



Article

Tri-CNN: A Three Branch Model for Hyperspectral Image Classification

Mohammed Q. Alkhatib ^{1,*}, Mina Al-Saad ^{1,t}, Nour Aburaed ^{1,2}, Saeed Almansoori ³, Jaime Zabalza ², Stephen Marshall ² and Hussain Al-Ahmad ¹

¹ Collage of Engineering and IT, University of Dubai, Dubai P.O. Box 14143, United Arab Emirates

² Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow G1 1XW, UK

³ Mohammed Bin Rashid Space Center, Dubai P.O. Box 211833, United Arab Emirates

* Correspondence: mqalkhatib@ieee.org

† These authors contributed equally to this work.

Abstract: Hyperspectral Image (HSI) classification methods that use Deep Learning (DL) have proven to be effective in recent years. In particular, Convolutional Neural Networks (CNNs) have demonstrated extremely powerful performance in such tasks. However, the lack of training samples is one of the main contributors to low classification performance. Traditional CNN-based techniques under-utilize the inter-band correlations of HSI because they primarily use 2D-CNNs for feature extraction. Contrariwise, 3D-CNNs extract both spectral and spatial information using the same operation. While this overcomes the limitation of 2D-CNNs, it may lead to insufficient extraction of features. In order to overcome this issue, we propose an HSI classification approach named Tri-CNN which is based on a multi-scale 3D-CNN and three-branch feature fusion. We first extract HSI features using 3D-CNN at various scales. The three different features are then flattened and concatenated. To obtain the classification results, the fused features then traverse a number of fully connected layers and eventually a softmax layer. Experimental results are conducted on three datasets, Pavia University (PU), Salinas scene (SA) and GulfPort (GP) datasets, respectively. Classification results indicate that our proposed methodology shows remarkable performance in terms of the Overall Accuracy (OA), Average Accuracy (AA), and Kappa metrics when compared against existing methods.

Keywords: hyperspectral classification; Convolutional Neural Networks; deep learning; feature fusion



Citation: Alkhatib, M.Q.; Al-Saad, M.; Aburaed, N.; Almansoori, S.; Zabalza, J.; Marshall, S.; Al-Ahmad, H. Tri-CNN: A Three Branch Model for Hyperspectral Image Classification. *Remote Sens.* **2023**, *15*, 316. <https://doi.org/10.3390/rs15020316>

Academic Editors: Sidike Paheding and Ashraf Saleem

Received: 4 December 2022

Revised: 30 December 2022

Accepted: 2 January 2023

Published: 5 January 2023



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

1. Introduction

Hyperspectral Imaging (HSI) systems collect spectral and spatial image data simultaneously in hundreds of narrow contiguous spectral bands, from visible to infrared wavelengths. Applications for HSI remote sensing are numerous, such as environmental monitoring [1,2], agriculture [3–6], mineralogy [7–9], and surveillance [10–12]. Each pixel in the image consists of hundreds of elements measuring the reflected or emitted energy as a function of wavelength, known as the spectral response. Unlike common or normal images, hyperspectral ones carry information about the material's chemical and physical properties since the interaction of a substance with light depends on its atomic and molecular structure, HSIs are rich in spectral information that reflect the material's chemical and physical properties, which makes image classification tasks easier.

HSI classification has become one of the most active research topics in the field of HSI analysis [13,14] and has drawn the attention of researchers in the remote sensing field. It is defined as the task of assigning each pixel of the HSI to one certain class based on its spectral characteristics. There are challenges faced in most of the HSI classification tasks due to high-dimensional characteristics of HSI data and the similarity of spectra and mixed pixels. The following challenges need to be addressed:

1. The lack of ground-truth data or labeled samples: A typical challenge in remote sensing is that images are acquired from a far distance, which makes it difficult to

distinguish the materials by the a simple observation. In many applications, scientists need to go to the field of study to observe the materials in the scene from a close distance.

2. HSI has high dimensionality: This is related to the large number of channels (or bands) that HSI has. As the number grows, the data distribution becomes sparse and hard to model, which is also known as the *curse of dimensionality* problem, a term that was introduced by Bellman et al. [15]. However, multiple adjacent bands are similar and present redundant information, which enables the ability to use dimensionality reduction techniques to reduce the amount of involved data and speed up the classification process.
3. Low spatial quality: Sensors suffer from a trade-off that allows capturing images either with high spatial resolution or high spectral resolution. Thus, HSIs generally have relatively low spatial resolution when compared to natural images.
4. Spectral variability: The spectral response of each observed material can be significantly affected by atmospheric variations, illumination, or environmental conditions.

Machine Learning (ML) techniques were first used in the early research studies of HSI classification. These techniques use pixel-wise classification methods to classify each pixel in HSI data based on spectral information. Support Vector Machine (SVM) [16], Multinomial Logistic Regression (MLR) [17], Random Forest (RF) [18], and K-Nearest Neighbor (KNN) [19] are the most common examples of traditional approaches. However, these methods ignore the spatial information and rely solely on the spectral information of the pixels, which affects the classification performance. In order to tackle this challenge, Deep Learning (DL) techniques were utilized to obtain the spatial features along with the spectral ones. The DL approaches used for HSI classification are mainly Convolutional Neural Networks (CNNs), Autoencoders [20], Deep Belief Networks [21], Generative Adversarial Networks (GANs) [22,23], and Recurrent Neural Networks (RNNs) [24,25].

Related Work

As previously mentioned, the high-dimensionality characteristic of HSI data makes HSI classification a challenging task. Therefore, dimensionality reduction is a necessary pre-processing step prior to the classification task in order to reduce the redundancy and complexity of data. Numerous studies have been conducted to find a solution to the HSI's curse of dimensionality problem. So far, these dimensionality reduction techniques can be grouped into two main categories; feature extraction and band selection. Feature extraction approaches can be divided further into two groups; supervised and unsupervised feature extraction. Supervised feature extraction approaches utilize label information to maximize between-class separation, such as Linear Discriminant Analysis (LDA) [26], and Fisher's Linear Discriminant Analysis (FLDA) [27]. As for unsupervised feature extraction methods, they include Principal Component Analysis (PCA) [28,29] and Kernel PCA (KPCA) [30,31] that extract useful linear and nonlinear features, respectively, from HSI data. As for band selection dimensionality reduction, it is implemented by selecting the most descriptive and distinctive band subset from the original band space of the HSIs according to certain criteria or search strategy. Band selection approaches can be categorized into supervised, semi-supervised and unsupervised.

In recent years, DL methods, specifically, Convolutional Neural Networks (CNNs), have received an increased attention and have been widely employed in various computer vision tasks. CNNs can learn and extract deep features from raw data through a sequence of hierarchical layers; shallower layers can extract edges and texture features, whilst the deeper layers extract more complex features. Thus, CNNs have been extended and introduced for HSI classification [32]. The authors in [33] utilized PCA as pre-processing technique to handle data dimensionality. Then, the reduced components were fed to three consecutive 2D convolutional layers for feature mapping, followed by a flattening layer to convert all 2D feature maps into 1D feature vectors, which were then passed to fully connected layers to perform the final classification of the HSI. Similarly, in [34], the authors presented a CNN

approach based on information measure and enhanced spectral information where both spectral and spatial information are measured and used as an input to the CNN classifier. Dimensionality reduction is also employed using entropy calculation in order to select the three most informative spectra from HSI bands. Accurate classification of HSI was achieved under the reduced computational complexity.

In another study [35], Li et al. proposed a two-stream spectral and spatial feature extraction and fusion network based on Two-dimensional CNN (2D-CNN) architecture. The purpose of using the two extraction networks, shallow and deep, with different depth is to obtain multi-scale spectral, local, and global spatial information. Furthermore, an attention mechanism based on Squeeze and Excitation (SE) networks was adopted in this study to improve the capability of extracting useful spectral–spatial features and to overcome the lack of labeled training data. In addition to 2D-CNNs, Three-dimensional Convolutional Neural Networks (3D-CNNs) have been extensively studied for HSI classification as well. The authors in [36] presented a novel 3D-CNN approach for HSI classification, in which both spectral and spatial information are utilized to boost the classification performance. In a similar approach, the researchers in [37] introduced a 3D-CNN model for HSI classification, the HSI is first split into small overlapping 3D patches, which are then processed to create 3D feature maps utilizing a 3D kernel function over multiple contiguous bands of the spectral information in an effective manner. Since then, there has been an increasing number of HSI studies based on 2D-CNN or 3D-CNN [38–40].

However, there are a few drawbacks to utilizing either 2D-CNN or 3D-CNN for HSI classification. In the case of 2D-CNN architectures, even though they excel at obtaining spatial information, they fall short with regards to extracting good discriminating or informative feature maps from the spectral dimensions. On the other hand, 3D-CNNs are considered computationally expensive as a large number of 3D convolution operations are invoked. Deep 3D-CNNs require more training samples, which is not possible due to the limited dataset size provided by the publicly available HSI datasets. Furthermore, most of the common 3D-CNN-based approaches consist of stacked 3D convolutions within their architecture. Hence, they cannot optimize the estimation loss directly through such a nonlinear structure [41].

