

(19) 水中マッピングデータを用いた マルチモーダル深層学習に基づく物体分類

大川 博史¹・八木 笙太²・大本 茂之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 中野区本町 5-33-11)

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

²学生会員 中央大学大学院 理工学研究科 都市人間環境学専攻 (〒112-8551 文京区春日 1-13-27)

E-mail: a16.365j@g.chuo-u.ac.jp

³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

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 文京区春日 1-13-27)

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

本論文は、小型の自律航行型無人潜水機 (Autonomous Underwater Vehicle : 以後「AUV」とする) 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したものである。水中音響反射強度画像に加え、水深データおよび散乱反射強度データを入力データとして用い、これらの異種データの特徴量を融合したマルチモーダル深層学習に基づく畳み込みニューラルネットワーク物体分類手法を提案した。本手法の有効性を検証するために、実測した水中マッピングデータへの適用を行った。

Key Words: *mapping data, AUV, deep multimodal classification, convolutional neural network*

1. はじめに

近年の測量・測位技術や無人航空機 (Unmanned Aerial Vehicle : UAV) に代表される自動制御ロボット技術の発展により、陸上マッピング技術は飛躍的に向上し、水中マッピング技術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、危険な作業を伴う。

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可能となった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、直ちに自動的かつ迅速に判別するといった行為は困難である。

これらの問題を解決するために、本研究では小型 AUV を用いて取得可能なデータから、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。

著者らは、既往の研究において作業の自動化を目的とし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の適用を行い、物体の局所的な特徴が分類に影響を与えていることを確認した²⁾。本論文は、従来手法の更なる精度向上を目指し、これまで使用してきた水中音響反射強度画像に加え、水深データおよび散乱反射強度データを用いた。これらの異種データの各モダリティを融合し、マルチモーダル深層学習を行う手法を提案する。本手法を用いて学習データに対するテストデータの正解率を比較・評価することにより、本手法の有効性を示す。

2. マッピングデータと取得方法について

(1) 使用機材

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、水中で取得した画像や地形、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。水中マッピングデータの取得方法として、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、AUVをプラットフォームとして、任意のセンサーを搭載することにより、必要なデータを取得する。コースや深度を設定することにより、オペレータが近くで待機することなく、自動的にデータを取得することが可能となる。本論文に、使用するAUV(図-1参照)に搭載されたセンサーのうち、インターフェロメトリ音響測深機(EdgeTech; 2205AUV)を用いて、3種類のデータ(音響反射強度画像/水深データ/後方散乱強度)を同時に取得することが可能である。



図-1 使用AUV諸元

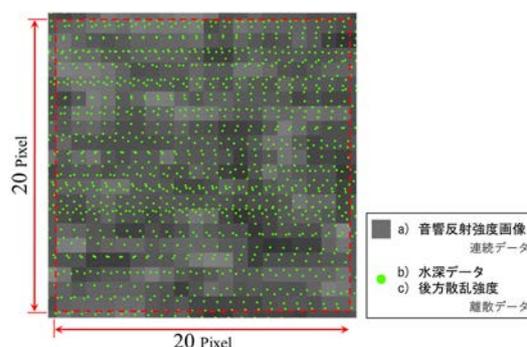


図-2 使用マッピングデータ分布

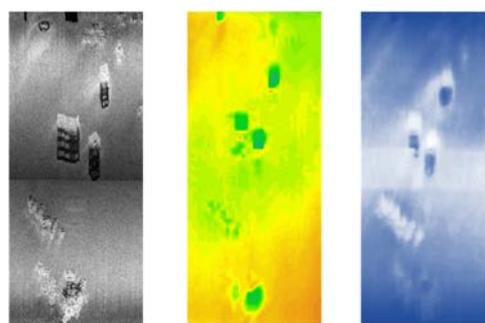


図-3 取得データ (※比較のためラスタ化したもの)

(2) 水中音響測深機を使ったマッピングデータの取得

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⁾が、本研究では、間接的手法の1つで、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方法によるデータを採用する。AUVに搭載した音響測深機によって音響反射強度画像/水深データ/後方散乱強度を取得する。なお、取得したマッピングデータの分布図を図-2に示す。音響反射強度画像は各セルに値をもつ連続データであり、図中の緑のポイントで示される水深データおよび後方散乱強度は共通の平面座標に特徴量を持った離散データである。また、特徴比較のためにラスタ表示したものを図-3に示す。

各データの詳細については次項にて説明する。

a) 水中音響反射強度画像

水中では反射する音波の強弱が水底面の地質等によって異なるため、強弱を濃淡でラスタ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可能となる。本論文では、次項に示す水深データ計測システムとは独立した2周波同時発振型のイメージングサイドスキャンソナーを用いる。これにより、特徴量の異なる独立したデータを有した高解像度なデータを取得している。

