# Detection of Ships Cruising in the Azimuth Direction Using Spotlight SAR Images with a Deep Learning Method

**Takero Yoshida** [1,2,]*[ ] **and Kazuo Ouchi** [2]

1   Department of Ocean Sciences, Tokyo University of Marine Science and Technology, Minato-ku, Tokyo 108-8477, Japan
2   Institute of Industrial Science, The University of Tokyo, Kashiwa-shi, Chiba 277-8574, Japan
*   Correspondence: tyoshi3@kaiyodai.ac.jp or tyoshida@iis.u-tokyo.ac.jp; Tel.: +81-3-5463-0639

**Abstract:** Spotlight synthetic aperture radar (SAR) achieves a high azimuth resolution with long integration times. Meanwhile, the long integration times also cause defocused and smeared images of moving objects such as cruising ships This is a typical imaging mechanism for moving objects in Spotlight SAR images. Conversely, ships can be classified as stationary or moving from the amount of smearing, and this classification method is, in general, based on manual observation. This paper proposes an automatic method for detecting cruising ships using deep learning known as the "You Only Look Once (YOLO) v5 model", which is one of the frameworks of the YOLO family. In this study, ALOS-2/PALSAR-2 L-band Spotlight SAR images over the waters around the Miura Peninsula, Japan, were analyzed using the YOLO v5 model with a total of 53 ships' images and compared with Automatic Identification System (AIS) data. The results showed a precision of approximately 0.85 and a recall rate of approximately 0.89 with an *F*-measure of 0.87. Thus, sufficiently high values were achieved in the automatic detection of moving ships using the deep learning method with the YOLO v5 model. As for false detections, images of breakwaters were classified as ships cruising in the azimuth direction. Further, range moving ships were found to be difficult to detect. From the present preliminary study, it was found that the YOLO v5 model is limited to ships cruising predominantly in the azimuth direction.

**Keywords:** Spotlight SAR image; ship detection; deep learning; YOLOv5

## 1. Introduction

In recent years, deep learning has been applied in various fields including image analyses. It is composed of a multi-layer neural network and is often based on a convolutional neural network (CNN) [1,2]. In the case of optical and synthetic aperture radar (SAR) remote sensing, deep learning is used mainly for target detection, recognition and classification on land and ocean.

It is well known that SAR has all-weather and day-and-night imaging capability, and, in ocean applications, many studies have been focused on ship detection [3–11]. In addition, studies using deep learning have also been reported for ship detection [12–19] and classification [20–22].

Since these studies focused on all types of ships, regardless of whether they were moving or stationary, the current study differs from them in that it focuses on detecting only cruising ships. Information on cruising ships can be useful for maritime domain awareness, such as for the detection of suspicious ships and illegal fishing ships. In previous studies, cruising ships were detected using ship wakes [23,24]. However, ship wakes may not always be observed in SAR images in high sea states.

This study focuses on Spotlight SAR images, which have the finest imaging mode with long azimuth integration times in comparison with other modes. If targets move during the long integration time, the images are defocused and smeared in the azimuth

direction. These extended image features are less significant in other imaging modes such as the Stripmap and ScanSAR modes of shorter integration times. In the present study, the deep learning algorithm is used for detection of cruising ships by taking into account the motion-induced large image smearing in Spotlight SAR images.

In a previous study [25], a method using sub-look processing for Spotlight data was proposed to estimate the azimuth velocity components of cruising ships and the velocity vector from the inter-look position difference. This approach requires first to detect the candidate images of ships. As discussed in the following sections, both the detection and velocity estimation of cruising ships can also be carried out by applying deep learning, as proposed in this study. Among the different algorithms of deep learning, the You Only Look Once (YOLO) model was considered in the present study. YOLO is a well-known model with a high accuracy and fast learning speed [26]. Specifically, YOLOv5 [27], which is a relatively newer version of the YOLO family, was used to demonstrate the automatic detection of cruising ships. The present study aims to demonstrate the automatic detection of cruising ships in Spotlight SAR images with the YOLOv5 model and velocity estimation by the sub-look algorithm [25].

In the following sections, the Spotlight SAR data used for training, validation and test are described first, followed by the deep learning and evaluation methods. The results are then presented and discussed with concluding remarks.

