Data Paper

The International Journal of Robotics Research
2024, Vol. 0(0) 1–14
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/02783649241266853
journals.sagepub.com/home/ijr
Sage

Machine learning for shipwreck segmentation from side scan sonar imagery: Dataset and benchmark

Advaith V. Sethuraman[1,*], Anja Sheppard[1,*], Onur Bagoren[1], Christopher Pinnow[2], Jamey Anderson[2], Timothy C. Havens[2] and Katherine A. Skinner[1]

Abstract

Open-source benchmark datasets have been a critical component for advancing machine learning for robot perception in terrestrial applications. Benchmark datasets enable the widespread development of state-of-the-art machine learning methods, which require large datasets for training, validation, and thorough comparison to competing approaches. Underwater environments impose several operational challenges that hinder efforts to collect large benchmark datasets for marine robot perception. Furthermore, a low abundance of targets of interest relative to the size of the search space leads to increased time and cost required to collect useful datasets for a specific task. As a result, there is limited availability of labeled benchmark datasets for underwater applications. We present the AI4Shipwrecks dataset, which consists of 28 distinct shipwrecks totaling 286 high-resolution labeled side scan sonar images to advance the state-of-the-art in autonomous sonar image understanding. We leverage the unique abundance of targets in Thunder Bay National Marine Sanctuary in Lake Huron, MI, to collect and compile a sonar imagery benchmark dataset through surveys with an autonomous underwater vehicle (AUV). We consulted with expert marine archaeologists for the labeling of robotically gathered data. We then leverage this dataset to perform benchmark experiments for comparison of state-of-the-art supervised segmentation methods, and we present insights on opportunities and open challenges for the field. The dataset and benchmarking tools will be released as an open-source benchmark dataset to spur innovation in machine learning for Great Lakes and ocean exploration. The dataset and accompanying software are available at https://umfieldrobotics.github.io/ai4shipwrecks/.

Keywords

Marine robotics, side scan sonar, deep learning, segmentation, benchmark datasets

Received 17 January 2024; Revised 3 April 2024; Accepted 21 May 2024

1. Introduction

It is estimated that over 3 million undiscovered shipwrecks lie on the ocean floor (Gonzalez et al., 2009). Locating these submerged archaeological sites enables research into important maritime assets of historical significance. However, searching over large areas and vast depths of the sea requires expensive and time-consuming surveys, which inhibits new discovery of shipwreck sites. Marine robotic platforms, including autonomous underwater vehicles (AUVs), have demonstrated potential to carry out efficient, cost-effective large-area surveys of marine environments returning hundreds of gigabytes worth of data. Still, the interpretation of sonar imagery to identify sites of interest requires manual expert review. This can take many months to complete, often requiring multiple surveys to verify potential new discoveries.

Automated processing of sonar data collected over large-area surveys has the potential to accelerate the discovery of new sites of interest. On land, machine learning has led to great advances in computer vision and robot perception tasks, including object detection, semantic segmentation, and scene understanding. State-of-the-art machine learning methods rely on labeled datasets for supervised training of neural networks to learn pixel-

[1]Department of Robotics, University of Michigan, Ann Arbor, MI, USA
[2]Great Lakes Research Center, Michigan Technological University, Houghton, MI, USA

*denotes equal contribution

Corresponding author:
Advaith V. Sethuraman, Department of Robotics, University of Michigan, 2505 Hayward St., Ann Arbor, MI 48109, USA.
Email: advaiths@umich.edu

wise segmentation predictions. However, for underwater domains, there is limited availability of public, labeled datasets for sonar data due to challenges, time, and expense associated with data collection (Ochal et al., 2020; Singh and Valdenegro-Toro, 2021).

In this paper, we present a new dataset, AI4Shipwrecks, and we present benchmarking results for segmentation of shipwrecks from side scan sonar (SSS) imagery collected on an AUV (Figure 1). The AI4Shipwrecks dataset contains 286 high-resolution side scan sonar images in .PNG format, with accompanying pixel-wise segmentation labels of shipwreck sites. This dataset was collected over the span of 5 weeks in Thunder Bay National Marine Sanctuary (TBNMS) in Lake Huron, Michigan, USA. Spanning 4300-square-miles, TBNMS contains almost 100 known shipwreck sites and over 100 undiscovered sites. We leverage the unique abundance of known shipwrecks to curate a rich dataset of 28 distinct shipwrecks for investigating the application of machine learning for this task. In this paper, we present details of dataset collection and preparation for easy indexing of our dataset by machine learning pipelines. To encourage further advances in segmentation for SSS, we report extensive results on our dataset from modern, state-of-the-art segmentation models. Lastly, we include discussion on lessons learned from our field expeditions and experiments to provide insight on future challenges and opportunities for machine learning for processing sonar data. The resulting dataset and benchmarking tools will be made publicly available as a benchmark dataset for segmentation from sonar imagery to enable future research in machine learning for ocean exploration.

2. Background

Sonar is a popular perception sensor underwater due to its long operational range (Lin et al., 2023). As a result, sonar image understanding algorithms have the potential to enable autonomous underwater vehicles to operate in unstructured environments.

2.1. SSS imagery versus RGB imagery

Converting the raw acoustic data to sonar imagery makes it more easily viewable by humans. However, there are key phenomena found in SSS imagery that are not present in RGB imagery. SSS imagery primarily measures the intensity of sound returned at a given range. Although it may look similar to an optical birds-eye-view of the underwater terrain, the SSS sensor model is dissimilar to RGB camera models. First, the sonar's beam pattern unevenly distributes energy towards the seafloor, leading to varying image intensities along the horizontal axis (Burguera and Bonin-Font, 2020). Next, SSS is a time-of-flight sensor that exhibits acoustic shadows due to occlusions (Burguera and Bonin-Font, 2020). Finally, the resolution of SSS imagery depends on several parameters such as the altitude of the vehicle, range, and grazing angle, all of which can change within a single image (Grzadziel, 2023).

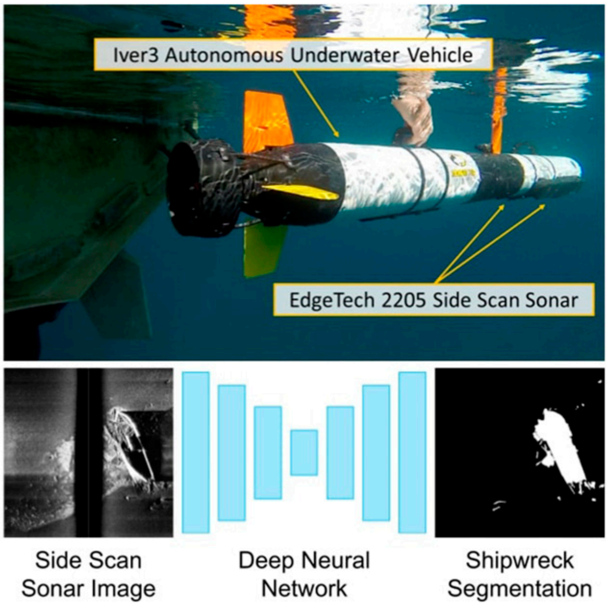

Iver3 Autonomous Underwater Vehicle
EdgeTech 2205 Side Scan Sonar
Side Scan Sonar Image
Deep Neural Network
Shipwreck Segmentation

Figure 1. Our AI4Shipwrecks dataset aims to accelerate the development of shipwreck segmentation algorithms for sonar data collected onboard autonomous systems.

2.2. Object segmentation in sonar imagery

Since many sonar images can be manipulated as single-channel images after post-processing, sonar image understanding algorithms have many similarities to techniques intended for RGB images from computer vision.

Recent work focuses on the application of data-driven deep neural networks for object detection and segmentation in sonar imagery. The majority of work in object detection for sonar imagery involves retraining or fine-tuning existing object detection algorithms on task-specific datasets of sonar imagery (Einsidler et al., 2018; Yang et al., 2022b). The unique waterfall data format of SSS data has also motivated the development of specialized network architectures that enable real-time inference on streamed SSS data (Burguera and Bonin-Font, 2020).