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。砂澁や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可能である。

b) 水深データ

送波器から放射された音波が、海底に到達して後方散乱し、その音波を複数の受波器で受信する。それぞれの受波器での音波の到達時刻が到来角によって微妙に違うことを利用して、その到来角と散乱点までの往復距離を算出することで測深点の情報を取得し、各点において後方散乱強度も同時に保存可能である。

c) 水中後方散乱強度

水深データを計測する発射された音波が水底面等で散乱した後、発射した方向に戻る音波の強さを後方散乱強度と呼ぶ。インターフェロメトリ測深の場合、前述の水深データと同時に副次的に後方散乱係数として取得可能である。後方散乱係数は散乱面の粗さに応じた入射角依存性がある。後方散乱強度と水中音響反射強度画像は類似の関係にあり、位置座標におけるスカラー値自体には意味を持たず、常に周囲との相対的な差において意味をなすものである。本論文において、前者は受波ビームごとの信号強度を収録している数値データとし、後者は前者と異なる周波数帯2波を用いて高解像度なサンプリングを実施した音響画像と定義する。

3. マルチモーダル深層学習による物体分類

(1) マルチモーダル深層学習

深層学習とは、学習により特徴抽出過程を自動で獲得し、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が行えるアプローチである。マルチモーダル深層学習は、それに加え、複数のモダリティからの情報を統合的に処理することで、より精度の高い特徴抽出を可能とする。また、深層ニューラルネットワークにおいて複数のモダリティを扱うアプローチは、モダリティごとのネットワークの最上位の隠れ層を共有することが主要となっている。

本研究では、入力層でテンソルの連結を行った後、畳み込みニューラルネットワークを特徴抽出器として用いることにより、複数のモダリティを扱うこととした。

(図-4 入力層を参照) この方法の特徴として、各モダリティ間で同期の取れたデータを扱う場合、複数のモダリティに共通する特徴表現を獲得可能であることが挙げられる⁴⁾。つまり、水中マッピングデータのように地理空間座標を有するデータとは親和性が高いということであり、衛星画像を代表とするリモートセンシングにおいても広く利用されている。

(2) 物体分類

畳み込みニューラルネットワーク (Convolutional Neural Network : 以後「CNN」とする) は、画像を入力データとする際に使用される。CNNによる画像分類⁵⁾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⁶⁾。

本研究では、画像分類の高精度化を目指し、複数種的水中マッピングデータと CNN を組み合わせたマルチモーダル深層学習を行うものであり、任意のピクセルサイズ単位に分割されたデータを分類するものである。本論文にて提案するモデル構成を図-4に示す。

本モデルは、入力層・中間層・出力層に分類され、入

力層では各モダリティの入力データに対し連結層において、音響反射強度画像および水深データ・後方散乱強度の各テンソルを融合する。中間層では、融合したデータに関して特徴量の抽出を行った後、出力層にて特徴量に基づいたクラス分類を行う。なお、中間層では、図-2に示すとおり、畳み込み層・プーリング層・ドロップアウト層・アクティベーション層の4層を1セットとし、2度繰り返すことにより特徴量の抽出を行っている。また、アクティベーション層では、活性化関数として正規化線形関数を採用した。一方、出力層においては、全結合層およびドロップアウト層での処理後、活性化関数を用いることによりクラス分類を行っている。活性化関数には、多クラス分類問題に一般的に使用されるソフトマックス関数を採用し、クラス分類を行った。なお、本深層学習モデルにおける最適化アルゴリズムには、適応モーメント推定を採用している。

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島根県隠岐の島沖を取り上げ、前述の水中マッピングデータを取得した。水中音響反射強度画像は、540/1,600 kHzの2周波同時発振にて取得し、水深データおよび後方散乱強度については、周波数500kHzにて取得した。可能な限り安定して高解像度なデータを取得するために、水面から50m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200m×150mとし、海域上を往復させることにより、水中マッピングデータを取得した。

(2) 計算条件

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ンクゾーンを除去・編集の後、GeoTiff形式にて保存

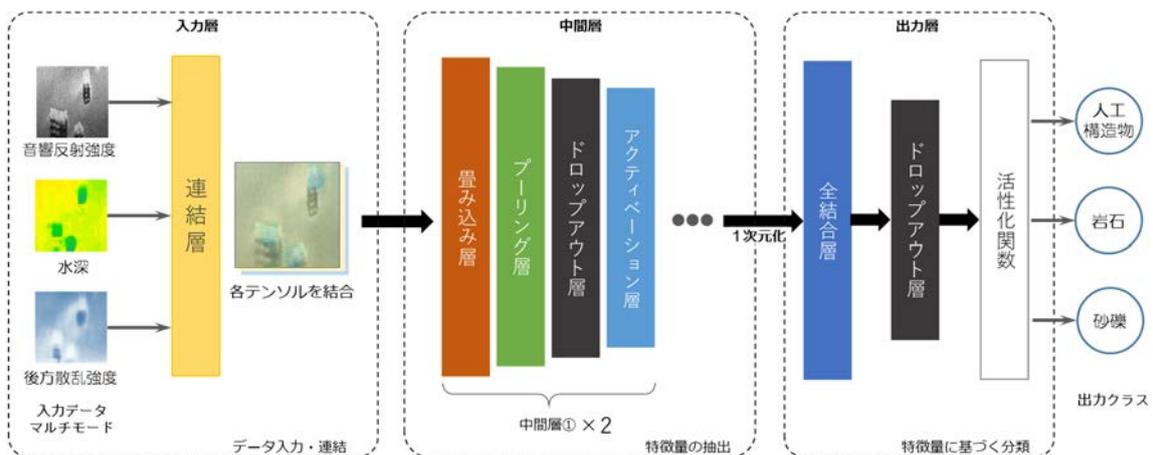


図-4 CNNモデル構成

した。編集した水中音響反射強度画像・水深データ・散乱反射強度データを用いて、前章の CNN を用いたマルチモーダル深層学習モデルにより学習させることにより物体分類を行う。

