Article

Cloud Detection for High-Resolution Satellite Imagery Using Machine Learning and Multi-Feature Fusion

Ting Bai [1], Deren Li [1,2], Kaimin Sun [1,*], Yepei Chen [1] and Wenzhuo Li [2]

[1] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan 430079, China; baitinghzau@gmail.com (T.B.); drli@whu.edu.cn (D.L.); chenyp@whu.edu.cn (Y.C.)

[2] School of Remote Sensing and Information Engineering, Wuhan University, 129 Luoyu Road, Wuhan 430079, China; alvinlee@whu.edu.cn

* Correspondence: sunkm@whu.edu.cn; Tel.: +86-27-6877-8512

Abstract: The accurate location of clouds in images is prerequisite for many high-resolution satellite imagery applications such as atmospheric correction, land cover classifications, and target recognition. Thus, we propose a novel approach for cloud detection using machine learning and multi-feature fusion based on a comparative analysis of typical spectral, textural, and other feature differences between clouds and backgrounds. To validate this method, we tested it on 102 Gao Fen-1(GF-1) and Gao Fen-2(GF-2) satellite images. The overall accuracy of our multi-feature fusion method for cloud detection was more than 91.45%, and the Kappa coefficient for all the tested images was greater than 80%. The producer and user accuracy were also higher at 93.67% and 95.67%, respectively; both of these values were higher than the values for the other tested feature fusion methods. Our results show that this novel multi-feature approach yields better accuracy than other feature fusion methods. In post-processing, we applied an object-oriented method to remove the influence of highly reflective ground objects and further improved the accuracy. Compared to traditional methods, our new method for cloud detection is accurate, exhibits good scalability, and produces consistent results when mapping clouds of different types and sizes over various land surfaces that contain natural vegetation, agriculture land, built-up areas, and water bodies.

Keywords: multi-feature fusion; cloud detection; high-resolution; machine learning

1. Introduction

Clouds often obscure on the ground target objects and thus have significant impact on remote sensing image interpretation [1,2]. Further, the accurate detection of clouds directly affects inversion of biophysical variables in other applications [3]. Even though a particular image might have excessive cloud coverage, the image can still be useful as long as specific areas of interest are not obscured. Therefore, the presence or absence of clouds in a particular image needs to be accurately determined.

Cloud detection is an active area of research in the field of remote sensing with many proposed cloud detection methods. These methods can be divided into two categories: those using and not using thermal infrared bands. Methods that use thermal infrared bands often provide highly accurate cloud detection results. Zhu and Woodcock proposed an object-oriented Function of Mask (Fmask) cloud and shadow detection method for Landsat images using the Top of Atmosphere (TOA) reflectance and brightness temperature. This method achieves a detection accuracy reaching 96.4% [4]. Huang et al. used clear view forest pixels as a reference to define cloud boundaries when separating clouds from clear view surfaces in a spectral-temperature space for forest change analysis [5]. Most high-resolution satellite images, however, lack a thermal infrared band; so researchers often use temporal and spatial

correlation information from images in cloud detection algorithms. Sedano et al. [6] proposed a method that first use pixel value comparison between high resolution images and cloud free composites at lower spatial resolution from almost simultaneously acquired dates to identify cloud pixels and then obtain cloud mask using object-based regional growth. The advantage of this method is that it does not rely on thermal bands and can be used to extract clouds in the highest-resolution remote sensing images, but the disadvantages is that it relies on lower spatial resolution from almost simultaneously acquired dates and if there are a large cloud in image, the method will fail. Tseng et al. [7] applied multi-temporal cloud detection algorithms for cloud detection. This method have good accuracy, but the biggest disadvantage is that acquired dates of the reference image and the observed image must be very close. Han et al. proposed an automatic cloud detection method for high-resolution images that first detects clouds in a target image including cloud cover, by applying a simple threshold. Using the difference between the target image and the reference image, clouds can be extracted from the peripheral region of high-resolution images [8]. This method also needs to use temporal correlation among pixels.

On the other hand, methods that do not use thermal bands and rely on the attribute information from single images have been proposed. Maximum and minimum filters and second-moment texture measures have been tested to detect clouds and cloud shadows based on an analysis of different spectral and texture features of clouds, thin clouds, vegetation, bare ground, and roads [9]. Marais et al. [10] have presented a simple image transform that optimally combines four image channels into a greyscale image for threshold-based cloud detection on Quickbird images. Fisher used reflectance feature and automated morphological feature extraction to detect cloud in SPOT5 HRG Imagery [11]. Wavelet analysis, pattern recognition, machine learning, and computer vision technologies have also been used for cloud detection [12–16]. These methods use spectral information to extract cloud areas; however, these methods only apply in situations when spectral features of other objects such as buildings are not similar to clouds. In high-resolution images, the spectral features of city buildings are similar to those of clouds [17]. Therefore, additional features found in clouds and buildings are needed to discriminate them accurately. Clouds have high reflectance with relatively uniform, smooth texture characteristics and have irregular natural shapes, therefore integration of different features might provide a better solution for differentiating between clouds and high reflective objects in high-resolution images.

Several classification methods have been proposed that are based on the integration of spatial and spectral features; these can be divided into two types: multi-classifier algorithms [18], and single classifier multi-feature algorithms [19,20]. The multi-classifier methods use a variety of classifiers for decision fusion and processing multi-features in parallel. The single classifier multi-feature approach normalizes the different features and then uses a single classifier on several related features to process vectors of hybrid features. We therefore selected a single classifier multi-feature approach as the basis for our new cloud detection method; however, one key issue facing single classifier multi-feature algorithms is the choice of classification method. Machine learning as an artificial intelligence algorithm has been widely used in remote sensing image classification [21]. The application of artificial neural networks, support vector machine, intelligent optimization algorithms such as genetic algorithms, partial swarm optimization, and artificial immune systems in remote sensing image classification reflect this trend [22–27].

Some researchers have studied cloud features and machine learning for cloud detection. Wang et al. [28] compared Support Vector Machine (SVM) with Back Propagation Neural Network (BP-NN) method with training set numbers for cloud detection, finding that when sample size is small, SVMs perform better than BP-NN method. Hughes et al. [29] developed a neural network approach to detect clouds in Landsat scenes using spectral and spatial information. Chenthan [30] has compared Linear, Polynomial, and Gaussian-RBF kernel in SVMs and concluded the combination of texture and Linear SVM with correct classification rate of 92.30% in cloud detection. Başeski [31] used color and line based elimination as well as texture based SVM classification to detect cloud on RGB images. Latry considered SVM cloud detection within PLEIADES-HR framework [32]. Although many researchers

have studied cloud detection classification methods based on features, many of these approaches are based on spectral and texture information, and do not exploit other features, such as shape and NDVI or combine multi-features in machine learning. The integration of multi-feature and machine learning strategies to detect clouds could not only allow features to complement each other, but also make full use of the remote sensing image information.

In this paper, we propose a new cloud detection approach based on machine learning and multi-feature fusion. It does not use the thermal band or rely on time information and is suitable for different clouds types, sizes, and densities, and different underlying surface environments. This paper is organized as follows. The first section describes the experimental data, discusses typical cloud features in Gao Fen-1/2(GF-1/2) satellite images, details the selection of the typical characteristics of various images, and discusses the construction of feature space and multi-feature fusion. The second section details our results. The third section includes a discussion and the last section includes conclusions drawn from our research.

## 2. Materials and Methods

2.1. Experimental Data

A total number of 102 GF-1/2 images were acquired from National Disaster Reduction Center of China to evaluate our proposed method. The testing 102 images were acquired between March 2014 and May 2016. The training data of each scene are randomly selected from within the scene. Each testing image is an inset of the scene with a size of 2000 × 2000 pixels in order to speed up the detection process since not the entire image is covered by clouds. The testing insets are carefully selected, covering various underlying surface environment. Eight representative images are shown in Table 1, as it is not possible to show all of the images. Only the blue (0.45–0.52 μm), green (0.52–0.59 μm), red (0.63–0.69 μm), and near infrared (0.77–0.89 μm) bands of the GF images were used in our experiments, panchromatic bands were excluded. GF-1/2 image resolution is eight meters and four meters, respectively, while the correction level of the tested images was Level 1A. All images have not been resampled. The majority of the images in the datasets contain both cloud and non-cloud regions. Cloud regions include small, medium, and large size clouds; the backgrounds are common underlying surface environment including mountains, buildings, roads, agriculture, and rivers. In addition, we do not consider impact of season on the clouds.

Table 1. Experimental data sources and descriptive information for eight selected images.

Abbreviation/Image Information	High Resolution Image Date	Solar Zenith Angle (Degree)	Cloud Types	Scene Description
GF-1-145 L1A0000232145	22 May 2014	73.6156	Medium, low density clouds	Agriculture, high reflectance buildings and road
GF-1-270 L1A0000268270	6 July 2014	74.5897	Medium, low density clouds	vegetation, high reflectance buildings and road
GF-1-419 L1A0000232419	22 May 2014	74.1973	Medium, high density clouds	Mountainous area with high reflectance bare soil slope
GF-1-963 L1A0000264963	2 July 2014	75.0930	Small, high density clouds	Mountainous area with high reflectance bare soil slope
GF-2-865 L1A0001028865	21 May 2015	17.6541	Large clouds	Urban area with high reflectance buildings and road
GF-2-193 L1A0001451193	5 March 2016	36.4954	Large clouds	Mountainous area with high reflectance bare soil slope
GF-2-042 L1A0000836042	31 May 2015	16.4647	Medium, low density clouds	Agriculture, high reflectance buildings and road
GF-2-975 L1A0000835975	31 May 2015	16.4336	Medium, high density clouds	Agriculture, high reflectance buildings and road

The GF-1/2 level-1A product only contains DN information, and does not directly contain reflectivity data. For spectral feature extraction, we needed surface reflectance values. The radiometric calibration is a fundamental process to eliminate radiometric errors and distortion [33,34]. Atmospheric correction eliminates the effects of the atmosphere on the reflectance values of images taken by satellite or airborne sensors [35]. Radiometric calibration and atmospheric correction parameters were obtained from the attached XML file along with GF data. Radiometric calibration was carried out using ENVI software and we used the Fast Line-of-sigh Atmosphere Analysis of Spectral Hypercubes (FLAASH) [36] for atmospheric correction.

