

CNNを用いた音響画像に基づく定置網内の魚種推定の精度向上 Improvement in Accuracy of Fish Species Estimation in Set-Net Based on Sonar Image Using CNN

平間 友大¹⁾ 横山 想一郎²⁾ 山下 倫央²⁾ 川村 秀憲²⁾
Yudai Hirama Soichiro Yokoyama Tomohisa Yamashita Hidenori Kawamu
鈴木 恵二³⁾ 和田 雅昭³⁾
Keiji Suzuki Masaaki Wada

1 はじめに

2017年の定置網漁に割り当てられたメジマグロ(体重30kg以下の産卵前のクロマグロ)の漁獲枠は、全国で580トン、北海道は57トンである。しかしながら、北海道では2018年5月に大幅に割り当てを超える602トンを漁獲した[1]。産卵前の魚を獲ってしまうことは母数の減少に繋がるため、水産庁は2018年の北海道のメジマグロの漁獲枠を実質ゼロとする考えを発表した[2]。北海道の定置網漁業者らは、これまでに自主休漁や定置網の解放により漁獲制限を超えないようにしてきたが、効果は得られなかった。この原因は、網をあげるまで定置網内にどんな魚が入っているかわからない点にある。定置網漁場のサイズは円周150メートル・水深約40メートル、海中の定置網内全域を目視で把握することは難しい。また、一度網をあげて定置網内の魚を獲ってしまうと魚に傷がつき衰弱するため、漁業者が網をあげてから特定の魚種だけを逃がすといった対策を取ることができない。一方、適切な漁獲量の場合に網を開放してしまうと、大きな経済的損失を生むことになる[3]。現在、この定置網漁場に魚群探知機を設置し、リアルタイムで定置網内を観察するシステムが開発されている[4]。このシステムは2015年から現在まで稼働しており、24時間観察する際に得られる音響画像(3.4で述べる)とその漁場で獲れた魚種と漁獲量をまとめた漁獲リストを蓄積している。本研究ではこのシステムを利用し、2015年から2016年で漁獲されたマグロを対象に、音響画像から魚種を分類する推定モデルの開発を行う。この推定モデルを用いることで、網をあげることなく定置網内の魚種を推定し、定置網漁業者が漁獲管理を行うための判断基準を提供するシステムを構築することが目的である(図1)。

2 関連研究

水中の魚種を判別するアプローチとして、ビデオカメラを用いたZionらの研究[6]がある。この研究は水槽内の1匹の魚を撮影し、手動による特徴量設計を行っていた。しかし、重なり合う魚は判別できない他、自然光の影響を受けやすいことが課題としてあげられる。定置網内は水深40mほどあり、漁獲する直前は午前四時と暗く、海中が濁った場合日光が十分に届かない可能性

- 1) 北海道大学大学院情報科学院. Graduate School of Information Science and Technology, Hokkaido University, Sapporo, Hokkaido, Japan
- 2) 北海道大学大学院情報科学研究院. Faculty of Information Science and Technology, Hokkaido University, Sapporo, Hokkaido, Japan
- 3) 公立はこだて未来大学. Future University Hakodate.

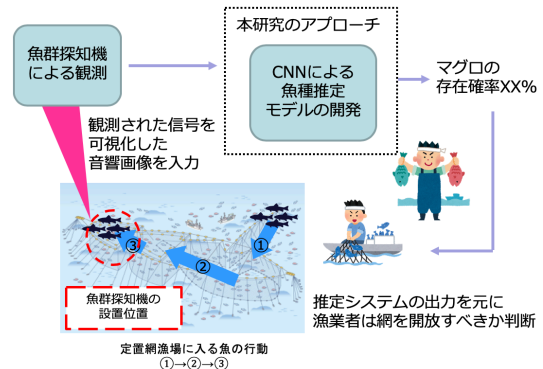


図1 提案システムの概要

がある。魚群探知機を用いた魚種判別には、Langkauら[7]やStationら[8]の研究があげられる。DIDSON[9]と呼ばれる2つの異なる周波数の超音波を用いることで、Zion[6]らが用いているビデオカメラと同様な画像を得ることができ、魚影を認識するのに特別な訓練は必要ない。しかし、DIDSONは高い周波数帯域の超音波を発射するため、音圧の減衰が低周波よりも早く、狭い範囲しか観察できない。また、高価であるため、本研究で解決するためには導入コストが高いことがデメリットとして挙げられる。

