

多種少量データを対象とした隣接花卉の重なり順推定

中谷 友哉[†] 内海ゆづ子^{††} 藤本 仰一^{†††} 岩村 雅一^{††} 黄瀬 浩一^{††}

[†] 大阪府立大学 工学域 情報工学課程 〒599-8531 大阪府堺市中区学園町 1-1

^{††} 大阪府立大学 大学院工学研究科 電気・情報系専攻 〒599-8531 大阪府堺市中区学園町 1-1

^{†††} 大阪大学 大学院理学研究科 生物科学専攻 〒560-0043 大阪府豊中市待兼山町 1-1

E-mail: [†]nakatani@m.cs.osakafu-u.ac.jp, ^{††}{yuzuko,masa,kise}@cs.osakafu-u.ac.jp, ^{†††}fujimoto@bio.sci.osaka-u.ac.jp

あらまし 花の器官の発生順序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）を画像だけから自動で推定することが望まれている。しかし、花の中にはイチリンソウとその近縁種のように、個体ごとに見た目が大きく異なるため、見た目が近い（均質な）画像を多量に集めにくいものもある。本稿では、このような多種少量データを用いる場合の対処法として、2つの方策をとる。1つ目はメタ学習の導入である。すなわち、多数の個体に共通の特徴をメタ学習時にあらかじめ求めておき、認識（メタテスト）時にはそれを初期値として、学習モデルを認識対象の個体に適応させる。2つ目は、重なっていない花卉の画像から花卉の重なりを模した画像を合成して学習に用いることである。実験により、上記2つの方策を用いる提案手法は、両方用いない場合に比べて、認識精度が11.8%向上することを確認した。

キーワード 植物計測, 花卉の発生順序, メタ学習, 画像合成

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）の枚数や形状の対称性の有無などがある。これらの特徴のうち、花卉の対称性は、近い種でも異なる場合や、遠い種でも同じ場合があるなど、容易に切り替わることが知られている。そのため、花の発生過程には、種を超えて共通した原理が存在することが予測されている。植物が花の形状の多様性をどのように表現しているかを解明するため、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どのように花の器官が形成されていくかを解明する研究が行われている [1]。このような花の発生モデルの考案や検証を行うには、実際に咲いている花の器官の対称性や花卉の発生順序を観察して、モデルと照らし合わせる必要がある。

本研究では、花卉の発生順序に焦点を当てる。花卉の発生順序は、図1のように、花卉同士が重なっている開ききる前の花を観察したとき、隣接する花卉の重なり順からある程度推定できる。ただし、図1の4番と5番の順序関係のように、重なり順だけでは分からない順序関係もある。このような足りない情報は花卉の位置関係の事前知識をもとに推定される。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。大量の花の画像を目視で調査するのは非常に労力がかかるため、その軽減のために、本研究ではこの作業の自動化を目指す。

本稿では、その第一歩として、隣接する花卉の順序を画像から推定する。対象とするのは、花の形状の発生モデルの研究の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、



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。

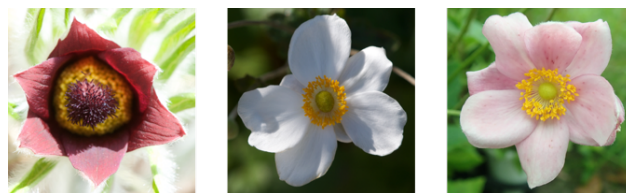


図2: 本稿で用いるイチリンソウとその近縁種の多様性。花卉の形状、色がそれぞれ異なる。

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た。そのため、画像認識で頻繁に用いられる ImageNet [2] や MS COCO [3] といった大規模データセットとは異なり、画像

枚数は高々 200 枚強である。

画像枚数の少なさも然る事ながら、このデータを用いる上で問題となるのが、花卉の見た目の多様性である。図 2 に示すように、花卉の見た目が個体によって大きく異なり（多種）、また画像総枚数の少なさからわかるように、見た目が類似する花の画像が少ない（少量）。

