

Article

Ship Detection in Multispectral Satellite Images Under Complex Environment

Xiaoyang Xie , Bo Li and Xingxing Wei *

Beijing Key Laboratory of Digital Media, School of Computer Science and Engineering, Beihang University, Beijing 100191, China; elizabeth.xie@buaa.edu.cn (X.X.); Boli@buaa.edu.cn (B.L.)

* Correspondence: xxwei@buaa.edu.cn

Received: 27 January 2020; Accepted: 22 February 2020; Published: 2 March 2020



Abstract: Ship detection in multispectral remote-sensing images is critical in marine surveillance applications. The previously proposed ship-detection methods for multispectral satellite imagery usually work well under ideal conditions. When meeting complex environments such as shadows, mists, or clouds, they fail to detect ships. To solve this problem, we propose a novel spectral-reflectance-based ship-detection method. Research has shown that different materials have unique reflectance curves in the same spectral wavelength range. Based on this observation, we present a new feature using the reflectance gradient across multispectral bands. Moreover, we propose a neural network called lightweight fusion networks (LFNet). This network combines the aforementioned reflectance and the color information of multispectral images to jointly verify the regions with ships. The method utilizes a coarse-to-fine detection framework because of the large-sense- sparse-targets situation in remote-sensing images. In the coarse stage, the proposed reflectance feature vector is used to input the classifier to rule out the regions without ships. In fine detection, the LFNet is used to verify true ships. Compared with some traditional methods that merely depend on appearance features in images, the proposed method takes advantage of employing the reflectance variance in objects between each band as additional information. Extensive experiments have been conducted on multispectral images from four satellites under different weather and environmental conditions to demonstrate the effectiveness and efficiency of the proposed method. The results show that our method can still achieve good performance even under harsh weather conditions.

Keywords: ship detection; spectral feature; convolutional neural networks; remote sensing

1. Introduction

Ship detection in remote-sensing imagery is of significant importance in maritime security and transportation surveillance applications, such as vessel salvage and fisheries management [1–3]. As revisit periods have decreased, along with the improvements in image resolution for optical satellites, continuous monitoring over a vast area via images has become a reality through sensors embedded within these satellites. According to the statistics from Reference [4], ship detection in remote-sensing images received increasing attention in the period from 1978 and 2016, and this trend continues. Among them, the adoption of multispectral remote-sensing images allows for a more detailed analysis of the reflectance variance in different objects in different wavelength intervals [5,6], thus offering a significant advantage by presenting discriminate reflectance when the target is occluded by shadows, mist, or clouds. Moreover, it provides an aerial view mode that makes it more convenient to intuitively observe fleets over vast ranges.

Previous research has shown the productive results in ship detection under the limited resolution of spaceborne multispectral images, as presented in References [5,7,8]. Specifically, in Reference [5], a ratio image of two bands in the multispectral image was used to make a land mask, and a Bayes

classifier was then applied to the three bands of the multispectral remote-sensing image to remove the islands from the potential ships. In Reference [7], six bands in multispectral images were treated as multidimensional data and were processed as biquaternions. In References [1,9], a ship-detection method that can be applied to both panchromatic imagery and multispectral imagery is proposed. These two methods perform ship detection in only one band. However, because optical remote-sensing images are susceptible to bad weather and have a negative impact on detection tasks [4], these algorithms can barely resist false alarms under complex weather conditions.

The powerful feature representation of convolutional neural networks (CNNs) on computer vision detection tasks has been demonstrated in Reference [10]. Compared with traditional features designed by human-subjective accumulated knowledge, CNN features rely on a large amount of positive and negative sample data to perform detection tasks, and its deep structured neural networks can automatically generate feature representations from original image pixels. Since then, this technique was utilized to pursue higher accuracy by making deeper and more complicated networks (e.g., References [11–14]). If these advances to the improvement of accuracy are used directly in remote-sensing images, they might not necessarily make networks more efficient with respect to the detection tasks of large scenes and sparse targets but requires high computational and storage expenses. For utilizing CNNs in missions with restricted space and time, researchers have suggested using lightweight CNNs [15,16]. The rise of CNNs technology has also led to its application research in high-resolution optical satellite image target detection (e.g., References [17–21]). Due to the powerful feature representation of CNNs, Reference [17] combined a singular-value decomposition algorithm with a three-layered CNN operated on 2-m or even higher resolution panchromatic remote-sensing images. The method in Reference [18] has modified AlexNet CNN [10], the winner of the ImageNet [22] competition in year 2012, to obtain rotation-invariant appearance features in remote-sensing images with a spatial resolution ranging from 0.5 m to 2 m; the method in Reference [20] built a new network architecture based on Reference [23] to adapt to capturing the clustered small objects in remote-sensing images. Both methods were applied in RGB color space images. The methods in References [3,19–21] combined the CNN technique with the coarse-to-fine strategies and believed that this would perform better in remote-sensing image-detection tasks. They all took full advantage of the spatial feature representation of remote-sensing images and benefitted from a large data set or the large size (increased depth) of the network to improve their performance under different situations (e.g., References [11–13]) but ignored the essential difference between remote-sensing images and natural scene images. Among these works, Reference [9] is the only work to conduct experiments under different weather conditions. To address this problem, this paper investigates and exploits the spectral properties in multispectral images to compensate for the insufficient appearance features on targets in multispectral remote-sensing images in different environment conditions.

Unlike the RGB channel in the natural colorful images, each spectral band in a multispectral remote-sensing image records the ratio of the reflected over the incoming radiation in a specific spectral wavelength range falling on each pixel, which can be used for surface material identification. The studies in References [5,6] showed that the reflectance varieties among bands in multispectral satellite images would benefit ship detection. Figure 1 shows a grayscale image of the blue, green, red, and near-infrared bands containing clouds, ships, and islands. In Figure 1, the vegetation on the island clearly has a higher intensity in the fourth spectral band than in the other three bands. However, the cloud in the blue band seems larger though they are the same cloud in these four images. The ships seem to have a higher contrast in the fourth band. Different surface materials exhibit varying degrees of reflection intensity in each spectral band.

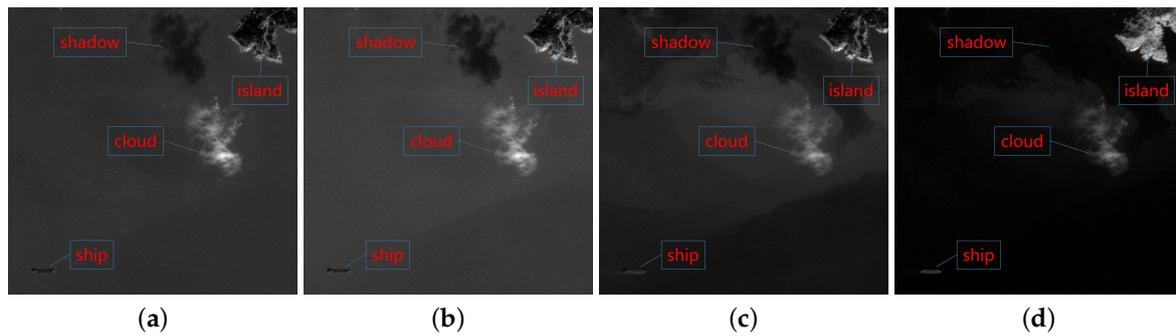


Figure 1. Four bands of a multispectral satellite image with cloud and wave backgrounds: (a) the grayscale image of the blue band; (b) the grayscale image of the green band; (c) the grayscale image of the red band; and (d) the grayscale image of the near-infrared band.

Research on spectrum imaging shows that different materials have unique spectral reflectances in remote-sensing images. Researchers have assembled digital spectral libraries recording the reflectance spectra of seawaters, minerals, mineral mixtures, and many other materials [24] for hyperspectral images. These spectral properties in hyperspectral remote-sensing images are widely applied for analyzing object composition. For example, the normalized difference vegetation index (NDVI) [25] is used to calculate vegetation coverage and the normalized difference water index (NDWI) is used to delineate water features [26]. Spectral properties are also generally used for the classification of scenes and land usage in hyperspectral satellite images [27,28]. While spectral feature extraction in hyperspectral data is often performed to reduce dimensionality for classification, it is rarely applied to multispectral data with lower spectral resolution [28]. However, as indicated in Reference [6], the reflectance of water and ships differs among bands with different wavelength spectra and could serve as efficient auxiliary information in ship detection. Although spectral bands show a pronounced advantage in performing ship detection, these characteristics have neither been comprehensively analyzed nor been efficiently utilized [4].

