, 5 施設の市販前データを使用して 学習を行う. 続いて, 注目施設以外と注目施設の再 学習用のデータセットのうちの一つを用いて市販 後の再学習, その結果を同じ施設の評価用データ を用いて評価する. 続いて, 注目施設の二つ目のデ ータセットも加えて再学習し, 評価する. なお, 市 販後データの分割法は, 注目施設ごとに, ランダム に10 回試行した. 上記の実験を, 注目施設を順番 に変えて, 6 通り行った.

実験試料には前後面の骨シンチグラム1,032症例 (512*1024[pixel], 2.8[mm/pixel]),および,対応する 高集積ラベルを使用した(東京農工大学倫理委員 会承認番号 No.200904-0238). ミニバッチサイズは 128, Optimizer には Adam を使用し,パラメータは , $\alpha = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$ とした.テスト には,検証データで最も FP+FN[pixel]が少ないネッ トワークを使用した.なお, FP と FN は,陽性高 集積かそれ以外かの2クラス分類として計算した. 評価値として,テストデータに対する Sensitivity=90%における FP(P)+FN(P)を使用する. 検定法は Wilcoxon の符号付順位和検定である.

3.2 実験結果

再学習により各施設の FP+FN[pixel]は図 2 のように変化した.



(点線:回帰直線 数値:相関係数)

この図より,再学習によってもっとも有意に FP+FN[pixel]が減少したのは施設 004 であり,逆に FP+FN に検定が保留,あるいは有意に増加したの は施設 000 であった.

また、再学習前後の FP+FN の変化を説明するために、再学習前後での BSI(骨転移の評価値)0の 症例割合、bone-to-soft tissue ratio(画質の評価値) の変化との関係を求めた(図3).この図からは、 いずれの評価値も再学習前後の FP+FN の変化と関 係があることを示唆している.

4. まとめ

本稿では、骨シンチグラムを用いた高集積検出 システムの市販後の再学習による性能の変化の実 験を行った.再学習によって施設ごとに性能変化 が見られたが、その結果はデータセット内の BSI の 変化と bone-to-soft tissue ratio の変化に左右される ことがわかった.今後は、データセットの増やし方 を工夫した再学習の実施、今回の実験結果を踏ま えたシステムの性能改善を行う予定である.

謝辞

日頃から熱心にご指導・ご討論いただく東京農工大学 清水昭伸研究室の各位に感謝いたします.なお本研究は, AMEDの「人工知能等の先端技術を利用した医療機器プ ログラムの薬事規制のあり方に関する研究(課題管理番 号:20mk0102148h0302)」の支援を受けています.

利益相反の有無

無

- **文 献** [1] 小野慈:骨シンチによる骨転移診断. 南江堂, 東京, 2002, pp1-30.
- [2] 金森巧: "骨シンチグラムからの解剖学的骨格認識 処理の改良".東京農工大学修士論文,2019
- [3] H. Wakabayashi, A. Saito, S. Higashiyama, J. Kawabe, S. Shiomi, A. Shimizu, "Detection of bone metastasis in a scintigram using U-Net," 2018 Proc. of Computer Assisted Radiology and Surgery(CARS 2018), Berlin, Germany, S260 June 2018.
- [4] 野村行弘, 増谷佳孝, 三木総一郎, 花岡昇平, 根本 充貴, 吉川健啓, 林直人, 大友邦: 遠隔読影環境に おける他施設連携型 CAD 開発, 実運用, および継 続的性能改善. Medical Imaging Technology, 32(2): 98-108, 2014.
- [5] A. Sekuboyina, M. Rempfler, J. Kukačka,
 G. Tetteh, A. Valentinitsch, J. S. Kirschke,
 B. H. Menze, "Btrfly Net: Vertebrae Labelling with Energy-based Adversarial Learning of Local Spine Prior", arXiv:1804. 01307, 2018
- [6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition", arXiv:1512.03385, 2015

Performance transition of a positive hot-spot detection system of bone

scintigram by post-market training

Kaho SHIMADA^{*1}, Atsushi SAITO^{*1}, Hiromitsu DAISAKI^{*2}, Shigeaki HIGASHIYAMA^{*3}, Joji KAWABE^{*3}, Ryusuke NAKAOKA^{*4}, Akinobu SHIMIZU^{*1}

- *1 Tokyo University of Agriculture and Technology
- *2 Gunma Prefectural College of Health Sciences
- *3 Osaka City University Hospital
- *4 National Institute of Health Sciences

Abstract:

This paper reports the performance transition of the hot-spot detection network of bone scintigrams by the re-training using a dataset different from the initial training dataset, e.g. post-marketing data. Original dataset is virtually divided into pre-marketing and post-marketing data. Then, the hot-spot detection network was trained by the pre-marketing data. Subsequently, the network is re-trained using pre-marketing and post-marketing data followed by the performance evaluation using the test data of post-marketing data. We iterate the above process by changing the division of the data of multiple hospitals and discuss factors for the performance transition as well as the effectiveness of the re-training. **Key words:** bone scintigram, hotspot detection, retraining, butterfly network

CycleGAN を用いた甲状腺組織画像のドメイン変換と

識別システムへの応用

一氏 良仁*1 間普 真吾*1 八田 聡美*2,3

稻井 邦博*2 木戸 尚治*4

要旨

近年,深層学習を用いた医用画像の識別について多くの研究がなされている.深層学習には大量の教師ラベル付きデータが必要であるが,一つの医療施設のみのデータでは不十分であることも多いため,複数施設から得たデータの活用が必要である.複数施設のデータを使用する場合,施設ごとのデータの取得条件等の違いにより特徴が異なるため,データを一定の基準にそろえる正規化を行うことで,データの特徴の分布を一致させる必要がある.本研究では、2つの施設から取得した甲状腺組織画像を対象に、甲状腺癌の識別を行うことを目的とし、甲状腺組織画像に CycleGAN によるドメイン変換を適用することで、2つのデータセットの正規化を行い、深層ニューラルネットワークの学習および識別に利用した.識別性能を評価する実験の結果,施設Aのデータセットと、CycleGANで正規化を行った施設Bのデータセットを混合し訓練データとした場合に学習の効果が高まることを明らかにした.

キーワード:深層学習, CycleGAN, 甲状腺組織, 識別, 正規化

1. はじめに

近年,深層学習を用いた医用画像の識別について,多くの研究がなされている.医用画像の 一種である甲状腺組織画像は,現在医師によって1枚ずつ識別されているが,大量の画像の識別を行う医師の負担は大きく,さらに熟練の経験が必要である.医師の負担を軽減するため,

連絡先:

間普真吾,山口大学大学院創成科学研究科 〔〒755-8611 宇部市常盤台 2-16-1〕

- Tel/Fax0836-85-9519, mabu@yamaguchi-u.ac.jp
- *1 山口大学大学院創成科学研究科
- *2 福井大学医学部病因病態医学講座分子病理学

*3 福井大学医学部附属病院病理診断科/病理部

*4 大阪大学大学院医学系研究科人工知能画像診 断学共同研究講座 本論文では深層学習によって甲状腺組織画像 の識別を行うシステムを構築し,診断の補助に 応用すること目指した研究を行う.甲状腺組織 画像識別システムを構築する際の課題として, 深層ニューラルネットワーク(今後,単にネッ トワークと呼ぶ)の学習のための大量の教師デ ータを用意するのが難しいことが挙げられる. これは,1つの施設で収集できる教師データの 数には限界があり,また,疾患の発症頻度差か ら,種別(ラベル)ごとのデータ数も異なるた めである.特に,まれな組織型や希少症例のデ ータを収集するには複数施設の協力が不可欠 である.

本研究では、教師データの不足を補うため、 2 つの施設で採取した甲状腺組織を用いて識別 システムを構築することにした.しかし、これ らの甲状腺組織は異なる施設で採取されてい るため、採取した施設ごとにデータの特徴が異

		正常	乳頭癌	低分化 癌	濾胞癌	濾胞型 乳頭癌	髄様癌	未分化 癌
訓練	ドメインA	8353	109442	5423	1425	1360	1597	6537
	ドメイン B	0	0	0	785	215	830	433
テスト	ドメインA	3105	33175	1901	503	462	516	2228
	ドメイン B	0	0	0	271	86	303	147

表1 癌の種類ごとのパッチ画像数(枚).

なる.したがって,施設ごとのデータの特徴を 同じ分布にそろえるため,データを一定の基準 に正規化する必要がある.このように,2 つの 施設の医用画像を正規化した研究として,ある 施設で撮影した胸部 CT 画像を用いて学習した 画像の識別に応用する研究が行われている[1]. この研究では,別施設で撮影された画像を CycleGAN により正規化することで,識別性能 が向上することが示されている.そこで,異な る施設で採取された甲状腺組織画像を効果的 に深層学習に用いるため,CycleGAN によりデ ータの正規化を行い,データ間の特徴を同じ分 布にすることで,識別ネットワークの学習およ び識別に活用した.

本研究では、ある施設の甲状腺組織画像のデ ータセットに、CycleGAN により正規化した別 施設のデータを追加して学習することでネッ トワークの識別性能が向上するかを検証した. また、ある施設の甲状腺組織画像で学習したネ ットワークで、CycleGAN により正規化した別 施設の甲状腺組織画像を識別できるかについ ても調査した.

2. データと手法

2.1. 使用データ

本研究では福井大学医学部附属病院で染色・ 撮影された甲状腺組織画像を使用する.染色は 細胞核と核以外を染め分ける HE 染色という染 色方法を用いた.実験では,まず HE 染色を行 った組織をスライドスキャナーで撮影し,組織 全体画像を得た.この全体画像をさらに 1024× 1024 pixel のパッチ画像に分割して識別の対象 とした.パッチ画像は医師によって癌の種類ご とに7クラスにラベル付けされている.



図 1 ドメインごとの HE 染色画像の違い.(左) ドメイン A:福井大学の症例であり,全体的に 赤が強い.(右)ドメイン B:別施設の症例であ り,全体的に紫が強い.

本研究では、一部のクラスのデータが不足し ていたため、福井大学で採取した甲状腺組織に 加えて、別施設で採取した甲状腺組織も使用す ることにした.福井大学で採取した組織をドメ インA、別施設で採取した組織をドメインBと し、図1に各ドメインのパッチ画像の典型例を 示す.ドメインA、ドメインBのどちらも同様 の手順でHE染色、撮影、パッチ画像化が行わ れたが、ドメインごとに画像の特徴が異なる. 画像の特徴の違いはネットワークの学習およ び識別を困難にする.

表1に各ドメインの癌の種類ごとのパッチ画 像数を示す.訓練データセットとテストデータ セットは、リークによりネットワークが過度な 性能を示さないように症例ごとに分割した.ド メインAは濾胞癌、濾胞型乳頭癌、髄様癌のデ ータが不足している.また、ドメインBの正常、 乳頭癌、低分化癌の画像数は0枚である.

2.2. CycleGAN によるデータの正規化

ドメイン間の特徴の分布を一致させるには, 画像をなんらかの基準に合わせる正規化を行



図 2 CycleGAN の構造. A, B はそれぞれドメイ ン A, B のデータである. G, Fはそれぞれ A→B 変換, B→A 変換を行う Generator であり, D_B は 本物の B とGで変換された偽の B を見分ける Discriminator, D_A は本物の A とFで変換された 偽の A を見分ける Discriminator である.

う必要がある.本研究では、ドメインAを基準 として、ドメインBの画像をドメインAに類似 する画像に変換することで正規化を行った.つ まり、全体的に紫が強いドメインBの画像を、 ドメインAのように赤が強い画像に変換する ことを目的として正規化を行った.

本研究ではデータの正規化に CycleGAN 用いた. CycleGAN はペアとなる画像なしに、入力 画像と出力画像のマッピングを行う生成モデ ルの一種である[2]. つまり、CycleGAN は 2 つ のドメイン間の変換を行うことを可能とする.

CycleGAN は図 2 のように, 2 組の Generator と Discriminator によって構成される. CycleGAN の目標は 2 つのドメインAとBの間の変換を行 う Generator *G*,*F*を学習することである. この目 標のため, Generator が生成した画像の特徴と変 換対象となるドメインの画像の特徴を一致さ せる Adversarial Loss と, 2 つの Generator *G*,*F* が 互 い に 矛 盾 し な い よ う に す る Cycle Consistency Loss, 入力と出力の間で色情報を維 持する Identity Mapping Loss の 3 つの損失関数 が定義される.

1 つ目の Adversarial Loss は一般的な GAN に 用いられる損失関数であり、Generator Gと Discriminator D_B に対しては、各ドメインのデー タ分布 $b \sim p_{data}(b)$ 、 $a \sim p_{data}(a)$ を用いて、式(1)で 表される.

 $\mathcal{L}_{GAN}(G, D_B, A, B)$

 $= E_{b \sim p_{data}(b)}[\log D_B(b)]$

+ $E_{a\sim p_{data}(a)}[\log(1 - D_B(G(a)))]$ (1) GはドメインBに類似する画像G(a)を生成しよ うとし、 D_B は偽物画像G(a)と本物画像bを識別 しようとする. CycleGAN では、モデルの訓練 の安定性の向上およびより質の高い画像の生 成のため、式(1)を式(2)に置き換えている.

 $\mathcal{L}_{GAN}(G, D_B, A, B)$

 $= E_{b \sim p_{data}(b)}[(D_B(b) - 1)^2]$

$$+ E_{a \sim p_{data}(a)}[(D_B(G(a)))^2]$$
 (2)

同様に、Generator FとDiscriminator D_A に対する 損失関数は、式(3)で与えられる.

 $\mathcal{L}_{GAN}(F, D_X, A, B)$ = $E_{a \sim p_{data}(a)}[(D_A(a) - 1)^2]$

 $+ E_{b \sim p_{data}(b)}[(D_A(F(b)))^2]$ (3)

Adversarial Loss のみでもGとFの学習を行う ことはできるが、様々な入力画像に対して同じ 出力画像を生成する可能性がある.したがって、 2つ目の Cycle Consistency Loss は、同じ出力画 像を生成する問題を解決する手段として導入 される. Cycle Consistency Loss はドメイン A か らの入力画像と再構成画像との L1 ノルムを求 めることであり、式(4)で表される.

 $\mathcal{L}_{cyc}(G,F) = E_{a \sim p_{data}(a)} [\|F(G(a)) - a\|]$

+ $E_{b\sim p_{data}(b)}[||G(F(b)) - b||](4)$ この制約により、ドメイン間に共通する構造を 可能な限り保つように学習することができる.

3 つ目の Identity Mapping Loss は入力と出力 の色情報を保持するためのもので,特に絵画⇔ 写真の変換などで有用である. Identity Mapping Loss は,入力画像と出力画像との L1 ノルムを 求めることであり,式(5)で表される.

 $\mathcal{L}_{identity}(G,F) = E_{b \sim p_{data}(b)}[\|G(b) - b\|]$

 $+ E_{a \sim p_{data}(a)}[||F(a) - a||] (5)$

ドメイン B の画像bが生成器Gに入力された場 合,G(b)がbに近づくように正則化する.同様に ドメイン A の画像aが生成器Fに入力された場 合,F(a)がaに近づくように正則化する.

これら3種類の損失関数を用いた, CycleGAN の最終的な損失関数を式(6)に示す.

$$\mathcal{L}(G, F, D_X, D_Y) = \mathcal{L}_{GAN}(G, D_Y, X, Y) + \mathcal{L}_{GAN}(F, D_X, X, Y) + \lambda_1 \mathcal{L}_{cyc}(G, F) + \lambda_2 \mathcal{L}_{identity}(G, F)$$
(6)

ここで、 λ_1, λ_2 はそれぞれ $\mathcal{L}_{cyc}, \mathcal{L}_{identity}$ の相対的 な重要度を決定する係数である.また、この損 失関数を用いて、式(7)で $G \geq F$ を最適化する.

$$G^*, F^* = \arg\min_{G,F} \max_{D_X, D_Y} \mathcal{L}(G, F, D_X, D_Y)$$
(7)

2.3. 実験条件

本研究で行う3つの実験および実験条件について述べる.

実験1:CycleGANによるデータ正規化を行う. 福井大学の甲状腺組織画像をドメインA,別施 設の甲状腺組織画像をドメインBとして,各ド メインの訓練画像でCycleGANの学習を行った. その後,学習を終えたCycleGANにドメインB の画像を入力し,ドメインBの画像の特徴分布 をドメイン A の特徴分布とするように変換し た.CycleGANの訓練時に読込む画像サイズは 224×224 pixel,読込んだ画像をクロップするサ イズは 196×196 pixel,テスト時は読込む画像 サイズは 224×224 pixel,クロップは行わない ものとした.

実験2:ドメインAの訓練データにCycleGAN で正規化したドメインBの訓練データを混合し 識別ネットワークの学習を行った後,ドメイン Aのテストデータの識別性能を評価した. CycleGANと識別ネットワークの学習は,各ド メインの訓練データで行い,ドメインAのテス トデータで評価を行った.表2に識別ネットワ ークの学習に使用した訓練データセットの比 較を示す.条件1はドメインAのみで学習,条 件2はドメインA+正規化前のドメインBで学 習,条件3はドメインA+正規化後のドメイン Bで学習した.また,CycleGANの学習時は,ド メインBは元々4クラスしか存在しないため, 7クラス全てのデータを用いたドメイン変換は

表 2 識別ネットワークの学習に使用するデー タセットの比較

		訓練データセッ	Ь
	۸	正規化前	正規化後
	Λ	В	В
条件 1	\bigcirc		
条件 2	\bigcirc	\bigcirc	
条件 3	\bigcirc		\bigcirc

表 3 CycleGAN の学習に使用するデータセッ トの比較

	ドメインA	ドメイン B
条件 i	全データ	全データ
条件 ii	Bと同じ4クラ スのデータのみ	全データ

むしろノイズになると考えた. そのため表3の ように, 条件 i ではドメイン A, B の全てのデ ータで学習,条件 ii ではドメイン B は全てのデ ータ, ドメイン A はドメイン B に存在する 4 ク ラスのデータのみで学習する場合を比較検証 する.識別ネットワークは ImageNet で事前学習 済みの ResNet18[3]を FineTuning して使用した. また, FineTuning するにあたって, ネットワー クの構造は, ResNet の特徴抽出層, 40%の結合 をランダムに無効化するドロップアウト層[4], 出力を行う全結合層の順につなぎ構成した. ネ ットワークに入力する画像サイズは 224×224 pixel, データオーギュメンテーションとして, 0, 90, 180, 270 度のランダム回転, 上下左右 反転, ImageNet の全画像の画素値の平均と標準 偏差で標準化を行った.損失関数は交差エント ロピー, 最適化手法は Adam [5], Adam のパラ メータとして,特徴抽出層の学習率を0.00001, 全結合層の学習率を 0.0001, weight decay を 0.0001 とした. バッチサイズは 32, エポック数 は10である.また,使用したデータは癌の種類 ごとにデータ数が異なる不均衡データである ため, 癌の種類ごとにアンダーサンプリングを 行い,ネットワークの学習を行った.評価指標 は Recall であり, 10 回の独立試行の平均によっ て評価した.

		- / / / / /)	
識別ネット ワーク条件	CycleGAN 条件	正常	乳頭 癌	低分 化癌	濾胞 癌	濾胞型 乳頭癌	髄様 - 癌	未分 化癌	平均	
条件 1		97.9	90.9	99.2	39.8	43.9	33.1	83.4	69.8	
条件 2		98.0	90.7	99.2	27.2	34.2	40.9	83.5	67.7	
条件 3	条件i	97.9	87.0	92.7	45.7	56.3	78.9	80.3	77.0	
条件 3	条件 ii	98.0	88.6	95.7	49.1	51.5	61.6	80.2	75.0	

表4 Recall(%)を用いた7クラス分類における訓練データセットごとの比較(実験2)



図 3 CycleGAN によるドメイン B の正規化の 結果. (左)入力画像, (中) $\lambda_1 = 10$ の正規化画 像, (右) $\lambda_1 = 100$ の正規化画像である.

実験3:ドメインAで学習したネットワークを 用い,CycleGANで正規化したドメインBのテ ストデータの識別性能を評価した.比較のため, 正規化しないドメインBの識別性能を同様に調 査した.CycleGANの学習は各ドメインの訓練 データで行い,識別ネットワークの学習はドメ インAの訓練データで行い,テストはドメイン Bのテストデータで評価した.識別ネットワー クとCycleGANの学習や評価方法は実験2と同 様である.

3. 実験結果と考察

3.1. 実験 1 : CycleGAN によるデータの正規 化

CycleGAN によるデータ正規化の結果を述べる. 図 3 は CycleGAN により、ドメイン B のテスト画像をドメイン A に変換した結果である. λ_1 は式(6)の L_{cyc} の係数であり、値が大きいほど

入力画像の特徴を残しつつ変換する. $\lambda_1 =$ 10,100ともに、ドメインAのように赤が強い画像に変換された.しかし、 $\lambda_1 =$ 10の場合、上段のように画像全体に細胞が写っている画像は入力画像の特徴を残しつつ正規化できたが、下段のように細胞部分と背景部分が見られる画像は背景部分が細胞のように変化,または、細胞部分が背景のように変化する傾向がみられたが、 $\lambda_1 =$ 100の場合も同様の傾向がみられたが、 $\lambda_1 =$ 10よりも少ない変化であった.直感的には、背景や細胞が大きく変化するのは好ましくないと考えられるため、以降の実験では $\lambda_1 =$ 100で正規化したドメインBの画像を用いることにした.

3.2. 実験 2: 正規化したデータを追加した学 習

正規化したデータを追加して学習を行うこ とでネットワークの識別性能が向上するか検 証した.表4は訓練データセットごとの7クラ ス分類の結果である.

Recall の平均値は条件 1 の場合が 69.8%, 条件 2 で学習した場合が 67.7%, 条件 3+i で学習 した場合が 77.0%, 条件 3+ii で学習した場合が 75.0%となった. ドメイン B にデータが存在す るクラスに注目すると, 濾胞癌, 濾胞型乳頭癌, 髄様癌はドメイン A と正規化後のドメイン B で学習した場合の Recall が最も高い値を示した が, 未分化癌は Recall の値が低下した. ドメイ ン B にデータが存在するクラスの内, 未分化癌 のみ Recall が減少した原因として, ドメイン A

211() =	. /		///	• /	/ /			
ドメインBの テストデータセット	正常	乳頭癌	低分化 癌	濾胞癌	濾胞型 乳頭癌	髄様癌	未分化 癌	平均
正規化なし	-	-	-	29.9	68.6	11.2	93.9	50.9
条件 i で学習済みの CycleGAN で正規化	-	-	-	14.8	19.8	57.8	62.6	38.8
条件 ii で学習済みの CycleGAN で正規化	-	-	-	36.2	43.0	77.9	63.3	55.1

表5 Recall(%)を用いた7クラス分類におけるテストデータセットの比較(実験3)

において未分化癌の訓練データが比較的多い ことが挙げられる.画像枚数の多い未分化癌に 正規化後のドメインBを追加しても、学習のノ イズになったと考えられる. 一方, ドメイン A のみではデータ数が不足しているため, 識別が 困難なクラスが存在したと考えられる. ドメイ ン A に正規化前のドメイン B を追加するだけ では,ネットワークの識別性能は低下するが, ドメイン A に正規化後のドメイン B を追加す ることで、ドメイン B が存在するクラスの識別 性能が向上する.特に、ドメインAのデータ数 が少ない濾胞癌,濾胞型乳頭癌,髄様癌では, 正規化後のドメイン B を追加した場合のみ Recall が高い値を示した. よって, CycleGAN で 正規化したドメイン B からもドメイン A の識 別に有用な特徴を抽出できることを意味し, CycleGAN による正規化は有用であると言える.

次に,条件3+iで学習した場合と,条件3+ iiで学習した場合を比較すると,濾胞癌は条件 iiの方がRecallは高い値を示したが,濾胞型乳 頭癌と髄様癌は条件iの方がRecallは高い値を 示した.よって,条件iiのように4クラスのみ でCycleGANの学習,正規化を行うことは,識 別ネットワークの学習に必須ではないことが 分かる.

一方,ドメインBにデータが存在しない正常, 乳頭癌,低分化癌に注目すると,ドメインAと 正規化後のドメイン B を追加した場合のみ, Recall の値は変わらないか減少した.正規化後 のドメインBの追加は,ドメインBが存在しな いクラスにとってノイズとなり,学習を困難に させると考えられる.

3.3. 実験3:正規化したデータの識別

ドメイン A の訓練データで学習したネット ワークを, CycleGAN で正規化したドメイン B のテストデータの識別に適用した結果を述べ る.表5はテストデータセットごとの7クラス 分類の結果である.なお,正常,乳頭癌,低分 化癌はドメイン B の画像枚数が 0 枚のため, Recall を表記していない.

初めに, Recall の平均値を比較すると,正規 化前のドメインBを識別した結果が 50.9%,条件 i で学習後正規化したドメイン B を識別した 結果が 38.8%,条件 ii で学習後正規化したドメ インBを識別した結果が 55.1%となった.また, 濾胞型乳頭癌,未分化癌に注目すると,正規化 前と比べて条件 i, ii ともに Recall の値が低下 した.一方,髄様癌は向上した.また,濾胞癌 は条件 i で正規化後は Recall の値が低下し,条 件 ii で正規化後は Recall の値が低下した.

正規化前と条件 i で正規化後の結果について 考察する. 従来研究[1]ではテスト画像を正規化 したことで識別性能が向上したことに反して, 本実験では癌の種類によって Recall の値に違い が見られる. この原因として 2 つ考えられる. 1 つ目は, ドメインA で学習した識別ネットワ ークの性能差である. 従来研究ではドメインA で学習したネットワークでドメイン A の検証 データを識別した際の識別率は約 90 %であり, Recall も高い値を示したと考えられるが,本研 究では Recall の平均値は 69.8%である. 従来研 究ではドメイン A のみの学習で識別に必要な 特徴を十分に抽出しているため,正規化したド

メイン B も識別できたと考えられる. しかし, 本実験において、訓練データ、テストデータと もドメイン A とした場合の Recall の平均値が 69.8%であり、識別に必要な特徴を十分に抽出 しているとは言えない.2つ目は, CycleGANの 学習の際のデータセットの違いである.従来研 究では、 ドメイン A・B ともにすべてのクラス において約 1000 枚ずつのデータセットで学習 を行っていたが、本実験では、ドメインAは7 クラスで計 134137 枚, ドメイン B は 4 クラス で計 2263 枚, 癌の種類ごとにデータ数が異な る不均衡データで学習を行った.このデータセ ットの違いにより、CycleGAN による正規化結 果に差が生じた可能性がある.本研究では,2つ 目の原因に対処するため,条件 ii で正規化後の ドメイン B も同様に検証した.条件 i で正規化 後の結果と比較して、すべてのクラスで Recall の値が向上した.しかし、実応用の場では、ド メインBのテストデータセットが4クラスであ るとは限らない. そのため, 診断対象となるク ラスのデータはドメイン A, B ともに必要であ る.

4. 結論

本研究では、甲状腺組織画像の識別において、 CycleGAN を用いた 2 施設のデータの活用につ いて研究した. CycleGAN は異なるドメイン間 のデータの相互変換を行うGANの一種であり、 これによりデータの正規化を行った.実験の結 果、別ドメインのデータの識別にCycleGAN を 適用するという点では改善が必要なものの、正 規化した別ドメインのデータを学習データに 追加した場合、特にデータ数の少ないクラスの 識別性能が向上することを明らかにした.

今後の課題として、CycleGAN による正規化 性能を向上するため、CycleGAN で生成する画 像の高精細化および識別に有用な特徴を抽出 可能な構造へ改良することが挙げられる. ま た、本研究では不均衡データである問題に対し て、アンダーサンプリングなど単純な方法を適 用したのみであるため、不均衡データに対して、 より適切な決定境界を与える学習方法を適用 することが考えられる.

謝辞

本研究はJSPS 科研費 JP19K12120, 20H03908 の助成を受けた.

利益相反の有無

なし.

文 献

- [1] 三宅将司,間普真吾,木戸尚治,呉本尭, 平野靖: CycleGAN を用いた胸部 CT 画像 のドメイン変換とその識別システムへの 応用,第 38 回日本医用画像工学会大会, 2019, pp. 108-1154
- Zhu, Jun-Yan and Park, Taesung and Isola, et al: Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. In IEEE International Conference on Computer Vision, 2017, pp. 2223-2232
- [3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, et al: Deep Residual Learning for Image Recognition, In IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770-778
- [4] Geoffrey E. Hinton, Nitish Srivastava, Alex Krizhevsky et al: Improving neural networks by preventing co-adaptation of feature detectors, arXiv:1207.0580, 2012
- [5] Diederik P. Kingma, Jimmy Lei Ba: ADAM:
 A method for stochastic optimization, In International Conference on Learning Representations, 2015

Domain Transformation of Thyroid Tissue Images Using CycleGAN

and Its Application to Classification Systems

Yoshihito ICHIUJI*1, Shingo MABU*1, Satomi HATTA*2, Kunihiro INAI*2, Shoji KIDO*3

*1 Yamaguchi University*2 University of Fukui*3 Osaka University

In recent years, a large number of research has been done on the classification of medical images using deep learning, and the performance improvement for clinical applications has been expected. However, it is difficult to prepare a large number of training data with class labels for deep learning. Especially in medical field, data prepared in only one medical institution may not be enough, therefore, data should be collected by multiple medical institutions. Even in that case, since the features of the data vary from institution to institution, it is necessary to match the features of the data by normalizing the data. In this study, thyroid tissue images obtained from two medical institutions are used and a thyroid cancer classifier is designed. By applying domain transformation of CycleGAN to the thyroid tissue images, the features of the data in the two datasets are normalized, and the normalized data are used to train deep neural networks and build classifiers. The experimental results showed that the classification performance was improved by combining the data prepared in institution A and the normalized data prepared in institution B. On the other hand, we also found that the network trained on the dataset of institution A needs to be improved to classify the normalized dataset of institution B.

Key words: Deep Learning, CycleGAN, Thyroid Tissue Image, Classification, Normalization

アテンション誘導を利用した胸部 CT 画像による COVID-19

の診断支援

高手山 雄亮*1, 斉藤 篤*1, 橋本 正弘*2, 大竹 義人*3,

明石敏昭*4,清水 昭伸*1

要旨

本報告では、COVID-19 の症例を含む胸部 CT 画像を対象にした疾病クラスの分類について述べる.単純な Convolutional Neural Network ベースの手法では、広範囲の特徴を抽出することが難しく、COVID-19 の典型 的な特徴である両側性の病変などを見落としてしまう可能性がある.提案手法は、病変を含むスライスの みを選択するスライス選択ネットワークと、選択されたスライスからクラス分類を行う病変分類ネットワ ークで構成される.本報告では病変分類ネットワークについて述べるが、これは、ResNet18 をベースとし て Attention 誘導機構を組み合わせることで、ネットワークの注目部位を COVID-19 の特徴的な部位に誘導 可能なモデルである.247 症例を用いて性能を評価した結果を示し、提案法の有効性について議論をする. キーワード: COVID-19、胸部 CT、深層学習、ResNet18、Attention 誘導

1. はじめに

新型のウイルス感染症である Coronavirus Disease 2019 (COVID-19)の臨床検査には逆 転写ポリメラーゼ連鎖反応 (RT-PCR)検査が 用いられている.しかし,RT-PCR 検査には 感度に限界があることが知られている[1].そ こで,この検査を補完する目的で,単純 X 線 像や CT 画像などを用いた放射線画像診断が 注目されている [2].しかし,専門の読影医 の数は不足し,コンピュータによる診断支援 システムが必要とされている.

COVID-19 の分類を行う先行研究としては, 胸部 CT 画像の各スライスを 2 次元的に解析

*1 東京農工大学大学院工学研究院

〔〒184-8588 東京都小金井市中町 2-24-16〕

e-mail: s216965t@st.go.tuat.ac.jp

*2 慶応義塾大学医学部放射線科(診断)

〔〒160-8582 東京都新宿区信濃町 35〕

*3 奈良先端科学技術大学院大学

〔〒630-0192 奈良県生駒市高山町 8916-5〕

*4 順天堂大学放射線診断学講座

〔〒113-8421 東京都文京区本郷 2-1-1〕

して分類する手法[3]や,胸部 CT 画像から肺 の3次元的なモデルを作成して解析,分類を 行う手法[4]などがある.また,胸部 CT 画像 のスライス単位の特徴抽出を行い,その特徴 をまとめて症例ごとに分類を行う手法[5]が ある.この文献[5]の手法は Convolutional Neural Network (CNN)を用いているが,広範囲の特徴 を抽出することが難しく,COVID-19 の典型 的な特徴である両側の肺に現れる病変などを 見落としてしまう可能性がある.

本稿では, ResNet18 に対して Attention 誘導 機構を組み込むことで,ネットワークの注目 部位を COVID-19 の特徴的な部位に誘導する 手法を提案し,より高い性能を発揮したこと を報告する.

2. 提案手法

提案手法は、CT スライスの病変の有無を 判定するスライス選択ネットワークと、選択 されたスライスのクラス分類を行う病変分類 ネットワークで構成される.本報告では後者 の病変分類ネットワークについて述べる.こ のネットワークは、図 1 に示した ResNet18 を並列に並べたベースのモデルにおいて、

Non-Local Block(NLB)[6]を追加した ResNet18 (図 2(a)) に置き換えたものである. Attention を誘導する位置は, COVID-19 の病変の両側 性という特徴を考慮して, 図 2(b)に示した Non-Local map に対して,赤い四角の画素と 左右対称の位置にある図 2(c)に示した注目マ スクを足し合わせた.



(a) ResNet18 + NLB



(b) Non-Local map
 (c)注目マスク
 図 2 NLB に対する Attention 誘導

3. 実験

実験試料として、国内の医療機関から収集 されたスライス厚 5[mm]の胸部 CT 像 247 症 例を用いた(東京農工大学倫理委員会承認番 号 No.200705-0226).本実験では、1. Typical appearance, 2. Indeterminate appearance, 3. Atypical appearance, 4. Negative for pneumonia という4クラスのラベルを用い、3-fold CV で 評価した.評価値として、4-class、COVID-19, Pneumonia の3種のAccuracyを用いる. Accuracy は以下の式であり、それぞれのTと Fの関係を図3に示す.



(a) 4-class (b)COVID-19 (c)Pneumonia 図3 3種類の Accuracy における TとFの関係

ベースのネットワークと提案ネットワーク に関して3種類の Accuracy を表1に示す.

表1 実験結果

衣 I 天厥阳木								
ベース		ベース	提案手法(ベース+					
		+NLB	NLB の Attention 誘導)					
4-class	0.5628	0.5506	0.6154					
COVID-19	0.7449	0.7409	0.8016					
Pneumonia	0.8664	0.8462	0.9069					

表1より,Attention 誘導機構を追加した提 案手法は、3種すべてのAccuracy に対して性 能の向上がみられる.また、帰無仮説を二つ の手法の性能が等しいとしたマクネマー検定 を、ベースと提案手法、ベース+NLBと提案 手法に対して行い、帰無仮説を有意水準0.01 で棄却した.

4. まとめ

本稿では、COVID-19 の CT スライスの病 変分類ネットワークにおいて、Attention 機構 を追加した. 247 症例を用いた実験結果から は、4-class、COVID-19、Pneumonia のすべて の Accuracy に対して性能が向上しており、 Attention 誘導が COVID-19 の典型例に対して 有効であることを確認した. 今後は、 COVID-19 の典型例のみでは無く、他のクラ スに関する Attention 誘導を検討し、性能の向 上を目指す.

謝辞

本研究では日本医学放射線学会のナショナルデータ ベース J-MID(臨床研究等 ICT 基盤構築・人工知能実装 研究事業/画像診断ナショナルデータベース実現のた めの開発研究(課題管理番号:JP20lk1010025))を利用 しました.データベース構築に関わっていただいた先生 方に感謝いたします.また,本研究は,AMEDの医療ビ ッグデータ利活用を促進するクラウド基盤・AI 画像解析 に関する研究(課題管理番号:JP20lk1010036)の支援を 受けています.

利益相反の有無

なし

[1]

Fang. Y, "Sensitivity of chest CT for COVID-19: comparison to RT-PCR", Radiology, 19 Feb 2020.

文 献

- [2] Xie, X, "Chest CT for typical 2019-nCoV pneumonia: relationship to negative RT-PCR testing", Radiology, 12 Feb 2020. 1983
- [3] M. Rahimzadeh, "A Fully Automated Deep Learning-based Network ForDetecting COVID-19 from a New And Large Lung CT ScanDataset", medRxiv, 8 Jun 2020.
- [4] Stephanie A. Harmon, "Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets", Nature Communication, July 2020.
- [5] L. Li, "Artificial intelligence distinguishes

covid-19 from community acquired pneumonia on chest ct", Radiology, Mar 2020.

[6]

6] X. Wang, "Non-local Neural Networks," CVPR, 2018.

Computer-aided diagnosis of CT images of COVID-19 using deep

learning with attention induction

Yusuke TAKATEYAMA^{*1}, Atsushi SAITO^{*1}, Masahiro HASHIMOTO^{*2}, Yoshito OTAKE^{*3}, Toshiaki AKASHI^{*4}, Akinobu SHIMIZU^{*1}

- *1 Institute of Engineering, Tokyo University of Agriculture and Technology
- *2 Department of Radiology, Keio University School of Medicine
- *3 Nara Institute of Science and Technology
- *4 Department of Radiology, Juntendo University

This paper describes the classification of chest CT images of COVID-19. Conventional convolutional neural network (CNN) based approach might not detect typical lesions of COVID-19, such as bilateral lesions, because a CNN is not suitable to capture features distributed in distant areas. The proposed classification method consists of a slice selection network to select slices with lesions followed by a lesion classification network to classify the selected slices. This paper focuses on the lesion classification network that induces attention to distant areas with COVID-19 specific features by combining an attention mechanism with ResNet18. We present the results of applying the proposed classification method to CT images of 247 cases and discuss the effectiveness.

Key words: COVID-19, Chest CT, Deep learning, ResNet18, Attention induction

口腔診断支援のための舌の色特徴に基づく

機械学習を用いた疾患予測

野口 桂冴*1 斎藤 一郎*2 吉村 裕一郎*3 並木 隆雄*4

渡辺 悠紀*4 中口 俊哉*5

要旨

舌診は舌特徴を捉え, 簡便かつ非侵襲的に健康状態を把握する診断法であるが, 医師の主観や経験に依存するため, 定量化が求められている.本研究ではドライマウスや口腔疾患を誘発するシェーグレン症候群(SS)について, 診断支援を目的に舌画像に基づく解析を行った. SS 患者 19 名を含む, 計 60 名において舌画像を取得した.舌形状規格化のため,舌を5角形に近似し同一の解析領域を指定した.舌色をCIELAB空間で取得し Support Vector Machine, Random Forest, Logistic Regression を訓練した. SVM が最も高い性能を示し, 交差検証における平均精度 71.3%, 感度 57.5%, 特異度 78.1%となった.本手法は現行の診断基準の一つである唾液量検査と比較し同等の診断性能を示しており,衛生面と簡便性において優位であるため, コロナ禍に適した新たな診断法となることが示唆された.

キーワード:舌診,画像分類,機械学習

1. 背景

漢方の診断法の一つである舌診は舌の色, 形状,湿潤,舌苔などの所見から患者の体調 を診断する.我々は舌撮影の信頼性向上を目 的に,図1に示すような舌撮影装置 TIAS

(Tongue Image Analyzing System)を開発し, 舌色診断支援システムを構築した[1,2]. 舌画 像による舌診は簡便かつ非侵襲的であり,疾 患診断への利用が期待されている.

舌所見に変化を与える疾患の一つとして, シェーグレン症候群 (SS: Sjogren's syndrome)

- *1 千葉大学大学院融合理工学府
- 〔〒263-8522 千葉市稲毛区弥生町 1-33〕
- *2 鶴見大学歯学部 病理学講座
- *3 富山大学附属病院放射線部
- *4 千葉大学大学院医学研究院和漢診療学
- *5 千葉大学フロンティア医工学センター

がある[3,4]. SS はドライマウスを主徴とす る自己免疫疾患であり、日本においては指定 難病に位置付けられている.免疫異常、ホル モン分泌、遺伝等の因子が複雑に関係し発症 すると考えられているが、原因は未だ明らか にされていない.そのため、様々なアプロー チによって、診断及び治療法の研究が進めら れている.



図1 舌撮影装置 TIAS と舌色診断支援システム

SS の診断では口腔乾燥を定量的に評価す る方法として, 唾液を採取し計量するサクソ ンテストが用いられている[5]. しかし, 近年 の COVID-19 感染予防の観点から, 唾液に関 する検査は実施しづらいものとなっており, 口腔状態を評価する客観的な代替手法が必要 である.

本研究では、舌とSSの関係を調査し、舌 診によるSSの診断の可能性を示す.診断基 準であるサクソンテストの分類結果と比較を 行うことで、コロナ禍に適した診断手法を舌 診が提供可能であること示す.

2. 手法

SS 患者を含む 60 名において, TIAS を用い て舌を可視光撮影し,舌画像(1024× 1280 画素)を取得した.

舌診において特に重要な所見である,舌色 について抽出を行う.撮影条件に依存しない 色特徴を取得するため,図2(a)に示す校正用 カラーチャートを用いて,RGBから CIE1976L*a*b*へ変換した.また,大きさや 形の異なる舌写真から舌部位を決定するため, 図2(b)に示すように,舌形状を手動入力によ って5角形に近似した.その後,図3に示す ように舌辺,舌根,舌中,舌尖の各領域で色 特徴を抽出した.抽出した特徴について統計 解析を行った.検定には対応のないt検定及 びU検定を用いた.

舌色特徴を用いて機械学習アルゴリズムで ある Logistic Regression(LR), Support vector machine(SVM), Random Forest(RF) の3つを 学習した. LR は最も古くから使用されてき た分類のためのアルゴリズムであり、入力変 数をシグモイド関数で確率変数へ写像した後, 閾値によって分類を行う. SVM は, カーネル 法を用いることで、LR と比べ非線形的な分 類が可能となっている. 一方で RF は, 簡単 な条件分岐に基づく弱学習器を複数学習し, 組み合わせることによって高性能な分類を可 能とする.入力する特徴量は平均を 0,分散 を1となるようにスケーリングした. 学習で はデータの不均衡性を解消するため、ランダ ムに学習サンプルをオーバーサンプリングし てから学習した.また,主成分分析によって 入力次元の削減を行った.パラメータはそれ ぞれグリッドサーチによって総当たり的に最 適化を行い、モデルの検証には層化5分割交 差検証を用いた.更に、分割の偏りによる影 響を低減するため,交差検証を10回繰り返し て平均スコアを算出した.





(a) カラーチャート (b) 舌形状近似結果 図2 取得された舌画像の例



図3 (1)舌辺, (2)舌根, (3)舌中, (4)舌根の4領域の定義

3. 実験と結果

抽出された各領域における舌色について, SS 患 者群と非 SS 患者群間で比較及び有意差検定を行 った. a*値に関する結果を図 4, b*値に関する結 果を図 5 に示した. 舌根部の a*(p<0.05), 舌中部 の b*値(p<0.05)において統計的有意差を認めた. これにより, SS 患者と非 SS 患者間では舌色が異 なることが示唆された.

最適な分類アルゴリズムと特徴量を決定するた め、比較実験を行った. 各モデルは舌色、舌色主 成分, 舌色主成分と患者の基礎情報の3つで訓練 された. また, 目標値として, サクソンテストに よって得られた唾液量を入力にモデルを訓練した 結果も示した.比較結果を表1に示す.分類性能 を示す AP では、どの特徴量を用いた場合でも SVM の結果が LR, RF よりも優位であった. LR とRFの分類結果ではいくつかの特徴量において、 感度が100%に近く、特異度が0%に近いような結 果が得られたが、これは検証サンプルのほとんど を陽性と予測した結果であり、入力した特徴とク ラスとの関係を考慮できていない場合であった. SVM に舌色主成分に加え, 患者の性別と年齢を入 力として与えた場合に,最も高い性能を示し,AP についてはサクソンテストを入力した場合とほぼ 同等であった.

分類器から出力される信頼度に対し,全ての値 で感度と適合率を計算しプロットする,PR曲線に よる比較結果を図6に示す.全感度帯において舌 色主成分を用いた場合のほうが,舌色を用いた場 合よりも適合率が良く,主成分分析による次元削 減の効果が示された.また,年齢と性別を入力に 加えることで,更に性能が向上し,全感度帯にお いてサクソンテストと同程度の適合率を示してい ることが確認された.



図4 非SS 患者とSS 患者における舌色 a*値の比較







図6 SVM を用いた入力特徴量ごとの PR 曲線の比 較

4. 結論

本研究では舌とSSの関係を明らかにするため, SS 患者群から舌画像を取得し解析を行った.舌色 についてSS 患者群と非SS 患者群間で舌根部のa* 値,舌中部のb*値において統計的有意差があり, SS が舌色に変化を与える可能性が示唆された.関 連研究[3,4]によれば,SS 患者ではドライマウスに 起因する舌粘膜の炎症により,舌全体の色味が赤 方向へ偏移することが示唆されている.本研究の 結果も舌根部のa*値が有意に大きく,これを支持 している.一方で舌中部のb*値においても有意差 が得られたが,これは舌表面の苔による影響によ り,舌中央が黄色味を帯びている可能性がある. ドライマウスでは口内の雑菌繁殖により,舌に苔

が厚く乗る症状が見られることから[7], 今回の結 果が得られたと考えられる.

舌画像を用いて SS を予測する機械学習分類器 を学習したところ, SVM を用いた場合に最もスコ アが優れていた.患者の基礎情報も追加すること で,更に診断能が向上することを示し, SS 診断基 準の一つであるサクソンテストを用いた分類器の 診断能(mAP=0.670)と比較して,舌色に基づく 分類器も同等の性能(mAP=0.664)が得られた.

唾液を採取するサクソンテストと比べ、舌画像 を用いた予測は衛生面、簡便性において有意であ る.SS 患者では舌の初見に変化が現れることが統 計的・機械学習的手法によって明らかとなり、舌 診が SS 診断の一助となる可能性が示唆された. 今後は更に症例数を収集し、信頼性と精度を向上 することが必要であると考えられる.

利益相反の有無 無し.

文 献

[1] Yamamoto S, Ishikawa Y, Nakaguchi T, et al., Temporal Changes in Tongue Color as Criterion for Tongue Diagnosis in Kampo Medicine. Forsch Komplementmed 2012: 80-85, 2012.

- [2] Nakaguchi T, Takeda K, Ishikawa Y, et al., Proposal for a new non-contact method for measuring tongue moisture to assist in tongue diagnosis, and development of the Tongue Image Analyzing System, which can separately record the gloss components of the tongue. BioMed Research International Article ID 249609, 10 pages, 2015.
- [3] Maeda M., Dermoscopic patterns of the filiform papillae of the tongue in patients with Sjögren's syndrome. J Dermatol., Feb 33(2):96-102, 2006.
- [4] Błochowiak K, Olewicz-Gawlik A, Polańska A, et al. Oral mucosal manifestations in primary and secondary Sjögren syndrome and dry mouth syndrome. Postepy Dermatol Alergol. 33(1):23-27., 2016.
- [5] Tsuboi H, Hagiwara S, Asashima H, et al., Comparison of performance of the 2016 ACR-EULAR classification criteria for primary Sjögren's syndrome with other sets of criteria in Japanese patients. Ann Rheum Dis. 2017 Dec;76(12):1980-1985, 2017.
- [6] López-Pintor RM et al., Oral involvement in patients with primary Sjögren's syndrome.
 Multidisciplinary care by dentists and rheumatologists. Reumatol Clin. Nov-Dec;11(6):387-94, 2015.
- [7] Soto-Rojas AE, Villa AR, Sifuentes-Osornio J, Alarcón-Segovia D, Kraus A. Oral manifestations in patients with Sjögren's syndrome. J Rheumatol. May;25(5):906-10, 1998.

分類器	特徴量	精度	感度	特異度	AP
LR	サクソンテスト	0.665	0.735	0.633	0.670
	舌色	0.330	0.985	0.026	0.514
	舌色主成分	0.317	1.000	0.000	0.562
	舌色主成分,性別,年齡	0.432	1.000	0.169	0.560
	サクソンテスト	0.675	0.730	0.650	0.670
	舌色	0.472	0.895	0.281	0.510
5 V W	舌色主成分	0.625	0.567	0.652	0.580
	舌色主成分, 性別, 年齡	0.713	0.575	0.781	0.664
RF	サクソンテスト	0.598	0.643	0.578	0.601
	舌色	0.370	0.980	0.089	0.502
	舌色主成分	0.317	1.000	0.000	0.568
	舌色主成分,性別,年齡	0.615	0.692	0.580	0.609

表1 診断予測結果の比較,各スコアは交差検証の平均値

Oral disease prediction using machine-learning based on color feature

of tongue image for diagnostic support system

Keigo NOGUCHI^{*1}, Ichiro SAITO^{*2}, Yuichiro YOSHIMURA^{*3}, Takao NAMIKI^{*4}, Yuki WATANABE^{*4}, Toshiya NAKAGUCHI^{*5}

*1 Graduate School of Science and Engineering, Chiba University

*2 Department of Pathology, Tsurumi University School of Dental Medicine

*3 Department of Radiology, Toyama University Hospital

*4 Dept. of Japanese-Oriental (Kampo) Medicine, Graduate School of Medicine,

Chiba University

*5 Center for Frontier Medical Engineering, Chiba University

Tongue diagnosis is a simple and non-invasive diagnostic method to estimate the health condition of a patient by detecting tongue features, but since it depends on the subjectivity and experience of the physician, quantification is required. In this study, we analyzed tongue images to support the diagnosis of Sjogren's syndrome (SS), which is characterized by dry mouth. 60 tongue images were obtained, including 19 SS patients. To standardize the tongue shape, the tongue was approximated as a pentagon and the same analysis area was designated. We trained Support Vector Machine, Random Forest, and Logistic Regression on tongue color in CIEL*a*b* space. SVM showed the best performance, with an average accuracy of 71.3%, sensitivity of 57.5%, and specificity of 78.1% in cross-validation. Our method showed the comparable performance as the Saxon test, which is one of the diagnostic criteria, and was superior in terms of hygiene and simplicity, suggesting that it will be a new diagnostic method suitable during the COVID-19 pandemic.

Key words: Tongue diagnosis, Image classification, Mathine-Learning

P1-11

腸閉塞・イレウスの病変箇所特定における

診断支援システムの精度評価

小田 紘久*1 林 雄一郎*2 北坂 孝幸*3

玉田 雄大*4 滝本 愛太朗*1 檜 顕成*1 内田 広夫*1

鈴木 耕次郎*5 小田 昌宏*6,2 森 健策*2,7,8

要旨

本稿では,腸閉塞・イレウス診断支援のための腸管抽出手法に関して,抽出結果からの病変箇所の見つけ やすさに関する精度評価を行う.腸閉塞ないしイレウスの画像診断において,医師は CT 像において腸管 の走行を追跡して病変箇所を探す.診断支援システムは,CT 像における腸管のうち,ある程度の太さをも った区間をそれぞれ抽出する.このとき,いずれかの抽出結果の端点が病変箇所を表す可能性が高い.そ こで本稿では,病変箇所の探索における診断支援システムの有効性を評価する.抽出結果の端点が病変箇 所に近接しており,また端点の数が少ないほど,医師は容易に病変箇所を探索できる.そのため,抽出結 果と病変箇所との最短距離,ならびに抽出結果の端点数(病変候補数)を計測する.CT 像 104 例において 腸管領域を抽出したところ,最短距離の平均は 19.3 mm,病変候補数の平均は 23.0 個であった.多くの症 例において,病変箇所の至近まで腸管領域の抽出が可能であることを確認した.

キーワード:小腸抽出,3次元可視化,グラフ表現

*1 名古屋大学大学院医学系研究科

〔〒466-8560 愛知県名古屋市昭和区鶴舞 町 65〕

e-mail: hoda@mori.m.is.nagoya-u.ac.jp

- *2 名古屋大学大学院情報学研究科
- *3 愛知工業大学情報科学部
- *4 名古屋大学医学部医学科
- *5 愛知医科大学放射線医学講座
- *6 名古屋大学情報連携推進本部情報戦略 室
- *7 名古屋大学情報基盤センター
- *8 国立情報学研究所医療ビッグデータ研 究センター

1. はじめに

腸閉塞 (intestinal obstruction) とイレウス (ileus) はいずれも腸の疾患である.両者に共通 するのは,腸管の内容物が肛門へ向かって正常 に流れなくなることである.腸閉塞は閉塞や絞 扼といった機械的な閉塞によるものを指すの に対し,イレウスは麻痺などによって蠕動運動 が正常に行われない状態によるものを指す [1]. これらの疾患においては,血管の造影された CT 像を用いた画像診断を用いて,病変箇所の 特定が行われる.しかしながら,腸管を追って 病変部位を見つける作業は,慣れない医師にと っては容易でない.

腸閉塞・イレウスにおける画像診断を支援す るため, Oda らは腸管抽出手法 [2] を提案し た.この手法を用いて CT 像から腸管を抽出す ると,腸管のうち内容物によって太くなった区 間が抽出される (この個々の抽出結果を腸管セ グメント領域とよぶ).このとき,1個以上の腸 管区間領域の端点が病変箇所付近に存在する ことが期待される.医師は腸管セグメント領域 の端点ごとに,その付近に病変箇所が存在する かどうかを確認することで,病変箇所に辿り着 くことが可能である.

しかしながら Oda ら [2] は,病変部位の見 つけやすさの観点から手法を定量評価してい なかった.そこで本稿では,腸管セグメント領 域の端点を"病変候補"とみなし,その病変候 補の数ならびに病変部位からの最短距離を求 め,病変部位の見つけやすさの定量評価を行う.

2. 手法

1) 腸管抽出手法

ヒトの腸管は数メートルに及び,腹部におい て複雑に走行する.単なる腸管領域抽出では, 腸管のうち離れた2区間が互いに近接している 場合,抽出結果が誤って接続する場合がある (この問題を Shin ら [3] は "touching issue"

と呼称). そこで Oda ら [2] は, 腸管の中心線 付近のみを抽出した. 2 セグメントが近接して いる場合も,中心線付近のみの抽出結果ならば, 互いに距離が確保されるため,誤って接続され なくなる.以下, Oda らの手法 [2] の概略を述 べる.

まず 3D U-Net [3] を用いて,入力画像の腸 管領域内における距離画像の推定を行う.この 距離画像は,腸管の中心線付近において1.0 程 度,腸管の辺縁付近では 0.0 となる.この距離 画像の極大点から Watershed アルゴリズム [4] を適用して得られた領域を,本稿ではそれぞれ 「腸管ブロック」と呼称する.Watershed アルゴ リズムの開始点が距離画像の極大点であるた め,腸管ブロックは腸管の中心線に沿うように 多数生成される.互いに接触する腸管ブロック の組ごとに,枝葉を取り除き,端点から端点ま でのルートを求める.腸管セグメント領域は, このルート上に位置する腸管ブロックの領域 の和として定義される.最終的な腸管抽出結果 は、すべての腸管セグメント領域の組となる. 2) 評価法

腸管抽出手法[2]は腸閉塞およびイレウスの 診断支援システムとして用いることを目指し ている.抽出結果から,患者の病変箇所を正確 かつ簡便に検出できることが肝要である.その ため本稿では,抽出結果および手動でアノテー ションされた病変箇所をもとにして,以下の2 指標に基づき各症例の評価を行う.

● 最短距離

病変箇所と, すべての腸管区間領域の両端付 近との最短距離. この最短距離が短いほど, いずれかの腸管抽出結果の端点が病変箇所 を的確に示していることを表す.

● 病変候補数

腸管区間領域数の2倍(端点の数).値が小さ いほど,提示される病変候補の数が減るため, 個々の病変候補が実際に病変箇所であるか どうか,目視で確認する負担が減る.

3. 実験および考察

腸閉塞またはイレウスの患者の CT 像 110 症例を用いて,4 分割交差検定を実施した.3D



(a) 最短距離 11 mm, 病変候補数 14 個.



(b) 最短距離 41 mm,病変候補数 24 個.
 図 1 結果例.腸管区間領域それぞれをレインボーカラーで表示.

U-Net により距離画像を推定するため,腸管領 域のラベルを医学生が手動で作成したものを 学習に使用した.病変箇所は 110 症例のうち 104 例に対して,外科医が手動でアノテーショ ンした.

実験の結果,104 例における最短距離および 病変候補数の平均・標準偏差は,それぞれ19.3 ±16.4 mm および23.0±8.1 個となった.図1に 結果の例を示す.

図 1 (a) の例では, 最短距離が 11 mm と比較的短く, 抽出結果から病変箇所を見つけるの は容易といえる. 全症例の最短距離の平均は 19.3 mm に留まっており, 多くの症例で良好な 結果が得られたものと考えられる. ただし, 本 症例の病変候補数は 14 個にのぼっている. た とえば鋭敏なカーブにおいて腸管区間領域が 複数に分かれており, 病変候補の増加の一因と なっている. 今後, このような誤った病変候補 の修正が必要と考えられる.

図 1 (b) の症例の最短距離は 41 mm であった.図 1 (a) の症例と比較すると、本症例では 全体に腸管が細く、病変箇所の近くまで腸管ブ ロックが生成されなかったものと考えられる.

4. むすび

本稿では,腸閉塞・イレウス診断支援のため の腸管抽出手法 [2] に関して,抽出結果からの 病変箇所の見つけやすさに関する評価を行っ た.抽出結果から病変箇所への最短距離の平均 は 19.3 mm であり,大多数の症例で腸管付近 までの抽出が可能であった.しかしながら,腸 管の細い症例では最短距離が長くなった.また, 鋭敏なカーブにおいて腸管区間領域の誤った 分断が多数発生し,病変候補数の平均は23.0個 に上った.今後,このような分断を修正するこ とで,病変候補数を削減する予定である.

謝辞

本研究の一部は, 堀科学芸術振興財団, MEXT/JSPS 科研費 (21K19898, 17H00867, 17K20099, 26108006, 26560255), JSPS 二国間交 流事業, AMED (JP19lk1010036, JP20lk1010036), JST CREST (JPMJCR20D5) によった.

利益相反の有無

なし

文 献

- 急性腹症診療ガイドライン作成委員会:
 急性腹症診療ガイドライン 2015. 医学書院,東京,pp.16.
- [2] Oda H, Hayashi Y, Kitasaka T, et al.: Intestinal region reconstruction of ileus cases from 3D CT images based on graphical representation and its visualization. SPIE Medical Imaging 2021: Computer-Aided Diagnosis, 115971N, 2021.
- [3] Çiçek Ö, Abdulkadir A, Lienkamp S, et al.: 3D U-Net: learning dense volumetric segmentation from sparse annotation. MICCAI 2016, LNCS 9901, 424-432, 2016.
- [4] Beucher S: IMAGE SEGMENTATION AND MATHEMATICAL MORPHOLOGY. https://people.cmm.minesparis.psl.eu/users/be ucher/wtshed.html (2021年6月14日閲覧)

Evaluating Diagnosis Assistance System's Efficacy for

Locating Diseased Parts of Intestinal Obstruction or Ileus

Hirohisa ODA*1, Yuichiro Hayashi*2, Takayuki Kitasaka*3, Yudai Tamada*4, Aitaro Takimoto*1,

Akinari Hinoki^{*1}, Hiroo Uchida^{*1}, Kojiro Suzuki^{*5}, Masahiro Oda^{*6,2}, Kensaku Mori^{*2,7,8}

- *1 Nagoya University Graduate School of Medicine
- *2 Graduate School of Informatics, Nagoya University
- *3 School of Information Science, Aichi Institute of Technology
- *4 Nagoya University School of Medicine
- *5 Department of Radiology, Aichi Medical University Hospital
- *6 Strategy Office, Information & Communications, Nagoya University
- *7 Information Technology Center, Nagoya University
- *8 Research Center for Medical Bigdata, National Institute of Informatics

In this paper, we evaluate the accuracy of our intestine segmentation method for diagnosis assistance of finding the diseased points of the intestinal obstruction or the ileus. In diagnosing those diseases, clinicians find the diseased points by manually tracking the intestines on the CT volumes. The diagnosis assistance system offers segmentation results of the intestine sections enlarged by the difficulty to move contents towards the anus. Endpoints of some connected components among the segmentation results may be locating the diseased point. The system is helpful if those endpoints are close to the diseased points and the numbers of the endpoints (candidates of diseased points) are low. We performed experiments using 104 CT volumes. The mean minimum distance between endpoints and diseased points was 19.3 mm. 23.0 endpoints were produced on average. We confirmed that the system could produce segmentation results to point the diseased points in most CT volumes accurately.

Key words: Small bowel segmentation, 3D visualization, Graphical representation

手術標本および生検標本の H&E 染色組織画像を用いた 畳み込みニューラルネットワークによる肝細胞癌の識別

高木 弦^{*1} 山口 雅浩^{*1}

阿部 時也^{*2} 橋口 明典^{*2} 坂元 亨宇^{*2}

要旨

本研究では、H&E 染色病理標本のデジタルスライド画像に対して畳み込みニューラルネットワーク (CNN) による肝細胞癌と非癌の識別を行った. CNN として EfficientNetB3 を採用し、高解像度の病理画像を分割し て学習させ、その結果を統合する方法について検討するとともに、学習率などについて改善行い精度向上 を図った. 2048 × 2048 画素の手術標本及び生検標本の画像それぞれ約 1,000 枚に対して 5-fold の交差検証を 行った結果、先行研究において特徴量ベースの SVM で約 90%の精度であったデータに対して、正解率約 97% と極めて高い精度を得ることかができた. このとき、生検標本と手術標本の両者の画像を合わせて教 師データとした際に、高精度でロバストな識別器が作成できることが示された. さらに、生検標本と手術 標本の両方を学習させた際に、CNN が両者に共通する特徴を抽出しているのかについても考察を行った.

キーワード: 肝細胞癌, 深層学習, CNN

1. はじめに

病理診断では、人体から採取された病変組 織や細胞から作成された標本を顕微鏡で観察 し診断を行う.スライドガラス上の標本を超 高精細なデジタル画像化し、モニターを用い て診断する方法も普及しており、その画像に 対する人工知能技術の応用も活発に検討され ている.

病理診断で用いる標本のうち生検標本は治 療前に腫瘍の一部を採取し、病変の種類や悪 性度の判定を行うために用いられる.また、 手術により切除された検体による標本(以下、 手術標本)は、病変部を全て取り除けるかを 確認するとともに、予後の予測や述語の治療

*1 東京工業大学工学院情報通信系

〔〒226-0026 神奈川県横浜市緑区長津田 町 4259〕

e-mail: takagi.g.ab@m.titech.ac.jp

*2 慶應義塾大学医学部病理学教室

方針を決めるために必要な診断を用いられる。



図1 標本画像 (左)生検標本 (右)手術標本

図1にH&E 染色された生検標本と手術標本の画像例を示す.生検標本は針による穿刺などの手段で組織を採取するので,標本が非常に小さく,組織の変形や崩れなども生じやすい.反対に手術標本では摘出された大きなブロックから標本を作製するので,大きな標本で組織構造も崩れることなく観察できる.

深層ニューラルネットなどによる機械学習 を用いた診断支援システムは,特に生検標本 において必要性が高いが,生検標本において 膨大かつ品質の良い教師データを集めること は容易ではない。

病理医が病変の特徴を理解するプロセスを 考えると、変形や崩れが少なく大面積で観察 できる手術標本を用いて病変部の基本的な組 織構造を理解した上で、構造の崩れなどを含 む生検標本の観察を行っていると考えられる. このため、生検標本を対象とした深層学習に 基づく識別器においても、手術標本に基づく 教師データを活用することが望まれる.

本研究では生検材料と手術材料から作成さ れた H&E 染色病理標本のデジタルスライド 画像に対して,畳み込みニューラルネットワ ーク(CNN)による肝細胞癌と非癌の識別を行 った.関心領域(ROI)をタイル上に分割して CNN に入力し、得られた結果を統合化するこ とにより高解像度の画像に対する識別を行 う.

はじめに、今回構築した CNN による識別 器の基本性能を評価し、先行研究^[1]の特徴量 ベースの SVM による方法との比較を行った.

また、生検標本と手術標本における特徴量 の差異を検討するために、各々の材料による 標本で学習した場合と、2種類の標本を混合 して学習した場合に対して、識別精度の比較 を行った.そして、生検標本の識別を行う場 合に、手術標本から得られる教師画像データ を使用することの効果を検討するとともに、 その意味に関して考察を行った.

2. データセットと識別器

本研究で用いた画像は肝細胞癌の H&E 染 色病理組織標本をデジタルスライドスキャナ で 20x の倍率で全スライド画像(Whole Slide Image:WSI)化したものである.ここから約 1mm 角の ROI(2048×2048pixel)が切り出され てアノテーションが付与されている.生検と 手術標本における画像枚数・患者数を表1に 示す.非癌部は主に肝細胞癌患者の背景組織 から採取されたものである.アノテーション された 2048×2048pixel の画像を直接 CNN へ

表1 生検・手術標本における画像枚数と患者数

	生検標本	手術標本
画像枚数	1067 枚	1104 枚
	悪性:557 枚	悪性:554 枚
	良性:510 枚	良性:550 枚
患者数	108 人	111 人

の入力とすることは深層学習の実装上困難で 亜あるため,本研究では,元の画像(2048× 2048pixel)を縦横8分割し,256×256pixelの 画像をCNNの入力とした.図2にROIに対す るCNNによる識別器の適用法を示す.ROIに 対する識別結果を得る際は各分割画像の結果 を統合する.そのときの統合方法は図2のよ うに,分割画像にスコアの合計値に基づいて 行った.



図2 分割画像を用いて元の画像を評価する方法

癌か非癌かを学習させた CNN に評価目的 の画像に入力すると,0(非癌)から1(癌)のスコ アが出力される.1 枚の ROI 画像を分割して それぞれ CNN に入力すると,一枚の画像に 対して 64 個のスコアが出力される.この 64 個のスコアを合計し,ある閾値よりも高い場 合に癌と,低い場合に非癌と識別することに よって,元の分割前の画像を評価する.なお 分割画像内に組織が含まれる面積が小さい亜 場合には判定から除外した.この閾値は検証 用データを用いて決定される.

3. CNN の学習

本研究では手術標本と生検標本,その二つ を混ぜた標本の3種類のデータを学習させた. そして,それぞれの学習された学習器に対し て手術標本と生検標本を入力し識別を行った. また,交差検証を行う場合には5交差検証で 行い3splitを学習データ,1splitを検証データ,

表3 ROI 画像に対する識別精度 (それぞれの評価データに対して最大の精度の値を赤,最低の精度を青で着 色している)

評価データ	学習データ	教師データ数	閾値	正答率	感度	特異度	AUC	SVM(正答率)
手術標本	手術標本	47296	22	0.9737	0.9711	0.9764	0.9934	0.9189
	生検標本	47083	56	0.8967	0.8863	0.9072	0.9554	0.6953
	手術標本と生検標本	94379	31	0.9620	0.9477	0.9764	0.9970	0.9050
生検標本	手術標本	47296	23	0.7835	0.7584	0.8098	0.8715	0.8463
	生検標本	47083	31	0.9831	0.9785	0.9882	0.9975	0.8985
	手術標本と生検標本	94379	31	0.9850	0.9749	0.9961	0.9971	0.8994

1split を評価データとして全学習過程を行った.

表 2 CNN のパラメータ

Parameter	値
学習モデル	EfficientNetB3 ^[2]
学習済みモデル	ImageNet
学習率	初期値 0.001 で epoch ごとに
	減衰
Epoch 数	20(学習が進まないと停止)
Batch 数	128
損失関数	交差エントロピー誤差
最適化アルゴリズム	Adam

表 2 は本研究で用いた CNN の各パラメー タである. Python の機械学習ライブラリであ る pytorch の EfficientNetB3-pretrained モデル を使用し,記載したパラメータで学習を行っ た.検証データで学習の進みを管理し,損失 関数が 3epoch の間改善しないと学習が終了 する.使用マシンに TITAN-RTX(24GB)を用 いることにより,全交差検証合わせて 40 分 ほどで学習を終えた.

4. 結果·考察

4.1 識別結果

はじめに、識別対象として手術標本および 生検標本それぞれについて、教師データに手 術標本、生検標本およびそれら両方を用いた 場合の識別精度を評価した.その結果を表 3 に示す.手術標本と生検標本のいずれにおい ても、それぞれ単体で学習した場合と両方を 用いて学習した場合はほぼ同等に高い精度が 得られている.それに対して、手術標本のデ ータを用いて学習し、生検標本を識別した場 合,およびその逆の組み合わせでは精度が低い.これは,同一の病変であっても,標本の 作製方法によって画像の特徴が大きく異なる ためだと考えられる.二つの標本を混ぜて学 習した識別器は,どちらの標本も癌か非癌か を識別できている.これは識別器が両者の共 通した特徴量を抽出しているか,それとも学 習の過程で両者を分けてから特徴量を抽出し ているかの二つのパターンが考えられる.

表の最右列には、従来研究において画像か ら算出した細胞核及び組織構造に関する特徴 量を用いて SVM で識別した際の精度を示し ている.特徴量抽出と SVM による識別器と 比較して、格段に優れた精度が得られている ことが確認できる.

4.2 グレード別の結果

表4 グレード別の結果

	本	F法	先行研究(SVM)		
Edmondson Grade	確率	TP/(TP+FN)	確率	TP/(TP+FN)	
1	0.7160	58/81	0.7750	62/80	
2	0.9860	282/286	0.9303	267/287	
3	1.0000	62/62	1.0000	64/64	
4	1.0000	23/23	0.8636	19/22	
overall	0.9477	525/554	0.9134	496/543	
precision	0.9758	525/538			

表4は,4.1節の識別結果のうち手術標本に ついて,病理医により付与されたグレード (Edmondson Grade)別に表記したものである. SVM を用いた先行研究^[1]の結果も合わせて 示している. SVM を用いた手法では,グレー ド 3,4 といった病理医であれば誤るはずのな い症例におけるエラーがなくなり,信頼性が 高まっているといえる.

<u>4.3 生検標本が少ない場合に手術標本を用い</u>ることの効果

4.1 の結果では生検標本だけを教師データ として学習させた場合でも高い精度が得られ たため、それに手術標本を加えても精度の向 上はほとんど見られなかった.それでは、教 師データとして十分な生検標本の画像が得ら れない場合に、手術標本を教師データとする ことの意味はあるのだろうか.

ここでは、教師データ中における生検標本 の数を変化させた場合に識別精度がどのよう に変化するか、同じデータセットを用いて検 討を行った.なお手術標本のデータ数は一定 とした.



図 3 生検標本の量を変えた時の精度(評価 データは生検標本)

その検討の結果が図3のグラフである.図 3のグラフは、教師データとして混ぜたデー タを用い、生検標本の量を変更した際の精度 を示している.評価データは生検標本のみで ある.縦軸は正答率・感度・特異度の割合を 示しており、横軸は最大47000枚の生検画像 のうち何割を使用したかについて示している. 例えば、biopsy rate が0.2の場合、約9400枚 の生検標本画像と47000枚の手術標本画像を 混ぜて学習を行なっている。ただし、この実 験は1-foldのみで行なっている.

全く生検標本を学習せずに生検標本を評価 すると精度が低いことは表3で示されていた. しかし、図3の結果から生検標本のデータを 少量加えるだけでも正答率が95%を超えるこ とがわかった.生検標本は一度に採取できる 量が手術標本に比べて少なく、生検標本の画 像を大量に集めるのは難しい.しかし、この 結果から少量の生検標本でも大量の手術標本 と混ぜ合わせることで、生検標本を高精度で 評価できる識別器を作成できる可能性が示さ れた.

5. まとめ

本研究では CNN による肝細胞癌画像の識別を行った.画像を分割し学習を行い,統合 を行うことで病理画像のような大きいサイズ の画像を CNN で識別することができた.ま た,手術標本と生検標本を混ぜて学習するこ とにより,両者を識別できるロバストな学習 器を作成することができた.また,手術標本 のデータに生検標本のデータを少量加えて学 習することで、生検標本を 95%以上の精度で 評価する識別器が作成することができた。

謝辞

本研究の一部は、NEDO の助成事業 (JPNP20006)として実施されました。

利益相反の有無

なし

文 献

- [1] Aziz MA, Kanazawa H, Murakami Y, et. al., Enhancing automatic classification of hepatocellular carcinoma images through image masking, tissue changes and trabecular features. J Pathol Inform 6:26, (2015).
- [2] Tan, M. & Le, Q., EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. Proc. 36th International Conference on Machine Learning, PMLR 97:6105-6114, (2019).

Identification of Hepatocellular Carcinoma using CNN with H&E Stained Histological Images of Surgical and Biopsy Specimens

Gen Takagi^{*1}, Masahiro Yamaguchi^{*1}, Tokiya Hashiguchi^{*2}, Akinori Hashiguchi^{*2} Michiie Sakamoto^{*2}

*1 School of Engineering, Tokyo Institute of Technology*2 Department of Pathology, Keio University School of Medicine

In this study, we used a convolutional neural network (CNN) to discriminate hepatocellular carcinoma (HCC) from non-cancerous ones on digital slide images of H&E-stained pathological specimens. EfficientNetB3 was adopted as the CNN, and the method of training the CNN by dividing the high-resolution pathological images into segments and integrating the results was investigated. The learning rate and other factors were improved to improve the accuracy. As a result of 5-fold cross-validation on approximately 1,000 images of surgical specimens and biopsy specimens of 2048×2048 pixels each, we obtained an extremely high accuracy of approximately 97%, compared to the accuracy of approximately 90% obtained by feature-based SVM in previous studies. The results show that a highly accurate and robust discriminator can be created when the images of both biopsy specimens and surgical specimens are combined as the teacher data. We also discussed whether CNNs extract features common to both biopsy specimens and surgical specimens when both specimens are trained.

Key words: Hepatocellualr Carcinoma, CNN, Deep Learning

骨粗鬆症 CT 検診における

コンピュータ診断支援システムの開発

宮崎 優太*1 鈴木 秀宣*1 河田 佳樹*1 仁木 登*2

松本 裕司*3 土田 敬明*3 楠本 昌彦*3 金子 昌弘*4

要旨

骨粗鬆症とは、高齢化に伴い骨の強度が低下して骨折しやすくなる状態を指す.前腕部、大腿骨付近の骨 が特に骨折しやすくなり、寝たきりの大きな原因となっている.骨粗鬆症の予防には早期発見・治療が最 も重要となる.我々の研究グループでは胸部 CT 画像から肺がん・COPD・骨粗鬆症などの多疾患を自動検 出・診断するシステムを開発している.このシステムをスタンドアロンアプリケーションから Web アプリ ケーションに改良し、院内ネットワークから多疾患の診断支援モジュールに効率的にアクセスできる診断 支援環境を提供することを目指している.本報告は、骨粗鬆症の診断支援に焦点を当て、システムの設計・ 実装と結果について述べる.

キーワード:骨粗鬆症,診断支援システム,Webアプリケーション

1. はじめに

現在、日本において人口の急激な高齢化に伴い骨粗鬆症の患者数は年々増加しており、その数は現時点で約 1300 万人と推測されている. 骨粗鬆症では、錐体、前腕骨、大腿骨近位部などの骨折が生じやすく、生活の質や日常生活で行う動作の低下を引き起こす.骨粗鬆症による骨折を起こすと、骨折のリスクが骨折していない人に比べ高くなり、骨折を繰り返し起こしやすくなる.骨折を繰り返すたびに健康状態は悪

*1 徳島大学創成科学研究科理工学専攻 [〒770-0814 徳島市南常三島2丁目1] <u>c612036024@tokushima-u.ac.jp</u>

- *2(株)医用科学研究所
- *3 国立がんセンター中央病院
- *4 東京都予防医学協会

化し,骨折前の状態に戻すことは容易ではなく, 要介護状態や寝たきりの大きな原因となる.大 腿骨近位部骨折の予後に関する報告では,骨折 した人のうち約10%強の患者が1年以内に死亡 するなど,その対策が社会的にも重要な問題に なっている^[1].骨粗鬆症の予防には早期発見・ 早期治療が重要となっている.CT検診では,骨 の3次元形状の解析や,短時間でのCT 画像が



図1 現行の CAD システム

撮影可能となっている. 医師が高精度かつ大容 量の画像を読影することが困難となっている ため, 膨大な画像の読影を支援するためのコン ピュータシステム(CADシステム)が求められて いる. 胸部 CT 検査では, 呼吸器, 骨, 循環器 等が撮影される.我々はこれらの画像情報を活 用し,肺がん・COPD・骨粗鬆症などの多疾患を 自動検出・診断するシステム(CAD システム)を 開発している.このシステムを図1に示す.こ のシステムはスタンドアロンアプリケーショ ンであり、使用において限定される.本研究で は、このシステムを Web アプリケーションに改 良し, 院内ネットワークのどこからでも, 必要 な患者の多疾患の診断支援モジュールにアク セスできるシステムとすることを目標として いる. ここでは、骨粗鬆症の診断支援に焦点を 当て、システムの設計・実装と結果について述 べる.

2. Web アプリケーション型 CAD システムの 設計

ド,ユーザが入力したデータを処理するバック エンドからなる. バックエンドは、ユーザの入 力のインターフェースを担うサーバーと, **DICOM** 画像を処理する CAD エンジン, 患者情 報やユーザ情報を保存するデータベースから なる. フロントエンドは Web ブラウザである. 図 2 に骨粗鬆症の診断支援モジュールを利用 するシナリオのシーケンス図を示す.シーケン ス図は、 プログラムの処理の流れや要素を設計 する際に使用される[2]. システムはブラウザで ユーザからの入力を受けると、その処理がフロ ントエンドで完結する処理であるなら処理結 果をブラウザに2次元表示(4.1)する. バックエ ンドを使用する際, ユーザから入力された DICOM データをサーバーを介し CAD エンジン に3次元処理のリクエスト(5.1)が送信される. CAD エンジンで処理されたデータは再びサー バーを介しブラウザに表示(5.4)される.診断結 果入力・保存(6)する際は、ブラウザで入力され た内容をサーバーを介してデータベースに保 存する.



本システムはユーザが操作するフロントエン

図2Webアプリケーション型CADシステムのシーケンス図



図3Webアプリケーション型CADシステムの画面遷移図

図 3 にシステムの画面遷移図を示す. 画面遷 移図は, 画面の表示される順序や画面同士の関 連性などを図解したものである.

読影医はシステムにログインした後,患者と 疾患を選択する.診断画面が表示されるので読 影医は診断をスムーズに行うことができる.

3. Web アプリケーション型 CAD システムの 実装

フロントエンドは JavaScript のライブラリで ある React.Js^[3]を用いて開発した.バックエンド は C#を用いた. React.Js は UI を開発すること に特化した JavaScript のライブラリである. React.Js はコンポーネント指向といい,Web 上 に表示する構成要素をそれぞれ独立したオブ ジェクトとして扱う概念がある.そのため、保 守性が高くシステムのメンテナンスも容易に 行うことができる.また,オブジェクトを再利 用して別の要素を作成することもできるため, 汎用性が高い.React.Js を用いないWeb システ ムでは,画面が表示されるまでの時間が長い. しかし,React は画面のレンダリングが高速な ためスムーズな操作を実現することが可能と なる^[4].

4. 実装結果

Chrome 上で DICOM の 2 次元処理のテストを 行った. 開発に使用したソフトは Visual Studio Code, 使用した React のバージョンは 6.14.9 で ある.

5. まとめ

サーバーを用いず読み込んだ DICOM データ を 2 次元処理できる機能の開発を行っている. また,多疾患の診断支援が行えるシステムを開 発する.

利益相反の有無

なし

文 献

- [1] 折茂 肇:「骨粗鬆症の予防と治療のガ イドライン 2015 年度版」、ライフサイ エンス出版、東京、2015
- [2] Sinan Si Albir 著:入門 UML オライリー ジャパン
- [3] React 公式ドキュメント https://ja.reactjs.org/
- [4] 穴井宏幸,石井直矢,柴田和祈,三宮肇著:React入門 翔泳社

Development of a computerized diagnosis support system for CT examination of osteoporosis

Yuta MIYAZAKI^{*1}, Hidenobu SUZUKI^{*1}, Yoshiki KAWATA^{*1}, Noboru NIKI^{*2} Yuji MATSUMOTO^{*3}, Takaaki TSUTIDA^{*3}, Masahiko KUSUMOTO^{*3}, Masahiro KANEKO^{*4}

- *1 University of Tokushima
- *2 Medical Science Institute, Ltd.
- *3 National Cancer Center Hospital
- *4 Tokyo Health Service Association

Osteoporosis refers to a condition in which the strength of bones decreases with age, making them more susceptible to fracture. The bones of forearm and proximal femur are particularly prone to fracture and are major cause of bedridden patients. Early detection and treatment is the most important faction in preventing osteoporosis. We are developing a system to automatically detect and diagnose multiple diseases such as lung cancer, COPD, and osteoporosis from chest CT images. We aim to improve this system from stand-alone application to a web application to provide a diagnosis support environment that allows efficient access to diagnostic support modules for multiple diseases from the hospital network. This report focuses on the diagnostic support for osteoporosis, and describes the design, implementation, and result of the system.

Key words: Osteoporosis, Diagnostic support system, Web application

Diagnosis Assisted by Machine Learning to Differentiate Benign

and Malignant Soft Tissue Tumors in MRI

Hasnine HAQUE^{*1,*4}, Toru HIROZANE^{*3}, Masahiro HASHIMOTO^{*1}, Naoki TODA^{*1}, Yuki ARITA^{*1}, Kanta SHONO^{*2}, Yurie KANAUCHI^{*2}, TDQ KHANH^{*1}, Nozomu UETAKE ^{*4}, Masahiro JINZAKI^{*1}

Abstract

Standardized protocol for how MR imaging should be used for patients with suspected soft tissue sarcoma (STS) disease is widely varied among institutions. Image-level detection of malignancy of STS is a challenging task for radiologists and orthopedic surgeons. In this research work, we investigate which MR imaging protocol or protocols contains most malignancy signature of STS. Retrospective Gd-T1W(gadolinium-enhanced), T2W, FS-T1W, DWI, T1W MRI of 80 patients with histologically proven subtypes of STS were examined. We grouped UPS, myxofibrosarcoma, myxoid liposarcoma, leiomyosarcoma, synovial sarcoma, dedifferentiated liposarcoma subtypes as malignant and schwannoma, lipoma, ALT, myxoma, desmoid subtypes as benign. After manual segmentation and preprocessing texture-based radiomics features and deep CNN features from pretrained VGG19 model along with feature reduction methods and machine learning modeling ware used for malignancy detection performance comparison. Only combined imaging features of Gd-T1W and T2W were used due to its most sample size on the cohorts. 14 patient Images scans were used as test dataset to evaluate our claims. Detection accuracies of the model using radiomics feature of Gd-T1W and T2W+Gd-T1W were 86% and 78% respectively. Whereas CNN features of same MR images was found to be 93%. In our investigation Gd-T1W scanned MR image contains most malignancy signature and CNN features of image can most accurately detect STS malignancy. This research can optimize the MR scan protocol for STS disease and the risk of oversight of malignant disease by radiologists and orthopedic surgeons.

Keywords : Machine Learning, Soft Tissue Sarcoma, Magnetic Resonance Image, Malignancy

1. Introduction

Soft Tissue Sarcomas (STS) is at type of cancer that begins in the tissues that connect, support and surround other body structures. STS occur anywhere in your body, but the most common types occur in the arms and legs, and in the trunk wall, the retroperitoneum. In USA overall, 5-year survival rate for sarcoma is 65% (Feb. 2021) and it becomes only 15% when the it turns into metastatic stage [1]. If it is found at an early stage and has not spread from where it started, surgical treatment is often very effective, and many people are cured. However, if the sarcoma has spread to other parts of the body, treatment can usually control the tumor, but it is often incurable. There exist more than 50

^{*1} Department of Radiology, Keio University School of Medicine

^{*2} Faculty of Science and Technology, Keio University

^{*3} Department of Orthopedic Surgery, Keio University School of Medicine

^{*4} GE Healthcare Japan [4-7-127 Asahigaoka, Hino-shi, Tokyo 191] e-mail: hasnine.haque@ge.com Receive: January 31, 2013

subtypes of this cancer and those can be further divided into benign (non-cancerous) and malignant (cancerous) group. Invasive biopsy procedure is the gold standard to confirm the subtype of STS [2]. In standard clinical practice MR imaging scan protocols were commonly used to learn detail of STS size, location and extent and its progression over time. There were many efforts to find non-invasive approaches to classification of STS from multi parametric MR images were invested [3,4]. In MRI examination, there is a trade-off between the time required for imaging and the image quality. Due to time constraints, it is difficult to use all types MR scan protocols with high image quality, and there is a clinical diagnosis need of high-quality imaging with selected MR scan protocols. In this research, we investigate which scan protocol or protocols of MR Imaging contains most malignancy features and can be used to distinguish between benign and malignancy subtypes. With recent of advancement of machine learning technology features present on the images which human eye cannot correlate with the diseases finding. We will use widely used standard "Radiomics" and "Deep learning" techniques to extract radiomics and imaging features for the multi parametric MR scan images often scan in this institution. Finally, select imaging features or radiomic features or combination of both features and produce best performance in detecting the malignancy group of STS. By this, it may be possible to improve diagnostic performance by taking time to scan important protocols for diagnosis.

2. Materials and Methods

The institutional review board of the hospital approved this study and waived the need for informed consent. For this retrospective study the Radiological Information System (RIS) and Pathology Laboratory Information Systems were used to search for patients with biopsy proven 11 subtypes of soft tissue sarcoma (Fig.1a) in the lower extremities who underwent MR imaging from September 2005 to December 2020. Subtypes were divided in malignant and benign groups. In malignant group contains MFS/UPS (myxofibrosarcoma with undifferentiated pleomorphic sarcoma), myxofibrosarcoma, myxoid liposarcoma, leiomyosarcoma, synovial sarcoma, dedifferentiated liposarcoma subtypes and benign group contains schwannoma, lipoma, ALT (atypical lipomatous tumor), myxoma,



Fig. 1 From 80 patient datasets (a) shows the distribution of image volume count for each subtype of STS. (b) shows the image volume counts for each type of MR scan protocol
desmoid subtypes. An orthopedic surgeon with 14 years' experience and a board-certified radiologist with 15 years' experience collected and annotated the retrieved MR data in consensus. We used the MR scanned image datasets which were collected prior to biopsy or any treatment for this study. Total of 80 patient was designated for this research. Among these patents near about 700 image volumes were found with different types of MR scan protocols mainly contrast enhanced Fat sat T1 weighted (GD), T1W, Fat sat T1 weighted (FST1), Fat sat T2 weighted(T2) and Diffusion weighted (DWI). STS regions of all those image volumes were manually segmented and used to develop and test machine learning model to detect malignancy. Image volume count distribution of each subtypes of STS is shown on Fig. 1b.

2.1 Analyzing 3D Radiomics Features of Multiple Scan Protocol Image

Image volumes of each scan protocol along with its manual segmented tumor were processed according to the Fig. 2 to classify the STS tumor. Radiomics features were extracted using an open source python package PyRadiomics [5]. Shape based radiomic features were omitted for simplification as it has less information in classification task. Table 1 shows the 88 standard texture based radiomic features after application 13 different filters generates total of 88×13=1144 features.

•]	ſable	e 1	Font	sizes	and	formats.
-----	-------	-----	------	-------	-----	----------

Image Filter (Total 13)	Radiomics Features (Total 88)
Original	First Order Statistics (19 features)
z-score normalization	Gray Level Cooccurrence Matrix (24 features)
Box Sigma Image Filter (3 level)	Gray Level Run Length Matrix (16 features)
Wavelet decompositions (8)	Gray Level Size Zone Matrix (16 features)
	Gray Level Dependence Matrix (14 features)

Two hybrid feature selection methods were chosen 1. Spearman + Support Vector Machine (SVM) + Random Feature Extraction (RFE) 2. Speraman + LASSO to reduce the dimension of input features to the classifier model. There were five classifiers models used in the pipeline and determine the best classifier model and feature selection method to predict correct prediction. Five standard classifiers ware 1) Decision tree, 2) Naive Bayes, 3) k-nearest neighbors (KNN), 4) Logistic Regression and 5) SVM. Performance of this classifier for individual MR scan protocol image (GD, T2, FST1, DWI, T1, and FST2) where evaluated with 5-fold cross validation with 15% of total datasets as validation data. Receiver operating curves (ROC) for each types of protocol image are shown in Fig.2 with blue color curve as training and green color curve as validation results. This classifier pipeline tells that best prediction performance only using texture based radiomic feature present in which MR protocol image. However, model pipeline did not confirm if there were other imaging features like convolutional neural network (CNN) imaging features or combination of different features of multiple scan protocol image improve the performance or not. Next section explored this opportunity.



Fig. 2 Top shows the classifier model pipeline to predict malignancy from individual MR protocol image volume, bottom shows the performance of the model as ROC and AUC value of respective image protocol type. Blue curve indicate training and green curve indicates the validation plot

2.2 Analyzing Imaging Feature and 2D Radiomic Features from Contrast Enhanced & T2 weighted image

From Fig.2 shows that top three scan protocol images that has texture based radiomics features (RF) and can act as malignancy biomarkers. Top batter is contrast enhanced fat sat T1 weighted image (GD). We tried to investigate an additional CNN feature of the image and features from top performed protocol image (GD and T2) can improve the classification performance. As we learned that scan protocols used in this study was not same for all patient. Therefore, all the patient did not have GD and T2 protocol image. Input image data sets was further filter out and keep only those datasets that contains same patient GD and T2 scanned images. With recent advancement of deep neural network ImageNet [6] model and utilize CNN imaging features to classify natura image. We chose pretrained VGG-19 model as CNN feature extractor model. We used the fully connected feature vectors of 4096 length from intermediate layer just before final layer as CNN imaging feature (DL). Input of VGG-19 is two-dimensional image, therefore, to combine both DL and RF features we propose 2D image input framework to feed all the slices of the image volume both in DL model and RF model pipeline shown in Fig. 3. In this framework one set of input image is formed 2D image slice and two of its neighboring image slices form 3 channel input of the VGG-19 model and for radiomics feature model with same 2D image slice and four of its neighboring image slices forming total of 5 slices and calculate the radiomics feature. From all patient we found 726 set of images. For our investigation we used of such datasets for 10 patients as hold out validation datasets and remaining datasets were used for training the model. We confirmed the malignant-benign image types ware balanced on both training and validation datasets. Same radiomics features selection methods and classifier

models mentioned in previous section were chosen for the investigation. For the DL feature selection technique, we implemented ANOVA (Analysis of Variance) [7]. Fig.3 shows the pipeline of classifier model which was trained to predict either malignancy from RF or GD features from both protocol image and individual protocol image.



Fig. 3 2D input dataset to VGG-19 model for DL feature extraction and Radiomic feature extraction method

For simplification we selected fixed number of significant 150 DL feature vectors and pass to the classifier model. In this pipeline DL features from GD image and concatenation of DL features from both GD and T2 images were fed to the classifier model and validate the model performance. Similar combination of RF features were used to evaluate the performance of RF features. It was found that combination of RF and DL features to the current set of classifier model did not show any improved performance



Fig. 4 Pipeline of malignancy detection classifier model which can get input image from both contrast enhanced protocol(GD) image and T2W image SVC model gives the best performance for DL imaging features and Classification And Regression Tree

(CART) for RF features

3. Results

From the left-hand side of Fig. 5 ROC plot indicated that validation performance of the classifier model was improved when radiomics feature from both GD and T2 image were used. Area under curve AUC value improved from 0.7 to 0.81. Right hand side plot of Fig.5 shows the comparison of classifier model performance when DL features of both imaging protocols were used and when DL features of GD image. Here the performance of the model degraded when DL features of both image type was used as input to the classifier. AUC value was reduced from 0.89 to 0.82 when additional DL features from T2 image were included. DL feature of the GD image time shows the best validation performance. To test the model, we have included additional four new patient datasets to the 10 patient validation datasets, in total 14 datasets. Prediction was done by using the features form images slice which has the maximum area of the



Fig. 5 ROC plots of the classifier for RF feature and DL features of validation datasets of 10 patient

tumor region and its neighboring slices. Some of test image data were shown in Fig.6 along with its STS subtype. Detection accuracies of the model using radiomics feature of GD and GD+T2W were 86% and 78% respectively. Whereas when only DL features of same types of MR images were used the 93%. accuracy was achieved. Among all test images only image data shown in Fig.6e where the model with DL features from GD image predicts wrongly.

4. Conclusions

In this research we found that 3D radiomics features from GD image type gives the best malignancy detection performance flowed by radiomics features from T2W image. We further investigated contribution of detection performance by the DL imaging features of both GD and T2W type with our proposed framework. With that we conclude that DL features of image scanned by contrast enhance fat sat T1 weighted (GD) along would be enough to predict malignancy of soft tissue sarcoma. Thereby contrast enhance fat sat T1 weighted image contains more biomarkers than that of other scan type image for STS malignancy detention.



Fig.6 Shows some test image datasets of contrast enhanced (GD) image slice, T2 weighted image slice and tumor mask. STS subtypes of image and overall malignancy detection accuracy of the classifier from 14 datasets when applied to the slice that has the maximum size of the STS region.

Competing interests

Hasnine Haque and Nozomu Uetake disclose a financial relationship with GE Healthcare Japan Ltd, because they are employee of that company

References

- [1] Rebecca L, Siegel MPH, Kimberly D, et.al.: Cancer statistics, 2020.CA: A Cancer Journal for Clinicians. 70: 7-30, 2020
- [2] Domagoj A, Vodanovich, Peter C: Soft-tissue Sarcomas: Indian Journal of Orthopedics 52(1): 35-44, 2018
- [3] Haithem H, Olfa M, Ezzeddine Z: Deep feature learning for soft tissue sarcoma classification in MR images via transfer learning, Expert Systems with Applications. 120: 116-127, 2019
- [4] Fang Z, Shivani A, Sahar J, et.al : Can MR Imaging Be Used to Predict Tumor Grade in Soft-Tissue Sarcoma? Radiology. 272: 192-201, 2014
- [5] Griethuysen M, Fedorov A, Parmar C, et.al.: Computational Radiomics System to Decode the Radiographic Phenotype. Cancer Research 77(21): 104–107, 2017
- [6] Deng J, Dong W, Socher R, et.al.: Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition, 2009, pp248–255
- [7] Maloof A: On machine learning, ROC analysis, and statistical tests of significance. Object Recognit. Support. by user Interact.
 Serv. Robot. 2: 204–207, 2002

. Biography











Hasnine Haque

Senior Scientist in STO-J Research Division, GE Healthcare Japan. He received Ph.D. (Eng.) from Tokyo Institute of Technology in 2000. He is also a research collaborator with department of Radiology, Keio Univ. School of Medicine. He received poster award (2nd place) at ISMRM (2012) in Interventional MR category. His research fields include machine learning with medical imaging in clinical research and image processing.

Toru Hirozane

He is assistant professor at Department of Orthopedic Surgery, Keio Univ. School of Medicine. His research interests are musculoskeletal tumor and cellular signaling. He is a member of Japanese Orthopedic Association

Masahiro Hashimoto

He is specially appointed assistant professor at Department of Radiology Keio Univ. School of Medicine. His research interests are MR imaging methods, machine learning, and ultrasonography. He is a member of JAMIT and Japan Radiological Society.

Naoki Toda

He is assistant professor at Department of Radiology Keio Univ. School of Medicine. His research interests are MR imaging methods, machine learning, and ultrasonography. He is a member of Japan Radiological Society.

Yuki Arita

He is assistance professor at Department of Radiology Keio Univ. School of Medicine. His research interests are MR imaging methods. He is a member of JAMIT and IEEE

Tran Duy Quoc Khanh

Biography: Graduated General Practitioner specialty at Hanoi Medical University - Vietnam in 2016, he continued his study with a Radiologist specialize course at Bach Mai Hospital - Vietnam. After completing being an exchange student at Medical School - University of Sydney, he is pursuing a PhD degree at Graduate School of Medicine - Keio University, specialized in researching AI for medical imaging diagnosis.

Yurie Kanauchi

Graduated from Keio University in 2020. She is a graduate student at Keio University. Her research field is deep learning of medical images.

Nozomu Uetake

He is General Manager STO-J Research Division, GE Healthcare Japan. His research interests are MR imaging methods. He is a member of JAMIT and IEEE

Jiro Suzuki

He is professor at Department of Radiology Keio Univ. School of Medicine. His research interests are MR imaging methods. He is a member of JAMIT and Japan Radiological Society



Improving Classification Accuracy of Hands' Bone Marrow

Edema by Transfer Learning

Dongping PAN^{*1}, Masahiro ODA^{*2,1}, Kou KATAYAMA^{*3}, Takanobu Okubo^{*3}, and Kensaku MORI^{*1,4,5}

Abstract

We propose a method using transfer learning to improve the classification accuracy of hands' bone marrow edema. Bone marrow edema (BME) is the main symptom of early rheumatoid arthritis (RA). By finding bone marrow edema, early diagnosis and treatment for RA will be possible. Although we can observe BME from Magnetic Resonance Imaging (MRI) scans, the cost of taking MRI images is quite high. To develop an economical solution, we aim to use X-ray images to recognize BME. While previous work verified the feasibility of using CNN to classify BME, we encountered the problem of the shortage of data. To overcome this problem and achieve a better performance, we propose a new method that uses transform learning from the RSNA bone age dataset. We use a pre-trained weight from RSNA bone age regression task. In our experiments, the additional knowledge from the RSNA bone age dataset improved our f-score by 17%.

Keywords : bone marrow edema, rheumatoid arthritis, CNN, classification

1. Introduction

In this paper, we propose a method to help diagnose rheumatism by classifying the bone marrow edema (BME). The BME syndrome is a diagnosis of exclusion that is characterized by pain and increased interstitial fluid within bone marrow without an obvious cause [1]. In recent studies, BME is proved to be closely related to rheumatism [2]. To diagnose rheumatism at an early stage, detecting BME is very important. Although BME can be easily observed by Magnetic Resonance Imaging (MRI), it is much more expensive than other medical imaging methods such as X-ray. In the course of rheumatism treatment, the patient should regularly take MRIs, which can be an economic burden. Furthermore, not every clinic has an MRI imaging device. If there are no alternatives diagnosis methods, patients can only go to those clinics with better equipment. Therefore, diagnosing rheumatism from X-ray images can be very meaningful. Fukae et al. [3] converted clinical information to a 2D image and used a convolutional neural network to classify RA or nonRA. They are pioneers to use a convolutional neural network (CNN) for RA diagnosis.

^{*1} Graduate School of Informatics, Nagoya University [Furou-cho, Chikusa-ku, Nagoya, 464-8601, Japan] e-mail:dpan@ mori.m.is.nagoya-u.ac.jp

^{*2} Information Strategy Office, Information and Communications, Nagoya University

^{*3} Katayama Orthopedic Rheumatology Clinic

^{*4} Information Technology Center, Nagoya University

^{*5} Research Center of Medical Bigdata, National Institute of Informatics







(b) Examples of BME classification dataset

Fig. 1 In (a), there are some examples of RSNA bone age dataset. The left one in (a) is a similar image to our classification dataset. The middle one in (a) is a lot darker. The right one in (a) is a hand from a baby which is very rare in our classification dataset. In (b), there is a typical example of our BME classification dataset. BME classification dataset is our main target. There are images diagnosed as bone marrow edema and normal images in this dataset.

However, they have not used any information from medical images like X-ray images. From X-ray images, swelling in soft tissue can usually be observed from BME cases, which may help diagnosis. In our previous work [4], we used X-ray images for BME's classification and used Grad-cam[5] to interpret our result.

In this work, we focused on solving the problem of the shortage of data. Using transfer learning from other similar datasets, we succeeded to gain more knowledge about hand bone's X-ray images. Benefitting from this knowledge, we got a better classification accuracy on our BME classification dataset.

In the method section, we will introduce the intuitive reason of choosing transfer learning, details of our dataset and details about training. In the experiment section, we will show the experiment details and results. In the discussion part, we will discuss our experiment result.

2. Method

2.1 Overview

We aim to acquire knowledge from RSNA bone age dataset [6] to boost the classification accuracy of our private BME dataset. Because the RSNA bone age dataset is not designed for BME diagnosis, also, it's not designed for classification but for regression. We should share a certain part of the network to transfer knowledge from regression task to classification task. We proposed a method to share the feature extraction part. The only difference between the classification network and regression network is the later fully connected layers. From Fig. 2, we choose ResNet18 [7] as our backbone because a shallow network is better for small datasets. We will use the pre-trained weight from the regression task as the initial weight of our classification task.

2.2 Dataset



Fig. 2 Classification network and regression network share the feature extractor which is a ResNet18 network. There are adaptive average pooling layer and different fully connected layers following the feature extractor. Numbers 512, 2, 1 are the output dimension of adaptive average pooling layer and fully connected layers. For classification network, we use cross entropy loss to optimize our parameters. Cross entropy loss evaluates the difference between prediction distribution and real distribution. Mean squared error is used in our regression network which measure the distance between ground truth and prediction.

We use our private BME dataset of 577 images. 450 of them are non-BME and the rest are BME. BME dataset is prepared for binary classification which is the main task in this paper. To solve data shortage, we also use the public RSNA bone age dataset with more than 12000 images in it. RSNA bone age dataset provides pairs of images and numbers. It's designed for regression tasks. One image should be predicted to one number which is the age of the bone. Age is recorded in the month. The average age in RSNA dataset is 127 months old, while people who suffer from BME are 50 years old on average. Although the hands in the RSNA bone age dataset are much younger than those in the BME dataset, we assume both datasets will require the network to focus on bone gap and bone density characteristics. Because they have common parts, transfer learning is promising to succeed. There are some examples in Fig. 1. Some of these pictures are very similar to our dataset for classification, while others are slightly different. For example, the middle image of Fig. 1 (a) is much darker than our dataset, and the right image of (a) is an image from a baby whose bone gap is significant. We designed a data ablation experiment to test which part of data is beneficial.

2.3 Regression method

The lower part of Fig. 2 is the regression network. It is trained in advance and will provide the initial weight for the classification network. We use ResNet18 as the backbone to train the regression network. The regression network is trained on RSNA bone age dataset. The input is the images of hand X-ray images.

They have ground truth which is their age. The age is the number showing how many months old they are. The ResNet18 extract features from the images and passes the feature vector to the adaptive average pooling layer. The final fully connected layer reduces the dimension to 1, which is the predicted age.

10000 images in 12600 images are used for training. The rest is used for testing. Images are resized to 128128. There is no data augmentation. Adam optimizer and learning rate of 0.00003 are used for optimizing. Training epoch is set to 100 epochs.

When predicting the bone age of an X-ray image, the gap between the bones is an important clue to judge whether the child is young [8]. We assume that utilizing gap information is suitable for our BME classification because swelling in the soft tissues can be observed in our cases. This is visually similar to the gap between bones.

2.4 Classification method

Our proposed method use shared the feature extraction part. The only difference between the classification and regression networks is the later fully connected layers, drawn in Fig. 2. The ResNet18 block consists of a single convolution layer and four residual blocks. Each residual block has two convolutional layers and a shortcut. The shortcut will carry the previous information to later layers, which enable faster training and better performance. Besides the image encoder, we use an adaptive average pooling layer and a fully connected layer to reduce the dimensionality to classify.

119 images, including 93 non-BME images and 26 BME images, are used for testing. 378 images, including 297 non-BME images and 81 BME images, are used for training and evaluation. For data augmentation, Horizontal flips and vertical flips will happen at a probability of 50%. Adam optimizer and learning rate of 10-3 are used for optimizing. Training epoch is set to 200 epochs.

We saved the weight of ResNet18 from the previous regression network. The classification network is the same as the regression network except for the final fully connected layer. In the classification experiment, we update the initial weight of ResNet18 with pretrained weight and randomly initialize the final layer. Then we train the classification network with our private BME dataset.

3. Experiments

To prove that our pretrained weight has absorbed knowledge from RSNA bone age dataset, we evaluated the regression result using mean square error. To choose the most suitable backbone for both our regression network and classification network, we designed the backbone comparison experiment among ResNet18, ResNet50, ResNet101. To get the best result of transfer learning, we designed the experiment to find whether we should freeze certain layers. To confirm which part of RSNA bone age dataset is helpful, we designed the data ablation study.

3.1 Regression task



Fig. 3 Relation between actual age (ground truth) and predicted age from ResNet18. Actual age is drawn in blue line. There are multiple test cases whose ground truth values are the same so that there are multiple red points corresponding to one horizontal coordinate. Red points are predictions from each test case. The vertical coordinate is the prediction value. From this figure, we can find a lot of predictions (red points) are close to the ground truth.

Our goal is to obtain some knowledge from RSNA bone age dataset through this regression task. To prove that the pretrained weight from regression task is meaningful, we evaluated the regression network.

RSNA bone age dataset [6] consist of 12611 X-ray images in jpeg format. The average age is 127 months. The ground truth is made by six separate estimates from pediatric radiologists. The example images are shown in Fig. 1 (a). Every image has a ground truth which is the age in month. We use RSNA bone age dataset to achieve a good initial weight for our later classification task.

As we mentioned before, the network was a compromise because it should work both on our BME classification dataset and RSNA bone age dataset. Therefore, we chose ResNet18 as our backbone. The final fully connected layer reduced the dimension to 1 so that regression task could be completed.

Table 2	Result	of	transfer	learning	using
	ResNet	18 a	nd previo	us method	

	F-score	
ResNet 18	0.63	
ResNet 50	0.62	
ResNet 101	0.59	

	F-score
Without transfer learning	0.63
No freeze	0.80
Freeze first 2 layers	0.82
Freeze first 3 layers	0.74
ModaNet [4]	0.73
(without transfer learning)	

From Fig. 3, we can check our regression result. In most cases, the error could be controlled under 35 months. The average month error is 33.3 months. Although our result was worse than the champion solution from [6], we still got a reasonable regression result. We believe our feature extraction layers have grasped certain characteristics of X-ray images.

3.2 Backbone choice

We assumed that shallower network would be more suitable for our dataset of 577 images. However, we were not sure how complex our network should be. We designed this experiment to compare the performance between ResNet18, RestNet50, ResNet101. We summarized the result in Table 1. In Table 1, we tried different ResNet to classify our BME dataset, we calculated F-score to compare among them. F-score is the harmonic mean of the precision and recall.

3.3 Transfer learning and learning scheme

We stored the weight from the regression network and used them as the initial weight of the classification network. Since we used exactly the same feature extractor, the network structure would be the same. Then we added fully connected layers to our pre-trained ResNet18 and trained the network on the BME classification dataset.

As shown in Table 2, we can find that F-1 score was improved. Since the only difference was the initial weight, transfer learning played an important role in this improvement. We assumed that the shallower layers had learned about corner, gap or lines, so we tried to freeze certain shallower layers to maintain its ability to focus on the characteristics mentioned above. Therefore, we added some comparison experiments about freezing certain layers. The result can also be seen in Table 2. Moreover, there is a result of the previous method [4] in Table 2. Using transfer learning, ResNet18 achieved a better performance than the previous method.

3.4 Data ablation

Because we were not confident that all the 12611 images would do good to our classification, we

Table 3 Result of data ablation. RSNA bone age dataset is split into three parts. Normal image means images similar to our private BME dataset. Baby images means images from baby less than 50 months. Dark images mean dark X-ray images which is difficult for human to watch. Three experiments use different data to complete the regression task and classification task.

	Normal images	Baby images	Dark images	MSE (regression)	F-score (classification)
1	\checkmark	\checkmark	\checkmark	1110	0.80
2	\checkmark		\checkmark	718	0.82
3	\checkmark	\checkmark		1194	0.79

Table 4 Confusion matrix of result obtained when only using normal images and dark images for classification (f-score = 0.82)

Predict

		BME	Non-BME
	BME	20	6
Real	Non-BME	3	90

designed several comparative experiments. There are several kinds of images in this dataset. For example, the middle of Fig. 1(a) is too dark and the right of Fig. 1(b) is a hand from a baby. We aim to learn a discriminator which focuses on gap between bones and other low-level features. Maybe baby's data can be useful. For those too dark images, although they are very difficult to see clearly by human, they may be useful information for convolutional neural network. Thus, we utilized (1) all the data or (2) data removing kids under 50 months or (3) data removing too dark images (judged manually) to check the contribution of each kind of images.

In Table 3, we find which part of the data is important for our transfer learning. In this table, MSE means mean square error which is the mean of the squares of the errors. F-score is the harmonic mean of the precision and recall. From the result, removing young kids led to the best result. We show the confusion matrix of the best case with F-score of 0.82 in Table 4.

4. Discussion

In our previous work [4], we mentioned that initial weight may greatly influenced our final classification accuracy. In this experiment, we believe the pretrained weight will carry the common sense of X-ray images and boost the performance.

We have known that our pretrained weights keep some knowledge about hand bone because it achieved a reasonable regression error in Fig. 3. Using this pretrained weight, our transfer learning boosted the Fscore for our classification task. To find the optimal setting, we designed experiments to explore whether all the data from RSNA bone age dataset was important and whether we should freeze certain layers to maintain the ability of recognizing low-level feature such as bone gap.

The most valuable try in this paper is to utilize regression task to strengthen our classification ability. From our results, we believe this strategy is possible and useful.

5. Conclusions

In previous work, we encountered the problem of data shortage and instability caused by initial weight. Transfer learning will provide an initial weight which absorbed the knowledge from other dataset, intuitively it should be a feasible solution for our trouble. From our experiment, transfer learning boosted our F-score by 17% and the qualitative experiment also proofed the usefulness of transfer learning. Our future work is to make our transfer learning more interpretable. We will focus on visualization method to prove the effectiveness of this method.

Competing interests

None

Acknowledgement

Part of this research was supported by the MEXT/JSPS KAKENHI (26108006, 21K19898).

Reference

- Davis DD, Kane SM: Bone Marrow Edema Syndrome. [Updated 2020 Dec 2]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2021 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK559176/
- [2] Jimenez Boj E, Nöbauer Huhmann I, Hanslik Schnabel B, et al.: Bone erosions and bone marrow edema as defined by magnetic resonance imaging reflect true bone marrow inflammation in rheumatoid arthritis. Arthritis & Rheumatism, 56(4): 1118-1124, 2007.
- [3] Fukae J, Isobe M, Hattori T, et al.: Convolutional neural network for classification of two-dimensional array images generated from clinical information may support diagnosis of rheumatoid arthritis. Scientific reports, 10: 1-7, 2020.
- [4] Pan D, Oda M, Katayama K, et al.: Preliminary Study on Classification of Hands' Bone Marrow Edema Using Xray Images. The 39th Japanese Society of Medical Imaging Technology. 488-493, 2020
- [5] Selvaraju R R, Cogswell M, Das A, et al.: Grad-cam: Visual explanations from deep networks via gradient-based localization. Proceedings of the IEEE international conference on computer vision. 618-626, 2017.
- [6] Halabi SS, Prevedello LM, Kalpathy-Cramer J, et al.: The RSNA Pediatric Bone Age Machine Learning Challenge. Radiology 290(2):498-503, 2018.
- [7] He K, Zhang X, Ren S, et al.:Ddeep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition. 770-778, 2016.
- [8] Westerberg E: AI-based Age Estimation using X-ray Hand Images: A comparison of Object Detection and Deep Learning models. Undergraduate Thesis from Faculty of Computing, Blekinge Institute of Technology, 2020.

肺胞蛋白症 CT 画像における重症度分類器構築の検討

山浦 賢太郎*1, 小田 未来*2, 石井 晴之*2,

北村 信隆*3, 田中 崇裕*3, 中田 光*3, 野中 尋史*1

要旨

肺胞蛋白症では肺胞腔内及び終末気管支内にサーファクタント様物質が異常に貯留することで労作時呼吸 困難等の症状が現れる.診断時の HRCT 画像の主要所見としては、すりガラス影及び小葉間隔壁肥厚等が ある. 画像所見の割に症状が比較的軽微で、約3割いる無症状者でも肺野異常陰影は明瞭に観察される特 徴があり、画像所見は他の臨床パラメーターよりも病態を鋭敏に反映している可能性がある.そこで本研 究では、主として CT 値に基づく重症度分類器の構築を検討した. 肺野を様々な症例及び撮影条件で事前 学習した U-net でボリューム抽出した後、ボクセル処理してボクセルごとに CT 値の一次統計量、ヘシアン 固有値を求め、これらを特徴量として bag of visual words を導出しサポートベクトルマシン(SVM)で分類を 行ったところ精度は 52.5%となった.

キーワード:肺胞蛋白症,HRCT 画像,重症度分類,CT 値の一次統計量,ヘシアン固有値

1. はじめに

肺胞蛋白症とは肺胞腔内及び終末気管支内 にサーファクタント様物質が異常に貯留する ことで、労作時呼吸困難等の症状が現れる希 少肺疾患である.HRCT 画像の画像所見の割 に症状が比較的軽微で、約3割いる無症状者 でも肺野異常陰影は明瞭に観察される特徴が ある.また、本症の主要所見の一つであるす りガラス影は幅広い CT 値をとり、濃度の濃 さと広さは Tokura et al.[1]の報告によれば動 脈血酸素分圧(PaO2)と相関がある.本症の重 症度は表1に示す通り、症状の有無とPaO2

*1 長岡技術科学大学

〔〒940-2188 新潟県長岡市上富岡町 1603-1〕

e-mail: yamaura@konomi.nagaokaut.ac.jp

- *2 杏林大学医学部
- *3 新潟大学医歯学総合病院

で定義されている.

びまん性肺疾患の CT 画像を用いた陰影の 検出及び分類に関する先行研究[2][3][4]では, すりガラス影は一意に定まっていることが多 い.本症のように幅広い CT 値をとる疾患で は,特に濃度の薄いすりガラス影が評価対象 とならず病態を過小評価する恐れもある.

そこで本研究では、主として CT 値に基づ く重症度分類器の構築を検討した.まず、肺 野をボリューム抽出した後、ボクセル処理し、 ボクセルごとに CT 値の一次統計量とヘシア ン固有値を求め bag of visual words を導出し た.その後、SVM を用いて分類を行った.

表1 重症度(DSS)の定義

DSS	症状	PaO2 [Torr]
2	有り	PaO2 ≧ 70
3	不問	$70 > PaO2 \ge 60$
4	不問	60 > PaO2 ≧ 50
5	不問	50 > PaO2

2. 実験

2016年5月から2017年11月までに中田ら の研究グループにより実施された医師主導治 験(Pulmonary Alveolar Proteinosis GM-CSF Inhalation Efficacy Trial in Japan.: PAGE 試験) [5]で測定された60人のHRCT 画像を2回分, 合計で120枚を長岡技術科学大学生命倫理委 員会の承認を得て使用した.多施設における 共同治験のため,撮影条件はシーメンス,東 芝,フィリップス,GE 各社の機種で管電圧 は120kVp,スライス厚は0.5mm,0.625mm, 1mm,1.25mm,2mm,再構成関数は縦隔関数 から肺野・骨関数まで存在する.重症度ごと のHRCT 画像数の内訳は表2の通りである.

表2 重症度(DSS)ごとの内訳

DSS	HRCT 画像数		
2	55		
3	43		
4	18		
5	4		

HRCT 画像を用いた重症度分類器を以下の 通り構築する.

1) bag of visual words の導出

HRCT 画像から Hofmanninger ら[6]の手法 を使って肺野をボリューム抽出した後,バイ キュービック法を用いて画素間隔及びスライ ス厚を 1mm 立方で等方ボクセル化処理した. その後,全肺野を $32 \times 32 \times 32$ でボクセル処理 し,ボクセルごとに CT 値の期待値,標準偏 差,歪度,尖度とヘシアン固有値を求めた. さらに,ヘシアン固有値 $\lambda_1, \lambda_2, \lambda_3$ についてもそ れぞれ期待値,標準偏差,歪度,尖度を求め CT 値の 4 次元の特徴量に結合して 16 次元の 特徴量を作成した.この特徴量について k-means クラスタリングすることで bag of visual words を導出した.

2) SVM を用いた分類器の構築

肺野ボリュームからボクセル処理してボク セルごとに求めた 16 次元の特徴量について, bag of visual words を使って肺野ボリューム ごとの 400 次元のヒストグラム特徴量に直し, 正則化パラメータ C を 2, 直径 gamma を 17 とした RBF(Radial Basis Function, 放射基底関 数)カーネルの SVM で重症度の分類を行う.

3. 実験結果と考察

10 分割交差検証した結果,混同行列は表 3 の通りとなり,精度は 52.5%となった.原因 としては,HRCT 画像の撮影条件がそろって いないために,CT 値のバラつきやノイズの 偏りが生じていることや,重症度間の画像数 の偏りが大きいこと,個人差を考慮していな いことが考えられる.

予測 2 5 3 4 2 44 10 0 Æ 1 3 24 19 0 0 解 4 5 13 0 0 5 0 3 1 0

表3 混同行列

4. まとめ

本研究では PAGE 試験で測定された HRCT 画像を用いて重症度の分類器構築を検討した. 多様な撮影条件の HRCT 画像を用いているこ とや重症度間の画像数のバラつきが大きいこ と等から,精度は 52.5% となった.今後は撮 影条件等も考慮しつつ,GAN を用いた重症度 分類に取り組んでいきたい.

