# Comprehensive Approach to Hyperspectral Image Analysis: Multiscale Feature Extraction and Spectral Imaging for Classification

## B.Nancharaiah[1], M.V.Subrahmanyam[2], R.Manasa[3],J.snehitha[4],K.Jagan Mohan Achari[5]

[1]B.Nancharaiah:Professor, Dept. of Electronics and Communication Engineering,
Usharama College of Engineering and Technology,Telaprolu,Andhra Pradesh-521109,India.
[2]M.V.Subrahmanyam,[3]R.Manasa,[4]J.Snehitha,[5]K.Jagan Mohan Achari- Students of Usharama College of Engineering
and Technology,Telaprolu,Andhra Pradesh-521109,India.

---------------------------------------------------------------------***---------------------------------------------------------------------

**Abstract -** *We present a novel approach for hyperspectral image classification, utilizing a dual-branch architecture for concurrent spatial and spectral feature extraction. Prior to feature extraction, we employ principal component analysis (PCA) to reduce data dimensionality, with varying degrees of downsampling across the two branches. Spatial information is captured through a multiregion piecewise Gaussian pyramid downsampling method, generating multiscale and multiresolution image data. Enhanced ResNet networks are then employed to extract spatial features, enabling the extraction of contextually specific features inherent to hyperspectral images. Meanwhile, spectral information is processed using a unique imaging spectral data technique, involving initial PCA-based dimensionality reduction followed by expansion into N×N images. A dedicated ResNet network, tailored with a distinct number of layers, is utilized for spectral feature extraction, addressing the challenge of variations in spectral data. Subsequently, the spatial and spectral features extracted from the dual-branch network are integrated and fed into a fully connected network for classification, resulting in significantly enhanced classification accuracy. Experimental validation on two benchmark datasets demonstrates the effectiveness of our proposed method, showcasing substantial improvements in classifier accuracy compared to existing approaches.*

*Keywords*:Deep learning, Gaussian pyramid, hyperspectral image (HSI) classification, multiscale feature extraction.

## 1. INTRODUCTION

With the continuous advancements in science and technology, hyperspectral imaging, also referred to as imaging spectroscopy, has experienced rapid progress. This technology involves remote sensing satellites capturing tens of thousands of narrow spectral bands emitted or reflected from a given area. This capability enables the acquisition of more detailed spatial and spectral information compared totraditional panchromatic and multispectral remote sensing images, allowing for improved differentiation between various materials. Due to its advantages in attribute recognition, hyperspectral image (HSI) processing technology has seen significant development and widespread application in various domains as rural planning, environmental monitoring, urban planning, vegetation coverage assessment, mineral extraction, national defense infrastructure and precision agriculture.

Consequently,classification techniques in the realm of hyperspectral imagery have also advanced rapidly, playing a crucial role in remote sensing applications. HSI classification, a focal point of recent research, entails assigning specific class labels to individual pixels based on their spatial and spectral characteristics. However, the complexity of HSIs, characterized by numerous spectral bands with high correlation information redundancy, poses computational challenges. Moreover, distinguishing between different materials with similar or identical spectra further complicates classification tasks.

Currently, methods for hyperspectral image (HSI) classification can be broadly categorized into those leveraging spectral information alone and those incorporating joint spatial–spectral features. Spectral information-based classification methods rely solely on the spectral dimension of HSIs, disregarding spatial pixel correlations. Examples include support vector machine, random forest, sparse representation, and similar techniques. In contrast, approaches integrating joint spatial–spectral features have shown improved classification performance, utilizing methods such as edge-preserving filtering, multiscale adaptive strategies, low-rank Gabor filtering, and hierarchical guided filtering with nearest-neighbor regularization subspaces. However, these methods heavily rely on manually crafted features, leading to limited classification performance as they may not fully capture the complex content within HSIs. In comparison to traditional shallow methods, deep learning techniques offer enhanced representation and generalization capabilities, capable of extracting deeper image features and achieving more discriminative representations. Consequently, deep learning methods have gained significant traction in HSI classification, including convolutional neural networks (CNNs), two-channel networks, spectral spatial attention networks, and related approaches. Despite their

effectiveness, CNN-based models often suffer from redundancy in spatial dimension information, affecting performance. To address this, some researchers have proposed innovative solutions like the octave convolution network for preprocessing hyperspectral data, followed by spectral and spatial attention networks for feature extraction, fusion, and softmax classification. Additionally, efforts such as the spectral–spatial unified network (SSUN) combine deep and shallow convolutional layers to mitigate information loss, while techniques like the deformable HSI classification network (DHCNet) introduce adaptively adjustable deformable convolution sampling positions. Other methodologies like the contextual deep CNN (CD-CNN) optimize local contextual information exploration, and methods like R-PCA and CNN-PPF incorporate CNNs for spectral and spatial information encoding and classification tasks. Meanwhile, approaches such as multiscale CNN (MS-CNN) and those employing support vector machines (SVMs) and random feature selection (RFS) aim to synergize high-dimensional data and multimodal classes. These methods encompass both unsupervised and supervised learning paradigms, with deep neural networks witnessing rapid advancements, especially in feature extraction and classification tasks within HSI analysis.

In recent years, the widespread adoption of deep learning methods, facilitated by rapid advancements in hardware, has revolutionized various fields.

Following the introduction of convolutional neural networks (CNNs), neural networks have experienced remarkable progress in image processing, giving rise to a multitude of CNN-based image classification models. This surge in deep learning has significantly advanced image classification capabilities and has inspired the exploration of intelligent and efficient hyperspectral classification methods discussed in this paper. Unlike traditional machine learning algorithms, which typically focus on learning low-level features such as object outlines or textures, deep learning excels in capturing complex, high-dimensional features essential for hyperspectral classification tasks. As such, deep learning emerges as the optimal choice for addressing the intricate demands of hyperspectral classification, where precise recognition of numerous features is paramount. Deep learning not only encompasses the learning of low-level features but also extends to capturing subtle nuances of target objects, presenting a robust foundation for hyperspectral classification tasks. Moreover, hyperspectral imaging (HSI) data exhibit high-dimensional spectral characteristics with distinct waveband responses. Leveraging the residual network architecture, which effectively addresses the challenge of network degradation as networks deepen, allows for the extraction of profound features. This makes residual networks particularly well-suited for capturing the intricate features present across various hyperspectral

