# OpenSARShip: A Dataset Dedicated to Sentinel-1 Ship Interpretation

Lanqing Huang , Bin Liu , *Member, IEEE*, Boying Li, Weiwei Guo, Wenhao Yu , Zenghui Zhang, *Member, IEEE*, and Wenxian Yu

*Abstract*—With the rapid growth of Sentinel-1 synthetic aperture radar (SAR) data, how to exploit Sentinel-1 imagery and achieve effective and robust marine surveillance are crucial problems. In this paper, we present the OpenSARShip, a dataset dedicated to Sentinel-1 ship interpretation. The OpenSARShip, providing 11 346 SAR ship chips integrated with automatic identification system messages, owes five essential properties: specificity, large scale, diversity, reliability, and public availability. These properties make sure that the OpenSARShip achieves its objectives. The first is to provide researchers a benchmark dataset to develop applicable and adaptive ship interpretation algorithms and push the performance ceilings of data analysis. The other is to provide a dataset for performing application-oriented quality assessment for Sentinel-1 imagery, which can boost their applications in a targeting way. The construction and the organization of the OpenSARShip are discussed, which show the inside of the dataset and ensure the essential properties. The elaborate geometric and scattering analyses, the benchmark for classification, and the imagery applicability assessment by using the OpenSARShip all demonstrate the applicability and potential of the dataset.

*Index Terms*—OpenSARShip dataset, ship interpretation, Sentinel-1, synthetic aperture radar (SAR).

## I. INTRODUCTION

FOCUSING on marine surveillance, the Sentinel-1 satellites which are designed by the European Space Agency are regarded as well-suited sensors for their capabilities of wide coverage [1]. The Sentinel-1 satellites operate C-band synthetic aperture radars (SARs) and provide continuous all-weather day/night imagery at medium to high resolution.

The data are publicly available and provide sufficient resources for marine application [2], such as sea ice detection [3], sea wind retrieval [4], [5], ship detection, and classification [1], [2], [6]–[9]. In order to achieve more sophisticated and robust algorithms for marine surveillance, efficiently exploiting such enormous explosion of data is essential. However, exactly how such data can be indexed, organized into a dataset, and

utilized for the specific application is a crucial problem yet to be solved.

Several applicable datasets were constructed in the field of remote sensing. For optical imagery, the well-known UCMerced Land-Use dataset [10] contains 2100 aerial scene images acquired from different US locations and manually labeled as 21 land categories [11]. International Society for Photogrammetry and Remote Sensing labeling contest dataset [12], a standard dataset for evaluating object extraction, provides 71 image patches from two scenes with 6 manually labeled categories. For hyperspectral imagery, several publicly available datasets, such as Indian Pines Scene and Pavia Centre and University Scene, are widely used [13], [14]. For interferometric SAR (InSAR) imagery, Western North America interferometric SAR (WinSAR) Consortium [15] acquires SAR imagery aiming to promote the development and use of InSAR technology. Moving and stationary target dataset [16] composed of ten types of ground targets with different aspect angles, depression angles, and target configurations is widely used [17]–[19]; however, this dataset is applicable for ground object extraction and automatic target recognition [18], [19] rather than for SAR marine surveillance.

In this paper, aiming at marine surveillance, we introduce a new SAR dataset called *OpenSARShip* [20], which is collected from 41 Sentinel-1 images with various environmental conditions, providing 11 346 SAR ship chips integrated with automatic identification system (AIS) messages. The OpenSARShip owes five essential properties: specificity, large scale, diversity, reliability, and public availability. The construction of the OpenSARShip follows four main steps. First, Sentinel-1 SAR data are collected and preprocessed. Next, an online-coordination labeling tool (OCLT) is designed to achieve semi-automatic labeling. Then, the SAR ships are accurately integrated with AIS messages. Finally, postprocessing is done to each ship chip so as to guarantee the reliability and good quality. The OpenSARShip is accurately verified in every step of construction and demonstrated in several data formats so as to satisfy different requirements of users. Furthermore, based on the geometric and scattering properties of the OpenSARShip, we analyze the factors which adversely affect ship detection and classification performance, present a promising benchmark for designing advanced algorithms, and exploit the applicable potential of Sentinel-1 imagery. Preliminary exploitations are achieved to illustrate the application of the OpenSARShip for marine surveillance.

The main contribution of this paper is that via organizing and exploiting rapidly growing Sentinel-1 SAR imagery, we

Fig. 1.    Flow of constructing OpenSARShip.



Fig. 2.    Semi-automatic extraction of SAR ship signatures.



Fig. 3.    Flow of integration of SAR ships and AIS messages.
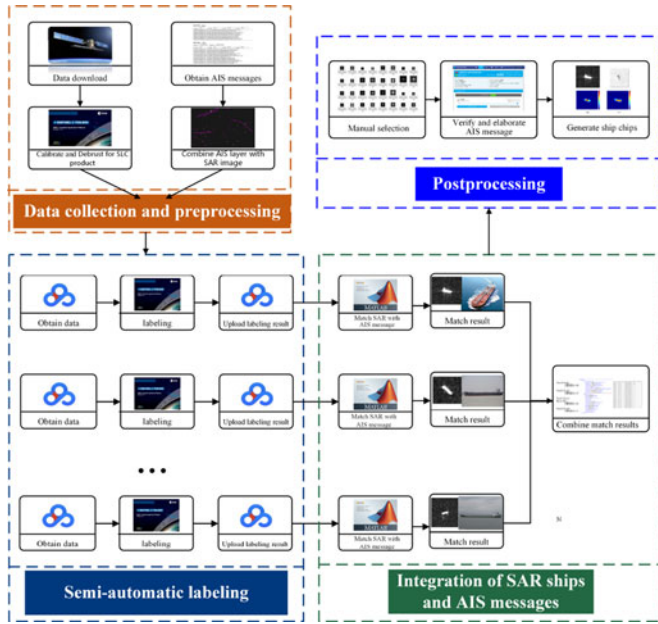
construct a dataset which is particularly applicable for ship interpretation. There are two main objectives of the OpenSAR-Ship. The first is to provide researchers a benchmark dataset to develop applicable and adaptive ship interpretation algorithms and push the performance ceilings of data analysis. The other is to provide a dataset for performing application-oriented quality assessment for Sentinel-1 imagery, which can boost their applications in a targeting way.

The remainder of this paper is organized as follows. Section II introduces the construction of the OpenSARShip. The properties and the organization of the OpenSARShip are presented in Sections III and IV, respectively. Section V gives the geometric and scattering analyses of the OpenSARShip. Preliminary applications are given in Section VI. Finally, the conclusions are drawn and the future work is introduced in Section VII.

## II. OPENSARSHIP CONSTRUCTION

Constructing the OpenSARShip is an ambitious project. In this section, we describe how we construct the OpenSARShip, and how this procedure guarantees the properties which are introduced in Section III. Fig. 1 illustrates the flow of the dataset construction. There are four main steps: data collection and preprocessing, semi-automatic labeling, integration of SAR ships and AIS messages, and postprocessing.

### A. Data Collection and Preprocessing

The OpenSARShip contains 11 346 ship chips for 41 Sentinel-1 SAR images. The original data are downloaded from the Sentinels Scientific Data Hub [21] and the corresponding AIS messages are acquired at the same zone in a 20 min window whose center is the SAR acquisition time.

In the preprocessing, we employ the SNAP 3.0 [22] to perform the radiometric calibration and the terrain observation with progressive ScanSAR (TOPSAR) deburst. The radiometric cal-
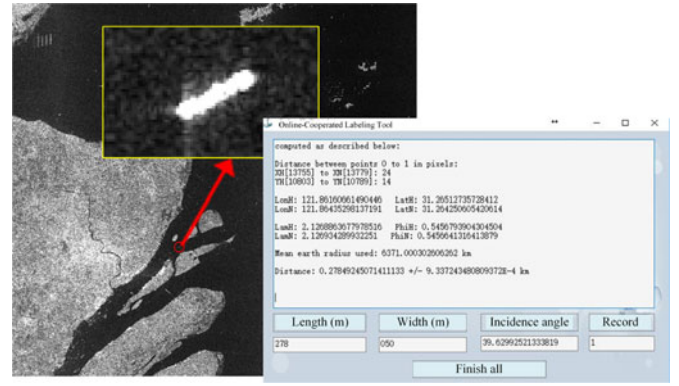
ibration is carried out for both single look complex (SLC) and ground range detected (GRD) products while the deburst is employed merely for SLC products.
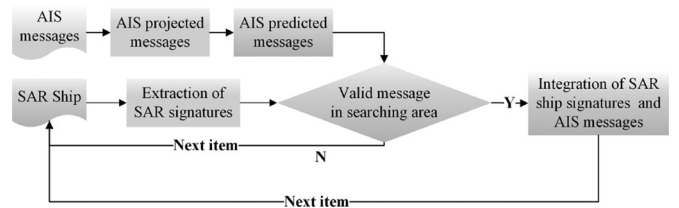
### B. Semi-Automatic Labeling

In this step, we manually extract the positional information (latitude and longitude), the geometric information (length, width, and direction), and the SAR imaging parameters (incidence angle, elevation angle, and slant range time) of each ship via the SNAP 3.0. These signatures can be automatically recorded by a self-designed tool called online-coordinate labeling tool (OCLT). Fig. 2 presents this semi-automatic step by using the SNAP 3.0 and the OCLT. This step plays an essential role in ensuring reliable and accurate OpenSARShip construction because the semi-automatic labeling provides not only important SAR imaging parameters, but also key target parameters (positional and geometric information) which are essential in the next integration step.

### C. Integration of SAR Ships and AIS Messages

The goal of this part is to establish reliable links between information provided by AIS messages and SAR ship signatures. However, positional differences between the SAR signatures and the AIS messages due to ship motion should be taken into consideration in the matching process. For a moving ship, positional differences exist both along the track of the ship and in the azimuth direction of the image, resulting from the temporal difference between AIS and SAR acquisition, and a Doppler frequency shift of the SAR ship, respectively. Motivated by [23], the AIS messages are corrected for position shifts; therefore, the reliability of matching is improved. The processing in the OpenSARShip in Fig. 3 consists of the following steps:

1) Project the AIS positions to the exact SAR acquisition time by interpolation, and generate a set of new positions defined as AIS-projected.

