algorithms

MDPI

Article

Improved Ship Detection Algorithm from Satellite Images Using YOLOv7 and Graph Neural Network

Krishna Patel [1], Chintan Bhatt [2,*] and Pier Luigi Mazzeo [3,*]

[1] Department of Computer Science & Engineering, Devang Patel Institute of Advance Technology and Research (DEPSTAR), CHARUSAT Campus, Charotar University of Science and Technology (CHARUSAT), Changa, Anand 388421, Gujarat, India
[2] Department of Computer Science & Engineering, School of Technology, Pandit Deendayal Energy University, Gandhinagar 382007, Gujarat, India
[3] Institute of Applied Sciences and Intelligent Systems, National Research Council of Italy, 73100 Lecce, Italy
* Correspondence: chintan.bhatt@sot.pdpu.ac.in (C.B.); pierluigi.mazzeo@cnr.it (P.L.M.)

Abstract: One of the most critical issues that the marine surveillance system has to address is the accuracy of its ship detection. Since it is responsible for identifying potential pirate threats, it has to be able to perform its duties efficiently. In this paper, we present a novel deep learning approach that combines the capabilities of a Graph Neural Network (GNN) and a You Only Look Once (YOLOv7) deep learning framework. The main idea of this method is to provide a better understanding of the ship's presence in harbor areas. The three hyperparameters that are used in the development of this system are the learning rate, batch sizes, and optimization selection. The results of the experiments show that the Adam optimization achieves a 93.4% success rate when compared to the previous generation of the YOLOv7 algorithm. The High-Resolution Satellite Image Dataset (HRSID), which is a high-resolution image of a synthetic aperture radar, was used for the test. This method can be further improved by taking into account the various kinds of neural network architecture that are commonly used in deep learning.

Keywords: deep learning; GNN; object detection; HRSID dataset; high-resolution satellite images

check for updates

Citation: Patel, K.; Bhatt, C.; Mazzeo, P.L. Improved Ship Detection Algorithm from Satellite Images Using YOLOv7 and Graph Neural Network. *Algorithms* **2022**, *15*, 473. https://doi.org/10.3390/a15120473

Academic Editors: Lucia Maddalena and Laura Antonelli

Received: 3 October 2022
Accepted: 7 December 2022
Published: 12 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The management of marine security relies significantly on remote sensing images for the automatic ship detection. Its primary duties include keeping an eye on traffic, finding illicit fishing, and stopping maritime pollution. Military organizations use the automatic ship detection system to enhance maritime security. This process can be carried out through various activities, such as reconnaissance, surveillance, and intelligence. One of the most common technologies used in this field is advanced remote sensing. This type of technology is used to gather various types of data. It can gather various data points such as radar, electro-optical cameras, and electronic support systems. This research focuses on analyzing satellite photos. Deep learning is a process that requires a lot of training data to develop.

All commercial and passenger ships weighing over 300 tons are required to have an automatic identification system (AIS) transponder. This type of device transmits information about the vessel's location and destination. However, it can be easily manipulated. For instance, if a fishing boat wants to pretend to be another vessel, it can alter the type of information that the ship transmits. Convolutional neural networks (CNNs), a tiny subset of machine learning (ML), are among the more recent technologies that have seen more successful implementations [1]. Additionally, they have been integrated with a multi-layered network architecture created using conventional neural network techniques. CNNs consist of various components, such as activation function, input layers, output layers, and convolutional layers.

Ship identification and classification have also been accomplished using a deep learning strategy [2], a process of deep learning influenced by the human brain's structure and function. It can be used to process the data collected by the SAR system, which include monitoring plants and diseases, mapping various trajectories, and analyzing the data collected from various sources. Therefore, the primary goal of this study is to identify the presence of ships using satellite photos with high accuracy.

In this study, we suggest going a step further in addressing the issue in the automatic identification of ships. When compared to handcrafted features, our method, which is based on the well-known CNN architecture You Only Look Once (YOLO), can determine the most distinctive features for the given task [3]. In the suggested framework, picture features are extracted through Graph Neural Networks (GNNs) and then categorized using a YOLOv7 detector. The HRSID dataset was used to train the automatic ship detection system; a comparison was made with various CNNs (YOLOv3, YOLOv4, YOLOv5 and YOLOv6) presented by other authors [4–13]. We evaluated our approach using a publicly accessible ship dataset made up of around 16 K satellite photos that also included moving ships.

