

氏名	曲 佳		
学位の種類	博 士 (工学)		
学位記番号	博 甲 第 9 0 0 1 号		
学位授与年月日	平成 3 1 年 3 月 2 5 日		
学位授与の要件	学位規則第 4 条第 1 項該当		
審査研究科	システム情報工学研究科		
学位論文題目	Studies on Stepwise Transfer of Domain Knowledge for Computer-aided Diagnosis in Pathology Using Deep Neural Networks (深層ニューラルネットワークによる病理診断支援のためのドメイン知識の段階的転移に関する研究)		
主 査	筑波大学 教授 (連携大学院) (産業技術総合研究所)	博士 (工学)	坂無英徳
副 査	筑波大学 教授	工学博士	水谷孝一
副 査	筑波大学 教授 (連携大学院) (産業技術総合研究所)	博士 (工学)	村川正宏
副 査	筑波大学 教授 (連携大学院) (産業技術総合研究所)	博士 (工学)	喜多泰代
副 査	東邦大学 准教授	博士 (医学)	蛭田啓之

論 文 の 要 旨

本審査対象論文は、深層畳み込みニューラルネットワーク (DCNN) を病理診断支援へと適用するにあたって、少数の訓練データで効率よく学習させるための多段階転移学習を導入し、さらに専門家 (病理診断医) の労力を抑制しつつ効果的に教師データを作成する方法論を提案している。第 2 章において、病理診断分野の現状を概観し、診断支援技術に関する先行研究のサーベイを行っている。第 3 章では、病変部位の自動検出における DCNN の優位性と、その課題を明らかにしている。

第 4 章では、上記課題を克服するための提案手法として、段階的転移学習のフレームワークが紹介されている。提案手法では、大量に入手可能な公開自然画像データセット (ImageNet) を用いて DCNN を最初に訓練し、訓練用データセットを最終的に解決すべき目的のドメインへと段階的に近づけてゆくことで、目的ドメインの訓練用データが少ない場合でも効率的に学習が可能となる。第 5 章では、目的ドメインと同じ病理画像を使用しながらも、高度な専門性を要しないラベル付け作業を非専門家が行ったデータセットを中間段階の訓練に用いることで効果を得られることを示している。第 6 章では、汎用の画像認識技術として優れた実績を有する特徴抽出を利用して、模様としての相似性に基づいて病理画像をクラスタリングすることで中間段階の訓練用データセットを人手に依らず構築する手法が提案されている。第 7 章では、上記 2 手法を組み合わせた更なる多段階化や、他の臓器のデータを用いたマルチタスク学習を導入することで性能が一層向上することを示している。

第 8 章では本論文の総括を行い、提案手法が他臓器の病理画像だけでなく、他のモダリティへと展開できる可能性について言及している。

審査の要旨

【批評】

転移学習は、高品質なアノテーションが施された訓練用データを大量に準備することができないドメイン（タスク）へと機械学習を適用する際に有効な手法として知られているが、これを医療分野（病理診断）において段階的に適用して優れた効果を得るための方法論を提案していることが本論文の新規性である。段階的転移学習を実現するための、3つのアプローチが示されている。

第1のアプローチでは、目的ドメインと同じデータに基づきながらも異なる観点でアノテーション（ガン／非ガンの分類ではなく、上皮／非上皮組織の分類）されたデータセットを中間段階で学習させている。知識転移の観点から考察すると、公開済み大規模自然画像データセット（ImageNet）を用いて「一般的な画像の見方」を獲得した DCNN に、中間段階にて「病理画像の見方」を学ばせ、最後に目的の「ガン／非ガンの区別」を学習させており、最終的な目標へ向かってタスクを段階的に高度化・複雑化する方法論は理に適っている。実際、計算機実験でも提案手法の優位性が示されている。

第2のアプローチでは、提案者が以前に考案した画像認識技術を用いて自動生成したデータセットを中間段階として使用している。第1のアプローチでは人間が病理画像を分類した「結果」を教師データとして DCNN に与えているのに対し、ここでは分類の「仕方」という、より高次の知識を転移させようとしている。近年注目されている自己学習型人工知能の枠組みで捉え直すことで、新しい学術的発見を生み出しうると期待される。本論文中では、汎用性の高い特徴抽出技術に基づくクラスタリング手法によって、人間による介入なしに中間段階の教師データを自動生成できることを実証しており、医療分野のみならず、データ収集コストの高いドメインにおける本提案手法の有効性や可用性の高さが示されたと言ってよい。

第3のアプローチでは、異なる臓器の病理画像データセットを利用したマルチタスク学習を中間段階で実施している。具体的には、胃とリンパの病理画像群をそれぞれ上記第2のアプローチに基づいて2クラスに分割し、合計4クラスにラベル付けされたデータセットを生成し、これを中間段階の教師データとして使用することで更なる性能向上を果たした。異なるドメインのデータを混ぜて教師データとして用いる発想は大変ユニークである。2クラス分類だけで段階の多い転移学習を行う場合と、多クラス分類で少ない段数の知識転移をさせた場合との厳密な性能比較や使い分け指針の検討など、学術的に考察を深めるべき要素は残されているが、転移学習という普及した技術に更なる発展の可能性を示した点で本論文の意義は高い。

【最終試験の結果】

平成31年2月13日、システム情報工学研究科において、学位論文審査委員の全員出席のもと、著者に論文について説明を求め、関連事項につき質疑応答を行った。その結果、学位論文審査委員全員によって、合格と判定された。

【結論】

上記の学位論文審査ならびに最終試験の結果に基づき、著者は博士（工学）の学位を受けるに十分な資格を有するものと認める。