

FUSAR-Ship: building a high-resolution SAR-AIS matchup dataset of Gaofen-3 for ship detection and recognition

Xiyue HOU¹, Wei AO¹, Qian SONG¹, Jian LAI², Haipeng WANG¹ & Feng XU^{1*}

¹Key Lab for Information Science of Electromagnetic Waves (MoE), Fudan University, Shanghai 200433, China;

²Shanghai Center for Gaofen Data and Applications, Shanghai 201109, China

Received 7 October 2019/Revised 13 December 2019/Accepted 1 January 2020/Published online 10 March 2020

Abstract Gaofen-3 (GF-3) is China's first civil C-band fully polarimetric spaceborne synthetic aperture radar (SAR) primarily missioned for ocean remote sensing and marine monitoring. This paper proposes an automatic sea segmentation, ship detection, and SAR-AIS matchup procedure and an extensible marine target taxonomy of 15 primary ship categories, 98 sub-categories, and many non-ship targets. The FUSAR-Ship high-resolution GF-3 SAR dataset is constructed by running the procedure on a total of 126 GF-3 scenes covering a large variety of sea, land, coast, river and island scenarios. It includes more than 5000 ship chips with AIS messages as well as samples of strong scatterer, bridge, coastal land, islands, sea and land clutter. FUSAR-Ship is intended as an open benchmark dataset for ship and marine target detection and recognition. A preliminary 8-type ship classification experiment based on convolutional neural networks demonstrated that an average of 79% test accuracy can be achieved.

Keywords FUSAR-Ship, Gaofen-3, SAR-AIS matchup, automatic target recognition, multi-scale CFAR, deep learning

Citation Hou X Y, Ao W, Song Q, et al. FUSAR-Ship: building a high-resolution SAR-AIS matchup dataset of Gaofen-3 for ship detection and recognition. Sci China Inf Sci, 2020, 63(4): 140303, <https://doi.org/10.1007/s11432-019-2772-5>

1 Introduction

Synthetic aperture radar (SAR) is one of the most powerful tools of Earth observation (EO) owing to its unique ability of all-time imaging in all-weather conditions [1]. SAR plays an increasingly important role in ocean surveillance such as oil spill detection, fishery management and ship detection [2]. During past decades, great progress has been made in ship detection of low- and medium-resolution SAR images [3–5]. Recently, some high-resolution SAR satellites have been successfully launched, which enables us to classify ship targets into categories, like cargo, tanker, tug, and fisher. Gaofen-3 (GF-3) satellite is China's first civil C-band high-resolution quad-pol SAR satellite specifically missioned for ocean remote sensing. The nominal highest resolution of GF-3 data is up to 1 m. GF-3 data have been widely used on applications such as ship recognition [6–10], aircraft detection [11], and image translation between optical and SAR images [12].

Radar target recognition is of great interests [13, 14]. A typical SAR automatic target recognition (ATR) system comprises three stages, i.e., detection, discrimination and classification [15]. Constant false alarm rate (CFAR) [15], including many variations of it [16], is the most widely-used algorithm for

* Corresponding author (email: fengxu@fudan.edu.cn)

ship detection. CFAR is tailored for ship detection in homogeneous sea clutters. Discrimination is to eliminate false alarms in order to decrease the computational load of the subsequent classification. So, discrimination can be seen as a binary classification between ship targets and false alarms. Classification algorithms are usually composed of two phases, namely feature extraction and classification. Common features [17] used in SAR ATR include length, width, and polarization. Conventional classification algorithms [18] include support vector machine, decision tree, and neural networks. Recently, data-driven deep learning (DL) technology has revolutionized the computer vision (CV) regime. With a large volume of annotated training data, e.g., [19, 20], a deep learning-based algorithm can achieve an unprecedented level of accuracy in the classical CV tasks such as object recognition and scene understanding. DL has also been applied to SAR ATR, e.g., [15, 21]. Apparently, those DL algorithms heavily rely on the size and distribution of annotated training data. A large-scale annotated dataset is critical to the development of practically useful DL algorithms. This is also true for the automated interpretation of remote sensing images. To accelerate the application of DL in SAR ATR, this study presents a large-scale GF-3 ship image dataset with matched label information from ship automatic identification system (AIS) data.

There are several datasets available in the field of remote sensing. For example, DOTA is a large-scale optical dataset for object detection [22], which contains 2806 aerial images from different sensors and platforms. It includes a total of 188282 instances of 15 object categories. UCAS-AOD is another dataset of object detection in aerial images [23], containing automobile, aircraft and background negative samples. NWPU VHR-10 is a target detection dataset and has a total of 800 images [24, 25], including 10 categories [26]. Note these datasets are all optical images. SAR datasets are much more difficult to build mainly because of two reasons: (1) SAR data is much less publically accessible; (2) SAR images are difficult for human to understand and difficult for manual annotation. MSTAR is a classical dataset for SAR ATR consisting of foot-resolution X-band SAR image chips of 10 types of vehicles observed at various azimuth angles [27]. It has enabled many deep learning-based SAR ATR studies [15, 28] which achieved state-of-the-art performance. OpenSARShip is a medium-resolution ship dataset consisted of 11346 image chips cropped from a total of 41 Sentinel-1 images. It also contains ship information extracted from co-registered AIS data [29, 30]. It covers mainly 5 ports in Asia. It has 17 types of ships in total, but the population is mainly concentrated in a few common types, i.e., the largest portion is cargo type containing more than 9470 samples. On the other hand, the highest azimuth resolution in the dataset is 22 m, which is less useful for high-resolution ship recognition. Thus two high-resolution ship datasets have been presented in [31, 32], respectively. A total of 43819 ship ships with resolution ranges from 1.7 m to 22 m are constructed in [32]. However, because the dataset is designed for ship detection, detail information about the type of ships is not given. HR4S [31] (high-resolution SAR ship sample set) is constructed using 25 RADARSAT-2 and 28 GF-3 SAR images, which contains 22 types and a total of 1962 image chips.

This paper proposes a fully automatic annotation framework of SAR ship targets, designs a SAR ship taxonomy and builds a high-resolution SAR-AIS matched-up ship dataset named FUSAR-Ship. The annotation framework comprises two modules, a multi-scale CFAR algorithm [15] and an AIS-SAR matchup algorithm [29, 30]. In order to present a standard benchmark, we design an extensible SAR ship taxonomy which takes into account both appearance and functionality of ships. Based on the annotation framework and the SAR ship taxonomy, we build the FUSAR-Ship dataset covering various ship ATR scenarios such as near-shore regions, islands and dynamic sea conditions. It is intended as an open standard dataset for the purpose of ship SAR ATR algorithm development, evaluation and benchmarking. It has 15 ship categories, 98 ship subcategories, consisted of more than 5000 high-resolution ship samples and thousands of non-ship samples cropped from a total of 126 GF-3 images.

The major contribution of this paper is in threefold:

- Develop a fully automatic matchup and annotation procedure, which can automatically get rid of land areas, detect ships, filter out false alarms, matchup with AIS ground truths and finally generate labeled samples.
- Establish a high-resolution GF-3 SAR-AIS dataset of ship targets, which is intended as a standard benchmark dataset and hence built under three basic principles, namely reliability, consistency and

Table 1 GF-3 technical parameters as compared with Sentinel-1 and RadarSat-2

	GF-3	Sentinel-1	RadarSat-2
Frequency (km)	755	693	798
Peak power (kW)	1.5	4.7	1.27
Incident angle	10°–60°	20°–45°	10°–60°
Antenna size	15 m×1.5 m	12.3 m×0.84 m	15 m×1.37 m
Bandwidth (MHz)	240	100	100
Polarization mode	Single/double/quad polarization	Single/double polarization	Single/double/quad polarization
Elevation sweep angle	±20°	±11°	±20°
No. of imaging mode	Spotlight Mode, Strip Mode, Sweep Mode, Wave Mode, Ultrafine Mode, etc	Strip Mode, TOPS Mode, Wave Mode, Ultra-width Mode, etc	Strip Mode, Sweep Mode, Ultrafine Mode, etc
Resolution (m)	0.5–500	5–20	1–100
Imaging swath (km)	10–650	20–400	20–500
Lifespan (year)	8	7.5	7.5

extensibility.

- Propose an extensible ship taxonomy for SAR ATR, which can facilitate ship recognition and help us determine the relationship between a specific ship with other ships.

The remainder of the paper is organized as follows. Section 2 introduces GF-3 SAR data and AIS data used for dataset construction. Section 3 describes the automatic procedure of matchup data construction, including sea-land segmentation, multi-scale CFAR detection, SAR-AIS matchup, and false-alarm discrimination based on CNNs. Section 4 presents the intermediate results and explains the final dataset. Section 5 introduces an example application of FUSAR-Ship dataset for ship recognition. Section 6 concludes the paper.

2 GF-3 and AIS datasets

2.1 GF-3 SAR data

GF-3 is the first civilian fully-polarimetric C-band SAR satellite of China, which is designed mainly for ocean remote sensing applications. It was launched on August 10, 2016 and started operation on January 23, 2017. It is on a polar orbit with an average orbital altitude of about 755 km, with a repeated visit period of 29 days. It has advantages such as multiple imaging modes, large imaging width and long lifespan (see Table 1 for specifications).

Some key parameters of GF-3 SAR image are as follows [33]: cross-polarization isolation >37 dB, polarization amplitude imbalance <0.3 dB, polarization phase imbalance <6.9°, the noise equivalent sigma naught (NESZ) below –20 dB, the relative radiometric accuracy 0.7–0.8 dB, the absolute radiometric accuracy 1.3–1.4 dB, and the radiometric resolution <3.0 dB. In this study, we process 126 GF-3 images with azimuth resolution of 1.124 m and slant range resolution ranging from 1.700 m to 1.754 m. These scenes are distributed in the estuary, the middle and lower reaches of the Yangtze River, the Huanghai Sea, and the East China Sea. Figure 1 shows an example GF-3 scene data located in Wuhan, China.