3 実験環境とデータ

本研究は、定置網内を観測するセンサとして魚群探知機を用いる。北海道函館市木直町の定置網漁業者に協力いただき、定置網内の海面上に魚群探知機を設置している。2015~16年までの毎年6~12月の漁期を対象に、音響画像と実際に漁獲した魚種・量が記された水揚げデータを用いて実験を行う。

3.1 定置網漁

定置網漁とは、主に沿岸で行われる漁法の一つである。海に網を固定し、魚の習性を利用して図2の1~3のように網の奥へと誘導し、漁獲を行う。3の網を箱網と呼び、箱網に入った魚は逃げにくい構造となっている。漁獲は通常1日1回、午前4時に行われ、箱網の魚を獲り終わると再度網をおろし、翌日まで待つ。この漁法は短い時間で漁獲を行うため燃料消費が少ない。また、定置網が魚介類の住処として機能するなど、環境に負担がかからない漁法となっている。

3.2 魚群探知機

魚群探知機は音響技術を用いた、海中を観測するための漁具である。本研究で用いる魚群探知機に搭載されたトランスデューサが、3秒に1回水中に50kHzの超音波を発射し、水と異なる密度の物体に反射した反射波を受信する。トランスデューサは、超音波の発射から受信ま

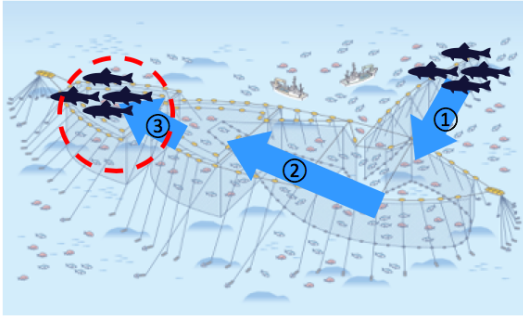


図2 定置網漁場のイメージ

での時間間隔を利用して対象物までの距離の測定できるほか、反射強度を0(弱)~255(強)の離散値で表現することで対象物の密度を測定できる。

超音波は人間の耳には聞こえない高い振動数の音波のことを指す。気体・液体・固体などの媒体中を伝搬し、減衰していく。音は均質な媒体中では直進し、音速が異なる媒体との境界面で反射する性質を持つ[10]。魚探の場合、この反射を受信し、物体との距離や密度を計測する。反射波はさらに減衰し、微弱であるため、海底に近いほど捉えにくい性質を持つ[11]。

3.3 音響画像

魚群探知機が受信した反射波の反射強度を輝度値とし、時系列に並べたものを音響画像(図3)と呼ぶ。魚群探知機から得られる信号は256階調のため、グレースケール画像が得られる。本稿では人の目で見やすいようOpenCVのカラーマップCOLORMAPJET[12]を用いて可視化しているが、実験ではグレースケールで処理を行う。魚群探知機を用いて漁業を行う漁業者は、音響画像に映る魚影の形状(図4)、季節、水温などからどのような魚種がどれくらいいるのかを経験と勘から推測している。