2) For all the AIS messages, based on the velocity vectors reported by AIS messages directly, compensate Doppler frequency shifts to the AIS-projected positions and generate a set of new positions defined as AIS-predicted. Note that the shift is calculated to be zero when the AIS reported velocity is zero (stationary ship). In this case, the AIS-predicted position is equal to the AIS-projected position.

3) Given an SAR ship, search all the AIS-predicted messages within a 300 m radius, and analyze if a valid message for matching is found.

If no valid message is found, this SAR ship is skipped and will not be included in the OpenSARShip. If more than one valid messages are found within the given 300 m radius, we calculate the length errors and distance errors of these valid messages. The length error is the defined as $\delta L = |L_{SAR} - L_{AIS}|$, where $L_{SAR}$ is the ship length manually extracted in Section II-B, and $L_{AIS}$ is the ship length reported by AIS message. The distance error $\delta D$ is the positional difference of the SAR ship and AIS-predicted position. Based on the calculation, if there exists a valid AIS message showing minimum values of both $\delta L$ and $\delta D$, we match the SAR ship with this AIS message; otherwise, we skip this SAR ship.

4) Integrate the SAR ship signatures with the corresponding AIS messages.

### D. Postprocessing

In order to further ensure the reliability of the OpenSARShip, each ship integrated with the AIS message is verified in the MarineTraffic Website [24]. Moreover, in order to guarantee the high diversity of ship types, we elaborate ship types via information provided by the MarineTraffic Website. For example, the ship-type cargo reported by the AIS message can be further categorized as container, bulk carrier, or general cargo.

## III. OPENSARSHIP PROPERTIES

### A. Specificality

The OpenSARShip essentially aims to achieve the application of marine surveillance based on Sentinel-1 SAR imagery. Particularly, the OpenSARShip is constructed in order to address two SAR image interpretation problems.

First, some tough problems, including how to develop sophisticated ship detection and classification algorithms under high interference and how to achieve a highly adaptive and robust ship detection and classification system capable of processing images from a wide range of imagery, are yet to be solved. Therefore, specifically aiming at these practical difficulties, we construct the OpenSARShip dataset and expect it to provide a powerful way to settle the aforementioned problems.

Second, according to [25], SAR imaging characteristics (resolution, noise level, etc.) are not comprehensive enough to determine whether the image is suitable for a particular marine
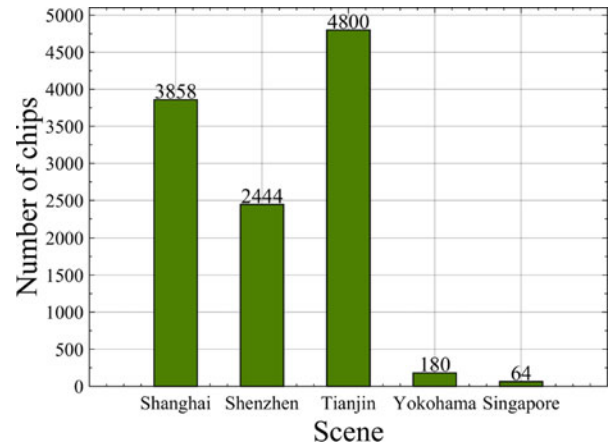


Fig. 4.    Numbers of ship chips in different scenes.

application. Therefore, we construct the OpenSARShip dataset, expect it to evaluate applicable potential of Sentinel-1, and provide effective methods for assessing SAR image quality from both imaging characteristics and marine applicability.

### B. Large Scale

The OpenSARShip is designed to continually collect ship chips integrated with AIS messages with the increasing amount of Sentinel-1 data. The OpenSARShip contains 11 346 ship chips covering 17 AIS types from 41 Sentinel-1 SAR images. These 41 Sentinel-1 SAR images are collected from five typical scenes because of their intense marine traffic: Shanghai Port (China), Shenzhen Port (China), Tianjin Port (China), Yokohama Port (Japan), and Singapore Port (Singapore). For the OpenSARShip, Figs. 4 and 5 show the distributions of the numbers of chips in different scenes and types, respectively. Fig. 5 shows an example the elaborated types of cargo. The elaborated types of all the AIS types are provided in the OpenSARShip.

### C. Diversity

Sentinel-1 are the first satellites exploiting the terrain observation with a TOPSAR technique for the interferometric wide swath mode (IW). The IW mode, with VV–VH dual polarization, is the default Sentinel-1 mode, covering almost global land and littoral zones [26]. Although several other acquisition modes for specific aims are available: the stripmap mode which only be used on request for emergency management, the extra-wide swath mode which is primarily applied to sea-ice monitoring over high latitude areas, and the wave mode which is designed for ocean wind field and swell spectra, yet these modes only cover specific areas with limited quantity and nearly none ship targets.

Therefore, in order to build up a consistent long-term dataset with various coverage, the OpenSARShip focuses on the IW mode, which covers almost global littoral zones and is suitable for ship detection and classification due to its appropriate properties [1], [6]–[9]. The OpenSARShip provides two available products of the IW mode: the SLC and GRD products, listed in Table I. For ship detection and classification, three key factors,
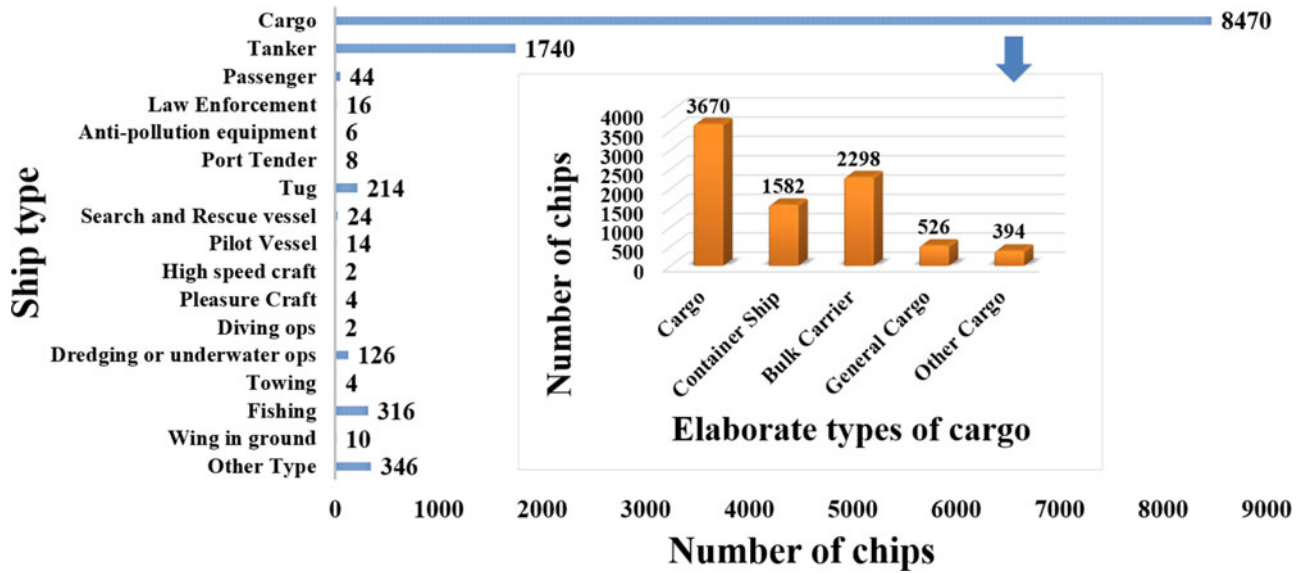
Fig. 5. Numbers of ship chips of different ship types.

TABLE I
PARAMETERS FOR SENTINEL-1 IW PRODUCTS

|  | Resolution rg × az (m) | Looks rg × az | Pixel spacing rg × az (m) |
|---|---|---|---|
| GRD | 20 × 22 | 5 × 1 | 10 × 10 |
| SLC | 2.7 × 22 to 3.5 × 22 | 1 × 1 | 2.3 × 17.4 |

i.e., the target (size, type, direction, etc.), the background (environmental conditions), and the sensor (polarization, mode, incidence angle, etc.) influence the performance of the algorithms for this application. In [8], the influences of polarizations, incidence angles, meteorological conditions, and ship lengths for Sentinel-1 data in ship detection are analyzed. Therefore, in order to guarantee covering different conditions widely, which is called diversity in this paper, the OpenSARShip is organized to contain ship chips with various ship lengths, ship direction angles (with $180°$ ambiguity), and incidence angles. Fig. 6(a)–(d) demonstrates diversity of the OpenSARShip. It should be noted that, in Fig. 6(a), 0 m represents the missing data of ship length reported by AIS messages.

*D. Reliability*

Achieving high precision for the integration of all the SAR ships and AIS messages is challenging because of the errors of matching SAR and AIS messages and the errors of AIS messages themselves. In order to guarantee the reliability of the OpenSARShip, we set strict restrains when matching AIS messages with SAR ships and validate all the AIS messages in MarineTraffic website [24]. Here, we define the labeling noise to describe the error between an SAR ship and its matched AIS message. In order to evaluate the labeling noise quantitatively, the distance error and the ship length error are calculated as measures. The distance error which indicates the two-dimensional distance in the image plane is the positional difference of the SAR ship and the matched AIS

message; and the ship length error is calculated by $|L_{AIS} - L_{SAR}|$, where $L_{AIS}$ is automatically reported by the AIS message and $L_{SAR}$ is manually measured from the SAR image. Fig. 6(e) and (f) shows the labeling noise of all the ship chips in the OpenSARShip. About 73% ship chips are within a 100 m distance error range and 63% within a 20 m length error range, which show high reliability.

*E. Public Availability*

We construct an open platform, *OpenSAR*, to contribute to sharing research inside the SAR image interpretation community. The OpenSARShip dataset is now publicly available at the OpenSAR platform [20].

It is worth noting that the platform merely provides the dataset for research use only, and the corresponding Sentinel-1 SAR images should be found at [21].

## IV. OPENSARSHIP ORGANIZATION

The OpenSARShip is organized in different folders for different images. For a specific image, Fig. 7 illustrates its content.

