

AI による画像認識を活用した野生動物の判別

Identification of Wildlife Using Image Analysis by AI

成岡道男*
(NARUOKA Michio)

I. はじめに

現在、スマートフォンやパソコンに送られる現場の画像を見ながら捕獲檻を遠隔操作して害獣を捕らえる「スマート捕獲」が脚光を浴びている。この技術により、高齢化が進む捕獲従事者の見回り負担が軽減し、親個体の選択的な捕獲で害獣の増加を抑制することが可能になった。しかし、多くの害獣は夜間に活動するので、夕食や就寝の時間に現場の画像が送られることが多く、夜間作業が増加した。この状況を改善するには、「人がスマートフォンなどを見ながら檻を操作する」工程を人工知能 (AI: Artificial Intelligence) に代替させることが有効となる。

筆者の所属する農研機構では、Society 5.0 の早期実現を組織目標に掲げてスマート農業や AI などを重点化し、研究開発を推進している^{1,2)}。その一環として、AI の画像認識能力を活用したスマート捕獲技術の開発を進めている。

本報では、スマート捕獲の「人がスマートフォンなどを見ながら檻を操作する」工程の中の「捕獲する野生動物の選別」機能を開発する目的で、AI の画像認識能力を活用した野生動物の判別について検討する。

II. 研究の方法

1. 判別する野生動物

筆者は神奈川県伊勢原市や大磯町で実施されている箱わなを使ったイノシシの有害鳥獣捕獲に参加している。この地域にはイノシシ、シカ、タヌキ、アナグマ、クマの5種類の野生動物が出没する。本報では、この5種類の野生動物を AI によって判別し、その画像認識能力について検討する。

2. 利用する AI

本報では、安田³⁾ が組み立てた AI モデル (interface-201812-3-umaibar-classifier.ipynb) を利用している。この AI モデルは表-1 に示すライブラリーなどから構成されており、①学習済み CNN

表-1 AI モデルを構成するソフトウェアやライブラリーなど

ソフトウェア名など	ソフトウェアなどの機能	
Ubuntu	OS	
Python	コンピューター言語	
Jupyter Notebook	エディター	
CUDA	GPU*1向けソフトウェア	
TensorFlow Keras NumPy Scikit-learn Matplotlib Pandas seaborn	ライブラリー*2	機械学習用 機械学習用 科学技術計算用 機械学習用 グラフ描画用 データ解析用 グラフ描画用

*1 GPU: Graphics Processing Unit

*2 ライブラリー: 汎用性の高い関数などをまとめたファイル

(Convolutional Neural Network: 畳み込みニューラルネットワーク) のファインチューニング (Fine Tuning) と、②データ拡張 (Data Augmentation) の2つの手法を活用することで比較的少ないデータでの画像データの学習が可能になっている。

(1) 学習済み CNN のファインチューニング 学習済み CNN とは、ImageNet⁴⁾ などの大規模画像データベースを使って学習された CNN であり、画像の学習や判別を行うプログラムである。大学などが開発し公開している学習済み CNN もある⁵⁾。ファインチューニングとは、既存の学習済み CNN を目的に応じて追加学習させることで、新しい CNN を再構築する手法である。すでに大量の画像データで学習した CNN を利用することで、少ない画像データでの追加学習が可能になる。安田による AI モデルでは、表-1 に示す Keras に含まれる VGG16⁶⁾ と呼ばれる学習済み CNN を機械学習に利用している。VGG16 は、2014年にオックスフォード大学の研究グループが提案した CNN であり、画像を1,000種類のカテゴリに分類できる。