利益相反の有無

なし

文 献

- Tokura S, Akira M, Okuma T et al:A Semiquantitative Computed Tomographic Grading System for Evaluating Therapeutic Response in Pulmonary Alveolar Proteinosis. Ann Am Thorac Soc, 14(9): 1403-1411, 2017
- [2] Chen A, Karwoski R, Gierada D et al: Quantitative CT analysis of diffuse lung disease. Radiographics, 40.1: 28-43, 2020

- Xu R, Hirano Y, Tachibana R et al: A bag-of-features approach to classify six types of pulmonary textures on high-resolution computed tomography.
 IEICE TRANSACTIONS on Information and Systems, 96.4: 845-855, 2013
- [4] Xu R, Cong Z, Ye X et al: Pulmonary textures Classification via a multi-scale attention network. IEEE journal of biomedical and health informatics, 24.7: 2041-2052, 2019
- [5] Tazawa R, Ueda T, Abe M et al: Inhaled GM-CSF for pulmonary alveolar proteinosis. New England Journal of Medicine, 381.10: 923-932, 2019
- [6] Hofmanninger, J, Prayer, F, Pan, J et al: Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem. European Radiology Experimental, 4(1): 1-13, 2020

•

A Study on the Construction of a Severity Classifier for CT Images of

Pulmonary Alveolar Proteinosis

Kentaro YAMAURA^{*1}, Miku ODA^{*2}, Haruyuki Ishii^{*2},

Nobutaka KITAMURA*3, Takahiro TANAKA*3, Koh NAKATA*3, Hirofumi NONAKA*1

*1 Nagaoka University of Technology

*2 Department of Respiratory Medicine, Kyorin University School of Medicine

*3 Niigata University Medical and Dental Hospital

Pulmonary alveolar proteinosis is a disease characterized by abnormal accumulation of surfactant in the alveoli, resulting in symptoms such as dyspnea on exertion. The main findings on HRCT images at the time of diagnosis include ground glass opacities and interlobular septal thickening. In contrast to the imaging findings, the symptoms are relatively mild, and even in asymptomatic patients (about 30%), abnormal shadows in the lung fields are clearly observed. In this study, we investigated the construction of a severity classifier based mainly on CT values. After the lung fields were extracted using U-net, which had been pre-trained for various cases and imaging conditions, voxelization was performed to obtain the first-order statistics of CT values and Hessian eigenvalues for each voxel, and the bag of visual words was derived using these features and classified using a support vector machine (SVM). The accuracy was 52.5%.

Key words: Pulmonary alveolar proteinosis, HRCT images, a severity classifier, the first-order statistics of CT values, Hessian eigenvalues P1-17

歯科パノラマX線画像における分岐を含むSSDと

ヒューリスティクスによる歯列の認識手法

森下 拓水*1 村松 千左子*2 清野 雄多*3 周 向栄*3 髙橋 龍*4

林 達郎*4 西山 航*5 原 武史*3 有地 淑子*6 有地 榮一郎*6

勝又 明敏*5 藤田 広志*3

要旨

本研究の目的は、歯科パノラマX線画像を解析し、読影に必要な情報を自動抽出して歯科医師の診断補助 に寄与することである.その初期段階として、950 症例を用いて画像内の歯牙の検出と歯種の分類を行っ た.Single Shot Multibox Detector (SSD) に分岐を与えたネットワークを提案し、16クラス(上顎と下顎で 区別した中切歯、側切歯、犬歯、第一小臼歯、第二小臼歯、第一大臼歯、第二大臼歯、第三大臼歯)での 検出を行った.また、偽陽性削除、追加分類、分類修正等の後処理を行った結果、32 歯種について、検出 率 99.40%、誤検出数 0.28 個/画像、認識率 97.29% となり、提案手法の有効性が示された.

キーワード:歯科パノラマ X 線画像,深層学習,コンピュータ支援診断, Single Shot Multibox Detector (SSD)

1. はじめに

歯科パノラマX線撮影装置は日本の歯科医 院全体の98%が保有しており、歯科領域にお いて最も普及している撮影装置である.歯科 パノラマX線画像には口内の全ての歯牙につ

*1 岐阜大学大学院自然科学技術研究科知 能理工学専攻

〔〒501-1194 岐阜県岐阜市柳戸 1-1〕
e-mail: tmori@fjt.info.gifu-u.ac.jp
*2 滋賀大学データサイエンス学部
*3 岐阜大学工学部電気電子・情報工学科
*4 メディア株式会社
*5 朝日大学歯学部口腔病態医療学講座歯
科放射線学分野
*6 愛知学院大学歯学部歯科放射線学講座

いての情報と、歯周病などの歯牙以外の病変 の情報が描出されるため、歯科医師が全ての 歯牙に関して正確な画像診断をするには多大 な時間と労力が必要となる.本研究の目的は、 歯科パノラマX線画像から読影に必要な情報 を自動抽出し、歯科医師の診断の補助に寄与 することである.本稿では、歯科パノラマX 線画像内の歯牙の検出と、検出された歯牙の 歯種の分類を行った.

歯科パノラマX線画像における歯列の認識 手法については以下が報告されている[1,2]. 本稿ではいずれの手法とも異なる,多クラス 検出と後処理による歯列の認識を行った.

2. 方法

2.1. 物体検出器

歯種の分類問題において,対象歯牙の周囲 の情報は非常に重要である.物体検出では, 局所的に対象の歯牙とその周辺の特徴を抽出 できるため,多クラス検出が歯列認識に適し ていると考えた.クラス数は16クラス(上顎 と下顎で区別した中切歯,側切歯,犬歯,第 一,第二小臼歯,第一,第二,第三大臼歯) とした.検出器には,Single Shot Multibox Detector (SSD)[3]に分岐を加えた提案モデル を用いた(図1).分岐先で32歯種の存在確 率を予測・学習し SSD の特徴マップに掛け合 わせることで,畳み込み層では得られないグ ローバルな情報の付与を行った.



図1 提案モデル

2.2. 後処理

検出器が予測した 2D-Bounding Box (以下:検出box)に対して,(1)重複削除,(2)左 右分類,(3)分類修正を順番に行う.

(1)では、ある2つの検出boxの一致率(IoU) が設定した閾値を超える場合に確信度の低い 方を削除する. Boxのクラスが同じ場合は0.3 を,異なる場合は0.5を閾値に用いた.

(2)では、中切歯と予測された検出 box を基 準にクラスを左右に分類する.本処理により、 16 クラスを 32 クラス (32 歯種) とした.

(3)では、クラスが重複する場合に確信度が 低い方の検出 box のクラスを修正、もしくは box を削除した.対象の box に隣接する box の確信度が高い方の box を基準とした.

2.3. 試料画像

9 つの施設で撮影された歯科パノラマ X 線 画像からなる 950 症例を用いた. 検出器には, 深層学習により歯科パノラマ X 線画像から歯 列範囲を抽出し上下のパディングと縮小によ って 512×512 pixels とした画像を入力した.

2.4. 評価方法

歯科パノラマX線画像950症例のうち,760 症例を学習データ,190症例をテストデータ として5 foldの交差検証を行う.

検出 box については, 正解 box との IoU が 0.5 以上であれば成功 box とし, それ以外は 誤検出 box とした. また, 認識性能の指標に は,検出率(成功 box の総数/正解の歯牙の総 数), 誤検出数(誤検出 box の総数/症例数), 認識率(クラスが正解と一致する成功 box の 総数/正解の歯牙の総数)を用いる.

3. 結果

3.1. SSD と提案モデルの比較

後処理を施す前の検出器の出力に対する各 指標を表1に示す.いずれも全症例に対する 予測の結果である.

	検出率	誤検出数	認識率		
	[%]	[個/画像]	[%]		
SSD	99.65	6.65	96.89		
提案モデル	99.63	5.81	97.17		

表1SSD と提案モデル(16 歯種)

3.2. 後処理の効果

提案モデルの結果に対する後処理の結果を 表2に示す.ここで,(3)の結果を本研究の最 終出力とする.最終出力例を図2に示す.

表2後処理((2)以降は32歯種)

	検出率	誤検出数	認識率
	[%]	[個/画像]	[%]
後処理前	99.63	5.81	97.17
(1)重複削除	99.46	0.36	97.08
(2)左右分類	99.46	0.36	96.92
(3)分類修正	99.40	0.28	97.29



図2 最終出力例(検出率100%,誤検出数0 個,認識率100%)

4. 考察

SSD と提案モデルを比較して,提案モデル について認識性能の若干の向上がみられた. 提案モデルについて,分岐先から特徴マップ にフィードバックを行う方法などを模索する ことでさらなる認識率の向上が期待される.

後処理(1)では検出率を維持しながら大き く誤検出数を減らすことに成功した.歯牙検 出において,過去の実験ではクラス数と検出 性能はトレードオフであったが,採用する検 出 box の確信度を下げて本処理を施すことで 高い検出性能を有した 16 クラス検出が可能 となった.

後処理(2)は多くの症例で成功しているが, 一部症例で分類失敗が発生し,認識率が低下 した.本処理は検出器の予測精度に依存する ため,検出器の認識性能の向上による改善が 期待される.

また,検出器で32クラスの物体検出を行ったところ,全症例に対する予測の結果(後処 理未施行)で検出率99.46%,誤検出数31.79 個/画像,認識率60.53%となった.主に歯牙 の左右の区別に失敗していたことから,本手 法の有用性が示された.

5. まとめ

SSD に分岐を与えた提案モデルでは, 認 識性能に若干の改善がみられた.また,16ク ラス検出に後処理を組み合わせることで,32 歯種について高い性能を得た.今後は提案モ デルのさらなる検討を行い,歯科医師の画像 診断の補助への貢献を行う.

謝辞

本研究の一部は, JSPS 科研費 JP19K10347, 及びアイテック株式会社から共同研究(プロ ジェクト NO: 501001258)助成を受けた.本 研究を行うにあたり,有益なご助言を頂いた 皆様に感謝の意を表します.

利益相反の有無

無

文 献

- [1] Chung M, Lee J, Park S, et al.:Individual Tooth Detection and Identification from Dental Panoramic X-Ray Images via Point-wise Localization and Distance Regularization. arXiv preprint arXiv:2004.05543, 2020
- [2] Tuzoff D V, Tuzova L N, Bornstein M M, et al.:Tooth detection and numbering in panoramic radiographs using convolutional neural networks. Dentomaxillofacial Radiology:20180051, 2019
- Liu W, Anguelov D, Erhan D, et al.:
 Ssd: Single shot multibox detector. In European conference on computer vision: 21-37, 2016

Tooth recognition by branched SSD and heuristics in dental

panoramic radiographs

Takumi MORISHITA^{*1}, Chisako MURAMATSU^{*2}, Yuta SEINO^{*3}, Xiangrong ZHOU^{*3}, Ryo TAKAHASHI^{*4}, Tatsuro HAYASHI^{*4}, Wataru NISHIYAMA^{*5}, Takeshi HARA^{*3}, Yoshiko ARIJI^{*6}, Eiichiro ARIJI^{*6}, Akitoshi KATSUMATA^{*5}, Hiroshi FUJITA^{*3}

- *1 Department of Intelligence Science and Engineering, Graduate School of Natural Science and Technology, Gifu University
- *2 Faculty of Data Science, Shiga University
- *3 Department of Electrical, Electronic and Computer Engineering, Faculty of Engineering, Gifu University
- *4 Media Co., Ltd
- *5 Department of Oral Radiology, School of Dentistry, Asahi University
- *6 Department of Oral and Maxillofacial Radiology, School of Dentistry, Aichi-Gakuin University

The purpose of this study is to analyze dental panoramic radiographs and contribute to the diagnostic aid of dentists by automatically extracting the information necessary for reading them. As the initial stage, we detected teeth in images and classified their tooth types using 950 cases. We proposed single-shot multibox detector (SSD) networks with branching for 16-class (Center incisor, lateral incisor, canine, first premolar, second premolar, first molar, second molar, third molar, distinguished by upper and lower jaws) detection. After post-processing, including false positive deletion, additional classification, and classification modification, the detection rate was 99.40%, the number of false detections was 0.28 per image, and the recognition rate was 97.29% for 32 tooth types. Therefore, the effectiveness of the proposed method was shown.

Key words: Dental panoramic radiographs, Deep Learning, CAD, Single Shot Multibox Detector

3D-Attention Branch Networks を用いた脳 MRI 画像における

低悪性度グリオーマの 1p/19g 共欠損推定法

田中 大貴^{*1} 檜作 彰良^{*2} 中山 良亚*2

要旨

本研究では、腫瘍領域の特徴マップに重み付けする Attention 機構を伴う 3D-Attention Branch Networks によ り, 脳 MRI (Magnetic Resonance Imaging) 画像の低悪性度グリオーマから、1p/19q 共欠損の有無を推定す る手法を開発する. 実験試料は, The Cancer Imaging Archive で公開されている T2 強調画像 159 症例 (1p/19q 共欠損有り:102 症例, 無し:57 症例)を用いる.提案手法では,まず, Feature Extractor により抽出した 特徴マップに対して Attention 機構を適用することにより,腫瘍領域に重み付けした特徴マップを生成する. そして,重み付けされた特徴マップを用いて, Perception Branch で 1p/19q 共欠損の有無を推定する.提案 手法の平均正答率,感度,特異度,Receiver Operating Characteristic 曲線下の面積は,それぞれ 76.7%, 81.4%, 68.4%, 0.804 で, Attention 機構なしの 3D-Convolutional Neural Networks (68.6%, 75.5%, 56.1%, 0.736; p < 0.014) より有意に高い結果が得られ、その有用性が示唆された.

キーワード: 3D-Attention Branch Networks, 1p/19q 共欠損, 脳 MRI 画像, 低悪性度グリオーマ

1. はじめに

1 番染色体短腕と 19 番染色体長腕の共欠損 (1p/19q 共欠損)を伴う低悪性度グリオーマ(LGG: Low-Grade Glioma) は, PVC 療法が著効を示し, 予 後を大きく改善することが明らかになっている[1]. したがって、1p/19q 共欠損を診断することは、患者 の治療方針を決定する上で重要である.1p/19q共欠 損は,遺伝子解析により診断されるが,その検査は 侵襲的で、多くの時間と費用を要する問題がある.

この問題の解決策の一つとして、畳み込みニュ ーラルネットワーク (CNN: Convolutional Neural Networks) を用いて, 脳 MRI (Magnetic Resonance Imaging) 画像上の LGG から, 1p/19q 共欠損の有無 を推定する手法が提案されている[2]. この手法で は、腫瘍全体を含む関心領域(ROI: Region of Interest) と対応する教師信号(1p/19q 共欠損有り/無し)の 関係を CNN に学習させる.しかし, ROI には, 腫 瘍だけでなく、他の組織も含まれるため、CNN が

*1 立命館大学大学院 理工学研究科 〔〒525-8577 草津市野路東 1-1-1〕 e-mail: ri0082vh@ed.ritsumei.ac.jp *2 立命館大学理工学部電子情報工学科 腫瘍領域だけに着目した学習を行っていない可能 性がある.

本研究では, 腫瘍領域の特徴マップに重み付け する Attention 機構を伴う 3D-Attention Branch Networks (3D-ABN) により, 脳 MRI 画像の LGG から、1p/19g 共欠損を推定する手法を開発する.

2. 方法

1) 実験試料

実験試料は, The Cancer Imaging Archive の LGG-1p19qDeletion データベースに含まれる脳 MRI 画像 159 症例(1p/19 共欠損有り:102 症例, 無し:57 症 例) である. これらは, 159 患者 (年齢: 13-84歳, 平均年齢: 42歳)から取得された T2 強調画像で構 成される.本研究では、線形補間法により、T2強 調画像のボクセルサイズを1mmに正規化した. そ して, 腫瘍全体を含む ROI を手動で切り出し, 3D-ABN の入力として用いた.

2) 3D-ABN による 1p/19q 共欠損の推定

図1に、3D-ABN のネットワーク構成を示す. 3D-ABN は, Feature Extractor, Attention Branch, Perception Branch で構成される. Feature Extractor は 3次元畳み込み層,バッチ正規化層, ReLU (Rectified Linear Unit)活性化関数,最大プーリング層からな る 4 つの Convolutional Block を持つ. Attention Branch は、3 次元畳み込み層、バッチ正規化層、 ReLU 活性化関数, Sigmoid 関数で構成される. Attention Branch 1 では, Convolutional Block 2 が出 力した特徴マップと腫瘍のマスク画像の関係を学 習することにより,注目すべき領域を示す Attention Map を生成する. そして, Convolutional Block 2の 特徴マップに, Attention Branch 1 の Attention Map を用いた Attention 機構を適用することにより, 腫 瘍領域を重み付けした特徴マップを生成する. Attention Branch 2,3 では, Convolutional Block 3,4の 特徴マップからそれぞれ Attention Map を生成し, それらの特徴マップに対して Attention 機構を適用 する. 最後に, 生成した特徴マップを Perception Branch に入力し、1p/19q 共欠損を推定する. Perception Branch は 3 つの Sub-network を持ち, 各 Sub-network は GAP 層, 全結合層, Dropout 層で構 成される.

3D-ABN の学習と評価には, *k*-分割交差検証法を 用い, *k*の値は3に設定した.

3. 結果と考察

図 2 に,提案手法によって生成された Attention Map を示す. Attention Map が, ROI の中心にある 腫瘍領域に注目していることが確認できる.提案 手法の平均正答率,感度,特異度,Receiver Operating Characteristic 曲線下の面積は,それぞれ 76.7%, 81.4%, 68.4%, 0.804 で, Attention 機構なしの 3D-CNN (68.6%, 75.5%, 56.1%, 0.736; p < 0.014) よ り高い結果が得られた. 腫瘍領域に注目した Attention Map を生成できたことにより,腫瘍領域 の特徴マップに重み付けする Attention 機構が有効 に働き,推定精度が向上したと考えられる.

4. まとめ

本研究では、Attention 機構を伴う 3D-ABN により、 脳 MRI 画像上の LGG から 1p/19q 共欠損の有 無を推定する手法を開発した. 提案手法は Attention 機構なしの従来の CNN より高い推定精度を示した.



図1 3D-ABN のネットワーク構成



謝辞

本研究はJSPS 科研費 19K20719 の助成を受けた.

利益相反の有無

なし

文 献

- Cairneross G, Wang M, Shaw E et al: Phase III trial of chemoradiotherapy for anaplastic oligodendroglioma: long-term results of RTOG 9402. J Clin Oncol, vol.31, no.3, 337-343, 2013.
- [2] Decuyper M, et al: Automated MRI based pipeline for segmentation and prediction of grade, IDH mutation and 1p19q co-deletion in glioma. Computerized Medical Imaging and Graphics, 88, 101831, 2021.

Computerized Method for Predicting 1p/19q Codeletion of Low-Grade Glioma in Brain MRI images Using 3D-Attention Branch Networks

Daiki TANAKA*1, Akiyoshi HIZUKURI*2, Ryohei NAKAYAMA*2

*1 Graduate School of Science and Engineering, Ritsumeikan University*2 Department of Electronic and Computer Engineering, Ritsumeikan University

In this study, we developed a computerized prediction method for the low-grade gliomas (LGG) with 1p/19q codeletion using 3D-attention branch networks (3D-ABN) with an attention mechanism that weight feature maps of a tumor region. Our database consisted of brain T2-weighted MRI images obtained in 159 patients (102 LGGs with 1p/19q codeletion and 57 LGGs without it) from The Cancer Imaging Archive. The 3D-ABN was constructed from a feature extractor, an attention branch, and a perception branch. The feature extractor generated feature maps. Using these feature maps, the attention branch generated attention maps for weighting the tumor region. The attention maps were then applied to the feature maps in the attention mechanisms. Finally, the perception branch evaluated the likelihood of LGG with 1p/19q codeletion using the feature maps. The prediction accuracy, the sensitivity, the specificity, and the area under the receiver operating characteristic curve with the proposed 3D-ABN were 76.7%, 81.4%, 68.4% and 0.804, respectively. All evaluation indices for the proposed 3D-ABN were greater than those for the 3D-convolutional neural networks without the attention mechanism (68.6%, 75.5%, 56.1%, and 0.736; p < 0.014). The proposed 3D-ABN was shown to have a high prediction performance for 1p/19q codeletion.

Key words: 3D-Attention Branch Networks, 1p/19q codeletion, brain MRI image, low-grade gliomas

Radiomics Prediction for Progression in NSCLC Patients Treated

with Stereotactic Body Radiation Therapy

Takumi Kodama^{*1}, Hidetaka Arimura^{*2}, Kenta Ninomiya^{*1}, Yuko Shirakawa^{*3}, Tadamasa Yoshitake^{*2}, Yoshiyuki Shioyama^{*4}

Abstract

The treatment options for patients with stage I lung cancer are surgery and stereotactic body radiation therapy (SBRT), because their outcomes were almost the same according to a survey using a Japanese multi-institutional database (Onishi et al, 2011). Therefore, it is important to choose the more appropriate treatment for each patient based on pretreatment prediction of patients' prognoses. This study aimed to explore a radiomics prediction for progression using planning computed tomography (CT) images in non-small cell lung cancer (NSCLC) patients treated with SBRT. SBRT treatment planning CT images of 125 NSCLC patients with stage I lung cancer were selected for this study. A total of 486 radiomics features were calculated from original and wavelet decomposition images using histogram and texture analysis. Patients were divided into 88 training cases and 37 test cases, and time-to-progressions (TTPs) for the patients were analyzed by a Kaplan-Meier (KM) method. A signature was constructed with 8 significant features selected according to the frequency of feature selection and the absolute values of the regression coefficients based on Cox-net algorithm. Those patients were stratified into high-risk and low-risk groups using the medians of radiomics scores in training cases. Radiomics scores were derived from the signature feature values and their corresponding coefficients. The prognostic predictability was evaluated using statistically significant differences (p-values, log-rank test) between KM curves of high- and low-risk groups. The p-values between the KM curves of the two groups were 5.98×10^{-5} for training cases and 1.95×10-2 for test cases. This study demonstrated that the high- and low-risk of progression after SBRT in NSCLC patients can be predictable by using image features derived from pretreatment planning CT images.

Keywords : Lung cancer, Time to progression, Radiomics, Cox-net

1. Introduction

Lung cancer is one of the most common cancer around the world, with almost 2.2 million new cases (11.4 %) and 1.8 million new deaths (18.0 %) in 2020 [1]. Lung cancer has two main types that are small-cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC), and NSCLC accounts for approximately 85 % of all [2]. NSCLC is divided into adenocarcinoma, squamous cell carcinoma (SCC), and large cell carcinoma. In Japan, the standard treatment option for stage I NSCLC patients is surgery, and radiation therapy is applied to inoperable patients [3]. Onishi et al. investigated treatment outcomes for SBRT in operable patients with stage I NSCLC patients according to a survey using a Japanese

^{*1} Department of Health Sciences, Graduate School of Medical Sciences, Kyushu University [3-1-1 Maidashi, Higashi-ku, Fukuoka, 812-8582, Japan] e-mail: kodama.takumi.987@s.kyushu-u.ac.jp

^{*2} Department of Health Sciences, Faculty of Medical Sciences, Kyushu University

^{*3} National Hospital Organization Kyushu Cancer Center

^{*4} SAGA Heavy Ion Medical Accelerator in Tosu

multi-institutional database, and they showed the outcomes from surgery and SBRT were almost the same [4]. Hence, pretreatment prediction of patient's prognosis is important to choose the appropriate treatment option.

This study aimed to explore a radiomics prediction for progression using planning computed tomography (CT) images in NSCLC patients treated with SBRT.

2. Materials and methods

1) Clinical cases

Our study employed pretreatment planning CT images of 125 patients with NSCLC (male: 90, female: 35, age: 60-91) treated with SBRT at Kyushu University Hospital. Patients were categorized stage IA (n = 77; n: number of patients) and IB (n = 48). TNM classification was T1a-T2a/N0/M0, and the median of TTP was 29 months. The dimension of CT images was 512×512 pixels, pixel sizes were 0.98 mm, 0.88 mm, 0.84 mm, and 0.78 mm, and slice thicknesses were 2.0 mm and 3.0 mm.



Fig. 1 An overall scheme of this study

2) Overall procedure

Figure 1 shows the overall scheme of the proposed procedure. CT images were re-quantized to 8 bits and transformed into isotropic voxel images (isovoxel size: 0.98 mm) using cubic and shape-based interpolation, and the images were cropped with the gross tumor volume (GTV) used for treatment. The patients (n = 125) were divided into a training dataset (n = 88) and test dataset (n = 37). Radiomics features were extracted from original and three-dimensional wavelet

decomposition images analyzed by histogram and texture analysis. A signature was constructed with 8 features selected by a Cox-net algorithm, and a radiomics score was calculated from the signature. The signature was defined as a set of significant features. Patients were divided into high-risk and low-risk groups using the radiomics score, and the proposed approach was assessed with p-values between Kaplan-Meier curves of the two groups and concordance index (c-index).

3) Feature extraction

Three-dimensional wavelet decomposition was performed by applying low-pass filter (L) and high-pass filter (H) in the x-, y-, and z-directions [5]. The eight wavelet decomposition filters consisted of a combination of two filters in each direction, and radiomics features were derived from 9 types of images (original and 8 wavelet decomposition images). 486 features were calculated from 9 types of images based on histograms, gray-level co-occurrence matrices (GLCM), gray-level run-length matrices (GLRLM), gray-level size-zone matrices (GLSZM) and neighbor gray-tone difference matrices (NGTDM) [6, 7].

4) Signature construction

Signature was constructed by 8 selected significant features ($< n_t/10$; n_t : number of patients in training cases [8]) from Cox-net algorithm in the training cases. Significant features were selected based on more frequently than in 90 % of 100 iterations (over 90 times in 100 iterations) of Cox-net. If more than 8 features were selected over 90 %, then the 8 features with the greatest sums of the absolute coefficients in 100 iterations were selected [9].

In the Cox-net, the Cox proportional hazard model was optimized by maximizing the partial likelihood of the model using the L1 and L2 norm penalties (elastic net penalty). The Cox proportional hazard model $h_i(t|\mathbf{x}_i)$ for a patient *i* at a time *t* can be expressed by

$$h_i(t|\mathbf{x}_i) = h_0(t) \exp(\boldsymbol{\beta}^T \mathbf{x}_i), \qquad (1)$$

where $h_0(t)$ is baseline hazard function, β is a coefficient vector, and x is a feature vector. The Lagrangian formulation is constructed as follow

$$\widehat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\operatorname{argmax}} [kl(\boldsymbol{\beta}) - \lambda P_{\alpha}(\boldsymbol{\beta})], \qquad (2)$$

where

$$l(\boldsymbol{\beta}) = \sum_{i=1}^{n} d_i \left[\boldsymbol{\beta}^T \boldsymbol{x}_i - \ln \left\{ \sum_{t_j \ge t_i} \exp(\boldsymbol{\beta}^T \boldsymbol{x}_j) \right\} \right], \tag{3}$$

$$d_i = \begin{cases} 1 & (\text{event occured at } t_i) \\ 0 & (\text{censored at } t_i) \end{cases}$$
(4)

$$P_{\alpha}(\boldsymbol{\beta}) = \alpha \|\boldsymbol{\beta}\|_{1} + \frac{1-\alpha}{2} \|\boldsymbol{\beta}\|_{2}^{2}.$$
(5)

and *i*, *j* are the patient number, t_i, t_j ($t_i \le t_j$) are the survival time of a patient *i* and *j*, *k* is the scaling factor, λ is

the Lagrangian multiplier, and α ($0 \le \alpha \le 1$) is a blending parameter for adjusting the impact of the LASSO and ridgeregression penalties on an overall regularization [10].

5) Calculation of radiomics score and stratification into high-risk and low-risk groups

Radiomics score was calculated as follow

$$Radiomics\ score = \sum_{l=1}^{m} \beta_l \times Feature_l\,,\tag{6}$$

where l is the feature number, m is the number of significant features, β_l is the coefficient value from Cox proportional hazard model, and *Feature*_l is the value of significant feature.

Patients in both training and test cases were stratified into high-risk and low-risk groups by the median radiomics score of training cases.

6) Evaluation method

The Kaplan-Meier curves for time to progression (TTP) were used to evaluate the high-risk and low-risk groups. P-values of the survival curves were calculated from the log-rank test, and the statistically significant differences (p < 0.05) between two groups were validated.

3. Results and discussion

Figure 2 shows frequencies of radiomics features with non-zero coefficients optimized with the frequency of selection times in the Cox-net algorithm, and eight features were selected as signature (left side of orange line). Figure 3 depicts the plot of response score and the c-index. Response score is the minus value of radiomics score, and the c-index in training cases was 0.814 and in test cases was 0.721. Figure 4 shows Kaplan-Meier curves for high-risk and low-risk groups, and the p-values were 5.99×10^{-5} in training cases and 1.95×10^{-5} in test cases. There were statistically significant differences (p < 0.05) for both.



Fig. 2 Bar graph of the frequency of selection times for features chosen with the Cox-net in 100 times iteration. Hist: Histogram, Org: Original image, H, L: High (H) and Low (L) pass filter for x-, y-, and z-direction in wavelet decomposition



Fig. 3 Scatter plots of a response score and time to progression for a training and test datasets.



Fig. 4 Kaplan-Meier curves of high- and low-risk groups for a training and test datasets.

In this study, radiomics features were calculated from cropped CT images with GTV used for SBRT. This approach did not include surroundings of GTVs, and the marginal tumor information may have prognostic potential.

4. Conclusions

This study showed a potential that radiomics features calculated from pre-treatment planning CT images can be used to predict high risk and low risk of cancer progression after SBRT in NSCLC patients.

Competing interests

The authors do not have any competing interests to disclose

Acknowledgement

This study was supported by a grant from JSPS KAKENHI Grant Number 20K08084. The authors are grateful to all members in Arimura laboratory, whose valuable comments for this study.

References

- [1] Sung H, Ferlay J, Siegel RL, et al.: Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. A Cancer Journal for Clinicians71: 209-249, 2021.
- [2] Tan WW, Huq S, Windle ML, et al.: Non-Small Cell Lung Cancer (NSCLC). Medscape: <u>https://emedicine.medscape.com/article/279960-overview#a7</u> [accessed 21.06.09]
- [3] Takiguchi Y, Okumura M, Hotta K, et al.: Guidelines for the Treatment of Lung Cancer 2020. The Japan Lung Cancer Society. <u>https://www.haigan.gr.jp/modules/guideline/</u> [accessed 21.06.05]
- [4] Onishi H, Shirato H, Nagata Y, et al.: Stereotactic Body Radiotherapy (SBRT) for Operable Stage I Non-Small-Cell Lung Cancer: Can SBRT Be Comparable to Surgery?. IJROBP. 81(5):1352-1358, 2011.
- [5] Chaddad A, Daniel P, Niazi T.: Radiomics Evaluation of Histological Heterogeneity Using Multiscale Textures Derived From 3D Wavelet Transformation of Multispectral Images. Frontiers in Oncology 2018. <u>https://doi.org/10.3389/fonc.2018.00096</u>
- [6] Haralick RM, Shanmugam K, Dinstein I: Textural features for image classification. IEEE Trans Syst Man Cybern. 3: 610-621, 1973.
- [7] Galloway MM: Texture analysis using gray level run lengths. Comput Graph Image Process. 4: 172-179, 1975.
- [8] Arimura H, Soufi M, Ninomiya K, et al.: Potentials of radiomics cancer diagnosis and treatment in comparison with computeraided diagnosis. Radiol Phys Technol. 11:365-74, 2018.
- [9] Ninomiya K, Arimura H.: Homological radiomics analysis for prognostic prediction in lung cancer patients. Physica Medica.69:90-100, 2020.
- [10] Simon N, Friedman H, Hastie T, el al.: Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descend. J Stat Softw. 39(5): 1-13, 2011.

3D CNN を用いた小児脳 CT 画像からの脳成長年齢推定研究

森田 蓮*1 安藤 沙耶*2 藤田 大輔*1 新居 学*1

安藤 久美子*3 石藏 礼一*3 小橋 昌司*1

要旨

脳画像を用いて小児の脳疾患を診断する指標の一つとして,未発達や未熟児などの脳の正常な発達の進行 が評価される。しかし,脳の発達度合いを定量的に推定する方法はなく,現状では医師の経験に基づいて 診断が行われている.そのため,読影可能な医師の不足,定量性の欠如が問題である.本研究では,小児 の脳 CT 画像から脳の発達年齢を予測する手法を提案する.この手法では,CT 画像から頭蓋領域を抽出し, 姿勢と位置の補正を行う.本研究では、3次元畳み込みニューラルネットワーク(3D CNN)を用いて CT 画像から特徴を抽出し,全結合層で脳の発達年齢を予測する新しいネットワークモデルを提案した.この モデルの性能を、0歳から3歳までの脳神経医学的に異常のない小児 60人を用いて評価した.予測年齢と 患者の実年齢の間の平均平方根誤差は7.80(月)で、相関係数は 0.801 であった.

キーワード: 3D CNN, 小児脳, CT 画像, 脳成長年齢, 年齢推定

1. はじめに

小児医療では、脳疾患の診断に CT (Computed Tomography)や MRI (Magnetic Resonance Imaging)がよく用いられる.例え ば、脳疾患の一つである自閉症は、明確な臨 床症状が現れる前の生後1カ月から2歳の間 に小脳に過形成がみられる[1]など、脳画像診 断法の一つとして、未発達もしくは早熟とい った脳の正常成長が評価される.しかし、脳 の成長度合いを定量的に求める方法が明らか でなく、熟練医による経験に基づく診断であ る.そのため、読影可能な医師の不足、定量 性の欠如が問題である.

成人脳の解析には,主に白質,灰白質の組 織コントラストが高い MR 画像が用いられて おり,成人脳 MR 画像を用いた年齢推定に関

*1 兵庫県立大学 大学院 工学研究科 〔〒671-2280 兵庫県姫路市書写 2167〕 *2 兵庫県立尼崎総合医療センター *3 神戸市立医療センター中央市民病院 するいくつかの研究が行われている[2][3][4]. Lao らは, 脳画像から特徴ベクトルを抽出し, サポートベクターマシン (SVM)を用いて予 測することで, 脳年齢を推定する手法を提案 した[2]. Wang らは, 隠れマルコフモデル (HMM)を用いた年齢推定法を提案している [3]. Kondo らの提案した方法は, 統計的パラ メトリックマッピング 2 (SPM2)を用いて脳 領域を解剖学的に意味のある 90 の領域に分

割し,各領域で得られた局所特徴量を用いて 年齢を推定するものである[4].これらの研究 はいずれも,白質,灰白質,脳脊髄液の領域 を識別し,体積を正確に算出する前処理が必 要である.しかし,脳実質間のコントラスト が低い小児脳 CT 画像からそれらを識別して 解析することは困難であり,画像特性が大き く異なる小児脳 CT 画像を対象とした先行研 究はない.

いくつかの研究で, CNN(Convolutional Neural Network)を用いた脳年齢の推定方法が 報告されている[5][6]. 文献[5]では,成人 MRI の脳画像から 2D CNN を用いて年齢推定を行 っている. 文献[6]では, 3D 成人 MRI 画像デ ータから 3D CNN を用いて年齢推定を行って いる. これらの CNN ベースの手法は, WM, GM, CSF の識別を必要としないため, 髄質 の識別が困難な小児脳 CT 画像への適用が期 待されるが, これまで報告されていない.

本論文では、CNN を用いて小児脳 CT 画像 から脳の発達年齢を推定する新しい手法を提 案する.小児期の脳の成長の進み具合を定量 的に評価することで,成長遅延や未熟児を識 別することを目的としている.提案手法は 2 つのステップで構成される.まず,3次元 CT 画像から頭蓋領域を抽出する.CT 検査中に 静止することが難しい小児脳 CT 画像には、 大人の指や補助器具、子供の手が写っている ものが散見される.これらの不要領域を取り 除くことで,解析領域を頭蓋内に限定する. 次に,抽出された頭蓋領域から CNN を用い て脳成長年齢を推定する.

2. 使用データ

本研究では、神戸市立医療センター中央市 民病院に来院した0から3歳の小児計60例を 対象とした.本研究の実施は、同病院の倫理 委員会により承認された.

各被験者を撮影した 3 次元 CT 画像と修正 月齢を使用した. 図1に撮影された CT 画像 の一例を示す.

3. 提案手法

 本研究における脳成長年齢の定義 脳が正常に成長しているときの被験者の年 齢を脳成長年齢と定義する.すなわち成長遅



図1 CT 画像例 (0 歳). A; anterior, P; posterior, L; left, R; right.

滞においては実年齢より小さい年齢が,早期 成長においては実年齢より大きい年齢が推定 される.推定された脳成長年齢を実年齢と比 較することで,定量的に成長を評価すること が可能になり,診断の質向上に貢献する.

2)小児脳 CT 画像の自動正規化法の提案 各被験者において, CT 画像中の脳の位置, 姿勢は多様である.さらに,小児頭部固定の ため,図1に示すように補助者の手や補助器 具などが多く含まれる.提案法では,CNN に 入力される画像信号値から直接脳年齢を推定 するため,これらの要因が推定精度を低下さ せると考えられる.そこで,本研究ではこれ ら頭部領域外を,しきい値処理とモルフォロ ジー処理により削除する.また,姿勢をアキ シャル画像中の頭部輪郭に対する主成分分析 による主軸を求めることで,位置を頭部領域 の重心により校正する.また,ボクセルサイ ズを isovoxel 化する.

3) 3D CNN による脳成長年齢推定

本研究では2)で正規化した3次元脳画像から脳年齢を予測する回帰問題として扱う.回帰の出力を0から1とするために、0から47ヶ月(=3歳11ヶ月)を0から1の値に線形変換する.表1と図2に、本研究で用いたネットワーク構造を示す.フィルターの数は、最初の Conv3D では8とし、MaxPooling 層毎

表1 3D CNN モデルのネットワーク構造

÷ •		
	Layer	Output shape
	Input	218, 218, 218, 1
	3D Convolution+ReLU	216, 216, 216, 8
Dlash	3D Convolution+ReLU	214 214 214 9
Block	+Batch Normalization	214, 214, 214, 8
	Max Pooling	107, 107, 107, 8
	$Block \times 4$	2, 2, 2, 128
Fully Connected		1024
Fully Connected		6
	Output	1



図2 脳成長年齢推定モデル構造





(b) 抽出後 図 3 頭蓋領域抽出結果.



(b)正規化後 図4 CT 画像の位置姿勢正規化結果

に2倍にして特徴を抽出した後,2つの全結 合層で,脳の発達年齢を予測する.予測され た値を,0から47ヶ月の値に逆変換すること で脳成長年齢を算出する.

4. 実験結果

小児脳 CT 画像から頭蓋領域を抽出した結 果を図3に示し、自動正規化法の実験結果の 一例を図4に示す.同結果より正常に補助者 の手や補助器具を除外でき、位置、姿勢を正 規化し脳実質を明瞭に描写できていることを 確認した.

60 名を学習 45 名, 評価 15 名に分割し, 3D CNN モデルを学習した, 汎化性能の検証には 評価データを検証データとして用い, モデル の更新に学習データを用いた. 学習モデルに は, 直近のエポックで得られた学習モデルを 用いた. 図 5 にエポック毎の二乗平均誤差 (RMSE)の推移を示した。



図 5 学習曲線(青実線と赤実線はそれぞれ学 習データと検証データの各エポックにおける RMSEの推移)



図 6 学習データと評価データに対する脳成長 年齢の予測結果

本提案法による脳成長年齢推定結果を横軸 に真値,縦軸に予測値として,それぞれ月齢 を算出した値を用いて示したものを図6に示 す.学習データと評価データの真値と予測値 の相関係数(r)はそれぞれ,0.954 と0.801 であ った.また学習データと評価データの真値と 予測値の RMSE はそれぞれ 4.51 月と7.80 月 であった.

5. まとめ

本論文では、小児脳 CT 画像を用いた脳成 長年齢推定手法を提案した.提案手法は、小 児の脳の発育遅滞もしくは早期成長を定量的 に診断することを支援するものである.提案 手法は、頭蓋領域の抽出、正規化、3D CNN をもちいた脳成長年齢予測からなる.実験結 果より評価データに対して相関係数 r=0.801、 RMSE=7.80月で約8ヶ月の誤差で脳年齢推定 が可能であることを初めて明らかにした.急 速に脳が発達する小児期において8ヶ月の誤 差が生じていることは,現状の精度は不十分 であると考える.この精度を通常成長におい て発生する成長誤差の範囲内に収めることが できれば,本研究の目的を十分に満足する.

以上を踏まえた今後の課題は,データ数増 加及び,CNN 構造の改良による年齢推定精度 向上の検討である.

利益相反の有無

なし

文 献

- [1] Hazlett H C, Poe M, Gerig G et al: Magnetic resonance imaging and head circumference study of brain size in autism: birth through age 2 years. Archives of general psychiatry 62(12): 1366-1376, 2005
- [2] Lao Z, Shen D, Xue Z, Karacali B et al: Morphological classification of brains via high-dimensional shape transformations and machine learning methods. Neuroimage 21(1): 46-57, 2004
- [3] Wang B, Pham T D: MRI-based age prediction using hidden Markov models. Journal of neuroscience methods, **199**(1): 21(1): 46-57, 2004
- Kondo C, Ito K, Wu K, Sato K et al: An age estimation method using brain local features for T1-weighted images. In 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC): 666-669, 2015
- [5] Pan Y, Huang W, Lin Z et al: Brain tumor grading based on neural networks and

convolutional neural networks. In 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC): 699-702, 2015

[6] Cole J H, Poudel R P, Tsagkrasoulis Det
 al: Predicting brain age with deep learning
 from raw imaging data results in a reliable
 and heritable biomarker. NeuroImage 163:
 115-124, 2017

Brain Developmental Age Prediction from Pediatric Brain CT Images Using 3D CNN

Ren MORITA^{*1}, Saya ANDO^{*2}, Daisuke FUJITA^{*1}, Manabu NII^{*1} Kumiko ANDO^{*3}, Reiichi ISHIKURA^{*3}, Syoji KOBASHI^{*1}

*1 Graduate School of Engineering, University of Hyogo
*2 Hyogo Prefectural Amagasaki General Medical Center
*3 Kobe City Medical Center General Hospital

One of the indices for diagnosing brain diseases in children using brain imaging is the evaluation of the progression of normal brain development, such as underdevelopment and prematurity. However, there is no quantitative method to estimate the degree of brain development, and diagnosis is currently based on the experience of doctors. Therefore, the lack of doctors who can read images and the lack of quantitative methods are problems. In this study, we propose a method to predict the brain developmental age from CT images of the brain in pediatric patients. In this method, the cranial region is extracted from the CT image and corrected for posture and position. We propose a new network model that extracts features from CT images using a 3D convolutional neural network (3D CNN) and predicts the brain developmental age in all coupled layers. The performance of this model was evaluated using 60 patients with no neurological abnormalities between the ages of 0 and 3 years. The root mean square error between the predicted age and the actual age of the patients was 7.80 (months), and the correlation coefficient was 0.801.

Key words: 3D CNN, CT Image, Pediatric Brain, Developmental Age, Age Prediction
自己符号化器を用いた慢性閉塞性肺疾患における

身体活動性の予測に関する研究

中原 祥喜*1 間普 真吾*1 平野 綱彦*2

村田 順之*3 土居 恵子*2 松永 和人*2

要旨

慢性閉塞性肺疾患(COPD)は、タバコ煙などの有害粒子を長期に吸入することで生じる肺の慢 性炎症性疾患である. COPD に罹患すると、患者は身体活動性が低下し日常生活が制限される (身体活動性障害).そして身体活動性障害が COPD の悪化につながるという循環に陥ってし まう.しかし、身体活動性障害を引き起こしている根本的な原因は明らかになっていない.また、 検査にはコストがかかるという問題もあるため、限られた検査項目から高精度な身体活動性障 害の予測が可能になれば COPD の治療に大きな貢献となる.そこで本研究では、各患者の検査 項目から身体活動性を予測し身体活動性を低下させる要因を明らかにすることを目的とする.

キーワード:自己符号化器,慢性閉塞性肺疾患,身体活動性,識別

1. はじめに

慢性閉塞性肺疾患(COPD)は、主に喫煙により 生じた肺の慢性炎症性疾患であり、呼吸機能検査 では正常に復することのない気流閉塞をきたす疾 患である[1].気流閉塞やエアートラッピングが生 じると肺の静的、動的過膨張に至る.これにより 患者は労作時呼吸困難を感じるため、運動耐容能

*1 山口大学大学院創成科学研究科

〔〒755-8611 山口県宇部市常盤台 2-16-1〕 e-mail: mabu@yamaguchi-u.ac.jp

*2 山口大学医学部附属病院呼吸器・感染症内 科

*3 山口大学医学部附属病院器官病態内科学

および各運動強度での活動時間と身体活動量を指 標とした身体活動性が低下し,日常生活が制限さ れてしまう.これを身体活動性障害と呼ぶ.すな わち COPD 患者は運動能力が低下し,運動が行 えないことで身体活動性が低下し,さらに筋力が 低下する悪循環を引き起こし予後の悪化にもつな がってしまう.しかし,COPD の悪化の原因であ る身体活動性障害の機序や制御機構,および治療 標的は分かっていない.また,関連する検査項目 の量が多く,すべての検査を行うにはコストがか かる問題もある.そのため,身体活動性障害の高 精度な予測を可能にするとともに,重要な検査項 目を明らかにすることは COPD の治療に対して 重要である.本研究では,山口大学医学部附属病

) - I JH/(
	データ数
COPD 患者	143
非 COPD 患者	238
健常者	25

表1 各患者データの内訳

にて検査された COPD 患者および非 COPD 患者 のデータから身体活動性予測モデルを構築すると ともに身体活動性を低下させる要因の分析を行っ た.

2. データセット

1) データの概要

本研究で使用するデータセットは、山口大学医 学部附属病院呼吸器・感染症内科から提供された 患者データ (COPD 患者, 非 COPD 患者, 健常 者)である.それぞれの内訳を表1に示す.検査 項目は、年齢、Whole body Phase Angle、4 種類の 血液検査結果や8種類の肺機能など31種類であ る.また、実験の際にはこれらの値を最小値0最 大値1になるように正規化した.なお,目的変数 は T-Ex (身体活動性) である. 身体活動性は身体 活動の性能を表すもので一般に、3 より小さけれ ば身体活動性の低下と考えられるため, 本研究で は、3 未満なら異常、3 以上なら正常とし 2 クラス 分類を行う. また、本研究で扱ったデータセット には、欠損値が含まれていたので、0 埋め, Multiple Imputation by chained equations (MICE) [2], K近 傍法(KNN)のいずれかで欠損値処理を行った.以 下に、それらの詳細を示す.

2) 欠損値処理

0 埋め:

0 埋めでは,欠損している箇所すべてに 0 を代入 することで欠損値を処理した.

MICE :

MICE は、欠損値を他の特徴量から回帰計算して 予測する多重代入法と呼ばれる手法の一種である. MICE では、多重代入法の欠点であった、複数の 特徴量において、異なるパターンで個人ごとのデ ータに欠損値が発生した場合、データセットの解 析が困難であるという問題を解決した手法である [3]. ここで、異なるパターンとは、MCAR (missing completely at random), MAR (missing at random), MNAR (missing not at random)といった欠損値の パターンを指す.以下に、解析するデータセット の特徴量を $x_1, x_2, ..., x_q$ として MICE の手順を示す. 1. 欠損個所に初期値 (平均値など)を代入し疑似

- x₁が欠損値の場合, x₁を欠損状態に戻し 回帰 モデル(線形回帰など)で疑似的な完全データ x₂,x₃...,x_qを基にx₁の補完値を生成し置き換え る.
- 3. x2 以降も手順 2 と同様に補完値を生成し置き 換える.

4. 手順 2~3 を任意の回数繰り返す.

的な完全データを作成する.

MICE は補完値を逐次的に更新していく手法で ある.本研究では,初期値に平均値,回帰モデルに 線形回帰モデルを用いる.

KNN:

KNN は教師あり学習の分類アルゴリズムであ る.与えられた学習データを特徴空間上にプロッ トし、未知のデータが現れたら そこから距離が近 い順に任意のk個のデータを取得し、多数決で決め られたクラスに未知データを分類する手法である [4].距離として一般的にユークリッド距離を使う. 本研究ではk = 1として欠損値処理に使用する.す なわち、欠損が含まれるデータに最も近いデータ を探索し、探索されたデータが有する値で置き換 える.最も近いデータも欠損があった場合は、欠 損を含まないデータから最も近いデータを選択す る.なお、k = 1の場合は、特に最近傍法という.

3. 方法

1) 自己符号化器を用いた身体活動性予測モデル 本研究では、3 層と4 層の階層型ニューラルネ



図1 3層自己符号化器のモデル構造



図2 ファインチューニング時のモデル構造

ットワーク(Neural Network, NN)をベースとし, 自己符号化器による事前学習とファインチューニ ングを行うモデルを構築した.自己符号化器は重 みの初期値の学習だけでなく,入力の圧縮・復元 を目的とした学習によって検査項目に潜む重要な 概念を抽出することを目的としている. そして, ファインチューニングにより,身体活動性(T-Ex) の予測を目的とした学習を行う.3 層の自己符号 化器とファインチューニング後のモデル構造の概 観をそれぞれ図1と図2に示す.なお,自己符号 化器の誤差関数は平均二乗誤差,ファインチュー ニング後のモデルの誤差関数は交差エントロピー を用いて学習を行った.

なお本研究では,自己符号化器の訓練時におい て,モデルの入力層に目的変数となる身体活動性 を含めて学習を行った.これは,事前学習の際に 身体活動性を入力とすることで,身体活動性の予 測に必要となる特徴をより抽出できると考えたか らである.次にファインチューニングを行うが, その際,入力層に目的変数である身体活動性を含 めることはできないため,身体活動性に対応する ユニットには0を入力することにした.また,身 体活動性を表すユニットは正常と異常に対応して 2 つあるため(すなわち one-hot ベクトルを表す), ファインチューニング後の出力層のユニット数は 2 とした.

また、本研究で扱っているデータ数は多いとは 言えず、単純に層が深くなっただけでは効果がな い可能性がある。例として、学習パラメータ数が 多すぎるネットワークは過学習を引き起こす可能 性がある。そこで、それを防ぐために 4 層の自己 符号化器に Dropout を適用した。

2) 各検査項目の重要度の算出

本研究では、重要度の算出に Permutation importance (PIMP)[5]を用いた. PIMP は、ある 特徴量をデータ間でランダムに入れ替える手法で ある.この処理により、入れ替えられた特徴量は 目的変数との相関が取れない状態になる、つまり 目的変数を説明する能力を持たなくなる.この状 態で予測を行い、その識別率(*acc_p*)が入れ替え前 のデータで予測した識別率(*acc_p*)と比べて、向上 または低下したかどうかで当該特徴量がどれだけ 予測に貢献しているか判断する.本研究では、(1) 式で特徴量の重要度(importance)を算出する.

importance = $acc_o - acc_p \cdots (1)$

(1)式より, importance の値が大きいほど入れ替え られた特徴量が予測に貢献していると判断できる.

4. 実験と結果

1) 実験条件

データセットを訓練データとテストデータに分 割する方法について説明する.本データセットに は別日に検査した同じ患者の検査結果が含まれて おり,これを考慮せずに分割すると訓練データと テストデータに同じ患者の検査結果を含む可能性 がある.これでは信頼性の高い予測結果が得られ ないため,訓練データとテストデータにはすべて の試行で同じ患者の検査結果は含まれていない.

表 2	訓練デー	- タ	と	テス	\mathbb{P}	デー	タ	の内訳
1			\sim	/ / ·	•	/	/	· / J J H/ V

	データ数
訓練データ	377
テストデータ	29

表3 各手法の平均識別率

	識別率(%)
3層 AE なし NN (0)	63.55
3層AEなしNN (MICE)	66.26
3層AEなしNN (KNN)	67.49
4 層 AE なし NN (0)	63.79
4層AEなしNN (MICE)	66.50
4層AEなしNN (KNN)	67.24
3 層 AE あり NN (0)	64.53
3層AEありNN (MICE)	68.47
3層AEありNN (KNN)	70.70
4 層 AE あり NN (0)	64.78
4層AEありNN (MICE)	68.72
4層AEありNN(KNN)	70.20

それらを考慮した訓練データとテストデータの内 訳を表2に示す.

2) 識別結果

本実験では、自己符号化器による事前学習なし の3層、4層NN、および事前学習ありの3層、4 層NNの4つのNNに対して、0埋め、MICE、 KNNの3つの欠損値処理を行ったデータセット 用い、計12パターンの条件で比較を行った.なお、 各手法について、独立した試行を14回行い、その 平均で評価を行った.評価指標は識別率である. 各手法における平均識別率を表3に示す.なお表 3では、事前学習なしの3層(または4層)のNN を「3層(または4層)AEなしNN」、自己符号化 器による事前学習ありのNNを「3層(または4 層)AEありNN」と表している.表3より、KNN で欠損値を処理した3層AEありNNが最も高い 識別率を示した.また、欠損値処理に着目すると、



どのモデルも KNN で欠損値処理をしたモデルが 最も高い識別率を示した.これにより,本研究で 扱ったデータでは,自己符号化器による事前学習 と KNN による欠損値処理が効果的であったと言 える.

3) 各検査項目の重要度

図 3 に各検査項目の重要度を箱ひげ図で示す. なお,図 3 で示す結果は,最も高い識別率を示し た KNN で欠損値を処理した 3 層 AE あり NN に よるものである.

図3において,緑の三角形は平均値を表しており, 平均値でソートしている.実験方法でも記載したよ うに, importance の値が大きいほどその検査項目が 重要であるので,このグラフから性別や年齢,Whole body Phase Angle (栄養状態)が重要な項目であると いえる.また,平均値が0または,0を下回っている 検査項目は,予測するのにノイズになっている可能 性があると考えられる.

4. まとめ

本研究では、自己符号化器を用いたニューラル ネットワークを用いて身体活動性の予測とその予 測に最も貢献している検査項目の分析を目的とし た研究を行った. 実験の結果,自己符号化器による事前学習を行った手法が事前学習を行わない手法より高い識別 率を得た.また,欠損値処理は KNN が全ての手 法において,最も高い識別率を得ることができた. 各検査項目の重要度に関しては,年齡や性別, Whole body Phase Angle (栄養状態)が重要な項 目であることが確認できた.

今後の課題としては,自己符号化器の改良や, ノイズの可能性がある検査項目を除去することで, 識別率を向上させることが挙げられる.また本研 究では,特徴量の重要度の算出の際に,一つの特 徴量のみをランダムに入れ替え,その特徴量がど れだけ目的変数の予測に貢献しているかを判断し たが,この手法では,相関の高い特徴量を同時に 扱っている場合,それらの評価が適切に行われな い問題もある.今後は複数の特徴量を同時にシャ ッフルすることで,予測精度がどのように変わる か,特徴量の組合せによる識別への効果などを明 らかにしたい.

謝辞

本研究は JSPS 科研費 JP19K12120 の助成を受けたものです.

利益相反の有無

なし.

参考文献

- [1] 南方良章:身体活動性の重要性とその向上法, 日本内科学会雑誌 105 巻 6 号 pp.963-969, 2016
- [2] Melissa J. Azur, Elizabeth A. Stuart, Constantine Frangakis, Philip J. Leaf: Multiple Imputation by chained equations: what is it and how does it work?, PSYCHIATRIC RESEARCH Volume 20 Issue 1 pp.40-49, 2011
- [3] 野間久史:連鎖方程式による多重代入法,応 用統計学 46 巻 2 号 pp.67-86,2017
- [4] 浜本義彦: 統計的パターン認識, 森下出版株 式会社, 2012
- [5] Andre Altmann, Laura Tolosi, Oliver Sander, Thomas Lengauer: Permutation importance: a corrected feature importance measure, Bioinformatics Volume 26 Issue 10 pp.1340-1347, 2010

Study on the Prediction of Physical Activity of COPD

Patients Using Autoencoder

Yoshiki NAKAHARA^{*1}, Shingo MABU^{*1}, Tsunahiko HIRANO^{*1}, Yoriyuki MURATA^{*1}, Keiko DOI^{*1}, Kazuto MATSUNAGA^{*1}

*1 Yamaguchi University

COPD is lung inflammation which is caused by long-term inhalation of tobacco and other harmful substances. COPD patients have reduced physical activity and restricted daily life (physical activity disorder), which then aggravates COPD. However, the main reason of the physical activity disorder has not become clear. Also, the cost of the medical examination another problem. Therefore, all kinds of medical examination cannot be done every time, but only simple examination is carried out. The purpose of this study is to predict physical activity and identify the factors that reduce the physical activity based on several attributes of the medical examination of each patient.

Key words: Autoencoder, COPD, Physical activity, classification

機械学習・深層学習を用いた手 X 線画像における

リウマチmTSS 推定法

中津 康平^{*1} 盛田 健人^{*2} 藤田 大輔^{*1} 小橋 昌司^{*1}

要旨

現在,関節リウマチの国内患者数は 60 万人から 100 万人であると推定され,毎年数万人ずつ増加している.関節リウマチの診断には、手足の X 線画像から算出する mTSS (modified Total Sharp Score) が標準的な診断基準として用いられているが、診断の問題点として、医師の目視による主観的診断であること、スコアの判断箇所が多いことが挙げられている.そのため、医師の負担軽減や診断の迅速性・正確性向上のために mTSS 自動で評価するコンピュータ支援診断 (CAD) システムの開発が期待されている.本研究では、手関節リウマチの mTSS 推定法に関して、リッジ回帰 (RR)を用いた手法と、畳み込みニューラルネットワーク (CNN)の3 つのモデル (VGG16, DenseNet201, Xception)を用いた手法の比較を行う.4 手法の比較のために、90 名の RA 患者を対象に、手のX 線画像を用いた実験を行った.実験結果は、erosionの mTSS 予測は RR, JSN の mTSS 予測は VGG16 で最も良い結果が得られた.

キーワード:関節リウマチ,畳み込みニューラルネットワーク,X線画像,コンピュータ支援診断システム

1. はじめに

2014年の調査で,日本には60万から100万 人の関節リウマチの患者がいると推計されて いる.近年では薬の進歩により,薬による関節 リウマチの治療が一般的に行われるが,適切な 治療を行うために,関節リウマチの進行度を正 確に評価する必要がある.

関節リウマチの進行度の評価には,手または 足の X 線画像から計測可能な modified total sharp score (mTSS)が用いられる[1]. mTSS は手 関節 16 箇所または足関節 6 箇所においてびら ん度,関節裂隙狭小化度をそれぞれ 5 段階評価

*1 兵庫県立大学 工学研究科 [〒671-2280 兵庫県姫路市書写 2167] e-mail: kobashi@eng.u-hyogo.ac.jp *2 三重大学 工学研究科 [〒514-0102 三重県津市栗真町屋町 1577] し、そのスコアの合計値で進行度の評価を行う 評価方法である. 適切な治療のためには mTSS を年に数回計測し、その推移を観測する必要が あるが、mTSS の評価は医師の熟練性に依存し、 スコアを決定する客観的な方法は確立されて いない. そのため、客観的にスコアを決定する ために、X線画像解析に基づく mTSS 自動測定 システムが必要とされている.

我々は以前,手のX線画像を用いたmTSSの 推定法として,リッジ回帰によるmTSS推定法 を提案した[2]. その結果,びらん度を 53.3%, 関節裂隙狭小化度を 60.8%の精度で推定に成功 した.しかし,この推定精度は実用化には不十 分であり,更なる精度向上が必要である.

本研究では、手関節リウマチの mTSS 推定手 法として、リッジ回帰を用いた手法[2]と3 つの 畳み込みニューラルネットワーク (CNN) モデ ル (VGG16, DenseNet201, Xception)を用いた 手法の比較を行う.





(a) リウマチが進行 していない患者. 図1 手関節 X 線画像例.

(b) リウマチが進行 した患者.

表1 mTSSと進行度の関係[1].

スコア	びらん度	関節裂隙狭小化度
0	びらんなし	正常
1	小さなびらん	局所のみ、わずか
2	関節面の半分に	50%以上の残存
	満たないびらん	
3	関節面の半分以	50%以下が残存
	上のびらん	亜脱臼
4		関節裂隙消失
		完全脱臼
5	完全に圧壊	

2. 使用データ

本研究では、関節リウマチが進行した 90 名 の両手単純 X 線画像を使用した. 撮影された両 手 X 線画像を左右に分割し,右手画像を反転さ せることですべて左手のみの X 線画像として 扱うことで、データ数を 180 例とした. また、 学習用データおよび評価データ作成のため,各 関節に対し専門家により mTSS を手動で与えた. 図1に使用した X線画像の例を、表1にmTSS の評価基準を示す.

本研究は兵庫県立加古川医療センターの倫 理委員会で承認を受け実施した.

3. 提案法

mTSS は関節リウマチの進行度に応じた順序 尺度で定義される.提案法では,機械学習や深 層学習を用いた回帰アルゴリズムにより, mTSS を順序尺度として推定を行う. mTSS 推 定法として 1) リッジ回帰による推定法[2], 2) CNN による推定法の比較を行う.なお,リッジ 回帰による推定法については,我々の過去の研 究と同じ手法を用いる[2].

1) リッジ回帰を用いた mTSS 推定法[2]

手作業で抽出した関節位置を中心に、90×90

		推定誤差 RMSE	推定 精度 (%)	相関 係数
リッジ	びらん	0.824 ±0.537	51.2	0.578
回帰	関節裂隙 狭小化	0.694 ±0.420	52.4	0.602
VGG16	びらん	0.854 ±0.620	48.4	0.697
	関節裂隙 狭小化	0.708 ±0.541	56.8	0.741
DenseNet	びらん	0.948 ±0.682	45.7	0.567
201	関節裂隙 狭小化	0.782 ±0.569	50.7	0.619
Vantion	びらん	0.986 ±0.714	44.4	0.497
Xception	関節裂隙 狭小化	0.844 ±0.608	47.9	0.553

表 2 mTSS の 推定結果.

(ピクセル)の画像パッチをX線画像から切り 出す. 各パッチは, X 線撮影時の手の傾きや, リウマチの進行に伴う指の曲がりに対応する ため、-30~30°の範囲で 15°刻みで回転させ る.以上により、1人の被験者につき140個の 画像パッチが切り出され、その画像パッチから 抽出された HOG 特徴量[3]を入力とし、リッジ 回帰による学習を行う.

2) CNN を用いた mTSS 推定法

本手法では, CNN モデルとして VGG16, DenseNet201, Xception を使用し, 比較を行う. 回帰アルゴリズムに基づいて mTSS を予測する ために, 各 CNN モデルの出力層を1 ニューロ ンとする.学習データはリッジ回帰と同様の手 法で作成した画像パッチを用いる. また, CNN による予測精度を向上させるために、ImageNet データセットを用いて転移学習を行う. 再学習 は、VGG16モデルの場合は15層目以降のみ、 DenseNet201 と Xception の場合は全結合層に対 して行う. CNN の学習に用いる損失関数は二乗 平均平方根誤差(RMSE)とする.

4. 実験結果·考察

提案法に対し,90 例の被験者の X 線画像を

用いて実験を行った.ここで,90名の被験者を 学習:評価:テスト=8:1:1と分割する10分割 交差検証で評価を行った.

各手法の mTSS 推定結果を表 2 に, mTSS 推 定結果の混同行列を表 3 から表 10 に示す.こ こで,推定誤差は, mTSS の推定結果の真値と の RMSE,推定精度は mTSS の推定結果を四捨 五入した値の真値との正解率とした.

結果より,推定誤差に関してはびらん度と関 節裂隙狭小化度のいずれもリッジ回帰の結果 が最も小さく,推定精度に関してはびらん度は リッジ回帰の結果が,関節裂隙狭小化度では VGG16の結果が最も高い結果となった.また, 表2の相関係数,表3から表10の混同行列よ り,CNNによる推定結果は,リッジ回帰による 推定結果と比較して,スコア間の偏りが小さい ことが分かった.

5. まとめ

本研究では手関節リウマチの mTSS 自動推定 法に関して、リッジ回帰による手法と CNN の 3 つのモデルによる推定法の比較を行った.

結果より、びらん度推定に関してリッジ回帰 の手法が 51.2%の推定精度と最も高く、関節裂 隙狭小化度推定に関して VGG16 の手法が 56.8%の推定精度と最も高いという結果となっ た.

表 3 リッジ回帰による mTSS 推定の混同行列 (びらん).

				推定値		
		0	1	2	3	4
	0	18	332	53	0	0
	1	86	1094	223	8	0
真値	2	7	250	143	9	0
	3	0	33	77	28	0
	4	0	15	78	59	7

今後の課題としては, CNN による mTSS 推定 に関して転移学習のためのデータセットの最 適化,学習データ拡張による精度改善である.

利益相反の有無 なし

文 献

- D. M. Van der Heijde, "Plain X-rays in rheumatoid arthritis: overview of scoring methods, their reliability and applicability," Baillieres Clin Rheumatol, 10, pp. 435-453, 1996.
- K. Morita, P. Chan, M. Nii, et al. "Finger Joint Detection Method for the Automatic Estimation of Rheumatoid Arthritis Progression Using Machine Learning," 2018
 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 1315-1320, 2018.
- [3] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol.1, pp. 886-893, 2005.

表 4 リッジ回帰による mTSS 推定の混同行列 (関節裂隙狭小化).

			推定值							
		0	1	2	3	4				
	0	2	15	11	0	0				
	1	4	41	99	19	0				
真値	2	0	86	598	372	19				
	3	0	5	291	555	61				
	4	0	0	32	185	125				

表 5 VGG16 による mTSS 推定の混同行列 (びらん).

		推定值							
		0	1	2	3	4	5		
	0	191	172	38	2	0	0		
	1	407	756	204	41	3	0		
古体	2	29	170	149	55	6	0		
具但	3	0	14	33	67	23	1		
	4	0	3	28	58	62	8		
	(5)	0	0	0	0	0	0		

表 6 VGG16 による mTSS 推定の混同行列 (関節裂隙狭小化).

			推定值								
		0	1	2	3	4	5				
	0	10	18	0	0	0	0				
	1	16	102	45	0	0	0				
真値	2	21	221	609	214	10	0				
	3	0	13	272	544	83	0				
	4	0	0	15	156	167	4				
	(5)	0	0	0	0	0	0				

表 7 DenseNet201 による mTSS 推定の混同行列 (びらん).

(5)

(6)

真値

推定値

4 5 6

1 0 0

3 0 0

1 0

13 3 0

0 0 0

0 0

表 8 DenseNet201 による mTSS 推定の混同行列 (関節裂隙狭小化).

		推定值							
		0	1	2	3	4	5		
	0	5	12	11	0	0	0		
	1	8	63	81	11	0	0		
古庙	2	17	172	530	331	24	1		
具但	3	0	23	291	503	91	4		
	4	0	0	29	181	116	16		
	(5)	0	0	0	0	0	0		

表 9 Xception による mTSS 推定の混同行列 (びらん).

			推定値							
		0	1	2	3	4	5	6		
	0	38	248	100	13	4	0	0		
	1	96	819	435	57	3	1	0		
	2	10	168	180	41	8	2	0		
真値	3	0	17	81	24	12	4	0		
	4	0	18	55	66	10	9	1		
	(5)	0	0	0	0	0	0	0		
	(6)	0	0	0	0	0	0	0		

表 10 Xception による mTSS 推定の混同行列 (関節裂隙狭小化).

		_	推定値				
		0	1	2	3	4	5
真値	0	10	18	0	0	0	0
	1	16	102	45	0	0	0
	2	21	221	609	214	10	0
	3	0	13	272	544	83	0
	4	0	0	15	156	167	4
	(5)	0	0	0	0	0	0

Comparison of rheumatoid arthritis modified total sharp score prediction performance in x-ray images between machine learning and deep

learning

Kohei NAKATSU^{*1}, Kento MORITA^{*2}, Daisuke FUJITA^{*1}, Syoji KOBASHI^{*1}

*1 University of Hyogo *2 Mie University

Currently, the number of patients with rheumatoid arthritis(RA) in Japan is estimated to be between 600,000 and 1,000,000, with the number increasing by tens of thousands every year. The modified total sharp score (mTSS) calculated from X-ray images is a standard diagnosis method of RA progression, but it has many evaluation joints and the determination of the score depends on the experience of the physicians. Therefore, the development of a computer-aided diagnosis (CAD) system that automatically evaluates the mTSS is expected to reduce the burden on physicians. In this study, we compare a method using ridge regression and methods using 3 models (VGG16, DenseNet and Xception) of deep learning with respect to the mTSS prediction of hand joints. To compare the 4 methods, we conducted an experiment on 90 RA patients using hand X-ray images. The experimental results showed that the mTSS prediction of JSN was best with VGG16.

Key words: Rheumatoid arthritis, Convolutional Neural Network, X-ray image, Computer-aided diagnosis system

機械学習によるコルポスコープ画像における

白色上皮領域検出精度向上のための検討

篠原 寿広*1 藤井 一輝*1 村上 幸祐*2 中迫 昇*1 松村 謙臣*2

要旨

子宮頸癌の診断において、コルポスコープとよばれる拡大鏡を用いて、目視による検査が行われている. 本検査はコルポスコピーとよばれ、子宮頸癌の確定診断を行う組織診のための組織を採取する部位を決定 する非常に重要な検査である.しかしながら、コルポスコピーの精度は、検査者の知識と経験に大きく依 存する.そこで、われわれは、機械学習を用いたコルポスコピー支援を検討している.これまでに初期検 討として、子宮頸部の代表的な病変の一つである白色上皮を対象に、分割したコルポスコープ画像に対し て、Convolutional Neural Network を用いて、白色上皮の判別を行った.本研究では、判別精度向上のため、 画像の分割条件について詳細に検討した.また、分割した画像を元の位置に戻し、周辺の分割画像の判別 結果を重み付き平均することにより、判別精度が向上することを確認した.

キーワード:コルポスコピー,子宮頸癌,白色上皮,診断支援,機械学習

1. はじめに

子宮頸癌はヒトパピローマウイルスの持続 的な感染が原因となって発症する癌で,全世界 では,年間約 570,000 人が罹患し,約 311,000 人 が死亡しており,女性の癌による死亡原因の第 4 位となっている [1].日本においては,年間 約 10,000 人が子宮頸癌に罹患し,約 3,000 人が 死亡している [2,3].

子宮頸癌の診断において、コルポスコープと よばれる拡大鏡を用いて、酢酸で加工した子宮 頸部を目視により観察する検査が行われてい る.本検査はコルポスコピーとよばれ、子宮頸 癌の診断において重要な役割を担っている.し かしながら、コルポスコピーは検査者の知識と 経験に強く依存し[4]、経験の浅い医師には正

*1 近畿大学生物理工学部生命情報工学科
〔〒649-6493 和歌山県紀の川市西三谷 930〕
e-mail: sinohara@waka.kindai.ac.jp
*2 近畿大学医学部産科婦人科学教室

確な診断が難しい.そこでわれわれは,機械学 習によるコルポスコピー支援を検討している. 初期検討として,これまでに子宮頸部の代表的 な前癌病変の一つである白色上皮を対象に,コ ルポスコープに取り付けたデジタルカメラで 撮影した子宮頸部画像を小領域に分割し, Convolutional Neural Network (CNN)を用いて白 色上皮の判別を行った [5].

本研究では、判別精度向上のため、画像の分 割条件を詳細に検討する.また、判別精度向上 のため、分割した画像を元の位置に戻し、周辺 の分割画像の判別結果を重み付き平均する手 法を提案する.

2. 機械学習による白色上皮の判別 [5]

本研究では、機械学習による白色上皮の判別 のため、はじめにコルポスコープ画像を小領域 に分割し、医師による診断データを元に、分割 した画像の各画素にラベル付けを行ない、それ らを学習用画像とする.ラベルは白色上皮と白 色上皮以外の2つのラベルであり,2クラス分類問題を考える.判別対象のコルポスコープ画像(以下,テスト用画像とよぶ)も同様に分割し, それらを CNN により判別する.

3. 判別精度向上のための検討

前節で述べた機械学習による白色上皮判別 の精度向上のため、コルポスコープ画像の分割 条件の詳細な検討を行う.また、分割画像を元 の位置に戻し、周辺の分割画像の判別結果を重 み付き平均する手法を提案する.

3.1 画像分割条件の検討

これまでのわれわれの初期検討におけるコ ルポスコープ画像の分割条件は十分に検討さ れていなかった.そこで、本研究では、さまざ まな分割条件を試すことにより、最適な分割条 件を探索する.このとき、分割画像の大きさは 分割条件の一つとなる.また、本研究において、 コルポスコープ画像は重なりを許して分割す る.すなわち、隣り合う分割画像同士は重なる 部分が存在する.したがって、分割条件として、 分割画像の大きさ(以下、分割サイズとよぶ) と、隣の分割画像との重なりの程度(以下、重 なりとよぶ)の2つがパラメータとなる.

3.2 判別結果の重み付き平均化処理

分割画像を単純に判別する場合,1 枚の画像 のみで判別しなければならない.しかしながら, 照明の影響などにより,1 枚の画像のみでは判 別が難しい場合がある.病変部の多くは複数枚 の分割画像にわたって広がっているため,注目 する分割画像の周辺情報を利用することは有 効であると考える.そこで,白色上皮の判別後, 画像を元の位置に戻し,周辺の分割画像の判別 結果を重み付き平均することにより,精度向上 を図る.本処理を以下では単に平均化処理とよ ぶ.

4. 白色上皮判別実験

近畿大学病院においてコルポスコピーを実施した 20 代~40 代の患者のコルポスコープ画像 11 症例を対象に実験を行った.本研究は近畿大学医学部倫理委員会および同大学生物理

工学部生命倫理委員会によって承認され, コル ポスコープ画像の使用はすべての患者より書 面による同意を得ている.また,使用したコル ポスコープ画像はデジタルカメラ (E-PL6, OLYMPUS)により撮影した 4,806×3,456 pixels の RGB 画像である.実験に使用した計算機環 境を表1に示す.本研究で用いる CNN のネッ トワーク構造は VGG (Visual Geometry Group)16 [6] に基づく. CNN の学習パラメータなどの 実験条件を表2に示す.また,画像分割条件と 画像枚数を表3に示す.

表1 計算機環境.

CPU	Intel Core i7-4770K
メインメモリ	16GB
GPU	NVIDIA Quadro K5000
OS	Ubuntu 16.04 LTS
ソフトウェア	Chainer

表2CNNの学習パラメータ.

エポック数	500
バッチサイズ	128
最適化手法	慣性項付き確率的 勾配降下法
学習率	0.001
慣性項	0.9

表3 画像分割条件と画像枚数.

分割サイズ	重なり	白色上	白色上皮
(pixels)	(pixel)	皮枚数	以外枚数
128×128	なし	756	1005
128×128	64	3017	3947
128×128	96	12042	15723
64×64	なし	3146	4080
64×64	32	12643	16285
32×32	なし	12837	16502

これまでの初期検討により, コルポスコープ 画像の分割画像が128×128 pixels よりも大きい 場合は, 判別性能が良くなかったため, 128× 128 pixels を最大の分割サイズとしている.テス ト用画像も, 学習用画像と同様に作成して判別 した. なお, 学習用画像は回転などによる加増 は行っていない. 11 症例のうち, 学習用画像を 作成するために 10 症例を使用し, テスト用画 像を作成するために残り1症例を使用して,交 差検証により判別性能を算出した.また,平均 化処理においては,注目する分割画像を中心に 注目画像,上下左右の4近傍画像,4近傍を除 く8近傍画像のそれぞれのSoftmax 関数の出力 をそれぞれ4:2:1の割合で重み付けして平均 した.

実験結果の一例として,分割サイズが 128× 128 pixels,重なりが 64 pixel のときの判別結果 を図 1 に示す.図 1(a)は平均化処理しておらず, 図 1(b)は平均化処理を行っている.黄色が真陽 性,青色が偽陰性,緑色が偽陽性を表している.



(a) 平均化処理なし



(b) 平均化処理あり
図2 白色上皮判別結果の一例
(分割サイズ:128×128 pixels,重なり:64 pixel 黄色:真陽性,青色:偽陰性,緑色:偽陽性).

表4に平均化処理していないときの各画像分 割条件における判別性能を示す.また,表5に 平均化処理したときの各画像分割条件におけ る判別性能を示す.

表4 白色上皮の平均判別性能

(平均化処理なし).						
分割サイズ	重なり	感度	適合度	F値		
(pixels)	(pixel)					
128×128	なし	0.559	0.290	0.382		
128×128	64	0.592	0.318	0.414		
128×128	96	0.667	0.300	0.414		
64×64	なし	0.586	0.270	0.370		
64×64	32	0.625	0.284	0.391		
32×32	なし	0.649	0.247	0.358		

表5 白色上皮の平均判別性能

(平均化処理あり).					
分割サイズ	重なり	感度	適合度	F値	
(pixels)	(pixel)				
128×128	なし	0.561	0.363	0.441	
128×128	64	0.624	0.347	0.446	
128×128	96	0.688	0.320	0.437	
64×64	なし	0.728	0.292	0.417	
64×64	32	0.651	0.301	0.412	
32×32	なし	0.673	0.270	0.385	

表4より、平均化処理を用いない場合は、分 割サイズが128×128 pixels で重なりが64 pixel および32 pixel のとき、もっともF値が大きく なり、分割サイズが32×32 pixels のとき、F値 が最も小さくなった.また、表5より、平均化 処理を用いた場合は分割サイズが128×128 pixels、重なりが64 pixel のとき、F値がもっと も大きくなり、平均化フィルタを用いない場合 と同様に、分割サイズが32×32 pixels のとき、 F値が最も小さくなった.なお、平均化処理を 用いない場合と比べ、用いた場合は、すべての 指標においてに向上がみれらた.

以上の結果から,分割サイズについてはテク スチャがある程度認識できる大きさが判別に は適していることがわかる.また,重なりを許 して分割することで,学習用画像が増え,汎化 性能が向上したと考えられる.さらに,平均化 処理することにより,1 枚では照明の影響等に より判別が難しい場合でも周辺情報を効果的 に利用でき、判別性能が向上することがわかった.子宮頸部のテクスチャは個人差が大きいことに加え、照明の影響等により、コルポスコープ画像のバリエーションが大きく、判別精度は、あまり良い結果が得られていない.また、本実験で得られた分割サイズや重なりは、今回使用したデータセットのみに意味を持つという制限はあるものの、本研究のように分割した画像に対し、機械学習を行う問題において、有益な知見になり得るものと考える.

さらなる判別精度の向上のためには,学習用 画像を増やすことがまず考えられるが,子宮頸 部の個人差,照明の影響等によるコルポスコー プ画像のバリエーションを補正または分類す る前処理や,誤って判別された部分を修正する 後処理を追加することにより判別精度の向上 を図ることも期待できる.

5. まとめ

機械学習を用いたコルポスコピー支援とし て、これまで検討してきたコルポスコープ画像 からの白色上皮の判別に対し、本稿では、判別 精度向上の検討を行った.本研究では、コルポ スコープ画像を小領域に分割し、CNNを用いて 白色上皮かどうかを判別するが、画像分割条件 の詳細な検討、および分割画像を元の位置に戻 し、重み付き平均する手法の検討を行った.実 際のコルポスコープ画像 11 症例を用いて白色 上皮判別実験を行った結果、分割する画像の大 きさが 128×128 pixels、画像の重なりが 64 pixel のとき最も判別性能がよく、感度 62.4%、適合 率 34.7%となった.また、重み付き平均化処理 により、すべての画像分割条件において、判別 性能が向上することを確認した.

今後はさらなる判別精度向上のため、判別の ための前処理および後処理を考えるとともに、 VGG 以外の深層学習の適用を試みる.また、白 色上皮以外の異形成の判別や、異形成の軽度、 高度などの程度の判別についても検討を加え る予定である.

謝辞

コルポスコープ画像をご提供してくださっ た近畿大学医学部産科婦人科学教室の皆様に 感謝いたします.

利益相反の有無

なし.

文 献

- Bray F, Ferlay J, Soerjomataram I, et al.: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 68(6): 394-424, 2018.
- [2] 厚生労働省: 平成 29 年全国がん登録 罹患数・率報告, 2017.
- [3] 国立がん研究センターがん情報サービ ス: がん登録・統計(人口動態統計),2018.
- [4] Xue P, Ng M.T.A and Qiao Y: The challenges of colposcopy for cervical cancer screening in LMICs and solutions by artificial intelligence. BMC Med 18, 169 2020.
- [5] 篠原 寿広,藤井 一輝,村上 幸祐,他: 機械学習によるコルポスコープ画像に おける白色上皮領域検出の基礎的検討. 第39回日本医用画像工学会大会予稿集, オンライン,2020,P1-16.
- [6] Karen S and Andrew Z: Very Deep Convolutional Networks for Large-Scale Image Recognition. Proceedings of International Conference on Learning Representations, 2015.

Consideration for accuracy improvement of acetowhite epithelium

region detection in colposcopic image by using machine learning

Toshihiro SHINOHARA^{*1}, Kazuki Fujii^{*1}, Kosuke MURAKAMI^{*1}, Noboru NAKASAKO^{*1}, Noriomi MATSUMURA^{*2}

*1 Department of Computational Systems Biology, Faculty of Biology-Oriented Science & Technology, Kindai University

*2 Department of Obstetrics & Gynecology, Faculty of Medicine Kindai University

Gynecologists diagnosis cervical cancer by visual inspection with a magnifying lens called a colposcope. This diagnosis is called colposcopy, which is a crucial inspection for determining the place to obtain cells of the cervix for tissue diagnosis of cervical cancer. This inspection strongly depends on inspectors' knowledge and experience, so it is often difficult for inexperienced gynecologists to accurately diagnose lesions on the cervix. Thus, we are considering diagnosis assistance with machine learning. We discussed the detection of the acetowhite epithelium, one of the representative cervical lesions, for colposcopic images divided into small patches using a convolutional neural network in our previous study. In this study, the condition for image division was discussed to improve the discrimination performance. Furthermore, it was confirmed that the discrimination performance improved by weighted-meaning the discrimination results of the divided images around the divided image of interest.

Key words: Colposcopy, Cervical cancer, Acetowhite epithelium, CAD, Machine learning

覚醒下脳腫瘍摘出術における手術工程同定のための 深層学習を用いた顕微鏡画像からの特徴量抽出手法の提案

佐藤 生馬^{*1},田村 学^{*2},山口 智子^{*2},藤野 雄一^{*1}, 吉光 喜太郎^{*2},村垣 善浩^{*2},正宗 賢^{*2}

要旨

覚醒下脳腫瘍摘出手術では,最大限の腫瘍摘出と最小限の術後合併症リスクを実現するため,患者の脳構 造や機能を把握し,腫瘍摘出領域を決定することが重要である.しかし,脳腫瘍の種類・位置や脳機能位 置,患者の年齢などによって適した手術方針や工程は異なる.よって,執刀医は術中に患者の脳構造や機 能を把握し,適切な腫瘍摘出領域を決定する必要がある.この執刀医の意思決定は,執刀医の知識や経験 に依存しているため,手術スタッフの作業の最適化や若手医師の教育には手術工程の可視化と解析が重要 である.また,手術工程の可視化と解析の実現のためには,手術工程の自動同定が必要であると考えられ る.本研究では,手術工程の可視化と解析に向けた手術工程同定のために,手術顕微鏡動画から深層学習 を用いた特徴量抽出手法の提案を行う.

キーワード:深層学習,手術工程同定,術具検出,覚醒下手術,脳腫瘍

1. はじめに

覚醒下脳腫瘍摘出手術において,熟練医によ る手術では最大限の腫瘍摘出と最小限の術後 合併症リスクを実現するため,患者の脳構造や 機能を把握し,独自の判断プロセスにより腫瘍 を切除する.この判断プロセスは,熟練医が培 った知識や経験にもとづく暗黙知とされてお り,医療の質の向上や教育支援のためには若手 医師などと共有することが望まれる.また,手 術工程の可視化と解析の実現のためには,手術 工程の自動同定が必要であると考えられる.

覚醒下脳腫瘍摘出術における手術工程の自動同定を目標に機械学習を用いた手術工程同定システムを開発してきた[1]. このシステムは、 手術ナビゲーションシステムからの MRI 画像 上における術具位置情報と顕微鏡からの映像 情報を用いて、覚醒下脳腫瘍摘出術の工程を 3 階層の 13 工程に定義し、90%以上の精度で同定 可能とした.同定誤差要因の一つに顕微鏡動画 からの術具検出ミスがあった.また、より詳細 な工程を同定可能とするには、顕微鏡動画から 取得可能な特徴量を増やし、この特徴量による工程モデルを再構築する必要である.

本研究では,覚醒下脳腫瘍摘出手術におけ る手術工程の可視化と解析に向けた詳細な手 術工程同定のために,顕微鏡動画から深層学 習を用いた特徴量抽出手法を提案する.

2. 深層学習を用いた特徴量抽出手法

詳細な手術工程同定を目指し,顕微鏡動画から抽出する特徴量を選定する必要がある.これまでのシステム[1]では顕微鏡動画からの特徴量としてバイポーラ,電気刺激装置,剪刀の3種類の術具を検出していた.本手法では,対象術式で使用されることが多い吸引管,鉗子,鑷子を追加した6種類の術具を特徴量として検出する.また,脳表や脳内の血管への処置が工程を遷移する特徴量として使用可能であるため血管情報を抽出する.

深層学習を用いた特徴量抽出手法として,術 具検出には少ない画像で学習可能とするため, GANの一種である pix2pix[2]を用いる.また, 血管情報抽出には画像上では細い血管を抽出 するため,高解像度な画像に対応可能な pix2pixHD[3]を用いる.そして,顕微鏡動画に対 して, pix2pix の学習済みモデルを用いて術具の

^{*1} 公立はこだて未来大学 システム情報科学部 [〒041-8655 北海道函館市亀田中野町 116-2] *2 東京女子医科大学 先端生命医科学研究所

セグメンテーションを行う.同時に,pix2pixHD の学習済みモデルを用いて血管のセグメンテ ーションを行う.その後,pix2pix および pix2pixHD が生成した画像に対して,閾値を用 いて術具検出および血管情報の抽出を行う.

3. 特徴量抽出用学習データセット

顕微鏡画像を用いて術具検出および血管情 報の抽出を行うための学習データセットを構 築する.使用するデータは,臨床現場で得られ た顕微鏡動画(1080i)である.術具検出用の学習 画像は,6種類の術具が写っている画像(574枚) を手動でセグメンテーションしたものと術具 が何も写っていない画像(207枚)の計781枚を 用意した.セグメンテーションした術具はバイ ポーラが179個,吸引管424個,鑷子86個,鉗 子161個,剪刀114個,電気刺激プローブ70個 である.これらは複数の術具が1枚の画像に映 っていることがあるため,術具の合計数と作成 した画像の枚数に相違がある.このほかに,血 管情報抽出用の学習画像は,血管を手動でセグ メンテーションしたものを200枚用意した.

4. 術具検出および血管情報抽出精度評価

本手法による顕微鏡画像からの術具検出精 度評価を行う.本実験では,臨床現場で得た顕 微鏡動画に対して,術具ごとに約 20 個ずつを 抜き出し,術具検出を行い,精度を算出し評価 した.術具ごとの検出結果を表1に示す.術具 検出精度における平均正解率は,バイポーラが 98.5%,吸引管が 85.1%,鉗子が 91.0%,鑷子が 92.5%,剪刀が 87.3%,電気刺激装置が 88.8%と なった.

本手法による顕微鏡画像からの血管情報抽 出精度評価を行う.本実験では,顕微鏡動画に 対して,評価用の血管情報正解データを 10 枚 作成した.そして,正解データと生成データの

	正解率	再現率	適合率	F值
バイポーラ	98.5%	95.0%	95.0%	95.0%
吸引管	85.1%	96.2%	81.5%	88.2%
鉗子	91.0%	92.0%	69.7%	79.3%
鑷子	92.5%	72.0%	85.7%	78.3%
剪刀	87.3%	84.0%	61.8%	71.2%
電気刺激装置	88.8%	100%	57.1%	72.7%

双	表1	術具検出結果
----------	----	--------

	元画像	生成画像	正解画像	正解率	F值
1	あた	Por C		98.4 %	96.3 %
2	JE LE			99.4 %	95.0 %
3				99.4 %	96.1 %
4				99.0 %	96.1 %
5		100	100	99.5 %	97.2 %
6	125			99.6 %	96.2 %
7				99.5 %	95.0 %
8				99.6 %	96.0 %
9	1	11	1	99.4 %	94.6 %
10	1 Carlos	125	An	99.5 %	95.2 %
		全画像平均		99.3 %	96.0 %

図1 血管情報抽出精度評価結果

ー致度を画素単位で算出し,評価した.顕微鏡 画像からの血管情報抽出精度評価結果を図1に 示す.顕微鏡画像からの血管情報抽出精度にお ける平均正解率は99.3%であった.

5. 結語

本研究では、脳腫瘍摘出術における手術工程 同定のために、顕微鏡画像からの特徴量抽出手 法の提案を行った.術具検出精度における平均 正解率は 85%以上であった.また、血管情報抽 出精度における平均正解率は 99.3 %であった. 提案手法を用いることで従来手法[1]より、高精 度に術具を検出可能かつ新たな特徴量を抽出 可能であることから、詳細な手術工程の同定に 有効であると考えられる.今後は、学習データ の増加を行い、提案手法の精度を向上する.

- 謝辞:本研究成果の一部は,国立研究開発法人 情報通信研究機構の委託研究により得 られたものである.
- 利益相反の有無:「なし」

文 献

- [1] 永井智大,佐藤生馬,正宗賢,他:覚醒下脳腫瘍摘出 術における手術工程同定システム.日本コンピュー タ外科学会誌 22 (2): 87-101, 2020
- [2] Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, et al.: High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs. arXiv:1711.11585v2 [cs.CV] 20 Aug 2018.
- [3] T-C Wang, M-Y Liu, J-Y Zhu: High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs. arXiv:1711.11585v2 [cs.CV] 20 Aug 2018.

Microscopic Images Feature Extraction Method using Deep Leaning for Surgical Process Identification in Awake Surgery for Glioma

Ikuma SATO^{*1}, Manabu TAMURA^{*2}, Tomoko YAMAGUCHI^{*2}, Yuichi FUJINO^{*1}, Kitaro YOSHIMITSU^{*2}, Yoshihiro MURAGAKI^{*2}, Ken MASAMUNE^{*2}

*1 Future University Hakodate

*2 Tokyo Women's Medical University

In glioma surgery, it is important to grasp the brain structure and function of the patient and determine the tumor excision area to achieve maximal tumor removal and minimal risk of postoperative complications. However, suitable surgical policies and processes are different depending on the patient's status. Thus, it is necessary for the surgeon to grasp the brain structure and function of the patient during the operation and decide an appropriate tumor extraction region. Because this decision of the surgeon depends on the knowledge and experience of the surgeon, visualization and analysis of surgical processes are important for optimization of surgical staff work and education of young doctors. Therefore, automatic identification of the surgical process is necessary to realize the visualization and analysis of the surgical process. In this research, we propose a feature extraction method from surgical microscope video by deep learning for visualization and analysis of surgical processes.

Key words: Deep learning, Surgical process identification, Surgical tool detection, Awake surgery, Brain tumor

Transformer ベースのモデルを用いた

肺領域における所見文からの疾患名抽出

岡崎 真治*1 林 雄一郎*1 小田 昌宏*2,1 橋本 正弘*3

陣崎 雅弘*³ 明石 敏昭^{*4} 青木 茂樹^{*4} 森 健策^{*1,5,6}

要旨

本稿では、Transformer ベースの BERT モデルを用いた肺領域における所見文の疾患分類を行う方法を提案 する. これは疾患名が付与された画像データセット構築の初期検討である. 画像処理において、一般的に は機械学習モデルの性能向上のため多くの学習データが必要である. しかし、医用画像のアノテーション には膨大な手間がかかる問題がある. そこで、放射線科医が CT 画像の診断結果から作成し、患者の状態を 詳細に記載している所見文を利用する. 本研究では疾患名が付与された画像データセット構築の初期検討 として、Transformer ベースのモデルで Fine-tuning を行い、肺領域において 4 クラス分類を行った. 肺領域 における 400 例の疾患分類実験により Macro-F1 で 79.4%となった.

キーワード:自然言語処理,所見文,深層学習,肺疾患分類

1. はじめに

深層学習は飛躍的な進歩をもたらしており, 性能向上にはアノテーションデータが大量に 必要である.しかし,医用画像のアノテーショ ンには専門知識が必要であり,大量のアノテー ションデータセット構築には膨大なコストが かかる.そこで,医師が CT 画像をもとに,疾 患名,位置,大きさなど患者情報を詳細に記載 した,所見文を利用する研究が盛んに行われて いる[1,2].日本では,国立情報学研究所が医用

*1 名古屋大学大学院情報学研究科 [〒464-8601 名古屋市千種区不老町] e-mail: sokazaki@mori.m.is.nagoya-u.ac.jp *2 名古屋大学情報連携推進本部情報戦略室 *3 慶応義塾大学医学部放射線科 *4 順天堂大学医学部放射線医学講座 *5 名古屋大学情報基盤センター *6 国立情報学研究所医療ビッグデータ研究 センター 画像ビッグデータクラウド基盤[3,4]を構築して おり,クラウド基盤上には所見文,CT 画像など が含まれ研究に活用されている[5].

近年,自然言語処理では,Bidirectional Encoder Representation from Transformers (BERT) [6]と呼 ばれる Transformer ベースの手法で事前学習さ れたモデルを Fine-tuning する手法が取られて きた.事前学習モデルは大量のデータから学習 されており, Fine-tuning に利用すると解決した い課題ごとに少量のラベル付きデータを用い るだけで性能の高いモデルを構築できる. その ため事前学習モデルを Fine-tuning することで, 質問応答,固有表現抽出,分類などで有効な性 能を達成している. そこで、本稿では事前学習 済みの BERT モデルを医用画像ビッグデータク ラウド基盤上の所見文で Fine-tuning し, 肺領域 における疾患分類を行う.そして,分類した所 見文に紐づけられている CT 画像を収集するこ とで,疾患名が付与された画像データセット構 築の初期検討を行った.

表 1	疾患分類の実験結果
2.1	

モデル	Macro-F1(%)
BiLSTM	60.5
tohoku-BERT	79.4

表2 tohoku-BERT の混同行列.

		予測				
		肺癌	間質性 肺炎	COVID -19	正常 所見	
	肺癌	90	7	0	3	
解	間質性 肺炎	0	92	0	8	
비	COVID -19	1	0	92	7	
	正常所見	11	9	32	48	

2. 提案手法で使用する肺領域の所見文

本研究では、医用画像ビッグデータクラウド基 盤上の放射線画像データベースを利用する. 今 回は医用画像ビッグデータクラウド基盤上の データを利用するが、所見文とそれに紐づくCT 画像が格納されたデータベースであればどの ようなものでも構わない. 放射線画像データベ ースは6つの医療施設がデータを提供しており、 所見文とCT 画像が ID により紐づけられてい る. その中で、肺癌、間質性肺炎、COVID-19 の単語を含む所見文を無作為に取得し、正常所 見を含めそれぞれ150 例となるよう疾患の有無 をアノテートした. そして各症例の分布が等し くなるよう180 例、20 例、400 例に分割し、学 習データ、検証データ、テストデータとして使 用した.

3. 提案手法

本研究では、事前学習済みの BERT モデルの 最終層に全結合層を追加し、出力値から、肺癌、 間質性肺炎、COVID-19、正常所見の4クラス分 類をするよう Fine-tuning を行った.

BERT の事前学習では、入力文の先頭に [CLS]トークンを追加し次文予測を行う.この 次文予測により、隣接文の類似性を学習するた め、[CLS]トークンは文の表現を学習できると 言われている.一方 Fine-tuning では、[CLS]ト 2020年05月20日のCTと比較します。両 肺上葉、下葉末梢優位にすりガラス濃度上 昇~consolidationが出現しています。多巣性 分布、辺縁は一部円形で、COVID-19肺炎の 除外を要する像と思われます。右肺上葉 S2 の結節は今回評価困難。右胸膜病変はほぼ 同様。胸水は前回より増加しています。縦 隔に小リンパ節を認めますが非特異的で す。食道全摘術後、胸骨後食道再建術後の 状態。

図1 COVID-19 の正常所見例

ークンに対応する最終層の埋め込み表現を全 結合層の入力として利用する.また,BERT モ デルと全結合層を含めたネットワーク全体の パラメータを更新するよう学習を進めていく.

4. 実験

本研究で用いる日本語 Wikipedia で事前学習 済みの BERT モデルは、東北大学が公開してい るもの(tohoku-BERT¹)を利用した.また実装 には Transformers²を用いた.疾患名の分類実験 の結果を表 1 に示す.ベースラインとして、 Bidirectional Long Short Term Memory (BiLSTM) [7] モデルを使用した.単語ベクトルは CBOW[8]モデルを適用することで獲得してい る.また提案手法の評価指標には Macro-F1 を 用いた.表1より Transformer ベースのモデル は Macro-F1 が 79.4%であり、ベースラインであ る BiLSTM モデルよりも大幅に高い性能となっ ていることが確認できた.

5. 考察

Transformer ベースの手法では BiLSTM モデ ルより Macro-F1 が約 19%高く,有効であるこ とを示した.これは Transformer により文脈を 考慮した学習ができていると考えられるため

¹ https://github.com/cl-tohoku/bert-japanese/

² https://github.com/huggingface/transformers/

である.一方,表2に示す混同行列より,Recall に着目すると,肺癌,間質性肺炎,COVID-19 は90%を超えているが,正常所見は48%と低い 値であった.正解が正常所見であり,誤分類さ れた52例のうち51例は,所見文中に記されて いる疾患名と一致するクラスに分類されてい た.図1に示すCOVID-19の正常所見例では「多 巣性分布、辺縁は一部円形で、COVID-19 肺炎 の除外を要する像と思われます。」のように記 載されている.このように正常所見は,複雑で 学習データに含まれていない記述が多いこと から,BERTの事前学習に所見文を利用し, Wikipedia には含まれない所見文特有の表現を 学習する必要があると考える.

6. むすび

本稿では、疾患名が付与された画像データセット構築の初期検討として、事前学習済みの BERT モデルを Fine-tuning し、肺領域における 疾患分類を行った.疾患分類は Macro-F1 で 79.4%と高い性能を実現し、Transformer ベース のモデルが有効であることを示した.今後の課 題としては、複雑な所見文の学習ができるよう、 BERT の事前学習に所見文を利用し、所見文に 特化した事前学習済みモデルを開発すること が必要となる.

謝辞

日頃から熱心にご討論いただく名古屋大学 森研究室の諸氏に感謝する.本研究の一部は AMEDの課題番号 JP201k1010036, JSPS 科研費 JP17H00867, 堀科学芸術振興財団の研究助成事 業によった.

利益相反の有無

利益相反 無し

文 献

- [1] Biyong J, Wang B, Lyons T et al: Information Extraction from Swedish Medical Prescriptions with Sig-Transformer Encoder. ClinicalNLP Workshop: 41-54, 2020
- [2] Zhang Z, Liu J, Razavian N et al: BERT-XML: Large Scale Automated ICD Coding Using BERT Pretraining. ClinicalNLP Workshop: 24-34, 2020
- [3] 佐藤真一,合田憲人,村尾晃平 他:AMED プロジェクトの概要:クラウド基盤と AI 画 像解析. Medical Imaging Technology 37: 67-71, 2019
- [4] Murao M, Ninomiya Y, Han C et al: Cloud platform for deep learning-based CAD via collaboration between Japanese medical societies and institutes of informatics. SPIE Medical Imaging 11318: 113180T, 2020
- [5] 林雄一郎, 鈴村悠輝, 岡崎真治他:読影レ ポート解析を利用した医用画像データベー スからのアノテーション付きデータセット 作成に関する初期検討. JAMIT: 163-167, 2020
- [6] Devlin J, Chang M, Lee K et al: BERT: Pretraining of Deep Bidirectional Transformers for Language Understanding. NAACL: 4171-4186, 2019
- [7] Schuster M, Paliwal K.K: Bidirectional recurrent neural networks. IEEE Transactions on Signal Processing 45: 2673-2681, 1997
- [8] Mikolov T, Chen K, Corrado G et al: Efficient estimation of word representations in vector space. ICLR Workshop: 2013

Extraction of Disease Names from Radiology Reports

in the Lung Region Using a Transformer-based Model

Shinji OKAZAKI^{*1}, Yuichiro HAYASHI^{*1}, Masahiro ODA^{*2,1}, Masahiro HASHIMOTO^{*3} Masahiro JINZAKI^{*3}, Toshiaki AKASHI^{*4}, Shigeki AOKI^{*4}, Kensaku MORI^{*1,5,6}

*1 Graduate School of Informatics, Nagoya University

- *2 Information Strategy Office, Information and Communications, Nagoya University
- *3 Department of Radiology, Keio University School of Medicine

*4 Department of Radiology, Juntendo University School of Medicine

*5 Information Technology Center, Nagoya University

*6 Research Center for Medical Bigdata, National Institute of Informatics

This paper proposes a method for disease classification of radiology reports in the lung region using a Transformerbased model. This is an initial study on constructing an image dataset with disease names. In image processing, a large amount of training data is generally required to improve the performance of machine learning models. However, medical image annotation is a time-consuming process. To solve this problem, we use radiology reports written by radiologists about diagnosis results from CT images. As an initial study on constructing an image dataset with disease names, we conducted Fine-tuning on a Transformer-based model and performed four-class classification in the lung region. An experiment on disease classification of 400 radiology reports related to lung region showed 79.4% of Macro-F1.

Key words: Natural language processing, Radiology report, Deep learning, Lung disease classification

3次元 CT 画像を用いた模擬ボーリング調査に基づく

脆弱性骨盤骨折検出法による特徴量の検討

山本 侃利^{*1} 藤田 大輔^{*1} Rashedur Rahman^{*1} 八木 直美^{*2}

林 圭吾*3 圓尾 明宏*3 村津 裕嗣*3 小橋 昌司*1

要旨

骨粗鬆症による高齢者脆弱性骨盤骨折は、外傷に因らず、自覚症状が顕著でなく、CT 画像上でその検出が 容易ではないため、発見後の治療が遅れ、転位が進行し、機能的予後回復が得られない場合がある.そこ で、医師の診断能向上のため、CT 画像から骨盤脆弱性骨折を自動的に検出する医師の診断支援システムが 求められている.従来手法では、単純X線画像やCT 画像による2次元画像解析に基づくため、3次元的 に分布する微小な脆弱性骨折の検出が困難であった.そこで我々は、新しい手法として、3次元 CT 画像を 用いて、骨表から骨内部にかけて3次元的に骨折有無を探索するボーリング調査法を模した自動骨盤骨折 検出法(BSFD 法; boring survey based fracture detection)を提案した.本研究では、BSFD 法における特徴量抽 出法について検討する.BSFD 法では、3次元 CT 画像から骨表同値面を求め、同値面上の各点に CT 値で 構成される3次元特徴ベクトルを割り当て、学習済みの3次元畳み込みニューラルネットワーク(CNN) モデルにより、各点において骨折確率を求める.各点でアノテーションされた骨折領域からの3次元 Chamfer 距離から求められた骨折確率を用いて、CNNを学習する.ここで、3次元特徴ベクトルに関して、 領域範囲の拡大を比較検討して、検出性能を評価する.提案手法を110人の被験者のデータで検証した結 果、学習データではAUC0.90、評価データではAUC0.84を確認した. キーワード:CT 画像、骨盤骨折、CNN、医用画像処理、CAD

1. はじめに

骨粗鬆症による高齢者脆弱性骨盤骨折は,外 傷に因らず,自覚症状が顕著でなく,CT 画像上 でその検出が容易ではないため,発見後の治療 が遅れ,転位が進行し,機能的予後回復が得ら れない場合がある.脆弱性骨折の検出重要性が 高い事も報告されている[1].そこで,医師の診 断能向上のため,CT 画像から骨盤脆弱性骨折 を自動的に検出する医師の診断支援システム

*1 兵庫県立大学工学研究科電子情報工 学専攻

〔〒671-2280 姫路市書写 2167〕
e-mail: ei20q029@steng.u-hyogo.ac.jp
*2 姫路独協大学
*2 姫路独協大学

*3 製鉄記念広畑病院

が求められている.

これまでに CT 画像からの骨折検出に深層学 習が応用された成果[2][3][4][5]が数多く報告さ れている.しかし,今までに深層学習による骨 盤における骨折検出[2][3]はされておらず,骨盤 骨折検出の研究においても 2 次元画像の解析 [4][5]にとどまっており,表面上に現れない脆弱 性骨折の自動検出は困難である.そこで我々は, 新しい手法として,3次元 CT 画像を用いて,3 次元畳み込みニューラルネットワークによる 学習手法を提案し,2 値分類における検出精度 は 69.5%を達成している[6].

本研究では、ボーリング調査を模倣した boring survey based fracture detection (BSFD) 法 において、精度向上を目的に特徴量抽出法につ いて検討する.



図 2. BSFD 解析概要図

BSFD 法では、3 次元 CT 画像から骨表同値面を 求め、同値面上の各点に CT 値で構成される 3 次元特徴ベクトルを割り当て、学習済みの 3 次 元畳み込みニューラルネットワーク(CNN)モ デルにより、各点において骨折確率を求める. 各点でアノテーションされた骨折領域からの 3 次元 Chamfer 距離から求められた骨折確率を用 いて、CNN を学習する.ここで、3 次元特徴ベ クトルに関して、領域範囲の拡大を検討して、 検出性能を評価する.

2. 提案手法

図 2 に提案する BSFD 法解析概要図を示す. 以下に各手法の詳細を示す.

2.1 骨表抽出法

本手法では,撮影 CT 画像の骨表 CT 値を点 抽出し,抽出された頂点により骨表 3 次元デー

タを作成する.

2.2 Boring Survey based Fracture Detection (BSFD) 法

作成された各骨表面から骨内部にかけて座 標を算出し,座標点における CT 値をバイキュ ービック補間した 3 次元 CT 値分布を骨折検出 の特徴量とする.

2.3 骨折検出畳み込みニューラルネットワ ークモデル

CNN には抽出した 3 次元特徴量データを入 力し, 次節で定義する骨折確率を出力する回帰 問題として学習を行う.

CNN モデルの中間層には畳み込み層, Max プ ーリング層, Batch normalization を用い, Flatten を用いて最終出力は2ニューロンになるように 調整する.各出力は骨折確率と非骨折確率に相 当する.中間層では Relu 関数,最終出力層では Softmax 関数を活性化関数として用いる.

2.4 骨折確率算出

アノテーションされた骨折箇所からの Chamfer 距離[7]を各ボクセルにおいて算出し, 次式(1)により骨折確率に変換する.なお,骨折 確率を算出する範囲は骨折箇所から 25 ボクセ ル以内とし,その範囲外は全て非骨折,つまり 骨折確率は0と定義する.

$$P_f = \frac{Max(D_f) - D_f}{Max(D_f)}$$
(1)

P_f:骨折確率,D_f:骨折からの距離

3. 実験結果

3.1 使用データ

本研究では、社会医療法人製鉄記念広畑病院 にて撮影された横断面 CT 画像を使用した. CT 画像は全身にかけて撮影され、本研究では骨盤 付近の CT 画像のみを使用した. これらの撮影 断面は axial (体軸断) 面で、スライスあたりの 画素数はすべて 512×512 pixel である. axial 面 とは体軸に垂直な断面であり、人では体を上下 に分ける水平面である. 被験者は金属インプラ

	訓練 / 評価						
feature size	accuracy		precision		recall		
(mm)	(%)		(%)		(%)		
10	81.5	77.2	37.6	30.0	76.0	67.8	
20	88.1	82.3	50.6	39.1	84.9	76.2	
30	85.8	78.9	55.3	42.6	83.0	71.4	
40	84.9	75.8	43.4	29.3	80.8	71.2	
50	78.0	68.8	56.3	43.9	70.6	58.9	

表1 特徴量サイズによる検出精度比較



(a). 評価データ骨折確率分布真値 (b). 評価データ骨折確率分布推定 図 3. 骨折学習済みモデルによる骨折確率推定結果

ントを有しない 110 例(男性 59 例,女性 47 例, 不詳 4 例)である.各被験者は,1か所以上の骨 盤骨折を有する.撮影断面数は約 500 から 1000 枚程度,1ボクセルの大きさ 0.6 から 0.8mm 程 度で,各被験者の体の大きさに合わせて撮影時 に放射線技師により調整された.なお,本研究 の実施は,社会医療法人製鉄記念広畑病院倫理 委員会の承認を受けて行った.

3.2 実験条件

本研究では CNN の学習において最適化手法 には Adam を用い,学習率は 1×10⁻⁴ で学習す る.評価関数には MAE, RSME を使用する.

3.3 実験結果·考察

CNN に入力する特徴量範囲変更による検出 精度向上を検討した.特徴量範囲変更として, 10 mm 四方から 50 mm 四方で 10 mm 間隔で, 抽出制度への影響を調べた.評価では,各参照 点の骨折確率が 0.9 以上である箇所を骨折とし,

0.9 未満である箇所は非骨折と判定し二値分類 での性能評価を行う.評価値として文献[6]で用 いた Accuracy に加え, Precision, Recall を用い, この際の値は ROC 曲線で得られる点で Precision, Recall の精度を最適にする閾値で得 た値を算出する.比較した結果を表1に示す. 訓練データ、評価データどちらにおいても Accuracy, Recall は 20 mm で最も高くなり, Precision は 50mm において最も高い結果となっ た. また,特徴量サイズの拡大につれて, Accuracy, Recall においては 20 mm を起点に精 度が下降しており、Precision においては 40 mm を除いた全体を通して精度が上昇している傾 向がみられる.本研究においては骨折検出正確 性の性能である Precision に比べ, 実際の骨折が 検出される確率である Recall をより重視する. そのため、本研究においては特徴量サイズとし て 20 mm が最適であるといえる.

Accuracy に関して文献[6]において評価デー タに対して 69.5%であったが、本研究において は 82.3%を達成し、12.8%の精度向上を達成し た.これらの結果は抽出特徴量サイズの拡大に よる骨折特徴量の検出範囲拡大が大きく起因 し、加えて深層学習モデル構造の改良による推 定精度の向上が考えられる.

図3には骨折確率を推定した結果を示す.骨 折部に対して予測を大きく外す箇所は少ない が,非骨折箇所に対して確率を大きく予測して しまう箇所が多くみられる.特に仙骨と腸骨の 境界や腸骨と大腿骨の境界等は CT の変化を捉 えやすいため,健常であると判別しにくいと考 えられる.

4 まとめ

本研究は撮影 CT 画像から骨表を抽出し, BSFD 法によって骨表領域における CT 値を解 析し, CNN モデルを用いて骨折確率を算出する 手法を提案し,その特徴量抽出領域の比較検討 を行い,最適化を行った.その結果特徴量サイ ズ 20 mm において評価データにおいて Accuracy は 82.3%, Precision は 39.1%, Recall は 76.2%を達成した.

今後の課題として骨折検出精度の向上に伴 い,他の深層学習モデルの検討や骨折単位での 検出精度や分析を行うことが挙げられる.加え て骨の境界部分である箇所に対しては通常の 非骨折確率ではなく負の値を学習することで 健常な箇所に対してより高い精度が得られる ことが期待できる.

利益相反の有無

なし

文 献

[1] I. Dana Alexa, B. Veliceasa, O. Alexa et al: Importance if CT Scan in Fragility Fracture of the Pelvis, 2020 International Conference on e-Health and Bioengineering (EHB), 2020

[2] A. Bar, L. Wolf, O. B. Amitai et al:

Compression Fractures Detection on CT. Proc. SPIE Vol. 10134. Medical imaging, 2017: Computer-Aided Diagnosis, 1013440.

[3] C. M. Deniz, S. Xiang, R. S. Hallyburton et al: Segmentation of the Proximal Femur from MR Images using Deep Convolutional Neural Networks. Scientific Reports, 2018

[4] S. P. Singh, L. Wang, S. Gupta, H. Goli et al:3D Deep Neural Network on Medical Images. A Review Sensors, 2020

[5] N. Tomita, Y. Y. Cheung, S. Hassanpour et al: Deep neural networks for automatic detection of osteoporotic vertebral fractures on CT scans. Computers in Biology and Medicine, Vol. 98, pp. 8-15, 2018

[6] N. Yamamoto, R. Rahman, N. Yagi et al: An automated fracture detection from pelvic CT images with 3-D convolutional neural networks. 2020 International Symposium on Community-centric Systems (CcS), 2020

[7] M. Martinek, R. Grosso G. Greiner et al: Fast and Efficient 3D Chamfer Distance Transform for Polygonal Meshes. Vision, Modeling, and Visualization, 2011

Discussion on extracted features in fragility fracture detection with

boring survey method using 3D CT pelvis images

Naoto YAMAMAMOTO^{*1}, Daisuke FUJITA^{*1}, Rahman Rashedur^{*1}, Naomi YAGI^{*2}, Keigo HAYASHI^{*3}, Akihiro MARUO^{*3}, Hirotsugu MURATSU^{*3}, Shoji KOBASHI^{*1}

*1 University of Hyogo

*2 Himeji Dokkyo University

*3 Steel Memorial Hirohata Hospital

In the case of fragility fractures in the elderly due to osteoporosis, it is difficult for physicians to recognize the subjective symptoms and detect fractures on CT images, which is likely to lead delayed treatment, progressive dislocation, and no functional recovery. Therefore, it is an essential system that can automatically detect pelvic fragility fractures from CT images to support the diagnosis of physicians and improve the diagnosis ability of them. Previous studies are based on two-dimensional image analysis of simple X-ray images or CT images, making it difficult to detect small fragility fractures distributed in three dimensions. Then, we proposed a new method that uses 3D CT images to search for fractures from the bone surface to the bone interior called the boring survey based fracture detection (BSFD). In this study, we consider the feature extraction method in the BSFD method, which calculates the bone isosurface from 3D CT images, assigns a 3D feature vector consisting of CT values to each point on the equivalence plane, and calculates the fracture probability using a trained 3D convolutional neural network (CNN) model. The fracture probability is calculated at each point using a trained 3D convolutional neural network (CNN) model. A CNN is trained using the fracture probability calculated from the 3D Chamfer distance from the annotated fracture region at each point. We evaluate the detection performance of the proposed method by considering the following 3D feature vectors: intra-bone region, inter-bone region, and extended region. The proposed method was experimented on 110 subjects, and the results showed that the AUC of the training data was 0.90, and the AUC of the evaluation data was 0.84.

Key words: CT images, pelvic fracture, CNN, Medical image processing, CAD

膵臓の超音波内視鏡から作成した

2次元ヒストグラムを用いた膵癌自動分類

竹内 祐慈*1 原 武史*2*3 周 向栄*2 岩佐 悠平*4

上村 晋也*5 三田 直樹*5 岩下 拓司*5 清水 雅仁*5

要旨

現在最も有用性が高い診断方法とされる超音波内視鏡下穿刺吸引生検では、出血や膵炎、穿刺ラインへの 播種など、患者に負担がかかることが知られている.本研究では、撮影された超音波内視鏡から深層学習 を用いて膵癌の診断を行い、患者の負担の少ない膵癌分類を行う.以下の2種類の画像の特徴を組み合わ せることで精度の向上を試みた.1つ目は医師の作成した腫瘍領域をもとに超音波内視鏡映像から腫瘍領 域を切り出し、症例内で最も輝度値の高い1フレームである.2つ目は横軸を時間、縦軸を輝度値、色を 頻度とした2次元ヒストグラムである.前者の画像のみを使って分類した AUC が 0.69 であるのに対し、 本手法で提案した2次元ヒストグラムの特徴を追加したところ AUC が 0.80 に向上した.以上より、膵臓 の超音波内視鏡から作成した2次元ヒストグラムは膵癌の分類に有用であることが示唆された.

キーワード: 膵癌分類, 深層学習, 超音波内視鏡

1. はじめに

膵腫瘍は、 膵癌や神経内分泌腫瘍などの悪 性疾患と自己免疫性膵炎などの炎症性疾患に 分類される. 膵腫瘍の診断では、これらの鑑別 が治療方針を決定する上で重要となる. 超音波 内視鏡下穿刺吸引針生検の診断能は感度 85%、 特異度 98%と非常に高く、現在は膵腫瘍の確定

*1 岐阜大学大学院自然科学技術研究科 知能理工学専攻

〔〒501-11194 岐阜県岐阜市柳戸 1-1〕

e-mail: take.ug1219@gmail.com

*2 岐阜大学工学部電気電子・情報工学科 *3 東海国立大学機構医療健康データ 統合研究教育拠点

*4 岐阜市民病院消化器内科

*5 岐阜大学病院第一内科

診断において最も有用性が高い診断法とされ ている.しかし,採取できる組織が非常に小さ いため,腫瘍の細胞を採取できずに偽陰性とな る可能性があり,また観血的処置であるため, 出血や膵炎などの有害事象の発生,さらには穿 刺ラインへの播種が 3.4%程度生じたとの報告 もあるなど,臨床的に問題となる点が散見され る.本研究の目的は,撮影された超音波内視鏡 から深層学習を用いて膵癌の診断を行い,患者 の負担の少ない診断法を確立することである.

腫瘍内部の輝度値の経時的変化を表した Time Intensity Curve(TIC)による膵腫瘍の診断も されている[1].本研究では、TICのような輝度 値の経時的変化を視覚的に表した2次元ヒスト グラムを作成し、画像認識による分類に有効で あるかを検証した.

2. 試料画像

本研究では膵臓に疾患のある 500~600 フレー ムからなる 131 症例を用いる.内訳は膵癌が 86 症例,膵癌以外が 45 症例である.使用した画像 は 2 種類ある.1 つ目は症例内で腫瘍領域の輝 度値が最も高い1フレーム(図 1)である.これは 医師が作成した腫瘍領域を用いて超音波内視 鏡から腫瘍領域の切出しを行い,腫瘍領域の平 均輝度値の最も高い1フレームを用いた.2 つ 目は,2次元ヒストグラム(図 2)である.このヒ ストグラムは横軸を時間,縦軸を輝度値,色を 頻度としたカラーマップで,撮像された動画内 で腫瘍がどのように造影されていくかを1枚の 画像で視覚的に判断できることが特徴である.