For every Sentinel-1 SAR image, four subfolders provide four formats of ship chips: original data, visualized data in grey scale, visualized data in pseudo-color, and calibrated data, so as to satisfy different requirements of users. Ship chips are named by pixel coordinates for the convenience of retrieval.

The original data are stored in single 32-bit format which is accordant with Sentinel-1 original data format. For the GRD products, each ship chip is stored in a matrix which indicates the amplitude values of pixels for VH and VV polarizations. For the SLC products, each ship chip is stored in a matrix which indicates the real part values and the imaginary part values of pixels for VH and VV polarizations. Based on the original ship chips, we apply image enhancement programmed by the MATLAB to visualize ship chips with grey scale (stored in unsigned 8-bit integers) for VH and VV polarizations. The SNAP
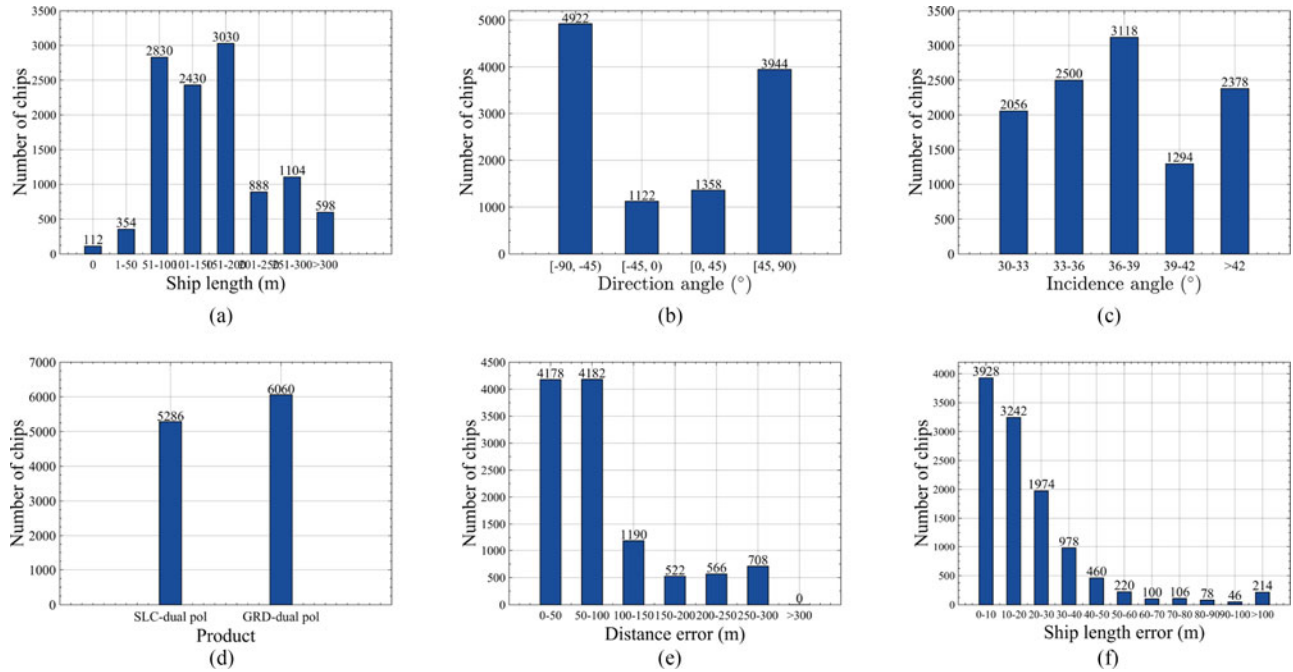
Fig. 6.    Statistics of the OpenSARShip. (a), (b), (c), and (d) are the distributions of the chip numbers in ship lengths, ship direction angles, incidence angles, and product types, respectively. (e) Distance error of matching. (f) Ship length error of matching.
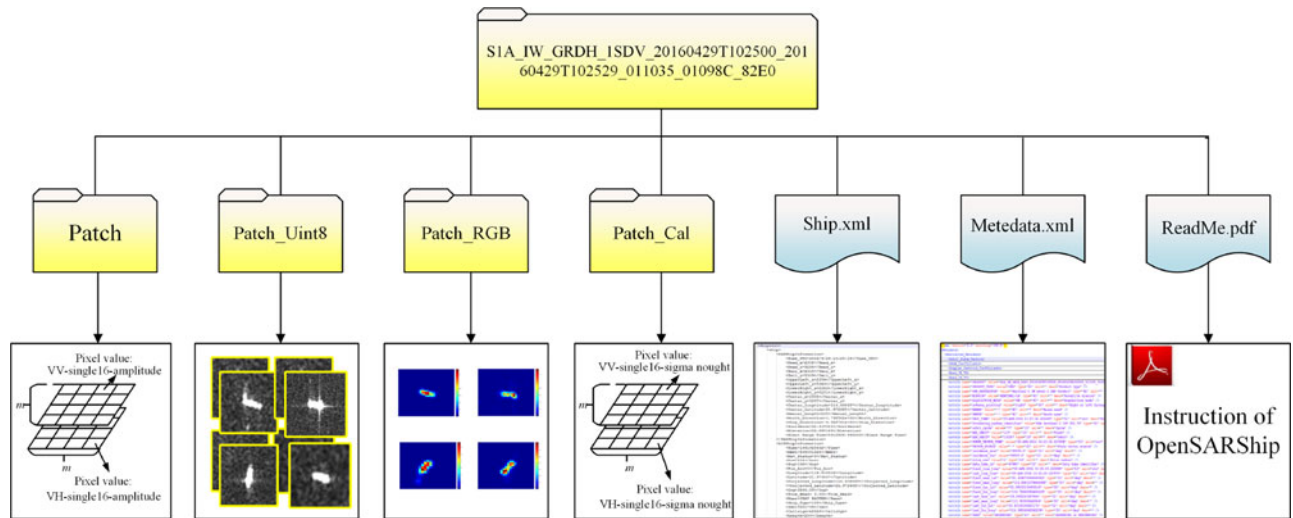


Fig. 7.    Organization of the OpenSARShip.

3.0 is employed to get radiometric calibrated products. The radiometric calibrated data of each ship chip are stored in a matrix which indicates the normalized radar cross section (NRCS) $\sigma_0$ of pixels for VH and VV polarizations. Based on the calibrated ship chip, MATLAB function "Imagesc" is employed to visualize ship chip with pseudo-color. For each ship chip, the RGB visualized chips are provided for VH and VV polarizations. Fig. 8 shows an example. The maximum and minimum NRCS values of this radiometric image are 17.5969 and 0.0029, respectively. The values of the elements of image are indexed into the current ColorMap (jet) that determines the color of pixels. The colorbar is provided to illustrate the radiometric boundary of each pseudo-color ship chip.
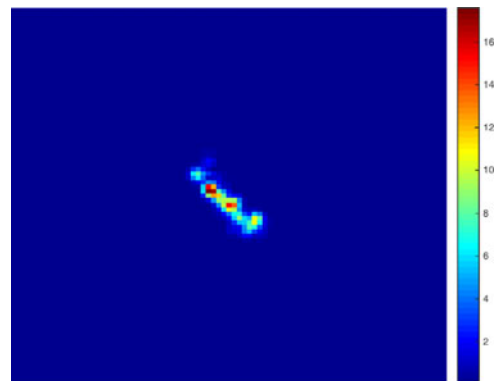


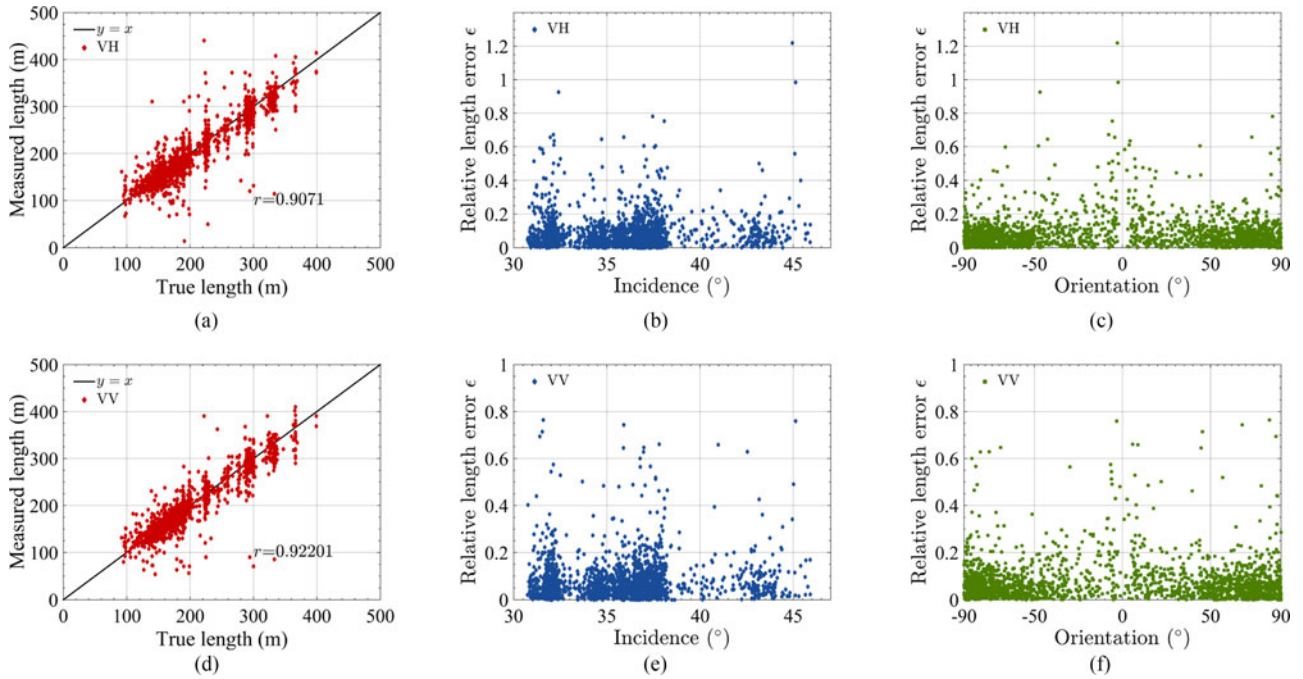Fig. 8.    Pseudo-color representation of ship chip.

