University of Tehran

Earth Observation and Geomatics Engineering

Online ISSN: 2588-4360

Homepage: https://eoge.ut.ac.ir/

Identification of Coastal and Non-Coastal Vessels in SAR Radar Images Using YOLOv8 Deep Learning Network

Farnaz Ashtari Mahini[1], Tayebe Managhebi[2]✉, Farhad Samadzadegan[3]

1. School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran. E-mail: f.ashtari@ut.ac.ir
2. Corresponding author, School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran. Email: t.managhebi@ut.ac.ir
3. School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran. E-mail: samadz@ut.ac.ir

Article Info

Article type:
Research Article

Article history:
Received 2024-09-01
Received in revised form 2024-09-25
Accepted 2024-09-30
Available online 2025-01-01

Keywords:
Ship detection,
SAR radar images,
SSDD dataset,
Deep learning network,
YOLOv8 network

ABSTRACT

Synthetic Aperture Radar (SAR) satellites, unlike optical imagery, can operate under any weather conditions and at any time, making them highly effective for ocean monitoring. Automatic ship detection in SAR images is critical for military and civilian applications. Traditional SAR-based methods rely on analyzing backscatter differences between ships and the sea surface, but challenges such as wave interference and the proximity of ships to the coastline often reduce accuracy. To address these limitations, deep learning-based approaches have emerged, offering advanced feature extraction and processing capabilities. However, many existing models are computationally intensive, limiting their application in real-time scenarios. To overcome these challenges, YOLOv8, a lightweight and efficient network, is proposed for ship detection in SAR images. This optimized architecture, incorporating multiple convolutional layers, enhances high-level semantic feature extraction, improving accuracy, speed, and robustness. The SSDD dataset, containing a variety of SAR images with different polarizations, resolutions, and coastal scenarios, was used to evaluate YOLOv8's performance. The results demonstrated exceptional accuracy of 99%, precision of 96%, mean Average Precision (mAP) of 98%, F1 score of 97%, and recall of 95%. YOLOv8 successfully detected small and large ships, even when located close to each other in crowded coastal backgrounds, showcasing its adaptability and reliability. These results highlight YOLOv8's potential for efficient and accurate ship detection in SAR images, addressing the challenges of traditional methods while enabling real-time monitoring for a wide range of applications. With its optimized design, YOLOv8 is a promising solution for improving ocean surveillance through SAR imagery.

Cite this article: Ashtari Mahini, F., Managhebi, T., & Samadzadegan, F. (2023). Identification of Coastal and Non-Coastal Vessels in SAR Radar Images Using YOLOv8 Deep Learning Network. *Earth Observation and Geomatics Engineering*, Volume 7, Issue 2, Pages 24-34. http//doi.org/10.22059/eoge.2024.381462.1160

Publisher: University of Tehran.
DOI: http//doi.org/10.22059/eoge.2024.381462.1160

1. Introduction

Synthetic Aperture Radar (SAR) is an active microwave imaging radar capable of observing the Earth under cloudy and foggy weather conditions, as well as throughout the day and night. Consequently, SAR images are widely used in target recognition, pattern detection, natural disaster monitoring, and environmental management. Among these applications, automatic ship detection in SAR images has garnered significant attention from researchers in recent years due to its numerous military and civilian applications, such as defense, fishing vessels monitoring, marine life surveillance, and maritime transportation oversight (Stasolla, Mallorqui, Margarit, Santamaria, & Walker, 2016). However, compared to optical images, SAR images obtained from satellite and airborne platforms generally have lower spatial resolution and require specialized knowledge for processing due to speckle noise. The nature of SAR images is significantly different from the optical images, as the amount and type of backscattering in SAR images are entirely influenced by the target's geometry. Therefore, in the category of target recognition, the impact of the recorded image's susceptibility to target geometry must be considered. In ship detection, especially for ships near the shoreline, some ships in SAR images exhibit scattering mechanisms similar to the surrounding areas, or many ships are densely scattered, easily leading to misidentification. Moreover, in SAR-based conventional ship detection methods, detecting small ships, particularly in cluttered backgrounds, often seems impossible. Additionally, identifying ships of various sizes poses another challenge, making accurate ship detection in SAR images with different spatial resolutions a significant challenge (Pang, Li, Zhang, Meng, & Zhang, 2022). Traditional SAR-based ship detection methods typically rely on texture features and polarimetric decomposition, which can be influenced by user-selected parameters. Furthermore, these traditional strategies are only suitable for detecting ships against simple backgrounds (Z. Wang, Wang, Ai, Zou, & Li, 2023; Xu & Liu, 2016). In other words, the conventional techniques usually adjust the threshold based on the contrast between the target (ship) and the background (water surface), which can perform well in high-contrast scenes. However, the presence of waves and the proximity of ships to the shoreline complicate the task, which leads to a decrease in the accuracy and reliability of the target detection algorithms. Therefore, when the surrounding environment is complex, using statistical data to describe the ship's backscattering mechanism becomes difficult, leading to a decrease in performance of target detection algorithms.

