

Ship Detection on Space borne Image Using Machine Learning Techniques

Prathima V R¹, Manasa C H²

¹Associate Professor, Computer Science and Engineering, RGIT College, Bangalore, India

²M. Tech, Computer Science and Engineering, RGIT College, Bangalore, India

Abstract: As technology has increased and so has led the attacks in water boundaries under national and international currents. 82% of world trade is dependent on water ways and thus monitoring and maintain of continues water vehicle is a challenge. In this paper have proposed a novel technique for ship monitoring under compressed domain for continues monitoring. The paper also focuses on mis-masking the actual location of ship with a difference of random allocation.

Keywords: Compressed domain, deep neural network (DNN), extreme learning machine (ELM), JPEG2000, optical space borne image, remote sensing, ship detection

1. Introduction

Ship detection and monitoring has become most important and focused agenda under water and navy management. International and national waters have increased cargo masking and mis-aligning of goods. Apart from monitoring, the proposed work has future scope towards trace monitoring of oil drops and fish illegal activities. Vessel monitoring from satellite images provides a wide visual field and covers large sea area and thus achieves a continuous monitoring of vessels' locations and movements. With known that optical space borne images have higher resolution and more visualized contents than other remote sensing images, which is more suitable for ship detection or recognition in the aforementioned applications.

However, optical space borne images usually suffer from two main issues: 1) weather conditions like clouds, mists, and ocean waves result in more pseudo targets for ship detection, and 2) optical spaceborne images with higher resolution naturally lead to larger data quantity than other remote sensing images, and thus, optical spaceborne images are more difficult to be tackled for real-time applications.

The paper is organized with an introduction followed by a brief literature reviews on focusing towards previous version of work and architectures, section III consist of proposed methodology and section IV consist of Implementation and results.

2. Proposed Method

The typical JPEG 2000 compression is shown in Fig. 1. To clearly illustrate the proposed approach, it is necessary to define compressed domain in advance. According to the working the *compressed domain* is anywhere in the compression or decompression procedure, after transform or before inverse transform. Therefore, object detection can be conducted in compressed domain from points 1 to 6 in Fig. 1. Unlike the other points, entropy coding (points 3 and 4 in Fig. 1) will obviously change the spatial distribution of the object features and destroy the structure information. Hence, points 1, 2 and 5, 6 are more suitable for ship detection. Furthermore, as points 5, 6 are symmetry to points 1, 2 in

code implementation, only points 1, 2 are discussed here in after. At the encoder side, DWT is first performed (point 1 in Fig. 1). Then, the resulting coefficients are mapped to different bit planes by quantization (point 2 in Fig. 1). The bit-plane encoding will not obviously change the properties of wavelet features and thus, the detection accuracy will not be severely affected. Based on this analysis, point 1 is viewed as the ideal place for ship detection.

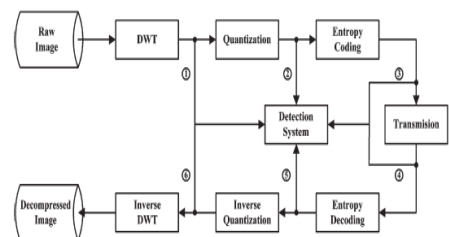


Fig. 1. JPEG2000 image compression.

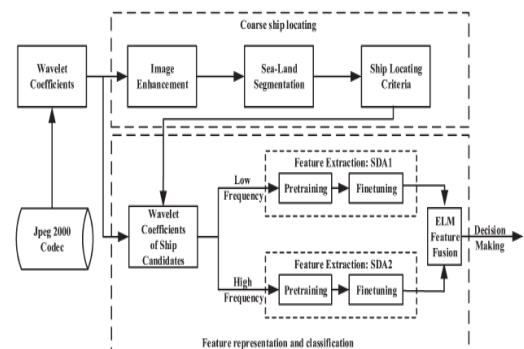


Fig. 2. Proposed ship detection framework.

