

Extracting Ship's Size from SAR Images by Deep Learning



Yibin Ren, Xiaofeng Li, and Huan Xu

1 Introduction

Ship detection is very important to marine transportation [5]. Space borne Synthetic Aperture Radar (SAR) has been one of the most critical data source for ship detection because it can penetrate the clouds and track objects in all kinds of weather [28]. In marine applications, ship recognition from SAR imagery has long been a hotspot [4, 9, 19]. With the advancement of image analysis technology, SAR images can be used to derive more detailed ship information [8]. The size of a ship provides basic information for ship classification [11]. And the size information can provide useful information for ship classification. The intricate geometric parameter estimate is also a part of the interpretation of SAR image. A method for extracting ship size that is both efficient and precise will bring a new concept for SAR image interpretation.

Ships, in general, are metallic objects that may reflect SAR sensor electromagnetic radiation significantly more strongly than the surrounding ocean. On SAR images, one ship can be identified as a bright back scattering intensity target with high normalized radar cross-section (NRCS) values. The minimum bounding rectangle (MBR) is a geometric characteristic of the ship's NRCS that offers a preliminary size for determining a ship's ground size. In the meantime, the ship's superstructure, sea-ship interaction, and imaging conditions all have an effect on the NRCS. Li et al. [11]. These factors lead to a large gap between the initial size and the ground size. Figure 1 shows several examples of ship's signature on SAR images, the size of the MBR,

Y. Ren · X. Li (✉)

CAS Key Laboratory of Ocean Circulation and Waves, Institute of Oceanology,
Chinese Academy of Sciences, Qingdao 266071, China
e-mail: lix@qdio.ac.cn

Key Laboratory of Earth Observations of Hainan Province, Sanya 572029, China

H. Xu

School of Geomatics and Marine Information, Jiangsu Ocean University,
Lianyungang 222005, China

© The Author(s) 2023

X. Li and F. Wang (eds.), *Artificial Intelligence Oceanography*,
https://doi.org/10.1007/978-981-19-6375-9_15

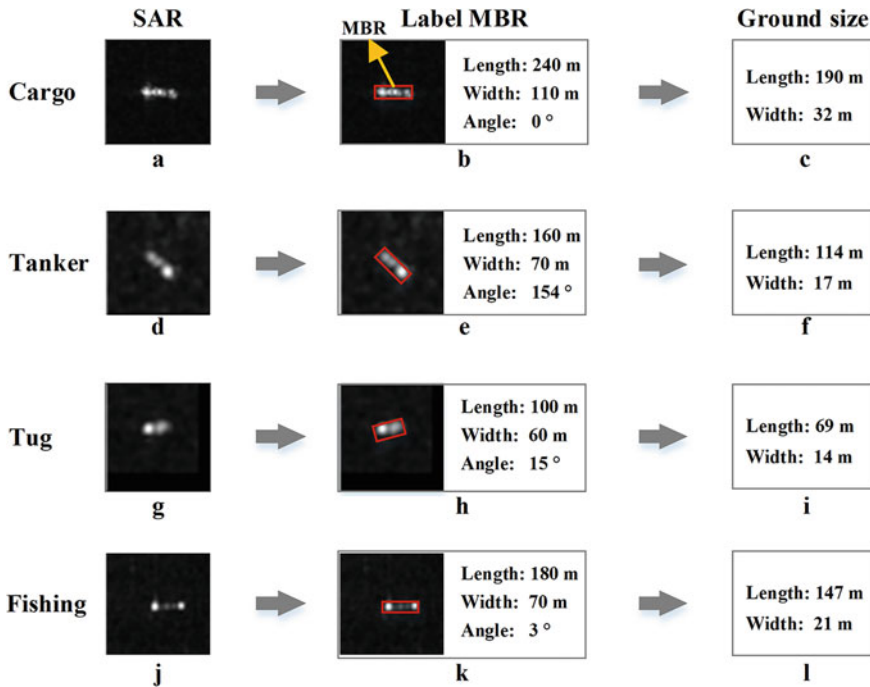


Fig. 1 Examples of ships on SAR images. **a/d/g/j** ship signature on SAR image; **b/e/h/k** labeled MBR of ship signature and the MBR's size; **c/f/i/l** the ship's ground size

and the ground size of the ship. The MBR is labeled by visual interpretation. The difference between the MBRs and the ground size appears to be clear. As a result, precisely extracting ship size from SAR images is difficult.

2 Traditional Methods

2.1 Typical Procedure of Traditional Methods

The majority of classic techniques for extracting ship size from SAR images have three stages (Fig. 2): (1) binarization, (2) initial size extraction, and (3) accurate size estimation. Binarization divides the pixels in the SAR image into two groups: ship signatures and non-ship signatures. The binary result is then converted into an MBR in the second phase. The length and width of the created MBR are used to determine the ship's starting size. Finally, a regression model is used to determine the accurate ship size using the initial size and other relevant factors such as the maximum and minimum NRCS of the ship signature. Statistical/machine learning

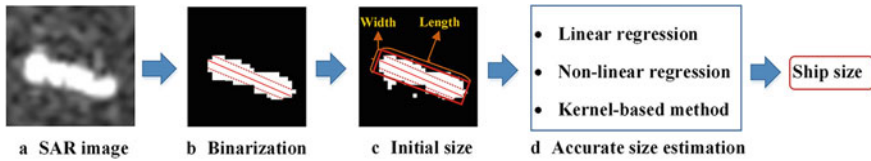


Fig. 2 The procedure of the traditional algorithm for ship size extraction from SAR images

(ML) methods, such as linear regression, non-linear regression, and kernel-based methods, are commonly used in regression models.

2.2 Representative Traditional Methods

Stasolla and Greidanus [26] used Constant False Alarm Rate (CFAR) to binary the SAR image. CFAR is a common method [21, 29, 30] that separates ship signatures and backgrounds. Further, to extract the ship's MBR, they used the mathematical morphology method to refine the signature. They adopted the MBR's length and width as the ship's final length and width without a third step. They tested their model with 127 available ship samples from Sentinel-1 images. The mean absolute error (MAE) of length is 30 m (relative error 16%), and the MAE of width is 11 m (relative error 37%). In 2018, Li et al. [11] estimated the ship's size of the OpenSARShip [7]. The ship signature was obtained using a threshold-based approach. They use an image segmentation procedure to refine the ship signature and determine the original ship size. Finally, the gradient boosting model is employed to estimate the accurate ship size. The MAE of the length and width, according to experiments, is 8.80 m (relative error 4.66 percent) and 2.17 m (relative error 7.01%), respectively.

