TKP-Net: A Three Keypoint Detection Network for Ships Using SAR Imagery

Xiunan Li , Peng Chen , Jingsong Yang , Wentao An , *Member, IEEE*, Gang Zheng , *Senior Member, IEEE*, Dan Luo , Aiying Lu , and Zimu Wang

Abstract—Remote-sensing ship monitoring is a crucial area of research with key applications in military and civilian fields. The ability to extract information, such as ship length, width, and heading from remote-sensing data, particularly from synthetic aperture radar (SAR) images, is of paramount importance. Current state-of-the-art SAR image ship monitoring focuses primarily on ship detection. Assessing the direction of ships usually relies on the observability of wake features. However, the observability of these wake features is often affected by factors, such as the SAR system parameters, ship attributes, and dynamic marine environments. This can make accurate direction assessments a challenging task. In response to these challenges, this study has presented a novel and effective algorithm for ship monitoring from SAR images based on an anchor-free framework and the powerful feature extraction capabilities of convolutional neural networks. The proposed method learned the scattering and morphological information of a ship's bow and stern from high-resolution SAR images to determine the ship's direction with a high level of accuracy using a rotating bounding box. The algorithm was tested on a dataset, achieving an average precision of 90.8% and bow classification accuracy of 92.5%, demonstrating its potential contributing to the advancement of remote sensing.

Index Terms—Convolutional neural network (CNN), keypoint, ship detection, ship head classification, synthetic aperture radar (SAR).

Xiunan Li is with the Ocean College, Zhejiang University, Zhoushan 316021, China, and also with the State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, Ministry of Natural Resources, Hangzhou 310012, China (e-mail: lixiunan@zju.edu.cn).

Peng Chen, Gang Zheng, Dan Luo, Aiying Lu, and Zimu Wang are with the State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, Ministry of Natural Resources, Hangzhou 310012, China (e-mail: chenpeng@sio.org.cn; zhenggang@sio.org.cn; luod@zju.edu.cn; luaiying@stu.ouc.edu.cn; zimuwang1878567@163.com).

Jingsong Yang is with the Ocean College, Zhejiang University, Zhoushan 316021, China, also with the State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, Ministry of Natural Resources, Hangzhou 310012, China, also with the School of Oceanography, Shanghai Jiao Tong University, Shanghai 200030, China, and also with Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai 519082, China (e-mail: jsyang@sio.org.cn).

Wentao An is with the National Satellite Ocean Application Service, Key Laboratory of Space Ocean Remote Sensing and Application, Ministry of Natural Resources, Beijing 100081, China (e-mail: anandyandrew@126.com).

Digital Object Identifier 10.1109/JSTARS.2023.3329252

I. INTRODUCTION

MONITORING ships is of considerable importance in military and civilian fields [1], including maintaining national marine security, supervising water traffic, maritime fishery management, and maritime rescue [2]. Optical remote-sensing images have clear details and textures, which have the key advantages for identifying ship types. However, this process is easily impeded by weather conditions, including clouds, rain, and fog, and can only be conducted during daylight hours. In contrast, synthetic aperture radar (SAR) can be used at night and has fewer limitations imposed by weather conditions. Therefore, it is widely used in ship detection [3]. With the development of high-resolution SAR satellites, finer scale ship information [4], such as the length, width [5], and direction, can be extracted from high-resolution SAR images.

Traditional SAR image detection algorithms first need to perform sea and land segmentation to remove land interference, and then perform ship detection. The ship detection results are often affected by the accuracy of land and sea segmentation [6]. The classic algorithm used in this context is the constant false alarm rate (CFAR) algorithm, such as the cell average CFAR detector [7], two-parameter CFAR detector [8], and order statistics CFAR detector [9]. The CFAR-based detection algorithm adaptively selects an appropriate detection threshold for ship detection based on the background clutter statistical distribution model (such as the K distribution [10] and gamma distribution [11]) [12]. This method works well when the sea surface clutter estimation model is consistent with the detection image distribution. However, this algorithm type requires manual analysis of the target and background. The parameter design is complex, lacks versatility, and makes it difficult to deal with complex sea conditions and ship motion blur, sidelobe effects, and interference generated by near-shore ships [12]. It is also difficult to extract the accurate length, width, and bow information using the CFAR method.

Recently, target detection algorithms based on deep convolutional neural networks (DCNN) have made considerable progress, and their powerful feature extraction and learning capabilities have strongly surpassed those of the traditional methods in terms of accuracy and robustness, and have received considerable attention from researchers. Depending on whether an anchor exists, this method can be anchor-based or anchor-free. Anchor-based methods predominantly include R-CNN [13], fast R-CNN [14], faster R-CNN [15], R-FCN [16], SSD [17], YOLOv2 [18], YOLOv3 [19], and YOLOv4 [20].

The anchor-based method needs to manually set many anchor boxes and introduce an excess number of hyperparameters, such as the scale, ratio, and quantity of anchors. A relatively small number of anchor boxes can match the target, causing an imbalance between the target and background that affects the convergence speed. The anchor-free method does not need to set the anchor box parameters. It is simple to use and its accuracy is comparable with that of the anchor-based algorithm and is now being used by an increasing number of researchers. CornerNet [21], ExtremeNet [22], and CenterNet [23] are representative anchor-free methods based on keypoints. CornerNet uses two keypoint in the upper left and lower right corners to locate the upper left and lower right corners of the object bounding box. It then uses the embedding vector to determine whether the two keypoints belong to the same object. ExtremeNet detects objects by predicting extreme and center points and grouping keypoints according to the geometric structure. CenterNet only uses the central keypoint to locate the target. The width and height parameters of the target are obtained through regression, which substantially improves the detection speed.

Based on a DCNN, the current SAR ship detection algorithm omits the sea–land segmentation strategy of the traditional algorithm and uses an end-to-end method to send SAR images to the DCNN for detection. It then directly outputs ship detection results, which improves the efficiency and accuracy of detection [24]. Based on CenterNet, Guo et al. [25] proposed CenterNet++, which introduced a feature refinement module to extract contextual information for detecting small ships. Considering the multiscale and multiangle characteristics of SAR image scenes and ship targets, Hu et al. [6] proposed a BANet-based anchor-free method for multiscale ship detection in SAR images. By introducing a deformable convolution, local ship features can be extracted effectively. Deng et al. [26] proposed a novel and effective method for learning deep ship detection from scratch, which achieved a high level of detection performance using Sentinel-1 data.

The aforementioned algorithms all employ horizontal bounding boxes (HBBs) for object detection. However, in the case of remote-sensing images captured from an aerial perspective, objects in the image can have arbitrary orientations, and the use of HBBs can result in poor detection performance for densely arranged, arbitrarily oriented, and variably scaled objects. This is partly due to the inclusion of a significant amount of irrelevant background information within the HBBs, which results in imprecise object representations, as well as the nonmaximum suppression (NMS) algorithm's tendency to remove highly overlapping ships, resulting in missed detections.

Recently, detection methods based on an oriented bounding box (OBB) have been proposed to solve these problems. Xia et al. [27] used a method of returning four vertex coordinates to generate an OBB based on a faster R-CNN algorithm. Gliding vertex [28] works on the assumption that the order of coordinates generated by this method of returning four vertices is easy to confuse. Therefore, they choose to return four ratios, which represent the relative sliding offsets on each side of the HBB. Yang and Yan [29] represented the OBB as a center point, length, width, and angle, and the angle was predicted by classification
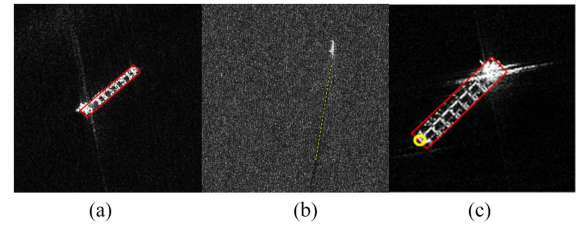

Fig. 1. Ship direction detection method. (a) Rotating box detection. (b) Vessel-wake detection. (c) Rotating box detection + Head point extraction.

rather than regression. He et al. [30] proposed a keypoint-based anchor-free algorithm to generate SAR ship OBBs by learning polar coordinate encoding. Although the OBB algorithm obtained a compact bounding box and had certain direction information, it had 180° ambiguity and could not be used to obtain the precise direction of each ship [see Fig. 1(a)].

Estimating the direction of motion of a ship from SAR images is usually performed indirectly using the ship-wake method [see Fig. 1(b)]. Ship wakes typically occur in many forms. The most common are turbulent wakes, Kelvin wakes, and narrow V-wakes [31]. Turbulent wakes occurring as dark lines along the longitudinal axis of a ship are the most commonly observed wakes in SAR images [32]. A narrow V-tail typically appears as a bright edge next to a darker trail [33]. The Kelvin wake is composed of shear, divergent, and sharp waves, and its V-shaped angle is approximately 39° [34], which is usually smaller than the angle in SAR images. Ship wakes are predominantly shown as linear structures in the image [35]. The most common method is to use the Radon or Hough transform to detect lines [36]. The lines and ships detected are combined to obtain the ship's navigation direction [37]. However, the visibility of wake information in SAR images is affected by radar parameters, ship parameters, and background sea conditions [38] and is difficult to observe in many cases. We proposed a keypoint-based OBB detection algorithm that could automatically extract high-resolution SAR position and bow information using DCNN and keypoints without wakes [see Fig. 1(c)]. The proposed dataset reflected a strong performance. The contributions of this study were given as follows.