The block diagram of the proposed framework is depicted in Fig. 2. It can be decomposed into two main steps: image preprocessing (for coarse ship locating), and feature representation and classification (for ship object detecting). In the preprocessing, CDF 9/7 wavelet coefficients are extracted from JPEG 2000 codec. The wavelet coefficients in different sub bands tend to reflect different properties of the original image. Generally speaking, the low frequency contains most of the global information, while the high frequency represents local or detail information. In the proposed model, the low-frequency sub band LL is exploited

for the extraction of the regions of ship candidates. On the other hand, the low-frequency coefficients and high frequency coefficients (HFCs) are individually processed for feature extraction by two DNNs, which are to be discussed in Section IV. Moreover, to fully exploit the information of the original image in wavelet domain, the resulting features from low and high sub bands are further fused by ELM, for more accurate feature classification (i.e., higher ship detection accuracy). The detailed implementations are listed as follows. 1) Wavelet singularities of LL are detected to train a stacked denoising auto encoder (SDA) 1. Note that SDA is one of the implementation strategies of DNN and will be introduced in Section IV-A. 2) The combination of the wavelet coefficients in high frequency sub bands (i.e., LH, HL, and HH) are used to train an SDA2. 3) The weight matrices of the trained SDAs are considered as feature extractors for low- and high-frequency sub bands, respectively. The obtained features are then combined to train an online sequential ELM (OS-ELM). It should be mentioned that the third step can be regarded as decision pooling of SDA1 and SDA2, or training a high performance classifier of ship features. Since we are to make our algorithm more robust to various environmental conditions, online training is adopted to further improve the network's performance. The experiments in [10] and [11] showed that ELM is fast and more accurate in large class training and the generalization performance of ELM turns out to be very stable.

3. Coarse Ship Locating

As shown in Fig. 2, fast ship locating (i.e., ship candidate extraction) is performed in LL sub band, which includes image enhancement, sea-land segmentation, and ship locating based on shape criteria.

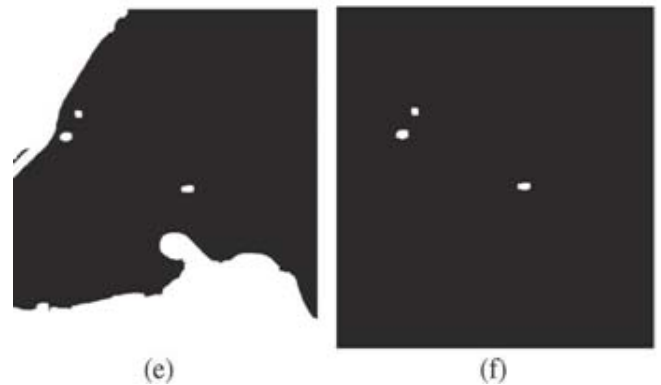
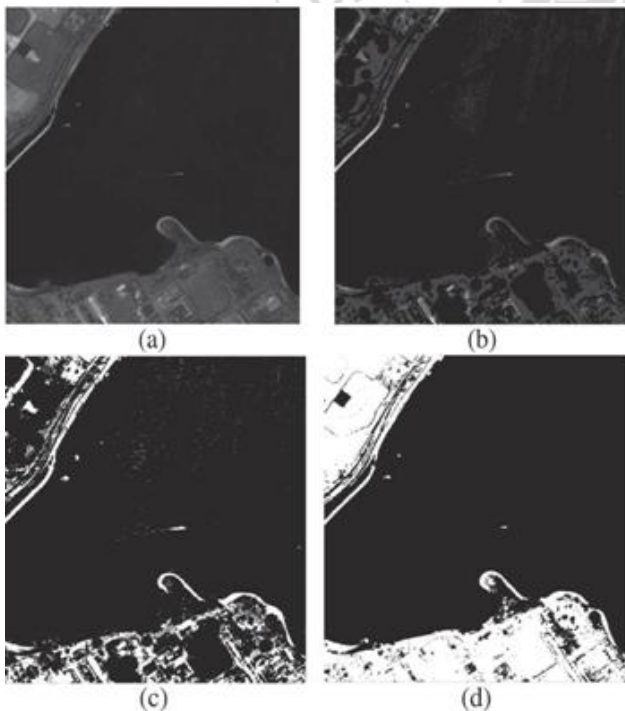


Figure 3: Coarse ship segmentation: (a) Input image (b) top-hat-transformed image, (c) binarized with the first threshold T , (d) corrected with the second threshold T_- , (e) refined by morphology dilation and erosion, and (f) coarse ship location.