本報では、AI を「学習・記憶・推理・判断など、人間のもつ知的機能を代行するコンピューターシステム」と定義し、画像の学習と判別を行う VGG16 を AI

*農研機構農村工学研究部門

Society 5.0, 鳥獣害対策, 人工知能, 機械学習, ニューラルネットワーク, 画像認識

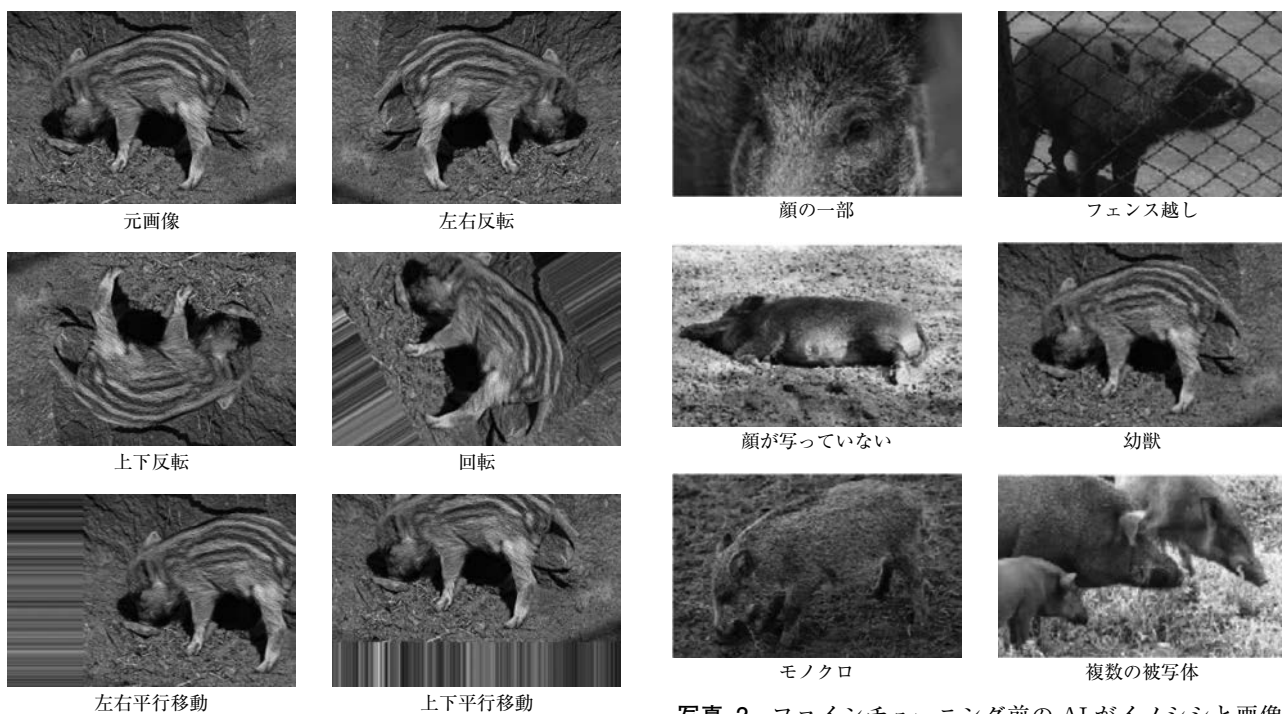


写真-1 画像の幾何学的変換

写真-2 ファインチューニング前の AI がイノシシと画像認識した例

と位置づける。以下、VGG16 を AI と表記し、ファインチューニング前と後の AI の画像認識能力について検討する。

(2) **データ拡張** データ拡張とは、元画像を幾何学的に変換することで、データ量を増やす技術であり、「水増し」とも呼ばれている。この技術を使うことで、準備するデータ量の削減が可能になる。写真-1 の幾何学的変換を活用した場合、左右反転、上下反転、回転、左右平行移動、上下平行移動の 5 つの変換により元画像を含めてデータ量が 6 倍になる。

3. 学習データとテストデータ

ファインチューニングするために使用する学習データおよびファインチューニング後に画像の判別に用いるテストデータは、Google の画像検索を使って収集した。これらは、上述した 5 種類の野生動物の画像である。収集に当たり、Google のフィルタリング機能を使用し、ライセンスが「改変後の再使用が許可された画像」から選んだ。

4. 利用する情報処理機材

ファインチューニングには、高速情報処理が必要になる。このため、本報では、HPC テック製コンピューター (HPC W115gs-DL, Intel Xeon Processor W-2100 family, NVIDIA 製 GPU 2 基) を利用した。

III. 検討結果

1. ファインチューニング前の AI の画像認識能力

ファインチューニングを行う前の AI の画像認識能

力を検討するために、100 枚のイノシシの画像を判別した。その結果、86 枚をイノシシと判定しており、写真-2 に示すように、顔の一部やフェンス越し、顔が写っていないもの、幼獣、モノクロ、複数頭が写っているものなどもイノシシと判別していた。一方、イノシシと判別しなかった 14 枚の画像は、イボイノシシ、ブタ、水牛などに間違えていた。