2.2 AIS data

AIS is a real-time tracking system for global ships [34]. AIS can be classified as two categories, i.e., ground-based AIS and satellite-based AIS. A ground-based AIS receiver can only detect inshore ships within 50 km, while an AIS satellite can receive from any ships within 1000 km [34]. A ship continuously broadcasts three types of information, dynamic information, static information and voyage information. Dynamic information includes latitude, longitude, course, speed, and navigation status. Static information includes ship name, call sign, International Maritime Organization (IMO), Maritime Mobile Service

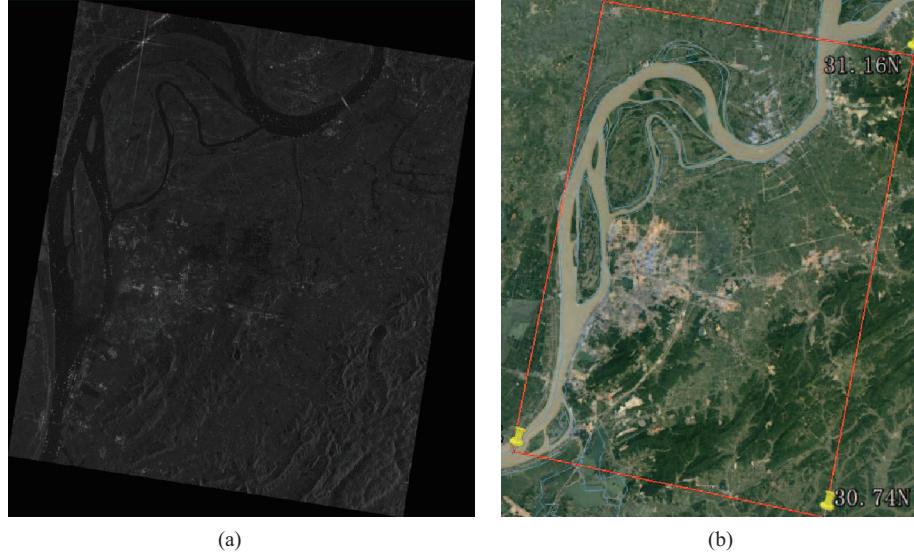


Figure 1 (Color online) (a) GF-3 example scene and (b) the corresponding aerial image.

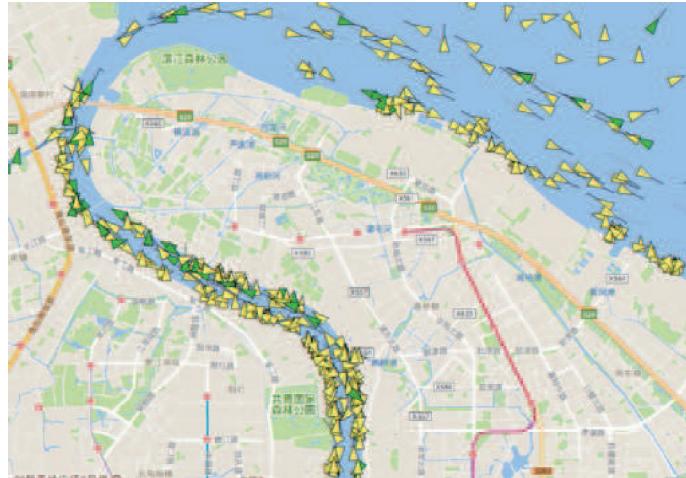


Figure 2 (Color online) Distribution of ships in the Huangpu River of Shanghai at 20:56 on November 18, 2018 as extracted from AIS dataset.

Identify (MMSI), length, width, and ship category. We can get its destination and estimated time of arrival (ETA) from voyage information.

With space-time coordinate information, we can co-register a SAR image with the corresponding AIS and further convert a latitude-longitude coordinate into a row-column image coordinate. After matching up, the category of a ship can be determined from static information, like fisher, cargo or bulk carrier, which enables us to accurately and reliably annotate a ship. Figure 2 shows an example distribution of ships in the Huangpu River of Shanghai [35] as obtained from the AIS data of 20:56 on November 18, 2018. We obtain the AIS data covering all space-time region of the 126 GF-3 SAR scenes.

3 Automatic matchup and annotation of SAR-AIS

3.1 Overview

Figure 3 shows the overall procedure of automatic matchup and annotation for SAR and AIS data. The mainstream of the framework comprises 5 modules as follows.

(1) Geo-coding. It maps a SAR image from the original range-azimuth coordinate to the latitude-

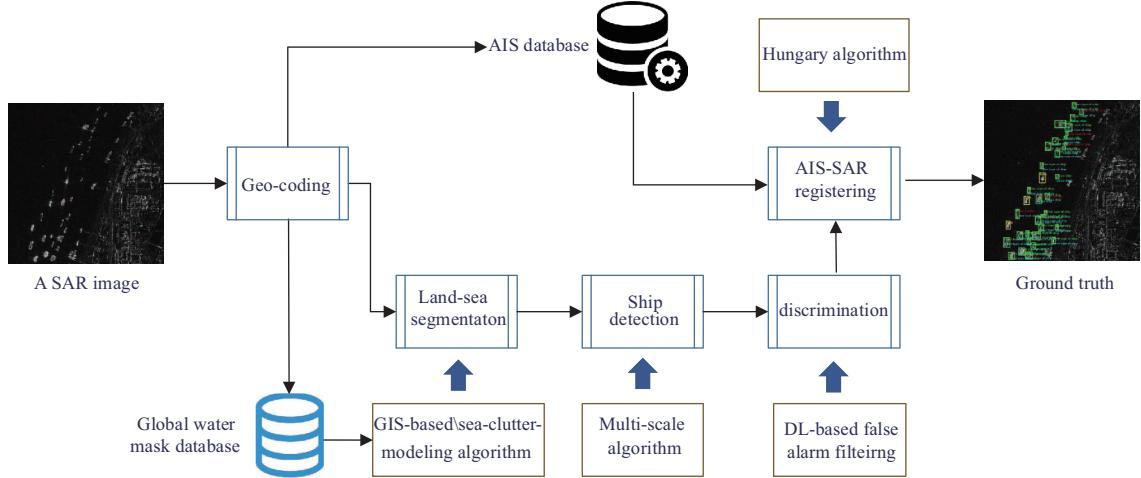
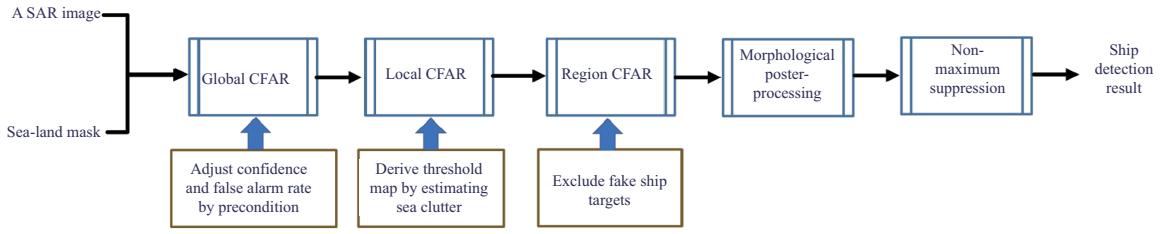


Figure 3 (Color online) The framework of SAR-AIS matchup and ship annotation.



longitude coordinate. This process can be done using commercial software, like PIE¹⁾ and ENVI²⁾.

(2) Sea-land segmentation. Sea-land segmentation is to remove land areas to accelerate subsequent stages and decrease false alarms. In this paper, we employ two sea-land segmentation algorithms, i.e., with or without the auxiliary watermask data. This module is detailed in Subsection 3.2.

(3) Ship detection. The previously proposed multi-scale CFAR [15,36] is adopted to detect ship targets in Figure 4. Furthermore, the registered AIS provides prior knowledge, including the number of ships of this area and the size of each ship, based on which it can be automatically adjusted for accurate estimation of sea clutter. AIS precondition can improve robustness, fit a statistical model of clutter distribution, smooth the disturbance of local outliers and eliminate sea clutter model errors under harsh environment or complex backgrounds, such as high density of ships or islands and high sea state of the ocean. Besides, the object profile of ships is used as prior knowledge, such as area, axis length ratio, and principal axis length, to exclude non-ship false alarms by the detection algorithm. It can be seen later from the results that the algorithm can effectively eliminate most of the false alarms through the three-scale modeling of sea clutter. Nevertheless, some inshore non-ship targets, such as wharves, waves, dams, islands or reefs, are difficult to eliminate not only because of their strong backscattering intensities but also their complex geometric shapes, which remains a challenge in ship detection [15].