2.3. Datasets

Large, labeled, and publicly available SSS datasets incur high costs of collection and require expert analysis for labeling. As a result, there are fewer options for evaluating the performance of object detection and segmentation algorithms meant for sonars. A summary

Table 1. Table comparing existing publicly available sonar datasets for object detection and segmentation (FLS denotes forward looking sonar, and SSS denotes side scan sonar). Note that our dataset is the only one to offer pixel-level segmentation labels for shipwrecks in a real-world environment.

Dataset	Environment	Sonar Type	# Total Images	# Shipwreck Images	Segmentation Labels
Marine Debris FLS (2021)	Indoor Tank	FLS	1868	0	✓
SeabedObjects-KLSG (2020)	Real World	SSS	1190	385	✗
AURORA (2022)	Real World	SSS	2	0	✗
SeabedObjects-KLSG-II (2022)	Real World	SSS	1296	487	✗
Marine-PULSE (2023)	Real World	SSS	719	0	✗
Burguera and Bonin-Font (2020)	Real World	SSS	10	0	✓
AI4Shipwrecks (ours) (2024)	Real World	SSS	286	161	✓

of existing public sonar datasets for machine learning applications is shown in Table 1.

There has been recent interest in deep learning datasets for forward looking sonars (FLS) (Choi et al., 2021; Santos et al., 2022; Singh and Valdenegro-Toro, 2021; Xie et al., 2022). Singh and Valdenegro-Toro present an FLS dataset of debris in marine environments, but FLS is shorter range and has distinct sensor geometry compared to SSS. Furthermore, the dataset is object-centric and captured in a controlled lab environment.

There are two recent SSS datasets that contain submerged objects, namely SeabedObjects-KLSG (Huo et al., 2020) and Marine-PULSE (Du et al., 2023). However, neither of these datasets contain pixel-wise semantic labels. Bernardi et al. present SSS and AUV data of terrain in the Greater Haig Fras Marine Conservation Zone; however, the dataset does not contain any shipwrecks and also does not have semantic labels (Bernardi et al., 2022). Similarly, the SSS dataset released by Burguera and Bonin-Font (2020) focuses exclusively on seafloor terrain. While this dataset does include pixel-wise labels, the labels are only for terrain types: "sand," "rock," and "other." There are no shipwrecks included in the surveys and the dataset only contains a total of ten waterfall images.

To the best of our knowledge, our AI4Shipwrecks dataset is the first publicly available dataset for shipwreck segmentation in SSS images in the marine autonomy community.

2.4. Synthetic datasets for sonar

Simulation can produce low-cost, diverse, and labeled datasets for sonar image understanding (Cerqueira et al., 2020; Lee et al., 2018; Liu et al., 2021; Sethuraman and Skinner, 2023). Many simulators use ray-tracing techniques to render arbitrary 3D meshes (Cerqueira et al., 2020; Sethuraman and Skinner, 2023), while others use style-transfer techniques to mimic the noise and sensor models of sonars (Lee et al., 2018; Sung et al., 2019). Methods that train solely on simulated data encounter a *sim-to-real gap* between simulated sonar imagery and real imagery, which leads to reduced performance when testing on real data (Sethuraman and Skinner, 2022). Prior work has studied the sim-to-real gap and techniques for reducing it (Sethuraman and Skinner, 2023). Although simulated data can help mitigate training data scarcity, it is still desirable to *evaluate* the performance of vision algorithms on a real test dataset for deployment in the real world. AI4Shipwrecks addresses this gap by leveraging the natural diversity of shipwreck sites in TBNMS. We present a dataset representative of real shipwreck sites an AUV could encounter during autonomous surveys. We also release pixel-wise segmentation labels of shipwreck sites to enable thorough evaluation of machine learning methods for the task of shipwreck segmentation in real environments.

3. Technical approach

Figure 2 provides an overview of the developed pipeline for acquiring and testing sonar data for machine learning applications. First, data is collected through deploying an AUV for a large-area survey. The AUV carries out a pre-programmed survey mission to acquire sonar data. Once the data are returned, post-processing converts the raw sonar data format (.JSF) to standard image format (.PNG). The standard image format is further processed to be input into a deep neural network. The network outputs a prediction in the form of a binary per-pixel segmentation mask, which can be visualized as an overlay on the input sonar image.

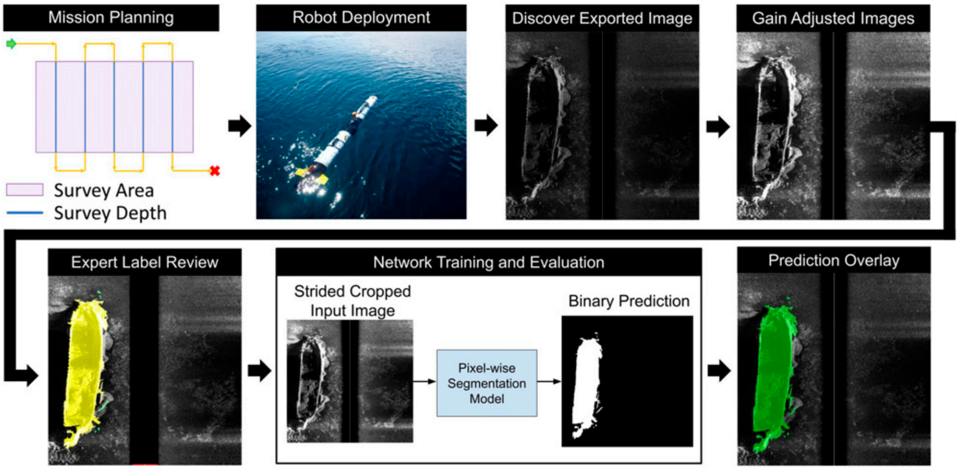


Figure 2. Data acquisition, processing, and network inference pipeline using the Iver3 autonomous underwater vehicle. The yellow lines in Mission Planning denote the AUV's trajectory, with blue segments occurring at survey depth.

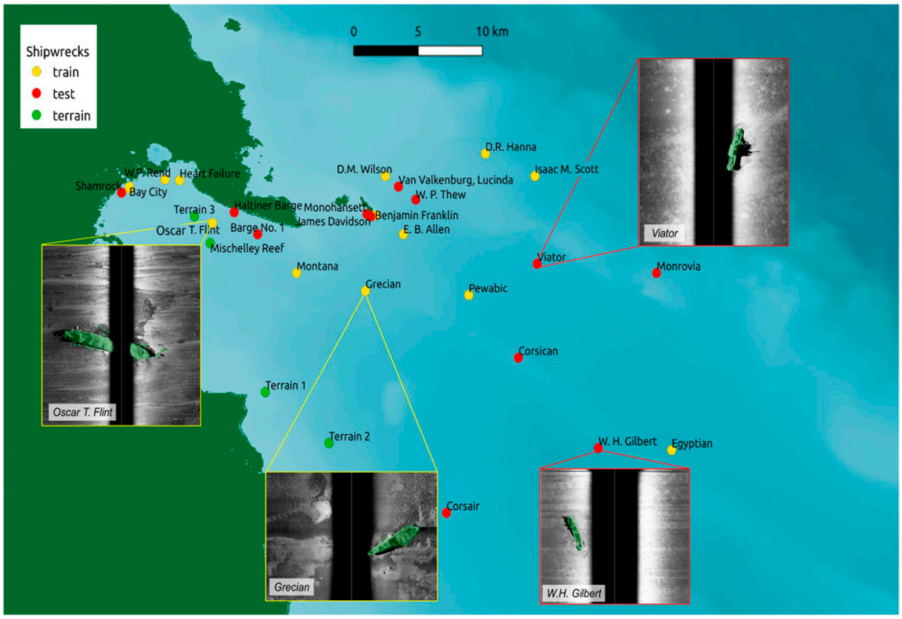


Figure 3. Map of survey sites in TBNMS, Lake Huron, MI. Callouts include example sonar data overlaid with ground truth labels. Color indicates sites that are included in testing (red) and training (yellow) splits, and locations of additional terrain surveys (green). Best viewed in color and zoomed in.

