

CNN を用いた魚群探知機による魚種判別 Discriminating Fish Species by Fish Finder in a Set Net using CNN

平間 友大† 横山 想一郎† 山下 倫央† 川村 秀憲† 鈴木 恵二‡ 和田 雅昭‡
北海道大学大学院情報科学研究科† 公立はこだて未来大学‡

1 概要

現在の漁業は、魚種の判別や漁獲量の予測をすることは難しく、漁業者の経験に基づいて行われているのが実状である。そこで、乱獲の防止や水産資源の管理という点からみても、正確に海中の様子を把握する必要があると考えられる。しかし現在のところ、コンピュータが魚種を判別するシステムは確立されておらず、魚群探知機による魚影の把握のみにとどまっている。そこで、コンピュータによる自動判別が可能になれば、漁業のみならず、資源保護にも役立てることができると思われる。本研究の目的は、定置網漁場に設置された魚群探知機の音響データから、魚種判別や漁獲量の推定を行うための手法を提案し、漁業を支援することである。

2 定置網漁業の魚群探知機と音響データ

定置網漁場に設置された魚群探知機からデータをサーバーに送信するシステムが開発されている [1]。この定置網の外周は約 150m であり、魚群探知機は網上に設置されている。

この魚群探知機は 3 秒ごとに音波を発射し、その時刻と受信した反射の強度を含めたデータを 24 時間サーバーに蓄積している。海中の様子を深さ 0.3125m 間隔で捉え、反応強度を 256 階調で表す。音響データはこの反応強度と音波発射時刻をまとめたものである。音響データを画像化 (音響画像) する際、反応強度を画素値と捉え、255 を白色、0 を黒色のグレースケールとした。この際に取れた魚種と漁獲量 (漁獲リスト) が記録され、現状では 2015 年 9~12 月の 93 日分ある。漁獲リストにある魚種は、必ず 1 度はこの音響画像内に映っている。この期間内に観測された魚種は、36 種類である。

漁獲リストを参考にすると 1 日に平均 12 魚種が漁獲されており、各魚種が混ざって定置網内にあることが予想される。音響画像から魚種を特定する際、複数の

魚種が映っていることを踏まえ、システムを構成する必要がある。

3 提案手法

本手法では DeepLearning 手法の一つ、CNN (Convolutional neural network) を用いる。音響画像を入力とし、その画像に映っている魚種を出力とする教師あり学習を行い、未知の画像に対して魚種判別を行う。また、各最適化手法ごとに精度を比較し、どの手法が有用か明らかにする。

3.1 教師データの作成

漁獲リストを用いて、音響データにラベル付けを行う。93 日分の漁獲リストを分析すると、1 日に複数の魚種が漁獲されていることが分かっている。本実験では問題をシンプルにするために、1 回の魚種判別で 1 魚種を判別する CNN を作成する。1 魚種を教師データとして抽出するために、漁獲リストから 1 日の漁獲量の中で、1 魚種の占める漁獲量割合が最大の日を 1 日選択する。さらに、魚群探知機が同魚種を捉えている可能性の高い 0 時から網起こしの午前 4 時までを入力データとして用いる。本研究では、漁獲量の多かったサケ、ブリ、イワシ、イカ、マグロの 5 魚種を判別対象とした。

3.2 音響画像のセグメント化

0 時から午前 4 時までの音響画像は、横 4600 × 縦 125 ピクセルとなる。入力データとして扱うために、音響画像をセグメント化する必要がある。本実験では、横 25 × 縦 125 ピクセルに加工した。現実の情報に例えると 75 秒間、水深約 39 メートルの範囲の海中の様子を捉えたものとなる。これを 1 枚の画像とし、1 つのラベル付けをする (図 1)。セグメントのスライド幅は 1 ピクセルであり、重なるように行っている。また、音響データにエラーがある時間帯は画像化しない。このような方法で生成した画像枚数は、サケ 3799 枚、ブリ 4600 枚、イワシ 4600 枚、イカ 4600 枚、マグロ 4599 枚の 22198 枚である。このうち 17758 枚を学習データとし、残り 4440 枚をテストデータとする。

