



Article

A Hybrid-Scale Feature Enhancement Network for Hyperspectral Image Classification

Dongxu Liu ^{1,2,3} , Tao Shao ^{1,2,3}, Guanglin Qi ^{1,2,3}, Meihui Li ^{1,2,3} and Jianlin Zhang ^{1,2,3,*}

- ¹ National Key Laboratory of Optical Field Manipulation Science and Technology, Chinese Academy of Sciences, Chengdu 610209, China; liudongxu18@mails.ucas.ac.cn (D.L.); taoshao@mail.nwpu.edu.cn (T.S.); qiguanglin@ioe.ac.cn (G.Q.); limeihui@ioe.ac.cn (M.L.)
² Key Laboratory of Optical Engineering, Chinese Academy of Sciences, Chengdu 610209, China
³ Institute of Optics and Electronics, Chinese Academy of Sciences, Chengdu 610209, China
* Correspondence: jlin@ioe.ac.cn

Abstract: Due to their devastating ability to extract features, convolutional neural network (CNN)-based approaches have achieved tremendous success in hyperspectral image (HSI) classification. However, previous works have been dedicated to constructing deeper or wider deep learning networks to obtain exceptional classification performance, but as the layers get deeper, the gradient disappearance problem impedes the convergence stability of network models. Additionally, previous works usually focused on utilizing fixed-scale convolutional kernels or multiple available, receptive fields with varying scales to capture features, which leads to the underutilization of information and is vulnerable to feature learning. To remedy the above issues, we propose an innovative hybrid-scale feature enhancement network (HFENet) for HSI classification. Specifically, HFENet contains two key modules: a hybrid-scale feature extraction block (HFEB) and a shuffle attention enhancement block (SAEB). HFEB is designed to excavate spectral-spatial structure information of distinct scales, types, and branches, which can augment the multiplicity of spectral-spatial features while modeling the global long-range dependencies of spectral-spatial informative features. SAEB is devised to adaptively recalibrate spectral-wise and spatial-wise feature responses to generate the purified spectral-spatial information, which effectively filters redundant information and noisy pixels and is conducive to enhancing classification performance. Compared with several sophisticated baselines, a series of experiments conducted on three public hyperspectral datasets showed that the accuracies of OA, AA, and Kappa all exceed 99%, demonstrating that the presented HFENet achieves state-of-the-art performance.

Keywords: convolutional neural network; hybrid-scale feature extraction; hyperspectral image classification; shuffle attention enhancement



Citation: Liu, D.; Shao, T.; Qi, G.; Li, M.; Zhang, J. A Hybrid-Scale Feature Enhancement Network for Hyperspectral Image Classification. *Remote Sens.* **2024**, *16*, 22. <https://doi.org/10.3390/rs16010022>

Academic Editor: Akira Iwasaki

Received: 27 October 2023

Revised: 11 December 2023

Accepted: 16 December 2023

Published: 20 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Hyperspectral imaging, a spectrum-image merging technology combining spectral detection and imaging techniques, utilizes diverse sensors to distinguish the electromagnetic waves reflected from objects and precisely describes the physical characteristics of objects [1,2]. A hyperspectral image (HSI) possesses plentiful spectral and spatial information, which has been widely adopted in extensive application areas, such as precision agriculture [3], environmental monitoring [4], mineral exploration [5], and urban planning [6]. HSI classification has become a research hotspot in pattern recognition and image processing, which is devoted to assigning a unique category label to each spatial pixel [7–10]. However, HSI classification is still a challenging issue, i.e., especially spatial variability and the curse of dimensionality, thereby increasing the difficulty of classification. The former is induced by factors such as light angle [11], and atmospheric interference [12], which leads to the same object presenting different characteristics. The latter is caused by the unbalance between high-dimensionality features and limited samples, which easily

results in overfitting. Consequently, how to capture more representative and discriminative features from the original data is a critical problem in HSI classification.

Initially, bountiful HSI classification techniques have been presented, which focus on two stages, i.e., feature engineering and classifier training. Feature engineering aims to reduce the spectral dimension of HSI data and capture informative features or bands. Feature engineering generally contains two regular methods, i.e., feature selection and feature extraction. Feature selection aims to retain important spectral bands for the following tasks and fade out unnecessary ones. Representative methods have a spectral angle mapper (SAM) [13], Jeffries-Matusita distance [14], Bhattacharyya distance [15], etc. Feature extraction can easily detect varying categories by converting HSI data from a high-dimension space to a low-dimension space. Typical methods include principal component analysis (PCA) [16], independent component analysis (ICA) [17], minimum noise fraction (MNF) [18], etc. Features generated by feature engineering are fed to classifiers for classification tasks. Common classifiers involve support vector machine (SVM) [19], manifold ranking (MR) [20], random forests (RF) [21], etc. However, the classification methods mentioned above only utilize spectral information and do not fully consider spatial information in the target area. Compared with methods-based spectral features, many researchers have demonstrated that making full use of spatial and spectral information helps to strengthen the classification results. In general, these methods exploit multi-kernel learning (MKL) [22], morphological profiles (MP) [23], sparse representation (SR) [24], etc., to extract spatial features. Nevertheless, both spectral-features-based and spectral-spatial-features-based classification methods depend on hand-crafted features with poor generalization ability and limited representation ability, which extremely degenerate the classification performance.

Of late, due to the powerful representation of learning potentials, deep learning (DL)-based methods have acquired tremendous advancements in HSI classification. For example, Chen et al. applied a multilayer stacked autoencoder (SAE) to extract deep features for HSI classification [25]. To obtain spatial-spectral features, Li et al. utilized a multilayer deep belief network (DBN) and a single restricted Boltzmann machine (RBM) [26]. Hong et al. designed a supervised mini graph convolutional network (GCN) for HSI classification [27]. To provide new insight into HSI classification, Hang et al. devised a multitask generative adversarial network (GAN) [28]. Hang et al. constructed a cascaded recurrent neural network (RNN) to fully excavate spectral information to achieve high-accuracy HSI classification [29]. Li et al. built a two-stream convolutional neural network (CNN) to simultaneously capture spectral and spatial features [30]. To model the global relationships of HSI, Zu et al. proposed a cascaded convolution-based transformer [31]. In the abovementioned network models, CNN is always a considerable and indispensable module [32–36]. Giving credit to the characteristics of weight-sharing and local connection, Hu et al. built a 1D CNN model to explore spectral information [37]. Xu et al. designed a pixel-to-pixel, end-to-end spectral-spatial fully convolutional network for HSI classification [38]. To tackle the information leakage of the training, Zou et al. constructed a spectral-spatial 3D fully convolutional network, which can exploit the spectral-spatial joint features and semantic information [39]. Zhang et al. presented a CNN based on varying region inputs to effectively extract contextual interactional information [40]. A multiscale and cross-level attention learning network was devised by Xu et al., which can use multiscale information from local and global views [41]. Although the classification methods-based CNN has demonstrated remarkable success, there are still some drawbacks. To be more specific, the squared region of convolutional kernel size gravely limits the capacity of methods based on CNN to acquire long-range dependencies. Additionally, the informative features captured by CNN commonly involve redundant features and noise, adverse to the classification performance. Consequently, it is an urgent problem to overcome the challenge of finding a way to obtain significant features to enhance HSI classification.

Lately, many promising tricks have been integrated into CNNs, such as neural network search strategy [42], multiple available receptive fields with varying scales [43],

sample augmentation [44], residual learning [45], attention mechanism [46], and dense connection [47]. From the perspective of imaging procedures, Chen et al. built a virtual sample augmentation approach to create training data [48]. Cao et al. constructed a compressed CNN to effectively enhance the classification performance of the student network by using virtual samples to describe the teacher network's classification boundary [49]. To sufficiently exploit information from varying scales of HSI, Xie et al. built a multi-scale densely-connected convolutional network [50]. Wang et al. used a multiscale ghost module to capture more distinguishable information using simple operations [51]. Zhu et al. designed a spectral attention block and spatial attention to adaptively emphasize necessary spectral bands and important spatial pixels [52]. Roy presented an improved spectral-spatial ResNet to obtain the spectral-spatial joint information [53]. Zhang et al. devised cascaded parallel improved residual blocks to capture spectral-spatial features [54]. To degrade the computation cost and obtain better classification accuracy, Dong et al. combined the dense connection with attention modules [55].

In the article, we present a hybrid-scale feature enhancement network (HFENet) for HSI classification. HFENet contains two important submodules: hybrid-scale feature extraction block (HFEB) and shuffle attention enhancement block (SAEB). HFEB is devised to extract spectral-spatial structure information of different types and scales, thereby modeling the global long-range dependencies of spectral-spatial features. HFEB consists of two parallel branches, and the nuclear component of each branch is a heterogenous feature refine block (HFRB), where the upper branch has two HFRBs, the latter branch has an HFEB, and the convolutional kernel size of each HFRB is different. HFRB is designed to capture the local dependencies of spectral-spatial features. SAEB is constructed to effectively dispel the redundant information and noisy pixels, further strengthening the discrimination ability of spectral-spatial informative features for HSI classification. In this context, the main contributions of the proposed work rely on the following:

- (1) We construct a heterogenous feature refine block (HFRB) to capture the internal correlations of different channels and the external interactions of all channels, which complement each other, thereby enhancing the local dependencies of spectral-spatial features.
- (2) Different from existing multiscale feature extraction strategies, our designed hybrid-scale feature extraction block (HFEB) exploits multiple HFRBs to obtain more discriminative and representative spectral-spatial structure information of distinct scales, types, and branches, which can not only augment the multiplicity of spectral-spatial features but also model the global long-range dependencies of spectral-spatial features.
- (3) To effectively fade out the redundant information and noisy pixels, we devise a shuffle attention enhancement block (SAEB) to adaptively recalibrate spectral-wise and spatial-wise feature responses to generate the purified spectral-spatial information, which is conducive to enhancing the classification performance.

The rest of this work is formulated as follows. Section 2 describes the proposed approach in detail. Section 3 provides the relevant experimental results and comparisons with several state-of-the-art methods. Section 4 concludes this work.

2. Methods

2.1. Framework of HFENet Model

Figure 1 graphically illustrates the framework of our presented HFENet, which is composed of an initial block, two HFEBs, a SAEB, and an output block. First, considering the classical curse of dimensionality issue of HSI, we conduct the PCA algorithm on raw HSI to reduce spectral band numbers and effectively alleviate the interference of high correlation between spectral bands, where 40 spectral bands remain. Second, to effectively reduce the training time and fully exploit the property of HSI containing both spectral and spatial data, a 3D data cube $x \in R^{7 \times 7 \times 40}$ consisting of the target pixel and its adjacent pixels is used as the input data of our presented HFENet, where 7, 7, and 40 represent the dimensions of height, width, and spectrum, respectively. Third, the 3D data cube is transmitted to the initial block to obtain general spectral-spatial features. The initial block contains a

3D convolutional layer with 128 filters of $1 \times 1 \times 40$ size, a 3D convolutional layer with 128 filters of $3 \times 3 \times 1$ size, 2 BN layers, and 2 PReLU activation functions. Then, the initial spectral–spatial features are transmitted to two HFEBs to extract more discriminative and representative global long-range dependencies of spectral–spatial features. Furthermore, these features are transmitted to a SAEB to filter unnecessary information and dispel interference of noises, thus achieving spectral–spatial feature purification. Finally, the output block is utilized to generate the probabilities of 16 categories. The output block involves a 2D GAP operation, two fully connected layers, two dropouts, and a softmax layer. In addition, to avoid the overfitting problem, L2 regularization is also introduced into the proposed HFENet. We will depict two primary submodules of our presented HFENet: hybrid-scale feature extraction block (HFEB) and shuffle attention enhancement block (SAEB).