1) We proposed an end-to-end ship detection method based on three keypoints for arbitrary direction SAR. The rotation angle-prediction problem was transformed into an estimation and matching problem for keypoints, which avoided the angle periodicity problem based on the regression method.

2) Aiming to address the problem of 180° ambiguity in the current OBB detection method for determining ship direction, we automatically obtained the shape and scattering information of the SAR bow and stern through two keypoints. We then assessed whether the two keypoints were the bow and obtained the accurate bow classification.

3) We exploited the difference in information extracted from two keypoints to construct a simple and practical analytical model to evaluate the bow classification results.
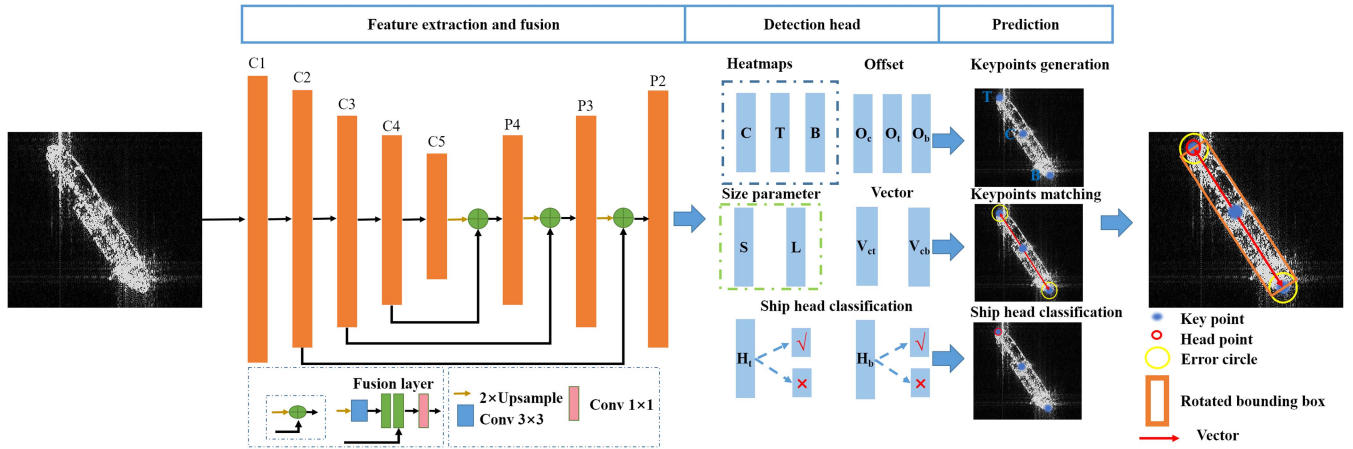
Fig. 2. Overall architecture of our method. The network structure can be divided into three parts, that is, feature extraction and fusion module, detection head module, and prediction part. After the data had passed through the feature extraction and fusion module, the feature map, offset, size parameter, vector, and ship head layers had five parameter parts. The feature map and the offset were combined to generate keypoint coordinates. The size parameters and vectors were used to match the keypoint coordinates to generate OBB, and the ship head layers were used to determine which keypoint the bow was on.

The rest of this article is organized as follows. In Section II, the proposed method is described in detail. Section III presents the experimental datasets and comparisons. Section IV discusses the limitations of our method. Finally, Section V concludes this article.

II. PROPOSED METHOD

Our proposed method is shown in Fig. 2 and has been divided into three main parts, that is, feature extraction and fusion, target detection head, and prediction module. The feature extraction and fusion module was used to extract the features of multiscale ships. The target detection head module was used to generate keypoint heatmaps, keypoint offsets, two side lengths of OBB boxes, direction vectors, and bow classification results. The prediction module combines the information from the detection head module to form the OBB and the head point position.

A. Feature Extraction and Fusion Module

The feature extraction and fusion modules comprised two parts. The first part used the classic ResNet50 [39] structure for feature extraction. The second part used feature maps from different scales of the ResNet residual module for layer-by-layer fusion. We employ skip connections to fuse multiscale features at different layers, aiming to capture both fine-grained details and high-level contextual information. Unlike feature pyramid network [40], our method does not adopt the multibranch prediction head because the multibranch prediction head is more computationally intensive. The resolution of the feature map of the last layer of the ResNet was 32 times lower than that of the original image. For small-scale objects, the loss of information after repeated downsampling is relatively high, and small objects are easily missed. The semantic information was combined with shallow, high-resolution morphological information. As shown in Fig. 3, C4 was fused with C5 to obtain P4; P4 and C3 were fused to obtain P3; and P3 and C2 were fused to obtain P2. Taking the fusion of C4 and C5 as an example, assuming that
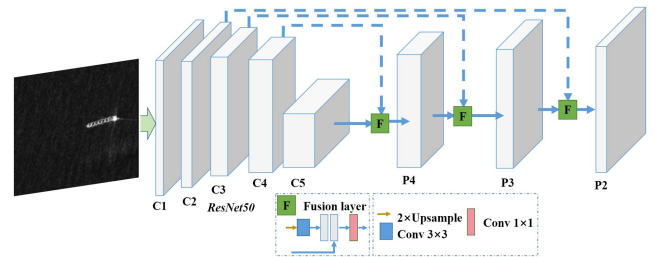

Fig. 3. Feature extraction and fusion module.

the input image size is $512 \times 512 \times 3$, the size of the C5 feature map is $16 \times 16 \times 2048$, and the size of the feature map obtained by double upsampling through bilinear interpolation was $32 \times 32 \times 2048$. Then, 3×3 convolution became $32 \times 32 \times 1024$, spliced with C4 to $32 \times 32 \times 2048$, and finally, 1×1 convolution becomes $32 \times 32 \times 1024$. P4, C3, P3, and C2 exhibited similar fusion processes. The detailed structural parameters of the feature extraction and fusion modules are listed in Table I.

B. Detection Head

1) Keypoints Estimation: Heatmaps are often used to locate keypoints in human bones. We use the same keypoint extraction method as CenterNet and CornerNet. In order to avoid the problem of close proximity of multiple keypoints, we set the parameters of the Gaussian function according to the size of each ship, and the maximum Gaussian kernel scale set does not exceed the width of the ship. Here, we use the center Keypoint C, upper Keypoint T, and lower Keypoint B to locate the key positions of the ship. The keypoints were divided, as shown in Fig. 4. The coordinate axis was established with the center Keypoint C as the origin. When the coordinates of keypoints B and T are on the X-axis, the left side of the Y-axis is Keypoint B, and the right side is Keypoint T. When Keypoints B and T are not on the X-axis,

TABLE I
FEATURE EXTRACTION AND FUSION MODULE STRUCTURE

Layer name	layer	output
C1	7×764, stride 2	256×256
	3×3, max pool, stride 2	
C2	$\begin{bmatrix} 1 \times 1,64 \\ 3 \times 3,64 \\ 1 \times 1,256 \end{bmatrix} \times 3$	128×128
C3	$\begin{bmatrix} 1 \times 1,128 \\ 3 \times 3,128 \\ 1 \times 1,512 \end{bmatrix} \times 4$	64×64
C4	$\begin{bmatrix} 1 \times 1,256 \\ 3 \times 3,256 \\ 1 \times 1,1024 \end{bmatrix} \times 6$	32×32
C5	$\begin{bmatrix} 1 \times 1,512 \\ 3 \times 3,512 \\ 1 \times 1,2048 \end{bmatrix} \times 3$	16×16
P4	$\begin{bmatrix} 3 \times 3,1024 \\ 1 \times 1,1024 \end{bmatrix} \times 1$	32×32
P3	$\begin{bmatrix} 3 \times 3,512 \\ 1 \times 1,512 \end{bmatrix} \times 1$	64×64
P2	$\begin{bmatrix} 3 \times 3,256. \\ 1 \times 1,256 \end{bmatrix} \times 1$	128×128

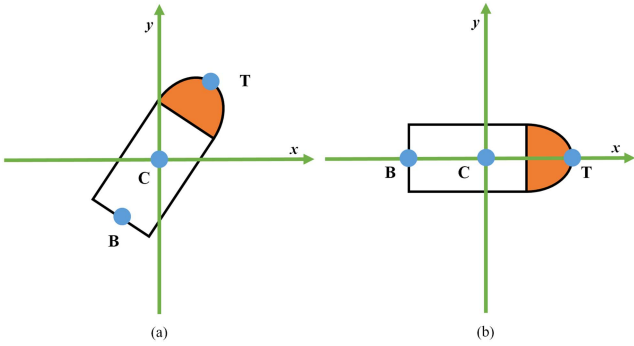Fig. 4. Division of keypoints. (a) When the coordinates of keypoints B and T are not on the X-axis, the keypoint above the X-axis is T, and the keypoint below the X-axis is B. (b) When the coordinates of points B and T are on the X-axis, the left side of the Y-axis is keypoint B, and the right side is the keypoint T.

