

イチリンソウとその近縁種における隣接花卉の重なり順推定

中谷 友哉^{1,a)} 内海 ゆづ子^{2,b)} 藤本 仰^{3,c)} 岩村 雅一^{2,d)} 黄瀬 浩一^{2,e)}

概要

本研究では、1 枚の花の画像から隣接する花卉（花びら）の重なり順の推定をする。花の発生を調べる取り組みの一環として、隣接する花卉の前後関係（重なり順）を画像から自動で推定することが望まれている。しかし、花の見た目が個体ごとに大きく異なる上、野生の花を対象とするため、推定に十分な量の画像を収集することは困難である。本研究では、このような多種・少量のデータで隣接する花卉の重なり順を推定するために、2 つの方策をとる。1 つは、メタ学習のフレームワークを用いて、多様なデータを効率的に学習すること、もう 1 つは、合成画像を用いて画像の不足を補うことである。実験により、上記 2 つの方策を用いる提案手法は、これらを用いない場合に比べて推定精度が 11.8% 向上することを確認した。

1. はじめに

花の形状は多種多様であり、その違いを特徴付けるものとして、花卉（花びら）の枚数や配置、対称性の有無などがある。中でも花卉の配置は、近い種でも異なるものや遠い種でも同じものが存在する。そのため、花の発生には種を超えた共通の原理が存在することが予測されている [5]。このような花の多様性がどのように生まれるのかを解明するため、数理モデルを用いて花の発生をモデル化し、計算機シミュレーションで発生過程を調査する研究が行われている [5]。このような花の発生モデルの考案や検証を行うには、野生の花を観察し、花卉の発生順序とモデルを照らし合わせる必要がある。

花卉の発生順序は、図 1 のような開ききる前の花画像から、隣接する花卉の重なり順を元に推定できる。花卉の発生順序を調べる作業は現在人手で行われている。そのため、労力の軽減を目的として、この作業の自動化を目指す。

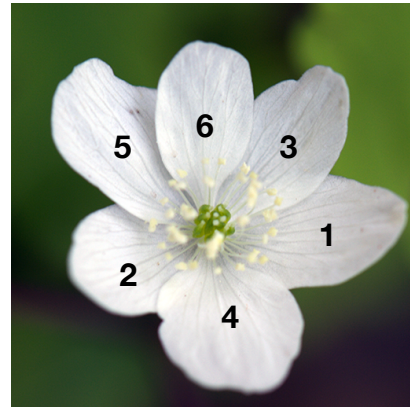


図 1 本稿で用いるイチリンソウとその近縁種の花の画像例。花卉上の数字は発生順序を示す。



図 2 本稿で用いるイチリンソウとその近縁種の多様性。花卉の形状、色がそれぞれ異なる。

本稿では、花卉の発生順序を自動で推定するための第一歩として、隣接する花卉の重なり順を画像から推定する手法を提案する。

今回、対象とするのは、花の発生モデルを研究するために収集されたイチリンソウとその近縁種の花の画像データである。このデータは、花の咲いているところに赴き、花卉の順序が分かるものを 1 つ 1 つ手で撮影して収集される。そのため、画像認識で頻繁に用いられる ImageNet [1] や MS COCO [4] といった大規模データセットとは異なり、画像枚数が少ない。画像枚数の少なさに加え、問題となるのが、花卉の見目の多様性である。図 2 に本研究で用いる花画像データの一部を示す。花卉の見目が個体によって大きく異なり、また画像の総枚数が少ないことからわかるように、見目が類似する画像はさらに少ない。

この多種・少量のデータに対応するため、本研究では 2 つの方策を導入する。まず、重なり順の推定にメタ学習手法の Model-Agnostic Meta-Learning (MAML) [2] を導入

¹ 大阪府立大学 工学域 情報工学課程

² 大阪府立大学 大学院工学研究科 電気・情報系専攻

³ 大阪大学 大学院理学研究科 生物科学専攻

a) nakatani@m.cs.osakafu-u.ac.jp

b) yuzuko@cs.osakafu-u.ac.jp

c) fujimoto@bio.sci.osaka-u.ac.jp

d) masa@cs.osakafu-u.ac.jp

e) kise@cs.osakafu-u.ac.jp

することで、多種の問題に対処する。MAML には、学習済みの識別器を新たなタスクに適応させる仕組みがあるため、多種多様な花画像の認識が可能である。また、大量の合成画像を作成して学習に用いることで少量の問題に対処する。合成画像は、学習データを大幅に増やせる上に、自動で正解ラベルが付与できることから、学習データを効率的に用意することができる。

