YOLO 法による動画からの鳥類等のリアルタイム検出の高精度化

中山　紘喜[1]・安枝　裕司[2]・久冨　学[3]・野上　敦嗣[4]・松本　亨[5]

[1]非会員　(株)エコプラン研究所（〒808-0146 福岡県北九州市若松区高須西 1-14-13）
E-mail: h-nakayama@kitakyu-u.ac.jp

[2]非会員　北九州市響灘ビオトープ（〒808-0021 福岡県北九州市若松区響町 1）
E-mail: yasueda@gmail.com (Corresponding Author)

[3]非会員　(株)エコプラン研究所（〒808-0146 福岡県北九州市若松区高須西 1-14-13）
E-mail: hisadomi@hotmail.co.jp (Corresponding Author)

[4]非会員　北九州市市立大学国際環境工学部（〒808-0135 福岡県北九州市若松区ひびきの 1-1）
E-mail: nogami@kitakyu-u.ac.jp (Corresponding Author)

[5]正会員　北北九州市市立大学国際環境工学部（〒808-0135 福岡県北九州市若松区ひびきの 1-1）
E-mail: matsumoto-t@kitakyu-u.ac.jp (Corresponding Author)

風力発電などの周辺開発が続く響灘ビオトープでは，チュウヒを頂点に希少な生態系が保全されており，定常的なモニタリングによる鳥類の生活行動の把握は重要である．本研究グループでは深層学習法 YOLO を用いて，鳥類のリアルタイム自動検出システムを開発し，運用することで 4K 動画に対し最小で 156 ピクセルの大きさの鳥類を検出できた．しかし風車，雲，植物，昆虫などの誤検出も見られ，学習モデルの改善が課題となった．本研究では学習データの整理，増強を行い検出精度の向上を図った．響灘ビオトープで得られた画像を学習することで，正しい鳥類の検出数は約 2.4 倍に増加した．一方で，風車の誤検出は大きく減少し，雲は大きなサイズの誤検出が減少した．また，昆虫はトンボなど一部で識別ができており，更なる検証を進めていく．

Key Words: *bird survey, YOLO, real-time detection*

1. はじめに

福岡県北九州市に位置する響灘ビオトープは，絶滅危惧種に分類されるチュウヒ（タカ科の鳥類，国・絶滅危惧ⅠB類，国内希少野生動植物）やベッコウトンボ（昆虫類，国・絶滅危惧ⅠA類，国内希少野生動植物）[注1)]の生息や繁殖が確認され，保全活動がなされている．中でも湿地生態系の頂点にいるチュウヒは，営巣，採餌の場所となる水辺のヨシ原，湿地，草地の面積減少や未管理により，繁殖が困難になりつつある状況であるため，その行動や生息環境を定量的に把握することは、今後本地区の生態系保全対策において極めて重要である．また本地区は産業用地の一部をビオトープとして整備したという経緯から，その周辺ではリサイクル促進工場や再生可能エネルギーをはじめとした発電施設の建設が進み，数年後には洋上風力発電施設の建設も計画されている．

このように開発地区が隣接した自然保護区といった性質を持つ本地区ではチュウヒをはじめ行動圏が広大な鳥類に対し継続的なモニタリングによる鳥類の生活行動の把握は重要となるが，人材やコストの面から実施されていない．この課題に対しビデオカメラを用いた観察は定常性おいて有用だが，膨大な映像データを人手で詳細に解析することは困難であり，近年では様々な機械学習，深層学習の手法を応用した画像解析の自動化研究が活発に行われている．中でも YOLO(You Only Look Once)法[1)]は高速性と高精度を両立しており[2),3)]，高精度な鳥類検出やリアルタイム運用によるデータ量の低減などが期待できる．

本研究グループでは，響灘ビオトープに設置された 4K ネットワークカメラと YOLO 法を用いて，リアルタイム鳥類自動検出システムを開発し，2022 年 1 月 22 日より運用を行い，その有用性について評価した[4)]．これ



