

YOLOv5 を利用したカモの識別

Duck Recognition with YOLOv5

石井 友基†
Yuki Ishi稲岡 歩望†
Ayumu Inaoka岡田 聖冬†
Masato Okada堀尾 将吾†
Shogo Horio宮下 恭一†
Kyoichi Miyashita

1. 研究の背景

現在、野鳥調査では、池ごとに写真を撮り、そこから、種族と個体数を数え上げている。その作業は、現在、人力によって行われている。私たちの居住地区の付近では、特にカモの調査が盛んである。

また、YOLOv5 は、2020 年 6 月に公開された最新の物体検出手法である。他の手法より高速の処理が可能で、リアルタイムでの実行も可能だ。さらに、自作データセットを YOLOv5 で簡単に学習できることも特徴である。YOLOv5 を利用することで、個人が、所有している PC での、物体検出が容易になった。^①

物体検出とは、画像・動画内に写っている物体の位置とクラス(人間、車といったカテゴリー)を特定する方法である。正解データを大量に収集して、未知の画像に対しても予測ができるモデルを作成する。

以上のことから、私たちは、YOLOv5 を利用して、AI によるカモの識別を可能にしようと考えた。

2. 研究の目的

PC を使って、専門家以外でも容易、かつ、少ない労力で、カモの種族判別、個体数調査ができるようにする。私たちが、今回対象にしたカモは、コハクチョウ、オシドリ、ヨシガモ、ヒドリガモ、オカヨシガモ、マガモ、トモエガモ、カルガモ、ホシハジロ、オナガガモ、ハシビロガモ、コガモ、キンクロハジロ、アカハジロ、ミコアイサ、カワアイサの 16 種のオスである。オスを対象とした理由は、カモのメスは、種族間での見た目の差異が小さく、識別が困難だからである。

3. 研究方法

3.1 サンプル画像収集

物体検出には、大量の正解データが必要である。私たちは、必要な分のカモの画像を、インターネット上から収集した。それぞれのカモの特徴をつかむために、数羽のクローズアップされた写真を多く収集した。

3.2 (1) CVAT (Computer Vision Annotation Tool) について^②

正解データを人手で付与するための、アノテーションツールの一つである。今回行った物体検出というタスクでは、検出したい対象(カモ)に正解の領域を付与する必要があった。

† 兵庫県立小野高等学校

CVAT は、INTEL が作成している OSS。基本はブラウザ上で利用できる、という利点がある。

YOLOv5 のための、クラスと画像上の座標をまとめたテキストドキュメントファイルに、エクスポートできる。

3.2 (2) サンプル画像のアノテーション

3.1 で収集したサンプル画像を、CVAT でアノテーションする。

カモのことを識別することができるものを、カモ全体を覆うように区切る。今回は、池の野鳥調査に使用できるモデルを作ることを目的にしていたので、泳いでいる姿を中心に、教師データとした。

図 1 は、実際に CVAT で処理した画像である。緑の四角内部が、正解領域であり、人手で付与された。



<図 1>

<表 1> カモの種類と教師データ数

カモの種類	教師データ数
コハクチョウ	61
オシドリ	126
ヨシガモ	77
ヒドリガモ	49
オカヨシガモ	47
マガモ	77
トモエガモ	33
カルガモ	75
ホシハジロ	127
オナガガモ	93
ハシビロガモ	53
コガモ	90
キンクロハジロ	117
アカハジロ	12
ミコアイサ	125
カワアイサ	42

3.3 学習

Google Colaboratory で学習を行った。

学習は、ある程度の範囲で、回数を多くするにつれて、モデルの精度が高くなる。また、利用するデータの数によって、一回の学習にかかる時間が変化する。

まず、3.1, 2で収集した画像と教師データを関連付けながら、Google Colaboratory にアップロードした。

次に、学習を開始した。学習は、回数などを指定して、PC が処理する。Google Colaboratory の性質上、pro 版を利用しない場合、一回の利用可能時間が限られているので、注意が必要である。

学習結果は、重みファイルとして、Google Colaboratory に保存される。この重みファイルを利用することで、次の学習を、既に得た結果を基に、行うことができる。

最後に、検証用の画像に、得たモデルを適用して、精度を確認した。このとき、どれほどの確率値以上で、対象オブジェクトと認識するかの設定も可能である。PC 上で表示することで、視覚的に確認できる。

