

機械学習による胸部 CT 画像からの COPD 診断支援に関する研究

鈴木 昭弘[†] 川上 敬[†] 木村 徹[‡] 菊池 明泰[†] 大江 亮介[†] 石田 崇[§]北海道科学大学[†] 北海道科学大学大学院[‡] (株)テクノフェイス[§]

1. 緒言

慢性閉塞性肺疾患 (COPD; Chronic Obstructive Pulmonary Disease) は, 2013 年から新たに厚生労働省が計画した「健康日本 21」の対象疾患に加えられた疾患である¹⁾。その病態は進行性で, 不可逆性の気流閉塞を生じる呼吸器疾患である。

COPD の診断はスパイロメトリー検査 (呼吸器検査) が一般的であるが, ある程度病状が進行していないと検出できない。そのため最近では胸部 CT 検査による検出方法が用いられ始めた²⁾。しかしこれは肺胞壁破壊病変がある程度出現後の画像解析であり, 病変が出現するまでのより早期 (以下, COPD 0 期とする) での変化は検査されない。加えて検査機器の充実やマルチモダリティ診断により医用画像や診断情報が増え, 放射線科専門医の読影負担は増す一方である。

そこで本研究では, COPD 0 期を対象として医師の読影負担を軽減し, 診断に有用な情報を系統的に提供するための診断支援システムを深層学習の手法を使って構築することを目指す。

2. COPD 診断支援システム

本研究では胸部 CT 画像から LAA 領域の判別を行い, COPD の病期を推定するプロセスを自動化し診断支援に役立つシステムの開発を目指す。LAA (Low attenuation area) とは肺胞実質領域の破壊により生じた, 拡大した気腔が低吸収領域のことである (Fig.1)。CT 画像において, 正常肺に比べて, COPD の肺気腫病変の部位では肺胞壁が破壊されているため肺比重が低くなっていることから, この比重差をみることで気腫性病変と正常肺の区別が可能である。

本研究での COPD 診断支援システムの概要は以下の通りである。

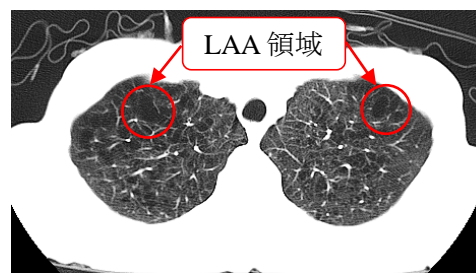


Fig.1 Law attenuation area in the lung

- 1) 全身の CT 画像を入力し, 部位分類システムにより胸部 CT 画像を抽出する。
- 2) 肺が撮像されている CT 画像から, 大動脈弓上縁位置の上肺野画像, 気管分岐部位置の中肺野画像, 横隔膜上縁位置の下肺野の代表画像 3 枚を抽出する。
- 3) 抽出された 3 枚の代表画像に対し, 肺野領域を認識し, 内部の各画素の CT 値を調べ, その分布を出力する。従来手法では LAA 領域を人為的なしきい値で判定しているが, 病気の初期ではこのしきい値に達しない。そこで, より精緻な評価を行う目的で本研究では CT 値分布をそのまま扱う。
- 4) 3) で出力された CT 分布に対して, Goddard 法³⁾という CT 画像における肺の気腫性病変の広がり注目した COPD の視覚的評価法に基づいた評価点を正解ラベルとして SVM により学習を行い, 病期識別器を構築する。
- 5) 病気の進行度合いを測るため, 同一対象者の過去と現在の CT 画像を入力し, 1) から 3) のプロセスを実行し, CT 値分布の差分を出力する。これに対して, 予後の進行予測を行うような SVM を学習させる。

3. COPD 診断支援システムの学習手法

1), 2) は深層畳み込みニューラルネットワーク (DCNN) により行う。この学習には, 予め分類クラスをタグ付けした学習データセットを使用する。このデータセットは全身 CT 画像であり, 全 53 個体の中から学習毎に必要な個体を抽

A Study on Diagnostic Assist Systems of COPD from CT Images by Machine Learning

[†]Hokkaido University of Science

[‡]Graduate School of Hokkaido University of Science

[§]Technoface

出し作成した。の1個体とは1回の撮影で得られる300~500枚の画像群であり、同一人物のCT画像であっても撮影日が異なれば別個体として扱う。学習を行うDCNNは、AlexNet³⁾と呼ばれる全11層のモデルを用いる。

