



Object-based classification of hyperspectral images based on weighted genetic algorithm and deep learning model

Davood Akbari¹ · Vahid Akbari²

Received: 14 October 2022 / Accepted: 20 February 2023 / Published online: 27 February 2023
© The Author(s), under exclusive licence to Società Italiana di Fotogrammetria e Topografia (SIFET) 2023

Abstract

Numerous uses of the hyperspectral remote sensing technology exist for identifying land cover and tracking its evolution. The classification of hyperspectral images must now take into account both spectral and spatial information due to recent advancements and the production of images with high spatial resolution. Convolutional neural networks (CNNs) have much employed in recent years to enhance the classification precision of hyperspectral images. The simultaneous use of spatial feature extraction methods in CNNs has not received significant attention in prior studies. In this study, a novel CNN architecture has been developed for classifying hyperspectral images. The weighted genetic (WG) algorithm is used in the proposed technique to minimize the hyperspectral image's dimensions. The WG algorithm keeps every band in the image and gives each one weight between zero and one based on how much information it contains. Following the expectation maximization (EM) method to the collected features, the segmented objects are then categorized using the CNN algorithm. Three benchmark hyperspectral images, Pavia, DC Mall, and Indiana Pine, were used to assess the proposed approach. The trials' findings demonstrate the proposed approach's superiority over the multilayer perceptron (MLP) algorithm in the Pavia, DC Mall, and Indiana Pine images by 14, 16, and 8% in the overall accuracy parameter, respectively.

Keywords Hyperspectral image · Object-based classification · Weighted genetic (WG) · Expectation maximization (EM) segmentation · Deep learning

Introduction

In the past two decades, there has been considerable advancement in hyperspectral remote sensing technology. The capacity to generate data with high spectral, spatial, and radiometric properties enables improved analysis and the effective identification of ground objects. It also introduces additional challenges not present in multispectral data. The first issue is the high volume of the data, which necessitates specialized technology and software for processing. Another issue is the time length needed to process the data (Homayouni & Roux, 2003). Reducing the number of bands is one way to address these issues. Several techniques have been proposed for this purpose, such as feature extraction and

feature selection (Chang, 2003). Genetic algorithms are a few of the applications for evolutionary computation techniques with success applications in fields such as the vehicle routing problem, feature selection, optimization, heart sound segmentation, or traveling salesman problem (Sivanandam & Deepa, 2008).

The classification of these images is the current focus of most hyperspectral remote sensing technology research. Hyperspectral image classification techniques include two groups (Akbari, 2017) – (Akbari, 2020a). The first group refers to spectral or pixel-based classification techniques in which each pixel is allocated to a particular class solely based on its spectral data without considering the data in nearby pixels. Support vector machine (SVM) (Cristianini & Shawe-Taylor, 2000) and multilayer perceptron (MLP) neural network (Atkinson & Tatnall, 1997) are two examples of these techniques. The spectral-spatial or object-based classification methods fall under the second group using surrounding pixels' spectrum data in addition to their own (Akbari et al., 2022) – (Pan et al., 2020). Three techniques are used to extract spatial information: nearest

✉ Davood Akbari
davoodakbari@uoz.ac.ir

¹ Department of Geomatic Engineering, Faculty of Engineering, University of Zabol, Zabol, Iran

² Department of Computing Science and Mathematics, Faculty of Natural Sciences, University of Stirling, Stirling, UK

neighborhoods (Huang & Zhang, 2009), morphological profiles (Pesaresi & Benediktsson, 2001), and segmentation (Tarabalka et al., 2011). According to qualities like homogeneity, segmentation algorithms identify objects in the image (a group of pixels with the same attribute). The benefits of adopting segmentation techniques are listed in (Bitam & Ameur, 2013) – (Tarabalka et al., 2010). In these approaches, each item is described as a spatial neighborhood for all pixels inside it. While maintaining regions with one or more pixels, this technique builds huge neighborhoods for vast and uniform areas. Therefore, the segmentation map will produce accurate and complete spatial information if a precise map of objects is to be constructed based on the spatial structures in the image. Hierarchical and expectation maximization (EM) (Celeux & Govaert, 1992) algorithms are among these methods. Golipour et al. (2015) reported a spectral-spatial classification approach based on hierarchical segmentation (Golipour et al., 2015). They employed multinomial logistic regression and SVM classification to calculate the conditional probability distribution for each class. In 2020, Akbari used the marker-based hierarchical method (MHS) to classify the hyperspectral image after first reducing its dimensions with the minimal noise fraction (MNF) technique (Akbari, 2020b). A spectral-spatial feature tokenization transformer approach with a Gaussian weighted feature marker for function transformation was described in (Sun et al., 2022), collecting spectral-spatial characteristics as well as sophisticated semantic features for the classification of hyperspectral images. Aletti et al. introduced a new semi-supervised approach for multilayer segmentation of hyperspectral images in order to compare the similarity indices of various spectra (Aletti et al., 2021). This method combines suitable linear discriminant analysis.

Convolutional neural networks (CNNs) have gained attention in recent years in areas including image classification, segmentation, target recognition, and video analysis, among others (Xu et al., 2015). Considering regional connections, CNNs may extract spatial information. Additionally, the weight-sharing method in these networks significantly lowers the network's trainable parameters (Hong et al., 2021a) – (Hong et al., 2021b). Li et al. classified hyperspectral images using CNNs (Li et al., 2019). Zhao et al. developed a collaborative classification system using hyperspectral and LiDAR data shown to be very effective at isolating features from multisource remote sensing data (Zhao et al., 2020). Ding et al. suggested a convolutional neural network based on diverse branch modules (DBB) (Ding et al., 2021). It enriches the spatial feature by merging branches with various scales and levels of complexity, such as convolution sequence, multiscale convolution, and average pooling. As a result, single convolution's capacity to extract features is enhanced. Ahmad et al. have expressed several strategies to improve the performance of the deep

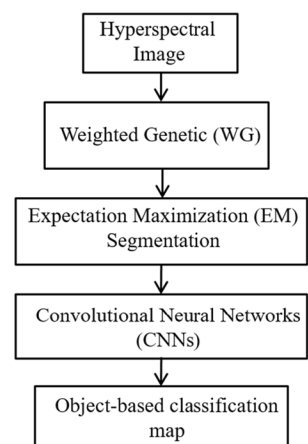
learning method in hyperspectral images (Ahmad et al., 2022). They examined the deep learning technique in three different ways for this purpose: spectral features, spatial features, and spectral-spatial features.

From the review of previous researchers, it is feasible to comprehend the significance of spatial features and dimensionality reduction in enhancing classification accuracy from a study of prior studies. The weighted genetic algorithm has achieved the best result among the dimensionality reduction methods, in which no information deleted, and each band assigned a weight between zero and one. So, this study tried to provide a new spectral-spatial classification method using dimensionality reduction techniques, spatial feature extraction, and CNN classification of the generated objects. This problem is significant because the proposed network can produce deep spectral-spatial features to reach high levels of accuracy in classification. It is true even when little training data is available. The main innovation of this paper is to provide a framework for using deep learning to increase the accuracy of the classification with the increase in the quality of the input data. To achieve this, first, we designed the architecture of CNNs as a suitable-based network. Then, the input raw image was reduced dimensionally with the help of the weighted genetic algorithm. After applying the EM algorithm, image objects were created and considered as network input.

