A Novel Image Fusion Method of Multi-Spectral and SAR Images for Land Cover Classification

Yinghui Quan 1, Yingping Tong 1, Wei Feng 1,*, Gabriel Dauphin 2, Wenjiang Huang 3 and Mengdao Xing 4

1 Department of Remote Sensing Science and Technology, School of Electronic Engineering, Xidian University, Xi’an 710071, China; yhquan@mail.xidian.edu.cn (Y.Q.); yptong@stu.xidian.edu.cn (Y.T.)
2 Laboratory of Information Processing and Transmission, L2TI, Institut Galilée, University Paris XIII, 93430 Paris, France; gabriel.dauphin@univ-paris13.fr
3 Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China; huangwj@aircas.ac.cn
4 Academy of Advanced Interdisciplinary Research, Xidian University, Xi’an 710071, China; xmd@xidian.edu.cn
* Correspondence: wfeng@xidian.edu.cn

Received: 3 October 2020; Accepted: 16 November 2020; Published: 19 November 2020

Abstract: The fusion of multi-spectral and synthetic aperture radar (SAR) images could retain the advantages of each data, hence benefiting accurate land cover classification. However, some current image fusion methods face the challenge of producing unexpected noise. To overcome the aforementioned problem, this paper proposes a novel fusion method based on weighted median filter and Gram–Schmidt transform. In the proposed method, Sentinel-2A images and GF-3 images are respectively subjected to different preprocessing processes. Since weighted median filter does not strongly blur edges while filtering, it is applied to Sentinel-2A images for reducing noise. The processed Sentinel images are then transformed by Gram–Schmidt with GF-3 images. Two popular methods, principal component analysis method and traditional Gram–Schmidt transform, are used as the comparison methods in the experiment. In addition, random forest, a powerful ensemble model, is adopted as the land cover classifier due to its fast training speed and excellent classification performance. The overall accuracy, Kappa coefficient and classification map of the random forest are used as the evaluation criteria of the fusion method. Experiments conducted on five datasets demonstrate the superiority of the proposed method in both objective metrics and visual impressions. The experimental results indicate that the proposed method can improve the overall accuracy by up to 5% compared to using the original Sentinel-2A and has the potential to improve the satellite-based land cover classification accuracy.

Keywords: land cover classification; image fusion; random forest; multi-spectral; remote sensing

1. Introduction

The application of remote sensing satellite images for land cover classification has received widespread attention [1]. Satellite images, with high spatial resolution and rich spectral information, could play a greater role in improving classification accuracy [2]. However, due to the limitation of sensor storage space or communication bandwidth, it is difficult for satellite data to take both spatial resolution and spectral resolution into account [3]. Sentinel-2 [4], the multi-spectral (MS) imaging satellite of the European Space Agency (ESA), contains relatively rich spectral information [5]. However, the maximum resolution of Sentinel-2 is 10 m, which makes it difficult to meet the requirements for ultra-high resolution. Synthetic aperture radar (SAR) sensors can capture images...
of the Earth all the time and in all weather conditions, and therefore play an increasingly important role in land cover classification [6]. SAR images provide richer surface information due to their imaging characteristics, which can compensate for the inability of other sensors to obtain sufficient information. For example, GF-3, a new high-resolution C-band fully polarized SAR image, has attracted more and more people [7]. The images obtained by GF-3 have been widely used and show excellent performance in many fields, such as target recognition and land cover classification [8–10]. However, this kind of data is easily affected by the multiplicative noise. The noise makes SAR images appear grainy and makes SAR interpretation a challenging task [11]. GF-3 data actively emit electromagnetic radiation, while Sentinel-2 data obtain information by passively measuring electromagnetic radiation. Hence, the fusion of these two types of data may provide supplementary information about the imaging area [12].

Data fusion is an effective way of synergistically combining information from various sources to better understand a given scene [13]. The fusion of MS and SAR data can sharpen the details of low-resolution data while providing complementary data observed from the same site [14]. In other words, the fusion of GF-3 and Sentinel-2 can not only help to generate images with rich spatial and spectral information but also contribute to better understand and explain the image area. In addition, a good fusion algorithm needs to ensure that the color content of the original MS image is not distorted while improving the spatial details [15]. With respect to the ordinary pan-sharpening, the fusion of the optical and SAR image is more difficult because the gray values of SAR image are not relevant to those of multi-spectral images [16].

Some researchers have proposed some effective algorithms to solve the problem in the fusion of SAR and optical images. For example, Montgomery et al. use multi-temporal and multi-modal data to develop a classification method based on decision-making and data fusion and apply it to wetland classification [17]. Iervolino et al. propose a multi-sensor data fusion method based on generalized intensity-hue-saturation transform and wavelet transform [18]. The above methods could remove some redundant information and achieved certain fusion results, however, there are noise components in the image surface after fusion [19]. In recent years, the fusion of GF-3 and Sentinel-2 has attracted the attention of some scholars. For example, Wang et al. realize polarization mode information fusion based on curvelet transform [9]. In [20], the full feature fusion method is proposed to classify crops. However, these studies only use part of the Sentinel data bands, not all bands. Consequently, the fusion algorithm of these two kinds of data is urgently needed.

Land cover classification is a vital assessment method to evaluate the performance of image fusion algorithms. For example, the maximum likelihood classifier is adopted to classify images before and after fusion in [18]. Experimental results show that the rich details of SAR texture can improve classification performance compared with MS images. In [21], authors integrate the multi-scale texture features of the image into the random forest (RF) classifier. Zhang et al. apply object-oriented multi-scale segmentation to the image after feature fusion and implement an improved support vector machine (SVM) model for classification [22]. According to algorithm comparison, RF is a powerful and highly recognized ensemble classification method [23,24]. In this paper, RF is selected to evaluate the fusion effect of the image.

Moreover, land cover is an important area of interest in Earth observation (EO) [25]. In recent years, there have been many works on EO data, and these works have proposed fusion algorithms based on multi-source images to obtain robust performance. For example, Rasaei et al. use the Bayesian method to fuse soil data from traditional soil surveys with direct soil information from remote sensing images [26]. In [27], authors merge information of different spectral indices with feature-level fusion technology to extract buildings from high-resolution satellites. Nowadays, powerful tools are still needed to process EO data and to generate EO products [25].

The objective of this paper was to propose a new algorithm to fuse GF-3 and Sentinel-2 images and apply the fusion result to land cover classification. The advantage of the algorithm lies in reducing the noise of GF-3 and keeping the spectral characteristics of Sentinel-2 during the fusion process.
The rest of this paper is organized as follows. At first, the image fusion methods and the random forest classification method are described in Section 2. Section 3 describes the studied area and the imagery being used for this study. In Section 4, the proposed method is presented. The fusion performance comparison and analysis of the experimental results obtained by the different transform fusion methods are made in Section 5. The discussion is presented in Section 6. Finally, the concluding remarks are given in Section 7.

2. Related Works

2.1. Image Fusion Methods

The purpose of image fusion is to integrate information of different spatial and spectral resolutions, so as to generate images with more detail [28]. Image fusion methods could be divided into the pixel, feature, and decision levels roughly [29]. There is information loss in the process of feature and decision level fusion [19]. The pixel-level fusion technology has higher accuracy and fewer data changes; hence, these methods are more suitable for remote sensing applications [30,31].