Therefore, in this paper, we propose a ship-detection method based on the reflectance of the multispectral satellite images. To address the large scene and sparse targets problem, we applied a coarse-to-fine detection framework. In the coarse stage, the reflectance image of the original multispectral image was obtained and a feature set was designed. We use the reflectance gradients to extract ship pixels and to obtain the locations of ship candidates. In the fine stage, the reflectance image and synthesized false color image obtained from the original image were combined to learn the weight of the proposed lightweight fusion network (LFNet). In this way, the network learned the shape, color, and reflectance features of the ships. Compared with other deep CNNs, LFNet does not require very large training samples. In addition, LFNet is performed on patches to improve the processing efficiency instead of the whole image. The entire detection process of the proposed method is shown in Figure 2. It included three stages: preprocessing stage, ship candidate extraction stage, and ship verification stage. In the preprocessing stage, the Geographic Information System (GIS) was utilized to obtain the sea-land mask. Besides, the reflectance images of the original input, four-band images were calculated and the 3-channel false color image was synthesized for further process. In the ship candidates extraction stage, spectral reflectance gradients were calculated from the reflectance images and combined with random forest classifier for classification of ship and nonship. In the ship verification stage, the obtained ship candidates from the previous stage were input into the lightweight fusion network for feature extraction and screened for false ship candidates.

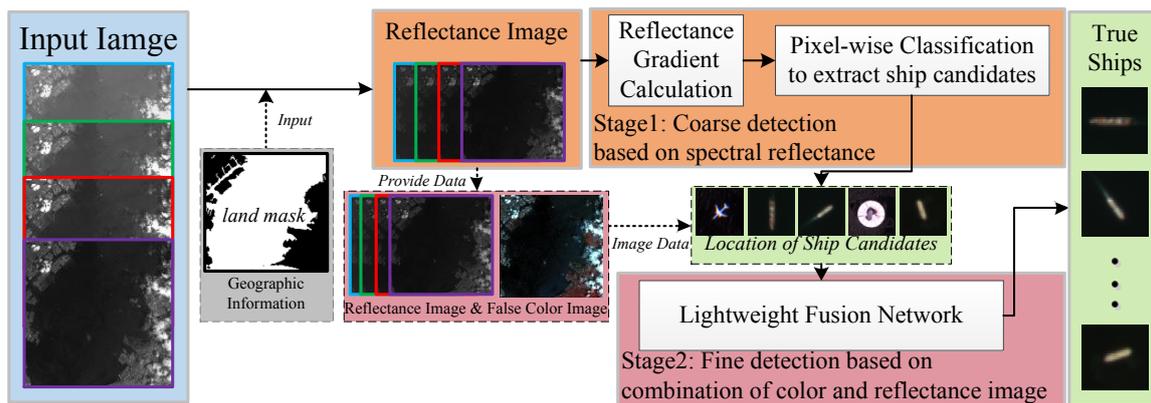


Figure 2. The flowchart of the proposed method: When considering a multispectral image, we use geographic information to obtain the ocean regions. Afterward, a coarse detection stage is performed, where the spectral reflectance image is computed based on the ocean image. This procedure is then followed by the calculation of the reflectance gradient across different bands for each pixel. The pixels are classified according to the designed features. The classification result enables us to obtain the ship candidate location. In the fine detection stage, we use the lightweight fusion network to further verify the ship candidates. After this step, the final ship-detection results are achieved.

The contributions of the proposed work are as follows:

- (1) To the best of our knowledge, this study might be the first to utilize spectral reflectance to perform ship detection tasks. Furthermore, a new type of feature to coarsely rule out the regions without ships is carefully designed;
- (2) A new CNN architecture, i.e., LFNet is designed based on References [16,17] to effectively combine the reflectance map with the color map of the multispectral image for extracting the ship reflectance features and appearance features.

The paper is structured as follows. Section 2 presents the analysis of the spectral properties of the targets and the task-related background objects in multispectral remote-sensing images. The spectral feature calculation formula and details of the structure of the LFNet used for ship detection are introduced in Section 3. Section 4 delineates the training process for the detection algorithm. The experimental results from four multispectral remote sensors are reported in Section 5, and a conclusion is drawn in Section 6.

2. Coarse Detection Based on Reflectance

In this section, the details of the coarse detection are introduced.

2.1. Spectral Characteristics Analysis

The multispectral satellite images discussed in this paper include four bands, namely the blue band (band 1, ranging from 0.45 to 0.52 microns), green band (band 2, ranging from 0.52 to 0.59 microns), red band (band 3, ranging from 0.63 to 0.69 microns), and near-infrared band (band 4, ranging from 0.77 to 0.89 microns). These images have six-meter, ten-meter, sixteen-meter, and twenty-meter spatial resolutions, respectively. In the following paragraphs, we use $B_i, i = 1, 2, 3, 4$ to refer to the values of each band within an multispectral satellite image (MSI).

To further investigate the reflectance of different materials, some statistics were processed. Specifically, pixels belonging to ships, clouds, islands, and seawater are collected from multispectral images to establish their spectral reflectance signatures in the four bands for classification. The image of 2000 ships (19,530 pixels) and 20,000 seawater pixels under the conditions of shadows, mist, clouds, and clear sky are selected; 15,422 island pixels and 20,500 cloud pixels are also gathered. Details of the environmental conditions classification are explained in Section 5.1. Here, we mainly applied

Fourier transform [29] to characterize clouds in a given tile of an image. All come from four different multispectral remote sensors and are converted to reflectance values.

To calculate the reflectance value from the digital number (DN) of the pixel value in MSIs, the first step is to convert the DN to radiance:

$$Rad(B_i) = DN(B_i) \cdot G(B_i) + b(B_i) \quad (1)$$

where $i \in 1, 2, 3, 4$ is the number of each spectral band, Rad is the radiance value ($W \times m^{-2} \times sr^{-1} \times \mu m^{-1}$), and G and b are the band-specific rescaling gain factor and bias factor in one radiometric, respectively. Then, the radiance is converted into reflectance:

$$Ref(B_i) = \frac{Rad(B_i) \cdot \pi \cdot d^2}{E(B_i) \cdot \cos(\theta)} \quad (2)$$

where Ref is the reflectance value, E is the band-specific exoatmospheric irradiance ($W \times m^{-2} \times sr^{-1} \times \mu m^{-1}$), d is the Earth–Sun distance on the day of acquisition for the specific scene in astronomical units, and θ is the solar zenith angle in degrees. Values for the band-specific rescaling gain factor, bias factor, and the exoatmospheric irradiance used in this paper (see Tables A1 and A2 for details) are detailed by CRESDA (China Centre for Resources Satellite Data and Application). Values for the scene acquisition day and the solar zenith angle are packaged together with the image data of that scene. The Earth–Sun distances can be calculated using the day of scene acquisition [30].

The average reflectance values for ships, seawater, islands, and clouds collected under different conditions and their average values are shown in Figure 3. The dots in the graph represent the reflectance values for each individual pixel, while the dashed lines represent the average values for each category. In the analysis, we found that the same substances show different reflection curve clusters. The seawater pixels concentrate in three sets of reflection clusters, as the pixels came from three distinct areas: clear seawater, shadowed seawater, and high sediment seawater. Pixels belonging to clouds were gathered into two sets of clusters due to their thickness differences. The island pixels were clustered according to whether they were covered by plants into two groups. The ships were grouped into five subcategories based on their reflection changes. Especially, the reflectance of some clear seawater was similar to that of ships. The nuances between different clear seawater were scrutinized by applying different linear enhancements in each band image. It turns out that this clear seawater, similar to the reflection of a ship, was actually covered with mist.

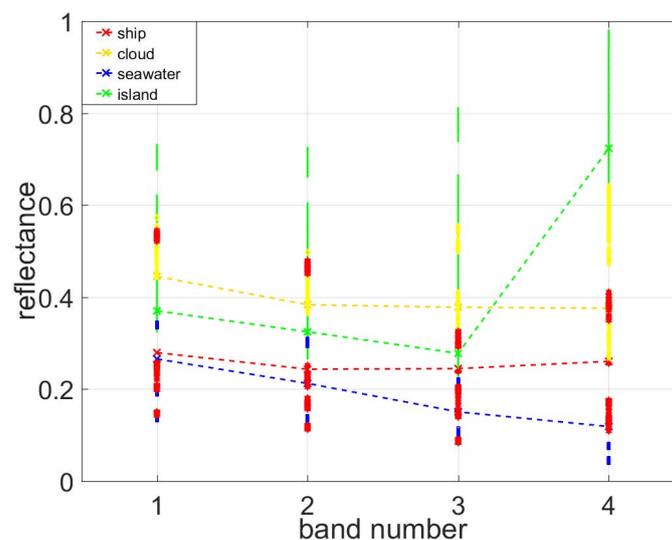


Figure 3. The reflectance values of seawater, clouds, islands, and ships collected from multispectral images in each band.