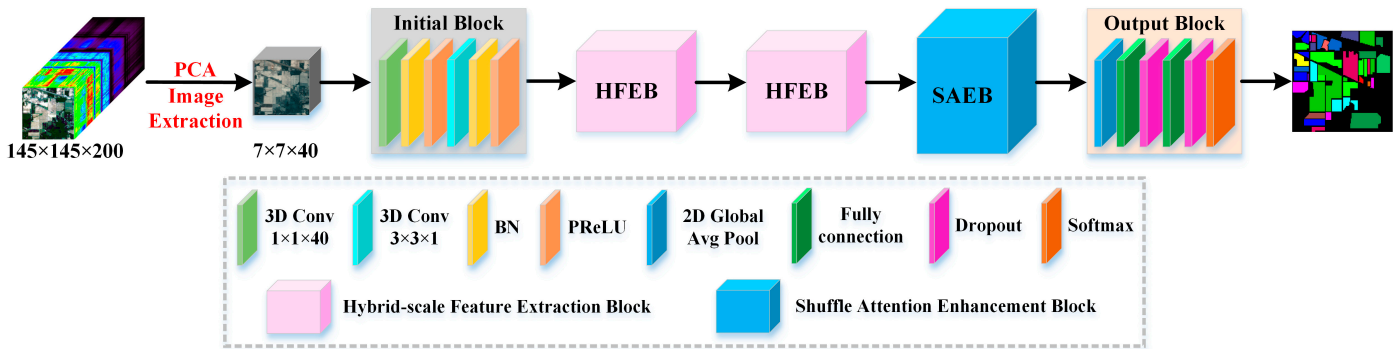


Figure 1. Overview of hybrid-scale feature enhancement network (HFENet).

2.2. Heterogenous Feature Refine Block

In recent years, to meet higher-quality computer vision task requirements, numerous researchers have ameliorated network performance by exploring features of different channels or layers. For example, Gao et al. constructed a triple-branch attention block to capture interactions across different spatial regions, spectral bands, and channels [56]. Wang et al. proposed an attention mechanism module to obtain the weight information of channel dimension, spectral dimension, and spatial dimension, respectively [57]. To improve SISR performance, Zhang et al. built a residual channel attention mechanism, which can not only reinforce interdependencies of varying channels but also adaptively discard plentiful low-frequency features [58]. Inspired by the above approaches, we construct an innovative, heterogenous feature refine block (HFRB) to strengthen internal and external interactions of different channels and layers while enriching the local dependencies of spectral–spatial features. HFRB exploits a heterogeneous architecture in a parallel manner, which contains a symmetric residual unit (SRU) and a complementary residual unit (CRU). The architecture of our devised HFRB is provided in Figure 2. The input data of HFRB can be referred to $X \in R^{H \times W \times C}$, where H , W , and C are dimension of height, width, and channel, respectively.

Symmetric Residual Unit: SRU adopts a pair of twin branches to boost internal relations of different channels. More precisely, we split the input data X into two sub-branches: $X_1 \in R^{H \times W \times (C/2)}$ and $X_2 \in R^{H \times W \times (C/2)}$. Each sub-branch involves two $Conv + BN + ReLU$ layers and a $Conv + BN$ layer, where 2D convolutional operation with $C/2$ filters of $n \times n$ size is utilized to excavate spectral–spatial features of varying channels, BN is utilized to strengthen the network performance and ReLU is utilized to obtained features map into a non-linearity. In addition, to enhance information propagation form shallow to deep layers and avoid loss, we apply the skip transmission to each sub-branch. The formulas of upper sub-branch can be explained as follows:

$$U_{11} = F_1(X_1) \quad (1)$$

$$U_{12} = F_1(U_{11}) \quad (2)$$

$$U_{13} = F_2(U_{12}) \quad (3)$$

$$U = U_{13} + X_1 \quad (4)$$

where X_1 and U are the input data and output data of upper sub-branch. $F_1(\cdot)$ represents the composite function containing $Conv + BN + ReLU$ layer, $F_2(\cdot)$ represents the composite function containing $Conv + BN$ layer, and $+$ is element-wise addition operation.

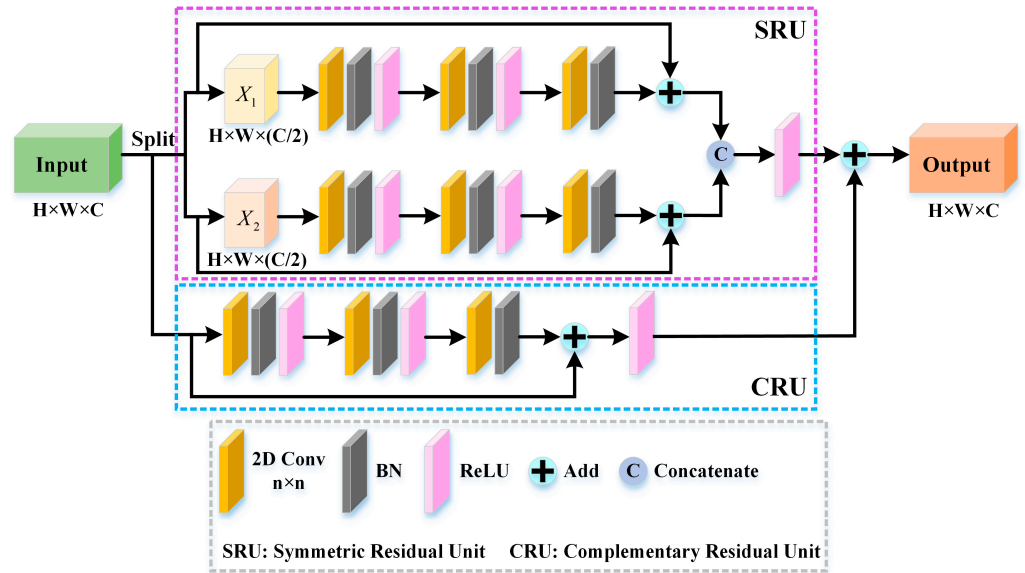


Figure 2. Schematic of heterogeneous feature refine block (HFRB).

The formulas of bottom sub-branch are the same as those of upper sub-branch, which can be explained as follows:

$$B_{21} = F_1(X_2) \quad (5)$$

$$B_{22} = F_1(B_{11}) \quad (6)$$

$$B_{23} = F_2(B_{12}) \quad (7)$$

$$B = B_{13} + X_2 \quad (8)$$

where X_2 and B are the input data and output data of bottom sub-branch. Finally, we exploit a plain concatenate operation to aggregate the output features of two sub-branches and use ReLU to strengthen the nonlinear ability of network. The formulas can be explained as follows:

$$O_1 = \sigma([U, B]) \quad (9)$$

where O_1 stands for the output data of SRU, $[]$ refers to the concatenate operation, σ and is ReLU activation function.

Complementary Residual Unit: CRU is devised to enhance the robustness of spectral-spatial features by learning external correlations of all the channels, which complements SRU. CRU contains two $Conv + BN + ReLU$ layers and a $Conv + BN$ layer, where 2D convolutional operation with C filters of $n \times n$ size is utilized to extract spectral-spatial features of entire channels, different from those of SRU. Uniformly, the skip transmission is also introduced to CRU to avert the loss of information. Finally, the ReLU activation function is exploited to boost the nonlinear ability of model. The formulas of CRU can be explained as follows:

$$C_{11} = F_1(X) \quad (10)$$

$$C_{12} = F_1(C_{11}) \quad (11)$$

$$C_{13} = F_2(C_{12}) \quad (12)$$

$$C_{14} = \sigma(C_{13}) \quad (13)$$

$$O_2 = C_{14} + X \quad (14)$$

SRU and CRU can obtain internal and external corrections of split channels and all channels, which complement each other. For that, we employ the element-wise addition operation to SRU and CRU to generate richer deep and wide local dependencies of spectral-spatial features. The formulas of HFRB can be explained as follows:

$$O = O_1 + O_2 \quad (15)$$

where O represents the output data of HFRB, and $+$ is element-wise addition operation.

2.3. Hybrid-Scale Feature Extraction Block

With the increased demand for HSI classification tasks, many scholars have focused on utilizing multiple receptive fields to explore luxuriant spectral-spatial information, thereby achieving remarkable performance. For example, Zhang et al. designed a multi-scale dense network, which can not only fully use the varying scale information of network structure but also aggregate the scale information of the entire network for HSI classification [59]. Xie et al. built a multiscale densely-connected convolutional network to effectively capture spectral-spatial features of multiple scales [50]. To tackle larger intraclass variability, Safari et al. constructed a multiscale deep learning by combining diverse CNNs for HSI classification [43]. Compared with the fixed-scale extraction manner, utilizing multiple different receptive fields contributes to enhancing HSI classification performance. Inspired by the above approaches, we devise an innovative hybrid-scale feature extraction block (HFEB) exploiting spectral-spatial structure information of distinct types and scales to increase the multiplicity of spectral-spatial features while modeling global long-range dependencies of spectral-spatial features. The structure of our proposed HFEB is provided in Figure 3.

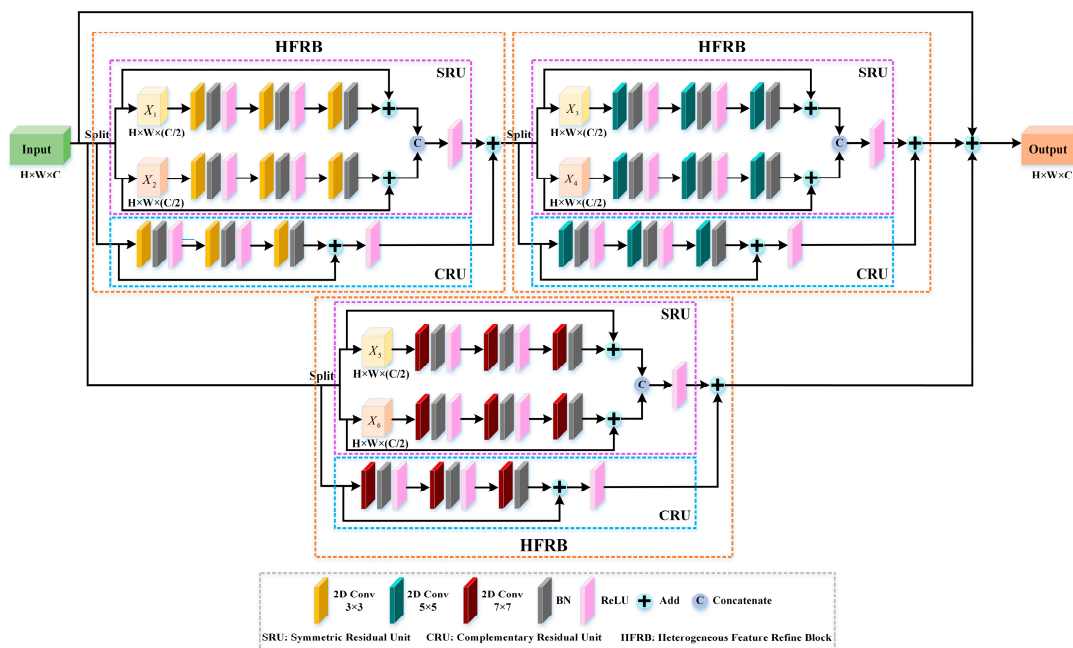


Figure 3. Schematic of hybrid-scale feature extraction block (HFEB).

