



# Optimal Feature Selection for the Classification of Hyperspectral Imagery Using Adaptive Spectral–Spatial Clustering

S. Chidambaram<sup>1</sup>  · A. Sumathi<sup>2</sup>

Received: 1 June 2018 / Accepted: 20 October 2018 / Published online: 30 October 2018  
© Springer Science+Business Media, LLC, part of Springer Nature 2018

## Abstract

Hyperspectral images captured through the hyperspectral sensors play an imperative part in remote sensing applications in the present context. Unlike traditional images sensed with few bands in the visible spectrum, the hyperspectral (HS) images are obtained with hundreds of spectral band ranges from infrared to ultraviolet regions. Because of its vast spatial and spectral data, it requires an extensive computational system for processing and its hidden features are needed to be unveiled in an effective manner specifically for the classification of HS imagery. This approach exploits the high spectral band correlation and rich spatial information of the HS images for the generation of feature vectors. To attain optimal feature space for the best probable classification, an adaptive approach is incorporated to adaptively choose spectral–spatial features for feature selection to classify the pixels effectively. Furthermore, the HS image encompasses several bands including noisy bands. To categorize the images with great accuracy, it is suggested to eradicate the noisy bands whilst retaining the informative bands. In this research, an adaptive spectral–spatial feature selection scheme is proposed for HS images where the extremely correlated representative bands are considered for analysis with uncorrelated and noisy spectral bands are judiciously discarded during its classification process. This hybrid approach not merely diminishes the computational time and also improves the general classification accuracy significantly. The empirical result displays that the proposed work surpasses the conventional approach of HS image classification systems.

---

✉ S. Chidambaram  
chidambaram.s@christuniversity.in

A. Sumathi  
hod\_ece@adhiyamaan.ac.in

<sup>1</sup> Department of Electronics and Communication Engineering, Christ University, Bengaluru, India

<sup>2</sup> Department of Electronics and Communication Engineering, Adhiyamaan College of Engineering, Hosur, India

**Keywords** Spectral–spatial features · Hyperspectral image · Clustering · Parallel classifiers · Classification

## 1 Introduction

RS stands as a technology that recognizes the attributes of an object of interest without having direct contact with that object. This technology makes utilization of the shape and size of an object. The reflected/radiated electromagnetic waves are obtained through sensors. The features of the reflected/radiated waves count on the condition of the objects. RS plays an imperative part in the prolonged observation of a specific area. It also enables the user to attain details of objects that are at a distant location. A HS image gathers and also processes the information over the electromagnetic spectrum. The objective of this is to get the spectrum of every pixel on the image. The reason is to find objects, identify materials or even detect disparate sorts of processes.

HS image sensor mounted in earth viewing airborne and spaceborne platforms for RS applications provide with abundance information concerning the region of our interest. These sensor types are operated in distinct wavelengths and capture the intent region with images of many narrow contiguous spectral bands called image cube. Generally, the HS sensors mounted in aircraft and satellite platforms could capture the images with the spectral extent of (380–12,700 nm and 400–14,400 nm) respectively. It is functioned by numerous government and commercial imaging agencies. A reflectance of specific spatial coordinates in all the existent spectral bands constitutes spectral signature which is unique for every object exists in the regions.

The compilation of spectral signatures of all the objects, materials and elements forms the spectral library furthermore it is utilized for classification of pixels with its similar classes. Generally, HS images would be classified effectively centered on its unique spectral–spatial features. Classification of pixels into various groups centered on spectral similarity approach is the vital tasks in RS domain predominantly in land classification, pattern recognition along with machine vision. The presence of vast spatial and spectral components in this type of image requires more memory and extensive computational efforts for analysis. The classification cost of HS image cube is reduced if bands or features that do not aid discrimination effectively are removed [1].

In machine learning, classification provides a way by which the computer program learns from the input data, after that this learning is utilized for classifying the new observation. The dataset might contain two classes or can contain many classes. Some of the commonly used types of classification algorithms in machine learning are linear classifiers, decision trees, SVM, random forest, neural network along with nearest neighbour algorithm.

This paper ponders on a proficient way of choosing the features that pave a means for the effective HS images classification. It is noted that in the feature selection (FS) method, the system lacked the capability to optimally select feature as per the feature weight. The proposed work aims at rendering a solution to resolve the problem.

Remaining paper is prearranged as; Sect. 2 manages the literature review. Section 3 summarizes the proposed method. Section 4 elaborates on the SVM and also the k-NN

classifiers. Section 5 depicts the investigational outcomes and discussions. Section 6 renders the computational expense analysis and Sect. 7 deduces the paper.

## 2 Related Works

Several classification methods [2–11] centred on spectral–spatial were developed. Numerous authors including spectral–spatial restraint suggested by Rongrong et al. [2] in which the pixels were classified centered on the constructed hypergraph including feature-based hyperedge and spatial based hyperedge. It was done based on the assumptions that the pixels, which were close to the feature region or space, and the pixels which were spatially closure maximally be a member of the same class. In Hanye et al. [3] incorporated spatial and spectral similarity measures for both dimensionality reduction and HS image classification. It was done centered on the observation that the pixels amid the close region were spatially related and spectral similarity measures exploited the redundancy for dimension reduction. In Ke et al. [4] suggested the ‘spectral frequency spectrum difference’ [SFSD] method which determines the spectral similarity in the frequency domain using Fourier transform with the concept that the attributes of the spectral signature can well be obtained clearly in the frequency domain. Erlei et al. [5] developed a sparse representation centered classifier with spectral information divergence [SID], exploiting the spectral discrepancy between the two pixels for the effective HS image classification. Hongzan et al. [6] developed an unsupervised spectral matching centred on artificial DNA calculating [UADSM] in which dynamic and artificial DNA calculating strategy was incorporated over a spectral signature for effective classification of HS images. HS image classification centered upon sparsity model [7] in which sparse representation was utilized to represent HS pixels. In this algorithm, two approaches were suggested to enhance the classification’s performance. Using explicit smoothing restraint, a pixel represented with similar spectral attributes and mixed pixels were represented by joint sparsity model. In Mahdi et al. [8] recommended a ‘spectral–spatial classifier’ that exclusively resolves the issue of combined pixels. Here the spectral information can well be characterized locally and globally for in-depth analysis for the determination of mixed components in every pixel. Xudong et al. [9] recommended an edge preserving filtering centred approach to improve classification accuracy exclusively for instantaneous applications.

Recently, the notion of manifold feature learning aimed at the classification of HS images proposed and reported in Jun et al. [10]. This framework was mainly to characterize linear and non-linear features exist in the data. Spectral signature/similar spectral statistical attributes [11] approach for the effectual classification of HS imagery was suggested in the previous work accounting of all the presented spectral bands. But all the aforementioned work primarily oriented toward for the development of optimal classification algorithm appropriate for various applications. In this work, to ameliorate the general classification accuracy of the classification system, the spectral bands with high correlation concerning reflectance alone needed to be regarded for experimentation in the test image data exclusively for the built out of real-time HS image classification system.