このような多種少量データを用いる識別問題を扱うにあたり、本稿では 2 つの方策を導入する。1 つは、メタ学習の枠組みの 1 つである Model-Agnostic Meta-Learning (MAML) [4] の導入である。従来の一般的な識別問題は学習とテストから成るのに対して、メタ学習の枠組みで従来の学習とテストに対応するのはメタ学習とメタテストである。MAML は、メタテストで与えられる未知の問題（タスク）のために、メタ学習で識別器の良い初期パラメータを推定する。つまり、多数の個体に共通の特徴をメタ学習時にあらかじめ求めておき、認識（メタテスト）時にはそれを初期値として、学習モデルを認識対象の個体に適応させる。そのため、MAML は多種多様なデータに向いていると考えられる。もう 1 つは、画像の合成による学習データの水増しである。本稿で用いる多種少量データでは、見た目が近い（均質な）画像が少ないため、例え MAML を用いたとしても、十分な認識精度が得られない懸念がある。そのため、重なっていない花卉の画像から花卉の重なりを模した画像を合成して学習に用いる。画像を合成して学習に用いれば、学習データを大幅に増やせる上に、合成した画像に自動で正解ラベルが付与できるので、手動でのラベル付けを避けられるという大きな利点がある。その反面、合成した画像が、本来の花の重なり領域の画像と乖離しており、学習に寄与しないという懸念もある。そこで、上記 2 つの方策を導入した提案手法の有効性を、前述のイチリンソウとその近縁種の花画像データセットを用いた実験で評価する。

2. 関連研究

ここでは、花卉の重なり順序推定の関連研究として、植物計測のうち、花の画像の観察に画像処理が利用された例を紹介する。また、本研究では、多種少量の画像データを対象としたパターン識別を行うことから、これに関係する Few-shot learning の関連研究も示す。

2.1 花を対象とした画像処理

花は身近な物体であることから、コンピュータビジョン、パターン認識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 [2] にも花のカテゴリ含まれている。また、花の形状や色が様々であることから、花の見た目の特徴を用いた種類の識別も研究されてきた。コンピュータビジョン分野では、花データセットである Oxford flower datasets [5], [6] が知られており、広く利用されている。このデータベースは 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。このデータベースを用いて、画像識別手法が様々提案されてきた [5]~[14]。また、識別だけでなくセグメンテーション [15]~[17] も行われている。他にも、独

自にデータセットを作成して識別を行う手法 [18] や、インタラクティブなセグメンテーション手法 [19] が提案されている。

農業の分野では、農作業の自動化や農作物の収量の予測を目的とし、画像からの花の検出が研究されている。温室などの室内環境では、トマトにおける摘花作業の自動化 [20] や、ラズベリーやブラックベリーの受粉作業を自動化 [21] を目的として花の検出手法が提案されている。また、Deep learning により、照明条件にロバストな物体検出が実現可能になったことで、屋外で栽培される作物の花の検出も盛んに行われている。イチゴ [22], [23]、リンゴ [24], [25]、ブドウ [26], [27]、キウイフルーツ [28] などの収量を予測したり、受粉や摘花の作業を自動化するため、画像中の花が検出されている。

花卉などの花の器官の詳細な形状情報をモデル化している研究も存在する。CT を用いて計測した画像から花の花弁を含む詳細な形状を復元する手法 [29]、花の開く様子を時系列でモデル化する手法 [30], [31]、RGB-D データとあらかじめ作成した花弁の色・形状を含むモデルから実際の画像の花形状を復元する手法 [32] などが提案されている。これらは CG での花の描画を目的としてモデルを構築したものであり、3 次元形状を計測したデータや 3 次元のモデル当てはめなどが必要となる。そのため、本研究のような屋外で咲いている花の形状の解析には向かない。

2.2 Few-shot learning

Deep learning をベースとする画像認識で良好な認識結果を得るためには大量の正解ラベル付き学習データが必要である。しかし、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、学習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。本研究で用いる花画像データも画像の収集に大きなコストがかかることから、大量に収集できないタスクの 1 つである。そのようなタスクの中には、カテゴリ毎に学習に使用できるデータ数が異なる場合がある。一部のカテゴリには多量のデータがあり、そのデータを用いる事で、少数のラベル付きデータしか存在しないカテゴリを学習しようという試みは 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。

1 つ目のアプローチとして、メタ学習が挙げられる。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 [33] では、新たなタスクに適応しやすいモデルのパラメータの初期値を推定することで、学習の効率化を図る。他に、モデルのパラメータではなく、最適化手法に着目し、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 [34] などが提案されている。特に MAML については、損失関数がモデルパラメータに関して微分可能であれば導入できるという手軽さから、顔認識のためのドメイン適応 [35]、画像品質評価 [36]、ビデオフレーム補間 [37] など様々な分野で応用されている。

Neural network に外部記憶装置とその読み書き機能を付加した Neural Turing Machine (NTM) を few-shot learning に適用する手法 [38]~[40] も提案されている。一般的な Neural Network で

