

深層学習による水中映像からのマグロ成魚の領域抽出

Region Extraction of Tuna Adults from Underwater Images by Deep Learning

竹長 慎太郎† 波部 斉† 阿部 孝司† 井口 信和†
Shintaro Takenaga Hitoshi Habe Koji Abe Nobukazu Iguchi

1. はじめに

水産養殖の生産性を向上させるためには、養殖魚の状況を観察し、適切な養殖環境になるように調整することが必要になる。養殖魚の成長状況を把握する上で基本となるのは尾数の把握である。特に、クロマグロは外部からの刺激に敏感であるため、その養殖においては尾数計測が求められている。しかし、多数の魚が洋上の生簀の中で遊泳している場合に、その尾数を非侵襲・非接触で把握することは困難である。

本稿では、この最終目標に近づくための第一歩として、生簀の中で撮影した映像から遊泳しているクロマグロ成魚の領域抽出を行う。洋上の生簀の中で撮影した映像は太陽の光や影、波といった外的要因の影響を受け、検出する対象もカメラとの距離によって様々な大きさをとることが尾数推定を行う上での課題となる。本稿ではこのうち前者の問題に対応した領域抽出を行う。提案手法は深層学習によってマグロ成魚に相当する領域（以下、成魚領域）とそれ以外の領域の識別を行う。深層学習を用いることで、先に述べた外的要因による変化と成魚領域を精度よく識別できると期待される。

2. 提案手法の概要

本研究では、実際の現場で役立つ技術を目指として、領域抽出を高速、省メモリで行える SegNet[2]を元にして今回の適用環境に適した領域抽出の手法を提案する。Badrinarayanan らが[2] で使用した映像に比べてここで対象とする映像は分類するクラス数が少ないため、SegNet を元に構造を簡略化したネットワークを用いる。そのため、Badrinarayanan らの SegNet に比べて使用する学習データが少なくても成魚を精度良く検出することが可能であり、実用面でもメリットが大きいと期待される。

先行研究として、クロマグロの稚魚を検出する研究が行われている[1]。この研究は画像上で数画素の大きさしかない小さい対象を検出することを目的としているが、本研究では図2のように大きなマグロ成魚を扱うため、[1]とは異なり領域分割処理を行うことにした。

3. 深層学習を用いた領域抽出

提案手法の流れを図1に示す。提案手法では、成魚領域の抽出したい画像を入力画像（図1左）として入力する。あらかじめ学習したモデルを使い入力画像に対して成魚領域の抽出を行う（図1右）

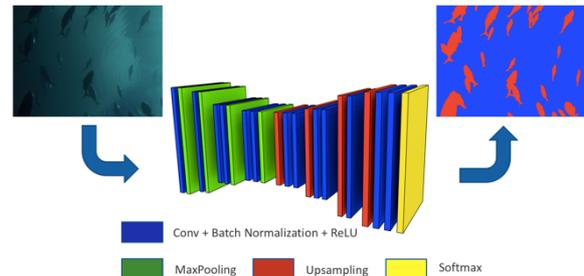


図1. 提案手法の流れ

3.1 使用するモデル

Badrinarayanan らのモデルの構造は、Encoder・Decoderではともに畳み込み層は13層になっている。今回使用するモデルのEncoderでは畳み込み層は6層、Decoderでは畳み込み層は7層で構成されており、参考にした SegNet の構造より少ない。EncoderではMaxPoolingを使用し、DecoderではUpsamplingを用いる。DecoderではMaxPoolingで採用した値の場所を記録し、upsampling時にその場所を使って特徴マップを拡大する。これにより位置情報を正確化させることができる。複合機には、各ピクセルのクラス確率を独立して生成するためにsoft-maxを使用する。

