

花画像データセットの構築と畳み込みニューラルネットワークによる分類

三浦 健太[†] 喜田 拓也[†]

[†] 北海道大学大学院情報科学研究科 〒060-0814 札幌市北区北14条西9丁目

E-mail: †{k-miura,kida}@ist.hokudai.ac.jp

あらまし 花画像について、精度のよい分類を行うシステムの開発に取り組んだ。システム構築のために、ウェブから花画像を収集し、花卉領域のみを抽出したおよそ4万枚の画像を作成した。さらに、回転と反転操作を加え、全体で100クラス約300万枚からなる花画像データセットを構築した。このデータセットに対し、畳み込みニューラルネットワークを用いた分類を行い、その精度を調査した。実験の結果、データ拡張を行うことで分類精度が2.8%向上し、テストエラー9.4%を達成した。また、花の特徴を捉えた絵であっても、正しく分類できることを確かめた。

キーワード 深層学習, CNN, 画像データベース

1. はじめに

見知らぬ草花について、その名前や性質を知りたいと思う場面は存在する。しかし、多数の種類がある草花を判別するには、草花についてある程度の知識が必要である。例えば、草花の知識を得る最も単純な方法は、図鑑の中から類似する草花の画像を手作業で見比べながら探すことであるが、草花の知識に乏しい素人にとって容易な作業とは言い難い。逆に、その草花の名前さえ判明すれば、その情報を見つけることは容易である。

一方で、スマートフォンなどカメラ機能を備えた携帯端末が広く普及し、それらを用いて草花の写真を撮ることは一般的に行われるようになった。そのような写真画像から草花の名前を同定できると非常に便利である。こうした目的を達成するための画像認識技術は既に多くの研究がなされているが、そのうち草花の自動認識に特化した研究がいくつか存在する。

齋藤ら[1]が行った野草の認識実験では、花卉画像から10種、葉画像から11種の画像特徴量を取得して認識に用いている。同種の研究としては、森本ら[2]の研究があり、葉の形状からP型フーリエ記述子と呼ばれる特徴量を抽出して認識に用いている。ただしこれらの研究では、背景の影響をなくすために、採取した草花を無地の紙の上に置いて正面から撮影した画像を入力として用いている。このような背景のない適切な画像を準備することは、手軽に出来るとは言いがたい。

自然に撮影された花画像を用いた既存研究では、前処理として花卉領域の抽出を行うものがある。武市ら[3]は、花卉の色に関するドメイン知識(花卉の色味)を用いて対象とする花卉領域を抽出する方法を提案している。彼らの手法においては、ドメイン知識はあらかじめシステムに登録されている。また、田畑ら[4]は、認識させたい花の特徴を画像とともにユーザに入力させる方式をとっている。一方、大藪ら[5]は、ユーザに画像中から花卉領域に含まれるピクセルを選択させることで、花領域の抽出に必要な情報を取得する手法を提案している。これらの手法はいずれも、抽出した花卉領域の色と形状から特徴量を設計し、画像認識を行っている。したがって、花卉の画像部分をどれだけ正確に抽出できるかが認識精度に影響する。こ

れは、花卉部分を含む矩形領域を特定するよりも困難な作業であり、自動化が難しい。

近年、画像認識研究では、深層学習の手法の一つである畳み込みニューラルネットワーク(CNN)が多くの成功を収めている[6]。渡邊ら[7]は、CNNを利用した花画像分類を試みており、花卉の形を抜き出すことなく、26種の花データベースに対して76.7%の精度で正しく分類できたと報告している。用いられたニューラルネットワークの構成はAlexNet[8]で、ILSVRC2012のデータセットで学習済みのモデルからFine-tuning[9]による学習を行っている。Liuら[10]は、自然な花画像の分類に取り組んでおり、クラス間の類似性とクラス内の多様性による難しさに言及している。8層構造のCNNを用いており、独自に作成した79クラスの花画像データセットに対し76.5%、花画像データセットOxford102に対し84.0%の分類精度を報告している。CNNによる画像認識のボトルネックは、学習のために大量のデータを必要とすることである。渡邊ら[7]の研究では、26種類で総計15,250枚の画像をネット上から収集し用いているが、より多数の種類を高精度に判別するためにはさらに多数の画像データが必要となる。そこで本論文では、100クラス約4.5万枚の花画像を収集し、それをもとにデータ拡張を行い、総計3,043,496枚の花画像データベースを作成した。これに対し、VGG[11]のモデルを用いた学習を行い、その分類精度について調査を行った。その結果、90%以上の精度で認識できることを確認した。