図-1　開発したシステムで検出した鳥類の一例

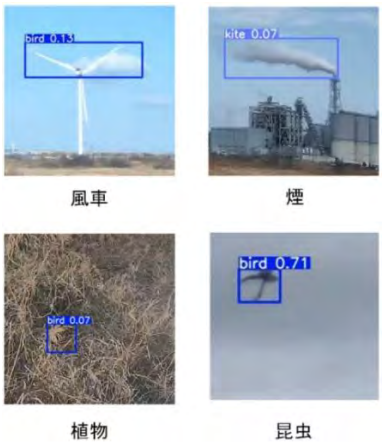


図-2　システム運用で確認された誤検出の例

により，**図-1**のような対象の位置を正確に捉えた検出を確認できたほか，4K動画に対し最小で156ピクセルの面積を持つ鳥類を検出できたが，**図-2**のように風車や雲，植物，昆虫類などの誤検出が見られ，学習モデルの改善が課題となった．検出に用いた学習済みモデルの学習にはCOCOデータセット[注2)]が使われており，誤検出された種はこのデータセットには含まれておらず，これが誤検出の一因と考えられる．これを受けて本研究ではシステム運用により得られた検出データを用いて，整理，増強した学習データセットによる学習を行い，学習済モデルの改善と誤検出の低減を図る．また本地区で保全されている希少種であるベッコウトンボを想定し，YOLOを用いたトンボ類の検出についてその有用性について検討を行う．



図-3　ネットワークカメラ設置地点と撮影方向

2.　研究方法

(1)　検出データの収集

本研究で用いたオリジナル画像データはすべて福岡県北九州市若松区にある響灘ビオトープで撮影した．撮影にはカメラ 4 台で全周撮影可能なパナソニック製ネットワークカメラ WV-X8570N を用いた．動画の解像度は 3840×2160 で，フレームレートは 15.00(fps)である．ネットワークカメラの設置地点と 4 台のカメラの撮影方向を**図-3** に示す．本研究では，西側を撮影している 3 番カメラの映像を主に用い，補助として北側を撮影している 4 番カメラの映像を用いた．これらのカメラはチュウヒの飛来が多く確認されている方向を撮影している．収集期間は 2022 年 2 月から 2022 年 12 月である．

(2)　深層学習システム

物体検出にはUltralytics社のYOLO-v5[注3)]を用い，学習済みモデルには同社より提供されているYOLOv5mを用いた．このモデルの性能は，パラメータ数が2140万，入力画像サイズが640ピクセル，mAP0.5が63.1%となっている．mAP0.5は物体検出の評価指標で，式(1)で与えられるIOU（Intersection Over Union）が0.5を超えたものを正解として求めた識別精度の全識別種平均である．この学習にはCOCOデータセット[注2)]が学習データセットとして用いられており，その内訳は識別種数80，画像数約33万，アノテーション数約150万となっている．本研究で用いた計算機環境はNVIDIA製GPUとGPUライブラリCUDAで構築し，pythonでプログラムを作成した．

$$IOU = \frac{\textit{正解領域と検出領域の重った面積}}{\textit{正解領域と検出領域の和集合の面積}} \qquad (1)$$

(2) 学習済みモデルの作成

学習には YOLOv5m のパラメータを初期値にした転移学習を用いた．学習に用いる画像とアノテーションのデータには，響灘ビオトープでのシステム運用で得られた検出結果を用いた．検出結果は動画ファイルとこれに紐づけられたテキストファイル群で纏められており，テキストファイルには，検出時フレーム，検出ボックスの矩形情報が含まれている．これらを用いてアノテーションデータを作成し，さらに動画ファイルから検出時の画像を抽出することで画像データを作成した．

