

Deep Convolutional Neural Network を用いた 高山植物認識システム

根岸 知誠[†]山梨大学 大学院医工農学総合教育部[†]服部 元信[‡]山梨大学 大学院総合研究部[‡]

1 はじめに

高山植物とは、高山帯に生息している植物を指し、日本に生息している種類は約 440 種(亜種・変種・品種を含まない)である。高山植物の種類を同定する方法として一般的なのは、図鑑などを用いて花の色、葉の形、植物の背丈などから図鑑と見比べて種類の見当をつけるものである。しかし、高山帯に実際に図鑑を持っていき、その場でたくさんの種類の中から植物を同定することはとても煩雑である。

本研究では、高山帯で携帯、撮影が可能な端末を使用し、撮影された高山植物画像を用いて種類を同定することを目指すものとし、その根幹となる高山植物画像の認識システムを Deep Convolutional Neural Network (DCNN)を用いて構築することを目的とする。

DCNN は大量のデータによって学習することで特徴表現を獲得し、高い認識性能を得ることができる。しかし、本研究で認識のドメインと設定した高山植物画像は、十分な枚数の実験データの確保が困難であることが予想される。そこで、ドメイン以外の植物画像を実験データの補填として用いる実験を行った。

2 DCNN

DCNN は、画像認識に広く応用される Convolutional Neural Network (CNN)を深層学習させた多層のニューラルネットワークである。本研究で用いる DCNN は特徴抽出を行う 10 層の中間層、中間層の出力情報から認識結果を出力する 3 層の識別層から構成される(図 1)。中間層は convolution 層 → pooling 層 → normalize 層という処理を繰り返し行い、入力された画像の特徴表現を獲得している。各層は、複数の画像状の平面構造を持つ素子から成り立っている。

- convolution 層: 畳み込み演算が行われ、画像上の特徴的な濃淡パターンの検出を行う
- pooling 層: 最大プーリングを行い、情報圧縮と局所変形に対する応答の頑健性の向上を図る
- normalize 層: 出力の正規化を行い、一連の処理のあとの出力が過剰に大きくなることを抑制する

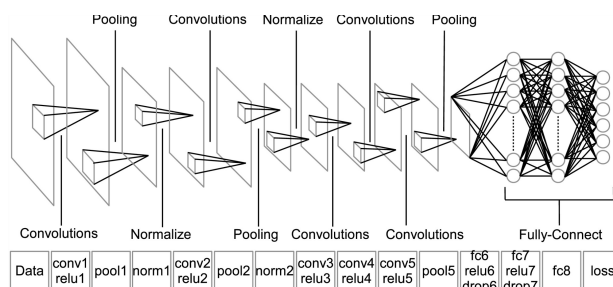


図 1: DCNN の概要

この DCNN では損失関数として softmax 関数を用い、誤差逆伝播法 (Back-propagation: BP) によって学習を行う。

3 Fine-tuning

鈴木らによる研究では、実験データの少ない「びまん性肺疾患画像」の特徴解析において、DCNN に対して取得が容易な自然画像とびまん性肺疾患の CT 画像を用いた Fine-tuning を適用し、評価実験、比較検討を行うことで、実験で構成された DCNN 特徴量が有効であることを示した [1]。

Fine-tuning では、自然画像を用いて一般的な画像特徴の学習を行った DCNN に対し、目的の画像を用いて再学習することで、中間層が目的の画像に対して有効な特徴を獲得できるように調整することができる。

本研究では、自然画像で学習した DCNN を高山植物画像と補助の植物画像を用いて再学習を行うことで、有効な特徴量を獲得することを考える。Fine-tuning 前後では、学習するデータの違いから、同じ識別層を用いた比較ができない場合がある。そこで、DCNN の識別層の代わりに、識別器として Support Vector Machine (SVM) を用いることで、Fine-tuning 前後の比較を行う。SVM は DCNN 中間層で得られる特徴表現(図 1 の DCNN 中間層の最終 pooling 層 (pool5) から識別層の fully-connected (fc6) に渡される 4096 次元の特徴量)により学習され、分類を行う。

4 実験

4.1 実験データ

認識のドメインに設定した高山植物は、高山植物図鑑に掲載されている植物の種類を参考に 71 種類をク

スに設定した [2]. 画像は, 非営利目的での再利用が許可されたものをインターネットから 2104 枚収集した.

ドメイン以外の補助のための植物は, ドメインに設定されている高山植物と同じ属から 55 種類をクラスに選び, 画像は, ImageNet データベースから 25255 枚を収集した.