In recent years, deep learning methods have been widely used in target recognition and identification, target positioning, image segmentation, and more (Ball, Anderson, & Chan, 2017; Zhu et al., 2017). These methods, with their lightweight architecture and high speed, can learn various objects without the limitations of traditional methods (Tang, Zhuge, Claramunt, & Men, 2021). Deep learning networks use deep layers to extract features, and the most relevant features are selected by the architecture. Among deep learning networks, Convolutional Neural Networks (CNNs) are the most common type, operating based on multiple convolutional layers. In these networks, images are fed into the initial convolutional layer and combined with different kernels to identify patterns, ultimately producing feature maps. Region-based CNN (R-CNN) and Spatial Pyramid Pooling Networks (SPPNet) are other types of convolutional networks known for their two-stage algorithms. R-CNN networks provide high accuracy and flexibility for object detection in images but require uniform image dimensions for use (Girshick, Donahue, Darrell, & Malik, 2014). In contrast, SPPNet networks, by utilizing pyramid pooling layers, allow for processing images with different dimensions and create fixed-dimension feature maps (He, Zhang, Ren, & Sun, 2015). Another category of deep learning networks is known for their one-stage algorithms and can typically perform detection operations in a single step, such as SSD and YOLO. The YOLO deep learning network is an advanced algorithm that not only detects objects but also provides detailed information about the bounding boxes containing objects, such as center, length, and width (Redmon, Divvala, Girshick, & Farhadi, 2016). Recently, several versions of the YOLO network have been introduced, with the fifth version quickly gaining attention due to its flexibility and favorable architecture (Sun, Zhang, Wang, & Du, 2022). Consequently, after the release of this version, several iterations were published, with the latest being version eight (Skalski, 2022). The YOLOv8 deep learning network is an advanced model built on the success of previous YOLO versions. This version introduces new features and advancements to enhance performance and flexibility, enabling high-precision classification and segmentation. With its multi-scale architecture, lightweight design, and sequential convolutional layers, this network excels at extracting high-level semantic information more accurately and efficiently than previous models. The YOLOv8 network achieves impressive speed by using a one-stage detection approach, where bounding boxes and class probabilities are predicted in a single pass from the input image. This eliminates the need for Region Proposal Networks (RPNs) used in two-stage detectors like Faster R-CNN, resulting in faster inference times. Furthermore, by combining both high- and low-level semantic information, the network is well-equipped to tackle challenges such as detecting small objects in crowded or complex backgrounds.

This study aims to detect ships in SAR remote sensing images using the YOLOv8 deep learning network, which has an optimal architecture with high accuracy in object detection. In this framework, a variety of SAR images with low and high speckle noise are used, and the ability of the model to detect ships in non-coastal areas with quiet backgrounds and coastal areas with cluttered backgrounds is investigated.

2. Related Works

In 2019, Chang et al. utilized the second version of the YOLO deep learning network for ship detection in radar images. They employed the SAR Ship Detection Dataset (SSDD), which includes images from Radarsat-2, TerraSAR-X, and Sentinel-1 sensors, comprising a total of 1,160 radar images and 2,456 ships. Their study achieved an accuracy rate of 90%, demonstrating the superior performance of YOLO compared to the Faster-RCNN in ship detection based on radar images (Chang et al., 2019).

In 2021, Hong et al. explored ship detection using the third version of the YOLO network, including the tiny version and an improved variant. They used two separate datasets, SAR and optical, to evaluate and compare the performance of both datasets. The SAR dataset included 102 images from Gaofen-3 and 108 images from Sentinel-1, with spatial resolutions ranging from 3 to 20 meters, containing a total of 43,819 ships. The optical dataset comprised 150,000 ships extracted from SPOT satellite images with a spatial resolution of 15 meters. The study emphasized the importance of examining ships of various sizes in cluttered backgrounds and capturing images under different weather conditions, such as rainy and cloudy scenarios, to ensure data comprehensiveness. The goal was to develop a multi-scale algorithm for detecting ships in SAR and optical images with varying spatial resolutions. They employ an improved YOLOv3 model for their study. In the first step, an enhanced k-means++ algorithm was utilized to obtain the precise identification of the anchor boxes for ships. Afterward, a Gaussian model was introduced to predict the uncertainty of bounding boxes, and four anchors per scale were allocated in the Gaussian-YOLO detection layer to manage the significant variations in object size and orientation across different images. This strategy improved the accuracy of YOLOv3 and YOLOv3-tiny by two to three percent (Hong et al., 2021).

Ren et al. focused on ship detection in radar images using the SSDD dataset by employing the fifth and seventh versions of the YOLO network. They compared the results with networks like SSD and Retina-net and ultimately proposed a new network based on YOLOv5, named YOLO-Lite. The proposed network incorporated a backbone module to enhance the feature extraction which led to reduce the model training time. Additionally, an attention module was embedded in the backbone to accurately determine the target locations by capturing positional information. Afterward, an Enhanced Spatial Pyramid Pooling (EnSPP) module was also developed to boost feature capabilities while preventing the loss of small ship location information in high-level features. Finally, an effective Multi-scale Fusion Network (MFFNet) with dual feature combination channels was created to obtain feature maps with enriched positional and semantic information. The study concluded that YOLO-Lite, YOLOv7, and YOLOv5 achieved accuracy rates of 94%, 93%, and 92%, respectively, outperforming other networks in accurate SAR-based ship detection with diverse backgrounds while maintaining a lightweight architecture with low computational cost (Ren, Bai, Liu, & Zhang, 2023).

This paper aims to address the challenges of detecting ships of various sizes and backgrounds using the SSDD dataset to evaluate the performance of the eighth version of the YOLO network and improve the accuracy of ship detection compared to the previous versions of YOLO networks.

2. Proposed Method

The proposed method stages involve data collection and preparation, examining the selected network architecture, and finally training and testing it.

