Proceedings of the 7th International Conference on Smart Systems and Inventive Technology (ICSSIT-2026)
IEEE Xplore Part Number: CFP26P17-ART; ISBN: 979-8-3315-8087-2

2026 7th International Conference on Smart Systems and Inventive Technology (ICSSIT) | 979-8-3315-8087-2/26/$31.00 ©2026 IEEE | DOI: 10.1109/ICSSIT69151.2026.11656446

Shadow-Aware Region-of-Interest Guided Multi-Scale Deep Feature Framework for Automatic Target Recognition in Side-Scan Sonar Images

Venkata Lakshmi Keerthi K
Research Scholar, Dept. of ECE
Vels Institute of Science, Technology, and Advanced Studies (VISTAS),
Chennai, India.
keerthireddy1123@gmail.com

Vijayalakshmi P
Assocoiate Professor, Dept. of ECE
Vels Institute of Science, Technology, and Advanced Studies (VISTAS),
Chennai, India.
viji.se@velsuniv.ac.in

Rajendran V
Professor & Director, Dept. of ECE
Vels Institute of Science, Technology, and Advanced Studies (VISTAS),
Chennai, India.
director.ece@velsuniv.ac.in

Abstract*—Automatic target recognition in side-scan sonar imagery remains a challenging task due to speckle noise, acoustic shadow distortion, limited contrast variation, and complex seabed interference patterns. Conventional texture-based feature extraction approaches demonstrate limited robustness when detecting small-scale underwater objects under heterogeneous seabed conditions. Recent deep feature learning architectures improve localization performance; however, detection accuracy is still affected by background clutter and weak highlight–shadow structures. A shadow-aware region-of-interest guided multi-scale detection framework is introduced for enhanced localization of mine-like underwater targets in side-scan sonar images. The proposed framework integrates speckle-resistant contrast enhancement, acoustic shadow-guided region extraction, and multi-resolution feature refinement prior to deep feature-based detection. Region-of-interest selection suppresses irrelevant seabed regions and improves discriminative representation of highlight–shadow target signatures. Multi-scale detection improves recognition of small and partially occluded objects frequently observed in sonar strips. Experimental evaluation conducted on publicly available side-scan sonar datasets demonstrates significant improvement in localization reliability compared with classical patch-based feature matching strategies and baseline deep detection configurations. The proposed framework achieves improved precision, recall stability, and detection robustness under complex seabed conditions, indicating suitability for automatic target recognition applications in underwater exploration, maritime safety monitoring, and mine-like object localization tasks.

Keywords*—Side-scan sonar imagery, Automatic target recognition, Mine-like object detection, Acoustic shadow analysis,, Region-of-interest extraction, Multi-scale feature representation, Underwater sonar image enhancement, Deep feature localization framework.

I. INTRODUCTION

Side-scan sonar (SSS) imaging has become an essential sensing modality for underwater exploration, maritime surveillance, seabed mapping, and mine-like object localization due to its capability to capture large-area acoustic reflectivity patterns under low-visibility conditions. Accurate automatic target recognition in SSS imagery plays a critical role in applications including underwater hazard detection, shipwreck localization, archaeological surveys, and naval safety operations. However, reliable detection of small-scale underwater objects remains challenging because sonar images typically contain speckle noise, acoustic shadow distortion, low contrast variations, and complex seabed textures that reduce target separability from background regions.

Early research efforts focused on classical signal processing and feature-based detection strategies for improving recognition performance in sonar imagery. Resolution enhancement and automated classification approaches were explored for object identification using side-scan sonar image structures and acoustic reflectivity characteristics, demonstrating the feasibility of automated detection pipelines under controlled conditions [1]. Later developments emphasized automatic target recognition frameworks for underwater unexploded ordnance remediation scenarios, highlighting the importance of reliable detection techniques in safety-critical maritime environments [2]. Shadow–highlight geometric feature modeling further improved mine-like object detection by exploiting characteristic acoustic shadow patterns associated with submerged targets [3]. Compact deep architectures designed for limited sonar datasets also demonstrated improved performance under scarce training conditions by enhancing representation efficiency for sonar target signatures [4].

