Ship Extraction using Post CNN from High Resolution Optical Remotely Sensed Images

Fuqiang Lei[1,2], Wenliang Wang[2], Wei Zhang[2]
1. CSSC System Engineering Research Institute
2. CSSC (ZHE JIANG) Ocean Technology Co., LTD
Zhoushan, China
lei@shiplinker.com, wangwl@shiplinker.com, zhangwei@shiplinker.com

***Abstract*—A post convolutional neural networks (CNN) method is proposed to extract ships from high resolution optical remotely sensed images. It consist of two parts: ship proposal detection and ship extraction based on CNN. The first part aims to locate possible ships through classification of water and no-water, seawater area extraction using mathematical morphology, and ship proposal extraction. Ships are extracted in the second part by implementing a trained CNN on the ship proposals. Experimental results on a high resolution optical image of San Francisco Bay show the efficiency and robustness of the proposed post CNN method.**

Keywords—high resoltuion; optical remotely sensed image; ship extraction; ship proposals; convolutional neural networks (CNN)

I. INTRODUCTION

The ship extraction from remotely sensed images has attracted much attention. It can supervise fisheries and manage marine traffics to ensure its safety. With the development of satellite and intelligence, automatic ship extraction has replaced the traditional manual ship detection.

Ship extraction methods can be divided into two groups by image sources: synthetic aperture radar (SAR) image based and optical image based methods. The SAR image based ones have advantages of all-weather time and big difference of ships and sea; therefore, it has extensively been studied [1, 2]. A constant false-alarm rate (CFAR) detector is a usually used ship extraction algorithm, which assumes a certain background distribution for SAR images ,such as k-distribution, Gamma distribution [3], Gauss distribution [1], and other combination. However, the ship extraction from SAR images also has its limitations, such as low resolution of SAR images, relatively long revisit cycle, and the complicated sea clutter. Optical remotely sensed images have the advantage of high resolution and relatively short revisit cycles, and have more detailed texture, spectral and shape information. Therefore, the ship extraction from optical remotely sensed images were widely studied [4, 5]. Shape features and gray intensity are two mainly used optical image based ship extraction methods [6].

Shape feature based methods use shape and edge information to extract ships. For example, most ships have narrow bow area and parallel hull edges, which are easily detected because of the big difference between ships and water. Many researchers have adopted this idea to detect ships [7, 8]. Due to the big gray intensity distinction between ships and water, gray intensity was exploited to detect ships by segmenting images [9, 10].

With the development of artificial intelligence, neural network has been widely studied in many researches, especially deep learning and deep neural network. Li et al. use a regional proposal network to generate ship candidates from feature maps produced by a deep convolutional neural network. A hierarchical selective filtering layer was proposed to map features in different scales to the same scale space to efficiently detect ships with various scales [11]. Yao et al. proposed a framework for ship detection based on deep convolutional neural networks (CNN). First, the deep CNN is designed to extract features. Then, a region proposal network is applied to discriminate ship targets and regress the detection bounding boxes, in which the anchors are designed by intrinsic shape of ship targets [12]. Yang et al. developed a ship detection algorithm to overcome cluttered scenes and variable appearances of ships. It consists of a simple region proposal network and a deep forest ensemble. The training process of deep forest ensemble is efficient and users can control training cost according to computational resource available [13]. Zuo and Shi we propose a novel ship detection method based on the recent popular convolutional neural networks and the singular value decompensation algorithm. It provides a simple but efficient way to adaptively learn features from remote sensing images [14]. Wang et al. combined CFAR and CNN based on the CFAR global detection algorithm and image recognition with the CNN model. It is more suitable for application to ship detection systems [15]. Tang et al. proposed a ship detection approach using wavelet coefficients extracted from JPEG2000 compressed domain combined with deep neural network (DNN) and extreme learning machine (ELM). Compressed domain is adopted for fast ship candidate extraction, DNN is exploited for high-level feature representation and classification, and ELM is used for efficient feature pooling and decision making [5].

