

令和 6 年 6 月 26 日現在

機関番号：31302

研究種目：基盤研究(C)（一般）

研究期間：2020～2023

課題番号：20K11871

研究課題名（和文）グリッドニューラルネットワークを用いた小規模データセットのための画像分類手法

研究課題名（英文）Image Classification Models based on Grid Neural Networks for Small Datasets

研究代表者

武田 敦志（TAKEDA, Atsushi）

東北学院大学・情報学部・教授

研究者番号：90424001

交付決定額（研究期間全体）：（直接経費） 3,100,000円

研究成果の概要（和文）：本研究課題では、小規模データセットを用いた場合でも高い画像認識性能を達成する画像認識モデルの実現を目指し、グリッドニューラルネットワークの構想を一般化して構成したニューラルネットワークであるSkipResNetを開発した。SkipResNetは、本研究課題で開発したGate Moduleを導入した画像認識のためのニューラルネットワークであり、計算経路を動的に変更できる仕組みを有している。標準的な画像データセットを用いた性能評価により、SkipResNetの画像認識性能は従来手法よりも高く、小規模なデータセットを用いた転移学習であっても高い画像認識性能を実現できることを確認した。

研究成果の学術的意義や社会的意義

本研究課題では、画像認識のための代表的な多層ニューラルネットワークの動作解析を行い、その解析結果にもとづき、Gate Moduleを導入した新しい多層ニューラルネットワークを開発した。標準的な画像認識のためのデータセットを用いた性能評価により、本研究課題で開発した多層ニューラルネットワークは従来手法よりも高い画像認識性能を発揮することを確認した。また、これらの多層ニューラルネットワークをbackboneとした画像認識モデルが従来手法よりも高い画像認識性能を示すことを確認した。

研究成果の概要（英文）：In this research project, aiming to realize an image recognition model that achieves high image recognition performance even when using small data sets, we developed SkipResNet, which is a neural network composed of a generalization of the grid neural network concept. SkipResNet is a neural network for image recognition that includes the Gate Modules developed in this research project, and has a mechanism that can dynamically change the calculation path. Performance evaluation using standard image datasets confirmed that the image recognition performance of SkipResNet is higher than that of existing models and that high image recognition performance can be achieved even with transition learning using small datasets.

研究分野：画像認識

キーワード：画像認識 深層学習 ニューラルネットワーク

## 様式 C-19、F-19-1 (共通)

### 1. 研究開始当初の背景

多層ニューラルネットワークを用いた画像認識技術に関する画期的な成果が多数報告されている。これらの画像技術には自動運転・ロボット制御・製品検査・異常検知などの様々な応用が考えられるため、多層ニューラルネットワークの画像認識性能の改善は今後の産業の発展にとって必要不可欠な課題である。一方、多層ニューラルネットワークのパラメータを調整するためには大規模なデータセットが必要となる。しかし、現実に存在する多くの課題では大量の訓練用データを収集することが困難であり、これらの課題に対して有効な多層ニューラルネットワークを実現することは容易ではない。そのため、小規模なデータセットに対しても高い画像認識性能を発揮する多層ニューラルネットワークを実現する手法を開発する必要がある。

訓練データが小規模であっても高い画像認識性能を有する多層ニューラルネットワークを実現する手法として転移学習と呼ばれる手法がある。転移学習は、あらかじめ大規模データセットを用いてパラメータを調整した多層ニューラルネットワークを作成しておき、この訓練済みの多層ニューラルネットワークに対して新たな訓練データを学習させる手法である。転移学習を用いることにより、少ない訓練データであっても高い画像認識性能の多層ニューラルネットワークを実現することが可能となるため、画像に含まれる物体の検知や領域の識別などの複雑な画像認識問題に対応できると考えられる。そのため、画像認識性能が高く、かつ、効果的に転移学習を行うことができる多層ニューラルネットワークの実現が求められている。

### 2. 研究の目的

転移学習では、大規模データセットを用いて学習を行った多層ニューラルネットワーク (backbone) として、VGGNet や ResNet などの多層ニューラルネットワークを用いることが多い。一方、近年の研究により、backbone として用いる多層ニューラルネットワークの構造を変更すると転移学習の性能も大きく変化することが判明している。そこで、本研究課題では、転移学習の backbone として従来手法よりも適切な多層ニューラルネットワークを実現することを目的とする。具体的には、高い汎化性能を有する多層ニューラルネットワークである「グリッドニューラルネットワーク」を backbone とした転移学習を行い、グリッドニューラルネットワークを用いることで転移学習後の多層ニューラルネットワークの画像分類性能が既存手法よりも改善することを検証する。これにより、訓練データが少ない場合であっても高い画像分類性能を発揮する多層ニューラルネットワークを実現する。