Pixel-level fusion technology directly acts on the pixel data of multi-source images. Commonly used methods in pixel-level fusion technology include intensity-hue-saturation (IHS) transform [32], Brovey transform [33], principal component analysis (PCA) [34,35] and Gram–Schmidt (GS) transform [36,37]. IHS transform extracts the spatial and spectral information from the original image. However, this method may have a serious problem of spectral distortion [38,39]. Brovey transform is to normalize each band of the MS image and perform a multiplicative operation with the high-resolution image, so as to add the required spatial information to the MS image [40]. The merged image generated by this method can not retain the radiance of the input MS image [41]. Moreover, Brovey transform is the same as IHS transform in that only three bands can be converted at once. PCA fusion method extracts the main components in the MS image, where the first principal component contains most of the spatial information, and the remaining components reflect the specific spectral details of other bands. PCA can reduce data redundancy in the fusion process, however, it may cause the loss of important information in the original image, especially the spectral information [42]. Gram–Schmidt transform is an improved version of the PCA fusion method. Different from PCA, the components of the image generated by GS are orthogonal. Since GS applies the spectral response function to estimate SAR data, the fusion result may be more accurate.

Moreover, MS and SAR fusion images have a wide range of applications in earth sciences. Ma et al. propose a SAR image change detection method based on improved wavelet fusion [43]. The experimental results show that, compared with the original SAR image, their change detection method can better reflect the real change trend and reduce the influence of speckle noise. In [44], Salentinig and Gamba jointly apply multiple pixel fusion methods to multi-scale SAR data sets.

2.2. Classification Methods

Different types of classification methods have been applied to map land cover [45]. Classification methods could be divided into unsupervised algorithms and parameter-supervised algorithms. Among them, unsupervised algorithms mainly include some typical methods, such as K-means and iterative self-organizing data analysis methods. Supervised algorithms mainly include maximum likelihood method, support vector machine [46], some ensemble learning algorithms, etc. However, the structure of some supervised methods (such as neural networks and support vector machines) is very complex and requires a large number of parameters to be adjusted, making it difficult to automate [47].

Ensemble model, an effective method, has been successfully used in remote sensing classification [48–52]. Random forest (RF), a typical ensemble learning method, has received widespread attention due to its excellent classification effect, avoidance of overfitting and fast processing speed. Each tree in RF is constructed using a subset feature and random sampling with substitution. The training samples are divided into left sub-tree and right sub-tree with the best
splitting attribute, selected out from original attributes, with a splitting criterion. The training dataset keeps splitting and divided until the stoppage condition reaches. When the trees are balanced, the RFs are very efficient, as the time complexity of classification is logarithmic in the number of nodes.

RF has been widely used in land cover classification. Tian et al. apply RF to classify wetland land cover in arid areas and find that the RF classification method fused with multi-sensor data can retrieve better land cover information than other classifiers [53]. In [54], RF is used in multi-source EO datasets. The experimental result shows that RF is more stable on the multi-source EO data than the support vector machine. Wu et al. use RF in the fusion data of the GF-2 images, the normalized digital surface model derived from Lidar data and their texture features [55].

3. Material

3.1. Datasets

The experimental data contain the GF-3 radar images and the Sentinel-2A multi-spectral images in five areas. The GF-3 images were acquired on 15 May 2018. The data (level 1A) are the Horizontal transmitting and Horizontal receiving (HH) polarization image in the sliding spotlight (SL) mode with a high spatial resolution (0.99 m in azimuth and 0.56 m in range). The incidence angle is 39.27°. In order to avoid any land cover inconsistencies between the acquisition dates of the image in the same areas, Sentinel-2A data imaged on 14 May 2018 were selected. The Sentinel-2A images were acquired from (https://scihub.copernicus.eu/dhus/) official ESA website. These images (level 1C) have 13 spectral bands, including four bands of 10m, six bands of 20 m and three bands of 60m spatial resolution. The test site is a bound part of Jiangsu province in China.

3.2. Data Preprocessing

The preprocessing of both GF-3 and Sentinel-2A images is implemented first. The speckle noise is randomly distributed in the SAR image and mixed with the ground objects, with a result of reducing the spatial resolution of the SAR image. To reduce the noise of the GF-3 image, $1 \times 1$ (range × azimuth) multi-look processing is applied to all the GF-3 images. The processed data have a range resolution of 0.89 m and an azimuth resolution of 0.99 m. Then, Frost filtering (window size of 3) is adopted to further reduce the negative effect of the noise. For a better registration between GF-3 image and Sentinel-2A image, the terrain radiation correction based on the digital elevation model is used to weaken the geometric deformation of GF-3 caused by the terrain. After preprocessing, the geometric resolution of GF-3 is 1.4 m. For the Sentinel-2A image, the geometric correction and atmospheric correction are implemented in the preprocessing step. Since there are three different resolutions between Sentinel bands, Sentinel Application Platform (SNAP) software is used to resample all the 20 m and 60 m bands to 10m resolution by the nearest neighbor method. After data preprocessing, all GF-3 images and the three-band color composite of Sentinel-2A images (bands 4, 3, 2) are shown in Figure 1.
Figure 1. Cont.
3.3. Evaluation Methods

Five statistical parameters are used to evaluate the quality of different aspects of the fused image, namely standard deviation (SD), average gradient (AG), spatial frequency (SF), peak signal-to-noise ratio (PSNR) and correlation coefficient (CC). Moreover, three classic evaluation indicators are used to test the effect of the classifier, namely overall accuracy (OA), Kappa statistics, and McNemar’s test [56]. In the following equations, $F(i, j)$ and $R(i, j)$ respectively represent the fused image and the reference image, $M$ and $N$ are the numbers of rows and columns of the whole image and $(i, j)$ is the index. $P = (p_{xy})_{c \times c}$ is the confusion matrix, $c$ is the number of classes, $p_{xy}$ represents the total number of pixels belonging to category $x$ but classified into category $y$. Sentinel-2A image is used as the reference image.

- Standard deviation (SD): It calculates the contrast in the image. The larger the standard deviation, the more scattered the gray level distribution, which means that more information is available.

$$SD = \sqrt{\frac{\sum_{m=1}^{M} \sum_{n=1}^{N} F(m, n) - \bar{F}}{M \times N}} \quad (1)$$

where $\bar{F}$ is the average gray value of the fused image.

- Average gradient (AG): It reflects the image’s ability to express small detail contrast and texture changes, as well as the sharpness of the image. It is desirable to have this value as high as possible.
\[ AG = \frac{\sum_{m=1}^{M} \sum_{n=1}^{N} (\Delta I_i^2 + \Delta I_j^2) / 2}{M \times N} \]  

where \( I_i = F(i, j) - F(i - 1, j) \) and \( I_j = F(i, j) - F(i, j - 1) \).