Different from prior works based on convolution operations with multiple different receptive fields, our presented HFEB utilizes promising functional HFRBs to obtain spectral-

spatial information of distinct types and scales. To be more specific, HFEB is composed of two parallel branches: the upper branch contains two HFEBs with 2D convolutional operations with 3×3 size and 5×5 size, respectively; the lower branch contains one HFEB with 2D convolutional operations with 7×7 size. The structure of HFRB is provided in Figure 2. Second, we utilize the element-wise addition operation to aggregate the output data of two branches, thereby obtaining the global long-range dependencies of spectral–spatial features. Furthermore, to avert the loss of information, the skip transmission is also applied to HFEB. The formulas of HFEB can be explained as follows:

$$y_1 = HFRB_{3 \times 3}(X) \quad (16)$$

$$y_2 = HFRB_{5 \times 5}(y_1) \quad (17)$$

$$y_3 = HFRB_{7 \times 7}(X) \quad (18)$$

$$y = X + y_2 + y_3 \quad (19)$$

where X and y are the input data and output data of HFEB. $HFEb(\cdot)$ refers to the entire treatment process of HFRB, and the subscripts refer to the convolutional kernel size of HFRB. y_1 , y_2 , and y_3 represent the output data of each HFRB, respectively.

2.4. Shuffle Attention Enhancement Block

The attention mechanism mimicking the perception system of humans is one of the most distinguished ideas in the DL domain, which is utilized to focus on the information regions more relevant to computer-vision tasks and filter irrelevant ones. For example, Dong et al. constructed an attention module composed of spatial and spectral axes to emphasize the salient spatial–spectral information [60]. Guo et al. devised a spectral–spatial connected attention mechanism, which integrates spatial attention module and spectral attention module to enhance the distinguishing capacity of spatial pixels and spectral bands [55]. Zhu et al. built a spectral attention module to obtain useful spectral bands and a spatial attention module for the adaptive selection of spatial pixels [52]. Inspired by the above approaches, we design a shuffle attention enhancement block (SAEB) to adaptively recalibrate spectral-wise and spatial-wise feature responses, which effectively eliminates redundant information and noisy pixels, thereby heightening the discriminative ability of spectral–spatial features. The structure of our proposed SAEB is provided in Figure 4. As seen in Figure 4, the SAEB is composed of four prominent parts: feature grouping, spectral enhancement branch, spatial enhancement branch, and feature aggregating.

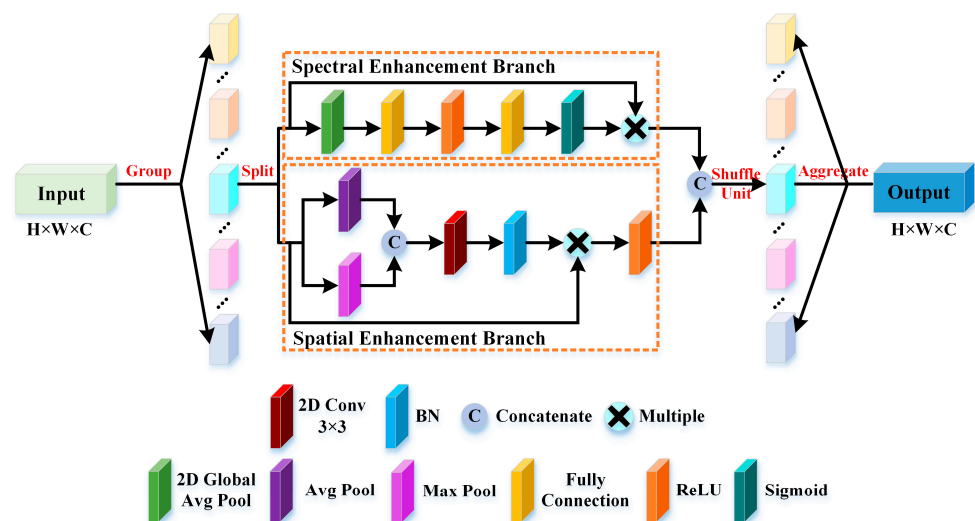


Figure 4. Schematic of shuffle attention enhancement block (SAEB).

Feature Grouping: $X \in R^{H \times W \times C}$ refers to the input data of SAEB, where H , W , and C are dimension of height, width, and channel, respectively. The input data X are divided into n subsets, and each subset is also divided into two parts along the spectral dimension: $x_1 \in R^{H \times W \times (C/2n)}$ and $x_2 \in R^{H \times W \times (C/2n)}$. x_1 and x_2 are fed into the spectral enhancement branch and spatial enhancement branch, respectively.

Spectral Enhancement Branch: This is built to reassign the weights to spectral bands, emphasizing the meaningful bands and fading out the irrelevant ones. Concretely, first, a 2D global average pooling is used to convert $x_1 \in R^{H \times W \times (C/2n)}$ to $x_1' \in R^{1 \times 1 \times (C/2n)}$. Second, two fully connected layers, a ReLU activation function and a sigmoid activation function, are adopted to generate the weights of spectral bands $W_{spectral}$. Finally, $W_{spectral}$ is multiplied by x_1 to obtain the local spectral-wise feature responses $X_{spectral}$. The formulas of spectral enhancement branch can be explained as follows:

$$W_{spectral} = \delta(FC_2(\sigma(FC_1(GAP(x_1)))))) \quad (20)$$

$$X_{spectral} = x_1 \otimes W_{spectral} \quad (21)$$

where σ and δ are ReLU and sigmoid activation functions. \otimes is the multiplication operation.

Spatial Enhancement Branch: This is constructed to reassign the weights to spatial pixels, strengthening pixels that are conducive for classification in the pixel-centered neighborhood or those from the same class as the center pixel and suppressing unimportant ones. Specifically, first, an average pooling operation and a max pooling operation are adopted to convert $x_2 \in R^{H \times W \times (C/2n)}$ to $x_{21}' \in R^{H \times W \times 1}$ and $x_{22}' \in R^{H \times W \times 1}$, respectively. Second, the average pooling feature and max pooling feature are aggregated by concatenation operation. Third, the aggregated features are transmitted to a 2D convolutional layer with 3×3 size to generate the weights of spatial pixels $W_{spatial}$. The $W_{spatial}$ is sent to a BN layer to strengthen the network classification performance. Finally, $W_{spatial}$ is multiplied by x_2 to obtain the local spatial-wise feature responses $X_{spatial}$. In addition, ReLU is utilized to turn feature map into a non-linearity. The formulas of spatial enhancement branch can be explained as follows:

$$W_{spatial} = Conv([AP(x_2), MP(x_2)]) \quad (22)$$

$$X_{spatial} = \sigma(BN(W_{spatial}) \otimes x_2) \quad (23)$$

where $AP(\cdot)$ and $MP(\cdot)$ are the average pooling operation and max pooling operation, respectively. $Conv(\cdot)$ is the 2D convolutional operation, and $BN(\cdot)$ is the BN layer. σ is the ReLU activation function. $[]$ and \otimes denote the concatenation operation and multiplication operation, respectively.

Feature Aggregating: The concatenation operation is utilized to integrate the local spectral-wise feature responses $X_{spectral}$ and the local spatial-wise feature responses $X_{spatial}$ into a new subset, thus obtaining the local spectral–spatial feature responses. To encourage the cross-information flow of local spectral–spatial feature responses between different subsets, we also introduce the shuffle unit into our proposed SAEB. Finally, we aggregate all local spectral–spatial feature responses to obtain global spectral–spatial feature responses.

3. Experimental Results and Discussion

3.1. Hyperspectral Datasets and Setup

To estimate the classification performance of our developed HFENet, we adopted three publicly available datasets, i.e., Pavia University (UP), Indian Pines (IP), and Houston 2013 datasets.

The UP dataset was captured by the ROSIS-3 sensor over the city of Pavia, Italy. The image possesses 610×340 pixels with a geometric resolution of 1.3 m. It is composed of 9 categories and 115 spectral bands ranging from about 0.43 to 0.86 μm . The corrected image contains 103 spectral bands after removing 12 noisy bands.

The IP dataset was collected by the AVIRIS sensor over northwestern Indiana, USA. The image involves 145×145 pixels with a geometric resolution of 20 m. It constitutes

16 categories and 224 spectral bands ranging from about 0.4 to 2.5 μm . The corrected image has 200 spectral bands; after removing 20 bands, it cannot be reflected by water.

The Houston 2013 dataset was obtained by the ITRES CASI-1500 instrument over the University of Houston campus, USA. The image contains 349×1905 pixels with a geometric resolution of 2.5 m. It has 15 categories and 144 spectral bands ranging from about 0.38 to 1.05 μm .

Tables 1–3 provide the number of samples of each category for training and testing. The validation experiments were performed in a TensorFlow 2.3, Keras 2.4.3, CUDA 10.1, and Python 3.6 environment utilizing an Intel(R) Core(TM) i7-9700F CPU made by the Intel corporation and an NVIDIA GeForce RTX 2060 SUPER 6 GB GPU made by the NVIDIA corporation, procuring from Chengdu, China. The epoch and batch size influence the classification performance of our proposed: if the epoch and batch size are too small, the training process of the model will be unstable and easily disturbed by noisy data; if the epoch and batch size are too large, the training time of model will be too long, and the learning ability of model will be limited. Therefore, setting the suitable epoch and batch size is vital for our proposed HFENet. For the UP, IP, and Houston2013 datasets, the training epochs were set to 100, 200, and 200; the batch size was set to 16, 16, and 16, respectively. The Adam's Optimizer was chosen as the optimizer, and the learning rate was defined as 0.0005. The average accuracy (AA), overall accuracy (OA), and Kappa coefficient (Kappa) were used as criteria metrics to evaluate the classification performance.

Table 1. Data description of UP dataset.




No.	Color	Class	Train	Test
1		Asphalt	1326	5305
2		Meadows	3729	14,920
3		Gravel	419	1680
4		Trees	612	2452
5		Metal sheets	269	1076
6		Bare Soil	1005	4024
7		Bitumen	266	1064
8		Bricks	736	2946
9		Shadows	189	758
Total			8551	34,225

Table 2. Data description of IP dataset.



















No.	Color	Class	Train	Test
1		Alfalfa	10	36
2		Corn-notill	286	1142
3		Corn-mintill	166	664
4		Corn	48	189
5		Grass-pasture	97	386
6		Grass-trees	146	584
7		Grass-pasture-mowed	6	22
8		Hay-windrowed	96	382
9		Oats	4	16
10		Soybean-notill	195	777
11		Soybean-mintill	491	1964
12		Soybean-clean	119	474
13		Wheat	41	164
14		Woods	253	1012
15		Buildings-grass-tree	78	308
16		Stone-steel-towers	19	74
Total			2055	8194

Table 3. Data description of the Houston 2013 dataset.

No.	Color	Class	Train	Test
1		Healthy grass	251	1000
2		Stressed grass	251	1003
3		Synthetic grass	140	557
4		Trees	249	995
5		Soil	249	993
6		Water	65	260
7		Residential	254	1014
8		Commercial	249	995
9		Road	251	1001
10		Highway	246	981
11		Railway	247	988
12		Parking Lot 1	247	986
13		Parking Lot 2	94	375
14		Tennis Court	86	342
15		Running Track	132	528
Total			3011	12,018

3.2. Classification Comparison with State-of-the-Art Models

Our proposed HFENet model is compared with eleven outstanding classification approaches to comprehensively demonstrate the superiority of the HFENet model. Eleven classification methods are broadly divided into two groups: one containing SVM¹, RF¹, KNN¹, and GaussianNB¹ belongs to traditional ML; the other involving HybridSN² [61], RSSAN³ [52], MSRN [62], MAFN⁴ [63], DCRN⁵ [64], DMCN [65], and MSDAN [57] belongs to DL. Specifically, HybridSN is composed of 2D CNN and a spectral–spatial 3D CNN to achieve maximum possible accuracy. RSSAN exploits a spectral–spatial attention learning module to filter unimportant information and strengthen beneficial information while using a spectral–spatial feature learning module to refine the learned features. MSRN utilizes the depthwise separable convolution with a mixed depthwise convolution layer to replace the convolutional layer to construct residual blocks, which can emphasize the feature representation ability. MAFN constructs a spatial feature extraction module, a spectral feature extraction module, and a spectral–spatial feature extraction module to obtain more representative features. DCRN designs two parallel branches and a spatial–spectral fusion structure to extract joint features. DMCN involves coordinate attention, a grouped residual 2D CNN, and a dense 3D CNN to mine fusion information. MSDAN applies three different scale modules with dense connections to achieve feature reuse while embedding spectral–spatial–channel attention to improve classification performance. For fairness, all experiments stochastically pick 20% of labeled data as the training size for three datasets. Obtained classification results are shown in Tables 4–6.