Fig. 9.　Factors influence geometric parameter extraction. (a) and (d) are relationships of the measured length and the ground true length for VH and VV polarizations, respectively. (b) and (e) are relationships of the error $\epsilon$ and the incidence angle for VH and VV polarizations, respectively. (c) and (f) are relationships of the error $\epsilon$ and the orientation angle for VH and VV polarizations, respectively.

The detailed information of each ship chip, containing the AIS messages, the SAR ship signatures, and the elaborated messages provided by the MarineTraffic Website, is listed in an XML file named Ship.xml. Via the pixel coordinates shown in the filename of each chip, users can retrieve the corresponding information in the XML file. Besides, an XML file of the metadata of the original SAR image is also provided. The detailed instructions about how to utilize the dataset are explained in the ReadMe.pdf.

## V. OPENSARSHIP ANALYSIS

Ships in SAR images are characterized by spatial patterns and polarization-influenced distributions of scattering centers [27]. Geometric and scattering analyses are the basis for ship target interpretation in SAR images. In this section, by using the OpenSARShip, we analyze the factors which influence the geometric and scattering characteristics. This can provide a deep look of the OpenSARShip and reveal its applicable potential.

### A. Geometric Analysis

First, we describe the approach for geometric parameter estimation, and then elaborate the factors impeding its use.

*1) Dataset:* 3740 GRD samples are selected from the OpenSARShip. All of the samples are in dual polarization and in grey scale.

*2) Parameter Estimation Method:* Ship contours in SAR images are distorted to some degree. For nonmoving objects, geometric distortion results from the side-looking geometry. For moving objects, there is one more reason for geometric distortion, i.e., azimuth smearing, induced by the azimuth component of target velocity and the range component of target acceleration, leading to significant impacts on the extraction of geometric

#### TABLE II
MEAN AND STANDARD DEVIATION OF THE RELATIVE LENGTH ERROR FOR DIFFERENT POLARIZATIONS

| | Mean $\mu$ of the relative length error $\epsilon$ | Standard deviation $\sigma$ of the relative length error $\epsilon$ |
|---|---|---|
| VH | 0.0873 | 0.1014 |
| VV | 0.0779 | 0.0899 |

parameters [28]. Objects on the oceans with strong scattering show artifacts propagating straight from them in the azimuth and range directions due to the existence of sidelobes [29]. The consideration of addressing sidelobe effects should also be taken into ship geometric parameter extraction.

In order to reduce these undesirable effects, two-times morphological filterings are applied. The first morphological filtering is erosion using a line mask whose heading is parallel to the main axis of the ship; and the second filtering, parallel to the auxiliary axis.

*3) Analysis:* In the analysis, the accuracy of ship geometric parameter extraction is measured by the mean $\mu$ and the standard deviation $\sigma$ of the relative length error $\epsilon$, which is the difference of the extracted length and the ground truth length reported by AIS divided by the ground truth.

The relationships between the ground truth length and the extracted length for VH and VV polarizations are illustrated in Fig. 9(a) and (d) and Table II. $\mu_{VV}$ is better than $\mu_{VH}$, and $\sigma_{VV}$ is better than $\sigma_{VH}$. This shows the geometric parameter extraction has smaller relative length error in the VV polarization since scattering is more sensitive to target structure in VH as shown in Section V-B3.
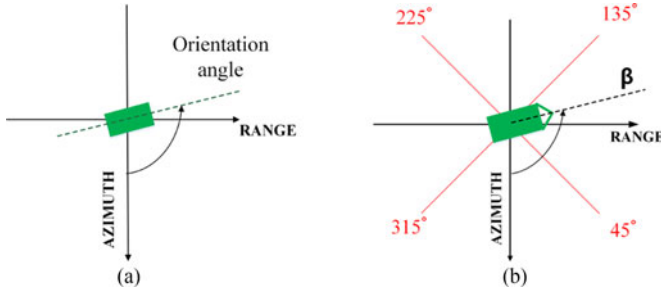
Fig. 10. Definitions of the orientation and the bearing angles.

Aiming to explore the influence of the incidence angle, the relationships between the relative length error $\epsilon$ and the incidence angle for VH and VV polarizations are illustrated in Fig. 9(b) and (e). Due to the unique TOPSAR technique for the IW mode, three subswaths are captured, denoted as IW1, IW2, IW3 band with incidence angles roughly ranging from $30°$ to $36°$, $36°$ to $41°$, $41°$ to $46°$, respectively. Therefore, the samples here are divided into three groups according to the range of incidence angles for IW1, IW2, and IW3 band, respectively. The relative length errors for the smaller incidence angles $(30°, 36°]$, medium incidence angles $(36°, 41°]$, and larger incidence angles $(41°, 46°]$ are $0.0856 \pm 0.1005$, $0.0849 \pm 0.0909$, and $0.1105 \pm 0.1481$ for VH, and $0.0752 \pm 0.0876$, $0.0783 \pm 0.0891$, and $0.0919 \pm 0.1060$ for VV, respectively. This indicates that the geometric parameter extraction generally shows smaller relative length error for the smaller incidence angle since scattering is more sensitive to target structures in the larger incidence angle as shown in Section V-B3.

The orientation angle measures the ship counter-clockwise orientation with respect to the azimuth direction, illustrated in Fig. 10(a). It should be noted that the orientation angle does not distinguish the bow and stern of a ship, so the range of the orientation angle is $(-90°, 90°)$. In order to explore the influence of the orientation angle, the relationships between the relative length error $\epsilon$ and the orientation angle for VH and VV polarizations are illustrated in Fig. 9(c) and (f). These figures show that the errors spread in a wider range if the orientation angle is closer to the $0°$. This is due to the distortion results from the ship motion in the azimuth direction. In order to analyze this relationship quantitatively, the mean $\mu$ and the standard deviation $\sigma$ of the relative length error are illustrated in Fig. 11. The orientation angle appears to be the most influential factor for geometric analysis.

As illustrated in Fig. 9(a) and (d) and Table II, the experiments show relatively good agreement between the extracted length and the ground truth. In [29], three ship detection methods: UPC-WTR, R&G, and SUMO algorithms (see [29] for details) are carried out to estimate lengths of 59 ships. The UPC-WT and R&G algorithms overestimate the ship length, whereas the SUMO algorithm shows underestimation. The three methods measure ship length directly on the detected target cluster without morphology, therefore showing undesirable overestimation and underestimation. In [30], Stasolla and Greidanus proposed a new method based on mathematical morphology for Sentinel-1 imagery. This mathematical morphology aims at better delineating the vessel outline in the cluttered SAR image, thus enabling the extraction of ship's actual dimensions. The mean relative length error is $15\%$ for 127 ships without overestimation or underestimation. This result is consistent with the experiment in this paper. In [31], the geometric/standard method tends to overestimate ship length. One possible explanation is that the method does not have morphological processing. Therefore, the ship's outline in the cluttered SAR image is not well delineated due to the effects of smearing and sidelobes. Hajduch *et al.* [31] fitted an NRCS model with respect to ship length and incidence angle. The model is inverted to estimate ship length and shows good result without overestimation. Based on this novel model [31], the estimation of ship length taking both incidence angle and orientation angle into consideration will be studied in our future work.

### B. Scattering Analysis

Factors, such as product mode, polarization, incidence angle, bearing angle, and velocity, have impacts on the scattering analysis of both ship targets and clutter backgrounds [27]. Based on the OpenSARShip, this part analyzes how these factors influence the use of ship scattering characteristics in classification.

*1) Dataset:* 4040 GRD samples are selected from the OpenSARShip, including three main types (tanker, container ship, and bulk carrier). Tanker is a typical ship designed to transport liquids in bulk, such as hydrocarbon products, chemicals, fresh water, and wine. Container ship is a cargo ship type which carries all of its loads in truck-size containers and is used for the carriage of miscellaneous goods. Bulk carrier is used to transport bulk cargo items such as coal, tin, and wheat. Bulk carrier usually has large box-like hatches on the deck, designed to load and discharge cargo easily. These three ship types cover around $80\%$ of the international shipping market [32]. All of these samples are in dual polarization and radiometrically calibrated into NRCS $\sigma_0$ format. Then, the whole samples are divided into several groups according to the different factors: polarization, velocity, incidence angle, and bearing angle. As illustrated in Fig. 10(b), the bearing angle $\beta$ refers to the angle between the azimuth direction and the ship-bow direction. The whole GRD samples are divided into four groups according to $\beta$: $(-45°, 45°]$, $(45°, 135°]$, $(135°, 225°]$, and $(225°, 315°]$.

In some cases, this direction can be estimated by examining ship wake [33] and the shape of ship, which depends on the type of ship. In this paper, the bearing angle is estimated with the help of AIS message.

