Object Detection in Side Scan Sonar Images using Local Binary Pattern features

Venkata Lakshmi Keerthi.K[1], Vijayalakshmi.P[2*], Rajendran. V[3]

[1]Research Scholar, Department of Electronics and Communication Engineering, Vels Institute of Science, Technology, and Advanced Studies (VISTAS), Chennai, India, keerthireddy1123@gmail.com
[2]Associate Professor, Department of Electronics and Communication Engineering, Vels Institute of Science, Technology, and Advanced Studies (VISTAS), Chennai, India, viji.se@velsuniv.ac.in
[3]Professor & Director, Department of Electronics and Communication Engineering, Vels Institute of Science, Technology, and Advanced Studies (VISTAS), Chennai, India, director.ece@velsuniv.ac.in

Detecting underwater objects using side-scan sonar images is highly complex and challenging. Despite several object detection techniques developed in the last decade, the absence of public datasets and reliance on collected datasets have made the process even more intricate. Several deep neural network-based object detection techniques were developed, but they show optimal results due to fewer datasets. This paper proposes side-scan image object detection using patch feature extraction and matching. Initially, patch features are extracted using the Local binary patterns (LBP). Further patch feature matching between non-overlapping patch features and collected side-scan sonar database images is performed. The maximum scored patches are merged to locate the object in the side-scan sonar image. Finally, the proposed and state-of-art techniques are applied to the collected database. The proposed object detection precision score is 81.6%, and the speed is 28.7 frames/ second, which shows robustness against state-of-the-art techniques.

Keywords: Side-scan sonar, object detection, feature extraction, feature matching, underwater object detection.

1. Introduction

Underwater object detection using side-scan sonar (SSS) and synthetic aperture sonar (SAS) systems is one of the complex tasks in computer vision. Applications such as Shipwrecks, airplane crashes, human body, and other mine-like object detection in the SSS is an essential tool [1]. During the scanning process, manual annotation and detection of objects is cumbersome and time-consuming due to huge scanning images, and objects may or may not

exist. To overcome difficulties, computer vision methods for object detection in SSS images have been developed with automatic target recognition (ATR) techniques. Mine-like objects (MLOs) detection in SSS images is essential for the military for safety and cost. The haar-like classifier methodology is used for object detection by rapid serial visual presentation (RSVP) of image chips [2]. Further, object recognition is improved by using a fine-tuned Support Vector Machine (SVM) with Haar-like features. Another methodology is introduced to detect mines underwater by segmenting the fish sonar images using Amplitude dominant component analysis (ADCA) [3]. Narrow band components are obtained using bandpass filters to analyze the SSS images. Gabor filters are utilized to extract spectral and optimum localization properties. Amit et al. (2016) developed image super-resolution and feature-matching techniques for identifying mine-like objects [4].

Shadow reduction techniques are introduced to improve the detection of submerged objects [4]. These techniques are based on shadow region separation by C fuzzy clusters by extracting pixels. Further, spare encoders are introduced for learning Haar and local binary pattern features of segmented images [5]. These techniques are time-consuming and ineffective in low resolution, noise, and blur. Further active contour algorithms were developed to detect mine-like objects [6]. The K-means method identifies mine-like objects. Segment restoration is enhanced and maybe sharpened using Chan-Vese active contour. Shadow images and highlight maps with geometric properties could indicate objects [6]. Kaeli et al. (2016) used saliency and rarity to identify SSS image abnormalities. SSS image anomalies may be detected by viewing higher resolutions and forwarding selected images for operator review [7]. Iterative diffusion maps recognize objects in large SSS data sets [8]. These methods sample and expand the function beyond the observed data, identify suspicious points, and find irregularities relative to usual data points. Optimal diffusion map anomaly detection for prolonged data requires well-chosen samples.

The Cubic smoothing spline and mine shadows method detect SSS landmarks using track signals. Fast object identification via graph-based segmentation [9]. After preprocessing using a graph-based approach called both-way spanning forest (BSF), exact segmentation is done using level sets and region-scalable fitting (RSF). The minimum/maximum spanning tree technique quickly analyzes sonar images to approximate segmentation in the first phase. The RSF model improves this result for final segmentation [10]. Einsidler et al. (2018) implemented transfer learning using a pre-trained CNN. This enabled pixel-intensity-based seabed anomaly detection using SSS images [11]. CNN models needed a large dataset to train and reduce false alarms. A pattern recognition tool by Kobenko et al. (2019) identified items in SSS images [12]. Since it lacks SSS images, this approach evaluates performance using a created dataset. Advanced and regular SSS help find drowning people, shipwrecks, and aircraft in underwater search and rescue activities. In protracted search missions, sonar operators may tire and miss objects. Therefore, automated object recognition and categorization systems are useful. Due to insufficient datasets, current methods for recognizing submerged things in SSS images focus on mines and ignore non-military objects. Huo et al. (2020) used semisynthetic data to create sonar images of drowning victims and aircraft. Segmenting optical images and simulating intensity distribution across areas are used in this method [13]. Thanh Le et al. (2020) also developed a Gabor filter-feature pyramid network detector. This approach uses lower training weights and merges weak and robust

