applied sciences

MDPI

Article

SF-ICNN: Spectral–Fractal Iterative Convolutional Neural Network for Classification of Hyperspectral Images

Behnam Asghari Beirami [1], **Mehran Alizadeh Pirbasti** [2] and **Vahid Akbari** [3,*]

[1] Department of Photogrammetry and Remote Sensing, Faculty of Geodesy and Geomatics, K.N. Toosi University of Technology, Tehran 15433-19967, Iran; b_asghari@email.kntu.ac.ir
[2] SFI Centre for Research Training in Machine Learning, School of Computer Science, University College Dublin, D04 V1W8 Dublin, Ireland; mehran.alizadehpirbasti@ucdconnect.ie
[3] Department of Computing Science and Mathematics, University of Stirling, Stirling FK9 4LA, UK
* Correspondence: vahid.akbari@stir.ac.uk

Abstract: One primary concern in the field of remote-sensing image processing is the precise classification of hyperspectral images (HSIs). Lately, deep-learning models have demonstrated cutting-edge results in HSI classification. Despite this, researchers continue to study and propose simpler, more robust models. This study presents a novel deep-learning approach, the iterative convolutional neural network (ICNN), which combines spectral–fractal features and classifier probability maps iteratively, aiming to enhance the HSI classification accuracy. Experiments are conducted to prove the accuracy enhancement of the proposed method using HSI benchmark datasets of Indian pine (IP) and the University of Pavia (PU) to evaluate the performance of the proposed technique. The final results show that the proposed approach reaches overall accuracies of 99.16% and 95.5% on the IP and PU datasets, respectively, which are better than some basic methods. Additionally, the end findings demonstrate that greater accuracy levels might be achieved using a primary CNN network that employs the iteration loop than with certain current state-of-the-art spatial–spectral HSI classification techniques.

Keywords: iterative convolutional neural network; fractal features; hyperspectral image; spatial–spectral features

check for updates

Citation: Asghari Beirami, B.; Alizadeh Pirbasti, M.; Akbari, V. SF-ICNN: Spectral–Fractal Iterative Convolutional Neural Network for Classification of Hyperspectral Images. *Appl. Sci.* **2024**, *14*, 7361. https://doi.org/10.3390/app14167361

Academic Editors: Hongmin Gao and Mingxiang Yang

Received: 25 June 2024
Revised: 5 August 2024
Accepted: 19 August 2024
Published: 21 August 2024

1. Introduction

Hyperspectral imaging is a powerful tool for studying surface phenomena because hyperspectral sensors record hundreds of spectral bands that correspond to the components that make up the Earth's surface. Hyperspectral image (HSI) classification aims to categorize similar materials and objects, traditionally using conventional classifiers like random forests and support vector machines (SVMs) [1,2]. However, due to the complex spectral behaviors of materials, these approaches based on spectral features often produce noisy results [3].

To enhance HSI classification, the spatial/textural information of pixels can be incorporated [4]. Various spatial features, including Gabor textural features, morphological profiles, attribute profiles, superpixels, and local binary patterns, have been widely used for this purpose [3,5,6]. Fractal features (FFs) represent a robust textural measure extensively employed in various image-processing applications [7,8]. Despite their considerable potential, FFs have yet to be thoroughly explored for remote-sensing HSI classification.

Deep-learning methods, in particular the convolutional neural network (CNN), have attracted a lot of attention recently. In addition to conventional handcrafted spectral and spatial features, a CNN can consider spatial and spectral information simultaneously, proving its effectiveness in HSI classification [9,10]. Additionally, certain studies have shown improved CNN performance by incorporating handcrafted spatial features alongside spectral features as inputs [11]. These advancements highlight the critical role of integrating multiple types of information to enhance the classification accuracy and robustness.

Despite the existing body of research on the subject, there remains a need for simpler and more robust deep-learning approaches in HSI classification. To address this, our study aims to propose a novel deep-learning method named spectral–fractal iterative CNN (SF-ICNN) that leverages spectral and spatial information at different stages of the classification process. Our motivation arises from the demonstrated benefits of using spectral features along with two kinds of spatial information—one from FFs and the other from probability map (PM) images—in an iterative manner to enhance the performance of CNN classifiers. We hypothesize that CNNs can significantly improve performance by repeatedly utilizing PMs [12]. The key novelty of our proposed method lies in integrating these two types of spatial features (FFs and PMs), resulting in remarkable performance improvements in CNN-based HSI classification. The following are the suggested method's primary contributions:

- The method combines FFs with spectral features as inputs to a CNN classifier. Experiments indicate that this novel strategy significantly enhances the classifier's performance.
- The method employs a novel combination of two types of spatial features, one from FFs and the other from PMs, iteratively to improve the CNN performance. This strategy enhances the classifier's performance statistically.
- Based on two benchmark HSI datasets, our experiment shows that the suggested deep-learning technique, with its more straightforward structure, outperforms certain state-of-the-art (SOTA) methods in terms of the classification accuracy within a reasonable processing time.

Section 2 of our study includes a review of recent research pertinent to the topics we are studying. Section 3 consists of a detailed explanation of our proposed method for HSI classification. In Section 4, the HSI datasets that we used are introduced, and then the results of the experimental tests and comparisons are analyzed. Section 5 includes our concluding remarks.

2. Literature Survey

Recent work has thoroughly examined the integration of spatial information for HSI categorization. Prior to classification, one popular method is to generate different spatial feature vectors. Rich spatial–spectral feature vectors, which are then fed into classifiers like SVM, are created by stacking these spatial feature vectors with spectral feature vectors. Based on past knowledge and research goals, users choose the kind of spatial feature generation method and its parameters. Some important spatial features include morphological profiles and their related topics [13–15], Gabor filters [16,17], local binary patterns [18,19], and superpixels [19]. Post-processing is another strategy to enhance HSI classification by incorporating spatial information after classification. For instance, Cao et al. and Tan et al. employed Markov random fields as a post-processing technique to improve the classification performance [20–22]. Kang et al. proposed a novel post-processing technique based on edge-preserving filters (EPFs), in that an EPF is applied to classified images to incorporate spatial features and reduce noise [23]. Zhong et al. developed a strategy combining spatial information from probability maps derived from SVM classifiers with EPFs iteratively to enhance the classification performance [24,25]. In addition to the mentioned methods that consider the spatial information in the pre-process or post-processing of the classification procedure, some recent classification approaches integrate spatial information innovatively during the classification process. In order to properly exploit spatial–spectral information, Jiang et al., for instance, developed a spatially aware collaborative representation approach that directly includes spatial information through a spatial regularization term in the representation objective function [26]. Wang et al. offered a unique low-rank representation model that is regularized by both locality and structure [27]. It incorporates a new distance metric that combines spectral and spatial characteristics and a structure constraint to increase the classification accuracy. Zhou et al. provided a brand-new spatial peak-aware collaborative representation technique for HSI classification that uses superpixel clusters' spectral–spatial data to input regularization terms for a brand-new closed-form solution [28].

