International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 03 Issue: 06 | June -2019 ISSN: 2590-1892

Ship detection using Faster R-CNN techniques

Harshitha M
Student
Department of EIE, Signal Processing
Bangalore Institute of Technology (BIT)
Bengaluru, India
harshithagowda994@gmail.com

Dr. H. N. Suresh
Professor and PG Coordinator
Department of EIE, Signal Processing
Bangalore Institute of Technology (BIT)
Bengaluru, India
hn.suresh@rediff.com

Abstract **Ship detection is of great value for fishing activity control, military defence, maritime transport, etc. Satellite-based synthetic aperture radar (SAR) can provide high-resolution images, allowing surveillance over massive water bodies to be possible. However, traditional ship detection algorithms like CFAR (Constant False-Alarm Rate), cannot produce convincing results. In recent years, detectors based on convolutional neural networks have made great progress, and among them, Faster R-CNN is one of the best in performance. In this paper, we propose an improved algorithm based on Faster R-CNN for ship target detection and adopt four strategies to improve the performance. The four strategies are replacing the VGG16 backbone network with ResNet101, implementing online hard example mining, replacing the default non-max suppression algorithm with soft-nms, and changing the aspect ratio of the anchors. Experimental results show that our improved algorithm can boost the performance of the traditional Faster R-CNN detector (72.7% mAP) by about 5% in mAP (mean average precision) on the HRSC2016 (High Resolution Ship Collection) dataset, showing the effectiveness of the proposed approach.**

Key words – Ship detection, faster R-CNN, SAR images.

.

I. Introduction

Humans can easily detect and identify objects present in an image. The human visual system is fast and accurate and can perform complex tasks like identifying multiple objects and detect obstacles with little conscious thought. With the availability of large amounts of data, faster GPUs, and better algorithms, we can now easily train computers to detect and classify multiple objects within an image with high accuracy.

Object detection technology has seen a rapid adoption rate in various and diverse industries. It helps self-driving cars safely navigate through traffic, spots violent behavior in a crowded place, assists sports teams analyze and build scouting reports, ensures proper quality control of parts in manufacturing, among many, many other things. And these are just scratching the surface of what object detection technology can do!

For an example, consider how much time have you spent looking for lost room keys in an untidy and messy house? It happens to the most of us and till date remains an incredibly frustrating experience. But what if a simple computer algorithm could locate your keys in a matter of milliseconds? That is the power of object detection algorithms. While this was a simple example, the applications of object detection span multiple and diverse industries, from round-the-clock surveillance to real-time vehicle detection in smart cities. In short, these are powerful deep learning algorithms. On the other hand, it takes a lot of time and training data for a machine to identify these objects. But with the recent advances in hardware and deep learning, this computer vision field has become a whole lot easier and more intuitive.

With the development of marine traffic, the number of ships on the high seas has increased greatly. The increased ships can improve seaborne trade. In the report of UNCTAD/RMT/2017, total volumes of seaborne trade reached 10.3 billion tons in 2016, Especially the strong import demand in China continued to support world maritime seaborne trade. However, in recent years, illegal activities such as smuggling, sea-jacking and maritime terrorism have seriously affected maritime trade. Therefore, maritime security is vital to global, regional and national economies.

In order to guarantee the safe navigation of the vessel and the safety of sea activities, the surveillance of ocean ships has become a very important issue for coastal countries. As a main application of maritime surveillance, ship detection has received more and more attention. At present, the most common and effective ship detection technologies are based on various sensors, such as radar and infrared sensor.

In practice, the ship detection system should have the capacity of all-time, all-weather and have a wide-area observation. As synthetic aperture radar (SAR) can provide high-resolution images of the observed ocean day and night and independent of weather condition, ship detection from SAR images is an effective technology and has become a hot research field.

Faster R-CNN for ship target detection adopt some strategies to improve the performance replacing the VGG16 backbone network with ResNet101 implementing online hard example mining replacing the default non-max suppression algorithm with Soft-NMS, the performance of the traditional Faster R-CNN detector (72.7% mAP) by about 5% in mAP.

.

