Journal of Network Intelligence
Taiwan Ubiquitous Information
 ISSN 2414-8105 (Online)
Volume **8**, Number **2**, May **2023**

Edge Fine Classification Optimization Network for Hyperspectral Image Classification

Hai-Feng Sima*

College of Computer Science and Technology
Henan Polytechnic University
Jiaozuo 454000, China
smhf@hpu.edu.cn

Feng Gao

College of Computer Science and Technology
Henan Polytechnic University
Jiaozuo 454000, China
gaofeng1011@163.com

Lan-Lan Liu

School of Liberal Arts and Law
Henan Polytechnic University
Jiaozuo 454000, China
liulanlan@hpu.edu.cn

Jun-Ding Sun

College of Computer Science and Technology
Henan Polytechnic University
Jiaozuo 454000, China
sunjd@hpu.edu.cn

Chao-Sheng Tang

College of Computer Science and Technology
Henan Polytechnic University
Jiaozuo 454000, China
tcs@hpu.edu.cn

*Corresponding author: Hai-Feng Sima

Received October 14, 2022, revised December 5, 2022, accepted February 5, 2023.

Abstract. *With the continuous updating and development of deep learning theory, the research of hyperspectral image classification has entered a new era. The emergence of convolutional neural networks has greatly improved the representation and recognition capabilities of hyperspectral images. The problem of misclassification arising from the underutilization of the spatial-spectrum features and the complexity of the high-dimensional spectral space is still a difficult area for accurate classification. To solve the above problems, we propose a novel network model. First, according to the high-dimensional characteristics of hyperspectral images, the channel attention mechanism can be used in the network to better extract spatial-spectral features. Secondly, a convolutional layer that adapts to the edges of irregular object contours in the image is incorporated to capture more complete object features. Finally, experiments are performed using authoritative datasets such as Pavia University, Indian Pines, and KSC, and compared with other current state-of-the-art network models. The results show that our network can provide more accurate classification results and is more robust.*
Keywords: Hyperspectral image classification; Convolutional network; Attention mechanism

1. **Introduction.** Hyperspectral imagery (HSI) is a three-dimensional data block consisting of multiple continuous spectra taken at high altitudes by specific sensing devices, and the data features in HSI can accurately reflect the land properties of the near-surface layer. Based on the acquired hyperspectral images, the objects in the images are precisely analyzed, which can provide effective data support for several fields, such as agricultural inspection [1], smart cities [2] and geological surveys [3], etc. The HSI classification task is a fundamental task among the many tasks of HSI, requiring extremely high accuracy of the classification results of images to provide accurate data support for other fields. However, there are some problems in the HSI classification task, such as the high dimensionality of a small number of samples leads to the Hughes phenomenon, the "same subject with different spectra" phenomenon makes the classification task more difficult, and the small sample size also makes the classification task difficult. Therefore, it is extremely difficult to classify them efficiently and accurately. This task has been studied in academia and industry for many years.

Machine learning includes classical algorithms such as random forest algorithms, regression trees, and SVMs, and is used in a wide variety of fields, such as options trading [4], capital structure analysis [5], and transport pattern detection [6]. The machine learning approach also shows excellent performance in processing hyperspectral image classification tasks. Wang et al. [7] used feature transformation methods such as PCA and ICA to reduce the dimensionality of the original image, followed by a random forest algorithm to finely classify the ground objects in the image. Shape adaptive reconstruction was used for preprocessing in the work of Li et al. [8], each pixel in the image was preprocessed to make full use of spatial-spectrum information. Wang and Zeng [9] considered that the traditional hyperspectral remote sensing image classification methods only use the spectrum features of the image and ignore the texture features in the image, so they proposed to use the above two features together and extract the most prominent texture features using the grayscale contribution matrix to form texture-spectrum features, which greatly improves the efficiency of feature usage in HSI. To deal with the linear indistinguishability problem in hyperspectral image classification, a KCRC-based packing algorithm was discussed by Su and Hongjun et al. [10], using bootstrapping to increase the diversity of the underlying classifiers to improve the classification performance and generalization ability. Su et al. [11] proposed a novel RCR method based on band weighting with superpixel

segmentation for HSI classification, which can effectively improve the classification accuracy by using the L2 norm of band coefficients and global average coefficients to ensure the similarity, while the variance determines the specific correlation weights of each band.

An M-3DCNNAttention-based HSI classification method was introduced in the work of Sun and Kun [12], where an expanded dataset is used to highlight spectral features in image features in a CNN structure. To alleviate the problem of small samples in classification tasks, Xue et al. [13] used a minimized fully convolutional network and a semi-supervised loss function for pixel-level classification of hyperspectral images. A semi-supervised spatial-spectrum method based on 3D Gabor and co-selective self-training are discussed by Pan et al. [14]. This method is used for HSI classification to make full use of spatial-spectrum features. Regarding the overfitting problem in training, it is solved in the 3D asymmetric concept network AINet [15]. A lightweight convolution operator was used in CNN networks by Lv et al. [16] to optimize the high computational cost of the network during training and the high redundancy of the convolution kernel. In HSI, data features are unique representations of objects, and therefore, the recognition of image objects is highly dependent on feature data. However, in feature data, phenomena such as data redundancy and data insufficiency can occur from time to time. To overcome these challenges, an integrated convolutional neural network (C-CNN) was proposed by Chang et al. [17] . This approach uses a 3DCNN to represent spatial-spectral features, a 2DCNN to learn abstract spatial features, and then a dimensionality reduction algorithm to reduce the redundancy of the spectral bands. A high-level semantic perspective to learn spatial-spectrum features was used in the work of Xu et al. [18] to obtain two feature descriptions of the same data from different perspectives, which can better utilize the feature information of spectrum images. To further capture the spatial-spectrum information in HSI, Guo et al. [19] proposed a spatial feature obtained using CNN-enhanced multilevel Haar combined with Haar wavelet decomposition features for better information interaction.

In this paper, we propose a new network model to relieve the impact of underutilized edge information and spatial-spectrum information on HSI classification results. The model uses convolutional neural networks in a deep learning approach to effectively classified for the complex curvilinear edge information of objects in HSI. The main contributions of this paper are as follows.

(1) In our network, convolutional layers that can adapt to object edge contours in images are used in the network, i.e., deformation convolution. During the convolution process, more complete object information is extracted based on the offset information of the target object pixels.

(2) To use the high-dimensional information in HSI more effectively, the channel attention mechanism is used in our network. The attention mechanism targets the multichannel information of HSI and can accurately extract the spectral features.