2.2. Feature Construction

Some objects in Figure 3 have overlapping ranges of reflectance values in these four bands, so direct application of reflectance values for classification may have limited performance. However, the identical material collected under equal situations has the same gradient for the reflectance values in the four bands. Considering that each spectral band for a satellite sensor has a distinctive central wavelength and that the same spectral bands for different satellites also have nuances of central wavelength, the gradient calculation formula is designed as follows:

$$grad(B_i, B_j) = \frac{Ref(B_i) - Ref(B_j)}{\lambda_{B_i} - \lambda_{B_j}} \tag{3}$$

where $i \neq j, i, j \in 1, 2, 3, 4$, and λ_{B_i} and λ_{B_j} are the central wavelengths of the i th and j th spectral bands, respectively. To amplify the result of the gradient, λ_{B_i} used units of μm . In this way, the gradient obtained was less affected by the parameters of the sensor. As shown in Figure 4, the reflectance gradient values for ships, seawater, clouds, and islands in groups $grad(B_1, B_3)$ and $grad(B_2, B_3)$ and in groups $grad(B_1, B_2)$, $grad(B_1, B_4)$ and $grad(B_2, B_4)$ have analogous distributions. Thus, combining them all into one feature set would be redundant. Consequently, one of these two groups will be chosen to compose a feature set of three reflectance gradients. Figure 5 shows an example of the values for the seawater, clouds, islands, and ships in three reflectance gradients. It displays that the island pixels are linearly separable from ship, seawater, and cloud in these three-dimensional reflectance gradient features. Compared to the one-dimensional reflectance gradient in Figure 4, higher-dimensional features are needed in order to distinguish among ship, seawater, and cloud. Moreover, a proper classifier is needed to accurately classify each category.

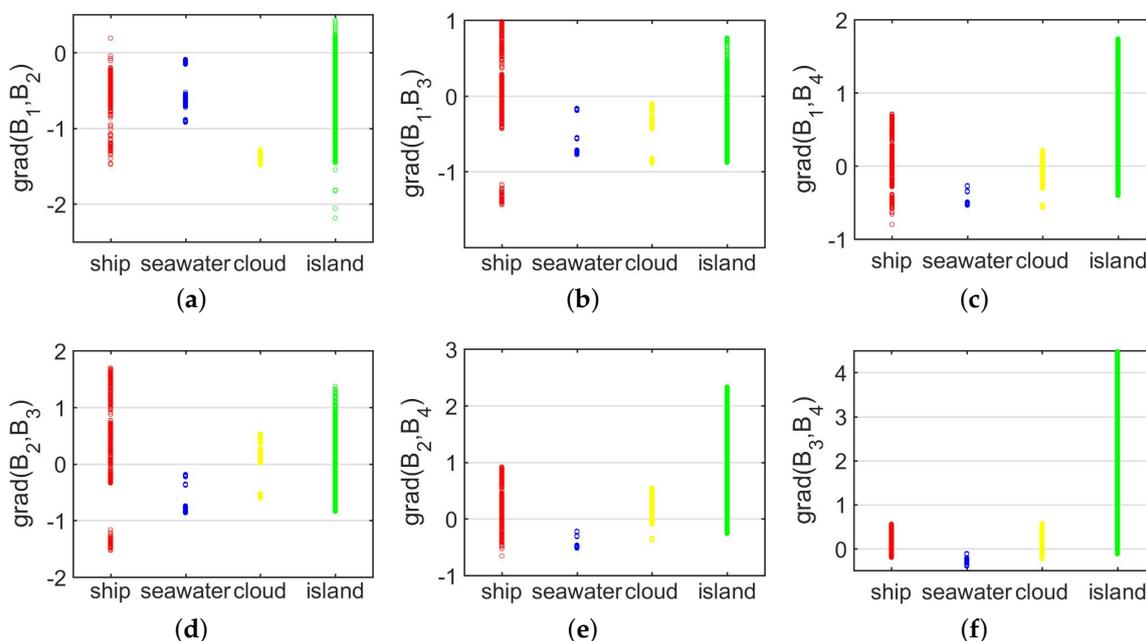


Figure 4. The reflectance gradients between two different bands for ships, seawater, clouds, and islands: (a) the reflectance gradients between band 1 and band 2; (b) the reflectance gradients between band 1 and band 3; (c) the reflectance gradients between band 1 and band 4; (d) the reflectance gradients between band 2 and band 3; (e) the reflectance gradients between band 2 and band 4; and (f) the reflectance gradients between band 3 and band 4.

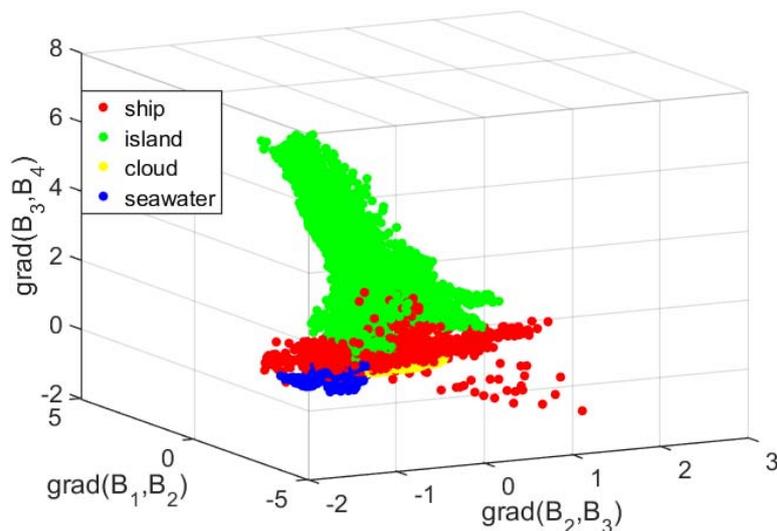


Figure 5. The 3-dimensional reflectance gradient values of seawater, clouds, islands, and ships.

2.3. Classification of Ship Candidates

To separate the ship from the background, support vector machine (SVM) with Gaussian kernel and random forest (RF) are adopted as basic classifiers based on two classes and four classes for analysis and comparison.

The first two methods employ SVM as the basic classifier. One considered the classification as a two-class recognition problem, named SVM-2c; the other one was a general multiclass classification strategy, named SVM-mc. The last two methods used RF as the classifier; one was binary classification (RF-2c), and the other was multiple classification (RF-mc). In SVM-2c and RF-2c, all samples were classified into two classes: ship and nonship. In SVM-mc and RF-mc, the nonship class includes several subclasses: seawater, cloud, and island. In these four methods, the classifier used all of the above reflectance gradients as classification features and trained the classifiers on training set I. The results in Section 5 show that RF-2c has won the competition. Thus, this paper uses random forest as the basic classifier with reflectance gradient features to classify ship candidates. After a close look at the distribution of the reflectance gradient, we find that some reflectance gradients are unnecessary to include. Therefore, the reflectance gradient feature set with a higher recall rate was employed finally. In this step, pixels that did not belong to ships were eliminated and set to zero to obtain a binary map of the ship. Afterwards, the connected component-labeling algorithm was used to label eight connected objects (areas larger than 3 pixels were retained). Then, the bounding boxes of the ship candidates were drawn according to the binary map for further verification.

3. Fine Detection Based on LFNet

In this section, we introduce the details of fine detection.

3.1. The Basic Idea of Feature Fusion

As shown in Equation (2), the reflectance image is nonlinearly related to the digital numbers of the original MSI when pixels come from different scenes. Therefore, the combination of the reflectance image with the original image will provide more useful information. To learn the spatial features with color information along with the spatial features with reflectance information, we propose a feature fusion based on a synthesized false color image and a reflectance image. The methods in References [31,32] demonstrated that the blue band supplies less effective information than the other three bands in detection tasks. After a comprehensive exploration of the combination with different spectral bands, a three-channeled false color image synthesized from bands 4, 3, and 2, arranged

as red, green, and blue in the three channels, respectively, was used for feature fusion with the reflectance image.

