Journal of Computer Science and Electrical Engineering
Print ISSN: 2663-1938
Online ISSN: 2663-1946
DOI: https://doi.org/10.61784/jcsee3087

SEMANTIC SEGMENTATION OF SHIP HULLS FOR UNDERWATER CLEANING ROBOTS

Lu Li#, Wei Fang#, YiGe Shang, Zhi Zhang, MingFu Li, DaoYi Chen*
Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518000, Guangdong, China.
#Lu Li and Wei Fang contributed equally to this work and they are both first authors.
Corresponding Author: DaoYi Chen, Email: chen.daoyi@sz.tsinghua.edu.cn

Abstract: Underwater barnacle cleaning robots require a semantic segmentation algorithm deployable onboard to effectively segment barnacles, obstacles, and background on ship hulls. This study independently constructed a dedicated ship hull semantic segmentation dataset and implemented an improved BiSeNet network. The approach integrates features from both a spatial path and a context path. Building upon the cross-entropy loss function, class weights were strategically assigned during training according to the distribution of different target categories within the dataset. This method achieved significantly high Intersection over Union (IOU) and F1 scores.
Keywords: Spatial features; Contextual features; Semantic segmentation; Ship hull dataset; Weighted loss function

1 INTRODUCTION

During prolonged oceanic transits, submerged hull surfaces inevitably develop complex biofouling communities comprising diverse marine organisms. Crustaceans within the order Cirripedia, particularly barnacles, constitute the most economically consequential fouling organisms due to their exceptional adhesive properties and rapid reproductive cycles. Barnacles form permanent calcareous attachments to hull substrates shortly after larval settlement through proteinaceous cement secreted from specialized glands. Hydrodynamic investigations confirm that extensive barnacle colonization beyond certain threshold levels substantially increases hydrodynamic drag by fundamentally disrupting boundary layer flow regimes. This hydrodynamic resistance directly compromises propulsion efficiency, necessitating considerable engine power augmentation to sustain operational speeds.

Beyond drag-related energy penalties, barnacles induce severe structural deterioration via biocorrosion pathways[1]. Their basal structures release concentrated organic acids that electrochemically compromise protective coatings. This process generates microscale electrochemical corrosion cells where barnacle-covered zones function as anodes relative to adjacent cathodic bare steel, markedly accelerating metal oxidation rates. Industry assessments verify that such biodeterioration significantly curtails vessel service lifetimes, elevates dry-dock maintenance frequency, and generates substantial latent asset depreciation not routinely accounted for in operational expenditures.

Conventional biofouling management predominantly employs hazardous diver operations or scheduled dry-docking—both approaches exhibit critical constraints. Human divers confront safety hazards and limited operational endurance, resulting in restricted cleaning coverage with inconsistent outcomes. Dry-dock procedures, while comprehensive, immobilize vessels for extended periods at considerable expense, triggering logistical disruptions throughout supply network.

Recent advancements in remotely operated and autonomous underwater vehicles have redirected industry attention toward robotic cleaning platforms[2], offering dramatic reductions in human risk exposure and substantial operational efficiency gains. However, achieving genuinely intelligent robotic cleaning requires overcoming perceptual limitations in turbid marine environments. Conventional optical imaging systems suffer severe performance degradation underwater due to light attenuation and particulate scattering, causing traditional edge detection methodologies to misclassify sedimentary particulates as biological fouling boundaries at elevated error rates[3].

This research addresses these sensory constraints through specialized semantic segmentation frameworks, enabling precise real-time differentiation of biofouling, structural obstacles, and background elements under validated turbidity conditions. This capability establishes an automated inspection-cleaning-validation cycle that drastically reduces hull maintenance durations while significantly suppressing biofouling recurrence through optimized cleaning protocols. Collectively, these innovations transform robotic systems from basic mechanized tools into intelligent marine maintenance platforms, delivering fundamental improvements in global maritime operational efficiency and environmental stewardship.