(3) Extensive experiments show that adding deformable convolution and channel attention mechanisms to the network can effectively improve the classification of HSI. After deformable convolution, more complete object features are obtained. The feature map of the target object is salient through the channel attention mechanism. It provides a better feature map for the subsequent feature extraction network part and improves the classification accuracy of the network.

2. **Related work.**

A convolutional neural network (CNN) is a major approach for processing images in deep learning, and its main structure includes an input layer, convolutional layer, pooling layer, fully connected layer, and output layer. It has better classification performance in the field of hyperspectral image classification compared to

the traditional approaches [20]. The convolution layer is the core part of CNN, which is responsible for extracting feature information from images.

2D convolution was discussed in Zhang's work [21], targeting specific regions in the image for feature extraction, effectively synthesizing the features of objects in the regions, and improving the classification of the network. MDCNN was proposed by Kumar et al. [22], which uses mathematical morphology in CNN not only to describe the spatial features of HSI directly but also significantly reduce the computational workload. Ren et al. composed MSAC by combining superpixels with conventional convolution [23], which increases the flexibility of the convolution kernel during convolution and fully suppresses the effect of edge noise. To increase the usage of different scale features in the network, FBMB was proposed by Shi et al. [24] to use multiple branching multiscale modules in the network to fuse features of different network depths. Wang and Fan [25] proposed the densely connected network (MSDAN), a model that enhances the classification performance and mitigates the generation of overfitting and gradient disappearance phenomena.

Based on two principles of pixel-level classification and hyperspectral image spatial information feature focus mechanism (SDPCA), a scaled dot product center focus mechanism was designed by Liu et al. [26] for extracting spatial-spectrum information similar to the center pixel and the center pixel in HSI. However, in CNNs, a deeper and broader architecture and a large amount of training data often need to be designed to obtain good performance, and the over-reliance on forwarding connectivity leads to no efficient use of data information. To overcome these drawbacks, Pande and Banerjee [27] presented a self-looping network scheme in which each layer in the self-looping unit has forward and backward connections, and each layer has inputs and outputs from other layers, maximizing information exchange and significantly improving the classification performance of the network. A residual attention network (SpaAG-RAN) for guiding spatial attention in HSI classification is discussed by Li and Wang [28], which utilizes spectrum similarity to capture relevant spatial regions and maximizes the promotion of classification accuracy. ConvNeXt is the latest CNN model designed by Liu et al. [29]. It has a faster inference speed as well as a higher accuracy rate. However, the huge number of parameters is a prominent problem of this network.

3. **Methodology.**

Based on the characteristics of multiple channels in HSI, we design a neural network model. The model draws on ShuffleNetV2, a lightweight network with high classification accuracy, and optimizes it. The details are as follows: (1) To address the problem of low classification accuracy caused by curved edges of objects in hyperspectral images, the normal convolution in the shallow part of the network is replaced by deformable convolution. According to the adaptive object shape of deformable convolution, the classification effect of object edges is effectively improved. (2) Due to the complex spatial structure of object edges, we believe that more attention to spectrum features is needed. The spectrum features in the image are extracted using the channel attention mechanism as a way to compensate for the disadvantage of using spatial features prone to classification errors. This is shown in Figure 1.

3.1. **Backbone network.**

ShuffleNetV2 [30] is a fast classification network model with a particularly lightweight feature. However, in the hyperspectral image classification task, classification results did not achieve comparatively excellent results. However, the classification results were not particularly outstanding in the hyperspectral image classification task. Based on this issue, ShuffleNetV2 was optimized and improved for the characteristics of HSI in our work.

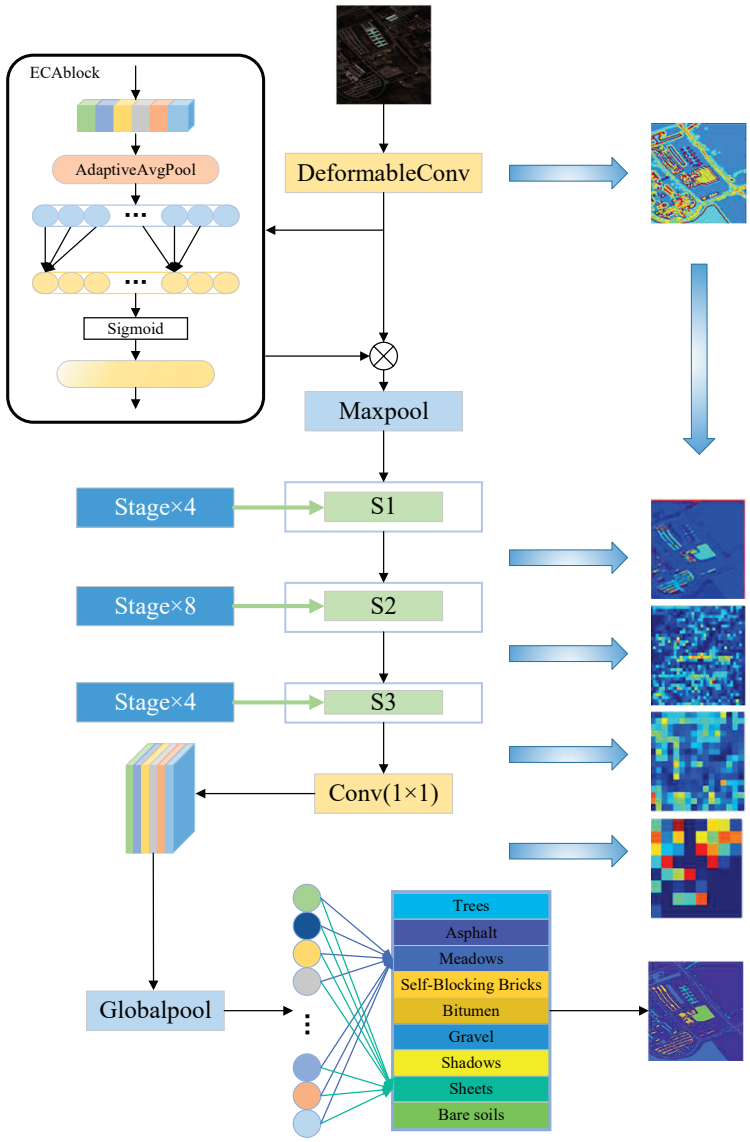

FIGURE 1. Network structure

The Stage structure is used as the main component of ShuffleNetV2, as shown in Figure 2. Figure (B) contains a 3×3 convolutional stage with a convolutional kernel step size of 1, which can downsample the input feature map. The convolutional layer in figure (A) is all convolutional kernels with step size 1, which can extract more detailed feature information and enhance the characterization ability, and has a separate branch to realize the fusion of the input initial feature map and the processed feature map. At the end of the network, a 1×1 convolution kernel improves the interaction between channels in the feature map, and this convolution can be well used in hyperspectral image classification tasks. However, the process of convolution consumes huge computational resources, so Channel shuffle is used instead of 1×1 convolution to break up the channels and enhance the overall generalization ability of the network.