A. Image Enhancement In order to remove uneven illumination, a morphological operator, i.e., top-hat transform (THT), is used for ship extraction and background suppression. As ships are usually brighter than their surroundings, the white THT is employed in the proposed work [shown in Fig. 3(b)]. The mathematical definition of white THT is as follows: $T_w(f) = f - f \circ b$ (1) where f is the input LL coefficients of the original image, \circ denotes opening operation, and T_w is the enhanced image. In the simulations, b is set as a circular structuring element with a radius of 12.

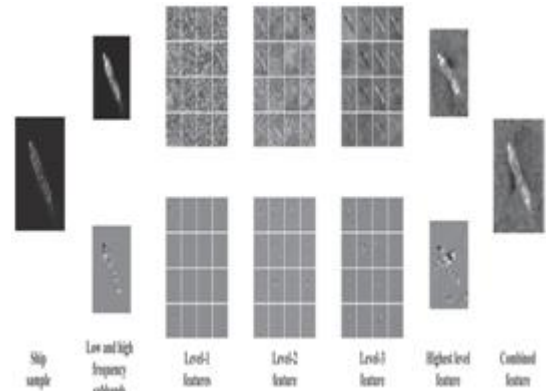
B. Sea-Land Segmentation Different from the traditional intensity histogram and maximum variance segmentation, here, a statistical Gaussian model is adopted to adaptively estimate the probabilistic distribution of the sea regions [36], and the algorithm is as follows. 1) Binarize the input image by the Otsu algorithm [37], and then label the connected regions. 2) Find the geometrical center P of the largest connected region R . 3) Use point P as the starting point; traverse R to obtain another set of points P_- satisfying that the $A \times A$ (empirically set as 60 in the experiments) neighboring regions of P_- are inside the region R . Label the points P_- as all-sea region S . 4) Compute the mean μ and variance σ of S , and use them as the statistical parameters of the Gaussian model. The resulting μ and σ are used to compute a threshold (T) for image binarization, as follows: $T = \mu + \lambda\sigma$ (2) where λ is the weight of variation (σ) and set as three according to the Gaussian distribution. The binarized image obtained by T [shown in Fig. 3(c)] usually remains *holes* in large lands or clouds. In this case, a new threshold (T_-) for the elimination of *hole regions* and incorrectly marked lands is chosen as $T_- = \lambda_- \sigma$ (3) where λ_- is a parameter to control the similarity of land and sea, empirically set as four. After thresholding with T_- [shown in Fig. 3(d)], the median filtering (with size of [3 3]), morphology dilation, and erosion (circular structuring element with a radius of three) are applied to fill the isolated holes. Then, the masks of land, cloud, and ship candidates are segmented [shown in Fig. 3(e)]. In the following, ship candidates will be further extracted by using the unique shape properties of ships. Note that some of the pseudo targets may be included in the extracted regions; however, they can be removed in the process of feature fusion and classification.

C. Ship Locating Criteria In the previous section, several connected regions are extracted from the resultant masks by labeling the eight-connected neighbors. Geometric properties of the connected regions are then used for the locating of ship candidates, which are listed as follows

- 1) **Area:** It equals the number of pixels in the corresponding connected region. Area is used to cut off the lands, clouds, and other obviously large/small false targets.
- 2) **Major minor axis ratio:** It is defined as $Rls = LaxisL / LaxisS$ (4) where $LaxisL$ and $LaxisS$ are the length of long and short axes of the bounding rectangle, respectively.
- 3) **Compactness:** Compactness measures the degree of circular similarity, and it is defined as $Compactness = Perimeter^2 / Area$. (5) By using these shape criteria, we can obtain the coarse locations of ship candidates [shown in Fig. 3(f)]. In the experiments, the size of testing images is 2000×2000 (in pixels) with a resolution of 5 m. The size of ship candidates is supposed to be smaller than 100×100 (or larger than 10×10). In this case, the regions with area larger than 10 000 (or smaller than 100) would be removed. Moreover, as the long axis of ship should be longer than the minor one, the major minor axis ratio is selected as 1.5. Compactness is set as 40 to exclude the regions which are obviously irregular. It is also worth to note that, compared with original images, using a low-frequency sub band (LL) for coarse ship locating would decrease the detection accuracy by 0.32% (in statistical average), but the detection speed is improved by more than 60%.

