



## Article

# An Unsupervised Feature Extraction Using Endmember Extraction and Clustering Algorithms for Dimension Reduction of Hyperspectral Images

Sayed Hamed Alizadeh Moghaddam <sup>1</sup>, Saeed Gazor <sup>1</sup>, Fahime Karami <sup>2</sup>, Meisam Amani <sup>3,4,\*</sup> and Shuanggen Jin <sup>3,5</sup>

- <sup>1</sup> Department of Electrical and Computer Engineering, Queen's University, Kingston, ON K7L 3N9, Canada; h.alizadeh@queensu.ca (S.H.A.M.); gazor@queensu.ca (S.G.)
- <sup>2</sup> Department of Geography and Planning, Queen's University, Kingston, ON K7L 3N6, Canada; fahime.karami@queensu.ca
- <sup>3</sup> School of Surveying and Land Information Engineering, Henan Polytechnic University, Jiaozuo 454000, China; sgjin@shao.ac.cn
- <sup>4</sup> WSP Environment and Infrastructure Canada Limited, Ottawa, ON K2E 7L5, Canada
- <sup>5</sup> Shanghai Astronomical Observatory, Chinese Academy of Sciences, Shanghai 200030, China
- \* Correspondence: meisam.amani@wsp.com

**Abstract:** Hyperspectral images (HSIs) provide rich spectral information, facilitating many applications, including landcover classification. However, due to the high dimensionality of HSIs, landcover mapping applications usually suffer from the curse of dimensionality, which degrades the efficiency of supervised classifiers due to insufficient training samples. Feature extraction (FE) is a popular dimension reduction strategy for this issue. This paper proposes an unsupervised FE algorithm that involves extracting endmembers and clustering spectral bands. The proposed method first extracts existing endmembers from the HSI data via a vertex component analysis method. Using these endmembers, it subsequently constructs a prototype space (PS) in which each spectral band is represented by a point. Similar/correlated bands in the PS remain near one another, forming several clusters. Therefore, our method, in the next step, clusters spectral bands into multiple clusters via K-means and fuzzy C-means algorithms. Finally, it combines all the spectral bands in the same cluster using a weighted average operator to decrease the high dimensionality. The extracted features were evaluated by applying an SVM classifier. The experimental results confirmed the superior performance of the proposed method compared with five state-of-the-art dimension reduction algorithms. It outperformed these algorithms in terms of classification accuracy on three widely used hyperspectral images (Indian Pines, KSC, and Pavia Centre). The suggested technique also showed comparable or even stronger performance (up to 9% improvement) compared with its supervised competitor. Notably, the proposed method exhibited higher accuracy even when only a limited number of training samples were available for supervised classification. Using only five training samples per class for the KSC and Pavia Centre datasets, our method's classification accuracy was higher than that of its best-performing unsupervised competitors by about 7% and 1%, respectively, in our experiments.

**Keywords:** hyperspectral remote sensing; dimension reduction; feature extraction; curse of dimensionality; supervised classifier; landcover classification



**Citation:** Alizadeh Moghaddam, S.H.; Gazor, S.; Karami, F.; Amani, M.; Jin, S. An Unsupervised Feature Extraction Using Endmember Extraction and Clustering Algorithms for Dimension Reduction of Hyperspectral Images. *Remote Sens.* **2023**, *15*, 3855. <https://doi.org/10.3390/rs15153855>

Academic Editor: Danfeng Hong

Received: 15 June 2023

Revised: 31 July 2023

Accepted: 1 August 2023

Published: 3 August 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 sensors, which capture images using hundreds of spectral bands, provide a rich source of spectral information and pave the way for a wide range of remote sensing applications, particularly precise landcover classification [1,2]. However, the inclusion of many spectral bands poses challenges in the transmission, storage, and processing

of hyperspectral images (HSIs) [3]. Because supplying adequate training samples could be difficult in real-world applications, HSI classification may suffer from the curse of dimensionality, which has a detrimental influence on final results by increasing the likelihood of overfitting [4]. More significantly, the high dimensionality increases computational time and complexity. As a solution, dimension reduction (DR) approaches are used as a critical preprocessing step, particularly for supervised landcover classification, to decrease the high dimensionality of HSIs while preserving discriminative information [5,6].

DR approaches are broadly divided into feature extraction (FE) and feature selection (FS). FS approaches select a desired band subset from the original HSI while removing redundant bands. The sequential information of the original data is kept in FS models, which is critical for physical interpretation. These FS approaches are further classified as supervised [7,8] and unsupervised [9,10].

The FE approaches combine spectral bands using mathematical processes, such as principal component analysis (PCA), and map the original HSI into a new feature space with a lower dimension [11]. FE methods have some advantages over FS ones, which contribute to their stronger performance [12]. FE techniques can extract the most information from the original high-dimensional spectral data by generating new features that indicate the most significant variations. Furthermore, FE approaches can generate new features that are not clearly present in the original spectral bands. These newly generated features are capable of capturing complex relationships and latent patterns in the data. As a result, they could be able to increase the representation of spectral characteristics, resulting in improved discrimination and classification accuracy [13,14]. Similar to FS approaches, FE is categorized into supervised and unsupervised models. Unsupervised algorithms do not require training samples and are thus preferred, because collecting and preparing ground truth training samples is time-consuming, expensive, and sometimes impossible [15].

According to [14], there are five types of FE methods: knowledge-based, statistical, wavelet-based, clustering, and deep learning. Knowledge-based strategies improve certain spectral band characteristics by performing basic mathematical operations on the original bands to differentiate the objects of interest. This category includes spectral indices, such as the normalized difference index (NDVI). Using hyperspectral data, researchers in [16] examined different vegetation indices, such as NDVI and the soil adjusted vegetation index (SAVI), and then modified these indices to forecast the green leaf area index. Although knowledge-based features have a direct relationship with physical characteristics, the goal of this type of FE is not to reduce dimensionality; instead, these approaches are utilized to gather extra information to facilitate the differentiation of land cover classes.

In statistical FE, the original HSI is transformed into a lower dimensional feature space without considerable loss of information, in which the classes of interest are more separable [17]. Nonparametric weighted feature extraction (NWFE) is an example of statistical FE [18]. Another statistical FE approach is discriminant analysis feature extraction (DAFE) [19]. Although this supervised technique is frequently used for dimension reduction, it can only extract a limited number of features (i.e.,  $M - 1$ , where  $M$  is the total number of classes), which may not be enough for an accurate image classification. Another disadvantage of DAFE is that it suffers greatly from the ill-conditioning problem when just a small number of training samples are available.

The wavelet-based FE methods are based on the fact that wavelet transform decomposes a signal into constituent wavelets of varying sizes and positions. The majority of methods in this category employ the discrete wavelet transform.

Deep learning FE performs admirably when dealing with nonlinear input data. Deep learning approaches, which are well reviewed in [13], use a hierarchical learning framework to extract high-level features [14]. The use of autoencoders is a basic method for extracting deep features in a hierarchical manner. Deep features were extracted from hyperspectral images in a work by Chen et al. (2014) using a stacked autoencoder with five layers [20]. The images were then classified using logistic regression.

Minibatch graph convolutional networks (miniGCN), a novel method for training large-scale graph convolutional networks in a minibatch fashion, was introduced by researchers in [21]. The feature extraction and classification of hyperspectral images were the primary purposes for the creation of this technique. An advantageous feature of miniGCN is its ability to handle out-of-sample data inference without the need for retraining the networks, leading to improved classification performance. In addition, SpectralFormer, a novel backbone network, was suggested in [22] for the deep feature extraction and classification of HSIs. Leveraging image transformer models, SpectralFormer is adept at learning spectrally local sequence information from adjacent bands within HSI images, thus generating groupwise spectral embeddings. Additionally, researchers in [23] addressed the need for finely classifying features in complex scenes by introducing a comprehensive multimodal deep learning framework, which serves as a baseline solution.

Deep learning algorithms for HSI dimension reduction have the benefit of learning hierarchical representations and capturing complex spectral-spatial features automatically. However, for hyperspectral datasets, supervised deep learning models require a substantial quantity of labeled training data, which might be difficult to obtain. When the training dataset is small, they are prone to overfitting, which can lead to poor generalization performance. Furthermore, developing deep learning models for dimension reduction can be computationally challenging, necessitating significant computer resources and longer training durations.

Clustering-based algorithms are the final type of FE that has shown promising results [24]. In this category, related spectral bands are first grouped into several clusters, and then the grouped bands are merged to extract a single feature from each cluster. This category includes unsupervised band correlation clustering (BCC) [25]. The BCC method involves determining correlation coefficients between spectral bands, clustering the bands using a k-means algorithm based on the correlation coefficient matrix, and extracting new features by computing the mean value of the bands in each cluster. The results of BCC experiments demonstrated that BCC outperformed well-known FE methods including PCA, independent component analysis, and minimum noise fraction. While band correlation is useful for understanding spectral band relationships, relying entirely on band correlation in BCC for dimension reduction may result in limited information capture and poor feature discrimination. By overlooking specific spectral characteristics and variations within classes, this method may result in less discriminative and less accurate feature representations, especially in complex hyperspectral datasets.

In [26], an iterative technique for clustering and merging neighboring bands based on the Pearson correlation coefficient was developed. Although the suggested technique is entirely unsupervised and has consistent behavior for different data sets, it requires time to fine-tune its parameters. Using an expectation-maximization clustering methodology and weighted average fusion, researchers at [27] developed a feature extraction method for hyperspectral image analysis. One of the most significant advantages of this technique is that it automatically identifies the optimal number of clusters based on the Bhattacharya distance. The extracted features are also highly discriminative, which resulted in superior classification results compared with its competitors.

Lu et al. (2014) described a feature extraction approach for spectral-spatial classification of hyperspectral data that is based on clustering [28]. The suggested method grouped the high-dimensional data via clustering, efficiently extracting significant information while decreasing redundancy. The technique is able to improve the extraction of hidden relationships between neighboring pixels that are not easily visible in the raw data. In [3], adjacent bands were optimally clustered, using a particle swarm optimization (PSO) algorithm, and merged using a mean-weighted operator. Although the method is capable of identifying optimal clusters, it is developed based on a metaheuristic optimization algorithm that is computationally expensive and time-consuming.

Another FE method employing a clustering-based algorithm is clustering-based feature extraction (CBFE) [29]. CBFE constructs a prototype space (PS) in a supervised manner

where each band is represented as a point/vector. CBFE searches for similar/correlated bands in the PS using a clustering algorithm and combines them to minimize the dimensionality of the HSIs. The fact that CBFE only uses first-order statistics from training samples (i.e., mean values) is one of its advantages. As a result, it could perform better than other methods which employ second-order statistics in ill-posed situations, especially when there are few training samples available for training. CBFE outperformed various conventional methods, including PCA, NWF, and linear discriminant analysis, according to the experimental results [29].

CBFE can show desirable results, but it is a supervised method and requires training samples. However, as mentioned before, acquiring accurate and sufficient training samples is expensive, time-consuming, difficult, and sometimes impossible. To address this issue, in this paper, we provide an unsupervised version of CBFE by introducing the concept of endmember extraction into the dimension reduction discipline. In the first step of our method, pure pixels, called endmembers, of the input HSI are extracted. Using these endmembers, we construct a PS in an unsupervised manner in which each axis represents an endmember. Second, the proposed method groups similar and correlated spectral bands in the PS. Finally, all bands in each cluster are combined using a weighted average operator to decrease the HSI's high dimensionality. More details about the proposed method are provided in the following section. The primary contribution of our research can be summarized as follows:

- Introducing an unsupervised breakthrough: By removing the requirement for training data, our method revolutionizes the CBFE into a genuinely unsupervised approach.
- Endmember extraction meets dimension reduction: By combining the strength of endmember extraction with dimension reduction, we have developed a unique method that uses pure pixels to build a new PS. Within this PS, similar/correlated spectral bands remain near one another and form several clusters, which lets us identify and reduce redundant spectral bands.
- Unleashing the potential of unsupervised learning: We use unsupervised learning principles to automatically group and combine similar and correlated spectral bands, opening the way for rapid and accurate dimension reduction.

## 2. Methodology

The proposed method for dimension reduction in hyperspectral imagery has five steps. The flowchart of the proposed method is illustrated in Figure 1, and details of each step are explained below.

**Step 1:** Estimate the input HSI's virtual dimensionality (VD) using HySime, a well-known VD estimation technique [30]. This step detects the number of existent endmembers in the input image, denoted by  $n$ . Further details and additional information about HySime can be found in the publication available at [31].

**Step 2:** Extract existing endmembers ( $E$ ) from the image using a state-of-the-art vertex component analysis (VCA) [32]. We utilized an open-source MATLAB Hyperspectral Toolbox for the practical application of VCA and HySime. The toolbox can be accessed and downloaded from [33]. The VCA algorithm can be outlined as follows:

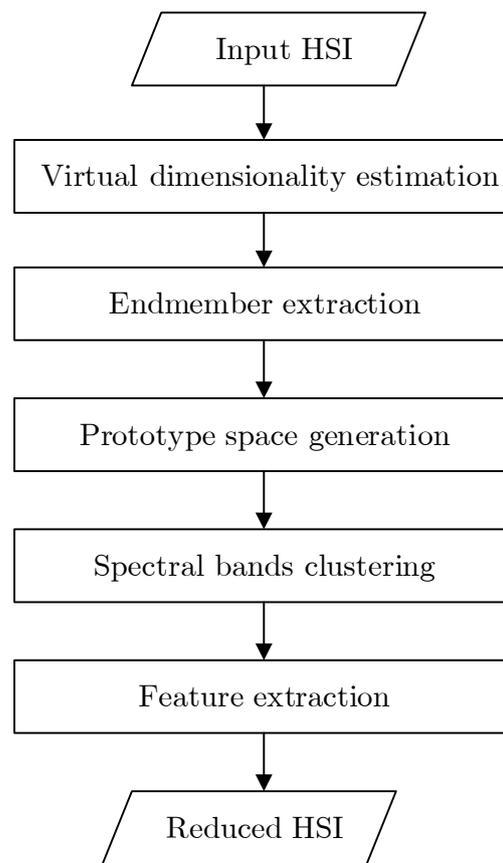
**Step 2.1.** Consider the desired number of endmembers to be generated as  $n$  and the HSI data to be  $H_{p \times d}$ , where  $p$  is the total number of pixels and  $d$  is the total number of spectral bands.

**Step 2.2.** Set the initial vector  $e_1 = [0, 0, \dots, 1]$ , and consider  $E^{(1)} = [e_1^T, \mathbf{0}, \dots, \mathbf{0}]$

**Step 2.3.** Produce a Gaussian random vector,  $w^{(k+1)}$ , which is utilized to generate  $f^{(k+1)}$ :

$$f^{(k+1)} = ((I - E^{(k)}E^{(k)\#})w^{(k+1)}) / \left\| (I - E^{(k)}E^{(k)\#})w^{(k+1)} \right\| \quad (1)$$

where  $E^{(k)\#}$  is pseudoinverse of  $E^{(k)}$ , and  $f^{(k+1)}$  is a vector orthonormal to the subspace spanned by  $E^{(k)}$ .



**Figure 1.** The flowchart of the proposed method for dimension reduction of HSIs.

**Step 2.4.** Determine  $e_{k+1}$ , which is a pixel (i.e., row) in  $H$  having the maximum length after projection using  $f^{(k+1)}$ . In other words,  $e_{k+1}$  is a pixel in  $H$  that maximizes the following expression:

$$e_{k+1} = \underset{i}{\operatorname{argmax}}(\|f^{(k+1)T}[H]_{i,:}\|) \quad (2)$$

where the symbol  $[H]_{i,:}$  means the  $i$ -th row of matrix  $H$ ;  $(\cdot)^T$  denotes the transpose operator, and  $\|\cdot\|$  is the norm operator. Argmax represents the operation of finding the argument that maximizes the value of a given function.

**Step 2.5.** Replace the  $k + 1$ -th column of  $E^{(k+1)}$  with  $e_{k+1}^T$ . Repeat the last three steps to extract  $n$  endmembers in  $E = [e_1^T, e_2^T, \dots, e_n^T]$ . These endmembers are pure pixels of the image and are considered as representatives of existing phenomena.

**Step 3:** Generate a prototype space (PS) using the spectra of the extracted endmembers (i.e.,  $E$ ) in the second step. In this space, each axis is associated with an endmember. Therefore, the dimension of the PS is equal to the total number of endmembers (i.e., VD). Matrix  $E$  generates the PS in our method. Each row of  $E$  is a point in the PS, and it is representative of a spectral band. For example, the  $j$ -th row of  $E$  represents the  $j$ -th spectral band in the PS. As mentioned, each spectral band is represented by a point in PS. As a result, similar bands in PS remain near one another, forming several clusters.

**Step 4:** Group the spectral bands' representatives in the PS into  $k$  clusters by applying a clustering algorithm. Each of these clusters contains similar and correlated spectral bands.  $k$  equals the number of extracted features (i.e., the dimension of the reduced HSI) in our method.

**Step 5:** Extract a new feature for each cluster to decrease the dimensionality of the input HSI. In order to achieve this, we present two FE versions in this study. The fourth step of the first method, known as weighted feature extraction (WFE), uses a  $k$ -means clustering algorithm [34]. Then, we selected the spectral bands that had a representative in

a similar cluster. Finally, a weighted average of the selected spectral bands was calculated as described below (Equation (3)).

$$B'_l = \frac{\sum_{j=1}^h w_j B_j}{\sum_{j=1}^h w_j} \quad (3)$$

where  $B'_l$  is the extracted feature for the  $l$ -th cluster;  $B_j$  is the  $j$ -th spectral band whose representative exists in cluster  $l$ , and  $w_j$  is its corresponding weight, which is the inverse of distance to the  $l$ -th cluster's centroid; and  $h$  denotes the total number of the spectral band's representatives in the  $l$ -th cluster.

In the second method, called fuzzy feature extraction (FFE), we first apply the renowned fuzzy  $C$ -means clustering [35] in the fourth step, and then use Equation (4) to reduce the high dimension of input HSI:

$$R_{p \times k} = H_{p \times d} \times U_{d \times k}^T \quad (4)$$

in which  $H$  is the input HSI matrix with a dimension of  $k \times d$ , where  $p$  is the total number of pixels, and  $d$  is the total number of spectral bands;  $R$  is the dimensionality deduced HSI;  $U = [u_{i,j}]$  is the fuzzy partition matrix, with a dimension of  $k \times d$ , where  $k$  is the total number of clusters; and  $u_{i,j}$  denotes the degree of membership of the  $j$ -th spectral band's representative to the  $i$ -th cluster. It should be mentioned that the columns of  $U^T$  are normalized by dividing each of its columns by their sum.

### 3. Experimental Results and Discussions

#### 3.1. Datasets

Three hyperspectral datasets were used in this study to evaluate the effectiveness of the proposed method as an unsupervised FE. The datasets are explained in detail in the following subsections.

##### 3.1.1. Pavia Centre

The first dataset is a hyperspectral image of Pavia Centre, captured by the ROSIS sensor (Figure 2a). Originally, this high-resolution image (1.3 m per pixel) had 115 spectral bands that covered a spectral range of 0.43 to 0.86  $\mu\text{m}$ . After removing all noise-affected bands, the remaining 102 bands were used in the experiments. According to the existing ground truth data, the nine classes of interest are water, tree, meadow, brick, soil, asphalt, bitumen, tile, and shadow (Figure 2b and Table 1).