3.1. Site selection

Surveys were conducted in TBNMS in Lake Huron, MI. Figure 3 shows shipwreck sites that were imaged in our surveys during 5 weeks over the course of 2 years. The abundance of known shipwreck targets in TBNMS was a crucial factor in our selection of this field site, as this enabled us to maximize the number of targets observed within a relatively constrained area and short timespan. We timed underwater surveys during late May through early June to mitigate the effects of thermocline. We did not conduct surveys in inclement weather. These proposed survey regions were selected in coordination with scientists at TBNMS to cover a wide range of ship types, site relief, wreck characteristics, and water depth, prioritizing sites within a reasonable distance from the Port of Alpena. This ensured that we could survey a maximum number of sites while still capturing variation across samples, providing a unique and valuable dataset for training machine learning methods. Furthermore, this large and diverse dataset allows us to validate and thoroughly evaluate the accuracy and generalizability of developed methods.

3.2. Data collection platform

The surveys were conducted using the Michigan Technological University (MTU) Iver3 AUV equipped with an EdgeTech 2205 dual frequency 540/1610 kHz SSS and 3D bathymetric system (pictured in Figure 4). Note that the EdgeTech 2205 uses CHIRP and not continuous wave

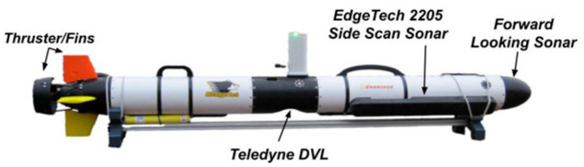


Figure 4. Iver3 data collection platform equipped with advanced localization and high-resolution seafloor mapping capabilities.

pulses (CW), and as a result there is a CHIRP start frequency (C_s) and CHIRP end frequency (C_e). Relevant sensor parameters are reported in Table 2. Each survey mission was conducted by pre-programming a route to survey target sites. The AUV was programmed to capture SSS data while driving at a constant velocity of 2.5 knots. The released imagery is produced from the low frequency sonar, as a lower frequency is able to cover a wider swath of area compared to high frequency sonar. The AUV's Teledyne RDI Explorer Doppler Velocity Log (DVL) provides a long-term accuracy of ± 0.3% ± 0.2 cm/s (Teledyne, 2024). The AUV's DVL has a maximum range of 81 m (Teledyne, 2024). The speed of sound in water is measured by an AML Oceanographic Sound Velocity Sensor, with a range of 1375-1625 m/s, precision of ± 0.006 m/s, accuracy of ± 0.025 m/s, and resolution of 0.001 m/s. The AUV is also equipped with a forward-looking obstacle avoidance sonar.

3.3. Side scan sonar sensor model

The EdgeTech 2205 SSS used in our surveys emits low and high frequency acoustic CHIRPs from two transducers aimed at the seafloor. The signal travels through the water column and reflects off the terrain or other objects in the swath area before being received by the sensor. After a sonar CHIRP has been emitted, the sensor receives and bins the intensity of returns according to time-of-flight. On the horizontal axis (across-track), an SSS image is a histogram of return intensity at equally spaced intervals in time. Each histogram is accumulated in the vertical axis (along-track) as the transducer moves to produce an image with two dimensions. The higher the returned signal intensity, the higher the pixel value in the resulting sonar image. Raw SSS imagery is single channel grayscale but can be viewed with various color palettes for improved visibility.

The sensor model is depicted in Figure 5(a) and illustrates an acoustic beam that encounters an object at point $p(P_x, P_y, P_z)$. SSS typically has two identical transducer arrays mounted symmetrically on either side of the AUV at a fixed tilt angle, θ_t. The first return from the sonar is called the first bottom return (p_{fbr}) and is used as an estimate of AUV altitude. There is typically a sonar deadzone called a *nadir* due to the transducer beam pattern and geometry of the mounted sensor on the AUV. This causes a black stripe down the center of sonar images, as seen in Figure 6.

Although SSS can produce detailed images of the environment, view-dependent shadowing effects, self-occlusion, material-dependent acoustic noise, and distortion make object detection a difficult task for both humans and automated algorithms. Figures 6(d)–(f) illustrate common distortions and noise found in SSS imagery.

Table 2. Table with side scan sonar parameters.

Sensor Parameter	Value
Center Frequency (f)	520 kHz
CHIRP Start Freq. (C_s)	488.5 kHz
CHIRP End Freq. (C_e)	551.5 kHz
Sweep Length (τ)	1 ms
Horizontal Beam Width (θ_h)	0.26 degrees
Survey Speed (V_s)	2.5 knots
Sonar Ranges (R)	30–150 m
DVL Maximum Range	81 m
DVL Long Term Accuracy	±0.3% ± 0.2 cm/s
Sound Velocity Sensor Range	1375–1625 m/s
Sound Velocity Sensor Precision	±0.006 m/s
Sound Velocity Sensor Accuracy	±0.025 m/s

Note sonar range indicates the selected ranges during our surveys, not the min/max ranges for the sensor. The beam width θ_h is two-way.

3.4. Side scan sonar resolution

SSS resolution is dependent on a variety of sensor and environmental parameters (Grzadziel, 2023). These include the horizontal beamwidth (θ_h), sonar range (R) reported in Table 2, and the speed of sound (c). Along-track resolution (R_x) is defined as the resolution in the direction of vehicle travel (vertical image axis):

$$R_x = R \times sin(\theta_h) \tag{1}$$

Similarly, across-track resolution (R_y) refers to the resolution along the range axis (horizontal image axis). This is calculated based on the bandwidth (BW = $C_s - C_e$) of the sonar CHIRP used.

$$R_y = \frac{c}{2 \times BW} \tag{2}$$

Using these formulas, we can calculate the resolutions of sonar imagery captured. We report along-track resolutions at the maximum sonar range and across-track resolutions at the measured speed of sound (c) in Table 3.

3.5. AUV survey mission planning

The Iver3 AUV has two standard patterns for pre-planned missions: lawnmower (LM) and object identification (OID). The LM pattern is a common down-and-back structure, as depicted by the dotted lines in Figure 5(b). In one leg of a LM survey, there are three stages: descent, survey, and ascent. The AUV must ascend, turn around, and then descend before collecting sonar data for each leg in order to geolocalize itself, as global positioning system (GPS) data is not available underwater. The survey occurs at either a

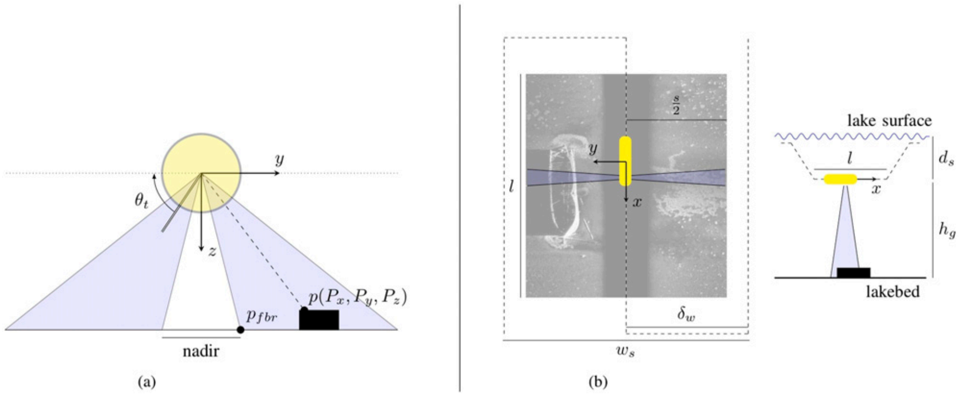
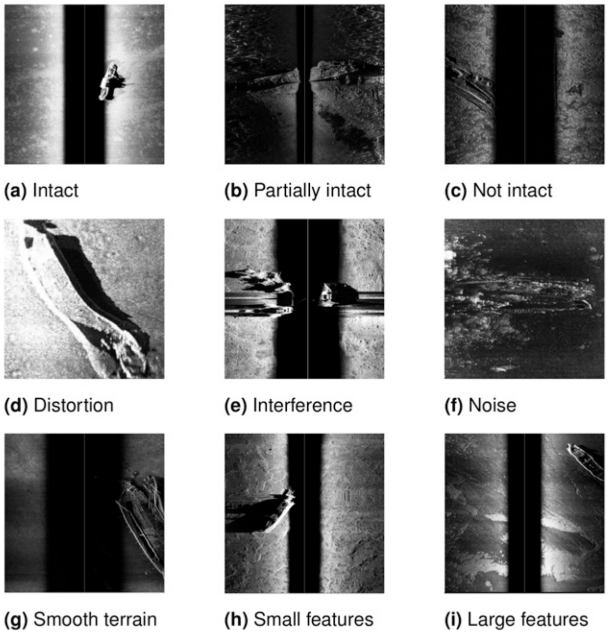


