An Effective Multimodel Fusion Method for SAR and Optical Remote Sensing Images

Wenmei Li, *Member, IEEE*, Jiaqi Wu, Qing Liu, Yu Zhang, Bin Cui, Yan Jia, and Guan Gui, *Senior Member, IEEE*

***Abstract*—Remote sensing images acquired by different sensors exhibit different characteristics due to their distinct imaging mechanisms. The fusion of Synthetic Aperture Radar (SAR) and optical remote sensing images is valuable for specific remote sensing image applications, as it enables the extraction of texture features from SAR images while preserving the spectral information of optical images. Several existing fusion approaches have been proposed in recent years, including the Nonsubsampled Shearlet Transform Pulse Coupled Neural Network (NSST-PCNN), which is a typical and effective fusion method. However, it suffers from the inconsistency in regional edge information. To address this issue, we propose a new method called MS-NSST-PCNN for multi-model fusion of SAR and optical remote sensing images. This method incorporates the multiScale morphological gradient (MSMG) into NSST-PCNN to detect edges and enhance the utilization of edge characteristics. The fusion results of two polarization modes, VV and VH are evaluated in combination with existing methods, using image fusion accuracy and visual interpretation criteria. The results demonstrate that for Sentinel 1 and Landsat 8 OLI image fusion the proposed MS-NSST-PCNN method achieves higher correlation coefficients and lower spectral distortion with VV polarization compared to traditional methods in two study areas. Moreover, the proposed method also exhibits better performance for GF3 and GF2 images with higher spatial resolution. In subsequent applications of land classification in urban and rural scenarios, the fusion results of the proposed method achieve higher accuracy than those of other fusion methods or source images applied directly.**

Manuscript received 24 July 2022; revised 30 March 2023; accepted 18 June 2023. Date of publication 21 June 2023; date of current version 7 July 2023. This work was supported in part by the Key Laboratory of Land Satellite Remote Sensing Application, Ministry of Natural Resources of the People's Republic of China under Grant KLSMNR-K202201 and Grant KLSMNR-202305, in part by the National Natural Science Foundation of China under Grant 42071414, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20191384, in part by the Open Fund of State Key Laboratory of Remote Sensing Science under Grant OFSLRSS202202, in part by the China Postdoctoral Science Foundation under Grant 2019M661896, and in part by the Qinglan Project in Jiangsu Province. *(Corresponding author: Guan Gui.)*

Wenmei Li, Bin Cui, and Yan Jia are with the School of Geographic and Biologic information, Nanjing University of Posts and Telecommunications, Nanjing 210023, China, and also with the Health Big Data Analysis and Location Services Engineering Laboratory of Jiangsu Province, Nanjing University of Posts and Telecommunications, Nanjing 210023, China (e-mail: liwm@njupt.edu.cn; bincui@njupt.edu.cn; jiayan@njupt.edu.cn).

Jiaqi Wu, Qing Liu, and Yu Zhang are with the School of Geographic and Biologic information, Nanjing University of Posts and Telecommunications, Nanjing 210023, China (e-mail: 1326464135@qq.com; 1021173512@njupt.edu.cn; 1021173513@njupt.edu.cn).

Guan Gui is with the College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China (e-mail: guiguan@njupt.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2023.3288143

***Index Terms*—Multiscale morphological gradient (MSMG), nonsubsampled shearlet transform (NSST), optical data fusion, pulse-coupled neural network (PCNN), synthetic aperture radar (SAR).**

I. Introduction

With the development of remote sensing technology, new satellite sensors continue to emerge, making it possible to obtain multisensor, multitemporal, and multiresolution remote sensing image data sequences in the same area. Images acquired by different satellite sensors usually have different characteristics and application scenarios due to different imaging principles, such as visible light sensors and synthetic aperture radar (SAR) sensors [1], [2]. Visible light sensors can obtain black-and-white and color images of objects on Earth, which are easy to be interpreted by the human eyes. Meanwhile, due to its unique imaging principle and the ability to penetrate terrestrial objects to a certain extent, SAR sensors can acquire surface information all-weather. To fully understand the characteristics and changes of ground objects, it is necessary to comprehensively utilize the respective advantages of different satellite sensors for remote sensing applications [3], [4], [5], and multisatellite sensor data fusion came into being. The fusion of SAR and optical remote sensing image data is one of the most typical and important research topics.

According to the principle of SAR and visible light image fusion [6], the existing fusion algorithms can be divided into two categories: 1) component substitution (CS); and 2) multiresolution analysis (MRA). The principle of CS [7], [8] can also be called a fusion method in the spatial domain, which fuses visible light and SAR images through its spatial transformation method. The intensity information and spectral acknowledge are first separated by projecting the visible light image into a new feature space, and then the intensity information is replaced by the SAR image. These methods can improve the spatial resolution of the fused images, and the fusion enhances the display of spatial details. However, the fusion results have poor spectral retention ability and are prone to spectral distortion. The representative methods of this category mainly include IHS transform, PCA transform, and Brovey transform. The IHS transform [9] divides the visible light image into three components: 1) intensity (I), 2) hue (H), and 3) saturation (S), and uses the grayscale stretched SAR image to replace the "I" component. This method does not consider the difference between the dynamic gray range of SAR and visible light images, which causes spectral distortion and affects the accuracy of object interpretation. Hong et al. [10]

combined the characteristics of panchromatic images and SAR images to establish feature selection criteria to adjust the luminance components of multispectral images, which avoided the mixing of redundant information.

The MRA method [11], [12] is a kind of fusion algorithm based on the transfer domain, which decomposes and fuses the multiscale coefficients of the source image, and finally performs inverse transformation to obtain the fused image. This type of method can extract the edge and texture of the image at different scales, while retaining more spectral information, and the fused image quality is higher. However, the fusion effect and operation time of this method mostly depend on the selection of multiscale decomposition approaches and fusion criteria. The representative methods of multiscale analysis methods mainly include discrete wavelet transform (DWT), Laplace pyramid, contourlet transform, and curvelet transform. The contourlet transform proposed by Do and Vetterli [13] in 2002 achieved multiscale and multidirectional decomposition of images. On this basis, the nonsubsampled contourlet transform (NSCT) was proposed in 2006 [14], which enhanced the translation invariance and orientation selectivity and avoided the pseudo-Gibbs phenomenon. In 2007, Guo and Labate [15] proposed shearlet transform with better localization characteristics and more direction-sensitive based on the idea of affine system. Similarly, the nonsubsampled shearlet transform (NSST) also greatly enhanced the direction selectivity and translation invariance. Based on these methods, many fusion methods have been proposed in the transfer domain [16], [17], [18]. However, after multiscale analysis, the low-frequency components cannot be sparsely represented and cannot effectively reflect salient features.

In recent years, more researchers have focused on the fusion method combining spatial domain and transform domain [19], [20], [21], [22], [23], [24]. Jiahuan et al. [25] proposed an image fusion method based on NSCT transform and adaptive gamma correction, changed the low-frequency fusion rule to adaptive gamma correction, and achieved better fusion results. Liu et al. [26] combined discrete wavelet and multiscale morphological transform in the IHS color space to fuse PAN and multispectral images. Liu et al. [27] proposed a convenient task-inspired multiscale nonlocal-attention network (MNAN) for remote sensing image fusion which focuses more on enhancing the multiscale targets in in the scene when improving the resolution of the fused images.

Sparse representation is another representative image fusion algorithm based on the transform domain [28], it has been a hot research topic in the field of signal processing in recent years. And, it has exerted its advantages in many fields, such as target tracking, pattern recognition, image denoising, audio processing, etc., [29], [30], [31], [32]. Based on this, researchers have introduced the principle of sparse representation into the field of image fusion, and achieved a series of research results [33], [34], [35]. Liu et al. [36] combined IHS transform and joint nonnegative sparse representation method to obtain fused images with higher image quality. Zhang et al. [37] presented a fusion framework to integrate the information from MS and SAR images based on the Laplacian pyramid and sparse representation theory.

