

大深度自律サンプリングに資する水中画像の自動認識・解釈手法の開発

課題責任者

金子 達哉

海洋研究開発機構 技術研究開発部門 海洋技術研究センター
海洋ロボティクス技術研究グループ

著者

金子 達哉^{*1}, 各務 均^{*1}, 岩下 和矢^{*1}, 植田 亮^{*2}

^{*1} 海洋研究開発機構 技術研究開発部門 海洋技術研究センター 海洋ロボティクス技術研究グループ

^{*2} 東京大学工学部システム創成学科

キーワード: 物体検出, 水中画像, 自律型無人探査機, Open-vocabulary object detection, Salient object detection

1. 緒言

自律型無人探査機 (AUV) を用いた広範囲・高効率の物体・生物の自律的観測・サンプリングが求められる。対象分野と対象物は多様であり、具体的には、海底地質学における岩石や、地球生命科学における生物、海底鉱物資源などが対象となる。

自律サンプリングにおいて、まず、水中画像などから対象を自動的に識別する必要がある。深層学習の台頭により、水中画像における物体検出に関する様々な研究が進められている。これらの研究の多くは教師あり学習を採用している。しかし、水中画像データとそのラベルが不十分なため、教師あり学習では十分な性能を達成できない。さらに、教師あり学習では未見の対象 (例えば、新種生物) を発見できず、これは科学研究において重要な課題である。

本研究では、第一に、水中画像から、自動的に特徴的な領域の抽出を行うことを目的とし、水中画像において、salient object detection (SOD) 手法や open-vocabulary object detection (OVD) 手法などの性能評価を行う。

2. 水中画像データ

本研究で性能の検証や fine tuning に用いた水中画像データは以下二種類である。一つは、深海生物検出用の公開データセット「J-EDI Organism Detection Dataset (JODD) [1, 2]」である。これには、計 8,151 枚の水中画像と、19 種類の生物及び機械の bounding box のアノテーションが含まれる。本研究では、このデータを主に定量的な性能評価に用いた。もう一つは、著者らが自律サンプリング用に開発中の小型 AUV [3] で撮影されたデータである。このデータは、定性的な評価に用いた。

3. Salient object detection

Salient object detection (SOD) は、画像の中から、顕著な領域をセグメンテーションする手法であり、追加学習無しで水中画像から対象を検知できる可能性がある。本研究では、多くのデータセットに対して高い性能を示している BiRefNet [4] を用いて検証した。その結果、SOD は対象生物をセグメンテーションできるケースがある一

方で、重なった生物を分離することができなかった。また、SOD ではどのようなオブジェクトを関心領域として抽出するかを指定できず、それは、事前学習データに大きく依存することが考えられる。このようなことが課題となり、定量評価の結果、次節の OVD 手法の性能を上回ることにはできなかった。しかし、SOD はセグメンテーションを可能にすることは利点であり、OVD との組み合わせにより有効となる可能性もある [5]。

4. Open-vocabulary object detection

Open-vocabulary object detection (OVD) は vision-language model を用いて、任意のテキストに対応するカテゴリーの物体を検出する手法である。特に、学習に用いられたカテゴリー以外の未知カテゴリーにも有効であることが特徴である。(このようなタスクは zero-shot object detection とも呼ばれる。) 本研究では、多くのデータセットに対して高い性能を示す LLMDet [6] を用いて、まず、OVD を追加学習無しで水中画像に適用し検証した。

まず、水生生物の検知を JODD データを用いて検証した。JODD には 'fish' や 'crab' など 19 種類のカテゴリーの生物が含まれるが本研究では、それらを区別せず、'marine organism' のテキストを入力し検出を行った。これにより、より広範な生物を検出することが可能となり、新種生物を検出する可能性も見込める。結果として、物体検出で一般的に用いられる AP (average precision) で定量評価を行った結果、AP@[0.5:0.95] が 58.24%、AP@0.5 が 79.16% の性能を示した [5]。

次に、小型 AUV で撮影されたデータを用いて検証した。その結果、同様に生物を精度よく検出することを確認した。それに加え、著者らは小型 AUV のドッキング制御をマーカーの画像認識により行っているが、そのようなマーカーの検知もテキストの調整次第で行ことが可能であった。このように小型 AUV に対する適応性だけでなく、OVD モデルの更なる応用可能性を示すことができた [5]。

7. 結言

本研究では、大深度自律サンプリングに資する水中画像の自動認識・解釈手法のため、従来のような教師あり学

習に依らない手法の検討を行った。その結果、水中画像に対する追加学習を行わなくても、高い性能で物体検出を行うことができた。このような手法を活用することで、将来的な自律サンプリングに資することが期待される。

文献

- [1] Takaki Nishio, Yuki Kawae, “J-EDI Organism Detection Dataset (JODD)”, JAMSTEC, 2025.
- [2] Takaki Nishio, Yuki Kawae, “Deep-sea image dataset for organism detection”, Data in Brief, Volume 65, 112462, 2026.
- [3] Yosaku Maeda, Hitoshi Kakami, Yoshitaka Watanabe, Tatsuya Aso, Kazuya Iwashita, Tatsuya Kaneko, “Towards an Autonomous Sampling System for the Hadal Zone”, 2025 IEEE Underwater Technology (UT), 2025.
- [4] Peng Zheng, Dehong Gao, Deng-Ping Fan, Li Liu,

Jorma Laaksonen, Wanli Ouyang, Nicu Sebe, “Bilateral Reference for High-Resolution Dichotomous Image Segmentation ” , CAAI Artificial Intelligence Research, Volume 3, 9150038, 2024.

- [5] Tatsuya Kaneko, Hakan Bilen, Tomoya Inoue, Hitoshi Kakami, Kazuya Iwashita, “Underwater Object Detection for Autonomous Observation and Sampling Using General-Purpose Object Detection Techniques” , ASME 45th International Conference on Ocean, Offshore & Arctic Engineering (OMAE2026), 2026.
- [6] Shenghao Fu, Qize Yang, Qijie Mo, Junkai Yan, Xihan Wei, Jingke Meng, Xiaohua Xie, Wei-Shi Zheng, “LLMDet: Learning Strong Open-Vocabulary Object Detectors under the Supervision of Large Language Models”, Proceedings of the Computer Vision and Pattern Recognition Conference, 2025.