

テーマ番号	2EP07		
プロジェクト テーマ	和文	深層学習に基づくカメラトラップ画像からの野生動物認識システムに関する研究	
	英文	A Study on Deep Learning Based Wildlife Camera Trap Image Recognition System	
プロジェクト メンバー	4EP1-63 大淵 香斐 (OHBUCHI Kai)	4EP2-27 高 慎之助 (TAKA Shinnosuke)	指導教員 長田 茂美 教授
	4EP3-46 加藤 響 (KATOU Hibiki)		

Abstract It is important to understand the habitat information of wild animals to better understand and properly manage ecosystems. One method of doing this is through the use of camera trap images. However, as the number of images increases, it becomes difficult for researchers to examine a vast number of images. In this study, we proposed a new model and learning method to improve the accuracy of classifying, identifying species and estimating the number of wild animals. As a result, the accuracy was improved and it was confirmed that the proposed method can greatly reduce the workload of investigators by using the system with the introduced method.

Keywords wildlife, deep learning, camera trap, computer vision, gridmask, attention.

1. まえがき

生態系の複雑性を理解し、より適切に保護管理を行うためには、生息する野生動物の場所、種、数に関する生態情報を調査し、得られた情報を活用することが重要である^[1]。調査法の1つとして、森林内にカメラトラップを設置し、撮影した画像から生態情報を分析する方法がある。しかし、カメラトラップの活用が進むにつれ、大量に撮影された画像を目視して生態情報を読み取る調査者の労力の大きさが問題となってきている。

この問題を解決するために、当研究室では、深層学習を用いて、カメラトラップ画像から、野生動物の在/不在認識、頭数推定、領域検出、種認識を行い、野生動物の生態情報を自動で認識、記録する DeepRecoWild (DReW) システムを開発してきた^[2]。本研究では、さらなる調査者の労力軽減を図るために、深層学習モデルとデータ拡張手法の改良により、DReW システムの精度向上を目指す。さらに、これらの手法を導入した DReW システムの評価実験を行い、提案手法の有効性を示す。

2. DReW システムの概要

図1に、DReW システムの概要を示す。本システムは、入力にカメラトラップ画像を受け取り、在/不在認識を行う在/不在認識モデル、在と認識された画像を入力として受け取る頭数推定モデル、および種認識モデルの3つのモデルで構成される。システムの出力は、在と認識された画像1枚に対する野生動物の領域、種、頭数を撮影時間/場所と紐づけた生態情報である。認識対象の種は、頭胴長が約50cm~1mの中型6種、1m以上の大型4種とした。

本研究では、さらなる調査者の労力軽減を図るために、DReW システムの認識/推定精度向上のためのアプローチを検討した。

3. 認識/推定精度向上に向けたアプローチ

3.1 在/不在認識のための提案手法

従来システムの在/不在認識には、体の一部が見切れた野生動物が写る画像の認識精度が低いという課題があった。そこで、学習過程において GridMask^[3]を適用し、精度向上を図った。当手法は、特定の画像に対して、ランダムな縦横一定の間隔で矩形領域内の画素値を0にするデータ拡張手法である。これにより、野生動物の体の一部の画素を欠損させることで、部分的な見え方をした野生動物の特徴からも正しく認識が行えるようになることが期待できる。また、GridMaskを導入した在/不在認識には、背景領域を野生動物と認識する割合が高いという課題があった。そこで、明示的に野生動物が写る画像領域に注目して学習する Attention^[4]機構をモデルに導入し、背景と野生動物を区別する能力を向上させることで、精度向上を図った。

3.2 種認識のための提案手法

従来の種認識モデルは、NAS (Neural architecture search) で最適化されたモデルを、Compounded Coefficientに基づいてスケールアップした EfficientNet-B6^[5]を採用していた。本研究では、EfficientNet-B6をスケールアップした EfficientNet-B7を種認識モデルとして採用し、より複雑な入力画像の特徴を学習させることで精度向上を図った。また、EfficientNet-B7の導入後に、体の一部が見切れた野生動物が写る画像の認識精度が低いという課題に対応し、学習過程で GridMask を適用することで精度向上を図った。