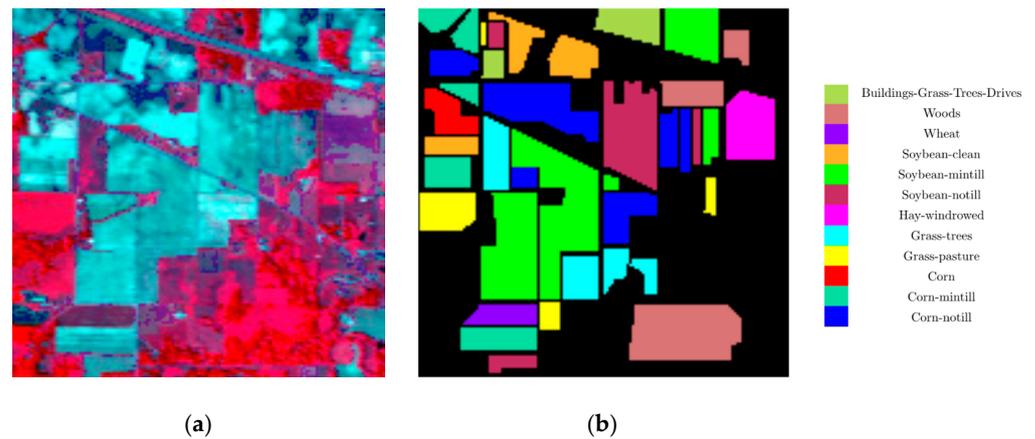
**Table 1.** Pavia Centre ground truth samples.

Classes	Number of Samples
Water	65,971
Trees	7598
Asphalt	3090
Self-Blocking Bricks	2685
Bitumen	6584
Tiles	9248
Shadows	7287
Meadows	42,826
Bare soil	2863

##### 3.1.2. Indian Pines

The AVIRIS sensor obtained the second data set over a vegetated area in Indiana, USA. The AVIRIS images contain 220 spectral bands (0.4–2.45  $\mu\text{m}$ ). Due to the 20 noisy water absorption bands (i.e., 104–108, 150–163, and 220) in the original image, the experiments were conducted using 200 spectral bands. This dataset has 16 classes, according to the published ground truth map. In this study, four classes were ignored due to their small sample size.

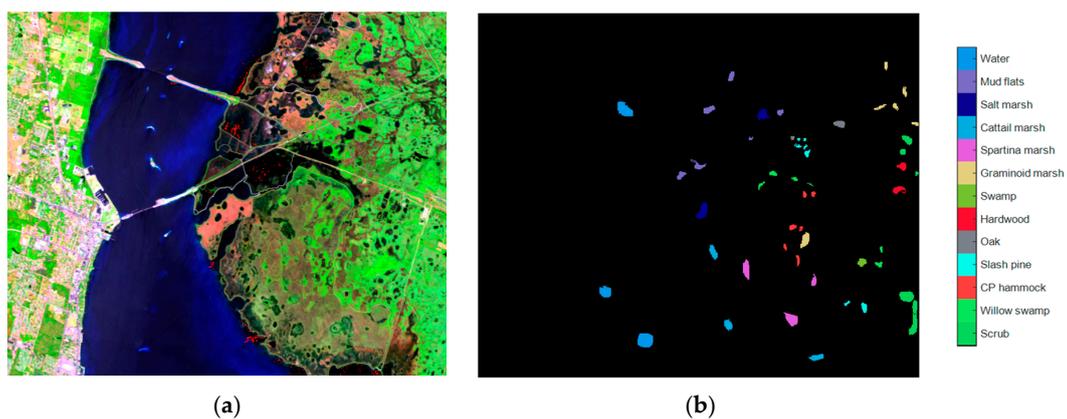




**Figure 3.** Indian Pine case study. (a) A false-color composite image. (b) Ground truth map.

**Table 3.** KSC ground truth samples.

Classes	Number of Samples
Scrub	347
Willow swamp	243
CP hammock	256
Slash pine	252
Oak	161
Hardwood	229
Swamp	105
Graminoid marsh	390
Spartina marsh	520
Cattail marsh	404
Salt marsh	419
Mud flats	503
Water	927



**Figure 4.** Kennedy Space Center case study. (a) A false-color composite image. (b) Ground truth map.

### 3.2. Experiment Design

To test and compare the quality of extracted features using FFE, WFE, and competing methods, we employed support vector machines (SVMs) [36] with an RBF kernel. The SVM was created with the LibSVM package [37]. We also employed a grid search method and fivefold cross-validation to fine-tune SVM's hyperparameters (i.e.,  $C$  and  $\gamma$ , which are "Cost" and RBF kernel parameters, respectively). For the KSC and Pavia Centre datasets, we randomly selected 5 and 20 training samples per class from available ground truth data, and 50 and 100 samples for the Indian Pines dataset. It should be noted that testing was conducted using the remaining ground truth samples.

To gain deeper insights into the capabilities of the methods in classifying the extracted or selected attributes, we employed the T-Distribution Stochastic Neighbour Embedding (T-SNE) [38] algorithm, effectively reducing their multidimensional nature to a mere pair. T-SNE stands as a nonlinear method of reducing dimensionality, tailor-made for effectively visualizing data with high dimensions. The primary strength of T-SNE lies in its ability to preserve local structure. In essence, this ensures that points with comparable distances in the original high-dimensional data space retain their similarity even after projection into a lower-dimensional data space. In this experiment, we utilized 18 features obtained from each method as input for T-SNE.

### 3.3. Model Comparison

Our proposed method was compared to five competing methods: hyperbolic clustering-based band hierarchy (HCBH) [24], clustering-based feature extraction (CBFE) [29], band correlation clustering (BCC) [25], prototype space-based feature selection (PFS) [39], and maximum tangent discrimination (MTD) [39]. The reasoning behind selecting these methods for comparison is that they are all based on clustering algorithms for reducing the dimensionality of HSIs.

A dimension reduction technique based on deep learning models was also added. In this regard, an autoencoder (AE) [13] was trained using four encoder layers (with 128, 64, 32, and X neurons) and three decoder layers (with 32, 64, and 128 neurons) with leaky ReLU activation functions. X represents the desired number of features for dimension reduction using AE. The AE was trained to reconstruct the input data by compressing it into a lower-dimensional representation and then expanding it back to its original dimension. The encoder layers gradually reduce the dimensionality of the input data, while the decoder layers aim to reconstruct the original input from the compressed representation. All the competing methods are unsupervised except for CBFE, which benefits from a training sample for dimension reduction. BCC and CBFE are FE algorithms, similar to our proposed methodology, whereas the remaining methods are FS algorithms.

Although HCBH is a FS method, we compared our method with it because, similar to our method, it was developed based on clustering algorithms. HCBH began by calculating the hyperbolic distance between spectral bands in a hyperbolic space. Then, a bottom-up hierarchical clustering was performed to cluster comparable bands. Finally, HCBH chose a band as a representative from each cluster using a cluster-center ranking algorithm.

PFS and MTD were chosen because they both employed PS in their techniques. They, however, generate the PS in a different way than we did. PFS and MTD first employ the k-means algorithm to cluster the input HSI, which has a computational burden. The mean value of each cluster is then determined to create the PS. Thus, each axis in their PS represents a cluster, which is in contrast to our method, where each axis represents an endmember in our PS. Finally, in their PS, PFS uses the orthogonal distance from the PS's diagonal to select informative bands, while MTD applies the tangent of the angles between pairs of bands to choose the bands with the lowest correlations.

In this study, the quality of the extracted features from different methods was evaluated using the kappa coefficient ( $\kappa$ ), overall accuracy (OA) and average accuracy (AA), all of which measure classification accuracy from different viewpoints. It should be mentioned that each experiment was repeated ten times, and the average values of  $\kappa$ , OA and AA were finally used for validation of the models in the following tables and figures.

### 3.4. Results and Discussion

#### 3.4.1. Indian Pines

Table 4 shows the obtained result for the Indian Pines dataset using 50 training samples per class. The best result for each row of this table is bolded. According to Table 4, our proposed methods (i.e., WFE and FFE) outperformed the other models in all cases (i.e., with different number of features). This demonstrates the high quality of the extracted features, using our method.

**Table 4.** Classification accuracy, i.e., mean  $\kappa$ /OA/AA, of different dimension reduction methods using 50 training samples per class for the Indian Pines dataset. The best results are highlighted in bold. Values of standard deviation (STD) are given in parentheses. Hyperbolic clustering-based band hierarchy (HCBH), clustering-based feature extraction (CBFE), band correlation clustering (BCC), autoencoder (AE), prototype space-based feature selection (PFS), maximum tangent discrimination (MTD), weighted feature extraction (WFE), and fuzzy feature extraction (FFE).

No. Features	Methods							
	BCC	AE	CBFE	MTD	PFS	HCBH	WFE	FFE
3	49.47(±0.00)/	49.95(±0.75)/	55.87(±0.00)/	39.90(±0.52)/	44.04(±1.83)/	54.04(±0.33)/	<b>57.31(±1.95)/</b>	56.75(±0.66)/
	55.77(±0.00)/	55.43(±0.74)/	60.81(±0.00)/	46.55(±0.52)/	50.59(±1.56)/	59.52(±0.53)/	<b>62.28(±1.72)/</b>	61.67(±0.53)/
	60.77(±0.00)	62.08(±1.01)	67.47(±0.00)	51.49(±0.67)	53.22(±2.11)	66.67(±0.03)	67.83(±2.18)	<b>68.75(±0.78)</b>
4	47.20(±0.18)/	50.86(±1.92)/	56.45(±0.00)/	48.16(±1.82)/	43.03(±0.21)/	61.19(±0.38)/	<b>67.14(±0.81)/</b>	62.62(±3.30)/
	53.25(±0.23)/	56.31(±1.68)/	61.32(±0.08)/	54.16(±1.69)/	49.63(0.05±)/	65.91(±0.58)/	<b>71.08(±0.74)/</b>	66.95(±3.05)/
	60.49(±0.12)	62.24(±1.84)	68.72(±0.00)	60.12(±1.00)	52.85(±0.45)	72.43(±0.38)	<b>75.81(±0.46)</b>	73.27(±1.70)
5	49.89(±0.63)/	52.53(±2.07)/	59.19(±0.10)/	52.04(±1.77)/	38.24(±0.00)/	62.06(±0.88)/	<b>67.70(±0.25)/</b>	66.86(±0.77)/
	55.85(±0.48)/	57.75(±1.87)/	63.91(±0.25)/	57.70(±1.61)/	45.27(±0.00)/	66.73(±1.06)/	<b>71.52(±0.21)/</b>	70.77(±0.68)/
	62.45(±0.43)	63.84(±1.57)	70.60(±0.16)	63.29(±1.13)	48.67(±0.00)	72.74(±0.68)	<b>77.75(±0.40)</b>	76.61(±0.57)
6	52.69(±0.96)/	54.15(±1.28)/	59.96(±0.26)/	56.28(±0.00)/	44.63(±0.00)/	63.72(±0.60)/	<b>69.54(±1.16)/</b>	67.84(±1.39)/
	58.59(±0.66)/	59.23(±1.14)/	64.66(±0.06)/	61.52(±0.00)/	52.01(0.00±)/	68.02(±0.33)/	<b>73.14(±1.02)/</b>	71.63(±1.23)/
	64.16(±1.14)	64.94(±0.96)	71.70(±0.28)	67.67(±0.00)	52.77(±0.00)	74.81(±0.10)	<b>78.73(±0.69)</b>	77.63(±0.98)
7	60.13(±1.00)/	55.84(±1.32)/	63.50(±0.07)/	52.62(±2.55)/	47.16(±3.89)/	65.89(±1.09)/	<b>70.60(±1.52)/</b>	69.45(±1.58)/
	64.82(±0.80)/	60.76(±1.18)/	67.69(±1.09)/	58.00(±2.30)/	53.29(±3.52)/	69.97(±1.03)/	<b>74.10(±1.33)/</b>	73.07(±1.38)/
	70.82(±1.15)	66.26(±1.56)	74.23(±0.06)	64.52(±2.30)	59.03(±3.38)	76.99(±0.05)	<b>79.66(±1.31)</b>	78.39(±1.02)
8	60.26(±1.47)/	56.10(±1.98)/	65.04(±1.15)/	56.44(±0.46)/	47.38(±0.00)/	68.51(±0.66)/	70.73(±1.57)/	<b>70.83(±1.12)/</b>
	64.88(±1.32)/	60.94(±1.80)/	69.07(±1.18)/	61.47(±0.48)/	53.41(±0.00)/	71.95(±0.74)/	74.22(±1.40)/	<b>74.29(±0.98)/</b>
	71.17(±1.21)	67.38(±1.44)	75.44(±0.61)	67.94(±0.13)	59.71(±0.00)	78.55(±0.42)	<b>79.64(±1.43)</b>	79.63(±1.01)
9	60.26(±2.27)/	55.63(±2.31)/	64.76(±1.29)/	56.14(±0.66)/	52.40(±2.42)/	69.14(±0.37)/	70.34(±2.87)/	<b>71.54(±1.47)/</b>
	64.94(±2.01)/	60.57(±2.18)/	68.86(±1.78)/	61.14(±0.62)/	57.95(±2.11)/	72.79(±0.67)/	73.88(±2.54)/	<b>74.93(±1.30)/</b>
	71.09(±1.75)	66.99(±1.36)	75.25(±0.67)	67.80(±0.41)	64.10(±2.87)	78.76(±0.02)	79.24(±2.57)	<b>80.37(±0.84)</b>
10	59.74(±1.34)/	56.70(±2.19)/	65.48(±1.99)/	55.94(±1.52)/	52.06(±1.40)/	69.82(±0.86)/	69.93(±3.50)/	<b>72.11(±1.07)/</b>
	64.47(±1.16)/	61.55(±1.99)/	69.47(±1.56)/	60.95(±1.44)/	57.76(±1.32)/	73.36(±0.42)/	73.52(±3.11)/	<b>75.45(±0.94)/</b>
	70.67(±1.25)	67.19(±2.03)	75.73(±1.43)	67.78(±1.02)	62.44(±0.99)	79.61(±0.21)	78.67(±2.88)	<b>80.74(±0.78)</b>
11	58.33(±1.99)/	57.18(±2.55)/	66.60(±1.75)/	59.91(±0.81)/	53.98(±0.94)/	70.18(±1.15)/	70.00(±3.18)/	<b>73.57(±0.44)/</b>
	63.22(±1.83)/	61.99(±2.32)/	70.52(±1.86)/	64.57(±0.73)/	59.34(±0.81)/	73.71(±0.60)/	73.57(±2.81)/	<b>76.74(±0.41)/</b>
	69.61(±1.48)	67.34(±2.10)	76.46(±1.27)	70.71(±0.47)	65.26(±1.78)	79.42(±0.37)	79.14(±2.88)	<b>81.69(±0.15)</b>
12	61.67(±1.68)/	56.49(±2.10)/	67.21(±2.09)/	59.70(±0.40)/	54.19(±0.68)/	71.34(±0.71)/	70.17(±2.89)/	<b>73.41(±1.04)/</b>
	66.27(±1.47)/	61.32(±1.99)/	71.04(±1.74)/	64.38(±0.37)/	59.49(±0.61)/	74.76(±1.14)/	73.73(±2.56)/	<b>76.61(±0.93)/</b>
	72.25(±1.80)	67.66(±1.34)	77.40(±1.52)	70.64(±0.29)	65.72(±1.60)	80.85(±0.43)	79.52(±2.25)	<b>81.36(±0.60)</b>
13	61.54(±2.16)/	58.32(±0.95)/	66.64(±1.97)/	60.27(±0.80)/	55.01(±1.34)/	71.16(±0.36)/	70.75(±2.54)/	<b>72.89(±0.91)/</b>
	66.11(±1.98)/	63.02(±0.93)/	70.59(±1.09)/	64.87(±0.72)/	60.29(±1.26)/	74.56(±0.66)/	74.25(±2.24)/	<b>76.15(±0.79)/</b>
	72.43(±1.49)	68.52(±0.22)	76.36(±1.83)	71.16(±0.73)	66.21(±1.37)	80.33(±0.46)	79.97(±2.18)	<b>81.34(±0.70)</b>
14	61.34(±1.52)/	56.99(±1.88)/	67.22(±1.23)/	60.51(±0.25)/	55.67(±1.48)/	72.23(±1.02)/	68.60(±2.95)/	<b>72.72(±0.90)/</b>
	65.97(±1.39)/	61.78(±1.71)/	71.11(±2.34)/	65.13(±0.23)/	60.88(±1.35)/	75.50(±0.43)/	72.34(±2.60)/	<b>75.99(±0.79)/</b>
	72.04(±1.31)	67.35(±1.58)	76.80(±1.06)	71.15(±0.43)	66.94(±1.93)	<b>81.31(±0.15)</b>	78.15(±2.68)	81.25(±0.67)
15	62.41(±1.40)/	59.31(±1.37)/	66.89(±2.34)/	59.42(±0.81)/	55.99(±1.87)/	72.85(±0.80)/	68.44(±3.25)/	<b>73.11(±0.97)/</b>
	66.92(±1.22)/	63.90(±1.26)/	70.84(±2.08)/	64.18(±0.78)/	61.20(±1.63)/	76.11(±0.35)/	72.16(±2.91)/	<b>76.35(±0.85)/</b>
	73.05(±1.42)	69.56(±0.76)	76.43(±1.97)	69.86(±0.25)	67.08(±2.48)	81.43(±0.13)	78.36(±2.63)	<b>81.58(±0.73)</b>
16	62.69(±1.46)/	57.99(±0.84)/	66.74(±1.19)/	60.46(±1.89)/	56.49(±0.74)/	70.33(±0.40)/	67.72(±1.96)/	<b>73.26(±0.70)/</b>
	67.14(±1.30)/	62.75(±0.71)/	70.69(±1.05)/	65.11(±1.68)/	61.60(±0.68)/	73.91(±0.56)/	71.51(±1.77)/	<b>76.46(±0.63)/</b>
	73.33(±0.89)	68.13(±1.20)	76.42(±1.18)	70.54(±1.55)	68.10(±0.85)	80.13(±0.28)	77.94(±1.53)	<b>81.97(±0.59)</b>
17	63.38(±1.80)/	58.34(±2.13)/	66.64(±0.74)/	60.84(±2.56)/	57.20(±1.73)/	73.14(±1.00)/	67.89(±2.59)/	<b>73.64(±1.06)/</b>
	67.79(±1.60)/	63.02(±1.94)/	70.58(±0.66)/	65.46(±2.26)/	62.24(±1.54)/	76.39(±0.46)/	71.66(±2.30)/	<b>76.80(±0.96)/</b>
	73.59(±1.45)	68.91(±1.39)	76.59(±0.72)	70.97(±2.14)	68.44(±1.88)	81.48(±0.86)	77.94(±2.27)	<b>82.19(±0.72)</b>
18	62.19(±1.44)/	58.95(±1.69)/	66.35(±1.15)/	60.03(±2.06)/	57.33(±2.13)/	73.11(±0.70)/	67.72(±1.61)/	<b>73.63(±0.70)/</b>
	66.68(±1.32)/	63.56(±1.55)/	70.31(±1.04)/	64.75(±1.82)/	62.36(±1.90)/	76.34(±0.57)/	71.50(±1.47)/	<b>76.79(±0.63)/</b>
	73.25(±1.13)	69.23(±1.26)	76.35(±0.96)	70.19(±2.03)	68.62(±2.36)	81.75(±0.34)	77.86(±0.97)	<b>82.23(±0.49)</b>
Mean	58.32(±1.33)/	55.96(±1.71)/	64.03(±1.08)/	56.17(±1.18)/	50.92(±1.29)/	68.04(±0.71)/	68.41(±2.16)/	<b>70.26(±1.13)/</b>
	63.29(±1.17)/	60.87(±1.56)/	68.22(±0.97)/	61.25(±1.08)/	56.71(±1.15)/	71.85(±0.63)/	72.16(±1.92)/	<b>73.79(±1.01)/</b>
	69.45(±1.13)	66.73(±1.35)	74.50(±0.86)	67.24(±0.91)	61.82(±1.50)	77.95(±0.31)	77.89(±1.83)	<b>79.31(±0.77)</b>

In another experiment, the number of training samples was increased to 100 per class for the classification task. The acquired findings are shown in Table 5. Based on the results, this experiment did not affect the results, compared with the first experiment (i.e., Table 4). This is because most of the approaches in our trials are unsupervised and do not require training data for dimension reduction. In fact, the proposed WFE and FEE models

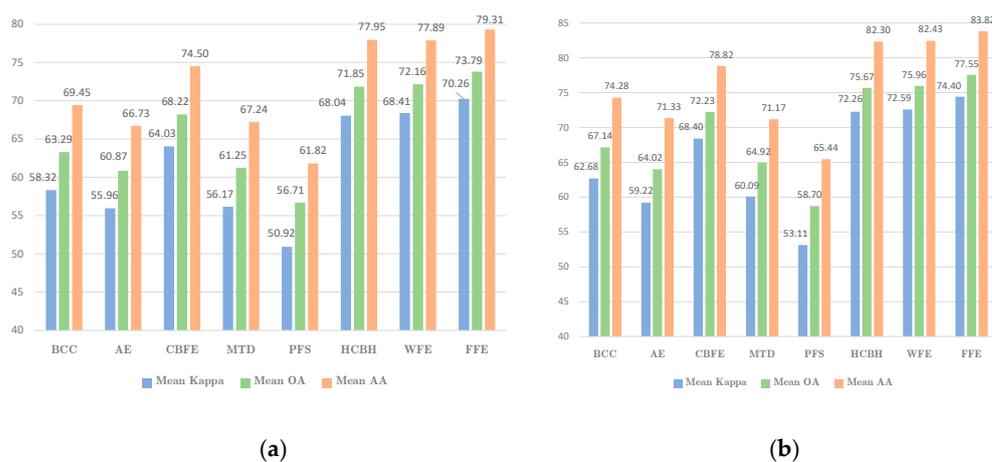
outperformed existing approaches in terms of accuracy, demonstrating the high quality of their derived features.

