

令和 2 年 4 月 24 日現在

機関番号：24403

研究種目：挑戦的研究（萌芽）

研究期間：2018～2019

課題番号：18K19785

研究課題名（和文）学習の阻害による深層学習の学習促進

研究課題名（英文）Promotion of learning process of deep learning by interference

研究代表者

岩村 雅一（Iwamura, Masakazu）

大阪府立大学・工学（系）研究科（研究院）・准教授

研究者番号：80361129

交付決定額（研究期間全体）：（直接経費） 4,900,000円

研究成果の概要（和文）：本研究では、我々が提案した「学習の阻害」を利用して物体認識（画像分類）の認識精度を向上させる正則化手法であるShakeDropについて、(1)学習能力の向上、(2)学習に使用するサンプル数を減らす方法の開発、(3)原理の解明を実施した。(1)と(2)については、様々な実験を通して有効性を検証した。(3)については、ShakeDropの「学習の阻害による学習の促進」という現象が特徴空間内でのデータ拡張であるという解釈を得た。

研究成果の学術的意義や社会的意義

ShakeDropは、物体認識のためのデータベースであるCIFAR-100において、一時世界最高精度を達成した正則化手法である。現在は他の手法がより良い精度を達成しているが、現在最高精度を達成している手法も最高精度を達成するためにShakeDropを使用している。本研究では、ShakeDropのメカニズムを解明し、更にShakeDropが前述のCIFAR-100データベース以外においても高い認識精度を達成できることを実験的に示した。

研究成果の概要（英文）：In this research, on the regularization method we have proposed and named ShakeDrop which improves the accuracy of object recognition (aka, image classification) using "interference of learning process," we have completed (1) improvement of learning ability, (2) development of a method that requires less training data, and (3) unraveling the mechanism. Through various experiments, we have shown (1) and (2) hold. Regarding (3), we have got an explanation that "promotion of learning process of deep learning by interference" is achieved by data augmentation in the feature space.

研究分野：パターン認識

キーワード：深層学習 物体認識 学習の阻害 正則化

## 様式 C - 19、F - 19 - 1、Z - 19 (共通)

### 1. 研究開始当初の背景

深層学習（ディープラーニング；従来よりも多層のニューラルネットワークを使った学習により、様々な分野で従来手法を越える性能を発揮）の発展には目覚ましいものがある。人と同等以上の画像認識能力を持つニューラルネットワークが報告されたり、囲碁や将棋で機械が人に勝ったりと、これまでは実現が難しい、または実現に時間がかかると考えられてきた事が深層学習を使って次々に実現されている。さらに、近い将来には機械が人の職業を奪うと予想され、また少し遠い将来（25～30年後）には機械が人の能力を完全に越える（シンギュラリティ）と考える研究者もいて、深層学習は我々の生活に大きな影響を及ぼし得る技術として、社会的に極めて高い関心を集める。社会の各所に眠っている潜在的なビッグデータの利活用により、新たな知見の獲得を目指すという社会的ニーズの高まりや、モバイルデバイスの小型化と普及によるIoT（Internet of Things）の進展もこの傾向に拍車を掛け、深層学習への期待は膨らむばかりである。

これらの関心の源泉は、深層学習が持つ「データから自動で学習する」能力への期待である。画像認識の分野では、深層学習は2012年頃に表舞台に登場し、コンペティションで従来手法に大差を付ける圧倒的な性能差を見せつけた。この性能差は、認識に効果的な画像特徴を、研究者が「経験と勘」で設計する従来の方式から、深層学習を用いてデータから自動で抽出する方式に変更したことで生じた。その後の深層学習の快進撃も「データからの自動学習」に依るところが大きい。この図式は画像認識分野以外にも当てはまる。このように、深層学習の学習能力は、人々が深層学習に期待する最たるものであるため、その向上は深層学習研究における中心的課題である。その一方で、深層学習の強力な学習能力を発揮するには大量の「ラベル付きデータ」が必要という難点もある。