2 METHODOLOGY

2.1 Multi-Level Supervised Dual-Path Network Framework

BiSeNet is a lightweight neural network specifically designed for real-time semantic segmentation[4]. Its dual-path architecture balances spatial detail and contextual semantic information, reconciling speed and accuracy. The core design comprises a Spatial Path (SP) and a Context Path (CP). A Feature Fusion Module (FFM) merges features extracted from these paths. To accelerate convergence, a multi-level supervision strategy was employed, incorporating four auxiliary segmentation heads into the intermediate layers of the Context Path. The auxiliary loss from each head is calculated, contributing to the total loss function, as defined in Formula (1):

$$L_{\text{total}} = L_{\text{main}} + \sum_{\text{i}=1}^{4} L_{aux}^{i} \tag{1}$$

2.2 Weighted Cross-Entropy Loss Function

To address class imbalance (e.g., barnacle pixels being substantially fewer than background pixels), a class weighting strategy was introduced into the standard cross-entropy loss function. This elevates the learning weight for minority classes. The standard cross-entropy loss is expressed mathematically as Formula (3)[5]. The class-weighted variant is given in Formula (4), where C represents the total number of classes, yi is the one-hot encoded ground truth label for each pixel, pi is the predicted probability of the pixel belonging to each class, and wi is the class weight, inversely proportional to the class's pixel frequency ai.

$$w_i = \frac{1}{a_i} \tag{2}$$

$$L_{CE} = -\sum_{\text{i}=1}^{c} y_i \log(p_i) \tag{3}$$

$$L_{CE} = -\sum_{\text{i}=1}^{c} w_i y_i \log(p_i) \tag{4}$$

3 EXPERIMENTAL RESULTS

3.1 Dataset Description

A custom dataset simulating the underwater ship hull environment was constructed. Images were captured using an RGB camera from a simulated robotic perspective. Three target categories were annotated: Background, Barnacle Region, and Hull Obstacle. The original dataset comprised 408 images, annotated using LabelMe. Through augmentation techniques including rotation, cropping, and brightness adjustment, the dataset was expanded to 1,872 images. Pixel distribution analysis revealed: Background (Category 0) covered 52.87%, Barnacle Region (Category 1) covered 39.32%, and Hull Obstacle (Category 2) covered 7.80%. The dataset was split into a training set (1,498 images) and a test set (374 images) using a 4:1 ratio. Representative dataset samples are shown in Figs. 1 to 3.

Figure 1 Image Containing Seawater Background and Barnacle Region

Figure 2 Image Containing Cleaned Hull Surface and Barnacle Region

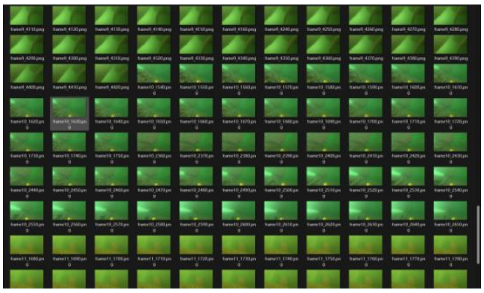

Figure 3 Image Containing Hull Obstacle, Barnacle Region, and Seawater Background

3.2 Analysis of Results

A significant imbalance exists in the pixel proportion of the three target classes within the dataset, as summarized in Table 1.

Table 1 Dataset Class Distribution and Description

Pixel Ratio	Description
52.87%	Clean hull areas & seawater background
39.32%	Barnacle adhesion regions
7.80%	Fins, propellers, etc.

Post-implementation of the class weighting strategy, segmentation accuracy for barnacles and obstacles markedly improved. Performance was evaluated using the F1 score and IoU metrics[6,7]. The F1 score represents the weighted harmonic mean of precision and recall Formula(5), while IoU denotes the average Intersection over Union Formula(6), A represents the predicted area, and B represents the ground truth area.

$$F1\_Score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{5}$$

$$Iou = \frac{A \cap B}{A \cup B} \tag{6}$$

Results before weighting application are detailed in Tables 2 and 3.

Single Scale(SS) denotes an inference approach where predictions are generated using the original image resolution or a single fixed scale, without any multi-scale augmentation.

Multi-Scale Fusion (MSF) involves aggregating predictions from multiple scaled versions of the input image,followed by fusion mechanisms such as averaging or max-voting to consolidate results across scales.

Multi-Scale Fusion with Cropping (MSFC) combines the principles of MSF with image cropping, applying multi-scale fusion to cropped sub-regions to further enhance segmentation accuracy, especially for fine-grained details in high-resolution scenarios.