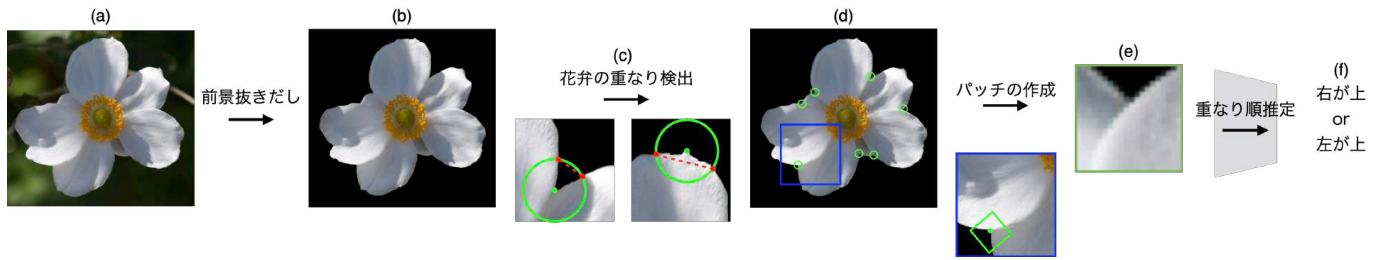


図 3: ベースライン手法の概要.

は、新たな学習データが入力された場合、パラメータの更新が非常に効率が悪く、一方、NTM では、外部記憶装置に新たな学習データの知識を素早く記録できることから、少量のデータで学習を行う few-shot learning に適した手法であるといえる。

Metric learning も few-shot learning で頻りに用いられる手法の 1 つである。Metric learning では、様々なタスクにおいて学習データの埋め込み空間への写像を学習する。この写像は、類似したもの動詞の距離が埋め込み空間上で近く、類似しないもの同士の距離が遠くなるよう学習される。そして、新しいタスクに対して、学習した写像を適用し、埋め込み空間上で識別を行う。識別には、簡単な線形識別器や最近傍法を用いる場合 [41]~[43] もあれば、CNN を用いる場合 [44] もある。

3. 手 法

本稿では、花卉の重なり順推定に MAML [4] と画像の合成を導入する。最初に 3.1 で、これらの方策を導入する前のベースライン手法を述べ、ベースライン手法に適用する形で MAML と画像の合成を 3.2 と 3.3 でそれぞれ説明する。

3.1 花卉の重なり順推定のベースライン手法

花卉の重なり順推定のベースライン手法の概要を図 3 に示す。この手法では、まず認識処理を頑健にするために、撮影された花の画像 (図 3 (a)) から背景領域を削除して、図 3 (b) のように花の領域のみを切り出す。続いて、図 3 (d) のように、花卉の重なっている点を検出し、その周りの領域を切り出す。切り出した領域は、3.3 で作成する「合成パッチ」との対比のために、「自然パッチ」と呼ぶ。最後に、切り出した自然パッチを回転させて向きを揃えた (図 3 (e)) 後、左右のどちらの花卉が手前かを推定する (図 3 (f))。以降、各処理の詳細を説明する。

まず、花領域の切り出しについて説明する。図 1 や図 2 に示す通り、本研究で用いる花の色は背景と異なっている。また、撮影時の照明環境も比較的安定していることから、色情報を用いた手法で花の領域が切り出せる。そこで、本研究では、Grabcut [45] を用いて花の領域を切り出す。Grabcut は Graph cut をベースとして画像の前景と背景を分割する手法であり、ユーザが大まかに前景領域を指定することで、切り出しを行う。今回、手動で前景領域を決定し、Grabcut を適用することで、図 3 (b) のように花領域を切り出した。