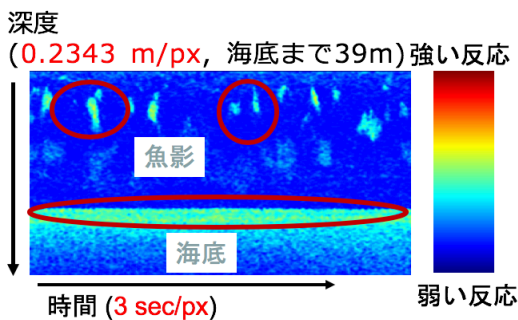


図3 45分間、水深0~30メートルの音響画像のイメージ

3.4 実験データの特性と課題

音響画像は水中を進む超音波からのみデータを得ているため、超音波の速度が変化すると同じ魚種が音響画像内に映っていても形状が異なることが分かっている。超音波は、以下の要因がその速度の減衰度合いを決定する[13]。

- 温度 T
- 塩度 S
- 圧力 D

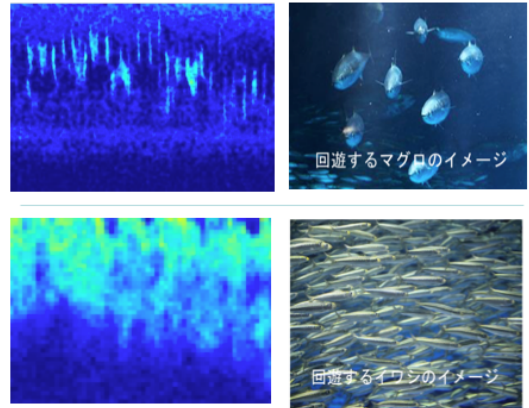


図4 25分間の音響画像に映る魚影と魚群のイメージ

この外的要因は、短い時間間隔では変化は小さいが、長い時間間隔、例えば昼と夜では温度に差があるため、音響画像への影響が想定される。外的要因による魚影の具体的な変化量は明らかにされていないが、本研究が提案するシステムの推定モデルはこの変化に頑健であることが求められる。

4 実験

本研究は、現在問題になっているメジマグロ(以下、マグロ)を対象に絞った推定モデルを構築する。そのため、マグロの魚影の有無による音響画像の2クラス分類問題を解く。音響画像から特定の魚種の有無を推定するためには、魚影の特徴量抽出を行う必要があるが、経験を積んだ漁業者でも推定が難しいことや、3.4で述べた課題を考慮すると、著者らが特徴量設計を手動で行うことは容易ではないことが予想される。そのため、本研究では自動的に特徴抽出が可能であり、高い物体認識精度を誇る畳み込みニューラルネットワーク:Convolutional Neural Network(以下、CNN)[5]を適用する。

4.1 モデルとパラメータ

CNNのモデル構造にはResnet[14]を用いる。出力層のユニット数を2に変更し、損失関数にはソフトマックスクロスエントロピーを適用した。最適化手法は、予備実験を行いAdam[15], SGD, AdaDelta[16], MomentumSGDの4手法を比較した結果、MomentumSGDを使用することとした。

4.2 教師データセットの作成

教師データとして音響画像とその画像にマグロの魚影が映っているか否かのラベルを組み合わせたデータセットを作成する。2015年から2016年までの221日分のデータを教師データの作成に使用した。現在、音響画像に映る魚影に対する正しいラベルは不明であり、経験を積んだ漁業者でも魚種の特定は難しいといった現状がある。そこで、ラベルの推定を行う必要がある。本研究では、マグロの漁獲がある日の音響画像には、マグロが映っている可能性が高いことから、漁獲量を裏付けにしたラベル付を行う。マグロの漁獲が250kg以上あった10日分の音響画像から魚影のある部分を目視で抽出して、約28時間分の音響画像を得た。

さらに、定置網に入ったマグロが定置網内を回遊する周期が約75秒である傾向から、1枚の画像に必ずマグロが含まれるように横25px・縦165pxのウィンドウを設定し、横に1pxずつ矢印の方向にスライドして分

割する(図 5)。この画像の横 25px が 75 秒に相当し、縦 165px が水面から水深約 39m の海底までの範囲を示している。このサイズは、およそ定置網内を周遊する魚介類の周期を捉えられるようにするためである。