3.3 頭数推定のための提案手法

従来システムの頭数推定には、体の一部が見切れた野生動物が写る画像の認識精度が低いという課題があった。そこで、頭数推定モデルの学習過程において GridMask を適用し、精度向上を図った。

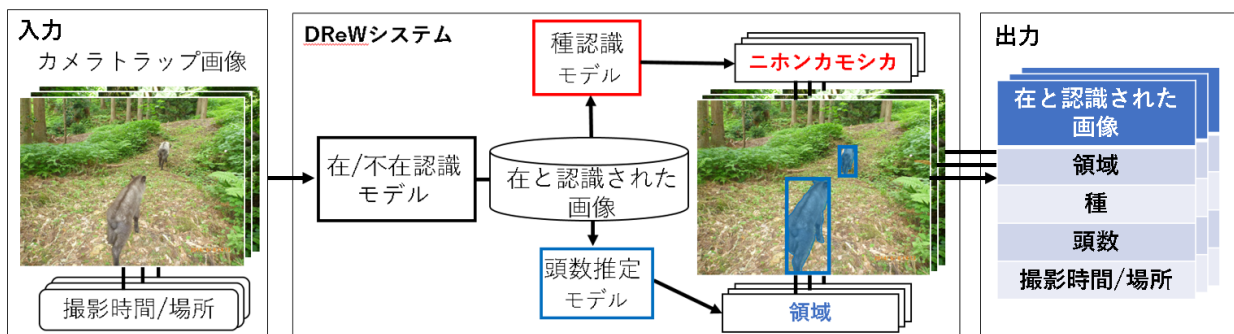


図1 DReW システムの概要
Fig. 1 Overview of the DReW system.

4. 評価実験

精度向上に対する提案手法の有効性を検証するため、石川県農林総合研究センター林業試験場で取得されたカメラトラップ画像を使用し、予め在/不在認識と種認識では ImageNet、頭数推定では MS COCO で学習されたモデルを用いて各タスクを学習するファインチューニングの形式を採った上で、提案手法の有無での精度と、体の一部が見切れた野生動物が写る誤認識画像の枚数を比較した。

4.1 在/不在認識の評価実験

在/不在認識の検証実験には、10,834 枚の学習データと、1,357 枚の検証データ、2,976 枚の評価データを用いた。表 1 に在/不在認識精度を示す。また、GridMask を適用することで、体の一部が見切れた野生動物が写る誤認識画像は 22 枚から 13 枚に削減された。この結果から、提案手法の有効性が確認できた。他の誤認識画像は、「在/不在の判定に迷う」、「野生動物が小さく写っている」ものが多かった。

この誤認識画像は、学習データに少なく、今後は、そのような画像の収集やデータ拡張など、学習データの拡充を行うことで、さらなる精度向上が期待できる。

表1 在/不在認識の評価実験結果

Table 1 Experimental results of 'presence/absence' recognizer.

GridMask	Attention	再現率[%]	適合率[%]
未適用	未適用	95.7	86.3
適用	未適用	96.3	87.1
適用	適用	97.8	97.6

4.2 種認識の評価実験

種認識の検証実験には、32,511 枚の学習データと、1,355 枚の検証データ、1,194 枚の評価データを用いた。GridMask 未適用/適用の場合、モデルを EfficientNet-B6/EfficientNet-B7 を使用した場合で、各々の精度を比較した。表 2 に各々の正解率を示す。また、EfficientNet-B7 に GridMask を導入することで、体の一部が見切れた野生動物が写る誤認識画像は 20 枚から 15 枚に削減されていた。この結果から、提案手法の有効性が確認できた。他の誤認識画像は、「背景と野生動物の色が酷似している」、「野生動物が小さく写っている」ものが多かった。

この誤認識画像は、学習データに少なく、今後は、学習データにこのような画像の拡充を行うことや、野生動物のエッジ検出を行い、野生動物の輪郭を強調した画像を学習させることで、さらなる精度向上が期待できる。

表2 種認識の評価実験結果

Table 2 Experimental result of species recognizer.

