

Detection of ocean internal waves based on Faster R-CNN in SAR images*

BAO Sude¹, MENG Junmin^{2, **}, SUN Lina², LIU Yongxin¹

¹ Inner Mongolia University, Hohhot 010021, China

² First Institute of Oceanography, Ministry of Natural Resources, Qingdao 266061, China

Received Feb. 19, 2019; accepted in principle Apr. 8, 2019; accepted for publication Apr. 26, 2019

© Chinese Society for Oceanology and Limnology, Science Press and Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract Ocean internal waves appear as irregular bright and dark stripes on synthetic aperture radar (SAR) remote sensing images. Ocean internal waves detection in SAR images consequently constituted a difficult and popular research topic. In this paper, ocean internal waves are detected in SAR images by employing the faster regions with convolutional neural network features (Faster R-CNN) framework; for this purpose, 888 internal wave samples are utilized to train the convolutional network and identify internal waves. The experimental results demonstrate a 94.78% recognition rate for internal waves, and the average detection speed is 0.22 s/image. In addition, the detection results of internal wave samples under different conditions are analyzed. This paper lays a foundation for detecting ocean internal waves using convolutional neural networks.

Keyword: ocean internal waves; faster regions with convolutional neural network features (Faster R-CNN); convolutional neural network; synthetic aperture radar (SAR) image; region proposal network (RPN)

1 INTRODUCTION

Internal waves are a type of kinetic phenomenon in the ocean. To form internal waves, the seawater density must be stably stratified, and disturbance energy must be present (Du et al., 2001). Field observations have shown that internal waves can reach a maximum amplitude of 240 m (Huang et al., 2016), which can seriously threaten the navigability of underwater submarines. Therefore, the study of internal waves is of great significance. Internal waves contain a substantial amount of energy and can, therefore, change the depth of the seawater transition layer, causing changes in the underwater sound field and affecting underwater communication. In addition, ocean internal waves can have far-reaching effects on research into marine sediments, marine fisheries, ocean acoustics, and other disciplines in the shelf sea area.

Remote sensing constitutes the primary means with which internal waves are monitored over a wide range. Ocean internal waves appear as bright and dark irregular stripes in remote sensing images;

unfortunately, these stripes are easily confused with other features in remote sensing images (vortices, ship wakes, wind, wave, etc.). Given the richness and extent of existing satellite remote sensing data, traditional manual interpretation methods are time-consuming and laborious, and thus, it is difficult to detect internal waves manually, which is accompanied by a large workload. Therefore, it is necessary to develop an automated technology to detect internal waves, thereby accelerating the processing of relevant data through the extraction and recognition of target features.

To detect internal waves, Hogan et al. (2002) used the Hough transform to extract the fringes of ocean internal waves from synthetic aperture radar (SAR) images. However, because the Hough transform is

* Supported by the National Natural Science Foundation of China (No. 61471136), the Special Project for Global Change and Air-sea Interaction of Ministry of Natural Resources (No. GASI-02-SCS-YGST2-04), and the Chinese Association of Ocean Mineral Resources R&D (No. DY135-E2-4)

** Corresponding author: mengjm@fio.org.cn

used to detect mainly straight lines, this approach is not ideal for identifying ocean internal waves in remote sensing images. As an alternative, the wavelet transform has multiresolution characteristics; with this technique, image noise (sea clutter) is concentrated at high frequencies, while the signals of internal waves are distributed at lower frequencies. Accordingly, Rödenas and Garello (1997) and Rodenas and Garello (1998) used the wavelet transform method to detect ocean internal waves in remote sensing images. Subsequently, Marghany (1999) and Kang et al. (2008) used two-dimensional empirical mode decomposition (2D-EMD) to decompose remote sensing images into different intrinsic mode functions (IMFs) to distinguish internal waves.

The convolutional neural network (CNN) is a type of artificial neural network. CNNs have been shown to exhibit excellent feature learning abilities, and the features learned by these networks can more accurately describe the nature of data, which is conducive to performing visualization or classification tasks. As a consequence, CNNs boast great advantages in terms of their target detection accuracy and speed. Girshick et al. (2014) initially proposed the regions with CNN features (R-CNN) framework structure, which first uses the selective search (SS) algorithm to extract the proposed region and then employs a CNN model to extract the target features from the proposed region and classifies those regional features with a classifier, such as a support vector machine (SVM); then, this algorithm uses bounding box regression on the classified proposed region, thereby increasing the accuracy of the bracketed target. However, the training of the R-CNN model requires multiple steps, and R-CNN models suffer from a slow detection speed. To ameliorate these problems, Girshick (2015) further proposed the Fast R-CNN framework, which performs only feature extraction for each picture using the proposed region mapped onto the feature map after the convolutional layer, after which the classification and regression tasks are completed. This greatly reduces the number of redundant calculations, which improves both the detection speed and detection performance. Subsequently, Ren et al. (2015) proposed the Faster R-CNN framework, with which the extraction of the region proposal network (RPN) is also completed by a CNN, but the extraction network and the target detection network share a feature extraction layer to improve the detection speed and achieve a better detection performance.

