

(55) 小型 AUV による水中マッピングデータ取得と 深層学習に基づく物体判別

大川 博史¹・大本 茂之²・木村 一朗³・宮本 崇⁴・檜山 和男⁵

¹正会員 株式会社エイト日本技術開発 技術本部 (〒164-8601 東京都中野区本町 5-33-11)

E-mail: ookawa-hi@ej-hds.co.jp

²非会員 株式会社エイト日本技術開発 技術本部 (〒700-8617 岡山県岡山市北区津島京町 3-1-21)

E-mail: oomoto-shi@ej-hds.co.jp

³正会員 株式会社エイト日本技術開発 技術本部 (〒164-8601 東京都中野区本町 5-33-11)

E-mail: kimura-i@ej-hds.co.jp

⁴正会員 山梨大学准教授 工学部 土木環境工学科 (〒400-8511 山梨県甲府市武田 4-3-11)

E-mail: tmiyamoto@yamanashi.ac.jp

⁵正会員 中央大学教授 理工学部 都市環境学科 (〒112-8551 東京都文京区春日 1-13-27)

E-mail: kaz@civil.chuo-u.ac.jp

本論文は、小型 AUV により取得した水中マッピングデータを用いて、水中にある物体の判別を迅速かつ正確に行う手法の構築について検討を行ったものである。AUV により取得した音響反射強度画像を入力データとして、深層学習の一つである畳み込みニューラルネットワークを用いて水中の物体判別を行った。物体判別には画像分類タスクを適用し、学習データの画像分割サイズを変化させることにより、水中の人工構造物・岩石・砂礫の分類結果に及ぼす影響について検討を行った。

Key Words: mapping data, side-scan sonar, AUV, image recognition, deep learning

1. はじめに

近年の土木・水産業界における ICT の積極的な導入により、情報化施工や BIM/CIM 活用をはじめとした一連の建設プロセスにおいて様々な取組が行われており、水中マッピング技術の分野においても発展の一途をたどっている。

その一方で、水中マッピングは陸上マッピングとは異なり、水中においては、容易にアクセスできないだけでなく電波や光を通しにくいといった事情から、水中マッピングは非常に難易度が高く、40m 以深では潜水行為自体が危険な作業となっているのが現状である。

近年の計測機器や水中ロボットの発達により分解能の高い音響探査装置を利用した高品質な水中マッピングが実現できるようになった¹⁾。これらによって得られる水中マッピングデータは、音の反射の強弱によって表されるため、色彩情報は持たず、これまで獲得したデータから物体判別を行う際、専門技術員による判読に頼ってきた。

また、ハードウェアにおいてもサイズが大きいくだけでなく、調査領域の近辺において、オペレータによる操作が必要な機器の利用が主流である。そのため、津波被災海域のような調査員が容易にアクセスできない場所において、水中マッピングデータを取得した後、がれきに代表されるような対象物を直ちに自動的かつ迅速に判別するといった行為は難しく非常に危険を伴う。

これらの問題を解決するために、小型 AUV を用いることで、オペレータが対象エリアに近接することなく水中マッピングデータを取得し、そのデータから水中に存在する物体を自動的に判別することで、迅速かつ正確に水中の現況把握を行う手法の構築を目標とする。本研究では、小型 AUV に搭載されたサイドスキャンソナーにより取得した音響反射強度画像を使用した。また、自動物体判別手段として、深層学習の一つである畳み込みニューラルネットワークによる画像分類タスクを適用し、学習データの画像分割サイズの変化による分類結果への影響について検討を行った。

2. マッピング手法およびデータについて

(1) マッピング手法

マッピングデータとは、地理的な位置座標に基づき、様々な情報を付加させたデータと定義され、水中マッピングデータは、水中で獲得した画像や地形、水質や時間などのデータと位置座標を付加したデータである。また、水中のマッピング手法として、船舶にサイドスキャンソナーやマルチビームエコーサウンダーなどの音響探査装置を搭載し、計測するのが一般的である。

本研究では、Autonomous Underwater Vehicle (AUV) とよばれる自律航行型無人潜水ロボットをプラットフォームとして、コースと深度を設定することにより、自動的かつオペレータが近くで待機することなくデータを取得することが可能となる。使用する AUV を図-1 に示す。

(2) サイドスキャンソナーと音響反射強度画像

水底面を把握する手法には、直接的手法と間接的手法がある²⁾。本研究では、間接的手法の1つで、AUV 搭載のサイドスキャンソナーによって獲得した音響反射強度画像を使用する。サイドスキャンソナー(図-2)は、音響を使った水底の画像探査システムであり、音響探査データを精密に処理・解析を行うことにより、沈船や落下物の捜索、水底地形や地質分布情報を獲得できるものである。