(1) オナガガモ単種

YOLOv5 の使い方を確認するために、オナガガモで学習を行った。

(2) 16種

対象とした16種の学習を行った。

実用性を高めるために、教師データの追加や、正解領域の変化をさせた。

3.4 検証

YOLOv5 では、学習の経過の様子、正解領域の確率値がわかる。都度、検証をして、より正しい学習ができるように、改良をした。

以降、信頼できる精度のモデルが得られるまで、3.4と3.5を繰り返した。

4. 実験結果

自作のデータセットで学習した結果、得られたモデルで、カモが写っている画像を検出した。

YOLOv5 では、検出された領域は、それぞれのクラスに対応した色で、自動に囲まれる。さらに、クラスに対応させた名前と、信頼度となる確率値も、同時に表記される。

4.1 オナガガモ単種

初めは、複数回検出される個体や、一度も検出されない個体が多かった。さらに、木の上部など、カモと全く関係ない領域も、オナガガモとして検出される部分があった。(図2)

学習回数を増加させることで、モデルの精度は、大幅に上がった。オブジェクトと検出が、一対一でより対応するようになった。また、カモの存在しない領域を、検出することも減った。しかし、完璧ではなかった。(図3)

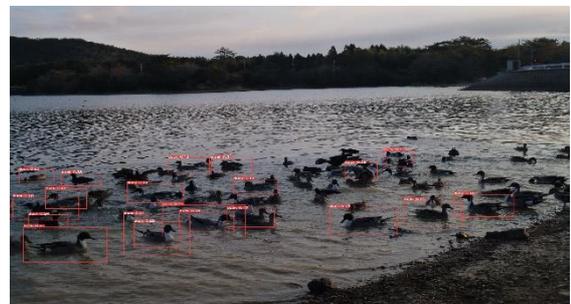
しかし、学習回数を過剰に増やすと、モデルの精度が、かえって下がってしまった。(図4)後述するが、過学習によるものと考えられる。



<図2>



<図3>



<図4>

なお、検証済み画像には、クラスによって囲む線の色が変わる。左上に、クラス名と、確率値が書かれる。

4.2 16種(クローズアップ写真)

初めは、クラス定義が間違っていたので、全ての種類のカモが、アカハジロに誤検出されてしまった。(図5, 6) 図5はアカハジロ、図6はマガモの画像である。

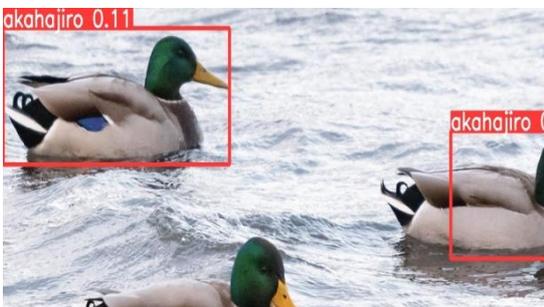
訂正以降、適切な回数の学習を繰り返すと、それぞれのカモについて、正確な検出ができるようになった。