続いて、切り出された花領域から花卉の重なり部分を検出する。今回用いた花の花弁の形状は概ね楕円であり、花卉が重

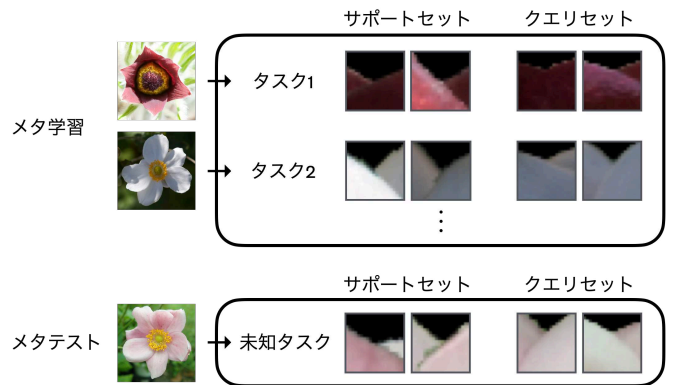


図 4: メタ学習の枠組み。本稿で扱う問題と相性が良い。

なっている部分は鋭く尖った形をしている。そのため、花卉の重なり部分を Harris のコーナー検出手法を用いて検出を行った。但し、画像によっては花卉の重なり部分以外も検出されることがある。花卉の重なり部分のみを取り出すため、検出点の周りの画像の様子から判別を行う。図 3 (c) の緑線で示すように検出点を中心とする半径数ピクセルの円を考え、この円と花の領域の境界線が交差する 2 点を求める。そして、図 3 (c) 右側のように、赤い破線で示された、2 点を結んだ線が花領域の上であれば、重なり部分でないと判別し、図 3 (c) 左側のように、背景領域の上であれば、花卉の重なり部分であると判別する。

花卉の重なり部分から、重なり順推定に用いる正方領域を切り出す。このとき、正方領域の中心は花卉の重なり部分であると判断された検出点であり、図 3 (c) 左図の赤い破線と正方領域の辺が平行になるよう、また、花卉と背景の境界領域が上になるよう、 20×20 pixels の正方領域を切り出す。この切り出した領域のことを自然パッチと呼ぶこととする。このように切り出すことで、図 3 (e) に示すように常に画像の下側に重複した花卉が位置することになる。こうすることで、重なり順序推定の際に左が上か、右が上かの 2 クラス判別問題に帰着させることができる。

そして、最後に、自然パッチを Neural network ベースの識別器に入力することで、左が上か、右が上かを判別する。

3.2 MAML における初期パラメータの学習

本研究では、花の重なり識別の枠組みとして MAML [4] を用いる。MAML は、様々なタスクに素早く適応することのできる「良い初期パラメータ」を獲得する手法である。本研究では、識別対象の花の見た目が異なることから、花の個体ご

とに識別モデルのパラメータをチューニングする必要がある。そのため、MAML を用いてパラメータチューニングを行いやすい初期値を得ることで、個体に適したパラメータチューニングが可能となり、高精度な重なり推定を実現できる。

MAML を説明する前に、初めにメタ学習の枠組みを説明する。従来の一般的な識別問題は学習とテストから成り、データも学習データとテストデータに分ける。それに対して、メタ学習の枠組みは2階建てであり、図4に示すように、従来の学習とテストの上位概念として、メタ学習とメタテストが存在する。メタ学習は複数のタスク、メタテストは1つのタスクから構成される。タスクは、機械学習の個々の問題を指し、各タスクはサポートセットと呼ばれるモデルの学習用データと、クエリセットと呼ばれるテスト用データから構成される。本稿では、1つの花の個体の識別問題（花卉の重なり部分の順序推定）を1つのタスクとして、サポートセット、クエリセット共にその花の重なり位置から切り出した自然パッチから成る。

続いて、MAML における初期パラメータの学習方法について説明する。MAML では、学習用に複数のタスクが与えられた時に、まず、各タスクで損失関数を最小化しようモデルのパラメータの更新を行った後、それらを統合して全てのタスクにおいて損失関数が最小化するように、パラメータを更新する。これを繰り返すことで、様々なタスクでもパラメータチューニングが容易に行える良い初期値のパラメータを求める。識別問題を扱う場合、各タスクでのパラメータ更新にはサポートセットを用い、各タスクを統合したパラメータの更新はクエリセットを用いて行う。

3.3 合成パッチ生成

3.1 で得られた自然パッチの数は、画像1枚あたり多くて数枚であり、学習に用いるには不十分である。それに加えて、実際のタスクでは学習に用いる自然パッチに正解ラベルを付与する必要があり、大変な時間と労力を要する。そこで、本研究では自動的に正解ラベルが付与でき、かつ、学習に十分に必要枚数を増やせるという利点がある合成データを作成し、学習に利用する。合成されたパッチを自然パッチとの対比で「合成パッチ」と呼ぶ。

合成パッチの生成方法の詳細は、図5の通りである。まず、3.1 で検出された花卉の重なり位置（図3(c)の左側の画像の緑色の点）から、それぞれの花卉の輪郭を追跡して、十分離れた位置で合成パッチを構成する2枚のパッチ（これを便宜上、元パッチと呼ぶ）を切り取る。ここで十分離れた位置とは、輪郭上で所定の大きさの正方領域を切り取ったとき、切り取った領域中に花卉の重なり位置が入らない位置を言う。そして、それらの2つの元パッチを重ねることで、合成パッチが生成される。合成パッチの右前、左前の正解ラベルは、左右のどちらかを先に重ねて画像を合成させたかが明らかなので、自動的に付与可能である。

