Vol: 3, Issue: 1, 2025
Pages:17-27
Received: 11 February 2025
Accepted: 15 April 2025

CURRENT TRENDS IN COMPUTING
Research Article
DOI: 10.71074/CTC.1636434

DETECTION OF SMALL AND MEDIUM SIZED SHIPS IN SATELLITE IMAGES USING YOLO MODELS

MOHAMED EMARA[1*], YASMEEN S.A.ABUSHAWAREB[2] AND EFTAL SEHIRLI[3]

[1,2] *Biomedical Engineering Department, Karabük University, 78050, Karabük, Türkiye*
[3] *Computer Engineering Department, Karabük University, 78050, Karabük, Türkiye*

ABSTRACT. **Ship detection in satellite images is an essential part of maritime security and surveillance. This work presents a helper tool for its growth. The three models presented in this paper use YOLOv8 and YOLOv5 with colored and grayscale satellite images to recognize small and medium-sized ships, which are frequently hard to spot in satellite photographs. Performance parameters such as mean average precision (mAP), recall, and precision were assessed. The models' mAP50 and recall ranged from 0.90 to 0.93 and 0.82 to 0.86, respectively. This work uncovered a comparable performance and acceptable detection accuracy for both grayscale and colored images. In satellite images, Model 2 and Model 3 demonstrated efficacy in identifying small and medium-sized ships.**

1. INTRODUCTION

Ships are required to use the Automatic Identification System to transmit identification signals. A timestamp, geographic coordinates, ship type, size, and the vessel's identifying are among the crucial details this technology offers [1]. Authorities cannot, however, track the ship or keep an eye on its actions, including fishing, smuggling, or any other illegal activity, if the transmitter is turned off.

Satellite images are an essential resource for enhancing marine situational awareness since they may identify ships that attempt to conceal. Constant False Alarm Rate (CFAR) detection techniques are the mainstay of conventional ship detection techniques in these kinds of images [2] . When applied to new satellite data, these conventional methods perform worse since they are sluggish in detecting medium- and small-sized ships and mostly rely on statistical patterns of sea clutter [3] .

Researchers working on satellite images have begun looking into methods for ship recognition utilizing deep learning-based target detection algorithms in computer vision as these algorithms have become more sophisticated. These techniques have made significant strides in fields like autonomous navigation and marine monitoring by demonstrating exceptional efficacy in identifying and categorizing objects

E-mail address: 2428142510@ogrenci.karabuk.edu.tr [(1)] abushawarebyasmeen@gmail.com, eftalsehirli@karabuk.edu.tr .
Key words and phrases. Ship detection, Satellite images, Image preprocessing, Maritime surveillance, YOLO.

from satellite images. This capability is essential for identifying illegal activity and guaranteeing compliance with legal obligations [4] . For several applications, such as satellite-based optical images and the analysis of various datasets, deep learning approaches have emerged as a crucial standard [5] .

In maritime surveillance, deep learning has greatly improved ship identification, increasing precision and effectiveness in a range of applications. In order to enhance detection skills, especially in difficult contexts, recent research has emphasized novel strategies that make use of deep learning architectures and multi-modal data integration [6] . YOLO-based strategies have advanced significantly among them. Notably, the YOLOv9-Adan model outperforms earlier YOLO versions with a mean Average Precision (mAP) of 65.5% for ship recognition utilizing drone images [7] . Similarly, by using attention techniques to improve the identification of small and distant boats, the custom architecture ScatYOLOv8+CBAM obtains a mAP of 75.46% for real-time ship recognition [8] . Aside from these developments, multi-modal data integration frameworks, such those that integrate SAR, optical, infrared, and AIS data, attain an impressive accuracy of 95.4%, successfully overcoming the drawbacks of conventional single-source approaches. With a lightweight model, detection skills are further improved by integrating complicated information from SAR images, such as phase and time-frequency characteristics, which yield an amazing 99.5% accuracy [9] . Notwithstanding these advances, there are still issues to be resolved, such as creating reliable models that can function well in a variety of environmental settings and manage imbalanced datasets. However, when used directly for small-target ship detection in datasets, these general target detectors encounter difficulties.