3.2. Network Configuration

A binary tree structure consisting of two input paths, called lightweight fusion network (LFNet), was used because it ensures that each input path is mixed with features first from its own input channel. As suggested in Reference [33], we use the grouped convolutional filter in the first two layers because it learns better representations and reduces the number of parameters. In each branch of the LFNet, there are two layers of grouped convolutional filters and a max pooling layer. The features extracted by these two branches were then concatenated and shuffled for fusion and finally flowed into one mainstream for higher feature representation and classification. In the mainstream of the LFNet, the depthwise separable convolution technology [16] was applied, followed by a global pooling and a fully connected convolution for final prediction. The complete network structure is shown in Figure 6. As displayed in the graph, the pink colored branch of the LFNet receives the synthesized false color image for feature extraction while the orange colored branch receives the reflectance image.

Table 1 illustrates the overall parameters of the proposed network. LFNet consists of seven weight layers, including six convolutional layers and one fully connected layer. During training, the output of the last fully connected layer was fed to the sigmoid function, which produced a prediction probability over the target category label.

Table 1. Detail parameters in the lightweight fusion network. Key: Output groups are the number of group convolutions in each layer and each group; FC denotes fully convolutional layer.

Layer	Output Size	Kernel Size	Output Groups (g)			
			$g = 1$	$g = 2$	$g = 3$	$g = 4$
Color Image	32×32	–	3			
<i>GConv1_1</i>	26×26	7×7	11	11	13	–
<i>GConv1_2</i>	20×20	7×7	2	2	1	–
MaxPool	10×10	2×2	57			
Reflectance Image	32×32	–	4			
<i>GConv2_1</i>	26×26	7×7	8	12	11	12
<i>GConv2_2</i>	20×20	7×7	1	3	2	3
MaxPool	10×10	2×2	102			
<i>DWConv2_4</i>	6×6	5×5	159			
<i>GConv2_5</i>	6×6	1×1	159			
GlobalPool	1×1	6×6	159			
FC	1×2	159×2	2			

For a comparative analysis, a similar network was also implemented, removing the branch that receives the synthesized false color image of the proposed LFNet. In this network, only the reflectance images of the original four-band multispectral image were received. By implementing this network, the benefits of synthesized false color images in feature fusion can be explored.

Assume that one of the input groups \mathfrak{S} has a size of $W \times H \times D$ and that one grouped convolutional layer exists in any branch of LFNetwork, denoted as \mathcal{W} , we have the output feature \mathcal{F} , represented by the following:

$$\mathcal{F} = g(\mathfrak{S} * \mathcal{W}) \quad (4)$$

where $g(\cdot)$ is the sigmoid activation function defined as

$$g(\theta) = \frac{1}{1 + e^{-\theta}} \quad (5)$$

and $*$ is the convolution operation.

The overall structure and design of the LFNet is a brand new idea, but some of its components were learned from other inventors. The idea behind the design and structure of the first two layers of the LFNet branch is inspired by SVDNet [17], but it has entirely different convolutional filter components. Instead of using ordinary convolution, the proposed network utilizes the grouped convolution to learn a more sparse representation. Additionally, a max pooling layer is added after the second convolutional layer to retain as much texture information as possible. The component after the concatenation layer in the mainstream of LFNet is a simplified version of the shuffle unit from ShuffleNet [15] and the depthwise separable convolution unit from MobileNets [16].

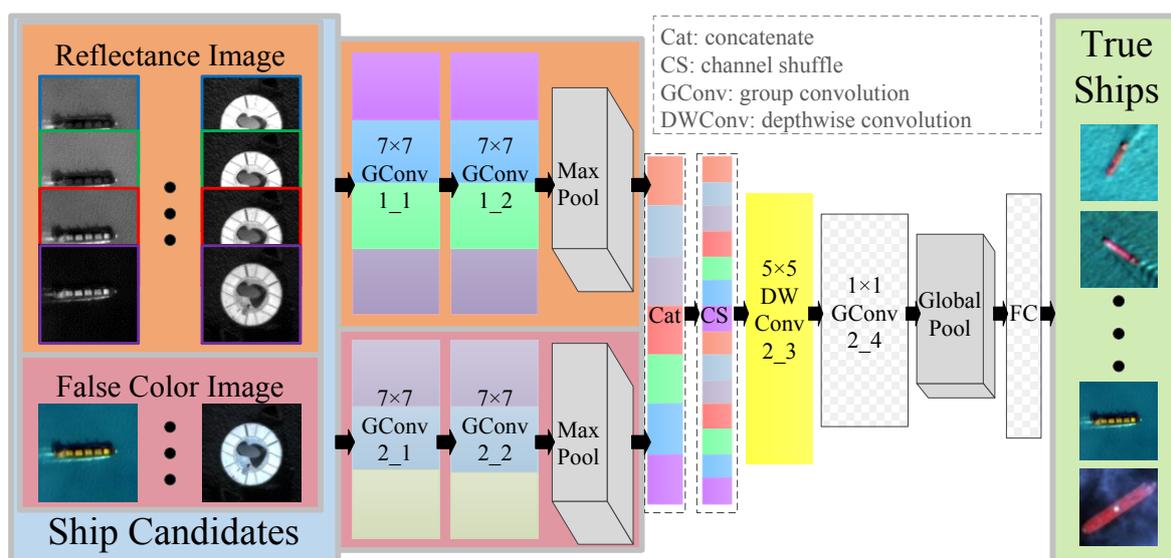


Figure 6. The architecture of the proposed network: The pink colored branch of the network receives the synthesized false color image input with two grouped convolution (GConv) layers and a max pooling layer, and the orange colored branch receives reflectance image input having the same structure as the pink colored branch. The mainstream consists of GConv, depthwise convolution (DWConv), global pooling layer, and a fully connected layer.

4. Training for Detection

In the coarse stage, we adjusted the two parameters in the random forest classifier: number of trees to grow n_{tree} and number of variables randomly selected at each node m_{try} . After experiments on different combinations of $n_{tree} = \{20, 30, 35, 40, 50\}$ and $m_{try} = \{1, 2, 3\}$, we set the random forest with parameters $m_{try} = 2$ and $n_{tree} = 35$. In fact, the increase in precision and recall begins to slow down when $n_{tree} = 30$ and $m_{try} = 2$. This classifier trains on training set I.

In the fine stage, a training set Π (consisting of patches with positive and negative samples) is used to learn through an unsupervised learning algorithm to first obtain the parameters of the first two layers of LFNet's two branches. After that, all of the filter banks were trained on the same training set by a supervised learning algorithm. Hence, LFNet's learning process is divided into two steps: unsupervised learning and supervised learning.

For the two grouped convolutional layers in the branches of LFNet, the efficient singular value decomposition (SVD) unsupervised learning algorithm was used to represent the eigenvectors [17,34]. Assume that the randomly selected patches I_i for group g of the input layer with size $k \times k$ can be reshaped into a vector and $I = [I_1, I_2, \dots, I_n] \in R^{k^2 \times n}$. Then, the SVD of the matrix I is

$$I = U\Sigma V^T \quad (6)$$

where $U \in R^{k^2 \times k^2}$ and $V \in R^{n \times n}$ are orthogonal unitary matrices and $\Sigma \in R^{k^2 \times n}$ is a rectangular diagonal matrix with the singular value σ_i of I on its diagonal. Then, we choose u_i from the U with the largest singular values that fulfill the following condition:

$$N = \{ \min(i) \mid \frac{\sigma_1 + \dots + \sigma_i}{\Sigma} > 0.99 \} \quad (7)$$

where N is the number of filter banks. The filter banks belonging to group g are then obtained from U by forming u_i into 2_D filters. Different from the usage of SVDs in Reference [17], this unsupervised learning algorithm is applied to the input layers of more than one channel of the grouped convolutional layers. The number of filters in each layer is determined by Equation (7).

In the second step of LFNet training, each ship target sample in the training set is labeled with a ground-truth class u (+1 for ships and -1 for nonships). The log loss L is applied on each labeled ship target sample for training for final prediction as follows:

$$L(y, u) = \log(1 + e^{-yu}) + \beta \|w\|_2^2 \quad (8)$$

where β is a weight decay parameter, y is the output from the last layer, and w is the weight of the fully connected layer. Then,

$$y = w^T x \quad (9)$$

where x is the output from the upper layer. The prediction error calculated by the log loss will propagate back by the gradient from the last layer to the front layers except for the unsupervised-learning layer. Finally, they are fine-tuned together.

5. Experiment and Discussion

In the following section, extensive experiments are carried out to verify the effectiveness of each stage in the proposed method. First, images with pixels containing ships, seawater, islands, and clouds are prepared to verify the performance of the proposed ship candidate-extraction method. Second, the performance of LFNet with different parameters is evaluated and the results are given in detail. Third, an experiment comparing several other methods with the proposed algorithm is conducted.