Among the fusion methods based on sparse representation, NSST-PCNN has attracted extensive attention for its ability to preserve details and energy without any training process. NSST is simple in operation and there is no requirement for the directionality of the image decomposition. Consequently, NSST is utilized in the image fusion field, and a series of fusion methods were proposed based on NSST. However, the performance of NSST-based fusion methods is subject to effective activity level metrics and fusion rules. Pulse coupled neural network (PCNN) is a simplified neural network model proposed by Eckhorn, and the signal form and processing mechanism are more in line with the physiological basis of the human visual nervous system [38]. It is usually applied as an effective activity level measurement tool for image fusion. But it could not determine the parameters adaptively, which limits the performance of PCNN-based fusion methods. In this context, NSST-PCNN is proposed to deal with image fusion problems, especially medical image fusion. Gai et al. [39] proposed an IQPSO-based PCNN in the NSST domain for medical image fusion, and their experiments showed that the method outperforms the existing state-of-the-art methods in terms of visual performance and objective evaluation. The NSST-PCNN performs well in medical image fusion, but when it is used for the fusion of remote sensing images with more complex textures and colors, it often cannot identify small edge information, resulting in discontinuous edges of the fused image.

Therefore, this article will continue to study the fusion method for SAR and optical images by combining the spatial domain and the transform domain, selecting NSST-PCNN as the basic method, and improving its performance by introducing features that can better express and preserve edge information. The accuracy evaluation of the fusion results will be carried out from two aspects: 1) the evaluation of various indexes of the fusion image; and 2) the classification accuracy of ground objects based on multiscale segmentation. The main contribution of this article could be divided into three parts as follows.

1) A method called MS-NSST-PCNN is proposed for SAR and optical image fusion. It is built by introducing MSMG and IHS transforms to traditional NSST-PCNN methods to deal with multimodel fusion issues. And the fusion results not only retain the spectral characteristics of optical images but also inherit the texture characteristics of SAR images.
2) The MS-NSST-PCNN method proposed in this article can better combine the texture features of SAR and optical images and has better ability to distinguish object edges and small objects, which verifies the effectiveness of the proposed method in object-oriented classification.
3) The fusion method proposed in this article helps to improve the classification accuracy. Compared with the existing fusion methods and the direct application of the original optical image for classification, MS-NSST-PCNN has achieved better classification results in both urban and rural application scenarios.

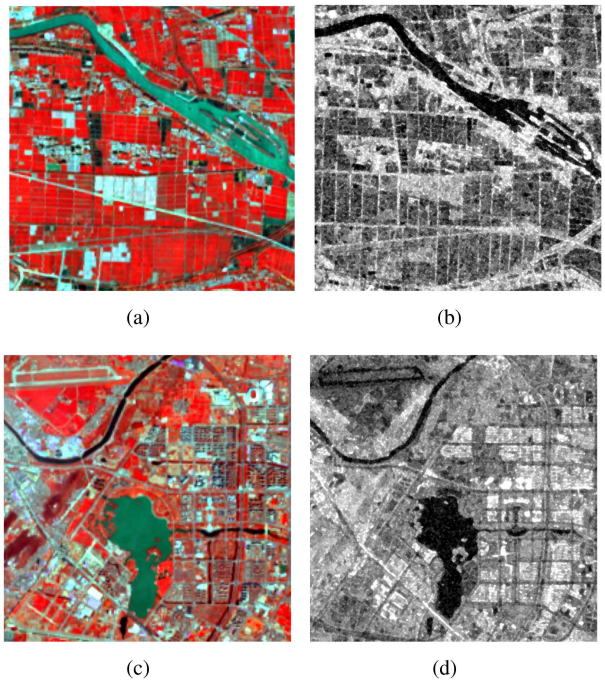

Fig. 1. Image used in fusion. (a) Landsat-8 OLI image (false color composite Red:5, Green:4, Blue:2) of study area A. (b) Sentinel-1 A VH polarization of study area A. (c) Landsat-8 OLI image (false color composite Red:5, Green:4, Blue:2) of test site B. (d) Sentinel-1 A VH polarization of test site B.

II. Dataset and Preprocessing

A. Test Sites

Two typical study areas (Fig. 1), rural and urban, are selected to carry out research on the fusion method of SAR and optical imagery, and ground object classification experiments are carried out based on them. Study area A shown in Fig. 1(a) and (b), lies on Yangbei Street, southeast of Suqian City, Jiangsu Province ($33°47'N \sim 33°51'N, 118°23'E \sim 118°27'E$). The main types of land features are farmland, and land for construction, while the main water body is the Beijing–Hangzhou Canal. Study area B is shown in Fig. 1(c) and (d), and it is located in Yunlong District, Xuzhou ($34°10'N \sim 34°14'N$, $117°14'E \sim 117°21'E$). Green land (park vegetation) and construction land (houses, roads) are the main types. The water bodies are mainly the old Yellow River, Dalongkou Reservoir in Dalong Lake Scenic Area and surrounding rivers.

B. Dataset and Preprocessing

Since the ratio of remote sensing image resolutions used for fusion determines the quality of the fused image to some extent [40]. Sentinel-1 A data in the IW mode with a resolution of 10 m is selected, so that the resolution ratio with the Landsat-8 OLI multispectral data can reach 1:3. And the parameters of the optical and SAR data used in the experiment are shown in Table I.

The optical image preprocessing [41] in this experiment mainly includes three steps: 1) radiation correction, 2) geometric correction, and 3) band selection. Since three bands need to be selected from the multispectral image for color synthesis

TABLE I
Data and Parameters Setting

Data	Type	Parameter
Landsat-8 OLI	Resolution	30m
	Band	2,4,5
	Revisit period	16d
	Acquisition Date	2018-04-10
Sentinel-1A IW GRD SAR	Resolution	10m
	polarization	VV,VH
	Revisit period	6d
	Acquisition Date	2018-04-09

to participate in the fusion, it is necessary to select the band combination that is most conducive to the fusion effect. According to the principle of color synthesis of remote sensing images, objects in false-color composite images are easier to identify. The Landsat-8 OLI includes nine bands, appropriate band selection to identify the ground objects is particularly important. Considering the need for vegetation classification, bands 2, 3, 4, 5, and 6 are selected as alternative bands.

For the fusion algorithm dealing with three bands in multispectral images, the optimal band combination is calculated according to the optimal index formula proposed by Chavez et al. [42]. And the calculation formula is given as

$$O_{IF} = \frac{\sum_{i=1}^{n} S_i}{\sum_{i=1}^{n} \sum_{j=j+1}^{n} |R_{ij}|} \tag{1}$$

where S_i is the standard deviation of $i-$band, R_{ij} is the CC between i-band and j-band, and the optimal O_{IF} parameter of each three-band combination sequence of Landsat-8 multispectral is listed in Table II.

The larger the O_{IF} parameter, the greater the information content of the band combination. Therefore, the combination of band 2, 4, and 5 was selected as the Landsat-8 multispectral data for follow-up research. The preprocessing process of Sentinel-1 A data mainly includes filtering, geocoding, and radiometric scaling. The GRD data used in this article had already undergone multilook processing before it was released. Then refined Lee filter was applied for speckle filtering, and the last step is to perform geometric registration and cropping of SAR and optical images to meet the fusion requirements. Finally, SAR and optical image pairs with a size of 600×600 were produced for subsequent experiments.

III. Our Proposed Methods

A. NSST

Shearlet transform [43] is a multiscale geometric analysis method based on the construction of a 2-D affine system with compound dilation, which was introduced to image processing by Easley et al. [44] in 2007. The affine system can be expressed in 2-D space as

$$M_{AB}(\Psi) = \{\Psi_{j,l,k}(x) = |detA|^{j/2}\Psi(B^l A^j x - k) : j, l \in Z; k \in Z^2)\} \tag{2}$$

TABLE II
LANDSAT-8 BAND COMBINATION AND EVALUATION INDEX

Band Combination	Standard Deviation	Correlation Coefficient	O_{IF} Parameter
456	6.55595	2.7447	2.38858
345	7.92911	2.7658	2.86684
245	8.46776	2.75908	3.06905