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 20×20 pixel とした。人工構造物/岩石/砂礫の物体スケールは概ねそれぞれ、1.0m 以上/20cm 以上/20cm 未満と定義する。学習データセット作成には、アノテーションソフト (Labelme) を用い、水中音響反射強度画像中の物体形状に応じて任意に指定した後、学習データのラベル付けを行った (図-5 参照)。また、水深データおよび後方散乱強度の離散データは再近傍法によりグリッド化処理を行った上で連結層にてデータを融合する。

(3) 分類結果および考察

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を用いてハイパーパラメータを設定することで、クラス分類モデルを構築する。繰り返し学習回数は 200 回に設定した。作成した分割画像データ数は、それぞれ人工構造物 176 枚/岩石 208 枚/砂礫 352 枚であり、その半数をトレーニングデータ、残りの半数をテストデータとして使用した。

なお、比較検証のためにシングルモード時：水中音響反射強度画像のみを入力データとした時とマルチモード時の分類における各クラスの分類結果を表-2 に示し、各セルの上段には分類数/クラス総数、下段に各クラスの解答率を示す。

全体正解率はシングルモード時が 73%であったのに対し、マルチモード時は 97%となり、特に、人工構造物および岩石の正解率が大幅に向上した。また、全体の正解率が大幅に向上しただけでなく、各クラスにおける正解の様相も変化した。シングルモード時では、人工構造物を岩石と分類し、岩石を人工構造物に分類する誤答傾向が多くみられた。一方、マルチモード時では、砂礫と岩石の間で誤答がみられた。水深データあるいは後方散乱強度により、人工構造物の特徴を補足しやすくなったためであると考えられる。

5. おわりに

本論文では、水中音響反射強度画像に加え、水深データおよび散乱反射強度データを用いたマルチモーダル深層学習による分類手法を提案し、学習データに対するテストデータの正解率の評価を行った。本分類手法における正解率は、水中音響反射強度画像のみを使用した場合

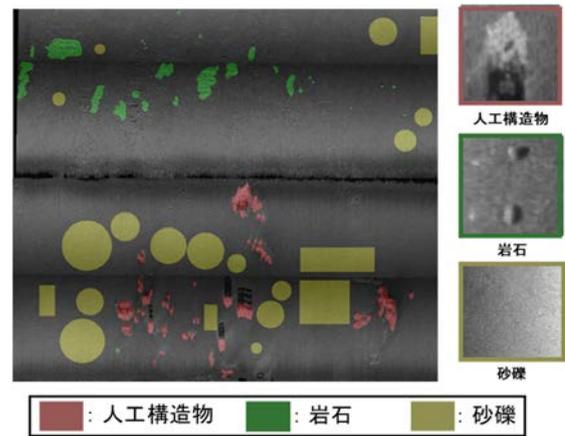


図-5 音響反射強度画像および学習データ例

表-2 画像分類結果

	【 シングルモード 】 水中音響画像のみ			【 マルチモード 】 水中音響画像+水深+散乱強度		
	人工 構造物	岩石	砂礫	人工 構造物	岩石	砂礫
人工 構造物	37/88 42 %	39/88 44 %	12/88 14 %	88/88 100 %	0/88 0 %	0/88 0 %
岩石	37/104 36 %	39/104 54 %	12/104 11 %	0/104 0 %	97/104 93 %	7/104 7 %
砂礫	0/176 0 %	0/176 0 %	176/176 100 %	0/176 0 %	4/176 2 %	172/176 98 %
	269/368			357/368		
	全体正解率： 73 %			全体正解率： 97 %		

と比較して大幅に向上し、本手法の有効性が確認できた。

今後の課題として、モード数による正解差の原因分析、分類クラスの細分化、本研究によって得られた学習済モデルの利用について検討していく予定である。

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) 大川博史, 宮本崇, 樫山和男: 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, 土木情報学シンポジウム講演集, Vol.45, pp.217-220, 2020.
- 3) 海洋調査協会: 海洋調査技術マニュアル—海洋地質調査編一, pp.29-30, 2004.
- 4) Jiquan, N., Aditya, K., Mingyu, K., Juhan, N., Honglak, L., and Andrew, Y.N.: Multimodal deep learning. *In Proceedings of the 28th international conference on machine learning (ICML-11)*, pp. 689-696, 2011.
- 5) Haralick RM, Shanmugam K, Dinstein IH.: Textural features for image classification, *Man and Cybernetics, IEEE Transactions on Systems*, pp.610-621, 1973
- 6) 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-I_1446, 2018.