### 3. 研究の方法

まず、大規模データセットを用いて訓練した多層ニューラルネットワークの動作解析を行い、画像認識問題で利用される多層ニューラルネットワークの動作特性を明らかにする。具体的には、ResNet・ResNeXt・EfficientNet などの代表的な多層ニューラルネットワークの動作を解析することで、多層ニューラルネットワークを順伝播する信号や逆伝播する誤差の特性を明らかにする。次に、この解析結果をもとに、入力信号や誤差信号を効率的に伝播する多層ニューラルネットワークを設計する。また、標準的な画像分類問題のためのデータセットである ImageNet を用いて新たに設計した多層ニューラルネットワークの性能を評価し、この研究課題で開発した多層ニューラルネットワークの画像分類性能が従来手法よりも高いことを確認する。さらに、標準的な物体検知問題のためのデータセットである MS-COCO や領域分割問題のためのデータセットである ADE20k を用いて、新たに設計した多層ニューラルネットワークを backbone とした物体検知モデルや領域分割モデルの画像認識性能を検証する。この性能検証により、この研究課題で開発した多層ニューラルネットワークを backbone として採用することにより、従来手法よりも効果的な転移学習を実現できることを確認する。

### 4. 研究成果

ImageNet などの画像分類データセットを用いて ResNet・ResNeXt・EfficientNet などの代表的な多層ニューラルネットワークの動作を解析したところ、入力信号や誤差信号がすべての畳み込み層に対して安定的に伝播していることや、入力信号が異なると計算に使用される畳み込み層が異なることが判明した。これらの特徴が ResNet などの多層ニューラルネットワークが画像認識問題に対して高い性能を発揮する理由であると考え、これらの特徴をより活用するための仕組みとして新しい注意機構である Gate Module を開発した。多層ニューラルネットワークに対して Gate Module を導入することにより、入力信号によって多層ニューラルネットワークの構造を動的に変更し、入力画像を認識するために必要となる畳み込み層をより重点的に使用した計算が可能となった。また、Gate Module を導入した多層ニューラルネットワークは、これまでに開発してきたグリッドニューラルネットワークに対して注意機構を導入した場合と同じ構造となっており、複数の計算経路が混在する多層ニューラルネットワークとなっているため、従来の多層ニューラルネットワークよりも汎化性能が高い構造となる。

ResNet や ResNeXt に Gate Module を導入した多層ニューラルネットワークとして SkipResNet

や SkipResNeXt を実装した。また、最新の画像分類のための多層ニューラルネットワークである ConvNeXt に Gate Module を導入した多層ニューラルネットワークとして SkipConvNeXt を実装した。これらの多層ニューラルネットワークに対して標準的な画像分類データセットである ImageNet の学習データを用いた学習を行い、ImageNet のテストデータを用いてそれぞれの多層ニューラルネットワークの画像分類性能を評価した。表 1 にそれぞれの多層ニューラルネットワークの画像分類性能の比較を示す。この研究課題で開発した SkipResNet・SkipResNeXt・SkipConvNext は従来手法である ResNet・ResNeXt・ConvNeXt よりも高い画像分類性能を示した。また、ImageNet を学習した ResNet が誤分類する画像で構成されるデータセットである ImageNet-A を用いた性能評価では、Gate Module を導入することによって多層ニューラルネットワークの汎化性能が向上することを確認した。

SkipResNet や SkipResNeXt を backbone とした多層ニューラルネットワークに対して転移学習を行ったときの画像認識性能について、標準的な物体検知問題のデータセットである MS-COCO と領域分割問題のデータセットである ADE20k を用いた性能評価を行った。表 2 に SkipResNet や SkipResNeXt を backbone とした物体検知モデルの画像認識性能の比較を示し、表 3 に SkipResNet や SkipResNeXt を backbone とした領域分割モデルの画像認識性能の比較を示す。物体検知問題と領域分割問題のいずれにおいても、Gate Module を導入した多層ニューラルネットワークを backbone とした画像認識モデルは従来手法よりも高い画像認識性能を発揮した。これは、Gate Module を導入した多層ニューラルネットワークは、従来の多層ニューラルネットワークよりも高い汎化性能と高い表現力を有しており、転移学習のための backbone として従来手法よりも適しているためと考えられる。

本研究課題では、画像認識のための代表的な多層ニューラルネットワークである ResNet や ResNeXt などの動作解析を行い、その解析結果にもとづき、Gate Module を導入した SkipResNet や SkipResNeXt を開発した。標準的な画像認識のためのデータセットを用いた性能評価により、本研究課題で開発した SkipResNet や SkipResNeXt は従来手法よりも高い画像認識性能を発揮することを確認した。また、これらの多層ニューラルネットワークを backbone とした画像認識モデルが従来手法よりも高い画像認識性能を示すことを確認した。

本研究課題の成果について、研究会等において成果を発表し、GitHub にて実装を公開した (<https://github.com/takedarts/skipresnet>)。ただし、2020 年からの世界的な感染症の流行や大学組織の変更による研究環境の変化により、研究成果の発表は当初計画より大幅に遅れた。そのため、今後は本研究課題の成果をまとめた学術論文を執筆して発表する予定である。

