

畳み込みニューラルネットワークを用いた花画像の分類

渡邊 葵† 櫻井 彰人‡

慶應義塾大学大学院理工学研究科開放環境科学専攻† 慶應義塾大学理工学部‡

1. はじめに

道端に咲く花を調べる方法として、現在では植物図鑑やインターネットの利用があげられる。しかし、これらは花の名前を知らないと調べるのが難しい。もしスマートフォンなどを活用し、花を見つけたその場で撮影した写真を使って検索するシステムがあれば、便利だろう。

花の画像分類については、これまで研究がなされているが、正面など、向きを定めて撮影した画像を用いた研究[1]が多い。実際には同じ花でも、様々な状況で撮影した花の画像があり、環境に左右されない分類手法が必要である。

そこで本研究では、畳み込みニューラルネットワークを用いて花画像の分類を行い、多くのデータを集めることが困難な花に対し、上記の環境に対応する画像の拡張手法を考案した。

2. 畳み込みニューラルネットワーク

畳み込みニューラルネットワークは深層学習(Deep Learning)の手法の1つであり、特に画像認識の分野で顕著な成績を収めている[2]。

脳の神経科学の知見を基礎とし、画像の局所的な特徴抽出を行う畳み込み層と、局所毎に特徴をまとめるプーリング層を繰り返していく構造となっている。この構造により、微小な変位に対する不変性を獲得している[3]。

3. 実験方法

本研究の主な内容は「データ収集・モデルの構築・花の種類判定」である。

まず、Google の画像検索を利用して、花画像を収集した。本研究で主に使用したのは、26 種類 15250 枚の画像である。

次に、これらの画像を用いてモデルを構築した。Deep Learning の学習には、フレームワークである Caffe を利用した。ニューラルネットの構成は AlexNet[4] とし、学習の際は、ILSVRC2012 のデータセットで学習済みの CaffeNet のモデル[5]を用いて、Fine-tuning による学習を行った。学習の最大繰り返し数を、

10000 回とした。

そして構築したモデルで、ある花の画像が何の花だと判定されるのかを調べた。その際、上位 5 つを所属確率と共に出力した。また、第 4 節では 26 クラスの分類、第 5 節・第 6 節では花 a~c をそれぞれ加えて 27 クラスの分類をした。

4. 予備実験

まず、様々な環境下での画像で構成した訓練データ(データ 1)と、正面から撮影した 1 枚の花が写った画像からなる訓練データ(データ 2)に基づく 2 つのモデルを作成した。それぞれのモデルの訓練データの例を図 1 に示す。

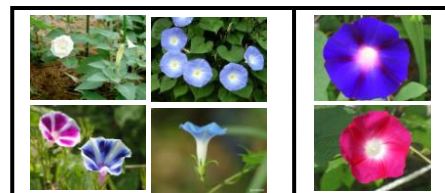


図 1 訓練データの例(左:データ 1, 右:データ 2)

テストデータには、データ 1 のような画像を 3000 枚程、訓練データとは別に用意した。結果を表 1 に示す。

表 1 訓練データの違いによる正答率の違い

データ 1	86.6%
データ 2	49.0%

表 1 からわかるように、データ 1 では正答率が高いが、データ 2 では正答率が落ちてしまった。そのため、多くの画像を集めることができない花を訓練データに入れる場合は、データを増やさなければ、様々な環境下での画像で構成した訓練データと同程度に正解することはできないとわかった。

5. 訓練データ拡張の評価

第 4 節の結果を受けて、入手できる画像が少ない花の訓練データにおける対策を提案する。3 種類の花(ハマボウ(花 a)、シラネアオイ(花 b)、フシグロセンノウ(花 c))を用いて実験を行った。

5.1 訓練データの拡張方法

撮影された花に考えられる状態として、「小さい・複数・位置・傾き・日向日陰」などが考えられる。そこで、入手できた画像を正面から撮影した 3 枚(i)のみとし、これを以下のように拡張した(図 2)。また、すべての画像において

Classification of flower images using a convolutional neural network

†Watanabe Aoi (Keio University Graduate School)

‡Sakurai Akito (Keio University)