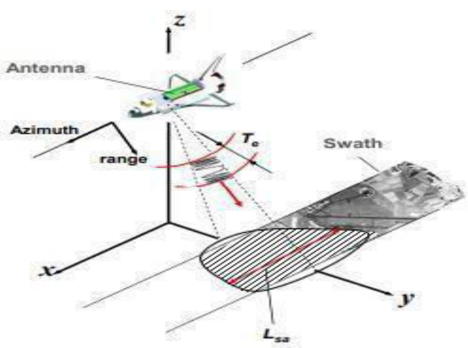


Fig 1. SAR imaging

Synthetic Aperture Radar(SAR) images can be obtained from satellites such as ERS, JERS and RADARSAT. Since radar interacts with the ground features in ways different from the optical radiation, special care has to be taken when interpreting radar images. It provides high-resolution, day-and-night and weather-independent images for a multitude of applications ranging from geoscience and climate change research, environmental and Earth system monitoring, 2-D and 3-D mapping, change detection, 4-D mapping (space and time), security-related applications up to planetary exploration.

In synthetic aperture radar (SAR) imaging, microwave pulses are transmitted by an antenna towards the earth surface. The microwave energy scattered back to the spacecraft is measured. The SAR makes use of the radar principle to form an image by utilizing the time delay of the backscattered signals.

In real aperture radar imaging, the ground resolution is limited by the size of the microwave beam sent out from the antenna. Finer details on the ground can be resolved by using a narrower beam. The beam width is inversely proportional to the size of the antenna, i.e. the longer the antenna, the narrower the beam. It is not feasible for a spacecraft to carry a very long antenna which is required for high resolution imaging of the earth surface. To overcome this limitation, SAR capitalises on the motion of the space craft to emulate a large antenna (about 4 km for the ERS SAR) from the small antenna (10 m on the ERS satellite) it carries on board.

The rest of the thesis follows as, second section explains literature survey, third section gives explanation of the methodology used, fourth section explains the experimental results and comparison of parameters and fifth section is all about the conclusion and future scope.

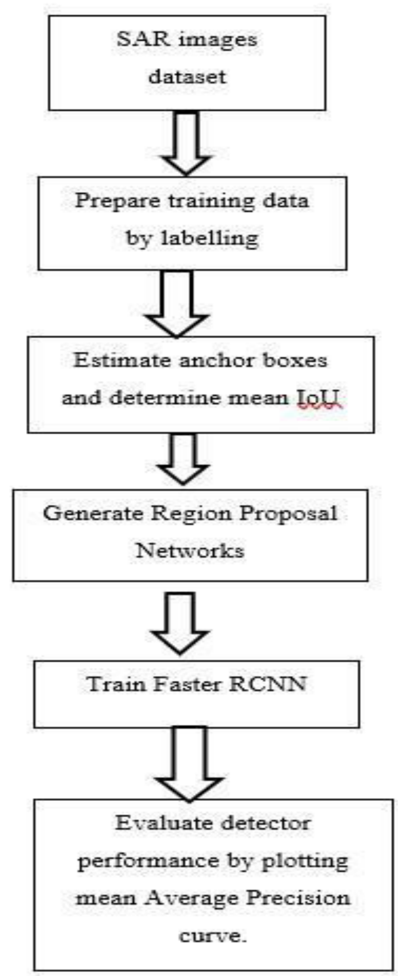
II. Methodology

A performance study by training Faster RCNN with the SAR images dataset. The following are the assumptions and dependencies taken into consideration while doing the comparative study of all algorithms MATLAB must be installed Deep learning toolbox Computer Vision toolbox Signal processing toolbox Image processing toolbox must be installed.

Traditional ship detection algorithms like CFAR (Constant False Alarm Rate), cannot produce convincing results. In recent years, detectors based on convolutional neural networks have made progress and among them, faster RCNN is one of the best in performance.

Advances in this field are mainly driven by the improvements in backbone networks and detection frameworks. Backbone networks include ZF-Net (Zeiler and Fergus Network), VGG16 (Visual Geometry Group), ResNet (Residual Network) etc. And common frameworks have developed from R-CNN to Fast R-CNN , Faster R-CNN,etc

The existing methodology uses VGG16 backbone network to improve the performance.



