

# ニューラルネットワーク及びLMMを用いた魚類の検出と分類

AquaWiz

e-kagaku 松本 慎太郎 大塚 雄日 神山 瑠星 小松 琢磨



## 1.概要と目的

近年、**地球温暖化**などの気候変動により**海洋生物の分布が変化**しつつあり、その結果として生態系に**悪影響**を与えている。

したがって、広範な海洋生物の**生息地や動態を迅速に把握**し、気候変動が生態系に及ぼす影響をより正確に評価することが必要であると考えられる。

そこで、本研究では水中の映像から自動的に魚類を検出および分類を行い、**高速かつ容易に効率的な魚類の生態系分布の調査**が実施できる手法を模索した。

具体的には、水中の映像から**深層学習ネットワーク(YOLOv4)**を用いて魚類の検出を行い、検出した魚の種類をOpenAIの**LMM(大規模マルチモーダルモデル)**を用いて識別する。

## 2.YOLOv4を用いた魚の検出と追跡

魚類の検出には**YOLOv4**を使用し、トラッキングには**カルマンフィルタ**を使用した。これらを組み合わせることで、検出およびトラッキングを**同時**に行うことに成功した。

YOLOとは、物体を短時間で検出するために開発された**物体検出のアルゴリズム**であり、画像に映った物体を検出することができる。

今回は、**CSP-DarkNet-53**ネットワークをバックボーンとし、転移学習を実施した。転移学習には約**8000枚**の画像を使用し、**ほぼすべての魚類を検出**できるようなネットワークを構築した。

## 3.LMMを用いた魚の分類

魚の分類には、OpenAIが開発した**大規模マルチモーダルモデル(LMM)**、GPT-4-Turboの**イメージキャプション機能**を使用した。

処理はOpenAIのサーバー上で行っており、**約3秒**で分類が完了する。

CNN等の手法を使用した魚類の分類では、**大量のデータセットと訓練**が必要となり、個人での開発には多くの時間と労力を要するが、LMMを用いることでこの問題を**解決**した。

## 4.手法

### YOLOv4による魚の検出

計8000枚の画像データセットを使用し、転移学習したVRAMが不足していたため、3回に分けて学習実施した

csp-darknet-53-coco

1回目  
Validation:  
Training:  
Epoch:10

2回目  
Validation:  
Training:  
Epoch:10

3回目  
Validation:  
Training:  
Epoch:10

画像データ

### オブジェクトトラッキングによる魚の追跡

#### 境界ボックス予測

KalmanFilter

前フレームの境界ボックスの幅と高さを予測サイズとし、重心を予測

現在のフレームにおける各トラックの重心

#### トラックへの検出の割り当て

bboxOverlapRatio関数  
境界ボックスからオーバーラップ率を計算

コスト行列

bboxOverlapRatio関数  
コスト行列から検出をトラックに割り当て

割り当てられたトラック  
割り当てのない検出

新しいトラックを作成

#### トラックの削除

追跡時間が短い

信頼度が低い

大部分のフレームが不可視とマークされた

これらのフレームを削除

画像データ  
境界ボックス

#### トラックの更新

割り当てられたトラック

対応する検出によって更新

KalmanFilter correct  
メソッド推定位置を訂正

直近の4つの境界ボックスの平均から新しい境界ボックスの平均を保存  
トラックの持続時間と合計可視カウント、信頼度スコアを更新

更新されたトラック

#### LMMによる魚の分類

追跡時間が長く  
信頼度が高い  
まだ分類されていないトラック

トラックから画像を切り取り、保存する

GPT4-Turbo  
保存した画像を分類し、画像ファイルの名前を分類した結果に変更

魚種名

切り抜いた画像の例(錦鯉)