- Spatial frequency (SF): It measures the overall activity level of the image. Fusion images with high spatial frequency values are preferred. SF is expressed in row frequency (RF) and column frequency (CF).
  \[ SF = \sqrt{RF^2 + CF^2} \]  

where RF and CF are given by

\[ RF = \sqrt{\frac{1}{M \times N} \sum_{m=1}^{M} \sum_{n=2}^{N} (F(i, j) - F(i, j - 1))^2} \]  

\[ CF = \sqrt{\frac{1}{M \times N} \sum_{m=2}^{M} \sum_{n=1}^{N} (F(i, j) - F(i - 1, j))^2} \]  

- Peak signal-to-noise ratio (PSNR): It calculates the visual error between the fused image and the reference image. A large value of PSNR indicates that the distortion of the fused image is small.
  \[ PSNR = 10 \log_{10} \frac{MAX}{MSE} \]  

where \( MAX \) is the maximum pixel value of the image and \( MSE \) represents the mean square error of the image.

\[ MSE = \frac{1}{M \times N} \sum_{m=1}^{M} \sum_{n=1}^{N} (F(i, j) - R(i, j))^2 \]  

- Correlation coefficient (CC): It measures the correlation between the reference image and the fused image.
  \[ CC = \frac{\sum_{m=1}^{M} \sum_{n=1}^{N} [R(m, n) - \bar{R}] [F(m, n) - \bar{F}]}{\sqrt{\sum_{m=1}^{M} \sum_{n=1}^{N} [R(m, n) - \bar{R}]^2 \sum_{m=1}^{M} \sum_{n=1}^{N} [F(m, n) - \bar{F}]^2}} \]  

where \( \bar{R} \) is the mean value of reference image.

- Overall accuracy (OA): It represents the proportion of samples that are correctly classified among all samples.
  \[ OA = \frac{\sum_{x=1}^{c} p_{x,y}}{\sum_{y=1}^{c} \sum_{x=1}^{c} p_{x,y}} \]  

- Kappa statistics: It is a multivariate statistical method for classification accuracy.
  \[ \kappa = \frac{\tau \sum_{x=1}^{c} p_{x,y} - \sum_{x=1}^{c} (p_{x} \cdot p_{y})}{\tau^2 - \sum_{x=1}^{c} (p_{x} \cdot p_{x})} \]  

where \( \tau \) is the total number of test samples, \( p_{x} \) and \( p_{y} \) respectively represent the sum of the \( x \)-th row and the \( y \)-th column of the confusion matrix.

- McNemar’s test: It is used to determine whether a learning algorithm is superior to another in a particular task with acceptable the probability of false detection of differences when there is no difference. Suppose there are two algorithms A and B.
  \[ Z = \frac{e_{01} - e_{10}}{\sqrt{e_{01} + e_{10}}} \]  

where \( e_{01} \) is the number of samples misclassified by algorithm A but not B, and \( e_{10} \) represents the number of samples misclassified by algorithm B but not A. The \( p \)-value of the Chi-squared
value is calculated by McNemar’s test. Since a small $p$-value indicates that the null hypothesis is unlikely to be established, we may reject the null hypothesis if the probability that $Z \geq 1.96$ is less than 0.05.

### 3.4. Parameter Setting

In order to ensure that each datum pixel in the high-resolution GF-3 image has a value based on the Sentinel-2 image pixel, each Sentinel-2 image is adjusted to the same size as the corresponding GF-3 image through the nearest neighbor resampling. Two parameters need to be set in the RF classifier: the number of decision trees $T$ and the number of features $K$. In this study, $T$ is set to 100 and $K$ is the square root of the number of inputs. Through visual interpretation, it is found that the classes of woodland, residential land, crops, roads and water are obvious in the study areas. Therefore, all areas are divided into these five categories through random forest classification. A total of 1000 samples are extracted from each study area through visual interpretation. The number of samples in each class is the same. The training samples and test samples are randomly selected from all samples.

### 4. The Proposed Method

The experiment in this paper consists of three steps, namely data preprocessing, image fusion and result evaluation. Firstly, the Sentinel-2A image and GF-3 image are respectively subjected to different preprocessing processes. Secondly, the proposed fusion algorithm is used in the preprocessed image. Meanwhile, PCA and GS are used in the comparison. Finally, the quality assessment function and random forest classification are used to evaluate the fusion image. The flowchart of the experiment is shown in Figure 2.

![Figure 2. Flowchart of algorithm design.](image)
4.1. Weighted Median Filter

Weighted median filter (WMF) mathematically corresponds to global optimization. It can effectively filter images while not strongly blurring edges [57]. It is an operator that replaces the current pixel with the weighted median of neighboring pixels within a local window. Consider only pixels within the local window \(R(p)\) of radius \(r\) centered at \(p\) in processing \(\tilde{p}\) in image \(I\). For each pixel \(q \in R(p)\), WMF associates it with a weight \(\omega_{pq}\) in corresponding feature map \(f\), i.e.,

\[
\omega_{pq} = g(f(p), f(q))
\]  

(12)

where \(f(p)\) and \(f(q)\) are features at pixels \(p\) and \(q\), \(\omega\) is a typical influence function between neighboring pixels.

In [57], joint-histogram, which is a two-dimensional histogram structure for storing pixel count, is used to regenerate weights every time the window shifts. Each pixel \(q\) in the filter is associated with its value \(I_q\) and feature \(f(q)\). Suppose the total value of pixels is \(N_I\) and denote the pixel value index of \(I(q)\) as \(i\). Then, list all possible pixel values in ascending order as \(L = \{I_0, I_1, \ldots, I_{N_I-1}\}\). Similarly, suppose there are \(N_f\) different features and denote the feature index for \(f(p)\) as \(f\) after ordering all features as \(F = \{f_0, f_1, \ldots, f_{N_f-1}\}\).

In the joint-histogram \(H\), the pixel \(q\) is put into the histogram bin \(H(i, f)\) in the \(f\)-th row and the \(i\)-th column.

\[
H(i, f) = \sum \{q \in R(p) \mid I(q) = I_i, f(q) = f_f\}
\]  

(13)

where \(\Sigma\) counts the number of elements. For any pixel belonging to bin \(H(i, f)\), due to the filter center \(p\), its weight is calculated as \(g(f_f, f_q)\). All weights can be obtained by traversing the joint histogram. The total weight for all pixels with value index \(i\) is

\[
\omega_i = \sum_{f=0}^{N_f-1} H(i, f) g(f_f, q_f)
\]  

(14)

In the case where all \(i\) have the weight \(\omega_i\), the weighted median can be found.

By combining the joint histogram and sliding window strategy, the input images are processed in a scan line order. Traverse the joint histogram to calculate the weight \(\{\omega_k\}_{k=0}^{N_I-1}\), and add them to get the total weight \(\omega_i\) for the first pass. In the second pass, the sum of the weights is at most half of \(\omega_i\), and the pixel value is output. Initially, at the first position of the image, the joint histogram calculates the number of the pixel whose index \((i, f)\) is within the filter range. The number is placed in the corresponding cell.

4.2. Gram–Schmidt Transform (GS)

Gram–Schmidt (GS) transform is one of the pan-sharpening methods for multi-spectral images.