In this study, a post CNN is proposed to extract ships from optical remotely sensed image to improve efficiency and accuracy. The rest of this paper is organized as follows. The proposed method is detailed presented in the next section. Section three shows experimental results, and conclusion is drawn in the last section.

II. SHIP EXTRACTION USING POST CNN

The proposed method consists of two main parts. The first one aims to detect ship proposals, including classification of

water and no-water, seawater area extraction using mathematical morphology, and ship proposal extraction. The second part includes the construction and training of CNN using a large dataset of ship and no-ship, and ship extraction using trained CNN. The flowchart of the proposed method was shown in Figure 1.

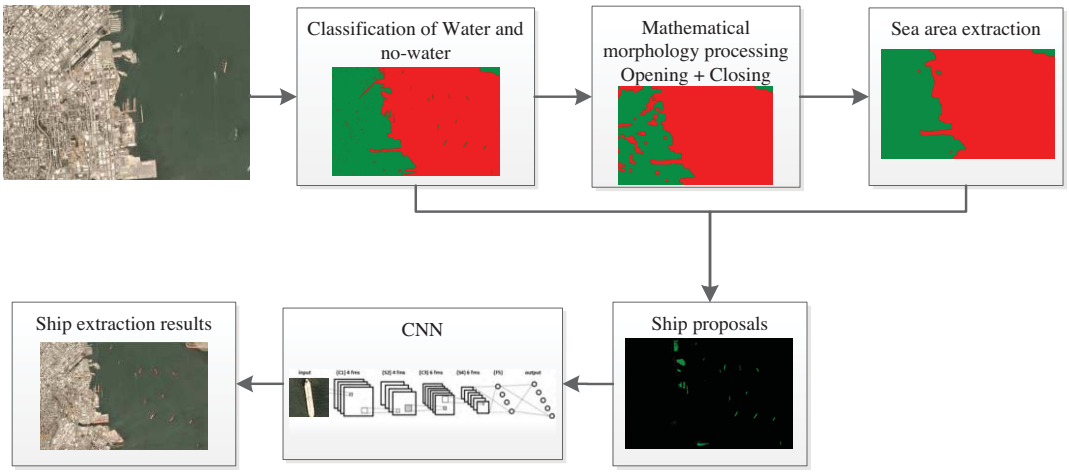

Fig. 1. Flowchart of the proposed method.

A. *Ship proposal extractio*

First, the remotely sensed image was classified with SVM into water and no-water.

The SVM is a nonparametric supervised classifier based on statistical learning theory, which is robust to high dimensional datasets and to ill-posed problems and has been widely used for classification [16]. Assuming a training data set with k sample is given, and it can be represented as $\{\mathbf{x}_i, y_i\}$, $i=1,2,...,k$, where $\mathbf{x}_i$ represents the spectral response of the case, and $y_i \in \{1,-1\}$ is the class label. The optimal separating hyperplane was found that positions the samples of a class in one side of it and makes the distance between it and the closest training samples in both classes as large as possible. The equation $\mathbf{w}\cdot\mathbf{x}+b=0$ represents a hyperplane in feature space, where vector $\mathbf{w}$ is normal to the hyperplane, $\mathbf{x}$ is a point on the hyperplane, and the scalar b is the bias of the hyperplane from the origin. A separating hyperplane was defined by $f(x)=\mathbf{w}\cdot\mathbf{x}+b$. The optimal hyperplane will be achieved by maximizing the margin using the constrained optimization problem

$$\min\left\{\frac{\|\mathbf{w}\|^2}{2}\right\} \tag{1}$$

subject to $y_i(\mathbf{w}\cdot\mathbf{x}_i+b)\geq 1$, $i=1,2,...,k$.

