Bridging the Domain Gap: Enhancing Underwater Laser Stripe Segmentation with Synthetic Data

Javiera Fuentes-Guíñez
Monterey Bay Aquarium Research
7700 Sandholdt, Moss Landing, CA
jfuentes@mbari.org

Giancarlo Troni
Monterey Bay Aquarium Research
7700 Sandholdt, Moss Landing, CA
gtroni@mbari.org

Abstract

Underwater 3D mapping using low-cost scanning systems relies on accurate laser stripes segmentation. However, the scarcity of annotated data and the variability of underwater environments limit the model's generalization and scalability. To address this issue, we introduce a synthetic dataset specifically designed for laser stripe segmentation. Created using a custom laser scanner module integrated into Blender and the Infinigen procedural generator. The dataset contains over 1,200 high-resolution images across 23 diverse terrains, each with ground truth. We evaluate the impact of synthetic data using a segmentation network trained under different field-to-synthetic data ratios. Our results show that augmenting field datasets with synthetic images significantly improves performance on unseen domains—achieving up to 10% higher recall and 7% higher precision on deep-sea imagery from the Salas y Gómez Ridge, a location with different lighting, seafloor composition, and depth. Our findings highlight the value of synthetic data for domain diversity, reducing annotation costs and enhancing model generalization, supporting broader and more robust deployment of underwater laser mapping systems.

1. Introduction

High-resolution seafloor mapping is crucial for studying ocean dynamics and benthic ecosystems [14]. It typically relies on remotely operated vehicles (ROVs) or autonomous underwater vehicles (AUVs) equipped with sensors such as echo sounders or underwater LIDARs to capture detailed information about underwater terrains. However, these systems are often bulky, expensive, and energy-intensive.

An emerging alternative for underwater 3D reconstruction is the use of one-stripe laser scanners, due to their high precision, energy efficiency, and relatively low cost [1, 6]. The method consists of projecting a laser stripe onto the seafloor and capturing its deformation using a monocular camera. With proper system calibration, 3D coordinates of the seabed surface can be computed from the observed laser stripe. Fig. 1 illustrates the scanning process.

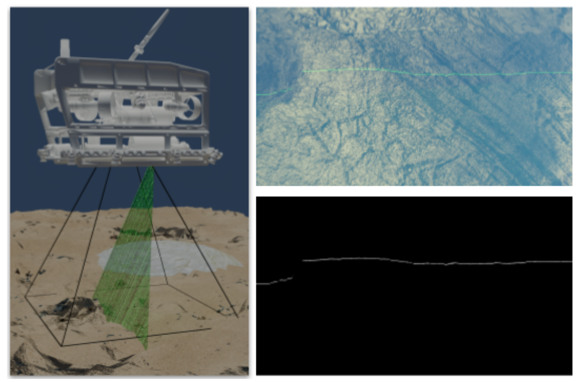


Figure 1. The image shows (left) a simulation of MBARI's MiniROV performing laser scanning, (top right) a simulated image captured during the survey, and (bottom right) the corresponding ground truth mask.

Laser stripe segmentation is a critical step in this process. Existing methods can be broadly divided into two categories: classical approaches based on intensity thresholding [8, 13] or color cues [1, 3]; and more recent approaches using deep learning. Among these, a novel transformer-based architecture has been proposed [5], significantly outperforming classical methods—achieving up to 64% improvement in accuracy under illuminated conditions.

Despite these advancements, deep learning performance remains limited by the size and diversity of the training data. This challenge is especially pronounced in underwater settings, where annotated datasets are scarce due to the high cost of data collection and the labor-intensive nature of labeling high-resolution images [7].

Moreover, underwater imagery varies widely due to differences in terrain, water clarity, sensor characteristics, and lighting—often influenced by depth and vehicle-mounted illumination. For instance, a model trained on data from Monterey Bay, California and a test tank performed well locally but showed a sharp drop in accuracy when tested on data from Easter Island, Chile, where environmental conditions differ significantly. This limited generalization restricts the usability and scalability of the method across diverse marine regions.

To address these limitations and to complement scarce annotated field data, we propose a synthetic dataset designed to improve the robustness and adaptability of underwater laser segmentation systems.

