

Received April 14, 2022, accepted May 2, 2022, date of publication May 3, 2022, date of current version June 2, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3172501

Hyperspectral Classification of Two-Branch Joint Networks Based on Gaussian Pyramid Multiscale and Wavelet Transform

YIGANG TANG¹, XIAOLAN XIE², (Member, IEEE), AND YOUHUA YU¹, (Member, IEEE)

¹College of Information Science and Engineering, Guilin University of Technology, Guilin 541000, China

²Guangxi Key Laboratory of Embedded Technology and Intelligent System, Guilin 541000, China

Corresponding author: Xiaolan Xie (xie_xiao_lan@foxmail.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61762031, in part by the Science and Technology Major Project of Guangxi Province under Grant AA19046004, and in part by the Natural Science Foundation of Guangxi under Grant 2021JJA170130.

ABSTRACT Due to its high spectral resolution, Hyperspectral remote sensing data can provide practically continuous spectral curves for target objects and fully reflect the detailed characteristics of ground objects. However, the data redundancy generated by a large number of bands poses challenges to the feature extraction of target objects. Hence, the spectral data are processed by wavelet transform to reduce the influence of intra-class spectral variation on classification. The multi-scale image data were collected by Gaussian pyramid multi-scale transformation. Then the multi-scale spatial information was captured through the feature extraction network to improve the classification accuracy. We propose a dual-branch feature extraction network. The first branch adopts Gaussian pyramid multi-scale transformation to obtain multi-scale images and then applies the feature extraction module to gain multi-scale spatial features. The second branch employs wavelet transform to process spectral data to reduce the impact of abnormal spectral data on classification and then applies a feature extraction module to acquire spectral features. Finally, the spectral and spatial features obtained by the two branches are fused in the full connection layer to achieve an accurate classification. This method can effectively capture the fine features of hyperspectral images by combining spectral features and spatial features of different scales. Simultaneously, it can capture the interaction between spectral and spatial features by combining spatial and spectral features through joint learning. Experimental results on hyperspectral image datasets indicate that the method outperforms other traditional deep learning-based and other advanced classifiers.

INDEX TERMS Hyperspectral image, Gaussian pyramid, wavelet transforms, hyperspectral image classification.

I. INTRODUCTION

Hyperspectral image (HSI) is a three-dimensional image captured by an aerospace vehicle carrying a hyperspectral imager, and each pixel in the image contains hundreds of different wavelengths of reflection information, which makes it suitable for many practical applications such as military target detection, mineral exploration, environmental monitoring, and agricultural production. In recent years, hyperspectral remote sensing techniques have received great attention in various applications of Earth observation [1]–[6]. Hyperspectral imagery (HSI) provides hundreds of continuous narrow

spectral bands [7]–[10], which can distinguish different substances more accurately than conventional panchromatic and multispectral remote sensing images. With its high spectral resolution, HSI is uniquely advantageous for finer classification [11], [12], as it can detect subtle spectral features that cannot be resolved by conventional images. In the early stages of HSI classification, many machine learning-based methods have been applied, such as the nearest neighbor, decision tree, and linear function algorithms. Among these methods, k-nearest neighbor (K-NN) [13] can be considered as the simplest classifier, which uses the nearest neighbor algorithm and is the simplest classifier that employs the Euclidean distance to measure the similarity between the test and training samples. Support vector machine (SVM) [14] hyperspace

The associate editor coordinating the review of this manuscript and approving it for publication was Donato Impedovo¹.

images (HSI) can classify data in high-dimensional space by dividing hyperplanes based on different spatial structures and spectral features of different materials. Some popular classifiers, such as support vector machine (SVM) [14] and correlation vector machines, have been widely used for HSI analysis.

Recently, sparse representation (SR) has been applied to HSI classification as a powerful image processing tool [15]. Sparse representation relies on the assumption that pixels in the same class should have similar spectral features. Test samples can be linearly represented by a small number of training samples from the same class. However, traditional SR methods consider only the spectral information of the test pixels and ignore the spatial neighbors around the test pixels. Based on the assumption that pixels in a local region usually have similar spectral materials and features, Chen [16] proposed a JSR-based classification (JSRC) algorithm, which considers spatial information by jointly representing pixels in a local window to obtain better classification performance. In [17] and [18], some space-based classifiers, such as kernel-based SR and l_2 -norm regularized sparse subspace clustering, were proposed, and these classifiers showed better performance.

However, these local window-based methods have a single limitation. A common limitation of the methods is that the narrow windows may include different classes of pixels. In other words, pixels in the edge regions are appropriate for small-sized windows, while large-sized windows are suitable for smooth regions. To solve this problem, Fang [19] proposed the multi-scale adaptive SR (MASR) method, which obtained better performance. Other advanced tools, such as the adaptive mean shift analysis method [20], are also effective ways to solve this problem.

The main contributions of this paper are as follows:

- 1) Compared with the traditional feature extraction method, the multi-scale spatial feature extraction method proposed in this paper pays more attention to the detailed information and local information of HSI in the model. And the local information can effectively preserve the local structure inherited from the original data, reducing the loss of key information. The spectral feature extraction method proposed in this paper uses wavelet transform to process spectral data, which can reduce the impact of abnormal data on the overall classification results.
- 2) This paper uses the method of double-branch feature extraction and fusion. The spatial features and spectral features are extracted by different feature extraction networks, and then the feature fusion is performed to obtain the classification results.

II. RELATED WORK

HSI not only contains abundant ground image information but also has rich spectral information. Therefore, various spatial-spectral classification methods have been developed [21]–[23]. For example, Kang *et al.* proposed

a spectral-spatial classification-based method of edge-preserving filtering (EPF) [24]. EPF is a hot research topic in the field of image processing and computer vision in recent years. In recent years, EPF is a research hotspot in image processing and computer vision, which not only holds the function of image smoothing but also can import spatial structure information into the input image. Peng *et al.* proposed a region kernel to measure region-to-region distance similarity for HSI classification [25]. The region kernel was designed as a linear combination of multiscale box kernels, which can handle HSI regions with arbitrary shapes and sizes. These methods consider information about the spatial environment of the pixel and the surrounding pixels while using the spectral information of ground objects for classification. Therefore, they can effectively reduce the influence of the phenomenon of the same object with different spectra and distinct objects with the same spectrum on the classification accuracy. At the same time, the classification accuracy can be significantly improved.

With the advancement of artificial intelligence and deep learning, various CNN-based HSIC methods have attracted a lot of attention in solving the nonlinear structure of hyperspectral data [8], [26]–[32]. Convolutional networks with multilayer CNN for HSIC were first designed [33] to extract spectral features. Yu [34] proposed a CNN structure embedded with extracted hash features to improve the accuracy of HSIC. In [35], a recurrent neural network (RNN) was proposed to process spectral data using the following information of the spectral correlation and band variability.

Nonetheless, the range of variation in spectral magnitude or pixel shape can vary widely within the same ground level. Therefore, foreign objects may have similar spectra, while objects within a class may have discrete spectra. This is the bottleneck hindering the improvement of classification accuracy. In addition, spectral variations complicate the statistical distribution of sample points and exacerbate problems associated with small sample sizes. Therefore, reducing the effects of spectral changes is one of the key issues for HSI classification. To address this problem, various HSI classification methods have been proposed.

Xue *et al.* [36] proposed a completely different approach from the perspective of sub-pixel target detection. They used band selection followed by nonlinear expansion (BSNE) and iterative constrained energy minimization to classify HSIs. It is not difficult to implement and has advantages over other methods. Meanwhile, with the development of remote sensing imaging technology, the spatial resolution of HSI has been increased. The joint spectral-spatial features have also attracted more attention [37]. For example, many features in [38], such as multiple features, texture features, grayscale coexistence features and statistical features are used as computational parameters to obtain better classification results. Moreover, inspired by tensor learning, spatial and spectral features can also be fused to the 3-D tensor [39]. It can narrow the loss of structural information intrinsic to HSI. Guo [40] proposed a tensor-based technique for HSI classification and

used multilinear principal component analysis to preprocess the tensor. However, most of these tensor methods build the tensor physically and ignore those features that follow a pre-determined logical arrangement. In [41], a novel generalized tensor regression (GTR) method extended from a simple but effective classifier, was used for HSI classification.