Deep learning has attracted much interest in HSI analysis, in addition to conventional techniques, because of its innate capacity to take spatial pixel properties into account during categorization. Several deep-learning strategies have been put forth to improve the HSI classification accuracy using CNNs [29,30]. For example, ensemble deep-learning methods utilize local covariance descriptors from different window sizes around each pixel as inputs to a CNN classifier [31]. Another method proposed a new deep feature extraction method, named a random patches network, to extract deep features, and then the SVM classifier was fed with shallow and deep features to classify HSI [32]. For few-shot learning tasks, a multidimensional CNN model with convolution operators and an attention mechanism has also been utilized [33]. For HSI classification, a novel deep-learning model based on bidirectional long short-term memory (LSTM) and 3D convolutions has been presented [34]. To enhance the CNN performance, a hybrid deep-learning technique known as an attention-fused hybrid network combines an attention mechanism with a 3D and 2D CNN inception net [35]. Another approach involves first extracting spatial and spectral features and then using them as input to a CNN classifier [36]. A ghost module-based spectral network utilizing Ghost3D and Ghost2D modules, along with non-local operations, achieves superior performance in HSI classification [37]. A multi-feature fusion algorithm combines 3DCNN and a graph attention network to enhance the classification accuracy in HSI [38]. Finally, a neural network combines 3D CNN, 2D CNN, and Bi-LSTM to jointly learn spatial–spectral features, reduce parameters, and achieve superior performance in HSI classification [39].

This literature review shows that both traditional methods incorporating spatial features (either pre-processing or post-processing) and CNNs, which inherently consider spatial features, have been successfully utilized in HSI classification. Unlike traditional HSI methods that rely solely on spectral features, often leading to noisy results, or conventional deep-learning models that may not efficiently balance complexity and performance, this proposed innovative method of integrating spectral–fractal features and iteratively refining the CNN classification using PMs is unique and has not been explored in previous studies. In other words, this study leverages the capabilities of CNNs along with textural FFs and PMs for HSI classification—a combination not previously explored in the literature. Unlike the iterative SVM approach proposed by Zhong et al. [25], our spectral–fractal ICNN presents a more flexible and powerful framework for incorporating spatial and spectral information in HSI classification. Through two real HSI datasets, we show that our proposed technique achieves superior performance not only over [25] but also over numerous other state-of-the-art HSI classification methods.

3. Methodology

We initially present our SF-ICNN technique in this section. Subsequently, the principal elements of our suggested approach—PCA dimension reduction, fractal-based spatial features, and CNN—are presented in the following three subsections.

The suggested method's flowchart is displayed in Figure 1. The phases of this strategy are as follows:

1. The HSI's dimensionality is reduced in the first stage using the well-known principal components analysis (PCA) method. This method efficiently eliminates data redundancy and extracts compact, informative 3D data cubes.
2. The first three PCA characteristics are used to create FFs in the second stage.
3. The PCA approach is then used to reduce the fractal data cube's dimensions.
4. In the fourth step, the reduced fractal and spectral PCA features with five dimensions are stacked to create a combined 10-band spatio-spectral data cube. This number of

bands is empirically chosen as input to balance speed and accuracy. The important aspect is that the 10-band input with spectral and fractal features is a compact and information-rich representation that is used as input to the CNN.

5. This dense and rich spectral–fractal data cube is fed to a CNN with a simple structure, and primitive PMs are generated. The generated PMs have dimensions equal to the number of classes. The accuracy of the classification process is increased by the two-step use of spatial information resulting from the FFs and PMs.
6. The PMs are then stacked onto the combined data cube of spectral and fractal features. This new data cube is then used again as input for the CNN to create new PMs.
7. Step 6 is repeated until it reaches the desired results. It is worth noting that in each step of the iterative process, old PMs are replaced by new PMs. Also, the best iteration number is determined based on the classification accuracy of the validation samples.

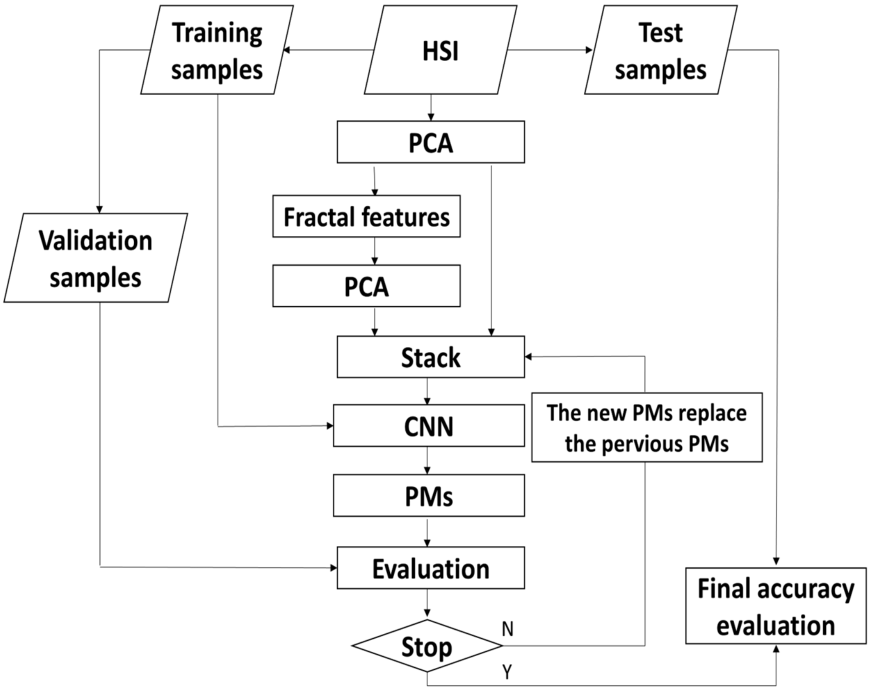

Figure 1. The suggested SF-ICNN classification method's flowchart. The proposed method consists of several key steps: spectral–fractal feature generation, CNN classification, and iterative integration of PMs into the feature data cube to improve the classification accuracy.

3.1. PCA Dimensionality Reduction

Finding a lower-dimensional representation of the original data while maintaining the greatest variance is the goal of PCA's popular linear dimensionality reduction approach [40]. The fundamental principle of PCA is finding the best transformation matrix W^* to project the original feature space X (spectral or fractal characteristics) onto a lower-dimensional subspace. The PCA transformation matrix W^* may be calculated mathematically by working through the following formula:

$$W^* = argmax_{\ W^T W = I}\ Tr\left(W^T\ Cov\ (X)\ W\right) \tag{1}$$

where:

- Tr stands for the trace operation, which is the sum of a matrix's diagonal members.
- $Cov\ (X)$ is a representation of the original data X's covariance matrix.