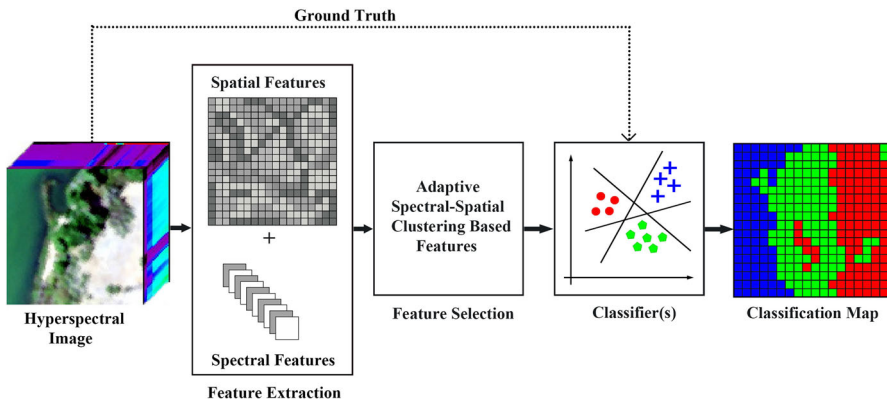


Fig. 1 The flowchart of the proposed hyperspectral image classification scheme

### 3 Proposed Methodology

In this study, an adaptive spectral–spatial clustering is employed with optimal FS centered on the predefined criteria by exploiting the high spectral band correlation along with rich spatial information of HS image dataset. It is done to generate concise and prime feature vectors which are to be incorporated for the built out of an effective HS image classification system aimed at instantaneous applications.

This proposed system framework considering spectral and spatial features for classification is delineated in Fig. 1. The HS image classification system generally focused on the assigning of every pixel to appropriate classes of interest. The classification of disparate classes in an image is primarily centered on the features extorted from spatial data, spectral data or both spectral–spatial data. In literature, the earlier approach of HS image classification is to allocate every pixel into one of the classes centered on the consideration of spectral characteristics alone for FS [12]. However, for an effective classification, the spatial features are as well considered along with spectral features reported in [13] in which each pixel is assigned to individual classes centered on spectral and spatial features. HS image classification with spectral–spatial features centered on binary tree representation has been proposed by [14]. The concept of 3D convolutional neural network based framework has been proposed by [15] for the extraction of combined spectral–spatial features for effective classification.

The FS method in the existent literature doesn't provide an optimal approach for choosing the relevant features. The proposed one facilitates the optimum FS. In certain instances of classification when considering both spectral–spatial features, there may be an existence of equal weight for more than two classes. Concerning feature space where the threshold function needed to be altered locally to determine the pixel under test is assigned to intend the class of interest. The solution to this point of context is to incorporate an adaptive nature of FS approach when the aforementioned situation emerged during the classification process.

Consider that the input  $n$ -pixel image together with  $m$ -spectral bands is signified as  $X = \{x_i \in \mathbb{R}^m; i = 1, 2, \dots, n\}$ . To lessen the dimensionality of the HS image for

an effective classification, representative bands and prototype pixels for every class are generated by making utilization of the redundancy between intra band and inter bands. And then spectral–spatial feature vectors are constructed and optimized for the development of appropriate classification model with various classifiers to assess the efficacy of the presented work.

The efficiency of a particular classifier mainly relies on the training data required for proper assignment of pixels into appropriate category accordingly. This proposed approach is well-matched for the case of in-sufficient training data when at least 2 classes ended-up with identical feature weight lead to spectral classes with the false assignment. Consider that the  $k$ -classes available in the HS image and it is represented as  $C_i ; i = 1, 2, \dots, k$ . With conditional probability [1], it can well be determined that the classes to which a pixel vector  $x$  belong furthermore it is written as,

$$P(w_i|x); i = 1 \dots \dots k \tag{1}$$

In which,  $w_i$  is the spectral class and  $x$  is the pixel vector.

The classification of the pixel vector  $x$  into a particular category is centered on the general rule that,

$$x \in w_i; \text{ if } P(w_i|x) > P(w_j|x) \text{ for all } j \neq i \tag{2}$$

where  $P$  denotes the probability,  $w_i$  is the spectral class and  $x$  is the pixel vector.

The above stated expression can as well be represented with discriminant functions,  $g_i(x)$  with maximal likelihood classification it is specified by,

$$x \in w_i; \text{ if } (g_i(x)) > (g_j(x)) \text{ for all } j \neq i \tag{3}$$

and

$$g_i(x) = \ln P(w_i) - \frac{1}{2} \left| \sum_i \right| - \frac{1}{2} (x - m_i)^t \sum_t^{-1} (x - m_i) \tag{4}$$

where  $m_i$  and  $\sum_t^{-1}$  are mean vector and covariance matrix respectively for the data in class  $w_i$ .

There might be predefined criteria to the extent at which each class to be categorized with independent threshold limit intended for all the classes. The decision was made with the threshold value furthermore it is devised as,

$$x \in w_i; \text{ if } (g_i(x)) > (g_j(x)) \text{ for all } j \neq i \text{ and } (g_i(x)) > T_i \tag{5}$$

and

$$g_i(x) = \ln P(w_i) - \frac{1}{2} \left| \sum_i \right| - \frac{1}{2} (x - m_i)^t \sum_t^{t-1} (x - m_i) > T_i \tag{6}$$

where  $T_i$  denotes the threshold for the spectral class  $w_i$ . In certain cases of insufficient training data, the feature space needs to be selected adaptively between spectral and spatial perspective when the given conditions are satisfied for at least 2 classes during the classification. Furthermore it is outlined in Algorithm 1 in which adaptation technique with optimal FS has been incorporated whenever uncertainty exists for the classification of pixels to disparate classes.

---

**Algorithm 1 : Adaptive Spectral-Spatial Feature Selection**

---

**Input :** A Hyperspectral image with n-pixels and m-spectral bands

$$X = \{x_i \in \mathbb{R}^m; i = 1, 2, \dots, n\}$$

**Output :** A classified image that represents k-classes

$$Y = \{C_j; j = 1, 2, \dots, k\}$$

**1 :** Compute spectral-spatial correlation ,  $\rho_{xy}, \rho_z$

**2 :** Generate representative bands ,  $M \ll m$

**3 :** Extract spectral-spatial features

**4 :** Compute optimal feature weight,  $\{w_i | w_j\}$

**5 :** begin

**6 :** if  $\{W_i \geq T_i\}$

**7 :**  $W \leftarrow W_i$

**8 :** else

**9 :**  $W \leftarrow W_j$

**10 :** end

**11 :** Response ,  $R \leftarrow k$

---

The input of the adaptive spectral-spatial FS is an HS image that contains n-pixels and m-spectral bands. The spectral-spatial correlation is computed for the input Hyperspectral image. Various representative bands are generated. This is followed by the extraction of spectral-spatial features. The optimal feature weights are determined. Centered upon these weights, the classified image representing k-classes are obtained as outlined in Fig. 2.

The proposed system of adaptive spectral-spatial FS aimed at HS image classification is tested on the SVM (Support Vectors Machine) due to the actuality that it yields a good result for the big dimensional data encompassing less number of training samples [16–19]. In input, a Hyper-spectral image with n-pixels and m-spectral band  $X = \{x_i \in \mathbb{R}^m; i = 1, 2, \dots, n\}$  along with extorted spectral-spatial features were given. With SVM classifier, the probability of the misclassification of classes is greatly reduced [20] and also proved that it is the best classifier on basis of minimizing classification error [21]. The notion of support vector introduced aimed at the classification of hyper-spectral RS image reported in [17] and also reported in [22]. Numerous variants of spectral and spatial classification schemes was suggested in earlier studies includes based on the mass voting rule proposed by [23] and with local and global probabilities developed by [8]. Further, the proposed scheme is also tested on the k-NN (k-Nearest Neighbor) along with its variants. SVM along with k-NN aimed at HS image classification is studied widely in [22, 24].