4. 実験

本研究では、二つの実験を行った。まず、合成パッチとMAMLの有効性を確認するため、合成パッチを用いた場合と

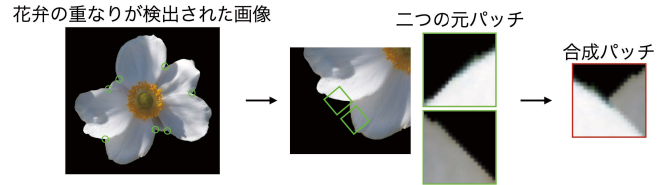


図5: 合成パッチの作成方法

	サポートセット (合成パッチ)	クエリセット (自然パッチ)
メタ学習用 (花画像199枚分)	class R (124,375枚) class L (124,375枚) 	class R (648枚) class L (655枚)
メタテスト用 (花画像40枚分)	class R (25,000枚) class L (25,000枚) 	class R (136枚) class L (136枚)

(a) 合成パッチを学習に用いる場合。

	サポートセット (自然パッチ)	クエリセット (自然パッチ)
メタ学習用 (花画像199枚分)	class R (449枚) class L (456枚) 	class R (199枚) class L (199枚)
メタテスト用 (花画像40枚分)	class R (96枚) class L (96枚) 	class R (40枚) class L (40枚)

(b) 合成パッチを学習に用いない場合。

図6: 本実験で用いるサポートセットとクエリセット

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四つの場合で精度の比較を行った。また、合成パッチの数や、タスクのサポートセットに用いる合成パッチの数を変化させ、識別率に与える影響について検証を行った。

4.1 実験データ

学習及び評価に用いる隣接花卉の重なり部分のパッチはイチリンソウとその近縁種の花が写った画像266枚をもとに作成した。このデータは、2016年から2020年に滋賀県、京都府、兵庫県、岡山県、広島県にて撮影されたものであり、撮影には、Olympus社製デジタルカメラTG-5と、Nikon社製一眼レフカメラD5200を用いた。

3.1で説明した通り、初めに、GrubCutを用いて画像から前景部分を抜き出した。抜き出しに成功した239枚の花画像に対して、Harrisのコーナー検出を用いて重なり部分の検出を行った。得られた重なり部分から、自然パッチが1つの花あたり4枚から16枚、合計で1,575枚得られた。また、合成パッチ

を花ごとに作成し、一つの花あたり 1,250 枚、合計で 298,750 枚得た。タスクに含まれるパッチは全て同一の花から選んだ。合成パッチを用いる際には、2 クラスから 10 枚ずつサンプリングを行い、自然パッチを用いる際 2 クラスから 1 枚ずつサンプリングを行った。全てのパッチでサイズは 20×20 pixels である。これらを花卉画像が 5:1 の割合になるように、図 6 に示すようにメタ学習用のデータセットとメタテスト用のデータセットに割り当てた。これらの画像処理の実装には、全て Python 版の OpenCV を利用した。

4.2 実験条件

全ての実験を通してネットワークのモデルは MAML [4] と同じ、3×3 ので 64 フィルタの畳み込み層の後にバッチ正規化、ReLU、2×2 の最大値プーリングが続くようなモジュールを 4 つ連結させたものを用いた。事前学習は行っていない。メタ学習には MAML を用いてモデルのパラメータの初期値を与え、メタテスト用のサポートセットで fine-tuning を行ってからメタテスト用のクエリセットを用いて評価をした。非メタ学習では、メタ学習を使わず、従来の学習と同様に利用するデータセットを学習データとテストデータに分割し、学習データでモデルを学習した後、テストデータでの評価を行った。

合成パッチとメタ学習手法をともに用いる提案手法では、図 6(a) に示す通り、サポートセットに合成パッチ、クエリセットに自然パッチを割り当て学習、評価を行った。合成パッチを用いた非メタ学習手法では、図 6(a) のメタ学習用サポートセットの合成パッチを用いて学習し、メタ学習用クエリセットの自然パッチで fine-tuning をした後にメタテスト用クエリセットで評価を行った。合成パッチを用いないメタ学習では、図 6(b) の通り、サポートセット、クエリセットの両方に自然パッチを割り当て学習、評価を行った。合成パッチを用いない非メタ学習では、自然パッチのみを用いて学習、評価を行った。この場合、図 6(b) のメタ学習用のクエリセットを学習画像、メタテスト用のクエリセットをテスト画像として用いた。MAML を用いた場合、パラメータのメタ更新を行う際に用いるタスクの数は 20 とした。