Currently, convolutional neural network (CNN) has shown excellent performance in HSI classification [8], [32]. This is because it can naturally deal with the problem that HSIs are often in a nonlinear and complex feature space [42]. Some researchers have proposed multi-stream CNN for HSI classification. Another aspect [43], [44]. In most of CNN, feature extraction and classifier training are separated. To overcome this drawback [45], a spectral-spatial unified network (SSUN) [45] was designed and combined with shallow and deep convolutional layers to deal with the information loss problem. In [46], a dual-channel deep CNN is proposed. Discriminative information is captured by the spectral domain and deep convolutional layer, respectively. The information is available from the spectral and spatial domains respectively and can be efficiently utilized and fused. Xu [47] proposed a novel two-branch CNN (MS-CNN) based on multi-source data for the classification and fusion of HSI and data from multiple other sensors. data from multiple other sensors, such as Light Detection and Ranging (LIDAR) [48] data. They help the two networks to place emphasis on different features separately and obtain excellent classification performance. Cao [49] proposed integrate active learning and deep learning into a unified framework and leverage Markov random fields (MRF [50]) to enhance the smoothness of class labels to further improve classification performance (CNN-AL-MRF). Han [51] proposed a method to select pixel blocks of different scales around the central pixel as the basic unit for processing. Then, a spatial augmentation strategy is intended to obtain various spatial location information under limited training samples through spatial rotation and row-column transformation to obtain better accuracy. Dong [52] propose weighting feature fusion based on convolutional neural network and graph attention network (WFCG) for HSI classification. The GAT is first built with the help of superpixel-based encoder and decoder modules, and then coupled with the attention mechanism to build the CNN. Finally, features are weighted and fused with the features of the two neural network models. Ortac [53] proposed to using 1D, 2D and 3D convolutional neural networks to classify samples from widely used hyperspectral datasets by extracting spatial, spectral and spatial-spectral features.

The research of hyperspectral image classification faces some challenges, each pixel of the hyperspectral image contains hundreds of different bands of reflection information, the number of bands and the correlation between adjacent bands of the hyperspectral image is large, there is high information redundancy, as well as a large amount of hyperspectral image data, needs a lot of operations, using data dimensionality reduction algorithm can not only reduce data redundancy but also reduce the number of operations.

PCA (Principal Component Analysis) [54] is a principal component analysis method, which is a technique to analyze and simplify data sets. The principal component analysis is often reduce the dimensionality of a dataset while maintaining the features in the dataset that contribute the most to each other's variance. This is achieved by retaining lower-order principal components and ignoring higher-order principal components so that the lower-order components tend to retain the most important aspects of the data. It is particularly suitable for dimensionality reduction processing in hyperspectral images, thus achieving a reduction in data redundancy and computational effort. Moreover, there is a strong correlation between similar neighboring pixel points in hyperspectral images, and the small size or low contrast of a certain type of object in the image requires a higher resolution to be observed, while on the contrary the strong contrast or large size of a certain type of object in the image requires just a lower resolution, and both situations exist in hyperspectral images, which then require multi-resolution processing. The bottom of the Gaussian pyramid [55] is a high-resolution representation of the image to be processed, while the top is a low-resolution representation, and as the pyramid moves up the pyramid, both size, and resolution decrease. In this way, a multi-resolution and multi-size image can be calculated by Gaussian pyramid, and a better performance can be obtained for its feature extraction and classification.

With the efforts of scientists, wavelet transform has been well developed in various fields. Discrete wavelet transform (DWT) [56] is a time-frequency analysis method in signal processing improved based on Fourier transform, which can represent local features in both time and frequency domains, and is an important algorithm in image coding. Barzegar [57] improved the model's prediction accuracy on the dataset by combining boundary correction (BC) [58] and maximum overlap discrete wavelet transform (MODWT) [59] preprocessing data with a hybrid convolutional neural network (CNN). The current hyperspectral remote sensing image classification method does not apply to all types of images. The hyperspectral remote sensing image data can be regarded as three-dimensional cube data, and we adopt multi-scale and multi-resolution feature extraction for the hyperspectral image in pixel space feature to get more accurate classification features and in spectral feature, we try to process the image spectral curve of hyperspectral remote sensing image and then introduce Fourier transform and wavelet transform in signal processing to realize the image processing. The processed image spectral data are put into LSTM neural network for classification, while the multi-scale and multi-resolution images are put into the resNet [60] network for classification, and the two network output layers are connected together into a fully connected network for classification to obtain the final results.

III. METHODOLOGY

Our HSI classification method consists of two networks whose overall architecture is shown in Fig.1 The upper part

is processed by PCA for dimensionality reduction, and then the dimensionality-reduced data is downsampled through the Gaussian pyramid to obtain images of three scales, which are divided into two resNet networks with different layers for feature extraction, as shown in Fig.2 and Fig.3. Each module in turn consists of two different residual blocks as shown in Fig. 4 and Fig. 5. Use a shallow network for small-sized images and a deep network for large-sized images to prevent overfitting. The following part performs discrete wavelet processing on the spectral data of each pixel point, and obtains the spectral data after denoising and de-anomaly, which is input to the LSTM network for spectral feature extraction. The first network LSTM requires discrete wavelet processing of the data to obtain finer denoised data. The second network resNet needs to segment the data, place each pixel to be classified in the center of a rectangular area, and then perform PCA data down scaling and Gaussian pyramid processing on the segmented image to obtain a multi-scale and multi-resolution image. Using two separate networks to extract spatial features and spectral features respectively, and finally combining the two parts of the network through a fully connected network to obtain the classification results. Details are described in Section IV-A-IV-B-IV-C.

A. REGION MULTI-SCALE SPATIAL FEATURE EXTRACTION

Based on the above analysis, hyperspectral data are prone to dimensional disasters because of their data redundancy as well as a large number of operations and the high-dimensional features involved. As the dimensionality of the dataset increases, the number of samples required for algorithm learning increases exponentially. In hyperspectral classification, it is very disadvantageous to encounter such large data because redundant data and certain unimportant features reduce the classification accuracy, and more memory and processing power are required to learn from large datasets. In addition, the sparsity of the data increases as the level. It is very difficult to explore the same dataset in a high-dimensional vector space than in an equally sparse dataset. Therefore, this paper first uses PCA to reduce the dimensionality of HIS data to alleviate the dimensional catastrophe and data redundancy. PCA algorithm can make the samples more spatially relevant after rounding off some information to remove some noise and data that affect the classification effect. The PCA algorithm is utilized to partition the data into regions, and the pixels to be classified are placed at the center of the partitioned region. The segmented data are then down sampled using the Gaussian pyramid algorithm to obtain multi-scale and multi-resolution image data. The Gaussian pyramid is used for down sampling. First, the original image is used as the bottom image G_0 (layer 0 of the Gaussian pyramid), and the gaussian kernel is used to convolve it, and then the convolved image is down sampled to get the upper-layer image G_1 , and this image is used as the input, and the convolution and down sampling operations are repeated to get the upper-layer image, iterating several times to form a pyramid-shaped image data structure as shown

in Fig.6. L_1 indicates the upper layer to be generated, and L_{l-1} indicates the lower layer of the pyramid to obtain the upper layer by F-function calculation. The image is continuously computed by iterating through the following equation:

$$L_1(m, n) = \sum_{u=-c}^c \sum_{v=-c}^c F(u, v) * L_{l-1}(2m-1-u, 2n-1-v) \quad (1)$$

The * is the convolution operation. $\mathbf{X} \in \{1, 2, 3, \dots, l\}$, l is the total number of layers of the Gaussian pyramid. $F(u, v)$ is $(2c+1) \times (2c+1)$ Gaussian window, which can be defined as:

$$F(u, v) = \frac{1}{2\pi\Upsilon^2} e^{-\frac{(u^2+v^2)}{2\Upsilon^2}} \quad (2)$$

where Υ is the variance of the Gaussian filter. The Gaussian pyramid consists of a series of images $\{L_1, L_2, L_3, \dots, L_l\}$, which is generated by the above equation.

