remote sensing

MDPI

Article

Hyperspectral Remote Sensing Images Feature Extraction Based on Spectral Fractional Differentiation

Jing Liu [1,*], Yang Li [1], Feng Zhao [2] and Yi Liu [3]

[1] School of Electronic Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710121, China; lyaa981030@stu.xupt.edu.cn
[2] School of Computer Science and Technology, Shandong Technology and Business University, Yantai 264005, China; zhaofeng1016@126.com
[3] School of Electronic Engineering, Xidian University, Xi'an 710071, China; yiliu@xidian.edu.cn
* Correspondence: jingliu@xupt.edu.cn

Abstract: To extract effective features for the terrain classification of hyperspectral remote-sensing images (HRSIs), a spectral fractional-differentiation (SFD) feature of HRSIs is presented, and a criterion for selecting the fractional-differentiation order is also proposed based on maximizing data separability. The minimum distance (MD) classifier, support vector machine (SVM) classifier, *K*-nearest neighbor (*K*-NN) classifier, and logistic regression (LR) classifier are used to verify the effectiveness of the proposed SFD feature, respectively. The obtained SFD feature is sent to the full connected network (FCN) and 1-dimensionality convolutional neural network (1DCNN) for deep-feature extraction and classification, and the SFD-Spa feature cube containing spatial information is sent to the 3-dimensionality convolutional neural network (3DCNN) for deep-feature extraction and classification. The SFD-Spa feature after performing the principal component analysis (PCA) on spectral pixels is directly connected with the first principal component of the original data and sent to $3DCNN_{PCA}$ and hybrid spectral net (HybridSN) models to extract deep features. Experiments on four real HRSIs using four traditional classifiers and five network models have shown that the extracted SFD feature can effectively improve the accuracy of terrain classification, and sending SFD feature to deep-learning environments can further improve the accuracy of terrain classification for HRSIs, especially in the case of small-size training samples.

Keywords: fractional differentiation; convolutional neural networks (CNNs); feature extraction; hyperspectral remote-sensing images (HRSIs)

check for updates

Citation: Liu, J.; Li, Y.; Zhao, F.; Liu, Y. Hyperspectral Remote Sensing Images Feature Extraction Based on Spectral Fractional Differentiation. *Remote Sens.* **2023**, *15*, 2879. https://doi.org/10.3390/rs15112879

Academic Editor: Salah Bourennane

Received: 25 April 2023
Revised: 23 May 2023
Accepted: 30 May 2023
Published: 1 June 2023

1. Introduction

Hyperspectral remote-sensing images (HRSIs) contain abundant spatial and spectral information simultaneously. The spectral dimension reveals the spectral curve characteristics of each pixel, while the spatial dimension reveals the spatial characteristics of the ground surface, and the organic fusion of spatial and spectral information is realized by HRSIs [1–3]. However, HRSIs have the characteristics of information redundancy and high dimensionality that bring difficulties and challenges to feature extraction and terrain classification [4,5].

For the feature extraction of HRSIs, the dimensionality reduction methods are usually utilized to project the HRSIs' spectral pixels to a low-dimensionality feature subspace [6,7]. Principal component analysis (PCA) and linear discriminant analysis (LDA) are representative approaches [8,9]. PCA calculates the covariance matrix of the original data, then, the eigenvectors corresponding to the first several largest eigenvalues of the covariance matrix are selected, and the original spectral pixels are projected to the orthogonal subspace supported by these eigenvectors to achieve the feature extraction and dimensionality reduction. LDA projects the original spectral pixels into a low-dimensional subspace, which has the

largest between-class scatters and the smallest within-class scatters so that the data have the best separability in the subspace.

In addition to reducing the dimensionality of HRSIs by feature extraction, discriminant features that can enhance the spectral differences of different terrains can also be achieved by other data analysis methods. Bao et al. have demonstrated that the derivatives of the spectral feature of HRSIs can capture the salient features of different land-cover categories, and have shown that in the case of small samples or poor data quality, combining the spectral first-order differentiation rather than second-order differentiation with the original spectral pixel can avoid the curse of dimensionality and improve the recognition rate [10]. Ye et al. extracted the spectral first-order differentiation in HRSIs and then used locality preserving nonnegative matrix factorization (LPNMF) and locality Fisher discrimination analysis (LFDA) to reduce the dimensionality of the original spectral pixel and spectral derivative, respectively, and, finally, performed feature fusion, which can effectively improve the classification performance [11].

At present, fractional differentiation is usually used in spectral analysis to estimate the contents of some elements or ions in soil or vegetation and is rarely used in spectral classification [12]. Lao et al. calculated the fractional differentiation of the soil spectral pixel in visible near-infrared spectroscopy to estimate the soil contents of salt and soluble ions [13]. Hong et al. used fractional differentiation to estimate soil organic carbon (SOC), wherein the spectral parameters derived from different spectral indices based on spectral fractional differentiation are combined to obtain the best estimation accuracy of SOC [14].

In recent years, convolutional neural networks have achieved remarkable results in the terrain classification of HRSIs [15]. Hu et al. applied a 1-dimensionality convolutional neural network (1DCNN) to HRSIs, which only used spectral information without considering spatial information [16]. Zhang et al. used PCA to reduce the dimensionality of spectral pixels and then used a 2-dimensionality convolutional neural network (2DCNN) for feature extraction and classification, which considered the spatial information of HRSIs [17]. To achieve full use of both the spectral and spatial information of HRSIs, Chen et al. used PCA to reduce the dimensionality of spectral pixels and sent the dimensionality-reduced data into a 3-dimensionality convolutional neural network (3DCNN), and, simultaneously, extracted the spatial and spectral deep features of HRSIs [18].

This paper uses fractional differentiation to perform feature extraction on the pixel spectral curves of HRSIs from the aspect of data analysis, because fractional differentiation can retain part of the original characteristics of the data while obtaining the characteristics that express the differences in the data, and the order of the fractional differentiation can change with the different data. In this paper, a spectral fractional-differentiation (SFD) feature of HRSIs is presented, and a criterion for selecting the fractional-differentiation order is also proposed based on maximizing data separability. The minimum distance (MD) classifier, support vector machine (SVM) classifier, *K*-nearest neighbor (*K*-NN) classifier, and logistic regression (LR) classifier are used to verify the effectiveness of the proposed SFD feature, respectively. The obtained SFD feature is sent to the full connected network (FCN) [19] and 1DCNN for deep-feature extraction and classification, and the SFD-Spa feature cube containing spatial information is sent to 3DCNN for deep-feature extraction and classification. The SFD-Spa feature after performing PCA on spectral pixels is directly connected with the first principal component of the original data and sent to $3DCNN_{PCA}$ and hybrid spectral net (HybridSN) [20] models to extract deep features. Compared with integer-order differentiation, the advantage of fractional-order differentiation is that it has memory and globality. When the order of the integer differentiation is just larger than that of fractional differentiation, fractional-order differentiation can preserve more low-frequency components of the signal, while the high- and middle-frequency components are also obviously enhanced [21]. The advantages of the presented HRSIs SFD feature are as follows:

(1) The presented SFD feature preserves both the overall curve shape and local burrs characteristics of the pixel spectral curves of HRSIs, which is very suitable for ter-

rain classification. The overall curve shapes of spectral curves correspond to the low-frequency components and the local burrs correspond to the high-frequency components of the pixel spectral curve. For HRSIs terrain classification, the shape characteristics of spectral curves contribute most to the discriminant of quite different terrains, such as water, soil, and plants; while the local burr characteristics contribute most to the identification of the different terrains which have similar spectral curves, such as wheat and soybean. These two characteristics are both important, however, the integer differentiation invariably enhanced the high-frequency components, i.e., local burrs while losing most of the low-frequency components, i.e., the shape characteristics of spectral curves. Fractional differentiation preserves the low-frequency components sufficiently while amplifying the high-frequency components remarkably, thus, the presented SFD feature contains both the overall curve shape and local burr characteristics of the spectral curves;

(2) The order of the fractional differentiation of the presented SFD feature can be selected by achieving the best separability. With the increase in the differentiation order, the shape characteristics of the original spectral curve are less preserved, while the local burr characteristics are enhanced more significantly. In view of such character, a criterion for selecting the appropriate fractional-differentiation order is presented based on achieving the best data separability, which guarantees that the overall curve shape characteristics and the local burr characteristics are properly preserved in the resulting SFD feature, such that the quite different terrains and the similar terrains are all easy to identify.

Experimental results on four real HRSIs using four traditional classifiers and five network models have shown that the extracted SFD feature can effectively improve the terrain classification accuracy, and sending the SFD feature to deep-learning environments can further improve the terrain classification accuracy, especially in the case of small-size training samples [22,23].

2. Spectral Fractional-Differentiation (SFD) Feature

2.1. Fractional Differentiation

Among the many definitions of fractional differentiation, the commonly used three forms are Riemann–Liouville, Grümwald–Letnikvo, and Caputo [24]. In this paper, the Grümwald–Letnikvo definition is used to generalize the differentiation of continuous functions from integer order to fractional order, and the fractional-order differential expression is deduced by using the difference equation of integer-order differentiation.

According to the definition of integer-order differentiation, for a differentiable function $f(x)$, its g-th integer-order differentiation is

$$f^{(g)}(x) = \lim_{h\to 0}\frac{1}{h^g}\sum_{j=0}^{n}(-1)^j\binom{g}{j}f(x-jh) \tag{1}$$

where $g \in \mathbb{N}$, the binomial coefficient $\binom{g}{j} = \frac{\Gamma(g+1)}{\Gamma(j+1)\Gamma(g-j+1)} = \frac{g!}{j!(g-j)!}$, $\Gamma(\cdot)$ is Gamma function, $\Gamma(x) = \int_0^\infty e^{-t}t^{x-1}dt$, and h represents the differential step size. Extending the order g to any real number v, the Grümwald–Letnikvo differentiation of $f(x)$ is defined as

$$\begin{aligned} {}_aD_x^v f(x) &= \lim_{h\to 0}\frac{1}{h^v}\sum_{j=0}^{[(x-a)/h]}(-1)^j\binom{v}{j}f(x-jh) \\ &= \lim_{h\to 0}\frac{1}{h^v}\sum_{j=0}^{[(x-a)/h]}(-1)^j\frac{\Gamma(v+1)}{\Gamma(j+1)\Gamma(v-j+1)}f(x-jh) \end{aligned} \tag{2}$$

where a represents the lower limit of $f(x)$, and $[(x-a)/h]$ represents taking the integer part of $(x-a)/h$ [25].

Fractional differentiation defined in Equation (2) is a generalization of integer differentiation. When the order v is a positive integer, Equation (2) still holds, thus, the integer-order differentiation can be regarded as a special case of the fractional-order differentiation. Fractional differentiation is different from integer differentiation in numerical computing. During the calculation of the integer-order differentiation, the differential result of a point is only related to the information of the nearby points and unrelated to the information of other points; when the fractional differentiation is calculated, the differential result of a point is related to the information of all points before that point, and the points closer to it have greater weights in the calculation, thus, results in that fractional differentiation have memory and globality [25].

2.2. Spectral Fractional-Differentiation (SFD) Feature

The classical integer-order differentiation is a tool to describe the characteristics of the Euclidean space samples and is often utilized for signal extraction and singularity detection in signal analysis and processing. Fractional differentiation is a generalization of integer differentiation. Pu has pointed out that when the fractional-differentiation operation of a signal is performed, the high-frequency and middle-frequency components of the signal will be greatly improved, while the low-frequency components are retained nonlinearly [26]; and with the increase in the differentiation order, the improvement of the high-frequency and middle-frequency components will be enhanced, but fewer low-frequency components will be preserved [27]. In this paper, fractional differentiation is performed on the spectral pixel of HRSIs, and the resulting spectral fractional-differentiation (SFD) feature is used for terrain classification. Since the definition of Grümwald–Letnikvo fractional differentiation is generalized from the definition of integer differentiation and is expressed in discrete form, which is convenient for numerical calculation. Therefore, the presented SFD feature is defined according to the Grümwald–Letnikvo formula.

For a unary function $f(x)$, let the differential step $h = 1$, then the expression of the v-th order fractional differentiation of $f(x)$ is

$$f^{(v)}(x) = f(x) + (-v)f(x-1) + \frac{(-v)(-v+1)}{2}f(x-2) + \cdots + \frac{\Gamma(-v+1)}{n!\Gamma(-v+t+1)}f(x-t) \tag{3}$$

which has $(t + 1)$ terms.