To date, the use of deep learning to detect ocean

internal waves has not been documented. In this paper, ocean internal waves in SAR images are detected within the Faster R-CNN framework, and a sample library of SAR images containing internal waves is constructed. Accordingly, CNNs are successfully employed to detect ocean internal waves in SAR images, and the results of detecting internal waves at different scales and stripes are analyzed.

2 ALGORITHM INTRODUCTION

The Faster R-CNN framework consists of two networks, an RPN and a Fast R-CNN. In the latter, an image segmentation method (such as the SS method) is not used to extract the candidate regions from the input image; rather, an RPN is added to the convolutional network. To share the convolutional layer between both networks, a pre-trained model is used to form a unified network through cross-training to complete the target detection task.

2.1 Fast R-CNN

The Fast R-CNN network structure is shown in Fig.1. An image of any size is obtained through the convolutional layer; then, after the feature is extracted by the convolutional layer, a region of interest (RoI) (He et al., 2015) pooling layer is added to the network the coordinates of the proposed region to the feature map. Because the fully connected layer requires a fixed-size input, the parameters 'H' and 'W' are set in the RoI pooling layer, and the feature map of each proposed region is fixed to a uniform scale. These feature maps obtain a fixed-size feature vector through the fully connected layer. Additionally, the Fast R-CNN network replaces the last fully connected layer with two output layers: a classification layer and a regression layer. The former outputs the categorical probability of each border, while the latter outputs the corresponding coordinates, uses non-maximum suppression (NMS) to remove overlapping bounding boxes, and finally outputs the bounding boxes with the highest score after applying a regression correction in each category.

2.2 RPN

An RPN is a fully convolutional network (FCN) constructed through backpropagation and stochastic gradient descent (SGD) end-to-end training. The network structure of an RPN, which is shown in Fig.2, uses three box areas (128^2 , 256^2 , 512^2) and three aspect ratios (1:1, 1:2, 2:1) for a total of nine sliding

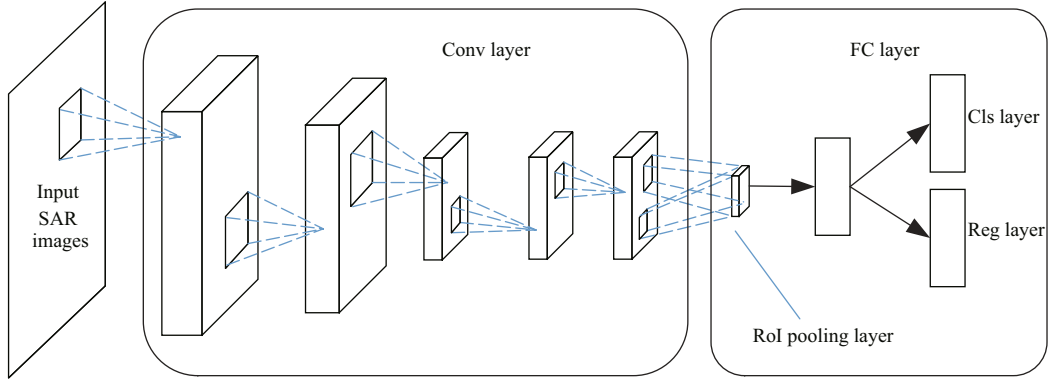


Fig.1 Fast R-CNN framework

Fast R-CNN consists of input images, convolutional layer, and fully connected layer. The input image is outputted to the feature map by convolution and pooling at the convolutional layer, then Faster R-CNN output the target score at “cls layer” and position at “reg layer”. “Conv layer” represents the convolutional layer, “FC” represents the fully connected layer.

windows centered on each point within the feature map. The RPN takes an image with an arbitrary size as the input, extracts the target region, and outputs a target score for each region. To generate the proposed region, a small network is slid over the convolutional feature map outputted by the last layer of the convolutional layer, and the features collected by different types of windows are reduced to a fixed dimension as the input of two congruence layers of the same level (the box regression layer and the box-classification layer).

To train the RPN, we assign a binary label (yes or no) to each anchor, and we assign positive labels to both types of anchors. One type of anchor exhibits an intersection-over-union (IoU) overlap with a ground-truth box region, while the other type demonstrates an IoU overlap with any ground-truth box region higher than 0.7. Additionally, the anchor that overlaps with the IoU of the ground-truth box area by less than 0.3 is assigned a negative label.