In order to tackle these limitations, many authors have introduced a hybrid approach. For example, in [42], the authors introduced a hybrid model called HybridSN, which fuses 2D-CNN with 3D-CNN to effectively extract both spectral and spatial feature maps from HSI. This improves the accuracy of the classification results. Similarly, in [43], the authors used 2D- and 3D-CNNs to extract spatial and spectral features in HSI, respectively. An attention mechanism is also utilized to combine these two kinds of features. In another study [44], the researchers presented a 3D fast learning block followed by 2D convolution for HSI data. The 3D block consists of 3D depthwise separable convolution and fast convolution blocks. The proposed approach shows superior performance in terms of training time and learning parameters when compared to existing 2D-CNN and 3D-CNN architectures. In another study [41], a hybrid model, named Synergistic CNN (SycNN) is proposed. The model also combines 2D and 3D convolution to extract the spectral and spatial information in turn. Moreover, the least important features are removed by introducing a 3D attention mechanism before the fully-connected layer. To tackle the issue of a limited number of labeled samples, Zhang et al. [45] proposed a 3D lightweight CNN model by introducing two different transfer learning strategies; cross-sensor strategy and cross-modal strategy. The first strategy transfers models between HSI datasets captured by the same or different sensor, whereas the second one uses a pre-trained 3D-CNN model on 2D RGB, or natural, image data sets to train 3D HSI through fine tuning the model. The proposed approach also reduces the number of trainable parameters while improving the classification's accuracy. Xu et al. [46] introduced a robust self-ensembling network (RSEN) that consists of two subnetworks; base network and an ensemble network. Experiments show good performance with very limited labeled training samples when compared to state-of-the-art methods. Although the CNN-based approaches have proven their efficiency

in HSI classification task, they still encounter a series of challenges such as, the receptive field is limited, information is lost in down-sampling layer, and a lot of computing resources are consumed for training deep networks. To overcome this challenge, transformer-based network is used to replace the traditional CNN to perform HSI classification tasks. The authors in [47] introduced a multilevel spectral–spatial transformer network (MSTNet) for hyperspectral classification. The proposed model incorporates an image-based classification framework for HSI in order to extract richer global information. Experimental results show that the proposed approach outperforms other CNN-based methods for classification.

Many researchers employ a multi-scale approach. It is an effective way for improving the classification accuracy due to the different sizes of land covers, which can capture more important information [48]. The spatial spectral unified network (SSUN) proposed in [49] utilizes this strategy. This approach combines the long short-term memory (LSTM) model as the spectral feature extractor with 2D-CNN for extracting spatial features and integrating the spectral Finite Element (FE), spatial FE, and classifier training into one unified network. By extracting multi-scale spectral–spatial features simultaneously (since diverse layers in CNN will produce different scale features) the Multi-Scale Convolutional Neural Network (MSCNN) [49] encompasses multi-scale characteristics. The experiments showed that the full use of both spectral and spatial information can considerably boost the classification accuracy. Although the existing spectral–spatial based methods show excellent performance in HSI classification but most of them are focusing on processing the pixels adjacent to the central pixels and neglect the implicit spatial information between features which in turn affects the performance of the classification model. To tackle this problem, Parallel Multi-Input Mechanism CNN (PMI-CNN) is introduced by Zhong et al. [50]. This model employs a parallel multi-input approach by using four parallel convolution branches to extract spatial features with different levels or window sizes in HSI data. Each branch of the network adopts four different convolution kernels to carry out multi-scale convolution on the inputs, of which 1×1 convolution kernel is used to extract spectral features of HSIs that will be fused with spatial information to classify HSI data. Their approach shows superiority in terms of overall accuracy and kappa metrics.

The following is a list of this paper’s main contributions:

1. A new model that incorporates feature extraction at different filter scales is proposed, which effectively improves the classification performance.
2. A deep feature-learning network based on various dimensions with varied kernel sizes is created so that the filters may concurrently capture spectral, spatial, and spectral–spatial properties in order to more effectively use the spatial-spectral information in HSIs. The proposed model has a better ability to learn new features than other models that are already in use, according to experimental results.
3. The proposed model not only outperforms existing methods in the case of smaller number of samples, but also achieves better accuracy with enough training samples, according to statistical results from detailed trials on three HSIs that will be reported and discussed in the next sections.
4. The impact of different percentages of training data on model performance in terms of OA, AA and kappa metrics are also examined.

The rest of the paper is organized as follows: Section 2 provides a description of the proposed model, Section 3 explains the experimental procedure and analyzes the findings, and Section 4 summarizes this paper and presents the future direction of this research.

2. Methodology

2.1. Framework of the Proposed Model

The methodology consists of two parts; data pre-processing and the proposed CNN. The details of these parts are explained in the next subsections.

2.1.1. Data Preprocessing

As stated in Section 1, HSI data must be pre-processed to reduce their dimensionality before processing them with DL models. In this research, PCA [28] is used to fulfill this role. PCA is one of the widely used unsupervised techniques for dimensionality reduction and feature extraction. It is based on the fact that adjacent bands of HSI exhibit strong correlation and frequently represent the same information about the object. It is carried out using an orthogonal transformation, which considers a set of variables that could be correlated and transforms or projects them into a set of values that are linearly uncorrelated and are known as Principal Components (PCs). Assuming the size of raw HSI cube is $H \times W \times C$, where H and W are the height and width of the HSI, respectively, and C is the depth (number of spectral bands). PCA works by downsizing the raw HSI cube's depth from C to D in such a way that the reduced HSI cube will be of size $H \times W \times D$, where D is the new spectral dimension and $D \ll C$.

2.1.2. Architecture of the Proposed Tri-CNN

Currently, the majority of the models use 2D-CNN or 3D-CNN architectures for HSI classification tasks. 2D-CNN has the ability to capture the spatial information; however, it ignores the rich spectral information provided by the HSI. On the other hand, 3D-CNN can extract both spectral and spatial information simultaneously, which may lead to insufficient extraction of features. In order to solve these shortcomings, the spatial feature extractor and the spectral-spatial feature extractor are combined along with spectral-only feature extractor to form a three-branch feature fusion network. The purpose of adding this path to the architecture is to enhance the extraction of spectral feature characteristics and improve the overall feature extraction process.

It is worth mentioned that the spatial feature extractor was designed with multiple $3 \times 3 \times 1$ 3D convolution filters. A number of smaller convolution kernels can extract features more effectively with lower computational cost [51], when compared to large convolution kernels. To extract the spectral features, multiple $1 \times 1 \times 3$ 3D convolution filters were used. Using a convolution kernel with of this size makes the model only focus on a specific pixel when extracting the spectral features, This can deal with the problem of adding unrelated information and improve the model's classification performance.

The flattening layer is used to convert each feature extracted by each branch into a one-dimensional vector. Then, all vectors are fused to produce the final classification results. This is achieved by concatenating the three one-dimensional features into a new one-dimensional vector, whose dimension number equals the total of the three feature dimensions. The flattened features at each branch are of different length (as shown in Section 3.3, adding them directly is not possible. Hence, concatenation was the most suitable choice for feature fusion.

The detailed distribution and parameters of each layer in the proposed approach are shown in Figure 1. The first branch has 2 layers of 3D-CNN with 64 filters and the kernel size is set to $1 \times 1 \times 3$ to capture spectral features. The second branch captures the spatial features by setting the kernel size to $3 \times 3 \times 1$ with the same number of layers and filters in the first branch. While the third branch is designed to capture both spectral and spatial features simultaneously by having 3D-CNN layers and using 64 filters and setting the kernel size of each filter to $3 \times 3 \times 3$. Due to the fact that the pooling layer causes loss of information, it has not been used within the network architecture. In addition, the network structure and hyperparameters described above perform well on all HSI datasets used in this research.

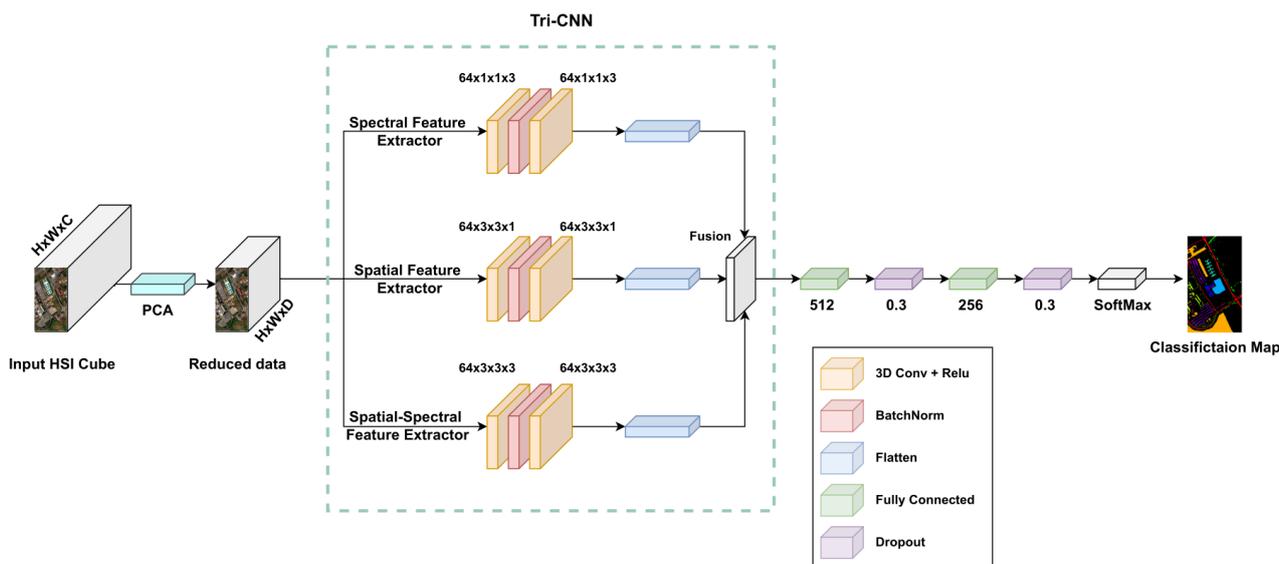


Figure 1. Framework of the proposed model.