3.1. Data Preparation

Collecting challenging datasets is one of the most important parts of object detection projects and a prerequisite for studying the performance of deep learning algorithms. In this study, 1,160 radar images from coastal areas containing 2,456 ships of various sizes were analyzed. The dataset used is the SSDD radar dataset, which includes full and dual polarimetric images from Radarsat-2 and Sentinel-1 with a spatial resolution of 1 to 20 meters (Zhang et al., 2021). After collecting the data, all images were labeled using LabelImg module in python (Tzutalin/LabelImg, 2015) (Figure 1).

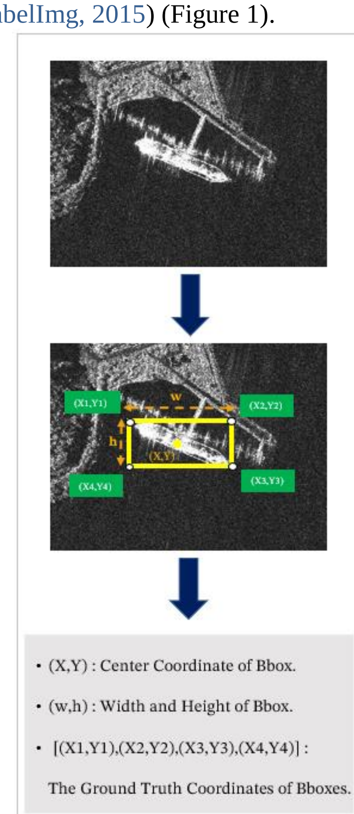

Figure 1. Data labeling process for network training.

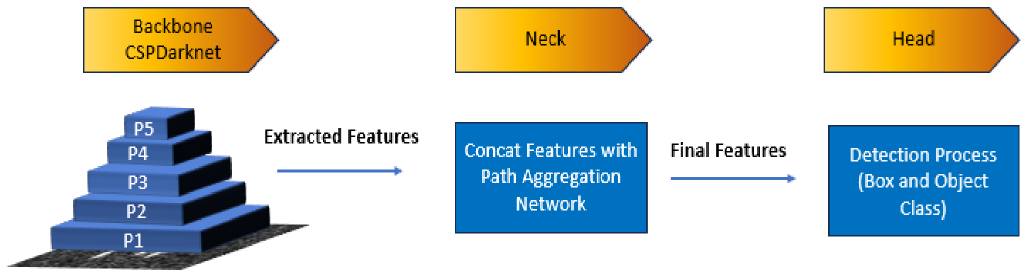

Figure 2. The Network Architecture.

As addressed in Figure 1, the labels contain the object's class, the coordinates of the bounding box's center, and the dimensions of the box. Subsequently, 80% of the samples were allocated for training the model, while the remaining 20% were reserved for testing.

3.2. Network Architecture and Training

At this stage, the deep learning architecture used is YOLOv8, with modifications made from version five. Afterward, the network was trained using 80% of the available data to identify and classify ships in the radar images. The integration of multi-scale architecture and lightweight design reduces complexity and the number of parameters while significantly increasing accuracy. As depicted in Figure 2, the implemented deep neural network structure consists of three parts: the backbone, neck, and head (C.-Y. Wang et al., 2020). In a CNN, the backbone is responsible for generating feature maps. The neck consists of layers aimed at receiving features from the backbone, combining them, and transferring them to the head, which predicts the object's class.

The proposed deep learning network is offered in five variations: nano, small, medium, large, and very large, which are different in depth and the number of layers (Rath, 2022). As the depth of layers in different versions of this network increases, both memory usage and final accuracy improve, while execution time decreases. In the neck of the proposed network, the Path Aggregation Network (PANet) is used, which reduces the number of parameters. The head of the network is used for the detection process and directly predicts the object center. This detection method speeds up the Non-Maximum Suppression (NMS) algorithm, resulting in fewer predicted boxes and ultimately reducing the model's overall complexity.

The goal is to select the most accurate bounding box that encloses the object, ensuring no overlapping with critical areas of interest. Finally, the deep learning algorithm use the NMS algorithm to select the final bounding boxes containing the object. At this stage, boxes with low confidence scores are removed, and the final box containing the vessel is selected (Hosang, Benenson, & Schiele, 2017). As shown in Figure 3, the yellow bounding box represents the optimal and most definitive choice.

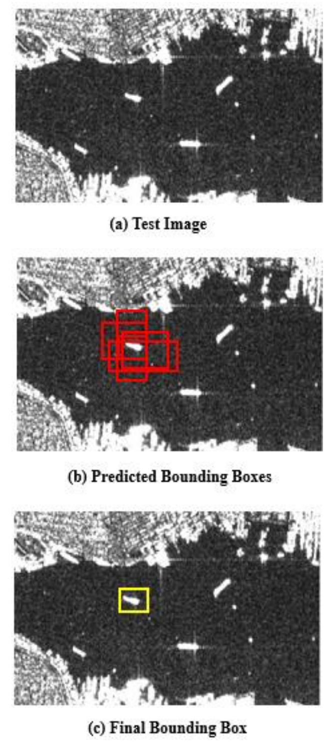

Figure 3. Non-Maximum Suppression (NMS) algorithm.