Figure 5. (a) SSS sensor model detailing sensor tilt angle (θ_t), first bottom return (p_{fbr}), nadir gap, and ensonified point on object ($p(P_x, P_y, P_z)$). The SSS field of view is shown in blue. (b) The AUV (in yellow) performs a survey with half-swath width ($s/2$), leg width (δ_w), leg length (l), and total survey width (w_s). For each survey leg, the AUV dives to the depth from surface (d_s) or height over ground (h_g) and resurfaces between legs to acquire a GPS update. The SSS is only collecting data once submerged at depth (d_s) for length (l).


(a) Intact **(b)** Partially intact **(c)** Not intact

(d) Distortion **(e)** Interference **(f)** Noise

(g) Smooth terrain **(h)** Small features **(i)** Large features

Figure 6. Images of wrecks highlighting levels of the three major categories used to split the dataset. Figures 6(a)–(c) represent the shipwreck conditions, Figures 6(d)–(f) represent different sonar quality conditions, and Figures 6(g)–(i) represent the terrain types.

constant height from bottom (h_b) or depth from surface (d_s), and this parameter may change from site to site based on the depth of the wreck. Given that the maximum range of the DVL is 81 m, and the maximum depth of the shipwreck sites surveyed was 79.2 m, we may conclude that bottom-tracking was not lost during repeated descent/ascent stages. The leg length (l) in Figure 5(b) denotes the distance that the AUV was underwater collecting SSS data, and the half-swath width ($s/2$) describes the distance on the seafloor covered by the sonar beams from one of the transducer arrays. For a planned LM mission, w_s is the total survey width and δ_w is the width of each individual leg.

3.6. Post-processing

The Iver3 AUV stores each leg of a sonar survey separately as a .JSF file, which is a proprietary format (EdgeTech JSF Spec., 2023). These files are readable by the free Edgetech Discover software. We used the Discover software for limited sonar image processing and exporting to .PNG. The Discover software normalizes the sonar image intensity and applies Time Varying Gain (TVG). As sound moves through water, absorption loss and spreading cause decreased signal strength. As a result, the returned echoes from objects will also have reduced intensity. TVG addresses this shortcoming and ensures the intensity of the return on the sonar image is not range-dependent (MacLennan, 1986). We do not perform slant-range correction within the Discover software (Chang et al., 2010). No offline post-processing of the .JSF data was performed using GPS information.

Pixel-wise labeling of the post-processed sonar images was conducted by a team of three researchers with shared labeling guidelines. The labels were then reviewed in-depth by a marine archaeologist from the State of Michigan who is an expert on the shipwrecks at TBNMS. There are two labels: "shipwreck" and "other." "Shipwreck" consists of the primary wrecks as well as any debris. If part of the ship is obscured in an acoustic shadow, the label was extrapolated into the shadowed region to follow the expected shape of the wreck based off of expert knowledge. The labels are exported to a standard binary mask format where `0` represents the "other" class and `1` represents the "shipwreck" class.

As survey lengths (l, in meters) can vary for each mission, the resulting sonar images are of different heights (h, in pixels). The dataset provides full-sized images of dimension $h \times w$, where w is image width in pixels. We square-cropped these images into overlapping tiles to be input to deep neural networks.

Images are generated by padding the full-size images with black pixels from $h \times w$ to $h_{\text{padded}} \times w$, according to

Table 3. Detailed survey parameter information including sonar range, leg length, survey depth (from surface) or survey height (from bottom).

Survey	2*Sonar Range (m)	Leg Dist. (m)	Depth (m)	Height (m)	R_x Range (cm)	R_y (cm)
20220523-141805–EBAllenLM	60	110	25	–	13.61	1.16
20220523-154917–EBAllenLM15mDFS	60	110	15	–	13.61	1.15
20220523-185512–Thew15DFS	60	150	15	–	13.61	1.17
20220523-204910–Thew15DFS54oa	60	150	15	–	13.61	1.17
20220524-134748–VanValkenburgLM10mDFS	60	100	10	–	13.61	1.17
20220524-150629–VanValkenburgOID10mDFS	60	150	10	–	13.61	1.13
20220524-161604–WilsonOID2mDFS	60	150	2	–	13.61	1.13
20220524-174036–WilsonBroadSearch2mDFS	60	1000	2	–	13.61	1.13
20220524-201003–MontanaOID8mDFS	60	150	8	–	13.61	1.13
20220527-134851–ViatorOID12mDFS	120	400	12	–	27.23	1.17
20220527-180547–MonroviaOID20mDFS	250	400	20	–	56.72	1.17
20220527-211838–HaltinerLM2mDFS	250	150	2	–	56.72	1.15
20220531-132042–Flint	120	400	2.5	–	27.23	1.15
20220531-145419–Flint_terrain	240	400	2.5	–	54.45	1.16
20220531-162857–Flint_120m	240	400	2.5	–	54.45	1.16
20220531-180338–Heart_Failure	260	200	1	–	58.99	1.17
20220601-143442–BargeNo1	250	600	2	–	56.72	1.16
20220601-171552–Rend	120	200	1	–	27.23	1.16
20220601-185052–Heart_Failure2	120	200	2	–	27.23	1.13
20220602-131155–Grecian	260	250	5	–	58.99	1.15
20220602-141250–Pewabic	240	400	22	–	54.45	1.15
20230605-141930–Egyptian1_100m_DFS45 m_LM	200	400	45	–	45.38	1.14
20230605-173643–Egyptian2_130m_DFS32m_OID	260	300	32	–	58.99	1.14
20230605-193324–Gilbert1_130m_DFS15m_LM	260	200	15	–	58.99	1.13
20230605-202803–Gilbert2_160m_DFS25m_LM	320	400	25	–	72.61	1.13
20230606-181949–NellieGardner1_HFB_LM	120	400	–	3	27.23	1.13
20230606-195902–NellieGardner2_HFB_LM	70	400	–	3	15.88	1.13
20230607-122726–Monohansett1_HFB_LM	100	800	–	3	22.69	1.13
20230607-144904–MischelleyReef1_3mDFS_LM	80	500	3	–	18.15	1.13
20230607-164552–MischelleyReef2_3mDFS_LM	80	400	3	–	18.15	1.13
20230607-175515–NearShore1_2mHFB_LM	60	500	3	–	13.61	1.14
20230608-125411–Hanna1_100m_DFS20m_LM	300	600	10	–	68.07	1.14
20230608-154854–Terrain1_3mDFS_LM	160	2000	2.5	–	36.30	1.14
20230609-130155–Corsican1_100m_DFS5m_LM	300	400	5	–	68.07	1.14
20230609-155043–Corsican2_100m_DFS5m_LM	300	700	5	–	68.07	1.14
20230609-173705–Corsair1_700m_DFS12m_LM	300	500	12	–	68.07	1.14
20230609-194421–Exploratory1_15mDFS_LM	160	2000	15	–	36.30	1.14
20230614-121856–Rend	80	400	1	–	18.15	1.16
20230614-155731–Monohansett1_HFB_LM	70	300	–	3	15.88	1.14
20230614-173130–Davidson	70	420	1	–	15.88	1.14
20230615-124230–Scott1	300	600	15	–	68.07	1.16
20230615-140855–Scott2	300	400	10	–	68.07	1.16
20230615-152709–Wilson-2mDFS	120	600	3	–	27.23	1.16
20230615-165126–VanValkenburg3mDFS	240	600	3	–	54.45	1.17
20230616-125457–Shamrock1_1m_DFS	180	300	1	–	40.84	1.16
20230616-133407–Shamrock2_1m_DFS_2	140	300	1	–	31.76	1.16
20230616-143922–BargeNo1_23	220	600	2	–	49.92	1.16
20230616-155616–BargeNo1_Terrain	200	600	3	–	45.38	1.16
20230616-173938–Haltiner_Terrain	70	600	1.5	–	15.88	1.16
20230616-193812–Haltiner_Bilge	100	75	1.5	–	22.69	1.16

Note that sonar range and leg length are related to the width and height of a captured SSS image in meters respectively. All values were recorded prior to deployment of the AUV and cross-referenced with information from the .JSF Files. Note we are describing each individual survey in this table rather than aggregated per site, as multiple surveys at one location may have had different parameters.