図1 輝度値の最も高いフレーム



図2 2次元ヒストグラム

3. 実験

腫瘍領域の特徴は ResNet50[2]の畳み込み部 分で抽出し、2次元ヒストグラムの特徴は VGG16[3]の畳み込み部分で抽出した.その2つ の特徴をチャネル方向に結合し、全結合層を追 加することで分類を行なった. 誤差関数には Binary Cross Entropy を用いた. 評価は症例数が 少ないため leave-one-out を用いた.

また2次元ヒストグラムの有用性を検証するために,症例内で最も輝度値の高い1フレームのみを ResNet50 で分類する実験も行った.

4. 結果

2 次元ヒストグラムを使用せずに輝度値の最 も高いフレームのみで分類をした結果が AUC0.69,確信度 0.5 以上を膵癌としたときの 分類率が膵癌 69.8%, 膵癌以外 33.3%という結 果だったのに対し,本手法では AUC0.81,確信 度 0.5 以上を膵癌としたときの分類率が膵癌 80.2%, 膵癌以外 71.1%という結果となった.

膵癌以外の症例は膵癌に比べて造影のピーク の持続時間が長いため、経時的な特徴を持つ2 次元ヒストグラムの特徴を与えることで、膵癌 以外の分類精度が向上したと考えられる.

5. まとめ

超音波内視鏡映像の輝度値の経時的変化を 表す2次元ヒストグラムは、その特徴を用いな いモデルより高い分類率を得たことから、膵癌 の分類に有用であることが示唆された.今後は 医師のラベルを使用しない手法を検討し、医師 の負担の少ない画像診断を行う.

謝辞

本研究を進めるにあたり,有益なご助言をい ただきました研究室の方々に感謝の意を表し ます.

利益相反の有無

なし

文 献

- [1] Matsubara H, Itoh A, Kawashima H, et al,
 "Dynamic quantitative evaluation of contrast-enhanced endoscopic ultrasonography in the diagnosis of pancreatic diseases" *Pancreas*. Oct 2011;40(7):1073-1079.
- [2] Kaimimng He, Xiangyu Zhang, Shaoqing Ren, et al.: Deep Residual Learning for Image Recognition. arXiv:1512.03385
- [3] Karen Simonyan, Andrew Zisserman.:Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv: 1409.1556v6

Automatic Classification of Pancreatic Cancer Using

2D Histograms Generated from Pancreas Ultrasound Endoscopy

Yuji TAKEUCHI^{*1}, Takeshi HARA^{*1*2}, Zhou XIANGRONG^{*1}, Yuhei IWASA^{*3}, Shinya UEMURA^{*4}, Naoki MITA^{*4}, Takuji IWASHITA^{*4}, Masahito SHIMIZU^{*4}

*1 University of Gifu

*2 Tokai National Higher Education and Research System Medical and Health Data Integration Research and Education Center *3 Gifu Municipal Hospital *4 Gifu University Hospital

Endoscopic Ultrasound-Fine Needle Aspiration, which is currently considered the most useful diagnostic method, is known to cause bleeding, pancreatitis, and dissemination to the puncture line, all of which are burdensome to patients. In this study, we use deep learning to diagnose pancreatic cancer from ultrasound endoscopy image and classify pancreatic cancer with less burden to patients. I tried to improve the accuracy by combining the features of the following two types of images. The first is a frame with the highest brightness in the case. The second is a 2D histogram with time on the horizontal axis, luminance values on the vertical axis, and color as frequency. The AUC of classification using only the former images was 0.69, whereas the addition of 2D histogram features improved the AUC to 0.80. These results suggest that the 2D histogram generated from ultrasound endoscope images of the pancreas is useful for classification of pancreatic cancer.

Key words: Pancreatic Cancer Classification, Deep Learning, ultrasound endoscop

多段の画素異常検知による FDG-PET/CT 上の

がん骨転移候補検出

山口 明乃*1 根本 充貴*2 甲斐田 勇人*3*4 木村 裕一*1

永岡 隆*1 山田 誉大*4 花岡 宏平*4

北島 一宏*5 槌谷 達也*6 石井 一成*3*4

要旨

FDG-PET/CT 画像診断の支援技術として、マハラノビス距離を用いた異常検知と one-class support vector machine (OCSVM) による 2 種類の異常検知を骨画素に適用したがん骨転移候補の検出法を提案する.提案法ではまず CT 値と SUV の 2 次元空間上で正常骨画素からのマハラノビス距離を計測し、距離の閾値処理により骨転移が疑われる領域を粗抽出する.次に CT 値, SUV, CT 値と SUV の平均曲率とガウス曲率、マハラノビス距離の計 7 特徴量を入力とした OCSVM による異常検知を行う.得られた異常の極大点を骨転移候補とする.FDG-PET/CT 画像 44 例を用いて学習した異常検知処理を 46 の骨転移病変を含む FDG-PET/CT 画像 20 例に適用した結果、感度 80.4%(38/46)、FP 数 314.6 点/例の性能が得られた.今後の課題は、病変検出感度の改善、骨領域の抽出方法の改善、過検出候補点の削減処理の追加による性能向上である.

キーワード:コンピュータ診断支援,FDG-PET/CT,マハラノビス距離, one-class SVM, 異常検知

*1 近畿大学大学院生物理工学研究科 〔〒649-6493 紀の川市西三谷 930〕
e-mail: yamaguchi.h917@gmail.com
*2 近畿大学生物理工学部
*3 近畿大学医学部放射線医学教室放射線
診断学部門
*4 近畿大学高度先端総合医療センター
PET 分子イメージング部

- *5 兵庫医科大学放射線医学講座
- *6 兵庫医科大学病院放射線技術部

1. はじめに

近年の医用画像を用いた画像診断技術の向 上により,1回の検査で全身を撮像した際の 患者1人あたりの医用画像データ量が非常に 多くなっている.その結果,診断に要する時 間の増加に伴い読影医の身体的及び精神的負 担が増加しており,見落としや診断結果のば らつきなどによる診断精度の低下が懸念され ている.また,診断には読影医の経験量や主 観が影響するため,診断精度を統一すること も非常に重要である.そこで,病変候補の自 動検出を行い,セカンドオピニオンとして医 師に情報を提示する画像診断支援(CAD)シ ステムへの期待が高まっている[1].本研究で は、解剖学的形態情報を示す CT と局所的糖 代謝情報を示す FDG-PET (以下 PET)を併用 した FDG-PET/CT (以下 PET/CT)の CAD シ ステムに着目する.PET/CT は,悪性腫瘍の 発見手段として普及し臨床的には非常に有用 である.しかし,全身を対象とした撮像とな るため,すべての病変に対しての読影が必要 となり,読影にかかる負担が大きい.

PET/CTのCADシステムの研究例[2-4]では, その大多数が有病症例を機械学習した多クラ ス識別処理を用いることで病変検出を行って いる. 寺本らの研究[2]では, CT を特殊なフ ィルタにかけ, PET の閾値処理により初期病 変候補検出を行い, ルールベース識別や Support Vector Machines (SVM) によって過検 出 (False positive : FP) の削減処理を施すこ とで肺病変の病変を検出していた. Rui らの 研究[3]では、テクスチャー解析により多数の 特徴量を算出し、その中から6個の適切なパ ラメータを客観的に選択し SVM に組み合わ せることで軟部組織病変の悪性・良性の区別 を行っていた. Jilong らの研究[4]では, 閾値 処理及びラベリング処理により、肺病変の関 心領域を抽出し、その領域内でテクスチャー 解析により特徴量を算出し, k 近傍法を使用 して肺病変と正常組織を区別していた.

我々はこれまで、健常症例のみを用いた異 常検知処理に基づく PET/CT 上の頚胸部領域 の原発性病変と、リンパ節転移性病変の自動 検出 CAD システムについて検討を行ってき た.医用画像では、疾患によっては有病デー タの収集は容易ではなく、CAD システムの開 発において常に十分な量の学習データが得ら れるとは限らない.提案法は検診等で収集が 比較的容易である健常データのみを学習に用 いるため、システムの開発が容易である.先 行研究では one-class SVM (以下 OCSVM)を 用いて画素異常度計測による頚胸部領域・骨 領域病変の強調[5]画像から、病変候補点を検 出していた.骨転移病変候補の検出では、病 変検出感度 91.1%、過検出候補点 653 点/症例 という性能が得られているが、検出精度の改善(過検出の低減)が求められている.本研究さらなる精度向上を目的とし、本研究ではマハラノビス距離と OCSVM の2種類の画素 異常検知処理を段階的に組み合わせた骨転移 病変候補検出法を提案する.提案法の有効性 を検証するため、臨床 PET/CT を用いた実験 を行う.

2. 手法

2.1 提案手法

提案する手法を図1のフローチャートに示 す.



図1 骨転移病変の候補点検出処理.

本提案法は、1) 骨領域の抽出、2) 関心領 域の粗抽出、3) 異常度画像の生成、4) 極大 点の算出による病変候補画素抽出の4段階で 構成されている.

1) 骨領域の抽出

CT と PET の解像度が異なるため,スケー リング処理により共に 2.40 mm 解像度に等方 化する.補間法には,バイキュービック補間 を用いる.次に CT 値の閾値処理及びモルフ オロジー処理により骨領域の抽出を行う.

頭部領域は,歯科治療に伴う金属製インレ ー等によるメタルアーチファクトの影響を考 慮し,今回は対象外とした.

2) 関心領域の粗抽出

抽出された骨領域の各画素で健常骨画素からのマハラノビス距離を算出する.距離の推定に使用する正規分布モデルの作成には,健常 PET/CT データにおける骨領域内の CT 値とSUV の2特徴量を使用する.全骨画素でのマハラノビス距離算出後,距離値降順の蒸気420000 voxelを閾値処理により抽出し,骨転移と疑われる領域(粗抽出関心領域)とする. この閾値順位は実験的に定めた値であり,す

べての病変候補領域を必ず1画素以上検出す ることができる値である.

3) 異常度画像の生成

抽出された粗抽出関心領域画素において, OCSVMによる異常度検知を行う.OCSVM学 習には,健常 PET/CT データにおける骨領域 内の正常画素からのマハラノビス距離を計測 し,距離値が大きい上位420000 voxelのHU, SUV,HUとSUVのそれぞれの平均曲率とガ ウス曲率およびマハラノビス距離値の計7 特徴量を用いる.HUとSUVの曲率は空間的 3次元に輝度値軸を加えた4次元空間上の 輝度曲率であり,周辺の輝度勾配分布の情報 から得られる特徴量である[6].

OCSVM では ν と γ の 2 種類のハイパーパ ラメータを最適な値に設定する必要がある. ν は以下のソフトマージン最大化の目的関数 の式(1)で示される[7].

 $\min_{w \in F, \xi \in \mathbb{R}^{n}, b \in \mathbb{R}} \frac{1}{2} \|w\|^{2} + \frac{1}{\nu n} \sum_{i} \xi_{i} - b$ (1)

また,本研究ではカーネル関数として RBF カ ーネルを使用している.以下の RBF カーネル の式(2)に示されるのが γ である[8].

 $K(\mathbf{x}_i, \boldsymbol{\theta}) = exp(-\gamma || \mathbf{x}_i - \boldsymbol{\theta} ||^2)$ (2) 使用したハイパーパラメータはそれぞれ $\gamma =$ 10, $\nu = 0.01$ である. 4) 極大点算出による病変候補画素抽出

異常度画像を,スケールσ=1.0 及び2.0 の ガウス関数で畳み込んだ平滑化画像を作成す る.それぞれの平滑化画像において,ラスタ スキャンにより注目画素を中心とした 26 近 傍を比較し,注目画素が最大値となる画素を 病変候補点として検出する.病変候補点が両 方のスケールの平滑化画像の同座標に存在す る場合は,2つの病変候補点を統合する.

2.2 実験試料

本実験では、兵庫医科大学病院(以下兵庫 医大病院)及び、近畿大学病院(以下近大病 院)にて撮像された全身 FDG-PET/CT データ 64 症例を使用する.うち、健常症例は兵庫医 大病院で20症例、近大病院で24症例であり、 有病症例は兵庫医大病院で10 症例、近畿大学 病院で10 症例(骨転移病変計46 病変含む) であった.撮像機器は兵庫医大病院がGE Discovery IQ,近大病院がGE Discovery PET/CT 710 scanner である.2 施設のPET/CT の解像度について以下の表1に示す.

表12 施設の PET/CT 解像度.

	近大病院	兵庫医大病院
СТ	0.977×0.977×3.27	0.977×0.977× 3.26
[mm]	1.37×1.37×3.27	1.17×1.17×3.26
PET	2.60×2.60×3.27	2.60×2.60×3.26
[mm]	3.65×3.65×3.27	3.13×3.13×3.26

3. 結果

病変候補点検出結果と平均過検出画素数を 以下の表2に示す.

表2病変候補点検出結果

	感度	平均過検出数		
	[%]	[点/症例]		
提案手法	84.8	322.5		
先行研究[5]	91.1	653.0		

病変候補点検出成功例と失敗例を図2と図3に示す. (a)はCT, (b)はPET, (c)は候補点 画像を示す.



(a)CT
 (b)PET
 (c)候補点画像
 図2病変候補点検出成功例.病変は胸椎 Th10 転移であり,画像中の赤丸でそれぞれ示す.



 (a)CT
 (b)PET
 (c)候補点画像
 図3 病変候補点検出失敗例.病変は第5肋骨転移 であり,画像中の実線の赤丸でそれぞれ示す. 点線の赤丸は候補点の検出できなかった病変 部分を示す.

4. 考察

表2及び図2より,平均過検出数が先行研 究の結果よりも約50%削減でき,病変部分の 画素候補が検出できていることから,提案手 法の有用性が確認できた.

病変候補点が検出できなかった病変は 46 病変中 7 病変であった. このうち 6 病変は OCSVM を用いた異常検知による異常度画像 の生成の時点では検出されていたが,その後 の極大点の算出で失敗し,病変候補点として 検出されなかった. この原因として,病変サ イズが非常に小さいことと,病変の周囲に他 の病変や異常度の高い過検出画素が存在する ことが考えられる.この6病変は46病変の 中でも病変サイズが特に小さいものであり, 抽出された異常度画素も少なくなっていた. そのため,病変候補点の算出で周囲の他の病 変や過検出画素の影響を大きく受け,病変候 補点が算出されなかったと結論付けた.病変 候補点の算出方法の見直しにより,感度の向 上が期待できると考えられる.

図3に示す検出失敗例は,骨領域の抽出で 失敗し,病変画素が抽出できなかった例であ る.この病変は溶骨型の骨転移であり,CT 値が低下してしまい,骨領域抽出の際の閾値 処理で漏れてしまったためであると考えられ る.転移病巣の骨の反応型によらない骨領域 抽出法を用いることにより,感度の向上につ なげることができると考えられる.

5. 結論

マハラノビス距離と One-class Support Vector Machine の2種類の画素異常検知処理 を段階的に組み合わせた骨転移検出法を提案 した.提案法の有効性を検証するため,臨床 PET/CT を用いた実験を行った.骨領域にて 感度 84.8% (39/46 病変),平均過検出数 322.5 点/症例という結果が得られた.平均過検出数 が先行研究の結果よりも約 50%削減するこ とができたため,提案手法の有効性が確認で きた.今後の課題は,病変候補点の算出方法 と骨領域の抽出方法の見直しによる感度の向 上と平均過検出数のさらなる削減.検出した 病変候補点を基準点とした病変領域の推定と 高精度な識別である.

謝辞

本研究の遂行において,医用画像ビューワ VOLUME-ONEの使用,及び画像処理ライブ ラリの使用をご快諾いただいた,広島市立大 学の増谷佳孝教授に深く感謝申し上げます. 本研究の一部は,JSPS 科研費 17K17680 お よび 17H05284,近畿大学学内研究助成金 SR06 により行われた.

利益相反の有無

利益相反なし.

文 献

- [1] Ronald A. C: Computer aided detection(CAD): an overview. Cancer Imaging, 5(1):17-19, 2005.
- [2] 寺本篤司: PET/CT 画像を用いたコン ピュータ支援診断技術, 医用画像情報学 会雑誌 31(2), 2014 年.
- [3] Rui Xu : Texture analysis on 18F-FDG PET/CT images to differentiate malignant and benign bone and soft-tissue lesions, The Japanese Society of Nuclear Medicine(2014) 28, pp.926-935.
- [4] Jilong Z: Computed Diagnosis System for Lung Tumor Detection based on PET/CT Images, 3rd International Conference on Biomedical Engineering and Informatics, 2010, pp. 166-170.

- [5] Tanaka A, ,et al.: Automatic detection of cervical and thoracic lesions on FDG-PET/CT by organ specific one-class SVMs. Int J Computer Assisted Radiology and Surgery 15(Suppl 1): S208-S209.
- [6] 平野靖:4次元超曲面の曲率を用いた
 3次元濃淡面像に対する追跡型細線化の
 一手法,電子情報通信学会論文誌 D-II
 vol.J83 D-II.
- [7] Bernhard S, Robert W, Alex S et al: Support Vector Method for Novelty Detection. Proceedings of the 12th International Conference on Neural Information Processing Systems, 582-588, 1999
- [8] Han S, Cao Q, Han M: Parameter selection in SVM with RBF kernel function.
 World Automation Congress 2012, 24-28, 2012
Detection of bone metastases on FDG-PET/CT

using multistage voxel anomaly detection

Haruno YAMAGUCHI^{*1}, Mitsutaka NEMOTO^{*2}, Hayato KAIDA^{*3*4}, Yuichi KIMURA^{*1}, Takashi NAGAOKA^{*1}, Takahiro YAMADA^{*3}, Kohei HANAOKA^{*3}, Kazuhiro KITAJIMA^{*5}, Tatsuya TSUCHITANI^{*6}, Kazunari ISHII^{*3*4}

*1 Graduate School of Biology-Oriented Science and Technology, Kindai University

*2 Faculty of Biology-Oriented Science and Technology, Kindai University

*3 Division of Positron Emission Tomography Institute of Advanced Clinical Medicine, Kindai University

*4 Department of Radiology, Kindai University Faculty of Medicine

*5 Department of Radiology, Hyogo College of Medicine

*6 Department of Radiology, Hyogo College of Medicine Hospital

We propose a method to detect bone metastases on FDG-PET/CT images by combining a Mahalanobis distance-based anomaly voxel detection and an anomaly voxel detection using a one-class support vector machine (OCSVM). At the first step of the proposed method, the Mahalanobis distance-based anomaly detection with CT value and SUV is applied to each bone voxel. The bone voxels with a large distance are extracted coarsely as suspicious regions of bone metastasis. Next, the anomaly detection using the OCSVM with seven voxel features (CT value, SUV, mean curvatures and Gaussian curvatures of CT value and SUV, and Mahalanobis distance in the last step) is applied to each voxel within the suspicious regions. The local maximum points of anomaly degree calculated by the OCSVM are detected as candidates of bone metastasis. The proposed metastasis detection was trained by 44 volumes of normal FDG-PET/CT cases and was evaluated by 20 volumes of FDG-PET/CT cases with 46 bone metastasis lesions. The evaluation result showed the 80.4 % (38/46) sensitivity with 314.6 false positives / case. Future works improve the lesion detection sensitivity, improve bone area extraction, and develop a false positive candidate reduction.

Key words: CAD, FDG-PET/CT, Mahalanobis distance, one-class SVM, anomaly detection

PSA density 算出のためのセグメンテーションによる

前立腺体積計測法の検討

安倍 和弥*1 武尾 英哉*1 永井 優一*2 黒木 嘉典*3 縄野 繁*4

要旨

現在,医療の様々な分野において画像支援診断(CAD: Computer Aided Diagnosis)の研究・開発が行われている.また,AI技術を用いた臓器・病変検出においても積極的な検討が行われている.

前立腺がんは欧米人に羅漢率の高い病変であり、日本においても 2025 年には男性の羅漢率1位の病変と なると予想されている.その発見には PSA 測定を用いた検査の精度が高く用いられているが、PSA 値は前 立腺肥大症などでも高値が出ることがある.そこで、感度の向上を目的として PSA 値を前立腺の体積で除 する PSA density(PSAD)を用いている.しかし、現在の PSAD 算出に用いる体積は楕円体法により楕円に見 立てて縦径、横径、上下径から算出するやや正確性に欠けるものである.

本研究では、セグメンテーションの手法として DeepLab を用いた領域抽出手法により前立腺の体積を計 測, PSAD 算出に用いる手法について研究を行った.これにより従来法よりも正確な体積算出と PSAD 算 出が行え、前立腺がんのスクリーニング精度向上が期待できる.64 例で DeepLab の学習を行い 188 例で評 価を行った結果、生検データでは 58% だったスクリーニング精度が、DeepLab を用いることで 63% へと向 上する結果となり、本手法の有効性が確認できた.

キーワード:前立腺,体積計測,DeepLab, PSA density, セグメンテーション

1. はじめに

近年, 医療の現場では CT や MRI など撮影 装置の高精細化により生成される医用画像の 量が増大している.撮影される画像の枚数は 患者一人あたり数百枚にもおよび,それを読 影する医師の負担も急増している.そのため 診断の補助を行うコンピュータ画像支援診断

(Computer Aided Diagnosis/Detection: CAD) の開発が非常に重要視されている[1]. その CAD の研究分野においても人工知能, AI を 用いた研究が進められている. その中でも画 像認識手法である CNN (Convolutional Neural

*1 神奈川工科大学工学部電気電子情報 工学科

〔〒243-0292 厚木市下荻野 1030〕
e-mail: abex0930@ele.kanagawa-it.ac.jp
*2 国立がん研究センター東病院
*3 にいむら病院
*4 新松戸中央総合病院

Network)は積極的に用いられるようになった [2].

前立腺がんは欧米人に羅漢率の高い病変で あり,日本においても2025年には男性の羅漢 率1位の病変となると予想されている.その 発見には一般的に PSA (Prostate-Specific Antigen)値が用いられ,数値が高い場合に前 立腺がんが疑われるが,それ以外にも前立腺 肥大症や前立腺炎などでも高値が出る場合が ある.そのため,感度の向上を目的としてPSA 値を前立腺の体積で除して算出する PSAD

(PSA density)が用いられている[3].しかし, 現在 PSAD 算出に用いられている体積は,前 立腺を楕円体法により楕円に見立てて縦径, 横径,上下径から算出するやや正確性に欠け るものである. PSAD が 0.15 を超えるとがん の疑いがあると評価される.

本研究では、深層学習の一種である DeepLabを用いたセグメンテーションにより、 前立腺自体を抽出,それをもとに体積を計算 し PSAD 算出に用いる手法を検討する.これ により,従来法よりも高精度での PSAD 算出 が行え,前立腺がんスクリーニングの精度向 上が期待できる.

セグメンテーションの手法については様々 な手法が提案されている. 代表的なものでも FCN や U-Net, PSPNet などがある.本研究で は、Chen らが提唱した DeepLab v3[4]に着目 し、それを用いた前立腺セグメンテーション 法を検討した. DeepLab v3 は v2 より組み込 まれたアトラス空間ピラミッド型プーリング (ASPP: Atrous Spatial Pyramid Pooling)を改良 した, 改良型 ASPP に特徴がある. まず ASPP はサイズの異なった4つのカーネルを使い, 得られる4種の特徴を並列にプーリングを行 い集約することにより広範囲の特徴を密に集 約している. 改良型は ASPP のカーネルのサ イズを変更し、さらに画像プーリングの層を 並行して追加することにより、特徴マップの 高精度化を測っている. この改良は物体境界 部分での精度向上が見られるため、周辺との 情報差異が乏しい前立腺の抽出に有効的であ ると期待できる.

前立腺の認識においては,他部位を含めて であるが,周ら[5]が CT を用いて CNN で行 っている.ただ近年 MRI の機能向上により前 立腺の局所診断に非常に有効とされているた め,MRI から体積を求める本手法が有効であ ると考えられる.

本論文のポイントを以下に示す.

 DeepLab を用いた前立腺のセグメンテー ション

② 上記検出結果を用いた PSAD の精度向上本論文では、2章で DeepLabを用いた前立腺セグメンテーションについて、3章で提案手法でのセグメンテーション結果と PSAD 算出結果を示す、4章にて考察、5章にて本研究のまとめについて述べる。

本論文で使用する画像および生検データは, TCIA (The Cancer Imaging Archive) のデータ セット(Prostate-MRI-US-Biopsy)[6]を用いて いる. 生検データは, 検査及び Artemis biopsy system を用いて算出されたものであり, 検討 に用いた PSA 値は検査データ, 体積はこのシ ステムで仮想的に算出されたものである.

また,学習データおよび評価データで用い た症例画像は,すべて前立腺がんのある症例 である.

2. 前立腺の体積計測

DeepLab で用いる学習データの条件を表 1 に示す.

私 I 于目 /	アの木口
データセット数	64
画像枚数	19~72[枚]
画像サイズ	$256{ imes}256$ \sim
	432×432[pixel]
スライス厚	1.5~3.0[mm]

表1 学習データの条件

2.1 前処理

本研究で使用した画像はMRIのraw画像で あるため,画像に明暗があり研究に用いるデ ータとして不適格なデータであった.そのた め前処理として濃度とコントラストの正規化 を行い,画像を均一にする調整を行っている.

本研究で用いるラベル画像は手動にて前立 腺を抽出した画像を使用した.まず医師の確 認の下で,数例のスケッチを行った.以降は その時の指摘を参考に筆者が抽出を行った.

前立腺領域が含まれていない画像を除外した結果,学習に使用する画像数は1680枚となった.元画像とラベル画像の例を図1に示す.



図1 ラベル画像の作成例

2.2 DeepLab を用いた領域抽出処理

今回用いた DeepLab は Chen らの DeepLab 構造[4]を参考に構築した.開発環境は Anaconda を用いた Python. Torchvision 内の事 前学習モデルである deeplabv3_resnet101 をも とに転移学習を行った.出力は前立腺領域の みなので1チャンネルに変更, resnet101 モデ ルであるため特徴ベクトルを 2048 とした.

2.3 前立腺体積計測と PSAD 算出

DeepLab の検出結果には,前立腺と連結し ていない微細領域も発生するため,まず微細 オブジェクトの除去を行った.その画像から 総ピクセル数を算出,スライス厚と画素間隔 の2乗を乗算することにより体積を算出する. その体積と PSA 値から PSAD を算出,生検デ ータと比較する.

3. 実験と結果

本実験で評価データとして用いる画像デー タの条件を表2に示す.

データセット数	188
画像枚数	20~112[枚]
画像サイズ	$256 imes256$ \sim
	432×432[pixel]
スライス厚	1.5~3.0[mm]
PSA 値	4.0~124.5[ng/ml]

表2 評価データの条件

評価データの症例は PSA 値がグレーゾー ンの下端といわれる 4[ng/ml]以上であるもの を対象とした.

3.1 DeepLab の教師付き学習

2 章で述べた画像を用いて DeepLab の学習 を行った.パラメータとしては,入力画像を 256×256 に均一化,学習率はデフォルトの 0.0001,バッチサイズは 16,エポック数は 25 で学習を行った.学習時の最終的な正解率は 学習データで 98.8%,検証データで 95.2%で あった. 3.2 体積計測と PSAD によるがんスクリー ニング結果

本手法と生検データでの体積及びがんスク リーニング結果を表3に示す.本手法と生研 データでの PSAD ヒストグラムの比較を図2 に、PSAD の基準である PSAD=0.15 に着目し たグラフを図3に示す.緑と紫はともにスク リーニング結果が変わらなかったもの.青は 本手法ではスクリーニングできなくなってし まったもの、赤が本手法でのみスクリーニン グに成功したものである.

表3より,体積の計測結果は,生検データ と比較して4章で後述する1例を除き,最大 38.7%の増加,最小で79.8%の減少となった. また,平均すると12.9%の減少となった. PSAD によるスクリーニングの結果は,生検 データで188 例中109 例の58%だったのに対 し,本手法では119 例の63%と精度の向上が 見られた.

生検データの体積を基準に3領域に分けた 結果を表4に示す.前立腺は一般的に健康で あれば20ml 程度といわれているが,多少の 個人差や高齢になると健常者でも肥大が始ま るなどを加味し,40ml 未満を肥大がないと思 われるグループ,40ml 以上かつ60ml 未満を グレーゾーン,60ml 以上は肥大も疑われるグ ループとして,各々結果の評価を行った.す べての領域で本手法でのスクリーニング精度 の向上が確認できる.

表3 生検データと本手法の比較

	生検データ	本手法						
休巷	$14.4 \sim$	14.2~						
14~1俱	222.4[ml]	160.0[ml]						
体積平均	51.5[ml]	46.2[ml]						
	$0.05\sim$	$0.05 \sim$						
FSAD	2.51[ng/ml]	2.90[ng/ml]						
PSAD 平均	0.259[ng/ml]	0.288[ng/ml]						
スクリーニ	58%	63%						
ング精度	(109/188)	(119/188)						

表4 肥大を考慮した体積ごとの性能比較

前立腺体積	生検データ	本手法
40ml 未満	86.4% (70/81)	91.4% (74/81)
40ml 以上	50.0%	52.2%
60ml 未満	(23/46)	(24/46)
60ml D/ E	26.2%	34.4%
の加め上	(16/61)	(21/61)



図 2 本手法と生研データでのPSAD 値ヒストグラムの比較



図3 閾値 0.15 付近での生検データと 本手法での PSAD 値の比較

図2より,細かな差ではあるが全体的に本 手法の方が PSAD 値を高く算出できている結 果となった.

図3より, PSAD が0.15 付近で精度の向上 が見られた.3 例検出できなくなったデータ があったものの,13 例は本手法で検出が可能 となり精度の向上が確認できる.

抽出結果の例を2例,図4に示す.赤枠の スケッチ領域に対し,右の白い領域が本手法 での抽出領域である.どちらも本手法によっ てスクリーニングが可能となったものである. 抽出はしっかりと行えており,本手法での前 立腺体積の抽出精度向上が寄与している例で ある.



図4 セグメンテーション成功例

4. 考察

3 章の表 3 より, DeepLab を用いた体積計 測でのスクリーニング精度の向上が確認でき る.また検出結果の画像を見ても,前立腺領 域が検出できていることが確認できる.

スクリーニングの精度は生検データの58% に対して, DeepLab を用いた手法では63%と 判別精度の向上が見られた.

また,3章の表4より肥大を考慮した体積 のグループ分けでの評価結果でも全領域で本 手法での性能向上が確認できる.PSAD は体 積に極端に左右される指標でもあるため,こ の結果からもセグメンテーションにより楕円 体法より細かく体積を算出できることの有効 性が確認できる. 体積計測については、1 例検出不足による 著しい体積減少が見られたが、この例を除け ば最大 38.7%の増加、最小で 79.8%の減少と いう計測結果となった.

増加したものの要因としては,図5の緑枠 で示すように前立腺領域が外に向けて張り出 しており,ここを抽出してしまったため体積 が増加したことが考えられる.こちらの症例 に関しては,PSA が極端に高くPSADによる スクリーニングには影響がなかった.また, 楕円体法では近似しにくいこういった形状の ものに対しても検出できうる可能性が確認で きた.



図5 侵食している前立腺

1 例,数スライスだけではあるがあからさ まな抽出不足が見られた.その例を図6に示 す.前立腺領域が特に不鮮明になっている例 であり,この症例ではこのスライスを含む 3 スライスで抽出ができていないため体積値が 低くなっていた.また,前述のような著しい 抽出不足とはならなかったものの,図7に示 すように明らかに特異な形状に膨張している ものが前立腺領域と判断されず,全体を抽出 できなかったことによる計測体積の減少が見 られた.今回はこの検出不足による体積減少 がPSAD値算出に影響はなかったものの,検 出不足によるPSAD値上昇は本意で無いため 対処が必要となる.

これらに対しては、より多彩な学習データ を用いて開発を行うことにより検出精度は向 上するものと考えられる.



図6 不鮮明で抽出ができなかった例



図7 肥大化した前立腺

5. まとめ

本論では、DeepLab を用いての前立腺の体 積計測とそれを利用した PSAD によるがんス クリーニング精度向上について研究を行った. DeepLab を用いた体積計測により、楕円体法 より精密な前立腺形状抽出が可能となり、結 果 PSAD によるスクリーニングが、生検デー タでの 58%から 63%へと向上が見られた.こ れにより、本手法での前立腺体積計測の有効 性と PSAD 精度向上への有効性が示唆された.

利益相反の有無

利益相反 なし

文 献

[1] 土井邦雄:医用画像とコンピュータ
 支援診断 ~現状と将来の可能性~.映像情報メディア学会誌,65(4),pp.427-431,2011

- [2] Robertson S, Hartman J, Robertson S et al: Digital image analysis in breast pathology-from image processing techniques to artificial intelligence. Translational Research Vol.194, pp.19-35, 2018
- [3] 松岡光明:前立腺癌診療ガイドライン 2016 年版.メディカルレビュー社, 大阪, 2016, pp.63-73
- [4] Chen L, Papandreou G, Schroff F, et
 al: Rethinking atrous convolution for
 semantic image segmentation.
 arXiv:1706.05587, 2017

- [5] 周向栄,藤田広志:深層学習に基づく CT 画像からの複数の解剖学的構造の
 同時自動認識と抽出. Med Imag Tech, Vol.35(4), pp.187-193, 2017
- [6] Natarajan S, Priester A, Margolis D, et al: Prostate MRI and Ultrasound With Pathology and Coordinates of Tracked Biopsy (Prostate-MRI-US-Biopsy) [Data set]. The Cancer Imaging Archive. DOI: 10.7937/TCIA.2020.A61IOC1A, 2020

A study of prostate volume measurement by segmentation

for PSA density calculation

Kazuya ABE*1, Hideya TAKEO*1, Yuuichi NAGAI*2, Yoshihumi KUROKI*3, Shigeru NAWANO*4

- *1 Kanagawa Institute of Technology
- *2 National Cancer Center Hospital East
- *3 Niimura Hospital
- *4 Shinmatsudo Central General Hospital

Currently, research and development of Computer Aided Diagnosis (CAD) is being conducted in various fields of medicine. In addition, AI technology is also being actively studied for organ and lesion detection. Prostate cancer is a disease that has a high incidence among Westerners, and it is expected to become the most common disease among men in Japan by 2025. PSA measurement is used to detect this disease with high accuracy, but PSA levels can also be high in benign prostatic hyperplasia. Therefore, PSA density (PSAD), which is the PSA value divided by the volume of the prostate, is used to improve sensitivity. However, the current volume used to calculate PSAD is somewhat inaccurate, as it is calculated using the ellipsoid method, based on the vertical, horizontal, and vertical diameters of the ellipse. In this study, we investigated a method for measuring prostate volume and calculating PSAD using a segmentation method based on DeepLab. This method is expected to improve the accuracy of screening for prostate cancer by providing more accurate volume and PSAD calculation than conventional methods. As a result of training DeepLab in 64 cases and evaluating it in 188 cases, the screening accuracy increased from 58% using biopsy data to 63% using DeepLab, confirming the effectiveness of this method.

Key words: prostate, volume measurement, DeepLab, PSA density, segmentation

Synthesized Perforation Detection from Endoscopy Videos Using

Model Training with Synthesized Images by GAN

Kai Jiang^{*1}, Hayato Itoh^{*1}, Masahiro Oda^{*1}, Taishi Okumura^{*2} Yuichi Mori^{*2}, Masashi Misawa^{*2}, Takemasa Hayashi^{*2}, Shin-Ei Kudo^{*2}, Kensaku Mori^{*1,3,4}

Abstract

Endoscopic submucosal dissection (ESD) is a standard treatment for early gastric cancer. Although ESD is already a remarkably reliable treatment, a high-incident perforation still can cause peritonitis. To prevent potential peritonitis, there is a need for a Computer-Aided Diagnosis (CAD) system, since accurate perforations detection is an essential characteristic in the practical CAD system. We aim to construct an accurate perforation detection model using machine learning techniques and a large dataset of perforation scenes. For ESD videos collection and annotation, physicians have to spend much time. We only collected ESD videos with the annotation of 17 patients at this point. Therefore, in our dataset, the number of perforation images extracted from ESD videos is limited. To achieve a better classification accuracy with the limited-size dataset, we present the Generative Adversarial Network-based data augmentation of the training dataset. To evaluate this data-augmentation method, we used a four-layer convolutional neural network for perforation-scene classification. We trained the perforation detection models with the original and augmented training dataset achieved higher accuracy, recall, precision, and AUC score than the model training by original training dataset.

Keywords : CAD, Deep learning, Colonoscopic video, GAN, Perforation detection

1. Introduction

Gastric cancer is the fourth most general cancer and the second leading cause of death in all cancers, which has to be a widespread health problem in the world [1]. In 2012, about 720,000 patients were killed by gastric cancer [1]. Endoscopic submucosal dissection (ESD) [2] is a treatment for early gastric cancer, which allows the submucosa's direct dissection, and large lesions can be resected en bloc without the restriction of cancer size [3]. Although ESD is already a remarkably reliable treatment, it still associated with a high incidence of perforation [3]. In ESD, the flush knife may accidentally pierce the colonic wall and generate a perforation on it. If perforation happened during ESD, the patient might need emergency open surgery since perforation can easily cause peritonitis [4].

*1 Graduate School of Informatics,

Nagoya University [Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan]

- e-mail: kjiang@mori.m.is.nagoya-u.ac.jp
- *2 Digestive Disease Center, Showa University Northern Yokohama Hospital

^{*3} Information Technology Center, Nagoya University

^{*4} Research Center for Medical Bigdata, National Institute of Informatic

Our research project aims to construct a Computer-Aided Diagnosis (CAD) system that prevent perforations by predicting perforation in ESD. In addition, accurate perforations detection is also helpful for preventing expansion of existing perforations to prevent the potential peritonitis in the ESD. Furthermore, to construct the CAD system, we need to analyze many ESD videos that contain many perforation-exist scenes. However, physicians have to spend a much time collecting and annotating ESD videos. We only collected ESD videos with the annotation of 17 patients at this point. Therefore, it is very challenging to achieve accurate perforation prediction due to the limited-size dataset. This limited data with machine learning techniques makes it very hard to construct a perforation detection model. Currently, perforation detection research progresses slowly; automatic image-based perforation detection is very challenging with imbalanced datasets because of limited perforation-exist images. Data augmentation is a standard preprocessing method for the limited-size dataset. The effect of classical data augmentation methods for a small number of perforation scenes is limited. The Generative Adversarial Network (GAN) [5] is a network for estimate generative models, which can learn to generate images with similar features as the training set. This characteristic is helpful for augmented features in the training dataset. This paper presents a method to augment the training dataset by using GAN to generate perforation-exist images.

2. Method

To construct an accurate perforation detection model, we present the GAN-based data augmentation of training images. Figure 1 shows an overview of the perforation-exist image synthesizing network using generative models. We set $X_i \in \mathbb{R}^{H \times W \times 3}$, $i = 1, 2, ..., n_i$ to be images extracted from ESD scenes. We set $Y_j \in \mathbb{R}^{H \times W \times 3}$, $j = 1, 2, ..., n_j$ to be images extracted from perforation scenes in ESD videos. In addition, we define a random noise vector $z = (z_k) \in \mathbb{R}^{100}$, $z_k \sim N(0,1)$, whose elements are randomly sampled from the normalized Gaussian distribution N(0,1) with i.i.d. condition. We define a prior probability distribution over a noise vector as p(z). We set $p(X_i)$ to be a prior probability distribution over images extracted from ESD scenes, and $p(Y_j)$ to be a prior probability distribution over images extracted from perforation scenes. We define a four-layer perceptron G(z) that generates an image of an artificial perforation scene from an noise vector z. For an input image $I \in \mathbb{R}^{H \times W \times 3}$, we define two three-layer perceptrons D_x and D_y as follows,

$$D_x(I) = P_x,\tag{1}$$

$$D_{y}(I) = P_{y}, \tag{2}$$

where $D_x(I)$ outputs a likelihood P_x of $I \sim p(X_i)$, and $D_y(I)$ outputs a likelihood P_y of $I \sim p(Y_j)$. The discriminator D_x determines whether the generated image is similar to ESD scenes. The discriminator D_y determines whether the generated image has perforations or not. We use two discriminators to ensure that the generated image is as realistic as the ESD scene, while preventing the generated image from being the same as the image in the training data set. With a prior probability distribution p(I) over image I, we define an adversarial loss

$$L_{GAN}(G,D) = \mathbb{E}_{I \sim p(I)}[\log D(I)] + \mathbb{E}_{G(z) \sim p(z)}\left[\log\left(1 - D(G(z))\right)\right],\tag{3}$$

which evaluates how accurately a given discriminator D classifies real and fake images, where fake images are



Fig. 1 Overview of the perforation-exist image synthesizing network.

generated by G. For two discriminators D_x and D_y , we set two losses $L_{GAN}(G, D_x)$ and $L_{GAN}(G, D_y)$. By using these losses, we define the min-max problem

$$\min_{G} \max_{D_{\mathcal{X}}, D_{\mathcal{Y}}} L(G, D_{\mathcal{X}}, D_{\mathcal{Y}}) = L_{GAN}(G, D_{\mathcal{X}}) + L_{GAN}(G, D_{\mathcal{Y}}).$$
(4)

In this min-max problem, we maximize the classification ability of the discriminators by training D_x and D_y to correctly discriminate fake ESD and fake perforation images, respectively. Furthermore, we minimize the discrimination performance of the discriminators by training *G* to generate realistic perforation images of ESD scenes. The first and second losses contribute to generate realistic ESD-image and perforation-image generations, respectively.

3. Experiment and Result

We collected ESD videos of 17 patients in the digestive center of Showa University Yokohama Northern Hospital with IRB approval. Expert endoscopists annotated perforation-existing scenes of these videos. Based on this annotation, we extracted perforated and non-perforated images from colonoscopic videos by 30 and 1 fps, respectively. We collected 74,714 no-perforated and 29,754 perforated images from ESD videos. We resized all extracted images to 224 × 224 pixel with Lanczos interpolation. We split all resized images of 17 patients into training, validation, and test dataset

Dataset	Number of non- perforated images	Number of perforated images
Training	57,341	20,903
Validation	2,809	2,817
Test	16,564	6,034

Table 1 Number of images of training, validation and test data



(b)

Fig. 2 Comparison of original perforation-exist images and generated perforation-exist images. (a) perforation-exist images generated by our perforation-exist image synthesizing network. (b) original perforation-exist images.

without the duplication of patients. Table 1 shows the number of extracted images of training, validation, and test data. Note that an individual patient image was assigned to a unique set. For experiments, we use the perforation-exist image synthesizing network generated 3,305 images for the detection model training. We added all generated images in training dataset. We defined training dataset with generated images as augmented dataset, and training dataset without generated images as original dataset. Figure 2 shows the comparison of original perforation-exist images and generated images examples.

We constructed a four-layer convolutional neural network with a global average pooling layer as a detection model. We trained the perforation detection model by using the augmented dataset and original dataset, respectively. We used an Adam as an optimizer for this detection model and set the learning rate to be 1.0×10^{-6} . The model was trained from scratch on NVIDIA Tesla V100 PCIe 32 GB with CUDA 10.0, and the size of a minibatch was 512. For data augmentation, we applied flipping in both vertical and horizontal directions.

	1		U	
Dataset	Accuracy	Recall	Precision	AUC
Real and generated images	0.789	0.328	0.621	0.837
Real images	0.769	0.308	0.612	0.819

Table 2 Detection performance of detection models training with different dataset



Fig. 3 ROC curve in perforation image detection.

To evaluate the performance of the perforation detection model, we introduced the accuracy, the recall, the precision, the Receiver Operating Characteristic (ROC) curve and the Area under the ROC curve (AUC) score as quantitative evaluation values. Table 2 reports all quantitative evaluation of the perforation detection models trained by different training data. Figure 3 shows the ROC curve drawn by these sensitivity and 1-specificity values of perforation detection models. The detection model trained by the augmented training dataset achieved 0.789 accuracy and 0.823 AUC score.

5. Discussion

The model training by augmented training dataset achieved higher accuracy, recall, precision, and AUC score than the model only training by original training dataset in detecting perforations in scenes of the colonoscopic video. Figure 3 shows that the model training by augmented dataset has better AUC than the model training by original dataset. In addition, sensitivities and 1-specificities of the ROC curve of the model training by augmented dataset are better than the model training by original dataset. Table 2 reports that accuracy, recall, and precision are no significant improvement. This result implies the insufficient size of generated data. We will use more generated data for training in the future. Furthermore, we can not assure that each generated image has a perforation. Thus, we consider using a trained YOLOv3 [6] as a discriminator to ensure the perforation generation in the future.

6. Conclusions

This work presents the GAN-based data augmentation of training images for the training perforation detection model in the colonoscopic video to construct a CAD system for ESD. We extracted 104,468 images from 17 colonoscopic videos and split them into training, validation and test dataset for the experiment. We generated 3,305 perforation-exist images for the augmented dataset. We evaluated detection models training by different datasets, respectively. The model

training by augmented dataset achieved higher accuracy, recall, precision, and AUC score than the model training by original dataset.

Competing interests

None

Acknowledgement

Parts of this research was supported by JST CREST, MEXT JSPS KAKENHI (21K19898, 17H00867, 17K20099), the JSPS Bilateral International Collaboration Grants, the AMED (19hs0110006h0003) and the Hori Sciences & Arts Foundation.

References

- [1] Ferlay, J., Steliarova-Foucher, E., et al. Cancer incidence and mortality patterns in europe: estimates for 40 countries in 2012.
 European journal of cancer, 49(6): 1374-1403, 2013
- [2] Probst A, Schneider A, Schaller T, et al. Endoscopic submucosal dissection for early gastric cancer: are expanded resection criteria safe for Western patients?. Endoscopy, 49(09): 855-865, 2017
- [3] Gotoda, T., Yamamoto, H., Soetikno, R.M.: Endoscopic submucosal dissection of early gastric cancer. Journal of gastroenterology, 41(10): 929-942, 2006
- [4] Saito Y, Uraoka T, Yamaguchi Y, et al.. A prospective, multicenter study of 1111 colorectal endoscopic submucosal dissections (with video), Gastrointestinal Endoscopy 2010, 72(6):1217-1225, 2010
- [5] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks[J]. arXiv preprint arXiv:1406.2661, 2014
- [6] Redmon J, Farhadi A. Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767, 2018

Vascular Structure Segmentation in Stereomicroscope Image

Yunheng WU^{*1}, Masahiro ODA^{*2,1}, Yuichiro HAYASHI^{*1} Takanori TAKEBE^{*3,4,5} and Kensaku MORI^{*1,6,7}

Abstract

This paper proposes a scheme to segment vascular structures from low-contrast stereomicroscope images. In recent years, researchers have cultivated some three-dimensional tissues in-vitro which are similar to human organs in biological functions called organoid. Organoids can be used in many fields such as drug development and regenerative medicine. Because stereomicroscope can observe living organisms and tissues without cutting into slices, it is often used to observe organoids in culture. However, it will take researchers a lot of time and energy to observe organoid culture and transplantation. It is necessary to automatically analyze stereomicroscope images to assist researchers to improve their research efficiency. The blood vessel in a cranial window is an important factor to analyze organoid transplantation results. Therefore, in this paper, we propose a scheme to segment vascular structure of cranial window in low-contrast stereomicroscope images. We preprocess a stereomicroscope image to remove noises, and then enhance contrast of the image so that the vascular structure which is difficult to be observed by the naked eye in the original image can also be well observed. Then we use a weighted line detector to extract the vascular structure from the image. The experimental results showed that the vascular structure can be well segmented from a low-contrast and noisy image. This study enables researchers to better observe the vascular structure in low-contrast stereomicroscope images and lays a foundation for the subsequent research for stereomicroscope images in organoid research.

Keywords : Stereomicroscope images, Vascular Structure Segmentation, Organoid, Line detector

1. Introduction

In medical research, research samples of human organs are not easy to obtain, which is an obstacle for researches about human organs. However, in recent years, researchers have cultivated in-vitro three-dimensional tissues that have similar functions as human organs, which is called organoid [1]. Organoids play an important role in drug development, disease simulation, and regenerative medicine [2]. In [3], Takebe et al. effectively regenerated avascular elastic cartilage from in-vitro cultured mesenchymal condensation, which reappears in the early stage of chondrogenesis. After transplanting vascularized condensed progenitor cells into the mice's cranial window because of the optical access [4], a unique intravital monitoring system like using stereomicroscope can be provided to track maturation of cartilage and

^{*1} Graduate School of Informatics, Nagoya University [Furo-cho, chikusa-ku, Nagoya, 464-8601, Japan]

e-mail: yunhengwu@mori.m.is.nagoya-u.ac.jp

^{*2} Information and Communications, Nagoya University, Japan

^{*3} Institute of Research, Tokyo Medical and Dental University (TMDU), Tokyo, Japan

^{*4} Division of Gastroenterology, Hepatology & Nutrition, Developmental Biology, and Center for Stem Cell and

Organoid Medicine (CuSTOM), Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA

^{*5} Department of Pediatrics, College of Medicine, University of Cincinnati, Cincinnati, OH, USA

^{*6} Information Technology Center, Nagoya University, Japan

^{*7} Research Center for Medical Bigdata, National Institute of Informatics, Japan



Fig. 1 Vascular structure in mice's cranial window was observed by stereomicroscope after cartilage organoids were transplanted. As shown in figure, the contrast between the vascular structures and the background is low. Some small vascular structures are difficult to be observed.

it can allow us to dissect the previously uncharacterized phenomenon in cartilage organ cultivation [3][5]. In this process of chondrogenesis, blood vessels in the mice's cranial window play a certain role in maturation of cartilage organoids. Therefore, it is necessary to observe the blood vessels in the cranial window after transplantation of vascularized condensed progenitor cells by stereomicroscope. Stereomicroscopes can be used to observe living tissues. However, the blood vessels in the cranial window are located in the cerebral cortex. Therefore, in the stereomicroscope images, the color of blood vessels is similar to the background, resulting in a low-contrast problem. Some blood vessels are not easy to be observed, especially those very small blood vessels, as shown in Fig. 1. If researchers need to analyze a large number of low-contrast stereomicroscope images, it will cost them a lot of time and energy. Therefore, in order to improve the efficiency of organoid research, a method that can automatically extract vascular structures from stereomicroscope images is needed. In the image processing field, image segmentation is an important task. We can extract the region we want to observe in the image and lay a foundation for the subsequent image registration. When stereomicroscope images can be registered with other microscope images, different information captured by different microscopes can be fused, which can help researchers to analyze organoids.

In this paper, we propose a scheme to extract the vascular structure in the mice's cranial window from a low-contrast stereomicroscope image after transplantation of vascularized condensed progenitor cells. For preprocessing, we use the Bilateral filter to remove noise and a Contrast Limited Adaptive Histogram Equalization (CLAHE) to enhance contrast. Next, we extract a green channel image to further improve the contrast between the vessel and the background. Finally, we utilize a weighted line detector method to extract the vascular structure.

2. Method

1) Overview

As shown in Fig. 2, we use some effective methods to extract vascular structure in mice's cranial window from a lowcontrast stereomicroscope image. Our method for automatic vascular structure segmentation is mainly divided into two parts: image preprocessing and line detection. Firstly, the noise in the image is removed, and then the contrast of the image is enhanced and the green channel of the image is extracted. After preprocessing, we use a weighted line detector to extract the linear structure of blood vessels and achieve the automatic vascular structure segmentation from a lowcontrast stereomicroscope image. In the following part, we will introduce each part of the automatic vascular structure segmentation method for low-contrast stereomicroscope images in detail.

2) Preprocessing

Noise Removal. When a stereomicroscope image is taken, the sensor and other factors will lead to noises in the image.



Fig. 2 Workflow of our scheme for vascular structure segmentation in low-contrast stereomicroscope image.

Some filters can be used to smooth the image. However, they also make the edge of some small vascular structures more blurred. If small blood vessels become more blurred, it is not conducive to the next vascular structure segmentation. Considering the problem, in this paper, we use the Bilateral filter for stereomicroscope denoising. Bilateral filter [6] is a kind of nonlinear filtering method. Bilateral filtering considers not only the spatial proximity of the image but also the similarity of pixel values. There is the weight of pixel value in the Bilateral filter, the pixel values are closer, the weight is greater. Therefore, near the edge, the pixels far away will not affect the pixels on the edge too much, and the pixels with a great difference on both sides of the edge will not be smoothed too much. In this way, the edge can be well preserved while the noise is removed.

Contrast Enhancement. In the stereomicroscope image of mice's cranial window, the contrast of intensities between the vessel and the background is low. Contrast enhancement will seriously affect subsequent segmentation. Therefore, we use the Contrast Limited Adaptive Histogram Equalization (CLAHE) method [7] to enhance the contrast of the image. The histogram of low-contrast images is concentrated in a certain range, which is not conducive to distinguish the foreground and background. Therefore, histogram equalization is used to enhance contrast. CLAHE is a kind of adaptive histogram equalization (AHE). Compared with the general histogram equalization, AHE can calculate multiple histograms which corresponding to different sub-regions of an image to change image contrast. However, AHE can enhance the contrast but also enlarge noises in images. Before applying adaptive histogram equalization, CLAHE clips the height of the histograms of sub-regions to make the height lower than the preset threshold, and then the pixels that exceed the threshold are evenly distributed in the histogram. In this way, contrast of image can be relatively enhanced and noises are not amplified at the same time.

Green Channel Extraction. By splitting the three channels (RGB) of the image, we can find that the contrast between the blood vessel and background is more obvious in the green channel image, and the vascular structure is more prominent than images of blue and red channels. In [8], E. Ricci et al. converted images into the green channel to extract retinal blood vessels and achieved an excellent result. Therefore, in this study, we use the green channel image to extract the vascular structure in mice's cranial window.

3) Line detector

In [8] and [9], line detector method based on a linear operator is proposed for retinal fundus vessels segmentation. The line detector method can segment linear region well. Hence, it has a good performance in retinal fundus vessels segmentation task. The blood vessels in mice's cranial window are similar to retinal fundus vessels in structure. Therefore, we use a weighted line detector [9] to extract the blood vessels in mice's cranial window from a low-contrast stereomicroscope image.

Next, we will introduce how the weighted line detector segments the vascular structure. This method is built based on



Fig. 3 The results after each processing can be shown as (a) input image, (b) image after noise removal, (c) image after contrast enhancement, (d) green channel image, (e) segmented image, and (f) ground-truth image.

the idea that the blood vessels regions have higher pixel value than the surrounding region after green channel extraction. A square window of size L is defined for every pixel in the green channel image, and then the average pixel value in the square window is calculated as I_A . Next, an average pixel value is calculated along each line in the square window and then these lines in the square window are through the center pixel of the window at 12 different orientations, and angle between the lines is 15 degrees. According to [9], in fact, pixels which are closer to the center pixels play an important role in judging whether the center pixel belongs to a vessel. Therefore, weight is defined according to the distance between the center pixels and other pixels in a line, and the average pixel value on each line is calculated based on the weight. These pixels which are closer to the center pixel will be given greater weight when calculating the average pixel value on the lines. The line with the highest average pixel value among the 12 lines is defined as 'winning line' and the highest average pixel value in the line is defined as I_M . Finally, we make a mask image with the same size as the green channel image, and then the difference between pixel value of the 'winning line' and the average pixel value of the square window is assigned to a pixel in the mask image whose position corresponds to the position of center pixel of the square window in the green channel image, and the difference is defined as $I_D = I_M - I_A$. If the central pixel belongs to a blood vessel, the winning line is parallel to the blood vessel, and I_D is larger, because the difference between the pixel value of the blood vessel and the background is very large. On the contrary, when the central pixel does not belong to a blood vessel, I_D is smaller. In this way, we can distinguish the blood vessels and the background to extract the blood vessel structure. As in [9], we combined the I_D of each pixel in the mask image with the pixel value of the green channel image. Finally, we perform normalization and set a threshold t to determine whether pixels belong to blood vessels.

3. Experiments and Results

The data used in the experiment is taken by a camera on a stereomicroscope after vascularized condensed progenitor cells were transplanted into the mice's cranial window. The size of the raw image is 5184×3456 pixels after imaging. In order to reduce the interference to the experiment, we used a mask to block unnecessary areas, such as the mouse hair and bone cement. Next, we cut some blacked areas to reduce computation costs. Finally, the size of the image used in the experiment was 2900×2800 pixels. In addition, ground-truth image as shown in Fig. 3 (f) was made by a technologist.



Fig. 4 Many small blood vessels in the input image are difficult to be observed by naked eyes, but they can correctly be segmented and some non-vascular areas can be identified. (a) is the input image. Column (b) are sub-images cut from the input image, and column (c) are sub-images of segmented image corresponding to column (b), respectively. Column (d) are sub-images of ground-truth image corresponding to column (c), respectively. Three images in every row, respectively, are sub-images cut from the same areas of the input image, segmented image, and ground-truth image.

For our experiment, firstly, we used Fig. 3 (a) as an input image. Figure 3 (b) was obtained after noise removal, and then the contrast enhancement operation was performed to obtain Fig. 3 (c). Through contrast enhancement operation, blood vessels were easier to be observed. Then, a green channel image is extracted as shown in Fig. 3 (d). The blood vessels in the green channel image are distinct and it is easier to observe some small blood vessels. Finally, the green channel image was input into the weighted line detector to obtain a segmented image as shown in Fig. 3 (e). With regard to the parameter setting, the size of window *L* was set as 55, and *t* was set as 0.12.

We used some performance metrics to evaluate our scheme by comparing the segmented image with ground-truth image. Firstly, we used dice coefficient to evaluate segmentation results. Dice coefficient is mainly used to measure the similarity between the segmented image and ground-true image. Dice coefficient is defined as

$$D = \frac{2|\mathbf{X} \cap \mathbf{Y}|}{|\mathbf{X}| + |\mathbf{Y}|},\tag{1}$$

where X and Y represent prediction and the ground-truth mask, respectively, while |X|, |Y| represent the number of pixels on foreground (vessel, in our case), for prediction mask and ground-truth mask, and $|X \cap Y|$ indicates the number of intersection pixels of foreground between prediction mask and ground-truth mask. Next, we used precision and recall to evaluate the experimental results. Precision represents the percentage of pixels predicted as blood vessels that are true blood vessels. Recall represents the percentage of true blood vessels that are successfully predicted. In this study, dice coefficient is 0.76. The precision is 0.66, and recall is 0.91.

4. Discussion

As shown in row (1) of Fig. 4, our scheme can successfully segment small blood vessels which are not easily observed by naked eyes in a low-contrast stereomicroscope image. As shown in row (3) of Fig. 4, our scheme can successfully identify the circular blocked area. The edges of these areas are not mistakenly segmented as blood vessels. However, as shown in row (2) of Fig. 4, although our scheme can extract the blood vessels connected with the blocked area, it may be mistakenly identified that the linear blocked area is a blood vessel. This is because the weighted line detector used in our scheme is based on a linear operator to segment blood vessels. In the future, we will further improve our scheme to solve this problem.

5. Conclusion

The vascular structure segmentation in mice's cranial window from stereomicroscope image is of great significance for organoid analysis. In this paper, we proposed a scheme to extract the vascular structure from a low-contrast stereomicroscope. This scheme is based on contrast enhancement and line detector method, which does not request a lot of images to train a model, and just a single image can finish the vascular structure segmentation task. The results showed that the method can effectively extract the vascular structure in a low-contrast stereomicroscope image, even some small vessels can also be extracted. In addition, our scheme can identify some non-vascular areas in the image.

In the future, we will solve the problem that some linear occluded areas are wrongly segmented. In addition, based on the results of blood vessels segmentation, 3D vessels could be reconstructed, which is more beneficial to analyze the role of cerebral cortex blood vessels in organoid culture and transplantation.

Competing interests

None

Acknowledgement

This work was supported by JST Moonshot R&D Grant Number JPMJMS2033.

References

- [1] Rossi G, Manfrin A, Lutolf M P: Progress and potential in organoid research. Nature Reviews Genetics 19(11): 671–687, 2018
- [2] Garreta E, Kamm R D, Chuva de Sousa Lopes S M, et al.: Rethinking organoid technology through bioengineering. Nature Materials 20(2): 145–155, 2021
- [3] Takebe T, Kobayashi S, Suzuki H, et al.: Transient vascularization of transplanted human adult-derived progenitors promotes self-organizing cartilage. Journal of Clinical Investigation 124(10): 4325–4334, 2014
- [4] Takebe T, Sekine K., Enomura M, et al.: Vascularized and functional human liver from an iPSC-derived organ bud transplant. Nature 499: 481–484, 2013
- [5] Takebe T, Zhang RR., Koike H, et al.: Generation of a vascularized and functional human liver from an iPSC-derived organ bud transplant. Nature Protocols 9: 396–409, 2014
- [6] Tomasi C, Manduchi R: Bilateral filtering for gray and color images. Proceeding of Sixth International Conference on Computer Vision: 839-846, 1998
- Zuiderveld K: Contrast Limited Adaptive Histogram Equalization. In: Heckbert P S eds. Graphics Gems IV. Academic Press Professional, San Diego, 1994, pp474–485
- [8] Ricci E, Perfetti R: Retinal Blood Vessel Segmentation Using Line Operators and Support Vector Classification. IEEE Transactions on Medical Imaging 26(10): 1357-1365, 2007
- [9] Zhou C, Zhang X, Chen H: A new robust method for blood vessel segmentation in retinal fundus images based on weighted line detector and hidden Markov model. Computer Methods and Programs in Biomedicine 187: 105231, 2020

深度情報を利用した FCN による

腹腔鏡映像からの血管領域自動抽出の検討

榎本 圭吾^{*1} 林 雄一郎^{*1} 北坂 孝幸^{*2} 小田 昌宏^{*3,1} 伊藤 雅昭^{*4} 竹下 修由^{*4} 三澤 一成^{*5} 森 健策^{*1,6,7}

要旨

本稿では,深度情報を利用した腹腔鏡映像からの血管領域自動抽出手法について述べる.腹腔鏡下手術は 低侵襲であるため患者への負担は小さいが,手術難易度が高く術者にとっての負担は大きい.そのため, 術者を支援するシステムの開発が求められている.特に,血管は手術を進めるうえで重要な体内構造であ り,損傷により出血や合併症を引き起こす可能性があるため,腹腔鏡映像から血管領域を自動抽出する研 究が行われている.従来の血管領域自動抽出手法では,FCNによって RGB 画像から色情報に基づき血管 領域を抽出していた.本稿では,腹腔鏡映像から深度情報を推定し,RGB-D 画像を用いた FCN による血管 領域抽出手法を提案する.実験の結果,従来手法と比べて,鉗子とのオクルージョンがみられるシーンの 抽出精度が向上した.

キーワード: 腹腔鏡下手術,血管領域抽出,深度推定,RGB-D セグメンテーション

1. はじめに

がんによる死亡率は年々増加しおり,死因の 中でがんが占める割合も高い値となっている [1-3]. がんの外科的治療法の一つとして腹腔鏡 下手術があり,低侵襲で患者への負担が小さい

*1	名古屋大学大学院情報学研究科
[=	〒464-8601 名古屋市千種区不老町 IB
電	子情報館南棟4階466号室〕
e-m	nail: keno@mori.m.is.nagoya-u.ac.jp
*2	愛知工業大学情報科学部
*3	名古屋大学情報連携推進本部情報戦
略量	Ê
*4	国立がん研究センター東病院
*5	愛知県がんセンター
*6	名古屋大学情報基盤センター
*7	国立情報学研究所医療ビッグデータ
研究	究センター

という点において注目されている.しかし一方 で、内視鏡から得られる映像が二次元であるこ とや鉗子操作が制限されていることによる手 術難易度の高さから、術者にとっては負担の大 きい手術である.そのため、術者を支援するシ ステムの開発が求められている.

腹腔鏡下手術において,血管領域は手術工程 を理解する手がかりとなる体内構造物である. また,損傷により出血や合併症を引き起こす可 能性もある.さらに,血管は脂肪等の組織に覆 われている場合が多いため,術中に術者が血管 領域を見極めることは簡単ではない.そこで, コンピュータにより血管領域を自動で抽出し, 術中にリアルタイムで術者に提示する技術が 必要である.

従来の血管領域自動抽出手法[4]では,血管の 境界部分や鉗子とのオクルージョン領域,出血 領域において安定した抽出ができていないと いう問題点があった.そこで提案手法では,血 管や鉗子の持つ立体的な構造に注目して,RGB の色情報に加えて深度情報を利用することで 血管領域自動抽出を行う.

2. 提案手法

提案手法では,腹腔鏡画像から深度画像の生成を行い,色情報と深度情報を組み合わせた RGB-D の4 チャンネル画像を作成する.そして,RGB-D 画像を入力とした深層学習モデルにより血管領域自動抽出を行う.

まず深度情報の生成は, CycleGAN [5] により RGB 画像ドメインから深度画像ドメインへの 変換器を学習することにより行った.データセ ットには実腹腔鏡画像と公開データセット[6] である仮想腹腔鏡画像および仮想腹腔鏡深度 画像を利用した.ここで,仮想腹腔鏡深度画像 における濃度値と距離値の関係は不明である. なお,公開データセットの画像のみで学習を行 うと,仮想腹腔鏡画像から仮想腹腔鏡深度画像 への変換を学習してしまうため, RGB 画像ドメ インには仮想腹腔鏡画像に実腹腔鏡画像を加 えたデータセットを利用した.そして,これに より得られた学習済みの変換器を用いて,腹腔 鏡画像を対応する深度画像へ変換することに より深度情報の生成を行った.

次に血管領域自動抽出は、CycleGAN により 生成した深度情報の値を RGB 画像の 4 チャン ネル目に追加することで作成した RGB-D 画像 を利用して、この画像を入力とした 4 チャンネ ル入力の深層学習モデルにより行った.深層学 習モデルには、従来手法[4]において用いられて いた U-Net [7] に dilated convolution [8] を行う dilated convolution 層を追加した Dilated U-Net を 利用した.ここで、U-Net および dilated convolution 層は、Dilated LSTM U-Net [9] の構 造を参考にした.

3. 実験ならびに結果

本実験で利用した腹腔鏡映像は,37 症例の腹 腔鏡下S状結腸癌切除術または腹腔鏡下前方切 除術の映像である.血管領域自動抽出の実験で は,37 症例の腹腔鏡映像から,国立がんセンタ

表 1: Dilated U-Net による抽出の認識精度

入力画像	適合率 (%)	再現率 (%)	F值(%)
RGB 画像	60.9 ± 0.6	45.7 ± 1.3	47.0 ± 0.9
RGB-D 画像	61.1 ± 0.8	45.0 ± 1.0	47.5 ± 1.2







(d)

図 1: 生成された深度画像と血管領域自動抽出の結果 の一部. (a) は正解ラベル画像と元画像の合成画 像, (b) は CycleGAN により生成された深度画像, (c) は RGB 画像を利用したときの抽出結果, (d) は RGB-D 画像を利用したときの抽出結果を表す.

ー東病院の医師の協力のもとアノテーション を行った合計 2566 枚の腹腔鏡画像を利用した. また,認識精度の評価には 5 分割交差検証法に より適合率および再現率, F 値を算出し, RGB 画像を利用した場合と RGB-D 画像を利用した 場合で比較を行った. なお,学習毎の認識精度 の誤差を考慮するためにそれぞれの実験で学 習を 10 回ずつ行い, その平均を認識精度とし た.

実験の結果として,評価用データに対する認 識精度を表1に,生成された深度画像および抽 出結果である予測ラベルを元の腹腔鏡画像と アルファブレンドした画像を図1にそれぞれ示 す.表1から,深度情報を利用することにより 認識精度がわずかに向上することが分かった. また,図1の抽出結果を見ると,特に血管の境 界部分や鉗子とのオクルージョン領域におい て正しく抽出できていることが知られた.

4. 考察

図1を見ると、血管の境界部分や血管と鉗子 の境界部分が正しく距離値の差として表され ていることが分かる.そのため,深度情報が追 加された RGB-D 画像を利用した場合では, RGB 画像を利用した場合と比べてより正確な抽出 が可能となっているのだと考えられる.また, 図1右列のような出血のみられるシーンにおい て RGB-D 画像を利用した場合では正解領域の 一部を抽出することに成功している.しかし, 生成された深度画像を見ると,出血領域の深度 情報は明らかに誤っており、この深度情報を頼 りに抽出がされているとは考えにくい.この画 像においては血管領域の抽出に成功している が, 誤った深度情報が抽出結果に影響している 画像もみられ、結果として RGB 画像を利用し た場合との認識精度の差が小さくなったと考 える.

5. むすび

本実験では、RGBの色情報に加えて深度情報 を利用することにより血管領域自動抽出を行 った. RGB-D 画像を利用した場合には、RGB 画 像を利用した場合と比べて血管の境界部分や 鉗子とのオクルージョン部分において抽出精 度の向上がみられた.しかし、深度情報が不正 確な場合、認識精度にも影響が及ぶという問題 も知られた.そのため、今後は誤った深度情報 に対しても有効な抽出が可能となるよう、深層 学習モデルを工夫することが必要であると考 える.

謝辞

本研究は JST CREST JPMJCR20D5, JSPS 科研 費 17H00867, 21K19898 の支援を受けた.

利益相反の有無なし

文 献

- [1] 厚生労働省:平成 29 年(2017)人口動 態統計(確定数)の概況.
 https://www.mhlw.go.jp/toukei/saikin/hw/j inkou/kakutei17/index.html(2021年6月 6日閲覧)
- [2] 厚生労働省:平成 30年(2018)人口動 態統計(確定数)の概況.
 https://www.mhlw.go.jp/toukei/saikin/hw/j inkou/kakutei18/index.html(2021年6月 6日閲覧)
- [3] 厚生労働省:令和元年(2019)人口動態
 統計(確定数)の概況.
 https://www.mhlw.go.jp/toukei/saikin/hw/j
 inkou/kakutei19/index.html(2021年6月
 6日閲覧)
- [4] 盛満慎太郎,小澤卓也,北坂孝幸,他.: 腹腔鏡動画像からの Fully Convolutional Network による血管領域抽出.第38回日 本医用画像工学会大会予稿集:395-398, 2019
- [5] Zhu J, Park T, Isola P, et al.: Unpaired image-to-image translation using cycleconsistent adversarial networks. ICCV 2017: 2223-2232, 2017
- [6] Open CAS: Laparoscopic Image to Image Translation. http://opencas.dkfz.de/image2image/ (2021 年6月6日閲覧)
- [7] Ronneberger O, Fischer P, Brox T: U-Net: Convolutional Networks for Biomedical Image Segmentation. MICCAI 2015 9351: 234-241, 2015
- [8] Yu F, Koltun V: Multi-scale context aggregation by dilated convolutions. ICLR 2016
- [9] 盛満慎太郎,山本翔太,北坂孝幸,他.:
 Dilated convolution を用いた FCN による腹腔鏡動画像からの血管領域抽出.第
 39回日本医用画像工学会大会予稿集:
 230-233,2020

A Study on Segmentation of Blood Vessel Regions from

Laparoscopic Videos Using FCN with Depth Information

Keigo ENOMOTO^{*1}, Yuichiro HAYASHI^{*1}, Takayuki KITASAKA^{*2},

Masahiro ODA *3,1, Masaaki ITO *4, Nobuyoshi TAKESHITA *4, Kazunari MISAWA 5, Kensaku MORI *1,6,7

*1 Graduate School of Informatics, Nagoya University

- *2 Faculty of Information Science, Aichi Institute of Technology
- *3 Information Strategy Office, Information and Communications, Nagoya University
- *4 National Cancer Center Hospital East
- *5 Aichi Cancer Center Hospital
- *6 Information Technology Center, Nagoya University
- *7 Research Center for Medical Big Data, National Institute of Informatics

In this paper, we describe an automatic segmentation method of blood vessel regions from laparoscopic surgery videos using depth information. Since laparoscopic surgery is minimally invasive, the burden to the patient is small. On the other hand, because it is a difficult surgery for the surgeon, there is a need to develop a system to support surgeons. Blood vessel regions are important internal structures in the surgery. There is a possibility of causing bleeding and complications by damaging them during the surgery. To reduce such complications, a lot research on automatic segmentation of blood vessel regions from laparoscopic videos is being conducted. In the previous automatic segmentation methods, blood vessel regions are extracted based on color information from RGB images using an FCN. In this paper, we propose a new segmentation method that extracts blood vessel regions using RGB-D images. We estimate depth information from laparoscopic videos. The blood vessel regions are extracted from the RGB-D images using an FCN. As a result of our experiment, the segmentation accuracy was improved compared to the previous method especially in scenes that contain occlusion caused by forceps.

Key words: Laparoscopic surgery, Blood vessel segmentation, Depth estimation, RGB-D segmentation

Attention 機構を導入した

グラフニューラルネットワークによる

腹部動脈血管名自動対応付け

出口 智也*1 林 雄一郎*1 北坂 孝幸*2

小田 昌宏*3,1 三澤 一成*4 森 健策*1,5,6

要旨

本稿では、3 次元腹部 CT 像から抽出した動脈領域に対し、Attention 機構を導入したグラフニューラルネットワークを用いて血管名自動対応付けの検討を行ったので報告する.腹部動脈は個人差が大きく複雑な構造を持ち、種類も多様であるため、構造の把握は困難である.外科手術において医師が患者の血管構造を把握する際に医師の負担を軽減するため、血管領域に血管名を対応付けすることが提案されている.これまで、血管構造をグラフ構造と捉え、血管の位置や太さ、長さなどを特徴量としてグラフ畳み込みネットワークを用いて腹部動脈領域に対して血管名自動対応付けを行う手法が提案されており、その有用性が示されている.本稿では、Attentionを用いるグラフニューラルネットワークを利用した血管名自動対応付けの検討を行った.100 症例の CT 像に対して血管名自動対応付けを行った結果、平均 F 値は 89.6%であった.

キーワード:血管,グラフニューラルネットワーク,3次元 CT像,解剖学的名称認識,血管構造解析

1. はじめに

タ研究センター

近年広く行われている腹腔鏡下手術は,低侵 襲であるが得られる視野が狭く,高難度である. そのため,個人差が大きい血管構造を術中に正 確に把握することは難しい.そのため術前に3 次元腹部 CT 像から血管構造を把握することが 重要である.これまでにも腹部動脈領域に対し て血管名の自動対応付けを行う血管名自動命

*1 名古屋大学大学院情報学研究科
〔〒464-8601 名古屋市千種区不老町〕
e-mail: tdeguchi@mori.m.is.nagoya-u.ac.jp
*2 愛知工業大学情報科学部
*3 名古屋大学情報連携推進本部情報戦略室
*4 愛知県がんセンター
*5 名古屋大学情報基盤センター
*6 国立情報学研究所 医療ビッグデー

名手法がいくつか提案されてきた.鉄村らは血 管領域を木構造と捉え,条件付確率場を用いた 機械学習により命名を行った[1].日比らは血 管木構造をグラフと捉えることでグラフニュ ーラルネットワーク(Graph Neural Network, GNN)を用いた命名を行った[2].しかし,こ れらの手法は稀な分岐を持つ症例が存在する 肝動脈や,複雑な分岐をとる部分の末端に位置 する血管の命名精度が低い.そこで,本手法で はAttentionに基づくネットワークを導入し血 管名自動命名を行う手法を提案する.また,ネ ットワークの層を増やすことでネットワーク の表現能力を向上させると同時に広範囲の特 徴を利用できるようにして実験を行った.

2. 手法

本手法は,従来手法 [2] と同様の血管木構造 構築を行う.木構造は3次元腹部 CT 像から抽 出した血管領域画像から構築され,特徴量とし て血管の直径や長さ,走行方向,血管以外の臓 器との位置関係を抽出する.

推定には Attention を利用する GNN である Graph Attention Networks (GAT) [3] を用いる. GAT は注目ノードに隣接するノードに対し Attention に基づいて重みを付与し,特徴量の集 約を行う.1 層で隣接するノードのみを集約す るため,層を深くすることにより暗黙に広範囲 の特徴を利用可能である.本稿では広範囲の特 徴を利用するために6層に拡張した GAT を用 いた.また,比較対象は従来手法 [2] で用いら れた ChebNet [4] を用いる.

本稿では、血管木構造は重みなし無向グラフ とみなし各手法を適用する.血管木構造は、血 管構造の分岐から分岐までを1つの枝と見みな し1つの枝を1つのノードとする.グラフ構造 の特徴量は各血管枝の特徴量、接続関係は血管 枝の接続関係に対応させる.ノードの接続関係 及び特徴量を入力とし、学習にはノードの正解 ラベルデータを使用する.

3. 実験と結果

自動対応付けの主要な対象は、手法 [2] と同 様の、腹腔鏡手術において重要となる 22 種類 の動脈とする.学習と推定には GAT と手法[2] による重みづけされた ChebNet を使用する. い ずれも学習率は 0.001 とし、1000epoch 学習す る.実験は 3 次元腹部 CT 像から抽出された血 管領域画像 100 症例に対して 10 分割交差検定 法により行った.評価には F 値を使用した.

実験の結果,全体に対する F 値は,GAT は 89.6%,手法 [2] は91.3%となり,従来手法が GAT を約1.5 ポイント上回った.血管ごとの評 価では,一部の血管でGAT が上回ったものの, 従来手法が GAT を大きく上回る血管がみられ た.血管ごとのF 値の比較を表1に示す.

各手法による対応付け結果の例を図1に示す. GAT では一部命名できていない血管があり,正 解データよりもラベルが減ってしまったが,従 来手法で命名できなかった RGA を命名できて いる.(b)に示す従来手法では,おおむね命名に 成功しているものの,存在しない PHA が命名 されている.また,いずれの手法も LGEA を SA と誤っている.

5. 考察

GAT は隣接するノードのみの特徴量を集約 することから、層を深くし広範囲の特徴を利用 することを試みたが、従来手法に及ばなかった. これは、遠い位置にある特徴も隣接するノード の特徴として扱うことが、明示的に広範囲から 集約できる従来手法に比べ不利に働いたと考 えられる. 特に, LGEA のように1種類の長い 血管から分岐する場合や Ao のように近隣に多 くの種類の血管がある場合,暗黙に広範囲の特 徴量を利用することが誤りの原因となってい ると考える.一方で LCIA や RCIA, CHA のよ うに他の血管に挟まれる位置にある血管のF値 は従来手法を上回っており、このような血管に 対しては Attention の利用は有効であったとい える.また、差が開いた種類の血管はあったも のの,全体のF値は1.5ポイントの差と大きな 差ではなく、向上した血管もみられたことから 血管名自動対応付けにおいて Attention 機構は 有用であると考える.

4. まとめ

本稿では、Attention 機構を利用する GNN を 用いて3次元腹部 CT 像から抽出した動脈領域 に対する血管名自動対応付けを行った.その結 果、全体のF値では従来手法に及ばず、明示的 に拾い隣接関係を用いることが重要であるこ とが示された.また、一部の血管の精度におい てGATが上回ったことから Attention は GNN に よる血管名自動対応付けにおいて有用である ことが示された.

謝辞

日頃から熱心に御討論頂く名古屋大学森研 究室の諸氏に感謝する.本研究の一部は科研費 17H00867,17K20099,26560255,26108006, 21K19898,AMED 19lk1010036h0001,JST CREST JPMJCR20D5,堀科学芸術振興財団,JSPS 二国 間交流事業によった.

利益相反の有無

なし.

文 献

[1] 鉄村悠介,林雄一郎,小田昌宏 他:機

械学習を用いた腹部動脈血管名自動命 名におけるデータ拡張法の適用に関す る検討.信学技報118:191-196,2019

[2] 日比裕太,林雄一郎,北坂孝幸他:
 Spectral-based Convolutional Graph
 Neural Networks を用いた腹部動脈領域
 の血管名自動命名に関する研究.信学

技報 120:176-181,2019

- [3] Veličković P, Cucurull G, Casanova A et al.: Graph Attention Networks. ICLR, 2018
- [4] Defferrard M, Bresson X, Vandergheynst P: Convolutional neural networks on graphs with fast localized spectral filtering. NIPS, 2016



(a) GAT

GAT(b) 手法 [2](c) 正解データ図1血管名自動対応付け結果の例.同じ色は同じラベルを示す.赤色の丸は誤りを,黄色の丸は改善を示す.

表	1	GAT	と手法	[2]	によ	る血	管の利	重類ご	と0	の自重	力対応	「付け	け結果	のF	値の	比較	(%).
į	約当	半数の	⊃種類の	血管	では	GAT	が手	法 [2] を	上回	るも	しく	は同業	等と	なっ	ている).

血管名	Ao	CA	СНА	РНА	LHA	RHA	GDA	LGA	RGA	LGE A	RGE A
GAT	76.9	85.1	79.4	42.6	73.0	84.2	67.1	83.0	50.3	55.3	53.0
手法 [2]	88.8	84.2	74.9	47.2	69.9	80.2	77.2	85.0	58.9	60.8	74.8
血管名	SA	SMA	IMA	LRA	RRA	LCIA	RCIA	LEIA	REIA	LIIA	RIIA
GAT	92.0	96.5	80.8	98.3	98.3	74.4	74.3	97.5	98.0	96.3	98.0
手法 [2]	91.1	97.6	77.0	98.0	98.6	71.2	71.1	97.3	97.3	96.5	96.5

Automated Anatomical Labeling of Abdominal Arteries Based on

Graph Neural Networks Using Attention Mechanism

Tomoya DEGUCHI*1, Yuichiro HAYASHI*1, Takayuki KITASAKA*2,

Masahiro ODA*3,1, Kazunari MISAWA*4 and Kensaku MORI*1,5,6

*1 Graduate School of Informatics, Nagoya University

*2 Faculty of Information Science, Aichi Institute of Technology

*3 Information Strategy Office, Information and Communications,

Nagoya University

*4 Aichi Cancer Center Hospital

*5 Information Technology Center, Nagoya University

*6 Research Center for Medical Bigdata, National Institute of Informatics

In this paper, we report an automated anatomical labeling method for the abdominal artery regions extracted from 3D abdominal CT volumes. In this method, we adopt graph neural networks with an Attention mechanism. The abdominal artery has a complex structure with large individual differences, and thus surgeons have difficulty in understanding the structure. Therefore, assistance by computer is desired to help surgeon's understanding of vascular structure by presenting the results of automated anatomical labeling of abdominal arteries. Several automated anatomical labeling methods for the abdominal artery regions using machine learning have been proposed. Among them, the methods based on Graph Neural Network (GNN) were examined. These methods represent vascular structure as graph structure and use features of vascular like thickness, position and length as node's features. Automated anatomical labeling method with the GNN showed high precision and demonstrated its usefulness. In this paper, we adopt an attention-based GNN model, Graph Attention Networks (GAT) to automated anatomical labeling. In our experiment, a 10-fold-cross-validation was performed on 100 artery regions extracted from CT volumes. As a result of the experiment, the GAT achieved an F-value of 89.6% in the evaluation of all cases.

Key words: Blood vessel, Graph neural network, CT volume, Anatomical names recognition, Blood vessel structures analysis

3D Kidney Tumor Semantic Segmentation using

Cascaded Convolutional Networks

Wuyang Zhao^{*1}, Chen Shen^{*1}, Masahiro Oda^{*2,*1},

Yuichiro Hayashi^{*1}, Naoki Higashida^{*3}, Msahiro Hashimoto^{*3},

Masahiro Jinzaki^{*3}, Kensaku Mori^{*1,*4,*5}

Abstract

In this paper, we propose a kidney tumor segmentation method using the cascaded convolutional network. Kidney tumor segmentation from CT volumes is an essential step in kidney tumor surgery planning, and accurate segmentation results can assist doctors in making surgical planning. However, the irregularity of tumor shape and low contrast of lesion boundaries in CT volumes make kidney tumor segmentation especially challenging. This paper described a two-stage pipeline for kidney tumor segmentation in CT volumes using convolutional networks. In the first stage, we get the coarse location of the kidney using the lightweight network and then crop VOI (Volume-of-interest) from CT volumes. Then the second stage, we precisely segment the kidney tumor from the cropped patch in the first stage. We used our datasets for the experiment and obtained 71.77% of the Dice score in kidney tumor segmentation. The experimental results showed that our method for kidney tumor segmentation is effective.

Keywords: kidney tumor, computed tomography, segmentation, deep leaning

1. Introduction

The kidney tumor is one of the most rapidly increasing cancer, with more than 4,000,000 cases of kidney tumors each year, and kidney cancer accounted for 175,000 deaths worldwide in 2018[1]. In medical image segmentation, the main task is to segment organs or lesions from medical images. We can obtain information about size, shape, and position of organs or lesions through semantic segmentation. Semantic segmentation is beneficial for doctors to make diagnosis plans; however, morphological heterogeneity makes segmentation especially challenging.

Various deep learning methods [2] have been applied to kidney tumor segmentation in the last decade. Kim et al. [3]

- *1 Graduate School of Informatics, Nagoya University, Nagoya, Japan
- *2 Information Strategy Office, Nagoya University, Nagoya, Japan
- *3 Keio University School of Medicine, Tokyo, Japan
- *4 Information Technology Center, Nagoya University, Nagoya, Japan
- *5 Research Center for Medical Bigdata, National Institute of Informatics, Tokyo, Japan

used gray level thresholding method for kidney tumor segmentation. However, the thresholding segmentation result is greatly affected by image noise and not the automatic segmentation. Anis et al. [4] proposed construction of 3D kidney tumor using 2D ultrasound images. However, this method ignores the contexts on the z-axis, which would lead to limited segmentation accuracy. Guanyu et al. [5] proposed a three-dimensional, Fully Convolutional Neural Network model for kidney tumors segmentation. However, for the small size of kidney tumor, this method cannot extract the tumor extract.

This paper proposes a cascaded segmentation framework for kidney tumor segmentation to overcome the limitation in kidney tumor segmentation accuracy. We first extract the kidney region using the Residual ConvUNet3D [6], in which we replace the max-pooling layers with the convolution and add residual blocks which can effectively learn small-sized kidney tumors. Then a fine kidney tumor segmentation is conducted on the extracted kidney regions. Furthermore, we employ a sequence of data augmentation methods to increase data to improve the model's robustness. We describe details of our pipeline in the following section.

2. Method

The proposed method extracts the kidney tumor from a CT volume using a two-stage pipeline. An overview of the proposed method is shown in Fig. 1. In the first stage, we extract kidney regions from the CT volume. Then, we extract tumor regions from the cropped CT volume. The new method can extract lesions' boundaries and segment small-sized kidney tumor. We use many data augmentation methods which can effectively alleviate the problem of insufficient data.

1) Data preprocessing and augmentation

Data preprocessing plays a vital role in the workflow. We clip the intensity values of CT volume into [-79, 304] H.U. to remove the irrelevant detail and use the max-min normalization method to normalize the image intensity values into the range of [0, 1] for the entire CT volume. Various data augmentation techniques are applied to avoid the model being overfitted in the training phase. We apply rotation, scaling, flipping, translation, Gaussian noise, Gaussian blur, and gamma correction on input volumes. The inputs of Residual ConvUNet3D are randomly cropped CT volumes.



Segmentation result

2) Network architecture

Fig. 1 Kidney tumor inference flow of our propose method.

We used Residual ConvUNet3D model, which is made by replacing some parts in the 3D U-Net [7]. In our model, max-pooling layers are replaced by down convolution layers, transposed convolution layers replace upsampling layers, and residual blocks [8] are used in the encoder part of the network architecture.

3) Network training and estimation

Network training includes two stages. We use the ACC kidney dataset (which contains CT volumes and their corresponding kidney regions) and the RCC kidney tumor dataset (which contains CT volumes and their corresponding kidney tumor regions). In the first stage, we resample the ACC dataset to a low resolution with voxel spacing of 1.426 \times 1.426 \times 1.839 mm, representing the voxel's length, width, and height. We train a coarse resolution Residual ConvUNet3D model using the ACC dataset with a patch size of $192 \times 192 \times 64$ voxels. In the second stage, we resample the RCC dataset to a high resolution with a voxel spacing of 0.957 \times 0.714 \times 1.243 mm, which represents the length, width and height of the voxel. We use a trained coarse Residual ConvUNet3D model in the first stage to get the coarse location of the kidney and crop kidney part from a CT volume of the RCC dataset. Then we train a fine Residual ConvUNet3D with a patch size of $224 \times 96 \times 96$ voxels.

In the estimation stage, we input the test data and obtain the coarse location of the kidney using the coarse resolution Residual ConvUNet3D model and get kidney segmentation results. Then we crop the kidney part from the kidney segmentation result, and input cropped volumes into the fine resolution Residual ConvUNet3D model. Finally, we get the tumor segmentation result.



Ground truth Residual ConvUNet3D 3D U-Net Fig. 2 Comparisons of tumor segmentation results of Residual ConvUNet3D and 3D U-Net. Red parts in ground truth images are kidney tumor, and red parts in Residual ConvUNet3D and 3D U-Net result images stand for segmentation results.

Model	Dice score of tumor	Dice score of background
3D U-Net	69.75%	99.97%
Cascaded Residual UNet3D	71.77%	99.97%

Table 1 Mean Dice score of different models on test dataset.

3. Experiments and Results

We extracted kidney tumor regions from CT volumes. Also, segmentation accuracies were calculated for two different datasets.

1) Dataset

Our experiments used two different datasets: the ACC kidney dataset and the RCC kidney tumor dataset. The ACC dataset includes 137 cases of kidney CT volumes, and the RCC dataset consists of 304 cases of kidney tumor CT volumes. The size of each image was 512×512 pixels, the number of slices was 101 to 561, pixels spacing was 0.59 to 0.96 mm, and slice thickness was 0.5 to 1.5 mm. The ACC dataset is collected in the Aichi Cancer Center. We split the RCC dataset into the training set, validation set, test set with 240 cases, 30 cases, and 34 cases, respectively. The Keio University School of Medicine provided the RCC dataset.

2) Implementation and result

Ground truth

The total training epoch was 300 epochs, and one epoch took about 500 seconds. We trained the model on a NVIDIA Tesla V100 (single GPU) and used Pytorch framework to implement our network. The network in our experiments was trained with stochastic gradient descent. Dice score was used as the evaluation criterion.

Table 1 shows the kidney tumor extraction results of different models. We obtained 71.77% and 69.75% Dice scores of kidney tumors in the Residual ConvUNet3D and 3D U-Net. The segmentation results showed that our approach is effective. Figure 2 shows the comparisons of the tumor segmentation results. From the comparisons, our method can extract lesions' boundaries and segment small-sized kidney tumor well. Figure 3 is an example of failed segmentation result of kidney tumor. Due to the difference of two datasets, we cannot obtain the location of the kidney in some cases that affect the tumor segmentation results.



Fig. 3 An example of failed tumor segmentation result using our method. Due to the difference of two datasets, we can't get the kidney location in some cases that failed to extract kidney tumor.

Residual ConvUNet3D

3D U-Net

4. Discussion

The experimental results showed that the Residual ConvUNet3D can learn small-sized kidney tumors well. The twostage strategy contributed to improving lesions' boundary extraction accuracy. However, we used two different datasets in the experiments. Due to the difference of two datasets such as the range of CT value and image spacing, we cannot obtain the kidney location in some cases in the first stage that affects the experiment results. A two-stage strategy for tumor segmentation needs more training and inference time. In the future, we will try to investigate a new method to solve the problems.

5. Conclusion

In conclusion, we used Residual ConvUNet3D model and described a two-stage semantic segmentation for kidney tumor segmentation from a CT volume. Simultaneously, we used many data augmentation methods to increase the data. The proposed model is trained to segment small-sized kidney tumor well. The proposed method achieved promising segmentation results. However, the current experimental results do not guarantee that our method can achieve good results on other datasets. We will conduct further research on applicability to other datasets.

Competing interests

None.

Acknowledgement Part of this research was supported by the MEXT/JSPS KAKENHI (17H00867, 21K19898) and AMED (JP20lk1010036).

References

- [1] Bray F, Ferlay J, Soerjomataram I, et al: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA **68**: 394-424, 2018
- [2] Taghanaki S A, Abhishek A, Cohen J P, et al: Deep semantic segmentation of natural and medical images: a review. Artificial Intelligence Review 54: 1–42, 2021
- [3] Kim D Y, Park J W: Computer-aided detection of kidney tumor on abdominal computed tomography scans. Acta radiologica **45**(7): 791-795, 2004
- [4] Ahmad A, Cool D, Chew B H, et al: 3D segmentation of kidney tumors from freehand 2D ultrasound. Medical Imaging 2006:Visualization, Image-Guided Procedures, and Display. International Society for Optics and Photonics: 6141: 61410S, 2006
- [5] Yang G, Gu J, Chen Y, et al: Automatic kidney segmentation in CT images based on multi-atlas image registration. Conference of the IEEE Engineering in Medicine and Biology Society 2014: 5538-5541, 2014
- [6] Zhao W, Shen C, Oda M et al A Study of Automatic Kidney Tumor Segmentation with Deep Learning. IEICE Technical report 121: 3-8, 2021
- [7] Çiçek Ö, Abdulkadir A, Lienkamp S S, et al: 3D U–Net: Learning Dense Volumetric Segmentation from Sparse Annotation. In MICCAI 2016, LNCS 9901: 24–432, 2016

[8] He K, Zhang X, Ren S, et al: Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition: 770–778, 2016

Uncertainty Prediction of Vertebrae Segmentation Using Bayesian

U-Net: Towards Age- and Gender-dependent Statistical Modeling

in a Large-scale CT Database

Zhuo CHENG^{*1}, Yoshito OTAKE^{*1}, Mazen SOUFI^{*1}, Keisuke UEMURA^{*1}, Hideki SHIGEMATSU^{*2}, Masaki IKEJIRI^{*2}, Kento AIDA^{*3}, Shin'ichi SATOH^{*3}, Masahiro HASHIMOTO^{*4}, Toshiaki AKASHI^{*5}, Yoshinobu SATO^{*1}

Abstract

Quantifying segmentation uncertainty has become an important task for anatomical structures with large diversity, such as vertebrae. A previously proposed Bayesian U-Net demonstrated a correlation between the Monte Carlo (MC) dropout sampling-based uncertainty prediction and the segmentation accuracy in an application of the muscle segmentation. However, the effectiveness of this approach in vertebrae segmentation has not been validated. This work integrates the MC dropout sampling in a previously proposed framework that achieved high vertebrae segmentation accuracy and landmark detection rate in the MICCAI 2019 challenge. A 6-fold cross-validation study using the public CT dataset VerSe 2019 was conducted. The proposed method achieved an average Dice coefficient (DC) of 0.922 ± 0.112 , with a correlation coefficient between DC and corresponding uncertainty of -0.533 among the correctly identified vertebrae (landmark detection rate was 92.00%, n=793/862). Furthermore, we validated our approach on 100 CT volumes obtained from a large-scale CT database collected independently from the training dataset, on which an average DC of 0.982 ± 0.106 and a correlation coefficient of -0.502 are achieved among the correctly identified vertebrae (landmark detection rate was 93.91%, n=1821/1939). The results suggest the feasibility of the uncertainty estimated by Bayesian U-Net as a predictive measure of vertebrae segmentation accuracy, which would be helpful for the age- and gender-dependent statistical modeling in a large-scale CT database for our future work.

Keywords : Bayesian Deep Learning, Convolutional Neural Network, U-Net, Vertebrae Segmentation

*1 Division of Information Science, Graduate School of Science and Technology,

Nara Institute of Science and Technology [8916-5, Takayama-cho, Ikoma, Nara 630-0192, Japan]

E-mail: cheng.zhuo.br5@is.naist.jp

- *2 Department of Orthopedic Surgery, Nara Medical University
- *3 National Institute of Informatics
- *4 Department of Radiology, Keio University School of Medicine
- *5 Department of Radiology, Juntendo University Graduate School of Medicine

1. INTRODUCTION

In recent years, remarkable success has been shown by Convolutional Neural Network (CNN) in medical image segmentation. However, although CNN-based methods have achieved high performance in various segmentation tasks, the reliability of results still needs to be quantified for anatomical structures with large diversity, such as vertebrae.

Gal et al. [1] showed that uncertainty could be predicted by Bayesian approximation that incorporates Monte Carlo (MC) dropout and deep learning model. Hiasa et al. [2] also demonstrated a correlation between the MC dropout sampling-based uncertainty prediction and the segmentation accuracy in an application of the muscle segmentation. However, the effectiveness of Bayesian deep learning in vertebrae segmentation has not been validated yet. This work applies the MC dropout sampling to a previously proposed CNN-based framework that achieved high vertebrae segmentation accuracy and landmark detection rate in the VerSe 2019 challenge [3]. We evaluate the proposed approach on both VerSe2019 dataset and the dataset randomly selected from a large-scale database (noted as "J-MID database" in this paper), including over 10,000 CT volumes [4]. The results show that the predicted uncertainty of Bayesian 3D U-Net is feasible to estimate the vertebrae segmentation accuracy, which would help build age- and gender-related statistical model in a large-scale CT database.

2. METHODS

Vertebrae localization and segmentation are performed by a three-stage fully automatic approach: 1) spine localization, 2) vertebrae localization and identification, and 3) binary segmentation of each localized and identified vertebrae [5]. The overview of the proposed method is shown in Figure 1. Due to the different field-of-view of CT volumes, in the first stage, 3D U-Net is used to regress a heatmap of the spinal centerline to ensure that the spine is centered in the input for the subsequent vertebrae localization step, which can help reduce memory cost and processing time. Then, the SpatialConfiguration-Net is employed to identify each vertebra and localize their centroids. Finally, 3D U-Net with sigmoid cross-entropy loss is used for binary segmentation of individual vertebrae.

In order to create the final multi-class segmentation mask, the predicted identification labels in the second stage are merged with the binary segmentation results. Additionally, the segmentation network in the final stage is reformed to the Bayesian 3D U-Net to estimate segmentation uncertainty by multiple test-time MC dropout samples.

3. EXPERIMENTS

We used 80 CT volumes from the public dataset of VerSe 2019 challenge [3] with two types of annotations: 1) coordinates of vertebral centroids and 2) multi-class segmentation mask. It consists of CT volumes with various field-of-view (FoV), including the cervical spine, the thoracolumbar spine, and the cervico-thoracolumbar spine. A 6-fold cross-validation experiment was conducted for the evaluation. Furthermore, we validate the approach on randomly selected 100 CT volumes of the thoracolumbar and cervico-thoracolumbar spine from the J-MID large-scale CT database collected independently from the training dataset. The annotations of VerSe 2019 dataset are provided by the organizers. For the J-MID dataset, the model trained by VerSe 2019 dataset generated initial segmentation masks, and



Fig. 1 Overview of the proposed approach for vertebrae segmentation and segmentation uncertainty prediction and the data annotation approach for J-MID dataset.

Dataset	Cervical Vertebrae	Thoracic Vertebrae	Lumbar Vertebrae
Verse 2019	93.75%	91.74%	91.40%
J-MID Dataset	98.35%	92.66%	95.81%

Table 1 Success rate of vertebral centroid detection [%].

non-expert annotators revised the masks, then an expert surgeon verified them.

The segmentation accuracy and the uncertainty metric for predicting the accuracy in the absence of the ground truth are assessed using the Dice coefficient (DC) and Pearson correlation coefficient between DC and uncertainty of individual vertebra. The vertebra-wise uncertainty metric is computed as the average of the estimated pixel-wise variance in the estimated labels. To reduce the interference from landmark detection, we only evaluated the vertebrae that are correctly identified and localized. To define the failure in landmark identification, we followed the criteria used in the MICCAI 2019 challenge, i.e., the landmark identification was considered as successful if the error at the vertebral centroid is smaller than 20 millimeters. The network configurations we used is same as [5], and 10 test-time MC dropout samples are performed.

4. RESULTS

Table 1 shows the landmark detection rate of vertebral centroids. DC of correctly detected and localized vertebrae on the VerSe 2019 dataset and J-MID dataset was 0.922 ± 0.112 and 0.982 ± 0.106 , respectively. Figure 2 shows the segmentation accuracy of each vertebra and the quantitative results of correlation between uncertainty and segmentation accuracy. The middle column is the scatter plots of DC as a function of uncertainty, and the right column is the box plots of DC of two groups separated by a predefined uncertainty threshold (0.3×10^{-3} in this case). The mean DC for the group exhibiting the mean uncertainty less than the threshold was 0.949 ± 0.042 and 0.995 ± 0.056 on the VerSe 2019 dataset


Fig. 2 Evaluation of segmentation accuracy and uncertainty prediction. (a) DC of the correctly identified and localized vertebrae.
(b) Scatter plots of DC as a function of the mean of uncertainty. (c) Box-and-whisker plots of DC for two groups separated as a predefined uncertainty threshold (0.3×10⁻³ in this case).

and J-MID dataset, respectively, while DC for the other group was 0.801 ± 0.204 and 0.832 ± 0.285 . The segmentation and uncertainty metric of three representative cases from the VerSe2019 dataset is visualized in Figure 3. The first row shows a case with high DC and low uncertainty and the second row shows a case including the vertebra with low DC and high uncertainty due to vertebral fracture.

5. DISCUSSION AND CONCLUSION

This paper demonstrates the feasibility of the uncertainty estimated by Bayesian U-Net as a predictive measure of vertebrae segmentation accuracy. However, mis-detected landmarks may lead to overlapping segmentations with low uncertainty but low DC, such as Th10 in case 3, as shown in Figure 3. In order to analyze the variations among different age and gender groups, our future work mainly focuses on building age- and gender-dependent statistical models of vertebral centroids and shapes. These structures will be obtained by applying the proposed approach on over 10000 CT volumes in J-MID large-scale CT database.



Fig. 3 Visualization of the segmentation and uncertainty metric of representative cases from the VerSe 2019 dataset.

COMPETING INTERESTS

The authors declare that they have no competing interests.

ACKNOWLEDGEMENT

This study was supported by the Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research (KAKENHI) Numbers 19H01176 and 20H04550, and Japan Agency for Medical Research and Development (AMED) Number JP20lk1010036. The CT database used in this study is a part of the Japan Medical Image Database (J-MID) constructed by the Japan Radiology Society (JRS)

References

- [1] Gal Y, Ghahramani Z: Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In Proceedings of the 33rd International Conference on Machine Learning (ICML), New York, 2016, pp1050-1059
- [2] Hiasa Y, Otake Y, Takao M, et al.: Automated muscle segmentation from clinical CT using Bayesian U-Net for personalized musculoskeletal modeling. IEEE Transactions on Medical Imaging 39.4: 1030-1040, 2019
- [3] Sekuboyina A, Husseini M E, Bayat A, et al.: Verse: a vertebrae labeling and segmentation benchmark for multi-detector CT images. arXiv preprint arXiv:2001.09193, 2020
- [4] Nishie A, Kakihara D, Machitori A, et al.: Japan Safe Radiology 2020. European Congress of Radiology (ECR), Vienna, 2020.
- [5] Payer C, Stern D, Bischof H, et al.: Coarse to Fine Vertebrae Localization and Segmentation with SpatialConfiguration-Net and U-Net. In: Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, Valletta, 2020, pp123-133

弱教師あり学習を用いた MR 画像における血管奇形腫瘍の

領域の自動抽出

深谷 航生*1 原 武史*1 野崎 太希*2 周 向栄*1 藤田 広志*1

要旨

本研究の目的は、MRI を用いて撮像した画像において、深層学習を利用した血管奇形腫瘍の領域の自動抽 出手法を開発し、定量的評価につなげることである.本研究が対象とするクリッペルトレノネーウェーバ 一症候群は正解画像の入手が困難である.そこで、病変が左右の大腿のどちらに発生しているかのラベル のみを用いて領域抽出を行った.画像単位のラベルを用いて分類を行い、得られた Attention Map に前後処 理を施して、疑似ラベルとして U-Net でセグメンテーションを行った.医師が手動で算出した病変の体積 値と比較したところ、提案手法の有効性が示された.

キーワード:血管奇形腫瘍,弱教師あり学習,セグメンテーション, Grad-cam

1. はじめに

血管奇形腫瘍の完治には複数回の治療が必要である。治療の前後で血管腫の領域の正確な把握が必要になるが,現在は目視による診察や,MR 画像から医師が主観で治療の効果 判定を行っている。

本研究が対象にするにしているクリッペル トレノネーウェーバー症候群は、日本での患 者数が約3000人と非常に少なく、大量のデー タを集めることが困難である.また、びまん 性を有するため、正解画像の作成に多大な労 力と時間を要する.そこで本研究では画像単 位で付与されたラベルのみを用いて、血管 奇形腫瘍の領域の自動抽出手法の開発を行っ た.

*1 岐阜大学大学院自然科学技術研究科知 能理工学専攻

- 〔〒501-1194 岐阜県岐阜市柳戸 1-1〕
- e-mail: kfukaya@fjt.info.gifu-u.ac.jp
- *2 聖路加国際病院放射線科

2. 方法

クリッペルトレノネーウェーバー症候群は 片側肥大の特徴を有する.そこで,深層学習 を用いて病変が大腿の左右どちらにあるかの 2分類を行う.分類にはvgg16[1]の最後のMax Pooling と畳み込み層を除いたモデルを用い た.そして,作成した分類モデルの3箇所の 畳み込み層に,改良した Grad-cam[2][3]を適 用する.得られた3枚のアテンションマップ の和をとり,血管腫の持つびまん性に対応す るために確信度に強弱をつけた疑似ラベルを 作成する.疑似ラベルで,MR 画像の画素値が 70以下の座標を背景として削除する.出来上 がった疑似ラベルを用いて U-Net[4]で学習し, 最終結果を得る.



図1 試料画像, アテンションマップ, 疑似ラベルの例

3 実験

3.1. 試料画像

岐阜大学医学部附属病院,国立成育医療研

究センター,神戸大学医学部附属病院,神戸 MR 画像 13 症例 450 枚を用いる.13 症例の うち、岐阜大学医学部附属病院の2 症例と, 国立成育医療研究センターの4 症例について は医師が手動で病変の体積値を計算した正解 データがある.画像サイズは撮像された施設 によって異なるため,256×256pixels にリサ イズして実験に用いる.

3.1. 評価方法

疑似ラベルの作成は,13 症例で4 fold の交 差検証を行う.セグメンテーションは,10 症 例を学習データ,3 症例をテストデータとし て2回学習とテストを行う.評価には症例ご との体積値を用い,正解データとの誤差を求 める.

4. 結果

本手法によって求められた病変の体積値と 正解データの誤差を表1に示す. 平均誤差は 27.3%だった. 症例ごとで見ると最小誤差は 11.6%,最大誤差は56.2%だった. また、疑似 ラベルの病変の体積値と正解データの誤差も 表1に示す. MR 画像と最終出力の例を図2 に示す.

表 1	正解デー	\$	との	体積値の語	誤差(症例	ごと	:)
-----	------	----	----	-------	-----	----	----	----

	抽出結果	疑似ラベル
	×	×
	正解データ	正解データ
平均誤差(%)	27.3	29.3
最小誤差(%)	11.6	10.4
最大誤差(%)	51.6	47.1



図 2 MR 画像と最終出力例

5. まとめ

画像単位のラベルのみを用いて, 血管奇形腫

瘍の領域の抽出を行った.出力結果と正解デ ータとの平均の誤差は 27.3%だった.疑似ラ ベルと正解データの誤差は 29.3%であり, Grad-Cam を用いて得られたアテンションマ ップを疑似ラベルとして用いることで抽出精 度の向上が見られた.また,抽出結果,疑似 ラベルと正解ラベルで最小・最大誤差が計測 されたのは同一症例であった.より適切な疑 似ラベルを作成することで精度の向上が見込 まれる.

謝辞

本研究を進めるにあたり,有益なご助言を いただきました研究室の方々に感謝の意を表 します.本研究の一部は,文部科学省科研費・ 新学術領域研究(26108005)および基盤研究 C(17K10455,18K12102)の補助によって行 われました.

利益相反の有無

無し

文 献

- Simonyan,K. Zisserman, A. "Very deep convolutional networks for large-scale image recognition", arXiv:1409.1556, April 2015
- [2] Ramprasaath R.Selvaraj, Michael Cogswell "Grad-Cam:Visual Explanations from Deep Networks via Gradient-based Localization", IEEE, ICCV, pp618-626, July 2017
- [3] 坂井創一,竹中要一"畳み込みニュー ラルネットワークの学習過程の可視化", The 33rd Annual Conference of the Japanese Society for Artifical Intelligence, 2019
- [4] Olaf Ronneberger, Philipp Fischer, Thomas Brox "U-Net:Convolutional Networks for Biomedical Image Segmentation", Medical Image Computing and Computer-Assisted Intervention – MICCAI2015, pp.234-241, November 2015

Automatic Extraction of Hemangiomas Regions in MR Images Using Weakly Supervised Learning

Koki Fukaya*1, Takeshi HARA*1, Taiki Nozaki*2, Xiangrong ZHOU*1, , Hiroshi FUJITA*1

- *1 Department of Intelligence Science and Engineering, Graduate School of Natural Science and Technology, Gifu University
- *2 Department of Radiology, St.Luke's International Hospital

The purpose of this study is to develop a method to automatically extract regions of vascular malformation tumors using deep learning in MRI images for quantitative evaluation. In this study, we focus on Klippel-Trennone-Weber syndrome, for which it is difficult to obtain correct images. Therefore, we extracted the regions using only the label of whether the lesion occurred on the left or right thigh. We classified the lesions using labels for each image, and segmented them using U-Net as pseudo-labels by applying pre/post processing to the obtained attention maps. The effectiveness of the proposed method was demonstrated by comparing it with the volume values of lesions calculated manually by doctors.

Key words: Hemangioma, Weakly supervised learning, Segmentation, Grad-Cam

頭部 thick-slice CT 画像からの

階層的脳内血腫領域抽出法の提案

岡 和範*1 藤田 大輔*1 野原 康信*2 井上 創造*3

有村 公一*4 飯原 弘二*5 小橋 昌司*1

要旨

脳血管疾患は日本の死因第4位であり、2019年の死亡者数は約10万人に上る.その要因の一つは脳内血 腫(ICH; intracerebral hemorrhage)である. ICH は専門医でも読影が難しく,診断に時間を要する為,医師の 負担軽減や診断の迅速性・正確性向上の為に,脳CT 画像からのICH 領域自動抽出法が必要とされている. 従来研究として,FCN による手法が研究されているが、実用化には不十分で、特に高吸収を示す脳室内や クモ膜下出血領域の誤抽出が多い.本研究では、高吸収領域の過抽出を削減する脳CT 画像からのICH 領 域自動抽出法を提案する.提案法は、ICH の解剖構造に注目し、高吸収領域の抽出、深層学習を用いたク ラス分類による階層的手法を提案する.このモデルの性能を、脳室内出血を有する被験者30名を用いて評 価し、感度83%、特異度91%を記録した.

キーワード:脳内血腫,脳CT 画像,領域抽出,深層学習,階層的手法

1. はじめに

現在の脳内血腫における画像診断の問題点と して,救急現場では脳 CT 画像が常に専門の放 射線科医に読影されるとは限らず,脳内血腫読 影に時間を要することと,経験豊富な放射線科 医でも脳内血腫読影の難易度が高いことが挙 げられる.その為,脳内血腫を自動抽出するこ とで画像診断の観点から医師の負担を軽減さ せ,他の診断プロセスに時間を充てられる為, 脳内血腫診断の迅速性・正確性が向上し,診断 支援となることが期待される.

脳内血腫領域は単に脳実質や脳脊髄液と比較

*1 兵庫県立大学大学院工学研究科

〔〒671-2280 兵庫県姫路市書写 2167〕

- e-mail: kobashi@eng.u-hyogo.ac.jp
- *2 熊本大学
- *3 九州工業大学
- *4 九州大学
- *5 国立循環器病研究センター

して高い CT 値を持つ領域ではなく,出血領域 を除外した高吸収域と,辺縁部の低吸収域で構 成される.従来研究[1][2][3][4]では FCN による 脳内血腫領域抽出法が提案されているが,頭蓋 内の高吸収域を血腫としており,低吸収域の抽 出を考慮していないため,臨床的に不十分な抽 出精度であった.

本研究においては, CT 画像上の脳内血腫の信 号値と空間的な構造に基づき, Level-Set 法や深 層学習によるクラス分類によって, 階層的に脳 内血腫領域を抽出する.

2. 使用データ

本研究では、J-ASPECT study[5]で収集され、 脳内血腫が確認された 30 名を対象とした.各 被験者について、救急搬送直後の CT 撮影で得 られた thick-slice CT 画像を使用する.本研究は 九州大学の倫理委員会の承認を受けて実施し た.

各撮影において, multidetector computed tomography(MDCT)装置により約20枚の横断面





(a) 正常な脳 CT 画像(b) 脳内血腫図 1 頭部 CT 画像例

(アキシャル)画像を撮影した.各画像は医用画 像の一般的画像フォーマットである DICOM 形 式で保存され,画像左上が原点で右方向が X 軸, 下方向が Y 軸である.画素数は 512×512(pixel) で,空間解像度は縦横共に 0.47mm/画素,スラ イス幅は,9.0mm である.図1に頭部 CT 画像 の一例を示す.全ての画像に対し,専門医によ る血腫領域のアノテーションを行い,抽出した 領域を血腫領域の正解データとした.

3. 提案法

1) CT 値の正規化と Level-Set 法を用いた頭蓋 内領域の抽出

CT 値を式(1)で正規化する.

$$\mu' = max\left(0, min\left(\frac{\mu - TH_{min}}{TH_{max} - TH_{min}} \times 255, 255\right)\right)$$
(1)

ここで、 μ は注目ボクセルの CT 値[HU]、 μ' は変 換後の値[a.u.]で、 TH_{min} は 0、 TH_{max} は 100 と実 験的に定めた.次に、頭蓋骨外側輪郭を Level-Set 法を用いて抽出する.頭蓋骨領域の 2 値画 像に対して、初期形状を画像端にした Level-Set 法を適用することで、頭蓋骨領域輪郭を抽出す る.得られた輪郭内部で、閾値以下の領域を頭 蓋内領域として抽出する.閾値は実験的に求め た 230 を用いた.

2) 高吸収血腫候補領域抽出法

頭蓋内領域において,高吸収域の脳内血腫候 補領域を閾値処理で抽出する.閾値を実験的に 120[a.u.]とし,閾値以下の領域を高吸収血腫候 補領域として抽出する.

3) 深層学習を用いた脳内血腫領域識抽出法

高吸収血腫候補領域には,真の脳内血腫と,





図3 頭蓋内領域と高吸収領域の抽出結果

それ以外の脳室内出血等の領域を含む.そこで, 本手法は深層学習を用いて,高吸収域血腫候補 領域の各ボクセルに対して,脳内血腫クラスか, 出血クラスかの2クラス分類を行うことで,高 吸収の脳内血腫領域を抽出する.

高吸収血腫候補領域の各ボクセルについて, 同ボクセルを含む撮像断面であるアキシャル 断層画像中において,注目ボクセルを中心に高 さ W,幅 W の正方形領域をイメージパッチと し,各ボクセルの正規化 CT 値を同点の属性値 とする.本実験では W=129 とする.これを,2D-CNN によりクラス分類する.本研究で用いる 2D-CNN は,図2に示すネットワーク構造を使 用する.

4. 実験結果と考察

1) Level-Set 法による頭蓋内領域抽出結果と高 吸収領域抽出結果

提案した Level-Set 法による頭蓋内領域抽出 法と, 閾値処理による高吸収領域抽出法を, 脳 内血腫を有する被験者へ適用した結果例を図 3 に示す. 同図より, 頭蓋骨内部の頭蓋内領域が 良好に抽出できることが確認できた. また, 出 血領域を含む血腫の候補領域抽出されている ことが確認できる. なお同候補領域には, 頭蓋 骨付近に同等の正規化 CT 値を持つ画素など, 血腫以外の領域を含むことが存在しているこ とが確認できる.

汎化性能		비	解		H L 田
		脳内 血腫	高吸収	感度	度
予測	脳内 血腫	58,770	18,298	0.92	0.01
	高吸 収	5,704	61,698	0.83	0.91

表1 検証データに対する脳内血腫領域抽出結果



2) 高吸収血腫候補領域からの血腫分類結果

被験者 30名を、学習データ 27名、検証デー タ3名として実験を行った、学習済みモデルに よる検証データに対する分類性能をそれぞれ 表1に示す、同表で混同行列はボクセル数を示 す、表1より感度 0.83、特異度 0.91 であり、未 知データに対して良好な精度が得られる可能 性を確認できる、また、評価関数と損失関数の 学習過程を図 4,5 に、抽出結果を画像に描出し た例を図 6 に示す.

5. 結論

本論文では、先行研究である CNN による血 腫領域抽出法を拡張し、階層的領域抽出法を提 案した.今後の課題は、前処理における輪郭抽 出に対して、バルーン法等を用いた3次元的な 頭蓋骨の連続性を考慮することによる頭蓋骨 輪郭抽出法の検討、新たな閾値決定法の検討、 深層学習における交差検証の検討、低吸収域に おける血腫領域抽出法の検討が挙げられる.

謝辞

本研究の一部は,厚生労働科学研究費補助金 (政策科学総合研究事業(臨床研究等 ICT 基盤 構築・人工知能実装研究事業)課題番号 19AC1003の支援を受けて実施した.





図 6 高吸収領域中の脳内血腫領域抽出結果 左から元画像,脳内血腫領域抽出結果(赤; TP, 黄; FP), 真値(緑).