Put the data processed by dimensionality reduction and Gaussian pyramid into the resNet network for feature extraction. The resNet [60] network can extract deep-level features and networks that can deepen the number of network layers without affecting the effect. It is especially suitable to use it to extract hyperspectral image features, because hyperspectral images have many dimensions, even after the PCA algorithm reduces the layer. The number is not low, so use the resNet network to extract its spatial features.

The overall structure of the image is achieved by performing the same steps for down-sampling to obtain the processed image. The same processing is applied to the HIS image so that the HSI can be analyzed at multiple scales, and different spatial features can be available at different scales. It classifies with distinct layers of the resNet [60] network. Each layer of the pyramid is obtained by the following formula. L_1 represents the upper layer that needs to be generated, and L_{l-1} means the lower layer of the pyramid is calculated by the F function to obtain the upper layer representation. Iteratively calculate the value of the image pixel by the following formula.

B. SPECTRAL FEATURE EXTRACTION

For spectral features, we sample wavelet transform for each image element of the hyperspectral image, because of weather, atmosphere, light or satellite sensor imaging process, the uncompensated atmosphere, uncompensated error of the sensor and the angle of the sun relative to the zenith and other objective factors may cause the spectral curve of similar image elements of the hyperspectral image will be different so that the spectral curve of similar image elements is very different, resulting in The recognition rate is reduced, and the image elements cannot be correctly classified by spectral curves alone. Therefore, this paper uses wavelet transform to reduce the gap of spectral curves as well as to decrease the noise points. The processed spectral data are put into the LSTM network to learn and extract the spectral features. The wavelet transformed spectral curve data are used as

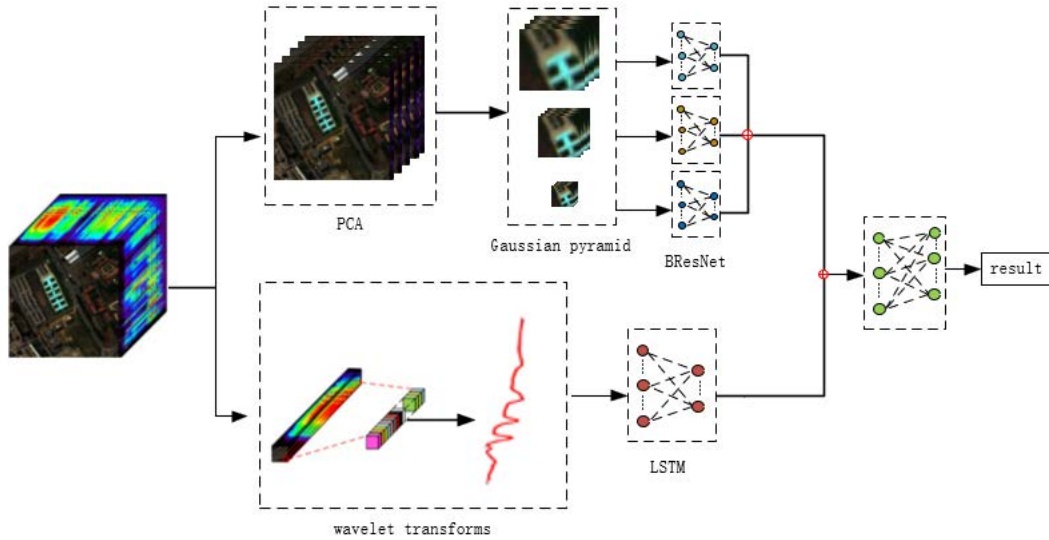


FIGURE 1. The overall flowchart of the proposed GWJ-Net.

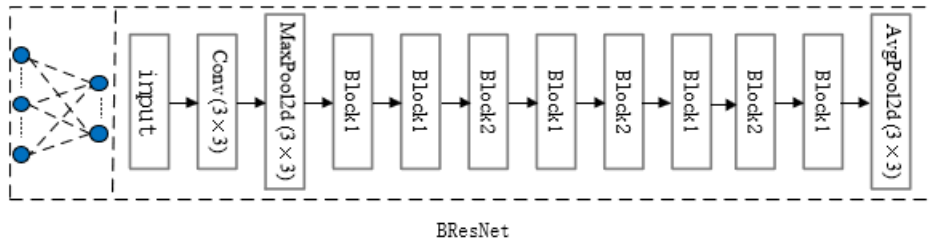


FIGURE 2. The overall flowchart of the BResNet(11).

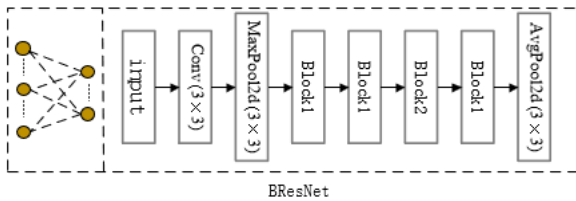


FIGURE 3. The overall flowchart of the BResNet(7).

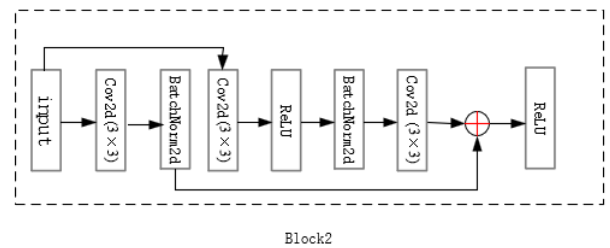


FIGURE 4. The BResNet residual block1 structure.

sequence data and then used to train the LSTM network, which can remember the previous input data and influence the next input, so that it can learn the long-term dependency. Moreover, the forgetting gate can determine what information we will discard from the cell state, and discarding some unimportant data can improve the accuracy of the model.

C. THE PROPOSED MODEL

The models we use are residual network and long-short memory neural network, which perform region segmentation on the data after PCA dimensionality reduction and segment the pixel points to be classified into a rectangular block, which we

segment into 11×11 rectangular blocks. Then the rectangular block data are processed by a Gaussian pyramid to get multi-scale multi-resolution hyperspectral region image data and use the residual network for spatial feature extraction. Data normalization is performed before in putting into the network so that each batch of data is similarly distributed and gradient vanishing can be avoided. We use the following formula to make the mean of the data close to 0 and the standard deviation close to 1. The formula is as follows:

$$Z = \gamma \frac{Z - E(Z)}{\sqrt{Var(Z) + \xi}} + \beta \tag{3}$$

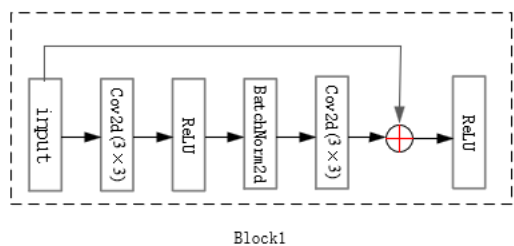


FIGURE 5. The BResNet residual block2 structure.

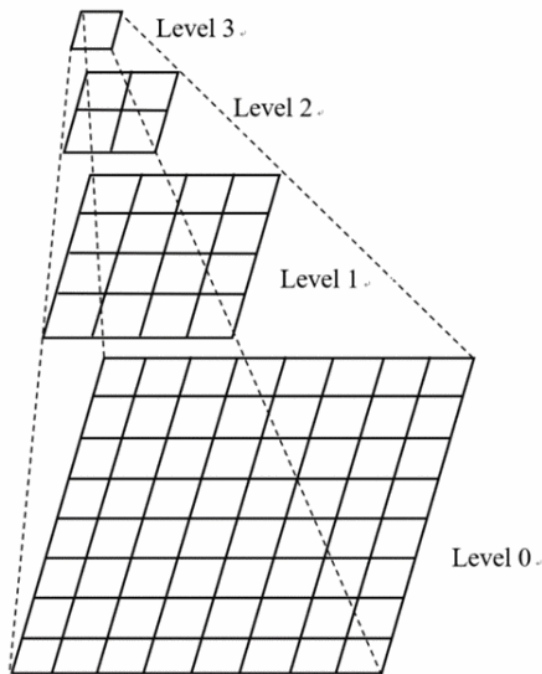


FIGURE 6. The Gaussian pyramid structure.

