

論文審査の要旨

博士の専攻分野の名称	博 士 （ 情報科学 ）	氏名	野村 祐一郎
学位授与の要件	学位規則第 4 条第 1 項・ 2 項該当		
論 文 題 目			
Methods for Robust Training of Deep Neural Networks in the Presence of Noisy Labels (誤りを含む教師信号からの深層学習の頑健な訓練法)			
論文審査担当者			
主 査	教 授	栗田 多喜夫	印
審査委員	教 授	向谷 博明	印
審査委員	教 授	本田 直樹	印
審査委員	准教授	Raytchev Bisser Roumenov	印
審査員	准教授	宮尾 淳一	印
〔論文審査の要旨〕			
<p>深層学習が、昨今の人工知能の発展を牽引している。深層学習モデルを安定に学習させるためには、正確にラベルが付与された大量の学習データが必要である。一般には、匿名の素人にアンケートを依頼してデータを生成するクラウドソーシングなどのデータ収集技術が開発されている。しかしながら、これらのデータ収集法では、付与されたラベルに誤りが含まれている可能性が高く、学習されたモデルの性能は著しく低下してしまう。</p> <p>本論文は、ラベルノイズの影響を低減し、深層学習モデルを頑健に学習させるための3つの新たな学習法を提案している。</p> <p>第2章では、まず、画像分類と深層学習の概要を説明し、深層学習の構造と学習アルゴリズムについて解説している。続いて、ラベルノイズの定義と、深層学習を用いた分類問題の問題設定について説明している。</p> <p>第3章では、深層学習モデルの学習中に訓練データのラベルをクリーンなラベルに更新する手法を提案している。深層学習がラベルノイズに過適合する前に、深層学習モデルの中間層から全ての学習サンプルの特徴ベクトルを抽出し、学習サンプル間の類似性グラフを構築し、グラフ上でラベル伝搬を行うことで、ラベルノイズを除去する。ラベルの更新とモデルパラメータの更新を交互に行うことで、ラベルノイズに対して頑健な学習手法を実現した。手書き数字文字データセット (MNIST) と 10 クラスある画像データセット (CIFAR-10) に対して人工的なラベルノイズを付与したデータに対して、既存手法よりも高い分類精度であることを確認した。</p>			

第4章では、**Self-Supervised Learning** の手法を応用した頑健な訓練手法を提案している。先行研究である **DivideMix** は深層学習から得られる各サンプルの誤差値に着目し、誤差値の分布に対して2成分混合ガウスモデルを学習することで、データセットをクリーンなデータセットとラベルノイズのあるデータセットに分割することに成功した。**DivideMix** は誤差値の小さいサンプルをクリーンなサンプルとして扱うが、誤差値の小さいサンプルは分類しやすい単純なパターンを持つ傾向があり、このようなサンプルを重点的に学習すると、モデルは単純なパターンに過適合してしまい、ラベルはクリーンだが分類が困難なサンプルを見落としてしまう。そこで **DivideMix** によりクリーンであると選択された入力サンプルに対してノイズを加える **Consistency Regularization(CR)**を導入し、単純なパターンへの過適合を防ぎつつ、分類が困難かつクリーンなサンプルを学習することに成功した。画像分類データセット **CIFAR-10** と **CIFAR-100** に対して、テストセットでは既存手法よりも高い分類精度が得られることを示した。

第5章では、深層学習を頑健に学習させるための新たなサンプル選択法を提案した。深層学習は単純なパターンを最初に学習するため、学習の初期段階ではラベルノイズに対して頑健である。学習の初期段階では、ラベルノイズのあるサンプルに対しては与えられたラベルと矛盾するラベルを予測する。そのためラベルノイズのあるサンプルに対しては、学習中に発生する誤った予測の総数が、クリーンなサンプルよりも多くなる。そこで、サンプルに対するモデルの誤予測数を利用した新たなサンプル選択手法を提案した。**CIFAR-10** と **CIFAR-100** を用いた比較実験では、従来のサンプル選択手法よりも優れた分類精度を達成し、提案手法は有効なサンプル選択手法であることを確認した。

これらの手法は、いずれもラベルノイズに対して頑健な学習が実現できることを示している。これらの結果は、未学習データに対する性能（汎化性能）とも関連しており、深層学習の本質とも関連し、今後のさらなる進展が期待される。

以上、審査の結果、本論文の著者は博士（情報科学）の学位を授与される十分な資格があるものと認められる。