The objective function in Equation (1) seeks to find the transformation matrix W^* that maximizes the trace of the transformed covariance matrix $W^T\ Cov\ (X)\ W$. Accordingly, PCA aims to identify the directions in the original feature space that maximize the variance of the data. The main components, or the orthogonal directions in the original feature space that account for the most variation in the data, are found via PCA by solving this equation. Applications for the transformed data in the lower-dimensional subspace produced by PCA include feature extraction, dimensionality reduction, and data visualization.

3.2. Fractal Features

There are different methods to quantify the fractal textural measure of images. One notable method is the Pentland method, which is a directional method known for its excellent performance in generating spatial features for remote-sensing images [41]. This approach considers the spectral fluctuations in the region around each pixel to extract FFs. The self-similarity parameter (H) is calculated as the difference in the gray level between two pixels at a distance r, represented by E_r, and a constant value, represented as E_0 [42]:

$$E_r\, r^{-H} = E_0 \tag{2}$$

The following formula is then used to obtain the fractal dimension (FD):

$$FD = 1 - H \tag{3}$$

The following is what we obtain when we take the logarithm of Equation (2) on both sides:

$$log(E_r) - H\, log(r) = C_0 \tag{4}$$

Fitting a linear regression line between $log(E_r)$ and $log(r)$ yields the values of H and C_0, where H is the slope and C_0 is the intercept. Applying the Pentland technique involves enclosing each pixel in a moving window of size $L \times L$. Then, a random variable r, representing the difference between two random variables, a and b, is chosen within the range of $(0, L)$, meeting the constraint $0 < a < b \leq L$. The parameter e_r is then calculated as follows:

$$e_r = \frac{|S_b - S_a|}{r} + 1 \tag{5}$$

In this equation, the gray values of pixels at distances r and in the presumptive direction are represented by S_b and S_a. The average of the parameter e_r can be computed as follows, given the positive integer a:

$$\bar{e}_r = \frac{\sum_{a=0}^{L-r} e_r}{L - r + 1} \tag{6}$$

E_r is then given as:

$$E_r = \bar{e}_r \frac{L}{r} \tag{7}$$

The Pentland method considers four directions (vertical, horizontal, and two diagonals) to capture the spatial properties of pixels. This method produces distinct values for H and C_0 in every direction, providing exact information on the textures of pixels. In this study, FFs are generated using window sizes of 9 × 9, 17 × 17, and 25 × 25 to capture multiscale spatial information. We can enhance our understanding of image textures by employing this method with diverse orientations and window sizes.

3.3. CNN

CNNs are models inspired by the human visual system, making them highly effective for image-processing tasks. These deep-learning models are particularly adept at extracting complex features from HSI because they use spatial kernels. The weight-sharing technique of CNNs is a significant benefit as it drastically lowers the number of parameters and increases the computing efficiency. Convolutional layers, activation functions, pooling layers, and fully connected layers are the common layer types seen in a CNN design (Figure 2) [36].

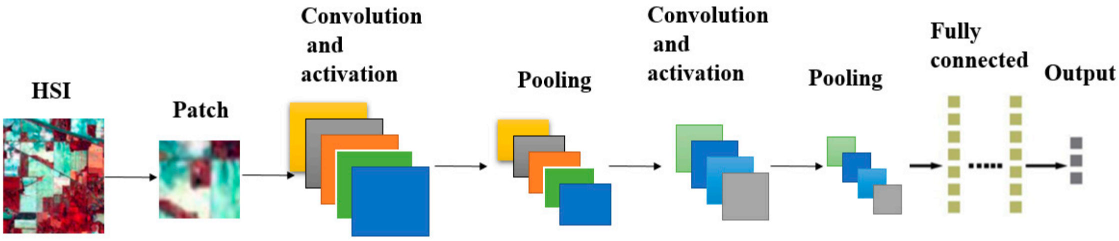

Figure 2. The CNN's main layers.

Convolutional layers work by moving a small filter across the input image, calculating the dot product between the filter and each image patch to detect specific features such as edges or textures. This process produces feature maps, in which the image content features are highlighted and enable the network to learn complex patterns. Multiple convolutional layers are usually used to extract deeper features. Early layers consider basic features such as edges and gradient information, while deeper layers detect more abstract features. By adding non-linearity, activation functions enable the network to simulate intricate connections found in the data. Tanh, sigmoid, and ReLU (rectified linear unit) are examples of standard activation functions. Pooling layers downsample the feature maps, reducing their size and complexity, which improves the network's efficiency and robustness to minor input variations. Standard pooling methods include max pooling, which selects the maximum value within a region, and average pooling, which computes the average value. These layers are exceptionally well suited for HSI classification since they generate position- and rotation-invariant characteristics. Typically, convolutional layers are followed by activation and pooling layers. Following a number of these layers, the obtained spectral–spatial information is combined into a 1-D vector by a fully connected layer, which is subsequently input into a softmax function.

The unknown parameters of a CNN, represented by the convolutional kernel values, are determined through training [43]. During the forward propagation phase, random weights are assigned to the kernels, which are convolved with the inputs and passed through the activation functions before reaching the pooling layer. By employing average or maximum operators, the pooling layer shrinks the feature maps' size. The final layer transforms the output into a deep feature vector, which is then processed by a fully connected layer to determine the pixel labels. The weight of the convolution kernel is modified in the backpropagation stage to minimize the loss performance. Batch normalization (BN) is also used to increase the accuracy and speed up network learning [44].

4. Datasets and Experimental Results

4.1. Datasets

Our study employed two widely recognized his datasets: Indian Pines (IP) and Pavia University (PU).

(1) The AVIRIS hyperspectral sensor gathered the IP dataset used in this study. This dataset includes 220 spectral bands covering the wavelength range of 0.4 to 2.5 µm, organized in 145 by 145 pixels with a spatial resolution of 20 m. A total of 200 bands remained for examination after 20 bands (104–108, 150–163, and 220) displaying water absorption and noise were eliminated to improve the data quality. A ground truth (GT) image and a color composite image created using this dataset are shown in Figure 3. According to Table 1, 10% of the labeled samples of each class are used to train the network, while the remaining data are used to evaluate the model performance.

Table 1. Sample count specification for the IP his dataset.

