

深層学習法における植物識別精度改善の試み

曙 瑠美¹・藤崎 莉純夢²・中山 紘喜³・西野 友子⁴・野上 敦嗣⁵

¹非会員 北九州市立大学 国際環境工学部環境生命工学科 (〒808-0135 福岡県北九州市若松区ひびきの1-1)

E-mail: w5551001@eng.kitakyu-u.ac.jp

²非会員 前北九州市立大学 国際環境工学部環境生命工学科 (〒808-0135 福岡県北九州市若松区ひびきの1-1)

³非会員 北九州市立大学博士後期課程 国際環境工学研究科 (〒808-0135 福岡県北九州市若松区ひびきの1-1)

E-mail: y7dac004@eng.kitakyu-u.ac.jp

⁴非会員 北九州市立大学研究員 国際環境工学部環境生命工学科 (〒808-0135 福岡県北九州市若松区ひびきの1-1)

E-mail: t-nishino@kitakyu-u.ac.jp

⁵正会員 北九州市立大学教授 国際環境工学部環境生命工学科 (〒808-0135 福岡県北九州市若松区ひびきの1-1)

E-mail: nogami@kitakyu-u.ac.jp

植生調査や生き物調査において撮影画像からの種の識別は重要な要素であり、近年急速に発展している機械学習法による画像識別技術はこの分野への応用が期待されている。学習精度向上のためには訓練データの整備や、学習パラメータを試行錯誤で経験的に決定することが必要となる。本研究では、植物の花画像を対象に訓練画像の加工方法や学習パラメータの最適化を試みた。画像はImageNetから用意した。特徴的な4種の識別を基本に、56種の多種の識別、同じ科の識別としてキク科14種、ラン科84種で学習モデルを作り評価を行った。4種の識別において、パラメータの調整、訓練画像の加工、再学習処理により高い正答率を得た。しかし、56種の分類では種により正答率が大きく異なり、キク科とラン科の正答率は著しく低くなった。今後は訓練データの質も再検討する。

Key Words : deep learning, vegetation investigation, ICT, image classification

1. 緒言

植生調査や生き物調査において撮影画像からの種の識別は極めて重要である。撮影方法はITの応用が進み、高精度のデジタル画像やドローンを用いた空中撮影が広く行われているが、正確な分類には専門知識と経験と積んだ専門家が必要となる。ITによる画像認識技術の研究は自動運転車への応用も契機となっており、近年、急速に発展し¹⁾、機械学習法、特に深層学習法 (Deep Learning; 以下DLと記す) による画像識別の精度が対象により人間に匹敵するようになった。1000種類画像認識コンテストILSVRC²⁾では、2012年にトロント大のAlexNetが認識精度85%を実現してから、2014年のGoogLeNetが93.3%、2015年のResNetが96.4%を達成し、訓練した人間の精度94.9%を超えている。機械学習の開発環境は公開されており、ILSVRCで用いられている画像データベースImageNetや

機械学習ライブラリCaffe、Torch、Tensorflowに加え、GPUによるCNN高速演算ライブラリcuDNNや開発プラットフォームDIGITSもNVIDIAが無償で提供されており、機械学習DL法の応用研究の活性化に大きく貢献している。

植物分野への応用研究例は、米スミソニアン学術協会らによる葉、花、果実、種子を識別するLeafSnapや、千葉工大が開発した406種類の花を分類するハナノナなど図鑑的システムが先行して開発されている。

このように画像識別では極めて有望なDL法だが、植物の画像認識精度は17種の花画像に対してAlexNetやGeogLeNetで80%程度であり³⁾、1000種類で95%以上に達するILSVRCに比べて劣っている。学習精度向上のためには、訓練画像データの整備 (量と質) に加えて、選択した機械学習手法の様々な制御パラメータを試行錯誤で経験的に決定することが必要となる。

本研究では、ImageNetから集めた花画像を対象とし、

DLモデルはAlexNetおよびGeogLeNetを用いて科が異なる花および同じ科の花それぞれに対して、学習精度改善のための訓練画像の加工方法や学習パラメータの最適化を試みた。

2. 研究方法

(1) 基本学習性能の評価

本研究で用いた計算機の構造を図-1に示す。ハードウェア性能は、CPUにはIntel製Core i7-3960X(3.30GHz)を、GPUにはNVIDIA製GeForce GTX1080(3584コア11GB)を、用いており、計算機のメモリは64GBある。使用した機械学習ライブラリはCaffe、Torch、Tensorflowで、各ライブラリに対しAlexNetとGoogLeNetによる学習を行った。

