

テーマ番号	1EP15		
プロジェクト テーマ	和文	ニューラルネットワーク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	指導教員 長田 茂美 教授
	英文	A studies on wildlife recognition system from camera trap images based on neural network.	
プロジェクト メンバー	4EP4-78 福田 一之 (FUKUDA Kazuyuki) 4EP3-43 中村 龍 (NAKAMURA Ryu)	4EP4-51 平田 大貴 (HIRATA Daiki)	

**Abstract** Damage to agriculture and forestry caused by wildlife has become one of the most serious social problems in Japan. In order to solve this problem, camera trap images have been used to understand wildlife habitat information. However, with the increase in the number of images, it has become difficult for researchers to examine a huge number of images. In this study, we propose a new DNN model for wildlife type recognition, and aim to improve the recognition accuracy by augmenting the training dataset for type recognition and head count estimation using GAN. Evaluation experiments of the proposed method are conducted to show the usefulness of the proposed method.

**Keywords** wildlife, computer vision, deep neural network, conditional generative adversarial network.

### 1. まえがき

野生鳥獣がもたらす農林業の被害は、我が国の深刻な社会問題の一つとなっている<sup>[1]</sup>。こうした野生動物の被害対策を講じるためには、どのような種がいつ、どこに、何頭現れるのかといった生息情報の調査が必要である。一般に、この生息情報の調査には、カメラトラップ（自動撮影カメラ）画像が活用されているが、大量に撮影されたカメラトラップ画像から目視で生息情報を読み取る調査者の労力が大きな問題となっている。

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、種認識、頭数推定を行い、野生動物の生息情報を自動認識する DeepRecoWild (DRew) システムを提案してきた<sup>[2]</sup>。本研究では、DRew システムのさらなる認識/推定精度の向上を目指して、特に、種認識および頭数推定における課題解決に取り組み、種認識のための Deep Neural Network(DNN)の改良、種認識および頭数推定のための GAN を用いたデータ拡張手法を提案する。さらに、提案手法を導入した DRew システムの評価実験を実施し、それらの手法の有用性を示す。

### 2. DRew システムの概要

図 1 に DRew システムの概要を示す。DRew システムはカメラトラップ画像を入力として受け取る 3 つの DNN から成る。まず、入力されたカメラトラップ画像に対して、野生動物の在/不在認識を行う。次に、在と認識された画像に対してのみ、種認識、頭数推定、および領域検出の処理を行い、最後に、これらの認識/推定結果に基づいて、野生動物の生息情報（頭数、領域、種、在/不在、撮影時間・場所）を出力する。

### 3. 精度向上に向けたアプローチ

#### 3.1 種認識のための DNN の改良

従来の DRew システムでは、種認識のための DNN として ResNet50 を用いていたが、外見が酷似するニホンアナグ

マやニホンタヌキなどの種間相互の誤認識が発生しやすいことや、蜘蛛の巣や濃霧などの強いノイズを含む画像に対する認識精度が低いといった課題が残されていた。

本研究では、これらの課題解決を目指して、種認識のための DNN として Vision Transformer(ViT)<sup>[3]</sup>と EfficientNet<sup>[4]</sup>の 2 つのモデルを検証し、改良を図った。ViT は、畳み込みではなく Attention を用いることで、局所的な特徴のみならず全体を俯瞰する大域的な特徴をも考慮することによって、種認識にとって本質的な特徴の抽出が可能となり、種認識精度の向上が期待できる。一方、EfficientNet は、NAS (Neural Architecture Search) により最適化された構造の EfficientNet-B0 を Compound Coefficient に基づいてスケールアップしたモデルであり、種認識に有効な本質的な特徴の抽出、認識精度の向上が期待できる。

#### 3.2 GAN によるデータ拡張

従来の種認識および頭数推定のクラスごとの学習データには偏りがあり、DNN が本質的な特徴を学習により獲得するのは難しく、DRew システムの種認識および頭数推定精度の低下を招く要因となっていた。本研究では、この課題解決に対しては、GAN によるデータ拡張を検討し、種認識および頭数推定の学習データの拡充を図った。