3.3. Network Testing Process

As previously mentioned, 20% of the data are employed in the evaluation phase uses, which includes vessels of various sizes. Clearly, the accuracy indices are used to quantify the network's performance in accurately detecting types of vessels. To calculate the accuracy indices, the values of the number (or percent) of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) of the model, which are included in the confusion matrix, are used (Simonyan & Zisserman, 2014). As is evident from Figure 4, the confusion matrix is $n \times n$ in size (where n is the number of classes) to represent model performance in detecting targets. The columns of this matrix indicate the actual classes of objects, which in this research include two classes: vessel and non-vessel. The matrix's rows indicate the classes predicted by the deep learning model. The positive class in Figure 4 relates to vessels, while the negative class refers to the background or non-vessels.

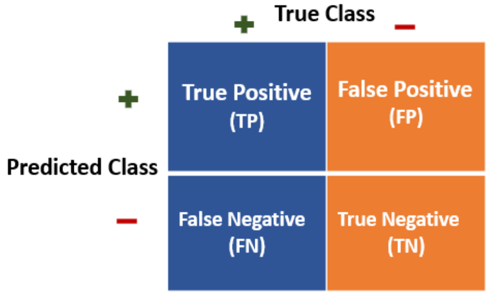

Figure 4. Confusion Matrix.

The precision index calculates the percentage of correct predictions for positive samples (TP) among all samples predicted as either true or false positive, as Equation 1.

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

The recall metric measures the level of missed detections, calculating the ratio of correct predictions for positive samples (TP) to the total positive samples correctly identified (TP) and those incorrectly predicted as negative (FN), as

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

In addition, the F1 score, which is the harmonic mean of precision and recall, is also used to evaluate the algorithm's performance as Equation 3.

$$F1 = \frac{2TP}{2TP + FP + FN} \quad (3)$$

Accuracy is another evaluation metric that measures the model's performance across all classes, showing the ratio of correct predictions (both positive and negative) to the total predictions, as

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

Mean Average Precision (mAP) is also used to compute the average precision across all classes of a model, with values ranging from 0 to 1.

$$mAP = \frac{1}{|n|}\sum_{i=1}^{n}\frac{TP(i)}{TP(i) + FP(i)} \quad (5)$$

4. Study Area

The SAR Ship Detection Dataset (SSDD) is the first open dataset widely used for advanced ship detection research from SAR images using deep learning techniques. The SSDD contains SAR images with various spatial resolutions ranging from 1 to 20 meters and different polarization bases. As shown in Figure 5, SSDD covers various maritime conditions with simple and complex backgrounds, including different types of ships, both inshore and offshore, of various sizes. In other words, many real-world challenges in ship detection are represented in this dataset (Zhang et al., 2021). It includes small ships, dense groups of ships docked in ports, large vessels, ships affected by significant speckle, and those in complex backgrounds. Thus, SSDD is a strong data source for studying these topics in SAR images with different resolutions and polarimetric features.

Table 1 represents the various characteristics of this dataset and details about the type of SAR imaging sensors, polarizations, environmental conditions, and the different types of ships. This highlights the diversity of data utilized and the identification problem at hand.

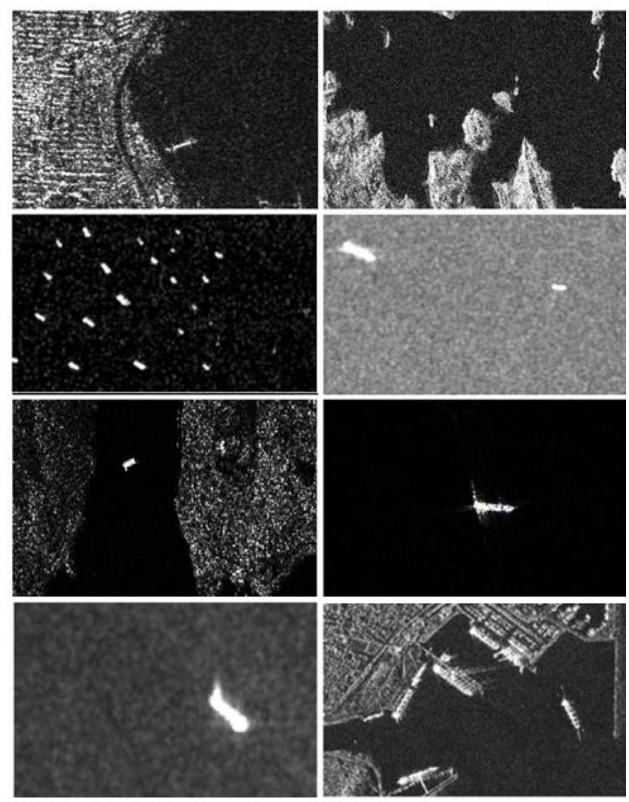

Figure 5 - Samples of SSDD dataset images (Zhang et al., 2021).

Table 1. Characteristics of the SSDD dataset (Zhang et al., 2021).

Sensors	RadarSat-2, TerraSAR-X, Sentinel-1
Polarization	HH, VV, VH, HV
Resolution	1 m-15 m
Places	Yantai,China; Visakhapatnam,India
Scale	1:1, 1:2, 2:1
Ship	Different sizes and materials
Sea Condition	Good and bad conditions
Scenes	Inshore and offshore
Image Number	1160
Ship Number	2456

5. Implementation and Result

The proposed deep learning network was trained using an Nvidia GeForce RTX 3050 Ti with 1000 training iterations. The input image size is set to 416 × 416, with one class and a batch size of 16. The proposed network relies on the PyTorch framework (version 1.9.0). For training and evaluation, CUDA (version 10.0) and deep neural network libraries (version 8.2) were used.