the keypoints above the X-axis are T and the keypoints below the X-axis are B. The heat map of the keypoint labels is generated using a two-dimensional (2-D) Gaussian function $e^{-\frac{c_x^2 + c_y^2}{2\sigma^2}}$. The loss function uses a Gaussian focal loss [41], as shown in the following equation:

$$L_{\text{hm}} = \frac{-1}{N} \sum_{h=1}^{H} \sum_{w=1}^{W}$$
$$\times \begin{cases} (1 - \text{hm}_{hw})^{\alpha} \log(\text{hm}_{hw}) & \text{if } \widehat{\text{hm}}_{hw} = 1 \\ \left(1 - \widehat{\text{hm}}_{hw}\right)^{\beta} (\text{hm}_{hw})^{\alpha} \log(1 - \text{hm}_{hw}) & \text{otherwise} \end{cases}$$
$$(1)$$

where N is the number of objects in the image, and H and W are the height and width, respectively. After downsampling the image four times, α and β are the hyperparameters ($\alpha = 2$ and
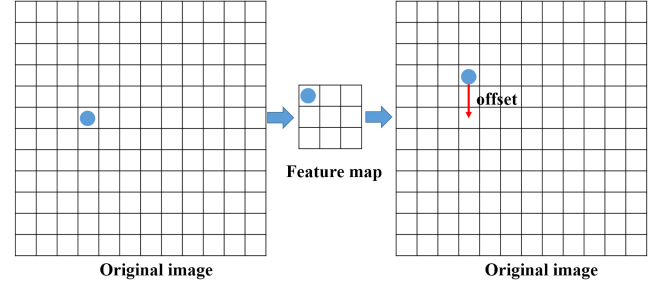

Fig. 5. Offset of the feature map to the original image.

$\beta = 4$), hm_{hw} is the predicted heatmap value, and $\widehat{\text{hm}}_{hw}$ is the heatmap value of the real label. The coordinate position of the original image to the coordinate position of the heatmap is d times smaller than that of the original image and is rounded down. Therefore, there is a deviation between the coordinates of the keypoints extracted from the heatmap and those of the original image. An offset parameter was used to quantify this error, as shown in Fig. 5. The deviation at any keypoint is given as follows:

$$o = \left(\frac{x}{d} - \left\lfloor \frac{x}{d} \right\rfloor, \frac{y}{d} - \left\lfloor \frac{y}{d} \right\rfloor \right) \qquad (2)$$

where x and y are the real coordinates of the keypoint, and d is the scaling factor.

The loss function uses the $L1$ smoothing loss function as follows:

$$L_{\text{off}} = \frac{1}{N} \sum_{k=1}^{N} \text{Smooth}L1\text{Loss}(o_k, \hat{o}_k) \qquad (3)$$

where o_k is the predicted bias and $\hat{o}_k$ is the true bias.

2) Scale Parameters Regression: The scale parameters of the rotating box are represented by the long side l and the short side s. The loss function uses a smooth $L1$ loss function, as shown in the following equations:

$$L_l = \frac{1}{N} \sum_{k=1}^{N} \text{Smooth}L1\text{Loss}\left(l_k, \widehat{l_k}\right) \qquad (4)$$

$$L_s = \frac{1}{N} \sum_{k=1}^{N} \text{Smooth}L1\text{Loss}(s_k, \widehat{s_k}). \qquad (5)$$

3) Keypoints Matching: When two or more keypoints are used to locate the target, a keypoint matching problem occurs. The proposed method uses a vector from the center point to the endpoint, and the distance from the endpoint coordinates meets a certain threshold to confirm whether they belong to the same target, especially for the regression vector numerical stability. We mapped the prediction vector $\overrightarrow{v_k}$ between 0 and 1. The loss function of the vectors is given as follows:

$$L_v = \frac{1}{N} \sum_{k=1}^{N} \text{Smooth}L1\text{Loss}\left((2\overrightarrow{v_k} - 1) \frac{l_k}{2}, \overrightarrow{\hat{v}_k} \right). \qquad (6)$$

4) Bow Classification: To facilitate the extraction of information related to the bow and stern of the ship, we use two
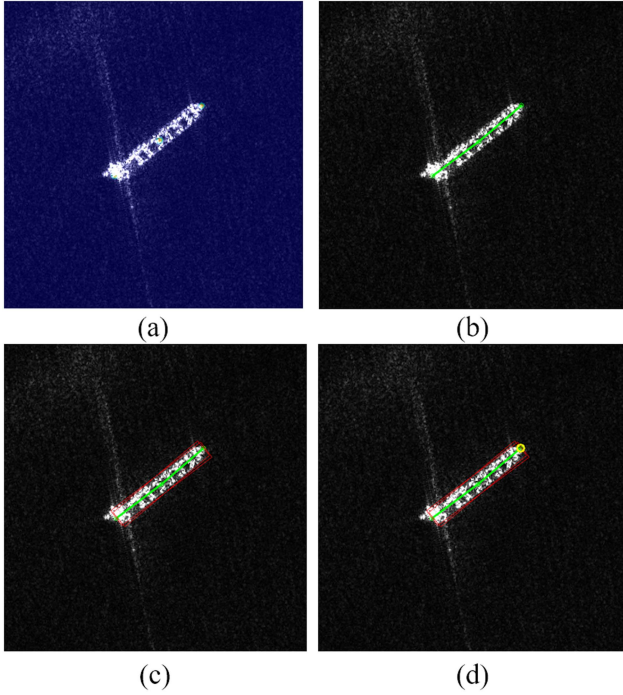
Fig. 6.   Prediction in the test stage. (a) Keypoint heatmap. (b) Keypoint matching. (c) Rotating box generation. (When two boxes are generated, then a low score can be removed using NMS). (d) Bow classification.

independent classification branch modules in the architecture. These classification modules are designed to extract information from positions corresponding to the upper and lower keypoints. The purpose of introducing these two independent classification modules is to capture features specifically associated with the ship's bow and stern regions. The upper and lower keypoints serve as reference points, allowing the classification modules to focus on extracting relevant information in proximity to these keypoints. There are two keypoints at the bow and stern of the ship. We need to use the BCE loss function at these two endpoints to determine the probability that the point is the bow. The L_{hc} loss function was designed as follows:

$$L_{hc} = -\frac{1}{N} \sum_{k=1}^{N} \left(\widehat{hc_k} \log\left(hc_k\right) + \left(1 - \widehat{hc_k}\right) \log\left(1 - hc_k\right)\right)$$
(7)

where hc_k is the predicted result and $\widehat{hc_k}$ is the real label.

C. Prediction in Test Stage

We extracted the top-k coordinates from the predicted heat maps and then added the corresponding offset, as shown in (8), to obtain the k keypoint coordinates. We utilize a relatively large value of k, such as 500, to perform an initial selection in each patch. This ensures a comprehensive coverage of ship candidates, even in scenarios where the number of ships may be high. Next, we refine the selection process by considering the confidence scores associated with the detected ships. As shown in Fig. 6(a), the keypoint heatmap is superimposed on the original image. The position of the keypoint heatmap corresponds

closely to the position of the ship

$$\{(xi + xi_{\text{off}}, yi + yi_{\text{off}}) | i = 1, 2, 3, \ldots, k\}.$$
(8)

An image contains multiple targets, and many keypoints are generated simultaneously. It is necessary to confirm the keypoints belonging to the same target. Taking the center keypoint as a reference, the coordinates of the center point and vector predicted were added to obtain the coordinates of the two endpoints. We then compared them with the coordinates of the two endpoints extracted from the heat map. If the Euclidean distance between them was less than a certain threshold, the same goal was considered, as in (9)–(11), the coordinates of the central keypoint $C\left(Cx, Cy\right)$, the coordinates of the upper Keypoint $T\left(Tx, Ty\right)$, and the coordinates of the lower keypoint $B\left(Bx, By\right)$. The vector from the central keypoint to the upper endpoint is Vct, that from the lower endpoint is Vcb, and s_k is the threshold. The keypoint matching is shown in Fig. 6(b)

$$Vct = (2v_{tk} - 1)\frac{l_k}{2}$$
(9)

$$Vcb = (2v_{bk} - 1)\frac{l_k}{2}$$
(10)

$$\begin{cases} d\left(C + Vct, T\right) \leq s_k \\ d\left(C + Vcb, B\right) \leq s_k. \end{cases}$$
(11)

We only required the coordinates of the center point and any endpoint, plus the length of the short side of the ship, to obtain the coordinates of the OBB. If a center point matches the upper and lower keypoints simultaneously, two OBBs appear simultaneously. The OBB box score is the average of the center point and endpoint scores. Finally, NMS was used to remove the box with an excessive overlap rate. Fig. 6(c) shows the OBB calculated using the keypoint coordinates and the length of the short side.

The upper and lower keypoints set by this algorithm were only related to the spatial position of the center point and did not contain category information. This method decoupled the target detection and bow classification. The two keypoints T and B matched the spatial positions of the bow and stern. This was beneficial for extracting the features of the head and tail. Two classifiers were used at the two endpoints to predict whether a point is a bow. When the center point only matched one endpoint and the classification threshold of this point was greater than 0.5, it was the bow. Otherwise, it was the stern. When two endpoints were matched simultaneously, the scores of the two keypoints were compared, and the classification threshold of the point was used to calculate the position of the bow if the score was relatively large. In Fig. 6(d), the yellow circle represents the ship's bow.