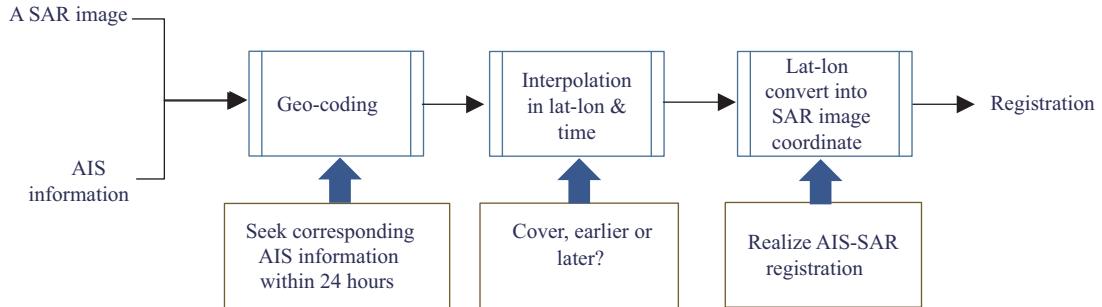
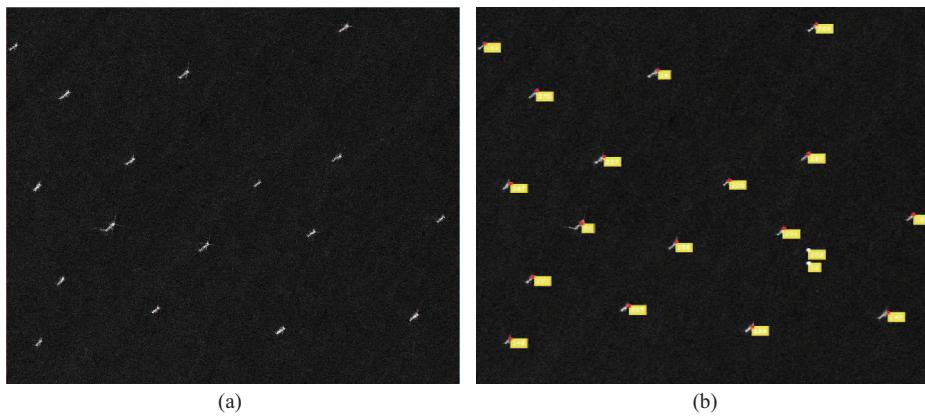
(4) False-alarm discrimination. As a simple intensity-based algorithm, CFAR is efficient but often results in some false alarms, like harbors, and islets or waves. In this study, we further design a deep convolutional neural network (DCNNs) to discriminate between ships and false alarms, which is detailed in Subsection 3.3.

(5) AIS-SAR co-registration. AIS-SAR co-registration comprises temporal matching, space matching and coordinate transformation (Figure 5). To obtain the optimal temporal-spatial matchup, different situations of AIS are carefully considered as follows.

Step 1. After geo-coding, the corresponding AIS of a SAR image can be retrieved according to longitude

1) <http://www.piesat.com.cn/alticle/PIE-GeoImage.html>.

2) <http://www.enviidl.com/>.

**Figure 5** (Color online) AIS-SAR co-registration.**Figure 6** (Color online) Result of AIS-SAR co-registration. (a) SAR image and (b) AIS reports co-registered with SAR image.

and latitude and time t of the SAR image. A temporal matchup window of 12 hour is used here, i.e., AIS information ranging from $t - T$ to $t + T$ is all retrieved, with the half-window time T set to 12 hours.

Step 2. The corresponding AIS information can be interpolated in the time domain to determine the position of a ship at time t . However, three different situations should be addressed respectively. Given the corresponding AIS information sequence $[A_{t_0}, A_{t_1}, A_{t_2}, \dots, A_{t_N}]$ where A_t and N denote the AIS information of a ship at time t and the number of AIS information of the ship from time t_0 to t_N , if $t_0 \leq t \leq t_N$, the AIS information of the ship at time t , i.e., A_t , can be derived using interpolation; if $t \leq t_0$, $A_t = A_{t_0}$; if $t_N \leq t$, $A_t = A_{t_N}$.

Step 3. The ship positions (longitude and latitude) obtained in Step 2 are converted to the SAR image coordinate.

Step 4. We get all kinds of categories of ships from the marine traffic website³⁾ according to the MMSI from AIS data.

However, there is a slight difference between the AIS-derived position and the SAR-detected position of the ship because of the Doppler shift [34] or latitude-longitude data error. A simple greedy algorithm [37, 38] is adopted to matchup between the results of Ship Detection and AIS. It is a variant of the Hungarian assignment algorithm where the cost is defined as the Euclidean distance between the matched-up pair. This algorithm works well for the case of open sea ships, about 94.49% accuracy can be reached. However, if there are near-shore ships (moored), owing to high density, the accuracy drops to 70%, which needs to be corrected manually. Figures 6 and 7 show an example of AIS-SAR matchup, i.e., red circles in Figure 6 are the positions of AIS matchup, while yellow boxes are the ship's MMSI. In Figure 7, the green bounding boxes are the results of ship detection of Subsection 3.1(3). As a result, we can get the results of AIS-SAR co-registration according to the results of Figures 6 and 7 with greedy algorithm.

3) <https://www.marinetraffic.com/en/ais/home/centerx:120.0/centery:30.9/zoom:7>.

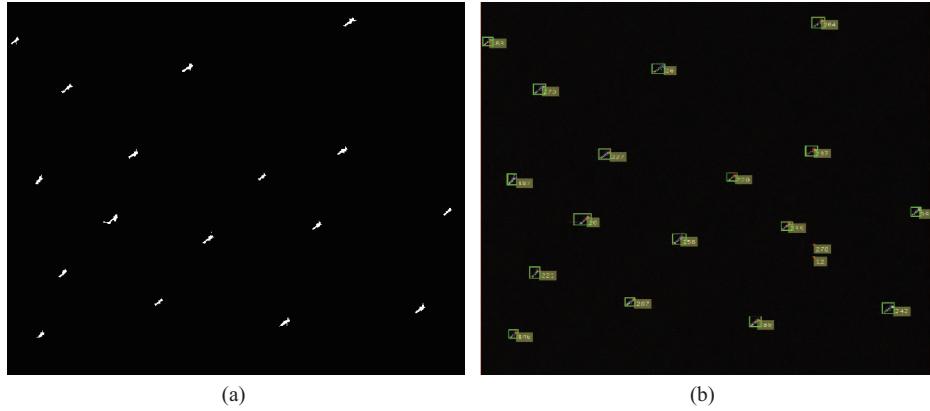


Figure 7 (Color online) SAR-AIS matchup via Hungarian algorithm. (a) Original SAR image; (b) SAR detected ships matched-up with AIS reports.

3.2 Sea-land segmentation

Sea-land segmentation can significantly reduce false alarm rate and improve efficiency of the follow-up processing. The flowchart is shown in Figure 8. It has two alternative approaches depending on whether the SAR image is geo-coded. For geo-coded SAR images, it employs the water-mask-aided sea-land segmentation by using the Global 250 m MODIS Water Mask⁴⁾ data as an initial segmentation. Subsequently, sea clutter distributions can be obtained from outer sea pixels which provide thresholds for pixel-wise CFAR binary segmentation in the coastal areas [15]. If the SAR image is not geo-coded or such a water mask is not available, the alternative approach is proposed as follows.

First, the logarithmic Gaussian mixture model (LGMM) is used to fit SAR image distributions [39]. Note that global outliers, i.e., the highest intensity 1% pixels, are excluded from model fitting. It automatically determines whether the mixture Gaussian distribution belongs to a single-peaked or bimodal-peaked distribution as [39]

$$L(\theta_k, \pi_k, z_{ik}|x) = \sum_{i=1}^n \sum_{k=1}^M Z_{ik} [\log \pi_k f_k(x_i|\theta_k)]. \quad (1)$$

While x represents the observation samples, $f(x)$ represents the probability density function (PDF) of x , θ_k represents the conditional probability, π_k represents the probability that the observed value belongs to the k^{th} class, and $z_i \in 1, 2, \dots, n$ is the implicit variable.

As shown in Figure 9(a), if it is a multi-peaked distribution, the image is determined as a mixed image, and clustering is conducted as sea-land segmentation according to the parameters of LGMM. It provides the initial sea mask M1. However, if it is single-peaked distribution, as in Figure 9(b), it implies that the image is an all-sea or all-land, which does not need to be segmented.

Next, Otsu method [35] is performed to obtain a second type of segmentation mask M2 which maximizes inter-class variance in the logarithmic domain and complex domain respectively. Last, on top of the Otsu result, the C-V model [40] is further conducted to adjust the M2 mask and finally obtain the third mask M3. Finally, a majority voting mechanism is applied to merge M1, M2 and M3 to reach the final segmentation mask.

The SAR image is classified as foreground and background where the foreground indicates the potential land area, while background means potential water area. In the binary image, morphological post-processing is used to further refine the segmentation result, which consists of the following operations:

- Hole filling and closing. Small-scale isolated sea and ship-size regions are eliminated by hole filling and morphological closing;
- Opening and dilation. The mask is further refined by removing the isolated pixels by opening and dilation operations, which achieves the desired segmentation result;

4) <http://glcf.umd.edu/data/watermask/>.

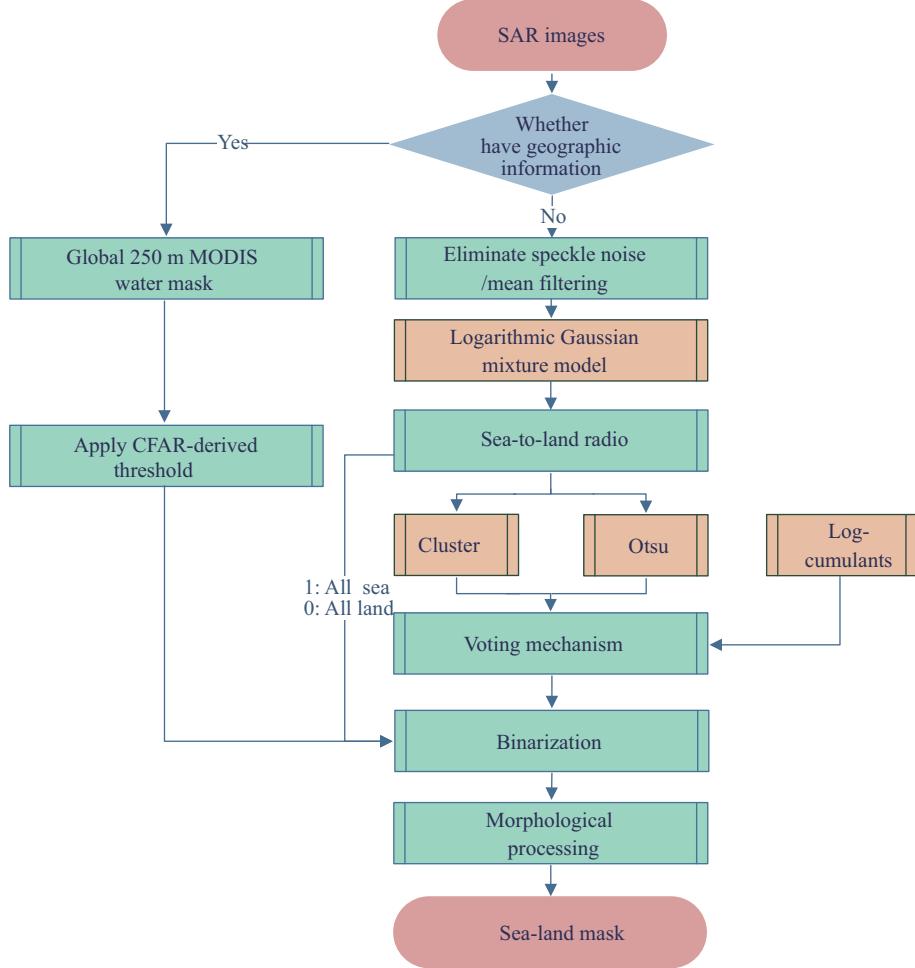


Figure 8 (Color online) The flowchart of sea-land segmentation.