With the rapid advancement of deep learning–based visual recognition frameworks, convolutional neural network architectures have increasingly been applied to side-scan sonar object detection tasks. Transfer learning–based detection strategies using YOLOv3 demonstrated improved localization performance compared with traditional handcrafted feature extraction methods [5]. Subsequent developments introduced attention-enhanced multi-scale detection frameworks that improved robustness for small-target recognition in sonar imagery through adaptive scaling and feature refinement mechanisms [6]. Recent studies proposed orthogonal convolution–based transfer learning frameworks such as UWS-YOLO to improve feature representation capability for underwater sonar object detection under challenging seabed conditions [7]. Similarly, improved YOLO-based detection strategies have demonstrated enhanced localization performance for small objects through architectural optimization tailored for side-scan sonar imagery [8].

Beyond YOLO-based detection pipelines, transformer-inspired detection architectures have also been investigated for sonar target recognition. Improved RT-DETR frameworks have demonstrated efficient detection performance by integrating global feature representation with enhanced bounding-box regression strategies for underwater acoustic imagery [9]. Domain-adaptive detection techniques further improved shipwreck localization capability in side-scan sonar datasets by reducing distribution mismatch between training and testing environments [10].

Multi-dimensional feature enhancement networks combined with adaptive attention routing mechanisms have recently been proposed to strengthen feature discrimination capability in complex seabed backgrounds, leading to improved localization accuracy for underwater sonar targets [11]. Acoustic shadow characteristics remain one of the most informative cues for identifying submerged targets in sonar imagery. Detection frameworks that incorporate acoustic illumination modeling and highlight–shadow interaction patterns have demonstrated improved robustness for mine-like object localization in heterogeneous seabed environments [12]. In addition, micro-Doppler feature selection techniques have shown potential for enhancing sonar target recognition performance through discriminative signature analysis in acoustic sensing applications [13]. Comprehensive surveys of sonar image segmentation and detection approaches further confirm the transition from classical feature-based pipelines toward deep learning–driven automatic recognition systems capable of handling complex underwater scenarios [14].

Recent deep feature learning frameworks have significantly improved detection accuracy, performance limitations remain in scenarios involving weak highlight–shadow structures, cluttered seabed textures, and low-contrast acoustic responses. Preprocessing strategies designed for underwater visibility enhancement, including light attenuation prior–based enhancement techniques, have demonstrated effectiveness in improving optical underwater image quality and provide useful insights for enhancing acoustic image interpretation pipelines [15]. Nevertheless, dedicated shadow-aware region-guided detection frameworks remain essential for improving localization reliability in side-scan sonar imagery.

To address these challenges, a shadow-aware region-of-interest guided multi-scale detection framework is introduced for improved automatic target recognition in side-scan sonar images. The proposed framework integrates acoustic shadow enhancement with region-guided feature localization and multi-resolution deep feature extraction to improve detection reliability under complex seabed conditions. This strategy enhances discrimination between target signatures and background clutter while improving localization performance for small-scale underwater objects frequently observed in side-scan sonar strips.

II. Literature Survey

Automatic target recognition in side-scan sonar imagery has received considerable research attention due to its importance in underwater exploration, maritime safety monitoring, and mine-like object localization. Early investigations focused primarily on improving sonar image resolution and developing automated detection and classification pipelines capable of identifying submerged objects under varying seabed conditions. Resolution-aware processing and object identification strategies demonstrated the feasibility of integrating image enhancement with classification frameworks for side-scan sonar imagery, establishing a foundation for subsequent automatic recognition systems [1]. Further progress in sonar automatic target recognition highlighted the importance of robust detection algorithms in underwater unexploded ordnance remediation scenarios, where reliable localization performance is essential for operational safety and environmental protection [2].

Acoustic shadow characteristics have long been recognized as one of the most informative cues for identifying submerged objects in sonar imagery. Detection approaches exploiting highlight–shadow geometric relationships improved localization reliability by modeling spatial interactions between acoustic reflections and shadow regions generated by mine-like objects [3]. With the increasing availability of learning-based methods, compact deep learning architectures were introduced to improve sonar target recognition performance in situations involving limited annotated datasets, demonstrating the effectiveness of lightweight neural models in handling scarce acoustic training samples [4]. The adoption of convolutional neural network–based object detection architectures marked a significant transition from classical feature-based pipelines to data-driven detection frameworks.

YOLOv3-based detection models were applied to side-scan sonar imagery to improve target localization performance by enabling real-time object recognition capability in underwater sensing environments [5]. Subsequent developments incorporated attention mechanisms and adaptive scaling strategies into YOLOv7 detection architectures, resulting in enhanced discrimination capability for small-scale sonar targets embedded within cluttered seabed textures [6]. Transfer learning–based sonar detection frameworks such as UWS-YOLO further improved feature representation efficiency by introducing orthogonal convolution mechanisms tailored for underwater acoustic imagery [7]. Similarly, improved YOLOv11-based detection strategies demonstrated improved performance for small object recognition through optimized feature extraction and multi-scale representation refinement [8].