The summary of the contribution is as follows:

(1) For ship detection, a high-resolution SAR dataset is used. It was not able to take into account the various flaws in the previous SAR ship dataset, which is mainly used for CNN-based detectors.
(2) The goal of this paper is to analyze the effects of ship detection on the images captured by the SAR system. A large-sized image of the ship is used to test the model's performance.
(3) A comprehensive evaluation of ship detection is performed using MS COCO metrics. The IoU threshold of objects is evaluated using an average precision. An HRSID comparison between different YOLO versions is also carried out.

The organization of paper is as follows. The related work and state of the art on ship detection from satellite images, including those that employ DL algorithms for classification, is briefly summarized in Section 2. Then, in Section 3, we provide our DL-based YOLOv7_GNN ship detection approach. The examined datasets and a thorough examination of the stated outcomes resulting from the executed experiments are both provided in Section 4. The key pertinent conclusions from this study are presented in Section 5, along with a list of future research topics.

2. Related Work and State of the Art

2.1. Ship Detection Data Collections Platform

According to Kanjir et al. [14], optical, infrared, and radar sensors are the most frequently utilized sensors in sea surveillance applications. Since the 1990s, radar has been a common technology for ship surveillance and detection. One of the most common types of ship detection that can be performed is by using a satellite-based sensor. This method involves collecting data from various sources.

2.2. Improvement in Deep Learning (DL)

The amount of data that are now being saved grows daily. As there are more datasets available, researchers are constantly attempting to improve the algorithms that are currently in use. There are typically many layers of deep learning that are used in the processing of complex information. These include the input and output layers of the deep learning classifier. In order to compare the three machine learning algorithms, Chua and colleagues presented a comparison of the SVM, the histogram of oriented gradient, and the latent SVM.

In order to detect objects with complex backgrounds and scale variation, Kanjir et al. [14] developed a fully convolutional neural network algorithm. The algorithm uses a combination of box regression and CNN to label the class [15].

Through a combination of deep learning and CNN, Jaafar Alghazo [16] was able to create two models that can detect ships in the Airbus Satellite dataset. These models

can be used to deal with various maritime-related problems, such as illegal fishing and resource surveillance.

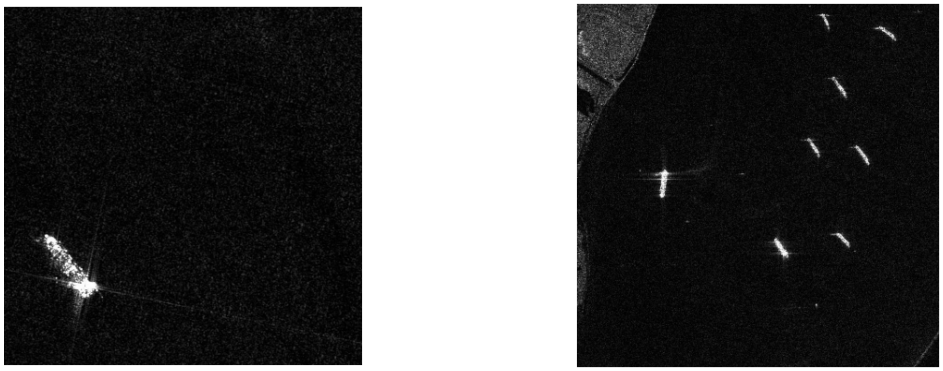
Researchers have developed a new technique that allows them to recognize ships using satellite images. The method, known as R-CNN, is mainly used for analyzing the images taken by the RADARSAT-2 and Sentinel-1 satellites [17]. The models were able to recall and improve their accuracy by 89.14% and 89.23%, respectively.

In order to create a generative transfer learning framework that can be used for ship detection, X. Lou [18] proposed a method that combines knowledge transfer and ship recognition. The output of this module was fed into a detector model. The goal of the experiment was to analyze the characteristics of the ship detection datasets taken from the Air-SAR-Ship-1.0 and SAR Ship Detection datasets.