traits to find mine-like things of various sizes. To improve computational performance, an optimized Gabor layer with customizable parameters is suggested [14]. Additionally, sophisticated neural networks have been developed for object identification, including Faster R-CNN [15], Darknet-53, YOLOv1, ResNet18 [16], and YOLOv2. The Faster R-CNN model takes a long time to identify tiny SSS image items. Darknet-53, YOLOv1, ResNet18[16], and YOLOv2 spotted microscopic objects [15]. Yulin et al. (2020) used Darknet-53 and YOLOv3's frozen convolutional layer. The research used multi-scale training and shallow characteristics to recognize smaller things. Shallow-feature Feature Pyramid Networks (FPN) enhanced shipwreck target identification. Manonmani et al. (2021) presented HOG and edge-based feature extraction [17]. Information is ready for grey-scale image processing and scaling. Features are extracted, and patch matching is used to assess whether the image includes a mine. Połap et al. (2022) suggested a CNN-based hybrid image-processing approach for deep sea item detection [18]. Li et al. (2023) introduced shadow-based target recognition in SSS images [19]. The Gaussian function second derivative is employed for weighted items with l2- and l1-norm in weighted sparse detection model building. This work's ultimate solution employs a lengthy shadow and target detection. According to extensive SSS image analysis, sonar image noise is mostly caused by internal and external causes. SSS images become noisier, reducing image quality and target recognition. Proper SSS image preprocessing improves target detection. The SSS image algorithm's YOLO V5-based target object recognition has also been improved. Multiple comparative studies show that the modified method improves SSS image target identification accuracy and speed. It solved problems, including failure detection and limited accuracy in detecting microscopic things [20]. In 2023, Tang et al. created a single-cycle consistency network with spatial attention, channel, and generative adversarial networks employing least square errors [21]. This was created to convert optical and SSS acoustic samples [22] efficiently. The detection model was used to identify underwater targets using training from the produced samples to test this approach's ability to provide high-quality samples. SSS image target object segmentation has been achieved using various balancing frameworks. A multi-headed balance module controls output in these systems. Zhang et al. (2024) developed a level-set-based heterogeneity filter-based image segmentation method [23]. This combo method segments sonar target images. The approach removes sonar speckle noise using a nonlocal means filter. A super-pixel method groups areas with similar textures to simplify calculation. Two heterogeneity filters improve target outlines and minimize sonar image variations. An adjustable threshold captures bright and dark areas' first outlines. Acoustic shadows, speckle noise, and geometric distortion reduce sonar image quality. SSS landmark detection is essential for underwater activities. Side-scan image object identification via patch feature extraction is suggested in this article. LBPs of non-overlapping SSS image patches are used to extract features. Features of non-overlapping patches and database objects are matched. Merging maximum-scored feature-matching similarity patches detects items.

The main contributions of this work are summarized as follows.

1. Patch feature matching using Local binary patterns (LBP) improves object detection performance.

2. Feature matching, similarity estimation, most similarity patch extraction, and merging effectively mine like objects and other objects under complex seafloors.

3. An object detection technique has been developed to reduce computational complexity using patch feature matching and maximum-scored patch merging.

4.. Experimental results on the SSS dataset show that the developed object detection can effectively detect objects under complex seafloors.

The rest of the paper is organized as follows: Sect. 2 discusses the proposed object detection technique and Sect. 3 illustrates the experimental results using an SSS dataset.

2. The proposed Object Detection Framework

Figure 1 shows the proposed two-stage feature extraction-based object detection framework in SSS images. Initially, the given SSS image is divided into non-overlapping patches. Patch feature vectors are extracted using an LBP. Further feature matching uses cosine similarity between the database images and overlapping patches. The maximum-scored patches using LBP feature matching are merged to locate the object. A bounding box in the SSS image bounds the object matched with the database.