- Mean filtering. We adopt median filtering and sigma filtering. In addition, Sobel, Robert and Laplace edge detection can effectively detect most coastlines.

In Figure 10, SAR images are presented in the top row, and the segmentation results are shown in the second row, while the difference between the ground truths and segmented results is in the bottom row. Green means the truth segmented area, red means the missing sea area, and yellow means the missing land. Experimental results demonstrate that the proposed method is robust even for narrow and curved rivers, isolated islands, coastal contour and other complex backgrounds.

3.3 False-alarm discrimination

After sea-land segmentation, multi-CFAR ship detection is conducted and an example result is shown in Figure 11(a), while Figure 11(b) is the result after filtering out false alarms. It can be seen that many nearshore false alarms should be further removed. It is hard to design handcrafted features to discriminate ships from false alarms. DCNNs can abstract multi-level features automatically by end-to-end training [28]. We first manually label some false alarms as negative samples while the matchup ships as positive samples, which can be used to train a binary classifier to discriminate ship targets from false alarms.

In this study, the false alarms are further categorized into 6 types: ship-like strong false alarms, bridges, coastal lands & islands, sea clutter waves, random sea and land (as shown in Figure 12 from the second row to bottom). Thus, a 7-category CNN is designed to classify these false-alarm maritime targets and ship targets.

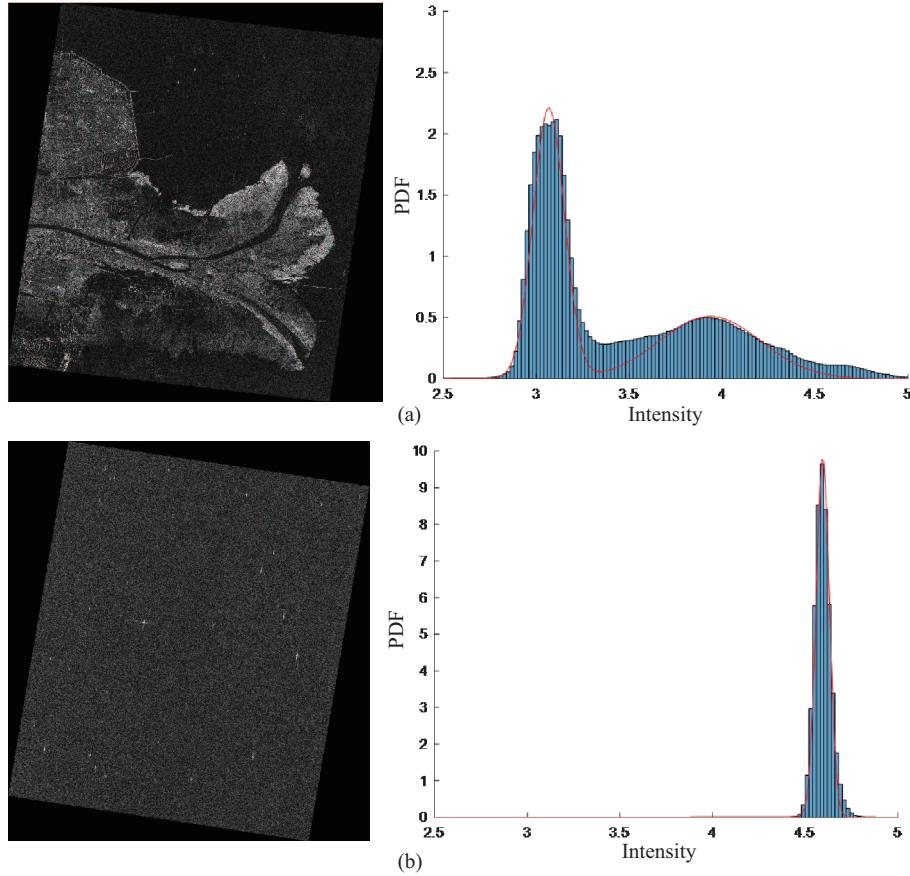


Figure 9 (Color online) The experimental results of LGMM. (a) Bimodal-peak vs. (b) single-peak.

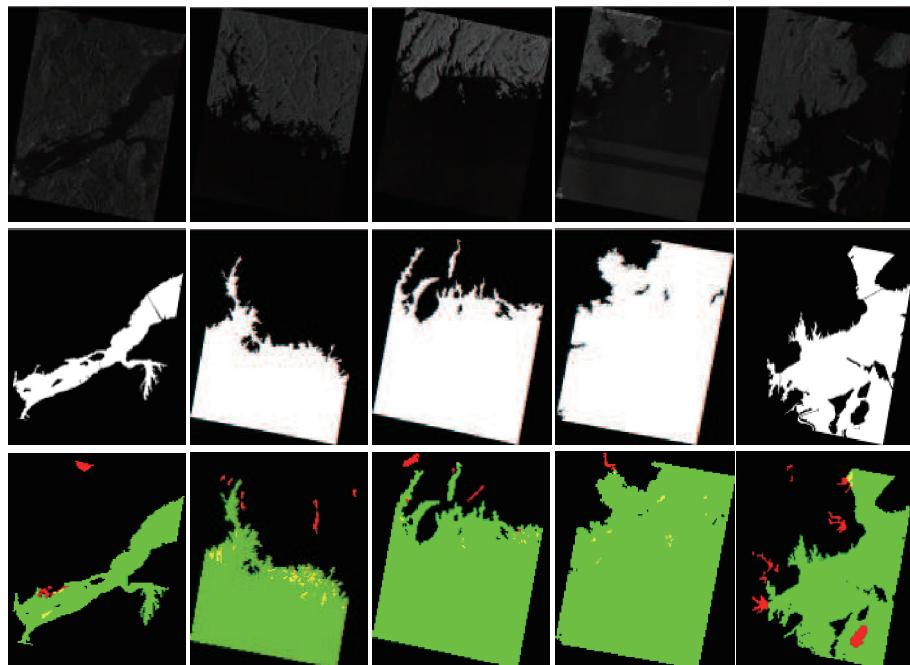


Figure 10 (Color online) The results of sea-land segmentation.

Table 2 shows the training dataset and validation dataset of the 7-category CNN. Figure 13 shows the structure of discrimination CNN. It employs the conventional iterating convolution and max pooling

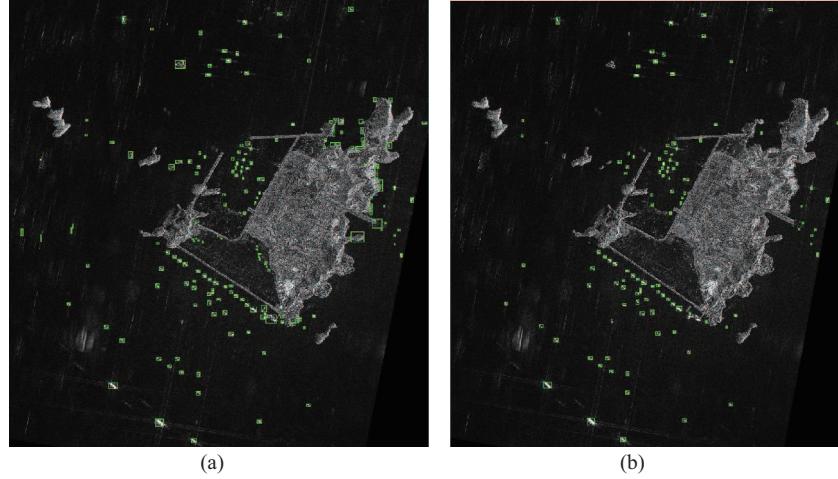


Figure 11 (Color online) Ship detection result of multi-scale CFAR algorithm (detailed information is shown in Table S1). (a) Before and (b) after false alarm discrimination.