Methodology

Figure 1 shows the steps of the proposed spectral-spatial classification method. As observed, the proposed technique begins by using the WG algorithm to decrease the image's dimensions, followed by the EM algorithm to extract the spatial information, and CNNs to classify the segmented objects. The three algorithms used in the proposed technique are described in the following section.

Fig. 1 Schema of the proposed approach



A. Weighted genetic (WG) algorithm

The most popular evolutionary algorithm that includes recurring procedures and no single method is genetic algorithms, a subset of meta-heuristic optimization approaches. The individuals in the current population are sorted according to their worth throughout each iteration of the algorithm (generation), a new population of solutions created by employing the genetic operators select, crossover, and mutation. This process continues until the algorithm’s termination condition is satisfied (Zhuo & Zheng, 2008). In the binary genetic algorithm, each subset of characteristics is represented using a binary string as an n-dimensional chromosome, where 1 and 0 denote the presence or absence of a particular feature, respectively (Zhuo & Zheng, 2008). However, in the WG method, this string has values between 0 and 1. The fitness function is used to determine the likelihood of each chromosome’s survival and transferred to the following generation. In this study, the value of each chromosome was calculated using the kappa coefficient parameter of the MLP classification. Additionally, the select operator uses the roulette wheel approach. According to this strategy, the likelihood of choosing each chromosome is inversely correlated with its merit score (Huang & Wang, 2006). To avoid the selection of local optimums, the crossover operator with a single point and the mutation operator was applied. The algorithm will stop if the fitness function does not improve before a specific repetition, equivalent to 100 generations. If it does, the repetition will continue until the 100th generation. This halting condition was also considered as a dynamic condition in this study.

B. Expectation maximization (EM) segmentation

The EM algorithm, as a member of the statistical algorithms, operates under the presumption that the data are described by a statistical model (Celeux & Govaert, 1992). We assume that the pixels in a cluster are drawn from a multivariate Gaussian probability distribution in order to cluster the hyperspectral image using the EM technique. Equation (1) states that the probability distribution function may be used to statistically model each individual image pixel.

$$P(x) = \sum_{c=1}^C w_c Q_c(x; \mu_c, \Sigma_c) \tag{1}$$

where $w_c \in [0, 1]$ is the mixing ratio (weight) of cluster C with $\sum_{c=1}^C w_c = 1$ and $Q(\mu, \Sigma)$ is the multivariate Gaussian density with mean vector μ and covariance matrix is Σ . In this relation, P is the probability distribution function and C is the number of clusters.

$$Q_c(x; \mu_c, \Sigma_c) = \frac{1}{(2\pi)^{\frac{d}{2}}} \frac{1}{|\Sigma_c|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (x - \mu_c)^T \Sigma_c^{-1} (x - \mu_c) \right\} \tag{2}$$

In this relation, d is equal to the dimensions of the variable x and $|\Sigma|$ means the determinant of the matrix Σ and T also indicates the transpose of the vector. The estimation of the distribution parameters $\Psi = \{C, w_c, \mu_c, \Sigma_c; c = 1, 2, \dots, C\}$ is done iteratively, similar to the EM classification technique. The pixels are given the cluster C designation during the parameter estimate phase. As a result, the clustering of pixel vectors into C clusters is accomplished after the algorithm convergence.

C. Convolutional neural networks (CNNs)

A variety of tasks relating to image processing, machine vision, signal processing, and natural language processing may be solved using deep learning, a potent machine learning technique. CNNs, applied in disciplines involving remote sensing images, is one of the most well-known deep learning architectures (Lyu & Mou, 2016). This technique classifies remote sensing images more accurately than previous deep learning systems through supervised learning and runs images through several layers, such as neurons or tiny kernels (convolution). Each neuron examines a little portion of the image before producing an output. This output contains the class or possibly the best description of the classes. Three layers make up the convolutional neural network architecture employed in this study. Convolution is the first layer, where weights (kernels) progressively travel across the image to extract various information. Three convolution layers are part of the architecture that employed in this study. The second layer is the pooling layer. This layer shrinks the dimensions of location-related data to lower the number of parameters, cut expenses, and avoid overfitting. The third layer, fully coupled, reduces the input layers to a single-dimensional layer (Dutta et al., 2017).

In the classification process with CNNs, the training samples are randomly divided into several groups because network processing is time-consuming and heavy. Each group contains an equal number of training samples. For each iteration, only one set of pieces is sent to the network for training, and the output value determined after operating each layer. Then, it is trained by the stochastic gradient descent with momentum (sgdm) method and using the cross-entropy cost function of the network. The process completes after entering all small groups generated from training samples into the network. The training process will not stop until reaching the maximum number of steps. The structure of CNN layers is examined next.

Convolution layer

The convolution layer uses linear convolution filters with dimensions $k_w \times k_h \times d$ to extract features. K_w is the length of the filter, k_h the width of the filter, and d the height of the filter, equal to the number of bands of the input image. Convolution kernels with size $1 \times 1 \times d$ can only extract spectral features not spatial ones. If the purpose of classification is only based on spatial information, the third dimension that indicates the number of input bands is one, and the kernel is two-dimensional. In the spectral-spatial classification mode, a three-dimensional kernel is used. A feature map is created after applying each kernel.

Pooling layer

In CNNs, there is usually a pooling layer after the convolution layer, so the number of feature maps in this layer equals the convolution layer. The purpose of the pooling layer is to reduce the computing power required for data processing through dimensionality reduction. In addition, due to the combination of the output results of several neurons of the convolution layer, the pooling layer produces features stable against rotation and position change. These features, known as dominant features, increase the speed of network convergence and cause better network training. Different functions can be used in the pooling layer. The

most used are maximum and average functions. As the maximum one produces better results, according to the research, it is used in the pooling layer (Chen et al., 2016).