3. Dataset and Our Method

3.1. Dataset

The HRSID [19,20] is a repository for ship detection in high-resolution SAR images. It contains over 16,951 instances of ships, as well as 5604 high-resolution SAR photos. The HRSID was created using the COCO datasets, which include images with varying marine areas, resolutions, water conditions, and coastal ports. These allow researchers to compare their methods against those of other researchers. Figure 1a,b show the images from HRSID repository. The three resolutions of the HRSID's high-resolution SAR images are: 1 m, 3 m, and 0.5 m. Table 1 displays some detailed information of the images from HRSID dataset.



(**a**) (**b**)

Figure 1. (**a**,**b**) A few images of 800 × 800 pixels from the HRSID dataset.

Table 1. Information of the images from the HRSID dataset.

Satellite	Resolution	No. of Images	No. of Ships
Sentinel-1B	3	99	13,471
TerraSAR-X	3	31	2924
TerraSAR-X	0.5	4	190
TerraSAR-X	1	1	280
TanDEM	1	1	86

3.2. Our Approach

Based on the YOLOv4, Scaled YOLOv4, and YOLO-R YOLO model architectures, the YOLOv7 architecture was developed.

The YOLOv7, an Extended Efficient Layer Aggregation Network (E-ELAN) architecture, is a framework that enables the continuous improvement of the learning capabilities of the network by implementing various features such as shuffle, expand, and merge cardinality. This allows the network to maintain its learning performance even when the gradient route is changed.

Compound model scaling is a process that involves modifying the characteristics of a model to improve its performance in various applications. For instance, it can help to improve the model's depth, width, and resolution. Different scaling considerations in conventional techniques with concatenation-based architectures (such as ResNet or PlainNet) must be taken into account collectively rather than separately. For instance, increasing model depth will affect the ratio between a transition layer's input and output channels, which may result in less hardware being used by the model. For a concatenation-based model, YOLOv7 introduces compound model scaling as shown in Figure 2. The compound scaling method can preserve the model's original design elements. It can be used to modify the output channel and depth factor of a computational block. This process can be performed in order to maintain the model's ideal structure. For instance, changing the block's depth factor can affect the output channel.

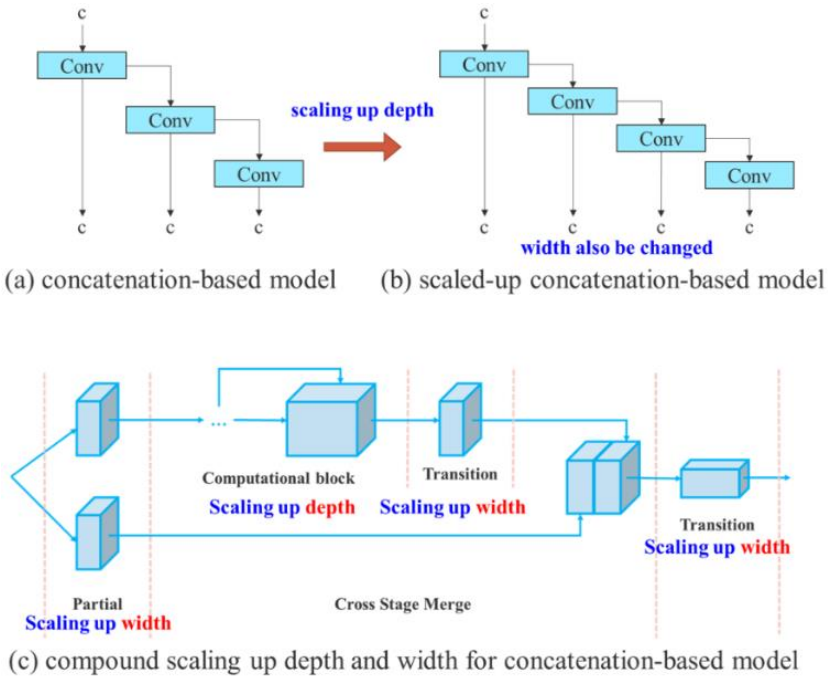


Figure 2. Working of YOLOv7: compound model scaling [21].