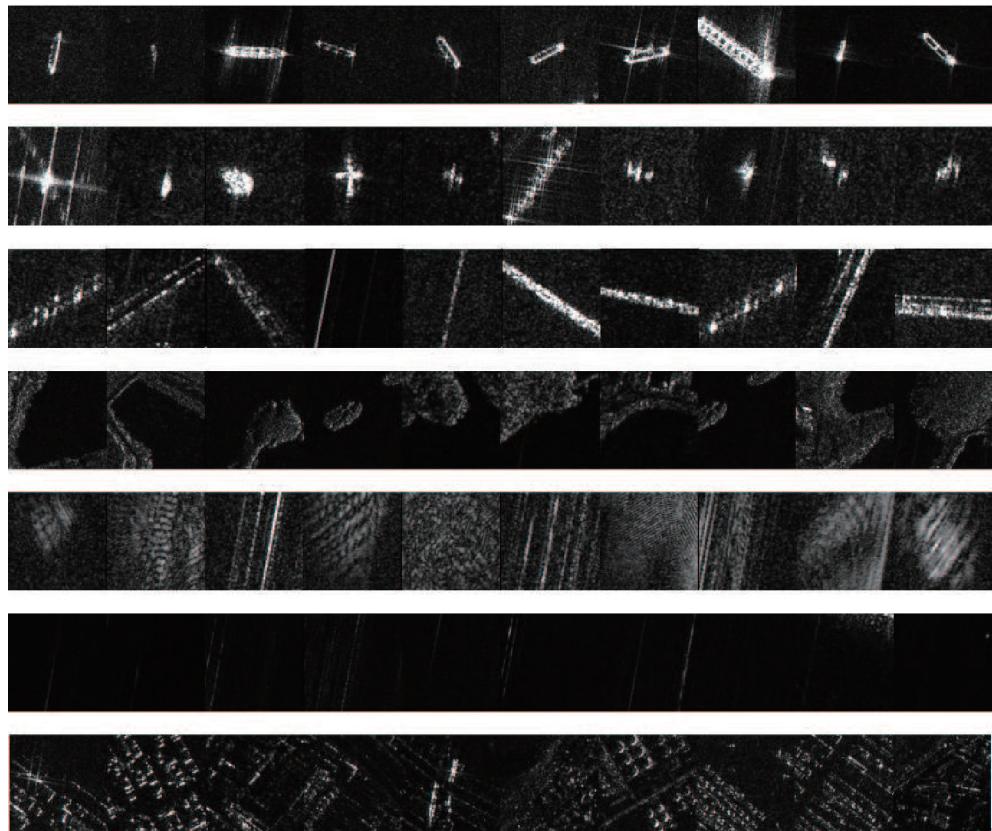


Figure 12 FUSAR-Ship of marine objects. From top to bottom, they are ships, strong scatterers, bridges, coastal lands & islands, sea clutter waves, random sea, and land sample patches, respectively.

structure for feature extraction [27]. The experiment setups are as shown in Table S2.

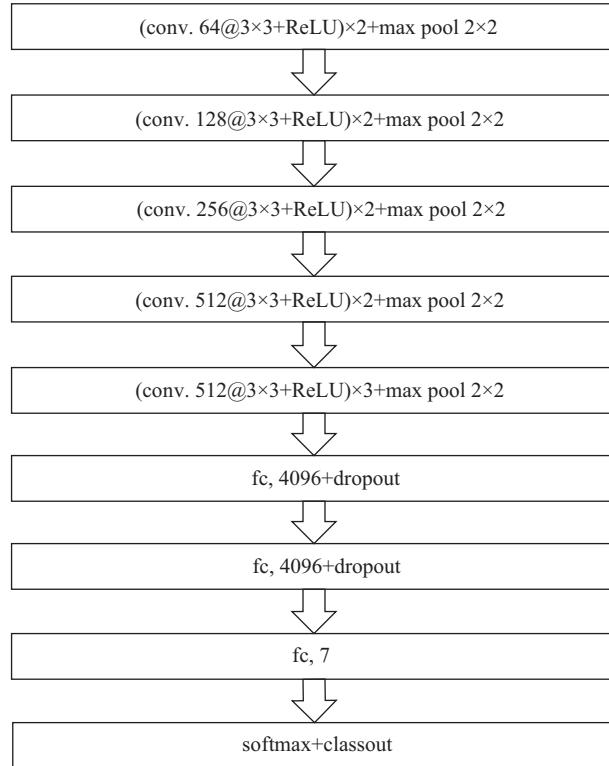
Table 3 is the confusion matrix of classifying 7 maritime targets. The ship miss rate is 4.7%, and the average accuracy is 96.15%.

3.4 Case study

Figure 14 shows the process of dataset construction algorithm. The memory size of a GF-3 SAR image such as the one in Figure 14 is about 20 GB, while the processing time is approximately 50 s on a regular

Table 2 Maritime targets of the FUSAR-Ship dataset

	Ships	Strong scatterers	Bridges & coastlines	Coastal lands & islands	Sea clutter waves	Sea patches	Land patches
Training data	1296	229	1023	707	1377	1250	1137
Validation data	555	128	438	303	590	535	487
Dataset	1851	427	1461	1010	1967	1785	1624

**Figure 13** The structure of 7-category CNN.**Table 3** The confusion matrix of marine objects classification

Category	Ships	Strong scatterers	Bridges & coastlines	Coastal lands & islands	Sea clutter waves	Random sea	Random land
Ships	529	15	0	0	1	0	0
Strong scatterers	14	103	0	0	0	0	2
Bridges & coastlines	3	1	428	3	1	0	0
Coastal lands & islands	4	3	6	287	4	1	3
Sea clutter waves	2	6	0	3	575	2	0
Random sea	0	0	1	1	4	528	13
Random land	3	0	3	9	5	2	469
Sum	555	128	438	303	590	535	487
Accuracy (%)	95.32	80.47	97.72	94.72	97.48	98.69	96.30
Overall accuracy = 96.15%							

Core i5-4590 CPU. Figure 14(a) is a test GF-3 SAR image, Figures 14(b)–(e) are the experimental results of Subsections 3.1(3), 3.1(5), 3.2, 3.3. Figure 14 is the extraction of ship target profile, and Figure 14(g) is the final detection result from Section 3. In addition, there are two red bounding boxes in Figure 14(g), in which Figures 14(h) and (i) are the enlarged figure. Moreover, the global and local information is combined to smooth the influence of local outliers. The experimental results demonstrate that the dataset construction above has a good performance, with higher robustness, higher adaptively

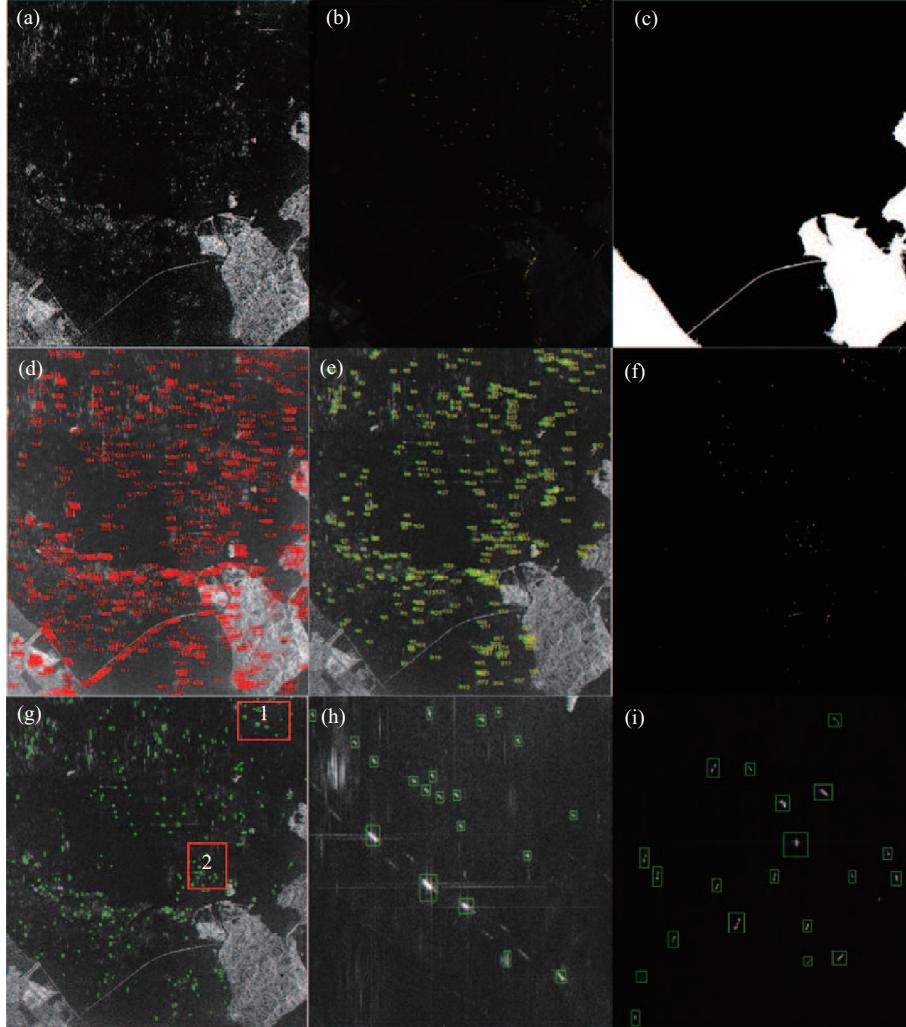


Figure 14 (Color online) The results of one example image (from left to right, top to bottom). (a) GF-3 SAR images, (b) AIS-SAR co-registration, (c) sea-land segmentation, (d) ship detection, (e) false-alarm discrimination, (f) ship target contour extraction, (g) final result of ship detection, (h) local amplification result of (g)-1, (i) local amplification result of (g)-2.

and lower computational complexity.