Fully connected layer

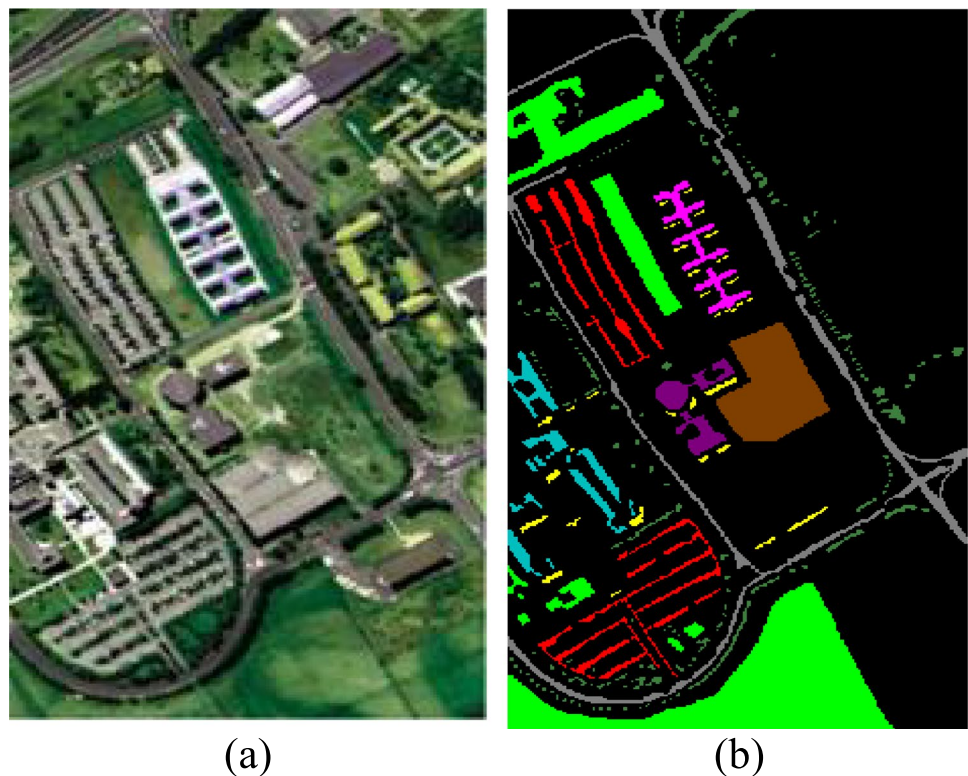
This layer combines the spectral and spatial features extracted in the previous layers to classify the input data. As the name suggests, for two successive layers to be fully connected, all neurons in this layer must be connected to all neurons in the next layer. The output value of the fully connected layer can be calculated from Eq. (3) using weight and bias values (Srivastava et al., 2014).

$$o^{(k)} = (o^{(k-1)})^T w^{(k)} + b^{(k)} \quad (3)$$

In Eq. (3), $w^{(k)}$ is the weight of the position k , and b is the bias value. In the discussion of classification with CNNs, the number of neurons of the last fully connected layer is selected as the number of classes using the maximum smoothing function calculated from Eq. (4). Possible output is produced that specifies the probability of the sample belonging to each of the classes. The sample belongs to the class with the highest probability.

$$y = \frac{1}{\sum_{K=1}^M e^{W_{L,K}^T X_L + b_{L,K}}} \begin{bmatrix} e^{W_{L,1}^T X_L + b_{L,1}} \\ \vdots \\ e^{W_{L,M}^T X_L + b_{L,M}} \end{bmatrix} \quad (4)$$

Fig. 2 Pavia dataset: **a** RGB color composite, **b** reference map



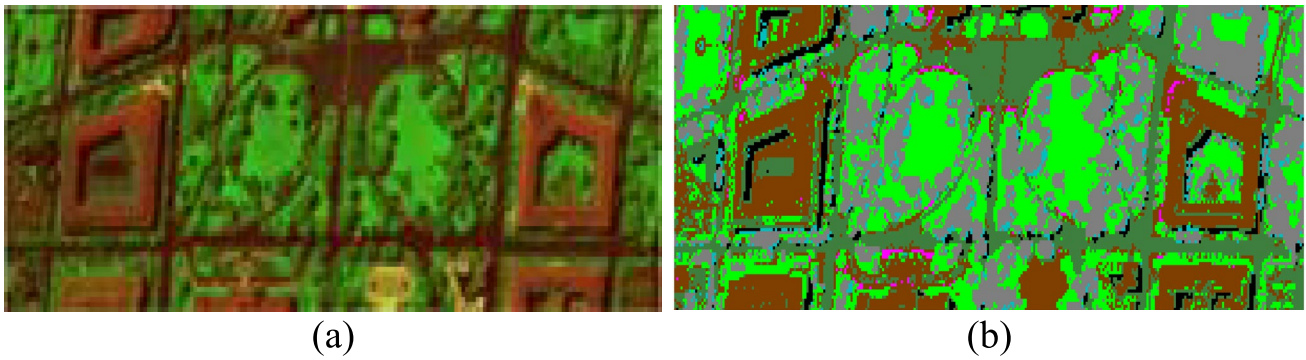
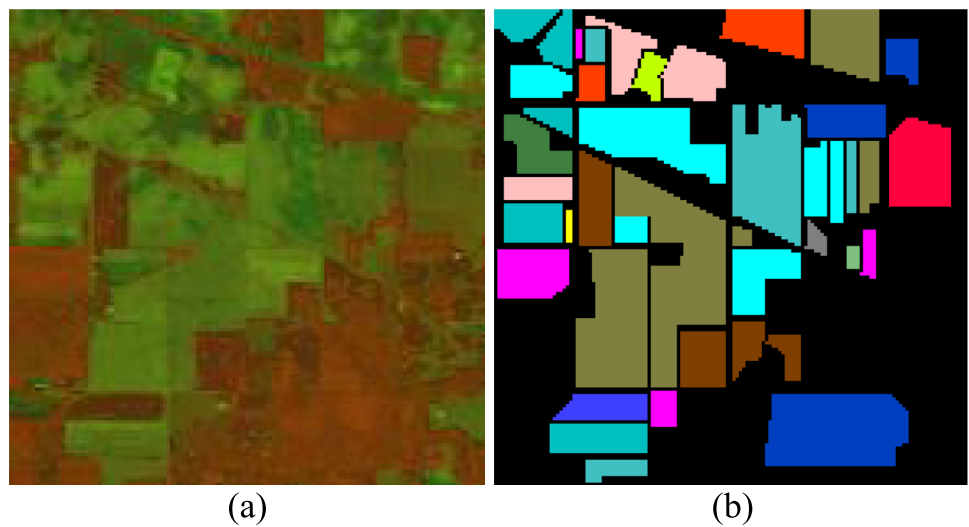


Fig. 3 DC Mall dataset: **a** RGB color composite, **b** reference map

Fig. 4 Indiana Pine dataset: **a** RGB color composite, **b** reference map



In relation (4), M is the number of classes and the denominator term of the fraction $\sum_{K=1}^M e^{W_{L,K}^T X_L + b_{L,K}}$ is only for the normalization of the softmax function.

Hyperspectral data

In this research, Pavia, DC Mall, and Indiana Pine, known as the benchmark images in the hyperspectral remote sensing, were used to evaluate the proposed method.

A. Pavia dataset

The image of Pavia was taken by the ROSIS-03 sensor from the urban area of Pavia in Italy (Chi et al., 2009). The images in this collection includes a spatial resolution of 1.3 m and nine classes. Figure 2 shows color-false combination and reference map of Pavia. The classes of asphalt, meadows, gravel, trees, metal sheets, bare soil, bitumen, bricks, and shadows were chosen from this image’s 610 by

340 pixels and 103 spectral bands to assess the classification findings.

B. DC Mall dataset

The HYDICE sensor with a spatial resolution of 3 m resulted in the DC Mall dataset (Camastra, 1995). The above image includes 210 bands in the spectral range of 0.4 to 2.4 μm and seven classes of shadow, trees, grass, water, roads, rooftops, and trails. Figure 3 displays the color-false combination of this image and its reference map.