.

Fig 2. Block diagram

The data set of sample images is labeled using the This is done by Machine learning algorithm. This differentiates noisy pixels; easy application of de-noising technique on the image. step for de-noising of SAR image is as follows, Input the original noise free SAR image to the summation block. The second input is impulse noise of anticipated noise density. The output of summation is now the SAR image corrupted with impulse noise of certain noise density. SAR image corrupted by impulse noise is now fed to minmax filter for preprocessing of the image. Here around 20% to 30% of the noise is removed. This pre-processed image endures training and classification of Random Forest or Support Vector Machine. Now from the classification and training we are obtaining corrupted pixels, on which we apply SWT and perform de-noising. In final stage Averaging filter is applied on to the image to further reduce the amount of noise in the image.

1. Faster R-CNN

The Faster R-CNN detector. Instead of using an external algorithm like Edge Boxes, Faster R-CNN adds a region proposal network (RPN) to generate region proposals directly in the network. The RPN uses Anchor Boxes for Object Detection. Generating region proposals in the network is faster and better tuned to your data.

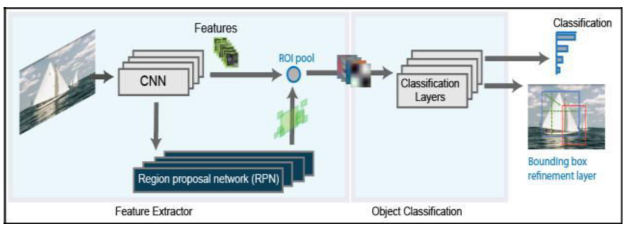


Fig 3. Faster R-CNN Detector

2. ResNet101

A backbone network refers to the convolutional layers It is used to extract features from an image. The generated feature maps are the input for both RPN and Fast R-CNN. Thus, the choice of backbone networks is essential for overall performance The default backbone used in Faster R-CNN is ZF-Net, a slightly improved version of AlexNet. In the openly published code of Faster R-CNN there are two other backbones that are optional, VGG_M and VGG16 However, years after Faster R-CNN was proposed, fair amounts of better backbones are made open. ResNet101 is one of the most outstanding representatives. ResNet101 is considered to be far more resistant to correlation decay between gradients allowing it to go much deeper than other backbones, which eventually helps the performance. Thus, the first improvement is to change the backbone to a ResNet101 network.

.

3. Soft-NMS

Non-maximum suppression (NMS) algorithm is crucial in Faster R-CNN. PRN often generates multiple region proposals around a single target. NMS is used whenever we need to select the best region proposal around the same object—a local optima. NMS picks the proposal with the highest confidence score and zeros out the confidence score of regions that have an IoU (Intersection over Union) greater than a certain threshold with the previously picked region. And the same procedure is performed on the rest untouched proposals until all is processed. However, NMS does not always work as it should.

NMS will not be aware when there are two objects staying close enough in one image. The situation can be described by Fig. 4. In this case, NMS is likely to eliminate all proposals around the other object and leave only one proposal for two objects, which causes missed detection. Instead of setting the score to zero, which makes the region never used since, we may assign a lower score to 'suppress' the region. That is the core idea of Soft-NMS We still stick to the intuition that a higher IoU means stronger suppression.

4. Aspect Ratios of Anchors

The original Faster R-CNN have anchors that are predefined with aspect ratios of 1:2, 1:1, 2:1. These ratios are suitable for PASCAL VOC dataset, which mainly consists of images taken in ordinary life. Ship targets is usually long and narrow. Nevertheless, this is not the case for bounding boxes. As ships can sail in any direction when the picture was captured, the aspect ratio for a rectangular bounding box can also be any value within a certain range.

Thus, adding more aspect ratios for anchors should give the performance a boost.