Despite being an excellent VGG architecture, the planned re-parameterized version of RepConv loses significant accuracy when applied to either DenseNet or ResNet. In Figure 3, YOLOv7, the architecture is presented without an identity connection. Concatenation or residual is used to replace a previously created layer with a re-parameterized convolutional layer.

A head, a neck, and a backbone are parts of a YOLO architecture as shown in Figure 4. The projected model outputs are located in the head. YOLOv7 is suitable for lead loss and auxiliary. The aim of this paper is to create a neural network that is capable of training various types of neural networks. It is inspired by deep supervision, a method that involves training a neural network. The generation of final product is the responsibility of the lead head, whereas the auxiliary head supports middle-layer training. Figure 5 demonstrates the general flow of the ship detection model from satellite images. In the dataset preprocessing step in Figure 5, first, the division of the dataset is made with a ratio of 6:4 for training set and testing set. For the identification of the ship, we consider the area of the bounding box and aspect ratio of the bounding box. The shape of the bounding box matches the aspect ratio of the bounding box, which is helpful in adopting an anchor for generating the bounding boxes.

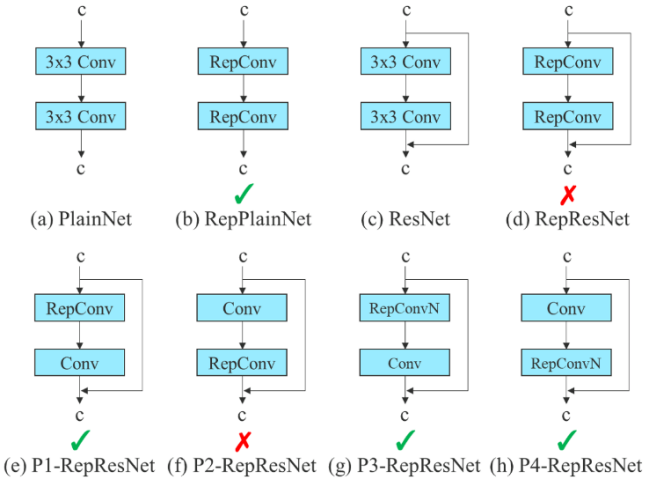


Figure 3. The YOLOv7 architecture, the planned re-parameterized version of RepConv. RepConv is replaced by RepConvN as RepConv does not have an identity connection [21].

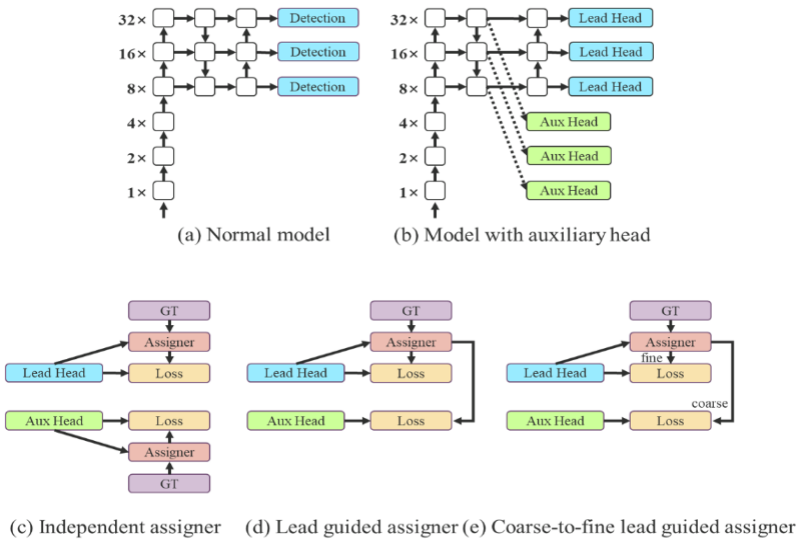


Figure 4. The YOLOv7 framework features a lead head and an auxiliary head that are guided by a label assigner [21].