Table 1 Values of WG parameters in the three datasets used

Parameters	Value
Population	100
Crossover probability	80%
Mutation probability	0.9%
K-tournament	2
K-elitism	2

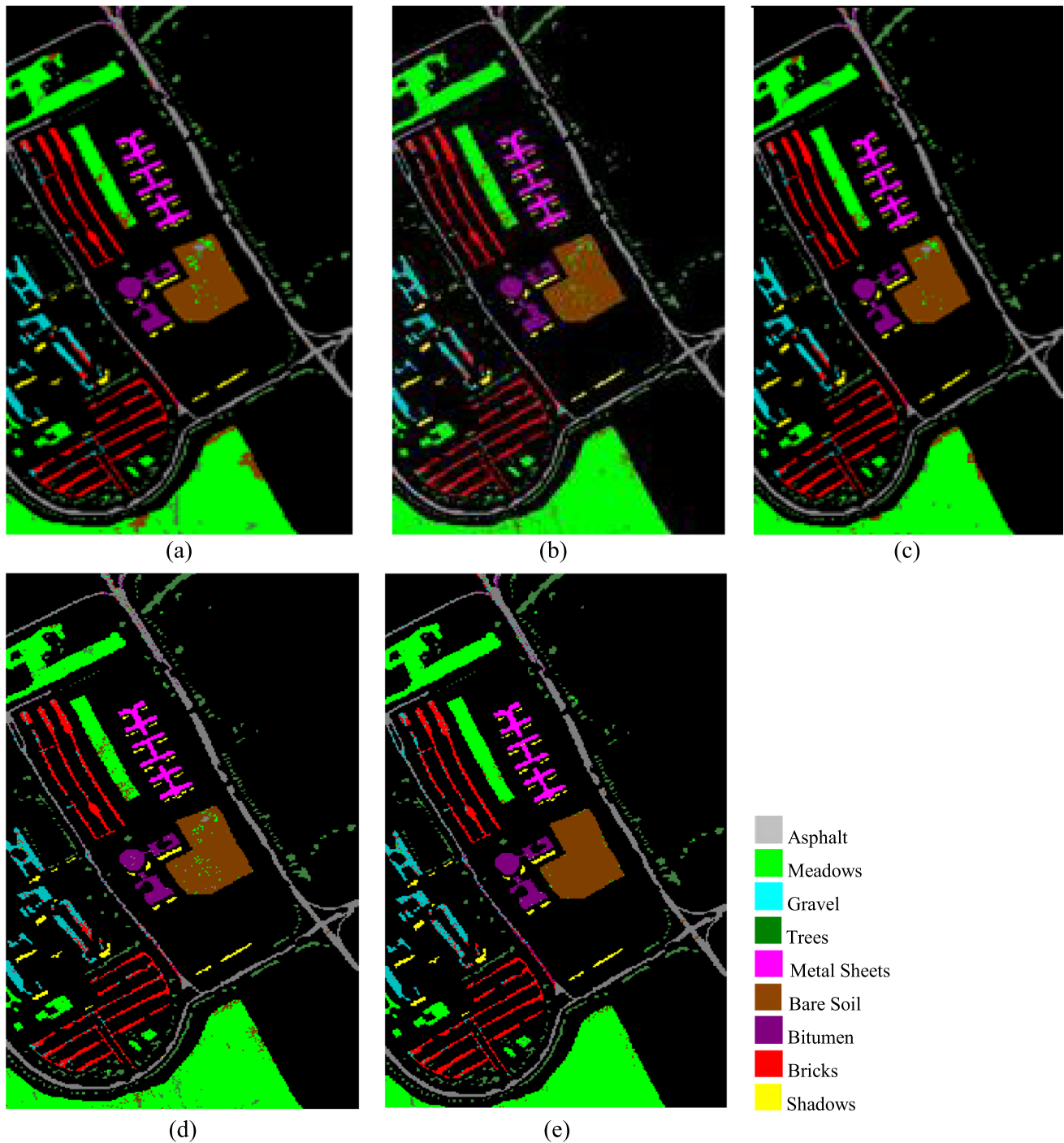


Fig. 5 Pavia dataset classification maps: **a** SVM, **b** MLP, **c** MNF-MHS, **d** CNNs, **e** proposed approach

C. Indiana Pine dataset

The AVIRIS sensor with a spatial resolution of 20 meters captured the Indiana Pine image of an agricultural region in northwest Indiana in 1992 (Landgrebe, 2003). The image consists of 220 spectral bands in the range of 0.4

to 2.5 μm , each with a width of 10 nm. Figure 4 displays the reference map and the color-false combination of the Indiana Pine image. This image has 16 classes, as you can see.

Each image's defined classes are inversely correlated with its nature and complications. In all three image datasets, 10%

Table 2 Accuracy values obtained for the Pavia image

	SVM	MLP	MNF-MHS	CNNs	Proposed approach
OA (%)	79.3	81.9	85.4	86.0	95.3
κ (%)	76.0	77.3	82.7	83.2	92.1
Asphalt	80.5	83.0	88.1	86.1	92.2
Meadows	84.0	86.1	87.0	87.2	95.9
Gravel	64.1	69.6	72.1	73.3	90.7
Trees	78.4	86.2	87.3	88.4	92.8
Metal sheets	86.4	85.3	87.1	87.8	96.6
Bare soil	84.4	86.9	85.9	89.4	94.4
Bitumen	80.6	85.2	83.7	82.8	91.0
Bricks	77.5	80.4	83.7	84.2	90.8
Shadows	80.4	79.2	83.3	86.2	95.9

Values in bold indicate the maximum accuracy value among the classification algorithms used

of the labeled samples were randomly chosen as training data and the remaining 90% as test data.

Results and discussion

The values of the parameters derived from the WG method implementation, which are the same for all three image data, are displayed in Table 1.

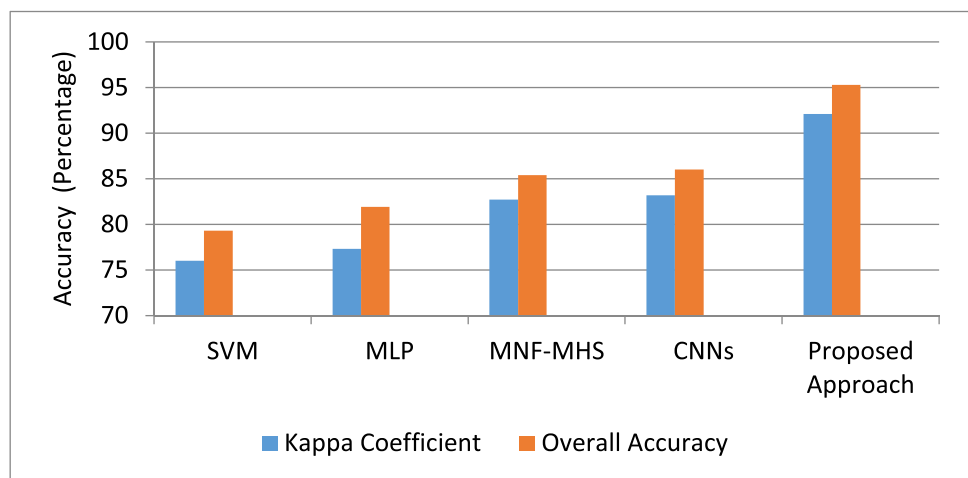
The proposed classification method was compared with the algorithms from SVM, MLP, MNF-MHS (Akbari, 2020b), and CNNs. Cross-validation approach was used to calculate the values of the two penalty parameters (C) and the Gaussian kernel (γ) in the SVM algorithm (Cristianini & Shawe-Taylor, 2000). Thus, the final values of the aforementioned parameters were determined to be C=168 and $\gamma = 0.02$ for the Pavia image, C=240 and $\gamma = 0.01$ for the DC Mall, and C=156 and $\gamma = 0.01$ for the Indiana Pine image. Five hundred evaluations were performed on the MLP

