

深層学習法による高精度植物識別システムの構築と里山保全への適応

中山 紘喜¹・曙 瑠美²・井畔 実穂³・西野 友子⁴・野上 敦嗣⁵

¹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

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)

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

³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)

E-mail: w5551005@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

環境アセスメント等による生態系調査や、里山等の環境保全活動において種の識別は重要な要素だが、正確な識別には専門家の知識が必要となっている。一方で深層学習法（DL法）の確立以来AIによる画像認識技術の応用研究が進み、種の識別においてもその効果が期待される。本研究ではこれまでDL法を用いた植物識別システムを構築してきたが、花画像56種の識別において精度は62%に留まっていた。本研究では植物の識別において、転移学習等の調整を行いその高精度化を図った。結果47種の識別において90%を越す精度が得られ、植物公園で独自に撮影した53種の識別においても同等の精度を得た。そのうえで環境保全地区となっている里山の樹木を対象に、デジカメやドローンの撮影画像や樹冠画像の識別精度を検証し、里山保全への適応について検討する。深層学習システムにはKerasを、樹冠画像の作成にはMeanShift法によるスーパーピクセル分割を用いた。

Key Words : deep learning, vegetation investigation, IOT, image classification, OpenCV

1. 緒言

環境アセスメント等による生態系調査や、里山等の環境保全活動において種の識別は重要な要素だが、正確な識別には専門家の知識が必要となっている。一方でDL法の確立により画像認識技術の応用研究が進み、ILSVRCではAIによる1000種の識別において一部で人間の正答率95%を超える精度を得ている¹⁾。画像による種の識別において極めて有望なDL法だが、高い学習精度を得るには、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。本研究でも

ILSVRC が利用する画像データベースImageNetから集めた56種の花画像の識別精度は、高いものでも60%に留まっており²⁾、精度向上が課題となっている。また一枚の写真からいくつかの領域を切り出し、学習に用いる画像数を増やすことで識別率の向上が望めることが分かっている³⁾。本研究では植物の識別において、学習画像の選別、転移学習や神経回路網モデルの更新、パラメータの最適化によって識別精度改善を図った。そのうえで里山保全に着目した際、樹木種の特定、育成規模の把握に対し本システムの適応について検討する。今回はドローン空撮による樹冠画像からの植物種の識別を試みる。

2. 研究方法

(1) 植物の撮影

本研究で用いた植物画像は、ILSVRCで用いられている画像データベースImageNetの他、北九州市門司区の白野江植物公園と、北九州市立大学ひびきのキャンパス校内にある里山で撮影した写真を用いている。北九州ひびきの学研都市内には自然保護地区として整備された区画がいくつかあり、そのうちの一つは北九州市立大学ひびきのキャンパス校内にある。本研究ではこの地区を研究対象の里山に選んだ。撮影にはSONY製デジタルカメラDSC-HX60Vを用い、ひびきのキャンパスでは上記に加え、DJI製ドローン MAVIC2 Proを用いた。画像サイズはデジカメが5184×3888、ドローンが5472×3648となっている。またドローンの空撮写真とwebサービスDroneDeployを用いて19968×12544のパノラマ写真を作成した。

(2) 深層学習による多種類植物識別の精度改善

今回深層学習に使用した計算機システムを図-1に示す。従来はDIGITSにおいてCaffe, Torch, Tensorflowの3種の機械学習フレームワークと、AlexNet, GoogLeNetの2種の神経回路網モデルを用いて深層学習を行っていたが、転移学習の導入に伴い使用する深層学習システムをKerasに変更した。KerasではTensorflowを機械学習フレームワークに用い、神経回路網モデルとしてVGG16, InceptionV3, Xceptionの3種を用いた。表-1、表-2に各モデルの層数やパラメータ数といった特徴を示す。

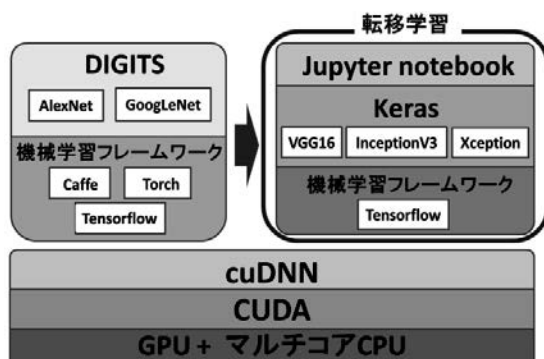


図-1 計算機システム構成図

表-1 DIGITSで使用したモデルの特徴

DIGITS	
AlexNet	入力:224*224 層数:18
	パラメータ数:22.0M
GoogLeNet	入力:224*224 層数:87
	パラメータ数:6.79M