学習させる花の画像はImageNetのflowerデータセットからダウンロードした。最初に、花の特徴が大きく異なるボタン、シクラメン、ポピー、ヒマワリの4種類の花を用いた。ダウンロードした枚数はボタン1275枚、シクラメン1449枚、ポピー1224枚、ヒマワリ1183枚であったが、このままの画像では学習率が低く、写真サイズが極端に違うものやデータが反映されていないものなど、学習の妨げになるであろう写真を削除した。削除後の写真の枚数は、ボタン1237枚、シクラメン1443枚、ポピー1192枚、ヒマワリ1008枚になり、これを用いて、機械学習ライブラリやCNNモデルの組み合わせに対して正答率や計算性能を調べた。また正答率の低い組み合わせに対しては、パラメータの調整や再学習を行い、正答率の向上を図った。

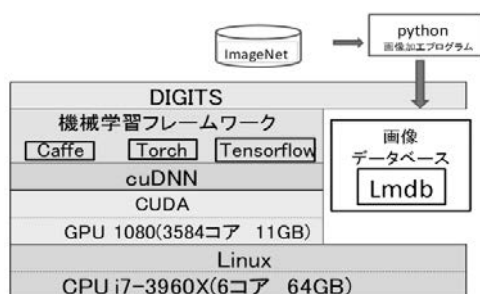


図-1 計算機の構造

(2) 多種類の花画像識別性能の評価

上記の学習では識別数が4つと少ない状態での学習であった。そこでより実践的な学習モデルを作成するため識別数を増やした状態で学習モデルを作成した。具体的にはImageNetから「flower」にカテゴリ化されている56種類の画像をダウンロードし、学習させその性能を検討した。また、同じ科に属するような見た目が似ている画

像をを識別できるのか調べるため、ImageNetからキク科のみを集めたデータセット、ラン科のみを集めたデータセットを作り、それぞれに対して機械学習を行った。

(3) 画像加工

DIGITSでは、学習をかける際に、全ての写真が、256*256に圧縮・拡大される。そのため、アスペクト比が変化しないよう画像のアスペクト比を1:1に加工するプログラムをpythonで作成し、すべての画像に適応した。以降この加工をアス比加工と呼ぶ。

3. 研究結果

(1) 基本学習性能の結果

基本学習性能の正答率を表-1に示す（ () 内は Loss の値）。ほぼすべてのパッケージで70~90%での学習で正答率を得ることができた。しかし

Torch/GoogLeNetの正答率は極端に低いものだった。

計算性能は表-2, 3のようになった。AlexNetでは、計算時間が最も短かったのはTorchであった。これはGPU、CPU使用率がどちらも高く、効率よく使えているからだと考えられる。一方CaffeではGPUもCPUも使用率が低く、TorchやTensorflowの3倍近くの時間がかかってしまった。

表-1 Imageneflower2-2の正答率

	AlexNet	GoogLeNet
Caffe	88.78% (0.312)	88.56% (0.310)
Torch	85.81% (0.398)	29.53% (1.38)
Tensorflow	85.80% (0.416)	76.06% (0.605)

表-2 計算性能 (AlexNet)

	計算時間	GPU (メモリ量)	CPU (メモリ量)
Caffe	9m15s	54%(3.04GB)	106%(2.27GB)
Torch	3m3s	97%(4.13GB)	409%(6.51GB)
Tensorflow	3m44s	80%(10.5GB)	415%(4.19GB)

表-3 計算性能 (GoogLeNet)

	計算時間	GPU (メモリ量)	CPU (メモリ量)
Caffe	9m14s	70%(4.02GB)	116%(1045GB)
Torch	11m52s	99%(5.24GB)	162%(5.22GB)
Tensorflow	19m41s	85%(10.5GB)	224%(3.59GB)

Torch/Googlenetに対して行った再学習の結果を表-4に示

す。6回目の学習で91.71%が最高値となった。このように再学習を行うことで正答率をある程度向上できることがわかった。画像にアス比加工を行い再度学習させた結果が表-5である。この結果に対しても再学習を行い、正答率の最大値とその時の学習回数を表-6にまとめた。