利益相反の有無なし

文 献

- A. Arab, B. Chinda, G. Medvedev et al: A fast and fully-automated deep-learning approach for accurate hemorrhage segmentation and volume quantification in non-contrast whole-head CT. Scientific Reports 10: Article No. 19389, 2020
- [2] W. Kuo, C. Hane, P. Mukherjee: Expertlevel detection of acute intracranial hemorrhage on head computed tomography using deep learning. PNAS 116: Article No.45, 2019
- [3] J. Cho, K. Park, M. Karki et al: Improving Sensitivity on Identification and Delineation of Intracranial Hemorrhage Lesion Using Cascaded Deep Learning Models. Journal of Digital Imaging 32: 450-461, 2019
- [4] M. Islam, P. Sanghani, A. A. Q. See et al: ICHNet: Intracerebral Hemorrhage (ICH) Segmentation Using Deep Learning. BrainLes 2018: 456-463, 2019
- [5] J-ASPECT study, <u>https://j-aspect.jp/</u>

Hierarchical Method for Extracting Intracerebral Hematoma Regions from Thick-slice CT Images of the Head

Kazunori OKA^{*1}, Daisuke FUJITA^{*1}, Yasunobu NOHARA^{*2}, Sozo INOUE^{*3} Koichi ARIMURA^{*4}, Koji IIHARA⁵, Syoji KOBASHI^{*1}

- *1 University of Hyogo
- *2 Kumamoto University
- *3 Kyushu Institute of Technology
- *4 Kyushu University
- *5 National Cerebral and Cardiovascular Center

Cerebrovascular disease is the fourth leading cause of death in Japan, with approximately 100,000 deaths in 2019. ICH (intracerebral hemorrhage) is difficult and time-consuming to diagnose even for specialists, so an automatic method for extracting ICH regions from brain CT images is needed to reduce the burden on physicians and improve the speed and accuracy of diagnosis. In this study, we developed an automatic extraction method for ICH regions from brain CT images. In this study, we propose an automatic ICH region extraction method from brain CT images that reduces the overextraction of high-absorption regions. The proposed method focuses on the anatomical structure of ICH, extracts the high-absorption regions, and proposes a hierarchical method based on class classification using deep learning. The performance of the model was evaluated using 30 subjects with intraventricular hemorrhage, and a sensitivity of 83% and a specificity of 91% were recorded.

Key words: Deep learning, Intracerebral hemorrhage, Computed tomography, Region extraction, Hierarchical approach

Cross-modality segmentation by CycleGAN and Bayesian U-net for muscle volumetry in MRI using CT training data set

Ganping LI^{*1}, Yoshito OTAKE^{*1}, Naoki NAKANISHI^{*1},

Masashi TANIGUCHI^{*2}, Masahide YAGI^{*2}, Noriaki ICHIHASHI^{*2},

Mazen SOUFI^{*1}, Keisuke UEMURA^{*1, 3}, Masaki TAKAO^{*4},

Nobuhiko SUGANO^{*3}, Yoshinobu SATO^{*1}

Abstract

Generative methods based on neural networks have been proved efficient in cross-modality medical image translation and segmentation tasks. Meanwhile, the scarcity of medical datasets with high-quality annotations has become the bottleneck of the segmentation performance of neural networks. The complex anatomical structures make the annotation task extremely time-consuming. In this work, we investigate the feasibility of the image translation from a specific modality (e.g., MRI) to another (e.g., CT) with CycleGAN for the cross-modality segmentation, where annotations performed on one modality are used for training of the segmentation model on another modality. We first trained a CycleGAN model using training data set of 462 CTs and 136 MRIs, and a Bayesian U-net segmentation model using training data set of 20 manually annotated CTs. The CycleGAN translated the target MRI to CT-like images, which were automatically segmented by the Bayesian U-net. Additionally, we discuss the possible relationship between mutual information and dice coefficient, together with uncertainty and dice coefficient. The experiments on quadriceps muscles using fully-manually segmented MRIs of 80 subjects showed dice coefficient of 0.768 ± 0.063 (mean \pm std).

Keywords : CycleGAN, Bayesian U-net, Image translation, Image segmentation

1. Introduction

A medical dataset with corresponding labels is an essential prerequisite for many applications such as detection, classification, and semantic segmentation. Meanwhile, the monotonous and time-consuming nature of manual

*1 Imaging-based Computational Biomedicine Lab, Nara Institute of Science and Technology

[8916-5 Takayama-cho, Ikoma, Nara 630-0192, Japan]

- e-mail: li.ganping.lc2@is.naist.jp
- *2 Human Health Sciences, Graduate School of Medicine, Kyoto University
- *3 Department of Orthopaedic Medical Engineering, Osaka University Graduate School of Medicine
- *4 Department of Orthopaedics, Osaka University Graduate School of Medicine

annotation made acquiring high-quality medical images with annotations a much laborious work [1]. Given this dilemma, approaches like image synthesis and cross-modality segmentation based on existing medical data have been proposed [2].

Computed Tomography (CT) is praised for its high image acquisition speed and highlighting bone boundaries as a routine inspection. However, the radiation used in CT scans might lead to cell damage, which triggers the researches on image translation from MR images, a considerably more innocuous modality with superior performance in soft-tissue contrast, to CT images [3]. In this respect, medical image synthesis using generative models, precisely, generative adversarial networks (GANs) like CycleGAN [4], have exposed great potential on domain shifting (e.g., from MR domain to CT domain).

We performed cross-modality segmentation using CycleGAN and CT-trained Bayesian U-net. The contribution of this paper is two-fold: 1) analysis of the metrics to predict accuracy in cross-modality segmentation without using the ground truth, namely mutual information in image translation and uncertainty in segmentation, 2) extensive evaluation of the segmentation accuracy using 80 annotated data set, including comparison with single-modality segmentation as the upper limit of the accuracy in this data set.

2. Methods

This study implemented two methods: cross-modality segmentation using CycleGAN + CT-trained Bayesian U-net, and MR-trained Bayesian U-net. In cross-modality segmentation, we used only unlabeled CT and MR images to train an extended CycleGAN with gradient consistency (GC) loss, which helps contour alignment between images in two domains, to generate synthesized CTs [3]. Then, we prepared a 23-calss Bayesian U-net [5] trained with labeled CT dataset to perform synthesized CTs segmentation. Mutual information between MRIs and synthesized CTs, dice coefficient between cross-modality predictive labels and ground truth of MRI (quadriceps muscles), and uncertainty generated by CT-trained U-net were calculated and preserved as metrics. Finally, we trained an MR-trained Bayesian U-net model to identify the upper limit of the CycleGAN + CT-trained Bayesian U-net using a k-fold cross-validation.



Fig. 1 Workflow of implemented methods. The evaluation process holds all metrics data.

3. Experiments

We prepared three compound datasets in this experiment for image translation and segmentation. CycleGAN dataset consists of 462 CT cases and 136 MR cases without labels. Bayesian U-net of CT training. dataset consists of 20 annotated cases associated with manual segmentation labels of 19 muscles around the hip and thigh, pelvis femur, and sacrum bones. The field of view is the quadriceps femoris region from the iliac spine to the knee joint space, and the slice thickness are 1mm and 4mm for all CT slices and all MRI slices, respectively. All slices were resized from 512 by 512 to 256 by 256 matrix after a WL/WW setting. Additionally, 80 out of 136 cases MR cases with quadriceps labels were organized as 4-fold dataset for MR-only training using Bayesian U-net, where 60 cases were used for training, 20 cases for testing in each fold. Figure 1 gives a more intuitive demonstration of the workflow.

4. Results

The averaged accuracies of quadriceps muscles over each method can be inferred at figure 2, showing the segmentation accuracy using Bayesian U-Net for the cross-modality segmentation and MR-only segmentation. For the CycleGAN + CT method, it reaches a DC of 0.768 ± 0.0625 (mean \pm std) without using MR ground truth for training. Our CycleGAN + CT method does not require any annotation in the MR domain, allowing us to perform cross-modality segmentation when facing a new image-only MR dataset using current CT labels.



Fig. 2 Segmentation accuracy of quadriceps muscles for 80 patients with CycleGAN + CT method and MR-only method ((mean \pm std).). "MR" denotes the Bayesian U-net results of MR + Ground truth, while "SYN" infers that of CycleGAN and CT ver. Bayesian U-net using unlabeled CT + MR dataset and labeled CT dataset.

Figure 3 shows scatter plots of dice coefficient for the muscle structures as a specific function of its PSV and the consistency between increased mutual information trend and increased segmentation accuracy (i.e., dice coefficient). The structure of the rectus femoris shows a high correlation between Dice and uncertainty, also a relatively high correlation between Dice and mutual information, while the rest show relationships of intermediate level. Representative samples are manifested in figure 4.



Fig. 3 Relationship between the uncertainty metric and segmentation accuracy presented by dice coefficient. (a-c) Scatter plots with the 95% confidence ellipses, of which DC of each structures act as a function of PV. (d-f) That of MI for each structure as the variable of DC. The symbol "*r*" indicates Pearson's correlation coefficient.



Fig. 4 Sample for a representative case (Patient #131). (a) Original MRI volume consists of three series with overlay. (d-e) predicted labels from U-net were selected and merged into labels of three classes. (h-l) Representative thigh sample with low (Dice < 0.5) cross-segmentation accuracy per slice.

5. Conclusions

This research investigates the feasibility of the image translation from MR domain to CT domain using CycleGAN and Bayesian U-net for the cross-modality segmentation over quadriceps muscles group without any MR annotaion. Moreover, it also infers a possible relationship between cross-modality segmentation accuracy and mutual information among original MR images and their synthesis. Subsequently, we trained an additional Bayesian U-net for a SOTA segmentation benchmark, revealing a considerable potential improvement of the CycleGAN-CT method. Inspired by the study by Hiasa et al. [5], the cross-modality segmentation results of 23-class labels can be used as the start point of a new 23-class dataset, reducing the manual annotation cost significantly.

Competing interests

The authors have no competing interests.

Acknowledgement

This study was supported by the Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research (KAKENHI) Numbers 19H01176 and 20H04550.

References

- [1] Kazeminia, Salome & Baur, Christoph & Kuijper, Arjan & Ginneken, Bram & Navab, Nassir & Albarqouni, Shadi & Mukhopadhyay, Anirban. (2020). GANs for Medical Image Analysis. Artificial Intelligence in Medicine. 109. 101938. 10.1016/j.artmed.2020.101938.
- [2] Wolterink, J.M., et al.: Deep MR to CT synthesis using unpaired data. In: International Workshop on Simulation and Synthesi in Medical Imaging, Springer (2017) 14–23
- [3] Hiasa, Yuta & Otake, Yoshito & Takao, Masaki & Matsuoka, Takumi & Takashima, Kazuma & Prince, Jerry & Sugano, Nobuhiko & Sato, Yoshinobu. (2018). Cross-modality image synthesis from unpaired data using CycleGAN: Effects of gradient consistency loss and training data size.
- [4] Zhu, J.Y., et al.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2017) 2223–2232.
- [5] Hiasa, Yuta & Otake, Yoshito & Takao, Masaki & Ogawa, Takeshi & Sugano, Nobuhiko & Sato, Yoshinobu. (2019). Automated Muscle Segmentation from Clinical CT Using Bayesian U-Net for Personalized Musculoskeletal Modeling. IEEE Transactions on Medical Imaging. PP. 1-1. 10.1109/TMI.2019.2940555.

ポータブル型トモシンセシスにおける

X線源と検出器の位置関係推定法

堀 拳輔*1 小池 貴久*1 只野 喜一*1 橋本 雄幸*1

要旨

トモシンセシスは制限角度内で取得した投影像を用い、断層像を再構成する技術である. ポータブル X 線 撮影は病室から撮影室まで移動できない患者の体内画像を得るために用いられているが、現在の装置では 投影画像しか得られない. また、ポータブル X 線撮影において検出器と X 線源の位置関係は一意に定まら ないため、トモシンセシス画像再構成により断層像を得ることは困難である. 我々は、ポータブルトモシ ンセシスを実現するために、金属マーカーとスタイロフォーム板で作成したマーカーパネルを検出器に装 着するだけで検出器と X 線源の位置関係を推定できる手法を考案した. この手法により、検出器と X 線源 の位置関係を幾何学的に計算することが可能となる. 本研究では、提案手法の妥当性について数値シミュ レーションにより検討した. 推定誤差は数 mm 以内であることがわかり、ポータブル型でもトモシンセシ スを実現できる可能性が示唆された.

キーワード:トモシンセシス,ポータブル撮影,幾何学的位置推定

1. はじめに

トモシンセシスは制限角度内で得られた複数の投影画像から断層画像を再構成する技術であり、乳房、胸部および整形外科領域のX線撮影で利用されている[1].一方、患者の経過観察や処置後の確認のためにベッドサイドまで小型の装置を運び撮影する、ポータブルX線撮影が利用されている[2].一般的なポータブルX線撮影では、X線束に沿った経路の線積分である2次元投影画像が得られるが、画像コントラストが悪いことや病変や解剖学的な構造が重なることで診断情報の低下を招く恐れがある[3].ベッドサイドでトモシンセシスが可

*1 杏林大学大学院保健学研究科保健学 専攻

〔〒181-8612 三鷹市下連雀 5-4-1〕
e-mail: hori1911n@std.kyorin-u.ac.jp
投稿受付: 2003 年 1 月 31 日

能となれば、これまでのポータブルX線撮影に 加えて、より多くの診断情報が得られると考え られる. 先行研究では, X線 CT 画像から仮想 的に再構成したトモシンセシス画像を用いて, ポータブルトモシンセシスの有用性について 報告されている [4]. しかしながら, ポータブ ル X 線撮影では、撮影毎に検出器と X 線源を 配置するため、これらの位置関係を一意に定め ることは困難である.一方で、トモシンセシス 画像再構成において検出器と X 線源の幾何学 的位置関係は必要不可欠な条件である.これま で、検出器とX線源の位置関係を推定するため の研究がなされているが, 推定するための X 線 撮影を検査のための X 線撮影と同時に行うこ とができないため、臨床現場でのポータブルト モシンセシスの実現には至っていない [5,6]. 本研究では、ポータブルトモシンセシスを実 現するため、検出器とX線源の幾何学的な位置 関係を撮影と同時に推定できる新たな手法を

考案し,数値シミュレーションでその精度を検 証した.

2. 方法

1)提案手法

提案するポータブルトモシンセシスでは,4 枚のタングステンマーカー(10 mm×10 mm×1.0 mm)をスタイロフォーム板(420 mm×500 mm×10 mm)の四つ角に固定した「マーカーパ ネル」をフラットパネル検出器に装着して利用 する.提案手法では,第一にマーカー位置の推 定(マーカー位置推定)を行い,第二にポータ ブルトモシンセシス撮影で得られる投影画像 を用いて,X線源と検出器の位置関係を推定(X 線源位置推定)する.

マーカー位置推定では,y座標のみ異なる Pair 1のX線源位置で撮影することにより投影され た2つのマーカー画像の重心を用い,この2つ の重心を通る直線を求める.さらに,Pair2の X線源でも同様に直線を求め,この2直線の交 点を求めれば,マーカー位置のx,z座標を示す ことになる(図1(a)).一方,マーカー位置のy 座標はスタイロフォーム板の厚みで決定され るため既知となる.

X線源位置推定では、まずxおよびz座標を 求め、最後にy座標を求める.投影されたマー カーの重心 P と、前述の方法で推定したマーカ



図1 マーカー位置推定(a)と X 線源位置推定(b).

ー重心の xz 平面の位置 H'を用いて直線を求め ることができる(図 1 (b)).同様に,他のマー カーを用いれば,新たな直線を求めることがで き,この2直線の交点 Hは X線源の x および z 座標を示すことになる.次に,三角形 MPH'と 三角形 SPH が相似であることを利用して y座 標を求めることができる.この2段階の手法に より,検出器と X線源の位置関係を推定するこ とができる.X線源位置推定のための X線撮影 を検査のための X線撮影と同時に行うことが できるため,通常のトモシンセシスと比較して 追加撮影なしでポータブルトモシンセシスを 実現することが可能となる.

2) 推定方法の妥当性の検証

X線源を以下で示す座標に設置し,提案手法 によりX線源位置推定を行った.

$$\begin{cases} f_n^x [mm] = 40 \, j - 200 \\ f_n^y [mm] = 1200 \\ f_n^z [mm] = 40 \, i - 200 \\ n = 10 \, i + \, j \\ i = \{0, 1, \cdots, 10\} \\ j = \{0, 1, \cdots, 10\} \end{cases}$$
(2)

$$\operatorname{err}^{a}[\operatorname{mm}] = \frac{1}{N} \sqrt{\sum_{n=1}^{N} \left(g_{n}^{a} - f_{n}^{a}\right)^{2}}$$
 (3)



(b) X線源位置推定

ここで、 $f_n(f_1, f_2, \dots, f_n)$ は真の X 線源座標,

g_n(g₁,g₂,…,g_n)は推定された X 線源座標, N は X 線源数, a は x, y,または z で評価する方向を 示す.

3. 結果と考察

xz 方向における推定された X 線源位置を図 2 に示す.格子点が X 線源の実際の位置で,小円 で示した位置が推定位置である.x, yおよび z 方向における推定誤差はそれぞれ 6.98 ± 4.44 mm, 0.05 ± 0.07 mm および 6.97 ± 4.45 mm であった. 提案手法では推定アルゴリズムにおいて,マー カー位置推定誤差が X 線管位置推定誤差に伝 播する特徴がある.つまり, x および z 方向に おいてはマーカーから X 線管までの距離に比 例して伝播し, y 方向においては伝播しづらい 性質があるため, x および z 方向と比較して, y 方向において X 線源位置推定誤差が明らかに 小さい結果になったと考えられる.



図2X線源位置推定.格子点にX線源を配置したとき,提案手法により推定されたX線源位置を点で示す.

4. 結論

マーカーパネルを既存の検出器に装着する という簡単な手順と考案した手法を用いるこ とで、本研究で検討した一般的なポータブル撮 影で想定されるポータブル X 線撮影での検出 器と X 線源の幾何学的な位置関係を、誤差数 mm 精度で推定できることが示され、ポータブ ルトモシンセシスを実現できる可能性が示唆 された.

利益相反の有無

なし

文 献

- Dobbins JT III, Godfrey DJ: Digital x-ray tomosynthesis: current state of the art and clinical potential. Phys Med Biol 48: R65-R106, 2003
- [2] Henschke CI, Pasternack GS, Schroeder S et al: Bedside chest radiography: diagnostic efficacy. Radiology 149: 23-26, 1983
- [3] Dobbins III JT, McAdams HP, Song JW et al: Digital tomosynthesis of the chest for lung nodule detection: Interim sensitivity results from an ongoing NIH-sponsored trial. Med. Phys. Lett. 35: 2554-2557, 2008
- [4] Cant J, Snoeckx A, Behiels G et al: Can portable tomosynthesis improve the diagnostic value of bedside chest X-ray in the intensive care unit? A proof of concept study. Eur Radiol Exp 1 1: 20, 2017
- [5] Rougee A, Picard C, Ponchut C et al: Geometrical calibration of X-ray imaging chains for three-dimensional reconstruction. Comput Med Imaging Graph 17: 295-300, 1993
- [6] Cho Y, Moseley DJ, Siewerdsen H et al: Accurate technique for complete geometric calibration of cone-beam computed tomography systems. Med Phys 32: 968-983, 2005

Estimation method of the geometrical relationship

between X-ray tube and the detector in portable tomosynthesis

Kensuke HORI*1, Takahisa KOIKE*2, Kiichi TADANO*2, Takeyuki HASHIMOTO*2

*1 Kyorin University Graduate School of Health Sciences

The portable X-ray imaging is used for acquiring the X-ray radiograph of the patient who cannot go to the examination room from hospital room. However, only the 2D projection image can be obtained by the conventional portable X-ray imaging. If tomosynthesis can be performed in the bedside, the more diagnostic information can be obtained. However, it is difficult to make the geometrical relationship between X-ray tube and the detector be constraint. We proposed the novel method that estimate the geometrical relationship with only attaching special "marker panel" to the detector. In this method, the geometrical relationship can be calculated using the centroid of the metal marker and that of projected metal marker. Since the estimation error was within a few milli meter in numerical simulation, it is suggested that the portable tomosynthesis would be possible.

Key words: Tomosynthesis, Portable imaging, Geometric estimation

深層学習とディジタルファントムを用いた

骨陰影低減技術の開発

五島 風汰*1 田中 利恵*2 小田 昌宏*3 森 健策*4

高田 宗尚*5 田村 昌也*5 松本 勲*5

要旨

呼吸過程を撮影した胸部 X 線動画像による肺癌の胸壁浸潤判定の実現可能性が報告されているが, 骨陰影 が定量的な動態解析の妨げとなっている.本研究の目的は, 様々な投影方向で取得された胸部 X 線動画像 に適用可能な骨陰影低減技術を開発することである.実人体の 4 次元 CT 画像をモデル化したディジタル ファントム 20 体から疑似胸部 X 線動画像(骨陰影あり/なし)を 7200 枚作成し, 敵対的生成ネットワーク (pix2pix)を訓練した.ディジタルファントムから作成した推論用データセット(骨陰影あり/なし)を用 いて評価した結果, ピーク対雑音比(PSNR)42.4 dB, 構造的類似度指標(SSIM)0.987 と, いずれも高値 が得られた.また,臨床画像に適用したところ,時間軸方向の低減効果にばらつきがみられたものの,す べての投影方向の画像の骨陰影を選択的に低減できることを確認した.今後の課題として、動画像の処理 が可能となるよう,時間軸方向の情報を加味したネットワーク構造の改良があげられる. キーワード: 胸部 X 線動画像, pix2pix, 骨陰影低減, Bone suppression

1. はじめに

近年,一般撮影室での動画撮影を可能にする ディジタル X 線動態撮影システムが実用化さ れ,胸部 X 線動画像による肺癌の浸潤判定の実 現可能性が報告されている^[1].しかし,骨陰影 が定量評価の妨げとなっており斜位像や動画 像対応の骨陰影低減技術の開発が急務である. 本研究の目的は,斜位像や動画像対応の骨陰影

*1 金沢大学大学院 医薬保健学総合研 究科

〔〒920-0942 金沢市小立野 5-11-80〕

*2 金沢大学 医薬保健研究域附属 AI ホ スピタル・マクロシグナルダイナミクス 研究開発センター

*3 名古屋大学情報連携推進本部情報戦 略室

*4 名古屋大学大学院情報学研究科

*5 金沢大学呼吸器外科

低減技術の開発である.

胸部 X 線画像の骨陰影低減技術として,画像 回帰モデルによる手法が数多く報告されてい るが,学習データに Dual energy subtraction 法で 作成された骨陰影あり/なしのペア画像を必要 とするため正面像の処理に限定されている.そ こで,我々はディジタルファントム (4D extended cardiac-torso phantom: XCAT phantom) に注目した^[2]. XCAT phantom は様々な体格の 人体 4DCT を生成可能であり,これを疑似的に 2 次元投影することで人体を被ばくさせること なく任意の投影方向の胸部 X 線画像が作成で きる.

本稿では, Generative adversarial frameworks の ひとつである pix2pix^[3]とディジタルファント ムを用いた骨陰影低減技術を提案する.

2. 方法

2.1 データセットの作成

正常な心拍数(60 回/分)で10 秒間の努力呼

吸(呼気 6 秒+吸気 4 秒)をするという設定
で,Body Mass Index (BMI) が 22~32 [kg/m²]
の 20 体の男女の XCAT phantom の1 呼吸を
10 位相分作成した.骨陰影がない X 線画像を
作成するために肋骨・鎖骨・胸骨・肩甲骨など
の骨構造を除外した XCAT phantom も作成した.
それらをさまざまな方向から疑似的に 2 次元投影することで骨陰影あり/なしのペア画像を
3600 枚ずつ,合計で 7200 枚作成した.画像の
マトリクスサイズは 256×256 pixels, 8 bit, PNG
形式とした.作成した画像例を図 1,2 に示す.



図 1 XCAT phantom から作成した骨陰影ありの 2 次元投影像



図 2 XCAT phantom から作成した骨陰影なしの 2 次元投影像



図3 臨床画像

2.2 pix2pixの学習と評価

作成した7200枚の画像から学習用に5880枚, 検証用に1080枚,評価用に240枚にそれぞれ 使用した.pix2pixは骨陰影ありから骨陰影なし の2次元投影像を推定するように学習用画像を 与え300 epochの学習を行った.そして,検証 用データでの構造的類似度指標(SSIM)が最も 高い epochの重みを用いて評価した.本研究で は,推論画像の評価指標として,ピーク対雑音 比(PSNR)とSSIMを用いた.さらに,学習済 みモデルに診療目的で撮影された胸部 X線動 画像5例(図3)を入力し,骨陰影低減効果を 視覚評価した.なお,臨床画像の利用について は,金沢大学医学倫理審査委員会の承認を得て 行われた(承認番号1729).

3. 結果と考察

図 4,5 に提案手法から得られた骨陰影低減 画像の一部を示す.XCAT phantom から作成し た2次元投影像においては,PSNRが42.4 [dB], SSIM が 0.987 と,いずれも高値を示し,どの方 向から投影された画像においても高い骨陰影 低減効果を確認できた.また,臨床画像におい ても高い骨陰影低減効果を視覚的に確認でき た.本研究で提案した手法が,胸部X線画像の 骨陰影低減に有効であったと考える.ただし, XCAT phantom から作成した2次元投影像,臨 床画像の両方において,時間軸方向に骨陰影低 減効果のばらつきがみられた.今後,時間軸方 向の情報も加味したネットワークモデルに改

良することで改善が期待できる.



図4 生成した骨陰影低減画像(XCAT phantom)



図5 生成した骨陰影低減画像(臨床画像)

4. 結語

本研究では、XCAT phantom から作成したデ ータセットで学習させた pix2pix による胸部 X 線画像の骨陰影低減手法を提案した. XCAT phantom と深層学習を用いた斜位像や動画像対 応の骨陰影低減技術開発の実現可能性が示さ れたと考える.時間軸を加味した処理が可能と なるよう、モデルの改良が今後の課題である.

利益相反の有無 なし

文 献

[1] Tanaka R, Samei E, Segars WP, et al. Preoperative assessment of pleural invasio n and adhesion of lung tumors with dyna mic chest radiography: A virtual clinical tri al. MedPhys, 2021; 48(4):1616-1623.

- [2] Segars WP, Sturgeon S, Mendonca S, et al.
 4D XCAT phantom for multimodality imaging research. Med. Phys. 2010; 37(9): 4902-4915.
- [3] Isola P, Zhu JY, Zhou T, et al. Image-toimage translation with conditional adversarial networks. Proceedings of the IEEE conference on computer vision and pattern recognition, 2017: 1125-1134.

Development of Bone Suppression Technology

Using Deep Learning and Digital Phantom

Futa GOSHIMA^{*1}, Rie TANAKA^{*2}, Masahiro ODA^{*3}, Kensaku MORI^{*4}, Munehisa TAKATA^{*5}, Masaya TAMURA^{*5}, Isao MATSMOTO^{*5}

*1 Graduate School of Health Sciences, Kanazawa University

*2 AI Hospital/Macro Signal Dynamics Research and Development Center, Kanazawa University

*3 Information Strategy Office, Information and Communications, Nagoya University

*4 Graduate School of Informatics, Nagoya University

*5 Department of Thoracic Surgery, Kanazawa University

Pleural invasion of the tumors could be assessed by dynamic chest radiography, but negatively affected by the rib shadow in conventional images. The purpose of this study was to develop the bone suppression technique that can be applied to dynamic chest radiographs acquired in various projection directions. A total of 7200 chest X-ray images (paired images with and without bone shodows) were created from 20 digital phantoms modeled from 4-dimensional computed tomography of real patients, and were used for training an adversarial generative frameworks (pix2pix). The resulting images of the test dataset showed 42.4 [dB] in peak signal-to-noise ratio (PSNR) and 0.987 in structural similarity index measure (SSIM), suggesting pix2pix model effectively worked for projection images in various directions. Although there was a temporal variation in bone suppression effect, the bone shadow was selectively suppressed both on projection images of XCAT phantom and clinical images. Further studies are required to modify the network architecture to take into account of information in the time axis direction so that it could work well on time-series images.

Key words: Dynamic chest radiography, pix2pix, Bone suppression

解剖学的情報を併用した教師なし深層学習による PET 画像のノイズ除去

大西 佑弥^{*1} 橋本 二三生^{*1} 大手 希望^{*1} 大庭 弘行^{*1} 大田 良亮^{*1} 吉川 悦次^{*1} 尾内 康臣^{*2}

要旨

教師あり CNN を用いて PET 画像のノイズ除去を行う場合,一般的に患者侵襲を伴う高線量または長時間 計測および低線量または短時間計測の画像ペアを多数準備する必要がある.そこで,我々は教師なし CNN を導入し,さらに MR 画像の解剖学的情報を併用することでノイズ除去性能の向上を検討する.ネットワ ークは MR 画像から解剖学的特徴を抽出する CNN と PET 画像のノイズ除去を行う CNN が 2 種類のアテ ンション機構によって結合された構造を有する.数値シミュレーションおよび実測データを用いた評価の 結果,従来手法と比較し定量性を維持しつつノイズが抑制されたことから,提案手法を用いることで PET 計測時間の短縮や RI 投与量の削減による被ばく線量の低減が期待される.

キーワード:PET, MR, ノイズ除去, 教師なし畳み込みニューラルネットワーク, アテンション機構

1. はじめに

陽電子放出断層撮影(PET) 画像のノイズ除 去に畳み込みニューラルネットワーク(CNN) を応用する場合,一般的に高線量または長時間 計測および低線量または短時間計測の画像ペ アを多数準備する必要があり,多大な労力を要 する.さらに,学習セットに含まれていない未 知症例や新規 PET 薬剤といった異なるドメイ ンに対する堅牢性が問題となっている.これら の課題を解決するために,我々は教師なし深層 学習技術の1つである deep image prior (DIP) を導入する[1]. DIP は学習データセットを必 要としないノイズを含む単一データのみでノ イズ除去を行うことが可能であり,CNN の構造 自体を事前情報と捉えた最適化問題として考 えることができる.

PET 画像のノイズ除去における先行研究では, DIP の初期入力にランダムノイズではなく磁気

〔〒434-0041 浜松市浜北区平口 5000〕

e-mail: yuya.onishi@hpk.co.jp

*2 浜松医科大学光先端医学教育研究センター生体機能イメージング研究室

共鳴(MR) 画像などの解剖学的情報を用いるこ とでノイズ除去性能が向上することが報告さ れている [2-4].しかし,解剖学的情報の特徴 や微細構造を十分に活用できているかは未だ 不明確である.そこで本研究では,MR 画像の 解剖学的情報をより明示的に利用するために, DIP アルゴリズムのネットワーク構造にアテン ション機構を加えた MR-guided deep decoder (MR-GDD) を考案する.

2. 方法

MR-GDD は encoder-decoder subnetwork と deep decoder subnetwork の2つの CNN から構成 される [5]. Encoder-decoder subnetwork は、MR 画像から複数スケールでの階層的特徴を抽出 し, deep decoder subnetwork は、初期入力され たランダムノイズからノイズ除去された PET 画像を復元する.双方のネットワークは2種類 のrefinement unit によって結合されており, MR 画像から抽出された特徴がアテンション機構 を介して deep decoder subnetwork に重み付けさ れる.アテンション機構を介した解剖学的情報 の入力は MR 画像の微細構造やセマンティック 特徴をより明示的に利用することが可能であ

^{*1} 浜松ホトニクス株式会社中央研究所

る. MR-GDD の損失関数を以下に示す.

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \|x_0 - f_{\theta}(z,g)\|, \ x^* = f_{\theta^*}(z,g) \quad (1)$$

||・||は L2 損失であり, ネットワーク f_θ にランダ ムノイズ z と MR 画像 g を入力したときの出力 画像 f_θ(z,g) とノイズを含む PET 画像 x₀ の平 均二乗誤差を最小化する. 損失関数の最適化に は準ニュートン法である L-BFGS アルゴリズム を用いることでネットワーク重みの更新を行 い, ノイズが復元される前に更新を停止させる ことでノイズ除去された PET 画像を得る.

3. 結果

¹⁸F-FDG, ¹¹C-raclopride, ¹⁸F-florbetapir による モンテカルロシミュレーション及び実測デー タを用いて提案手法の評価を行った. ノイズを 含む画像は得られたリストデータの全カウン トをダウンサンプリングすることで作成した.

シミュレーションによる定量評価の結果,従 来のノイズ除去アルゴリズムであるガウシア ンフィルタ(GF),MR 画像を DIP に直接入力 する MR-DIP と比較し,ピーク信号対雑音比 (PSNR),構造的類似性(SSIM),コントラス ト対ノイズ比(CNR)が向上した(図1).人の 実測データにおいても,従来手法と比べ統計ノ イズを抑制しつつ,組織コントラストが保存で きている(図2).よって,提案手法を用いるこ とで PET 計測時間の短縮や RI 投与量の削減に よる被ばく線量の低減が期待される.



図1 模擬腫瘍を挿入したシミュレーションデ ータによるノイズ除去結果



図2 人の実測データによるノイズ除去結果

4. まとめ

本研究では、教師なし CNN を用いた PET 画 像のノイズ除去に対し、アテンション機構を介 して解剖学的情報を加える MR-GDD を考案し た.様々な PET 薬剤を統一的に解けるフレーム ワークであり、従来手法を上回るノイズ除去性 能が示された.提案手法を用いることで PET 計 測時間の短縮や RI 投与量の削減による被ばく 線量の低減が期待される.

利益相反の有無 無し

文 献

- Ulyanov D, Vedaldi A, and Lempitsky, V:
 Deep image prior. In Proc. IEEE CVPR: 9446-9454, 2018
- [2] Hashimoto F, Ohba H, Ote K et al: Dynamic PET Image Denoising Using Deep Convolutional Neural Networks Without Prior Training Datasets. IEEE Access 7: 96594-96603, 2019
- Cui J, Gong K, Guo N et al: PET image denoising using unsupervised deep learning. Eur. J. Nucl. Med. Mol. Imaging 46: 2780-2789, 2019
- [4] Gong K, Catana C, Qi J et al: PET Image Reconstruction Using Deep Image Prior. IEEE Trans. Med. Imag 38: 1655-1665, 2019
- [5] Uezato T, Hong D, Yokoya N et al: Guided Deep Decoder: Unsupervised Image Pair Fusion. In Proc. ECCV: 87-102, 2020

PET Image Denoising Using Unsupervised Learning with Anatomical Information

Yuya ONISHI^{*1}, Fumio HASHIMOTO^{*1}, Kibo OTE^{*1}, Hiroyuki OHBA^{*1}, Ryosuke OTA^{*1}, Etsuji YOSHIKAWA^{*1}, Yasuomi OUCHI^{*2}

*1 Central Research Laboratory, Hamamatsu Photonics K. K.
*2 Hamamatsu University School of Medicine.

Although supervised convolutional neural networks (CNNs) often outperform conventional alternatives for denoising positron emission tomography (PET) images, they require many low- and high-quality reference PET image pairs. Herein, we propose an unsupervised 3D PET image denoising method based on attention-guided anatomical information. Our proposed magnetic resonance-guided deep decoder (MR-GDD) utilizes the spatial details and semantic features of MR image more effectively by introducing encoder-decoder and deep decoder subnetworks. Monte Carlo simulation using the ¹⁸F-fluoro-2-deoxy-D-glucose (FDG) shows that the proposed method outperforms other denoising algorithms in terms of the highest peak signal-to-noise ratio and structural similarity. For preclinical (using ¹⁸F-FDG and ¹¹C-raclopride) and clinical (using ¹⁸F-florbetapir) studies, the proposed method demonstrates state-of-the-art denoising performance while retaining spatial resolution and quantitative accuracy, despite using only a single architecture for various noisy PET images. These results suggest that the proposed MR-GDD can reduce PET scan times and PET tracer doses considerably without impacting patients.

Key words: Positron emission tomography, Magnetic resonance, Denoising, Unsupervised deep learning, Attention

ピンホール SPECT システムにおける

散乱線の定量と除去法の比較

島田 良*1 村田 一心*2 尾川 浩一*3

要旨

マルチピンホール SPECT システムは、検出器の投影方向数を減らしデータの収集時間を短縮できるため、臓器の動態イメージングを目的として研究が進んでいる.しかし、ピンホールコリメータは、開口角が広いため周囲の臓器で散乱した不要な光子を通過させてしまう恐れがある.心筋イメージングの場合、 肝臓からの散乱光子の多くが心筋の画質に影響を与える.これらの散乱光子の影響を調査するために、 MCAT ファントムを使用してモンテカルロシミュレーションを実行した.さらに、散乱補正方法である Triple Energy Window(TEW)法と散乱カーネルを使用したデコンボリューション法の散乱線除去精度の 比較を行った.シミュレーションでは、再構成された画像とそのプロファイルから精度を評価した.その 結果、TEW 法がデコンボリューション法よりも正確に散乱光子を除去できることが明らかとなった.

1. はじめに

近年、マルチピンホール SPECT システム は、検出器の投影数を減らして臓器の動態を 観察することが可能となるため研究が進んで いる. このピンホールコリメータは、様々な 角度から検出器に入射する光子を検出できる ため、データ取得時間を短縮することができ る. しかし、ピンホールコリメータは広い角 度範囲から入射する光子を通過させるため、 従来のパラレルホールコリメータに比べてよ り多くの散乱光子が検出器に混入してしまう 可能性があり、再構成画像のコントラスト分 解能が低下してしまうことが予想される.本 論文では、マルチピンホール SPECT システム における散乱線を除去することを目的として、 Triple Energy Window(TEW)法と散乱カーネル

e-mail: ryo.shimada.7z@stu.hosei.ac.jp

*2 法政大学理工学部

*3 国立天文台

を使用したデコンボリューション法の両方の 散乱補正法の性能を比較した.

2. シミュレーション方法

図 1 に, 使用した SPECT システムの検出 器のジオメトリを示す.検出器のサイズを 512×256 pixel (ピクセルサイズ: 0.8 mm×0.8 mm), 検出器の有効面積を 400 mm × 200 mm, エネルギー分解能を 10% FWHM (@140 keV) とした. また, 簡単のため検出器の固有分解 能は無視した.本研究ではピンホール数を3 とした. 図2は,3個のピンホールを備えた マルチピンホールコリメータの平面図である. コリメータプレートの厚さは1cm, ピンホー ルの直径は 5 mm で, 各ピンホールにはコリ メータの深さの半分にナイフエッジがあるも のとし、ピンホールの開口角は48度とした. この3個のピンホールは、対象物の投影像の 重なりが起こらないように配置されている. 物体の中心とコリメータの間の距離は25 cm,

^{*1} 法政大学大学院理工学研究科

^{〔〒184-8584} 東京都小金井市梶野町

³丁目 7-2〕

コリメータと検出器の間の距離は 7.6 cm と し、このシミュレーションでは、MCAT ファ ントムを使用した. このファントムの線源分 布と吸収係数を図 3 に示す.放射能の集積比 は心筋:血液:肝臓で 10:2:1 に設定した. 再構 成領域は 128 × 128 × 128 voxel、ボクセルサ イズは 2.0 mm × 2.0 mm である.







図2マルチピンホールコリメータ

モンテカルロ法を使用してファントム内の 光子を輸送した.考慮した相互作用はコンプ トン散乱,コヒーレント散乱,および光電効 果であり,散乱の最大数は5回とした.まず, ピンホールの感度の不均一性を投影データに おいて補正した[2].画像再構成はML-EM 法で行い,反復回数は50回とした.この反 復画像再構成の過程でガンマ線の吸収を補正 した.また,7rays法を使用して空間分解能を 向上させた. 散乱補正法には TEW 法 [3] とデコンボリ ユーション法 [5] を用いた. TEW 法では, 検出器のメインウィンドウ幅を 140 keV ± 10%, サブウィンドウ幅を 7 keV とした. デコンボリューション法では,厚さ 1 ~ 35 cmの水の平板ファントムを想定して1 cm 間 隔で散乱カーネルを作成した. このカーネル のサイズは 65 × 65 pixel (ピクセルサイズ: 0.8 cm × 0.8 cm), エネルギー分解能は 10% FWHM (@140 keV) で, ウィンドウ幅は 140



 keV ± 10% とした.補正は反復画像再構成の プロセスにおいて、この散乱カーネルを計算 投影データに畳み込むことによって実現した.
 図 3 MCAT ファントム.放射能分布(左)、 吸収係数分布(右)

3. 結果と考察

シミュレーションの結果, MCAT ファント ムを使用した場合, FWHM 10%のエネルギー 分解能を持つ NaI 検出器ではメインウィンド ウ内に約 21%の散乱線が混入することが確 認された.

再構成画像を図 4 に, 再構成画像の心筋の 短軸に沿って設定されたプロファイルを図 5 に示す. TEW 法による散乱補正を行った再構 成画像のプロファイルは, プライマリのみを 再構成した際のプロファイルとほとんど一致 していることが確認された. 一方, デコンボ リューション法による補正画像は, 心筋の放 射能を過大評価していた. この結果から, TEW 法の方がデコンボリューション法より も正確であることがわかる. これは, TEW 法 は散乱体の分布に応じて散乱光子の量を推定 できるためであると考えられる.





4. まとめ

本論文では, ピンホール SPECT システムに おける散乱線除去を目的として TEW 法とデ コンボリューション法の散乱補正方法の評価 を行った.その結果, TEW 法がデコンボリ ューション法よりも正確な散乱補正をもたら すことが明らかとなった.

謝辞

本研究の一部は文部科学省科学研究費補助 金基盤研究(c) 19K12849 による助成を受けた.

利益相反の有無

なし

文 献

- Jansen FP, et al., "Uniformity correction using non-uniform floods," *IEEE Nucl. Sci. Symp. Conf. Record (NSS/MIC)*, pp. 2314-2318, 2010.
- [2] Andreyev A, et al., "Pinhole SPECT reconstruction using blobs and resolution recovery," *IEEE Trans. Nucl. Sci.*, vol. 53, no. 5, pp. 2719- 2728, 2006.
- [3] Ogawa K, et al., "A practical method for position-dependent Compton-scatter correction in single photon emission CT," *IEEE Trans. Med. Imag.*, vol. 10, pp. 408–412, 1991.
- [4] Fufishiro Y, et al., "Comparison of a Multi-pinhole Stationary SPECT System and a Parallel-hole High Speed Rotational SPECT System," *IEEE Nucl. Sci. Symp. Conf. Record (NSS/MIC)*, 2019.
- [5] Shao L, et al., "Cross-plane scattering correction-point source deconvolution in PET," *IEEE Trans. Med. Imag.*, vol. 10, pp. 408–412, 1991.

Comparison of Scatter Correction Methods and Quantification in a

Multi-pinhole SPECT

Ryo SHIMADA*1, Kazumi MURATA*2, Koichi OGAWA*3

*1 Graduate School of Science and Engineering, Hosei University

*2 National Astronomical Observatory Japan

*3 Faculty of Science and Engineering, Hosei University

Multi-pinhole SPECT systems have been studied for dynamic imaging of organs, because they can reduce the number of projection angles of the detector and the data acquisition time. However, due to the wide aperture angle of a pinhole collimator, unnecessary photons scattered by surrounding organs may pass through the collimator. In the case of myocardial imaging, the scattered photons from the liver affect the image quality of the myocardium. To investigate the effect of these scattered photons, Monte Carlo simulations were performed using the MCAT phantom. In addition, we compared the accuracy of the scatter correction methods; the Triple Energy Window (TEW) method and the deconvolution method using a scattering kernel. In the simulation, the accuracy was evaluated with the reconstructed images and their profiles. The results showed that the TEW method can remove scattered photons more accurately than the deconvolution method.

Key words: multi-pinhole collimator, SPECT system, scatter correction, Monte Carlo simulation, myocardial imaging.

デコンボリューション法による静止型マルチピンホール

SPECT システムの空間分解能改善

大越 迪*1 村田 一心*2 尾川 浩一*3

要旨

従来の SPECT システムでは、データを取得する際に検出器を回転させるため、データ収集に時間がかかっ てしまうという課題がある.そこで近年、検出器を回転させない静止型ピンホール SPECT システムの開発 が進められている.このシステムでは、検出器を回転させない代わりに、複数のピンホールを有するコリ メータを用いているが、ピンホール径を大きくすると、画質の劣化が起こってしまう.これを補正する方 法として、従来 7rays 法などが提案されてきたが、ピンホール径が大きい場合には近似精度が悪く、様々な 補正との絡みから十分な方法とは言えなかった.そこで本研究では、点拡がり関数を推定し、マルチピン ホールコリメータで取得した投影データに対して実空間でデコンボリューションを行うことで、空間分解 能の改善を試みた.再構成画像の画質を評価したところ、従来の方法よりも空間分解能の改善, PSNR 値の 向上を確認することができた.

キーワード:SPECT, ピンホールコリメータ, デコンボリューション, 点拡がり関数

1. はじめに

医療画像診断装置の一つである SPECT は, 体内に投与した放射性物質を含む薬剤から放 出される放射線を検出し, X線 CT や MRI など では得ることのできない, 臓器の機能情報を得 ることができる.しかし,この SPECT システム では,検出器を回転させることでデータの取得 を行っているため,データ収集に時間がかかり 動態情報を得ることが難しい.そこで近年,検 出器を回転させない静止型 SPECT システム

*1 法政大学大学院理工学研究科応用情 報工学専攻

〔〒184-8584 東京都小金井市梶野町 3-7-2〕

e-mail: michi.okoshi.6u@stu.hosei.ac.jp

*2 国立天文台

*3 法政大学理工学部応用情報工学科

[1][2]の開発が進められている. このシステムで は、検出器を回転させない代わりに、複数のピ ンホールを有するマルチピンホールコリメー タを使用することで、投影方向数を増加させる とともに、感度を改善している. しかし、この ようなピンホールコリメータを用いたデータ 収集では、ピンホール径が有限の大きさを持つ ため、空間分解能が大きく劣化する.この劣化を 補正する方法として、従来より 7rays 法[3]など が提案されてきた. しかし、この方法は本来無 数の方向から飛来する光子をわずか7本の投影 線で近似するため、特にピンホール径が大きい 場合に空間分解能の劣化を十分に補正するこ とができない.

そこで本研究では、実空間における逆畳み込 み法を用いた空間分解能改善手法を提案する. SPECT システムにおける投影データは、ピンホ ール径に応じたボケの影響を受けるため、点拡 がり関数 (point spread function; PSF)でボケ具合 を推定し、投影データに対して逆畳み込みを行 うことで、ピンホール径が無限小の場合の投影 画像に近似することができる.この投影データ を用い再構成を行うことで、ピンホール径が有 限であることによる空間分解能の劣化を防ぎ, 高画質な再構成画像を取得できると期待され る. 本研究では、その有用性を検証するため、 静止型 SPECT システムにおけるシミュレーシ ョンを行った.

2. シミュレーション

本研究では,静止型ピンホール SPECT シス テムに対して、PSF による逆畳み込みを用いた 補正方法の有用性を示すためにシミュレーシ ョンを行った.本シミュレーションでは、図1 に示すように、 ターゲットである脳を模擬した ファントム(図 3(a))の周囲に検出器を 6 個配置 した静止型 SPECT システムを用い, コリメー タ上に3つのピンホールを配置したマルチピン ホールコリメータを適用した.また,物体中心 からコリメータ面の中央までの距離を 25 cm, コリメータ面から検出器までの距離を7.5 cmと し, 検出器のサイズは 512×256 pixels, ピクセ ルサイズを 0.8 mm×0.8 mm とした. ピンホー ルは、コリメータ中心から 10cm 間隔で三つ配 置し, その際のピンホール径は, 本研究におけ る提案手法の効果を検証するため、大きめの5 mm とした. この時の開口角は, 先行研究[2]に 合わせ24度,形状はナイフエッジとした.

これらの条件の下,モンテカルロ法による光 子輸送シミュレーションを行った.図3(a)の灰 色領域から1×10⁷個,白色領域から3×10⁷個 の光子を放出した.また本シミュレーションで は簡単のため,検出器の固有分解能は考慮して おらず,ターゲットによる散乱も計算に入れて いない.すなわち理想的な検出器を想定し,散 乱線は完璧に除去できたものとした.光子の入 射方向による感度不均一性は,モンテカルロ法 を用いた手法[4]により補正した.

このようにして得られた投影データ(図 2(b)) は、ピンホール径に応じたボケが発生している. これに対して PSF の逆畳み込みを適用するのが 本研究の提案手法である.本研究では、再構成 領域の中心に置かれた線源による PSF を、ジオ メトリから解析的に推定し、それを投影データ に対し適用した.簡単のために、3 つのピンホ ール全てに同じ PSF を適用し、逆畳み込みは実 空間上で行った.実空間上での逆畳み込みは、 Richardson-Lucy アルゴリズム[5]を実装した.画 像再構成は、ML-EM 法[6]を用い、反復回数は 50 回とした.提案手法と比較を行うために ML-EM 法に 7rays 法を組み込んだものも実装した.



図1 シミュレーションジオメトリ.脳ファントムの周りに、検出器を6つ配置した.物体と検出器の距離などは、先行研究とほぼ同じになるようにした.

3. 結果

図2に処理後の投影データの比較を示す.無限小ピンホールで得られた理想的な投影データ(a)と,ピンホール径の影響のある投影データ(b)を比較すると,ピンホール径の大きさが原因でボケが生じていることが分かる.一方,この投影データ(c)は,(b)の投影データと比較すると,理想的な投影データ(a)により近くなっており,またより細かい構造が確認できることが分かる.このことから,逆畳み込みをすることで,ピンホール径の大きさによる影響を改善することができていることが分かる.

図3に、再構成画像の比較を示す.ボケの補 正を行っていない投影データと提案手法によ り補正を行った投影データから再構成を行い、 脳ファントムの原画像(a)との比較を行った.今 回提案した、逆畳み込み後の投影データを再構 成に適用した結果の画像(c)は、従来手法である 7rays 法を適用して空間分解能の改善を行った 結果の画像(b)より, (a)に近づいた画像となって おり,細部の構造まで再構成できていることが 分かる.また, PSNR 値の比較を行うと,従来 手法では12.7 dB,提案手法では13.5 dB となっ た.この結果からも,提案手法が,空間分解能 の改善に寄与していることが分かる.



図2 投影データの比較. (a)無限小ピンホール で得られた理想的な投影データ, (b)シミ ュレーションで得られた投影データ(ピ ンホール径5mm), (c) (b)の投影データ に逆畳み込みを行った投影データ(提案 手法).



 図3 再構成画像の比較. (a)脳ファントムの原 画像, (b)従来の 7rays 法を用いた再構成 画像, (c)逆畳み込み後の投影データを再 構成した画像.

4. 考察

図2の投影画像の比較より,再構成領域の中 心に線源を置いた場合のPSFを推定し,それを 用いて逆畳み込みを行った結果,ピンホール径 の大きさによるぼけを補正することができ,よ り無限小ピンホールでの投影データに近づく ことが分かった.また,図3の再構成画像の比 較より,逆畳み込み後の投影データを用いた結 果の方が,従来の7rays 法を用いた結果より, 視覚的にも,また PSNR という定量的尺度を用 いた結果からも画質が改善した.したがって, 点拡がり関数を用いた逆畳み込み手法が,静止 型 SPECT システムにおける空間分解能改善に 有用であることが示された.

しかし、本研究では、点拡がり関数があくま で再構成領域の中心にある場合のみでしか行 っていない.実際には、コリメータ上のピンホ ールの位置や、ファントム内の3次元的位置に よって、点拡がり関数は大きく変わると考えら れるが、現段階では考慮できていない.そこで 今後は、コリメータ上のピンホールの位置や再 構成領域の場所による点拡がり関数の変化も 考慮したシミュレーションを行っていく予定 である.

5. まとめ

本研究では、コリメータ上に複数のピンホー ルを配置したマルチピンホールコリメータを 用いた静止型 SPECT システムにおいて点拡が り関数の逆畳み込みを適用し、空間分解能の改 善を行う新たな手法を提案し、その有用性を視 覚的ならびに定量的に示すことができた.

謝辞

本研究の一部は文部科学省科学研究費補助 金基盤研究(c) 19K12849 による助成を受けた.

利益相反の有無

なし

文 献

- H. Kubota, et al., "Multi-pinhole SPECT system with a triple head gamma camera," IEEE Nuclear Science Symposium, Medical Imaging Conference and Room-Temperature Semiconductor Detector Workshop(NSS/MIC/RTSD), DOI:10.1109 /NSSMIC.2016.8069609, 2016
- [2] Y. Fujishiro, et al., "List Mode Image Reconstruction With a Multi-pinhole Triple Head SPECT System," IEEE Nuclear Science Symposium and Medical Imaging

Conference Proceedings(NSS/MIC), DOI: 10.1109/NSSMIC.2018.8824476, 2018

- [3] A. Andreyev, et al., "Pinhole SPECT Reconstruction Using Blobs and Resolution Recovery," IEEE Transaction on Nuclear Science, vol. 53, No. 5, pp. 2719-2728, DOI:10.1109/TNS.2006.878004, 2006.
- F. P. Jansen, et al., "Uniformity correction using non-uniform floods," IEEE Nuclear Science Symposium and Medical Imaging Conference, pp. 2314-2318, DOI:10.1109/ NSSMIC.2010.5874197, 2010
- [5] W. H. Richardson, "Bayesian-Based Iterative Method of Image Restoration," J. Opt. Soc. Am., vol. 62, no.1, pp. 55-59, 1972
- [6] 篠原 広行,小島 慎也,中世古 和真, 橘 篤志,橋本 雄幸, "ML-EM 法と Maximum a Posterior-EM(MAP-EM)法," 断層映像研究会雑誌 第40巻 第2号, 2013

Improvement of the spatial resolution with a deconvolution

method for a multi-pinhole SPECT system

Michi OKOSHI*1, Kazumi MURATA*2, Koichi OGAWA*3

*1 Graduate School of Science and Engineering, Hosei University

*2 National Astronomical Observatory of Japan

*3 Faculty of Science and Engineering, Hosei University

As a conventional SPECT system rotates detectors to acquire a projection data, it takes long time to acquire data for a patient study. Recently, a static SPECT system with a multi-pinhole collimator which does not rotate detectors in data acquisition has been studied. In this system, instead of rotating detectors, a collimator with multiple pinholes is used, but the quality of an image becomes worse when we use a large pinhole collimator. To improve the spatial resolution of reconstructed images, a 7-rays method has been proposed, however it is not enough in terms of image quality in the case of a large pinhole collimator. In this study, we proposed a new method to improve the spatial resolution by estimating the point spread function, and performing a deconvolution of the projection data in the real space. We evaluated the image quality of reconstructed images with the simulations, and confirmed that the spatial resolution of deconvolved projection images and the PSNR values improved compared with the conventional method.

Key words: SPECT, pinhole collimator, deconvolution method, point spread function

Deep Image Prior による SPECT 画像の画質改善の試み

矢部 暁登^{*1} 對馬 結太^{*1} 福原 健太^{*1} 松原 佳亮^{*2} 越野 一博^{*3} 渡部 浩司^{*4} 銭谷 勉^{*1}

要旨

SPECT(Single photon emission computed tomography)装置は目的臓器に集積した放射性薬剤から放出されるガ ンマ線を検出し,体内の薬剤の分布を画像化することで,体の機能が可視化できる.我々のグループでは高解 像度 SPECT 装置を開発しており,通常の SPECT 装置より高解像度の SPECT 画像を得ることができるが, 検出効率が低く十分なカウントが得られないため,画質が劣化してしまう.そこで,事前学習が不要で,ターゲ ット画像のみで画像を復元できる DIP(Deep Image Prior)の利用を検討した.本研究では DIP を用いて SPECT 画像の画質改善を目的とする.評価実験では数値ファントムを用いて統計ノイズを付加した投影デ ータを生成し,DIP 処理をした投影データと DIP 処理をしていない投影データを OS-EM 法で画像再構成 した.作成された再構成画像を数値ファントムと比較評価した結果,DIP 処理によって PSNR は 32.26dB か ら 39.16dB,SSIM は 0.860 から 0.928 に向上した.また,実測データに DIP 処理した結果,再構成画像全体のノ イズが低減され,ファントムの構造が鮮明になった.DIP による SPECT 画像の画質改善効果が示された.

キーワード: SPECT, 画像再構成, Deep image prior

1. はじめに

核医学診断装置の一つである SPECT(Single Photon Emission Computed Tomography)装置は, 放射性同位元素を用いた放射性薬剤を体内に 投与し,目的臓器に集積した薬剤から放出され るガンマ線を検出し,体内の薬剤の分布を画像 化することで,体の機能が可視化できる.そのた め,薬剤の分布を画像化することで,体の機能情

*1 弘前大学大学院理工学研究科 〔〒036-8561 青森県弘前市文京町3〕 e-mail: h21ms425@hirosaki-u.ac.jp *2 秋田県立循環器・脳脊髄センター研究 所脳血管研究センター放射線医学研究部 *3 北海道情報大学経営情報学部システ ム情報学科 *4 東北大学サイクロトロン・ラジオアイ

ソトープセンター

報を評価できる.また,我々のグループでは,解像 度 3.5mm 程度の高解像度 SPECT 装置の開発を 行っており、通常の SPECT 装置よりも高解像 度で放射性薬剤の3次元分布を得ることがで きる[1]. しかしながら,高解像度化に伴い,ガ ンマ線の検出効率が低下し、十分なカウントを 得られず, SPECT 画像の画質が劣化してしま う. 近年,ディープラーニングを用いた画像のノ イズ除去の研究が盛んにおこなわれている[2]. その中で DIP(Deep Image Prior)という CNN (Convolutional Neural Network)を用いた画像処 理法がある. DIP は画像の自然な成分とノイズ 成分の学習の早さの違いを利用して画像のノ イズ除去、超解像、復元を行うことができ、未 学習のニューラルネットワークとターゲット 画像のみで画像処理を行うため、ネットワーク の事前学習や大量のデータセットが必要ない という利点がある[3]. また, DIP は PET(Positron Emission Tomography)の画像再構

成に対して有用であることが示唆されている [4].本研究は,DIPを用いて、ノイズ成分が多 い画像のノイズ除去を行い、SPECT 画像の画 質を改善することを目的とする.

2. 方法

2. 1 Deep Image Prior

ターゲット画像を x_0 ,入力として与えるノイ ズ画像をz,ネットワークのパラメータを θ ,ニュ ーラルネットワークを f_θ とすると,以下の式か らネットワークパラメータを更新する.

$$\theta^* = \operatorname{argmin} E(f_{\theta}(z); x_0) \tag{1}$$

また,得られる画像x*は,

$$x^* = f_{\theta^*}(z) \tag{2}$$

と表すことができる.ロス関数はターゲット画 像*x*₀と得られた画像*x*との MSE (平均二乗誤差) をとるため,以下の式で表せられる.

$$E(x; x_0) = ||x - x_0||^2$$
(3)

そのため,勾配降下法によってパラメータの更 新をしていくと,MSE の値が小さくなっていき, 得られる画像がターゲット画像に近づいてい く,図1に DIP の概要図を示す.

このネットワークの特性として,ノイズのようなランダム画像よりも,規則的な構造を持つ 画像の方が早く復元されることが示唆されて いる[3].そのため,ノイズ成分が復元される前に パラメータの更新を止めることで,ノイズ成分 が取り除かれた自然な画像を得ることができ る.



図1 Deep Image Prior の概要図

今回 DIP に用いたネットワークは図 2 (a) の ように、ダウンサンプリングモジュールとアッ プサンプリングモジュール及び U-Net のような スキップモジュールによって構成され、それぞ れのモジュール内の処理は図 2 (b) の順番で行 われる.ネットワークの構築にはニューラルネ ットワークライブラリの一つである Keras を用 いて勾配降下法には Adam を用いた.プログラ ムの実行には Google が提供している、ブラウザ から Python の記述、実行できるサービスの Google Colaboratory を使用した.



図2 ネットワークの構造の概要図

2. 2 DIPの SPECT 画像再構成への応用

SPECT 装置では複数方向からの 2 次元投影 画像データが収集される.投影画像をターゲッ ト画像として DIP で処理する.今回はメモリ容 量の制約から各方向の投影画像毎に DIP 処理を 行った. DIP 処理した投影画像データが画像再 構成に用いられる.画像再構成にはパラレルホ ールコリメータ用 OS-EM(Ordered Subsets-Expectation Maximization)が使用された.

3. 実験

シミュレーションデータと実測データを用
いて DIP の効果を評価するための実験が行われた.シミュレーションデータ,実測データともに投影画像は画素サイズ $0.81 \text{ mm} \times 0.81 \text{ mm}$,画像サイズ 256×256 ,投影数 180 cbas。OS-EM画像再構成は 5 subsets, 10 iterations で行われ,

得られた再構成画像は画素サイズ 0.81 mm× 0.81 mm×0.81 mm, 画像サイズ 256×256×256 である.

3.1 シミュレーションデータ

物理脳ファントム (図3)の数値ファントム 画像(図4)を用いた[5].図4の白い領域に放射 性薬剤が満たされる.数値ファントム画像から 投影データを作成するプログラムを用いて,投 影データを作成した.投影データのトータルカ ウントを実測データのトータルカウントに合 わせてから,統計ノイズを付加した.画像のボケ、 吸収や散乱線の影響は考慮ぜず,統計ノイズの 影響のみを考慮した.この投影データをターゲ ット画像として,DIP の処理を行った.DIP の処 理をした投影データと DIP の処理をしていない 投影データ画像を,画像再構成を行い,視覚的評 価および,作成された再構成画像と原画像であ る数値ファントムとの PSNR(Peak Signal-to-Noise Ratio),SSIM(Structual Similarity)の値を計 算し,定量的に画質を評価した.



図3 物理脳ファントムの外観



3.2 実測データ

我々の高解像度 SPECT 装置でパラレルホー ルコリメータを使用し,447MBq の Tc-99m 溶液 で満たされた物理脳ファントム(図 3)を撮像 した.得た投影データを用いた.この投影データ 画像をターゲット画像として DIP の処理を行 い,DIP の処理を行った投影データと DIP の処 理を行っていない投影データを画像再構成し, 比較した.

4. 結果

4.1 シミュレーションデータ

図 5 は横軸が DIP の epoch 数,縦軸が再構成 画像と数値ファントムとの SSIM,PSNR の値を 示したグラフである.また,図 5 の黒の点線と黒 の直線はそれぞれ DIP の処理を行っていない投 影データの再構成画像の PSNR,SSIM の値であ る.DIP によって,PSNR は 32.26dB から 39.16dB に,SSIM は 0.860 から 0.928 に向上した.



図5 数値ファントムの評価

図 6 の(a),(b)はそれぞれ DIP の処理を行って いない投影データと,epoch 数 700 で DIP の処理 を行った投影データの再構成画像である.図 6 の(a),(b)から,DIP により再構成画像全体の低ノ イズ化が確認できた.



(a)DIP 処理なし(b)DIP 処理あり図 6 数値ファントム再構成画像

4.2 実測データ

図 7(a)は投影データに DIP の処理をせずに再 構成した画像,図 7(b)は投影データに epoch 数 700 で DIP の処理をして再構成した画像である. 図 7 の(a),(b)を比較してみると,DIP の効果によ り,バックグラウンドノイズが軽減され,脳ファ ントムの構造が鮮明になっていた.



(a)DIP 処理なし(b)DIP 処理あり図 7 実測データ再構成画像

5. まとめ

Deep Image Prior を用いた SPECT 画像再構成 は,再構成画像のノイズ軽減に有効であること を確認でき,数値評価において改善を示した.ま た,実測データに対しても画像の低ノイズ化に 効果的であることが確認できた.

謝辞

本研究は JSPS 科研費 JP17H04116 の助成 を受けて実施された.

利益相反の有無

なし

文 献

- Zeniya T, Gaku K, Hori Y, et al.: Development of high-resolution brain SPECT system using full-digital gamma camera with multiple position-sensitive PMTs. In 2015 IEEE Nuclear Science Symposium Conference Record, San Diego, 2016, pp1-3
- [2] Chen H, Zhang Y, Zhang W, et al.: Low-dose CT via convolutional neural network. Biomed Opt Express 8: 679-694, 2017
- [3] Ulyanov D, Vedaldi A, Lempitsky V: Deep image prior. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, 2018, pp9446-9454
- [4] Gong K, Catana C, Qi J, et al.: PET Image Reconstruction Using Deep Image Prior. IEEE Trans Med Imaging 38: 1655-1665, 2019
- [5] Iida H, Hori Y, Ishida K, et al.: Threedimensional brain phantom containing bone and grey matter structures with a realistic head contour. Ann Nucl Med 27: 25-36 2013

Improving Image Quality of SPECT by Deep Image Prior

Akito Yabe^{*1}, Yuta Tsushima^{*1}, Kenta Fukuhara^{*1}, Keisuke Matsubara^{*2}, Kazuhiro Koshino^{*3}, Hiroshi Watabe^{*4}, Tsutomu Zeniya^{*1}

*1 Hirosaki University

*2 Akita Cerebrospinal and Cardiovascular Center

*3 Hokkaido Information University

*4 Tohoku University

Single photon emission computed tomography (SPECT) can visualize the functional information of the body. We are developing SPECT system with high spatial resolution of approximately 3.5 mm, like positron emission tomography (PET). However, image acquired by our SPECT system is noisy due to insufficient sensitivity. We aimed to improve the image quality of SPECT by using deep image prior (DIP), which does not require prior learning and can recover images using only the target image. To demonstrate usefulness of DIP for our SPECT system, simulation using a numerical phantom and measurement with physical phantom using a high-resolution SPECT system were performed. Projection image data processed by DIP were reconstructed using ordered subset expectation maximization (OS-EM) method. The reconstructed SPECT images with and without DIP were compared. In the simulation, DIP processing improved peak signal-to-noise ratio (PSNR) from 32.26 dB to 39.16 dB and structural similarity (SSIM) from 0.860 to 0.928. Visual evaluation also demonstrated that the noise in the reconstructed image was reduced and the structure of the phantom became clearer. In an experiment using a physical phantom, DIP also significantly suppressed the noise and clearly depicted the structure of the brain phantom. These results suggest that DIP is effective in reducing noise and improving image quality for high spatial resolution SPECT.

Key words: SPECT, Image reconstruction, Deep image prior

近接マルチスプリッティングを用いた

Row-Action 型 PET 画像再構成法の体系的な構成

定方 数弥*1 金 喜正*1 工藤 博幸*1

要旨

現在,PETの画像再構成には Maximum Likelihood Expectation Maximization (MLEM)法や,それを高速化した Ordered Subsets EM (OSEM)法が広く用いられている.前者は収束が遅いという欠点が、後者はコスト関数の最小化に収束しないという欠点がある事が知られている.近年,近接スプリッティングという数学的枠組みを用いた手法が盛んに研究されているが、用いられてきた近接スプリッティングの多くはコスト関数を2項に分割する事に基づいてきた.このような従来の手法では高速に収束する反復法を得る事が困難である.本論文ではこの欠点を克服する為に、3種類の異なる近接マルチスプリッティングの枠組みを用いて高速なアルゴリズムを構築する.シミュレーション実験の結果、提案法はいずれも従来法の性能を上回っており、特に提案法2、3はステップサイズコントロールが不要であるという大きな利点が存在する.

キーワード: PET, 画像再構成, 近接スプリッティング

1. はじめに

現在, PET や SPECT の画像再構成には統計的手法が一般的となっており[1], X 線 CT においても,低線量データから画質を向上させるために統計的手法が用いられており,解 析的手法が統計的手法に完全に置き換えられ る可能性も指摘されている.統計的手法は各 モダリティの物理モデルから構築されたコス ト関数を,反復法を用いて最小化する事で画 像再構成を行う.PET や SPECT で使われて いる反復法の多くは 1982 年に提案された MLEM 法[2]に基づいている.MLEM 法の欠 点は収束が遅い事であり,この欠点を克服す る為にこれまでに様々な高速化手法が研究さ れてきた.その中の一つにブロック反復法と いうものがある.これは,投影データをいく

*1 筑波大学システム情報工学研究群 情報 理工学位プログラム

〔〒305-8573 つくば市天王台 1-1-1〕 e-mail: kudo@cs.tsukuba.ac.jp

つかのサブセットに分割し, 適切なアクセス オーダーに従ってサブセットのみを使用して 画像更新を行う手法である.この方法は1994 年に初めて提案され, OSEM 法[3]と呼ばれて いる.しかしながら,OSEM 法には,投影デ ータにノイズが乗っている場合にコスト関数 の最小化に収束しないという欠点が存在する. 近年では近接スプリッティングという数学的 な枠組みを用いた画像再構成法が様々な理由 から研究されており、代表的なものに FISTA[4]や ADMM[5], CP algorithm[6]等が挙 げられる.しかしながら、これらの手法はコ スト関数を2項に分割する近接スプリッティ ングを使用しており、それでは OSEM 法のよ うな高速な反復法を導出する事が困難である. そこで、本研究では画像再構成の分野で使わ れた実績の少ない近接マルチスプリッティン グという新たな枠組みを用いてアルゴリズム の構築を行う.

2. 提案手法

初めに、近接作用素の定義を式(1)に示す.

$$\operatorname{prox}_{\gamma f}(\vec{x}) = \operatorname{argmin}_{\vec{z} \in \mathbb{R}^{J}}(f(\vec{z}) + \frac{1}{2\gamma} \| \vec{z} - \vec{x} \|^{2}), \quad (1)$$

ここで, $\vec{x}, \vec{x} \in R^J$, $\gamma > 0$, $f(\vec{x})$ は微分可能と は限らない凸関数を表す.式(1)より,近接作 用素は写像と最小化を混ぜ合わせた演算子で あることが確認できる.次に近接スプリッテ ィングについて説明する.式(2)に示すような 凸最小化問題を考える.

$$\min_{\vec{x}\in R^J} f(\vec{x}) \equiv f_1(\vec{x}) + \dots + f_n(\vec{x}), \tag{2}$$

ここで, $f_1(\vec{x})$, ..., $f_n(\vec{x})$ は凸関数であるとする. 近接スプリッティングは基のコスト関数 $f(\vec{x})$ の近接作用素を計算するのは困難であるが, 分割されたサブコスト関数 $f_1(\vec{x})$, ..., $f_n(\vec{x})$ の近 接作用素を容易に計算できる場合に効果的な 手法である.本研究では Passty[7], Boyle and Dykstra[8], Han[9]によって提案された 3 つの 近接マルチスプリッティングのフレームワー クを PET 画像再構成に適用する事で,高速か つ安定的なアルゴリズムの構築を目指す.

3. シミュレーション実験

提案手法の性能を評価するシミュレーショ ン実験を Vardi によって提案された数値ファ ントム[10]を用いて行った. 画像のサイズは 256×256[pixel]で投影データの方向数は 256 である. 図1に OSEM 法と提案法の再構成結 果を示す. Method1-3 は順に Passty, Boyle and Dykstra, Han のフレームワークを利用したも のである. 図1より, OSEM 法では構造物が 消失しているが,提案法はいずれも全ての構 造物を再構成する事ができている事が確認で きる.

4. まとめ

本研究では、これまで画像再構成の分野で 利用された実績の少ない近接マルチスプリッ ティングという枠組みを用いて PET 画像再 構成のアルゴリズムの構築を行った.シミュ レーション実験の結果,提案法は従来法と比 較して良好な画質に収束する事が確認できた. また,提案法1はリミットサイクルを回避す る為にステップサイズコントロールを行った が,提案法 2,3 はステップサイズコントロー ルなしでリミットサイクルを回避する事に成 功した. 今後は 3D-PET への拡張等を予定し ている.

謝辞

本研究は JST-CREST (JPMJCR1765) の助 成を受けた.

利益相反の有無

なし

文 献

- [1] 日本医用画像工学会:医用画像工学ハンドブック.東京,2012
- [2] L. A. Shepp and Y. Vardi: Maximum likelihood reconstruction for emission tomography. IEEE Trans Med Imaging 1: 113-122, 1982
- [3] H. M. Hudson and R. S. Larkin: Accelerated image reconstruction using ordered subsets of projection data. IEEE Trans Med Imaging 13: 601-609, 1994
- [4] A. Beck and M. Teboulle: A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems.
 SIAM J. IMAGING SCIENCES. 2(1): 183-202, 2009
- [5] S. Boyd, N. Parikh, E. Chu et al: Distributed optimization and statistical learning via the alternating direction method of multipliers. FTML. 3(1): 1-122, 2010
- [6] A. Chambolle and T. Pock: A first-order primal-dual algorithm for convex problems with applications to imaging. J. Mathmetical Imaging and Vision. 40: 120-145, 2011
- [7] G.B.Passty: Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. J. Math. Anal. Appl. 72: 383-390, 1979

- [8] J. P. Boyle and R. L. Dykstra: A method for finding projections onto the intersection of convex sets in Hilbert spaces. Lecture Notes in Statistics 37: 28-47, 1986
- [9] S. P. Han: A Decomposition Method and Its Application to Convex Programming.

Math. Oper. Res. 14: 237-248, 1989

 Y. Vardi. L. A. Shepp and L. Kaufman: A Statistical Model for Positron Emission Tomography, Journal of the American Statistical Association. Mar. 80(389): 8-20, 1985



図1OSEM 法と提案法の比較

Unified Approach to Fast Convergent Row-Action-Type Iterative

Methods for PET Image Reconstruction Using Multi Proximal Splitting

Kazuya SADAKATA*1, Heejeong KIM*1, Hiroyuki KUDO*1

*1 University of Tsukuba

Currently, Maximum Likelihood Expectation Maximization (MLEM) method and its accelerated version called Ordered-Subsets EM (OSEM) method have been used for image reconstruction in PET. It is known that the former has a drawback of slow convergence and the latter does not converge to a minimizer of cost function. Recently, image reconstruction methods using a new mathematical framework called proximal splitting have been actively studied. So far, most of the proximal splitting frameworks used for image reconstruction are based on splitting the cost function into two terms. With these conventional frameworks, it is impossible to obtain iterative methods that converge fast such as OSEM method and row-action-type iterative methods. To overcome this drawback, in this paper, we propose a unified approach to construct row-action-type iterative methods using three different types of multi proximal splitting frameworks. Results of simulation studies show that all the iterative methods obtained from the proposed approach can reduce the effect of statistical noise well and converge to a minimizer of the cost function with a high speed comparable to that of OSEM method. In addition, proposed method 2 and 3 have an advantage in that they do not require step size control.

Key words: PET, Image Reconstruction, Proximal Splitting 著者紹介

和文でも英文でも可.





Kazuya Sadakata

received the B.E. degree from the Division of Information and Electronic System Engineering, NIT Sendai college, Japan, in 2020. He is a student in the Degree Programs in Systems and Information Engineering Graduate school of Science and Technology, University of Tsukuba, Japan. His major research field is PET image reconstruction.

Heejeong Kim received the B.E. degree from the Department of Advanced Robotics, Chiba Institute of Technology, Japan, in 2020. She is a student in the Degree Programs in Systems and Information Engineering Graduate school of Science and Technology, University of Tsukuba, Japan. Her major research field is CT image reconstruction.



Hiroyuki Kudo received the B.Sc. degree from the Department of Electrical Communications. Tohoku University, Japan, in 1985, and the Ph.D. degree from the Graduate School of Engineering, Tohoku University, in 1990. In 1992, he joined the University of Tsukuba, Japan. He is currently a Professor with the Division of Information Engineering, Faculty of Engineering, Information and Systems, University of Tsukuba, Japan. His research areas include medical imaging, image processing, and inverse problems. In particular, he is actively working on tomographic image reconstruction for X-ray CT, PET, SPECT, and electron tomography. He received best paper awards more than 10 times from various international and Japanese societies. He received the IEICE (The Institute of Electronics. Information, and Communication Engineers, Japan) Fellow award for his contributions on "cross-sectional image reconstruction methods in medical computed tomography". In 2018, he obtained Commendation for Science and Technology by the Minister of Education, Culture, Sports, Science and Technology for his contributions on "research on design method and image reconstruction method for new CT". For 2011-2016, he was an Editor-in-Chief of the Journal of Medical Imaging Technology (MIT). From 2020, he is a president of Japanese Society of Medical Imaging Technology (JAMIT).

Higher SNR PET Image Prediction Using a Deep Learning Model

with Efficient Channel Attention

Tuo YIN^{*1}, Takashi OBI^{*2}

Abstract

In this study, a deep residual efficient channel attention (RCA) network is proposed for positron emission tomography (PET) image denoising. Since adjoining slices share similar brain structure while having different random noise, we take the concatenation of adjoining three slices as an input volume to learn distinguishing noise from consistent structure. Deep RCA network consists of a cascade of five residual groups. Each residual group contains three repetitions of RCA block. RCA module adaptively rescales channel-wise features by learning channel statistics. Short skip connection and long skip connection are employed to bypass abundant low-level features and reach a very deep network. 20 brain models are utilized to simulate PET emission data in the real case. The noisy activity maps reconstructed using maximum likelihood attenuation correction factors (MLACF) algorithm with 50 iterations are used as training data. Gaussian filter, nonlocal mean (NLM) algorithm, and deep RCA network with a single slice as input are selected as reference methods. Qualitative results show that proposed method with adjoining three slices as input is capable of preserving fine details of tissues and restoring high-quality PET images. Quantitative analysis further confirms the superior performance of this approach over the reference methods with structural similarity index measure (SSIM) of 0.89, peak signal-to-noise ratio (PSNR) of 21.64 dB, and normalized root mean square error (NRMSE) of 3.32%.

Keywords : positron emission tomography, deep learning, image analysis, efficient channel attention

1. Introduction

Positron emission tomography (PET) is an essential medical imaging modality which has been widely used in neurology, oncology, and cardiology. PET images suffer from poor signal-to-noise ratio (SNR) due to physical degradation factors, low counts of detected photons, and ill-poseness of the reconstruction algorithms [1]. In order to address this problem, many approaches were proposed to improve PET image quality. Maximum likelihood expectation maximization (MLEM) algorithm is the baseline for statistical reconstruction. Taking advantage of time-of-flight (TOF) information, maximum likelihood activity and attenuation reconstruction (MLAA) algorithm and maximum likelihood attenuation correction factors (MLACF) algorithm were proposed for joint estimation of activity distribution images and attenuation correction coefficients. The post-processing techniques aim at denoising PET images after reconstruction. These include Gaussian filtering, non-local mean (NLM) algorithm [2], and the block-matching 3-D method [3]. Some recent works applied deep learning to PET image denoising and demonstrated superior performance [1], [4]. In this

^{*1} Department of Information and Communications Engineering, Graduate School of Engineering,

Tokyo Institute of Technology [Yokohama 226-8503, Japan] e-mail: yintuo.scarlett@gmail.com

^{*2} Institute of Innovative Research, Tokyo Institute of Technology [Yokohama 226-8503, Japan]

work, we investigate the use of efficient channel attention (ECA) mechanism to a convolutional neural network (CNN) for PET image denoising.

2. Materials and Methods

We used 20 brain models from Brainweb database [5] to simulate TOF PET emission data. We assumed there was a lutetium-oxyorthosilicate (LSO) PET scanner with time resolution of 520 ps. Activity maps and attenuation maps were created by assigning activity and attenuation coefficient to each tissue class [6]. Activity maps were recognized as the ground truth. TOF sinograms were generated by forward projection. TOF sinogram was organized in five sinograms and each sinogram was for one TOF bin. We added a uniform background that accounts for 20% of the noise-free data to simulate scatter and random events. Poisson noise was then added to noise-free sinograms. The resulting noisy TOF sinograms were assumed to be the PET emission data obtained in the real case. For each brain model, 490 noisy TOF sinograms were created. We reconstructed PET images using MLACF algorithm with 50 iterations. 490 noisy PET images for each model were in the dataset. We used 15, 1, and 4 brain models for training, validation, and testing, respectively.





We propose a deep RCA network to improve PET image quality. The network architecture is shown in Fig. 1. In order to distinguish random noise from the consistent brain structure, the concatenation of three adjoining slices is utilized as the input of the network. To predict the kth slice in a brain model, the k-1th, kth and k+1th slices are concatenated as an input volume. Deep RCA network starts with a 2D Convolution (Conv) layer following five residual groups. Each residual group consists of three RCA blocks and a 2D Conv layer. Short skip residual connections add the output of the previous residual group and the current group together. A long skip residual connection adds the input and output of the last residual group together. The efficient channel attention (ECA) mechanism is illustrated in the RCA block. Different from channel attention, ECA extracts the direct correspondence among channels using a 1D Conv layer. Let the input of ECA be $X \in \mathbb{R}^{H \times W \times C}$, where $H \times W$ is the spatial dimension of feature maps and *C* is channel dimension. A global average pooling (GAP) layer extracts the channel statistics $g(X) \in \mathbb{R}^C$ by averaging each feature map. The final channel statistics after a 1D Conv layer and a sigmoid gating is

$$s = \sigma(Wg(X)), \tag{1}$$

where σ denotes sigmoid function and W is the weight set of 1D Conv layer. The output of RCA is obtained by multiplying the input X and the final statistics s using an element-wise product operation. The last layer of the

network is designed as a 2D Conv layer with 1 kernel to produce an image.

We used structural similarity index measure (SSIM), peak SNR (PSNR), and normalized root mean square error (NRMSE) to evaluate the image quality. The mean and standard deviation of three metrics among all the slices in the test dataset were calculated. Gaussian filtering and NLM algorithm were selected as the reference methods for comparison. To illustrate the effectiveness of using three adjoining slices as inputs, one single slice and three slices were used as inputs of the proposed method, respectively. The networks were trained for 200 epochs. The mean square error (MSE) between the prediction and the label was selected as the cost function. The adaptive moment estimation method (Adam) was used to minimize the loss with a learning rate of 0.001. The networks were implemented using the deep learning platform Tensorflow 2.0.

3. Results

As indicated in Fig. 2, the proposed method with adjoining three slices as inputs better preserves the details of tissues. Gaussian filtering and NLM produce over-smooth images, which suffer from resolution losses and blurred details. Deep RCA with one single slice as input is capable of predicting images with fine detail. However, some blurred area can be observed. The quantitative results are shown in Table 1, where the same trend as qualitative results can be seen. Our approach with three slices as inputs achieves the superior performance with the highest SSIM of 89% and PSNR of 21.64 dB, and the lowest NRMSE of 3.32%. Compared with Gaussian and NLM methods, both qualitative and quantitative results have an obvious improvement using deep RCA network.



Fig. 2 Three representative cases of qualitative results using different methods

Method	SSIM	PSNR (dB)	NRMSE (%)			
MLACF	0.65 ± 0.11	15.23±1.47	8.53±9.28			
Gaussian	0.67 <u>±</u> 0.07	16.78±2.91	6.19 <u>+</u> 4.88			
NLM	0.70 ± 0.08	17.04 <u>+</u> 2.34	6.30±5.79			
Proposed-single	0.72 ± 0.07	19.27±3.93	3.97±3.19			
Proposed-three	0.89 <u>+</u> 0.03	21.64±5.32	3.32±2.15			

Table 1 SSIM, PSNR, and NRMSE of all the slices in the test of	lataset
---	---------

4. Discussions and Conclusions

The analysis of the above experimental results found that deep RCA network can improve PET image quality by postprocessing the reconstructed images. The superior performance of deep RCA with three slices as inputs over using one single slice, Gaussian, and NLM illustrates the effectiveness of this approach. A limitation of this method is that there are 2D slices in the dataset. The emission process was assumed to be 2D. In future work, we plan to explore 3D models from real patients and expand the network into 3D.

In conclusion, a deep residual efficient channel attention network was proposed to denoise PET images after reconstruction. The data analysis shows that compared with the Guassian and NLM post-processing methods, this approach provides high-quality PET images from activity maps reconstructed using MLACF algorithm. Additionally, the effectiveness of using adjoining three slices as inputs is demonstrated.

Competing interests

The authors declare no competing interests.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Number 18H02761. The authors would like to thank Dr. Hideaki Tashima for helpful advice and appropriate instructions.

References

- [1] Liu C C and Qi J: Higher SNR PET image prediction using a deep learning model and MRI image. Phys. Med. Biol. 64(11): 115004, 2019
- [2] Arabi H and Zaidi H: Non-local mean denoising using multiple PET reconstructions. Ann. Nucl. Med. 35(2): 176–186, 2021
- [3] Feruglio P F, Vinegoni C, Gros J, et al: Block matching 3D random noise filtering for absorption optical projection tomography.
 Phys. Med. Biol. 55(18): 5401–5415, 2010
- [4] Gong K, Guan J, Liu C C, et al: PET image denoising using a deep neural network through fine tuning. IEEE Trans. Radiat. Plasma Med. Sci. 3(2): 153–161, 2018
- [5] Aubert-Broche B, Griffin M, Pike G B, et al: Twenty new digital brain phantoms for creation of validation image data bases. IEEE Trans. Med. Imaging 25(11): 1410–1416, 2006
- [6] Yang B, Ying L, and Tang J: Artificial neural network enhanced Bayesian PET image reconstruction. IEEE Trans. Med. Imaging 37(6): 1297–1309, 2018

Generation of delayed PET/CT images for pancreatic cancer using CNN-LSTM model

Ritu Bhusal Chhatkuli^{*1}, Yasuhiko Tachibana¹, Takayuki Obata¹, Ryuichi Nishii¹ and Tatsuya Higashi¹

Abstract

Initial and delayed scans (dual-time-point scan: DualS) are widely used to improve the diagnostic efficacy of PET/CT for accurate delineation and staging. The acquisition process however is time consuming. To eliminate the need of these delayed scans, we propose a Convolutional Neural Network (CNN) - Long Short-Term Memory (LSTM) based approach for generating delayed phase image series images from the 5 min dynamic scan (DynS) images acquired right after the early phase scan. The network was trained to generate delay phase images from the DynS images to predict the delayed phase image. Qualitative and quantitative analysis of the results obtained in this primary study shows significant accuracy with relative image error not more than 2.44% and the maximum image intensity error based on maximum standard uptake value (SUVmax) up to 1.8% when comparing the actual delayed scan and image generated by our model.

1. Introduction

Positron Emission Tomography/Computed Tomography (PET/CT) is one of the popular modalities for diagnosis. Dual time point scans (DualS) are widely used to improve the diagnostic efficacy of PET/CT for accurate delineation and staging [1]. However, the acquisition process is time consuming, causing extra burden both on patients and the clinicians.

In this pilot study, we report a deep learning-based regression analysis to show the potential of DynS image to generate the virtual delayed image. We take a CNN-LSTM approach for possible elimination of the need of the delayed scan from the conventional DualS acquisitions for malignant pancreatic tumors by generating the delayed scans using Initial PET/CT (Early) images and dynamic list-mode PET images.

2. Material and Method

13 patients underwent the whole-body PET/CT scans by PET scanner (Biograph 16 SIEMENS) after the single intravenous injection of PET tracer 18F-FDG according to the standard clinical protocol in our institute. Initial image acquisition for the whole-body scans started 60 mins after the intravenous administration. Additional list mode PET images series (DynS) were acquired at the rate of 10 sec per image for additional 5 minutes. Acquisition time was 3 mins per bed positions. For each patient 2460 list mode PET images were acquired corresponding to 82 CT slices over 5mins. Approximately 120 mins after the intravenous injection delayed scans are performed for the volume of interest. With the obtained images, we prepared a dataset consisting of training, validation, and testing with the ratio of 10-2-1 for the available number of patients.

Department of Molecular Imaging and Theranostics, National Institute for Quantum and Radiological Science and Technology [Chiba, Japan]

^{*} e-mail: bhusal.ritu@qst.go.jp

The proposed network model is shown in Fig.1. It consists of time distributed LSTM -convolutional layers for dynamic input and multiple convolution layers for Early PET images. Maxpooling layers, upsampling layers, dropout, batch normalization is also added with rectified linear unit (ReLU) and sigmoid as activation functions.



Fig. 1 Proposed network for the generation of delayed images.

Training was performed using Keras (version 2.1.1) with TensorFlow backend (GPU -version 1.4.0) platform.

3. Results and Discussion

Fig.2 shows the target and the predicted generated images for same slice for the same patient. Visually, the generated slices indicate a good match with the respective target images.



Fig. 2 The target image (left) and the Generated image (right) for two random slices of a patient.

Relative image error (I_{RE}), the difference between the target and the predicted image is calculated using equation [1] in two separate regions of interest (ROI).

$$I_{RE} = \frac{\left(I_{Target} - I_{Gen}\right)}{I_{Target}} \times 100\% \qquad [1]$$

The smaller ROI covers inside the tumor pixels whereas

the larger ROI covers the tumor and the surrounding area. Fig.3 shows the relative image error with average 2.44% in small ROI and 1.45% in Large ROI.



Fig. 3 Relative Image Error (IRE) for small ROI and Large ROI Difference in maximum image intensity is also calculated by pixel based SUVmax analysis by converting the pixel intensity value from ROI to the activity concentration in Becquerel, the error ranges from 0.21-1.8%.

4. Conclusion

The proposed CNN-LSTM model generated delayed images with significant error both qualitatively and quantitively. In future, we aim to add more patient's data in this analysis and further validations will be performed for its clinical use.

Competing Interest

The authors have no conflict of interest to declare in this research.

Acknowledgement

This work is supported by JSPS grant-in-aid support for scientific research KAKENHI grant number 19K20685.

Refrences

[1] Nakajo, Masatoyo, et al. "The efficacy of whole-body FDG-PET or PET/CT for autoimmune pancreatitis and associated extra pancreatic autoimmune lesions.

" European journal of nuclear medicine and molecular imaging 34.12 (2007)

腹部超音波スクリーニング支援のための

深層学習による撮影断面推定に関する初期検討

目加田慶人*1 道満恵介*1 小川眞広*2 西田直生志*3 工藤正俊*3

要旨

本稿では、腹部超音波スクリーニングにおいて撮影された画像に対して、それが25 断面撮影法のどの断面 に対応しているかを推定する深層学習手法の初期的検討について述べる.25 断面撮影法は腹部スクリーニ ングにおける診断の網羅性を保証するものであり、画像から撮影された断面が推定できることで腫瘍等が 腹部のどの位置に存在しているのかを把握できる.25 の断面画像には比較的類似した見えの画像も含まれ ているため、画像特徴の類似した断面をまとめたクラスとして扱う分類器に加えて、同一クラスと判定さ れた画像をいずれかの断面に分類する分類器を利用する2段階の分類アルゴリズムを開発した.25 断面を 記録した267 例の検査データセットを対象とした評価実験の結果、25 クラス分類を単純に適用した場合の 正解率が0.790 であったのに対して、2 段階の分類をする本手法により正解率が0.836 に向上したことを確 認した.

キーワード:腹部超音波検査,深層学習

1. はじめに

腹部超音波スクリーニングにおいて,対象 臓器を網羅的に確認するための断面撮影法が 提案されている.日大方式の腹部超音波スク リーニング 25 断面撮影法[1]は,超音波検査 の客観性を向上と,円滑な二重読影を目指し, 腹部の複数の臓器を対象とした撮影断面およ びその撮影順番を定められたものである.こ の撮影法に準拠すれば,どの断面が臓器のど の部分を撮像したものかを容易に理解できる. 本稿では,腫瘍など関心部位の解剖学的な部

*1 中京大学 工学部
〔〒470-0393 豊田市貝津町床立 101〕
e-mail:y-mekada@sist.chukyo-u.ac.jp
e-mail: kdoman@sist.chukyo-u.ac.jp
*2 日本大学医学部
e-mail: echo.m.ogawa0922@gmail.com
*3 近畿大学医学部
e-mail: naoshi@med.kindai.ac.jp
e-mail: m-kudo@med.kindai.ac.jp

位の把握や,撮影された断面画像の適切性の 判断を目的として,この25断面の分類を試み る.25種類の断面画像は,見えが明確に異な るものばかりではなく,類似したものも存在 するため,これらを2段階に分けて分類する 方法を検討する.

2. 25 断面分類手法

本稿で扱う撮影法は, 肝臓, 胆嚢, 膵臓, 脾臓, 腎臓, 腹部大動脈のスクリーニング検 査を目的としており, 左肋間走査に始まり, 右肋間走査に終わる一連の撮影方法を定めた ものである.実験に用いたデータセットは, 日本大学病院より提供された 267 名分の画像 データセットである.主に肝臓を対象とした 断面の例(断面 17 から 20)を図1に示す. ゲイン調整が不適切な画像を除外したところ, 各断面で少なくとも 257 画像は存在すること が確認された.予備実験として, WideResNet[2]を利用して25 断面を分類した ところ, 断面の見えの類似性から断面 2 と 3,



図1 断面画像の例(断面17から20)

断面 15 と 23 がそれぞれに誤分類する傾向で あったため,以下の分類手法を構築した.

2.1. 提案手法

提案手法は、次のような2段階の階層的分 類とした.第1段階では、断面2と3、断面 15と23をそれぞれ1つのクラスにした23ク ラス分類器を作成する.断面15と23につい ては、当該クラスに分類された画像を細分す ることで最終的な分類結果を得るものとする. なお、断面2と3については、画像上での見 えの差はほとんどないことから分類すること を断念した^[注 1].

2.2. 実験

学習,検証,テストのためのデータは,被 験者をランダムに 8:1:1 の割合で分割した. 各画像は,画像中心を切り出して使用し,学 習時にはランダム切り出しのデータオーギュ メンテーションを適用した.前段の 23 クラス 分類器のネットワークモデルとその最適化手 法はそれぞれ WideResNet, SAM[3]とし,後 段の 2 クラス分類については VGG16, SDG とした.なお,前段の分類器のみを本手法と 同条件で適用した場合を比較手法とした.

3. 結果と考察

混同行列を図2に示す.前段の分類器で断面15と23に分類された53枚の画像に対して, 2段階目で41枚を正しく分類することができた.全体の平均分類正解率は,本手法が0.836, 比較手法が0.790であった.このことから,2 段階で分類する本手法により,分類正解率を

注1:実際の現場においてもプローブを反転して撮影するこ とで画像が判別できるようにしている



約 4%向上させることができ、本手法の有効 性を確認できた.

4. まとめ

本稿では,超音波スクリーニングの客観性 を向上させるために提案されている腹部超音 波スクリーニング25断面撮影法を対象に,撮 影された画像の断面分類に取り組んだ.見え の類似性から分類が困難な断面をまとめたク ラス分類の後に,まとめられたクラス内で断 面を細分する2段階の分類手法を検討した. 実験の結果,本手法の平均分類正解率は0.836 であった. 今後は断面2と3の分類,分類結 果を用いた対象臓器の解剖学的な情報の抽出 手法の検討を行う予定である.

謝辞

本研究の一部は,日本学術振興会科研費補 助金の援助による.

利益相反の有無

なし

文献

- [1] 小川眞広: あなたも名医!腹部エコーの "みかた",日本医事新報社,2019
- [2] S. Zagoruyko and N. Komodakis: Wide Residual Networks, Proc. of the British Machine Vision Conf. (BMVC), pp.87.1-87.12. BMVA Press, Sept. 2016.
- [3] P. Foret, A. Kleiner, H. Mobahi, B. Neyshabur: Sharpness-Aware Minimization for Efficiently Improving Generalization, https://arxiv.org/abs/2010.01412

A Preliminary Study on Estimating Cross-section Category

by DCNN for Abdominal Ultrasound Screening Images

Yoshito MEKADA*1, Keisuke DOMAN*1, Masahiro OGAWA*2, Naoshi NISHIDA*3, Masatoshi KUDO*3

*1 Chukyo University

*2 Nihon University Hospital

*3 Kindai University Hospital

This manuscript presents a preliminary study of a DCNN for estimating the cross-sectional classification of a 25-sectional imaging method for abdominal ultrasound screening. The 25-sectional imaging method guarantees the comprehensiveness of the diagnosis in abdominal screening. Considering that some of the 25 cross-sectional images are relatively similar in appearance, we treat such cross-sections as one group class. We developed a two-step classification algorithm that uses a classifier that treats cross-sections with similar image features as a group of classes, followed by a classifier that classifies images determined to be in the grouped class into their respective cross-sections. Two hundred sixty-seven inspection datasets with 25 cross-sections were examined, and the accuracy was improved to 0.836 by using our two-step algorithm, compared to 0.790 by simply applying the 25-class classification. **Key words:** Abdominal ultrasound screening, deep learning

深層学習を用いた超音波画像における腎嚢胞の計測補助機 能の開発

金内 友里恵^{*1} 橋本 正弘^{*2} 東田 直樹^{*2} Hasnine Haque^{*2,3}

陣崎 雅弘^{*2} 榊原 康文^{*1}

要旨

超音波検査において, 臓器や腫瘍の計測のために2点1組の計測マーカーが付与される. 計測マーカー に基づいて対象所見の位置やサイズが計測される.本研究は腹部超音波検査における計測マーカーの付与 をサポートすることを目的として, 深層学習を用いた超音波画像における計測補助機能の開発を行なった. 対象は,超音波検査の中でも計測頻度の高い腎嚢胞とした.慶應義塾大学病院で2019年1月から2020年 5月に撮影された腹部超音波画像170,538枚のうち,腎嚢胞が計測されている2,664枚を使用して,画像を 入力すると計測マーカーの位置を表す distance map を予測するようにモデルを訓練した.モデルはUNet++ を用いた.出力された distance map から決定した計測マーカーの位置について,座標誤差を用いて予測精 度を評価したところ,平均8.13 mm,中央値4.63 mmという結果になった.

キーワード:超音波,深層学習,計測マーカー

1. はじめに

超音波検査において, 臓器や腫瘍の計測の ために2点1組の計測マーカーが付与される. 計測マーカーに基づいて対象所見の位置やサ イズが計測される.計測マーカーは診断にお いて重要な役割を担っている一方で,マーカ ーを置く作業は手動で行われているため,医 師の負担になっている.また,計測マーカー の置き方には決められたルールがないため経 験差や個人差が出てしまうという問題もある. そこで,機械学習を用いて計測マーカーを付 与する作業を自動化することができれば,医 師の負担軽減や検査時間の短縮,および個人 差をなくすことにつながると考えられる.そ こで,本研究は腹部超音波検査における計測 マーカーの付与をサポートすることを目的と して,深層学習を用いた超音波画像における 計測補助機能の開発を行なった.対象は,超 音波検査の中でも計測頻度の高い腎嚢胞とした.

2. 方法

本研究は倫理委員会の承認のもと実施された(承認番号 20170018).

1 組の計測マーカーが写ったスクリーンキ ャプチャ画像から,計測マーカーからの距離 に応じて画素値を変化させることで計測マー カーの位置を表現するヒートマップを作成し, 損失関数を定義する正解画像とした.前処理 によって計測マーカーを消したスクリーンキ ャプチャ画像を入力画像とした.入力画像を 学習モデルに入力し,正解画像に近いヒート マップを出力するように学習させた.学習モ デルが出力したヒートマップから計測マーカ ーの予測座標を決定し,正解の計測マーカー

^{*1} 慶應義塾大学理工学部

^{〔〒223-8522} 神奈川県横浜市港北区日 吉 3-14-1〕

の座標との誤差を求めることで精度評価を行った.

1) データセット

慶應義塾大学病院で 2019 年 1 月から 2020 年 5 月に撮影された腹部超音波画像 170,538 枚のうち, 腎嚢胞が計測されている 2,664 枚 を使用した.

2) 学習モデル

二乗誤差とした.

UNet++[1]のモデルを使用した.UNet++は, 画像セグメテーションというタスクのための モデルで,画像内で個別物体の領域を認識す る.入力画像を畳みこむことで特徴抽出を行 う Encoder と,抽出された特徴マップを逆畳 み込みしてセグメテーション結果を出力する Decoder で構成されている.図1のように, Encoder の各階層で出力される特徴マップを decode してから skip connection で Decoder に 連結している点が UNet++の特徴である.こ れにより Encoder/Decoder の表現差異が減り, 最適化問題を簡単にすることができると言わ れている.損失関数はヒートマップ上の平均



図 1 UNet++のアーキテクチャ[1]

3) データの前処理

スクリーンキャプチャ画像には撮影日や撮 影条件のパラメータなどが書かれた部分も含 まれているので,超音波画像のみを含むよう にトリミングした.また,計測マーカーは座 標情報ではなく画像の一部として含まれてい るため,テンプレートマッチングによって計 測マーカーを同定した.その後,周囲の画素 の情報を用いて補間することで,超音波画像 から計測マーカーを削除した.

4) 計測マーカー位置画像

計測マーカー位置を表す画像を作成した. 計測マーカー位置を中心とするガウス分布を 用いて、計測マーカーから各ピクセルまでの 距離に応じて画素値が変化するヒートマップ を作成した.計測マーカー位置の画素値を 0 とし、この位置から離れたピクセルほど画素 値が大きくなるようにした. 最大値は 255 と した. ガウス分布の半径は 50 pixel とした. 1 組の計測マーカーの2点それぞれについてヒ ートマップを作成し, 左のマーカーのヒート マップを RGB 画像の B チャネル,右のマー カーのヒートマップを G チャネルに入れた. R チャネルは全て, 画素値0とした. 図2の 左の超音波画像上に黄色い十字で示した計 測マーカーの位置を表す画像は図 2の右の ようになる、緑の部分が左のマーカー、青 い部分が右のマーカーの位置を表している.



図 2 超音波画像(左)とその計測マーカー位置 画像(右)

5) 精度評価用の入力画像

入力画像は以下の2通りの形式を用いた.(1) 超音波画像のみ

前処理によってトリミングおよび計測マー カーの削除を行った画像を入力に用いた.

(2) 超音波画像と腎嚢胞のヒートマップ 前処理をした画像に、腎嚢胞の位置を表現 するヒートマップを追加した. 腎嚢胞の位置 は次のようにして決定した. 計測マーカーの 座標を $p_1 = (x_1, y_1), p_2 = (x_2, y_2)$ とし,式(1) で腎嚢胞の半径rを求めた.

$$r = \frac{\|p_1 - p_2\|}{2} \tag{1}$$

腎嚢胞の座標p_{cvst}を、-r以上、r以下の乱数

x_{random}, *y_{random}*を用いて,式(2)によって決 定した.

$$p_{cyst} = \left(\frac{x_1 + x_2}{2} + x_{random}, \frac{y_1 + y_2}{2} + y_{random}\right) (2)$$

実際の超音波検査で使用する場合は腎嚢胞の 中心位置を医師が指定することになるため, 手作業によるぶれを想定し,乱数を用いて, 計測マーカー2 点の中点から腎嚢胞の半径の 範囲内で座標をずらした.計測マーカー位置 画像と同様にして,この座標を中心としたヒ ートマップを作成した.超音波画像はモノク ロであるため,RGBのRチャネル,Gチャネ ルに超音波画像の同じ画素値を入力し,Bチ ャネルをヒートマップに置き換えた.図3 の黄色い部分が腎嚢胞の位置を表している.



図 3 入力画像(2)

6) 座標誤差

モデルが出力したヒートマップにおいて, 画素値が最小となっている点を予測座標とした.この座標と正解の座標のユークリッド距離を座標誤差として評価に用いた.座標誤差の単位はmmとした.

3. 結果

データセットを 7:3 の割合で学習用とテス ト用に分割して評価を行った.テストデータ で評価を行った結果,座標誤差は表 1,図 4, 図 5のようになった.腎嚢胞のヒートマップ を入力に追加することで精度が改善した.

表 1 座標誤差

	座標誤差(mm)			
入力画像	平均	中央値		
(1)	8.13	4.63		
(2)	4.99	3.61		



図 4 入力画像(1)の座標誤差の分布(横軸:座標誤 差(mm),縦軸:データ数)



図 5入力画像(2)の座標誤差の分布

図 6, 図 7, 図 8 は入力画像(1)の予測結果 である.黒い円形で写った腎嚢胞の輪郭上に 黄色の十字で示した正解の計測マーカーが置 かれている.水色の十字はモデルが予測した 計測マーカーの位置である.図 6 は正しく予 測ができた例である.図 7, 図 8 は予測が正 しくない例である.計測マーカーは腎嚢胞の 長径を計測するように配置されるが,図 7 で は短径が予測されている.図 8 では,腎杯付 近に計測マーカー位置が予測されている.



図 6 入力画像(1)の予測結果① (黄色:正解座標,水色:予測座標)



図 7 入力画像(1)の予測結果②



図8入力画像(1)の予測結果③

4. 終わりに

超音波画像を用いて UNet++で構築された モデルの学習を行うことで,腎嚢胞の計測マ ーカー位置を予測するシステムを構築した. 腎嚢胞の位置を表すヒートマップを入力に追 加することで座標誤差が減少することが確認 できた. 今後は,中心座標の特定も含めて自 動化することを検討している.

利益相反の有無

本研究の一部は GE ヘルスケア・ジャパン の研究助成金を得て実施された.

文 献

[1] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. Unet++: A nested u-net architecture for medical image segmentation. In Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, pp. 3–11. Springer, 2018.

Development of a Deep Learning-based Method to Support the

Measurement of Renal Cysts in Ultrasound Images

Yurie KANAUCHI^{*1}, Masahiro HASHIMOTO^{*2}, Naoki TODA^{*2}, Hasnine Haque^{*2,3}, Masahiro JINZAKI^{*2}, Yasubumi SAKAKIBARA^{*1}

- *1 Faculty of Science and Technology, Keio University
- *2 Keio University School of Medicine
- *3 GE Healthcare Japan

In ultrasonography, two sets of measurement markers are given to measure organs and tumors by a radiologist. Based on the measurement markers, the location and size of the target findings are measured. In this study, we developed a measurement support system for ultrasound images using deep learning to support the assignment of measurement markers in abdominal ultrasonography. The target of this study is renal cysts, which are frequently measured in ultrasonography. Among 170,538 abdominal ultrasound images taken at Keio University Hospital from January 2019 to May 2020, we used 2,664 images in which renal cysts were measured, and trained the model to predict a distance map representing the location of the measurement marker. UNet++ was used as the deep learning model. We evaluated the prediction accuracy using the coordinate error of the positions of the measurement markers determined from the output distance map. The mean error was 8.13 mm and the median error was 4.63 mm.

Key words: Ultrasonography, Deep Learning, Measurement Marker

著者紹介



金内 友里恵 (かなうち ゆりえ) 2020 年慶應義塾大学卒.現在,同大学院在 学.研究領域は医用画像の深層学習.



Hasnine Haque

Senior Scientist with the STO-J Research Division, GE Healthcare Japan. He received Ph.D. (Eng.) from Tokyo Institute of Technology in 2000. He is also a research collaborator with department of Radiology, Keio Univ. School of Medicine. He received poster award (2nd place) at ISMRM (2012) in Interventional MR category. His research fields include machine learning with medical imaging in clinical research and image processing



橋本 正弘 (はしもと まさひろ) 2006 年慶應義塾大学医学部卒業, 2006-2008 年 済生会中央病院で初期臨床 研修, 2008-2010 年 慶應義塾大学医学部放 射線科, 2010-2012 年 日本鋼管病院放射線 科, 2012 年- 慶應義塾大学医学部放射線 科. 放射線診断専門医, IVR 専門医, 核医 学専門医、情報処理安全確保支援士.専門 は超音波画像診断, 医療情報学, 深層学習.







東田 直樹 (とうだ なおき) 2015 年慶應義塾大学医学部卒業,2015 年 -2017 年 独立行政法人国立病院機構東京 医療センターで初期臨床研修,2017-2018 年 慶應義塾大学医学部放射線科,2018 年 独立行政法人国立病院機構東京医療セン ター 放射線科,2019 年-慶應義塾大学大 学院医学研究科.放射線科専門医.研究領 域は深層学習.



榊原 康文 (さかきばら やすぶみ) 東京工業大学大学院修了後,富士通研究所, 東京電機大学を経て,慶應義塾大学理工学部 教授.博士(理学).バイオインフォマティ クスを専門とし,ゲノムデータなどのオミッ クスデータに機械学習を適用したがんゲノ ム解析・腸内細菌叢解析,バーチャルスクリ ーニングによる創薬,深層学習を用いた医療 画像解析などの研究に従事.

ピットパターン特徴量の解析に向けた 超拡大内視鏡画像の再構成法に関する初期的検討

伊東 隼人*1,小田 昌宏*2,1,森 悠一*3,4,三澤 将史*4, 工藤 進英*4,森 健策*1,5,6

要旨

超拡大内視鏡画像に含まれる識別的なピットパターンを可視化するための画像再構成方法を検討する.ポ リープ表面の拡大観察で得られる画像パターン,つまりピットパターンを表すハンドクラフト特徴量を用い ることで高精度な病理類型分類が可能であると判明している.さらには特徴量選択処理による尤度推定・識 別の高精度化も報告されている.一方,ピットパターンと特徴量の詳細な対応関係は未だ不明瞭であり,分 類においてピットパターンのいずれの部分が特に重要であるかは未だ明らかになっていない.選択した特 徴量を画像として再構成する可視化技術が求められているものの,前述のハンドクラフト特徴量と画像の間 には恒等写像が存在しないため,特徴量抽出の逆変換は存在しない.したがって特徴量から識別的なピット パターンの可視化は不良設定問題であり,一般的に可視化は困難である.本稿では画像と抽出したピットパ ターン特徴量のペアが予め学習データとして与えられていることを仮定し,識別的な画像パターンの抽出・ 可視化を達成した.

キーワード:超拡大内視鏡,病理類型分類,ホワイトボックス化,特徴選択,特徴の可視化

1. はじめに

内視鏡による大腸ポリープ表面の拡大観察に 基いてポリープの組織学的な類型分類を行う診 断方法が研究されている.[1] この診断方法は 細胞核の分布パターン(ピットパターン)や血 管の形状といったポリープ表面のテクスチャに 基づく診断方法である. 生体内の拡大観察に基 づく診断方法は生検を介さない実時間病理診断 を実現しうるが、高度な知識と十分な経験を要 するために非熟練者には難しい. そこで実時間 病理診断の実現に向けた非熟練医の教育・補助 を目的とした計算機支援診断(CAD)システム の開発が長らく行われている [2-4]. 超拡大大 腸内視鏡 (Olympus, Tokyo) の 520 倍超拡大観察 に対して開発された CAD システムは 5mm 以 下のポリープに対して熟練者と同程度かそれ以 上の病理類型分類精度を達成した [3,4].

超拡大大腸内視鏡に対して開発された CAD システムは Haralick 特徴量 [5] をベースに高階 テンソル分解とマルチスケール化を利用して設 計されたハンドクラフト特徴量, つまりピット パターン特徴量に対してサポートベクターマシ ンにより分類を行う [4]. このピットパターン特 徴量に基づく分類に関しては, 複数施設にて収 集された約 55,000 枚の超拡大内視鏡画像(EC 画像)を利用した評価実験において, 深層学習 手法よりも高い汎化性能が得られることが報告 されている [4]. さらには多様体間の距離に基づ く特徴選択による尤度推定・識別の高精度化も 報告されている [6]. しかし, ピットパターン特 徴量と分類に重要なピットパターンとの関係は 不明瞭である.

本稿では EC 画像に含まれる識別的なピット パターンを可視化することを目的に 2 つのアプ ローチから検討を進める. 図1 が示すように,そ もそも特徴の選択には 2 通りの計算方法が考え らられる. 1 つ目の合成射 $I \rightarrow H \rightarrow H$ では特 徴抽出後に特徴空間にて特徴選択を行う. 2 つ 目の合成射 $I \rightarrow \check{I} \rightarrow \check{H}$ では特徴選択前に画像 空間にて画像の特徴選択を行う. 前者の場合は 図 1(a) が示すように,特徴空間にて選択した特 徴量を画像空間に写す g を求め,画像再構成を 行う必要がある. 一方,後者の場合は図 1(b) が 示すように画像空間で特徴選択を行うため,特 徴空間から画像空間への写像を求める必要はな

^{*&}lt;sup>1</sup>名古屋大学大学院情報学研究科知能システム 専攻 [〒 464-0822名古屋市千種区不老町] e-mail: hitoh@mori.m.is.nagoya-u.ac.jp *²名古屋大学情報連携推進本部情報戦略室 *³オスロー大学臨床効果研究グループ

^{*4} 昭和大学横浜市北部病院消化器センター

^{*5} 名古屋大学情報基盤センター

^{*6}国立情報学研究所医療ビッグデータ研究センター



(a)

図1EC 画像から特徴抽出,特徴選択と可視化を 行う経路.以下の写像を用いることで,EC 画像 I に対して (a) と (b) の 2 通りの経路を定義で きる. f:特徴量抽出, g:画像再構成, s_H :特徴量 空間での特徴選択, s_I :画像空間での画像パター ンの選択. \check{I} : 選択されたパターンで再構成され た画像. \check{H} : 特徴空間で選択された特徴量. (a) および (b) の対角方向の射はそれぞれ $s_H \circ f$ と $f \circ s_I$ を示す.

(b)

い.本稿ではそれぞれの場合に対して以下を実施する.前者に対しては画像再構成を行うgを 求め,選択した特徴量を可視化できるかどうか を検討する.後者に対しては画像空間で腫瘍・ 非腫瘍カテゴリ間に存在する共通の画像パター ンを除去することで識別的なピットパターンの みを再構成できるかどうかを検討する.

2. 手法

1) ピットパターン特徴量からの画像再構成

Haralick 特徴量は画像全体の濃度共起行列の 統計量を表すが、ピットパターン特徴量はさら に回転不変となるように多方向の統計量の平均 となっている [4]. 画像とピットパターン特徴量 の間には1対1対応は存在せず、特徴抽出の逆 変換を直接求めることも不可能である.

HoG 特徴量のように画像の局所領域の勾配 を直接方向統計として表しているような単純な 場合は,辞書画像集合に対応する HoG 特徴量を 基底としたスパースモデリングによって局所領 域ごとの部分的な画像再構成が可能である [7]. しかし,ピットパターン特徴量は画像全体の濃 度共起行列の統計量からなるため,HoG 特徴量 の場合に用いたスパースモデリングによる画像 再構成は不適切である.

ー様乱数を要素に持つベクトルから特定ドメ インの画像を生成する手法として DC-GAN が ある [8]. DC-GAN は画像生成器と識別器のペ アからなる敵対性生成ネットワーク(GAN)の 1種である.さらには、文書から抽出した特徴量 を入力として文書の記述する画像を生成する研 究が活発に行われている[9–11].これらの手法 のいずれもが、様々な構造改変や学習工夫を凝 らしているにせよ、DC-GANの画像生成器をそ のベースにしている.乱数や抽出した特徴ベク トルと画像の構造を対応付ける学習が可能であ るならば、DC-GANの画像生成器を応用するこ とでピットパターン特徴量から画像を再構成で きる可能性がある.

本稿では DC-GAN の生成器の層を深くし、出 カサイズ $64 \times 64 \times 3$ から $256 \times 256 \times 1$ へ変更し た生成器を用いる. さらに、高解像度な画像生成 は困難であることが知られているため [9–11]、 超解像のための深層学習手法に向けて提案され た EDSR residual block [12] ならびに pixel shuffler [13] を利用した画像生成器を新たに定義し た. 加えて、batch normalisation の代わりに instance normalisation [14] を、EDSR residual unit 以外の ReLU を Leaky ReLU [15] で置き換え、 全結合層に dropout [16] を追加した. 図 2(a) お よび (b) は上記の 2 つの構造を示す.

与えられた N 枚の濃淡値 EC 画像のうち i 番目の EC 画像を $X_i \in \{0, 1, 2, ..., 255\}^{H \times W}$ とするとき, X_i から抽出したピットパターン特徴量を $y_i \in [0, 1]^{312}$ とおく.上記の2つの画像生成器の構造に対し, フロベニウスノルム $\|\cdot\|_F$ を用いて定義される次の最小化問題

$$\arg\min\|\boldsymbol{X}_i - \boldsymbol{g}(\boldsymbol{y}_i)\|_{\mathrm{F}}^2 \tag{1}$$

の解としてピットパターン特徴量から EC 画像 を再構成する写像 g を得る.

2) 識別的な画像パターン抽出と画像再構成

離散画像の画素値をある決まった順番で縦に 並べ直して得られる列ベクトルで画像パターン を表す場合,列ベクトルの集合がなす線形部分 空間を用いて画像パターンの集合を表すことが できる.このとき,ある2つのカテゴリの部分 空間を直交させることで,カテゴリ同士に共通 する画像パターンを取り除ける場合がある[17]. 先行研究では抽出した特徴ベクトルのなすカテ ゴリ部分空間から文献[18]の手法を拡張する



図2ピットパターン特徴量(ハンドクラフト特徴量)からEC画像再構成を行う深層学習アーキテク チャ. (a) 拡張 DC-GAN generator. (b) EDSR-residual block を利用した提案手法.

ことで共通の特徴を取り除いた [6]. 本稿では抽 出した特徴ベクトルではなく,離散画像の画素 値からなる列ベクトル集合のなす2つの部分空 間から共通成分を取り除くことで識別的な画像 パターンの抽出を行う.

離散画像の各列を左から順に行方向に連結す る演算子 vec を用いて, 画素値を並べ替えて得ら れるベクトルを $x_i = \text{vec}(X_i) \in \{0, 1, 2, \dots, 255\}^d$ とする. このとき d は $H \times W$ の画素数を表 す. N 個の列ベクトルの平均ベクトルを $\mu = \frac{1}{N} \sum_{i=1}^{N} x_i$ と置くとき, 共分散行列は

$$\boldsymbol{C} = \frac{1}{N} \sum_{i=1}^{N} (\boldsymbol{x}_i - \boldsymbol{\mu}) (\boldsymbol{x}_i - \boldsymbol{\mu})^{\mathsf{T}}$$
(2)

であり,次の固有値分解

$$\boldsymbol{C}\boldsymbol{u}_j = \lambda_j \boldsymbol{u}_j \tag{3}$$

より得られる固有ベクトル u_j , j = 1, 2, ..., mは画像空間を張る直交基底となる. ここで固有 ベクトルは固有値 λ_j に対応し, $\lambda_1 \ge \lambda_2 \ge \cdots \ge$ $\lambda_m \ge 0$ ならびに $u_j^{\mathsf{T}} u_{j'} = \delta_{jj'}$ を満たす. な お, $d \le N$ ならばm = d であり, d > N ならば m = N である.