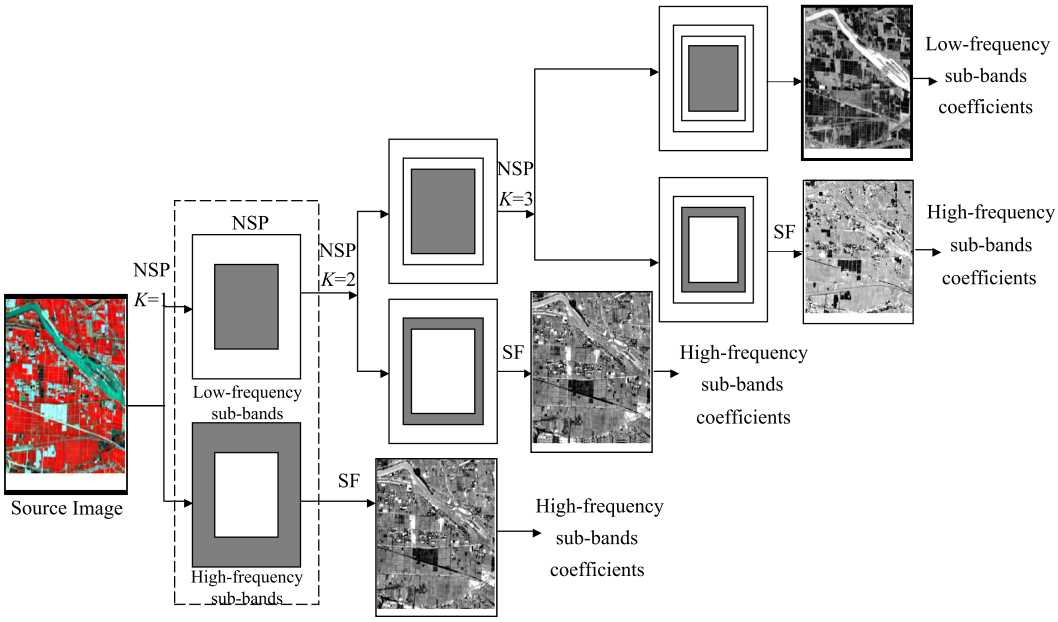

Fig. 2. NSST model.

where $\Psi \in L^2(R^2)$, if $\forall f \in \varphi$ satisfies the Parseval compact frame, the elements of the affine system are called synthetic wavelets. A represents an anisotropic expansion matrix, which is related to the transformation scale. B represents a shearlet matrix, which is related to geometric scales, such as rotation and shearing. A and B are 2×2 invertible matrices, and $|\det B| = 1, j, l, k$ represents the scale parameter, direction parameter, and translation parameter of the decomposition, respectively. When $A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, synthetic wavelet turns to a special form which is shearlet.

NSST also eliminates the pseudo-Gibbs phenomenon of shearlet waves during image reconstruction, and it can obtain sparse representations of images in different orientations and scales. The decomposition process of NSST is shown in Fig. 2, and the whole process consists of two steps, multiscale decomposition through nonsubsampled pyramid filter (NSP) and directional decomposition based on shearlet filter (SF). NSP could decompose the source image into low-frequency subbands (LF1) and high-frequency subbands (HF1), then the LF1 could continue to be decomposed into low-frequency subbands (LF2) and high-frequency subbands (HF2). The HF1 and HF2 are reserved, and the LF2 continue to be decomposed. This cycle continues until the set number of decompositions (3) is reached or other set requirements are met. Then HF1, HF2, and HF3 ($K = 3$) are filtered by SF to obtain high-frequency subbands coefficients, respectively. And LF3 ($K = 3$) is retained as low-frequency subbands coefficients. It is worth noting that the number of decompositions can be set according to the actual decomposition needs. In Fig. 2, $K = 3$ is taken as an example.

By decomposing with the NSP, the low-frequency subbands describing most of the information and the high-frequency subbands that describe the detailed information can be obtained. After K times of NSP decomposition, K high-frequency subbands and one low-frequency subband that are consistent with the size of the source image will be generated. NSST [45] uses the Meyer's window function with a variable aspect ratio to form an improved SF for the directional decomposition of high-frequency components. Meanwhile, the directional subbands are obtained by the convolution operation of the new shearlet filter in which high-frequency subbands are transferred from the pseudo-polarization grid network system to the Cartesian coordinate system. In this way, 2-D convolution can be realized by inverse Fourier transform, avoiding the translation invariance of shearlet coefficients obtained in the down sampling step. Compared with the wavelet structure, shearlet has anisotropy, and is more sensitive to the orientation information of the image and has the best sparse representation ability.

B. PCNN

PCNN [46] is constructed based on iterative computation and requires no training process, which distinguishes it from most existing artificial neural networks. PCNN models applied

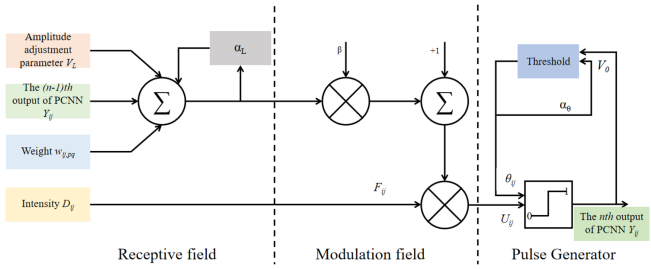

Fig. 3. PCNN model.

in image processing tasks are generally single-layer networks with 2-D array inputs. There is a correspondence between input image pixels and PCNN neurons, so the number of neurons is equal to the number of pixels. Each neuron is connected to its neighbors for information transfer and coupling. It is composed of multiple interconnected neurons including receptive fields, modulation fields, and impulse generators, and is continuously adaptively adjusted through self-learning and communication between neurons. However, the standard PCNN model has many parameters that need to be set through a lot of experiments or experience, such as link strength, attenuation coefficient, and so on. To reduce the complexity of the algorithm as much as possible, Blasch [47] proposed a simplified PCNN model, and the mathematical expression is illustrated as

$$\begin{cases} F_{ij}(n) = D_{ij} \\ L_{ij} = L_{ij}(n-1) \times \exp(-\alpha_L) + V_L \Sigma_{pq} \omega_{ij,pq} Y_{pq}(n-1) \\ U_{ij} = F_{ij}(1 + \beta L_{ij}(n)) \\ \theta_{ij}(n) = \theta_{ij}(n-1) \times \exp(-\alpha_\theta) + V_\theta Y_{ij}(n-1) \\ Y_{ij}(n) = sgn(U_{ij}(n) - \theta_{ij}(n)) \end{cases} \tag{3}$$

where $F_{ij}(n), L_{ij}(n), U_{ij}(n)$, and $\theta_{ij}(n)$ are the feedback input, link input, internal activity, and dynamic threshold of the neuron at (i, j) in the nth iteration, respectively. $F_{ij}(n)$ is fixed as the input image intensity throughout the iteration process. α_L represents the time decay parameter, V_L is the parameter of the amplitude adjustment, β is the link strength, and $Y_{ij}(n)$ represents the output result of the nth iteration of PCNN. If $U_{ij}(n)$ is greater than or equal to the dynamic threshold $\theta_{ij}(n)$, the neuron is activated. The simplified PCNN model structure is shown in Fig 3, it also contains three parts, namely, the receptive fields, modulation fields, and impulse generators. In receptive field, amplitude adjustment parameter V_L, the last iteration output result of PCNN $Y_{ij}(n-1)$, and the wight $w_{ij,pq}$ are comprehensively considered. During the modulation field, the link input $L_{ij}(n)$, link strength β, and the information extracted in the receptive field are comprehensively calculated to obtain $U_{ij}(n)$. In the pulse generator, $U_{ij}(n)$, and dynamic threshold $\theta_{ij}(n)$ are applied to generate the output $Y_{ij}(n)$.

C. MS-NSST-PCNN

The schematic diagram of the MS-NSST-PCNN model is shown in Fig 4. The first step is to apply the IHS transform to extract the intensity components from the three-bands pseudo color optical image. And then SAR image is used to extract high-frequency $H_S^{l,k}$ and low-frequency L_S parts. The low-frequency parts of L_V and L_S are then used to generate low-frequency L_p through WSEML algorithm. And the high frequency of H_V and H_S is tackled with PCNN and MSMG techniques to obtain high-frequency $H_p^{l,k}$. After the fusion frequency is completed, the inverse IHS transform is applied to obtain the fused image.