Figure 6 displays recall values over 1000 iterations on the x-axis and corresponding accuracy values on the y-axis. Clearly, as the curve moves towards the top right corner (values close to 1), the model's misclassification rate decreases. A larger area under the curve (AUC) signifies better model performance for that class.

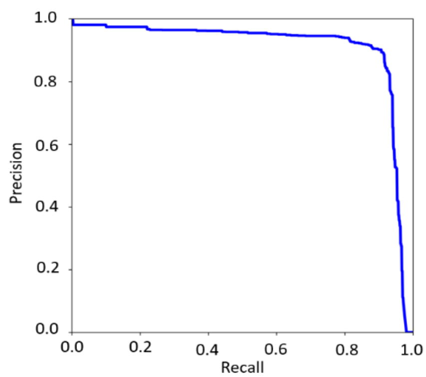

Figure 6. Chart of recall and precision values.

Finally, the model's final testing phase utilizes 20% of the remaining ground truth data and the confusion matrix will be generated as shown in Figure 7. Also, Figure 8 shows a summary of the final training results, in the form of changes in positive and negative evaluation metrics in the network training process. The upward trend in positive evaluation metrics (precision, recall and mAP) and the downward trend in negative metrics (box, objectness, val objectness and val box) indicate improved results over 1000 training iterations. This process reduced classification error and matched identified boxes to ground truth boxes.

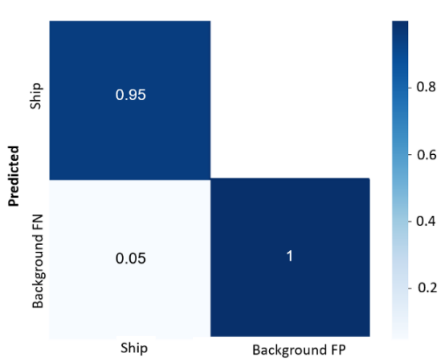

Figure 7. The final confusion matrix of the network.

Table 2 shows the final evaluation metric values. The experimental results show that 95% of the samples identified as ships were correctly classified. As shown in this Table, the evaluation results highlight the effectiveness and precision of the YOLO8 network in detecting ships in SAR images. It's important to note that the precision and mAP metrics derived from this network have improved by over 6% compared to the performance of the YOLO7 and YOLO-Lite networks on the SSDD.

Table 2. Evaluation criteria values.

Index	Precision	Recall	F1-Score	Accuracy	mAP
Value	0.99	0.95	0.97	0.96	0.98

5.1. Final Ship Detection Map

To better analyze the obtained results and show the ability of the presented network in ship identification, the final results in two separate categories of ship identification in coastal and non-coastal images are presented in Figures 9 and 10, respectively.

Figure 9 shows examples of detected ships in coastal scenes of different sizes. It is worth stressing that, the complex background is a significant challenge in object detection, so multiple types of backgrounds are illustrated in Figure 9. For example, Figures 9-b, 9-g, 9-i, and 9-n show images with different spatial resolutions and complex backgrounds. As expected, the trained network successfully detects ships despite the complex environments. In addition, as shown in Figures 9-a, 9-c, 9-h,9-m, and 9-o, small boats, which are likely to be vessels, have been successfully detected near the shore. This indicates the network's ability to detect ships in complex conditions, especially close to the dock with other ships nearby. Moreover, all images contain speckles, yet the results show that the model can still identify various types of ships of different sizes, even in complex backgrounds. This demonstrates the model's accuracy and performance in challenging situations.

Figure 10 presents examples of ship detection results in non-coastal images. These images do not show docks, and the target ships are farther from the shore. Additionally, wave effects and speckles are more pronounced compared to the coastal images. The model is capable of identifying the ships even with severe noise, as seen in Figures 10-a,10-f, 10-j, and 10-k, which confirms that the proposed model also detects various ship types despite the presence of speckles, accurately. The proposed model is also effective in detecting ships in simpler backgrounds compared to coastal images (Figures 10-a, 10-c, 10-e, and 10-f). Furthermore, in images with many ships moving together, the model successfully identifies different types of vessels with high accuracy, despite the speckle (Figures 10-b, 10-d, 10-e, and 10-n).

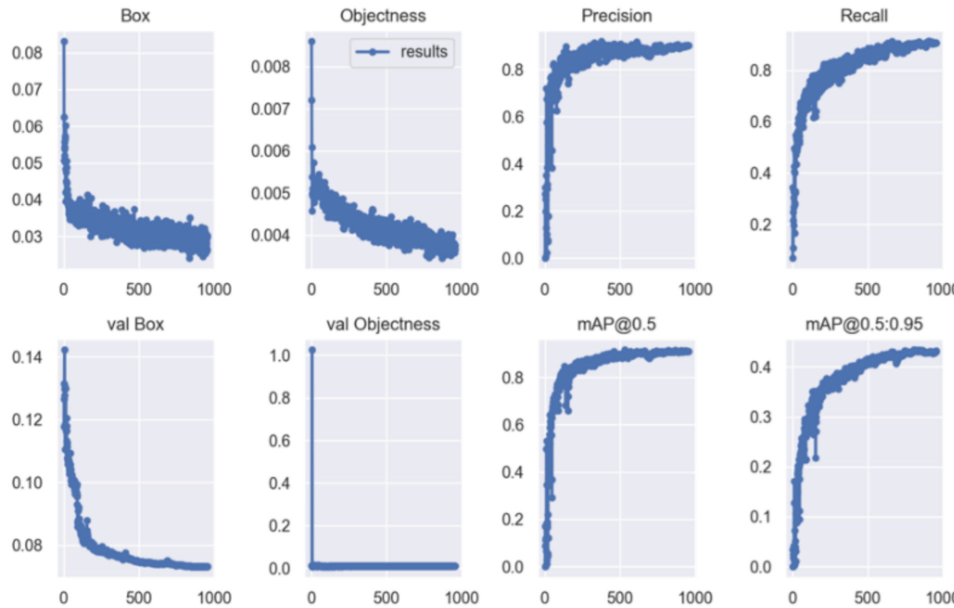

Figure 8. The Results of Network Training.