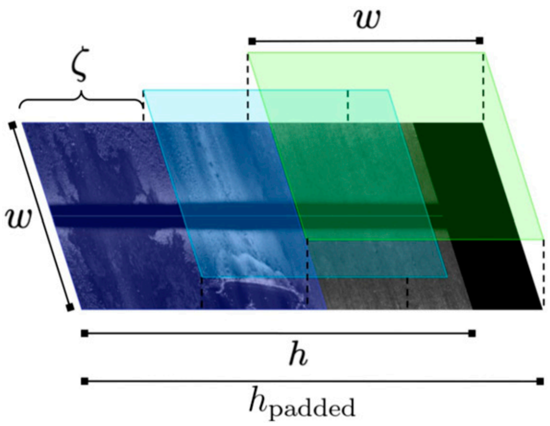


Figure 7. Visualization of the strided cropping of the original side scan sonar image to process before inputting into neural networks. Each colored square represents a $w \times w$ crop of the image, with a stride length of $\zeta = 100$ px.

$$h_{\text{padded}} = h + \zeta - [(h - w) \bmod \zeta] \tag{3}$$

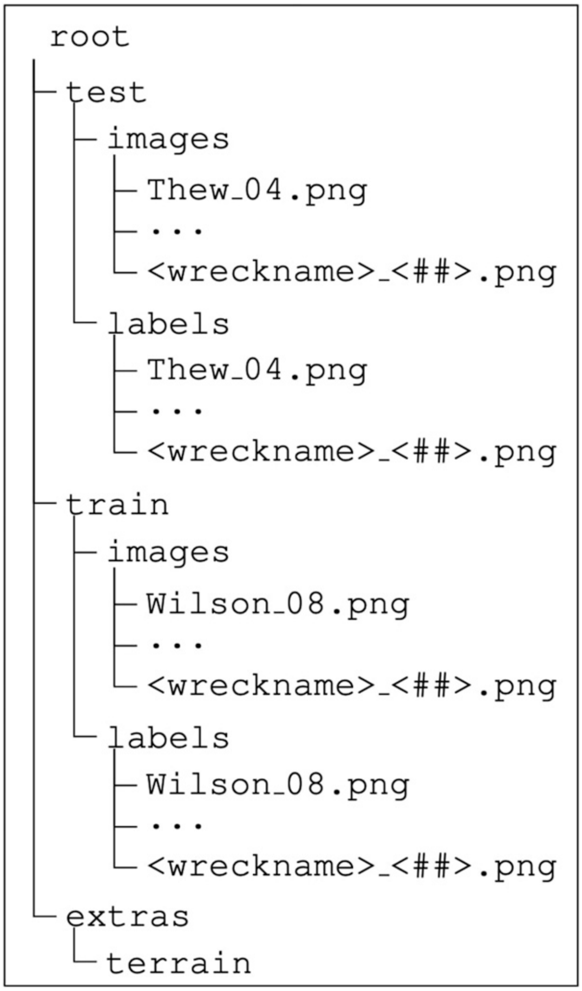
where $\zeta = 100$ is the stride length in pixels. Then, images of dimensions $w \times w$ are cropped every ζ pixels down the length of the padded image (h_{padded}). The image width w output by the Discover software is 1728 pixels, so the resulting square cropped images are of dimension 1728 × 1728. A visualization of this process is shown in Figure 7. The only other data processing step applied to the images is a normalization around 0 based on the mean and standard deviation of the pixel values of the entire dataset.

4. Dataset organization

The dataset is organized by the SSS image and their corresponding labels. We provide the labels as segmentation masks of objects belonging to a shipwreck in the SSS image, including the shipwreck itself and any debris associated with the shipwreck. This includes parts of the shipwreck that may have come off from the shipwreck and parts of the ship that have been dislodged. The images and labels are separated into test and train sets using a 50/50 split. We select a 50/50 split to ensure that the test set is large enough to capture the variance of the shipwrecks sites, ensuring that the robustness of deep learning models can be evaluated through our dataset. In addition, the images are grouped on a per-site basis. This is done to ensure that SSS images that see the same terrain, shipwreck, or debris are not shared between the test and train sets of our dataset.

4.1. Released dataset file structure

In Figure 8, the released data directory structure is shown. The data is split into test and train, according to our iterative approach as described in this paper. Within the test and train directories, the sonar images and their corresponding binary labels are stored. Each sonar image (.PNG format) corresponds to one leg of a survey mission. The site of the survey is indicated in the image file names. We also include a directory with four sites of just surveyed terrain in the `extras` directory. The raw .JSF sonar files—which contain AUV GPS and positioning data—will be released on the NOAA National Centers for Environmental Information database and linked to the project webpage.



```
root
├test
│ ├images
│ │ ├Thew_04.png
│ │ ├...
│ │ └<wreckname>_<##>.png
│ └labels
│   ├Thew_04.png
│   ├...
│   └<wreckname>_<##>.png
├train
│ ├images
│ │ ├Wilson_08.png
│ │ ├...
│ │ └<wreckname>_<##>.png
│ └labels
│   ├Wilson_08.png
│   ├...
│   └<wreckname>_<##>.png
└extras
  └terrain
```

Figure 8. Directory structure of released dataset. Sonar images are sorted into test and train, and named according to their site.

4.2. Train-test split selection

Table 4 provides details of each site, organized by train and test split. Our goal with the train-test split is to ensure that each set provides enough diversity of data for both the training and evaluation of deep learning models. We inform our split selection through both expert-informed and data-driven metrics, with the target of capturing the underlying distribution of the data as uniformly as possible. The atomic unit we use to split up the dataset into train and test sets is each *survey* site location. This was to ensure that no sonar

Table 4. Details of site characteristics, image quality, and number of images per site. Depth refers to the distance from the lake surface to the seafloor at the shipwreck site. Condition is in increasing order with 8 corresponding to most intact. SSS quality is in increasing order with 3 corresponding to highest quality. Terrain corresponds to categorical terrain types as defined in Section 4.2.3.