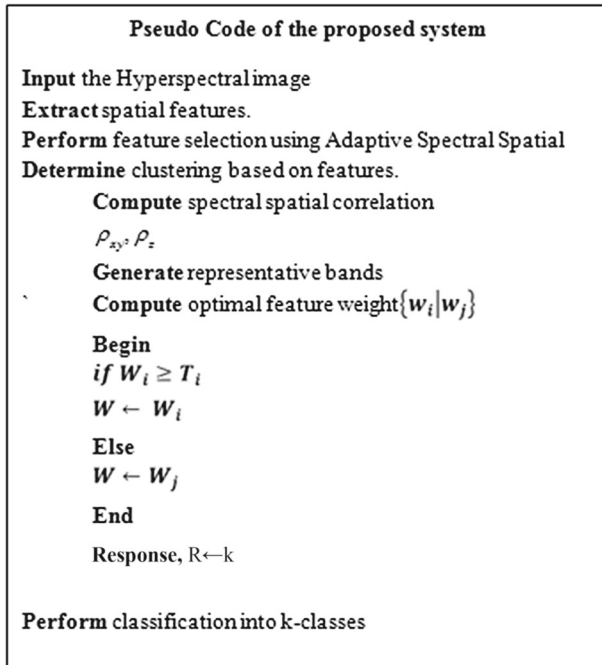


Fig. 2 Pseudo code of the proposed system

## 4 SVM And K-NN Classifiers

In this section, SVM along with K-NN classifiers are briefly discussed which were adopted to demonstrate the proposed FS scheme for the HS image classification.

### 4.1 Using Support Vector Machine Classifier

SVM stands as a discriminative classifier built out by Cortes and Vapnik [25] for binary classification defined by a separating hyperplane. The objective of SVM in HS image classification is to map the input vectors in to the higher-dimensional features space and an optimum separating hyper-plane is formed in this space [26]. However, the feature space's dimensionality is huge in HS image datasets and the pixels are linearly inseparable, the SVM incorporates kernel function for the mapping of pixels into higher-dimensional features space [27].

The widely used type of kernels is the 'radial basis function kernel'. This kernel function is popular in several 'kernelized learning algorithms'. In SVM classification, this kernel plays an imperative role. This kernel can well be implemented on 2 samples. These can well be signified as features vectors of some input space. This vector contains the 'squared Euclidean distance' between 2 feature vectors. This kernel lessens with distance and ranges between 0 and 1.

In HS image classification, it is often required discriminating above two classes and can be constructed using n-class classifier which selects the class centered on the maximal value functions specified in the kernel function. The methodology of the SVM for the input vector  $x = (x^1, x^2 \dots x^n)$ , which transforms the input pixel into higher dimensional features space centered on support vectors  $(x_1, x_2, \dots x_N)$  with the kernel function  $K(x, x_i)$ ,  $i = 1, 2, \dots, N$  and the decisions rule of any pixels under test is written as,

$$y = \sum_{i=1}^N y_i \alpha_i K(x, x_i) + b \quad (7)$$

where  $y_1 \alpha_1, y_2 \alpha_2 \dots y_N \alpha_N$  represents weights.

In SVM, the classification of data is centered on the determination of best hyperplane which takes apart data points of 1 class as of the other class. The most excellent hyperplane intended for an SVM stands as the one with the greatest margin between the 2 classes. The ‘support vectors’ denote the data points which are nearby the parting hyperplane and these are on the border of the ‘slab’.

The SVM classifier utilizes an iterative training algorithm that minimizes an ‘error function’. As per the type of ‘error function’, the models of SVM can well be split into 4 distinct groups. The first group refers to the C-SVM classification. The other groups of SVM are the classification SVM type two, Regression SVM Type one along with the Regression SVM Type two.

## 4.2 Using k-Nearest Neighbor Classifier

The k-NN stands as a method to classify an unknown data centered upon the known classification of its neighbors. When a compilation of samples with known classification is presented, each sample is classified likewise to its nearby samples and if the categorization of a sample is unknown, then the prediction is by regarding the classification of its NN samples. It is a learning algorithm that performs classification centered on the similarity of the surrounding data sets. “K” represents the number of data set items that are considered for the classification. The ‘similarity measure’ that is utilized in this technique quantifies the relationship amongst the different items.

There are 2 main considerations on the k-NN classifier. Firstly, the distance function on k-NN plays an indispensable role in the success of the classification. Generally, the smaller distance function used in the classification algorithm provides a greater likelihood for samples to be a member of the same class. Secondly, the option of the value for the parameter ‘k’ as it signifies the numbers of nearest neighbors to be regarded for the classification of unknown samples [28]. Typically, the unknown sample is classified centered on the samples of its ‘k-NN’ by majority vote method and it acts upon well with multimodal classes [29].



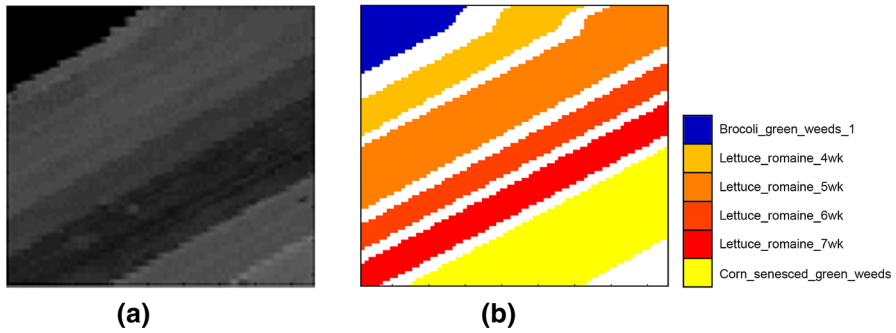


Fig. 3 Salinas-A image, **a** sample band and **b** ground truth

## 5 Experimental Results and Discussion

This section summarizes the findings of the proposed FS that was tested on two benchmark classifiers of SVM and k-NN with MATLAB's statistics along with machine learning toolbox. First, the experimental datasets are introduced and then experimental outcomes and discussion are given.

### 5.1 Experimental Datasets

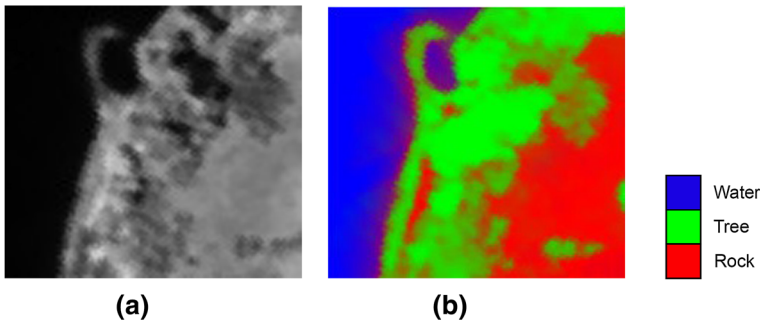
In this research, the proposed work is empirically tested on 2 well known HS datasets namely, Salinas-A and Samson available online [30] using the above-stated classifiers. The spectral and spatial information of these datasets is given as follows.

#### 5.1.1 Salinas-A Scene Dataset

The Salinas image comprising of  $512 \times 217$  pixels with 224 spectral bands attained by Airborne Visible Infrared Imaging Spectrometer sensor over Salinas Valleys, California with the spatial resolution of 3.7 m along with spectral wavelengths extending as of 400 nm through 2500 nm. The Salinas-A scene comprising of  $83 \times 86$  pixels is a small sub scene situated within the Salinas image at (samples, lines) = (591–676, 158–240) and it includes six distinct classes namely: Brocoli\_green\_weeds\_1, Corn\_senesced\_green\_weeds, Lettuce\_romaine\_4wk, Lettuce\_romaine\_5wk, Lettuce\_romaine\_6wk and Lettuce\_romaine\_7wk. There are twenty water absorption bands on the 'Salinas-A scene' including the bands 108 through 112, 154 through 167 and 224 are discarded with the remaining of 204 bands are considered for the experimentation. Figure 3a shows a sample band of Salinas-A and its ground truth displays in Fig. 3b. Table 1 demonstrates the Salinas-A reference data which consists of six classes and its number of samples.