*2) Feature Selection:* First, the ship bearing angle is estimated from the SAR image ship chip, with the help of the corresponding AIS message. Then, motivated by [34], the principal axis feature vector is employed for classification. In this paper, the principal axis feature is defined as $F = \{F_{axis}\}$, where $F_{axis}$ denotes the values directly extracted from the bow-to-stern axis and normalized to 50 dimensions. Aiming at analyzing how aforementioned factors influence the use of ship scattering characteristics in classification, the bow-to-stern axis is chosen as the feature vector for two reasons. First, most of the ships are designed to be symmetric along their centerlines [34];
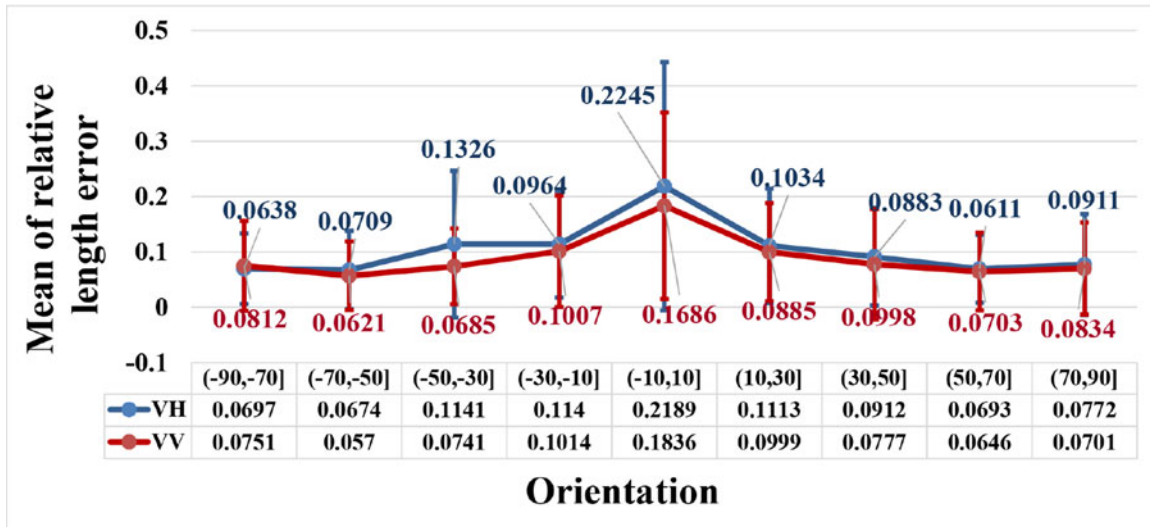
Fig. 11.     Quantitative analysis of the influence of the orientation angle in geometric parameter extraction.

| Paramaters | Incidence angle (degree) | | | Velocity | | Bearing angle (degree) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Product-Polarization | (30,36] | (36,41] | (41,46] | Velocity=0 | Velocity>0 | | (-45,45] | (45,135] | (135,225] | (225,315] |
| GRD-VH | 78.23% | 77.84% | 49.25% | 78.30% | 73.19% | Velocity=0 | 76.79% | 78.98% | 73.61% | 74.44% |
| | | | | | | Velocity>0 | 73.57% | 78.02% | 64.53% | 72.39% |
| GRD-VV | 79.17% | 77.07% | 51.13% | 78.62% | 74.33% | Velocity=0 | 80.71% | 80.88% | 68.06% | 74.32% |
| | | | | | | Velocity>0 | 76.67% | 78.20% | 63.96% | 74.02% |

Fig. 12.     Classification accuracy $P_C$ for experimental datasets.

therefore, the bow-to-stern axis can represent the mainly disparate structures of ships. Besides, compared with the local radar cross section (LRCS) (three sections) feature [35] defined in Section VI-A, the principal axis is a relatively reliable feature to reflect the scattering characteristics since it does not lose too much scattering information.

*3) Analysis:* In the analysis, the k-nearest neighbors (k-NN) algorithm [36], a classic method sensitive to the local structure of the data [37], is used to perform classification. Aiming to analyze the scattering characteristics and estimate the potential classification capability, 4040 samples (in NRCS $\sigma^0$ format) with 70% for training and 30% for testing including three aforementioned types (tanker, container ship, and bulk carrier) are used.

In the analysis, the results of classification are measured by classification accuracy $P_C = N_{correct}/N_{total}$, where $N_{correct}$ denotes the number of correctly classified ships and $N_{total}$ denotes the total number of ships. We use $P_C$ as a measure to show the influential factors and application potential of the scattering analysis by using the OpenSARShip.

The classification accuracies for different incidence angles are shown in the green rectangular in Fig. 12. Due to the unique TOPSAR technique for the IW mode, three subswaths are captured, denoted as IW1, IW2, IW3 band with incidence angles

roughly ranging from 30° to 36°, 36° to 41°, 41° to 46°, respectively. Therefore, the samples here are divided into three groups according to the range of incidence angles for IW1, IW2, and IW3 band, respectively. We observe that the samples in the smaller incidence angles show better classification performances for both polarizations. The explanation is that the scattering characteristics are more stable in the smaller incidence angles than that in the larger incidence angles [34]. Specifically, under the condition of smaller incidence angles, the scattering characteristics maintain relatively stable with respect to the changes of factors (i.e., velocity, bearing angle). The stable scattering characteristics guarantee better performances for classification. Otherwise, when the incidence angles are larger, the scattering characteristics tend to be more sensitive to the changes of factors; thus. the classification performances become worse.

The classification accuracies for different polarizations and velocities are listed in the yellow rectangular in Fig. 12. For VH and VV polarization, the differences of $P_C$ between the groups covering stationary and moving samples are 5.11% and 4.29%, respectively. It indicates that VH is more sensitive to the changes of velocities. We illustrate this visually. The first and third rows in Fig. 13 show the stationary ship chips. According to the photos, it can be inferred that the scattering centers are in
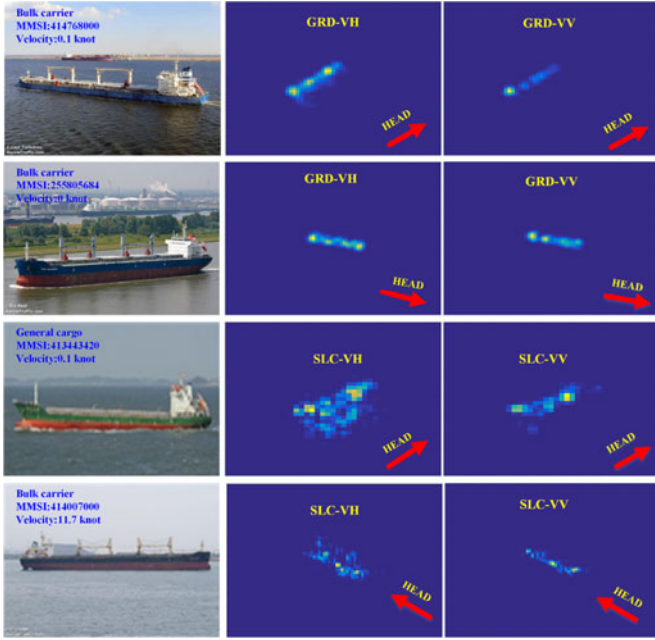
Fig. 13. Impacts of movement and polarization on ship scattering analysis.

the decks, which are better demonstrated in VH. The second and fourth rows in Fig. 13 show the moving ship chips. Smearing effects induced by movement are more noticeable in VH, due to better reflection of scattering structures in this polarization; thus, we can understand that the VH polarization is more sensitive to movement.

It is unsound to directly conclude which polarization is more suitable for ship classification. For one thing, VH is more an indication of volume scattering [6]. For another, although double or direct bounce structures cause intense backscattering in both polarizations, VV polarization performs larger intensity values of double bounce due to its higher signal-to-noise ratio [6], [38]. Therefore, to some extent, classification performances of VV and VH polarizations are related to the different scattering mechanisms corresponding to different ship types.

The classification accuracies for different bearing angles and velocities are listed in the blue rectangular in Fig. 12. Two observations can be made. First, the groups covering stationary ship samples perform better than those covering moving ship samples in both polarizations and four bearing angle ranges. The differences of $P_C$ between the stationary and moving ship samples result from the different spatially-distributed features, which are affected by smearing due to the ship movement [27], [39], [40]. Besides, the mean differences of $P_C$ in the groups of $(-45°, 45°]$ and $(135°, 225°]$ are 5.11%; otherwise, the differences of $P_C$ in the groups of $(45°, 135°]$ and $(225°, 315°]$ are 1.50%. Obviously, the differences between the stationary and moving samples are larger when the bearing angles of ships are parallel to the azimuth direction (groups denoted as $(-45°, 45°]$ and $(135°, 225°]$). This is because the smearing is mainly caused by the azimuth component of velocity [41].

In this section, the factors influencing the geometric and scattering analyses are discussed. This provides an inside and deep look at the OpenSARShip, and demonstrates its

application potential. Led by these analyses, the users of the OpenSARShip can develop suitable and adaptive models for geometric and scattering analyses, and continuously push the performance ceilings with the help of the OpenSARShip.

## VI. APPLICATION AND DISCUSSION

In this section, aiming at the two objectives introduced in Section III-A, we give two promising applications to demonstrate the applicability of the OpenSARShip.

### A. Benchmark

The geometric and scattering analyses of the OpenSARShip are elaborated in Section V. Based on the detailed analyses, the first application of the OpenSARShip is to be a promising benchmark.

The OpenSARShip is a benchmark for ship geometric and scattering analyses, and detection and classification algorithms. According to the analyses presented in Section V, the extraction of geometric and scattering information still suffers from limited performance; this impedes the development of effective and robust detection and classification algorithms.

The goal is to construct the OpenSARShip as a benchmark for researchers to develop applicable and adaptive models, and push the performance limitations of detection and classification algorithms. Therefore, in the following experiment, we choose to illustrate the usefulness of the OpenSARShip by achieving ship classification based on different features.

*1) Feature Selection:* Due to the medium-to-high resolution of the Sentinel-1 products, seven features are employed for classification. These features can be divided into three categories: geometry, moments, and scattering statistics. Motivated by the recent researches of ship classification, some effective features are selected and improved to be more suitable to the OpenSARShip. The definitions of these features are described as follows.

The geometric feature, motivated by [42], is defined as $F_1 = \{G_i\}_{i=1}^8$, where $\{G_i\}_{i=1}^8$ denote the ship length, width, aspect ratio, area, perimeter, shape complexity, $x$ coordinate of centroid, and $y$ coordinate of centroid.

The Hu moment invariants feature [43] is defined as $F_2 = \{H_i\}_{i=1}^7$, where $\{H_i\}_{i=1}^7$ denote the Hu moment invariants.

The scattering center feature uses the scattering centers extracted from the principal axis, defined as $F_3 = \{N_{\text{center}}, \mu_{\text{center}}, \sigma_{\text{center}}, M_{\text{center}}\}$. Here, $N_{\text{center}}$ denotes the number of scattering centers, and $\{\mu_{\text{center}}, \sigma_{\text{center}}, M_{\text{center}}\}$ denote the mean, the standard deviation, and the maximum values extracted from the scattering centers, respectively.

The principal axis feature, motivated by [34], is defined as $F_4 = \{F_{\text{axis}}\}$, where $F_{\text{axis}}$ denotes the values directly extracted from the bow-to-stern axis and normalized to 50 dimensions.