2. 準備

本節では、実験に用いる畳み込みニューラルネットワークと、その構成の一つであるVGGについて説明する。

2.1 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(CNN)は、深層学習の手法の一つであり、画像認識の分野で従来の方法を上回る高い性能を見せている。従来のような特徴量の設計を行わず、認識に必要な特徴を自動的に発見することができる[12]。

画像を入力する入力層から出力層へ向けて、畳み込み層(conv)とプーリング層(pool)を繰り返した後、何層かの全結

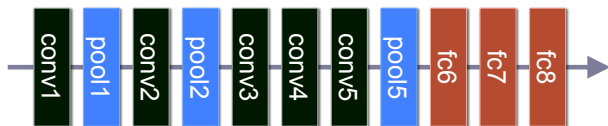


図 1 VGG-CNN_M の構成

合層 (fc) を繋げた構造となっている。出力層は、分類を行うクラスに対応したユニットからなり、各クラスの分類確立が出力される。畳み込み層では、フィルタの畳み込みが行われ、特徴抽出に対応する働きをする。プーリング層では、局所領域の値を集約することで、位置情報の汎化を行う。学習アルゴリズムに従って、自動的にパラメータを更新することで、訓練データに適した分類器として利用することができる。

2.2 VGG

VGG [11] は、2014 年に開催された画像認識のコンテスト ImageNet Large Scale Visual Recognition Challenge (ILSVRC) に参加したチームの名称であり、そこで使用された CNN の構成モデルの名前にも用いられている。

本実験では、VGG-CNN_M [13] を使用した。このモデルは ILSVRC2012 のデータセットで、テストセットの top-5 エラー 13.7% の性能を示している。5 層の畳み込み層と 3 層の全結合層からなる 8 層構造である (図 1)。

3. データセットの構築

本節では、ポケット花図鑑 [14] の花の名前一覧を参考に、Google 画像検索から 128 種の花画像を取得し、このうちの 100 種について花卉領域の抽出を行っている。また、データ数を増やすために、画像の回転と反転によるデータオーグメンテーション [8] を行う。

3.1 花卉領域の抽出

武市ら [3] の花卉の色に関するドメイン知識 (花卉の色味) に基づいて、花卉領域の抽出を行った。花卉の色味としてありうる色空間を青、赤～黄、紫、白の 4 つの色空間に分けた。それぞれの色を画像から抜き出し、連続領域ごとにそれがちょうど収まる矩形領域を考える。連続領域のラベリングには、井村 [15] のラベリングクラスを利用した。この矩形領域の縦幅と横幅を比較し、長辺の長さを画像を切り出す正方領域の 1 辺の長さとする。ただし、この正方領域が画像からはみ出す場合には、画像内に収まる可能な限り大きな正方領域を切り出した。最後に、抽出した画像は、100 × 100 画素にリサイズした。

このようにして得られた画像は、それが花であるかどうかを認識していない。そこで、抽出した画像を人の目で確認し、花の画像だけを取り分けた。この時、アジサイのように小花が集まって咲く種は、その全体を 1 つの花として扱った。また、画像データ中には花が 1 つ、あるいは同じ種の花が複数写っている。

この一連の処理によって、1 つの画像から複数の花卉領域を抽出することが可能である。また、4 つの色空間による抽出を行うために、1 つの花が複数の異なる矩形領域で切り出される場合もある。例えば、図 2 のヒマワリは背景の空の青色で抽出



図 2 異なる色空間による抽出例

された画像と、花卉の黄色で抽出された画像がある。

3.2 画像の回転・反転

画像の回転と反転によるデータ拡張を次の方法で行った。領域抽出前の元の画像を 0° から 80° まで 10° ずつ回転し、3.1 節の方法を行って、花卉領域を抽出した。さらに、これらの画像とこれらを 90° 回転した画像を上下、左右、上下左右にそれぞれ反転した画像を生成した。これにより、360° 分の回転とその鏡像を考慮した画像セットができる。

以上の操作は、元の画像を 0° から 350° まで 10° ずつ回転して領域を抽出し、それらの画像を上下反転 (あるいは左右反転) した画像を生成する操作と同等である。しかし、色情報のみを用いて抽出した画像は、それが花であることを確認する必要がある。作業の効率化のために、80° までの回転画像からの抽出と 90° 回転、反転操作を組み合わせた手順を用いた。