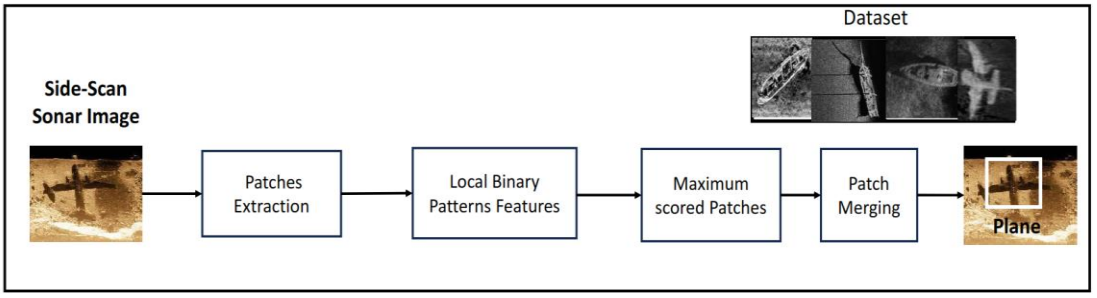

Figure.1 The proposed object detection framework

2.1 Patches Extraction

SSS object detection using non-overlapping patches feature matching is employed to determine the exact location and features of the identified object. This method uses patch-matching to find the target object in the input SSS image by comparing it to the SSS database object patch. Non-overlapping sections are extracted by splitting the image equally horizontally and vertically.

The non-overlapping patches set T of the SSS image with resolution M×N is given by

$$T = (t_1, t_2, t_3, \ldots . t_K) \quad (1)$$

where K=(M/16) ×(N/8) =number of overlapping patches, $t_1, t_2, t_3, \ldots . t_K$ are the nonoverlapping patches extracted, M=Number of rows, and N=Number of columns of SSS input image.

2.2. Local binary pattern features

Feature extraction is crucial to computer vision object recognition. Feature extraction in visual applications is vital and complex. Patch, segment, and patch similarity structures are assessed

during feature-matching. The recommended object detection system uses two-stage feature extraction and matching. LBP features perform well in texture classification, surface inspection, and segmentation [24,25]. The typical Local Binary Pattern (LBP) operator classifies image pixels by comparing the center pixel's value against a 3-by-3 proximity threshold. Consider the result a binary integer. To see an illustration of LBP calculation, refer to Figure 2. A texture descriptor may be obtained by calculating the 256-bin histogram of the labels derived from the image. Each bin in the histogram of Local Binary Pattern (LBP) codes may be interpreted as a micro-texton, which captures local features like curving edges, spots, flat regions, and other similar patterns.

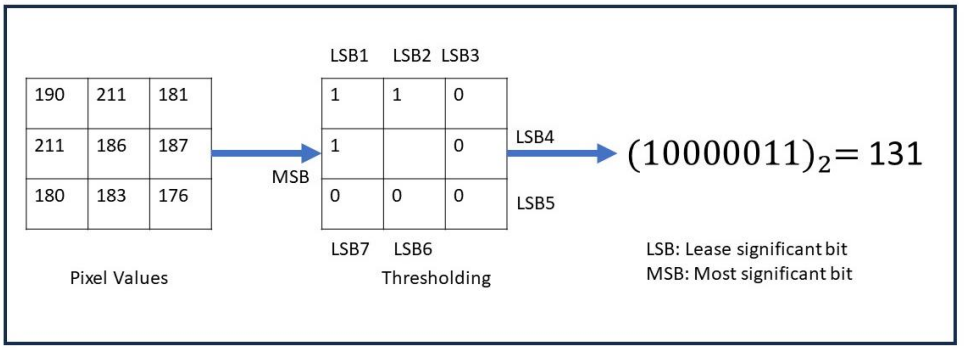

Figure 2 Example of LBP calculation.

The SSS images comprise micro-patterns that the LBP operator can effectively detect. They are partitioned into M distinct and non-overlapping areas denoted as R0, R1, ..., and RM to examine their specific characteristics, as seen in Figure 3. The LBP histograms obtained from each sub-region are merged into a unified feature histogram incorporating spatial information.

$$H_{i,j} = \sum_{x,y} I(f_l(x,y) = i) I\left((x,y) \in R_j\right) \quad (2)$$

where i = 0, ... , L-1, j = 0, ..., M-1. The extracted feature histogram describes SSS images' local texture and global shape.

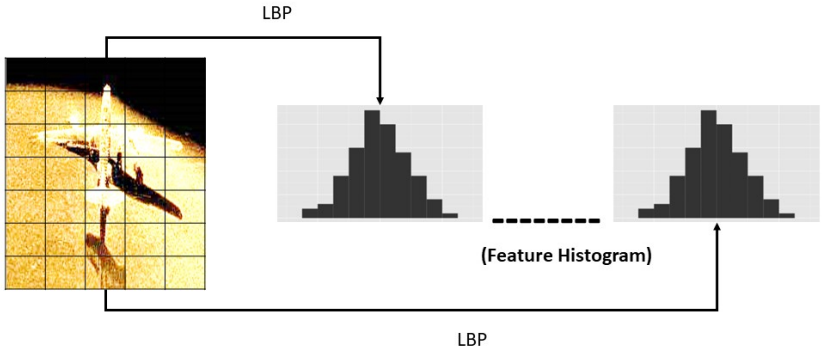

Figure 3 LBP-based feature representation.

