

学位論文及び審査結果の要旨

横浜国立大学

氏名	小林 雅幸
学位の種類	博士(情報学)
学位記番号	環情博甲第2219号
学位授与年月日	令和3年3月25日
学位授与の根拠	学位規則(昭和28年4月1日文部省令第9号)第4条第1項及び横浜国立大学学位規則第5条第1項
学府・専攻名	環境情報学府 情報環境専攻
学位論文題目	画像分類器の説明性向上に関する研究 A Study on Improving the Interpretability of Image Classification Models
論文審査委員	主査 横浜国立大学 教授 長尾 智晴 横浜国立大学 教授 田村 直良 横浜国立大学 教授 森 辰則 横浜国立大学 准教授 富井 尚志 横浜国立大学 准教授 原下 秀士 横浜国立大学 講師 白川 真一

論文及び審査結果の要旨

近年、深層学習をはじめとした機械学習による処理の自動化に関する技術が注目されており、様々な分野で研究が活発に行われている。画像分類を例に挙げると、深層学習モデルの一つである Convolutional Neural Network (CNN) は様々な画像分類タスクにおいて高精度な手法として知られている。しかし、これらの構築された処理は複雑で解析は困難であり、処理過程は未だにブラックボックス化されている場合がほとんどである。CNNなどの獲得された処理の多くは社会や産業応用の現場で使用されることが期待されているが、これまでは精度だけが注目され、構築された処理を利用する人の立場で注目していることは少ないと考えられる。

そこで本論文では、分類精度と可読性の両方に優れる画像分類器の提案を行う手法を提案し、それらの有効性を実験的に検証している。本論文は、序論(第1章)と結論(第6章)を含めて6章から構成されている。

第2章「関連研究」では、本研究に関連する従来研究について述べている。はじめに、機械学習の説明性向上に関する従来研究として、進化的条件判断ネットワークについて述べ、続いて学習済み深層学習の解析に関する先行研究として、いくつかの従来研究について述べ、それぞれの特徴と問題点を示すことで、本研究の意義を明らかにしている。

第3章「If-then ルールを用いた分類器の精度と可読性の向上」では、分類器の中でも比較的人間が理解しやすい手法とされる if then ルールで分類を行う分類器の精度と可読性を向上させる方法の提案を行なっている。提案手法では分類に用いられている特徴量分布を考慮したヒートマップの作成と分類器内の各条件分岐における特徴量の可視化を行うことで、分類器が分類の際に画像のどの領域に注目して分類を行ったかを利用者に提示することが可能になると述べている。実験では提案手法を一般画像分類に適用し、獲得した可視化画像の有効性の検証を行なって有効性を確認している。

第4章「学習済み深層学習モデルの特徴量の可視化」では、学習済み深層学習モデルの可視化手法の提案を行なっている。先行研究では学習済み CNN の特定ユニットの活性化と自然画像への変換を同時に満たすための制約を人手で設計した上で生成を行うが、提案手

法では Generative adversarial networks; GAN の枠組みを用いることで、自然画像への制約を用いることなく End to end で画像生成器を学習することが可能になると述べている。実験では提案手法を学習済み AlexNet に適用し、モデルの内部で獲得された特徴量の可視化を行なっている。また、任意のクラスに対して強く発火する特徴量の可視化を行うことで、モデル内部で階層的な特徴量が獲得されていることを示している。

第5章「高精度かつ高い可読性をもつ深層学習」では、高い分類精度と可読性をもつ新しい深層学習モデルである Evolutionary generative contribution mappings (EGCM)を提案している。先行研究では精度と可読性のトレードオフの問題のため、可読性向上の機構を導入することで大幅な分類精度の低下が見られる。これに対して提案手法は可読性の高いモデルに適したモデル構造を進化計算法で自動最適化することで、高い精度と高い可読性をもつ分類器を自動獲得することが可能になると述べている。実験では提案手法を複数のデータセットに適用し有名な深層学習モデルの性能比較を行うことで EGCM が高い可読性を保ちながら先行研究と同等精度以上の性能もつことを示している。

第6章「結論」では、本研究の成果を総括している。

以上のように、本論文では、昨今必要性が益々高くなりつつある機械学習の説明性を向上させる問題について、特に画像認識を対象とした「説明できる AI (XAI: explainable AI)」手法を提案してその有効性を実験的に示している。本論文で提案された手法は学術的新規性が高く、かつ産業応用可能性が高い実用性も兼ね備えた手法であり、学術的新規性が高いとともに実用性が高い手法であると言える。本論文の成果は、査読付き論文誌の正論文2編、フルペーパー査読付き国際会議論文2編、並びに関連正論文2編、関連国際会議論文5編、5件の口頭発表などを通して公表され、いずれも高く評価されている。

以上から、本論文は博士(情報学)の学位論文として十分な価値を有すると審査員全員一致して認め、令和3年1月20日(水)13時00分から14時30分まで、オンライン(Teams)にて博士論文公聴会を開催した。博士論文公聴会では充実した質疑応答がなされた。

その後、同日14時30分から引き審査委員全員出席のもと、小林雅幸氏の学力の確認を行った。始めに、同氏より同氏の博士論文「画像分類器の説明性向上に関する研究(A Study on Improving the Interpretability of Image Classification Models)」の概要と、博士論文に関連する学術論文の公表状況、大学院講義の単位取得状況などについて説明があった。

これに引き続き質疑応答を行った。審査員からは博士論文に関する質問、一般専門知識に関する質問などがあり、学力及び博士号取得資格の確認を行なった。発表論文については、学術論文(正論文)2編、正論文相当のフルペーパー査読国際会議論文2編が採録・公表されていることから、博士号取得の資格を十分に満たしていることを確認した。外国語については、本人が国際会議において英語で発表していることから学力を確認した。また、課程博士修了のための単位を取得していることを確認した。

以上により、同氏は学力の確認に合格であると、審査委員全員一致で判定した。

以上の論文審査委員会の結論に基づき、令和3年2月15日(月)に開催された環境情報学府 情報環境専攻会議にて審議し、全員一致で本論文を博士(情報学)の学位論文としての価値があるものとして環境情報学府教授会に付議することを決定した。その後、環境情報学府学務委員会での確認を経て、令和3年3月1日(月)に開催された環境情報学府教授会において審議を行い、無記名投票により、小林雅幸氏に博士(情報学)の学位を授与することを決定した。