ラベルと魚種名を表示

## 5.結果

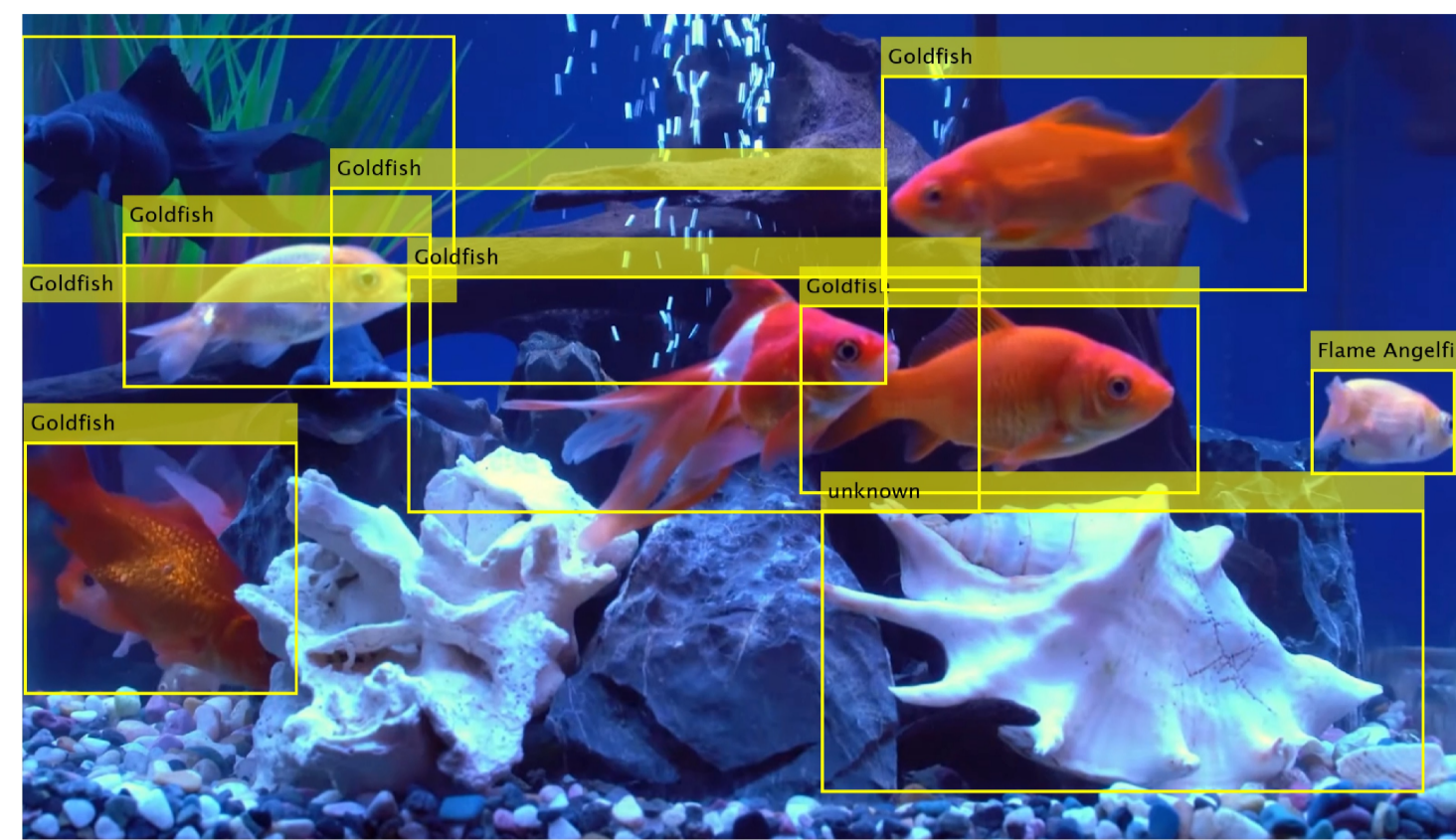


Fig.1 検出結果1

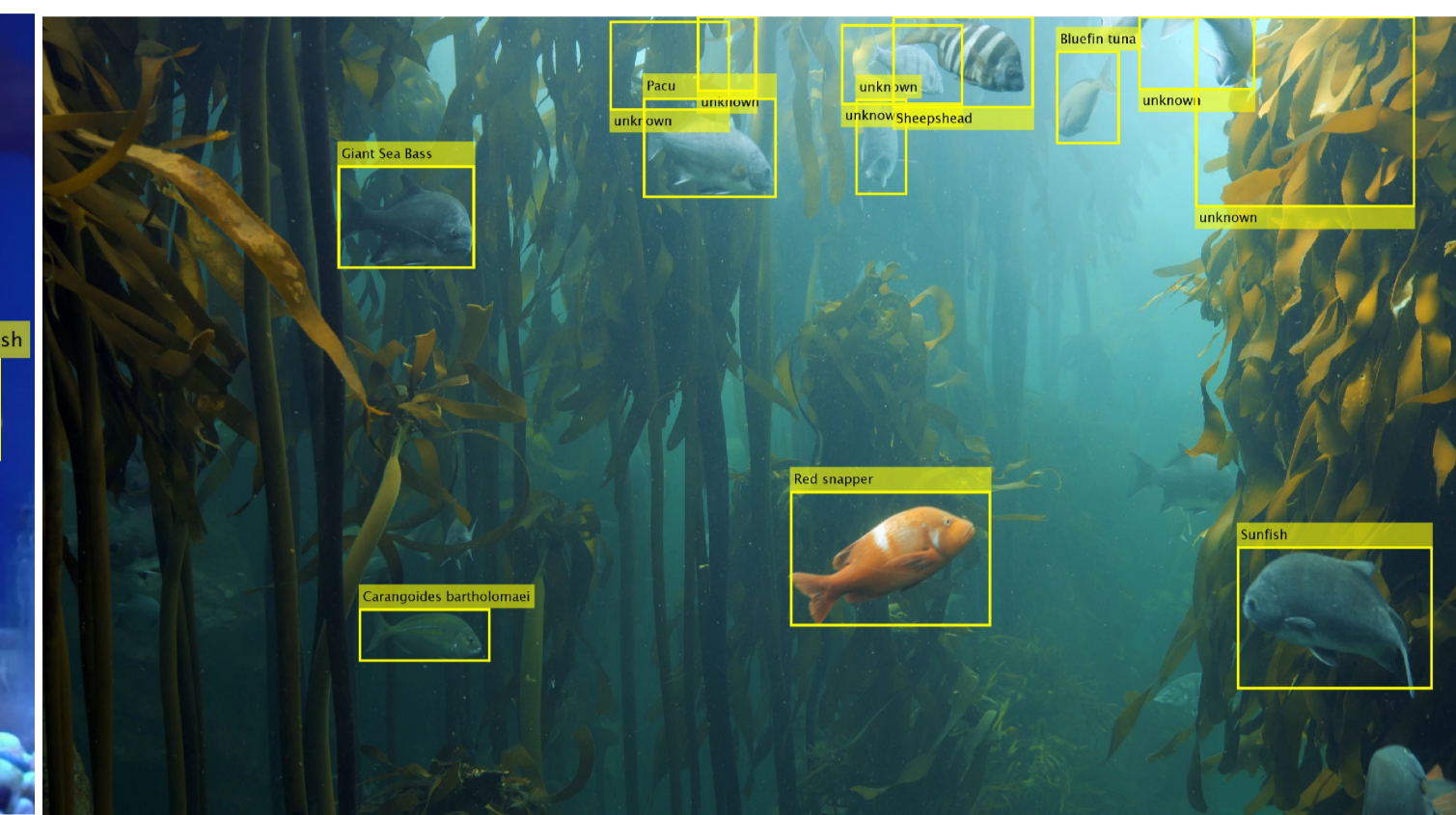


Fig.2 検出結果2

検出および分類結果は以上の画像のようになった。

魚の検出について、Fig.1, Fig.2の正解率はそれぞれ**76%**と**77%**、再現率は**共に100%**となり、**高い精度**で魚の検出を行うことに成功した。

魚の分類について、Fig.1のような比較的有名な魚類の分類の正答率は**90%以上**であったが、Fig.2のようなあまり一般的ではない魚類の分類の正答率は**5割程度**となった。

トラッキングモデルの精度(Multi-Object Tracking Accuracy; MOTA)、は**約73%**となった。

トラッキングを実際に動作させている動画は以下のQRコードより観覧可能である。精度は以下の式より算出した。(Fig.3)

$$MOTA = 1 - \frac{\sum_t (FP_t + Miss_t + IDSW_t)}{\sum_t g_t}$$

$t$ : t番目のフレーム  
 $g_t$ : t番目のフレームにおける正解データの数  
 $FP_t, Miss_t, IDSW_t$ : t番目のフレームのFP, Miss, IDSW

また、魚の検出、分類及びトラッキングを**MATLAB App Designer**を用いることで一つのアプリに統合し、**効率的に分類およびトラッキング**を行うことに成功した。



Fig.3 精度検証に使用した動画



Fig.1の動画



Fig.2の動画



Fig.3の分類を行っている動画

## 6.考察

YOLOv4による魚類の検出精度については、さらに多様なデータセットを用いた学習を行うことで**精度が向上**すると考えられる。

しかし、そのようなデータセットは少ないため、**自身で作成**する等の工夫が必要である。

オブジェクトトラッキングについては、YOLOv4による魚類の検出精度が向上すれば、オブジェクトトラッキングの精度も向上すると考えられる。

LMMによる分類については、**ファインチューニング**等を実施し、魚の分類に特化したモデルを作成することで精度の向上が期待できる。

## 7.今後の展望

今後、以下の取り組みを実施したい。

- 実際の環境調査における精度検証。
- データセットの増加や、より時間を変えて学習を行うことで、オブジェクトトラッキング及び検出の**精度の向上**。
- 魚の検出、種類の判別、魚の数がカウントできるアプリケーションを開発したため、今後アプリケーションとしての**実用性を向上**させる。