2.3 Issue to be Further Addressed

The accuracy of ship size extraction is improving as years roll on. The standard three-step procedure is quite complicated. Binarization and initial size extraction need advanced image processes in order to meet the next estimation stage [11]. The third stage is similarly difficult [20]. The inaccuracies caused in each stage will add up and eventually compromise the accuracy of the final size extraction. It is possible to build new approaches to increase ship size extraction accuracy and efficiency in the era of big data.

Deep learning (DL), as the cutting-edge AI technology, has made great achievements in computer vision [10]. Multiple neural network layers make up a typical DL model. It accepts raw data as input and learns the essential characteristics automatically to perform classification or prediction [25]. End-to-end learning is the term

for this process. DL simplifies feature engineering and is well suited to modeling massive data and complex interactions when compared to traditional machine learning. DL has been successfully employed in oceanography, geography, and remote sensing in recent years [12, 13, 22, 24, 31, 32]. DL proposes novel approaches to the problem of estimating the size of a ship.

3 Deep Learning Method

3.1 Ship Detection Based on DL

A deep convolution neural network (CNN) is a subtype of DNN that is made up of CNN layers. CNN-based models have had a lot of success in target detection. Researchers proposed CNN-based ship detection models, such as models based on faster region-based convolutional network (Faster-RCNN) [23], single-shot multi-box detector (SSD) [15], and you only look once (YOLO) [2]. Orientation is an important characteristic of a ship. Several researchers suggested a rotatable bounding box (RBB) to replace the usual non-rotating RBB, such as DRBox [14] and DRBox-v2 [1].

For the ship detection task, DL has become the first choice. DL-based models achieve end-to-end detections with higher accuracy and robustness over conventional models. However, for ship size extraction, there is almost no application of deep learning. Therefore, developing an end-to-end DL model is necessary.

3.2 SSENet: A Deep Learning Model to Extract Ship Size from SAR Images

SSENet is a new end-to-end DL model that replaces the previous three-step process for extracting ship size from SAR data. The model uses DRBox-v2 to create the ship's RBB from the SAR image and a DNN-based regression model to estimate the accurate ship size. The DNN-based regression model is proposed using a hybrid input and a loss function termed mean scaled square error (MSSE), which considerably increases ship size estimation accuracy.

3.2.1 Overall Structure of SSENet

SSENet's overall structure consists of three phases (Fig. 3): (1) RBB generation; (2) accurate ship size estimation; (3) MSSE loss calculation and overall model optimization. The SAR chip is used as input in the first stage, which uses a deep CNN model called DRBox-V2 to automatically detect the ship's RBBs. The RBB with

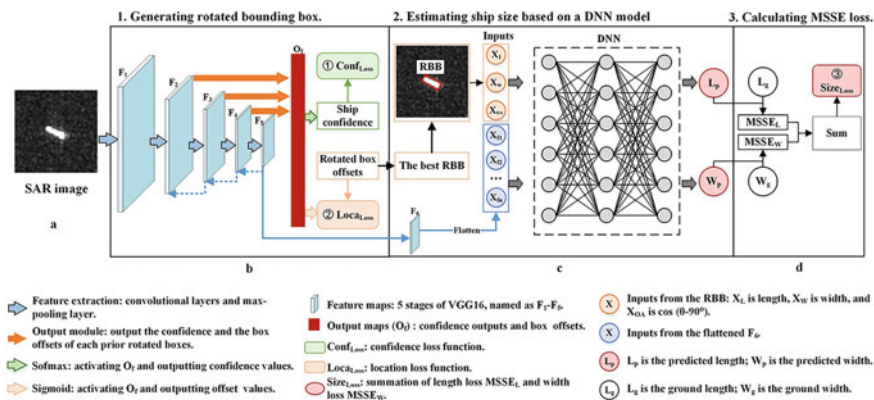


Fig. 3 Structure of SSENNet

the highest confidence is chosen as the initial RBB. A DNN model is used in the second stage to estimate ship size. The DNN model takes two types of data as inputs: (1) the initial length, width, and orientation angle, and (2) the SSD feature map. The accurate ship's length and width are generated using the DNN model.

3.2.2 Generating RBB for the Ship

The DRBox-v2 is used to generate RBB for the ship [1]. Its input is a 300×300 pixels SAR image, and its output is a series of RBBs. DRBox-v2 contains two sub-modules: a feature extraction module and an output module. The feature extraction module extracted abstracted features. Here, the VGG16 is employed as the feature extraction module. The VGG16 consists of five feature extraction units. Two stacking CNN layers make up the first feature extraction unit, while a max-pooling layer and two stacking CNN layers make up the others. Each feature extraction unit produces a three-dimensional feature map as its output. Five feature maps named F_1, F_2, \dots, F_5 are generated. The number of channels in the F_1-F_5 feature maps is 64, 128, 256, and 512. The pooling kernel is 2×2 . After on max-pooling layer, the spatial size of a feature map is downscaled as $1/2$ size of its original size. As the input SAR image is 300×300 pixels, the spatial size of F_1-F_5 feature maps is $300 \times 300, 150 \times 150, 75 \times 75, 38 \times 38,$ and 19×19 pixels.

The output module generates output maps by convolutioning feature maps O_f , Fig. 2b. There are two outputs for one SAR image: the confidence of being a ship, as well as the geographic offsets of prior RBBs. A softmax function activates the O_f to obtain the confidence output. A sigmoid function activates the O_f to obtain the location offsets. Three feature maps ($F_2, F_3,$ and F_4) are fused to generate O_f . FPN is used to combine different feature maps. The cross-entropy and the smooth L1 loss [15] are used as the confidences loss and geographic loss for DRBox-v2.

Following the first process, a ship's candidate RBBs are collected, providing beginning references for the future exact size estimation.