For HRSIs, a spectral pixel can be regarded as a discrete form of a unary function. Assuming that each pixel has N spectral bands, for a spectral pixel $\boldsymbol{x} = [x_0, x_1, x_2, \cdots, x_{N-1}]$, the v-th order fractional-differentiation vector of $\boldsymbol{x}$, i.e., the presented SFD feature, is

$$\boldsymbol{x}^{(v)} = [a_0x_1 + a_1x_0, a_0x_2 + a_1x_1 + a_2x_0, \cdots, a_0x_{N-1} + a_1x_{N-2} + a_2x_{N-3} + \cdots + a_{N-1}x_0] \tag{4}$$

where $a_0, a_1, \cdots, a_{N-1}$ are the first $(N-1)$ coefficients on the right side of Equation (3) and

$$\begin{cases} a_0 = 1 \\ a_1 = -v \\ a_2 = [(-v)(-v+1)]/2 \\ a_3 = [(-v)(-v+1)(-v+2)]/6 \\ \cdots \\ a_{N-1} = \Gamma(-v+1)/[(N-1)!\Gamma(-v+N)] \end{cases} \tag{5}$$

The dimensionality of the SFD feature $\boldsymbol{x}^{(v)}$ is $(N-1)$, and the components of $\boldsymbol{x}^{(v)}$ correspond to the v-th order fractional differentiation of bands $x_1, x_2, \cdots, x_{N-1}$. In particular, when the order v equals 1, the expression of spectral fractional differentiation is the same as that of the first-order differentiation.

2.3. Criterion for Selecting Optimal Fractional-Differentiation Order

In the terrain classification of HRSIs based on the spectral pixel, when there only exist quite different terrains, such as water, soil, and plants, the overall curve shape characteristics of spectral curves, which correspond to the low-frequency components, contribute most

to the discriminant. However, the phenomenon of different subjects with similar spectra commonly exists in HRSIs scenes [28,29]; in this case, the local burrs characteristics, which correspond to the high-frequency components, contribute most to the identification. In real HRSIs scenes, the phenomena of different objects with quite different spectra and different subjects with similar spectra both exist. This required that the feature extracted from the spectral curve should properly contain the low-frequency and high-frequency components simultaneously.

Shown in Figure 1 is the amplitude spectrum of the SFD feature of the corn-no-till class in the Indian Pines dataset, where the fractional-differentiation order v varies from 0 to 1.6 with step 0.4, the amplitude spectrum is taken logarithm for observation and the bases of the logarithm is 2. Figure 1 shows that as the fractional-differentiation order increases, the low-frequency components of the amplitude spectra decrease, while the mid-frequency and high-frequency components significantly increase. It can be concluded that the low-order SFD feature can enhance the high-frequency components while sufficiently retaining the low-frequency components of the spectral pixel, which is beneficial for preserving both the overall curve shape characteristics and the local burrs characteristics. However, how to select an appropriate differentiation order is a problem worth considering.

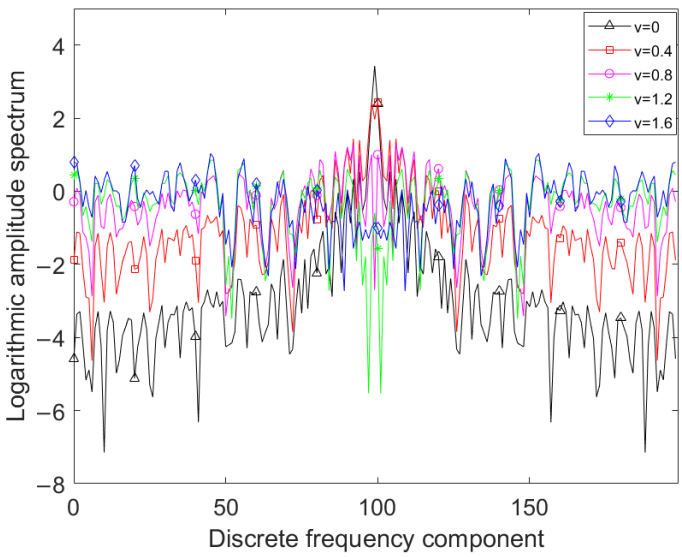


Figure 1. Amplitude spectra of SFDs with different fractional-differentiation orders.

To study how the presented SFD feature is influenced by the fractional-differentiation order and to select the appropriate SFD order, the spectral curve of the corn-no-till class in the Indian Pines dataset is selected to extract the SFD feature with the fractional-differentiation order varies from 0 to 1.9 at step 0.1, thus, a total of 20 SFD curves are obtained, as shown in Figure 2.

It can be seen intuitively from Figure 2 that, as the differentiation order increases, the SFD values corresponding to the slowly changing parts of the original spectral curve gradually approach 0, and the SFD values corresponding to the local sharply changing parts dramatically increase. When the differentiation order increases from 0 to 1, the SFD curves still retain lots of shape characteristics of the original spectral curve, and the local sharp characteristics are enhanced. When the differentiation order increases from 1 to 1.9, the SFD curves lost most of the shape characteristics of the original spectral curve, and the local sharp characteristics are further enhanced. Therefore, for HRSIs terrain classification, when the differentiation order $0 < v < 1$, the presented SFD feature contains the discriminant information benefit for classifying the different objects with quite different spectra and different subjects with similar spectra simultaneously and is very suitable for real HRSIs scenes. However, how to achieve more precise ranges of appropriate differentiation orders

for different HRSIs is a problem worth further considering. In this paper, a criterion for selecting the SFD order is proposed based on maximizing the separability.

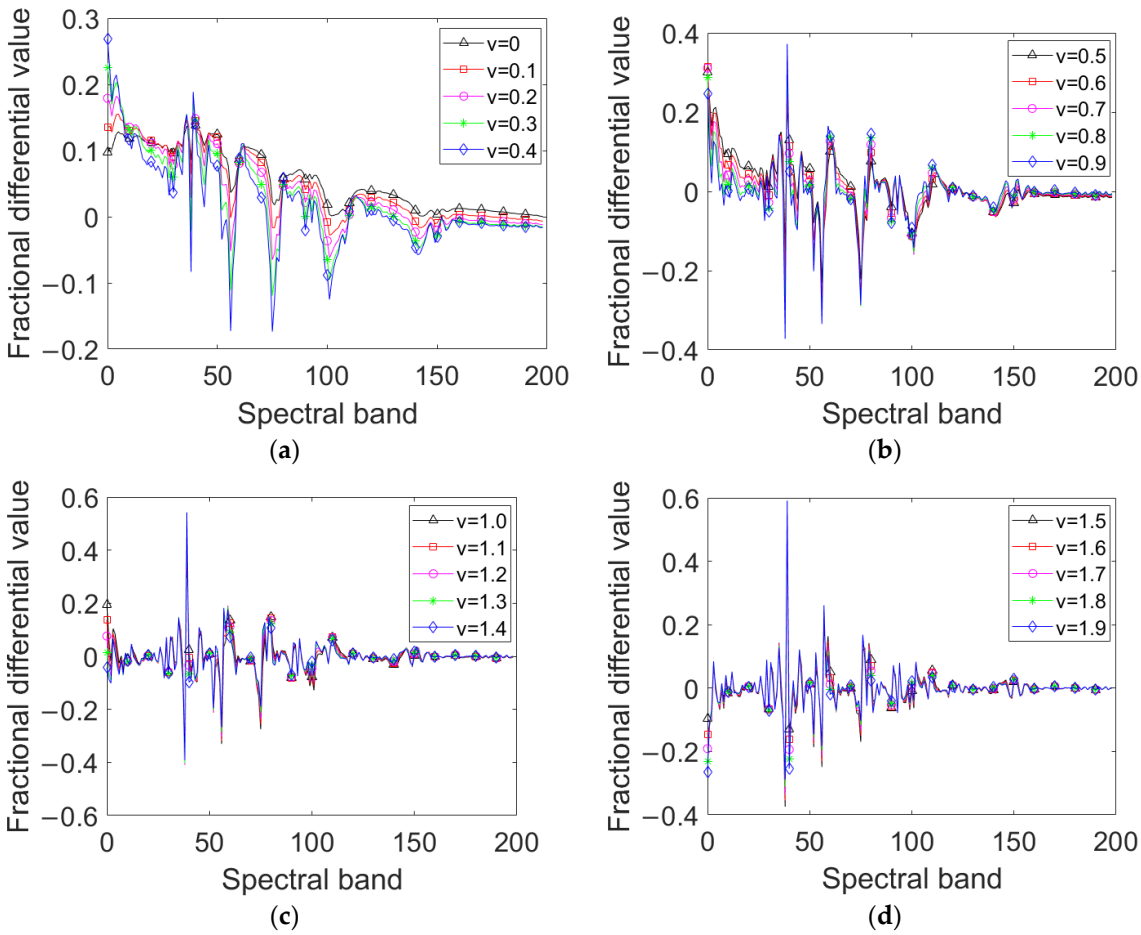


Figure 2. SFD curves of corn-no-till with fractional-differentiation order varying from 0 to 1.9: (**a**) v = 0~0.4; (**b**) v = 0.5~0.9; (**c**) v = 1~1.4; (**d**) v = 1.5~1.9.

Assuming that the number of classes is C, let v denote the order of SFD, and the dimensionality of the SFD feature is $(N-1)$, the within-class scatter matrix $\boldsymbol{S}_w^{(v)}$ and the between-class scatter matrix $\boldsymbol{S}_b^{(v)}$ in the $(N-1)$-dimensionality SFD feature space are

$$\boldsymbol{S}_w^{(v)} = \sum_{i=1}^{C} P_i \frac{1}{n_i} \sum_{k=1}^{n_i} \left(\boldsymbol{x}_{ik}^{(v)} - \boldsymbol{m}_i^{(v)}\right)\left(\boldsymbol{x}_{ik}^{(v)} - \boldsymbol{m}_i^{(v)}\right)^{\mathrm{T}} \tag{6}$$

and

$$\boldsymbol{S}_b^{(v)} = \sum_{i=1}^{C} P_i \left(\boldsymbol{m}_i^{(v)} - \boldsymbol{m}^{(v)}\right)\left(\boldsymbol{m}_i^{(v)} - \boldsymbol{m}^{(v)}\right)^{\mathrm{T}} \tag{7}$$

, respectively, where n_i represents the number of samples of class i, $\boldsymbol{x}_{ik}^{(v)}$ represents the v-th order fractional differentiation of the k-th sample of class i, P_i represents the prior probability of class i, $\boldsymbol{m}_i^{(v)}$ represents the mean of the v-th order fractional differentiations of class i, and $\boldsymbol{m}^{(v)} = \sum_{i=1}^{C} P_i \boldsymbol{m}_i^{(v)}$ represents the overall mean of the v-th order fractional differentiations.

The presented criterion for optimizing SFD order is

$$J(v) = Tr\left(\boldsymbol{S}_b^{(v)}\right) - Tr\left(\boldsymbol{S}_w^{(v)}\right), \tag{8}$$

where "*Tr*()" represents the trace of a matrix. The principle of the SFD order selecting criterion is that the data separability should be maximized in the SFD feature space. $Tr\left(\boldsymbol{S}_b^{(v)}\right)$ measures the variance of the class means in the v-th order SFD feature space, the larger $Tr\left(\boldsymbol{S}_b^{(v)}\right)$ is, the greater the between-class separability is. $Tr\left(\boldsymbol{S}_w^{(v)}\right)$ measures the within-class divergence in the v-th order SFD feature space, the smaller the $Tr\left(\boldsymbol{S}_w^{(v)}\right)$ is, the smaller the within-class divergence is. Therefore, $J(v)$ evaluates the data separability in the v-th order SFD feature space, by maximizing J, the data separability in the SFD feature space is maximized, thus, the optimal SFD order is

$$v^* = \underset{v}{\operatorname{argmax}} J(v). \tag{9}$$

Shown in Figure 3 are the variations of criterion J with the SFD order v on 4 real HRSIs datasets.

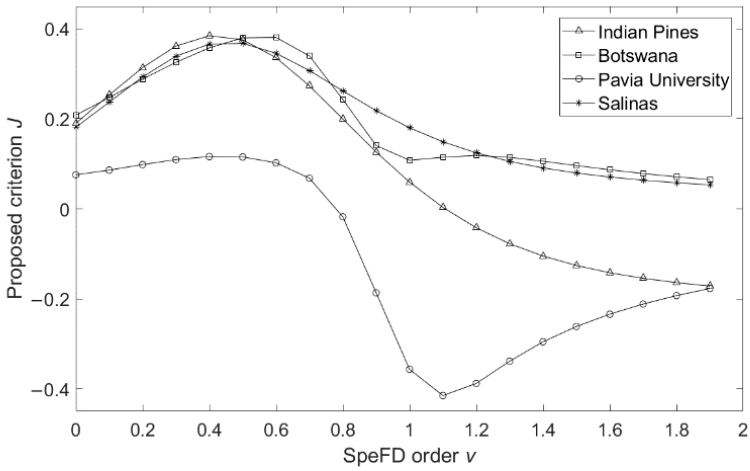


Figure 3. Comparison of variations of J with SFD order v.

Figure 3 shows that, when the SFD order $0 < v < 1$, Botswana, Indian Pines, and Salinas datasets have the same variation trend, they all have obvious peaks as $0.4 \leq v \leq 0.6$, as the SFD order v increases, the criterion J becomes smaller, which means smaller data separability, this is consistent with the analysis of Figure 1. For the Pavia University dataset, criterion J has an obvious peak as $0.5 \leq v \leq 0.7$ and then decreases with the increase in v, and an inflection point occurs at $v = 1.1$, but the general trend is still the same as the other three datasets. Therefore, it is confirmed again that the SFD order v ranges between 0 and 1 is more conducive to improving the classification accuracy of HRSIs, and the precise appropriate SFD order range for each dataset is given.