我々が独自に収集したイチリンソウとその近縁種のデータを用いた実験で提案手法を評価したところ、メタ学習と合成画像の両方を用いた提案手法が、どちらも用いなかった場合と比べて 11.8% 精度が向上することが確認できた。

2. 関連研究

花卉の重なり順推定の関連研究として、花画像の観察に画像処理が利用された例を紹介する。また、多種・少量の画像データを対象としたパターン認識を行うことから、これに関する few-shot learning のためのメタ学習の研究も示す。

2.1 花を対象とした画像処理

花は身近な物体であり、コンピュータビジョン、パターン認識の分野において認識対象の 1 つとなっている。代表的な一般物体画像データセットの 1 つである ImageNet [1] にも花のカテゴリが含まれている。また、花の形状や色が様々であることから、花の種類の識別も行われてきた。コンピュータビジョンの分野では、花画像データセットである Oxford flower datasets [6, 7] が知られており、広く利用されている。このデータセットは 2 種類あり、1 つ目は 17 種類、各 80 枚の画像、2 つ目は 102 種類、各種 40 枚から 258 枚の画像から構成されている。このデータセットを用いて、画像を識別する手法が様々提案されてきた [6, 7, 11, 12]。また、セグメンテーション [8, 13] も行われている。

2.2 few-shot learning のためのメタ学習

Deep learning をベースとする画像認識で良好な結果を得るためには大量のラベル付きデータが必要となる。しかし、本研究のように画像の収集が容易でなかったり、学習用のラベルを付与するコストが高すぎたりして、データを大量に収集することが困難なタスクも多く存在する。そのような、限られた少数データを用いて行われる学習は few-shot learning と呼ばれ、高精度な認識を実現するため、様々な試みが提案されてきた。

few-shot learning のアプローチの 1 つにメタ学習がある。メタ学習は、様々なタスクの学習から、新しいタスクの学習の効率化を図る枠組みである。効率化する対象は手法により異なっている。MAML [3] では、新たなタスクに適応しやすいモデルの初期パラメータを推定することで、学習の効率化を図る。他に、パラメータではなく、最適化

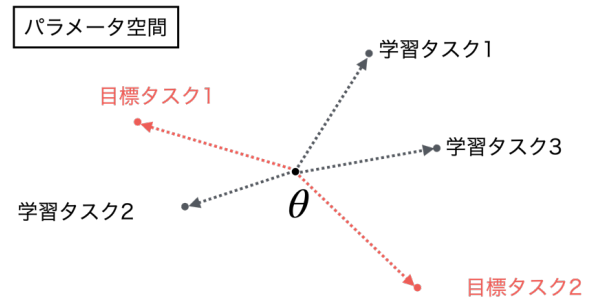


図 3 MAML における最適な初期パラメータ θ とタスクの関係。初期パラメータは特定のタスクに依存しない。矢印は追加学習によるモデルの適応を表す。

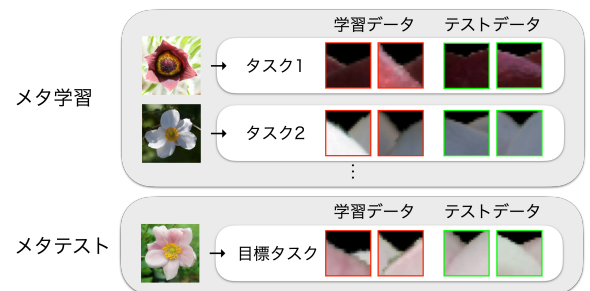


図 4 本研究でのタスクの構成とメタ学習の枠組み。学習データに合成パッチ、テストデータに自然パッチを用いる。

手法に着目し、Long Short-Term Memory (LSTM) ベースのネットワークを用いて対象のモデルの最適化を学習する手法 [9] などが提案されている。

