

YOLOv5 モデルを用いた動物識別器 —小さく写る動物に対する識別精度の向上—

森岡楓介[†] 藤嶋教彰[‡]

松江工業高等専門学校専攻科電子情報システム工学専攻[†] 松江工業高等専門学校電気情報工学科[‡]

1. はじめに

農作物被害で困っている声は各所で聞かれ、特に対策にかかる人的負担が農家を悩ませている。そこで被害防止のため、小型コンピュータ搭載の無人ドローンが定期的に飛行し、動物を識別するとスプレーを噴射して害獣を畑から追い払う方法が求められる。人が常に田畑を監視する必要がなく、農家の負担軽減につながる。

この場合、小型コンピュータに搭載可能な対象動物の場所を識別する AI 学習済みモデルが必要となる。著者は識別対象を猿、猪、鹿に絞り、YOLOv5l [1]を用いて、画像に写る各害獣の位置と種類を特定する動物識別器を開発している。

予備実験から、Web サイトで得た動物の実写画像（自然画像）3000 枚で識別器を開発すると、小さく写った動物が識別しづらいと確認した。

これは、小さく写る画像が学習用画像に少ないことが主原因と推察される。だが、求める画像データを大量に外部から取得することは難しい。

そこで、背景画像に面積の小さい動物画像を重ねた合成画像を学習画像に追加することで小さく写る動物の識別精度向上を目指す。

2. 開発環境

識別器は Python 3.7.3, IPython 7.34.0, numpy 1.21.6, CUDA 11.2, opencv-python 4.6.0.66, PyYAML 6.0 などで開発する。実験は松江高専ディープラーニングサーバ (Intel Xeon Gold 6234×2, Tesla V100S×3, メモリ 256GB) を利用する。

合成画像の作製は Photoshop 2022 の Python スクリプト機能を用いる。4 章で述べる内容の画像取得以外はスクリプトにより自動で行われる。

3. YOLOv5

YOLOv5 は Github で公開されている。5 つのモデ

ルがあり、高精度モデルほど処理速度が遅い。小型コンピュータ利用では、識別精度と処理速度を同時に考慮しなければならない。そこで、予備実験を行い、5 つのモデルを学習させ精度を確かめた。処理速度と識別精度の関係を考え、今回は 2 番目に精度が高い YOLOv5l を使用する。

4. 合成画像の作製

合成画像は畑の画像に対して、背景が透明の動物画像を重ねることで作製する。畑画像はまず短辺が 640 ピクセルとなるようにリサイズする。その後、加工前後で中心画素が同じになるように切り抜き、縦横 640 ピクセルの正方形画像とする。動物画像は縦横比変更しないように拡大縮小してサイズ調整する。動物画像の重畳位置はランダムとする。

画像は Web サイトから取得しており商用利用可である。畑画像は photo AC から、自然画像と動物画像の素材は Flickr が取得元である。

5. 実験方法

自然画像のみを学習させたモデルと、動物が小さく写る合成画像も学習させたモデルの識別精度を比較する。各学習方法で 5 つのモデルを学習させた。学習済みモデルにテスト用画像を入力し、出力画像から識別成功率を求めた。

2 種類の実験をした。1 つは合成画像を学習させても大きく動物が写ることが多い自然画像の動物識別に影響が出ないかを調査するものである。もう 1 つが合成画像を学習させることで小さく動物が写る画像の動物識別の精度が向上するかを調査するものである。

自然画像は動物ごとに学習用 1000 枚とテスト用 100 枚を用意した。動物画像は元画像の長辺が 640 ピクセルになるようにリサイズし、短い辺に余白をつけ加えて正方形画像を作成した。学習用のうち 2100 枚を train データ、900 枚を validation データとした。学習用画像に対して、フリーソフトの labelImg を用いてラベルデータを作製した。

An Animal Identifier using YOLOv5 model -Study on increasing identification accuracy of animals with small area in the pictures-

[†]Fusuke Morioka, Advanced Electronic and Information Systems Course, NIT, Matsue College

[‡]Noriaki Fujishima, Department of Electrical and Information Engineering, NIT, Matsue College

合成画像は学習用とテスト用で各 750 枚用意した。畑領域に対する動物の面積割合が 1%から 5%まで 1%刻みで異なるものがあり、各割合で 150 枚ある。学習用は 10 枚の畑画像と各動物 5 枚の動物画像を合成して作製した。そのうち、525 枚を train データ、125 枚を validation データとした。テスト用は学習用画像作製で使用したものと別の畑画像 10 枚と各動物 5 枚の動物画像から作製した。