3.2.3 Estimating Ship Size Based on a DNN Model

There are two elements to the DNN model's inputs, as shown in Fig. 3c. The initial ship size and orientation angle, which are determined from the best RBB and give primary and direct information for correct ship size regression, are the first part. The DRBox-v2 generates a sequence of ship RBBs. As the best RBB, the RBB with the highest confidence value is chosen. The initial ship size is the length and width of the best RBB. Furthermore, the best RBB's orientation angle is the ship's orientation angle, as shown in Fig. 4. It has an impact on the SAR image's ship signature [7, 11]. As the orientation does not distinguish between the bow and the stern of one ship, we transform the angle's range to $(-90^\circ, 90^\circ]$.

The other component of the inputs is the feature map derived from the input SAR image. In typical environmental conditions, the ship's signature in the SAR image reflects the sea clutter. It indicates whether the ship is moving or stationary. During the SAR integration time, a moving target is frequently found in several resolution cells. Smearing and brightness loss in the SAR image are caused by the dispersion of backscattered energy. A moving ship's signature reveals an azimuth displacement. The SAR system receives the Doppler signal from the scatter in the azimuth direction. A stationary ship's azimuth position is identical to the azimuth position of a SAR platform. The Doppler shift, on the other hand, has an extra component for a moving ship, resulting in an azimuth change in the ship signature. The environmental conditions during satellite imaging, such as wind fronts, ocean waves, and rain cells, alter the ship's signature on the SAR image. Under typical conditions, the sea-ship interaction produces a complicated ship motion in the real world and a polarimetric

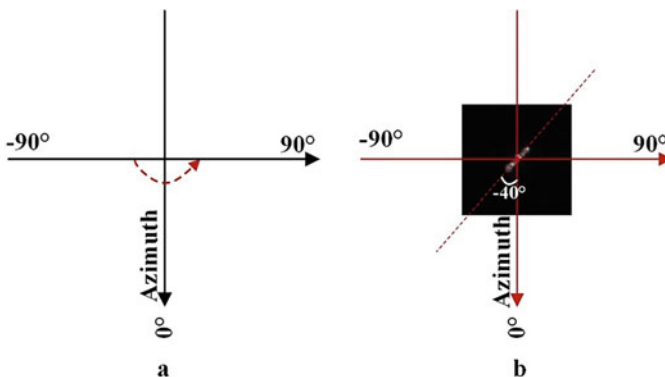


Fig. 4 Illustration of the ship orientation. **a** Coordinate system; **b** An example of a ship chip

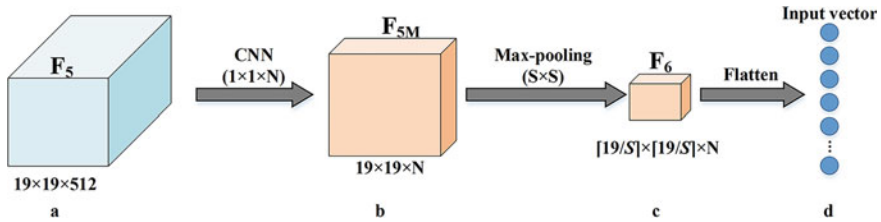


Fig. 5 Transforming the feature map F_5 as inputs. **a** F_5 feature map. **b** Compressing F_5 in the channel dimension and obtain the F_{5M} . **c** Compressing F_{5M} in the spatial dimension and obtain F_6 . **d** Flattening F_6 as one-dimensional input vector

scattering signature with a wide range of polarimetric scattering processes [14, 16, 17]. In reference [11], the relationship between the status of the ship, the surroundings, and the ship's size has been demonstrated. The abstracted feature map derived from the input SAR image contains the factors stated above. Therefore, the feature map F_5 in Fig. 3b is employed as the other component of the input.

F_5 is a three-dimensional feature map with 512 19×19 pixels channels. The input vector contains 184,832 ($512 \times 19 \times 19$) elements, which brings training difficulties for the fully connected DNN regression model. It is necessary to make some transformations to reduce the dimension of F_5 .

As shown in Fig. 5a, b, we transform F_5 by a CNN layer with $1 \times 1 \times N$ convolutional kernels, obtaining F_{5M} . Compared with F_5 , the channel number of F_{5M} is reduced from 512 to N , Fig. 4b. F_5 is compressed in channel dimension. Then, an S size max-pooling is performed on the new feature map F_{5M} , and a new feature map F_6 is obtained, Fig. 5c. The spatial size of the F_6 is $\lceil 19/S \rceil$. The values of N and S are defined by experiments. Finally, F_6 is flattened as a one-dimensional feature vector. The flattened vector is concatenated with the initial width, length, and orientation to form the inputs of the DNN model, Fig. 3c.

As shown in Fig. 3c, to perform regression, three hidden NN layers are used. There are 256 neurons in each NN layer. The parameter-tuning experiment produces the number of hidden NN layers and the number of neurons. The rectified linear unit is the activation function of each layer. Two neurons are stacked on the last hidden NN layer to form an output layer. A sigmoid function is stacked one the output layer to transform the estimated values to 0–1 and output the estimated width W_p and the estimated length L_p , Fig. 3c.

3.2.4 Calculating MSSE Loss and Optimizing SSENet

The MSSE loss function is used in the DNN regression model. For most regression issues, the mean square error (MSE) is a commonly used loss function. The definition of MSE is shown in Equation (1): y_i represents the ground truth, y'_i represents the prediction value, and N means the number of values to be predicted. The loss value calculated by MSE and the ground truth value have no relation. Assume a ship's

ground length and width are 100 and 50 m, respectively, and the predicted length and width are 80 and 30 m, respectively. Both the length and width MSE values are 400. Because the model is optimized based on loss values, both the length and width losses contribute equally to the model's optimization. In practice, a ship's length is much greater than its width. In most cases, the length is more concerning than the width. In order to increase the length estimate accuracy, we hope that the length loss helps to optimize the model more than the width loss.

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - y'_i)^2 \quad (1)$$

$$MSSE = \frac{1}{N} \sum_{i=1}^N y_i \cdot (y_i - y'_i)^2 \quad (2)$$

$$Size_{Loss} = MSSE_L + MSSE_W \quad (3)$$

MSSE loss function solves the mentioned issue. MSSE incorporates the ship length and width ground truth into the traditional MSE. The ground truth is utilized as a dynamic parameter to scale the square error. The definition of MSSE is shown in Eq. (2): y_i , y'_i and N is the number of all samples. The MSSE length and width losses in the example are 40,000 and 20,000, respectively. The loss in length is substantially greater than the loss in width. As a result, the penalty for the model's length will be increased during the training phase. Therefore, the optimization procedure is more conducive to length estimation. Based on Eq. (2), the loss of length $MSSE_L$ and the loss of width $MSSE_W$ are calculated. The size loss ($Size_{Loss}$) is the summation of $MSSE_L$ and $MSSE_W$, Eq. (3).