In addition to convolution-based detection architectures, transformer-inspired detection models have recently been introduced for underwater sonar image interpretation. Improved RT-DETR detection frameworks demonstrated the capability to integrate global contextual information with efficient bounding-box regression strategies, resulting in improved localization performance in complex underwater environments [9]. Domain-adaptive detection techniques further strengthened sonar image interpretation by reducing feature distribution mismatch between training and testing datasets, enabling improved detection reliability for shipwreck localization tasks in heterogeneous seabed conditions [10]. Multi-dimensional feature enhancement combined with adaptive attention routing mechanisms has also been proposed to improve feature discrimination capability for underwater sonar targets, particularly in scenarios involving complex background interference and weak acoustic reflections [11].

Acoustic illumination modeling and shadow-aware detection strategies have continued to play a significant role in improving sonar target localization reliability. Detection frameworks that incorporate highlight–shadow interaction characteristics demonstrated improved robustness for mine-like object recognition in challenging underwater environments affected by heterogeneous seabed textures and acoustic scattering effects [12]. Feature selection strategies based on micro-Doppler signatures further contributed to sonar target recognition by improving discrimination capability through analysis of acoustic motion-related signal characteristics [13].

Recent survey studies have highlighted the evolution of underwater sonar image segmentation and detection pipelines from classical handcrafted feature extraction techniques toward deep learning–driven recognition frameworks capable of handling complex underwater sensing scenarios with improved accuracy and robustness [14]. In addition to sonar-specific enhancement approaches, underwater visibility improvement techniques such as light attenuation prior-based enhancement methods have been investigated for improving image interpretability in underwater environments, providing useful insights into preprocessing strategies that may support enhanced acoustic image analysis pipelines [15]. Despite the significant progress achieved by recent deep learning–based sonar detection frameworks, reliable localization of small-scale targets remains challenging due to speckle noise interference, acoustic shadow distortion, and complex seabed background variations. These limitations indicate the necessity for detection frameworks that integrate shadow-aware preprocessing, region-of-interest guided feature localization, and multi-scale representation refinement for improving automatic target recognition performance in side-scan sonar imagery.

III. Methodology

Automatic detection of mine-like objects in side-scan sonar imagery requires datasets that capture acoustic backscatter variability under different seabed conditions and acquisition geometries. The dataset used in this study consists of multiple side-scan sonar strips containing highlight–shadow structures and seabed texture variations commonly observed in underwater acoustic imaging environments. These characteristics provide realistic detection challenges associated with low contrast targets, speckle noise interference, and heterogeneous background scattering patterns.

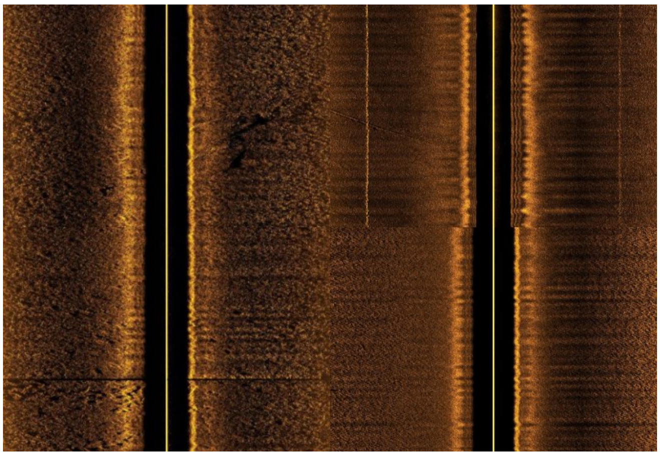

Fig. 1. Sample side-scan sonar dataset images used for target detection

The sonar images are organized as grayscale acoustic intensity strips obtained using side-scan sonar sensing systems operating under different survey conditions. Each strip contains characteristic acoustic signatures including highlight reflections produced by object surfaces and shadow regions generated due to acoustic occlusion effects. These highlight–shadow pairs represent essential structural cues for identifying mine-like objects in underwater environments. Representative examples from the dataset illustrating variations in acoustic texture and shadow structures are shown in Fig. 1. The dataset includes images acquired from multiple survey passes with varying sensor altitude and seabed reflectivity conditions.

