Research Article
Mixed Pixel Wise Characterization Based on HDP-HMM and Hyperspectral Image Shape Detection using Hybrid Canny Edge Detection and WPDF

[1]D. Regan and [2]S.K. Srivatsa
[1]St. Peter's University,
[2]Prathyusha Institute of Technology and Management, Chennai, Tamilnadu, India

Abstract: Hyper Spectral Imaging (HSI) gathers and processes information from across the electromagnetic spectrum. The information enclosed in hyperspectral data permits the characterization, recognition and classification of the land-covers with enhanced accuracy and robustness. On the other hand, quite a lot of vital complications must be considered during the classification process of hyperspectral data, among which the maximum quantity of spectral channels, the spatial unevenness of the spectral signature, shape discovery of the images and the value of data. Above all, the maximum quantity of spectral channels and low number of labeled training samples pose the setback of the curse of dimensionality and, accordingly, result in the possibility of overfitting the training data. With the aim of solving all these complications, in this study presented the framework of Support Vector Machine (SVM) together with Fuzzy Sigmoid Kernel Function (SVM-FSK) in the circumstance of HSI classification and analyzing their features in the hyperspectral domain. A Kernel Fisher Discriminant Analysis (KFDA) model is employed for the purpose of dimensionality reduction of HSI. The KFDA dimensionality reduction scheme depends on the selection of the kernel in a higher-dimensional HSI feature space. In order to enhance the gradient level of spatial information, employed Improved Empirical Mode Decomposition (IEMD) with Gaussian Firefly Algorithm (GFA) (IEMD-GFA) to boost the mixed pixel wise SVM-FSK classification accuracy. During the process of IEMD scheme, the identifiable of Intrinsic Mode Functions (IMFs) of spectral band, weight values of IMFs are computed with the help of GFA. In order to identify the shape of HSI, novel hybrid scheme depending on the canny operator and fuzzy entropy theory is formulated. This scheme computes the fuzzy entropy of gradients from an image to make a decision on the threshold for the canny operator. For the purpose of detecting the edges and to discover the shape of the object Weibull Probability Density Function (WPDF) scheme is used. The obtained both spectral and spatial pixels are classified using SVM-FSK and estimated by using Hierarchical Dirichlet Process (HDP)-Hidden Markov Model (HMM). The proposed SVM-FSK is assessed with hyperspectral AVIRIS Indian Pine dataset. It shows that the proposed dimensionality reduction with SVM-FSK classification shows improved classification accuracy in terms of parameters like overall accuracy, standard deviation and mean.

Keywords: Canny Edge Detection, Gaussian Firefly Algorithm (GFA), Hierarchical Dirichlet Process (HDP), Hyperspectral image classification, Hidden Markov Model (HMM), Hidden Markov Model (HMM), Hyper Spectral Images (HSI), Pixel wise characterization, Support Vector Machine (SVM), Spectral Gradient enhancement, shape detection, Weibull Probability Distribution Function (WPDF)

INTRODUCTION

Hyper Spectral Images (HSI) are made up of hundreds of bands with an extremely high spectral resolution, from the perceptible to the infrared region. The extensive spectral range, combined with constantly increasing spatial resolution, permits to better differentiate materials and provides the capability to locate ground between spectrally close ground classes, making hyperspectral imagery appropriate for land cover classification. Owing to their characteristics, hyperspectral data have, at present, gained an incessantly growing interest among the remote sensing group of people (Yang and Li, 2012; Landgrebe, 2003). The one most important concern in the extremely high spectral resolution of remotely sensed hyperspectral data (Du, 2013), is the high dimensionality and it introduces a new challenge in the spectral-spatial feature extraction classification of the HSI. In order to overcome this complication, the proposed system makes use of a high dimensionality spectral image for the purpose of classification in HSI processing.

The enormous amount of information and the high spectral resolution of HSI offer the chance to resolve

complications which typically cannot be solved through multispectral images. On the other hand, quite a lot of significant issues required to be considered during the classification process for this category of images. During the classification of HSI, the higher dimensionality of the data boosts the potential to identify and differentiate several classes with better accuracy. HSI classification is the process which is employed to generate thematic maps from remote sensing image. Classification in remote sensing includes the process of clustering the pixels of an image to a set of classes in order that pixels in that particular class are having comparable features. Kernel-based classification scheme like Support Vector Machines (SVMs) has been famous and utilized broadly in the process of HSI classification in recent years (Anthony *et al*., 2007). The most important motivation of the attractiveness of kernel-based schemes and SVM specifically, are relatively higher classification accuracies and lower sensitivity to the high dimensions of data (Bazi and Melgani, 2006). One more useful feature of SVMs is their better generalization capability, resulting in sparse solutions (Bazi and Melgani, 2006). In case of kernel-based schemes, kernel functions are fundamentally utilized to discriminate among classes that are not linearly separable, by means of mapping the data to a higher dimensional space.