According to the above definition, the multitask loss function of the RPN is expressed as,

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{\text{cls}}} \sum_i L_{\text{cls}}(p_i, p_i^*) + \lambda \frac{1}{N_{\text{reg}}} \sum_i p_i^* L_{\text{reg}}(t_i, t_i^*),$$

where i is the index of a reference area in the mini-batch and P_i is the prediction probability that the reference area i is the target. If the reference area is positive, the ground-truth box tag P_i^* is 1; if the reference area is negative, P_i^* is 0; t_i is a vector containing the four parameter coordinates of the prediction region, and t_i^* is the coordinate vector of the ground-truth box region corresponding to the positive reference region. λ represents the balance factor used

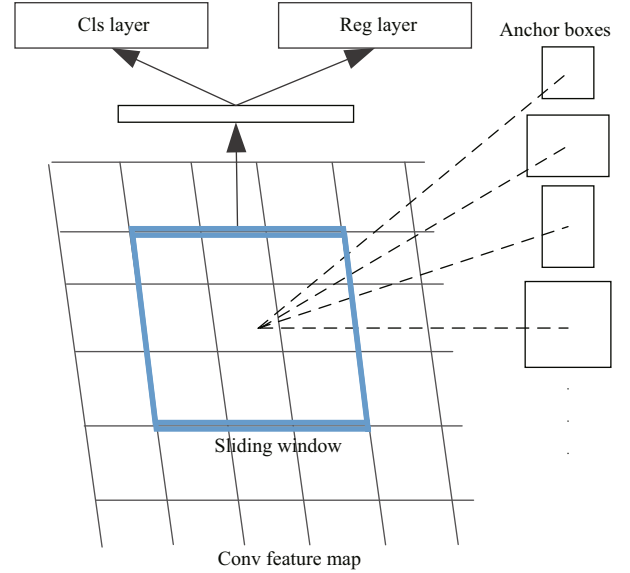


Fig.2 Region proposal network (RPN) framework

RPN use different scale sliding windows by anchor boxes to extract proposed region on the feature map, and output target score at classification layer (“cls layer”) and target position at region layer (“reg layer”). The proposed region generated by RPN is brought into the fully connected layer of Fast R-CNN.

to balance the weights of the two loss functions.

The classification loss L_{cls} is a logarithmic loss function of two categories (target/nontarget) determined by the following equation,

$$L_{\text{cls}}(p_i, p_i^*) = -\log[p_i p_i^* + (1-p_i^*)(1-p_i)].$$

For regression loss, the above formula is expressed as,

$$L_{\text{reg}}(t_i, t_i^*) = \sum_{i \in \{x, y, w, h\}} \text{smooth}_{L_1}(t_i - t_i^*),$$

in which,

$$\text{smooth}_{L_1}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise} \end{cases}.$$

The main purpose of the above formula is to be more robust both to outliers and to the magnitude of the control gradient during training. The term $(P_i^* L_{\text{reg}})$ indicates either that there is only a positive reference region ($P_i^*=1$) for regression loss or that there is no other case ($P_i^*=0$). The outputs of the classification layer and the regression layer are composed of $\{P_i\}$ and $\{t_i\}$, respectively, which are normalized by N_{cls} and N_{reg} and a balance coefficient λ . For regression, the following four coordinates are used:

$$\begin{aligned} t_x &= x - x_a / w_a, t_y = y - y_a / h_a, \\ t_w &= \log(w / w_a), t_h = \log(h / h_a), \\ t_x^* &= x^* - x_a / w_a, t_y^* = y^* - y_a / h_a, \\ t_w^* &= \log(w^* / w_a), t_h^* = \log(h^* / h_a), \end{aligned}$$

where x, y, w , and h indicate that the center coordinates (x, y) of the zone are wide and high. The variables x, x_a , and x^* refer to the coordinates of the prediction region, the reference region, and the ground-truth box region (likewise for y, w , and h).

The candidate frames extracted by the RPN have overlapping portions; these overlapping candidate blocks are removed by NMS. Finally, the N candidate frames before the score are output as the input of the RoI layer.

2.3 Sharing convolutional features for RPN and Fast R-CNN

Both RPN and Fast R-CNN are convolutional networks. The following four steps are used to implement two network shared convolutional layer to form a unified network. In the first step, the pre-trained model is used to train RPN to generate proposed region. In the second step, Fast R-CNN using the regions generated by the RPN and pre-trained model completes the detection task. In the third step, the shared conv layers are used to initialize RPN training, and only fine-tune the layers unique to RPN. Now the two networks share a convolutional layer. Finally, the fc layer of the Fast R-CNN is fine-tuned by keeping the shared convolutional layer fixed. As such, both networks share the same convolutional layer and form a unified network.