Our main contributions are:

(i) A novel *Underwater Laser Simulation Module*, integrated with Blender and Infinigen, for rapid data generation;

(ii) A 1,200-image dataset spanning 23 underwater terrains, with pixel-level annotations;

(iii) A study quantifying synthetic data's impact on underwater segmentation performance.

The remainder of this paper is organized as follows: Section 2 describes the synthetic data generation pipeline; Section 3 presents the experimental setup; Section 4 discusses the results; and Section 5 concludes the paper.

2. Underwater Laser Simulation Module

2.1. System

Implemented in Python 3.10 within Blender 3.6 [2], our pipeline simulates a monocular camera, laser stripe, and lighting system for underwater scanning.

The camera, enclosed in a rigid camera rig, has a 1920×1080 resolution and is randomly placed at altitudes ranging between 0.5 and $3m$ above the terrain, pointing downward. The camera's focal length and sensor size are user-configurable, allowing realistic replication of specific optical setups.

The laser is modeled using a `SPOT` light rigidly mounted to the camera, with a configurable baseline (between $0.4\,m$ and $0.6\,m$) and either fixed or altitude-dependent inclination. The beam intensity is sampled from 500-1000 lumens to emulate realistic laser output variation. The spot size is constrained to control the spread of the projected laser stripe across the scene - between $30°$ and $45°$.

To render a physically plausible thin laser stripe, we implement a custom node-based shader. This setup isolates the Y-component of the light's surface normal using a `Separate XYZ` node, followed by a `Math > Compare` operation with a narrow threshold (e.g., 1×10^{-4}) to ensure a sharp-edge beam with minimal soft falloff, effectively restricting emission to a single narrow plane.

Two lighting configurations were implemented to emulate different AUV/ROV setups:

- A dual-spotlight system with left and right lights mounted on the center camera rig.
- A quad-light system with four spotlights (forward-left, forward-right, rear-left, rear-right).

Light intensities are randomly selected in the range of 500–1000 W to simulate operational variability and environmental lighting conditions.

2.2. Terrain Generation

The simulated images were generated using two complementary approaches for underwater terrain formation: field 3D models and procedurally generated environments.

2.2.1. MBARI's 3D Survey Models

We used pre-existing 3D models from previous MBARI (Monterey Bay Aquarium Research Institute) seafloor surveys, publicly available on $Sketchfab$[10]. Three distinct models were selected, each representing geological features observed in the Monterey Bay area. These meshes were imported into Blender and placed over a flat sandy plane to provide environmental continuity and context. Additional props such as rocks were added to increase scene complexity.

2.2.2. Procedural Generation with Infinigen

To diversify terrain geometry, we used Infinigen [12], a photorealistic procedural scene generator built on Blender. We extended Infinigen's source code to integrate our custom laser scanner module, enabling automatic placement of the camera-laser rig within procedurally generated scenes. A total of 20 unique terrains were generated using the coral reef configuration, each derived from a different random seed to ensure variation in structural features (Fig. 2).

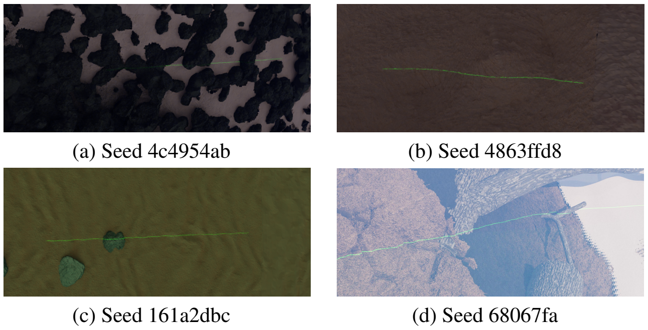
Underwater Visual Effects

Two common optical effects in underwater environments are *absorption* and *scattering* [9]. To simulate these effects, we applied a volumetric absorption shader to the background and surrounding water using Blender's `Principled Volume` node. The volume was tinted with cool blue to approximate the spectral absorption of light underwater. The density of the volume was randomly varied between 0.08 and 0.2 to simulate different water clarity conditions.