**Table 5.** Classification accuracy, i.e., mean  $\kappa$ /OA/AA, of different dimension reduction methods using 100 training samples per class for the Indian Pines dataset. The best results are highlighted in bold. Values of standard deviation are given in parentheses. Hyperbolic clustering-based band hierarchy (HCBH), clustering-based feature extraction (CBFE), band correlation clustering (BCC), autoencoder (AE), prototype space-based feature selection (PFS), maximum tangent discrimination (MTD), weighted feature extraction (WFE), and fuzzy feature extraction (FFE).

No. Features	Methods							
	BCC	AE	CBFE	MTD	PFS	HCBH	WFE	FFE
3	47.01(±0.00)/	51.61(±0.80)/	57.28(±0.00)/	42.10(±0.42)/	46.07(±1.57)/	56.74(±0.31)/	56.78(±1.68)/	<b>58.27(±0.33)/</b>
	52.61(±0.00)/	56.97(±0.76)/	62.28(±0.00)/	48.67(±0.36)/	52.20(±1.38)/	61.79(±0.31)/	61.93(±1.48)/	<b>63.12(±0.28)/</b>
	63.75(±0.00)	65.68(±0.84)	68.39(±0.00)	55.38(±0.38)	58.58(±1.46)	<b>70.97(±0.70)</b>	69.50(±0.97)	70.57(±0.21)
4	48.96(±0.42)/	52.52(±2.10)/	60.35(±0.00)/	51.38(±0.69)/	43.50(±1.61)/	62.61(±0.40)/	<b>69.40(±1.01)/</b>	65.42(±4.39)/
	54.64(±0.41)/	57.89(±1.98)/	65.14(±0.00)/	57.05(±0.68)/	49.85(±1.59)/	67.04(±1.19)/	<b>73.13(±0.91)/</b>	69.55(±3.98)/
	64.11(±0.24)	66.06(±1.74)	70.85(±0.00)	63.84(±0.23)	56.47(±1.35)	75.60(±0.20)	<b>79.80(±0.58)</b>	76.64(±2.45)
5	55.09(±0.27)/	54.82(±2.53)/	63.75(±0.00)/	53.86(±0.18)/	41.34(±0.52)/	64.92(±0.48)/	<b>71.89(±0.65)/</b>	69.99(±0.21)/
	60.42(±0.27)/	60.05(±2.35)/	68.13(±0.00)/	59.39(±0.05)/	48.24(±0.52)/	69.07(±0.42)/	<b>75.32(±0.59)/</b>	73.62(±0.19)/
	67.09(±0.09)	67.70(±1.70)	74.25(±0.00)	66.21(±0.58)	53.45(±0.46)	77.32(±0.55)	<b>82.30(±0.26)</b>	81.15(±0.19)
6	57.43(±0.41)/	57.08(±0.84)/	67.00(±0.37)/	56.81(±1.24)/	44.94(±0.77)/	66.85(±0.36)/	<b>74.38(±2.80)/</b>	71.15(±0.50)/
	62.50(±0.40)/	62.16(±0.81)/	71.04(±0.34)/	62.09(±1.02)/	51.52(±0.72)/	70.84(±0.51)/	<b>77.54(±2.46)/</b>	74.67(±0.46)/
	69.30(±0.19)	69.44(±0.76)	77.40(±0.23)	68.18(±1.61)	56.81(±0.74)	78.88(±0.21)	<b>83.36(±2.25)</b>	81.32(±0.40)
7	62.76(±2.05)/	58.80(±1.10)/	68.15(±0.17)/	57.98(±0.64)/	50.93(±2.09)/	67.72(±0.29)/	<b>75.46(±1.07)/</b>	73.09(±1.38)/
	67.27(±1.84)/	63.67(±1.00)/	71.94(±0.16)/	63.07(±0.42)/	56.82(±1.88)/	71.58(±0.98)/	<b>78.49(±0.96)/</b>	76.40(±1.25)/
	73.94(±1.72)	70.77(±0.97)	79.19(±0.01)	69.28(±1.04)	62.88(±2.42)	80.09(±0.54)	<b>84.31(±0.70)</b>	82.80(±0.74)
8	62.39(±1.06)/	58.78(±3.24)/	67.91(±0.56)/	60.02(±2.31)/	53.26(±1.06)/	72.29(±0.27)/	<b>75.93(±2.09)/</b>	73.97(±1.31)/
	66.91(±0.95)/	63.60(±2.99)/	71.75(±0.51)/	64.85(±2.11)/	58.80(±1.01)/	75.73(±0.62)/	<b>78.90(±1.85)/</b>	77.19(±1.18)/
	73.91(±0.84)	71.07(±2.34)	78.98(±0.49)	71.14(±1.78)	65.77(±0.55)	81.75(±0.97)	<b>84.76(±1.79)</b>	83.43(±0.54)
9	64.44(±1.42)/	59.96(±1.76)/	68.55(±1.31)/	61.69(±2.84)/	54.59(±0.86)/	73.97(±0.75)/	75.98(±3.29)/	<b>76.16(±1.32)/</b>
	68.75(±1.25)/	64.72(±1.59)/	72.35(±1.15)/	66.38(±2.59)/	59.99(±0.80)/	77.23(±0.56)/	78.96(±2.92)/	<b>79.13(±1.18)/</b>
	75.51(±1.31)	71.78(±1.16)	79.34(±1.21)	72.43(±2.18)	67.13(±1.02)	83.17(±0.41)	84.70(±2.55)	<b>85.15(±0.83)</b>
10	64.33(±1.45)/	59.57(±1.84)/	70.37(±1.17)/	62.75(±2.42)/	54.88(±0.81)/	74.58(±0.37)/	74.75(±3.12)/	<b>76.82(±0.91)/</b>
	68.67(±1.30)/	64.36(±1.65)/	74.00(±1.06)/	67.36(±2.16)/	60.21(±0.76)/	77.78(±0.46)/	77.88(±2.76)/	<b>79.71(±0.80)/</b>
	75.29(±1.06)	71.60(±1.78)	80.56(±0.83)	73.15(±2.09)	67.29(±1.04)	83.59(±0.27)	83.92(±2.65)	<b>85.79(±0.73)</b>
11	65.68(±2.44)/	61.01(±3.28)/	70.83(±1.30)/	64.30(±1.92)/	56.03(±1.36)/	74.79(±0.22)/	75.46(±3.18)/	<b>77.44(±0.60)/</b>
	69.86(±2.16)/	65.64(±3.00)/	74.42(±1.15)/	68.71(±1.77)/	61.36(±1.16)/	77.97(±0.97)/	78.51(±2.82)/	<b>80.26(±0.52)/</b>
	76.36(±2.10)	72.90(±2.08)	80.96(±1.08)	74.49(±1.28)	68.31(±1.26)	83.40(±0.53)	84.62(±2.37)	<b>86.31(±0.64)</b>
12	66.36(±2.27)/	60.83(±0.86)/	70.83(±1.02)/	65.70(±1.64)/	55.99(±1.03)/	75.80(±0.26)/	74.72(±2.69)/	<b>77.47(±0.52)/</b>
	70.46(±2.03)/	65.53(±0.83)/	74.43(±0.91)/	69.91(±1.44)/	61.31(±0.95)/	78.83(±0.41)/	77.86(±2.39)/	<b>80.29(±0.46)/</b>
	76.98(±1.91)	72.82(±0.86)	80.75(±0.58)	75.92(±1.42)	68.59(±0.95)	84.99(±0.68)	83.95(±1.81)	<b>86.32(±0.43)</b>
13	67.88(±1.39)/	61.91(±1.51)/	72.21(±1.20)/	65.93(±1.42)/	56.49(±0.94)/	76.60(±0.34)/	72.32(±4.24)/	<b>77.86(±0.68)/</b>
	71.84(±1.22)/	66.45(±1.37)/	75.66(±1.06)/	70.10(±1.25)/	61.82(±0.88)/	79.56(±0.84)/	75.73(±3.77)/	<b>80.64(±0.59)/</b>
	78.37(±1.30)	73.73(±1.03)	81.79(±0.94)	76.03(±1.32)	69.10(±0.85)	85.28(±0.75)	82.26(±3.27)	<b>86.66(±0.72)/</b>
14	67.36(±2.05)/	59.93(±2.02)/	72.41(±1.10)/	63.81(±2.94)/	57.12(±1.56)/	78.40(±0.20)/	73.02(±3.68)/	<b>78.50(±0.62)/</b>
	71.35(±1.83)/	64.68(±1.79)/	75.81(±0.98)/	68.23(±2.59)/	62.33(±1.38)/	81.16(±0.85)/	76.34(±3.28)/	<b>81.19(±0.55)/</b>
	78.18(±1.79)	71.59(±2.13)	82.11(±0.76)	74.58(±2.37)	69.40(±1.31)	86.59(±0.82)	82.82(±2.58)	<b>87.15(±0.44)</b>
15	67.25(±1.85)/	62.66(±1.63)/	71.27(±1.82)/	63.77(±3.19)/	57.73(±1.63)/	77.65(±0.77)/	72.99(±2.60)/	<b>78.42(±0.55)/</b>
	71.28(±1.66)/	67.12(±1.49)/	74.80(±1.63)/	68.18(±2.82)/	62.92(±1.49)/	80.49(±1.06)/	76.31(±2.33)/	<b>81.12(±0.49)/</b>
	78.00(±1.67)	74.30(±1.03)	81.41(±1.12)	74.62(±2.59)	70.09(±1.44)	85.96(±0.67)	83.14(±1.77)	<b>86.94(±0.43)</b>
16	69.11(±1.31)/	61.87(±1.87)/	71.19(±1.36)/	63.75(±2.83)/	58.13(±1.60)/	77.33(±0.30)/	73.04(±2.57)/	<b>78.47(±0.66)/</b>
	72.93(±1.14)/	66.45(±1.69)/	74.70(±1.23)/	68.22(±2.50)/	63.24(±1.40)/	80.19(±0.57)/	76.36(±2.28)/	<b>81.17(±0.59)/</b>
	79.75(±1.06)	73.16(±1.59)	81.61(±0.99)	74.31(±2.36)	70.31(±1.41)	86.15(±0.85)	83.31(±1.88)	<b>86.86(±0.45)</b>
17	68.41(±1.02)/	63.18(±1.53)/	71.35(±1.59)/	63.28(±2.48)/	58.86(±1.31)/	77.96(±0.16)/	72.32(±1.12)/	<b>78.70(±0.94)/</b>
	72.30(±0.89)/	67.60(±1.39)/	74.85(±1.42)/	67.79(±2.25)/	63.87(±1.19)/	80.76(±0.68)/	75.72(±0.98)/	<b>81.39(±0.83)/</b>
	78.96(±0.73)	74.39(±0.83)	81.79(±1.09)	73.96(±1.90)	70.89(±1.14)	86.18(±0.68)	82.78(±0.94)	<b>86.96(±0.65)</b>
18	68.49(±1.55)/	62.91(±1.49)/	70.95(±1.30)/	64.33(±2.64)/	59.89(±0.89)/	78.02(±0.03)/	73.02(±1.38)/	<b>78.69(±1.03)/</b>
	72.38(±1.36)/	67.36(±1.35)/	74.46(±1.17)/	68.71(±2.40)/	64.80(±0.84)/	80.77(±0.60)/	76.34(±1.23)/	<b>81.37(±0.92)/</b>
	78.98(±1.26)	74.32(±1.02)	81.72(±0.95)	75.18(±2.21)	72.01(±0.46)	86.91(±0.52)	83.32(±1.14)	<b>86.99(±0.66)</b>
Mean	62.68(±1.31)/	59.22(±1.78)/	68.40(±0.89)/	60.09(±1.86)/	53.11(±1.23)/	72.26(±0.34)/	72.59(±2.32)/	<b>74.40(±1.00)/</b>
	67.14(±1.17)/	64.02(±1.63)/	72.23(±0.80)/	64.92(±1.65)/	58.77(±1.12)/	75.67(±0.69)/	75.96(±2.06)/	<b>77.55(±0.89)/</b>
	74.28(±1.08)	71.33(±1.37)	78.82(±0.64)	71.17(±1.58)	65.44(±1.12)	82.30(±0.58)	82.43(±1.72)	<b>83.82(±0.66)</b>

To compare the overall performance of our proposed method with others, all the obtained kappa coefficients, OA and AA were averaged, and the results were reported in

Tables 4 and 5, as well as in Figure 5. As shown in Figure 5, the proposed FFE and WFE models had better performance compared with the other models. In the experiments with 50 and 100 training samples per class, FFE produced the highest accuracy, with kappa coefficient/OA/AA values of 70.26/73.79/79.31 and 74.40/77.55/83.82, respectively. This was higher than the values for WFE (68.41/72.16/77.89 and 72.59/75.96/82.43). In this dataset, the best competing approach was HCBH, which had kappa coefficient/OA/AA values of 68.04/71.85/77.95 and 72.26/75.67/82.30 in the 50 and 100 training sample experiments, respectively. When the kappa coefficients are compared, we discover that the FFE's results were 3.26% and 2.96% more accurate than those of HCBH, on average. More importantly, FFE results were, on average, about 9% more accurate than its supervised version (i.e., the CBFE method), according to Figure 5.

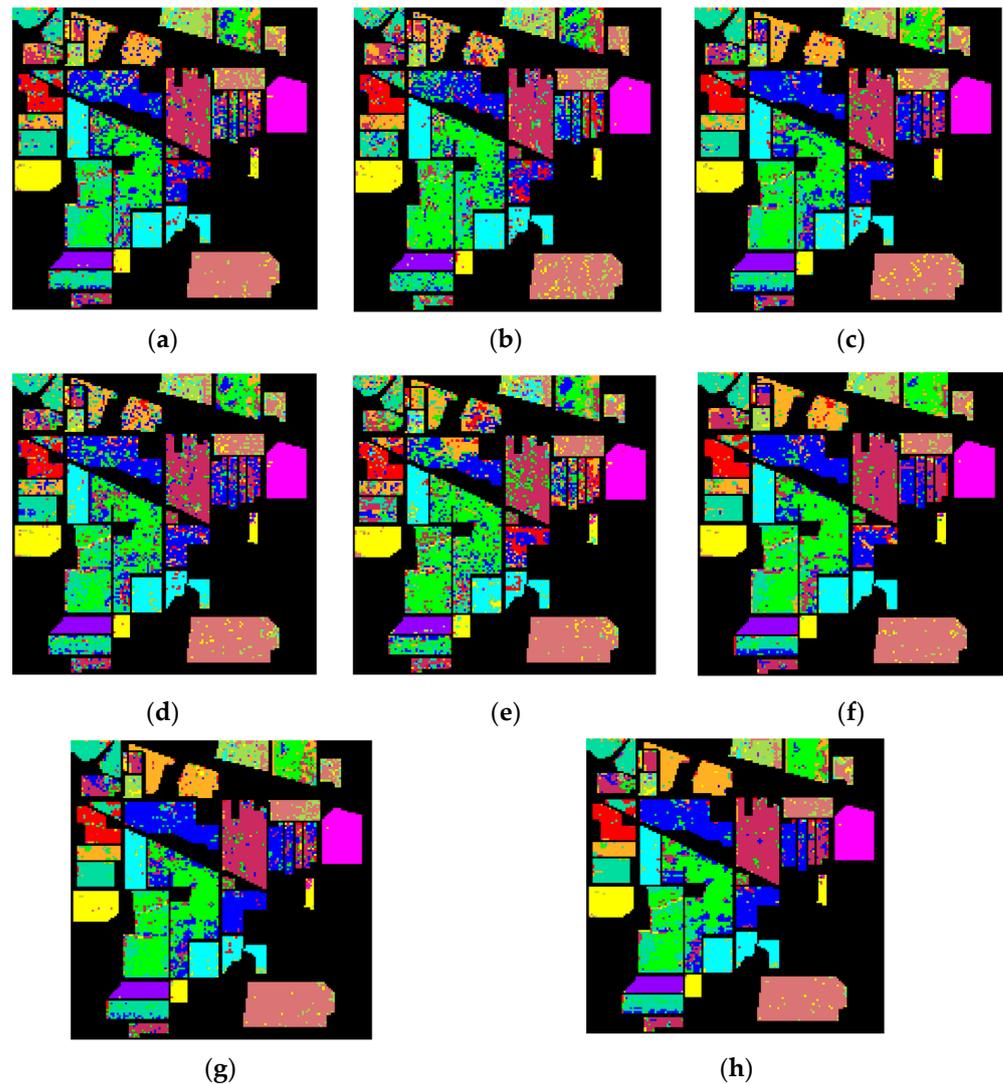


**Figure 5.** Indian Pines dataset: classification accuracy for different methods, including the proposed models in this study, which are fuzzy feature extraction (FFE) and weighted feature extraction (WFE), and competing methods: hyperbolic clustering-based band hierarchy (HCBH), clustering-based feature extraction (CBFE), autoencoder (AE), band correlation clustering (BCC), prototype space-based feature selection (PFS), and maximum tangent discrimination (MTD); (a,b) illustrate the mean kappa coefficient (multiplied by 100), OA and AA for the cases of 50 and 100 training samples per class, respectively.