1)の肺野画像抽出プロセスでは既に先行研究⁴⁾で開発した、臓器分類抽出システムを利用している。これは1個体分の画像データの中から、指定した臓器(心臓, 肺, 肝臓, 胃, 脾臓, 腎臓)が撮像されている画像群を分類・抽出する。しかし、胸部や腹部の画像等では1枚の画像に複数の臓器が同時に写り込むため、特定の臓器の存在の有無を識別する必要がある。そこで出力層のユニット数を分類したい臓器の数7とし、それぞれのユニットからの出力値をしきい値により判別することとした。実験には各個体からの学習データに各臓器のタグをつけ学習を行った。

Fig.2は学習データに対する学習曲線を表しており、学習後は99.9%以上の精度で正しく臓器検出ができることが確認できる。学習後の検出器を用いて未知の個体の臓器検出実験を行うと、95.24%の精度で検出が可能となっている。

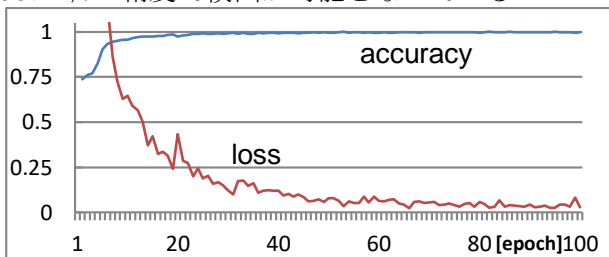


Fig. 2 Learning curves of the extraction system for learning data

2)の代表画像抽出プロセスもAlexNetを用いたDCNNにより学習を行う。学習データは肺野画像に対して、代表画像3種とそれ以外の4クラス分類としてタグ付けを行っている。したがって出力層と最後の全結合層は4ユニットの構成としている。この学習もFig.3に示す通り、99%以上の学習精度が出ており、テスト画像に対して96%程度の精度で検出が可能となっている。

3)のCT値分布の算出は、一般的な画像処理により肺野領域を抽出し、その内部のCT値を読み取り、分布を作成する。Fig.4に示す通り、代表的な肺野画像に対して、体外をマスクで除去したうえでCT値の分布が計算され、出力されていることがわかる。この結果を基にしてGoddard法による病態評価はすでに可能となっている。

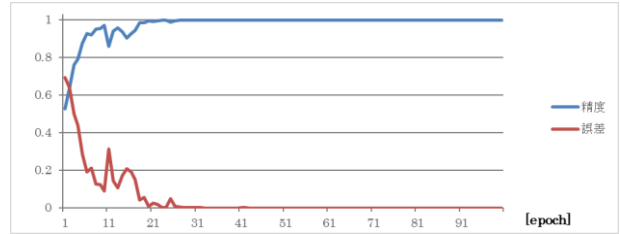


Fig.3 Learning curves of the classification system for three representative lung CT images

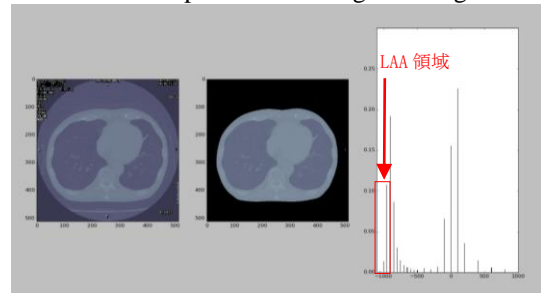


Fig.4 Probability distribution of CT values for representative lung CT images

4)および5)のSVMによる病期識別器については原稿執筆時点で未実装であるため、早急に開発を進めている。

4. 結言

本報告では、COPD 0期を対象とした診断支援システムを深層学習によって構築する手法を提案した。全身CT画像からend-to-endで学習でき、診断支援情報を提示できるシステムとなっている。特にLAA領域を判別する手法よりも精緻な判別が可能ないようにCT値分布を直接扱うことで高い有用性を持つものと考えられる。しかしSVMによりCT分布を認識するプロセスが未実装であるため、この検証を早急に行う予定である。

謝辞

本研究を進めるにあたり、群馬がんセンターから多くの貴重な資料を提供して頂きました。ここに感謝申し上げます。

引用文献

- 1) 日本呼吸器学会 COPD ガイドライン第3版作成委員会 編：COPD(慢性閉塞性肺疾患)診断と治療のためのガイドライン第3版2009年
- 2) Goddard PR, et.al. :computed tomography in pulmonary emphysema. Clin.Radiol 33:379-387, 1982.
- 3) A. Krizhevsky, I. Sutskever, GE. Hinton, "Imagenet classification with deep convolutional neural networks," Adv. Neur. In., vol.25, pp.1106-1114, 2012.
- 4) 川上敬, 橋本大夢, 大江亮介, 丹羽孔明, 菊池明泰：深層学習による医用画像読影支援システムー肺野 CT の異常検出ー, 2017 年度ロボティクス・メカトロニクス講演会論文集, 2P2-E03(1)-(4), 2017.