3. Networks Structure and Parameter Settings

To further extract deep features, five network models are used for deep-feature extraction and terrain classification. The five network models used are fully connected network (FCN), one-dimensional convolutional neural network (1DCNN), three-dimensional convolutional neural network (3DCNN), three-dimensional convolutional neural network after spectral PCA dimensionality reduction ($3DCNN_{PCA}$), and hybrid spectral net (HybridSN). Tables 1 and 2 show the parameters and the number of output feature maps for each layer of the networks, respectively. N represents the dimension of the input dataset. C represents the number of classes. *I*, *Conv*, *Po*, and *FC* represent the input layer, convolutional layer, pooling layer, and fully connected layer, respectively. For example, *Conv6* indicates that this layer is a convolutional layer located in the sixth layer of the network structure. "$\surd$" means there exists an *FC* layer. <*> represents rounding up the calculation result.

Table 1. Parameter settings for five network models.

	FCN	**1DCNN**	**3DCNN**	**3DCNN$_{PCA}$**	**HybridSN**
I1	$1 \times N$	$1 \times N$	$1 \times N$	$11 \times 11 \times N$	$25 \times 25 \times N$
Conv2	-	<$1 \times N/9$>	$7 \times 7 \times 3$	$3 \times 3 \times 7$	$3 \times 3 \times 7$
Po3	-	$1 \times$ <<$N/9$>/5>	-	-	-
Conv4	-	-	$3 \times 3 \times 3$	$3 \times 3 \times 5$	$3 \times 3 \times 5$
Po5	-	-	-	-	-
Conv6	-	-	-	$3 \times 3 \times 3$	$3 \times 3 \times 3$
Po7	-	-	-	-	-
Conv8	-	-	-	$3 \times 3 \times 3$	3×3
Po9	-	-	-	-	-
FC1	√	√	√	√	√
FC2	√	√	-	√	√
FC3	√	-	-	√	√
FC4	√	-	-	-	-

Table 2. Number of output feature maps for five network models.

	FCN	**1DCNN**	**3DCNN**	**3DCNN$_{PCA}$**	**HybridSN**
I1	1	1	1	1	1
Conv2	-	20	2	8	8
Po3	-	20	-	-	-
Conv4	-	-	4	16	16
Po5	-	-	-	-	-
Conv6	-	-	-	32	32
Po7	-	-	-	-	-
Conv8	-	-	-	64	64
Po9	-	-	-	-	-
FC1	2048	100	*C*	256	256
FC2	4096	*C*	-	128	128
FC3	2048	-	-	*C*	*C*
FC4	*C*	-	-	-	-

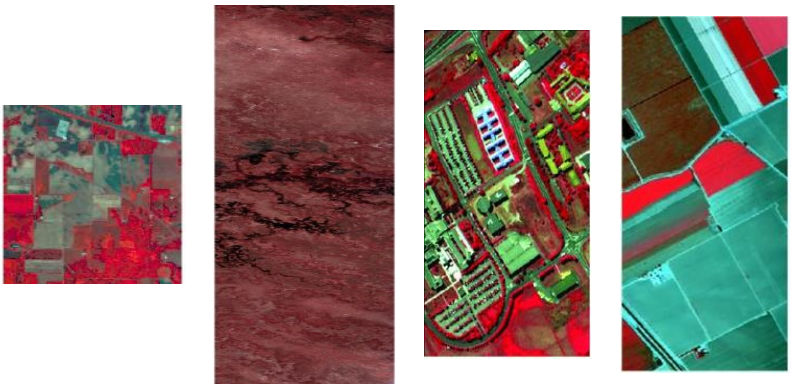
4. Experimental Results

Firstly, using the proposed criterion J to select the appropriate SFD order for each dataset, the selected SFD order for Indian Pines, Botswana, Pavia University, and Salinas are 0.6, 0.3, 0.6, and 0.4, respectively. Additionally, then perform fractional differentiation on the pixel spectral curves with the selected order and achieve the SFD feature. Four traditional classifiers and five networks are used to verify the effectiveness of the resulting SFD feature. Among the five network models, the inputs of FCN and 1DCNN models are SFD feature vectors without spatial information, while the inputs of 3DCNN, 3DCNN$_{PCA}$, and HybridSN contain spatial information. The input of 3DCNN is the SFD-Spa feature cube, and the input of 3DCNN$_{PCA}$ and HybridSN is the data cube by connecting the SFD-Spa feature after PCA with the first principal component of the original data. To unify the forms, the experimental results on five network models are all represented by "SFD".

4.1. Experimental Datasets

Four real HRSIs, namely, Indian Pines, Botswana, Pavia University, and Salinas, are used for the experiments. The Indian Pines dataset includes 16 classes, and the image size is 145×145, and a total of 10,249 pixels can be used to classify. After removing the bands 104–108, 150–163, and 220 affected by noise factors, 200 bands were finally left for the experiment. The Botswana dataset was obtained by NASA's EO-1 satellite in the Botswana area. 14 terrain classes are included, and the image size is 1476×256, and 3248 of them are terrain pixels. After removing the bands affected by noise and water vapor, the bands 10–55, 82–97, 102–119, 134–164, and 187–220 were retained, i.e., a total of 145 bands were

finally selected. Pavia University dataset contains 9 classes, the image size is 610 × 340, including 42,776 terrain pixels, and 103 bands were finally selected. The Salinas dataset has 16 classes, and the image size is 512 × 217. Bands 108–112, 154–167, and 220 were affected by noise and water vapor and were removed. 204 bands are reserved for research, and a total of 54,129 pixels can be used for terrain classification. Table 3 shows the specific sampling results of the experimental data. Figures 4 and 5 show the false-color image and ground truth of these datasets.



(**a**) (**b**) (**c**) (**d**)

Figure 4. The false-color image of four HRSIs datasets: (**a**) Indian Pines; (**b**) Botswana; (**c**) Pavia University; (**d**) Salinas.

Table 3. Category number, name, and sample number of each dataset.

Category Number	Category Name	Sample Number	Category Number	Category Name	Sample Number
			Indian Pines		
1	Alfalfa	46	9	Oats	20
2	Corn-notill	1428	10	Soybean-notill	972
3	Corn-mintill	830	11	Soybean-mintill	2455
4	Corn	237	12	Soybean-clean	593
5	Grass-pasture	483	13	Wheat	205
6	Grass-trees	730	14	Woods	1265
7	Grass-pasture-mowed	28	15	Building-Grass-Trees-Drives	386
8	Hay-windrowed	478	16	Stone-Steel-Towers	93
			Botswana		
1	Water	270	8	Island interior	203
2	Hippo grass	101	9	Acacia woodlands	314
3	Floodplain grass1	251	10	Acacia shrub lands	248
4	Floodplain grass2	215	11	Acacia grasslands	305
5	Reeds1	269	12	Short mopani	181
6	Riparian	269	13	Mixed mopani	268
7	Firescar2	259	14	Exposed soils	95
			Pavia University		
1	Asphalt	6631	6	Bare Soil	5029
2	Meadows	18,649	7	Bitumen	1330
3	Gravel	2099	8	Self-Blocking Bricks	3682
4	Tress	3064	9	Shadows	947
5	Painted metal sheets	1345			

Table 3. *Cont.*

Category Number	Category Name	Sample Number	Category Number	Category Name	Sample Number
			Salinas		
1	Brocoli_green_weeds_1	2009	9	Soil_vinyard_develop	6203
2	Brocoli_green_weeds_2	3726	10	Corn_senesced_green_weeds	3278
3	Fallow	1976	11	Lettuce_romaine_4wk	1068
4	Fallow_rough_plow	1394	12	Lettuce_romaine_5wk	1927
5	Fallow_smooth	2678	13	Lettuce_romaine_6wk	916
6	Stubble	3959	14	Lettuce_romaine_7wk	1070
7	Celery	3579	15	Vinyard_untrained	7268
8	Grapes_untrained	11,271	16	Vinyard_vertical_trellis	1807

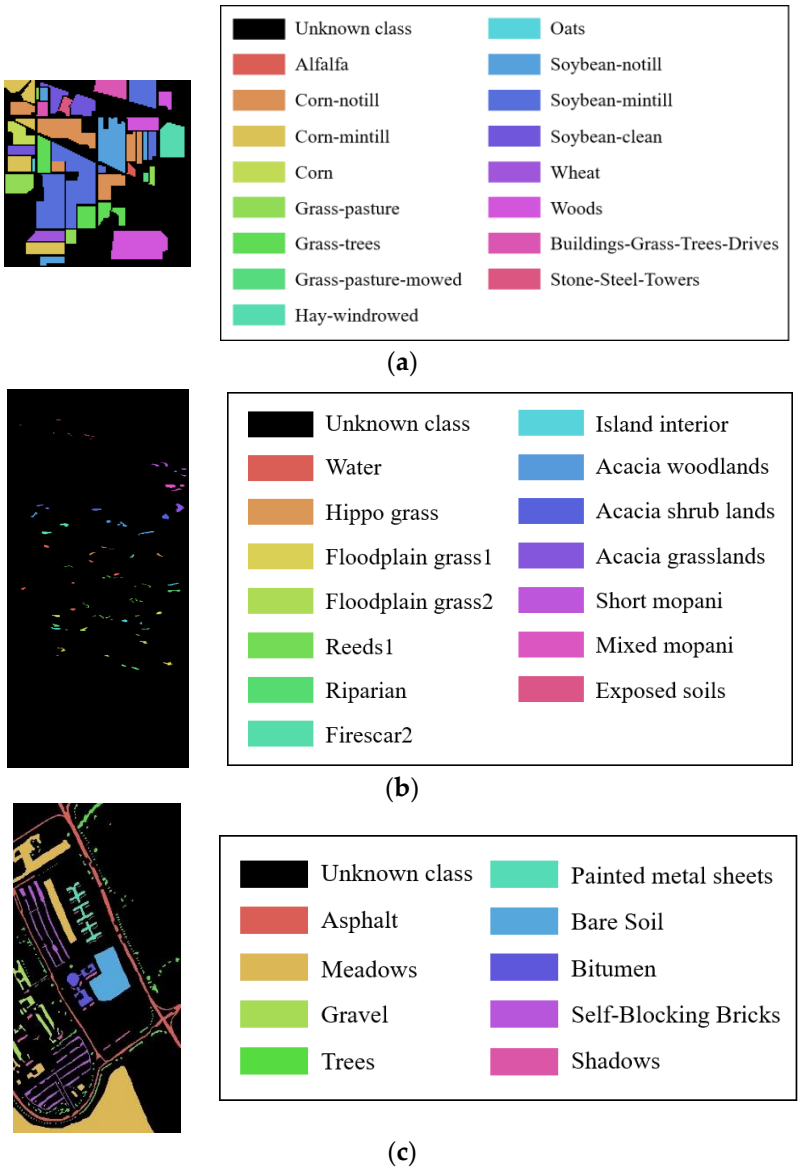
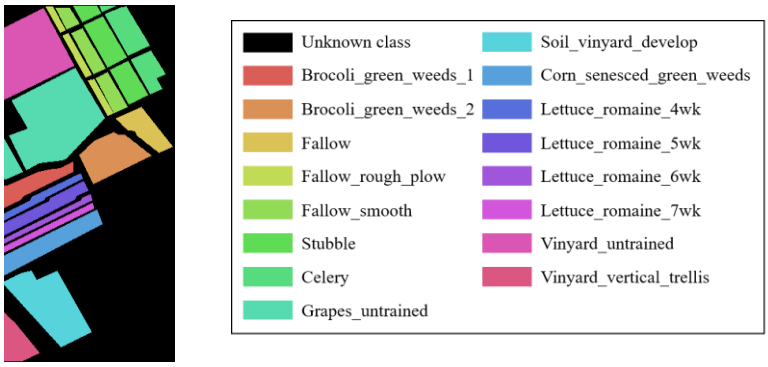


Figure 5. *Cont.*



(**d**)

Figure 5. The ground truth of four HRSIs datasets: (**a**) Indian Pines; (**b**) Botswana; (**c**) Pavia University; (**d**) Salinas.

4.2. Classification Results of Traditional Shallow Classifiers

The presented SFD feature will be compared with the spectral (Spe) feature, spectral first-order differential (Spe-1st) feature, and spectral second-order differential (Spe-2nd) feature. The above four features will be further compared through LDA dimensionality reduction to form SFD_{LDA}, Spe_{LDA}, $Spe\text{-}1st_{LDA}$, and $Spe\text{-}2nd_{LDA}$ features. The comparison process will be validated using four traditional classifiers, namely, the MD classifier, the SVM classifier, the *K*-NN classifier, and the LR classifier. For each dataset, 20% of each class data are randomly selected as training samples and the rest as testing samples. Considering the randomness of the experiment, the average overall accuracy (AOA) and standard deviation (SD), average Kappa coefficient of 10 runs are used to describe the classification results. The experimental results on four real HRSIs datasets are shown in Tables 4–7, "Average Kappa" is the abbreviation of "Average Kappa coefficient", the optimal classification results are shown in bold in each column.