The large dimensionality of the data in the spectral domain guides to theoretical and practical complications for classification of HSI. This complication is solved with the help of a signal-analysis scheme like Empirical Mode Decomposition (EMD), which will produce a collection of Intrinsic Mode Functions (IMF) (Kim and Oh, 2009). The decomposition process of EMD is completely based on the magnitude of the original signal with a range of intrinsic time scales, explicitly; it decomposes the signal into dissimilar frequency elements. The EMD has been extensively employed in recent decades for the purpose of time-domain signal processing and was also employed to decompose the time-sequence signal in order to find out the intrinsic information (Chen and Jegen-Kulcsar, 2006). For effective functioning of EMD, the variations in both frequencies and amplitude have to be adequate for decomposition analysis. In case if the physical criteria for the variations between two signals are not satisfied, the sifting process obtains an IMF with single tone modulated in amplitude in place of a superposition of two unimodular tones. As a result, the modulated signal would no longer follow the features of the original signals. Hence, it is essential to overcome this complication of mode mixing and for this purpose proposed an improved Empirical Mode Decomposition.

EMD has been initially formulated for 1-D and 2-D signals in Amira-Biad *et al*. (2015) and Shi *et al*.

(2009), which makes use of geodesic operator-based morphological processes to identify the extremum points and Radial Basis Function (RBF) interpolation for the purpose of extracting the envelopes. A Kernel Fisher Discriminant Analysis (KFDA) scheme utilized a Weibull Probability Density Function (WPDF) for the purpose of extraction and estimates the edge characteristics of the HSI. 3-D EMD was formulated in Zemzami *et al*. (2013). In this study, the 3-D improved EMD method is presented.

However, the above mentioned classification scheme doesn't work under mixed wise categorization for HSI images and shape based edge detection is also not done at some stage in dimensionality reduction process in HSI images. On the other hand, none of the existing schemes pays attention to shape detection results and spectral information embedding in terms of mixed pixel representation and efficient characterization of spatial data representation, which is one of the foremost challenges in the HSI classification. In view of the fact that edge information from an image can be utilized to determine or identify objects in images. Edges signify considerable transformations in the image, preferably at the boundary among two different regions. Also, false edges are commonly detected and (parts of) vital edges are omitted. As a result, subsequent to edge detection there remains the setback of obtaining significant information regarding object boundaries with edges. Hence, it is believed that the classification of edges in, for instance, geometry edge, shadow edge, material edge or highlight edge, is helpful. Intensity-based edge detectors cannot differentiate the physical cause of an edge. A number of effective schemes for edge detection in normal (one-band) images are available (Verzakov *et al*., 2006). There are numerous methods to integrate the edge gradients computed from the several color bands.

In this study, hyper-spectra are therefore used to classify the edges. In addition, it is essential to find the way to employ the spatial information of the pixels to enhance performance of HSI classification still requires more investigation. In this study, the present work to hybrid Canny Edge Detection (HCED) schema is extended to detect the edges and shapes of the HSI images and classification scheme to operate on shape derived from hyper-spectra. Given the hyper-spectral gradients, subsequently the 3-D improved EMD (IEMD) is executed to each spectral band to get hold of a finite number of IMFs. Then Improved Empirical Mode Decomposition (IEMD) scheme is employed to enhance gradient level of spatial data in order to separate identifiable of Intrinsic Mode Functions (IMFs) of each band of the spectral data and optimum weights of the band obtained IMFs for reconstruction. Gaussian Firefly Algorithm (GFA) scheme is utilized to map all the optimum weights of the band obtained IMFs. The IMFs are summed by means of these

weights to reconstruct the features that are employed in mixed pixel wise classification framework is carried out in accordance with the semi hidden markov model in Support Vector Machine (SVM)-Fuzzy Sigmoid Kernel (FSK). It is shown in results that the proposed scheme provides a significant increase in accuracy for HSI classification. The main objective of this study is to completely use of spectral and spatial data for classification task to get exact land cover and land use class results. Classification of fewer number of HSI data samples has been more concentrated in a wide range of investigation recently.

LITERATURE REVIEW

Chen *et al.* (2011) formulated a new sparsity-based approach for the classification of hyperspectral imagery. Two different schemes were formulated to incorporate the contextual information into the sparse recovery optimization problem with the aim of improving the classification accuracy. In the first scheme, an explicit smoothing constraint is enforced on the problem formulation by means of forcing the vector Laplacian of the reconstructed image to become zero. In this scheme, the reconstructed pixel of significance has comparable spectral features to its four nearest neighbours. The second scheme is using a joint sparsity model in which hyperspectral pixels in a small neighbourhood in the region of the test pixel are concurrently represented by linear combinations of a small amount of common training samples, which are weighted with a diverse set of coefficients for each pixel.

Chen *et al.* (2013) formulated a novel nonlinear scheme for HSI classification. For every test pixel in the feature space, a sparse representation vector is acquired by means of decomposing the test pixel over a training dictionary, in addition in the same feature space, with the help of a kernel-based greedy pursuit approach. The recovered sparse representation vector is then employed directly to find out the class label of the test pixel. By means of projecting the samples into a high-dimensional feature space and kernelizing the sparse representation considerably develops the data separability among different classes, offering better classification accuracy.