N 枚の画像を非腫瘍性病変画像 X_{1,k_1} , $k_1 = 1, 2, ..., N_1$ ならびに腫瘍性病変画像 X_{2,k_2} , $k_2 = 1, 2, ..., N_2$ に分割し, カテゴリラベル $l \in \{1, 2\}$

に対して $x_{l,k_l} = \text{vec}(X_{l,k_l})$ とする. 画像の平均 ベクトル μ を原点としたカテゴリごとの自己相 関行列は

$$\boldsymbol{A}_{l} = \frac{1}{N_{l}} \sum_{k_{l}=1}^{N_{l}} (\boldsymbol{x}_{l,k_{l}} - \boldsymbol{\mu}) (\boldsymbol{x}_{l,k_{l}} - \boldsymbol{\mu})^{\mathsf{T}}, \quad (4)$$

である. 式(4) とカテゴリ*l* である確率 *P*(*Cl*) を 用いると式(2) は

$$C = \frac{N_1}{N} A_1 + \frac{N_2}{N} A_1 = \sum_{l=1}^{2} P(C_l) A_l$$
 (5)

となる.

式(3)より得られる固有ベクトルを対角成分と する行列 Λ = diag($\lambda_1, \lambda_2, ..., \lambda_m$)と固有ベクト ルを並べて得られる行列 U = [$u_1, u_2, ..., u_m$] より, 白色化行列 $W = \Lambda^{-\frac{1}{2}}U^{\top}$ を得る. 白色化 行列と式(5)より

$$\sum_{l=1}^{2} \boldsymbol{W} \boldsymbol{P}(\boldsymbol{C}_l) \boldsymbol{A}_l \boldsymbol{W}^{\top} = \sum_{l=1}^{2} \check{\boldsymbol{A}}_l = \boldsymbol{I}$$
(6)

となる. 上式で得られた Ål に対する固有値分解

$$\check{\boldsymbol{A}}_{l}\boldsymbol{v}_{l,j} = \lambda_{l,j}\boldsymbol{v}_{l,j} \tag{7}$$

より得られる固有ベクトル $v_{l,j}$, j = 1, 2, ..., mは固有値 $1 \ge \lambda_{l,1} \ge \lambda_{l,2} \ge ..., \ge \lambda_{l,m} \ge 0$ に対 応し、 $v_{l,j}^{\mathsf{T}} v_{l,j'} = \delta_{jj'}$ を満たす. さらに式 (8) を 考慮することにより、

$$v_{1,j} = v_{2,m-j+1},$$
 (8)

$$\lambda_{1,j} = 1 - \lambda_{1,m-j+1},$$
 (9)

を満たす [6,18].

式(8)と(9)は、1に近い値である $\lambda_{1,j}$ と $\lambda_{2,j}$ に 対応する固有ベクトルそれぞれが非腫瘍性病変 または腫瘍性病変カテゴリのみをよく表す画像 パターンであることを示す.そこで、 $\lambda_{2,j}$ のうち 1または0に近い固有値に対応する固有ベクト ルをn本選んで並べた行列 $V_n = [v_1, v_2, ..., v_n]$ とおき、画像 X_i に含まれる識別的な画像パター ンを

$$\boldsymbol{z}_i = \boldsymbol{V}_n^\top \boldsymbol{W} (\boldsymbol{x}_i - \boldsymbol{\mu}) \tag{10}$$

として抽出する. さらに $W^* = U\Lambda^{\frac{1}{2}}$ とし,

$$\tilde{\boldsymbol{x}}_i = \boldsymbol{W}^* \boldsymbol{V}_n \boldsymbol{z}_i + \boldsymbol{\mu} \tag{11}$$

とすることで画像パターンを再構成する. ここ で vec に対する逆変換 Mat を用いてベクトル \tilde{x}_i の要素を並べ直した 2 次元配列 $\tilde{X}_i = Mat(\tilde{x}_i)$ が再構成画像である.

3. 実験

1) データセット

実験においては通常診察時に収集された超拡 大大腸内視鏡画像の集合 (IRB 承認済) を利用し た. 画像収集においては染色液でポリープ表面 のコントラスト強調を行い、ポリープ表面に接 着した超拡大内視鏡の位置を少しずつ,1つのポ リープから 10 枚から 200 枚ほどの画像を撮影 した.これらの収集した画像には生検・病理診断 を経て腫瘍・非腫瘍ラベルのアノテーションを 施した. 収集した画像から、実臨床における腫 瘍・非腫瘍のパターンの分布の分散を反映した 構成となるように熟練内視鏡医2名による選別 を行った. 収集した画像を患者レベルで重複が ないように訓練・検証データに分割することで 実験用データセットを構築した.表1はデータ セットにおける画像数の内訳を示す.表1にお いてはテストデータの症例数は画像数の比と対 応するように算出した. これら学習・検証データ

表1 実験用データセット

	訓練	検証	計
腫瘍 [枚]	26,462	3,213	29,675
非腫瘍 [枚]	13,180	1,592	14,772
計[枚]	39,642	4,805	44,447
計 [症例数]	2,822	342	3,164

の画像それぞれから文献 [3,4] と同様に高階テ ンソル分解を用いた回転不変かつマルチスケー ル化された Haralick 特徴量を 312 次元のピット パターン特徴量として抽出した.

2) 深層学習手法による画像再構成

図2で示した2種類の構造を Tensorflow バッ クエンドの keras を用いて実装し, NVIDIA 製 GPU V100 (32GB) を1 枚使用してg を求めた. ピットパターン特徴量に対して値域が [0,1] と なるスケーリングのみを行う場合と白色化を行 う場合の2通りに対し、gを求めるための最適 化を行った. 最適化には基礎学習率 lr = 0.0001 およびミニバッチサイズ 256 とし、表1の訓練 データ約 40,000 枚に対して Adam [19] を用い て250エポック分の学習を行った.このとき、提 案手法に対する dropout の割合はr = 0とした. 図 3(a) は得られた4つモデルを用いて訓練デー タに対して画像再構成を行った結果を示す. 図 3(b) は白色化を施したピットパターン特徴量を 用いて検証データに対して画像再構成を行った 結果を示す.

白色化を施したピットパターン特徴量と提案 手法の組み合わせに対する最適化においては, dropout の割合を r = 0.25,0.50 の 2 通りの最適 化を追加で計算し,提案手法に対して 3 つの学 習済みモデルを用意した.ここで,訓練データの ピットパターン特徴量に対して主成分分析を行 い,固有値の大きな固有ベクトルを M 本を用い て表した特徴量を白色化し,提案手法を用いて 画像再構成を行った.図 3(c) は固有ベクトルの 数と dropout の割合に応じた再構成結果を示す.

さらに, M 個の固有値 λ_j に対する累積寄与率 を $\sum_{j=1}^{M} \lambda_j / \sum_{j'=1}^{m} \lambda_{j'}$ として計算し, M 個の固有 値に対応する固有ベクトルで再構成した結果を 図 4(a) として示す. 加えて, 画像の画素値を並 べて列ベクトルとして表した訓練データの画像



(a)



(b)

Reconstruction with M principal components



(c)

図3 深層学習手法による EC 画像再構成. (a) 訓練データに対する再構成結果. (b) 検証データに対する再構成結果. (c) 主成分からの再構成結果. 深層学習時に dropout の割合を r = 0,0.25,0.50 とした.



図4画像空間と特徴空間における主成分の比較. (a) と (b) はそれぞれ画像空間と特徴空間における主成分の累積寄与率と主成分より再構成された EC 画像を示す. *M* = 0 は平均画像を示す. (a) では特徴 空間において選択した主成分からなる特徴量を深層学習手法を用いて EC 画像を再構成した.

集合に対して主成分分析を行い,選んだ固有ベクトルの数に応じて画像を再構成した.図4(b) は累積寄与率と選択した固有ベクトルで再構成 された画像を示す.

3) 画像パターンの解析と選択

訓練画像のピットパターン特徴量,128×128 画素および256×256 画素のEC 画像を表す列べ クトルを対象に、式(7)で定義した固有値分解 を行った.図5は腫瘍性病変カテゴリの白色化 済み自己相関行列の固有値分布を示す. 図 6(a) は128×128 画素の画像サイズに対し、非腫瘍性・ 腫瘍性病変の自己相関行列において閾値以上の 固有値に対応する固有ベクトルで訓練データの 画像再構成した例を示す. 図 6(b) は 256 × 256 画素の画像サイズに対し, 非腫瘍性・腫瘍性病 変の自己相関行列において固有値の大きなもの から上位 K%を用いて訓練データの画像再構成 した結果を示す.図 6(c)は,256×256 画素の画 像サイズに対し,腫瘍性病変の自己相関行列に おいて固有値の大きなものから上位 K%を用い て訓練データの画像再構成した結果を示す.図 6(d)は256×256 画素の画像サイズに対し, 非腫 瘍性・腫瘍性病変の自己相関行列において固有 値の大きなものから上位 K%を用いて検証デー タの画像再構成した結果を示す.

4. 考察

図 3(a) が示すように、 ピットパターン特徴量 に白色化を加えてから学習を行った提案手法の モデルが最も細かいテクスチャ等を再構成でき た. DC-GAN の生成器の場合も、 白色化を前処 理に加えることで改善されたが、細かいテクス チャは再構成できずにボケた再構成画像となっ た. 図 3(a) より訓練データの画像に関しては ピットパターン特徴量からの画像再構成が可能 であることがわかるが、一方で図 3(b) が示すよ うに、検証データの画像再構成には失敗してい る. 図 3(c) より、 dropout による訓練データの過 学習防止を試みると、 細かいテキスチャは再構 成されない. これらの結果は訓練データに対す る過学習を意味している.

図 4(a) では, 特徴空間では 30 本程度の固有 ベクトルに 95%以上の累積寄与率が集中してい るが, 100 本の固有ベクトルで表現したピット パターン特徴量をgに与えても元の画像パター ンは再構成されていない. その上, 図 4(b) が示 す画像空間の主成分で画像再構成を行った場合 と再構成結果の傾向が大きく異なる. 以上の結 果より, 深層学習に基づいた学習済みモデル g は学習で与えられた特徴量と元画像のペアに対 する個別の関係を学習しているに過ぎず, 特徴



図5腫瘍性病変カテゴリの白色化済み自己相関 行列の固有値分布.(a)特徴空間(312次元).(b) 画像空間(16,384次元=128×128画素).(c)画 像空間(65,536次元=256×256画素).

量の分布と元画像の分布間にある構造の対応を 学習しているわけではない.

図5より、312次元から65.536(256×256)次 元という高次元空間に特徴選択のための空間を 変えることで2つのカテゴリのいずれかのみを 表す基底, つまり識別的な画像パターンの抽出 に成功している. 図 5(b) より, 0.99 より大きい 固有値を持つのは腫瘍性病変を表す基底だけで あるが、図 6(a) では上段の再構成画像で非腫表 病変のテクスチャが再構成されている. 図 6(c) は腫瘍性病変を表す基底のみを用いた画像再構 成で腫瘍性病変のパターンのみが再構成されて いる. 図 6(b) の 25-50% 程度の基底で画像再構 成した結果を観察すると、図7で示すような非 腫瘍・腫瘍病変の診断で注目される特徴的な形 状が存在することが確認できる.これらの特徴 的な形状は図 6(a) 中の λ_i > 0.95 の場合のよう に少数の基底で再構成した場合でも確認できる.

一般的な主成分分析として少数の主成分から 画像再構成を行っても図 4(b) のようにボケた 画像となることに対し,識別的な画像パターン を表す基底を求めることで図 5(a)-(c) のように 少数の基底で細かいテクスチャが再構成できる. これは、EC 画像の識別的な画像パターンが特 定のピットパターンにあることを示す.ただし、 図 6(d) が示す腫瘍性病変を表す基底のみを用い た検証データに対する画像再構成結果では非腫 瘍・腫瘍性病変の両方のテクスチャが表れてい る.これは高次元の画像空間が広く、高次元空 間で EC 画像パターンの分布を表現するに十分 な訓練データが揃っておらず、訓練データのみ に有効な特徴選択となっていることを示す.

高次元の画像空間に対しては訓練データ数は 不十分であり、図 4(b)の累積寄与率が示すよう に少数の基底で画像パターンの分布を表せてい ない. しかし図 4(a) が示すように,特徴空間で は画像パターン分布を少数の基底で表現できる. さらにピットパターン特徴量に対しては十分な 汎化性能を与える訓練データとなっていること が報告されており [4], 訓練データ以外に対して も本稿と同様の数理で識別的な特徴量表現が可 能であることが報告されている [6]. この特徴空 間における識別的な特徴選択に関する報告では、 図4(a)と同様に一方のみを表す基底を求めるこ とはできないが、本稿と同じ数理に基いて共通 成分を取り除くことによって分類精度向上を達 成している. 画像空間と特徴空間を直接対応付 けた可視化は困難であるが、2つの空間におい て同じ方法で特徴選択を行うことで、どのよう な特徴選択が行われているかという意味をアナ ロジーとして説明できるものであると考える.

5. むすび

本稿では EC 画像の病理類型分類における識 別的なパターンを可視化するための初期的検討 を行った. ピットパターン特徴量からの EC 画像 再構成は不良設定問題であり,回帰問題として 深層学習ベースで写像を求めても,意味のある 画像再構成が得られないことが判明した. 画像 空間における識別的な画像パターンの抽出・再 構成においては,少数の識別的な画像パターン による画像再構成結果に,実際の診断で重要視 されるピットパターンが表れることを確認した.

謝辞 本研究の一部は JSPS/MEXT 科研費 (No. 26108006, 17H00867, 21K19898), 堀科学芸術振

Reconstruction with major discriminative components

(a)

Reconstruction with major *K*% discriminative components

Teeonstruction with high K/0 discriminative components				Original		
K = 5	K = 10	K = 25	K = 50	K = 75	K = 90	Original
					-	1

(b)

Reconstruction with major K% discriminative components of neoplasia

Reconstruction with major K% discriminative components of neoplasia				Original		
<i>K</i> = 5	K = 10	K = 25	<i>K</i> = 50	K = 75	K = 90	Oliginai

(c)

Reconstruction with majo K% principal components of neoplasia

Reconstruction with high K/0 principal components of neoplasta				Original		
<i>K</i> = 5	K = 10	K = 25	K = 50	K = 75	K = 90	Original

(d)

図6選択した識別的な画像パターンを用いた EC 画像再構成結結果. (a) 16,384 次元の画像空間で求めた識別的な画像パターンから再構成した学習データの画像. (b) 65,536 次元の画像空間で求めた識別的な画像パターンから再構成した学習データの画像. (c) 65,536 次元の画像空間で求めた腫瘍性病変に関する識別的な画像パターンのみから再構成した学習データの画像. (d) 65,536 次元の画像空間で求めた識別的な画像パターンから再構成した検証データの画像. (a)-(d) において上段および下段はそれぞれ非腫瘍性病変と腫瘍性病変の EC 画像を示す.



non-neoplasia

adenoma

図7 選択した識別的な画像パターンにより再構成した EC 画像中で確認できる特徴.(左)非腫瘍病変に見られる腺腔小円形(矢印).(右)腫瘍性病変(腺腫)における腺腔スリット(破線内)と紡錘形細胞核(矢印).2つの EC 画像はともに図 6(b)の k = 50 の場合を示す.

興財団そして AMED (No. 18hk0102034h0103, 19hs01100 06h0003, 19lk1010036h0001)の支援 を受けた。

利益相反 なし

文献

- Kudo S-E, et al.: The Diagnosis of Colorectal Tumors with Endocytoscopy (Japanese), Gastroenterological endoscopy, 55(4): 1510-1517, 2013.
- [2] Wimmer G, et al.: Directional wavelet based features for colonic polyp classification, Medical Image Analysis, **31**: 16-36, 2016.
- [3] Mori Y, et al.: Real-Time Use of Artificial Intelligence in Identification of Diminutive Polyps During Colonoscopy: A Prospective Study, Annals of Internal Medicine, 169(6): 357-366, 2018.
- [4] Itoh H, et al.: Robust endocytoscopic image classification based on higher-order symmetric tensor analysis and multi-scale topological statistics, IJCARS, 15: 2049-2059, 2020
- [5] Haralick, RM, et al.: Textural Features for Image Classification, IEEE Transactions on Systems, Man, and Cybernetics, SMC-3(6): 610-621, 1973.
- [6] Itoh H, et al.: Discriminative Feature Selection by Optimal Manifold Search for Neoplastic Image Recognition, Proc. ECCV workshps, 4: 534-549, 2018
- [7] Vondrick C, et al.: Visualizing Object Detection Features, International Journal of Computer Vision, 119: 145-158 (2016)

- [8] Radford A, et al.: Unsupervised representation learning with deep convolutional generative adversarial networks, Proc. ICLR, 2016
- [9] Zhang H, et al.: StackGAN: Text to Photo-Realistic Image Synthesis with Stacked Generative Adversarial Networks, Proc. CVPR, :5908-5916, 2017
- [10] Xu T, et al.: AttnGAN: Fine-Grained Text to Image Generation with Attentional Generative Adversarial Networks, Proc. CVPR, :1316-1324, 2018
- [11] Zhang Z, et al.: Photographic Text-to-Image Synthesis with a Hierarchically-Nested Adversarial Network, Proc. CVPR, :6199-6208, 2018
- [12] Lim B et al.: Enhanced Deep Residual Networks for Single Image Super-Resolution Proc. CVPR workshops, :1132-1140, 2017
- [13] Shi W et al.: Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network, Proc. CVPR, :1874-1883, 2016
- [14] Ulyanov D, et al.: Improved Texture Networks: Maximizing Quality and Diversity in Feed-forward Stylization and Texture Synthesis, Proc. CVPR, :4105-4113, 2017
- [15] Maas AL, et al.: Rectifier Nonlinearities Improve Neural Network Acoustic Models, Proc. International Conference on Machine Learning, 2013.
- [16] Srivastava N, et al.: Dropout: A Simple Way to Prevent Neural Networks from Overfitting, Journal of Machine Learning Research, 15: 1929-1958,2014.
- [17] Itoh H, et al.:Dimension Reduction and Construction of Feature Space for Image Pattern Recognition, Journal of Mathematical Imaging and Vision, 56:1-31, 2016
- [18] Fukunaga K, et al.: Application of the Karhunen-Lo[´] eve expansion to feature selection and ordering, IEEE Transactions on Computers C-19(4):311-318, 1970
- [19] Kingma DP and Ba J: Adam: A Method for Stochastic Optimization, Proc. International Conference on Learning Representations, 2015.

Preliminary Stduy of Endocytoscpic Image Reconstruction for Pit-Pattern Feature Analysis

Hayato Itoh^{*1}, Masahiro Oda^{*2,1}, Yuichi Mori^{*3,4}, Masashi Misawa^{*4}, Shin-Ei Kudo^{*4}, Kensaku Mori^{*1,5,6}

*¹ Graduate School of Informatics, Nagoya University *² Information Strategy Office, Information

and Communications, Nagoya University

^{*3} Clinical Effectiveness Research Group, University of Oslo

*4 Digestive Disease Center, Showa University Northern Yokohama Hospital

*5 Information Technology Center, Nagoya University

^{*6} Research Center for Medical Bigdata, National Institute of Informatics

We explore visualisation methods of discriminative pit patterns in pathological pattern classification for endocytoscopic images. Pit pattern is an texture pattern in a magnified view of polyp's surface. Previous works demonstrated accurate pathological pattern classification by using a handcraft feature extractoin of pit patterns. Another previous work also suggested discriminative feature selection from the handcraft features for the achievement of precise pathological pattern recognition. On the other hand, a relation between discriminative pit pattern and discriminative handcraft feature is still unclear. Unfortunately, however, the handcraft feature extraction does not have its inverse operation, even though visualisation technique of discriminative handcraft feature is required for the interpretation of feature selection. Therefore, an image reconstruction from a handcraft feature is an ill-posed problem. In this work, by asuming that pairs of endocytoscopic images and their extracted handcraft features are given, we tackled with a visualisation of discriminative pit patterns with two approaches: image reconstruction from extracted feature via deep learning and direct extraction of discriminative pit pattern via algebraic image analysis.

Keywords: Endocytoscope, pathological pattern classification, white boxing, feature selection, feature visualisation

X線暗視野法を用いた乳腺における乳管内腺腔の自動抽出

および3次元解析手法の開発

谷口 華奈*1 砂口 尚輝*1 市原 周*2 西村 理恵子*2

岩越 朱里*2 村上 善子*2 島雄 大介*3 湯浅 哲也*4

安藤 正海*5

要旨

乳腺疾患の病理診断において、2次元で観察する組織標本上は、非浸潤性乳管癌(DCIS)の一部は特徴的な篩 状構造を示すが、良性の通常型乳管過形成(UDH)も篩状様の構造を示すことがある.そのため、2次元的観 察では DCIS との鑑別が困難なことがある.しかし、これらの3次元構築は異なるため、腺腔の3次元構 築の可視化は貴重な情報をもたらす.我々はこれまでX線暗視野法による屈折コントラスト CT 像から腺 腔を手動抽出し、3次元可視化・解析を行い、DCIS と UDH 各1症例において3次元腺腔構造に差異があ ることを示した.しかし、乳管内には膨大な数の腺腔があるため、手動抽出では労力が大きく狭い領域の 観察に限られる.本研究では、短時間での腺腔抽出を実現するために U-Net を用いた自動抽出および解析 手法を提案する.7症例のCT 像から自動腺腔抽出を試み、手動による抽出結果を正解として比較し、高い 類似度が得られた.

キーワード: U-Net, セグメンテーション, 屈折コントラストX線CT, X線暗視野法, 乳癌

1. はじめに

女性の癌罹患数の中で最も高い割合を占め るとされる乳癌の多くは、乳管に発生する乳管 癌である.中でも、乳管以外の組織への浸潤が みられず癌細胞が乳管系にとどまっているも のを非浸潤性乳管癌(Ductal carcinoma in situ;

*1 名古屋大学大学院医学系研究科
〔〒461-8673 名古屋市東区大幸南 1-1-20〕
e-mail:taniguchi.kana@f.mbox.nagoya-u.ac.jp
*2 名古屋医療センター病理診断科
*3 北海道科学大学保健医療学部
*4 山形大学大学院理工学研究科
*5 高エネルギー加速器研究機構 (KEK)

DCIS)といい,組織を薄くスライスし染色した ものを顕微鏡で観察する病理診断が行われる.

病理診断において,篩状腺腔構造(正円形に 近く,大きさが規則的,多孔性)を示す一部の DCIS と,不規則な非円形でスリット状の腺腔 構造が特徴的にみられるが,時に篩状様の腺腔 構造を示すことのある良性の通常型乳管過形 成(Usual ductal hyperplasia; UDH)について,顕 微鏡による2次元観察では鑑別が困難な場合が あると指摘されている.しかし,これらは3次 元構築が異なることも報告されており[1],腺腔 構造の3次元的な観察は貴重な情報をもたらす.

これまでにも,連続する複数の病理切片から 腺腔の3次元像を作成する研究が報告されてお り,2次元の病理観察では得られなかった3次 元構造が明らかになった[1~3].しかし,病理像 による3次元化手法は膨大な数の病理切片の作 成が必要となり,多大な労力がかかることから, 少ない症例数でのみ議論が行われてきた.

そこで我々は X 線暗視野法[4]に基づく屈折 コントラスト CT [5]で撮像された乳腺組織 CT 画像から腺腔を手動で抽出し、3次元可視化・ 解析を行った. 屈折コントラスト CT は放射光 X 線を利用して物質の密度分布を画像化する. 非破壊・未染色で染色組織像に匹敵するコント ラストの3次元像を得ることができ、近年病理 診断学への応用に向けた研究が進められてい る[6]. 我々が手動抽出で行った DCIS と UDH 各1症例の3次元解析により、DCISではバブ ル状, UDH ではチューブ状の腺腔構造を示すこ とが明らかになり、これらの3次元的な特徴が DCIS と良性の乳管内増殖病変の鑑別において 重要な指標になり得ることを示した.しかし、 乳管内に存在する膨大な数の腺腔を全て手作 業で抽出するために要する時間と労力は大き く, 少ない症例数かつ狭い領域の観察に限られ ることが課題であった.

本研究では、短時間での腺腔抽出を実現する ために U-Net [7]を用いた自動セグメンテーシ ョン手法および3次元解析手法を提案する.U-Net は画像のセグメンテーションを推定する全 層畳み込みネットワークモデルの1種である. 本手法により、多数症例を用いた解析と解析領 域の拡大が可能となる.

2. 方法

1) 撮像試料

試料は乳癌と診断された日本人女性から摘 出された乳腺組織ブロックである. DCIS また は UDH を含む計 7 症例(症例 A~G)を解析対 象とした.7 症例のうち 5 症例(症例 A~E)は DCIS,2 症例(症例 F~G)UDHである.これ らは 10%緩衝ホルマリン液による固定後,エタ ノールに浸して保存されており,撮像の 5~6時 間前,エタノールを除去するために蒸留水に移 された.撮影時はアガロースで満たしたアクリ ル容器内に試料を固定した.本研究は,名古屋 大学および名古屋医療センターの生命倫理審 査で承認が得られている.

2) 屈折コントラスト CT 画像の取得

高エネルギー加速器研究機構(茨城県つくば 市)フォトンファクトリーBL14Bビームライン に構築された屈折コントラスト CT を用いて試 料の撮像を行った.撮像システムの概念図を図 1に示す.加速器から発生した放射光 X線は 2 結晶モノクロメーターにより単色化,非対称ブ ラッグケースモノクロコリメーター結晶(MC) で回析し,試料に入射する.試料の電子密度に 従って微小に屈折した X線はラウエケース角 度アナライザー結晶(LAA)でX線の強度情報 に変換され,X線カメラで測定される.この測 定を試料の回転毎に繰り返すことで投影デー タを得た.その他,投影データ撮像,CT 再構成 に使用したパラメータを表1に示す.

3) 腺腔の自動抽出および3次元解析手法

本稿では、U-Net を用いて腫瘍内の腺腔を自 動でセグメンテーションし、腺腔構造の3次元 解析を行う手法を提案する.図2に提案手法の 概要を示す.



図1 屈折コントラスト CT 撮像システム.

表1. CT 撮像,再構成パラメータ

X-ray energy	19.8keV	
Pixel size	5.5 µm	
Number of	2500	
projections		
	4608×4608×2592 pixels	
Volume size	$(25 \times 25 \times 14 \text{ mm}^3)$	
Slice thickness	5.5 µm	
Reconstruction	Filtered back projection	
algorithm	(FBP)	



図2提案手法の概要.

1. データ作成

CT 画像から乳腺腫瘍部分を含む領域を解析 対象領域として選択する.大きさは 256×256 [pixels] (1.41×1.41 [mm²])で,連続する 100 ス ライスを選択する.次に U-Net のトレーニング 用データを作成する.解析対象領域 100 スライ スのうち,等間隔に 10 枚のスライスについて 手動で腺腔抽出を行い,腺腔領域の情報のみを 持つ2 値化画像を作成する.なお,腫瘍内に存 在する腺腔の内部は他組織に比べ密度の低い 組織液で満たされていることから,屈折コント ラスト CT 画像上で腺腔は黒く表現される.手 動抽出については,画像処理ソフト Image J [8] の Drawing Tools を使用して,病理医に指導を 受けた診療放射線技師 1 名が実施した.

2. セグメンテーション

手動抽出した 10 スライスの腺腔像をトレー ニング用画像として,U-Net でトレーニングを 行う.オプティマイザーはAdamを用い,学習 回数を10000とする.腺腔抽出精度を評価する ため,正解画像と出力画像の類似度を評価する Jaccard index (1) [8]を計算する.Jaccard index が 大きいほど類似度が高いと評価され,抽出精度 が高いといえる.U-Net へ画像をインプットす る際は,画像全体の画素値が 0~255 となるよう に正規化を行っている.

$$Vaccard\ index = \frac{|A \cap B|}{|A \cup B|} \tag{1}$$

トレーニングを行った U-Net モデルを用いて, 手順 1.で作成した解析対象領域の CT 画像 100 スライスをインプットとした腺腔自動セグメ ンテーションを行う.出力された画像を目視で 確認,微修正を行う.

3. 3 次元解析

セグメンテーションで得られた腺腔像デー タを用いて,腺腔構造の3次元解析を行う.腺 腔構造の3次元可視化像を作成し,腺腔数,腺 腔体積,腺腔経路長を計測する.3次元可視化 像の作成では,画像解析ソフト Avizo[9]を使用 してボリュームレンダリングを行った.

3. 結果と考察

7 症例において自動セグメンテーションを行った結果,Jaccard index は平均値 82.2%となった.図3に本手法でトレーニングした U-Net での腺腔抽出結果を示す.左から,入力画像,出力画像,セグメンテーション画像である.全て手作業で腺腔抽出を行っていた場合と比較して,短時間で高い精度の自動セグメンテーションを行うことができた.具体的には,以前の手動抽出手法では30 スライスの3 次元化に約2時間の作業時間を要していたが,本研究では自

動抽出手法の採用により 100 スライスの 3 次元 化においても手動抽出作業を行ったのは 10 ス ライスであり、3 分の1の作業時間で実施でき た.しかし、一部の腺腔について過剰な抽出や 欠損も確認された.本研究では、正確な解析を 行うために、出力結果を目視で確認し必要な箇 所を微修正しているが、入力画像に先鋭化処理 やモロフォロジカル処理などの前処理を行う ことで、より高精度な抽出を実現し、微修正を 必要としないセグメンテーションに改善でき ると考える.

セグメンテーションで得られた腺腔像デー タからボリュームレンダリングした3次元可視 化像(症例 A)を図4に示し,腺腔数,腺腔体 積,腺腔経路長の計測結果を表2にまとめた. なお,腺腔体積,腺腔経路長については各症例 における1 腺腔あたりの平均値を示している. 例えば,DCIS である症例 A と UDH である症 例 F を比較すると,症例 A は腺腔数が大きく, 腺腔経路長が小さい.これは DCIS で特徴的な バブル状構造と UDH で特徴的なチューブ状構 造を示しており,以前の手動抽出による解析と 矛盾の無い結果となった.しかし,今回7症例 の解析を行ったことにより,同一疾患の症例に おいても解析結果に大きな差異があることが 明らかとなった.本稿では,癌のグレードや進 展度,発生箇所などの違いによって,腺腔数, 腺腔体積,腺腔経路長に差異が生じていると考 えるが,症例数を増やし,多数症例で統計的に 分析を行って考察する必要がある.



図3 腺腔抽出結果



図4 症例Aにおける腺腔の3次元可視化像.

		腺腔数	腺腔体積 [µm 3] (平均値)	腺腔経路長 [µm] (平均值)
DCIS	症例 A	263	46764.7	19.4
	症例 B	315	41693.0	53.2
	症例 C	257	44426.7	50.9
	症例 D	388	37223.0	34.3
	症例 E	896	112774.4	72.9
UDH	症例 F	152	27292.1	59.2
	症例 G	658	228824.8	217.4

表 2. 3 次元解析結果

4. まとめ

DCIS または **UDH** を含む計7 症例の乳腺組織 試料の屈折コントラスト **CT** 画像を撮影し,

U-Net を用いた自動腺腔抽出および3次元解析 手法を試みた.その結果,従来の手動腺腔抽出 手法と比較して,短い時間で多症例かつ広い領 域におけるセグメンテーションと解析を実現 できた.

今後は、トレーニング手法の改良により腺腔 の抽出精度を高めていくとともに、症例数を増 やして統計的な解析を行っていきたい.さらに、 より詳細な解析を行うために解析項目を増や し、新たな病理組織学的情報の取得を目指して いきたい.

謝辞

本研究の試料を準備していただいた名古屋 医療センター病理診断科の皆様に感謝申し上 げます.本研究は JSPS 科研費 16K01369, 16K08654, 22591353, 15H01129, 26286079, 18K13765,平成 29 年度国立病院機構共同臨床 研究 H29-NHO(癌般)-01, H29-NHO(多共)-02,知 の拠点あいち重点プロジェクト第 3 期および KEK 放射光実験課題 2016G0625 の支援を受け ています.

利益相反の有無

なし.
文 献

- Ohuchi N, Abe R, Takahashi T et al. : Three-dimensional atypical structure in intraductal carcinoma differentiating from papilloma and papillomatosis of the breast. Breast Cancer Research and Treatment 5: 57-65, 1985
- [2] Norton KA, Namazi S, Barnard N et al. : Automated Reconstruction Algorithm for Identification of 3D Architectures of Cribriform Ductal Carcinoma In Situ. PLoS ONE 7:e44011, 2012
- Booth ME, Treanor D, Roberts N et al. : Three-dimensional reconstruction of ductal carcinoma in situ with virtual slides . Histopathology, 66, 966-973, 2015
- [4] Ando M, Maksimenko A, Sugiyama H et al. : Simple X-Ray Dark- and Bright-Field Imaging Using Achromatic Laue Optics. Jpn. J. Appl. Phys. 41 No.9A, B, L1016-L1018 2002.
- [5] Sunaguchi N, Yuasa T, Huo Q et al. : X-ray refraction-contrast computed tomography images using dark-field imaging optics. Appl Phys Lett, 107, 253701, 2015.
- [6] 砂口尚輝,島雄大介,市原周他:X線暗 視野法に基づく屈折コントラストX線 CTを用いたヒト乳頭の三次元可視化お よび解析. Med Imag Tech **38**: 33-39, 2020
- [7] O. Ronneberger, P. Fischer, T. Brox : U-Net: Convolutional Networks for Biomedical Image Segmentation. The Medical Image Computing and Computer Assisted Intervention Society (MICCAI), vol.9351, pp.234-241, 2015
- [8] National Institutes of Health (NIH): image J Available: https:// imagej.nih gov/ij/
- [9] Jaccard, P: Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines. Bulletin de la Société Vaudoise des Sciences Naturelles 37, 241-

272, 1901

[10] Thermo Fisher Scientific, Avizo Available:https://www.thermofisher.com/jp /ja/home/industrial/electronmicroscopy/electron-microscopyinstruments-workflow-solutions/3dvisualization-analysis-software.html

Automatic extraction and 3D analysis methods of lumens in breast tissue

using X-ray Dark-Field Imaging

Kana TANIGUCHI^{*1}, Naoki SUNAGUCHI^{*1}, Shu ICHIHARA^{*2}, Rieko NISHIMURA^{*2}, Akari IWAKOSHI^{*2}, Yoshiko MURAKAMI^{*2}, Daisuke SHIMAO^{*3}, Tetsuya YUASA^{*4}, Masami ANDO^{*5}

- *1 Nagoya University
- *2 Nagoya Medical Center
- *3 Hokkaido University of Science
- *4 Yamagata University
- *5 High Energy Accelerator Research Organization (KEK)

In the pathological diagnosis of breast diseases, some Ductal carcinoma in situ (DCIS) show a characteristic cribriform structure in 2D histological specimens, while benign Usual ductal hyperplasia (UDH) may also show cribriform-like structures. Therefore, it is sometimes difficult to differentiate DCIS from UDH by two-dimensional observation. However, because their 3D structures are different, visualization of the 3D structure of the glandular lumen provides valuable information. We have manually extracted lumen from refraction-contrast CT images based on X-ray dark field imaging of DCIS and UDH in one case each, visualized and analyzed them in 3D, and showed that there were differences in three-dimensional lumen structures between DCIS and UDH. However, manual extraction of the large number of lumen in the ducts is labor-intensive and limited to observation area. In this study, we propose an automatic extraction and analysis method using U-Net to enable quick extraction of lumen.

Key words: U-Net, segmentation, refraction-contrast CT, X-ray dark field imaging, breast cancer

圧縮センシング CT 画像再構成のための FBP 法組み込み型

主双対反復法とコーンビーム CT への拡張

工藤 博幸*1 王 婷*1

要旨

スパースビューCT や低線量 CT の画像再構成はトータルバリエーション (TV) と最小2 乗誤差データ項の 和を最小化する圧縮センシングにより行われ,評価関数最小化には Chambolle-Pock アルゴリズムや FISTA などの反復法が用いられる.しかし、これらの反復法は収束が非常に遅く計算量が膨大であり、オーダー ドサブセット(OS)法などの経験的な手法を組み合わせ高速化が行われているのが現状である.しかし, OS 法は評価関数の厳密な最小値に収束せず並列化が困難である.筆者らは,既に新しい別の高速化の数学 的枠組みとして, FBP 法の構造を主双対反復法に埋め込んで高速化を実現する FBP 法組み込み型主双対反 復法を平行ビーム CT とファンビーム CT に対して提案している (Kudo et al, Proc.SPIE, 2016, Wang et al, Phys.Med.Biol., 2019). 本手法は、厳密に評価関数の最小値に収束し FBP 法の構造が基本であるため並列化 が容易で利便性が高い. 本研究では, FBP 法組み込み型主双対反復法のよく知られる反復法構築の数学的 枠組みである Chambolle-Pock アルゴリズムに前処理を導入した枠組みに基づき新しい数学的導出を試み、 鞍点問題を解く前処理付き近接点法と見ることができることを示し、収束性を満たす FBP 法前処理フィル タの条件を導出し前論文では経験的で曖昧であった収束性の問題をすっきり解明した.また,FBP 法組み 込み型主双対反復法を円軌道コーンビーム CT 画像再構成に拡張した.実験では,コーンビームのスパー スビューCT 及び低線量 CT の画像再構成のシミュレーション実験を行い、本手法が有効に動作することを 示した.更に,256列検出器のキャノンコーンビーム CT 装置実機投影データの画像再構成を行い,良好な 結果を得た.

キーワード: CT, 画像再構成, 圧縮センシング, トータルバリエーション, コーンビーム

1. はじめに

スパースビューCT や低線量 CT の画像再構 成はトータルバリエーション(TV)と最小2乗 誤差データ項の和を最小化する圧縮センシン グにより行われ,評価関数最小化には Chambolle-Pock アルゴリズムや FISTA などの反 復法が用いられる [1]-[4].しかし,これらの反 復法は収束が非常に遅く計算量が膨大であり, オーダードサブセット(OS)法などの経験的な 手法を組み合わせ高速化が行われているのが 現状である.しかし,OS 法は評価関数の厳密な 最小値に収束せず並列化が困難である.筆者ら は,既に新しい別の高速化の数学的枠組みとし

*1 筑波大学システム情報系 〔〒305-8573 つくば市天王台 1-1-1〕 e-mail: kudo@cs.tsukuba.ac.jp 投稿受付: 2021 年 6 月 14 日

て,FBP 法の構造を主双対反復法に埋め込んで 高速化を実現する FBP 法組み込み型主双対反 復法を平行ビーム CT とファンビーム CT に対 して提案している [5],[6]. 本手法は, 厳密に評 価関数の最小値に収束し FBP 法の構造が基本 であるため並列化が容易で利便性が高い.本研 究では,FBP 法組み込み型主双対反復法のよく 知られる反復法構築の数学的枠組みである Chambolle-Pock アルゴリズムに前処理を導入し た枠組みに基づき新しい数学的導出を試み、鞍 点問題を解く前処理付き近接点法と見ること ができることを示し, 収束性を満たす FBP 法前 処理フィルタの条件を導出し前論文では経験 的で曖昧であった収束性の問題をすっきり解 明した.また,FBP 法組み込み型主双対反復法 を円軌道コーンビーム CT 画像再構成に拡張し た. 実験では、コーンビームのスパースビュー CT 及び低線量 CT の画像再構成のシミュレー

ション実験を行い、本手法が有効に動作するこ とを示した.更に、256 列検出器のキャノンコ ーンビーム CT 装置実機投影データの画像再構 成を行い、良好な結果を得た.

2. 提案手法

1) FBP 法組み込み型主双対反復法の新導出

FBP 法組み込み型主双対反復法は文献[5],[6] では全く別の方法で導出されたが、収束性を保 証させる FBP 法前処理フィルタの条件が必ず しも明確でなかった.そこで、トータルバリエ ーション (TV) 正則化項つきの画像再構成問題 を解く反復法構築に用いられる Chambolle-Pock アルゴリムに高速化の前処理を導入した枠組 みを利用して、再導出を行った.

本章では,投影データ収集の幾何学系として, 2 次元の平行ビーム方式でスパースビューCT や低線量 CT を行う状況を想定する.画像を x, 投影データを b,システム行列を A で表し,次 式の TV 正則化項つき画像再構成問題を考える.

$$\min_{\vec{x}} f(\vec{x}) \equiv G(\vec{x}) + \frac{1}{2} \left\| A\vec{x} - \vec{b} \right\|^2$$

$$G(\vec{x}) \equiv \beta \left\| \vec{x} \right\|_{\text{TV}}$$
(1)

ただし,式(1)では正則化項はTVとしているが, 提案手法は一般性があり *G*(**x**)は近接写像が計 算可能な凸関数の正則化項であればTV以外で もよい(Nonlocal TV や核ノルムなど).提案手 法では, Chambolle-Pock アルゴリズムと同様に, 式(1)を以下の鞍点問題の形に定式化した問題 を解く.

$$\min_{\vec{x}} \max_{\vec{y}} G(\vec{x}) - F^{*}(\vec{y}) + \vec{y}^{T}(A\vec{x})$$

$$F^{*}(\vec{y}) \equiv \min_{\vec{z}} (\vec{y}^{T}\vec{z} - \frac{1}{2} \|\vec{z} - \vec{b}\|^{2}) = \frac{1}{2} \|\vec{y} + \vec{b}\|^{2}$$
(2)

式(2)を(x,y)について劣微分してゼロとおくと, 解の鞍点が満たす式は以下となる.

$$F\begin{pmatrix}\vec{x}\\\vec{y}\end{pmatrix} \equiv \begin{pmatrix}\partial G(\vec{x}) + A^T \vec{y}\\-A \vec{x} + \partial F^*(\vec{y})\end{pmatrix} = 0$$
(3)

式(3)を直接解くことは、G(x)が微分不可能なこ

と、xとyが分離(スプリッティング)されて いないことなどの理由で困難である.そこで, Chambolle-Pock アルゴリズムでは次式で表され る近接点法を用いて反復的に解く.

$$k = 0,1,2,\cdots$$

$$F\left(\frac{\vec{x}^{(k+1)}}{\vec{y}^{(k+1)}}\right) + M\left(\frac{\vec{x}^{(k+1)} - \vec{x}^{(k)}}{\vec{y}^{(k+1)} - \vec{y}^{(k)}}\right) = 0 \quad (4)$$

$$M \equiv \left(\frac{1/\tau - A^{T}}{-A P^{-1}/\sigma}\right)$$

ただし、kは近接点法の反復回数、Mはxの更 新とyの更新を分離するために導入された近接 点法の前処理行列である [2].また、Mの中に ある行列 P は通常の Chambolle-Pock アルゴリ ズムには存在しない、FBP 法組み込み型反復法 で高速化を実現するために新しく導入した前 処理行列である.以降では、MとPを区別する ためPを「フィルタ基礎行列」と呼ぶ.式(4)か ら、xとyの具体的な更新式を求めると次の鞍 点問題を解く反復法が得られる.

Mail Loop: $k = 0, 1, 2, \cdots$

[Step 1] 主変数更新

$$\vec{x}^{(k+1)} = \Pr_{\tau_i}(\vec{x}^{(k)} - \tau A^T \vec{y}^{(k)})$$
(5)

[Step 2] 主変数 Extrapolation

$$\bar{x} = 2\bar{x}^{(k+1)} - \bar{x}^{(k)} \tag{6}$$

[Step 3] 双対変数更新

$$\vec{y}^{(k+1)} = \arg\min_{\vec{y}} \left(F^{*}(\vec{y}) + \frac{1}{2\sigma} \left\| \vec{y} - (\vec{y}^{(k)} + \sigma P A \bar{x}) \right\|_{P^{-1}}^{2} \right)$$
$$= \vec{y}^{(k)} + Q(A \bar{x} - \vec{b} - \vec{y}^{(k)})$$
$$Q = \sigma (I + \sigma P)^{-1} P$$
(7)

式(5)~(7)が提案手法の反復式であるが,フィル タ基礎行列 P を単位行列に選べば通常の Chambolle-Pock アルゴリズムと一致する. 直感 的には,式(5)~(7)は以下の構造をしていると考 えると,通常の逐次近似画像再構成法との関係 が分かりやすい.式(7)では,更新画像を順投影 した投影データ $A\bar{\mathbf{x}}$ と実測投影データ \mathbf{b} の誤差 を計算し、それに収束高速化のため(平滑化し た)FBP 法のランプフィルタ Q をかけている. 式(5)では、その誤差を A^T で逆投影して画像更 新を行った後 TV の近接写像で平滑化を行う. 双対変数 \mathbf{y} はノイズを表す変数である.以降で は、Q を「フィルタ前処理行列」と呼ぶ.

2) 収束が保証されるフィルタ前処理行列

上述の解釈からフィルタ前処理行列 Q はラ ンプフィルタの意味を持つ High Pass フィルタ であることは分かるが,提案手法が理論的に収 束するためには P 及びステップサイズ(r,o)は以 下の 2 つの条件を満足しなくてはならない. [条件 1] 式(4)が前処理つき近接点法になるた めには, M は正定値対称行列である必要があり, よって次の行列 D は正定値対称行列でなくて はならない.

$$D \equiv \frac{1}{\tau\sigma} P^{-1} - AA^{T} \tag{8}$$

[条件 2] フィルタ基礎行列 *P* 自体が正定値対称行列でなくてはならない.

また,上述の Q の意味から P は Q がランプフィルタに近くなる行列である必要がある.以上の条件を満たす自然な P とステップサイズ(τ,σ)の選び方は次式である.

$$P^{-1} = AA^T, \tau \sigma < 1 \tag{9}$$

式(9)が満足されれば上述の条件1と条件2が満 足されることは容易に示せる.更に,CT画像再 構成分野で,AA^Tは投影データに周波数特性が 1/|ω|のランプフィルタ逆特性を持つフィルタを かけてぼかす演算と等価なことが知られる.こ の性質をフーリエ変換の記号Fを用いて次式の ように表す.

$$P^{-1} = KF^{-1} \frac{1}{|\omega|}F$$
 (10)

ただし,Kはサンプリングの条件から定まる既 知の定数である.このとき,Pに対応するQは 次式となる.

$$Q = \sigma (1 + \sigma P)^{-1} P = F^{-1} \frac{\sigma}{1 + \sigma |\omega| / K} \frac{|\omega|}{K} F$$
$$= F^{-1} \frac{|\omega|}{|\omega| + K / \sigma} F$$
(11)

式(11)が提案手法で用いる前処理フィルタ行列 *Q*の周波数特性である.ステップサイズ σの値 によって平滑化の度合いが変化する平滑化ラ ンプフィルタになっている.更に,もう一つの ステップサイズ τ は式(9)にある τσ<1 の条件を 満足するように選ぶ必要がある.最終的な,前 処理フィルタの選択に関する結論を以下にま とめる.

[収束条件] 提案手法は,前処理フィルタ行列*Q* とステップサイズ(*r*,*o*)を次式の条件を満足する ように選べば,画像 **x**^(k)は式(1)の問題の解に収 束する.

$$Q = F^{-1} \frac{|\omega|}{|\omega| + K/\sigma} F, \tau\sigma < 1$$
⁽¹²⁾

Qはランプフィルタに近いものなら何でも収束 する訳ではなく、収束を保証するには(r,o)も合 わせて式(12)を満足するように選ばなければな らない. 先行研究の文献[5],[6]ではこの点が明 確でなく、本研究ではそれを明確に示し収束性 の問題を解決したと言える.

3) コーンビーム CT への拡張

本研究のもう一つの成果は,FBP 法組み込み 型主双対反復法を円軌道コーンビーム CT でス パースビューCT や低線量 CT を行う場合に拡 張して,実装と実験まで行ったことである.理 論的には,逐次近似画像再構成法は汎用性が高 いためコーンビームに拡張することは容易に できるが,提案手法では幾何学系に特化した前 処理フィルタ Q が使われているため,コーンビ ーム CT に用いられる代表的な画像再構成法の Feldkamp 法に基づき拡張した.また,TV の正 則化項は 3D に拡張したものを用いる.コーン ビームにおける変更箇所を以下にまとめる. [変更 1] システム行列 A はコーンビーム幾何 学系のものを用いて順投影 A と逆投影 A^Tの計 算を行う.

[変更 2] 前処理フィルタ *Q* は, Feldkamp 法と 同様に,各方向のコーンビーム投影 *g_u(X,Y)*に対 して,X線源軌道の接線方向(*X*方向)に式(11) の前処理フィルタを一行一行作用させる. [変更 3] TV は(*x*,*y*,*z*)の全ての方向に等方的に平 滑化を行う次式の 3D TV を用いる.

$$\|\vec{x}\|_{3\text{DT}V} = \sum_{i} \sum_{j} \sum_{k} \sqrt{(x_{ijk} - x_{i+1,j,k})^2 + (x_{ijk} - x_{i,j+1,k})^2 + (x_{ijk} - x_{i,j,k+1})^2}$$
(13)

3. 実験結果

1) シミュレーション実験

円軌道コーンビーム CT において低線量の状 況とスパースビューの状況を想定して、シミュ レーション実験を行った.実験には門脈相の腹 部 CT 実画像を用い,設定を表1にまとめて示 す. 比較する手法として, Chambolle-Pock アル ゴリズム,FISTA,OS-SART-TV 法を実装して, 収 束速度と画質評価指標(RMSE 値)の比較を行 った.紙面の制約から全ての実験結果は紹介で きないため、代表してスパースビューCT の実 験結果のみを示す.図1に,スパースビューCT の設定で、各手法の反復回数 10.30 回における 再構成画像を示す.図2に,評価関数値とRMSE 値の収束の様子を示す. Chambolle-Pock アルゴ リズムと FISTA は収束が非常に遅く, 提案手法 は OS-SART 法(64 サブセット)とほぼ同等に 高速に収束している.

夜」 ンミュレーンヨノ夫釈の	り設定
----------------	-----

X 線 源 軌 道	65.28(cm)
半径	
視野	直径 38.4(cm), 体軸方向 20(cm)
各投影デー	384×384×200 (画素)
タ画素数	
方向数	低線量 700(方向)
	スパースビュー 64(方向)
ノイズ	低線量 1.0×10 ⁵ (counts/bin)
	スパースビュー5.0×10 ⁶ (counts/bin)

2) キャノン 256 列検出器コーンビーム CT 装

置実機投影データへの適用

キャノン社が 2000 年頃に開発した 256 列検 出器を用いた汎用コーンビーム CT 装置(4 次 元 CT)実データへの適用を行った[7].投影デ ータの測定方向数は 900 方向であったが,300 方向に削減して画像再構成を行った.撮影対象 は腹部である.実画像には TV 正則化は向かな いという批判もあるため,TV 正則化に加え Nonlocal TV 正則化も実装した.図3に,反復回 数30回における提案手法(TV 及び Nonlocal TV) の再構成画像と FBP 法の再構成画像を示す.

4. まとめ

本研究では、筆者らが文献[5],[6]で提案した 主双対問題を解く反復法内部に FBP 法の構造 を組み込んで TV 正則化つき画像再構成問題を 解く枠組みである FBP 法組み込み型主双対反 復法に関する,更なる展開研究を行った.主要 な成果は以下の2つである.一つ目は, Chambolle-Pock アルゴリズムに前処理を導入し た数学的枠組みを用いて FBP 法組み込み型主 双対反復法の新導出を行い, 収束性が保証され る前処理フィルタ(とステップサイズ)の条件 を数学的に厳密に示した.二つ目は、コーンビ ーム CT の画像再構成に拡張して実装と評価実 験を行い高性能の逐次近似画像再構成法の一 つであることを示した.特に,現在主流の高速 化手法である OS 法は並列化が困難である欠点 があり, OS 法に代わる並列化が容易な高速化 手法として期待される.

謝辞

本研究は, JST-ERATO (JPMJER1403) 及び JST-CREST (JPMJCR1765) の助成を受けた.また, 本研究で使用したコーンビーム CT 実機投影デ ータは,「放射線医学総合研究所 4 次元 CT 装置 開発研究班」から提供を受けた.

利益相反の有無

なし.

文 献

- [1] E.Y.Sidky, J.H.Jorgensen, and X.Pan, Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm, Phys Med Biol, 57, 3065-3091, 2012
- [2] A.Chambolle and T.Pock "A first-order primaldual algorithm for convex problems with applications to imaging, J Math Imaging Vis, 40, 120-145, 2011
- [3] Q.Xu, D.Yang, J.Tan, A.Sawatzky, and M.A. Anastasioa, Accelerated fast iterative shrinkage thresholding algorithms for sparsity-regularized cone-beam CT image reconstruction, Med Phys, 43, 1849-1872, 2016
- [4] A.Beck and M.Teboulle, A Fast iterative shrinkage-thresholding algorithm for linear inverse problems, SIAM J Imaging Sci, 2, 183-

202, 2009

- [5] H.Kudo, F.Yamazaki, K.Takaki, and T.Nemoto, A very fast iterative algorithm for TVregularized image reconstruction with applications to low-dose and few-view CT, Proceedings of SPIE 2016 Optics + Photonics (Development in X-ray Tomography), Paper No. 9967-37, 2016
- [6] T.Wang, H.Kudo, F.Yamazaki, and H.Liu, A fast regularized iterative algorithm for fan-beam CT reconstruction, Phys Med Biol, 64, Paper No. 145006, 2019
- M.Endo, T.Tsunoo, S.Kandatsu, S.Tanada,
 H.Aradate, and Y.Saito, Four-dimensional computed tomography (4D CT) Concepts and preliminary development, Radiat Med, 21, pp.17-22, 2003



図2 投影方向数64のスパースビューCTにおける代表的な反復法の収束速度比較 (上段:評価関数値,下段:RMSE値,両方とも縦軸は対数スケールで表示)



図1 投影方向数64のスパースビューCT における代表的な反復法の再構成画像比較 (*z*=6.4(cm)の transaxial スライスと *y*=0の coronal スライスを表示)



図3 キャノン 256 列検出器コーンビーム CT 実機データの提案手法による再構成画像(投影データ 方向数 300 方向, *z*=-2.2cm の transaxial スライスと *x*=0 の sagittal スライスを表示)

Primal-Dual FBP-Embedded Iterative Method for Compressed Sensing

CT Reconstruction and Its Extension to Cone-Beam CT

Hiroyuki KUDO^{*1}, Ting WANG^{*1}

*1 University of Tsukuba

Image reconstruction in sparse-view CT and low-dose CT is performed by Compressed Sensing (CS), in which a sum of least-squares data fitting term and Total Variation (TV) regularization term is minimized by using an iterative method such as Chambolle-Pock algorithm and FISTA. However, convergence of these iterative methods is very slow leading to a large computational cost. To overcome this drawback, an empirical approach called Ordered-Subsets (OS) technique is often used, but the OS technique is not guaranteed to converge to an exact minimizer of cost function and does not match to parallel computations. We have already developed an alternative mathematical framework to accelerate the convergence named as FBP embedded primal-dual iterative method for 2-D parallel-beam and fan-beam CT (Kudo et al, Proc.SPIE, 2016, Wang et al. Phys.Med.Biol., 2019). An iterative method derived using this framework converges to an exact minimizer of cost function, and it matches to parallel computations thanks to its parallerizable structure. In this paper, we provides a new mathematical derivation of this method based on a well-known mathematical framework called Chambolle-Pock algorithm modified by including a preconditioning. The key of derivation is the fact that FBP embedded primal-dual method can be interpreted as a proximal point method to solve a saddle-point problem, and we have succeeded in obtaining conditions of a preconditioning FBP filter to guarantee exact convergence. This theoretical result solves the convergence issue of FBP embedded primal-dual method, which was not clear in our previous work. Furthermore, we have extended FBP embedded iterative method to cone-beam CT with a circular x-ray source trajectory. We demonstrated performances of the proposed method by simulation studies for image reconstruction in sparse-view CT and low-dose CT. Furthermore, we have applied the proposed method to real data acquired with Canon generalpurpose cone-beam CT scanner equipped with a 256-row detector.

Key words: CT, Image Reconstruction, Compressed Sensing, Total Variation, Cone-Beam

生体軟組織の屈折コントラスト CT に適用するための

機械学習によるリングアーチファクト除去法の開発

黄 卓然*1 砂口 尚輝*1 島雄 大介*2 市原 周*3

湯浅 哲也*4 安藤 正海*5

要旨

我々は屈折コントラスト CT (RCT) のリングアーチファクト (RA) を効果的に除去するために機械学習 を用いた手法の開発を進めている.この手法は,RCT の投影像が微分像であることに着目し,機械学習に よるネットワークの最適化の際に,その疎の性質を取り入れているところに特徴がある.以前,我々は512 ×512 画素のラット胎児の CT にこの手法を適用し優れた RA 除去効果を得た.一方で,本手法を実用する には,少なくとも 2000×2000 画素以上のサイズを持った画像に適用できることが必須であるが,これまで の開発環境では制約があり試すことが困難あった.

本研究では,独自に開発した RCT の撮像を模擬したシミュレーターを用いて 2000×2000 画素のシミュレーションデータを生成し,本手法内の条件付き敵対的生成ネットワークで学習させる. そのネットワークを用いて,様々な生体軟組織の RCT に適用した結果,高い RA 除去効果が得られた.

キーワード: 屈折コントラスト X線 CT, リングアーチファクト, 機械学習, 敵対的生成ネットワーク

1. はじめに

屈折コントラスト X線 CT (Refraction-contrast X-ray computed tomography, RCT) は、物体内の 屈折率分布を再構成する撮像法である[1]. 従来 の吸収コントラストに基づく CT(Absorptioncontrast X-ray computed tomography, ACT)と異な る点は、CT の再構成に必要な投影が線減弱係 数の線積分ではなく、物体を透過した X線の屈

*1 名古屋大学大学院医学系研究科

〔〒461-8673 名古屋市東区大幸南 1-1-20〕

e-mail: huang.zhuo.ran@j.mbox.nagoya-u. ac.jp

- *2 北海道科学大学保健医療学部
- *3 名古屋医療センター病理診断科
- *4 山形大学大学院理工学研究科
- *5 高エネルギー加速器研究機構

折角を表すところである. この屈折角は 10⁻⁷ ラ ジアンと非常に微小な量であり,本手法では Laue 型 Si 単結晶でできた Laue 型角度分析板 (Laue angle analyzer, LAA)の回折を利用するこ とで,角度解析が可能となる.

本手法の課題は,結晶表面に傷やゴミがある とき,屈折角に誤差が生じるため,投影像上の ある画素領域に常にアーチファクトが乗って しまうことである.これは,CTを再構成した際 に,リング状のアーチファクト (RA) が発生す る要因となる.

RA の除去法は、ACT の RA を除去する目的 で多数提案されている.例えば、Prell et al [2]が 提案した画像フィルタ法、Liang et al [3]が提案 した Relative total variation に基づいた反復法、 敵対的生成ネットワーク[4]を用いた機械学習 法が存在する.これらの中で、機械学習法は最 近提案された強力な手法であり、元の構造に影 響を与えることなく, RA のみを効果的に除去 できる. RCT の RA の除去についても, 機械学 習法は期待できる方式であるが, RCT は CT を リング状の構造をファントムの中心に追加した.各ファントムは,我々の先行研究[6]で開発したX線光線追跡に基づく屈折X線CTシミュ



図 1 cGAN 学習のフローチャート

再構成する際に,投影の積分処理が必要である ことから屈折角の誤差が投影全体に広がって しまっているという状態を考慮しなくてはな らない.

本研究では, RCT の物理的性質から導いた損 失関数を取り入れた条件付き生成的敵対ネッ トワーク (Conditional generative adversarial network, cGAN)[5]を RA の除去に用いる. 学習 には,独自に開発した RCT の撮像を模擬したシ ミュレーターから取得した 2000×2000 画素の シミュレーションデータを利用する. この RA 除去法が実用できるか評価するために, 2000× 2000 画素以上のサイズを持った様々な生体軟 組織の RCT 像に適用し, RA 除去効果を検証す る.

2. 機械学習を用いた RA 除去手法

cGAN を用いた RA 除去の学習手法を図1に 示す.まず,屈折率δ,半径,中心が異なるラ ンダムな円を含む数値ファントムを 1000 セッ ト用意した.生体サンプルはリング状構造があ る場合を考慮して,半径がランダムに変化する

レーターを用いて測定し、投影回数は0~180度 の 720 回とした. シミュレーションで得られた サイノグラムは、図1の左上の例に示すように、 δの線積分の1階微分であり、水平方向と垂直 方向はそれぞれ検出器素子の配列と試料への X 線の入射角を示す.一般的には、サイノグラム を水平方向に積分した後,フィルタ補正逆投影 法(Filter back projection, FBP)を適用して RCT 画像を再構成される.しかし、本手法では、後 述する RCT の正弦/余弦微分成分を生かすため, 積分ではなく微分投影の正弦と余弦成分を生 成し、微分投影像とリングアーチファクトの原 因となるノイズを含んだ微分投影像にも同様 の処理を行って, RCT の正弦成分と余弦成分を 再構成する. 微分投影の正弦成分と微分成分は 以下の式で生成する.

$$P_{\cos}(r,\theta) = P(r,\theta) \times \cos\left(\frac{\theta\pi}{N}\right) (0 \le \theta \le N) \quad (1)$$

$$P_{sin}(r,\theta) = P(r,\theta) \times sin\left(\frac{\theta\pi}{N}\right) (0 \le \theta \le N)$$
 (2)

ここで、 $P_{sin}(r,\theta)$ と $P_{cos}(r,\theta)$ は微分投影の正弦と 余弦成分. $P(r,\theta)$ は微分投影、N は投影数である. 得られた RCT の正弦成分や余弦成分と,それ に対応する RA を含む画像は,2つの cGAN ネ ットワークに入力する. cGAN の損失関数とし ては,生成する画像と正解画像の L2 ノルムに 加えて,正弦成分または余弦成分の画素値の和 が理論的にゼロであることを利用し,L1 ノルム も追加される.これにより,画素値が過度に平 坦化されることを防ぎ,画像の中央部など RA が密集している領域でのノイズ除去効果を高 めることができる.

三角関数の0に近い部分があるため, RCT の 余弦成分と正弦成分の一部の情報が欠落され る. 生成された2つの画像の投影画像をラドン 変換で得た後,それぞれの情報の欠落部分を補 うために,投影像を加算し, RA を除去した一 次微分投影像が得られる.最後に一般的の RCT 再構成法で, RA を除去した通常の RCT, つま りδ画像に変換することができる.

3. 実験

本研究を利用するネットワークは Python3.6 と TensorFlow に実装された. 学習用の GPU は Nvidia TITAN V であり, RAM は 256GB である. 学習のハイパーパラメータでは, learning rate = 0.001, iteration = 60,000 とした.

本手法の有効性を評価するために、学習に使 用していない 100 枚のシミュレーション画像を 用いて、提案法、画像フィルタ法[2]、反復法[3] の結果を比較した.評価方法としては、各手法 を適用した画像と正解画像との間の平均平方 根誤差(RMSE)と構造的類似度(SSIM)を用 いた.また、実際の高空間分解能を持つ生体試 料に適用し、RA 除去効果を検証する.

4. 結果

シミュレーション画像の実験結果を図2に示 す.シミュレーションによる RMSE と SSIM は, 画像フィルタ法では 0.3882 と 0.9231, 反復法で は 0.3893 と 0.9209, 提案法では 0.0917 と 0.9972 であった. どちらの指標においても, 提案手法 が最も優れていた.

これらの手法の RA 除去結果の画像を観察す



(a) 正解画像



(b) 入力画像



(c) 画像フィルタ法



(d) 反複法



(e) 提案法

図 2 2048×2048 のシミュレーションデー タの結果



(a) 入力画像

(b) 画像フィルタ法



(c) 反復法

(d) 提案法

図 3 2048×2048の生体試料の結果

ると、図3に示すように、画像フィルタ法では RAの残存量が多く、リング状の構造を誤って 除去してしまうことがある.反復法は画像フィ ルタ法よりも優れているが、RAを完全に除去 することはできない.提案された方法は、RAを 除去することに成功し、RAではないリング状 の構造がはっきりと残っている.しかし、唯一 の欠点は、ネットワークの学習能力が高すぎて、 処理中に画像が平坦になったり、低コントラス トの構造が部分的に失われたりする.ネットワ ークのパラメータを減らし、学習回数を増やす ことで、この問題を改善することができる.

5. まとめ

本研究では, RCT の正弦成分と余弦成分の画 素値の総和が理論的にゼロとなる RCT の物理 的特性を利用し,機械学習に基づくリングアー チファクト除去法を提案した.提案手法は,他 の手法に比べて RA の除去効果が高いことがわ かった.また,2000×2000以上の生体試料での 検証に成功し,高空間分解能の RCT でも同じよ うに優れた RA 処理能力を持つ. 今後,様々な 生体試料に本手法を適用していく予定である.

謝辞

本研究の試料を準備していただいた名古屋 大学大学院医学系研究科榎本篤教授に感謝申 し上げます.本研究は JSPS 科研費(16K01369, 15H01129,26286079,18K13765),平成29年度国 立病院機構共同臨床研究 H29-NHO(癌般)-01, H29-NHO(多共)-02,知の拠点あいち重点プロジ ェクト第3期および KEK 放射光実験課題 (2016G0625,2019G598,2020G565,2020G583) の支援を受けています.

利益相反の有無

なし.

文 献

- [1] Sunaguchi N, Yuasa T, Huo Q et al: X ray refraction contrast computed tomography images using dark field imaging optics Applied Physics Letters 97, 2010
- [2] Prell D, Kyriakou Y, Kalender W A: Comparison of ring artifact correction methods for flat-detector CT. Physics in Medicine & Biology 54: 3881-3895, 2009
- [3] Liang X, Zhang Z, Niu T et al: Iterative

image-domain ring artifact removal in cone-beam CT. Physics in Medicine & Biology 62: 5276-5292, 2017

- [4] Goodfellow I J, Pouget Abadie J, Mirza M et al: Generative adversarial nets. In Proceedings of NIPS, 2672-2680, 2014
- [5] Isola P, Zhu J Y, Zhou T et al: Image-to-Image translation with Conditional

Adversarial Networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR): 1125-1134, 2017

[6] Sunaguchi N, Yuasa T, Sun F et al: Limited view reconstruction for differential phasecontrast computed tomography. Optics Express 23: 9717-9729, 2015

Development of a ring artifact removal method using machine learning

for high-resolution refraction-contrast CT of biological soft tissue

Zhuoran HUANG^{*1}, Naoki SUNAGUCHI^{*1}, Daisuke SHIMAO^{*2}, Shu ICHIHARA^{*3},

Tetsuya YUASA^{*4}, Masami ANDO^{*5}

- *1 Nagoya University
- *2 Hokkaido University of Science
- *3 Nagoya Medical Center
- *4 Yamagata University
- *5 High Energy Accelerator Research Organization

We are developing a machine learning-based method to effectively eliminate ring artifacts (RA) in refraction-contrast computed tomography (RCT). This method is characterized by the fact that the projected image of RCT is a differential image, and the sparse nature of the image is incorporated in the optimization of the network by machine learning. Previously, we applied this method to a 512x512 pixel CT of a rat fetus and obtained excellent RA removal results. On the other hand, for the practical use of this method, it is necessary to be able to apply it to images with a size of at least 2000×2000 , but we could not try it before due to the limitation of the development environment. In this study, we generate 2000×2000 simulated data using our own simulator that simulates RCT imaging, and train it in the conditional generative adversarial network in our method. The network was applied to RCTs of various soft tissues in the body, and the results showed high RA removal efficiency.

Key words: refraction-contrast CT, ring artifact, machine learning, generative adversarial network

Image Quality Improvement for Chest Four-Dimensional Cone-Beam Computed Tomography by Cycle-Generative Adversarial Network

Keisuke USUI^{*1, *2}, Koichi OGAWA^{*3}, Masami GOTO^{*1}, Yasuaki SAKANO^{*1}, Shinsuke KYOGOKU^{*1}, Hiroyuki DAIDA^{*1}

Abstract

Four-dimensional cone-beam computed tomography (4D-CBCT) can visualize moving tumors, thus the 4D-CBCTbased adaptive radiation therapy (ART) may improve the quality of radiation therapy. The aim of this study is to improve the quality of 4D-CBCT images using cycle-generative adversarial network (Cycle-GAN) and evaluate these images by a quantitative index. In this study, unpaired thoracic 4D-CBCT images and four-dimensional multislice computed tomography (4D-MSCT) images in 20 patients were used for training, and synthesis of 4D-CBCT (sCT) images with improved quality was tested in another 10 patients. The mean error (ME) and mean absolute errors (MAE) were calculated to assess CT number deviation, and peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) were used to evaluate image similarity. The sCT image generated by our Cycle-GAN model effectively reduced artifacts on 4D-CBCT image. The ME and MAE were 46.5 and 61.9 in lung regions, whereas soft tissue and bone regions insufficiently restored CT number. Results of the SSIM and PSNR were significantly improved in the sCT image. The proposed Cycle-GAN method generates sCT images with a quality close to 4D-MSCT image, particularly in the lung region; however, anatomical regions with soft tissue and bone still require further improvement.

Keywords : Cycle-GAN, 4D-CBCT, radiotherapy, deep learning

1. Introduction

Accurate radiotherapy can be accomplished with high precision irradiation position using a cone-beam computed tomography (CBCT) attached with the linear accelerator [1, 2]. A four-dimensional CBCT (4D-CBCT) can visualize moving tumors as a series of 3D-image. Therefore, this 4D image can be used to treat a non-small cell lung cancer under free breath condition [3]. On the other hands, adaptive radiation therapy (ART) is one of the newest technologies in which the quality of radiation therapy is improved by modifying an initial treatment plan with morphological changes during fractionated treatment courses [4, 5]. This treatment technique can change the irradiation area or x-ray intensity

e-mail: k-usui@juntendo.ac.jp

^{*1} Department of Radiological Technology, Faculty of Health Science, Juntendo University [Bunkyo-ku 2-1-1, Tokyo 113-8421, Japan]

^{*2} Department of Radiation Oncology, Faculty of Medicine, Juntendo University

^{*3} Department of Applied Informatics, Faculty of Science and Engineering, Hosei University

in each treatment day, and has shown clinical benefits [6, 7]. Conversely, a conventional linear accelerator-based ART is a major challenge due to insufficient quality of CBCT image affected by a detector scatter, patient-specific scatter, image lag, and beam hardening [8]. On the other hands, a respiration-correlated 4D-CBCT is achieved with cone-beam projection subgroups into different respiratory phases. However, insufficient projections each phase bin cause severe imaging artifacts, resulting in deterioration of irradiation position accuracy and preventing the creation of an accurate ART plan using this 4D-CBCT images [9, 10]. Many CBCT image correction methods have existed such as a scatterblocker, image processing algorithms, and deformable image registration (DIR) [11-13]. However, the sparseness of projection data is a dominant factor associated with 4D-CBCT image quality deterioration, and the DIR deformation accuracy also deteriorated. In addition, scattering photons and respiratory motion also cause strong artifacts on reconstructed images. Therefore, improving the quality of 4D-CBCT images is still difficult using these methods. Recently, with the development of deep learning in the medical imaging field, several approaches have been proposed in reducing artifacts and generating synchronized multimodal images [14-17]. Kida et al. applied the convolutional neural network (CNN) to improve the quality of CBCT images [14]. Additionally, Jiang et al. improved the undersampled CBCT images in over-smoothed edge regions with a total variation regularized by the CNN-based method [15]. However, these studies implement a deep learning architecture with supervised data, and a small difference in the teacher image causes a conversion process error. Furthermore, 4D-CBCT cannot be combined with supervised data, since respiratory movements cannot be exactly similar between CBCT and four-dimensional multislice CT (4D-MSCT) images. Therefore, an accurate image correction cannot be made through a supervised learning with 4D-CBCT images.

In this study, we implement a cycle-generative adversarial network (Cycle-GAN) framework to improve the quality of 4D-CBCT images and evaluate the quality by quantitative indices. The Cycle-GAN model enforces an inverse transformation and achieves high accurate consistency when underlying structures are similar, even mapping in nonlinear domains [18, 19]. The Cycle-GAN method is expected to optimally improve the quality of 4D-CBCT images, which cannot create accurate supervised data, and no previous study has reported the correction of projection data sparseness in the 4D-CBCT. We investigated the quality of synthetic 4D-CBCT (sCT) images created using the Cycle-GAN method in which the mean CT number error (ME) and mean absolute CT number error (MAE) in each organ region were used. Moreover, the peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) were used to evaluate the image quality and similarity of the entire image.

2. Materials and Methods

1) Data objects

In this study, thoracic 4D CT images acquired with a CBCT and MSCT were used as training and evaluation data. These CT images were obtained from a publicly available dataset (the Cancer Imaging Archive), an open-access information source created by the US National Cancer Institute [20]. All 4D-CBCT datasets were acquired using the onboard imager equipped with a kilo-voltage x-ray source and flat panel detector (Varian Medical systems, Palo Alto, CA, USA). Moreover, 4D-MSCT data were acquired using a multi-detector CT with a helical scanning. The thickness of each CT image was 3 mm, matrix size 512 × 512 pixels, and field-of-view (FOV) of 50 × 50 cm in 4D-MSCT and

 45×45 cm in 4D-CBCT. These 4D images were composed of 10 phases divided by one respiratory motion. Only the first-phase image (0% phase) was used in each 4D image for training and evaluation of image quality. The 4D-CBCT image was acquired with the center position of the lung cancer region, including the upper, middle, and lower lung areas. Moreover, the 4D-CBCT image was acquired during the course of radiation treatment, and the 4D-MSCT was acquired at the treatment planning before radiation therapy.

2) Image synthesis with Cycle-GAN

The sCT image was generated based on 4D-CBCT using the 2D Cycle-GAN method. Training datasets were 50 slice images per one patient with 20 patients, in a total of 1,000 images in 4D-CBCT and 4D-MSCT. These 4D images were not completely paired because of non-simultaneous data acquisition. Therefore, organ positions are possibly inconsistent due to respiration. Before feeding images into the model, each CBCT datum was first resampled to match each resolution and a size of 256 × 256 to avoid learning from unnecessary image information. The Cycle-GAN model has four CNNs and relies on two subnetworks, one generator and one discriminator, which work in opposite function. In two training datasets, one generator converts 4D-CBCT images into 4D-MSCT images, whereas the second generator converts 4D-MSCT images from 4D-CBCT and 4D-MSCT. To train the Cycle-GAN, all four networks were trained simultaneously to maximize the performance (generate data that "fools" the discriminators and distinguish between real and generated data). As these networks are pitted against each other, each capability improves, resulting in accurate 4D-CBCT image generation [21, 22]. Figure 1 shows the schematic flowchart of the Cycle-GAN model, and Table 1 shows structural details of the generator and discriminator and the gradient optimization method. The total loss function in this Cycle-GAN training is given in Eq. (1).

$$Loss = (Adversarial \ Loss) + \lambda \times (Cycle \ Consistency \ Loss)$$
(1)

Parameter λ controls the relative importance of two types of losses; therefore, we set this value 10 in this study. The adversarial loss is the loss function of discriminator and generator using the sum of square difference calculated using Eqs. (2) and (3) and minimize between expected and predicted output for 4D-MSCT and 4D-CBCT images.

$$loss Discriminator = (D(X) - 1)^2 + D(X')^2$$
⁽²⁾

$$loss Generator = (D(X') - 1)^2$$
(3)

where X is a real input, X' a generated input, D(X) the result of the discriminator network applied to X, and D(X') the result of the discriminator network applied to the generated output X. The cycle consistency loss attempts transformation of an image from domain X to domain Y using generator X, and then transforming the result back to domain X using generator Y provides a good approximation of the original image. Therefore, the cycle consistency loss can be expressed as

$$loss Cycle = |Y - G_Y(G_X(Y))| + |X - G_X(G_Y(X))|$$
(4)

where X and Y are images from two domains (4D-CBCT and 4D-MSCT). $G_X(Y) = X'$ generates an image in domain X from a real image in domain Y. The cycle consistency loss is defined as the difference between original and generated images reconstructed back into the original domain. As the training progresses, the reconstructed image closely matches with the original images. In this study, we generated the sCT images from 4D-CBCT images using the

trained Cycle-GAN model.

We conducted experiments with a personal computer equipped with two GPUs (Quadro RTX 5000, NVIDIA Corporation) and a CPU (Intel Xeon Sliver 4210R) with 96 GB memory, and implemented our algorithm with MATLAB 2020b (Math Works Inc., Natick, MA, USA).

3) Image quality evaluation

Synthesis images were constructed using 10 patients in each of 4D-CBCT and 4D-MSCT images, which are different from the training data. To evaluate the quality of sCT images by generating the Cycle-GAN model and to compare them with the original 4D-CBCT image, CT number differences were analyzed and compared with those of the 4D-MSCT image. We set region of interests (ROIs) in lung, soft tissue, and bone regions and measured the ME and MAE as follows:

$$ME(X,Y) = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} (X(i,j) - Y(i,j))$$
(5)

$$MAE(X,Y) = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} |X(i,j) - Y(i,j)|$$
(6)

where *M* and *N* indicate the width and height of pixels within a ROI, X(i, j) the CT number of i-th and j-th pixels in a sCT image or original 4D-CBCT image, and Y(i, j) the CT number of i-th and j-th pixels in a 4D-MSCTimage. Sizes of ROIs were 35×35 , 25×25 , and 15×15 pixels in the lungs, soft tissues, and bones, respectively. ME and MAE differences between a sCT image and the original 4D-CBCT image were considered statistically significant when p < 0.05 (two-tailed t-test). Moreover, as a comprehensive image evaluation of the whole image quality, PSNR and SSIM were calculated for the sCT image and original 4D-CBCT image based on the 4D-MSCT image. The SSIM of image *X* and *Y* was defined as follows:

$$SSIM(X, Y) = \frac{(2\mu_X\mu_Y + C_1)(2\sigma_{XY} + C_2)}{(\mu_X^2 + \mu_Y^2 + C_1)(\sigma_X^2 + \sigma_Y^2 + C_2)}$$
(7)

where μ is mean, σ variance, C terms are regularization constants. Furthermore, the PSNR was defined as follows:

$$PSNR = 10 \log_{10} \frac{\max[X(i,j)]^2}{MSE}$$
$$MSE = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} (X(i,j) - Y(i,j))^2,$$
(8)

The PSNR is defined by the maximum value in an input image X(i, j) divided by the mean squared error between image X (sCT or original 4D-CBCT) and image Y (4D-MSCT). In addition, M and N indicate the width and height of images, respectively. To reduce the geometric mismatch between sCT image and 4D-MSCT image, sCT images were performed linear-registered to approach the corresponding pixel value of 4D-MSCT image. SSIM and PSNR differences between sCT image and original 4D-CBCT image were considered statistically significant when p < 0.05 (two-tailed ttest).

3. Results

1) Image correction performance

Figure 2 shows representative axial slice images of three patients in the corresponding 4D-MSCT image (left), the original 4D-CBCT image (center), and the sCT image (right) in the upper, middle, and lower thoracic regions. Figure 3 shows CT number distribution and profile of each image in thoracic areas. For the original 4D-CBCT, significant image

degradation was caused by the limited number of projection data. In contrast, we observed that the sCT image generated by our Cycle-GAN model effectively reduced the artifacts in all thoracic images. The CT number distribution in the lung region with a number near -500 was significantly restored in sCT image in all thoracic areas. However, the improvement of soft tissue in CT number near 0 was slight, and no appropriate distribution agreement was obtained in the bone region at CT number of 500. Moreover, for three thoracic images, profiles of CT number in the 4D-CBCT image were not match, in contrast, profiles of sCT image were similar to that of the 4D-MSCT image in lung region.

2) Image quality analysis

Table 2 shows validated mean results for the ME and MAE in 10 patients for ROIs of the lung, soft tissue, and bone regions. After the sCT image conversion, the ME and MAE were close to 0 Hounsfield Unit (HU) in lung and soft tissue regions, which were statistically significant in the lung (ME and MAE) and soft tissue (ME). In contrast, in the bone region, the error reduction of the CT number deviation was not observed by generating the sCT image. Table 3 shows results of the comprehensive image evaluation by the SSIM and PSNR. These results were shown in the mean improvement rate in 10 patients and calculated as ratio of sCT image to 4D-CBCT image for SSIM and PSNR values. SSIM and PSNR results were significantly improved in the sCT image.

4. Discussion

In this study, we generated the sCT image with improved quality of 4D-CBCT image with Cycle-GAN method using unpaired 4D-CBCT and 4D-MSCT images. Moreover, the quality of sCT image was evaluated by quantitative image quality index as compared with the original 4D-CBCT image. In order to perform accurate image-guided irradiation and to realize 4D-CBCT-based ART, sufficient image quality is necessary by restoring CT number.

The sCT image generated by the Cycle-GAN was close to the ideal 4D-MSCT image and was able to significantly reduce artifacts in the original 4D-CBCT image. In the results shown in Fig. 2, the 4D-CBCT image remarkably had artifacts, which were caused by sparse angular sampling, respiratory motion and photon scatterings. In contrast, sCT images significantly reduced the strike artifact and image noise; therefore, visibility was improved in the whole image, which was also represented at each part of the upper, middle, and lower thoracic regions, because the training data were composed 4D-CBCT and -MSCT images that included all lung areas. Improving the visibility of 4D-CBCT images enhances the tumor and surrounding organ visibility in radiotherapy, increasing the accuracy of target localization. In previous studies, the Cycle-GAN methods improved the quality of CBCT images, and our results were in agreement with literatures [19, 22]. The image quality in the lung region in sCT image was remarkably restored as compared to soft tissue and bones regions. Conversely, the improvement in the soft tissue region was small, and the bone region had not been sufficiently restored. Based on the results of the histogram and profile analysis in Fig. 3, CT number distribution in the sCT image can be close to that of the 4D-MSCT image. The CT number distribution also shows remarkable improvement, especially in the lung region in each thoracic area. Moreover, as shown in the results of CT number deviation in Table 2, the pixel value of sCT image was close to that of 4D-MSCT image, improving a significant difference from 4D-CBCT image, especially in the lung regions. The problem is whether the algorithm can essentially identify the distinction of the air region outside the body for low CT number such as the lungs; however, our results

showed good composite images without loss of microvascular structure in the lung region. ME and MAE values of the lung region were 46.5 and 61.9 HU, respectively, i.e., ME was estimated to be large because of biased deviation of pixel values due to the influence of respiratory motion in the 4D-CBCT image. However, these results were similar range presented in previous studies [19]. Moreover, the identified larger standard deviation of up to 100 HU for the lung region might be caused by the respiratory motion inherent in 4D images. In the results of soft tissue, the CT number tended to be close to that of the 4D-MSCT image and the ME had a significant improvement with sCT image. However, CT number deviation was not decreased using the results in the bone region. Since the CT number deviation of the bone region in the original 4D-CBCT image was smaller than that in the lung region, the improvement effect on the sCT image could not be obtained. A direct comparison remains difficult as referenced studies differed from the presented work with the number of training data, image acquisition parameters, and architecture of the GAN model; however, results of MA and MAE in the bone region were similar to those of previous studies [18]. Accurate dose calculation relies on the accurate HU measurement and conversion to electron density. CT number restoration to 4D-MSCT improves the dose calculation accuracy of radiotherapy planning. Therefore, sCT image is expected to have sufficient image quality for accurate dose calculation in the thorax areas. Because, most of the thorax areas is the lung and soft tissue regions, we considered that the deviation of the CT number of the bone region could not be caused a large dose calculation error. Furthermore, CT number restoration improves the warping accuracy to daily 4D-CBCT image in the organ contour region and dose distribution. As accurate warping technique is essential for realizing the online ART, generation of synthetic CT image by Cycle-GAN enhances the ART feasibility using a 4D-CBCT. In the results of Table 3, SSIM and PSNR values were significantly increased in sCT image as compared with original 4D-CBCT image by 59.7 times in SSIM and 24.2 times in PSNR. In our study, as restoration in the lung region was remarkably good, the evaluation of the whole image was greatly improved, demonstrating that the entire sCT image could resemble the 4D-MSCT image.

The 4D-CBCT image cannot train with paired supervised data, as it depends on the reproducibility of respiratory movements. The Cycle-GAN method can be used in non-paired training data and is quite advantageous for this synthetic image generation, and no other study has been conducted to restore the 4D image with a GAN to date. In our method, image artifacts and low-quality 4D-CBCT images could be corrected by the sCT image generation, which resulted in a reliable assignment of CT numbers. To realize 4D-CBCT-based ART, reducing streaking and motion artifacts to reveal material boundaries is an important step to create accurate treatment planning on 4D-CBCT images. However, effects of warping accuracy with organ contouring and dose distribution were not evaluated in this work. Moreover, dose calculation accuracy using sCT images should be demonstrated. Therefore, implementation of a larger FOV and geometric mismatch at body boundaries are needed to accomplish the Cycle-GAN process in future works. In a previous study, the rigid registration in training data could effectively prevent the geometric mismatch in generation images [18, 19]. However, respiratory motion affected the training data and was not a perfect pair; thus, how much pre-registration method can improve the geometric distortion and remains a future task.

5. Conclusions

The proposed Cycle-GAN method improves the quality of 4D-CBCT images, making it comparable to that of 4D-MSCT images. CT numbers of the sCT image were significantly restored in the lung region; however, obtaining imaging in soft tissue and bone anatomical regions still requires special caution.

Figures



Fig. 1 Overview of flowchart in cycle-generative adversarial network (Cycle-GAN) framework. The training model consisted of two generators and two discriminators. To train the Cycle-GAN, the whole network's performance is enhanced through acting bidirectionally with each other. The sCT image is generated by a network that maps from a source domain (4D-CBCT) to target domain (4D-MSCT).



Fig. 2 4D-MSCT, 4D-CBCT, and synthetic 4D-CBCT (sCT) images of three patients in the corresponding axial slice images. (a) - (c) represent each thoracic region in the upper, middle, and lower areas. All images are shown with the same window width and level.



Fig. 3 Histogram of CT numbers (left column) and line profiles along the white arrow (right column) in each image. (a) - (c) show each thoracic region corresponding to Fig. 2.

Tables

Table 1 Structure of the generator and discriminator and gradient optimization conditions.

Optimizer	Adam gradient descent method
Minibatch size	1
Initial learning rate	0.0002
Epochs	100
Encoder depth	Generator: 3, Discriminator: 4
Convolution filter	Generator: 32, Discriminator: 128
Residual block	6

 Table 2 Mean error (ME) and mean absolute error (MAE) in CT number of three organ sites for lung, soft tissue, and bone regions. These values are calculated with reference of the CT number in 4D-MSCT and are shown in terms of mean ± standard deviation within 10 patients for each organ site. * indicates p < 0.05, showing a significant difference with respect to the results between 4D-CBCT and sCT.</th>

Body site	Image	ME (HU)	MAE (HU)
Ŧ	4D-CBCT	455.5 ±117.9	455.5 ± 117.9
Lung	sCT	46.5* ±144.5	61.9* ±137.9
	4D-CBCT	87.9 ±124.6	120.0 ± 90.0
Soft tissue	sCT	$-34.8^* \pm 109.7$	$74.9 \hspace{0.2cm} \pm \hspace{0.2cm} 84.5$
D	4D-CBCT	17.9 ± 185.2	147.4 ± 102.5
Bone	sCT	-40.9 ± 131.8	99.4 ± 90.8

Table 3 Improvement rate of the structural similarity index (SSIM) and peak signal-to-noise ratio (PSNR). Thesevalues are calculated as ratio of sCT image to 4D-CBCT image for SSIM and PSNR values, and are shown as mean \pm standard deviation within 10 patients. * indicates p < 0.05, showing a significant difference with respect to the results</td>

of between 4I	D-CBCT and	d sCT.
---------------	------------	--------

Improvement rate of SSIM (%)	Improvement rate of PSNR (%)
59.7* ± 26.4	24.2* ±13.8

Competing interests

The authors declare to have no affiliations with or involvement in any organization or entity with any financial interest in the subject matter or materials discussed in this manuscript.

References

- [1] Barker JL Jr, Garden AS, Ang KK et al: Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system. Int J Radiat Oncol Biol Phys 59: 960-970, 2004
- [2] Hong TS, Tomé WA, Chappell RJ et al: The impact of daily setup variations on head-and-neck intensity modulated radiation therapy. Int J Radiat Oncol Biol Phys 61: 779-788, 2005
- [3] Purdie TG, Moseley DJ, Bissonnette J et al: Respiration correlated cone-beam computed tomography and 4DCT for evaluating target motion in stereotactic lung radiation therapy. Acta Oncol 45: 915-922, 2006
- [4] Qin A, Sun Y, Liang J et al: Evaluation of online/offline image guidance/adaptation approaches for prostate cancer radiation therapy. Int J Radiat Oncol Biol Phys 91: 1026-1033, 2015
- [5] Moller DS, Holt MI, Alber M et al: Adaptive radiotherapy for advanced lung cancer ensures target coverage and decreases lung dose. Radiother Oncol 121: 32-38, 2016
- [6] Yoo S, Yin FF: Dosimetric feasibility of cone-beam CT-based treatment planning compared to CT-based treatment planning. Int J Radiat Oncol Biol Phys 66: 1553-1561, 2006
- [7] Acharya S, Fischer-Valuck BW, Kashani R et al: Online magnetic resonance image guided adaptive radiation therapy: first clinical applications. Intl J Radiat Oncol Biol Phys 94: 394-403, 2016
- [8] Schulze R, Heil U, Grob D et al: Artefacts in CBCT: a review. Dentomaxillofac Radiol 40: 265-273, 2011
- [9] Lu J, Guerrero TM, Munro P et al: Four dimensional cone beam CT with adaptive gantry rotation and adaptive data sampling. Med Phys 34: 3520-3529, 2007
- [10] Zhang Y, Huang X, Wang J: Advanced 4-dimensional cone-beam computed tomography reconstruction by combining motion estimation, motion compensated reconstruction, biomechanical modeling and deep learning. Vis Comput Ind Biomed Art 2: 23, 2019
- [11] Stankovic U, Ploeger LS, van Herk M et al: Optimal combination of anti-scatter grids and software correction for CBCT imaging. Med Phys 44: 4437-4451, 2017
- [12] Bootsma G, Verhaegen F, Jaffray D: Efficient scatter distribution estimation and correction in CBCT using concurrent Monte Carlo fitting. Med Phys 42: 54-68, 2015
- [13] Disher B, Hajdok G, Wang A et al: Correction for 'artificial' electron disequilibrium due to cone-beam CT density errors: implications for online adaptive stereotactic body radiation therapy of lung. Phys Med Biol 58: 4157-4174, 2013
- [14] Kida S, Nakamoto T, Nakano M et al: Computed tomography image quality improvement using a deep convolutional neural network. Cureus 10: e2548, 2018
- [15] Jiang Z, Chen Y, Zhang Y et al: Augmentation of CBCT Reconstructed from Under-sampled Projections using Deep Learning. IEEE Trans Med Imaging 38: 2705-2715, 2019
- [16] Sahiner B, Pezeshk A, Hadjiiski LM et al: Deep learning in medical imaging and radiation therapy. Med Phys 46: e1-e36, 2019
- [17] Mori S, Hirai R, Sakata Y: Using a deep neural network for four-dimensional CT artifact reduction in image-guided radiotherapy. Phys Med 65: 67-75, 2019
- [18] Harms J, Lei Y, Wang T et al: Paired cycle-GAN-based image correction for quantitative cone-beam computed tomography.

Med Phys 46: 3998-4009, 2019

- [19] Eckl M, Hoppen L, Sarria GR et al: Evaluation of a cycle-generative adversarial network-based cone-beam CT to synthetic CT conversion algorithm for adaptive radiation therapy. Phys Med 80: 308-316, 2020
- [20] Clark K, Vendt B, Smith K et al: The Cancer Imaging Archive (TCIA): Maintaining and operating a public information repository. J Digit Imaging 26: 1045-1057, 2013
- [21] Zhu JY, Park T, Isola P et al: Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. IEEE International Conference on Computer Vision, 2017, pp2242-2251
- [22] Liang X, Chen L, Nguyen D et al: Generating synthesized computed tomography (CT) from cone-beam computed tomography (CBCT) using CycleGAN for adaptive radiation therapy. Phys Med Biol 64: 125002, 2019

Biography



Keisuke Usui

Lecture with the Dept. of Radiological Technology, Univ. of Juntendo. He received Ph.D. (Eng.) from Univ. of Hosei in 2015. He is a winner of JSMP grand president's award (2012, 2015) and JAMIT best presentation award (2015). His research fields include medical imaging and radiation therapy. He is a member of JAMIT, JSMP, JSRT, JASTRO and IEEE.

Koichi Ogawa

Professor Koichi Ogawa received his B.Sc., M.Sc., and Ph.D. in Electrical Engineering in 1980, 1982, and 1989 from Keio University, Tokyo, Japan. He is with the Dept. of Applied Informatics, Faculty of Sci. and Engineering, Hosei University, Tokyo, Japan. He is a former president of JAMIT and awarded the outstanding medical physicist award from the Asia-Oceania Federation of Organizations for Medical Physics (AFOMP) in 2020. He is a member of JAMIT, JSMP and IEEE.

Masami Goto

Yasuaki Sakano

Associate professor with the Dept. of Radiological Technology, Univ. of Juntendo. He received Ph.D. from Kanazawa University, Japan, in 2012. His research fields include medical imaging. He is a member of JSMRM and JSRT.

Professor with the Dept. of Radiological Technology, Univ. of Juntendo. He received Ph.D. from Tokyo Metropolitan University in 2018. His research fields include radiobiology and

environmental radiology. He is a member of JSRT, JSERT and JAHS.



Shinsuke Kyogoku



Professor with the Dept. of Radiological Technology, Univ. of Juntendo. He received M.D. in 1985 from Juntendo University. He is department dean of Dept. of Radiological Technology in Juntendo University. His research fields include general diagnostic imaging. He is a member of JRS, JSNM, JSPR, JSMRM and BMFSA.

Hiroyuki Daida

Professor with the Dept. of Radiological Technology, Univ. of Juntendo. Dean, Faculty of Health Science. He received M.D. in 1979 from Juntendo University. He had been the chair and professor, Department of Cardiovascular Medicine in 2000-2019. His research fields include Atherosclerosis, coronary artery disease and cardiovascular imaging. He is a fellow of JCS, JCC, ESC and ACC and a member of JAS.

近接マルチスプリッティングを用いた

Row-Action 型 TV 正則化 CT 画像再構成法の体系的な構成

金 喜正*1 定方 数弥*1 工藤博幸*1

要旨

圧縮センシングによる低線量 CT やスパースビューCT の画像再構成にはトータルバリエーション (TV) 正則化項と最小 2 乗誤差のデータ項の和の形をした評価関数を反復法で最小化することにより再構成画像 を得ている.評価関数の反復法には同時反復型の Chambolle-Pock アルゴリズムや FISTA 法がよく用いられ ているが、画像の収束が非常に遅いことが課題となっている.本研究では高速に評価関数を最小値に収束 する Row-Action 型の構造を持つ反復法を、近接マルチスプリッティングと呼ばれる数学的枠組みを用いて 体系的に示した.アルゴリズムに組み込む枠組みは Passty のスプリッティング,Dykstra 型スプリッティン グ、近接最小化トリック付き Dykstra 型スプリッティングの3つであり、2つ目以降の手法はCT 画像再構 成に使われた先行研究例はなく、十分な新規性のある手法となっている.3手法での腹部 CT 実画像を用 いた低線量 CTを想定したシミュレーション実験を行った結果、多くの研究で用いられている OS-SIRT 法 に優れることを確認した.

キーワード: Row-Action, 画像再構成, 低線量 CT, トータルバリエーション

1. はじめに

コンピュータ 断層 撮影 (C T: Computed Tomography)は,被写体を切断することなく物体 の重なりを含まない鮮明な断面図を取得する ことができることから,この技術は工業から医療までの幅広い分野で応用されている装置である.[1]

医療分野における CT 技術は,精密検査やス クリーニングなどで利用されているが,長年取 り扱われている課題として医療用 CT における 被ばく量に関する問題が挙げられる.現代の C T 装置は放射される X 線の放射角度や,放射量 の削減により被ばく量の軽減を行うことで低

*1 筑波大学大学院システム情報工学研 究群情報理工学位プログラム 〔〒305-8573 つくば市天王台 1-1-1〕 e-mail: s2020582@s.tsukuba.ac.jp

投稿受付:2021年月日

線量での診断を実現している.しかし,これら の工夫の弊害として再構成画像には多くのア ーティファクトを含くまれてしまうことが問 題となっている.また,画像再構成技術におけ る同時反復手法を用いたアルゴリズムでは,C T 画像を再構成するまでの収束に時間がかかる ことがある.そこで本研究では低線量CTにお ける画像再構成において収束スピードが速く, 高精度な画像を再構成する手法を体系的に示 す.本研究はCT 画像再構成の分野において初 めて再構成アルゴリズムに組み込む3つの手法 を実験結果と共に示しており,その新規性は高 いものとなっている.

2. 提案手法

CTの画像再構成は一般的に反復法を用いて 評価関数を最小化することによって再構成画 像を得ている.本研究では最小化の対象となる 評価関数を,同時反復型の手法ではなく,評価 関数の収束が速いとされている Row-Action 型 の構造を持つ反復法を用いる. さらに近接マル チスプリッティングと呼ばれる数学的枠組み を組み合わせることによって微分が不可能な 項を含む評価関数に対しても最小化問題を解 き,再構成画像を得ることができる.本手法の アルゴリズムに組み込まれている数学的枠組 みは3つあり,(1)第1の手法(Passyのスプリ ッティング),(2)第2の手法(Dykstra 型スプリ ッティング),(3)第3の手法(近接最小化トリッ ク付き Dykstra 型スプリッティング)の3つの枠 組みを用いて,評価関数を最小化する[2],[3], [4]. これらはCT 画像再構成において先行研究 例がなく,新規性の富んだ研究となっている.

はじめに本研究では CT における再構成画像 を取得するために,最小化の対象となる評価関 数を式(1)に示す.

 $f_{\beta}(\vec{x}) = L(\vec{x}) + \beta U(\vec{x})$ (1) ここで $L(\vec{x})$ はデータ項を指し、 $\beta U(\vec{x})$ は正則化 項を表している。データ項には測定データの二 乗誤差を用おり、正則化項には Total Variation

(TV)を用いている. これらを踏まえて, 式(1)を 式(2)へ変換する.

$$f_{\beta}(\vec{x}) = \frac{1}{2} ||A\vec{x} - \vec{b}||^2 + \beta ||\vec{x}||_{TV}$$
(2)

式(2)において
$$\frac{1}{2}$$
|| $A\vec{x} - \vec{b}$ ||²がデータ項を表し,

 $\beta || \vec{x} ||_{TV}$ が正則化項を表している.

ここで, TV については式(3)のように定めると し, *x_{i,j}*は画像*x*上のピクセル(*i,j*)に位置する画 素値を指しているとする.

 $U(\vec{x}) = ||\vec{x}||_{TV} =$

$$\sum_{i} \sum_{j} \sqrt{\left(x_{i,j} - x_{i+1,j}\right)^2 + \left(x_{i,j} - x_{i,j+1}\right)^2}$$
(3)

次に式(2)を近接スプリッティングへ組み込む. 一般的に, *f*(*x*)に対応する近接写像は式(4)のように定義される.

 $\operatorname{prox}_{\gamma f}(\vec{x}) = \arg \min_{\vec{z} \in \mathbb{R}^N} (f(\vec{x}) + \frac{1}{2\gamma} ||\vec{z} - \vec{x}||^2)$ (4)

ここで $\gamma > 0$ はステップサイズのパラメータを 指している. $f(\vec{x})$ は凸関数であれば TV のよう な微分が不可能な関数も組み込むことが可能 である.次に式(4)について近接スプリッティン グを用いて部分評価関数の和として再定義す る.その結果が式(5)にあたる.

$$\begin{cases} \vec{x}^{(k,i+1)} = \operatorname{prox}_{\gamma^{(k)}f_i} (\vec{x}^{(k,i)}) \ (i = 1, \dots, n) \\ \vec{x}^{(k+1,1)} = \vec{x}^{(k,n+1)} \end{cases}$$
(5)

式(5)に対して式(2)との関連付けを行うとそれ ぞれ式(6),(7)として表すことができる.

$$f_{\beta}(\vec{x}) = \frac{1}{2} ||A\vec{x} - \vec{b}||^2 + \beta ||\vec{x}||_{TV}$$

$$\equiv \sum_{i=1}^{I} f_i(\vec{x}) + f_{I+1}(\vec{x})$$
(6)

$$f_{i}(\vec{x}) = \frac{1}{2} \left(\vec{a}_{i}^{T} \vec{x} - \vec{b}_{i} \right)^{2} (i = 1, 2, ..., I),$$

$$f_{I+1}(\vec{x}) = \beta ||\vec{x}||_{TV}$$
(7)

以上のように関連づけることで、微分不可能な 正則化項を含む問題に関しても最小化問題を 解くことができる.

最後に本提案手法で用いる評価関数と最小 化モデルについて述べる.第1の手法(Passty+ TV)についての評価関数は式(8)で表し,アルゴ リズムの重要箇所について式(9)に示す.

$$\min \equiv \sum_{i=1}^{l} f_i(\vec{x}) \tag{8}$$

$$\vec{x}^{(k,i+1)} = \operatorname{prox}_{\alpha^{(k)}f_i(\vec{x})}(\vec{x}^{(k,i)})$$

$$\equiv \min_{x} \left(\frac{1}{2\alpha} ||\vec{x} - \vec{x}^{(k,i)}||^2 + f_i(\vec{x}) \right) \quad (9)$$

ここで*a*^(k)はステップサイズを指し,経験則が 必要なパラメータとなっている.そのため,収 束の速い Row-Action 型のアルゴリズムではあ るが,実用性は低い手法となっている.続いて 第2の手法 (Dykstra 型スプリッティング+TV) についても同様に式(10), (11)を通して示す.

$$\begin{split} \min_{x} &\equiv \frac{1}{2\alpha} ||\vec{x} - \vec{m}||^{2} + \sum_{i=1}^{I} f_{i}(\vec{x}) \quad (9) \\ \vec{x}^{(k,i+1)} &= \operatorname{prox}_{\alpha f_{i}(\vec{x})}(\vec{x}^{(k,i)} + \vec{y}_{i}^{(k)}) \\ &\equiv \min_{x} \left(\frac{1}{2\alpha} ||\vec{x} - \vec{x}^{(k,i)} + \vec{y}_{i}^{(k)}||^{2} + f_{i}(\vec{x}) \right) \quad (10) \end{split}$$

第2の手法は式(9)にあるように,作為的に摂動 項 $\frac{1}{2\alpha}||\vec{x} - \vec{m}||^2$ を加える.これによって経験則が 必要なパラメータは不要となるが,摂動項を考 慮したパラメータαの調節が必要となる.最後 に第3の手法 (DLPM+TV) について述べる.評 価関数とアルゴリズムの重要部分を式(11), (12) に示す.

$$\begin{split} \min_{x} &\equiv \sum_{i=1}^{l} f_{i}(\vec{x}) \quad (11) \\ \vec{x}^{(k,i+1)} &= \operatorname{prox}_{\alpha f_{i}(\vec{x})} (\vec{x}^{(k,i)} + \vec{y}_{i}^{(k)}) \\ &\equiv \min_{x} \left(\frac{1}{2\alpha} ||\vec{x} - \vec{x}^{(k,i)} + \vec{y}_{i}^{(k)}||^{2} + f_{i}(\vec{x}) \right) \quad (12) \end{split}$$

本手法は摂動項の影響を最小限にするための 工夫がアルゴリズムに施されている.具体的に は、アルゴリズムの反復において、1反復の前 の画像との差分を摂動項内に取り入れること によって、その影響を最小化している.この手 法によって、経験則の必要なパラメータは不要 となり、摂動項の影響を抑えた状態で再構成画 像を取得することが可能となる.

3. 実験結果

提案する3手法を本研究では腹部の実画像を 用いてシミュレーション実験を行った.画像の サイズは384×384とし,CTにおいて想定され るノイズとしてポアソンノイズを付与してい る.再構成の対象となる腹部の元画像を図1に 示す.



図1 シミュレーションの正解画像(腹部の実画像)

本実験では、第1~3の手法との比較実験のため、OS-SIRT with TV による再構成画像をともに示す.OS-SIRT は収束の速い Row-Action 型の 再構成アルゴリズムであるが、ノイズに対しての耐性が低いため、安定した画像の収束は望めない.これらの腹部実画像を用いたシミュレーション画像を図2に示す.



図2 各手法における 10 回反復後の再構成画像

表1 論文の各部分で使用するフォント.

	RMSE	SSIM
OS-SIRT	0.00825	0.767
第1の手法	0.00548	0.951
第2の手法	0.00241	0.972
第3の手法	0.00241	0.972

さらに定量評価のため、評価指標として Root Mean Squared Error (RMSE) & Structural Similarity (SSIM)を用いた評価結果を表1に示 す. 表 2, 図 2 の結果から, 提案手法である第 1~3の手法は OS-SIRT と比べて高精度な画 像が得られていることがわかる. 定量的な結果 から第2,3の手法が最も良い結果となってお り、定量的な結果からは、第1~3の手法につ いて高精度な再構成画像が得られていると言 える.実用性について考慮してこれらの結果を 考察すると、経験則の必要なパラメータ調節が なく、摂動項の影響等も考慮する必要のない第 3 の手法は、大変実用性の高い手法であると判 断することが可能である.続いて、収束速度に ついての結果画像を図3に示す.こちらもOS-SIRT については収束がリミットサイクルへ陥 てしまっているが、第1~3の手法は1回の反 復で高精度な画像に収束できていることが分 かる.

3. まとめ

本研究では、新たな数学的枠組みを初めて CT 画像再構成のアルゴリズムに組み込むことに



図3 各手法における1,5,10回反復ごとの再構成画像

より,実用的に高精度な再構成画像を速い収束 スピードで取得する手法を体系的に示した.今 後は本手法をフォルトトレラントなどの他の 再構成問題へ適用することを展望としている.

利益相反の有無

なし

文 献

- [1] 日本医用画像工学会:医用画像ハンドブ ック,190,2012
- [2] Boyle J.P and Dykstra R.L: A method for Finding Projections onto the Intersection of

Convex Sets in Hillbert Spaces. Lecture Notes in Statistics 37, 1986, pp28-47.

- [3] Han S.P: A Decomposition Method and its Application to Convex Programming. Math Oper Res 14, 1989, pp237-248.
- [4] Combettes P.L and Pesquet J.C: Proximal Splitting Methods in Signal Processing. In Bauscheke H.H, Burachik R.S, Combettes P.L et al. eds. Fixed-Point Algorithm for Inverse Problems in Science and Engineerign. 2011, pp185-212.

Systematic Construction of

Row-Action Type TV Regularized CT Image Reconstruction Method

Using Proximal Multi-Splitting

HeeJeong Kim^{*1}, Kazuya Sadakata^{*1}, Hiroyuki Kudo^{*1}

*1 University of Tsukuba

For image reconstruction of low-dose CT and sparse-view CT using compressed sensing, the reconstructed image is obtained by minimizing an evaluation function in the form of a sum of a total variation (TV) regularization term and a least-squares error data term using an iterative method. The Chambolle-Pock algorithm and the FISTA method, which are simultaneous iterative methods, are often used for iterating the evaluation function, but the problem is that the convergence of the image is very slow. In this research, we systematically present iterative methods with a Row-Action type structure that converges the evaluation function to a minimum value quickly, using a mathematical framework called proximity multi-splitting. The three frameworks incorporated into the algorithm are Passty's splitting, Dykstra's splitting, and Dykstra-type splitting with the proximity minimization trick. The second and subsequent methods have not been used for CT image reconstruction in previous studies and are sufficiently novel. Abdominal CT using the three methodsSimulation experiments using real images for low-dose C T confirmed the superiority of the OS-SIRT method, which has been used in many studies.

Key words: Row-Action, Reconstruction, Low-does CT, Total Variation.

インテリア CT 厳密解法における安定な画像再構成が可能な先

験情報領域配置の解明

伊藤 優太*1 藤井 克哉*2 工藤 博幸*2

要旨

インテリア CT とは CT 画像再構成の一つで, 心臓や乳房などの検査の関心領域(ROI)に対してのみ X 線を照射し, 関心領域の画像を再構成する手法である. 全身に X 線を投影する従来の CT と異なり, インテリア CT では ROI を通る直線上にしか X 線を投影しないため, 被曝量が少なくなるという利点がある一方で, 関心領域を通らない投影データは測定されないため, 得られる投影データは不完全なものになる.

この不完全データに対し, ROI 内部の任意の小領域において値が既知(先見情報領域)であれば解が一意に定 まることが発見された[1,2]. しかし, 解が一意に定まる一方で解が安定して求まるのは先見情報領域の近傍のみ であり, 先見情報領域から離れるほど安定性が下がる. 本研究では従来手法を拡張し, 先見情報領域の数を複 数にすることで, その先見情報領域で囲まれた多角形内部という広い領域でが安定に再構成可能になることを示 した.

キーワード: CT, 画像再構成, インテリア CT,

1. はじめに

コンピュータ断層撮影(Computed Tomography: CT)は対象物体の内部情報を非破壊で得ること ができる技術であるため、工業・医療の分野で 広く用いられており、医療現場では体内の病変 を発見する手段の一つとして用いられる.しか し、2004年にCTで用いるX線被曝によるガン 発生リスクに関する論文[3]が発表され、以降 低被曝を実現するCTに重点が置かれて研究が 行われている.インテリアCTは低被曝を実現 する手法の一つである.従来のCTは物体断面 全体を通るようにX線を照射し,物体を透過し たX線の減衰を検出器で測定し,得られたデー

*1 筑波大学大学院システム情報工学研究群情 報理工学位プログラム [〒305-8573 つくば市天王台 1-1-1] e-mail: s2120561@s.tsukuba.ac.jp *2 筑波大学システム情報系 投稿受付:2021年6月17日

タから物体内部を再構成する. それと比較し、 インテリア CT では物体内部にある検査の関心 領域のみを通るように X線を照射し、それによ って得られたデータから物体内部の画像を再 構成する. 医療現場における CT イメージングで も心臓病や乳がんの診断といった場合、心臓や 乳房などの特定の領域の画像のみ再構成でき れば十分である.このような場合にインテリア CT を用いると X 線の被曝量が減少するという 利点があるが、その一方で検査の関心領域を通 らないデータは測定されないため、得られるデ ータは不完全なものになる. そのため、インテ リア CT は角度欠損問題, エクステリア問題と 並び CT における不完全データからの再構成問 題の一つにあげられる.こうした不完全データ からの画像再構成に従来の CT で用いられてい た画像再構成法であるフィルタ補正逆投影法

(Filterd Back Projection : FBP) を適用して も正しく画像を再構成することはできない.

インテリア CT の画像再構成に関して Natterer は、インテリア CT の解は一意に定ま らないことを数学的に証明した[4]. ここで一 意とは逆問題において解が数学的に唯一に定 まることを指す.そのため,近似的手法が多く の研究 [5,6] で発表されてきたが 2007 年-2008 年にかけて Ye らや Kudo らによりインテ リア CT の厳密解法が発見された [1,2].これ らの論文では関心領域(ROI)内部の任意の小領 域において値が既知 (先見情報領域)であれば, インテリア CT の解が一意に定まることが証明 されている.しかし解の一意性は示された一方 で,解が安定的に求まるのは先見情報領域の近 傍領域のみであり,先見情報領域から離れるほ ど安定性が下がる.

本研究では従来の研究を拡張し,先見情報領 域の数を複数個にし,先見情報領域で結ばれた 幾何形状内の安定性の評価を行なった.従来研 究は先見情報領域を通る直線上のみの安定性 の評価に留まっている.そのため、安定的に解 が求まる領域が狭いという問題があった.本研 究では安定的に求まる領域をより大きく取れ るため,ノイズに強い画像再構成が可能である.

2. インテリア CT の従来解法

1) 微分逆投影法

インテリア CT で得られる投影データは欠損 したデータとなるため、これまで CT 画像再構 成で用いられてきた FBP 法を適応することは できない. そこで考案された手法が微分逆投 影 (Differentiated Backprojection: DBP) で ある.対象物体をf(x,y)、ROI をS、f(x,y)の値 が既知である先見情報領域をBとする.DBP 法 ではまず、ROI Sをヒルベルト直線と呼ばれる 直線の集合L(u)に分解する.そして、このヒル ベルト直線ごとにヒルベルト変換と呼ばれる 積分変換の逆変換をすることで画像再構成を 行う.この時、ヒルベルト直線は以下の二つの 条件を満たすように選択する.

- 1. ROI S内部の全ての点が少なくとも一つの ヒルベルト直線上に含まれる.
- 全てのヒルベルト直線が先見情報領域 Bと 交わる.

ここで,ある *L(u)*上に一次元座標 t を定義する. 座標 *t* 上の元画像*f(x,y)*を*f(t)*とすると逆変換は次のように表せる.

$$\sqrt{1-t^2}f(t) = C + p.\nu \int_{-1}^{1} \frac{Hf(s)}{t-s} \sqrt{1-s^2} ds \quad (1)$$

2) 先見情報領域が一箇所の時の安定性

詳細は[1]に示されているため,ここでは次の ような簡略化した式を考える.ノイズを含む再 構成画像には下付きの r をつけることにする.

 $|f_r(t) - f(t)| < K(t)\varepsilon^{1-\omega(t)}$ (2)

 $\mathbb{CCC}\lim_{t\to a} 1 - \omega(t) = 1 \mathcal{CB} \mathcal{V}, \quad \lim_{t\to R_1} 1 - \omega(t) =$

0である. また $\lim_{t\to a} K(t) < \infty$ であり, $\lim_{t\to B_t} K(t) =$

∞となる.このことから、(2)は先験情報領域 に近いほど安定となり、離れるほど不安定に なることがわかる.図1に先見情報領域を一箇 所与えた時の問題設定を示す.



図 1: 先見情報領域一箇所の時の問題設定.

3) 先見情報領域が二箇所の時の安定性

先見情報領域が二箇所の時の安定性は以下 のように示せる.

$$|f_r(t) - f(t)| < K(t)\varepsilon^{1-\omega(t)}$$
(3)

先見情報領域が一箇所の時との違いは、 $1-\omega(t) = \delta(t)$ とすると、 $\delta(t), K(t)$ はそれぞれ、

$$\lim_{t\to b,c}\delta(t)=1,\lim_{t\to b,c}K(t)<\infty$$

となることである. これは b, c が FOV の端から 十分に遠い時に成立する. FOV とは完全なイン テリア CT の時, ROI の同様である. よって δ, K を

$$\delta = \min_{b,c} \delta(t), K = \max_{b,c} K(t) < \infty$$

と定めると,任意の*x* ∈ (*b*,*c*)で

 $|f_r(t) - f(t)| < K\varepsilon^{\delta}$

となり、二箇所の先見情報領域間では座標 t に よらず安定して求まる.しかしながらヒルベル ト直線が一箇所の先見情報領域しか通らない 場合は先見情報領域を一箇所与えた場合と同 様に、先見情報領域から離れるほど不安定にな るという問題がある.図 2 に先見情報領域を二 箇所与えた時の問題設定を示す.



図 2:先見情報領域二箇所の時の問題設定.

3. 提案手法

この節では、先験情報領域が3つの場合の安 定性を2)、3)をもとに示す.これまでとの相 違は考えうる領域が2次元平面に拡張される ため、図3に示すように安定性を評価する方 程式が無数に存在する.これら無数のパスの 中で最大となる不等式を、その点における安 定性評価式とする.また、3)より、先験領域 で囲まれた領域は位置に依らず安定になるこ とから、囲まれた領域を新たな先験情報領域 とすることで安定性領域を拡張することが可 能となる[7].このとき、点Xを通るヒルベル ト直線は大きく分けて2つ存在する.

 真値の先験領域と誤差を含む先験領域を 通る直線

2. 両方とも誤差を含む先験領域を通る直線

1, 2 はともに先験情報領域で囲まれているため不等式は有界となることに注意する.

1,2に関して,考えられる主な不等式は以下で 与えられる.

$$|f_r(t) - f(t)| < \varepsilon^j + \kappa^j \left(\frac{\varepsilon + \varepsilon_i}{\kappa^j}\right)^{1 - \omega(t)} \tag{4}$$

ここで、各 ϵ_i は誤差を含む先験情報領域のずれ であり、 κ^j は先験情報領域の配置から決まる 定数である.また、 ϵ は、対象のヒルベルト直 線上で、投影データ取得領域の誤差を表す. 不等式(4)はxが先験情報の場所とそこを通る パスを定めるパラメータ*i*,*j*により評価式が変 わることを意味している.従って,不等式(4) に対して,最大値を取った以下の式が点Xに おける安定式となる.

$$|f_r(t) - f(t)| < \max_{i,j} \varepsilon^j + \kappa^j \left(\frac{\varepsilon + \varepsilon_i}{\kappa^j}\right)^{1 - \omega(t)}$$
(5)

式(5)は3箇所の先験領域で囲まれた三角形の 内部の点に対して成り立つため、少なとも三 角形領域では安定になることがわかる.



図 3:三箇所の先験領域を通るヒルベルト直線.

4. 実験

提案手法の有効性を確認するために,400x400 ピクセルの頭部ファントム画像の再構成を行 い,再構成画像を比較する.それぞれ,先見情 報領域を与えずに再構成した時,先見情報領域 を二箇所与えた時,三箇所与えた時で比較する. 図 4に ROI とそれぞれの先見情報領域を示す.



図 4:実験の問題設定. (a)ROI の範囲, (b)先 見情報領域を二箇所与えた時, (c)先見情報領 域を三箇所与えた時.

1)ファントム画像の再構成

本実験では安定性の変化の要因である先見 情報領域の数を変えて実験を行った.画像の再
構成はサブセット分割した ART 法を用いた.投 影数は1200,サブセット数は16,イテレーショ ンは4000回である.再構成画像を図 5に示す. また図 6 には再構成画像を詳細に見るために 再構成画像と元画像の差分を示した.