4.3 実験結果

提案手法とその他 3 つの手法で実験を行った結果を表 1 に示す。提案手法は自然パッチのみを用いた場合よりも 9.4%、非メタ学習手法を用いた場合よりも 11.5% 高い識別精度を達成している。これらの結果から、合成画像の導入とメタ学習の両方が重なり順の推定に有効であることが確認できた。一方で、メタ学習も合成パッチも用いなかった場合とメタ学習のみを用いた場合を比較すると 2.4%、合成画像のみを用いた場合と比較すると 0.3% とそれほど精度の向上は見られなかった。

続いて、タスクを生成する際のサポートセットをサンプリングを行う合成パッチの集合の数を変化させた場合、また、タスクのサポートセットの合成パッチの数を変化させた時に、精度にどのように影響を与えるか、評価を行った。実験の結果を図 7 に示す。図 7 において、“MAML” は MAML を利用した場合、“Non-MAML” は MAML を用いなかった場合である。“Natural patches” と “Synthetic pathces” は、それぞれ学習に

表 1: 隣接花卉の重なり順の推定精度 (%). 数値は 15 回の実験で得られた値の平均と標準偏差を表す。

	非メタ学習	メタ学習
合成パッチ無	84.1 ± 2.1	86.5 ± 3.4
合成パッチ有	84.4 ± 1.9	95.9 ± 1.6

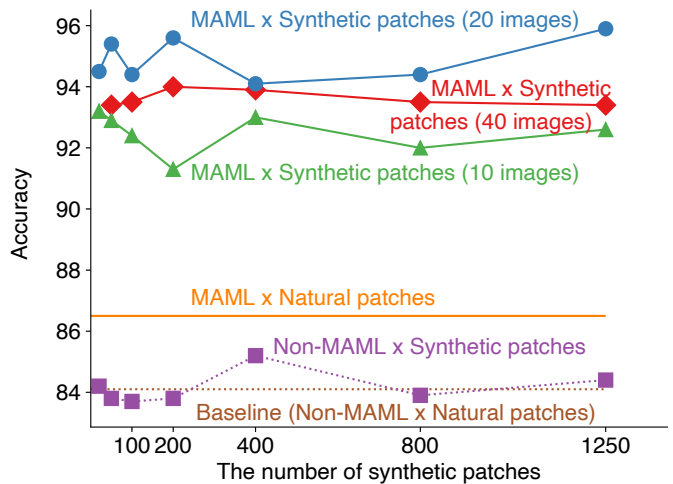


図 7: 隣接花卉の重なり順の推定精度 (%). 横軸は各画像における生成したパッチの数を示す。

自然パッチを用いた場合と合成パッチを用いた場合を表す。“Baseline” は 3.1 で示したベースライン手法に相当する。括弧内に画像の枚数が書いているものは、サポートセット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十分な数の学習データを与えることで、自然パッチに対する識別精度の向上が見られることがわかった。一方で、サンプリングを行う合成パッチの集合の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。識別精度にばらつきが生じたが、これはランダム性によるところが大きいと考えられる。

5. むすび

本稿では、多種少量データを対象とした識別問題（隣接花卉の重なり順推定）に対して、2 つの方策を導入することで、精度向上を図った。一つ目は、メタ学習手法の一つである MAML を導入である。1 つの花に含まれる花卉の見た目は均質であると考えられるので、花の画像毎に 1 つのタスクを構成し、複数のタスクにまたがるメタ学習で、多数の個体に共通の特徴を求めた。認識（メタテスト）時には、こうして得られた学習モデルを認識対象の個体に適応させることで、多種多様なデータへの対応を図った。2 つ目は、画像の合成による学習データの増強である。すなわち、重なっていない花卉の画像から花卉の重なりを模した画像を合成して学習に用いた。各タスクで用いる学習データを大幅に増やすことで、精度向上を狙った。

上記 2 つの方策を導入した提案手法の有効性を、イチリンソウとその近縁種の花画像データセットを用いた実験で評価した。実験の結果、まず、上記の方策のいずれか片方のみを

用いる場合は認識精度がほとんど向上しないことがわかった。一方、両方の方策を用いることにより、両方用いない場合に比べて、認識精度が 11.8% 向上することを確認した。このことから、多種少量データに対する識別問題では、メタ学習と合成パッチの両方の導入が認識精度向上に寄与することが推察される。ただし、本稿ではイチリンソウとその近縁種の花画像データセットに対して検証しただけであるため、今後さらなる検証が必要である。