classification method, which was developed using 3 hidden layers with 5, 6, and 8 neurons. SVM classification map and Gaussian radial basis kernel were utilized to pick markers in the marker-based HSEG method. For this purpose, the labeling of the connected components was analyzed based on eight nearby pixels. In regions with more than 20 pixels, 5% of the pixels with the highest probability of belonging to a class were selected as marker pixels. For small regions, fewer than 20 pixels, pixels with a degree of probability more than a threshold were chosen as marker pixels. The threshold corresponds to the likelihood that is 2% of the highest for the entire image (Van der Meer, 2006). The value of the parameter S_{wght} was 0.2 for the HSEG method due to the complexity of the hyperspectral images. Reference data was used to create the confusion matrix and to extract the parameters for overall accuracy (OA), kappa coefficient (κ), and producer accuracy for each class to assess the experiment’s accuracy (Rosenfield & Fitzpatric-Lins, 1986). The difference between the proposed approach and other classification methods was also assessed using the Z-statistic (Akbari et al., 2014).

Figure 5 shows the classification maps produced by the proposed methods via SVM, MLP, MNF-MHS, and CNNs. Compared to the previous methods, the map produced by the proposed technique has uniform areas.

The values of the accuracy parameters of the classification maps derived from the Pavia hyperspectral image are displayed in Table 2 and Fig. 6. As shown, compared to the SVM, MLP, MNF-MHS, and CNN algorithms, the proposed technique has raised the kappa coefficient parameter by 16, 15, 10, and 9%, respectively. Additionally, the proposed strategy has improved all classes’ accuracy, which now stands above 90%. The simultaneous use of deep learning methods and spatial information in the classification process may cause the proposed method’s improved accuracy. The accuracy of the findings can also be improved by dimensionality reduction using a weighted genetic algorithm.

Fig. 6 Comparison of the values of the two parameters of overall accuracy and kappa coefficient for the classification algorithms used in Pavia image



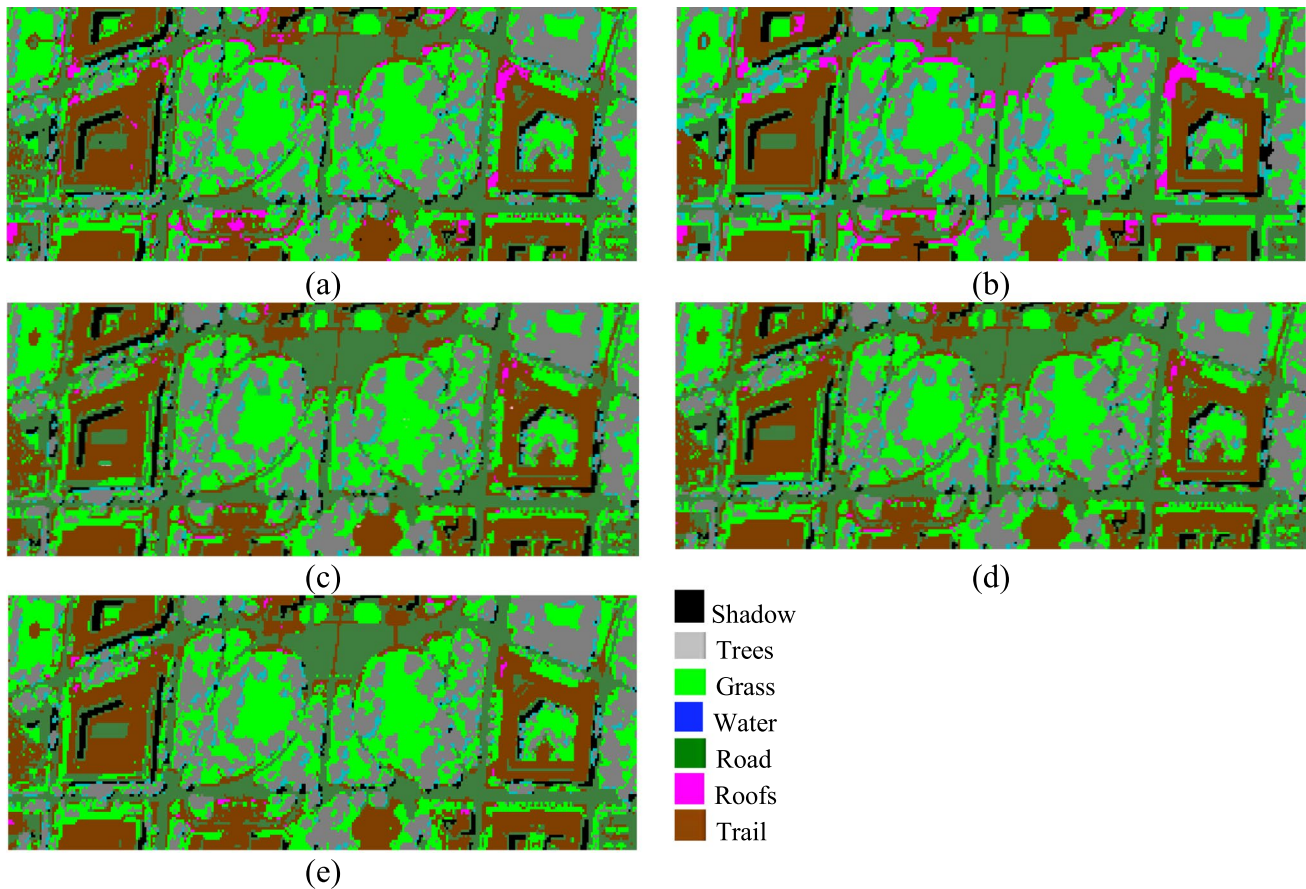


Fig. 7 DC Mall dataset classification maps: **a** SVM, **b** MLP, **c** MNF-MHS, **d** CNNs, **e** proposed approach

The classification maps for the hyperspectral image of the DC Mall are displayed in Figure 7. As can be observed, the proposed method's map is more uniform than the previous methods' maps.

The values of the classification maps' accuracy parameters, derived from the hyperspectral image of the DC Mall, are displayed in Table 3 and Fig. 8. Compared to SVM,

Table 3 Accuracy values obtained for the DC Mall image

	SVM	MLP	MNF-MHS	CNNs	Proposed approach
OA	78.2	77.9	85.2	84.1	93.8
K	75.1	74.1	82.8	81.2	91.0
Shadow	86.1	86.4	87.4	85.5	93.6
Trees	81.1	83.1	86.2	84.3	90.8
Grass	72.3	72.1	81.1	78.6	87.6
Water	80.5	79.6	82.2	79.4	89.3
Road	77.7	80.3	81.5	81.5	90.1
Roofs	85.3	84.3	89.9	88.8	89.8
Trail	76.4	73.2	78.2	79.0	88.1

Values in bold indicate the maximum accuracy value among the classification algorithms used

MLP, MNF-MHS, and CNNs, the proposed technique raised the kappa coefficient parameter in this image by 16, 17, 9, and 10%, respectively. Additionally, all classes—except for roofs—have improved accuracy according to the proposed strategy. The class roofs may have a low density and large dispersion in the image, contributing to this decline.