図 5:元画像と再構成画像. 左上:元画像. 右上: 先見情報領域なし. 左下:先見情報領域二箇所 与えた時. 右下:先見情報領域三箇所与えた時.



図 6:再構成画像と元画像の差分.

5. 考察

1) 再構成の安定性について

図 5より,先見情報領域なしの時,二箇所の 時と比較し,三箇所与えた時の再構成画像が 最も真値と近くなっている.逐次近似再構成 の場合,徐々に誤差が生じるため,この結果 から雑音に強く安定性が高いと言える.

2)先見情報領域の面積について

本研究は先見情報領域の数に関する問題で ある.しかし,実験での問題設定から先見情 報領域の数ではなく,先見情報領域の面積が 影響していることも考えられる.そこで先見 情報領域の数は変えずに,先見情報領域を二 箇所与えた時と同等になるように面積を変 えて比較した.二箇所与えたときはそれぞれ 10x10 ピクセルの領域であった.今回は 8x8 ピクセルの領域三箇所与えた時と比較する. 結果を図 7 に示す.図からわかるように,先 見情報領域の面積が同等でも三箇所与えた 時の方が真値に近くなっていることがわか る.ここから,再構成のパラメータとして有 用なのは先見情報領域の面積ではなく,数で あると言える.



図 7:領域の面積が同等になるように与えた場 合の再構成画像と元画像の差分.

6. まとめ

本研究ではインテリア CT 画像再構成の従来 手法を拡張し,新たな安定性の評価式を提案し た.本評価式では先見情報領域の数を3個とし, その先見情報領域で結ばれた領域内部では安 定して解が求まる.これにより,従来手法より もノイズに強い画像再構成が可能になる.

実験では,逐次近似法を用いて画像再構成を 行い,再構成画像を比較した.この時,先見情 報領域の数を変化させて実験を行い,従来手法 と比較して本提案手法が最も安定して再構成 できることを示した.今後は実データを用いて 実験を行い本提案手法の有用性を示したい.

謝 辞

本研究は JST-CREST (JPMJCR1765) の助成を受けた.

利益相反の有無

なし.

文 献

[1] Hiroyuki Kudo, Matias Courdurier, Frederic Noo, and Michel Defrise. Tiny a prioriknowledge solves the interior problem in computed tomography. *Physics in Medicine and Biology*, Vol. 53, No. 9, pp. 2207–2231, apr 2008.

[2] Yangbo Ye, Hengyong Yu, Yuchuan Wei, and Ge Wang. A general local reconstruction approach based on a truncated hilbert transform. *International Journal of Biomedical Imaging*, Vol. 2007, , 2007.

[3] Amy Berrington de Gonzalez and Sarah Darby.
Risk of cancer from diagnostic x-rays: estimates for the uk and 14 other countries. *The lancet*, Vol. 363, No. 9406, pp. 345-351, 2004.

[4] Frank Natterer. *The mathematics of computerized tomography*. SIAM, 2001.

[5] K. Ogawa, M. Nakajima, and S. Yuta. A reconstruction algorithm from truncated projections. *IEEE Transactions on Medical Imaging*, Vol. 3, No. 1, pp. 34–40, 1984

[6] B. Ohnesorge, T. Flohr, K. Schwarz, J. P.
Heiken, and K. T. Bae. Efficient correction for ct image artifacts caused by objects extending outside the scan field of view. *Medical Physics*, Vol. 27, No. 1, pp. 39-46, 2000.

[7] Courdurier M, Noo F, Defrise M and Kudo H:
Solving the interior problem of computed
tomography *a priori* knowledge. *Inverse Problems*24: 065001, 2008

A prior knowledge arrangement that enables stable and exact ROI

reconstruction for the interior problem in computed tomography

Yuta ITOH^{*1}, Katsuya FUJII^{*2}, Hiroyuki KUDO^{*2}

*1 Degree programs in Systems and information Engineering, Graduate School of Science and Technology, University of Tsukuba
*2 Department of Computer Science, University of Tsukuba, Japan

We propose a prior knowledge regions arrangement technique that enables stable and exact region of interest (ROI) reconstruction for the interior problem in computed tomography. In the previous method, if one or two tiny prior knowledge regions in the ROI are available, the uniqueness result can be obtained. However, the further from the prior knowledge regions, the stability is decreased. To overcome this drawback, we show that the solution of the interior problem is stable within a broader region with the prior knowledge region as a vertex than the region of previous methods, if three tiny prior knowledge regions are available. We demonstrate that the proposed method works well with simulation studies

Key words: Interior problem, Computed tomography, Image reconstruction (maximum 5 words)

CT 画像再構成におけるインテリア問題とエクステリア問題

の統一解法と安定性解明

藤井 克哉*1 工藤 博幸*1

要旨

近年,被ばく量を軽減するための新たな CT 画像再構成法が精力的に研究されている.不完全投影データ からの再構成でよく知られた問題は投影データ欠損が互いに外側と内側に起きている場合の再構成問題で あり,それぞれインテリア問題,エクステリア問題と呼ばれる.安定性解析において前者は Defrise や Coudurier らの先行研究によって安定な問題となることが肯定的に示されているが,後者は Cormack によ り不安定な問題となることが知られている.両者の相違は投影データの欠損が内側か外側かのみではある が用いる数学的枠組みは大きく異なるため,扱いが難しい.そこで本研究では,エクステリア問題の安定 性も含めた両問題の統一的な理論を提案する.そのための鍵は Cormack の第二再生公式と先験情報領域の 幾何的な配置であり,これらによりインテリア問題,混合インテリアエクステリア問題,更には未解決で あったエクステリアの安定性を統合的に保証することが可能になることを示した.またこれらのシミュレ ーション実験を行い理論の妥当性を検証した.

キーワード: CT, インテリア CT, エクステリア CT, 画像再構成

1. はじめに

X線コンピューター断層撮影(CT)は、さま ざまな方向から X線を照射することにより 対象物の断層画像を取得する手法である.近 年、CTへの応用では、制限された投影データ からの再構成問題に関心が集まっている. このとき数学的に問題となるのは、解の一意 性と安定性である.本稿では再構成の安定性 問題を扱う.安定性とは、投影データに雑音 がある場合、再構成画像の雑音の伝搬性を表 す.雑音の伝搬が著しく顕著になるとき、そ の再構成法は不安定な問題となり、実際の再 構成に深刻な影響をもたらす.そこで、理論 的な側面から安定性を測る必要がある.

制限された投影データからの再構成問題に は、内部の問題(インテリア問題),外部の 問題(エクステリア問題),制限された角度の 問題(Limited Angle problem)などがよく知ら れている[1].本稿では、インテリア問題とエ

〔〒305-8573 つくば市天王台 1-1-1〕 e-mail: fujii@imagelab.cs.tsukuba.ac.jp
 投稿受付: 2021 年 6 月 18 日

クステリア問題, さらにはこれらの混合問題 である混合インテリアエクステリア問題を扱 う.インテリア問題とエクステリア問題は、そ れぞれオブジェクトの内側と外側にあるとき に関心領域 (ROI) を通過する投影データの みに基づく再構成方法である.前者は医療現 場において,対象の一部を復元することによ り低被ばくを実現し、後者は金属など X 線 を完全に吸収するような系に対し(例えば歯 科用 CT) 有効である. インテリア問題は解 の一意性,安定性ともに ROI の中の任意の 先験領域が既知であれば成り立つことが知ら れている[2-4]. エクステリア問題に関して は,解の一意性が成り立つことは,Cormack の解析的逆変換公式を用いて示すことができ るが,特異値分解を行うことで逆変換(再構 成)は、著しく不安定となる[5].

本稿では、これら2つの問題を統一するような理論を紹介する.まず[6]でも紹介したように対象物には、リング状の先験情報領域が存在すると仮定する.インテリア、エクス

^{*1} 筑波大学システム情報系

テリア問題は互いに投影データの欠損が外部 か内部であるかの違いにも関わらず,扱う数 学が異なるため,異なる問題として捉えられ てきた.提案手法で紹介する Cormack の第 二公式と先験用法領域を用いることで,これ ら2つの問題の安定性の評価式を構成する.

2. 提案手法と安定性解析

インテリア問題の安定性を示す評価式は, Coudurier や Defrise らによって,先験領域付 近では安定となり,離れるほど不安定になる ことを主張している.

[インテリア問題の安定性評価]

 $|f - f_{err}| < \varepsilon + K(x)\varepsilon^{\delta(x)}$ (1) ここで ε は,投影データ取得領域での誤差を 表し,K(x)は,先験領域付近では有界で ROIの縁では非有界となるような関数であり, $\delta(x)$ は $0 \le \delta \le 1$ であり先験領域で1を取り, 縁 で 0 と な る よ う な 関 数 で あ る [3]. Coudurier らの導出した評価式の基となる再 生公式は有限 Hilbert 逆変換である.エクス テリア問題に対しては,この逆変換では ROI を完全に覆えるような Hilbert 直線が構 成できない.そこで今回,我々は以下の再生 公式を用いる[7].

[Cormack の第二再生公式]

 $f_{\ell}(r) = -\frac{1}{\pi} \int_{r}^{1} \frac{\left(s/r + \sqrt{s^{2}/r^{2} - 1}\right)^{-\ell}}{\sqrt{s^{2} - r^{2}}} g'_{\ell}(s) ds + \frac{1}{\pi r} \int_{0}^{r} U_{\ell-1}\left(\frac{s}{r}\right) g'_{\ell}(s) ds$ (2)

ここで、*g*_ℓは投影データであり、*U*_ℓはチェ ビシェフ第二多項式を表す. Cormack の第二 公式は投影データが完全(欠損のない)の場 合の解析的逆変換であるが、図 1 のように ROI の内部にリング型の先験領域が存在すれ ば、投影データ取得領域から得られた関数を 解析接続することにより、欠損のある投影デ ータからの再構成解は一意であることが解る. 一方、安定性は次のような評価が得られた. [インテリア・エクステリア問題における安

定性評価]

先験領域の外側で,

$$\left| f_{\ell} - f_{\ell,err} \right| < \varepsilon + K_1(r)\varepsilon^{\delta_1(r)}$$
(3)
先験領域の内側で,

 $\left|f_{\ell} - f_{\ell\,err}\right| < \varepsilon + K_2(r)\varepsilon^{\delta_2(r)} \tag{4}$

となる.ここで, $K_1, K_2, \delta_1, \delta_2$ は, インテリア 問題とほぼ同じ関数である.動径方向のパラ メータを調整することによって,上記の式は インテリア問題,エクステリア問題,混合問 題を含む安定性評価式となり,それ故エクス テリア問題に対しても先験領域の存在によっ て安定性を測ることができる.



図 1: 混合インテリアエクステリア問題

3. 数値シミュレーション

提案手法を数値シミュレーション実験によ り確認する. データセットは医療頭部 CT 画 像 512x512pixels を用いた. ROI は円環領域 であるとする.また投影方向数は 512 方向, 不完全投影データは内部と外部の欠損が起き ているサイノグラムとする. 先験領域の場所 により ROI は先験領域付近では再構成でき ていることが確認できる(図2参照).また、 先験領域で囲まれた領域は安定して再構成で きており、先験領域から離れるほど不安定に なることが図3からわかる. 図4では、ノ イズを含む投影データからのエクステリア CT の画像再構成である. y = 256の青線のプ ロファイルは先験領域ありの再構成であり, 赤線がなしのプロファイルである. 先験領域 が機能して、先験領域がない場合よりも安定 な再構成が可能となっている.

4. まとめ

本研究では、インテリア問題とエクステリ ア問題に関する安定性理論を統一的に考察し た.その結果、リング型先験領域の有無が一 意性のみならず安定性も評価できることを示 し、シミュレーションにて理論の妥当性を検 証し、エクステリア **CT** に対しても先験情報 領域の存在が良好な画像再構成をもたらすこ とが判明した.

謝辞

本研究は, JSTCREST (JPMJCR1765)の 助成を受けたものです.

利益相反の有無

なし.

文 献

- [1] 工藤博幸:第2章 投影からの画像
 再構成.日本医用画像工学会編:医
 用画像工学ハンドブック.国際文献
 社.東京.2012, pp39-84.
- [2] H. Kudo, M. Courdurier, et al: Tiny A Priori Knowledge Solves the Interior Problem in Computed Tomography.

Phys Med Biol. 53:2207-31, 2008.

- Y. Ye, H. Yu, et al: A general local reco- nstruction approach based on a truncated Hilbert transform.
 Int.J.Biomed.Imaging. 2007:Article ID 63634, 2007.
- [4] M. Coudurier, F. Noo, et al: Solving the interior problem of computed tomography using *a prior knowledge*. Inverse Problems. 24: 065001(27pp), 2008.
- [5] A. M. Cormack: Representations of a function by its line integrals, with some radiological applications. J Appl Phys 34:2722–7, 1963.
- [6] 藤井克哉,金鎔采,工藤博幸: CT 画像再 構成における混合インテリアエクス テリア問題の 解の一意性,第 39 回日 本医用画像工学会大会(JAMIT 2020) 予稿集, Paper No P3-14, 2020.
- [7] F. Natterer: The Mathematics of Computerized Tomography (second edition),
 SIAM (Classics in Applied Mathematics, Vol32), Philadelphia, PA, 2001.



図2:混合インテリアエクステリア CT 画像再構成



図3:図2の水平成分 y=256 プロファイル



Unified Methods and Stability Estimate of Interior and Exterior Problems

in CT Image Reconstruction

Katsuya FUJII*1, Hiroyuki Kudo*1

*1 University of Tsukuba

In recent years, new CT image reconstruction methods for reducing radiation exposure have been energetically studied. A well-known problem in reconstruction from incomplete projection data is the reconstruction problem when projection data truncations outside and inside each other and is called the interior problem and the exterior problem, respectively. In the stability analysis, the former is positively shown to be a stable problem by previous studies by Defrise and Coudurier. However, the latter is known to be an unstable problem by Cormack. The only difference between the two is whether the projection data is missing inside or outside, but it is not easy to handle because the mathematical framework is very different. Therefore, in this study, we propose a unified theory of both problems, including the stability of the exterior problem. The key to this is the geometric arrangement of Cormack's second reconstruction formula and the prior region, which comprehensively guarantees the stability of interior problems, mixed interior, exterior problems, and unsolved exteriors. It was shown that it is possible. We also conducted these simulation experiments to verify the validity of the theory.

Key words: CT, Image Reconstruction, Interior CT, Exterior CT

異常データ検知トモグラフィ ABD-CT の概念提唱と事例紹介

工藤 博幸*1

要旨

トモグラフィ(CT)において、測定した投影データの一部に異常データが含まれることがよく発生する. 例えば、医療用 X 線 CT において被写体内部に金属が含まれる場合に発生するビームハードニング、検出 感度特性のむらで発生するリングアーティファクト,暗視野 CT で固い結晶構造の被写体で発生する回折 コントラストアーティファクト,などがこれに相当する.このような状況において,フィルタ補正逆投影 (FBP) 法で画像再構成を行うと異常データの影響が画像全体に広がり強いアーティファクトが発生する ことがよく知られる。本研究では、異常データを画像再構成の過程で自動的に検知して除外して画像再構 成を行う,新しい CT の枠組み ABD-CT (Abnormal Data Detected CT) を提案する. ABD-CT では,画像再 構成の評価関数として異常データを除外する効果がある L⁰ ノルム(または L¹ ノルム)をデータ項に用い, トータルバリエーション(TV)と L^0 ノルム(または L^1 ノルム)の和 $f(\mathbf{x})=\beta |||\mathbf{x}||_{TV}+||A\mathbf{x}-\mathbf{b}||_0^0$ (または $f(\mathbf{x})=\beta$ |||x||rv+||Ax-b||1)を最小化することにより画像再構成を行う.異常データの発生の仕方に制約がありまとま って異常データになる場所が既知であれば、その情報も加味して評価関数 f(x)は異常データになりやすい 場所をまとめてグループL⁰ノルム(またはグループL¹ノルム)により設計する.本手法の有効性を示す実 データの再構成例を紹介する.投影角度毎にX線のエネルギーと強度が異なる極端なビームハードニング を発生するミリ秒トモグラフィ、医療用コーンビーム CT と放射光 CT におけるリングアーティファクト、 暗視野 CT で問題となる回折コントラストアーティファクト, 医療用 CT のメタルアーティファクトであ る. 全てシミュレーションではなく実データであるが、ABD-CT の枠組みは強力でいずれも高精度の画像 再構成に成功している.

キーワード: CT, 画像再構成, スパースモデリング, 異常データ検知, メタルアーティファクト

1. はじめに

トモグラフィ (CT) において, 測定した投影 データの一部に異常データが含まれることが よく発生する. 例えば, 医療用 X 線 CT におい て被写体内部に金属が含まれる場合に発生す るビームハードニング, 検出感度特性のむらで 発生するリングアーティファクト, 暗視野 CT で固い結晶構造の被写体で発生する回折コン トラストアーティファクト, などがこれに相当 する. このような状況において, フィルタ補正 逆投影 (FBP) 法で画像再構成を行うと異常デ ータの影響が画像全体に広がり強いアーティ ファクトが発生することがよく知られる. 本研

*1 筑波大学システム情報系 〔〒305-8573 つくば市天王台 1-1-1〕 e-mail: kudo@cs.tsukuba.ac.jp 投稿受付:2021年7月2日 究では, 異常データを画像再構成の過程で自動 的に検知して除外して画像再構成を行う、新し いCTの枠組み ABD-CT (Abnormal Data Detected CT)を提案する. ABD-CT では, 画像再構成の 評価関数として異常データを除外する効果が ある L⁰ ノルム(または L¹ ノルム)をデータ項 に用い、トータルバリエーション(TV)と L⁰ノル ム(または L^1 ノルム)の和 $f(\mathbf{x})=\beta |||\mathbf{x}||_{\mathrm{TV}}+||A\mathbf{x}-\mathbf{b}||_0^0$ (または *f*(**x**)=β |||**x**||_{TV}+||*A***x**-**b**||₁¹) を最小化する ことにより画像再構成を行う.異常データの発 生の仕方に制約がありまとまって異常データ になる場所が既知であれば、その情報も加味し て評価関数 f(x)は異常データになりやすい場所 をまとめてグループL⁰ノルム(またはグループ L¹ノルム)により設計する.本手法の有効性を 示す実データの再構成例を紹介する.投影角度 毎に X 線のエネルギーと強度が異なる極端な ビームハードニングを発生するミリ秒トモグ ラフィ, 医療用コーンビーム CT と放射光 CT に

おけるリングアーティファクト, 暗視野 CT で 問題となる回折コントラストアーティファク ト, 医療用 CT のメタルアーティファクトであ る. 全てシミュレーションではなく実データで あるが,ABD-CTの枠組みは強力でいずれも高 精度の画像再構成に成功している. 比較的簡単 なアイデアで驚くべき画質改善を実現してい る.参考のため、同じ問題やデータに深層学習 の適用も試みているが、正直に述べると現状で は ABD-CT と同等の高画質の画像を再構成す ることには成功しておらず深層学習は学習デ ータ収集などはるかに手間がかかる.

2. 提案手法

1) ABD-CT の原理

CT 計測の多くの場面において、測定した投 影データの一部に異常値が含まれる状況が発 生する. 例えば、幾つか典型的な例を表1にま とめて示す.

表1	CT 測定において投影データー部に異常	킭
	データが含まれる例	

X線の照射異常、検出	メタルアーティファ
器の特定素子の異常	クト
ビームハードニング	リングアーティファ
アーティファクト	クト
測定中の被写体一部	位相 CT における位
の急激な変化	相ラップ
暗視野 CT における	投影データ位置合わ
回折コントラスト	せのエラー

このような場合,フィルタ補正逆投影(FBP)法 や SIRT 法/SART 法などの逐次近似法で画像再 構成を行うと,異常データの影響が画像全体に 広がり再構成画像は著しく劣化したものとな る.具体的な例を図1に示す.本研究では、こ のような状況において,画像再構成の過程で異 常データの場所を自動的に検知して除外して 画像再構成を行う新手法を提案する. この枠組 みをその原理から ABD-CT (Abnormal Data Detected CT) と呼ぶ.

少数検出器の投影データが 少数角度の投影データが 異常の場合の再構成画像

異常の場合の再構成画像



図1 少数検出器や少数角度の投影データに異常があ る場合の CT 画像劣化の例(SIRT 法で画像再構 成)

次に、ADB-CTの原理を説明する.画像をx、投影 データを b, 画像と投影データを関係づけるシステ ム行列をAで表すと、通常の逐次近似画像再構成法 では次式の評価関数 f(x)を最小化して画像再構成が 行われる.

$$f(\vec{x}) \equiv \frac{1}{2} \left\| A\vec{x} - \vec{b} \right\|_{2}^{2} (L^{2} / \nu \, \Delta \, \Pi \, \mathrm{H} \, \mathrm{H} \, \mathrm{H} \, \mathrm{H})$$
 (1)

しかし、L²ノルム再構成は投影データに含まれ る誤差が白色ガウス雑音のモデルに基づくた め、投影データの一部に異常値が含まれる場合 除外する効果を持たず,再構成画像は著しく劣 化してしまう. そこで, ADB-CT では, 異常値 を検知して除外する効果を持つことが知られ る L¹ ノルムと L⁰ ノルムを画像再構成の評価関 数に用いる.

$$f(\vec{x}) \equiv \left\| A\vec{x} - \vec{b} \right\|_{1}^{1} (L^{1} / \mu \, \Delta \, \mathrm{時構成})$$
(2)

$$f(\vec{x}) \equiv \left\| A\vec{x} - \vec{b} \right\|_{0}^{0} (L^{0} / \mu \, \Delta \, \mathrm{ቐ構} \, \mathrm{K})$$
 (3)

 L^1 ノルムと L^0 ノルムが異常値を除外する効果 を持つことは機械学習や統計学の分野ではよ く知られ,既に「ロバスト推定」や「Outlier 検

出」などの名称で様々な問題に応用されている. しかし,筆者が知る限り,この性質を CT 画像 再構成に応用した研究は存在せず本研究が初 めてである [1],[2]. また, L^1 ノルムと L^0 ノルム のどちらが優位かは、以下の通りである.投影 データ中の異常データの場所を検知し除外し てデータフィッティングを行う効果は L¹ ノル ムより L⁰ノルムの方が数学的に優れるが, L⁰ノ ルムは非凸関数であり正確な評価関数の最小 化が困難である. 逆に, L¹ノルムは凸関数であ り評価関数の最小化は容易だが、異常データの 場所を検知し除外する効果はL⁰ノルムに劣る. 筆者らが幾つかの事例で実装した結果によれ ば, L⁰ノルムを用い近似解を求める手法が L¹ノ ルムで厳密解を求めるよりも高性能のようで ある.更に,提案手法の異常データ検知し除外 する性能を向上させるため、式(2)、(3)に圧縮セ ンシングの TV 正則化項を加えた以下の評価関 数も提案する.

$$f(\vec{x}) \equiv \beta \|\vec{x}\|_{\text{TV}} + \|A\vec{x} - \vec{b}\|_{1}^{1} \quad (L^{1}\text{-}\text{TV} \, \text{Fr} \, \text{KR})$$
(4)
$$f(\vec{x}) \equiv \beta \|\vec{x}\|_{\text{TV}} + \|A\vec{x} - \vec{b}\|_{0}^{0} \quad (L^{0}\text{-}\text{TV} \, \text{Fr} \, \text{KR})$$

(5)

更なる改良として,実際によく起こる異常デ ータの発生の仕方に事前に既知の制約がある 場合についても検討を行った.例えば,1)同じ 角度の投影データはまとまって異常データに なる,2)同じ検出器で測定したデータはまとま って異常データになる,3)近くのデータは同時 に異常データになりやすい,などの場合である. より具体的には,リングアーティファクトは2) の場合,メタルアーティファクトは3)の場合に 相当する.このような場合,グループ LASSO と 同様の考え方に基づき,評価関数を次式のよう に設計する.

(グループ L¹-TV 再構成)

$$f(\vec{x}) \equiv \beta \|\vec{x}\|_{\text{TV}} + \sum_{m=1}^{M} \|A_m \vec{x} - \vec{b}_m\|_2^1$$
(6)

$$f(\vec{x}) \equiv \beta \|\vec{x}\|_{\rm TV} + \sum_{m=1}^{M} \|A_m \vec{x} - \vec{b}_m\|_2^0$$
(7)

ただし, *A*₁,*A*₂,---,*A*_Mは同時に異常データになる (なりやすい)行を*A*から抜き出して作成した *A*の部分行列, **b**₁,**b**₂,---,**b**_Mは*A*₁,*A*₂,---,*A*_Mに対応 する部分投影データである.即ち,残差ベクト ル**r**=*A***x**-**b**を次式のように*M*個の同時に異常デ ータになる(なりやすい)成分のグループに分 割して,グループ LASSO でデータ項を定義す る.

$$\vec{r} = \begin{bmatrix} A_1 \vec{x} - \vec{b}_1 \ (\mathcal{I}') \mathcal{V} - \mathcal{T}' \mathbf{1}) \\ A_2 \vec{x} - \vec{b}_2 \ (\mathcal{I}') \mathcal{V} - \mathcal{T}' \mathbf{2}) \\ \vdots \\ A_M \vec{x} - \vec{b}_M \ (\mathcal{I}') \mathcal{V} - \mathcal{T}' M) \end{bmatrix}$$
(8)

どのようにグループに分割するかは、問題によ って決まる異常データ発生の仕方を考慮して 決定する.

2) 反復解法

次に,評価関数の最小化を行う反復解法について述べる.もちろん,式(2)~(7)の評価関数の 最小化を正確に行える反復解法であれば,どれ でも ABD-CT に用いることができる.ただし, 高速収束性からブロック反復型(Row-Action型) や FBP 法組み込み型の構造を持つ手法が望ま しい.例えば,筆者らの文献[1],[2]では,ART 法 と類似した構造を持ち実装が容易な Row-Action型の反復法をL¹ノルム再構成とL¹-TV 再 構成に対して提案した.その後継続的に研究を 実施しており,現時点で式(2)~(7)の最小化に問 題なく使用可能な幾つかの反復解法を構築し ている.

3) 事例紹介

既に, ABD-CT を以下の事例に対して適用し て, 異常データを除外して高精度で画像再構成 を行うことに成功している.主には物理計測を 目的とした CT 装置で発生する事例であるが, リングアーティファクトやメタルアーティフ ァクトなど医療用 CT 装置にも応用可能な事例 も含まれる.また,全て計算機によるシミュレ ーションではなく実データでの成功である.

[事例1] 投影角度毎にX線のエネルギーと強度 が異なる極端なビームハードニングを発生す るミリ秒トモグラフィ

[事例 2] 放射光 CT におけるリングアーティフ ァクト

[事例 3] 医療用コーンビーム CT のリングアー ティファクト

[事例 4] 暗視野 CT で問題となる回折コントラ ストアーティファクト

[事例 5]医療用 CT のメタルアーティファクト

紙面の都合上全ては紹介できないため,図2 にリングアーティファクト除去の事例を示す. アルミニウム合金小片の試料を,Spring-8の放 射光を用いて Propagation ベースの位相コント ラスト CT で撮影した実データである.戸田裕 之教授と清水一行助教(九州大学)から提供を 受け,彼らが顕著なリングアーティファクトの 例として文献[3]で紹介しているものである.リ ングアーティファクトの場合,同じ検出器の測 定データがまとまって異常になる性質がある ため,異常データはサイノグラム上では縦線状 のエラーになる.この性質を用いて,式(7)のグ ループ L⁰-TV 再構成を行っている.リングアー ティファクトは上手く除去できている.

3. まとめ

通常,ABD-CTと同様な異常データに影響されない CT 画像再構成を実現しようとすれば, サイノグラム上で異常データの場所を問題(異 常データ要因)に依存する何らかの手法で推定 し、その後異常データを除外して画像再構成を 行う、2 段階の処理が必要となる. ABD-CT は 圧縮センシングと同等な単一の最適化問題を 解くのみでこれを実現しており、異常データ要 因に特化した処理も必要なく様々な異常デー タ要因に統一的に同じ数学的枠組みで対応で きる.

謝辞

本研究は、JST-ERATO(JPMJER1403)及びJST-CREST (JPMJCR1765)の助成を受けた.図2で 使用した実データは、九州大学の戸田裕之教授 と清水一行助教から提供を受けた.

利益相反の有無

なし.

文 献

- H.Kudo, K.Takaki, F.Yamazaki, and T.Nemoto, Proposal of fault-tolerant tomographic image reconstruction, Proceedings of SPIE 2016 Optics + Photonics (Development in X-ray Tomography), Paper No. 99671K, 2016
- H.Kudo, J.Dong, K.Chigita, and Y.Kim, Metal artifact reduction in CT using fault-tolerant image reconstruction, Proceedings of SPIE 2019 Optics + Photonics (Development in X-ray Tomography), Paper No. 111130A, 2019
- [3] 戸田裕之, X線 CT: 産業・理工学でのトモ グラフィー実践活用, 共立出版, 2019



図2 グループ L⁰-TV の評価関数を用いた ABD-CT でリングアーティファクトの除去を行った例 (試料:アルミニウム合金の小片,放射光を用いた Propagation ベースの位相コントラスト CT)

Proposing Concept of Abnormal Data Detected Tomography ABD-CT and Introducing Some Examples

Hiroyuki KUDO*1

*1 University of Tsukuba

In many applications of CT, a part of measured projection data becomes abnormal (erroneous) due to several physical reasons. For example, beam-hardening which arises when the object contains metals inside, ring-artifact which arises due to non-uniform sensitivity of detector, etc. In these cases, reconstructed images by the standard Filtered Backprojection method suffer from severe artifact because the effect of abnormal projection data spreads out over the whole image. In this work, we propose a new framework of CT reconstruction called ABD-CT (Abnormal Data Detected CT) in which locations of the abnormal data in the sinogram are automatically detected and excluded from the data fitting during the image reconstruction. In ABD-CT, we use L^0 norm (or L^1 norm) regularized with Total Variation (TV) penalty $f(\mathbf{x})=\beta |||\mathbf{x}||_{TV}+||A\mathbf{x}-\mathbf{b}||_1^0$, which possesses an effect of excluding the abnormal data. Furthermore, we also propose the following extension. In some cases such as ring artifact and metal artifact, we know that some specified group of projection data becomes abnormal concurrently in advance. For such a case, we design the cost function of image reconstruction by using group L^0 norm (or group L^1 norm) founded by the idea of group LASSO. We introduce several example reconstructed images which demonstrate the power of ABD-CT. These include very strong beam-hardening in millisecond CT, ring artifact in synchrotron radiation CT and medical CT, diffraction-contrast artifact in dark-field CT, and metal artifact in medical CT. In all of the examples, ABD-CT succeeds in suppressing the artifact well.

Key words: CT, Image Reconstruction, Sparse Modeling, Outlier Detection, Metal Artifact

距離マップを利用した肺マイクロ CT 像からの肺胞抽出

椎名 健*1 小田 紘久*2 鄭 通*1 中村 彰太*2

林 雄一郎*1 小田 昌宏*3,1 森 健策*1,4,5

要旨

本稿では、肺のマイクロ CT 像からの肺胞抽出手法を提案する.末梢肺には多様な微細構造が存在するが、 肺胞と呼ばれる袋状の構造が最も多く見られる.肺胞は、切除肺標本からマイクロ CT 像として画像化で きるが、肺胞は複雑かつ三次元的に分布しており、直感的な理解には自動抽出による可視化が重要である. 従来手法では、1) 二値化やオープニング処理による肺胞壁の抽出、2) 肺胞壁からの距離画像をもとにした 肺胞粗抽出 ののち、3) 小成分除去により過分割の修正を行い肺胞を抽出したが、過分割ラベルの残存が 課題であった.本手法では、接触面が大きいラベルの統合を行うことで、過分割をさらに抑制する.切除 肺標本 1 例のマイクロ CT 画像から肺胞領域を切り出した 3 画像に本手法を適用し、定量・定性評 価を行った.正規化相互情報量の平均値は最大で 0.917 であった.本手法により、従来手法に比べ過分割 を抑制できることを確認した.

キーワード:呼吸器,マイクロ解剖学,セグメンテーション,肺胞

1. はじめに

本稿では、肺マイクロ CT 像からの肺胞抽 出について述べる.末梢肺には、終末細気管 支や肺胞といった微細な構造が三次元的に分 布する[1].これらの構造は、マイクロ CT と よばれる高解像度 CT を利用することで三次 元画像化できる.しかし、三次元画像のスラ イスを表示するのみでは、肺の微細構造を直 感的・立体的に理解するのは難しい.解剖構 造の詳しい理解や、病態解明につながる医学 的知見を獲得するために、肺マイクロ CT 画 像中に存在する各々の解剖構造を抽出・可視 化することが望まれる.

*1 名古屋大学大学院情報学研究科
〔〒464-8601 名古屋市千種区不老町〕
e-mail: tshiina@mori.m.is.nagoya-u.ac.jp
*2 名古屋大学大学院医学系研究科
*3 名古屋大学情報連携推進本部
*4 名古屋大学情報基盤センター
*5 国立情報学研究所医療ビッグデータ
研究センター

末梢肺には、「肺胞」とよばれる小胞がもっ とも普遍的にみられる.しかしながら、肺胞 は三次元的に複雑な分布をしており、マイク ロ CT 画像そのものからの観察は困難である. さらに、肺胞壁には解剖学的に孔が存在する ことから、CT 画像中で壁は途切れて映る. そのため、単純な塗りつぶし手法で肺胞を抽 出することは困難である.

肺マイクロ CT 画像からの構造抽出に関す る研究としては、Saito らによる肺胞房抽出[2], Moriya らによる肺がん領域抽出[3], 佐藤らに よる細気管支, 肺胞管, 肺胞嚢, 肺胞の抽出 [4]などが挙げられる. 個々の肺胞に特化した ものとしては、Shiina らによる肺胞抽出[5]が ある. Shiina らは肺胞壁の領域を距離マップ に変換し, Watershed アルゴリズム[6]を用い て肺胞抽出を行う手法を提案している. この 手法では小成分除去で体積が小さいラベルを 削除し, Watershed を再実行することで過分割 を修正した.しかし、Shiina らの手法[5]では, 体積が大きい肺胞の過分割が依然として見ら れた. そこで本稿では、さらに過分割を抑制する 手法を検討する.具体的には、接触面が大き いラベルを統合する処理を追加する.あるラ ベルが別のラベルと一定割合以上の面積で接 する場合、これらのラベルを統合する.この 処理により、肺胞壁が存在しない箇所でのラ ベル分割を抑制できたので報告する.

2. 手法

本手法は、ヒト肺マイクロ CT 画像からク ロップした肺胞画像を入力とし、個々の肺胞 の領域を示すラベルを出力とする.本手法は (1)肺胞壁領域抽出,(2)肺胞壁からの距離画 像をもとにした肺胞粗抽出,(3) 過分割修正 の3 手順からなる.

(1) 肺胞壁領域抽出

入力画像を大津の2値化[7]で2値画像に変換し、おおよその肺胞壁領域を得る. 半径 r[µm]の球を構造要素とするオープニング処理により、肺胞壁領域に含まれる False Positive (肺胞壁に対応しないボクセル)を削減する. 得られた領域のうちの最大連結成分 を最終的な肺胞壁領域とする.

(2) 肺胞粗抽出

入力画像において肺胞壁領域以外を内腔領 域とする.内腔領域に対して肺胞壁領域から のユークリッド距離変換[8]を行い,距離画像 を得る.各々の肺胞中心を,距離画像におけ る極大点と仮定して求める.ただし,極大点 間の最小距離をm [µm]と定める.極大点を 開始点,距離値を高さとして,Watershed アル ゴリズム[6]により内腔領域の領域分割を行 う.

(3) 過分割修正

(2) で得られたラベルのうち,体積 ν [μm³] 未満のラベルを削除する.残存するラベルを 開始点として, Watershed アルゴリズムを再び 実行する.削除されたラベルの存在した領域 を,残存するラベルが埋めることとなる.

さらに、ラベルの統合処理を行う.隣接する2個のラベルが大きく接触しているとき、 両者の間に肺胞壁領域が少ないにも関わらず 誤って分割されているものとみなし、両者を 統合する. ラベルに 1, 2, ..., n の整数値をイ ンデックスとして割り振る. 整数 $i,j \in$ {1,2,...,N} (Nはラベルの総数)に対し, i 番 目と j 番目の各ラベルが接する面積を s(i, j) (ただし $i \neq j$) と表す. また, i 番目の ラベルの表面積を $s_t(i)$ とする. このとき, 過分割とみなす接触割合の最小値を c (0 < c < 1) として,

$$\frac{s(i,j)}{s_{\rm t}(i)} \ge c \tag{1}$$

であれば *j* 番目のラベル を *i* 番目のラベ ルに統合する. すべての *i* に対して, 式(1) をみたす *j* がある限り統合を繰り返し, 過分 割領域の修正を行う.

最後に,画像辺縁に存在するラベルは肺胞 全体を含まないため,削除する.1 画素以上 が画像の辺縁に含まれるラベルを除去し,残 ったラベルを最終的な肺胞ラベルとする.

3. 実験と結果

肺がんの切除肺標本 1 例のマイクロ CT 画 像から肺胞領域をクロップした 3 画像 A, B, C に対し,本手法を適用した.ボクセルは等方 性で,1辺は 5.00 µm であった.パラメータ は経験的に r = 5.00 µm, m = 25.0 µm, v =1.25×10⁵ µm³ とした.そのうえで c を 0.2,0.3,0.4,0.5,0.6 と変化させた.クロップし た画像ごとに 1 スライスを対象として Ground-truth を手動で作成し,正規化相互情 報量 (NMI)により抽出結果の定量評価を行 った.図 2 に示すように,画像 A, B では c =0.3,画像 C では c = 0.4 のとき NMI は最大 であった.3 画像の NMI の平均値は c = 0.3で最大となり,最大値は 0.917 であった.

c = 0.3,0.4 における抽出結果を図1に示す.
また, c = 0.4 における画像 A の抽出結果の ボリュームレンダリング画像を図3に示す.
c = 0.3,0.4 における従来手法[5]との正規化 相互情報量の比較結果を表1に示す.



図1 肺胞抽出結果の例.正解ラベルと各抽出結果でラベルの色は対応しないことに注意されたい.(左)入 力画像の1スライスと正解ラベル.(中)従来手法及び提案手法による抽出結果.(右)抽出結果で破線 により示した領域の拡大図.

4. 考察

図2の通り,3 画像のいずれでも *c* = 0.3 な いし *c* = 0.4 において正規化相互情報量 (NMI) が最大となり, *c* がそれより大きく ても小さくても NMI は低下した.これは, 過 少分割の防止と過分割の修正がトレードオフ であるためだと考える.すなわち, *c* が小さ すぎるとラベル同士のわずかな接触でも統合 が起き, 過少分割を引き起こす.反対に, *c* が 大きすぎれば過少分割は起きにくい一方, 過 分割を修正できない.

図1に示すように、本提案手法により、肺胞 壁に沿って各々の肺胞ラベルが得られること を確認した.図3からは、本手法で得られた ラベルのボリュームレンダリング表示によっ て、肺胞の立体構造や分布を直感的に観察可 能であることがわかる.

図1右に示した拡大図から,従来手法で過 分割だったラベルの一部が統合されているこ とが確認できる.加えて,表1に示すように. 定量評価では,3 スライスのいずれも正規化 相互情報量の値が向上したことから,提案手 法のラベル統合処理により,従来手法より過 分割を抑制できたと考える.

一方,図1からは、依然として1つの肺胞

が複数にラベリングされる箇所も見て取れる. 今回提案したラベル統合処理においては,注 目ラベルに対する各ラベルの接触割合を基準 としている.同程度の大きさの2個を統合す る場合ならば提案手法により対応できるが, 統合すべき抽出結果がさらに細かく多数に分 割されている場合,1つのラベルから他の1 つのラベルへの接触割合が小さくなるため, 統合されにくくなるものと考えられる.

また,既存手法で正しく分割されていたラ ベルが,提案手法では誤って統合された例も みられた(図1において画像Aの抽出結果に 示した赤矢印部分など).肺胞壁が不明瞭な箇 所では肺胞壁領域の欠け(False Negative 領域) が多く,ラベルの接触割合が増加するため, 誤って統合の対象に含まれたものと考える. 対策として,肺胞壁領域抽出の感度を向上さ せ,内腔領域における肺胞どうしの接触面積 の割合を減らすことが考えられる.

5. むすび

本稿では、肺マイクロ CT 画像からの肺胞 抽出手法を提案した.従来手法[5]に接触面が 大きいラベルを統合する処理を追加すること で、過分割が抑制され抽出精度が向上するこ とを確認した.今後は抽出精度のさらなる向 上のため,肺胞壁領域の抽出手法を改良して 感度を向上させることや,ラベルの形状特徴 を考慮することを検討する.

謝辞

日頃から熱心に御討論頂く名古屋大学森研究室の諸氏に感謝する.本研究の一部は堀科学芸術振興財団,JST CREST JPMJCR20D5,
JSPS 二国間事業ならびに JSPS 科研費(17H00867, 17K20099, 21K19898), AMED (JP20lk1010036)によった.

利益相反の有無なし

文 献

- [1] 牛木辰男:入門組織学.南江堂,東京. 2013,pp203-205
- [2] Saito K, Fuketa S, Shimatani R et al: 3D microstructure analysis of human pulmonary emphysema using a synchrotron radiation CT, Proc. SPIE 11317, Medical Imaging 2020, 113170H, 2020
- [3] Moriya T, Roth HR, Nakamura S, et al: Unsupervised pathology image segmentation using representation learning with spherical k-means. SPIE Medical Imaging 2018 Digital Pathology 10581: 1058111, 2018
- [4] 佐藤嘉晃,長尾慈郎,北坂孝幸他:マイ クロCT 画像を用いた肺微細構造からの 肺胞領域の抽出.信学技報 MI,医用画 像 104(90): 49-54, 2004
- [5] Shiina T, Oda H, Zheng T et al: Watershed-based Alveoli Segmentation from Micro-focus X-ray CT Volumes of Dissected Human Lungs. IEICE Tech Rep 121(21): 9-10, 2021
- [6] Vincent L, Soille P: Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations, IEEE Trans Pattern Anal Mach 13(6): 583-598, 1991
- [7] Otsu N: A Threshold Selection Method from

Gray-Level Histograms, IEEE Trans Syst Man Cybern **9**(1): 62-66, 1979

 [8] 斉藤豊文,鳥脇純一郎:3 次元ディジタ ル画像に対するユークリッド距離変換. ICICE J76-D-II:445-453,1993







図3 画像Aの抽出結果のボリュームレンダリング 画像.

表12手法のセグメンテーション結果の 正規化相互情報量の比較

	画像A	画像 B	画像C	平均
従来手法[5]	0.904	0.874	0.891	0.890
提案手法 (<i>c</i> = 0.3)	0.932	0.934	0.884	0.917
提案手法 (<i>c</i> = 0.4)	0.914	0.882	0.892	0.895

Alveoli Segmentation from Micro-CT Volumes of

Lung Specimen using Distance Map

Takeru SHIINA^{*1}, Hirohisa ODA^{*2}, Tong ZHENG^{*1}, Shota NAKAMURA^{*2}, Yuichiro HAYASHI^{*1}, Masahiro ODA^{*3, 1}, Kensaku MORI^{*1, 4, 5}

*1 Graduate School of Informatics, Nagoya University

*2 Nagoya University Graduate School of Medicine

*3 Information and Communications, Nagoya University

*4 Nagoya University Information Technology Center

*5 Research Center for Medical Bigdata, National Institute of Informatics

In this paper, we propose a segmentation method of the alveoli from lung micro-CT volumes. The alveoli are the most common sac-like structures among various microstructures in the peripheral lung. The alveoli can be imaged as micro-CT volumes from dissected lung specimens. The alveoli are complex and three-dimensionally distributed. Visualization by automatic segmentation is important for intuitive understanding of them. In the conventional method, the alveoli segmentation was performed by 1) extraction of the alveolar walls by thresholding and morphological opening, 2) rough segmentation of the alveoli using the distance map from the alveolar walls, and 3) correction of over-division by removing small components. However, there was a problem that over-divided labels were created. In the proposed method, over-division is further suppressed by integrating labels with large contact surfaces. The proposed method was applied to three cropped volumes of the alveolar regions from a lung micro-CT volume for quantitative and qualitative evaluation. The best average score of normalized mutual information was 0.917. We confirmed that the proposed method can suppress over-division as compared with the conventional method.

Key words: Respiratory System, Microanatomy, Segmentation, Alveoli

非ラグランジアン解法を用いた少数方向投影からの

Micro-CT 画像再構成

熊切 俊夫*1 岡本 尚之*2 工藤 博幸*3 羽石 秀昭*4

要旨

病理診断は微細構造の観察に有用であるが、2次元情報しか得られないといった問題点が存在する.そこで 数μmの分解能で3次元構造が観察可能な Micro-CT を導入することで、網羅的な検査が可能になる.一方 で、高画質な Micro-CT 像の取得には長時間撮影が必要となる. CT 撮影の高速化には少数方向からの撮影 が有効であるが取得データ削減により、解析的画像再構成手法(FBP 法)ではスポーク状のアーチファクト が発生してしまう.本研究では少数方向投影条件下における画像再構成手法を Micro-CT に適用することで 画質向上を目指す. 具体的には Total Variation を用いた圧縮センシング再構成手法を Micro-CT 撮影データ に適用することとし、この際の解法としてラグランジアンや双対変数を用いない非ラグランジアン解法に 基づいた手法を用いた. この手法は新型近接勾配法を Ordered-Subsets 反復法に拡張することで厳密解に高 速で収束することが可能であり、尚且つ実装も簡便であるといった特徴がある.実験の結果、FBP 法と比 較しスポーク状のアーチファクトを低減し、鮮明な物体構造が得られた.

キーワード:画像再構成, Micro-CT, スパースビュー, 圧縮センシング

1. はじめに

病理検査は摘出組織から病理標本を顕微鏡 下で観察する検査であり、微細構造の取得に有 用である.しかし標本作成は薄切を伴うため観 察領域が2次元に限定されてしまうという問題 点がある.そこでマイクロメートルオーダで3 次元構造が観察可能な Micro-CT の導入により がん診断の確からしさ向上が期待されている [1].しかしながら Micro-CT には高画質な画像

*1 千葉大学大学院融合理工学府数学情報科学専攻情報工学コース

〔〒263-8522 千葉市稲毛区弥生町 1-33〕 e-mail: t-kumakiri@chiba-u.jp

*2 千葉大学大学院融合理工学府基幹工 学専攻医工学コース

- *3 筑波大学システム情報系
- *4 千葉大学フロンティア医工学センター

取得に長時間撮影が必要であるという課題が ある.解決策としてスパースビュー撮影が挙げ られるが,取得データ削減により解析的画像再 構成手法である FBP 法ではスポーク状のアー チファクトが発生する問題点も存在する.そこ で本研究ではスパースビュー撮影条件下にお ける高画質な再構成画像取得を目的に,Total Variation(TV)を用いた圧縮センシング再構成手 法を Micro-CT 撮影データに適用できるよう検 討を行った.

2. 手法

本研究では工藤が提案した再構成手法を採 用した[2][3]. この手法はラグランジアンや双対 変数を用いずに,厳密な制約条件付き最適化問 題の解を求めることを可能にした.特徴は近接 勾配法を安定して厳密解を求められるよう改 良し,尚且つ OS-SART 法を埋め込むことによ り高速に収束することを可能にした.

Micro-CT の撮影データはノイズが多いため,

従来の最適化問題を解くと過剰適合してしま うという問題点がある.そこでまず最適化問題 を式(1)のように変更した.

 $min_{x,n} \|\mathbf{n}\|^{2} + \beta f_{TV}(\mathbf{x}) \quad s.t. \mathbf{Ax} + \mathbf{n} = \mathbf{b} \quad (1)$ ここで n はノイズ変数, $f_{TV}(\mathbf{x})$ は画像 x におけ
る TV, β は TV パラメータ, A はシステム行列,
b は観測データを示している.

またノイズは OS-SART 法における projection 演算に影響を及ぼすことが知られており, pixel driven や ray driven などが主流であるが, 今回 はより正確な演算が可能である distance driven を適用した[4].

3. 実験·結果

データは public data であるクルミの Micro-CT 像[5]を使用した. このデータは 1201 方向から 撮影されたものである.実験では等間隔に 64 方 向の投影データを選択し, FBP 法, OS-SART 法, OS-SART-TV 法(適用手法)の間で比較を行った. 評価指標は PSNR と SSIM を使用し,正解画像 はフルサンプリングデータの FBP 法再構成画 像に設定した.

各手法の再構成画像を図1 に示す. 正解画像 と比較し FBP 法や OS-SART 法ではスポーク状 のアーチファクトが顕著に現れている.一方, OS-SART-TV 法ではアーチファクトが低減され, より鮮明な画像を出力した.





(c) OS-SART

図1 再構成画像

(d) OS-SART-TV

表1 では各手法における評価値を示す. 各評 価値においてもOS-SART-TV法における大幅な 評価値向上が見られた.

表1 各手法の評価値

	FBP	OS-SART	OS-SART- TV
PSNR	31.58	31.84	46.59
SSIM	0.2813	0.3171	0.4307

4. まとめ

Micro-CT 撮影データに対して圧縮センシン グに基づいた再構成手法を適用した.結果大幅 な画質改善を達成した.

謝辞

本研究の一部は、文部科学省科学研究費補助 金基盤研究(A)(課題番号:19H01172)により行わ れた.

利益相反の有無

なし

文 献

- Katsamenis, O. L., Olding, M., Warner, J.
 A., et al.: X-ray micro-computed tomography for nondestructive three-dimensional(3D) X-Ray histology. The American journal of pathology 189:1608-1620, 2019
- [2] 工藤博幸,制約付き最適化問題をとく 新しい数学的枠組み「非ラグランジアン解 法」の考案,私信
- [3] 遠藤隼人,工藤博幸,制約条件付き凸 最適化問題の非双対解法と画像処理逆問 題への応用,映情学技報 44:55-56,2020
- [4] De Man, B., and Basu, S.: Distance-driven projection and backprojection in three dimensions. Phys Med Biol 49:2463-2475, 2004
- [5] Der Sarkissian H, Lucka F, van Eijnatten, M et al.: A cone-beam X-ray computed tomography data collection designed for machine learning. Sci Data 6: 215 2019

A sparse-view micro-CT reconstruction using non-Lagrangian solution

Toshio KUMAKIRI*1, Takayuki OKAMOTO*1, Hiroyuki KUDO*2, Hideaki HANEISHI*3

*1 Graduate School of Science and Engineering, Chiba University

*2 Information and Systems, University of Tsukuba

*3 Center of Frontier Medical Engineering, Chiba University

Current pathological diagnosis is performed by observing a limited number of 2D images. To complement the 2D only diagnosis, micro-computed tomography (Micro-CT) can be used for comprehensive examination of specimen because it can provide 3D structure with high spatial resolution. However, it takes long scanning time to obtain high quality images. Although reducing the number of projection directions can help shorten scanning time, conventional analytical reconstruction methods such as filtered back projection (FBP) produces serious streak artifacts under such a poor imaging condition. In this report, we apply a total variation-based compressed sensing reconstruction method to sparse-view Micro-CT data. The method includes a non-Lagrangian solution and realizes fast convergence to the exact solution and easy implementation. The experimental results showed that streak artifacts were reduced compared to the FBP method and clearer object structures were obtained by the proposed method.

Key words: Image reconstruction, Micro-CT, sparse-view, compressed sensing

圧縮センシングと深層学習を組み合わせた低線量 CT 画像 再構成法 CSDL-Net のヘリカルコーンビーム CT への拡張

川又 健太郎*1 数尾 昴*1 工藤博幸*1

要旨

低線量 CT の画像再構成に有効な手法として, 圧縮センシング(CS) とディープラーニング(DL) が精力 的に研究され実用化が進んでいる. CS の欠点は画像の滑らかな濃度変化やテクスチャーが消失しやすいこ と, DL の欠点は学習データから学びにくい構造物が消失しやすいことであるとされている. 各々の欠点を 相補に補い合い高画質の画像を得るため, 森ら(工藤,森,JAMIT2020)は, CS と DL を組み合わせて高画 質の画像を生成する CSDL-Net を提案した. 本研究では, CSDL-Net をヘリカルコーンビーム CT の 3 次元 画像再構成に拡張する検討を行った. 拡張した 3D-CSDL-Net では, まず 3 次元画像を FDK 法で再構成し た後に 3DTV でデノイジングを行い, デノイジングなし画像とデノイジングあり画像のペアをスライスご とに出力が1 チャンネルの 2D 畳み込みニューラルネットワーク(CNN)に入力し, 再構成画像を求める. また, デノイジングしたいスライスのみを CNN に入力する場合と, デノイジングしたいスライスとその前 後のスライスを CNN に入力する場合を検討した. シミュレーション実験を行った結果, 3DCT の場合にお いても CSDL-Net が有効であることが明らかになった.

キーワード: CT, 画像再構成, 圧縮センシング, 深層学習

1. はじめに

現時点で、低線量 CT の再構成において有 効だとされている手法に圧縮センシング (Compressed Sensing : CS)と深層学習(Deep Learning : DL)がある. CS は画像がある空間 でスパースであるという仮定のもと再構成を 行う手法であり、CT 再構成にはトータルバ リエーション(TV)などが用いられる. 一方, DL を CT 再構成に用いる典型的な手法[1,2] は投影データbから解析的再構成少法を用い てノイズ等で劣化した画像xを求め、xを畳み 込みニューラルネットワーク(Convolutional Neural Network : CNN)に入力して画質が改善

*1 筑波大学システム情報系

〔〒305-8573 つくば市天王台 1-1-1〕

e-mail: kawamata@imagelab.cs.tsukuba.ac.jp

投稿受付: 2021 年 6 月 18 日

された画像を得る. つまり, CNN はノイズを 除去する後処理として利用される. また, CNN は多数の劣化画像と正解画像のペアか ら学習を行う.

ただし, CS には測定条件が悪い場合に画像 の滑らかな濃度変化やテクスチャーが消失し て階段状アーティファクトが発生しやすいと いう欠点があり, DL には個人差が大きい学 習データから学びにくい構造物(特に血管な どの低コントラスト構造物)が消失しやすい という欠点があるとされている[3].

そこで森らは CS と DL の欠点を相互に補 いより高画質な画像を生成する CSDL-Net を 提案した[3].

2. 提案手法

本研究では, CSDL-Net を現在普及している CT 装置の構成方式であるヘリカルコーン ビーム CT の3次元画像再構成に拡張する. 拡張した CSDL-Net は次の手順からなる.

まず,投影データから FDK 法を用いてNス ライスを持つ三次元画像 $X = \{x_1, x_2, \dots, x_N\}$ を 再構成する. その後, Xに 3DTV デノイジン グを適用しデノイジングされた画像Y= {y₁, y₂,..., y_N}を得る. X, Yを得た後にX, Yか らスライスごとに取り出し, チャンネル方向 に結合して CNN に入力して再構成スライス z;を得る. 最後に得られたz;を結合し, 再構 成画像 $Z = \{z_1, z_2, \dots, z_N\}$ を得る. このとき, CNN に入力するスライスが1スライスのみの 場合, CSDL-Net の単純な拡張である. ただ し,3次元 CT では2次元 CT と異なり前後の スライスの情報を取得できるため, CNN の入 力に再構成を行いたいスライスとその前後の 計3スライスを使用することでさらなる画質 改善を期待できる.よって本研究では CSDL-Net の拡張として 6 チャンネルの画像 を CNN に入力する方法を提案する.

3. シミュレーション実験

提案手法の性能を評価するシミュレーショ ン実験を、医療用腹部 CT40 例の実画像デー タセットを用いて行った. この CT 画像につ いて、低線量ヘリカルコーンビーム CT のシ ミュレーションを行い肝臓の上から 84 スラ イスを FDK 法で再構成したものと TV デノイ ジングをしたものを用意し、元の画像と合わ せてデータセットとした. このデータセット のうち 30 例を学習データ、10 例を評価に用 いるテストデータとした. また、CNN として Unet[4] 、学習の際の損失関数は絶対値誤差、 最適化アルゴリズムは Adam[5]を使用した. この実験では次のような再構成法と比較した.

- FDK 法
- FDK 法+3DTV (CS 再構成法)
- CNN に FDK 法で再構成した画像を1ス
 ライスずつ入力
- CNN に FDK 法と CS 再構成法で再構成 した画像を1スライスずつ入力
- CNN に FDK 法で再構成した画像を 3 スライスずつ入力

 CNN に FDK 法と CS 再構成法で再構 成した画像を 3 スライスずつ入力 (提 案手法)

表1にテストデータにおける各手法の平均 PSNRとSSIMの値を示す.数値的評価では, CNNを使用した手法のほうがFDK法やCS 再構成法より良い画質が得られている.また, CNNにFDK法で再構成した画像のみを入力 した場合よりCS再構成法で再構成した画像 も加えて入力した場合のほうが良い画質が得 られている.そして,再構成したいスライス のみを入力とした場合よりその前後のスライ スも入力した場合のほうが良い画質が得られ ていることも確認できる

図1にテストデータを再構成した画像のうちの1スライスのみを示す.再構成結果を見ると,FDK法,CS再構成法を用いて再構成 をした場合よりDLを用いて再構成した場合の方が構造を忠実に再現できていることがわかる.また,DLを用いた際にもCS再構成法を用いて再構成した画像を入力した場合の方が,細かい構造物を復元できていることが確認できる.また,入力するスライスを増やすことでさらなる画質改善効果が得られている.

4. まとめ

本研究では CSDL-Net を三次元 CT である ヘリカルコーンビーム CT に拡張した.その 拡張の際に三次元 CT は前後のスライスが取 得可能であることを利用し,さらなる画質改 善を試みた.ヘリカルコーンビーム CT にお いても CSDL-Net は有効であり,前後のスラ イスも加えて入力することでさらなる画質改 善ができることが確認できた.

謝辞

本研究は JST-CREST (JPMJCR1765) の助 成を受けた.

利益相反の有無

なし

	PSNR	SSIM
FDK 法	22.878	0.8021
CS 再構成法	26.413	0.8870
CNN+FDK 法 1 スライス	31.022	0.9512
CNN+FDK 法+CS 1 スライス	31.541	0.9545
CNN+FDK 法 3 スライス	31.963	0.9578
CNN+FDK 法+CS 3 スライス	32.474	0.9606

文 献

- [1] K.H.Jin, M.T.McCann, E.Froustey et al, Deep Convolutional Neural Network for Inverse Problems in Imaging, IEEE Trans Image Process, 26, 4509-4522, 2017.
- [2] H.Chen, Y.Zhang, W.Zhang et al, Low-dose CT via Convolutional Neural Network,

Biomed Opt Express, 8, 679-694, 2017.

- [3] 工藤博幸,森和希:深層学習を用いた圧縮
 センシング CT 画像再構成の画質改善,
 第 39 回日本医用画像工学会(JAMIT)
 大会,2020年
- [4] O.Ronneberger, P.Fischer, T.Brox, U-net: Convolutional networks for biomedical image segmentation, Medical Image Computing and Computer-Assisted ntervention, Springer, 2015 (2015).
- [5] D.P.Kingma and J.Ba, Adam: A method for stochastic optimization, In Proceedings of the 3rd International Conference on Learning Representations (ICLR) (2015).



図1 再構成画像例

Extension of Combined Compressed Sensing and Deep Learning Approach for Low-dose CT Image Reconstruction to Helical

Cone-beam CT

Kentaro KAWAMATA*1, Subaru KAZUO*1, Hiroyuki KUDO*1

*1 University of Tsukuba

Compressed sensing (CS) and deep learning (DL) have been energetically studied and put into practical use as effective methods for image reconstruction of low-dose CT. The drawbacks of CS are that smooth density changes and textures of images are likely to disappear, and the drawbacks of DL are that structures that are difficult to learn from training data are likely to disappear. Mori et al. (Kudo, Mori, JAMIT2020) proposed CSDL-Net, which produces high-quality images by combining CS and DL to complement each defect and obtain high-quality images. In this study, we investigated the extension of CSDL-Net to 3D image reconstruction of helical cone-beam CT. In the expanded 3D-CSDL-Net, a 3D image is first reconstructed by the FDK method and then denoized on a 3DTV. Then a pair of images without denoising and images with denoising is inputted by 2D convolutional neural network and obtain the reconstructed image. We also examined the case where only the slices to be denoized are input to the CNN and the cases where the slices to be denoized and the slices before and after it are input to the CNN. As a result of simulation experiments, it became clear that CSDL-Net is effective even in 3DCT.

Key words: CT, Image Reconstruction, Compressed Sensing, Deep Learning

Pix2pix を用いた任意形状の肺がん CT 画像生成及び

CT 画像様ノイズ再現の試み

遠田 涼*1 寺本 篤司*1 辻本 正和*2 外山 宏*3

近藤 征史*3 今泉 和良*3 齋藤 邦明*1 藤田 広志*4

要旨

現在,人工知能による画像生成技術である敵対的生成ネットワーク(GAN)の応用が,医用画像に 対しても進められている.現状は,生成画像の多様性の問題から,応用は症例の豊富な疾患に留 まっている.任意の病変形状の画像生成が可能であれば,希少疾患に対する応用も実現し得ると 考えられる.そこで本研究では,GANの派生型である Pix2pix を利用し,スケッチにより病変 形状を制御した画像生成を行うことを目的とした.また,GANの生成画像においては,本来あ るべきノイズが消失する課題もあるため,ノイズの再現も検討事項とした.自然画像の場合,画 像を複数のパッチに分割して識別器に入力することで,生成画像の品質向上が確認されているた め,この手法の有無で比較を行った.初期的な検討として,症例が豊富かつ多様な病変形状を呈 す肺がんのCT 画像を対象とし,画像生成が可能であることを確認した.また,パッチ非分割時 にCT 画像様ノイズが良好に再現された.

キーワード:肺がん, CT, 敵対的生成ネットワーク, 画像生成

1. はじめに

医用画像を対象として応用が進む人工知能 技術の1つに,精巧な画像を生成可能な敵対 的生成ネットワーク(GAN)が挙げられる.そ の主な目的は,他の人工知能タスクの性能を 向上させるためのデータ拡張であり,大量の 画像を生成することに焦点が当てられている. 一方で,GANの画像生成は不安定なことも多 く,類似した被写体形状の画像に偏った生成

〔〒470-1192 愛知県豊明市沓掛町田楽 ヶ窪 1-98〕

e-mail: 82020305@fujita-hu.ac.jp

- *2 藤田医科大学病院
- *3 藤田医科大学医学部
- *4 岐阜大学工学部

がなされる,モード崩壊に代表される課題を 抱えている [1]. そのため、データ拡張を目 的とはしつつも,その応用の範囲は症例数の 豊富な疾患に留まっている「2-47. 任意の被 写体形状の画像生成が可能であれば,従来 GAN が応用されてこなかった,希少疾患への 適用も実現しうると考えられる. そこで本研 究では、GANの派生技術であり、画像間の変 換が可能な Pix2pix [5] を用いて、スケッチ により直接病変形状を制御した画像生成を行 うことを目的とする.スケッチに基づく病変 像生成の実現可能性を探るべく、比較的症例 数が豊富かつ多様な形状を呈す肺がんの CT 画像を初期的な解析対象として選定した.ま た,実画像に存在するノイズや摂動が適切に 再現されない[6] という GAN の課題を踏ま え、ノイズの再現も検討事項とする.

^{*1} 藤田医科大学大学院保健学研究科

2. 方法

133 症例の肺がん患者の胸部 CT 画像から, 腫瘍領域とその周辺のみを 3 次元関心領域 (VOI)として抽出した.各 VOI から, 腫瘍中 心の Axial 断面を取り出し,全体の 20%をテ スト,残りを学習に使用した.Axial 断面を基 準に,スライス角度を変化させた 2 次元画像 を追加作成し,学習用画像を増量した.

Pix2pix の入力用に, 膨大な枚数のスケッチ を作成することは現実的ではないため, その 代替として, Canny 法により 2 次元画像から 検出したエッジ像を用いた.エッジ像を入力 し, Pix2pix による画像生成を行った.通常, GAN の生成器は単純なデコーダ構造をとる が, Pix2pix においては一般に U-net [7] が用 いられる.また,自然画像を対象とした研究 において,識別器入力時,画像を複数のパッ チに分割する PatchGAN を採用することで, 生成画像の品質向上が報告されている.本研 究でもこの手法の有無による比較を行った.

3. 結果と考察

テスト用データでの生成結果を図1に示す. 腫瘍の形状は概ね良好に再現された.一方で, 胸壁などの周辺構造や,腫瘍内部に空洞を含 むパターンの再現は一部困難であった.入力 画像が単純な線画であるため,スケッチの各 領域と解剖学的構造の対応付けが困難であっ たと考えられる.本研究の主たる目的である, 任意形状の病変像の生成は実現可能であると 考える.また,PatchGANの採用により,生 成画像から CT 様のノイズが消失したため, 本研究においてはパッチ分割の優位性は確認 されなかった.



図1 テスト用データでの画像生成結果

4. 結論

本研究では、スケッチにより形状を制御した、肺がんCT画像をPix2pixにより生成した. 周辺構造などの再現には改善の余地が残されているものの、腫瘍自体は明瞭に再現がなされ、任意形状の病変像生成の実現可能性が示された.また、本手法におけるノイズ再現において、PatchGANの優位性はなかった.

利益相反の有無

なし

文 献

- Srivastava A, Valkov L, Russell, C et al: VEEGAN: Reducing Mode Collapse in GANs using Implicit Variational Learning. Adv Neural Inf Process Syst 30: 3308-3318, 2017
- [2] Frid-Adar M, Klang E, Amitai M et al: Synthetic data augmentation using GAN for improved liver lesion classification. ISBI 2018: 289-293
- [3] Han C, Kitamura Y, Kudo A et al: Synthesizing Diverse Lung Nodules Wherever Massively: 3D Multi-Conditional GAN-Based CT Image Augmentation for Object Detection. 3DV 2019: 729-737
- [4] Shorten C and Khoshgoftaar TM: A survey on Image Data Augmentation for Deep Learning, J Big Data 6:60, 2019
- [5] Isola P, Jun-Yan Z, Tinghui Z et al: Image-to-Image Translation with Conditional Adversarial Networks. CVPR 2017: 1125-1134
- [6] Karras T, Laine S, and Aira T: A Style-Based Generator Architecture for Generative Adversarial Networks. CVPR 2019: 4396-4405
- [7] Ronneberger O, Fischer P, and Brox T:
 U-Net: Convolutional Networks for Biomedical Image Segmentation. MICCAI 2015: 234-241

Shape-Manipulable Lung Cancer CT Image Synthesis

and Representation of CT-Specific Noise Using Pix2pix

Ryo TODA^{*1}, Atsushi TERAMOTO^{*1}, Masakazu TSUJIMOTO^{*2}, Hiroshi TOYAMA^{*3} Masashi KONDO^{*3}, Kazuyoshi IMAIZUMI^{*3}, Kuniaki SAITO^{*1}, Hiroshi FUJITA^{*4}

- *1 Graduate School of Health Sciences, Fujita Health University
- *2 Fujita Health University Hospital
- *3 School of Medicine, Fujita Health University
- *4 Faculty of Engineering, Gifu University

Generative Adversarial Networks (GAN), an artificial intelligence-based image generation technique, is now being applied to medical image processing. Currently, due to problems with the diversity of generated images, the application is limited to diseases with abundant cases. If it is possible to generate images of arbitrary lesion shapes, it can be possible to apply GAN to rare diseases. In this study, we aimed to generate images with controlled lesion shape by sketching, using Pix2pix, a variant of GAN. In addition, since there is a problem that the noise included in real images disappear in generated images of GAN, the reproduction of the noise was also considered. In the case of natural images, it has been confirmed that the quality of the generated image is improved by dividing the image into multiple patches and inputting them to the discriminator. Therefore, the quality of generated images was compared with and without this method. As an initial study, CT images of lung cancer, which has many cases and various lesion shapes, were used for the analysis, and it was confirmed that image generation was possible. In addition, the CT image-like noise was well reproduced when the image was not divided into patches.

Key words: Lung cancer, Computed tomography, Generative adversarial networks, Image synthesis

コーンビーム CT の散乱線除去のための機械学習ネットワーク

村田 一心*1 尾川 浩一*2 臼井 桂介*3

要旨

本研究では、コーンビーム CT 撮影で発生する散乱線を除去するために適した機械学習ネットワ ークを提案する. コーンビーム CT は散乱線の影響を強く受けるため、再構成画像の画質が劣化 しやすい. そのため、散乱線を精度良く除去するためのソフトウェアが必要とされている. 先行 研究で我々は、機械学習が散乱線除去に有効であることは示したものの、ネットワーク構成の最 適化は考慮できていなかった. そこで本研究では、散乱線の特性とネットワークの性質との関係 を考察することにした. 散乱線の特性を考慮するため、通常の投影画像と垂直方向の画像を同時 に入力することで、擬似的に3次元空間を考慮できるようにした. さらに,通常の畳み込み層に 加え、全結合層を導入した. ネットワークの学習および評価には、モンテカルロ法で作成した画 像を用いた. その結果、提案ネットワークは従来より高精度で散乱線を除去できることがわかっ た.

キーワード:コーンビーム CT, 散乱線, 機械学習

1. はじめに

コーンビーム CT は2次元検出器アレイを 用いることで、効率よく被写体の断層画像を 取得することができる.そのため、画像誘導 放射線治療に活用されている.しかし、コー ン角が広がることで、散乱線の混入量が大き くなり、画質が劣化し、治療計画 CT のよう な定量性が重視される場合には適用できな い.

そのため、これまでコーンビーム CT の散 乱線を除去する研究は多く行われてきた. モ

*1 国立天文台

〔〒181-8588 東京都三鷹市大沢 2-21-1〕
e-mail: kazumi.murata@nao.ac.jp
*2 法政大学 理工学部
*3 順天堂大学 保健医療学部
投稿受付: 2021 年 6 月 18 日

ンテカルロ法[1]やカーネルを用いた方法[2] だけでなく、機械学習を用いた方法も開発さ れている[3,4].特に機械学習を用いた方法は、 精度が高いだけでなく、計算時間を極めて短 くできるため、非常に有用であると言える.

先行研究において我々は、簡単な機械学習 ネットワークを作成し、その散乱線除去性能 を調査した.モンテカルロ法による散乱線あ り・なしのデータを使用し、機械学習が高精 度な散乱線除去に有効であることを示した. しかし、散乱線除去におけるネットワーク構 成の最適化については考慮していなかった.

そこで本研究では、散乱線の除去性能とネットワーク構成との関係を調査する.特に、 垂直方向投影を組み込むことで性能の向上を 試みる.散乱線のパターンは被写体の3次元 構造に強く依存するはずであるが、3次元デ ータを学習に用いると計算コストが膨大になってしまう.一方,2次元投影画像だけを使用する場合,立体構造を考慮することはできない.そこで本研究では,投影画像における散乱除去に,垂直方向の投影データも用いることで,擬似的に立体構造を考慮できるようにした.

2. 散乱線除去ネットワーク

ネットワーク構成は U-net をベースに設計 した.散乱線パターンは高周波成分が小さく, 圧縮による情報量抽出が向いているためであ る.投影画像のサイズは 200×200 画素とし, 3 つの畳み込み層により 25×25 に圧縮した. 圧縮に伴い,各層でチャンネル数は倍増させ た.その後,全結合層に接続し,同サイズの 画像を出力させた.この層では,散乱線のパ ターンが画像全体と相関することを想定して いる.さらに,同じ処理を垂直方向の投影画 像にも適用し,両者の画像を結合させた.結 合したデータは3つの畳み込み層で復元させ, 200×200の画像を出力させた.3つの畳み込 み層および全結合の後には,活性化関数・ PReLUを適用した.

本ネットワークの有効性を確認するため, 垂直方向のデータを使用しないものも作成した.両者の違いは垂直方向データの結合がないことだけとした.

3. シミュレーション方法

構築したネットワークの有効性は、シミュ レーションデータを用いて検証した.モンテ カルロ法により散乱線あり・なしのデータセ ットを作成し、散乱線の除去性能を調査した. データ生成のシミュレーション条件は以下の 通りである.X線管はファントムの中心から 32 cm離した位置に、検出器はその反対側12 cmの位置に設置した.投影方向は1度ずつず らし、360 方向のデータを取得した.管電圧 はデータ数の確保のために90,100,130 kVの 3 つとし、一投影あたりの光子数は10億個と した.検出器のピクセルサイズは0.2 cm×0.2 cmとし、ピクセルあたりの光子数は16000 個 程度となった.

使用したファントムは実際の CT 画像で得 られた断層画像を用いた. CT 値から軟部組織 と骨に分離し,その減衰係数は水とカルシウ ムで近似した.

ネットワークの学習は、プライマリ光子だけで作成された投影画像を教師、散乱線を含む画像を入力、損失関数は MSE として行った.最適化はAdam、エポック数は3000、バッチサイズは 36 とした.データセットの総数は1080枚、そのうち108枚をテストデータとした.画質評価は再構成画像上で行った.再構成にはFDK 法を用いた.

4. 結果および議論

図1に再構成画像,図2に元画像との差分 を示す.散乱線を補正する前には CT 値に大 きな差異がある一方,ネットワークを通して 散乱線を除去した画像(b, c)は元画像との残 差が小さいことがわかる.しかし,垂直方向 の投影データを用いなかった場合(b)には,そ の補正精度は低く,残差にパターンが見られ る.それに対し提案手法による補正(c)には目 立った残差パターンは見られず,高精度で補 正できていることを示している.補正精度を 定量的に示すと,補正前,垂直データのない 補正,提案手法の MAE がそれぞれ,ΔHU = 17.7, 10.1, 8.3 となり,提案手法の有効性が確 認された.

5. まとめ

本研究では、コーンビーム CT の散乱線の 除去性能と機械学習ネットワークの構造との 関係を調査した.散乱線のパターンをより正 確に推定するため、垂直方向の投影データを ネットワークに組み込んだ.モンテカルロ法 で作成したデータセットを用いて検証したと ころ、本手法が散乱線を高精度で除去できる ことを視覚的かつ定量的に確認した.

利益相反の有無

なし



図 1 再構成画像の比較.(b) は通常のU-netベースのネッ トワークで散乱線を補正した結果.(c) はそれに加え, 垂 直方向の投影データをネットワークに入力した結果を示 す.



図 2 元画像と再構成画像との差分. その他の条件は図 1 と同じである.

文 献

- Xu Y., Bai T., Yan H., et al: A practical Cone-beam CT Scatter Correction Method with Optimized Monte Carlo Simulations for Image-Guided Radiation Therapy. Phys. Med. Biol., 60 (9), pp.3567-3587, 2015
- [2] Sun M, Star-Lack JM. Improved scatter correction using adaptive scatter kernel superposition. Phys Med Biol. 55, 6695–6720, 2010
- [3] Nomura Y., Xu Q., Shirato H., et al., Projection-domain scatter correction for cone beam computed tomography using a residual convolutional neural network. Med. Phys. 46 (7), 2019
- [4] Liu Y., Lei Y., Wang T., et al, CBCT-based synthetic CT generation using deep-attention cycleGAN for pancreatic adaptive radiotherapy, Med Phys., 47(6), 2020

A neural network for scatter correction in corn-beam CT

Kazumi MURATA*1, Koichi OGAWA*2, Keisuke USUI*3

*1 National Astronomical Observatory of Japan

*2 Faculty of Science and Engineering, Hosei University

*3 Department of Radiation Oncology, Juntendo University

We propose a neural network for correcting scattered photons in a cone beam CT system. Since the images are severely degraded due to the scattered photons, methods that accurately remove them are required. In our previous research, we showed the efficacy of a machine learning based method, but we did not optimize the network structure. Hence, in the current work, we investigate the relation between the network structure and performance of scattered-photon removal. Considering the characteristics of scattered photons, we input two projection images whose directions are perpendicular to each other into the network at the same time. We used both convolution and fully connected layers. Learning and evaluation of the network were conducted with a dataset produced with a Monte-Carlo simulation. We confirmed that the proposed network more accurately removes the scattered photons compared with a conventional network.

Key words: Cone beam CT, scattered photon, machine learning

Clinical CT Super-resolution Utilizing

Registered Clinical – Micro CT Database

Tong ZHENG^{*1}, Hirohisa ODA^{*1}, Yuichiro HAYASHI^{*1} Shota NAKAMURA^{*2}, Masaki MORI^{*3}, Hirotsugu TAKABATAKE^{*4}, Hiroshi NATORI^{*5}, Masahiro ODA^{*6,1}, and Kensaku MORI^{*1,7,8}

Abstract

This paper proposes a super-resolution (SR) method, for performing SR of clinical CT images utilizing newly-built micro CT – clinical CT images database. Precise non-invasive diagnosis of lung cancer typically utilizes clinical CT images for making surgical plans. Due to the resolution limitations of clinical CT (about $500 \times 500 \times 500 \ \mu\text{m}^3$ / voxel), it is difficult to obtain enough pathological information, such as the invasion condition at alveoli level. On the other hand, micro CT scanning allows the acquisition of volumes of lung specimens with much higher resolution ($50 \times 50 \times 50 \ \mu\text{m}^3$ / voxel) of upages are required for training SR models. Thus, we build a paired low resolution (LR) and high-resolution (HR) images are required for training SR models. Thus, we build a paired clinical CT (LR) and micro CT (HR) dataset by conducting non-rigid registration of clinical CT images that is trained stably on the newly-built dataset. With PSNR score of 23.26, experimental results showed that proposed method qualitatively and quantitatively outperformed conventional supervised SR method.

Keywords : super-resolution, micro CT, lung micro-anatomy

1. Introduction

Super-resolution (SR) is a common topic of medical image processing. The SR aims to reconstruct a high-resolution (HR) image from its corresponding low-resolution (LR) image. In medical imaging, due to hardware and physics

*1 Graduate School of Informatics, Nagoya University

[Furo-cho, chikusa-ku, Nagoya, 464-8601, Japan]

e-mail: tzheng@mori.m.is.nagoya-u.ac.jp

- *2 Nagoya University Graduate School of Medicine
- *3 Sapporo-Kosei General Hospital
- *4 Sapporo Minami-sanjo Hospital
- *5 Keiwakai Nishioka Hospital
- *6 Information Strategy Office, Information and Communications, Nagoya University
- *7 Information Technology Center, Nagoya University
- *8 Research Center for Medical Big Data, National Institute of Informatics



Fig. 1 Workflow of building aligned LR-HR dataset. The workflow consists of three steps: 1) scanning lung specimens using a clinical CT scanner and a micro CT scanner, respectively; 2) selecting feature points from corresponding clinical CT – micro CT volumes and perform affine transform – based registration to obtain roughly aligned clinical CT – micro CT volumes and 3) utilizing deedsBCV [1] to obtain finely-aligned clinical CT – micro CT volumes from the poorly aligned clinical CT – micro CT volumes.

limitations, HR medical imaging comes at the cost of long scan time [2]. Thus, since SR could reconstruct HR images only with LR images, SR of LR images is an optional approach.

Previous SR method commonly need aligning paired LR-HR images to train a fully convolutional network (FCN) for SR [3]. Thus, we firstly obtain paired LR (clinical CT) – HR (micro CT) images by conducting registration between clinical CT and micro CT volumes. Then we propose an SR framework, training with registrated LR-HR images. We consider that an image could be decomposed into two parts: one is a content code (Ccode) that captures the content of the image, another is a style code (Scode) that captures the image style and anatomy difference between LR and HR images. Based on this assumption, we designed an SR framework that could train on LR-HR images from different modalities (clinical CT and micro CT). We performed qualitative and quantitative evaluation using aligned clinical CT-micro CT image pairs.

2. Methods

1) Overview

Our proposed method performs SR of a given lung clinical CT (LR) image. Our method employs a step-by-step framework. The first step is building a cross-modality paired clinical CT (LR) –micro CT (HR) dataset by registration of corresponding clinical CT and micro CT volumes. The second step is training an FCN for SR, using the aligned LR-HR image pairs. There is registration error between clinical CT and micro CT images caused by their modality difference and registration algorithm's low precision, thus the trained FCN outputs a coarse SR image while given a clinical CT (LR) image. Thus, we propose the third step. Input of the third step is an LR image and a latent variable. We optimize the latent variable, and output is an optimized latent variable. The optimized latent variable and input LR image outputs an optimal SR result. Note that the optimal SR image is the final SR result.


Fig. 2 The network structure of proposed method. We utilize poorly-aligned LR-HR images for training. Note that aligned LR-HR images have the same content code (Ccode) but different style code (Scode). The structure of this network was inspired by MUNIT [6].

2) Building aligned LR-HR dataset

Datasets commonly used for SR, such as DIV2K [4] are consisted of color photographs; and there is no existing crossmodality LR-HR dataset consisted of medical images. Therefore, we build a clinical CT (LR) – micro CT (HR) dataset by scanning human lung specimens by a clinical CT scanner and a micro CT scanner, respectively. For obtaining paired clinical CT – micro CT images, we marked feature points of corresponding clinical CT – micro CT images and calculated affine-translation matrix using corresponding feature points. Then we conducted affine-translation of the whole clinical CT volume using the affine-translation matrix. Further, we utilized a non-rigid registration tool named deedsBCV [1] to refine the registration result. Registrated clinical CT – micro CT volumes constitutes the aligned LR-HR dataset. The workflow of building aligned LR-HR dataset is illustrated in Fig. 1.

3) Training FCN for SR of clinical CT images

Given a pair of LR-HR images, we assume that the LR and HR images could be decomposed into two latent representations: one is a CCode that captures domain-invariant characteristics of both LR and HR image. Another is a SCode that captures domain-specific characteristics of LR and HR imag. The domain-specific characteristics consist of image deformation, registration error, etc. Thus, we could assume CCode of paired LR and SR images to be similar to each other, while their SCode are different. Due to this assumption, we propose an FCN to perform super-resolution of LR images while trained on poorly-aligned LR-HR dataset.

The structure of proposed FCN is shown in Fig. 2. First, a style encoder network E_1^s consists of four Resblocks [5] and downsampling convolutional layers that encodes a LR image into Scode $s^{LR} = E_1^s(x^{LR})$. Simultaneously, a content code encoder E_1^c with average pooling layers and four Resblocks encodes the LR image into Ccode $c^{LR} = E_1^c(x^{LR})$. Next, the Ccode c^{LR} will input into a decoder network D_1 . D_1 is an FCN with four residual blocks and convolutional layers that upsamples the ccode to an SR image. The residual blocks are equipped with adaptive instance normalization (AdaIN) [6]. A Scode \hat{s}^{HR} sampled from a multidimensional normal distribution $N_k(0,1)$, passes through a multilayer perceptron and generates parameters of the AdaIN layers. Finally we obtain a HR output $\hat{x}^{HR} = D_1(c^{LR}, \hat{s}^{HR})$. Note that \hat{x}^{HR} is actually the SR result x^{LR} we desire. Symmetrically, we use a style encoder network E_2^s to encode a HR image x^{HR} into Scode $s^{HR} = E_2^s(x^{HR})$, and we use a content encoder network E_2^c consists of four Resblocks and downsampling convolutional layers to encode x^{HR} into Ccode $c^{HR} = E_2^c(x^{HR})$. Then c^{HR} and a Scode \hat{s}^{LR}



Fig. 3 Qualitative results of real clinical CT. (a) blood vessels, (b) bronchus. While EDSR [9]'s result looks blurred and noisy, result of proposed method has sharper bronchus edges and higher sampled from a multidimensional normal distribution $N_k(0,1)$, are input into a decoder network D_2 that has four Resblocks and upsampling convolutional layers to obtain a LR image \hat{x}^{LR} .

Our loss function contains a code loss that ensures the Ccode of paired LR and HR images to be similar to each other and a SR loss that ensures the SR image x^{SR} to be similar enough with its HR counterpart x^{HR} . Other loss terms contain adversarial loss and consistency loss.

The code loss is formulated as

$$L_{\rm Co} = \mathbb{E}[\|E_1^{c}(\mathbf{x}^{\rm LR}) - E_2^{c}(\mathbf{x}^{\rm HR})\|_2^2],\tag{1}$$

where $||||_2^2$ is the square of 2-norm. While poorly-aligned LR and HR images have differences, they share the same anatomical structure which could be projected to the Ccodes $c^{LR} = E_1^c(x^{LR})$ and $c^{HR} = E_2^c(x^{HR})$. Thus, we need to design a loss term to let c^{HR} and c^{LR} approximate each other. Thus we design the code loss L_{Co} . Code loss is utilized to optimize weight of encoders E_1^c and E_2^c .

The SR loss is formulated as

$$L_{\rm SR} = \mathbb{E}[\|\boldsymbol{x}^{\rm HR} - D_1(E_1^c(\boldsymbol{x}^{\rm LR}), \hat{\boldsymbol{s}}^{\rm HR})\|_2^2], \qquad (2)$$

where all the notations are mentioned above. The *SR loss* is commonly used in supervised super-resolution: given an LR image \mathbf{x}^{LR} as input, we desire its SR output $D_1(E_1^c(\mathbf{x}^{LR}), \hat{\mathbf{s}}^{HR})$ to match the corresponding HR image \mathbf{x}^{HR} . SR loss is utilized to optimize weight of encoder E_1^c and D_1 .

Additionally, we also utilize *adversarial loss* L_A and *consistency loss* L_C for training. The adversarial loss forces the SR image $\mathbf{x}^{SR} = D_1(E_1^c(\mathbf{x}^{LR}), \mathbf{\hat{s}}^{HR})$ to have close perceptual quality with its HR counterpart \mathbf{x}^{HR} , and the *consistency loss* ensures encoder E_1^c and decoder D_1 are inverse to each other, as well as encoder E_2^c and decoder D_2 . The *adversarial loss* L_{SR} and *consistency loss* are of the same definition with MUNIT [7] to enhance result's perception quality and stability of training. Adversarial loss is used to optimize encoders E_1^c , E_2^c ; decoders D_1 and D_2 .



Clinical CT image Proposed method Micro CT image (ground truth)

Fig. 4 Quantitative results. Inputting clinical CT image of lung specimen (left side), we obtain an SR result (middle). Then we compared the SR result with its corresponding micro CT image (right side) for quantitative evaluation.

We combined the *code loss*, *SR loss*, *adversarial loss*, *consistency loss* to train the FCN in Fig. 2 for SR. We add losses by weights λ_1 , λ_2 , λ_3 , λ_4 to obtain the total loss L_{Total} to train the network,

$$L_{\text{Total}} = \lambda_1 L_{\text{Co}} + \lambda_2 L_{\text{SR}} + \lambda_3 L_{\text{C}} + \lambda_4 L_{\text{A}}.$$
(3)

4) Super-resolution by latent space exploration

After training the FCNs for SR in section 3), we obtain the trained encoder E_1^c and decoder D_1 . Therefore, given an LR image \mathbf{x}^{LR} , we obtain its SR counterpart by $\mathbf{x}^{SR} = D_1(E_1^c(\mathbf{x}^{LR}), \mathbf{\hat{s}}^{HR})$. Since $\mathbf{\hat{s}}^{HR}$ is the Scode that determines the "style" (image appearance) of the image, we need to choose an $\mathbf{\hat{s}}^{HR}$ that minimizes such function,

$$L_{\rm SR} = \|\boldsymbol{x}^{\rm HR} - D_1(E_1^{\rm c}(\boldsymbol{x}^{\rm LR}), \hat{\boldsymbol{s}}^{\rm HR})\|_2^2 = \|\boldsymbol{x}^{\rm HR} - D_1(\boldsymbol{c}^{\rm LR}, \hat{\boldsymbol{s}}^{\rm HR})\|_2^2, \tag{4}$$

in which D_1 and E_1^c are trained an cannot be optimized anymore. Thus, the only variable we could optimize is $\hat{\mathbf{s}}^{\text{HR}}$ and \mathbf{c}^{LR} . The problem is, a critical variable \mathbf{x}^{HR} for optimizing is unknown. And since

$$\|\boldsymbol{x}^{\text{HR}} - D_1(\boldsymbol{c}^{\text{LR}}, \hat{\boldsymbol{s}}^{\text{HR}})\|_2^2 \approx \|Rs(\boldsymbol{x}^{\text{HR}}) - Rs(D_1(\boldsymbol{c}^{\text{LR}}, \hat{\boldsymbol{s}}^{\text{HR}}))\|_2^2 \approx \|\boldsymbol{x}^{\text{LR}} - Rs(D_1(\boldsymbol{c}^{\text{LR}}, \hat{\boldsymbol{s}}^{\text{HR}}))\|_2^2,$$
(5)

where *Rs* is a resample function which resamples $\mathbf{x}^{SR} = D_1(\mathbf{c}^{LR}, \mathbf{\hat{s}}^{HR})$ to be the same size as \mathbf{x}^{LR} . And since \mathbf{x}^{LR} is a known variable, we optimize \mathbf{c}^{LR} and $\mathbf{\hat{s}}^{HR}$ using stochastic gradient descent (SGD) by minimizing the following function,

$$L_{\rm SR} = \| \mathbf{x}^{\rm LR} - Rs(D_1(\mathbf{c}^{\rm LR}, \hat{\mathbf{s}}^{\rm HR})) \|_2^2.$$
(6)

And after optimization, we obtain optimized $\mathbf{c}^{\prime LR}$ and $\mathbf{\hat{s}}^{\prime HR}$. The final SR output is $\mathbf{x}^{SR} = D_1(\mathbf{c}^{\prime LR}, \mathbf{\hat{s}}^{\prime HR})$.

3. Experiments and results

1) Datasets and parameter settings for experiment

We prepared five resected lung specimens resected from lung cancer patients. We took clinical CT volumes of these lung specimens using a clinical CT scanner (SOMATOM Denition Flash, Siemens Inc., Munich) in the resolution of $625 \times 625 \times 600 \ \mu\text{m}^3$ / voxel. We took micro CT volumes of these specimens using a micro CT scanner (inspeXio SMX90CT Plus, Shimadzu, Kyoto) in the resolution of $(42-52) \times (42-52) \times (42-52) \ \mu\text{m}^3$ / voxel. We chose 6-10 feature points manually and used affine transformation [7] to obtain coarsely-aligned clinical (LR) - micro CT (HR) volumes. We utilized deedsBCV [1] for precise registration of coarsely-aligned clinical CT – micro CT volumes. For training, we

utilized four aligned clinical - micro CT volumes. For inference, we utilized one aligned clinical - micro CT volume.

For training the FCN for SR, we randomly cropped 2000 2D patches from each case. We experimented 4×SR (size of SR image is 4-times of LR image by height and width). The size of patches was of 48×48 pixels extracted from the clinical CT volumes and 192×192 pixels extracted from the micro CT volumes. We set the learning rate of 10⁴ and halves every 30 epochs. The epoch number is 200 and optimize algorithm is Adam [8]. The size of Scode is 8 (\hat{s}^{HR} is sampled from 8-dimensional normal distribution). Weight of each loss term: $\lambda_1 = 0.2$, $\lambda_2 = 1.0$, $\lambda_3 = 1.0$, $\lambda_4 = 1.0$.

2) Results

Figure 3 and 4 illustrate the result. We compared our method with a supervised SR method named EDSR [9]. For quantitative evaluation, we evaluated our method using 100 precisely-aligned clinical-micro CT image pairs. Proposed method obtained PSNR score of 23.26 (db) and SSIM of 0.28, while EDSR obtain PSNR score of 22.80 (db) and SSIM score of 0.02. For quantitative evaluation, we evaluated our method using 500 clinical CT images from a lung cancer patient. Our method outputted satisfying SR result while EDSR outputted very blurry and nearly arbitrary results as in Fig. 3. Our method also reconstructed micro anatomical structures such as small bronchus and edge of tumor better than EDSR.

4. Discussion and conclusion

Due to the reason that obtaining resected lung specimens is difficult, we could only scan five resected lung specimens. Results of proposed method also has some artifacts and noises as illustrated in Fig. 3. We consider this is because the HR (micro CT) images have severe corresponding noises and artifact. Proposed method optimizes the SR output to as close as its HR (micro CT) counterpart. Thus SR images are likely to appear some noise and artifacts like micro CT images.

In this paper, we constructed an aligned LR (clinical CT) and HR (micro CT) dataset for SR of clinical CT images. Further, we have proposed a novel SR approach for performing the SR of clinical CT images, while training on poorlyaligned LR-HR dataset. The proposed method outperfromed supervised SR method qualitatively and quantitatively. The results demonstrate that our proposed method successfully performed SR of lung clinical CT four-times by height and width. Future work contains building a larger LR-HR dataset and denoising of HR (micro CT) images. In the future, we consider constructing a larger dataset by collecting more lung specimens, and we also plan to reduce noises and artifacts of μ CT images by denoising algorithms.

Competing interests

None

Acknowledgement

Parts of this research was supported by MEXT/JSPS KAKENHI (26108006, 17H00867, 17K20099), the JSPS

Bilateral International Collaboration Grants, the AMED (JP19lk1010036 and JP20lk1010036) and the Hori Sciences & Arts Foundation.

References

- Mattias P, Heinrich M, Jenkinson M, et al.: NTIRE 2017 MRF-Based deformable registration and ventilation estimation of lung CT. IEEE Transactions on Medical Imaging 32: 1239-1248, 2013
- [2] Goldman, W: Principles of CT: radiation dose and image quality. Journal of Nuclear Medicine Technology 35: 213-225, 2007
- [3] Dong C, Loy C, He K, et al.: Image super-resolution using deep convolutional networks. IEEE transactions on pattern analysis and machine intelligence **38**: 295-307, 2015
- [4] Sotiras, A., Davatzikos, C., Paragios, N.: NTIRE 2017 Challenge on Single Image Super-Resolution: Dataset and Study. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops: 1122-1131, 2017
- [5] He, K., Zhang, X., Ren, S.: Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition: 770-778, 2016
- [6] Huang, X, Serge B.: Arbitrary style transfer in real-time with adaptive instance normalization. Proceedings of the IEEE International Conference on Computer Vision: 1501-1510, 2017.
- [7] Mattias P. Heinrich, M. Jenkinson, M, et al.: Multimodal unsupervised image-to-image translation. Proceedings of the European conference on computer vision: 172-189, 2018
- [8] Kingma, Diederik P, and Jimmy Ba.: ADAM: A method for stochastic optimization, proceedings of ICLR, 2015.
- [9] Lim, Bee, et al.: Enhanced deep residual networks for single image super-resolution. Proceedings of the IEEE conference on computer vision and pattern recognition workshops: 1132-1140, 2017.

敵対生成ネットワークによる

COVID-19のCT 画像の生成モデル

川崎 良*1 斉藤 篤*1 橋本 正弘*2 大竹 義人*3

明石 敏昭*4 清水 昭伸*1

要旨

本報告では、COVID-19の典型症例の胸部 CT 画像を対象に、敵対的生成ネットワーク(Generative Adversarial Networks; GAN)を用いた生成モデルの構築について述べる. COVID-19の典型症例の CT 画像は、病変が 末梢優位性と両側性を持つと報告されている.本研究では、この特徴を示す胸部 CT 画像の生成が可能な モデルを2種類の GAN を使って開発した.まず、胸部 CT 画像は Axial 面のスライス画像のサイズが 512 ×512pixel と大きく、そのままでは学習が困難である.そこで高解像度の画像を安定して正確に生成可能な Progressive growing を取り入れた PGGAN を用いた.また、2 つのネットワークで構成された Generator を組 合せることで高品質な画像生成を実現した StyleGAN を用いた.それぞれのモデルによって生成された画 像に対して、Fréchet Inception Distance (FID)による定量評価や、血管、気管支、病変の特徴に関する視覚的 な評価指標による定性評価を行い、二つのモデルの優劣について議論する. キーワード: COVID-19、CT 画像、画像生成、Generative Adversarial Network

1. はじめに

2019 年末から 2021 年現在に至るまで, COVID-19 (新型コロナウイルス感染症)の世界 的感染拡大が続いている. COVID-19 に対する 診断法として現在, RT-PCR 法が主流である.し かし, RT-PCR 法は COVID-19 の診断には不十 分であるという報告がある [1]. そこで, RT-PCR 法を用いたウイルス検査だけでなく, CT 画 像を活用した画像診断への関心が高まってい る. 深層学習を用いた COVID-19 の診断支援の 研究 [2]はすでに進められているが, COVID-19 のラベル付学習データを大量に収集すること は容易ではない.

*1	東京農工大学大学院工学研究院
----	----------------

- 〔〒184-8588 東京都小金井市中町 2-24-16〕
- e-mail: s211085t@st.go.tuat.ac.jp
- *2 慶應義塾大学医学部放射線科
- *3 奈良先端科学技術大学院大学
- *4 順天堂大学放射線診断学講座

本研究では、診断支援システムの学習データ の増強を行うための COVID-19 を含む肺炎画像 の生成モデルを、GAN (Generative Adversarial Network) [3]を用いて構築する.

2. 手法

本研究では、生成モデルを構築するためのネ ットワークとして PGGAN [4]および StyleGAN [5]を使用する. PGGAN は徐々にネットワーク の層を追加して段階的に学習する解像度を上 げる Progressive Growing という手法を用いた GANである.これにより,本研究で使用する512 ×512pixel の CT 画像のような解像度の高い画 像を安定して正確に学習することができる.ま た, StyleGAN は従来まで1つの CNN で構成さ れていた Generator を Mapping network と Synthesis network の2 つのネットワークに改良 し、スタイル変換の分野で提案された AdaIN [6] という正規化手法を取り入れることで、さらに 表現能力を向上させた GAN である.本研究で は、それぞれの GAN により独立に生成モデル を構築し、両者を比較する.

3. 実験条件および実験結果

実験試料には、日本医学放射線学会が収集し た 80 症例の CT 画像 (東京農工大学倫理委員会 承認番号 No.200705-0226)を用いた.この 80 症 例は北米放射線学会 (Radiological Society of North America; RSNA)の報告 [7]に基づいて、 放射線医が COVID-19 典型例と診断したもので ある.実験では、CT 画像から取り出した病変の 写った Axial 面スライス画像 (512×512pixel) 1510 枚を使用した.前処理として肺野条件に基 づく[-1,1]の正規化を実施した.Optimizer は Adam [8] ($\beta_1 = 0.0$, $\beta_2 = 0.99$)を用いた.生成 画像は Fréchet Inception Distance (FID) [9]に よる定量評価に加えて、ランダムに選んだ 20 枚 の生成画像に対して以下の定性評価を行った. (I)肺の輪郭形状が自然である

(Ⅱ)血管や気管支の解剖学的構造が自然である (Ⅲ)その他解剖学的に不自然なノイズがない

(IV)COVID-19 典型例の特徴を表現できている (両側性・末梢優位, 病変部位が不自然でない)

実験結果と評価値をそれぞれ図 1,表1に示 す.(↓)は値が低いほど,(↑)は値が高いほ ど良い評価であることをあらわす.表1より, 定量評価,定性評価ともに StyleGAN の生成画 像の方が優れていた.また,図1のように COVID-19 典型例の特徴である両側性・末梢優 位の病変を表現することができたが,実画像に 見られないような不自然なノイズや解剖構造 が散見された.



(a) PGGAN



(c) 実画像

実験結果 評価値

Mathad	FID	定性評価(↑)			
Method	(\downarrow)	(I)	(11)	(Ⅲ)	(IV)
PGGAN	0.138	0.3	0.275	0.25	0.375
StyleGAN	0.107	0.575	0.5	0.25	0.4

(b) StyleGAN

図1

事 1

4. まとめ

本稿では、PGGAN および StyleGAN を用いた COVID-19 典型例の CT 画像生成について述べ た. 今後は、不自然なノイズや解剖構造が見ら れる点が課題であり、これらを抑制するための 制約を検討する. また、典型例以外の症例につ いても生成モデルを構築して、その生成画像を 用いた診断支援モデルの学習を行う予定であ る.

謝辞

研究遂行にあたり貴重なご助言を賜った清水昭伸 研究室各位に深謝いたします.本研究では日本医学 放射線学会のナショナルデータベース J-MID(臨床研 究等 ICT 基盤構築・人工知能実装研究事業/画像診 断ナショナルデータベース実現のための開発研究 (課題管理番号: JP20lk1010025))を利用しました. データベース構築に関わっていただいた先生方に感 謝いたします.また,本研究は,AMDの医療ビッグ データ利活用を促進するクラウド基盤・AI 画像解析 に関する研究(課題管理番号: JP20lk1010036)の支 援を受けています.

利益相反の有無

無

文 献

- Y. Fang, H. Zhang, J. Xie, et al. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. Radiology, 2020
- [2] 村尾晃平:特集/医療画像ビッグデータクラウド 基盤を活用した医療系学会・大学間連携による COVID-19 肺炎画像診断支援 AI 開発. Medical Image Technology Vol.39, 2021
- [3] I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al. Generative Adversarial Nets. In *Proc. NeurIPS*, 2014
- [4] T. Karras, T. Aila, S. Laine, et al. Progressive Growing of GANs for Improved Quality, Stability, and Variation. In Proc. ICLR, 2018
- [5] T. Karras, S. Laine, T. Aila. A Style-Based Generator Architecture for Generative Adversarial Networks. In *Proc. CVPR*, 2018
- [6] X. Huang and S. Belongie. Arbitrary style transfer in real-time with adaptive instance normalization. *CoRR*, 2017
- [7] S. Simpson, F. Kay, S. Abbara, et al. Radiological Society of North America Expert Consensus Statement on Reporting Chest CT Findings Related to COVID-19. Radiology, 2020
- [8] D. P. Kingma, J. Ba. Adam: A Method for Stochastic Optimization. In Proc. ICLR, 2015
- [9] M. Heusel, H. Ramsauer, T. Unterthiner, et al. GANs trained by a two time-scale update rule converge to a local Nash equilibrium. In *Proc. NeurIPS*, 2017.

Generative Models of COVID-19 CT Images

Using Generative Adversarial Networks

Ryo KAWASAKI^{*1}, Atsushi SAITO^{*1}, Masahiro HASHIMOTO^{*2}, Yoshihito OTAKE^{*3}, Toshiaki AKASHI^{*4}, Akinobu SHIMIZU^{*1}

- *1 Institute of Engineering, Tokyo University of Agriculture and Technology
- *2 Department of Radiology, Keio University School of Medicine
- *3 Nara Institute of Science and Technology
- *4 Department of Radiology, Juntendo University

In this paper, we propose generative models using Generative Adversarial Networks (GANs) for chest CT images with typical appearances of COVID-19. Typical radiological findings of COVID-19 pneumonia are bilateral and peripheral lesion. We trained two generative models to generate CT images with these specific features using two different GANs. The size of CT images is large, or 512 x 512, which makes training of a generative model difficult. We employed PGGAN that has progressive growing to generate high resolution images stably and precisely. Another GAN is StyleGAN that combines two networks to generate high quality images. Synthesized images by the two different models were quantitatively evaluated using Fréchet Inception Distance (FID) and qualitatively evaluated by visual assessment of blood vessels, bronchi, and lesions, followed by the discussion about the comparison of the two models.

Key words: COVID-19, computed tomography, image synthesis, generative adversarial network

Non-contrast to Artery Contrast CT Translation Via Representation-Aligned Generative Model

Tao HU^{*1}, Masahiro ODA^{*1,2}, Yuichiro HAYASHI^{*1}, Zhongyang LU^{*1}, Toshiaki AKASHI^{*3}, Shigeki AOKI^{*3}, Kensaku MORI^{*1,4}

Abstract

This paper proposes a deep representation-based network to improve the CT modality translation from non-contrast CT to artery contrast CT (ACT). ACT has been widely applied to the diagnosis of abdominal vascular diseases. For the imaging of ACT, the contrast agent is administrated to the patients' body before the scan. To avoid the potential risk led by contrast agent, a deep learning technique has been utilized to simulate ACT from non-contrast CT. In this paper, we propose to improve the translation performance via a representation-aligned network. Unlike conventional generative adversarial networks (GANs), the proposed generator is trained to fool the discriminator simultaneously and approximate the real samples in the embedding space extracted by a representation learning network. The results show that the proposed network is able to generate ACT with more semantic information and higher accuracy.

Keywords : generative adversarial network, modality translation, contrast CT analysis, representation learning

1. Introduction

Artery contrast computed tomography (ACT) is widely used to diagnose vascular diseases in the abdomen. Compared with non-contrast CT (NCT), ACT produces enhancements in vascular structures and provides a non-invasive way to observe blood vessels and lesions. ACT enables precise three-dimensional reconstruction of the aorta in the clinical field, thus being widely accepted as key criteria for diagnosing and managing aorta aneurysms and thrombus [1].

However, for the imaging of ACT, the intake of the contrast agents is necessary. Contrast agent administration may bring damages to the patients with potential contrast agent allergy. A previous study finds that the rate of side effects of ionic contrast materials is 4% to 12%, and 1% to 3% for non-ionic contrast materials [2]. The severe symptoms of allergy may even result in death. To avoid the potential risks of contrast agent allergy and protect patients from excessive radiation, researchers propose using deep learning techniques to estimate ACT from NCT [3].

The investigated NCT to ACT modality transfer can be referred as an image-to-image translation task in the related works [4,5,6]. In these models, the generative networks are optimized to learn the intensity distribution in the target

*1 Graduate School of Informatics, Nagoya University, Aichi, Japan,

[Furo-cho, Chikusa-ku, Nagoya, 464-8601, Japan]

e-mail: thu@mori.m.is.nagoya-u.ac.jp

^{*2} Information Strategy Office, Information and Communications, Nagoya University

^{*3} Department of Radiology, Juntendo University, Japan

^{*4} Research Center for Medical Bigdata, National Institute of Informatics, Japan

domain by minimizing an overall cost function. Specifically, for the generative networks of these models, the cost function is either designed to quantify the difference between the images in source and target domains in a pixel to pixel manner like L1 or L2 norm [4]. Alternatively, the difference between the real and the synthesized images is measured by the discriminator [5,6]. This mechanism means that the networks are primarily optimized to learn the whole translation. No sufficient priority on the aorta pixels is introduced. This process is very different from the visual inspection by human experts. A trained physician pays attention to a specific region of interest (ROI) such as lesion and aorta. In the training of the conventional networks, the identification of ROI is nevertheless expected to be learned implicitly. This may lead to decrease accuracy in the ROI and a lack of practical information in the generated images.

This paper proposes a novel network combining the NCT to ACT translation and representation learning to address the problems mentioned above. First, we train a fully convolutional network (FCN) with a narrow bottleneck block to learn the reconstruction of ACT and the aorta segmentation. The output of the bottleneck block can be regarded as the deep representation of input ACT. Then we define the L2 norm of the estimation error calculated from the deep representations of estimated and actual ACT as a *representation-align* loss. The representation-align loss is integrated to the final loss of the generative model during training.

2. Methods

As shown in Fig. 1, our proposed network comprises two modules. The first is the image translation module, whose input is NCT and output is ACT. We implemented the translation module using the unsupervised method CycleGAN [6]. CycleGAN is proposed to solve the unpaired image translation problem. There are two generators in CycleGAN. The first generator G_1 is designed to map a sample in the source domain (i.e. NCT in our study) to a sample in the target domain (i.e. ACT in our study). Also, a discriminator is introduced to distinguish the synthesized ACT from the real ACT. The well-behaved adversarial training of this generator and discriminator will result in very similar distributions over the synthesized ACT and real ones. Thus the well-trained generator can perform translation from desired NCT to ACT.

To guarantee the structural consistency between synthesized ACT and input NCT, an additional generator G_2 is introduced to reconstruct the input NCT from synthesized ACT. This backward-forward translation will bring a consistency loss. By minimizing this consistency loss, the consistent shapes of organs in NCT and synthesized ACT can be better maintained.

The second module is the deep representation learning module to extract salient features of estimated and true ACT. The deep representation module is implemented as FCN with a narrow bottleneck block. In the training phase, the input of this FCN is true ACT, and the output is the reconstructed ACT and aorta segmentation mask. This FCN architecture consists of an encoder and a decoder. The output of this encoder is the feature map of the narrow bottleneck block.



Fig. 1 Proposed network architecture. (a): Proposed FCN for representation learning. (b): Proposed network for NCT to ACT translation.

Through this encoder, the input ACT of size 512×512 pixels is compressed to a tensor of size $32 \times 32 \times 4$ (length×width×channel) elements. Then this tensor is fed to the decoder to reconstruct the original input ACT and segment aorta. To ensure the efficiency of the compressed tensor, we did not implement the encoder-to-decoder feature concatenation or summation, which is commonly adopted in U-Net. In the training phase, to accurately reconstruct the original ACT and identify the aorta, the encoder is forced to learn the abstract and high-dimensional representation, which keeps the most prominent information and aorta location of input ACT, and the decoder is optimized to correctly recover the input ACT and identify aorta from the compressed tensor. After the training step, the encoder is used as a deep representation extractor, and the output feature map is regarded as the deep representation of input ACT.

To maintain the high-dimensional semantic information and encourage attention on aorta regions, we integrate the deep representation loss into the final estimation loss of the generative network. The deep representation loss shown in Eq.(1) is defined as the L2 norm of the representation estimation error.

$$L_z = \|\boldsymbol{r}_e - \boldsymbol{r}_t\|_2^2, \tag{1}$$

where r_t is the representation of true ACT, r_e is the representation of estimated ACT by G_1 .

Based on the deep representation loss, we define the overall loss of generative network as the linear combination of the original CycleGAN loss and the deep representation loss as Eq.(2).

$$L_3 = \lambda_1 L_c + \lambda_2 L_z, \tag{2}$$

where L_c is the loss of CycleGAN, L_z is the representation loss, λ_1 and λ_2 are weights to balance the losses.

Method	MSE	PSNR
CycleGAN [6]	2703.60	15.70
Proposed	2692.02	15.82

 Table 1 Mean MSE and PSNR for the proposed model and baseline model (CycleGAN). Pixels in aorta regions were included for the calculation of the two metrics.

3. Experiments

We collected 13 cases (7 normal cases, 6 patients with AAA) with both non-contrast and artery contrast CT volumes for the present study. The total number of CT slices is 1,316, and each of them has a size of 512 by 512 pixels. The pixel spacing is 0.665mm by 0.665mm, and the slice spacing is 5.0mm. We randomly selected 11cases for model training, and the remained two cases are used for testing.

B-spline registration [7] between NCT and ACT volumes were applied to eliminate the influence of misalignment and distortions. Poorly registered slices were excluded by manually checking to ensure only paired images were included. Intensity normalization was then used to linearly map the original CT values to the range of 0 and 1. In addition, we used random affine transformation for data augmentation based on the assumption that the proposed network should be sensitive to the scale variation.

In the present study, the training data are processed in a slice-wise way. Instead of using 3D convolution layers, we utilize 2D filters with less trainable parameters to implement the models. In order to involve 3D information in the models, we use three adjacent axial slices as input, and the output is the estimation result of the central slice. Batch size is set to 2, and the Adam optimizer is used to train all the models. For the training of the proposed network, the weights λ_1 , λ_2 were set to 1.5, 1 respectively.



Fig. 2 Estimation results by CycleGAN and proposed network. Red boxes show the region of aorta. (a): Input NCT.(b): Estimated ACT by CycleGAN. (c): Estimated ACT by proposed network. (d): True ACT.

4. Results

To evaluate the quantitative performances of the models, we computed two widely used metrics: mean square error (MSE) and peak signal-to-ratio (PSNR) on test cases. Since there are still location misalignments and organ shifts in many CT slices, directly calculating the whole slice's metrics will disturb evaluation accuracy. We computed these slicewise metrics only on aorta regions. Only pixels in aorta regions are included for calculating the metrics. The mean MSE and PSNR of estimated ACT by the baseline and proposed models are shown in Table 1. As shown in Table 1, the proposed network could produce better estimation performance with lower mean MSE and higher mean PSNR in aorta regions compared to baseline CycleGAN without representation loss.

Figures 2 and 3 show the quantitative evaluations of the baseline and proposed model. Figure 2 shows the estimation results of an abdomen slice by CycleGAN and the proposed method. Red boxes indicate aorta regions. We can see that the estimated intensity by the proposed method is higher than CycleGAN in the aorta region.

Figure 3 presents the volume rendering results by baseline and proposed methods. We set the opacity curve setting for four volumes the same. The presented case is a patient with AAA. The second row of each subfigure shows the zoomed-in region of AAA. Yellow arrows point to the regions where different estimation results can be observed. From Fig. 3, we can see that the aorta and AAA regions were visible in the results by baseline and proposed networks. Additionally, the artifacts presented near AAA are severe in the result by CycleGAN and could be reduced by the proposed network.



Fig. 3 Volume rendering results by CycleGAN and proposed network. The second row in each subfigure shows the zoom-in AAA region. Yellow arrows point to the location of AAA. Circles show the regions with artifacts in estimated ACT. (a): Input NCT. (b): estimated ACT by CycleGAN. (c): Estimated ACT by the proposed network. (d): True ACT.

5. Discussion

This paper proposes a novel network for NCT to ACT translation. In comparison to the baseline method, our network achieved better quantitative and qualitative evaluations. By introducing representation loss in the training of the generative model, our proposed network can pay more attention to the accurate semantic information of the estimated ACT. The false enhancements with no real information can be thus reduced. Also, the aorta regions with significant semantic features can be better estimated.

Besides, the proposed FCN with auxiliary aorta segmentation enables convenient measurement of an aorta from NCT. Since the maximum axial radius is reported to be able to serve as a biomarker of AAA [1], this automated measurement of aorta provides primary evidence of AAA diagnosis. It may be helpful for the future development of an explainable computer-aided diagnosis system for AAA.

Competing interests

None.

Acknowledgement

Thank for the help and advices from the members of Mori laboratory and AMED NII team. A part of this research was supported by the AMED Grant Numbers JP19lk1010036, JP20lk1010036, and JSPS KAKENHI 26108006, 17K20099, 17H00867.

References

- [1] Metcalfe D, Holt P, Thompson M: The management of abdominal aortic aneurysms. BMj 342, 2011
- [2] Jean M, Emmanuelle P, Philippe P, et al: Allergy-like reactions to iodinated contrast agents: A critical analysis. Fundamental & clinical pharmacology 19: 263-281, 2005
- [3] Oda M, Kumamaru K, Aoki S, et al.: Non-contrast to contrasted abdominal CT volume regression using fully convolutional network. Int Journal of Computer Assisted Radiology and Surgery 14: s103–s104, 2019
- [4] Ronneberger O, Fischer P, Brox T: U-Net: Convolutional networks for biomedical image segmentation. Int. Conference on Medical image computing and computer-assisted intervention, LNCS 9351: 234–241, 2015
- [5] Isola P, Zhu J, Zhou T, et al.: Image-to-image translation with conditional adversarial networks. In: Proc. IEEE Conference on Computer Vision and Pattern Recognition, 1125–1134, 2017
- [6] Zhu J, Park T, Isol P, et al.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE international conference on computer vision, 2223-2232, 2017
- [7] Balci S, Golland P, Wells W: Non-rigid groupwise registration using B-spline deformation model. Int. Conference on Medical image computing and computer-assisted intervention, 105–121, 2007

多重解像度画像のグループ型学習を利用した

新たな MR 深層学習再構成の検討

大内 翔平*1*2, 伊藤 聡志*1

要旨

近年, CNN を用いた MR 画像の再構成法が数多く提案されているが,信号の間引きに由来する高周波成分の復元と 折り返しアーチファクトの除去には依然として課題が残っている.また,位相情報を持つ複素 MR 画像に対応した CNN 再構成法の提案は少ない.本研究では,多重解像度解析手法である eFREBAS 変換と画像空間型の CNN 再構成法を併 用した新たな深層学習再構成法として eFREBAS-CNN を提案する.提案法では, eFREBAS 展開により,画像を解像度 別に分解した複数枚のサブ画像を生成する.これらのサブ画像を被写体構造の保存度合に応じてグループ分割し,グル ープ毎に別個の CNN で学習と再構成を行う. eFREBAS 変換は複素変換であるため,提案法は位相を含む MR 画像に 適用可能な特長を有する.再構成シミュレーションの結果,既存の位相考慮型の再構成法と比べて位相画像の PSNR を 改善でき,さらに画像の詳細構造の復元性能を高めることができた.

キーワード: 深層学習, 圧縮センシング, 多重解像度解析, eFREBAS 変換

1. はじめに

医用画像診断装置 MRI(Magnetic Resonance Imaging)の撮像時間の短縮を目的として,圧縮セ ンシング(Compressed Sensing: CS)[1][2]を応用し た画像再構成法(CS-MRI)が提案されている[3]. しかし, CS-MRI による再構成では数理的な反復 処理を伴うため,従来のフーリエ変換を基本と した再構成処理と比較して処理が長時間に及ぶ こと,また,画像を変換した空間でスパース性を 仮定することによる人工的な様相が像上に生じ る場合があることなどの課題がある.これらの 課題の解決を目的に,近年では CS 再構成処理に CNN(Convolutional Neural Network)を応用した CNN-CS が注目されている[4][5]. CNN-CS は従

*1 宇都宮大学大学院 工学研究科 システム創成工学専攻 [〒321-8585 栃木県宇都宮市陽東 7-1-2] e-mail: ultuchiboy@gmail.com *2 日本学術振興会特別研究員 DC 来の CS-MRI と比較して, 再構成像の品質と再構 成時間の点で優れた性能を示す一方で, CS にお ける MR 信号の間引き収集に起因する, 折り返 しアーチファクトの除去性能と失われた高周波 成分の復元性能が依然として課題である.

