



HOKKAIDO UNIVERSITY

Title	Study on Robust Large-scale Neural Network Optimization [an abstract of dissertation and a summary of dissertation review]
Author(s)	張, 恩智
Degree Grantor	北海道大学
Degree Name	博士(情報科学)
Dissertation Number	甲第16502号
Issue Date	2025-03-25
Doc URL	https://hdl.handle.net/2115/94990
Rights(URL)	https://creativecommons.org/licenses/by/4.0/
Type	doctoral thesis
File Information	Enzhi_Zhang_review.pdf, 審査の要旨



学位論文審査の要旨

博士の専攻分野の名称 博士 (情報科学) 氏名 張 恩智

審査担当者 主 査 教 授 棟朝 雅晴

副 査 教 授 山本 雅人

副 査 准教授 杉木 章義

学位論文題名

Study on Robust Large-scale Neural Network Optimization

(堅牢な大規模ニューラルネットワーク最適化に関する研究)

深層ニューラルネットワークにおける過学習の検出と回避は、機械学習、人工知能の分野において堅牢なニューラルネットワークを求める最適化のため非常に重要な課題であり、学習率の制御やドロップアウト、敵対的学習、データ拡張など様々な研究がなされてきた。この問題に対処するため、メタラーニング、特にメタレベルでのハイパーパラメータ最適化手法が提案されており、そこでは内部と外部の2つのレベルで目的関数を定義している。通常、このような2レベルの最適化には2つの一般的なアプローチがある。第1のアプローチは、外部のターゲットを最適化し、コントローラがパラメータ空間を探索する。その後、内部の最適化から得られたメタ情報を使用してコントローラを改善する。この方法は実装が容易であり、例えばグリッドサーチなどが用いられるが、計算集約的な内部最適化が必要となるため、探索に要する計算コストが高い。第2のアプローチでは、外部の目的関数を直接更新する。メタレベルにおける勾配ベースのアプローチでは、ハイパーパラメータに関する高次の勾配を計算する。しかしながら、高次のヘシアン行列の計算にコストを要するにもかかわらず、メタレベルでの勾配の劣化も問題となる。したがって、内部モデルを直接かつ効率的に訓練しながらメタレベルでの勾配を分離することが必要となる。

本論文では、勾配降下のプロセスにおいて検証損失の情報を活用することで、上記の課題を解決することを試みた。この目的を達成するために、訓練プロセス、データ拡張、およびモデル構造の観点から検討を行った。訓練プロセスの文脈では、与えられた状態の下で最大値を持つアクションを選択する形で過去の検証知識を利用することを意味しており、この概念は強化学習に適している。類似の枠組みとして、勾配ベースの最適化を強化学習のタスクまたはマルコフ決定プロセスとして扱うことが一般的である。特に、重みと勾配の空間をそれぞれ状態空間とアクション空間として扱うこととなる。次に、データ拡張の文脈では、逆勾配を使用して偽のサンプルを生成し、それを訓練中のデータ拡張に使用する。モデル構造の側面では、特にトランスフォーマーをベースとした事前訓練モデルに対して、ガウス平滑化とクアッドツリー分割に基づいて画像パッチを動的に調整し、大規模な問題に対してもセグメンテーション精度を失うことなく、トランスフォーマーモデルのシーケンス長を大幅に削減することができた。

これを要するに、著者は、堅牢な大規模ニューラルネットワークの最適化手法に関して新知見を得

たものであり、機械学習ならびに人工知能分野の発展に貢献するところ大なるものがある。よって著者は、北海道大学博士 (情報科学) の学位を授与される資格があるものと認める。