Table 2 F1 Scores (Baseline - No Class Weighting)

F1 Score	ratio	ss	msf	msfc
cat 0	0.528740	0.852262	0.866208	0.871694
cat 1	0.393239	0.811935	0.845999	0.847779
cat 2	0.078021	0.4704	0.442935	0.531036
macro_F1	~	0.711532	0.718381	0.75017
micro_F1	~	0.818883	0.837874	0.845198

Table 3 IoU Scores (Baseline - No Class Weighting)

Iou	ratio	ss	msf	msfc
cat 0	0.528740	0.742559	0.763993	0.77257
cat 1	0.393239	0.68341	0.733101	0.735779
cat 2	0.078021	0.422108	0.543133	0.532094
mIous	~	0.577834	0.593854	0.623284
fw_mIous	~	0.685358	0.714432	0.72603

Results after applying class weighting are presented in Tables 4 and 5.

Table 4 F1 Scores (With Class Weighting)

F1 Score	ratio	ss	msf	msfc
cat 0	0.528739	0.872693	0.891672	0.891887
cat 1	0.393239	0.846798	0.870601	0.867972

cat 2	0.078022	0.68807	0.74849	0.75926
macro_F1	~	0.80252	0.836921	0.839706
micro_F1	~	0.85155	0.874783	0.874697

Table 5 IoU Scores (With Class Weighting)

Iou	ratio	ss	msf	msfc
cat 0	0.528739	0.77414	0.804521	0.804871
cat 1	0.393239	0.734302	0.770854	0.766741
cat 2	0.078021	0.524472	0.59807	0.611942
mIous	~	0.677638	0.724481	0.727851
fw_mIous	~	0.738994	0.775174	0.774824

Following the implementation of the class weighting strategy,significant gains were observed during single-scale evaluation:Barnacle Region (Category 1): F1 score increased from 0.811 to 0.846; IoU increased from 0.683 to 0.734.
Obstacle Region (Category 2): F1 score increased from 0.470 to 0.688; IoU increased from 0.30 to 0.524.
On the Jetson Orin NX embedded platform, the model inference speed reaches 24 FPS, meeting the real-time control frequency requirements.

3.3 Inference Visualization

Qualitative results generated by the trained model (with weighting) on different test image types are depicted below (see Figure 4-9), confirming effective segmentation of all three classes:

Figure 4 An Image Containing a Cleaned Hull Surface and Barnacler

Figure 5 Segmentation Map of an Image Containing a Cleaned Hull Surface and Barnacler

Figure 6 An Image Containing a Hull Obstacle, Barnacle Region, and Seawater Background

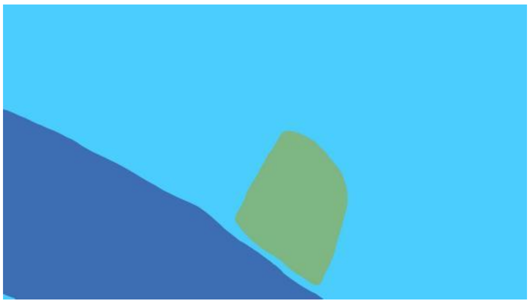

Figure 7 Segmentation Map of an Image Containing a Hull Obstacle, Barnacle Region, and Seawater Background

Figure 8 An Image Containing Seawater Background and Barnacle Region

Figure 9 Segmentation Map of an Image Containing Seawater Background and Barnacle Region

4 CONCLUSION

This research confronts the fundamental technological barrier of precise hull-adhering contaminant segmentation. This precision is essential for deploying truly intelligent underwater cleaning robotics across global maritime operations. It proposes an architecturally refined lightweight semantic segmentation paradigm, anchored in an optimized BiSeNet framework. This framework implements a sophisticated bilateral feature fusion mechanism, which coordinates high-resolution spatial pathways with deep contextual streams. This coordination overcomes inherent precision-context trade-offs in turbid subaquatic environments. The paradigm also strategically deploys dynamically class-weighted loss functions. These functions probabilistically recalibrate learning priorities to counteract extreme categorical imbalances. These imbalances exist among densely aggregated barnacle colonies, sparse structural obstructions, and heterogeneous background interfaces.