D. Bow Classification Analysis Model

We established a simple bow classification analysis model, as shown in (12). h_t represents the bow classification threshold at point T, h_b represents the bow classification threshold at point B, h_u represents the uncertainty of bow classification, and a value
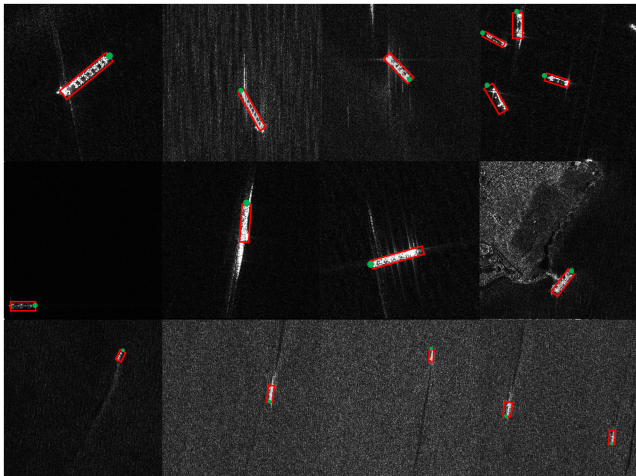
Fig. 7.   Annotation example.

closer to zero indicates greater uncertainty

$$h_u = |h_t - h_b|. \tag{12}$$

III. EXPERIMENT

A. Data

The data used in this study were obtained from the RSDD-SAR [42] and FUSAR-ship [43] datasets. RSDD-SAR, which includes Gaofen-3 and TerraSAR-X satellite data, has multiple imaging modes, polarization modes, and resolutions. The FUSAR-ship high-resolution ship dataset contained 15 main ship classes. The data slices were taken from 126 GF-3 remote-sensing images. The polarization mode included HH and VV, the resolution was 1.124 m $\times$ 1.728 m, and the imaging mode was the UFS mode, covering various sea, land, coast, river, and island scenes. A total of 1237 data slices were selected from the RSDD-SAR dataset, and 533 data slices were selected from the FUSAR-ship dataset. We selected 80% of the data from the two subsets as the training set and 20% as the test set. We relabeled the selected data using four corner points and four head points for labeling purposes (see Fig. 7).

In this study, the identification process of the SAR bow and stern is shown in Fig. 8. For SAR ship slices with maritime mobile service identification information, the bow and stern are determined by comparing real photos. For the existence of an optical image matching the SAR, the bow and stern of the ship can be determined by comparing the optical image. For ship slices without additional auxiliary information, the discriminative knowledge accumulated through a large number of comparisons is used for discrimination. This is also the basis for the slice selection of the two datasets.

There are three main bases for assessing the bow and stern of a ship in high-resolution SAR images, that is, shape, scattering, and wake information. The inclusion of wakes in our data was
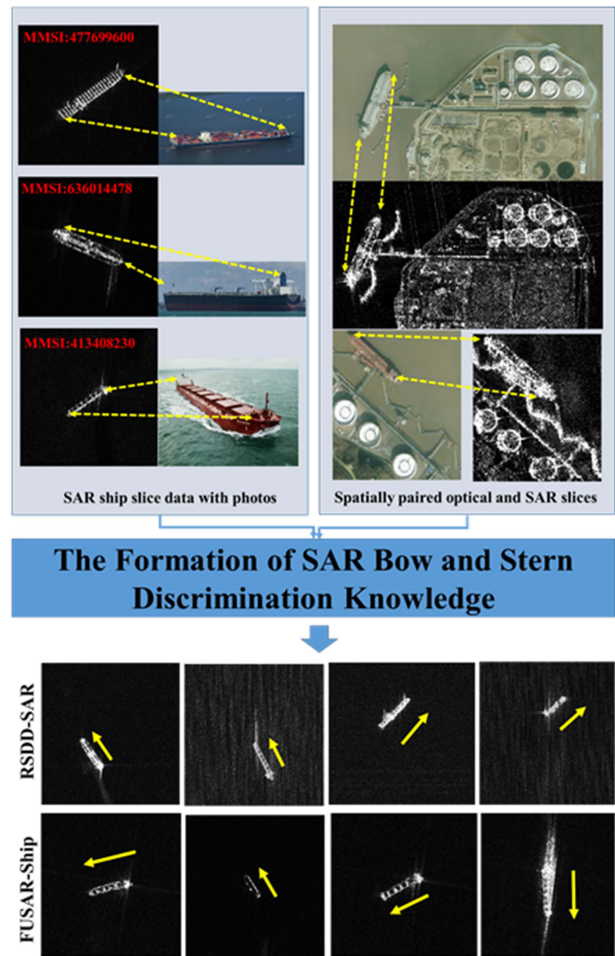

Fig. 8.   Process of determining the direction of the ship.

TABLE II
SIZE DISTRIBUTION OF TRAINING DATASET AND TEST DATASET

	Ship size	Small (area<625)	Medium (625<=area<=7500)	Large (area>7500)
Train dataset				
	Number	246	1327	16
	Ratio	15.48%	83.51%	1%
Test dataset	Number	61	333	6
	Ratio	15.25%	83.25%	1.5%

sparse and predominantly relied on morphological and scattering information. Fig. 9 shows the key components needed for identifying the bow and stern of a ship.

Table II presents the size distributions of the training and test sets and the size division references [44]. The proportions of large, medium, and small particles in the training and test sets were relatively consistent (see Table II). The proportion of medium objects was the largest, followed by small objects and large objects. In Fig. 10, an interval of 10° was used to obtain statistics on the distribution of the bow angle of the training and test sets, and the heading of the bow in the dataset is distributed in all directions. Fig. 11 shows the width and height distributions of the OBB. The width and height distribution is uneven, and the aspect ratio is relatively large.
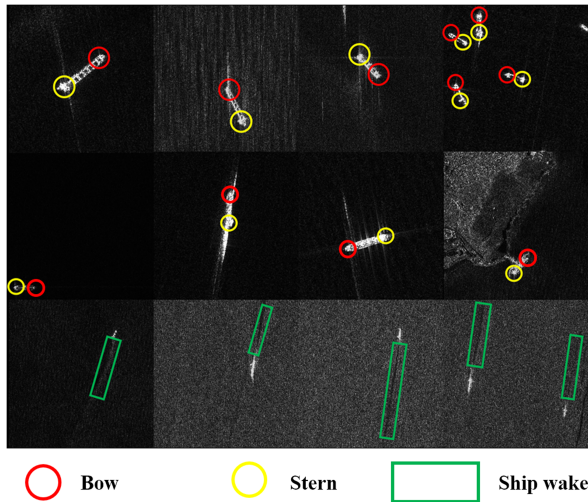
Fig. 9.   Key components needed for differentiating between the bow and stern.
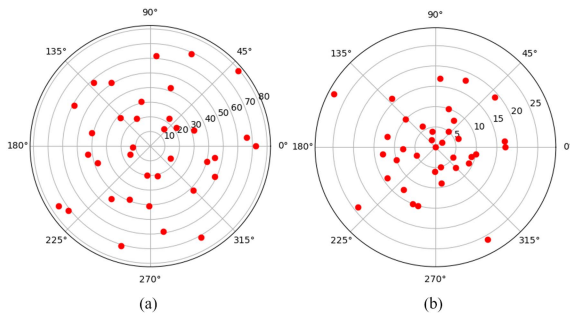

Fig. 10.   Distribution of bow angles in the training set and test set. (a) Number distribution of ships from different angles in the training set. (b) Number distribution of ships from different angles in the test set.

B. Experimental Details

Our deep-learning framework was based on Pytorch1.6, and the GPU model was an NVIDIA GTX 1080Ti. The training and testing image sizes were 512×512 pixels. The images were randomly rotated and flipped during training. The initial learning rate was set to 0.000125, the optimizer was Adam, the learning rate decay strategy was exponential decay, and iterative training was performed 200 epochs.

C. Comparison of OBB Ship Detection Algorithms

In this section, the OBB detection performances of our proposed algorithm and five other algorithms (gliding vertex [28], oriented RCNN [45], R3Det [46], YOLOv5 + CSL [29], and faster R-CNN (OBB) [15]) were compared. For a fair comparison, other algorithms, except for the YOLOv5 + CSL model, including the algorithm proposed in this study, used ResNet50 as the backbone network for feature extraction. The evaluation indicators used were precision rate (Precision), recall rate (Recall), $F1$-score ($F1$-score), transmission frames per second (FPS), and average accuracy rate (AP), where AP50 means merge, and the average accuracy rate when the intersection over union (IoU) threshold was 0.5.
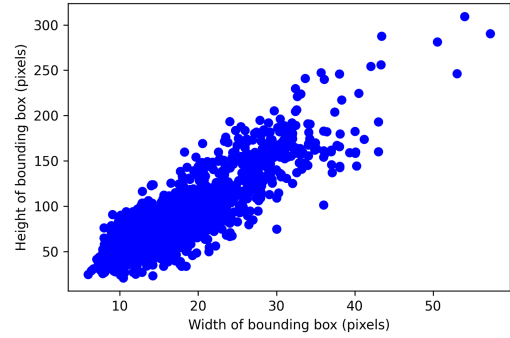

Fig. 11.   OBB width and height distribution.