この方法でマグロの魚影が映っている正例約 1.6 万枚を作成した。マグロの漁獲がなかった 211 日の音響画像からマグロの映っていない負例約 23 万枚をランダムに作成した。また、本論文ではカラーマップを用いて音響画像を RGB で表記しているが、CNN への入力の際はグレースケールのままとする。

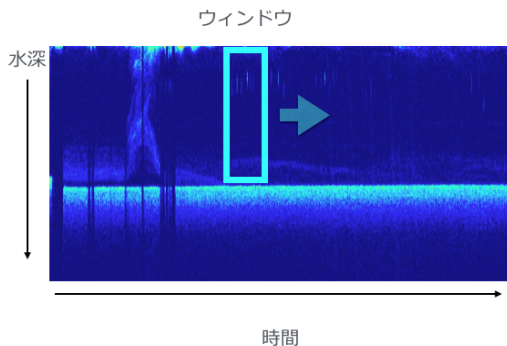


図 5 音響画像の分割 (60 分間の音響画像)

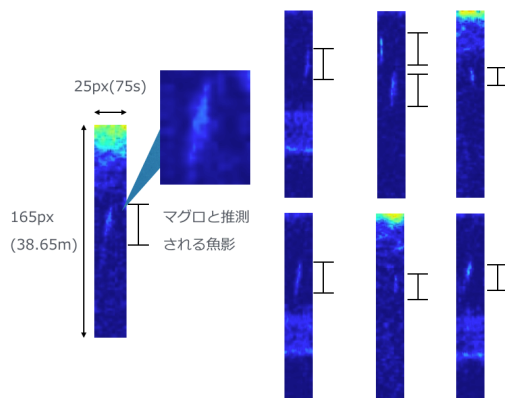


図 6 分割後の正例の音響画像

4.3 不均衡データに対する補正

機械学習における学習データの数が均等でない場合、推定モデルは学習データの多いクラスに偏った学習・推定を行ってしまう。これを防ぐための方法として

- アンダーサンプリング
- オーバーサンプリング
- Loss の重み調整

を Wyss ら [17] が提案している。アンダーサンプリングは全データからランダムサンプリングする方法であり、データのサンプルが限られている今回の研究データの状況から選択しない。オーバーサンプリングは、データの水増しによってクラス間の均衡を取る手法だが、通常用いられるいくつかの水増し手法を音響画像の適用を提案するには、実環境で起こりうる範囲内の加工に留める必要がある。画像の拡大や縮小などは魚影の大きさを変化させることにつながるため、実環境で変化する範囲内にパラメータを調整するのは困難である。Loss の重みを

表 1 混同行列

	正例	負例
正例	14103	2277
負例	935	230435

表 2 5 分割交差検証ごとの TPR, FPR, FNR, TNR

k=	TPR	FPR	FNR	TNR
0	0.863	0.005	0.995	0.137
1	0.819	0.002	0.998	0.181
2	0.878	0.003	0.997	0.122
3	0.872	0.005	0.995	0.128
4	0.869	0.005	0.995	0.131
平均	0.860	0.004	0.996	0.140

調整は、学習時の Loss に対して重み付けすることで角層の重みの更新幅を調整し、不均衡データに対する偏った推定モデルの学習を防ぐ手法である。本研究では、式 (1)、式 (2) を用いて、重み値を設定する。

$$\text{負例にかかる重み } w1 = \frac{\text{正例データ数}}{\text{正例データ数} + \text{負例データ数}} \quad (1)$$

$$\text{正例にかかる重み } w2 = \frac{\text{負例データ数}}{\text{正例データ数} + \text{負例データ数}} \quad (2)$$

5 結果と考察

音響画像というデータの特性を考慮し、CNN への適用の課題を明らかにするため、データのサンプリング方法を比較した以下の実験を行い、その結果について考察し、推定モデルの検証をした。