AI は、ImageNet で学習したデータを元に画像を判別するため、ImageNet に含まれていない画像を判別できない。本報で判定する 5 種類の野生動物のうちイノシシとクマ、アナグマは ImageNet に含まれている。しかし、シカとタヌキは含まれていなかった。このため、5 種類の野生動物の判定には、AI をファインチューニングで追加学習させる必要があった。

2. ファインチューニング後の AI の画像認識能力

判別する 5 種類の野生動物に対して、動物ごとに 50 枚の画像を準備し、40 枚を学習データ用とし、残り 10 枚をテストデータ用に利用した。合計 200 枚の学習データで AI をファインチューニングし、50 枚のテストデータを使って野生動物を判定した。

判定結果の評価には、混同行列 (Confusion Matrix) を使用した。混同行列とは、2 値分類問題で出力された結果をまとめた表であり、実際の分類結果に対する判定の結果を TP (True Positive, 真陽性)、TN (True Negative, 真陰性)、FP (False Positive, 偽陽性)、FN (False Negative, 偽陰性) の 4 つの観点で分類している (表-2)。TP と TN が正解であり、

表-2 混同行列

		判定の結果	
		P (真)	N (偽)
実際の分類結果	P (真)	TP (真陽性)	FN (偽陰性)
	N (偽)	FP (偽陽性)	TN (真陰性)

FP と FN が不正解を示している。ここで、これらの分類を使って、正解率(Accuracy)を計算した。

$$\text{正解率(Accuracy)} = \frac{TP + TN}{TP + TN + FN + FP} \times 100\%$$

..... (1)

表-3 は、5種類の野生動物のテストデータに対するAIによる判別結果を混同行列で示している。判定には各動物のテストデータを10枚ずつ使用した。判別の結果、アナグマ、イノシシ、クマのそれぞれ10枚、シカの9枚、タヌキの8枚がTPで正解となった。一方、FPで不正解の判定が3枚あり、シカをイノシシと、タヌキをアナグマと、タヌキをイノシシと間違えていた。この結果から、ファインチューニングしたAIで5種類の野生動物を判別した正解率は94%となった。AIが間違えなかったアナグマ、イノシシ、クマとAIが間違えたシカ、タヌキの大きな違いは、前者がImageNetに含まれていたのに対し、後者はImageNetに含まれていなかったことである。これは前者がImageNetとファインチューニングの両方で学習したのに対し、後者はファインチューニングでしか学習していないことを示す。このことから、AIが学習したデータ量の差が判別の正解に影響を与えたと考えられる。

3. AIの判定根拠

AIが画像の判別を行う判断根拠を可視化する技術として、Grad-CAM (Gradient-weighted Class Activation Mapping)⁷⁾がある。Grad-CAMは、画像のどの部分がAIの最終的な判別の決め手になったかをヒートマップで表示する技術である。ファインチューニング後のAIの判別にGrad-CAMを使用した結果、以下に示す部分をAIが判断根拠にしていることが分かった。

(1) **アナグマ** アナグマの画像は10枚中8枚が顔にヒートマップが集中し(写真-3 a)、残りの2枚は顔から背中にかけて集中していた。顔から背中にかけてヒートマップが示された画像は、画像中の被写体が小さいものであった。また、顔をアップで写した画像では、顔全体ではなく目の周りや鼻先などにヒートマップが集中している場合もあった。これらの結果から、AIはアナグマの顔を判断根拠にしているようである。

表-3 混同行列によるAIの性能判定

		AIによる判別結果				
		アナグマ	イノシシ	クマ	シカ	タヌキ
テストデータ	アナグマ	10				
	イノシシ		10			
	クマ			10		
	シカ		1		9	
	タヌキ	1	1			8

a. アナグマ



元画像



ヒートマップ

b. イノシシ



元画像



ヒートマップ

c. クマ



元画像



ヒートマップ

d. シカ



元画像



ヒートマップ

e. タヌキ



元画像



ヒートマップ

写真-3 Grad-CAMによるヒートマップへの変換

(2) **イノシシ** イノシシの画像は10枚中6枚が顔と四肢にヒートマップが集中し(写真-3 b)、残りの4枚は顔だけに集中していた。顔だけにヒートマップが集中している画像は、正面や顔のアップ、寝転ん