The feature vector F_L of the patch using the LBP feature extraction is represented by

$F_L=\{F_{LBP}\}$ (3)

2.5 Object detection

The proposed object detection method involves matching patches with the earlier features. This matching process uses a cosine similarity score to compare the feature vector of the object patch in the database with the feature vectors of nonoverlapping patches in the side-scan sonar image. The feature extraction process involves using LBP features to obtain features for comparison. The similarity scores of patches are compared with features, and the patch maximum score value is considered. The patch scores higher than the threshold value are considered object patches. Finally, all patches are merged to locate the object. The LBP-based K^{th} non-overlapping patch feature vector set F_{LBP} (k) formed using Equations 1 and 3.

The similarity score uses the cosine similarity between the dataset feature vector F_{DLBP} using LBP and the K^{th} overlapped patch feature vector. F_{LBP} (k) is expressed by Equation 4:

$$SLBP(m,k) = similarityscore(FDLBP(m), F_{LBP}(k)) \quad (4)$$

where FDLBP(m) is the m^{th} image feature vector in a dataset using LBP. SLBP(m,k) represents the similarity score between a dataset's m^{th} image feature vector and the k^{th} overlapped patch feature vector.

$Smax=max(SLBP \geq Threshold)$ (5)

where T=threshold score.

Further, if maximum-scored patches exist, the object location is obtained using maximum-scored patches merging. A bounding box identifies the patches with maximum scores after merging and the object. The object detection procedure is explained in detail in algorithm 1.

Algorithm 1: Proposed Object Detection Algorithm

Input: SSS image I(n), SSS dataset, Threshold T=0.7.

Output: Object Location, Bounding box(I(n)

For i =1 to m

Extract non-overlapping patches

T=t(i)

Extract LBP features for non-overlapping patches

$FL = F_{LBP}(i)$

For j = I to n

Extract LBP features for data set patches.

$FDLBP = F_{DLBP}(j)$

Apply cosine similarity between nonoverlapping patch features and dataset using LBP.

$SLBP(i,j) = similarity(FDLBP(j), F_{LBP}(i))$

Extract maximum scored patches.

Smax(i, j) =maximum(SLBP(i, j)≥T)

If count(Smax(i, j)) ≠0

Merge maximum scored patches

Object bounded by bounding box

Else

Object not detected

End

End

End

3. Experimental Results

The following section comprehensively evaluates the proposed object detection technique using a specifically curated dataset. Given the absence of publicly available datasets for side scan sonar, our collected dataset comprises 1079 images from diverse sources-[26-29], and publicly shared websites. Each image in the dataset has been meticulously annotated with bounding boxes utilizing label boxes [30]. The dataset visualization is presented in Figure 4. Qualitative and quantitative analyses were carried out using MATLAB software on a meticulously configured system featuring an Intel Core i7 CPU, 32 GB of RAM, and an NVIDIA GeForce RTX 4060 GPU to assess the proposed object detection performance rigorously. The experimental validation of the proposed technique incorporated a two-stage feature extraction process, wherein patches with a cosine similarity score exceeding the threshold of 70 were extracted. This stringent approach ensures the extraction of the most similar patches, thereby substantiating the robustness of the proposed technique.

Figure 4. The collected dataset.