## 2. Materials and Methods

### 2.1. Spotlight SAR Data

Three sets of ALOS-2/PALSAR-2 Spotlight SAR images were used for training and validation of the YOLOv5 model, as shown in the upper row in Figure 1. The area of interest was Tokyo Bay, east coast of the Miura Peninsula, Japan, where many cruising ships were observed. The scene IDs are ALOS2047162919-150408, ALOS2052782911-150516 and ALOS2063722919-150729 from left to right. The first and third images are in HH-polarization, while the second image is in VV-polarization. All the data were in the descending mode with integration times of approximately 20–30 s. The resolution was 1 m × 3 m in the azimuth and range directions, with the same pixel sizes in the corresponding directions. Prior to applying the deep learning algorithm, the images were divided into 25 sections, as shown in the upper-right section of Figure 1 (each image size is approximately 8000 × 7850 pixels). Then, the sections including the images of cruising ships were selected, and these images were used as the training datasets.
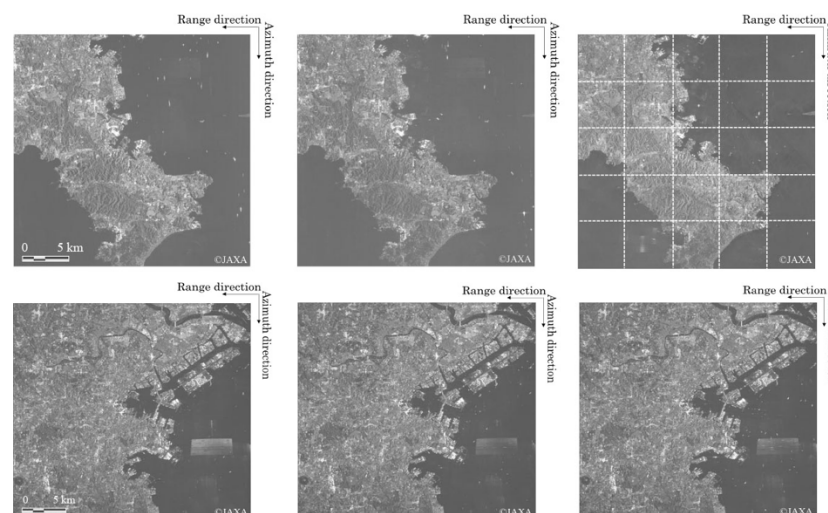


**Figure 1.** Spotlight SAR images of the waters around the Miura peninsula, Japan. The images of ships in Tokyo Bay on the right in the upper images are used for training and validation of the deep learning algorithm. For computational purposes, each image is divided into 25 small sections, as shown in the upper-right image. The lower 3 images are the samples used for the test.

The four left images in Figure 2 are examples of cruising ships' images extracted from the training datasets, where images are defocused and extended in the azimuth direction. An example of the image of a stationary ship is also shown on the right of Figure 2. Unlike the cruising ships, the image of a stationary ship is well focused and often contains a strong point-like image. The difference between these imaging characteristics can be used to classify moving and stationary ships in the Spotlight SAR images.
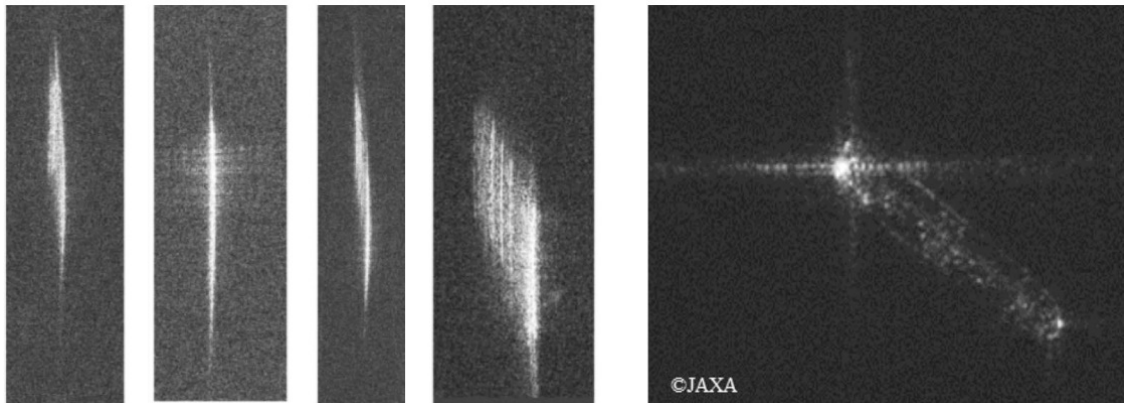


**Figure 2. Left**: Examples of 4 defocused and azimuth-extended images of cruising ships due to long integration times extracted from Figure 1. **Right**: Focused image of a stationary ship. The directions are the same as those in Figure 1.