The principal axis (three sections) feature, motivated by [34], is defined as $F_5 = \{\mu_{\text{bow}}, \mu_{\text{middle}}, \mu_{\text{stern}}\}$, where $\mu_{\text{bow}}, \mu_{\text{middle}}, \mu_{\text{stern}}$ are the mean values extracted from the bow, middle, and stern sections of the ship, respectively.

The LRCS feature, motivated by [44], is defined as $F_6 = \{S_{\text{local}}, \mu_{\text{local}}, \sigma_{\text{local}}\}$, which denotes the sum, the mean, and the standard deviation of the values extracted from the ship target.

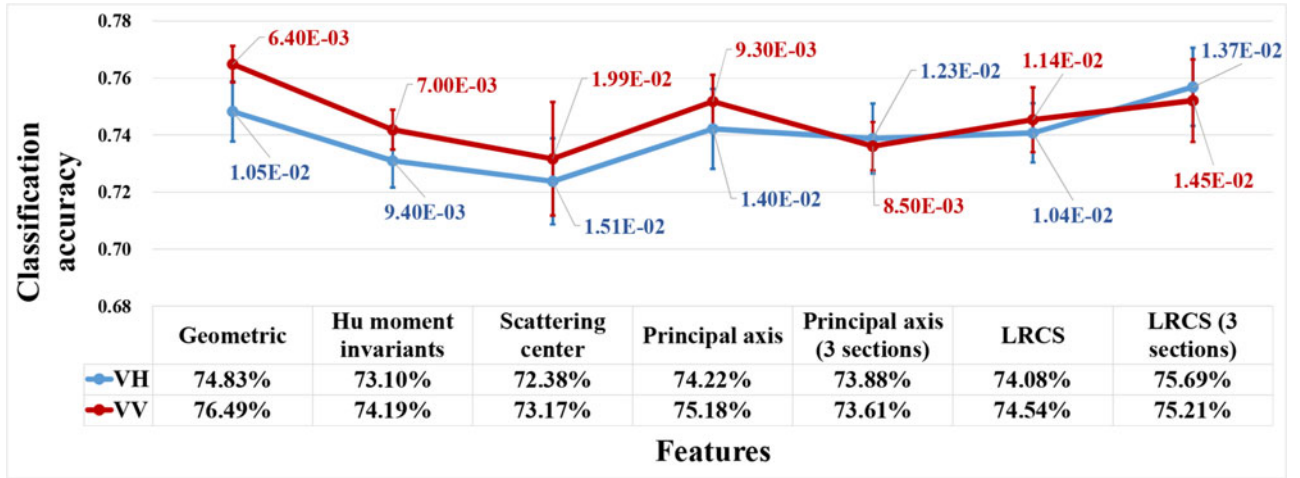| | Geometric | Hu moment invariants | Scattering center | Principal axis | Principal axis (3 sections) | LRCS | LRCS (3 sections) |
|---|---|---|---|---|---|---|---|
| VH | 74.83% | 73.10% | 72.38% | 74.22% | 73.88% | 74.08% | 75.69% |
| VV | 76.49% | 74.19% | 73.17% | 75.18% | 73.61% | 74.54% | 75.21% |

Fig. 14. Performance comparison of ship classification based on different features.

The LRCS (three sections) feature, motivated by [35], is defined as $F_7 = \{B_i, M_i, S_i\}_{i=1}^3$, where $\{B_i\}_{i=1}^3$, $\{M_i\}_{i=1}^3$, $\{S_i\}_{i=1}^3$ denote the sum, the mean, and the standard deviation of the values extracted from the bow, middle, and stern sections of the ship, respectively.

*2) Analysis:* As described in Section V-B, we use the same 4040 experimental data (in NRCS $\sigma^0$ format) with 70% for training and 30% for testing, and the same classification method k-NN in this section.

The classification performances based on different features are illustrated in Fig. 14. This indicates that the geometric feature, the principal axis feature, and the LRCS (three sections) feature show better classification performances. For the Sentinel-1 products with medium-to-high resolution, the categories of geometric and scattering statistical features show better classification performances. However, the distinctions of the classification accuracies among different features are not significant. Therefore, for ship classification, the feature selection is still a challenging problem needing further research.

The OpenSARShip serves as a benchmark for deep analysis with the help of a large and diverse dataset. Therefore, the OpenSARShip has the potential for researchers to develop meaningful and robust analysis due to big data, and determine which information is vital and how large is enough.

### B. Assessment of Imagery Applicability

Aiming at evaluating the applicability of Sentinel-1 imagery, a model is designed to estimate the minimum detectable ship length by using the OpenSARShip [45]. The minimum detectable ship length is a measure of ship detectability assessment and a practical application-oriented way of image quality evaluation.

*1) Ship Detectability Model:* Modeling the ship detectability is a statistical problem which considers imaging characteristics and environmental conditions. Motivated by [46], Fig. 15 shows the framework of the proposed algorithm, which consists of four steps.
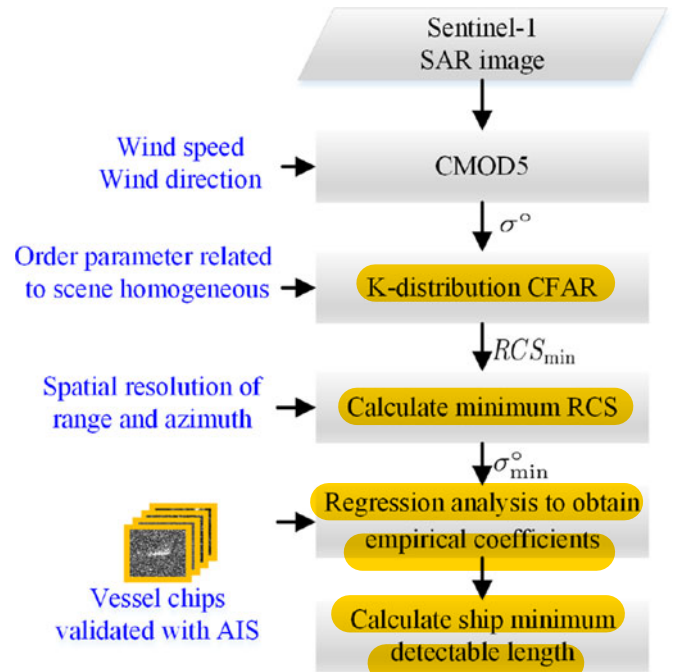


Fig. 15. Flow of the proposed algorithm.

First, as described in [47], we calculate the NRCS $\sigma^0$ using the geophysical model function CMOD5 for the given ranges of incidence angle, wind speed, and wind direction.

Then, the $\mathcal{K}$-distribution [48] is employed in the constant false alarm rate (CFAR) algorithm [49]. For a given false alarm rate $P_{\text{FA}}$, the minimum value of detection $\sigma_{\min}^0$ (= threshold $T$) is obtained. In the $\mathcal{K}$-distribution, the order parameter $\nu$ is related with scene homogeneity [50]. For homogeneous clutter, the $\mathcal{K}$-distribution tends to be the Gaussian with large $\nu$, while for inhomogeneous clutter, the $\mathcal{K}$-distribution tends to be long tail with small $\nu$.

Third, given the spatial resolutions in range $r_{\text{rg}}$ and azimuth $r_{\text{az}}$, the minimum RCS ($\text{RCS}_{\min}$) is calculated according to $\text{RCS}_{\min} = r_{\text{rg}} \times r_{\text{az}} \times \sigma_{\min}^0$.

<div style="text-align:center">

TABLE III
EXPERIMENTAL PARAMETERS AND RESULTS

</div>

|        | $P_{\mathrm{FA}}$ | Wind speed | Wind direction | $\nu$ | $L_{\min}$ |
|--------|-------|------------|----------------|----|------------|
| GRD-VH | $10^{-5}$ | $3 - 5\,\mathrm{m/s}$ | $5°$ | 21 | $43\,\mathrm{m}$ |
| GRD-VV | $10^{-5}$ | $3 - 5\,\mathrm{m/s}$ | $5°$ | 26 | $35\,\mathrm{m}$ |
| SLC-VH | $10^{-5}$ | $3 - 5\,\mathrm{m/s}$ | $145°$ | 17 | $23\,\mathrm{m}$ |
| SLC-VV | $10^{-5}$ | $3 - 5\,\mathrm{m/s}$ | $145°$ | 20 | $19\,\mathrm{m}$ |

Finally, based on the large samples in the OpenSARShip, the empirical function described in [51] is employed to determine the link between the $\mathrm{RCS}_{\min}$ and the minimum ship detectable length $L_{\min}$.

*2) Ship Detectability Estimation:* We evaluate the ship detectability of two Sentinel-1 IW images. Both images are acquired at Shanghai Port. One is a GRD image acquired at 2016-01-16 09:54, and another is an SLC image acquired at 2016-03-28 09:54.

Based on the calibrated products, the parameters listed in Table III are inputted into the model. $P_{\mathrm{FA}}$ is set by empirical value, $\nu$ is estimated from the homogeneous area of the experimental SAR images by the method of log-cumulants [52], and wind speed and wind direction are set according to the data provided by Shanghai Pudong airport meteorological station. Following the steps shown in Fig. 15, the minimum detectable ship lengths for both images are estimated as listed in the last column of Table III.

As we can see, the SLC image performs better than the GRD for the smaller values of minimum detectable ship lengths. Because the range resolution of the SLC products in range is better than the GRD products. For different polarizations, the VV polarization images perform smaller values of minimum detectable ship lengths than the VH polarization. One possible explanation is that due to the larger signal-to-noise ratio of VV polarization [38], the NRCS $\sigma^0$ in VV polarization is larger than that in VH polarization. Therefore, the NRCS $\sigma^0$ of targets in VV polarization are larger than those in VH polarization, resulting in smaller minimum detectable ship length values.

*3) Validation:* The model of estimating minimum detectable ship length combines statistical and semi-empirical modeling; thus, it is impossible to take all factors into consideration and give a generalized validation. Here, the CFAR algorithm is employed to give an approximate validation of the proposed method. The probability of detection is calculated as $P_D = N_D/N_{\mathrm{total}}$, where $N_D$ is the number of detected ships, and $N_{\mathrm{total}}$ is the total number of ships integrated with AIS messages.