1) Low-Frequency Fusion: For the fusion rule of low-frequency coefficients, using the WESML (weighted sum of eight-neighborhood-based modified Laplacian) instead of the traditional average low-frequency fusion rule can reduce loss

$$E_P(i,j) = \Sigma_{m=-r}^{r} \Sigma_{n=-r}^{r} \omega(m+r+1, n+r+1) \cdot DL_P(i+m, j+n)^2 \tag{4}$$

$$\text{WSEML}_P(i,j) = \Sigma_{m=-r}^{r} \Sigma_{n=-r}^{r} \omega(m+r+1, n+r+1) \cdot \text{EML}_P(i+m, j+n)^2 \tag{5}$$

$$\begin{aligned} \text{EML}_P(i,j) = &|2DL_P(i,j) - DL_p(i-1,j) - DL_P(i+1,j)| \\ &+ |2DL_P(i,j) - DL_p(i,j-1) - DL_P(i+1,j)| \\ &+ \frac{1}{\sqrt{2}} |2DL_P(i,j) - DL_p(i-1,j-1) - DL_P(i+1,j+1)| \\ &+ \frac{1}{\sqrt{2}} |2DL_P(i,j) - DL_p(i-1,j+1) - DL_P(i+1,j-1)| \end{aligned} \tag{6}$$

where $p \in \{V, S\}$ and W is a weighting matrix size of $(2r+1) \times (2r+1)$. For each element in W, its value is set as $2^{(2r-d)}$, where d is its four-neighborhood distance to the center. The activity level of the low-frequency subband is defined as the multiplication of E_p and WSEML_P. The final low frequency subband coefficients can be calculated as follows:

$$L_p(i,j) = \begin{cases} L_V(i,j), \text{if } EW_V(i,j) \geq EW_S(i,j) \\ L_S(i,j), \text{otherwise} \end{cases} \tag{7}$$

where $EW_V(i,j) = E_V(i,j) \cdot \text{WSEML}_V(i,j)$, $EW_S(i,j) = E_S(i,j) \cdot \text{WSEML}_S(i,j)$.

2) High-Frequency Fusion: If the gray value of certain pixels in the image changes rapidly and the gradient value is large, it can be judged as an edge point. MSMG is an efficient operator that extracts gradient information through intensity difference of the pixels at the boundary part in the original image. It is an efficient method mostly used for edge detection and image segmentation in medical images [48], [49]. Therefore, we propose a method based on the conventional NSST-PCNN fusion method and aided by MSMG to improve the performance of multimodel fusion. The single-scale morphological gradient of the tth level $G_t(x,y)$ can be obtained by the dilation–corrosion operator of the source image $f(x,y)$

$$G_t(x,y) = (f(x,y) \bigoplus g_t(x,y)) - (f(x,y) \bigodot g_t(x,y)) \tag{8}$$

where $\bigoplus$ and $\bigodot$ represent morphological dilation and erosion operations, respectively. $g_t(x,y)$ denotes the structural elements of the tth level. From multiscale structuring elements and gradient features, the MSMG can be obtained by computing the

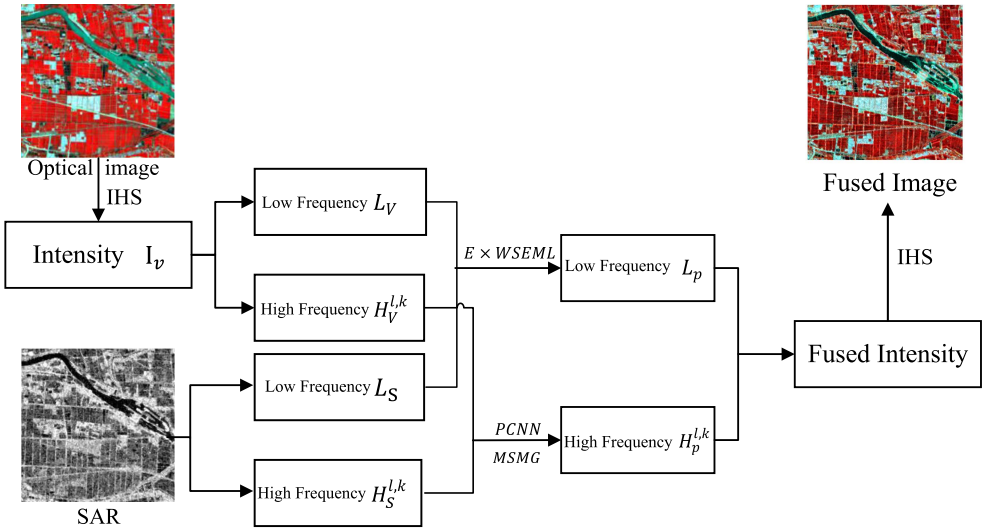

Fig. 4. MS-NSST-PCNN.

weighted sum of all scale gradients

$$M(x,y) = \sum_{t=1}^{N} W_t \cdot G_t(x,y) \tag{9}$$

where W_t represents the weight of the gradient in the tth scale, and it can be represented as $W_t = \frac{1}{(2t+1)}$. The MSMG of the two input images is applied to replace the link strength β of the PCNN in (3) to strengthen the edge information extraction.

IV. Experimental Results

A. Fusion Results

1) Evaluation Standard: After the fusion process of SAR and optical images is completed, the quality of fusion results needs to be evaluated to ensure that the multimodel fusion quality meets the requirements for subsequent experiments. According to the image fusion quality evaluation standard, the fusion reconstruction results can be evaluated with the following parameters.

- *Average value (μ)* reflects the average reflection intensity of the ground objects in remote sensing images, and M and N represent the number of rows and columns of remote sensing images, respectively

$$\mu = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} P(i,j). \tag{10}$$

- *Information entropy (E):* It is a statistical form of features, which is an indicator of the richness of information in an image. The unary gray entropy of a gray-scale image is defined as

$$E = -\sum_{i=0}^{L-1} P_i \log P_i \tag{11}$$

where P_i represents the proportion of pixels whose gray value is i in the image, and $L-1$ is the maximum gray value of the image.

- *Standard deviation (σ):* The standard deviation reflects the discrete degree of the gray value set of images, and to a certain extent, it can also be used to evaluate the contrast of the images. μ is the average value, and $P(i,j)$ represents the pixel value of the ith row and jth column in remote sensing images

$$\sigma = \sqrt{\frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} ((P(i,j) - \mu))}. \tag{12}$$

- *Degree of distortion (DD):* The degree of spectral distortion reflects the ability of the fusion image to preserve the spectral information of the source image. The lower the degree of spectral distortion, the lower the degree of spectral distortion caused by fusion

$$DD_{F,S} = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} |F(i,j) - S(i,j)| \tag{13}$$

where $F(i,j)$ and $S(i,j)$ are the pixel gray values at (i,j) in the fused image and source image.

- *Correlation coefficient (CC):* The CC is a statistical indicator to evaluate the linear relationship between the fused image and the source image, reflecting the similarity of spectral features between the two

$$CC_{F,S} = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} [F(i,j) - \mu_F][S(i,j) - \mu_S]}{\sqrt{\sum_{i=1}^{M} \sum_{j=1}^{N} [F(i,j) - \mu_F]^2 [S(i,j) - \mu_S]^2}} \tag{14}$$

where μ_F and μ_S are the average value of the fused image and source image, respectively.

2) Experiment Procedure: To verify the fusion effect of the proposed method MS-NSST-PCNN, a fusion experiment was carried out with Landsat-8 and Sentinel-1 A images in the test site A. Classic methods including the Brovey transform, IHS transform, PCA transform, NSCT transform, and NSST-PAPCNN were used for comparison. The decomposition scale of NSCT and NSST transforms is unified to 4, the directional filter

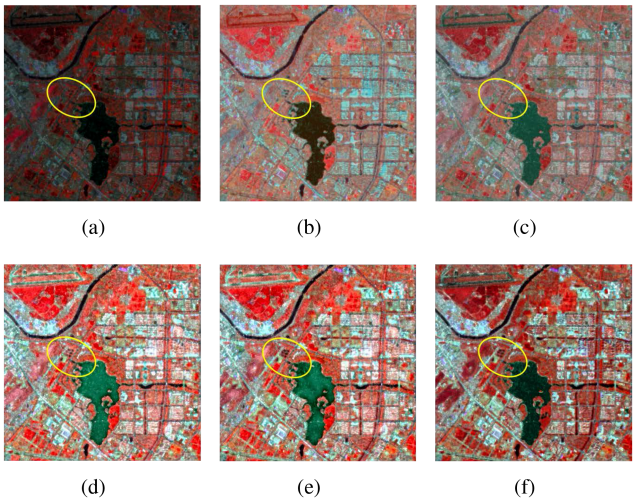

Fig. 5. Fusion results of study area A. (a) Brovey. (b) PCA. (c) IHS. (d) NSCT. (e) NSST-PAPCNN. (f) MS-NSST-PCNN.