The test data are shown in the lower row in Figure 1. The scene IDs are ALOS2018020809-140922, ALOS2068602912-150831 and ALOS2062392912-150720 from left to right. These data are all in the descending mode and HH-polarization, with the same resolution and pixel sizes as those of the training data. In order to examine the detection accuracy, the images of look-alikes, such as the ghost images from land areas, RFI (Radio Frequency Interference) and the images of breakwaters are discussed.

In the analysis, the effect on polarization was not considered since the defocused and azimuth extended images of moving ships appeared to be almost the same, regardless of the HH- and VV-polarizations.

*2.2. Deep Learning Method*

YOLO is a real-time object detection model that has a high accuracy and fast learning speed [26]. In the current research, YOLOv5 [27], a comparatively newer type than the past YOLO series [28–30], was used to build a model for the detection of cruising ships. YOLOv5 also has short learning times and is easy to use. In addition, YOLOv5 has better memory requirements and exportability. In particular, the GPU memory efficiency is high in YOLOv5 during the learning process. The models of S, M, L, and X are the options that exist in YOLOv5, and, in general, the accuracy of the X-model is the highest. In this study, the YOLOv5 X-model was selected for the detection of cruising ships.

The network architecture of YOLOv5 is briefly explained as follows: YOLOv5 consists of three parts: the backbone, neck, and head. The data are processed to the backbone for feature extraction, to the neck for feature fusion, and to the head for the output. Backbone extracts the key features from the input images by incorporating a cross-stage partial network [31]. This improves the processing time and ensures a reduced model size. The neck is used to generate feature pyramids, which are helpful for detecting the same object at different scales. A path aggregation network [32] was used as the neck. It enables the enhanced localization of the targets. The head generates different-sized feature maps. It can obtain the final output with class probabilities (confidence scores) and bounding boxes.

In order to prepare the information on training data for cruising ships, the annotation tool of LabelImg (Available at: https://github.com/tzutalin/labelImg (accessed on 18 September 2022)) was applied to label each moving ship with a rectangular box in the

YOLO text format. To increase the amount of training data, the images were augmented in terms of brightness, contrast, and rotated in the opposite azimuth direction. The image rotation took into account the extended image of a cruising ship on both sides from the center.

Figure 3 shows an example of an annotation with rectangular boxes. In total, 84, and 304 labels were used for the moving ships. The data were trained with eight batches and 1000 epochs using YOLOv5. The number of batches was determined by the GPU memory. A size of 640 × 640 pixels within each small image section in the upper-right section in Figure 1 was used during the training process. The computer specifications with CPU: AMD RYZEN 5 2600, GPU: NVIDIA Geforce RTX 2080 Ti GPU, and RAM: 32 GB DDR4-2666 (TSUKUMO, Chiyoda-ku, Tokyo) were used. With this system, the time for the training process was approximately 12 h. The evaluation parameters were convergence by 1000 epochs with some fluctuations. Thus, the number of epochs appeared to be sufficient for accurate training. The model with the highest performance during the training process of 1000 epochs was selected as the best model.
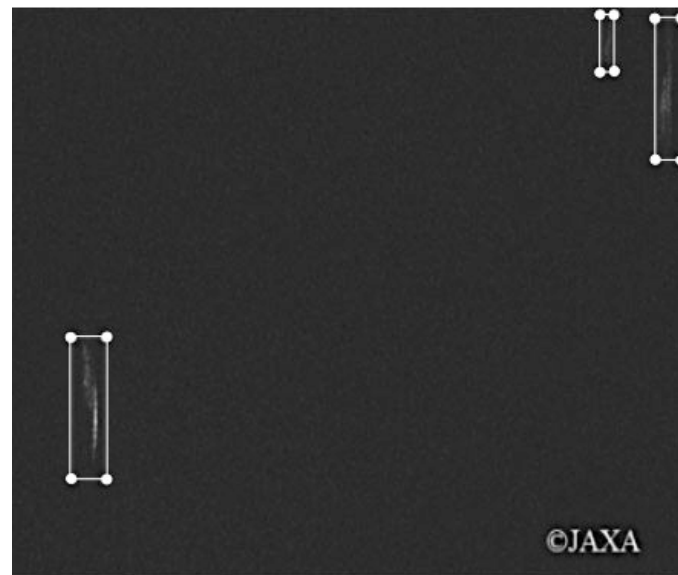


**Figure 3.** Annotation for the images of moving ships. The white bounding boxes are the annotation for creating the training model.