bands. Additionally, the residual network's inherent capacity to handle increasingly complex data representations enables it to effectively navigate the rich and diverse spectral information inherent in HSI datasets. This capability enhances the model's ability to discern subtle differences among spectral signatures, consequently bolstering the accuracy and reliability of hyperspectral classification outcomes.In addition to its prowess in capturing subtle features of target objects, deep learning's application in hyperspectral classification benefits from the unique characteristics of spectral data within hyperspectral images (HSIs). These datasets are distinguished by their high dimensionality, often comprising hundreds of narrow spectral bands, each exhibiting distinct waveband response features. The residual network architecture, known for its ability to mitigate the challenge of degradation in deep networks, proves highly advantageous in this context. By leveraging residual connections, which enable the direct propagation of information across network layers, residual networks excel in extracting deep-level features from hyperspectral bands. This capability is particularly valuable in hyperspectral analysis, where discerning nuanced variations in spectral signatures is crucial for accurate classification. Furthermore, the adaptability of residual networks to intricate data representations facilitates the exploration of complex hyperspectral datasets, empowering the model to uncover subtle patterns and distinctions across spectral bands. As a result, the integration of deep learning techniques, especially through architectures like residual networks, offers a promising avenue for enhancing the precision and robustness of hyperspectral classification methodologies.The extensive array of spectral bands, characterized by their narrowness and multitude, poses challenges in hyperspectral imaging (HSI).
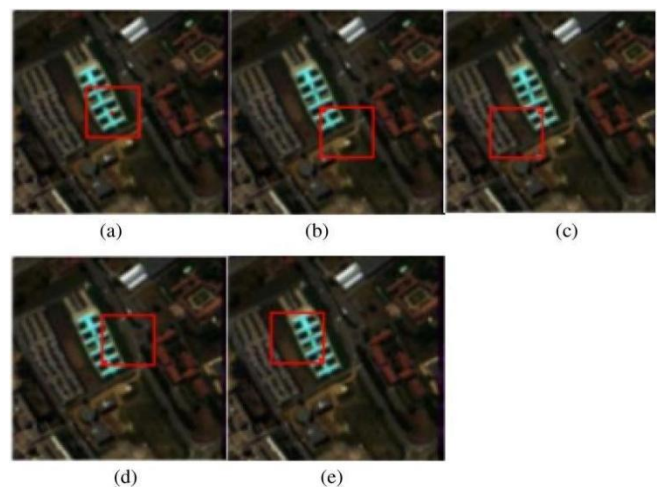


**Fig. 1.** Diverse regions of a pixel of interest. (a) Global region. (b) Upper-left region. (c) Upper-right region. (d) Bottom-left region. (e) Bottom-right region.
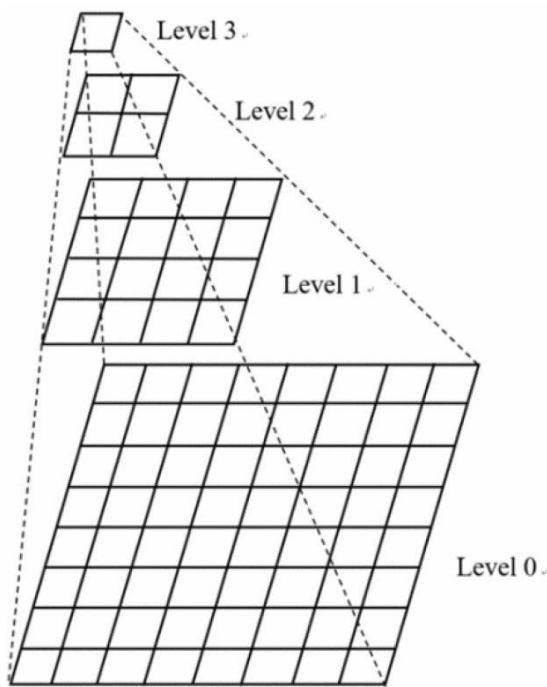
**Fig. 2.** Gaussian pyramid structure.

These challenges include the high correlation among adjacent bands, resulting in information redundancy, as well as the substantial volume of HSI data necessitating computationally intensive operations. Employing data dimensionality reduction techniques becomes imperative in mitigating these challenges. Notably, such techniques not only alleviate data redundancy but also curtail the computational burden associated with processing large-scaleHSIdatasets.
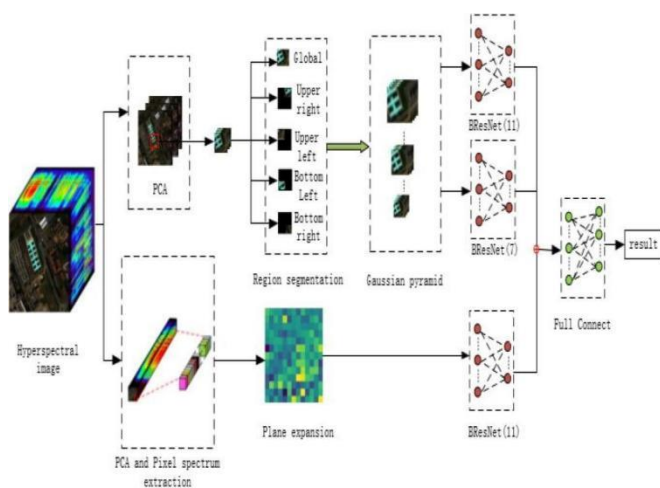


**Fig. 3.** Overall flowchart of the proposed DRB-ResNet.
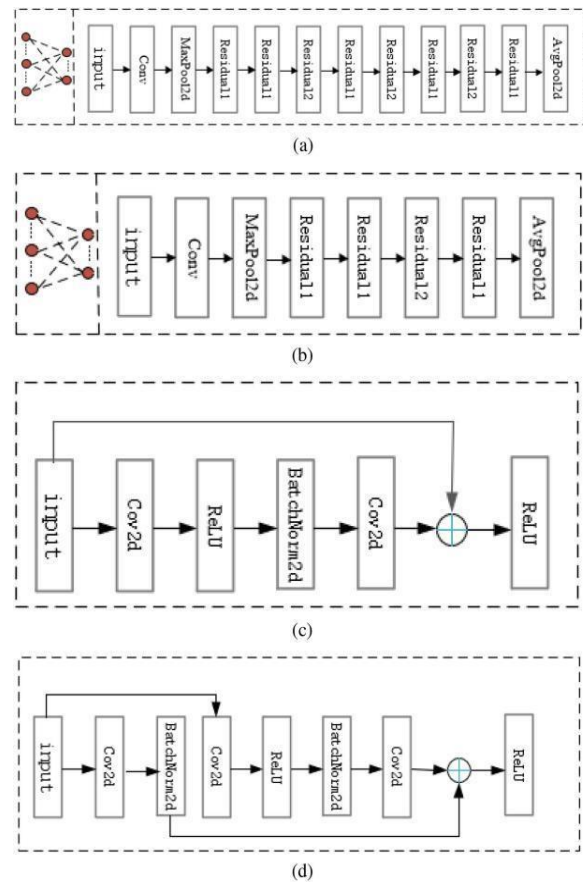


**Fig. 4.** DRB-ResNet structure and residual block structure. (a) DRB-ResNet(11). (b) DRB-Resnet (7). (c) Residual block 1. (d) Residual block 2.