Our algorithms were applied to distinguish clouds in different situations with different clouds types, cloud sizes, densities, and different underlying surface environments, except for very thin clouds, which are shown in Figure 1. In order to verify the accuracy of the cloud detection, we obtained the manual cloud mask of all tested images from National Disaster Reduction Center of China. Very thin clouds in the manual cloud mask are transparent and therefore not treated as cloud objects; we consider the ground targets under the thin cloud cover as useful in applications.

(**a**)　　　　　　(**b**)
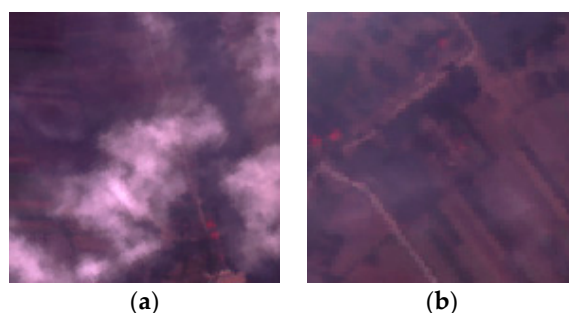
Figure 1. Clouds on Gao Fen-1(GF-1) images: (**a**) thin clouds and (**b**) very thin clouds.

2.2. Typical Cloud Features

Cloud types vary widely, but despite differences, their inherent physical properties are essentially the same; and, specifically, the overall radiation characteristics of cloud particles are similar [37]. In the visible band (0.4–0.8 μm), the optical properties of water droplets and ice particles not only make clouds very reflective, but also increase light absorption thus reducing the transmittance of the radiation information from the surface of materials [38]. To reflect typical spectral characteristics of the clouds and other backgrounds in GF-1/2 satellite images, we selected sample areas from 102 GF-1/2 images randomly to calculate the mean reflectance spectra of thick clouds, thin clouds, vegetation, highly reflective buildings, general buildings, and water covered areas. Figure 1 presents the mean reflectance spectra of those targets commonly found in GF-1/2 satellite images.

As shown in Figure 2, cloud reflectance in the visible band is generally higher than that of most natural surface features, but is similar to buildings with the highest reflectance. Therefore, spectral characteristics by themselves are insufficient for discriminating clouds from highly reflective buildings.

In contrast to low spatial resolution satellite images, high spatial resolution remote sensing images show richer texture information, clearer geometry shape attributes, and a more distinct spatial distribution of gray features. Figure 2 illustrates textures of the cloud, water, building, and mountain textures as found in our dataset.

In Figure 3, the gray value changes very little between adjacent pixels of clouds and water, but the textures of buildings and mountains are very different. Thus, cloud texture is uniform and smooth, with low contrast in these high-resolution images, making clouds easily distinguishable from buildings and mountains. At the same time, however, clouds can be easily confused with water.

In geometry shape attributes, cloud shapes are more natural and irregular; road shapes are relatively narrow and linear; and building shapes have relatively, regular shape outlines. Compared to low-resolution satellite images, shape features of high-resolution satellite images are highly reflective.

In other attributes such as the Normalized Difference Vegetation Index (NDVI), pixels from vegetated areas yield high positive value for NDVI due to high near-infrared reflectance and low red reflectance. Bare soil and rocks generally show similar reflectance in the near-infrared band and visible bands, generating positive but lower NDVI value close to 0. The visible bands reflectance of water, clouds, and snow are larger than their near-infrared reflectance, producing negative NDVI value.
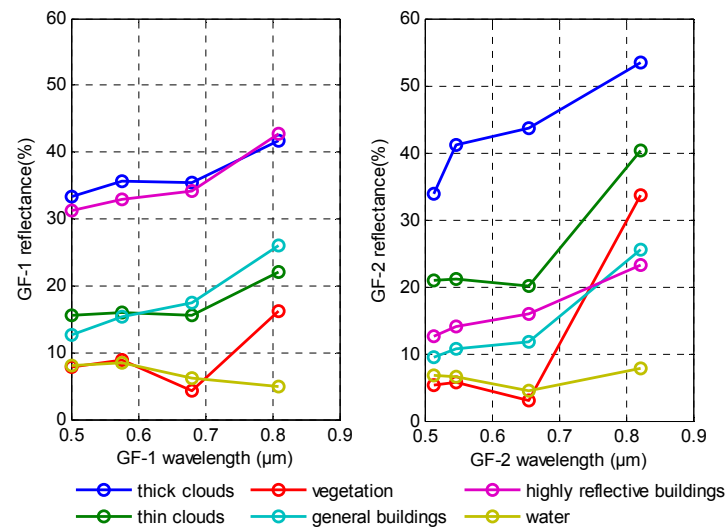
Figure 2. The mean reflectance spectra of thick clouds, thin clouds, vegetation, general buildings, highly reflective buildings, and water covered areas commonly found in GF-1/2 satellite images.
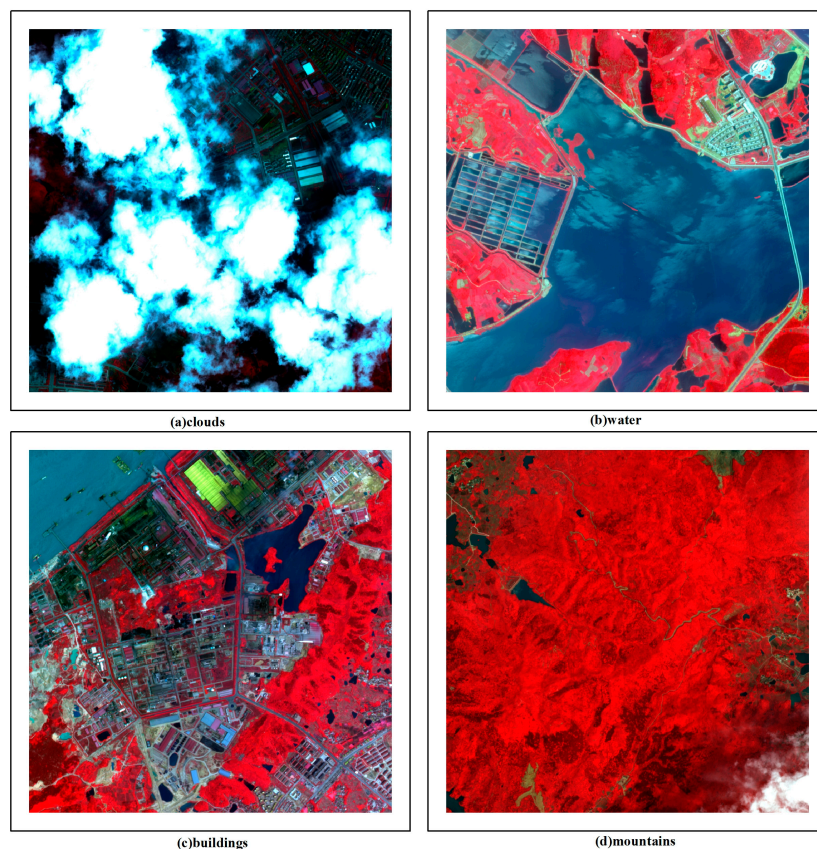
Figure 3. Image textures for: (**a**) clouds; (**b**) water; (**c**) buildings and (**d**) mountains. Near infrared red, red and green bands are chosen for making these false color images.

2.3. Feature Selection

According to analysis of the typical cloud features, we cannot use a single feature to separate clouds from backgrounds, we must select features from the spectral, texture, and other features simultaneously.

2.3.1. Selection of Spectral Features

In multispectral images, clouds and background exhibit different spectral characteristics [39]. Clouds generally scatter various wavelengths evenly, so clouds have high reflectivity in the visible and near infrared bands, but the cloud spectral reflectance decreases slowly with increasing wavelength. Accordingly, the reflectance values of blue, green, red, and near infrared bands are used to obtain spectral features of clouds.

2.3.2. Selection of Texture Features

The use of a Grey Level Co-occurrence Matrix is a standard texture extraction method [40], and has demonstrated validity in dealing with high-resolution images [41,42], hence we use Grey Level Co-occurrence Matrix (GLCM) to extract texture features. We need consider three key issues including band selection, window size, and direction in GLCM method.

1. Band selection

We obtain the first and second principal component (PC) comprising 95% of the image information from the GF-1/2 images using principal component analysis, and then conduct texture analysis by utilizing PC1 and PC2.

2. Window size

Window size for texture analysis is related to image resolution and the contents within the image [43]. The Variation Coefficient (VC) was used to measure if window size of the texture measures was suitable or not [43,44]. If the VC is large, the variation amplitude is large, and unstable. Therefore, the chosen optimal window size is at the point where the value of the coefficient of variation starts to stabilize and has the smallest value. The VC is defined as follows [43]:

$$\text{VC} = \text{S}/\overline{SV} \tag{1}$$

$$\text{S} = \sqrt{\frac{\sum sv^2 - \frac{\sum sv^2}{n}}{n-1}} \tag{2}$$

$$\overline{SV} = \frac{1}{n}\sum_{1}^{n} sv_n \tag{3}$$

where S denotes standard deviation, $\overline{SV}$ is an average value, sv is the sample value, and n is the number of samples.