水底面に向かって10～2,000kHzの周波数を扇形に発射し、水底で散乱・反射した音波を受信する。この音波の強弱が水底面の地質によって異なるため、この強弱を濃淡表示することにより、白黒画像のように水底面の様子を面として捉えることができる。水中の地形・底質の判読は、陸上と比較してはるかに時間がかかり作業量も大きい。底質土や地形の判読については、採取された表層土を解釈の指標とし、一般には判読基準³⁾を基に専門技術員の目視によって判読される。リップル(砂漣)やデューン(砂堆)といった波型の模様は砂質土に限定された模様であり、そのような模様があれば、容易に砂質土と断定できる。また、人工構造物や水底から突出した物体に関しては、反射強度とその後方に伸びる音響的な影から物体の形状を判断しやすくなることが多い。



図-1 使用 AUV 諸元

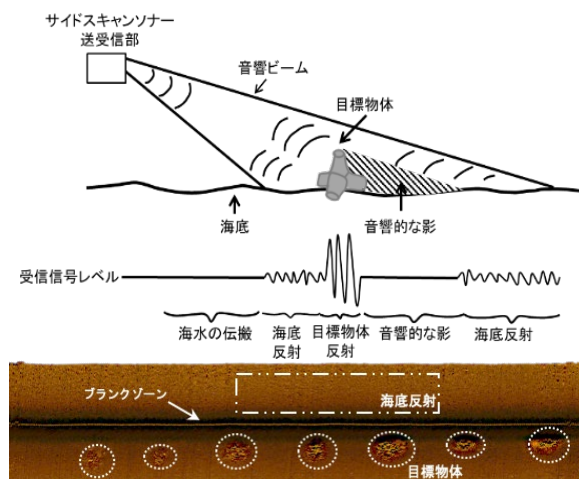


図-2 サイドスキャンソナー(片側)の原理と音響画像例

3. 物体判別と深層学習

(1) 物体判別について

音響反射強度画像から物体判別を行うアルゴリズムを大別すると①反射強度を特徴量として判別する方法 ②テクスチャ(模様)に着目して判別する方法の2種類に分類される。①は、平坦な地形の中に異なった反射特性をもつ場合に効果を発揮する一方、地形による反射強度の陰影を異なった材質と誤認しやすい。②は、フィルタで抽出した特徴量に基づいて分類を行い、本研究もその1種である。代表的なアルゴリズムの1つとして、白黒画像から計算された同時正規行列: Gray Level Co-occurrence Matrix (GLCM)³⁾から求められる特徴量を用いる方法があり、実際にこれらのアルゴリズムを利用して画像分類が行える汎用ソフトウェアも存在するものの、分類パターンから物体を直接判断するものはない。

(2) 深層学習について

深層学習とは、学習により特徴抽出過程を自動で獲得するものであり、研究者の知見に頼らず認識に有効な特徴量の抽出処理を自動化できるアプローチである。深層学習を用いた画像解析は、現在幅広い分野において実用化されており、海砂海域での底質判別を対象とした研究も行われている⁴⁾。本研究においては、画像解析技術のうち、「分類(Recognition)」タスクを採用した。

深層学習による画像解析方法は、まず、ニューラルネットワーク: NN のモデルを選択し、未学習の NN に「正解」のラベル付けをした大量の学習用画像を入力後、NN の結果の認識誤差により各ニューロンの重み係数を微調整する。次に、得られた NN に「正解」のラベルを付けた学習用画像とは別に、用意されたテスト用画像を与えて、正解率を調べる。これを繰り返すことにより、適切な学習曲線を持つ学習済み NN を得るものである⁵⁾。

4. 適用例

(1) 調査対象

本研究の調査対象として、人工構造物・岩石・砂礫が混在する海域である三重県熊野灘沖を取り上げた。使用 AUV にはサイドスキャンソナー EdgeTech2205 を搭載し、海底の音響反射強度画像を獲得した。低周波 75kHz から高周波 1,600kHz までの 2 周波同時に送受信できるものであるが、本研究においては 230kHz および 850kHz と設定した。可能な限り安定して高解像度な画像を取得するために、海底から 2m の高度を保つように自律航行するよう設定した。また、調査対象領域は 210m×170m とし、海域上を 6 往復させることにより、音響反射強度画像を獲得した。なお、インターフェロメトリ測深機能により同時測深を行っているが、今回は音響反射強度画像のみを学習データとした。