III. EXPERIMENTAL RESULTS

The dataset on the level 2 classification, which has five classes—submarines, aircraft carrier, merchant ship, warship and others. The use the mAP criterion defined in the PASCAL VOC challenge 2007 as evaluation metric The metric mAP is the mean value of the average precisions for the five given classes. Note that there is also a mAP metric which is the mean value of the average precisions for the first four classes, excluding class 'others'. This class is special because ships in these images cannot be identified accurately by human eyes. The objects in these images are mostly blurry and as a result, this class should contain various kinds of targets. Thus, bad performance on class 'Others' is acceptable. Together with mAP, we also take mAP metric into consideration when evaluating performance.

.

A. Performance improvement by changing the backbone network to Resnet101

The existing model used VGG16 as the backbone network to train faster RCNN. In our project, we have changed the backbone network to ResNet101 and could see the improvement in the performance of faster RCNN network.

Backbone Netwrok	mAP
VGG16	0.7763
ResNet101	0.7873

Table 1. Results of backbone network

B. *Average Precision Metric*

Computer Vision Toolbox™ provides object detector evaluation functions to measure common metrics such as average precision and log-average miss rates. Here, the average precision metric is used. The average precision provides a single number that incorporates the ability of the detector to make correct classifications (precision) and the ability of the detector to find all relevant objects (recall).

C. *Non-Max Suppression(NMS)*

Non-maximum suppression (NMS) algorithm is crucial in Faster R-CNN. PRN often generates multiple region proposals around a single target. NMS is used whenever we need to select the best region proposal around the same object—a local optima.

Fig. 4. Situation where NMS is applied

D. *Aspect Ratios of Anchors*

Some tests were conducted with different predefined anchor points and aspect ratios to increase the accuracy of the detector. Anchor boxes are important parameters of deep learning object detectors such as Faster R-CNN. The shape, scale, and number of anchor boxes impact the efficiency and accuracy of the detectors.

Load the training data. Data is prepared in image labeller by labelling the set of images used for training. After labelling data is imported as .mat file.Our dataset contains 200 labeled images and associated box labels.The labeled boxes are plotted as Box area vs. Aspect Ratio to better understand the range of object sizes present in the dataset.

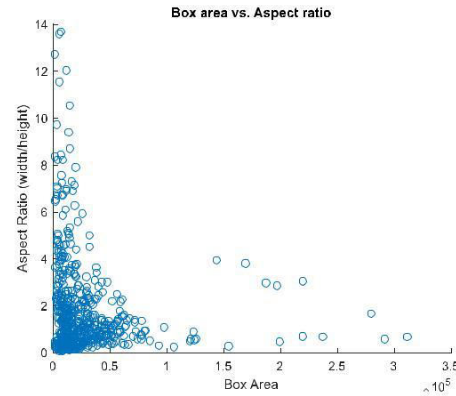


Figure 5. Plot of Box area vs. Aspect ratio

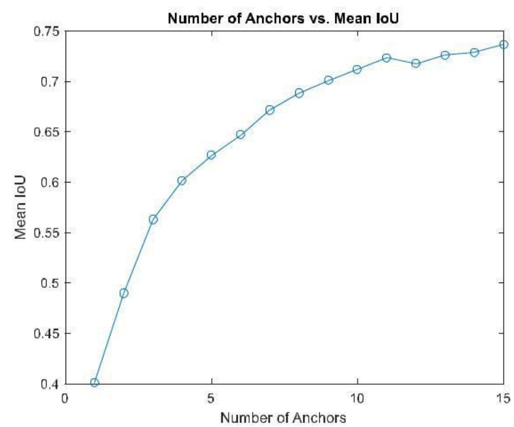


Figure 6. Plot of Number of anchors vs. Mean IoU

Figure 7. Tested image with rectangle labels around detected objects

E. *Measurement of average precision metric*

To measure the performance of the trained detector, the detector needs to be evaluated on the large set of images. For this evaluation, Computer Vision Toolbox provides object detector evaluation functions such as average precision and log-average miss rate Here we have used average precision metric.