Window size of the texture measures is shown as follows:

$$W = \underset{w}{\text{argmin}}\{\text{VC}\} \tag{4}$$

3. Direction

Pesaresi et al. demonstrated that rotational invariance properties can be obtained when texture features in different directions are selected [41]. Therefore, the mean, variance, uniformity, contrast, correlation, and entropy values for 0°, 45°, 90°, and 135°, respectively, were calculated to derive the average value of the each parameter of four directions.

2.3.3. Selection of Other Features

NDVI can help researchers to distinguish between clouds and other surface features. Therefore, NDVI was chosen as a supplement to spectral and texture features during cloud detection.

NDVI values of all pixels are calculated after radiometric and atmospheric correction and serve as another feature.

2.4. Methods

In this paper, we used a single classifier method for multi-feature fusion, based on SVM [45]. SVMs are particularly appealing in the remote sensing field due to their ability to generalize even with limited training samples [46]. SVMs perform better than neural networks, particularly in the case of small samples [23]. SVMs are very effective tools in many classification methods incorporating machine learning and multiple feature fusion to separate clouds from backgrounds.

2.4.1. SVM Classification

The basic principle is as follows:

In the given trainings $T\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, where $x_i \in X = R^n$, X is called the input space. Each point x_i in the input space is composed of spectral characteristics, texture features and NDVI feature. When $y = 1$, the sample is a cloud; when $y = -1$, the sample is part of the background. The classification function $f(x)$ is used to infer the corresponding value of y to any pattern x.

For linearly separable cases, seeking the optimal hyper-plane on two class patterns involves finding the hyper-plane that can separate classes from each other with a maximum margin, solving $y_i(\omega^* * x + b^*) \geq 1$, when $\frac{1}{2}\| \omega \|^2$ is minimized. The optimal hyper-plane defined by the weight vector $\omega^* \epsilon R^n$ and the bias $b^* \epsilon R^n$ is the one that minimizes a cost function that expresses a combination of two criteria, namely: (1) margin maximization; and (2) error minimization.

The Lagrange function is introduced, solving the optimal solution for the equation: $a_i(y_i(\omega^* * x + b^*) - 1) = 0$, where a_i is a Lagrange multiplier of corresponding samples and not zero, and the corresponding sample is the support vector. When this meets the constraints of $\sum_{i=1}^{n} y_i a_i = 0$, solving for the maximum value of the following functions can be calculated as:

$$\omega(a) = \sum_{i=1}^{n} a_i - \frac{1}{2} \sum_{i,j=1}^{n} a_i a_j y_i y_j (x_i x_j) \tag{5}$$

where $a_i \geq 0$. Then, to find the optimal function:

$$f(x) = \text{sign}\{(\omega x) + b\} = sign\left\{\sum_{i=1}^{n} a_i^* y_i (x_i x) + b^*\right\} \tag{6}$$

where $f(x)$ represents the discriminant function associated with the hyper-plane.

Where summation is only for support vectors, b^* is the classification threshold; and can be obtained by any support vector. When the linear SVM classification is extended to the case of a non-linear surface, the kernel function is used. The principle idea is that: (1) the problem is transformed into a linear problem in high-dimensional space by a non-linear change; and, therefore; (2) the optimal hyper-plane in a high-dimensional space is found. First, the input space sample is mapped in high-dimensional space features by $\psi : H : \psi : R^d \rightarrow H$ thus avoiding a direct mapping of the samples to the high-dimensional space, by constructing a kernel function. In the SVM, the choice of kernel is critical, we use the radial basis function (RBF) kernel based on its performance characteristics in several applications [47,48].

The RBF kernel is defined as [49]: $k(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2}$.

2.4.2. Feature Fusion

Feature fusion is a popular term that has been used in a variety of remote sensing projects. Texture features have been fused in remote sensing classification applications [50,51] and image segmentation [52]. Classification performance can benefit from proper feature fusion [53]. Features fusion involves normalizing a variety of features and then merging the normalized features into a single feature set; this allows the user to focus on feature selection from this feature set to obtain the best feature subset. The basic principle of feature fusion is to select the features that are related to categories and to exclude redundant features. Here, we focus on a methods used to fuse features using different typical features of clouds.

Feature fusion methods can be broadly divided into two categories: multiple sets of feature vectors are created from head to tail to generate a new feature vector, and we use the new vector to classify the target features; and, multiple sets of features are combined by complex vectors, and we use the complex vector space to classify the target features. The first method is used to classify objects. Spectral, textural and NDVI features are selected as multi-type features. The multi-type features are normalized to [0, 1] and thus could be input into SVMs to perform classification. For comparative purposes, other experiments use different features combinations. The accuracy assessment can be seen below.

$\alpha_i = \{\alpha_{i1}, \alpha_{i2}, \alpha_{i3}, \alpha_{i4}\}$ are spectra features and $\beta_i = \{\beta_{i1}, \beta_{i2}, \beta_{i3}, \beta_{i4}, \beta_{i5}, \beta_{i6}, \beta_{i7}, \beta_{i8}, \beta_{i9}, \beta_{i10}, \beta_{i11}, \beta_{i12}\}$ are texture features. $\varepsilon_i = \{\varepsilon_{i1}\}$ is NDVI feature. We use the normalization equation to each feature first. Then, Features are combined from head to tail $t = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}, x_{13}, x_{14}, x_{15}, x_{16}, x_{17}\}$. We construct feature space based on SVM.

The normalization equation is shown as below:

$$x = \frac{(rf - rf_{\min})}{rf_{\max} - rf_{\min}} \tag{7}$$

where rf represents raw feature data, and x denotes feature data after normalizing.

2.4.3. SVM-Based Multi-Feature Fusion

SVM-based multi-feature fusion steps and diagram (Figure 4) are as follows:

Step 1: Based on our analysis of typical cloud characteristics, spectral, textural and other features are extracted, and then normalized to construct a feature space.

Table 2 shows the feature selection and feature space construction:

**Table 2.** Feature selection indicators.

Features	Feature Space
Spectral features	Band 1, Band 2, Band 3, Band 4
Texture features	Mean, Variance, Homogeneity, Contrast, Correlation, Entropy Window Size $= \underset{w}{\mathrm{argmin}}\{VC\}$ Band selection: PC1 and PC2
Other features	NDVI

Step 2: Extract training samples. Samples are extracted randomly from cloud and non-cloud regions.

Step 3: Spectral, textural and NDVI features are extracted from both cloud and non-cloud regions. The spectral, textural and NDVI features of a sample are referred to as a characteristic description for image pixel content information.

Step 4: A Support Vector Machine-Radial Basis Function (SVM-RBF) classification model is constructed to train samples of cloud and non-cloud regions to obtain a group of non-parallel hyperplanes parameters.

Step 5: Spectral, textural and NDVI feature for an obtained high-resolution image are then extracted. The extracted dimensional feature vector is input into the SVM-RBF classifier for classification, and cloud detection results of all the pixels are obtained.
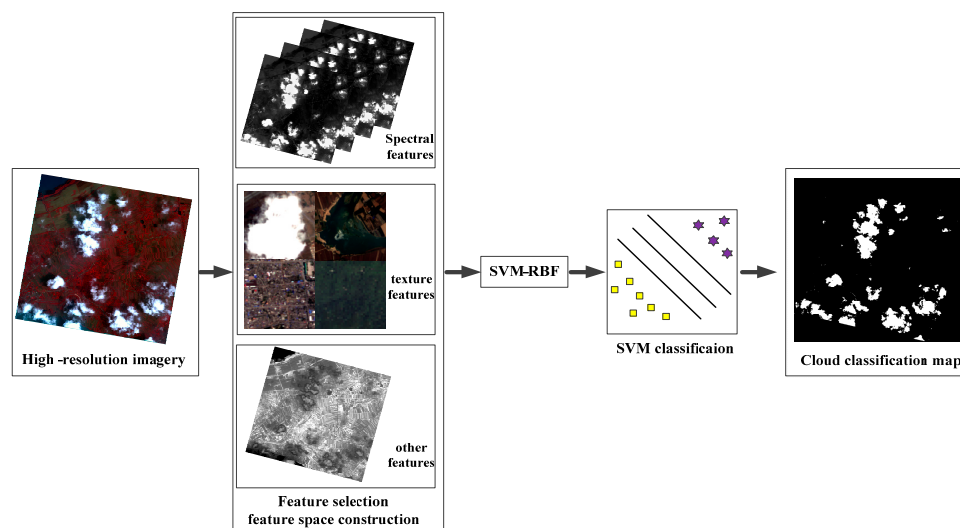
Figure 4. Diagram showing the method used to combine multiple features.

3. Results

3.1. Cloud Detection Experiment

Cloud detection experiment results based on SVM and feature fusion are as follows.

The training samples in cloud detection experiment for eight selected images include both cloud and non-cloud regions, as shown in Table 3.