Figure 9 displays the classification maps for the hyperspectral image of Indiana Pine. As can be observed, the proposed method's map is cleaner than the maps produced by existing methods.

Table 4 and Figure 10 provide the accuracy parameter values for the hyperspectral image of the Indiana Pine. The proposed strategy has improved the accuracy of this image. Comparing the kappa coefficient parameter to the SVM, MLP, MNF-MHS, and CNN algorithms, the increase is 15, 9, 4, and 1%, respectively. The proposed strategy has also boosted the accuracy of all classes. As previously said, the proposed method's simultaneous use of deep learning techniques, dimensionality reduction, and spatial information may blame for the findings' growth.

To identify the position of the proposed method, we compared the results with some other deep learning methods. These methods are as follows:

Fig. 8 Comparison of the values of the two parameters of overall accuracy and kappa coefficient for the classification algorithms used in DC Mall image

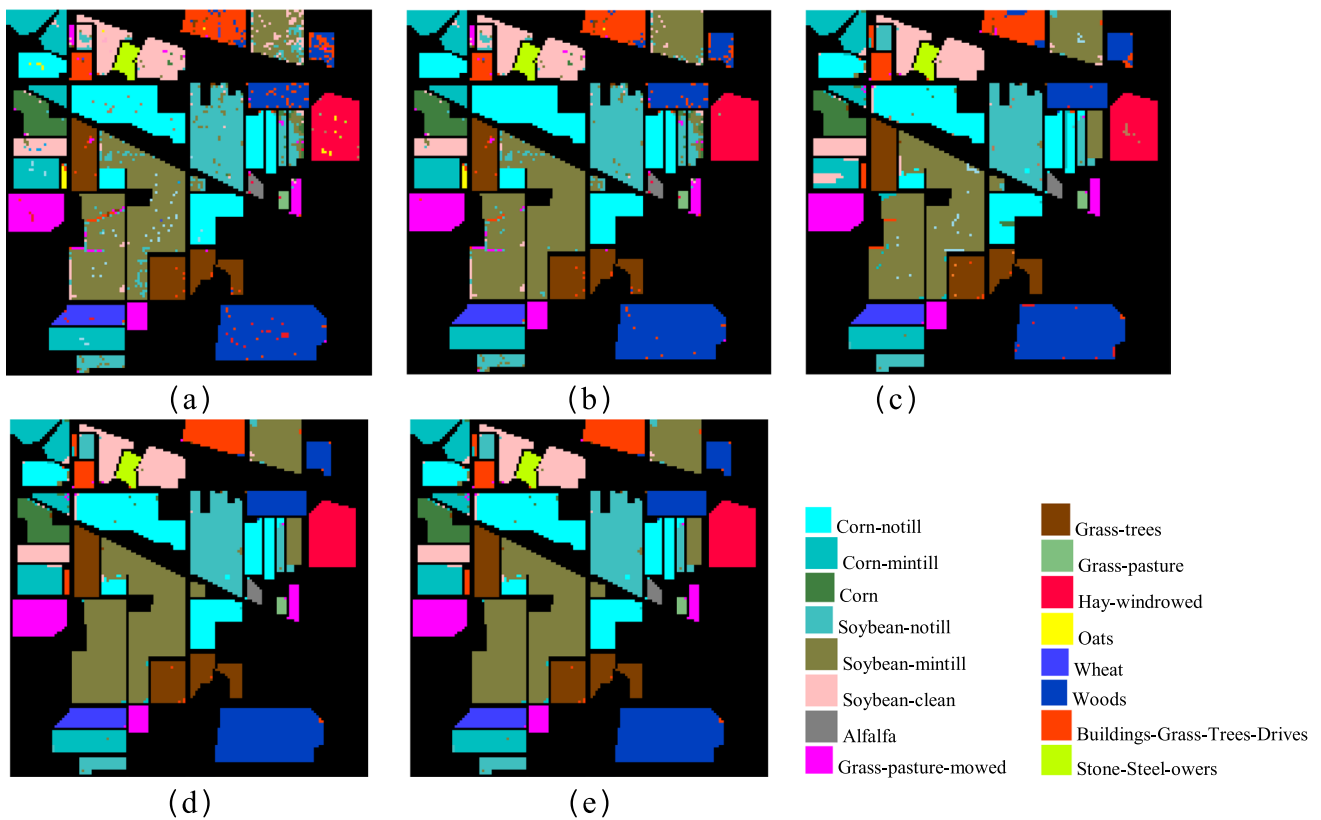
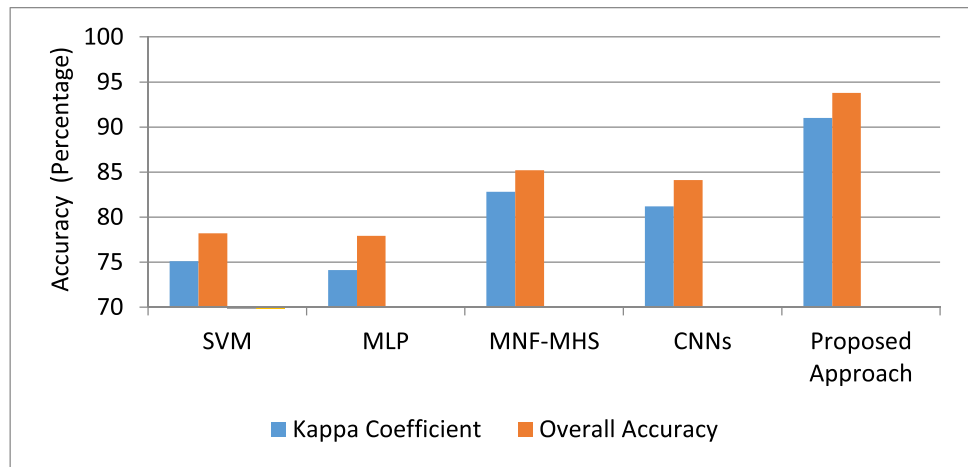


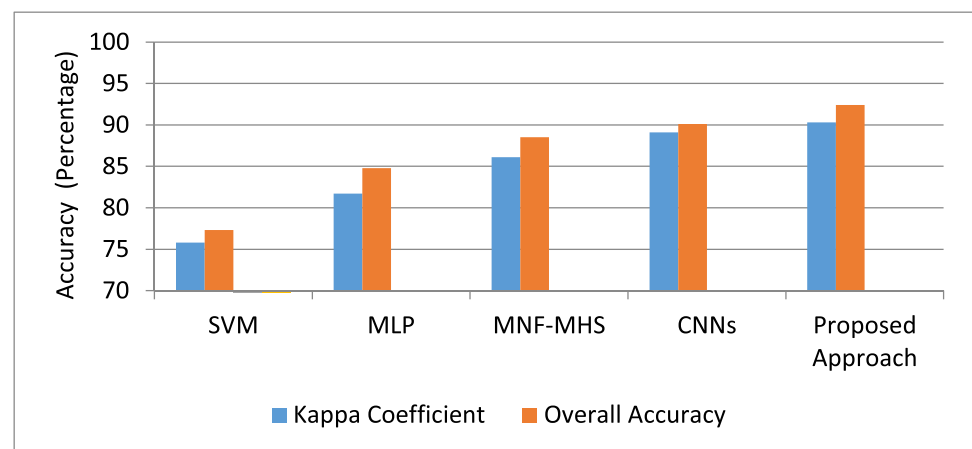
Fig. 9 Indiana Pine dataset classification maps: **a** SVM, **b** MLP, **c** MNF-MHS, **d** CNNs, **e** proposed approach

