

Pythonによる転移学習を活用した画像分類モデルの作成 Creating an image classification model using transfer learning in Python

○生野慎太郎, 堀泰史, 内田空美子, 西昇悟, 羽多宏彰
(Shono Shintaro, Hori Yasushi, Uchida Kumiko, Nishi Shogo, Hata Hiroaki)

1. はじめに

野生鳥獣による農作物被害が全国的に発生しており, 農作物の金額的な被害だけでなく, 農家の営農意欲の減少など数字以上に深刻な影響を及ぼしている。Y県S市のS地区でも, 主にニホンジカによる鳥獣被害が発生しており, 早急な鳥獣害対策が求められている。S地区では, シカの農地への侵入経路や出現時刻を把握するために5台の監視カメラ(Fig.1)を農地周辺に設置した。設置した監視カメラは, モーションセンサにより起動し, 動く物体を検知し撮影される仕組みである。そのため, 野生鳥獣以外にも, 人や車だけでなく, 風で揺れた草木にも反応し撮影しており, 1地点の1ヶ月の撮影枚数が10,000枚になることがあり, 結果の整理に膨大な時間を要している。本稿では, データ整理の簡略化を目的として, 画像認識技術を活用し, 動物が撮影された画像と動物以外が撮影された画像を区別する分類モデルの作成を試みた。分類モデルの作成では, 学習済みモデルを活用し転移学習させることにより, 簡便な作成方法を検討した。



Fig.1 監視カメラ
Surveillance camera

2. 転移学習を活用した分類モデルの作成と精度の検証

2.1 転移学習とは

転移学習は, 近年注目されている機械学習の手法の一つであり, 既に学習済みのモデルを転用することで, 新たなモデルを生成する方法である(Fig.2)。転移学習の特徴は, 学習済みのモデルを再利用するため, 必要なデータ数と学習時間を大幅に削減でき, 手軽にモデルを作成できる点にある。しかし, 転移先と転移元との関連性が低い場合には, 精度が低いモデルになってしまう場合がある点にも留意が必要である。

2.2 転移学習に活用したモデル

本稿では, シンプルな構造で, 実装も容易であり, 比較的良好な精度が得られるとされているVGG16¹⁾を採用し, VGG16の転移学習によって画像分類モデルを作成した。VGG16とは, ImageNetの120万枚の画像を1,000カテゴリに分類した16層の畳み込みニューラルネットワーク(CNN)モデルである。ImageNetは, 画像認識の研究で使用するために作成された大規模画像データセットである。

2.3 モデルの作成と精度検証

分類モデルは, 教師データ数の異なる3つのモデル(モデルa, モデルb, モ

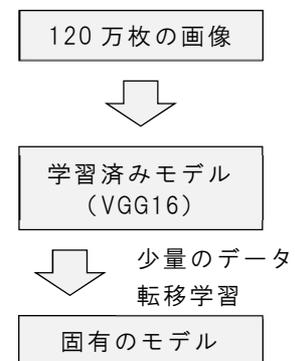


Fig.2 転移学習のイメージ
Illustration of transfer learning

Table1 教師データ数 Number of training data

モデル	モデルa	モデルb	モデルc
教師データ数	200	400	600

デルc)をVGG16の転移学習によって作成した。本稿では、分類モデルの簡便な作成を目的としたため、教師データの前処理等の工程を省略し、与える教師データ数のみを変更した(Table1)。モデルaは、動物が撮影された画像(Fig.3-1)100枚と動物以外が撮影された画像(Fig.3-2)100枚を教師データとし、計200枚からモデルを作成した。モデルbは、それぞれを100枚追加し、計400枚、モデルcはさらにそれぞれ100枚追加し、計600枚から分類モデルを作成した。



Fig.3-1 画像例 A
Image example A

精度の検証は、1,000枚(動物が撮影された画像418枚、動物以外が撮影された画像582枚)の画像を用意し、1,000枚の画像の分類結果をもとに正答率を算定した。モデルの作成と精度の検証には、S地区において、2022年の2月から5月にかけて撮影された画像を使用した。



Fig.3-2 画像例 B
Image example B

3. 作成したモデルの分類精度の結果と考察

3.1 各モデルにおける分類精度

各モデルの正答率は、モデルaで89.5%、モデルbで91.0%、モデルcで92.9%となり、教師データ数が多いほど、正答率が高くなった(Table2)。3つのモデルとも正答率は90%程度となっており、比較的良好な精度で画像を分類できた。

3.2 結果の考察

本稿では、転移学習を活用することにより、少量の学習データでも比較的良好な精度で分類モデルが作成できることが判明した。一方で不正答画像の特徴を確認すると、夜間に何も撮影されていない画像を動物が撮影された画像と誤認するケースが多かった。S地区では、野生動物は人気のない夜間に撮影されていたため、夜間の撮影画像は動物が撮影されている画像であるという単純な分類をしている可能性が挙げられた。しかし、3つのモデルで教師データ数が最も多かったモデルcでは、夜間の何も撮影されていない画像を誤認する傾向が低くなっており、この要因について今後さらなる精査が必要である。例えば、画像中で判断に用いたとされる特徴部分を可視化するヒートマップ手法の活用が挙げられる。一般的に機械学習はデータ数が多いほど精度が向上するとされている。単純な分類であれば、少量の教師データでも比較的良好な精度で分類できるが、少ない教師データ数でより高度な分類モデルを作成するには、教師データの前処理やモデルの最適化を検討する必要が認められた。

Table2 各モデルの評価結果
Evaluation results of each model

モデル	モデルa	モデルb	モデルc
サンプル数	1,000	1,000	1,000
正答数	895	910	929
不正答数	105	90	71
正答率	89.5%	91.0%	92.9%

4. おわりに

本稿では、学習済みモデルの転移学習を活用することにより、分類モデルを作成した結果、少量のデータでも単純な分類であれば比較的良好な精度で分類モデルを作成することが可能であることが示唆された。

謝辞：画像データの収集に協力いただいた関係者の方々に心より御礼申し上げます。

参考文献1)Karen Simonyan, Andrew“ZissermanVery Deep Convolutional Networks for Large-Scale Image Recognition”, 2015