

# ディープラーニングを利用した水中音による船舶分類

木原 将吾

OKIは、水中音に関わる水中音響システムを研究開発、設計、製造し、音を使用してどこに船舶や海中生物が存在しそうかなどの情報をお客様に提供してきた<sup>1),2)</sup>。それに加えて現在、その音が何の音なのかの分類に取り組んでいる。本稿では、水中音を利用した船舶分類にAI技術を応用した最近の取組みについて述べる。

## はじめに

水中では、電波は減衰が大きく、光波は散乱が大きいため、遠方までエネルギーが伝達される音波が船舶の探知・分類に用いられている。OKIは長年、水中マイクで受波した音がどんな特徴をもっているかを確認するシステムを開発してきた。従来はそのシステムを通して、音から得られる特徴を可視化した画像（以降、時間-周波数分析画像と表記する）で強い応答が出ている周波数やその組合せなどを見て、経験を積んだ熟練者が、その音がどのような船の音なのかを分類していた。

図1はソーナーシステムで用いられる時間-周波数分析画像のイメージ図である。この図は音波のスペクトルの時間変化を二次元画像として可視化したものである。

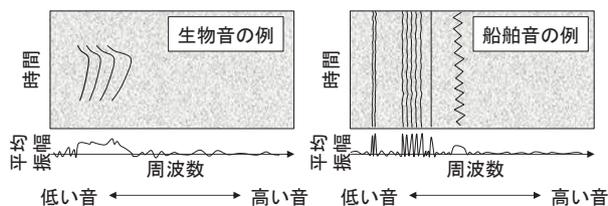


図1 時間-周波数分析画像のイメージ図

しかし、近年省人化の需要が高まり、水中音の分類も自動化が求められている。我々が目指していることは熟練者に代わり、受波した水中音が何の音でありそうなのかを自動的に認識することである。そのために画像認識などで注目されているディープラーニングモデルを利用した水中音分類手法の開発を進めている。

## 水中音から船舶を分類する手法

図2は、我々が開発を進めている水中音分類手法の構成である。この手法は、水中マイク、前処理部、特徴抽出部、分類部から構成される。水中マイクで受波した水中音は、分類したい船舶の音のほかに、周囲の生物や環境から放射される雑音も多く拾ってしまっている。

前処理部では、受波した音から分類したい船舶の音の特徴を強調する処理を行う。音の特徴は、音の高さを表す周波数や音の持続時間を使って表現するのが一般的である。例えば生物の音は、図1左のように持続時間が短く、周波数も広い幅を持つような特性になることが多い。一方で船舶の音は、図1右のように持続時間が長く、周波数も幅の狭い複数の高さの音の組合せといった傾向があり、この特性の差異を利用することで、分類したい船舶の音を強調したものが時間-周波数分析画像として表示される。

次に、図1のような時間-周波数分析画像から強く表れている周波数はどのあたりで、それらはどのような組合せで現れているかなどを特徴として認識し、その特徴がどの船の音であるかを、過去に分類済みの船名つきの時間周波数画像と見比べて判断するが、従来は経験を積んだ熟練者が実施していたこの特徴抽出と分類を、図2の特徴抽出部と分類部でAI技術の一つであるディープラーニングを使って実現した。

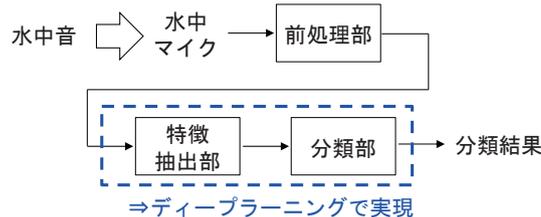


図2 水中音分類手法の構成の概要

## ディープラーニングの適用

ディープラーニングとは、近年社会実装が進められているAI技術であり、人間の脳神経の働き方を模倣したニューラルネットワークと呼ばれるシステムをベースとしている。

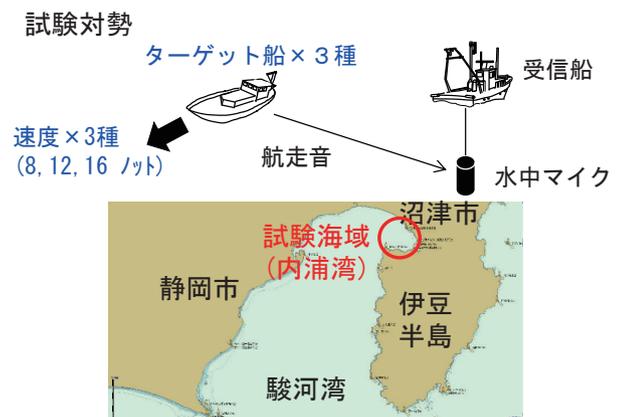
大量のデータとその正解情報(以降ラベルと表記する)のペアを使って何度も学習することで、最終的には、新しい未知のデータに対しても正確に判断できるようになる。運用時には、この学習済みモデルを使って新たに取得したデータを分類する。