Name	Depth (m)	Length (m)	Beam (m)	Material	Condition	SSS Quality	Terrain	# Imgs.
Training								
Alpena Steamer†	3.4	–	–	Wood	4	2	C	10
Bay City†	3.4	44.5	8.8	Wood	4	2	C	10
D.M. Wilson	12.2	54.6	9.8	Wood	6	3	B	22
D.R. Hanna	41.1	162.2	17.1	Steel	8	2	A	5
E.B. Allen	30.5	40.8	7.9	Wood	6	3	D	24
Egyptian	76.8	70.7	11.0	Wood	5	2	D	15
Grecian	30.5	90.2	12.2	Steel	7	2	B	5
Harvey Bissell†	4.6	49.4	10.1	Wood	4	2	C	10
Heart Failure	5.5	–	–	Wood	2	1	B	12
Isaac M. Scott	53.3	159.7	16.5	Steel	8	2	C	6
Montana	19.2	71.9	11.0	Wood	6	3	A	8
Oscar T. Flint	9.1	66.4	11.3	Wood	5	2	A	22
Pewabic	50.3	61.0	9.4	Wood	6	2	A	5
W.P. Rend	5.2	87.5	12.2	Wood	5	3	A	11
Testing								
Artificial Reef	3.0	–	–	–	1	1	B	6
Barge No. 1	12.8	94.2	13.4	Wood	5	2	A	15
B. Franklin‡	4.6	41.1	5.8	Wood	2	1	B	10
Corsair	55.5	40.8	7.3	Wood	5	2	C	4
Corsican	48.8	34.1	7.6	Wood	5	1	C	6
James Davidson‡	10.7	30.5	9.1	Wood	4	1	A	10
John F. Warner*	2.7	38.4	7.9	Wood	3	1	B	6
Haltiner Barge	5.2	24.4	10.1	Wood	3	2	A	13
L. van Valkenburg	18.3	39.0	7.9	Wood	5	3	B	19
Monohansett‡	5.5	48.8	9.1	Wood	5	1	B	10
Monrovia	42.7	136.6	17.1	Steel	8	3	A	8
Shamrock*	3.4	44.5	9.1	Wood	2	1	B	6
W.H. Gilbert	77.7	100.0	12.8	Steel	8	3	C	6
W.P. Thew	25.6	40.2	7.3	Wood	5	3	A	17
Viator	57.3	70.7	10.1	Steel	8	2	C	11

*,†,‡ denote shipwrecks that were imaged in a single survey and thus considered as a single "site." Number of images refers to the number of survey legs collected for each site, where each survey leg is processed into one sonar image. Site properties including depth, length, and beam are obtained from Thunder Bay National Marine Sanctuary (2024). We do not present length and beam information for *Heart Failure* because the wreck is very dispersed. Although the main feature of artificial reef is not a wreck, there is one wreck imaged in a single scan at that site, whose length and beam information is unknown. Length and beam information for the *Alpena Steamer* is also not known.

images overlapping the same region are present in both the test and train sets. Please note in Table 4 some sites contain multiple shipwrecks, as indicated by the superscripts. To inform the train-test split selection, we group the sites based on three major categories: shipwreck condition, sonar image quality, and terrain type. For each site, a vector score of length three is assigned based on these quantitative and qualitative metrics. A 50/50 train-test split is then performed using an iterative stratification on our multi-labeled data (Sechidis et al., 2011).

4.2.1. Wreck condition. We leverage domain expert knowledge from scientists at TBNMS in order to assign wreck condition labels. Experts who have conducted dive surveys to inspect these ships provided a scale of 1-8 for wreck condition. There are five sites labeled 8 for complete and intact, one site labeled 7 for complete but collapsed, four sites labeled 6 for semi-complete, nine sites label 5 for partially intact, four sites labeled 4 for fragmented, two sites labeled 3 for fragmented and disarticulated, three sites labeled 2 for widely dispersed debris fields, and one site labeled 1 for non-shipwreck cultural materials.

4.2.2. Sonar image quality assessment. The sonar image quality assessment is done based on well-developed metrics for optical image quality assessment (IQA). We utilize the no-reference, completely blind IQA metric Natural Image Quality Evaluator (NIQE), which constructs a quality score from a natural scene statistic model trained on undistorted images (Mittal et al., 2013). Although there are some sonar quality metrics (Chen et al., 2020; Liu et al., 2023), they generally rely on reference images and measure signal degradation during transmission, which is not suitable for our dataset. There are many challenges in evaluating the quality of a sonar image, the first of which being that many IQA metrics are targeted at measuring quality deterioration from compression. However, our sonar images are not compressed—rather, any distortion in the images is a result of when the Iver3 is surveying shallower sites, making it more vulnerable to drift from waves (Figure 6(d)). If the shipwreck partially falls into the nadir, this can cause significant interference and acoustic shadows that disrupt the geometry of the wreck (Figure 6(e)). Additionally, if the shipwreck falls into the outer edges of the sonar image, the increased noise of more distant acoustic returns causes significant noise in the final sonar image (Figure 6(f)). Informed by the NIQE score, we assign classes of 1–3 to the images where 1 indicates poorer sonar image quality and 3 indicates higher image quality.

4.2.3. Terrain type assessment. We select the terrain type as a category for determining the train-test split, as the varying conditions of terrain can play a crucial role in the detection and segmentation of shipwrecks. This is most prominent in terrains that have a large amount of texture, making it difficult even for humans to discern the shipwreck from the terrain, leading to either false negatives or false positives. We hence aim to have the terrains split evenly based on their characteristics observed from the SSS imagery. In order to determine the most even split across the data based on terrain, we utilize clustering methods.

Specifically, the clustering is done on the N-dimensional latent space of a Variational Auto Encoder (VAE) pre-trained on a large color image (RGB) dataset. The input images are pre-processed by applying white balancing. This is done in order to prevent the clustering based on image intensity; otherwise, the intensity differences across the dataset are too prominent. We use *k*-means clustering, yielding a partitioning on an image-to-image basis. Each image is grouped based on the site it belongs to, which is then used to inform the train-test split. We assign sites a class of A for terrains with small-scale texture, a class of B for terrains with large-scale texture, a class of C for smooth terrain, and a class of D for patchy terrain texture.

4.3. Dataset statistics

The dataset consists of 286 images exported from the Discover software as high-resolution .PNG images. Out of these 286 images, 161 contain shipwrecks. There are 1.49e7 shipwreck pixels and 4.39e11 background pixels. Note that despite the number of images being low, the average dimension of the images is 3480 × 1728 pixels. The width of each exported image is fixed at 1728 pixels. The height varies across the 286 images due to it being dependent on the leg length of each survey. We note that by applying the strided-cropping, as described in Section 3.6, we can obtain 2539 images in the train set and 1722 images in the test set of size 1728 × 1728 pixels. Table 4 provides further details of each site, including material, ship length and ship beam (width) obtained from Thunder Bay National Marine Sanctuary (2024), as well as classification details according to our data categorization for terrain type, sonar image quality, and wreck condition.

5. Experiments and results

5.1. Baseline comparison

We perform benchmark evaluations on a series of deep learning-based segmentation methods to evaluate the ability of state-of-the-art deep learning methods to segment shipwrecks from SSS imagery. Benchmark methods are selected to reflect the variety of deep learning-based segmentation models used in the vision community: Yang et al. (Yang et al., 2022a), ViT Adapter (Chen et al., 2023), DeepLabV3 (Chen et al., 2017), HRNet (Wang et al., 2020), UNet (Ronneberger et al., 2015), and Salient Object Detection (SOD) InSPyReNet (Kim et al., 2022). The model with the longest training time was SOD, which took 14 h on a single NVIDIA A100 GPU with 80 GB of memory. All baselines were trained on 512 × 512 images, which are downsampled from the original square cropped size of 1728 × 1728 described in Section 3.6. Code to replicate our cropping process and to set up a PyTorch dataset can be found in the repository linked on the project website.

We train each baseline according to the training parameters (learning rate, epoch, batch size) suggested by their original papers. This includes the choice of encoder backbones, for which we use the pre-trained backbone associated with the highest performance reported in the original papers. Although we use pre-trained backbones, we train each model from scratch on our dataset. We use the code released from the baseline papers, except for Yang et al., which we re-implemented ourselves following the architecture discussed in (Yang et al., 2022a).

5.2. Evaluation metrics

Each model is evaluated on an aggregate level with Intersection over Union (IoU) for the “shipwreck” class (IoU_{ship}), F1 score, True Positive Rate (TPR), and True Negative Rate (TNR). All metrics are calculated in pixel-space, and are therefore measures of how well network output corresponds to expert opinions in pixel-space. Table 5 provides results across baselines, averaged across each site. Table 6 shows IoU_{ship} per site for each baseline. Segmentation predictions for all six baselines on three example sites are shown in Figure 9. Across this comparison, the SOD model consistently outperformed or performed comparably against all other networks for the aggregate metrics. We performed the subsequent experiment with the SOD model.