Concurrently, it integrates hierarchical multi-level supervision modules. These modules inject auxiliary gradients across complementary feature hierarchies to stabilize convergence pathways and amplify discriminative feature representation capabilities. Collectively, these innovations enable unprecedented pixel-accurate real-time identification of biofouling distributions and hazardous protrusions. This identification occurs across complex curved hull surfaces under real-world visibility constraints. The system sustains real-time computational efficiency, which is essential for continuous hull scanning during robotic transit. Thereby, it establishes an industrial-grade perception backbone for autonomous underwater maintenance systems. These systems are capable of executing millimeter-precision cleaning trajectories around thrusters, sensors, and anodes with zero collision tolerance thresholds. This capability fundamentally redefines vessel husbandry paradigms by transitioning from reactive labor-intensive scrubbing toward strategically optimized biofouling management systems. These new systems directly mitigate billions in global shipping fuel penalties and coating degradation costs annually. They also provide environmentally sustainable alternatives to toxic anti-fouling chemical treatments.

Future evolutionary research vectors explicitly target multi-physics convolutional architectures intrinsically resilient to severe scattering noise prevalent in harbor silting conditions through wavelength-adaptive optical modeling coupled with hybrid optical-acoustic sensor fusion frameworks that synergize RGB imaging with bathymetric laser scanning and ultrasonic thickness mapping to circumvent spectral attenuation limitations. They also advance predictive maintenance cognition through temporal consistency networks that correlate sequential segmentation maps into 4D hull corrosion progression models facilitating preventative structural interventions and lifecycle durability forecasts. Together, these developments collaboratively position next-generation marine robotic platforms as comprehensive vessel health guardians reconciling operational safety compliance, economic sustainability objectives, and oceanic ecosystem preservation imperatives within unified autonomous service frameworks[8-10].

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Trif László, Abdul Shaban, Judit Telegdi. Electrochemical and surface analytical techniques applied to microbiologically influenced corrosion investigation. Corrosion Reviews, 2018, 36(4): 349-363.
[2] Lander Karl. Proactive Cleaning, Maintenance and Inspection of Vessel Hulls via Autonomous Robots. Paper presented at the AMPP Annual Conference + Expo, Denver, Colorado, USA, 2023.
[3] Huang Q, Huang J. Comprehensive review of edge and contour detection: From traditional methods to recent advances. Neural Computing and Applications, 2025, 37(4): 2175-2209.
[4] S Kim, K Jo. BiSeNet with Depthwise Attention Spatial Path for Semantic Segmentation. 2022 International Workshop on Intelligent Systems (IWIS), Ulsan, Korea, Republic of. 2022, 1-4. DOI: 10.1109/IWIS56333.2022.9920717.
[5] N Das, S A Begum. An Empirical Study of Loss Functions for Aspect Category Detection in Imbalanced Data Scenario. 2025 10th International Conference on Signal Processing and Communication (ICSC), Noida, India. 2025, 247-252. DOI: 10.1109/ICSC64553.2025.10968230.
[6] G R S N Kumar, R S Sankuri, S P K Karri. Multi Scale aided Deep Learning model for High F1-score classification of fundus images based Diabetic Retinopathy and Glaucoma. 2023 International Conference on Computer, Electronics & Electrical Engineering & their Applications (IC2E3), Srinagar Garhwal, India. 2023, 1-6. DOI: 10.1109/IC2E357697.2023.10262434.
[7] Y Chen, X Li, Z He, et al. IoU_MDA: An Occluded Object Detection Algorithm Based on Fuzzy Sample Anchor Box IoU Matching Degree Deviation Aware. IEEE Access, 2024(12): 47630-47645. DOI: 10.1109/ACCESS.2024.3375109.
[8] Hennigh O, Narasimhan S, Nabian M A, et al. NVIDIA SimNet™: An AI-accelerated multi-physics simulation framework. International conference on computational science. Cham: Springer International Publishing. 2021, 447-461.
[9] Ferreira I O, Andrade L C, Teixeira V G, et al. State of art of bathymetric surveys. Boletim de Ciências Geodésicas, 2022, 28(01): e2022002.
[10] Wang X, Lin M, Li J, et al. Ultrasonic guided wave imaging with deep learning: Applications in corrosion mapping. Mechanical Systems and Signal Processing, 2022(169): 108761.