3 RESULT AND DISCUSSION

3.1 Data introduction

In this paper, we construct an internal wave database consisting of 466 Environmental Satellite (ENVISAT) advanced synthetic aperture radar

(ASAR) images from the South China Sea region acquired from 2003 to 2012. First, the remote sensing images are preprocessed, and the brightness is adjusted. Then, we extract each image containing the effective area of internal waves. Finally, 946 partial images with different internal wave morphologies and scales are extracted as samples for the database. The minimum image size in the database is approximately 240×240 , and the maximum image size is approximately 1400×1300 . Among the data samples, 58 partial images extracted from 42 remote sensing images are used as the test set evaluation network, and 888 partial images are input as training data to the convolutional network to learn the internal wave characteristics.

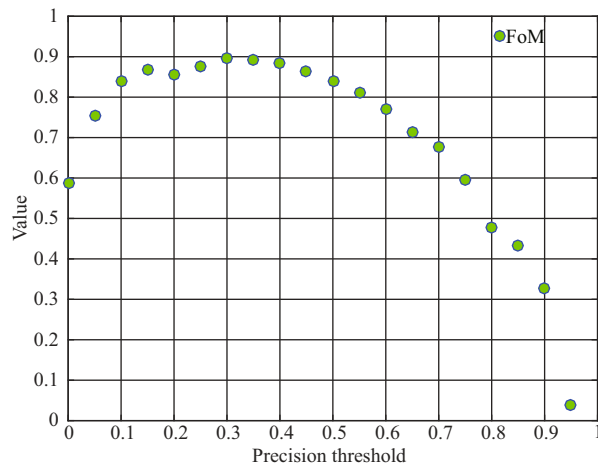
3.2 Threshold determination

The Faster R-CNN framework has a good feature learning ability. This paper aims mainly to detect ocean internal waves by debugging various network parameters, the selection of which affects the quality of the overall test results. Accordingly, it is necessary to debug these parameters several times to optimize the network and thus to optimize the detection results. In this paper, the Zeiler and Fergus net (ZFnet) (Zeiler and Fergus, 2014) is used to train the data. Through multiple experiments, the test set demonstrates the best detection effect when the ratio of training data to verification data is set to 0.5. In addition, the final network parameters determined in this paper are as follows: the momentum is 0.9, the weight attenuation is 0.0005, the learning rate is 0.0001; in addition, the learning rate must be divided by 10 after 20 k iterations, and 10 k iterations must be performed thereafter.

The Faster R-CNN outputs the detected target class and precision in the input image through the precision-recall (PR) curve, as shown in Fig.4b, in which the internal waves have a precision of 0.915. Additionally, the output shows a target border that is larger than the precision threshold. If the precision threshold is set too low, although more targets will be displayed, the target area in which the error is displayed will be included, and the number of false alarms will also increase. If the precision threshold is set too high, although the number of false alarms may be reduced, the number of detected targets may also decrease. For the trained network, the test results must not only ensure a high recognition rate but also have a low false alarm rate. Therefore, to meet the above conditions, the precision threshold interval was

Table 1 Changes in the FoM value with the test precision threshold of 0.30–0.35

Number of image	Number of target	Precision threshold	Number of target detected	Missing target number	Number of false alarm target	FoM
58	134	0.30	127	7	8	0.89
58	134	0.31	127	7	8	0.89
58	134	0.32	127	7	8	0.89
58	134	0.33	127	7	7	0.90
58	134	0.34	126	8	7	0.89
58	134	0.35	126	8	7	0.89

**Fig.3** Changes in the FoM curve with the test precision threshold

To evaluate the test results, the precision threshold interval was determined to be 0.05 after conducting 20 sets of experiments, then the FoM values are obtained from the curve to determine the ideal precision threshold interval.

determined to be 0.05 after conducting 20 sets of experiments, and the figure of merit (FoM) curve was drawn under different precision threshold conditions (Ai et al., 2009). The expression for calculating the FoM is $FoM = N_{it} / (N_{fa} + N_{gt})$, where N_{it} is the correct number of detection targets in the detection results, N_{fa} is the number of false alarm targets, and N_{gt} is the actual number of targets. According to Fig.3, the detection result was ideal when the precision threshold was between 0.30 and 0.35; furthermore, the number of false alarms was small while satisfying the number of targets needed to obtain the correct detection results.

Further comparison shows that the overall results were optimal with a precision of 0.33; the results are shown in Table 1. However, the internal wave shape was more complicated. In the test results, some internal waves were significantly different from the samples in the training database, resulting in a lower detection precision. However, among the results below the threshold, most of the detected targets

either were false alarms or bordered surrounding inaccurate targets, and the correctly detected targets composed a minority. In addition, by calculating the time from the input to the output of the image in the network, including the time spent extracting the candidate region, extracting the feature from the convolutional network, classifying the feature, and displaying the target region after regression, the average detection speed was 0.22 s/image with a GPU (this model was trained on an HP Z640, Intel Xeon CPU, NVIDIA Quadro K2200 GPU using the Caffe framework).

3.3 Method suitability

In this paper, internal waves of different shapes, including fringes at both large and small scales as well as wave packet groups, were analyzed. According to the marked region and precision, some types of internal waves were accurately detected.