2.2. Loss Function

Cross-Entropy (CE) is used as the loss function in this paper. The formula is given by:

$$Loss_{CE} = -\frac{1}{M} \sum_{m=1}^M \sum_{l=1}^L y_l^m \log(\hat{y}_l^m) \quad (1)$$

where y_l^m and \hat{y}_l^m are the reference and predicted labels, respectively, M and L are the overall number of small batch samples and land cover categories, respectively.

3. Experiments and Analysis

In this section, we will present details of the HSI datasets used in this paper. Secondly, we will introduce the the experimental configuration and parameter analysis. Then, we will apply the ablation experiments on the proposed model. Finally, we will perform a comparison between the proposed model against existing methods to prove the model superiority and its effectiveness both quantitatively and qualitatively.

3.1. Datasets

In our experiments, two of the most commonly used HSI datasets are adopted, namely, Pavia University and Salinas. Additionally, the Gulfport of Mississippi dataset is also used as well, although that it has not been widely used for HSI classification tasks, it is of great interest as it is small in size and consists of 72 spectral bands only. Each of these dataset has its own specifications, which are listed as follows:

1. Pavia University (PU): This is a scene acquired with the Reflective Optics Imaging Spectrometer Sensor (ROSIS) during a flight campaign over Pavia, northern Italy. The image consists of 103 spectral bands with wavelengths ranging from 0.43 to 0.86 μm and a spatial resolution of 1.3 m. The size of Pavia University is 610×340 pixels. Figure 2a shows RGB color composite of the scene. The reference classification map in Figure 2b shows nine classes, with the unassigned pixels colored in black and labeled as Unassigned. Table 1 shows the number of pixels per each class in the data set.
2. Salinas (SA): This scene was collected by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor over Salinas Valley, California. The image consists of 224 spectral bands with wavelengths ranging from 0.4 to 2.45 μm and a spatial resolution of 3.7 m. The size of Salinas is 512×217 pixels. Bands 108–112, 154–167, and 224 were removed due to distortions from water absorption. Figure 3a shows

RGB composite of the scene. The reference classification map in Figure 3b shows that Salinas ground truth class map contains 16 classes. Similar to PU, unassigned pixels are colored in black and labeled as Unassigned. Table 2 shows number of pixels per class in the data set.

- Mississippi Gulfport (GP): The dataset was collected over the University of Southern Mississippi's-Gulfpark Campus [52]. The image consists of 72 bands with wavelengths ranging from 0.37 to 1.04 μm and a spatial resolution of 1.0 m. The size of GP is 185×89 pixels. Figure 4a shows RGB color composite of the scene. The reference classification map in Figure 4b shows six classes. As with PU and SA datasets, the unassigned pixels in the image are colored in black and labeled as Unassigned. Table 3 shows the number of pixels per class in the dataset.

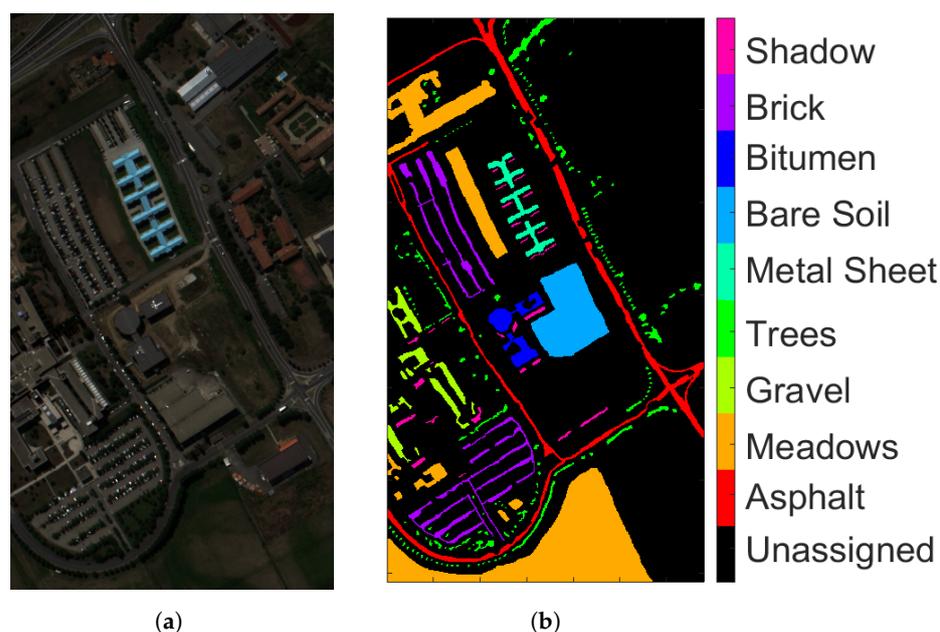


Figure 2. ROSIS Pavia University (a) true-color composite; (b) reference class map.

Table 1. Groundtruth classes for PU scene and their respective samples number.

Class	Name	Total Samples
1	Asphalt	6631
2	Meadows	18,649
3	Gravel	2099
4	Trees	3064
5	Painted Metal Sheet	1345
6	Bare Soil	5029
7	Bitumen	1330
8	Self-Blocking Bricks	3682
9	Shadows	947
Total		42,776

3.2. Evaluation Metrics

To evaluate the performance of classification results, predicted class maps are compared with the available reference or ground truth data. Verifying the correctness of pixel assignment in the image with visual inspection is subjective and might not be comprehensive. Therefore, quantitative evaluation is more reliable. The most common example is the

Overall Accuracy (OA), which computes the number of correctly assigned HSI pixels over the number of overall samples, and it is expressed as Equation (2):

$$Overall\ Accuracy\ (OA) = \frac{Number\ of\ Correctly\ Assigned\ Pixels}{Overall\ Number\ of\ Pixels} \tag{2}$$

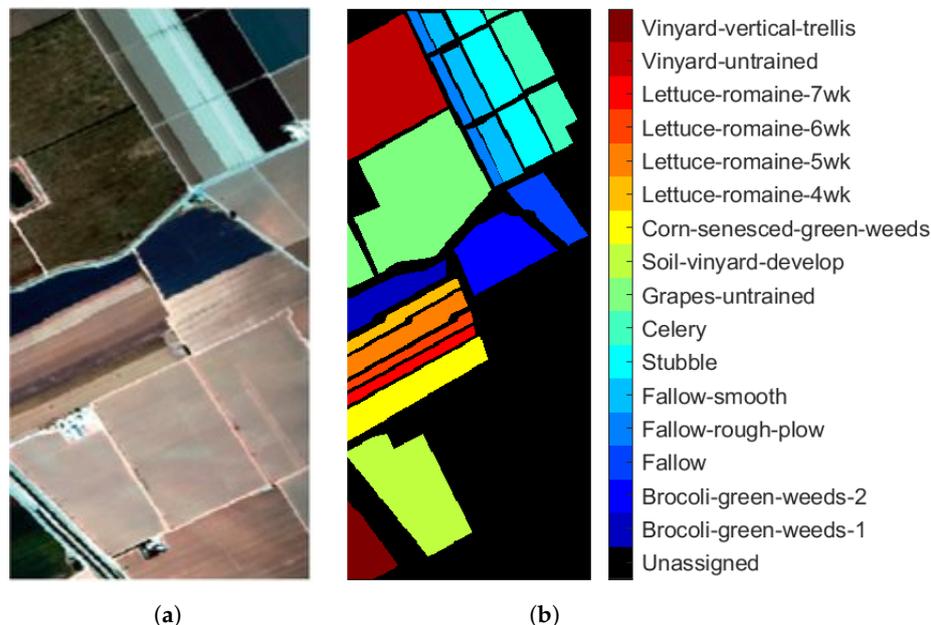


Figure 3. AVIRIS Salinas (a) true-color composite; (b) reference class map.

Table 2. Groundtruth classes for SA scene and their respective samples number.

Class	Name	Total Samples
1	Brocoli-green-weeds-1	2009
2	Brocoli-green-weeds-2	3726
3	Fallow	1976
4	Fallow-rough-plow	1394
5	Fallow-smooth	2678
6	Stubble	3959
7	Celery	3579
8	Grapes-untrained	11,271
9	Soil-vinyard-develop	6203
10	Corn-senesced-green-weeds	3278
11	Lettuce-romaine-4wk	1068
12	Lettuce-romaine-5wk	1927
13	Lettuce-romaine-6wk	916
14	Lettuce-romaine-7wk	1070
15	Vinyard-untrained	7268
16	Vinyard-vertical-trellis	1807
Total		54,129

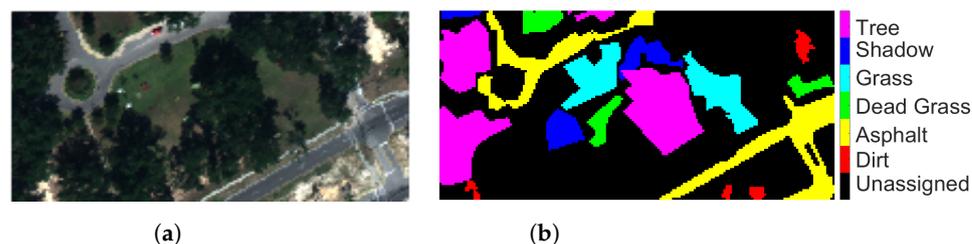


Figure 4. Mississippi Gulfport (a) true-color composite; (b) reference class map.

Table 3. Groundtruth classes for Gulfport of Mississippi scene and their respective samples number.

