

# 学位論文内容の要旨

北海道科学大学大学院工学研究科  
工学専攻  
申請者氏名 木村 徹

## 機械学習を用いた胸部 CT 画像からの 肺気腫拡大予測に関する研究

昨今、医療分野、特に医用画像診断領域の分野での人工知能 (Artificial Intelligence :AI) 技術の活用が進んでいる。機械学習や Deep Learning に代表されるような手法を含む、AI 技術の利用はあらゆる分野に大きな影響を及ぼすと予想されており、音声認識、画像認識、自動運転技術等、産業の構造までも大きな変化が期待されている。特に、生命科学および医療と情報工学の融合、すなわち医工連携によって AI の活用の場はゲノム医療、画像診断、診断支援、創薬、看護・介護、医療情報など多岐に広がることが期待されている。その結果、医療画像診断を機械学習の一手法である多層ニューラルネットワークを用いた深層学習の活用によって自動化しようという研究が盛んに行われている。多層ニューラルネットワークの中でも、畳み込みニューラルネットワーク (CNN:Convolutional Neural Network) は画像分類や画像認識などに長けており、たとえば胸部 X 線画像から所見を分類したり、皮膚病変写真から良性・悪性を判別するなど、多くの研究報告がされている。加えてその分類・判別結果は診断医と同等な能力であることも報告されている。

本学位論文は胸部 CT 検査画像を解析することによって、肺気腫から COPD (chronic obstructive pulmonary disease : 慢性閉塞性肺疾患) に進行しそうな者を早期発見することで、COPD の発症を未然に防ごうとするためのものである。特に、肺気腫の早期発見過程において、この医学・医療的な課題に対し、工学的なアプローチによって解決を試みた一連の研究成果をまとめたものである。

COPD とは厚生労働省発表「平成 22 年人口動態統計」において、日本人の死因順位第 9 位の疾病である。総死亡者数も年々増加傾向にあり、日本人 40 歳以上の 8.6% (約 530 万人) が COPD に罹患していると考えられている。COPD は不可逆性病変とされており、一度発症すると完治を期待することは難しく、一般的な治療方法は薬剤を用いての対症的療法しかない現状である。高齢化社会も相まって、このまま COPD の患者数が増え続けることで医療費高騰が懸念されており、厚生労働省は「がん、循環器疾患、糖尿病および COPD」の予防対策・早期発見、特に COPD に関

しては一步進んで「禁煙による発症予防と早期発見, および認知度の向上」を進めている. 認知度向上への対応や COPD 原因物質の特定, 治療法の確立等各分野での研究が進められている中, 重要な問題となるのは, 労作時の息苦しさを主訴として医療機関に受診してきた場合, すでに COPD がかなり進行した状態であることである. これは COPD の病型で日本人に多い肺気腫が広範囲に広がらなければ, 本人の自覚症状が出ないためである. 疫学調査により日本人の 40 歳以上では, 530 万人の COPD 患者数であると推定されているが, このうち病院で実際に COPD 患者であると診断された人は 26 万人しかいない. これはすでに COPD であるにも関わらず, 何ら医療行為を受けていない人が 500 万人以上存在する可能性を示している.

そこで本研究の大きな特徴は, 不可逆性病変である COPD を発症していない, すなわち「異常なし」と診断されている者の胸部 CT 画像をメインに解析することによって, 肺気腫の発生とその拡大進行に着目し経年変化の追跡を行うことで, COPD の早期発見に必要な要素を調査する. 次に, この解析結果から明らかになった事実を利用することで, 肺気腫が拡大し, COPD へ進行しそうなタイプを予測するシステムを構築することとした. 胸部 CT 画像の解析結果から, COPD の発症を未然にかつ早期に発見することができれば「COPD が発症する前や進行を防ぐように警告や生活習慣の指導」を行うことが可能となり COPD に対する予防の観点から大変有益なものとなる.