Besides $Size_{Loss}$, the confidence loss ($Conf_{Loss}$) and the location loss ($Loca_{Loss}$) are another two losses calculated in the first stage, Fig. 3b. $Conf_{Loss}$ is the cross-entropy loss, and $Loca_{Loss}$ is the smooth L1 loss [1, 23]. Their definitions are as follow:

$$Conf_{Loss} = \sum_{i=1}^N c_i \log c'_i + (1 - c_i) \log (1 - c'_i) \quad (4)$$

$$Loca_{Loss} = \frac{1}{N} \sum_{i=1}^N \text{smooth}_{L1}(x_i) = \begin{cases} 0.5x_i^2, & \text{if } |x| < 1 \\ |x| - 0.5, & \text{otherwise} \end{cases} \quad (5)$$

where N is the number of predicted targets, c_i is the ground confidence of a sample, c'_i is the predicted confidence of a sample, and x_i is the element-wise difference between the ground RBB and the predicted RBB. The three losses, $Size_{Loss}$, $Conf_{Loss}$, and $Loca_{Loss}$, are added to form the final loss that optimizes SSENNet integrally.

3.3 Experiments on SSENet

3.3.1 Experiments Data

The OpenSARShip dataset (<http://opensar.sjtu.edu.cn/>) is a Sentinel-1 ship interpretation dataset that includes 11,346 SAR ship chips and automatic identification system (AIS) messages. The ground size for each ship is provided via the AIS. The ground range detected (GRD) of IW is the picture mode of Sentinel-1. The spatial resolution of the SAR image is around 20m, with a pixel spacing of 10 m. SNAP 3.0 performs radiometric calibration and terrain correction. The amplitude values of pixels for VH (vertical emitting and horizontal receiving) and VV (vertical emitting and vertical receiving) polarizations are stored on each SAR chip, which has one ship and two channels. The experiment set for SSENet includes 1,890 samples in the VV mode. Figure 6 shows the distributions of ground ship’s length and width. The length ranges from 28 to 399 m. The width ranges from 6 m to 65 m. Each SAR chip is 300×300 pixels in size. We transform the values of SAR images to $[0, 255]$. The training set consisted of 1,500 SAR chips chosen at random. The remaining 390 chips will be used for testing.

The ground truths for the experimental set include two parts: the ground ship size and the RBB for each ship. The ground size is obtained from the OpenSARShip. The RBB for each ship is labeled manually by a Matlab tool shared in DRBox-v2. The DRBox-v2 is trained to generate accurate RBB based on the ground RBB.

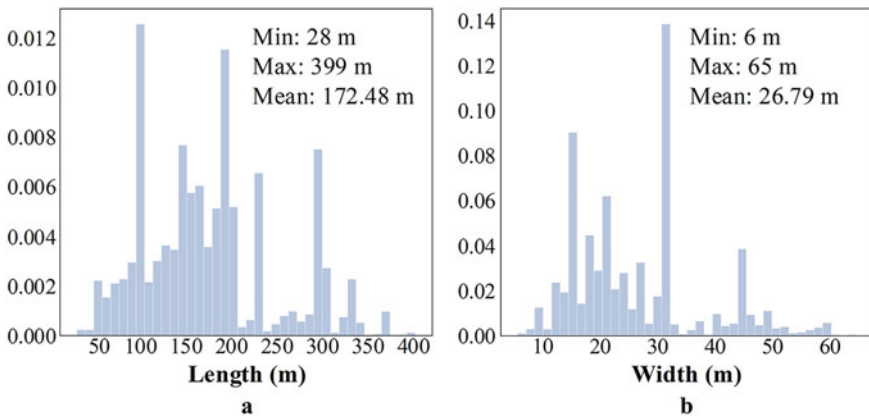


Fig. 6 The range of length and width of the testing set

3.3.2 Experiments Setting

A workstation with one GeForce RTX 2070 8GB GPU is used in the experiment. Python 3.6 is the programming language used. TensorFlow is a deep learning package. For training, the batch size is six. 0.0002 is the initial learning rate. The learning rate reduces by half every 5,000 training epochs during the training procedure. When the $Size_{Loss} < 0.001$, the $Loca_{Loss} < 0.005$, and the composite loss < 0.01 , the training procedure stops.

MAE and the mean absolute percentage error (MAPE) are employed as metrics. MAE is a typical absolute error, and MAPE is a widely used relative error. Assuming y_i is the ground truth, y'_i is the estimation value, and N is the number of samples, the definitions of MAE and MAPE are as follow:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - y'_i| \quad (6)$$

$$MAPE (\%) = \frac{100}{N} \sum_{i=1}^N \left| \frac{y_i - y'_i}{y_i} \right| \quad (7)$$

3.3.3 Performance of SSENet

The hyper-parameters of SSENet are determined by parameter tuning, and a well-trained model is picked up to be evaluated. The 390 samples of the testing set are fed into the well-trained SSENet. The outputs are the scaled lengths and widths estimated by the model. The scaled values are rescaled to normal values.

The estimated ship sizes are shown in Fig. 7a, b. The length and width MAEs are 7.88 and 2.23 m, respectively. The MAEs of the estimated length and width are pushed under 0.8-pixel spacing. The MAPE of estimated length and width are 5.53 and 8.93%, respectively. The R^2 score are 0.9773 and 0.9093. This indicates that the estimated ship length/width is quite close to the ground length/width. The R^2 score of widths is smaller than that of length, which means the width is difficult to estimate than the length. There are two factors that contribute to this phenomena. A ship's width is far smaller than its length. The width of the ship's signature on the SAR image is more ambiguous than the length [26], which causes random errors in the width of the labeled RBB. Second, the MSSE loss function makes the model fit the length better.