モデル	GridMask	正解率[%]
EfficientNet-B6	未適用	91.4
EfficientNet-B7	未適用	91.8
EfficientNet-B7	適用	92.2

4.3 頭数推定の評価実験

頭数推定の検証実験には、2,026 枚の学習データ、7,725 枚の評価データを用いた。表 2 に頭数別の頭数推定の正解率を示す。また、提案手法を導入することで、体の一部が見切れた野生動物が写る誤推定画像は 210 枚から 176 枚に削減され、野生動物が小さく写る誤推定画像は 60 枚から 75 枚に増加していた。この結果から、頭数全体の推定精度の向上、体の一部が見切れた野生動物が写る画像の精度向上に対する提案手法の有効性が確認できたが、複数頭の推定精度の低下、野生動物が小さく写っている画像の推定精度が低下したことが新たな課題として確認できた。これは、小さな野生動物の領域を GridMask が欠損させた

状態で学習を進めたためだと考える。特に、画像に複数頭が写る場合、野生動物はカメラトラップから遠い位置で小さく写ることが多く、その影響が大きく受けたと考える。

今後は、野生動物の領域を保持しながら画像を欠損させることが出来る KeepAugment を適用し、体の一部のみを欠損させることで、複数頭の精度向上が期待できる。

表3 頭数推定の評価実験結果

Table 3 Experimental results of the evaluation of head count.

単位：正解率[%]

GridMask	1頭	2頭	3頭	4頭以上	全体
未適用	88.7	64.5	49.4	41.6	84.8
適用	91.3	60.0	46.9	29.2	86.4

5. DReW システムによる調査者の労力削減の検証

調査者が提案手法を導入した DReW システムを使用することで、システムを使わない場合、従来システムを使用する場合と比べてどの程度労力が削減できたかを検証するための実験を行った。調査者はカメラトラップ画像を目視することで、野生動物の正しい種、領域を判別できるとして、表 4 で示した 3 つの場合で、4.1 で使用した評価データから正しい生態情報を獲得するまでにかかる調査時間を推定し、比較した。調査時間は、種ラベルの書き換え(3 秒)/領域の移動(6 秒)、削除(1 秒)、追加(1 秒)、野生動物を探す時間(1 頭当たり 7 秒)を単位として推定した。

表 4 にシステムの出力に対する推定調査時間を示す。この結果から、提案手法が導入されたシステムを使用することで、従来システムと比較して大幅な調査者の労力を軽減出来なかったが、システムを使用しない場合と比較して大幅に調査者の労力を軽減できたことが確認できた。

表4 推定調査時間

Table 4 Estimated modification time.

調査方法	推定調査時間[分]
全て目視	214.8
従来システムを使用	49.1
提案システムを使用	42.5

6. あとがき

本研究では、カメラトラップ画像から野生動物の生体情報を自動で認識/推定するシステムを改良し、評価実験により、その有効性を確認し、提案手法を導入したシステムを使用することで、調査者の労力は大幅に軽減されたことを確認した。今後も、認識/推定精度向上、特に複数頭への推定精度の向上により労力軽減を目指していきたい。

参考文献

- [1] Harris G, Thompson R, Childs JL, Sanderson JG: "Automatic Storage and Analysis of Camera Trap Data", The Bulletin of the Ecological Society of America 91(3), pp.352-360,(2010).
- [2] 上原祐介, 大谷直也, 西本圭佑, 川崎邦将, 松井康浩, 長田茂美ほか: "深層学習に基づく野生動物認識システム", 中部森林研究, No. 68, pp. 25-30, (2020).
- [3] Pengguang Chen, Shu Liu, Hengshuang Zhao, Jiaya Jia: "GridMask Data Augmentation", arXiv:2001.04086,(2020).
- [4] Sanghyun Woo, Jongchan Park, Joon-Young Lee, In So Kweon: "CBAM: Convolutional Block Attention Module", arXiv:1807.06521, (2018).
- [5] Mingxing Tan, Quoc V. Le: "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks", arXiv:1905.11946, (2020).