To further enhance visual realism, a scattering component was added to the volumetric shader to emulate the diffusion of light caused by suspended particles in the water column.

2.3. Data Collection

To simulate realistic AUV/ROV surveys, a script was developed to capture multiple frames across a $5 \times 5m$ grid, with a step size parameter controlling the spatial resolution



(a) Seed 4c4954ab (b) Seed 4863ffd8

(c) Seed 161a2dbc (d) Seed 68067fa

Figure 2. Samples from Infinigen-based terrains

of the survey. The system follows a standard *mowing-the-lawn* trajectory, starting from the bottom-left corner. At each step, one image is captured, with Gaussian noise added to simulate vehicle motion and localization uncertainty

For each frame, two renders are generated: one with all lights active (RGB) and one isolating the laser, used to create the target segmentation mask. In total, 23 survey scenes produced 1,277 images with corresponding masks.

3. Evaluation Methodology

3.1. Dataset Composition

The simulated dataset consists of 1,277 images generated across 23 distinct underwater survey scenarios. To evaluate the impact of synthetic data on segmentation performance, we also incorporated 564 manually annotated field images. These field annotations were created using Roboflow.

The field dataset is divided as follows:

- **Monterey Canyon, California:** 207 images collected during field deployments using an ROV equipped with a laser scanner.
- **MBARI's Test Tank:** 187 images captured under controlled conditions with varying lighting configurations.

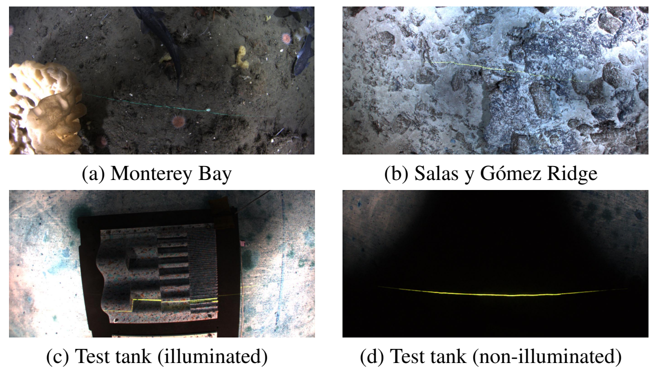
These two subsets form the **training set**, representing environments with similar sensor setups and moderate lighting variability.

To evaluate generalization, a separate **test set** of 156 images was curated from a scientific expedition to the Salas y Gómez Ridge, Chile. These images were acquired at greater depths (approximately 500m deeper than the Monterey sites) using a different vehicle and lighting setup, providing a rigorous test of the model's adaptability to unseen underwater conditions.

Fig. 3 shows representative examples of the field annotated images used in both the training and test sets.

3.2. Network Implementation Details

To assess the impact of the simulated dataset on segmentation performance, we used the same architecture across all experiments: **SPLASH-SegFormer** [5], trained with varying proportions of synthetic and field data. Experiments were run on an NVIDIA GeForce RTX 3090 GPU using Full HD resolution images (1920×1080). For each training run, 90% of the data was used for training and 10% for validation. A OneCycle learning rate scheduler and early stopping were applied to prevent overfitting. The model with the lowest validation loss was saved and evaluated on the test set.



(a) Monterey Bay (b) Salas y Gómez Ridge

(c) Test tank (illuminated) (d) Test tank (non-illuminated)

Figure 3. Samples from field images

3.3. Training Modes

To establish a baseline, we trained three models using only field data: one with the full dataset (Monterey Canyon + test tank), one with only test tank images, and one with only Monterey Canyon images. This allowed us to assess how the training domain affects generalization.

To evaluate the contribution of synthetic data, we trained models under three additional settings: (i) using only synthetic images, (ii) combining synthetic and field data in a 1:1 ratio, and (iii) combining them in a 2:1 ratio. To analyze the effect of underwater visual realism, we repeated the first two settings using a variant of the synthetic dataset with no underwater visual effects. Finally, for reference, we trained a model including the test set (Salas y Gómez Ridge) in the training data. This result is shown as an upper-bound comparison, not used for evaluation.