As shown in Table III, the proposed algorithm obtained the most successful results for the $F1$-score and AP, which were 0.979 and 0.908, respectively. The R3Det algorithm had the highest recall rate and a lower precision rate, resulting in the lowest $F1$-score. The AP values of the R3Det and oriented RCNN algorithms were second only to those of the proposed algorithm, and the fastest algorithm in terms of speed was the YOLOv5 + CSL algorithm, which reached 50FPS. The speeds of the other algorithms did not differ significantly. The slowest algorithm was R3Det because it uses multiple refinement schemes to improve accuracy. This can be observed in Fig. 12 that the gliding vertex, oriented RCNN, and faster R-CNN (OBB) algorithms produced more false alarms near the shore, whereas R3Det and YOLOv5 + CSL produce more missed detections. Our proposed method showed better performance results near the shore. In the third row, the oriented RCNN algorithm missed small targets, while other algorithms detected small targets. In the fourth row, oriented RCNN and YOLOv5 + CSL missed the target. The rotating box obtained using the gliding vertex algorithm did not correctly surround the target. Although the faster R-CNN (OBB) algorithm detected the correct target, it also detected the noise generated by the ship's motion. R3Det and the proposed algorithm correctly detected objects, and the proposed algorithm obtained bounding boxes that were more consistent with the ground truth.

D. Bow Discriminant Analysis

Table IV compares the three methods relying on keypoints for bow differentiation. The $C +$ head method, such as CHPDet [47], is an anchor-free method that uses the center point and head point to determine the OBB and bow. $C + T + B$ used the center point and the up and down points to determine the OBB. It then compared the thresholds of the upper and lower keypoints. A point with a larger threshold uses the bow classifier where the point is located to determine whether the point is the bow. The $T + B$ method uses only the upper and lower endpoints for detection. Its bow classification strategy is consistent with that of $C + T + B$. In the $C + T + B$ method, the accuracy rate of bow differentiation was the highest, and the accuracy rate of bow classification (accurate) was the ratio of the number of objects correctly classified as bows to the total number of objects. The $T + B$ scheme was lower than the $C + T + B$

TABLE III
COMPARISON OF DIFFERENT OBB ALGORITHMS

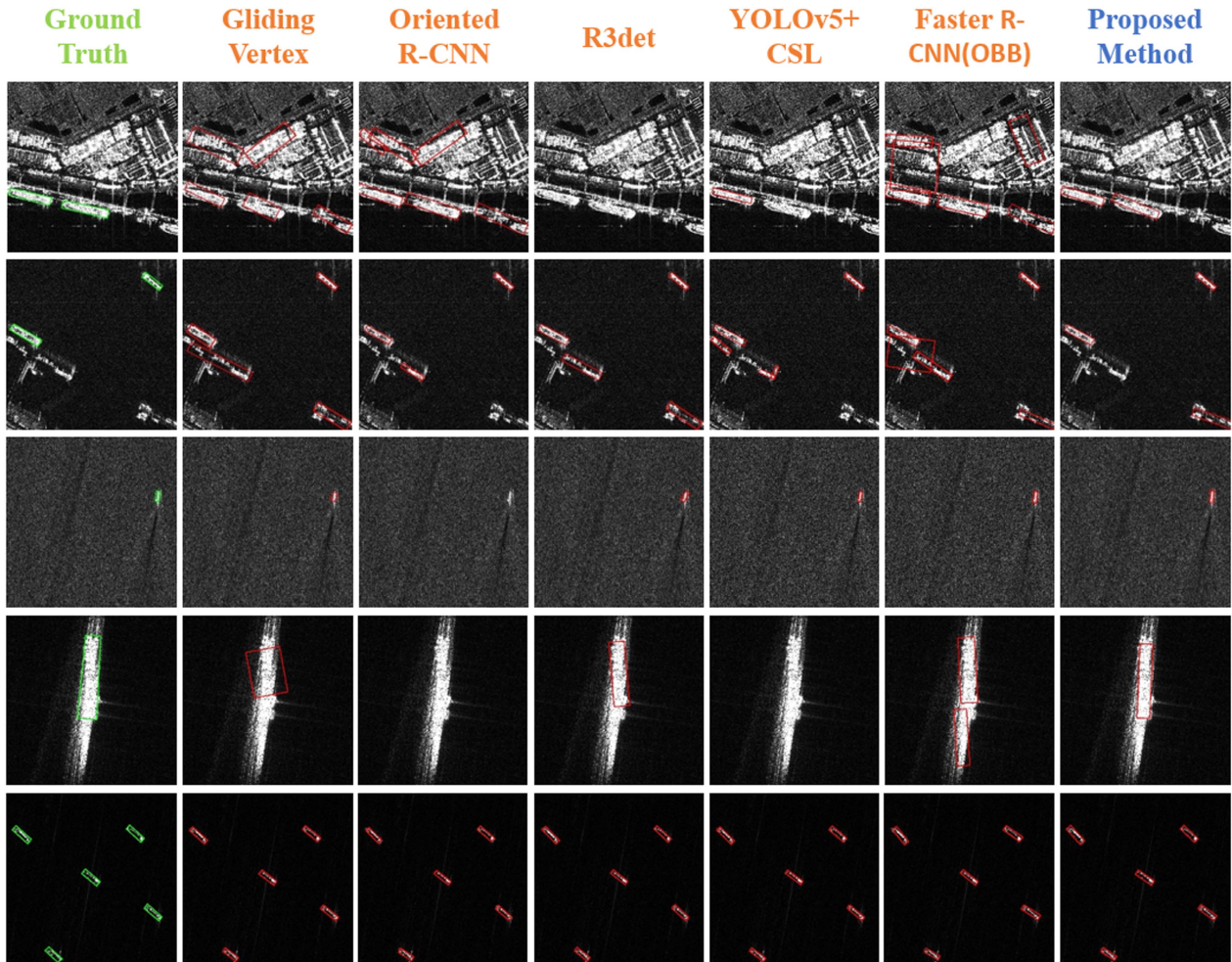| Method	Precision	Recall	F1	AP$_{50}$	FPS
Gliding vertex	0.874	0.902	0.887	0.898	19.9
Oriented RCNN	0.981	0.935	0.957	0.907	20
R3Det	0.787	**0.980**	0.873	0.907	16.6
YOLOv5+CSL	0.896	0.955	0.924	0.884	**50**
Faster R-CNN(OBB)	0.869	0.965	0.915	0.900	20.3
TKP-Net (Our)	**0.982**	0.977	**0.979**	**0.908**	18.1

Fig. 12.   Visualization comparison of different methods.

TABLE IV
COMPARISON OF TWO KEYPOINTS AND THREE KEYPOINTS

Method	Precision	Recall	F1	AP$_{50}$	Accurate
C + head	**0.994**	0.937	0.964	**0.908**	0.910
T + B	0.905	0.882	0.893	0.788	0.725
C + T + B	0.982	**0.977**	**0.979**	**0.908**	**0.925**

scheme for each accuracy index. The C + head scheme had the highest accuracy rate, but a low recall rate. The $F1$-score was not as high as that of the $C + T + B$ scheme. The main defect of the C + head scheme was that when the model could not identify the bow of the ship, and the detection target was not recognized,

resulting in missed detection. The proposed method divided bow differentiation and ship detection into two stages such that the bow discrimination result does not affect the detection result. Fig. 13 showed the OBB detection results of SAR ships with bow marks.