TABLE III
VV POLARIZATION FUSION RESULTS OF THE TEST SITE A

Method	μ	E	σ	DD	CC
Brovey	44.12	6.78	22.57	62.41	0.79
PCA	118.61	7.35	23.36	39.64	0.8
IHS	116.74	7.35	30.02	32.63	0.78
NSCT	129.37	7.61	55.45	42.08	0.84
NSSTPAPCNN	133.46	7.62	55.81	40.56	0.85
Proposed	95.79	7.76	55	30.68	0.89

TABLE IV
VH POLARIZATION FUSION RESULTS OF THE TEST SITE A

Method	μ	E	σ	DD	CC
Brovey	51.22	6.93	23.09	55.97	0.84
PCA	136.31	7.42	24.86	38.58	0.83
IHS	123.14	7.34	29.98	32.63	0.8
NSCT	136.34	7.68	52.92	41.82	0.86
NSSTPAPCNN	140.9	7.64	54.07	38.42	0.86
Proposed	99.29	7.78	53.44	27.22	0.92

TABLE V
VV POLARIZATION FUSION RESULTS OF THE TEST SITE B

Method	μ	E	σ	DD	CC
Brovey	47.91	6.7	19.72	74.96	0.68
PCA	117.07	7.3	29.47	33.75	0.7
IHS	122.04	7.26	29.43	34.76	0.68
NSCT	136.99	7.72	47.37	34.88	0.78
NSSTPAPCNN	143.36	7.73	47.17	33.34	0.79
Proposed	110.88	7.79	47.46	26.68	0.87

TABLE VI
VH POLARIZATION FUSION RESULTS OF THE TEST SITE B

Method	μ	E	σ	DD	CC
Brovey	46.34	6.61	18.53	76.26	0.73
PCA	125.63	7.34	30.43	32.21	0.71
IHS	115.96	7.16	26.89	32.24	0.7
NSCT	134.35	7.74	45.93	28.84	0.83
NSSTPAPCNN	140.45	7.76	45.04	27.75	0.83
Proposed	110.68	7.75	47.04	24	0.9

of NSCT is "pkva" and the multiscale decomposer is "9–7." The pyramid filter of NSST-PAPCNN is "maxflat." The images of study areas will also be evaluated in a fusion effect comparison experiment based on the selection of SAR image polarization mode, as the basis for the next step of ground object classification application.

3) Fusion Results Evaluation: The Brovey transform [Fig. 5(a)] performs a ratio operation on the gray value of the pixels of the fused image, so the gray difference between SAR data and multispectral data results in a large color difference between the fused image and the optical data. Although the PCA transform [Fig. 5(b)] has improved the preservation of spectral features, the boundary of the objects is blurred, which makes it the worst texture effect among all methods. In IHS transform [Fig. 5(c)], the "I" component is replaced by the texture of the SAR image. It can also be seen from the figure that the distinction of the ground objects is not obvious enough in fused image. The unavoidable pseudo-Gibbs phenomenon in NSCT leads to low edge preservation in fused images. The comparison of the yellow circles in Fig. 5(d) and (e) clearly shows that the description of the road edge is more ambiguous than that of NSST-PAPCNN. The MS-NSST-PCNN [Fig. 5(f)] method proposed in this article, owing to the introduction of MSMG, can process the edge of objects more clearly, with higher image contrast and stronger readability.

From the comparison between Tables III–VI, Brovey, PCA, and IHS are unable to fully meet the requirements of fusion experiments from quantitative aspect. Compared with the first three methods, the NSCT transform has gained higher standard deviation and CC. However, when faced with the application scenario of farmland in study area A, the pseudo-Gibbs phenomenon leads to spectral distortion, due to a large number of finely fragmented plots, showing a fusion result with a high spectral distortion index. Due to the introduction of PCNN, the fusion results of NSST-PAPCNN have improved in spectral preservation ability, CC, and obtained the highest average value of 140.45. Compared with NSCT and NSCT-PAPCNN, the proposed method gets a greater correlation between the fusion image and the source image, and the CC reaches the highest 0.901, which better retains the information of the source image. In terms of the evaluation index of spectral distortion, it also achieves the lowest 24.00 with less spectral loss. Meanwhile, we also found that the performance of our method is not superior in all evaluation metrics, such as average value μ and standard deviation σ. It may relate with the MSMG, which focuses more on the preservation or extraction of edge information, but does not pay more attention to the mean or standard deviation.

On the other hand, the change of polarization mode also makes a slight difference in the aspect of evaluation indexes of the fused image. However, when the VV polarimetric SAR data are used as the source image for fusion in the application scenarios of study area A and study area B, the evaluation indicators related to the source image of the six methods all present various degrees of improvement. This phenomenon may relate to the backscatter echo energy, that is the energy obtained in VV is larger than that of HV, which leads to a stronger representation of ground objects

TABLE VII
CLASSIFICATION ACCURACY OF STUDY AREA A

Data Types	Accuracy	Farmland	Construction	Water	Other	OA	KC
Landsat-8 OLI	User	0.84	0.74	0.88	0.84	0.8	0.73
	Producer	0.84	0.85	0.71	0.73		
NSCT	User	0.76	0.78	0.71	0.79	0.77	0.67
	Producer	0.74	0.82	0.69	0.75		
NSST-PAPCNN	User	0.85	0.78	0.88	0.82	0.82	0.75
	Producer	0.85	0.82	0.75	0.86		
Proposed	User	0.97	0.8	0.91	0.82	0.87	0.81
	Producer	0.85	0.93	0.79	0.84		

TABLE VIII
CLASSIFICATION ACCURACY OF STUDY AREA B

Data Types	Accuracy	Vegetation	Construction	Water	OA	KC
Landsat-8 OLI	User	0.86	0.52	0.94	0.71	0.53
	Producer	0.61	0.81	0.89		
NSCT	User	0.94	0.67	0.76	0.79	0.66
	Producer	0.67	0.93	1		
NSST-PAPCNN	User	0.9	0.79	0.82	0.8	0.65
	Producer	0.67	0.87	1		
Proposed	User	0.92	0.83	0.94	0.88	0.76
	Producer	0.77	0.94	0.94		

in the VV polarimetric channel. Therefore, in the following part of the ground object classification experiments, the VV polarization NSCT, NSST-PAPCNN, and MS-NSST-PCNN methods with better fusion effects are selected for comparison, and the ground object discrimination ability of the fusion results of each method is verified.

B. *Land Classification Based on Multiscale Segmentation*

The random forest classifier is selected to perform ground object classification experiments with the fused images after multiscale segmentation. The scale is set to 30, the shape factor is set to 0.6, the compactness factor is set to 0.5, and the weight of each band is set to 1. In terms of verification data, point vector data are randomly generated in the study area, and then the actual types of ground objects in the verification points are counted according to the visual interpretation of the image and the Google Earth satellite image. Lastly, the results of the ground object classification are evaluated by calculating the confusion matrix.

From Table VII and Fig. 6, it is found that the NSCT transform [Fig. 6(a)] classification image has negative optimization compared with the Landsat-8 source image classification, which may be caused by the defects in its antidistortion ability. The classification result of NSST-PAPCNN [Fig. 6(c)] misclassifies some other types of ground objects as water bodies, resulting in unsatisfactory classification accuracy of water bodies and other types of ground objects. The proposed MS-NSST-PCNN method [Fig. 6(d)] improves the classification effect for various ground objects, especially for the road details in construction areas. Meanwhile, alternating farmland is also the closest to the actual situation. Due to Others and Water both appearing black on the fused image, part of Others type of ground objects are misclassified (as shown in the blue circles). In the rural application scenario of the final study area A, the overall accuracy (OA) and Kappa coefficient (KC) of the proposed method are 0.87 and 0.81, respectively. Compared with Landsat-8 source image classification, the overall accuracy and Kappa coefficient have been improved by 8.75% and 10.96%, respectively.