Class	Total Samples	Train/Validation	Test
1	46	4	42
2	1428	142	1286
3	830	83	747
4	237	23	214
5	483	48	435
6	730	73	657
7	28	3	25
8	478	47	431
9	20	3	17
10	972	97	875
11	2455	245	2210
12	593	59	534
13	205	20	185
14	1265	126	1139
15	386	38	348
16	93	9	84

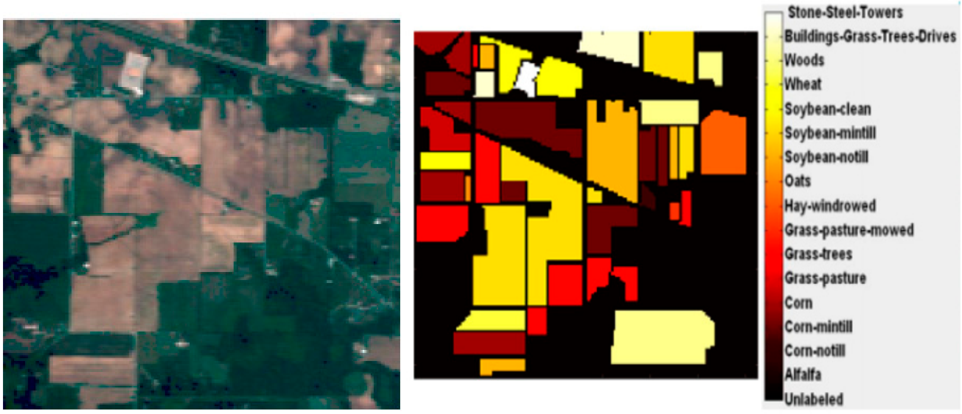

Figure 3. The IP dataset.

(2) The PU dataset was collected by the airborne hyperspectral sensor ROSIS-03 from the University of Pavia campus in Italy. This dataset has 103 spectral bands and 610 × 340 pixels with a spatial resolution of 1.3 m. The GT for this scene includes the nine distinct urban classes. Figure 4 shows a color composite of the HSI and the GT corresponding to the PU dataset. Following the specifications in Table 2, 1% of the labeled samples from each class were utilized for training, while the remainder samples were used to evaluate the model performance.

Table 2. Sample count specification for the PU his dataset.

Class	Total Samples	Train/Validation	Test
1	6631	66	6565
2	18,649	186	18,463
3	2099	20	2079
4	3064	30	3034
5	1345	13	1332
6	5029	50	4979
7	1330	13	1317
8	3682	36	3646
9	947	9	938

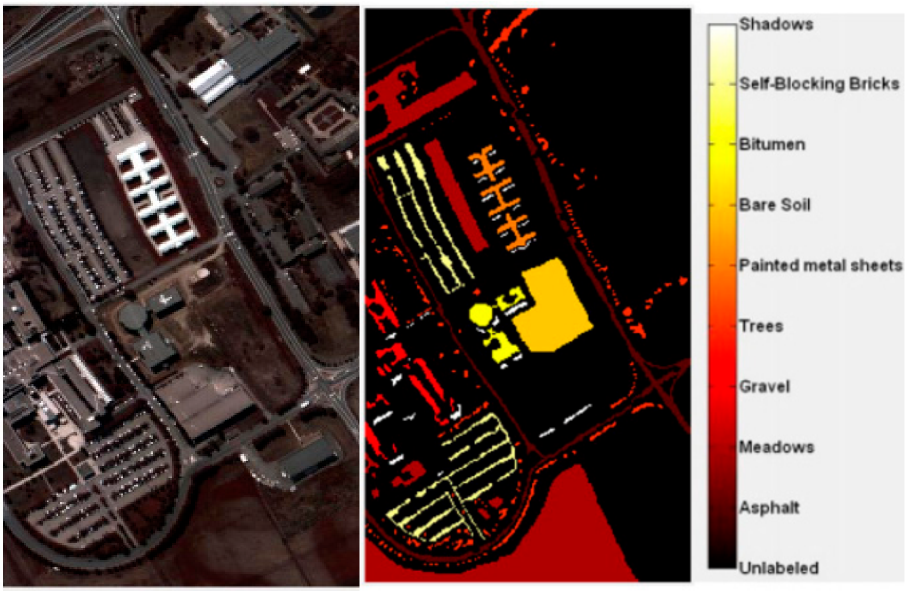

Figure 4. The PU dataset.

4.2. Analysis of the Results

The accuracy assessment in this study includes the use of two main classification indices [45], overall accuracy (OA) and kappa coefficient, both obtained from the confusion matrix, along with class-wise classification accuracy. These measures provide a comprehensive evaluation of the classifier's performance.

In order to strike a balance between accuracy and processing time, a 9 × 9 window size was chosen as the input patch size around each pixel for the CNN classifier. In our proposed SF-ICNN, we employed a simple CNN architecture with two main layers. In the first layer, 3 × 3 convolutional filters are used to generate 32 feature maps, followed by BN and ReLu activation and a 3 × 3 pooling layer. In the second layer, 3 × 3 convolutional filters are again used to generate 16 feature maps, followed by BN and ReLu activation and a 3 × 3 pooling layer. Finally, we employ a single layer of fully connected neurons, succeeded by a SoftMax function, to ascertain the labels of individual pixels. We also use ADAM as the optimizer, with the maximum number of epochs equal to 50. All the tests are performed with MATLAB 2020b on a fourth-generation Core i5 3.3 GHz CPU and 8 GB RAM, with the other CNN parameters set to their default values in MATLAB 2020b.

The effectiveness of the suggested approach is demonstrated by comparing the SF-ICNN method with a few baseline techniques in the following portion of this section. The next two subsections examine the effects of the training sample ratio and number of iterations on the classification outcomes of the suggested SF-ICNN. In the last part, we compare the performance of the suggested SF-ICNN with several SOTA techniques.

4.3. Comparison to Baselines

In this subsection, we compare our proposed SF-ICNN strategy with several baseline techniques, including spectral SVM (S-SVM), spectral–fractal SVM (SF-SVM), conventional spectral CNN (S-CNN), and conventional spectral–fractal CNN (SF-CNN). More information about these methods is as follows:

- S-SVM: In this method, spectral features are classified using kernel SVM, a commonly employed classifier in the literature.
- SF-SVM: In this method, stacked spectral–fractal features are classified using kernel SVM.
- S-CNN: In this method, spectral features are classified using the conventional CNN approach without iteration loops.

- SF-CNN: In this method, stacked spectral–fractal features are classified using the conventional CNN approach without iteration loops.