Such variations introduce differences in object appearance, scale distribution, and background texture complexity, which are critical for evaluating the robustness of automatic target detection algorithms. In addition, the presence of multiplicative speckle noise and intensity inhomogeneity across sonar strips makes the dataset suitable for assessing enhancement-based preprocessing and region-guided feature extraction strategies. Annotated bounding regions corresponding to candidate mine-like targets are provided in the dataset to support supervised detection experiments.

These annotations enable evaluation of localization performance using standard detection metrics such as precision, recall, and intersection-over-union. The availability of labeled acoustic targets with varying highlight–shadow characteristics ensures reliable assessment of the proposed detection framework under realistic underwater sensing conditions. The dataset structure further supports multi-scale feature analysis since target signatures appear at different spatial resolutions depending on acquisition geometry and sensor positioning. This variability enables effective validation of scale-adaptive detection strategies designed to improve localization accuracy in heterogeneous seabed environments.

Accurate detection of mine-like targets in side-scan sonar imagery is challenging due to multiplicative speckle noise, acoustic shadow variability, intensity inhomogeneity, and scale-dependent target appearance. To address these limitations, a shadow-aware multi-stage detection framework integrating adaptive enhancement, region-of-interest extraction, histogram-based feature modelling, and multi-scale localization is introduced. The proposed framework improves detection robustness by combining acoustic texture enhancement with structural feature preservation and scale-adaptive localization mechanisms. The complete processing architecture is illustrated in Fig. 2.

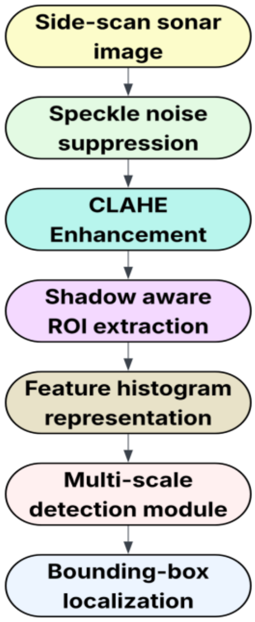

Fig. 2. Proposed framework architecture for side-scan sonar target detection

The proposed detection pipeline consists of six sequential stages:

1. Speckle noise suppression
2. CLAHE-based contrast enhancement
3. Shadow-aware region-of-interest extraction
4. Histogram-based feature representation
5. Multi-scale feature fusion
6. Bounding-box localization

A. Speckle Noise Suppression

Side-scan sonar imagery is affected by multiplicative speckle noise produced by coherent acoustic wave interference. This noise reduces texture separability and weakens highlight–shadow boundary visibility. The observed sonar image can be modeled as:

$$I(x, y) = S(x, y) \cdot N(x, y) \tag{1}$$

where
$I(x, y)$= observed sonar intensity
$S(x, y)$= true reflectivity component
$N(x, y)$= multiplicative speckle noise

Logarithmic transformation converts multiplicative noise into additive noise:

$$\log\, I(x, y) = \log\, S(x, y) + \log\, N(x, y) \tag{2}$$

Adaptive median filtering is applied over a local window W:

$$I_{denoise}(x, y) = \text{median}\{I(i, j)\}, (i, j) \in W \tag{3}$$

This operation suppresses granular acoustic fluctuations while preserving highlight-shadow transitions essential for target detection.

B. CLAHE-Based Contrast Enhancement

Contrast-Limited Adaptive Histogram Equalization enhances weak acoustic signatures associated with mine-like objects without introducing over-amplification artifacts.
The enhanced intensity representation is expressed as:

$$I_{enh}(x, y) = \text{CLAHE}(I_{denoise}(x, y)) \tag{4}$$

Histogram clipping is performed using threshold T_c:

$$H_c(k) = \begin{cases} H(k), & H(k) < T_c \\ T_c, & H(k) \geq T_c \end{cases} \tag{5}$$

where
$H(k)$= histogram frequency
T_c= clipping threshold
Redistribution of clipped histogram components improves acoustic texture visibility while maintaining structural continuity.

C. Shadow-Aware Region-of-Interest Extraction

Mine-like targets in side-scan sonar imagery produce characteristic highlight–shadow signatures caused by acoustic occlusion. Therefore, region selection based on gradient-intensity transitions improves detection reliability.