(2) 画像分類方法

画像分類には実際に取得した音響反射強度画像からブラックゾーンを除去・編集の後、GeoTiff 形式にて保存した。編集した音響反射強度画像（図-3）を用いて、画像分割サイズ [pixel] を変更することによる画像分類の正解率の変化について考察する。なお、あらかじめ学習データを作成する必要があるが、本研究では「人工構造物・岩石・砂礫」の 3 クラスに分類を行い、画像分割サイズはそれぞれ 10×10, 20×20, 40×40 [pixel] とした。学習データセット作成には、アノテーションツールである LabelMe を用い、物体形状に応じて指定・作成した（図-4）。これらの画像群を深層学習ライブラリ（Keras および TensorFlow）により、畳み込みニューラルネットワーク（Convolutional Neural Network : CNN）を用いて学習させた。なお、本研究で利用したネットワークの各層の構成は図-5 のとおりである。

なお、計算開発環境について、ハードウェア環境は CPU : intel Core i9 7900X 3.30GHz / GPU : GeForce GTX 1080-Ti / OS : Windows 10 Pro / RAM : 64.0 GB であり、ソフトウェア環境は、Anaconda3 上の Python3.6 / Keras / TensorFlow にてプログラムを実行している。

(3) 画像分類結果

本研究は、作成したデータからトレーニングデータとバリデーションデータの精度を検証することでハイパーパラメータを設定し、その検証を行うものである。繰り返し学習回数（エポック数）を 10～80 回の間で試行計算し、過学習の影響を鑑みて 40 回が最適と判断した。

画像分割サイズ 3 パターン : ①10×10 pixel ②20×20 pixel ③40×40 pixel におけるクラス別の解答率と判定画像例を表-1～3 に示す。なお、ピンク色に着色されたセル

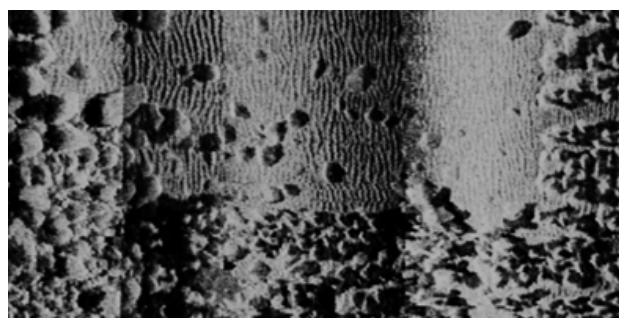
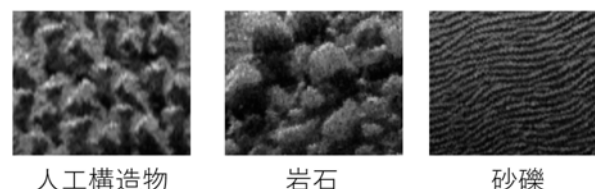


図-3 対象領域の音響反射強度画像



人工構造物

岩石

砂礫

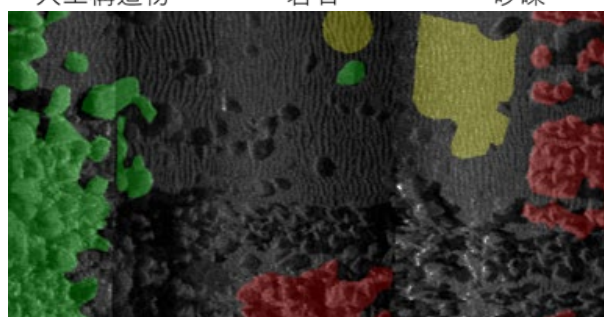


図-4 学習データ作成例

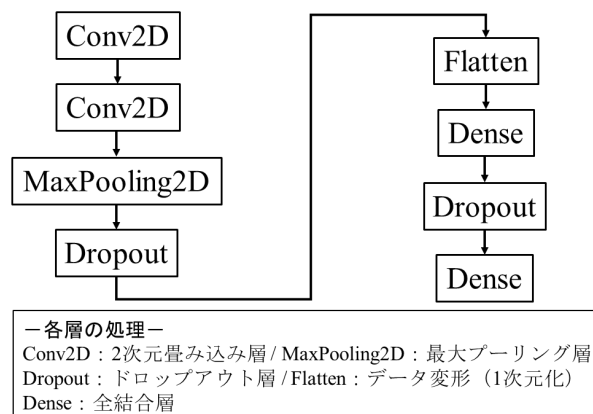


図-5 本CNNのネットワーク構成

は各クラスの正解率を表す。

正解率は、総合すると①が高いといった結果となったが、分割サイズと正解率の傾向はそれぞれ一致しない結果となった。砂礫に関しては、全ての分割サイズにおいて正解率は 95% を超え、特に③においては正解率 100% であった。一方、人工構造物は、分割サイズが大きくなるにしたがって、正解率は下がっていくといった結果となった。また、岩石に関しては、人工構造物とは逆の現象がみられ、分割サイズが大きくなるにしたがって、正解率が上がるという傾向がみられた。