*2.3. Evaluation Method*

The detected results are classified into three patterns: The first pattern defines the bounding box that made the correct detection; i.e., it is a true positive (*TP*), which means the samples identified are positive. The second pattern defines the bounding box that made the incorrect detection; i.e., it is a false positive (*FP*), meaning that the samples were incorrectly identified as positive. The third pattern defines that the object was not detected by the bounding box; i.e., it is a false negative (*FN*), which means that the samples were incorrectly identified as negative. Based on the aforementioned patterns, two types of evaluation factors, precision (*P*) and recall (*R*), can be defined. *P* and *R* can be calculated by the following equations.

$$P = \frac{TP}{TP + FP} \tag{1}$$

$$R = \frac{TP}{TP + FN} \tag{2}$$

Using these parameters, the *F*-measure (*F*), which is commonly used to compare detection performances, can be expressed by the following equation:

$$F = 2\frac{PR}{P + R} \tag{3}$$

Another evaluation factor is average precision (AP). It refers to the area enclosed by the precision–recall curve. It can be expressed by the following equation:

$$AP = \int_0^1 P(R)dR \tag{4}$$

The average precision represents the identification accuracy of a single category. A higher AP value indicates better detection performance using the deep learning model.

### 3. Results

A detection model was built using the aforementioned training data shown in the upper row in Figure 1. A three-fold cross-validation was carried out by combining the training and validation image datasets. Cross-validation is widely used to reduce the effect of over-fitting, which may occur during training in deep learning models. In this study, 56 images were used for training and 28 images were used for validation from a total of 84 images (304 labels).

Table 1 shows the precision, recall, and *F*-measure values of the S- and X-model during the training process. The values are the best parameters, which showed the highest *AP* during the entire training process using the YOLOv5 model, and the mean values of three best parameters from the three-fold cross-validation are listed in Table 1.

**Table 1.** Mean values of precision, recall, and *F*-measure for 3-fold cross-validation with the YOLOv5 S- and X-models.

|  | *P* | *R* | *F*-Measure |
|---|---|---|---|
| S-model | 0.819 | 0.844 | 0.831 |
| X-model | 0.854 | 0.894 | 0.874 |

From the results, the X-model showed a higher detection capability with precision and recall rates of approximately 0.85 and 0.89, respectively. Thus, the recall rate was sufficiently high to facilitate the accurate detection of moving ships in Spotlight SAR images with approximately a 10% error, i.e., 10% misdetection, and therefore the recall rate higher than approximately 0.9 can be used for automatic detection. In addition, the mean value of the *F*-measure was approximately 0.87 during the three-fold cross-validation, and it is reasonable to conclude that these values are sufficiently high to ensure the effectiveness of the detection model.

Figure 4 shows an example of the result of the automatic detection, showing both the almost stationary and moving ships. The detected bounding box indicates the moving ship on the bottom-left side of the image, whereas the image of an almost stationary ship on the upper-right is not classified as a moving ship. As mentioned, the images of moving ships have the unique characteristic of an extended outline in the azimuth direction, which is used for automatic detection. However, other images also show similar characteristics, including images of breaking waves known as azimuth streaks [33]. Such images are not considered in the present model, and are left for further study.

**Figure 4.** Example of detecting cruising ships and avoiding stationary ships. The left bottom image within the white bounding box is a cruising ship successfully detected with a high confidence score of 0.94. The right two images of ships are defined as stationary.

## 4. Discussion

Following the training and validation of the YOLOv5 model, further detection tests were conducted using the Spotlight SAR images shown at the bottom of Figure 1. Here, the test images were close to the port areas that included not only moving ships but also other objects such as azimuth ambiguity from land, breakwaters, and moored ships at the port.

Figure 5 shows examples of detection results using the test images with the best model developed in the previous section. A comparison was made between the detected images and AIS data. Note, again, that the test images were not used in the training and validation processes. In the following, the results of correct detected and undetected cruising ships are described. Misdetection and detection in the presence of ghost images and RFI are discussed along with the detection of range moving ships.