Where Z denotes the output of samples in the batch after applying batch normalization, $E(Z)$ and $Var(Z)$ represent the expectation and variance of Z , respectively, and γ and β are hyper parameters to be learned. We down sampled the data three times to obtain four different scales of images and trained the network with four residual networks without layers to prevent overfitting. The wavelet-transformed spectral data are constituted into the long-short memory neural network for feature extraction.

IV. EXPERIMENTAL RESULTS

In this section, we establish the effectiveness of the proposed method on three datasets and compare it with current state-of-the-art methods. All programs in the experiments are run on Python, and the network models are built using the PyTorch deep learning framework, an open-source Python machine learning library that allows custom deep learning models that can be trained and used flexibly. The networks used as well as the dataset processing procedures are implemented using the Python language. The computer configuration used in our

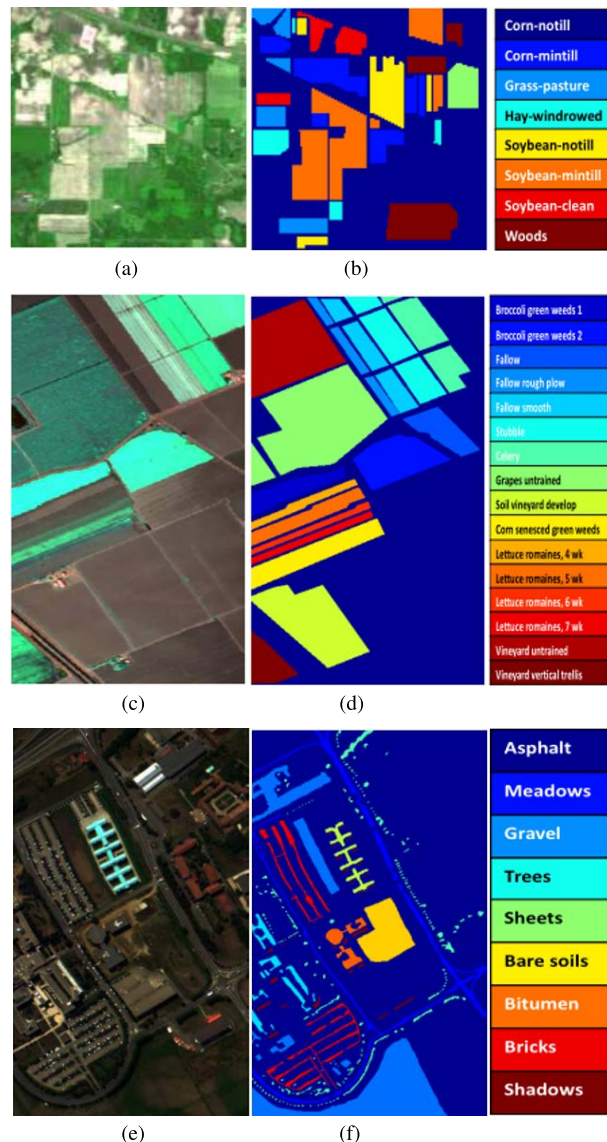


FIGURE 7. For three experimental datasets: (a) False-color image of the Indian Pines data, (b) Ground truth of the Indian Pines data, (c) False-color image of the Salinas data, (d) Ground truth of the Salinas data, (e) False-color image of the University of Pavia data, and (f) Ground truth of the University of Pavia data.

TABLE 1. The numbers of training and testing samples for the Indian Pines dataset.

#	Class	Training	Test
1	Corn-notill	200	1228
2	Corn-mintill	200	630
3	Grass-pasture	200	283
4	Hay-windrowed	200	278
5	Soybean-notill	200	772
6	Soybean-mintill	200	2255
7	Soybean-clean	200	393
8	Woods	200	1065
-	Total	1600	6904

experiment is as follows: CPU AMD3990X, GPU3090 and computer memory are 64G.

TABLE 2. The numbers of training and testing samples for the Salinas dataset.

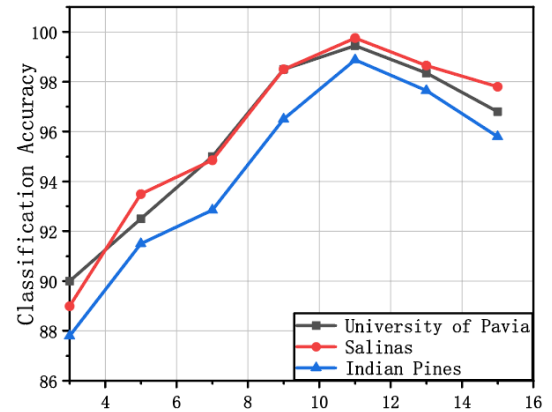
#	Class	Training	Test
1	Broccoli green weeds 1	200	1809
2	Broccoli green weeds 2	200	3526
3	Fallow	200	1776
4	Fallow rough plow	200	1194
5	Fallow smooth	200	2478
6	Stubble	200	3759
7	Celery	200	3379
8	Grapes untrained	200	11071
9	Soil vineyard develop	200	6003
10	Corn senesced green weeds	200	3078
11	Lettuce romaines, 4 wk	200	868
12	Lettuce romaines, 5 wk	200	1727
13	Lettuce romaines, 6 wk	200	716
14	Lettuce romaines, 7 wk	200	870
15	Vineyard untrained	200	7068
16	Vineyard vertical trellis	200	1607
-	Total	3200	50929

TABLE 3. The numbers of training and testing samples for the university of Pavia dataset.

#	Class	Training	Test
1	Asphalt	200	6431
2	Meadows	200	18449
3	Gravel	200	1899
4	Trees	200	2864
5	Sheets	200	1145
6	Baresoil	200	4829
7	Bitumen	200	1130
8	Bricks	200	3482
9	Shadows	200	747
-	Total	1800	40976

A. EXPERIMENTAL DATA

To evaluate the usability of our method, the performance of our proposed dual-branch feature network is evaluated on three commonly used datasets, namely Indian Pines dataset, Salinas dataset and University of Pavia dataset. The classification results are shown in Fig.7. For each dataset, we randomly selected 200 labeled pixels of each class for training and all other pixels in the ground truth graph for testing. The Indian Pine dataset consists of 145×145 pixels and was collected by an airborne visible infrared imaging spectrometer (VIRIS) located in northwest Indiana. 220 spectral channels are covering $0.4 \sim 2.5 \mu\text{m}$ with a spatial resolution of 20 m. The Indian Pine dataset initially has 16 different land cover classes. However, from a statistical point of view, we did not select classes with a low number of samples and selected eight classes with a high number of samples. The numbers of training and test samples are listed in Table 1. are shown in Fig. 7. Three experimental data sets: (a) pseudo-color image data of Indian pine trees, (b) ground truth data of Indian pine trees, (c) pseudo-color image data of Salinas, (d) ground truth data of Salinas, (e) pseudo-color image data of Pavia University, and (f) ground truth data of Pavia University. The second dataset is the Salinas dataset collected by the sensor on AVIRIS consists of 512×217 pixels, the image contains 224 spectral bands with a spatial resolution

**FIGURE 8.** Varying window size performances.

of 3.7 m. In total, there are 16 classes and the number of training and test samples are listed in Table 2.

The third dataset is that of the University of Pavia (containing 610×340 pixels) was collected by the Reflective Optics System Imaging Spectrometer covering the city of Pavia, Italy. The dataset consists of 103 spectral bands covering $0.43 \sim 0.865 \mu\text{m}$ with a spatial resolution of 1.3 m. Approximately 42776 pixels in the ground truth map have been labeled and divided into 9 categories, and the number of training and test samples is shown in Table 3.

B. LEARNING THE PROPOSED METHOD

For each training pixel, we use the surrounding 11×11 pixels, and diverse regions are extracted from the square-based region and then pour into a sequence of convolutional layers. Note that the proposed diverse-region strategy can be viewed as a flexible representation of the square-shaped region, hence the region size affects the ultimate performance of the proposed GWJ-Net. Here, we empirically set the global region size to be 11×11 .