Additionally, hyperspectral imaging (HSI) presents numerous challenging issues, including band redundancy and significant noise interference. To address these challenges, various dimensionality reduction techniques have been developed to extract pertinent information from images while simultaneously reducing computational complexity. These methods encompass linear principle analysis, probabilistic principal component analysis (PPCA), kernel principal component analysis (KPCA),and neighbor-independent component analysis (ICA). In this study, we opted for the PCA method, a widely utilized dimensionality reduction approach. PCA transforms a set of potentially correlated variables into a set of linearly uncorrelated variables through orthogonal transformation, with the resulting variables termed principal components. Notably, PCA stands out as one of the simplest and most efficient dimensionality reduction methods available. Hence, this paper employs the PCA algorithm to achieve dimensionality reduction in HSI processing.

Furthermore, there exists a significant correlation between adjacent pixels in an image, encompassing both texture and grayscale levels. Objects with small sizes or low contrasts often necessitate higher resolutions for

accurate observation, whereas larger objects or those with high contrasts may suffice with lower resolutions. Such variations are commonplace in hyperspectral images (HSIs), necessitating multiresolution processing strategies. Gaussian pyramids, a type of multiscale representation in images, offer an effective and intuitively simple framework for analyzing images across multiple resolutions. Initially employed in machine vision and image compression, an image pyramid comprises a series of images arranged in a hierarchical structure with progressively decreasing resolutions, all derived from the same original image collection. This hierarchical structure is achieved through iterative downsampling until a predefined termination condition is met. At the base of the pyramid lies a high-resolution representation of the image, while the apex corresponds to a low-resolution approximation. Analogously, each layer of the pyramid can be likened to a tier, with a higher number of layers yielding smaller images and lower resolutions. Gaussian pyramids facilitate the scaling of images to obtain multiscale representations, enabling subsequent multiscale analyses.To address the high-dimensional nature of features in hyperspectral images (HSIs), we can employ the PCA technique to reduce the dimensionality of the data before engaging Gaussian pyramids for multiscale processing of HSI data.
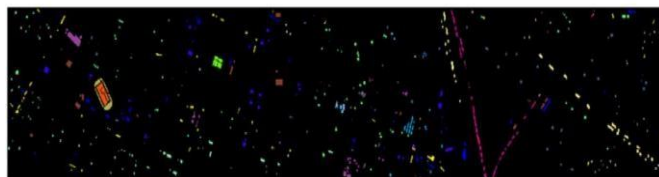


**Fig. 5.** For experimental datasets. (a) False-color image of the Houston data. (b) Ground truth of the Houston data.

We employ Gaussian pyramids and multiregion pixel segmentation to preprocess the original data. Initially, a small rectangular area traverses the original image, ensuring that points for classification reside at the center of each rectangular

area. By moving the rectangulararea to positionthe points at the corners, five rectangular images are generated. Subsequently, Gaussian pyramid processing is applied to the segmented regions to yield multiscale image data across different regions. For spectral data, we utilize the PCA algorithm to reduce data dimensionality to a predefined level before expanding it intoa 2- D image format. Notably, we introduce spectral imaging methodology for the first time to represent spectral data as image data, followed by spectral feature extraction via convolutional networks. Moreover, employing multiregion segmentation and Gaussian pyramids allows for the extraction of spatial information from pixels, facilitating the fusion of local and global spatial features to capture more nuanced spatial characteristics. The classification of HSIs using this approach demonstrates outstanding performance. The primary contributions of this studycan besummarizedas follows: Firstly, we partition the data into multiple regions to capture diverse spatial information corresponding to the same pixel across these regions. Subsequently,we apply the Gaussian pyramid downsampling method to the segmented data, enabling the extraction of multiscale data and spatial information across different scales. For spectral data processing, we introduce a novel spectral imaging technique, transforming spectral data into image format, and utilize a feature extraction network to extract intricate spectral details effectively. Secondly, we introduce a dual-branch feature extraction network architecture, leveraging two separate branches to extract spectral and spatial features, respectively. This dual-branch approach enhances the classification accuracyand robustness of the network by enabling the extraction of distinct features through parallel pathways.The above part is combined with the extracted features of the below part and made as the fully connected layer and give n to the classifier to enhance the classification accuracy of theclassifier.

## 2. METHODOLOGY

Our hyperspectral image (HSI) classification method comprises two key components. The first component employs the PCA algorithm to reduce the dimensionality of the original data and subsequently divides the resulting data into five regions. Pixels designated for classification are positioned at the center of rectangular areas, while additional pixels for classification are placed at the four corners of each rectangular area upon movement. Gaussian pyramid downsampling is then applied to the resulting image data from the five rectangular areas, yielding multiscale and multiresolution image data.The second component begins with PCA dimensionality reduction on the original data. Subsequently, spectral data for each pixel slated for classification is extracted and expanded into an image format. Distinct ResNet networks with varying structures are utilized to extract spatial and spectral features, respectively. Finally, the spatial and spectral feature extraction networks are combined, and

the resultant features are fed into a fully connected network to obtain the final classification result.The proposed DRB-ResNet model leverages the aforementioned five regions as input, extracting spatial and spectral features via the "multiscale" module within a meticulously designed ResNet network and the "spectral imaging" module. This innovative approach enhances feature extraction capabilities, ultimately improving classification accuracy and robustness Details are described in the following Section II-A–C.

## 2.1. Multiscale Feature Extraction of Multiregion Gaussian Pyramid.

We employ the PCA algorithm to reduce the dimensionality of hyperspectral image (HSI) data. PCA extracts the primary feature components from the data while discarding redundant information, thereby reducing data dimensionality. This process is essential for mitigating the curse of dimensionality and enhancing data density.
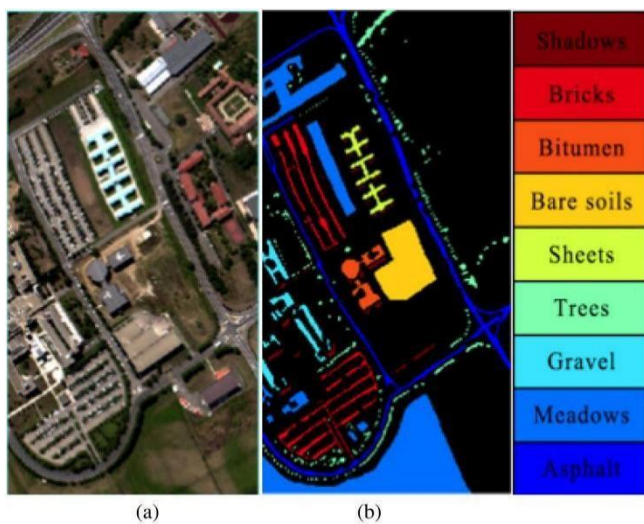


**Fig. 6.** For experimental datasets. (a) False-color image of the University of Pavia data. (b) Ground truth of the University of Pavia data.

