

PSPNet の質感に対するクラス分類精度の評価と考察

岡村 一矢[†] 松村 遼^{††} 北風 裕教^{††}

[†] 大島商船高等専門学校専攻科電子・情報システム工学専攻

^{††} 大島商船高等専門学校情報工学科

1. はじめに

近年、ドローンを用いた害獣・害鳥追い払いシステムの開発が注目されている。我々はこれまでに画像領域分割アルゴリズム Semantic Segmentation 手法である Pyramid Scene Parsing Network (PSPNet) [1]を用いて、カラスとイノシシの2種類の動物と、それらに誤検出されやすい飛行機と車、および背景画像を加えた5クラス分類の認識実験を行っている。しかし、イノシシは生活環境の影響で毛の質感に個体差が存在するため、多様な学習用画像の準備が必要となるが、愛玩動物ではないため入手困難であり、認識精度の向上にはつながらなかった。

そこで本研究では、写真画像ではなくイノシシの3Dモデルを構築し、仮想空間上で撮影カメラの位置を変更して2D画像を再取得するData Augmentationを行うと同時に、イノシシの3Dモデルの高品質化(毛の質感の有無)を行うことで、毛並みの質感情報が認識精度に与える影響について検討したので報告する。

2. 学習用データセット作成

本研究では、3Dモデルまたは実写画像から学習用データセットを作成する。イノシシの3Dモデルは毛の質感の有無を分け、それぞれ別のデータセットを作成する(図1)。学習データは、カラー画像とラベル画像の2種類から構成される。また、ラベル画像は、カラーパレット形式でインデックス番号とRGB情報が対応している。

飛行機、カラス、イノシシの3クラスは3Dモデルから2D画像を再取得した画像を利用する。3Dモデルを、IDEを内蔵するゲームエンジンであるUnityに取り込み、物体から一定距離に保ったカメラの位置を垂直方向、水平方向に移動しながら画像を撮影する。この一連の処理を物体とカメラの距離を変え4か所分行う。飛行機、カラスは全方位から撮影し4,080枚、イノシシは垂直方向60°から-10°の範囲内で撮影し2,880枚の画像を取得した。撮影画像はランダムに選択した背景画像と合成しカラー画像を作成、物体領域のピクセルにインデックス番号を記憶したラベル画像を作成する。

一方、車は2,909枚の実写画像を用意し、そのままカラー画像として使用する。ラベル画像は車を事前に学習したモデルDeepLabv3の出力画像を使用し、車のピクセルにインデックス番号を上書きしてラベル画像を作成する。



図1 イノシシの毛の質感



図2 結果例

(左:質感なし, 右:質感あり)

表1 IoUによる認識実験の結果

クラスIoU	質感なし	質感あり
背景	0.936	0.995
飛行機	0.669	0.726
車	0.941	0.953
カラス	0.841	0.919
イノシシ	0.543	0.871
mIoU	0.749	0.869

3. 学習と評価方法

イノシシの毛の質感の有無が異なる2つのデータセットを、それぞれPSPNetで学習する。また、入力画像のサイズは473×473とし、ハイパーパラメータにはバッチサイズ16、エポック数14、学習率0.01を設定する。

各クラス20枚のテスト画像から得られた出力画像と正解ラベルと比較した領域の重なり度合いを表す評価指標 Intersection over Union (IoU)を用いて評価を行う。

4. 実験結果と考察

IoUによる実験結果を表1、テスト画像の結果画像を図2に示す。質感ありデータセットを学習したモデルはイノシシの認識精度をIoUにおいて0.543から0.897へと大幅に向上させることが確認できた。また、イノシシの毛の特徴を学習できたことにより、背景を誤認識するなど他のクラスと間違えることが減少し認識精度が向上したと考えられる。さらにイノシシの学習過程が改善されたことにより、学習パラメータが最適化され全クラスの認識精度向上に影響したと考えられる。一方で、体の一部が草葉に隠れている場合には、その部分より下を認識できない例が見られた(図2)。この問題は、似た状況を再現し学習することで解決できると考えられる。

5. まとめ

イノシシの学習用画像の質感をより現実的にすることで、認識精度の飛躍的向上を確認することができた。今後は状況や動作再現を行い更なる向上を目指す。

参考文献

[1] H. Zhao, *et al.*, In CVPR, 2017, pp. 2881-2890.