3. 研究結果

(1) 風車，植物，煙・雲を加えた学習

誤検出された物体のうち，風車，植物，煙・雲を識別種に追加した学習データセットを用いて作成した学習済みモデルを用いて，誤検出の低減効果の検証を行った [4]．検証に用いたデータセットは 4 つでその内容は COCO データセット（A），A から飛行機，鳥類，馬，凧を抽出したもの（B），B に誤検出の目立った風車，植物，雲・煙を追加したもの（C），C の鳥類にビオトープで検出された鳥類画像を追加したもの（D）である． B での抽出種は，ビオトープでの運用において，鳥類と誤答した種である．

学習結果の比較を**表-1** に，検出結果の比較を**表-2** にそれぞれ示す．検出には 400 秒の動画を用いた．**表-1** において train，val はそれぞれ学習と検証に用いたアノテーション数で，**表-2** において検出数は鳥類と凧として検出された検出物の合計値で，各種の検出数は人が判定した内訳である．**表-1** より識別種が 80 ある A に対し種数の少ない他のモデルは mAP0.5 が約 10%向上している．**表-2** より種数を限定した B は A に対し風車の誤検出が減少し，鳥類の検出が増加した．C では誤検出が劇的に改善した一方で鳥類の検出数も A と B の検出数の 30%～46%に止まり，検出漏れが発生している．D では風車と煙・雲の誤検出は低いままだが，植物の誤検出が劇的に増加した．これは草陰を草地に降り立った鳥類と判断した結果と考えられる．鳥類の検出に関しては B と比較して約 2.3 倍に増加し，鳥類の検出性能が向上している．このことから，誤検出された物体を新たな識別種に追加することは，鳥類の検出精度向上に寄与することが分かる．

(2) 昆虫を加えた学習

データセット C 及び D は昆虫を識別種に含めておらず，追加した画像は冬に撮影されたものを使っており，他の季節の画像は使われていない．そこで飛行機，鳥類，馬，凧，風車，植物，煙・雲，昆虫の 8 種を識別種に持ち，追加するオリジナル画像は季節を網羅した学習データセットを作成した．データセットの内容は，B に鳥類，風車，植物，雲・煙，昆虫を追加したもの（E），C に E で追加した鳥類，植物，昆虫を追加したもの（F）である．F については D の検出結果を考慮し，検出数が増加した鳥類と，誤検出が増加した植物を追加した．

E，F の学習結果の比較を**表-3** に，B，C，E，F の検出結果の比較を**表-4** にそれぞれ示す．検出には昆虫の検出を検証するため前節とは別の 700 秒の動画を用いた．**表-4** において検出数は鳥類として検出された検出物の合計値である．**表-3** より mAP0.5 は**表-1** の学習済みモデルより低い値になったが，**表-4** より鳥類の検出数は B，C よりも多くなっており，響灘ビオトープでの主要な鳥類誤検出種を加えた学習済みモデルでの鳥類検出数を増加させることができた．一方で，誤検出については，風車の検出数が E，F ともに 0 で，B，C と比較して風車と鳥類を高い精度で識別できていると言える．しかし植物と煙・雲については E，F ともに B，C と比較して検出数が大幅に増加しており，昆虫についても同様に増加している．

煙・雲については，誤検出の傾向として検出面積の小ささが考えられる．**表-5** に E と F について，煙・雲と

表-1 誤検出種を追加した学習済みモデルの学習結果

モデル名	種数	train	val	mAP0.5
A	80	849942	36336	0.641
B	4	31040	1169	0.765
C	7	32311	1488	0.744
D	7	35003	2162	0.765

表-2 誤検出種を追加した学習済みモデルの検出結果

モデル名	検出数	風車	植物	煙・雲	鳥類
A	1328	1039	26	15	231
B	1336	776	165	29	353
C	116	4	1	0	107
D	2468	5	1613	3	802

表-3 昆虫を追加した学習済みモデルの学習結果

モデル名	tain	val	mAP0.5
E	32390	1319	0.662
F	33121	1578	0.739

表-4 昆虫を追加した学習済みモデルの検出結果

モデル名	検出数	風車	植物	煙・雲	昆虫	鳥類
B	1106	990	0	0	27	89
C	359	104	107	5	29	114
E	1841	0	719	647	117	358
F	1210	0	693	224	79	214

