



HOKKAIDO UNIVERSITY

Title	ニューラルネットワークの効果的な訓練のための探索と収束の制御 [論文内容及び審査の要旨]
Author(s)	高瀬, 朝海
Degree Grantor	北海道大学
Degree Name	博士(情報科学)
Dissertation Number	甲第13298号
Issue Date	2018-09-25
Doc URL	https://hdl.handle.net/2115/71804
Rights(URL)	https://creativecommons.org/licenses/by-nc-sa/4.0/
Type	doctoral thesis
File Information	Tomoumi_Takase_abstract.pdf, 論文内容の要旨



学 位 論 文 内 容 の 要 旨

博士の専攻分野の名称 博士（情報科学） 氏名 高瀬 朝海

学 位 論 文 題 名

ニューラルネットワークの効果的な訓練のための探索と収束の制御
(Controlling Exploration and Convergence for Effective Training of Neural Networks)

近年、様々な問題に対する機械学習手法として、ニューラルネットワークが広く利用されている。ニューラルネットワークの学習において、モデルの汎化性能を向上させることが重要であるが、そのためには誤差関数上の望ましくない局所解にとらわれないようにパラメータ（すなわち、ニューラルネットワークを構成する各ニューロンの重み）を反復的に更新する必要がある。一般的には、誤差の小さい解に収束することが汎化性能の向上につながると考えられているが、フラットな解に収束することもまた重要であると言われている。

現在広く利用されている確率的勾配降下法 (SGD 法) や、それを発展させた AdaGrad 法などのパラメータ更新手法を用いると、誤差の大きな解やシャープな解による捕捉を避けつつ、誤差の小さな解やフラットな解に収束することが困難である。そこで、本論文ではそれらの課題を解決するための効果的な学習戦略として、学習の前半は広い範囲を探索し、後半は安定した収束を行うという戦略を考えている。具体的には、ニューラルネットワークのハイパーパラメータである学習率とバッチサイズを動的に変化させるアプローチに着目し、学習率の動的変化と誤差の小さな解への収束との関係、およびバッチサイズの動的変化とフラットな局所解への収束との関係について、それぞれ研究を行っている。そして、上述の戦略に基づき、ニューラルネットワーク学習のための2つの手法として、適応的学習率調整法 (ALR 法) およびバッチサイズ増加法を提案している。ALR 法は、学習中の訓練誤差が最小となるようにエポックごとに学習率を決定する手法であり、一般には学習率が増加から減少に転じることで上述の戦略に従う。バッチサイズ増加法は学習中のバッチサイズを逐次的に増加させる手法であり、それが学習戦略に従うことを、誤差関数の非定常的な特性に基づいて説明し、学習中の探索範囲の広さの動的な変化を調べることで実証する。本論文は、それらの手法がその戦略に従っていることの説明と実証、およびそれらの手法を用いることで汎化性能が向上することの実証を目的としている。

本論文では、上記の内容を次の構成で述べている。

第1章では、研究背景、研究目的、および本論文の構成について述べている。

第2章では、ニューラルネットワークおよびその学習について説明している。まず代表的なニューラルネットワークである多層パーセプトロンや畳み込みニューラルネットワークを紹介する。その後、SGD 法や AdaGrad 法などのパラメータ更新手法を紹介し、本論文の提案手法に関係するミニバッチ学習や学習率についても説明を加え、さらに非凸性を考慮した学習戦略について述べている。

第3章では、学習率の動的変化と誤差の小さな解への収束との関係、および提案手法である ALR 法について説明している。まず、学習率決定に関する先行研究を紹介した後、訓練誤差最小化に基づく ALR 法を説明するとともに、本手法を用いた学習アルゴリズムおよび3つの主要パラメータの影響について説明している。その後、ALR 法を用いた実験結果に基づき、手法の有効性について議論している。ALR 法と一般的なニューラルネットワークの更新手法である SGD 法を用いた場合とで、

訓練誤差およびテスト精度の結果を示し,ALR 法が汎化性能を向上させることを示している. また, 学習中の学習率の推移から, 探索と収束の制御に関する本手法の効果について考察を行っている. さらに, 更新手法の比較や, 主要パラメータの影響, 畳み込みニューラルネットワークを用いた結果, 複数のデータセットの利用結果などについても, 実験結果に基づき考察を行っている. 最後に ALR 法の問題点と今後の課題について述べている.

第 4 章では, バッチサイズの動的増加とフラットな局所解への収束との関係, および提案手法であるバッチサイズ増加法について説明している. まず, 非凸最適化問題の関連研究を紹介し, その論文で示された, 大きなバッチサイズを用いたときに汎化性能が悪化するという結果について述べている. 本論文では, この理由を損失関数の非定常的な特性に基づき説明し, それを発展させ, 効果的な学習を行うバッチサイズ増加法を導いている. その後, バッチサイズ増加法を用いた実験結果に基づき, 手法の有効性について議論している. 訓練誤差やテスト精度の結果を示し, バッチサイズ増加法がバッチサイズを固定した場合に比べて, 過学習を抑制できていることを示している. また, 学習の前半および後半の探索範囲を示し, バッチサイズ増加法を用いることによって探索と収束を制御できていることを確かめている.

第 5 章では, 本論文のまとめおよび今後の課題や展望について述べている.