##### 3.2.1 種認識に対するデータ拡張

従来の種認識に対するデータ拡張には、左右上下の平行移動、左右反転、輝度変化を基本とするデータ拡張手法が用いられていたが、本研究では、さらなる精度向上を目指して、StyleGAN2 に Adaptive Discriminator Augmentation (ADA) というデータ拡張を適用した GAN<sup>[5]</sup>によるデータ拡張手法を提案する。StyleGAN2 は細部の表現能力に優れた GAN であり、ADA を適用することで、学習データが少数の場合でも、過学習を抑制し、優れた品質の画像生成が可能となる。この GAN により、種認識の本来の学習データと遜色のない画像を生成し、学習データを拡充することで、学習データが少数のクラスの精度向上が期待できる。

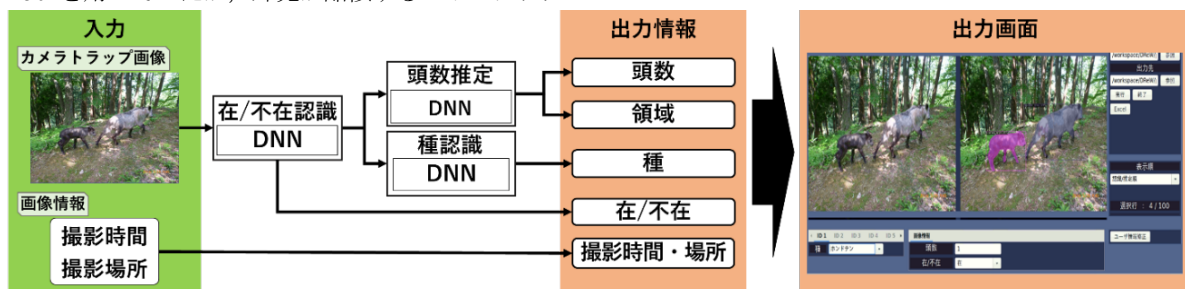


図 1 DRew システムの概要  
Fig. 1 Overview of the DRew system.

### 3.2.2 頭数推定に対するデータ拡張

頭数推定に対しては、図2に示すように、ラベル画像から本物に酷似した画像を生成できる GAN の一種である GauGAN<sup>[6]</sup>を用いたデータ拡張手法を提案する。また、GauGAN は、Spatially-Adaptive (DE)normalization(SPADE)という正規化手法を用いることで、ラベル画像の意味情報が失われることを防げるため、ラベル画像の意味情報の生成画像への反映度合いが高く、本物の画像と遜色のない画像生成が可能であり、カメラトラップではあまり写らない複数頭の動物が存在する画像を生成し、頭数推定の学習に用いることで、複数頭の認識精度の向上が期待できる。



ラベル画像 生成画像  
図2 GauGANによる画像生成

Fig.2 Image Generation with GauGAN.

## 4. 評価実験

### 4.1 種認識の評価実験

種認識に対する提案手法の有用性を確認するために、10,834 枚の学習データ、1,355 枚の検証データ、1,194 枚のテストデータに基づいて、従来手法、提案手法の2種類のデータ拡張手法で拡充した2種類の学習データセット(従来手法版、提案手法版)を構築し、3種類のDNN(ResNet50, ViT-B/16, EfficientNet-B6)の認識精度の評価実験を実施した。提案手法のGANによるデータ拡張では、すべての11種のクラスの学習データが3,000枚になるように画像を生成することで、従来手法版とほぼ同数の元の学習データの約3倍の学習データセットを構築した。

表1に各々のDNNの認識精度(正解率)、表2に従来手法で誤認識が多かった2クラスのF値、表3に強いノイズが載った画像に対する正解率を示す。EfficientNet-B6がすべてにおいて最も良い結果を示しており、Compound Coefficientに基づくモデル構造の最適化が種認識においても有効に機能したものと考える。一方、ViTはEfficientNet-B6よりも劣った結果となったが、種認識において、帰納バイアスが弱いという特徴をもつViTを用いるには、学習データが少なすぎたことが原因であろう。

2種類のデータ拡張手法で拡充した学習データセットについては、データ拡張自体の有用性は確認できたが、従来手法版の結果の方が良く、提案手法のGANによるデータ拡張は有効ではなかったことがわかる。DNNがGANによる生成画像に過度に適合しすぎたことで、本来の学習データに見られる特徴ではなく、生成画像特有の特徴を重視するようになったことが原因として考えられる。