We plot the relationship between the labeled RBB's size and the ship's ground size, as shown in Fig. 7c, d. The labeled RBB is treated as the RBB closest to the ship's signature for visual interpretation. As shown in Fig. 7c, the MAE of length is nearly 40 m, and the MAE of width is more than 50 m. The gap between the labeled RBB's size and the ground size is large. By adding the regression model, SSENet pushes the MAEs under 8 m. Therefore, the proposed regression model based on

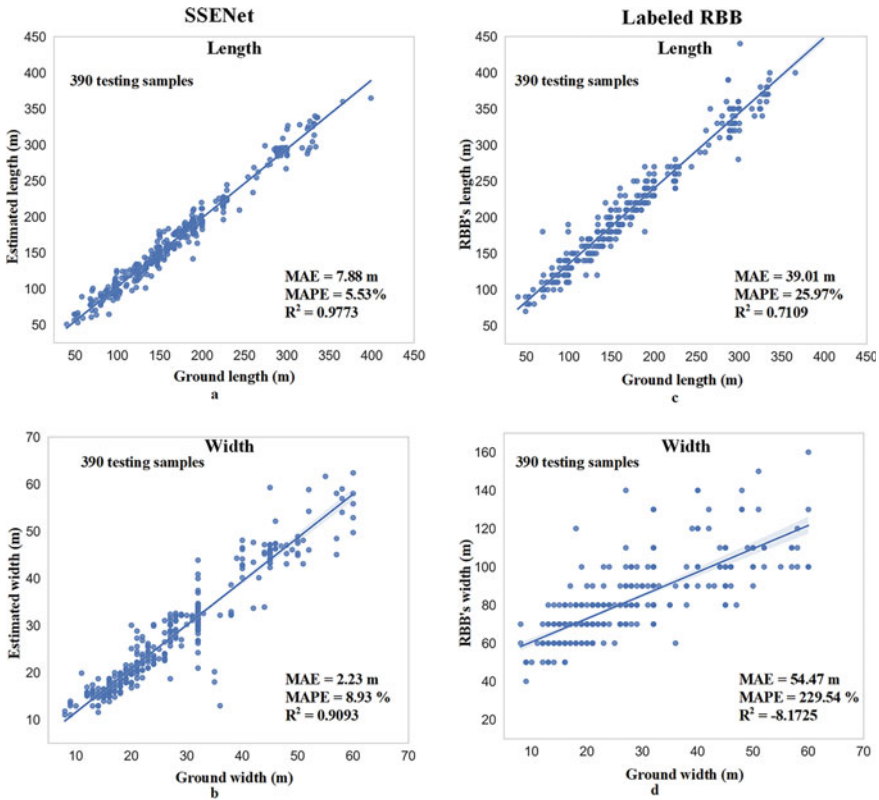


Fig. 7 Relationships between the SSENet’s estimated size, and the size of the labeled RBB. **a** and **b** The relationships between the ground size and the SSENet’s size. **c** and **d** The relationships between the ground size and the labeled RBB’s size

DNN is necessary and effective. Figure 8 shows some examples of SSENet’s results. The outputs of one sample include the detected RBB, the confidence score to be a ship, and the estimated ship size. For most ship samples, the estimated sizes are consistent with the ground sizes.

3.3.4 Effectiveness of the Inputs

The efficiency of the inputs for the DNN regression model is tested. The results are shown in Table 1. Three compared models employ different inputs. The inputs for SSENet₁ include initial ship size, without feature map F_6 . For SSENet₂, the inputs are initial ship size and F_6 . Based on the three inputs, SSENet₃ adds the initial orientation as another input.

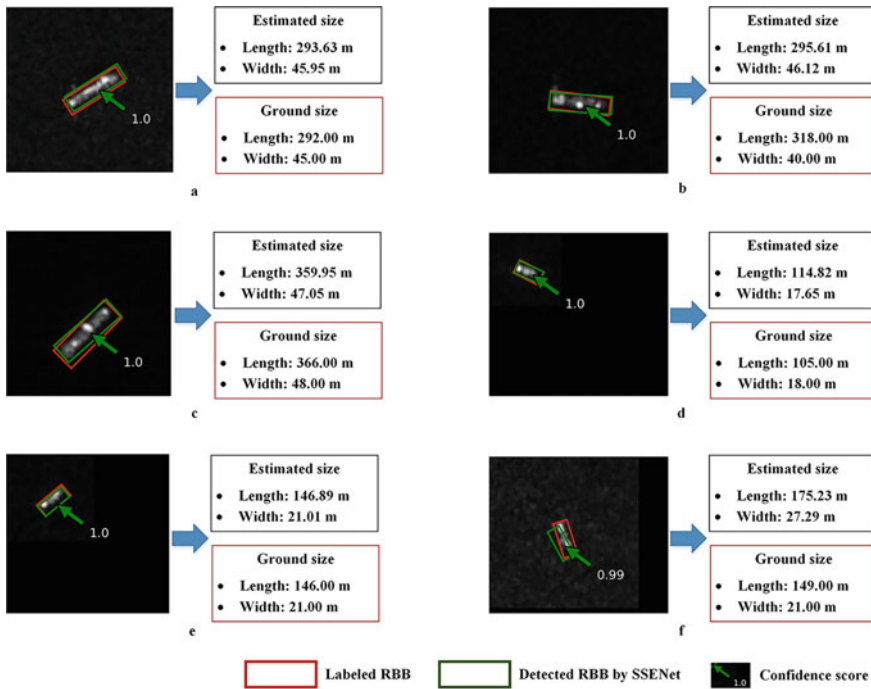


Fig. 8 Some examples of SSENNet, the outputs include the detected RBB, the confidence score to be a ship, and the estimated ship size

Table 1 Model performance with different inputs

Models	Inputs of DNN	MAE (m)		MAPE (%)	
		Length	Width	Length	Width
SSENNet ₁	X_l, X_w	10.01	2.65	6.87	9.89
SSENNet ₂	X_l, X_w, F_6	8.14	2.27	5.82	9.22
SSENNet ₃	$X_l, X_w, \cos \theta, F_6$	7.88	2.23	5.53	8.93

The results are displayed in Table 1. SSENNet₁ obtains the largest MAE and MAPE among the three models. By adding F_6 , SSENNet₂ reduces the length’s MAE about 2 m compared with SSENNet₁. This finding illustrates that the feature map of a SAR image is an important input for estimating ship size. Adding the feature map as an input improves the accuracy of size estimation. Finally, by explicitly including the ship’s initial orientation as another input, the estimation errors are significantly minimized. Therefore, each element of the inputs for SSENNet shows contributions to the final size estimation. Figure 8 shows several results of SSENNet₃, and the red/green rectangle is the labeled/detected RBB. The estimated confidence score to be a ship and estimated the size by SSENNet are also displayed.