Table 4. Quantitative comparison on the UP dataset.

No.	SVM ¹	RF ¹	KNN ¹	GaussianNB ¹	HybridSN ²	RSSAN ³	MSRN	MAFN ⁴	DCRN ⁵	DMCN	MSDAN	HFENet
1	76.52	93.17	91.34	96.01	99.53	99.79	99.83	100.00	99.46	97.86	99.89	99.94
2	85.94	89.70	88.74	80.20	99.99	99.91	99.99	100.00	99.95	99.94	99.76	99.98
3	83.78	85.32	71.81	28.09	99.58	99.13	100.00	100.00	97.99	93.11	73.50	99.82
4	95.94	94.83	96.61	50.01	100.00	99.88	96.61	92.81	84.55	86.88	99.71	100.00
5	99.81	99.27	99.33	80.49	99.72	99.91	98.63	100.00	86.46	65.72	100.00	100.00
6	95.86	91.48	82.30	37.77	100.00	99.33	99.62	98.72	99.88	86.54	99.63	100.00
7	0.00	86.90	74.87	40.61	98.88	99.53	99.91	100.00	96.89	59.21	100.00	99.91
8	67.39	83.01	80.68	69.98	99.86	97.33	93.49	97.18	96.08	98.95	99.32	100.00
9	99.87	99.87	100.00	100.00	99.45	99.21	95.61	97.04	96.84	92.60	99.06	99.74
OA (%)	83.89	90.38	87.63	67.46	99.83	99.53	98.94	98.98	97.55	92.53	98.00	99.96
AA (%)	70.90	87.71	85.14	73.03	99.35	99.23	98.21	99.19	94.83	90.08	97.13	99.94
Kappa × 100	77.82	87.05	83.34	58.48	99.78	99.38	98.59	98.66	96.76	90.17	97.35	99.95

¹ <http://scikit-learn.org>, ² <https://github.com/gokriznastic/HybridSN>, ³ <https://github.com/mmhzhu/RSSAN>, ⁴ <https://github.com/Li-ZK/MAFN-2021>, ⁵ <https://github.com/Li-ZK/DCRN-2021>. The bold font highlights which mechanic works best.

Table 5. Quantitative comparison on the IP dataset.

No.	SVM	RF	KNN	GaussianNB	HybridSN	RSSAN	MSRN	MAFN	DCRN	DMCN	MSDAN	HFENet
1	0.00	86.67	36.36	31.07	97.06	97.30	90.32	92.31	100.00	100.00	100.00	100.00
2	61.51	82.02	50.38	45.54	98.86	98.00	97.45	96.20	77.30	97.46	98.95	99.21
3	84.04	78.66	61.95	35.92	97.04	99.54	98.74	100.00	86.25	93.50	99.54	100.00
4	46.43	72.87	53.26	15.31	98.86	99.46	99.39	96.89	95.94	96.81	98.85	98.42
5	88.82	90.16	84.71	3.57	98.47	98.22	92.54	100.00	96.00	98.69	98.70	99.23
6	76.72	82.61	78.08	67.87	100.00	99.83	99.65	99.49	97.27	100.00	99.49	100.00
7	0.00	83.33	68.42	100.00	100.00	100.00	100.00	95.65	100.00	100.00	86.96	100.00
8	83.49	87.16	88.55	83.78	96.46	99.48	80.08	100.00	100.00	98.70	99.74	100.00
9	0.00	100.00	40.00	11.02	76.19	100.00	0.00	100.00	62.50	100.00	100.00	100.00
10	70.89	83.61	69.40	27.07	99.74	99.48	88.93	100.00	84.39	99.87	98.46	99.48
11	58.51	75.16	69.49	60.60	98.77	99.19	97.57	99.90	92.64	99.69	99.74	99.49
12	59.38	66.74	62.13	23.95	98.34	98.13	91.52	97.90	92.46	92.74	91.30	98.34
13	82.23	92.53	86.70	84.38	100.00	99.39	94.58	100.00	90.30	96.91	97.02	100.00
14	87.39	89.78	91.76	75.08	99.90	99.80	100.00	99.90	100.00	99.40	99.90	99.80
15	86.30	72.00	64.127	53.17	94.12	98.72	100.00	99.68	97.60	92.92	95.00	100.00
16	98.36	100.00	100.00	98.44	98.67	97.33	94.37	94.87	100.00	91.14	98.53	98.67
OA (%)	70.21	89.91	70.95	50.88	98.58	99.07	95.56	99.07	90.36	97.86	98.61	99.51
AA (%)	53.06	66.77	62.39	52.65	96.87	96.53	86.25	98.60	82.22	93.47	94.85	99.70
Kappa × 100	65.07	78.01	66.63	44.07	98.39	98.94	94.94	98.94	88.97	97.57	98.41	99.44

The bold font highlights which mechanic works best.

Table 6. Quantitative comparison on the Houston 2013 dataset.

No.	SVM	RF	KNN	GaussianNB	HybridSN	RSSAN	MSRN	MAFN	DCRN	DMCN	MSDAN	HFENet
1	81.98	98.49	97.74	93.97	95.88	98.52	99.80	99.70	98.88	99.78	99.00	99.40
2	98.85	98.40	98.44	98.31	98.57	99.80	100.00	99.11	97.84	94.80	99.90	99.50
3	96.68	99.81	98.37	91.35	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
4	98.53	98.00	99.40	98.81	99.69	99.60	99.50	99.80	95.51	98.12	99.50	99.90
5	88.51	95.07	92.42	71.57	100.00	100.00	100.00	99.80	99.80	99.90	100.00	100.00
6	100.00	99.17	100.00	35.89	100.00	100.00	100.00	100.00	97.65	85.81	100.00	100.00
7	68.66	88.87	89.39	54.52	96.41	97.37	99.70	99.10	94.51	97.92	98.82	99.51
8	84.53	92.23	88.06	79.41	97.64	98.70	100.00	97.85	98.90	96.47	98.03	99.70
9	59.65	80.72	79.22	43.03	96.23	96.52	97.46	99.19	96.44	99.68	99.19	100.00
10	58.38	85.59	84.59	0.00	99.09	97.48	99.90	97.98	100.00	93.25	99.90	99.29
11	59.17	80.72	83.40	35.92	100.00	97.23	100.00	98.40	100.00	98.31	99.20	99.60
12	63.07	76.86	79.79	24.31	99.49	97.47	99.39	96.54	99.39	98.00	99.29	99.90
13	100.00	87.74	93.62	17.54	100.00	99.70	86.78	98.88	100.00	100.00	98.93	99.72
14	78.57	96.50	98.56	68.72	100.00	99.13	100.00	100.00	100.00	100.00	98.84	100.00
15	99.08	99.62	99.43	99.79	99.81	100.00	98.32	97.24	95.83	100.00	100.00	100.00
OA (%)	77.90	90.23	90.50	61.21	98.55	98.53	99.11	98.80	98.19	97.60	99.33	99.73
AA (%)	77.15	89.17	88.87	63.67	98.39	98.38	99.08	98.58	97.95	97.31	99.23	99.62
Kappa × 100	76.07	89.86	89.72	58.1595	98.43	98.41	99.04	98.70	98.04	97.41	99.28	99.70

The bold font highlights which mechanic works best.

By comparing the devised HFENet model with diversiform approaches, we can draw the following conclusions:

- (1) According to Tables 4–6, it is obvious that ML-based methods obtain inferior classification results compared with DL-based methods. For example, for the UP dataset, GaussianNB has the worst OA, AA, and Kappa values, which are 33.37%, 26.32%, and 41.30% lower than those of HybridSN, respectively. For the IP dataset, SVM has the second worst OA, AA, and Kappa values, which are 25.35%, 33.19%, and 29.87% lower than those of MSRN, respectively. This is because ML-based methods only utilize the spectral information and ignore fertile spatial information. Meanwhile, they devilishly rely on hand-crafted features with poor generalization ability and limited representation ability, which damage the classification accuracies. Due to the hierarchical structure and power feature extraction ability, DL-based methods can adaptively capture features and obtain good classification values.
- (2) Table 4 provides the classification results for the UP dataset. This scenario contains many smaller areas of the species and possesses rich spatial information; most of the methods yield good classification results. Tables 5 and 6 provide the classification results for the IP and Houston2013 datasets. In the early stages of growth, the crop areas of the former have been imaged and induce strong mixing phenomena; the latter has highly similar spectral characteristics between categories, which increases the classification difficulty. Nevertheless, our proposed HFENet still achieves impressive results on the three datasets. For example, for the Houston2013 dataset, HFENet obtains 99.73% OA, 99.62% AA, and 99.70% Kappa, which are 2.13%, 2.31%, and

2.29% higher than those of DMCN, respectively. For the UP dataset, HFENet obtains 99.96% OA, 99.94% AA, and 99.95% Kappa, which are 1.96%, 2.81%, and 2.60% higher than those of MSDAN, respectively. For the IP dataset, HFENet obtains 99.51% OA, 99.70% AA, and 99.44% Kappa, which are 9.15%, 17.48%, and 10.47% higher than those of DCRN, respectively. These sufficiently prove the superiority and stability of our proposed HFENet.

- (3) From the point of view of the attention mechanism, RSSAN devises a spectral–spatial attention learning module to refine the learned features. MAFN uses a spatial attention module and a band attention module to relieve the influence of noisy pixels and redundant bands. Our constructed SAEB adaptively recalibrates spectral-wise and spatial-wise feature responses to generate the purified spectral–spatial information. In Tables 4–6, we can clearly see that our presented method obtains superb values for three datasets. For example, for the UP dataset, HFENet obtains 99.96% OA, 99.94% AA, and 99.95% Kappa, which are 0.43%, 0.71%, and 0.36% higher than those of RSSAN, 0.98%, 0.75%, and 1.39% higher than those of MAFN, respectively. For the IP dataset, HFENet obtains 99.51% OA, 99.70% AA, and 99.44% Kappa, which are 0.44%, 3.17%, and 0.5% higher than those of RSSAN, 0.44%, 1.1%, and 0.5% higher than those of MAFN, respectively. This is because SAEB can effectively dispel the interference of redundant bands and noises from local and global views.
- (4) From the point of view of the multiscale strategy, MSRN devises a multiscale residual block with mixed depthwise convolution to achieve multiscale feature learning. MSDAN designs three different scales modules to enhance feature reuse. Our proposed HFEB can exploit spectral–spatial structure information of distinct types and scales, which is composed of two parallel branches: the upper branch contains two HFEBs with 2D convolutional operations with 3×3 size and 5×5 size, respectively; the lower branch contains one HFEB with 2D convolutional operations with 7×7 size. In Tables 4–6, the classification results indicate that HFENet is advantageous in extracting multiscale features. For example, for the UP dataset, HFENet obtains 99.96% OA, 99.94% AA, and 99.95% Kappa, which are 1.02%, 1.73%, and 1.36% higher than those of MSRN, and 1.96%, 2.81%, and 2.6% higher than those of MSDAN, respectively. This is because our proposed HFEB exploits multiple HFRBs to obtain more discriminative and representative spectral–spatial structure information instead of simple concatenation convolutional layers with different sizes. HybridSN, RSSAN, MAFN, DCRN, and DMCN utilize the fixed-scale convolutional kernels to extract spectral–spatial features. Although these methods obtain good classification performance, they lack an exploration of the diversity of spectral–spatial features. Compared with the aforementioned methods, our proposed HFENet uses multiple HFRBs with diverse sizes to augment the multiplicity of spectral–spatial features and model the global long-range dependencies of spectral–spatial features. For example, for the UP dataset, HFENet obtains 99.96% OA, 99.94% AA, and 99.95% Kappa, which are 2.41%, 5.11%, and 3.19% higher than those of DCRN, respectively. For the IP dataset, HFENet obtains 99.51% OA, 99.70% AA, and 99.44% Kappa, which are 1.65%, 6.23%, and 1.87% higher than those of DMCN, respectively.
- (5) Figures 5–7 provide the ground-truth map and the visual classification result map of each comparison method for the three datasets. By comparison, the visual classification result map of our proposed HFENet is closest to the ground truth and the cleanest. Four ML-based methods are inclined to produce salt and pepper noises on the visual classification result maps for three datasets. The visual classification result maps of seven DL-based methods are relatively smooth but may result in the misclassification of pixels at edges. Particularly, our proposed HFENet can effectively avoid the oversmoothing of edges and achieve more precise classification with fine details and more realistic features.