5.3. Train set size experiment

Intuitively, a larger training set should lead to better learning conditions for the network and ultimately result in a more accurate model. We conducted an experiment where we gradually increased the size of the train set back up to the full set in order to observe the relationship between train set size and model performance for our dataset. As we increase the train set size, we evaluate model performance on a frozen test set. As shown in Figure 10, we observe that test performance increases with increased training dataset size and the network performance plateaus at around six sites in the train set.

6. Conclusion and future work

This work contributes AI4Shipwrecks, an open-source dataset for comparison of state-of-the-art deep neural networks for shipwreck segmentation from SSS imagery. While recent advances in deep learning have revolutionized the field of computer vision for terrestrial robotics, adoption of similar methods across marine applications is limited due to a lack of widely available data for development and direct comparison of results. We establish a benchmark for semantic segmentation of shipwrecks on the AI4Shipwrecks dataset, and we provide comparison of current state-of-the-art deep neural networks for segmentation. The dataset and code for evaluation will be open-source to enable future research in machine learning for ocean exploration.

The AI4Shipwrecks dataset is a powerful new tool in the field of object segmentation in the wild. Field datasets are often smaller, and can help encourage discoveries in few-shot learning or simulated data generation. A promising direction for future work includes leveraging synthetic data to augment real sonar

Table 5. Aggregate baseline performance averaged across sites: metrics are computed per site and then the average is taken across all test sites for each metric. Metrics include IoU$_{ship}$, F1 Score, TPR, and TNR. TPR and TNR are calculated assuming the *shipwreck* class is positive. ↑ indicates higher is better.

Baseline	IoU$_{ship}$ ↑	F1 Score ↑	TPR ↑	TNR ↑
Yang et al.	0.212	0.310	0.296	0.995
ViT-Adapter	0.283	0.395	0.488	0.995
DeepLabv3	0.363	0.473	0.485	0.996
HRNet	0.372	0.490	0.516	**0.998**
UNet	0.411	0.526	0.592	0.997
SOD	**0.445**	**0.594**	**0.652**	0.997

Table 6. Per-site IoU$_{ship}$ compared across each baseline for each site. Higher is better (↑).

	Artificial Reef	Barge No. 1	Corsair	Corsican	Davidson	W.H. Gilbert	Haltiner Barge	L. Van Valkenburg	Monohansett	Monrovia	Shamrock	W.P. Thew	Viator
Yang et al.	0.002	0.471	0.231	0.056	0.000	0.276	0.031	0.367	0.020	0.478	0.102	0.161	0.563
ViT-Adapter	0.003	0.412	0.542	0.088	0.000	0.006	0.368	0.127	**0.496**	0.469	0.354	0.176	0.631
DeepLabv3	0.003	0.670	0.531	0.042	0.031	0.658	**0.459**	0.645	0.034	0.411	0.151	0.419	0.667
HRNet	0.011	0.703	0.293	**0.174**	**0.067**	0.641	0.448	0.641	0.207	0.566	0.003	0.437	0.646
UNet	0.006	0.736	0.564	0.116	0.074	0.726	0.242	0.641	0.042	**0.572**	**0.428**	0.545	0.646
SOD	**0.017**	**0.775**	**0.583**	0.077	0.039	**0.749**	0.442	**0.713**	0.467	**0.572**	0.001	**0.580**	**0.776**

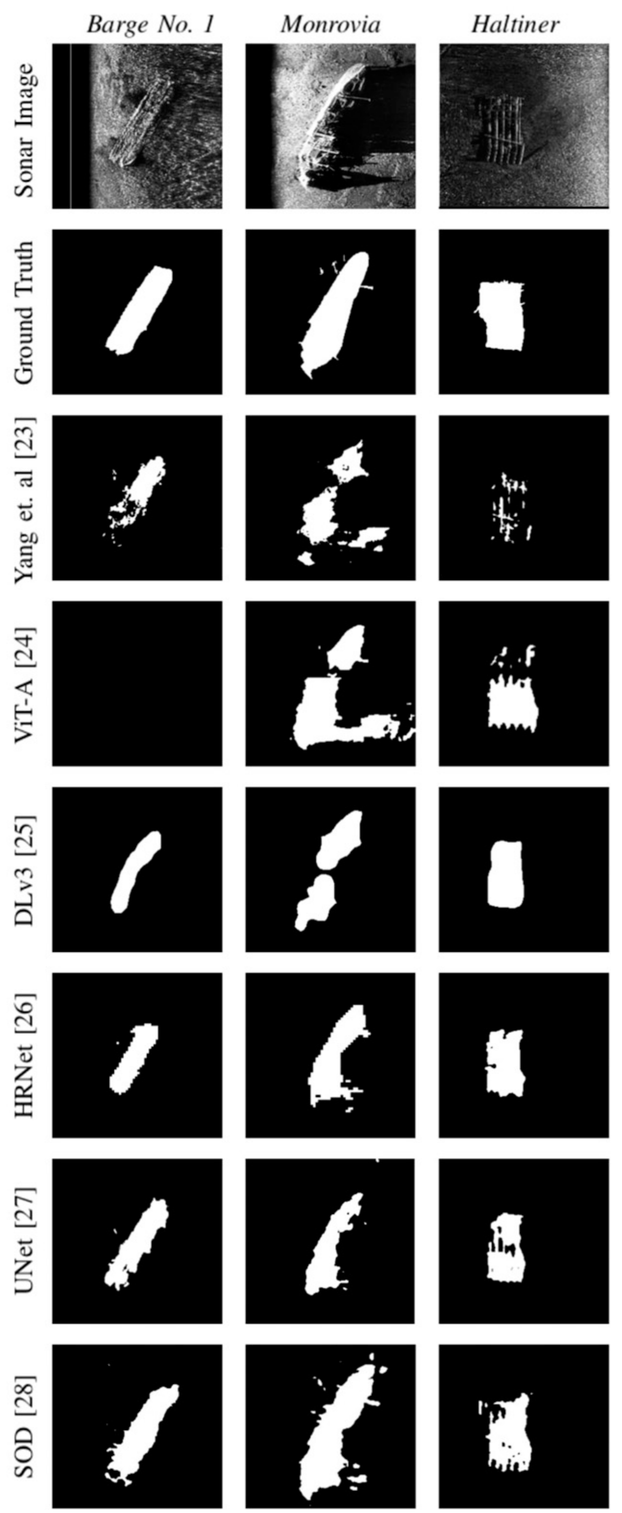


Figure 9. Shipwreck segmentation predictions from the baselines on *Barge No. 1*, *Monrovia*, and *Haltiner*. Zoomed in for better detail.

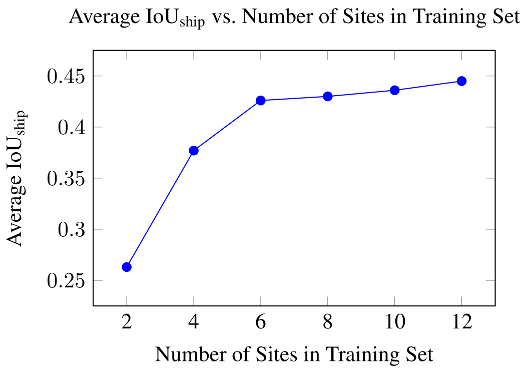


Figure 10. Results of the train set size experiment. Note the plateau in performance after six training sites.

datasets for learning-based detection and segmentation tasks (Lee et al., 2018; Sethuraman and Skinner, 2022). Additionally, AI4Shipwrecks narrows a focus on advancing network architecture to enable deep neural networks to learn from limited training data. Of notable interest is recent work that has demonstrated the potential for few shot learning for object detection from marine optical and sonar imagery (Ochal et al., 2020). Few shot and one shot learning approaches aim to effectively learn to represent a class of objects after seeing a few or single instance of that class, which is ideal for datasets with relatively low abundance of samples per class. We hope the AI4Shipwrecks dataset will enable future work on training from limited training data for marine applications.