3段階の明度をつけた。

(ii)1枚のサイズを変えた画像を載せたもの

(iii)サイズを変えた画像を敷き詰めたもの

(iv)複数枚をランダムに配置させたもの

(v)縦又は横の長さだけ変えたもの

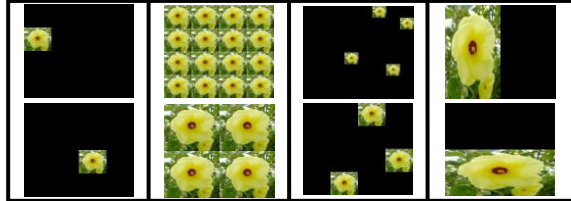


図2 拡張した画像の例(左から(ii)~(v))

上記の(i)~(v)を以下のように組み合わせて学習させ、5つのモデルを構築した。

Model1...(i)

Model2...(i), (ii), (iii)

Model3...(i), (iii), (iv)

Model4...(i), (iii), (iv), (v)

Model5...(i), (ii), (iii), (iv), (v)

5.2 モデル毎の結果

花 a~c について、20枚ずつテストデータを用意した。判定された花を所属確率順に上位5つを出力し、1位の所属確率が80%以上であった画像の数を、モデル毎に表2に示す。

表2 モデル毎の正解した数

	Model1	Model2	Model3	Model4	Model5
花 a	2	11	7	9	8
花 b	0	2	1	0	2
花 c	2	8	5	8	3

表2より、Model2が最も正解しやすいことがわかった。

6. テストデータ編集の評価

第5節で正解できなかったものに対して、テストデータを編集することで正しく判定できるのか、検証した。

6.1 データの編集方法

花卉領域を正方形に切り取った。具体例を、図3に示す。



図3 テストデータの編集方法

この編集をカスケード分類器[6]によって行い、編集後の画像をModel2で判定した。

6.2 編集を含めた結果

6.2.1 予備実験（手作業による編集）

カスケード分類器を用いる前に、手作業で第

6.1節の編集を行った。第5.2節のModel2において、1位の所属確率が80%又は90%以上の場合はその花を答えとし、そこで答えが出なかったものを第6.1節の方法で判定し、所属確率に関係なく1位だと判定されたものを答えとした。この方法で、最終的にテストデータ20枚のうち、正解した割合の結果を表3に示す。

表3 最終的に正解した割合(手作業)

	80%以上	90%以上
花 a	95%	95%
花 b	55%	65%
花 c	70%	70%

表3より、第5.2節において1位の所属確率が90%以上の場合に答えを確定し、その他の画像は第6.1節の編集をしてから判定する方法が、最も正答率が高かった。

6.2.2 カスケード分類器による編集

次に、カスケード分類器により第6.1節の編集をした。結果を表4に示す。

表4 最終的に正解した割合(カスケード分類器)

	80%以上	90%以上
花 a	85%	85%
花 b	25%	30%
花 c	55%	60%

カスケード分類器においても、第5.2節における1位の所属確率の判断基準については、手作業と結果が同じであった。表3,4の右列を見ると、手作業では全体で76.7%の正答率だが、カスケード分類器においては58.3%の正答率であり、花卉領域検出の精度改善が必要であることがわかった。

7. まとめ

訓練データの拡張とテストデータの編集により、入手できる画像が少ない花でも、手作業であれば76.7%の精度で正しく判定することができた。しかし、カスケード分類器での検出失敗、似た花があるときの正答率の低下、クラス数が変わった場合の影響など、更に多くのデータで実験する必要がある。また、学習器のアーキテクチャの変更も検討したい。

参考文献

- [1] 藤賀裕美, 西尾孝治, 小堀研一 “花画像認識のための一手法” 電子情報通信学会総合大会.2013.
- [2] 中山英樹 “深層畳み込みニューラルネットワークによる画像特徴抽出と転移学習” 電子情報通信学会音声研究会 7月研究会.2015.
- [3] 岡谷貴之 “深層学習” 講談社. 2015.
- [4] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems. 2012.
- [5] Caffe Model Zoo, http://caffe.berkeleyvision.org/model_zoo.html
- [6] カスケード型分類器, http://opencv.jp/opencv-2.2/c/objdetect_cascade_classification.html