The assessment criteria in Reference [35] are used here, and the definitions are as follows:

$$precision = \frac{n_{tp}}{n_{tp} + n_{fp}} \quad (10)$$

$$recall = \frac{n_{tp}}{n_{tp} + n_{fn}} \quad (11)$$

$$F_1 = \frac{2 \times precision \times recall}{precision + recall} \quad (12)$$

where n_{tp} is the number of true positives, n_{fp} represents the number of false positives, and n_{fn} denotes the number of false negatives. *precision* and *recall* are the precision rate and recall rate, respectively. F_1 is the F1-measure. So far, no public multispectral satellite ship detection dataset contains related satellite parameters used for reflectance calculation, and few researchers have published their source code or software. Therefore, we do not directly compare this method with other methods.

5.1. Data Description

We tested our method on forty-eight images, twelve each from GaoFen-1 (GF1), ShiJian-9 (SJ9), ZiYuan-3 (ZY3), and CBERS-04 (CB04) MSIs taken at different times and locations and containing coastal landscapes. The images from the first three satellites have varying spatial resolutions, while images from the fourth have a different resolution and quantization level from the first three sets of images. Detailed information on these images is presented in Table 2. As the proposed method takes advantage of the spectral features of the original images, the wavelength of each band and the quantization level of each satellite sensor are provided for comprehensive analysis.

Table 2. Detailed information on experimental images. Key: Qt. = Quantization.

Satellite	Resolution	Size (Pixel)	Wavelength (μm)				Qt.
			Band1	Band2	Band3	Band4	
GF1	8 m	12,000 \times 13,400	0.450~0.520	0.520~0.590	0.630~0.690	0.770~0.890	10 bit
SJ9	10 m	4548 \times 4544	0.450~0.520	0.520~0.590	0.630~0.690	0.770~0.890	10 bit
ZY3	6 m	8816 \times 8792	0.450~0.520	0.520~0.590	0.630~0.690	0.770~0.890	10 bit
CB04	20 m	6500 \times 6500	0.450~0.520	0.520~0.590	0.630~0.690	0.770~0.890	8 bit

Two training sets were collected from images produced by all the satellites referenced above: details of training set I can be found in Section 2.1: training set II including 60,780 ship target samples and 69,475 negative samples after data augmentation. The data augmentation performed here is 90° rotations each time and scaling each sample by downsampling to 0.8, 0.5, and 0.2 scales of its original size.

In terms of testing, we prepared a test set for each stage and a test set for a comprehensive comparison. Test set I includes 4660 ship pixels, 7120 seawater pixels, 5764 cloud pixels, and 5544 island pixels; test set II contains 6379 ship target patches and 11,256 nonship-target patches; and test set III contains 40 MSIs, including 1372 ship targets in various weather conditions, i.e., clean, wavy, and cloudy sea situations. The experimental images in test set III were divided into three categories based on the environment on the images: “wavy”, “cloudy”, and “clean” refers to the images that contain environments such as heavy waves, clouds, or clean sea conditions. When classifying the images based on environmental conditions, we first characterized the clouds in a given image and then characterized waves. If the results indicated that no cloud and wave conditions were found, the given image was classified as a clean image. Specifically, we first traversed the image based on the ocean background characterization algorithm in Reference [29] to determine whether a given image contains clouds. If the tiles (200 by 200 pixels suggested by Reference [29]) in the ocean part of a given image containing clouds exceed thirty percent of its ocean part, the given image was classified as “cloudy”; otherwise, we traversed the image to determine the sea surface roughness measured by the standard deviation (recommend by References [36,37]). If the tiles (512 by 512 pixels suggested by Reference [38]) in the ocean part of a given image has a standard deviation larger than 15 (empirically set), exceeding fifty percent of its ocean part, the given image was classified as “wavy”. A given image that dissatisfied both of the above conditions was classified as “clean”. To apply the original one-band-based clouds and sea states characterization algorithm in multispectral images, we used the average result from each band as the final result for a given tile. After automatically classifying environmental conditions, we manually checked and corrected the results.

In the statistics, ships with a length of more than five pixels (ships with a length of more than 100 m in CB04, ships with a length of more than 80 m in GF1, ships with a length of more than 50 m in SJ9, and ships with a length of more than 30 m in ZY3) were counted because smaller pixel clusters have very limited object details/features to extract and are barely recognizable by the human eye.

5.2. Reflectance Gradient Analysis and Classification

The experiment in this section was performed on test set I. All sampled pixels were divided into four categories: ship, cloud, seawater, and island. In the following experiments in this section, all classifiers were trained from the training set I. Considering that the reflectance gradient classification was the first step in the whole process, we pursued a higher recall rate than the precision rate in the following experiments in this section.

The first experiment analyzed the performance of the classification ability in which each reflectance gradient was treated as a classification feature. The experiment was based on the selection of each different reflectance gradient combined with the classifier. Figure 7 clearly shows that each feature had a lower classification ability for ships than for seawater. That is, each reflectance gradient combined with SVM was a weak classifier for ship classification, which leads to the second experiment.

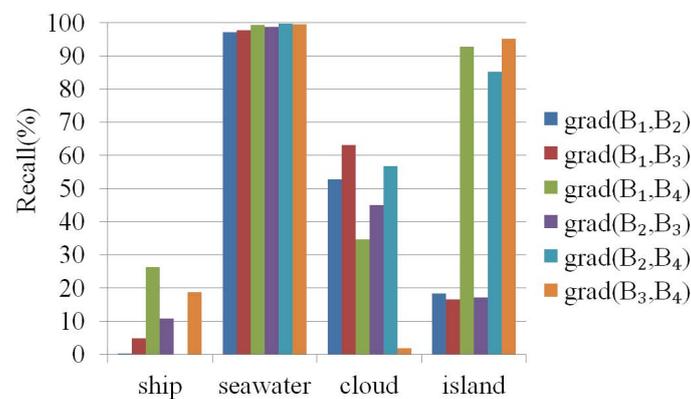


Figure 7. Performance of classification based on every single feature.

The second experiment adopted all reflectance gradients as classification features to compare the classification capabilities of the SVM and RF. In this experiment, the aforementioned four classification approaches were compared, and their results are listed in Table 3. It is easy to find that the precision and recall of RF was much higher than SVM, which is inferable from the results of the first experiment. The classification ability of each feature for different objects varies, which leads to their divergent importance in filtering ships. Random forests classifier has the capability to estimate the importance of variables [39], which makes it more suitable for classification used in combination with these features than support vector machine. Moreover, the recall of RF-mc based on four categories was lower than the recall of RF-2c based on two-class classification. Therefore, RF-2c was adopted in the following experiment.

Table 3. The classification result of different approaches: Testing time is the classification time per megapixel, including the calculation of reflectance gradients; for definitions of abbreviations, see Section 2.3.

Method	SVM-2c	SVM-mc	RF-2c	RF-mc
Recall (%)	77.79	84.52	99.05	98.71
Precision (%)	92.36	91.83	97.92	98.08
Testing time (s)	18	22	19	19
Training time (s)	67	109	34	38

As mentioned previously, the distribution of reflectance gradient values for ships, seawater, clouds, and islands in groups $grad(B_1, B_3)$ and $grad(B_2, B_3)$ and in groups $grad(B_1, B_2)$, $grad(B_1, B_4)$, and $grad(B_2, B_4)$ was similar. To reduce the redundancy of classification features and to maintain good performance, a third experiment was conducted. We selected one feature from $grad(B_1, B_3)$ and $grad(B_2, B_3)$ and the other from $grad(B_1, B_2)$, $grad(B_1, B_4)$, and $grad(B_2, B_4)$ and then combined them with a fixed feature $grad(B_3, B_4)$ to compose a feature set. Table 4 listed the details of feature set combinations. Then, we obtained six feature sets listed in Figure 8a. These feature sets based on RF were adopted to analyze the performances of different combinations of feature sets, as shown in Figure 8a. set4 achieved the highest recall rate when using feature sets as classification features and 0.1% lower than the recall rate using all reflectance gradients as classification features. The low precision rate mainly came from the misclassification of mist-covered seawater and other man-made objects at sea. Most misclassified mist-covered seawater scattered on the binary map of ship candidates can be screened out easily by area. Thus, we adopted the set4 combined with RF in the following experiments.

Table 4. The feature sets for contrast.