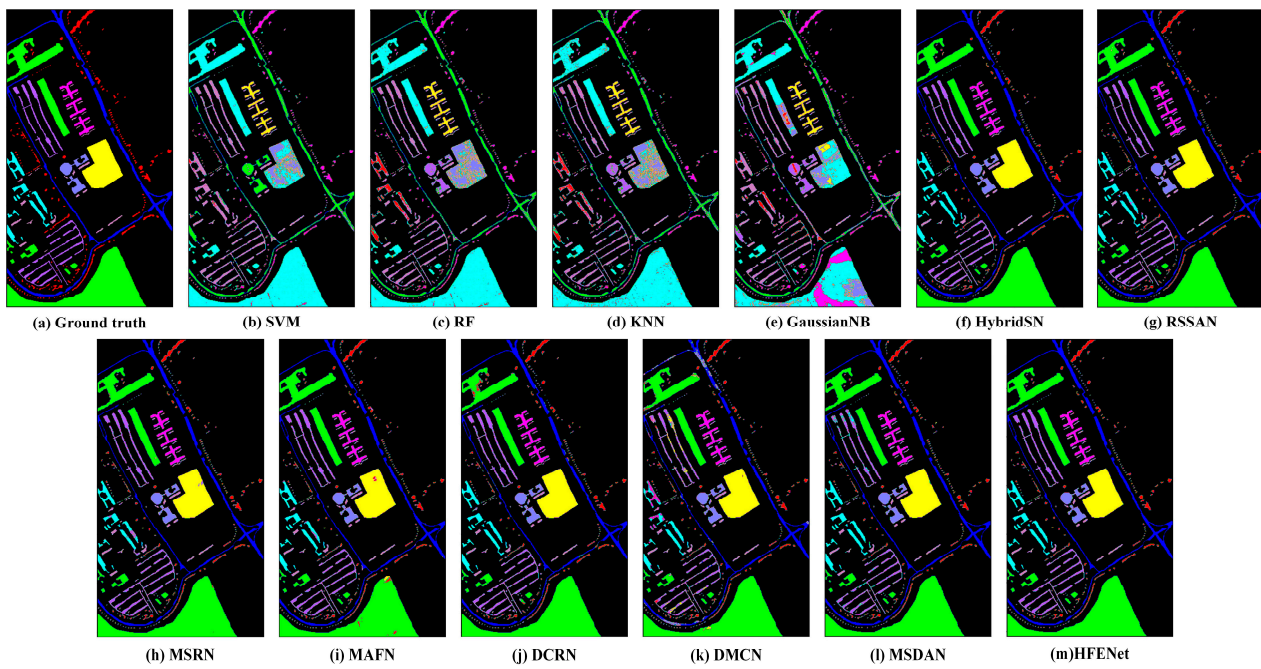


Figure 5. Classification maps on the UP dataset.

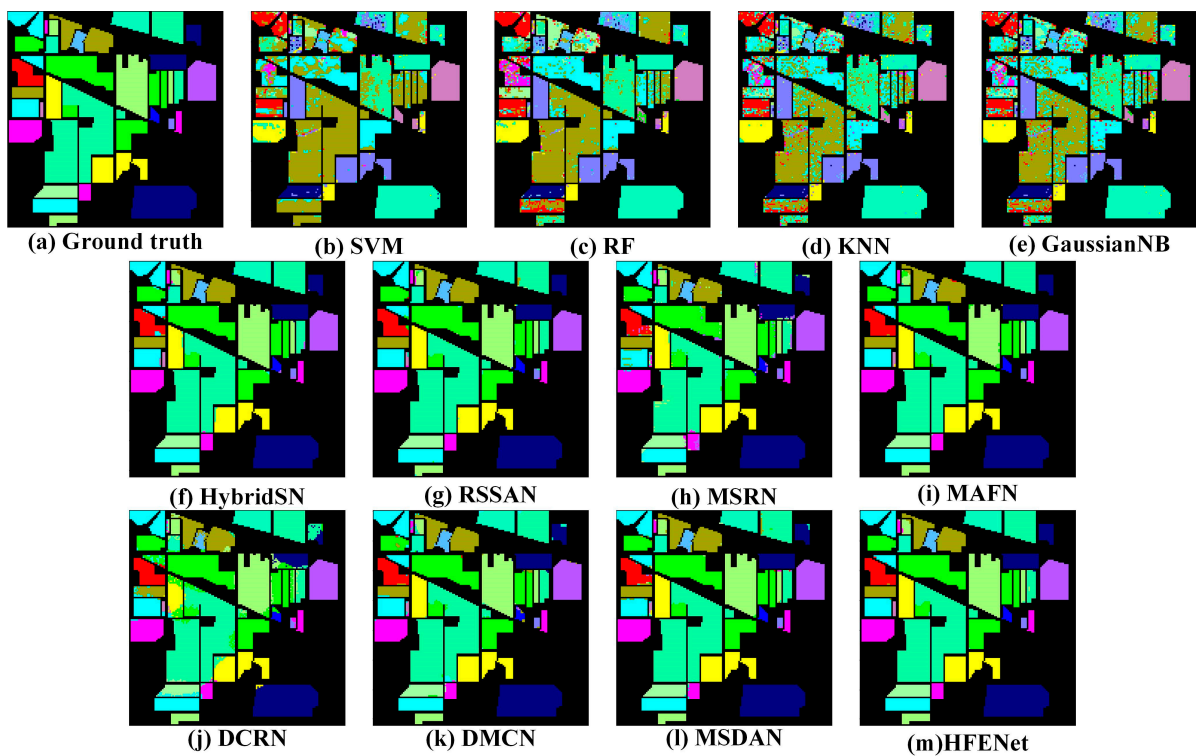


Figure 6. Classification maps on the IP dataset.

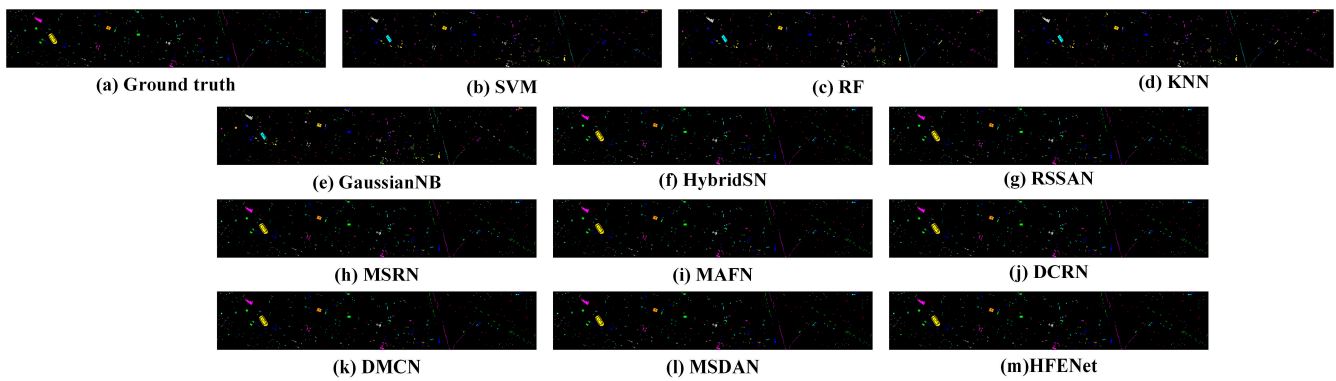


Figure 7. Classification maps on the Houston2013 dataset.

3.3. Parameter Analysis

3.3.1. Varying Proportions of Training Samples

The classification performance of our proposed method is extremely affected by the proportions of training samples. We randomly pick labeled samples in the grid of $\{1\%, 3\%, 5\%, 7\%, 10\%, 20\%, 30\%\}$ as the training size, which aims to analyze the classification performance under the varying proportions of training samples. The classification results of three experimental datasets are provided in Figure 8. In Figure 8, for the three datasets, it can be obviously observed that as the training size increases, the OA, AA, and Kappa values gradually increase for the three datasets. When the training size is 20%, the values of criteria metrics are the most impressive. As the training size outweighs 20%, the values of criteria metrics gradually decline. This is because although a large number of training samples contribute to the training process of HFENet, it may introduce more background information and noisy pixels, which weaken the effect of label pixels and are adverse to the classification performance. In addition, we can also find that the proportions of training samples have a great impact on the IP and Houston2013 datasets. This is because, in the early stages of growth, the crop areas of the IP dataset have been imaged and induced strong mixing phenomena, which increase the classification difficulty. The Houston 2013 dataset has highly similar spectral characteristics between categories, where only a small number of categories are labeled, and most of the samples are unlabeled. The two datasets involve a relatively large number of labeled data for training to fulfill decent classification accuracy. In comparison, the UP dataset contains a mass of labeled samples and utilizes a small number of labeled samples as the training set to achieve good classification accuracies. To make the proposed method generalized and obtain excellent classification results, we set the appropriate proportion of training samples to 20% for three datasets.

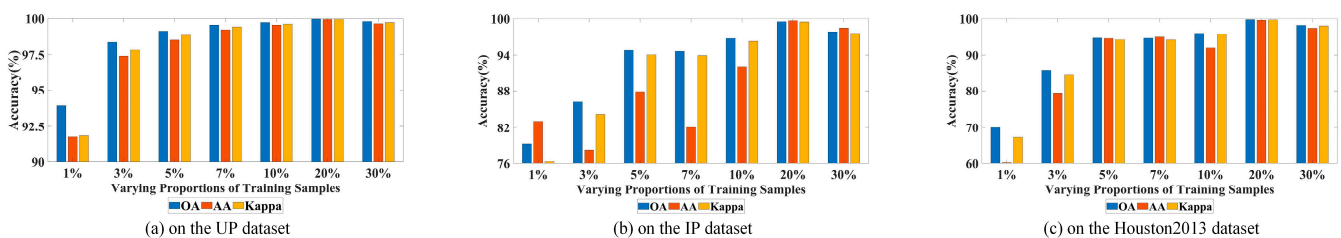


Figure 8. Varying proportions of training samples.

3.3.2. Different Spatial Sizes of Input Image Patches

The too-small spatial size of the input image patch, due to the insufficient receptive field, results in important information loss. The too-large spatial size of the input image patch involves many noisy pixels and is persecuted for the interclass field. Therefore, we set the spatial size of the input image patch in the grid of $\{5 \times 5, 7 \times 7, 9 \times 9, 11 \times 11, 13 \times 13, 15 \times 15\}$

to analyze the classification performance under the different spatial sizes. The classification results of three experimental datasets are shown in Figure 9. In Figure 9, in the UP dataset, it can be readily found that the values of criteria metrics attained are unexceptionable when the spatial size is 15×15 . The OA, AA, and Kappa values are optimal when the spatial size is 7×7 for the IP and Houston2013 datasets. The above results show that when the spatial size is optimal, the input image patch contains less background information and noisy pixels, and the label pixel can play an important role in the classification task. Hence, to achieve splendid classification results, we set the definitive spatial size of the input image patch for three datasets to 15×15 , 7×7 , and 7×7 , respectively.