Firstly, to make the explanation simpler, two assumptions are proposed. The first concerns the mean of the intensity of each wavelength. It is first subtracted on both of the GF-3 and on the Sentinel-2A image before any processing,

\[
I'_{i,j,k} = I_{i,j,k} - \frac{1}{N_I N_J \sum_{i,j} I'_{i,j,k}}
\]  

(15)

where \(k\) indicates the wavelength of which \(I'_{i,j,k}\) is the intensity; \(i, j\) are the pixel coordinates and the summation on \(i, j\) is assumed to cover the whole image; \(N_I \times N_J\) is the size of image \(I\). After all processing, these means are added back.

\[
I''_{i,j,k} = I'_{i,j,k} + \frac{1}{N_I N_J \sum_{i,j} I'_{i,j,k}}
\]  

(16)
The second assumption states that describing pixel locations using lines and columns is here not an issue. A raster scanning order is chosen and images are considered as column vectors where the \( n \)th-component is the \( n \)th-image-intensity read when processing the image along with the chosen raster scanning order. Hence, images are represented as matrices, where each column is assigned with wavelength-specific intensities and each line is assigned with pixel-specific intensities.

\[
I_{i \times N_j + j, k} = I_{i,j,k}
\]  

(17)

when \( N_j \) is the number of columns of the image \( I \), and \( I \) is the matrix associated to \( I \).

The following notations will be used.

- \( E, F \): number of wavelength available in Sentinel-2A and in GF-3.
- \( N_e, N_f \): total number of pixels in Sentinel-2A and in GF-3.
- \( F \): matrix related to the GF-3 image.
- \( E \): matrix related to the Sentinel-2A image.
- \( S \): matrix related to Gram–Schmidt transformed image.

Secondly, panchromatic images are extracted from Sentinel-2A. The single panchromatic wavelength image \( B \) can be computed on each pixel of the Sentinel-2A image with weights denoted as \( w_k \).

\[
B_n = \sum_{k=0}^{E-1} E_{n,k} w_k
\]  

(18)

Using the new notations, this can be rewritten as

\[
B_n = EW^T
\]  

(19)

where \( W \) is a line-vector of size \( E \) whose components are \( w_k \) and \( ^T \) stands for the action of transposing a matrix.

A vector \( v \) is projected onto a line containing vector \( u \):

\[
P_u(v) = \frac{u \cdot v}{u \cdot u} u
\]  

(20)

where \( \cdot \) is the dot product defined as the sum of the products of corresponding components.

As \( F_k \) is a column of \( F \), \( F_k \) is a vector and it can be used to define a projection operator that can be applied to \( B \).

\[
P_{F_k}(B) = \frac{F_k^T \cdot B}{F_k^T \cdot F_k} F_k
\]  

(21)

GF-3 image is an SAR image not using any wavelength within or near to the visible band. From a statistical viewpoint, it makes sense to compute a panchromatic image whose pixel values are linear combinations of GF-3 pixel-values and the weights take into account how each wavelength intensity is statistically correlated to a Sentinel-2A panchromatic intensity. This computation using one weight is actually equivalent to a projection onto a wavelength-specific image and the whole computation is equivalent to adding all projections onto all \( F \) wavelengths.

\[
\hat{B} = \left[ \begin{array}{c} F_k^T B \ F_k \end{array} \right] k = FF^T B \sum_D^{-1}
\]  

(22)

where \( \sum_D \) is a diagonal \( F \times F \) matrix whose diagonal components are \( F_k^T F_k \).

Let us denote \( \hat{W} \) the line-vector of size \( F \) equal to \( (F^T B \sum_D^{-1})^T \). The panchromatic GF-3 image is defined as

\[
\hat{B} = FW^T
\]  

(23)

By replacing the values of the first wavelength with those of \( \hat{B} \), a modified GF-3 image is obtained.
\[ F' = F\hat{W}^T = F \begin{bmatrix} \hat{W} \\ 010 \cdots 0 \\ 001 \cdots 0 \\ \vdots \\ 000 \cdots 1 \end{bmatrix} \]  

(24)

Actually, \( \hat{W} \) is an identity-matrix except for the first line equal to \( \hat{W} \).

The Gram–Schmidt is an algorithm performing a Q-R decomposition, this transform is applied to \( \Sigma' = QR \) with the following properties.

- \( Q \) is an orthogonal matrix, \( Q^TQ = 1_F \) where \( 1_F \) is the identity matrix.
- \( R \) is an upper triangle matrix having an upper triangle matrix inverse and \( QQ^{-1} = 1_F \).

The \( R \) matrix yielded from the Q-R decomposition of the variance-covariance matrix when applied to \( F' \) yields a new image.

\[ S = F'R^{-1} \]  

(25)

An important property of the Q-R decomposition here is that the first column of \( S \) is proportional to that of \( F' \) and hence can still be interpreted as a panchromatic wavelength. This stems from the fact that \( R^{-1} \) is an upper triangle matrix having only one non-zero component in the first column.

The first column of \( S \) having this panchromatic interpretation is now replaced with the true panchromatic image \( B \). All other columns remain unchanged.

\[ S' = SW_F^T + EW_F^T \text{and} W_S = \begin{bmatrix} 00 \cdots 0 \\ 10 \cdots 0 \\ 01 \cdots 0 \\ \vdots \\ 00 \cdots 1 \end{bmatrix}, \quad W_F = \begin{bmatrix} W_0 \cdot w_1 \cdots w_{E-1} \\ 00 \cdots 0 \\ 00 \cdots 0 \\ \vdots \\ 00 \cdots 0 \end{bmatrix} \]  

(26)

Using the \( R \) matrix, the modified GF-3 image is obtained. The first column of the image has this SAR interpretation, while the other columns have the classical multi-spectral interpretation.

\[ \hat{F} = S'R \]  

(27)

5. Experimental Results

This section shows the results obtained by different image fusion methods and evaluates the results.

5.1. Visual Comparison

Figure 3 shows the fusion results of GF-3 images and Sentinel-2A images obtained by PCA method, GS transform method and WMF-based Gram–Schmidt transform (WMFGS). All bands in the Sentinel image are used for fusion, and each fused image contains 13 bands. Figure 3 is displayed in the band combination red, blue and green. The geometric resolution of the fused images is 1.4 m, which is the same as the resolution of the preprocessed GF-3 image.

In all fused images, it is obvious that some characteristic features such as roads, field boundaries (i.e., linear features) and residential features are inherited from SAR images. For PCA fusion images, although most of the multi-spectral image information is retained, the SAR image information is not well utilized, especially the edge information of water bodies and buildings. Compared with the PCA method, the GS fusion method makes better use of the advantages of SAR images. The characteristics of the water body in the GS fusion image can be seen in Figure 3h,k. The images fused by WMFGS retain the advantages of GS fusion. As a result of the use of WMF, these images can not only effectively filter out the noise in the sentinel images, but also protect the details of the images.
In all results, WMFGS fused results have slightly higher quality as image brightness and sharpness. Visually, the WMFGS fusion algorithm (Algorithm 1) has the best fusion ability compared with the PCA fusion algorithm and the traditional GS fusion algorithm.

