

深層学習を用いた動物分類に関する既存研究

2.1 赤外線画像に対する既存研究

このような課題を解決するため、少数の赤外線画像を用いた動物分類の研究が行われている [1]。しかし、既存研究では、評価時に分類対象となる動物種は全て学習済みであると仮定しているが、実運用において、深層学習モデルを特定の地域に適用する際、モデルが対象地域に生息する全ての動物種を学習しているとは限らない。このような状況において、未学習の動物種は学習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 (Open-Set Problem) と呼ばれ、この問題に対処するため、学習済みクラスのカテゴリを行いつつ、未学習クラスを検出するオープンセット認識 (Open-Set Recognition, OSR) 手法が提案されている [2, 3]。

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) [4] が注目を集めている。代表的な FSOSR 手法として、少数データ学習 (Few-Shot Learning, FSL) 分野で有効な手法とされているメタ学習を OSR に拡張することにより FSL と OSR を同時に実現した PEELER [4] や、変換の一貫性に基づき未学習クラスを検出することによって、擬似的な未学習クラスサンプルを必要としない SnaTCHer [5] が挙げられる。

しかし、これらの手法は可視光画像を対象としており、赤外線画像に対して性能評価がなされていない。また、FSOSR では未学習クラスを単一種として扱っているが、未学習クラスのアノテーションや追加学習を考慮すると、実用的には未学習クラスも複数種に分類できることが望ましい。

本論文では、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。IFOR 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学習済みの動物種を正確に分類し、かつ、未学習の動物種の検出を可能にすることを目指す。

加えて、IFOR ではドメインシフトに対する頑健性の評価も行う。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生じる課題であり、背景や撮影環境の違い、同じ動物種の地域差による外見の違いによってモデルの性能が低下する現象を指す。ドメインシフトを考慮することにより、地理的条件に依存せず、様々な場所に適用可能な汎用性の高いシステムの実現が期待される。

IFOR の実現に向けて、赤外線画像に有効な既存手法の特定に加え、既存の FSOSR 手法の 1 つであるメタ学習フレームワークが IFOR に対して効果的であるか検証を行う。まず、赤外線画像に有効な特徴抽出器を特定するため、テクスチャ特徴に焦点を当てている CNN や、形状特徴 [6] を重視することで知られている Vision Transformer (ViT) [7] などの代表的な特徴抽出器の有効性を赤外線画像に対して評価する。次に、IFOR フレームワーク内の FSL タスクに有効なアプローチの 1 つである転移学習について検証する。転移学習では、事前学習のタスクと本番環境でのタスクの類似度が重要だと考えられている。そこで、一般的な ImageNet データセットを用いた事前学習と並行して、事前学習に色情報を持たないフラクタル画像を用いる Formula-Driven Supervised Learning (FDSL) [8] の有効性を探る。最後に、赤外線画像を分類する際、小規模データセットから汎用的な特徴抽出を行うための学習戦略であるメタ学習の IFOR における有効性を、ドメインシフトの条件下で評価する。特に、IFOR においては学習済みクラスの正確な画像分類と未学習クラスの検出が不可欠であるため、メタ学習による有効性を従来の学習方法であるミニバッチ学習と比較する。

さらに、IFOR を発展させ、未学習クラスに対する多クラス分類の精度向上にも取り組む。特徴空間上で各クラス分布がコンパクトに表現されることによって多クラス分類が容易になると仮定し、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・クラス間分散の最大化を図る。クラス内分散の最小化では、異常検知タスクで用いられている k-means 損失 [9] を導入する。クラス間分散の最大化では、k-means クラスタリングによって得られる各クラスタ中心を利用した損失関数、Between-Class 損失 (BC 損失) を提案し、性能評価を行う。

以下、第 2 章では深層学習を用いた動物分類に関する既存研究について述べる。第 3 章では夜間の野生動物モニタリングの実現に向けてより実用的な問題設定を提案し、様々な手法の有用性について述べる。第 4 章では評価実験を行い、その結果及び考察を多面的な方向性から述べる。最後に第 5 章では結

論を述べる.

2.1.1 Few-Shot Open-Set Recognition に関する既存研究

参考文献

- [1] Koki Kishi, Masako Kishimoto, Sulfayanti Situju, Hironori Takimoto, and Akihiro Kanagawa, “Few-shot learning for CNN-based animal classification in camera traps using an infrared camera,” In *Proceedings of the 10th IIAE International Conference on Intelligent Systems and Image Processing (ICISIP)*, pp. 11–17, 2023.
- [2] Xin Sun, Zhenning Yang, Chi Zhang, Keck-Voon Ling, and Guohao Peng, “Conditional gaussian distribution learning for open set recognition,” In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 13477–13486, 2020.
- [3] Sagar Vaze, Kai Han, Andrea Vedaldi, and Andrew Zisserman, “Open-set recognition: A good closed-set classifier is all you need,” In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2022.
- [4] Bo Liu, Hao Kang, Haoxiang Li, Gang Hua, and Nuno Vasconcelos, “Few-shot open-set recognition using meta-learning,” In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8795–8804, 2020.
- [5] Minki Jeong, Seokeon Choi, and Changick Kim, “Few-shot open-set recognition by transformation consistency,” In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 12561–12570, 2021.
- [6] Shikhar Tuli, Ishita Dasgupta, Erin Grant, and Thomas L. Griffiths, “Are convolutional neural networks or transformers more like human vision?,” *arXiv preprint arXiv:2105.07197*, 2021.
- [7] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby, “An image is worth 16x16 words: Transformers for image

- recognition at scale,” In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2021.
- [8] Hirokatsu Kataoka, Kazushige Okayasu, Asato Matsumoto, Eisuke Yamagata, Ryosuke Yamada, Nakamasa Inoue, Akio Nakamura, and Yutaka Satoh, “Pre-training without natural images,” *International Journal of Computer Vision (IJCV)*, 2022.
- [9] Chin-Chia Tsai, Tsung-Hsuan Wu, and Shang-Hong Lai, “Multi-scale patch-based representation learning for image anomaly detection and segmentation,” In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 3992–4000, 2022.