ディープラーニングの中でも、入力された画像が何であるかを認識する「画像認識」と呼ばれるタスクの研究が特に盛んに行われている。従来熟練者が時間-周波数分析画像を見て分類していたことから、我々の取組みも「画像認識」タスクと捉えることができ、これまでに成功が確認されている画像認識用のディープラーニングモデルを活用して水中音分類に取り組んでいる。

次に、我々の水中音分類システムを実際に収録したデータで評価した結果について説明する。

### 実験① ディープラーニングの適用結果

評価に使う船舶音のデータは、OKIコムエコースの協力のもと、図3に示すように、静岡県内浦湾で、3船種の漁船の航走音を水中マイクまでの距離や船の速度を変えて収録したデータであり、今回3種類の船種を分類する問題を考える。



※日本水路協会発行 航海用電子参考図 NP01「東京湾及び周辺」複製

図3 試験収録の概要

上記のように収録された船舶音データに対し、学習用の906個のデータ全てをディープラーニングモデルの学習に使用した場合と、その3割にあたる270個のデータのみ、及び全データの1割にあたる90個のデータのみをそれぞれ学習に使用した場合に対し、学習データに含まれない273個のテストデータを使って、ディープラーニングモデルがどれだけ正しく分類できたかを評価した結果が表1である。ディープラーニングの適用により9割程度の分類精度

が得られているが、学習に使用できるデータ数が270、90と少量になるほど精度が劣化していく。

一般の動物や風景などの画像データセットと比較して、公開されている水中船舶音データは限定されている。さらに、水中音データの収集には専門的な機器などが必要であり、時間がかかり費用も高い。さらに収集したデータにラベルを付ける「ラベリング」作業にも大きな課題がある。音の種類を正確に識別するには水中音の伝搬知識や各船舶はどんな音を放射するかの知識が求められるため、膨大な労力と時間が求められる。

これらの要因から、水中船舶音データの収集とラベリングは非常に困難であり、学習に使用できるデータ数が少なくなりやすい。このように学習データが少量であっても高精度化を実現可能な少量データ学習が課題である。

表1 実験①の分類精度

ラベル付データ数	ラベル無データ数	テストデータ正解率
906	0	92.8±1.1%
270	0	84.2±1.6%
90	0	65.8±2.3%

### 少量データ学習の工夫

少量データ学習という課題の解決に対し、「データ拡張」「半教師あり学習<sup>3)</sup>」「転移学習」の大きく3つのアプローチをとっている。

まずデータ拡張について説明する。図4はデータ拡張の例である。運用時に新たに分類したい実際のデータと学習に使用したデータとの間では、特徴パターンが異なる場合が多く、学習時の性能を発揮できない可能性がある。そのため、ある船舶が放射した音を受波した場合の運用時に対するデータのバリエーションを人為的に作成することで、学習に用いるデータ量を拡張することを考える。

データのバリエーションとしては、船舶が接近したり離れたりとといった相対的な動き方に起因する周波数の変化(ドップラーシフト)や、相対的な位置関係に起因して音の干渉(強め合いや弱め合い)による振幅(音の大きさ)の変化、音源である船舶内部の機器の稼働状況に起因する周波数の組合せの変化、所望船舶以外の妨害音の発生などの周囲環境の変化などが考えられるが、収集したデータセットがあらゆる条件バリエーションを網羅していることは稀である。

そこで、図4に示すように、オリジナルの実際に収録されたデータに対して、周波数を変化させる処理、振幅を変化させる処理、一部の周波数成分を除去・強調させて周波数の組み合わせを変化させる処理、周囲環境の変化を模した白色雑音を追加させる処理を行ったデータを人為的に作成することで、データのバリエーションを増やし、オリジナルの学習データには含まれていないような運用時のデータに対しても、性能を発揮できるように学習を行うことをデータ拡張では目指している。