Based on the mean STD values in the last rows of Tables 4 and 5, it is observed that HCBH had the highest stability and the lowest mean STD values. For example, according to Table 5, which shows the classification accuracy using 100 training samples per class, the HCBH's STD values are 0.34, 0.69 and 0.58 for the  $\kappa$ , OA and AA values, respectively. This suggests that HCBH is a reliable and stable algorithm for classification purposes. WFE, on the other hand, exhibits the least stability, with STD values of 2.32, 2.06, and 1.72 (Table 5). Other methods in this dataset have close STD values, indicating that they are almost the same from a stability point of view. However, when it comes to accuracy, these methods show significant variations. This suggests that while they may provide similar levels of stability, they may not necessarily yield equally precise results. Therefore, it is important to consider both stability and accuracy when choosing a method for any given analysis.

Overall, there are three reasons for our method's superiority: (1) Despite our technique being unsupervised and having no knowledge of the existing classes in the image, we were able to discover pure pixels of the image that reflect existing phenomena in the image using an endmember algorithm. These pure pixels provided significant information, equivalent to training samples for building a PS. (2) The prototype space helped identify clusters of correlated and similar spectral bands, which were then combined using a weighted-mean operator to decrease the high dimension of the HSIs without sacrificing much spectral information. (3) The extracted features have a greater signal-to-noise ratio since they are the weighted average of several correlated bands in the original HSI, which should reduce

random noise. Figure 6 shows the final results of each method, using 100 training samples per class and 10 features.

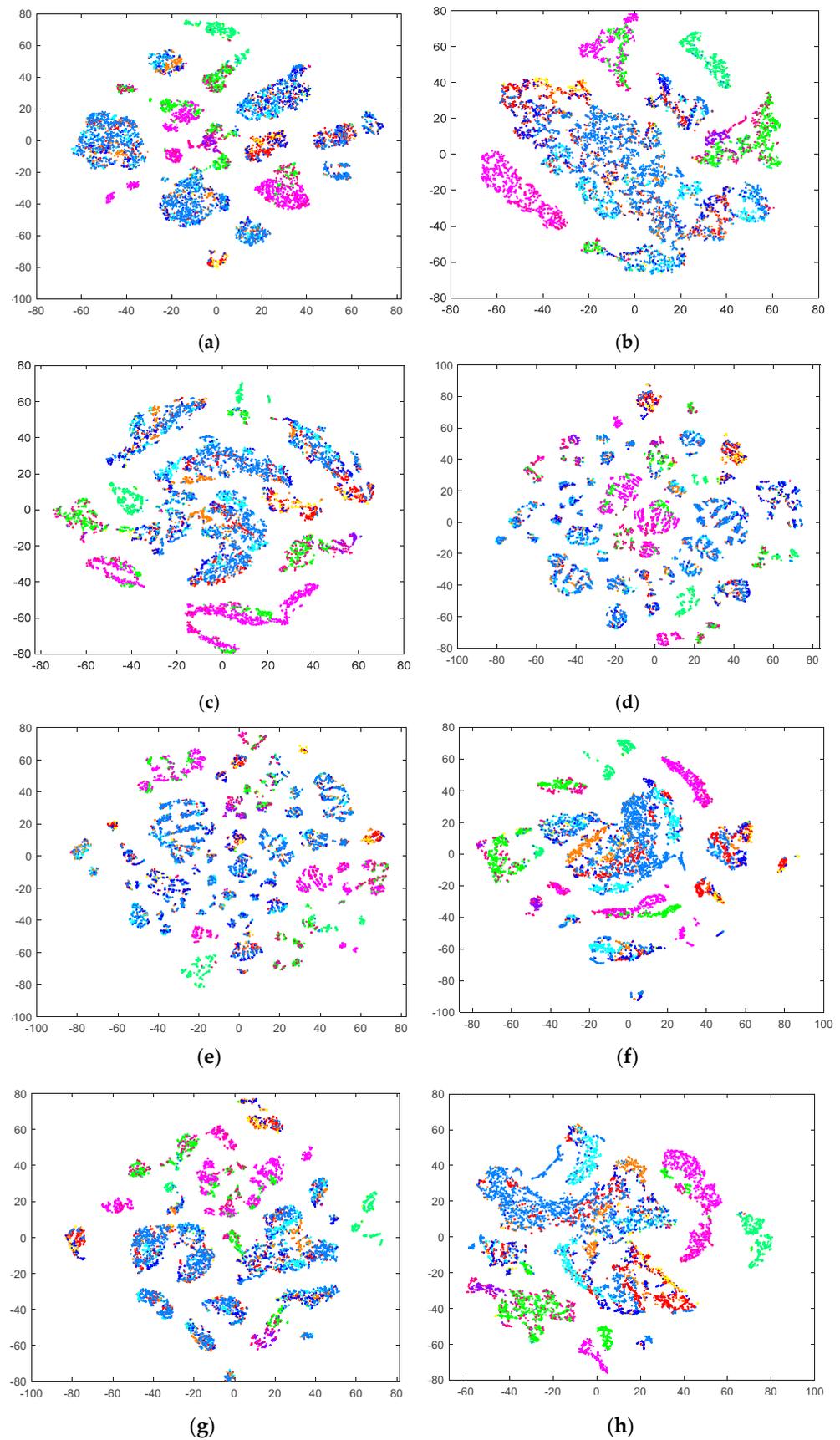


**Figure 6.** The classified image of Indian Pine dataset using 100 training samples per class and 10 features: (a) BCC with  $\kappa = 64.95$ , (b) AE with  $\kappa = 63.18$ , (c) CBF with  $\kappa = 69.82$ , (d) MTD with  $\kappa = 62.10$ , (e) PFS with  $\kappa = 54.58$ , (f) HCBH with  $\kappa = 73.20$ , (g) WFE with  $\kappa = 74.99$ , and (h) FFE with  $\kappa = 76.70$ .

The outcomes of the T-SNE experiment are visualized in Figure 7 where distinct colors represent the 12 classes within the Indian Pine dataset. As depicted in this figure, FFE and HCBH exhibit superior performance compared with the other methods. Notably, points belonging to the same class clustered together, forming distinct groupings. The superior performance of these two methods was further validated through image classification, as indicated by the highest classification accuracy for the 18 features, as presented in Tables 4 and 5.

### 3.4.2. Pavia Centre

The results for the Pavia Centre dataset with five training samples per class are provided in Table 6. The best outcome for each row of this table is bolded. The classification accuracies of all approaches were mostly similar, although the proposed WFE model had the highest  $\kappa$  value in eight cases, which was greater than those of other methods. This demonstrates the superiority of WFE's extracted features.



**Figure 7.** T–SNE results of Indian Pines dataset using 18 features; each color represents a class in this dataset; (a) BCC, (b) AE, (c) CBFE (d) MTD (e) PFS (f) HCBH (g) WFE, and (h) FFE.

**Table 6.** Classification accuracy, i.e., mean  $\kappa$ /OA/AA, of different dimension reduction methods using five training samples per class for the Pavia Centre dataset. The best results are highlighted in bold. Values of standard deviation are given in parentheses. Hyperbolic clustering-based band hierarchy (HCBH), clustering-based feature extraction (CBFE), band correlation clustering (BCC), autoencoder (AE), prototype space-based feature selection (PFS), maximum tangent discrimination (MTD), weighted feature extraction (WFE), and fuzzy feature extraction (FFE).

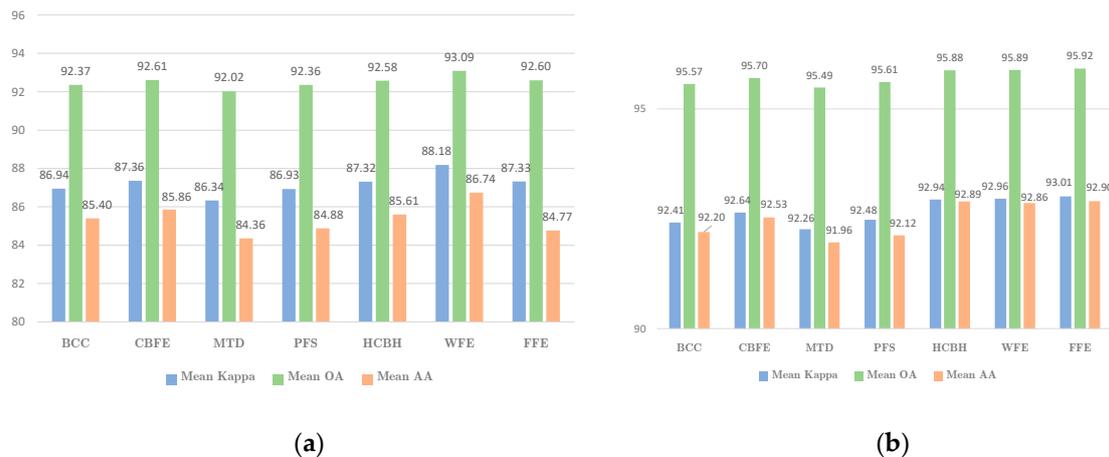
No. Features	Methods							
	BCC	AE	CBFE	MTD	PFS	HCBH	WFE	FFE
3	87.51(±0.00)/	86.30(±0.84)/	88.41(±0.00)/	84.25(±3.02)/	<b>88.93</b> (±0.10)/	85.36(±0.41)/	85.90(±1.25)/	86.05(±0.70)/
	92.70(±0.00)/	91.96(±0.49)/	93.23(±0.00)/	90.82(±1.75)/	<b>93.53</b> (±0.06)/	91.43(±0.32)/	91.74(±0.73)/	91.85(±0.41)/
	86.10(±0.00)	85.49(±0.88)	86.18(±0.00)	80.72(±5.10)	<b>87.00</b> (±0.02)	82.46(±0.59)	84.04(±1.66)	83.19(±0.87)
4	87.32(±0.00)/	84.76(±3.03)/	87.33(±0.00)/	84.20(±1.06)/	86.83(±0.06)/	83.72(±0.45)/	<b>87.57</b> (±0.82)/	86.32(±0.08)/
	92.58(±0.00)/	91.07(±1.78)/	92.59(±0.00)/	90.78(±0.61)/	92.31(±0.03)/	90.45(±0.36)/	<b>92.72</b> (±0.49)/	92.01(±0.04)/
	86.23(±0.00)	83.32(±3.71)	86.27(±0.00)	81.06(±1.84)	83.75(±0.19)	80.78(±0.63)	<b>86.53</b> (±0.91)	84.27(±0.23)
5	87.38(±0.03)/	86.09(±1.59)/	88.07(±0.25)/	86.01(±1.68)/	86.09(±1.72)/	82.02(±0.22)/	<b>88.56</b> (±0.28)/	85.86(±0.22)/
	92.62(±0.01)/	91.84(±0.93)/	93.02(±0.14)/	91.83(±0.98)/	91.87(±1.02)/	89.47(±0.13)/	<b>93.31</b> (±0.16)/	91.74(±0.13)/
	85.90(±0.07)	85.19(±2.15)	86.76(±0.27)	84.07(±1.98)	82.97(±1.92)	76.91(±0.40)	<b>87.80</b> (±0.42)	82.19(±0.42)
6	<b>88.84</b> (±0.46)/	87.05(±0.70)/	88.59(±0.96)/	85.71(±4.01)/	87.09(±0.52)/	87.89(±0.22)/	88.66(±0.15)/	86.77(±0.10)/
	<b>93.47</b> (±0.27)/	92.41(±0.42)/	93.33(±0.56)/	91.61(±2.44)/	92.45(±0.31)/	92.92(±0.13)/	93.36(±0.09)/	92.27(±0.06)/
	87.58(±0.53)	86.06(±0.71)	87.03(±1.51)	84.21(±3.98)	85.11(±0.64)	85.90(±0.40)	<b>88.15</b> (±0.20)	84.40(±0.19)
7	86.69(±0.81)/	85.69(±2.65)/	<b>88.97</b> (±0.26)/	86.80(±1.62)/	87.14(±0.52)/	87.31(±0.21)/	88.72(±0.08)/	87.23(±0.07)/
	92.22(±0.47)/	91.61(±1.57)/	<b>93.55</b> (±0.15)/	92.30(±0.94)/	92.49(±0.30)/	92.58(±0.12)/	93.40(±0.05)/	92.54(±0.04)/
	84.80(±1.02)	84.08(±4.04)	87.71(±0.26)	84.84(±1.92)	84.74(±0.77)	85.65(±0.39)	<b>88.03</b> (±0.06)	85.47(±0.29)
8	86.56(±0.08)/	86.19(±1.13)/	86.42(±0.87)/	86.74(±0.45)/	87.49(±0.20)/	87.69(±0.26)/	<b>88.38</b> (±0.14)/	87.46(±0.52)/
	92.14(±0.05)/	91.91(±0.66)/	92.06(±0.51)/	92.26(±0.26)/	92.69(±0.12)/	92.80(±0.17)/	<b>93.20</b> (±0.08)/	92.67(±0.31)/
	84.79(±0.17)	84.43(±2.33)	83.89(±1.67)	84.78(±0.65)	85.86(±0.73)	86.17(±0.44)	<b>87.92</b> (±0.19)	86.23(±0.45)
9	86.51(±0.12)/	87.57(±1.15)/	86.06(±0.80)/	86.57(±0.71)/	87.47(±0.26)/	87.38(±0.32)/	<b>88.04</b> (±0.26)/	88.00(±0.28)/
	92.11(±0.07)/	92.73(±0.67)/	91.84(±0.47)/	92.16(±0.41)/	92.67(±0.15)/	92.61(±0.23)/	<b>93.00</b> (±0.16)/	92.98(±0.16)/
	84.82(±0.24)	86.29(±1.06)	83.91(±1.26)	84.63(±1.11)	86.25(±0.30)	86.05(±0.50)	<b>87.51</b> (±0.34)	86.20(±1.05)
10	86.78(±0.17)/	87.40(±1.08)/	86.68(±0.82)/	86.90(±0.37)/	87.39(±0.15)/	87.55(±0.21)/	<b>88.39</b> (±0.21)/	87.38(±0.36)/
	92.27(±0.10)/	92.63(±0.63)/	92.21(±0.48)/	92.35(±0.22)/	92.63(±0.09)/	92.72(±0.12)/	<b>93.22</b> (±0.12)/	92.63(±0.21)/
	85.48(±0.30)	85.72(±1.78)	84.94(±1.38)	84.97(±0.65)	85.99(±0.10)	86.21(±0.39)	<b>86.49</b> (±0.21)	84.00(±0.80)
11	86.75(±0.21)/	86.34(±3.30)/	87.34(±0.16)/	86.86(±0.54)/	87.29(±0.28)/	89.45(±0.20)/	88.47(±0.20)/	87.57(±0.26)/
	92.25(±0.12)/	92.00(±1.95)/	92.59(±0.09)/	92.33(±0.32)/	92.58(±0.16)/	<b>93.83</b> (±0.11)/	93.26(±0.12)/	92.74(±0.15)/
	85.49(±0.38)	84.24(±4.49)	86.06(±0.51)	85.32(±0.50)	85.70(±0.31)	<b>88.24</b> (±0.38)	86.67(±0.30)	84.24(±0.58)
12	86.79(±0.06)/	87.29(±0.27)/	87.30(±0.23)/	86.76(±0.36)/	87.25(±0.24)/	<b>89.82</b> (±0.24)/	88.52(±0.33)/	87.43(±0.10)/
	92.27(±0.03)/	92.56(±0.16)/	92.57(±0.13)/	92.26(±0.21)/	92.55(±0.14)/	<b>94.05</b> (±0.15)/	93.29(±0.19)/	92.66(±0.06)/
	85.43(±0.18)	86.11(±0.58)	86.11(±0.47)	85.23(±0.39)	85.55(±0.15)	<b>88.77</b> (±0.42)	86.87(±0.43)	83.86(±0.18)
13	86.78(±0.25)/	86.42(±1.36)/	87.07(±0.22)/	86.68(±0.34)/	86.57(±0.18)/	<b>89.19</b> (±0.24)/	88.67(±0.24)/	87.81(±0.41)/
	92.27(±0.14)/	92.04(±0.80)/	92.43(±0.13)/	92.22(±0.20)/	92.16(±0.10)/	<b>93.68</b> (±0.15)/	93.38(±0.14)/	92.88(±0.24)/
	85.32(±0.50)	85.26(±1.73)	85.85(±0.41)	84.82(±0.31)	84.87(±0.21)	<b>88.37</b> (±0.42)	87.05(±0.28)	84.81(±0.91)
14	86.87(±0.22)/	86.20(±1.85)/	87.14(±0.15)/	86.48(±0.23)/	86.50(±0.29)/	<b>89.27</b> (±0.22)/	88.27(±0.63)/	87.65(±0.09)/
	92.32(±0.13)/	91.91(±1.09)/	92.48(±0.09)/	92.10(±0.13)/	92.12(±0.17)/	<b>93.72</b> (±0.13)/	93.15(±0.36)/	92.79(±0.05)/
	85.53(±0.54)	85.07(±2.18)	85.93(±0.26)	84.74(±0.46)	84.64(±0.27)	<b>88.45</b> (±0.40)	86.17(±1.14)	84.43(±0.19)
15	86.39(±0.65)/	86.52(±0.94)/	87.10(±0.14)/	86.52(±0.31)/	86.32(±0.26)/	87.51(±0.24)/	<b>88.32</b> (±0.82)/	87.68(±0.09)/
	92.04(±0.38)/	92.10(±0.57)/	92.45(±0.08)/	92.12(±0.18)/	92.01(±0.15)/	92.69(±0.14)/	<b>93.17</b> (±0.48)/	92.81(±0.05)/
	84.29(±1.48)	85.30(±0.97)	86.07(±0.27)	84.70(±0.36)	84.28(±0.28)	86.40(±0.41)	<b>86.43</b> (±1.68)	84.43(±0.17)
16	86.24(±0.69)/	87.16(±1.05)/	87.17(±0.18)/	86.85(±0.49)/	86.30(±0.20)/	87.64(±0.23)/	<b>88.04</b> (±0.56)/	87.81(±0.32)/
	91.96(±0.40)/	92.48(±0.62)/	92.49(±0.11)/	92.32(±0.28)/	92.00(±0.11)/	92.77(±0.13)/	<b>93.01</b> (±0.32)/	92.88(±0.19)/
	83.90(±1.63)	85.53(±1.35)	85.86(±0.39)	85.04(±0.71)	84.09(±0.31)	<b>86.59</b> (±0.40)	85.83(±1.25)	85.07(±1.15)
17	86.79(±0.11)/	87.67(±0.55)/	87.05(±0.26)/	87.05(±0.38)/	86.05(±0.17)/	87.71(±0.25)/	88.08(±0.56)/	<b>88.23</b> (±0.35)/
	92.27(±0.07)/	92.78(±0.32)/	92.43(±0.15)/	92.43(±0.22)/	91.85(±0.10)/	92.80(±0.16)/	93.04(±0.33)/	<b>93.12</b> (±0.21)/
	85.28(±0.45)	86.46(±1.01)	85.68(±0.79)	85.46(±0.76)	83.80(±0.60)	<b>86.66</b> (±0.43)	86.11(±0.90)	86.59(±0.17)
18	86.92(±0.22)/	87.02(±0.79)/	87.10(±0.49)/	87.04(±0.43)/	86.11(±0.20)/	87.63(±0.24)/	<b>88.30</b> (±0.38)/	87.96(±0.29)/
	92.35(±0.13)/	92.40(±0.46)/	92.45(±0.29)/	92.43(±0.25)/	91.89(±0.11)/	92.73(±0.14)/	<b>93.16</b> (±0.22)/	92.96(±0.17)/
	85.53(±0.27)	85.31(±1.62)	85.54(±1.04)	85.16(±0.59)	83.55(±0.34)	86.09(±0.41)	86.31(±0.69)	<b>86.90</b> (±0.19)
Mean	86.94(±0.25)/	86.60(±1.39)/	87.36(±0.36)/	86.34(±1.00)/	86.93(±0.33)/	87.32(±0.26)/	<b>88.18</b> (±0.43)/	87.33(±0.26)/
	92.37(±0.15)/	92.15(±0.82)/	92.61(±0.21)/	92.02(±0.59)/	92.36(±0.20)/	92.58(±0.17)/	<b>93.09</b> (±0.25)/	92.60(±0.15)/
	85.40(±0.49)	85.24(±1.91)	85.86(±0.66)	84.36(±1.33)	84.88(±0.45)	85.61(±0.44)	<b>86.74</b> (±0.67)	84.77(±0.49)

We also increased the number of training samples for the classification task to 20 per class, and the results are provided in Table 7. Similar to Table 6, all the methods had high accuracies. However, the proposed WFE and FFE models outperformed the others. WFE and FFE had the greatest kappa coefficient five and seven times, respectively, according to

Table 7. To evaluate the overall performance of the models, the kappa coefficients, OA and AA values obtained from each method were averaged (see Figure 8).