謝辞 本研究は、JSPS 科研費 JP20H05423 の補助による。

文 献

- [1] A. Nakagawa, M.S. Kitazawa, and K. Fujimoto, "A design principle for floral organ number and arrangement in flowers with bilateral symmetry," *Development*, vol.147, no.3, , 2020.
- [2] J. Deng, W. Dong, R. Socher, L.-J. Li, Kai Li, and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," *Proc. CVPR*, 2009.
- [3] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C.L. Zitnick, "Microsoft COCO: Common objects in context," *Proc. ECCV*, 2014.
- [4] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," *Proc. ICML*, 2017.
- [5] M.E. Nilsback and A. Zisserman, "A visual vocabulary for flower classification," *Proc. CVPR*, 2006.
- [6] M.-E. Nilsback and A. Zisserman, "Automated flower classification over a large number of classes," *Proc. Indian Conf. on Comp. Vis., Graph. & Image Proces.*, pp.722–729, 2008.
- [7] Z. Wang, K. Wang, X. Wang, and S. Pan, "A convolutional neural network ensemble for flower image classification," *Proc. ICCV*, 2020.
- [8] L. Xie, Q. Tian, M. Wang, and B. Zhang, "Spatial pooling of heterogeneous features for image classification," *IEEE Trans. on Image Proces.*, vol.23, no.5, pp.1994–2008, 2014.
- [9] B. Fernando, E. Fromont, and T. Tuytelaars, "Effective use of frequent itemset mining for image classification," *Proc. ECCV*, 2012.
- [10] M. Yang, L. Zhang, X. Feng, and D. Zhang, "Sparse representation based fisher discrimination dictionary learning for image classification," *IJCV*, vol.109, no.3, pp.209–232, 2014.
- [11] W. Hu, R. Hu, N. Xie, H. Ling, and S. Maybank, "Image classification using multiscale information fusion based on saliency driven nonlinear diffusion filtering," *IEEE Trans. on Image Proces.*, vol.23, no.4, pp.1513–1526, 2014.
- [12] B. Fernando, E. Fromont, and T. Tuytelaars, "Mining mid-level features for image classification," *International Journal of Computer Vision*, vol.108, no.3, pp.186–203, 2014.
- [13] B. Fernando, E. Fromont, D. Muselet, and M. Sebban, "Discriminative feature fusion for image classification," *Proc. CVPR*, 2012.
- [14] A.B. Mabrouk, A. Najjar, and E. Zagrouba, "Image flower recognition based on a new method for color feature extraction," *Proc. VISAPP*, 2014.
- [15] Yuning Chai, V. Lempitsky, and A. Zisserman, "BiCoS: A bi-level co-segmentation method for image classification," *Proc. ICCV*, 2011.
- [16] Y. Chai, E. Rahtu, V. Lempitsky, L. Van Gool, and A. Zisserman, "TriCoS: A tri-level class-discriminative co-segmentation method for image classification," *Proc. ECCV*, 2012.
- [17] M.E. Nilsback and A. Zisserman, "Delving deeper into the whorl of flower segmentation," *Image and Vision Computing*, vol.28, no.6, pp.1049–1062, 2010.
- [18] D.S. Guru, Y.H. Sharath Kumar, and S. Manjunath, "Textural features in flower classification," *Mathematical and Computer Modelling*, vol.54, no.3-4, pp.1030–1036, 2011.
- [19] J. Zou and G. Nagy, "Evaluation of model-based interactive flower recognition," *Proc. ICPR*, 2004.
- [20] U.F. Rahim and H. Mineno, "Tomato Flower Detection and Counting in Greenhouses Using Faster Region-Based Convolutional Neural Network," *Journal of Image and Graphics*, vol.8, no.4, pp.107–113, 2020.
- [21] N. Ohi, K. Lassak, R. Watson, J. Strader, and et al., "Design of an autonomous precision pollination robot," *Proc. IROS*, pp.7711–7718, 2018.
- [22] Y. Chen, W.S. Lee, H. Gan, N. Peres, C. Fraisse, Y. Zhang, and Y. He, "Strawberry yield prediction based on a deep neural network using high-resolution aerial orthoimages," *Remote Sensing*, vol.11, no.13, pp.1–21, 2019.