Particularly in the presence of noise,PCA identifies eigenvectors associated with smallest eigenvalues which often corresponds to noise components.By discarding these eigen vectors,PCA effectively reduces the noise levels in the data.Furthermore,PCA not only compress the data into lower dimensions but also ensures that the features of the data,post dimensionalityreduction mutually independent.Dimensionality reduction serves to preserve the most essential features within high-dimensional data while filtering out noise and irrelevant features, leading to enhanced data processing efficiency. In practical applications, reducing the dimensionality of data within an acceptable range of information loss can significantly reduce storage costs and processing time. Following PCA

dimensionality reduction, each pixel designated for classification undergoes segmentation, wherein a small area of 11×11 centered around the pixel is extracted. Subsequently, these segments are repositioned within a square bounding box, ensuring that the pixels to be classified occupy the upper-left, upper-right, lower-left, and lower-right corners of the box, along with the inclusion of the initially extracted global block. This approach facilitates the extraction of spatial information from hyperspectral imaging (HSI) data.For regional blocks in different directions, more detailed spatial relationships can be obtained, and even the dissimilarities or similarities between different categories shown in the Fig. 1(a) indicates the global segmentation area of the pixels to be classified, Fig. 1(b) indicates that the pixels to be classified are in the upper-left corner of the segmentation area, Fig. 1(c) indicates that the pixels to be classified are placed in the upper-right of the segmentation area corner, Fig. 1(d) indicates that the pixels to be classified are placed in the lower-left corner, and Fig. 1(e) indicates that the pixels to be classified are placed in the lower-right corner of the segmentation area.
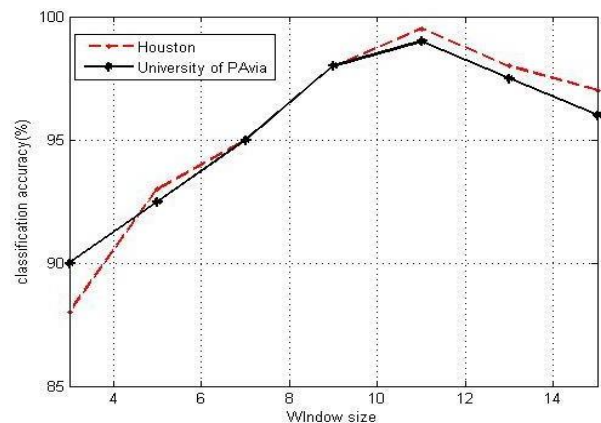


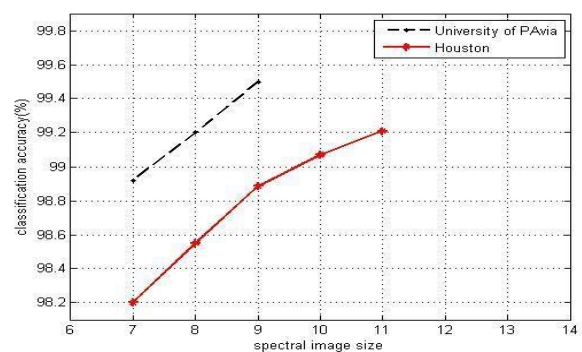**Fig. 7.** Classification accuracy (%) of different window size.



**Fig. 8.** Classification accuracy (%) of three datasets at different spectral imaging sizes.

By situating the feature extraction network in this manner, it can acquire heightened spatial intricacies around individual pixels, particularly those residing at the interface of two distinct classes. Pixels belonging to different categories at these boundaries often exhibit high similarity. However, through the implementation of multiregion segmentation, disparities between pixels within various regions are accentuated, thereby facilitating the network's ability to discern differences between the two categories. Although multiregion segmentation does impose an increased computational load, it concurrently enhances classification accuracy.Gaussian pyramid decomposition entails generating a sequence of images using a low-pass sampling filter, resulting in progressively lower sampling densities compared to the preceding image. In this context, Gaussian filtering and downsampling are applied to the data derived from the PCA algorithm to diminish dimensionality and segment the initial hyperspectral imaging (HSI) data. Initially, the base image is convolved with a Gaussian kernel, followed by downsampling to produce the subsequent higher-level image. Subsequently, this image serves as input for further iterations of convolution and downsampling operations, yielding successive higher-level images.

The Guassian pyramid is formed through repeated iterations. As shown in the Fig. 2, it is a Gaussian pyramid structure.The overall structure of the image is obtained by performing the same steps for downsampling to obtain the processed image. The same processing is applied to the HSI so that the HSI can be analyzed at multiple scales, and different spatial features can be available at different scales.It classifies with different layers of the ResNet Network. Each layer of the pyramid is produced by the following formula L1 represents the upper layer that has to be generated and Ll-1 is the lower part 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:

$$L_1(\mathrm{h},k)=\sum_{u=-d}^{d}\sum_{v=-d}^{d}F(b,v)*L_{x-1}(2h-1-b,2k-1-v).$$

The * is the convolution operation. $\mathbf{x} \in \{1, 2, 3, \cdots, X\}$, $X$ is the total amount of layers of the Gaussian pyramid. $F(b, v)$ is $(2d + 1) \times (2d + 1)$ Gaussian window, which can be defined

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

where $\Upsilon$ is the variance of the Gaussian filter in the Gaussian pyramid operation. The Gaussian pyramid is formed by stacking one picture of different scales $\{L_1, L_2, L_3, \cdots, L_l\}$, which is generated by the above equation.

Spectral feature changes cannot reflect the similarity between features or distinguish the boundaries of features, and multiscale feature analysis can effectively solve this problem and improve the accuracy of HSI classification. The Gaussian pyramid can effectively transform the HSI into a multiscale structure, which can extract features at multiple scales by stepwise sampling or downsampling. We obtain multiscale images by performing Gaussian pyramid processing on each segmented area, and then arrange the images of the same scale into a new dataset, and finally put them into the feature extraction network for multiscale feature extraction.

## 2.2. Spectral Feature Extraction

Each material exhibits distinct spectral characteristics in terms of absorption, reflection, and radiation. Objects within the same spectral region manifest varying degrees of reflection and absorption of radiation spectra, and even identical objects exhibit noticeable differences in reflection across different spectra. Furthermore, environmental conditions such as sunlight angles contribute to variations in reflection and absorption spectra for the same object at different times and locations. During satellite sensor imaging processes, practical factors such as uncompensated atmospheric effects, sensor errors, and solar zenith angles can induce alterations in hyperspectral imaging (HSI) spectra . Additionally, biological dynamics represent another significant factor. For instance, benthic organisms inhabiting underwater environments may display differing reflectance spectra in water . Moreover, the presence of pests and diseases can also cause distinct spectral variations in otherwise similar crops. In these cases, even when the sensor is utilized to image
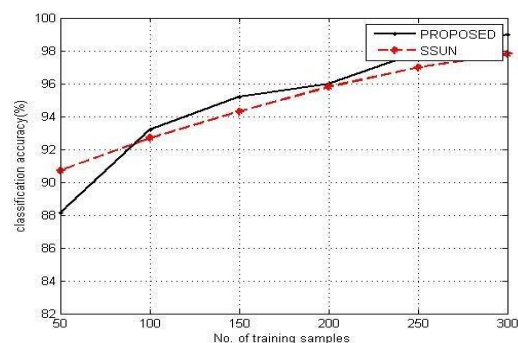
**Fig. 9.** Classification accuracy (%) of different methods in the Houston dataset over different numbers of training samples

may vary. For this kind of problem, this article divides each pixel in the dataset into $1 \times 1 \times L$ ($L$ represents the number of bands) and then performs the dimensionality

reduction. Expand the reduced image into a picture. By unfolding it into a picture, discrete spectra of similar crops due to anomalies in certain spectral bands can be avoided. And by unfolding it into a picture, the network model can ignore some features that are unique in the spectrum of similar crops owing to spectral anomalies, which can improve its robustness and accuracy. Expand it into a picture, make each band value the value of a pixel on the picture, and then extract features through a feature extraction network. Feature extraction in multiple regions can reduce the incidence of abnormal spectral segments on classification, thereby improving classification accuracy.