(図7~9) 図7はアカハジロ、図8はマガモ、図9はキンクロハジロの画像である。

表2は、ランダムに抽出した画像の検証結果である。検出精度は、(個体数) / (重複検出+未検出+誤検出) で定義した。



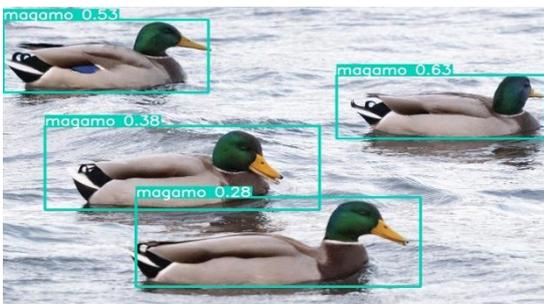
<図5> アカハジロ



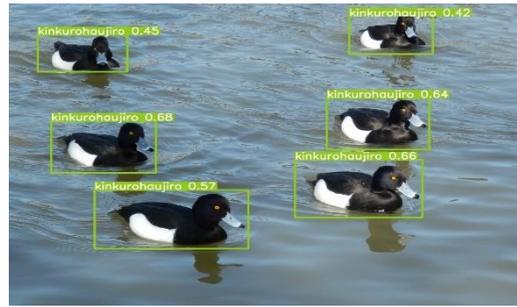
<図6> マガモ



<図7> アカハジロ



<図8> マガモ



<図9> キンクロハジロ

なお、図 5、7で検出されていないカモがいるが、中段の二羽以外は、メスなので、検出していないのが正しい。

種類	個体数	重複	未検出	誤検出	精度
コハクチョウ	12	2	2	1	70.6%
オシドリ	10			3	77.0%
ヨシガモ	10	2			83.3%
ヒドリガモ	14	2			87.5%
オカヨシガモ	14		6		57.1%
マガモ	13	2			86.7%
トモエガモ	10	3		1	71.4%
カルガモ	12	4		1	70.6%
ホシハジロ	12	4		2	66.7%
オナガガモ	12	3	2		70.6%
ハシビロガモ	14	2			87.5%
コガモ	12	3		1	75.0%
キンクロハジロ	10	5		4	52.6%
アカハジロ	12	7			63.2%
ミコアイサ	11	1	1		84.6%
カワアイサ	12	4			75.0%

<表2>

4.3 調査用写真

今回の研究では、野鳥調査に実際に利用できるモデルを製作することだった。そこで、学習に利用したクローズアップの写真ではなく、遠方から、複数のカモが写っている写真で、検証を行った。

しかし、利用した教師データとの齟齬により、期待ほどの精度がなかった。(図10)よって、遠方からの写真を、教師データに追加して、再び学習を行う方法と、遠方からの写真を、トリミング、クローズアップして、検証する方法を試した。後者の方が効果があった。(図11)



<図10>



<図11>

図 11 は、図10の中央寄りの右下を、切り取った後、再検出した画像である。図 10 では左のオナガガモは検出されていないのに対し、図 11 では二羽とも検出されている。

5. 考察

5.1 過学習について

図3と図4より、過剰な回数の学習をすると、過学習が起きた。

過学習とは、学習をし過ぎて、一般的ではなく、一つ一つの要素にフィットしすぎたモデルができることである。特殊な事例を取りすぎて、精度が下がってしまう。

対策として、正則化、ハイパーパラメーターチューニング、ドロップアウトなどがある。単純な方法として、モデルの精度を上げるために、学習回数を増やすのではなく、教師データを増やした方が良い。

今回の件では、学習モデルが 50 程度、学習回数が 1 万回程度だった。

5.2 実用性について

4. 3にあったように、調査に実際に使われると思われる画像を検証すると、検出の精度は大幅に下がった。理由は、教師データとの乖離によるものと考えられる。

現状の対策として、教師データの画像と類似するように編集している。しかし、手間がかわってしまうので、自動に処理するように設定した方が良い。

6. 結論と展望

今回の研究は、野鳥調査を補助するための、画像検出モデルを作成するものだった。4. 2より、識別の精度は十分に高くなった。

しかし、実用するには、改善すべき課題がある。今後は、学習を進めるのではなく、利便性を高めるために、プログラミングを実行する、

謝辞

ご指導いただいた株式会社 BRAIN の方々に、深く感謝申し上げます。

参考文献

- (1) ALPHA SYSTEMS INC. “YOLOv5 を使った物体検出”, アルファテックブログ. 2021/8/20, https://www.alpha.co.jp/blog/202108_02, (2022/6/20)
- (2) PA Lab メディア “Intel 製のアノテーションツール「CVAT」とアノテーションの活用事例”, PA Lab 技術ブログ, 2021/7/12, <https://pa-laboratory.com/2021/07/12/cvat-annotation-tool/>, (2022/7/10).
- (3) 酒井隆成: “過学習とは? 初心者向けに理由から解決法までわかりやすく解説”, DATA VIZ LAB, 2021/7/14, <https://data-viz-lab.com/overfitting>, (2022/7/11)

- (4) Tian, Ming, および ZhihaoLiao. 「YOLOv5 に基づく花の画像分類法に関する研究」。 *Journal of Physics : Conference Series*. 巻 2024. No. 1. IOP Publishing, 2021.

引用コード

<https://github.com/ultralytics/yolov5>

CVAT

<https://github.com/openvinotoolkit/cvat>