6. Discussion and Conclusion

Automatic ship detection in SAR images is crucial for both military and civilian applications. Ship target information from SAR images is often unclear, and complex backgrounds due to the sea and land interactions make detecting and monitoring ships more challenging. Compared to the traditional methods, deep learning approaches offer strong capabilities in data processing and feature extraction. However, many existing deep learning methods require complex models and heavy computations, making them less suitable for real-time ship detection.

Among the research, deep learning networks, especially the YOLO deep learning network, have made significant progress in the field of object recognition and have become a practical model in object recognition. Several methods have been used in the field of ship identification, including YOLOv3, YOLOv3-tiny, and other networks. In 2023, research was conducted by Ren et al. that could identify ships using different YOLO deep learning methods and provide a new method. They concluded that YOLO-Lite, YOLOv7, and YOLOv5 outperformed other networks with an accuracy of 94, 93, and 92% respectively in accurately detecting ships with different backgrounds while maintaining a lightweight architecture with low computational cost. However, to date, no research has been conducted on the use of the latest version of the YOLO network, YOLOv8, for ship identification in SAR images.

The YOLOv8 deep learning network, on the other hand, is characterized by its multiscale architecture, lightweight design, and sequential convolutional layers, extracting high-level semantic information more accurately and efficiently than previous models. The YOLOv8 network achieves impressive speed by using a one-step recognition approach in which bounding boxes and class probabilities are predicted from the input image in a single pass. This eliminates the need for Region Proposal Networks (RPNs) used in two-stage detectors such as the Faster R-CNN, resulting in faster inference times. By combining high- and low-level semantic information, the network is also well-equipped to handle challenges such as detecting small objects in crowded or complex backgrounds. Therefore, in addition to a higher accuracy than YOLOv7 and other networks (see Table 2), it was able to detect all types of small and large ships (close together and far apart) in images with crowded backgrounds. For this reason, it can be said that the YOLOv8 network has a more acceptable and effective performance and can achieve significant success in image recognition applications in SAR images.

In this study, the YOLOv8 deep learning network, featuring a lightweight and multi-scale architecture, was proposed for detecting and identifying ships in various SAR images, even with speckles. Using the SSDD dataset comprising 1160 radar images, the model's performance was evaluated using metrics such as mAP, F1-score, accuracy, precision, and recall, achieving values of 98%, 97%, 99%, 96%, and 95%, respectively. For instance, the mAP value of 98% indicates what percentage of all entries predicted as ships are actually ships. The F1-score of 97% indicates the harmonic mean of recall and precision values, which indicates adequate performance of the model. The accuracy rate of 99% indicates that the deep learning model correctly recognizes the percentage of the ship's input data class.