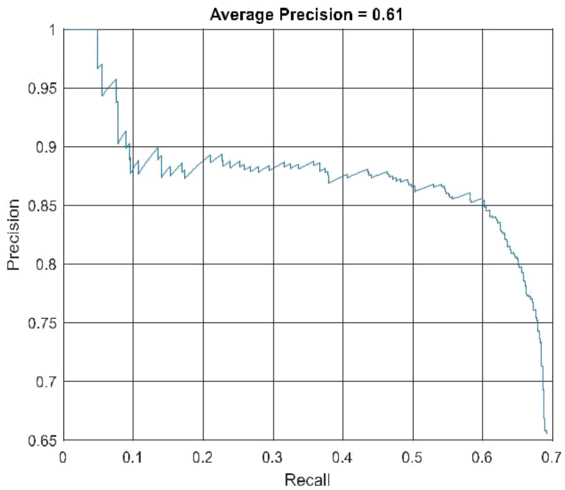


Figure 8.. Average Precision Metric

IV Conclusion and Future Scope

This project proposes a method for ship detection in SAR images by training the faster RCNN detector using ResNet101 as the backbone network. Using of pretrained network ResNet101 for detection improves the accuracy of the detector. By implementing NMS (Non-Max Suppression), probability of detection of ship in the presence of multiscale images has also been increased. The experiment has also been conducted by using different anchor boxes and the optimum value of anchor boxes has been selected and faster RCNN has been trained.

Many factors, such as incidence angle, ship size, wind speed, and metocean parameters influence the ship detectability. Typically, it is easy to detect ships in high incidence angle, low wind speed, or low sea state. Here the incidence angle, wind speed, and other ocean parameters mainly influence the contrast between the ocean and the ship, thus making the background of ships complex. Future work will be conducted regarding the evaluation of ship detectability on various specific conditions.

Ships with complex backgrounds are multiscale and are relatively small. Some ships in the dataset labeled by SAR experts may be incorrect, to some extent. Hence correction of dataset can be carried out to improve the accuracy of the detector.

Using the same dataset, the existing algorithm can be improved, or new ship detection algorithms can be developed which can improve accuracy of the detector. Also the existing detector can be improved in terms of detection by increasing accuracy and reducing the training time.

References

[1] H. Rohling, “Radar CFAR Thresholding in Clutter and Multiple Target Situations,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-19, no. 4, pp. 608-621, Jul. 1983.

[2] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Kauai, HI, USA, Dec. 2001, pp. I511-I518.

[3] C. Gong and J. Han, “A survey on object detection in optical remote sensing images,” *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 117, pp. 11-28, 2016.

[4] Y. LeCun, B. Boser, J. S. Denker, and D. Henderson, “Backpropagation Applied to Handwritten Zip Code Recognition,” *Neural Computation*, vol. 1, no. 4, pp. 541-551, 1989.

[5] K. Alex, I. Sutskever, and G. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in neural information processing systems*, Lake Tahoe, Nevada, USA, Dec. 2012, pp. 1097-1105..

[6] R. Girshick, “Fast R-CNN,” in *Proceedings of the 2015 IEEE International Conference on Computer Vision*, Santiago, Chile, Dec. 2015, pp. 1440-1448.

[7] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1137-1149, 2017.

[8] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask R-CNN,” in *Proceedings of the 2017 IEEE International Conference on Computer Vision* , Venice, Italy, Oct. 2017, pp. 2980-2988.

[9] Z. Liu, L. Yuan, L. Weng, and Y. Yang, “A High Resolution Optical Satellite Image Dataset for Ship Recognition and Some New Baselines,” in *Proceedings of the International Conference on Pattern Recognition Application and Methods*, Portugal, Porto, Aug. 2017, pp. 324-331.

[10] S. Ren, (2016, Aug.).faster_rcnn [Online]. Available: https://github.com/ShaoqingRen/faster_rcnn.

[11] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556*, 2014.

[12] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, Jun. 2016, pp. 770-778.

[13] D. Balduzzi, M. Frean, L. Leary, JP. Lewis, K. W-D. Ma, and B. McWilliams, “The Shattered Gradients Problem: If resnets are the answer, then what is the question?,” *arXiv preprint arXiv:1702.08591*, 2017.

[14] A. Shrivastava, A. Gupta, and R. Girshick, “Training Region-Based Object Detectors with Online Hard Example Mining,” in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, Jun. 2016, pp. 761-769.