表-2 Kerasで使用したモデルの特徴

Keras	
VGG16	VGG19
入力:224*224 層数:23 パラメータ数:138.3M	入力:224*224 層数:26 パラメータ数:143.6M
InceptionV3	Xception
入力:299*299 層数:159 パラメータ数: 23.8M	入力:299*299 層数:126 パラメータ数: 22.9M

上記システムを用いてImageNetより集めた56種の花の識別について、学習画像の整理、転移学習の導入、神経回路網モデルの更新、パラメータの最適化、を行い識別精度の改善を図った。転移学習については学習画像数と正答率の相関も検証した。パラメータの最適化については、転移学習の際に重要となるパラメータ群、ハイパーパラメータの最適化を行った。今回は学習率、学習率ステップ数、学習率デックイレート、バッチサイズを最適化項目とし、特に学習率を中心に最適値を探索するプログラムを作成し最適値を求めた。さらに得られた学習モデルを用いて、個人で撮影した写真に対する識別テストを行うことで学習モデルの精度を検証した。

(3) 里山を対象とした植物識別システムの構築

北九州市立大学ひびきのキャンパス校内にある里山を対象にデジタルカメラとドローンで撮影した写真を用いて深層学習による識別システムを構築した。精度検証には画像はデジカメとドローンによるものの他、里山全域を映したパノラマ写真から切り出した樹冠画像を用いた。樹冠画像の切り出しはMeanShift法によるスーパーピクセル画像を基に行った⁴⁾。

3. 研究結果

(1) 高精度識別システムの構築

ImageNetより集めた花画像56種の精度改善を行った。はじめに学習に適さない画像を除去し学習画像の質を整理した。その結果56種69,321枚あった画像は47種54,148枚となった。この47種について学習率(lr)を調整するパラメータ調整を行って学習させたところ、最大で82.5%の正答率を得た。その結果を表3に、学習率調整の推移を図-2に示す。この時の深層学習システムにはDIGITSを用い、機械学習フレームワークはTensorflow、神経回路網モデルはGoogLeNetをそれぞれ用いた。また表-4に最終的に識別した47種の植物の内訳を示す。

次に更なる精度向上のため深層学習システムをKerasに移して転移学習や神経回路網モデルの更新に取り組んだ。

表-3 各学習率調整手法による正答率の最高値

lr減少手法	正答率(%)
Fixed	80.0
StepDown(arbitrary steps)	82.49
Exponential Decay	77.75
Sigmoid Decay	81.65

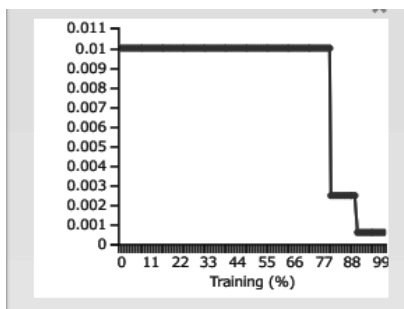


図-2 学習率減少の推移

表-4 47種の花画像の内訳

名前	枚数	名前	枚数
ageratum	1016	marigold	1156
anemone	1162	nigella	1093
aquilegia	1143	orchid	1674
begonia	1520	peony	1207
bellwort	942	petunia	1764
brompton_stock	550	phacelia	1355
browallia	305	poppy	1069
calceolaria	864	portulaca	1121
candytuft	611	rue_anemone	876
catananche	521	sandwort	774
centaury	484	scabious	1520
cornflower	863	shortia	300
commelina	1091	silene	1691
coneflower	1110	spathiphyllum	1016
corydalis	1094	streptocarpus	873
cyclamen	1396	sunflower	1435
dahlia	1256	sweet_allyssum	957
daisy	4523	tidytips	881
dames_violet	692	toadflax	1339
delphinium	1751	tuberosa	786
four_oclock	1011	verbena	1529
gazania	1425	veronica	1424
gentian	1312	zinnia	1054
lesser_celandine	612		