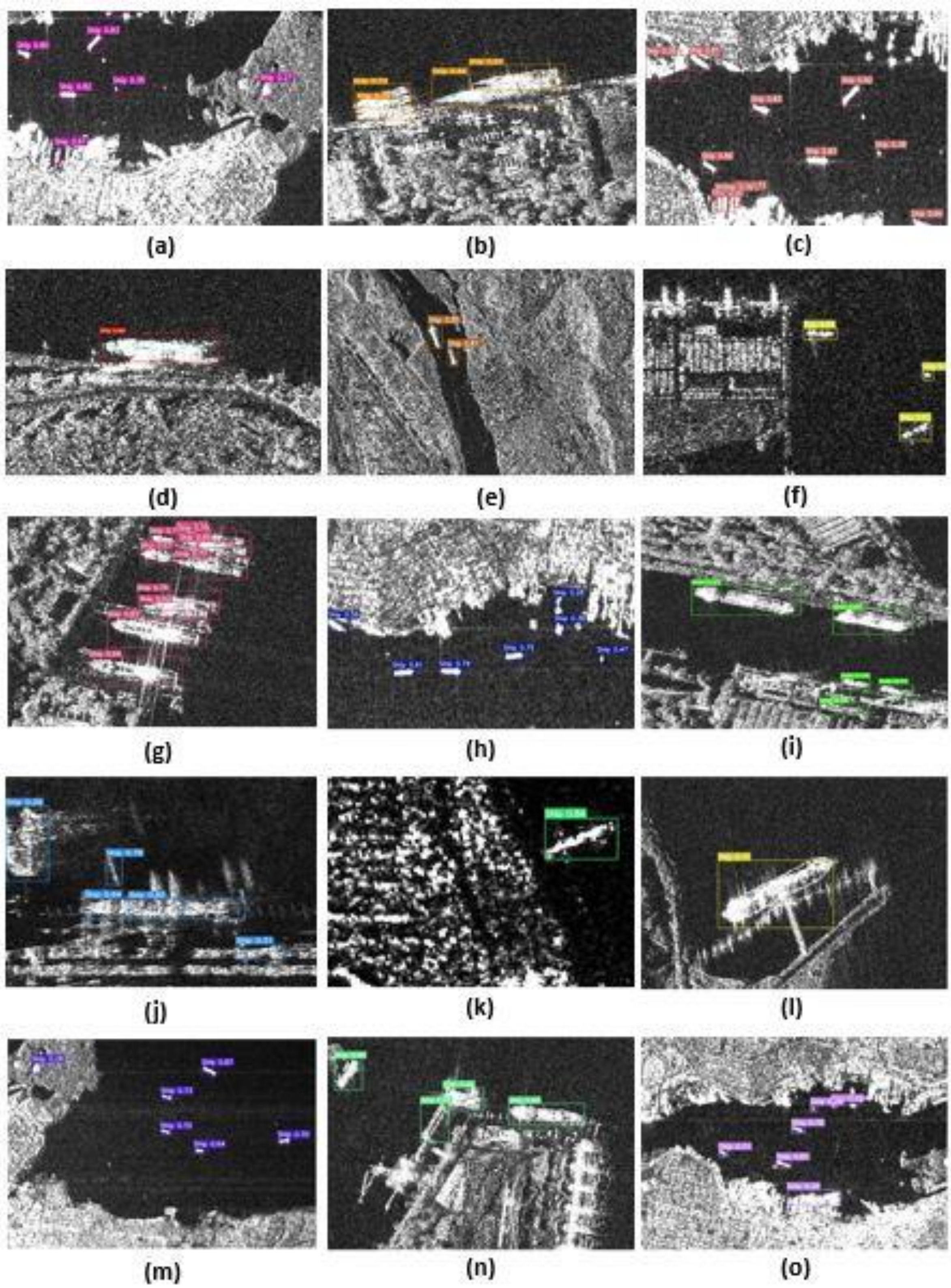

Figure 9. The results of network testing in coastal images.

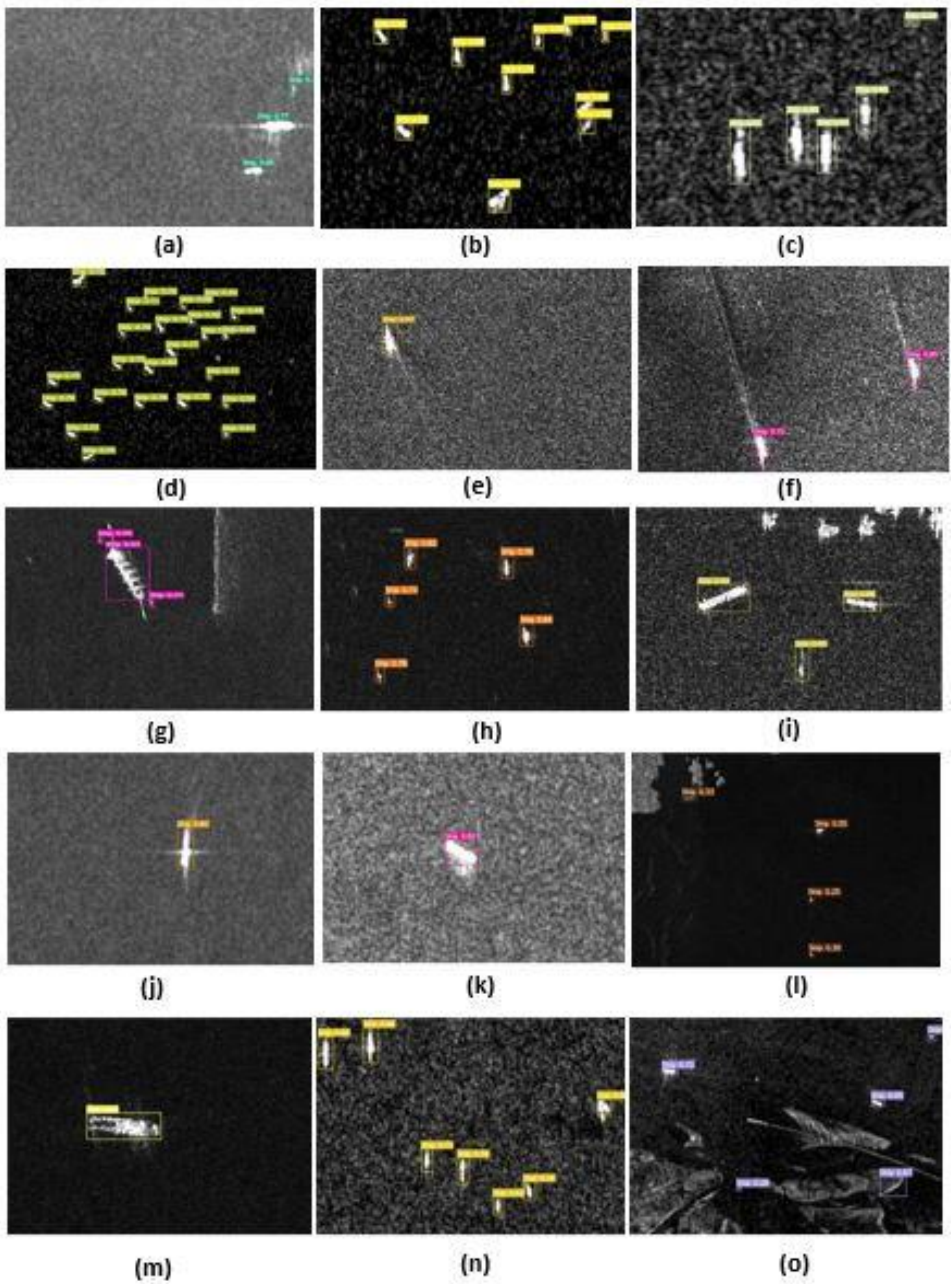

Figure 10. The results of network testing in non-coastal images.