Table 3. Training samples for eight selected images.

| Sensor	Types	Training Samples
GF-1-145	Cloud	2207
	Non-cloud	5444
GF-1-270	Cloud	2043
	Non-cloud	5722
GF-1-419	Cloud	2714
	Non-cloud	5671
GF-1-963	Cloud	2565
	Non-cloud	5672
GF-2-865	Cloud	2041
	Non-cloud	5365
GF-2-193	Cloud	2423
	Non-cloud	5325
GF-2-042	Cloud	2356
	Non-cloud	5671
GF-2-975	Cloud	2456
	Non-cloud	5432

From Table 3, we see that in the training sample there were about 2000–3000 cloud pixels, while non-cloud regions including buildings, vegetation, water and roads numbered about 5000–6000 pixels. Our cloud detection model was constructed to train samples of cloud and non-cloud regions. Origin

images, classification results and insets marked by red frame region for eight selected GF-1/2 images are shown in Figures 5 and 6, respectively. The white regions in the classification results represent our cloud detection results. The green frame refers to cloud detection misjudgments. To make the extracted feature correlation index more accurate, radiometric calibration and atmospheric corrections were carried out prior to cloud detection with the high-resolution imagery data set.

As can be seen in Figures 5 and 6, the overall cloud detection results for the tested GF-1/2 images were visually accurate for different clouds types, cloud size, density, and different underlying surface environments, suggesting that the method has good scalability and adaptation. As can be seen from local classification maps, there are still some buildings and roads often are mistaken for clouds. Cloud detection has better results on mountainous area than high reflective urban area. High reflectance buildings and road, cloud edges and haze have the potential to reduce the algorithm performance.
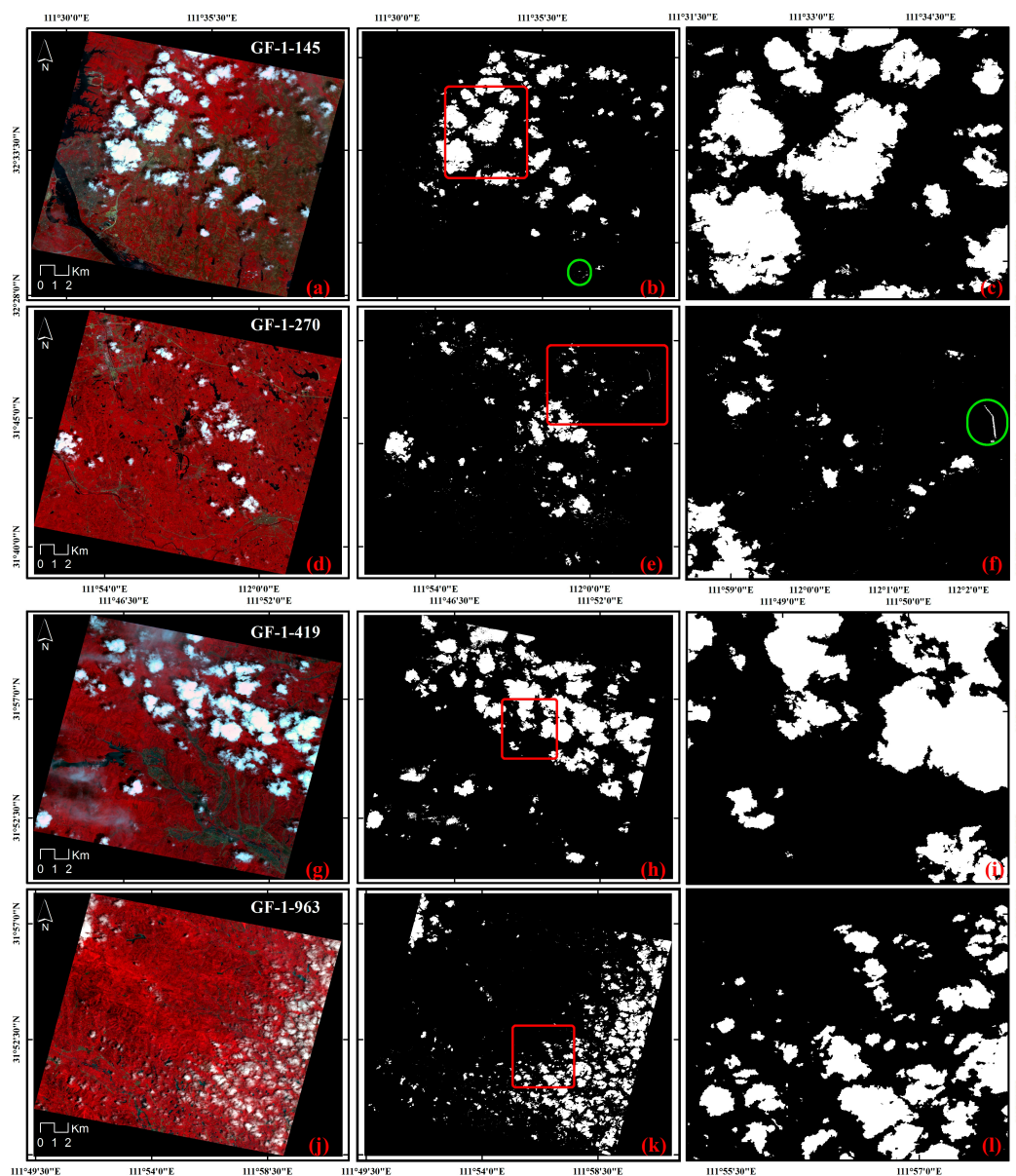
Figure 5. The original map and classification results of four selected GF-1 images: (**a**–**c**) the original map, cloud classification map, and the magnified area corresponds to the red line regions of cloud classification map for GF-1-145, respectively; (**d**–**f**) for GF-1-270; (**g**–**i**) for GF-1-419; (**j**–**l**) for GF-1-963. The white regions in the classification results represent our cloud detection results. The green frame refers to cloud detection misjudgments.
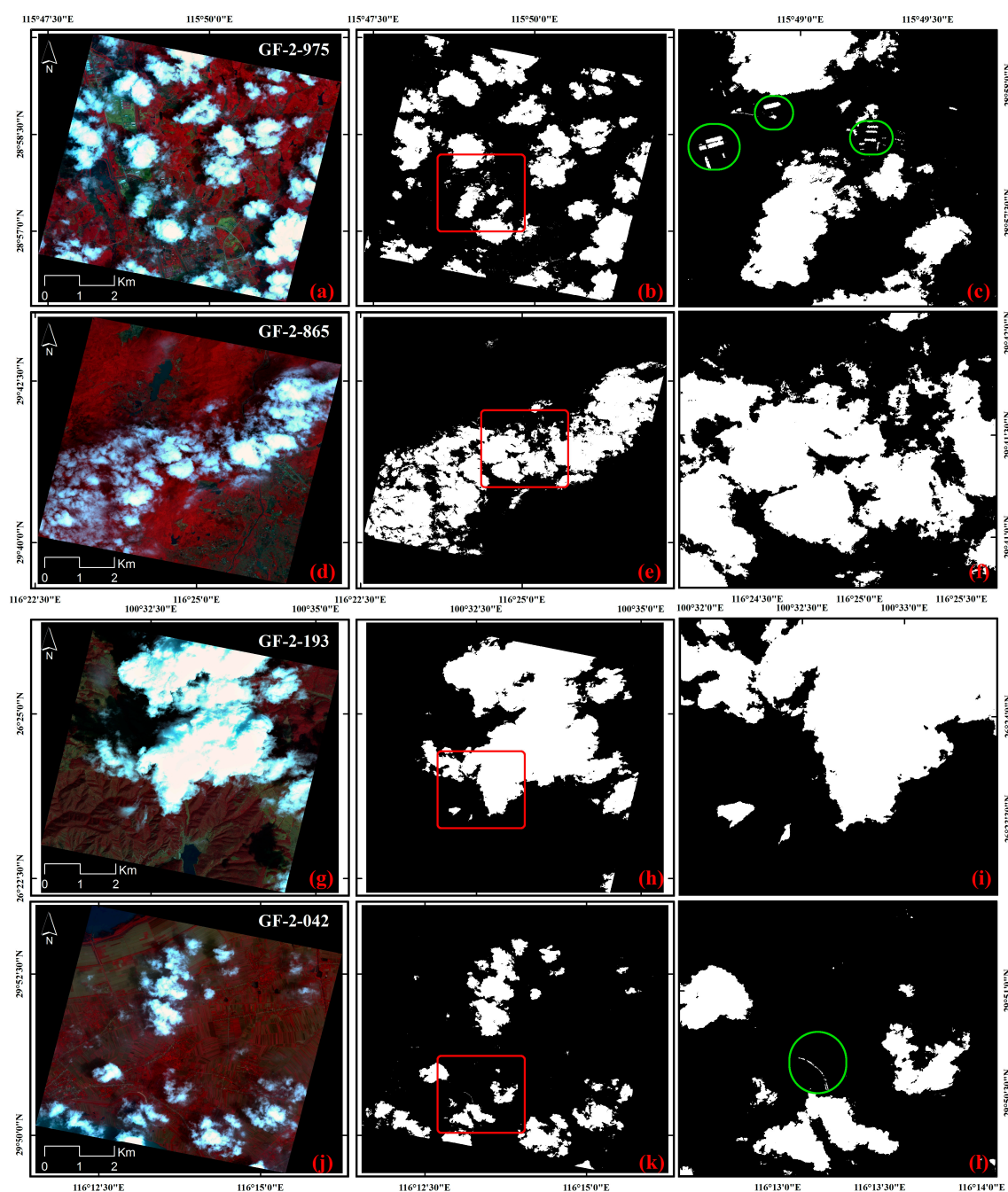
Figure 6. The origin map and classification results of four selected GF-2 images: (**a**–**c**) the original map, cloud classification map, and the magnified area corresponds to the red line regions of cloud classification map for GF-2-975, respectively; (**d**–**f**) for GF-2-865; (**g**–**i**) for GF-2-193; (**j**–**l**) for GF-2-042. The white regions in the classification results represent our cloud detection results. The green frame refers to cloud detection misjudgments.