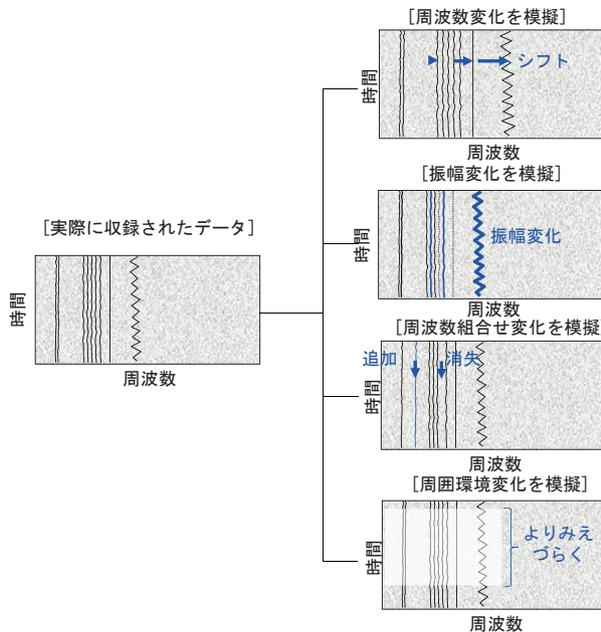


図4 データ拡張の例

次に、半教師あり学習について説明する。図5は半教師あり学習の概要の一例である。前述のようにラベリングされた水中音データを大量に取得することは困難であるが、ラベルが付与されていない、水中音だけのデータセットなどは大量に存在する場合がある。半教師あり学習では、正解ラベル付与済の一部のデータセットと、正解ラベル未付与のデータセットの両方を学習に使用する。図5の1ステップ目に示すように、限られたデータで学習した分類モデルを使って、正解ラベル未付与データに推論結果を疑似ラベルとして付与する。続いて、疑似ラベルが付与されたデータを正解ラベル付与済データと同様に学習に使用することで、より分類モデルが洗練される。この洗練された分類モデルで元々正解ラベル未付与だったデータの再度推論結果を新たな疑似ラベルとして付与する。この疑似ラベルの付与と学習を繰り返すことで正解ラベル未付与データの特徴を分類モデルの学習で有効活用する手法である。

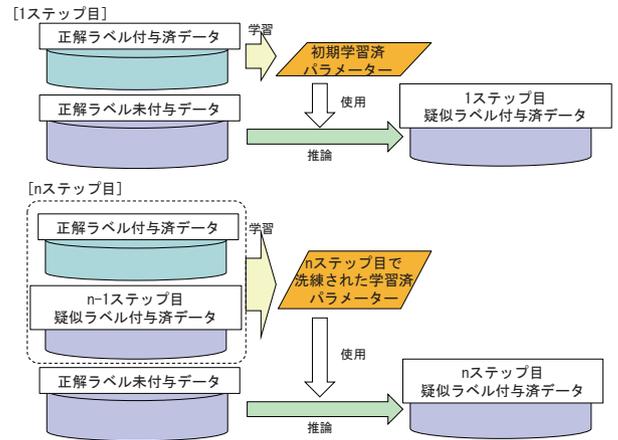


図5 半教師あり学習の概要

最後に図6に示す転移学習について説明する。例えばデータ数が限定的な船舶の水中音を分類する問題と、データ数が十分にあって入手が容易な、類似の音を分類する問題を考える。

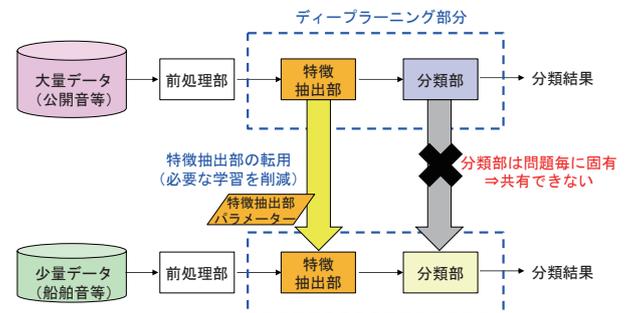


図6 転移学習の概要

それぞれ別のデータセットであり、別の音の種別であるが、どちらも音を使って分類する問題である。そのため、熟練者が分類に用いる特徴(強く表れているのはどの周波数でどんな組合せなのかという観点)の抽出部は共通的で、抽出した特徴の分類方法(○Hzと□Hzの組合せが強ければ種類Aなど)が問題ごとに異なると考えられる。そのため、この共通する特徴抽出部はデータ数が十分にある問題で学習したものを転用し、データ数が限定的な水中音は分類部の学習に注力させることで、特徴抽出部の学習の必要量を削減し、分類部の学習を効率的に行うというのが転移学習である。