## 2.3. Proposed Model DRB-ResNet

This study utilizes a residual network model augmented with batch normalization. This technique standardizes batch data distribution, averting gradient vanishing and expediting calculations. It adjusts activation means towards 0 and standard deviations towards 1. The formula is as follows:

$$Z = \gamma \frac{Z - E\left(Z\right)}{\sqrt{\text{Var}\left(Z\right) + \xi}} + \beta.$$

Among them, Z represents the output after batch normaliza- tion, E (Z) and Var (Z) represent the expectation and variance of Z, respectively, and $\gamma$ and $\beta$ are the hyperparameters that need to be learned to get the best. For HSI data with different window sizes, DRB-ResNet networks with different structures are used. After multiscale processing of the data in a Gaussian pyramid, images of different sizes are generated. In order to prevent the use of the same network structure, the obtained features cannot be changed due to the small size of the picture, and different network structures are used for pictures of different sizes. For large-sized images, a deeper network is used to extract features, while for small-sized images, a shallower network is used to extract features to prevent the network from overfitting.
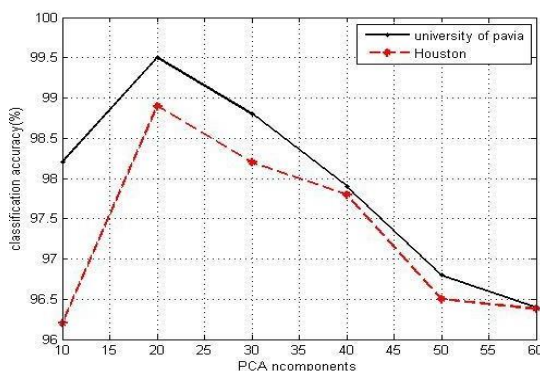


**Fig. 10.** Classification accuracy (%) of three data sets at different n_components of PCA.

This article mainly uses two networks of different depths. The overall network structure is shown in Fig. 3. The upper part performs PCA dimensionality reduction on selected parts of the original data and then divides the data into regions. The pixels that need to be classified are placed in the four corners and the center of the selected size area, and finally, a multiscale image is obtained through Gaussian filtering and downsampling. Images of the same size are sequentially arranged into a DRB-ResNet network with a specified depth for training, the feature extrac- tion data of each region are added, and finally, the results are output through the fully connected network. This article uses a four-layer Gaussian pyramid for multiscale analysis. The first and second layers use a deep network, such as Fig. 4(a), the third layer uses a shallow network, such as Fig. 4(b), the last layer is directly connected to the output because the image size is too small. The lower half of the entire network is used to extract spectral features. First, the spectral information of each pixel to be classified is dimensionally reduced, and then expanded into $N{\times}N$ pictures, and then put into the deep DRB-ResNet network for feature extraction, and finally combined The feature extraction results of the upper half and the lower half is used to obtain the final classification result using the fully connected network. Fig. 4(a) and (b) are deep and shallow feature extraction networks used to extract features, consisting of two residuals in Fig. 4, which are combined from (c) and (d). In practical scenarios, orientation-based regions exhibit minimal cancellation owing to the existence of positional anisotropy properties. Nonetheless, it's evident that classification outcomes vary significantly across different block types, suggesting diverse contextual distributions possess varying generalization capabilities for classification tasks. The experimental findings on fusion strategies undeniably highlight potential for improvement. Thus, the proposed DRB-ResNet model can enhance accuracy by providing a more resilient feature representation, considering multiple potential distributions.

## 3. EXPERIMENTAL RESULTS

In this section, we validate the effectiveness of our pro- posed DRB-ResNet on two generic datasets and compare it with contemporary state-of-the-art HSI classification methods. The experiments in our study utilize Python for programming, with network models constructed using the PyTorch deep learning framework within the Python environment. PyTorch is a widely used open-source machine learning library known for its flexibility and efficiency in customizing deep learning models.

TABLE 1-NUMBER OF TRAINING AND TESTING SAMPLES FOR THE HOUSTON DATASET

| # | Class | Training | Test |
|---|-------|----------|------|
| 1 | Health grass | 300 | 951 |
| 2 | Stressed grass | 300 | 965 |
| 3 | Synthetic grass | 300 | 398 |
| 4 | Trees | 300 | 920 |
| 5 | Soil | 300 | 966 |
| 6 | Water | 300 | 32 |
| 7 | Residential | 300 | 894 |
| 9 | Commercial | 300 | 858 |
| 10 | Road | 300 | 886 |
| 11 | High Way | 300 | 977 |
| 12 | Rail Way | 300 | 941 |
| 13 | Parking lot 1 | 300 | 995 |
| 14 | Parking lot 2 | 300 | 214 |
| 15 | Tennis court | 300 | 140 |
| 16 | Running track | 300 | 392 |
| - | Total | 4500 | 17 029 |

TABLE 2-NUMBER OF TRAINING AND TESTING SAMPLES FOR THE PAVIA DATASET

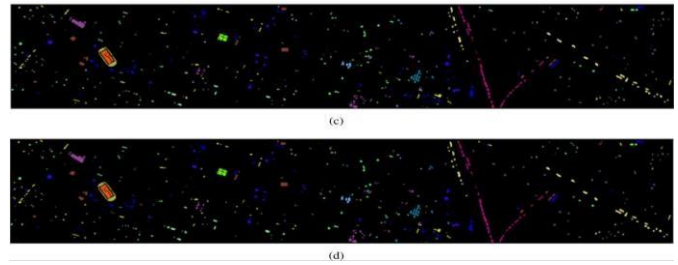| # | Class | Training | Test |
|---|-------|----------|------|
| 1 | Asphalt | 500 | 6131 |
| 2 | Meadows | 500 | 18 149 |
| 3 | Gravel | 500 | 1599 |
| 4 | Trees | 500 | 2564 |
| 5 | Sheets | 500 | 845 |
| 6 | Baresoil | 500 | 4529 |
| 7 | Bitumen | 500 | 830 |
| 8 | Bricks | 500 | 3182 |
| 9 | Shadows | 500 | 447 |
| - | Total | 4500 | 38 276 |

## 3.1. Datasets

To validate the effectiveness of our proposed method, we verified the classification performance of our proposed DRB- RESNET on three generic datasets (the Houston dataset, Uni- versity of Pavia dataset, and), as shown in Figs. 5 and 6. The pseudocolor maps, label maps, and their category representation colors for the two datasets. For these two datasets. the Houston dataset consists of 349 × 1905 pixels. Houston dataset, provided by the 2013 IEEE GRSS Data Fusion

Competition, was acquired on the University of Houston campus and adjacent areas with a compact airborne spectral imager (manufactured by ITRE Canada). It collects a dataset of 144 spectral channels covering 0.364–1.046 $\mu$m. For the 15 different classes in the Houston dataset, the sample sizes for the training and test sets are shown in Table I.