表 1 ImageNet を用いて学習した画像分類モデル性能比較。それぞれの数値は、ImageNet の訓練データを用いて 3 個ずつの画像分類モデルを作成し、それらの画像分類モデルについて ImageNet のテストデータに対する正解率を計測したときの中央値である。

model	top-1 acc. (224x224)	top-1 acc. (256x256)	top-1 acc. (288x288)	top-1 acc. (320x320)
ResNet-34	0.7553	0.7622	0.7654	0.7665
<b>SkipResNet-34</b>	0.7675	0.7759	0.7778	<b>0.7782</b>
ResNet-50	0.7901	0.7953	0.7964	0.7954
SE-ResNet-50	0.7991	0.8055	0.8081	0.8072
<b>SkipResNet-50</b>	0.8041	0.8103	<b>0.8120</b>	0.8104
ResNet-101	0.8036	0.8100	0.8095	0.8086
SE-ResNet-101	0.8102	0.8157	0.8184	0.8177
<b>SkipResNet-101</b>	0.8139	0.8217	<b>0.8234</b>	0.8208
ResNeXt-50-32x4d	0.7973	0.8015	0.8030	0.8011
SE-ResNeXt-50-32x4d	0.8041	0.8093	0.8117	0.8110
<b>SkipResNeXt-50-32x4d</b>	0.8067	0.8125	<b>0.8131</b>	0.8126
ResNeXt-101-32x4d	0.8066	0.8102	0.8112	0.8101
SE-ResNeXt-101-32x4d	0.8085	0.8137	0.8165	0.8188
<b>SkipResNeXt-101-32x4d</b>	0.8139	0.8203	<b>0.8216</b>	0.8210
ConvNeXt-T	0.8157	0.8171	0.8157	0.8094
<b>SkipConvNeXt-T</b>	0.8158	0.8205	<b>0.8224</b>	<b>0.8224</b>
ConvNeXt-S	0.8314	0.8344	0.8341	0.8307
<b>SkipConvNeXt-S</b>	0.8333	0.8367	0.8374	<b>0.8375</b>
ConvNeXt-B	0.8355	0.8376	0.8372	0.8350
<b>SkipConvNeXt-B</b>	0.8337	0.8391	0.8399	<b>0.8400</b>

表 2 ImageNet を用いて学習した画像分類モデルを backbone とする画像検知モデルを MS-COCO の学習データを用いて転移学習したときの画像認識性能の比較。

Model	Backbone	mAP (box)	mAP (seg)
Cascade Mask RCNN	ResNet-50	0.44806	0.38977
	Skip-ResNet-50	0.47716	0.41421
Cascade Mask RCNN	ResNet-101	0.44813	0.38817
	Skip-ResNet-101	0.48224	0.42012
Cascade Mask RCNN	ResNeXt-50-32x4d	0.45450	0.39470
	Skip-ResNeXt-50-32x4d	0.47445	0.41304
Cascade Mask RCNN	ResNeXt-101-32x4d	0.46085	0.39817
	Skip-ResNeXt-101-32x4d	0.48862	0.42241

表 3 ImageNet を用いて学習した画像分類モデルを backbone とする領域分割モデルを ADE20k の学習データを用いて転移学習したときの画像認識性能の比較。

Model	Backbone	aAcc	mIoU
UPerNet	ResNet-50	0.8019	0.4194
	Skip-ResNet-50	0.8016	0.4244
UPerNet	ResNet-101	0.8101	0.4402
	Skip-ResNet-101	0.8138	0.4443
UPerNet	ResNeXt-50-32x4d	0.8036	0.4269
	Skip-ResNeXt-50-32x4d	0.8074	0.4339
UPerNet	ResNeXt-101-32x4d	0.8114	0.4407
	Skip-ResNeXt-101-32x4d	0.8185	0.4565

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 Takeda Atsushi	4. 巻 526
2. 論文標題 Detection and Analysis of Intrusion Attacks Using Deep Neural Networks	5. 発行年 2022年
3. 雑誌名 Lecture Notes in Networks and Systems	6. 最初と最後の頁 258 ~ 266
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-031-14314-4_26	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Takeda Atsushi, Nagasawa Daichi	4. 巻 1
2. 論文標題 A Simple Deep Learning Approach for Intrusion Detection System	5. 発行年 2021年
3. 雑誌名 Proceedings of the 13th International Conference on Mobile Computing and Ubiquitous Network (ICMU)	6. 最初と最後の頁 1-2
掲載論文のDOI (デジタルオブジェクト識別子) 10.23919/ICMU50196.2021.9638850	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	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 0件）

1. 発表者名 武田 敦志
2. 発表標題 Gate Moduleを導入した画像認識のためのResidual Neural Networkの提案と評価
3. 学会等名 The 25th Meeting on Image Recognition and Understanding (MIRU2022)
4. 発表年 2022年

1. 発表者名 武田 敦志
2. 発表標題 画像分類のためのDense Residual Networkの提案
3. 学会等名 第23回 画像の認識・理解シンポジウム
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

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
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------