Class	Name	Total Samples
1	Dirt	175
2	Asphalt	1657
3	Dead Grass	438
4	Grass	970
5	Shadow	527
6	Trees	2330
Total		6097

Average Accuracy (AA) [53] is another evaluation criteria used to assess the classification performance. It estimates the mean of the classification accuracy of all categories or classes. It is defined as Equation (3):

$$\text{Average Accuracy (AA)} = \frac{1}{c} \sum_{i=1}^c (x_i) \quad (3)$$

where c is the number of classes, and x is the percentage of correctly assigned pixels in a single class. Additionally, *Kappa coefficient (k)* [54] is also adopted to evaluate the quality of the classification results of HSI data. It is expressed as Equation (4):

$$\text{Kappa Coefficient (k)} = \frac{OA - P_e}{1 - P_e} \quad (4)$$

where, k determines the agreement between the predicted classified map and the ground truth, and P_e represents the expected agreement between the model classification map and the ground truth map by chance probability. Kappa values range between 0 and 1, such that 1 shows complete agreement between the predicted output and the ground truth, and vice versa for 0. Usually, $\text{Kappa} \geq 0.80$ indicates a substantial agreement, while $k < 0.4$ indicates poor model performance.

3.3. Experimental Configuration

All the experiments have been operated using the compiler and deep learning frameworks; Python 3.8 and TensorFlow 2.4.0, respectively. Adam is adopted as the optimizer algorithm, where the learning rate is 1×10^{-3} , the batch size is 16 and the training epoch is set to 100. The network configuration of the proposed model using PU dataset is shown in Table 4. The number of samples used for training the model is set to 1% for all three datasets to guarantee a fair comparison.

3.4. Experimental Results

In this section, classification performance of the proposed model is evaluated quantitatively and qualitatively using the three aforementioned datasets: Pavia University (PU), Salinas (SA) and Mississippi Gulfport (GP). Furthermore, the proposed methodology is compared with six state-of-the-art algorithms. In order to reduce the fluctuation of classification results caused by the randomness of sample selection, the experiments were repeated 10 times, and the average value of these experiments is used as the final result. Additionally, classification results for each category are reported.

3.4.1. Analysis of Parameters

In this section, we examine how the spatial dimension and number of principal components of different datasets affect the proposed model's ability to classify the HSI data and determine the window size and spectral dimension that are best suited for the dataset.

Table 4. The network configuration for the proposed model on the Pavia University dataset.

Proposed Model Configuration		
Spectral Feature Learning	Spatial Feature Learning	Spectral–spatial Feature Learning
Input:(13 × 13 × 15 × 1)		
3DConv-(1,1,3,64), stride = 1, padding = 0	3DConv-(3,3,1,64), stride = 1, padding = 0	3DConv-(3,3,3,64), stride = 1, padding = 0
Output10:(13 × 13 × 13 × 64)	Output20:(11 × 11 × 15 × 64)	Output30:(11 × 11 × 13 × 64)
3DConv-(1,1,3,64), stride = 1, padding = 0	3DConv-(3,3,1,64), stride = 1, padding = 0	3DConv-(3,3,3,64), stride = 1, padding = 0
Output11:(13 × 13 × 11 × 64)	Output21:(9 × 9 × 15 × 64)	Output31:(9 × 9 × 11 × 64)
Flatten	Flatten	Flatten
Output12:(118,976)	Output22:(77,760)	Output32:(57,024)
Concat(Output12,Output22,Output32)		
FC-(253,760,512)		
Dropout(0.3)		
FC-(512,256)		
Dropout(0.3)		
FC-(256,9)		
Output:(9)		

The window size indicates how much of the retrieved 3D patch’s spatial information can be used to assign a label to the HSI patch. A large window may include a lot of neighborhood data that may contain information from other classes, which hinders the feature extraction process. On the contrary, if the chosen window is too small, the model’s ability to extract features will suffer from a significant loss of spatial information. This research validates the impact of window size on model performance over the three aforementioned datasets. In the experiment, the spatial size was set to $\{5 \times 5, 7 \times 7, 9 \times 9, 11 \times 11, 13 \times 13, 15 \times 15, 17 \times 17\}$, while the spectral dimension was uniformly set to 30. It is clear from Figure 5 that for the PU, SA and GP datasets, the suitable window size for the proposed model was 13×13 , 11×11 and 9×9 , respectively. As seen from the Figure, the overall accuracy for the GP dataset varies in a notable manner with respect to the window size. Since the model was trained with only 1% of the data (60 training samples), we assume that any small change in the initial settings may lead to a large impact on the model performance.

Moreover, the spectral dimension (or number of principal components) shows how much of the 3D patch’s collected spectral data can be used to assign a label to the HSI patch. It is intuitive that the model performance increases with the increase of the spectral dimension. However, having many principal components may contain redundant information, which causes the OA to fall after a certain value. This paper validates the impact of spectral dimension on model performance on the three datasets. In the experiment, the spectral dimension was set to 10, 15, 20, 25, 30, 35, 40, 45, 50, and the spatial size of the three datasets of PU, SA and GP were set to 13×13 , 11×11 , 9×9 , respectively. As can be seen from Figure 6, for PU, SA and GP datasets, the most suitable spectral dimensions for the proposed model were 15, 35, and 45, respectively.

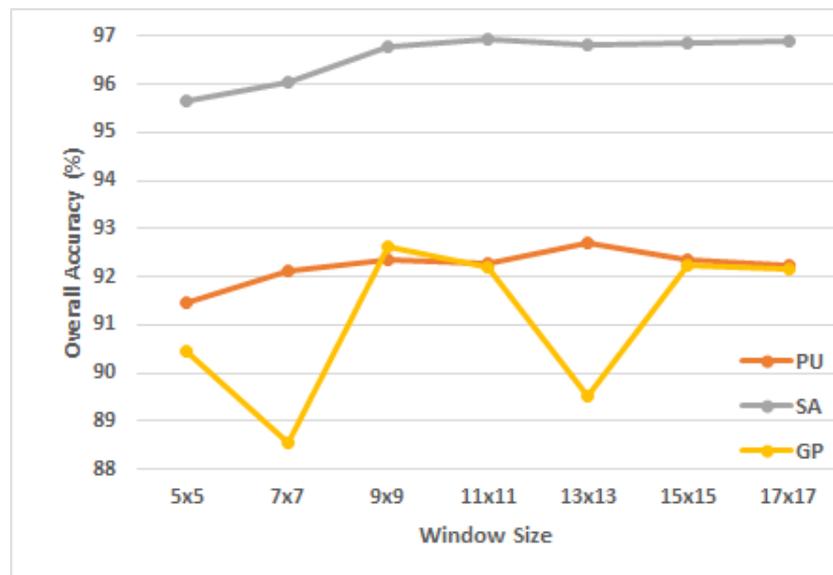


Figure 5. The OA of the proposed model using different values of window size in the three datasets.

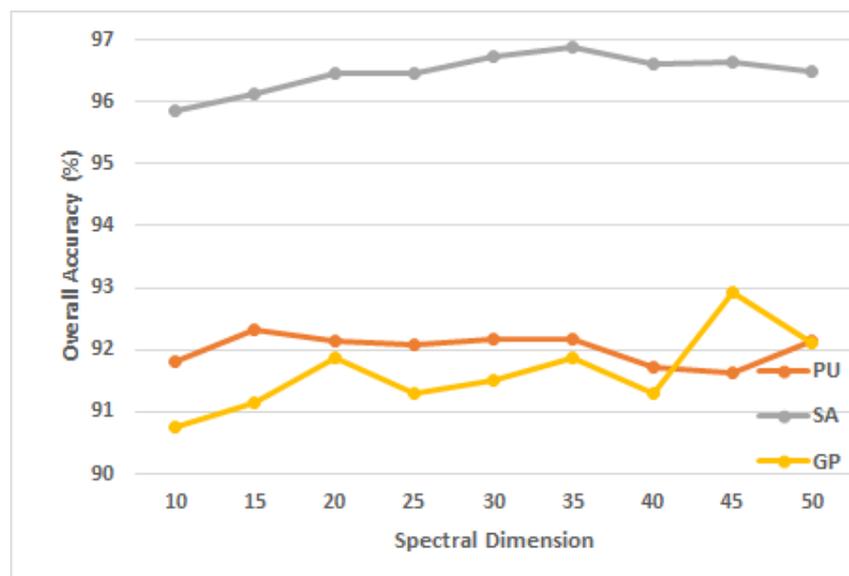


Figure 6. The OA of the proposed model using different spectral dimension in the three datasets.

Based on the above parameter analysis, Table 5 summarizes the optimal values of window size and spectral dimension for the proposed model.

Table 5. The optimal window size and spectral dimension of the proposed model on three HSI datasets.

Dataset	Window Size	Spectral Dimension
PU	13 × 13	15
SA	11 × 11	35
GP	9 × 9	45

3.4.2. Ablation Studies

To fully demonstrate the effectiveness of the proposed model, the influence of different combinations of the model components that belong to the network on the GP dataset is investigated and reported in Table 6. The results show that our proposed fusion technique

obtains the best performance compared with other combination or fusion approaches such as spectral, spatial, spectral–spatial, spectral+spatial, spectral+spectral–spatial, and spatial+spectral–spatial in terms of AA, OA, and Kappa metrics.

Table 6. Results of ablation studies on different combination of model branches over the GP dataset.

