

## 水中音響反射強度画像を用いた深層学習に基づく画像分類

(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

近年の土木業界においては、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 ICT の積極的な導入が行われている。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<sup>1)</sup>。これらによって得られる水中マッピングデータの 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、色彩情報は持たず、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<sup>2)</sup>。

筆者らは、これらの作業の自動化を目指し深層学習の 1 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の適用を行った結果、物体の局所的な特徴が分類へ影響を与えていることが確認できた<sup>3)</sup>。本論文は、従来手法に入力データの局所的な特徴を際立たせるよう変換を行うことにより、分類結果の更なる精度向上を目指すものである。

### 2. 使用ハードウェアと音響反射強度画像

本研究では、音響反射強度画像の取得には、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) を使用する。なお、サイドスキャンソナーは、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、沈船や落下物の搜索、水底地形や地質分布状況の獲得など広く利用されている。

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。

### 3. 深層学習に基づく画像分類

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いるが、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。本手法では取得した音響反射強度画像を任意のサイズに分割したものを入力データとし、畳み込みニューラルネットワーク<sup>4)</sup>による画像分類タスクを適用した。本システムの流れを図-1 に示す。

画像分類は取得した音響反射強度画像からブランクゾーンを除去後、GeoTiff 形式にて保存した。また、局所的な特徴を際立たせるために画像の閾値処理 (図-2 参照) 後の音響反射強度画像を用いて、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を算出した。なお、学習データ作成について本研究では「人工構造物・岩石・砂礫」の 3 クラスの分類を行い、物体スケールは概ねそれぞれ、1.0m 以上 / 20cm 以上 / 20cm 未満と定義した。

学習データセット作成には、アノテーションツ

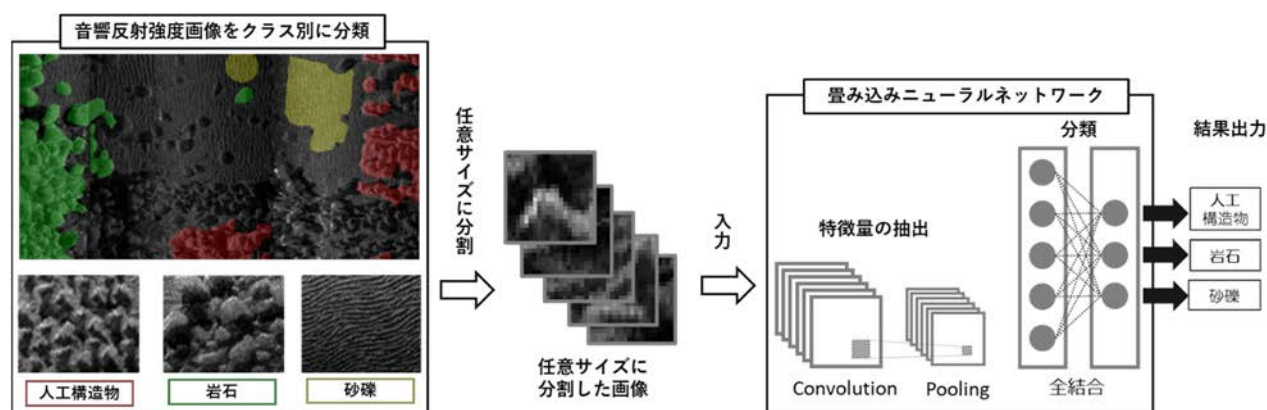


図-1 画像分類のフローチャート

キーワード 音響反射強度画像, 深層学習, Deep Learning, 画像分類, 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

ルを用い、物体形状に応じて指定・作成した。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 CNN の構成は、表-1 のとおりであり、計算速度向上のため可能な限り単純な構成としている。

また、ハードウェア環境は CPU : intel Core i9 10900KF 3.70GHz / GPU : GeForce RTX 2080 Ti / OS : Windows 10 Pro / RAM : 64.0 GB である。

#### 4. 適用例および分類結果

本論文では、人工構造物・岩石・砂礫を採用し、これらの混在する水域である島根県江の川流域を対象とし、前章の手順により画像分類を行った。画像分類に使用した音響反射強度画像および学習データは図-3 のとおりである。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検証からハイパーパラメータを設定し、その検証を行った。