表-4 Torch/GoogLeNet再学習結果

学習回数	正答率(%)
1	29.53
2	56.36
3	70.3
4	86.79
5	90.73
6	91.71
7	91.55

表-5 Imageryflower2-2アス比加工後正答率

	AlexNet	GoogLeNet
Caffe	88.54% (0.313)	89.61% (0.304)
Torch	87.45% (0.368)	29.53% (1.37)
Tensorflow	84.50% (0.434)	76.96% (0.607)

表-6 アス比加工画像の再学習正答率 (学習回数)

	AlexNet	GoogLeNet
Caffe	90.06% (2回目)	91.07% (2回目)
Torch	91.63% (3回目)	94.67% (8回目)
Tensorflow	88.96% (4回目)	89.22% (11回目)

(2) 多種類の花の画像での学習

花の種類を56種類に増やしてアス比加工を行い学習させた。この56種類はImageNet上に画像が500枚以上あったものである。結果を表-7に示す。このときのTensorflow/GoogLeNetにおける各花の正答率の上位3種と下位3種を表-8, 9にまとめた。

表-7 56種類の花の正答率

	AlexNet	GoogLeNet
Caffe	60.67%	60.01%
Torch	54.36%	36.26%
Tensorflow	54.03%	62.43%

表-8 正答率上位3種

	名前	正答率(%)
1	Sunflower	86.21
2	Bellwort	85.87
3	Commelina	84.21

表-9 正答率下位3種

	名前	正答率(%)
1	Bloomer	13.49
2	Browallia	32.48
3	Daisy	33.42

また、キク科14種、ラン科61種に対し学習させた。キク科はTorch/GoogLeNetでは18.48%, Caffe/AlexNetでは17.18%, ラン科はTorch/GoogLeNetでは52.46%, Caffe/AlexNetでは1.97%とどれも低い学習率にとどまってしまった。これらの学習率の向上については今後の課題となる。

4. 考察

(1) バッチサイズの効果

DISITSではバッチサイズというものを学習の際設定することができる。バッチとはまとまりのある入力データのことで、バッチ処理をすると画像1枚当たりの処理時間を短縮する、ニューラルネットワークの計算においてデータ転送がボトルネックになる場合はバス帯域の負荷を軽減するなどといった利点がある。そこでバッチサイズを8, 16, 32にそれぞれ設定しそれ以外の設定を同一にして学習させた。用いたデータセットはff53 (後述)で、Tensorflow/GoogLeNetを用いた。結果を表-12にまとめた。表-10を見るとバッチサイズが大きいほど計算時間は短いが正答率も低くなるということがわかる。よって以降の学習にはバッチサイズを8に設定して行う。

表-10 各バッチサイズによる計算結果

バッチサイズ	計算時間	正答率(%)
8	6h38m	68.83
16	2h50m	58.64
32	2h48m	57.79

(2) 画像データセットの質の影響

56種類に花の種類を増やしたところ正答率は60%台にとどまる結果となった。このような結果になった原因は用いた画像にあると考える。ここで正答率を上昇させるために正答率の低かったbloomer (正答率13.49%), daisy (正答率33.42%), rocket larkspur (正答率35.34%) 3種を削除した53種類の画像データセットを用いてTorch/GoogLeNetで学習させた。結果は図-2に示すように

正答率は70%を超え、最高値は71.76%であった。このことから正答率の低い画像を削除することで正答率を上げることができることを確認できた。

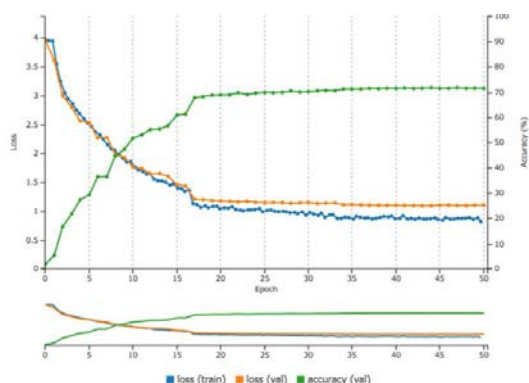


図-2 f153の学習結果

さらに、正答率を向上させる試みとして全ての画像データセットに対して、より厳密に画像のチェックを行った。①植物以外のもの（人，虫，文字など）が大きく写り込んでいるもの，②画像サイズが著しく小さい，また画像が荒いもの（目視で何の花か識別できない程度），③素人目で見ても違う花であるとわかるもの，④多種の花が一緒に写っているもの，⑤咲いた花が1輪も写っていないもの，⑥イラストや加工の激しいもの，以上6項目を判断基準とし，該当するものをデータ内から削除した。削除した画像の例は図-3のようなものである。また以下のような一部の画像ファイルは全削除した。aster（daisyやconeflowerといったキク科全体を含むため）bloomer（咲いた花の写真の寄せ集めであるため）chrysanthemum（菊属を指し，図-4のように形が大きく異なるため）lychnis（センノウ属を指し，図-5のように形が大きく異なるため）pink（ピンク色の花の写真の寄せ集めであるため）。さらに同じ花の種類で名前だけ違うものは一つに統合した。具体的にBlue daisy, oxeye daisy, painted daisyはdaisyへ統合し，Rocket larkspurはdelphiniumへ統合した。最終的に花の種類は47種類となった。