本研究では、画像空間上で再構成を行う CNN-CS である Image Domain Learning を基本とした CNN と,多重解像度解析手法である eFREBAS 変 換(enhanced Fresnel band-split transform)[6]を融合 した新たな CNN-CS である eFREBAS-CNN を提 案する. eFREBAS 変換によれば、入力画像を解 像度別に任意数のサブ画像に展開可能であり、 展開されたサブ画像は、それぞれが入力画像中 の異なる周波数成分をもとに構成される. この 特徴を利用して、eFREBAS-CNN では eFREBAS 変換を経て得られたサブ画像単位で CNN による 再構成処理を行う方法により、周波数帯域ごと にアーチファクト除去と間引きにより消失した

被写体構造の復元を図る. サブ画像は間引きに よる構造の消失度合いに応じて複数のグループ に分割し、グループごとにマルチチャンネルを 持つ複数の CNN によって再構成する.また, eFREBAS変換は複素変換であるため, eFREBAS-CNN は位相情報を持つ MR 画像の再構成に対応 する. 画像の周波数帯域に着目した CNN-CS と して,画像を高周波成分と低周波成分に分離し て再構成を行う CNN-CS[7]や、CNN 中にウェー ブレット変換を導入した CNN-CS[8]等が提案さ れているが、マルチチャンネル CNN かつグルー プ型学習を利用した,位相画像対応型の CNN-CS は我々の知る限りこれまでに提案されていない. 本研究では、提案する eFREBAS-CNN により位 相画像の再構成を行い、既存の位相対応型の再 構成法との比較を通して, eFREBAS-CNN の特徴 について検討した.

2. eFREBAS-CNN

2.1 eFREBAS 変換

eFREBAS 変換は、2 つの異なるアルゴリズム を持つフレネル変換によって構成される多重解 像度解析手法である.入力画像を任意数かつ任 意のサイズの画像群(サブ画像)に展開でき、展 開されたサブ画像の位置とk-spaceの領域には対 応関係があるため、各サブ画像によりフィルタ バンクが構成される.ここで、簡単のため、一次 元信号について考えると、eFREBAS 変換による 展開像 $\rho_m(x)$ は、式(1),(2)のように、バンドパスフ ィルタによって畳み込まれた形で表される.

 $\rho_m(x) = \left[\rho(x)e^{-jcx^2}e^{-j2cmXx}\right] * g(x) \quad (1)$

 $g(x) = \operatorname{sinc}(2cXx) \tag{2}$

ここで,式中の $\rho(x)$ は入力画像, $X(= N\Delta x)$ は入力画像の FOV, Nは入力画像のサイズ, Δx は ピクセル幅, cは eFREBAS 変換のスケーリン グ係数である $D = \pi/(cN\Delta x^2)$ に影響する定数を



表している. また, mはフィルタバンクのインデ ックスを表しており, サブ画像群はそれぞれ入 力画像の特定の周波数成分から構成されている. 図 1 にスケール係数 D を 1,3,5 としたときの eFREBAS 変換像および, (f)の信号量 30%のパタ ーンでアンダーサンプリングした際のゼロフィ ル再構成像(D=3)を示す. なお, D=1 では画像を 展開しないため, eFREBAS 変換を適用しないオ リジナルの画像に等しい.

図 1. eFREBAS 変換像の例

2.1 eFREBAS-CNN による再構成処理

eFREBAS-CNN では、MR 信号のアンダーサン プリングを経て得られた、画像上に折り返しア ーチファクトが発生しているゼロフィル再構成 像と、全信号から得たフルデータ像の eFREBAS 空間上での関係を Image Domain Learning を基本 とする CNN によって学習する. eFREBAS 展開 像の各サブ画像は異なる解像度を有するため、 図 1 のように、ゼロフィル再構成像のサブ画像 では、被写体構造の消失やアーチファクトの重 畳にはサブ画像間で特徴の差が存在する. これ らの異なる特徴を持つサブ画像を効率的に学習 するため、eFREBAS-CNN では展開像をサブ画像 ごとに分割し、図 2 のように、フルデータ像との 近似の程度に応じてサブ画像を Group A, Group B, Group C の 3 グループに分割した. このうち、



図2. サブ画像のグループ分割の概要 Group A は低周波成分であり、アンダーサンプリ ングを受けても情報損失が少ないサブ画像群で ある. Group C は画像の高周波成分であり、アン ダーサンプリングにより多くの情報を消失する サブ画像群である. Group Bは2つのグループの 中間に位置する特徴を持つサブ画像群である. これらの3グループは、対応する3つのCNNに よって独立に学習および再構成が行われる. 各 CNN による再構成後は、サブ画像を組み合わせ て1枚の展開像を生成し,逆 eFREBAS 変換の適 用によって画像空間で再構成像を得る. 最後に, 再構成像の k-space に対して Data Consistency を 行うことで、k-space の誤差を減少し再構成像の 高品質化を図る. eFREBAS-CNN による再構成処 理の手順を図3に示す.なお、本研究では eFREBAS 変換のスケール係数を D=3 に設定し たため、生成されるサブ画像は計9枚となる.

2.3 eFREBAS-CNN のアーキテクチャ

eFREBAS-CNN は、図3に示すように3つの U-NetベースのCNN(U-Net1, U-Net2, U-Net3)から 構成されている.これらのU-Net は文献[9]のLee らのU-Netを基本としている.eFREBAS変換は 複素変換のため、出力される画像は複素数を持 つが、本研究で使用するCNN は実関数型である ため、画像の実部と虚部を分離してそれぞれを CNN に与える.そのため各グループ用の3つの CNN は実部像と虚部像の再構成を行うために2 回用いられる.

eFREBAS-CNN では再構成の品質向上のため、 次の3つの手法を導入した.最初に、すべての CNN の入出力をマルチチャンネルとした. 各サ ブ画像は周波数成分が異なるものの、被写体構 造に関しては類似しているため、マルチチャン ネル化によりそれぞれのサブ画像が持つコンテ キスト情報を相互に有効活用した再構成が行え るようにした. 2つめに, Group Bの再構成を行 う U-Net2 については、Group A の画像も共に入 力として与え, Group C の再構成を行う U-Net3 については、Group A と Group B の画像も共に入 力として与えた. 最後に, U-Net2 と U-Net3 の入 力を1ch分増やし、サブ画像と同一のサイズに縮 小した画像空間のゼロフィル再構成像を入力と して与えた. Group B と Group C は, 被写体構造 の消失の程度が比較的大きいため、2つめと3つ めの手法によって CNN に与える情報量を増やす ことで、構造の復元性能の向上を図った.以上の 手法の導入により, 各 CNN の入力チャンネル数 は, U-Net1 が 3ch, U-Net2 が 6ch, U-Net3 が 10ch となる.

3. 画像再構成シミュレーション

3.1 学習およびテスト用画像

CNN の学習には、キヤノンメディカルシステム ズ社製の Vantage Titan 3T で撮像された T1 強調 像を計 25 枚, T2 強調像 5 枚を使用した. このう ち, T1 強調像 25 枚を学習用とし,残りの 10 枚 をテスト用とした. これらの画像のマトリクス サイズは256 × 256である. また, eFREBAS-CNN で用いる U-Net ベースの CNN は、中間層での特 徴マップのスケーリング処理にプーリングと逆 プーリングを使用している. そのため、全ての画 像の k-space に対して外周部にゼロデータを充填 することで、マトリクスサイズを264 × 264とし、 eFREBAS 変換後のサブ画像サイズを88 × 88と





することで,スケーリング処理後のサイズが偶 数となるようにした.この処理により,再構成像 のマトリクスサイズは264×264となるため, Data Consistencyの適用前にk-space上で充填部の 信号をカットし,256×256への縮小を行った.

3.2 学習条件と計算機環境

eFREBAS-CNN の学習時の適用条件と,計算機 環境をそれぞれ表 1,表 2 に示す.本研究では全 てのグループ用の CNN に対して同一の条件を適 用した. CNN の学習には MATLAB および MatConvNet[10]を使用した.

衣 I. CNN の子首条件			
Optimizer	Adam		
Learning Rate	$1 \times 10^{-3} \sim 1 \times 10^{-5}$		
Loss function	画像空間の MSE		
Epoch	300		

ま1 ODIの 単羽々 他

3	表 2. 計算機環境
CPU	Intel Core i7-9700K
RAM	DDR4-2666 64GB
GPU	NVIDIA GeForce RTX 2080ti
	MATLAB R2017b
C - Arren arr	CUDA toolkit 10.2
Sonware	cuDnn 7.6.5
	MatConvNet 1.0-beta25

3.3 比較手法

提案する eFREBAS-CNN との比較のため,画 像空間上で学習を行う Image Domain CNN, Zhao らが提案した位相歪みを考慮する CS-MRI[11]お よび,Yang らが提案した ADMM-CSNet[12]によ る再構成を行った.Image Domain CNN では,単 一の実関数型 U-Net により,実部像と虚部像を 分割した画像の画像空間上での関係性を表 1 と 同一の条件で学習した.また,再構成像に対して Data Consistency を適用した. ADMM-CSNet は, CS-MRI の一種である ADMM[13]の処理を CNN で実装した手法である.本研究では,ADMM-CSNet で 25 枚の複素画像を用いて 300 イテレー ションの学習を行った.再構成像の品質評価に は PSNR(Peak Signal-to-Noise Ratio)を使用した. なお,PSNR は複素画像を用いて算出した.

3.4 サンプリングパターン

本研究では,提案法及び比較手法による再構 成に,信号収集量が 20,30,40%のサンプリング パターンを使用した.

3.5 再構成シミュレーション結果

各種の再構成法によってテスト画像の再構成 を行った.図1に示した収集信号量 30%のサンプ リングパターンを適用した際の再構成像を図 4 に,図4の注目箇所の拡大画像を図5に示す.ま



図 4. 信号収集量 30%のときの各手法による再構成像: (a)フルデータ像, (b)ゼロフィル再構成像, (c)Image Domain CNN, (d)ADMM-CSNet, (e)eFREBAS-CNN, (f)CS-MRI

た、再構成像の平均 PSNR を図 6 に、各手法の学 習及びテストの所要時間を表 3 に示す. 図 4 よ り、eFREBAS-CNN では、収集信号量が 20-30% のときの PSNR と SSIM が、他手法と比較して最 も高くなった. また、信号量の増加に伴う数値の 上昇は次第に緩やかになった. 図 5、6 より再構 成像に注目すると、(b)のゼロフィル再構成像に 発生していたアーチファクトのほとんどは、(f)の CS-MRI を除くすべての手法で除去できた. また、 CS-MRI ではアーチファクトの重畳に伴って消 失した被写体構造を復元できない箇所が目立っ た. (c)、(d)、(e)の Image Domain CNN、ADMM-CSNet、eFREBAS-CNN では、いずれの手法でも 消失した構造の復元は良好に行えた. しかし、 ADMM-CSNet では、拡大画像で示したような細



図 5. 信号収集量 30%のときの各手法による再構成像の拡大画像: (a)フルデータ像, (b)ゼロフィル再構成像, (c)Image Domain CNN, (d)ADMM-CSNet, (e)eFREBAS-CNN, (f)CS-MRI

部構造において平滑化の傾向が見らた. eFREBAS-CNNと Image Domain CNN は、いずれ も平滑化の傾向が小さかったが、eFREBAS-CNN のほうがより鮮鋭な再構成を行えた.



図 6. 各種再構成法による再構成像の平均 PSNR

表	3.	各種再構成法の学習および再構成時間
---	----	-------------------

	学習時間[h]	再構成時間[s]	
eFREBAS-	0.65	0.150	
CNN	0.05 0.159		
Image Domain	0.5	0.036	
CNN	0.5		
ADMM-	0.5	1.06	
CSNet	9.5	1.00	
CS-MRI	-	35.5	

4. 考察

図4,5,6および表3で示したように、提案す る eFREBAS-CNN では、アーチファクトを除去 した鮮鋭な再構成像を短時間で取得できた.本 研究では eFREBAS 変換のスケーリング係数を D=3に設定し、3つのCNNで3つのグループの 再構成を行ったため、各 CNN が異なるアプロー チから再構成像の高品質化に寄与したと言える. Group A のサブ画像は MR 信号の低周波成分を多 く有しているため, アンダーサンプリングによ る影響はアーチファクトの重畳の方が強く、被 写体構造の消失の影響は小さい.一方の Group B, Cに含まれるサブ画像は、高周波成分を多く有し ており、アンダーサンプリングによる影響はア ーチファクトの重畳よりも構造の消失の方が大 きい. したがって, CNN の学習の結果, U-Net1 はアーチファクトの除去を主として動作し、U-Net2 と U-Net3 は消失した情報の復元を主とし て動作した. これは, 3 つの CNN がそれぞれ特 定の周波数成分の再構成に特化していると言え る.以上の要因により, eFREBAS-CNN ではアー チファクトの除去と高周波成分の復元を同時に 行えた. その一方で,図6より信号収集量が40% においては、ADMM-CSNetの PSNR が eFREBAS-CNN を上回った. ADMM-CSNet は、いずれの信 号量でもアーチファクトの除去性能が高い.信 号収集量が20%と30%の場合は、アーチファク

トの除去とともに画像の細部構造の損失が大き いが、40%の場合は細部構造の損失が少なくな ったため PSNR の向上につながったものと考え る.一方、提案法では、特徴として細部構造の保 存性に優れる特徴があり、信号量が少ない場合 は、優れた細部の構造保存特性が PSNR に貢献し たが、信号収集量 40%では、その特徴が一部の アーチファクトの残留に繋がり PSNR に影響し たと考える.

また,本研究では3つのCNNを同時に学習させたが,比較的短時間で処理が完了した.これは,各 CNN に与える画像が88×88のサブ画像であったため,通常のフルサイズで画像を与える場合よりもCNNのパラメータ数が減少したことによる.この点は,計算機のGPUおよびメモリの負荷の低減にもつながる.

現状での提案法の課題点として、さまざまな 位相分布を持つ画像を用いた位相への頑健性の 検討および、サンプリングパターンに応じた適 切なサブ画像のグループ分割法の検討の不足、 CNN アーキテクチャの改良等が挙げられる.

5. まとめ

本研究では、多重解像度解析手法と CNN-CS を 利用した、MR 画像のグループ型学習を基本とす る新たな深層学習再構成法を提案し、その有効 性を検討した.提案法では、位相を含んだ MR 画 像に対応する実用的手法であり、かつ、画像の細 部構造の保存性に優れる.信号収集量が少ない 場合に従来法を上回る信号対雑音比の画像を得 ることができた.

謝 辞

本研究の一部は、科学研究費補助金(19K0442 3,21J14120)および柏森情報科学振興財団の助成 により実施されました.また、本研究に使用し た MR 画像をご提供いただいた、キヤノンメデ ィカルシステムズに感謝の意を表します.

利益相反の有無

なし

文 献

- [1] Donoho DL: Compressed sensing. IEEE Transactions on Information Theory **52** : 1289-1306, 2006
- [2] Candès EJ, Wakin MB: An Introduction To Compressive Sampling. IEEE Signal Processing Magazine 25: 21-30, 2008
- [3] Lustig M, Donoho DL, Pauly JM: Sparse MRI: The application of compressed sensing for rapid MR imaging. Magnetic Resonance in Medicine 58 : 1182-1195, 2007
- [4] Wang S, Su Z, Ying L, et al. : Accelerating magnetic resonance imaging via deep learning. IEEE 13th International Symposium on Biomedical Imaging, Prague, Czech Republic 2016, 514–517
- [5] Hyun CM, Kim HP, Lee SM, et al. : Deep learning for undersampled MRI reconstruction. Physics in Medicine and Biology 2018, 63
- [6] Ito S, Yamada Y: Multiresolution Image analysis using dual Fresnel transform Pairs and Application to Medical Image Denoising. IEEE International Conference on Image Processing, Barcelona, Spain 2003, Map8.7.
- [7] Nakarmi U, Joseph YC, Edgar PR, et al.: Exploiting Coarse-Scale Image Features for Transfer Learning in Accelerated Magnetic Resonance Imaging. Proceeding of 2020 International Society for Magnetic Resonance in Medicine, Online 2020, 3595.
- [8] Ramanarayanan S, Murugesan B, Ram Km et al. : DC-WCNN: A Deep Cascade of Wavelet Based Convolutional Neural Networks for MR Image Reconstruction. IEEE 17th International Symposium on Biomedical Imaging, Iowa City, IA, USA 2020, 1069-1073
- [9] Lee D, Yoo J, Tak S, et al. : Deep Residual Learning for Accelerated MRI Using Magnitude

and Phase Networks. IEEE Transaction on Biomedical Engineering 65: 1985-1995, 2018

- [10] Vedaldi A, Lenc K : MatConvNet: Convolutional Neural Networks for MATLAB. Proceedings of the 23rd ACM international conference on Multimedia, Brisbane, Australia 2015, 689-692
- [11] Zhao F, Noll DC, Nielsen JF, et al. : Separate Magnitude and Phase Regularization via Compressed Sensing. IEEE Transactions on Medical Imaging 31 : 1713 - 1723, 2012
- [12] Yang H, Sun J, Li H, et al. : ADMM-CSNet: A Deep Learning Approach for Image Compressive Sensing. IEEE Transactions on Pattern Analysis and Machine Intelligence 42 : 521-538, 2020
- [13] Liu Y, Cai JF, Zhan Z, et al. : Balanced sparse model for tight frames in compressed sensing magnetic resonance imaging. PLoS ONE 2015, 10:e0119584

Study on Novel MR CNN-based Image Reconstruction

using Grouped Learning of Multi-resolution Images

Shohei OUCHI^{*1*2}, Satoshi ITO^{*1}

*1 Graduate School of Engineering, Utsunomiya University, *2 JSPS Research Fellow

Recently, it has been reported that CNN-CS outperforms traditional CS iterative method (CS-MRI) in terms of image quality and reconstruction time. In CS-MRI using CNN, it is difficult to restore the fine structure of images since some of those information is lost in signal acquisition stage. In this study, a novel transformed image domain CNN-CS method based on equal-sized multi-resolution image decomposition (eFREBAS transform) [1] was used to restore the detail of images. Since decomposed sub-images by eFREBAS transform have similarity each other, group of sub-image based CNN learning were proposed. The eFREBAS based CNN consists of three U-Net CNNs, each with multi-channel input and output of eFREBAS sub-images. Proposed method was applied to spatially phase varied images and compared with preceding methods. It was shown that proposed method is excellent in preserving the detail of the subjects. **Key words**: Deep Learning, Compressed Sensing, eFREBAS transform

信号空間の深層学習を利用した位相拡散フーリエ変換に基

づいた MR 超解像

山登 一輝*1 若槻 泰迪*1 伊藤 聡志*1

要旨

位相拡散フーリエ変換法(PSFT)の信号は、被写体が実関数であるという拘束条件を利用することで、撮像 後の数値処理によって、信号の外挿が可能であり、高解像度化した再構成像を取得できる.これまで反復 的手法により MR 信号を外挿し、再構成像の空間分解能改善を行ってきた.しかしながら、反復的方法で は分解能改善に伴って雑音の増幅が生じ、画像の再生には 20 回程度反復する必要がある.本稿では深層学 習再構成の一種である Generic-ADMM-Net を利用し、PSFT 信号の信号外挿と空間分解能の改善を図った. また.位相成分を含む複素数画像でも本手法が拘束条件として利用する実関数の条件を満足させるために、 位相推定と位相補正を行う手法を提案した.提案手法の有効性を確認するために信号収集率を 25%に制限 した PSFT 信号に適用した結果、高分解能画像が得られるとともに増幅雑音の低減を確認できた.

キーワード:超解像,位相拡散フーリエ法,MRI,深層学習

1. はじめに

核磁気共鳴映像法(Magnetic Resonance Imagi ng: MRI)における取得画像の高分解能化は,画 像診断精度に直結する重要課題の一つである. MRI の画像再構成で一般的に用いられるフー リエ変換に基づいた方法では,サンプリング対 象信号に対して,エイリアス回避のためサンプ リング定理を満足するように収集され,取得画 像の分解能が決定する.一般に,信号のサンプ リング長を大きくすることで,取得画像の分解 能を向上させることができる.しかしながら, MRI の撮像時間は 30 分から 1 時間と長いため 患者の負担になりやすく,高分解能化のための 信号取得時間の増加は望ましくない.

この問題に対して、先行研究では、位相拡散

*1 宇都宮大学工学部基盤工学科 〔〒321-8585 栃木県宇都宮市陽東 7-1-2〕 e-mail: yamato@is.utsunomiya-u.ac.jp

投稿受付: 2003 年 1 月 31 日

フーリエ変換(Phase-Scrambling Fourier Transfo rm: PSFT) [1] [2] に基づいた MR 画像の超解 像手法を提案した [3] [4]. PSFT は, 信号のダ イナミックレンジを低減するために提案され た方法であり,フーリエ変換でのパルスシーケ ンスの位相エンコードに同期して, 被写体のス ピンに二次の位相変調を与えることにより、す べてのスピンが同相になる時相を排除し、大振 幅の信号の出現を押さえることができる.この 二次の位相を与える処理は信号ダイナミック 低減以外に, エイリアスレス再構成, 単一信号 パラレルイメージング, エコープラナー法との 併用による高画質化など,多くの優れた特性が 明らかになっている [5]. 提案された MR 超解 像は, Gerchberg の超解像アルゴリズム [6] と 類似の反復アルゴリズムを利用し,低分解能な 信号に対して外挿することにより、事後的に再 構成画像の分解能改善を行う方法であり、その 反復回数は 10~20 回程度である. この方法は 少数の観測信号から画像を復元する点で圧縮 センシングに似ているが、ランダムパターンな どの特定のサンプリングパターンで信号を間 引き収集するのではなく、信号空間の収集領域 を制限することで実質的に撮像時間を短縮し た効果を得る.また、信号を間引くことが無い ので位相分布を求めることが可能であり、画像 再構成は圧縮センシングよりも比較的低演算 量で済むため計算時間が長くなることはない. しかしながら、この方法では鮮鋭化とともに画 像上の雑音が増幅する課題があった.

一方で,近年,医用画像に深層学習を用いる研究が注目されている.深層学習を,MR 画像の超解像や圧縮センシング MRI (CS-MRI) [7] に応用すると,従来の反復的手法よりも高品質な画像を再生できることが報告されている [8] - [10].特に,凸最適化手法の一種である交互方向乗数法(Alternating Direction Method of M ultipliers: ADMM)の計算ステップをモデルとして深層学習の層に組み込んだ Generic-ADMM-N et は,低収集率信号からの画像再構成に成功している [10].

本研究では、PSFT 信号を Generic-ADMM-Ne t を使用して外挿し,超解像的な画像拡大手法 を提案する.深層学習の利用により,エイリア スや雑音の抑制,高分解能化,再構成時間の短 縮化が期待できる [11].さらに,位相成分を含 む複素数画像に対して位相推定と位相補正を 導入することによって,複素数画像に対しても 実関数画像の拘束条件を適用した超解像を図 る.

2. PSFT 信号の外挿に基づいた反復超解像

2.1 **PSFT**

PSFT は、位相エンコードに同期して二次関 数状の磁界を一定の期間印加するか、あるいは 線形勾配磁界を印加のもと周波数変調された R F 励起パルスを照射するなどの方法により、被 写体内のスピンの位相を拡散させ、読み出し用 の勾配磁界によりエコー信号を読み出す方法 である.信号式を式(1)に示す.



図 1 PSFT 信号の反復的外挿の例 $v_{PSFT}(k_x, k_y) =$

 $\iint_{-\infty}^{\infty} \{\rho(x,y)e^{j\gamma b\tau(x^2+y^2)}\}e^{-j(k_xx+k_yy)}dxdy \qquad (1)$

ここで、 $\rho(x,y)$ はxy平面のスピン密度分布、bは 2 次関数状磁界の係数、 τ は印可時間、 γ は 核磁気回転比である.式(1)に示すように、位 相変調 $e^{-jyb\tau(x^2+y^2)}$ を施したスピン密度関数の フーリエ変換として表せる.PSFT 信号 $v_{PSFT}(k_x,k_y)$ からの画像再構成は、式(2)に示す ように、 $v_{PSFT}(k_x,k_y)$ に逆フーリエ変換 F^{-1} を 施した後に、二次の位相変調効果の逆位相関 数を乗じることで実現できる.

$$\rho(x,y) = e^{j\gamma b\tau \left(x^2 + y^2\right)} F^{-1} \left| v_{\text{PSFT}(k_x,k_y)} \right| \qquad (2)$$

2.2 PSFT 信号の外挿

本節では、PSFT 信号の反復的外挿に基づく 超解像手法について述べる.被写体関数が実関 数であるという拘束条件を与えた場合に、反復 的に信号を外挿することにより、少量サンプリ ングされた信号の分解能の向上を図る.図1に PSFT 信号の反復的外挿の例を示す.Gerchberg の超解像アルゴリズムと類似の反復アルゴリ ズムが利用され、以下の手順で実行される.

 MRI から取得されるサンプリング長 N の PSFT 信号の外側にゼロデータを充填し, サンプリング長を N から 2N に拡張する.



図 2 Generic-ADMM-Net のネットワーク構成

- 2. ゼロデータが充填された PSFT 信号に対し て逆フーリエ変換を適用し,補間処理を含 んだ画像再構成処理を行う.
- 再構成像に付与された位相変調を復調す るため、再構成像に二次の位相変調の逆位 相を乗ずる.
- 実関数の拘束条件により、復調した像の実 部を再構成画像とする.
- 5. 再構成画像に再び二次の位相変調を行い, 位相変調された像を計算する.
- 6. 位相変調像にフーリエ変換を適用することにより、サンプリング長 2N の PSFT 信号を計算する.
- サンプリング長 2N の PSFT 信号の中央部 分(サンプリング長 N)に対して、MRI から 取得された長さ N の PSFT 信号を真値とみ なし、置換する.
- 8. 真値置換が行われた PSFT 信号に逆フーリ エ変換することにより再構成を行う.
- (3)から(8)を反復して,再構成画像の空間分 解能の改善を図る.

本研究では、上記のアルゴリズムに深層学習 を導入し、さらなる空間分解能の改善を目指す.

3. Generic-ADMM-Net

本研究で利用する Generic-ADMM-Net につい て述べる.まず,Generic-ADMM-Net の基盤と なる計算アルゴリズムである ADMM について 説明する.ADMM とは,近接分離に基づく凸最 適化手法[12][13]の一つで,圧縮センシング, ロバスト統計,画像復元,機械学習など幅広く 利用されている. 一般的なCS-MRIモデルは式(3)の最適化問題 を求めることで再構成画像が推定される.

$$\hat{x} = \operatorname{argmin}_{x} \left\{ \frac{1}{2} | PFx - y|_{2}^{2} + \sum_{l=1}^{L} \lambda_{l} g(D_{l}x) \right\}$$
(3)

ここで、 $x \ge MR$ 再構成像、 $y \ge h$ 限して収集し た信号を示す. P は信号に制限を与える関数で あり、F はフーリエ変換を表している. D_l は離 散コサイン変換などの直交変換行列、 λ_l は正則 化パラメータ、 $g(\cdot)$ は正則化関数、L はフィルタ 数を表している. ADMM では拡張ラグランジュ 乗数法と変数分離を利用して、式(3)を式(4)に示 す 3 つの最適化問題に分割する.

$$\begin{cases} \arg \min_{x} \frac{1}{2} |PFx - y|_{2}^{2} + \sum_{l=1}^{L} \frac{\rho}{2} |x + \beta - z|_{2}^{2} \\ \arg \min_{z} \sum_{l=1}^{L} \lambda_{l} g(D_{l}z) + \frac{\rho}{2} |x + \beta - z|_{2}^{2} \\ \beta \leftarrow \beta + \tilde{\eta} (x - z) \end{cases}$$
(4)

ここで、 ρ はペナルティパラメータで、 $\beta = \alpha / \rho$ はラグランジュ未定乗数である。 $\hat{\eta}$ は、更新レートを示す. *I*を単位行列とすると、式(4)の解が式(5)のように表せる.

$$\begin{cases} x^{(n)} = F^{T} (P^{T} P + \rho I)^{-1} \\ [P^{T} y + \rho F (z^{(n-1)} - \beta^{(n-1)})] \\ z^{(n,k)} = \mu_{1} z^{(n,k-1)} + \mu_{2} (x^{(n)} + \beta^{(n-1)}) \\ - \sum_{l=1}^{L} \lambda_{l} D_{l\mathcal{H}}^{T} (D_{l} z^{(n,k-1)}) \\ \beta^{(n)} = \beta^{(n-1)} + \tilde{\eta} (x^{(n)} - z^{(n)}) \end{cases}$$
(5)

式(5)は n 回目の反復を表しており, $z^{(n,k)}$ は関数の最小値を探索する勾配法の一種である最急降下法を用いて算出されており, k は最急降下法の反復回数である. $\mu_1 = 1 - l_r$, $\mu_2 = l_r$ は, 最急降下法のステップサイズ l_r で決定され



図 3 提案する MR 超解像の概要

る値であり、 \mathcal{H} は $g(\cdot)$ の勾配に対する非線形変換である.

式(5)で再構成画像は取得できるが、最適なD₁ や \mathcal{H} の選択や、 ρ 、 $\tilde{\eta}$ などのパラメータを適切に 設定することは困難である. Generic-ADMM-N et は、これらのニューラルネットワークによる 学習によって、最適解を求めることができる. 図 2 に示す Generic-ADMM-Net の構造は, AD MMアルゴリズムの各処理に対応した層と処理 間のデータの流れに対応した矢印で表される. ADMM アルゴリズムの n 回目の反復は,図2 のn番目のステージに対応しており,各ステー ジは再構成層(X), 雑音除去層(Z), 乗数更新層 (M)から成り、Z層は加算層(A)、2つの畳み込み 層(C1, C2), そして非線形変換層(H)で構成される. n番目のステージを構成する $X^{(n)}$ 層, $Z^{(n,k)}$, $M^{(n)}$ 層が、式(5)の $x^{(n)}, z^{(n,k)}, \beta^{(n)}$ にそれぞれ対応し ている. Generic-ADMM-Net では, サンプリン グ長を制限された少数の信号を入力とし、各ス テージを通過することによって,最終的な再構 成画像を取得できる.

4. PSFT と Generic-ADMM-Net を利用した MR 超解像

2節で述べた手法は、サンプリング長 Nの P SFT 信号に対して外挿を行うことで、2Nに拡大 した画像を生成するアルゴリズムを提案し、真 値置換を反復することによって生成画像の分 解能向上を図った.この手法では、信号の外挿 処理の初期段階として、信号の外側にゼロを充 填する.ゼロが充填されることで、サンプリン グ長 2N で信号の中心部分のみ値のあるような 信号となる.この信号は一種のスパース性を満 たすため, 圧縮センシングによる復元処理を適 用できる. すなわち, ゼロを充填した信号を G eneric-ADMM-Net を用いることで, 解像度が 2 Nの再構成画像を取得できる. 図 3 に本稿で提 案する MR 超解像の概要を示す.

本研究での超解像を適用する対象は PSFT 信 号であることから, Generic-ADMM-Net へ入力 する信号も PSFT 信号である. それゆえ, Gene ric-ADMM-Net を PSFT 信号に対応できるよう に, 再構成層 $X^{(n)}$ を変更する必要がある. 具体 的には,式(5)のフーリエ変換 F を PSFT に変更 することで対応可能となる. 式(6),(7)に PSFT に対応した $X^{(n)}$ の順伝播式及び逆伝播式を示す.

$$\begin{aligned} x^{(n)} &= S^* F^T (P^T P + \rho I)^{-1} \\ \left[P^T y + \rho F S (z^{(n-1)} - \beta^{(n-1)}) \right] (6) \\ \frac{\partial E}{\partial \rho^{(n)}} &= \frac{\partial E}{\partial x^{(n)T}} S^* F^T (P^T P + \rho^{(n)} I)^{-1} \\ \left[F \left[S (z^{(n-1)} - \beta^{(n-1)}) \right] - \left(P^T P + \rho^{(n)} I \right)^{-1} \right] \end{aligned}$$

 $[P^T y + \rho^{(n)} FS(z^{(n-1)} - \beta^{(n-1)})]$ (7) なお、Sは二次の位相変調であり、式(8)で表さ れ、S*はSの複素共役を示す.

$$S = e^{-j\gamma b\tau \left(x^2 + y^2\right)} \tag{8}$$

5. 位相画像の超解像のための位相推定

2 節で説明した PSFT 信号の外挿に基づく反 復超解像手法では,被写体関数が実関数である という拘束条件があった.しかしながら,MRI によって撮像される画像は一般的に位相成分 を含む複素数画像であり,反復超解像手法の拘 束条件を満足できない.そこで,本研究では複 素数画像から位相推定することで実関数化す る処理を追加することにより複素数画像に対 応させた MR 超解像を構築した.図 4 に提案す る位相推定手法の流れを示す. 主磁界に由来す る位相分布は一般に MR 信号の帯域制限された 画像より推定できるが, PSFT 信号はその分布 が画像のおおまかな分布に対応するので PSFT 信号を帯域制限すると再生像の辺縁部が削れ てしまう.本研究では, PSFT 信号に 2 次の位 相変調を施すとフレネル変換の記述式を含ん だ信号v_{Frenel}に変形できる(フレネル変調と称 する)特徴を利用し,フレネル変換信号のフー リエ変換空間で帯域制限を行う方法を使用し た.フレネル変調は次式で表される.

$$v_{\text{Frenel}}(k_x, k_y) = e^{-j\frac{k_x^2 + k_y^2}{4\gamma b\tau}} v_{\text{PSFT}}(k_x, k_y) \qquad (9)$$

次に、フレネル信号に対してフーリエ変換を適 用することで得られる信号に対して主磁界の 不均一性に由来する位相分布を求めるために 帯域制限を施す.帯域制限した制限した信号に 対して、逆フーリエ変換、フレネル復調、逆フ ーリエ変換、位相復調と順に実行することによ って、帯域制限した信号から再構成された画像 ρoffが得られる.ここで、フレネル復調とは、式

(9)の $e^{-j\frac{k_{s}^{2}+k_{y}^{2}}{4ybt}}$ の複素共役を,逆フーリエ変換を 適用した信号に乗じることを指し示す. 位相推 定に必要な位相推定項 S_{pe} は, ρ_{off} の位相成分で あり,次式によって取得できる.

 $S_{\rm pe} = e^{-j \tan^{-1} \frac{\rm Im(\rho_{off})}{\rm Re(\rho_{off})}}$ (10)

ここで、Re(·), Im(·)は、それぞれ複素数の実部、 虚部を表す.取得された位相推定項*S*_{pe}は、式(6) の位相変調および位相復調と共に乗じる.位相 推定が加わったX⁽ⁿ⁾を次式に示す.

 $\begin{aligned} x^{(n)} &= S_{\rm pe} S^* F^T (P^T P + \rho I)^{-1} \\ \left[P^T y + \rho F S S_{\rm pe}^{*(z^{(n-1)} - \beta^{(n-1)})} \right] (11) \\ ここで, S_{\rm pe}^* は S_{\rm pe} の 複素共役である. \end{aligned}$

6. 計算機実験

6.1 実験条件

提案した MR 超解像手法の有効性を確認する ために計算機実験を実施した.本研究では,既 存の MR 画像から計算された PSFT 信号に対し



て,図3のように外側にゼロデータを充填され

た信号を適用して得られる信号を MRI から取 得される信号と想定して行った.ここで、図 3 の処理に位相推定を加えた手法を提案手法と して PSFT-PE-DL とした. 比較手法として,フ ーリエ変換信号を Generic-ADMM-Net を用いて 超解像を行う手法(FT-DL), 位相推定なしで PS FT 信号を超解像する手法(PSFT-DL),反復手法 に位相推定を組み合わせた手法(Iter-PE)とした. なお, Iter-PE の反復回数は 30 回とした. MR 画 像は、キヤノンメディカルから提供された健常 ボランティアの頭部 T1 強調像 80 枚とし, 学習 に65枚,テストに15枚をランダムに選択した. なお,使用画像は位相成分を含む複素数画像で ある. 再構成画像の定量評価には, Peak Signal -to-Noise Ratio(PSNR) & Structural SIMilarity in dex(SSIM) [14] を使用した. 計算機環境を表1 に, Generic-ADMM-Net の学習条件を表 2 に示 す. なお,表1,2の実験環境下で学習には約3 2時間を要した.

位相推定の効果

図 5 に位相推定により変化した複素数画像 を示す. 位相推定前の位相成分は変化が大きい ことが確認できる. 一方, 位相推定の適用結果 は, 実部が強度像に類似した像となり, 虚部は 被写体の輪郭部分以外を以外はほぼ0になって いる. 位相成分から, 被写体の位相がほぼ補正

表 1 計算機環	境
入力データサイズ	256×256
畳み込みフィルタサイズ	5×5
畳み込みフィルタ数	128
ステージ数	10
エポック数	300
パラメータ更新手法	L-BFGS-B [15]

表 2 学習条件				
CPU	Intel(R) Core(TM) i7-9700K(3.6GHz)			
RAM	32GB			
GPU	NVIDIA GeForce RTX 2070 SUPER			
	MATLAB R2019b			
Software	CUDA Toolkit 10.2			
	cuDNN 7.6.5			

表	3	定	量	評	価	結	果
1	2		- E	41	ΙШ	/PH	~1\

	PSNR[dB]	SSIM
FT-DL	30.64	0.943
PSFT-DL	27.56	0.913
Iter-PE	31.34	0.963
PSFT-PE-DL	31.53	0.963

されていることが確認できる.

また, PSFT 信号による超解像では被写体関 数が実関数であるという制限がある. 位相推定 後の複素数画像の実部は強度画像と類似して おり, 虚部はほとんど0となっている. 本稿で は, 位相推定を適用した複素数画像の実部を超 解像対象として扱うことで, 被写体関数に対す る拘束条件を満足させた. なお, 位相推定を適 用した場合, 学習時の教師データには, 位相推 定を適用したフルデータ像の実部を利用した.

6.3 超解像実験

15 枚のテスト画像に各手法を適用し再構成 した画像の PSNR,および SSIM を表 3 に示す. 位相推定を適用した Iter-PE および PSFT-PE-DL の 2 手法では,再構成画像に対して S_{pe}^* を乗し, 位相推定効果をキャンセルした画像で定量評 価を行った.FT-DL と PSFT-DL を比較すると, 定量評価結果は FT-DL のほうが高い結果とな った.一方で,Iter-PE は,深層学習に基づいた FT-DL および PSFT-DL を超える PSNR, SSIM を取得することができた.さらに,位相推定と





虚部(位相推定前)



強度像(位相推定前)









位相成分(位相推定前)



図 5 位相推定の適用による複素数画 像の変化.

深層学習を組み合わせた PSFT-PE-DL は, Iter-P Eを上回ることが確認された.以上の結果から, 定量的には位相推定の有効性が確認できた.

また、図 6に PSFT-PE-DL と比較手法による 超解像を行った結果を示す.図 6から、定量評 価が最も低かった PSFT-DL は、主観的にも他の 結果と比較して低品質であることが確認され る.FT-DL については、PSFT-DL ほどではない が、わずかに平滑化が加わっているように見え る.一方で、Iter-PE および PSFT-PE-DL は、主 観的に(a)の目標画像に近い画像を取得できる ことが確認できた.また、図 6 (k)-(o)に示すプ ロファイルを比べても、提案手法と Iter-PE は目

実部(位相推定後)



虚部 (位相推定後)



図 6 再構成結果.1行目は各手法による再構成画像,2行目は1行目の画像の拡大像,3行目は1 行目の画像の128列目のプロファイルをそれぞれ示す.

標画像に近いプロファイルを得られた.

6.4 考察

Amp.

今回の実験では、位相推定を用いることによ り、位相画像に対する再構成処理で高い PSNR と SSIM を得ることができた.図 6 では,再構 成結果として,再構成画像とその拡大画像,お よび画素プロファイルを示した. ここで, プロ ファイルをより詳しく考察するために、目標画 像と各手法間の差分プロファイルを測定した. 図 7 に目標画像のプロファイルと各手法のプ ロファイル間の差分プロファイルを示す.図7 から、位相推定を導入した Iter-PE および PSFT -PE-DLの差分値は,位相推定を導入していない 手法の差分値よりも全体的に小さいことが確 認できる.目標画像との差分値は,再構成画像 に含まれる雑音であると考えれば、位相推定の 導入により雑音を抑制しつつ再構成できると 考えられる.

なお、今回の実験のみでは分解能の向上効果 検証には不十分であり、今後の課題である.具 体的には、ファントム像などを使用した検証用 画像を使用した実験が必要である.

7. おわりに

本研究では、再構成画像を高分解能化するために、深層学習手法を利用する信号外挿法を提案した.本稿では、位相を含んだ複素数画像での分解能改善を図るため、新たに位相推定と位相補正処理を組み込む方法を提案した.実験では、位相補正処理を組み込むことにより、再構成像のPSNR改善が確認できた.このことから、位相画像に置いても超解像の可能性が示された.

謝辞

本研究は,科学研究費助成金(19K04423),およ び柏森情報科学振興財団,キヤノンメディカル システムズの支援を受けて実施されました.

利益相反の有無

なし

文 献

[1] Maudsley, A. A.: Dynamic range improve



図 7 目標画像と再構成画像間の差分プロファ イル

ment in NMR imaging using phase scram bling. Journal of Magnetic Resonance **76**: 287-305, 1969.

- [2] Wedeen, Van J., Yong Sheng Chao, and Jerome L. Ackerman.: Dynamic range co mpression in MRI by means of a nonline ar gradient pulse. Magnetic resonance in medicine 6: 287-295, 1988.
- [3] Ito, S, and Yamada. Y.: Compressed sens ing with super-resolution in magnetic res onance using quadratic phase modulation. Proc. Asia Pacific Signal and Information Processing Association Annual Summit a nd Conference.: 2012.
- [4] Ito, S. and Yamada, Y.: Improvement of spatial resolution in magnetic resonance i maging using quadratic phase modulation. 2009 16th IEEE International Conference e on Image Processing (ICIP).: 2009.
- [5] 伊藤聡志,中村俊,劉娜,山田芳文:位相 拡散フーリエ変換映像法を用いた MR ア ンチエイリアス画像再構成.電子情報通 信学会論文誌 D 90: 1149-1157, 2007.
- [6] Gerchberg, R. W.: Super-resolution throug h error energy reduction. Optica Acta: In ternational Journal of Optics 21: 709-720, 1974.

- [7] Pham, C-H, Ducournau, A. Fablet, R. an d Rousseau, F.: Brain MRI super-resoluti on using deep 3D convolutional networks. 2017 IEEE 14th International Symposiu m on Biomedical Imaging (ISBI 2017).: 2017.
- [8] Donoho, D. L.: Compressed sensing. IEEE Trans. Inf. Theo. 52; 1289-1306, 2006.
- [9] Yang, Y. Sun, J. Li, H. and Xu, Z.: Dee p ADMM-Net for compressive sensing M RI. Proc. the 30th international conferenc e on neural information processing syste ms: 2016.
- [10] Yang, Y. Sun, J. Li, H. and Xu. Z.: A DMM-CSNet: A deep learning approach for image compressive sensing. IEEE TPAMI, 42: 521-538, 2020.
- [11] 山登一輝,若槻泰廸,伊藤聡志:位相拡 散フーリエ法を用いた信号一画像間学 習に基づいた MR 超解像法の基礎検討, 信学技報 MI2021-6: pp.14-19, 2021.
- [12] Combettes, P. L., and Jean-Christophe.
 P.: Proximal splitting methods in signal processing. Fixed-point algorithms for i nverse problems in science and enginee ring. Springer: 185-212, 2011.
- [13] Parikh, N. and Stephen, B.: Proximal al gorithms. Foundations and Trends in op timization 1: 127-239, 2014.
- [14] Wang, Z., Bovik, A. C., Shelkh, H. R., and Simoncelli, E.P.: Image quality asse ssment: from error visibility to structura 1 similarity. IEEE TIP, 13: 600-612, 20 04.
- [15] Zhu, C., Byrd, R. H., Lu, P., and Noce dial, E. P.: Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale boun d-constrained optimization. ACM TOMS 23: 550-560, 1997.

MR Super-resolution based on Phase Scrambling Fourier Transform

Imaging Using Deep Learning for Signal Domain

Kazuki YAMATO^{*1}, Hiromichi WAKATSUKI^{*1}, and Satoshi ITO^{*1}*1 School of Engineering, Utsunomiya University

The signal of the phase-spread Fourier transform (PSFT) imaging can be extrapolated by numerical processing using the constraint that the subject is a real-value image, which results in the improvement of spatial resolution of obtained images. To extrapolate PSFT signals, iterative reconstruction method has been used, however, the iterative method amplified the noise on the images, and it took about 20-time iterations to reconstruct images. In this study, Generic-ADMM-Net, which is a kind of deep learning reconstruction, is used to expand the PSFT signal. Furthermore, in order to apply the proposed method to phase varied images, phase estimation and phase correction was incorporated in the deep learning network. Simulation experiments were conducted to verify the effectiveness of the proposed method. PSFT signals with the signal sampling rate of 25% were used for the input of the deep neural network and trained the network. It was shown that, spatial resolution on the images could be improved with smaller noise amplification.

Key words: Phase-scrambling Fourier transform, MRI, Deep learning, and Super-resolution

位相拡散フーリエ変換法を導入した MR 深層学習における

信号間引きパターンの検討

金澤 崇大*1 山登 一輝*1 伊藤 聡志*1

要旨

MRI に圧縮センシングを応用する場合に 2 次関数状の位相変調を利用する位相拡散フーリエ変換法を使用 すると再構成像の画質を改善できることが報告されている.これまでに交互方向乗数法 (ADMM) を CNN 化した ADMM-Net を使用し, 信号収集比が 33%の場合に汎用のフーリエ変換映像法に比べ高品質な画像を 再生できることを示した.本研究では, 信号収集比を 25%から 40%と変え, また, 信号間引きパターンを 等間隔, ランダム,ポアソン分布等に変えた場合の検討を行った.結果, 信号収集率によって最適な位相 拡散係数が変化すること, および, フーリエ変換法では実現不可能な等間隔間引きを使用しても画像再構 成が可能であることが示された.

キーワード: MRI, 圧縮センシング, 深層学習, 位相拡散フーリエ変換法

1. はじめに

MRIでは,信号が回復する緩和過程を待って から次の信号を取得するためにX線CTに比べ ると撮像時間が長くなる.撮像自体は無痛であ るが,閉塞した空間に長時間留まることは患者 にとって大きな負担となる.そのため,撮像時 間を短縮する研究が行われている.圧縮センシ ング (Compressed Sending:CS)[1]を MRI に 応用した CS-MRI [2]は,少数の信号から良質 な MR 画像を再構成する方法であり,数理的に 撮像時間の短縮を図ることができるため,様々 な研究が行われている.

CS-MRI は L1 ノルムを含む最適化問題を解 くことで再構成像を推定する. CS-MRI の最適 化問題を解く方法として, 交互方向乗数法

*1 宇都宮大学大学院地域創生科学研究 科工農総合科学専攻情報電気電子システ ム工学プログラム

〔〒321-8585 宇都宮市陽東 7-1-2〕 e-mail: mc216821@cc.utsunomiya-u.ac.jp (alternating direction method of multipliers:
 ADMM) [3] がある. ADMM は収束保証を伴う 変数分割アルゴリズムであり, Bregman 法[4] と比べて高画質な画像を再構成できることが 報告されている[5]. 一方で, ADMM における再 構成像に影響を及ぼすパラメータは経験則に 基づいて決定する場合が多く, 最適なパラメー タの決定は困難である.

近年では、CS-MRI に深層学習を導入した研 究が注目されており、従来の反復的手法よりも 高速かつ高品質な画像再構成を行えることが 報告されている[6-9]. ADMM の反復アルゴリズ ムを深層学習に導入した Generic-ADMM-Net [10] は、ADMM における最適な重み係数やパ ラメータを学習するため、従来の反復再構成よ りも高速かつ良質な画像再構成を行うことが できる.

また, CS-MRI の高画質化の一手法として, 汎用のフーリエ変換映像法に非線形磁界によ る位相エンコーディングを導入した位相拡散 フーリエ変換法がある[11,12]. 位相拡散フーリ エ変換法は, 画像の縮小スケーリング機能を使

用し,標本化定理を満足しない信号から折り返 しのない画像を再構成するエイリアスレス再 構成や[13]、パラレルイメージングの適用など 柔軟な再構成処理をとることができる[14,15]. また,位相拡散フーリエ変換法の式に含まれる 2 次の位相変調は、信号間引きを行った場合に 再生像の誤差を拡散する効果があり、CS-MRI に応用することで汎用のフーリエ変換映像法 よりも高画質な再構成像が得られることや [16,17], 一般的な CS-MRI では困難な等間隔信 号間引きによる画像再構成が可能であること [18]が報告されている. 再構成処理に深層学習 を使用する方法として, ADMM をネットワーク 化した方法により,信号収集率 33%の場合につ いて評価を行った[19]. その結果, ランダム間引 きよりも等間隔間引きにおいて信号対雑音比 の高い画像を再生できることを示した.

本研究では、信号間引きパターンとして等間 隔とランダムにポアソン・ディスクサンプリン グ[20]を加え、また、信号収集率を25%、33%、 40%とした場合の再構成実験を行った.信号対 雑音比、構造的特徴などの観点から提案法と汎 用法であるフーリエ変換映像法との比較を行 ったので報告する.

2. 位相拡散フーリエ変換法

位相拡散フーリエ変換法(Phase Scrambling Fourier Transform imaging: PSFT)とは、2次関 数状の磁界を一定期間印加するか、あるいは線 形勾配磁界を印加のもと周波数変調された RF 励起パルスを照射するなどの方法により、被写 体内のスピンの位相を拡散させてから汎用の フーリエ変換映像法によって撮像シーケンス を実行する手法である.PSFTの信号式を式(1) に示す.

$$v(k_x, k_y) =$$
$$\iint_{-\infty}^{\infty} \{\rho(x, y) e^{-j\gamma b\tau(x^2 + y^2)}\} e^{-j(k_x + k_y)} dx dy \qquad (1)$$

ここで, k_x , k_y はx, y方向の空間周波数, $\rho(x,y)$ は被写体内のスピン密度分布, γ は磁気



図1 位相拡散係数と信号拡がりの関係

回転比, bは2次関数状磁界の係数, rはその印 加時間である.PSFT 信号からの画像再構成は, 式(1)の信号を逆フーリエ変換した後, 与えた2 次の位相変調項の逆位相を乗じることで実現 できる.PSFT 信号からの画像再構成式を式(2) に示す.

$$\rho(x, y) = e^{j\gamma b\tau (x^2 + y^2)} F^{-1}[v(k_x, k_y)]$$
(2)

標本化定理より、画像上で隣接する画素間 Δx の 位相変化が π のとき、被写体に与えることがで きる位相変化量が最大となる.画像空間上のx方向で考えるものとし、標本化定理を満足する 最大座標を $m\Delta x$ とすると、 $\gamma b \tau$ は式(3)の条件を 満足する必要がある.

$$\Delta x \frac{\partial (\gamma b \tau x^2)}{\partial x} |_{x = m \Delta x} = 2\gamma b \tau m \Delta x^2 \le \pi$$
(3)

ここで,被写体のサイズによらない条件として 視野幅全域で標本化定理を満足する場合を考 える.式(3)について, *m = N/2*とおくと式(4) を得る.

$$\gamma b\tau \le \frac{\pi}{N\Lambda r^2} \ (= \overline{\gamma b\tau})$$
 (4)

撮像パラメータγbτを, $\overline{\gamma b\tau}$ を基準として表現す るために係数hを導入し,新たにγbτ = h $\overline{\gamma b\tau}$ とす る. hと信号拡がりの関係を示すために 256 点 のデータ列の中央に 200 点の幅を持つ矩形のス ピン密度分布 $\rho(x,y)$ を置き, hと PSFT 信号 $v(k_x,k_y)$ の分布を調べた.図1に示す.ここで, *h* = 0の場合は汎用のフーリエ変換映像法に相当する.
 *h*の値が大きいほど拡散効果が大きくなり、PSFT 信号は k 空間の広い領域に分布する関係がある.

汎用のフーリエ変換映像法において信号を ランダムに収集すると、アーティファクトのコ ヒーレント性が低下する. PSFT では 2 次関数 状の位相変調によりアーティファクトのコヒ ーレント性が低下し、誤差成分が信号空間上と 画像空間上で拡散する. このため、PSFT では必 ずしもランダムに信号を間引く必要がなくな る.

3. Generic-ADMM-Net

3.1. 交互方向乗数法(ADMM)

Generic-ADMM-Net のアルゴリズムの基盤と なる交互方向乗数法 (ADMM) について説明す る. ADMM とは、2 つの関数による最適化問題 に対し、拡張ラグランジュ法を適用する方法で ある. 一般的な CS-MRI 最適化問題を式(5)に示 す. ここで、xは推定したい MR 再構成像、Pは 信号収集を行う関数、Fはフーリエ変換、Lはフ ィルタ数、 λ は正則化パラメータ、Dはスパース 変換行列、 $g(\cdot)$ は正則化関数を表す.式(5)を解 くことで MR 再構成像を推定する.

$$\underset{x}{\operatorname{argmin}} \left\{ \frac{1}{2} \| PFx - y \|_{2}^{2} + \sum_{l=1}^{L} \lambda_{l} g(D_{l}x) \right\}$$
(5)

ADMM は式(5)について, *x* = *z*と置き, 拡張ラ グランジュ法を導入することにより, 式(6)に示 す 3 つの最適化問題へと帰着する.

$$\begin{cases} \arg\min_{x} \left\{ \frac{1}{2} \| PFx - y \|_{2}^{2} + \frac{\rho}{2} \| x + \beta - z \|_{2}^{2} \right\}, \\ \arg\min_{z} \left\{ \sum_{l=1}^{L} \lambda_{l} g(D_{l}z) + \frac{\rho}{2} \| x + \beta - z \|_{2}^{2} \right\}, \\ \arg\max_{\beta} \langle \beta, x - z \rangle. \end{cases}$$
(6)

ここで、 $\beta = \alpha / \rho$ はラグランジュ乗数、 ρ はペナ ルティパラメータを表す.式(6)の最適化問題に ついて、xは微分または平方完成により 0 とな る条件から、zは最急降下法によって導出でき、 式(7)のように表せる.

$$\begin{cases} x^{(n)} = F^{T} (P^{T} P + \rho I)^{-1} \\ [P^{T} y + \rho F (z^{(n-1)} - \beta^{(n-1)})], \\ z^{(n,\xi)} = \mu_{1} z^{(n,\xi-1)} + \mu_{2} (x^{(n)} + \beta^{(n-1)}) \\ - \sum_{l=1}^{L} \lambda_{l} D_{l}^{T} \mathcal{H} (D_{l} z^{(n,\xi-1)}), \\ \beta^{(n)} = \beta^{(n-1)} + \tilde{\eta} (x^{(n)} - z^{(n,\xi)}). \end{cases}$$
(7)

式(7)のnは回目の反復処理を表しており、Iはサ イズ $N \times N$ の単位行列、 μ_1, μ_2 は重みパラメータ、 Hはgに対する非線形変換、 $\hat{\eta}$ は更新率を表す. また、 ξ は最急降下法の反復回数を表し、最急降 下法のステップサイズを l_r とすると $\mu_1 = 1 - l_r \rho, \mu_2 = l_r \rho と$ なる.

3.2. PSFT-ADMM-Net

汎用法であるフーリエ変換映像法に対し, ADMM をネットワーク化した構造(Generic-ADMM-Net)を図2に示す.ステージnはADMM のn番目の反復処理を表しており,再構成層 $X^{(n)}$, 誤差除去層 $Z^{(n,\xi)}$, 乗数更新層 $M^{(n)}$ で構成される. また, $Z^{(n,\xi)}$ は N_t 回の反復を行うサブステージと なり,畳み込み層 $C_1^{(n,\xi)}, C_2^{(n,\xi)}$, 非線形変換層 $H^{(n,\xi)}$,加算層 $A^{(n,\xi)}$ で構成される.ここで, $X^{(n)}, Z^{(n,\xi)}, M^{(n)}$ は式(7)の $x^{(n)}, z^{(n,\xi)}\beta^{(n)}$ の処理に 対応している.ネットワークのフローは, k空 間上で間引いた信号を入力とすると,各ステー ジの処理を繰り返し,最終的に再構成層によっ て MR 画像が出力される.本研究では,図2の



図2 Generic-ADMM-Netの構造

ネットワークに 2 次の位相変調と復調処理を 加えた PSFT-ADMM-Net を構成した. PSFT 信 号を入力とした場合の再構成層*X*⁽ⁿ⁾の処理を式 (8)に示す.

$$x^{(n)} = S^* F^T (P^T P + \rho I)^{-1} \left[P^T y + \rho SF (z^{(n-1)} - \beta^{(n-1)}) \right]$$
(8)

ここで、位相変調項 $S = e^{-jyb\tau(x^2+y^2)}$ であり、 S^* はSの複素共役である、式(8)は順伝播の式であるが、逆伝播も同様の処理を追加する.

4. シミュレーション条件

4.1. 実行環境

MR 数値モデルより PSFT 信号を合成し,計 算機内で間引き処理を行った. PSFT 信号の $\overline{\gamma b \tau}$ 値は $N = 256, \Delta x = 0.1 \text{cm}$ のとき $\overline{\gamma b \tau} = 1.227 \text{rad}/$ cm²となり, $h = 0.1 \sim 1.0$ の範囲で 0.1 間隔で再 構成を行った.使用した MR 画像は,IXI-dataset [20] に含まれる,Philips 社製 MRI (1.5T) によ って撮像された 256×256 の頭部プロトン密度 強調像である.学習用に 100 枚,テスト用に 25 枚を使用した.学習とテストに使用した計算機 環境を表1に示す.学習及びテストには GPU を 使用し,1 回の学習には 15 時間を要した.表 2 に学習条件を示す.再構成像の定量評価には, ピーク信号対雑音比 (Peak Signal-to-Noise Ratio: PSNR)を使用した.

	PK = PT JT DAVR JE
CPU	Intel Corei7-8700
GPU	NVIDIA GeForce GTX1070Ti
RAM	32GB (DDR-240)
Software	MATRAB 2017b
	CUDA Toolkit 9.0 / cuDNN 7.0.5

表1 計算機環境

表 2 学習条件

エポック数	150
オプティマイザ	Adam [22]
フィルタ数	128
フィルタサイズ	5×5
ステージ数	10

4.2. 信号間引きパターン

本研究で使用した信号間引きパターンを図 3 に示す.臨床で使用されることの多いカルテシ アン座標系を使用するものとし,信号収集率 25%,33%,40%とした等間隔間引き,ランダム 間引き,ポアソン・ディスク間引きをそれぞれ 使用した.

5. 再構成実験

5.1. 信号収集率と信号間引きパターンの比較 位相拡散係数h = 0.1~1.0において,信号収集率 25%,33%,40%の各信号間引きパターンを使用 して得られた PSNR をプロットしたものを図 4 に示す.各信号間引きパターンにおいて信号収 集率 25%の場合はh = 0.7,33%の場合はh = 0.5, 40%の場合はh = 0.3で比較高い PSNR を示す傾 向が見られた.信号収集率 25%から 40%におい て,どの信号収集率においても等間隔間引きが 最も高い PSNR を示した.

5.2. FT 信号と PSFT 信号の比較

FT 信号と, PSFT 信号より再構成した像を比較する.FT 信号の再構成に用いた信号間引きパターンを図 5 に示す.FT 信号の再構成では,中央 30 行を連続収集した信号間引きパターンを使用し, PSFT 信号の再構成には図 3 の等間隔間引きを使用した.PSFT 信号は信号収集率25%,33%,40%でそれぞれ*h* = 0.7,0.5,0.3とした.表







━━ 等間隔 ━━ ポアソン ━━ ランダム

(c) 40%



3 に PSNR, 図 6 に再構成像を示す. 表 3 より, 信号収集率 25%, 33%の場合は PSFT 信号が FT 信号の PSNR を大きく上回った. また, 図 6 よ り FT 信号の再構成像は信号収集率が低い場合 に構造の消失が見られる. 対して, PSFT 信号の



図5 FT 信号で使用した信号間引きパターン

表3 FT 信号と PSFT 信号の PSNR

	25%	33%	40%
FT信号	30.24	33.03	43.16
PSFT信号	35.78	40.06	41.70

再構成像は信号収集率が低い場合でも詳細構 造を保存できているが,雑音状のアーティファ クトが表れていることが確認できる.

6. 考察

図4より、信号収集率によって PSNR が最 大になるhの値が変化した. 再構成像に誤差が 生じる理由に, 主に信号間引きによる再生誤 差と信号打ち切りにより実関数にならない要 因の2つがある.信号の収集率が小さい場合 は,信号間引きによる再生誤差が支配的な誤 差となる. PSFT 信号の再構成処理に含まれる 2次の位相変調と復調処理は再生誤差を拡散さ せる効果があるため, 誤差の拡散効果が大き いhの値が大きいときに PSNR が高くなったも のと考えられる.一方,信号収集率が大きい 場合は、hの値が比較的小さいときにピークの 値が得られた.これは、信号間引きによる誤 差が減り、相対的に信号が実関数にならない ことによる誤差が大きくなったことが原因と 考える.図1のように信号拡がり幅が小さい 信号では、信号の打ち切りによる信号損失が 小さく,再生像のうち虚部は小さくなるた め,実関数化した場合の誤差が小さくなる. そのため、PSNR 最大値はhが小さい方向にシ フトしたものと考える.

図4より,いずれの信号収集率においても 等間隔間引きが最も高い PSNR を示した.ラ ンダム間引きは,信号収集点が乱数によって


図 6 再構成像: (a) 目標画像, (b)-(d) FT 信号の 再構成像, (e)-(g) PSFT 信号の再構成像, (b,e) 25%, (c,f) 33%, (d,e) 40%

決定される.そのため、収集点間隔が広い領域 で収集信号の損失があるのに対し、等間隔間引 きは信号空間の低域から高域まで等密度に信 号を収集できるため、収集信号の損失が小さく なったことが考えられる.

7. まとめ

PSFT 信号を使用した深層学習画像再構成を 行い,信号収集パターンと再構成像との関係を 検討した.その結果,いずれの信号収集率でも 等間隔間引きが最も良い結果が得られた.また, 信号取得率が25%と33%の場合に提案法はフー リエ変換映像法に比べて大幅に高いPSNRを示 した.今後は位相を含む画像の再構成について 検討する予定である.

謝辞

本研究の一部は,科学研究費補助金 (19K04423),および柏森情報科学振興財団 の助成によって実施された.また,MR 画像デ ータを提供いただいた Information eXtraction from Images プロジェクトに感謝の意を表しま す.

利益相反の有無

なし.

文 献

- [1] Donoho DL: Compressed sensing. IEEE Trans Inform Theor52: 1289-1306, 2006
- [2] Lustig M, Donoho DL, Santos JM, et al: Compressed sensing MRI. IEEE Signal Process Mag 25: 72-82, 2008
- [3] Boyd S, Parikh N, Chu E, et al: Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. Foundation and Trends in Machine Learning3: 1-122, 2010
- [4] Yin W, Osher S, Goldfarb D, Darbon J et al: Bregman Iterative Algorithm for 11-Minimization with Applications to Compressed Sensing. SIAM Journal on Imaging Sciences1: 143-268, 2008
- [5] Wang Y, Yin W, Zeng J: Global convergence of ADMM in noncnvex nonsmooth optimization arXiv: 1511.06324, 2015
- [6] Yang Y, Sun J, Li H,, et al: Deep ADMM-Net for Compressed Sensing MRI. Advances in Neural Information Processing Systems: 10-18, 2016
- [7] Wang S, Su Z, Ying L, et al: Accelerating

magnetic resonance imaging via deep learning. IEEE International Symposium on Biomedical Imaging: 514-517, 2016

- [8] Yang G, Yu S, Dong H, et al: DAGAN: Deep De-Aliasing Generative Adversarial Networks for Fast Compressed Sensing MRI Reconstruction. IEEE Transactions on Medical Imaging37: 1310-1321, 2018
- [9] Han YS, Lee D, Yoo J, et al: Accelerated projection reconstruction MR imaging using deep residual learning. In proceedings of 25 Annual Meeting of International Society for Magnetic Resonance in Medicine, Hawaii, 2017, O690
- [10] Yang Y, Sun J, Li H, et al: ADMM-CSNet: A Deep Learning Approach for Image Compressive Sensing. IEEE Transactions on Pattern Analysis and Machine Intelligence42: 521-538, 2020
- [11] Maudsley AA: Dynamic range improvement in NMR imaging using phase scrambling: J Magn Reson 76: 287-305, 1988
- [12] Wedeen VJ, Chao YS, Ackerman JL: Dynamic range compression in MRI by means of a nonlinear gradient pulse. Magn Reson Med6: 28-295, 1988
- [13] 伊藤聡志,川和康弘,山田芳文:位相拡 散フーリエ法に仮想的パラレルイメージングを適用したエイリアスレス画像 再生法の検討.MEDICAL IMAGING TECNOLOGY28: 118-126, 2010
- [14] Ito S, Yamada Y: Parallel Image Reconstruction using a Signal In Magnetic Resonance Imaging. IEEE International Conference on Image Processing 2008, San Diego, 2008, 2964-2967
- [15] 伊藤聡志,山田芳文:位相拡散フーリエ 変換法を使用した単一信号 PILS.日本 磁気共鳴医学会論文誌 31: 20-31, 2011.
- [16] Ito S, Osada K, Yamada Y: Compressed Sensing using FREBAS transform in Phasescrambled Fourier Transform Imaging

Technique. The37th Annual Meetings of Japanese Society for Magnetic Resonance in Medicine, P-1-26, p316, Japan, 2009.

- [17] Wiaux Y, Puy G, Gruetter R, et al: Spread Spectrum for Compressed Sensing Techniques in Magnetic Resonance Imaging. IEEE ISBI, 756-759, Rotterdam, The Netherlands, 2010
- [18] 八坂俊吾,伊藤聡志:位相拡散フーリエ 変換映像法を利用した等間隔間引きを 基本とする MR 圧縮センシング.電子 情報通信学会論文誌 J101-D: 27-35, 2018
- [19] 植松駿,伊藤聡志:位相拡散フーリエ 変換映像法の導入による MR 深層学習 再構成の高画質化.MEDICAL IMAG-ING TECHNOLOGY39: 59-67, 2021.
- [20] Cook RL: Stochastic sampling in computer graphics. ACM Transactions on Graphics5: 51-72, 1986
- [21] IXI Dataset. <u>https://brain-development</u>. org/ixi-dataset/
- [22] Kingma DP, Ba J: Adam: A method for stochastic optimization. arXiv: 1412.6980

Study on Signal Under-sampling Pattern in MR Deep Learning based Image Reconstruction Using Phase Scrambling Fourier Transform Technique

Sota KANAZAWA^{*1}, Kazuki YAMATO^{*1}, Satoshi ITO^{*1}

*1 Graduate Program in Information, Electrical and Electronic Systems Engineering, Utsunomiya University

Reducing image acquisition time is one of the major issues for clinical use of MRI. An effective way to speed up MRI is the under-sampling of acquisition signal. It was reported that image quality could be improved in MRI compressed sensing (CS) by adopting the phase scrambling Fourier transform imaging (PSFT) that uses quadratic phase modulation to the subject. In this paper, a deep learning based CS image reconstruction using PSFT signal was examined to improve the quality of CS images.

Simulation results showed that image sharpness, structural similarity and image contrast were superior to standard Fourier transformed imaging in proposed method. In the comparison of random under-sampling and non-random under-sampling in proposed method, higher PSNRs and SSIMs were obtained in non-random under-sampling based CS. These studies indicate that the application of quadratic phase scrambling Fourier transform imaging is the promising method to reduce the image acquisition time or to improve the image quality in CS.

Key words: MRI, Compressed Sensing, Deep Learning, Phase Scrambling Fourier Transform

MR画像からの3次元導電率分布の回帰によるTMSのため

の誘導電場の回帰精度の改善

牧 豊大*1 横田 達也*1 平田 晃正*2 本谷 秀堅*1

要旨

TMS とはうつ病の治療などを目的として導入された治療法であり、頭部にコイルを当て磁場を発生させる ことで、脳内の神経細胞を非侵襲的に電気刺激する手法である. TMS を行う上で重要なことは電気刺激に 最適なコイルの位置を定めることであり、そのためにはコイルが脳内に誘導する電場を高速に推定する必 要がある.本稿では、コイルが脳内に誘導する電場を実時間で推定するために、患者の頭部 MRI 画像から 誘導電場を回帰するニューラルネットワークを二通りの手法で実現し、その性能を比較する.一つ目の手 法は頭部 MRI から誘導電場を直接回帰し、もう一つは中間表現として頭部画像を導電率の異なる領域へと 分割する.実験の結果、後者の、誘電率分布を求めた上で電場を推定することにより、画像から直接推定 するより推定精度が改善出来ることを確認した.

キーワード:経頭蓋磁気刺激法,深層学習,MRI 画像,電場推定

1. はじめに

非侵襲的な電気刺激は、神経科学や脳神経外 科学、神経疾患の治療に使われている [1]. 経頭 蓋磁気刺激法(TMS)はその1つである. TMS で は、電流を流したコイルを頭部に当て、脳の神 経細胞を刺激する. 電気刺激された脳の部位と 電気刺激による身体の反応との対応で、脳機能 を特定することができる.

臨床の場における問題は,頭部に当てたコイ ルが脳のどの部位を刺激しているのか正確に 知る手段がないことである.脳内の誘導電場を 計測することは困難である.このため刺激部位 を知るためには,コイルが脳内に誘導する電場

*1 名古屋工業大学大学院情報工学専攻 [〒466-8555 名古屋市昭和区御器所町] e-mail: t.maki.985@nitech.jp *2 名古屋工業大学大学院電気機械工学専攻 投稿受付: 2021年6月18日 をなんらかの方法で推定しなければならない. 代表的な推定法の1つは電磁気学に基づく誘導 電場のシミュレーションである[2].

誘導電場は磁場と電場ならびに媒質の導電 率の関係を示すマクスウェルの方程式に基づ きシミュレーション計算される.有限要素法に よる高速シミュレーション法は頭部の誘導電 場を約3秒で計算する [2].患者頭部のMRI 画 像を撮影し,仮想的にコイルを配置したときの 誘導電場をシミュレーションするためには,ま ず頭部の3次元導電率分布画像を作成しなけれ ばならない.導電率分布画像の作成に際しては, 頭部 MRI 画像を部位ごと (e.g. 灰白質,白質) に領域分割し,各部位に対応する導電率を割り 当てる.脳の各部位の導電率は既知である [3]. 導電率分布画像を求める際のボトルネックの1 つは領域分割の精度が十分ではなく、その間違



図 1: コイル直下の領域で切り出し、それを学習データとして用いる

いを手動で修正しなければならないことであ る. 著者らは近年, 頭部 MRI の領域分割を行わ ず, 頭部 MRI より直接誘導電場を回帰するニュ ーラルネットワークを構築した[4]. 一方, [5] では, ニューラルネットワークを用いて頭部 MRI 画像を領域分割することにより導電率分 布画像を構築し,物理シミュレーションをする ことで誘導電場を高速推定している.

そこで、本稿では、頭部 MRI 画像の領域分 割及び導電率画像の推定を行うニューラルネ ットワークに続けて、導電率画像から誘導電 場を回帰するネットワークを接続する.前段 のニューラルネットワークによる領域分割と 導電率分布画像の推定精度は、手作業による 修正を含む従来法と比べて推定精度が劣る可 能性がある一方で、電場に影響を与える導電 率分布の情報を後段のニューラルネットワー クに提供する. MRI 画像から直接電場を回帰 する手法と、導電率分布画像の推定を経て回 帰する手法のいずれが高精度となるかは、自 明では無い.本研究では、実験によりこれら の手法の精度を比較する.

2. 手法

本研究では、GE 3 T with fast SPGR sequence で 撮像された 18 人の被験者の頭部 MRI 画像をデ ータセットに用いた. 画像のサイズは256 × 256×176であり,分解能は1.00mm×1.00mm× 1.00mmである. 導電率分布画像は,FreeSurfer [6]を用いて頭部 MRI 画像を脳の部位ごとに領 域分割し,誤りを手動で修正することにより構 築した.TMS で広く使われている8の字コイル の位置を様々に変化させて,誘導電場を手法[2] で計算した.コイルの位置はそれぞれの被験者 に対し,左右の脳両方に配置した14×14のグ リッド上の各点である.コイルの方向について は各点を中心に頭皮の法線に対して10度ごと に360度回転させた.従って1症例あたり, 14×14(グリッド)×2(左右)×36(角度)= 14.112通りの誘導電場が得られる.

8 の字コイルはコイル直下に大きな電場を誘 導することが知られている [7]. そこで文献[4] と同様に,8の字コイル直下のみの MRI 画像の 分割や誘導電場の回帰を行う.具体的には,図 1 に示すように,72mm×144mm×24mmの領 域のみを入出力とする.

被験者を学習データに 15 人, テストデータ に3人に分けて,以下の実験を行った.

- i. MRI 画像から誘導電場を直接回帰
- ii. シミュレーションに用いた導電率画像 から誘導電場を回帰
- iii. ニューラルネットワークにより推定したた導電率分布画像から電場を回帰



図 2:回帰結果とその推定精度の可視化

i. では頭部 MRI 画像の領域分割を行わず,頭部 MRI 画像から誘導電場の回帰を直接行う.ii. で は、参照のためにシミュレーションで用いた正 確な導電率分布画像から誘導電場を推定する. iii. では、異なる導電率を持つ17の領域で頭部 MRI 画像を分割することにより導電率分布画 像を作成し、その導電率画像をニューラルネッ トワークの入力として、誘導電場を回帰した.

電場の回帰には先行研究[4]で提案されている U-Net を採用した.ただし,[4]では電場強度 のスカラ場の回帰であったが本研究ではベク トル場を回帰する.すなわち,脳の各位置で回 帰される量はスカラではなく,3次元ベクトル である.領域分割する際には,U-Net の最後3 つの畳み込み層の出力のチャンネルを17とし た.誤差関数は,電場の回帰では2乗誤差,領 域分割では交差エントロピーである.最適化ア ルゴリズムには Adam[8]を用いた.

回帰したボクセル全体と大脳皮質表面で回 帰精度を評価する.評価指標には,真値と推定 値の強度の2乗誤差,角度の差を用いた. さらに、誘導電場を実時間回帰できる利点を 活かし、コイルの位置とパラメータ空間を全探 索することにより、最適なコイル位置を求めた. 最適なコイル位置には、コイルで刺激したい部 位を最も強く刺激し、なおかつそれ以外の部位 を刺激しないことが求められる.加えて、誘導 電場と神経細胞の向きが一致していることが 望ましい.本研究では神経細胞の向きを大脳皮 質の法線とする.そのため、指標として、以下 の2つを用いた. S_1 は電場の強度の比を示し、 S_2 は電場と法線の内積の比を表す.

$$S_{1}(\mathbf{y}) = \frac{\sum_{i \in \Omega} \|\mathbf{y}_{i}\|}{\sum_{i \in \overline{\Omega}} \|\mathbf{y}_{i}\|}$$
$$S_{2}(\mathbf{y}) = \frac{\sum_{i \in \Omega} \|\mathbf{n}_{i}^{T} \mathbf{y}_{i}\|}{\sum_{i \in \overline{\Omega}} \|\mathbf{n}_{i}^{T} \mathbf{y}_{i}\|}$$

yは回帰した誘導電場で、 Ω は刺激したい部位、 $\overline{\Omega}$ は Ω 以外の部位、 n_i は大脳皮質の表面の法線を 表す. S_1 、 S_2 ともに値が大きいほど Ω の刺激に 適していることを示す.本研究では、 Ω は運動野 とした.



図 3:誘導電場の推定精度

3. 結果

ニューラルネットを用いて誘導電場の回帰 を行った.図2にその結果を大脳皮質にプロッ トしたものを示す.また,その推定精度を図3 に示す.図2より,MRI画像と導電率分布画像 どちらから回帰しても回帰が成功しているこ とが確認できる.図3より,MRI画像よりも導 電率分布画像から誘導電場を回帰したほうが, 推定精度が高いことが分かる.表1にニューラ ルネットを用いて領域分割した結果のDice係 数を示す.

ニューラルネットで推定した誘導電場の精 度がコイル位置を特定する上で十分かどうか を確認するために、学習に用いていな被験者の 運動野を局所的に刺激する位置を求めた結果 を表2に示す.運動野の位置は手動で事前に指 定した.表2より、MRI 画像と導電率分布画像 どちらから回帰しても、コイルの位置は大きく 変わらない.そのため、提案しているニューラ ルネットの推定精度が十分であることが確認 できた.

4. まとめ

ニューラルネットを用いて,誘導電場を回帰 する手法を提案した. MRI 画像から直接回帰す るよりも, MRI 画像を領域分割して導電率を割 り当て,導電率分布画像を作り,その導電率分 布画像から誘導電場を回帰する方が,推定精度 が高いことが分かった.また,最適なコイル位 置の全探索を通して,誘導電場の推定精度が十 分であることが分かった.

利益相反の有無

なし

文 献

- [1] Rossi S, Hallett M, Rossiniet P. M, et al: Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. Clinical neurophysiology, **120**(12): 2008-2039, 2009
- [2] Laakso I, Hirata A: Fast multigrid-based computation of the induced electric field for transcranial magnetic stimulation. Physics in Medicine & Biology, 57(23):7753, 2012
- [3] Gabriel S, Lau R. W, Gabriel C: The dielectric properties of biological tissues: Iii. parametric models for the dielectric spectrum of tissues. Physics in Medicine & Biology, 41(11):2271, 1996
- [4] Yokota T, Maki T, Nagata T, Murakami T, et al: Real-time estimation of electric fields induced by transcranial magnetic stimulation with deep neural networks. Brain Stimulation 12(6), 1500–1507 (2019)
- [5] Rashed E. A, Gomez-Tames J, Hirata A: Development of accurate human head models for personalized electromagnetic dosimetry using deep learning. NeuroImage, 202(116132), 2019
- [6] Fischl B: Freesurfer. NeuroImage, **62**(2), 774-781
- [7] Çan M. K, Laakso I, Nieminen J. O, et al: Coil model comparison for cerebellar transcranial magnetic stimulation. Biomedical Physics & Engineering Express, 5(1):015020, 2018
- [8] Kingma D. P, Ba J: Adam: A method for stochastic optimization. CoRR, abs/1412.6980, 2015

表 1: 領域分割した結果の表: 非記入のセルは切り出したデータセットに 含まれなかった部位である

	Dice 係数		Dice 係数		Dice 係数
脳の外	0.99	椎間板,粘膜,小		前庭	
		脳灰白質			
皮質骨	0.86	髄膜の骨	0.67	一部小脳灰白質	
海綿骨	0.89	灰白質, 視床, 尾状	0.89	一部灰白質	0.73
		核, 被殻淡蒼球, 海			
		馬,扁桃体,側坐核			
白質, 脳幹, 間脳	0.88	脂肪	0.05	眼球	0.24
頭蓋, 脂肪, 骨膜	0.95	髄膜,筋肉	0.73	脳脊髄液	0.87
白質と灰白質の平均	0.70	血液	0.55		

表 2:コイルの最適な位置

運動野		シミュレーションで 得られた誘導電場	MRI 画像から回帰	推定した導電率画像 から回帰
	<i>S</i> ₁			
	<i>S</i> ₂			

Improvement of Regression Accuracy of Induced Electric Field for TMS

by Regression of 3D conductivity Image from MR Image

Toyohiro MAKI^{*1}, Tatsuya YOKOTA^{*1}, Akimasa HIRATA^{*2}, Hidekata HONTANI^{*1}

*1 Department of Computer Science, Nagoya Institute of Technology*2 Department of Electrical and Mechanical Engineering, Nagoya Institute of Technology

Transcranial Magnetic Stimulation (TMS) is a non-invasive stimulation method that stimulate neurons in subject's brain by using a coil. TMS is used treatment of depression and identification of brain function (e. g. motor cortex). It is important for TMS to identify coil position and orientation. Thus, electric filed induced by the coil is an important criterion. In this paper, we develop two approaches by using neural network. First approach, we regress inductive electric field from conductivity image. The conductivity image is constructed by MR image. Second approach, we regress inductive electric field by MR image without construction of conductivity image. As a result, we confirmed that first approach has better accuracy than second one.

Key words: Transcranial Magnetic Stimulation, Deep Neural Network, MR Image, Electric Field Estimation

PROPELLER MRI における

CS と CNN を利用した再構成法の検討

松元 祐太*1 堀 拳輔*1 只野 喜一*2

久原 重英*1,*2 遠藤 祐太*2 橋本 雄幸*1,*2

要旨

PROPELLER MRI は、k 空間の原点を中心に回転した帯状領域(ブレード)の収集データから再構成する 方法である.この方法では、ブレード間の位相と回転を利用して被写体の動きを補正できる.一方で、k 空間のサンプリング率が低いと、再構成画像に線状のアーチファクトが発生する.本研究では、圧縮セン シング(CS)と畳み込みニューラルネットワーク(CNN)を使用して、腹部 PROPELLER MRIの画質向上 を試みた.ブレードの幅と数を変化させてサンプリング率を 11~54%とした k 空間データにおいて、CS で再構成する方法、CS を利用せずに再構成した画像を CNN で処理する方法、CS と CNN を併用する方法 の3つの手法を比較した.すべてのサンプリング率において、CS と CNN を併用する方法の評価値が最良 となった.サンプリング率が低い場合においても、CS と CNN を併用することで画質向上の可能性が示唆 された.

キーワード: 圧縮センシング, 畳み込みニューラルネットワーク, 画像再構成, MRI, PROPELLER

1. はじめに

Magnetic Resonance Imaging (MRI) 検査に おいて被写体の体動, 呼吸, 拍動, および蠕 動はモーションアーチファクトを引き起こし, 画質を低下させる. モーションアーチファク トの低減策として, Periodically Rotated Overlapping ParallEL Lines with Enhanced Reconstruction (PROPELLER) が挙げられる [1]. PROPELLER MRI は原点を中心にブレー

*1 杏林大学大学院保健学研究科保健学 専攻

〔〒181-8612 三鷹市下連雀 5-4-1〕

e-mail: hashi@ks.kyorin-u.ac.jp

*2 杏林大学保健学部診療放射線技術学 科

ドと呼ばれる帯状領域を回転させて充填した k 空間データから画像再構成する方法である [2,3]. 原点付近においてブレードが重複する ことを利用してブレード間での被写体の動き を推定し、モーションアーチファクトを補正 できる. 臨床では呼吸や蠕動運動による動き を補正するため、腹部領域の撮像に多く用い られている [45]. PROPELLER MRI では、よ り多くの位相エンコード数とブレード数を用 いて収集することで,動き補正の精度が向上 するが撮像時間が延びる.また、収集するブ レード数を減らすと撮像時間を短縮できるが, k 空間データが疎(スパース)になるため再 構成画像に線状のアーチファクトが発生する. 圧縮センシング (CS) は、k 空間の充填率 (サンプリング率)が低い場合でも、データ

のスパース性を利用することで信号を復元し, 画像再構成できる手法として知られており, MRI において有用であると報告されている [6,7].

また,畳み込みニューラルネットワーク (CNN)は深層学習モデルの1つであり,医 用画像処理において広く応用されている [8, 9]. CNN は複数のフィルタで畳み込み演算処 理を行い,特徴を抽出する畳み込み層とその 特徴をまとめるプーリング層を繰り返した構 造となっており,CNN によって肝臓 MRI に おけるモーションアーチファクトの低減が可 能であることが報告されている [9,10].

本研究では、PROPELLER MRI においてブ レードの幅を決定する位相エンコード数とブ レード数を減らしたサンプリング率が低い k 空間データから CS と CNN を用いて再構成す ることで、腹部 PROPELLER MRI の画質向上 の可能性を検討した.

2. 方法

位相エンコード数とブレード数を変化させ て取得したスパースな k 空間データから CS で再構成する方法場合, CNN で処理する方法, CS と CNN を併用する方法の3つの手法で再 構成し,画像を比較した.本研究では,k 空 間全体からフーリエ逆変換して再構成した画 像を原画像とし,サンプリング率は式(1)で 算出した.

Sampling rate [%] = $\frac{A}{(R/2)^2 \times \pi} \times 100$ (1)

ここで, A はサンプリングした面積, R はブ レードの横幅を示す.

1) CS で再構成する方法

本研究では CS の 1 つである Projections Onto Convex Sets (POCS) 法を用いた [11]. POCS 法は, Soft threshold 処理によって微小 なノイズ成分を取り除くことで k 空間のデー タを復元できる. POCS 法で用いた評価関数 を式 (2) に示す.

$$\underset{\mathbf{x}}{\arg\min} \left\| F\mathbf{x} - \mathbf{y} \right\|_{2}^{2} + \left\| \Psi \mathbf{x} \right\|_{1}$$
(2)

ここで, H₁およびH₂は L1 ノルムおよび L2 ノルムであり, x は目的画像, y は測定画像, F はフーリエ変換演算子, Y はウェーブレッ ト変換演算子を表す. さらに, アルゴリズム を以下に示す.

Initialization $\lambda = 0.001$ k = 0 $X_0 = y$ Iterations for(k < m) {

$$\boldsymbol{x}_k = \boldsymbol{\Psi} \boldsymbol{F}^{-1} \, \boldsymbol{X}_k$$

$$x_{k+1}[i] = \begin{cases} 0 & \text{if } |x_k[i]| < \lambda \\ \frac{|x_k[i]| - \lambda}{|x_k[i]|} & \text{otherwise} \end{cases}$$

$$\boldsymbol{X}_{k+1} = F \boldsymbol{\Psi}^{-1} \boldsymbol{x}_{k+1}$$

$$X_{k+1}[i] = \begin{cases} X_{k+1}[i] & \text{if } y[i] = 0\\ y[i] & \text{otherwise} \end{cases}$$

k = *k* +1

Output

$$\hat{\boldsymbol{x}} = \boldsymbol{F}^{-1} \boldsymbol{X}_m$$

ここで \hat{x} は再構成画像,*i*は画素番号, F^{-1} は フーリエ逆変換演算子, Ψ^{-1} はウェーブレット 逆変換演算子を表す.今回 Soft threshold 処理 の閾値は 0.001 とした.また繰り返し回数 *m* は 30 回以上にしても画像に大きな変化は見 られなかったため、30 回に設定した.

2) CNN で処理する方法

肝臓の MRI 画像において CNN を用いると モーションアーチファクトを低減できると報 告されている [9].本研究では,改良したネ ットワーク構造をモーションアーチファクト の低減処理に使用した(図1).入力層はカー ネルサイズが 3×3 の 64 個のフィルタからな る 2 次元の畳み込みと、活性化関数 Rectified



図1CNNの構造.入力画像はアーチファクトあり画像で、出力画像は入力画像から原画像を差し引いたア ーチファクト成分画像である.

Linear Unit (ReLU) から構成されている. 畳 み込み層はカーネルサイズが 3×3 の 64 個の フィルタからなる 2 次元の畳み込み、Batch Normalization,および活性化関数 ReLU の組 み合わせを1層としたとき,これを7層連ね て構成している.出力にはカーネルサイズが 3×3 の1 個のフィルタからなる2次元の畳み 込みを使用した.

トレーニングデータセットには、オープン アクセスデータセットである TCGA-LIHC の 1660 枚の T1 および T2 強調画像を 256×256 pixels にリサイズして使用した [12]. さらに, 回転と縮小を行うことで、データセットを 4980 枚に増強した.このとき,回転範囲は0 から 360 度, 縮小範囲は 0.6 から 1 倍に設定 した. 被写体が画像の外側にはみ出すことを 防ぐため、平行移動と拡大は行わなかった. 各画像をフーリエ変換して得られるk空間デ ータにおいて, 位相エンコード数およびブレ ード数を7から16の範囲でランダムに変化さ せ, 収集できる領域をサンプリングすること で PROPELLER 法での撮像を想定した.その k 空間データをフーリエ逆変換で再構成する ことでアーチファクトあり画像を作成し、原 画像を差し引くことでアーチファクト成分画 像を生成した(図 2). CNN の入力画像には



図2 トレーニングデータセット画像:(a) 原画像,
 (b) アーチファクトあり画像,(c) アーチファクト
 成分画像(教師画像).

アーチファクトあり画像を、教師画像として アーチファクト成分画像を使用して学習を行 った(図 1). 学習の評価関数として式(3) に示す Mean Squared Error (MSE)を用い、そ れを Adaptive moment estimation (Adam)で最 適化した.

$$MSE = \frac{1}{N} \sum_{i=0}^{N-1} (m_i - f_i)^2$$
(3)

ここで $m_i \ge f_i$ は評価画像と原画像における i番目の画素値, N は画像の画素数を示す. バ ッチサイズおよびエポック数は, それぞれ 32 ×32 pixels および 100 回とした.

3) CS と CNN を併用する方法

入力画像として 1)で示した POCS 法で k 空 間データを復元させて再構成した画像を使用 し,教師画像としてアーチファクト成分画像 を使用して2)で示したCNNの学習を行った. このとき,位相エンコード数およびブレード 数を7から16の範囲でランダムに変化させ, アーチファクト成分画像はPOCS法で再構成 した画像から原画像をサブトラクションする ことで生成した.CNNのパラメータは2)と 同じである.

4) 数値シミュレーション

評価画像には TCGA-LIHC に含まれる T2 強調画像を 2) と同様, 256×256 pixels にリ サイズして使用し, フーリエ変換することで k 空間データを得た [12]. 位相エンコード数 およびブレード数を7から16の範囲で変化さ せ, 収集できる領域をサンプリングすること で100 種類作成した. さらに, 等間隔にサン プリングした各ブレードを±4 度の範囲でラ ンダムに回転させることで, 動き補正を行っ たときを想定した.

本研究で使用した計算機環境を表1に示す. また, CNN の学習およびテストには Graphics Processing Unit (GPU) を用いた.

表1 計算機環境.

CPU	Intel Core i7-6950X (3.0GHz)
GPU	NVIDIA GeForce GTX 1080
Software	CUDA Toolkit 9.0/cuDNN 7.1.4

5) 評価指標

画像の定量的評価指標として式(4) に示す Normalized root mean squared error (NRMSE) を用いた.

NRMSE[%] =
$$\sqrt{\frac{\sum_{i=0}^{N-1} (m_i - f_i)^2}{\sum_{i=0}^{N-1} (f_i)^2}} \times 100$$
 (4)

ここで $m_i \ge f_i$ は評価画像と原画像における i番目の画素値, N は画像の画素数を示す. NRMSE は値が 0 に近いほど優れていること を意味する.

3. 結果

図3に原画像および位相エンコード数が8 でブレード数が13(サンプリング率:24.2%) のときを想定し、各手法で再構成した画像を 示した.また、上段の画像において左下に記 載されている数字は NRMSE を表しており, CS と CNN を併用した方法が最も良い結果と なった. CS で再構成しても、あまりアーチフ ァクトが低減しなかったのに対し, CNN で処 理を行うことで、ほとんどのアーチファクト を取り除くことができた.しかし、コントラ ストが大きい部分では、図3(i)の矢印に示す ようにストリーク状のアーチファクトが低減 できなかった.CS と CNN を併用することで, それらも低減することができ、視覚的にも最 もアーチファクトを低減できることがわかっ た.

図4に各サンプリング率における CNN で 処理した画像とCS と CNN を併用した画像を 示す.図5に各サンプリング率における NRMSEの値をプロットしたグラフを示す. どのサンプリング率においても,CS と CNN を併用した再構成法が最も優れた値を示した.

4. 考察

CNN のみを用いた場合,視覚的にアーチフ ァクトは低減したが,皮下脂肪など低コント ラスト領域の微小構造が消失した(図4).ま た,定量的にはサンプリング率が40%以上に おいて NRMSE の改善が緩やかになった(図 5).原因として,サンプリング率を11%から 54%と広範囲に変化させてデータセットを作 成し,一度に学習させることでサンプリング 率が低く,アーチファクトが目立つデータセ ットに学習パターンが大きく影響を受けたと 考えられる.改善策として,再構成の対象と するサンプリング率に合わせたデータセット を作成し,サンプリング率ごとに学習し直す 方法が挙げられる.

CS と CNN を併用した場合、視覚的にもコ ントラストを維持しつつ、線状のアーチファ クトが低減し、定量的にも NRMSE が最も改



図3 原画像とサンプリング率24.2%の各手法の再構成画像:(a) 原画像,(b) CS と CNN を用いていない 画像,(c) CS で再構成した画像,(d) CNN で処理した画像,(e) CS と CNN を併用した画像. 下段 (f,g,h,i,j) は上段(a,b,c,d,e)の拡大画像である. 矢印は, CS を用いない場合において高輝度なストリーク状のアー チファクトが低減できなかった部分を示す.



図4 各サンプリング率における CNN で処理した画像(上段)とCS と CNN を併用した画像(下段): (a,e) サンプリング率 19.9%, (b,f) サンプリング率 30.2%, (c,g) サンプリング率 39.9%, (d,h) サンプリング 率 51.0%.

善した.これは, CS でデータを復元して再構 成することで,サンプリング率の違いによる アーチファクトパターンの差がなくなったた めと考えられる.

今回は, 被写体の動きを全て補正できたと

仮定してサンプリング率が低い画像の再構成 法を検討した.今後は動きを補正しきれなか った場合における再構成画像の画質向上につ いての検討を進める予定である.



図5 各サンプリング率における NRMSE.

5. まとめ

サンプリング率が低い場合において, CS で再構成する方法, CS を利用せずに再構成し た画像を CNN で処理する方法, CS と CNN を併用する方法の 3 つの手法ついて比較した. CS と CNN を併用した場合,視覚的にもアー チファクトが軽減し,定量的にも数値が最良 となった.したがって, CS と CNN を併用し て再構成することで,サンプリング率の低い 腹部 PROPELLER MRIの画質を改善できるこ とが示唆された.

謝辞

本研究で使用した画像は全てオープンアク セスデータセットである TCGA-LIHC を使用 した.

利益相反の有無

なし.

文 献

- Zaitsev M, Maclaren J, Herbst M: Motion artifacts in MRI: A complex problem with many partial solutions. J Magn Reson Imaging 42: 887-901, 2015
- [2] Pipe JG: Motion correction with PROPELLER MRI: Application to head

motion and free breathing cardiac imaging. Magn Reson Med **42**: 963-969, 1999

- [3] Forbes KPN, Pipe JG, Bird CR et al: PROPELLER MRI: Clinical testing of a novel technique for quantification and compensation of head motion. J Magn Reson Imaging 14: 215-222, 2001
- [4] Deng J, Miller FH, Salem R et al: Multishot diffusion weighted PROPELLER magnetic resonance imaging of the abdomen. Invest Radiol 41: 769-775, 2006
- [5] Hirokawa Y, Isoda H, Maetani YS et al: Evaluation of motion correction effect and image quality with the periodically rotated overlapping parallel lines with enhanced reconstruction (PROPELLER) (BLADE) and parallel imaging acquisition technique in the upper abdomen. J Magn Reson Imaging 28: 957-962, 2008
- [6] Lustig M, Donoho DL, Pauly JM et al: Sparse MRI: The application of compressed sensing for rapid MR imaging. Magn Reson Med 58: 1182-1195, 2007
- [7] 町田好男: MR イメージングにおける
 圧縮センシングとその周辺技術. Med
 Imag Tech 34: 198-202 2016
- [8] 大手希望,得居葵,橋本二三生,他: 畳み込みニューラルネットワークによる PET 画像ノイズ除去. Med Imag Tech 37: 35-45,2019
- [9] Tamada D, Kromrey ML, Ichikawa S et al: Motion artifact reduction using a convolutional neural network for dynamic contrast enhanced MR imaging of the liver. Magn Reson Med Sci 19: 64-76, 2020
- [10] Kromrey ML, Tmada D, Johno H et al: Reduction of respiratory motion artifacts in gadoxetate-enhanced MR with a deep learning-based filter using convolutional neural network. Eur Radiol **30**: 5923-5932, 2020
- [11] Shah J, Qureshi I, Omer H et al: A

modified POCS-Based Reconstruction method for Compressively sampled MR imaging. Int J Imaging Syst Technol **24**: 203-207, 2014

[12] Erickson BJ, Kirk S, Lee Y et al.: Radiology Data from The Cancer Genome Atlas Liver Hepatocellular Carcinoma [TCGA-LIHC] collection. The Cancer Imaging Archive, 2016

Examination of the reconstruction method using CS and CNN for

PROPELLER MRI

Yuta MATSUMOTO^{*1}, Kensuke HORI^{*1} Kiichi TADANO^{*1}, Shigehide KUHARA^{*1} Yuta ENDO^{*1}, Takeyuki HASHIMOTO^{*1}

*1 University of Kyorin

PROPELLER MRI is a method of reconstructing images from k-space data acquired in a band (blade) rotated around the origin in k-space. This method compensates for the subject's motion by using the phase and rotation between the blades. However, if the sampling rate is low, high intensity linear artifacts will appear in the reconstructed image. In this study, we attempted to improve the image quality for PROPELLER MRI in the abdomen using Compressed Sensing (CS) and Convolutional Neural Network (CNN). For the k-space data with sampling rates of 11%–54%, which varied in width and the number of blades, we compared three methods: (1) reconstructing the image via CS; (2) processing the image via CNN after reconstruction using inverse Fourier transform; and (3) processing the image via CNN after reconstruction using rates, combining CS and CNN achieved the best NRMSE, which may improve the image quality even when the sampling rate is low.

Key words: Compressed sensing, Convolutional neural network, Image reconstruction, MRI, PROPELLER

著者紹介

和文でも英文でも可。



松元 祐太 (まつもと ゆうた) 2020 年杏林大学保健学部放射線技術学科 卒業.現在,同大大学院保健学研究科博士 前期課程に在学中.MRI における深層学習 の研究に従事.日本医用画像工学会会員.





2017年北里大学医療衛生学部卒業.2019 年北里大学大学院医療系研究科修士課程

堀 拳輔 (ほり けんすけ)



只野 喜一(ただの きいち) 2009 年北里大学医療衛生学部卒業.2011 年筑波大学大学院人間総合科学研究科修 士課程修了.2012 年株式会社バリアンメデ ィカルシステムズ,2015 年杏林大学保健 学部診療放射線技術学科助教.MR spectroscopy, Diffusion Tensor Imaging の研究に従事.



久原 重英 (くはら しげひで) 1983 年九州大学大学院修士課程工学研究 科情報工学専攻修了. 工学博士. 東芝メデ ィカルシステムズ (株) MRI 事業部等にて 超高速 MRI システムの開発等に従事. 現 在, 杏林大学保健学部診療放射線技術学科 教授. MRI のデータ収集ならびに画像再構 成法等に関する研究に従事. ISMRM, 日本 磁気共鳴医学会, 各会員.



速藤 祐太 (えんどう ゆうた) 2017 年杏林大学保健学部診療放射線技術 学科卒業. 2019 年同大大学院保健学研究科 保健学専攻博士前期課程修了.現在,杏林 大学保健学部診療放射線技術学科助教.診 療放射線技師.循環器領域を中心とした MRI 画像による生体定量化技術に関する研 究に従事.



橋本 雄幸 (はしもと たけゆき) 1994 年筑波大学工学研究科博士課程修 了.工学博士.1994 年横浜創英短期大学情 報処理学科講師,2012 年横浜創英大学こど も教育学部教授,2016 年杏林大学保健学部 診療放射線技術学科教授.ラドン変換の逆 問題,画像処理,非破壊検査の応用研究に 従事.

AD-MIL を用いた頭部 MR 画像における

回復期脳梗塞患者の帰結予測

東 優大*1 寺本 篤司*1 岡崎 英人*2

武田 湖太郎*3 前田 繁信*4

要旨

現在, 脳梗塞は要介護状態に至る疾患のなかで主要な疾患の1つである. 要介護状態の患者が日常動作の 改善や在宅復帰を目指すに際して, リハビリテーションが重要となる. その際に各々の患者に適したリハ ビリテーション計画を立案するには, 正確な予後予測が必要不可欠である. そのため本研究では, 回復期 脳梗塞患者の MR 画像を CNN に入力し, 弱教師あり学習手法である Attention-based deep multiple instance learning (AD-MIL)を用いて帰結良悪の予測を行った. まず, リハビリテーションを実施する前に撮影した MR 画像を入力し, CNN モデルの1つである VGG-16 の全結合層から 1024 個の特徴量を抽出した. そして, アテンション機構に基づいて, その特徴量より重み付けされた特徴ベクトルを算出し, 症例ごとに帰結良 好および不良の予測を行った. 結果は, 帰結良好の予測精度が 86.7%, 帰結不良では 61.5%となり, 総合 正解率が 75.0%であった. 以上の結果より, 本手法が回復期脳梗塞患者の帰結予測に有効であることが示 唆された.

キーワード:ディープラーニング,脳梗塞,MR 画像,リハビリテーション,帰結予測

1. はじめに

国内の要介護状態に至った主な原因におい て,脳血管疾患は認知症に次いで2位であ る.さらに,介護度が高い患者ほど脳血管疾 患が原因である場合が多いため,発症した患 者は一命を取り留めたとしても重度後遺症が 残る可能性が高い[1].特に脳梗塞は脳血管疾

- *1 藤田医科大学大学院保健学研究科〔〒 470-1192 豊明市沓掛町田楽ヶ窪1番地
- 98] e-mail: teramoto@fujita-hu.ac.jp
- *2 藤田医科大学医学部連携リハビリテーション医学
- *3 藤田医科大学保健衛生学部リハビリテ ーション学科
- *4 藤田医科大学七栗記念病院

患の大半を占めており、運動麻痺や感覚障害 など多岐にわたる後遺症をもたらす疾患であ る.そのため、患者が脳梗塞発症前の生活を 獲得するうえでリハビリテーション(以下リ ハビリ)が重要となる.

リハビリは実施時期やその目的によって急 性期リハビリ,回復期リハビリ,維持期リハビ リの3つに分けられる.その中でも回復期リハ ビリは,専用施設で発症後1~6カ月の時期に 日常生活活動の改善や在宅復帰を目指して行 われる.その際必要となるリハビリ実施計画は 患者の状態に応じて立案されているが,より具 体的な計画立案には正確な予後予測が必要不 可欠だと考えられてきた.そのため,予後予測 に関する研究が多数存在するが,その多くは患 者の基本情報や機能評価指標などを入力とし て,人工ニューラルネットワークや重回帰分析 を用いることで予測を行うものである[2][3]. ここで後遺症が残る場所やその重症度は梗塞 部位に大きく依存しているため,脳梗塞の画像 診断で有用性が高い MR 画像は,予後予測にお いても有意的な情報だと考えられる.それにも かかわらず, MR 画像を入力としたリハビリ領 域における予後予測に関する研究は,我々が知 る限りでは行われていない.

現在, コンピュータ支援診断技術 (CAD) に 関する研究において,人工知能技術の医用画像 への応用が盛んに行われている.その人工知能 技術の一つに Multiple instance learning (MIL)が ある[4]. MIL は bags・instance という概念を用 いて,学習および検証を行う弱教師あり学習手 法である.そして,その派生技術に Deep learning Attention 機構を導入した Attentionbased deep multiple instance learning (AD-MIL)が ある[5].本研究では,この AD-MIL を用いた 頭部 MR 画像における回復期脳梗塞患者の帰結 予測を目的とした.

2. 手法

本手法では、図1に示すようなリハビリ実施 前の頭部 MR 画像を入力として、AD-MIL を用 いて帰結良好・帰結不良の予測を行った.まず、 リハビリ領域において機能評価指標として利 用されている Functional independence measure (FIM)を用いて、帰結良好および帰結不良の 選定を行った.次に MR 画像を Convolutional neural network (CNN)モデルの一つである VGG-16[6]に入力し、その全結合層から画像 1 枚あ たり 1024 個の特徴量を抽出した.そして、AD-MIL が有するアテンション機構に基づいて、そ の特徴量から重み付けされた特徴ベクトルを 算出し、症例ごとに帰結良好および不良の予測 を行った.



図1 リハビリ実施前の MR 画像 (FLAIR)

3. 結果および考察

藤田医科大学七栗記念病院で収集された脳 梗塞患者 56 症例に対して予測を行った.検証 方法として 10 分割交差検証法を採用し,帰結 良好・不良それぞれの予測精度およびその総合 予測精度を算出した.

その結果, 帰結良好の予測精度が 86.7%, 帰結不良では 61.5%となり, 総合正解率が 75.0%であった.以上の結果より,本手法が回 復期脳梗塞患者の帰結予測に有効であること が示唆された.

利益相反の有無

なし.

文 献

- [1] 2019年国民生活基礎調査の概況: <u>https://www.mhlw.go.jp/toukei/saik</u> <u>in/hw/k-tyosa/k-tyosa19/dl/14.pdf</u> (2021/06/07)
- Fujita T, Sato A, Narita A et al: Use of a multilayer perceptron to create a prediction model for dressing independence in a small sample at a single facility. J Phys Ther Sci 31: 69–74, 2019
- [3] Jeong S, Inoue Y, Kondo K et al: Formula for predicting FIM for stroke patients at discharge from an acute ward or convalescent rehabilitation ward. Jpn J Compr Rehabil Sci 5: 19-25, 2014
- [4] Dietterich T.G., Lathrop R.H., Lozano-Pérez T: Solving the multiple instance problem with axis-parallel rectangles. Artificial Intelligence 89: 31-71, 1997
- [5] Ilse M, Tomczak J, Welling M: Attentionbased Deep Multiple Instance Learning. Proc Mach Learn Res 80: 2127-2136, 2018
- [6] Simonyan K, Zisserman A: Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556, 2014

Predicting Outcomes of Patients with Cerebral Infarction in the Recovery Stage Using Attention-Based

Deep Multiple Instance Learning in MR image.

Yudai HIGASHI^{*1}, Atsushi TERAMOTO^{*1}, Hideto OKAZAKI^{*2} Kotaro TAKEDA^{*3}, Shigenobu MAEDA^{*4}

*1 Graduate School of Health Sciences, Fujita Health University

*2 Department of Cooperation Rehabilitation Medicine, School of Medicine, Fujita Health University

*3 Faculty of Rehabilitation, School of Health Sciences, Fujita Health University

*4 Fujita Health University Nanakuri Memorial Hospital

The cerebral infarction is one of the major diseases that lead to the need for nursing care. A rehabilitation is important for patients who require nursing care to improve their daily activities and be discharged from hospitals. Accurate prediction of functional outcome is essential to develop a rehabilitation implementation plan that is appropriate for each patient. Therefore, in this study, we aimed to predict good and poor functional outcome using Attention-based deep multiple instance learning (AD-MIL), a weakly supervised learning method, by inputting MR images of patients with cerebral infarction in the recovery stage into convolutional neural network (CNN). First, using MR images acquired before rehabilitation,1024 features were extracted from all the combined layers of VGG-16, one of the CNN models. The weighted feature vectors were then calculated from the features based on the attention mechanism to predict good and poor outcome for each case. As a result, the prediction accuracy for good and poor outcome was 86.7% and 61.5%, respectively. The overall correct prediction rate was 75.0%. These results indicate that the proposed scheme may be useful to predict the outcome of patients with cerebral infarction in the recovery stage.

Key words: Deep learning, cerebral infarction, MR image, rehabilitation, prediction of outcome

Cine-MRI における小腸蠕動運動の

オプティカルフローによる定量評価および腹膜炎の鑑別

川原 稔暉*1 井上 明星*2*4 岩本 祐太郎*1 古川 顕*3

Bolorkhand Batsaikhan^{*3} 茶谷祥平^{*4} 陳 延偉^{*1}

要旨

消化管は食物の運搬,吸収,消化機能を担う臓器である.内容物を口側から肛門側に運搬する役割を担う 消化管の蠕動運動は,食事内容,薬剤,閉塞ならびに腹部の炎症など様々な要因により変化しうる.Cine-MRI は小腸全体の蠕動運動を俯瞰的に描出することが可能で臨床現場で用いられているが,主観的に評価 されており,簡便で客観性・再現性のある解析方法が期待される.先行研究では深層学習を用いて局所領域 の小腸領域をセグメンテーションし,その面積変動を評価していたが,広範囲の小腸領域の動きを評価で きていない.また,小腸の動きを数値的に評価している研究はあるものの,医師が事前に ROI を指定した 範囲だけでしか評価できず,小腸の全体的な運動を可視化する研究は行われていない.本研究では Cine-MRI に対してオプティカルフローを計算し,広範囲の腸の運動を可視化し,その運動量から腹膜炎の鑑別 を行った.

キーワード:深層学習,腹膜炎,オプティカルフロー,可視化,診断支援

1. はじめに

消化管蠕動運動は食事内容,薬剤,腸管閉塞, 炎症性疾患による影響を受けることから,この 動態変化から病態をとらえる試み[1-3]がなさ れている.本邦で広く行われている内視鏡検査 は,内腔の観察による形態変化に基づき診断が 行われるが,小腸蠕動運動のような動態変化の

*1 立命館大学大学院情報理工学研究科

*2 東近江総合医療センター

- 〔〒527-8505 滋賀県東近江市五智町 255〕
- *3 東京都立大学
- 〔〒192-0397 東京都八王子市南大沢 1-1〕*4 滋賀医科大学放射線科
- 〔〒520-2192 滋賀県大津市瀬田月輪町〕

評価には不向きである. X 線を用いた消化管造 影は小腸全体の蠕動運動を観察する検査法で あるが,放射線被曝を伴うことが欠点として挙 げられる.一方, Cine magnetic resonance imaging (Cine-MRI) は放射線を用いず,俯瞰的に小腸 蠕動運動を描出できるモダリティであるが,画

□ 場動運動を描出でさるモタリティであるか、画像評価の客観性・再現性の担保および評価に費やす時間的負担が課題である.

上記課題解決へのアプローチとして, Cine-MRIによる小腸蠕動運動の解析が若宮らにより 提案されている[4].この手法では, Cine-MRI から小腸蠕動運動の腸管口径を手動で計測し, その周波数を解析することにより,小腸機能の 異常の有無を低侵襲に診断することを目的に 提案された手法である. Cine-MRIでの検査に より,患者の検査は低侵襲になり患者の負担軽

^{〔〒525-8577} 滋賀県草津市野路東 1-1-1〕



減に貢献したが,手動による計測のため,解析 に膨大な時間を要し,医師の負担が非常に大き く,診断効率が悪い.そのため,先行研究では super-pixel法[5]や深層学習(U-Net)[6,7]を用い て医師が指定した局所領域(ROI)内の小腸領 域を自動セグメンテーションし,小腸領域の面 積の時間変化から蠕動運動の自動計測と定量 評価を行なってきた[8,9].しかし,先行研究で は小腸全体の蠕動運動を把握することは困難 であった.

そのため、本研究では、小腸領域全体を対象 として、物体追跡手法であるオプティカルフロ ーを計算することで、より広範囲の腸の運動を 定量評価する手法を提案する.この手法では、 必要に応じて既存研究と同様に、注目領域の運 動も評価することができる.そして、臨床応用 を目指し、腸の蠕動運動の解析結果を用いて小 腸疾患の代表的な症例の一つである腹膜炎の 鑑別法について検討する.

2. 提案法

2.1. 提案法の概要

本研究では、広範囲の腸の運動を解析する手 法と蠕動運動の定量評価方法、腹膜炎の鑑別法 について提案する.提案法の全体図を図1に示 す. 本研究で利用する Cine MRI は,前処理とし て,小腸領域以外を除外する目的で,画像診断 医の指導のもと,左右の中心は画像の中心,上 下の中心は上前腸骨棘とし,その交点を中心と した画像サイズ 128(pixel)×128(pixel)に切り出 しを行った.その後,連続するフレーム間にお いて Farneback 法[10]を用いてオプティカルフ ローを計算し,小腸の蠕動運動を可視化した. 既存方法[8,9]では,ROI の大きさが 16(pixel)× 16(pixel)と局所領域の解析しかできなかったが, 本手法では,小腸領域全体の運動を可視化し, 蠕動運動の解析が可能となった(図 1③).

蠕動運動の定量評価では、従来手法と同様に 小腸の局所領域を切り出し、その領域のオプテ ィカルフローの値から蠕動運動の定量評価を 行う(図 1⑥). 膨張・収縮運動を評価するため、 ROI の中心を原点として、オプティカルフロー が中心に向いているのか外側に向いているの かを計測し、その合計量から蠕動運動を客観的 に評価した.

最後に腹膜炎の鑑別法について検討した.提 案法では Long Short-Term Memory (LSTM)[11]を 用いて蠕動運動量の時系列データから腹膜炎 患者と非腹膜炎患者の鑑別を行った(図1⑦).



2.2. オプティカルフロー

オプティカルフローは、動画中の二枚の時間 的に隣接したフレーム間から、対象物体の移動 やエッジの変化などによる移動前の画素とそ の画素が移動した後の画素の対応ペア間の差 分ベクトルが形成するベクトル場である.本研 究では、連続する2フレーム間の近傍画素を2 次多項式で近似することで濃度勾配を頑健に 求めることができ、OpenCV[12]でも実装されて いる Farneback 法を用いてオプティカルフロー を計算した.

2.3. 蠕動運動成分の算出

本研究では、ROI の中心を原点として同心円 状に膨張や収縮を繰り返すと仮定する. 膨張す る方向(ROI の中心 O から外側に向かう向き)を 正の方向、収縮する方向(ROI の外側から中心 O に向かう向き)を負の方向とし、各オプティカ ルフローがどちらの向きを向いているか計測 する. オプティカルフロー計算時点では、各画 素のオプティカルフローは様々な方向を向い ているため、図 2 に示すように、原点中心 O か ら各オプティカルフロー始点 A を通る直線l上 に射影する. 射影されたオプティカルフローの 大きさ[f]は式(1)で表される.

$$\left| \mathbf{f}' \right| = \mathbf{f} \cdot \mathbf{e} \tag{1}$$

ここで, e は直線lの単位ベクトルである. そし

て, ROI 内の各画素の射影されたオプティカル フローの値を合計し, 各フレームの蠕動運動量 を算出した.

2.4. 腹膜炎の鑑別法

腹膜炎患者の小腸領域の評価において蠕動 運動が緩やかになるという医学的知見を基に, 本研究では Cine-MRI を用いて腹膜炎の鑑別法 を検討する.提案手法では,図 1⑥の蠕動運動 量の時系列データを LSTM に入力し,腹膜炎の 鑑別を行った.

本研究で用いた LSTM は, 再帰型ニューラル ネットワークである RNN(Recurrent Neural Network)の 1 つで, 音声認識や言語処理など, 主に時系列データなどの連続的なデータの処 理に強みを持つモデルである[11].

3. 実験結果

本研究に用いた Cine-MRI は東近江総合医療 センターの倫理委員会から承認(承認番号:30-8)された腹部 Cine-MRI データセットを用いた. Cine-MRI の撮像には 1.5-T MRI 装置(Magnetom Aera, Siemens Helthineers, Forchheim, Germany)お よびフェーズドアレイコイル(1.5-T Tim Coil, Siemens Healthineers)を使用した. 被験者が伏臥 位の状態で、位置決め画像から小腸が最大限に 描出される任意の冠状断を決定し, steady-state coherent sequence(True-FISP: TR, 4.26 ms; TE, 2.13 ms; flip angle, 60°; slice thickness, 5 mm; matrix, 320×320 ; the field of view, $400 \text{ mm}) \ge 0.5$ 秒ごとに1フレーム, 35 秒間取得した(計70フ レーム). 同一断面に対して複数の画像を連続 取得することで、小腸蠕動運動を動画のように 描出することが可能である.本研究では,腹膜 炎症例 6 症例, 非腹膜炎症例 29 症例を解析し た. 患者の呼吸等の動きを排除したため, 症例 ごとにフレーム数が異なる.



図 3 MRI に対してオプティカルフローをマッピング. (a), (b): 非腹膜炎患者;(c)腹膜炎患者



3.1. オプティカルフロー可視化結果

本提案手法で生成されたオプティカルフロ ーを MRI に重ねて描写した画像を図 3 に示す. 図 3(a), (b) に示すように,小腸領域で蠕動運動 領域が緑の矢印で表示されていることが確認 できる.また,図 3(c)の症例は蠕動運動量が少 なく腹膜炎の典型的な症例である.オプティカ ルフローを用いることで広範囲の腸の運動を 可視化し,定量的に解析することができた.

3.2. 蠕動運動の定量評価結果

本研究では既存研究[9]との比較を行い, ROI

領域内における腸の蠕動運動の定量的評価を 行った.既存研究では,MRIのROIデータに対 して 3D U-Net を用いてセグメンテーションを 行い,フレーム間での面積変化を計算すること で腸の蠕動運動の定量評価を行っていた[9].図 4 に手動で小腸領域を抽出した面積変動 (Ground Truth,GT)と3D-UNetで計算されたフレ ーム変化による面積変動,図5に同ROI領域の 2.3 節で計算したオプティカルフローの算出結 果を示す.図4(a)と図5(a)ではGT(青線)と既存 研究[9]の3DU-Netの波形(橙),我々の提案手法 であるオプティカルフローを用いた波形(橙)の 波形がおおむねー致していることが確認でき る.一方,図4(b)と図5(b)ではGTの波形と一致 しなかった例を示している.

GT の面積変化の波形に対して,既存研究で ある 3D-UNet の面積変化の波形と,提案手法で あるオプティカルフローの解析結果がどれだ け再現されているかを確認するために正規化 相互相関による評価を行った.

正規化相互相関は,2つの時間波形の類似度 を図る指標である. 相関値が 1.0 に近くなる 程,解析結果がより GT に近い波形であること が示唆される. 正規化相互相関を式(2) に示す.

$$Correl(X,Y) = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2\sum(y-\bar{y})^2}}$$
(2)

既存研究の結果である GT の面積変動と 3D-UNet の面積変動の正規化相互相関係数の平均 値が 0.40, 提案法である GT の面積変動とオプ



図 7:時刻 t=1~7におけるオプティカルフロー(上列)非腹膜炎の腸,(下列)腹膜炎の腸



図6提案手法とGTの相関係数が既存研究よりも相関係数が低かった例(a)と既存研究よりも相関係数が低かった例(b)

ティカルフローによる蠕動運動との正規化相 互相関係数の平均値が 0.33 という結果であっ た.平均値で比較すると面積変化同士を比較し ている既存研究の方が精度は高い結果となっ た.