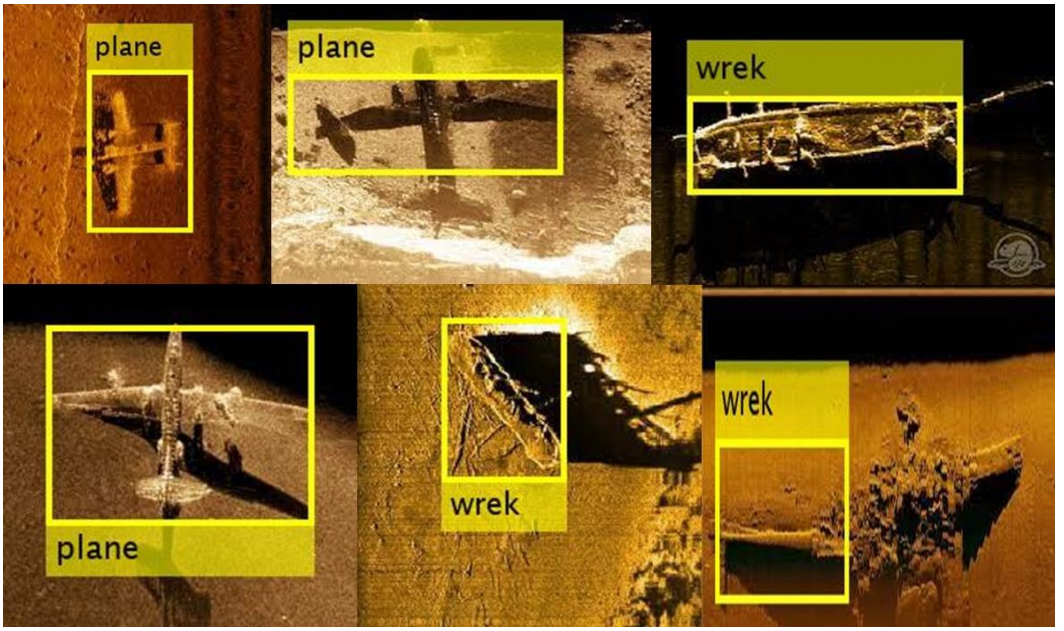

Figure 5 The object detection framework results on the collected dataset.

Figure 5 displays the qualitative results of the object detection method on the collected database. The suggested object detection approach in each SSS image accurately detects the target object within the scene with a ground truth overlap greater than 0.5. In addition, the object detection approach may be further improved by using multi-feature extraction techniques to enhance detection accuracy and reduce computing complexity.

The proposed object detection approach's quantitative evaluation and estimation were assessed using a confusion matrix. Qualitative analysis of the performance included performance parameters including accuracy, recall, average precision, and detection speed [29]. The performance of the object detection is shown in Table 1. The proposed -feature extraction and matching technique outperforms LBP features, achieving a precision score of 83.5%, an accuracy of 74.6%, and an F1 score of 0.85. The suggested approach's overall performance indicators align with the most recent standards for visual detection of objects.

Table 1 Performance of two-stage feature-based object detection framework.

Performance Measure	Aircraft	Ship	Human Body	others	Overall
Number of SSS images	432	356	203	88	1079
True Positive	326	247	135	46	754
False Positive	64	56	26	24	170
False Negative	31	43	16	14	104
True Negative	11	10	26	4	51
Precision	83.5	81.5	83.9	65.7	81.6
Accuracy	78	72.1	78.3	56.8	74.6
F-1 Score	0.873	0.83	0.865	0.71	0.85

The performance of the latest object detection techniques, including Mask R-CNN [30], MS R-CNN [31], and SOLO [32], was compared. The results were then compared with the proposed technique in Table 2 to evaluate its overall performance against state-of-the-art techniques.

Table 2 The object detection performance comparison with state-of-art techniques.

	Precision				
Method	Aircraft	Ship	Human Body	Other objects	Average Precision
Mask R-CNN [37]	70.29	78.32	67.54	64.57	70.2
MS R-CNN [38]	72.27	81.29	67.65	68.4	72.4
SOLO [39]	76.78	77.33	72.82	69.7	74.2
Proposed	83.6	83.3	86.5	70.7	81.6

The object detection proved more accurate than Mask R-CNN, MS R-CNN, and SOLO at identifying the target object in SSS images. Its average accuracy in spotting items in the dataset was assessed using a confusion matrix. The suggested method outperformed Mask R-CNN (70.29%), MS R-CNN (72.27%), and SOLO (76.78%) in aircraft object detection with 83.6%. Compared to state-of-the-art methods, the suggested technology detected shipwrecks and human corpses more accurately. The proposed method outperformed Mask R-CNN (70.2%), MS R-CNN (72.4%), and SOLO (74.2%) in average accuracy at 81.6%.