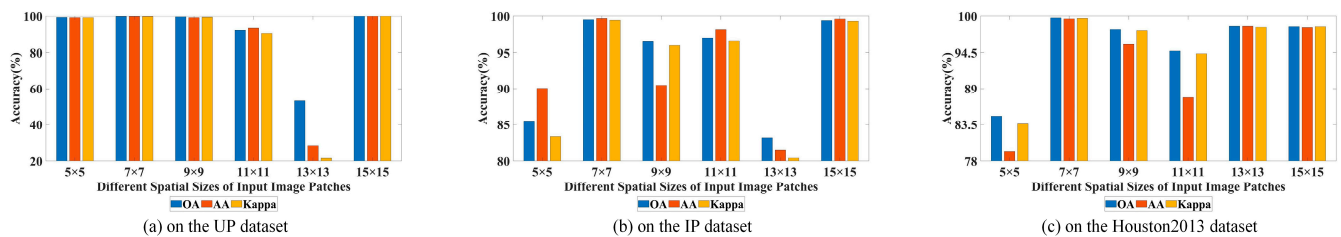


Figure 9. Different spatial sizes of input image patches.

3.3.3. Diverse Numbers of Principal Components

HSI contains abundant spectral information from hundreds to thousands of narrow bands, but these bands have a high correlation with each other and easily trigger the Hughes phenomenon, which prejudices the classification performance. Therefore, before extracting general spectral–spatial information, the PCA was performed on the raw HSI. We set the principal component numbers in the grid of $\{5, 10, 20, 30, 40\}$ to analyze the classification performance under the diverse number of principal components. The classification results of three experimental datasets are shown in Figure 10. In Figure 10, for the UP dataset, the accuracies of OA, AA, and Kappa rise as the number of principal components increases; it can be clearly seen that when the principal component number is 40, our proposed HFENet achieves impressive results. For the IP dataset, except the number of principal components is 10, the accuracies of OA, AA, and Kappa increase monotonically as the number of principal components increases. This is because, compared with other conditions, when the number of principal components is 10, these spectral bands have a high correlation with each other and are adverse to the classification task. For the Houston2013 dataset, the accuracies of OA, AA, and Kappa fluctuate significantly. When the principal component number is 30, our proposed HFENet obtains competitive accuracies. These phenomena indicate that the number of principal components has a great impact on the classification performance for the Houston2013 dataset. Hence, to obtain the best classification performance, we set the most suitable number of principal components for the three datasets to 40, 40, and 30, respectively.

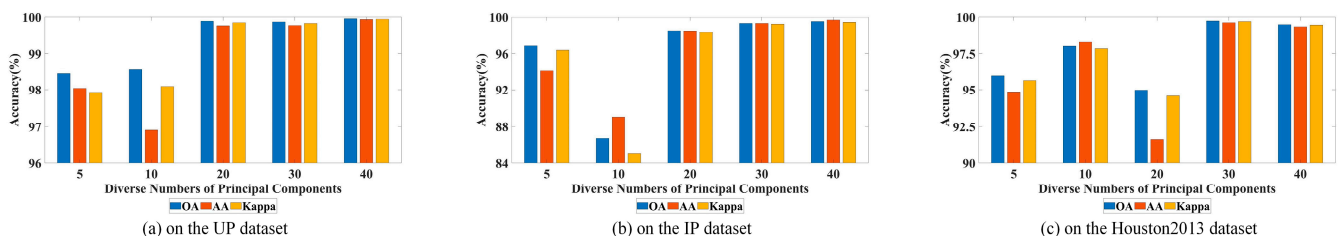


Figure 10. Diverse principal component numbers: 3, 3, and 4. Varying numbers of HFEBs.

HFEB is built to capture the spectral–spatial structure information of distinct types, scales, and branches. When the number of HFEBs is too small, obtained spectral–spatial information is inadequate; when the number of HFEBs is too large, the number of parameters and model complexity exacerbate. Both of these situations do not contribute

to the classification task. Hence, setting a pertinent number of HFEBs is crucial for our developed method. The number of HFEBs is set in the grid of $\{2, 3, 4, 5, 6\}$ to analyze the classification performance under the varying numbers of HFEBs. The classification results of three experimental datasets are shown in Figure 11. In Figure 11, for the UP and IP datasets, it can be obviously observed that when the number of HFEBs is 2, our proposed HFENet can sufficiently exploit spectral–spatial structure information of distinct types and scales and obtains excellent classification performance. For the Houston2013 dataset, when the number of HFEBs is 4, the values of criteria metrics are perfect. Hence, we set the most appropriate HFEb numbers for three datasets to 2, 2, and 4, respectively.

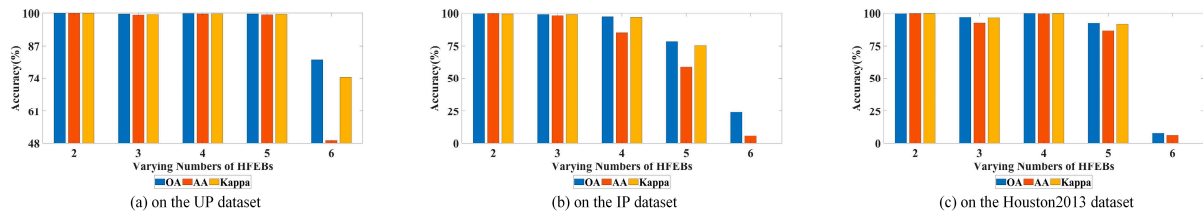


Figure 11. Varying numbers of HFEBs.

3.3.4. Different Numbers of Groups for SAEB

SAEB can adaptively recalibrate spatial-wise and spectral-wise responses to generate the purified spectral–spatial information. The number of groups has a significant effect on the classification results of the proposed method. When the number of groups is too small, redundant information and interference pixels are inadequately filtered; when the number of groups is too large, the number of parameters and model complexity exacerbate. Both of these situations are not conducive to the classification task. Therefore, the number of groups is set in the grid of $\{2, 4, 8, 16, 32\}$ to analyze the classification performance under the varying numbers of groups. The classification accuracy of three experimental datasets is provided in Figure 12. In Figure 12, it can be readily found that for three datasets, our proposed HFENet achieves the outstanding values of criteria metrics, as the number of groups for SAEB is 8. Hence, we set the definitive numbers of groups for SAEB for three datasets to 8, 8, and 8, respectively.

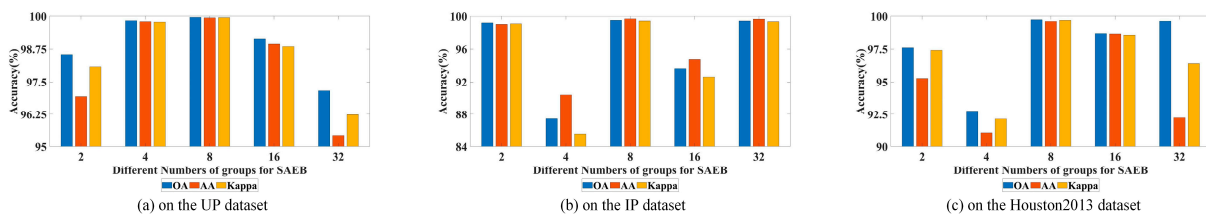


Figure 12. Different numbers of groups for SAEB: 3, 3, and 6. Diverse channel ratios of spectral enhancement branch.

The spectral enhancement branch can emphasize the meaningful bands and fade out the irrelevant ones, which can model the interdependencies between features and enhance the expressive ability of the model. The channel ratio r decides the number of neurons in the first fully connected layer, which is utilized to reduce computation. We set the channel ratios in the grid of $\{1, 2, 4, 8, 16, 32\}$ to analyze the classification performance under the diverse channel ratios. The classification results of three experimental datasets are shown in Figure 13. In Figure 13, it can be clearly seen that, for the UP dataset, when the r is 4, the classification performance is the worst; when the r is 2, the accuracies of AA, OA, and Kappa are best. For the IP dataset, when the r is 1 or 8, the classification ability does not manifest well; when the r is 2, the classification accuracies are excellent. For the Houston2013 dataset, when the r is 2 or 16, the classification perform well; the classification accuracies of other conditions are inferior. In addition, we can also find that the accuracies of OA, AA, and

Kappa do not increase monotonically as r increases but fluctuate significantly for the IP and Houston2013 datasets. A possible reason for this is that the spectral enhancement branch overfits the feature spectral-wise correlations. Compared with the above two datasets, the accuracies of OA, AA, and Kappa slightly degrade as r increases. This is because that spectral enhancement branch underfits the feature spectral-wise correlations. Therefore, to obtain the outstanding classification results, the most suitable channel ratios are 2, 2, and 16 for three datasets, respectively.

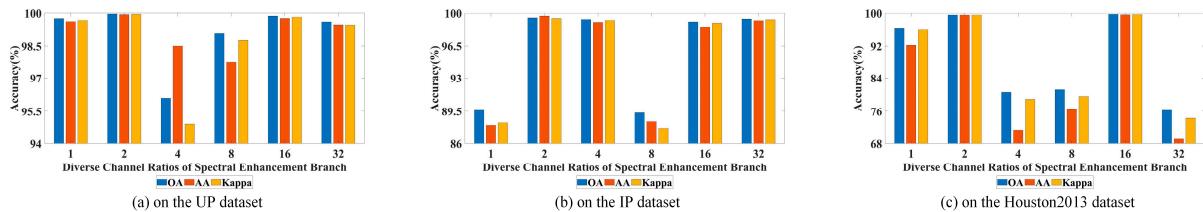


Figure 13. Diverse channel ratios of spectral enhancement branch.

3.3.5. Varying L2 Regularization Parameters

L2 regularization effectively avoids the overfitting problem and was applied to our proposed method. We set the spatial size of the input image patch in the grid of $\{0, 0.0005, 0.002, 0.01, 0.02, 0.03, 0.1, 1\}$ to analyze the classification performance under the varying L2 regularization parameters. The classification results of the three experimental datasets are shown in Figure 14. In Figure 14, it can be obviously observed that the most proper L2 regularization parameters are 0.002, 0.002, and 0.03 for the three datasets, respectively.

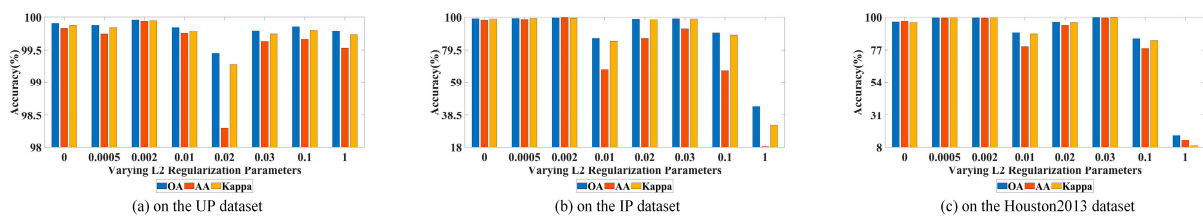


Figure 14. Varying L2 regularization parameters.