3.3.1 Test results for different stripes

Figure 4b shows a SAR image from May 22, 2008 at 1401 UTC around the northern part of the South China Sea, and Fig.4c shows a SAR image from August 28, 2006 at 1407 UTC near the Natuna Islands. The internal waves of two figures propagate from east to west. In Fig.4b–c, “iw” denotes internal waves, and 0.758 represents the precision of the borders of the detection frame. Because internal solitary waves are small stripe and the scale is large, it is easy to distinguish the internal waves from the background, thereby increasing the detection precision. However, compared with a more defined target such as a ship, the internal waves do not have clear edges in this image, and the edges of the internal waves are easily confused with the background. Therefore, for the detection of internal waves, the deviation of the detection frame within a certain range is reasonable. Figure 4b has a higher precision result, while the area marked in Fig.4c is more accurate. In general, due to

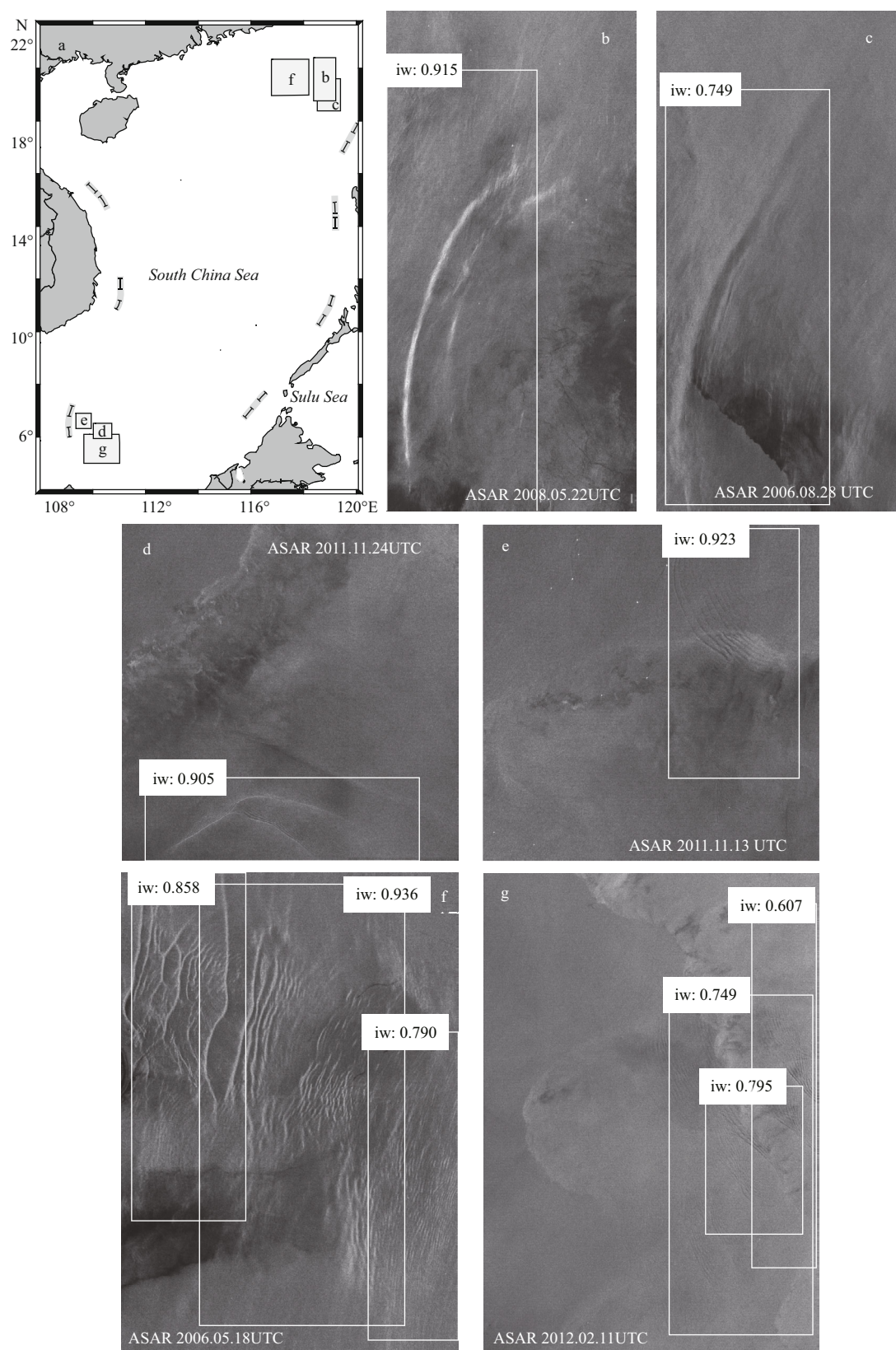


Fig.4 Images location and test results for different stripe

Remote sensing images are a quick look in this paper, so there is no latitude and longitude information. “iw” denotes internal waves; Numbers and white borders represent the precision and the borders for test results. a. images location of Fig.4b–g; b, c. test results of internal solitary waves; d, e. test results of wave packet; f, g. test results of wave packet group.