3. 提案手法

3.1 タスクの定義

花卉の重なり順推定のタスクを具体的に説明する。まず、図 6 の右から 2 つ目の画像のように、花卉の重なっている点を検出し、その点の周囲を切り出す。この切り出した正方領域をパッチと呼ぶ。そして、図 6 の右端の図のように、花卉の重なり部分が下部になるようにパッチを回転した時に、左右どちらの花卉が手前かを推定する。そのため、本研究で取り扱うタスクは、パッチ画像に対し、左右どちらの花卉が手前かを推定する 2 クラス分類問題となる。

3.2 メタ学習を用いた花卉の重なり順推定

1 で述べた通り、本研究の対象であるイチリンソウとその近縁種は、個体により見た目が大きく異なる。そのため、一般的に機械学習で行われているような、学習データから最適な推定器のパラメータを学習し、得られたパラメータを用いて推定をする方法では、見た目の違いをうまく学習できず、推定精度が低下する可能性がある。そこで、本研究では、メタ学習を用いることで、見た目の違いに合わせた推定を実現する。具体的には、MAML [2] を用いる。

図 3 に MAML の概要を示す。MAML は、最適な推定器の初期パラメータ θ を求め、追加学習を行うことで目

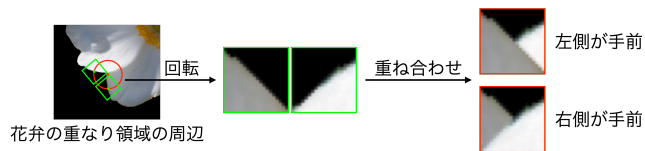


図 5 合成パッチの作成方法.

標タスクでの良好な推定を行う手法である．本研究では，図 4 のように，タスクをそれぞれ 1 つの花での重なり推定とする．そして，タスク内での学習データとして，人工的に作成した合成パッチを利用し，推定を行うテストデータを花の画像からそのまま切り出した自然パッチとする．合成パッチの作成方法は 3.3 で説明する．

3.3 合成パッチ

本研究では，学習用のデータを増やすための処置として，花卉の重なり部分を合成し利用する．合成パッチは，花卉の数でその数が決まる自然パッチとは異なり，自由に枚数を調整出来る上，自動で正解ラベルを付与することができるため，低コストで学習データを増やすことができる．

合成パッチの作成方法を，図 5 に示す．はじめに，花卉の重なり領域の付近から，左右どちらか一方の花弁を含む自然パッチを 2 枚切り出す．それらを重ね合わせることで，左側が手前のものと右側が手前のもの，2 種類の合成パッチを得る．

4. 実験

メタ学習と合成パッチ導入の有効性を検証するため，合成パッチを用いた場合と用いない場合，メタ学習を用いた場合と用いない場合の組み合わせ計 4 通りで自然パッチの識別率を評価する．また，合成パッチの総数や，タスク内の合成パッチの数を変化させ，合成パッチが識別精度に与える影響を検証する．

4.1 実験データ

実験に用いた花卉のパッチは，イチリンソウとその近縁種の画像 239 枚をもとに作成した．これらの画像は，2016 年から 2020 年に滋賀県，京都府，兵庫県，岡山県，広島県他にて，Olympus 社製デジタルカメラ TG-5 と，Nikon 社製一眼レフカメラ D5200 を用いて撮影した．

花画像からパッチを得るまでの流れを図 6 に示す．初めに Grabcut [10] による花領域の切り出しを行った．続いて，Harris のコーナー検出手法により花卉の重なり候補点を検出し，検出点の付近の様子から重なり部分であるかを判別した．最後に，検出された重なりからパッチの切り出しを行い，1 つの花あたり 4 枚から 16 枚，合計で 1,575 枚の自然パッチが得られた．また，3.3 で説明した手順により合成パッチを 1 つの花あたり 1,250 枚，合計で 298,750 枚作成した．全てのパッチの大きさは 20×20 pixels である．

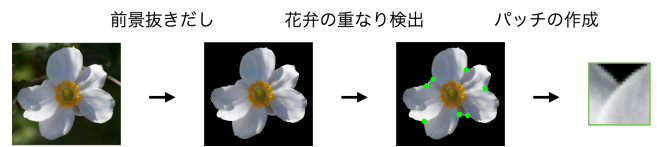


図 6 元画像から自然パッチを得るまでの手順.