### 2. 研究の目的

本研究では、今後深層学習の性能を飛躍的に向上し得る、「学習の阻害による学習の促進」という、従来の常識とは異なるアイデアに焦点を当てる。これは、ニューラルネットワークの学習を敢えて邪魔することで、邪魔しないときよりも高い学習効果が得られるという、一見矛盾しているかのような方法である。この方法は、2017年4月に最初に提案され、原理は未解明ながら最近の深層学習の性能向上に大きな役割を果たしている。本研究はこのアイデアを中心に据えて、以下の3つの課題に貢献する。

- (1) 深層学習の学習能力の向上
- (2) ラベル付きデータの収集コスト削減に繋がる、少ないデータからの高精度な学習の実現
- (3) それらの原理の解明

### 3. 研究の方法

「学習の阻害による学習の促進」というアイデアは、2017年4月の国際会議で発表された Shake-Shake という手法で最初に提案された[1]。2017年10月には、このアイデアを発展させることで申請者らは ShakeDrop という手法を提案し、2度目の世界一の認識精度を達成した[2]。本研究では、この ShakeDrop に関する研究を通して、前述の目的の達成を目指した。具体的には、本研究の提案段階では計算機資源の関係で小規模なデータセットを用いた実験しかできていなかったため、当該分野で標準的に使用されている大規模なデータセットを用いて ShakeDrop の性能評価実験を行った。その際、どのようなパラメータが有用であるかや、ShakeDrop を適用するベースネットワークの層の深さなどが変化すると ShakeDrop の効果がどう変わるかなどを詳細に調べた。更に、「学習の阻害による学習の促進」がどうして深層学習に有効なのかが未解明であったので、この方法が効果を発揮するメカニズムの解明にも取り組んだ。

### 4. 研究成果

ShakeDrop の有効性については、従来は CIFAR-10/100 という比較的小規模なデータセットにおいてのみの実験をしていたが、当該分野で標準的に使われている ImageNet というデータセットでの実験を実施して、提案手法の有効性を確認した。その際、CIFAR-100 データセットを用いた実験で、ShakeDrop が効果的に働くパラメータセットを2つあることが分かった。これまでは、ShakeDrop が効果的に働くパラメータセットを1つだけを発見しており、実験結果から、ネットワークの Residual Block と呼ばれる処理ブロックの最後に Rectified Linear Unit (ReLU) がある場合には ShakeDrop は有効に働かないと考えていた。しかし、新しく発見したパラメータセットを使うことで、最後に ReLU があっても良いということがわかった。これらのパラメータは、ImageNet の実験でも有効であったため、これらのパラメータセットは汎用的に効果を発揮する可能性が高いと考えている。また、提案手法は層が深いベースネットワークに適用した方がより高い性能を発揮することが実験的に確認した。これは、ShakeDrop の元になっている正則化手法の1つである RandomDrop と類似の傾向である。更に、物体検出ならびにインスタンス・セグメンテーションを対象として、これらのタスクで標準的に用いられる MS COCO データセットでの実験を通して有効性を確認した。これらの実験結果の詳細については、[3]を参照されたい。

「学習の阻害による学習の促進」の原理の解明については、以下の解釈に至った。ShakeDrop は、従来手法である Shake-Shake と同様に、ニューラルネットワークのフォワードパスとバックワードパスで異なる乱数を用いて学習を阻害する。最近の研究成果（例えば[4]）によって得ら

れた知見から、いずれの手法もフォワードパスでは特徴に乱数による外乱を加えることで新たなデータを合成していると考えられる。また、ShakeDrop は外乱を与えるのにニューラルネットワークの1本の分岐しか使わないために様々なネットワークに適用可能である反面、2本の分岐を用いる Shake-Shake の方が提案手法よりも処理が安定しているという、実験結果に基づく解釈が得られた。ShakeDrop は1本の分岐のみを用いるために、本来不安定であるが、前述の RandomDrop を確率的に2つのネットワークを切り替える確率的なスイッチとして使用している。これは通常とは異なる用途である。このことを概念的に表したのが図1である。図1の青色の矢印は、ベースネットワーク(分岐1本を仮定)の学習過程を表している。ベースネットワークは正しく学習するが、局所解に陥っている。しかし、緑色の矢印のように強い摂動を加えると、摂動が強すぎるため、正しく学習が進まない。そこで、ShakeDrop では、青色の正しい学習をするベースネットワークと緑色の(強い摂動により)誤った学習をするネットワークを確率的に切り替えることにより、青色のベースネットワークでは到達できなかった、より良い解に到達できることが期待できる。