**Algorithm 1** Weighted Median Filter-Based Gram–Schmidt Image Fusion Method (WMFGS).

**Input:**
- \( F \): GF-3 image
- \( E \): Sentinel-2A image
- \( F \): number of wavelength available in GF-3
- \( E \): number of wavelength available in Sentinel-2A
- \( N_f \): total number of pixels in GF-3
- \( N_e \): total number of pixels in Sentinel-2A

1. for \( i = 1 : N_e \) do
2. Compute \( \omega_i \) with (13) and (14)
3. Compute \( P_{k_l}(B) \) using \( P_u(v) \) with (19)–(21)
4. end for
5. for \( j = 1 : N_f \) do
6. Compute \( \hat{B} \) with (23))
7. Compute \( F' \) using \( \hat{W} \) with (24)
8. Compute \( S \) using \( \sum' \) and \( F' \) with (25)
9. end for
10. for \( i = 1 : N_e \) do
11. Compute \( S' \) using \( S \) and \( W \) with (26)
12. Compute \( \hat{F}' \) with (27)
13. end for

**Output:**
- \( \hat{F}' \): fusion image

**Figure 3.** Cont.
5.2. Image Histogram

The image histogram can describe the frequency of image intensity through a two-dimensional distribution curve [58]. Generally speaking, the image intensity histogram can reflect the image’s lighting, contrast, and saturation effects. Histograms are used to evaluate different image fusion methods, which can overcome the shortcomings of conventional evaluation indicators. The histogram of the fusion images and the Sentinel image in the gray level is shown in Figure 4.
Generally, the abscissa of the histogram is between 0 and 255. Since the grayscale range of images is concentrated between 0 and 12, in order to make the image easier to observe, the grayscale range with a value of 0 is cropped in Figure 4. The curve of the Sentinel grayscale image has a single peak. Only in Area 1, based on the curves of the PCA fusion image and GS fusion image, two different peaks can be seen. The WMFGS-based fusion method can provide a larger image range in the peak. In all study areas, images based on WMFGS fusion have the most abundant spectral information.

5.3. Quantitative Analysis

Visual evaluation of fusion products is relatively subjective, therefore, the quality metric of image fusion needs to be evaluated objectively. SAR and multi-spectral image fusion is a special category of remote sensing image fusion, and the same quality indicators are used for evaluation [19].

To evaluate the performance of the fusion methods, the fused images are compared with the original Sentinel-2A image. The performance parameters are also evaluated and tabulated in Table 1.

In all research areas, the WMFGS fused images have the highest SD value. Since SD reflects the dispersion of the image gray relative to the average value, the larger the SD, the more information available in the image and the better the fusion quality. AG and SF are indicators to measure image spatial information. These two values of the WMFGS fused image are the largest in other regions, except that the GS fused image in Area 3 has a higher AG and SF. Since PSNR calculates the visual error between the images, it shows that the visual error between the WMFGS fusion image and the Sentinel image is the smallest in Areas 1, 2 and 5. The value of the CC ranges from 0 to 1. The correlation coefficients of all images have little difference.
Table 1. Quantitative analysis on the fused images.

| Area  | SD        | AG        | SF        | PSNR      | CC        |
|-------|-----------|-----------|-----------|-----------|-----------|
|       | PCA fused image | 301.3385  | 8.4444    | 59.0375   | 89.6709   | 0.6908    |
|       | GS fused image | 362.5207  | 8.5403    | 132.8691  | 94.5200   | 0.7078    |
|       | WMFGS fused image | 365.7824  | 9.4150    | 133.5344  | 94.5325   | 0.7107    |
|       | PCA fused image | 288.3498  | 9.5105    | 78.5329   | 90.0781   | 0.6040    |
|       | GS fused image | 324.7555  | 9.5290    | 177.7829  | 92.0075   | 0.6077    |
|       | WMFGS fused image | 329.9461  | 10.4368   | 178.6462  | 92.0360   | 0.6360    |
|       | PCA fused image | 620.3025  | 10.2363   | 131.9687  | 95.5631   | 0.7313    |
|       | GS fused image | 646.8171  | 10.7526   | 132.6416  | 95.5631   | 0.7344    |
|       | WMFGS fused image | 657.6024  | 9.9434    | 120.6161  | 95.6224   | 0.7344    |
|       | PCA fused image | 497.7268  | 9.1275    | 48.9170   | 76.2159   | 0.6041    |
|       | GS fused image | 453.9458  | 9.9560    | 70.0637   | 83.0082   | 0.6278    |
|       | WMFGS fused image | 641.7992  | 9.9691    | 71.6464   | 82.9793   | 0.6303    |
|       | PCA fused image | 274.4397  | 9.2216    | 80.3165   | 94.1552   | 0.7368    |
|       | GS fused image | 277.0106  | 9.2422    | 109.9329  | 97.1569   | 0.7775    |
|       | WMFGS fused image | 282.7229  | 9.4953    | 110.7430  | 98.8695   | 0.7810    |

5.4. Random Forest Classification Results

The purpose of remote sensing image fusion is to better apply in land cover. To further compare the fusion image in land cover, RF is used to classify the Sentinel image and all fusion images obtained from the new method. The same training samples are used for random forest classifications on both Sentinel-2A and the fusion images.

Figure 5 shows the random forest classified maps when tested on each dataset. Samples classified as residential land, forest, road, water and crop are respectively represented by red, dark green, brown, blue and light green.

The Sentinel classification shows some considerable speckling throughout the study area, whereas the WMFGS classification map is much more homogeneous. The color maps obtained by WMFGS fusion method are more clear and smooth, especially the details of road and water bodies are accurately extracted. The classification result of the sentinel image has low pixels because the original image has a resolution of only 10 meters. Although the PCA fusion method improves the spatial resolution of the image, there is a certain loss of spectral information. Therefore, the classification map based on this algorithm has relatively many wrong pixels.

The overall accuracy (OA) and Kappa coefficient of the classification maps are calculated to further compare the classification results. Tables 2 and 3 respectively show the OA and Kappa coefficient of Sentinel images and the fused images classification map. Due to the loss of spectral information caused by the PCA fusion algorithm, the OA and Kappa values of the PCA fused images are the lowest in all studied areas. The classification of the GS fusion image is generally better than the original Sentinel image, especially in Area 5, the OA is increased by 3.58%, and the Kappa value is increased by 0.0457. The classification accuracy of sentinel image is higher than that of GS fused image in Area 2 and Area 4. Among all images, WMFGS fused images have the highest classification accuracy. Especially in Area 1, the OA of WMFGS fused image is improved by 5.36% compared with sentinel image, and Kappa is improved by 0.07.

Table 4 shows the results of Sentinel and PCA, GS, WMFGS classification comparison. All of McNemar’s test results in the table are greater than 1.96, indicating that all \( p \)-values are less than 0.05. The classification results of Sentinel images are significantly different from those of all other fused images.
Figure 5. Cont.
Figure 5. Cont.
Figure 5. Cont.
Figure 5. Cont.
Figure 5. Random forest classification results of Sentinel images, PCA fused images, GS fused images and WMFGS fused images.