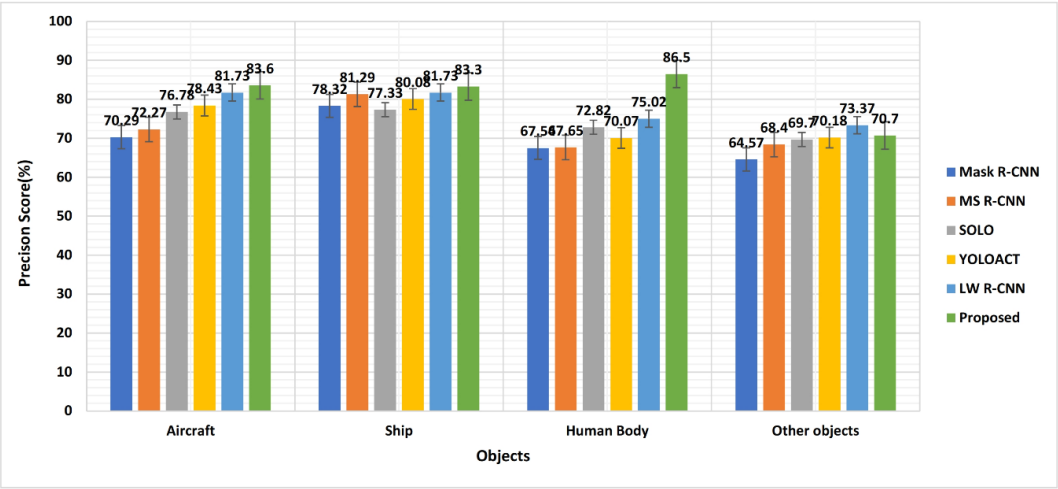

Figure 6 The overall performance of the proposed technique in different objects detection compared with the state-of-the-art techniques.

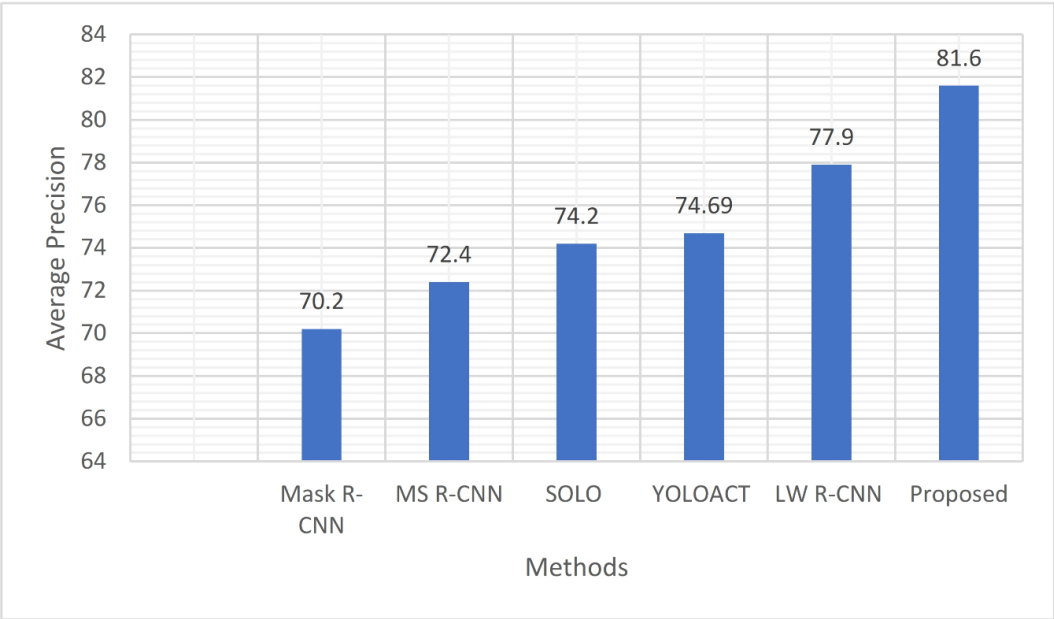

Figure 7. The overall performance of the proposed technique compared with the state-of-the-art techniques.

A dataset is used to quantify the number of frames processed per second to assess object detection speed. Table 3 shows that SOLO detects objects at 41.7 frames per second, whereas the suggested detector detects at 28.7. The proposed object detection system's speed of 29.4 frames per second, achieved through multi-stage feature extraction and matching,

demonstrates its robustness compared to other state-of-the-art techniques. Additionally, further performance and detection speed improvement could be achieved by implementing a faster feature extraction technique in the proposed object detection approach.

Table.3 The proposed object detection speed performance with state-of-the art-techniques,

Method	Mask R-CNN	MS R-CNN	SOLO	Proposed
Speed (Frames/second)	18.5	24.4	41.7	28.7

4. CONCLUSION

This paper introduces a Two-stage feature extraction and feature matching-based object detection framework with LBP Features. The process consists of four key steps: patch extraction, feature extraction, similarity estimation, and maximum-scored patch merging. Our proposed framework identifies objects by matching their patch features. In the feature extraction stage, we utilize LBP features for feature matching. Subsequently, the detected object patch is determined as the patches merging among the top-scored patches from feature-matching similarity scores. The quantitative and qualitative results demonstrate that our technique achieves a precision score of 81.6% while maintaining a detection speed of 28.7 frames per second, underscoring the effectiveness and robustness of our approach. Nevertheless, our experimental results also demonstrate that the illumination effect might influence the process of identifying an object. Hence, it is essential to address these problems in the next investigation. Moreover, our object detection system may be modified to identify numerous objects on the seafloor by using distinguishing features.