3.3 テストデータとトレーニングデータ

3.1 節の方法によるデータ拡張を含まないデータセット DS1 と、3.2 節の方法による回転や反転を考慮したデータセット DS2 を用意した。テストデータは 2 つのデータセットで共通で、それぞれ異なる画像から抽出された変形を加えていない画像を用いた。異なる個体の画像が 50 枚にも満たないクラスが存在したため、テストデータは各クラス 5 枚とし、100 クラス 500 枚からなるテストセットを構築した。トレーニングデータには、テストデータが抽出された画像から生成された画像を除外し、残った全ての画像を用いた。トレーニングセットは 100 クラス合計で、DS1 では 45,896 枚、DS2 では 3,043,496 枚からなる。

4. 実験

CNN を用いて、作成したデータセットの学習と分類実験を行う。

4.1 データ

ネットワークの学習とテストに用いるデータは、3.3 節で述べた通りである。

4.2 実験方法

前節で作成した画像データセットを用いて、畳み込みニューラルネットワークによる学習を行った。学習には、ディープラーニングのフレームワークである Caffe を利用した。

作成したデータのみを用いて一から行う学習 (FT 無し) と、ILSVRC2012 のデータセットで学習済みのモデルを用いた Fine-tuning による学習による結果とを比較した。Fine-tuning による学習では、最上位の全結合層 1 層のみを学習する方法 (FT1layer) と、3 層ある全結合層全体を学習する方法 (FT3layer) の 2 つを行った。学習の最大繰り返し回数を

表 1 学習終了時のテストセットに対するエラー率（データ拡張無し）

FT 無し	74.2%
FT1layer	19.4%
FT3layer	12.2%

表 2 学習終了時のテストセットに対するエラー率（データ拡張）

FT 無し	14.4%
FT1layer	17.0%
FT3layer	9.4%

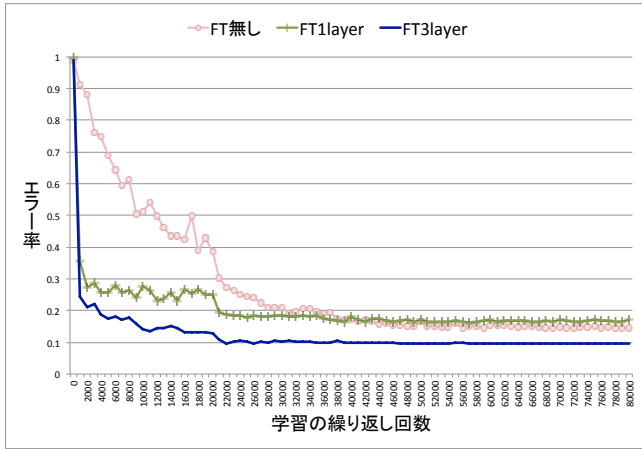


図 3 テストセットに対するエラー率の変化（データ拡張）

80,000 回とし、20,000 回ごとに学習率を 1/10 に変化させた。

4.3 データ拡張なしの実験

表 1 にテストデータに対する学習終了時の分類エラー率を示す。Fine-tuning を用いた方法ではうまく学習が進んだが、FT 無しでは良い結果が得られなかった。

4.4 データ拡張ありの実験の結果

データセット DS2 を用いて、VGG_CNN_M の学習を行った。表 2 にテストデータに対する学習終了時の分類エラー率を、図 3 に学習による分類エラー率の変化を示す。FT3layer では、データ拡張無しの時と比べて、分類精度が 2.8% 向上し、エラー率 9.4% という結果が得られた。3 つの方法それぞれの学習終了時のパラメータを用いて、分類精度が低いクラスを表 3 にまとめる。

FT3layer での分類精度が特に悪かった 2 クラスについて、それぞれ正しく分類されなかった 3 枚の画像の分類先クラスと出力層の Probability の値を上位 5 件まで調べた（表 4、5）。アルストロメリアは、上位 5 件の分類先に正しいクラスが 1 つも入っていなかったが、カーネーションは 3 枚とも上位 5 件までに正しいクラスが選ばれていた。