4 The FUSAR-Ship dataset

4.1 Taxonomy and data structure

FUSAR-Ship is intended as an open benchmark dataset for the purpose of ship SAR ATR algorithm development, evaluation and benchmarking. An extensible SAR ship taxonomy is first proposed in Figure 15. It has 15 ship categories, 98 ship subcategories, consisted of more than 5000 high-resolution samples cropped from a total of 126 GF-3 scenes in Figures 12 and 16. The root node is Maritime Object which is split into two branches, namely Ship and NonShip. The Ship node comprises nearly all categories of ships, e.g., Cargo, Tanker, and Fisher. From Figure 15, some of these nodes have subcategory nodes, e.g., the Cargo node is composed of Bulk Carrier, General Cargo, Container and Other Cargo nodes. On the other hand, the NonShip node has three childnodes, i.e., Land, Sea and Strong Scatterers (e.g., as buoy, windmill, and sea ranch). For instance, the Land node includes typical nature and manmade nearshore architectures including Bridges & Coastlines, Coastal Lands & Islands and Random Land Surfaces. In sum, the maritime object taxonomy is composed of four levels. The proposed multi-scale

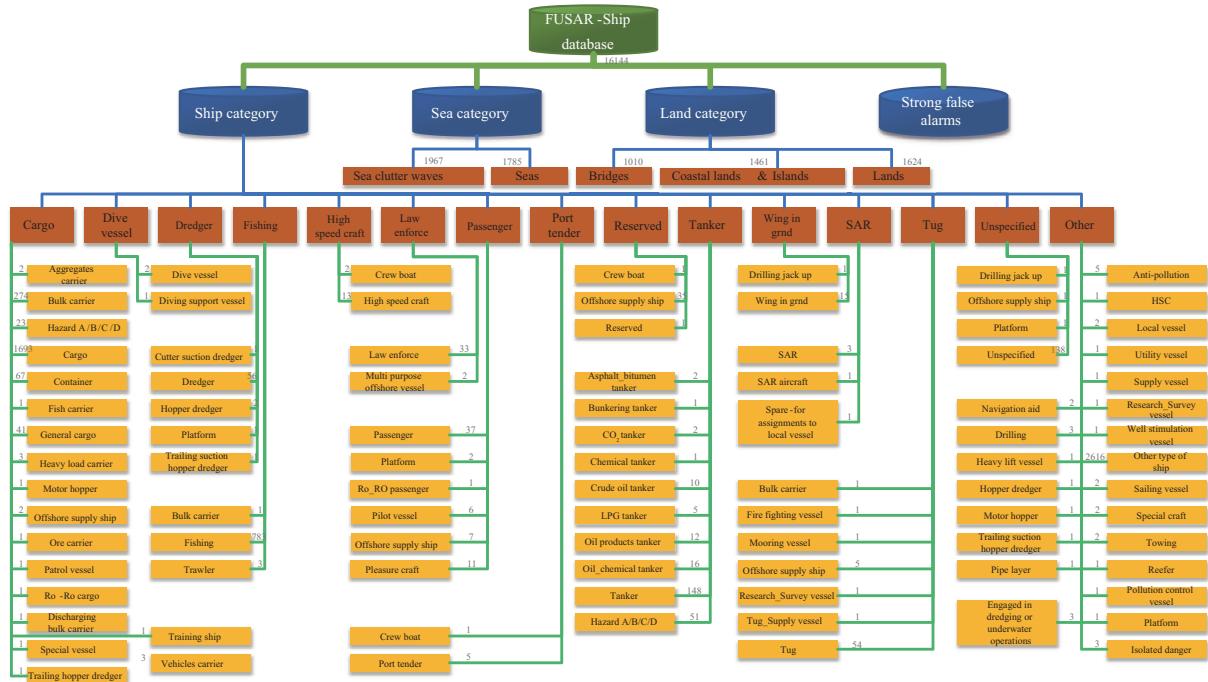


Figure 15 (Color online) The taxonomy of SAR ships.

CFAR detector can effectively detect nearly all maritime objects which matches with the root node. Then, a binary classifier is able to differentiate ships from nonships, which matches with the second level nodes. Finally, ships can be coarsely classified as the third level nodes, like cargos, tankers, and fishers. Ships can be further classified into subcategories, like bulk carriers, general cargos, and containers. In this way, this taxonomy can adapt to different application scenarios, like maritime object detection, ship detection, coarse ship classification, and precise ship classification.

4.2 Dataset composition

GF-3 FUSAR-Ship⁵⁾ contains marine targets chips of many categories, which provides strong support for high-resolution and multi-dimensional SAR marine imagery interpretation.

Example GF-3 SAR images are shown here (the first row images of Figure 10, Figure 14(a)), as well as the AIS matching images (Figure 14(b)), sea-land segmented masks (Figure 10), ship profiles (Figure 14(f)) and ship detection results (Figures 14(g)–(i)). Example AIS coordinates, ship detecting coordinates, and ship-AIS index matching results are shown in Table 4. The tables contain ships' MMSI, position, category, length and width, speed, and navigation status. “DH/DV” means “double aperture HH/VV polarization”, “pos_y, pos_x” represent the coordinates of pixels of AIS data point, while “ship_y, ship_x” represent the coordinates of the results of ship detection in the SAR images. The dataset contains 10125 marine target chips. Mentioned that the dataset includes 126 GF-3 images with annotation files and the ship patches, in which the GF-3 images and the patches are all amplitude data. The imaging mode of those 126 images is ultrafine strip-map (UFS) mode.

Compared with other datasets, FUSAR-Ship has the following advantages.

- High-resolution. It uses high resolution space-borne high-resolution GF-3 SAR images to build the dataset.
- Consistency. The annotation of the images is consistent, in terms of the definition of the classes, how bounding boxes are placed, and how viewpoints and truncation are defined. And we crop chips around the center points to a fixed size of 512×512 pixels.

5) The GF-3 FUSAR-Ship dataset is available at www.emwlab.fudan.edu.cn/resources.

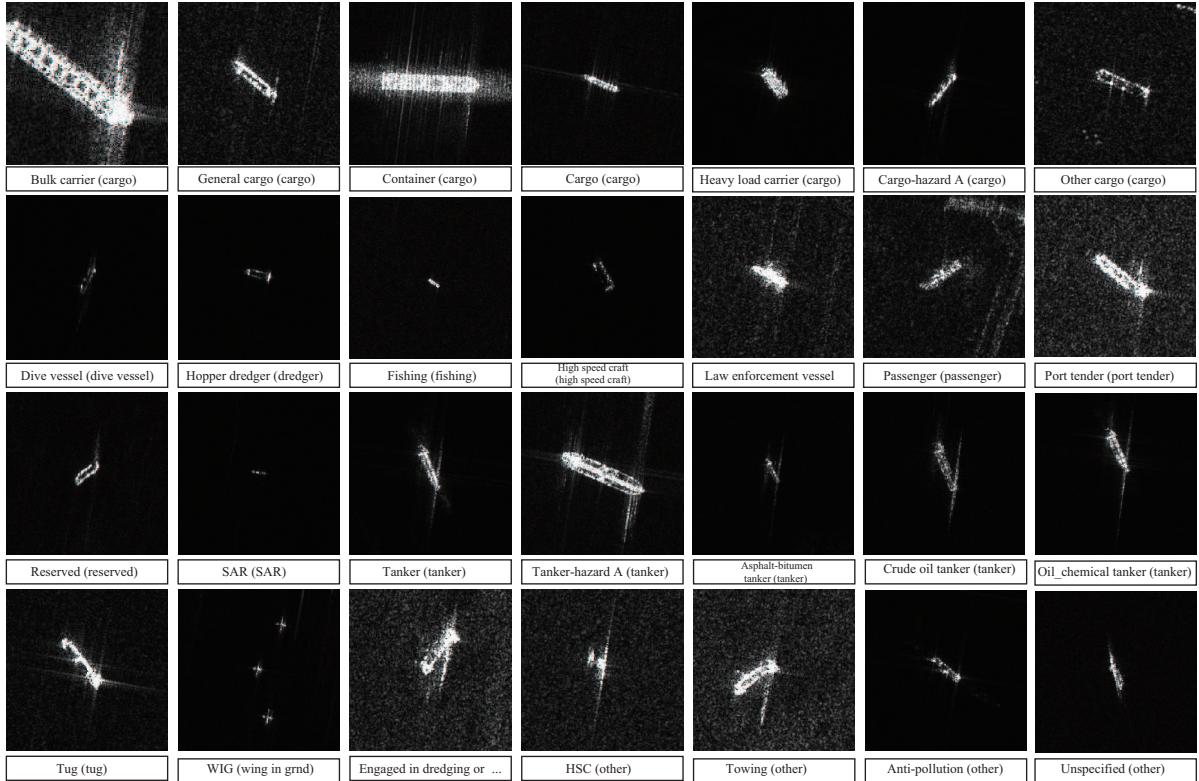


Figure 16 Different categories of ships in FUSAR-Ship.

- **Diversity.** The database includes 6 maritime targets and 98 ship categories, and all maritime object instances are labeled; and it covers various scenarios such as near-shore regions, islands and dynamic sea conditions.
- **Extensibility.** Based on the annotation framework and the SAR ship taxonomy, the dataset can be easily expanded by running the automatic procedure on more SAR data as they are becoming available in the future.
- **Large-scale.** FUSAR-Ship has a total of 15 ship categories, 98 ship subcategories, consisted of more than 5000 high-resolution ship samples and more non-ship samples cropped from a total of 126 GF-3 scenes covering various scenarios.

FUSAR-Ship can be used for target detection and recognition, wake tracking, change detection, semantic segmentation, training detector and related applications, which provides strong support for studies on maritime traffic monitoring, ship search and rescue, fisheries management, and maritime situational awareness.

5 Preliminary ship classification experiment

A simple 8-type classification experiment is carried out on the FUSAR-Ship dataset. The eight universal types to classify include bulk carriers, general cargos, containers, other cargos, fishing, tankers, other ships and false alarms as shown in Figure 17. From top to bottom, the figures are bulk carriers, general cargos, containers, other cargos, fishing, tankers, other ships, and false alarms, respectively. The first column is the optical images, and the rest are SAR image chips. A convolutional neural network is designed to automatically extract target features and describe the targets obtained from GF-3 SAR images in a compact and robust manner [27].