Method	Spectral Only	Spatial Only	Spectral–Spatial Only	Spectral + Spatial	Spectral + Spectral–Spatial	Spatial + Spectral–Spatial	Proposed
Overall Accuracy (%)	85.63 ± 4.42	89.45 ± 2.58	91.43 ± 1.93	90.88 ± 3.74	91.41 ± 1.57	91.51 ± 2.18	92.96 ± 1.05
Average Accuracy (%)	82.44 ± 3.28	85.21 ± 3.30	88.23 ± 2.32	87.70 ± 2.55	88.67 ± 2.84	89.48 ± 3.90	89.53 ± 1.39
Kappa × 100	79.31 ± 4.28	87.90 ± 2.40	89.26 ± 2.43	88.89 ± 3.74	90.06 ± 1.13	90.16 ± 1.23	90.56 ± 2.53

3.4.3. Comparison with Other Methods

For the sake of validating the proposed classification algorithm, SVM [16], 2D-CNN [55], 3D-CNN [56], PMI-CNN [50], HybridSN [42], and MSCNN [49] are used as comparison algorithms. As mentioned earlier, the experiments were conducted and repeated 10 times. For all algorithms, only the classification results with the highest accuracy in 10 trials are recorded. The quantitative comparisons of these compared methods are shown in Tables 7–9, and the best results in each Table are shown in bold.

Table 7. Classification results on the PU dataset with 1% labeled training samples. The best accuracy is shown in bold.

Class	Train	Test	SVM	2DCNN	3DCNN	PMI-CNN	HybridSN	MSCNN	Proposed
1	66	6565	83.59	92.91	94.7821	99.17	94.48	91.96	81.01
2	186	18,463	88.88	98.83	99.10	99.31	99.66	99.22	97.04
3	21	2078	63.22	76.22	71.89	77.37	77.46	70.98	86.08
4	31	3033	85.73	85.08	87.46	76.72	83.25	82.99	85.47
5	13	1332	99.40	99.85	99.33	99.22	99.85	98.51	99.43
6	50	4979	79.39	76.21	78.38	89.93	89.73	84.19	99.32
7	13	1317	61.42	79.69	90.30	59.47	90.00	74.66	73.30
8	37	3645	75.58	77.59	83.70	55.48	74.47	85.95	95.00
9	10	937	99.23	83.10	94.19	94.82	91.34	82.36	99.36
Overall Accuracy (%)			84.03 ± 3.33	90.42 ± 4.80	92.13 ± 2.19	90.40 ± 2.42	92.34 ± 2.74	91.48 ± 2.07	92.66 ± 2.24
Average Accuracy (%)			81.89 ± 4.41	85.50 ± 2.50	88.29 ± 2.41	83.59 ± 4.82	88.80 ± 2.86	85.65 ± 3.16	90.65 ± 2.37
Kappa × 100			78.98 ± 3.44	87.13 ± 2.35	89.44 ± 1.25	87.14 ± 2.57	89.17 ± 1.59	88.57 ± 1.09	90.31 ± 1.31

From Tables 7–9, we conclude that the proposed Tri-CNN model outperforms other methods. Among the datasets used in this paper, PU has 42776 labeled samples of which 1% (428 samples) was used for training the model, it has only 9 classes, which makes it easier to classify. SA has a large number of labeled samples where 542 of them were used for training, so it is no surprising that the classification performance on SA is higher. GP is relatively small dataset where it has only 6079 labeled samples, which results in lower classification accuracy since the training has been conducted with only 1% of the labeled samples.

On the PU dataset, SVM method had the lowest overall accuracy, since it utilizes only spectral information. 2D-CNN further improved the accuracy by almost 6% when compared to SVM, 2D-CNN that uses 2D filters to extract spatial information for classification. 3D-CNN had higher accuracy with 92.13%, it uses 3D filters to extract spatial and spectral information simultaneously to further improve classification accuracy. PMI-CNN extracts spatial features by using four parallel branches of two dimensional kernels, extracted features are then fused and used as the input of the classifier; however, it did not show any improvement over 3D-CNN and the accuracy was less by 1.73%. HybridSN uses 3 layers of 3D-CNN followed by one layer of 2D-CNN to extract both spectral–spatial (3D-CNN) and spatial (2D-CNN), it had higher accuracy when compared to the aforemen-

tioned models. The multi-scale CNN (MSCNN) Fuses the spatial features of different scales by using multiple two dimensional kernels to classify HSIs. Our proposed method was approximately 0.32% higher than HybridSN in terms of overall accuracy and outperformed other methods in 5 categories. The results of comparison are shown in Table 7. To have a visual comparison of the results, Figure 7 shows the classification results of the seven methods along with the reference map. It is clearly observed from Figure 7a that the results of SVM has a large number of incorrectly assigned pixels and the classification accuracy is low. The classification map of our proposed Tri-CNN method was the closest to the reference map, specially at the area bounded with the white box.

Table 8. Classification results on the SA dataset with 1% labeled training samples. The best accuracy is shown in bold.

Class	Train	Test	SVM	2DCNN	3DCNN	PMI-CNN	HybridSN	MSCNN	Proposed
1	20	1989	98.65	99.75	100.00	98.35	100.00	99.15	100.00
2	37	3689	99.38	100.00	100.00	100.00	99.91	99.94	100.00
3	20	1956	93.01	99.69	100.00	99.54	98.48	99.94	97.72
4	14	1380	96.05	98.63	99.56	99.13	98.85	95.33	96.05
5	27	2651	97.16	96.19	99.25	98.80	100.00	97.49	97.16
6	39	3920	99.36	99.57	100.00	100.00	99.92	99.94	99.44
7	36	3543	99.05	99.18	99.77	99.87	99.63	99.69	99.94
8	113	11,158	80.03	87.88	90.61	85.67	88.20	90.95	98.65
9	62	6141	99.41	99.98	99.88	100.00	99.90	100.00	100.00
10	33	3245	95.79	96.27	97.52	99.20	96.70	96.36	98.16
11	11	1057	95.22	95.31	99.15	98.22	96.25	98.40	99.53
12	19	1908	88.16	95.53	97.40	99.74	96.26	98.33	95.95
13	9	907	64.19	63.75	79.58	79.47	61.68	96.83	99.78
14	11	1059	89.81	86.16	97.28	96.54	98.13	97.19	98.41
15	72	7196	60.29	77.43	85.23	84.35	90.54	77.49	82.42
16	18	1789	97.45	97.95	99.16	99.28	99.05	97.67	99.88
Overall Accuracy (%)			88.08 ± 2.56	92.68 ± 1.69	95.30 ± 1.43	94.22 ± 2.85	95.54 ± 1.12	94.29 ± 2.23	96.68 ± 2.11
Average Accuracy (%)			90.81 ± 2.11	93.33 ± 2.77	96.52 ± 9.26	96.14 ± 1.47	95.31 ± 1.17	96.54 ± 1.14	97.69 ± 1.09
Kappa × 100			86.70 ± 1.10	91.85 ± 1.10	94.76 ± 1.59	93.57 ± 2.95	94.49 ± 1.24	93.63 ± 2.23	96.30 ± 2.10

Table 9. Classification results on the GP data set with 1% labeled training samples. The best accuracy is shown in bold.

Class	Train	Test	SVM	2DCNN	3DCNN	PMI-CNN	HybridSN	MSCNN	Proposed
1	2	173	1.71	2.28	16.00	3.42	13.42	2.28	86.85
2	16	1641	84.12	96.25	98.61	99.63	92.43	84.73	99.69
3	4	434	1.14	26.48	10.04	12.10	46.74	5.47	86.98
4	10	960	29.38	91.34	99.27	99.07	93.37	86.59	81.64
5	5	522	13.47	94.87	76.47	79.50	67.46	49.33	85.95
6	23	2307	98.71	93.13	94.03	95.83	97.95	94.72	96.05
Overall Accuracy (%)			66.55 ± 3.39	86.45 ± 1.29	86.32 ± 3.32	87.30 ± 4.95	87.79 ± 2.48	77.72 ± 3.28	92.96 ± 1.39
Average Accuracy (%)			38.09 ± 2.13	67.39 ± 3.57	65.74 ± 4.92	64.93 ± 3.72	69.06 ± 1.59	53.85 ± 2.73	89.53 ± 2.58
Kappa × 100			49.44 ± 3.53	81.56 ± 2.42	81.24 ± 3.41	82.58 ± 4.29	82.41 ± 2.69	68.69 ± 3.44	90.56 ± 1.53

Table 8 shows the results of SA dataset. Our proposed method was the highest in terms of OA, AA and Kappa. The model had the highest accuracy in nine out of the sixteen classes, three of which reached 100% accuracy. Figure 8 shows the classification maps of the methods used in this research along with the reference map. The classification maps of our method are closest to the ground-truth map. The area in the white box (Class #8) shows the part that easily distinguished, which matches the results shown in Table 8.