Apart from YOLO, deep learning has been a popular tool for ship detection [10] . CNN structure-based methods [11] such as baseline method [12] , RR-CNN [13] , RRD [14] , RBox-CNN [15] , R-DFPN [16] , R-FCN [17] , RoI Transformer [18] and LAFCR [19] were implemented for ship detection and classification on various datasets in the literature. It was uncovered that RBox-CNN achieved a 91.9% AP value as the best classifier method [11].

Firstly, the majority of research mostly use satellite remote sensing images of ships, however these images lack important characteristics that are present in satellite color images. These color images offer more information, such the color of the ships and other distinctive features, which are essential for improving ship recognition and helping authorities identify the ships more precisely.

Secondly, ships with small or ambiguous targets are frequently missed by conventional detection methods. Furthermore, occlusions or missed detections might result from using less-than-ideal angles or resolutions, therefore camera location is essential for producing high-quality images.

Finally, background noise from objects like lighthouses on the ocean's surface, islands, and fish farming nets along the coast can frequently impair remote sensing datasets, leading to false alarms in ship detection. Using particular traits, conventional algorithms try to differentiate small ships from other disturbances, but they frequently fall short of the necessary precision and effectiveness

This study proposes three models contributing to the challenges of small and medium-sized ship detection in sea environments by using a dataset constructed with 2231 high-quality satellite images, especially those with advanced single-stage detectors named YOLOv5 and YOLOv8 that increase the mAP and detection speed of tiny and hard-to-see ships.

The main contributions of this study can be summarized as follows:

- A high-quality dataset of captured small and medium-sized ships from satellite images.
- Solving the problem of unbalanced ship sizes by rotating the images 90 degrees, leading to a more uniform class distribution.
- A comparative analysis using three datasets: two in color and one in grayscale with adjusted brightness. While color datasets offer richer information, the grayscale version reduces noise, making ship detection easier.

The article is divided into main sections: the Materials and Methods section, which accounts for dataset preparation, processing techniques, and the implementation of YOLOv5 and YOLOv8 models; the Results section presents models on three datasets, of which two are colored and one grayscale; the Discussion gives an analysis of the importance of results and their practical applications. At the last section, the Conclusion summarizes the findings, benefits, and future objectives.

2. MATERIALS AND METHODS

2.1. **Material:**

A collection of ship-related satellite images was employed in this investigation. 85,000 satellite images of ships in all are included in the dataset; they were taken from the publicly available Kaggle collection with name ship data. The resolution of each photograph is 768x768 pixels, and they were all taken at different times from the same height. An analysis was conducted using carefully chosen 1115 images. Following preprocessing, there are 1043 images with one or more small ships and 1188 images with one or more medium-sized ships in the research dataset. Two classes labeled as ”There is a ship” and ”There is no ship” were identified from the dataset.

These images were modified by normalizing brightness and turning them into grayscale images. where some images remained grayscale, while others with brightness levels higher than 128 were reversed. This action lessened noise from clouds and reflected water surfaces. Additionally, it balanced the dataset and rotated at a 90-degree angle to increase the number of images, especially in the third model where the objective was to balance the distribution of small and medium-sized ships. They were stored in both color and grayscale forms so they could be used in two distinct models to confirm the outcomes.

2.2. **Method:**

Three separate sets were created from the dataset: the first set had colored images of medium-sized ships, the second set contained grayscale photographs of the same images totaling 1188, and the third set contained 2231 colored images of small and medium-sized ships.

To assess how well ship detection performed in various scenarios, three models were created. The original model only included medium-sized ships in colorful format and made use of the YOLOv5 architecture. A training set of 950 images, a validation set of 119 images, and a test set of 119 images comprised the dataset for this model. In order to maintain consistency and reduce noise, the second model used the same YOLOv5 architecture and the same images as the first model, but in grayscale. In particular, during preprocessing, all images were converted to grayscale, and brightness was then adjusted. While some images stayed grayscale, others with brightness levels higher than 128 were reversed. Cloud interference and reflecting water surface artifacts, which frequently manifest as false positives, were successfully minimized by this preprocessing technique. Additionally, the distribution of the dataset was the same as in the first model. With a training set of 1735 images, a validation set of 250 images, and a test set of 246 images, the third model, which was based on YOLOv8, included small and medium-sized ships in a colorful format.