**Table 7.** Classification accuracy, i.e., mean  $\kappa$ /OA/AA, of different dimension reduction methods using 20 training samples per class for the Pavia Centre dataset. The best results are highlighted in bold. Values of standard deviation are given in parentheses. Hyperbolic clustering-based band hierarchy (HCBH), autoencoder (AE), clustering-based feature extraction (CBFE), band correlation clustering (BCC), prototype space-based feature selection (PFS), maximum tangent discrimination (MTD), weighted feature extraction (WFE), and fuzzy feature extraction (FFE).

No. Features	Methods							
	BCC	AE	CBFE	MTD	PFS	HCBH	WFE	FFE
3	91.23(±0.00)/	91.61(±0.78)/	92.43(±0.01)/	89.94(±3.30)/	90.92(±0.36)/	<b>92.68(±0.05)/</b>	91.37(±0.96)/	91.92(±0.19)/
	94.88(±0.00)/	95.10(±0.46)/	95.58(±0.00)/	94.14(±1.93)/	94.70(±0.21)/	<b>95.72(±0.00)/</b>	94.96(±0.56)/	95.28(±0.12)/
	91.16(±0.00)	91.58(±0.68)	92.32(±0.00)	89.08(±4.08)	90.48(±0.30)	<b>92.55(±0.04)</b>	90.71(±1.00)	91.51(±0.02)
4	92.62(±0.00)/	92.29(±0.71)/	92.65(±0.07)/	91.02(±0.56)/	89.64(±1.34)/	92.77(±0.15)/	92.48(±0.54)/	<b>93.06(±0.02)/</b>
	95.70(±0.00)/	95.49(±0.42)/	95.71(±0.04)/	94.76(±0.33)/	93.95(±0.78)/	95.78(±0.02)/	95.61(±0.32)/	<b>95.95(±0.01)/</b>
	92.59(±0.00)	92.33(±0.46)	92.65(±0.10)	90.64(±0.71)	88.98(±1.85)	<b>92.82(±0.05)</b>	92.21(±0.75)	92.79(±0.06)
5	92.61(±0.08)/	92.03(±0.32)/	92.59(±0.01)/	91.12(±1.35)/	92.14(±0.80)/	92.91(±0.10)/	92.96(±0.18)/	<b>93.31(±0.14)/</b>
	95.66(±0.02)/	95.35(±0.19)/	95.68(±0.01)/	94.82(±0.79)/	95.41(±0.46)/	95.86(±0.02)/	95.89(±0.10)/	<b>96.10(±0.08)/</b>
	92.60(±0.02)	92.03(±0.29)	92.78(±0.02)	90.70(±1.74)	91.72(±1.39)	<b>93.09(±0.06)</b>	92.96(±0.11)	93.04(±0.04)
6	92.60(±0.09)/	92.42(±0.30)/	92.52(±0.02)/	92.30(±0.47)/	92.91(±0.18)/	92.95(±0.06)/	92.80(±0.15)/	<b>93.49(±0.10)/</b>
	95.65(±0.02)/	95.57(±0.18)/	95.64(±0.01)/	95.51(±0.28)/	95.86(±0.10)/	95.89(±0.03)/	95.80(±0.09)/	<b>96.20(±0.06)/</b>
	92.75(±0.04)	92.42(±0.33)	92.72(±0.02)	91.92(±0.71)	92.70(±0.18)	93.16(±0.08)	92.75(±0.19)	<b>93.24(±0.17)</b>
7	92.93(±0.16)/	92.51(±0.33)/	92.67(±0.04)/	92.41(±0.26)/	92.84(±0.34)/	92.79(±0.07)/	<b>93.27(±0.10)/</b>	92.88(±0.31)/
	95.83(±0.11)/	95.62(±0.20)/	95.75(±0.04)/	95.58(±0.16)/	95.82(±0.20)/	95.79(±0.13)/	<b>96.07(±0.06)/</b>	95.84(±0.18)/
	92.61(±0.27)	92.46(±0.25)	92.62(±0.11)	92.16(±0.26)	92.67(±0.20)	92.67(±0.16)	<b>93.29(±0.09)</b>	92.98(±0.22)
8	92.82(±0.25)/	92.27(±0.43)/	92.69(±0.13)/	92.76(±0.21)/	93.01(±0.21)/	92.73(±0.08)/	<b>93.26(±0.10)/</b>	92.78(±0.02)/
	95.82(±0.14)/	95.49(±0.25)/	95.68(±0.08)/	95.78(±0.12)/	95.92(±0.13)/	95.76(±0.14)/	<b>96.07(±0.06)/</b>	95.79(±0.01)/
	92.49(±0.12)	92.37(±0.37)	92.58(±0.12)	92.49(±0.50)	92.70(±0.15)	92.62(±0.17)	<b>93.32(±0.06)</b>	92.90(±0.02)
9	92.30(±0.40)/	92.61(±0.15)/	92.77(±0.17)/	92.54(±0.32)/	92.90(±0.33)/	92.74(±0.13)/	<b>93.11(±0.30)/</b>	92.84(±0.11)/
	95.59(±0.09)/	95.69(±0.09)/	95.72(±0.09)/	95.65(±0.19)/	95.86(±0.19)/	95.76(±0.04)/	<b>95.98(±0.17)/</b>	95.82(±0.07)/
	92.29(±0.03)	92.62(±0.15)	92.57(±0.22)	92.16(±0.24)	92.67(±0.44)	92.58(±0.08)	<b>92.58(±0.33)</b>	92.87(±0.03)
10	92.23(±0.08)/	92.51(±0.18)/	92.60(±0.04)/	92.52(±0.39)/	92.55(±0.37)/	92.90(±0.08)/	<b>93.34(±0.30)/</b>	92.93(±0.05)/
	95.49(±0.27)/	95.62(±0.11)/	95.73(±0.05)/	95.64(±0.23)/	95.65(±0.21)/	95.86(±0.05)/	<b>96.12(±0.17)/</b>	95.87(±0.03)/
	92.10(±0.35)	92.73(±0.17)	92.58(±0.04)	92.30(±0.41)	92.31(±0.34)	92.46(±0.12)	<b>93.25(±0.34)</b>	92.99(±0.07)
11	92.24(±0.43)/	92.06(±0.51)/	92.72(±0.20)/	92.83(±0.22)/	92.88(±0.31)/	<b>93.10(±0.12)/</b>	92.93(±0.37)/	93.06(±0.04)/
	95.56(±0.22)/	95.36(±0.30)/	95.70(±0.16)/	95.82(±0.13)/	95.85(±0.18)/	<b>95.98(±0.01)/</b>	95.88(±0.21)/	95.95(±0.02)/
	92.19(±0.16)	92.26(±0.41)	92.47(±0.30)	92.74(±0.10)	92.66(±0.34)	92.75(±0.09)	92.85(±0.44)	<b>93.08(±0.07)</b>
12	92.30(±0.42)/	92.59(±0.38)/	92.63(±0.10)/	92.69(±0.20)/	92.67(±0.33)/	<b>93.21(±0.07)/</b>	93.17(±0.29)/	93.19(±0.03)/
	95.50(±0.25)/	95.67(±0.22)/	95.74(±0.10)/	95.74(±0.12)/	95.73(±0.19)/	<b>96.04(±0.01)/</b>	96.02(±0.17)/	96.03(±0.02)/
	91.92(±0.44)	92.56(±0.39)	92.56(±0.20)	92.63(±0.18)	92.45(±0.36)	93.06(±0.04)	93.14(±0.37)	<b>93.19(±0.02)</b>
13	92.27(±0.30)/	92.39(±0.28)/	92.72(±0.28)/	92.56(±0.43)/	92.98(±0.12)/	92.89(±0.19)/	93.15(±0.28)/	<b>93.26(±0.03)/</b>
	95.54(±0.22)/	95.56(±0.16)/	95.66(±0.13)/	95.66(±0.25)/	95.91(±0.07)/	95.85(±0.12)/	96.01(±0.16)/	<b>96.07(±0.02)/</b>
	92.08(±0.31)	92.42(±0.48)	92.35(±0.34)	92.40(±0.63)	92.75(±0.20)	92.60(±0.17)	93.11(±0.31)	<b>93.21(±0.05)</b>
14	92.56(±0.36)/	92.42(±0.32)/	92.58(±0.28)/	92.63(±0.39)/	92.58(±0.35)/	93.15(±0.09)/	92.92(±0.37)/	<b>93.20(±0.33)/</b>
	95.50(±0.03)/	95.57(±0.19)/	95.73(±0.16)/	95.70(±0.23)/	95.68(±0.21)/	96.00(±0.06)/	95.87(±0.21)/	<b>96.03(±0.19)/</b>
	91.90(±0.09)	92.32(±0.25)	92.49(±0.35)	92.39(±0.39)	92.13(±0.27)	<b>93.37(±0.10)</b>	92.91(±0.36)	93.04(±0.40)
15	92.75(±0.53)/	92.43(±0.37)/	92.53(±0.32)/	92.74(±0.31)/	92.81(±0.34)/	93.09(±0.16)/	93.03(±0.26)/	<b>93.10(±0.42)/</b>
	95.49(±0.21)/	95.58(±0.22)/	95.65(±0.19)/	95.76(±0.18)/	95.81(±0.20)/	95.97(±0.06)/	95.94(±0.15)/	<b>95.98(±0.24)/</b>
	91.90(±0.42)	92.52(±0.40)	92.28(±0.42)	92.46(±0.43)	92.37(±0.37)	<b>93.34(±0.14)</b>	92.96(±0.34)	92.92(±0.52)
16	92.23(±0.60)/	92.27(±0.50)/	92.81(±0.20)/	92.63(±0.34)/	92.89(±0.32)/	93.04(±0.08)/	93.27(±0.25)/	<b>93.31(±0.22)/</b>
	95.65(±0.29)/	95.49(±0.29)/	95.68(±0.25)/	95.70(±0.20)/	95.85(±0.18)/	95.93(±0.10)/	96.08(±0.14)/	<b>96.10(±0.13)/</b>
	92.18(±0.50)	92.13(±0.57)	92.29(±0.47)	92.33(±0.39)	92.43(±0.39)	<b>93.26(±0.15)</b>	93.22(±0.36)	93.11(±0.31)
17	92.72(±0.30)/	92.53(±0.39)/	92.48(±0.49)/	92.67(±0.39)/	92.91(±0.29)/	93.08(±0.21)/	<b>93.27(±0.27)/</b>	92.98(±0.10)/
	95.67(±0.32)/	95.64(±0.23)/	95.77(±0.21)/	95.73(±0.23)/	95.87(±0.17)/	95.96(±0.16)/	<b>96.08(±0.15)/</b>	95.91(±0.05)/
	92.20(±0.59)	92.44(±0.49)	92.48(±0.47)	92.43(±0.27)	92.44(±0.38)	<b>93.27(±0.22)</b>	93.19(±0.37)	92.90(±0.24)
18	92.20(±0.37)/	92.37(±0.38)/	92.89(±0.27)/	92.84(±0.29)/	<b>93.04(±0.18)/</b>	92.99(±0.20)/	92.95(±0.35)/	92.87(±0.46)/
	95.63(±0.28)/	95.54(±0.22)/	95.84(±0.13)/	95.82(±0.17)/	<b>95.95(±0.11)/</b>	95.92(±0.07)/	95.89(±0.20)/	95.84(±0.27)/
	92.18(±0.59)	92.37(±0.57)	92.77(±0.33)	92.57(±0.29)	92.53(±0.20)	92.65(±0.12)	<b>92.84(±0.45)</b>	92.70(±0.60)
Mean	92.41(±0.27)/	92.33(±0.40)/	92.64(±0.17)/	92.26(±0.59)/	92.48(±0.38)/	92.94(±0.12)/	92.96(±0.32)/	<b>93.01(±0.16)/</b>
	95.57(±0.15)/	95.52(±0.23)/	95.70(±0.10)/	95.49(±0.35)/	95.61(±0.22)/	95.88(±0.06)/	95.89(±0.18)/	<b>95.92(±0.09)/</b>
	92.20(±0.25)	92.35(±0.39)	92.53(±0.22)	91.96(±0.71)	92.12(±0.46)	92.89(±0.11)	92.86(±0.37)	<b>92.90(±0.18)</b>



**Figure 8.** Pavia Centre Data set: classification accuracy for different methods, including the proposed models in this study, which are fuzzy feature extraction (FFE) and weighted feature extraction (WFE), and competing methods: hyperbolic clustering-based band hierarchy (HCBH), clustering-based feature extraction (CBFE), autoencoder (AE), band correlation clustering (BCC), prototype space-based feature selection (PFS), and maximum tangent discrimination (MTD); (a,b) illustrate the mean kappa coefficient (multiplied by 100), OA and AA for the cases of 5 and 20 training samples per class, respectively.

Figure 8a shows that WFE outperformed the other methods, with an average kappa coefficient/OA/AA of 88.18/93.09/86.74. FFE, HCBH, and CBFE produced similar results. However, it should be noted that CBFE is a supervised method, in contrast to the other two models. According to Figure 8a, when the number of training samples is limited for a supervised classifier, the feature quality of WFE outperforms both CBFE and HCBH, which are the two competing methods with the best performance in this experiment, by 0.94% and 0.98%, respectively.

FFE, WFE, and CBFE all achieved about the same classification accuracy for the case of 20 training samples per class (Figure 8b). However, FFE and WFE outperformed HCBH, yielding kappa coefficients of 93.01 and 92.96, respectively. Figure 8a further demonstrates that our suggested methodology worked more effectively for HSI dimension reduction prior to a classification task in the case of small training samples.

Given the mean STD values in the last rows of Tables 6 and 7, we can conclude that, while HCBH and WFE have the highest and lowest stability in the Indian Pines dataset, all the methods tested with Pavia Centre dataset have nearly the same stability because their STD values are close to each other. The classified image of each method with 20 training samples per class and 10 features is illustrated in Figure 9.

In a manner similar to the Indian Pine dataset, T-SNE was employed to visually evaluate the quality of the resultant features (Figure 10). As depicted in Figure 10, there is an absence of noticeable superiority in the T-SNE results across the methods. This is because points were visually well-clustered for all methods. These well-clustered points in Figure 10 justify the high accuracy values of the methods for this dataset. Although certain points may overlap, it is evident that points belonging to the same class maintain proximity, forming distinctive clusters in all methods. Consequently, T-SNE did not offer a visual basis for comparing the methods effectively. However, it is important to note that T-SNE is only a visual evaluation tool and does not provide quantitative measures of the quality of the resultant features. Therefore, it may not be sufficient to solely rely on T-SNE for comparing the methods. Additional analysis methods such as classification accuracy or clustering performance should be utilized to further assess the effectiveness of each method. These quantitative measures would provide a more comprehensive understanding of the differences between the methods and aid in making a well-informed comparison.

### 3.4.3. Kennedy Space Center (KSC)

Tables 8 and 9 provide the results of the experiments for the KSC dataset using 5 and 20 training samples per class, respectively. The best outcome is marked in bold in each table.

According to Table 8, our proposed methods (i.e., WFE and FFE) had the highest kappa coefficient values in eight cases, outperforming other methods. However, while employing 20 training samples per class, CBFE significantly surpassed our WFE in terms of the number of maximum kappa coefficient values (Table 9). According to Table 9, CBFE achieved the highest kappa coefficient values in nine cases, which was greater than WFE (i.e., five cases). Considering Figure 11, which provides the mean kappa coefficient/OA/AA values of all cases, it is clear that although CBFE and WFE competed closely, CBFE outperformed the other approaches.

CBFE, WFE, and FFE show comparable classification accuracy, according to Figure 11a, which demonstrates the overall performance of extracted/selected features in the case of five training samples. CBFE, on the other hand, had a slightly higher accuracy. In the case of 20 training samples, CBFE (with a  $\kappa$ /OA/AA of 80.36/82.37/78.29) outperformed WFE (79.01/81.13/76.11) and FFE (78.40/80.59/75.41), as shown in Figure 11b. This superiority can be explained by the fact that CBFE is a supervised approach that employs training samples to reduce dimensions. However, the proposed WFE and FFE models are unsupervised approaches, which is preferred since providing training samples is expensive and time-consuming. As a result, if we ignore the supervised CBFE, our proposed models considerably outperformed other unsupervised methods. For example, in the case of limited training samples (i.e., five samples per class), WFE and FFE had the kappa coefficient/OA/AA values of 71.17/74.08/66.07 and 70.90/73.85/65.82, respectively, which were greater than those of HCBH (65.74/69.22/62.11), PFS (65.74/69.14/62.24), MTD (61.32/65.18/56.22), and BCC (66.50/69.88/60.89). In other words, considering kappa coefficient, WFE's performance was 7.02% higher than the best unsupervised technique (i.e., BCC) in this experiment. This demonstrates that adopting our dimension reduction method is a good choice when few training samples are available for classifying HSIs.