The Adam algorithm is different from the traditional stochastic gradient descent, which maintains a single learning rate to update all weights, and the learning rate does not change during the training process, while Adam designs independent adaptive learning rates for different parameters by calculating the first-order moment estimates and the second-order moment estimates of the gradients. Authors of the Adam algorithm describe it as a collection of the advantages of two stochastic gradient descent extensions by calculating the first-order moment estimates and second-order moment estimates of the gradient to design independent adaptive learning rates for different parameters. In this paper, we set the initial learning rate to 0.001 and the batch size to 200. Fig.8 shows the classification performance of the rectangular region for different window sizes, from 3×3 to 15×15 . The performance tends to be satisfactory when the window size is 11×11 . When the window size is 11×11 is not the optimal window size for all experimental datasets.

TABLE 4. Comparison of the classification accuracy(%) among the proposed method and the baselines using the Indian Pines data.

Class	Performance								
	SVM-RBF [61]	SVM-RFS [62]	SVM-MRF [50]	CNN [33]	R-PCA CNN [63]	CNN-PPF [62]	CD-CNN [64]	SS-CNN [65]	GWJ-NET
1	76.14	88.73	93.55	78.58	82.39	92.99	90.1	96.28	97.72±0.012
2	85.40	91.20	90.41	85.24	85.41	96.66	97.1	92.26	97.77±0.003
3	97.88	97.52	95.80	96.10	95.24	98.58	100	99.3	99.55±0.02
4	99.28	100	100	99.64	100	100	100	100	100±0
5	83.94	91.67	91.12	89.64	82.76	96.24	95.9	92.84	99.58±0.03
6	73.48	78.79	97.72	81.55	96.2	87.8	87.1	98.21	99.56±0.02
7	92.11	93.76	91.71	95.42	82.14	98.98	96.4	92.45	99.39±0.03
8	97.28	98.74	99.84	98.59	99.81	99.81	99.4	98.98	99.52±0.01
AA	88.19	92.55	95.02	90.60	90.49	96.38	95.75	96.29	99.13±0.001
OA	82.98	88.68	95.34	87.01	91.13	93.9	94.24	96.63	99.06±0.12

TABLE 5. Comparison of the classification accuracy(%) among the proposed method and the baselines using the Salinas data.

Class	Performance								
	SVM-RBF [61]	SVM-RFS [62]	SVM-MRF [50]	CNN [33]	R-PCA CNN [63]	CNN-PPF [62]	CD-CNN [64]	SS-CNN [65]	GWJ-NET
1	96.81	99.55	100	97.34	98.84	100	100	100	100±0
2	94.67	99.92	99.70	99.29	99.61	99.88	100	99.89	100±0
3	90.27	99.44	98.94	96.51	99.75	99.60	100	99.89	99.69±0.002
4	98.61	99.86	98.44	99.66	98.79	99.49	99.3	99.25	99.58±0.001
5	94.82	98.02	99.47	96.97	99.84	98.34	98.5	99.39	99.66±0.002
6	97.61	99.7	99.95	99.60	99.7	99.97	100	100	100±0
7	99.24	99.69	100	99.49	79.05	100	99.8	99.82	100±0
8	54.69	84.85	87.64	72.25	99.17	88.68	83.4	91.45	99.74±0.004
9	98.32	99.58	99.45	97.53	96.88	98.33	99.6	99.95	100±0
10	81.91	96.49	94.41	91.29	99.31	98.6	94.6	98.51	99.84±0.003
11	90.57	98.78	99.91	97.58	100	99.54	99.3	99.31	100±0
12	92.43	100	99.64	100	100	100	100	100	100±0
13	98.07	99.13	100	99.02	98.97	99.44	100	99.72	100±0
14	90.39	98.97	98.79	95.05	82.24	98.96	100	100	100±0
15	60.06	76.38	83.37	76.83	97.57	83.53	100	96.24	98.73±0.029
16	90.87	99.56	97.34	98.94	99.61	99.31	98	99.63	100±0
AA	89.33	96.87	97.32	94.84	96.83	97.73	98.28	98.94	99.82±0
OA	81.55	93.15	94.59	89.28	92.39	94.8	95.42	97.42	99.74±0.11

TABLE 6. Comparison of the classification accuracy(%) among the proposed method and the baselines using the university of Pavia data.

Class	Performance								
	SVM-RBF [61]	SVM-RFS [62]	SVM-MRF [50]	CNN [33]	R-PCA CNN [63]	CNN-PPF [62]	CD-CNN [64]	SS-CNN [65]	GWJ-NET
1	84.01	87.95	98.22	88.38	92.43	97.42	94.6	97.4	99.65±0.005
2	88.9	91.17	98.90	91.27	94.84	95.76	96	99.4	99.96±0
3	87.57	86.99	88.97	85.88	90.89	94.05	95.5	94.84	98.42±0.03
4	96.09	95.5	93.64	97.24	93.99	97.52	95.9	99.16	97.86±0.003
5	99.91	99.85	99.11	99.91	100	100	100	100	99.63±0.02
6	93.33	94.31	80.13	96.41	92.86	99.13	94.1	98.7	99.85±0.04
7	93.98	94.74	82.79	93.62	93.89	96.19	97.5	100	99.26±0.003
8	82.94	85.89	91.88	87.45	91.18	93.62	88.8	94.57	98.27±0.002
9	99.6	99.89	100	99.57	99.33	99.6	99.5	99.87	98.67±0.002
AA	91.82	92.92	94.04	93.36	94.38	97.03	95.77	98.22	99.06±0.001
OA	89.24	91.1	92.63	92.27	93.87	96.48	96.73	98.41	99.45±0.253

For example, the red curve indicates that the square area 11×11 is the best window size for the Pavia University dataset, the blue curve indicates that the best window size for the Indian Pines dataset is 11×11 , and the best window size for the Salinas dataset is 11×11 , while when converting spectral data to images, the best image size for the Indian

Pines dataset is, Therefore, we choose a relatively large size among the allowed hardware resources to prevent wasting hardware resources and exceeding hardware usage.

Different initial learning rates are tested on the Indian Pines dataset as showed in Table 8. It can be observed that a larger learning rate may degrade the classification performance; the

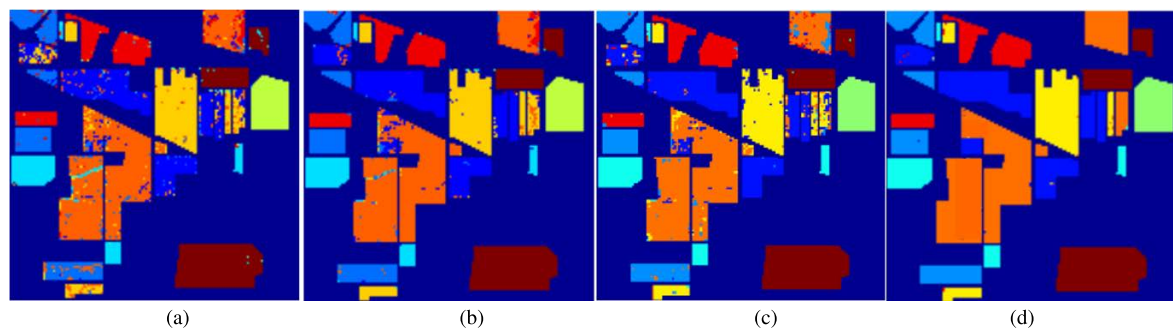


FIGURE 9. Classification maps from the proposed GWJ-NET and the baselines on the Indian Pines Data, (a) CNN:97.01%, (b) CNN-PPF:93.9%, (c) CD-CNN:94.24%, (d) GWJ-NET:99.06%.