Both CPU- and GPU-based systems were used to train the models in order to maximize training efficiency and guarantee computational efficiency. While the third model used a GPU-based system to increase performance on a bigger and more complicated dataset, the first two models were trained on a CPU-based system. Every model went through a rigorous training, validation, and testing process, during which hyperparameters were meticulously modified to get the best possible outcomes. Important variables were chosen to match the unique needs of each model, including the number of epochs, batch size, learning rate, and optimization technique.

YOLOv5s were used in the first and second models, which were trained on a CPU-based system with an emphasis on medium-sized ships. With a batch size of 16 and an image resolution of 768×768 pixels, these models were trained across 50 epochs. With adjustable learning rates and momentum settings, Data augmentation methods, such as rotation and brightness normalization, were used to take into consideration changes in image brightness and angle. An AMD Ryzen 7 4800H CPU system served as the training environment for these models, offering enough processing power to manage the medium-sized dataset.

The third model, on the other hand, was trained on a GPU-based system with an NVIDIA GeForce RTX 2060 and used YOLOv8s. In order to handle a more complicated dataset that included medium- and small-sized ships, this model was created. 50 epochs, a batch size of 16, and an initial learning rate of 0.001 were all part of the training setup. Higher efficiency and accuracy were ensured by the GPU system's ability to process the larger dataset and enable faster training. Table 1 presents the features and the models used.

All models' performances were assessed using common measures for object detection, such as Average Precision (AP), Precision, and Recall. The accuracy of bounding box predictions was assessed using the Intersection over Union (IoU) metric, which has a threshold of 0.5 IoU. Confusion matrices were also produced in order to examine the models' classification performance in more depth.

TABLE 1. **Features and the models used**

Feature	First Model (YOLOv5s)	Second Model (YOLOv5s)	Third Model (YOLOv8s)
System Used	AMD Ryzen 7 4800H	AMD Ryzen 7 4800H	NVIDIA GeForce RTX 2060
Dataset Size	1188	1188	2231
Dataset Composition	medium	medium	medium and small
Train/Val/Test	950/119/119	950/119/119	1735/250/246
Image Type Used	colored	grayscale	colored
Augmentation Techniques	rotation	rotation, brightness	rotation

YOLOv8 is a single-stage object detection framework developed for fast and efficient real-time detection. It does so by predicting the bounding boxes along with class probabilities in a single forward pass of the network. This architecture divides an input image into grid cells, then predicts for each cell a set of bounding boxes, confident scores of the bounding boxes, and class labels as shown in Figure 1.