3.2. Object-Oriented Post-Processing

After the rough detection of clouds, buildings and roads are often mistaken for clouds; morphological operators and shape features can be used to address this situation.

Step 1: Because clouds are natural features with irregular shapes when compared with the regular shapes of roads and buildings, and because roads are narrow, we chose rectangles and the length–width ratio shape index for classification.

Step 2: Handle false positives through erosion and dilation operators.

A rectangle [54] is classified as follows:

$$R = A_0 / A \tag{8}$$

where A_0 denotes a regional area, and A is the minimum bounding rectangle area for this region.

When R = 1 or R = $\pi/4$, the target area is very probably rectangular or circular, respectively.

The length–width ratio [54] is classified as follows:

$$T = a/b \tag{9}$$

where *a* and *b* are the length and width for the minimum bounding of a rectangle of this region respectively.

When T = 1, this area is square or circular; and, when T > 1, the area is a slender shape.

Five GF-1/2 images including buildings and roads are selected to do object-oriented post-processing. Figure 7 shows rectangle and length–width ratio histogram of rough cloud detection for five GF-1/2 images.
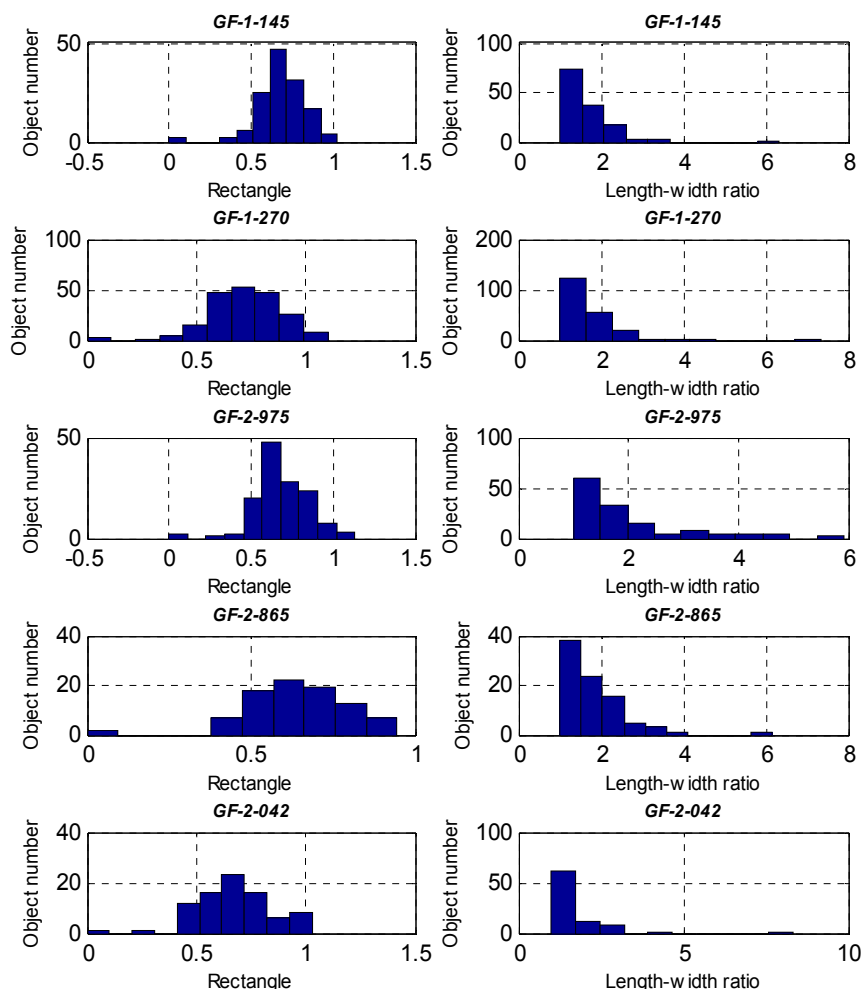
Figure 7. Rectangle and length–width ratio histogram of rough cloud detection for five GF-1/2 images.

In Figure 7, the rectangular values of most objects in the GF-1/2 images were concentrated in 0–0.8, and length–width ratio focused on 1–3.5. High-frequency targets are clouded in the images. Therefore, low frequency objects whose rectangle values are not in the range of 0–0.8 and where the

length–width ratio values are not in the range of 1–3.5, were removed. Then the map was examined to detect erosion and dilation and to remove slight misjudgments so that the cloud detection results could be finalized. Object-oriented post-processing results and accuracy for five GF-1/2 images are shown in Figure 7 and Table 4, respectively.

Figure 8 illustrates that for five GF-1/2 images, Object-oriented post-processing removed buildings and roads effectively using shape features.

Shown in Table 4, Object-oriented post-processing further improved the accuracy of our method. The average accuracy reached 96.5%, which demonstrates the effectiveness of object-oriented post-processing.

Table 4. Rough cloud detection accuracy and object-oriented post-processing accuracy for five GF-1/2 images.

Sensor	Overall Accuracy	Kappa Coefficient	Post Processing Accuracy
GF-1-145	96.6791%	0.8992	97.5263%
GF-1-270	95.4259%	0.8936	97.6589%
GF-2-975	95.2362%	0.8363	96.6701%
GF-2-865	96.6252%	0.9109	96.7713%
GF-2-042	96.4454%	0.9179	97.6235%
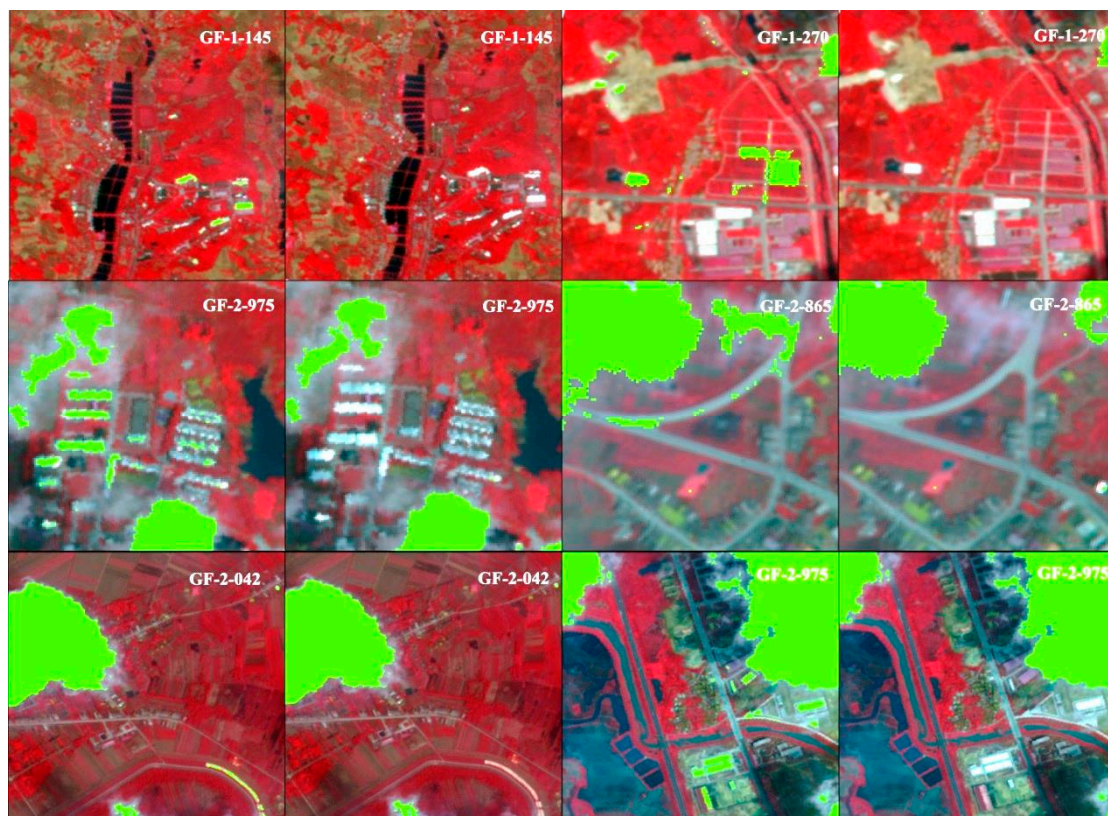
Figure 8. Rough cloud detection and object-oriented post-processing results for five GF-1/2 images (the fluorescent green regions in the picture represent our cloud test results).

4. Discussion

4.1. Accuracy Analysis of Multiple Features

To validate multi-feature fusion cloud extraction, different features were fused including Spectra (S), Texture (T), Spectra + Texture (ST), Spectra + NDVI (SN), Texture + NDVI (TN),

and Spectra + Texture + NDVI (STN) of 102 GF-1/2 images for cloud accuracy assessment and analyzed. Considering cloud and non-cloud areas as two classes, the cloud producer accuracy and cloud user accuracy, the overall accuracy and Kappa coefficients were calculated as below [55]:

$$CPA = \frac{GA}{GA + OC} \tag{10}$$

$$CUA = \frac{GA}{GA + CC} \tag{11}$$

where *CPA* denotes cloud producer accuracy; *CUA* denotes cloud user accuracy; *GA* denotes agreement between ground truth and the algorithm mask; *OC* represents pixels that belong to clouds in the ground truth but the classification technique has failed to classify them into clouds; and *CC* represent pixels that belong to backgrounds that are labeled as belonging to clouds.

Table 5 shows the quantity of images of the six features combination grouped by producer accuracy, 76%–80%, 80%–84%, 84%–88%, 88%–92%, 92%–96%, and 96%–100%, and the average producer accuracy of each feature combination. Figure 9 is a corresponding histogram.