だ姿であり、四肢が見えにくい画像であった。これらの結果から、AIはイノシシの顔と四肢を判断根拠にしているようである。

(3) **クマ** クマの画像は10枚中6枚が顔から後背にかけてヒートマップが集中し(写真-3 c), 3枚は顔だけに集中し, 1枚は顔と足に集中していた。顔だけまたは顔と足にヒートマップが集中している画像は, 正面からの画像や顔だけの画像であり, 後背が見えにくい構図であった。これらの結果から, AIはクマの顔から後背を判断根拠にしているようである。

(4) **シカ** シカの画像は10枚中5枚が顔にヒートマップが集中し(写真-3 d), 4枚は全身に広がっており, 1枚は顔と四肢に集中していた。全身にヒートマップが広がっていた画像は, 画像の中の被写体が小さいものであった。これらの結果から, AIはシカの顔を判断根拠にしているようである。しかし, 画像の中の被写体が小さいと顔も小さくなるので, 顔以外の部分にもヒートマップを広げる傾向があった。また, AIがシカをイノシシと間違えた画像は, シカの顔を拡大した画像であり, 目の周辺にヒートマップが集中していた。

(5) **タヌキ** タヌキは10枚中10枚が顔にヒートマップが集中しており(写真-3 e), AIはタヌキの顔を判断根拠にしているようである。また, タヌキをアナグマと, タヌキをイノシシと間違えた画像は, それぞれ顔の拡大画像であり, 鼻先や目の下など顔の一部へのヒートマップの集中が顕著であった。シカをイノシシと間違えた画像と同様に, 顔の拡大写真を判別すると顔の細部にヒートマップが集中するため, AIが判別を誤る可能性が生じるようである。

IV. おわりに

本報では, スマート捕獲の「人がスマートフォンなどを見ながら檻を操作する」工程の中の「捕獲する野生動物の選別」機能を開発する目的で, AIの画像認識能力を活用した野生動物の判別について検討した。200枚の画像データを使ってAIをファインチューニングし, 5種類の野生動物の画像を判別した。その結果を混同行列で分類したところ, 正解率が94%となった。そして, それぞれの動物の判別の精度は, 学習するデータ量に影響されることが推測された。Grad-CAMを利用して, AIによる画像の判別の判断根拠をヒートマップで可視化したところ, 野生動物の判別に際して, 顔の部分が重要な判断根拠になってい

ること, 顔以外の判断根拠になる部分が動物ごとに異なる場合があることが分かった。また, 顔の拡大画像を判別した場合, 顔の細部にヒートマップが集中するため, AIが誤る可能性が生じていた。これらのことから, 野生動物の判別のためにAIの学習データを作成するには, 顔が鮮明に写っていること, 顔以外の動物ごとの特徴が分かるように全身が写っていることなどが重要であると推察された。

今後, AIを活用した有害鳥獣の捕獲技術を実用化していくには, AIの判別能力の強化とともに, AIを「研究室」から「捕獲現場」へ持ち出す工夫が重要になる。捕獲現場は山間地も多く, 既存の通信施設が活用できない地域もあることから, IoT技術の活用とともにAIを搭載した小型コンピュータの活用などについても検討していく必要がある。

引用文献

- 1) 白谷栄作: 農業の Society 5.0 に向けた技術開発, 水土の知 87(9), pp.27~30 (2019)
- 2) 土居邦弘: スマート農業を支える基盤研究の現状と将来像, 水土の知 87(9), pp.23~26 (2019)
- 3) 安田国裕: My 画像認識人工知能のうまい育て方, Interface 44 (12), pp.20~37 (2018), <https://github.com/yasudakn/umaibar/> (参照 2019 年 11 月 6 日)
- 4) ImageNet: ImageNet, <http://www.image-net.org/> (参照 2019 年 11 月 6 日)
- 5) Keras Documentation: 利用可能なモデル, Applications, <https://keras.io/ja/applications/> (参照 2019 年 11 月 6 日)
- 6) Simonyan, K. and Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition, arXiv: 1409.1556 (2015), <https://arxiv.org/abs/1409.1556> (参照 2019 年 11 月 6 日)
- 7) Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D. and Batra, D.: Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization, arXiv: 1610.02391 (2017), <https://arxiv.org/abs/1610.02391> (参照 2019 年 11 月 6 日)

[2020.2.5.受理]

略歴

成岡 道男 (正会員・CPD 個人登録者)



1964年 大阪府に生まれる
1995年 鳥取大学大学院博士課程修了
2014年 農研機構農村工学研究所
現在に至る