提案法の正規化相関値が低く既存研究の結 果の方が高かった症例の特徴としては, ROI の 中で腸が平行移動するような症例(図 6(a))が多 く,対象とする小腸が ROI の中心と一致してい ないときに GT の波形と異なることが多かった. 一方,提案手法の方が,既存手法より相関係数 が良かった例としては,ROI 中に解析したい腸 だけでなく,連動して動く周りの腸が映る症例 が多かった(図 6(b)).既存手法の面積変動で計 測した場合,複数の腸が映りこむことにより, 解析したい腸以外のセグメンテーション結果 が不安定になり,蠕動運動を正確に評価できて いないことが原因であると考えられる.

3.3. 腹膜炎の鑑別結果

腹膜炎の鑑別のために、本研究ではプログラ



[腹膜炎, 非腹膜炎] 図 8 本研究で用いたネットワーク

表1 開発環境

CPU	Core (TM) i7-5960X CPU
	@ 3.00GHz×16
GPU	GeForceGTX TITAN X 12GB
RAM	64GB(8GB*8)
OS	Ubuntu 16.04 LTS

ミング言語に python3.6 を利用し, オプティカ ルフローの生成には OpenCV [12]を, LSTM によ る分類では深層学習フレームワークとして, Tensorflow[13], Keras[14]を用いて開発を行った. 開発環境は表1に示す.

本実験に用いるデータセットには腹膜炎 6 症 例,非腹膜炎 29 症例が含まれ,画像診断医の指 導のもと各症例あたり 1~5 個の ROI を手動で 抽出し,合計 133 個の ROI を用いて実験を行っ た.本実験では,各 ROI に対して 2.3 節で計算 したオプティカルフローによる蠕動運動量の 時系列データを LSTM に入力した.ここで, LSTM への入力は 20 次元(20 フレームの蠕動運 動量)である.

本研究で用いたモデルを図8に示す.入力さ

	真が 腹膜炎	真が 非腹膜炎
予測が 腹膜炎	5	10
予測が 非腹膜炎	1	19

表2 オプティカルフローの波形による 分類結果(症例数)

れた 20 次元ベクトルに対し、128 次元で出力さ れる LSTM で特徴抽出を行う. その出力を全結 合層に入力し、1次元ベクトルを得る、最後に出 力層の活性化関数 sigmoid で2 値分類結果(腹膜 炎 or 非腹膜炎)を得る. また, 学習の安定性の 向上のため全結合層に dropout(50%)を用いてい る. 本実験のロス関数にはバイナリークロスエ ントロピーを,最適化アルゴリズムには Adam[15]を設定し, Leave One Out 法(34 症例を 学習,1症例をテスト)で学習とテストを行った. 実験では、1 症例あたり複数箇所の ROI を持つ ため、1 症例あたり複数の判定結果が得られる が,腹膜炎の初期段階では炎症が部分的に現れ, 悪化すると、時間経過により全体的に広がる性 質を考慮し、本実験では、複数個の ROI の内 1 つでも腹膜炎の判定結果であればその症例は 腹膜炎であると判定した.

LSTM による分類結果を表 2 に示す. Accuracy((TP+TN)/(TP+TN+FP+FN))が 0.69, Recall が 0.83 という結果を得ることができた. 本実験結果より, 非腹膜炎であるが腹膜炎と判 定された症例には, 画像中の移動はあるものの, 同心円状の運動ではなく, 画像中の小腸が平行 移動する特徴があった(図 6(a)). 本研究では

ROIを中心とした同心円状の蠕動運動を前提 としているため、こうした特徴の症例は ROI を中心とした膨張・収縮運動が方向のオプティ カルフローがうまく検出できず、結果として誤 検出した.

4. おわりに

本研究では、画像診断医が手動で行っていた 局所領域の蠕動運動解析を、オプティカルフロ ーを用いることで小腸領域全体を定量的に解 析する手法を提案した.また、オプティカルフ ローによる蠕動運動量から腹膜炎を鑑別する 手法の検討を行った.今後の展望として、現在 医師が手動で行っている ROI の決定を自動化 することや、蠕動運動をロバストに評価可能な オプティカルフローの正規化方法などの開発 が挙げられる.

謝辞

本研究の一部は JSPS 科研費 20K21821 の助成 を受けたものです.

利益相反の有無

「なし」

文 献

- [1] Yasemin A, Mehmet B, Omer A. Eur J Radiol 2020. Assessment of the diagnostic efficacy of abdominal ultrasonography and cine magnetic resonance imaging in detecting abdominal adhesions: A double-blind research study PMID: 32145598
- Wnorowski AM, Guglielmo FF, Mitchell DG.
 J Magn Reason Imaging 2015. How to perform and interpret cine MR enterography.
 PMID: 26179244
- Inoue A, Furukawa A, Yamamoto H, et al. Acceleration of small bowel motility after oral administration of dai-kenchu-to (TJ-100) assessed by cine magnetic resonance imaging PLoS One 2018 Jan 10;13(1):e0191044. doi: 10.1371/journal.pone.0191044. eCollection 2018.
- [4] Wakamiya M, Furukawa A, Kanasaki S, et al.: "Assessment of Small bowel Motility Function with Cine-MRI Using balanced Steady-State Free Precession Sequence", JOURNAL OF MAGNETICRESONANCE IMAGING 33:1235-1240,2011.
- [5] Achanta, R, et al. "SLIC superpixels compared to state-of-the-art superpixel methods." IEEE Transactions on Pattern

Analysis & Machine Intelligence Vol.34, pp.2274-2282, 2012.

- [6] Ronneberger O, Fischer P, and Brox T, "U-Net: convolutional networks for biomedical image segmentation.", MICCAI2015, pp.243-241,2015.
- [7] Cicek O, Abdulkadir A, Lienkamp S, et al.
 "3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation", MICCAI2016, pp424-432, 2016
- [8] Nguyen D, Furukawa A, Taniguchi A, et. al. "Computerized assessment of small bowel motility function using cine-MR imaging: preliminary results in Super-Pixel Segmental Method," Journal of Transportation Medicine, Vol.71, pp.88-96, 2017.
- [9] Otsuki K, Iwamoto Y, Chen Y-W, et. al.
 "Cine-MR Image Segmentation for Assessment of Small Bowel Motility Function Using 3D U-Net," Journal of Image and Graphics, Vol.7, No.
 4, pp. 134-139, 2019. doi: 10.18178/joig.7.4.134-139

- [10] Farneback G. "Two-Frame Motion Estimation Based on Polynomial Expansion." In Proceedings of the 13th Scandinavian Conference on Image Analysis, 363 - 370. Halmstad, Sweden: SCIA, 2003.
- [11] Hochreiter, Sepp, and Jürgen Schmidhuber.
 "Long short-term memory." Neural computation 9.8 (1997): 1735-1780.
- [12] OpenCV, <u>https://opencv.org/</u>, アクセス日: 2021/6/17.
- [13] Abadi M, Agarwal A, Barham P, et al. "Tensorflow: Large-scale machine learning on heterogeneous distributed systems." arXiv preprint arXiv:1603.04467 (2016).
- [14] Keras, <u>https://keras.io/</u>, アクセス日: 2021/6/17.
- [15] Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980 (2014).

Cine-MRI Based Assessment of Small Bowel Motility and

Differentiation of Peritonitis Using Optical Flow

Toshiki KAWAHARA^{*1}, Akitoshi INOUE^{*2*4}, Yutaro IWAMOTO^{*1}, Akira FURUKAWA^{*3}, Bolorkhand BATSAIKHAN^{*3}, Shohei CHATANI^{*4}, Yen-Wei Chen^{*1}

- *1 Graduate School of Information Science and Engineering, Ritsumeikan University
- *2 National Hospital Organization Higashi-Ohmi General Medical Center
- *3 Tokyo Metropolitan University
- *4 Shiga University of Medical Science

The alimentary tract plays important role in the transport, absorption, and digestion of food. Bowel peristalsis that conveys contents from oral to the anal side is accelerated or decelerated by meals, drugs, bowel obstruction, and inflammatory disease. Cine-MRI provides overall small bowel modality and is used in clinical practice. However, cine-MRI is subjectively evaluated and it is expected that straightforward, objective, and reproducible analysis. The existing method, which evaluates the intestine by area variation using segmentation with deep learning, crops a local region of the intestine and evaluates the intestine, but does not evaluate the global range of intestinal movements. In this study, we evaluate global range of small bowel motility by calculating optical flow on Cine-MRI to solve these problems and also to differentiate peritonitis.

Key words: deep learning, peritonitis, optical flow, visualization, computer-aided diagnosis

著者紹介

和文でも英文でも可.



川原 稔暉 (かわはら としき) 2020 年立命館大学・情報理工学群・メデ ィア情報学科卒.現在,立命館大学大学 院・情報理工学研究科・情報理工学専 攻・博士前期課程人間情報科学コースに 在籍.



攻・博士前期課程人間情報科学コースに 在籍. **井上 明星**(いのうえ あきとし) 2007 年滋賀医科大学医学部医学科卒.2017 年滋賀医科大学・大学院博士課程・生体 情報解析系・高度専門医養成コース修了 後,東近江総合医療センター・放射線科・

医員を経て、Mayo Clinic, research fellow と

して在籍.







岩本 祐太郎 (いわもと ゆうたろう) 2017 年 3 月立命館大学情報理工学研究科 博士後期課程修了.工学博士.現在立命館 大学情報理工学部助教.

古川 顕 (ふるかわ あきら) 1984年3月滋賀医科大学医学部卒業。滋 賀医科大学放射線科助手,ロングアイラ ンドジューイッシュメディカルセンター 放射線科研究員,滋賀医科大学医学部放 射線講座准教授を経て,現在,東京都立 大学健康福祉学部放射線

Bolorkhand Batsaikhan

I graduated Health Science University of Mongolia in 2007-2013. Then studied Mongolian National University of Medical Sciences in 2013-2015 for Radiologist. I am working in Breast Care Center in Mongolia. Finally I joined in Tokyo Metropolitan University in 2020.



茶谷 祥平(ちゃたに しょうへい) 2014 年滋賀医科大学医学部医学科卒現在, 愛知県がんセンター病院・放射線診断・ IVR 部,および滋賀医科大学・大学博士課 程・先端医学研究者コースに在籍.

陳 延偉



2000年3月大阪大学工学研究科博士後 期課程修了。工学博士。(財)レーザー技 術研究所研究員,琉球大学工学部講 師,助教授,教授を経て,現在立命館 大学情報理工学部教授。

2方向フーリエ変換を用いた

リングアーチファクト低減フィルターの特性評価

米山明男*1 馬場理香 1,2 河本正秀 1

要旨

2方向フーリエ変換を用いたリングアーチファクト低減フィルター(以下 2Dフィルター)は、縦方向の ハイパスフィルターと横方向のローパスフィルターの組み合わせにより、サイノグラム上で縦ストライプ となって現れるリングアーチファクトを低減するフィルターである。単純なフーリエ変換で構成されてい るため、パラメータの設定が単純で、高速かつ汎用性が高いことが特徴である。これまでに放射光マイク ロ CT で計測した種子、木材、及び微化石の各データに適用した結果、CT 値を保持したままリングアーチ ファクトを大幅に低減可能なことがわかった。一方、CT 値が大きく変化する領域で、円周方向に画像が滲 むことが判明した。本稿では、円周方向の滲みについて、2Dフィルターのパラメータとの関連性を評価し、 最適な値を検討した結果について報告する。

キーワード: X線CT、リングアーチファクト、フーリエ変換

1. はじめに

放射光マイクロ CT は、放射光の大強度、 平行性、及び単色性を利用して、ミクロン以 下の高い空間分解能でサンプルの高精細な三 次元像を非破壊に取得する撮像法であり、バ イオメディカルから材料開発に至る幅広い分 野で利用されている。計測には高い空間分解 能を有した X 線画像検出器が不可欠で、一般 には入射 X 線を可視光に変換する蛍光体、可 視光用の顕微レンズ系、及び可視光用の高精 細カメラ (CCD 或いは sCMOS カメラ)から 構成された間接型 X 線画像検出器が利用され ている[1, 2]。蛍光体として、CsI や GAGG

*1 九州シンクロトロン光研究センター [〒841-0005 鳥栖市弥生ヶ丘 8-7] e-mail: yoneyama@saga-ls.jp *2 ㈱日立製作所研究開発グループ 投稿受付: 2003 年 1 月 31 日 (Gd₃Al₂Ga₃O₁₂: Ce) などの単結晶が利用され ているが、結晶欠陥が存在すると入射 X 線強 度と出力された可視光強度が比例関係から外 れ、背景画像の除算処理だけでは正常に規格 化できず、再構成像にリングアーチファクト が現れてしまう。

これまでに本リングアーチファクトを低減 するフィルターとして、サイノグラム上の縦 方向のハイパスフィルターと横方向のローパ スフィルターを組合せたフィルター(2 方向 フーリエ変換リングアーチファクト低減フィ ルター、以下 2D フィルター)を新たに開発 し、各種の CT データに適用した結果、リン グアーチファクトを大幅に低減することがで きた[3]。一方で、CT 値が大きく変化する領 域で、円周方向に画像が滲む(尾を引く)こ とが判明した。そこで、本研究では像の滲み と 2D フィルターのパラメータとの関連性を 評価し、最適なパラメータ値を検討した。



図1 2Dフィルターの構成。フィルタリングしたサイノグラムに対して、縦と横の周波数フィルターを適用した後に加算する。

2. 2D フィルターの構成

2D フィルターは、従来のフィルタード・バ ックプロジェクション法において、 Shepp-Logan 等のフィルターを適用した後の サイノグラムに、2 種類の周波数フィルター を別々に適用し、最後に再度加算する構成に なっている(図1)。2 種類のフィルターは、 縦方向のハイパスフィルターと、横方向のロ ーパスフィルターからなり、周波数空間にお いて各々個別に、

$$f_1(x) = a_1 \exp(-x/a_2)$$
(1)

 $f_2(x) = 1 - b_1 \exp(-x/b_2)$ (2)

に基づいてフィルタリングを行う。この際、

各フィルターのパラメータを等しくすること で $(a_1 = b_1, a_2 = b_2)$ 、再構成像の CT 値への影 響を抑えることができる。

3. 2D フィルターの課題とパラメータの最 適化結果

図2に九州シンクロトロン光研究センター (SAGA LS)のBL07において、放射光マイ クロCT[4](エネルギー10keV、露光時間2秒、 1000投影/360度)で取得した切り餅の断面像 について、2Dフィルターを適用した結果(図 2中央)を示す。適用前の断面像(図2左) と比べてリングアーチファクトを大幅に低減



2Dフィルター未適用

2Dフィルター適用後

サイノグラムの差分像

図2 2Dフィルターの適用例。リングアーチファクトを大幅に低減できているが、円周方向 に像の滲みが現れる。縦方向のハイパスフィルターによりサイノグラム(右)で像が上下方 向にぼけることが原因。



図3 各パラメータの適用結果。a1の増加に伴い、滲みは減少するが、リングアーチファクトは増加する。本例では a1=20 程度が最適と考えられる。

できているが、気泡(黒い穴)の周辺で像が 円周方向に滲んでいることがわかる。これは、 2Dフィルターのうち、縦方向のハイパスフィ ルターによって、サイノグラム上で気泡の信 号成分が縦方向(角度方向)に尾を引いてし まうことが原因である(図2右)。

2D フィルターではサンプルとリングアー チファクトの両信号について、同時にフィル タリング動作を行っている。このため、リン グアーチファクトの低減と画像の滲み(歪み) はトレードオフの関係になる。(原理的に無歪 でリングアーチファクトを低減することはで きない。) そこで、2D フィルターのパラメー タ (a 及び b) と滲みとの関連性を評価し、パ ラメータの最適化を試みた。

図3に同サンプルを対象として、カットオフ周波数であるa2とb2が5,10,20,30,40の2 Dフィルターを適用した結果を示す。a2及び b2が大きくなるにしたがって、滲みが少しず つ減少し、40ではかなり抑えられていること がわかる。一方、リングアーチファクトは予 想通り次第に増加してしまっている。この結 果から、本サンプル(切り餅)のように CT 値の変化が大きい場合は、両者のバランスが 取れている a₂=b₂=20程度が最適値と考えられ る。

図4には、植物(キウイ)の種子を対象と した同様の評価結果を示す(利用した放射光 施設や計測条件等は図3と同じ)。切り餅に比 べて構造が非常に複雑で個々のサイズが小さ く、且つ、密度差(CT値の変動)も少ない ために、断面像には滲みがほとんど現れてい ない。一方で、a2及びb2が大きくなるにした がって、リングアーチファクトによるCT値 のずれが右下から画像の中心に向かって(橙 色で示した領域)広がっていることがわかる。 このため、本サンプルのように CT 値の変化 が小さい場合は、リングアーチファクトの低 減効果が最大となる a2=b2=5 が最適値と考え られる。





3. まとめ

2D フィルターで発生する円周方向の滲み に関して、カットオフ周波数との関連性を実 データを対象として評価した。この結果、サ ンプル内の気泡など CT 値が大きく変動する 領域を含む場合は、フィルターの時定数を大 きくすることで滲みをある程度は抑えられる ことがわかった。一方、CT値の変動が小さ いサンプルの場合は、リングアーチファクト の低減効果をより強くするために、時定数を できるだけ小さくする方が良いことがわかっ た。

謝辞

放射光を用いたマイクロCTの計測は九州 シンクロトロン光研究センターの課題番号 190202I及び 200202Iにより行った。

利益相反の有無

なし

文 献

- [1] Yoneyama A, Baba R, Hyodo K, Takeda T, et al., AIP Conference Proceedings: 1696, 020007, 2016.
- [2] Yoneyama A, Baba R and Kawamoto M, Optical Materials Express 11: 398-411, 2021.
- [3] 米山明男、馬場理香、河本正秀、第39 回日本医用画像工学会大会:P4-3, 2020.
- [4] 米山明男、馬場理香、河本正秀、第 81回応物秋季学術講演会: 9a-Z25-2, 2020

Quantitative analysis of 2-directional Fourier transform filter

for rig artifact reduction

Akio YONEYAMA^{*1}, Rika BABA^{1,21}, MAsahide KAWAMOTO¹

*1 SAGA Light Source *2 Hitachi, Ltd.

The ring artifact reduction filter using 2D Fourier transform (2D filter) is a filter that reduces ring artifacts by combining a vertical high-pass filter with a horizontal low-pass filter, taking advantage of the fact that ring artifacts appear as vertical stripes on a sinogram. Since the 2D filter consists of simple Fourier transforms, it is characterized by its simple parameter settings, high speed, and high versatility. The filter has been applied to seeds, wood, and fossils acquired by synchrotron radiation based micro-CT, and has significantly reduced the ring artifacts caused by X-ray imager's defects while maintaining the CT values. In this paper, we report the results of the evaluation of the relationship between the parameters of the 2D filter and the circumferential image blurring caused by the filter on the reconstructed sectional image.

Key words: X-ray CT, Ring artifact, Fourier transform

GP-GPU を用いた光子輸送シミュレーションの高速化

瀬在 翔太*1 村田 一心*2 乳井 嘉之*3 尾川 浩一*4

要旨

現在、乳がんによる死亡者数は増加傾向にあり、早期発見のために高性能なマンモグラフィが必要となっている.マンモグラフィは低エネルギーのX線を使用しているため、コヒーレント散乱による散乱線の除去は、画質やコントラストの向上につながる.しかし、コヒーレント散乱のような複雑な計算には膨大な時間を要する.そこで散乱線を除去するために、GPUで高速計算が可能なシミュレーションコードを開発した.さらに計算時間を評価するために簡単なシミュレーションを行い、EGS5と比較をした.その結果、計算時間はEGS5の約 1/13となった.次に、散乱線量の測定を行うためにマンモグラフィ装置の形状を模したシミュレーションを行い、ファントムの厚さ、X線の入射角度による散乱割合を測定し、散乱線割合が変化していることを確認することができた.

キーワード:マンモグラフィ,モンテカルロシミュレーション,X線

1. はじめに

今日,世界的に乳がんの罹患率が増加してお り,それに伴う死亡が問題となっている.この 問題を解決するために有効なのが,マンモグラ フィによる乳がんの早期発見であり,高性能な マンモグラフィ装置が必要になる.マンモグラ フィでは低エネルギーの X 線を使用している ため,コヒーレント散乱がコントラストに大き く影響し,散乱光子を除去することが画像の質 を向上させるために重要となる.本研究では, 散乱線除去のために必要となる散乱線量の定 量評価を行った.一般的に散乱光子の計算には, EGS などの汎用シミュレーションソフトが広 く使われている.しかし,計算時間が長くマン

*1 法政大学大学院理工学研究科応用情 報工学専攻

〔〒184-8584 東京都小金井市梶野町 3-7-2〕

e-mail: shota.sezai.2a@stu.hosei.ac.jp

*2 国立天文台

*3 東京都立大学健康福祉学部放射線学 科

*4 法政大学理工学部応用情報工学科

モグラフィの画像処理法などを開発するには 実用的ではない.そこで,独自のシミュレーシ ョンコードを開発し,GPUに実装することで計 算を高速化した.また,このコードを用いて, 散乱線除去に必要となる撮影条件によって変 化する散乱線量の評価を行った.

2. 手法

1) 高速化コードの作成

今回作成した GPU コードでは,1つの光子輸送計算をエネルギービンごとに1つのスレッド に割り当てた.使用した GPU は,Tesla P100 (Pascal)であり,グリッドサイズは 100,000, ブロックサイズは 1,000 に設定した.これらの パラメータは試行錯誤によって求めた計算に 最も効率的なパラメータであり,理想的なスレ ッド数で計算を高速化した.使用した CPU は Kabylake-Y で,クロック周波数は 1.2 GHz,メ モリは 8 GB である.

1つ目のシミュレーションとして,計算時間を 評価するために,簡単なモデルをシミュレーシ ョンし, EGS5 の計算時間と比較した.図1に シミュレーションのジオメトリと条件を示す. 評価を簡略化するために 30 keV のγ線源を使 用し,光子は原点から x-y 平面に対し垂直に放
出した.また,線源と検出器の間の空間は水で 満たしたものを想定した.このシミュレーショ ンファントムを用いて,自作コードで得られた 結果と EGS5 で得られた結果から計算時間を比 較した.

2) 散乱線量の評価

散乱線量を測定するため, マンモグラフィ装 置を模したジオメトリでシミュレーションを 行った.図2にシミュレーションのジオメトリ を示す. θは X線の照射範囲であり,実際の装 置を模して24°に設定した.また、Lは線源と 媒質の間の距離であり、こちらも実際の装置を 模して 65 cm に設定した. T は媒質の厚さであ り,測定毎に設定した.また,線源は管電圧30 kVのX線管とし、モリブデンターゲットと、 厚さ 0.03 mm のモリブデンフィルターを装着し たものを想定した.以上の条件で ファントムの 厚さ,X線照射角度による散乱線量を測定した. 照射角度による測定ではファントム厚 T を 4.5 cm とした. 定量的な評価として, プライマリ 光子数に対する散乱光子数で表される散乱率 を算出した.



図 1. シミュレーションのジオメトリと条件 (EGS との計算時間比較のため, 簡易的なジ オメトリを設定した.)



図 2. シミュレーションのジオメトリ(実際の 装置を模したジオメトリを想定した.)

3. 結果と考察

1) 自作コードと EGS5 による結果の比較 1 つ目のシミュレーションの結果を図 3,4 に 示す.図3は,自作したコードで計算した x 軸 方向に沿って検出された光子の分布で,図4 は EGS5 で計算したものである.これらの結果は 概ね一致している.光子数のわずかな違いは, クロスセクションデータの相違と考えられる. すなわち,我々の使用したものは X-com のデー タであり,EGS5 は独自に用意された断面積デ ータを使っている.計算時間については,EGS5 が 2867 秒であるのに対し,我々の自作コード は 231 秒であり,EGS5 の約 12.4 倍の速さで計 算することができた.



図 3. 自作コードによるシミュレーション結 果(x軸に沿って検出された光子数を示す.)



図 4. EGS5 によるシミュレーション結果(x 軸に沿って検出された光子数を示す.)

2) 自作コードによる散乱線量の測定

実際のマンモグラフィに即したジオメトリ で、散乱線量を測定した結果を以下に示す.図 5は、水の厚さT[cm]に応じた散乱割合を表す. 水の厚さと散乱割合が比例することがわかる. 通過距離が長いほど散乱が発生する確率が高 くなるので、水の厚さが増すほど散乱光子の比 率が高くなっていると考えられる.図6はX線 入射角度に応じた散乱割合を表す.入射角度が 大きくなるにつれ、散乱割合が大きくなること がわかる.こちらも入射角が大きいほど、光子 が物体内を通過するパスが長くなり、散乱が発 生する確率が高くなるため、散乱光子の比率が 高くなっていると考えられる.

これらの図に示すように散乱光子の定量的な 評価を自作コードで行うことができた.また, これらの結果から,ファントムの厚さ,入射角 によって散乱線量を推定し,散乱線を除去でき る可能性を見出せた.



図 5. 水の厚さと散乱線量の関係(厚さごと の散乱割合を示している.)



図 6. X 線入射角度による散乱割合(x 線の 入射角度ごとの散乱率を示している.)

4. 結論

本研究では、GPU を用いたモンテカルロシミ ュレーションコードを開発した.このコードを 用いて、汎用ソフト EGS5 と比べ、約 12.4 倍の 高速化を達成することができた.また、散乱線 除去のための、ファントムの厚み、入射角度に よる、散乱割合の変化を確認した.

利益相反の有無

無し

参考文献

[1] 栗原孝史, 尾川浩一"モンテカルロ法を用いた光子輸送計算において高速計算を可能とする物体記述法",

Medical Imaging Technology, Vol.26, No.1, January 2008

- [2] J. H. Hubbell, W. J. Veigele, E. A. Briggs, etc. "Atomic form factors, incoherent scattering functions, and photon scattering cross sections ", J. Phys. Chem. Ref. Data, 4:471– 538, 1975.
- [3] H. Hirayama, Y. Namito, A.F. Bielajew, etc. "THE EGS5 CODE SYSTEM", January 13, 2016

Acceleration of a photon transportation simulation with a GP-GPU

Shota SEZEI*1, Kazumi MURATA*2, Yoshiyuki Nyui*3, Koichi OGAWA*4

*1 Graduate School of Science and Engineering, Hosei Univ., Tokyo
*2 National Astronomical Observatory of Japan
*3 Faculty of Health Science, Tokyo Metropolitan Univ., Tokyo, Japan

*4 Faculty of Science and Engineering, Hosei Univ., Tokyo, Japan

Currently, the number of deaths from breast cancer is increasing, and high-performance mammography is needed for early detection of this disease. Since mammography uses low-energy X-rays, the removal of scattered photons due to coherent scattering improves quality of images. However, complex calculations for the coherent scattering require a long calculation time. To reduce a computation time, we have developed a simulation code with a GP-GPU. And to evaluate the performance, we compared our code with EGS5. As a result, the calculation time was reduced to about 1/13 of EGS5. Then the distribution of scattered photon was analyzed with a geometry of an actual mammography system. The simulation results showed that the scatter/primary ratio depended on the thickness of the phantom and the angle of incidence of the X-rays.

Key words: mammography, Monte Carlo simulation, X-rays, EGS5

深層学習を利用した信号の間引きパターンに頑健な

圧縮センシング再構成の基礎検討

渋井 雅希*1 伊藤 聡志*1 山登一輝*1

要旨

圧縮センシング(Compressed Sensing: CS) を MRI に応用する CS-MRI は少数の信号から MR 画像を再 構成できるため撮像時間の短縮化が図れるが, 画像再構成には多くの時間を要することなどの問題がある. 近年,ディープラーニング(DL)を用いた CS の再構成は高速,かつ高画質であることが報告されており, 大きな注目を集めている.一方で,学習に多くの時間を費やすこと,また,信号の間引きパターンが変わ ると再学習が必要であるなどの課題がある.そこで,本研究では,信号-画像間学習による深層学習再構成 を利用し,信号の間引きパターンによらず良質な画像を再構成できる頑健な方法について検討を行った. ADMM をネットワーク化した ADMM-CSnet を使用し,複数の信号間引きパターンをネットワークに学習さ せたところ単一の間引きパターンを学習したネットワークより高い PSNR を得ることができた.

キーワード: 圧縮センシング, Generic-ADMM-Net, 深層学習

1. はじめに

MRI は生体のプロトン密度に由来した生体 の断層像を与えるだけでなく,解剖・機能含め た多種パラメータの画像として多次元情報の 提供を可能とする.MRIの課題は撮像時間の短 縮であり,これまで多くの方法が提案されてき た. 圧縮センシング(Compressed Sensing)[1]を MRI に応用した圧縮センシング MRI(CS-MRI)[2] は,サンプリング定理を満たさない少 数のデータから良質な MR 画像再構成が可能 であり,撮像時間の短縮が可能となる.

近年, CS-MRI の再構成に深層学習を応用する 研究が行われている.再構成像の画質改善と再 構成による時間のいずれの点において従来の 反復的再構成法より優れていることが知られ ており[3][4], 深層学習を利用した画像再構成に

*1 宇都宮大学大学院 地域創生科学研究科 工農総合科学専攻

[〒321-8585 宇都宮市陽東 7-1-2]

期待が高まっている.

深層学習再構成において,高精度かつ高速な 再構成を行う手法に Generic-ADMM-Net があ る[4]. Generic-ADMM-Net は効率的で収束保証 が証明されているアルゴリズムである交互方 向乗数法(Alternating Direction Method of Multipliers : ADMM)[5][6]の反復手順をもとにネット ワーク化され,反復回数に対応する複数のステ ージで構成される.

Generic-ADMM-Net は間引き MR 信号を入力 とし、再構成された MR 画像を出力する.従来 の ADMM を用いた反復再構成では、CS-MRI の 再構成の精度と速度に影響を与えるパラメー タ値は経験的に決定され、最適値を選択し使用 することは容易ではなかった.Generic-ADMM-Net は間引き MR 信号と、元の MR 信号を使用 して再構成された画像のペアを用い、学習を行 うことで最適なパラメータの選択が行える. そのため Generic-ADMM-Net は再構成の精度 と速度の両方で従来法より良好な画像再構成 が期待できる[4].

e-mail:mc216838@cc.utsunomiya-u.ac.jp

深層学習を利用する CS 再構成法では,単一 の信号間引きパターンで得られた MR 信号と教 師像との関係を学習することが多い.よって, 画像再構成時に学習時とは異なる信号間引き パターンを使用すると再構成像は乱れること になる.学習時とは異なる信号間引きパターン を使用したい場合は,再学習を行う必要があり, ネットワークの適用範囲に制限があった.そこ で,本研究では再構成画像の品質向上とネット ワークの有用性を拡張することを目的とし,信 号間引きパターンに依存せず良質な再構成像 を与える信号間引きパターンに頑健な深層学 習再構成法についての検討を行った.

2. Generic-ADMM-Net

2.1. 交互方向乗数法(ADMM)

ADMM とは、二つのコスト関数の組み合わせ による最適化問題を制約付きの最適化問題と 考え直し、その制約付き最適化問題に対して拡 張ラグランジュ法を適応する方法である[7]. x を再構成画像, y を MR 信号空間の間引き信 号とする.

一般的なCS-MRIモデルは式(1)の最適化問題 を解くことで再構成画像が得らえる.

 $\hat{x} = \underset{x}{\operatorname{argmin}} \left\{ \frac{1}{2} \|Ax - y\|_{2}^{2} + \sum_{l=1}^{L} \lambda_{l} g(D_{l} x) \right\},$ (1) ここで A=PF は観測行列, P は間引き行列, F はフーリエ変換, D_{l} はスパース性を促す変換行 列g(・)は正則化関数, λ_{l} は正則化パラメータ, L はフィルタ数を表す.

ADMM では式(1)の最適化問題に新たに補助 変数を用いて制約条件を加えた以下の最適化 問題について考える.

$$\min_{x, z} \left\{ \frac{1}{2} \|Ax - y\|_{2}^{2} + \sum_{l=1}^{L} \lambda_{l} g(D_{l} z) \right\}$$
(2)
s.t. $z = x$

式(2)に拡張ラグランジュ法を適応する.

$$L_{aug}(x, z, h) = \frac{1}{2} ||Ax - y||_{2}^{2} + \sum_{l=1}^{L} \lambda_{l} g(D_{l} z) + \langle a, x - z \rangle + \frac{\rho}{2} ||x - z||_{2}^{2}$$
(3)

ここで α はラグランジュ乗数, ρ はペナルティ パラメータである. $\beta = \alpha/\rho$ とし,式(3)に変数 分離を施し式(4)の3つの最適化問題を解くこと で元の変数 x,新たに導入した z, ラグランジ ュ乗数 α を最適化する.

$$\begin{cases} \arg \min_{x} \frac{1}{2} \|Ax - y\|_{2}^{2} + \frac{\rho}{2} \|x + \beta - z\|_{2}^{2}, \\ \arg \min_{z} \sum_{l=1}^{L} \lambda_{l} g(D_{l}z) + \frac{\rho}{2} \|x + \beta - z\|_{2}^{2} \qquad (4) \\ \arg \max_{\beta} \langle \beta, x - z \rangle, \end{cases}$$

2.2. Generic-ADMM-Netの構造

本研究で使用する Generic-ADMM-Net は上記 の ADMM をもとに構成されている. ネットワ ークの各層の処理(式(5))は ADMM の式(4)より 導かれる.

$$X^{(n)}: x^{(n)} = F^{T} (P^{T} P + \rho I)^{-1} [P^{T} y + \rho F(z^{(n-1)} - \beta^{(n-1)})],$$

$$Z^{(n)}: z^{(n,k)} = \mu_{1} z^{(n,k-1)} + \mu_{2} (x^{(n)} + \beta^{(n-1)})$$

$$-\sum_{l=1}^{L} \widetilde{\lambda}_{l} D_{l}^{T} \mathcal{H} (D_{l} z^{(n,k-1)}),$$

$$M^{(n)}: \beta^{(n)} = \beta^{(n-1)} + \widetilde{\eta} (x^{(n)} - z^{(n)}),$$

(5)

ここでIはサイズN×Nの単位行列, μ_1 , μ_2 は 重みパラメータ, $\mathcal{H}(\cdot)$ は正則化関数g(・)の勾 配に対する非線形変換, $\hat{\eta}$ はパラメータの更新 率を表す.k \in 1,2,… N_t は勾配降下法の反復回数 である.

本研究で使用する Generic-ADMM-Net の構造 を図1に示す.



図1 Generic-ADMM-Net の構造

図1は式(5)の各層の処理の流れを表している. n回目の反復はn番目のステージに対応し,n番 目のステージでは再構成層(X⁽ⁿ⁾),ノイズ除去層 Z⁽ⁿ⁾,乗数更新層M⁽ⁿ⁾からなる3つの層により構 成される.ノイズ除去層Z⁽ⁿ⁾は加算層(A),畳み 込み層(*C*₁, *C*₂), 非線形変換層*H*からなるサブス テージで構成され, *N*_t回の反復処理を行う. 入力された間引き MR 信号は各ステージで連続 的に処理され, 再構成層(*X*^{(*N*+1})により再構成M R画像が出力される.

3. シミュレーション設定

3.1. 再構成実験に使用した画像

学習及びテスト画像には, IXI Dataset[8]に含 まれる頭部T₁強調像 75 枚を使用した.使用した 画像 75 枚のうち学習に 50 枚, テストに 25 枚 に分割し, 再構成実験を行った.

3.2. 信号間引きパターン

信号間引きパターン,画像データのサイズは, 256×256 とし,信号収集はカルテシアン軌道と し,信号空間の中心 30 行を間引きせず連続収 集するものを使用した.また,使用した信間引 きパターンの信号収集率と信号収集分布を図 2 に示す.

3.3. 計算機環境·学習条件

本検討の学習と再構成に使用した計算機環 境を表1に示す。学習は表2の学習条件で GPU を用いて行い約13時を要した.

表1計算機環境				
CPU	Intel Core i7-4790 (3.60GHz)			
GPU	NVIDIA GeForce GTX 1080			
RAM	32.0 GB			
Software	MATLAB 2017b			
	CUDA Toolkit 9.0 / cuDNN 7.0.5			

|--|

エポック数	300
ステージ数	10
Z ⁽ⁿ⁾ 層での反復回数	1
畳み込みフィルタサイズ	5×5
畳み込みフィルタ数	128
パラメータの更新手法	adam[9]





図2 使用した信号間引きパターン

4. 学習と再構成方法

学習と再構成の方法を図3と図4に示す.従 来型の学習では図3のように1種類の信号収 集パターン(固定)で間引いた信号を Generic-ADMM-Net の入力としていた.本研究では図2 の信号間引きパターン A~E の中からランダム に1つを選択して間引き処理を行った信号を入 力として学習を行うように変更した.再構成テ ストでは、学習に用いた図2A~Eの5種の信号 間引きパターンと学習に使用しない信号間引 きパターンを用いた.画像の評価にはピーク信 号対雑音比(Peak Signal-to-Noise Ratio: PSNR)と 輝度値・コントラスト・構造の3つをもとに評 価を行う SSIM(Structural SIMilarity index)[10]を 用いた.



図31通りの間引きパターンを使用する従来法



図45通りの間引きパターンを使用する提案法

5. 再構成結果

表 3,表 4 に再構成像の PSNR と SSIM,図 5 に再構成像を示す.

5 種類の信号間引きパターンで学習したネッ トワークを Net-all,単一の信号収集パターンで 学習したネットワークを信号間引きパターン の記号を利用して Net-A~E とする.表3と表4 より Net-all はすべての信号パターンに対し Net-A~E より高い PSNR, SSIM が得られた.ま た Net-A~E は再構成テスト時と学習時の信号間 引きパターンが異なる場合,再構成性能が低下 しているのに対し,Net-all ではすべての信号収 集パターンに対し,良好な PSNR, SSIM を得ら れる.

図5は再構成時の信号間引きパターンにB を用いた再構成像であり,図5(a)は全信号を利 用した再構成像,(b)はNet-allによる再構成 像,(c)はNet-Bによる再構成像である.以後 簡略化のため,学習とテストに使用する信号 間引きパターン名を使用し,Net(学習時のパタ ーン)-Test(再構成時のパターン)と表記する. 図5において(b)Net-all-Test-Bでは,(c)に示す Net-B-Test-Bに比べて矢印部分の輪郭部が全信 号画像に忠実であり,高い再構成像を有して いることが確認できる.

表3 再構成像評価: PSNR

Net/test-pat	(a)	(b)	(c)	(d)	(e)	(f)
Net-all	36.84	35.11	32.23	29.86	39.83	32.10
Net-A	36.41	34.16	31.01	28.20	38.85	30.44
Net-B	36.31	34.51	31.52	28.91	38.85	31.09
Net-C	35.80	34.19	31.51	29.12	38.04	31.22
Net-D	34.67	33.11	30.72	29.47	35.62	-
Net-E	36.07	33.80	30.74	28.10	39.36	-

表 4 再構成像評価: SSIM

Net/test-pat	(a)	(b)	(c)	(d)	(e)	(f)
Net-all	0.969	0.956	0.929	0.894	0.981	0.926
Net-A	0.967	0.949	0.916	0.870	0.978	0.907
Net-B	0.966	0.951	0.921	0.879	0.978	0.914
Net-C	0.963	0.949	0.922	0.882	0.974	0.915
Net-D	0.955	0.939	0.910	0.886	0.962	-
Net-E	0.965	0.946	0.913	0.870	0.980	-



図 5 再構成結果: (a) 全信号画像, (b) Net-all 再 構成像, (c) Net-B 再構成像

次に学習に使用していない間引きパターン (f)を使用した再構成実験においても Net-all が より高い PSNR, SSIM を得ている.

6. 考察

Generic-ADMM-Net において複数の信号間引 きパターンを用いて学習することが,汎用性・ 再構成性能の向上に有効であった.ここで図 6

474

は各ネットワークにおいてフィルタ数を変更 した際の再構成結果である.図6よりフィルタ 数が8の場合, Net-all-Test-BとNet-B-Test-Bで は PSNR の差がわずかであることがわかる.し かしフィルタ数が増えるにつれ Net-all におい て PSNR の向上がみられ,フィルタ数が 128 で 単一の信号間引きより大幅に高い PSNR が得ら れた.これは学習に使用していない間引きパタ ーンFを用いた場合でも同様の結果が得らえた. これは3の手法では信号間引きパターンが1種 類のため、画像1枚に対するアーチファクトの 生じ方は規則的であったのに対し,図4の手法 では複数の信号間引きパターンを用いるため より多くのアーチファクトのパターンを学習 することになる.これは一種のデータ拡張が行 われていると考えられ、フィルタを有効に活用 した学習を行うことで, 汎用性・再構成性能が 向上したと考えられる. 今後は, 使用する信号 収集パターンをさらに増加させ,再構成性能と 汎用性・実用性の更なる向上を検討する予定で ある.

7. まとめ

本研究では,信号間引きパターンに頑健な深層 学習再構成ネットワークについて検討を行っ た.複数の信号間引きパターンを学習に使用し た結果,単一の信号間引きパターンによる学習 に比べ,再構成性能の改善が示された.そして, この性能向上には,各ステージに含まれるフィ ルタ数が関係していることが明らかになった. 今後は信号間引きパターンやフィルタ数を増 や,さらなる汎用性・再構成性能の向上を図る 予定である.

謝辞

本研究の一部は,科学研究費助成金 (16K06379, 19K04423)により実施された.また, 本研究を遂行する際に利用した IXI Dataset を 提供する Information eXtraction from Images プロ ジェクトに感謝の意を表します.



(a) 学習と再構成に使用する間引きパターンが 一致する場合



(b) 学習に使用しない間引きパターンを再構成
 に利用する場合
 図 6 フィルタ数と PSNR の関係

利益相反の有無

なし

文 献

- D.L.Donoho, "Compressed sensing," IEEE Transactions on Information Theory, vol.52, no.4, pp.1289-1306, 2006
- [2] M.Lustig, D.L.Donoho, J.M.Santos, J.M.Pauly, "Compressed Sensing MRI," IEEE Signal Processing Magazine, vol.25, no.2, pp.72-82, 2008
- [3] Y.Yang, J.Sun, H.Li, Z.Xu, "Deep Admm-Net for Compressive Sensing MRI," Advances in Neural Information Processing Systems, pp.10-18, 2016.
- [4] Y.Yang, J.Sun, H.Li, Z.Xu, "ADMM-CSNet: A Deep Learning Approach for Image Compressive Sensing", IEEE Transactions on Pattern Analysis and Machine Intelli- gence, Vol.42, pp.521-538, 2020.

- [5] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statical learning via the alternating direction method of multipliers," Foundation and Trends in Machine Learning, vol. 3, no. 1, pp. 1–122, 2011.
- [6] Y. Wang, W. Yin, and J. Zeng, "Global convergence of admm in nonconvex nonsmooth optimization," arXiv preprint arXiv:1511.06324, 2015
- [7] Afonso MV, Bioucas-Dias JM, Figueiredo MAT : An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems. IEEE Trans Image Process. 2011;20:681–95.
- [8] IXI Dataset," https://brain-development.org/ixi-dataset/
- [9] D.P.Kingma, J.Ba, "Adam: A Method for Stochastic Optimization," International Conference on Learning Representations, San Diego, 2015
- [10] Z.Wang, A.C.Bovik, H.R.Sheikh, E.P.Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Transactions on Image Processing, vol.13, no.4, pp.600-612, 2004

Study on Deep Learning based MR Compressed Sensing Image Reconstruction Robust to Signal Under-sampling Pattern

Masaki SIBUI^{*1}, Satoshi ITO^{*1}, Kazuki YAMATO^{*1}

*1 Graduate Program in Information, Electrical and Electronic Systems Engineering Utsunomiya University

Compressed sensing (CS) has attracted attention as the method that can shorten the scanning time of MRI(CS-MRI), however, it requires an extensive calculation time and it sometimes gives artificial appearances to reconstructed images. Recently, image reconstruction using deep learning has attracting attentions that can solve those problems. In the learning of deep convolution neural network (CNN), one sampling pattern of k-space is sometimes used for learning of deep CNN and an additional learning will be required when a different sampling pattern is used for image scanning. In this study, an improvement of robustness to sampling patterns used with CS reconstruction was examined by learning multiple sampling patterns in the learning of the deep-ADMM-CSNet. Simulation results learning various sampling patterns fixing the sampling rate showed that higher PSNRs and SSIMs were obtained with deep CNN that learn multiple sampling patterns compared to single-pattern-learned CNNs as shown in Fig.1. These studies indicate that the robustness to sampling pattern can be improved by learning multiple sampling patterns in deep-ADMM-CSNet.

Key words: Compressed sensing, Generic-ADMM-Net, deep learning

WGI コンプトン画像再構成条件の実験的検討

田島 英朗^{*1} 吉田 英治^{*1} 仁科 匠^{*1,*2} 菅 幹生^{*1,*2} 脇坂 秀克^{*1} 高橋 美和子^{*1} 永津 弘太郎^{*1} 辻 厚至^{*1} 鎌田 圭^{*3,*4} 吉川 彰^{*3,*4} Katia Parodi^{*5} 山谷 泰賀^{*1}

要旨

我々が開発した WGI(Whole Gamma Imaging)の小動物実証機は,世界初のフルリング型コンプト ンイメージングシステムであり,これまでに高精細なマウス画像が得られたことを報告した.フ ルリングにより高感度化を達成したことが要因の一つとして考えられるが,理論的にコンプトン 画像再構成の条件としては必ずしもフルリング型である必要性はなく,検出器の削減により,感 度は犠牲になるが,コスト削減や設計の柔軟性向上が期待できる.しかしながら,実際には散乱 角度の検出限界やブロック型の検出器配置の影響等で,再構成条件は保証されていない.そこで, 本研究では,試作装置の測定データを限定することで部分リング化し,フルリングの場合と比較 することで,コンプトン画像再構成に必要なジオメトリ条件を実験的に検討した.円筒型ファン トムの測定データを用い,再構成画像を視覚的に評価した結果,散乱検出器もしくは吸収検出器 に囲まれる範囲が180°未満の領域が物体にある場合,画像にアーチファクトが生じたが,物体 全域が散乱検出器と吸収検出器の共に180°以上囲まれている場合には、フルリングと同様にア ーチファクトのない画像が得られた.よって,WGIコンプトン画像再構成の完全性条件として, 散乱検出器,吸収検出器共に,対象視野全域を180°以上共通の角度範囲として囲む必要がある ことが示唆された.

キーワード:WGI, コンプトン画像再構成, 画像再構成条件

1. はじめに

我々が開発を進めている WGI (Whole

*1 量子科学技術研究開発機構
〔〒263-8555 千葉市稲毛区穴川 4-9-1〕
e-mail: tashima.hideaki@qst.go.jp
*2 千葉大学
*3 東北大学
*4 株式会社 C&A

*5 ミュンヘン大学

Gamma Imaging) は、PET (Positron Emission Tomography) のリングの内側に散乱検出器リ ングを挿入し、シングルガンマ線をコンプト ンカメラ (コンプトンイメージング)の原理 で検出することで、様々なガンマ線をイメー ジングに利用する独自のコンセプトである [1]. これまでに、小動物用の小型実証機を開 発し、高精細なマウス画像が得られたことを 報告した [2,3]. ここで、コンプトンイメー ジングの再構成条件に着目してみると、散乱 検出器をコーンビーム CT の線源軌道と考え, 各素子からの線積分をコンプトンイメージン グの測定データから再構成できるとすれば [4],完全再構成のためには必ずしもフルリン グである必要性はない [5].しかしながら, 実際には,検出器性能により検出可能な散乱 角度の制限があったり,ブロック型の検出器 の配置による影響があったりすることから, 理論的な再構成条件は必ずしも保証されてい ない.そこで,本研究では,WGI 実証機の測 定データを制限することで散乱検出器と吸収 検出器を部分リング化し,フルリングの場合 と比較することで,3 次元画像再構成が可能 なコンプトンイメージングジオメトリの検証 を実験的に行った.

2. 方法

WGI 小動物実証機は、1 リングあたり 40 個の4層 DOI 検出器 (2.8×2.8×7.5 mm³ GSO 結晶, 16×16×4 配列) を体軸方向に 4 列並 べた直径 660 mm の PET リング(吸収検出器 リング)と、1 リングあたり 10 個の散乱検出 器ブロック (0.9×0.9×6.0 mm³ GAGG 結晶, 24×24 配列)を体軸方向に2列並べた直径94 mm の散乱検出器リングから構成される. 実 証機を用いて 10.2 MBq の⁸⁹Zr 溶液を封入し た直径 38 mm の円筒ファントムを 60 分間測 定し, 909 keV のシングルガンマ線をコンプ トンイベントとして抽出した.また、検出器 を部分リング化したジオメトリを模擬するた めに,画像再構成を行う際に使用する測定デ ータをジオメトリとして設定した検出器で得 られたもののみに限定した.また、その際、 検出器が物体を覆う最小の角度範囲として, 円筒の再下点に対する検出器リングの角度範 囲を求めた.まず,吸収検出器をフルリング, 散乱検出器リングのリング内検出器の数を図 1 (b)のように 1, 3, 5, 7, 9 個とし, 10 個 (フ ルリング)の場合と比較した.次に、散乱検 出器は最小角度範囲が180度以上となる7個 とし、図1(c)のように吸収検出器のリング内 個数を3,9,15,23,33とした場合を模擬した.



図1WGI 実証機を用いた円筒ファントム測定(a), 部分リングとして設定した散乱検出器(b)と 吸収検出器(c)のリング内検出器数と物体を リングが覆う最小角度範囲.

なお,吸収検出器の個数は,最小角度範囲が 散乱検出器リングで設定した各値と近くなる 数として設定した.画像再構成には WGI コン プトン画像再構成用に開発した 3D list-mode ordered subset maximization expectation 法を適 用し[3],ボクセルサイズは1×1×1 mm³,サ ブセット数は16,反復回数は50とした.各 再構成像について,視覚的にアーチファクト を両リングともフルリングの場合と比較する ことで評価した.

3. 結果

図2に散乱検出器の数を変化させた場合の 結果を示す.リングあたりの散乱検出器数が 1の場合,冠状断では円筒ファントムの形状 がわかる画像となっていたが,軸位断と矢状 断では散乱検出器の位置から放射状に激しい ストリークアーチファクトが現れていた.検 出器数を3に増やすと,ストリークアーチフ ァクトは低減されたものの,散乱検出器の反 対側に分布が伸びるように歪んでいた.5 個 まで増やすとほぼ円筒形が再現され,わずか に下側に形状が歪む程度となった.そして,7 個まで増やすとフルリングと同等の画像が得 られた.つまり,物体を覆う最小角度範囲が 180°を超えた際にアーチファクトがなくなった.図3に,散乱検出器数をリングあたり7とし,吸収検出器数を変化させた場合の再構成像を示す.リングあたりの吸収検出器数が3の場合,縦方向に縮み,ストリークアーチファクトが現れた.吸収検出器数を増やしていくと,円筒の形状が復元され,23個以上(180°を超えた場合)ではほぼフルリングと同等の画像となった.これは,散乱検出器リングと吸収検出器リングの両方が,測定対象上のすべての領域に対して180°以上,共通の角度範囲を覆っている条件であった.

4. まとめ

WGI コンプトンイメージングの再構成条 件として,物体を覆う散乱検出器リング,及 び吸収検出器リングの角度範囲の要件につい て実験的に検討した.円筒ファントムの測定 を行い、物体中で最小の角度範囲となる最下 点から見たリングの角度範囲を最小角度範囲 とし, 散乱検出器リングと吸収検出器リング の最小角度範囲を変化させながら再構成画像 の変化を視覚的に評価した結果、両方のリン グがともに 180°以上の場合にアーチファク トのない画像を得ることができた.よって, WGI コンプトンイメージングにおいて3次元 画像再構成を行うための条件として, 散乱検 出器,吸収検出器共に,測定対象上のすべて の領域を共通の角度範囲として 180°以上囲 うことが必要であると示唆された.

謝辞

本研究は JSPS 科研費 20K12683 及び 20H05667の助成を受けたものです.

利益相反の有無

なし

文 献

[1] Yoshida E, Tashima H, Nagatsu K, et al.:Whole gamma imaging: a new concept of PET combined with Compton



図 2 散乱検出器リングの検出器数を変化させた 場合の再構成画像



図3 散乱検出器リングの検出器数を7(最小角度 範囲が180度以上となる数)とし,吸収検出 器リングの検出器を変化させた場合の再構 成画像

imaging. Phys Med Biol 65: 125013, 2020

- [2] 田島英朗,吉田英治,脇坂秀克,他: WGI 試作機によるフルリングコンプトンイメージングの実証.第39回日本医用画像工学会大会予稿集: 454-458,2020
- [3] Tashima H, Yoshida E, Wakizaka H, et al.: 3D Compton image reconstruction method for whole gamma imaging. Phys Med Biol 65: 225038, 2020
- [4] Smith B: Line-reconstruction from Compton cameras: data sets and camera design. Opt Eng 50: 053204, 2011
- [5] Tuy H K: An inversion formula for cone-beam reconstruction. SIAM J Appl Math 43: 546-52, 1983

Experimental Assessment of the WGI Compton Imaging Reconstruction

Condition

Hideaki TASHIMA^{*1}, Eiji YOSHIDA^{*1}, Takumi NISHINA^{*1, *2}, Mikio SUGA^{*1, *2}, Hidekatsu WAKIZAKA^{*1}, Miwako TAKAHASHI^{*1}, Kotaro NAGATSU^{*1}, Kei KAMADA^{*3, *4}, Akira YOSHIKAWA^{*3, *4}, Katia PARODI^{*5}, Taiga YAMAYA^{*1}

- *1 National Institutes for Quantum and Radiological Science and Technology
- *2 Chiba University
- *3 Tohoku University
- *4 C&A Corporation
- *5 Ludwig-Maximilians-University in Munich

We developed a small animal whole gamma imaging (WGI) prototype, and, to the best of our knowledge, it is the world's first full-ring Compton imaging system. Last year, we reported a high-definition mouse image obtained with the prototype. However, the theory in the reconstruction condition for Compton imaging states that it is unnecessary to be full-ring. Thus, reducing the number of detectors expect cost reduction and design flexibility at the cost of sacrificed sensitivity. In practice, however, the reconstruction condition is not guaranteed due to the detection limit of the scattering angle and the effect of the block-type detector arrangement. In this study, we experimentally investigated the necessary geometry condition for Compton image reconstruction by limiting the measurement data of the prototype system to simulate partial ring geometries and comparing them with the full ring system. We visually evaluated reconstructed images of a cylindrical phantom measured by the prototype. As a result, when the object was surrounded by less than 180° of scattering or absorption detectors, artifacts appeared in the images. On the other hand, when the entire object was surrounded by more than 180° of both scattering and absorption detectors, we obtained images comparable to the full-ring case. In conclusion, it is suggested that the completeness condition for WGI Compton image reconstruction requires that both the scatterer and absorber detectors surround the entire object field of view by a common angular range of at least 180°.

Key words: WGI, Compton image reconstruction, image reconstruction condition

膝半月板の 2.5 次元統計的形状モデル構築における

主成分分析と非負値行列因子分解の比較

網手 就矢^{*1} 增谷 佳孝^{*1} 北村 直幸^{*2}

要旨

膝半月板は外傷や加齢によって損傷し,厚さのみならずトポロジーの変化も含む多様な形態変化が観察される.3次元形状の評価に基づく診断には MRI 検査が使用され,抽出形状の定量的な解析は診断に有用な 情報をもたらすと期待される.我々は,膝半月板形状の厚さのみの情報を抽出した2.5次元の統計的形状モ デルの構築を研究している.同モデルの構築では,半月板形状サンプルが2次元の濃淡画像で表現され, 各画素は厚さの値を持つ.本研究では,その画像サンプル群の統計解析に主成分分析と非負値行列因子分 解を用い,基底パターンへの分解結果の視覚的比較,形状再現における残差などの定量比較を行ったので 報告する.使用する基底数の影響についても検討し,比較結果に基づいて両手法の特性および得失を議論 し,診断に有効な情報をもたらす2.5次元の統計的形状モデルの構築を目指す.

キーワード:膝半月板, 2.5 次元統計的形状モデル, 主成分分析, 非負値行列因子分解

1. はじめに

膝半月板は大腿骨と脛骨の間に存在し,正常 な場合 C型の形状をしている.両肢共に内側と 外側にそれぞれ一つずつ存在しており,膝にか かる荷重を分散させることや衝撃を吸収する 役割を担っている.膝半月板の損傷は主にスポ ーツや加齢により起こり,それらの症例として 縦断裂,バケツ柄状断裂や円板上半月板断裂な どが挙げられる.これらの症例は正常と比べて 厚さのみならずトポロジーの変化も含む多様 な形態変化が観察される.膝半月板の3次元形 状の評価に基づく診断には MRI 検査が使用さ れ,膝半月板の抽出形状の定量的な解析は診断 に有用な情報をもたらすと期待される.

*1 広島市立大学大学院情報科学研究科 [〒731-3194 広島市安佐南区大塚東 3-4-1] e-mail: mf69005@e.hiroshima-cu.ac.jp *2 株式会社エムネス

膝半月板に限らず臓器形状の定量解析には 統計的形状モデル(Statistical Shape Model: SSM) がしばしば用いられる[1]. 臓器の SSM は, 画 像解析や統計処理の手法を組み合わせ, 個体差 のある解剖構造を統計的に表現し,形状のバリ エーションを記述するもので,領域抽出,診断 や鑑別などに利用できる. 例えば肝臓の SSM は, 肝臓疾患の一つである肝硬変症の識別に用い られ、診断支援への可能性を示している[2].ま た, 膝半月板の SSM は, ニューラルネットワー クと合わせて利用することで,膝半月板領域の 抽出精度を向上させている[3]. 損傷症例に対し ても SSM を用いれば、膝半月板形状の平均や 分散,正常な分布からのずれも定量的に評価を 行うことができ,診断支援への応用が可能とな る.しかし、先述した通り、膝半月板損傷の中 でも特にバケツ柄状断裂や円板上半月板断裂 などは,正常と比べて大きく形状に差異が生じ ている. その結果, トポロジーが大きく異なる などにより, SSM で広く採用されている点分布 モデルではサンプル間の対応付けが難しくな

る場合がある.よって我々は,各画素が厚さの 情報を持った濃淡画像による 2.5 次元の統計的 形状モデル(2.5D-SSM)の構築を研究している. また,多くの SSM は主成分分析 (Principal Component Analysis: PCA)により構築されるが, 同様の解析手法として非不値行列因子分解

(Nonnegative Matrix Factorization: NMF) が挙げ られ,負の値をとらない厚さ画像を表現するの に適していると考えられる.よって本研究では, 膝半月板厚さ画像の統計解析に PCA と NMF を 用い,基底パターンへの分解結果の視覚的比較, 形状再現における残差などの定量比較を行っ たので報告する.

2. データセット

本研究では、The Osteoarthritis Initiative (OAI) [4]が公開している MRI の DESS(double-echo steady-state) データセットを用いた.これは、患 者のオリジナル DESS 画像と外側、内側半月板 それぞれのマニュアルによるセグメンテーシ ョン画像がセットになっているものである.本 研究では、このうち外側半月板セグメンテーシ ョン画像 118 サンプルを用いて統計解析を行っ た.DESS データセットの概要を表 1 に示す.

表1DESS データセットの概要

FOV	140 mm
matrix	384×384
spatial resolution	$0.365 \text{ mm} \times 0.365 \text{ mm}$
slice thickness	0.7 mm
number of slice	160

3. 膝半月板の 2.5D-SSM

2.5D-SSM構築のために外側半月板セグメンテ ーション画像を各画素に厚さの情報を持った 2 次元画像に変換する処理を行った.まずセグメ ンテーション画像のボクセルが各辺 0.365mm になるよう等方化した.次に,バウンディング ボックスを設定し,2次元画像に変換する領域 を決定した.その際に,注意すべき点として膝 半月板の前節と後節の先端部分,すなわち meniscal root 部分におけるセグメンテーション の曖昧さがある. meniscal root は輝度値が類似 した靭帯と繋がっており、 MRI 画像からマニ ュアルによるセグメンテーションを行う際に, 膝半月板をどこまでとするか明確に決定する のは難しい. そのため各データで前節と後節の 先端部分にバラツキが生じる可能性を考慮し, meniscal root 側にのみ余白が生まれるようにバ ウンディングボックスの横幅Wを縦幅Hの1.5 倍の長さに設定して region of interest (ROI)とし て切り出した (図 1). 次に, ROI を 48×32×D 画素のサイズに縮小する処理を行った. ここで D は ROI の体軸方向の長さである. 最後に, 2 次元の各座標でそれぞれの体軸方向のボクセ ル数を数えて厚さを計算し,各画素が厚さの情 報を持った2次元濃淡画像(48×32)を作成し た. これを厚さ画像とする. 作成した厚さ画像 のサンプルデータを図2に示す.



W:Hの1.5倍

図1 バウンディングボックスの設定



図2 厚さ画像のサンプルデータ(48×32)

以上より, PCA および NMF で統計解析する各 サンプルデータのベクトル**x**_iは式(1)のように 表すことができる.

 $\mathbf{x}_i = (x_{i1}, ..., x_{in})^T$ (1) ここで, *i*は*N*個のサンプルのうちi番目のサンプ ルであることを示す. x_{ij} は*j*番目の画素値(厚 さ), *n*はサンプルデータの画素数で 48× 32=1536 である.

1) PCA による 2.5D-SSM の構築

サンプルデータ群の平均ベクトル**x**と分散共 分散行列**S**を式(2)と式(3)により求める.

$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}_i \tag{2}$$

$$\mathbf{S} = \frac{1}{N} \sum_{i=1}^{N} (\mathbf{x}_i - \bar{\mathbf{x}}) (\mathbf{x}_i - \bar{\mathbf{x}})^T$$
(3)

次に分散共分散行列Sの固有値と固有ベクトル を求め、固有値の大きい順にM個の固有ベクト ルを用いて、式(4)により任意の厚さ画像を表現 する.

$$\hat{\mathbf{x}} \cong \bar{\mathbf{x}} + \sum_{i=1}^{M} c_i \, \mathbf{e}_i \tag{4}$$

ここで、 \mathbf{e}_i は*i*番目の固有ベクトル、すなわち主 成分ベクトルである.

2) NMF による 2.5D-SSM の構築

NMF は非負値(0か正の値)のみからなる数 値群を解析する手法である[5-7].本研究では, 式(5)のようにサンプルベクトルを横に並べた 非負値行列Xを定義し,式(6)により分解を行う.

 $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N] \tag{5}$

$$\mathbf{X} \cong \mathbf{T}\mathbf{V} \tag{6}$$

分解結果は、 $n \times K$ の非負値行列Tと $K \times N$ の 非負値行列Vの積の形になる.行列Tの各列は 基底ベクトル、Vの各列はエンコーディングと 呼ばれ、係数で構成されている. Kは基底の数 であり、分解を行う前に決めておく必要がある. 分解の計算は、XとTVの距離D(X,TV)を定義し、 これを最小化するように行列TとVを求める.本 研究ではユークリッド距離を Multiplicative update rules により最小化する手法を用いた.そ して、求めたK個の基底ベクトルを用いて、式 (7)により任意の厚さ画像を表現する.

$$\hat{\mathbf{x}} \cong \mathbf{T}\mathbf{v} \tag{7}$$

ここで、vは近似させる対象となるベクトルxが 与えられた際に、そのベクトルとTの一般化逆 行列の積から求められる係数ベクトルである.

4. PCAとNMFの比較

本研究では、PCA と NMF の比較を 2 つの実験 を用いて行った.ここで行う 2 つの実験の際に、 PCA では平均および使用した主成分ベクトル を合わせて基底パターンとし、NMF では使用し た基底ベクトルを基底パターンとする.実験 1 では、膝半月板厚さ画像サンプル群の基底パタ ーンへの分解結果の視覚的比較を行った.実験 2 では、同サンプル群の再構成を行い、表現能 力の定量的な評価を行った.

実験1)基底パターンの視覚的比較

PCA と 16, 24, 32, 40, 48, 56, 64 の基底数で得 られた NMF の基底パターンを図 3 に示す. な お, PCA の主成分画像(図 3(a))は負の値を含 むため, 0 がグレーとなるよう調整している. また, PCA の基底パターンは固有値の大きい順, NMF はサンプル再構成時の係数の分散が大き い順に並べ, それぞれ上位から示している.

PCA の第1主成分画像では,前節と後節の先端部分を中心とした構成要素が観察される.これは,半月板全体の厚さと前節と後節の先端領域の厚さにおける個体間の差異,およびセグメンテーションのばらつきに関わる成分と思われる.第2,第3の主成分では,外周部分の厚さに関した成分が観察される.各主成分ベクトルは平均に対する正と負の値による複雑な相殺を伴う線形結合で使用されるため,各主成分の意味を直観的に理解することは困難である.

一方,NMFにより得られた基底パターン(図 3(b))は全て非不値であり,データサンプル群に おいて類出するパーツで構成される傾向が顕 著である.例えば,基底数48の第1,第2基底 はそれぞれ,前節,後節領域に関係があるパー ツと示唆される.すなわち上位の基底は各形状 における特徴的なパーツであり,下位は各形状 に多く含まれているパーツであると理解でき る.ただし,基底数64の結果のように基底数を 多くすると分解が細かく行われ,各パーツの理 解が困難になった.以上より,適切な基底数の 設定により各基底がどの領域の厚さに関係が あるパーツを含むか直観的な解釈が可能と示 唆される.



(b) NMF: 各基底数における基底画像図 3 PCA および NMF による分解結果

実験2) 2.5D-SSM 再構成における定量比較

用いる基底パターン数を 16,24,32,40,48,56, 64 として,モデル構築に用いた全サンプルを再 構成した.再構成サンプルごとに元サンプルと の厚さの誤差を,二乗平均平方根誤差(Root Mean Square Error: RMSE),平均絶対誤差(Mean Absolute Error: MeanAE),最大絶対誤差(Max Absolute Error: MaxAE)により評価した.ただ し,PCA による再構成画像には負の値が生じる ため,負の値を 0 に修正した.さらに,厚さ以 外の 2 次元での形状評価として,Dice 係数(Dice Similarity Coefficient: DSC)による類似度,FP

 (False Positive)率, FN(False Negative)率を計 測した.これは、サンプル画像と再構成画像を 厚さ0の閾値で2値化して得られる形状の集合 演算による評価である.これらの定量評価の結 果を表2に示す.また、元サンプル画像、再構 成画像, 誤差ヒストグラム, 誤差画像を 16,32, 48,64 の各基底パターン数に対して作成した (図 4). ここで, 誤差ヒストグラムは正の誤差 と負の誤差を含む.

表2に示すように、同じ基底パターンによる 厚さ誤差を比較した際、どの基底パターンでも PCA による再構成が NMF よりも誤差が小さい 結果となった.これは、各主成分が直交してい る PCA に対して、基底が直交していない NMF の表現能力が劣るためであると考えられる.ま た、使用する基底数を増やせば両手法とも 118 サンプルの平均厚さ誤差が小さくなり, 近似の 精度が上がっている.また, PCA による再構成 では FP 率が NMF よりも高く, 前節と後節の先 端部付近に FP が多く確認された(例えば図 4(b) 黄色矢印).これは,第1主成分画像を構成する 要素である前節と後節の個体間における分散 が大きいためと示唆される. それによって DSC も,NMFによる再構成が PCA よりも高い数値 であることが確認された. FN 率は全再構成で0 であった.これは考察で述べる領域抽出に応用 する際に重要な性質であると示唆される.

図4より, PCA の厚さ誤差において, 正の誤 差と負の誤差の範囲はどの基底パターン数で も類似していることが分かる.対して NMF で は,多くの基底パターン数で正の範囲が広くな っている.つまり,加算の組み合わせのみが許 される NMF の特徴により過度な厚さを表現し ていると言える.

我 2 门阱然记忆了 3 定重計圖							
		厚さ評価			2次元形状評価		
基底裂	評価指標	RMSE (mm)	MeanAE (mm)	MaxAE (mm)	DSC (%)	FP (%)	FN (%)
16	PCA	0.604	0.4	3.306	76.1	64.0	0
10	NMF	0.736	0.463	4.069	77.7	58.1	0
24	PCA	0.517	0.344	2.877	76.8	61.6	0
24	NMF	0.685	0.432	3.876	79.9	51.0	0
32 PCA NMF	PCA	0.457	0.309	2.496	77.2	60.2	0
	NMF	0.644	0.403	3.763	80.6	48.7	0
40 PCA NMF	PCA	0.408	0.278	2.221	77.4	59.7	0
	NMF	0.618	0.388	3.614	82.1	43.9	0
40	PCA	0.363	0.25	1.954	77.7	58.7	0
40	NMF	0.578	0.356	3.455	81.5	45.7	0
56	PCA	0.323	0.224	1.713	77.9	58.0	0
	NMF	0.566	0.349	3.428	82.6	42.7	0
64	PCA	0.287	0.201	1.483	77.9	57.8	0
	NMF	0.543	0.336	3.317	83.6	39.7	0

表2再構成における定量評価

*各値は118サンプルの平均値



図4 サンプル再構成画像と誤差分布

5. 考察

本研究で得られた結果のように, PCA, NMF ともにそれぞれ特徴があり, 今後もさらに詳細 な比較評価を行う予定である. 特に, 本研究で はモデル構築に使用したサンプルデータを再 構成し形状表現の評価を行ったため, 新規の厚 さ画像に対しての再構成の評価を行う必要が ある. また, 厚さ画像のサイズや NMF におけ る誤差最小化の計算に関しても検討を行う予 定である.

また,厚さのみを評価する本研究の2.5D-SSM は,疾患等による半月板の全体および部分の摩 耗や先天性の異常などの定量評価に有用であ ると考えられるが,新規 MRI データでの解析に は領域抽出が必要である.これまでの膝関節内 部位の領域抽出に関する先行研究では,ユーザ ーが最初に ROI の設定のみを行う半自動方式 や,ディープラーニングを活用した全自動方式 による手法などが提案されている[8-12].本研 究では,厚さに加えて半月板中央曲面の位置を 統合した 3D-SSM の構築を行い領域抽出への応 用を検討している.

6. おわりに

本研究では、MRI より抽出した外側半月板 118 サンプルのセグメンテーション画像を 2 次 元の厚さ画像に変換し、PCA および NMF の 2 つの統計手法を用いて 2.5D-SSM を構築した. 各基底パターンへの分解結果の視覚的評価と 表現能力の定量的評価を行い比較した.視覚的 比較において、NMF は各基底が対応してるもの の解釈がより容易であることが示唆された.再 構成結果の厚さ誤差の定量比較において、PCA が NMF よりも誤差が小さいことが確認された. 一方、2 値化後の 2 次元形状における定量比較 においては、NMF による再構成が PCA よりも 高い類似度を示した.今後、さらに詳細な評価 および 3D-SSM への拡張による領域抽出への応 用を目指す.

謝辞

本研究の成果の一部は, JST CREST (JPMJCR15D1)の助成によるものです. OAI-DESS データセットの入手に関し,ご助言およ びご助力いただいたスタンフォード大学の Akshay Chaudhari 博士と Sandy Napel 教授に感 謝します.

利益相反の有無:なし

文 献

- [1] Heimann T, Meinze HP, "Statistical shape models for 3D medical image segmentation: A review," Med Imag Anal, 13: 543-563, 2009
- [2] 上谷芽衣,健山智子,小原伸哉,他,"肝臓の統計的形状モデルの構築と肝硬変症支援診断への応用,"電気学会論文誌C,133:2027-2043,2013
- [3] Tack A, Mukhopadhyay A, Zachow S,
 "Knee menisci segmentation using convolutional neural networks: data from the Osteoarthritis Initiative," Ost Cart 26: 680-688, 2018
- [4] The Osteoarthritis Initiative URL: <u>https://nda.nih.gov/oai/</u>
- [5] Lee DD, Seung HS, "Learning the parts of objects by nonnegative matrix

factorization," Nature, 401: 788-791, 1999

- [6] 澤田宏, "非不値行列因子分解 NMF の 基礎とデータ/信号解析への応用,"電 子情報通信学会, 95: 829-833, 2012
- [7] 亀岡弘和, "非不値行列因子分解,"計測 と制御, **51**: 835-844, 2012
- [8] Rahman MM, Dürselen L, Seitz AM,
 "Automatic segmentation of knee menisci -A systematic review," Art Int Med, 105: 101849, 2020
- [9] Swanson MS, Prescott JW, Best TM, et al., "Semi-automated segmentation to assess the lateral meniscus in normal and osteoarthritic knees," Ost Cart 18: 344-353, 2010
- Panfilov E, Tiulpin A, Nieminen M, et al.,
 "Automatic Segmentation of Knee Cartilage and Menisci from MRI Data: Efficient Multiclass Solution Based on Deep Learning," proc. ISMRM2019, #1137, 2019
- [11] Chaudhari AS, Stevens KJ, Wood JP, et al.,
 "Utility of deep learning super-resolution in the context of osteoarthritis mri biomarkers,"
 J Magn Reson Imag 51: 768-779, 2020
- Xu Z, Niethammer M, "DeepAtlas: Joint Semi-Supervised Learning of Image Registration and Segmentation," proc. MICCAI 2019 (LNCS 11765), pp420-429, 2019

Comparison of PCA and NMF in construction of 2.5D SSM for meniscus

Shuya NAWATE^{*1}, Yoshitaka MASUTANI^{*1}, Naoyuki KITAMURA^{*2}

*1 Hiroshima City University Graduate School of Information Sciences*2 Medical Network Systems MNES Inc.

Knee meniscus is damaged by trauma and aging, and various morphological changes including not only thickness, but also topological variation are observed. MRI scan is used for diagnosis based on the evaluation of 3D shape, and quantitative analysis of extracted shape is expected to bring useful information for diagnosis. We have been studying the construction of the 2.5-dimensional statistical shape model (SSM) that has information only on the thickness of the knee meniscus shape. In the construction of this model, a meniscus shape sample is expressed in a two-dimensional gray scale image, of which pixel values express thickness. In this manuscript, we report the results of the statistical analysis of the segmented sample group using the principal component analysis and non-negative matrix factorization, quantitative comparison of the decomposition results to the basis pattern and the residual in sample shape reconstruction with various basis numbers. We also discuss the characteristics and loss of both methods based on the comparative results, and our future work aimed to extension to 3D SSM for the purpose of segmentation.

Key words: knee meniscus, 2.5-dimensional statistical shape model, principal component analysis, non-negative matrix factorization

著者紹介

縄手 就矢 (なわて しゅうや) 2020 年広島工業大学情報学部知的情報シ ステム学科卒業 2021 年現在,広島市立大学大学院情報科 学研究科医用情報科学専攻博士前期課程 在学中

増谷 佳孝 (ますたに よしたか) 1997 年東大・大学院・工学系・精密機械 工学専攻博士課程了. ハンブルク大医用 数学・情報処理研究所,シカゴ大病院放 射線科,東大病院放射線科などを経て 2021 年現在,広島市立大学大学院医用情 報科学専攻・教授. 2003 年日本医用画像 工学会論文賞および奨励賞受賞. 博士 (工学)・博士(医学). 拡散 MRI を中 心とした多次元の医用画像の処理・解析 の研究に従事. 医用画像工学会,電子情 報通信学会,磁気共鳴医学会, IEEE な ど,各会員. 北村 直幸 (きたむら なおゆき) 1993 年広島大学医学部卒業. 放射線診断 専門医. 病院勤務を経て, 2000 年に遠隔 画像診断センター「エムネス」設立. 2007 年~2020 年 3 月末まで株式会社エム ネス代表取締役を務める. 2015 年, 検査 センター「霞クリニック」開院, 院長を 兼務.

腹腔鏡映像からの手術箇所の認識による

手術ナビゲーションの検討

林 雄一郎*1 三澤 一成*2 森 健策*1.3

要旨

本稿では,腹腔鏡映像からの手術箇所の認識による腹腔鏡下手術ナビゲーションの検討について述べる. 近年,手術中の術野映像を解析して手術を支援する研究が行われている.本研究では,腹腔鏡下手術支援 のために腹腔鏡から得られる腹腔鏡映像から手術箇所を認識して手術ナビゲーションを行う手法を検討す る.まず,腹腔鏡映像から畳み込みニューラルネットワークを用いて手術を行っている解剖学的な領域を 認識する.次に,認識結果を用いて手術箇所付近の解剖学的構造が確認できる仮想腹腔鏡画像を手術支援 画像として提示する.提案手法を腹腔鏡下胃切除術において使用し,手術箇所の認識と手術支援画像の提 示が可能であることを確認した.

キーワード:腹腔鏡下手術、ディープラーニング、手術ナビゲーションシステム、胃、手術工程解析

1. はじめに

低侵襲手術の一つである腹腔鏡下手術は, 従来の開腹手術と比較して難易度の高い手術 である.そのため,腹腔鏡下手術を支援する ために手術ナビゲーションシステム [1,2] などの手術支援システムの研究開発が行われ ている.近年は,深層学習の発展に伴い,腹 腔鏡映像を深層学習により解析する研究が盛 んに行われるようになってきた [3-6].ここ では,腹腔鏡映像中の解剖学的な構造を認識 する研究 [3,4],腹腔鏡映像中の手術器具を 認識する研究 [6],などが行われている.

*1 名古屋大学大学院情報学研究科 [〒464-8601 名古屋市千種区不老町] e-mail: yhayashi@mori.m.is.nagoya-u.ac.jp *2 愛知県がんセンター消化器外科 *3 国立情報学研究所医療ビッグデータ 研究センター 我々も腹腔鏡下胃切除術の手術工程を解析す るために,腹腔鏡映像から手術を行っている 解剖学的な領域を認識する手法を検討してき た[6].腹腔鏡映像から解剖学的な領域を認 識することができれば,術野に対応した手術 支援画像を提示することができると考えられ る.そこで本稿では,腹腔鏡映像の認識手法 を利用した手術ナビゲーションの検討を行っ たので報告する.

2. 提案手法

胃がんに対する腹腔鏡下胃切除術では,胃 を切除する前に胃の周囲の血管を順に処理す る.血管構造は個人毎に異なるため,手術箇 所に応じた血管の分岐構造の情報を提示する ことができれば手術ナビゲーションとなる. そこで提案手法では,腹腔鏡映像から手術を 行っている箇所を認識し,手術箇所の血管分 岐構造を確認できる仮想腹腔鏡画像を提示す る.

手術箇所の認識には深層学習を用いた腹腔

鏡下胃切除術の手術工程解析のための腹腔鏡 映像解析手法を用いる [6]. この手法は,腹 腔鏡映像の各フレームを DenseNet により,(1) 左胃大網動静脈付近のシーン,(2)右胃大網 動静脈付近のシーン,(3)右胃動脈付近のシ ーン,(4)左胃動静脈付近のシーン,(5)腹腔 全体を観察しているシーン,(6)トロッカー 内部のシーン,(7)腹腔鏡が体外に存在して いるシーン,の7つのシーンに分類する手法 である.

仮想腹腔鏡画像は術前に撮影した造影 CT 画像から生成する.CT 画像から血管とその 周囲の臓器を抽出し,仮想化内視鏡システム を用いてそれぞれの手術箇所に対応する仮想 腹腔鏡画像をあらかじめ生成しておく.

手術ナビゲーションは,腹腔鏡からリアル タイムに映像を入力して,DenseNetにより腹 腔鏡映像から現在手術を行っている箇所を認 識し,認識結果に対応した仮想腹腔鏡画像を 手術支援画像として提示することで行う.

3. 結果と考察

提案手法を実際の胃がんに対する腹腔鏡下 胃切除術において使用した.手術室において 提案手法を用いて手術ナビゲーションを行っ ている様子を図1に示す.図には手術ナビゲ ーションシステムの画面を示しており,腹腔 鏡から取り込んだ腹腔鏡映像とその認識結果 ならびに仮想腹腔鏡画像が表示されている.

外科医は,システムが提示する仮想腹腔鏡画 像を確認することで,現在の手術箇所の血管 処理に必要な解剖学的構造の情報を得ること が可能であった.

提案手法では,現在の手術箇所を認識して 手術箇所に応じた解剖学的構造情報を提示で きることから手術支援として役立つと考えら れる.さらに今後は,腹腔鏡映像から手術箇 所だけではなく,手術の様々な状況を認識理 解できるようになれば,手術状況に応じた手 術支援情報の提示が可能になり,より高度な 手術ナビゲーションが実現できると考える.

4. むすび

本稿では,腹腔鏡映像からの手術箇所の認 識による手術ナビゲーションの検討について 述べた.提案手法を実際の腹腔鏡下胃切除術 において使用し,手術ナビゲーションが可能 であることを確認した.今後の課題としては, より多くの症例での検討や,様々なシーン認 識手法を用いた手術ナビゲーションの検討な どが挙げられる.

謝辞

日頃ご討論いただく名古屋大学森研究室諸 氏に感謝する.本研究の一部は JST CREST JPMJCR20D5, JSPS 科研費 17H00867, 21k19898, JSPS 二国間交流事業によった.

利益相反の有無

なし

文 献

- Bernhardt S, Nicolau SA, Soler L et al: The status of augmented reality in laparoscopic surgery as of 2016. Med Image Anal 37: 66-90, 2017
- [2] Hayashi Y, Misawa K, Oda M et al: Clinical application of a surgical navigation system based on virtual laparoscopy in laparoscopic gastrectomy for gastric cancer. Int J CARS 11: 827-836, 2016
- [3] Kitaguchi D, Takeshita N, Matsuzaki H et al: Computer-assisted real-time automatic



図1 手術ナビゲーションの様子

prostate segmentation during TaTME: a single-center feasibility study. Surg Endosc **79**: 88-94, 2020

- [4] Tokuyasu T, Iwashita Y, Matsunobu Y et al: Development of an artificial intelligence system using deep learning to indicate anatomical landmarks during laparoscopic cholecystectomy. Surg Endosc 35: 1651-1658, 2021
- [5] Yamazaki Y, Kanaji S, Matsuda T et al. Automated surgical instrument detection

from laparoscopic gastrectomy video images using an open source convolutional neural network platform. J Am Coll Surg **230**: 725-732.e1, 2020

[6] 林雄一郎、杉野貴明、小田昌宏他: ディープラーニングを用いた腹腔鏡映像からの腹腔鏡下胃切除術の手術工程解析の検討.日本コンピュータ外科学会誌 20:368-369,2018

A study on surgical navigation in laparoscopic surgery

by recognition of surgical area from laparoscopic videos

Yuichiro HAYASHI^{*1}, Kazunari MISAWA^{*2}, Kensaku MORI^{*1,3}

*1 Graduate School of Informatics, Nagoya University

*2 Department of Gastroenterological Surgery, Aichi Cancer Center Hospital

*3 Research Center for Medical Bigdata, National Institute of Informatics

In this paper, we describe a study on surgical navigation in laparoscopic surgery by recognition of surgical area from laparoscopic videos. In resent year, there are many studies on analysis of laparoscopic videos. We propose a method for performing surgical navigation by recognizing surgical area from laparoscopic video during surgery. The proposed method classifies operative field views from laparoscopy based on surgical area using convolutional neural networks. Using the recognition results, virtual laparoscopic views generated from preoperative CT images are displayed as surgical assistance image. We performed surgical navigation using the proposed method during laparoscopic gastrectomy. The proposed method presented the 3D virtual laparoscopic images corresponding to operative field during surgery.

Key words: Laparoscopic surgery, Deep learning, Surgical navigation system, Stomach, Surgical phase analysis

VR Organ Puzzle: A Virtual Reality Application for the Education of Human Anatomy

Siqi LI^{*1}, Yuichiro HAYASHI^{*1}, Michitaka FUJIWARA^{*2},

Masahiro ODA^{*3,1}, Kensaku MORI^{*1,4,5}

Abstract

We developed a VR organ puzzle application for the education of human anatomy towards the public. Human body is a complicated system and consists of numerous organs. Knowing human anatomy is essential for elementary education and science popularization. Until now, the education and popularization are mainly implemented by textbooks and educational materials. A more immersive and intuitive way is necessary. Modern virtual reality (VR) technology based on head mounted display (HMD) can provide an immersive 3D environment and an intuitive interactive method, which has already shown the potential in educational field, like medical and military training. In this paper, we developed a virtual reality application, VR Organ Puzzle, to help with the education of human anatomy. Some abdominal organs are modeled from the segmentation result of CT volumes and rendered in the VR environment. Users can grasp and rotate organ models by controllers in an intuitive way. By returning organ models to proper positions and orientations, users can learn correct human anatomy knowledge.

Keywords : virtual reality, anatomy education, medical education, public education

1. Introduction

The human body is a complex and sophisticated machine. Learning anatomy is not only fundamental to medical students, but also important to the public. Patients' lack of anatomy knowledge damages the efficiency of a consultation with doctor as well as the patient compliance [1]. A more immersive and intuitive method will help with the anatomy education for the public. Meanwhile, since head mounted display (HMD) based virtual reality (VR) becomes accessible to the wider public from 2013, it has been increasingly attractive to researchers and developers in training and education area. HMD based VR can provide a fully immersive experience in a virtual environment and an intuitive interaction

*1 Department of Intelligent Systems, Graduate School of Informatics,

Nagoya University [Furo-cho, Chikusa-ward, Nagoya-City, 464-8601, Japan] e-mail: siqili@mori.m.is.nagoya-u.ac.jp

^{*2} Department of Medical Devices, Nagoya University Hospital, Nagoya University

^{*3} Information Strategy Office, Information and Communications, Nagoya University

^{*4} Information Technology Center, Nagoya University

^{*5} Research Center for Medical Bigdata, National Institute of Informatics

method with an extremely low learning cost. HMD based VR has been proven helpful with the acquisition of cognitive skills related to remembering and understanding spatial and visual information and knowledge [2], which exactly coincides with the requirements of anatomy education. Therefore, in this work, we developed a VR organ puzzle application to help with the anatomy education to the public. Several abdominal organs are rendered as 3D models in a virtual environment. Users can interact with organs with their hands in an intuitive way. By returning shuffled organs to correct positions and orientations, users can learn correct anatomy knowledge.

2. Methods

Before the details, we will give a quick interpretation of our application. The user is equipped with HMD and controllers. A virtual operation room is rendered. In the middle of the room, semi-transparent hint organ models are placed at the correct positions and orientations as a guidance for the user. Puzzle piece organ models are placed around the middle. The user needs to return puzzle piece organ models to correct positions and orientations hinted by hint organ models. The overview is shown in Fig. 1.



Valve INDEX controller

Fig. 1 The overview of our application. The left side is the overview in reality. The right side is the virtual view.



Fig. 2 Adopted VR devices.

2.1 Development environment

We adopted Unity and SteamVR plugin for fast development. On the hardware side, we adopted VIVE Pro Eye (HTC Corporation, Taiwan) HMD and Valve INDEX (Valve Corporation, US) controllers, combined with SteamVR Base Station 2.0 (Valve Corporation, US) units. Besides premium visual fidelity, the VIVE Pro Eye HMD can also provide precise eye tracking, which can be used for data analysis after further development. The Valve INDEX controller allows users to interact with objects in the virtual scene in a more intuitive way without having to always hold on to the controller. SteamVR Base Station 2.0 uses the SteamVR tracking system to provide extreme accuracy and low latency to enhance the experience and reduce VR sickness. Adopted devices are shown in Fig. 2.

2.2 System overview

1) Virtual scene

The virtual scene is set in an operation room. A start button is set on a table to start the application. A timer is set on another table to show remaining time. The scene is shown in Fig. 3.



Fig. 3 Virtual scene.



Fig. 4 A liver model. From the top left to the bottom right are visual mesh, convex collider mesh, nameplate, shaded full model.

2) Puzzle piece organ models

A puzzle piece organ model consists of three parts, visual mesh, convex collider mesh, and nameplate. Fine visual mesh segmented from the real CT volumes is used to render the appearance of the organ for a better visual experience. Rough convex collider mesh generated from the visual mesh is used to detect the grab behavior and not rendered in the scene, which can accelerate the collision detection process. When a controller is inside an organ's convex collider mesh, the user can grab the organ and the organ will move and rotate with the movement of the user. Nameplate shows organ's name and always faces to the position of the user's head. The user can identify different organs and learn anatomy knowledge in this way. For example, a puzzle piece organ model of liver is shown in Fig. 4.

3) Hint organ models

To help users learn correct anatomy, organ visual meshes are rendered semi-transparently at correct positions and orientations, which we call hint organ models. When the user holds an organ more than a given time, the corresponding hint organ model will glow and guide the user. Hint organ models and the glowing effect are shown in Fig. 5.

4) Judgement

The correct positions and orientations are pre-determined for each organ. When a puzzle piece organ model's distance between the correct position and current position and angle between the correct orientation and current orientation are under thresholds at the same time, the model will be judged successfully returned. The returned model will be fixed to the correct position and orientation and unable to be grabbed again. If all puzzle piece organ models are successfully returned in a given time, the user will be considered to have successfully completed the application.



Fig. 5 Semi-transparent hint organ models for the guidance. The hint organ model of liver is glowing.



Fig. 6 Visual effects. From left to right are visual effects for a successful return of a single puzzle piece, a successful return of all puzzle pieces in time, and time up.

5) Visual effects

To enhance the feedback, a firework-like visual effect will be played after each successful return at the correct organ position. After all organs are successfully returned in time, a larger firework-like visual effect and a success text will be shown. Or a dust-like visual effect and a failure text will be shown. All these visual effects are shown in Fig. 6.

3. Results

We implemented our method with several abdominal organs. Liver, stomach, spleen, pancreas, kidney, gallbladder, arteries, veins, and portal vein are included in the implementation. All correct positions and orientations of organs are pre-determined by their relative spatial relations obtained from CT volume.

Once the user is equipped with the HMD and controllers, the application will start. The user will feel like standing in a virtual operation room. After the start button is pressed, puzzle piece organ models and semi-transparent hint organ models will be rendered in the middle of the room. The user needs to recognize each organ and return it to the correct position and orientation by grabbing and moving it. If all organs are returned in time, the application is completed. A full process is shown in Fig. 7.

4. Conclusion

We developed a VR organ puzzle application for the anatomy education towards the public. Organ models with nameplates are rendered in the virtual environment. Users can grab and move them in an intuitive way and return them to the correct positions and orientations following the guidance. In this way, users can learn anatomy knowledge and develop an interest in anatomy. For future, we will add more organ models and do more evaluation work.



Fig. 7 The full process.

Competing interests

None.

Acknowledgement

This work was supported by JST CREST Grant Number JPMJCR20D5, Japan.

References

- Taylor A M, Diggle P, and Wessels Q: What do the public know about anatomy? Anatomy education to the public and the implications. Anatomical Sciences Education 11: 117–123, 2018
- Jensen L, Konradsen F: A review of the use of virtual reality head-mounted displays in education and training. Educ Inf Technol 23: 1515–1529, 2018

複数術野視点映像を用いた機械学習による

外科手術工程の自動分類

藤井 亮輔*1 梶田 大樹*2 高詰 佳史*3 青木 義満*1

要旨

撮影による手術映像の記録は手術技能の共有に効果的である一方,それらは長時間に渡る点,また管理の コストや労力の点からあまり実用されない現状にある.この問題の解決手段として機械学習を用いた手術 映像の工程分類による自動要約が挙げられるが,形成外科手術においては空間的な撮影難度が原因で行わ れていなかった.本研究ではマルチカメラ搭載型無影灯を用いた形成外科手術野の撮影による動画データ セットの作成を行った上で,それらを用いた形成外科手術の工程分類を行う際の多視点映像利用の有効性 の検証を行う.

キーワード:機械学習,医用画像処理

1. はじめに

動画の自動要約技術は長時間に渡る手術映 像においてそれらの管理コストや人的労力を 減らす上で有効な手段である[1].しかしながら 手術映像の撮影は内視鏡手術や顕微鏡手術な どのモニター表示が容易な手術では普及して いる一方で,形成外科手術においては術者の頭

*1 慶應義塾大学理工学研究科 総合デザ イン工学専攻 電気電子工学専修

[〒223-8522 神奈川県横浜市港北区日吉 3-14-1]

e-mail: rfujii@aoki-medialab.jp *2 慶應義塾大学医学部 形成外科

*3 慶應義塾大学医学部 解剖学教室

や体に遮られて質の高い術野の撮影が難しく 広く普及していない。

そこで本研究では術野を照らす無影灯にマ ルチカメラを搭載して多視点撮影された形成 外科手術映像から, 3D ResNet[4]と呼ばれる畳 み込みニューラルネットワーク(CNN)による手 術工程の自動認識を行うためのデータセット を構築し,それらを用いた認識精度の評価を 行う.

2. データセット構築

本研究で用いるデータセットは形成外科手 術12件を対象とし、マルチカメラ搭載型無影 灯を用いて5視点から撮影した映像を元にす る.

その後にラベルアノテーションソフト[3]を 用いて各映像に対して形成外科手術における 主要工程 anesthesia, closure, design, disinfection,

dressing, hemostasis, incision, dissection, irrigation

の9種類のクラスラベル付与を行う.

図1に各クラスの例を示す.





(a)design





(c)hemostasis(d)incision図1 手術工程クラスの画像例

ラベル毎に分割された映像は、施術により大 きな時間差があるため、加えて 10 秒毎に短く 分割した.

表1に各ラベルの動画数を示す.

Class	Number of labeled movies
anesthesia	82
closure	12736
design	371
disinfection	12
dressing	22
hemostasis	151
incision	383
dissection	603
irrigation	32

表1 各ラベルの動画本数

表1より, disinfection, dressing, irrigation クラ スのデータが他クラスと比較して少なく, closure クラスが過多である事がわかる.

そこで本研究においては closure クラスの動画 を 1/20 に抽出し, disinfection, dressing, irrigation クラスは使用しなかった. 最終的に用いる各ラ ベルのデータを表 2 に示す.

実験・考察

本実験では1視点入力(1channel),5視点入 力(5channel)に加え Hachiuma ら[2]の提案した5

表2 各ラベルの動画本数

Class	Number of labeled movies
anesthesia	82
closure	634
design	371
hemostasis	151
incision	383
dissection	603

視点の各映像の特徴に全視点の特徴を足し合わせ他視点情報を加味する事で各フレームにおいて最も術野を視認できる映像を選択し再構成する手法を用いた映像の入力(reconstruct)での比較を通して多視点映像利用の有効性の検証を行う.検証は各ラベルのF1値の比較により行う(式1).また評価の際には手術動画毎で学習データ,検証データ,テストデータに7:2:1の比率でデータをふり分け,完全未知な動画に対する評価を行った.

$$F1score = \frac{2Precision \times Recall}{Precision + Recall} \tag{1}$$

Ichannel, 5channel, reconstruct の各実験下で検 証を行い, 各ラベルの F1 値の平均をとる事で 平均認識精度として算出した. 表 3 にその結 果を示す.

Ichannel 入力での平均認識精度は 11.4%であった. 5channel 入力での平均認識精度は
 12.5%であり、1channel 入力と比較しての精度向上はわずか 1.1%であった. Reconstruct 入力での平均入力精度は 16.6%であり、1channel 入力の場合よりも 5.2%の精度向上が見られた.

1channel 入力と 5channel 入力での精度変化 が小さかったのは、各映像内に術野を遮るオ クルージョンが存在しており、映像切り替え としての機能が果たされていないためだと考 えられる.reconstruct 入力においては他視点情 報を加味した映像切り替えが果たされ、入力 データとしての質が他より優れている事がう かがえる.

	F1-score (%)			
Class	1channel	5channel	reconstruct	
anesthesia	1.5	1.6	5.4	
closure	15.2	16.2	22.1	
design	22.1	22.5	27.1	
hemostasis	1.8	2.4	2.6	
incision	16.6	17.8	23.3	
peeling	11.2	14.3	19.1	
micro ave	11.4	12.5	16.5	

表3 各実験下における評価結果

また,全体を通して6クラスでの認識精度期 待値である16.6%を下回る結果となった. これは形成外科手術が一様ではなくそれぞれ 多様な手術工程を踏んでいるため,各手術で 行われる施術のデータ数のばらつきが大きす ぎることが原因に挙げられる.

4. まとめ

本研究は形成外科手術を対象にしたマルチ カメラ搭載型無影灯での撮影によるデータセ ットの作成および,それらを用いた手術工程 分類における複数視点入力の有効性の検証を 目指した.

完全未知の手術に対する1視点入力での平 均認識精度が11.4%,5視点入力での精度が 12.5%,他視点情報を加味した再構成映像入力 での精度が16.5%であった.複数視点を用いた 入力それぞれにおいて,1視点入力と比較した 精度向上が見られたため,外科手術の工程分 類において複数視点の入力は有効な手法であ ると証明された.

今後の課題としては多様な工程を踏む外科 手術のそれぞれに対応できる頑強性を備えた モデルを構築することが必要である.未知の 手術においても高い工程分類精度を達成でき れば手術映像に対する自動要約技術として活 用できると考えられる.

謝辞

本研究の遂行にあたり,貴重なご協力やご 助言を賜りました多くの方々に深謝致しま す. また、本研究は構成労働科学研究費補助金政 策科学総合研究事業(臨床研究等 ICT 基盤構 築・人工知能実装研究事業)20AC1004 の助成, 及び総務省 SCOPE(受付番号 201603003)の委託 を受けたものです.

利益相反

なし

文 献

[1] Ting Y, Tao M, Yong R : Highlight Detection with Pairwise Deep Ranking for First-Person Video Summarization. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016

[2] Hachiuma R, Shimizu T, Saito H et al :Deep Selection: A Fully Supervised Camera Selection Network for Surgery Recordings. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2020

[3] マックス・プランク心理言語学研究所, ELAN-Linguistic Annotator,

https://www.mpi.nl/corpus/html/elan/ (2021 年 6 月 9 日閲覧)

[4] Hara K, Kataoka H, and Satoh Y: Learning spatio-temporal features with 3d residual networks for action recognition. In Proceedings of the IEEE International Conference on Computer Vision Workshops, pp. 3154-3160,2017

Verification of Auto Classification for Surgery Process

by Machine Learning Using Multi-View Videos

Ryosuke Fujii^{*1}, Hiroki Kajita^{*2}, Yoshifumi Takatsume^{*3}, Yoshimitsu Aoki^{*1}

*1. Department of Integrated Design Engineering, Graduate School of Science and Technology, Keio University

*2. Department of Plastic and Reconstructive Surgery, School of Medicine, Keio University

*3. Department of Anatomy, School of Medicine, Keio University

Recording surgery is effective for sharing operating techniques. However, it is hard to use recorded videos practically because of their length and management cost. Auto summarization by machine learning is major approach to solve this, but in plastic surgery, it is almost impossible to capture the surgical field because of occlusion by the surgeon's body and head. In this paper, we create supervised video data of plastic surgery process by using a surgical lamp system which installed multiple cameras and verify the effectiveness of using multi-view videos in process classification.

Key words: Machine Learning, Medical Image Processing

腹腔鏡の視野外における術具誤接触防止システムの開発

岸 知哉*1 中口 俊哉*2 長尾 俊也*3 駒村 賢三*4 西澤 裕吏*5

要旨

腹腔鏡下手術は視野が腹腔鏡映像のみのため視野が狭い.また,特にロボット支援腹腔鏡下手術 の場合は,腹腔鏡の視野外において術具と近接組織との誤接触が発生し医療事故に繋がる.この 腹腔鏡視野外での術具誤接触を防止するため,本研究では術具先端と近接組織間の距離を計測し, 危険と判断されれば術者に音声や画面表示などで警告を行うシステムを提案する.このシステム を実現するため先端に距離センサと小型カメラを装着した専用トロッカーを考案した.まず距離 センサを用いてトロッカー・組織間の距離を計測し,小型カメラで撮影した画像よりトロッカー から突出する術具の長さを計測する.これらの距離の差分から術具先端と組織間の距離を算出す る.また,警告対象となるのは腹腔鏡視野外であるため,術具が腹腔鏡の視野外に位置している か判定する手法を提案した.ドライボックス環境にて提案システムの評価実験を行った結果,有 用性が示唆された.

キーワード:腹腔鏡,誤接触防止,警告

1. はじめに

近年の医療では患者の QOL が重視されて おり,低侵襲手術分野の一種である腹腔鏡下 手術は手術創および術後の痛みが小さく,早 期退院・社会復帰が可能であるなど従来の開 腹手術に比べて患者の負担が少ない利点があ ることから普及が進んでいる[1].一方で,視 野が腹腔鏡映像のみのため術者にとっては視 野が狭い[2].また,特にロボット支援腹腔鏡

*1 千葉大学大学院融合理工学府基幹工 学専攻医工学コース

[〒263-8522 千葉市稲毛区弥生町 1-33]

*2 千葉大学フロンティア医工学センタ

- *3 京セラ株式会社東京事業所
- *4 株式会社ソフケン
- *5 国立がん研究センター

下手術の場合は,腹腔鏡の視野外において術 具と近接組織との誤接触が発生し,医療事故 に繋がる[3].このため,腹腔鏡下手術の安全 性の向上が強く求められている.

この解決策の一つとして,腹腔鏡下手術の 利点である低侵襲性を維持したまま,開腹手 術のような広い視野の確保が望まれている. 現在の腹腔鏡下手術では,図1のように単視 点の映像をモニターに表示する術式が主流で あり,広い視野を得るためには腹腔鏡以外の 新たな視点から腹腔内を観察する必要がある. そこで,鉗子の挿入口であるポートを確保す るトロッカーに,図2に示すような格納展開 の機構を有する小型カメラを取り付けたカメ ラ付きトロッカーの研究開発が進められてい る[4].これにより,腹腔鏡下手術の利点であ る低侵襲性を維持したまま,複数視点の腹腔 内映像を得ることが可能となる.


図1 腹腔鏡下手術



(a) 格納した状態(b) 展開した状態図 2 カメラ付きトロッカー

本研究では,腹腔鏡下手術において術具先端 と近接組織間の距離を計測し,危険と判断さ れれば術者に音声や画面表示などで警告を行 うシステムを提案する[5].

2. 提案手法

術具と近接組織間の距離を計測するために 本研究では、カメラ付きトロッカーの先端に 新たに距離センサを装着する.距離センサに よりトロッカー・組織間の距離D_Tを計測し、 小型カメラにより撮影された画像よりトロッ カーから突出する術具の長さD_Fを計測する. これらの距離の差分から術具先端と組織間の 距離D_Nを算出する.提案システムの構成を図 3に示す.



D_T:トロッカーから組織までの距離
 D_F:トロッカーから突出する鉗子長
 D_N: 鉗子先端から組織までの距離

図 3 鉗子先端から組織までの距離を算出する 原理

2.1 距離D_Tの計測

本研究では、距離センサに ToF センサを使 用した[6]. ToF は赤外光を使って距離を計測 する技術である. ToF センサのトランスミッ タから照射された赤外光が物体表面に衝突し、 その反射光が ToF センサのレシーバに到達す る. 光速と赤外光が照射されてから受光する までの時間から、センサ・物体表面間の距離 を計測する.

図4に示すように、本研究では、3Dプリン タで作成した固定具にToFセンサを取り付け、 カメラ付きトロッカーに装着した.また、ToF センサと接続したマイコン(Arduino Uno)によ って計測したセンサ値を取得する.図5にToF センサとマイコンを接続した様子を示す.

2.2 突出鉗子長D_Fの計測

トロッカーから突出する鉗子長D_Fを計測 するために,鉗子の胴体(シャフト)部に二色 の色付けを行ったラベルを貼付することで目 盛りを設けた.トロッカーカメラで撮影した 画像から目盛りの帯を検出しトラッキングす る.帯の間隔は 10mm に設定しているため, 検出した帯から鉗子先端部までの距離が把握 できる.画像端の基部座標からy座標最大の 目盛り検出点の座標までの距離を三次元距離



図4 ToF センサの装着



図 5 ToF センサと Arduino Uno

AX 1	ノーシャのと向	戶一內 世
ラベル	平均距離	最小距離
ラベル1	91.4	11.3
ラベル2	89.0	32.3
ラベル3	88.6	24.7
ラベル4	82.0	19.1

表1 ラベルの色間距離

に変換し、目盛り番号を足し合わせることに よって、トロッカーから突出する鉗子長 D_F を 計測する.

2.2.1 二色の最適な組み合わせ

目盛りラベルに用いる二色には,鉗子先端 の金属部分,生体組織の赤系色と識別性の良 い色を選択する.

そこで以下に示す四組のラベル色の組み合わ せを検討した.トロッカーカメラで撮影した RGB 値を HSV 色空間に変換し,H(色相)S(彩 度)の2次元面における距離を算出した.目盛 りラベルの2領域,鉗子先端金属部,背景(生 体組織)部の各平均色間の差の最小値と平均 を算出したところ表1のようになった.

この結果から、緑・黄色を用いたラベルが最 適な識別性を示したため、以後この2色ラベ ルを用いることとする.



図7 鉗子中心線上のH 微分値

(ラベル 1)
背景:黒,帯:黄色
(ラベル 2)
背景:黒,帯:緑
(ラベル 3)
背景:緑,帯:黒
(ラベル 4)
背景:緑,帯:黄色

2.2.2 目盛り検出

目盛りラベルを貼付した鉗子の中心線上の H値を取得し,その微分値を算出する.それ ぞれ図 6,7 のようになる.

図7のように,色の境界でH値の微分値が極 大値・極小値を取る.極大値と極小値の座標 の平均値を目盛りの帯の検出点とした.

2.2.3 目盛り検出点のトラッキング

鉗子はトロッカーに挿入・抜去の動作がな される.そのため突出鉗子長D_Fは常に変化す る.従って,帯の検出点をトラッキングし新 たな検出点に正確に目盛り番号を割り当てる 必要がある.

前後フレーム間で検出点の座標間の距離が最 小である組み合わせを対応付け,その後目盛 り番号の引継ぎを行う.引き継がれた目盛り 番号における最大番号を取得し,この番号と 検出点の個数から新たな検出点に目盛り番号 を割り当てる.このトラッキングの流れを図 8に示す.

2.2.4 二次元・三次元間の距離変換

画像の基部座標とy座標が最大の検出点ま での距離L(pixel)を三次元距離に変換する.距 離Lを図9に示す.

鉗子をトロッカーに挿入した状態の横から見 た模式図を図 10 に示す. ピンホールカメラモ デルを仮定し, ピンホール・トロッカー先端



図8 トラッキングの流れ



図10 トロッカー・鉗子の模式図 間の距離をC,画像の範囲をa(a = 1),画像下 端から対象までの画像上の距離をL,画像上 端から対象までの画像上の距離をL',焦点距 離をf,世界座標系における対象の位置を (Y,Z),トロッカー・目盛り検出点の距離をZ' とする.L'の正規化スケールをbとすると,Lは 式(1)のように表せる.また,Zは式(2)のよう に定義される.

$$L' = b(a - L)$$

= b(1 - L) (1)
$$Z = Z' + C$$
 (2)

図 10 における赤色三角形と青色三角形にお いて相似が成り立つので, L'は式(3)のように 定義する.

$$L' = \frac{fY}{Z} \tag{3}$$

(1),(2),(3)式から, Z'は式(4)のように定義される.

$$Z' = \frac{fY}{b(1-L)} - C \tag{4}$$

次に,式(4)中の未知数である*b,C*を算出し, その結果から得られた変換式の精度実験を行



図 11 トロッカー・目盛り検出点の距離Z'の 実測値と予測値

った.

2.5.1 変換式の精度実験

式(1)の透視変換モデルの未知数b,Cを決定 するためキャリブレーションを行った.トロ ッカー・目盛り検出点の距離Z'を 5mm から 100mm まで,5mm 間隔で変えたときの距離Lをそれぞれ求めた.実測した 21 組の D_F とLか ら未知数b,Cを決定したところ,b =349.1,c = 9.075が得られた.決定係数は $r^2 =$ 0.9986となり,モデルの確かさが示された. 実測値とモデル近似の様子を図 11 に示す.

3. 実験

3.1 ピポット運動評価実験

腹壁に設置されたトロッカーは鉗子操作に 伴ってピポット運動を行う.このとき、トロ ッカーカメラ映像の背景は変化し続けており このような動きのある状況における突出鉗子 長D_Fの算出の安定性を評価する.

本研究では腹腔鏡下手術用ドライボックスの ポートにトロッカーを設置する.その後,鉗 子を挿入し突出鉗子長*D_F*を 20mm に固定し, トロッカーをピポット運動させる.ピポット 運動は体軸に直交する面上を角度 10°~40° の範囲で往復運動させた.往復の周期を 1.0,2.0,3.0 秒間の三段階とし,各試行は 10 秒





(a) ピポット運動の図
 (b) 実験の様子
 図 12 ピポット運動評価の実験環境

間継続した.実験環境は図12に示す.

評価のため、各試行における突出鉗子長D_Fの 標準偏差を算出する. 鉗子はトロッカーに固 定されているため、計測が安定していればD_F は一定であり,標準偏差は0に近い値となる. 表2に結果を示す.

各試行において,標準偏差が非常に小さく, 変動は 0.5%以下であることが示された. つま りカメラ画像の背景がどのような速度で変化 しても、安定した突出鉗子長D_Fの算出が可能 であることが示唆された.

3.2 警告音発生実験

鉗子先端から組織までの距離 D_N が、 $D_N \leq$ 60mm, $D_N \leq 40mm$, $D_N \leq 20mm$ の三段階に 応じて警告音を発生するように設計した.ま た,距離D_Nが警告範囲内の場合,D_Nの値を赤 色に変化させた.

その結果、段階に応じて異なる警告音を発生 が可能であることを確認した. 警告音発生の 様子を図 13 に示す.

4 まとめ

本研究では,腹腔鏡の視野外で術具が近接 組織に誤接触することによって発生する医療 事故を防ぐことを目的として, センサ計測と 画像解析を統合し, 警告を行うシステムを提 案した. ピポット運動評価実験により突出鉗 子長D_F安定した算出が可能であることを確

-	2 突出蚶于長 D_F の標準偏差(単位 mn					
			[5	sec/1往復	[]	
			1	2	3	
		10°	0.061	0.014	0.096	
	角度	20°	0.000	0.085	0.073	
		30°	0.075	0.068	0.065	
		1.00	0.000			

表 2	突出鉗子長D	の標準偏差(単(立 如)
~ -				,



認した.また、鉗子先端・組織間の距離D_Nに 応じて三段階の警告音の発生を実装した. 今後の課題として,警告対象となる術具は腹 腔鏡の視野外に位置している場合であるため, 術具が警告対象であるか判定を行うシステム の開発に取り組む.

利益相反の有無

なし

文 献

- [1] A. Darzi and S. Mackay, "Recent advances in minimal access surgery," Brit. Med. J., vol. 324, p. 31-34, 2002.
- [2] 永川裕一,土田明彦:腹腔鏡下・ロ ボット支援下膵切除術の現状と将 来,日本臨床外科学会雑誌, Volume 81, Issue 9, p.1683-1695,2020
- [3] Tsuyoshi Kondo, Tadashi Shionoiri, Isao Shiozawa: A case elucidating a new mechanism of thermal injury in laparoscopic surgery: spread of monopolar electric current in collecting bag, 日産婦内視鏡学会, Volume 27, Issue 2,2011
- [4] Okubo T, Nakaguchi T, Hayashi H, Tsumura N: Abdominal View Expansion by Retractable Camera, Journal of Signal Processing, Volume 15, Issue 4: 311-314, 2011
- [5] Ryoichi Nakamura, Yume Honda, Ikuma Sato: Development and Evaluation of "Distance Sensation" Indicator for Surgical Navigation System,日本コンピュータ外科学会 誌, Volume 14, Issue 2, p.39-45,2012
- 小枝正直,上田悦子,中村恭之: [6] OpenCV による画像処理入門 講 談社, 2017,p 201-202

Development of a system for preventing false contact with surgical

instruments outside the field of view of a laparoscope

Tomoya KISHI*1, Toshiya NAKAGUCHI*2, Toshiya NAGAO*3, Kenzo KOMAMURA*4, Yuji NISHIZAWA*5

*1 School of Engineering, Chiba University

*2 Center of Frontier Medical Engineering, Chiba University

*3 Kyocera Corporation Tokyo Office

*4 Sofuken Co.,Ltd

*5 National Cancer Center East Hospital

Laparoscopic surgery has a narrow field of view because the field of view is only laparoscopic images. Further, especially in the case of robot-assisted laparoscopic surgery, erroneous contact between the surgical instrument and nearby tissues occurs outside the field of view of the laparoscope, leading to a medical accident. In order to prevent accidental contact of the surgical instrument outside the field of view of the laparoscope, in this study, the distance between the tip of the surgical instrument and the nearby tissue is measured, and if it is judged to be dangerous, the operator is warned by voice or screen display. Propose a system. In order to realize this system, we devised a dedicated trocar equipped with a distance sensor and a small camera at the tip. First, the distance between the trocar is measured using a distance sensor, and the length of the surgical tool protruding from the trocar is measured from the image taken by the small camera. The distance between the tip of the surgical instrument and the field of view of the laparoscope, we proposed a method to determine whether the surgical instrument is located outside the field of view of the laparoscope. As a result of conducting an evaluation experiment of the proposed system in a dry box environment, its usefulness was suggested.

Key words: Laparoscope, prevention of false contact, warning

第 40 回日本医用画像工学会大会(JAMIT2021) 大会役員委員会名簿

五十音順·敬称略

大会役員				
大会長				
	陣崎	雅弘	慶應義塾大学	
大会長補佐				
	橋本	正弘	慶應義塾大学	
業務委員長				
	宮崎	靖	富士フイルムメディカル	
			プログラム委員会	
	中田	典生	東京慈恵会医科大学	
委員		<u> </u>		
	有村	秀孝	九州大学	
	伊藤	聡志	宇都宮大学	
	禰富	仁	法政大学	
	上村	幸司	国立循環器病研究センター	
	内山	良一	能本大学	
	大城		大阪大学	
	小田			
	小国	<u> 百</u> 二 二 二 二 二 二 二 二 二 二 二 二 二 二 二 二 二 二	□ □□□注入了	
		高又 佳樹		
		安去	《····································	
		<u> </u>	◎ 友州上未八子 自津観作正	
	北村	<u>王</u> 印 出 込	上版十岁	
	<u></u> 木材	<u> </u>		
		11日本		
		 	内 (八)	
		新田 	□ 示民兀响伴子仅他入子阮入子 □ 言報十尚	
	—————————————————————————————————————		□	
		- 咱們 	水示辰山八子 	
		四木		
	上野	此	■	
		野生 まご	十葉入子	
	杉本	<u> </u>	□	
	<u> </u>	想局		
	田田	利思		
		<u></u> 些 怿	□	
	비비		十某大字	
	根本	允貢		
	野村	行弘	十葉大字	
	畑中	裕司	大分大学	
	花崗	昇平		
	羽石	秀昭	千葉大学	
	原	武史	岐阜大学	
	平野	靖		
	本谷	秀堅	名古屋工業大学	
	增谷	佳孝	広島市立大学	
	湯澤	史佳	キヤノンメディカルシステムズ	
	目加日	日慶人	中京大学	
	森	健策	名古屋大学	
	山谷	泰賀	量子科学技術研究開発機構	
	湯浅	哲也	山形大学	

日本医用画像工学会(JAMIT) 役員名簿 (令和3年6月8日現在)

五十音順・敬称略

	氏名	所属	
会長	工藤博幸	筑波大学	
	佐藤 嘉伸	奈良先端科学技術大学院大学	
副会長	原武史	岐阜大学	
	森 健策	名古屋大学	
	伊藤 聡志	字都宮大学	
	北坂 孝幸	愛知丁業大学	
		筑波大学	
	中田典生	東京慈恵会医科大学	
	加中 裕司	大分大学	
理事	花岡 昇平	東京大学	
	道谷 佳孝	広島市立大学	
	宮崎 靖	富士フイルムヘルスケア	
	日加田慶人	中京大学	
	湯澤 史佳	キヤノンメディカルシステムズ	
	安藤裕	ICHO 埼玉メディカルセンター	
	石田 隆行	大阪大学	
	上村 幸司	国立循環器病研究センター	
	大松 広伸	網走刑務所	
	岡田 知久	京都大学	
	尾川 浩一	法政大学	
	小田 昌宏	名古屋大学	
	小尾 高史	東京工業大学	
	掛川 誠	中央システム技研	
	笠井 聡	新潟医療福祉大学	
	川岸 将実	キヤノン	
	河田 佳樹	徳島大学	
	木戸 尚治	大阪大学	
	黒木 嘉典	新村病院、博愛会 HPN 東京	
	陣崎 雅弘	慶應義塾大学	
化镁昌	杉本 直三	京都大学	
八哦貝	高橋 哲彦	群馬県立県民健康科学大学	
	武田 徹	筑波記念病院	
	田中 利恵	金沢大学	
	陳 延偉	立命館大学	
	土居 篤博	富士フイルム	
	縄野 繁	新松戸中央総合病院	
	野崎 太希	聖路加国際病院	
	長谷川純一	中京大学	
	羽石 秀昭	千葉大学	
	平野 靖	山口大学	
	藤田 広志	岐阜大学	
	<u> 細羽 実</u>	京都医療科学大学	
	森 雅樹	札幌厚生病院	
	山谷 泰賀	量子科学技術研究開発機構	
	湯浅 哲也	山形大学	
	吉川 悦次		
監 事	清水 昭伸	東京農工大学	
		局津製作所	