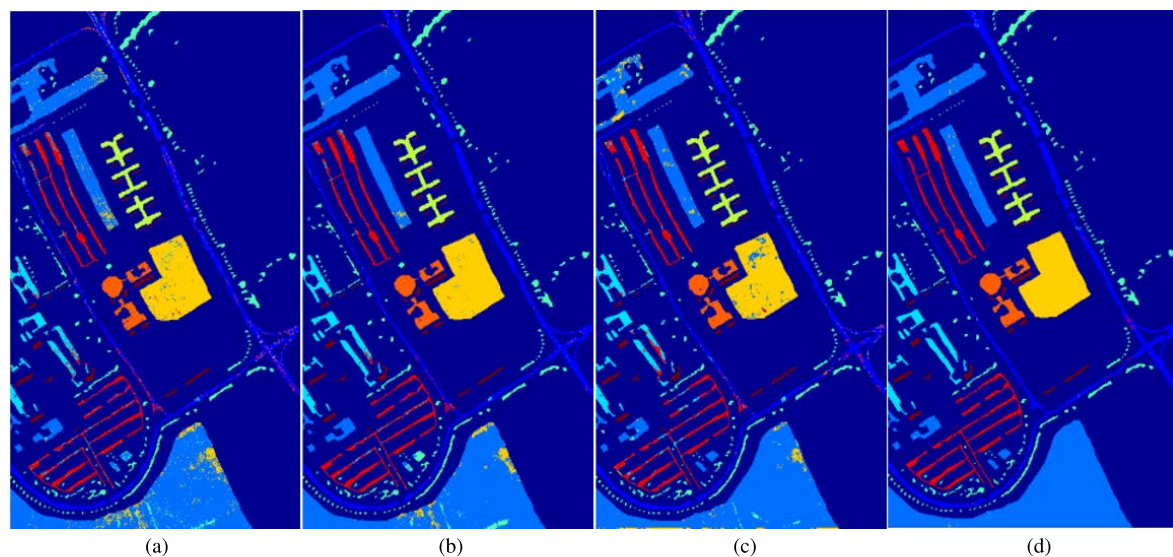


FIGURE 10. Classification maps from the proposed GWJ-NET and the baselines on the University of Pavia Data, (a) CNN:89.28%, (b) CNN-PPF:94.8%, (c) CD-CNN:95.42%, (d) GWJ-NET:99.74%.

best performance is achieved when the rate of learning is around 0.001. In follow-up experiments, we set the learning rate as 0.001 for GWJ-Net.

C. CLASSIFICATION PERFORMANCE

Our proposed GWJ-NET method is compared with current advanced hyperspectral image classification methods, such as SVM-RBF (SVM using radial basis function) and SVM-RF (SVM based on the random selection of features), as well as CD-CNN, SS-CNN, CNN-PPF, CNN, SVM-MRF, and R-PCA CNN. in Tables 4-6 Accuracy, OA, and AA indicate the classification performance of different methods on several datasets. We use a random selection method for dataset segmentation, for each category of data we randomly select 200 data for training and the remaining data for testing. The experimental results obtained from us show that the classification accuracy based on spatial-spectral features is higher than that using only one of the features, and our proposed GWJ-NET model is higher than other classifiers in terms of classification accuracy. In Table 5, we can obtain the accuracy of our proposed GWJ-NET as 99.74%, which is 7.35% higher

than the accuracy of R-PCA CNN (92.39%) compared to the table, and about 4% higher compared to CD-CNN (95.42%). A similar performance also exists for experiments conducted on other datasets. The classification performance of our proposed method on the University of Pavia dataset, Indian Pines dataset, and Salinas dataset is approximately 2%, 1%, and 1% better than the other classification methods compared to the other methods. Fig. 9-11 show the graphs of the classification results for each classifier, and the graphs show that the classification results are consistent with the classification results in Tables 4-6. By the presentation of the images, we can visually see that compared with the CNN, CNN-PPP, and CD-CNN in the table, our proposed GWJ-NET classification result graph has significantly fewer points of classification errors in many regions, such as the Bare soil region in Fig. 11 and the Soybean-clean region in Fig. 9, so we can conclude that our GWJ-NET classification method performs better than other classification methods.

Table 7 lists the classification performance with different numbers of training samples per class increasing from 50 to 200 with an interval of 50. Obviously, for all

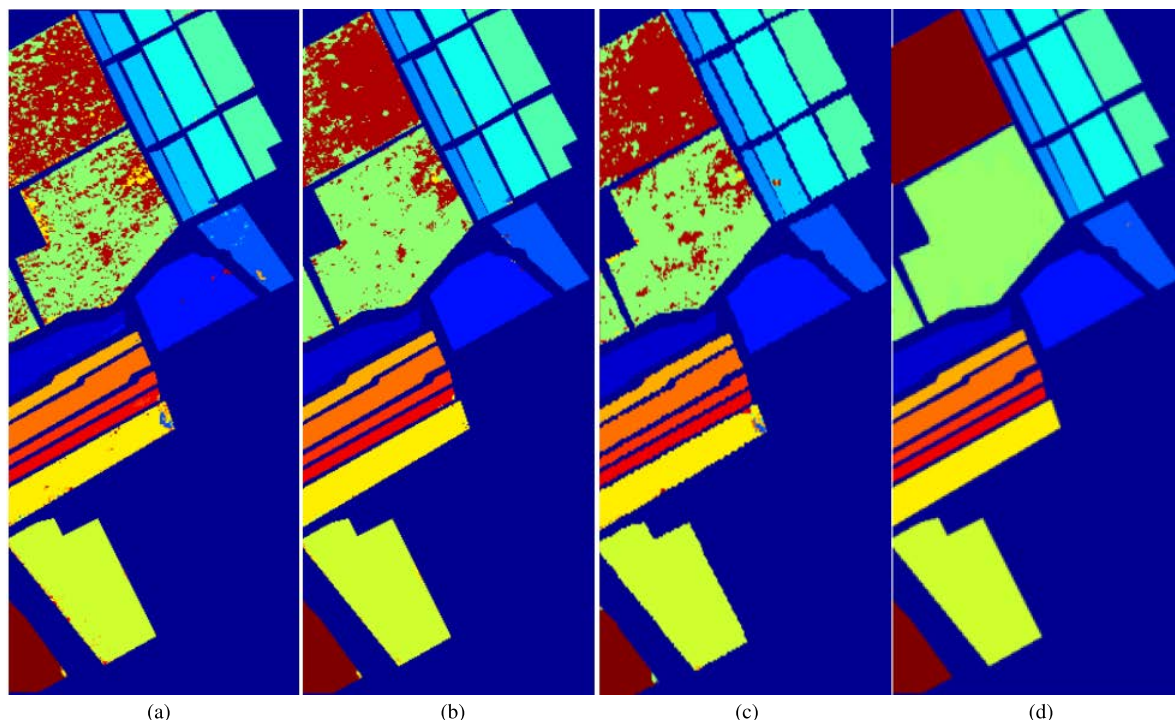


FIGURE 11. Classification maps from the proposed GWJ-NET and the baselines on the Salinas Data, (a) CNN:92.27%, (b) CNN-PPF:96.48%, (c) CD-CNN:96.73%, (d) GWJ-NET:99.45%.

TABLE 7. Classification accuracy (%) of different methods in the Pavia data set over different numbers of training samples.

Dataset	Method	50	100	150	200
Indian Pines	CNN	80.43	84.32	85.3	87.01
	CNN-PPF	88.43	91.72	93.14	93.9
	CD-CNN	84.43	88.27	90.01	94.24
	GWJ-Net	89.15	91.5	96.12	97.56
Salinas	CNN	89.2	89.58	89.6	89.72
	CNN-PPF	92.15	93.88	93.84	94.8
	CD-CNN	82.74	98.58	96.54	95.42
	GWJ-Net	89.0	94.25	96.15	98.24
University of Pavia	CNN	86.39	88.53	90.89	92.27
	CNN-PPF	88.14	93.35	94.97	96.48
	CD-CNN	92.19	93.55	94.15	96.73
	GWJ-Net	90.23	92.56	96.55	98.3

TABLE 8. Classification performance (%) of different initial learning rate for GWJ-NET on the Indian Pines data.

	OA (%)	AA (%)	Kappa Coefficient
0.1	96.57	96.81	0.9589
0.01	98.45	98.79	0.9821
0.001	99.02	99.15	0.9883

methods the accuracy can increase with the number of training samples. From the results, the proposed GWJ-NET still outperforms other methods, namely CNN, CNN-PPF and CD-CNN. Even with a small amount of training data, such as 50 or 100, our proposed network continues to have good classification performance.

Table 9 lists the computational complexity of training and testing for GWJ-NET, CNN, and CNN-PPF. During training,

TABLE 9. Elapsed time (h: Hours, s: Seconds) of training and testing for GWJ-Net method on three data.

		Indian Pines	Salinas	University of Pavia
CNN	Training(h)	0.50	1	0.60
	Testing(s)	0.21	0.26	0.37
CNN-PPF	Training(h)	6	12	1
	Testing(s)	4.46	20.97	16.92
GWJ-NET	Training(h)	0.81	1.52	0.74
	Testing(s)	35	200	112

TABLE 10. Overall accuracy (%) of resnet networks with different layers on Indian Pine data.