Figure 5a shows an example of a successful detection of three cruising ships with confidence scores higher than 0.9, as indicated in the brackets. One ship was not detected as it was stationary according to the AIS data. The results using 18 test images are summarized in Table 2, where $TP$, $FP$, and $FN$ the true positive ($TP$), false positive ($FP$) and false negative ($FN$), respectively, as explained earlier, and non-AIS shows the number of detected ships without AIS data. Table 2 also shows the $P$, $R$, and $F$-measure calculated by these parameters. Comparing Table 1 with Table 2, the values in the test case in Table 2 were lower than those of the training and validation cases in Table 1.

**Table 2.** Detection results using the test images confirmed by AIS data. *TP* and *FP* are the numbers of correctly detected moving ships and misdetected ships, respectively. *FN* is the number of moving ships but not detected, and non-AIS is that of detected ships without AIS data.

| TP | FP | FN | Non-AIS | P | R | F-Measure |
|----|----|----|---------|---|---|-----------|
| 28 | 10 | 5 | 10 | 0.7370 | 0.848 | 0.789 |

Figure 5b–d show examples of correct detection, misdetection and undetected. In Figure 5b, the breakwaters are classified as moving ships for their similar shapes to the azimuth extended images of cruising ships. There are several ghost images as a result of azimuth ambiguity shown in the elongated circles, but the model correctly disregarded these as moving ships. In Figure 5c, the cruising ship on the top-left was not detected for the reason that the image was overlaid over the RFI, and images of this type were not used in the training process. In Figure 5d, four azimuth moving ships were successfully detected, but the range moving ship on the left was not detected mainly due to the fact that the image was extended in the azimuth direction.
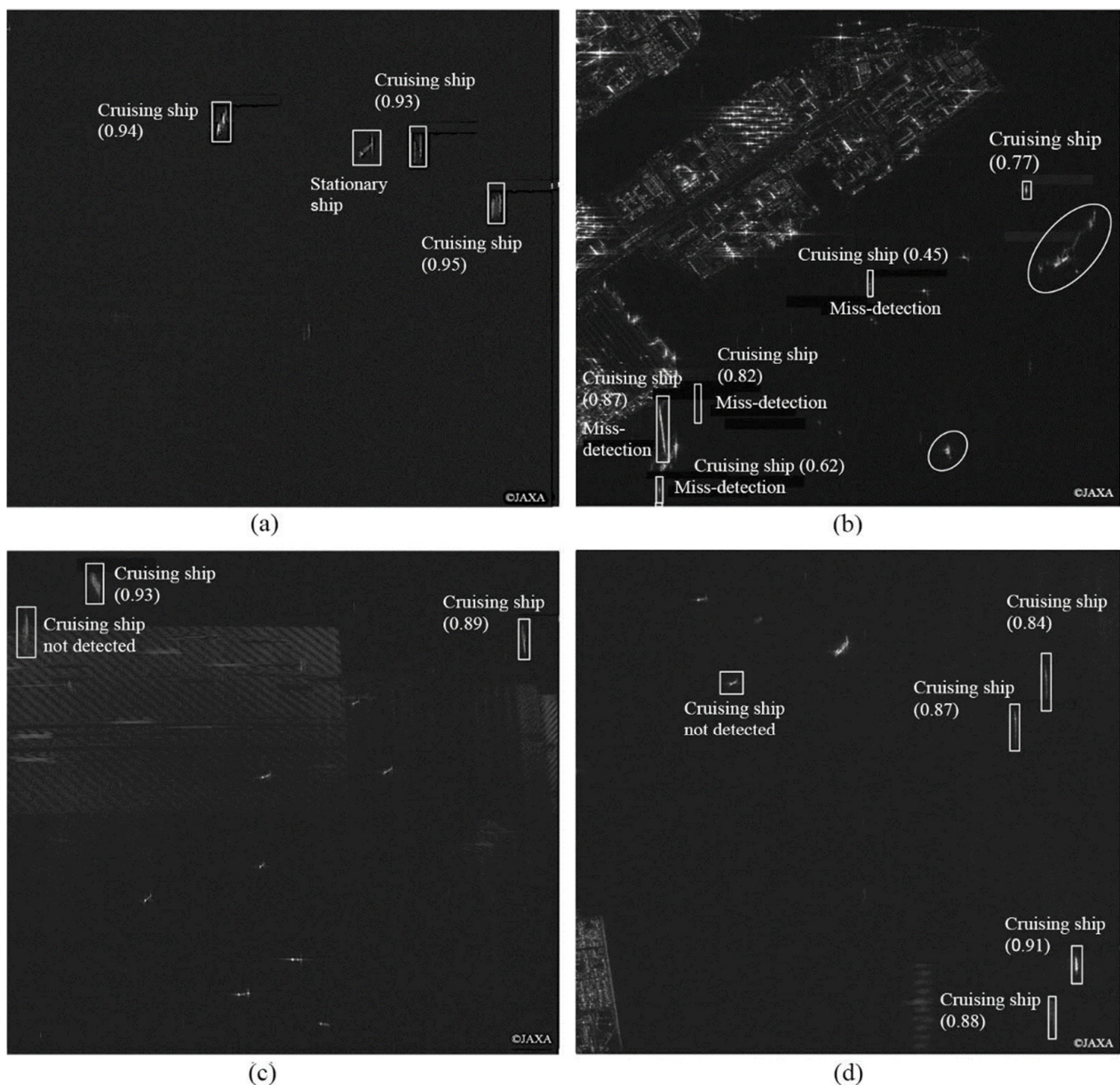
**Figure 5.** Examples of successful and misdetection results in the test images that were not used in the training and validation processes. The azimuth and range directions are from top to bottom and right to left, respectively. (**a**) Successful results for detecting moving ships. The numbers in brackets are the confidence score. (**b**) Results of detection and misdetection. The model could avoid the misdetection of ghost images in the white ellipses, while parts of breakwaters were incorrectly detected. (**c**) Two detected cruising ships and an undetected image over the Radio Frequency Interference (RFI). (**d**) Four detected cruising ships on the right-hand while a ship cruising in the range direction is not detected (image size is approximately 8000 × 7850 pixels).