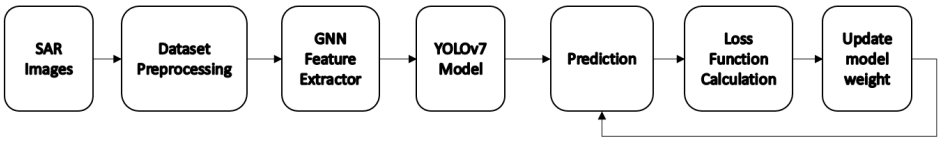


Figure 5. Flow diagram of the ship detection model.

4. Results

4.1. Setup

We develop a classifier in this paper using a particular YOLOv7 architecture. The layer number and parameter adjustments listed in Table 2 are used to finetune the method. Plotting the figures and handling data are carried out using the Python platform. These experiments also use the modules numpy, open cv, and pandas. The standard Keras library is used to download the pretrained DenseNet parameters. However, because our computer's RAM is limited and since no graphics processing unit is being employed in the trials, the batch size will also be altered.

Table 2. Hyperparameters for the ship detection model.

Parameters	Values
Optimizers	Adam, SGD
Pooling Layers	Average Pooling
Batch Size	16
Activation	ReLu
Learning rates	0.01 and 0.001

4.2. Optimization

To update the network's parameter in order to train it, an optimization algorithm is required. It focuses on the discrepancies between model predictions and the term used to refer to the actual situation. The Adam optimizer is a technique for stochastic objective function optimization using first-order gradients. It is a computationally effective optimizer that can be used with the majority of data types. It also uses a small amount of memory. One of the oldest optimizers, the stochastic gradient descent (SGD) optimizer, does not use momentum while determining the update weights. A performance comparison between the Adam and SGD optimizer for ship detection is shown in Table 3 and Figure 6.

Table 3. Comparison between Adam and SGD optimizer.

Batch Size	Learning Rate	Optimizer	No. of Epoch	Accuracy
16	0.001	Adam	12	93.4
16	0.01	SGD	12	91.3

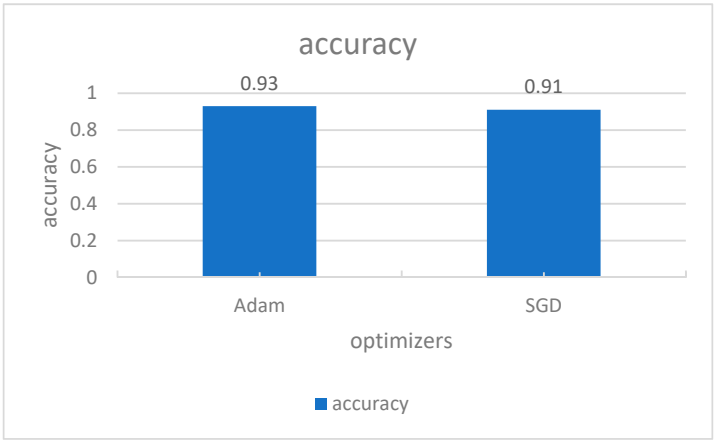


Figure 6. Comparison between Adam and SGD optimizer of ship detection.

According to Table 3, Adam optimizer is utilized with 93.4% accuracy to achieve the best classification performance. Compared with a learning rate of 0.01, a learning rate of 0.001 was used to achieve this performance. The batch size for various kinds of optimization is 16.

4.3. Divisioning of Dataset

Since this paper employs supervised learning, the training data must be labeled so that they can be used to train the algorithms. The set of non-overlapping photos used as testing data is used to gauge how reliable the classifier's predictions are. This batch of data has never been seen before. It is divided into three different scenarios. The findings in Table 4 demonstrate that there are only minor performance variations among all ratios of data splitting between training and testing, achieving the highest accuracy. The batch size and learning rate are both fixed at 16, while all other hyperparameters are set to 0.001.

Table 4. Accuracy comparison between training and testing images.

No. of Training Images	No. of Testing Images	Size of Batch	Epoch	Learning Rate	Accuracy
2200	1200	16	75	0.001	91.06%
3000	600	16	75	0.001	91.41%
3600	200	16	75	0.001	91.87%

4.4. Batch Size

The number of samples that are processed during one training iteration is determined by the hyperparameter known as batch size. The method will process the first 16 photos (from the first to the 16th) if the batch size is 16 samples, for instance. The algorithm will then obtain further 16 images (from positions 17 to 32), and the second iteration will continue with the same process until all of the images are processed for a particular epoch operation. A performance comparison between the different batches is shown in Table 5. A batch size of 16 for the 3000 training images and 600 testing images produces accuracy of 86.09%, which is higher than that of 84.01% for 2200 and 1200 training and testing images.