The classification accuracies of each method on the test samples are presented in Tables 3 and 4 for each dataset. The important findings of these tables are as follows:

- Comparing the OA and kappa of S-SVM and SF-SVM demonstrates that adding spatial FFs to spectral-based SVM can significantly improve the classification accuracy. This confirms the findings of previous studies about the efficacy of spatial features in enhancing the HSI classification outcomes.
- The results show that CNN-based methods outperform SVM-based methods in terms of the classification performance. This is due to the CNN's ability to consider spatial features during the classification procedures. However, the conclusion about the comparison of SF-SVM and S-CNN is mixed. In the IP dataset, SF-SVM performs better than S-CNN, but in the PU dataset, S-CNN performs better than SF-SVM. This may be because of the underlying complex land cover present in each scene.
- A comparison of the S-CNN and SF-CNN outcomes for the two datasets shows that although the CNN can take into account the spatial properties of pixels, its performance can still be enhanced by applying manually created FFs as inputs.
- Comparison of SF-ICNN and other techniques demonstrates that our proposed strategy outperforms all the baselines. Also, the results show that the proposed iterative strategy is efficient since SF-ICNN reaches higher classification accuracies than SF-CNN.
- The best baseline approach, SF-CNN, is compared with the suggested SF-ICNN method using the Z-score test [46]. The Z-score is a statistical measure used to compare the performance of two classification methods from a statistical point of view. The Z-score values between the SF-CNN approach (the best baseline method) and the SF-ICNN approach exceeded 1.96 for both datasets (2.87 for IP and 2.95 for PU), indicating significant accuracy improvements.
- In the context of class accuracies, the proposed SF-ICNN outperforms the other method in 15 of 16 classes for the IP dataset and 6 of 9 classes for the PU dataset. This further demonstrated the suggested method's effectiveness in raising the classification accuracy across the majority of classes.
- Figures 5 and 6 show the GT images and the final classification images of each classification method. As can be seen from the figures, our proposed SF-ICNN method shows a significant reduction in misclassified pixels in the classified images compared to the baseline methods.

Table 3. Results of the classification for the IP HSI dataset (the best results of each row are highlighted).

Class	S-SVM	SF-SVM	S-CNN	SF-CNN	SF-ICNN
1	50%	69.04%	69.04%	83.33%	97.61%
2	20.52%	82.03%	81.88%	93.23%	**98.83%**
3	70.01%	93.57%	86.07%	96.51%	**97.05%**
4	45.79%	94.85%	70.09%	98.13%	**99.06%**
5	66.20%	89.42%	82.98%	88.50%	**97.24%**
6	90.86%	96.49%	97.10%	99.84%	**99.84%**
7	76.00%	92.00%	84.00%	88.00%	**92.00%**
8	97.67%	93.03%	100%	98.37%	**100%**
9	41.17%	82.35%	58.82%	76.47%	**100%**
10	40.80%	82.74%	91.20%	98.97%	**100%**
11	39.45%	80.49%	94.84%	97.91%	**99.50%**
12	31.64%	75.09%	73.22%	88.76%	**99.62%**
13	95.67%	96.75%	98.37%	97.83%	**98.37%**
14	66.81%	97.27%	97.98%	98.15%	**99.73%**
15	29.59%	97.41%	75.00%	100%	**100%**
16	86.90%	**97.61%**	83.33%	89.28%	96.42%
OA	51.46%	87.32%	89.39%	96.33%	**99.16%**
Kappa	0.459	0.852	0.878	0.958	**0.99**

Table 4. Results of the classification for the PU HSI dataset (the best results of each row are highlighted).

Class	S-SVM	SF-SVM	S-CNN	SF-CNN	SF-ICNN
1	79.25%	84.03%	90.37%	**91.04%**	90.96%
2	94.71%	97.83%	97.71%	98.87%	**99.44%**
3	65.94%	76.91%	67.58%	72.10%	**81.14%**
4	70.79%	**97.82%**	84.97%	93.83%	94.75%
5	99.32%	99.77%	96.77%	96.32%	**100%**
6	44.90%	88.17%	80.91%	96.22%	**99.05%**
7	71.22%	70.15%	62.33%	64.00%	**82.30%**
8	77.15%	85.32%	87.35%	95.69%	**96.62%**
9	**99.89%**	92.21%	75.26%	79.53%	72.81%
OA	81.35%	91.53%	89.69%	93.80%	**95.50%**
Kappa	0.745	0.887	0.861	0.917	**0.943**

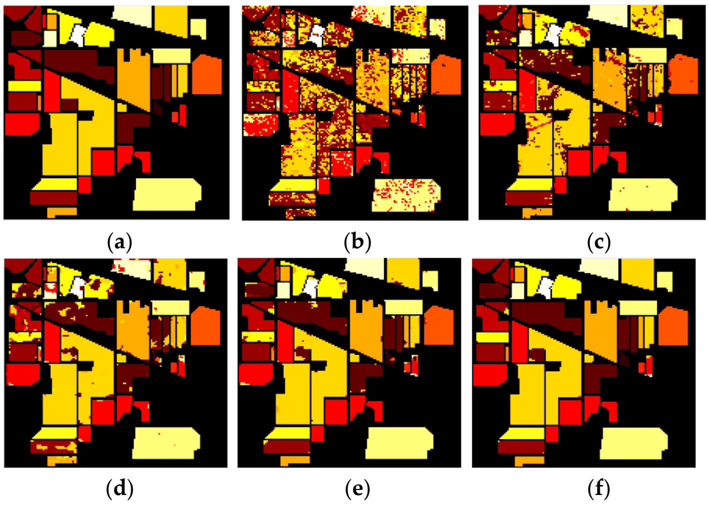

Figure 5. GT and other techniques' classification results for the IP HSI dataset. (**a**) GT. (**b**) S-SVM. (**c**) SF-SVM. (**d**) S-CNN. (**e**) SF-CNN. (**f**) Suggested SF-ICNN.

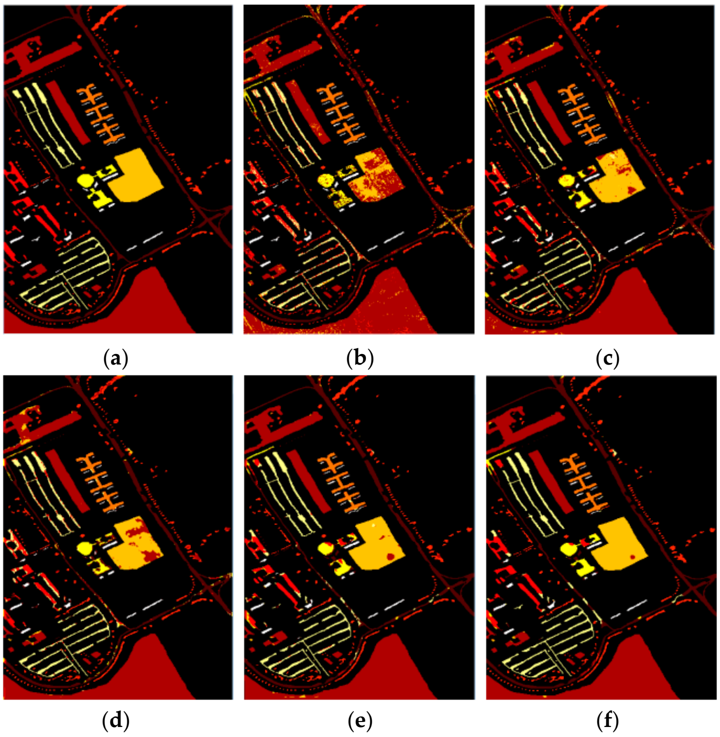

Figure 6. GT and other techniques' classification results for the PU HSI dataset. (**a**) GT. (**b**) S-SVM. (**c**) SF-SVM. (**d**) S-CNN. (**e**) SF-CNN. (**f**) Suggested SF-ICNN.