If the data set is not linearly separable, slack variables ξ_i ($i=1,2,...,k$) are introduced to relax the constraints [17]. The equation (1) can be written as

$$\min\left\{\frac{\|\mathbf{w}\|^2}{2}+C\sum_{i=1}^{k}\xi_i\right\} \tag{2}$$

subject to $y_i(\mathbf{w}\cdot\mathbf{x}_i+b)\geq 1-\xi_i$, ($\xi_i\geq 0$, $i=1,2,...,k$)

The input data are mapped into a high dimensional space by valid kernel function k(.,.) to solving a nonlinear separation [18]. The classification decision function in the high dimensional space is then defined by

$$f(x)=\mathrm{sgn}\left(\sum_{i=1}^{k}\alpha_i y_i k(\mathbf{x},\mathbf{x}_i)+b\right) \tag{3}$$

where α_i ($i=1,2,...,k$) is Lagrange multipliers for looking for the optimal separating hyperplane, and is a kernel function.

The Gaussian radial basis function, is adopted as the kernel function due to its interpretability and positive performances [19]. Finally, the optimal hyperplanes are generated for water and no-water classification after training processing.

Second, the classification result was processed by mathematical morphology of opening and closing algorithm. Given that, small water on mainland was removed and seawater area was generated by preserving the water with the largest area in the processed classification result. More details can be find in [20].

Finally, ship proposals were extracted by overlapping the initial classification results and the seawater area. The no-water class on the initial classification result and locating in the seawater area can be seen as ship proposals, which are used as the input layer in CNN of the second part.

B. *Ship extraction using CNN*

A multi-layer neutral network is usually trained through the back-propagation algorithm. Three parts exist in a network, namely input layer, multiple hidden ones and an output one.

CNN is a typical one of neutral network based on multi-layer networks. It has the input, hidden and output layers.

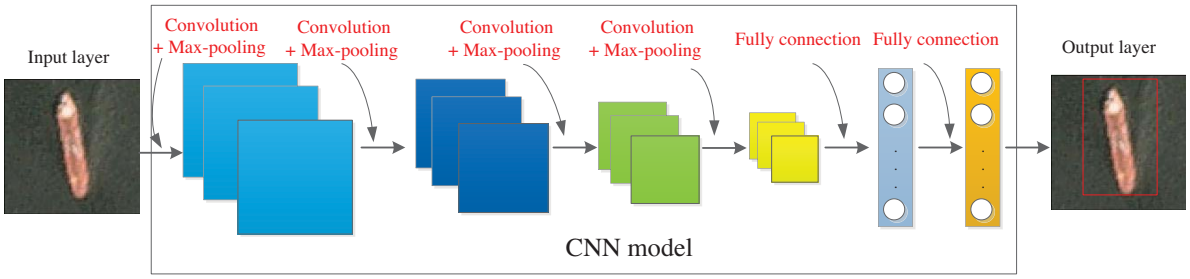

Fig. 2. Flowchart of ship extraction using CNN.

Figure 2 shows the flowchart of the CNN-based ship extraction, where convolution, max-pooling and fully connection are included. The input layer was processed under four times convolution and max-pooling. Each convolution includes 32 kernels with 3*3 size, and all max-pooling sizes are 2*2. The first fully connection includes 512 connections, and the latter one corresponds to ship and no-ship classes using softmax function. Activation functions for all layers except the last fully connection layer were relue.

The CNN model was first trained for ship extraction from a large dataset and then test the trained CNN model. Then the trained CNN model was applied to identify ships from ship proposals extracted in the last section. Proposals extracted by the first part improves the accuracy of CNN because these proposals show the possible locations of ship targets. It also reduces the computation time and improve the efficiency by getting rid of the sliding- window process.

III. EXPERIMENTAL RESULTS

A. *Data description*

The dataset consists of image chips extracted from Planet satellite imagery collected over the San Francisco Bay and San Pedro Bay areas of California. It includes 4000 80x80 RGB images labeled with either a "ship" or "no-ship" classification. Image chips were derived from PlanetScope full-frame visual scene products, which are orthorectified to a 3 meter pixel size [21].

