

氏名	えぐち しゅう 江口 脩
学位の種類	博士（理学）
報告番号	甲第 1905 号
学位授与の日付	令和 4 年 3 月 17 日
学位授与の要件	学位規則第 4 条第 1 項該当（課程博士）
学位論文題目	Output augmentation:A novel method of data augmentation (出力拡張：データ拡張の新しい手法)
論文審査委員	(主査) 福岡大学 教授 藤木 淳 (副査) 福岡大学 准教授 天羽 隆史 岡山理科大学 教授 青山 崇洋

## 内容の要旨

現在、ニュースなどで耳にすることが多くなった「人工知能 (AI)」に関する研究は歴史が長く、今までに二度のブームが起こったが実用的なレベルに達することができずにブームは終焉をむかえた。近年、計算機 (GPU) の性能の向上とクラウドの発達により大量のデータの入手が容易になったことで、人工知能の一種であるディープラーニングを用いた研究が盛んになり三度目のブームをむかえている。そのディープラーニングは、生物の脳の神経細胞 (ニューロン) をモデル化したニューラルネットワークを何層も階層的につなげた大量の計算が必要なモデルである。ディープラーニングを用いた研究が盛んになった理由の一つとしては、モデルへ与える特徴量を人間が考える必要がないことが挙げられる。すなわち、ディープラーニングは与えられたトレーニングデータから人間では理解できないような特徴量を自ら身つけ出す。ディープラーニングの中でも特に、画像データに対して畳み込みニューラルネットワークを用いた研究は、様々な分野において人間に劣らないレベルにまで性能が向上している。例えば、医療画像の診断をディープラーニングに行わせた場合、人間よりも高い精度で識別できたという報告が挙げられている。

一方で、畳み込みニューラルネットワークを用いて高い性能を得るためには多くのトレーニングデータが必要なことも知られており、十分なトレーニングデータは得られることが少ない。先ほど例で挙げた医療画像も症例数がない、または画像の権利の問題で利用が困難なことでディープラーニングの学習に必要なトレーニングデータが十分でない場合がある。

そこで一般的に用いられるのが「データ拡張」と呼ばれる手法である。データ拡張とは、手元にあるトレーニングデータを元に様々なバリエーションのデータを生成する手法で

ある。

データ拡張されたトレーニングデータは、元のトレーニングデータから少しだけずれたトレーニングデータになっているため、そのニューラルネットワークの出力も元のデータの出力から少しだけずれたものになる。データ拡張を用いることで、トレーニングデータを水増しすることで不足していたトレーニングデータが補えたり、トレーニングデータが多い場合でもバリエーションを増やすこと偏っていたトレーニングデータのバランスを調整したりすることでテストデータに対する性能の向上が期待できる。例えば、物体をある角度から撮影した画像を回転させたり歪ませたりすることで、様々な角度から撮影したかのような画像を生成することでモデルを様々な画像データに対応できるように調整することができる。

本論文では、データ拡張を行わずにテストデータに対する性能の向上を目的とする新しいアプローチである「Output Augmentation」を提案する。Output Augmentation は、ニューラルネットワークに入力されたデータから得られた出力を元に複数の出力を生成する。この拡張された出力は、本来ならば元のトレーニングデータを拡張されたトレーニングデータから得られるはずのものと考えることができる。まとめると、Output Augmentation は元のデータを与えたニューラルネットワークの出力を拡張することで、データ拡張されたトレーニングデータを与えずにその出力を得られたかのような状態を作り出す手法である。この手法のメリットは二つある。一つ目は、データ拡張と比べてニューラルネットワークにデータを入力する回数が少ないことである。二つ目は、データ拡張は用いるトレーニングデータの特徴を理解してデータ拡張を行う必要があるが、Output Augmentation はその知識が必要ないことである。