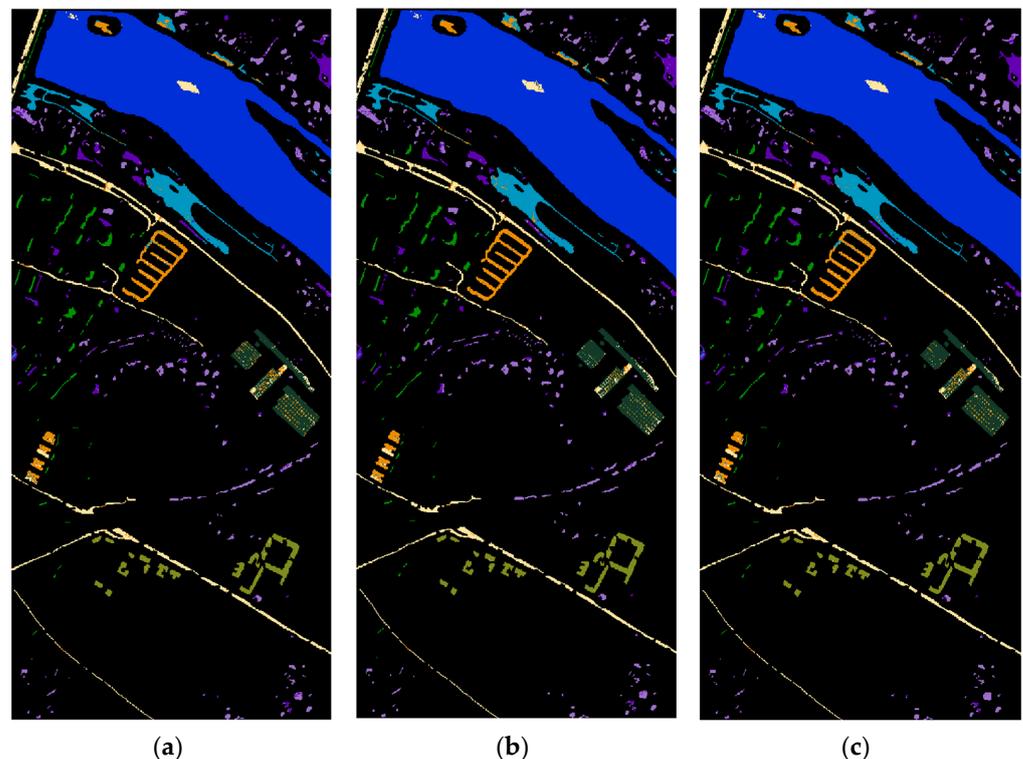
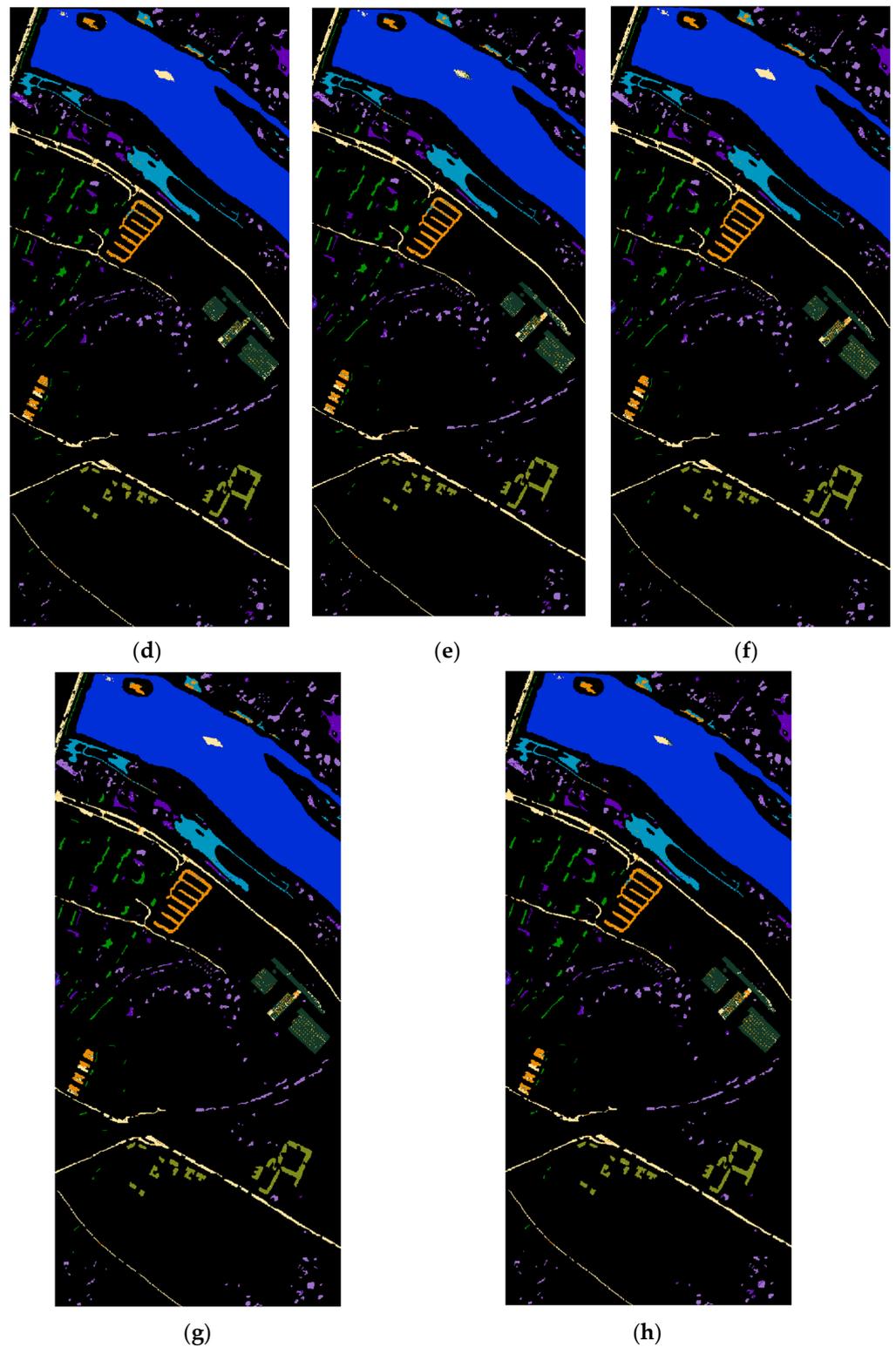
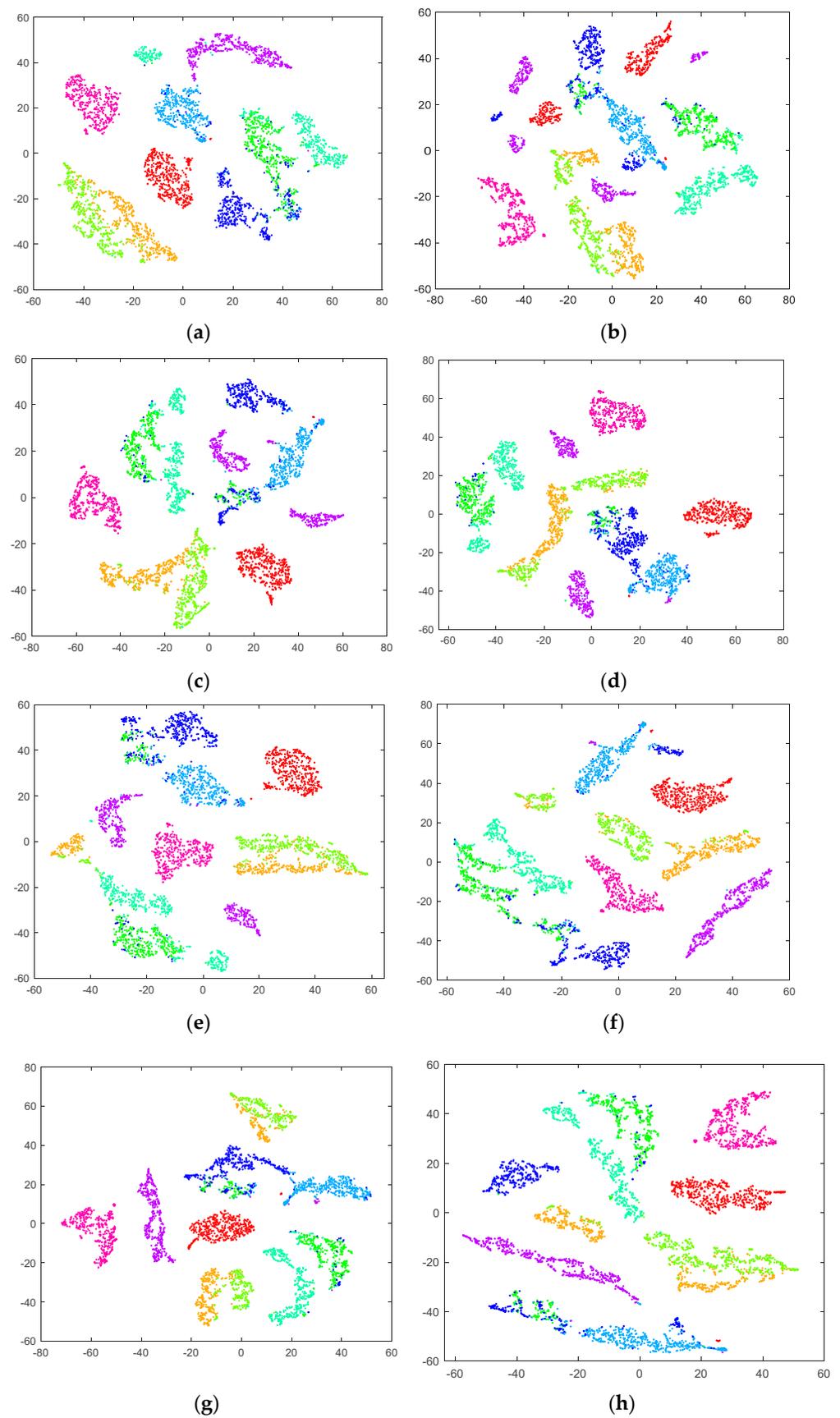


Figure 9. Cont.



**Figure 9.** The classified image of the Pavia Centre dataset using 20 training samples per class and 10 features; (a) BCC with  $\kappa = 92.29$ , (b) AE with  $\kappa = 92.36$ , (c) CBFEE with  $\kappa = 92.59$ , (d) MTD with  $\kappa = 92.70$ , (e) PFS with  $\kappa = 92.75$ , (f) HCBH with  $\kappa = 92.88$ , (g) WFE with  $\kappa = 93.48$ , and (h) FFE with  $\kappa = 92.97$ .



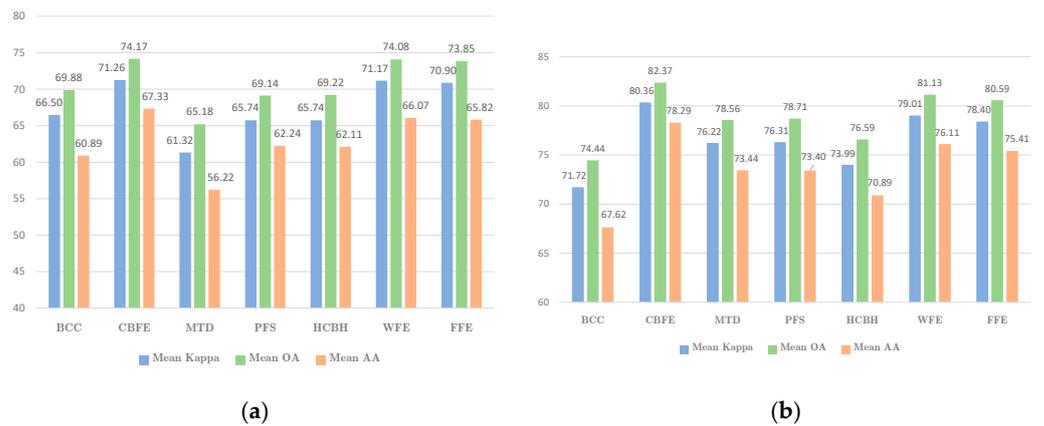
**Figure 10.** T-SNE results of Pavia Centre dataset using 18 features; each color represents a class in this dataset; (a) BCC, (b) AE, (c) CBF, (d) MTD, (e) PFS, (f) HCBH, (g) WFE, and (h) FFE.

**Table 8.** Classification accuracy, i.e., mean  $\kappa$ /OA/AA, of different dimension reduction methods using five training samples per class for the KSC dataset. The best results are highlighted in bold. Values of standard deviation are given in parentheses. Hyperbolic clustering-based band hierarchy (HCBH), clustering-based feature extraction (CBFE), band correlation clustering (BCC), autoencoder (AE), prototype space-based feature selection (PFS), maximum tangent discrimination (MTD), weighted feature extraction (WFE), and fuzzy feature extraction (FFE).

No. Features	Methods							
	BCC	AE	CBFE	MTD	PFS	HCBH	WFE	FFE
3	<b>64.49</b> (±0.00)/	61.27(±2.29)/	59.50(±0.00)/	39.65(±3.41)/	42.88(±9.97)/	27.87(±0.32)/	61.17(±1.41)/	62.52(±0.32)/
	<b>68.07</b> (±0.00)/	65.03(±2.11)/	63.56(±0.00)/	45.64(±3.17)/	48.31(±9.84)/	35.21(±0.01)/	65.05(±1.30)/	66.30(±0.29)/
	<b>56.83</b> (±0.00)	58.15(±3.79)	51.26(±0.00)	36.51(±4.70)	39.05(±8.68)	27.94(±0.49)	53.95(±0.99)	54.66(±0.25)
4	62.05(±0.00)/	61.59(±1.40)/	<b>71.71</b> (±0.00)/	46.64(±0.80)/	56.39(±2.89)/	27.80(±0.32)/	61.51(±1.17)/	61.13(±1.72)/
	65.84(±0.00)/	65.32(±1.29)/	<b>74.60</b> (±0.00)	52.01(±0.66)/	60.64(±2.68)/	35.15(±0.92)/	65.36(±1.07)/	65.03(±1.55)/
	54.32(±0.00)	58.22(±2.05)	<b>67.57</b> (±0.00)	42.92(±1.39)	50.74(±3.48)	27.88(±0.63)	54.29(±0.78)	53.40(±1.96)
5	62.06(±0.01)/	62.05(±2.14)/	69.83(±0.35)/	44.67(±1.12)/	56.45(±2.77)/	60.35(±0.49)/	<b>71.22</b> (±1.86)/	66.79(±3.66)/
	65.85(±0.01)/	65.78(±1.93)/	72.90(±0.31)/	49.78(±1.10)/	60.69(±2.58)/	64.46(±0.93)/	<b>74.14</b> (±1.67)/	70.14(±3.30)/
	54.33(±0.01)	58.89(±2.80)	<b>65.22</b> (±0.22)	39.91(±0.20)	50.82(±3.30)	52.62(±1.00)	64.70(±3.46)	59.91(±4.53)
6	62.07(±0.00)/	62.37(±2.05)/	<b>71.37</b> (±0.03)/	62.80(±0.36)/	61.23(±4.67)/	60.43(±0.46)/	70.38(±1.18)/	70.74(±2.09)/
	65.86(±0.00)/	66.04(±1.87)/	<b>74.28</b> (±0.03)/	66.54(±0.32)/	65.04(±4.22)/	64.54(±0.85)/	73.37(±1.07)/	73.70(±1.88)/
	54.35(±0.00)	60.00(±2.68)	<b>67.28</b> (±0.12)	55.12(±1.15)	56.67(±5.85)	52.68(±0.89)	63.98(±2.11)	65.39(±3.10)
7	65.23(±3.66)/	57.31(±9.34)/	71.40(±0.17)/	63.34(±1.11)/	58.61(±4.05)/	71.25(±0.46)/	71.51(±0.28)/	<b>71.92</b> (±0.80)/
	68.73(±3.32)/	61.47(±9.37)/	74.30(±0.15)/	67.03(±0.99)/	62.65(±3.73)/	74.15(±0.59)/	74.39(±0.25)/	<b>74.77</b> (±0.71)/
	59.00(±5.48)	54.49(±9.01)	67.58(±0.51)	56.19(±2.46)	53.28(±5.16)	<b>68.30</b> (±0.50)	65.98(±0.38)	67.11(±0.99)
8	65.78(±3.10)/	62.04(±3.94)/	71.49(±0.14)/	63.46(±1.54)/	65.29(±1.29)/	71.18(±0.34)/	<b>71.64</b> (±0.28)/	71.37(±0.44)/
	69.23(±2.82)/	65.76(±3.54)/	74.38(±0.12)/	67.14(±1.38)/	68.79(±1.20)/	74.10(±0.38)/	<b>74.51</b> (±0.25)/	74.28(±0.39)/
	59.66(±4.63)	59.95(±4.35)	<b>67.71</b> (±0.34)	55.92(±3.47)	62.24(±0.58)	68.21(±0.94)	66.25(±0.49)	66.60(±0.37)
9	65.67(±2.11)/	60.03(±7.86)/	72.15(±0.19)/	62.96(±0.50)/	67.17(±0.89)/	71.18(±0.55)/	<b>72.25</b> (±1.06)/	71.18(±0.14)/
	69.13(±1.93)/	64.01(±6.98)/	74.96(±0.17)/	66.69(±0.44)/	70.49(±0.80)/	74.10(±0.27)/	<b>75.07</b> (±0.95)/	74.11(±0.12)/
	59.92(±3.63)	56.66(±9.61)	<b>68.59</b> (±0.33)	54.61(±0.73)	65.05(±1.28)	68.21(±0.73)	66.95(±1.61)	66.37(±0.23)
10	68.67(±2.42)/	63.73(±1.79)/	<b>72.83</b> (±0.43)/	62.57(±2.74)/	67.74(±1.93)/	71.18(±0.43)/	71.38(±0.33)/	71.62(±0.41)/
	71.85(±2.20)/	67.28(±1.63)/	<b>75.57</b> (±0.38)/	66.35(±2.46)/	70.99(±1.73)/	74.10(±0.13)/	74.28(±0.30)/	74.50(±0.36)/
	64.07(±3.75)	60.51(±2.32)	<b>69.69</b> (±0.55)	57.42(±3.54)	65.47(±1.95)	68.21(±0.87)	65.66(±0.49)	66.99(±0.63)
11	66.79(±2.53)/	63.80(±1.62)/	72.74(±0.84)/	64.79(±2.25)/	70.61(±3.78)/	70.89(±0.48)/	<b>73.96</b> (±0.91)/	72.18(±0.31)/
	70.15(±2.30)/	67.36(±1.45)/	75.49(±0.75)/	68.35(±2.04)/	73.59(±3.41)/	73.77(±0.18)/	<b>76.61</b> (±0.82)/	75.01(±0.28)/
	61.47(±3.99)	60.92(±2.88)	69.42(±1.32)	60.07(±2.53)	67.81(±2.95)	68.92(±0.23)	<b>69.59</b> (±1.16)	67.75(±0.44)
12	67.93(±2.28)/	59.42(±7.92)/	73.35(±1.05)/	66.52(±1.56)/	70.57(±3.85)/	<b>75.39</b> (±0.46)/	72.93(±1.45)/	72.30(±0.24)/
	71.19(±2.07)/	63.47(±7.01)/	76.03(±0.94)/	69.90(±1.39)/	73.56(±3.47)/	<b>77.89</b> (±0.20)/	75.69(±1.31)/	75.11(±0.23)/
	63.17(±3.47)	55.12(±8.67)	70.26(±1.47)	61.93(±2.08)	67.88(±2.85)	<b>71.33</b> (±0.20)	68.15(±1.83)	67.92(±0.26)
13	68.09(±2.63)/	54.75(±9.54)/	71.58(±0.92)/	65.56(±0.14)/	72.87(±2.86)/	<b>75.39</b> (±0.54)/	73.56(±1.38)/	72.63(±0.07)/
	71.32(±2.38)/	59.31(±8.53)/	74.46(±0.82)/	69.04(±0.13)/	75.65(±2.56)/	<b>77.89</b> (±0.05)/	76.25(±1.24)/	75.41(±0.07)/
	63.41(±3.78)	50.80(±9.77)	67.72(±1.11)	60.61(±0.20)	69.70(±2.91)	<b>71.33</b> (±0.80)	69.28(±1.77)	68.30(±0.09)
14	69.28(±2.59)/	61.60(±6.39)/	72.32(±1.58)/	66.16(±2.24)/	73.96(±3.22)/	<b>75.39</b> (±0.69)/	74.03(±1.83)/	72.70(±0.46)/
	72.40(±2.35)/	65.45(±5.57)/	75.12(±1.42)/	69.58(±2.02)/	76.60(±2.90)/	<b>77.89</b> (±0.55)/	76.67(±1.65)/	75.48(±0.41)/
	64.94(±3.44)	58.39(±7.85)	68.67(±1.95)	61.83(±2.83)	70.45(±3.04)	<b>71.33</b> (±0.72)	70.49(±1.90)	68.28(±0.49)
15	68.59(±1.62)/	59.07(±9.16)/	72.21(±1.14)/	66.01(±1.81)/	73.15(±2.97)/	<b>75.39</b> (±0.46)/	74.12(±0.43)/	73.60(±1.06)/
	71.78(±1.46)/	63.10(±9.13)/	75.03(±1.02)/	69.43(±1.61)/	75.89(±2.68)/	<b>77.89</b> (±0.74)/	76.75(±0.40)/	76.29(±0.95)/
	64.33(±1.65)	56.70(±9.67)	68.53(±1.50)	61.64(±2.53)	69.72(±2.38)	<b>71.33</b> (±0.90)	70.19(±0.35)	69.32(±1.29)
16	68.88(±2.56)/	60.01(±7.31)/	72.35(±1.23)/	67.38(±2.97)/	71.59(±3.26)/	72.69(±0.47)/	73.22(±0.92)/	<b>73.86</b> (±0.95)/
	72.04(±2.32)/	63.97(±6.54)/	75.14(±1.12)/	70.66(±2.66)/	74.49(±2.94)/	75.46(±0.90)/	75.93(±0.84)/	<b>76.52</b> (±0.85)/
	64.36(±3.30)	56.24(±8.10)	68.91(±1.32)	63.20(±3.61)	69.04(±2.81)	68.50(±0.28)	69.20(±1.27)	<b>69.62</b> (±1.30)
17	68.94(±1.97)/	58.37(±8.83)/	72.34(±0.47)/	69.71(±2.65)/	71.48(±3.63)/	72.69(±0.41)/	72.41(±1.02)/	<b>75.12</b> (±0.40)/
	72.10(±1.78)/	62.53(±7.89)/	75.13(±0.41)/	72.75(±2.41)/	74.18(±3.27)/	75.46(±0.17)/	75.19(±0.92)/	<b>77.66</b> (±0.36)/
	64.79(±2.08)	55.09(±9.41)	69.05(±0.88)	66.06(±2.36)	69.02(±3.22)	68.50(±0.53)	68.89(±0.96)	<b>70.86</b> (±0.59)
18	69.43(±2.20)/	58.67(±7.71)/	73.02(±1.14)/	68.90(±1.86)/	71.79(±3.05)/	72.69(±0.38)/	73.37(±1.99)/	<b>74.70</b> (±0.86)/
	72.54(±1.99)/	62.62(±7.08)/	75.74(±1.03)/	72.01(±1.67)/	74.66(±2.76)/	75.46(±0.75)/	76.06(±1.80)/	<b>77.27</b> (±0.78)/
	65.33(±2.35)	57.59(±6.74)	69.80(±1.22)	65.56(±2.19)	68.92(±2.35)	68.50(±0.57)	69.52(±2.17)	<b>70.68</b> (±1.20)
Mean	66.50(±1.85)/	60.38(±5.71)/	<b>71.26</b> (±0.61)/	61.32(±1.69)/	65.74(±3.57)/	65.74(±0.45)/	71.17(±1.09)/	70.90(±0.87)/
	69.88(±1.68)/	64.28(±5.12)/	<b>74.17</b> (±0.54)/	65.18(±1.53)/	69.14(±3.24)/	69.22(±0.48)/	74.08(±0.99)/	73.85(±0.78)/
	60.89(±2.60)	57.36(±6.54)	<b>67.33</b> (±0.80)	56.22(±2.25)	62.24(±3.30)	62.11(±0.64)	66.07(±1.36)	65.82(±1.11)