- (i) Edge-preserving filters (EPF): in this method, first, the hyperspectral image is classified with the help of a classifier such as SVM, and then edge preservation filtering such as a guide filter is applied to the possible classification maps obtained. Finally, according to the filtered probability map, the class of each pixel is selected based on the maximum probability (Kang et al., 2013).
- (ii) R-VCANet: this method is designed to extract deep features from hyperspectral images. R-VCANet has a much simpler network structure than other deep methods. Because the parameters of the convolution kernel are obtained through vertex component analysis (VCA), they need less number of training samples (Ojala et al., 2002).

Table 4 Accuracy values obtained for Indiana Pine image

	SVM	MLP	MNF-MHS	CNNs	Proposed approach
OA (%)	77.3	84.8	88.5	90.1	92.4
κ (%)	75.8	81.7	86.1	89.1	90.3
Corn-notill	75.2	80.2	90.8	92.4	92.5
Corn-mintill	79.2	83.1	90.7	90.6	90.8
Corn	79.4	79.4	83.7	92.0	92.3
Soybean-notill	83.0	84.7	90.5	92.4	92.4
Soybean-mintill	92.2	95.3	96.0	97.2	98.2
Soybean-clean	91.1	96.0	97.2	98.8	99.7
Alfalfa	89.1	88.9	93.2	92.9	94.4
Grass-pasture-mowed	92.2	92.5	99.4	98.6	99.9
Grass-trees	78.0	78.0	82.0	89.0	90.0
Grass-pasture	77.7	80.1	94.7	94.5	96.6
Hay-windrowed	70.9	72.2	90.9	90.0	90.6
Oats	87.7	87.7	94.0	98.1	98.4
Wheat	90.1	91.5	96.7	97.5	99.9
Woods	80.9	82.8	89.1	92.0	92.3
Buildings-grass-trees-drives	82.8	88.1	92.4	96.1	97.5
Stone-steel-towers	88.8	90.5	97.8	97.8	98.2

Values in bold indicate the maximum accuracy value among the classification algorithms used

Fig. 10 Comparison of the values of the two parameters of overall accuracy and kappa coefficient for the classification algorithms used in Indiana Pine image**Table 5** Comparison of the method used in this research with competing methods

Dataset	EPF	R-VCANet	GFDN	RPNet	HybridSN	Proposed approach
Pavia	84.4	85.6	81.5	84.7	87.7	95.8
DC Mall	87	88.3	83.5	87.9	90.0	96.9
Indiana Pine	78.1	84.8	80.0	84.3	86.5	93.2

Values in bold indicate the maximum accuracy value among different methods

- (iii) Gabor filtering and deep network (GFDN): in the GFDN method, Gabor features are first extracted for the first three components of PCA and then placed next to the spectral ones, and the resulting spectral-spatial vector enters the deep SAE network (Kang et al., 2017).
- (iv) RPNet: the RPNet method is a new deep learning method, randomly selected pieces of the image without considering any training as convolution kernels. These kernels are applied to extract features on the images (Chen et al., 2016).

Table 6 Kappa analysis results for pairwise comparison of confusion matrices for different datasets

Dataset	Pairwise comparison	Z-statistic
Pavia	Proposed approach & SVM	13.2
	Proposed approach & MLP	10.8
	Proposed approach & MNF-MHS	7.7
	Proposed approach & CNNs	7.2
DC Mall	Proposed approach & SVM	12.7
	Proposed approach & MLP	13.0
	Proposed approach & MNF-MHS	6.6
	Proposed approach & CNNs	7.6
Indiana Pine	Proposed approach & SVM	12.3
	Proposed approach & MLP	5.7
	Proposed approach & MNF-MHS	2.4
	Proposed approach & CNNs	1.0

- (v) HybridSN: The HybridSN method uses the combination of 2D-CNNs with 3D-CNNs, in which deep spatial features are extracted with the help of 2D-CNNs and entered into 3D-CNNs along with spectral ones, and classification is done (Singhal et al., 2017). The results of the comparison between this research and the proposed method, shown in Table 5, proved that the results of the proposed algorithm were superior to other methods due to the new architecture considered for CNNs and the integration of spatial features with spectral features.

The confusion matrices for the proposed approach and other methods are compared using the third accuracy criterion, or Z-statistic, in addition to the previously mentioned criteria. Table 6 shows the outcomes of this test. As can be observed, for all three datasets, the proposed technique differs considerably from conventional classification methods.

Conclusion

The above research presents an object-based classification method with three consequences: (i) feature extraction based on a weighted genetic algorithm as a powerful tool to preserve hyperspectral image information, (ii) image segmentation interfering with the classification process through introducing spatial information, and (iii) deep learning classification method. No band, or information, is removed in the weighted genetic algorithm and assigned a weight between zero and one based on the information it contains. The EM and CNN algorithms used in the proposed method are also among the most accurate segmentation and classification algorithms in satellite images. Three benchmark

hyperspectral images, Pavia, DC Mall, and Indiana Pine, were used to test the proposed approach. The results proved the superiority of this technique, quantitatively and qualitatively, compared to the four classification algorithms of SVM, MLP, MNF-MHS, and CNNs. The results of the DC Mall image showed that, unlike the other two images, the MNF-MHS algorithm reached better results than the CNNs algorithm, proving the higher importance of using spatial information compared to spectral information due to the higher complexity of this image. Future research will concentrate more on spatial data. To this end, techniques like nearest neighborhoods based on spectral and texture characteristics can be mentioned to incorporate spatial correlations or by utilizing the principles of multiple discriminant analysis, for instance, by using the wavelet tool, to improve the classification.

Declarations

Conflict of interest The authors declare no competing interests.