Table 4. Classification results of the Indian Pines dataset on traditional shallow classifiers.

Classifier	SVM		MD		*K*-NN		LR	
	AOA (%) ± SD (%)	Average Kappa	AOA (%) ± SD (%)	Average Kappa	AOA (%) ± SD (%)	Average Kappa	AOA (%) ± SD (%)	Average Kappa
Spe	82.93 ± 0.36	0.805	46.15 ± 0.89	0.404	77.93 ± 0.44	0.748	56.30 ± 0.79	0.476
Spe-1st	67.54 ± 0.61	0.622	46.02 ± 0.80	0.402	50.41 ± 0.51	0.431	59.62 ± 1.33	0.523
Spe-2nd	54.83 ± 0.34	0.468	39.50 ± 0.85	0.332	39.67 ± 0.37	0.310	51.83 ± 0.67	0.428
SFD	**83.55 ± 0.38**	**0.812**	**48.95 ± 0.71**	**0.433**	**78.37 ± 0.54**	**0.753**	**62.89 ± 1.35**	**0.562**
Spe_{LDA}	79.53 ± 0.41	0.765	73.70 ± 0.46	0.704	78.75 ± 0.80	0.756	73.82 ± 0.52	0.699
$Spe\text{-}1st_{LDA}$	79.39 ± 0.40	0.764	73.39 ± 0.65	0.700	78.12 ± 0.69	0.749	73.84 ± 0.60	0.700
$Spe\text{-}2nd_{LDA}$	79.26 ± 0.46	0.762	72.97 ± 0.50	0.696	77.67 ± 0.68	0.744	73.33 ± 0.29	0.694
SFD_{LDA}	**79.67 ± 0.43**	**0.767**	**74.00 ± 0.51**	**0.707**	**78.96 ± 0.87**	**0.759**	**73.94 ± 0.41**	**0.701**

Table 5. Classification results of the Botswana dataset on traditional shallow classifiers.

Classifier	SVM		MD		*K*-NN		LR	
	AOA (%) ± SD (%)	Average Kappa	AOA (%) ± SD (%)	Average Kappa	AOA (%) ± SD (%)	Average Kappa	AOA (%) ± SD (%)	Average Kappa
Spe	91.79 ± 0.54	0.911	80.76 ± 0.68	0.793	90.71 ± 0.38	0.899	87.09 ± 1.20	0.860
Spe-1st	87.22 ± 0.81	0.862	79.69 ± 0.84	0.780	78.26 ± 1.08	0.765	86.13 ± 0.85	0.850
Spe-2nd	74.77 ± 1.13	0.726	65.74 ± 1.17	0.629	55.03 ± 0.74	0.515	71.96 ± 1.17	0.696
SFD	**92.55 ± 0.54**	**0.919**	**82.01 ± 0.53**	**0.805**	**91.69 ± 0.37**	**0.910**	**89.57 ± 0.92**	**0.887**
Spe_{LDA}	93.65 ± 0.45	0.931	92.85 ± 0.46	0.923	93.40 ± 0.53	0.929	89.09 ± 0.87	0.882
$Spe\text{-}1st_{LDA}$	93.22 ± 0.52	0.927	92.43 ± 0.58	0.918	92.87 ± 0.53	0.923	87.23 ± 0.57	0.862
$Spe\text{-}2nd_{LDA}$	92.42 ± 0.73	0.918	91.69 ± 0.50	0.910	92.05 ± 0.56	0.914	85.44 ± 0.94	0.842
SFD_{LDA}	**93.75 ± 0.44**	**0.944**	**93.01 ± 0.43**	**0.924**	**93.51 ± 0.46**	**0.930**	**89.37 ± 0.66**	**0.885**

Table 6. Classification results of the Pavia University dataset on traditional shallow classifiers.

Classifier	SVM		MD		*K*-NN		LR	
	AOA (%) ± SD (%)	Average Kappa	AOA (%) ± SD (%)	Average Kappa	AOA (%) ± SD (%)	Average Kappa	AOA (%) ± SD (%)	Average Kappa
Spe	89.96 ± 0.17	0.865	59.54 ± 0.44	0.501	84.77 ± 0.25	0.795	76.82 ± 1.37	0.681
Spe-1st	86.07 ± 0.14	0.811	58.66 ± 0.83	0.492	68.43 ± 0.21	0.574	81.53 ± 0.23	0.748
Spe-2nd	74.42 ± 0.18	0.643	30.82 ± 1.37	0.230	45.48 ± 0.26	0.244	74.62 ± 0.23	0.651
SFD	**91.43 ± 0.10**	**0.885**	**62.34 ± 0.35**	**0.544**	**85.46 ± 0.23**	**0.804**	**82.88 ± 0.50**	**0.767**
Spe_{LDA}	88.62 ± 0.30	0.848	71.50 ± 0.34	0.643	87.16 ± 0.41	0.827	81.91 ± 0.39	0.754
$Spe\text{-}1st_{LDA}$	88.71 ± 0.28	0.849	76.54 ± 0.37	0.699	87.54 ± 0.20	0.832	80.14 ± 0.18	0.729
$Spe\text{-}2nd_{LDA}$	87.05 ± 0.11	0.827	73.88 ± 0.47	0.669	85.58 ± 0.17	0.806	79.29 ± 0.15	0.718
SFD_{LDA}	**89.49 ± 0.29**	**0.859**	**77.52 ± 0.32**	**0.711**	**88.25 ± 0.21**	**0.842**	**81.97 ± 0.39**	**0.755**

Table 7. Classification results of the Salinas dataset on traditional shallow classifiers.

Classifier	SVM		MD		*K*-NN		LR	
	AOA (%) ± SD (%)	Average Kappa	AOA (%) ± SD (%)	Average Kappa	AOA (%) ± SD (%)	Average Kappa	AOA (%) ± SD (%)	Average Kappa
Spe	93.62 ± 0.09	0.929	75.57 ± 0.27	0.729	90.53 ± 0.16	0.895	84.51 ± 0.91	0.826
Spe-1st	91.39 ± 0.06	0.904	75.42 ± 0.14	0.727	86.68 ± 0.16	0.852	86.80 ± 0.92	0.852
Spe-2nd	88.16 ± 0.11	0.868	73.29 ± 0.19	0.704	81.71 ± 0.15	0.796	81.96 ± 0.58	0.797
SFD	**93.76 ± 0.10**	**0.930**	**76.80 ± 0.31**	**0.743**	**90.70 ± 0.17**	**0.896**	**86.92 ± 0.77**	**0.853**
Spe_{LDA}	94.26 ± 0.08	0.936	91.33 ± 0.09	0.903	93.57 ± 0.09	0.928	91.21 ± 0.08	0.902
$Spe\text{-}1st_{LDA}$	94.37 ± 0.09	0.938	91.34 ± 0.10	0.904	93.58 ± 0.10	0.928	91.22 ± 0.10	0.902
$Spe\text{-}2nd_{LDA}$	94.30 ± 0.08	0.937	91.33 ± 0.10	0.903	93.57 ± 0.08	0.928	91.22 ± 0.09	0.902
SFD_{LDA}	**94.39 ± 0.08**	**0.938**	**91.35 ± 0.09**	**0.904**	**93.59 ± 0.10**	**0.928**	**91.24 ± 0.08**	**0.902**

From Table 4, it can be seen that under 20% of the training samples in the Indian Pines dataset, compared to the original spectral feature Spe, the AOA of the extracted SFD features on SVM, MD, *K*-NN, and LR classifiers increased by 0.62%, 2.80%, 0.44%, and 6.59%, respectively; additionally, compared to the Spe-1st and Spe-2nd features, the AOA and average Kappa coefficient obtained by classification has significantly improved, indicating that the extracted SFD feature can achieve better accuracy in terrain classification. In addition, compared to the Spe_{LDA} feature by performing LDA on the original spectral feature Spe, the SFD_{LDA} feature has an AOA increase of 0.14%, 0.30%, 0.21%, and 0.12% on SVM, MD, *K*-NN, and LR classifiers; and compared to $Spe\text{-}1st_{LDA}$ feature and $Spe\text{-}2nd_{LDA}$ feature, the AOA and average Kappa coefficient have been improved to a certain extent, indicating that the extracted SFD feature can still retain their high separability after dimensionality reduction processing, enhancing the classification effect. In terms of the classification time, using the MD classifier as an example, the classification time for the Spe feature, Spe-1st feature, Spe-2nd feature, and SFD feature are 0.371 s, 0.361 s, 0.360 s, and 0.356 s, respectively. The result indicates that the extracted SFD feature can improve the accuracy of terrain classification while ensuring an almost constant classification rate.

Table 5 shows the classification results of the Botswana dataset under 20% of training samples. Compared to the original spectral feature Spe, the extracted SFD feature has significantly improved AOA and average Kappa coefficient on SVM, MD, *K*-NN, and LR classifiers compared to other features. Compared to the Spe feature, AOA has increased by 0.76%, 1.25%, 0.98%, and 2.48%, respectively, proving that the SFD feature can achieve better terrain classification accuracy. Meanwhile, compared to the Spe_{LDA} feature, the AOA of the SFD_{LDA} feature on SVM, MD, *K*-NN, and LR classifiers also increased by 0.10%, 0.16%, 0.11%, and 0.28%, respectively, indicating that the extracted SFD feature can still retain their high separability after dimensionality reduction processing, enhancing classification performance. In addition, the SD values of SFD and SFD_{LDA} features are smaller than those of other features, further proving that the extracted features have a more stable classification effect. In terms of the classification time, using the MD classifier as an example, the classification time for the Spe feature, Spe-1st feature, Spe-2nd feature, and

SFD feature are 0.132 s, 0.125 s, 0.136 s, and 0.126 s, respectively. The result indicates that the extracted SFD feature can effectively improve accuracy while maintaining runtime.

According to Table 6, it can be found that under 20% of training samples, the SFD feature extracted from the Pavia University dataset showed an increase in AOA on SVM, MD, *K*-NN, and LR classifiers by 1.47%, 2.80%, 0.69%, and 6.06%, respectively, compared to the original Spe feature. Moreover, compared to the Spe-1st and Spe-2nd features, the AOA and average Kappa coefficient of the SFD features were significantly improved, indicating that the extracted SFD feature can achieve better terrain classification accuracy. In addition, compared to the Spe_{LDA} feature, the AOA of the SFD_{LDA} feature on SVM, MD, *K*-NN, and LR classifiers increased by 0.87%, 6.02%, 1.09%, and 0.06%, respectively, and was also much greater than the AOA obtained from $Spe\text{-}1st_{LDA}$ and $Spe\text{-}2nd_{LDA}$ feature classification. Meanwhile, the SD values of the SFD feature and SFD_{LDA} feature have decreased to varying degrees compared to most other features, indicating that the SFD feature has a certain degree of stability in terrain classification compared to other features. In terms of the classification time, using the MD classifier as an example, the classification time for the Spe feature, Spe-1st feature, Spe-2nd feature, and SFD feature are 0.748 s, 0.778 s, 0.721 s, and 0.755 s, respectively. The result indicates that the extracted SFD feature can improve the accuracy of terrain classification while ensuring an almost constant classification rate.

From Table 7, it can be observed that when selecting 20% of the training samples in the Salinas dataset, the extracted SFD feature showed an increase in AOA on SVM, MD, *K*-NN, and LR classifiers by 0.14%, 1.23%, 0.17%, and 2.41%, respectively, compared to the Spe feature. Moreover, the AOA was significantly improved compared to the Spe-1st feature and Spe-2nd feature. In addition, the AOA of the SFD_{LDA} feature on SVM, MD, *K*-NN, and LR classifiers increased by 0.13%, 0.02%, 0.02%, and 0.03%, respectively, compared to the Spe_{LDA} feature. The AOA and average Kappa coefficient obtained by the SFD_{LDA} feature were also improved to some extent compared to the $Spe\text{-}1st_{LDA}$ feature and $Spe\text{-}2nd_{LDA}$ feature, indicating that the extracted SFD feature can still enhance the classification effect to some extent compared to other features. In addition, in terms of the classification time, using the MD classifier as an example, the classification time for the Spe feature, Spe-1st feature, Spe-2nd feature, and SFD feature are 1.680 s, 1.638 s, 1.666 s, and 1.642 s, respectively. The result indicates that the extracted SFD feature can effectively improve accuracy while maintaining runtime.

4.3. Classification Results of Networks

To extract deep features and verify the effectiveness of the SFD feature on different network structures, this paper sends the original spectral feature Spe, spectral first-order differential (Spe-1st) feature, spectral second-order differential (Spe-2nd) feature, spectral and frequency spectrum mixed feature (SFMF) [30], and extracted SFD feature into five different network structures for deep-feature extraction and classification, and compares the classification results. The experiments are conducted on a server with the RTX3080 graphical processing unit and 128 GB RAM, and the networks are implemented in Python. For each HRSIs dataset, 3%, 5%, and 10% samples of each class are randomly selected as training samples, and the rest are testing samples. Considering the randomness of the experimental results, the AOA and average Kappa coefficient of 10 runs were recorded to evaluate the classification effect. Tables 8–11 show the experimental results on four real HRSIs datasets, where "Avg. Kap." is the abbreviation of "Average Kappa coefficient", the optimal classification results are shown in bold.