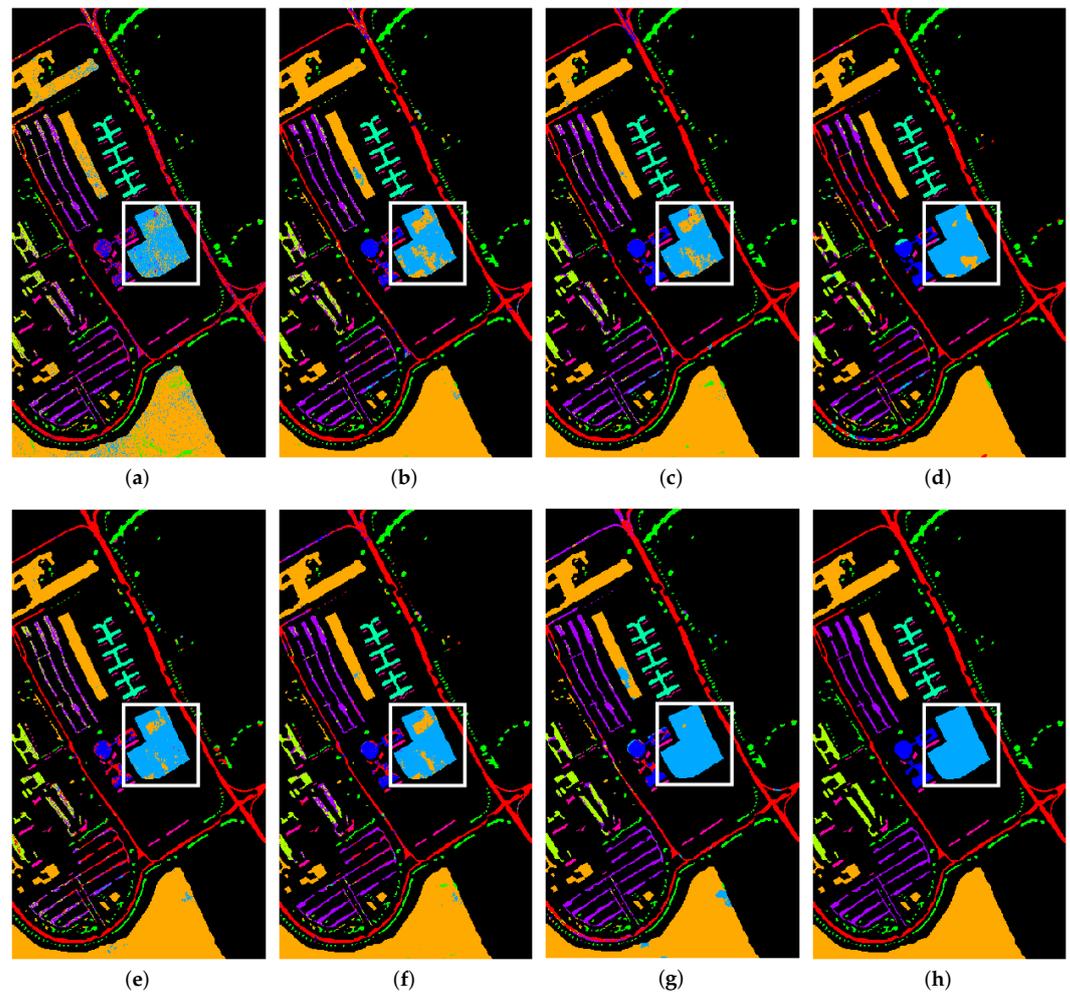


Figure 7. Classification maps for PU dataset. (a) SVM; (b) 2D-CNN; (c) 3D-CNN; (d) HybridSN; (e) PMI-CNN; (f) MSCNN; (g) Proposed method (h) Reference map.

The classification results of GP dataset are shown in Table 9. Our proposed method was the highest in terms of OA, AA and Kappa when compared to the other methods. The model had the highest accuracy in three classes. Figure 9 shows the classification results of the methods used in this research. It is very clear that our proposed methods has the closest classification map to the reference data, specially the area in the white box (Class #1). This particular class had only 2 samples for training, and yet, our model had achieved a classification score of 86.85%, which is extremely higher than the second best (3D-CNN) where the score was only 16%. A similar result is shown in Class #3 (labeled in Green color), where the training was conducted with only 4 samples. This shows the ability of our model to be trained on very small datasets and still good results.

Based on the experimental results, the proposed Tri-CNN produces better results when compared to the other classification methods used in this research. SVM has the worst classification performance since it utilizes spectral features only and loses the spatial information. 2D-CNN, PMI-CNN and MSCNN take into account the spatial information and the overall accuracy was improved when compared to SVM. 3D-CNN extracts spectral and spatial features simultaneously from the 3 dimensional image patch, which preserves the features in the data given. HybridSN combines both 2D-CNN and 3D-CNN to extract the spectral and spatial features to perform the image classification. Our proposed Tri-CNN method produced the best results when applied to the three datasets used in this research. Our model combines the advantage of each branch to extract features at each scale to improve the classification performance.

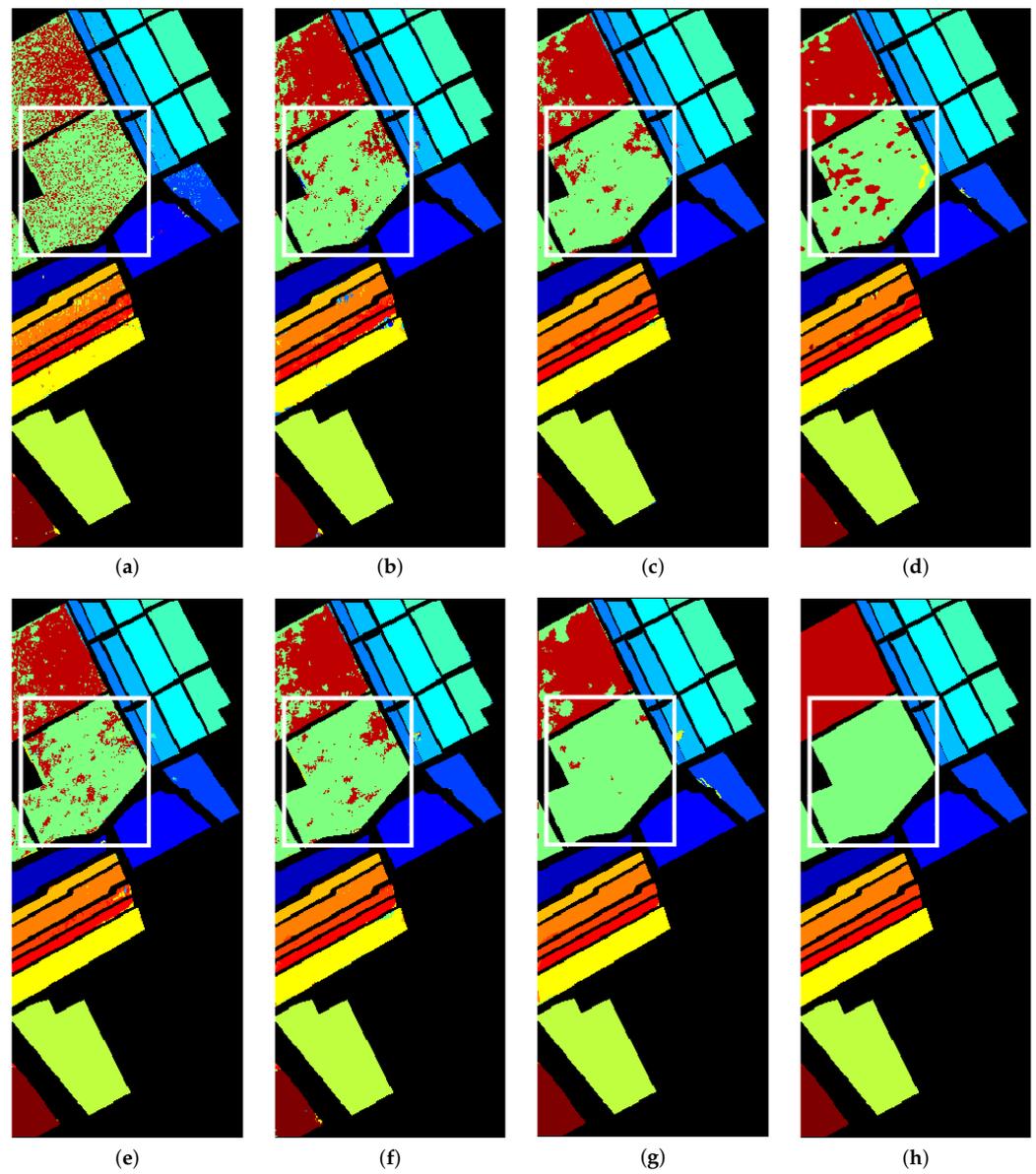


Figure 8. Classification maps for Salinas SA. (a) SVM; (b) 2D-CNN; (c) 3D-CNN; (d) HybridSN; (e) PMI-CNN; (f) MSCNN; (g) Proposed method; (h) Reference map.

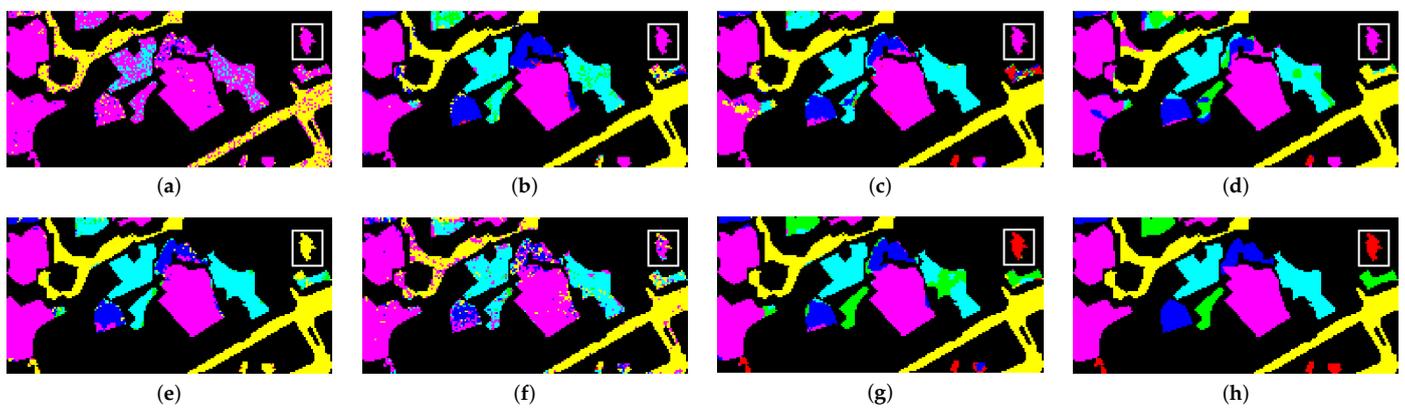


Figure 9. Class maps of GP. (a) SVM; (b) 2D-CNN; (c) 3D-CNN; (d) HybridSN; (e) PMI-CNN; (f) MSCNN; (g) Proposed method; (h) Reference map.