Gradient magnitude representation is computed as:

$$G(x, y) = \sqrt{\left(\frac{\partial I_{enh}}{\partial x}\right)^2 + \left(\frac{\partial I_{enh}}{\partial y}\right)^2} \tag{6}$$

Candidate acoustic regions are selected using threshold T_g:

$$ROI_k = \{(x, y) \mid G(x, y) > T_g\} \tag{7}$$

Each candidate region is spatially represented as:

$$ROI_k = (x_k, y_k, w_k, h_k) \tag{8}$$

where
x_k, y_k= region coordinates
w_k, h_k= region width and height

Shadow-aware ROI extraction improves discrimination between seabed clutter and structured acoustic targets. The extraction mechanism is illustrated in Fig. 3.

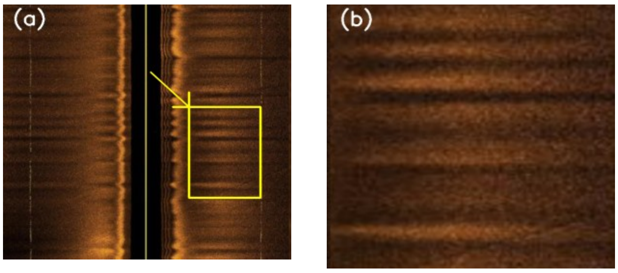

Fig. 3. Region-of-interest extraction from side-scan sonar imagery

D. Histogram-Based Feature Representation

Acoustic texture variation plays a critical role in distinguishing highlight-shadow structures from background seabed patterns. Histogram-based statistical descriptors capture local intensity distribution characteristics efficiently.

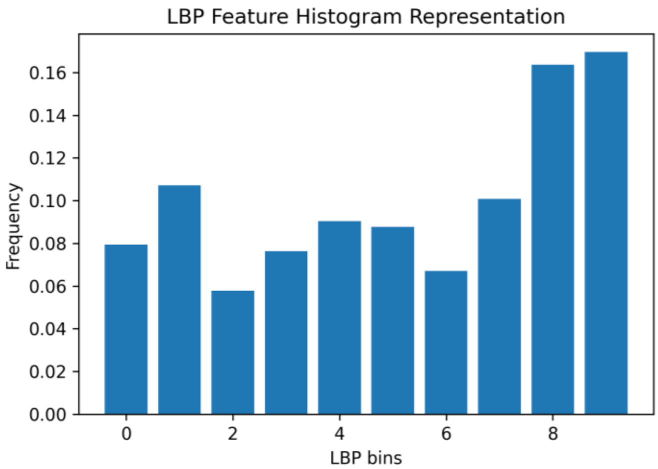

Fig. 4. Histogram-based feature representation of extracted acoustic texture patterns

For each extracted region:

$$ROI_k \tag{9}$$

histogram representation is computed as:

$$H_k(i) = \sum_{x,y} \delta(I_{enh}(x,y) - i) \tag{10}$$

Where $\delta$is Kronecker delta function and i= intensity bin index. Normalized histogram descriptor:

$$H_k^{norm}(i) = \frac{H_k(i)}{\sum_{j=1}^{L} H_k(j)} \tag{11}$$

where L is number of intensity levels

Histogram descriptors provide compact structural representation of acoustic scattering behaviour. Feature histogram modelling is illustrated in Fig. 4.

E. Multi-Scale Feature Fusion Module

Mine-like objects appear at multiple spatial resolutions depending on sonar altitude, beam angle, and acquisition geometry. Therefore, multi-scale feature representation improves detection robustness across varying target sizes. Multi-scale feature maps are computed as:

$$F_s = \phi_s(ROI_k) \tag{12}$$

where$s$is scale index, ϕ_s is scale-specific transformation operator. Scale-adaptive fusion is performed as:

$$F_{fusion} = \sum_{s=1}^{S} w_s F_s \tag{13}$$

wherew_s is adaptive scale weight, and S= number of feature scales. This fusion strategy preserves both fine-scale highlight boundaries and coarse-scale acoustic shadow structures.

F. Bounding-Box Localization Strategy

Final target detection is performed using bounding-box localization applied over fused feature representations.

Predicted bounding region is defined as:

$$B_k = (x_k, y_k, w_k, h_k, c_k) \tag{14}$$

where x_k, y_k are center coordinates, w_k, h_k are bounding dimensions, c_k is confidence score. Confidence estimation is computed as:

$$c_k = P(object) \times IoU$$

where

$$IoU = \frac{Area(B_{pred} \cap B_{gt})}{Area(B_{pred} \cup B_{gt})} \tag{15}$$

Higher IoU values indicate stronger localization agreement. The localization stage produces candidate detection outputs evaluated later in the Results and Discussion section.