Table 2. Overall accuracy of the random forest (RF) on test areas.

|       | Sentinel | PCA | GS | WMFGS |
|-------|----------|-----|----|-------|
| Area 1| 83.97%   | 79.58% | 85.64% | 89.33% |
| Area 2| 88.71%   | 86.50% | 91.78% | 92.96% |
| Area 3| 93.63%   | 92.76% | 92.21% | 95.45% |
| Area 4| 86.32%   | 82.87% | 83.78% | 89.13% |
| Area 5| 90.97%   | 88.08% | 94.55% | 95.42% |

Table 3. Kappa coefficient of the RF on test areas.

|       | Sentinel | PCA | GS | WMFGS |
|-------|----------|-----|----|-------|
| Area 1| 0.7910   | 0.7334 | 0.8136 | 0.8610 |
| Area 2| 0.8524   | 0.8236 | 0.8923 | 0.9076 |
| Area 3| 0.9168   | 0.9049 | 0.8977 | 0.9403 |
| Area 4| 0.8205   | 0.7755 | 0.7867 | 0.8558 |
| Area 5| 0.8835   | 0.8442 | 0.9292 | 0.9406 |
Table 4. Results on McNemar’s test.

|                | Area 1      | Area 2      | Area 3      | Area 4      | Area 5      |
|----------------|-------------|-------------|-------------|-------------|-------------|
| Sentinel & PCA | 3.725523    | 2.426208    | 2.110579    | 2.114182    | 3.323549    |
| Sentinel & GS  | 4.058187    | 14.41599    | 8.181650    | 4.221159    | 7.939799    |
| Sentinel & WMFGS| 2.994345    | 13.69692    | 9.777861    | 6.246950    | 7.559289    |

6. Discussion

Due to the limitation of imaging equipment, multi-spectral data have relatively rich spectral information, but its spatial resolution is relatively low. To meet social and economic applications, more flexible global land cover data products are needed. However, land cover classification maps obtained from multi-spectral images are not yet ready to meet this demand because of their low resolution. This paper aims to fuse multiple sensors to generate synthetic data with fine resolution. The proposed pixel-level fusion method and the classification method used can provide a cost-effective way to generate high spectral and spatial resolution for improving the land cover classification accuracy.

In this study, a novel image fusion method of Sentinel image and SAR image based on weighted median filter and Gram–Schmidt transform is proposed. Qualitative and quantitative indicators are used to evaluate the quality of the fused image. As a supplementary benefit, the training of random forest verifies the role of fusion images in land classification. The results of the five study areas demonstrate the effectiveness of the proposed method.

(1) In this study, a weighted median filter is used to combine with the Gram–Schmidt transform method. The result of Table 3 shows that compared with the traditional GS fusion method, the weighted median filter can improve the image quality. This is because WMF removes the noise in the Sentinel image while maintaining the details of the image. As described in reference [59], WMF could extract sparse salient structures, like perceptually important edges and contours, while minimizing the color differences in images regions. Since WMF is capable of identifying and filtering the noise while still preserving important structures, the classification map of WMFGS has higher OA and Kappa values.

(2) In Figures 1 and 3, comparing the fused image with the Sentinel image, the contrast between the colors of different classes in the fused image is obvious. In Areas 3 and 4, GS fused images and WMFGS fused images are relatively good at dividing urban areas and roads, while there is no obvious difference in the PCA fused images. In addition, the RF classification maps also show that the PCA fused images do not have better performance than the Sentinel images on the land cover. This is because the spectral information of the original image has a certain percentage of loss in the PCA algorithm fusion process. Comparing the details in Figure 5, the sentinel image has a roughter classification map due to its lower resolution. Although the resolution of PCA fusion images has been significantly improved, many tiny places are misclassified. Water bodies have special spectral characteristics and are easier to distinguish from other types in land cover classification. However, the inner city river is still difficult to distinguish. In the classification map of the WMFGS fusion image, small rivers also have relatively high classification accuracy.

(3) In this study, five indicators, including standard deviation, average gradient, spatial frequency, peak signal-to-noise ratio and correlation coefficient, are used to evaluate the quality of the fusion image. As shown in Table 1, WMFGS fused images have the best quality in most areas in terms of their own image quality and the correlation with the source image. In future work, more indicators will be used to evaluate the fusion image.

(4) In this paper, RF is used for land cover classification. The OA and Kappa coefficient of the five study areas are respectively displayed in Table 2 and 3. Compared with Sentinel images, when using WMFGS fusion method, the classification performance is superior for all study areas. The OA and Kappa coefficient have been significantly improved. The classification performance of GS fused images is also improved in Areas 1, 2 and 5. This is mainly due to the addition of
GF-3 image information. The classification results prove that the fusion image is very helpful for land cover.

(5) There are also many researchers using GF-3 or Sentinel-2 data for land cover classification. For example, Fang et al. propose a fast super-pixel segmentation method to classify GF-3 polarization data, which has an excellent performance in land cover classification [60]. However, their research also requires pre-classification based on optical data. In [61], Sentinel-2 image is used in the land cover study of the Red River Delta in Vietnam, and the balanced and imbalanced data sets are analyzed, but only the 10 band images with 20m resolution are used for classification. Different from the pixel-level algorithm proposed in this paper, Gao et al. use a feature-level fusion algorithm to fuse the covariance matrix of GF-3 with the spectral band of Sentinel-2 [20].

(6) The proposed method provides a feasible option for generating images with high spatial resolution and rich spectral information, thereby extending the application of EO data. The method is suitable for the efficient fusion of registered SAR images and MS images. Using the proposed method can make full use of existing SAR and MS satellite data to generate products with different spatial resolutions and spectral resolutions, thereby optimizing the results of land cover classification.

7. Conclusions

In this paper, a novel image fusion method of Sentinel-2A image and GF-3 image is proposed, which can simultaneously retain the spatial information in SAR images and the rich textural details in visible images. The proposed method is based on Gram–Schmidt transform, and can effectively avoid the loss of spectral information in the fusion process. In particular, weighted median filter is used to reduce noise. Five datasets are used for image fusion and evaluation. To examine performance, the proposed method is evaluated both by visual and quantitative methods and compared with various other fusion algorithms. By inputting the fused images into the random forest ensemble classifier, the image quality can be further checked. The classifying results show that the WMFGS fused images could effectively improve the classification accuracy and obtain accurate land cover maps. Recently, some approaches show great improvements in the remote sensing image fusion by using cloud computing [62,63], which will be exploited in our future work. In addition, we expect to extend and test the proposed method to other applications such as land cover change monitoring.

Author Contributions: Y.Q. and W.F. conceived and designed the experiments; Y.T. performed the experiments and wrote the paper. W.F. and G.D. revised the paper. W.H. and M.X. edited the manuscript. All authors reviewed and approved the final manuscript.

Funding: This work is supported by the National Natural Science Foundation of China (61772397,12005169), National Key R&D Program of China (2016YFE0200400), the Open Research Fund of Key Laboratory of Digital Earth Science (2019LDE005), science and technology innovation team of Shaanxi Province (2019TD-002), and Special fund for basic scientific research project in the central scientific research institutes (Institute of Grassland Research of CAAS) (1610332020026).

Acknowledgments: The authors would like to thank Prof. M. Xing from Xidian University for providing the GF-3 data.

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

References
1. Shiraishi, T.; Motohka, T.; Thapa, R.B.; Watanabe, M.; Shimada, M. Comparative Assessment of Supervised Classifiers for Land Use Land Cover Classification in a Tropical Region Using Time-Series PALSAR Mosaic Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2014, 7, 1186–1199.
2. Thomas, C.; Ranchin, T.; Wald, L.; Chanussot, J. Synthesis of Multispectral Images to High Spatial Resolution: A Critical Review of Fusion Methods Based on Remote Sensing Physics. *IEEE Trans. Geosci. Remote Sens.* 2008, 46, 1301–1312.
3. Cheng, M.; Wang, C.; Li, J. Sparse Representation Based Pan-sharpening Using Trained Dictionary. Geo-sci. Remote Sens. Lett. 2014, 11, 293–297.