The dataset contains .png image chips, and each individual image filename follows a specific information: label, scene id, longitude, and latitude. The label is 1 or 0, representing the "ship" class and "no-ship" class, respectively. The scene id represents the unique identifier of the PlanetScope visual scene the image chip was extracted from. The longitude and latitude coordinates of the image center point were also included.

The "ship" class consists of 1000 images. Images in this class are near-centered on the body of a single ship. Ships of different sizes, orientations, and atmospheric collection conditions are included. Example images of the "ship" class are shown below.

Fig. 3. Ship samples used to train CNN

The "no-ship" class includes 3000 images. The samples include different land covers, such as water, vegetation, bare earth, buildings, etc. The "partial ships" that contain only a portion of a ship were also included. Additionally, the images were also selected as "no-ship" samples that previously were mislabeled by machine learning models caused by bright pixels or strong linear features. Example images from this class are shown below.

Fig. 4. No-ship samples used to train CNN

B. Evaluation on ship extraction

The high resolution remotely sensed image used to extract ships was obtained from Planet's Open California dataset (https://www.planet.com/products/open-california/), which is an open source image as shown in Figure 5(a). The first part of the proposed post CNN method was implemented using ENVI and the results were shown in Figure 5. Figure 5(b) is the initial classification result generated by SVM, where red and green mean water and no-water respectively. Two parameters C and γ of SVM were set as 100 and 0.167, respectively. We can see that some ships were classified as no-water class in the water area. The mathematic morphology processing was then implemented to the classification result, where the widow size of opening and closing operator were 13*13 and 7*7. Therefore, more complete seawater area was detected as shown in Figure 5(c), and the final seawater area was extracted by only preserving the largest area water as shown in Figure 5(d). Ship proposals can be detected by overlapping the initial classification result and seawater area. For example, no-water areas in the initial classification result are identified as ship proposals that locates on the extracted seawater area, as shown in Figure 5(e).

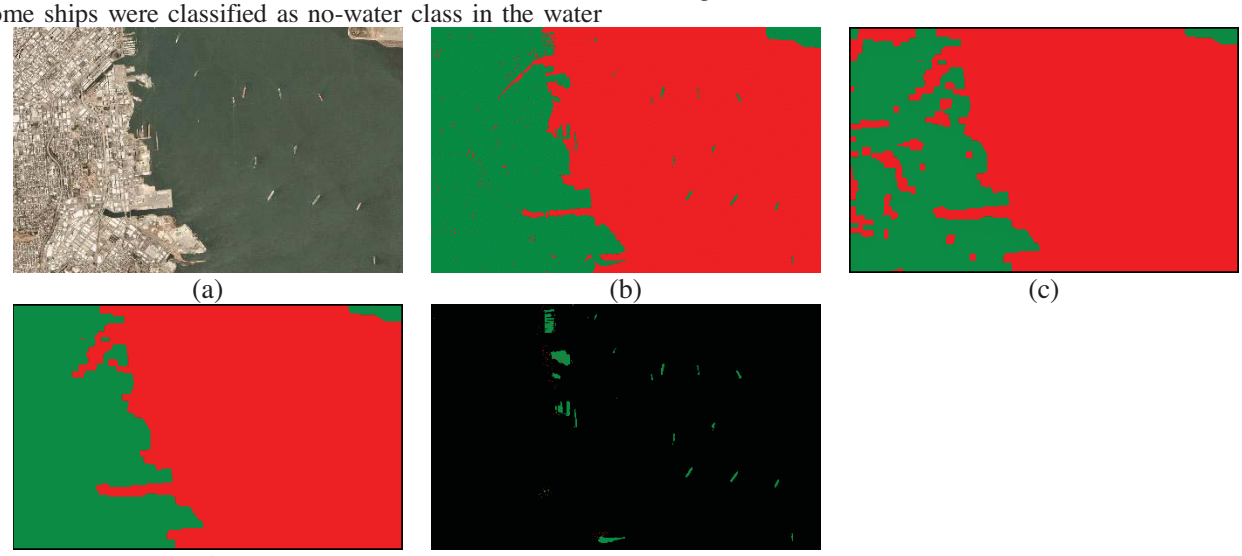

Fig. 5. Remotely sensed image (a) and initial results. (b) is the initial classification of water (red) and no-water (green), (c) is the processed result by mathematic morphology, (d) is the extracted seawater area (red), and (e) is ship proposals (green).