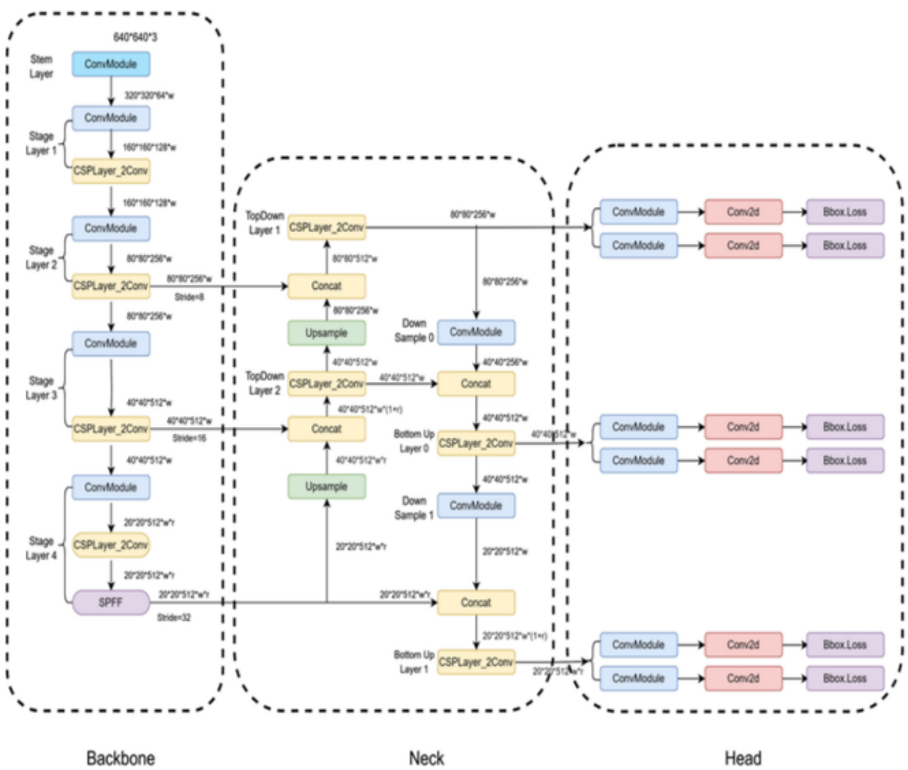


FIGURE 1. **YOLOv8 architecture.**

- Backbone: YOLOv8 is based on a Convolutional Neural Network that extracts hierarchical features from the input image. The backbone normally contains several convolutional and pooling layers to process the image data effectively.
- Neck: YOLOv8 refines, and fuses features from the different layers using a neck, such as PAN or FPN; it enhances the detection of objects at multiple scales, which is highly important to detect ships in a broad size range from the satellite images.
- Head: YOLOv8 directly predicts class probabilities, bounding box coordinates, and confidence scores for each cell in the detection head. This stage ultimately processes the feature maps to produce the detection output.

The YOLOv8 architecture functions as a single-stage detection model and was selected due to its effectiveness in real-time object identification applications. In a single forward pass, it predicts bounding boxes, confidence scores, and class probabilities after dividing the input image into a grid. The architecture consists of a detection head that generates the final predictions, a convolutional neural network (CNN) backbone for feature extraction, and a feature pyramid network (FPN) for multi-scale feature aggregation.

3. RESULTS AND DISCUSSION

3.1. **RESULTS:.**

Using datasets sorted by ship size and image color, three distinct YOLO-based models were trained and assessed for ship detection in this work. Precision (P), recall (R), mean average precision at IoU=0.5 (mAP50), and mean average precision at IoU=0.5:0.95 (mAP50-95) were used to evaluate each model's performance.

• First Model:Medium-Sized Ships

Using colored images, the first model was trained only on medium-sized ships. There were 950 images in the training dataset and 119 images in each of the validation and test datasets. Table 2 shows outcome obtained by the first model.

TABLE 2. **Outcome for the first model**

Dataset	**Precision**	**Recall**	**mAP@50**	**mAP@50:95**
Training Set	0.934	0.862	0.935	0.636
Test Set	0.910	0.855	0.925	0.578

• Second Model: Medium-Sized Ship

The second model was trained using the same medium-sized ship dataset, but it was transformed to grayscale in order to examine the impact of color information on ship detection. The distribution of the dataset (950 for training images, 119 for validation images, and 119 for test images) was unchanged. The findings are displayed in Table 3.

TABLE 3. **Outcome for the second model**

Dataset	**Precision**	**Recall**	**mAP@50**	**mAP@50:95**
Training Set	0.919	0.828	0.921	0.609
Test Set	0.928	0.838	0.924	0.588

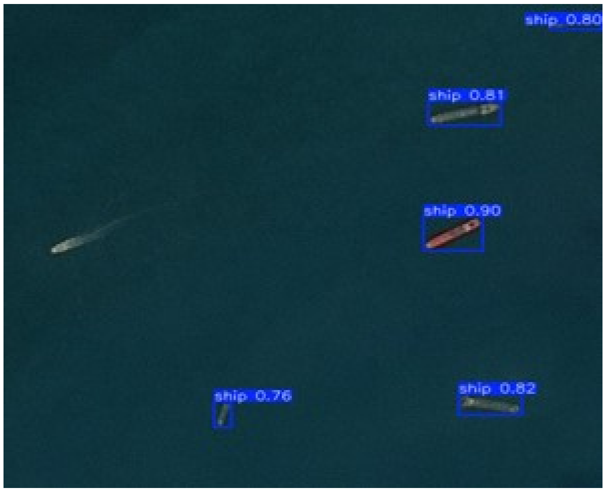
• Third Model: Mixed-Sized Ships

A dataset of colored images of small and medium-sized ships was used to train the third model. There were 246 test images, 250 validation images, and 1735 training images in the dataset. Table 4 displays the findings.