学習は学習用画像と作製したラベルデータを用い、合成画像を学習に加えた場合と加えない場合で各 5 つ識別モデルを作製した。バッチサイズは 256 とした。学習回数はアーリーストッピング法を用いて自動で決定した。

実験 1 として、Flickr で収集した各動物 100 枚ずつのテスト画像で自然画像に対する識別精度を調査した。総個体数に対する識別数の割合を、その動物の識別率とする。総個体数は猪 163 体、鹿 185 体、猿 181 体である。また、動物の群れのうち 1 匹でも識別できていれば追い払い動作は可能だと考えるため、実用上の精度として各画像で 1 匹でも害獣を識別できた割合を求めて存在検出率とした。

実験 2 として、テスト用の合成画像 750 枚を用いて、動物が小さく写る場合の識別成功率を調査した。総個体数は各面積割合に対し、各動物 50 体である。そのうち識別できた割合を動物識別率とする。各面積割合で、動物ごとに求めた。

6. 実験結果

モデルの学習回数と validation loss の関係を記録したグラフが自動で記録される。学習させたモデルは 10 個ある。その中で平均的なグラフを選んで図 1 に示す。

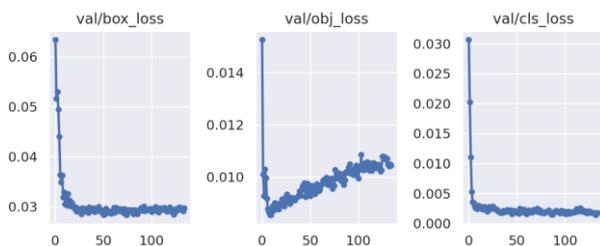
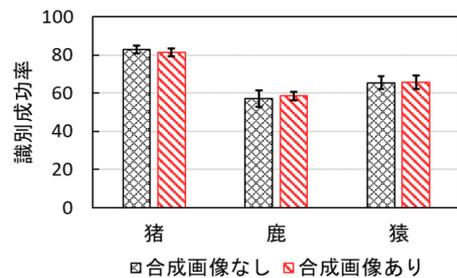


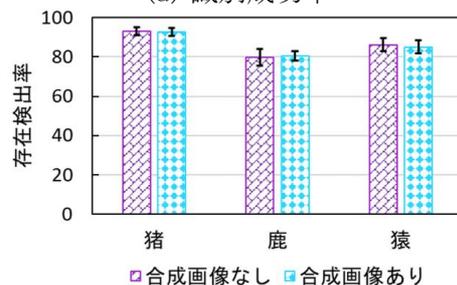
図1 validation loss の推移

実験 1 の結果を図 2 に示す。結果に対して F 検定および t 検定を実行したところ、識別と存在検出ともに有意差なしとなった。よって、動物が小さく写る合成画像を学習させても自然画像の識別精度低下は本実験では認められない。

実験 2 の結果を図 3 に示す。合成画像を学習したモデルの方が常に高い識別率である。すなわ



(a) 識別成功率

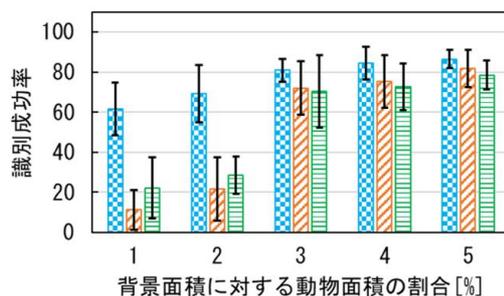


(b) 存在検出率

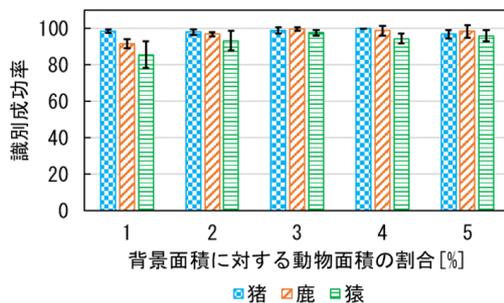
図 2 実験 1 (自然画像) の結果

ち写る動物の面積が小さい場合、有効性が認められる。

以上から、合成画像を学習用として画像に混ぜ込む手法で識別精度向上を達成したといえる。



(a) 合成画像未学習時



(b) 合成学習時

図 3 実験 2 (合成画像) の結果

参考文献

[1] University of Washington, Joseph Redmon, You Only Look Once: Unified, Real-Time Object Detection, <https://arxiv.org/pdf/1506.02640.pdf>