表-5　煙・雲の誤検出物の検出面積

データ名	検出数	面積900未満	面積900以上
E	647	527	120
F	224	224	0



図-4　面積の小さな鳥類学習画像例

誤検出した検出物の検出面積の比較を示す．**表-5** よりFでは全ての検出物が，Eでは約80%の検出物が900ピクセルに満たない面積であった．またEでも10000ピクセルを超す面積の検出物は全体の約4%にすぎなかった．このことから，煙・雲の誤検出について，図-2に示すような運用時に発生していた検出は低減させることができたと言える．この傾向の一因としては鳥類の学習画像の一部に**図-4** に示すような遠方の小さな画像を用いていることが考えられ，学習画像からこのような画像を取り除くことで面積の小さな煙・雲の誤検出は減少すると思われる．

植物については煙・雲のような傾向は見られない．植物の学習画像の特徴として，繁茂している時期と枯れている時期で植物の様子が大きく異なることである．E，Fともに学習に用いた画像のこれらの比率に大きな偏りはないため，学習アノテーション数の不足が誤検出の一因と考えられ，学習アノテーション数の増減による誤検出の変化を検証する必要がある．

昆虫については昆虫を鳥類と誤認した検出に加え，鳥類を昆虫と誤認した検出が確認された．**表-6** にEとFについて，昆虫の検出結果を示す．**表-6** より昆虫の検出において，E，Fともに鳥類が昆虫よりも多く，Eでは約1.6倍，Fでは約1.8倍の鳥類が昆虫として誤検出されていた．また**表-4** より鳥類の検出において，昆虫の誤検出数は鳥類の検出数の25%から37%の割合を占めている．これらの結果から，鳥類と昆虫の識別は十分にできておらず，適切な学習データセットの構築について検証を進める必要がある．昆虫種の学習が不十分な原因としては，昆虫が他の誤検出物と比較して鳥類と姿形が似ていることや，昆虫種によってその姿形が大きく異なること（**図-5**）が考えられるため，鳥類と明確な違いのある画像を優先的に学習させることや，特定の種に限定した学習データをを構築し，種の特徴を鮮明にするなどの方針が考えられる．

表-6　昆虫として検出された検出物の内訳

データ名	昆虫	鳥類
E	134	225
F	168	310



図-5　昆虫の学習画像例

以上より，主な誤検出種である風車，植物，煙・雲，昆虫を学習させた学習済みモデルの検出において，COCOデータセットによる学習済みモデルと比べ鳥類の検出数を約2.4倍に増加させ，風車と煙・雲の誤検出を大きく減少させることができた．一方で植物と昆虫の誤検出の減少については，適切な学習データセットの構築のための画像の選別やアノテーションの増強を検証する必要がある．

4.　議論

響灘ビオトープではチュウヒのほかにも絶滅危惧種であるベッコウトンボの生息が確認されている．本研究で運用しているシステムはチュウヒを対象に鳥類の検出を想定している．しかし鳥類の誤検出としてトンボを検出しており，本研究で昆虫を学習させた結果，トンボを昆虫として検出できたことから（**図-6**），YOLOとビデオカメラを用いてベッコウトンボの自動検出システムを構築し，モニタリングのために運用することは十分に可能であると考えられる．ただしベッコウトンボは積極的に飛翔する種ではなく，ヨシなどの葉先に止まっている時間が多いため，植物を背景とした検出が必要になる．**図-6** より検出できたトンボの背景は多くが空となっているが，これは学習に用いたトンボの画像は鳥類の誤検出

