

3 畳込みニューラルネットワークを用いた魚群探知機による魚種推定



平間友大 横山想一郎 山下倫央 川村秀憲 | 北海道大学大学院情報科学研究科

定置網漁の問題と解決すべき課題

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。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。

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、網起こし(漁獲)をするまで定置網内にどんな魚が入っているか分からない点にある。定置網漁場のサイズは円周150m・水深39mと大きいため、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、一度網起こしをし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、漁業者が特定の魚種だけを逃がすといった対策をとることができない。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。そこで、網起こしをすることなく定置網内の魚種を推定する技術が強く求められている。

魚群探知機と音響画像

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させていただいた。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での時間間隔を利用して対象物までの距離を測定できるほか、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。反射波の反射強度を輝度値として捉え、時系列に並べたものを音響画像と呼ぶ(図-1)。この音響画像の精度は、3秒に1回の時間分解能と0.23mの空間分解能で示されるため、定置網内のすべての環境の情報を表現できていない点に注意する必要がある。

本研究では、2015年から2018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保有している。

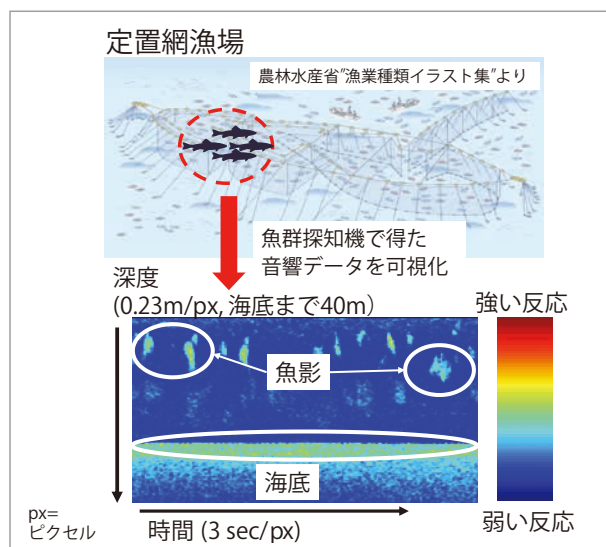


図-1 魚群探知機による定置網漁場の可視化

魚種推定モデルの構築

問題設定

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から、筆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込みニューラルネットワーク（Convolutional Neural Network, 以下, CNN）¹⁾を適用する。以下に、推定モデルの構築に向けた開発環境、教師データの作成、CNNの構成とパラメータ設定、モデルの推定結果、クラス推定に重要な領域の可視化、誤推定の原因調査に関する説明を行う。

開発環境について

本研究の開発環境として、ディープラーニングの

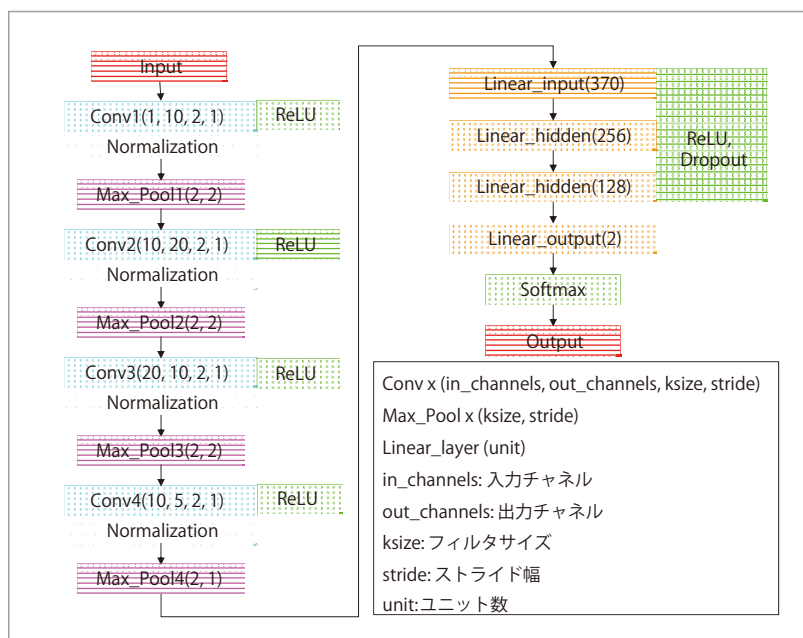
フレームワーク Chainer²⁾ を利用している。ほかにも TensorFlow, Caffe などさまざまな種類が開発されており、学習結果やパラメータを自動で可視化する機能を備えている。Chainer は日本企業である（株）Preferred Networks が開発を進めている。Chainer に関する日本語コミュニティが活発に活動しており、設定方法や使用方法を質問しやすい状況が整っているため、開発環境として採用した。モデルの学習計算には GPU NVIDIA TITAN X を使用している。