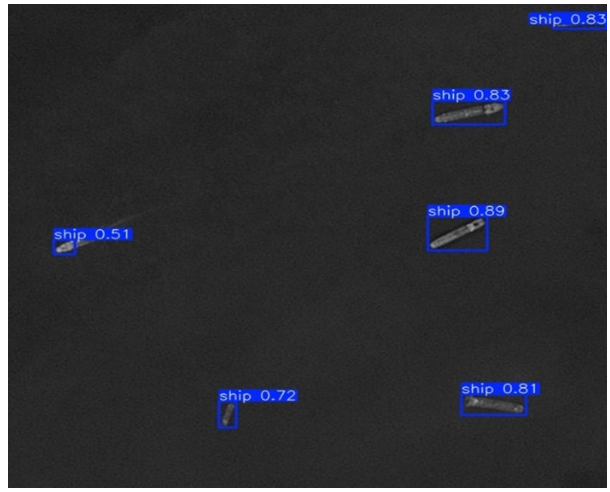
TABLE 4. **Outcome for the third model**

Dataset	**Precision**	**Recall**	**mAP@50**	**mAP@50:95**
Training Set	0.882	0.860	0.918	0.588
Test Set	0.866	0.855	0.908	0.528

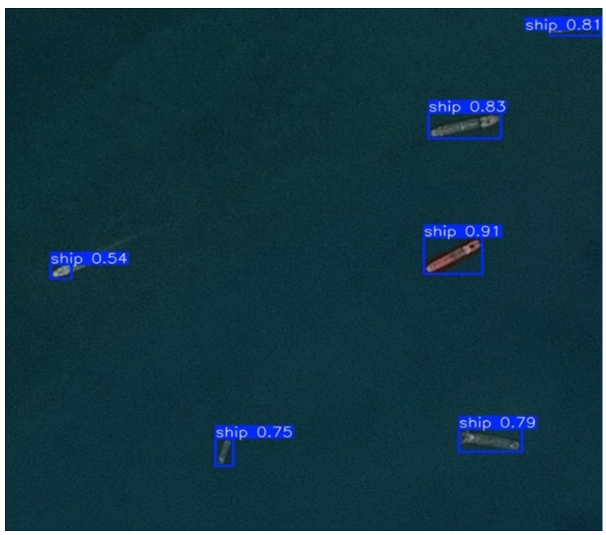
According to the outcome of the analysis, it was obtained that while **mAP@50** of the models took values between **0.90 to 0.93**, **recall** presented values between **0.82 to 0.86**. This work found sufficient and reliable performance and acceptable detection success for both grayscale and colored images. Figure 2 shows sample detected ships by the three models on a test image.



(A)



(B)



(C)

FIGURE 2. **A sample test image for: (A) First Model (Colored), (B) Second Model (Gray), (C) Third Model (Colored).**

3.2. **Discussion:**

• Color's Effect on Detection Accuracy

The comparison between the first and second models indicates that the presence of color information slightly improves detection performance. Nevertheless, competitive results were still achieved using grayscale images, suggesting that YOLO-based models are capable of learning structural properties, such as shape and size, even without color. This demonstrates the model's robustness in detecting ships based on spatial features alone.

• Impact of Variability in Ship Size

The third model, trained on both small and medium-sized ships, exhibited a moderate decrease in precision and recall, particularly in the mAP@50–95 metric. This decline is likely due to the increased variability in ship sizes, which makes accurate detection more challenging. Future improvements may involve focusing on specific ship size ranges or integrating strategies to better handle scale variations.