4.2 実験方法

鈴木らの実験と同様に, 自然画像のみを用いて学習した DCNN として, Caffe が提供している ImageNet データセットで学習済みのモデルを用いた [3]. ImageNet データセットはクラス数 1000, 画像枚数 150000 枚の自然画像データセットである. この学習済みのモデルに実験データを用いて再学習をし, Fine-tuning を行った. 再学習の手順として

- (a) 高山植物画像で再学習
 - (b) 補助植物画像で再学習
 - (c) 補助植物画像で再学習した後, 高山植物画像で再学習
 - (d) 高山植物画像と補助植物画像で再学習
- を考え, 再学習を行わない学習済みのモデルとともに手順(a)~(c)の比較を行った.

Fine-tuning では, 誤差が十分に下がると学習回数によってモデルの認識性能にばらつきがでることが実験的に確認できた. そこで, 本実験では Fine-tuning を学習回数 3000 回まで続け, 学習回数 100 回ごとのモデルに対してテストデータによって認識精度を確認した. その中で一番精度の高いモデルを Fine-tuning 後のモデルとして設定し, 実験を行った.

SVM のカーネル関数にはラジアル基底関数 (Radial Basis Function: RBF), 多クラス分類法として 1 対 1 分類法を用いた. SVM のパラメータは, ソフトマージンのコストパラメータ $C \in \{1.0e+1, 1.0e+2, 1.0e+3\}$, カーネルパラメータ $\gamma \in \{1.0e-2, \dots, 1.0e-6\}$ でグリッドサーチを行い, $C = 100$, $\gamma = 1.0e-3$ に設定した.

認識精度の評価として高山植物画像による 10 分割交差検証を行った. 実験では, ドメインに設定した高山植物画像をクラスごとに実験データを 10 分割し, 各検証でテストデータが重複しないように交差検証を行った. 補助の植物画像は, Fine-tuning 時の認識精度の確認のために各クラスから 10 枚をテストデータ, 残りの画像を学習データとして用いた. よって, 毎回の検証で用いた各画像の実験データは表 1 のように設定した.

表 1: 実験に用いた各画像のクラス数, 枚数

	高山植物	補助植物
クラス数	71	55
学習枚数	1919	24705
テスト枚数	185	550

4.3 実験結果

実験結果を表 2, 3 に示す. 表 3 は, 識別層が高山植

物クラスに対応している手順(a), (b)について比較したものである.

表 2: 交差検証実験の結果 (識別器: SVM)

再学習方法	平均認識率 (%)	標準偏差
再学習なし	73.56	4.87
手順(a)	78.37	4.55
手順(b)	78.54	4.25
手順(c)	80.21	5.13
手順(d)	73.02	3.70

表 3: Fine-tuning 後の識別層による認識率

再学習方法	平均認識率 (%)	標準偏差
手順(a)	83.62	4.06
手順(c)	85.73	4.06

5 考察

表 2 から, (b) は再学習なしと比べて精度が向上し, 高山植物の認識をサポートできている. (c) は (a) よりも良い結果 ($p < 0.05$) で, (a) の学習で得られなかった植物の特徴を (c) で補助植物から学習できていることが考えられる. 手順(d) が (c) よりも精度が低いことから, (c) のように高山植物のみによる最終的な調整がない場合, 高山植物のための特徴表現がぼやけてしまうものと考えられる.

表 3 の結果は表 2 と比較して高い認識率である. これは, DCNN の中間層が識別層に合わせて学習を行っているため, 得られる特徴が SVM よりも学習した全結合層に適していることが考えられる.

6 まとめ

本研究では, DCNN を用いた高山植物認識システムの構築を目的とし, Fine-tuning によって高山植物画像に対する DCNN の認識性能向上を図った. 少ない実験データの補填, 植物における特徴表現の獲得を期待して, ドメインに設定した高山植物画像とドメイン以外の補助植物画像を用いて再学習する実験を行った. 実験結果から, 補助植物画像でも, 高山植物画像の分類に有効な特徴を学習できることが確認できた.

参考文献

- [1] 鈴木聡, 庄野逸, 輝度尚治: Deep Convolutional Neural Network を用いたびまん性肺疾患画像の特徴解析, 信学技報, NC2014-114, pp.259-264 (2015).
- [2] 新井和也: 高山植物の基本 OUTDOOR POCKET MANUAL, エイ出版社 (2012).
- [3] Yangqing Jia, Evan Shelhamer: Caffe | Fine-tuning for style recognition, Caffe | Deep Learning Framework, <http://caffe.berkeleyvision.org/gathered/examples/finetune_flickr_style.html> (2015-12-8).