5.1 5 分割交差検証

音響画像という信号情報の集合によって表される画像を CNN に適用した際の学習・推定能力の検証を行う。

本実験データの学習・テストデータのサンプリングが独立同分布となるよう、全データからサンプリングを行う。これは、現在持っている音響画像が未知データの分布を十分に満たしていることを前提とした実験であるため、同じ日付の音響画像を学習とテストで分割する。ただし、スライドさせて分割しているため、重なりあった音響画像が学習・テストデータセットに含まれないよう分割する。

交差検証結果の TPR, FPR, FNR, TNR を表 2 に示す。各分割ごとの混同行列の合計値を表 1 に示す。表 2 からマグロの TPR は 0.8 を超えており、FPR は 0.02 となった。この結果は、同日の音響画像を学習とテストデータセットで分割しているため、未知データに対する評価はできない。しかし、十分なサンプルがあることを条件に、CNN の学習性能を検証することができた。

5.2 10 分割交差検証

外的要因による超音波の減衰により、音響画像が持つ特徴が同じ魚種においても異なる特性がある。そこで、この特性に対する CNN の汎化性能の検証するため、正例クラスにおいて抽出できた 10 日間の音響画像について、日付で分割しサンプリングする Leave one out 交差検証 (10 交差検証) を行う。9 日間を学習、残り 1 日をテスト日とする 10 パターンのデータセットを作成した。

表3 各テスト日を推定した際の精度

テスト日	画像枚数	TPR	FPR	TNR	FNR
2015/9/26	1961	0.614	0.019	0.981	0.386
2015/9/30	1131	0.381	0.006	0.994	0.619
2015/10/1	392	0.635	0.003	0.997	0.365
2015/10/3	5163	0.838	0.006	0.994	0.162
2015/10/4	2820	0.716	0.009	0.991	0.284
2015/10/5	142	0	0.002	0.998	1
2015/10/26	1099	0.591	0.008	0.992	0.409
2016/6/23	182	0	0.001	0.999	1
2016/10/5	3006	0.404	0.013	0.987	0.596
2016/10/7	489	0.045	0.001	0.999	0.955

テストの各日ごとのTPR, FPR, TNR, FNRを表3に示す。負例クラス分類精度を示すTNRの値は各データセットで平均99%を超えているが、正例クラスの精度指標であるTPRは、テスト日が異なることによってその精度にばらつきがあることがわかった。これは、日付が異なることによる外的要因の変化が魚影に影響しており、学習データセットと魚影の形状が異なるため推定できていないと考えられる。また、前実験で5分割交差検証した際にテストデータとなった各10日間の推定精度と、テストデータ枚数を図7に示す。この結果から、画像枚数の少ない日付において精度が低くなる傾向があり、少ない日付において比較的精度の高い2015-10-5の音響画像は全体の分布に近いことを示唆している。