一方、ウォーターポピーの訓練データ数は 3,296 枚であり、最も少ないクラスであるが、比較的良好に分類できていた。FT 無しと FT1layer、FT3layer での正答数はそれぞれ 4/5、3/5、4/5 であった。このクラスのデータを調べると、花卉の色は黄色あるいは白色の単色であり、花卉の形状も 2 種類ほどしかなかった。さらに、データのほとんどが花の正面方向から撮影された写真画像であったため、クラス内のデータは見た目の特徴がよく似ていた。したがって、色や形状がよく似たデータを集

表 3 分類精度が低いクラスと正答数（データ拡張）。

3 つの学習方法による分類の結果、いずれかで 2 枚以下しか正しく分類できなかったクラスの結果。正答数の合計が小さい順に並べている。

	訓練データ数	FT 無し	FT1layer	FT3layer
アンチューサ	8,264	1/5	1/5	3/5
キンギョソウ	7,968	1/5	1/5	3/5
アルストロメリア	21,832	2/5	2/5	2/5
カーネーション	10,704	1/5	3/5	2/5
スイートピー	12,264	2/5	2/5	3/5
アサガオ	57,544	3/5	2/5	3/5
アマリリス	34,272	2/5	2/5	5/5
アルメリア	34,712	4/5	2/5	3/5
デイジー	23,736	4/5	2/5	3/5
アイビーゼラニウム	28,704	2/5	3/5	4/5
クリスマスローズ	25,216	4/5	2/5	4/5

めて 1 つのクラスとすれば、学習がうまく進み、正しく分類できる可能性がある。

4.5 クラスを分割した実験

クラス内の特徴の分散が小さくなるように、表 3 の上から 5 つのクラスを花卉の色の特徴で分割し、106 クラスのトレーニングセットとテストセットを用意した。FT3layer の学習終了時のパラメータを利用し、4.2 節と同様の設定で Fine-tuning を用いて、全結合層 3 層の学習を行った。このネットワークを FT3layer106 と呼ぶ。

FT3layer106 での学習終了時のテストセットに対するエラー率は、11.1% であった。分割したクラスに関して、同じ花の異なる分割も正解クラスとしてみなすと、エラー率は 10.2% となる。4.4 節に比べ、いくつかのデータでは、正解クラスの分類確率の向上が確認された。表 5 のカーネーションは、3 つ全てのテストデータで正解クラスの確率が上がり、2 つのデータは正しく分類された。しかし、分割前の 100 クラスでのエラー率 9.4% には及ばず、全体の精度向上には至らなかった。

4.6 花の絵の分類

花の絵をテストデータとして用いる。花の絵は、データセット構築後に Instagram に投稿された画像を参考にして、色鉛筆を用いて描いたものである。カメラで撮影したものを、花卉領域に合わせた正方領域でトリミングし、100 × 100 画素にリサイズした。ハイビスカスと、ガゼニア、コモンマロウの 3 種について、周りの葉の有無などの違いがあるデータを合わせて 7 つ用意した。

学習済みの FT3layer での分類の結果、ハイビスカスの 1 枚は分類先 2 位で正解クラスが現れているが、残りの 6 枚は 1 位で正しく分類された。

5. おわりに

本研究では、畳み込みニューラルネットワークを用いて、花画像を精度よく分類するシステムの開発に取り組んだ。システム構築のために、ウェブから花画像を収集し、花卉領域のみを抽出したおよそ 4 万枚の画像を作成した。さらに、回転と反転操作を加え、全体で 100 クラス約 300 万枚からなる花画像

表 4 失敗したアルストロメリアの分類先クラス上位 5 件 .
FT3layer で分類に失敗したアルストロメリアの分類結果 . 表中 , 花の名前の下に記した数値
は , 出力層の Probability の値である .







テストデータ	1 位	2 位	3 位	4 位	5 位
	アンズリウム 76.1%	アマリリス 17.2%	クリスマスローズ 1.69%	ノウゼンカズラ 1.03%	グロリオーサ $8.34 \times 10^{-1}\%$
	アマリリス 34.9%	スイートピー 9.38%	インパチェンス 7.26%	ゴデチア 5.33%	クレマチス 4.90%
	カロライナジャズミン 56.2%	スイセン 30.9%	ウォールフラワー 5.29%	フリージア 2.06%	オシロイバナ $8.76 \times 10^{-1}\%$

表 5 失敗したカーネーションの分類先クラス上位 5 件 .
FT3layer で分類に失敗したカーネーションの分類結果 . 表中 , 花の名前の下に記した数値は ,
出力層の Probability の値である .