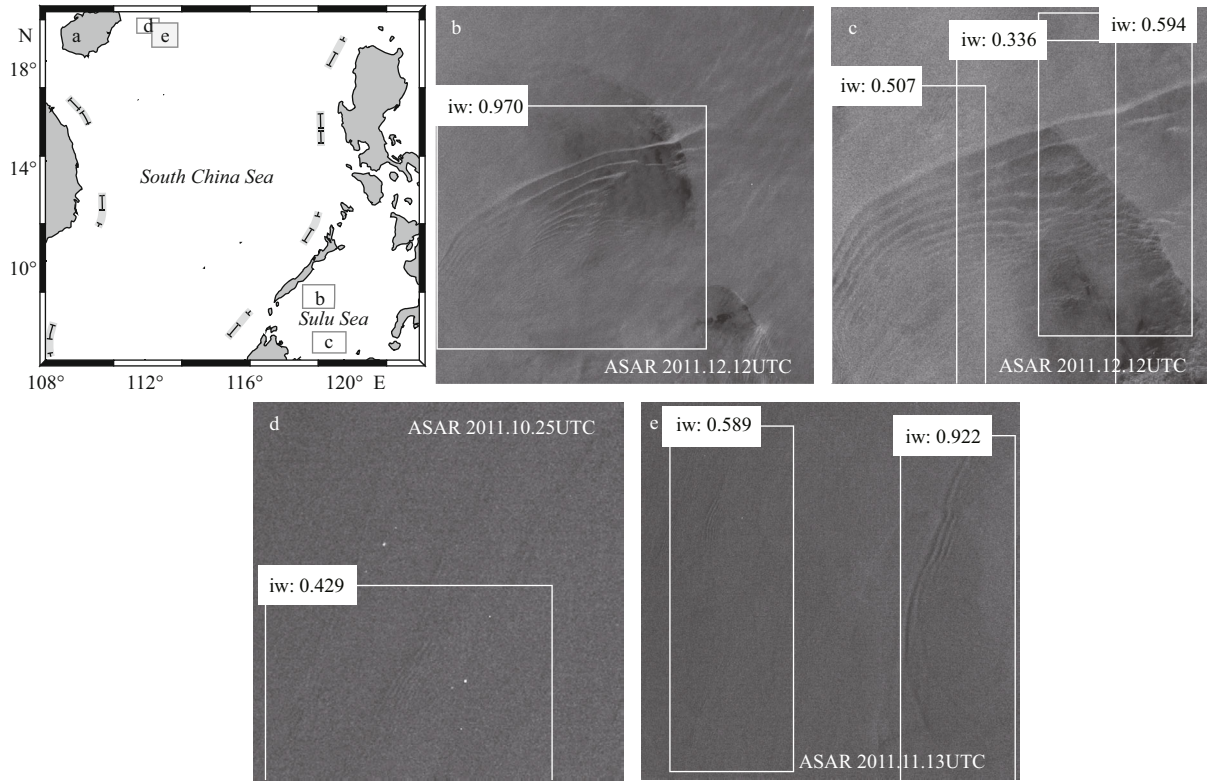


Fig.5 Images location and test results at different scale

a. images location of Fig.5b–e; b, c. test results of large-scale internal waves; d. test result of small-scale internal waves; e. test result of two scale internal waves.

the small stripes in Fig.4b–c, the internal wave shape is not complicated, and thus, the detection frame can more accurately surround the internal waves.

Figure 4d shows an ASAR image from November 24, 2011 at 1401 UTC around the northern part of the South China Sea, and Fig.4e shows an ASAR image from November 13, 2011 at 0240 UTC near the Natuna Islands. In Fig.4d–e, the internal waves form wave packet propagating eastward and westward, and the scale is smaller than that in Fig.4b–c. Since wave packets has many stripes, it is easy to distinguish the internal waves from the background.

Figure 4f shows an ASAR image from May 18, 2006 at 1407 UTC around Dongsha Island, and Fig.4g shows an ASAR image from February 11, 2012 at 0240 UTC near the Natuna Islands. The internal waves form wave packet group due to intersection and overlap in Fig.4f–g, and propagate in different directions. Due to the intersections of these wave packets, the wave packet groups cannot be distinguished according to the edges of each wave packet. As a consequence, when the database is created, several borders are created over the entire wave packet group. Therefore, overlapping borders appear in the detection results, but the detection precision is high.