教師データの作成

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。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。

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25ピクセル・縦165ピクセルで横に1ピクセルずつスライドしてセグメント化した。この画像の横25ピクセルが75秒に相当し、縦165ピクセルが水面から水深約39mの海底までの範囲を示している。この方法でマグロの魚影が映っている正例約3万枚を作成し、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3万枚をランダムに作成した。



■図-2 CNNモデルの構造

4CNNの構成とパラメータ設定

使用したモデルを図-2に示す。

畳込み層ではパディングを行わないため、入出力画像の縦横幅が保存される。

本稿で扱うタスクに最適なCNNの構造を探索するため、畳込み層・プーリング層を各3～5層、全結合層を4～5層、隠れ層のユニット数を128, 256, 512, 1,024とそれぞれ変更したモデルを7パターン作成し、学習を行ってその推定精度を比較する。最適化手法についても、Chainerで実装されているAdam, SGD, RMSpropGraves, AdaDeltaの4手法を比較する。

その結果、最も精度が高い図-2の構造を用い、最適化手法としてAdaDeltaを使用することとした。

モデルの推定結果

モデルの汎化性能を検証するため、マグロの教師データ10日分から得られる音響画像を日付により10種類に分割して正例とし、10交差検証を実施する。正例の画像枚数と同数をマグロの漁獲がない日からランダムで抽出し、負例とした。その結果、2015年10月1日の音響画像をテストデータとした際の、マグロの魚影を含む画像に対する推定精度が100%である一方、2015年9月26日については24%と、日にちの異なる音響画像の推定に大きな差が見られた(表-1)。

クラス推定に重要な領域の可視化

CNNが誤推定した傾向を明らかにするため、ク

■表-1 マグロ漁獲日ごとの推定精度

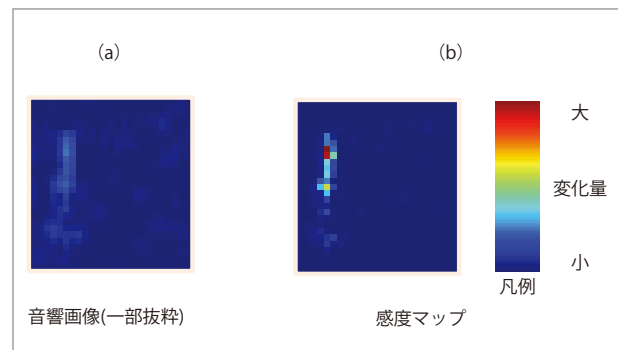
| 漁獲日 | 正例枚数 | 漁獲量(kg) | | 推定精度 | |
|-------------|------|---------|------|-------|-------|
| | | マグロ | その他 | マグロあり | マグロなし |
| 2015年9月26日 | 4387 | 949 | 2274 | 24% | 94% |
| 2015年9月30日 | 2564 | 432 | 866 | 70% | 88% |
| 2015年10月1日 | 652 | 257 | 1543 | 100% | 95% |
| 2015年10月3日 | 8240 | 2685 | 1206 | 91% | 92% |
| 2015年10月4日 | 7118 | 681 | 1520 | 75% | 90% |
| 2015年10月5日 | 280 | 300 | 1150 | 81% | 85% |
| 2015年10月26日 | 3200 | 1436 | 2835 | 36% | 91% |
| 2016年6月23日 | 762 | 474 | 184 | 26% | 84% |
| 2016年10月5日 | 5098 | 3192 | 303 | 88% | 85% |
| 2016年10月7日 | 1294 | 230 | 1815 | 34% | 92% |

ラス推定に重要な音響画像の領域を特定し可視化する手法としてSmoothgrad³⁾を適用する。重要な領域とは、画素値が変化することでCNNの出力値が大きく変化する領域である。Smoothgradは、入力画像にガウシアンノイズを加えてCNNの出力値が変化する量から、重要な領域を捉える手法である。この出力値の変化量を色として捉えて可視化した画像を感度マップと呼ぶ。図-3(a)は一部抜粋した音響画像であり、左に映る白い領域がマグロの魚影である。図-3(b)は図-3(a)にSmoothgradを適用した感度マップである。図-3(b)ではマグロの魚影に対応する領域が赤く示されており、マグロの魚影がマグロの有無の推定に重要であることが確認できる。