As in Figure 1, we have simply divided the network into 4 parts in our network. The first part (Base1) includes deformable convolution. The second part (Base2) is the Max-Pool layer. The third part (Base3) includes S1, S2, and S3 modules. The fourth part (Base4) includes the 1×1 convolution of the deeper layers of the network and the final classification module. In base 1, the original convolution is replaced by deformable convolution. Deformable convolution allows adaptive convolution in shallow networks based on object boundaries in the most original image, and the resulting feature maps contain more

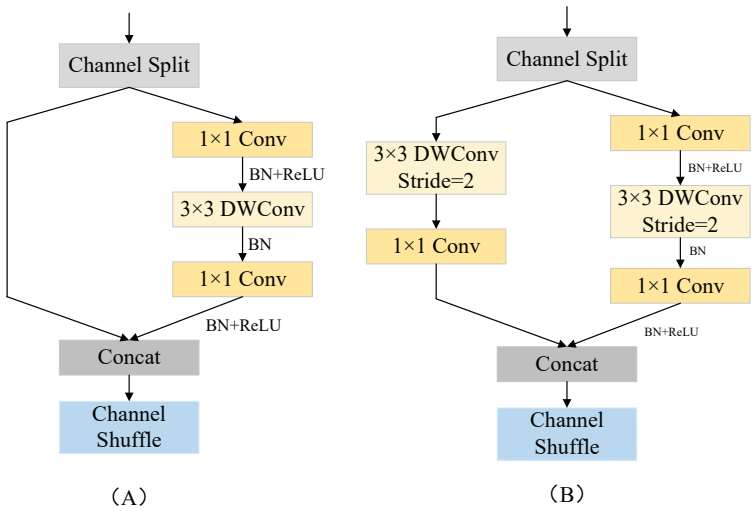

FIGURE 2. STAGE STRUCTURE

complete information about object features. And between Base1 and Base2, the channel attention mechanism is used in the network, which can efficiently process and extract spatial-spectrum features based on the feature maps outputted from the previous layer of the network. After that, the feature maps outputted by Base2 are further processed by Base3, and the feature maps are input to the classification module Base4 to get the final classification results.

3.2. **Adaptive object edge convolution module.** To alleviate the impact of the complex curved outline of objects in HSI on the classification results, we add deformable convolution to the network model. Make the convolution kernel adjust itself to the shape of the target object during the convolution process to obtain more effective feature maps.

The shape of convolutional kernels in traditional convolutional layers is square, especially the 3×3 convolution kernel structure, which can make better use of the GPU operation performance. But there is also a more obvious drawback that it cannot be adjusted well to the shape of the target in the image, which will affect the prediction effect of the network model to a certain extent. The deformable convolution is proposed to remedy this defect well, and the convolution kernel adds an orientation parameter to adapt to the feature changes of the target object during the computation, as in Figure 3.

In the deformable convolutional layer used in this network, the kernel is set to 3×3 and the stride is set to 1. It contains two conventional 3×3 convolution kernels, and the number of channels in the first convolution kernel is $2\times k^2$ (k=3 in 3×3 convolution, indicating the convolution kernel size), which is used to learn the amount of offset information on the x-axis and y-axis in the plane coordinate system for each position in the extracted perceptual field. Because the pixel positions after adding the offset information are not integers and cannot be corresponded to the pixels on the feature map, bilinear interpolation is used to calculate the pixel values at these positions, and the pixel values of the interpolated points are the weighted sum of the distances between the interpolated points and the pixel points in the four directions adjacent to them in the original feature map. At this point, we have completed the one-to-one correspondence between the interpolated points and the pixel points in the feature map. Finally, a standard 3×3 convolution kernel is used for the computation, and the resulting feature map is the one after deformable convolution processing. The feature map contains richer and more complete information about the object.

To process the original image with deformable convolution first in the network, deformable convolution is used in Base1 of the network. Based on the object contour

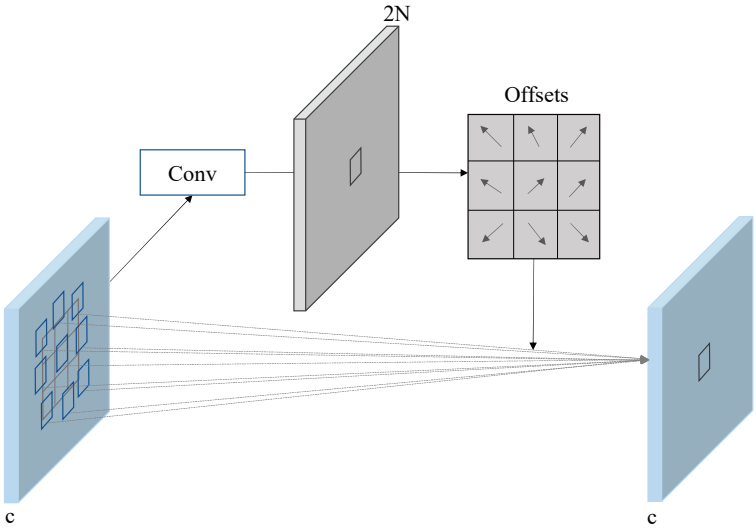

FIGURE 3. DEFORMABLE CONVOLUTION STRUCTURE

information in the original image, a deformable convolution operation is performed to greatly preserve the original information.

3.3. **Channel attention mechanism.** The attention mechanism in computer vision is inspired by the human visual attention mechanism and is applied in the field. It simply means that more attentional resources are devoted to a target object in a region, and therefore more detailed information can be obtained for that target object, which helps to optimize the prediction results.

SENet [31] is an attention mechanism that was first proposed to focus on important object features by capturing the relationship between channels in the feature map. However, the dimensionality reduction of channels in shallow networks makes a lot of feature information lost, which affects the final network results. Therefore, an ECA [32] module that focuses more on channel information is proposed and is also the attention mechanism used in this paper. The ECA module is based on SENet, avoiding the downscaling of the number of channels and maintaining some cross-channel interaction capability. In our network, the ECA attention mechanism is used to better process the image feature information for the characteristics of the HSI multi-channel.

