Towards Sim2Real for Shipwreck Detection in Side Scan Sonar Imagery

Advaith Sethuraman
Department of Robotics
University of Michigan
Ann Arbor, Michigan
advaiths@umich.edu

Katherine A. Skinner
Department of Robotics
University of Michigan
Ann Arbor, Michigan
kskin@umich.edu

***Abstract*—In underwater robotics, sonar is an important sensing modality for recognizing objects of interest. Application of deep learning to sonar imagery is difficult due to lack of large sonar datasets with ground truth labels. In this work, we investigate methods for generating large amounts of simulated side scan sonar data for training deep learning models. We perform experiments on simulated data and on real sonar data collected from the field. Results demonstrate potential for leveraging simulated data for sim2real for automated shipwreck detection in side scan sonar imagery.**

***Index Terms*—marine robotics, side scan sonar, segmentation, autonomous underwater vehicles**

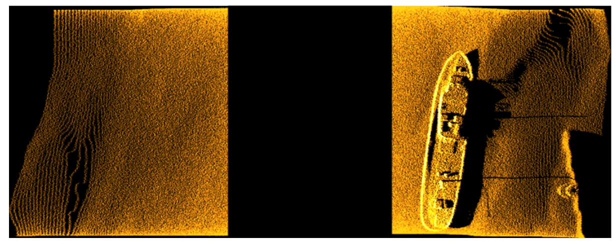

Fig. 1. Simulated side scan sonar data generated by our simulator. The simulation leverages a mesh of a simulated ship and real bathymetric data.

I. Introduction

Marine robotic platforms equipped with side scan sonar can return large amounts of data from automated surveys of subsea environments. Traditionally, sonar imagery is manually inspected by sonar operators. However, this is a time consuming task that requires expert training [1]. It would be optimal to equip robotic platforms with the ability to automatically detect and localize potential objects of interest. This would increase efficiency and decrease cost of data processing for large area marine robot surveys.

Deep learning has shown major performance gains in image processing [2], making it a promising method for automating detection in sonar imagery. However, there is a major problem faced by automated target recognition systems underwater. Traditional supervised deep learning requires large, labeled datasets to achieve optimal performance [3]. Given the difficulty of collecting data underwater, there is a lack of publicly available underwater datasets [4].

This work investigates techniques for generating large amounts of simulated side scan sonar data for sim2real in order to address this gap. We focus on the task of shipwreck detection from side scan sonar data, which is an important task for marine archaeology to ensure preservation of critical maritime assets. Since real datasets with a high number of diverse shipwrecks are scarce, this task is apt for sim2real.

II. Related Work

Prior work has leveraged autonomous underwater vehicles (AUVs) equipped with side scan sonar for surveying submerged archaeological sites [1]. This work developed a pipeline to perform automated detection of potential sites of interest with traditional hand-crafted computer vision methods.

Deep neural networks (DNNs) have proven to outperform hand-crafted and traditional machine learning methods by significant margins across a wide range of tasks [3]. One challenge for applying machine learning to side scan sonar data is the lack of large labeled datasets. Recent work has demonstrated the potential for few shot learning for object detection from marine optical and sonar imagery [5]. These methods aim to effectively learn to represent a class of objects after seeing a few instances of a class. In our work, we instead propose to leverage simulated data to develop learning-based methods for automated shipwreck detection without the need for any real labeled data.

Generating realistic sonar imagery is difficult given the uniquely noisy data from acoustic returns. Furthermore, modeling realistic environments requires consideration of terrain, aquatic life, and geological features. Previous approaches have used depth data from simulators and applied a style transfer network to produce realistic synthetic sonar images [6].

In this work, we instead leverage raytracing [7] to develop a side scan sonar simulator. We train a segmentation network on our simulated data and perform experiments with simulated and real data to demonstrate the potential to leverage our simulated data for training a data-driven model for automated shipwreck detection from side scan sonar imagery.

III. Technical Approach

A. Simulating Side Scan Sonar Imagery

We developed a side scan sonar simulator based on BLAINDER, a Blender-based simulation API [7]. Several modifica-

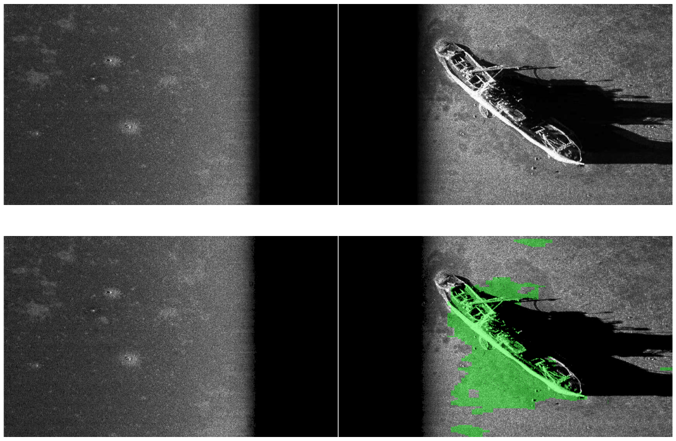

Fig. 2. (top) Real side scan sonar data of the *Viator* shipwreck in Thunder Bay National Marine Sanctuary collected with the Michigan Technological University Great Lakes Research Center IVER-3 AUV. This is input to the network during inference time. (bottom) Output of network inference for shipwreck segmentation. Shipwreck class segmentation is shown in green.

tions were made to more closely simulate collected side scan sonar data. First, an image renderer was created in order to export the sonar raytraces as images instead of the native LAS file format. The rendering process was changed in order to maintain a constant pixels-per-meter resolution in the output image. This resolution parameter was tuned to closely match the qualitative appearance of real side scan sonar data.