4. Ship Feature Representation and Classification

The state-of-the-art ship detection approaches extract complicated features and combine them with learning-based classification. These feature operators or descriptors, e.g., scaleinvariant feature transform [10], speeded up robust features [11], histogram of gradient [12], local multiple pattern (LMP) [13], and shape/texture features [14], are engineered to be invariant under certain rotations or scale variations and chosen for some specific vision tasks. Features extracted by these methods generally have some fundamental limitations in practical applications. For example, they may have poor performances when the images are corrupted by blur, distortion, or illumination, which commonly exist in the remote sensing images. Relatively, learning features from image would help to tackle these issues. Recent works in [15] and [16] have shown that the features extracted by the unsupervised learning outperform those manually designed ones on object detection or recognition. However, ship detection is usually under complicated environmental conditions, and the processed images may contain various pseudotargets, e.g., islands, clouds, coastlines, etc. Bengio *et al.* [17] indicated that traditional machine learning algorithms, e.g., SVM, may have difficulties in efficiently handling such highly varying inputs. These learning schemes usually use a few layers of computational units to establish the training model. When dealing with highly variant conditions, the computation is exponentially increased.



As mentioned in Section II and shown in Fig. 2, the low frequency (LL) and high-frequency (LL, LH, and HH) sub bands are trained by two SDAs, respectively. Singularities represent the sparse structures of LL, and therefore, they are extracted to train the SDA1. As the LH, HL, and HH already reflect the sparseness of the image, they are combined and used as the inputs of SDA2. Before training, the input data need to be initialized by a zero-mean or z-score normalization $Z = (M - \text{mean}(M)) / \text{std}(M)$.

C. Pretraining and Fine Tuning In this section, we introduce the details of SDA training for ship feature extraction in low and high frequencies. Generally speaking, SDA-based feature extractor involves two main steps: pretraining and fine tuning. The unsupervised layer-by-layer pretraining can help to achieve good generalization and low variance of testing error. Each layer is trained as a denoising autoencoder by minimizing the reconstruction of its input (which is the output code of the previous layer). Based on the recent works in [18], some additional parameters are set to further improve the performance of the SDA. Before training, the coefficients are scaled to $[0, 1]$, and the learning rate is set as 0.1. The number of training batches depends on the size of data set, usually between [10, 100]. Different training batches should contain different classes of training samples to achieve better performance. Compared with 5% noise that is typically used in SDA the simulations in [18] indicated that it is better drop out 20% inputs combined with 50% hidden units. Once all of the layers are pretrained, the network needs a second stage of supervised training called fine tuning. The supervised fine tuning is used to minimize the prediction error. Practically, a logistic regression layer is added on top of the pretrained network.

5. Experiments and Analysis

Extensive experiments are conducted in this section. Since SDA-based feature extraction, ELM-based feature fusion, and classification are adopted in this work, we term the proposed method as SDA-ELM, which is compared with the relevant state-of-the-art methods in [19] and [20]. In [19], multiple features are fused by SVM (denoted as MF-SVM), while in [20], salient regions are detected before SVM-based classification (denoted as SA-SVM). In addition, another method (SDA-based feature combined with SVM-based classification) is also tested (denoted as SDA-SVM). In the following sections, to verify the effectiveness of each component of the proposed method (i.e., ship locating, feature extraction, feature fusion, and classification), the performance of ship candidate segmentation is first tested; then, the proposed

SDA-based feature extraction is compared with other feature representation methods; classification performance of ELM is further evaluated against SVM, by using different combinations of extracted ship features; and finally, the overall ship detection accuracy is compared to demonstrate the advantages of the proposed scheme under practical testing conditions.