4.4. Analyzing the Impact of the Iteration Number

In this part, we look into how the number of iterations affected the SF-ICNN method's ultimate OA. By dividing the training samples, the accuracy results were acquired on validation samples, which helped identify the ideal number of iterations for the SF-ICNN approach. To accomplish this, we changed the number of iterations from zero, which means "no iteration applied," to six. Figure 7 reports the outcomes of the OA.

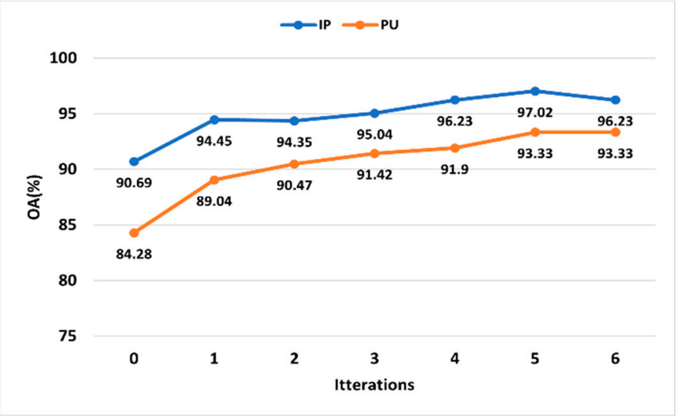

Figure 7. Variation of the suggested method's OAs according to the network's iteration number on validation samples.

As shown in Figure 7, the OA metric exhibited an initial increase as the number of iterations reached five. However, the next iterations did not result in any further improvement. Figure 7 shows that for both HSI datasets, the majority of the accuracy improvement happens in the first few rounds. It is important to note that the computation time increases by training more CNNs, resulting from more iteration loops. Therefore, it may be recommended to use a smaller number of iteration loops in studies where the computational time is critical.

4.5. Analyzing the Impact of the Training Sample Ratios

This section studies the sensitivity of the SF-ICNN technique to the number of training samples (#training). To accomplish this, the CNN classifier is trained using a different #training in the suggested SF-ICNN technique. The classification accuracy results for both HSI datasets are shown in Figure 8. This figure shows that, as expected, the classification accuracy of SF-ICNN increases with an increasing sample size. Furthermore, even with limited #training, the suggested approach continues to function well.

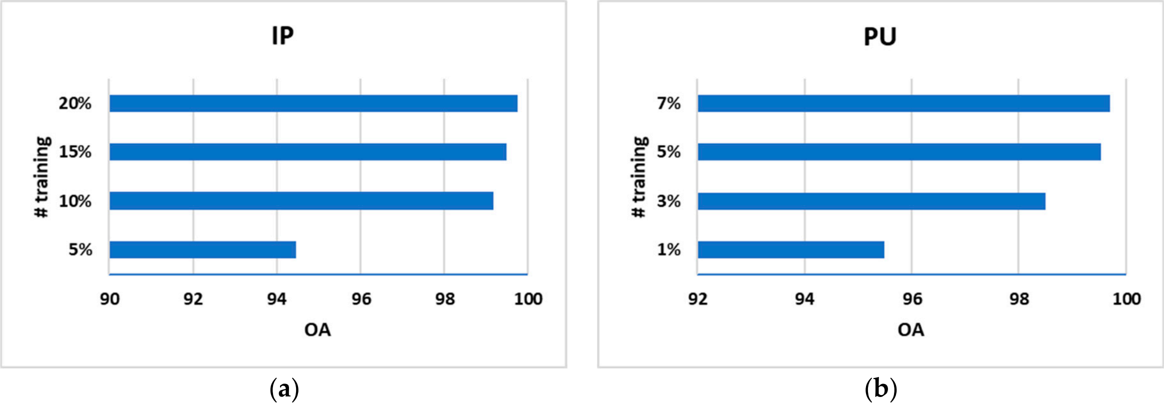

Figure 8. Variation of the OAs of the proposed SF-ICNN method in relation to the #training: (**a**) IP, and (**b**) PU.

4.6. Comparison to SOTA Methods

In this section's final experiment, we compare our proposed method to some previously introduced SOTA spatial–spectral HSI classification methods, namely Kang et al. [23], Xu et al. [32], Wang et al. [27], Liu et al. [33], Praveen et al. [34], Ren et al. [37], Sahoo et al. [39], and Zhong et al. [25]. The comparison experiment demonstrates the innovative strength and key advantages of the proposed SF-ICNN method. The results are shown in Table 5 and represent a fair comparison as each approach takes into account the same percentage of training and testing data. This table shows that on both the IP and PU datasets, the SF-ICNN methodology showed higher overall classification accuracies than various current SOTA spatial–spectral HSI classification approaches.

Table 5. Comparison with some SOTA methods (the best results of each column are highlighted).

Methods	OA	
	IP	PU
Kang et al. [23] *	91.33%	94.65%
Xu et al. [32] *	97.76%	94.21%
Wang et al. [27] **	95.63%	— ***
Liu et al. [33] **	—	92.23%
Praveen et al. [34] **	94.07%	—
Ren et al. [37] **	98.75%	95.19%
Sahoo et al. [39] **	98.62%	—
Zhong et al. [25] **	96.59%	—
Proposed SF-ICNN *	**99.16%**	**95.5%**

* The result is reported based on the implementation, ** The original paper's outcome is reported. *** The original study does not report the outcome.

The main innovative point of the SF-ICNN method lies in its iterative integration of spectral–fractal features and the CNN's PMs. By using spatial information in different stages of the classification, SF-ICNN can extract more discriminating features and refine the classification results iteratively. This is a significant improvement over previous SOTA methods, which only use spectral features, without considering the information contained in the FFs and PMs resulting from the CNN for HSI classification.