Table 5. Cloud producer accuracy for Spectra (S), Texture (T), Spectra + Texture (ST), Spectra + NDVI (SN), Texture + NDVI (TN), and Spectra + Texture + NDVI (STN) feature fusion on 102 GF-1/2 images.

| Cloud User Accuracy	76%–80%	80%–84%	84%–88%	88%–92%	92%–96%	96%–100%	The Average Accuracy
S	40	45	11	6	0	0	83.33%
T	65	30	7	0	0	0	81.76%
ST	23	34	18	10	15	2	84.17%
SN	43	28	20	6	5	0	86.67%
TN	49	24	4	16	9	0	84.55%
STN	2	5	7	22	41	25	93.67%
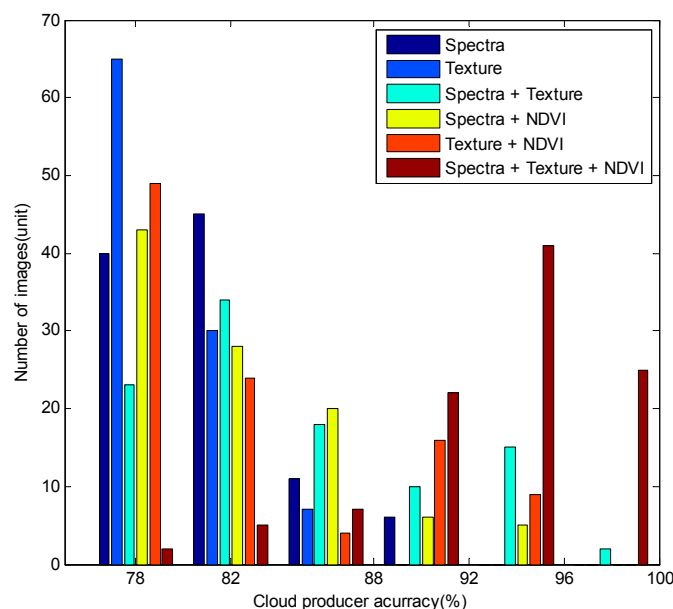
Figure 9. Bar of cloud producer accuracy for S, T, ST, SN, TN, and STN feature fusion on 102 GF-1/2 images.

Table 6 shows the quantity of images of the six features combination grouped by user accuracy, 86%–88%, 88%–90%, 90%–92%, 92%–94%, 94%–96%, and 96%–98%, and the average user accuracy of each feature combination. Figure 10 is the corresponding histogram.

Table 6. Cloud user accuracy for S, T, ST, SN, TN, STN feature fusion on 102 GF-1/2 images.

Cloud User Accuracy	86%–88%	88%–90%	90%–92%	92%–94%	94%–96%	96%–98%	The Average Accuracy
S	29	45	16	11	1	0	90.16%
T	57	28	15	2	0	0	89.25%
ST	26	29	18	12	15	2	91.35%
SN	27	35	25	9	6	0	90.67%
TN	45	23	9	15	10	0	90.47%
STN	0	4	8	22	35	33	95.67%
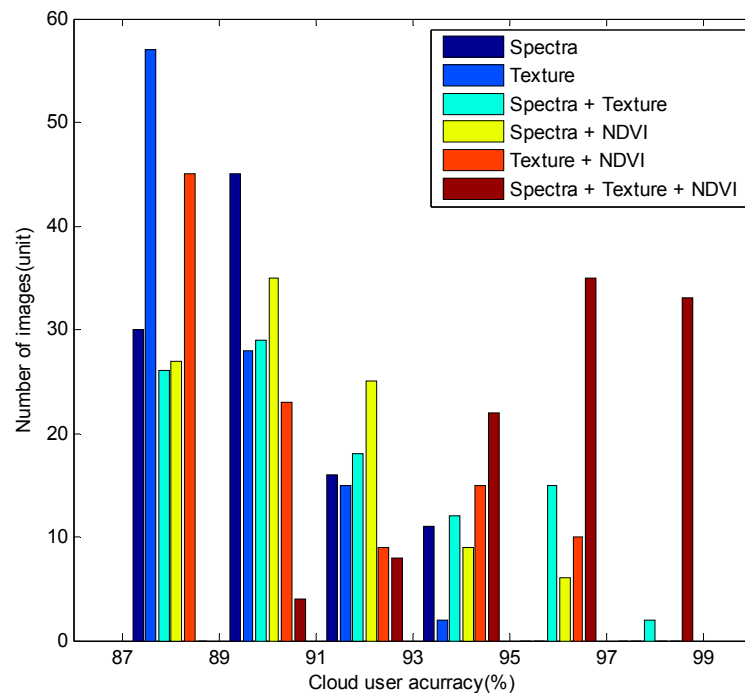
Figure 10. Bar of cloud user accuracy for S, T, ST, SN, TN, and STN feature fusion on 102 GF-1/2 images.

In order to compare the different feature fusion results, we show the partial enlarged view of total class error results of cloud detection for S, T, ST, SN, TN, and STN feature fusion on GF-1-145 images in Figure 11. Red pixel regions in the picture represent total classification error results in cloud detection.

From Figure 11, we see that total classification error results in cloud detection for STN was less than in S, T, ST, SN, TN feature fusion. The main reason is that STN performs better on houses, roads, and cloud edges in cloud detection than other feature fusion methods. These results suggest that multi-feature fusion is effective in cloud detection and allows features to complement each other.

The overall accuracy of our approach reached more than 91.45%; the Kappa coefficient of all images was more than 80%, which demonstrates the effectiveness of our method.

From Tables 5 and 6, and Figure 9, we see STN feature combinations delivered greater accuracy than other methods, the cloud producer accuracy reached 93.67% and cloud user producer accuracy reached 95.67%, indicating that feature combinations are effective in cloud detection. Our results show that multi-feature classification has higher accuracy than a single feature classification. Feature combinations with spectral features have a higher accuracy than feature combinations without spectral features. Our results demonstrate that a single feature cannot represent the image content. Using only a single feature to classify all categories does not make full use of all the discriminative information; hence, classification accuracy is low. Multi-features fusion can take advantage of every feature, and therefore delivers accurate results. Thus, we conclude that spectral feature is important for cloud detection.
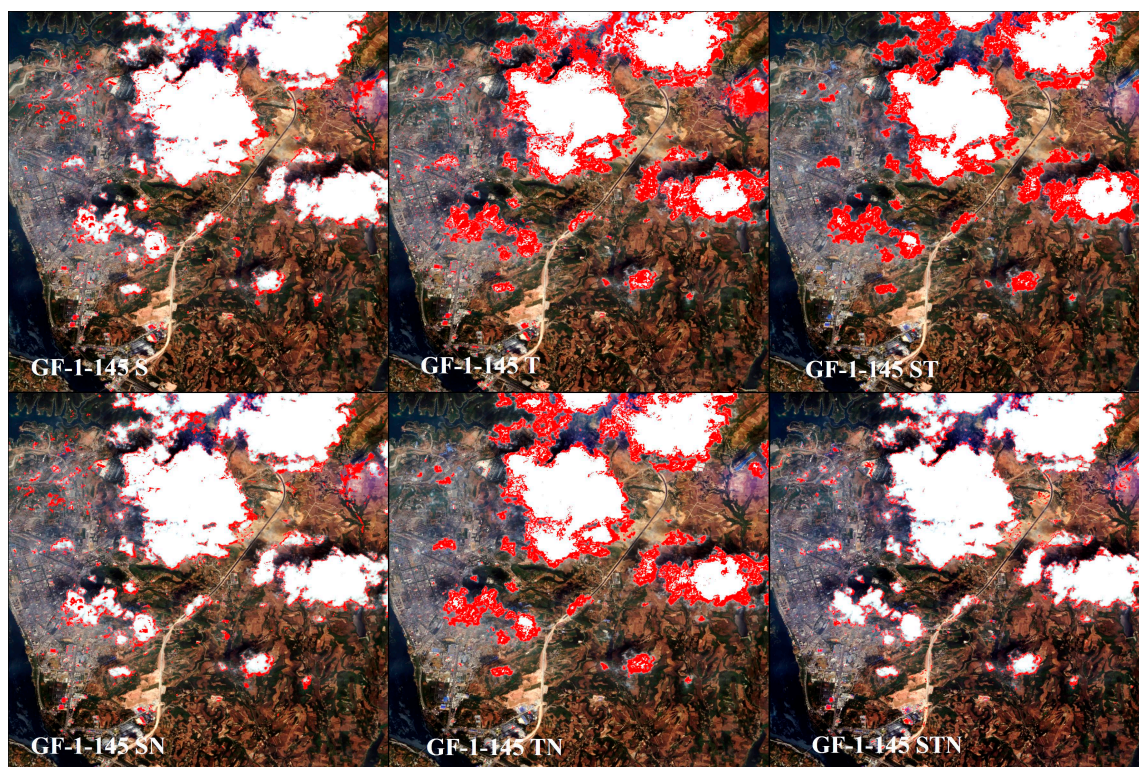
Figure 11. The partial enlarged view of total classification error results in cloud detection for Spectra (S), Texture (T), Spectra + Texture (ST), Spectra + NDVI (SN), Texture + NDVI (TN), and Spectra + Texture + NDVI (STN) feature fusion on GF-1-145 images (red regions in the picture represent total classification error results in cloud detection).