テストデータ	1 位	2 位	3 位	4 位	5 位
	ツバキ 99.1%	ラナンキュラス $7.30 \times 10^{-1}\%$	インパチェンス $1.14 \times 10^{-1}\%$	カーネーション $4.55 \times 10^{-2}\%$	スイセン $9.52 \times 10^{-4}\%$
	ラナンキュラス 28.9%	カーネーション 17.6%	デイジー 10.6%	セロシヤ 5.55%	ガーベラ 4.99%
	センニチコウ 85.9%	カーネーション 8.85%	アルメリア 2.78%	アスター $6.15 \times 10^{-1}\%$	ラナンキュラス $3.74 \times 10^{-1}\%$

データセットを構築した . このデータセットに対し , 畳み込みニューラルネットワークを用いた分類を行い , その精度を調査した . 実験の結果 , データ拡張と Fine-tuning を用いた方法で , テストエラー 9.4% を達成した . また , 花の特徴を捉えた絵であっても , 正しく認識されることを確かめた .

今後の課題は , より多くのクラスの分類についての性能評価と課題点の抽出である . また , 葉の形などの花卉の周りの情報も認識しているかを検証するために , 背景を塗りつぶしたデータやその花に関係のない背景を入れたデータを用意して分類実験を行う必要がある .

文 献

- [1] 齊藤剛史 , 金子豊久 . 花と葉による野草の自動認識 . 電子情報通信学会論文誌 . D-II, 情報・システム, II-パターン処理, Vol. 84, No. 7, pp.1419-1429, Jul. 2001.
- [2] 森元亜美, 向井彦彦, 小杉信. 葉形状特徴を用いた植物の識別. 映像情報メディア学会技術報告, Vol. 34, No. 15, pp. 69-72, Mar. 2010.
- [3] 武市寛之, 竹内義則, 山村毅, 松本哲也, 工藤博章, 大西昇. 画像入力による花図鑑検索. 電気学会論文誌. C, 電子・情報・システム部門誌, Vol. 123, No.12, pp. 2111-2119, Dec. 2003.
- [4] 田畑惣太郎, 岩崎慶, 高木佐恵子, 吉本富士市. 花の画像検索システムと検索方法の評価. 電子情報通信学会技術研究報告. PRMU, パターン認識・メディア理解, Vol. 104, No. 670, pp. 1-6, Feb. 2005.
- [5] 大藪貴志, 浅野晃. 画像をキーとする植物名称検索システム. 電子情報通信学会技術研究報告. SIS, スマートインフォメディアシステム, Vol. 109, No. 203, pp. 29-33, Sep. 2009.
- [6] 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 . Jul. 2015.
- [7] 渡邊葵, 櫻井彰人. 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. 情報処理学会第 78 回全国大会, 1N-01, pp. 197-198, Mar. 2016.
- [8] Alex Krizhevsky, Ilya Sutskever and Geoffrey E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems 25, pp. 1106-1114, Dec. 2012.
- [9] Ross Girshick, Jeff Donahue, Trevor Darrell and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, pp. 5805-5814, Jun. 2014.
- [10] Yuanyuan Liu, Fan Tang, Dengwen Zhou, Yiping Meng and Weiming Dong. Flower classification via convolutional neural network. In Proceedings of the International Conference on Functional-Structural Plant Growth Modeling, Simulation, Visualization and Applications (FSPMA), pp. 1101-1116. IEEE, 2016.
- [11] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. Computing Research Repository, <http://arxiv.org/abs/1409.1556>, Sep. 2014.
- [12] 岡谷貴之, 齋藤真樹 . コンピュータビジョン最先端ガイド 6 , 第 4 章 ディープラーニング . アドコムメディア . ISBN 978-4915851544. Dec. 2013 .
- [13] Ken Chatfield, Karen Simonyan, Andrea Vedaldi and Andrew Zisserman. Return of the Devil in the Details: Delving Deep into Convolutional Nets. Computing Research Repository, [abs/1405.3531](http://arxiv.org/abs/1405.3531), Nov. 2014.
- [14] COMERL.COM. ポケット花図鑑. <https://www.komeri.com/flower/index.html>.
- [15] IMURA LABORATORY ラベリングクラス. <http://imura-lab.org/products/labeling/>