References

- Ahmad M, Shabbir S, Roy SK, Hong D, Wu X, Yao J, Khan AM, Mazzara M, Distefano S, Chanussot J (2022) Hyperspectral image classification—traditional to deep models: a survey for future prospects. In: IEEE journal of selected topics in applied earth observations and remote sensing, vol 15. IEEE, p 968
- Akbari D (2017) Improving spectral–spatial classification of hyperspectral imagery using spectral dimensionality reduction based on weighted genetic algorithm. *J. Indian Soc. Remote Sens* 45(6):927–937
- Akbari D (2020a) Improving spectral-spatial classification of hyperspectral imagery by using extended minimum spanning forest algorithm. *Can. J. Remote. Sens* 46:146–153
- Akbari D (2020b) A novel method for spectral-spatial classification of hyperspectral images with a high spatial resolution. *Arab. J. Geosci* 13:1–10
- Akbari D, Safari A, Homayouni S (2014) Object-based hyperspectral classification of urban areas by using marker-based hierarchical segmentation. *Photogram. Eng. Remote Sens* 80(10):963–970
- Akbari D, Ashrafi A, Attarzadeh R (2022) A new method for object-based hyperspectral image classification. *J. Indian Soc. Remote Sens* 50:1761–1771
- Aletti G, Benfenati A, Naldi G (2021) A semi-supervised reduced-space method for hyperspectral imaging segmentation. *J. Imaging* 7:267
- Atkinson PM, Tatnall ARL (1997) Introduction neural networks in remote sensing. *Int. J. Remote Sens* 18:699–709
- Bitam A, Ameur S (2013) A local-spectral fuzzy segmentation for MSG multispectral images. *Int. J. Remote Sens* 34:8360–8372
- Celeux G, Govaert G (1992) A classification EM algorithm for clustering and two stochastic versions. *Comput. Stat. Data Anal* 14(3):315–332
- Chang CI (2003) Hyperspectral imaging: techniques for spectral detection and classification. Kluwer Academic, Orlando, FL
- Chen Y, Jiang H, Li C, Jia X, Ghamisi P (2016) Deep feature extraction and classification of hyperspectral images based on

- convolutional neural networks. *IEEE Trans. Geosci. Remote Sens* 54(10):6232–6251
- M. Chi, Q. Kun, J. A. Benediktsson, and R. Feng, Ensemble classification algorithm for hyperspectral remote sensing data,” *IEEE Geoscience and Remote Sensing Letters*, vol. 6, no. 4, pp. 762–766, 2009. IEEE
- Cristianini N, Shawe-Taylor J (2000) An introduction to support vector machines and other kernel-based learning methods. Cambridge University Press
- S. Dutta, B. Manideep, S. Rai, and V. Vijayarajan, “A comparative study of deep learning models for medical image classification,” *Materials Science and Engineering Conference Series.*, IOP Publishing 042097 263, no. 4, 2017.
- Golipour M, Ghassemian H, Mirzapour F (2015) Integrating hierarchical segmentation maps with MRF prior for classification of hyperspectral images in a Bayesian framework. *IEEE Trans. Geosci. Remote Sens* 54(2):805
- Homayouni S, Roux M (2003) Material mapping from hyperspectral images using spectral matching in urban area. In: *Proceedings. IEEE Workshop on Advances in Techniques for analysis of Remotely Sensed Data.* NASA Goddard center, Washington DC, USA
- Hong D, Gao L, Yao J, Zhang B, Plaza A, Chanussot J (2021a) Graph convolutional networks for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens* 59(7):5966–5978
- Hong D, Han Z, Yao J, Gao L, Zhang B, Plaza A, Chanussot J (2021b) Spectral former: rethinking hyperspectral image classification with transformers. *IEEE Trans. Geosci. Remote Sens* 60:1–5
- Huang C-L, Wang C-J (2006) A GA-based feature selection and parameter optimization for support vector machines. *Expert Syst. Appl* 38:231–240
- Huang X, Zhang L (2009) A comparative study of spatial approaches for urban mapping using hyperspectral rosis images over pavia city, northern Italy. *Int. J. Remote Sens* 30(12):3205–3221
- Kang X, Li S, Benediktsson JA (2013) Spectral–spatial hyperspectral image classification with edge-preserving filtering. *IEEE Trans. Geosci. Remote Sens* 52:2666–2677
- Kang X, Li C, Li S, Lin H (2017) Classification of hyperspectral images by Gabor filtering based deep network. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens* 11:1166–1178
- Landgrebe DA (2003) *Signal theory methods in multispectral remote sensing.* John Wiley & Sons, Inc
- Li S, Song W, Fang L, Chen Y, Ghamisi P, Benediktsson JA (2019) Deep learning for hyperspectral image classification: an overview. *IEEE Trans. Geosci. Remote Sens* 57:6690–6709
- Lyu H, Mou L (2016) Learning a transferable change rule from a recurrent neural network for land cover change detection. *Remote Sens* 8(6):506
- Ojala T, Pietikäinen M, Mäenpää T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell* 7:971–987
- Pan E, Mei X, Wang Q, Ma Y, Ma J (2020) Spectral-spatial classification for hyperspectral image based on a single GRU. *Neurocomputing* 387:150–160
- Pesaresi M, Benediktsson JA (2001) A new approach for the morphological segmentation of high-resolution satellite imagery. *IEEE Trans. Geosci. Remote Sens* 39(2):309–320
- Rosenfield GH, Fitzpatrick-Lins K (1986) A coefficient of agreement as a measure of thematic classification accuracy. *Photogram. Eng. Remote Sens* 52(2):223–227
- Singhal N, Singhal N, Kalaichelvi V (2017) Image classification using bag of visual words model with FAST and FREAK. In: *International Conference on Electrical, Computer and Communication Technologies (ICECCT).* IEEE, Coimbatore, India
- Sivanandam S, Deepa S (2008) *Introduction to genetic algorithms.* Springer, Berlin/Heidelberg, Germany, pp 1–442
- Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R (2014) Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn Res* 15:1929–1958
- Sun L, Zhao G, Zheng Y, Wu Z (2022) Spectral-spatial feature tokenization transformer for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens* 60:5522214
- Tarabalka Y, Benediktsson JA, Chanussot J (2009) Spectral-spatial classification of hyperspectral imagery based on partitionial clustering techniques. *IEEE Trans. Geosci. Remote Sens* 47(9):2973–2987
- Tarabalka Y, Benediktsson JA, Chanussot J, Tilton JC (2010) Multiple spectral-spatial classification approach for hyperspectral data. *IEEE Trans. Geosci. Remote Sens* 48(11):4122–4132
- Tarabalka Y, Tilton JC, Benediktsson JA, Chanussot J (2011) A marker-based approach for the automated selection of a single segmentation from a hierarchical set of image segmentations. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens* 5(1):262
- Van der Meer F (2006) The effectiveness of spectral similarity measures for the analysis of hyperspectral imagery. *Int. J. Appl. Earth Observation Geoformation* 8(1):3–17
- Xu C, Lu C, Liang X, Gao J, Zheng W, Wang T, Yan S (2015) Multi-loss regularized deep neural network. *IEEE Trans. Circuits Syst. Video Technol* 26:2273–2283
- Zhao X, Tao R, Li W, Li H-C, Du Q, Liao W, Philips W (2020) Joint classification of hyperspectral and LIDAR data using hierarchical random walk and deep CNN architecture. *IEEE Trans. Geosci. Remote Sens* 58(10):7355–7370
- Zhuo L, Zheng J (2008) A genetic algorithm based wrapper feature selection method for classification of hyperspectral image using support vector machine. *ISPRS Archives* 7147:397–402
- F. Camastra, 1995“Signal theory methods in multispectral remote sensing,” DC Mall image and band specifications for the HYDICE Washington D.C. Mall image provided on the CD with. [Online]. Available: http://www.lars.purdue.edu/home/image_data/hydice_dc_wavelengths.html.
- Ding X, Zhang X, Han J, Ding G (2021) Diverse branch block: building a convolution as an inception-like unit. *Comput. Vis. Pattern Recognit*:10886–10895

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.