3.4.4. Performance of Different Models at Different Percentages of Training Data

The performance of the model can be well inferred from the classification accuracy under various percentages of training data. We randomly select 1%, 2%, 3%, 4% and 5% of labeled samples for training and the remaining was used for testing. Classification results of each dataset are shown in Figure 10. It is seen that for all the methods used in this study, classification accuracy improves as the number of training samples increases. The proposed model, which performs the best on all training sample proportions on the three datasets, uses 3D-CNN with various kernel sizes to extract various degrees of features from the HSI.

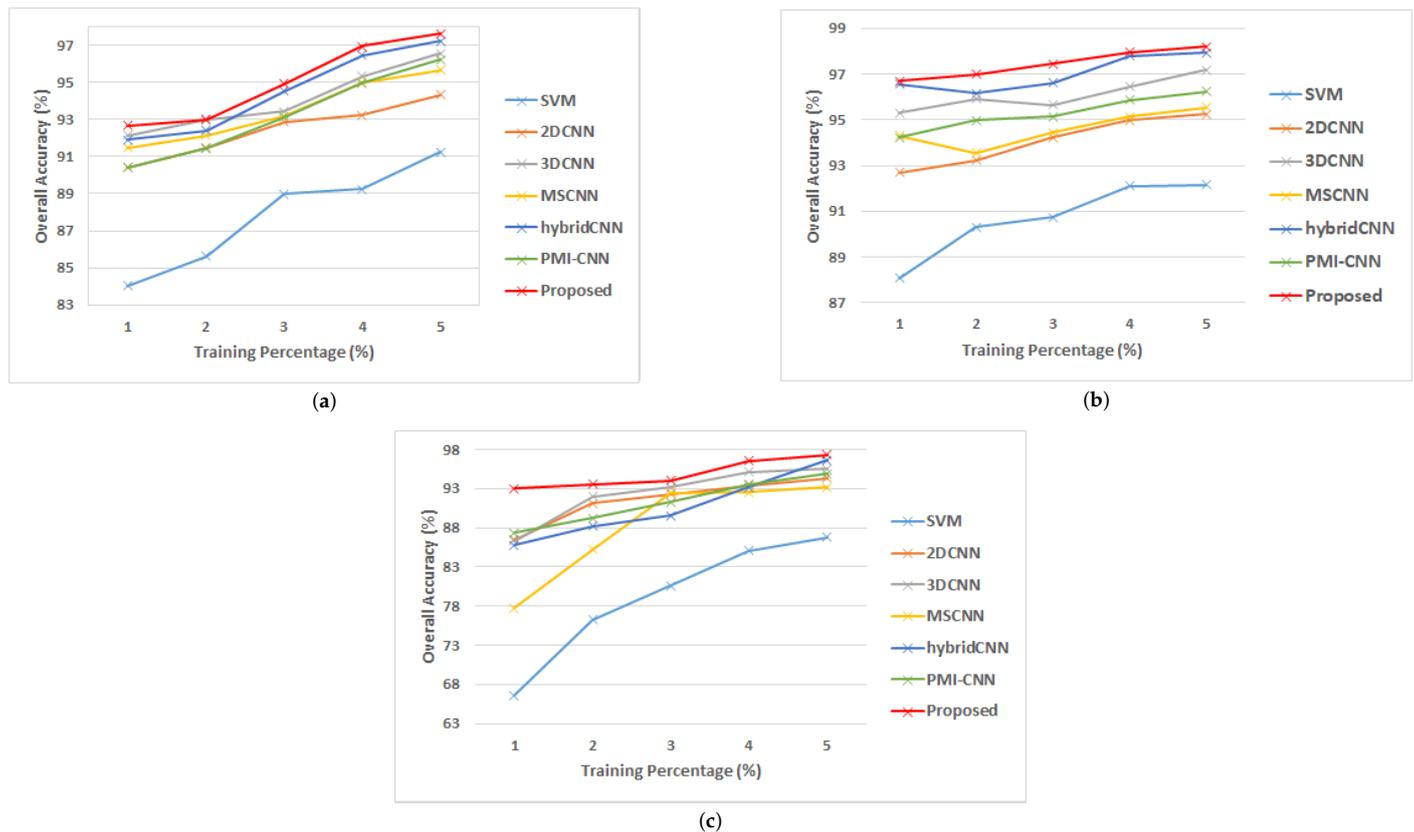


Figure 10. Classification accuracy at different percentages of training data (a) PU; (b) SA; (c) GP.

4. Conclusions

In this paper, we propose a new three-branch CNN model namely Tri-CNN for HSI classification. The proposed model builds a three-branch feature fusion structure and optimizes the creation of the CNN-based model in order to address the current issues with HSI classification. The model feeds the data into three branches of multiscale 3D-CNN after PCA, features generated of each branch are then flattened and fused, fully connected and dropout layers are utilized to produce the final classification result. Experimental results indicate that the proposed model performs well in terms of OA, AA, and Kappa. With just a little amount of training data, the model produces nearly identical classification results.

For future work, we will examine the use of a lighter architecture with less training parameters to reduce the computational complexity without weakening the model performance. In addition, the classification accuracy of the GP dataset was relatively lower compared to other two datasets. Therefore, the focus of our future study will be on investigating a more compact neural network model, broadening the generalization ability, and achieving adequate classification accuracy on different datasets.

Author Contributions: Conceptualization, M.Q.A., M.A.-S. and N.A.; Formal analysis, M.Q.A.; Funding acquisition, S.A. and H.A.-A.; Methodology, M.Q.A. and M.A.-S.; Project administration,

S.A. and H.A.-A.; Software, M.Q.A.; Supervision, M.Q.A. and H.A.-A.; Validation, M.Q.A. and N.A.; Writing—original draft, M.Q.A. and M.A.-S.; Writing—review & editing, N.A., J.Z. and S.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The PU and SA datasets can be obtained from: www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes (accessed on 28 November 2022). The GP dataset can be obtained from: <https://github.com/GatorSense/MUUFLGulfport> (accessed on 28 November 2022).

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

References

- Moroni, M.; Lupo, E.; Marra, E.; Cenedese, A. Hyperspectral image analysis in environmental monitoring: Setup of a new tunable filter platform. *Procedia Environ. Sci.* **2013**, *19*, 885–894. [[CrossRef](#)]
- Stuart, M.B.; McGonigle, A.J.; Willmott, J.R. Hyperspectral imaging in environmental monitoring: A review of recent developments and technological advances in compact field deployable systems. *Sensors* **2019**, *19*, 3071. [[CrossRef](#)] [[PubMed](#)]
- Ad ao, T.; Hruška, J.; Pádua, L.; Bessa, J.; Peres, E.; Morais, R.; Sousa, J.J. Hyperspectral imaging: A review on UAV-based sensors, data processing and applications for agriculture and forestry. *Remote Sens.* **2017**, *9*, 1110. [[CrossRef](#)]
- Dale, L.M.; Thewis, A.; Boudry, C.; Rotar, I.; Dardenne, P.; Baeten, V.; Pierna, J.A.F. Hyperspectral imaging applications in agriculture and agro-food product quality and safety control: A review. *Appl. Spectrosc. Rev.* **2013**, *48*, 142–159. [[CrossRef](#)]
- Park, B.; Lu, R. *Hyperspectral Imaging Technology in Food and Agriculture*; Springer: Berlin/Heidelberg, Germany, 2015.
- Lu, B.; Dao, P.D.; Liu, J.; He, Y.; Shang, J. Recent advances of hyperspectral imaging technology and applications in agriculture. *Remote Sens.* **2020**, *12*, 2659. [[CrossRef](#)]
- Bouabanga-Tombet, S.; Huot, A.; Vitins, I.; Heuberger, S.; Veuve, C.; Eisele, A.; Hewson, R.; Guyot, E.; Marcotte, F.; Chamberland, M. Thermal infrared hyperspectral imaging for mineralogy mapping of a mine face. *Remote Sens.* **2018**, *10*, 1518. [[CrossRef](#)]
- Kereszturi, G.; Schaefer, L.N.; Miller, C.; Mead, S. Hydrothermal alteration on composite volcanoes: Mineralogy, hyperspectral imaging, and aeromagnetic study of Mt Ruapehu, New Zealand. *Geochem. Geophys. Geosystems* **2020**, *21*, e2020GC009270. [[CrossRef](#)]
- Johnson, C.L.; Browning, D.A.; Pendock, N.E. Hyperspectral imaging applications to geometallurgy: Utilizing blast hole mineralogy to predict Au-Cu recovery and throughput at the Phoenix mine, Nevada. *Econ. Geol.* **2019**, *114*, 1481–1494. [[CrossRef](#)]
- Yuen, P.W.; Richardson, M. An introduction to hyperspectral imaging and its application for security, surveillance and target acquisition. *Imaging Sci. J.* **2010**, *58*, 241–253. [[CrossRef](#)]
- Stein, D.; Schoonmaker, J.; Coolbaugh, E. *Hyperspectral Imaging for Intelligence, Surveillance, and Reconnaissance*; Technical report; Space and Naval Warfare Systems Center: San Diego, CA, USA, 2001.
- Koz, A. Ground-Based Hyperspectral Image Surveillance Systems for Explosive Detection: Part I—State of the Art and Challenges. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 4746–4753. [[CrossRef](#)]
- Ghamisi, P.; Yokoya, N.; Li, J.; Liao, W.; Liu, S.; Plaza, J.; Rasti, B.; Plaza, A. Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 37–78. [[CrossRef](#)]
- 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](#)]
- Bellman, R.E. Dynamic programming. *Science* **1966**, *153*, 34–37. [[CrossRef](#)]
- Melgani, F.; Bruzzone, L. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 1778–1790. [[CrossRef](#)]
- Wang, X. Hyperspectral Image Classification Powered by Khatri-Rao Decomposition based Multinomial Logistic Regression. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 5530015. [[CrossRef](#)]
- Joelsson, S.R.; Benediktsson, J.A.; Sveinsson, J.R. Random forest classifiers for hyperspectral data. In Proceedings of the 2005 IEEE International Geoscience and Remote Sensing Symposium, IGARSS'05, Seoul, Republic of Korea, 25–29 July 2005; IEEE: Piscataway, NJ, USA, 2005; Volume 1, p. 4.
- Luo, K.; Qin, Y.; Yin, D.; Xiao, H. Hyperspectral image classification based on pre-post combination process. In Proceedings of the 2019 6th International Conference on Systems and Informatics (ICSAI), Shanghai, China, 2–4 November 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1275–1279.
- Windrim, L.; Ramakrishnan, R.; Melkumyan, A.; Murphy, R.J.; Chlingaryan, A. Unsupervised feature-learning for hyperspectral data with autoencoders. *Remote Sens.* **2019**, *11*, 864. [[CrossRef](#)]
- Li, Z.; Huang, H.; Zhang, Z.; Shi, G. Manifold-Based Multi-Deep Belief Network for Feature Extraction of Hyperspectral Image. *Remote Sens.* **2022**, *14*, 1484. [[CrossRef](#)]
- Zhu, L.; Chen, Y.; Ghamisi, P.; Benediktsson, J.A. Generative adversarial networks for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 5046–5063. [[CrossRef](#)]