Table 8. Classification results of the Indian Pines dataset on network models.

Model	FCN		1DCNN		3DCNN		$3DCNN_{PCA}$		HybridSN	
	AOA (%) ± SD (%)	**Avg. Kap.**	**AOA (%) ± SD (%)**	**Avg. Kap.**	**AOA (%) ± SD (%)**	**Avg. Kap.**	**AOA (%) ± SD (%)**	**Avg. Kap.**	**AOA (%) ± SD (%)**	**Avg. Kap.**
3% training samples										
Spe	62.15 ± 1.42	0.561	64.17 ± 1.13	0.585	74.12 ± 4.73	0.705	84.61 ± 0.90	0.823	87.57 ± 1.04	0.858
Spe-1st	53.83 ± 2.19	0.467	61.52 ± 1.46	0.555	76.68 ± 2.64	0.733	84.15 ± 0.49	0.819	88.77 ± 0.55	0.871
Spe-2nd	49.61 ± 1.36	0.415	49.33 ± 1.49	0.415	70.16 ± 2.41	0.656	82.86 ± 0.80	0.804	88.64 ± 0.71	0.870
SFMF	61.44 ± 0.92	0.555	68.47 ± 2.05	0.637	73.20 ± 2.28	0.693	84.90 ± 0.64	0.825	87.98 ± 0.86	0.863
SFD	**64.58 ± 1.94**	**0.589**	**73.10 ± 1.56**	**0.691**	**77.61 ± 3.61**	**0.743**	**85.74 ± 0.68**	**0.833**	**88.95 ± 0.93**	**0.874**
5% training samples										
Spe	67.62 ± 2.65	0.627	76.00 ± 1.27	0.724	79.41 ± 4.35	0.765	90.98 ± 0.49	0.895	93.20 ± 1.56	0.922
Spe-1st	59.53 ± 0.37	0.531	65.23 ± 1.12	0.601	81.89 ± 1.46	0.794	89.37 ± 1.05	0.879	94.61 ± 0.82	0.938
Spe-2nd	52.25 ± 1.43	0.449	51.61 ± 0.78	0.444	76.49 ± 1.12	0.731	89.59 ± 0.47	0.881	94.42 ± 0.47	0.936
SFMF	68.28 ± 1.45	0.634	74.24 ± 1.30	0.705	82.72 ± 1.90	0.803	90.86 ± 0.51	0.894	93.86 ± 0.93	0.927
SFD	**72.85 ± 1.88**	**0.687**	**78.79 ± 1.16**	**0.753**	**83.41 ± 2.59**	**0.811**	**91.24 ± 0.27**	**0.898**	**94.99 ± 0.51**	**0.935**
10% training samples										
Spe	73.80 ± 3.95	0.700	76.70 ± 1.43	0.733	87.75 ± 2.32	0.861	95.43 ± 0.55	0.943	98.04 ± 0.26	0.978
Spe-1st	63.61 ± 0.65	0.581	69.54 ± 0.98	0.650	87.46 ± 1.51	0.857	93.31 ± 0.47	0.924	98.15 ± 0.29	0.979
Spe-2nd	55.45 ± 0.52	0.487	53.33 ± 0.80	0.464	81.79 ± 1.18	0.792	92.41 ± 0.46	0.913	98.10 ± 0.19	0.978
SFMF	74.52 ± 1.35	0.708	79.09 ± 0.73	0.761	88.70 ± 1.45	0.871	95.72 ± 0.51	0.946	98.33 ± 0.21	0.980
SFD	**77.95 ± 3.19**	**0.749**	**81.64 ± 0.92**	**0.790**	**89.05 ± 1.65**	**0.869**	**96.03 ± 0.37**	**0.950**	**98.46 ± 0.19**	**0.982**

Table 9. Classification results of the Botswana dataset on network models.

Model	FCN		1DCNN		3DCNN		$3DCNN_{PCA}$		HybridSN	
	AOA (%) ± SD (%)	**Avg. Kap.**	**AOA (%) ± SD (%)**	**Avg. Kap.**	**AOA (%) ± SD (%)**	**Avg. Kap.**	**AOA (%) ± SD (%)**	**Avg. Kap.**	**AOA (%) ± SD (%)**	**Avg. Kap.**
3% training samples										
Spe	81.19 ± 1.41	0.796	76.26 ± 2.28	0.743	87.32 ± 4.78	0.863	97.14 ± 0.79	0.969	93.23 ± 1.75	0.927
Spe-1st	74.09 ± 2.33	0.719	55.74 ± 3.55	0.515	81.42 ± 6.27	0.799	97.35 ± 0.73	0.971	93.60 ± 2.03	0.931
Spe-2nd	59.89 ± 2.32	0.565	49.14 ± 2.70	0.444	77.73 ± 2.79	0.759	95.91 ± 0.67	0.956	91.73 ± 1.34	0.910
SFMF	82.03 ± 1.80	0.805	71.38 ± 1.55	0.689	88.47 ± 0.97	0.875	97.88 ± 0.82	0.974	93.87 ± 1.42	0.932
SFD	**86.44 ± 1.36**	**0.853**	**78.13 ± 1.87**	**0.763**	**90.24 ± 2.28**	**0.894**	**98.36 ± 0.31**	**0.981**	**94.33 ± 1.13**	**0.939**
5% training samples										
Spe	81.36 ± 2.37	0.786	81.78 ± 1.97	0.803	90.14 ± 4.13	0.893	98.70 ± 0.65	0.984	97.05 ± 1.06	0.968
Spe-1st	77.82 ± 2.10	0.759	79.63 ± 0.89	0.779	89.98 ± 1.61	0.891	97.45 ± 1.16	0.972	97.28 ± 1.02	0.971
Spe-2nd	65.30 ± 1.68	0.624	56.41 ± 2.82	0.524	86.26 ± 1.32	0.851	96.03 ± 0.98	0.957	95.61 ± 1.27	0.953
SFMF	83.22 ± 2.61	0.818	79.38 ± 1.92	0.776	90.16 ± 2.54	0.893	98.52 ± 0.75	0.981	96.41 ± 1.72	0.965
SFD	**87.02 ± 0.67**	**0.859**	**83.59 ± 1.13**	**0.822**	**91.03 ± 1.20**	**0.903**	**99.02 ± 0.15**	**0.987**	**97.49 ± 0.45**	**0.973**
10% training samples										
Spe	86.42 ± 0.83	0.853	86.13 ± 0.73	0.850	93.69 ± 2.80	0.932	99.09 ± 0.35	0.990	99.40 ± 0.58	0.994
Spe-1st	83.83 ± 0.80	0.825	86.55 ± 1.04	0.854	93.59 ± 2.81	0.931	99.01 ± 0.33	0.989	99.50 ± 0.25	0.995
Spe-2nd	73.86 ± 2.08	0.717	76.91 ± 1.46	0.750	92.42 ± 1.28	0.918	98.68 ± 0.22	0.986	99.18 ± 0.55	0.991
SFMF	88.08 ± 1.02	0.871	86.89 ± 1.16	0.858	94.37 ± 1.91	0.939	99.12 ± 0.28	0.991	99.51 ± 0.38	0.995
SFD	**89.94 ± 1.00**	**0.891**	**87.00 ± 0.76**	**0.859**	**95.24 ± 1.10**	**0.948**	**99.44 ± 0.11**	**0.993**	**99.61 ± 0.28**	**0.996**

Table 10. Classification results of the Pavia University dataset on network models.

Model	FCN		1DCNN		3DCNN		$3DCNN_{PCA}$		HybridSN	
	AOA (%) ± SD (%)	**Avg. Kap.**	**AOA (%) ± SD (%)**	**Avg. Kap.**	**AOA (%) ± SD (%)**	**Avg. Kap.**	**AOA (%) ± SD (%)**	**Avg. Kap.**	**AOA (%) ± SD (%)**	**Avg. Kap.**
3% training samples										
Spe	84.55 ± 1.58	0.794	79.13 ± 0.62	0.713	88.51 ± 5.47	0.848	98.17 ± 0.23	0.975	98.47 ± 0.62	0.980
Spe-1st	79.62 ± 1.27	0.727	72.51 ± 1.87	0.626	88.79 ± 1.89	0.853	98.14 ± 0.36	0.975	99.30 ± 0.21	0.991
Spe-2nd	72.60 ± 1.08	0.630	62.47 ± 0.43	0.477	87.67 ± 4.42	0.836	97.93 ± 0.58	0.973	98.81 ± 0.37	0.984
SFMF	84.54 ± 0.87	0.796	79.22 ± 0.36	0.720	88.59 ± 2.98	0.852	98.21 ± 0.44	0.976	98.71 ± 0.44	0.983
SFD	**85.02 ± 1.07**	**0.800**	**79.32 ± 1.56**	**0.725**	**91.49 ± 1.62**	**0.889**	**98.43 ± 0.14**	**0.978**	**99.33 ± 0.09**	**0.991**
5% training samples										
Spe	86.01 ± 2.50	0.815	82.28 ± 1.00	0.760	92.71 ± 0.91	0.904	99.01 ± 0.22	0.985	99.41 ± 0.13	0.992
Spe-1st	81.85 ± 0.96	0.754	79.30 ± 0.76	0.722	92.93 ± 0.88	0.907	98.46 ± 0.66	0.979	99.55 ± 0.20	0.994
Spe-2nd	73.48 ± 2.84	0.645	68.75 ± 0.82	0.576	92.77 ± 0.75	0.905	98.59 ± 0.13	0.981	99.58 ± 0.10	0.994
SFMF	86.55 ± 1.06	0.821	82.08 ± 0.82	0.764	91.34 ± 2.15	0.886	99.00 ± 0.35	0.985	99.42 ± 0.10	0.992
SFD	**87.00 ± 0.81**	**0.826**	**82.86 ± 1.02**	**0.765**	**93.66 ± 0.79**	**0.917**	**99.06 ± 0.06**	**0.985**	**99.59 ± 0.21**	**0.995**
10% training samples										
Spe	88.73 ± 1.64	0.850	84.90 ± 5.81	0.802	94.82 ± 0.56	0.932	99.43 ± 0.20	0.992	99.69 ± 0.07	0.996
Spe-1st	84.34 ± 2.79	0.794	81.10 ± 4.38	0.749	94.61 ± 0.40	0.929	99.39 ± 0.12	0.992	99.79 ± 0.07	0.997
Spe-2nd	78.11 ± 3.16	0.710	75.22 ± 0.89	0.668	94.92 ± 0.15	0.934	99.31 ± 0.08	0.991	99.79 ± 0.10	0.997
SFMF	89.05 ± 0.81	0.854	86.09 ± 1.06	0.821	94.15 ± 0.25	0.924	99.36 ± 0.09	0.992	99.71 ± 0.04	0.996
SFD	**89.13 ± 1.15**	**0.856**	**86.63 ± 2.17**	**0.829**	**95.31 ± 0.16**	**0.939**	**99.44 ± 0.04**	**0.992**	**99.81 ± 0.08**	**0.998**

Table 11. Classification results of the Salinas dataset on network models.