神経回路網モデルはVGG16, InceptionV3, XceptionといったILSVRCで近年好成績を収めたものやその改良モデル3種を使用した。まずは各モデルと転移学習を用いて47種の識別を行った。その結果を表-5に示す。ここでいう上位とは、上位3位を例にとると、識別時に候補を3つまで選びその中に正解がある場合は正しく識別できたとみなす、ということを表している。VGG16では1位の正答率

表-5 各モデルの正答率

	上位1位	上位2位	上位3位
VGG16	0.835	0.912	0.933
InceptionV3	0.898	0.946	0.967
Xception	0.901	0.951	0.968

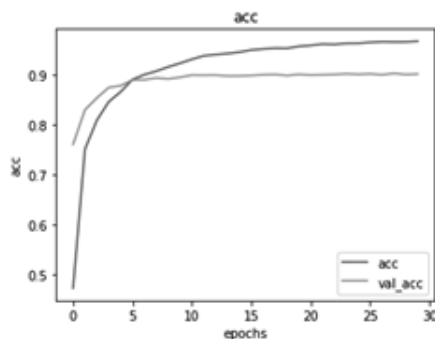


図-3 Xceptionの学習曲線

表-6 各モデルのハイパーパラメータ最適値

モデル名	lr	lrステップ数	lrデックイレート	バッチサイズ
VGG16	0.001	6	0.2323	16
Xception	0.01	6	0.6	16

表-7 最適化後の各モデルの正答率

	上位1位	上位2位	上位3位
VGG16	0.836	0.911	0.943
InceptionV3	0.899	0.947	0.964
Xception	0.916	0.958	0.972

が80%前半に留まっているが、InceptionV3やXceptionでは90%程度の正答率が得られた。正答率が最も高かったのはXceptionで図-3にその学習曲線を示す。さらにハイパーパラメータの最適値を、手動調整やプログラムによる最適値探索による複数回学習から探索した。求めたハイパーパラメータの値を表-6に、この値を用いた各モデルの正答率を表-7に示す。なおInceptionV3はXceptionの元になっているので、Xceptionの値を流用している。表-5と表-7を比べると、ほぼ全てにおいて正答率の向上が見られ、Xceptionが今回の識別において最も高い精度を得た神経回路網モデルとなった。以降の学習では全てXceptionと表-6の値を用いている。ここで構築したシステムはImageNetを使い学習モデルを作成したが、オリジナル画像に対しても同様の結果が得られるのが検証を行った。オリジナル画像は北九州市門司区の

表-8 オリジナル画像による学習モデルの正答率

モデル名	上位1位	上位2位	上位3位
Xception	0.993	0.998	0.999

表-9 画像枚数と正答率

画像数	上位1位	上位2位	上位3位
32470	0.914	0.958	0.972
27063	0.898	0.949	0.965
21640	0.906	0.949	0.964
16223	0.871	0.932	0.956
10812	0.882	0.932	0.954
5393	0.838	0.913	0.94
4308	0.821	0.898	0.927
3227	0.809	0.887	0.918
2142	0.717	0.826	0.877
1063	0.705	0.813	0.865

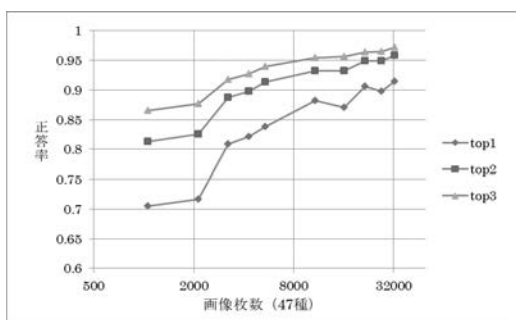


図-4 画像枚数と正答率の相関

白野江植物公園で撮影した植物 53 種を用いている。結果を表-8に示す。99.3%という高い正答率を得ることができ、オリジナルな撮影画像に対しても高い精度を持った植物識別システムを構築することができた。深層学習には大量の画像が必要であるが、ImageNetのようなデータベースを利用せず、オリジナルな画像で学習データを用意する場合その確保が問題となる。