In this ECA attention module, the feature map is first downsampled using an adaptive pooling function. Afterward, a convolutional layer with a kernel size of 1×1 is used to maintain and optimize the connections to other channels. In this process, the number of input and output channels of the convolution kernel is set to be the same, and the number of channels is not reduced, so that the object feature information in the image is retained to the maximum extent. Then the sigmoid activation function is used to increase the nonlinearity of the output data.

4. **Experimental results and analysis.** In this section, we introduce the dataset and the experimental environment used in this experiment, followed by the selection of the loss function and the optimizer and the setting of the hyperparameters of the network in the experiment. The network implementation is based on the Win10 operating system and utilizes the PyTorch 1.8 deep learning framework, with an RTX 8000 (48GB) GPU used to train the network. To reflect the classification effectiveness of the network, current advanced classification models are used in experiments for comparison. Finally, To ensure the stability of the experimental results, each set of experiments is repeated multiple times. We used overall accuracy (OA), average accuracy (AA), and Kappa coefficient as measures of experimental results.

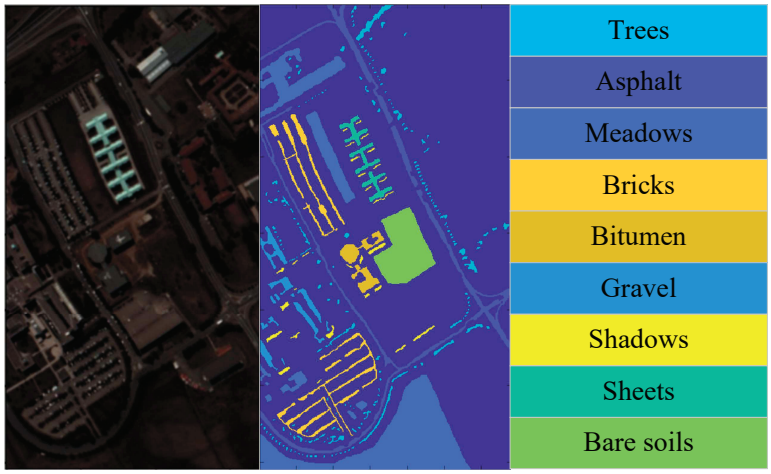

FIGURE 4. ORIGINAL, TRUE VALUE, AND CATEGORICAL REPRESENTATION OF THE PAVIA UNIVERSITY DATASET

4.1. **Hyperspectral dataset.** The datasets used in the experiments are those that are frequently used in HSI classification, namely: Pavia university, India pines, and KSC.

The Pavia University dataset is captured by the reflective-optics-spectrographic-imaging-system (ROSIS) sensor over the University of Pavia in northern Italy. It is illustrated in Figure 4. The spatial resolution of the image is 1.3 meters, and the spatial size is 610×340 pixels.The original image contains a total of 115 spectral bands, and the spectral range is in the range of 0.43-0.86μm. However, in our experiments, some of the noise bands were discarded and 103 spectral bands were selected as the dataset, with a total of 42776 sample data.

The India pine dataset is the first dataset taken by the airborne-visible-infrared-imager-spectrometer (AVIRIS) sensor in northwestern Indiana, USA. It is illustrated in Figure 5. The spatial resolution of the image is 20 m, the size is 145×145 pixels, and the spectral range of the image obtained is 0.4-2.5μm. There are 224 spectral bands, but some of the spectra are contaminated by noise such as water vapor, so 204 spectral bands are used as data sets for experimental analysis. The data set contains a total of 16 different categories, but from a statistical point of view, we discard 8 small categories and select 8 major categories for experiments [33, 34], a total of 8504 sample data.

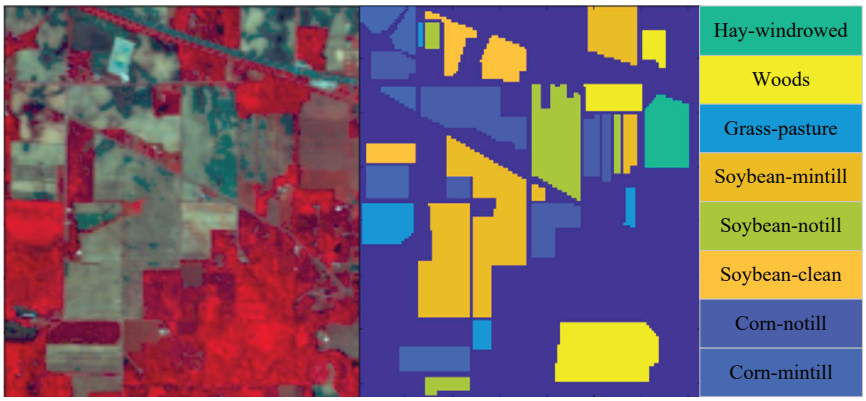

FIGURE 5. ORIGINAL, TRUE VALUE AND CATEGORICAL REPRESENTATION OF THE INDIAN PINES DATASET

The KSC dataset was taken in Florida in 1996 by the airborne-visible-infrared-imager-spectrometer (AVIRIS) sensor. It is illustrated in Figure 6. The spatial resolution of the image is 18 meters, and the spatial size is 512×614. The image consists of 176 bands and contains 13 categories.

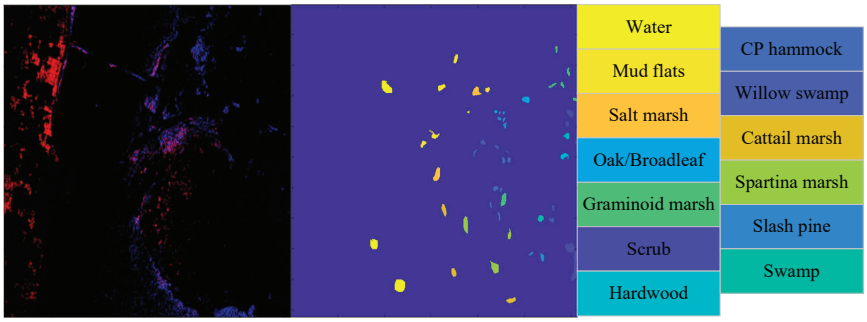

FIGURE 6. ORIGINAL, TRUE VALUE, AND CATEGORICAL REPRESENTATION OF THE KSC DATASET

4.2. **Experiment 1.** Convolutional neural networks are data-driven model networks that require a large amount of data to update and improve the weights of the network during the training process. We add Gaussian noise and level flipping to the original training set and double the number of samples to the original training set, which can better improve the generalization ability of the network.