As shown in Fig. 16(a), both the GRD-VH and the GRD-VV SAR images perform desirable ship detectability for ships whose lengths are greater than 100 m. For VH polarization, since $P_D$ for ships under 30 m and between 30 and 40 m are 0 and 0.06, respectively, these ships can be regarded as undetectable. For ships between 40 and 50 m, $P_D$ reaches 0.22, which indicates the detectability of some ships at this range of length. This approximately agrees with the estimation $L_{\min} = 43$ m. Similarly, for VV polarization, $P_D$ is 0.163 for ships
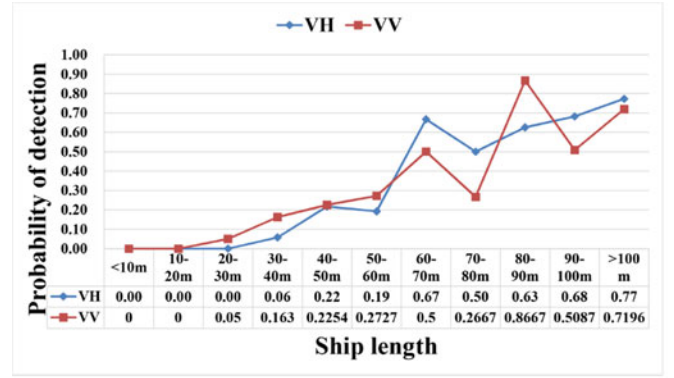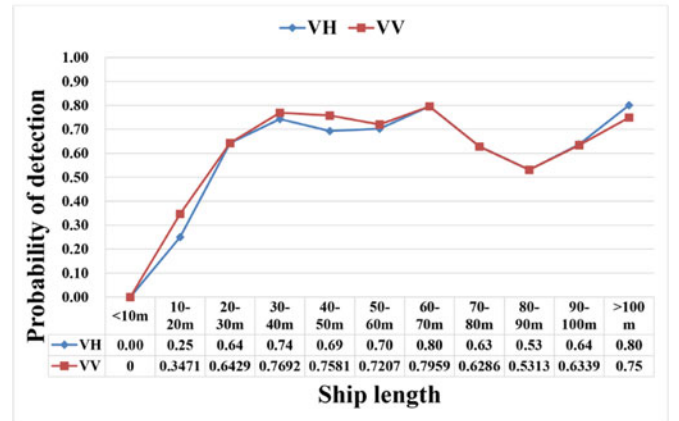




Fig. 16. Validation results. (a) and (b) are relationships between the ship length and the probability of detection for Sentinel-1 IW GRD and SLC modes, respectively.

between 30 and 40 m, which approximately agrees with the estimation $L_{\min} = 35$ m.

As shown in Fig. 16(b), both the SLC-VH and the SLC-VV SAR images perform desirable ship detection for ships whose lengths are greater than 100 m. For VV polarization, $P_D$ is 0.35 for ships between 10 and 20 m, which approximately agrees with the estimation $L_{\min} = 19$ m. For VH polarization, $P_D$ reaches 0.25 for ships between 10 and 20 m, and 0.64 for ships between 20 and 30 m. Although this does not exactly agree with the estimation $L_{\min} = 23$ m, $L_{\min}$ still can be regarded as an effective measure for image detectability assessment. After all, the minimum detectable ship length provides a statistical and general estimation of detectability considering imaging characteristics and environmental conditions, instead of a definite value to determine whether a ship can be detected or not.

## VII. CONCLUSION AND FUTURE WORK

In this paper, in order to exploit the rapidly growing Sentinel-1 SAR imagery, we present a constructed dataset dedicated to ship target interpretation. The OpenSARShip, providing 11 364 SAR ship chips integrated with AIS messages, owes five essential properties: specificity, large scale, diversity, reliability, and public availability. The method of constructing the OpenSARShip is introduced, showing how the properties are

guaranteed. The geometric and scattering analyses which provide an inside and deep look at the dataset, the feature benchmark, and the imagery applicability assessment all demonstrate the applicable potential of the OpenSARShip.

In the future, the OpenSARShip will be continuously enlarged. Furthermore, we expect to explore more challenging applications based on the OpenSARShip.

For one thing, ship detection and classification has been essential research for decades. New platforms such as Sentinel-1 and the planned RADARSAT Constellation request more challenging demands of ship detection and classification, such as robust performances in complex circumstances, reduced latency time, and full-automatic processing. For another, based on ship detection and classification, many applications in marine surveillance have been investigated, such as marine traffic monitoring by AIS messages [53], fishing regulation [54], and image quality assessment to marine surveillance [25]. In the era of big data, the OpenSARShip is promising to contribute to the development of more advanced SAR marine applications.

## REFERENCES

[1] D. Velotto, C. Bentes, B. Tings, and S. Lehner, "Comparison of Sentinel-1 and TerraSAR-X for ship detection," in *Proc. 2015 IEEE Int. Geosci. Remote Sens. Symp.*, 2015, pp. 3282–3285.

[2] C. Santamaria, M. Alvarez, H. Greidanus, V. Syrris, P. Soille, and P. Argentieri, "Mass processing of Sentinel-1 images for maritime surveillance," *Remote Sens.*, vol. 9, no. 7, 2017, Art. no. 678.

[3] J. Karvonen, "Baltic sea ice concentration estimation using Sentinel-1 SAR and AMSR2 microwave radiometer data," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 2871–2883, May 2017.

[4] F. Monaldo, C. Jackson, X. Li, and W. G. Pichel, "Preliminary evaluation of Sentinel-1A wind speed retrievals," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 9, no. 6, pp. 2638–2642, Jun. 2016.

[5] L. Huang, B. Liu, X. Li, Z. Zhang, and W. Yu, "Technical evaluation of Sentinel-1 IW mode cross-pol radar backscattering from the ocean surface in moderate wind condition," *Remote Sens.*, vol. 9, no. 8, 2017, Art. no. 854.

[6] D. Velotto, C. Bentes, B. Tings, and S. Lehner, "First comparison of Sentinel-1 and TerraSAR-X data in the framework of maritime targets detection: South Italy case," *IEEE J. Ocean. Eng.*, vol. 41, no. 4, pp. 993–1005, Oct. 2016.

[7] C. Schwegmann, W. Kleynhans, and B. Salmon, "Synthetic aperture radar ship detection using Haar-Like features," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 2, pp. 154–158, Feb. 2017.

[8] R. Pelich, N. Longépé, G. Mercier, G. Hajduch, and R. Garello, "Performance evaluation of Sentinel-1 data in SAR ship detection," in *Proc. 2015 IEEE Int. Geosci. Remote Sens. Symp.*, 2015, pp. 2103–2106.

[9] P. W. Vachon, J. Wolfe, and H. Greidanus, "Analysis of Sentinel-1 marine applications potential," in *Proc. 2012 IEEE Int. Geosci. Remote Sens. Symp.*, 2012, pp. 1734–1737.

[10] Y. Yang and S. Newsam, "Bag-of-visual-words and spatial extensions for land-use classification," in *Proc. 18th SIGSPATIAL Int. Conf. Adv. Geographic Inf. Syst.*, 2010, pp. 270–279.

[11] O. A. Penatti, K. Nogueira, and J. A. dos Santos, "Do deep features generalize from everyday objects to remote sensing and aerial scenes domains?" in *Proc. IEEE Conf. Comput. Vision Pattern Recog. Workshops*, 2015, pp. 44–51.

[12] I. W. III/4, "ISPRS 2D semantic labeling contest," 2015. [Online]. Available: http://www2.isprs.org/commissions/comm3/wg4/semantic-labeling.html

[13] R. O. Green *et al.*, "Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS)," *Remote Sens. Environ.*, vol. 65, no. 3, pp. 227–248, 1998.

[14] S. Holzwarth *et al.*, "HySens-DAIS 7915/ROSIS imaging spectrometers at DLR," in *Proc. 3rd EARSeL Workshop Imag. Spectrosc.*, 2003, pp. 3–14.

[15] UNAVCO, "Western North America interferometric synthetic aperture radar consortium," 2014. [Online]. Available: http://winsar.unavco.org/winsar.html

[16] E. R. Keydel, S. W. Lee, and J. T. Moore, "MSTAR extended operating conditions: A tutorial," in *Proc. Aerosp./Defense Sens. Controls*, 1996, pp. 228–242.

[17] F. Sadjadi, "Adaptive object classification using complex SAR signatures," in *Proc. IEEE Conf. Comput. Vision Pattern Recog. Workshops*, 2016, pp. 33–37.

[18] H. Wang, S. Chen, F. Xu, and Y.-Q. Jin, "Application of deep-learning algorithms to MSTAR data," in *Proc. 2015 IEEE Geosci. Remote Sens. Symp.*, 2015, pp. 3743–3745.

[19] S. Chen, H. Wang, F. Xu, and Y.-Q. Jin, "Target classification using the deep convolutional networks for SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4806–4817, Aug. 2016.

[20] Shanghai Jiao Tong University, "OpenSAR Platform," 2017. [Online]. Available: http://opensar.sjtu.edu.cn/

[21] ESA, "Copernicus Open Access Hub," 2015. [Online]. Available: https://scihub.copernicus.eu/

[22] ESA, "Step Science Toolbox Explotiation Platform," 2015. [Online]. Available: http://step.esa.int/main/

[23] P. W. Vachon, R. A. English, and J. Wolfe, "Ship signatures in RADARSAT-1 ScanSAR Narrow B imagery: Analysis with AISLive data," Defence Res. Develop. Canada-Ottawa, Ottawa, ON, USA, 2007.

[24] MarineTraffic, "Ship list with details and photos," 2007. [Online]. Available: http://www.marinetraffic.com/en/ais/index/ships/all

[25] M. Vespe and H. Greidanus, "SAR image quality assessment and indicators for vessel and oil spill detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4726–4734, Nov. 2012.

[26] R. Torres *et al.*, "GMES Sentinel-1 mission," *Remote Sens. Environ.*, vol. 120, pp. 9–24, 2012.