最後に画像数と正答率の相関を調べた。その結果を表-9に示す。表-8と図-4より画像数が5393枚を下回ると上位1位、上位2位において、上位3位においても4308枚を下回ると正答率の低下が激しくなっている。この結果から単純に一種あたり約115枚以上の画像があれば高精度な学習が行えると言える。

(2) 里山を対象とした植物識別システムの構築

北九州市立大学キャンパス校内にある里山を対象に、デジカメとドローンを用いて撮影した樹木画像について学習させた。撮影は樹木の葉を対象に行った。識別に用いた樹木の種名と撮影枚数を表-10に示す。これらの撮

表-10 里山で識別に用いた樹木の種名と画像数

種名	撮影数
アカマツ	142
アカメガシワ	99
アラカシ	102
ハチク	33
ヒメユズリハ	206
クロキ	34
ツブラジイ	115
ヤマモモ	66

表-11 里山樹木の正答率

モデル名	上位1位	上位2位	上位3位
Xception	0.999	1.000	1.000

表-12 里山樹木の識別テスト

種名	テスト数	上位1位	上位2位	上位3位
アカマツ	429	429	0	0
アカメガシワ	310	310	0	0
アラカシ	366	366	0	0
ハチク	104	104	0	0
ヒメユズリハ	669	668	1	0
クロキ	85	85	0	0
ツブラジイ	365	365	0	0
ヤマモモ	205	203	2	0

影画像を基に、部分画像を切り出し十分な学習データ数を確保し、学習を行った。その結果を表-11に示す。8種類と少数の識別なので非常に高い正答率が得られ、デジカメ画像、ドローン画像を用いた識別テストの結果も表-12に示すように制度の高いものとなった。これより里山のような自然に生えている植物においても高い精度が得られていることが分かった。

さらに里山保全におけるドローン調査を想定し、上空から里山全域を撮影したパノラマ写真に写っている樹冠画像の識別ができるかテストを行った。樹冠画像は図-5に示すパノラマ写真を基に、スーパーピクセル分割により切り出した。その過程を以下に示す。まずパノラマ写真を縮小しモザイク化と以降の処理の高速化を図る。縮小した画像にMeanShift法を用いて画素を平滑化しスーパーピクセルを生成する。これによりパノラマ写真を概ね樹冠単位で区分した画像に加工する。加工画像に対しSLICを用いて領域分割を行い、その結果を元のパノラマ写真へ適応させた。最後に領域分割されたパノラマ写真から分割区分の一つの画像として切り出し樹冠画像群を生成した。その結果を図-6に示す。図-6より樹冠画像を切り出した。この方法により切り出した樹冠画像数は



図-5 里山全域を収めたパノラマ写真



図-6 樹冠の分割結果

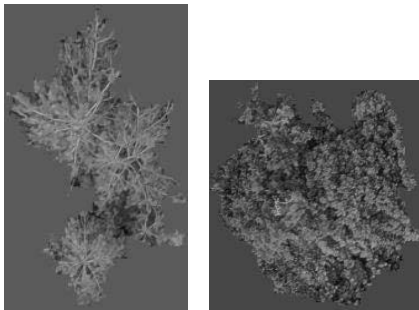


図-7 樹冠画像の一例

517枚だった。切り出した樹冠画像の一例を図-7に示す。

これらの樹冠画像の中から識別に用いた8種について、位置情報等から特定できたものを用いて構築した学習モデルが正しく識別できるかテストを行った。その結果を表-14に示す。アカメガシワは低木なため樹冠画像は確認できなかった。アカマツにおいては高い識別精度を示したが、他の種については、全体的に正しく識別できていなかった。ハチクにおいては上位3位まで含めると正しく識別できているが、上位1位の候補はアカマツでその確からしさは9割を超えていた。このことからハチクは正しく識別できているとはいえない。また517枚の樹

表-14 空撮画像による識別テスト

種名	テスト数	上位1位	上位2位	上位3位
アカマツ	9	8	0	1
アラカシ	5	0	1	0
ハチク	10	0	9	1
ヒメユズ リハ	8	0	0	1
クロキ	10	0	4	2
ツブラジ イ	8	0	0	1
ヤマモモ	14	0	1	4