4. 実験

本章では、深層学習を用いた成魚の領域検出を行う実験の手法と結果、比較について述べる。

4.1 用いる成魚のデータ

用いる画像データは、ダイバーが生簀の中に潜り下から遊泳しているクロマグロ成魚を撮影した映像から10枚切り出して使用する（解像度は2028×2704）。学習を行う際のデータ数が10枚では少ないため、使用する10枚の画像データを4,9,16等分に分割しデータ数を増やし利用する。学習時に使用するデータ数は、計290枚となる。290枚のデータの画像サイズが異なるため、画像サイズを（解像度360×480）にリサイズする。リサイズした画像290枚を学習データ、テストデータとして232枚、58枚に分けて使用する。本研究で使用する水中映像は、ダイバーや藻を含んでいない箇所の映像である。

4.2 評価方法

あらかじめ人手によって正解領域に色を塗った画像を用意し、領域検出された画像と画素単位で比較し精度を算出する。評価値として、再現率、適合率、それらの調和平均であるF値を用いる。

4.3 実験結果



図2 実験の入力画像

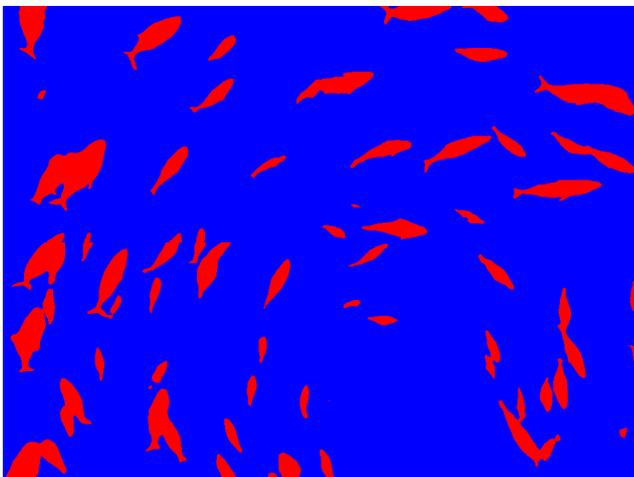


図3 実験の出力画像

Precision	Recall	F-measure
86.71%	89.01%	87.84%

表1 実験結果の精度

実験の入力画像を図2に示す。実験の出力結果を図3、実験結果の精度を表1に示す。実験結果（図3、表1）をみると、太陽の光や影、波といった外的要因があっても、高い精度で成魚領域を抽出できていることがわかる。