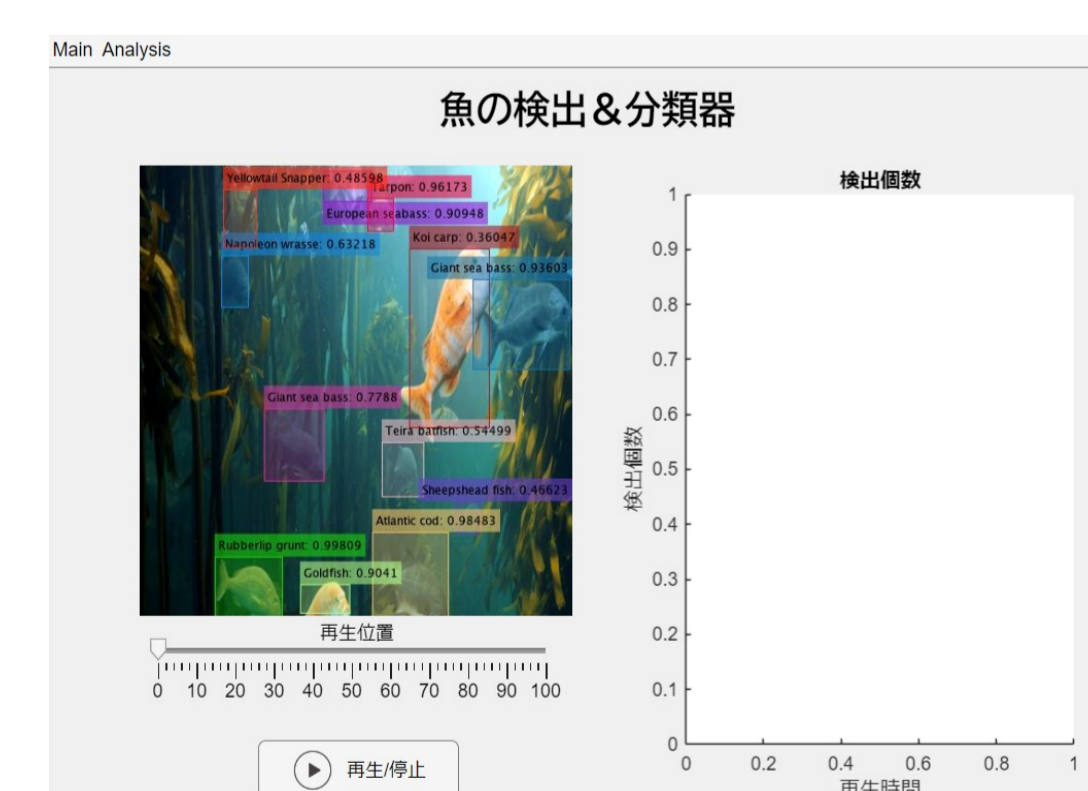


Fig.4 開発したアプリケーション

## 8.まとめ

**YOLOv4**を用いて魚類の検出を行い、**カルマンフィルタ**を用いてトラッキングを実施した結果、**非常に高い精度**で検出することに成功した。

LMMを用いることにより魚の種類を特定することが可能であったが、識別が困難な魚種もあったため、**ファインチューニング**などを実施し更なる精度向上に努めたい。

今後は今回得られた成果を元に、**水中ドローン**を使用した**リアルタイムでの解析**を行いたい。

## 9.出典

- [1]ARISE analytics Inc. Multi-Object Trackingの精度評価指標
- [2] Ditria, Ellen M; Connolly, Rod M; Jinks, Eric L; Lopez-Marcano, Sebastian (2021): Annotated video footage for automated identification and counting of fish in unconstrained marine environments [dataset]. PANGAEA, <https://doi.org/10.1594/PANGAEA.926930>
- [3] Saleh, Alzayat and Laradji, Issam H and Kononov, Dmitry A and Bradley, Michael and Vazquez, David and Sheaves, Marcus; Scientific Reports 10.1.14671 (2020): Nature Publishing Group UK London. <https://doi.org/10.1038/s41598-020-71639-x>
- [4]MathWorks 走行中の自動車からの歩行者の追跡
- [5]MathWorks YOLO v4 深層学習を使用したオブジェクトの検出
- [6]千葉県生物多様性センター研究報告 7:3-13. Feb. 2014 地球温暖化と南日本各地における魚類相の比較 須之部友基・川瀬裕司・坂井陽一・清水則雄・望岡典隆・田和篤史・竹垣毅・中村洋平・出羽慎一