

メタ学習に基づく Infrared Few-shot Open-set Recognition を考慮した動物分類

1 序論

気候変動や人間活動が生態系に与える影響を把握し、継続的な野生動物と人間との共存を実現するため、生態系モニタリングの重要性が世界的に高まっている。費用対効果の高い生態系モニタリングを行う上で、カメラトラップの使用は極めて重要である。赤外線カメラを用いたカメラトラップは昼夜ともに広範囲を撮影可能であり、膨大な画像や動画データが低コストで収集できるため、深層学習モデルを用いた野生動物の正確な検出と分類に期待が高まっている。

少数の赤外線動物画像に対して、畳み込みニューラルネットワーク (Convolutional Neural Network, CNN) を用いた手法がいくつか提案されている。これらの既存研究では学習用データセットと評価用データセットで同一クラスが用いられており、評価時に分類対象となる全ての動物種がモデルに登録済みであると仮定されている。しかし、深層学習モデルを特定の地域に適用する際、対象地域に生息する全ての動物種に登録しているとは限らない。このような状況において、未登録の動物種は登録済みの動物種に強制的に誤分類され、モデルの性能は著しく低下することが知られている。この問題はオープンセット問題と呼ばれ、これに対処するため、登録クラスの分類を行いつつ未登録クラスを検出するオープンセット認識手法 (Open-Set Recognition, OSR) が提案されている。

さらに近年では、少数データでもオープンセット認識を可能にする Few-Shot Open-Set Recognition (FSOSR) が注目を集めている。しかし、既存手法は可視光画像を対象としており、赤外線画像に対する性能評価は行われていない。また、既存研究では未登録種の検出に取り組んでおり、全ての未登録種は単一のクラスとして扱われる。実用性を考慮すると、新しい種に対するアノテーションコストや追加学習の手間を削減するため、未登録種についても各クラスに分類されている状態が望ましい。

そこで本研究では、夜間の野生動物モニタリングの実現を目的とした、より実用的な問題設定である Infrared Few-shot Open-set Recognition (IFOR) を提案する。本問題設定では少量の赤外線画像データのみを用いて、特定地域に生息するモデルに登録済みの動物種を正確に分類し、かつ、未登録の動物種の検出を目指すものである。

そして、IFOR の実現に向けて、特徴抽出器、転移学習、メタ学習の3つの既存手法の有効性を検証する。また、本研究では IFOR を発展させ、未登録の動物種に対する多クラス分類の精度向上に取り組んだ。特徴空間上の各クラス分布がコンパクトに表現されることにより、多クラス分類が容易になるという仮説のもと、クラスタリングに基づく損失関数を用いてクラス内分散の最小化及びクラス間分散の最大化を図った。

2 提案手法

これまでに、CNN はテクスチャ特徴に基づいて特徴抽出が行われ、ViT は形状特徴に重点を置き特徴抽出されることが明らかにされた。本研究の対象である赤外線画像は色情報を持たないため、動物分類時における形状特徴の重要性はより高まると考えられる。本研究では赤外線画像において、ResNet18 と ViT という2つの異なる特徴抽出アプローチを持つ特徴抽出器の有効性を評価する。

小規模なデータセットでモデルを効率的に学習させる FSL タスクでは、転移学習が重要な解決策として挙げられる。このアプローチは事前学習のタスクから得た知識を活用し、関連する新しいタスクに対して深層学習モデルの汎化性能を向上させる。本研究では IFOR フレームワークにおける転移学習の有用性を確認する。転移学習では事前学習のタスク (ソース) と本番環境でのタスク (ターゲット) の類似性が重要であると知られている。IFOR のターゲットタスクが赤外線画像であることを考慮し、本研究では、事前学習に色情報をもたないフラクタル画像を利用した Formula-Driven Supervised Learning (FDSL) の適用可能性を評価する。さらに、事前学習として一般的に様々な画像認識タスクに適用されている ImageNet データセットを用いた転移学習の有効性も評価する。

近年、FSOSR の分野において革新的なメタ学習フレームワークである PEELER [1] が提案された。この手法はモデルに登録済みのクラス分類とモデルに未登録のクラス識別において高い精度が示されている。PEELER は、高い精度を達成するために、FSL 損失、OSR 損失、分類損失の3つの異なる損失関数を採用している。FSL 損失は、限られた学習データのみを用いて登録クラスの分類を行うために利用される。一方、OSR 損失は、未登録クラスを登録クラスと区別する識別性能の向上に寄与する。最後に、分類損失は、ランダムな画像から適切な特徴を抽出し、モデルの分類能力を最適化するように設計されている。

本研究では、IFOR における未登録クラスに対する多クラス分類の精度向上を図り、クラス内分散の最小化を実現する損失関数として k-means 損失と、クラス間分散の最大化を目的とする損失関数として Between-Class 損失 (BC 損失) を PEELER に導入した。k-means 損失は、Chin ら [2] が異常検知タスクにおいて提案した損失関数であり、k-means クラスタリングによって得られる類似した性質を持つ特徴量から、より識別的な特徴表現の学習を可能にする。この損失関数の最小化により、各クラスタ中心とそのクラスタに属する特徴ベクトルとの距離が最小となることが期待される。