4.4 比較実験

4.3節の実験結果に対して次の三つの場合での比較を行う。

- ・ネットワークの層を減らす。
- ・用いる学習データ数を半分にする、
- ・学習回数を半分にする

上記の三つの比較結果を図4、図5、図6に示す。検出結果の精度を表2、表3、表4に示す。

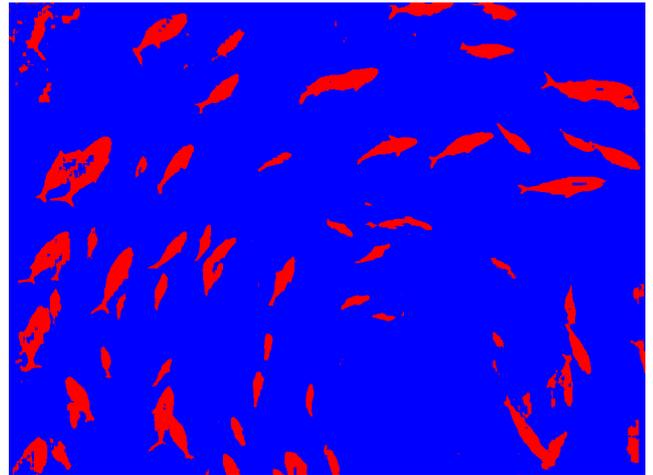


図4 ネットワークの層を減らした実行結果

Precision	Recall	F-measure
87.33%	81.3%	84.21%

表2 ネットワークの層を減らした精度評価

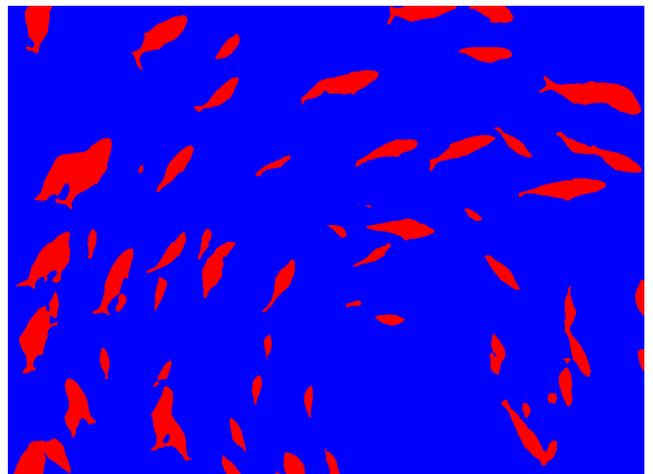


図5 用いるデータ数を半分にした実行結果

Precision	Recall	F-measure
85.70%	86.53%	86.11%

表3 用いるデータ数を半分にした精度評価

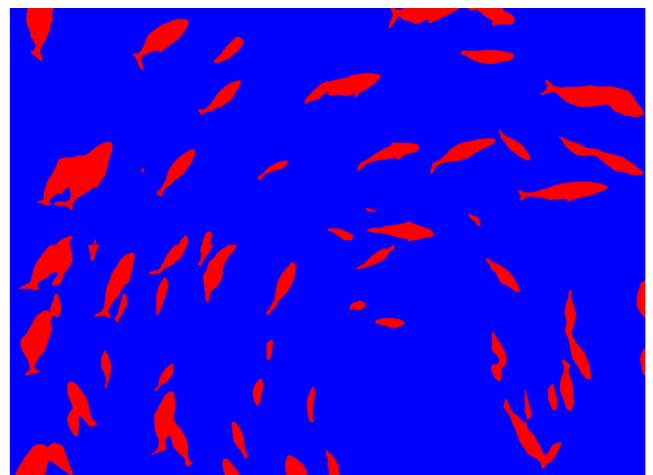


図6 学習回数を半分にした結果

Precision	Recall	F-measure
87.62%	86.18%	86.90%

表 4 学習回数を半分にした精度評価

この実験の結果をまとめると以下のようになる。

- ・図 4 は図 3 に比べ成魚領域の中が欠けている部分が顕著にみられる。表 1 と表 2 を比べると定量的な精度も落ちている。このことより、層が減ると詳細な明度パターンを示す特徴量が得られず精度が落ちる原因になっていることがわかる。
- ・図 5 も図 3 に比べ成魚領域に一部が欠けているが、図 4 とは異なり、輪郭部分が欠けている傾向が見てとれる。表 1 と表 3 を比べると精度が落ちていることもわかる。2 節で述べたように、本手法は少ない学習データでも学習できることを目指しているが、極端に少ないと精度が落ちることがこの結果からわかる。
- ・図 3 と図 6 を比較すると、ほぼ同じように成魚領域を検出できている。表 1 と表 4 を比べても差がほとんどないため、学習回数は半分にしてもあまり問題ないことがわかるが表 1 の方が精度は良い。

5. まとめ

本研究では、遊泳しているマグロ成魚の水中映像からマグロ成魚の領域抽出をする手法を提案した。検出する際に外的要因を考慮する必要があったため、ピクセル毎に推論処理を行える領域分割を用いることで外的要因を考慮し、高い精度で領域抽出を行い、期待通りの精度が得られた。

今後の課題として、マグロの尾数推定や追跡があげられる。尾数推定や追跡を行う際は、時系列データ、物体ごとのラベル付けなどが必要になると予想される。

謝辞

本研究を行う上で貴重なご意見を頂戴し、映像取得にご協力いただいた近畿大学水産研究所の升間主計所長、近畿大学水産養殖種苗センターの岡田貴彦センター長、ならびに奄美実験場の皆様に御礼を申し上げます。また、本研究の一部は JSPS 科研費 JP17H05981 の助成を受けて行った。

参考文献

- [1]. 佐藤僚太, 波部齊, 阿部孝司, 井口信和, CNN を用いた画像解析による大規模魚群の個体検出手法. SI2017, 2017.
- [2]. Vijay Badrinarayanan, Alex Kendall and Roberto Cipolla "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation." PAMI, 2017