[27] G. Margarit, J. J. Mallorqui, J. Fortuny-Guasch, and C. Lopez-Martinez, "Exploitation of ship scattering in polarimetric SAR for an improved classification under high clutter conditions," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 4, pp. 1224–1235, Apr. 2009.

[28] B. Tings, C. A. Bentes da Silva, and S. Lehner, "Dynamically adapted ship parameter estimation using TerraSAR-X images," *Int. J. Remote Sens.*, vol. 37, no. 9, pp. 1990–2015, 2016.

[29] M. Stasolla, J. J. Mallorqui, G. Margarit, C. Santamaria, and N. Walker, "A comparative study of operational vessel detectors for maritime surveillance using satellite-borne synthetic aperture radar," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 9, no. 6, pp. 2687–2701, Jun. 2016.

[30] M. Stasolla and H. Greidanus, "The exploitation of sentinel-1 images for vessel size estimation," *Remote Sens. Lett.*, vol. 7, no. 12, pp. 1219–1228, 2016.

[31] G. Hajduch, N. Longepe, J. Habonneau, and J. Le Bras, "Progress in automatic ship detection and classification," in *Proc. SeaSAR 2012*, vol. 709, 2013.

[32] H. Zhang, X. Tian, C. Wang, F. Wu, and B. Zhang, "Merchant vessel classification based on scattering component analysis for COSMO-SkyMed SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 6, pp. 1275–1279, Nov. 2013.

[33] P. W. Vachon and J. Wolfe, "Validation of ship signatures in envisat asar ap mode data using aislive: Data acquisition, processing, and analysis results," Defence Res. Develop. Canada-Ottawa, Ottawa, ON, USA, Tech. Rep., 2008.

[34] G. Margarit and A. Tabasco, "Ship classification in single-pol SAR images based on fuzzy logic," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 8, pp. 3129–3138, Aug. 2011.

[35] X. Xing, K. Ji, H. Zou, W. Chen, and J. Sun, "Ship classification in TerraSAR-X images with feature space based sparse representation," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 6, pp. 1562–1566, Nov. 2013.

[36] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *Amer. Statist.*, vol. 46, no. 3, pp. 175–185, 1992.

[37] M. Mohammed, M. B. Khan, and E. B. M. Bashier, *Machine Learning: Algorithms and Applications*. Boca Raton, FL, USA: CRC Press, 2016.

[38] P. W. Vachon and J. Wolfe, "GMES Sentinel-1 analysis of marine applications potential," Defence Res. Develop. Canada-Ottawa, Ottawa, ON, USA, vol. 218, 2008.

[39] K. Ouchi, M. Iehara, K. Morimura, S. Kumano, and I. Takami, "Nonuniform azimuth image shift observed in the Radarsat images of ships in motion," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 10, pp. 2188–2195, Jan. 2002.

[40] R. K. Raney, "Synthetic aperture imaging radar and moving targets," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-7, no. 3, pp. 499–505, May 1971.

[41] K. Ouchi, "On the multilook images of moving targets by synthetic aperture radars," *IEEE Trans. Antennas Propag.*, vol. 33, no. 8, pp. 823–827, Aug. 1985.

[42] W. Chen, K. Ji, X. Xing, H. Zou, and H. Sun, "Ship recognition in high resolution SAR imagery based on feature selection," in *Proc. 2012 Int. Conf. Comput. Vision Remote Sens.*, 2012, pp. 301–305.

[43] M.-K. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. Inf. Theory*, vol. 8, no. 2, pp. 179–187, 1962.

[44] C. Wang, H. Zhang, F. Wu, S. Jiang, B. Zhang, and Y. Tang, "A novel hierarchical ship classifier for COSMO-SkyMed SAR data," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 2, pp. 484–488, Feb. 2014.

[45] L. Huang, B. Liu, W. Guo, Z. Zhang, and W. Yu, "Preliminary evaluation of vessel detectability for Sentinel-1 SAR data," in *Proc. 2017 IEEE Int. Geosci. Remote Sens. Symp.*, 2017, pp. 1848–1851.

[46] C. Bentes, D. Velotto, and S. Lehner, "Analysis of ship size detectability over different TerraSAR-X modes," in *Proc. 2014 IEEE Geosci. Remote Sens. Symp.*, 2014, pp. 5137–5140.

[47] H. Hersbach, "CMOD5: An improved geophysical model function for ERS C-band scatterometry." Eur. Centre Medium-Range Weather Forecasts, Reading, U.K., No. 395, 2003.

[48] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*. Raleigh, NC, USA: SciTech Publishing, 2004.

[49] H. Greidanus, M. Alvarez, C. Santamaria, F.-X. Thoorens, N. Kourti, and P. Argentieri, "The sumo ship detector algorithm for satellite radar images," *Remote Sens.*, vol. 9, no. 3, 2017, Art. no. 246.

[50] D. J. Crisp, L. Rosenberg, N. J. Stacy, and Y. Dong, "Modelling X-band sea clutter with the K-distribution: Shape parameter variation," in *Proc. 2009 Int. Radar Conf. Surveillance Safer World*, 2009, pp. 1–6.

[51] P. Vachon, J. Campbell, C. Bjerkelund, F. Dobson, and M. Rey, "Ship detection by the RADARSAT SAR: Validation of detection model predictions," *Can. J. Remote Sens.*, vol. 23, no. 1, pp. 48–59, 1997.

[52] J. Nicolas, "Application de la transformée de mellin: étude des lois statistiques de limagerie cohérente," *Rapport de recherche*, 2006D010, 2006.

[53] C. Brekke, D. J. Weydahl, Ø. Helleren, and R. Olsen, "Ship traffic monitoring using multi-polarisation satellite SAR images combined with ais reports," in *Proc. 2008 7th Eur. Conf. Synthetic Aperture Radar*, 2008, pp. 1–4.

[54] F. Mazzarella, M. Vespe, D. Damalas, and G. Osio, "Discovering vessel activities at sea using AIS data: Mapping of fishing footprints," in *Proc. 2014 17th Int. Conf. Inf. Fusion*, 2014, pp. 1–7.

**Bin Liu** (M'16) was born in Hunan, China, in 1985. He received the B.Sc. degree in information engineering, the M.Sc. and Ph.D. degrees in signal and information processing from Shanghai Jiao Tong University, Shanghai, China, in 2007, 2009, and 2015, respectively.

From October 2012 to April 2013, he was a visiting Ph.D. student in the Department of Signal and Image Processing, Telecom ParisTech, Paris, France, under the supervision of Prof. F. Tupin. Currently, he is a Research Assistant Professor in the Shanghai Key Laboratory of Intelligent Sensing and Recognition, Shanghai Jiao Tong University. His main research interests include synthetic aperture radar (SAR) and polarimetric SAR image understanding and information mining, particularly spatial information analysis, segmentation and classification, multitemporal image analysis, target detection and recognition, and multisensor data joint interpretation.

Dr. Liu is a member of the IEEE Geoscience and Remote Sensing Society. He is a Reviewer of several journals, including the *Chinese Journal of Electronics* (both in Chinese and English), the *Canadian Journal of Remote Sensing*, *Remote Sensing Letters / International Journal of Remote Sensing*, IET RADAR, SONAR & NAVIGATION, the IEEE SIGNAL PROCESSING LETTERS, the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, and the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.



**Boying Li** received the B.Sc. degree in automation engineering from Northwestern Polytechnical University, Xi'an, China, in 2016. Since September 2016, she has been working toward the Ph.D. degree in the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China.

Her research interests include synthetic aperture radar image interpretation (ship detection and classification) and machine learning.



**Weiwei Guo** received the B.Sc, M.Sc., and Ph.D. degrees in information and communication engineering from the National University of Defense Technology, Changsha, China, in 2005, 2007, and 2014, respectively. He was a joint Ph.D. student at Queen Mary, University of London, London, U.K., from 2008 to 2010.

Since 2015, he has been a Post-Doctorate Researcher with Shanghai Jiao Tong University. He has coauthored many journal publications in a number of scientific journals, including the IEEE TRANSACTION ON IMAGE PROCESSING, *Pattern Recognition*, etc. His main research interests include areas of image and signal processing, computer vision, and pattern recognition.



**Lanqing Huang** received the B.Sc. degree in communication engineering from Xi'an Jiao Tong University, Xi'an, China, in 2015. Since September 2015, she has been working toward the Ph.D. degree in the Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China.

Her research interests include synthetic aperture radar (SAR) image interpretation, particularly SAR marine surveillance, SAR observation of ocean, and statistical analyses of SAR image.
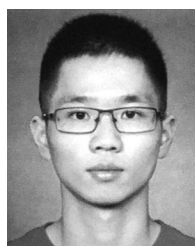


**Wenhao Yu** received the B.Sc. degree in automation engineering in 2016 from Shanghai Jiao Tong University, Shanghai, China, where he has been working toward the M.Sc. degree in the Department of Electronic Engineering since Sep. 2016.

His research interests include synthetic aperture radar image interpretation and deep learning.

**Zenghui Zhang** (M'13) received the B.Sc. degree in applied mathematics, the M.Sc. degree in computational mathematics, and the Ph.D. degree in information and communication engineering from the National University of Defense Technology (NUDT), Changsha, China, in 2001, 2003, and 2008, respectively.

From 2008 to 2012, he was a Lecturer in the Department of Mathematics and System Science, NUDT. He is currently an Associate Professor in the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China. His main research interests include radar signal processing and compressed sensing theory and applications.

**Wenxian Yu** was born in Shanghai, China, in 1964. He received the B.Sc. degree in radio measurement & control and data transmission, the M.Sc. degree in communication and electronic system, and the Ph.D. degree in communication and information system from the National University of Defense Technology (NUDT), Changsha, China, in 1985, 1988, and 1993, respectively.

From 1996 to 2008, he was a Professor in the College of Electronic Science and Engineering, NUDT, where he served as the Deputy Head of the College and the Assistant Director of the National Key Laboratory of Automatic Target Recognition. He is currently in the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, where he is a Yangtze River Scholar Distinguished Professor and the Head of Research and was the Executive Dean from 2009 to 2011. His current research interests include radar target recognition, remote sensing information processing, multisensor data fusion, integrated navigation system, etc. In these areas, he has published more than 200 research papers.