### 実験② 少量データ学習の工夫の適用結果

前章で説明した少量データ学習の工夫を取り入れた場合の分類精度を評価する。今回は、データ拡張と半教師あり学習の工夫を取り入れている。

ここでデータ拡張として、図7のように学習のために実際に収録されたデータに対して、例えば1ステップ目では振幅変化のデータ拡張がなされたものを学習に使用し、正解ラベル未付与データへ疑似ラベルを付与するという半教師あり学習のステップを実施する。別のステップでは周囲環境変化のデータ拡張がなされたものを学習に使用し、疑似ラベルを付与する。このようにデータ拡張の種類や割合を学習段階ごとに変化させることで、必要なメモリー量を抑えながら全ての学習段階を通してデータのバリエーションが増加し、効率的な学習を行うことができる。

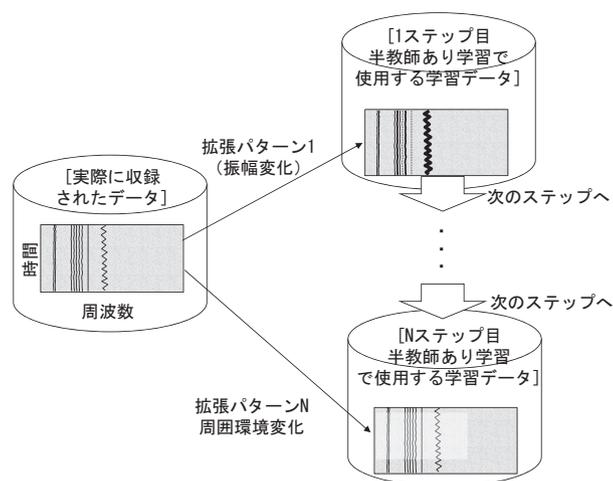


図7 データ拡張の例

表2に示すように、収録した906個の船舶音データへのラベリング数を実験①と同様906個全て、3割にあたる270個のみ、1割にあたる90個のみの3パターン用意し、実験①とは異なり、残りをそれぞれラベル無データとして使用した。用意した収録データ数が同じであっても、データ拡張による人為的なバリエーション増強と、半教師あり学習によるラベル無データからの有効活用により、少量データ学習の工夫を取り入れる前の表1の結果と比較して、それぞれ精度が向上し、最高で95%程度の正解率を実現し、工夫無に比べて2%pt程度の向上が確認できている。特にラベル付きデータ数が3割に絞られている場合であっても、ラベル未付与データと併用し、各種工夫も取り入れることで大幅に精度が改善し、全データがラベル付きデータとして使用できる場合と比較しても劣化を1%pt程度にまで低減できている。同様に、データ数が1割に絞られている場合であっても、全データがラベル付きで使用できる場合と比較しても劣化を3%pt程度にまで低減できていることを確認した。

今後、別途収録された船舶音データの識別の際には、今回使用したデータで学習した特徴抽出部のモデルを転移

学習に使用することで、別の船の分類であっても効率的な学習に寄与できると考えられる。

表2 実験②の分類精度

ラベル付データ数	ラベル無データ数	テストデータ正解率	(参考) 実験①結果
906	0	94.8±1.5%	92.8±1.1%
270	636	93.6±0.9%	84.2±1.6%
90	816	91.6±1.8%	65.8±2.3%

## まとめと今後の展望

本稿では、OKIが取り組んでいる水中音の分類に関連した技術の概要を述べた。現状では、限定的なデータで有効性を確認し、実証実験や音響データの分析で、実績を蓄積している段階である。今後更に実績を積み、お客様が解決したい実際のデータを使ってより実践的な実証実験を重ねること、納入後もお客様とともに更に学習を重ねて改良しながら分類精度を上げていくことが課題である。◆◆

## 参考文献

- 1) 川崎良道:パッシブソーナーの整相処理、OKIテクニカルレビュー 第224号、Vol.81 No.2、pp.72-75、2014年10月
- 2) 石原寛、稲葉稔智、石野田和英:水中音響技術による密漁対策IoTサービス、OKIテクニカルレビュー 第232号、Vol.85 No.2、pp.30-33、2018年12月
- 3) 玉井秀明、国定恭史、川村聡志、山本康平:AIエッジを強化するディープラーニング技術、OKIテクニカルレビュー 第239号、Vol.89 No.1、pp.28-31、2022年5月

## 筆者紹介

木原将吾:Shogo Kihara. 技術本部 先行開発センター 海洋技術先行開発部