4. Wang, Q.; Shi, W.; Li, Z.; Atkinson, P.M. Fusion of Sentinel-2 images. Remote Sens. Environ. 2016, 187, 241–252. doi:10.1016/j.rse.2016.10.030.

5. Shi, W.; Zhu, C.; Tian, Y.; Nichol, J. Wavelet-based image fusion and quality assessment. Int. J. Appl. Earth Obs. Geoinformation 2005, 6, 241–251. doi:10.1016/j.jag.2004.10.010.

6. Ma, X.; Wu, P.; Wu, Y.; Shen, H. A Review on Recent Developments in Fully Polarimetric SAR Image Despeckling. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2018, 11, 743–758.

7. Qingjun, Z. System design and key technologies of the GF-3 satellite. Acta Geod. Cartogr. Sin. 2017, 46, 269.

8. Jiao, N.; Wang, F.; You, H. A New RD-RFM Stereo Geolocation Model for 3D Geo-Information Reconstruction of SAR-Optical Satellite Image Pairs. IEEE Access 2020, 8, 94654–94664. doi:10.1109/ACCESS.2020.2991199.

9. Wang, J.; Chen, J.; Wang, Q. Fusion of POLSAR and Multispectral Satellite Images: A New Insight for Image Fusion. In Proceedings of the 2020 IEEE International Conference on Computational Electromagnetics (ICCEM), Singapore, 24–26 August 2020; pp. 83–84. doi:10.1109/ICCEM47450.2020.9219457.

10. Zhu, Y.; Liu, K.; W Myint, S.; Du, Z.; Li, Y.; Cao, J.; Liu, L.; Wu, Z. Integration of GF2 Optical, GF3 SAR, and UAV Data for Estimating Aboveground Biomass of China’s Largest Artificially Planted Mangroves. Remote Sens. 2020, 12, 2039.

11. Silveira, M.; Heleno, S. Separation Between Water and Land in SAR Images Using Region-Based Level Sets. IEEE Geosci. Remote Sens. Lett. 2009, 6, 471–475. doi:10.1109/LGRS.2009.2017283.

12. Fernandez-Beltran, R.; Haut, J.M.; Paoletti, M.E.; Plaza, J.; Plaza, A.; Pla, F. Multimodal Probabilistic Latent Semantic Analysis for Sentinel-1 and Sentinel-2 Image Fusion. IEEE Geosci. Remote Sens. Lett. 2018, 15, 1347–1351. doi:10.1109/LGRS.2018.2843866.

13. Ye, Z.; Prasad, S.; Li, W.; Fowler, J.E.; He, M. Classification Based on 3-D DWT and Decision Fusion for Hyperspectral Image Analysis. IEEE Geosci. Remote Sens. Lett. 2014, 11, 173–177. doi:10.1109/LGRS.2013.2251316.

14. Vohra, R.; Tiwari, K.C. Comparative Analysis of SVM and ANN Classifiers using Multilevel Fusion of Multi-Sensor Data in Urban Land Classification. Sens. Imaging 2020, 21, 17. doi:10.1007/s11220-020-00280-9.

15. Taha, L.G.E.d.; Elbeih, S.F. Investigation of fusion of SAR and Landsat data for shoreline super resolution mapping: the northeastern Mediterranean Sea coast in Egypt. Appl. Geomat. 2010, 2, 177–186. doi:10.1007/s11220-010-0033-x.

16. Byun, Y.; Choi, J.; Han, Y. An Area-Based Image Fusion Scheme for the Integration of SAR and Optical Satellite Imagery. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2013, 6, 2212–2220.

17. Montgomery, J.; Brisco, B.; Chasmer, L.; Devito, K.; Cobbaert, D.; Hopkinson, C. SAR and Lidar Temporal Data Fusion Approaches to Boreal Wetland Ecosystem Monitoring. Remote Sens. 2019, 11, 161, doi:10.3390/rs11020161.

18. Iervolino, P.; Guida, R.; Riccio, D.; Rea, R. A Novel Multispectral, Panchromatic and SAR Data Fusion for Land Classification. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2019, 12, 3966–3979.

19. Kulkarni, S.C.; Rege, P.P. Pixel level fusion techniques for SAR and optical images: A review. Inf. Fusion 2020, 59, 13–29. doi:10.1016/j.inffus.2020.01.003.

20. Gao, H.; Wang, C.; Wang, G.; Zhu, J.; Tang, Y.; Shen, P.; Zhu, Z. A crop classification method integrating GF-3 PolSAR and Sentinel-2A optical data in the Dongting Lake Basin. Sensors 2018, 18, 3139.

21. Rodriguez-Galiano, V.; Chica-Olmo, M.; Abarca-Hernandez, F.; Atkinson, P.; Jeganathan, C. Random Forest classification of Mediterranean land cover using multi-seasonal imagery and multi-seasonal texture. Remote Sens. Environ. 2012, 121, 93–107. doi:10.1016/j.rse.2011.12.003.

22. Zhang, R.; Tang, X.; You, S.; Duan, K.; Xiang, H.; Luo, H. A Novel Feature-Level Fusion Framework Using Optical and SAR Remote Sensing Images for Land Use/Land Cover (LULC) Classification in Cloudy Mountainous Area. Appl. Sci. 2020, 10, 2928.

23. Feng, W.; Huang, W.; Ren, J. Class Imbalance Ensemble Learning Based on the Margin Theory. Appl. Sci. 2018, 8, 815.

24. Feng, W.; Dauphin, G.; Huang, W.; Quan, Y.; Bao, W.; Wu, M.; Li, Q. Dynamic Synthetic Minority Over-Sampling Technique-Based Rotation Forest for the Classification of Imbalanced Hyperspectral Data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2019, 12, 2159–2169.
25. Gamba, P. Human Settlements: A Global Challenge for EO Data Processing and Interpretation. *Proc. IEEE 2013*, 101, 570–581. doi:10.1109/JPROC.2012.2189089.

26. Rasaei, Z.; Bogaert, P. Spatial filtering and Bayesian data fusion for mapping soil properties: A case study combining legacy and remotely sensed data in Iran. *Geoderma 2019*, 344, 50–62. doi:10.1016/j.geoderma.2019.02.031.

27. Puttinaovarat, S.; Horkaew, P. Urban areas extraction from multi sensor data based on machine learning and data fusion. *Pattern Recognit. Image Anal. 2017*, 27, 326–337.

28. Zhang, J. Multi-source remote sensing data fusion: status and trends. *Int. J. Image Data Fusion 2010*, 1, 5–24.

29. Pohl, C.; van Genderen, J. Multisensor image fusion in remote sensing: concepts, methods and applications. *Int. J. Remote Sens. 1998*, 19, 823–854.

30. Wenbo, W.; Jing, Y.; Tingjun, K. Study of remote sensing image fusion and its application in image classification. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 2008*, 37, 1141–1146.

31. Ghassemian, H. A review of remote sensing image fusion methods. *Inf. Fusion 2016*, 32, 75–89.

32. Tu, T.M.; Su, S.C.; Shyu, H.C.; Huang, P.S. A new look at IHS-like image fusion methods. *Inf. Fusion 2001*, 2, 177–186. doi:10.1016/S1566-2535(01)00036-7.