Table 2 Performance of with MSE or MSSE

Models	Loss Function	MAE (m)		MAPE (%)	
		Length	Width	Length	Width
SSENet _{MSE}	MSE	8.85	2.20	5.99	8.81
SSENet _{MSSE}	MSSE	7.88	2.23	5.53	8.93

3.3.5 Effectiveness of MSSE Loss

An experiment is conducted to test the effectiveness of the new loss function, MSSE. The results are shown in Table 2. SSENet_{MSE} is the model with MSE loss. SSENet_{MSSE} is the model with MSSE loss. The other parts of the two models are the same.

The results are shown in Table 2. The length MAE of SSENet_{MSSE} is nearly 1m less than that of SSENet_{MSE}, reducing by 11%. For the width, SSENet_{MSSE} performs slightly worse than SSENet_{MSE}. The reason for this is that MSSE emphasizes a significant loss and drives the model to focus on length rather than width. The difference in width between the two values, however, is only a few centimeters. The disadvantages of MSSE are not overshadowed by the aforementioned constraint. As a result, our MSSE loss is helpful, particularly when evaluating the ship’s length.

4 Discussions

4.1 ML versus DL

SSENet’s regression model is a DNN model. We choose three typical ML models, Gradient Boosting Regression (GBR) [6], Support Vector Regression (SVR) [3], and Linear Regression (LR) [18] to discuss their performances. GBR and SVR are applied in ship size extraction [8, 11]. LR is a baseline model [27]. Because these three ML models aren’t NN-based, they can’t be combined with the SSD to create an end-to-end model. The SAR images cannot be fed into the three ML models. The inputs for these three models are the initial ship size and orientation of the labeled RBB. The parameters of GBR, SVR, LR are tuned and the estimation results with the best metrics are recorded. The DNN model is used by SSENet.

The results are shown in Table 3. The result of SSENet is in the last row. GBR performs the best among four models (LR, SVR, GBR, and DNN). GBR is an ensemble learning model with good performance in the three-stage procedure [26]. However, GBR is unable to extract features from SAR images automatically. GBR cannot be combined with a DL-based ship detection model, such as DRBox-v2, to create an integrated ship size extraction model. The premise of using GBR is that the SAR image should be binarized accurately, and the initial RBB is well extracted by tra-

Table 3 MAE and MAPE of different models

Models	Inputs	MAE (m)		MAPE (%)	
		Length	Width	Length	Width
LR	RBB	10.60	2.86	6.97	10.47
SVR	RBB	10.16	2.48	6.61	9.50
GBR	RBB	9.69	2.39	6.93	9.27
DNN	RBB	10.79	2.83	7.27	10.22
SSENet	Images	7.88	2.23	5.53	8.93

ditional methods. As stated in Sect. 2, the traditional method faces big challenges. Practically, GBR is not an end-to-end model: feeding the SAR image and obtaining the ship size.

The error of DNN model is large. However, a DNN model can be combined with any deep learning models based on CNN or NN to extract size from the SAR image from beginning to end. In contrast to traditional techniques, the DL model optimizes all parameters globally. The DNN regression model can use the feature maps extracted by the DL model to increase the accuracy of the estimated ship size. As shown in Table 3, the SSENet reduces the MAE of length by nearly 2 m compared with the GBR, about 18.68 %. Therefore, compared with traditional methods, the ship size extraction model based on deep learning is more practical.

4.2 Errors's Sources

This section delves into the details of estimation errors and attempts to determine what causes large inaccuracies. The ship's direction and transit speed are two elements that need to be investigated, according to previous research [10, 26].

4.2.1 Ship Orientation

The estimated errors with respect to the ship's orientation angle is displayed. Figure 9a and b show the results of the length. Fig. 9c, d show the results of the width. The MAEs vary with the ship orientation variation. Large MAEs occur when the orientation angles are closer to 0° (0° means the azimuth direction) in the range of (−45°, 45°). The reasons for the above phenomenon include two aspects. The first one is the ship motion. When the ship moves in a direction that is near to the azimuth direction, the azimuth direction's speed component is large. Because of the large component, the ship signature appears to be smeared, increasing the estimation error. The other reason is the unequal resolution during imaging, 5 m × 20 m for range and azimuth directions, respectively. The low resolution in the azimuth enlarges the errors [26].

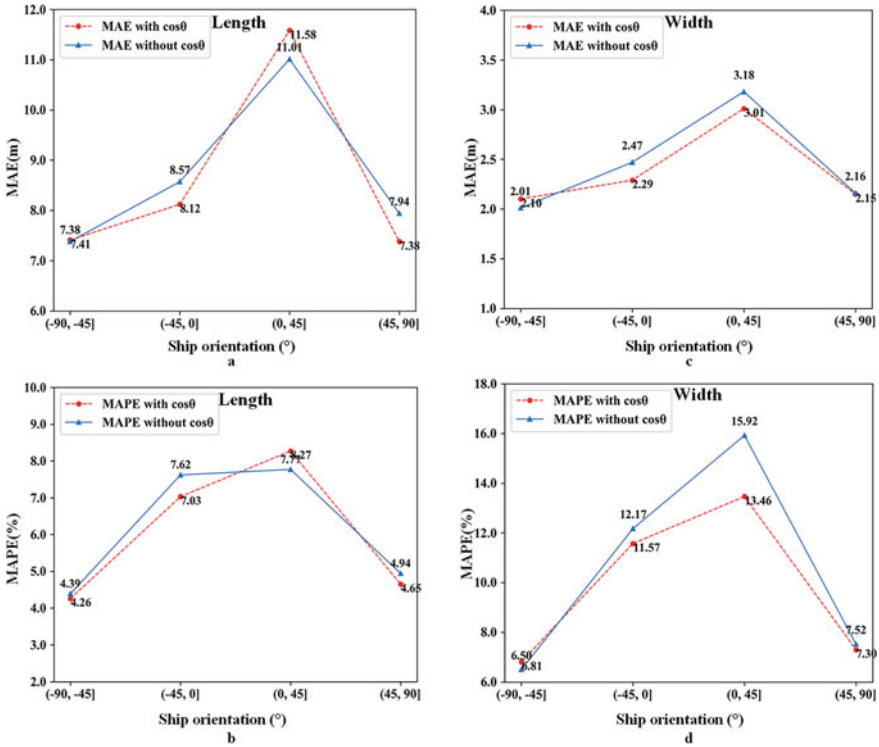


Fig. 9 The trend of errors with respect to the orientation angle. **a** Trend of length's MAE; **b** Trend of length's MAPE; **c** Trend of width's MAE; **d** Trend of width's MAE

As shown in Fig. 9, when the initial orientation angle ($\cos\theta$) is added to the DNN model, the errors are reduced. This finding also proves that using the original orientation angle as an input is valid.