- [23] P. Lin, W.S. Lee, Y.M. Chen, N. Peres, and C. Fraisse, "A deep-level region-based visual representation architecture for detecting strawberry flowers in an outdoor field," *Precision Agriculture*, vol.21, no.2, pp.387–402, 2020.
- [24] D. Wu, S. Lv, M. Jiang, and H. Song, "Using channel pruning-based YOLO v4 deep learning algorithm for the real-time and accurate detection of apple flowers in natural environments," *Computers and Electronics in Agriculture*, vol.178, no. August, p.105742, 2020.
- [25] P.A. Dias, A. Tabb, and H. Medeiros, "Multispecies Fruit Flower Detection Using a Refined Semantic Segmentation Network," *IEEE Robotics and Automation Letters*, vol.3, no.4, pp.3003–3010, 2018.
- [26] B. Millan, A. Aquino, M.P. Diago, and J. Tardaguila, "Image analysis-based modelling for flower number estimation in grapevine," *Journal of the Science of Food and Agriculture*, vol.97, no.3, pp.784–792, 2017.
- [27] S. Liu, X. Li, H. Wu, B. Xin, J. Tang, P.R. Petrie, and M. Whitty, "A robust automated flower estimation system for grape vines," *Biosystems Engineering*, vol.172, pp.110–123, 2018.
- [28] H. Williams, M. Nejati, S. Hussein, N. Penhall, and et al., "Autonomous pollination of individual kiwifruit flowers: Toward a robotic kiwifruit pollinator," *Journal of Field Robotics*, vol.37, no.2, pp.246–262, 2020.
- [29] T. Ijiri, S. Yoshizawa, H. Yokota, and T. Igarashi, "Flower modeling via x-ray computed tomography," *ACM Trans. on Graphics*, vol.33, no.4, pp.1–10, 2014.
- [30] T. Ijiri, M. Yokoo, S. Kawabata, and T. Igarashi, "Surface-based growth simulation for opening flowers," *Proc. Graphics Interface*, pp.227–234, 2008.
- [31] F. Yan, M. Gong, D. Cohen-Or, O. Deussen, and B. Chen, "Flower reconstruction from a single photo," *Computer Graphics Forum*, vol.33, no.2, pp.439–447, 2014.
- [32] C. Zhang, M. Ye, B. Fu, and R. Yang, "Data-driven flower petal modeling with botany priors," *Proc. CVPR*, 2014.
- [33] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," *Proc. ICML*, 2017.
- [34] S. Ravi and H. Larochelle, "Optimization as a model for few-shot learning," *Proc. ICLR*, 2017.
- [35] J. Guo, X. Zhu, C. Zhao, D. Cao, Z. Lei, and S.Z. Li, "Learning meta face recognition in unseen domains," *Proc. CVPR*, 2020.
- [36] H. Zhu, L. Li, J. Wu, W. Dong, and G. Shi, "MetaIQA: Deep meta-learning for no-reference image quality assessment," *Proc. CVPR*, 2020.
- [37] M. Choi, J. Choi, S. Baik, T.H. Kim, and K.M. Lee, "Scene-adaptive video frame interpolation via meta-learning," *Proc. CVPR*, 2020.
- [38] A. Santoro, S. Bartunov, M. Botvinick, D. Wierstra, and T. Lillicrap, "Meta-learning with memory-augmented neural networks," *Proc. ICML*, 2016.
- [39] T. Munkhdalai and H. Yu, "Meta networks," *Proc. ICML*, pp.2554–2563, 2017.
- [40] K. Tran, H. Sato, and M. Kubo, "Memory augmented matching networks for few-shot learnings," *IJMLC*, vol.9, no.6, pp.743–748, 2019.
- [41] G. Koch, R. Zemel, and R. Salakhutdinov, "Siamese neural networks for one-shot image recognition," *Proc. ICML Workshop*, 2015.
- [42] O. Vinyals, C. Blundell, T. Lillicrap, k. kavukcuoglu, and D. Wierstra, "Matching networks for one shot learning," *Proc. NIPS*, 2016.
- [43] R.Z. Jake Snell, Kevin Swersky, "Prototypical networks for few-shot learning," *Proc. NIPS*, 2017.
- [44] F. Sung, Y. Yang, L. Zhang, T. Xiang, P.H. Torr, and T.M. Hospedales, "Learning to compare: Relation network for few-shot learning," *Proc. CVPR*, 2015.
- [45] C. Rother, V. Kolmogorov, and A. Blake, "GrabCut: interactive foreground extraction using iterated graph cuts," *Proc. ACM SIGGRAPH*, New York, New York, USA, 2004.