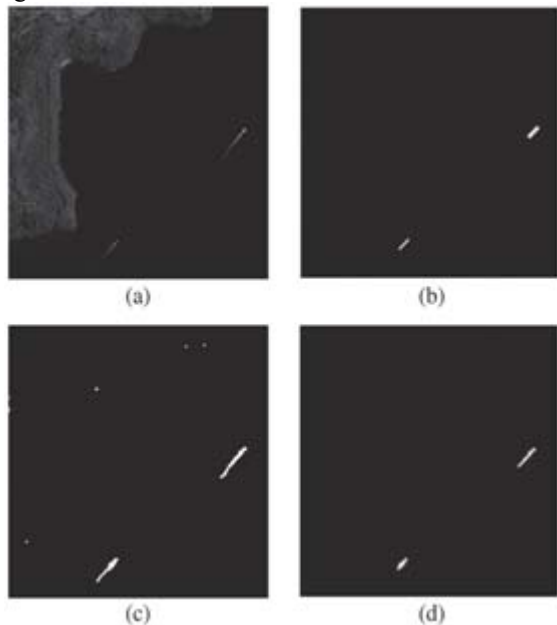


Figure 5: Performance comparison of ship candidate segmentation. (a) Original image. (b) Manually labeled ground truth (the white pixels indicate ship candidates, while the black pixels represent land/sea regions). (c) Results by Chan– Vese model in. (d) Results by the proposed method.

prepared to build an image data set to compare the performances of MF-SVM , SA-SVM , SDA-SVM, and SDAELM, and 1600 training samples are extracted for feature learning, shown in Table I. It should be emphasized that the images for the extraction of 1600 training samples have not been included in the testing images. The testing hardware and software conditions are listed as follows: Intel-i7 2.4 G CPU, 8 G DDR3 RAM, Windows 7, Matlab R2012b, and Microsoft Visual Studio 2010.

A. Comparison of Coarse Ship Locating The coarse ship locating is performed in the low-frequency sub band LL. Fig. 5 shows the comparison of segmentation results of ship candidates, and one can see that the proposed method achieves more accurate segmentation than the Chan– Vese model in. In addition, we also conducted objective comparisons to evaluate the performances of different methods. Three commonly used criteria were computed: false positive rate (FPR), false negative rate (FNR), and false error (FE). They are defined as follows: $FPR(SR, GT) = \frac{\#(SR \cap \overline{GT})}{\#(SR \cap \overline{GT}) + \#(GT)}$ * 100% $FNR(SR, GT) = \frac{\#(\overline{SR} \cap GT)}{\#(\overline{SR} \cap GT) + \#(GT)}$ * 100% $FE(SR, GT) = (FPR + FNR) * 100\%$