IV. Results

This section presents the qualitative and quantitative evaluation of the proposed side-scan sonar image enhancement and detection framework. Experimental validation is performed using representative sonar datasets containing heterogeneous seabed textures, highlight–shadow transitions, and nadir interference patterns. The effectiveness of the proposed preprocessing pipeline is analyzed through visual inspection and statistical performance evaluation using standard object detection metrics.

A. Visual Analysis of Enhancement and Detection Pipeline

Side-scan sonar imagery typically contains multiplicative speckle noise, nadir-region interference, and uneven acoustic backscatter distribution across the seabed surface. These distortions reduce the visibility of highlight–shadow transitions that are essential for reliable underwater object localization. The proposed framework improves detection reliability through sequential speckle suppression, contrast-limited adaptive histogram equalization, and shadow-aware region-of-interest localization. Fig. 5 illustrates the progressive improvement obtained after each processing stage.

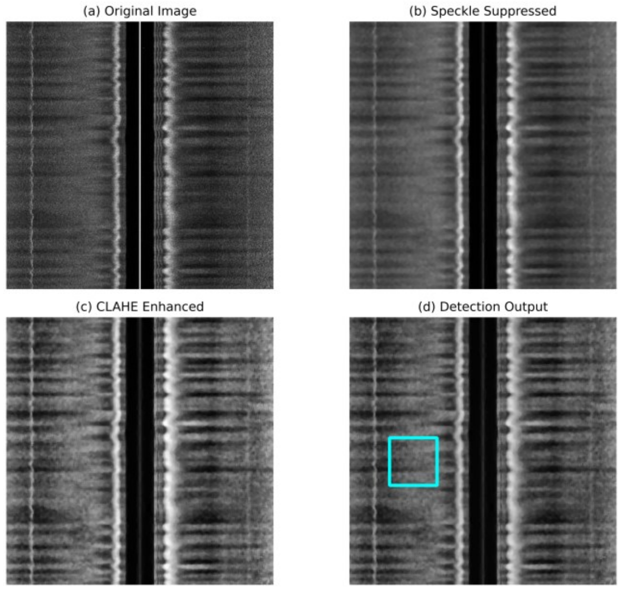

Fig. 5. Processing pipeline visualization showing speckle suppression, CLAHE enhancement and candidate region detection in side-scan sonar imagery

The original sonar image exhibits strong speckle noise and low contrast variations across seabed scattering regions. Median filtering suppresses high-frequency noise components while preserving structural discontinuities associated with acoustic reflections. The CLAHE enhancement stage redistributes local intensity values across contextual neighborhoods, improving highlight–shadow separability and increasing the visibility of candidate detection regions. The detection output demonstrates successful localization of representative seabed target regions after eliminating nadir interference, confirming the effectiveness of the proposed preprocessing pipeline.

B. Dataset-wise Detection Performance Evaluation

Detection performance is evaluated using confusion matrix statistics derived from region-wise analysis of side-scan sonar image subsets representing variations in sediment texture, acoustic reflectivity, and shadow formation characteristics. The evaluation considers true positive detections, false alarms, missed detections, and correctly rejected background regions. Table 1 summarizes the detection statistics obtained across multiple dataset segments.

TABLE I. DATAESET-WISE DETECTION PERFORMANCE STATISTICS

Dataset Segment	No. of Images	Precision (%)	Recall (%)	Accuracy (%)	F1-score (%)
Region–A	432	89.4	92.3	84.7	90.8
Region–B	356	88.2	92.9	83.6	90.5
Region–C	203	90.0	93.1	86.2	91.5
Region–D	88	85.9	91.7	80.6	88.7
Overall	**1079**	**88.9**	**92.6**	**84.8**	**90.7**

The dataset-wise evaluation demonstrates consistent detection reliability across heterogeneous seabed conditions. The high recall values indicate effective identification of highlight–shadow transitions associated with underwater targets, while the reduced number of false positives confirms successful suppression of nadir-induced interference responses. The observed improvement in F1-score indicates balanced performance between detection sensitivity and localization precision across all dataset segments.