Methods	OA(%)
ResNet-4	96
ResNet-6	95
ResNet-8	94
ResNet-10	93
GWJ-NET	98

CNN is faster than the other two because the network size and input size of CNN is smaller than the other two. In the testing process, GWJ-NET will be more time-consuming due to the increased computational burden of using multi-scale feature extraction methods.

In particular, we compare the classification performance by randomly selecting 10% of each class of training samples and ResNet networks with different layers on Indian Pines data. As showed in Table 10, GWJ-NET achieves the highest accuracy due to the combination of spectral feature extraction and “multi-scale” module feature extraction.

V. CONCLUSION

In this paper, we propose a new multi-branch network for hyperspectral image classification based on a multiscale region Gaussian pyramid and a wavelet transform of the spectrum with resNet combined with LSTM. First, the image is segmented so that the image elements to be classified are placed in the center of the rectangle, and then the segmented region is subjected to Gaussian pyramid processing to extract multi-scale features, where multi-scale spatial information and contextual interaction features in specific directions can be obtained to improve the model recognition rate. For spectral data, we perform wavelet transform to reduce the influence of some anomalous spectral segments on the classification and remove some noise that affects the recognition effect, so that finer spectral information can be captured. The advantages of our method come from two aspects: regional multi-scale feature extraction and wavelet transformation of the spectral information to reduce the influence of anomalous spectral values on the classification. Experimental results show that the proposed method based on regional Gaussian pyramid multiscale and spectral information wavelet transform outperforms other recent methods on three datasets.

REFERENCES

- [1] Y. Yuan, Y. Feng, and X. Lu, "Projection-based NMF for hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2632–2643, Jun. 2015.
- [2] W. Li, Q. Du, and B. Zhang, "Combined sparse and collaborative representation for hyperspectral target detection," *Pattern Recognit.*, vol. 48, no. 12, pp. 3904–3916, 2015.
- [3] X. Huang and L. Zhang, "An SVM ensemble approach combining spectral, structural, and semantic features for the classification of high-resolution remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 257–272, Jan. 2012.
- [4] H. Goldberg, H. Kwon, and N. M. Nasrabadi, "Kernel eigenspace separation transform for subspace anomaly detection in hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 4, pp. 581–585, Oct. 2007.
- [5] V. Sharma, A. Diba, T. Tuytelaars, and L. Van Gool, "Hyperspectral CNN for image classification & band selection, with application to face recognition," KU Leuven, ESAT, Leuven, Belgium, Tech. Rep. KUL/ESAT/PSI/1604, 2016.
- [6] K. Makantasis, A. D. Doulamis, N. D. Doulamis, and A. Nikitakis, "Tensor-based classification models for hyperspectral data analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 12, pp. 6884–6898, Dec. 2018.
- [7] D. Landgrebe, "Hyperspectral image data analysis," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 17–28, Jan. 2002.
- [8] Y. Li, W. Xie, and H. Li, "Hyperspectral image reconstruction by deep convolutional neural network for classification," *Pattern Recognit.*, vol. 63, pp. 371–383, Mar. 2017.
- [9] X. Zheng, Y. Yuan, and X. Lu, "Dimensionality reduction by spatial-spectral preservation in selected bands," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 9, pp. 5185–5197, Sep. 2017.
- [10] L. Lin and X. Song, "Using CNN to classify hyperspectral data based on spatial-spectral information," in *Advances in Intelligent Information Hiding and Multimedia Signal Processing*. Cham, Switzerland: Springer, 2017, pp. 61–68.
- [11] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. A. Chanussot, and J. C. Tilton, "Advances in spectral-spatial classification of hyperspectral images," *Proc. IEEE*, vol. 101, no. 3, pp. 652–675, Sep. 2012.
- [12] J. Li, P. R. Marpu, A. Plaza, J. M. Bioucas-Dias, and J. A. Benediktsson, "Generalized composite kernel framework for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 9, pp. 4816–4829, Sep. 2013.
- [13] E. Blanzieri and F. Melgani, "Nearest neighbor classification of remote sensing images with the maximal margin principle," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 6, pp. 1804–1811, Jun. 2008.
- [14] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, Aug. 2004.
- [15] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.
- [16] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Hyperspectral image classification using dictionary-based sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3973–3985, Oct. 2011.
- [17] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Hyperspectral image classification via kernel sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 217–231, Jan. 2013.
- [18] H. Zhai, H. Zhang, L. Zhang, P. Li, and A. Plaza, "A new sparse subspace clustering algorithm for hyperspectral remote sensing imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 1, pp. 43–47, Jan. 2017.
- [19] L. Fang, S. Li, X. Kang, and J. A. Benediktsson, "Spectral-spatial hyperspectral image classification via multiscale adaptive sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 12, pp. 7738–7749, Dec. 2014.
- [20] X. Huang and L. Zhang, "An adaptive mean-shift analysis approach for object extraction and classification from urban hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 12, pp. 4173–4185, Dec. 2008.
- [21] B. Tu, X. Zhang, J. Wang, G. Zhang, and X. Ou, "Spectral-spatial hyperspectral image classification via non-local means filtering feature extraction," *Sens. Imag.*, vol. 19, no. 1, pp. 1–25, Dec. 2018.
- [22] B. Tu, X. Zhang, X. Kang, G. Zhang, J. Wang, and J. Wu, "Hyperspectral image classification via fusing correlation coefficient and joint sparse representation," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 3, pp. 340–344, Mar. 2018.
- [23] X. Kang, S. Li, and J. A. Benediktsson, "Feature extraction of hyperspectral images with image fusion and recursive filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3742–3752, Jun. 2014.
- [24] X. Kang, S. Li, and J. A. Benediktsson, "Spectral-spatial hyperspectral image classification with edge-preserving filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2666–2677, May 2014.
- [25] J. Peng, Y. Zhou, and C. L. P. Chen, "Region-kernel-based support vector machines for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 9, pp. 4810–4824, Sep. 2015.
- [26] K. Zhang, W. Zuo, and L. Zhang, "FFDNet: Toward a fast and flexible solution for CNN-based image denoising," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4608–4622, Sep. 2018.
- [27] X. Lu, X. Zheng, and Y. Yuan, "Remote sensing scene classification by unsupervised representation learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 9, pp. 5148–5157, Sep. 2017.
- [28] X. Ma, H. Wang, and J. Geng, "Spectral-spatial classification of hyperspectral image based on deep auto-encoder," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 9, pp. 4073–4085, Feb. 2016.
- [29] W. Song, S. Li, L. Fang, and T. Lu, "Hyperspectral image classification with deep feature fusion network," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 6, pp. 3173–3184, Jun. 2018.
- [30] H. Huang and K. Xu, "Combing triple-part features of convolutional neural networks for scene classification in remote sensing," *Remote Sens.*, vol. 11, no. 14, p. 1687, Jul. 2019.
- [31] Y. Chen, K. Zhu, L. Zhu, X. He, P. Ghamisi, and J. A. Benediktsson, "Automatic design of convolutional neural network for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 7048–7066, Apr. 2019.
- [32] H. Huang, Y. Duan, H. He, and G. Shi, "Local linear spatial-spectral probabilistic distribution for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 2, pp. 1259–1272, Feb. 2020.
- [33] W. Hu, Y. Huang, L. Wei, F. Zhang, and H. Li, "Deep convolutional neural networks for hyperspectral image classification," *J. Sensors*, vol. 2015, pp. 1–12, Jan. 2015.
- [34] C. Yu, M. Zhao, M. Song, Y. Wang, F. Li, R. Han, and C.-I. Chang, "Hyperspectral image classification method based on CNN architecture embedding with hashing semantic feature," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 6, pp. 1866–1881, Jun. 2019.
- [35] L. Mou, P. Ghamisi, and X. X. Zhu, "Deep recurrent neural networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 7, pp. 3639–3655, Jul. 2017.