The second part of ship extraction with CNN was implemented using the open source CNN, which is depended on the TensorFlow and Keras. We trained and evaluated the CNN model with 3200 and 800 samples, respectively. Both precision and recall are often used to quantitatively evaluate extraction performance. The recall and precision can be calculated as follows:

$$\begin{aligned} Recall &= N_r / N_t , \\ Precision &= N_r / N_g \end{aligned} \tag{4}$$

where N_g represents the number of ground truth, N_t is the number of ships validated by CNN, and N_r represents the number of correctly detected ships. The extraction results from the trained CNN model are listed in Table 1. We can see that the recall and precision of no-ship and ships are both above 90%. Additionally, a high validation accuracy of 97% was achieved for both ships and no-ships.

The post CNN-based ship extraction results were shown with red rectangles in Figure 6. The ground truth was manually labeled as shown in Figure 6 with yellow rectangles. It can be seen that ships on the remotely sensed image were almost

exactly extracted. All ships on the seawater were extracted accurately, but some ships near the coast or parked at the dock were missed. Additionally, the land that extends into the sea is identified as a ship possibly. Therefore, more samples including ships near the coast or parked at the dock and lands or docks surrounded by sea should be added to improve the ship extraction accuracy.

TABLE I. PRECISION AND RECALL OF OUR PROPOSED SHIP DETECTOR ON DIFFERENT DATA SETS

	Recall (%)	*Precision (%)*	*Validation accuracy (%)*
No-ship	98%	98%	97%
Ship	95%	93%	

Fig. 6. Ship extraction result from high resolution remotely sensed image

IV. CONCLUSIONS

In this paper, we proposed a novel post CNN method for ship extraction from high resolution optical remotely sensed images. It includes ship proposal detection and ship extraction. In the ship proposal detection stage, SVM was first implemented to classify the images into water and no-water classes. Mathematic morphology algorithms including opening and closing operators were then implemented to the classification map to extract seawater area. Finally, no-water areas in the initial classification map were detected as ship proposals that locate on the seawater. In the second stage, a CNN model was trained based on a large dataset including ship and no-ship samples. The trained CNN was implemented to each ship proposal to extract ships. Experimental results on a high resolution optical image show the efficiency and robustness of the proposed post CNN method, which provides a satisfying ship extraction method.

REFERENCES

[1] K. Eldhuset, "An automatic ship and ship wake detection system for spaceborne SAR images in coastal regions," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 34, no. 4, pp. 1010-1019, 1996.

[2] X. Xing, K. Ji, L. Kang, and M. Zhan, "Review of ship surveillance technologies based on high-resolution wide-swath synthetic aperture radar imaging," *Journal of Radars,* vol. 4, no. 1, pp. 107-21, 2015.

[3] C. C. Wackerman, K. S. Friedman, W. G. Pichel, P. Clemente-Colón, and X. Li, "Automatic detection of ships in RADARSAT-1 SAR imagery," *Canadian Journal of Remote Sensing,* vol. 27, no. 5, pp. 568-577, 2001.

[4] C. R. Zhu, H. Zhou, R. S. Wang, and J. Guo, "A Novel Hierarchical Method of Ship Detection from Spaceborne Optical Image Based on Shape and Texture Features," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 48, no. 9, pp. 3446-3456, Sep 2010.

[5] J. X. Tang, C. W. Deng, G. B. Huang, and B. J. Zhao, "Compressed-Domain Ship Detection on Spaceborne Optical Image Using Deep Neural Network and Extreme Learning Machine," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 53, no. 3, pp. 1174-1185, Mar 2015.