今後は、GauGANのような条件付きGANなどを用いることで、生成画像の特徴を人がコントロールすることにより、本来の学習データがもつ特徴をより適切に近似し、さまざまなバリエーションの特徴をもつ画像を生成することで、さらなる精度の向上が期待できる。

### 4.2 頭数推定の評価実験

頭数推定に対する提案手法の有用性を確認するために、2,026枚の学習データ、7,725枚のテストデータに基づいて、提案手法で2頭以上の動物が写っている頭数クラスの学習データが500枚になるように拡充した学習データセット(提案手法版:13,730枚)を構築し、頭数推定の評価実験を実施した。なお、1頭クラスの学習データは1,730枚と数が多く、データ拡張は適用していない。

表4に、データ拡張未適用/提案手法適用の学習データに対する頭数クラスごとの正解率を示す。提案手法のデータ拡張により、すべての頭数クラスの精度が向上し、その有用性が確認できた。1頭クラスおよび全体の正解率については、2頭以上の頭数クラスの画像特徴の学習が重視され、相対的に1頭クラスの画像特徴の重要度が低下したこと、また、テストデータの約86%を1頭クラスが占めていることが、これらの精度低下の原因である。今後は、頭数以外の画像情報にも着目し、どのような画像特徴が精度に大きく影響するのかを考察し、データ拡張手法の改良、精度向上を図っていく予定である。

表1. 種認識の評価実験結果  
Table 1 Experimental results of type recognition evaluation.

学習データセット	精度:正解率(%)		
	ResNet50	EfficientNet-B6	ViT-B/16
従来手法版	86.3%	92.7% (+6.4%)	77.1% (-9.2%)
提案手法版	83.0% (-3.3%)	90.6% (+4.3%)	73.9% (-12.4%)

表2. ニホンタヌキとニホンアナグマのF値  
Table 2 F1score for Japanese badger and Japanese raccoon dog.

学習データセット	種	精度:F値(%)		
		ResNet50	EfficientNet-B6	ViT-B/16
従来手法版	ニホンタヌキ	85.2%	91.3% (+6.1%)	72.1% (-13.1%)
	ニホンアナグマ	79.0%	90.7% (+11.7%)	65.0% (-14.0%)
提案手法版	ニホンタヌキ	75.8% (-9.4%)	87.2% (+2.0%)	67.5% (-17.7%)
	ニホンアナグマ	71.1% (-7.9%)	78.0% (-1.0%)	58.1% (-20.9%)

表3. ノイズの強い画像に対する正解率  
Table 3 Accuracy for noisy images.

学習データセット	精度:正解率(%)		
	ResNet50	EfficientNet-B6	ViT-B/16
従来手法版	83.1%	89.7% (+6.6%)	73.5% (-9.0%)
提案手法版	81.5% (-1.6%)	89.2% (+6.1%)	69.7% (-13.4%)

表4. 頭数推定の評価実験結果  
Table 4 Experimental results of the evaluation of head count.

頭数	精度:正解率(%)	
	データ拡張未適用	提案手法適用
1頭	89.5%	88.0% (-1.5%)
2頭	62.7%	65.6% (+2.9%)
3頭	52.5%	56.9% (+4.4%)
4頭以上	38.2%	42.7% (+4.5%)
全体	85.3%	84.5% (-0.8%)

## 5. 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種と頭数を認識するシステムの評価・改良を行い、評価実験により、その有用性を確認した。今後は、さらなる認識精度の向上を目指し、野生動物の写り方の違いに影響されない実用的なシステムへと発展させていきたい。

### 参考文献

- [1] 農林水産省農村振興局:鳥獣被害の現状と対策,(2021).
- [2] 上原祐介,大谷直也ほか:“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] Alexey Dosovitskiy, Lucas Beyer et al.:“An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale,”ICLR,(2021).
- [4] Mingxing Tan, Quoc V. Le et al.:“EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,”(ICML 36), Vol. 97 of PMLR, pp. 6105-6114,(2019).
- [5] Tero Karras, Miika Aittala et al.:“Training Generative Adversarial Networks with Limited Data,”NeurIPS,(2020).
- [6] Taesung Park, Ming-Yu Liu et al.:“Semantic Image Synthesis with Spatially-Adaptive Normalization,”(CVPR 2019), pp. 2337-2346,(2019).

### 本プロジェクトに関する業績

- 1) 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).