Also, the precision rate of 96% indicates that 96% of all predicted inputs are ships. The recall rate of 95% means that of all the inputs that ship, what percentage of them were correctly detected and recognized as ship. These results indicate not only high accuracy in ship detection but also the model's adaptability in challenging scenarios, such as cluttered backgrounds.

Considering the patch-based labelLing process in the YOLO network and its effect on the accuracy of the network, methods based on semantic segmentation can be used in future works to directly detect the ship. For future work, several changes can be made to the network such as changing the activity function, the loss function, and the convolutional layers. Also, the investigation and application of deep learning models in Iranian ports can be an introduction to the development of an automatic ship detection system based on free radar and optical images.

References

Ball, J. E., Anderson, D. T., & Chan, C. S. (2017). Comprehensive survey of deep learning in remote sensing: theories, tools, and challenges for the community. *Journal of applied remote sensing, 11*(4), 042609-042609.https://doi.org/10.48550/arXiv.1709.00308

Chang, Y.-L., Anagaw, A., Chang, L., Wang, Y. C., Hsiao, C.-Y., & Lee, W.-H. (2019). Ship detection based on YOLOv2 for SAR imagery. *Remote Sensing, 11*(7), 786. https://doi.org/10.3390/rs11070786

Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). *Rich feature hierarchies for accurate object detection and semantic segmentation.* Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition. https://doi.org/10.1109/CVPR.2014.81

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE transactions on pattern analysis and machine intelligence, 37*(9), 1904-1916. https://doi.org/10.1109/TPAMI.2015.2389824

Hong, Z., Yang, T., Tong, X., Zhang, Y., Jiang, S., Zhou, R., . . . Liu, S. (2021). Multi-scale ship detection from SAR and optical imagery via a more accurate YOLOv3. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14*,6083-6101. http://dx.doi.org/10.1109/JSTARS.2021.3087555

Hosang, J., Benenson, R., & Schiele, B. (2017). *Learning non-maximum suppression.* Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition. https://doi.org/10.1109/CVPR.2017.685

Pang, L., Li, B., Zhang, F., Meng, X., & Zhang, L. (2022). A lightweight YOLOv5-MNE algorithm for SAR ship detection. *Sensors, 22*(18), 7088. https://doi.org/10.3390/s22187088

Rath, S. (2022). YOLOv8 Ultralytics: State-of-the-Art YOLO Models. Retrieved from https://learnopencv.com/ultralytics-yolov8/

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). *You only look once: Unified, real-time object detection.* Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition. https://doi.org/10.1109/CVPR.2016.91

Ren, X., Bai, Y., Liu, G., & Zhang, P. (2023). YOLO-Lite: An efficient lightweight network for SAR ship detection. *Remote Sensing, 15*(15), 3771. https://doi.org/10.3390/rs15153771

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*. https://doi.org/10.48550/arXiv.1409.1556

Skalski, P. (2022). How to Train YOLOv8 Object Detection on a Custom Dataset. Retrieved from https://blog.roboflow.com/how-to-train-yolov8-on-a-custom-dataset/

Stasolla, M., Mallorqui, J. J., Margarit, G., Santamaria, C., & Walker, N. (2016). A comparative study of operational vessel detectors for maritime surveillance using satellite-borne synthetic aperture radar. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9*(6),2687-2701. http://dx.doi.org/10.1109/JSTARS.2016.2551730

Sun, X., Zhang, Y., Wang, H., & Du, Y. (2022). *Research on ship detection of optical remote sensing image based on Yolo V5.* Paper presented at the Journal of Physics: Conference Series. http://dx.doi.org/10.1088/17426596/2215/1/012027

Tang, G., Zhuge, Y., Claramunt, C., & Men, S. (2021). N-Yolo: A SAR ship detection using noise-classifying and complete-target extraction. *Remote Sensing, 13*(5), 871. https://doi.org/10.3390/rs13050871

Tzutalin/LabelImg. (2015). Retrieved from Tzutalin/LabelImg. Available online: https://github.com/heartexlabs/labelImg

Wang, C.-Y., Liao, H.-Y. M., Wu, Y.-H., Chen, P.-Y., Hsieh, J.-W., & Yeh, I.-H. (2020). *CSPNet: A new backbone that can enhance learning capability of CNN.* Paper presented at the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops.

https://doi.org/10.1109/CVPRW50498.2020.00203

Wang, Z., Wang, R., Ai, J., Zou, H., & Li, J. (2023). Global and local context-aware ship detector for high-resolution SAR images. *IEEE Transactions on Aerospace and Electronic Systems*, *59*(4), 4159-4167. http://dx.doi.org/10.1109/TAES.2023.3237520

Xu, F., & Liu, J.-h. (2016). Ship detection and extraction using visual saliency and histogram of oriented gradient. *Optoelectronics letters*, *12*(6), 473-477. https://doi.org/10.1007/s11801-016-6179-y

Zhang, T., Zhang, X., Li, J., Xu, X., Wang, B., Zhan, X., . . . Su, H. (2021). SAR ship detection dataset (SSDD): Official release and comprehensive data analysis. *Remote Sensing*, *13*(18), 3690. https://doi.org/10.3390/rs13183690

Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE geoscience and remote sensing magazine*, *5*(4), 8-36. https://doi.org/10.1109/MGRS.2017.2762307