The second dataset (University of Pavia) was acquired by the ROSIS sensor. It has 610×340 pixels and the sensor has a total of 115 bands. after processing, the University of Pavia dataset contains 103 bands with a coverage of 0.43–0.86 $\mu$m and a spatial resolution of 1.3 m.

About 42 76 pixels in the dataset have been labeled and divided into nine categories, as shown in Table II.



**Fig. 11.** Classification maps from the proposed DRB-RESNET and the baselines on the Houston Data. (c) SSUN: 97.21%. (d) DRB-RESNET: 99.34%.

### 3.2. Learning the Proposed DRB-ResNet

For the pixels used for training and testing, we take a range of rectangular regions around them. The training and testing pixels are placed in the center of the rectangle, and on this basis, the rectangular area is moved to place the training and testing pixels in the four corners of the rectangle, to extract the spatial features around the pixels, and finally, the multi- region image is put into the convolutional layer for feature extraction, and the size of the region will affect the final classification performance of our proposed DRB-ResNet, so we empirically set in this article, we empirically set the region size to 11×11, and finally discuss the effect of different size rectangles on the performance of DRB-ResNet, as shown in Fig.7.



(a)                    (b)
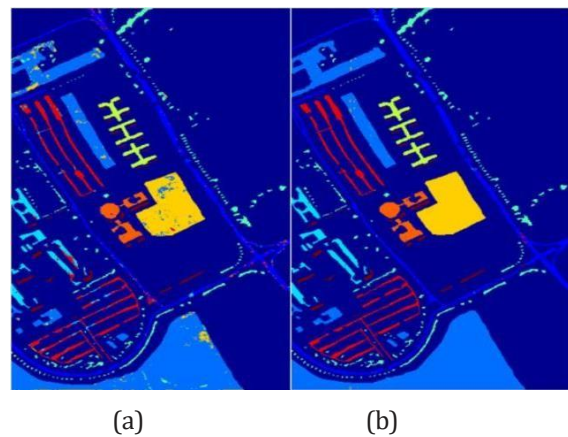
**Fig. 12.** Classification maps from the proposed DRB-RESNET and the baselines on the University of Pavia Data (a)SSUN: 99.45%. and (b)DRB-RESNET: 99.63%.
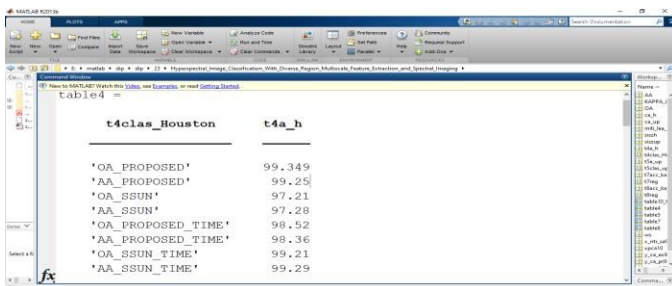
In addition, the optimizer in this article uses Adam, which combines the updated ideas of adaptive learning strategies Ada- grad and Momentum. The Adam algorithm is derived from traditional gradient descent. All weights are updated at a single learning rate. The rate does not change during the training process, and Adam designs

independent adaptive learning al- gorithms for different parameters by combining the first-order moment estimation with the second-order moment estimation of the gradient. The author of the Adam algorithm defines it as two kinds of gradient descent. An extended collection of advantages. In this article, we set the initial learning rate to 0.001 and different batch sizes. Fig. 7 shows the classification performance of the rectangular area under different window sizes.That is from 3×3 to 15×15.
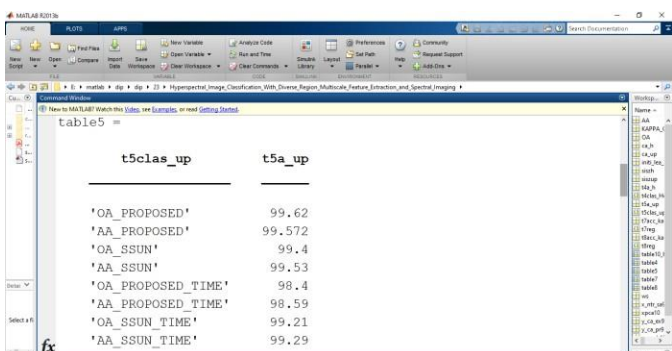
When the window size is 11×11, the performance tends to be satisfactory. When the window size is 11×11, it is not the best window size for all experimental datasets. For example, the red curve indicates that the square area 11×11 is the best window size of the University of Pavia dataset, and the blue curve indicates that the best window size of the Houston dataset is 11×11, and When converting spectral data into pictures, the best picture size of the Houston dataset is 11×11, and the best picture size of the University of Pavia dataset is 9×9. The number of bands in the two datasets in this article is different, the expansion size and the classification effect of different sizes are also different. For example, the Pavia dataset can only be expanded to 9×9, and the Houston dataset can only be expanded to 11×11, as shown in Fig. 8. Therefore, we choose a relatively large size (for example, 11×11, while the Pavia dataset expands to 9×9) among the allowed hardware resources to prevent the waste of hardware resources and beyond the scope of hardware use.

TABLE 3-COMPARISON OF THE CLASSIFICATION ACCURACY(%) AMONG THE PROPOSED METHOD AND THE BASELINES USING THE HOUSTON DATA



TABLE 4-COMPARISON OF THE CLASSIFICATION ACCURACY(%) AMONG THE PROPOSED METHOD AND THE BASELINES USING THE UNIVERSITY OF PAVIA DATA



We tested different initial learning rates on the Houston dataset for the trained hyperparametric learning rates, as shown in Table X, we can see the performance of different initial learn- ing rates for OA, AA, and Kappa Coefficient. We can conclude that the size of the initial learning rate has quite a big impact on the final results, and too large a learning rate will reduce the final classification results. In this article, the learning rate is set to 0.001 for the best classification performance. Therefore, in the subsequent experiments, our proposed DRB-ResNet will use 0.001 as learning rate.

The experimental results of the four methods based on the Houston dataset are shown in Fig. 9 The number of initial training samples for each class ranges from 50 to 300, with a pairwise interval of 50. We can get that, for all methods compared, the classification accuracy increases as the number of training samples increases. When the number of training samples is in the range of 150–300, our method can always maintain the best classification performance.