3.3.2 Test results at different scales

Figure 5b–c show an ASAR image from December 12, 2011 at 1414 UTC over the Sulu Sea, and the internal waves propagate from the southeast to the northwest. The test result in Fig.5b has high precision and accurate area, but the internal waves of Fig.5c is marked by three regions. Some internal waves occur along the edge of the remote sensing image; thus, only part of the area can be observed, and only a fragment of the entire wave packet can be extracted when the database is created. Inevitably, in the database, such internal wave samples account for a notable proportion of all samples; if these data are removed, the reduction in the sample number would be detrimental to the ability of the network to learn how to detect internal waves. In addition, because fewer internal waves were observed in the Sulu Sea area during the study period of this paper, most of the samples contain fragments of the entire wave packet. Figure 5c shows a wave packet that is delineated by three borders because the network produced through the training data is closer to the inner region of the database than to a partial detection region relative to the entire region. We can effectively solve this problem by increasing the number of data samples.

Figure 5d shows an ASAR image from October 25, 2011 at 0234 UTC over the northern part of the South China Sea. The propagation direction is from southeast to northwest, the detection precision is 0.429. According to the measured spacing between the internal wave stripes, the minimum spacing in Fig.5d is 2 pixels, which is the smallest cell size that can be detected in this paper. The internal waves have no clear edges compared with more defined targets such as airplanes or ships; moreover, at such a small scale, it is easy to confuse these types of internal waves with the background in remote sensing images, and it is difficult to observe the occurrence area. Figure 5e shows that the small-scale internal waves appear to

have lower precision than larger-scale internal waves, and the detection frame is larger than the actual occurrence area. Therefore, in the detection results, these small-scale internal waves appear to have lower precision than the larger-scale internal waves detected above, and the detection frame is larger than the actual occurrence area. We should further research to solve the problems posed by this situation.

3.3.3 Test results for ship wakes and front

Figure 6a shows a Sentinel-1 image from January 15, 2016 at 1133 UTC near the Straits of Malacca. The internal waves propagate eastward. The bright spots on the left side of the image are ships, which produces a straight trail originating from the stern. The internal wave detection results can accurately surround the internal wave-generating region. In addition, the network does not consider the traces originating from ship wakes on the left side of the image to be internal waves, showing that the Faster R-CNN can not only accurately detect an internal wave but also distinguish the features in a remote sensing image that are easily confused with internal waves. Figure 6b shows a Sentinel-1 image from January 30, 2016 at 2149 UTC over the Celebes Sea, and the precision of boat front is 0.251. We set the precision threshold to 0.33, so front is not detected after setting the precision threshold.

3.3.4 False alarm goal

Figure 7b shows an ASAR image from May 15, 2006 at 1401 UTC around the Philippines, and Fig.7c shows an ASAR image from August 02, 2005 at 0238 UTC around the northeastern Taiwan, China. The test results show that our algorithm in this paper incorrectly detected the river estuary and land error as

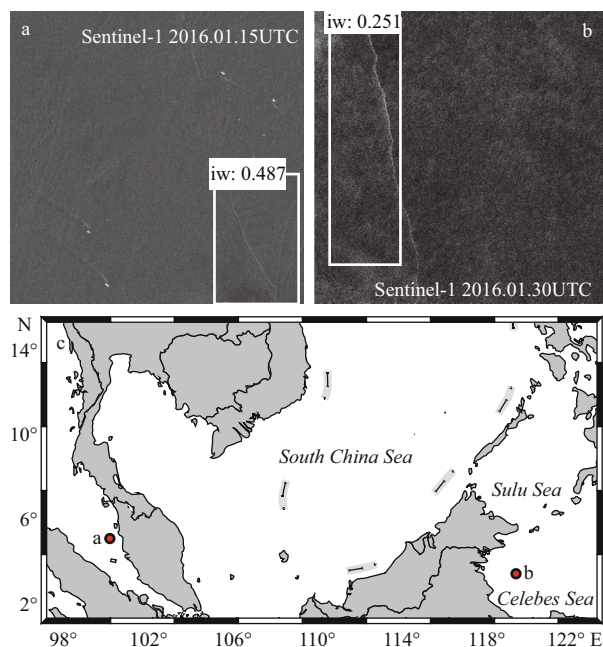


Fig.6 Test results for other features

a. test results of internal waves and ship wakes; b. test results of front; c. image location of Fig.6a–b.