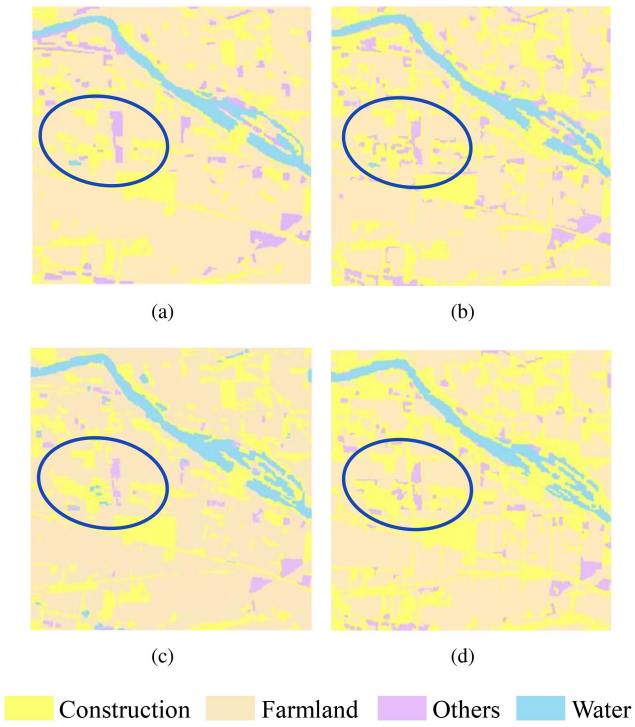

Fig. 6. Random forest classification results of (source)fused images in study area A. (a) Landsat 8 OLI. (b) NSCT. (c) NSST-PAPCNN. (d) MS-NSST-PCNN.

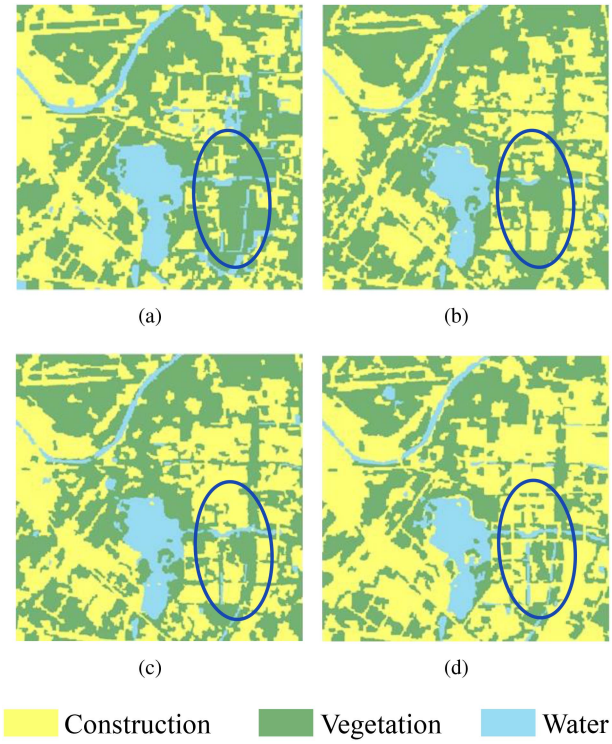

Fig. 7. Random forest classification results of (source)fused images in test site B. (a) Landsat 8 OLI. (b) NSCT. (c) NSST-PAPCNN. (d) MS-NSST-PCNN.

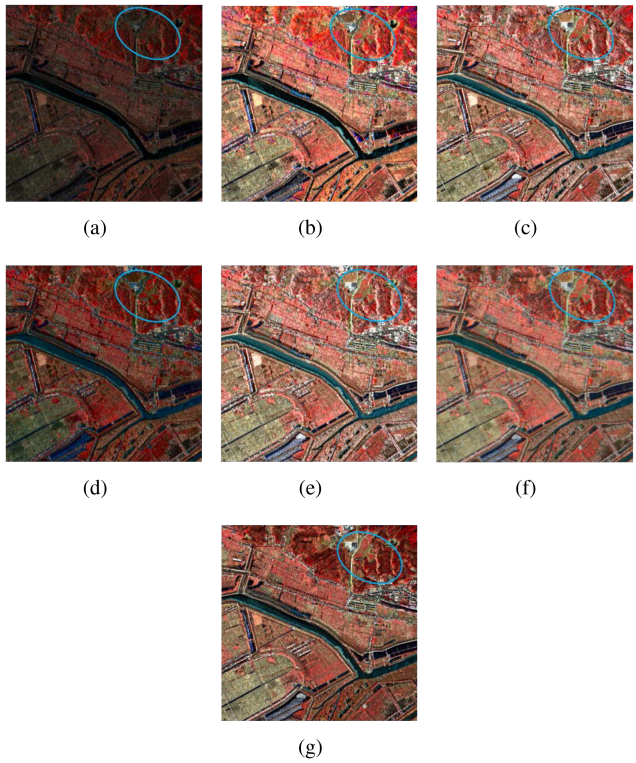

Fig. 8. Fusion results of GF2 and GF3 high-resolution images. (a) Brovey. (b) PCA. (c) IHS. (d) NSCT. (e) NSST-PAPCNN. (f) U2fusion. (g) MS-NSST-PCNN.

From Table VIII and Fig. 7, we can see that the Landsat-8 source image classification [Fig. 7(a)] cannot cope with the mixed ground features of buildings and vegetation, and there are a lot of objects misclassified. Due to the different components contained in different water bodies, different colors appear on the multispectral data, and many shaded areas between buildings are thus classified as water bodies. The classification results of NSCT fusion [Fig. 7(b)] cannot subdivide small river water bodies between buildings. However, the proposed MS-NSST-PCNN method [Fig. 7(d)] can well combine the texture features of SAR and solve the problem of not being able to accurately identify bridges on rivers and distinguish the small river next to the built-up area (as marked by the blue circles). Cases of misclassification have also decreased, and the user accuracy of vegetation and construction is also the highest among the four types of data. In the urban application scenario of study area B, the OA and KC of our method reach 0.88 and 0.76, respectively, which are 23.94% and 43.40% higher than those using the Landsat-8 source image classification alone.

C. High-Spatial Resolution SAR and Optical Image Fusion

With the rapid development of remote sensing techniques in recent years, tremendous progress has been made in spectral resolution, spatial resolution, and temporal resolution. It has formed hyperspectral, high spatial resolution, all-day, all-weather, real-time, and real-time Earth observation capabilities. To verify the feasibility and effectiveness of the fusion method MS-NSST-PCNN on high-spatial resolution SAR and optical images fusion, GF2(multispectral image) with spatial resolution of 3.3 m and GF3(SAR image) with spatial resolution of 3 m are employed for fusion. For high-spatial resolution satellite image fusion experiment, the test site lies in Liuhe district, Nanjing, Jiangsu province ($32°16'N \sim 32°13'N$, $119°1'E \sim 119°5'E$).

TABLE IX
FUSION RESULTS OF GF2 AND GF3

Method	μ	E	σ	DD	CC
Brovey	38.58	6.46	17.69	62.30	0.44
PCA	56.97	7.01	25.96	51.04	0.50
IHS	113.55	7.72	50.37	48.93	0.43
NSCT	115.70	7.76	48.37	37.13	0.62
NSSTPAPCNN	120.88	7.75	48.12	39.73	0.60
Proposed	89.99	7.66	45.63	22.44	0.85
U2fusion	107.11	7.54	35.07	25.711	0.79

The fusion results of GF2 and GF3 images by different methods are shown in Fig. 8, and the specific evaluation result is exhibited in Table IX. From Table IX, we can see that the proposed fusion method MS-NSST-PCNN can perform well in high-spatial resolution optical and SAR images fusion. It is superior to the other five fusion methods in terms of DD and CC, which is consistent with the conclusions mentioned above. In other words, the MS-NSST-PCNN method can retain the spectral information of the original image to the greatest extent, while maintaining the greatest similarity with the multispectral image. Compared with the Brovey's method, the proposed fusion

method in this article reduces the DD by 36.03% and improves the CC by 51.46%. In addition, visually, the proposed fusion method also has a good performance in reflecting complex terrain in high-spatial resolution image fusion, as illustrated in the blue circles on the upper right part in Fig. 8.

Meanwhile, the fusion performance of the deep learning method U2fusion was also compared with that of our multimodel method. From Table IX, we can see that compared with other five fusion methods (Brovey, PCA, IHS, NSCT, NSST-PAPPCNN), U2fusion performs well in high-spatial resolution multimodel remote sensing images fusion, especially DD and CC are significantly improved. However, it is not as good as the proposed fusion method in our article, compared with U2fusion, the proposed multimodel fusion model has 12.7% and 7.6% improvements in DD and CC, respectively.