Figure 18 shows the architecture of the network for ship classification. The previous few layers of the deep convolution network are composed of alternating convolution layers and max pooling layers, followed by several fully-connected layers [41]. Besides, dropout layers are inserted into full connected layers to

Table 4 The example of AIS ship information

Attribute	CHANGKUN7	HONGFAN6	SHUIWU	XINHAIHUA	ZHESHENGYU07817	HANGONGJIAO4
CallSign	BHNR2	0	BINW	0	0	0
IMO	0	0	0	413441210	0	0
MMSI	413363550	413784469	413370630	413441210	900307817	413824589
ShipTypeEN	Tanker	Cargo ship	LawEnforceVessel	Passenger ship	WIG	OtherTypeOfShip
NavStatusEN	Under way using engine	Under way using engine	Moored	Under way using engine	Unknown	Under way
Length	104	78	30	38	50	25
Width	15	14	4	7	6	7
Draught	5.2	3.9	1.6	1.5	0	0
Heading	511	249	511	298	511	178.2
Course	241.4	249.6	293.4	80	268	16.1
Speed	8.9	5.4	0	0	1.9	0.1
Lon	117.40E	117.36E	121.46E	122.18E	123.18E	118.3E
Lat	30.47N	30.46N	31.27N	29.56N	31.20N	31.11N
Dest	AN QING	SHANGHAI	SHENJIAM	—	—	—
UnixTime	1487200911	1487221804	1487086914	1487047384	1487080221	1487250424
Lon_d	117.66848	117.605323	121.771713	122.302167	123.303615	118.06088
Lat_d	30.792795	30.772788	31.466342	29.944667	31.339332	31.193668
pos_y	1861	10527	6100	6143	20705	1707
pos_x	13030	4070	20916	20144	21913	6004
ship_y	1850	10530	6094	6104	16636	1614
ship_x	13040	4084	20814	20164	21440	6000
Polarization mode	DH	DH	DH	DV	DV	DV
Incident angle	27.28	27.28	27.28	40.36	40.36	40.36
Resolution	1.726×1.124	1.726×1.124	1.726×1.124	1.736×1.124	1.736×1.124	1.736×1.124

Table 5 The confusion matrix of marine objects classification for ship discrimination

Category	Bulk carrier	General cargo	Container	Other cargo	False alarm	Fishing	Other ship	Tanker
Bulk carrier	240	0	0	0	0	0	2	1
General cargo	0	120	0	5	1	5	25	16
Container	0	0	65	4	12	4	9	0
Other cargo	10	5	3	166	18	3	37	7
False alarm	0	1	5	11	2406	8	21	2
Fishing	0	1	0	1	1	10	3	0
Other ship	82	51	1	116	102	76	519	35
Tanker	0	0	0	1	2	1	0	52
Accuracy (%)	72.28	67.42	87.84	54.62	4.65	9.35	84.25	46.18

avoid overfitting. Afterward, softmax nonlinear function is applied to the nodes of the final output layer. The experiment setups are as shown in Table S2.

The dataset is randomly divided into 70% training data and 30% testing data. Table 5 shows the confusion matrix of training dataset. The average training accuracy is 87.83%. The accuracy of bulk carriers, containers, other ships and false alarms are all more than 70%, but tankers and fishing are much lower, primarily owing to the unbalanced population distribution.

As shown in Figure 19, the trained classification neural network model is applied to test 4 randomly selected GF-3 SAR images. In Figure 19, the left panel of each row is GF-3 SAR scene images, and the right panel is the experimental results of testing GF-3 SAR scene images by trained ship classification networks. Note that the recognition accuracy from top to bottom is 77.78%, 80%, 81.67%, and 77.42%, respectively, with an averaged test accuracy of 79.18%. It suggests that a certain degree of overfitting has

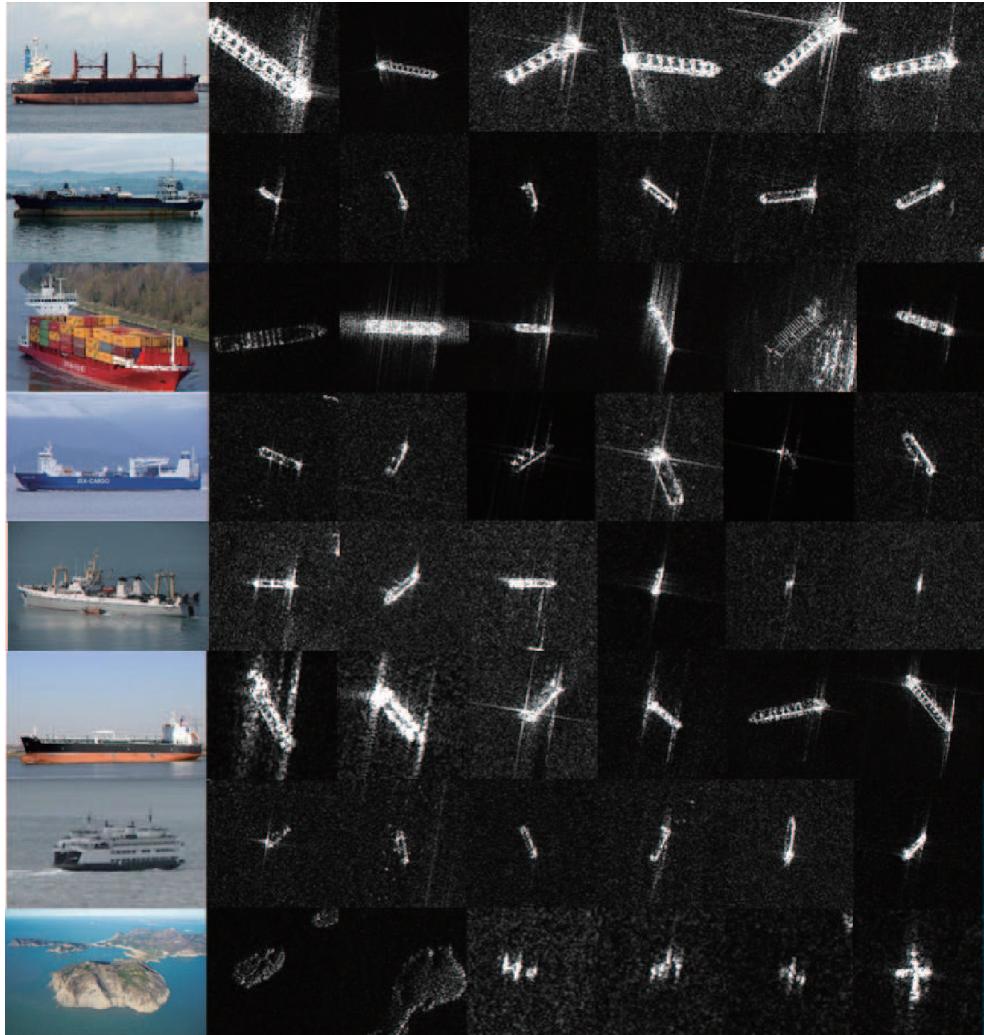


Figure 17 (Color online) The distribution of FUSAR-Ship datasets of ships. From top to bottom, they are bulk carriers, general cargos, containers, other cargos, fishing, tankers, other ships, and false alarms, respectively.

occurred during the training which deserves further investigation. The reasons for the overfitting might be the following. First, there are several different types of ships in one chip, while we only label this chip as the central ship type. So when test this chip, it will be misclassified into other categories. Second, the SAR-AIS registration is not necessarily correct, so there are several chips that may be assigned wrong category.

6 Conclusion

This paper presents a high-resolution GF-3 SAR marine target dataset automatically constructed by the proposed ship detection and SAR-AIS matchup procedure. The proposed automatic ship detection method consists of sea-land segmentation with or without auxiliary global water mask data and a multi-scale CFAR robust detector. Results demonstrate good performance on detecting offshore ships in complex coastal environments.

The proposed taxonomy of FUSAR-Ship is a hierarchical extensible classification tree structure. It contains ship and non-ship two main types. The Ship node contains 15 primary ship categories and then is divided into 98 sub-categories.

By running the procedure on a total of 126 GF-3 SAR images, we have established a large-scale FUSAR-Ship dataset covering a large variety of sea, land, coast, river and island scenarios.

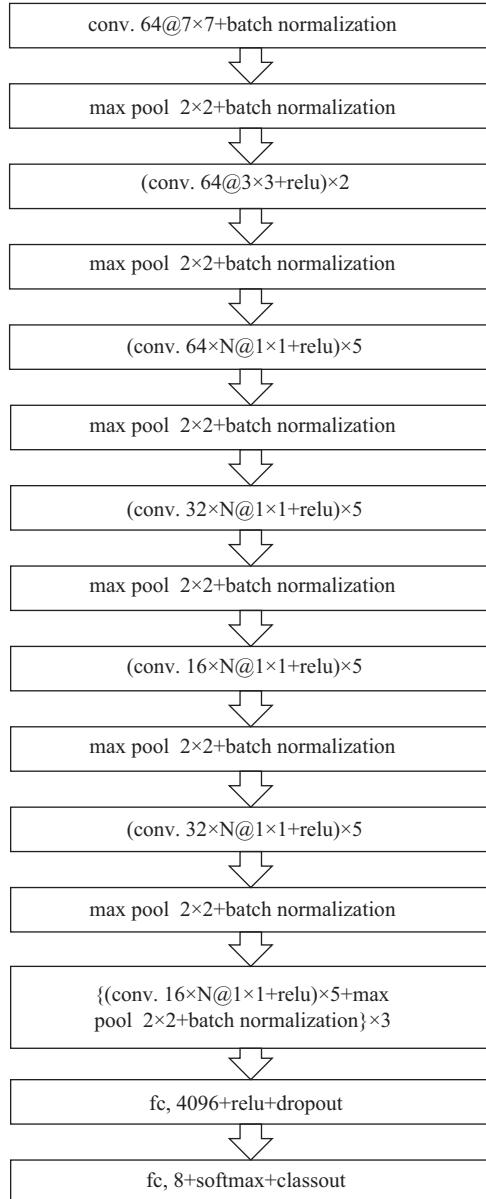


Figure 18 The structure of network layers of ship discrimination.