表-1 ① H×W : 10×10 [pixel] (epoch:40) の場合

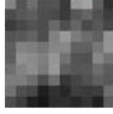
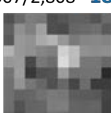
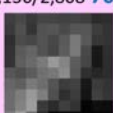
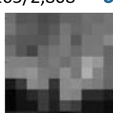
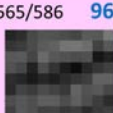
判断結果 正解	人工 構造物	岩石	砂礫
人工 構造物	611/628 97% 	6/628 1% 	11/628 2% 
岩石	507/2,808 18% 	2,136/2,808 76% 	165/2,808 6% 
砂礫	17/586 3% 	4/586 1% 	565/586 96% 

表-2 ② H×W : 20×20 [pixel] (epoch:40) の場合

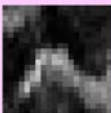
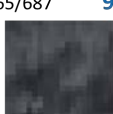
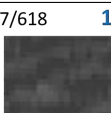
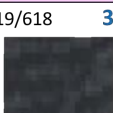
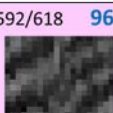
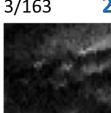
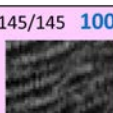
判断結果 正解	人工 構造物	岩石	砂礫
人工 構造物	62/145 43% 	78/145 54% 	5/145 3% 
岩石	1/687 1% 	621/687 90% 	65/687 9% 
砂礫	7/618 1% 	19/618 3% 	592/618 96% 

表-3 ③ H×W : 40×40 [pixel] (epoch:40) の場合

判断結果 正解	人工 構造物	岩石	砂礫
人工 構造物	11/34 32% 	15/34 44% 	8/34 24% 
岩石	3/163 2% 	159/163 98% 	1/163 1% 
砂礫	0/145 0% —	0/145 0% —	145/145 100% 

(4) 考察

画像分割サイズの変化によって、岩石と人工構造物の正解率の傾向が入れ替わる現象が確認された。このこと

から、クラスによって適切な分割サイズ (pixel) が異なり、その対象物体によって分割サイズを変化させる必要があると推測される。また、学習データ作成に際し、使用しているアノテーションツールが、対象画像内で最も大きい領域を占める物体のラベルを正解ラベルとして付与する処理を行っているため、これらの影響に関して検証を行う必要がある。一方、砂礫は人工構造物・岩石と比較して一貫して正解率が高かったが、その一因として、砂礫にしか現れない特有の現象 (リップルなど) をとらえやすかったためであると考えられる。逆に砂礫特有の現象が現れていない箇所に関しては、より判別が難しいといった結果になりかねないとも推測できる。また、サイドスキャンソナー搭載機器の異常な動作によって起こる、画像のゆれ・ゆがみ等のノイズの発生等につながるため、これらの処理に際しても注意が必要であると考えられる。

5. おわりに

本研究では、小型 AUV により取得した音響反射強度画像を入力データとして、深層学習による画像分類タスクを適用した。学習データの画像分割サイズ変化による分類結果への影響について検討し、以下の結論を得た。人工構造物・岩石・砂礫のクラスに分類を行う中で、総合的には 10×10[pixel]における正解率が高く、砂礫の正解率はその分割サイズにおいても高い結果となった。また、画像分割サイズによる正解率の変化の傾向はクラスによって異なることが明らかとなった。

今後の課題として、対象物による学習データの最適な画像分割サイズおよび学習データの作成方法による分類精度の検証、画像分類タスク以外の画像解析手法について取り組んでいく予定である。

参考文献

- 1) 澤隆雄, 大木健: 海底マッピングにおける海洋音響技術, 日本音響学会誌, 第 75 号 1 巻, pp.29-34, 2019.
- 2) 海洋調査協会: 海洋調査技術マニュアルー海洋地質調査編一, pp.29-30, 2004.
- 3) Haralick RM, Shanmugam K, Dinstein IH. Textural features for image classification, Man and Cybernetics, IEEE Transactions on Systems. 1973(6):610-21.
- 4) 長尾正之, 牟田直樹, 西島渉: 海砂海域での反射強度画像による底質判別の課題と深層学習による解決の試み, 土木学会論文集, Vol.74, No.2 pp. I_1441-1446, 2018.
- 5) Fukushima, K.: Neocognitron: A Self-organizing Neural Network Model for a Mechanism of Patte Unaffected by Shift in Position, Biological Cybernetics, Vol.36, pp.193-202, 1980.