• Localization Performance

The slight decrease in mAP@50–95 also suggests potential for improving the model's localization capabilities. More accurate bounding box predictions could be achieved by expanding the training dataset, optimizing anchor box sizes, or utilizing more advanced YOLO versions with enhanced localization features.

• Choosing a Model for Implementation

The first model, trained on color images of medium-sized ships, achieved the highest precision and recall values, making it the most suitable for high-resolution detection tasks. However, the grayscale model remains a viable alternative with minimal performance loss, especially in scenarios where computational efficiency is a priority.

Interestingly, while the grayscale model underperforms slightly in precision and recall, it performs better in detecting small-sized vessels compared to the color model. Although the third model achieved slightly lower performance overall, it is considered the most generalizable due to its exposure to a more diverse dataset. Its superior ability to detect ships of varying sizes—particularly smaller vessels—makes it ideal for practical ship detection applications in real-world maritime environments.

• Comparison with Existing Literature

According to the literature, RBox-CNN is among the most effective classification methods for ship detection, achieving an AP of 91.9% [**?**]. In comparison, the proposed third model achieved a test set mAP of 90.8%. While it does not surpass the state-of-the-art, the proposed method offers an acceptable level of success and serves as a competitive alternative.

4. Conclusion

The YOLOv5 and YOLOv8 deep learning models were applied to optical satellite images in this study in order to propose an effective ship detection method used on satellite images. Three experiments were carried out in order to (1) compare the accuracy of models trained on colored and grayscale images of medium-sized ships, and (2) analyze the performance of models trained on small and medium-sized ships.

By improving model training and combining both color and grayscale information, satellite images' limitations in spotting small and medium-sized ships may be mitigated. With notable gains in detection accuracy of $0.87 \leq$ mAP@0.5, especially for medium-sized ships, the study's findings therefore show the efficacy of YOLO-based models for ship detection in satellite images. Improving the algorithms' capacity to identify ships in intricate environmental settings and boosting detection efficiency for smaller vessels are two upcoming challenges.

DECLARATIONS

- **Conflict of Interest: The authors declare that they have no known competing financial interests or no conflict of interest that could have appeared to influence the work reported in this paper.**

REFERENCES

[1] P. Heiselberg, H. B. Pedersen, K. A. Sørensen, and H. Heiselberg, "Identification of Ships in Satellite Images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 17, pp. 6045–6055, 2024.

[2] B. Lebona, W. Kleynhans, T. Celik, and L. Mdakane, "Ship Detection Using VIIRS Sensor Specific Data," in *Proc. 2016 IEEE Int. Geoscience and Remote Sensing Symp. (IGARSS)*, Beijing, China, July 10–15, 2016, pp. 7729315. DOI: 10.1109/IGARSS.2016.7729315.

[3] G. Cheng, X. Xie, J. Han, L. Guo, and G.-S. Xia, "Remote Sensing Image Scene Classification Meets Deep Learning: Challenges, Methods, Benchmarks, and Opportunities," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 3735–3756, 2020.

[4] Z. Lin, K. Ji, X. Leng, and G. Kuang, "Squeeze and Excitation Rank Faster R-CNN for Ship Detection in SAR Images," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 5, pp. 751–755, May 2019.

[5] T. Zhang, X. Zhang, J. Li, X. Xu, B. Wang, X. Zhan, Y. Xu, X. Ke, T. Zeng, H. Su, I. Ahmad, D. Pan, C. Liu, Y. Zhou, and J. Shi, "SAR Ship Detection Dataset (SSDD): Official Release and Comprehensive Data Analysis," *Remote Sensing*, vol. 13, no. 18, article 3690, 2021. DOI: 10.3390/rs13183690.

[6] D. Barretta, L. M. Millefiori, and P. Braca, "Analytical Classification Performance Analysis of Machine-Learning-Based Ship Detection from Optical Satellite Imagery," in *Proc. 2024 IEEE Int. Geoscience and Remote Sensing Symp. (IGARSS)*, Athens, Greece, 2024, DOI: 10.1109/IGARSS53475.2024.10642217.