For Pavia University, we randomly selected 200 samples per class for training and other samples for testing. And the same operation was taken for the Indian Pines dataset. In the KSC dataset, due to the small total number of samples, we randomly selected 50 samples from each class as the training set of the network, and the remaining samples constitute the test set. In the experiments, Adam was used as the optimization in the network, the learning rate was set to 0.001, Epoch was set to 500 for the experiments, and Batch_size was set to 512. For the loss computation, CrossEntropy Loss is used in our experiments, for it is widely used in classification tasks. To better fit the local feature extraction network, we extracted a 25×25 square local image centered on the target pixel as the input to the network.

To know how well our proposed network model classifies, we ran other networks in the experiment to compare the results. In the comparison experiments, the parameter configuration was set to the same to ensure the same experimental environment for all networks. The comparison network models used in the experiments include ShufflenetV2 [30], HybridSN [35], DR-CNN [24], TBTA-D2Net [36], and SE-CNN [37]. HybridSN uses a combined 2D-3D convolution scheme to make fuller use of the spectral features of the image. DR-CNN performs spatial feature extraction for six regions of the spatial location where a pixel point is located, and captures spatial features more finely. TBTA-D2Net adds the Dense2Net bottleneck block and attention mechanism to the triple-branch ternary-attention mechanism to improve the effective classification accuracy of the TBTA-D2Net algorithm. SE-CNN combines squeezing and excitation networks with convolutional neural networks to increase its performance in extracting features and classifying HSI.

In Table 1, the classification accuracy of our proposed network is above 98% in all nine categories of the Pavia University dataset, and the majority of them remain above 99% with high stability. The overall classification accuracy of the dataset is 99.51%, which is 0.69 percentage points higher than the classification accuracy of the second-place DR-CNN, with a significant advantage. And in the Indian Pines dataset shown in Table 2, our network achieved an overall classification accuracy of 97.88%, all of which exceeded the other three network models. The accuracy in five categories is optimal, and the accuracy in four categories is 100%, which has a relatively good classification effect. The network also achieved the highest accuracy of 98.06% in the KSC dataset shown in Table 3, and both AA and Kappa coefficients exceeded those of the other comparison networks.

According to the above analysis, our network model has a good classification effect and is robust. The inclusion of deformable convolution in the network can be well adapted for the edges of objects in hyperspectral images. In the deformable convolution, the

Table 1. CLASSIFICATION ACCURACY ON THE PAVIA UNIVERSITY DATASET

Class	ShufflenetV2	TBTA-D2Net	SE-CNN	HybridSN	DR-CNN	Ours
1	93.59	99.73	98.56	97.59	99.47	99.22
2	99.97	99.96	99.97	99.57	99.73	99.86
3	91.25	97.88	98.53	99.51	95.65	98.43
4	99.87	87.81	93.70	96.99	97.86	99.28
5	100	100	99.93	99.06	99.39	99.41
6	91.62	98.07	99.90	99.58	99.01	99.92
7	99.22	98.88	99.47	71.04	98.43	98.95
8	98.84	99.37	99.70	95.05	95.12	98.85
9	98.85	96.51	97.76	93.16	99.20	99.05
OA(%)	97.26	98.45	98.85	97.30	98.82	99.51
AA(%)	95.82	99.13	99.39	94.86	98.84	99.17
Kappa×100	96.38	97.96	98.48	96.42	98.42	98.35

Table 2. CLASSIFICATION ACCURACY ON THE INDIAN PINE DATASET

Class	ShufflenetV2	TBTA-D2Net	SE-CNN	HybridSN	DR-CNN	Ours
1	96.91	98.92	96.84	97.97	91.65	92.55
2	99.72	88.94	94.75	88.99	92.52	100
3	99.79	95.04	99.79	97.23	98.77	100
4	99.79	100	100	99.30	100	100
5	91.92	98.74	97.76	99.29	99.89	100
6	92.73	98.28	97.93	98.51	100	98.05
7	96.23	98.96	96.15	99.14	97.35	99.16
8	100	100	99.61	99.46	99.68	98.98
OA(%)	96.01	97.62	97.78	97.63	97.37	97.88
AA(%)	95.70	97.81	97.58	97.15	97.95	98.16
Kappa×100	95.19	97.15	97.32	97.43	96.84	97.56

convolution kernel can be automatically adjusted according to the offset information of the object pixels to fit the contour of the object to the maximum extent. At the same time, the feature information of the object is obtained to the maximum extent in the receptive field. Based on the visualized images, it can be seen that there is a significant improvement in the classification of object contour edges. HSI is a high-dimensional data image, its multi-channel is one of its distinctive features. In this article, the use of the channel attention mechanism greatly enhances the information interaction between different channels in the image and improves the efficiency of using spectral features. Therefore, it can be optimized for some misclassification cases. And the channel attention mechanism in the network can be used for the characteristics of HSI with multiple channels. The results of the use are reflected in the visualized images, and there is a significant improvement in the misclassification of objects within the images.The visualized images are shown in Figures 7, 8 and 9.

TABLE 3. CLASSIFICATION ACCURACY ON THE KSC DATASET

Class	ShufflenetV2	TBTA-D2Net	SE-CNN	HybridSN	DR-CNN	Ours
1	95.84	97.44	99.48	100	100	99.61
2	89.92	95.51	97.01	97.10	94.42	95.22
3	100	99.61	98.81	98.05	95.85	93.73
4	83.64	95.45	93.90	96.86	96.09	95.85
5	95.78	98.10	98.73	88.95	90.00	99.29
6	97.54	98.23	97.41	98.25	100	99.10
7	100	99.04	100	93.75	87.50	100
8	96.40	97.47	98.84	96.41	98.40	97.70
9	99.43	99.81	100	99.05	100	99.62
10	97.96	95.63	96.50	96.79	98.72	98.28
11	99.76	99.52	99.52	100	100	99.76
12	91.54	95.02	93.54	97.16	96.22	95.59
13	99.78	99.78	99.46	99.57	98.72	99.67
OA(%)	96.35	98.06	98.14	98.06	97.93	98.25
AA(%)	93.85	97.00	97.20	97.30	97.09	97.34
Kappa×100	95.94	97.04	97.93	97.84	97.69	98.06