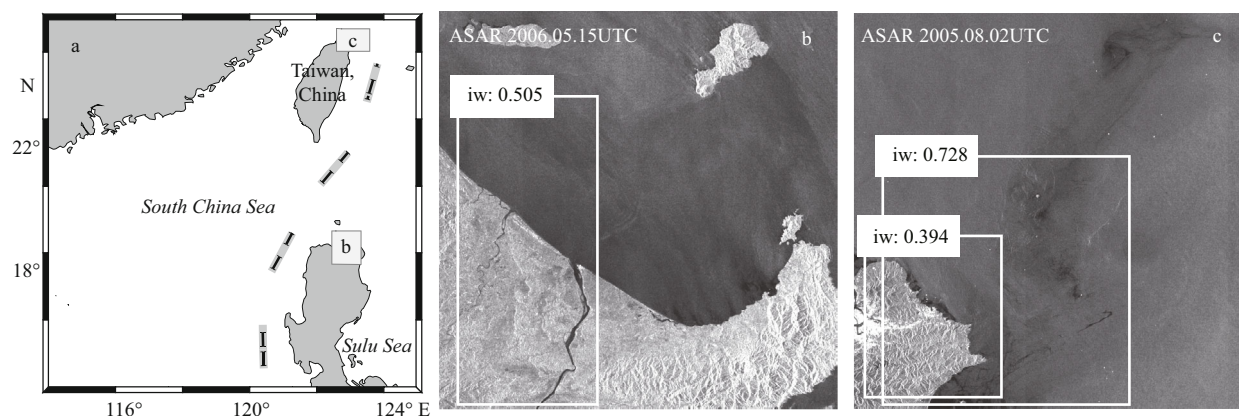


Fig.7 Test results of false alarm goals

a. image location of Fig.6a–b; b. test results of front; c. internal waves location of Fig.6a–b.

internal waves. For such errors, we should further research and resolution.

4 CONCLUSION

In this paper, ASAR remote sensing images are used to select internal wave samples at various scales, and these images are preprocessed to produce an internal wave database. Moreover, the detection of internal waves in SAR images within the Faster R-CNN framework is realized in combination with the constructed database. The experimental results show that the proposed algorithm can accurately frame the occurrence areas containing both internal wave stripes appearing at different scales and large-scale internal wave modes with high precision. At the same time, sufficient target detection can be achieved for small-scale internal waves. In some test results, the case in which a single large-scale wave packet is marked by two borders can be effectively resolved by expanding the number of data samples. In addition, in a remote sensing image, boat wakes can be easily confused with the shape of an internal wave; however, the Faster R-CNN can not only effectively avoid the identification of such aliasing features but also delineate internal wave-generating regions accurately. And we can also distinguish the front by the Faster R-CNN in this paper. However, some mis-detected targets have a morphology similar to those of internal waves, causing the network to incorrectly detect the internal waves. CNNs have broad application prospects in the detection of internal waves within SAR images; however, for target detection tasks, the Faster R-CNN framework achieves superior detection results relative to other methods.

5 DATA AVAILABILITY STATEMENT

The authors declare that the data supporting the findings of this study are available within the article.

References

- Ai J Q, Qi X Y, Yu W D. 2009. Improved two parameter CFAR ship detection algorithm in SAR images. *Journal of Electronics & Information Technology*, **31**(12): 2 881-2 885. (in Chinese with English abstract)
- Du T, Wu W, Fang X H. 2001. The generation and distribution of ocean internal waves. *Marine Sciences*, **25**(4): 25-28. (in Chinese)
- Girshick R, Donahue J, Darrell T, Malik J. 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. *In: Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, Columbus, OH, USA.
- Girshick R. 2015. Fast R-CNN. *In: Proceedings of 2015 IEEE International Conference on Computer Vision*. IEEE, Santiago, Chile.
- He K M, Zhang X Y, Ren S Q, Sun J. 2015. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **37**(9): 1 904-1 916.
- Hogan G G, Marsden J B, Henry J C. 2002. On the detection of internal waves in high resolution SAR imagery using the Hough transform. *In: Proceedings of the OCEANS 91 Proceedings*. IEEE, Honolulu, Hawaii, USA.
- Huang X D, Chen Z H, Zhao W, Zhang Z W, Zhou C, Yang Q X, Tian J W. 2016. An extreme internal solitary wave event observed in the northern South China Sea. *Scientific Reports*, **6**: 30041.
- Kang J, Zhang J, Song P J, Meng J M. 2008. The application of two-dimensional EMD to extracting internal waves in SAR images. *In: Proceedings of 2008 International Conference on Computer Science and Software Engineering*. IEEE, Hubei, China. p.953-956.
- Marghany M. 1999. Internal wave detection and wavelength estimation. *In: Proceedings of 1999*. IEEE, Hamburg, Germany.
- Ren S Q, He K M, Girshick R, Sun J. 2015. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**(6): 1 137-1 149.
- Rödenas J A, Garelo R. 1997. Wavelet analysis in SAR ocean image profiles for internal wave detection and wavelength estimation. *IEEE Transactions on Geoscience and Remote Sensing*, **35**(4): 933-945.
- Rödenas J A, Garelo R. 1998. Internal wave detection and location in SAR images using wavelet transform. *IEEE Transactions on Geoscience and Remote Sensing*, **36**(5): 1 494-1 507.
- Zeiler M D, Fergus R. 2014. *In: Proceedings of the 13th European Conference on Computer Vision*. Springer, Zurich, Switzerland. p.818-833.