

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の高精度化

(株) エイト日本技術開発 正会員 ○大川 博史 中央大学大学院 学生会員 八木 笙太
山梨大学 正会員 宮本 崇 中央大学 正会員 檜山 和男

1. はじめに

近年の計測機器や水中ロボットの発達により、高分解能の音響探査装置を利用した高品質な水中マッピングが可能となった¹⁾。これらの方法によって得られるデータの1つに音響反射強度画像があるが、音の反射の強弱によって表されるため、色彩情報は持たない。そのため、取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。

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 (Convolutional Neural Network : 以後「CNN」とする) にマルチモーダル深層学習による画像分類タスクを適用した³⁾。本論文は、使用する入力データに対し、データの特長値を解析し、新たな入力値として付加することで、さらなる正解精度の向上を目指すものである。

2. 使用ハードウェアと取得データ

本研究では、小型 AUV (Autonomous Underwater Vehicle : 自立航行型無人潜水艇) に搭載したサイドスキャンソナー (Edge Tech 2205) 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。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、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 (図-1 参照) である。特徴比較のために

ラスター表示したものを図-2 中の入力データに示す。

なお、図中には、砂礫中に人工構造物である漁礁が点在する様子が示されており、取得データによって各々異なった特徴が確認できる。

3. マルチモーダル深層学習による物体分類

深層学習とは、人間の知見による特徴抽出タスクをコンピューターに学習させることによって、自動的に特徴量を抽出するものである。マルチモーダル深層学習⁴⁾は、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。

本研究では、物体分類の高精度化を目指し、融合された複数種の水中マッピングデータ (図-1 参照) 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画像データの分類を行う。使用する CNN モデ

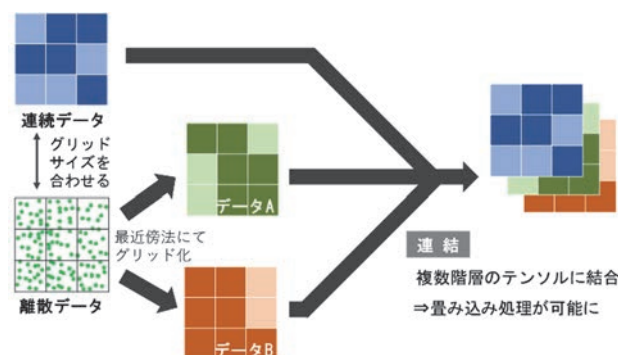


図-1 入力データの融合方法

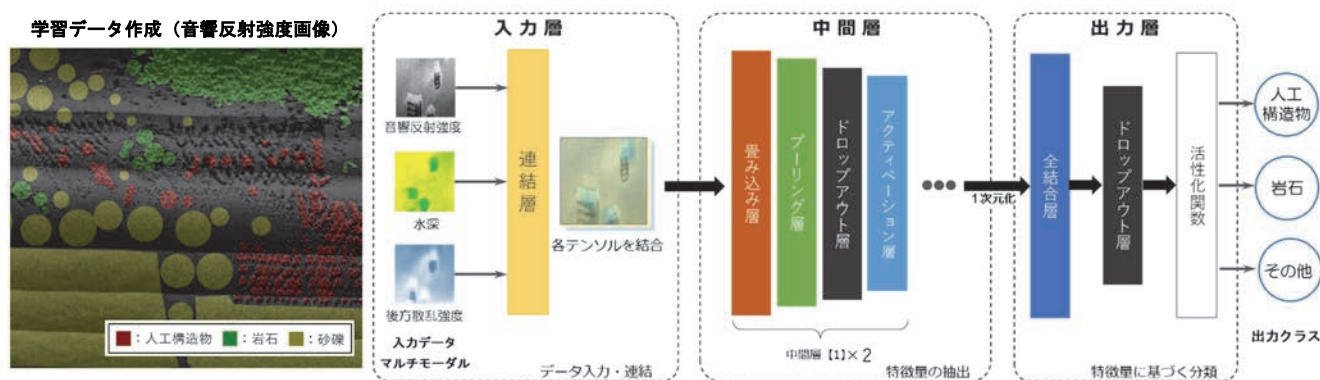


図-2 CNNモデルの概略

キーワード 音響反射強度画像, Deep Learning, 画像分類, マルチモーダル深層学習, 畳み込みニューラルネットワーク

連絡先 〒164-8601 東京都中野区本町 5-33-11 (株) エイト日本技術開発 TEL : 03-5341-5111 E-mail : ookawa-hi@ej-hds.co.jp

ル(図-2 参照)は、入力層・中間層・出力層に分類される。入力層では、各モダリティの入力データに対し連結層を設けることにより、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。これらの処理により、異種のデータであっても、位置座標に基づいたデータの特徴を表現でき、データ間に相関関係を持たせることが可能となる。