Model	FCN		1DCNN		3DCNN		$3DCNN_{PCA}$		HybridSN	
	AOA (%) ± SD (%)	**Avg. Kap.**	**AOA (%) ± SD (%)**	**Avg. Kap.**	**AOA (%) ± SD (%)**	**Avg. Kap.**	**AOA (%) ± SD (%)**	**Avg. Kap.**	**AOA (%) ± SD (%)**	**Avg. Kap.**
3% training samples										
Spe	87.43 ± 1.42	0.860	86.43 ± 1.48	0.848	90.42 ± 1.46	0.893	97.21 ± 0.31	0.968	99.70 ± 0.16	0.997
Spe-1st	87.40 ± 1.39	0.859	87.74 ± 1.29	0.863	90.82 ± 1.51	0.898	97.07 ± 0.55	0.967	99.52 ± 0.12	0.995
Spe-2nd	81.63 ± 1.16	0.795	82.48 ± 0.65	0.801	90.38 ± 2.94	0.893	96.32 ± 0.28	0.959	99.51 ± 0.15	0.995
SFMF	88.93 ± 0.52	0.878	87.94 ± 2.05	0.866	91.02 ± 0.77	0.900	97.27 ± 0.30	0.969	99.70 ± 0.08	**0.997**
SFD	**89.19 ± 0.98**	**0.879**	**88.40 ± 1.68**	**0.871**	**91.17 ± 1.56**	**0.902**	**97.62 ± 0.26**	**0.972**	**99.73 ± 0.06**	**0.997**
5% training samples										
Spe	88.26 ± 1.20	0.869	88.15 ± 1.26	0.868	91.54 ± 1.45	0.906	98.14 ± 0.30	0.979	99.85 ± 0.08	0.998
Spe-1st	88.81 ± 1.15	0.875	89.64 ± 0.74	0.884	92.00 ± 0.88	0.911	98.59 ± 0.35	0.985	99.83 ± 0.11	0.998
Spe-2nd	83.72 ± 1.65	0.819	84.73 ± 0.89	0.830	91.52 ± 1.81	0.906	98.20 ± 0.23	0.980	99.85 ± 0.07	0.998
SFMF	89.83 ± 1.17	0.887	88.39 ± 2.38	0.871	90.62 ± 2.17	0.896	98.42 ± 0.28	0.983	99.86 ± 0.10	0.998
SFD	**90.02 ± 1.18**	**0.889**	**88.90 ± 0.78**	**0.876**	**92.16 ± 1.37**	**0.913**	**98.70 ± 0.18**	**0.986**	**99.90 ± 0.04**	**0.999**
10% training samples										
Spe	90.67 ± 0.59	0.896	90.58 ± 0.75	0.895	92.94 ± 1.79	0.922	99.63 ± 0.14	0.996	99.97 ± 0.02	0.999
Spe-1st	90.59 ± 0.79	0.895	91.35 ± 0.23	0.906	92.47 ± 2.75	0.917	99.50 ± 0.23	0.995	99.93 ± 0.13	0.999
Spe-2nd	86.37 ± 0.69	0.848	86.99 ± 0.47	0.855	92.96 ± 1.67	0.922	99.48 ± 0.11	0.994	99.92 ± 0.12	0.999
SFMF	91.57 ± 0.84	0.897	91.31 ± 0.36	0.896	93.08 ± 1.50	0.922	99.60 ± 0.10	0.996	99.95 ± 0.06	0.999
SFD	**91.68 ± 0.79**	**0.907**	**91.58 ± 0.20**	**0.905**	**93.19 ± 1.72**	**0.924**	**99.64 ± 0.07**	**0.996**	**99.97 ± 0.01**	**0.999**

According to Table 8, it can be found that on the Indian Pines dataset, the AOA and average Kappa coefficient of the deep SFD feature are significantly higher than those of the deep Spe feature, deep Spe-1st feature, deep Spe-2nd feature, and SFMF on the five network models under 3%, 5%, and 10% training samples, and the deep Spe-1st and deep Spe-2nd features, generally, have lower AOA and average Kappa coefficient compared to the deep Spe feature. This indicates that the SFD feature extracted using fractional-

order differentiation can enhance recognition performance compared to features extracted using first-order differentiation and second-order differentiation. In addition, when the proportion of training samples is small, the deep SFD feature performs relatively better in terrain classification accuracy compared to other features, such as the results of $3DCNN_{PCA}$ and HybridSN models. Under the condition of 3% training samples, the number of training samples in each class is lower than 30 (except for classes 2, 11, and 14, the number of 3% samples per class is 42, 73, and 37, respectively), this indicates that even under the condition of small-size training samples, the SFD feature is superior to other features. Meanwhile, through comparison, it can be seen that the SD value of the deep SFD feature is also smaller compared to the deep Spe feature, indicating that the classification effect of the deep SFD feature has a better stability. In terms of running time, using the $3DCNN_{PCA}$ model with 5% training samples as an example, the testing times for the Spe feature, Spe-1st feature, Spe-2nd feature, SFMF, and SFD feature are 0.475 s, 0.476 s, 0.475 s, 0.531 s, and 0.476 s, respectively. The result indicates that the extracted SFD feature can effectively improve accuracy while maintaining runtime.

Figure 6 shows the classification maps of the Indian Pines dataset of deep Spe feature and deep SFD feature on five network models under 5% training samples. Through comparison, it can be found that the classification results of the deep SFD feature are, generally, better than those of the deep Spe feature on the five network models, with fewer misclassified pixels.

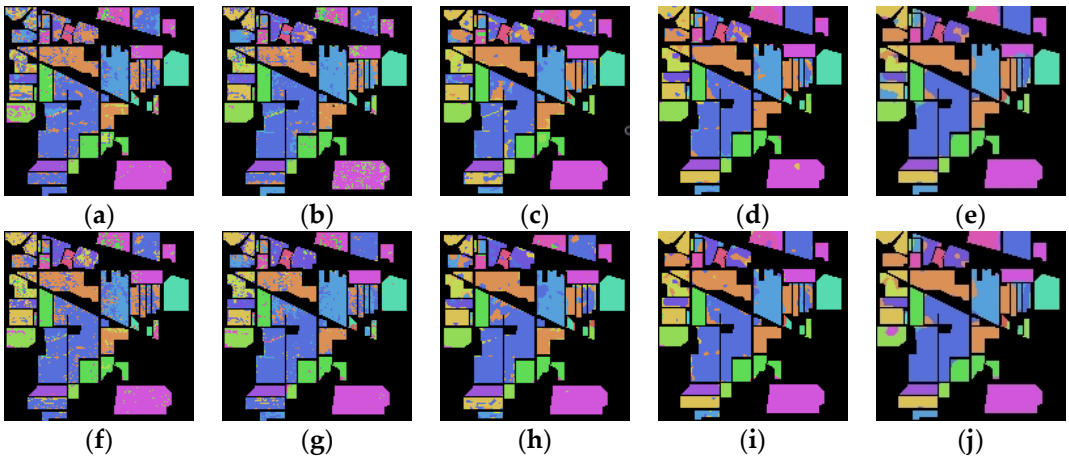


Figure 6. Indian Pines dataset classification map: (**a**) Spe feature in FCN model with 69.36% AOA; (**b**) Spe feature in 1DCNN model with 77.12% AOA; (**c**) Spe feature in 3DCNN model with 81.53% AOA; (**d**) Spe feature in $3DCNN_{PCA}$ model with 90.77% AOA; (**e**) Spe feature in HybridSN model with 93.81% AOA; (**f**) SFD feature in FCN model with 73.88% AOA; (**g**) SFD feature in 1DCNN model with 79.77% AOA; (**h**) SFD feature in 3DCNN model with 85.39% AOA; (**i**) SFD feature in $3DCNN_{PCA}$ model with 91.44% AOA; (**j**) SFD feature in HybridSN model with 95.47% AOA.

Table 9 shows the classification results of the presented SFD feature compared to the Spe feature, Spe-1st feature, Spe-2nd feature, and SFMF on five network models for the Botswana dataset with 3%, 5%, and 10% training samples. It can be found that the AOA of the SFD feature proposed in this paper has improved compared to the other three features on all five models, making it more effective for terrain classification. Additionally, when the proportion of training samples is smaller, the AOA and average Kappa coefficient of the SFD feature are significantly improved compared to other features. Under the condition of 3% training samples, the number of training samples in each class is far lower than 30, indicating that in the case of small-size training samples, the SFD feature can better exert its advantages compared to other features. At the same time, it can be found that the SD values of the SFD feature are, generally, smaller than those of other features, indicating that the SFD feature is more stable in the classification. In terms of running time, using

the $3DCNN_{PCA}$ model with 5% training samples as an example, the testing times for the Spe feature, Spe-1st feature, Spe-2nd feature, SFMF, and SFD feature are 0.349 s, 0.349 s, 0.348 s, 0.481 s, and 0.312 s, respectively. The result indicates that the extracted SFD feature not only improves the accuracy of terrain classification but also has a more efficient running rate compared to other features.

Figure 7 shows the classification results of the Spe and the presented SFD features of the Botswana dataset on five network models under 5% training samples. Through comparison, it can be seen that the classification results of the SFD feature are, generally, better than those of the Spe feature on the five network models, further demonstrating the effectiveness of the SFD feature in terrain classification.

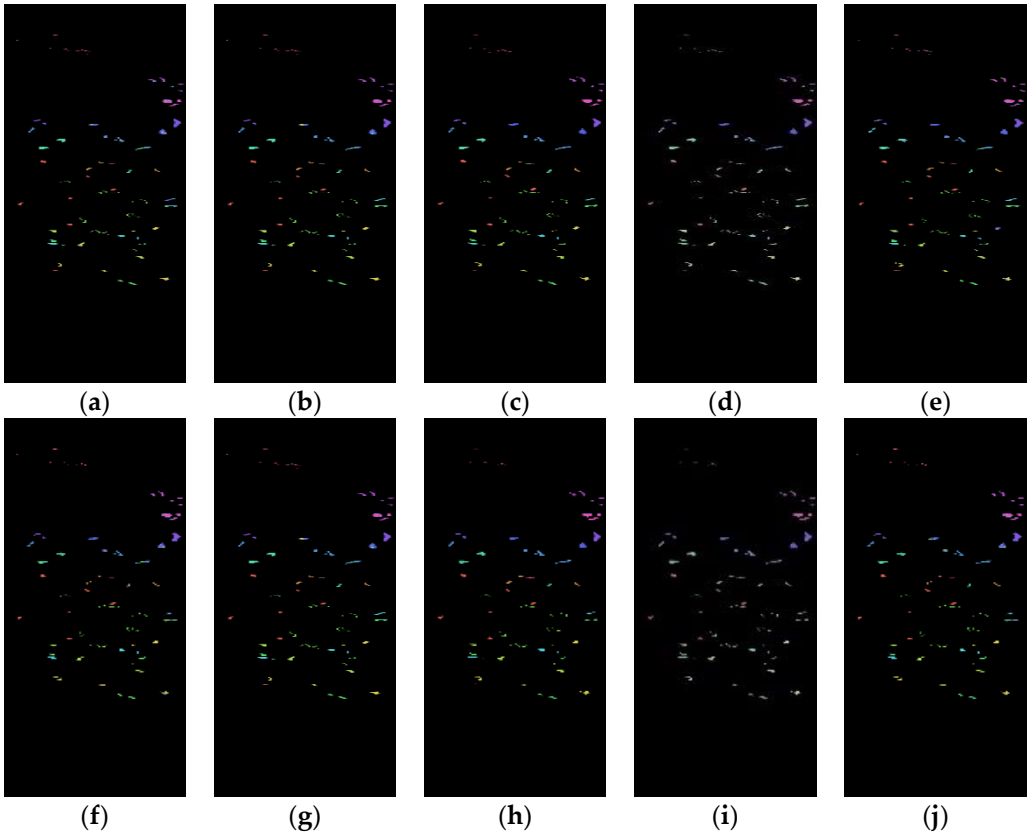


Figure 7. Botswana dataset classification map: (**a**) Spe feature in FCN model with 82.48% AOA; (**b**) Spe feature in 1DCNN model with 82.76% AOA; (**c**) Spe feature in 3DCNN model with 89.02% AOA; (**d**) Spe feature in $3DCNN_{PCA}$ model with 98.57% AOA; (**e**) Spe feature in HybridSN model with 96.45% AOA; (**f**) SFD feature in FCN model with 87.37% AOA; (**g**) SFD feature in 1DCNN model with 84.58% AOA; (**h**) SFD feature in 3DCNN model with 91.92% AOA; (**i**) SFD feature in $3DCNN_{PCA}$ model with 99.16% AOA; (**j**) SFD feature in HybridSN model with 97.93% AOA.

From Table 10, it can be seen that on the Pavia University dataset, the presented SFD feature has higher AOA and average Kappa coefficient compared to the Spe feature, Spe-1st feature, Spe-2nd feature, and SFMF on the five network models at 3%, 5%, and 10% of the training samples. Additionally, the smaller the proportion of training samples, the more significant the improvement in the AOA of the SFD feature on certain models. For example, on 3DCNN, the AOA of the SFD feature increased by 2.98%, 0.95%, and 0.49% compared to Spe feature under 3%, 5%, and 10% training samples, respectively. Meanwhile, the SD values of the SFD feature are also smaller than those of other features, indicating that the presented SFD feature is more stable in the classification compared to other features. In terms of running time, using the $3DCNN_{PCA}$ model with 5% training samples as an example, the testing times for the Spe feature, Spe-1st feature, Spe-2nd feature, SFMF, and

SFD feature are 1.648 s, 1.646 s, 1.647 s, 2.068 s, and 1.634 s, respectively. The result indicates that the extracted SFD feature can effectively improve accuracy while maintaining runtime.

Figure 8 shows the classification maps of the Spe feature and the presented SFD feature on five network models for the Pavia University dataset under 5% training samples. Through comparison, it can be found that the classification results of the SFD feature are, generally, better than those of the Spe feature, which further proves the effectiveness of the extracted SFD feature in terrain classification.