33. Gillespie, A.R.; Kahle, A.B.; Walker, R.E. Color enhancement of highly correlated images. II. Channel ratio and chromaticity transformation techniques. *Remote Sens. Environ. 1987*, 22, 343–365. doi:10.1016/0034-4257(87)90088-5.

34. Gonzalez-Audicana, M.; Saleta, J.L.; Catalan, R.G.; Garcia, R. Fusion of multispectral and panchromatic images using improved IHS and PCA mergers based on wavelet decomposition. *IEEE Trans. Geosci. Remote Sens. 2004*, 42, 1291–1299.

35. Palsson, F.; Sveinsson, J.R.; Ulfarsson, M.O.; Benediktsson, J.A. Model-Based Fusion of Multi- and Hyperspectral Images Using PCA and Wavelets. *IEEE Trans. Geosci. Remote Sens. 2015*, 53, 2652–2663.

36. Laben, C.A.; Brower, B.V. Process for Enhancing the Spatial Resolution of Multispectral Imagery Using Pan-Sharpening. U.S. Patent 6,011,875, 4 January 2000.

37. Aiazzi, B.; Baronti, S.; Selva, M. Improving Component Substitution Pan-sharpening Through Multivariate Regression of MS + Pan Data. *Trans. Geosci. Remote Sens. 2007*, 45, 3230–3239.

38. Chen, C.M.; Hepner, G.; Forster, R. Fusion of hyperspectral and radar data using the IHS transformation to enhance urban surface features. *ISPRS J. Photogramm. Remote Sens. 2003*, 58, 19–30.

39. Shao, Z.; Wu, W.; Guo, S. IHS-GTF: A Fusion Method for Optical and Synthetic Aperture Radar Data. *Remote Sens. 2020*, 12, 2796.

40. Zhijun Wang.; Ziou, D.; Armenakis, C.; Li, D.; Qingquan Li. A comparative analysis of image fusion methods. *IEEE Trans. Geosci. Remote Sens. 2005*, 43, 1391–1402.

41. Singh, D.; Garg, D.; Pannu, H.S. Efficient Landsat image fusion using fuzzy and stationary discrete wavelet transform. *Imaging Sci. J. 2017*, 65, 108–114. doi:10.1080/13682199.2017.1289629.

42. Kwarteng, P.; Chavez, A. Extracting spectral contrast in Landsat Thematic Mapper image data using selective principal component analysis. *Photogramm. Eng. Remote Sens. 1989*, 55, 1.

43. Ma, J.; Gong, M.; Zhou, Z. Wavelet Fusion on Ratio Images for Change Detection in SAR Images. *IEEE Geosci. Remote Sens. Lett. 2012*, 9, 1112–1126.

44. Salentinig, A.; Gamba, P. A General Framework for Urban Area Extraction Exploiting Multiresolution Data Fusion. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2016*, 9, 2009–2018.

45. Rodriguez-Galiano, V.F.; Ghimire, B.; Rogan, J.; Chica-Olmo, M.; Rigol-Sanchez, J.P. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens. 2012*, 67, 93–104.

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

47. Foody, G.M. Thematic map comparison. *Photogramm. Eng. Remote Sens. 2004*, 70, 627–633.

48. Feng, W.; Dauphin, G.; Huang, W.; Quan, Y.; Liao, W. New margin-based subsampling iterative technique in modified random forests for classification. *Knowl. Based Syst. 2019*, 182, 104845.

49. Feng, W.; Huang, W.; Dauphin, G.; Xia, J.; Quan, Y.; Ye, H.; Dong, Y. Ensemble Margin Based Semi-Supervised Random Forest for the Classification of Hyperspectral Image with Limited Training Data. In *Proceedings of the 2019 IEEE International Geoscience and Remote Sensing Symposium*, Yokohama, Japan, 28 July–2 August 2019; pp. 971–974. doi:10.1109/IGARSS.2019.8898415.
50. Feng, W.; Huang, W.; Bao, W. Imbalanced Hyperspectral Image Classification With an Adaptive Ensemble Method Based on SMOTE and Rotation Forest With Differentiated Sampling Rates. *IEEE Geosci. Remote Sens. Lett.* 2019, 16, 1879–1883.

51. Quan, Y.; Zhong, X.; Feng, W.; Dauphin, G.; Gao, L.; Xing, M. A Novel Feature Extension Method for the Forest Disaster Monitoring Using Multispectral Data. *Remote Sens.* 2020, 12, 2261.

52. Qiang Li, Wei Feng, and Yinghui Quan. *J. Infect.* 2020, 80, 469-496.

53. Tian, S.; Zhang, X.; Tian, J.; Sun, Q. Random forest classification of wetland landcovers from multi-sensor data in the arid region of Xinjiang, China. *Remote Sens.* 2016, 8, 954.

54. Xu, Z.; Chen, J.; Xia, J.; Du, P.; Zheng, H.; Gan, L. Multisource Earth Observation Data for Land-Cover Classification Using Random Forest. *IEEE Geosci. Remote Sens. Lett.* 2018, 15, 789–793.

55. Wu, Q.; Zhong, R.; Zhao, W.; Song, K.; Du, L. Land-cover classification using GF-2 images and airborne lidar data based on Random Forest. *Geosci. Remote Sens. Lett.* 2019, 40, 2410–2426.

56. Dietterich, T.G. Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms. *Neural Comput.* 1998, 10, 1895–1923. doi:10.1162/089976698300017197.

57. Zhang, Q.; Xu, L.; Jia, J. 100+ Times Faster Weighted Median Filter (WMF). In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 2830–2837.

58. Du, J.; Li, W.; Xiao, B.; Nawaz, Q. Union Laplacian pyramid with multiple features for medical image fusion. *Neurocomputing* 2016, 194, 326–339. doi:10.1016/j.neucom.2016.02.047.

59. Fan, Q.; Yang, J.; Hua, G.; Chen, B.; Wipf, D. A Generic Deep Architecture for Single Image Reflection Removal and Image Smoothing. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017.

60. Fang, Y.; Zhang, H.; Mao, Q.; Li, Z. Land cover classification with gf-3 polarimetric synthetic aperture radar data by random forest classifier and fast super-pixel segmentation. *Sensors* 2018, 18, 2014.

61. Thanh Noi, P.; Kappas, M. Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using Sentinel-2 imagery. *Sensors* 2018, 18, 18.

62. Amani, M.; Ghorbanian, A.; Ahmadi, S.A.; Kakooei, M.; Moghimi, A; Mirmazloumi, S.M.; Moghaddam, S.H.A.; Mahdavi, S.; Ghalremanlooo, M.; Farsian, S.; Wu, Q.; Brisco, B. Google Earth Engine Cloud Computing Platform for Remote Sensing Big Data Applications: A Comprehensive Review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2020, 13, 5326–5350. doi:10.1109/JSTARS.2020.3021052.

63. Stromann, O.; Nascetti, A.; Yousif, O.; Ban, Y. Dimensionality Reduction and Feature Selection for Object-Based Land Cover Classification based on Sentinel-1 and Sentinel-2 Time Series Using Google Earth Engine. *Remote Sens.* 2020, 12, 76.

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.