When a ship is traveling to the range direction, an azimuth image shift also occurs and the original position is shifted in the azimuth direction by the distance proportional to the slant-range velocity component, and the image is not well focused. Figure 6 shows the images of range moving ships in the Spotlight image, and the white broken lines are the trails of the corresponding ships from the AIS data. The size and speed are also depicted. The small ship of length 15 m is not clear, while the image of the large ship of length 113 m is clear although it is defocused. Thus, small ships moving fast in the range direction are difficult to detect using the present model.
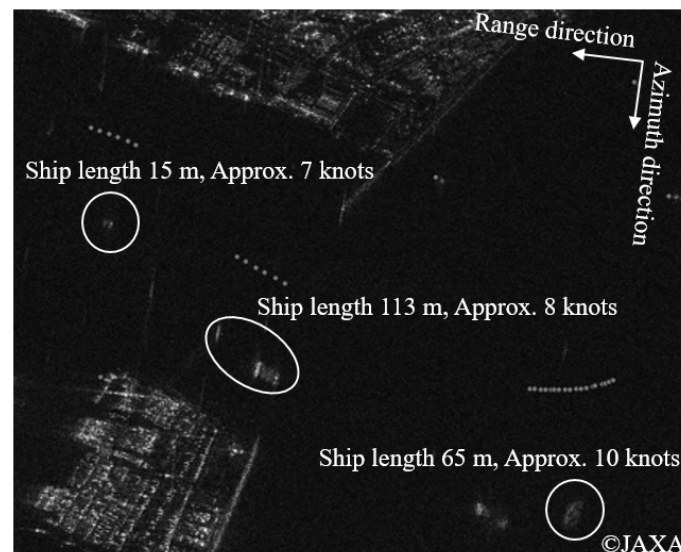
**Figure 6.** Example of range moving ships in a Spotlight SAR image compared with AIS data. The continuous dots show the trail of ship by AIS over one minute. The small ship moving to the range direction is difficult to image; meanwhile, the relatively large ships moving in the range direction can be imaged but not focused well.

Owing to the limited training data used in this study, additional training data of range moving ships will be required in future studies.