We create the simulated environment through randomization of several parameters. First, mesh files of ships and terrain are provided to the system. Mesh files are obtained from TURBOSQUID [8]. Terrain maps are obtained from real bathymetric data recorded from Thunder Bay National Marine Sanctuary in an effort to more closely mimic the testing environment [9]. The ships are replicated based on a desired shipwreck density parameter. They are then placed uniformly randomly on the terrain map. Finally, the roll, pitch, and yaw of each ship is sampled from a uniform distribution $U(-\pi/2, \pi/2)$. These parameters were chosen in an effort to generate diverse scenarios for the wrecks, since the orientation of a ship may change as it falls through the water column. The imagery is generated through a simulated survey in a lawnmower pattern, common in underwater robotics surveys. We vary the survey altitude across simulations. An example of generated imagery is shown in Fig. 1.

B. Deep Learning for Shipwreck Detection

We define our task as an object segmentation task with two classes: shipwreck and terrain. For this task, a U-Net network [10] with a ResNet-34 [2] backbone is used. U-Net networks were developed for biomedical image segmentation and have also been successfully applied to image segmentation. Since sonar imagery is single channel, the first convolutional layer is modified to accept single channel images.

IV. Experiments & Preliminary Results

The network is trained only with simulated data and tested on real data. The Adam optimizer is used with a learning rate of 0.001, $\beta_1 = 0.9$, and $\beta_2 = 0.999$. The model is trained for 800 epochs on a simulated dataset of 5000 images using an 80/20 split for training/validation images. Both the input and output resolutions for the network are 192×192 pixels. Training took 6 hours and 44 minutes on a single NVIDIA GeForce RTX 3080 Ti GPU.

We perform quantitative evaluation on the simulated test set for validation using mean intersection over union (mIOU). The network achieved mIOU of 0.9788 when trained and tested on simulated data.

We also perform experiments on real side scan sonar imagery for qualitative evaluation. These images were collected with Michigan Technological University Great Lakes Research Center's L3 OceanServer IVER-3 AUV equipped with an Edgetech 2205 side scan sonar in Thunder Bay National Marine Sanctuary in Lake Huron, Michigan. Figure 2 shows qualitative results of the network test on real side scan sonar imagery of the *Viator* shipwreck [11]. Note that segmentation outputs within the nadir (sonar deadzone) are removed.

V. Discussion and Future Work

These preliminary results show that training on simulated sonar data is promising for performing automated shipwreck detection on real side scan sonar data collected from AUV platforms. Still, the quality of the segmentation can be further improved.

One direction for potential improvement is to generate more realistic side scan sonar imagery. This simplistic randomization approach for generating shipwreck sites does not capture the true diversity of sites from the field. Factors such as local terrain features, biofouling, and cause of shipwreck yield sites that look vastly different from simulation data. Potential approaches to generating more realistic data include generative networks and neural rendering techniques to learn deformation transforms upon the ship mesh. Alternatively, domain adaptation techniques may enable the network to leverage low-fidelity simulated data and perform domain adaptation to the real sonar data. These directions will be investigated further in future work.

Future work will also include detailed quantitative evaluation on a real side scan sonar dataset collected from the field. This evaluation will be used to find failure modes in the segmentation algorithms to inform new methods for minimizing the sim2real gap for the side scan sonar modality.

VI. Acknowledgements

We would like to thank Professor Guy Meadows and Chris Pinnow of the Great Lakes Research Center at Michigan Technological University for IVER-3 AUV data collection, and Thunder Bay National Marine Sanctuary for supporting field experiments. This work is supported by a University of Michigan Robotics Institute Fellowship and by the NOAA Ocean Exploration program under Award NA21OAR0110196.

References

[1] J. Rutledge, W. Yuan, J. Wu, S. Freed, A. Lewis, Z. Wood, T. Gambin, and C. Clark, "Intelligent shipwreck search using autonomous underwater vehicles," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 6175–6182, 2018.

[2] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[3] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A Large-Scale Hierarchical Image Database," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.

[4] D. Singh and M. Valdenegro-Toro, "The marine debris dataset for forward-looking sonar semantic segmentation," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 3741–3749, 2021.

[5] M. Ochal, J. Vazquez, Y. Petillot, and S. Wang, "A comparison of few-shot learning methods for underwater optical and sonar image classification," in *Global Oceans 2020: Singapore–US Gulf Coast*, pp. 1–10, 2020.

[6] S. Lee, B. Park, and A. Kim, "Deep learning from shallow dives: Sonar image generation and training for underwater object detection," *CoRR*, vol. abs/1810.07990, 2018.

[7] S. Reitmann, L. Neumann, and B. Jung, "BLAINDER—A Blender AI Add-On for Generation of Semantically Labeled Depth-Sensing Data," *Sensors*, vol. 21, no. 6, 2021.

[8] TURBOSQUID, "3D Models for Professionals." https://www.turbosquid.com, Accessed online: 2022.

[9] NOAA National Centers for Environmental Information, "Report for W00319 ." https://www.ngdc.noaa.gov/nos/W00001-W02000/W00319.html, Accessed online: 2022.

[10] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241, Springer, 2015.

[11] Thunder Bay National Marine Sanctuary, "*Viator* ." https://thunderbay.noaa.gov/shipwrecks/viator.html, Accessed online: 2022.