Deep learning relies on a single categorical variable to make decisions which may be a problem of overconfidence [48]. In many remote-sensing applications, it is critical to estimate prediction uncertainty. The proposed method could simultaneously predict two bow classification results at the same time in most cases, and 82% of the test samples could simultaneously obtain two bow classification results at the same
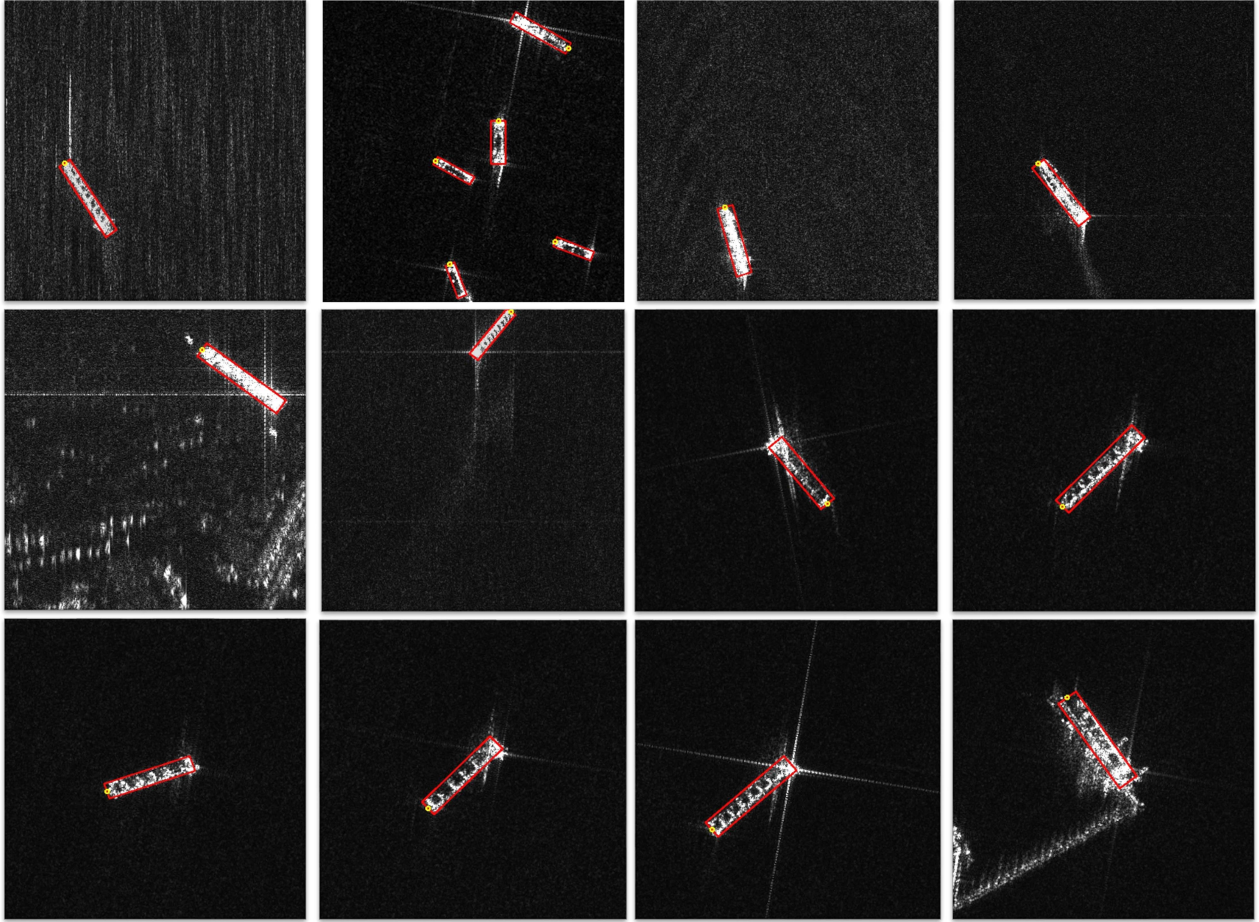
Fig. 13. TKP-net detection results. The yellow circle represents the bow.

TABLE V
ANALYSIS OF BOW DIRECTION

Threshold	Misclassified sample proportion	Correctly classified sample proportion
<0.1	0.115	0.009
<0.2	0.153	0.016
<0.3	0.230	0.016
<0.4	0.230	0.016
<0.5	0.307	0.019
<0.6	0.384	0.019
<0.7	0.461	0.039
<0.8	0.538	0.043
<0.9	0.615	0.046

time. The two classification results were used to determine whether the point was the bow of the ship. Owing to the influence of accidental and perceptual uncertainties, there are contradictions between the two classification results. Using this contradiction can reduce the inconsistencies to a certain extent. Table V presents the proportions of incorrectly and correctly classified samples in the bow classification test samples under different thresholds. As the threshold increases, the proportion of misclassified samples far exceeds the proportion of correctly classified samples. Assuming that we have higher requirements for the reliability of the bow classification, we can set a large threshold. With a threshold of 0.9, our method can label 61.5% of the wrong samples as low confidence samples, and only 4.6% of the correctly classified samples are affected. This approach provides an analytical tool for decision reliability.

E. Ship Detection in Large-Scale SAR Images

Large-scale SAR image testing can be used to validate and evaluate the performance of ship detection algorithms in real-world scenarios. Utilizing large-scale SAR images for testing enables a more accurate assessment of algorithm performance in handling complex backgrounds, multiple targets, occlusions, and other challenging situations.

We tested on two large-scale SAR images, as shown in Fig. 14(a), and ship detection had some false alarms on land or islands. As shown in Fig. 14(b), the detection result of large-scale SAR ships on the sea surface was better when there was no interference from land or islands. We selected ships with wakes on the image to verify the accuracy of our bow classification. As shown
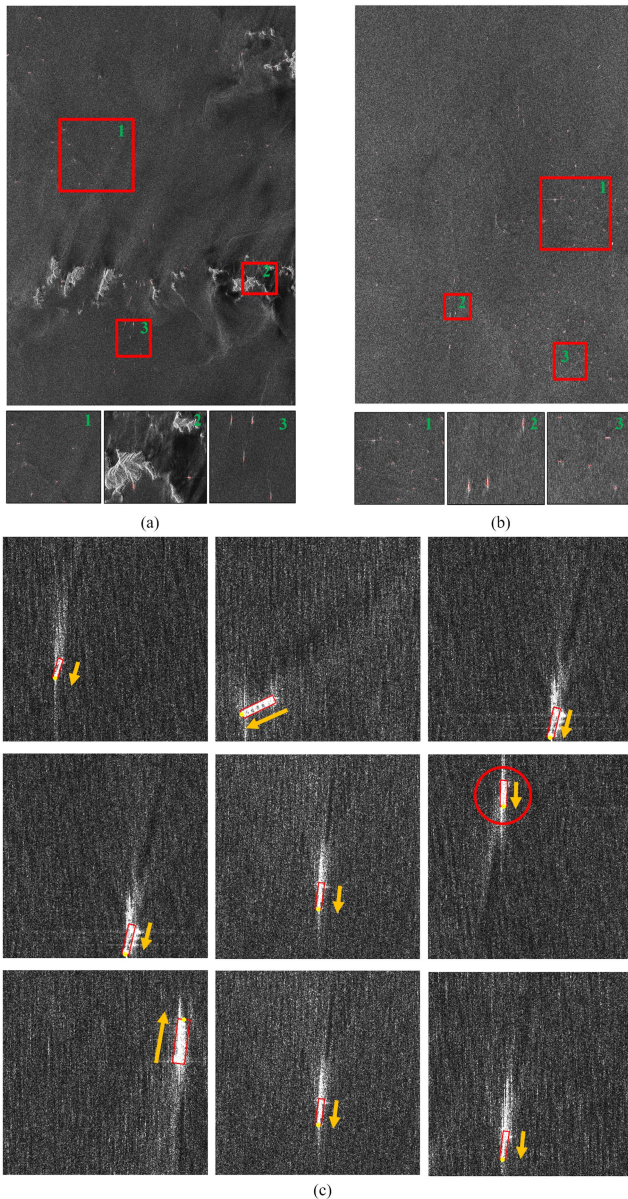
Fig. 14. Results in large-scale SAR images. (a) Large-scale SAR image with islands. (b) Sea surface SAR image without islands. (c) Bow extraction result with wake verification. The red circle represents the wrong classification result.

TABLE VI
DETECTION PERFORMANCE IN DENSE PORTS

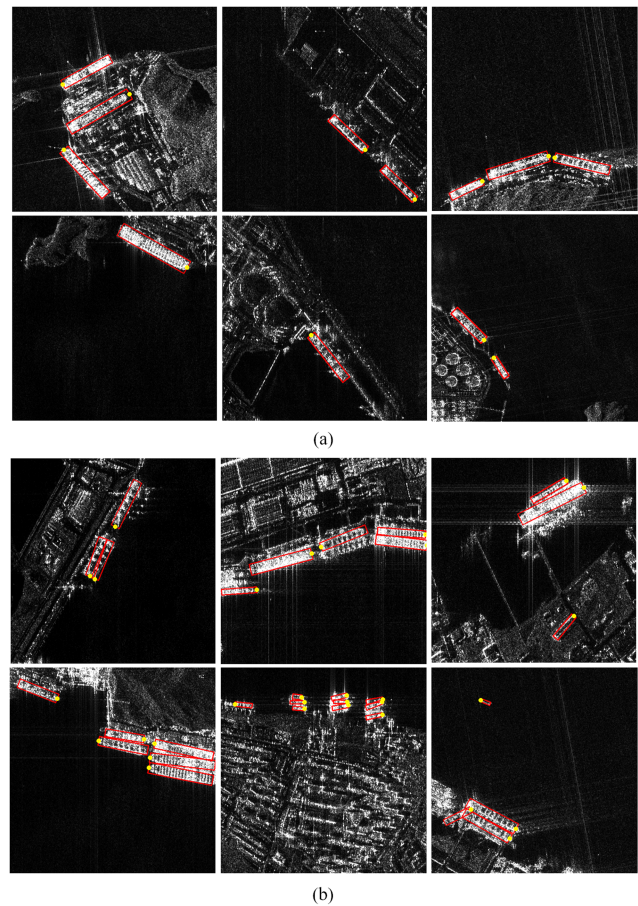| Scene	Precision	Recall	F1	AP$_{50}$	Accurate
Port	0.828	0.746	0.784	0.772	0.808

Fig. 15. Ship detection and bow extraction in dense port areas. (a) Ships docked on the shore alone. (b) Ships are densely packed together.

in Fig. 14(c), the direction of most ships was consistent with the direction verified by ship wakes. This can prove to a certain extent that our method has good performance in extracting ship direction from large-scale SAR images.