3.4. Ablation Study

3.4.1. Efficiency Analysis of HFRB

HFRB was constructed to strengthen internal and external interactions of different channels and layers while enriching the local dependencies of spectral–spatial features. HFRB is composed of SRU and CRU. The former is utilized to boost the internal relations of different channels; the latter is designed to enhance the robustness of spectral–spatial features by learning the external correlations of all the channels. To sufficiently verify the efficiency of HFRB, comparative experiments were performed under three conditions, i.e., case 1 (only using SRU), case 2 (only using CRU), and case 3 (namely, our presented method, using SRU and CRU). Figure 15 provides the classification results of the three experimental datasets.

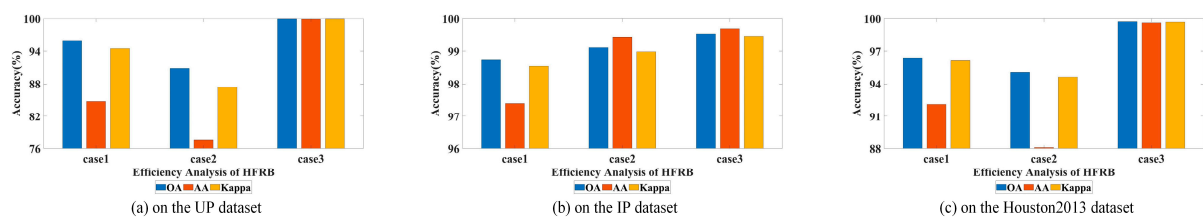


Figure 15. Efficiency analysis of HFRB.

According to Figure 15, it can be obviously observed that for the three datasets, case 3 obtains competitive values of the criteria metric. Regarding the UP dataset, case 2 achieves the worst values of the criteria metric, which are 9.19%, 22.26%, and 12.57% lower than those of case 3. For the Houston2013 dataset, the classification performance manifests a similar behavior with respect to the results obtained for the UP dataset. Regarding the IP dataset, case 1 has terrible values in terms of the criteria metric, which are 0.8%, 2.3%, and 0.91% lower than those of case 3. These results sufficiently demonstrate that SRU and CRU complement each other. Only when the SRU and CRU are utilized together can they fully strengthen internal and external interactions of different channels and layers and give play to the effect of $1 + 1 > 2$.

3.4.2. Efficiency Analysis of HFEB

HFEB utilizes multiple promising functional HFRBs to capture spectral–spatial structure information of distinct types and scales, where each HFRB exploits 2D convolutional operations with different sizes. To sufficiently verify the efficiency of HFEB, comparative experiments are performed under three conditions, i.e., case 1 (only using HFEB with 2D convolutional operations with 3×3 size), case 2 (using HFEBs with 2D convolutional operations with 3×3 and 5×5 sizes), and case 3 (namely, our presented method, using HFEBs with 2D convolutional operations with 3×3 , 5×5 , and 7×7 sizes). Figure 16 provides the classification results of the three experimental datasets.

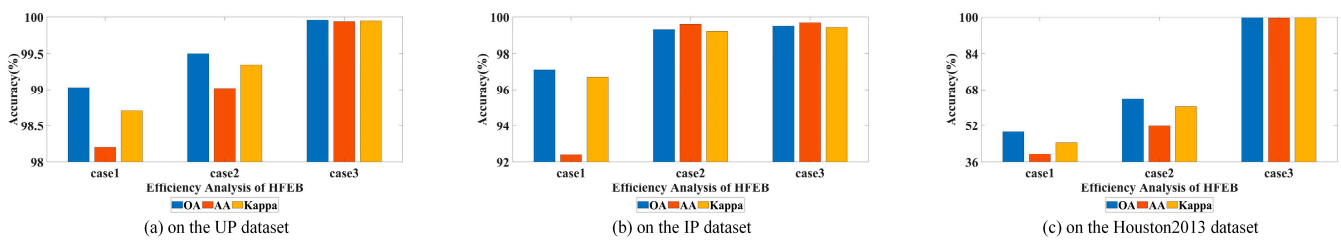


Figure 16. Efficiency analysis of HFEB.

As shown in Figure 16, it can be readily found that values of the criteria metric of case 1 are the lowest for three datasets. Regarding the UP dataset, case 1 obtains 99.03% OA, 98.21% AA, and 98.71% Kappa, which are 0.96%, 1.73%, and 1.24% lower than those of case 2, respectively. Regarding the IP dataset, case 1 obtains 97.10% OA, 92.40% AA, and 96.69% Kappa, which are, respectively, 2.41%, 7.3%, and 2.75% lower than those of case 2. Regarding the Houston2013 dataset, case 1 achieves 49.21% OA, 39.69% AA, and 44.63% Kappa, which are, respectively, 50.52%, 59.93%, and 55.07% lower than those of case 2. These numerical values effectively expound that utilizing HFEBs with 2D convolutional operations with 5×5 size is important. By comparison, the values of the criteria metric of case 3 are clearly better than the other two conditions. For example, for the houston2013 dataset, case 3 has 99.73% OA, 99.62% AA, and 99.70% Kappa, which are 36.03%, 47.93%, and 39.26% higher than those of case 2, respectively. These confirm that our constructed HFEB is successful and plays an important role.

3.4.3. Efficiency Analysis of HFENet Model

To analyze and demonstrate the impact of each component, comparative experiments were performed under three conditions, i.e., network 1 (only using HFEB), network 2 (only using SAEB), and network 3 (namely, our presented method, using HFEB and SAEB). Figure 17 provides the classification results of the three experimental datasets.

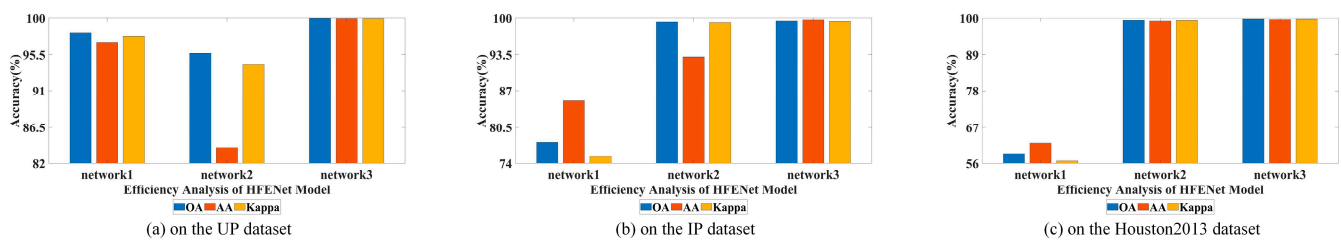


Figure 17. Efficiency analysis of HFENet model.

As shown in Figure 17, for the UP dataset, it can be clearly seen that the classification performance of network 2 is the worst. For the IP and UP datasets, the classification performance manifests a similar behavior; the values of the criteria metric of network 1 are the most terrible. This is because, compared with network 2, network 1 needs more parameters for the training process. The UP dataset has a relatively large number of labeled samples; although network 1 suffers from noisy pixels and redundant bands, it can extract spectral–spatial information from different types and scales and obtain good classification. The other two datasets contain a relatively small number of labeled samples; compared with network 1, the model architecture of network 2 is relatively not complicated and obtains good classification. Network 1 may lead to the overfitting problem. Among the three conditions, network 3 makes a big impression and obtains excellent classification results. For example, for the UP dataset, network 3 obtains 99.96% OA, 99.94% AA, and 99.95% Kappa, which are 1.73%, 3%, and 2.29% higher than those of network 1, respectively. For the IP and Houston2013 datasets, the obtained results exhibit very similar behavior with respect to the results obtained for the UP dataset. These results sufficiently prove that our designed SAEB is valid and effectively fades out the redundant information and noisy pixels, further generating the purified spectral–spatial information. Moreover, compared with network 2, network 3 achieves values of the criteria metric 4.29%, 15.97%, and 5.71% higher than those of network 2 for the UP dataset. The obtained results for the other two datasets have very similar behavior with respect to the results obtained for the UP dataset. These results abundantly verify that our devised HFEB is effective and can extract more discriminative and representative spectral–spatial structure information of distinct types, scales, and branches while modeling the global long-range dependencies of spectral–spatial features. To sum up, the constructed HFEB and SAEB of the proposed methods considerably contribute to the classification performance up to a point.

4. Conclusions

To remedy gradient disappearance and fully exploit spectral–spatial information, this article presents an innovative hybrid-scale feature enhancement network (HFENet) for HSI classification. Different from the classification methods focusing on utilizing the fixed-scale convolutional kernels or multiple available, receptive fields with varying scales, HFENet uses a hybrid-scale feature extraction block (HFEB) to model the global long-range spectral–spatial dependencies of different scales, types, and branches, enriching the diversity of informative features. In addition, to generate the purified spectral–spatial information, HFENet adopts a shuffle attention enhancement block (SAEB) to adaptively recalibrate spectral-wise and spatial-wise responses, effectively filtering redundant information and noisy pixels and is conducive to enhancing the classification performance. From an experimental point of view, our proposed HFENet possesses effectiveness and superiority and exhibits state-of-the-art performance compared with several advanced methods. In the future, we will be devoted to utilizing a neural search strategy to adaptively design the model architecture and apply the unsupervised, semi-supervised, or training mechanisms to our proposed method. Meanwhile, we will try to apply the proposed classification method to some computer vision tasks, such as target recognition, medical diagnosis, and urban planning.

Author Contributions: Conceptualization, D.L.; investigation, D.L. and J.Z.; formal analysis, D.L.; validation, J.Z.; original draft preparation, D.L.; funding acquisition, M.L. and J.Z.; review and editing, D.L., T.S., G.Q., M.L. and J.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China under grant number 62101529.

Data Availability Statement: The data presented in this study are available in this article.

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

References

1. Guo, Y.; Chanussot, J.; Jia, X.; Benediktsson, J.A. Multiple Kernel learning for hyperspectral image classification: A review. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 6547–6565. [[CrossRef](#)]
2. Wang, D.; Zhang, J.; Du, B.; Zhang, L.; Tao, D. DCN-T: Dual Context Network With Transformer for Hyperspectral Image Classification. *IEEE Trans. Image Process.* **2023**, *32*, 2536–2551. [[CrossRef](#)] [[PubMed](#)]
3. Zhang, X.; Sun, Y.; Shang, K.; Zhang, L.; Wang, S. Crop classification based on feature band set construction and object-oriented approach using hyperspectral images. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2016**, *9*, 4117–4128. [[CrossRef](#)]
4. Yang, X.; Yu, Y. Estimating soil salinity under various moisture conditions: An experimental study. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 2525–2533. [[CrossRef](#)]
5. Kruse, F.A.; Boardman, J.W.; Huntington, J.F. Comparison of airborne hyperspectral data and EO-1 hyperion for mineral mapping. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1388–1400. [[CrossRef](#)]
6. Lu, B.; He, Y.; Dao, P.D. Comparing the performance of multispectral and hyperspectral images for estimating vegetation properties. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2019**, *12*, 1784–1797. [[CrossRef](#)]
7. Wang, J.; Hu, J.; Liu, Y.; Hua, Z.; Hao, S.; Yao, Y. EL-NAS: Efficient Lightweight Attention Cross-Domain Architecture Search for Hyperspectral Image Classification. *Remote Sens.* **2023**, *15*, 4688. [[CrossRef](#)]
8. Yuan, D.; Yu, A.; Qian, Y.; Xu, Y.; Liu, Y. S2Former: Parallel Spectral-Spatial Transformer for Hyperspectral Image Classification. *Remote Sens.* **2023**, *12*, 3937. [[CrossRef](#)]
9. Liu, B.; Jia, Z.; Guo, P.; Kong, W. Hyperspectral Image Classification Based on Transposed Convolutional Neural Network Transformer. *Remote Sens.* **2023**, *12*, 3979. [[CrossRef](#)]
10. Fu, L.; Chen, X.; Pirasteh, S.; Xu, Y. The Classification of Hyperspectral Images: A Double-Branch Multi-Scale Residual Network. *Remote Sens.* **2023**, *15*, 4471. [[CrossRef](#)]
11. Borsoi, R.; Imbiriba, T.; Berumdez, J.C.; Richard, C.; Jutten, C. Spectral variability in hyperspectral data unmixing: A comprehensive review. *IEEE Geosci. Remote Sens. Mag.* **2021**, *9*, 223–270. [[CrossRef](#)]
12. Haut, J.M.; Paoletti, M.E.; Plaza, J.; Plaza, A.; Li, J. Visual attention driven hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 8065–8080. [[CrossRef](#)]
13. Keshava, N. Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 1552–1565. [[CrossRef](#)]
14. Bruzzone, L.; Roil, F.; Serpico, S.B. An extension of the Jeffreys-Matusita distance to multiclass cases for feature selection. *IEEE Trans. Geosci. Remote Sens.* **1995**, *33*, 1318–1321. [[CrossRef](#)]
15. Kailath, T. The divergence and bhattacharyya distance measures in signal selection. *IEEE Trans. Commun.* **1967**, *15*, 52–60. [[CrossRef](#)]
16. Kang, X.; Xiang, X.; Li, S.; Benediktsson, J.A. PCA-based edge-preserving features for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 7140–7151. [[CrossRef](#)]
17. Wang, J.; Chang, C.-I. Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1586–1600. [[CrossRef](#)]
18. Nielsen, A.A. Kernel maximum autocorrelation factor and minimum noise fraction transformations. *IEEE Trans. Image Process.* **2011**, *20*, 6612–6624. [[CrossRef](#)]
19. Melgani, F.; Bruzzone, L. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 1778–1790. [[CrossRef](#)]
20. Wang, Q.; Lin, J.; Yuan, Y. Salient band selection for hyperspectral image classification via manifold ranking. *IEEE Trans. Neural Netw. Learn. Syst.* **2016**, *27*, 1279–1289. [[CrossRef](#)]
21. Belgiu, M.; Dragut, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [[CrossRef](#)]
22. Camps-Valls, G.; Bruzzone, L. Kernel-based methods for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 1351–1362. [[CrossRef](#)]
23. Benediktsson, J.A.; Palmason, J.A.; Sveinsson, J.R. Classification of hyperspectral data from urban areas based on extended morphological profiles. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 480–491. [[CrossRef](#)]