## 5. Conclusions

A deep learning method for detecting cruising ships in the azimuth direction was presented using the YOLOv5 model and ALOS-2 Spotlight images. The model was trained with image characteristics of the azimuthally extended feature of azimuth moving ships. From the three-fold cross-validation, the results showed fairly high mean precision and recall rates of 0.85 and 0.89, respectively, and in addition, the mean *F*-measure was 0.87. By comparison of the results with AIS data, the confidence scores were found to be low in the case of almost stationary or slow-moving ships of speed less than approximately 0.5 knots. While the confidence scores were high for detecting moving ships with clear azimuth extended image features and avoiding ambiguous ghost images, the misdetection of objects such as breakwaters was found as their image shapes are similar to those of azimuth moving ships. As for range moving ships, the detection is difficult using the present model. This study showed the methodology and effectiveness of the YOLOv5 model for detecting the azimuth moving ships. Nevertheless, increasing data and application with the other deep learning models are desired to improve the detection model. For a future study, a new model is expected for both the detection and the velocity of moving ships by combining the present method and velocity estimation by the sub-look method [25]. In addition, our model was developed for water areas, and water–land mask could be an optional method to enhance the detection results.

**Author Contributions:** Conceptualization, T.Y. and K.O.; methodology, T.Y.; software, T.Y.; validation, T.Y. and K.O.; formal analysis, T.Y.; investigation, T.Y.; resources, T.Y.; data curation, T.YX.; writing—original draft preparation, T.Y.; writing—review and editing, K.O.; visualization, T.Y.; supervision, T.Y.; project administration, T.Y.; funding acquisition, T.Y. All authors have read and agreed to the published version of the manuscript.

**Data Availability Statement:** Not applicable.

## References

1. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014.
2. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *arXiv* **2015**, arXiv:1506.01497. [CrossRef] [PubMed]
3. Ouchi, K.; Tamaki, S.; Yaguchi, H.; Iehara, M. Ship Detection Based on Coherence Images Derived from Cross Correlation of Multilook SAR Images. *IEEE Geosci. Remote Sens. Lett.* **2004**, *1*, 184–187. [CrossRef]
4. Hwang, S.I.; Wang, H.; Ouchi, K. Comparison and Evaluation of Ship Detection and Identification Algorithms Using Small Boats and ALOS-PALSAR. *IEICE Trans. Commun.* **2009**, *E92-B*, 3883–3892. [CrossRef]
5. Hwang, S.I.; Ouchi, K. On a Novel Approach Using MLCC and CFAR for the Improvement of Ship Detection by Synthetic Aperture Radar. *IEEE Geosci. Remote Sens. Lett.* **2010**, *7*, 391–395. [CrossRef]
6. Brusch, S.; Lehner, S.; Fritz, T.; Soccorsi, M.; Soloviev, A.; Van Schie, B. Ship Surveillance with TerraSAR-X. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 1092–1103. [CrossRef]
7. Yin, J.; Yang, J.; Xie, C.; Zhang, Q.; Li, Y.; Qi, Y. An Improved Generalized Optimization of Polarimetric Contrast Enhancement and Its Application to Ship Detection. *IEICE Trans. Commun.* **2013**, *E96-B*, 2005–2013. [CrossRef]
8. Makhoul, E.; Baumgartner, S.V.; Jager, M.; Broquetas, A. Multichannel SAR-GMTI in Maritime Scenarios with F-SAR and TerraSAR-X Sensors. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 5052–5067. [CrossRef]
9. Marino, A.; Hajnsek, I. Ship Detection with TanDEM-X Data Extending the Polarimetric Notch Filter. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 2160–2164. [CrossRef]
10. Marino, A.; Sanjuan-Ferrer, M.J.; Hajnsek, I.; Ouchi, K. Ship Detection with Spectral Analysis of Synthetic Aperture Radar: A Comparison of New and Well-Known Algorithms. *Remote Sens.* **2015**, *7*, 5416–5439. [CrossRef]
11. Tello, M.; López-Martínez, C.; Mallorquí, J.J.; Greidanus, H. A Novel Algorithm for Ship Detection in ENVISAT SAR Imagery Based on the Wavelet Transform. *Eur. Space Agency* **2005**, *2*, 1557–1562.
12. Dechesne, C.; Lefèvre, S.; Vadaine, R.; Hajduch, G.; Fablet, R. Ship Identification and Characterization in Sentinel-1 SAR Images with Multi-Task Deep Learning. *Remote Sens.* **2019**, *11*, 2997. [CrossRef]
13. Chang, Y.L.; Anagaw, A.; Chang, L.; Wang, Y.C.; Hsiao, C.Y.; Lee, W.H. Ship Detection Based on YOLOv2 for SAR Imagery. *Remote Sens.* **2019**, *11*, 786. [CrossRef]
14. Mao, Y.; Yang, Y.; Ma, Z.; Li, M.; Su, H.; Zhang, J. Efficient Low-Cost Ship Detection for SAR Imagery Based on Simplified U-Net. *IEEE Access* **2020**, *8*, 69742–69753. [CrossRef]
15. Gao, F.; He, Y.; Wang, J.; Hussain, A.; Zhou, H. Anchor-Free Convolutional Network with Dense Attention Feature Aggregation for Ship Detection in SAR Images. *Remote Sens.* **2020**, *12*, 2619. [CrossRef]
16. Jiang, J.; Fu, X.; Qin, R.; Wang, X.; Ma, Z. High-Speed Lightweight Ship Detection Algorithm Based on YOLO-V4 for Three-Channels RGB SAR Image. *Remote Sens.* **2021**, *13*, 1909. [CrossRef]
17. Hong, Z.; Yang, T.; Tong, X.; Zhang, Y.; Jiang, S.; Zhou, R.; Han, Y.; Wang, J.; Yang, S.; Liu, S. Multi-Scale Ship Detection from SAR and Optical Imagery Via A More Accurate YOLOv3. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 6083–6101. [CrossRef]
18. Zhang, T.; Zhang, X. ShipDeNet-20: An Only 20 Convolution Layers and <1-MB Lightweight SAR Ship Detector. *IEEE Geosci. Remote Sens. Lett.* **2021**, *18*, 1234–1238. [CrossRef]
19. Jeong, S.; Kim, Y.; Kim, S.; Sohn, K. Enriching SAR Ship Detection via Multistage Domain Alignment. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 4018905. [CrossRef]
20. Jeon, H.K.; Yang, C.S. Enhancement of Ship Type Classification from a Combination of CNN and KNN. *Electronics* **2021**, *10*, 1169. [CrossRef]
21. Chen, S.; Wang, H.; Xu, F.; Jin, Y.Q. Target Classification Using the Deep Convolutional Networks for SAR Images. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 4806–4817. [CrossRef]
22. Raj, J.A.; Idicula, S.M.; Paul, B. One-Shot Learning-Based SAR Ship Classification Using New Hybrid Siamese Network. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 4017205. [CrossRef]
23. Karakuş, O.; Rizaev, I.; Achim, A. Ship Wake Detection in SAR Images via Sparse Regularization. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 1665–1677. [CrossRef]
24. Tings, B.; Pleskachevsky, A.; Jacobsen, S.; Velotto, D. Extension of Ship Wake Detectability Model for Non-Linear Influences of Parameters Using Satellite Based X-Band Synthetic Aperture Radar. *Remote Sens.* **2019**, *11*, 563. [CrossRef]
25. Yoshida, T.; Ouchi, K. Improved Accuracy of Velocity Estimation for Cruising Ships by Temporal Differences between Two Extreme Sublook Images of ALOS-2 Spotlight SAR Images with Long Integration Times. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 11622–11629. [CrossRef]
26. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014.