本研究では、似た性質を持つ特徴量をクラスタリングすることにより、特徴空間上における各クラスのク

表 1 赤外線画像に対する各特徴抽出器と転移学習の組み合わせによる実験結果

学習方法	メタ学習 (PEELER)					
特徴抽出器	ResNet18			ViT		
転移学習	ImageNet	FDSL	なし	ImageNet	FDSL	なし
Accuracy (%)	45.8	38.8	38.6	51.0	36.5	36.2
AUROC (%)	58.4	54.3	56.3	61.0	54.4	54.5

表 2 ImageNet 転移学習を用いた ViT による各学習方法の赤外線画像に対する実験結果

特徴抽出器	ViT			
転移学習	ImageNet			
学習方法	メタ学習 (PEELER)	従来法 (ミニバッチ学習)	なし	
Accuracy (%)	51.0	39.9	39.6	
AUROC (%)	61.0	53.3	55.0	

ラス内分散を最小化することを目指し、k-means 損失の有効性を検証する。

一方、BC 損失は、クラスタ中心間の距離を最大化することによって、クラス間分散の最大化を目指す損失関数である。

3 実験

本研究では、ドメインシフトを考慮したモデルの性能評価のため、学習用に WCS Camera Traps データセット、評価用に Caltech Camera Traps (CCT) データセットと異なるデータセットを用いて実験を行った。ドメインシフトとは、学習データと評価データが異なる地域で収集された場合にモデルの性能が低下する現象を指す。また、IFOR では特定の地域において野生動物の画像を大量に収集することが困難であるという現実的な制約を想定しており、5-Way, 1-Shot 問題という極めて厳しい条件下で評価を行った。評価では、分類精度と AUROC を指標として用いた。

本実験では、ResNet18 と ViT の 2 つの特徴抽出器に対して、ImageNet による転移学習、FDSL による転移学習、転移学習なしの 3 つの条件で評価を行った。IFOR における実験結果を表 1 に示す。表から、ViT は ResNet18 に比べて赤外線画像の分類精度が優れていることが分かる。さらに、ResNet18 と ViT は、ImageNet による転移学習を利用した場合に最も高い精度を達成した。このような傾向が見られた背景として、フラクタル画像のドメインが自然画像とは異なるため、赤外線動物画像では効果が現れなかった可能性がある。

次に、IFOR フレームワークにおけるメタ学習の有効性を検証した実験結果を表 2 に示す。表 2 から、メタ学習手法である PEELER は従来の学習方法や学習なしのシナリオと比較して、精度と AUROC の両方を有意に改善したことが分かる。ミニバッチ学習は学習なしのシナリオと比較して、分類精度はわずかに向上したが AUROC は低下した。これは、未学習クラスに対する識別を考慮していない 28 クラスの赤外線動物画像の分類に基づく学習が、AUROC の減少につながったことを示唆している。したがって、IFOR におけるメタ学習の有効性が示された。

最後に、表 3 に未登録の動物種の分類結果を示す。本実験では、距離学習に基づく FSL 損失と OSR 損失

表 3 k-means 損失と BC 損失のアブレーション結果

学習方法	KM	BC	Acc. (%)	AUROC
PEELER			38.1	0.497
	✓		38.1	0.497
		✓	38.1	0.497
	✓	✓	38.7	0.496
PEELER (w/o 分類損失)	✓		39.2	0.496
		✓	39.3	0.496
	✓	✓	39.2	0.496

に加え、分類損失を組み合わせた既存の PEELER フレームワークをベースラインとし、k-means 損失、BC 損失の有効性を検証した。結果として、k-means 損失と BC 損失の両方をベースラインの PEELER に組み合わせることによって、ベースラインより高い性能となることが確認された。次に、分類損失を用いない場合の PEELER に k-means 損失、BC 損失を組み合わせた結果を確認すると、いずれの組み合わせにおいてもベースラインの精度より高い精度が得られた。

4 結論

本研究では、IFOR フレームワークにおけるメタ学習の有効性を評価すると同時に、赤外線画像に適した特徴抽出器と転移学習手法を検討した。さらに、IFOR フレームワークを拡張し、未登録クラスに対する多クラス分類の精度向上に取り組んだ。結果として、特徴抽出器に ViT、転移学習に ImageNet、そしてメタ学習が IFOR において有効だと示された。また、分類損失の有無に関わらず、クラスタリングに基づく損失の有効性が確認された。

参考文献

- [1] B. Liu, H. Kang, H. Li, G. Hua, and N. Vasconcelos, “Few-Shot Open-Set Recognition Using Meta-Learning,” Proc. of CVPR, pp. 8795-8804, 2020.
- [2] C. C. Tsai, T. H. Wu, and S. H. Lai, “Multi-Scale Patch-Based Representation Learning for Image Anomaly Detection and Segmentation,” Proc. of WACV, pp. 3992-4000, 2022.