Featureset	Set1	Set2	Set3	Set4	Set5	Set6
features	$grad(B_1, B_3)$	$grad(B_1, B_3)$	$grad(B_1, B_3)$	$grad(B_2, B_3)$	$grad(B_2, B_3)$	$grad(B_2, B_3)$
	$grad(B_1, B_2)$	$grad(B_1, B_4)$	$grad(B_2, B_4)$	$grad(B_1, B_2)$	$grad(B_1, B_4)$	$grad(B_2, B_4)$
	$grad(B_3, B_4)$					

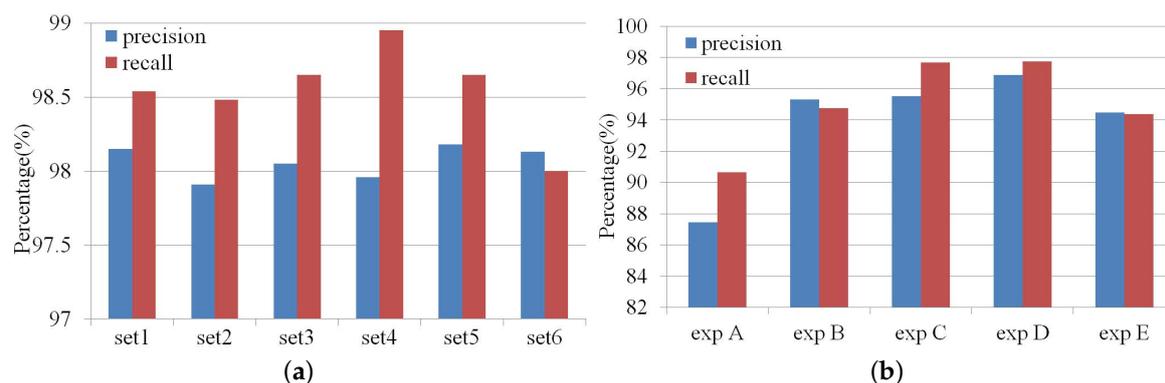


Figure 8. (a) Performance of classification based on every feature set and (b) performance of classification based on LFNNet with/without color image input with different band combination.

5.3. Color Image Fusion Experiment

In this experiment, we implemented an LFNNet with its branch that receives the synthesized false color image cutoff, which means that the reflectance image was the only input for the network, named *exp A*. The remaining control groups were all implemented through a fully structured LFNNet, but the input-synthesized false color image differed in band combination. The synthesized false color image of *exp B* treated band 4, band 2, and band 1 as red, green, and blue, respectively; the synthesized false color image of *exp C* used band 4, band 3, and band 1; group *exp D* synthesized from band 4, band 3, and band 2; and the last group *exp E* synthesized from a combination of band 3, band 2, and band 1. They were conducted under a maximum number of training iterations of $T = 1000$ on the training set II and tested on the test set II.

In Figure 8b, the result of *exp A* shows that the precision fell outside 90% without the input of color image. As we also give the predictions for other band combinations, the combined band 3, band 2, and band 1 as red, green, and blue, respectively, was not the best choice for ship detection. Group *exp D* achieved the highest precision and recall.

5.4. Parameter Optimization

We explored the performance of the LFNNet under a maximum number of iterations of $T = 1000$ by measuring the object proposal capability against our test set II under different learning rates and weight decay coefficients. Figure 9 shows the prediction results. The figure shows that the LFNNet obtained the highest performance when the learning rate was set to $\mu = 5 \times 10^{-4}$ and the weight decay coefficient was set to $\beta = 1 \times 10^{-3}$.

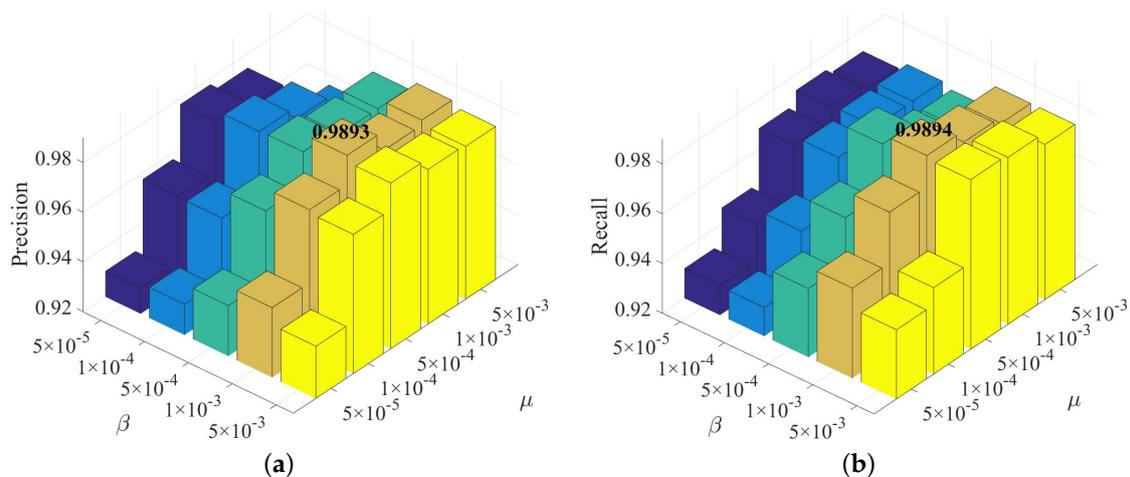


Figure 9. Ship-detection results of the LFNNet measured in terms of precision and recall over ship and nonship classes under different parameter settings: (a) precision rate and (b) recall rate. The highest precision and recall rate values are indicated with bold numbers on their corresponding bars.

5.5. Ship-Detection Performance

In this experiment, the proposed method was verified on satellite images with complex backgrounds, changing resolutions, and variable quantization levels, with the aim of demonstrating its robustness. LFNNet was fine-tuned for optimal performance when used here.

For comparison, six algorithms were implemented and trained with training set II and then tested on test set III. The algorithms used were: the dense scale-invariant feature transform (denseSIFT) [40] with SVM [41]; local binary pattern [42] with SVM; histogram of oriented gradient (HOG) with SVM, as some methods used new derivatives of HOG as the feature extractor [43,44]; the method of Reference [9] (denoted as cdDNN) and SVDNet [17] with band 4 as its input; and YOLT [20] with RGB color image as its input. Figure 10 displays the precision-recall curves of these four methods and the proposed method in this paper (denoted as Spec + LFN). The proposed algorithm was clearly superior to the other six algorithms in performance in test set III. Table 5 shows the F1-measurement of these algorithms and the proposed algorithm. The outcome indicates that the proposed method had relatively similar performances in different environments. Other algorithms achieved their best performance in clean sea situations.

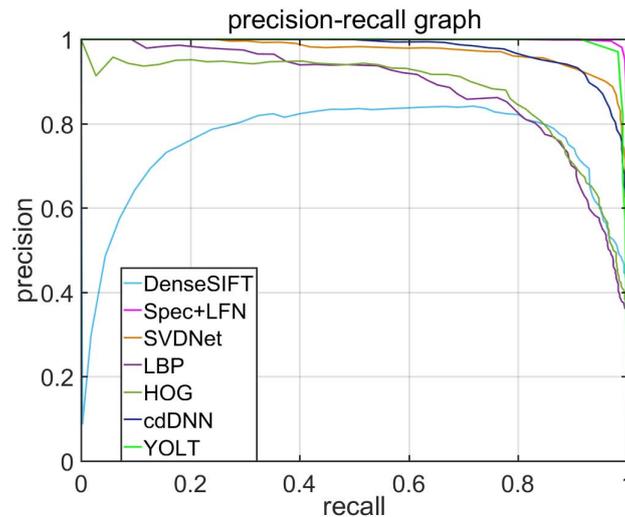


Figure 10. Precision-recall curves of six different feature descriptors: HOG, LBP DenseSIFT, SVDNet, cdDNN, YOLT, and Spec+LFN; for definitions of abbreviations, see Section 5.5.

Table 5. F1-measurement of performance on experimental images.

Condition	Clouds	Waves	Clean
DenseSIFT	0.830	0.841	0.865
LBP	0.810	0.806	0.839
HOG	0.824	0.826	0.855
cdDNN	0.934	0.924	0.954
SVDNet	0.749	0.986	0.959
YOLT	0.951	0.972	0.978
Spec + LFN	0.976	0.974	0.977

Some typical images from GF1, SJ9, ZY3, and CB04 are shown in Figure 11. The first column in Figure 11 presents a few detection results of multiscale ship targets in relatively calm seas with no clouds. The second column shows the ship-detection results under mist both with and without ocean waves. The pictures in the last column show successful detections of ships beneath dense clouds.

Based on the analysis of the experimental results, we can conclude that the performance of the proposed method in ship detection is rarely affected by cloudy weather, waves, resolution, or quantization level of the sensors. Furthermore, the performance of the proposed model can be related to the discussed parameters of the sensors except for the resolution and quantization level.