239 枚の花画像のうち，200 枚を学習用タスクに，残りの 39 枚をテスト用タスクに利用した．学習用タスク，テスト用タスクの花画像から生成されたパッチは，それぞれ学習用とテスト用のパッチプールに入れる．そして，後述するように，メタ学習を使う際には必要枚数をパッチプールからサンプリングして用いる．画像処理には，Python 版の OpenCV を利用した．

4.2 実験 1

合成パッチを用いた場合と用いない場合，メタ学習を用いた場合と用いない場合の組み合わせ計 4 通りで自然パッチの識別率を評価する．

合成パッチとメタ学習を併用する提案手法では，学習には 20 万タスク，テストには 100 タスク作成して使用した．ここで各学習タスクは，学習用 200 枚の花画像から 1 枚選び，その花画像に対応するパッチをパッチプールからランダムにサンプリングして作成する．テストタスクも，テスト用 39 枚の花画像から同様に作成する．各タスクでは，学習データとして 10 枚 $\times 2$ クラス，計 20 枚の合成パッチを，テストデータとして 1 枚 $\times 2$ クラス，計 2 枚の自然パッチを用いる．合成パッチを用いないメタ学習手法では，学習データを自然パッチ 1 枚 $\times 2$ クラスの合計 2 枚，テストデータを自然パッチ 1 枚 $\times 2$ クラスの合計 2 枚で各タスクを構成し，学習用のタスクを 20 万個，テスト用タスクを 100 個作成した．メタ学習を用いずに合成パッチを用いる場合は，学習用の花画像 200 枚から作成した合成画像 248,750 枚で学習した後，学習用の花画像に含まれる自然パッチ 1,303 枚で fine-tuning し，テスト用の花画像 39 枚に含まれる自然パッチ 272 枚を用いて精度を評価した．合成パッチ，メタ学習ともに利用しない場合，すなわちベースライン手法では，自然パッチ 1,303 枚を用いて学習し，272 枚のパッチを用いて評価を行った．

実験は NVIDIA TITAN V シングル GPU で行った．ネットワークのモデルは [2] と同じものを用いた．ネットワークの事前学習は行っていない．実験は 15 回行い，得られた精度の平均と標準偏差により評価を行った．

実験結果を表 1 に示す．ベースライン手法と比較すると，メタ学習を用いた場合は 2.4%，合成パッチを用いた場合は 0.3%，メタ学習と合成パッチの両方を用いた提案手法は 11.8% 精度が向上した．この結果から，合成画像とメタ学習の併用が重なり順の推定に有効であることが確認できた．

表 1 [実験 1] 隣接花卉の重なり順の推定精度.
15 回の平均と標準偏差.

手法	精度 (%)
ベースライン	84.1 ± 2.1
メタ学習	86.5 ± 3.4
合成パッチ	84.4 ± 1.9
メタ学習+合成パッチ (提案手法)	95.9 ± 1.6

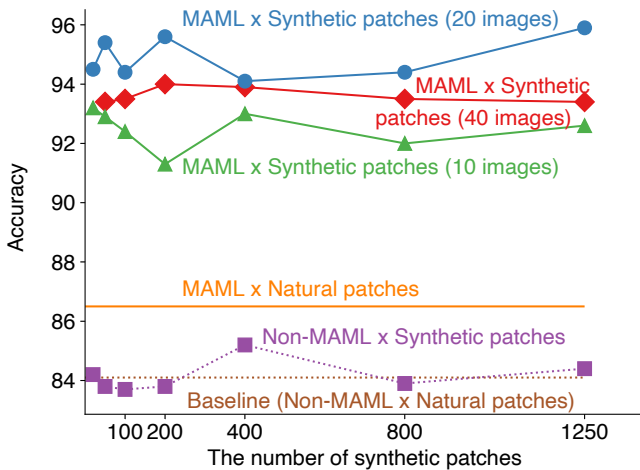


図 7 [実験 2] 隣接花卉の重なり順の推定精度 (%). 横軸は実験に用いた花個体ごとの合成パッチの数を示す.