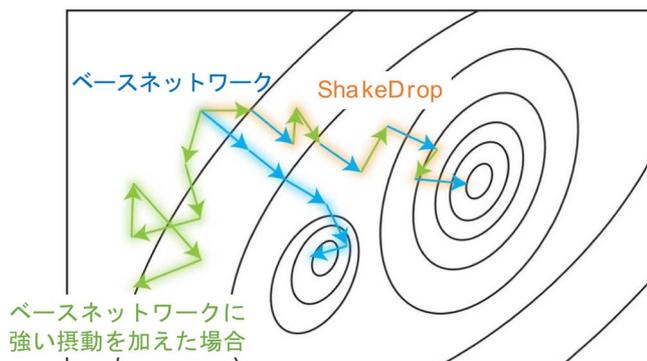


図1 学習過程の概念図

実施内容を当初の研究目的に照らすと、まず、(1)の同一のデータセットを用いた場合の学習能力の向上が、物体認識の大規模実験や他のタスクの実験を通して確認できた。また、これは相対的に(2)の同じ認識精度を達成するために必要な学習サンプル数の減少を意味する。そして、ShakeDrop の「学習の阻害による学習の促進」という現象は、特徴空間内でのデータ拡張であるという解釈を得ることができたため、(3)の原理の解明についても概ね達成することができた。したがって、本研究は当初の目的を達成できた。

最後に、本研究の波及効果について言及する。

- **受賞**：本研究の開始前に
  - ✓ 電子情報通信学会 パターン認識・メディア理解研究会 2017 年度 研究奨励賞を受賞していた(筆頭著者の山田良博が受賞者)。それに加えて、研究期間中に
  - ✓ 第21回 画像の認識・理解シンポジウム(MIRU2018) MIRU 学生優秀賞を受賞した。
  
- **引用**：ShakeDrop の最初の英語の文献は、2018 年 2 月にプレプリントサーバ arXiv に投稿したものである[5]。Google Scholar によると、この文献は現在までに 70 件引用されている。
  
- **有効性**：図2に一般物体認識データベースである CIFAR-100 の認識率の推移を示す。ShakeDrop は、一般物体認識データベースである CIFAR-100 において、一時世界最高精度を達成した。現在は他の手法がより良い精度を達成しているが、最近最高精度を達成した手法[6, 7]も最高精度を達成するために ShakeDrop を使用している。