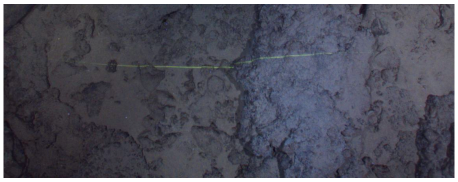
3.4. Metrics

We evaluate the performance over the test set of our model using recall, precision, F1-score and Mean Intersection Over Union (mIoU) [11].

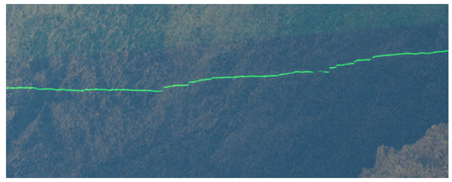
4. Results

4.1. Segmentation Performance with Synthetic Data

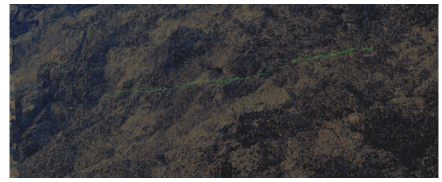
Table 1 presents the performance of each training mode on the Salas y Gómez Ridge test set. Models trained solely on test tank or Monterey data achieved the lowest scores,



(a) Field Image



(b) Synthetic Image without Effects



(c) Synthetic Image with Effects

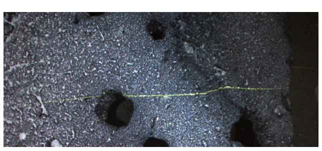
Figure 4. Visual comparison of (left) field image, (center) synthetic image without underwater effects, and (right) synthetic image with visual effects (e.g., volume absorption).

underscoring the limitations of domain-specific training and the importance of dataset diversity. While combining both field datasets led to improved generalization.

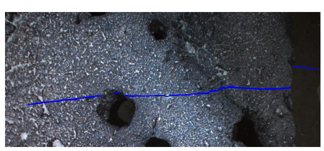
Incorporating synthetic data yielded further gains: a 1:1 synthetic-to-field ratio provided moderate improvements, while a 2:1 ratio resulted in a 10% increase in recall and a 7% increase in precision over the field-only baseline. Ablation experiments further indicate that simulating underwater visual effects enhances performance, as models trained without these effects underperformed, likely due to increased domain shift.

Table 1. Model comparison across key metrics. “No effects” excludes underwater visual effects. The best result is bolded.

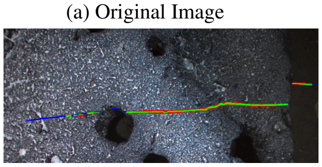
Model	Recall (%)	Precision (%)	F1-score (%)	mIoU (%)
Field-only training				
Test tank	27.31	66.93	34.79	24.53
Monterey	38.57	50.17	40.09	28.16
Previous Field	61.55	66.78	61.23	47.43
Synthetic + Field training				
Sim only	64.46	65.04	61.71	45.80
Sim only No effects	55.63	61.03	58.20	40.78
Ratio 1:1	64.52	70.03	63.80	49.65
Ratio 1:1 No effects	56.59	66.04	59.94	44.03
Ratio 2:1 (Ours)	**71.35**	**71.89**	**68.04**	**53.90**
*Baseline**	84.28	80.53	81.91	69.98



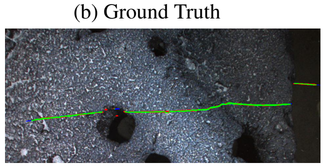
(a) Original Image



(b) Ground Truth



(c) Prediction with Previous Field



(d) Prediction with ratio 2:1

Figure 5. Effect of Synthetic Data on Segmentation Output

Fig. 5 presents qualitative results. Ground truth is shown in blue, true positives in green, and false positives in red. As illustrated, models trained with synthetic data produce predictions more closely aligned with the reference mask.