The utilization of spectral features in regular four-band multispectral images when performing ship detection has been analyzed. The reflectance value of the objects related to the ship-detection task itself is susceptible to weather conditions and may overlap each other in some cases; however, its gradient remains relatively stable under different conditions and therefore allows the relatively low spectral resolution of four-band multispectral images to apply this feature to coarse ship detection but with limited precision. By exploiting the gradients between different two bands, one can separate ships from most seawater, cloud, and island pixels. The difference among the mist-covered seawater, ships, and other man-made objects is indistinguishable using only the reflectance gradient features due to the low spectral resolution of the four-band multispectral image.

In the verification stage, the optimal synthesized false color image for feature fusion has been analyzed. The result showed that the false color image synthesized by selected bands (NIR–red–green) outperforms the traditional red–green–blue band-synthesized false color image in ship detection and other synthesized false color images. To our knowledge, this is the first evidence that three-channel false color images synthesized from different band combinations from the multispectral remote-sensing

image have different ship-detection performances. The explanation for this result may lie in the enhancement effect of the selected bands.

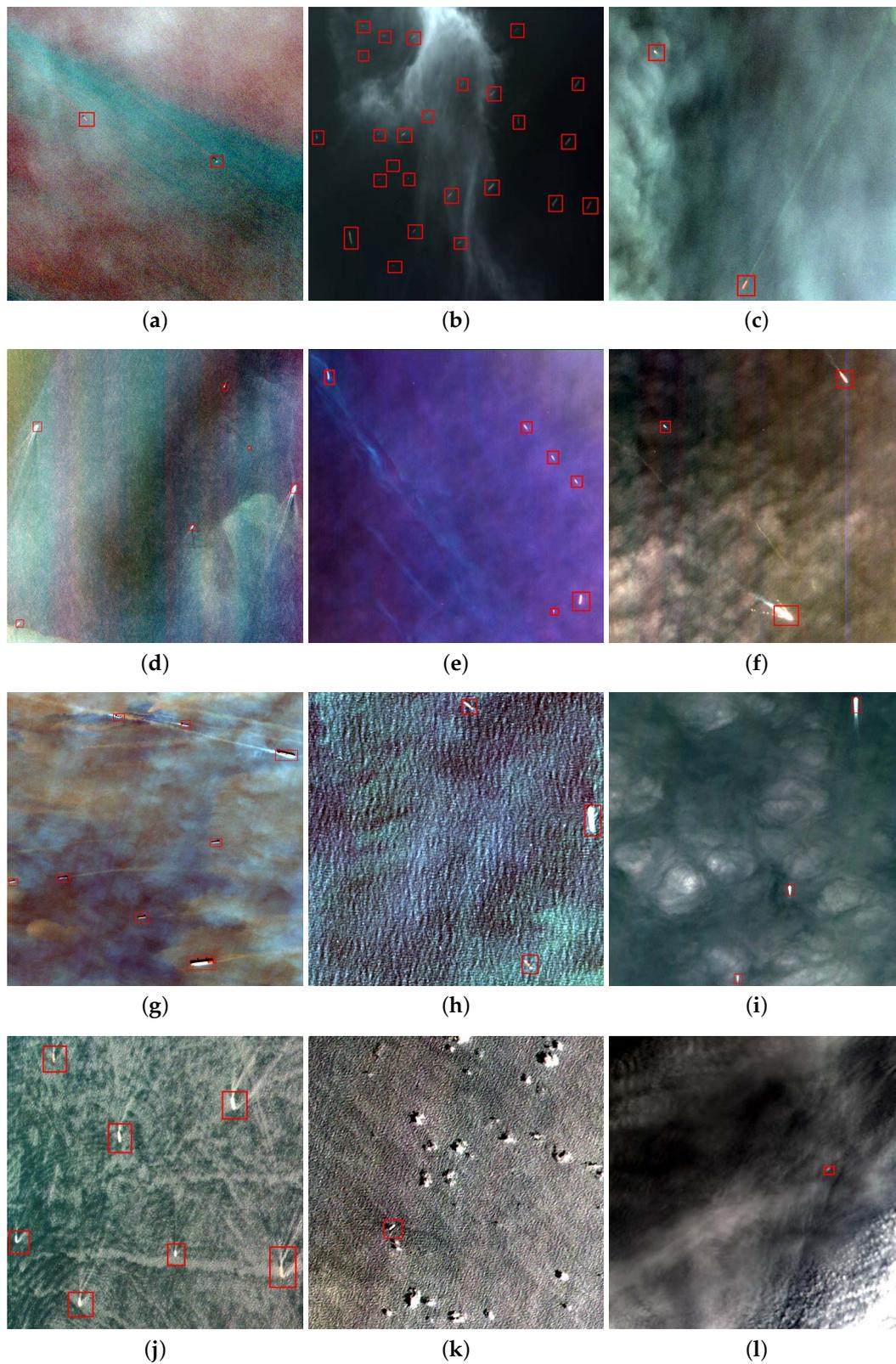


Figure 11. The performance of the proposed method under different conditions: **(a,d,g,j)** Calm, **(b,e,h,k)** waves and **(c,f,i,l)** mist and thick clouds. **(a–c)** represents images from GF1. **(d–f)** represents images from SJ9. **(g–i)** represents images from ZY3. **(j–l)** represents images from CB04.

6. Conclusions

This paper proposes a novel method of ship detection based on the spectral features and a CNN-based network with a binary tree structure known as LFNet in a coarse-to-fine manner. It analyzed the spectral reflectance relationship between the targets to be detected and objects that may obstruct the detection under various weather and environmental conditions. It is observed that their reflectance values may overlap in some cases; however, their reflectance gradients remain relatively unchanged under complex environment. The proposed features designed based on this characteristic combined with the random forests classifier to extract ship candidates with limited precision but high recall. In the second stage, the LFNet with two input paths received information from the synthesized false color image and the reflectance image for ship verification. It is observed that the synthesized false color image treating band 4, band 3, and band 2 as red, green, and blue, respectively, combined with the reflectance image in feature fusion performs better than other options.

The experiments conducted on a set of multispectral images from the GF1, SJ9, ZY3, and CB04 satellites demonstrate the effectiveness and robustness of method with 8-m, 10-m, and 6-m resolutions with a 10-bit quantization level and with a 20-m resolution with an 8-bit quantization level. The idea of spectral reflectance gradient features is transferable to object detection on multispectral images or hyperspectral images.

Author Contributions: Conceptualization, X.X. and B.L.; methodology, X.X. and X.W.; software, X.X.; validation, X.X., B.L. and X.W.; formal analysis, X.X.; investigation, X.X.; resources, B.L. and X.W.; data curation, X.X. and X.W.; writing—original draft preparation, X.X.; writing—review and editing, X.W.; visualization, X.X.; supervision, B.L.; project administration, B.L. and X.W.; funding acquisition, B.L. and X.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by NSFC grant number 61806109.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A

Table A1. Exo-atmospheric irradiance used in the experiment ($W \times m^{-2} \times \mu m^{-1}$).

Sensor Name	Band1	Band2	Band3	Band4
GF-1 PMS1	1944.98	1854.42	1542.63	1080.81
GF-1 PMS2	1945.34	1854.15	1543.62	1081.93
CB-04 MUX [45]	1958	1852	1559	1091
SJ-9A MUX	1942.93	1854.03	1543.25	1080.87

Table A2. Satellite parameter used in the experiment from CRESDA (China Centre for Resources Satellite Data and Application; $W \times m^{-2} \times sr^{-1} \times \mu m^{-1}$).