In respect of the computational time, the proposed SF-ICNN offers a processing time of around 100 s on the IP dataset, which may not be as fast as traditional non-deep approaches such as [23] or [32], which need a few seconds, but still better than some other non-deep approaches such as [27] by more than 300 s and deep-learning based techniques such as [34] that need more than 900 s for HSI classification. Our computation times can be improved by using advanced hardware such as GPU and parallel-processing techniques, improving its performance for larger-scale HSI classification tasks.

The proposed SF-ICNN method combines fractal–spectral features and CNNs with iterative loops to achieve high classification accuracy while maintaining computational efficiency. These results show the innovative nature and advantages of the SF-ICNN method and highlight its potential as a powerful and straightforward deep-learning framework for HSI classification.

5. Conclusions

Our study introduces a new deep-learning approach for the classification of HSI. Our proposed method combines spectral features and FFs as well as PMs in an iterative manner in the CNN classifier, resulting in highly accurate classification images. We conducted experiments using the IP and PU HSI datasets. We demonstrated that our SF-ICNN strategy outperformed basic methods by about 18% and 6.4% on average for the IP and PU datasets, respectively. The statistical analysis using the Z-score confirmed the statistical significance of this improvement. Additionally, we compared our SF-ICNN approach with some recently developed spatial and spectral classification methods and found that our method can achieve an acceptable processing time and superior classification accuracy. The proposed SF-ICNN framework also has potential applications beyond hyperspectral image classification. It can be used in various remote-sensing image analysis domains like land-cover mapping, urban planning, change detection, and environmental monitoring, where precise image classification is crucial. Even further, the proposed method can also be adopted to improve the CNN performance in texture analysis tasks such as medical image segmentation, anomaly detection, and face recognition. We intend to improve the SF-ICNN approach by making it less sensitive to the sample size by adopting a more advanced base model. Additionally, we aim to reduce the computational complexity by using advanced hardware such as GPUs and techniques such as parallel processing that increase efficiency while maintaining accuracy. These changes will enhance the performance and usability of the method.

Author Contributions: Conceptualization, B.A.B.; methodology, B.A.B.; software, B.A.B.; validation, B.A.B.; formal analysis, B.A.B.; writing—original draft preparation, B.A.B., M.A.P. and V.A.; writing—review and editing, B.A.B., M.A.P. and V.A.; supervision, V.A. All authors have read and agreed to the published version of the manuscript.

Funding: This publication's funding was provided by the University of Stirling Library.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: This work investigated publicly accessible datasets, which are available at http://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes (accessed on 4 August 2024).

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

References

1. Raczko, E.; Zagajewski, B. Comparison of support vector machine, random forest and neural network classifiers for tree species classification on airborne hyperspectral APEX images. *Eur. J. Remote Sens.* **2017**, *50*, 144–154. [CrossRef]
2. Sabat-Tomala, A.; Raczko, E.; Zagajewski, B. Comparison of support vector machine and random forest algorithms for invasive and expansive species classification using airborne hyperspectral data. *Remote Sens.* **2020**, *12*, 516. [CrossRef]
3. Asghari Beirami, B.; Mokhtarzade, M. Hyperspectral image classification using multiple weighted local kernel matrix descriptors. *Int. J. Remote Sens.* **2022**, *43*, 5280–5305. [CrossRef]
4. Ahmad, M.; Shabbir, S.; Roy, S.K.; Hong, D.; Wu, X.; Yao, J.; Khan, A.M.; Mazzara, M.; Distefano, S.; Chanussot, J. Hyperspectral image classification—Traditional to deep models: A survey for future prospects. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *15*, 968–999. [CrossRef]
5. Dundar, T.; Ince, T. Sparse representation-based hyperspectral image classification using multiscale superpixels and guided filter. *IEEE Geosci. Remote Sens. Lett.* **2018**, *16*, 246–250. [CrossRef]
6. Imani, M.; Ghassemian, H. An overview on spectral and spatial information fusion for hyperspectral image classification: Current trends and challenges. *Inf. Fusion* **2020**, *59*, 59–83. [CrossRef]
7. Sun, W.; Xu, G.; Gong, P.; Liang, S. Fractal analysis of remotely sensed images: A review of methods and applications. *Int. J. Remote Sens.* **2006**, *27*, 4963–4990. [CrossRef]
8. Panigrahy, C.; Seal, A.; Mahato, N.K. Image texture surface analysis using an improved differential box counting based fractal dimension. *Powder Technol.* **2020**, *364*, 276–299. [CrossRef]
9. Audebert, N.; Le Saux, B.; Lefèvre, S. Deep learning for classification of hyperspectral data: A comparative review. *IEEE Geosci. Remote Sens. Mag.* **2019**, *7*, 159–173. [CrossRef]
10. Paoletti, M.E.; Haut, J.M.; Plaza, J.; Plaza, A. Deep learning classifiers for hyperspectral imaging: A review. *ISPRS J. Photogramm. Remote Sens.* **2019**, *158*, 279–317. [CrossRef]
11. Zhao, W.; Li, S.; Li, A.; Zhang, B.; Li, Y. Hyperspectral images classification with convolutional neural network and textural feature using limited training samples. *Remote Sens. Lett.* **2019**, *10*, 449–458. [CrossRef]
12. Neshatpour, K.; Homayoun, H.; Sasan, A. Icnn: The iterative convolutional neural network. *ACM Trans. Embed. Comput. Syst. (TECS)* **2019**, *18*, 1–27. [CrossRef]
13. Amiri, K.; Imani, M.; Ghassemian, H. Empirical Mode Decomposition Based Morphological Profile For Hyperspectral Image Classification. In Proceedings of the 2023 6th International Conference on Pattern Recognition and Image Analysis (IPRIA), Qom, Iran, 14–16 February 2023; pp. 1–6.

