論　文　の　要　旨

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　）

氏　名　　野村　祐一郎

人工知能とは人間の脳を人工的に模倣する試みであり、近年、生活のさまざまな分野で急速に普及し始めている。人工知能を実現するための基礎的な技術の一つとして、学習データからパターンを発見し、モデルのパラメータを自動的に決定する「機械学習」がある。機械学習は非常に有用なツールであり、カメラやスマートフォン、ソーシャルメディアなど、現代社会に欠かせない技術やサービスに応用されている。機械学習の具体的な応用例としては、e-commerce の推薦システム、画像中の物体認識、顔検出、自動言語翻訳、音声合成などがある。これら全てのアプリケーションにおいて、大規模データセットの出現と計算機資源の増大により発展した深層学習（Deep Learning）と呼ばれる機械学習技術が重要な役割を担っている。

深層学習は複数の階層的な処理層からなる機械学習モデルであり、複数の抽象度でデータの表現を学習することができる。深層学習の各層は、下位の層から得られた入力データの表現に基づいて、最終層が目標とする信号を出力するようにモデルのパラメータを調整する。このような変換を膨大に組み合わせることで、深層学習は複雑な関数を学習し表現することができる。深層学習の強みは、手作業による調整や専門知識を必要とせず、データからそのような複雑な関数を自動的に学習できることである。深層学習は画像認識や音声認識など、人工知能が長年取り組んできた問題を解決し続けており、今後さらなる成果が期待されている。

しかし欠点として、深層学習モデルを安定的に学習させるためには、正確にラベルが付与された学習データを大量に必要とする点があげられる。データの各サンプルのラベルとは、そのサンプルが属するカテゴリを指し、主に分類学習を対象として利用される。通常それらのデータセットを用意するためには専門家に正確なラベル付けを依頼するが、全てのサンプルにラベルを付与するには膨大な時間と費用がかかるという問題がある。高コストなラベル付与問題を軽減するため、匿名の素人にアンケートを依頼してデータを生成する Amazon Mechanical Turk といったクラウドソーシングや、Web 上の画像周辺にあるキーワードからデータのラベルを自動的に収集する Web クローリングなど、非専門家によるデータ収集技術が開発されている。これらの手法で収集されたデータセットの例として、画像分類コンペティション ILSVRC2012 で用いられた ImageNet がある。

クラウドソーシングのような非専門家によるデータ収集技術は、容易に大規模なデータセ

ットを作成し、深層学習の発展に大きく貢献している。しかし、これらの非専門家が付与するラベルは、専門家が付与するラベルと比較して不正確であり、付与されたラベルに誤りが含まれている可能性がある。例えば、Amazon Mechanical Turk により収集された ImageNet では、作業者の知識不足やラベル割り当て方法の説明不足が原因で誤ったラベルが生成されている。また、専門家がラベルを付与した場合でも、質の低いサンプルや識別が困難なサンプルに対しては誤って付与されたラベルが発生する。さらに、単純にデータの符号化や通信の問題で誤ったラベルが発生することもある。このような真のラベルとは異なる誤ったラベルは「ラベルノイズ」と呼ばれる。人間がラベル付与に関わる場合、全てのデータに正確にラベルを付けることは困難であるため、ラベルノイズは実世界の機械学習応用において本質的に避けられない問題である。実際にラベルノイズを含む複数の実世界データセットを分析した結果、ラベルノイズを含む訓練サンプルの割合は 8.0%から 38.5%に及ぶと報告されている。

機械学習が特に成果を上げている課題の一つに分類学習がある。これは、学習データからカテゴリごとのパターンを学習し、テストデータのサンプルがどのカテゴリに属するかを予測する問題である。しかし機械学習モデルの性能はデータセットの質に依存し、教師信号（ラベル）が不正確な場合、性能が著しく低下してしまう。ラベルノイズは現実的に避けられない問題であるため、ノイズに対して頑健なモデルの学習手法が重要である。機械学習の分類器をラベルノイズのあるデータセットで学習した場合、真の事後確率ではなく、誤った事後確率を学習してしまう。その結果、分類器はテスト画像に対して誤った予測をしてしまい、汎化性能が低下する。本論文では、学習データは誤った分布からサンプルされ、ラベルにノイズはあるが、入力データの特徴にノイズはないと仮定する。一般的なロバスト学習では、入力空間に欠損やノイズのあるサンプルがあっても、教師信号の値を正しく出力するように分類器を学習させる。本論文ではノイズは入力空間ではなく、一部のサンプルの教師信号（ラベル）に含まれていると仮定する。

特に深層学習モデルは、モデルのパラメータ数が膨大であり、任意の複雑な関数を表現できるため、ラベルノイズに容易に過適合することが報告されている。深層学習がラベルノイズに過適合すると、汎化性能が低下し、テストデータにおける分類性能が低下する。このようにラベルノイズに対して過適合しやすい深層学習の性質は”memorization effect”と呼ばれている。memorization effect の解析により、深層学習は学習初期に入力データの単純なパターンを学習し、高い汎化性能を発揮することが報告されている。しかし、そのようなパターンを学習した後、深層学習はラベルノイズのあるサンプルに対して過適合を始め、分類性能が低下する。ラベルノイズが深層学習の汎化性能に悪影響を及ぼすことは明らかであり、深層学習の応用にはラベルノイズに頑健な学習手法が必要である。本論文では、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に提案した手法を説明する。

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する。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する。さらに、深層学習を用いたラベルノイズに関する先行研究についても言及する。第 3 章、第 4 章、第 5 章では、ラベルノイズのある画像分類問題において、汎化性能の低下を緩和するために提案した手法を紹介する。第 6 章では本研究の貢献を総括し、今後の課題について述べることで、本論文を締めくくる。

第 3 章で説明する手法は、ラベルノイズ問題を緩和するため、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。提案手法は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する。この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。その後、ラベル伝播によりラベルの更新されたデータセットを用いてモデルの学習を再開する。提案手法は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。比較実験では手書き数字文字データセットである 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、提案手法の性能を評価した。その結果、提案手法はラベルノイズに対して過適合せず、既存手法と比較して高い分類精度が得られることを示した。

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た。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る。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 に人工ラベルノイズを施し、提案手法と DivideMix や最先端手法との性能を比較した。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、CR が汎化性能の向上に寄与していることを示した。

第 6 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について述べる。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。つまり訓練モデルは学習の初期段階で各サンプルの真のラベルを予測する傾向があるため、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。つまり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。事前実験から、ラベルノイズのないサンプルはラベルノイズのあるサンプルより誤予測数が少ないことを確認した。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。
人工ラベルノイズを付与した CIFAR-10 と CIFAR-100 を用いた比較実験では、提案手法が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示した。