References

1. Yannik Steiniger, Dieter Kraus, Tobias Meisen, Survey on deep learning based computer vision for sonar imagery, Engineering Applications of Artificial Intelligence, Volume 114, 2022, 105157, ISSN 0952-1976, https://doi.org/10.1016/j.engappai.2022.105157.
2. C. Barngrover, A. Althoff, P. DeGuzman and R. Kastner, "A Brain-Computer Interface (BCI) for the Detection of Mine-Like Objects in Sidescan Sonar Imagery," in IEEE Journal of Oceanic Engineering, vol. 41, no. 1, pp. 123-138, Jan. 2016, doi: 10.1109/JOE.2015.2408471.
3. Y. Attaf, A. O. Boudraa and C. Ray, "Amplitude-based dominant component analysis for underwater mines extraction in side scans sonar," OCEANS 2016 - Shanghai, Shanghai, China, 2016, pp. 1-4, doi: 10.1109/OCEANSAP.2016.7485712.
4. R. Amit, G. Mishne and R. Talmon, "Improving resolution in supervised patch-based target detection," 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, China, 2016, pp. 1994-1998, doi: 10.1109/ICASSP.2016.7472026.
5. R. Chang et al., "Underwater object detection with efficient shadow-removal for side scan sonar images," OCEANS 2016 - Shanghai, Shanghai, China, 2016, pp. 1-5, doi: 10.1109/OCEANSAP.2016.7485696.
6. P. Zhang, L. Wang and S. Han, "Research on segmentation of sonar image based on features learning," 2016 IEEE International Conference on Mechatronics and Automation, Harbin, China, 2016, pp. 1897-1901, doi: 10.1109/ICMA.2016.7558855.
7. 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.

8. J. W. Kaeli, "Real-time anomaly detection in side-scan sonar imagery for adaptive AUV missions," 2016 IEEE/OES Autonomous Underwater Vehicles (AUV), Tokyo, Japan, 2016, pp. 85-89, doi: 10.1109/AUV.2016.7778653.
9. G. Mishne and I. Cohen, "Iterative diffusion-based anomaly detection," 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 2017, pp. 1682-1686, doi: 10.1109/ICASSP.2017.7952443.
10. M. Al-Rawi, A. Galdran, F. Elmgren, J. Rodriguez, J. Bastos and M. Pinto, "Landmark detection from sidescan sonar images," 2017 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), Aqaba, Jordan, 2017, pp. 1-6, doi: 10.1109/AEECT.2017.8257760.
11. Y. Bai, X. Xie, G. Li and Z. Wang, "A fast graph based method for object segmentation in sidescan sonar image," 2018 7th International Symposium on Next Generation Electronics (ISNE), Taipei, Taiwan, 2018, pp. 1-4, doi: 10.1109/ISNE.2018.8394652.
12. D. Einsidler, M. Dhanak and P. -P. Beaujean, "A Deep Learning Approach to Target Recognition in Side-Scan Sonar Imagery," OCEANS 2018 MTS/IEEE Charleston, Charleston, SC, USA, 2018, pp. 1-4, doi: 10.1109/OCEANS.2018.8604879
13. V. Kobenko, A. Koshekov, B. Koshekova, S. Latypov, N. Kalantayevskaya and M. Gavrilova, "Identification Pattern Recognition Technique for Small-size Underwater Objects," 2019 Dynamics of Systems, Mechanisms and Machines (Dynamics), Omsk, Russia, 2019, pp. 1-6, doi: 10.1109/Dynamics47113.2019.8944741.
14. G. Huo, Z. Wu and J. Li, "Underwater Object Classification in Sidescan Sonar Images Using Deep Transfer Learning and Semisynthetic Training Data," in IEEE Access, vol. 8, pp. 47407-47418, 2020, doi: 10.1109/ACCESS.2020.2978880.
15. H. Thanh Le, S. L. Phung, P. B. Chapple, A. Bouzerdoum, C. H. Ritz and L. C. Tran, "Deep Gabor Neural Network for Automatic Detection of Mine-Like Objects in Sonar Imagery," in IEEE Access, vol. 8, pp. 94126-94139, 2020, doi: 10.1109/ACCESS.2020.2995390.
16. T. Yulin, S. Jin, G. Bian and Y. Zhang, "Shipwreck Target Recognition in Side-Scan Sonar Images by Improved YOLOv3 Model Based on Transfer Learning," in IEEE Access, vol. 8, pp. 173450-173460, 2020, doi: 10.1109/ACCESS.2020.3024813.
17. Aklak, A.F., Vadamala, P.R. “Visual object tracking via adaptive deep feature matching and overlap maximization”, Pattern Anal Applic 26, 889–906 (2023). https://doi.org/10.1007/s10044-023-01157-9.
18. M. S, A. L, A. S. L, A. S. L and S. Rangaswamy, "Underwater Mine Detection Using Histogram of oriented gradients and Canny Edge Detector," 2021 International Carnahan Conference on Security Technology (ICCST), Hatfield, United Kingdom, 2021, pp. 1-6, doi: 10.1109/ICCST49569.2021.9717404.
19. D. Połap, N. Wawrzyniak and M. Włodarczyk-Sielicka, "Side-Scan Sonar Analysis Using ROI Analysis and Deep Neural Networks," in IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-8, 2022, Art no. 4206108, doi: 10.1109/TGRS.2022.3147367.
20. S. Li, J. Ma, Y. Wu, Z. Xiang, S. Bian and G. Zhai, "SSS Small Target Detection via Combining Weighted Sparse Model With Shadow Characteristics," in IEEE Transactions on Geoscience and Remote Sensing, vol. 61, pp. 1-11, 2023, Art no. 5911711, doi: 10.1109/TGRS.2023.3285436.
21. [21] Z. Yin, S. Zhang, R. Sun, Y. Ding and Y. Guo, "Sonar Image Target Detection Based on Deep Learning," 2023 International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE), Ballar, India, 2023, pp. 1-9, doi: 10.1109/ICDCECE57866.2023.10150970.
22. Y. Tang et al., "SSS Underwater Target Image Samples Augmentation Based on the Cross-Domain Mapping Relationship of Images of the Same Physical Object," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 6393-6410,