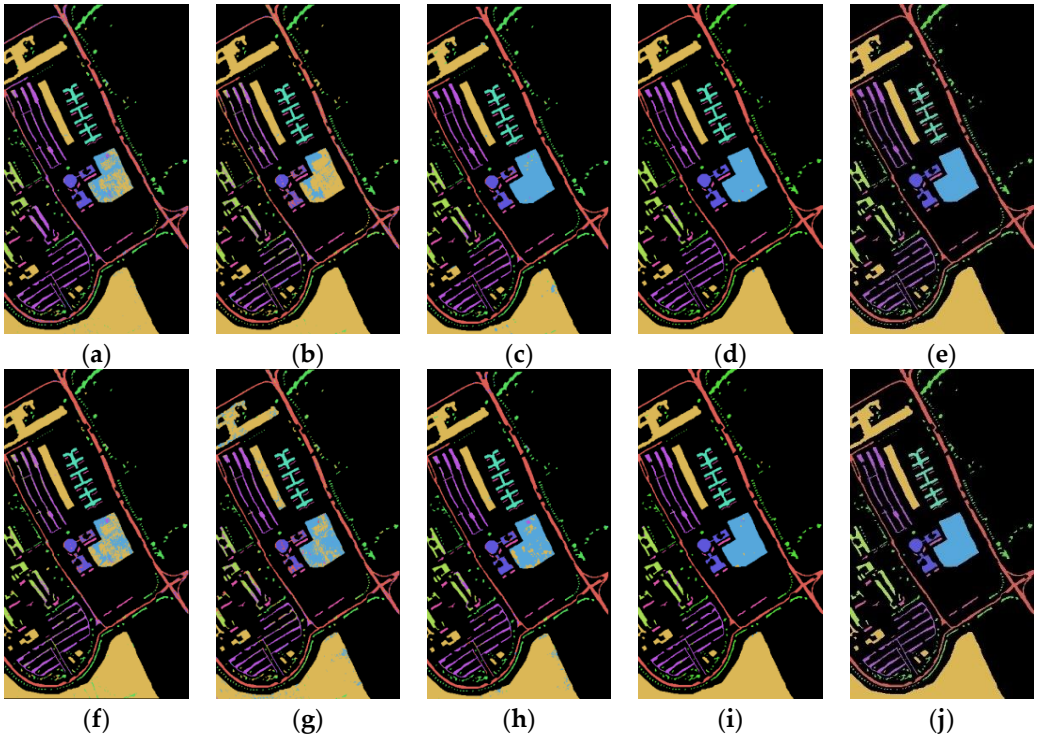


Figure 8. Pavia University dataset classification map: (**a**) Spe feature in FCN model with 85.36% AOA; (**b**) Spe feature in 1DCNN model with 81.66% AOA; (**c**) Spe feature in 3DCNN model with 93.03% AOA; (**d**) Spe feature in $3DCNN_{PCA}$ model with 98.89% AOA; (**e**) Spe feature in HybridSN model with 99.29% AOA; (**f**) SFD feature in FCN model with 87.73% AOA; (**g**) SFD feature in 1DCNN model with 83.39% AOA; (**h**) SFD feature in 3DCNN model with 94.34% AOA; (**i**) SFD feature in $3DCNN_{PCA}$ model with 99.12% AOA; (**j**) SFD feature in HybridSN model with 99.79% AOA.

Figure 9 shows the classification maps of the Spe feature and SFD feature of the Salinas dataset on five network models under 5% training samples. It can be found that the classification results of the presented SFD feature are, generally, better than those of the Spe feature on five network models, and the misclassification rate of the SFD feature is lower compared to the Spe feature, indicating that the extracted SFD feature can effectively improve the classification accuracy.

From Table 11, it can be seen that at 3%, 5%, and 10% of the training samples, the SFD feature extracted from the Salinas dataset has a certain improvement in AOA and average Kappa coefficient compared to the Spe feature, Spe-1st feature, Spe-2nd feature, and SFMF on the five network models. Moreover, when the proportion of training samples is small, the AOA of the presented SFD feature is more significantly improved. For example, on the 3DCNN model, when the proportion of training samples is 3%, 5%, and 10%, the AOA of the SFD feature increased by 0.75%, 0.62%, and 0.25% compared to the Spe feature, respectively. In addition, the SD values of the SFD feature are, generally, smaller compared to other features, further indicating that the SFD feature has higher stability in the classification. In terms of running time, using the $3DCNN_{PCA}$ model with 5% training samples as an example, the testing times for the Spe feature, Spe-1st feature, Spe-2nd feature, SFMF, and

SFD feature are 2.021 s, 2.136 s, 2.056 s, 2.499 s, and 2.093 s, respectively. The result indicates that the extracted SFD feature can effectively improve accuracy while maintaining runtime.

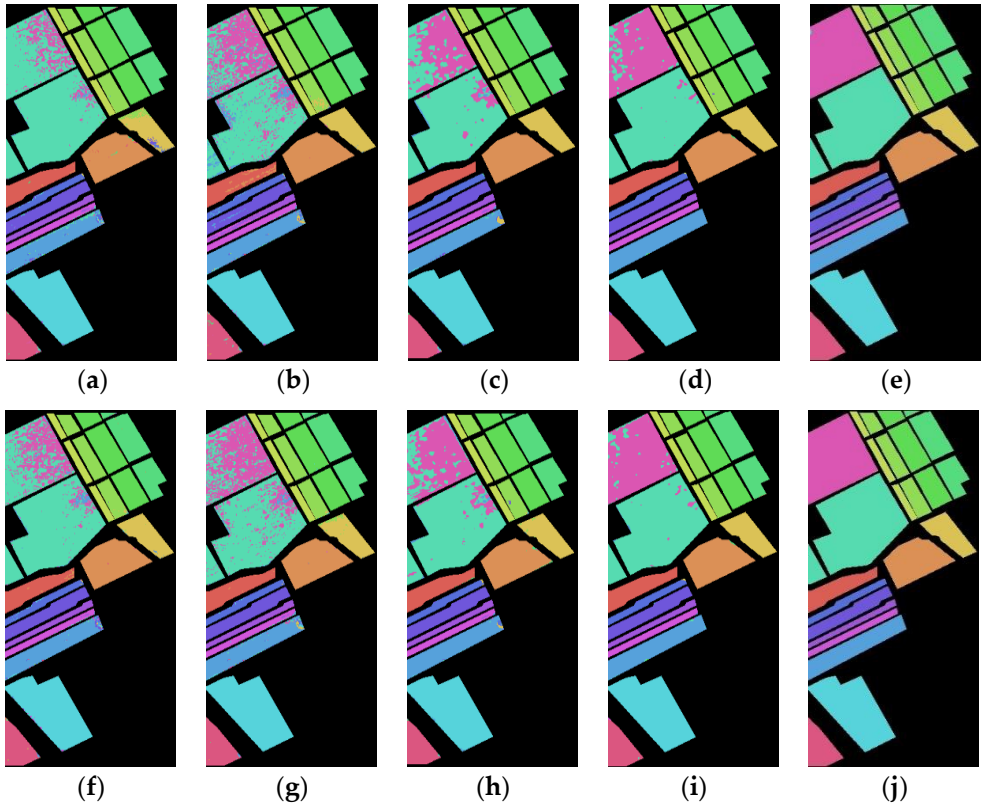


Figure 9. Salinas dataset classification map: (**a**) Spe feature in FCN model with 88.77% AOA; (**b**) Spe feature in 1DCNN model with 87.14% AOA; (**c**) Spe feature in 3DCNN model with 91.02% AOA; (**d**) Spe feature in $3DCNN_{PCA}$ model with 97.89% AOA; (**e**) Spe feature in HybridSN model with 99.78% AOA; (**f**) SFD feature in FCN model with 90.95% AOA; (**g**) SFD feature in 1DCNN model with 89.35% AOA; (**h**) SFD feature in 3DCNN model with 92.98% AOA; (**i**) SFD feature in $3DCNN_{PCA}$ model with 98.86% AOA; (**j**) SFD feature in HybridSN model with 99.94% AOA.

Table 12 shows the small-size training samples experiments on the Pavia University and Salinas datasets under the condition of 30 training samples per class, the optimal classification results are shown in bold. From Table 12, it can be concluded that, in the case of small-size training samples, the SFD feature has greater advantages compared to other features on the Pavia University and Salinas datasets.

Table 12. Classification results of Pavia University and Salinas datasets on network models under 30 training samples per class.

Model	FCN		1DCNN		3DCNN		$3DCNN_{PCA}$		HybridSN	
	AOA (%) ± SD (%)	**Avg. Kap.**	**AOA (%) ± SD (%)**	**Avg. Kap.**	**AOA (%) ± SD (%)**	**Avg. Kap.**	**AOA (%) ± SD (%)**	**Avg. Kap.**	**AOA (%) ± SD (%)**	**Avg. Kap.**
					Pavia University					
Spe	75.01 ± 2.00	0.654	74.86 ± 0.81	0.654	77.23 ± 5.10	0.705	72.63 ± 0.77	0.611	85.65 ± 1.63	0.817
Spe-1st	67.08 ± 1.54	0.554	64.63 ± 1.40	0.503	68.03 ± 3.02	0.566	74.96 ± 1.06	0.664	91.37 ± 1.78	0.885
Spe-2nd	55.05 ± 1.05	0.353	53.17 ± 3.90	0.274	58.33 ± 3.86	0.400	74.39 ± 0.45	0.662	90.70 ± 2.19	0.876
SFMF	76.94 ± 2.24	0.695	73.03 ± 1.70	0.632	79.12 ± 1.46	0.720	71.38 ± 1.12	0.609	87.79 ± 1.06	0.836
SFD	**77.28 ± 1.91**	**0.696**	**78.02 ± 0.72**	**0.704**	**82.04 ± 2.14**	**0.762**	**76.34** ± 1.05	**0.682**	**92.81** ± 1.58	**0.904**

Table 12. *Cont.*

Model	FCN		1DCNN		3DCNN		$3DCNN_{PCA}$		HybridSN	
	AOA (%) ± SD (%)	Avg. Kap.	AOA (%) ± SD (%)	Avg. Kap.	AOA (%) ± SD (%)	Avg. Kap.	AOA (%) ± SD (%)	Avg. Kap.	AOA (%) ± SD (%)	Avg. Kap.
Salinas										
Spe	83.10 ± 1.54	0.811	85.05 ± 0.41	0.834	83.16 ± 3.48	0.812	81.69 ± 1.18	0.795	98.92 ± 0.56	0.988
Spe-1st	83.65 ± 0.68	0.817	80.35 ± 3.16	0.780	81.25 ± 3.75	0.792	85.97 ± 0.44	0.844	98.97 ± 0.22	0.989
Spe-2nd	83.46 ± 2.47	0.805	82.43 ± 2.47	0.804	83.59 ± 4.30	0.818	85.09 ± 0.52	0.834	98.87 ± 0.36	0.987
SFMF	83.50 ± 3.18	0.817	84.03 ± 3.35	0.822	83.51 ± 3.20	0.816	79.89 ± 1.24	0.774	98.02 ± 0.75	0.978
SFD	**85.02 ± 2.39**	**0.834**	**85.78 ± 0.87**	**0.841**	**84.44 ± 3.99**	**0.828**	**86.93 ± 0.52**	**0.854**	**99.02 ± 0.11**	**0.989**

4.4. Discussion of Classification Results

From the above experimental results, it can be seen that the proposed SFD feature can effectively improve the classification accuracy of HRSIs. In the four HRSI datasets, the SFD feature has improved the accuracy of terrain classification to varying degrees. To demonstrate the effectiveness of the proposed criteria, Table 13 takes the MD classifier as an example and shows the AOA and SD vary with the SFD order v in the range of 0.1 to 0.9 at step 0.1 on four HRSIs datasets. For each dataset, 20% of each class data is randomly selected as a training sample and the rest are testing samples. The best result of each column is shown in bold.

Table 13. Classification results of MD classifier with SFD order in the range of 0~0.9.

Dataset	Indian Pines	Botswana	Pavia University	Salinas
SFD Order v	**AOA (%) ± SD (%)**			
0	46.15 ± 0.89	80.76 ± 0.68	59.54 ± 0.44	75.57 ± 0.27
0.1	46.60 ± 0.89	81.01 ± 0.58	59.84 ± 0.33	76.13 ± 0.28
0.2	47.02 ± 0.82	81.61 ± 0.66	60.35 ± 0.39	76.54 ± 0.29
0.3	47.67 ± 0.85	**82.01 ± 0.53**	60.77 ± 0.42	76.73 ± 0.35
0.4	48.29 ± 0.82	81.96 ± 0.47	61.32 ± 0.43	**76.80 ± 0.31**
0.5	48.85 ± 0.75	81.44 ± 0.48	61.86 ± 0.40	76.75 ± 0.26
0.6	**48.95 ± 0.71**	80.49 ± 0.41	**62.34 ± 0.35**	76.55 ± 0.23
0.7	48.56 ± 0.55	78.68 ± 0.47	61.93 ± 0.42	76.32 ± 0.21
0.8	47.72 ± 0.48	76.12 ± 0.63	60.16 ± 0.51	75.99 ± 0.19
0.9	46.54 ± 0.58	74.25 ± 0.71	57.21 ± 0.54	75.67 ± 0.15

Table 13 shows that the SFD order v corresponding to the highest AOA of each dataset is mainly within the range of the peaks of criterion J in Figure 3. Additionally, the variation trend of classification accuracy with SFD order is also similar to that of criterion J with SFD order, which proves the feasibility of the presented SFD order selection criterion. It can be concluded that the presented criterion J is an effective method to select appropriate SFD order v, and performing fractional differentiation on the pixel spectral curves with the selected order v will achieve the efficient SFD feature that can improve the classification accuracy.