4. 入力データ特性値の解析および画像分類

本論文では、従来の入力データに対し、さらに特性値の解析データを付加することにより、正解率の向上を目指す。前章に示した3種類の入力データのうち、水深データの特性値を解析し、これを新たに入力値として付加することとする。

地形の形態的特徴はそれぞれの地形の面積・傾斜・起伏・曲率などによって量的に表現される。このような形態的特徴を量的に表現した場合の数値を「地形特性値⁵⁾」と呼び、地形特性値は現地測量や地形計測によって得られたデータを用いて算出することができる。本研究では、これらの地形特性値を水深データへ適用し、格子点での標高値から傾斜の大きさ(ノルム)および曲率(ラプラシアン)を求める。地形特性値の定義・特徴について、表-1に示すとともに、実際に使用する20×20pixel(10cm/px)の水深データおよび傾斜の大きさ、曲率の解析例をそれぞれ図-3に示す。

本研究の調査対象として、人工構造物・岩石・その他(砂礫等)を採用し、対象地域としてこれらの混在する水域であり水深20m程度の海域を有する三重県熊野沖を採用した。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形式にて保存した。編集した水中音響反射強度画像・水深データ・後方散乱強度データおよび本稿にて解析した傾斜の大きさおよび曲率を用いて、

表-1 地形特性値の定義

地形特性値	定義	特徴および地形学的な意味
標高 u	格子点における標高値	基本となる数値データであり、スカラー量である。
傾斜の大きさ (ノルム) S	標高の一次微分ベクトルの絶対値 $S = \sqrt{\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2}$	標高の変化の割合を表す。対象格子点と隣接する点の距離と標高差から算出
曲率 (ラプラシアン) L	勾配の微分つまり、標高の二次微分 $L = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}$	方向に依存しないスカラー量。地形の凹凸を表す。

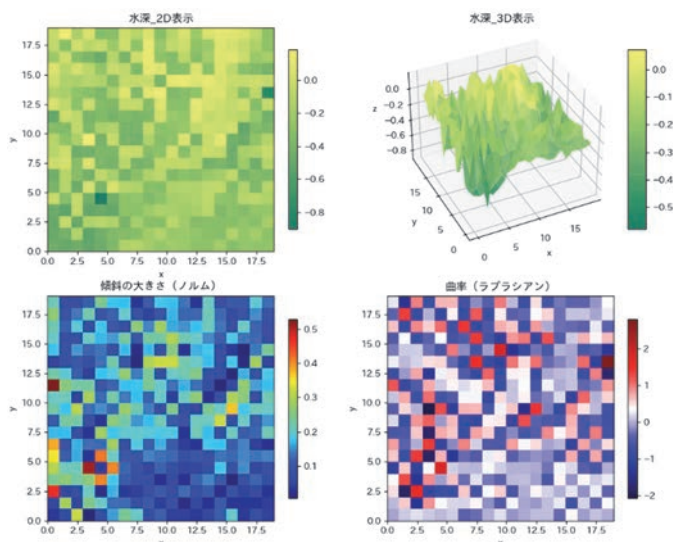


図-3 水深および傾斜の大きさ・曲率の解析例

前章のCNNによる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って物体分類を行う。また、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によりハイパーパラメータを設定した。なお、標高値の特性値解析を入力データとして付加した画像分類の結果および考察に関しては講演時に示すものとする。

5. おわりに

本研究では、水中の物体を対象としたマルチモーダル深層学習に基づく画像分類の正解率のさらなる向上を目指し、入力データの特性値を解析し、新たな入力値として付加した。特性値の解析には水深データを採用し、傾斜の大きさおよび曲率を算出し、これらが正解精度に与える影響を検証した。

今後の課題として、さらなる適用事例の増加および分類クラスの細分化を行うとともに転移学習による全ての水中領域を対象とした物体分類について検討を行う予定である。

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) Krizhevsky, A., Sutskever, I., and Hinton, G.E.: Imagenet classification with deep convolutional neural networks, *Advances in neural information processing systems*, pp.1097-1105, 2012.
- 3) 大川博史, 宮本崇, 檜山和男: 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, 土木情報学シンポジウム講演集, vol.46, pp.73-76, 2021.
- 4) Ngiam, J., Khosla, A., Kim, M., Nam, J., Lee, H., and Ng, A.Y.: Multimodal deep learning. *In Proceedings of the 28th international conference on machine learning (ICML-11)*, pp. 689-696, 2011.
- 5) 太田岳洋, 八戸昭一: 数値標高モデルによる地形計測の現状と応用例. 応用地質, vol.46, No.6, pp.347-360, 2006.