Acknowledgments

We would like to acknowledge the lives that were lost in shipwrecks throughout Thunder Bay National Marine Sanctuary. We would also like to thank the sanctuary scientists at Thunder Bay National Marine Sanctuary for supporting field experiments and labeling efforts.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work is supported by the NOAA Ocean Exploration Program under award #NA21OAR0110196.

ORCID iDs

Advaith V. Sethuraman https://orcid.org/0000-0002-0941-319X

Anja Sheppard https://orcid.org/0009-0009-5836-8899

References

Bernardi M, Hosking B, Petrioli C, et al. (2022) AURORA, a multi sensor dataset for robotic ocean exploration. *The International Journal of Robotics Research* 41(5): 461–469. DOI: 10.21227/nnms-te61.

Burguera A and Bonin-Font F (2020) On-line multi-class segmentation of side-scan sonar imagery using an autonomous underwater vehicle. *Journal of Marine Science and Engineering* 8(8): 557. DOI: 10.3390/jmse8080557.

Cerqueira R, Trocoli T, Albiez J, et al. (2020) A rasterized raytracer pipeline for real-time multi-device sonar simulation. *Graphical Models* 111: 101086. DOI: 10.1016/j.gmod.2020.101086.

Chang YC, Hsu SK and Tsai CH (2010) Sidescan sonar image processing: Correcting brightness variation and patching gaps. *Journal of Marine Science and Technology* 18(6): 785–789. DOI: 10.51400/2709-6998.1935.

Chen LC, Papandreou G, Schroff F, et al. (2017) Rethinking atrous convolution for semantic image segmentation. ArXiv: 1706.05587 [cs].

Chen Z, Duan Y, Wang W, Lu T, Dai J and Qiao Y (2023) Vision transformer adapter for dense predictions. *International Conference on Learning Representations*. Available at: https://openreview.net/forum?id=plKu2GByCNW.

Chen W, Gu K, Zhao T, et al. (2020) Semi-reference sonar image quality assessment based on task and visual perception. *IEEE Transactions on Multimedia* 23: 1008–1020. DOI: 10.1109/TMM.2020.2991546.

Choi WS, Olson DR, Davis D, et al. (2021) Physics-based modelling and simulation of multibeam echosounder perception for autonomous underwater manipulation. *Frontiers in Robotics and AI* 8: 706646. DOI: 10.3389/frobt.2021.706646.

Du X, Sun Y, Song Y, et al. (2023) Revealing the potential of deep learning for detecting submarine pipelines in side-scan sonar images: an investigation of pre-training datasets. *Remote Sensing* 15: 4873. DOI: 10.3390/rs15194873.

EdgeTech (2023) JSF file and message descriptions. Available at: https://www.edgetech.com/wp-content/uploads/2023/04/0023492_Rev_K.pdf (accessed January 2024).

Einsidler D, Dhanak M and Beaujean PP (2018) A deep learning approach to target recognition in side-scan sonar imagery. In: OCEANS 2018 MTS/IEEE Charleston, 22–25 October 2018, Charleston, SC, USA, pp. 1–4. DOI: 10.1109/OCEANS.2018.8604879.

Gonzalez AW, O'Keefe P and Williams M (2009) The UNESCO convention on the protection of the underwater cultural heritage: a future for our past? *Conservation and Management of Archaeological Sites* 11(1): 54–69.

Grzadziel A (2023) The impact of side-scan sonar resolution and acoustic shadow phenomenon on the quality of sonar imagery and data interpretation capabilities. *Remote Sensing* 15: 5599. DOI: 10.3390/rs15235599.

Huo G, Wu Z and Li J (2020) Underwater object classification in sidescan sonar images using deep transfer learning and semisynthetic training data. *IEEE Access* 8: 47407–47418. DOI: 10.1109/ACCESS.2020.2978880.

Kim T, Kim K, Lee J, et al. (2022) Revisiting image pyramid structure for high resolution salient object detection. *Proceedings of the 16th Asian Conference on Computer Vision* 16: 108–124.

Lee S, Park B and Kim A (2018) Deep learning from shallow dives: sonar image generation and training for underwater object detection. ArXiv:1810.07990 [cs].

Lin T, Hinduja A, Qadri M, et al. (2023) Conditional GANs for sonar image filtering with applications to underwater occupancy mapping. *2023 IEEE International Conference on Robotics and Automation (ICRA)* 2023: 1048–1054. DOI: 10.1109/icra48891.2023.10160646.

Liu D, Wang Y, Ji Y, et al. (2021) Cyclegan-based realistic image dataset generation for forward-looking sonar. *Advanced Robotics* 35(3–4): 242–254. DOI: 10.1080/01691864.2021.1873845.

Liu J, Pang Y, Yan L, et al. (2023) An image quality improvement method in side-scan sonar based on deconvolution. *Remote Sensing* 15: 4908. DOI: 10.3390/rs15204908.

MacLennan D (1986) Time varied gain functions for pulsed sonars. *Journal of Sound and Vibration* 110(3): 511–522. DOI: 10.1016/S0022-460X(86)80151-1.

Mittal A, Soundararajan R and Bovik AC (2013) Making a completely blind image quality analyzer. *IEEE Signal Processing Letters* 20(3): 209–212. DOI: 10.1109/LSP.2012.2227726.

Ochal M, Vazquez J, Petillot Y, et al. (2020) A comparison of few-shot learning methods for underwater optical and sonar image classification. *Global Oceans 2020: Singapore-US Gulf Coast* 1–10.

Ronneberger O, Fischer P and Brox T (2015) U-net: convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention-MICCAI Proceedings Part III* 18: 234–241.

Santos MMD, De Giacomo GG, Drews-Jr PLJ, et al. (2022) Cross-view and cross-domain underwater localization based on optical aerial and acoustic underwater images. *IEEE Robotics and Automation Letters* 7(2): 4969–4974. DOI: 10.1109/LRA.2022.3154482.

Sechidis K, Tsoumakas G and Vlahavas I (2011) On the stratification of multi-label data. In: *Machine Learning and Knowledge Discovery in Databases*. Berlin Heidelberg: Springer, 145–158.

Sethuraman A and Skinner KA (2022) Towards sim2real for shipwreck detection in side scan sonar imagery. In: *3rd Workshop on Closing the Reality Gap in Sim2Real Transfer for Robotics*, *Robotics: Science and Systems*. New York City, New York.

Sethuraman AV and Skinner KA (2023) Stars: zero-shot sim-to-real transfer for segmentation of shipwrecks in sonar imagery. In: 34th British Machine Vision Conference, 20-24 November 2023, Aberdeen, UK

Singh D and Valdenegro-Toro M (2021) The marine debris dataset for forward-looking sonar semantic segmentation. In: IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), 11-17 October 2021, Virtual.

Sung M, Kim J, Kim J, et al. (2019) Realistic sonar image simulation using generative adversarial network. *IFAC-PapersOnLine* 52(21): 291–296. DOI: 10.1016/j.ifacol.2019.12.322.

Teledyne (2024) Teledyne explorer DVL operational manual. Available at: https://www.teledynemarine.com/en-us/support/SiteAssets/RDI/Manuals_and_Guides/Explorer_DVL/Explorer_Operation_Manual.pdf (accessed January 2024).

Thunder Bay National Marine Sanctuary. Available at: https://thunderbay.noaa.gov/ (accessed January 2024).

Wang J, Sun K, Cheng T, et al. (2020) Deep high-resolution representation learning for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43(1): 3349–3364.

Xie K, Yang J and Qiu K (2022) A dataset with multibeam forward-looking sonar for underwater object detection. *Scientific Data* 9(1): 739. DOI: 10.1038/s41597-022-01854-w.

Yang D, Cheng C, Wang C, et al. (2022a) Side-scan sonar image segmentation based on multi-channel CNN for AUV navigation. *Frontiers in Neurorobotics* 16: 928206.

Yang D, Wang C, Cheng C, et al. (2022b) Semantic segmentation of side-scan sonar images with few samples. *Electronics* 11(19): 3002. DOI: 10.3390/electronics11193002.