

称号及び氏名 博士（工学） 山田 良博

学位授与の日付 2020年 3月 31日

論文名 「ニューラルネットワークの確率的正則化に関する研究」

論文審査委員 主査 黄瀬 浩一

副査 本多 克宏

副査 森 直樹

副査 岩村 雅一

論文要旨

近年、ニューラルネットワークの研究が飛躍的に進んでいる。ニューラルネットワークは、人間の神経回路網を単純化したモデルであり、あらゆる関数を表現可能なことから機械学習に利用される。ニューラルネットワークは1940年代初頭から研究が始まり、幾度か注目を集めた。しかし、他の手法に比べ計算量が多いことや精度が劣ることから、注目されない時代が長く続いた。ニューラルネットワークが改めて注目されるようになったのは2012年のことである。一般物体認識と呼ばれるタスクを対象としたコンペティションにおいて、従来手法を凌駕したことでニューラルネットワークは一躍脚光を浴び、以来ニューラルネットワークは重要な技術で在り続けている。

ニューラルネットワークは学習と呼ばれるパラメータの勾配の計算と調整を通して多くのタスクを実現する。しかし、どのような条件でも上手く学習するとは限らない。一定の条件下では、いくら学習しても学習に用いたデータの多くを上手く認識できないアンダーフィッティングと呼ばれる問題が発生する。アンダーフィッティングは主にパラメータ数が少ないことが原因で発生するが、パラメータ数が十分多い場合にもパラメータ以外の要因でアンダーフィッティングが起り得る。特に一般物体認識において有利と考えられていた多層な構造をもつニューラルネットワークは、その多層な構造が原因でアンダーフィッテ

ングが起こり得る。この問題は多層な構造によって多くのパラメータの勾配が非常に小さな値や非常に大きな値となり、学習が進まなくなることによって起る。この現象は勾配消失と呼ばれる。勾配消失を大幅に緩和したのが ResNet である。ResNet は Residual Block と呼ばれる特殊な構造によって、従来は学習が困難であった 100 層以上の多層な構造での学習を容易にし、当時の一般物体認識と呼ばれるタスクにおける最高精度を達成した。ResNet の登場は大きな衝撃を与え、画期的な構造である Residual Block の改良に焦点が当たるようになった。Residual Block を異なる形で改良した Wide ResNet, PyramidNet, ResNeXt は相次いで精度を改善し、これらの改良手法は現在の一般物体認識における標準的手法となっている。ResNet, Wide ResNet, PyramidNet, ResNeXt の 4 つをまとめて ResNet 及びその改良手法と呼称する。

ResNet 及びその改良手法はアンダーフィッティングの問題を緩和したが、依然としてオーバーフィッティングと呼ばれる大きな問題が残っている。オーバーフィッティングとは、学習に用いたデータはほぼ全て認識できるが、学習に用いなかったデータの多くが認識できない現象である。対策として正則化と呼ばれる処理がある。正則化は、学習の際に意図的にデータの一部を欠損させる等の工夫を施すことで、ニューラルネットワークに追加の情報を与え、学習に用いなかったデータの一部を認識可能にさせオーバーフィッティングを緩和する工夫である。実際に ResNet 及びその改良手法では、Data Augmentation, Weight Decay 等、既に様々な正則化が使われている。しかし、ResNet 及びその改良手法においてオーバーフィッティングの影響は依然として大きく、効果的な正則化が研究されている。

ResNet 及びその改良手法の精度を改善するためには、それらで基本的には用いられていない正則化を導入することが有効であると考えられる。実際に、確率的にニューラルネットワーク構造を変化させる確率的正則化と呼ばれる正則化の導入によって、それらの精度が改善することが知られている。ただし、ResNet 及びその改良手法に対する確率的正則化の研究には以下の 2 つの課題がある。1 つ目は、優れたニューラルネットワーク構造の多くで、有効な確率的正則化が検討されていない点である。2 つ目は、Residual Block 構造によらない、汎用的な確率的正則化が検討されていない点である。

そこで本研究では、前述の 2 つの課題に対応すべく、ResNet 及びその改良手法に導入する新たな確率的正則化を検討する。1 つ目の課題に対しては、ResNet に対する確率的正則化 RandomDrop を PyramidNet に適用する事例を通して検討する。PyramidNet に RandomDrop を適用すると、条件によっては精度が改善しない。これは PyramidNet のネットワーク構造が原因であると考えられる。ResNet と PyramidNet はどちらも先述の Residual Block を積層した構造をもつ。ResNet は殆どの Residual Block で特徴のサイズが変わらない。対して、PyramidNet は全ての Residual Block で特徴のサイズが変わる。このサイズの変化の違いが精度の改善を妨げていると考えられる。そこで、サイズの変化を考慮した確率的正則化として、サイズが変化する部分とそうでない部分で処理を分けた確率的正則化を新たに設計し、精度を改善する。2 つ目の課題に対しては、Residual Block 構