F. Ship Detection and Bow Discrimination in Dense Ports

This study addresses the task of ship detection and bow discrimination in dense port environments as one aspect of our brder research focus. We conduct an experiment to evaluate the proposed method and the obtained performance metrics are listed in Table VI. It can be seen from Table VI that the performance of ship detection and bow discrimination in dense ports has declined. Ship detection in dense port areas faces challenges,

such as high ship density, background interference, and ship diversity. There are usually a large number of ships berthing or passing through the port area, and the distance between the ships is relatively close, which increases the overlapping and occlusion of ships. In addition, various objects and structures in the port area, such as docks, buildings, cranes, and stacked cargo, will generate a lot of interference and occlusion, making the detection of ships more difficult. In addition, the port area frequently moors and passes the ships of various types and sizes, including cargo ships, passenger ships, fishing boats, etc., which adds to the complexity of ship classification and identification. As shown in Fig. 15(a), there is no dense arrangement of ships, and the detection effect is better when the ships are docked at the port. As shown in Fig. 15(b), missed detection occurs when ships are densely arranged.

Therefore, our proposed method suffers from some limitations in dense port areas.

Our proposed method exhibits an average inference time of 73.6 s for large-scale SAR images. The image dimensions

TABLE VII
BOW EXTRACTION ACCURACY OF DIFFERENT TYPES OF SHIPS

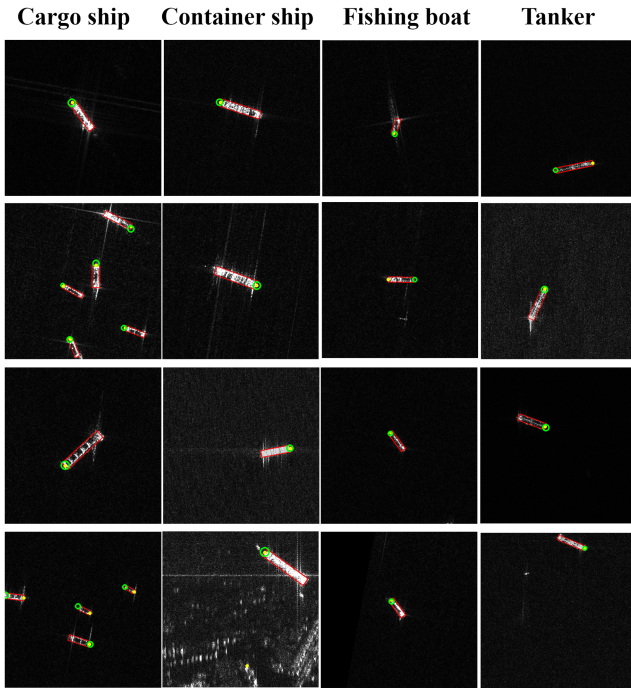| Ship types	Cargo ship	Container ship	Fishing boat	Tanker
Accurate	0.886	0.964	0.888	0.977

Fig. 16. Bow extraction results of different types of ships. The green circle is the ground truth label, and the yellow circle is the predicted bow coordinates.

in question are $15\,668 \times 21\,725$. These values represent the computational time our model requires when processing such large-scale SAR images.

G. Accuracy of Direction for Different Types of Ships

We selected data with class labels from the test set to test the accuracy of bow classification. It can be seen from Table VII that the classification accuracy of the bow is the highest for oil tankers and the lowest for fishing boats. It can be seen from Fig. 16 that the objects classified incorrectly by the bow are basically relatively small. Because the scattering structure of the small target is not clear enough, the amount of information provided is also small.

IV. DISCUSSION

The method we proposed for ship detection and bow classification has numerous advantages and innovations in the field. However, we also acknowledge certain limitations that need to be considered and addressed in further research and applications.

First, our algorithm requires high-resolution SAR data. Lower resolution SAR data may result in decreased algorithm performance, affecting the accuracy of ship detection and recognition. Therefore, our algorithm may be constrained by data resolution limitations in certain data sources or application scenarios. Second, our algorithm has certain requirements regarding the clarity of ship scattering structures. If the scattering structures of ships are too blurry or indistinct, the accuracy of the algorithm may be compromised. Polarization information provides valuable insights into the scattering behavior of objects, including ships, and helps enhance the detection and characterization of ships in SAR images [49], [50]. In the future, we can consider incorporating polarization information into our method to improve its performance.

In addition, our algorithm's performance may decline when dealing with densely packed port areas. This is due to potential occlusion, overlap, or interference between ships in such congested port regions. In these cases, our algorithm may experience false detections or missed detections, requiring further improvements to enhance its capability in handling dense port areas.

To address the aforementioned limitations, we can explore the following avenues for improvement.

1) Investigating techniques to adapt to low-resolution SAR data, such as image enhancement or multiscale processing methods, to enhance the algorithm's robustness.
2) Conducting further research on feature extraction and model design for ship scattering structures to address the challenges posed by blurry scattering structures.
3) Explore the utilization of prior knowledge or other sensor data to assist in ship detection and recognition tasks in dense port areas.

It is important to emphasize that our algorithm demonstrates good performance and robustness in most cases. However, further research and improvements are required to address the aforementioned limitations. An honest and comprehensive description of these limitations is crucial for readers to gain the accurate understanding of our research work and to provide guidance and inspiration for future studies.

In future research, we will strive to overcome these limitations and propose the corresponding improvement strategies to further enhance the performance and applicability of our algorithm.

V. CONCLUSION

In this study, a method for obtaining various ship information, including position, length, width, and direction from high-resolution SAR images, was proposed. A ship detection method for any direction based on three keypoints without an anchor box has been proposed. In the first step of the method, the angle-prediction problem of the rotating box was converted into an estimation and matching problem for the keypoint position to determine the rotating box. In the second step, bow discrimination was performed using classifiers placed at two keypoints.

The experimental results show that our method achieves good performance with an AP of 90.8, an F1-score of 97.9%, and an accuracy of 92.5% for bow classification in nondense port area ship detection. We also used the difference in information extracted from the two keypoints to establish a simple and effective uncertainty analysis method.

REFERENCES

[1] Y. Liu, J. Zhao, and Y. Qin, "A novel technique for ship wake detection from optical images," *Remote Sens. Environ.*, vol. 258, 2021, Art. no. 112375.

[2] H. Lang, Y. Xi, and X. Zhang, "Ship detection in high-resolution SAR images by clustering spatially enhanced pixel descriptor," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5407–5423, Aug. 2019.

[3] O. Pappas, A. Achim, and D. Bull, "Superpixel-level CFAR detectors for ship detection in SAR imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 9, pp. 1397–1401, Sep. 2018.

[4] H. Lang and S. Wu, "Ship classification in moderate-resolution SAR image by naive geometric features-combined multiple kernel learning," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 10, pp. 1765–1769, Oct. 2017.

[5] Y. Ren, X. Li, and H. Xu, "A deep learning model to extract ship size from Sentinel-1 SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Mar. 2022, Art. no. 5203414.

[6] Q. Hu, S. Hu, and S. Liu, "BANet: A balance attention network for anchor-free ship detection in SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Jan. 2022, Art. no. 5222212.

[7] Y. Cui, G. Zhou, J. Yang, and Y. Yamaguchi, "On the iterative censoring for target detection in SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 641–645, Jul. 2011.

[8] S. Brusch, S. Lehner, T. Fritz, M. Soccorsi, A. Soloviev, and B. van Schie, "Ship surveillance with TerraSAR-X," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 3, pp. 1092–1103, Mar. 2011.

[9] X. Sun, R. Tao, and X. Bai, "A fast order method on OS-CFAR detector in SAR images," in *Proc. 2nd Asian-Pacific Conf. Synthetic Aperture Radar*, 2009, pp. 725–728.

[10] A. C. Frery, H.-J. Müller, C. C. F. Yanasse, and S. J. S. Sant'Anna, "A model for extremely heterogeneous clutter," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 3, pp. 648–659, May 1997.

[11] G. Gao and G. Shi, "CFAR ship detection in nonhomogeneous sea clutter using polarimetric SAR data based on the notch filter," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4811–4824, Aug. 2017.

[12] J. Ai, R. Tian, Q. Luo, J. Jin, and B. Tang, "Multi-scale rotation-invariant Haar-like feature integrated CNN-based ship detection algorithm of multiple-target environment in SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 10070–10087, Dec. 2019.

[13] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 580–587.

[14] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 1440–1448.

[15] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

[16] J. Dai, Y. Li, K. He, and J. Sun, "R-FCN: Object detection via region-based fully convolutional networks," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2016, pp. 379–387.

[17] W. Liu et al., "SSD: Single shot multibox detector," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 21–37.

[18] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 779–788.

[19] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," 2018, *arXiv:1804.02767*.

[20] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal speed and accuracy of object detection," 2020, *arXiv:2004.10934*.

[21] H. Law and J. Deng, "CornerNet: Detecting objects as paired keypoints," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 765–781.

[22] X. Zhou, J. Zhuo, and P. Krahenbuhl, "Bottom-up object detection by grouping extreme and center points," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 850–859.

[23] X. Zhou, D. Wang, and P. Krähenbühl, "Objects as points," 2019, *arXiv:1904.07850*.