24. Chen, Y.; Nasrabadi, N.M.; Tran, T.D. Hyperspectral image classification using dictionary-based sparse representation. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 3973–3985. [[CrossRef](#)]
25. Chen, Y.; Lin, Z.; Zhao, X.; Wang, G.; Gu, Y. Deep learning-based classification of hyperspectral data. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2014**, *7*, 2094–2107. [[CrossRef](#)]
26. Li, T.; Zhang, J.; Zhang, Y. Classification of hyperspectral image based on deep belief networks. In Proceedings of the 2014 IEEE International Conference on Image Processing (ICIP), Paris, France, 27–30 October 2014; pp. 5132–5136.
27. Hong, D.; Guo, L.; Yao, J.; Zhang, B.; Plaza, A.; Chanussot, J. Graph Convolutional Networks for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 5966–5978. [[CrossRef](#)]
28. Hang, R.; Zhou, F.; Liu, Q.; Ghamisi, P. Classification of hyperspectral images via multitask generative adversarial networks. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 1424–1436. [[CrossRef](#)]
29. Hang, R.; Liu, Q.; Hong, D.; Ghamisi, P. Cascaded recurrent neural networks for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 5384–5394. [[CrossRef](#)]
30. Li, X.; Ding, M.; Pižurica, A. Deep Feature Fusion via Two-Stream Convolutional Neural Network for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 2615–2629. [[CrossRef](#)]
31. Zu, B.; Li, Y.; Li, J.; He, Z.; Wang, H.; Wu, P. Cascaded Convolution-Based Transformer With Densely Connected Mechanism for Spectral–Spatial Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 2615–2629. [[CrossRef](#)]
32. Zhang, C.; Li, G.; Lei, R.; Du, S.; Zhang, X.; Zheng, H.; Wu, Z. Deep Feature Aggregation Network for Hyperspectral Remote Sensing Image Classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 5314–5325. [[CrossRef](#)]
33. Xue, A.; Yu, X.; Liu, B.; Tan, X.; Wei, X. HResNetAM: Hierarchical Residual Network with Attention Mechanism for Hyper-spectral Image Classification. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2021**, *14*, 3566–3580. [[CrossRef](#)]
34. Xu, Q.; Xiao, Y.; Wang, D.; Luo, B. CSA-MSO3DCNN: Multiscale Octave 3D CNN with Channel and Spatial Attention for Hyperspectral Image Classification. *Remote Sens.* **2020**, *12*, 188. [[CrossRef](#)]
35. Gao, H.; Chen, Z.; Li, C. Sandwich Convolutional Neural Network for Hyperspectral Image Classification Using Spectral Feature Enhancement. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 3006–3015. [[CrossRef](#)]
36. Mei, S.; Ji, J.; Hou, J.; Li, X.; Du, Q. Learning Sensor-Specific Spatial-Spectral Features of Hyperspectral Images via Convolutional Neural Networks. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 4520–4533. [[CrossRef](#)]
37. Hu, W.; Huang, Y.; Wei, L.; Zhang, F.; Li, H. Deep Convolutional Neural Networks for Hyperspectral Image Classification. *J. Sens.* **2015**, *2015*, 258619. [[CrossRef](#)]
38. Xu, Y.; Du, B.; Zhang, L. Beyond the Patchwise Classification: Spectral-Spatial Fully Convolutional Networks for Hyperspectral Image Classification. *IEEE Trans. Big Data* **2020**, *6*, 492–506. [[CrossRef](#)]
39. Zou, L.; Zhu, X.; Wu, C.; Liu, Y.; Qu, L. Spectral–Spatial Exploration for Hyperspectral Image Classification via the Fusion of Fully Convolutional Networks. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *3*, 659–674. [[CrossRef](#)]
40. Zhang, M.; Li, W.; Du, Q. Diverse Region-Based CNN for Hyperspectral Image Classification. *IEEE Trans. Image Process.* **2018**, *27*, 2623–2634. [[CrossRef](#)]
41. Xu, F.; Zhang, G.; Song, C.; Wang, H.; Mei, S. Multiscale and Cross-Level Attention Learning for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2023**, *61*, 5501615. [[CrossRef](#)]
42. Zhang, H.; Gong, C.; Bai, Y.; Bai, Z.; Li, Y. 3-D-ANAS: 3-D Asymmetric Neural Architecture Search for Fast Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–19. [[CrossRef](#)]
43. Safari, K.; Prasad, S.; Labate, D. A Multiscale Deep Learning Approach for High-Resolution Hyperspectral Image Classification. *IEEE Geosci. Remote Sens. Lett.* **2021**, *18*, 167–171. [[CrossRef](#)]
44. Zhang, X.; Wang, Y.; Zhang, N.; Xu, D.; Luo, H.; Chen, B.; Ben, G. Spectral–Spatial Fractal Residual Convolutional Neural Network With Data Balance Augmentation for Hyperspectral Classification. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 10473–10487. [[CrossRef](#)]
45. Wu, P.; Cui, Z.; Gan, Z.; Liu, F. Residual Group Channel and Space Attention Network for Hyperspectral Image Classification. *Remote Sens.* **2020**, *12*, 2035. [[CrossRef](#)]
46. Guo, H.; Liu, J.; Yang, J.; Xiao, Z.; Wu, Z. Deep Collaborative Attention Network for Hyperspectral Image Classification by Combining 2-D CNN and 3-D CNN. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 4789–4802. [[CrossRef](#)]
47. Liu, D.; Li, Q.; Li, M.; Zhang, J. A Decompressed Spectral-Spatial Multiscale Semantic Feature Network for Hyperspectral Image Classification. *Remote Sens.* **2023**, *15*, 4642. [[CrossRef](#)]
48. Chen, Y.; Jiang, C.; Li, C.; Jia, X.; Ghamisi, P. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6232–6251. [[CrossRef](#)]
49. Cao, X.; Ren, M.; Zhao, J.; Li, H.; Jiao, L. Hyperspectral Imagery Classification Based on Compressed Convolutional Neural Network. *IEEE Geosci. Remote Sens. Lett.* **2020**, *17*, 1583–1587. [[CrossRef](#)]
50. Xie, J.; He, N.; Fang, L.; Ghamisi, P. Multiscale Densely-Connected Fusion Networks for Hyperspectral Images Classification. *IEEE Trans. Circuits Syst. Video Technol.* **2021**, *31*, 246–259. [[CrossRef](#)]
51. Wang, J.; Huang, R.; Guo, S.; Li, L.; Zhu, M.; Yang, S.; Jiao, L. NAS-Guided Lightweight Multiscale Attention Fusion Network for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 8754–8767. [[CrossRef](#)]
52. Zhu, M.; Jiao, L.; Liu, F.; Yang, S.; Wang, J. Residual Spectral–Spatial Attention Network for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 449–462. [[CrossRef](#)]

53. Roy, S.K.; Manna, S.; Song, T.; Bruzzone, L. Attention-Based Adaptive Spectral–Spatial Kernel ResNet for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 7831–7843. [[CrossRef](#)]
54. Zhang, X.; Sahng, S.; Tang, X.; Feng, J.; Jiao, L. Spectral Partitioning Residual Network with Spatial Attention Mechanism for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–14. [[CrossRef](#)]
55. Dong, Z.; Cai, Y.; Cai, Z.; Liu, X.; Yang, Z.; Zhuge, M. Cooperative Spectral–Spatial Attention Dense Network for Hyperspectral Image Classification. *IEEE Geosci. Remote Sens. Lett.* **2021**, *18*, 866–870. [[CrossRef](#)]
56. Gao, H.; Miao, Y.; Cao, X.; Li, C. Densely Connected Multiscale Attention Network for Hyperspectral Image Classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 2563–2576. [[CrossRef](#)]
57. Wang, X.; Fan, Y. Multiscale Densely Connected Attention Network for Hyperspectral Image Classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2022**, *15*, 1617–1628. [[CrossRef](#)]
58. Zhang, Y.; Li, K.; Li, K.; Wang, L.; Zhong, B.; Fu, Y. Image super-resolution using very deep residual channel attention networks. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 286–301.
59. Zhang, C.; Li, G.; Du, S. Multi-Scale Dense Networks for Hyperspectral Remote Sensing Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 9201–9222. [[CrossRef](#)]
60. Guo, W.; Ye, H.; Cao, F. Feature-Grouped Network with Spectral–Spatial Connected Attention for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5500413. [[CrossRef](#)]
61. Roy, S.K.; Krishna, G.; Dubey, S.R.; Chaudhuri, B.B. HybridSN: Exploring 3-D–2-D CNN Feature Hierarchy for Hyperspectral Image Classification. *IEEE Geosci. Remote Sens. Lett.* **2020**, *17*, 277–281. [[CrossRef](#)]
62. Gao, H.; Yang, Y.; Li, C.; Gao, L.; Zhang, B. Multiscale Residual Network with Mixed Depthwise Convolution for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 3396–3408. [[CrossRef](#)]
63. Li, Z.; Zhao, X.; Xu, Y.; Li, W.; Zhai, L.; Fang, Z.; Shi, X. Hyperspectral Image Classification with Multiattention Fusion Network. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 5503305. [[CrossRef](#)]
64. Xu, Y.; Li, Z.; Li, W.; Du, Q.; Liu, C.; Fang, Z.; Zhai, L. Dual-Channel Residual Network for Hyperspectral Image Classification with Noisy Labels. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5507916. [[CrossRef](#)]
65. Xiang, J.; Wei, C.; Wang, M.; Teng, L. End-to-End Multilevel Hybrid Attention Framework for Hyperspectral Image Classification. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 5511305. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.