造によらない汎用的な確率的正則化を検討する。その中で、特定の構造にしか使えないが効果的な手法である確率的正則化 **Shake-Shake** を分析し、**Residual Block** 構造によらない汎用的な確率的正則化 **ShakeDrop** を提案する。**Shake-Shake** は学習時の勾配の計算を敢えて狂わせることで効果を発揮するが、この原理を単純に他のネットワーク構造に適用すると、適切な学習が行えずかえって認識精度が低下する。そこで、適切な学習が行えるよう、確率的正則化 **RandomDrop** を導入した提案手法 **ShakeDrop** を提案する。それぞれ一般物体認識データセットを用いて、提案手法を用いた際の認識精度の改善を確認し、いずれも発表時点における最高精度を達成した。更に、**PyramidNet** において、汎用的な手法として設計された提案手法 **ShakeDrop** を適用した場合、正則化を用いない場合や **PyramidNet** に特化した提案手法を適用した場合よりも高い性能を示すことが分かった。

本論文は以下の全 5 章から成る。

第 2 章では、以降の章で参照する内容として、一般物体認識において代表的なニューラルネットワークと、それらに用いられる正則化について説明する。最初に多くの手法の基礎となった **ResNet** の画期的な構造である **Residual Block** に関する説明を行い、**ResNet** 及びその改良手法である **Wide ResNet**, **PyramidNet**, **ResNeXt** の差異や概要を説明する。次に、ニューラルネットワークに利用される正則化、ならびに本研究において重要な先行研究である確率的正則化の **RandomDrop** と **Shake-Shake** について、その仕組みや特徴について詳細に述べる。更に、それらの評価に用いられる一般物体認識用データセットの概要を説明する。

第 3 章では、**PyramidNet** と **RandomDrop** の単純な組み合わせである **PyramidDrop**, 更に工夫を施した **PyramidSepDrop** を提案して実験を通して比較、分析を行う。一般物体認識データセット **CIFAR-100** を用いた実験では、幾つかの条件の変化に伴って、確率的正則化の効果が向上することが確認された。更に、それぞれの条件で、**Residual Block** の構造を考慮した工夫を施した **PyramidSepDrop** が他の手法を上回るか同等の認識精度を達成することが確認できた。更に、**PyramidSepDrop** は一般物体認識データセット **CIFAR-10** 及び **CIFAR-100** のテストデータの分類精度において発表当時の主なニューラルネットワークの分類精度を上回り、最高精度を達成した。

第 4 章では、あらゆる **Residual Block** で効果的な確率的正則化について検討する。具体的には複数のネットワーク構造を対象とした汎用的な確率的正則化を目指し、事前実験を通して複数の構造において特に有効な提案手法 **ShakeDrop** の条件を探索し、更に本実験として一般物体認識用データセット **CIFAR-10**, **CIFAR-100**, **ImageNet**, **COCO** を用いた実験によって、提案手法 **ShakeDrop** の有効性を確認した。その結果、多くの実験において提案手法 **ShakeDrop** は他の手法を上回る認識精度を達成した。実験の中で **PyramidNet** に提案手法を適用することで **CIFAR-10** および **CIFAR-100** のテストデータの分類精度において発表当時の最高精度を達成した。本論文執筆時点においても提案手法 **ShakeDrop** は他の手法が **CIFAR-10** および **CIFAR-100** の最高精度を達成する際に用いられている。

最後に、第 5 章では本研究によって得られた知見をまとめ、今後の課題について述べる。

審査結果の要旨

本論文は、ここ数年、学界および産業界に大きな影響を与えてきたニューラルネットワークにおいて、その性能の向上に寄与する確率的正則化と呼ばれる技術についてまとめたものであり、以下の成果を得ている。

- (1) 特定のニューラルネットワークのために考案された確率的正則化手法を別の種類のニューラルネットワークに適用できることを実験により示した。さらに、ネットワーク構造の特徴を考慮することで、より性能の良い手法を提案した。提案した確率的正則化を用いることで、画像に写る物体の名前を言い当てる一般物体認識と呼ばれる問題において、当時の最高精度を達成した。
- (2) 効果が大きいものの、特定のニューラルネットワークにしか適用できない既存の確率的正則化手法に着目し、その要素を取り入れた新たな確率的正則化を提案した。そして、この確率的正則化が複数のニューラルネットワークに適用でき、より大きな効果が得られることを実験により示した。提案した確率的正則化を用いることで、上記(1)と同じ一般物体認識の問題において、当時の最高精度を達成した。

以上の研究成果は、従来より高い精度を達成するニューラルネットワークの構築を可能とするとともに、本分野の学術的・産業的な発展に寄与するところが大きい。また、申請者が自立して研究活動を行うに必要な能力と学識を有することを証したものである。