27. Jocher, G.; Stoken, A.; Borovec, J.; NanoCode012; ChristopherSTAN; Changyu, L.; Laughing; tkianai; Hogan, A.; lorenzomammana; et al. Ultralytics/Yolov5: V3.1—Bug Fixes and Performance Improvements. 2020. [CrossRef]
28. Redmon, J.; Farhadi, A. YOLO9000: Better, Faster, Stronger. *arXiv* **2016**, arXiv:1612.08242.
29. Redmon, J.; Farhadi, A. YOLOv3: An Incremental Improvement. *arXiv* **2018**, arXiv:1804.02767.
30. Bochkovskiy, A.; Wang, C.-Y.; Liao, H.-Y.M. YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv* **2020**, arXiv:2004.10934.
31. Wang, C.Y.; Mark Liao, H.Y.; Wu, Y.H.; Chen, P.Y.; Hsieh, J.W.; Yeh, I.H. CSPNet: A New Backbone That Can Enhance Learning Capability of CNN. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 14–19 June 2020. [CrossRef]
32. Wang, K.; Liew, J.H.; Zou, Y.; Zhou, D.; Feng, J. PANet: Few-Shot Image Semantic Segmentation with Prototype Alignment. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea, 27 October–2 November 2019; pp. 9196–9205. [CrossRef]
33. Ouchi, K.; Cordey, R.A. Statistical analysis of azimuth streaks observed in digitally processed CASSIE imagery of the sea surface. *IEEE Trans. Geosci. Remote Sens.* **1991**, *29*, 727–735. [CrossRef]