Sensor Name	Year	Band1		Band2		Band3		Band4	
		Gain	Bias	Gain	Bias	Gain	Bias	Gain	Bias
GF-1 PMS1	2015	0.211	0	0.1802	0	0.1806	0	0.187	0
	2016	0.232	0	0.187	0	0.1795	0	0.196	0
	2017	0.1424	0	0.1177	0	0.1194	0	0.1135	0
GF-1 PMS2	2015	0.2242	0	0.1887	0	0.1882	0	0.1963	0
	2016	0.224	0	0.1851	0	0.1793	0	0.1863	0
	2017	0.1460	0	0.1248	0	0.1274	0	0.1255	0
CB-04 MUX	2015	0.6575	0	0.6303	0	0.6145	0	0.5369	0
SJ-9A MUX	2014	0.1830	−2.1067	0.1660	−1.8065	0.1570	−0.4350	0.1428	−0.3639
	2015	0.1789	0.953	0.1608	−2.1306	0.1519	−2.176	0.1479	−1.2396

References

1. Zhu, C.; Zhou, H.; Wang, R.; Guo, J. A Novel Hierarchical Method of Ship Detection from Spaceborne Optical Image Based on Shape and Texture Features. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 3446–3456. [[CrossRef](#)]
2. Tian, T.; Pan, Z.; Tan, X.; Chu, Z. Arbitrary-Oriented Inshore Ship Detection based on Multi-Scale Feature Fusion and Contextual Pooling on Rotation Region Proposals. *Remote Sens.* **2020**, *12*, 339. [[CrossRef](#)]
3. Wu, Y.; Ma, W.; Gong, M.; Bai, Z.; Zhao, W.; Guo, Q.; Chen, X.; Miao, Q. A Coarse-to-Fine Network for Ship Detection in Optical Remote Sensing Images. *Remote Sens.* **2020**, *12*, 246. [[CrossRef](#)]
4. Kanjir, U.; Greidanus, H.; Ostir, K. Vessel detection and classification from spaceborne optical images: A literature survey. *Remote Sens. Environ.* **2018**, *207*, 1–26. [[CrossRef](#)] [[PubMed](#)]
5. Burgess, D.W. Automatic ship detection in satellite multispectral imagery. *Photogramm. Eng. Remote Sens.* **1993**, *59*, 229–237.
6. Wu, G.; de Leeuw, J.; Skidmore, A.K.; Liu, Y.; Prins, H.H. Performance of Landsat TM in ship detection in turbid waters. *Int. J. Appl. Earth Obs. Geoinf.* **2009**, *11*, 54–61. [[CrossRef](#)]
7. Ding, Z.; Yu, Y.; Wang, B.; Zhang, L. An approach for visual attention based on biquaternion and its application for ship detection in multispectral imagery. *Neurocomputing* **2012**, *76*, 9–17. [[CrossRef](#)]
8. Heiselberg, H. A Direct and Fast Methodology for Ship Recognition in Sentinel-2 Multispectral Imagery. *Remote Sens.* **2016**, *8*, 1033. [[CrossRef](#)]
9. Tang, J.; Deng, C.; Huang, G.B.; Zhao, B. Compressed-Domain Ship Detection on Spaceborne Optical Image Using Deep Neural Network and Extreme Learning Machine. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 1174–1185. [[CrossRef](#)]
10. Krizhevsky, A.; Sutskever, I.; Hinton, G. ImageNet Classification with Deep Convolutional Neural Networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 1097–1105. [[CrossRef](#)]
11. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. *arXiv* **2015**, arXiv:1512.03385.
12. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. *arXiv* **2016**, arXiv:1602.07261.
13. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. In Proceedings of the CVPR, Honolulu, HI, USA, 21–26 July 2017; Volume 1, p. 3.
14. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2014**, arXiv:1409.1556.
15. Zhang, X.; Zhou, X.; Lin, M.; Sun, J. Shufflenet: An extremely efficient convolutional neural network for mobile devices. *arXiv* **2017**, arXiv:1707.01083.
16. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *arXiv* **2017**, arXiv:1704.04861.
17. Zou, Z.; Shi, Z. Ship detection in spaceborne optical image with SVD networks. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 5832–5845. [[CrossRef](#)]
18. Cheng, G.; Zhou, P.; Han, J. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 7405–7415. [[CrossRef](#)]
19. Li, X.; Wang, S. Object Detection Using Convolutional Neural Networks in a Coarse-to-Fine Manner. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 2037–2041. [[CrossRef](#)]
20. Etten, A.V. You Only Look Twice: Rapid Multi-Scale Object Detection In Satellite Imagery. *arXiv* **2018**, arXiv:abs/1805.09512
21. Li, Q.; Mou, L.; Liu, Q.; Wang, Y.; Zhu, X.X. HSF-Net: Multiscale Deep Feature Embedding for Ship Detection in Optical Remote Sensing Imagery. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 7147–7161. [[CrossRef](#)]
22. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Li, F.-F. Imagenet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255.
23. Redmon, J.; Divvala, S.K.; Girshick, R.B.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. *arXiv* **2015**, arXiv:abs/1506.02640.
24. Clark, R.N.; Swayze, G.A.; Wise, R.; Livo, K.E.; Hoefen, T.; Kokaly, R.F.; Sutley, S.J. USGS Digital Spectral Library splib06a. 2007. Available online: <https://archive.usgs.gov/archive/sites/speclab.cr.usgs.gov/spectral.lib06/ds231/> (accessed on 25 February 2020).

25. DeFries, R.; Townshend, J. NDVI-derived land cover classifications at a global scale. *Int. J. Remote Sens.* **1994**, *15*, 3567–3586. [CrossRef]
26. McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [CrossRef]
27. Huang, X.; Zhang, L. An SVM ensemble approach combining spectral, structural, and semantic features for the classification of high-resolution remotely sensed imagery. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 257–272. [CrossRef]
28. Bernabe, S.; Marpu, P.R.; Plaza, A.; Dalla Mura, M.; Benediktsson, J.A. Spectral-spatial classification of multispectral images using kernel feature space representation. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 288–292. [CrossRef]
29. Buck, H.; Sharghi, E.; Bromley, K.; Guilas, C.; Chheng, T. Ship detection and classification from overhead imagery. *Appl. Digit. Image Process.* **2007**. 6696, 66961C.
30. USGS. Earth-Sun Distance (D) in Astronomical Units for Day of the Year. Available online: https://landsat.usgs.gov/sites/default/files/documents/Earth-Sun_distance.xls (accessed on 25 February 2020)
31. Teke, M.; Baseski, E.; Ok, A.O.; Yüksel, B.; Çağlar Ş. *Multi-spectral False Color Shadow Detection*; Springer: Berlin/Heidelberg, Germany, 2011.
32. Sun, L.; Mi, X.; Wei, J.; Wang, J.; Tian, X.; Yu, H.; Gan, P. A cloud detection algorithm-generating method for remote sensing data at visible to short-wave infrared wavelengths. *ISPRS J. Photogramm. Remote Sens.* **2016**, *124*, 70–88. [CrossRef]
33. Ioannou, Y.; Robertson, D.P.; Cipolla, R.; Criminisi, A. Deep Roots: Improving CNN Efficiency with Hierarchical Filter Groups. In Proceedings of the (CVPR), Las Vegas, NV, USA, 26 June–1 July 2016; pp. 5977–5986.
34. Dan, K. A Singularly Valuable Decomposition: The SVD of a Matrix. *Coll. Math. J.* **1996**, *27*, 2–23.
35. Powers, D.M.W. Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness and Correlation. *J. Mach. Learn. Technol.* **2011**, *2*, 2229–3981.
36. Yang, G.; Li, B.; Ji, S.; Gao, F.; Xu, Q. Ship Detection From Optical Satellite Images Based on Sea Surface Analysis. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 641–645. [CrossRef]
37. Kanjir, U.; Marsetič, A.; Pehani, P.; Oštir, K. An automatic procedure for small vessel detection from very-high resolution optical imagery. In Proceedings of the 5th GEOBIA, Thessaloniki, Greece, 21 May–24 May 2014; pp. 1–4.
38. Jubelin, G.; Khenchaf, A. A statistical model of sea clutter in panchromatic high resolution images. In Proceedings of the 2012 IEEE International Geoscience and Remote Sensing Symposium, Munich, Germany, 22–27 July 2012; pp. 424–427. [CrossRef]
39. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
40. Lindeberg, T. Scale Invariant Feature Transform. *Scholarpedia* **2012**, *7*, 10491. [CrossRef]
41. Amari, S.; Wu, S. Improving support vector machine classifiers by modifying kernel functions. *Neural Netw.* **1999**, *12*, 783–789. [CrossRef]
42. Ojala, T.; Pietikainen, M.; Maenpaa, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 971–987. [CrossRef]
43. Dong, C.; Liu, J.; Xu, F.; Liu, C. Ship Detection from Optical Remote Sensing Images Using Multi-Scale Analysis and Fourier HOG Descriptor. *Remote Sens.* **2019**, *11*, 1529. [CrossRef]
44. Qi, S.; Ma, J.; Lin, J.; Li, Y.; Tian, J. Unsupervised Ship Detection Based on Saliency and S-HOG Descriptor From Optical Satellite Images. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 1451–1455. [CrossRef]
45. Pinto, C.; Ponzoni, F.; Castro, R.; Leigh, L.; Mishra, N.; Aaron, D.; Helder, D. First in-Flight Radiometric Calibration of MUX and WFI on-Board CBERS-4. *Remote Sens.* **2016**, *8*, 405. [CrossRef]