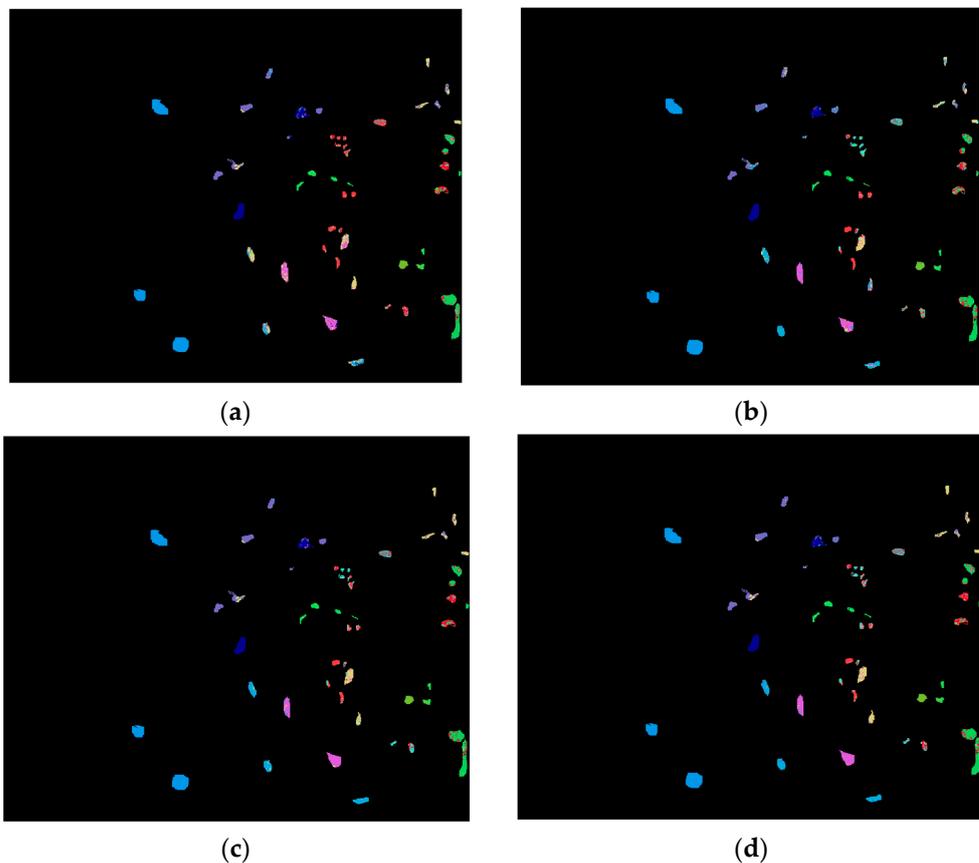
**Table 9.** Classification accuracy, i.e., mean  $\kappa$ /OA/AA, of different dimension reduction methods using 20 training samples per class for the KSC dataset. The best results are highlighted in bold. Values of standard deviation are given in parentheses. Hyperbolic clustering-based band hierarchy (HCBH), clustering-based feature extraction (CBFE), band correlation clustering (BCC), autoencoder (AE), prototype space-based feature selection (PFS), maximum tangent discrimination (MTD), weighted feature extraction (WFE), and fuzzy feature extraction (FFE).

No. Features	Methods							
	BCC	AE	CBFE	MTD	PFS	HCBH	WFE	FFE
3	65.72(±0.00)/	<b>72.53</b> (±2.64)/	66.92(±0.00)/	42.08(±0.02)/	65.13(±0.03)/	48.14(±0.98)/	63.66(±0.29)/	66.59(±0.61)/
	69.16(±0.00)/	<b>75.29</b> (±2.35)/	70.27(±0.00)/	47.33(±0.02)/	68.71(±0.03)/	53.30(±0.26)/	67.29(±0.27)/	69.96(±0.55)/
	59.86(±0.00)	<b>68.31</b> (±4.54)	60.81(±0.00)	39.20(±0.04)	59.24(±0.04)	42.21(±0.10)	58.52(±0.27)	<b>61.00</b> (±0.58)
4	66.94(±0.01)/	73.61(±1.41)/	<b>74.39</b> (±0.35)/	52.28(±9.87)/	65.17(±0.02)/	48.27(±0.85)/	65.00(±0.98)/	65.19(±0.82)/
	70.29(±0.01)/	76.26(±1.28)/	<b>76.91</b> (±0.00)/	56.74(±9.89)/	68.74(±0.02)/	53.38(±0.42)/	68.48(±0.87)/	68.69(±0.75)/
	60.72(±0.01)	70.50(±2.18)	<b>71.52</b> (±0.00)	48.50(±9.55)	59.32(±0.12)	42.55(±0.69)	59.71(±0.85)	59.68(±0.81)
5	66.89(±0.01)/	74.02(±0.62)/	<b>77.67</b> (±1.35)/	67.78(±5.39)/	65.17(±0.02)/	61.07(±0.80)/	76.60(±5.20)/	72.66(±4.42)/
	70.24(±0.01)/	76.62(±0.55)/	<b>80.21</b> (±1.15)/	70.98(±4.90)/	68.74(±0.02)/	64.90(±0.49)/	78.96(±4.68)/	75.43(±3.99)/
	60.64(±0.02)	71.04(±1.38)	<b>75.80</b> (±1.36)	62.49(±6.25)	59.32(±0.14)	56.34(±0.64)	73.20(±6.41)	68.09(±5.51)
6	66.94(±0.00)/	74.11(±1.05)/	<b>80.74</b> (±0.82)/	78.84(±0.78)/	72.47(±0.08)/	61.30(±0.72)/	75.51(±6.04)/	75.99(±2.33)/
	70.29(±0.00)/	76.70(±0.96)/	<b>82.16</b> (±0.32)/	80.99(±0.68)/	75.26(±0.07)/	65.10(±1.01)/	77.98(±5.44)/	78.43(±2.09)/
	60.73(±0.00)	71.40(±0.87)	<b>78.01</b> (±0.46)	75.76(±1.15)	69.81(±0.06)	56.58(±0.98)	71.92(±7.18)	72.27(±3.09)
7	68.91(±2.55)/	69.91(±9.41)/	80.53(±0.47)/	<b>80.71</b> (±0.57)/	76.52(±3.69)/	77.75(±0.79)/	79.81(±0.12)/	78.43(±1.93)/
	72.27(±2.53)/	72.93(±9.31)/	82.51(±0.15)/	<b>82.67</b> (±0.50)/	78.89(±3.30)/	79.96(±0.80)/	81.86(±0.10)/	80.62(±1.73)/
	63.91(±4.08)	66.87(±9.79)	<b>78.67</b> (±0.19)	77.34(±0.61)	74.09(±3.88)	76.01(±0.45)	77.06(±0.16)	75.31(±2.15)
8	71.78(±2.58)/	74.17(±1.47)/	<b>81.58</b> (±1.10)/	81.09(±0.25)/	77.08(±4.25)/	77.68(±0.82)/	80.64(±2.08)/	79.57(±1.08)/
	73.83(±1.97)/	76.74(±1.32)/	<b>83.61</b> (±0.95)/	83.01(±0.23)/	79.39(±3.81)/	79.90(±0.39)/	82.61(±1.87)/	81.65(±0.96)/
	67.27(±3.55)	72.26(±1.40)	<b>79.67</b> (±0.93)	77.75(±0.27)	74.61(±4.34)	75.92(±0.76)	77.41(±2.21)	76.53(±0.87)
9	73.13(±1.98)/	72.87(±5.29)/	<b>82.48</b> (±0.80)/	80.88(±0.08)/	78.07(±4.03)/	77.71(±0.75)/	81.57(±0.44)/	80.40(±1.41)/
	75.08(±1.67)/	75.61(±4.68)/	<b>84.26</b> (±0.60)/	82.82(±0.08)/	80.29(±3.61)/	79.92(±1.07)/	83.44(±0.39)/	82.39(±1.26)/
	68.98(±1.73)	70.21(±6.80)	<b>80.41</b> (±0.60)	77.62(±0.13)	75.41(±4.28)	75.94(±0.99)	78.78(±0.85)	77.50(±1.54)
10	70.62(±1.96)/	74.69(±1.42)/	<b>81.91</b> (±1.14)/	81.06(±0.18)/	78.54(±4.23)/	77.73(±0.73)/	81.42(±0.70)/	81.01(±0.61)/
	74.95(±1.63)/	77.22(±1.30)/	83.20(±0.93)/	82.99(±0.16)/	80.72(±3.79)/	79.94(±0.37)/	<b>83.31</b> (±0.63)/	82.93(±0.54)/
	68.96(±1.69)	72.40(±1.68)	<b>79.39</b> (±0.88)	78.23(±0.24)	75.92(±4.59)	75.97(±0.33)	78.98(±0.77)	78.17(±0.95)
11	72.20(±1.95)/	75.17(±0.36)/	<b>81.07</b> (±1.12)/	81.39(±0.33)/	77.74(±4.02)/	77.58(±0.84)/	<b>81.68</b> (±0.67)/	81.25(±0.67)/
	75.69(±1.93)/	77.63(±0.31)/	83.48(±0.90)/	83.28(±0.30)/	79.98(±3.61)/	79.82(±1.08)/	<b>83.54</b> (±0.60)/	83.14(±0.60)/
	69.56(±2.10)	73.11(±0.89)	<b>79.81</b> (±0.81)	78.76(±0.32)	75.53(±4.20)	75.74(±0.98)	79.24(±0.64)	78.47(±1.02)
12	74.33(±2.70)/	71.89(±7.07)/	81.83(±0.94)/	81.97(±0.51)/	79.93(±2.32)/	81.89(±0.78)/	<b>82.16</b> (±0.64)/	81.42(±0.29)/
	75.76(±2.44)/	74.71(±6.36)/	83.71(±0.56)/	83.79(±0.46)/	81.96(±2.09)/	83.72(±0.20)/	<b>83.97</b> (±0.58)/	83.30(±0.26)/
	70.17(±2.57)	69.45(±7.28)	<b>80.38</b> (±0.75)	79.58(±0.70)	77.77(±2.09)	78.96(±0.03)	79.79(±0.60)	78.84(±0.70)
13	73.86(±3.31)/	67.88(±5.81)/	81.80(±0.96)/	81.89(±0.31)/	79.96(±2.27)/	81.89(±0.77)/	<b>82.11</b> (±0.46)/	81.90(±0.53)/
	76.26(±2.59)/	71.09(±5.28)/	<b>84.04</b> (±0.41)/	83.72(±0.28)/	81.98(±2.04)/	83.72(±0.30)/	83.92(±0.42)/	83.73(±0.48)/
	70.61(±2.73)	65.62(±5.80)	<b>80.68</b> (±0.37)	79.65(±0.54)	77.82(±2.06)	78.97(±0.95)	79.69(±0.59)	79.82(±0.75)
14	73.16(±3.18)/	72.93(±4.43)/	82.13(±1.17)/	81.83(±0.43)/	80.93(±2.30)/	81.87(±0.80)/	<b>82.76</b> (±0.47)/	81.68(±0.28)/
	76.68(±3.18)/	75.66(±3.97)/	84.20(±0.73)/	83.66(±0.38)/	82.85(±2.08)/	83.70(±0.80)/	<b>84.51</b> (±0.43)/	83.53(±0.25)/
	71.15(±3.53)	70.57(±4.36)	<b>80.91</b> (±0.66)	79.68(±0.57)	78.95(±2.05)	78.95(±0.49)	80.64(±0.47)	79.66(±0.60)
15	75.48(±2.49)/	72.53(±6.44)/	<b>82.97</b> (±1.08)/	81.78(±0.26)/	81.18(±2.26)/	81.91(±0.89)/	82.86(±0.54)/	81.86(±0.28)/
	75.14(±2.67)/	75.29(±5.71)/	84.58(±0.92)/	83.62(±0.23)/	83.08(±2.04)/	83.74(±0.64)/	<b>84.60</b> (±0.48)/	83.69(±0.25)/
	69.53(±2.98)	70.60(±8.00)	<b>81.39</b> (±0.79)	79.75(±0.34)	79.24(±2.03)	79.02(±0.86)	80.90(±0.73)	80.18(±0.34)
16	75.26(±3.22)/	72.88(±2.17)/	<b>83.35</b> (±1.11)/	81.39(±0.15)/	81.19(±2.27)/	83.06(±0.73)/	82.49(±0.59)/	81.99(±0.40)/
	78.48(±2.91)/	75.59(±1.91)/	<b>85.04</b> (±0.95)/	83.27(±0.14)/	83.09(±2.04)/	84.77(±0.29)/	84.26(±0.54)/	83.81(±0.36)/
	73.11(±3.49)	71.01(±2.68)	<b>81.83</b> (±0.80)	79.62(±0.24)	79.26(±2.04)	80.39(±0.16)	80.34(±0.68)	80.25(±0.45)
17	75.41(±2.81)/	71.23(±6.08)/	<b>83.41</b> (±1.02)/	81.75(±0.48)/	81.03(±2.17)/	83.03(±0.75)/	82.83(±0.74)/	82.32(±0.37)/
	77.59(±3.17)/	74.08(±5.48)/	84.66(±1.21)/	83.59(±0.43)/	82.94(±1.96)/	<b>84.75</b> (±0.71)/	84.57(±0.67)/	84.11(±0.33)/
	72.24(±3.57)	69.66(±6.90)	<b>81.49</b> (±1.21)	80.02(±0.48)	79.12(±1.92)	80.37(±0.31)	80.72(±0.77)	80.52(±0.39)
18	76.88(±2.39)/	72.33(±4.53)/	82.96(±1.17)/	82.85(±0.43)/	80.84(±2.10)/	83.01(±0.75)/	<b>83.10</b> (±0.71)/	82.22(±0.54)/
	79.39(±1.01)/	75.07(±4.07)/	<b>85.07</b> (±0.99)/	84.58(±0.38)/	82.78(±1.89)/	84.73(±0.41)/	84.82(±0.64)/	84.02(±0.48)/
	74.39(±1.22)	71.08(±4.87)	<b>81.84</b> (±0.99)	81.17(±0.80)	78.99(±1.88)	80.35(±0.21)	80.85(±0.70)	80.30(±0.68)
Mean	71.72(±1.95)/	72.67(±3.76)/	<b>80.36</b> (±0.91)/	76.22(±1.44)/	76.31(±2.25)/	73.99(±0.80)/	79.01(±1.29)/	78.40(±1.03)/
	74.44(±1.73)/	75.41(±3.43)/	<b>82.37</b> (±0.67)/	78.56(±1.32)/	78.71(±2.02)/	76.59(±0.58)/	81.13(±1.16)/	80.59(±0.93)/
	67.62(±2.08)	70.26(±4.34)	<b>78.29</b> (±0.68)	73.44(±1.51)	73.40(±2.23)	70.89(±0.56)	76.11(±1.49)	75.41(±1.28)

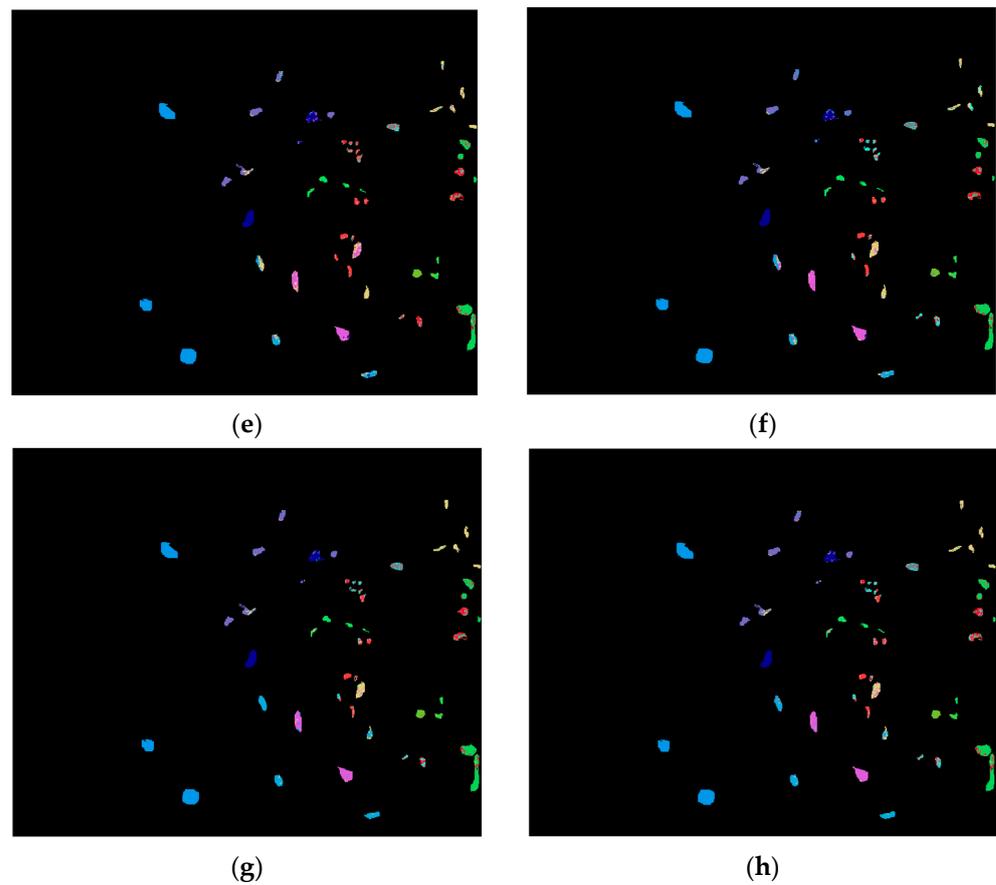


**Figure 11.** KSC dataset: classification accuracy for different methods, including the proposed models in this study, which are fuzzy feature extraction (FFE) and weighted feature extraction (WFE), and competing methods: hyperbolic clustering-based band hierarchy (HCBH), clustering-based feature extraction (CBFE), band correlation clustering (BCC), autoencoder (AE), prototype space-based feature selection (PFS), and maximum tangent discrimination (MTD); (a,b) illustrate the mean Kappa coefficient (multiplied by 100), OA and AA for the case of 5 and 20 training samples per class, respectively.