V. Conclusion

By introducing MSMG and IHS transforms into the NSST-PCNN algorithm, a new multimodel fusion method MS-NSST-PCNN is proposed for SAR and optical image fusion. Most of the evaluation indicators of the fused images by the proposed method are higher than those of traditional algorithms, especially in DD and CC indicators. The feasibility and effectiveness of the proposed fusion method for high-spatial resolution remote sensing images are verified as well. It also indicates that proposed method can greatly improve the DD and CC between the multispectral images and the fused images. The performance of the proposed fusion method in the fusion of high-spatial-resolution and medium-spatial-resolution images is slightly different, which is mainly related to the number of pixels available for fusion in the same cell, area selection, and land types at the test sites.

The ground object classification accuracy evaluation experiment in two application scenarios, urban and rural, was also carried out to further evaluate the fusion results. It is proved that our method can well extract the texture information of SAR while preserving the spectral information of the optical image. Moreover, the results of object classification show that the proposed method can efficiently distinguish the edge of small objects. The overall classification accuracy of study areas reaching 0.87 and 0.88 also verifies the effectiveness of the MS-NSST-PCNN method in object-oriented classification applications. Specifically, for rural test site A, our proposed method obtains 0.87 in OA, and 0.81 in KC. When compared with the Landsat-8 OLI, our proposed method improves OA by 8.75% and KC by 10.96%. For urban test site B, MS-NSST-PCNN gets 0.88 in OA and 0.76 in KC. When compared with the Landsat-8 OLI, our proposed method improves OA by 23.94% and KC by 43.40%.

In the future, the multimodel fusion technology of SAR and optical images will have the potential to achieve higher fusion precision and wider application scenarios. For example, the increased application of sparse representation methods, the deep learning algorithms and fusion source images with higher resolution and less noise can help improve the fusion performance.

References

[1] M. D. Mura, S. Prasad, F. Pacifici, P. Gamba, J. Chanussot, and J. A. Benediktsson, "Challenges and opportunities of multimodality and data fusion in remote sensing," *Proc. IEEE*, vol. 103, no. 9, pp. 1585–1601, Sep. 2015.

[2] H. Ghassemian, "A review of remote sensing image fusion methods," *Inf. Fusion*, vol. 32, pp. 75–89, Nov. 2016.

[3] C. Liao, J. Wang, T. Dong, J. Shang, J. Liu, and Y. Song, "Using spatio-temporal fusion of Landsat-8 and MODIS data to derive phenology, biomass and yield estimates for corn and soybean," *Sci. Total Environ.*, vol. 650, pp. 1707–1721, Feb. 2019.

[4] Q. Yin, M. Liu, J. Cheng, Y. Ke, and X. Chen, "Mapping paddy rice planting area in Northeastern China using spatiotemporal data fusion and phenology-based method," *Remote Sens.*, vol. 11, no. 14, Jan. 2019, Art. no. 1699.

[5] M. Ding, Q. Guan, L. Li, H. Zhang, C. Liu, and L. Zhang, "Phenology-based rice paddy mapping using multi-source satellite imagery and a fusion algorithm applied to the Poyang Lake Plain, Southern China," *Remote Sens.*, vol. 12, no. 6, Jan. 2020, Art. no. 1022.

[6] S. C. Kulkarni and P. P. Rege, "Pixel level fusion techniques for SAR and optical images: A review," *Inf. Fusion*, vol. 59, pp. 13–29, Jul. 2020.

[7] V. Pandit and R. Bhiwani, "Image fusion in remote sensing applications: A review," *Int. J. Comput. Appl.*, vol. 120, pp. 22–32, Jun. 2015.

[8] S. Li, X. Kang, L. Fang, J. Hu, and H. Yin, "Pixel-level image fusion: A survey of the State of the Art," *Inf. Fusion*, vol. 33, pp. 100–112, Jan. 2017.

[9] B. Aiazzi et al., "Context-sensitive pan-sharpening of multispectral images," in *Semantic Multimedia* (Lecture Notes in Computer Science Series), B. Falcidieno, M. Spagnuolo, Y. Avrithis, I. Kompatsiaris, and P. Buitelaar, Eds. Berlin, Germany: Springer, 2007, pp. 121–125.

[10] G. Hong, Y. Zhang, and B. Mercer, "A wavelet and IHS integration method to fuse high resolution SAR with moderate resolution multispectral images," *Photogrammetric Eng. Remote Sens.*, vol. 75, no. 10, pp. 1213–1223, Oct. 2009.

[11] Q.-G. Miao, C. Shi, P.-F. Xu, M. Yang, and Y.-B. Shi, "A novel algorithm of image fusion using Shearlets," *Opt. Commun.*, vol. 284, no. 6, pp. 1540–1547, Mar. 2011.

[12] C. S. Yilmaz, V. Yilmaz, and O. Gungor, "A theoretical and practical survey of image fusion methods for multispectral pansharpening," *Inf. Fusion*, vol. 79, pp. 1–43, Mar. 2022.

[13] M. Do and M. Vetterli, "Contourlets: A directional multiresolution image representation," in *Proc. Int. Conf. Image Process.*, 2002, pp. 497–501.

[14] A. D. Cunha, J. Zhou, and M. Do, "The nonsubsampled Contourlet transform: Theory, design, and applications," *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 3089–3101, Oct. 2006.

[15] K. Guo and D. Labate, "Optimally sparse multidimensional representation using Shearlets," *SIAM J. Math. Anal.*, vol. 39, no. 1, pp. 298–318, Jan. 2007.

[16] T. Chu, Y. Tan, Q. Liu, and B. Bai, "Novel fusion method for SAR and optical images based on non-subsampled Shearlet transform," *Int. J. Remote Sens.*, vol. 41, no. 12, pp. 4590–4604, Jun. 2020.

[17] W. Yi, Y. Zeng, and Z. Yuan, "Fusion of GF-3 SAR and optical images based on the nonsubsampled Contourlet transform," *Acta Optica Sinica*, vol. 38, no. 11, 2018, Art. no. 1110002.

[18] D. Anandhi and S. Valli, "An algorithm for multi-sensor image fusion using maximum a posteriori and nonsubsampled Contourlet transform," *Comput. Elect. Eng.*, vol. 65, pp. 139–152, Jan. 2018.

[19] X. L. Wang and C. X. Chen, "Image fusion for synthetic aperture radar and multispectral images based on sub-band-modulated non-subsampled Contourlet transform and pulse coupled neural network methods," *Imag. Sci. J.*, vol. 64, no. 2, pp. 87–93, Feb. 2016.

[20] B. Yan and Y. Kong, "A fusion method of SAR image and optical image based on NSCT and Gram-Schmidt transform," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2020, pp. 2332–2335.

[21] X. Sun, J. Zhang, Q. Yan, and J. Gao, "An IHS fusion method integrated by the nonsubsampled Contourlet transform to fuse the airborne X-band InSAR and P-band full PolSAR images," in *Proc. Int. Symp. Image Data Fusion*, 2011, pp. 1–4.

[22] Z. Liu, Y. Feng, Y. Zhang, and X. Li, "A fusion algorithm for infrared and visible images based on RDU-PCNN and ICA-bases in NSST domain," *Infrared Phys. Technol.*, vol. 79, pp. 183–190, Nov. 2016.

[23] Y. Zeng, W. Yi, J. Deng, W. Chen, S. Xu, and S. Huang, "Remote sensing image fusion using improved IHS and non-subsampled Contourlet transform," in *Urban Intelligence and Applications* (Studies in Distributed Intelligence Series), X. Yuan and M. Elhoseny, Eds. Cham, Switzerland: Springer, 2020, pp. 55–67.