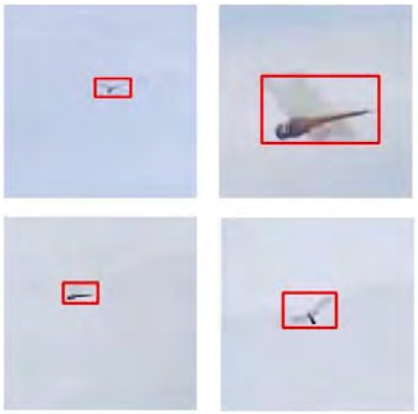


図-6　本研究で得られたトンボの検出画像例

が基となったため，学習画像の背景の多くが空になっていることや，体色が保護色となり植物を背景とした検出が困難であることが原因と考えられる．この検出結果を踏まえ，様々な背景に映るトンボ類を検出する学習済みモデルの構築とモニタリングへの適用性の評価は，今後の当地区での生物保全活動にとって有用であり，実現へ向け検証を進めることには大きな意義がある．

5.　まとめ

本研究では，響灘ビオトープを対象に運用している鳥類自動検出システムについて，誤検出が目立つ風車，植物，煙・雲，昆虫の誤検出低減のための学習済みモデル構築の検証を行った．誤検出画像を学習へ活用することで，鳥類の検出数の増加と，風車，煙・雲での誤検出を大きく減少させることができた．植物と昆虫の誤検出については，低減のための学習データセット構築の指針を得ることができ，精度向上へむけた検証を進めていく．また，トンボ類の検出が可能だったことから，響灘ビオトープで保全されているベッコウトンボのモニタリングへの本システムの応用についても検証を進める．

NOTES

注1) 環境省：環境省レッドリスト 2020，https://www.env.go.jp/press/107905.html，2020.
注2) COCOdataset, https://cocodataset.org/#home,
注3) YOLOv5, https://github.com/ultralytics/yolov5,

REFERENCES

1) Joseph, R., Santosh, D., Ross, G. and Ali, F.: You Only Look Once: Unified, Real-Time Object Detection, Computer Vision and Pattern Recognition Conference, pp.779-788, 2016.
2) Hong, S.J., Han, Y., Kim, S.Y., Lee, A.Y. and Kim, G.: Application of Deep-Learning Methods to Bird Detection Using Unmanned Aerial Vehicle Imagery, Sensors, Vol.19, No.7, pp.1651-1666, 2019.
3) Zhihao, L. and Ming, T.: A Bird Species Detection Method Based on YOLO-v5, 2021 International Conference on Neural Networks Information and Communication Engineering, https://doi.org/10.1117/12.2615310 , 2021.
4) 中山紘喜，安枝裕司，久冨学，野上敦，松本亨：深層学習法を用いた鳥類の動画自動検出追尾システムの開発と響灘ビオトープでの運用，環境共生，39 巻 1 号，pp.45-54，2023. [Nakayama, H., Yasue-da, Y., Hisadomi, M., Nogami, A. and Matsumoto, T.: Development of automatic bird video detection and track tracking system using deep learning method and its operation in Hibikinada biotope, *Journal of Human and Environmental Symbiosis*, Vol. 39, Issue 1, pp. 45-54, 2023.]

(Received June 19, 2023)
(Accepted August 21, 2023)

Improving the accuracy of real-time detection of birds, etc. from videos by the YOLO method

Hiroki NAKAYAMA, Yuji YASUEDA, Manabu HISADOMI, Atsushi NOGAMI, and Toru MATSUMOTO

In the Hibikinada Biotope, where the development of the surrounding area, such as wind power generation, continues, a rare ecosystem is conserved with the eastern marsh harrier at the top. It is also important to grasp the living behavior of birds through constant monitoring. Our research group developed a real-time automatic bird detection system using the deep learning method YOLO. By using it, we were able to detect birds with a minimum size of 156 pixels for 4K video. However, there were false detections of windmills, clouds, plants, insects, etc., and improvement of the learning model became an issue. In this study, we organized and augmented the learning data to improve the detection accuracy. By learning the images obtained in the Hibikinada biotope, the number of correct bird detections increased by about 2.4 times. On the other hand, false detections of wind turbines decreased significantly, and false detections of large clouds decreased. In addition, some insects, such as dragonflies, can be identified, and we will proceed with further verification.