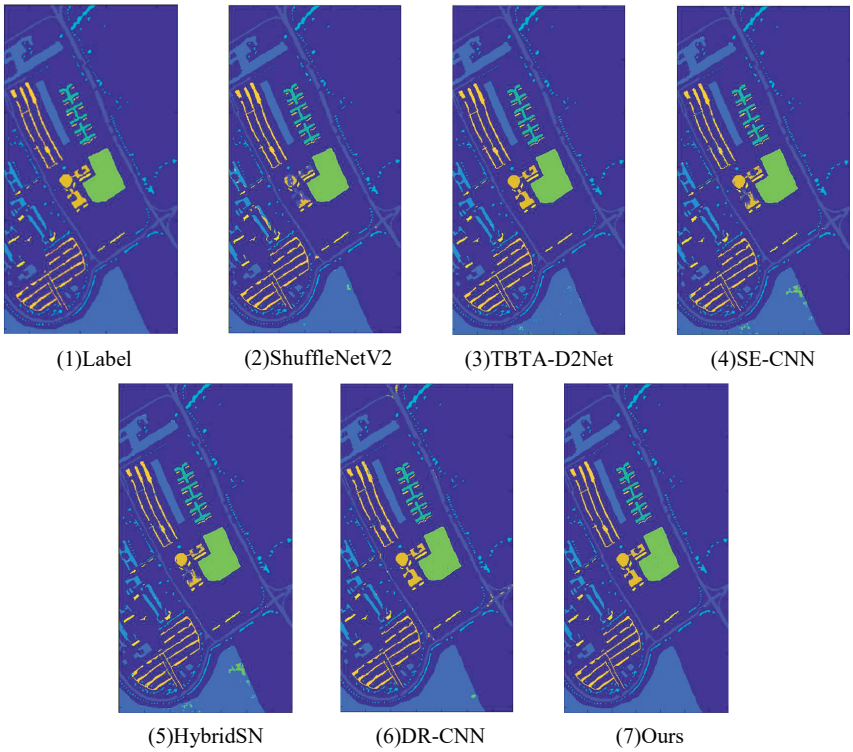

FIGURE 7. THE CLASSIFICATION MAPS ON THE PAVIA UNIVERSITY DATASET

4.3. **Experiment 2.** The data of the ablation experiments for the whole network are recorded in Table 4. The main ones include ShuffleNetV2, using the ECA attention mechanism in the backbone network (i.e., Using ECA), the backbone network using deformable convolutional modules at the base1 (i.e., Using DConv in base1), and our proposed network. The OA improves by 0.89%, 1.64%, and 1.33% on the three data sets for the backbone network with the ECA attention mechanism added compared to the backbone network. The backbone network with the addition of the deformable convolution module improves OA by 0.9%, 1.71%, and 0.71% on the three data sets, respectively, compared to the backbone network. In this stage, we can conclude that using the attention mechanism and the deformable convolutional module to the backbone network respectively can

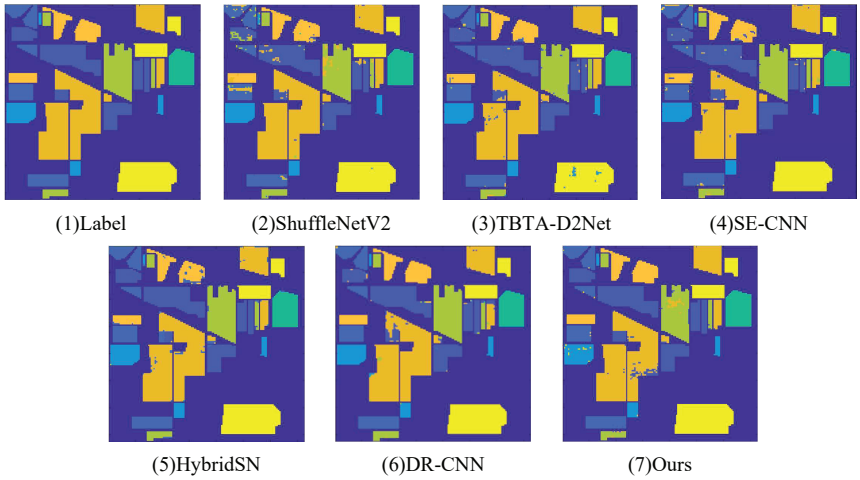

FIGURE 8. THE CLASSIFICATION MAPS ON THE INDIAN PINE DATASET

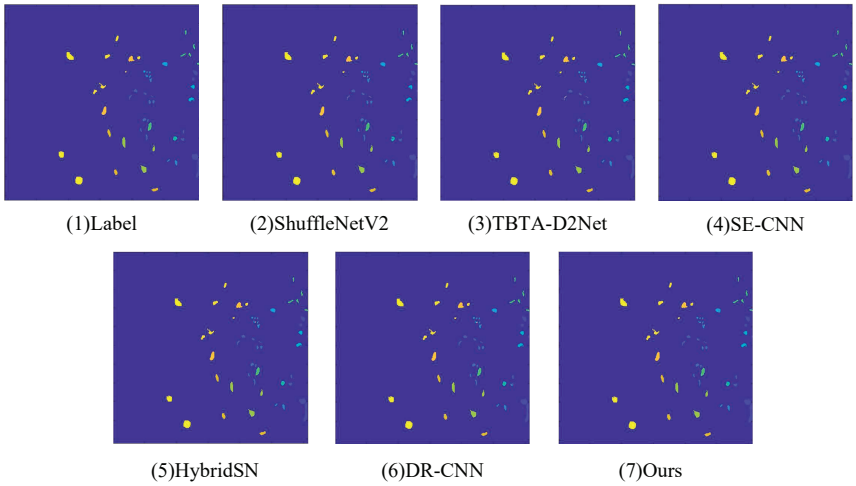

FIGURE 9. THE CLASSIFICATION MAPS ON THE KSC DATASET

improve the classification effect of the network to different degrees. In summary, using the ECA attention mechanism and deformable convolution in the backbone network can improve the classification effect of the network.

In the ECA attention mechanism, the size of the convolutional kernel affects the size of the receptive field, which affects the overall network classification effect. Therefore it is particularly important to find the most suitable convolutional kernel size for the network, and we performed ablation experiments based on different convolution kernel sizes on the datasets of Pavia University, Indian Pine, and KSC. In the experiments, the size of convolution kernel were respectively set to 1, 3, 5, 7, and 9. As can be known from Figures 10, 11, and 12, the OA keeps increasing when the convolution kernel size is 1 and 3 and reaches its best when it is 3. However, the OA keeps decreasing when taking 5, 7, and 9. Therefore, it can be concluded that the best classification of the whole network is achieved when the convolution kernel is taken as 3.