本論文の構成は以下の通りである. 第 1 章では緒言として研究全体の概要と目的を述べ, 第 2 章では COPD という疾病全般に関する背景として, COPD の患者数やその症状, 病態, 診断法, 病期など COPD という疾病を取り巻く概要についてその詳細を述べた. 特に, COPD は不可逆的な疾患であることから, その早期発見が重要な課題であることが認知されているにもかかわらず, 現状の COPD 画像診断は, 積極的な早期診断を目的にしていけないという問題点を提起している.

第 3 章では, 積極的な COPD の早期発見を目指すための新しい画像診断方法を提案し, そこで必須となる「CT 値解析」に着目する根拠を述べている. 第 4 章では, 現状で CT 検査画像に対し行われている Goddard 法や %LAA 法などの従来の画像診断方法に加え, 早期発見を目的に多角的に CT 値を解析する手法を組み合わせることで得られた情報の有用性を検証した. 特に, ここでは COPD の初期段階とも言える肺気腫の発生を検出することを目的として, 胸部 CT 検査の画像診断結果から「異常なし」すなわち COPD ではないとされた者を対象に, 肺 CT 値の階級度分布の解析および, 肺平均 CT 値の解析を行った. 以上の解析データは「異常なし」と診断されている者の中に, 肺平均 CT 値が 1 年当たり 40HU も低下した者が存在し, さらに詳細な解析を進めると, このタイプの CT 値の階級度分布は, COPD:0 期者 (極軽度の COPD 患者相当) と同様な分布を示していた. このことは今後 COPD へ進行する前段階である「COPD 境界群」の存在を明らかにした. このように, 肺の平均 CT 値データの取得, 肺 CT 値の階級度分布データ及び, それらの経年変化量を取得し解析するような, 本画像診断手法と同様な手法に関する報告例はこれまでにはない. そしてこれらは COPD の初期段階である肺気腫の検出に有用であり, COPD の早期発見の可能性を持った新しい画像診断法となりうると仮定し, 考察を行っている.

第 5 章では, 前章で提案した対象者の CT 値の階級度分布及び, 肺平均 CT 値それぞれの 1 年当たりの変化量を追跡する画像診断法を組込み, 工学的な問題解決手法である SVM 機械学習器のモデルを構築した. 「将来, 肺気腫を発症し COPD へ進行するのかを予測する」SVM 機械学習システムを構築し, システムの分類・予測の精度を検証する内容となっている. ここでは, 肺 CT 値の階級度分布の 1 年当たりの経年変化量を 14 次元の特徴量とし, SVM 機械学習器に入力する. 教師

信号には CT 値の経年低下量 $-30\text{HU}$  をしきい値とし、それよりも低下した対象を「CT 値の急速低下群：肺気腫拡大傾向」、それ以外を維持群として学習を行った結果、初回のシステム分類テスト精度は 90% となった。分類精度を向上させるために、PCA (Principal component analysis：主成分分析) 処理を行い、その 2 軸散布図を参考にしながら、学習データ数の増減や学習データ中の正解ラベルの比率引き上げ等、調整作業を繰り返した。結果、最終的な分類精度は最高 96% まで向上した。特に、臨床的には FN：False Negative (偽陰性)、「所見はあるのに見落とした」という結果は絶対にあってはならないものであり、この誤分類を減らすことは重要であった。

第 6 章では本論文の結論として、研究一連から導かれる事実及び、それらに対する考察を行った。また、本研究に関する今後の期待など、その応用の実現に必要な課題などに関しても付け加えている。本研究の最終的な目標は受診医療機関でたった 1 度の胸部 CT 検査を行っただけで「COPD に進行する可能性があります、精密検査を受けてください」などと、その危険性を警告・表示するシステムを完成させることである。これによって、COPD の予防措置、また次なる医療の提供につなげることができれば、COPD 患者の絶対数が減少すると考える。

以上を要するに、本論文は北海道地域において早急に対応が必要な慢性疾患である COPD に対するごく初期段階での診断支援に機械学習の手法を利用することで、より精緻な指針を与えようとするもので、学術的・社会的な意義は非常に大きいものである。