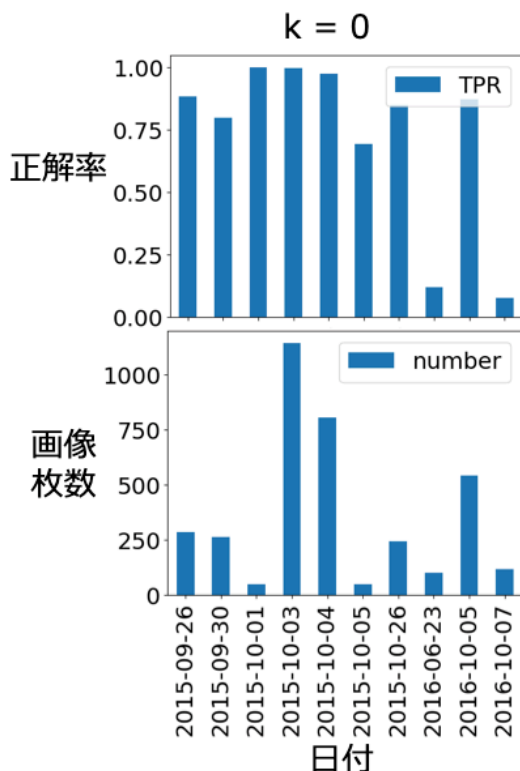


図7 5分割交差検証実験のk=0における正例のテストデータ精度と推定枚数

6 まとめと今後の展望

魚群探知機から得られる音響画像に対する、CNNを用いた推定モデルの検証を行った。実験の結果、十分なデータのサンプルがある場合、CNNを用いて音響画像からマグロの特徴量を学習し、推定できることを示した。ただし、漁獲日の異なる音響画像に対する汎化性能には改善の余地があった。未知データが学習データの分布と異なる場合、CNNはその認識精度を保証しない。そのため、汎化性能の向上にはデータの水増しを行うなど、学習データの分布を広げるための工夫が必要である。

参考文献

- [1] 水産庁:太平洋クロマグロの漁獲状況について, <http://www.jfa.maff.go.jp/j/tuna/maguro/gyogyo/gyokakujoukyou.html>.
- [2] 毎日新聞:クロマグロ来季の小型稚北海道と鹿児島は実質ゼロに, <https://mainichi.jp/articles/20180320/k00/00m/020/115000c>.
- [3] NHK:“逃がせ!クロマグロ”, https://www3.nhk.or.jp/news/web_tokushu/2018_0424.html.
- [4] R. Saville and K. Hatanaka: ICT application of real-time monitoring and estimation system for set-net fishery, OCEANS 2015 MTS/IEEE Washington, pp. 1–5 (2015).
- [5] A. Krizhevsky and G. E. Hinton: ImageNet Classification with Deep Convolutional Neural Networks, Advances in neural information processing systems, pp. 1–9 (2012).
- [6] B. Zion, V. Alchanatis, V. Ostrovsky, A. Barki and I. Karplus: “Real-time underwater sorting of edible fish species”, Computers and Electronics in Agriculture **56**, pp. 34–45 (2007).
- [7] M. C. Langkau: “Can acoustic shadows identify fish species? A novel application of imaging sonar data”, Fisheries Management and Ecology, pp. 313–322 (2012).
- [8] P. B. Station, P. Eumetrics and N. Bay: Classifying Sonar Image: Can a Computer-Driven Process Identify Eels?, North American Journal of Fisheries Management, pp. 1877–1886 (2008).
- [9] S. METRICS: “See more clearly know more definitively”, <http://www.soundmetrics.com/Products/DIDSON-Sonars>.
- [10] 本多電子株式会社: “超音波の正体<超音波ってなんだろう?>”, https://www.honda-el.co.jp/hb/1_1.html.
- [11] F. Style: “微弱な魚群からの反射波”, <https://www.furunostyle.jp/jp/mechanism/page10.html>.
- [12] OpenCV: “Colormaps”, <https://docs.opencv.org/2.4/modules/contrib/doc/facerec/colormaps.html>.
- [13] 間庭 愛信: “水中音の特性”, 2, pp. 61–64.
- [14] G. A. Hembury, V. V. Borovkov, J. M. Lintuluoto and Y. Inoue: An AcidBase Controlled Molecular Switch. Synanti Conformational Switching in a μ -oxo Bis (Iron Porphyrin), Chemistry Letters, pp. 428–429 (2003).
- [15] R. Shindjalova, K. Prodanova and V. Svechtarov: “Modeling data for tilted implants in grafted with bio-oss maxillary sinuses using logistic regression”, American Institute of Physics Conference Proceedings, pp. 58–62 (2014).
- [16] M. D. Zeiler: ADADELTA: An Adaptive Learning Rate Method, arXiv preprint, arXiv:1212.5701 (2012).
- [17] A. Wyss, M. Schamel, A. S. Sologubenko, R. Denk, M. Hohage, P. Zeppenfeld and R. Spolenak: Reflectance anisotropy spectroscopy as a tool for mechanical characterization of metallic thin films, Journal of Physics D: Applied Physics, pp. 0–5 (2015).