TABLE 4. EXPERIMENTAL RESULTS OF THE NETWORK USING DIFFERENT MODULES ON THREE DATASETS

	Index	ShuffleNetV2	Using ECA	Using DConv in base1	Ours
Pavia university	OA(%)	97.26	98.15	98.16	99.51
	AA(%)	95.82	95.95	97.75	99.17
	Kappa×100	96.38	97.55	97.57	99.35
Indian pines	OA(%)	96.01	97.65	97.72	97.88
	AA(%)	95.70	97.28	97.94	98.16
	Kappa×100	95.19	97.17	97.26	97.46
KSC	OA(%)	96.35	97.68	97.06	98.25
	AA(%)	93.85	96.52	95.46	97.34
	Kappa×100	95.94	97.41	96.73	98.06

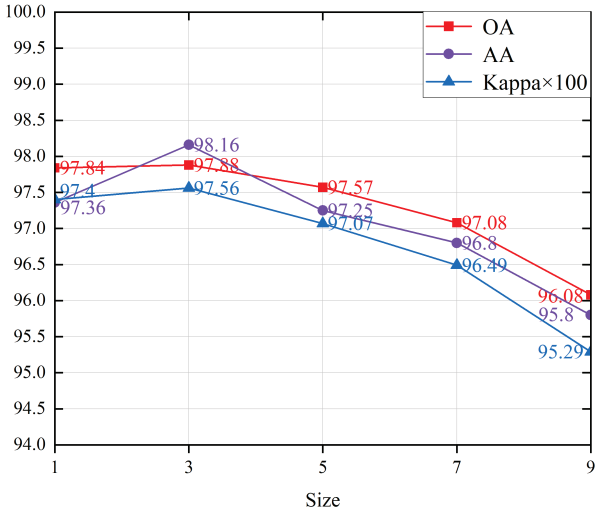

FIGURE 11. EXPERIMENTAL RESULTS OF CONVOLUTION KERNELS OF DIFFERENT SIZES IN CEA ON THE INDIAN PINE DATASET

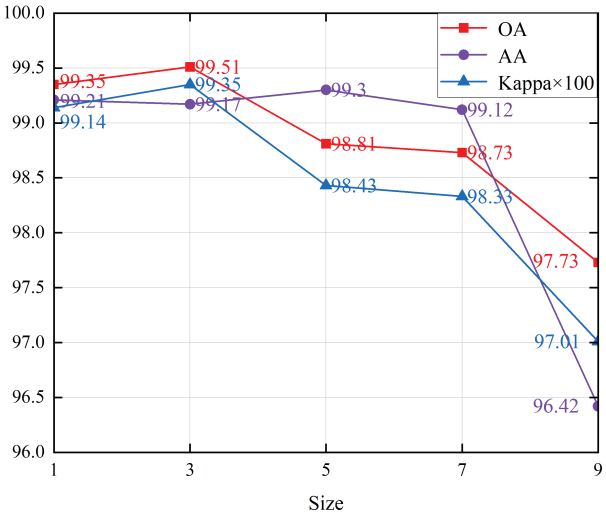

FIGURE 10. EXPERIMENTAL RESULTS OF CONVOLUTION KERNELS OF DIFFERENT SIZES IN CEA ON THE PAVIA UNIVERSITY DATASET

5. **Conclusions.** We propose a network model that effectively relieves the problem of error-prone object edge classification in hyperspectral images. The model uses the ECA channel attention mechanism for feature extraction of different channels of hyperspectral images. And the deformable convolution module is added, which can be adjusted based on

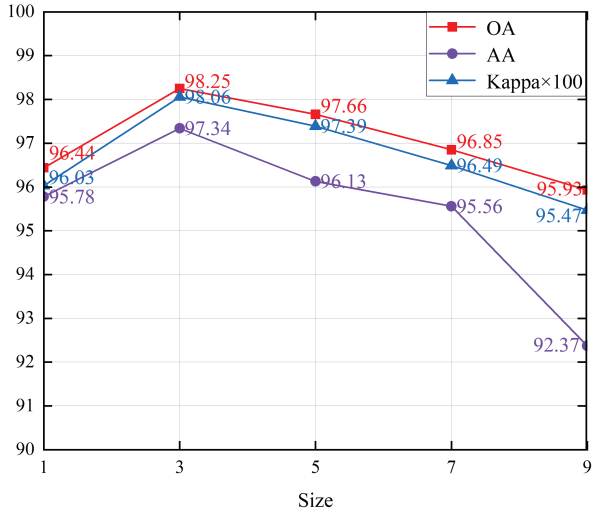

FIGURE 12. EXPERIMENTAL RESULTS OF CONVOLUTION KERNELS OF DIFFERENT SIZES IN CEA ON THE KSC DATASET

the complex and curved characteristics of the object contours, and the final classification results are optimized to a large extent. After comparing it with other current advanced network models, the practicality and robustness of our designed network are proven. However, network model training is still limited by hardware resources. Therefore, we will optimize the network model in the next work. We will further investigate the lightweight network model by reducing the number of parameters, provided that the network has a good classification effect.

Acknowledgment. This work is partially supported by the National Natural Science Foundation of China (No.61602157), and Key Scientific and Technological Projects in Henan Province, China (No.202102210167).

REFERENCES

[1] S. Singh, and S. KV, "Role of Hyperspectral imaging for Precision Agriculture Monitoring," *ADBU Journal of Engineering Technology*, 2022. [online]. Availabel:http://journals.dbuniversity.ac.in/ojs/index.php/AJET/article/view/3587

[2] E. Wang, F. Wang, S. Kumari, J.-H. Yeh, and C.-M. Chen, "Intelligent monitor for typhoon in IoT system of smart city," *The Journal of Supercomputing*, vol. 77, no. 3, pp. 3024–3043, 2021.

[3] A. Nisha, and A. Anitha, "Current Advances in Hyperspectral Remote Sensing in Urban Planning," *2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICICT)*, pp. 94–98, 2022.

[4] M.-E. Wu, J.-H. Syu, and C.-M. Chen, "Kelly-based options trading strategies on settlement date via supervised learning algorithms," *Computational Economics*, vol. 59, no. 4, pp. 1627–1644, 2022

[5] J. Tellez Gaytan, K. Ateeq, A. Rafiuddin, H. Alzoubi, T. Ghazal, T. Ahanger, S. Chaudhary, and G. Viju, "AI-Based Prediction of Capital Structure: Performance Comparison of ANN SVM and LR Models," *Computational Intelligence and Neuroscience*, vol. 2022, 8334927, 2022

[6] S. Kumar, A. Damaraju, A. Kumar, S. Kumari, and C.-M. Chen, "LSTM Network for Transportation Mode Detection," *Journal of Internet Technology*, vol. 22, no. 4, pp. 891–902, 2021

[7] Z. Wang, Zhao, Zhan and C. Yin, "Fine Crop Classification Based on UAV Hyperspectral Images and Random Forest," *ISPRS International Journal of Geo-Information*, vol. 11, no. 4, pp. 252, 2022.

