

多段ニューラルネットワークを用いた動物の品種識別

Animal Breed Identification using Multi-Stage Neural Network

渡部 宏樹[†]渡辺 裕[†]Hiroki Watabe[†]Hiroshi Watanabe[†][†] 早稲田大学大学院 基幹理工学研究科

Abstract In this paper, we propose a method that combine YOLO and CNN for pet breed identification. By the experiment, we confirmed the validity of our method that shows more than 10% better score compared with the one using only YOLO.

1. はじめに

近年、スマートフォンの普及により多くの人が気軽に写真をインターネット上にアップロードするようになってきている。それらの画像データを活用するために、写真中から特定の物体を検出し、切り出す必要が生じることがある。我々は、代表的な被写体の一つであるペットに焦点を当て、画像中のペットの顔の検出、品種識別手法を研究している。本稿では、物体検出分野で高い性能を発揮している You Only Look Once (YOLO) [1] と品種識別用の Convolutional Neural Network (CNN) を用いた二段階の手法を検討する。これを一段階で直接品種識別する手法と比較し、二段階手法の優位性を示す。

2. 提案手法

本稿では、YOLO を用いて画像中から猫・犬の顔領域を検出し、検出した顔領域を猫・犬それぞれに対応した CNN で品種識別するという二段階の方法でペットの品種識別を行う。

YOLO [1] は高速・高精度な物体検出ニューラルネットワークである。YOLO はまず始めに画像全体を $S \times S$ 個のセルに分割する。次に、セルごとに B 個の候補領域を予測する。各候補領域は領域内に物体が存在する確率を保持する。次に、セル内に存在する物体のクラスをセルごとに予測する。最後に、候補領域内の物体の存在確率とセル内の物体のクラス確率を掛け合わせ、閾値以上となった候補領域を物体領域として検出する。

3. 実験

YOLO による品種識別まで行う一段階の猫・犬検出と、提案手法である YOLO による猫・犬検出と CNN による品種識別を組み合わせた二段階の手法の評価実験を行った。本実験では、"The Oxford-IIIT-Pet dataset" [2] 内の猫・犬の画像各 12 品種を用いて実験を行った。訓練に 1 品種あたり 150 枚、テストに 50 枚の画像を使用した。

実験結果を表 1 に示す。表 1 より、提案手法が YOLO のみで品種識別まで行う手法より高い性能を示すことが確認できた。また、各手法が検出した物体数を表 2 に示す。提案手法の方が YOLO 単体よりも 100 個ほど多くの物体を検出した。この検出数の差が表 1 の性能

表 1: 実験結果

	Precision	Recall	F 値
提案手法	0.771	0.724	0.747
YOLO	0.667	0.563	0.610

表 2: 検出物体数

	提案手法	YOLO
検出物体数	1130	1016

差に現れたと考えられる。検出数の差が生じた原因は次のように考えられる。提案手法の YOLO は猫か犬かの 2 クラス識別を行うが、YOLO 単体の場合は猫犬各品種計 24 クラスの識別を行う。クラス数の多い YOLO 単体の場合の方が識別は難しく、識別の精度が低くなる。そのため、処理途中のプロセスで物体かどうかの判断をする際、セル内の物体のクラス確率が小さな値を取ることで候補領域内の物体の存在確率と掛け合わせた値が閾値を下回ってしまい、物体として検出されないということが起きるのだと考えられる。

4. まとめ

本稿では、ペットの品種識別のために YOLO と CNN を組み合わせた二段階手法を提案した。実験の結果、提案手法が YOLO 単体で識別するのに比べて高い性能を示すことを確認した。

文 献

- [1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi: "You Only Look Once: Unified Real-Time Object Detection", IEEE Conference on Computer Vision and Pattern Recognition, pp.779-788 (2016).
- [2] O.M. Parkhi, A. Vedaldi, A. Zisserman, and C.V. Jawahar, "Cats and Dogs", IEEE Conference on Computer Vision and Pattern Recognition, pp. 3498-3305, (2012).

[†] 早稲田大学 基幹理工学部 情報通信学科

〒169-0072 東京都新宿区大久保 3-14-9 早大シルマンホール 401

TEL: 03-5286-2509 E-mail: xps169@akane.waseda.jp