We conducted experiments on the parameter n_components (dimension reduction) of the PCA algorithm in three datasets, as shown in Fig. 10, we found that as the data dimension increased from 10 to 20, the data classification accuracy also increased, but when the data dimension increased from 20 to 60, the data classification accuracy begins to decrease. Therefore, when reducing the dimensionality of the dataset, set the n_components parameter to 20.
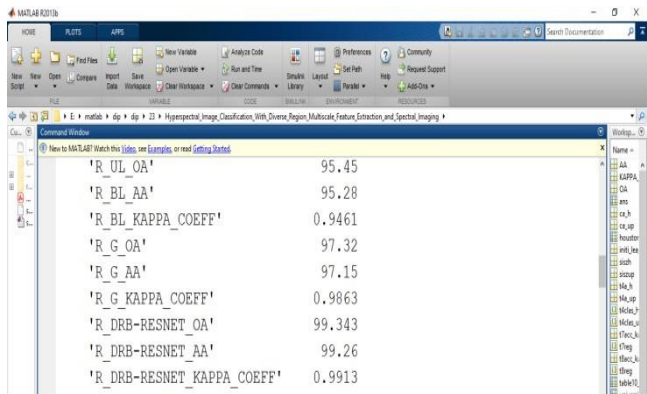
Tables VII–IX list the overall accuracy (OA), average accu- racy (AA), and Kappa coefficients using different regions and the proposed fusion strategy. $R_{UR}$ represents the region of upper right, $R_{UL}$ represents the region of upper left, $R_{BR}$ represents the region of bottom right, $R_{BL}$ represents the region of bottom left, and $R_G$ represents the region of global. The results show that due to the existence of positional anisotropy, the performance of single-region-based classification is very limited, and it is obvious that the classification results have great differences in different regions, which indicates that various context distribu- tions have different generalization capabilities for classification. As to which region classification works better, it largely depends on the actual distribution of the data, and our experimental results using the fusion strategy can definitely be improved. Therefore, the proposed DRB-ResNet achieves the best accuracy as it has a more robust feature representation considering multiple possible distributions.

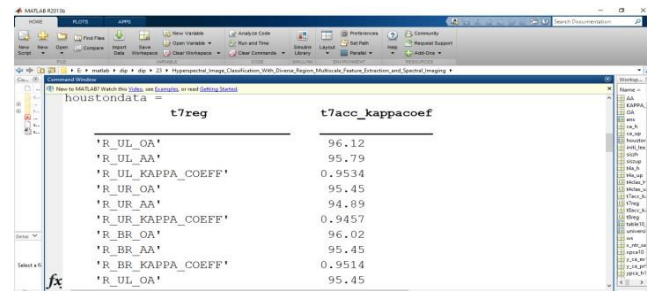### 3.3. Classification performance

Our proposed DRB-RESNET method is compared with cur- rent advanced HSI classification method, such as SSUN OA, and AA indicate the classification performance

of different methods on several datasets. Among them, the "Proposed (NS)" represents that the data without spectral imaging is input into a two-branch feature extraction network.And the "Proposed" represents that we spectral and spatial features.And selection method for dataset segmentation,for each category of data we randomly select different data for training and remaining data for testing.The experimental results we obtained show that the classification accuracy based on joint

**TABLE 5-**CLASSIFICATION PERFORMANCE OF DIVERSE REGIONS FOR THE HOUSTON DATA
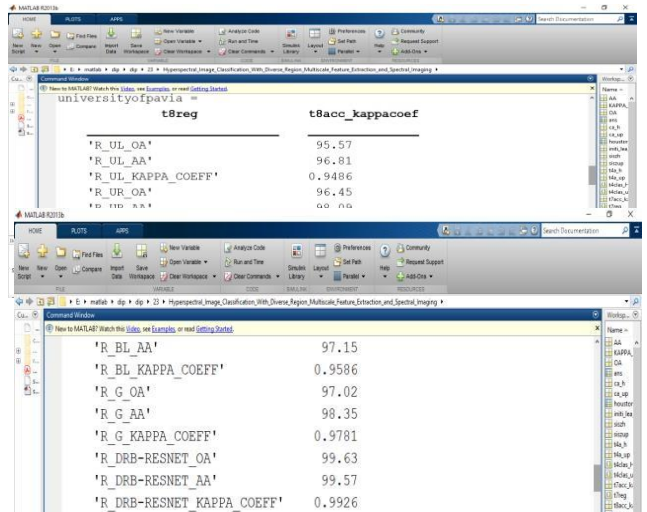




**TABLE 6-**CLASSIFICATION PERFORMANCE OF DIVERSE REGIONS FOR THE UNIVERSITY OF PAVIA DATA



spatial and spectral features is higher than that using only one of these features,and our proposed DRB-Resnet model out performs other classifiers in terms of classification accuracy.

**TABLE 7-**CLASSIFICATION PERFORMANCE (%) OF DIFFERENT INITIAL LEARNING RATE FOR DRB-RESNET ON THE HOUSTON 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            |

And selection method for dataset segmentation,for each category of data we randomly select different data for training and remaining data for testing.The experimental results we obtained show that the classification accuracy based on joint spatial and spectral features is higher than that using only one of these features,and our proposed DRB-Resnet model out performs other classifiers in terms of classification accuracy. we can obtain the accuracy of our proposed DRB-ResNet as 99.76%, which is about 2.13% higher compared to SSUN (97.21%). A similar performance also exists for experiments conducted on other datasets. The classification performance of our proposed method on the University of Pavia dataset, Houston, is approximately 2%, 1% better than the other classification methods compared to the other methods.Figs. 4–5 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 IV and V. By the presentation of the images, we can visually see that compared with SSUN our proposed DRB-ResNet classification result graph has significantly fewer points of classification errors in many regions, so we can conclude that our DRB - ResNet classification method performs better than other classification methods.

### 4.CONCLUSION

This article proposes a innovative dual-branch network utilizing multiscale multiregion Gaussian pyramid and spectral information imaging for hyperspectral image (HSI) classification. Initially, the image undergoes multiregion segmentation, followed by Gaussian pyramid processing on the graph to extract features across multiple scales. Subsequently, the dimensionality of each spectral data type is reduced to a predefined level and then expanded into an image. This process helps mitigate the influence of anomalous spectral data on classification outcomes. The dual-branch network is adept at extracting both spatial and spectral information effectively. Experimental findings demonstrate superior performance of the proposed dual-branch method, leveraging multiscale multiregion

Gaussian pyramid and spectral information imaging, compared to other state-of-the-art approaches across two datasets.

## REFERENCES

[1] D. Landgrebe, "Hyperspectral image data analysis," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 17–28, Jan. 2002.

[2] 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.

[3] H. Li, Y. Song, and C. L. P. Chen, "Hyperspectral image classification based on multiscale spatial information fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 9, pp. 5302–5312, Sep. 2017.

[4] 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*, J.-S. Pan, P.-W. Tsai, and H.-C.Huang, Eds., Cham, Switzerland: Springer, 2017, pp. 61–68.

[5] L. Zhang, W. Wei, Y. Zhang, C. Shen, A. van den Hengel, and Q. Shi, "Cluster sparsity field: An internal hyperspectral imagery prior for reconstruction," *Int. J. Comput. Vis.*, vol. 126, no. 8, pp. 797–821,2018.