†Yudai Hirama †Soichiro Yokoyama †Tomohisa Yamashita
†Hidenori Kawamura †Keiji Suzuki †Masaaki Wada
‡Graduate School of Information Science and Technology, Hokkaido University ‡Future University Hakodate

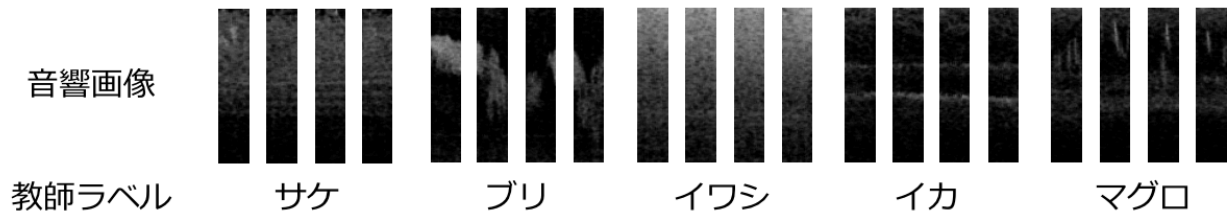


図 1: 画像のセグメント化

4 実験

4.1 CNN の構造とパラメータ設定

CNN の構築と学習には、フレームワーク「Chainer」を使用した。モデル構造には通常の CNN を用い、畳み込み層 2 つ、プーリング層 2 つである。フィルタサイズは畳み込み層がどちらも 2×2 、プーリング層 1 が 2×2 でスライド幅 2、プーリング層 2 が 3×3 でスライド幅 2 である。どの層もパディングは行っていない。全結合層は 4 つからなり、ユニット数は入力層 1500、512、208、出力層 5 である。Dropout 率は 0.5、活性化関数 (Relu) を用い、出力層はソフトマックス関数を用いて判別した。学習バッチサイズ 13、テストバッチサイズ 6、Epoch 数 20 で学習とテストを行った。

4.2 結果

最適化手法ごとの精度の比較検証を行うため、学習とテストを 20 回ずつ行い、平均と標準偏差を求めた (表 1)。さらに最も平均精度の高い最適化手法を用いて学習後の CNN をモデル化し、テストデータ 4440 枚をこのモデルに判別させた時の F 値を (表 2) に示す。各数値は小数点第 5 位以下切り捨てである。

4.3 考察と今後の展望

(表 1) から、RMSpropGraves 手法が音響画像を用いた魚種判別に最も有用であることが分かった。しかし、実験は問題をシンプルにするために実データの中でも特殊なデータを抽出して行っている。つまり、音

響画像に判別対象魚種が 1 魚種であることを前提としている。今後の研究においては、複数魚種を検出、判別できるシステムが求められるため、評価手法を検討したい。

5 まとめ

本研究では、定置網内に設置された魚群探知機のデータから CNN を用いて魚種判別する手法を提案した。音響データを画像化し、セグメントして学習に用いることで、約 99 % の正答率を得た。また、判別には RMSpropGraves 手法が有用であることも分かった。

謝辞

本研究は、公益財団法人北海道科学技術総合振興センター「地域産学官 AI/IoT 実証モデル委託事業」の支援により実施しています。ここに記して謝意を表します。

参考文献

- [1] Ramadhona Saville, Katsumori Hatanaka, Masaaki Wada, “ICT application of real-time monitoring and estimation system for set-net fishery,” Proceedings of OCEANS’15 MTS/IEEE Washington, 2015, 5pages.

表 1: 最適化手法ごとの精度検証

	Adam	SGD	RMSpropGraves	RMSprop	AdaDelta	AdaGrad	MomentumSGD	NesterovAG
平均精度	0.9811	0.9768	0.9903	0.4143	0.9848	0.9892	0.7679	0.7412
標準偏差	0.0160	0.0240	0.0110	0.3103	0.0149	0.0092	0.2992	0.3240

表 2: RMSpropGraves 手法モデルの 5 種の F 値

種類	サケ	ブリ	イワシ	イカ	マグロ
F 値	0.9786	0.9988	0.9828	0.9983	0.9983