

ディープラーニングを用いた擬態している生物の物体検出

坂本 悠哉† 大川 茂樹†

千葉工業大学 先進工学部 未来ロボティクス学科‡

1. はじめに

昨今、機械学習やディープラーニングによる画像認識技術が発展しており、リアルタイムでも高い精度での物体検出が可能であり、検出精度は人の能力以上になっている。一方で、人の目では判別が困難な例もある。その一例として擬態が挙げられる。擬態とは、自然界における生物の生存戦略であり、特定の物体に色や形を似せることである。またディープラーニングによる物体検出では、データ数が多くなると精度が上がる傾向にあるが、高い精度を得るには数千数万のデータが必要となる。しかし、人の目では判別が困難である擬態生物は大量のデータを集めるのが課題となる。

そこで本研究では、人でも判別が困難な擬態画像に対し、より少ない画像数でも検出可能な方法を検討する。

2. 目的

ディープラーニングを用いて少ないデータ数で擬態している生物の検出を可能にする方法の模索を目的とする。

3. 方法

3.1 データセット

本研究では、Fig. 1 のように検出対象の擬態生物にカエルを選択している。集めたカエルの画像はアノテーションツールである labelImg を用いてラベル付けをおこなう。また、擬態は生物が背景の色または特定の物体に模している状態とする[1]。

また、ディープラーニングでは大量のデータが必要となるため、データの増量をおこなう。本研究では、画像の回転処理、エッジ検出によりデータを増量する。



Fig. 1 擬態しているカエルの例

3.2 YOLOv5 による学習

擬態するカエルを検出する方法として、近年物体検出の主な手法として挙げられる YOLO (You Only Look Once)v5[2]を使用する。YOLO シリーズの初代である YOLOv1 をもとに改良が続き、YOLOv5 では、非常に速く、正確な検出が可能である。

3.3 SSD による学習

1枚の入力画像から複数クラス識別済みのN個の物体ボックスを、一度に検出するSSD(Single Shot MultiBox Detector)[3]を使用する。SSDにはボックスサイズのずれをクラスの予測情報を使用し、修正・深層学習していくため、AIの学習速度が速い点が挙げられる。

4. 結果

4.1 YOLOv5 と SSD の検出精度の比較

まず予備的な実験として、犬の画像を用いた検出をおこない YOLOv5 と SSD での比較をした。データ数として訓練画像である Train データが 413 枚、評価画像である Val データが 100 枚、テスト画像である Test データが 40 枚とし YOLOv5 と SSD でデータセットは共通として学習をおこなった。

Table 1 に検出結果を示す。Test データ内で犬として検出出来た画像の割合を Accuracy とした。YOLOv5 による検出では Accuracy が 0.57 であったが、SSD による検出では Accuracy が 0.90 となり、SSD の方が高い検出精度が得られた。

Detecting Mimicry by Deep Learning
†Sakamoto Yuya and Okawa Shigeki
‡Department of Advanced Robotics,
Chiba Institute of Technology

Table 1 2つの手法による比較の結果

使用した検出手法	Accuracy
YOLOv5	0.57
SSD	0.90

4.2 YOLOv5 によるカエルの画像の検出

データ数は前処理を行っていない画像 420 枚を Train データ, 120 枚を Val データとして学習をおこなった. また, そこに加え画像の回転, エッジ検出をした画像を含んだデータでも学習をおこなった. Test データは 40 枚とする.

学習した重みを用いて Test データ内の画像からカエルの検出をおこなう. ここで画像内の全てのカエルを検出が可能となった場合を Positive とし, カエルを検出出来ていない場合やカエル以外の似た形状のものを検出している場合, 画像内に複数カエルがいる際一部のみしか検出出来ていない場合は Negative とした. Test データの Positive の割合を Accuracy とした. 検出時の confidence は 0.5 以上のバウンディングボックスを表示する.

Table 3 は YOLOv5 で学習した重みで検出をおこなった結果であり, Accuracy が最高で 0.57 となった.

Table 2 YOLOv5 によるカエルの検出結果

Train (枚)	Val (枚)	Test (枚)	Accuracy
420	120	40	0.50
560*	160	40	0.50
840**	240	40	0.52
560***	160	40	0.57

* 画像の回転画像 140 含む

** 画像の回転画像 420 含む

*** エッジ検出画像 140 含む

4.3 SSD によるカエルの画像の検出

学習に使用したデータは 4.2 での YOLOv5 で使用したデータと同様のものを使用している. また検出時の Accuracy の導出方法は 4.2 と同様とした.

Table 4 は SSD で学習した重みで検出をした結果であり, Accuracy が最高で 0.82 となった.

Table 3 SSD によるカエルの検出結果

Train (枚)	Val (枚)	Test (枚)	Accuracy
420	120	40	0.70
560*	160	40	0.67
840**	240	40	0.70
560***	160	40	0.82

* 画像の回転画像 140 含む

** 画像の回転画像 420 含む

*** エッジ検出画像 140 含む

5. 考察

擬態しているカエルを検出対象とした際に, SSD が高い Accuracy を得た. この理由として, SSD の画像の畳み込み (入力画像の画素値に対して, 着目した画素値及びその周辺の重み付き総和を算出し, 出力画像の画素値を求める処理) をする際に読み込むグリッドのサイズをより小さくして, 検出した物体候補に対して複数のサイズの領域で枠取りをおこない, 大きさと縦横比が異なる複数種類の領域を準備して正確なバウンディングボックスの位置を予測するといった特徴を持つからではないかと考える. この特徴により, 擬態しているカエルやカエルが擬態対象とする物体に対しても細かい予測をおこなうことが可能なため, 高い Accuracy が得られたと推測する.

6. おわりに

本研究では, 擬態しているカエルを対象として少ないデータ数での検出方法を模索した. 今後はデータ数の増加や前処理による Accuracy の向上に着目しより実用的な精度への向上を目指す.

参考文献

- [1] 友野海, 延原肇, 河合新, “擬態によって学習・識別困難な対象の GAIN に基づく深層学習”, 人工知能学会, (2019)
- [2] Liu Yifan, Lu BingHang, Peng Jingyu, Zhang Zihao, “Research on the Use of YOLOv5 Object Detection Algorithm in Mask Wearing Recognition”, World Scientific Research Journal, 276-284, (2020)
- [3] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, Alexander C. Berg, “SSD: Single Shot MultiBox Detector”, Computer Vision – ECCV 21–37, (2016)