**Table 1** Description of six distinct classes of Salinas-A scene

ID	Class	Number of samples
C1	Brocoli_green_weeds_1	391
C2	Corn_senesced_green_weeds	1343
C3	Lettuce_romaine_4wk	616
C4	Lettuce_romaine_5wk	1525
C5	Lettuce_romaine_6wk	674
C6	Lettuce_romaine_7wk	799

**Fig. 4** Samson image, **a** sample band and **b** ground truth**Table 2** Description of three distinct classes of Samson Image

ID	Class	Number of samples
C1	Rock	2920
C2	Tree	3541
C3	Water	2564

### 5.1.2 Samson Image Dataset

The Samson Image consisting of  $952 \times 952$  pixels with 156 spectral bands obtained by Spectroscopic Aerials Mapping System with On-board Navigation (SAMSON) of Oregon State University with the spatial resolution of 3.13 m and with spectral wavelengths ranging from 401 through 889 nm. In this experiment, a small sub scene of an image comprising  $95 \times 95$  pixels located at (252, 332) in the actual dataset was used and it includes three distinct classes namely: rock, tree, and water. Figure 4a shows a sample band of Samson and its ground truth is displayed in Fig. 4b. Table 2 demonstrates the Samson reference data that comprises 3 classes and its number of samples.

**Table 3** Classification results of Salinas-A scene using SVM classifiers

	L-SVM (%)	Q-SVM (%)	C-SVM (%)	FG-SVM (%)	MG-SVM (%)	CG-SVM (%)
Broccoli_green_weeds_1	97	95	97	97	97	97
Corn_senesced_green_weeds	99	97	99	96	97	93
Lettuce_romaine_4wk	98	98	97	97	97	96
Lettuce_romaine_5wk	99	99	99	99	99	99
Lettuce_romaine_6wk	98	96	96	98	98	98
Lettuce_romaine_7wk	98	98	98	93	100	100
OA	<b>98.3</b>	97.5	97.8	96.6	98	96.9

Bold indicates high overall accuracy among this six classifiers

## 5.2 Performance Evaluation

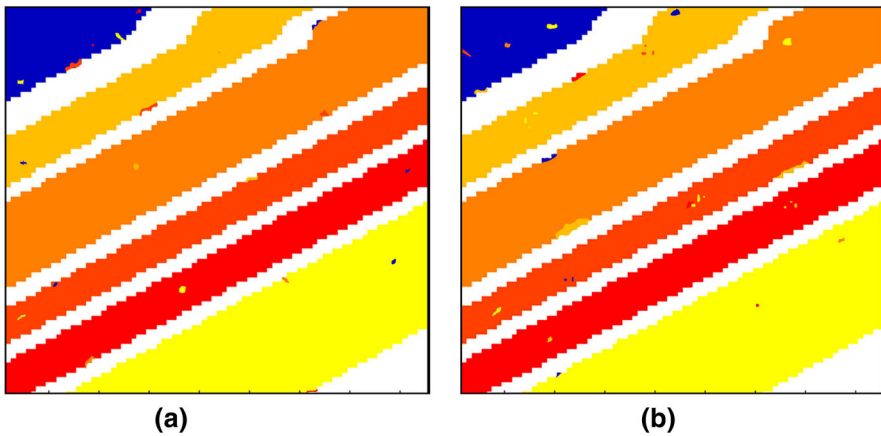
The classification outcomes demonstrated in this section made with different top-notch kernel functions of SVM and k-NN classifiers. Here, the outcomes of the proposed feature extraction technique are examined using various versions of SVM and K-NN with the metrics of classification accuracy, prediction speed and also training time. The effectiveness of classification with the variants of SVM such as linear (L), quadratic (Q), cubic (C), fine Gaussian (FG), medium Gaussian (MG) and coarse Gaussian (CG) is assessed. The comparative evaluation of these classifiers tested on Salinas-A scene. It shows that the linear SVM outperforms the other SVM classifiers with 98.3% considering overall accuracy. Similarly, the classification results is also compared with all the variants of k-NN which include fine (F), medium (M), coarse (C), cosine (Cos), cubic (Cu) and weighted (W). It specifies that the fine k-NN functions well with reference to its other types with 98.3% considering overall accuracy. The classification accuracy of the experimental datasets with different variants of these classifiers was assessed and compared. The classification outcomes of Salinas-A scene using the disparate versions of SVM and k-NN are tabularized in Tables 3 and 4 respectively. Figure 5a, b depicts the classification label of Salinas-A scene by L-SVM and F-kNN classifiers respectively. On Samson image, FG-SVM and C-kNN performed well with 96.9% and 96.5% respectively. The classification outcomes of Samson image using the disparate versions of SVM along with k-NN is tabularized in Tables 5 and 6 respectively. Figure 6a, b depicts the classification label of Samson image by FG-SVM and C-kNN classifiers respectively.

As HS image encompasses rich spatial and spectral components, its dimensionality has an effect on the consideration of spatial and spectral feature selection. This can well be solved and analyzed with the parallel coordinate plot, which is a technique proposed by Inselberg et al. [31]. And can be applied to a diverse set of multidimensional problems and for visualizing multivariate data and high dimensional geometry [32]. The Figs. 7 and 8 visualizing the spatial and spectral features of Salinas-A scene depicted as parallel coordinates which have been evaluated in different types of SVM and k-NN classifiers respectively. The images corresponding to these classifications

**Table 4** Classification results of Salinas-A scene using k-NN classifiers

	F-kNN (%)	M-kNN (%)	C-kNN (%)	Cos-kNN (%)	Cu-kNN (%)	W-kNN (%)
Broccoli_green_weeds_1	95	84	0	61	84	95
Corn_senesced_green_weeds	99	99	43	87	99	99
Lettuce_romaine_4wk	92	95	77	74	95	97
Lettuce_romaine_5wk	96	97	75	88	97	94
Lettuce_romaine_6wk	94	94	36	68	94	98
Lettuce_romaine_7wk	96	91	49	67	91	94
OA	<b>98.3</b>	94.4	51.1	75.8	94.4	96.3

Bold indicates high overall accuracy among this six classifiers

**Fig. 5** Classification results of Salinas-A scene by **a** L-SVM and **b** F-kNN**Table 5** Classification results of Samson image using SVM classifiers

	L-SVM (%)	Q-SVM (%)	C-SVM (%)	FG-SVM (%)	MG-SVM (%)	CG-SVM (%)
Rock	97	97	78	97	97	97
Tree	99	99	85	99	99	99
Water	94	94	76	94	94	94
OA	96.8	96.7	80.4	<b>96.9</b>	96.7	96.6

Bold indicates high overall accuracy among this six classifiers

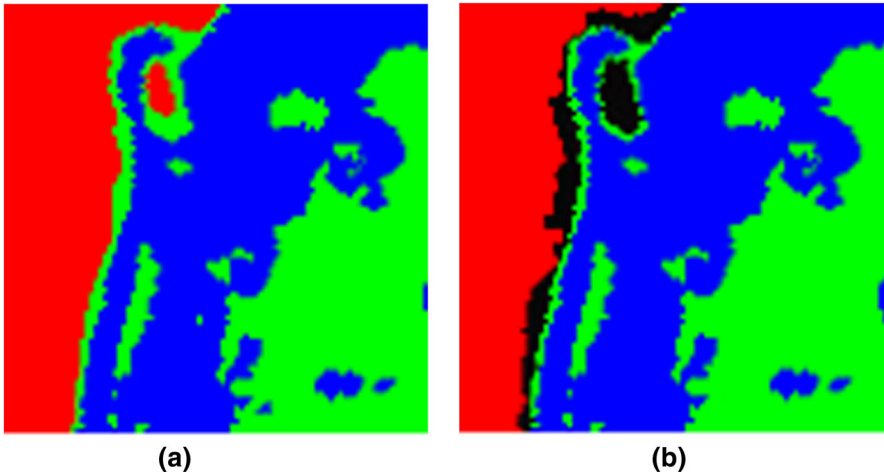
were depicted here but the other images were not shown here due to the space constraints. Figures 9 and 10 visualizing the spatial and spectral features of Samson image depicted as parallel coordinates which have been evaluated in different types of SVM and k-NN classifiers respectively.