[24] Z. Cui, Q. Li, Z. Cao, and N. Liu, "Dense attention pyramid networks for multi-scale ship detection in SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 11, pp. 8983–8997, Nov. 2019.

[25] H. Guo, X. Yang, N. Wang, and X. Gao, "A Centernet++ model for ship detection in SAR images," *Pattern Recognit.*, vol. 112, Apr. 2021, Art. no. 107787.

[26] Z. Deng, H. Sun, S. Zhou, and J. Zhao, "Learning deep ship detector in SAR images from scratch," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 6, pp. 4021–4039, Jun. 2019.

[27] G.-S. Xia et al., "Dota: A large-scale dataset for object detection in aerial images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 3974–3983.

[28] Y. Xu et al., "Gliding vertex on the horizontal bounding box for multi-oriented object detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 4, pp. 1452–1459, Apr. 2021.

[29] X. Yang and J. Yan, "Arbitrary-oriented object detection with circular smooth label," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 677–694.

[30] Y. He, F. Gao, J. Wang, A. Hussain, E. Yang, and H. Zhou, "Learning polar encodings for arbitrary-oriented ship detection in SAR images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 3846–3859, Mar. 2021.

[31] S. Ermakov, I. Kapustin, and T. Lazareva, "Ship wake signatures in radar/optical images of the sea surface: Observations and physical mechanisms," in *Proc. Remote Sens. Ocean, Sea Ice, Coastal Waters, Large Water Regions*, 2014, vol. 9240, pp. 156–161.

[32] R. Del Prete, M. D. Graziano, and A. Renga, "First results on wake detection in SAR images by deep learning," *Remote Sens.*, vol. 13, no. 22, Nov. 2021, Art. no. 4573.

[33] Z. Xu, B. Tang, and S. Cheng, "Faint ship wake detection in PolSAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 7, pp. 1055–1059, Jul. 2018.

[34] J. V. Wehausen and E. V. Laitone, "Surface waves," in *Fluid Dynamics/Strömungsmechanik*. Berlin, Germany: Springer, 1960, pp. 446–778.

[35] A. M. Reed, R. F. Beck, O. M. Griffin, and R. D. Peltzer, "Hydrodynamics of remotely sensed surface ship wakes," *SNAME Trans.*, vol. 98, pp. 319–363, 1990.

[36] J. Chong and M. Zhu, "Ship wake detection algorithm in SAR image based on normalized grey level Hough transform," *J. Image Graph.*, vol. 9, no. 2, pp. 146–150, 2004.

[37] A. Jiaqiu et al., "A novel ship wake CFAR detection algorithm based on SCR enhancement and normalized Hough transform," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 681–685, Jul. 2011.

[38] I. G. Rizaev, O. Karakuş, S. J. Hogan, and A. Achim, "Modeling and SAR imaging of the sea surface: A review of the state-of-the-art with simulations," *ISPRS J. Photogramm. Remote Sens.*, vol. 187, pp. 120–140, May 2022.

[39] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.

[40] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 2117–2125.

[41] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 2980–2988.

[42] C. Xu et al., "RSDD-SAR: Rotated ship detection dataset in SAR images," *J. Radars*, vol. 11, no. 4, pp. 581–599, Aug. 2022.

[43] X. Hou, W. Ao, Q. Song, J. Lai, H. Wang, and F. Xu, "FUSAR-ship: Building a high-resolution SAR-AIS matchup dataset of Gaofen-3 for ship detection and recognition," *Sci. China Inf. Sci.*, vol. 63, no. 4, Apr. 2020, Art. no. 140303.

[44] T. Zhang et al., "SAR ship detection dataset (SSDD): Official release and comprehensive data analysis," *Remote Sens.*, vol. 13, no. 18, Sep. 2021, Art. no. 3690.

[45] X. Xie, G. Cheng, J. Wang, X. Yao, and J. Han, "Oriented R-CNN for object detection," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2021, pp. 3520–3529.

[46] X. Yang, J. Yan, Z. Feng, and T. He, "R3Det: Refined single-stage detector with feature refinement for rotating object," in *Proc. AAAI Conf. Artif. Intell.*, 2021, pp. 3163–3171.

[47] F. Zhang, X. Wang, S. Zhou, Y. Wang, and Y. Hou, "Arbitrary-oriented ship detection through center-head point extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Oct. 2022, Art. no. 5612414.

[48] K. Ristovski, S. Vucetic, and Z. Obradovic, "Uncertainty analysis of neural-network-based aerosol retrieval," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 409–414, Feb. 2012.

[49] R. L. Paes, F. Nunziata, and M. Migliaccio, "On the capability of hybrid-polarity features to observe metallic targets at sea," *IEEE J. Ocean. Eng.*, vol. 41, no. 2, pp. 346–361, Apr. 2016.

[50] M. Adil, F. Nunziata, A. Buono, D. Velotto, and M. Migliaccio, "Polarimetric scattering by a vessel at different incidence angles," *IEEE Geosci. Remote Sens. Lett.*, vol. 20, Aug. 2023, Art. no. 4008605.

Gang Zheng (Senior Member, IEEE) received the B.Eng. degree in electronic information engineering from Zhejiang University, Hangzhou, China, in 2003, and the M.S. and Ph.D. degrees in radio physics from the University of Electronic Science and Technology of China, Chengdu, China, in 2006 and 2010, respectively.

From 2010 to 2013, he was an Assistant Researcher with the State Key Laboratory of Satellite Ocean Environment Dynamics, Second Institute of Oceanography, Ministry of Natural Resources, Hangzhou, where he was an Associate Researcher, from 2013 to 2020, and has been a Researcher, since 2020. His research interests include ocean microwave remote sensing, artificial intelligence (AI) applications, image processing, and electromagnetic numerical modeling.

Dr. Zheng is an Editorial Board Member of the ocean section of Remote Sensing and a Topic Editor for *Big Earth Data*. From 2018 to 2020, he also served as the Guest Editor for *Remote Sensing*, special issues on AI-Based Remote Sensing Oceanography, Synergy of Remote Sensing and Modelling Techniques for Ocean Studies, and Tropical Cyclones Remote Sensing and Data Assimilation.

Xiunan Li received the B.S. degree in geographic information system from Ludong University, Yantai, China, in 2016, and the M.S. degree in physical oceanography from the Second Institute of Oceanography, Ministry of Natural Resources, Hangzhou, China, in 2020. He is currently working toward the Ph.D. degree in multi-source remote sensing ship detection and recognition with Ocean College, Zhejiang University, Zhoushan, China.

His research interests include ocean microwave remote sensing, deep learning, and image processing.

Peng Chen was born in Hunan, China. He received the Ph.D. degree in geographic information system from Zhejiang University, Hangzhou, China, in 2011.

He is a Senior Engineer in marine remote sensing with the State Key Laboratory of Satellite Ocean Environment Dynamics, Hangzhou, China, working on the development of algorithms detecting marine targets (ship, oil rig, and oil slick).

Dan Luo is currently working toward the Ph.D. degree in AIS anomaly trajectory detection and classification with the College of Oceanology, Zhejiang University, Hangzhou, China.

His research interests include remote sensing image processing and AIS data analysis.

Jingsong Yang received the B.S. degree in physics and the M.S. degree in theoretical physics from Zhejiang University, Hangzhou, China, in 1990 and 1996, respectively, and the Ph.D. degree in physical oceanography from the Ocean University of China, Qingdao, China, in 2001.

Since 1996, he has been with the Second Institute of Oceanography (SIO), Ministry of Natural Resources, Hangzhou, China, where he is the Head of the Microwave Marine Remote Sensing, State Key Laboratory of Satellite Ocean Environment Dynamics. Since 2002, he has been a Supervisor of graduate students with SIO and, since 2011, an Adjunct Professor and a Doctoral Supervisor with Zhejiang University. He has more than 20 years of experience in microwave marine remote sensing. He has been a Principal Investigator and a participant of more than 20 research projects and authored or coauthored more than 100 scientific articles in peer-reviewed journals and international conference proceedings. His research interests include microwave marine remote sensing, data fusion, image processing, and satellite oceanography.

Aiying Lu received the B.S. degree in geoinformation science and technology from the Ocean University of China, Qingdao, China, in 2021. She is currently working toward the M.S. degree in physical oceanography with the Second Institute of Oceanography, Ministry of Natural Resources, Hangzhou, China.

Her research interests include hyperspectral remote sensing and maritime target recognition.

Wentao An (Member, IEEE) received the B.S. degree in communication engineering from Nankai University, Tianjin, China, in 2003, and the Ph.D. degree in electronic engineering from Tsinghua University, Beijing, China, in 2010.

He is currently an Associate Researcher with the Department of Systematic Engineering, National Satellite Ocean Application Service, Beijing, China. His research interests include polarimetric synthetic aperture radar data processing and target detection with SAR imagery.

Zimu Wang received the B.S. degree in remote sensing from the School of Remote Sensing Information Engineering, Wuhan University, Wuhan, China, in 2022. He is currently working toward the M.S. degree in multi-source remote sensing maritime target and environment sensing technology with the Second Institute of Oceanography, Department of Natural Resources, Hangzhou, China.

His research interests include ocean microwave imaging and object detection.