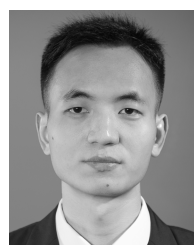
- [36] B. Xue, C. Yu, Y. Wang, M. Song, S. Li, L. Wang, H.-M. Chen, and C.-I. Chang, "A subpixel target detection approach to hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 9, pp. 5093–5114, Sep. 2017.
- [37] N. Falco, J. A. Benediktsson, and L. Bruzzone, "Spectral and spatial classification of hyperspectral images based on ICA and reduced morphological attribute profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 11, pp. 6223–6240, Nov. 2015.
- [38] K. Kavitha and S. Arivazhagan, "A novel feature derivation technique for SVM based hyper spectral image classification," *Int. J. Comput. Appl.*, vol. 1, no. 15, pp. 25–31, 2010.
- [39] H. Zhang, L. Liu, W. He, and L. Zhang, "Hyperspectral image denoising with total variation regularization and nonlocal low-rank tensor decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 5, pp. 3071–3084, May 2019.
- [40] X. Guo, X. Huang, L. Zhang, L. Zhang, A. Plaza, and J. A. Benediktsson, "Support tensor machines for classification of hyperspectral remote sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3248–3264, Jun. 2016.
- [41] J. Liu, Z. Wu, L. Xiao, J. Sun, and H. Yan, "Generalized tensor regression for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 2, pp. 1244–1258, Feb. 2020.
- [42] V. Slavković, S. Verstockt, W. De Neve, S. Van Hoecke, and R. Van de Walle, "Hyperspectral image classification with convolutional neural networks," in *Proc. 23rd ACM Int. Conf. Multimedia*, 2015, pp. 1159–1162.
- [43] H. Gao, Y. Yang, S. Lei, C. Li, H. Zhou, and X. Qu, "Multi-branch fusion network for hyperspectral image classification," *Knowl.-Based Syst.*, vol. 167, pp. 11–25, Mar. 2019.
- [44] W. Ma, Q. Yang, Y. Wu, W. Zhao, and X. Zhang, "Double-branch multi-attention mechanism network for hyperspectral image classification," *Remote Sens.*, vol. 11, no. 11, p. 1307, Jun. 2019.
- [45] Y. Xu, L. Zhang, B. Du, and F. Zhang, "Spectral-spatial unified networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 5893–5909, Oct. 2018.
- [46] J. Yang, Y. Zhao, J. C.-W. Chan, and C. Yi, "Hyperspectral image classification using two-channel deep convolutional neural network," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2016, pp. 5079–5082.
- [47] B. Waske, S. van der Linden, J. Benediktsson, A. Rabe, and P. Hostert, "Sensitivity of support vector machines to random feature selection in classification of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 7, pp. 2880–2889, Jul. 2010.
- [48] I. Kim, R. J. Martins, J. Jang, T. Badloe, S. Khadir, H.-Y. Jung, H. Kim, J. Kim, P. Genevet, and J. Rho, "Nanophotonics for light detection and ranging technology," *Nature Nanotechnol.*, vol. 16, no. 5, pp. 508–524, May 2021.
- [49] X. Cao, J. Yao, Z. Xu, and D. Meng, "Hyperspectral image classification with convolutional neural network and active learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 7, pp. 4604–4616, Jul. 2020.
- [50] T. Wu, M. H. Bae, M. Zhang, R. Pan, and A. Badae, "A prior feature SVM-MRF based method for mouse brain segmentation," *Neuroimage*, vol. 59, no. 3, pp. 2298–2306, Feb. 2012.
- [51] M. Han, R. Cong, X. Li, H. Fu, and J. Lei, "Joint spatial-spectral hyperspectral image classification based on convolutional neural network," *Pattern Recognit. Lett.*, vol. 130, pp. 38–45, Feb. 2020.
- [52] Y. Dong, Q. Liu, B. Du, and L. Zhang, "Weighted feature fusion of convolutional neural network and graph attention network for hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 31, pp. 1559–1572, 2022.
- [53] G. Ortac and G. Ozcan, "Comparative study of hyperspectral image classification by multidimensional convolutional neural network approaches to improve accuracy," *Expert Syst. Appl.*, vol. 182, Nov. 2021, Art. no. 115280.
- [54] S. Prasad and L. M. Bruce, "Limitations of principal components analysis for hyperspectral target recognition," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 4, pp. 625–629, Oct. 2008.
- [55] S. Li, Q. Hao, X. Kang, and J. A. Benediktsson, "Gaussian pyramid based multiscale feature fusion for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 9, pp. 3312–3324, Sep. 2018.
- [56] A. Osadchiy, A. Kamenev, V. Saharov, and S. Chernyi, "Signal processing algorithm based on discrete wavelet transform," *Designs*, vol. 5, no. 3, p. 41, Jul. 2021.
- [57] R. Barzegar, M. T. Aalami, and J. Adamowski, "Coupling a hybrid CNN-LSTM deep learning model with a boundary corrected maximal overlap discrete wavelet transform for multiscale lake water level forecasting," *J. Hydrol.*, vol. 598, Jul. 2021, Art. no. 126196.
- [58] C. Cook, J. L. McKinley, and G. J. O. Beran, "Modeling the α - and β -resorcinol phase boundary via combination of density functional theory and density functional tight-binding," *J. Chem. Phys.*, vol. 154, no. 13, Apr. 2021, Art. no. 134109.
- [59] B. Patnaik, M. Mishra, R. C. Bansal, and R. K. Jena, "MODWT-XGBoost based smart energy solution for fault detection and classification in a smart microgrid," *Appl. Energy*, vol. 285, Mar. 2021, Art. no. 116457.
- [60] Z. Zhong, J. Li, L. Ma, H. Jiang, and H. Zhao, "Deep residual networks for hyperspectral image classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2017, pp. 1824–1827.
- [61] A. P. Gopi, R. N. S. Jyothi, V. L. Narayana, and K. S. Sandeep, "Classification of tweets data based on polarity using improved RBF kernel of SVM," *Int. J. Inf. Technol.*, pp. 1–16, Jan. 2020.
- [62] W. Li, G. Wu, F. Zhang, and Q. Du, "Hyperspectral image classification using deep pixel-pair features," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 844–853, Feb. 2017.
- [63] K. Makantasis, K. Karantzaos, A. Doulamis, and N. Doulamis, "Deep supervised learning for hyperspectral data classification through convolutional neural networks," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2015, pp. 4959–4962.
- [64] H. Lee and H. Kwon, "Going deeper with contextual CNN for hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 26, no. 10, pp. 4843–4855, Oct. 2017.
- [65] S. Mei, J. Ji, J. Hou, X. Li, and Q. Du, "Learning sensor-specific spatial-spectral features of hyperspectral images via convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4520–4533, Aug. 2017.



YIGANG TANG received the B.E. degree from the Guilin University of Technology, Guilin, China, in 2020, where he is currently pursuing the M.S. degree. His research interests include remote sensing image process and hyperspectral image classification.



XIAOLAN XIE (Member, IEEE) received the M.S. degree in computer science from Shanghai Maritime University, in 2001, and the Ph.D. degree in mechanical manufacturing and automation from Xidian University, China, in 2009. She is currently a Senior Visiting Scholar at Middlesex University, U.K. She is also the Deputy Director of the Guangxi Key Laboratory of Embedded Technology and Intelligent System and the Dean of the School of Information Science and Engineering, Guilin University of Technology. She has published more than 100 scientific research articles, including more than 50 SCI/EI papers and more than 20 Chinese core journals. Her research interests include cloud computing, big data, intelligent computing, and manufacturing informatization. She is a fellow of the American Computer Society (ACM), the Chinese Computer Society (CCF), the Cloud Computing Expert Committee of the Chinese Institute of Communications (CIC), the High Performance Computing Expert Committee of CCF, and the Collaborative Computing, Distributed Computing and Processing Expert Committee of CCF. She is also the Director of the International Institute of Engineering and Technology (IETI).



YOUHUA YU (Member, IEEE) received the B.E. degree in digital media technology from the Guilin University of Electronic Technology, Guilin, China, in 2019. He is currently pursuing the master's degree in computer science and technology with the Guilin University of Technology, Guilin. His main research interests include hyperspectral image processing and cross-scene hyperspectral image classification.