2023, doi: 10.1109/JSTARS.2023.3292327.
23. R. Liu and X. Luo, "Image Segmentation of underwater side-scan Sonar under Long-Tail Distribution," 2023 8th International Conference on Intelligent Computing and Signal Processing (ICSP), Xi'an, China, 2023, pp. 432-437, doi: 10.1109/ICSP58490.2023.10248702.
24. M. Zhang, W. Cai, Y. Wang and J. Zhu, "A Level Set Method with Heterogeneity Filter for Side-Scan Sonar Image Segmentation," in IEEE Sensors Journal, vol. 24, no. 1, pp. 584-595, 1 Jan.1, 2024, doi: 10.1109/JSEN.2023.3334765.
25. C. Liu, F. Chang and Z. Chen, "Rapid Multiclass Traffic Sign Detection in High-Resolution Images," in IEEE Transactions on Intelligent Transportation Systems, vol. 15, no. 6, pp. 2394-2403, Dec. 2014, doi: 10.1109/TITS.2014.2314711.
26. S. Basar, M. Ali, G. Ochoa-Ruiz, A. Waheed, G. Rodriguez-Hernandez and M. Zareei, "A Novel Defocused Image Segmentation Method Based on PCNN and LBP," in IEEE Access, vol. 9, pp. 87219-87240, 2021, doi: 10.1109/ACCESS.2021.3084905.
27. L. Jiang, T. Cai, Q. Ma, F. Xu, and S. Wang, "Active Object Detection in Sonar Images," IEEE Access., vol. 8, no. 12, pp. 102540–102553, May. 2020, doi: 10.1109/ACCESS.2020.2999341.
28. Z. Wang, S. Zhang, W. Huang, J. Guo, and L. Zeng, "Sonar Image Target Detection Based on Adaptive Global Feature Enhancement Network," IEEE Sensors Journal., vol. 22, no. 2, pp. 1509–1530, Jan. 2022, doi: 10.1109/JSEN.2021.3131645
29. G. Huo, Z. Wu, and J. Li, "Underwater Object Classification in Sidescan Sonar Images Using Deep Transfer Learning and Semisynthetic Training Data," IEEE Access., vol. 8, no. 12, pp. 47407–47418, Mar. 2020, doi: 10.1109/ACCESS.2020.2978880.
30. B. C. Russell, A. Torralba, K. P. Murphy, and W. T. Freeman, "LabelMe: A Database and Web-Based Tool for Image Annotation," International Journal of Computer Vision., vol. 77, no. 3, pp. 157–173, May. 2008, doi: 10.1007/s11263-007-0090-8.
31. K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," in Proc. IEEE Conf. Comp. Vis. Patt. Recogn (ICCV), Oct. 2017, pp. 2980–2988. doi: 10.1109/ICCV.2017.322.
32. Z. Huang, L. Huang, Y. Gong, C. Huang, and X. Wang, "Mask Scoring R-CNN," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2019, pp. 6402–6411. doi: 10.1109/CVPR.2019.00657.
33. X. Wang, T. Kong, C. Shen, Y. Jiang, and L. Li, "SOLO: Segmenting Objects by Locations," in Proc. Eur: Conf. Comput. Vis. (ECCV), Dec. 2020, pp. 649–665, doi: 10.1007/978-3-030-58523-5 38.