14. Lu, Q.; Xie, Y.; Wei, L.; Wei, Z.; Tian, S.; Liu, H.; Cao, L. Extended Attribute Profiles for Precise Crop Classification in UAV-Borne Hyperspectral Imagery. *IEEE Geosci. Remote Sens. Lett.* **2024**, *21*, 2500805. [CrossRef]
15. Hong, D.; Wu, X.; Ghamisi, P.; Chanussot, J.; Yokoya, N.; Zhu, X.X. Invariant attribute profiles: A spatial-frequency joint feature extractor for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 3791–3808. [CrossRef]
16. Huang, K.-K.; Ren, C.-X.; Liu, H.; Lai, Z.-R.; Yu, Y.-F.; Dai, D.-Q. Hyperspectral image classification via discriminant Gabor ensemble filter. *IEEE Trans. Cybern.* **2021**, *52*, 8352–8365. [CrossRef]
17. Cruz-Ramos, C.; Garcia-Salgado, B.P.; Reyes-Reyes, R.; Ponomaryov, V.; Sadovnychiy, S. Gabor features extraction and land-cover classification of urban hyperspectral images for remote sensing applications. *Remote Sens.* **2021**, *13*, 2914. [CrossRef]
18. Huang, W.; Huang, Y.; Wu, Z.; Yin, J.; Chen, Q. A multi-kernel mode using a local binary pattern and random patch convolution for hyperspectral image classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 4607–4620. [CrossRef]
19. Huang, W.; Huang, Y.; Wang, H.; Liu, Y.; Shim, H.J. Local binary patterns and superpixel-based multiple kernels for hyperspectral image classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 4550–4563. [CrossRef]
20. Tan, X.; Xue, Z.; Yu, X.; Sun, Y.; Gao, K. Hyperspectral image classification with deep 3D capsule network and Markov random field. *IET Image Process.* **2022**, *16*, 79–91. [CrossRef]
21. Cao, X.; Xu, L.; Meng, D.; Zhao, Q.; Xu, Z. Integration of 3-dimensional discrete wavelet transform and Markov random field for hyperspectral image classification. *Neurocomputing* **2017**, *226*, 90–100. [CrossRef]
22. Cao, X.; Zhou, F.; Xu, L.; Meng, D.; Xu, Z.; Paisley, J. Hyperspectral image classification with Markov random fields and a convolutional neural network. *IEEE Trans. Image Process.* **2018**, *27*, 2354–2367. [CrossRef]
23. Kang, X.; Li, S.; Benediktsson, J.A. Spectral–spatial hyperspectral image classification with edge-preserving filtering. *IEEE Trans. Geosci. Remote Sens.* **2013**, *52*, 2666–2677. [CrossRef]
24. Zhong, S.; Chang, C.-L.; Zhang, Y. Iterative edge preserving filtering approach to hyperspectral image classification. *IEEE Geosci. Remote Sens. Lett.* **2018**, *16*, 90–94. [CrossRef]
25. Zhong, S.; Chang, C.-I.; Zhang, Y. Iterative support vector machine for hyperspectral image classification. In Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; pp. 3309–3312.
26. Jiang, J.; Chen, C.; Yu, Y.; Jiang, X.; Ma, J. Spatial-aware collaborative representation for hyperspectral remote sensing image classification. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 404–408. [CrossRef]
27. Wang, Q.; He, X.; Li, X. Locality and structure regularized low rank representation for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2018**, *57*, 911–923. [CrossRef]
28. Zhou, C.; Tu, B.; Ren, Q.; Chen, S. Spatial peak-aware collaborative representation for hyperspectral imagery classification. *IEEE Geosci. Remote Sens. Lett.* **2021**, *19*, 5506805. [CrossRef]
29. 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]
30. Ullah, F.; Ullah, I.; Khan, R.U.; Khan, S.; Khan, K.; Pau, G. Conventional to Deep Ensemble Methods for Hyperspectral Image Classification: A Comprehensive Survey. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2024**, *17*, 3878–3916. [CrossRef]
31. He, N.; Paoletti, M.E.; Haut, J.M.; Fang, L.; Li, S.; Plaza, A.; Plaza, J. Feature extraction with multiscale covariance maps for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2018**, *57*, 755–769. [CrossRef]
32. 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]
33. Liu, J.; Zhang, K.; Wu, S.; Shi, H.; Zhao, Y.; Sun, Y.; Zhuang, H.; Fu, E. An investigation of a multidimensional CNN combined with an attention mechanism model to resolve small-sample problems in hyperspectral image classification. *Remote Sens.* **2022**, *14*, 785. [CrossRef]
34. Praveen, B.; Menon, V. A bidirectional deep-learning-based spectral attention mechanism for hyperspectral data classification. *Remote Sens.* **2022**, *14*, 217. [CrossRef]
35. Ahmad, M.; Khan, A.M.; Mazzara, M.; Distefano, S.; Roy, S.K.; Wu, X. Attention mechanism meets with hybrid dense network for hyperspectral image classification. *arXiv* **2022**, arXiv:2201.01001.
36. Sharifi, O.; Mokhtarzadeh, M.; Asghari Beirami, B. A new deep learning approach for classification of hyperspectral images: Feature and decision level fusion of spectral and spatial features in multiscale CNN. *Geocarto Int.* **2022**, *37*, 4208–4233. [CrossRef]
37. Ren, Y.; Jin, P.; Li, Y.; Mao, K. An efficient hyperspectral image classification method for limited training data. *IET Image Process.* **2023**, *17*, 1709–1717. [CrossRef]
38. Bhatti, U.A.; Huang, M.; Neira-Molina, H.; Marjan, S.; Baryalai, M.; Tang, H.; Wu, G.; Bazai, S.U. MFFCG–Multi feature fusion for hyperspectral image classification using graph attention network. *Expert Syst. Appl.* **2023**, *229*, 120496. [CrossRef]
39. Sahoo, A.R.; Chakraborty, P. Hybrid CNN Bi-LSTM neural network for Hyperspectral image classification. *arXiv* **2024**, arXiv:2402.10026.
40. Beirami, B.A.; Mokhtarzade, M. Band grouping SuperPCA for feature extraction and extended morphological profile production from hyperspectral images. *IEEE Geosci. Remote Sens. Lett.* **2020**, *17*, 1953–1957. [CrossRef]
41. Pentland, A.P. Fractal-based description of natural scenes. *IEEE Trans. Pattern Anal. Mach. Intell.* **1984**, *PAMI–6*, 661–674. [CrossRef]
42. Beirami, B.A.; Mokhtarzade, M. Spatial-spectral classification of hyperspectral images based on multiple fractal-based features. *Geocarto Int.* **2022**, *37*, 231–245. [CrossRef]

43. Beirami, B.A.; Mokhtarzade, M. A new deep learning approach for hyperspectral image classification based on multifeature local kernel descriptors. *Adv. Space Res.* **2023**, *72*, 1703–1720. [CrossRef]
44. Bjorck, N.; Gomes, C.P.; Selman, B.; Weinberger, K.Q. Understanding batch normalization. *arXiv* **2018**, arXiv:1806.02375.
45. Beirami, B.A.; Mokhtarzade, M. SVM classification of hyperspectral images using the combination of spectral bands and Moran's I features. In Proceedings of the 2017 10th Iranian Conference on Machine Vision and Image Processing (MVIP), Isfahan, Iran, 22–23 November 2017; pp. 139–144.
46. Mirzapour, F.; Ghassemian, H. Moment-based feature extraction from high spatial resolution hyperspectral images. *Int. J. Remote Sens.* **2016**, *37*, 1349–1361. [CrossRef]

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