The ‘Receiver Operating Characteristics’ (ROC) curve demonstrates the relationship between ‘true positive rate’ and ‘false positive rate’. The ROC curve of Salinas-A

**Table 6** Classification results of Samson image using kNN classifiers

	F-kNN (%)	M-kNN (%)	C-kNN (%)	Cos-kNN (%)	Cu-kNN (%)	W-kNN (%)
Rock	92	95	96	94	95	94
Tree	96	98	98	99	98	96
Water	97	95	94	95	95	95
OA	94.9	96.2	<b>96.5</b>	96.3	96.2	95.4

Bold indicates high overall accuracy among this six classifiers



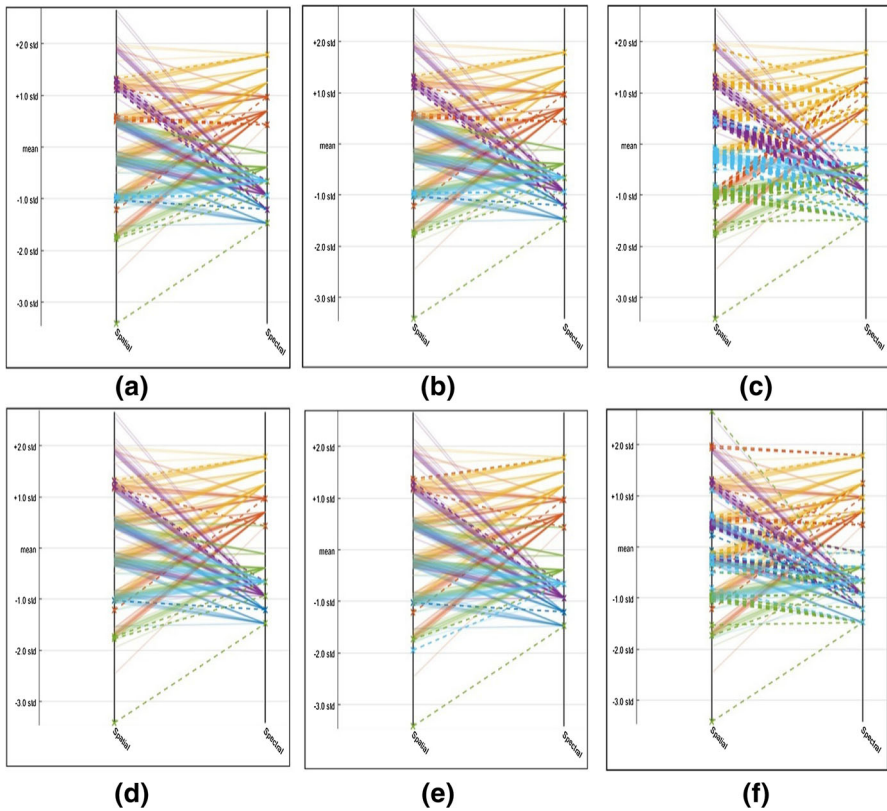
**Fig. 6** Classification results of Samson image by **a** FG-SVM and **b** C-kNN

scene obtained by SVM classifiers and also kNN classifiers are exhibited in Figs. 11 and 12 respectively.

The ‘ROC curve’ of Samson image obtained by SVM classifiers and also kNN classifiers are exhibited in Figs. 13 and 14 respectively.

## 6 Computational Cost Analysis

This section summarizes the computational cost analysis of the proposed work. The classification method is experimentally tested on an Intel Core i5-7200U 2.70 GHz with 8 GB RAM. The processing duration for the classification of two HS image datasets with distinct classifiers is reported. So as to increase the computational speed, we incorporated a parallel computing approach during the testing process. Moreover, in the process of HS image classification, simultaneous spectral–spatial FS and extraction enable the classification process ended up with less time consuming than traditional schemes. The parallel computing strategy can also be further upgraded using ‘Graphical Processing Unit’ (GPU) accelerated computing and the method of this kind has been proposed in [33]. The ‘GPU’ implementation of change detection

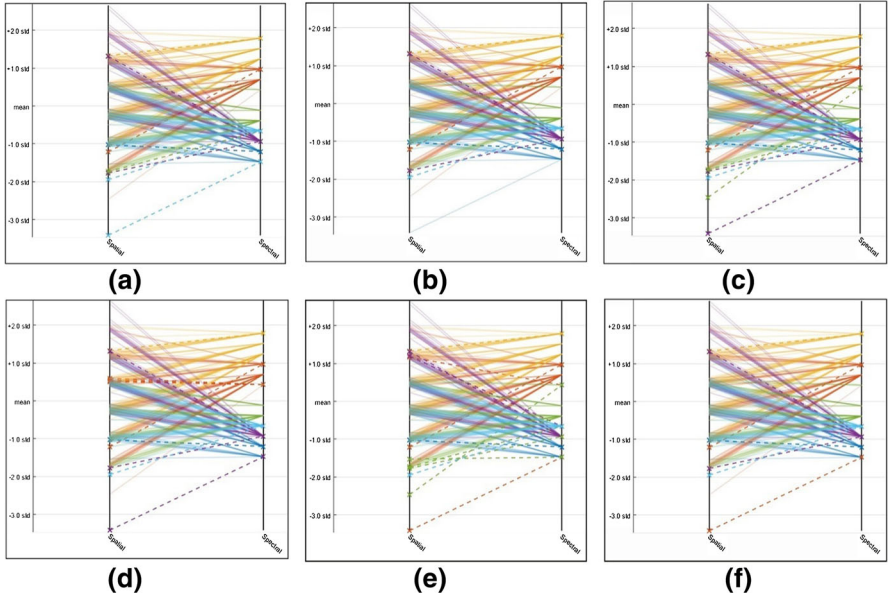


**Fig. 7** Spectral–spatial features of Salinas-A scene represented as parallel coordinates with SVM classifiers, **a** L-SVM, **b** Q-SVM, **c** C-SVM, **d** FG-SVM, **e** MG-SVM and **f** CG-SVM