[6] S. Li, Q. Hao, G. Gao, and X. Kang, "The effect of ground truth on per formance evaluation of hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 12, pp. 7195–7206, Dec. 2018.

[7] W. Wei, L. Zhang, Y. Jiao, C. Tian, C. Wang, and Y. Zhang, "Intracluster structured low-rank matrix analysis method for hyperspectral denoising " *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 866–880, Feb. 2019.

[8] Z. Wu, Y. Li, A. Plaza, J. Li, F. Xiao, and Z. Wei, "Parallel and distributed dimensionality reduction of hyperspectral data on cloud computing architectures," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 9, no. 6, pp. 2270–2278, Jun. 2016.

[9] K. Berger, C. Atzberger, M. Danner, M. Wocher, W. Mauser, and T. Hank, "Model-based optimization of spectral sampling for the retrieval of crop variables with the PROSAIL model," *Remote Sens.*, vol. 10, no. 12, 2018, Art. no. 2063.

[10] I. Aneece and P. Thenkabail, "Accuracies achieved in classifying five leading world crop types and their growth stages using optimal earth observing-1 hyperion hyperspectral narrowbands on Google Earth engine,"*Remote Sens.*, vol. 10, no. 12, 2018, Art. no. 2027.

[11] A. Marcinkowska-Ochtyra, A. Jaroci ´nska, K. BzdÄZga, and B. Tokarska-Guzik, "Classification of expansive grassland species in different growth stages based on hyperspectral and lidar data," *Remote Sens.*, vol. 10, no. 12, 2018, Art. no. 2019.

[12] A. Marcinkowska-Ochtyra, K. Gryguc, A. Ochtyra, D. Kope´c, A. Jaroci ´nska, and Ł. Sławik, "Multitemporal hyperspectral data fusion with topographic indices-improving classification of natura 2000 grassland habitats," *Remote Sens.*, vol. 11, no. 19, 2019, Art. no. 2264.

[13] C. Pelletier, S. Valero, J. Inglada, N. Champion, C. M. Sicre, and G. Dedieu, "Effect of training class label noise on classification performances for land cover mapping with satellite image time series," *Remote Sens.*, vol. 9, no. 2, 2017, Art. no. 173.

[14] Y. Zhong and L. Zhang, "An adaptive artificial immune network for supervised classification of multi-/hyperspectral remote sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 894–909, Mar. 2012.

[15] B. Tu, X. Yang, N. Li, C. Zhou, and D. He, "Hyperspectral anomaly detection via density peak clustering," *Pattern Recognit. Lett.*, vol. 129, pp. 144–149, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167 865519303423

[16] J. Sun *et al.*, "Deep clustering with intraclass distance constraint for hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 4135–4149, May 2021.

[17] X. Xu, W. Li, Q. Ran, Q. Du, L. Gao, and B. Zhang, "Multisource remote sensing data classification based on convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 937–949, Feb. 2018.

[18] Ham, Y. Chen, M. Crawford, and J.Ghosh,"Investigation of the random forest framework for classification of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 492–501, Mar. 2005.

[19] M. Belgiu and L. Dr˘agu¸t, "Random forest in remote sensing: A review of applications and future directions,"*ISPRS J. Photogrammetry Remote Sens.*, vol. 114, pp. 24–31, 2016.

[20] Y. Xu, B. Du, F. Zhang, and L. Zhang, "Hyperspectral image classification via a random patches network," *ISPRSJ. Photogrammetry Remote Sens.*, vol. 142, pp. 344–357, 2018.

[21] 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.

[22] 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.

[23] L. He, J. Li, A. Plaza, and Y. Li, "Discriminative low- rank Gabor filtering for spectral-spatial hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 3, pp. 1381–1395, Mar. 2017.

[24] D. Xu *et al.*, "Hyperspectral image classification based on hierarchical guidance filtering and nearest regularized subspace," *Acta Photonica Sinica*, vol. 49, no. 4, pp. 197–209, 2020.

[25] B. Xiaojun and Z. Zengyu, "Hyperspectral image classification algorithm based on two-channel generative adversarial network," *Acta Optica Sinica*, vol. 39, no. 10, pp. 297–308, 2019.

[26] E. Pan *et al.*, "Spectral-spatial classification of hyperspectral image based on a joint attention network,"in *Proc. IGARSS IEEE Int. Geosci. Remote Sens. Symp.*, 2019,pp. 413–416.

[27] R. Li, S. Zheng, C. Duan, Y. Yang, and X. Wang,"Classification of hyperspectral image based on double- branch dual-attention mechanism network," *Remote Sens.*, vol. 12, no. 3, 2020, Art. no. 582.

[28] X. Tang *et al.*,"Hyperspectral image classification based on 3-D octave convolution with spatial-spectral attention network," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 3, pp. 2430–2447, Mar. 2021.

[29] 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.

[30] Y. Chen *et al.*, "Drop an octave: Reducing spatial redundancy in convolutional neural networks with octave convolution," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2019, pp. 3435–3444.

[31] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, "Deep feature extraction and classification of hyperspectral images based on convolutional neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6232–6251, Oct. 2016.

[32] 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.

[33] J. Zhu, L. Fang, and P. Ghamisi, "Deformable convolutional neural net  works for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 8, pp. 1254–1258, Aug. 2018.

[34] 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.

[35] K. Makantasis, K. Karantzalos, A. Doulamis, and N. Doulamis, "Deep supervised learning for hyperspectral data classification through convolutional neural networks," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*,2015, pp. 4959–4962.

[36] 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.

[37] B. Waske, S. van der Linden, J. A. 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.

[38] 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.

[39] S. Li, Q. Hao, X. Kang, and J. A. Benediktsson, "Gaussian pyramid based multiscale feature fusion for hyperspectral image classification," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 11, no. 9, pp. 3312–3324, Sep. 2018.

[40] G. A. Shaw and H. K. Burke, "Spectral imaging for remote sensing," *Lincoln Lab. J.*, vol. 14, no. 1, pp. 3–28, 2003.

[41] S. Jay *et al.*, "Hyperspectral remote sensing of shallow waters: Consider  ing environmental noise and bottom intra-class variability for modeling and inversion of water reflectance," *Remote Sens. Environ.*, vol. 200, pp. 352–367, 2017.

[42] Y. Tarabalka, M. Fauvel, J. Chanussot, and J. A. Benediktsson, "SVM-and MRF-based method for accurate classification of hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 4, pp. 736–740, Oct. 2010.

[43] S. Mei, J. Ji, J. Hou, X. Li, and Q. Du, "Learning sensor-specific spatial  spectral features of hyperspectral images via convolutional neural net  works," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4520–4533, Aug. 2017.

[44] W. Hu, Y. Huang, L. Wei, F. Zhang, and H. Li, "Deep convolutional neural networks for hyperspectral image classification," *J. Sensors*, vol. 2015, 2015, Art. no. 258619.