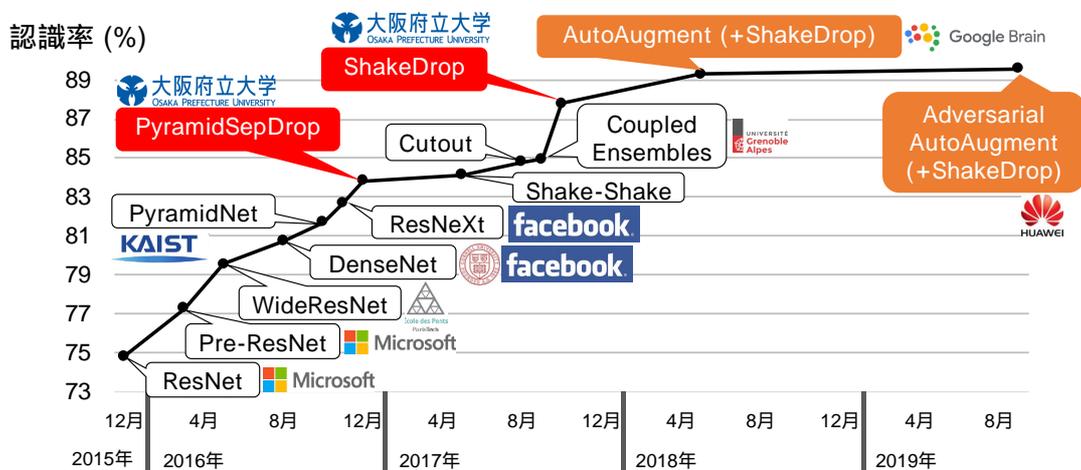


図2 一般物体認識の認識率の推移

#### 参考文献

- [1] X. Gastaldi, “Shake-Shake regularization,” Proc. International Conference on Learning Representations (ICLR) Workshop, 2017.  
<https://openreview.net/forum?id=Hk0-PCmYI>
- [2] Y. Yamada, M. Iwamura, K. Kise, “Shake-Drop Regularization,” Proc. International Conference on Learning Representations (ICLR) Workshop, 2018.  
<https://openreview.net/forum?id=Bymu6tJwz>
- [3] Y. Yamada, M. Iwamura, T. Akiba, and K. Kise, “ShakeDrop Regularization for Deep Residual Learning,” IEEE Access, vol. 7, pp. 186126-186136, 2019.  
 DOI: [10.1109/access.2019.2960566](https://doi.org/10.1109/access.2019.2960566).
- [4] T. DeVries and G. W. Taylor, “Dataset Augmentation in Feature Space,” in Proc. International Conference on Learning Representations (ICLR) Workshop, 2017.  
<https://openreview.net/forum?id=HyaF53XYx>
- [5] Y. Yamada, M. Iwamura, K. Kise, “Shake-Drop Regularization,” arXiv preprint arXiv:1802.02375, 2018.  
<https://arxiv.org/abs/1802.02375>
- [6] E. D. Cubuk, B. Zoph, D. Mane, V. Vasudevan, and Q. V. Le, “AutoAugment: Learning Augmentation Policies from Data,” Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019.  
 DOI: [10.1109/CVPR.2019.00020](https://doi.org/10.1109/CVPR.2019.00020)
- [7] X. Zhang, Q. Wang, J. Zhang, Z. Zhong, “Adversarial AutoAugment,” Proc. International Conference on Learning Representations (ICLR), 2020.  
<https://openreview.net/forum?id=ByxdUySKvS>

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件 / うち国際共著 0件 / うちオープンアクセス 1件）

1. 著者名 Yamada Yoshihiro, Iwamura Masakazu, Akiba Takuya, Kise Koichi	4. 巻 7
2. 論文標題 Shakedrop Regularization for Deep Residual Learning	5. 発行年 2019年
3. 雑誌名 IEEE Access	6. 最初と最後の頁 186126 ~ 186136
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/ACCESS.2019.2960566	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計2件（うち招待講演 0件 / うち国際学会 1件）

1. 発表者名 Yoshihiro Yamada, Masakazu Iwamura, Koichi Kise
2. 発表標題 ShakeDrop Regularization
3. 学会等名 6th International Conference on Learning Representation (ICLR) Workshop (国際学会)
4. 発表年 2018年

1. 発表者名 山田良博, 岩村雅一, 黄瀬浩一
2. 発表標題 ResNetsに対する新たな正則化手法ShakeDropの提案
3. 学会等名 第21回画像の認識・理解シンポジウム(MIRU2018)
4. 発表年 2018年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

2018年8月8日に、第21回 画像の認識・理解シンポジウム(MIRU2018)においてMIRU学生優秀賞を受賞した。この賞は、MIRUにおける筆頭発表者が投稿時点ならびに発表時点ともに学生である口頭研究発表のうち、優秀と認められる研究を選び、その発表者(連名者を含む)に贈呈されるものである。

受賞論文: ResNetsに対する新たな正則化手法ShakeDropの提案

受賞者: 山田良博, 岩村雅一, 黄瀬浩一

関連URL: <https://sites.google.com/view/miru2018sapporo/%E3%83%9B%E3%83%BC%E3%83%A0/%E8%A1%A8%E5%BD%B0>

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----