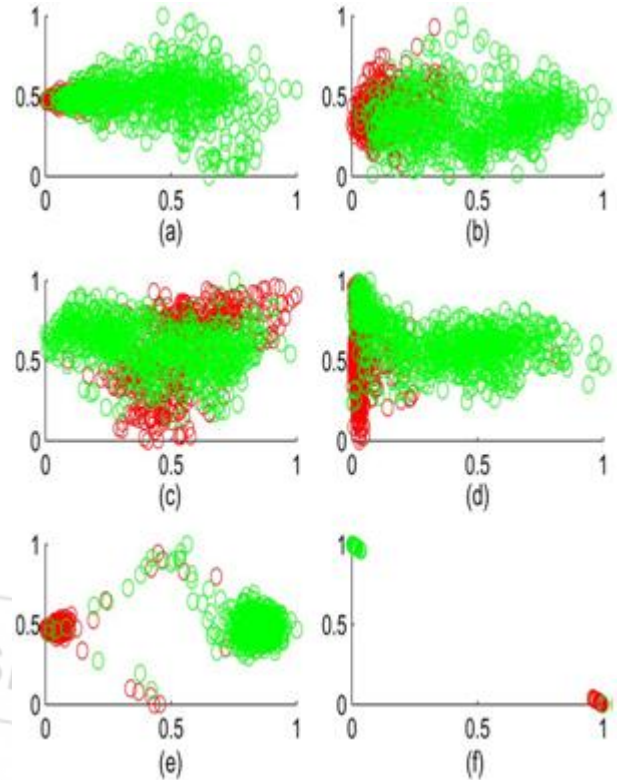


Figure 6: Visualized 2-D space distributions of the first two principal components of different features. (a) Singularities of LL. (b) HFCs. (c) Shape/texture features of SA-SVM . (d) LMPs of MF-SVM . (e) Proposed SDA-based features. (f) Two-dimensional outputs of ELM pooling. Fig. 6. Visualized 2-D space distributions of the first two principal components of different features. (a) Singularities of LL. (b) HFCs. (c) Shape/texture features of SA-SVM . (d) LMPs of MF-SVM . (e) Proposed SDA-based features. (f) Two-dimensional outputs of ELM pooling. where SR denotes the segmentation results (ship candidates), GT denotes the manually labeled ground truth, $\#(\cdot)$ is the number of pixels in the corresponding region, and \overline{GT} and \overline{ER} denote the regions which are not included in GT and ER, respectively. The averaging comparison results of the proposed method and Chan–Vese model in are demonstrated in Table II. It is shown that the proposed model has lower FPR, FNR, and FE (better performance).

B. Comparison of Feature Representations In this experiment, representation performances of different features are compared, including the LL singularities (LLSs), the HFCs, the LMPs in, the shape/texture features in , and the proposed SDA-based features. Principal component analysis is used for visualizing different features in 2-D space. Fig. 6 shows the first two principal components of each feature, where the red points represent ships and the green ones represent other subclasses shown in Table I. As can be seen, the distributions of LLS and HFC are completely blended together. Relatively, the distances of the feature points in expand little in the Cartesian coordinates, and still, a large amount of feature points are overlapped. LMP outperforms the aforementioned features; nearly half of the red points are separated from the green ones. The proposed

C. Comparison of Feature Fusion and Classification In this section, the classification performance of ELM is compared against that of SVM, by using different combinations of extracted features. The classification accuracy of each method is computed as $T = \text{Number of correctly classified}$

samples Number of tested samples * 100%. The experiments are based on k -fold cross-validation, which provides a better Monte Carlo estimate than simply randomly divided data set. First, the data set is randomly split into k mutually exclusive subsets S_1, S_2, \dots, S_k of approximately equal size. Then, the classifier is trained and

tested k times; for each testing $t \in \{1, 2, \dots, k\}$, it is trained on S without S_t and tested on S_t . As our training data set has 1600 samples, we select $k = 4$. The classification accuracy of each feature is shown in Fig. 7.

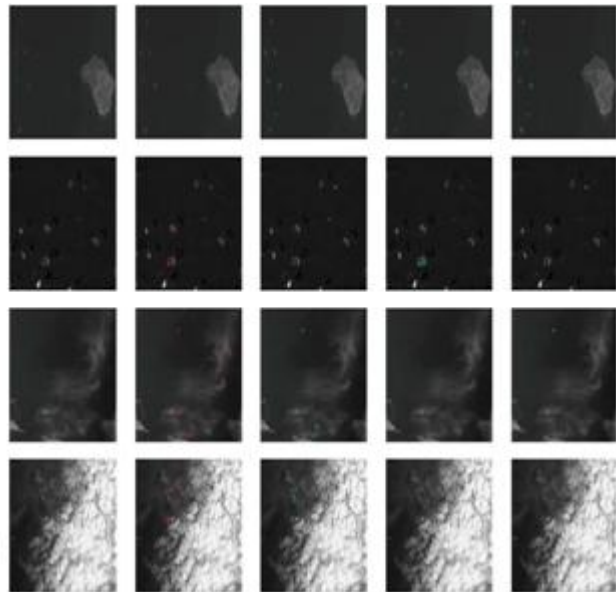
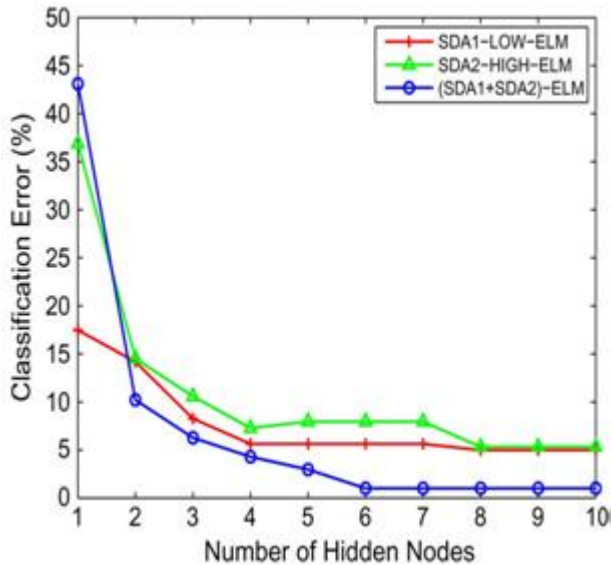