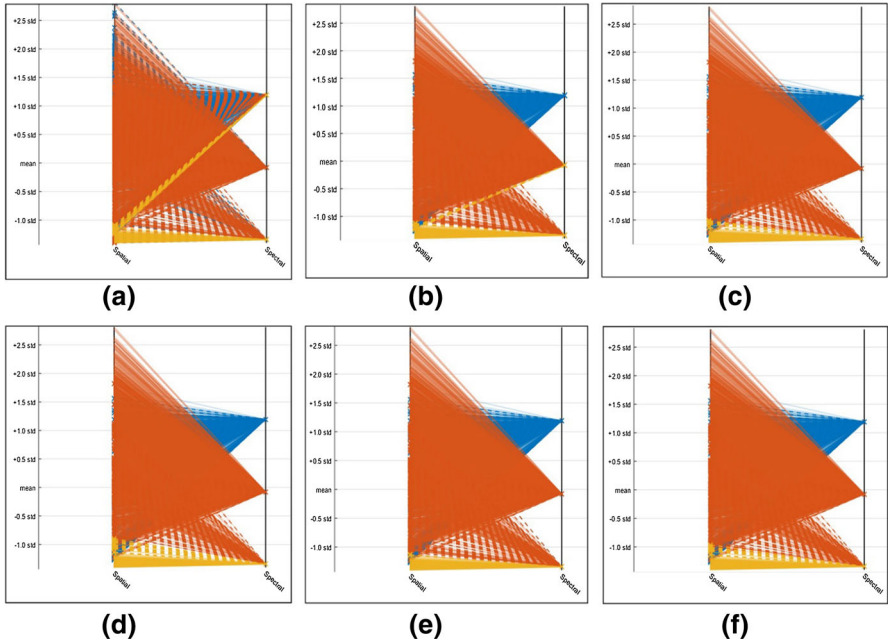
on multi-temporal HS images for instantaneous application was reported by [34]. In Tables 7 and 8, which reports the parallel computing performance of Salinas-A scene with prediction speed and training time evaluated by disparate types of SVM classifier in addition to k-NN respectively. It exhibits that the Q-SVM and F-kNN are outperforming the other variants of classifiers on Salinas-A scene with reference to training time. In Tables 9 and 10, which reports the parallel computing performance of Samson image with prediction speed and training time evaluated by disparate types of SVM and k-NN classifiers respectively. With Samson image, L-SVM and M-kNN performed well with reference to training time. This shows that the computational performance primarily relies on the characteristics of spatial–spectral feature vectors selected for categorization of HS images.

## 7 Conclusion

This work handles the feature selection and classification of the HS images. In this research, a new adaptive spectral and spatial centered FS is presented for an effective



**Fig. 8** Spectral–spatial features of Salinas-A scene represented as parallel coordinates with k-NN classifiers, **a** F-kNN, **b** M-kNN, **c** C-kNN, **d** Cos-kNN, **e** Cu-kNN and **f** W-kNN



**Fig. 9** Spectral–spatial features of Samson image represented as parallel coordinates with SVM classifiers, **a** L-SVM, **b** Q-SVM, **c** C-SVM, **d** FG-SVM, **e** MG-SVM and **f** CG-SVM

**Table 7** Parallel computing performance on Salinas-A scene using SVM classifiers

	L-SVM	Q-SVM	C-SVM	FG-SVM	MG-SVM	CG-SVM
Prediction speed	1800 obs/s	1700 obs/s	2700 obs/s	2800 obs/s	3400 obs/s	3500 obs/s
Training time	12.149 s	<b>10.398 s</b>	11.415 s	11.267 s	11.97 s	11.762 s

Bold indicates high speed among this six classifiers

**Table 8** Parallel computing performance on Salinas-A scene using k-NN classifiers

	F-kNN	M-kNN	C-kNN	Cos-kNN	Cu-kNN	W-kNN
Prediction speed	4700 obs/s	2900 obs/s	6300 obs/s	6000 obs/s	9100 obs/s	8400 obs/s
training time	<b>1.3241 s</b>	4.031 s	2.5493 s	2.4547 s	2.2902 s	2.5759 s

Bold indicates high speed among this six classifiers

**Table 9** Parallel computing performance on Samson image using SVM classifiers

	L-SVM	Q-SVM	C-SVM	FG-SVM	MG-SVM	CG-SVM
Prediction speed	56,000 obs/s	47,000 obs/s	80,000 obs/s	41,000 obs/s	49,000 obs/s	46,000 obs/s
Training time	<b>49.053 s</b>	195.07 s	55,805 s	200.44 s	205.37 s	210.75 s

Bold indicates high speed among this six classifiers

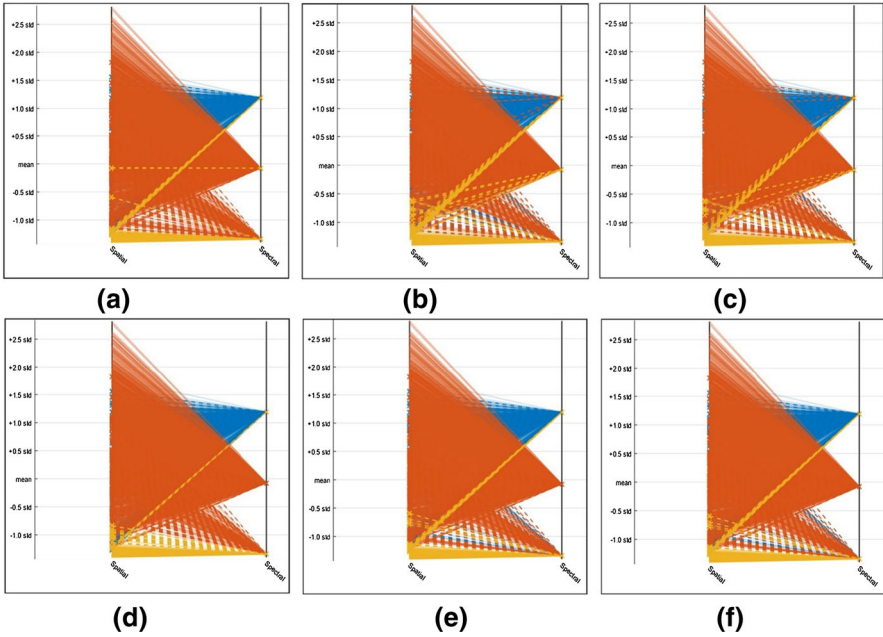
**Table 10** Parallel computing performance on Samson image using k-NN classifiers

	F-KNN	M-KNN	C-KNN	Cos-KNN	Cu-KNN	W-KNN
Prediction speed	66,000 obs/s	54,000 obs/s	28,000 obs/s	10,000 obs/s	45,000 obs/s	87,000 obs/s
Training time	2.0213 s	<b>1.7646 s</b>	2.5586 s	4.9984 s	3.4974 s	4.408 s

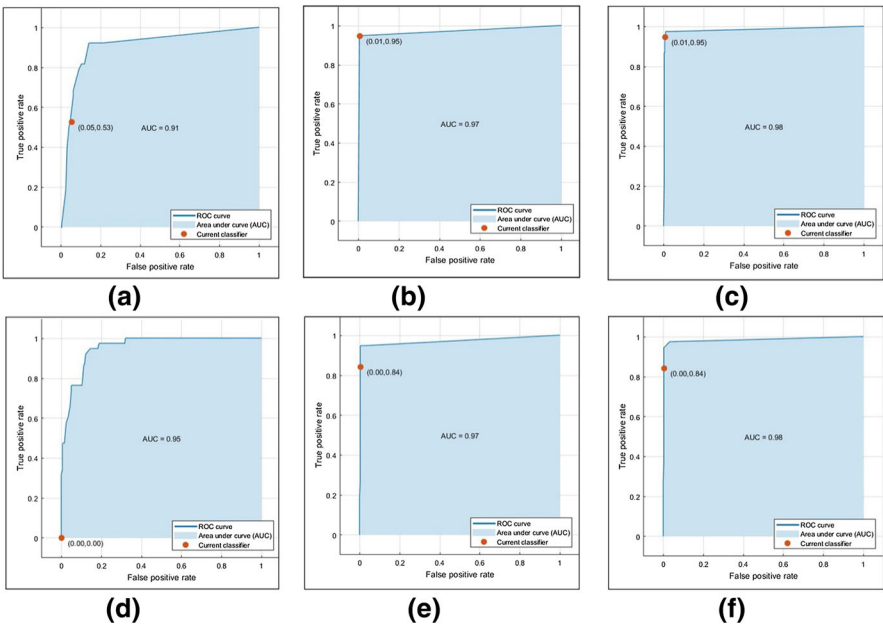
Bold indicates high speed among this six classifiers

classification of HS images. Unlike conventional FS methods reported in preceding works, the proposed work can adaptively choose the optimal spectral and spatial features as per the feature weight and pre-defined norm for the best potential classification of classes in the HS images. The investigational outcomes obtained by two benchmark classifiers on the 2 real HS images illustrate the efficiency of the proposed technique over traditional classifiers concerning the quantitative metrics like overall accuracy, prediction speed and training time with parallel computing approach. On the Salinas-A scene, L-SVM along with F-kNN individually yielded an accuracy of 98.3%. On Samson image, FG-SVM and C-kNN performed well with an accuracy of 96.9% and 96.5% respectively. This paper can well be extended to test on real-time images and the classification's accuracy can well be further enhanced.