23. Bai, J.; Lu, J.; Xiao, Z.; Chen, Z.; Jiao, L. Generative adversarial networks based on transformer encoder and convolution block for hyperspectral image classification. *Remote Sens.* **2022**, *14*, 3426. [[CrossRef](#)]
24. Mou, L.; Ghamisi, P.; Zhu, X. Deep Recurrent Neural Networks for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 3639–3655. [[CrossRef](#)]
25. Hang, R.; Liu, Q.; Hong, D.; Ghamisi, P. Cascaded recurrent neural networks for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 5384–5394. [[CrossRef](#)]
26. Fukunaga, K. *Introduction to Statistical Pattern Recognition*; Elsevier: Amsterdam, The Netherlands, 2013.
27. Zhou, M.; Samiappan, S.; Worch, E.; Ball, J.E. Hyperspectral Image Classification Using Fisher’s Linear Discriminant Analysis Feature Reduction with Gabor Filtering and CNN. In Proceedings of the IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 493–496.
28. Ali, U.M.E.; Hossain, M.A.; Islam, M.R. Analysis of PCA based feature extraction methods for classification of hyperspectral image. In Proceedings of the 2019 2nd International Conference on Innovation in Engineering and Technology (ICIET), Dhaka, Bangladesh, 23–24 December 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–6.
29. Sun, Q.; Liu, X.; Fu, M. Classification of hyperspectral image based on principal component analysis and deep learning. In Proceedings of the 2017 7th IEEE International Conference on Electronics Information and Emergency Communication (ICEIEC), Shenzhen, China, 21–23 July 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 356–359.
30. Fauvel, M.; Chanussot, J.; Benediktsson, J.A. Kernel principal component analysis for the classification of hyperspectral remote sensing data over urban areas. *EURASIP J. Adv. Signal Process.* **2009**, *2009*, 1–14. [[CrossRef](#)]
31. Ruiz, D.; Bacca, B.; Caicedo, E. Hyperspectral images classification based on inception network and kernel PCA. *IEEE Lat. Am. Trans.* **2019**, *17*, 1995–2004. [[CrossRef](#)]
32. Paoletti, M.; Haut, J.; Plaza, J.; Plaza, A. A new deep convolutional neural network for fast hyperspectral image classification. *ISPRS J. Photogramm. Remote Sens.* **2018**, *145*, 120–147. [[CrossRef](#)]
33. Tun, N.L.; Gavrilov, A.; Tun, N.M.; Trieu, D.M.; Aung, H. Hyperspectral Remote Sensing Images Classification Using Fully Convolutional Neural Network. In Proceedings of the IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), St. Petersburg, Russia, 26–29 January 2021; pp. 2166–2170. [[CrossRef](#)]
34. Lin, L.; Chen, C.; Xu, T. Spatial-spectral hyperspectral image classification based on information measurement and CNN. *EURASIP J. Wirel. Commun. Netw.* **2020**, *2020*, 1–16. [[CrossRef](#)]
35. Li, X.; Ding, M.; Pižurica, A. Deep Feature Fusion via Two-Stream Convolutional Neural Network for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*, 2615–2629. [[CrossRef](#)]
36. Kanthi, M.; Sarma, T.H.; Bindu, C.S. A 3d-Deep CNN Based Feature Extraction and Hyperspectral Image Classification. In Proceedings of the 2020 IEEE India Geoscience and Remote Sensing Symposium (InGARSS), Online, 1–4 December 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 229–232.
37. Ahmad, M.; Khan, A.M.; Mazzara, M.; Distefano, S.; Ali, M.; Sarfraz, M.S. A fast and compact 3-D CNN for hyperspectral image classification. *IEEE Geosci. Remote Sens. Lett.* **2020**, *19*, 5502205. [[CrossRef](#)]
38. Chang, Y.L.; Tan, T.H.; Lee, W.H.; Chang, L.; Chen, Y.N.; Fan, K.C.; Alkhaleefah, M. Consolidated Convolutional Neural Network for Hyperspectral Image Classification. *Remote Sens.* **2022**, *14*, 1571. [[CrossRef](#)]
39. Sun, K.; Wang, A.; Sun, X.; Zhang, T. Hyperspectral image classification method based on M-3DCNN-Attention. *J. Appl. Remote Sens.* **2022**, *16*, 026507. [[CrossRef](#)]
40. Li, W.; Chen, H.; Liu, Q.; Liu, H.; Wang, Y.; Gui, G. Attention Mechanism and Depthwise Separable Convolution Aided 3DCNN for Hyperspectral Remote Sensing Image Classification. *Remote Sens.* **2022**, *14*, 2215. [[CrossRef](#)]
41. Yang, X.; Zhang, X.; Ye, Y.; Lau, R.Y.; Lu, S.; Li, X.; Huang, X. Synergistic 2D/3D convolutional neural network for hyperspectral image classification. *Remote Sens.* **2020**, *12*, 2033. [[CrossRef](#)]
42. 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](#)]
43. Guo, H.; Liu, J.; Yang, J.; Xiao, Z.; Wu, Z. Deep Collaborative Attention Network for Hyperspectral Image Classification by Combining 2-D CNN and 3-D CNN. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *13*, 4789–4802. [[CrossRef](#)]
44. Ghaderizadeh, S.; Abbasi-Moghadam, D.; Sharifi, A.; Zhao, N.; Tariq, A. Hyperspectral Image Classification Using a Hybrid 3D-2D Convolutional Neural Networks. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 7570–7588. [[CrossRef](#)]
45. Zhang, H.; Li, Y.; Jiang, Y.; Wang, P.; Shen, Q.; Shen, C. Hyperspectral classification based on lightweight 3-D-CNN with transfer learning. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 5813–5828. [[CrossRef](#)]
46. Xu, Y.; Du, B.; Zhang, L. Robust self-ensembling network for hyperspectral image classification. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, 1–14. [[CrossRef](#)]
47. Yu, H.; Xu, Z.; Zheng, K.; Hong, D.; Yang, H.; Song, M. MSTNet: A Multilevel Spectral–Spatial Transformer Network for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–13. [[CrossRef](#)]
48. Wu, S.; Zhang, J.; Zhong, C. Multiscale spectral–spatial unified networks for hyperspectral image classification. In Proceedings of the IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, Japan, 28 July–2 August 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 2706–2709.
49. Xu, Y.; Zhang, L.; Du, B.; Zhang, F. Spectral–spatial unified networks for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 5893–5909. [[CrossRef](#)]

50. Zhong, H.; Li, L.; Ren, J.; Wu, W.; Wang, R. Hyperspectral image classification via parallel multi-input mechanism-based convolutional neural network. *Multimed. Tools Appl.* **2022**, 1–26. [[CrossRef](#)]
51. Pei, S.; Song, H.; Lu, Y. Small Sample Hyperspectral Image Classification Method Based on Dual-Channel Spectral Enhancement Network. *Electronics* **2022**, *11*, 2540. [[CrossRef](#)]
52. Gader, P.; Zare, A.; Close, R.; Aitken, J.; Tuell, G. *MuufI Gulfport Hyperspectral and Lidar Airborne Data Set*; Technical Report REP-2013-570; University Florida: Gainesville, FL, USA, 2013.
53. Nyasaka, D.; Wang, J.; Tinega, H. Learning hyperspectral feature extraction and classification with resnext network. *arXiv* **2020**, arXiv:2002.02585.
54. Song, H.; Yang, W.; Dai, S.; Yuan, H. Multi-source remote sensing image classification based on two-channel densely connected convolutional networks. *Math. Biosci. Eng.* **2020**, *17*, 7353–7378. [[CrossRef](#)] [[PubMed](#)]
55. 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](#)]
56. Hamida, A.B.; Benoit, A.; Lambert, P.; Amar, C.B. 3-D deep learning approach for remote sensing image classification. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 4420–4434. [[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.