[6] R. Zhang, J. Yao, K. Zhang, C. Feng, and J. Zhang, "S-CNN-based ship detection from high-resolution remote sensing images," presented at the International Archives of the Photogrammetry, Remote Sensing Spatial Information Sciences Prague, Czech Republic, July, 2016.

[7] G. Liu, Y. Zhang, X. Zheng, X. Sun, K. Fu, and H. Wang, "A new method on inshore ship detection in high-resolution satellite images using shape and context information," *IEEE Geoscience Remote Sensing Letters,* vol. 11, no. 3, pp. 617-621, 2014.

[8] J. Xu, X. Sun, D. Zhang, and K. Fu, "Automatic detection of inshore ships in high-resolution remote sensing images using robust invariant generalized Hough transform," *IEEE Geoscience Remote Sensing Letters,* vol. 11, no. 12, pp. 2070-2074, 2014.

[9] L. Huo, L. Jiao, S. Wang, and S. Yang, "Object-level saliency detection with color attributes," *Pattern Recognition,* vol. 49, pp. 162-173, 2016.

[10] C. Dong, J. Liu, and F. Xu, "Ship Detection in Optical Remote Sensing Images Based on Saliency and a Rotation-Invariant Descriptor," *Remote Sensing,* vol. 10, no. 3, Mar 2018, Art. no. 400.

[11] Q. P. Li, L. C. Mou, Q. J. Liu, Y. H. Wang, and X. X. Zhu, "HSF-Net: Multiscale Deep Feature Embedding for Ship Detection in Optical Remote Sensing Imagery," *IEEE Transactions on Geoscience and Remote Sensing,* Article vol. 56, no. 12, pp. 7147-7161, Dec 2018.

[12] Y. Yao, Z. G. Jiang, H. P. Zhang, D. P. Zhao, and B. W. Cai, "Ship detection in optical remote sensing images based on deep convolutional neural networks," *Journal of Applied Remote Sensing,* vol. 11, p. 12, Sep 2017, Art. no. 042611.

[13] F. Yang, Q. Z. Xu, B. Li, and Y. Ji, "Ship Detection From Thermal Remote Sensing Imagery Through Region-Based Deep Forest," *IEEE Geoscience and Remote Sensing Letters,* vol. 15, no. 3, pp. 449-453, Mar 2018.

[14] Z. Zou and Z. Shi, "Ship detection in spaceborne optical image with SVD networks," *IEEE Transactions on Geoscience and Remote Sensing* vol. 54, no. 10, pp. 5832-5845, 2016.

[15] R. Wang, L. Jie, Y. Duan, H. Cao, and Y. Zhao, "Study on the Combined Application of CFAR and Deep Learning in Ship Detection," *Journal of the Indian Society of Remote Sensing,* vol. 46, no. 9, pp. 1413-1421, 2018.

[16] H. Zhang, W. Shi, and K. Liu, "Fuzzy-Topology-Integrated Support Vector Machine for Remotely Sensed Image Classification," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 50, no. 3, pp. 850-862, Mar. 2012.

[17] C. Cortes and V. Vapnik, "Support-Vector Networks," *Machine Learning,* vol. 20, no. 3, pp. 273-297, 1995.

[18] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," presented at the Proc. 5th Annu. Workshop Comput. Learn Theory, Pittsburgh, PA, 1992.

[19] M. Volpi, D. Tuia, F. Bovolo, M. Kanevski, and L. Bruzzone, "Supervised change detection in VHR images using contextual information and support vector machines," *International Journal of Applied Earth Observation and Geoinformation,* vol. 20, no. S1, pp. 77-85, Feb. 2013.

[20] G. Sapiro, R. Kimmel, D. Shaked, B. B. Kimia, and A. M. Bruckstein, "Implementing continuous-scale morphology via curve evolution," *Pattern Recognition,* vol. 26, no. 9, pp. 1363-1372, 1993.

[21] https://www.kaggle.com/rhammell/ships-in-satellite-imagery.