**Fig. 10** Spectral–spatial features of Samson image represented as parallel coordinates with k-NN classifiers, **a** F-kNN, **b** M-kNN, **c** C-kNN, **d** Cos-kNN, **e** Cu-kNN and **f** W-kNN



**Fig. 11** ROC curve of Salinas-A scene obtained by SVM classifiers, **a** L-SVM, **b** Q-SVM, **c** C-SVM, **d** FG-SVM, **e** MG-SVM and **f** CG-SVM

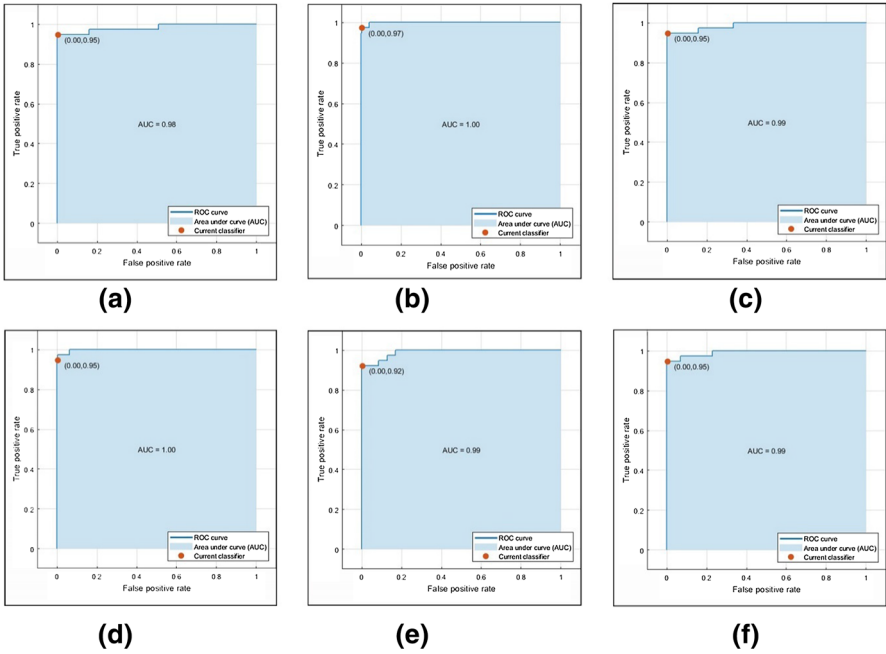


Fig. 12 ROC curve of Salinas-A scene obtained by k-NN classifiers, **a** F-kNN, **b** M-kNN, **c** C-kNN, **d** Cos-kNN, **e** Cu-kNN and **f** W-kNN

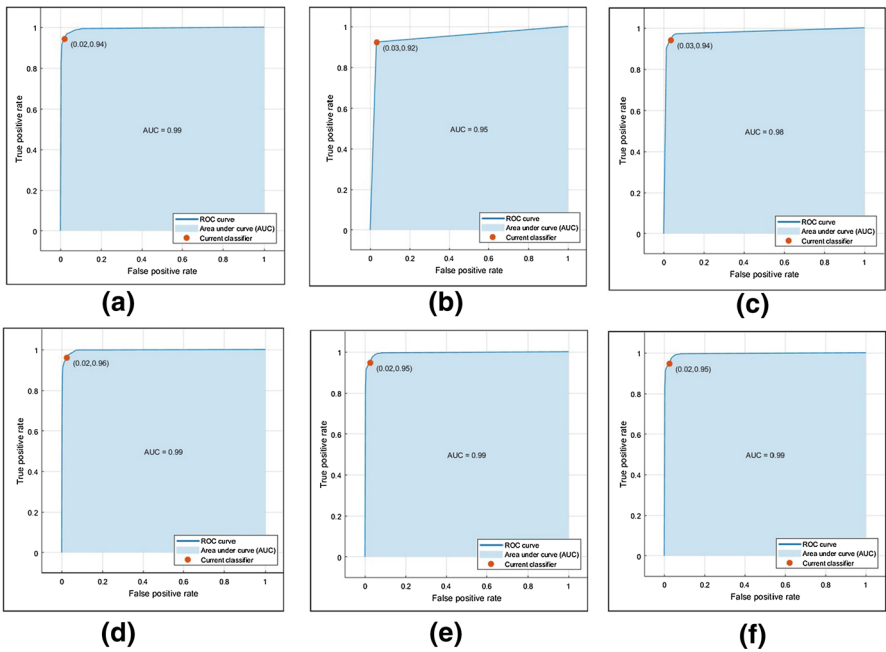
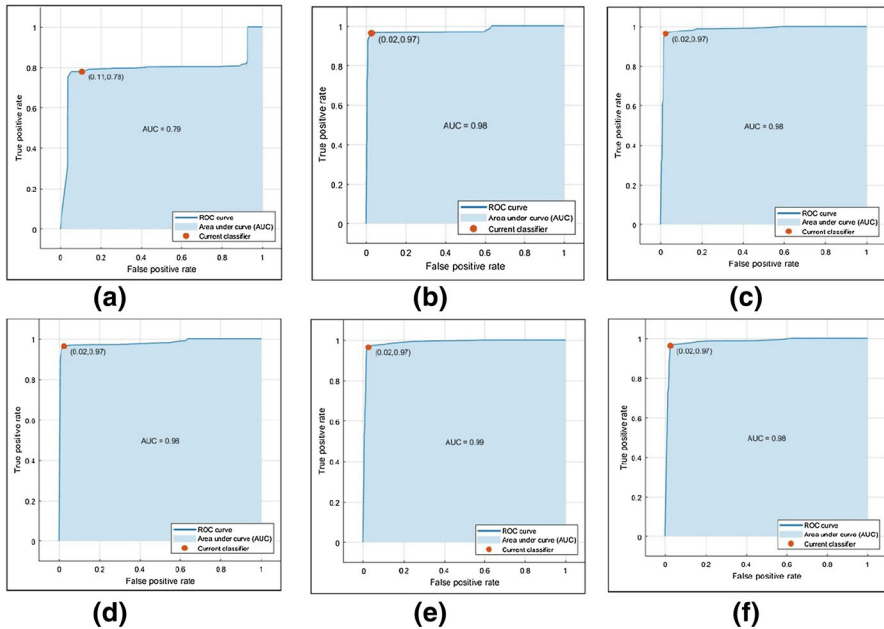


Fig. 13 ROC curve of Samson image obtained by SVM classifiers, **a** L-SVM, **b** Q-SVM, **c** C-SVM, **d** FG-SVM, **e** MG-SVM and **f** CG-SVM