4.2. The Scalability of the Algorithm

Because our method carry out spectral attributes, texture attributes and NDVI feature fusion based on machine learning to achieve cloud detection, our method is suitable for high-resolution image data include near-infrared bands. To demonstrate the scalability of the algorithm, the algorithm is applied to two OrbView-3 images and one IKONOS-2 image, and the cloud classification maps are shown in Figure 12. Figure 12a–c the original map, cloud classification map, and the magnified area corresponds to the red line regions of cloud classification map for one OrbView-3 image, respectively; Figure 12d–f for other OrbView-3 image; and Figure 12g–i for IKONOS-2 image;

As can be seen from the figure, our method has good cloud detection results on OrbView-3 and IKONOS-2 images, which indicates that our method is not only suitable for GF-1/2 satellite imagery, but also suitable for other high satellite image data with near-infrared bands.
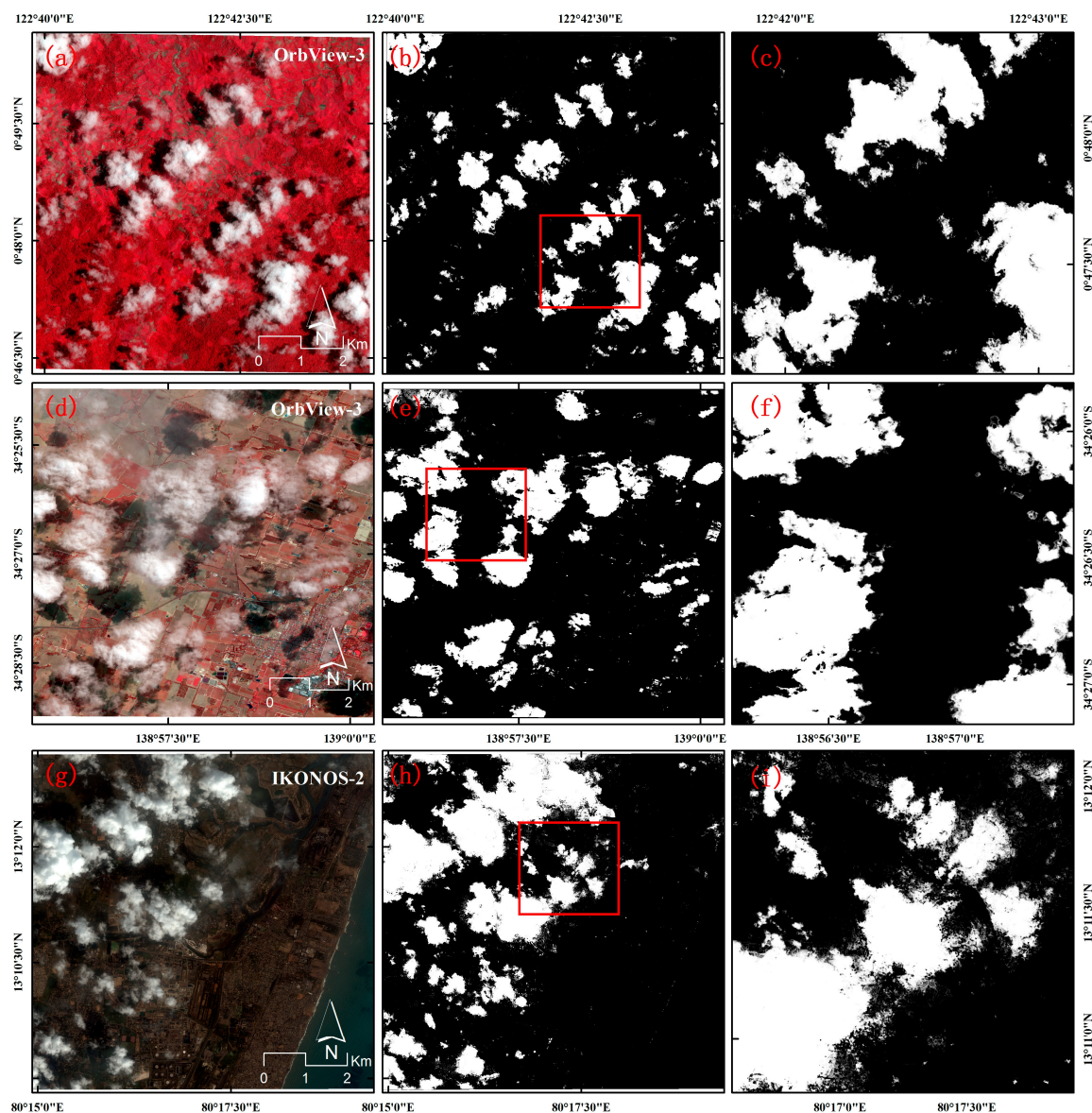
Figure 12. The origin map and classification results of two OrbView-3 images and one IKONOS-2 image: (**a–c**) the original map, cloud classification map, and the magnified area corresponds to the red line regions of cloud classification map for one OrbView-3 image, respectively; (**d–f**) for other OrbView-3 image; and (**g–i**) for IKONOS-2 image.

4.3. Summary of Advantages and Disadvantages

The proposed method has substantial advantages over conventional methods. In this paper, machine learning and multi-feature fusion are used for cloud detection. This method makes full use of the multi-feature information and has good accuracy than a single feature. Moreover, our method does not rely on thermal bands and multi-temporal information. It is suitable for various high-resolution earth observation sensors with near-infrared bands. Our proposed method is applicable in a variety of scenarios including high reflectance buildings, roads, agriculture and mountains. Furthermore, our method has good cloud detection results on large, medium, and small clouds and also has good effect on high, medium, and low density clouds.

However, our method has some disadvantages. We focus on machine learning and multi-feature fusion, but do not consider distinguishing between cloud and other objects including bare soil, desert, snow, and ice. In addition, we do not consider impact of season on the clouds. In future study, we will

try to apply machine learning and feature fusion method to distinguish between cloud and special underlying surface environments including bare soil, desert, snow, and ice.

5. Conclusions

In this paper, we propose a novel approach for cloud detection using machine learning and multi-feature fusion based on a comparative analysis of typical spectral, textural, and other feature differences between clouds and backgrounds in GF-1/2 images. The core principles are to select typical features and feature fusion based on machine learning for cloud detection. Cloud detection results showed that the proposed algorithm performed well on different clouds types, sizes, and densities, and different underlying surface environments. Furthermore, we experimentally demonstrate that multi-feature fusion yields more accurate cloud detection than a single feature approach and the spectral features in cloud detection are particularly important. We applied object-oriented post-processing in the final step by using rectangles and a length–width ratio shape index further reducing misclassification of highly reflective buildings and roads. Our proposed method is applicable in a variety of scenarios and is reliable in cloud detection. It can be applied to images from various high-resolution earth observation sensors with near-infrared bands. A shortcoming of our proposed method is that we need to select the training samples manually. Our future research will address automatic training samples selection and feature learning approaches. In addition, our research is designed to meet the needs of disaster reduction project of China, which focused on drought and floods in southern China, experimental data were selected from southern China. Thus, currently we do not consider other underlying surface environments including bare soil, desert, snow, and ice. There are other regional and seasonal data that need to be addressed in follow-up study, and we will improve methods to suit different underlying surface environments.

Author Contributions: Deren Li and Kaimin Sun conceived and designed the framework of this research; Ting Bai performed the experiments; Yepei Chen analyzed the data; Wenzhuo Li contributed analysis tools; and Ting Bai wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Roy, D.P.; Ju, J.; Kline, K.; Scaramuzza, P.L.; Kovalskyy, V.; Hansen, M.; Loveland, T.R.; Vermote, E.; Zhang, C. Web-enabled landsat data (Weld): Landsat ETM+ composited mosaics of the conterminous United States. *Remote Sens. Environ.* **2010**, *114*, 35–49. [CrossRef]

2. Zhu, Z.; Wang, S.; Woodcock, C.E. Improvement and expansion of the fmask algorithm: Cloud, cloud shadow, and snow detection for landsats 4–7, 8, and sentinel 2 images. *Remote Sens. Environ.* **2015**, *159*, 269–277. [CrossRef]

3. Shi, C.; Xie, Z. Operational method of total precipitable water retrieved from satellite multi-channels' infrared data. *J. Infrared Millim. Waves* **2005**, *24*, 304–308.

4. Zhu, Z.; Woodcock, C.E. Object-based cloud and cloud shadow detection in landsat imagery. *Remote Sens. Environ.* **2012**, *118*, 83–94. [CrossRef]

5. Huang, C.; Thomas, N.; Goward, S.N.; Masek, J.G.; Zhu, Z.; Townshend, J.R.; Vogelmann, J.E. Automated masking of cloud and cloud shadow for forest change analysis using landsat images. *Int. J. Remote Sens.* **2010**, *31*, 5449–5464. [CrossRef]

6. Sedano, F.; Kempeneers, P.; Strobl, P.; Kucera, J.; Vogt, P.; Seebach, L.; San-Miguel-Ayanz, J. A cloud mask methodology for high resolution remote sensing data combining information from high and medium resolution optical sensors. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 588–596. [CrossRef]

7. Tseng, D.-C.; Tseng, H.-T.; Chien, C.-L. Automatic cloud removal from multi-temporal spot images. *Appl. Math. Comput.* **2008**, *205*, 584–600. [CrossRef]

8. Han, Y.; Kim, B.; Kim, Y.; Lee, W.H. Automatic cloud detection for high spatial resolution multi-temporal images. *Remote Sens. Lett.* **2014**, *5*, 601–608. [CrossRef]

9. Lu, D. Detection and substitution of clouds/hazes and their cast shadows on ikonos images. *Int. J. Remote Sens.* **2007**, *28*, 4027–4035. [CrossRef]

10. Marais, I.V.Z.; Du Preez, J.A.; Steyn, W.H. An optimal image transform for threshold-based cloud detection using heteroscedastic discriminant analysis. *Int. J. Remote Sens.* **2011**, *32*, 1713–1729. [CrossRef]

11. Fisher, A. Cloud and cloud-shadow detection in spot5 hrg imagery with automated morphological feature extraction. *Remote Sens.* **2014**, *6*, 776–800. [CrossRef]

12. Lakshmanan, V.; DeBrunner, V.; Rabin, R. Texture-based segmentation satellite weather imagery. In Proceedings of the 2000 International Conference on Image Processing, Vancouver, BC, Canada, 10–13 September 2000.