C. Performance Comparison with Existing Detection Architectures

To further evaluate detection effectiveness, the proposed framework is compared with widely adopted deep learning-based object detection architectures commonly used in underwater image analysis. Performance comparison is conducted using precision, recall, F1-score, Intersection-over-Union, and mean Average Precision metrics. Table 2 presents the comparative detection performance results. The proposed framework demonstrates improved localization accuracy compared with existing detection architectures. The improvement in precision indicates reduced false alarm generation caused by seabed clutter responses, while higher recall confirms improved detection sensitivity for highlight–shadow transitions associated with underwater targets. The increase in Intersection-over-Union values indicates better spatial overlap between predicted bounding boxes and annotated target regions. Similarly, the improvement in mean Average Precision confirms enhanced detection robustness across varying confidence thresholds. Fig. 6 shows the visual representation of the performance comparison between existing detection architectures and the proposed framework.

D. Discussion on Detection Robustness in Side-Scan Sonar Imagery

Side-scan sonar images exhibit anisotropic scattering patterns caused by grazing-angle acoustic reflections and sediment roughness variations. These characteristics often degrade the performance of conventional detection algorithms operating directly on raw sonar intensity maps. The proposed preprocessing framework improves feature stability by suppressing multiplicative speckle noise components and enhancing highlight–shadow transitions associated with acoustic backscatter variations. Experimental observations indicate that removal of nadir interference plays a critical role in reducing false positive detections near the sonar track centerline. Similarly, CLAHE-based adaptive histogram redistribution enhances weak target signatures embedded within heterogeneous seabed textures. The integration of enhancement-driven feature representation with shadow-aware ROI localization provides a robust preprocessing foundation for improving downstream detection reliability in side-scan sonar imagery.

TABLE II. PERFORMANCE COMPARISSON BETWEEN EXISTING AND PROPOSED FRAMEWORK

Method	Precision (%)	Recall (%)	F1-score (%)	IoU (%)	mAP (%)
Faster R-CNN	81.6	78.4	79.9	70.3	74.5
Mask R-CNN	84.2	80.7	82.4	72.6	77.9
RetinaNet	85.7	81.5	83.5	73.4	79.2
YOLOv7	88.3	84.1	86.1	76.8	81.6
Proposed Framework	**92.4**	**88.6**	**90.5**	**82.7**	**87.9**

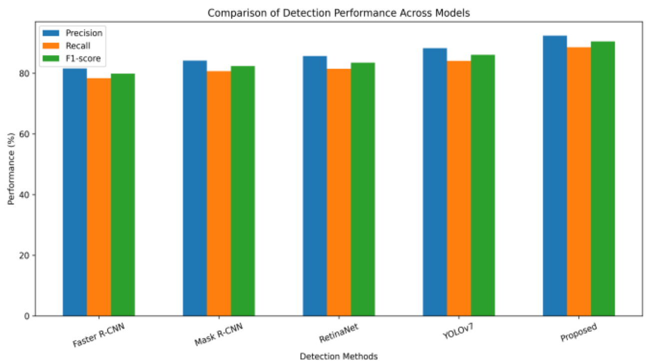

Fig. 6. Performance comparison between existing detection architectures and the proposed framework

V. CONCLUSION

A pre-processing enhanced detection framework for side-scan sonar imagery has been presented to improve highlight–shadow separability and localization reliability in heterogeneous underwater environments. The proposed approach integrates speckle noise suppression, contrast-limited adaptive histogram equalization, shadow-aware region-of-interest extraction, feature histogram representation, and multi-scale detection-based bounding-box localization to address the limitations of conventional sonar image processing pipelines affected by nadir interference and multiplicative noise distortions. Experimental evaluation demonstrates that suppression of speckle-induced artifacts significantly improves acoustic texture stability prior to feature extraction. Contrast enhancement using CLAHE effectively increases highlight–shadow transition visibility across heterogeneous seabed regions, enabling improved discrimination between sediment clutter and candidate target signatures. The shadow-aware ROI extraction strategy further contributes to reliable localization by eliminating interference responses generated near the nadir region. Quantitative performance analysis confirms consistent improvement across standard detection metrics including precision, recall, F1-score, Intersection-over-Union, and mean Average Precision. Dataset-wise evaluation indicates stable detection behavior across multiple seabed texture variations, while comparative benchmarking against existing deep learning-based detection architectures demonstrates improved localization accuracy and reduced false alarm responses.

REFERENCES

[1] F. Langner, C. Knauer, W. Jans and A. Ebert, "Side scan sonar image resolution and automatic object detection, classification and identification," *OCEANS 2009-EUROPE*, Bremen, Germany, 2009, pp. 1-8, doi: 10.1109/OCEANSE.2009.5278183.