繰り返し学習回数（エポック数）は過学習の影響を鑑みて学習回数 40 回が最適と判断し、画像分類結果を表-2 に示すとともに入力データを閾値処理した結果は本発表時に示す。また、深層学習のモデルが入力画像のどの部分を重要視して出力に至ったのかを Gradient-weighted Class Activation Mapping（Grad-CAM）<sup>5)</sup> と呼ばれる手法を適用することにより、着目度が高い箇所は赤色で表示される特微量ヒートマップで表示した。分類結果より、解像度および対象物体によって正解率が変化することが確認できる。とりわけ、解像度の高い人工構造物に対し着目度の高い箇所がみられ、局所的な境界部分を深層学習モデルが人工構造物と判断したものと推察できる。

#### 5. おわりに

本研究では、サイドスキャンソナーを利用して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。深層学習モデルの判断が水中物体によって異なることが確認でき、物体固有の特徴をより際立たせる処理を行うことで精度の高い物体分類を実現できると考えられる。

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、転移学習を利用した画像分類タスク以外の画像解析手法を導入する予定である。

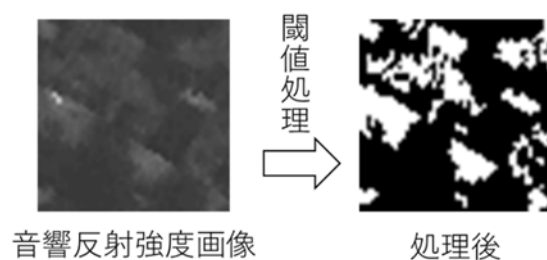


図-2 閾値処理例

表-1 CNN の構成と出力サイズ

| 層番号 | 層種       | 出力サイズ |
|-----|----------|-------|
| 1   | 畳み込み層    | 64×32 |
| 2   | 畳み込み層    | 36×64 |
| 3   | プーリング層   | 9×64  |
| 4   | ドロップアウト層 | 9×64  |
| 5   | 全結合層     | 128   |
| 6   | ドロップアウト層 | 128   |
| 7   | 全結合層     | 3     |

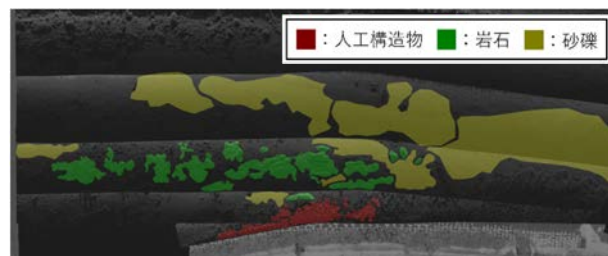


図-3 音響反射強度画像および学習データ例

表-2 解像度別 画像分類結果（上段：特微量ヒートマップ 下段：元画

| 10×10 [pixel] |     |         | 20×20 [pixel] |     |         | 40×40 [pixel] |     |         |
|---------------|-----|---------|---------------|-----|---------|---------------|-----|---------|
| 人工<br>構造物     | 岩石  | 砂礫      | 人工<br>構造物     | 岩石  | 砂礫      | 人工<br>構造物     | 岩石  | 砂礫      |
| 正解率<br>95%    | 54% | 97%     | 正解率<br>61%    | 99% | 84%     | 正解率<br>48%    | 41% | 99%     |
| 解像度<br>10×10  |     | [pixel] | 解像度<br>20×20  |     | [pixel] | 解像度<br>40×40  |     | [pixel] |

#### 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,
- 2) 植木俊明: サイドスキャンソナーによる海底の把握, 写真測量とリモートセンシング, 29 巻 Special 号, pp.84-93, 1990.
- 3) 大川博史, 宮本崇, 檜山和男: 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.
- 4) 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.
- 5) Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D. and Batra, D.: Grad-CAM: Visual explanations from deep networks via gradient-based localization, *IEEE International Conference on Computer Vision (ICCV)*, pp.618-626, 2017.