4.3 実験 2

メタ学習時の合成パッチの枚数が精度に与える影響を評価するため、提案手法でタスク内の学習データ（合成パッチ）の数を 10, 20, 40 と変化させ、評価した。これに加えて、合成パッチを用いる手法では、合成パッチプールのサイズを 20 から 1250 の範囲で変化させ、評価した。記載の無い条件は実験 1 と同じである。

実験結果を図 7 に示す。図 7 の横軸は、タスク生成の際に用いる合成パッチプールのサイズを示し、縦軸は推定精度を示す。また、“MAML”、“Non-MAML” は MAML の使用の有無を表し、“Natural patches” と “Synthetic patches” は、学習データに用いたパッチの種類を表す。括弧内に画像の枚数が書いているものは、タスク内の学習データに合成パッチを利用しており、その枚数を示している。タスクに含まれる合成パッチの数と識別精度の関係から、合成パッチであっても学習データを増やすと、自然パッチに対する識別精度が向上することがわかった。一方で、花 1 個体あたりの合成パッチの総数を変化させても識別精度に大きな変化は見られなかった。また、MAML を用いた場合、タスクの学習データである合成画像を 10 枚から 20 枚に増やした場合は精度が向上したが、40 枚の場合は 20 枚の時と比較して精度が悪化した。

5. まとめ

本稿では、花卉の発生順序を自動で推定するための第一歩として、花卉の重なりを含むパッチから、隣接する花卉

の重なり順を推定した。今回扱ったイチリンソウとその近縁種の花画像データは、見た目が多様で、同種のデータを大量に入手出来ない多種・少量のデータであるため、従来の認識手法の適用が難しい。そこで本研究では、メタ学習と合成画像、2 つの方策を導入することで精度向上を図った。具体的には、メタ学習のフレームワークを用いることで多種の問題に対処し、合成画像を導入することで、少量の問題に対処した。実験の結果、メタ学習と合成パッチの両方の導入により、認識精度が 11.8% 向上することを確認した。

謝辞

本研究は、JSPS 科研費 JP20H05423 の補助による。

参考文献

- [1] Deng, J., Dong, W., Socher, R., Li, L.-J., Kai Li and Li Fei-Fei: ImageNet: A large-scale hierarchical image database, *Proc. CVPR* (2009).
- [2] Finn, C., Abbeel, P. and Levine, S.: Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks, *Proc. ICML* (2017).
- [3] Finn, C., Abbeel, P. and Levine, S.: Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks, *Proc. ICML* (2017).
- [4] Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P. and Zitnick, C. L.: Microsoft COCO: Common Objects in Context, *Proc. ECCV* (2014).
- [5] Nakagawa, A., Kitazawa, M. S. and Fujimoto, K.: A design principle for floral organ number and arrangement in flowers with bilateral symmetry, *Development*, Vol. 147, No. 3 (2020).
- [6] Nilsback, M. E. and Zisserman, A.: A visual vocabulary for flower classification, *Proc. CVPR* (2006).
- [7] Nilsback, M.-E. and Zisserman, A.: Automated flower classification over a large number of classes, *Proc. Indian Conf. on Comp. Vis., Graph. & Image Proces.*, pp. 722–729 (2008).
- [8] Nilsback, M. E. and Zisserman, A.: Delving deeper into the whorl of flower segmentation, *Image and Vision Computing*, Vol. 28, No. 6, pp. 1049–1062 (2010).
- [9] Ravi, S. and Larochelle, H.: Optimization as a Model for Few-Shot Learning, *Proc. ICLR* (2017).
- [10] Rother, C., Kolmogorov, V. and Blake, A.: GrabCut: interactive foreground extraction using iterated graph cuts, *Proc. ACM SIGGRAPH* (2004).
- [11] Wang, Z., Wang, K., Wang, X. and Pan, S.: A Convolutional Neural Network Ensemble for Flower Image Classification, *Proc. ICCV* (2020).
- [12] Xie, L., Tian, Q., Wang, M. and Zhang, B.: Spatial pooling of heterogeneous features for image classification, *IEEE Trans. on Image Proces.*, Vol. 23, No. 5, pp. 1994–2008 (2014).
- [13] Yuning Chai, Lempitsky, V. and Zisserman, A.: BiCoS: A Bi-level co-segmentation method for image classification, *Proc. ICCV* (2011).