4.2.2 Ship Speed

Figure 10 shows the errors corresponding to the ship's speed. Because the OpenSARShip's SAR images are mostly from ports, around 83% of the ships are still there. Figure 10a shows that the MAEs are small in the range of (0, 1) knot (1.852 km/h). With the increase of ship speed, the MAE fluctuates slightly. When the speed is greater than 15 kn (27.780 km/h), the MAEs increase apparently: 19.04 and 4.71 m. These two values are far greater than those of other speed intervals. The ship's speed cannot be derived from the SAR image signature. Therefore, it is difficult to refine the estimated sizes of ships by pre-input the ship's speed parameter.

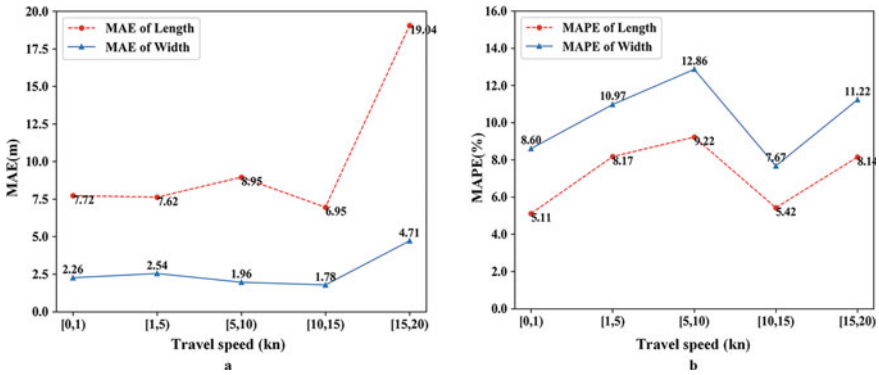


Fig. 10 Figure 10. The trend of errors with respect to the travel speed. a The trend of MAE; b Trend of MAPE

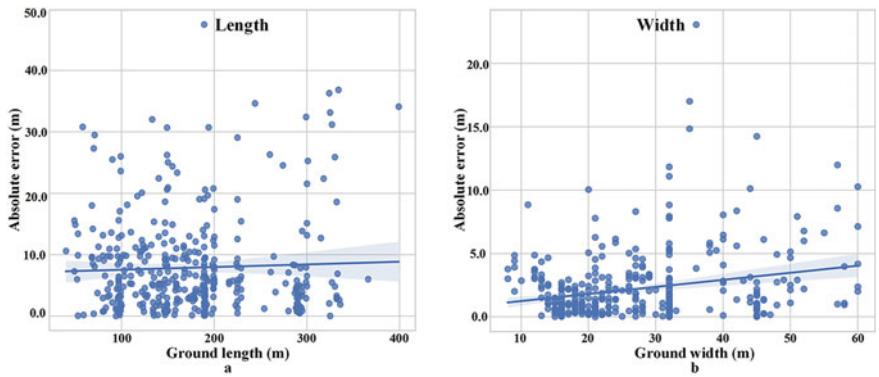


Fig. 11 The distribution of AE with respect to the ground size. a The AE of length; b The AE of width

4.2.3 Ship Size

Figure 11 shows the absolute error (AE) of each estimated and the ground size. The AE of a estimated size takes the absolute value of the difference between the predicted value and the true value. As shown in Figs 11a and b, there are no obvious relationships between AE and ground size. Therefore, the ship size is not a source of errors.

5 Conclusions

SSENet, a DL-based model for extracting ship size from SAR data, is proposed in this chapter. A DNN-based regression model and an SSD-based model make up the SSENet. The DNN model is fed the initial ship size and orientation angle derived from the RBB, as well as the high-level features extracted from the input SAR image. The OpenSARShip trains and validates SSENet. Experiments show that: (1) the SSENet can straight extract ship size from SAR images with MAE less than 0.8 pixels; (2) the new MSSE loss reduces the length's MAE nearly 1 m than the old MSE loss; (3) SSENet shows obvious advantage over the GBR model; (4) SSNet exhibits robustness over four separate data sets.

Acknowledgements The OpenSARShip is downloaded from “OpenSAR Platform” (<http://opensar.sjtu.edu.cn>).

References

1. An Q, Pan Z, Liu L, You H (2019) DRBox-v2: an improved detector with rotatable boxes for target detection in SAR images. *IEEE Trans Geosci Remote Sens* 57(99):8333–8349
2. Chang YL, Anagaw A, Chang L, Wang YC, Hsiao CY, Lee WH (2019) Ship detection based on YOLOv2 for SAR imagery. *Remote Sens* 11(7):786
3. Drucker H, Burges C, Kaufman L, Chris JC, Kaufman BL, Smola A, Vapnik V (1997) Support vector regression machines. *Adv Neural Inf Process Syst* 28(7):779–784
4. Eldhuset K (1996) An automatic ship and ship wake detection system for spaceborne SAR images in coastal regions. *IEEE Trans Geosci Remote Sens* 34(4):1010–1019. <https://doi.org/10.1109/36.508418>
5. Fingas M, Brown C (2001) Review of ship detection from airborne platforms. *Canadian J Remote Sens* 27(4):379–385. <https://doi.org/10.1080/07038992.2001.10854880>
6. Friedman JH (2001) Greedy function approximation: a gradient boosting machine. *Ann Stat* 29(5):1189–1232
7. Huang L, Liu B, Li B, Guo W, Yu W, Zhang Z, Yu W (2018) OpenSARShip: a dataset dedicated to Sentinel-1 ship interpretation. *IEEE J Sel Top Appl Earth Obser Remote Sens* 11(1):195–208. <https://doi.org/10.1109/JSTARS.2017.2755672>
8. Lang H, Wu S (2017) Ship classification in moderate-resolution SAR image by naive geometric features-combined multiple kernel learning. *IEEE Geosci Remote Sens Lett* 14(10):1765–1769. <https://doi.org/10.1109/LGRS.2017.2734889>
9. Lang H, Xi Y, Zhang X (2019) Ship detection in high-resolution SAR images by clustering spatially enhanced pixel descriptor. *IEEE Trans Geosci Remote Sens* 57(8):5407–5423. <https://doi.org/10.1109/TGRS.2019.2899337>
10. Lecun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436–444. <https://doi.org/10.1038/nature14539>
11. Li B, Liu B, Guo W, Zhang Z, Yu W (2018) Ship size extraction for Sentinel-1 images based on dual-polarization fusion and nonlinear regression: Push error under one pixel. *IEEE Trans Geosci Remote Sens* 56(8):4887–4905. <https://doi.org/10.1109/TGRS.2018.2841882>
12. Li X, Liu B, Zheng G, Ren Y, Zhang S, Liu Y, Gao L, Liu Y, Zhang B, Wang F (2020) Deep learning-based information mining from ocean remote sensing imagery. *Nat Sci Rev*
13. Liu B, Li X, Zheng G (2019) Coastal inundation mapping from bitemporal and dualcllarization SAR imagery based on deep convolutional neural networks. *J Geophys Res: Oceans* 124(12)