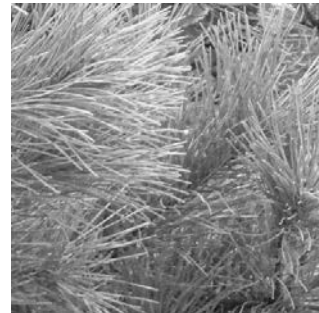


図-8 アカマツの学習画像一例

冠画像に対してこの学習モデルを用いて識別テストを行い8種の中で何と識別するのかを確認したところアカマツと答えた画像は399枚あった。これらの結果よりこの学習モデルは樹冠画像を全体的にアカマツと誤認していると言える。原因としてはアカマツが他の樹木と比べ特徴がはっきりしており、学習に用いた画像数も撮影の段階で数が確保できていたことが考えられる。また樹冠画像に用いたパノラマ写真は高度があり木々の特徴が潰れている。アカマツは図-8に示すように針状の葉が密集しており、学習時画像は299×299まで縮小されることから画像内に輪郭の不鮮明な緑が大部分を占め、特徴の潰れた樹冠画像を誤認しやすい状況にもなっていたと考えられる。さらに学習時に樹冠画像を用いておらず、樹冠の特徴を覚えこませていないことも考えられる。これらを考慮し識別精度の改善のためには、学習時に切り出した樹冠画像も用いること、パノラマ写真の高度を下げ樹冠の特徴を多く残せるようにすること、各樹木の撮影データを充実させること等が考えられる。

4. まとめ

学習画像データの充実、神経回路網モデルの更新、転移学習の導入、ハイパーパラメータの最適化により高い

精度を持つ学習モデルを構築することができた。この学習モデルはオリジナルな画像においても高い精度を示すことが確認できた。里山での識別ではデジカメや低空からドローンで撮影した画像については正しく識別できたが、パノラマ写真内の樹冠の識別はアカマツを除き精度が低い結果となった。精度改善には対象との距離も含めた様々なパターン画像の学習が必要になると考えられる。今後は学習データの充実や識別に用いるパノラマ写真の鮮明化を図り識別精度の向上を図っていく。また植物は季節により色合いや造形が変化するので、これについて同様に学習画像の多彩化による精度の変化を検証していく。さらにひびきの学研都市内の他の自然保護地区においてもDL法による高精度識別システムが構築できるのか検証していく。

参考文献

- 1) ILSVRC <http://www.image-net.org/challenges/LSVRC/>
- 2) 曙瑠美, 藤崎莉純夢, 中山紘喜, 西野友子, 野上敦嗣, : 深層学習法における植物識別精度改善の試み, 環境システム

研究論文発表講演集, 46 巻, 241-246, 2018

- 3) 井畔実穂, 福井すみれ, 中山紘喜, 西野友子, 野上敦嗣, : 機械学習ディープラーニング法による植物用の識別, 環境システム研究論文発表講演集, 46 巻, 253-258, 2018
- 4) 鈴木太郎, 土屋武司, 鈴木真二, 山場淳史 : 小型 UAV による空撮画像からのスーパーピクセル分割と機械学習を用いた植生分類手法の構築, Journal of The Remote Sensing Society of Japan, 36, 59-71, 2016

(Received June 19, 2019)

Construction of The Highly Precise Plant Identification System by The Deep Learning Method and Adaptation to Satoyama Conservation

Hiroki NAKAYAMA, Rumi AKEBONO, Miho IGURO, Tomoko NISHINO, and
Atsushi NOGAMI

Identification of the kinds is an important factor in an ecosystem investigation by the Environmental Impact Assessment and the environmental conservation activity of the Satoyama. However, the knowledge of the expert is necessary for accurate identification. On the other hand, the applied study of the image recognition technology by the AI advances since the establishment of the deep learning method (DL method), and the effect is expected in the identification of the kind. I built the plant identification system using the DL method until now in this laboratory, but, in the identification of 56 kinds of flowers, the precision remained in 62%. I coordinated the metastasis learning, and the high precision planned becoming it in the identification of the plant in this study. As a result, precision to exceed 90% in 47 kinds of identification was provided. In addition, I got equal precision in 53 kinds of identification that I photographed originally in a plant park. With that in mind, for the trees of a Satoyama becoming the environmental conservation district, I inspect the identification precision of an image and the top of the tree image using a digital camera and the drone and examine adaptation to Satoyama conservation. I used Keras for deep learning system and used the MeanShift method for the making of the top of the tree image.