[7] J. S. Rubikumar, "Efficient ship detection system for maritime surveillance using deep learning approach," *International Journal for Multidisciplinary Research (IJFMR)*, vol. 6, no. 5, pp. 1–11, Sept.–Oct. 2024. E-ISSN: 2582-2160. Available: `https://www.ijfmr.com`.

[8] B. C. Perez, "Real-time Ship Recognition and Georeferencing for the Improvement of Maritime Situational Awareness," Doctoral Thesis, Faculty 3 – Mathematics and Computer Science, University of Bremen, 2024. Available: arXiv:2410.04946 [cs.CV].

[9] C. Tian, Z. Lv, F. Xue, X. Wu, and D. Liu, "Multi-Domain Joint Synthetic Aperture Radar Ship Detection Method Integrating Complex Information with Deep Learning," *Remote Sensing*, vol. 16, no. 19, article 3555, 2024. DOI: 10.3390/rs16193555.

[10] M. J. Er, Y. Zhang, J. Chen, and W. Gao, "Ship detection with deep learning: a survey," *Artificial Intelligence Review*, vol. 56, no. 10, pp. 11825–11865, Mar. 2023. doi: 10.1007/s10462-023-10455-x.

[11] B. Li, X. Xie, X. Wei, and W. Tang, "Ship detection and classification from optical remote sensing images: A survey," *Chinese Journal of Aeronautics*, vol. 33, no. 10, pp. 1000-9361, 2020. DOI: 10.1016/j.cja.2020.09.022.

[12] Z. Liu, L. Yuan, L. Weng, and Y. Yang, "A High Resolution Optical Satellite Image Dataset for Ship Recognition and Some New Baselines," in *Proc. 6th Int. Conf. Pattern Recognition Applications and Methods (ICPRAM 2017)*, 2017, pp. 324-331. DOI: 10.5220/0006120603240331.

[13] Z. Liu, J. Hu, L. Weng, and Y. Yang, “Rotated Region Based CNN for Ship Detection,” in *Proc. 2017 IEEE Int. Conf. Image Processing (ICIP)*, Beijing, China, 2017, pp. 900-904. DOI: 10.1109/ICIP.2017.8296411.

[14] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in *Proc. 2014 IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Columbus, OH, USA, 2014, pp. 580–587. DOI: 10.1109/CVPR.2014.81.

[15] J. Koo, J. Seo, S. Jeon, J. Choe, and T. Jeon, “RBox-CNN: Rotated bounding box based CNN for ship detection in remote sensing image,” in *Proc. 26th ACM SIGSPATIAL Int. Conf. Adv. Geographic Inf. Syst. (SIGSPATIAL ’18)*, Seattle, WA, USA, Nov. 2018, pp. 420–423. doi: 10.1145/3274895.3274915.

[16] X. Yang, H. Sun, K. Fu, J. Yang, X. Sun, M. Yan, and Z. Guo, “Automatic ship detection in remote sensing images from Google Earth of complex scenes based on multiscale rotation dense feature pyramid networks,” *Remote Sens.*, vol. 10, no. 1, p. 132, 2018. DOI: 10.3390/rs10010132.

[17] M. Li, W. Guo, Z. Zhang, W. Yu, and T. Zhang, “Rotated region based fully convolutional network for ship detection,” in *Proc. IGARSS*, 2018, pp. 673. IEEE. DOI: 10.1109/IGARSS.2018.8517926.

[18] M. Liao, Z. Zhu, B. Shi, G. Xia, and X. Bai, “Rotation-sensitive regression for oriented scene text detection,” in *Proc. CVPR*, 2018, pp. 5909–5918. IEEE. DOI: 10.1109/CVPR.2018.00619.

[19] L. Liu, Y. Bai, and Y. Li, “Locality-aware rotated ship detection in high-resolution remote sensing imagery based on multiscale convolutional network,” *IEEE Geoscience and Remote Sensing Letters*, vol. 19, p. 3502805, 2022. DOI: 10.1109/LGRS.2021.3115732.