The mean STD values in the last row in Tables 8 and 9 reveal that while other methods have nearly identical stability, HCBH and CBFE exhibit higher stability with lower mean STD values. For example, according to Table 9, which shows the classification accuracy using 20 training samples per class, the HCBH STD values are 0.80, 0.58 and 0.56 for  $\kappa$ , OA and AA, respectively. The classified images for each method with 20 training samples per class and 10 features are depicted in Figure 12.

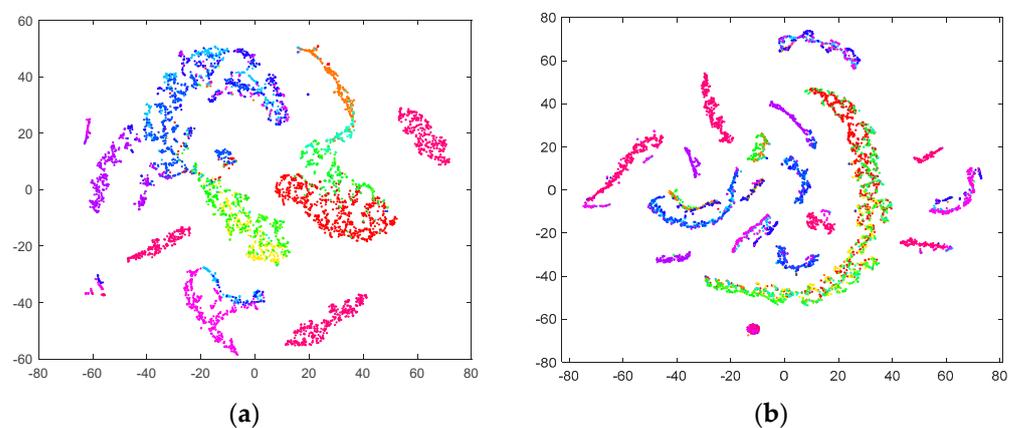


**Figure 12.** Cont.

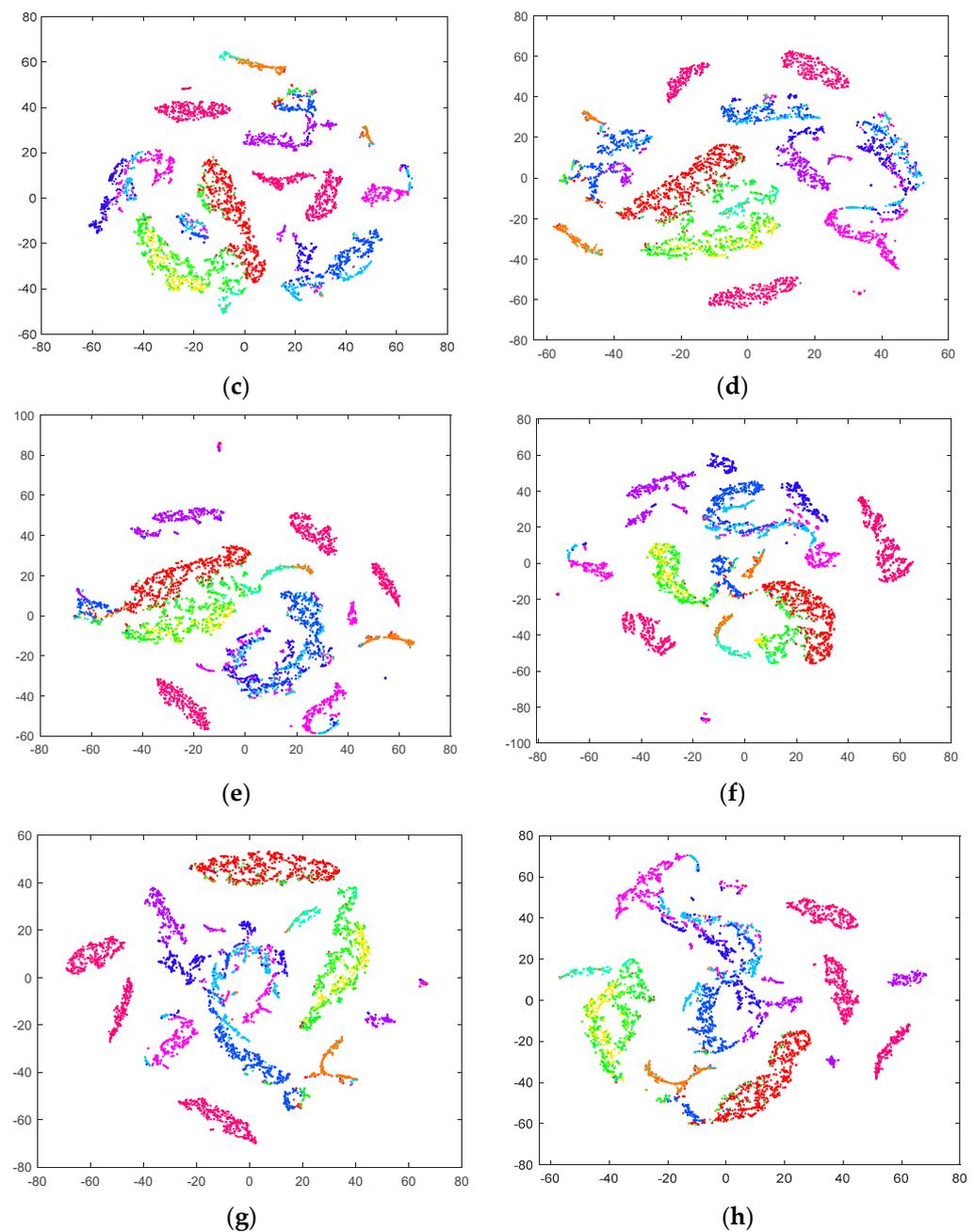


**Figure 12.** The classified images for the KSC dataset using 20 training samples per class and 10 features: (a) BCC with  $\kappa = 70.98$ , (b) AE with  $\kappa = 74.75$ , (c) CBF E with  $\kappa = 81.93$ , (d) MTD with  $\kappa = 80.81$ , (e) PFS with  $\kappa = 78.42$ , (f) HCBH with  $\kappa = 77.80$ , (g) WFE with  $\kappa = 81.73$ , and (h) FFE with  $\kappa = 81.27$ .

Furthermore, T-SNE served as a tool to visually assess the quality of the extracted or selected features (Figure 13). As illustrated in Figure 13, similar to the Pavia Centre dataset, it proves rather challenging to visually compare the T-SNE results due to the subtle differences. However, one noteworthy observation is that the sample points belonging to classes in CBF E, FFE, and WFE exhibit greater compactness than those in other methods. Consequently, this compactness facilitates the classifier to achieve higher accuracy during the image classification process (refer to Tables 8 and 9).



**Figure 13.** Cont.



**Figure 13.** T–SNE results of KSC dataset using 18 features; each color represents a class in this dataset; (a) BCC, (b) AE, (c) CBFE (d) MTD (e) PFS (f) HCBH (g) WFE, and (h) FFE.

#### 4. Conclusions

An HSI has numerous spectral bands, many of which are highly correlated and include high redundancy. As a result, a deterioration phenomenon known as the curse of dimensionality impacts the accuracy of supervised classifiers. In this study, we show that FE approaches could effectively be employed to mitigate this issue. We propose employing a simple but effective FE to decrease the high dimensionality of HSIs. Our unsupervised technique first estimates the number of existing endmembers in the input HSI. It then extracts endmembers using VCA. Using the extracted endmembers, the proposed method constructs a PS to cluster the comparable and correlated spectral bands' representatives in the PS. We finally use a weighted mean of existing bands in each cluster and extract a new feature. Our experiments demonstrate that our proposed technique outperforms five similar band clustering approaches for dimension reduction. Specifically, our technique

outperforms others when only a limited number of training samples are available to train the HSI classifier.

The proposed method only considers spectral information to reduce the high dimensionality of HSIs. Incorporating spatial information into the proposed methodology can be a topic for future research that will improve the quality of the extracted features. This could involve exploring spatial–spectral fusion techniques or integrating spatial context through the use of spatial filters. In addition, the robustness of the suggested FE approach against noise and outliers in hyperspectral data is another topic for future work. This may entail investigating strategies for improving its performance in the presence of sensor noise or atmospheric effects. Furthermore, integrating methodologies such as robust statistics or outlier identification might improve the method’s robustness, eventually leading to more reliable dimension reduction results in challenging circumstances.

**Author Contributions:** Conceptualization, S.H.A.M., F.K. and S.G.; methodology, S.H.A.M. and S.G.; software, S.H.A.M.; validation, S.H.A.M.; formal analysis, S.H.A.M. and F.K.; investigation, S.H.A.M.; writing—review and editing, S.H.A.M., F.K., S.G., M.A. and S.J.; visualization, S.H.A.M. and F.K.; supervision, S.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC). We express our sincere appreciation to the Government of Ontario for awarding us the Ontario Graduate Scholarship (OGS).

**Data Availability Statement:** For the sake of reproducibility, the source code for this study will be made available at [https://github.com/halizz821/unsupervised\\_dimension\\_reduction\\_HSI](https://github.com/halizz821/unsupervised_dimension_reduction_HSI).

**Acknowledgments:** We extend our sincere gratitude to He Sun for generously providing us with the source code of the HCBH method.

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

## References

1. Aghaee, R.; Mokhtarzade, M. Classification of hyperspectral images using subspace projection feature space. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 1803–1807. [[CrossRef](#)]
2. Moghaddam, S.A.; Mokhtarzade, M.; Moghaddam, S.A. A New Multiple Classifier System Based on a Pso Algorithm for the Classification of Hyperspectral Images. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *42*, 71–75. [[CrossRef](#)]
3. Moghaddam, S.H.A.; Mokhtarzade, M.; Beirami, B.A. A feature extraction method based on spectral segmentation and integration of hyperspectral images. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *89*, 102097.
4. Camps-Valls, G.; Marsheva, T.V.B.; Zhou, D. Semi-supervised graph-based hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 3044–3054. [[CrossRef](#)]
5. Fatemighomi, H.S.; Golarizadeh, M.; Amani, M. Object-based hyperspectral image classification using a new latent block model based on hidden Markov random fields. *Pattern Anal. Appl.* **2022**, *25*, 467–481. [[CrossRef](#)]
6. Seydi, S.T.; Shah-Hosseini, R.; Amani, M. A Multi-Dimensional Deep Siamese Network for Land Cover Change Detection in Bi-Temporal Hyperspectral Imagery. *Sustainability* **2022**, *14*, 12597. [[CrossRef](#)]
7. Shang, Y.; Zheng, X.; Li, J.; Liu, D.; Wang, P. A comparative analysis of swarm intelligence and evolutionary algorithms for feature selection in SVM-based hyperspectral image classification. *Remote Sens.* **2022**, *14*, 3019. [[CrossRef](#)]
8. Moeini Rad, A.; Abkar, A.A.; Mojaradi, B. Supervised distance-based feature selection for hyperspectral target detection. *Remote Sens.* **2019**, *11*, 2049. [[CrossRef](#)]
9. Bradley, P.E.; Keller, S.; Weinmann, M. Unsupervised feature selection based on ultrametricity and sparse training data: A case study for the classification of high-dimensional hyperspectral data. *Remote Sens.* **2018**, *10*, 1564. [[CrossRef](#)]
10. Xie, F.; Lei, C.; Li, F.; Huang, D.; Yang, J. Unsupervised hyperspectral feature selection based on fuzzy c-means and grey wolf optimizer. *Int. J. Remote Sens.* **2019**, *40*, 3344–3367. [[CrossRef](#)]
11. Landgrebe, D.A. *Signal Theory Methods in Multispectral Remote Sensing*; John Wiley & Sons: Hoboken, NJ, USA, 2003.
12. Serpico, S.B.; Moser, G. Extraction of spectral channels from hyperspectral images for classification purposes. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 484–495. [[CrossRef](#)]
13. Rasti, B.; Hong, D.; Hang, R.; Ghamisi, P.; Kang, X.; Chanussot, J.; Benediktsson, J.A. Feature extraction for hyperspectral imagery: The evolution from shallow to deep: Overview and toolbox. *IEEE Geosci. Remote Sens. Mag.* **2020**, *8*, 60–88. [[CrossRef](#)]
14. Kumar, B.; Dikshit, O.; Gupta, A.; Singh, M.K. Feature extraction for hyperspectral image classification: A review. *Int. J. Remote Sens.* **2020**, *41*, 6248–6287. [[CrossRef](#)]
15. Moghaddam, S.H.A.; Mokhtarzade, M.; Naeini, A.A.; Amiri-Simkooei, A. A statistical variable selection solution for RFM ill-posedness and overparameterization problems. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 3990–4001.

16. Haboudane, D.; Miller, J.R.; Pattey, E.; Zarco-Tejada, P.J.; Strachan, I.B. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sens. Environ.* **2004**, *90*, 337–352. [[CrossRef](#)]
17. De Backer, S.; Kempeneers, P.; Debruyn, W.; Scheunders, P. A band selection technique for spectral classification. *IEEE Geosci. Remote Sens. Lett.* **2005**, *2*, 319–323. [[CrossRef](#)]
18. Kuo, B.-C.; Landgrebe, D.A. Nonparametric weighted feature extraction for classification. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 1096–1105.
19. Richards, J.A.; Richards, J.A. *Remote Sensing Digital Image Analysis*; Springer: Berlin/Heidelberg, Germany, 2022; Volume 5.
20. Chen, Y.; Lin, Z.; Zhao, X.; Wang, G.; Gu, Y. Deep learning-based classification of hyperspectral data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 2094–2107. [[CrossRef](#)]
21. Hong, D.; Gao, L.; Yao, J.; Zhang, B.; Plaza, A.; Chanussot, J. Graph convolutional networks for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 5966–5978. [[CrossRef](#)]
22. Hong, D.; Han, Z.; Yao, J.; Gao, L.; Zhang, B.; Plaza, A.; Chanussot, J. SpectralFormer: Rethinking hyperspectral image classification with transformers. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5518615. [[CrossRef](#)]
23. Hong, D.; Gao, L.; Yokoya, N.; Yao, J.; Chanussot, J.; Du, Q.; Zhang, B. More diverse means better: Multimodal deep learning meets remote-sensing imagery classification. *IEEE Trans. Geosci. Remote Sens.* **2020**, *59*, 4340–4354. [[CrossRef](#)]
24. Sun, H.; Zhang, L.; Ren, J.; Huang, H. Novel Hyperbolic Clustering-based Band Hierarchy (HCBH) for Effective Unsupervised Band Selection of Hyperspectral Images. *Pattern Recognit.* **2022**, *130*, 108788. [[CrossRef](#)]
25. Ghorbanian, A.; Mohammadzadeh, A. An unsupervised feature extraction method based on band correlation clustering for hyperspectral image classification using limited training samples. *Remote Sens. Lett.* **2018**, *9*, 982–991. [[CrossRef](#)]
26. Rashwan, S.; Dobigeon, N. A split-and-merge approach for hyperspectral band selection. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1378–1382. [[CrossRef](#)]
27. Prabukumar, M.; Shrutika, S. Band clustering using expectation–maximization algorithm and weighted average fusion-based feature extraction for hyperspectral image classification. *J. Appl. Remote Sens.* **2018**, *12*, 046015. [[CrossRef](#)]
28. Lu, Q.; Huang, X.; Zhang, L. A novel clustering-based feature representation for the classification of hyperspectral imagery. *Remote Sens.* **2014**, *6*, 5732–5753. [[CrossRef](#)]
29. Imani, M.; Ghassemian, H. Band clustering-based feature extraction for classification of hyperspectral images using limited training samples. *IEEE Geosci. Remote Sens. Lett.* **2013**, *11*, 1325–1329. [[CrossRef](#)]
30. Bioucas-Dias, J.M.; Nascimento, J.M. Hyperspectral subspace identification. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 2435–2445. [[CrossRef](#)]
31. Remón, A.; Sánchez, S.; Bernabé, S.; Quintana-Ortí, E.S.; Plaza, A. Performance versus energy consumption of hyperspectral unmixing algorithms on multi-core platforms. *EURASIP J. Adv. Signal Process.* **2013**, *2013*, 68. [[CrossRef](#)]
32. Nascimento, J.M.; Dias, J.M. Vertex component analysis: A fast algorithm to unmix hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 898–910. [[CrossRef](#)]
33. Gerg, I. Open Source MATLAB Hyperspectral Toolbox. Available online: <https://github.com/isaacgerg/matlabHyperspectralToolbox> (accessed on 22 April 2015).
34. Hartigan, J.A.; Wong, M.A. Algorithm AS 136: A k-means clustering algorithm. *J. R. Stat. Soc. Ser. C* **1979**, *28*, 100–108.
35. Bezdek, J.C.; Ehrlich, R.; Full, W. FCM: The fuzzy c-means clustering algorithm. *Comput. Geosci.* **1984**, *10*, 191–203. [[CrossRef](#)]
36. Melgani, F.; Bruzzone, L. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 1778–1790. [[CrossRef](#)]
37. Chang, C.-C.; Lin, C.-J. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol. (TIST)* **2011**, *2*, 1–27. [[CrossRef](#)]
38. Liu, B.; Yu, X.; Zhang, P.; Yu, A.; Fu, Q.; Wei, X. Supervised deep feature extraction for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2017**, *56*, 1909–1921. [[CrossRef](#)]
39. Asl, M.G.; Mobasheri, M.R.; Mojaradi, B. Unsupervised feature selection using geometrical measures in prototype space for hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2013**, *52*, 3774–3787.

**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.