14. Liu P, Jin YQ (2017) A study of ship rotation effects on SAR image. *IEEE Trans Geosci Remote Sens* 1–13
15. Liu W, Angelov D, Erhan D, Szegedy C, Reed S, Fu CY, Berg AC (2016) SSD: single shot multibox detector. Springer, Cham
16. Margarit G, Mallorquí J, Fàbregas X (2006) Study of the influence of vessel motions and sea-ship interaction on classification algorithms based on single-pass polarimetric SAR interferometry. In: *IEEE international conference on geoscience and remote sensing symposium*
17. Margarit G, Mallorquí JJ, Fortuny-Guasch J, Lopez-Martinez C (2009) Exploitation of ship scattering in polarimetric SAR for an improved classification under high clutter conditions. *IEEE Trans Geosci Remote Sens* 47(4):1224–1235
18. Montgomery DC, Peck EA, Vining GG (2021) *Introduction to linear regression analysis*. Wiley
19. Ouchi K, Tamaki S, Yaguchi H, Iehara M (2004) Ship detection based on coherence images derived from cross correlation of multilook SAR images. *IEEE Geosci Remote Sens Lett* 1(3):184–187. <https://doi.org/10.1109/LGRS.2004.827462>
20. Pelich R, Longépé N, Mercier G, Hajduch G, Garello R (2015) Performance evaluation of Sentinel-1 data in SAR ship detection. In: *IEEE international geoscience and remote sensing symposium (IGARSS)*, pp 2103–2106 (2015). <https://doi.org/10.1109/IGARSS.2015.7326217>
21. Qin X, Zhou S, Zou H, Gao G (2013) A CFAR detection algorithm for generalized gamma distributed background in high-resolution SAR images. *IEEE Geosci Remote Sens Lett* 10(4):806–810. <https://doi.org/10.1109/LGRS.2012.2224317>
22. Reichstein M, Camps-Valls G, Stevens B, Jung M, Denzler J, Carvalhais N, Prabhat (2019) Deep learning and process understanding for data-driven Earth system science. *Nature* 566(7743):195
23. Ren S, He K, Girshick R, Sun J (2017) Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans Pattern Anal Mach Intell* 39(6):1137–1149
24. Ren Y, Cheng T, Zhang Y (2019) Deep spatio-temporal residual neural networks for road-network-based data modeling. *Int J Geogr Inf Sci* 1–19
25. Schmidhuber J (2015) Deep learning in neural networks: an overview. *Neural Netw* 61:85–117 <https://doi.org/10.1016/j.neunet.2014.09.003>, <http://dx.doi.org/10.1016/j.neunet.2014.09.003>
26. Stasolla M, Greidanus H (2016) The exploitation of Sentinel-1 images for vessel size estimation. *Remote Sens Lett* 7(12):1219–1228. <https://doi.org/10.1080/2150704X.2016.1226522>, <https://doi.org/10.1080/2150704X.2016.1226522>, <https://doi.org/10.1080/2150704X.2016.1226522>
27. Tings B, Da S, Bentes CA, Lehner S (2016) Dynamically adapted ship parameter estimation using TerraSAR-X images. *Int J Remote Sens*
28. Torres R, Snoeijs P, Geudtner D, Bibby D, Davidson M, Attema E, Potin P, Rommen B, Floury N, Brown M, Traver IN, Deghaye P, Duesmann B, Rosich B, Miranda N, Bruno C, L'Abbate M, Croci R, Pietropaolo A, Huchler M, Rostan F (2012) GMES Sentinel-1 mission. *Remote Sens Environ* 120:9–24. <https://doi.org/10.1016/j.rse.2011.05.028>, <https://www.sciencedirect.com/science/article/pii/S0034425712000600>, the Sentinel Missions—New Opportunities for Science
29. Wackerman C, Friedman K, Pichel W, Clemente-Colón P, Li X (2001) Automatic detection of ships in RADARSAT-1 SAR imagery. *Canadian J Remote Sens* 27(5):568–577. <https://doi.org/10.1080/07038992.2001.10854896>, <https://doi.org/10.1080/07038992.2001.10854896>, <https://doi.org/10.1080/07038992.2001.10854896>
30. Wang C, Liao M, Li X (2008) Ship detection in SAR image based on the Alpha-stable distribution. *Sensors* 8(8):4948–4960. <https://doi.org/10.3390/s8084948>, <https://www.mdpi.com/1424-8220/8/8/4948>
31. Zhang X, Li X (2020) Combination of satellite observations and machine learning method for internal wave forecast in the Sulu and Celebes Seas. *IEEE Trans Geosci Remote Sens* (99):1–11
32. Zheng G, Li X, Zhang RH, Liu B (2020) Purely satellite data-driven deep learning forecast of complicated tropical instability waves. *Sci Adv* 6(29):eaba1482

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits any noncommercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if you modified the licensed material. You do not have permission under this license to share adapted material derived from this chapter or parts of it.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