次に、Output Augmentation を用いてコンピュータビジョンの研究に広く用いられている日常的に見かけるような物体が写っているデータセットである CIFAR-10（10 種類の分類問題）、CIFAR-100（100 種類の分類問題）に対する画像分類の実験を行った。具体的には、データ拡張をしていないモデル、データ拡張（Auto-Augment）を用いたモデル、提案する手法を用いたモデルを使った実験を行い、画像分類の汎化性能を向上させることを、厳密な実証研究によって示し、次の二つの結果が得られた。一つ目は、データ拡張していないモデルに対して、Output Augmentation はテストデータに対する性能が大幅に向上した。二つ目は、データ拡張したモデルに対して、Output Augmentation はテストデータに対する性能があまり向上しなかった。

## 審査の結果の要旨

本博士学位申請論文は、パターン認識において、深層学習を用いた分類問題において識別精度を高めるために必要なデータ収集の手間を削減することを目的とし、追加のデータ収集を行わずに識別精度を高めることを実現したものであり、その結果、深層学習への参入の障壁を低くすることが期待される画期的かつ重要なものである。

2012年に深層学習、つまり層の数が多いニューラルネットワークを用いた機械学習が再評価されて以来現在まで、人工知能の実現に向けて深層学習を用いてパターン認識を行う研究が盛んに行われている。分類問題に深層学習を用いることの利点は、学習データ、すなわち、解きたい問題に対する入力とその入力に対して望まれる出力の組が十分な数だけ与えられれば、従来研究されてきたような分類問題を適切に解くために必要であった学習データの特徴を理解して入力データを加工する‘特徴抽出’という過程を必要とせず、分類問題の複雑さに見合うだけの深いニューラルネットワークさえ用意すれば、‘特徴抽出’に相当することがニューラルネットワークの内部で自動的に行われると考えられており、その結果、高い精度でその問題に対する適切な解答が得られるということである。つまり、深層学習の利点を活かすためには、十分多い学習データと十分深いニューラルネットワークを用意することが必要である。

その際、十分深いニューラルネットワークを設計するのはそれほど困難ではない。しかし、学習データを集めるためには集めるデータの個数に比例した時間や金銭が必要であり、その制約から、十分な数の学習データを集めるのには非常に困難をとまなうことが知られている。そこで、既存研究においては、例えば文字を少し斜めに書いても同じ文字であると認識されるように、入力を微小変形しても出力が変化しないことに着目することにより、現在得ている入力と出力の組を利用して、入力を微小変形することによって新しい学習データを人工的に作成して、十分な数だけの学習データを得ようとする試み、すなわちデータ拡張が行われてきた。

本論文ではニューラルネットワークが入力から出力への写像であることから、入力の微小変形が出力の微小変形として伝達されることに着目し、出力の微小変形による不変性を利用して、通常データ拡張のように学習データの数を実際に増やすことはせずに、あたかも学習データの数が増えたのと同じ働きをさせる出力拡張という、これまででない新しい考え方を提案し、実際の画像分類問題に出力拡張を適用することによって提案手法の有効性を示した。

本論文の構成は次の通りである。第1章に研究の背景と動機を述べて機械学習における本論文の位置付けを明確にした。第2章で先行研究であるデータ拡張について説明し、第3章においてニューラルネットワークとその学習方法について説明した。第4章において提案手法である出力拡張について詳細に述べ、第5章で提案手法の有効性を実画像の認識問題に適用することによって示し、第6章で本論文の結論を述べた。

上述のように、本論文では、深層学習の分類精度向上のために必要な十分な学習データの個数を、実際には増やさずに、あたかも増やしたのと同様な効果を得る手法を提案しており、深層学習を利用するために敷居を低くするという点で重要な知見を与えたと評価することができる。

本審査の内容は、令和4年2月1日午前10時より、理学部9号館941教室において、さらに新型コロナウイルス感染防止のため遠隔視聴するためWebexを用いて、公聴会形式にて発表された。公聴会において、会場における参加者は6名、Webexにおける参加者は17名、

合計 23 名が参加し、質疑応答が行われた。質疑においては提案手法で用いた出力拡張における出力の変形の限界など、本論文の内容を中心とした試問を行なったが、試問に対して適切に答えることができ、研究分野およびその周辺分野に関する知識や理解を十分に有することが認められた。

最終的には、主査1名および副査2名により審査され、全員一致で合格とし、本論文を学位論文に値するものと認め、博士(理学)の学位を授与するに値するものと認められた。