図-3 削除画像の例



図-4 chrysanthemumの画像



図-5 lychnisの画像

整理した47種類の画像データセットをTorch/GoogLeNetで学習をさせた結果は図-6に示す。バッチサイズは8である。正答率の最高値は79.32%で8割近い正答率が得られた。また花別の正答率の上位3位と下位3位を表-11, 12にまとめた。browalliaは表-9では32.48%であったのに対し，表-13では61.84%と約2倍の正答率となっている。上位の花も90%を超えていることから全体的に正答率が上がっていることが分かる。

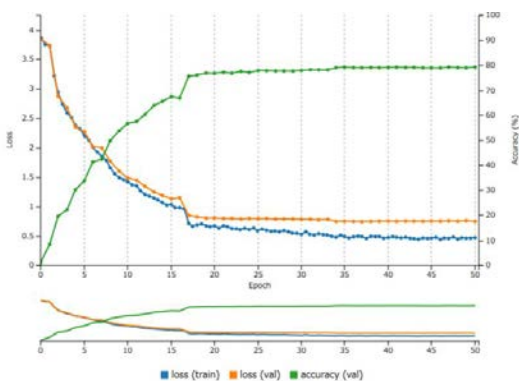


図-6 flower_gt500newの学習結果

表-11 正答率上位3種

順位	%	名前
1	92.34	bellwort
2	31.94	commelina
3	91.73	spathiphyllum

表-12 正答率下位3種

順位	%	名前
1	56.93	bromptonstock
2	57.93	anemone
3	61.84	browallia

がある。

5. まとめ

植物画像でデータセットを作成し、Deep Learningで学習させることができた。また、学習結果を考察し、パラメータの最適化や再学習、サイズ変換などを用いることで、正答率を上げることができ、それぞれのパッケージでの計算性能を評価することができた。また花の種類を増やした学習ではバッチの使用や画像の整理を行い、正答率を上昇させることができた。しかしこれらの学習で用いた花の画像は一般人でも判別できるレベルの種類分けであった。そのため、専門家でも使用できるレベルのツールにするにはさらに細かい種類での判別が必要で、それが可能であるのかをこれからの研究で確かめていく必要

参考文献

- 1) 神宮敏弘：深層学習，153-188，人工知能学会
- 2) ILSVRC <http://www.image-net.org/challenges/LSVRC/>
- 3) Keras：画像分類：GoogLeNet Inception v3
<http://caffe.classcat.com/2017/05/12/keras-image-classification-googlenet-inception-v3/>
- 4) ImageNet <http://www.image-net.org/>

(2018. 8. 24 受付)

ATTEMPT TO IMPROVE PLANT IDENTIFICATION ACCURACY IN DEEP LEARNING METHOD

Rumi AKEBONO, Atsushi NOGAMI, Tomoko NISHINO, Hiroki NAKAYAMA and
Rizumu FUJISAKI

Identification of the kinds from a photography image is an important factor, and, as for the image identification technology by the machine learning method that in late years develops rapidly, the application to this field is expected in a vegetation investigation and a creature investigation. It becomes necessary to decide the maintenance of training data and a learning parameter by trial and error from experience for the learning precision improvement. In this study, I tried the optimization of a processing method and the learning parameter of the training image for the images of the flower of the plant. The image prepared from ImageNet. I made a learning model with 14 kinds of Compositae, Orchidaceae 84 kinds as identification of the having many kinds of 56 kinds, identification of the same course and, on the basis of four kinds of characteristic identification, evaluated it. In four kinds of identification, I got a high correct answer percentage by adjustment, the processing of the training image, the re-learning handling of parameter. However, a correct answer rate greatly varied according to a kind by 56 kinds of classifications and the correct answer rate of Compositae and the Orchidaceae was remarkable and lowered. I will reexamine the quality of training data in future.