For two classes that are easily misclassified, the SFD feature shows its advantage and can enhance the separability between these two classes. Taking the Salinas dataset as an example, Table 14 shows the classification accuracy of each class and the overall accuracy, the significantly improved class at order 0.5 is shown in bold. It is shown in Table 14 that for most classes, the results of the SFD feature are better or equal to the original spectral feature. Because most classes in the Salinas dataset are vegetation and crops, which leads to different subjects with similar spectra, in this case, the local burrs characteristics of the pixel spectral curves, which correspond to the high-frequency components, contribute most to the identification. The extracted SFD feature can enhance the high-frequency components

while sufficiently retaining the low-frequency components of the spectral pixel, thus, the separability of these similar classes will increase and the classification accuracy will be improved, which confirms the results discussed in Section 2.3.

Table 14. Classification accuracy of each class and overall accuracy of Salinas dataset by MD classifier with SFD order in the range of 0~0.9.

SFD Order v	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Class 1	98.20	98.26	98.32	98.44	98.57	98.76	98.82	98.82	98.88	99.00
Class 2	79.44	80.41	80.81	80.61	80.64	**81.08**	81.15	81.08	80.98	80.68
Class 3	73.50	76.98	77.29	77.42	76.72	**76.03**	75.14	73.50	72.42	72.17
Class 4	98.57	98.57	98.57	98.57	98.57	98.48	98.39	98.30	98.21	98.21
Class 5	95.33	95.47	95.61	95.80	96.13	96.50	96.64	96.78	96.78	96.73
Class 6	96.68	97.00	97.19	97.16	97.16	97.16	97.19	97.19	97.22	97.10
Class 7	98.64	98.46	98.46	98.39	98.36	98.36	98.32	98.29	98.11	97.94
Class 8	60.67	61.18	61.57	62.02	62.19	**62.32**	62.35	62.37	62.29	62.07
Class 9	89.76	90.81	91.60	92.64	93.09	**93.53**	94.03	94.18	94.24	94.18
Class 10	23.11	23.42	23.23	21.85	20.90	20.06	19.03	18.31	18.12	18.12
Class 11	80.45	81.62	82.55	83.14	84.54	**86.18**	86.89	86.77	87.35	87.35
Class 12	89.82	91.76	93.19	93.19	92.54	90.73	87.87	84.89	81.00	78.92
Class 13	98.50	98.50	98.50	98.50	98.50	98.50	98.36	98.50	98.36	98.36
Class 14	88.79	88.43	88.32	88.32	88.32	88.32	88.43	88.32	88.32	88.20
Class 15	61.56	62.04	62.59	62.80	63.24	**63.26**	63.36	63.52	63.09	62.90
Class 16	52.35	52.49	52.90	53.46	54.22	**54.50**	53.53	51.11	49.31	47.51
OA	75.18	75.80	76.17	76.36	76.47	**76.51**	76.38	76.13	75.80	75.54

The experimental results have verified the validity of the proposed SFD feature-extraction method. The reason behind the experimental phenomenon is that the presented SFD feature-extraction method uses fractional differentiation to extract both the low-frequency components characteristics and high-frequency components characteristics of the pixel spectral curves of HRSIs, which can preserve both the overall curve shape and local burrs characteristics of the pixel spectral curves of HRSIs. On the other hand, the experimental results also show the effectiveness of the presented criterion for selecting the fractional-differentiation order. The network models perform deep-feature extraction based on importing the SFD feature and, thus, achieve efficient deep features that can further improve terrain classification accuracy. Especially under the condition of small-size training samples, the terrain classification accuracy is improved more significantly.

5. Conclusions

In this paper, a spectral fractional-differentiation (SFD) feature of HRSIs is presented, and a fractional-differentiation order selection criterion is proposed. The MD classifier, SVM classifier, K-NN classifier, and LR classifier are used to evaluate the performance of the presented SFD feature. The obtained SFD feature is sent to the FCN and 1DCNN for deep-feature extraction and classification, and the SFD-Spa feature cube containing spatial information is sent to 3DCNN for deep-feature extraction and classification. The SFD-Spa feature after performing PCA on spectral pixels is directly connected with the first principal component of the original data and sent to $3DCNN_{PCA}$ and HybridSN models to extract deep features. The experimental results on four real HRSIs show that the extracted SFD feature can effectively improve the accuracy of terrain classification, and sending SFD feature to deep-learning environments can further improve the accuracy of terrain classification for HRSIs, especially in the case of small-size training samples. The presented SFD feature-extraction method has limitations, such as the fact that the fractional-differentiation order needs to be selected, the SFD feature-extraction method cannot reduce the dimensionality of data, and the presented method should be performed on the datum one by one because there is no projection matrix that suits LDA or PCA.

Author Contributions: Conceptualization, Y.L. (Yang Li); formal analysis, Y.L. (Yang Li); funding acquisition, F.Z., Y.L. (Yi Liu), and J.L.; investigation, J.L. and Y.L. (Yang Li); methodology, J.L. and Y.L. (Yang Li); project administration, J.L., Y.L. (Yi Liu), and F.Z.; software, Y.L. (Yang Li); supervision, J.L.; validation, J.L.; visualization, Y.L. (Yang Li); writing—original draft, Y.L. (Yang Li); writing—review and editing, J.L. and Y.L. (Yi Liu). All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China (no. 62176140, no. 62077038, and no. 61672405) and the Natural Science Foundation of Shaanxi Province of China (no. 2021JM-459).

Acknowledgments: The authors would like to thank the editors and anonymous reviewers who handled our paper.

Conflicts of Interest: The authors declare that there are no conflict of interest regarding the publication of this paper.

References

1. Alcolea, A.; Paoletti, M.E.; Haut, J.M.; Resano, J.; Plaza, A. Inference in Supervised Spectral Classifiers for On-Board Hyperspectral Imaging: An Overview. *Remote Sens.* **2020**, *12*, 534. [CrossRef]
2. Ma, W.; Yang, Q.; Wu, Y.; Zhao, W.; Zhang, X. Double-Branch Multi-Attention Mechanism Network for Hyperspectral Image Classification. *Remote Sens.* **2019**, *11*, 1307. [CrossRef]
3. Li, S.; Song, W.; Fang, L.; Chen, Y.; Ghamisi, P.; Benediktsson, J.A. Deep Learning for Hyperspectral Image Classification: An Overview. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 6690–6709. [CrossRef]
4. Mei, X.; Pan, E.; Ma, Y.; Dai, X.; Huang, J.; Fan, F.; Du, Q.; Zheng, H.; Ma, J. Spectral-Spatial Attention Networks for Hyperspectral Image Classification. *Remote Sens.* **2019**, *11*, 963. [CrossRef]
5. Renard, N.; Bourennane, S.; Blanc-Talon, J. Denoising and Dimensionality Reduction Using Multilinear Tools for Hyperspectral Images. *IEEE Geosci. Remote Sens. Lett.* **2008**, *5*, 138–142. [CrossRef]
6. Dong, Y.; Du, B.; Zhang, L.; Zhang, L. Dimensionality Reduction and Classification of Hyperspectral Images Using Ensemble Discriminative Local Metric Learning. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 2509–2524. [CrossRef]
7. Li, N.; Zhou, D.; Shi, J.; Wu, T.; Gong, M. Spectral-Locational-Spatial Manifold Learning for Hyperspectral Images Dimensionality Reduction. *Remote Sens.* **2021**, *13*, 2752. [CrossRef]
8. Licciardi, G.; Marpu, P.R.; Chanussot, J.; Benediktsson, J.A. Linear Versus Nonlinear PCA for the Classification of Hyperspectral Data Based on the Extended Morphological Profiles. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 447–451. [CrossRef]
9. Bandos, T.V.; Bruzzone, L.; Camps-Valls, G. Classification of Hyperspectral Images with Regularized Linear Discriminant Analysis. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 862–873. [CrossRef]
10. Bao, J.; Chi, M.; Benediktsson, J.A. Spectral Derivative Features for Classification of Hyperspectral Remote Sensing Images: Experimental Evaluation. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* **2013**, *6*, 594–601. [CrossRef]
11. Ye, Z.; He, M.; Fowler, J.E.; Du, Q. Hyperspectral Image Classification Based on Spectra Derivative Features and Locality Preserving Analysis. In Proceedings of the 2014 IEEE China Summit and International Conference on Signal and Information Processing, Xi'an, China, 4 September 2014; pp. 138–142.
12. Tian, A.; Zhao, J.; Tang, B.; Zhu, D.; Fu, C.; Xiong, H. Hyperspectral Prediction of Soil Total Salt Content by Different Disturbance Degree under a Fractional-Order Differential Model with Differing Spectral Transformations. *Remote Sens.* **2021**, *13*, 4283. [CrossRef]
13. Lao, C.; Chen, J.; Zhang, Z.; Chen, Y.; Ma, Y.; Chen, H.; Gu, X.; Ning, J.; Jin, J.; Li, X. Predicting the Contents of Soil Salt and Major Water-soluble Ions with Fractional-order Derivative Spectral Indices and Variable Selection. *Comput. Electron. Agric.* **2021**, *182*, 106031. [CrossRef]
14. Hong, Y.; Guo, L.; Chen, S.; Linderman, M.; Mouazen, A.M.; Yu, L.; Chen, Y.; Liu, Y.; Liu, Y.; Chen, H.; et al. Exploring the Potential of Airborne Hyperspectral Image for Estimating Topsoil Organic Carbon: Effects of Fractional-order Derivative and Optimal Band Combination Algorithm. *Geoderma* **2020**, *365*, 114228. [CrossRef]
15. Gao, Q.; Lim, S.; Jia, X. Hyperspectral Image Classification Using Convolutional Neural Networks and Multiple Feature Learning. *Remote Sens.* **2018**, *10*, 299. [CrossRef]
16. Hu, W.; Huang, Y.; Wei, L.; Zhang, F.; Li, H. Deep Convolutional Neural Networks for Hyperspectral Image Classification. *J. Sens.* **2015**, *2015*, 258619. [CrossRef]
17. Xu, Y.; Du, B.; Zhang, F.; Zhang, L. Hyperspectral Image Classification Via a Random Patches Network. *ISPRS J. Photogramm. Remote Sens.* **2018**, *142*, 344–357. [CrossRef]
18. Chen, Y.; Jiang, H.; Li, C.; Jia, X.; Ghamisi, P. Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6232–6251. [CrossRef]
19. Audebert, N.; Le Saux, B.; Lefevre, S. Deep Learning for Classification of Hyperspectral Data: A Comparative Review. *IEEE Geosci. Remote Sens. Mag.* **2019**, *7*, 159–173. [CrossRef]

20. Roy, S.K.; Krishna, G.; Dubey, S.R.; Chaudhuri, B.B. HybridSN: Exploring 3-D–2-D CNN Feature Hierarchy for Hyperspectral Image Classification. *IEEE Geosci. Remote Sens. Lett.* **2020**, *17*, 277–281. [CrossRef]
21. Chen, W.; Jia, Z.; Yang, J.; Kasabov, N.K. Multispectral Image Enhancement Based on the Dark Channel Prior and Bilateral Fractional Differential Model. *Remote Sens.* **2022**, *14*, 233. [CrossRef]
22. Dong, S.; Quan, Y.; Feng, W.; Dauphin, G.; Gao, L.; Xing, M. A Pixel Cluster CNN and Spectral-Spatial Fusion Algorithm for Hyperspectral Image Classification with Small-Size Training Samples. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* **2021**, *14*, 4101–4114. [CrossRef]
23. Aydemir, M.S.; Bilgin, G. Semisupervised Hyperspectral Image Classification Using Small Sample Sizes. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 621–625. [CrossRef]
24. Sun, H.; Chang, A.; Zhang, Y.; Chen, W. A Review on Variable-order Fractional Differential Equations: Mathematical Foundations, Physical Models, Numerical Methods and Applications. *Fract. Calc. Appl. Anal.* **2019**, *22*, 27–59. [CrossRef]
25. Podlubny, I. *Fractional Differential Equations*; Academic Press: San Diego, CA, USA, 1999; pp. 43–48, 203.
26. Pu, Y. Fractional Calculus Approach to Texture of Digital Image. In Proceedings of the 8th International Conference on Signal Processing, Guilin, China, 16 November 2006; pp. 1002–1006.
27. Pu, Y.; Yuan, X.; Liao, K.; Chen, Z.; Zhou, J. Five Numerical Algorithms of Fractional Calculus Applied in Modern Signal Analyzing and Processing. *J. Sichuan Univ. Eng. Sci. Ed.* **2005**, *37*, 118–124.
28. Fauvel, M.; Tarabalka, Y.; Benediktsson, J.A.; Chanussot, J.; Tilton, J.C. Advances in Spectral-spatial Classification of Hyperspectral Images. *Proc. IEEE* **2013**, *101*, 652–675. [CrossRef]
29. Vali, A.; Comai, S.; Matteucci, M. Deep Learning for Land Use and Land Cover Classification Based on Hyperspectral and Multispectral Earth Observation Data: A Review. *Remote Sens.* **2020**, *12*, 2495. [CrossRef]
30. Liu, J.; Yang, Z.; Liu, Y.; Mu, C. Hyperspectral Remote Sensing Images Deep Feature Extraction Based on Mixed Feature and Convolutional Neural Networks. *Remote Sens.* **2021**, *13*, 2599. [CrossRef]