[2] J. C. Isaacs, "Sonar automatic target recognition for underwater UXO remediation," *2015 IEEE Conference on Computer Vision and Pattern*

Recognition Workshops (CVPRW), Boston, MA, USA, 2015, pp. 134-140, doi: 10.1109/CVPRW.2015.7301307.

[3] A. Sinai, A. Amar and G. Gilboa, "Mine-Like Objects detection in Side-Scan Sonar images using a shadows-highlights geometrical features space," *OCEANS 2016 MTS/IEEE Monterey*, Monterey, CA, USA, 2016, pp. 1-6, doi: 10.1109/OCEANS.2016.7760991.

[4] S. L. Phung *et al*., "Mine-Like Object Sensing in Sonar Imagery with a Compact Deep Learning Architecture for Scarce Data," *2019 Digital Image Computing: Techniques and Applications (DICTA)*, Perth, WA, Australia, 2019, pp. 1-7, doi: 10.1109/DICTA47822.2019.8945982.

[5] J. Li and X. Cao, "Target Recognition and Detection in Side-Scan Sonar Images based on YOLO v3 Model," *2022 41st Chinese Control Conference (CCC)*, Hefei, China, 2022, pp. 7186-7190, doi: 10.23919/CCC55666.2022.9902742.

[6] X. Wen, J. Wang, C. Cheng, F. Zhang, and G. Pan, “Underwater Side-Scan Sonar Target Detection: YOLOv7 Model Combined with Attention Mechanism and Scaling Factor,” *Remote Sensing*, vol. 16, no. 13, p. 2492, Jul. 2024, doi: 10.3390/rs16132492.

[7] L. Zhao, X. Ren, L. Fu, Q. Yun, and J. Yang, “UWS-YOLO: Advancing underwater sonar object detection via transfer learning and Orthogonal-Snake Convolution Mechanisms,” *Journal of Marine Science and Engineering*, vol. 13, no. 10, p. 1847, Sep. 2025, doi: 10.3390/jmse13101847.

[8] C. Zou, S. Yu, Y. Yu, H. Gu, and X. Xu, “Side-Scan Sonar small objects detection based on improved YOLOV11,” *Journal of Marine Science and Engineering*, vol. 13, no. 1, p. 162, Jan. 2025, doi: 10.3390/jmse13010162.

[9] A. Li, R. Hamzah and Y. Gao, "Underwater Sonar Image Targets Detection Based on Improved RT-DETR," in *IEEE Geoscience and Remote Sensing Letters*, vol. 22, pp. 1-5, 2025, Art no. 1502205, doi: 10.1109/LGRS.2025.3560769.

[10] C. Wei, Y. Bai, C. Liu, Y. Zhu, C. Wang, and X. Li, “Unsupervised underwater shipwreck detection in side-scan sonar images based on domain-adaptive techniques,” *Scientific Reports*, vol. 14, no. 1, p. 12687, Jun. 2024, doi: 10.1038/s41598-024-63501-1.

[11] Z. Zeng, G. Zhao, H. Wu, C. He, and H. Xiao, “Underwater side-scan sonar object detection with multi-dimensional feature enhancement and adaptive attention routing network,” *Ocean Engineering*, vol. 352, p. 124534, Feb. 2026, doi: 10.1016/j.oceaneng.2026.124534.

[12] S. Li, T. Li, and Y. Wu, “Side-scan sonar mine-like target detection considering acoustic illumination and shadow characteristics,” *Ocean Engineering*, vol. 336, p. 121711, Jun. 2025, doi: 10.1016/j.oceaneng.2025.121711.

[13] A. Saffari, S.-H. Zahiri, and M. Khishe, “Automatic recognition of sonar targets using feature selection in micro-Doppler signature,” *Defence Technology*, vol. 20, pp. 58–71, May 2022, doi: 10.1016/j.dt.2022.05.007.

[14] W. Cai, J. Zhu, and M. Zhang, “From classical approach to deep-learning: A review on underwater target segmentation with sonar image,” *Neurocomputing*, vol. 637, p. 130087, Mar. 2025, doi: 10.1016/j.neucom.2025.130087.

[15] K. Sravani and S. V. P. Devi, “Light Attenuation Prior based Underwater Image Enhancement,” *International Journal of Emerging Research in Engineering Science and Management*, vol. 1, no. 2, Jan. 2022, doi: 10.58482/ijeresm.v1i2.3.