4.2. Visual Fidelity of Simulated Data

Fig. 4 presents a comparison between a real field image, a synthetic image without underwater effects, and one with simulated effects. The addition of visual effects—such as absorption and scattering—produces images that more closely resemble real underwater conditions. This qualitative similarity supports the improved performance observed when these effects are included during training. To further assess visual fidelity, we compute the Fréchet Distance [4] between feature embeddings from the encoder of the best-performing model (Ratio 2:1) for synthetic and real-world datasets. The resulting matrix shows how simulation helps reduce the domain gap between diverse environments, such as Monterey Canyon and Salas y Gómez Ridge.

Figure 6. Fréchet Distance Matrix Between Domains

5. Conclusion

This work introduces a synthetic data pipeline for underwater laser segmentation, generating over 1,200 high-resolution images with corresponding segmentation masks across 23 terrains. Incorporating synthetic data into training significantly improved segmentation performance in previously unseen environments, highlighting its value as a scalable, cost-effective solution for underwater mapping.

Future work will focus on expanding the dataset, validating generalization across additional real-world sites, and releasing the simulation tools to the research community.

References

[1] Michael Bleier, Joschka van der Lucht, and A. Nüchter. SCOUT3D – An Underwater Laser Scanning System for Mobile Mapping. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W18:13–18, 2019. 1

[2] Blender Online Community. Blender - a 3D modelling and rendering package, 2023. Blender Foundation, Amsterdam. 2

[3] Adrian Bodenmann, Blair Thornton, and Tamaki Ura. Generation of High-resolution Three-dimensional Reconstructions of the Seafloor in Color using a Single Camera and Structured Light. *Journal of Field Robotics*, 34(5):833–851, 2017. 1

[4] Maike Buchin and Bernhard Kilgus. Fréchet distance between two point sets. *Computational Geometry*, 102:101842, 2022. 4

[5] Javiera Fuentes-Guíñez, Giancarlo Troni, and Hans Lobel. SPLASH-SegFormer Pipeline: A Transformer-Based Approach for High-Resolution and Low-Cost Laser Scanner Seafloor Mapping. Accepted to IEEE Robotics and Automation Letters (RA-L), May 2025. 1, 3

[6] Li Kong, Li Ma, Kai Wang, Xiong Peng, and Nan Geng. Three-Dimensional-Scanning of Pipe Inner Walls Based on Line Laser. *Sensors*, 24(11):3554, 2024. 1

[7] Chongyi Li, Chunle Guo, Wenqi Ren, Runmin Cong, Junhui Hou, Sam Kwong, and Dacheng Tao. An Underwater Image Enhancement Benchmark Dataset and Beyond. *IEEE Transactions on Image Processing*, 29:4376–4389, 2020. 1

[8] Haitao Lin, Hua Zhang, Yonglong Li, Jianwen Huo, Hao Deng, and Huan Zhang. Method of 3D reconstruction of underwater concrete by laser line scanning. *Optics and Lasers in Engineering*, 183:108468, 2024. 1

[9] Risheng Liu, Xin Fan, Ming Zhu, Minjun Hou, and Zhongxuan Luo. Real-world Underwater Enhancement: Challenges, Benchmarks, and Solutions, 2019. 2

[10] Monterey Bay Aquarium Research Institute. Sketchfab Profile - MBARI. `https://sketchfab.com/mbari`, 2017. Scientific Organization. 2

[11] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011. 3

[12] Alexander Raistrick, Lahav Lipson, Zeyu Ma, Lingjie Mei, Mingzhe Wang, Yiming Zuo, Karhan Kayan, Hongyu Wen, Beining Han, Yihan Wang, Alejandro Newell, Hei Law, Ankit Goyal, Kaiyu Yang, and Jia Deng. Infinite Photorealistic Worlds Using Procedural Generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12630–12641, 2023. 2

[13] Chris Roman, Gabrielle Inglis, and James Rutter. Application of structured light imaging for high resolution mapping of underwater archaeological sites. In *OCEANS'10 IEEE SYDNEY*, pages 1–9, 2010. 1

[14] Anne-Catherine Wölfl, Henry Snaith, Sam Amirebrahimi, Colin W Devey, Boris Dorschel, Vicki Ferrini, Veerle A Huvenne, Martin Jakobsson, James Jencks, Gary Johnston, et al. Seafloor mapping–the challenge of a truly global ocean bathymetry. *Frontiers in Marine Science*, page 283, 2019. 1