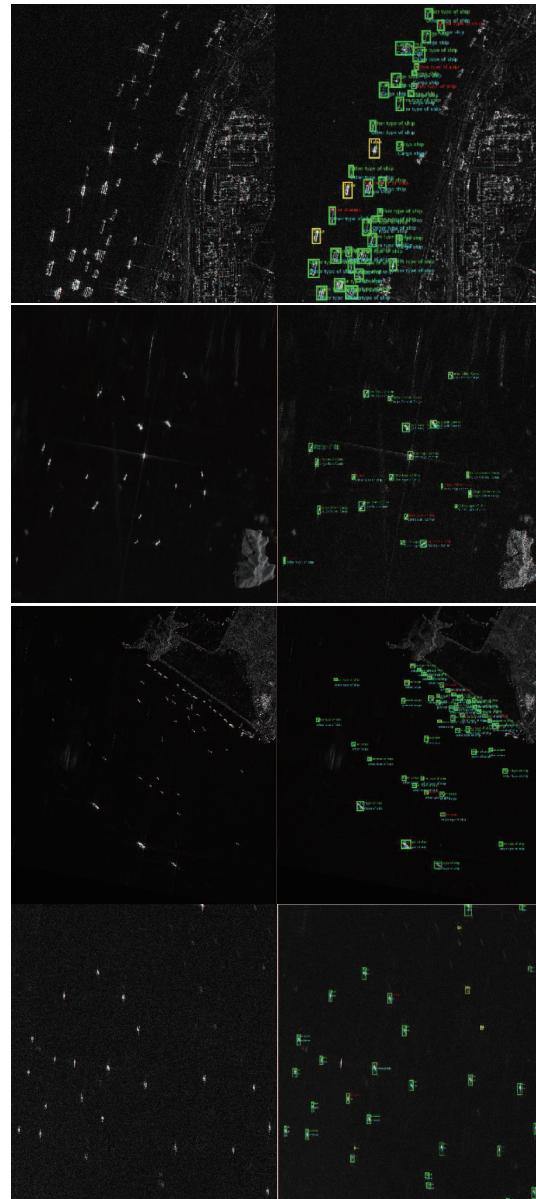


Figure 19 (Color online) Testing GF-3 SAR scenes by the trained CNN (detailed information is shown in Table S1). From top to bottom, there are four example scenes, where the left and right panel of each row is the original SAR image and the recognized ship targets respectively.

Then a preliminary experiment on 8-type ship classification experiment was carried out on the FUSAR-Ship dataset using regular CNNs which achieves on average 79% accuracy. Further studies are required to improve the performance in terms of accuracy, robustness and generalization capability. Moreover, other tasks and applications can be developed and evaluated on the FUSAR-Ship dataset.

Acknowledgements This work was supported in part by National Key R&D Program of China (Grant No. 2017YFB0502-703) and National Natural Science Foundation of China (Grant Nos. 61991422, 61822107).

Supporting information Tables S1 and S2. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

- 1 Xu F, Jin Y Q, Moreira A. A preliminary study on SAR advanced information retrieval and scene reconstruction. *IEEE Geosci Remote Sens Lett*, 2016, 13: 1443–1447
- 2 Paes R L, Lorenzzetti J A, Gherardi D F M. Ship detection using TerraSAR-X images in the campos basin (Brazil). *IEEE Geosci Remote Sens Lett*, 2010, 7: 545–548
- 3 Liu Y, Yao L, Xiong W, et al. Fusion detection of ship targets in low resolution multi-spectral images. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, 2016. 6545–6548
- 4 Shuai T, Sun K, Wu X, et al. A ship target automatic detection method for high-resolution remote sensing. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, 2016. 1258–1261
- 5 Wang W, Fu Y, Dong F, et al. Semantic segmentation of remote sensing ship image via a convolutional neural networks model. *IET Image Process*, 2019, 55: 1016–1022
- 6 Xu F, Wang H, Song Q, et al. Intelligent ship recognition from synthetic aperture radar images. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium, Valencia, 2018. 4387–4390
- 7 An Q, Pan Z, You H. Ship detection in Gaofen-3 SAR images based on sea clutter distribution analysis and deep convolutional neural network. *Sensors*, 2018, 18: 334
- 8 Ma M, Chen J, Liu W, et al. Ship classification and detection based on CNN using GF-3 SAR images. *Remote Sens*, 2018, 10: 2043
- 9 Pan Z, Liu L, Qiu X, et al. Fast vessel detection in Gaofen-3 SAR images with ultrafine strip-map mode. *Sensors*, 2017, 17: 1578
- 10 Wang Y, Wang C, Zhang H, et al. Automatic ship detection based on retinanet using multi-resolution Gaofen-3 imagery. *Remote Sens*, 2019, 11: 531
- 11 Guo Q, Wang H, Kang L, et al. Aircraft target detection from spaceborne SAR image. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2019. 1168–1171
- 12 Fu S, Xu F, Jin Y-Q. Reciprocal translation between SAR and optical remote sensing images with cascaded-residual adversarial networks. 2019. ArXiv:1901.08236
- 13 Shao Z Y, He J H, Feng S S. Extraction of a target in sea clutter via signal decomposition. *Sci China Inf Sci*, 2020, 63: 129301
- 14 Long T, Liang Z N, Liu Q H. Advanced technology of high-resolution radar: target detection, tracking, imaging, and recognition. *Sci China Inf Sci*, 2019, 62: 040301
- 15 Ao W, Xu F, Li Y, et al. Detection and discrimination of ship targets in complex background from spaceborne ALOS-2 SAR images. *IEEE J Sel Top Appl Earth Observ Remote Sens*, 2018, 11: 536–550
- 16 Bandiera F, Orlando D, Ricci G. Advanced radar detection schemes under mismatched signal models. In: *Synthesis Lectures on Signal Processing*. Williston: Morgan & Claypool, 2009. 1–105
- 17 Liu G, Zhang Y S, Zheng X W, et al. A new method on inshore ship detection in high-resolution satellite images using shape and context information. *IEEE Geosci Remote Sens Lett*, 2014, 11: 617–621
- 18 Chen H, Wang Q, Shen Y. Decision tree support vector machine based on genetic algorithm for multi-class classification. *J Syst Eng Electron*, 2011, 22: 322–326
- 19 Deng J, Dong W, Socher R, et al. ImageNet: a large-scale hierarchical image database. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Miami, 2009. 248–255
- 20 Everingham M, van Gool L, Williams C K I, et al. The pascal visual object classes (VOC) challenge. *Int J Comput Vis*, 2010, 88: 303–338
- 21 Liu N, Cao Z, Cui Z, et al. Multi-scale proposal generation for ship detection in SAR images. *Remote Sens*, 2019, 11: 526
- 22 Xia G, Bai X, Ding J, et al. DOTA: a large-scale dataset for object detection in aerial images. *Computer science: computer vision and pattern recognition*. 2017. ArXiv:1711.10398
- 23 Zhu H, Chen X, Dai W, et al. Orientation robust object detection in aerial images using deep convolutional neural network. In: Proceedings of 2015 IEEE International Conference on Image Processing (ICIP), Quebec City, 2015
- 24 Cheng G, Han J, Zhou P, et al. Multi-class geospatial object detection and geographic image classification based on collection of part detectors. *ISPRS J Photogrammetry Remote Sens*, 2014, 98: 119–132
- 25 Cheng G, Han J. A survey on object detection in optical remote sensing images. *ISPRS J Photogrammetry Remote Sens*, 2016, 117: 11–28
- 26 Cheng G, Zhou P, Han J. Learning rotation-invariant convolutional neural networks for object detection in vhr optical remote sensing images. *IEEE Trans Geosci Remote Sens*, 2016, 54: 7405–7415
- 27 Chen S, Wang H, Xu F, et al. Target classification using the deep convolutional networks for SAR images. *IEEE Trans Geosci Remote Sens*, 2016, 54: 4806–4817
- 28 Cho J H, Park C G. Multiple feature aggregation using convolutional neural networks for SAR image-based automatic target recognition. *IEEE Geosci Remote Sens Lett*, 2018, 15: 1882–1886
- 29 Huang L, Liu B, Li B, et al. OpenSARShip: a dataset dedicated to sentinel-1 ship interpretation. *IEEE J Sel Top Appl Earth Observ Remote Sens*, 2018, 11: 195–208
- 30 Li B, Liu B, Huang L, et al. OpenSARShip 2.0: a large-volume dataset for deeper interpretation of ship targets in Sentinel-1 imagery. In: Proceedings of 2017 SAR in Big Data Era: Models, Methods and Applications (BIGSAR-DATA), Beijing, 2017. 1–5
- 31 Bao M, Zhang J, Meng J M, et al. Construction and feature analysis of high resolution SAR ship sample set (in

- Chinese). *Chin J Radio Sci*, 2019, 34
- 32 Wang Y, Wang C, Zhang H, et al. A SAR dataset of ship detection for deep learning under complex backgrounds. *Remote Sens*, 2019, 11: 765
- 33 Liu J, Zhang Q. Overview of Gaofen-3 satellite applications (in Chinese). *Satell Appl*, 2018, (06): 12–16
- 34 Song J, Oh K, Kim I, et al. Application of maritime AIS (automatic identification system) to ADS-B (automatic dependent surveillance-broadcast) transceiver. In: Proceedings of International Conference on Control, Automation and Systems, Gyeonggi-do, 2010. 2233–2237
- 35 Pan J, Zheng X, Sun L, et al. Image segmentation based on 2D OTSU and simplified swarm optimization. In: Proceedings of 2016 International Conference on Machine Learning and Cybernetics (ICMLC), 2016. 1026–1030
- 36 Hou X, Ao W, Xu F. End-to-end automatic ship detection and recognition in high-resolution Gaofen-3 spaceborne SAR images. In: Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2019. 9486–9489
- 37 Miller M L, Stone H S, Cox I J. Optimizing Murty's ranked assignment method. *IEEE Trans Aerosp Electron Syst*, 1997, 33: 851–862
- 38 Munkres J. Algorithms for the assignment and transportation problems. *J Soc Industrial Appl Math*, 1957, 5: 32–38
- 39 Yue J, Wang S. EM algorithm and its initialization in Gaussian mixture model based clustering. *Microcomput Inf*, 2006, 22: 244–246
- 40 He H, Lin Y, Chen F, et al. Inshore ship detection in remote sensing images via weighted pose voting. *IEEE Trans Geosci Remote Sens*, 2017, 55: 3091–3107
- 41 Xu F, Wang H, Jin Y Q. Deep learning as applied in SAR target recognition and terrain classification (in Chinese). *J Radars*, 2017, 6: 136–148