13. Walder, P.; MacLaren, I. Neural network based methods for cloud classification on AVHRR images. *Int. J. Remote Sens.* **2000**, *21*, 1693–1708. [CrossRef]

14. Tian, B.; Azimi-Sadjadi, M.R.; Vonder Haar, T.H.; Reinke, D. Temporal updating scheme for probabilistic neural network with application to satellite cloud classification. *IEEE Trans. Neural Netw.* **2000**, *11*, 903–920. [CrossRef] [PubMed]

15. Simpson, J.J.; Gobat, J.I. Improved cloud detection in goes scenes over the oceans. *Remote Sens. Environ.* **1995**, *52*, 79–94. [CrossRef]

16. Chen, F.; Yan, D.; Zhao, Z. Haze detection and removal in remote sens.ing images based on undecimated wavelet transform. *Geo. Inf. Sci. Wuhan Univ.* **2007**, *1*, 71–74.

17. Addesso, P.; Conte, R.; Longo, M.; Restaino, R.; Vivone, G. Svm-Based cloud detection aided by contextual information. In Proceedings of the 2012 Tyrrhenian Workshop on Advances in Radar and Remote Sensing (TyWRRS), Wessling, Germany, 12–14 September 2012.

18. Dell'Acqua, F.; Gamba, P.; Ferrari, A.; Palmason, J.; Benediktsson, J.; Arnason, K. Exploiting spectral and spatial information in hyperspectral urban data with high resolution. *IEEE Geosci. Remote Sens. Lett.* **2004**, *1*, 322–326. [CrossRef]

19. Li, H.-Q.; Liu, Z.-K.; Lin, F. Aerial image classification method based on fractal theory. *J. Remote Sens* **2001**, *5*, 353–357.

20. Benediktsson, J.A.; Palmason, J.A.; Sveinsson, J.R. Classification of hyperspectral data from urban areas based on extended morphological profiles. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 480–491. [CrossRef]

21. Duro, D.C.; Franklin, S.E.; Dubé, M.G. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using spot-5 HRG imagery. *Remote Sens. Environ.* **2012**, *118*, 259–272. [CrossRef]

22. Foody, G.M. The effect Mis-Labeled Training Data on the Accuracy Supervised Image Classification by Svm. In Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, 26–31 July 2015.

23. 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]

24. Pasolli, E.; Melgani, F.; Tuia, D.; Pacifici, F.; Emery, W.J. SVM active learning approach for image classification using spatial information. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 2217–2233. [CrossRef]

25. Ciresan, D.; Meier, U.; Schmidhuber, J. Multi-Column Deep Neural Networks For image Classification. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Rama Chellappa, RI, USA, 16–21 June 2012.

26. Kussul, N.; Skakun, S.; Kussul, O. Comparative analysis of neural networks and statistical approaches to remote sensing image classification. *Int. J. Comput.* **2014**, *5*, 93–99.

27. Li, L.; Chen, Y.; Yu, X.; Liu, R.; Huang, C. Sub-pixel flood inundation mapping from multispectral remotely sensed images based on discrete particle swarm optimization. *ISPRS J. Photogramm. Remote Sens.* **2015**, *101*, 10–21. [CrossRef]

28. Wang, H.; He, Y.; Guan, H. Application support vector machines in cloud detection using EOS/MODIS. In Proceedings of the Remote Sensing Applications for Aviation Weather Hazard Detection and Decision Support, San Diego, CA, USA, 25 August 2008.

29.	Hughes, M.J.; Hayes, D.J. Automated detection of cloud and cloud shadow in single-date landsat imagery using neural networks and spatial post-processing. *Remote Sens.* **2014**, *6*, 4907–4926. [CrossRef]

30.	Chethan, H.; Raghavendra, R.; Kumar, G.H. Texture based approach for cloud classification using svm. In Proceedings of the International Conference on Advances in Recent Technologies in Communication and Computing, Kottayam, Kerala, India, 27–28 October 2009.

31.	Baseski, E.; Cenaras, C. Texture color based cloud detection. In Proceedings of the 2015 7th International Conference on Recent Advances in Space Technologies (RAST), Istanbul, Turkey, 16–19 June 2015.

32.	Latry, C.; Panem, C.; Dejean, P. Cloud detection with svm technique. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2007, Barcelona, Spain, 23–27 July 2007.

33.	Wyatt, C. *Radiometric Calibration: Theory Methods*; Elsevier: New York, NY, USA, 2012.

34.	Teillet, P. Image correction for radiometric effects in remote sensing. *Int. J. Remote Sens.* **1986**, *7*, 1637–1651. [CrossRef]

35.	Kaufman, Y.J.; Sendra, C. Algorithm for automatic atmospheric corrections to visible and near-ir satellite imagery. *Int. J. Remote Sens.* **1988**, *9*, 1357–1381. [CrossRef]

36.	Guide, E.U. Atmospheric correction module: Quac and flaash user's guide. *Version* **2009**, *4*, 1–44.

37.	Schiller, H.; Brockmann, C.; Krasemann, H.; Schönfeld, W. A method for detection and classification of clouds over water. In Proceedings of the 2nd MERIS/(A)ATSR User Workshop, Frascati, Italy, 22–26 September 2008.

38.	Xiao, Z. *Study on Cloud Detection Method for High Resolution Satellite Remote Sensoring Image*; Harbin Institute of Technology: Heilongjiang, China, 2013.

39.	Bajwa, I.S.; Naweed, M.; Asif, M.N.; Hyder, S.I. Feature based image classification by using principal component analysis. *ICGST Int. J. Graph. Vis. Image Process. GVIP* **2009**, *9*, 11–17.

40.	Haralick, R.M.; Shanmugam, K.; Dinstein, I.H. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* **1973**, *3*, 610–621. [CrossRef]

41.	Pesaresi, M.; Gerhardinger, A.; Kayitakire, F. A robust built-up area presence index by anisotropic rotation-invariant textural measure. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2008**, *1*, 180–192. [CrossRef]

42.	Ouma, Y.O.; Tetuko, J.; Tateishi, R. Analysis of co-occurrence and discrete wavelet transform textures for differentiation of forest and non-forest vegetation in very-high-resolution optical-sensor imagery. *Int. J. Remote Sens.* **2008**, *29*, 3417–3456. [CrossRef]

43.	Puissant, A.; Hirsch, J.; Weber, C. The utility of texture analysis to improve per-pixel classification for high to very high spatial resolution imagery. *Int. J. Remote Sens.* **2005**, *26*, 733–745. [CrossRef]

44.	Anys, H.; Bannari, A.; He, D.; Morin, D. Texture analysis for the mapping urban areas using airborne meis-ii Images. In Proceedings of the First International Airborne Remote Sensing Conference and Exhibition, Strasbourg, France, 12–15 September 1994.

45.	Chang, C.-C.; Lin, C.-J. Libsvm: A library for support vector machines. *ACM Trans. Int. Sys. Tech.* **2011**. [CrossRef]

46.	Mountrakis, G.; Im, J.; Ogole, C. Support vector machines in remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 247–259. [CrossRef]

47.	Paneque-Gálvez, J.; Mas, J.-F.; Moré, G.; Cristóbal, J.; Orta-Martínez, M.; Luz, A.C.; Guèze, M.; Macía, M.J.; Reyes-García, V. Enhanced land use/cover classification of heterogeneous tropical landscapes using support vector machines and textural homogeneity. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *23*, 372–383. [CrossRef]

48.	Anantrasirichai, N.; Achim, A.; Morgan, J.E.; Erchova, I.; Nicholson, L. Svm-Based texture classification in optical coherence tomography. In Proceedings of the 2013 IEEE 10th International Symposium onBiomedical Imaging (ISBI), San Francisco, CA, USA, 7–11 April 2013.

49.	Cristianini, N.; Shawe-Taylor, J. *An Introduction to Support Vector Machines Other Kernel-Based Learning Methods*; Cambridge University Press: Cambridge, UK, 2000.

50.	Christodoulou, C.; Michaelides, S.C.; Pattichis, C.S. Multifeature texture analysis for the classification of clouds in satellite imagery. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 2662–2668. [CrossRef]

51.	Wu, X.; Peng, J.; Shan, J.; Cui, W. Evaluation of semivariogram features for object-based image classification. *Geo-Spat. Inf. Sci.* **2015**, *18*, 159–170. [CrossRef]

52.	Dubuisson-Jolly, M.-P.; Gupta, A. Color and texture fusion: Application to aerial image segmentation and gis updating. *Image Vis. Comput.* **2000**, *18*, 823–832. [CrossRef]

53. Bai, X.; Liu, C.; Ren, P.; Zhou, J.; Zhao, H.; Su, Y. Object classification via feature fusion based marginalized kernels. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 8–12.

54. Hong, Z. *Digital Image Processing*; Science Press: Beijing, China, 2005.

55. Yuan, Y.; Hu, X. Bag-of-words and object-based classification for cloud extraction from satellite imagery. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 4197–4205. [CrossRef]