[8] R. Li, K. Cui, R.H. Chan, and Robert J. Plemmons, "Classification of hyperspectral images using SVM with shape-adaptive reconstruction and smoothed total variation," *arXiv preprint arXiv:2203.15619*, 2022.

[9] N. Wang and X. Zeng, "Hyperspectral Data Classification Algorithm considering Spatial Texture Features," *Mobile Information Systems*, vol. 2022, 9915809, 2022.

[10] H. Su, Y. Hu, H. Lu, W. Sun, and Q. Du, "Diversity-Driven Multikernel Collaborative Representation Ensemble for Hyperspectral Image Classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 2861–2876, 2022.

[11] H. Su, Y. Gao, and Q. Du, "Superpixel-Based Relaxed Collaborative Representation With Band Weighting for Hyperspectral Image Classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–16, 2022.
[12] K. Sun, A. Wang, X. Sun, and T. Zhang, "Hyperspectral image classification method based on M-3DCNN-Attention," *Journal of Applied Remote Sensing*, vol. 16, no. 2, pp. 026507, 2022.
[13] B. Xu, W. Hou, Y. Wei, Y. Wang, and X. Li, "Minimalistic fully convolution networks (MFCN): pixel-level classification for hyperspectral image with few labeled samples," *Optics Express*, vol. 30, no. 10, pp. 16585–16605, 2022.
[14] H. Pan, M. Liu, H. Ge, and S. Chen, "Semi-supervised spatial–spectral classification for hyperspectral image based on three-dimensional Gabor and co-selection self-training," *Journal of Applied Remote Sensing*, vol. 16, no. 2, pp. 028501, 2022.
[15] B. Fang, Y. Liu, H. Zhang, and J. He, "Hyperspectral Image Classification Based on 3D Asymmetric Inception Network with Data Fusion Transfer Learning," *Remote Sensing*, vol. 14, no. 7, pp. 1711, 2022.
[16] Z. Lv, X.-M. Dong, J. Peng, and W. Sun, "ESSINet: Efficient Spatial–Spectral Interaction Network for Hyperspectral Image Classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–15, 2022.
[17] Y.-L. Chang, T.-H. Tan, W.-H. Lee, L. Chang, Y.-N. Chen, K.-C. Fan, and M. Alkhaleefah, "Consolidated Convolutional Neural Network for Hyperspectral Image Classification," *Remote Sensing*, vol. 14, no. 7, pp. 1571, 2022.
[18] H. Xu, W. He, L. Zhang, and H. Zhang, "Unsupervised Spectral–Spatial Semantic Feature Learning for Hyperspectral Image Classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–14, 2022.
[19] W. Guo, G. Xu, B. Liu, and Y. Wang, "Hyperspectral Image Classification Using CNN-Enhanced Multi-Level Haar Wavelet Features Fusion Network," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.
[20] W. Hu, Y. Huang, L. Wei, F. Zhang, and H. Li, "Deep convolutional neural networks for hyperspectral image classification," *Journal of Sensors*, vol. 2015, 258619, 2015
[21] M. Zhang, W. Li, and Q. Du, "Diverse region-based CNN for hyperspectral image classification," *IEEE Transactions on Image Processing*, vol. 27, no. 6, pp. 2623–2634, 2018.
[22] V. Kumar, and R. Singh, and Y. Dua, "Morphologically dilated convolutional neural network for hyperspectral image classification," *Signal Processing: Image Communication*, vol. 101, pp. 116549, 2022.
[23] Q. Ren, B. Tu, Q. Li, W. He, and Y. Peng, "Multiscale Densely Connected Attention Network for Hyperspectral Image Classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 5115–5130, 2022
[24] C. Shi, J. Sun, and L. Wang, "Hyperspectral Image Classification Based on Spectral Multiscale Convolutional Neural Network," *Remote Sensing*, vol. 14, no. 8, pp. 1951, 2022
[25] X. Wang, and Y. Fan, "Multiscale Densely Connected Attention Network for Hyperspectral Image Classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 1617–1628, 2022
[26] H. Liu, W. Li, X.-G. Xia, M. Zhang, C.-Z. Gao, and R. Tao, "Central attention network for hyperspectral imagery classification," *IEEE Transactions on Neural Networks and Learning Systems*, vol.2022, 35271453, 2022.
[27] S. Pande, and B. Banerjee, "HyperLoopNet: Hyperspectral image classification using multiscale self-looping convolutional networks," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 183, pp. 422–438, 2022.
[28] N. Li, and Z. Wang, "Spatial Attention Guided Residual Attention Network for Hyperspectral Image Classification," *IEEE Access*, vol. 10, pp. 9830–9847, 2022.
[29] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A convnet for the 2020s," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11976–11986, 2022.
[30] N. Ma, X. Zhang, H.-T. Zheng, and J. Sun, "Shufflenet v2: Practical guidelines for efficient cnn architecture design," *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 116–131, 2018.
[31] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," *International Conference on Information Engineering and Computer Science*, pp. 7132–7141, 2018.

[32] Q.-L. Zhang, and Y.-B. Yang, "Sa-net: Shuffle attention for deep convolutional neural networks," *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2235–2239, 2021.

[33] W. Li, E. Tramel, S. Prasad, and J. Fowler, "Nearest regularized subspace for hyperspectral classification," *Computational Intelligence and Neuroscience*, vol. 52, no. 1, pp. 477–489, 2013.

[34] S. Ding, and L. Chen, "Classification of hyperspectral remote sensing images with support vector machines and particle swarm optimization," *International Conference on Information Engineering and Computer Science*, pp. 1–5, 2009.

[35] S. Roy, G. Krishna, S. Dubey, and B. Chaudhuri, "HybridSN: Exploring 3-D–2-D CNN feature hierarchy for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 2, pp. 277–281, 2019.

[36] T. Tang, X. Pan, X. Luo, X. Gao, and W. Yan, "TBTA-D2Net: a novel hyperspectral image classification method based on triple-branch ternary-attention mechanism and improved dense2Net," *IEEE Transactions on Geoscience and Remote Sensing*, 2022. [online]. Available: https://doi.org/10.21203/rs.3.rs-1989925/v1

[37] P. Valsalan, and C. Latha G, "Hyperspectral Image Classification Model Using Squeeze and Excitation Network with Deep Learning," *Computational Intelligence and Neuroscience*, vol. 2022, 9430779, 2022.