Table 5. Comparison between same batch size and different no of training and testing images.

No. of Training Images	No. of Testing Images	Size of Batch	Epoch	Learning Rate	Accuracy
2200	1200	16	10	0.001	84.01%
3000	600	16	10	0.001	86.09%

4.5. Learning Rate

The weights of the updated gradient error are affected by the learning rate. It specifically regulates the number of errors that the model's weights will take into account when it is updated. The learning rates of 0.001 and 0.01 are examined. In comparison to a learning rate of 0.01, the learning rate of 0.001 results in a substantially higher accuracy as shown in Table 6.

Table 6. Comparison between different learning rates.

No. of Training Images	No. of Testing Images	Size of Batch	Epoch	Learning Rate	Accuracy
2200	1200	16	10	0.001	93.4%
2200	1200	16	10	0.01	86.73%

Figure 7 demonstrates a few results of ship detection from HRSID using our approach, YOLOv7+GNN. A comparison of YOLOv7 and other object detection algorithms such as YOLOv4, YOLOv5, and many more with respect to speed and accuracy has been presented by Chien-Yao Wang [21].

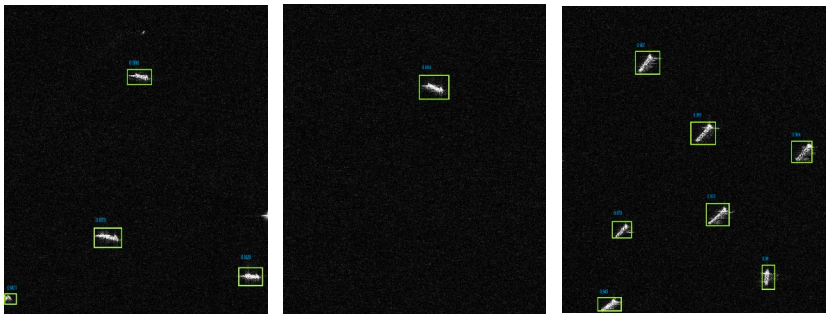


Figure 7. Ship detection result using our approach, YOLOv7 + GNN.

5. Conclusions

In this article, we provide a YOLOv7+GNN-based technique for automatic ship detection from the High-Resolution Satellite Image Dataset (HRSID), which is a high-resolution image of a synthetic aperture radar. This algorithm classifies ships with greater than 90% accuracy. The learning rate of the framework is as low as 0.001, and a batch size of 16 can produce the optimal results. For ship detection, the researchers used the Adam optimizer and YOLOv7 with GNN. The dataset's picture count can be enhanced to include a wider range of ship types in order to produce better results in the future.

Author Contributions: K.P.: data preprocessing, methodology investigation, writing—original draft, C.B.: writing—review and editing, methodology, project administration, P.L.M.: review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available in a publicly accessible repository that does not issue DOIs. Publicly available datasets were analyzed in this study. This data can be found here: https://github.com/chaozhong2010/HRSID (accessed on 26 September 2022).