**Fig. 14** ROC curve of Samson image obtained by k-NN classifiers, **a** F-kNN, **b** M-kNN, **c** C-kNN, **d** Cos-kNN, **e** Cu-kNN and **f** W-kNN

## References

- Richards, J.A., Jia, X.: Remote Sensing Digital Image Analysis: An Introduction, 4th edn. Springer, Berlin (2006)
- Ji, R., Gao, Y., Hong, R., Liu, Q., Tao, D., Li, X.: Spectral–spatial constraint hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **52**(3), 1811–1824 (2014)
- Pu, H., Chen, Z., Wang, B., Jiang, G.-M.: A novel spatial–spectral similarity measure for dimensionality reduction and classification of hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **52**(11), 7008–7022 (2014)
- Wang, K., Yong, B., Gu, X., Xiao, P., Zhang, X.: Spectral similarity measure using frequency spectrum for hyperspectral image classification. *IEEE Geosci. Remote Sens. Lett.* **12**(1), 130–134 (2015)
- Zhang, E., Zhang, X., Wang, S.: Improving hyperspectral image classification using spectral information divergence. *IEEE Geosci. Remote Sens. Lett.* **11**(1), 249–253 (2014)
- Jiao, H., Zhong, Y., Zhang, L.: An unsupervised spectral matching classifier based on artificial DNA computing for hyperspectral remote sensing imagery. *IEEE Trans. Geosci. Remote Sens.* **52**(8), 4524–4538 (2014)
- Chen, Y., Nasrabadi, N.M., Tran, T.D.: Hyperspectral image classification using dictionary-based sparse representation. *IEEE Trans. Geosci. Remote Sens.* **49**(10), 3973–3985 (2011)
- Khodadadzadeh, M., Li, J., Plaza, A., Ghassemian, H., Bioucas-Dias, J.M., Li, X.: Spectral–spatial classification of hyperspectral data using local and global probabilities for mixed pixel characterization. *IEEE Trans. Geosci. Remote Sens.* **52**(10), 6298–6314 (2014)
- Kang, X., Li, S., Benediktsson, J.A.: Spectral–spatial hyperspectral image classification with edge-preserving filtering. *IEEE Trans. Geosci. Remote Sens.* **52**(5), 2666–2677 (2014)
- Li, J., Huang, X., Gamba, P., Bioucas-Dias, J.M., Zhang, L., Benediktsson, J.A., Plaza, A.: Multiple feature learning for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **53**(3), 1592–1606 (2015)
- Chidambaram, S., Sumathi, A.: A novel spectral signature based classification approach for airborne and spaceborne hyperspectral imagery. *Asian J. Inf. Technol.* **15**(23), 4926–4933 (2016)

12. Landgrebe, D.: Hyperspectral image data analysis. *IEEE Signal Process. Mag.* **19**(1), 17–28 (2002)
13. Plaza, A., Benediktsson, J.A., Boardman, J., Brazile, J., Bruzzone, L., Camps-Valls, G., Chanussot, J., Fauvel, M., Gamba, P., Gualtieri, J.A., Marconcini, M., Tilton, J.C., Trianni, G.: Recent advances in techniques for hyperspectral image processing. *Remote Sens. Environ.* **113**(Supplement 1), S110–S122 (2009)
14. Maggiori, E., Tarabalka, Y., Charpiat, G.: Improved partition trees for multi-class segmentation of remote sensing images. In: *IEEE IGARSS*, pp. 1016–1019 (2015)
15. Li, Y., Zhang, H., Shen, Q.: Spectral–spatial classification of hyperspectral imagery with 3D convolutional neural network. *Remote Sens.* **9**, 67 (2017). <https://doi.org/10.3390/rs9010067>
16. Camps-Valls, G., Bruzzone, L.: Kernel-based methods for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **43**(6), 1351–1362 (2005)
17. Gualtieri, J.A., Cromp, R.F.: Support vector machines for hyperspectral remote sensing classification. *Proc. SPIE* **3584**, 221–232 (1998)
18. Fauvel, M., Chanussot, J., Benediktsson, J.A.: Evaluation of kernels for multiclass classification of hyperspectral remote sensing data. In: *Proceedings of ICASSP*, May 2006, pp. II-813–II-816 (2006)
19. BalaAnand, M., Karthikeyan, N., Karthik, S.: Designing a framework for communal software: based on the assessment using relation modelling. *Int. J. Parallel Program.* (2018). <https://doi.org/10.1007/s10766-018-0598-2>
20. Vapnik, V.N.: *Statistical Learning Theory*, vol. 2. Wiley, New-York (1998)
21. Lin, Y.: Support vector machines and the Bayes rule in classification. *Data Min. Knowl. Discov.* **6**, 259 (2002). <https://doi.org/10.1023/A:1015469627679>
22. Marconcini, M., Camps-Valls, G., Bruzzone, L.: A composite semisupervised SVM for classification of hyperspectral images. *IEEE Geosci. Remote Sens. Lett.* **6**(2), 234–238 (2009)
23. Tarabalka, Y., Benediktsson, J.A., Chanussot, J.: Spectral–spatial classification of hyperspectral imagery based on partitioning clustering techniques. *IEEE Trans. Geosci. Remote Sens.* **47**(8), 2973–2987 (2009)
24. Archibald, R., Fann, G.: Feature selection and classification of hyperspectral images with support vector machines. *IEEE Geosci. Remote Sens. Lett.* **4**(4), 674–677 (2007)
25. Cortes, C., Vapnik, V.: Support-vector network. *Mach. Learn.* **20**, 1–25 (1995)
26. Vapnik, V.: *The Nature of Statistical Learning Theory*. Springer, Berlin (1998)
27. Fang, L., Li, S., Duan, W., Ren, J., Benediktsson, J.A.: Classification of hyperspectral images by exploiting spectral–spatial information of superpixel via multiple kernels. *IEEE Trans. Geosci. Remote Sens.* **53**(12), 6663–6674 (2015)
28. Mucherino, A., Papajorgji, P.J., Pardalos, P.M.: k-nearest neighbor classification. In: Pardalos, P.M. (ed.) *Data Mining in Agriculture*. Springer Optimization and Its Applications, vol. 34. Springer, New York (2009). [https://doi.org/10.1007/978-0-387-88615-2\\_4](https://doi.org/10.1007/978-0-387-88615-2_4)
29. Kim, J., Kim, B.-S., Savarese, S.: 6th WSEAS International Conference on Computer Engineering and Applications, and Proceedings of the 2012 American conference on Applied Mathematics, pp. 133–138. Cambridge, Harvard, 25–27 January 2012. ISBN: 978-1-61804-064-0
30. AVIRIS Salinas Valley Hyperspectral Datasets. [http://www.ehu.es/ccwintco/index.php/Hyperspectral\\_Remote\\_Sensing\\_Scenes](http://www.ehu.es/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes). Accessed 19 Jan 2018
31. Inselberg, A.: The plane with parallel coordinates. *Vis. Comput.* **1**(2), 69–91 (1985)
32. Devi, G., Chauhan, C., Chakraborti, S.: Proceedings of the Sixth Symposium on Educational Advances in Artificial Intelligence (EAAI-16), pp. 4075–4080
33. Zebin, W., Wang, Q., Plaza, A., Li, J., Sun, L., Wei, Z.: Parallel spatial–spectral hyperspectral image classification with sparse representation and markov random fields on GPUs. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **8**(6), 2926–2938 (2015)
34. López-Fandiño, J., Heras, D.B., Argüello, F., et al.: GPU framework for change detection in multitemporal hyperspectral images. *Int. J. Parallel Program.* (2017). <https://doi.org/10.1007/s10766-017-0547-5>