[24] Y. Byun, J. Choi, and Y. Han, "An area-based image fusion scheme for the integration of SAR and optical satellite imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 5, pp. 2212–2220, Oct. 2013.
[25] L. Jiahuan, Z. Jian, and D. Yunfei, "A fusion method of multi-spectral image and panchromatic image based on NSCT transform and adaptive gamma correction," in *Proc. 3rd Int. Conf. Inf. Syst. Eng.*, 2018, pp. 10–15.
[26] D. Liu, F. Yang, H. Wei, and P. Hu, "Remote sensing image fusion method based on discrete wavelet and multiscale morphological transform in the IHS color space," *J. Appl. Remote Sens.*, vol. 14, no. 1, Mar. 2020, Art. no. 016518.
[27] N. Liu, W. Li, X. Sun, R. Tao, and J. Chanussot, "Remote sensing image fusion with task-inspired multiscale nonlocal-attention network," *IEEE Geosci. Remote Sens. Lett.*, vol. 20, 2023, Art. no. 5502505.
[28] F.-P. An, X.-M. Ma, and L. Bai, "Image fusion algorithm based on unsupervised deep learning-optimized sparse representation," *Biomed. Signal Process. Control*, vol. 71, Jan. 2022, Art. no. 103140.
[29] Z. Chen et al., "A novel sparse representation model for pedestrian abnormal trajectory understanding," *Expert Syst. Appl.*, vol. 138, Dec. 2019, Art. no. 112753.
[30] Z. Qin, J. Fan, Y. Liu, Y. Gao, and G. Y. Li, "Sparse representation for wireless communications: A compressive sensing approach," *IEEE Signal Process. Mag.*, vol. 35, no. 3, pp. 40–58, May 2018.
[31] J. Peng, W. Sun, and Q. Du, "Self-paced joint sparse representation for the classification of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 1183–1194, Feb. 2019.
[32] M. Yu, G. Dong, H. Fan, and G. Kuang, "SAR target recognition via local sparse representation of multi-manifold regularized low-rank approximation," *Remote Sens.*, vol. 10, no. 2, Feb. 2018, Art. no. 211.
[33] Z. Wang, Z. Cui, and Y. Zhu, "Multi-modal medical image fusion by Laplacian pyramid and adaptive sparse representation," *Comput. Biol. Med.*, vol. 123, Aug. 2020, Art. no. 103823.
[34] Z. Zhu, H. Yin, Y. Chai, Y. Li, and G. Qi, "A novel multi-modality image fusion method based on image decomposition and sparse representation," *Inf. Sci.*, vol. 432, pp. 516–529, Mar. 2018.
[35] H. Liu and S. Li, "Decision fusion of sparse representation and support vector machine for SAR image target recognition," *Neurocomputing*, vol. 113, pp. 97–104, Aug. 2013.
[36] C. Liu, Y. Qi, and W. Ding, "Airborne SAR and optical image fusion based on IHS transform and joint non-negative sparse representation," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2016, pp. 7196–7199.
[37] H. Zhang, H. Shen, Q. Yuan, and X. Guan, "Multispectral and SAR image fusion based on Laplacian pyramid and sparse representation," *Remote Sens.*, vol. 14, no. 4, Jan. 2022, Art. no. 870.
[38] R. Eckhorn et al., "Feature linking via synchronization among distributed assemblies: Simulations of results from cat visual cortex," *Neural Comput.*, vol. 2, no. 3, pp. 293–307, 1990.
[39] D. Gai, X. Shen, H. Chen, Z. Xie, and P. Su, "Medical image fusion using the PCNN based on IQPSO in NSST domain," *IET Image Process.*, vol. 14, no. 9, pp. 1870–1880, 2020.
[40] Z. Wang, W. Lv, and Z. Zhao, "Effect of spatial resolution ratio on optical remote sensing image fusion quality," *Opto-Electron. Eng.*, vol. 36, pp. 103–107, 2009.
[41] W. Li, X. Xiao, P. Xiao, H. Wang, and F. Xu, "Change detection in multitemporal SAR images based on slow feature analysis combined with improving image fusion strategy," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 3008–3023, 2022.
[42] P. Chavez et al., "'Statistical method for selecting LANDSAT MSS ratios," *J. Appl. Photographic Eng.*, vol. 8, pp. 23–30, 1982.
[43] B. Goyal, A. Dogra, and A. K. Sangaiah, "An effective nonlocal means image denoising framework based on non-subsampled Shearlet transform," *Soft Comput.*, vol. 26, pp. 7893–7915, Feb. 2022.
[44] G. Easley, D. Labate, and W.-Q. Lim, "Sparse directional image representations using the discrete Shearlet transform," *Appl. Comput. Harmon. Anal.*, vol. 25, no. 1, pp. 25–46, Jul. 2008.
[45] W. Kong, Q. Miao, Y. Lei, and C. Ren, "Guided filter random walk and improved spiking cortical model based image fusion method in NSST domain," *Neurocomputing*, vol. 488, pp. 509–527, Jun. 2022.
[46] N. Rafi and P. Rivas, "A review of pulse-coupled neural network applications in computer vision and image processing," 2021. [Online]. Available: https://www.rivas.ai/pdfs/rafi2021pcnn.pdf
[47] E. Blasch, "Biological information fusion using a PCNN and belief filtering," in *Proc. Int. Joint Conf. Neural Netw.*, 1999, pp. 2792–2795, ISSN: 1098–7576.
[48] W. Tan, P. Xiang, J. Zhang, H. Zhou, and H. Qin, "Remote sensing image fusion via boundary measured dual-channel PCNN in multi-scale morphological gradient domain," *IEEE Access*, vol. 8, pp. 42540–42549, 2020.
[49] T. Liu, L. Cui, and C. Zhang, "Study on fault diagnosis method of planetary gearbox based on turn domain resampling and variable multi-scale morphological filtering," *Symmetry*, vol. 13, no. 1, Jan. 2021, Art. no. 52.

Wenmei Li (Member, IEEE) received the M.S. degree in cartography and geography information system from Nanjing University, Nanjing, China, in 2010, and the Ph.D. degree in forest management specialized in remote sensing techniques and applications from the Chinese Academy of Forestry, Beijing, China, in 2013.

She is currently an Associate Professor with the School of Geographic and Biologic Information, Nanjing University of Posts and Telecommunications, Nanjing, China. She has been working for her postdoctoral studies with Nanjing University of Posts and Telecommunications, since 2018. Her research interests include deep learning, optimization, image reconstruct, and their application in land remote sensing.

Jiaqi Wu received the B.S. degree in geographic information science and the M.S. degree from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2019 and 2023, respectively.

His research interests include spatial- temporal fusion, time series, deep learning, and their applications in land remote sensing.

Qing Liu, photograph and biography not available at the time of publication.

Yu Zhang received the B.S. degree in communication engineering, in 2021, from the Nanjing University of Posts and Telecommunications, Nanjing, China, where he is currently working toward the M.A. degree in surveying and mapping science and technology.

His research interests include remote sensing, SAR data processing, inversion of forest parameters, terrain extraction, above ground biomass.

Bin Cui recieved the B.Eng. degree in surveying and mapping from Henan Polytechnic University, Jiaozuo, China, in 2011, the M.Eng. degree in surveying and mapping from the Beijing University of Civil Engineering and Architecture, Beijing, China, in 2015, and the Ph.D degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2020.

He is currently a Lecturer with the School of Geographic and Biologic Information, Nanjing University of Posts and Telecommunications, Nanjing, China. He is also enrolled as a Postdoctoral Researcher with the School of Geography and Ocean Science, Nanjing University, Nanjing, China. His research interests include remote sensing image processing, change detection, multimodality image fusion and its applications in remote sensing.

Yan Jia received the double M.S. degree in telecommunications engineering and computer application technology from Politecnico di Torino, Turin, Italy, and Henan Polytechnic University, Jiaozuo, China, in 2013, and the Ph.D. degree in electronics engineering from Politecnico di Torino, in 2017.

She is currently with the Nanjing University of Posts and Telecommunications, Nanjing, China. Her research interests include microwave remote sensing, soil moisture retrieval, global navigation satellite system reflectometry applications to land remote sensing and antenna design

Guan Gui (Senior Member, IEEE) received the Ph.D. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2012.

From 2009 to 2014, he joined Tohoku University as a Research Assistant and a Post-Doctoral Research Fellow. From 2014 to 2015, he was an Assistant Professor with Akita Prefectural University. Since 2015, he has been a Professor with the Nanjing University of Posts and Telecommunications, Nanjing, China. He has authored or coauthored more than 200 international peer-reviewed journal/conference papers. His current research interests include intelligent sensing and signal recognition. He has was the recipient of nine best paper awards, e.g., ICC 2017, ICC 2014, and VTC 2014.Spring. He was also the recipient of the IEEE Communications Society Heinrich Hertz Award in 2021, Clarivate Analytics Highly Cited Researcher in 2021, and Highly Cited Chinese Researchers by Elsevier in 2020 and 2021. He is serving or served on the Editorial Boards of several journals, such as IEEE Transactions on Vehicular Technology. In addition, he was the IEEE VTS Ad Hoc Committee Member in AI Wireless, Executive Chair of VTC 2021.Fall, Vice Chair of WCNC 2021, and the TPC Member of many IEEE international conferences. Since 2022, he has been a Distinguished Lecturer of the IEEE Vehicular Technology Society.