Figure 9: Detection results of different methods under different experimental conditions. (Rows 1–4) Land, little clouds, cotton shaped cloud with mist, and large area of floccus. (Column 1) Input images. (Column 2) Coarse location of ship candidates. (Column 3) Classification results of MF-SVM [1]. (Column 4) Classification results of SA-SVM [11]. (Column 5) Classification results of our method

6. Conclusion

In this paper, we have proposed a compressed-domain ship detection framework using DNN and ELM for optical spaceborne images. Compared with the previous works, the proposed approach achieves better classification by deeplearning- based feature representation model with faster detection in compressed domain. After ship candidates are extracted, the singularities in LL are detected to train the SDA1. Then, the combination of high-frequency components (i.e., LH, HL, and HH) is used to train the SDA2. The two SDAs are viewed as feature extractors to obtain high-level features, and the resultant features are fused by ELM to further improve the classification results. ELM learns extremely faster and has better generalization than other traditional learning algorithms. Extensive experiments demonstrate that our proposed scheme outperforms the state-of-the-art methods in terms of detection time and accuracy. As for the possible shortcomings of the proposed work, the parameters in coarse ship locating should be more adaptive to the image contents. In addition, due to the availability of image data sets, the simulations in the proposed work are conducted using panchromatic images, and other remote sensing image could be further tested or verified in a future work.

References

- [1] C. Zhu, H. Zhou, R. Wang, and J. Guo, "A novel hierarchical method of ship detection from spaceborne optical image based on shape and texture features," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 9, pp. 3446–3456, Sep. 2010.
- [2] C. Corbane, L. Najman, E. Pecoul, L. Demagistri, and M. Petit, "A complete processing chain for ship detection using optical satellite imagery," *Int. J. Remote Sens.*, vol. 31, no. 22, pp. 5837–5854, Jul. 2010.
- [3] F. Y. Lure and Y.-C. Rau, "Detection of ship tracks in AVHRR cloud imagery with neural networks," in *Proc. IEEE IGARSS*, 1994, vol. 3, pp. 1401–1403.
- [4] J.M.Weiss, R. Luo, and R.M.Welch, "Automatic detection of ship tracks in satellite imagery," in *Proc. IEEE IGARSS*, 1997, vol. 1, pp. 160–162.
- [5] A. N. de Jong, "Ship infrared detection or vulnerability," in *Proc. SPIE*, 1993, pp. 216–224, International Society for Optics and Photonics.
- [6] J.-W. Lu, Y.-J. He, H.-Y. Li, and F.-L. Lu, "Detecting small target of ship at sea by infrared image," in *Proc. IEEE Int. CASE*, 2006, pp. 165–169.
- [7] C. DeSilva, G. Lee, and R. Johnson, "All-aspect ship recognition in infrared images," in *Proc. Electron. Technol. Directions Year 2000*, 1995, pp. 194–198.
- [8] D. W. Burgess, "Automatic ship detection in satellite multispectral imagery," *Photogramm. Eng. Remote Sens.*, vol. 59, no. 2, pp. 229–237, 1993.
- [9] G. Wu, J. de Leeuw, A. K. Skidmore, Y. Liu, and H. H. Prins, "Performance of Landsat TM in ship detection in turbid waters," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 11, no. 1, pp. 54–61, Feb. 2009.
- [10] C. Corbane, F. Marre, and M. Petit, "Using SPOT-5 HRG data in panchromatic mode for operational detection of small ships in tropical area," *Sensors*, vol. 8, no. 5, pp. 2959–2973, 2008.

- [11] F. Bi, F. Liu, and L. Gao, "A hierarchical salient-region based algorithm for ship detection in remote sensing images," in *Advances in Neural Network Research and Applications*. New York, NY, USA: Springer-Verlag, 2010, pp. 729–738.
- [12] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, Jul. 2006. [13] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006.