誤推定の原因調査

CNNがマグロありと推定した際の音響画像の重要な領域をSmoothgradを用いて可視化し、音響画像に映るマグロの魚影部分と一致しているかを確認した。

マグロありと推定された音響画像に対する重要な領域の感度マップを図-4に示す。図-4(a)はマグロの魚影だけでなく、海面上のノイズや定置網の影も重要な領域であることを示している。図-4(b)は、マグロの魚影は映っておらず、海面上のノイズと定置網の影を重要な領域であると示している。このことから、マグロの魚影がなくとも海面上のノイズと定置網の影が映っている場合、CNNはマグロ



■図-3 Smoothgradの適用例



ありと誤推定していることが分かる。

これは、教師データの正例画像では、マグロの魚影とともに、海面上のノイズや定置網の影が映る頻度が高いため、マグロありと推定するよう学習してしまったと考えられる。

今後の展開

誤推定の原因調査から、教師データの正例に偏りがあったことが問題として挙げられる。教師データ作成の際、海面上のノイズや定置網の影を除去し、マグロの魚影のみを含む音響画像も加えて偏りを解消する。また、負例に関してもランダムに教師データを抽出しているため、同様の問題を含んでいると考えられる。今後はデータの偏りに注意して設計していく。

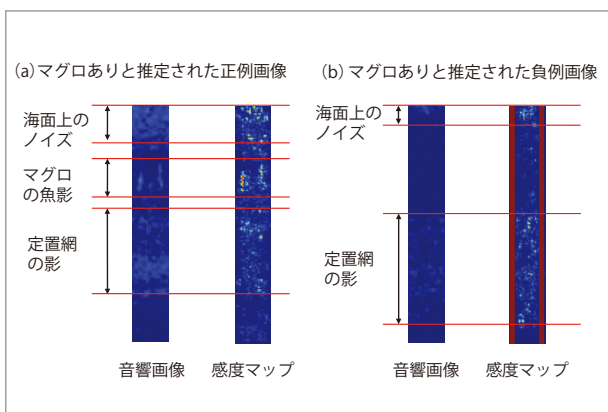
現状の推定モデルを用いて現実的なニーズに対応することを目指す。これまでは75秒間でセグメントされた1枚の音響画像からマグロの有無を推定するモデルを提案した。しかし、漁業者が実際に必要としているのは、網起こしをする直前に網の中にマグロがいるか、いないかという推定結果である。現状のマグロの有無の推定では、網起こしをする直前

に得られた音響画像に対してマグロがいるという推定結果が出たとしても、誤推定の場合には網の中にマグロがないことがあり得る。

網起こしをする直前の最終的なマグロの有無を高精度で推定するために、24時間分の推定結果を統合するモデルの開発を進める。そのモデルとしては、現状の推定器の24時間分の出力(1,152回分の推定結果)を入力とするニューラルネットワークや、現状の推定器の出力を時系列順に入力していく再帰型ニューラルネットワークの利用を予定している。

参考文献

- 1) Krizhevsky, A., Sutskever, I. and Hinton, G. E. : Imagenet Classification with Deep Convolutional Neural Networks, Advances in Neural Information Processing Systems (2012).
- 2) Tokui, S. et al. : Chainer : A Next-generation Open Source Framework for Deep Learning, Proceedings of Workshop on Machine Learning Systems (LearningSys) in the Twentieth Annual Conference on Neural Information Processing Systems (NIPS), Vol.5 (2015).
- 3) Smilkov, D., et al. : Smoothgrad : Removing Noise by Adding Noise, arXiv preprint arXiv:1706.03825 (2017).
(2018年8月1日受付)



■ 図-4 クラス推定に重要な領域の可視化

■ 平間友大 (学生会員) yhirama@complex.ist.hokudai.ac.jp

2016年公立はこだて未来大学システム情報学部情報アーキテクチャ学科卒業。同年北海道大学大学院情報科学研究科情報理工学専攻博士前期課程入学。

■ 横山想一郎 (正会員) yokoyama@complex.ist.hokudai.ac.jp

2016年北海道大学大学院情報科学研究科情報理工学専攻博士後期課程期間短縮修了。同年日本学術振興会特別研究員(PD)。2017年同大助教となり現在に至る。

■ 山下倫央 (正会員) tomohisa@complex.ist.hokudai.ac.jp

2002年北海道大学大学院工学研究科システム情報工学専攻博士後期課程期間短縮修了。2016年国立研究開発法人産業技術総合研究所人工知能研究センター主任研究員を経て、2017年より北海道大学大学院情報科学研究科准教授。

■ 川村秀憲 (正会員) kawamura@complex.ist.hokudai.ac.jp

2000年北海道大学大学院工学研究科システム情報工学専攻博士後期課程期間短縮修了。同年同大助手。2006年同大准教授、2016年同大教授となり現在に至る。