Acknowledgments: We are thankful to Pier Luigi Mazzeo for his insightful and helpful advice during the planning and writing of this research paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zulkifley, M.A.; Abdani, S.R.; Zulkifley, N.H. Pterygium-Net: A deep learning approach to pterygium detection and localization. *Multimed. Tools Appl.* **2019**, *78*, 34563–34584. [CrossRef]
2. Patel, K.; Bhatt, C.; Mazzeo, P.L. Deep learning-based automatic detection of ships: An experimental study using satellite images. *J. Imaging* **2022**, *8*, 182. [CrossRef] [PubMed]
3. Kothadiya, D.; Chaudhari, A.; Macwan, R.; Patel, K.; Bhatt, C. The Convergence of Deep Learning and Computer Vision: Smart City Applications and Research Challenges. In Proceedings of the 3rd International Conference on Integrated Intelligent Computing Communication & Security (ICIIC 2021), Bengaluru, India, 13 September 2021; Atlantis Press: Paris, France; pp. 14–22.
4. Laroca, R.; Severo, E.; Zanlorensi, L.-A.; Oliveira, L.-S. A Robust Real-Time Automatic License Plate Recognition Based on the YOLO Detector. *arXiv* **2018**, arXiv:1802.09567v6.
5. Zhu, X.; Lyu, S.; Wang, X.; Zhao, Q. TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-captured Scenarios. *CVF* **2021**, *2108*, 11539.
6. Lee, Y.-H.; Kim, Y.-S. Comparison of CNN and YOLO for Object Detection. *J. Semicond. Disp. Technol.* **2020**, *19*, 1.
7. Ultralytics/Yolov5. Available online: https://github.com/ultralytics/yolov5 (accessed on 25 June 2020).
8. Jia, W.; Xu, S.; Liang, Z.; Zhao, Y.; Min, H.; Li, S.; Yu, Y. Real-time automatic helmet detection of motorcyclists in urban traffic using improved YOLOv5 detector. *IET Image Process.* **2021**, *10*, 1049. [CrossRef]
9. Bohara, M.; Patel, K.; Patel, B.; Desai, J. An AI Based Web Portal for Cotton Price Analysis and Prediction. In Proceedings of the 3rd International Conference on Integrated Intelligent Computing Communication & Security (ICIIC 2021), Bengaluru, India, 13 September 2021; Atlantis Press: Paris, France; pp. 33–39.
10. Redmon, J.; Farhadi, A. YOLOv3: An Incremental Improvement. *arXiv* **2018**, arXiv:1804.02767.
11. Bochkovskiy, A.; Wang, C.-Y.; Mark Liao, H.-Y. YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv* **2020**, arXiv:2004.10934.
12. Zhang, H.; Tian, M.; Shao, G.; Cheng, J.; Liu, J. Target Detection of Forward-Looking Sonar Image Based on Improved YOLOv5. *IEEE Access* **2022**, *2022*, 3150339. [CrossRef]
13. Kasper-Eulaers, M.; Hahn, N.; Berger, S.; Sebulonsen, T.; Myrland, Q.; Kummervold, P.-E. Short Communication: Detecting Heavy Goods Vehicles in Rest Areas in Winter Conditions Using YOLOv5. *Algorithms* **2021**, *14*, 114. [CrossRef]
14. Kanjir, U.; Greidanus, H.; Oštir, K. Vessel detection and classification from spaceborne optical images: A literature survey. *Remote Sens. Environ.* **2018**, *207*, 1–26. [CrossRef] [PubMed]
15. Audebert, N.; Le Saux, B.; Lefèvre, S. Segment-before-detect: Vehicle detection and classification through semantic segmentation of aerial images. *Remote Sens.* **2017**, *9*, 368. [CrossRef]
16. Alghazo, J.; Bashar, A.; Latif, G.; Zikria, M. Maritime ship detection using convolutional neural networks from satellite images. In Proceedings of the 10th IEEE International Conference on Communication Systems and Network Technologies (CSNT), Bhopal, India, 18–19 June 2021; pp. 432–437.

17. Huang, X.; Zhang, B.; Perrie, W.; Lu, Y.; Wang, C. A novel deep learning method for marine oil spill detection from satellite synthetic aperture radar imagery. *Mar. Pollut. Bull.* **2022**, *179*, 113666. [CrossRef] [PubMed]
18. Lou, X.; Liu, Y.; Xiong, Z.; Wang, H. Generative knowledge transfer for ship detection in SAR images. *Comput. Electr. Eng.* **2022**, *101*, 108041. [CrossRef]
19. HRSID Dataset. Available online: https://github.com/chaozhong2010/HRSID (accessed on 26 September 2022).
20. Shunjun, W.; Xiangfeng, Z.; Qizhe, Q.; Mou, W.; Hao, S.; Jun, S. HRSID: A High-Resolution SAR Images Dataset for Ship Detection and Instance Segmentation. *IEEE Access* **2020**, *8*, 120234–120254.
21. Wang, C.Y.; Bochkovskiy, A.; Liao, H.Y.M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *arXiv* **2022**, arXiv:2207.02696.