Jabari and Zhang (2013) formulated a segmentation and fuzzy rule-based classification for extremely high resolution satellite images. The process of classifying extremely high resolution images is incredibly challenging because there are uncertainties in the location of the object borders. As a result, a fuzzy-rule based classification shows evidence of more promising solution to this challenging task. At first, the input image is segmented into shadows, vegetation and roads by means of eCognition software. Subsequently,

triangular and trapezoidal fuzzy functions are employed to allocate membership values to those segmented regions.

Ji *et al.* (2014) formulated a HSI classification scheme to consider both the pixel spectral and spatial parameters, in which the association among pixels is planned in a hypergraph structure. In the hypergraph, every vertex indicates a pixel in the HSI. And the hyperedges are composed from both the distance among pixels in the feature space and the spatial positions of pixels. To be exact, a feature-based hyperedge is produced with the use of distance among pixels, where each pixel is linked with its K nearest neighbors in the feature space. Subsequently, a spatial-based hyperedge is produced to model the outline among pixels by connecting where each pixel is connected with its spatial local neighbors. Both the learning on the combinational hypergraph is done through mutually examining the image feature and the spatial layout of pixels to look for their joint optimal partitions.

HSI classification approach by using pixel spatial relationship was formulated by Gao and Chua (2013). In case of HSIs, the spatial association among pixels has been revealed to be significant in the examination of pixel labels. In order to better employ the spatial information, it is essential to estimate the correlation surrounded by pixels in a hypergraph structure. In the constructed hypergraph, each pixel is indicated by a vertex and the hyperedge is built with the help of the spatial neighbors of each pixel. Semi-supervised learning on the constructed hypergraph is done for HSI classification.

Li *et al.* (2013) formulated a new structure for the improvement of generalized composite kernel machines for HSI classification. Build a new family of generalized composite kernels which show signs of enormous flexibility when integrating the spectral and the spatial information enclosed in the hyperspectral data, without any weight constraints. The classifier implemented in this scheme is the multinomial logistic regression and the spatial data is modeled from extended multiattribute profiles. With the intention of providing the better performance, SVM are also employed for evaluation purposes.

Demir and Erturk (2008) attempted to increase the classification accuracy of HSIs with the process of fusing spectral magnitude features and spectral derivative features. Principal Component Analysis (PCA) is employed as feature extraction scheme to decrease the final number of features of the hyperspectral data prior to feature fusion. PCA is executed independently to magnitude and derivative features to find out important constituents of each. Different fusion schemes of the important constituents of magnitude features and the important constituents of the first in addition to second spectral derivatives are

assessed to build the desired number of final features. SVM classification is employed for the purpose of classification of HSI following feature fusion.

Kalluri *et al.* (2010) formulated an effective scheme for the decision-level fusion of the spectral reflectance information with the spectral derivative information for robust land cover classification. This scheme is different from since an effective classification strategies is implemented to get rid of the increased over-dimensionality complication introduced by the addition of the spectral derivatives for hyperspectral classification.

Weng and Barner (2008) formulated a modified scheme for signal reconstruction depending on the EMD that improves the capability of the EMD to satisfy a specified optimality criterion. This reconstruction approach provides the best estimate of a particular signal in the minimum mean square error sense. Two different formulations are provided. The first formulation uses a linear weighting for the IMF. The second approach adopts a bidirectional weighting, specifically, it not only makes use of weighting for IMF modes, however also makes use of the correlations among samples in a specific window and performs filtering of these samples.

PROPOSED METHODOLOGY

In this study, largely concentrated on HSI images and demonstrates the most significant uniqueness of pixel for spatial and spectral domain. Having narrow band intervals allows the extension of discovery and classification actions to targets earlier not noticeable in multispectral images. For several applications, dimensionality reduction is an essential preprocessing phase to acquire a smaller set of characteristics that summarize the information in the HSI cube without losing any significant information and as a result circumvent 'the curse of dimensionality'. In HSI methods all of the methods extract spectral and spatial features pixels, which don't exactly extract shape of the image. This study proposes a novel canny edge detection methods to extract the edge features of the images and then the extracted edge features are estimated with the help of Weibull Probability Density Function (WPDF). Each observation from WPDF represents the shape of the image in accordance with the gradient value from hybrid canny edge detection method. In this study initially carry out dimensionality reduction with the assistance of Kernel Fisher Discriminant Analysis (KFDA). This method maps the entire hyperspectral data from the original feature space and maps the useful features. KFDA scheme is employed for the purpose of dimensionality reduction which represents a high dimensional feature space defined implicitly by means of a kernel function. Subsequently, Improved Empirical Mode Decomposition (IEMD) scheme is applied to improve gradient level of spatial data to separately identifiable of Intrinsic Mode Functions (IMFs) of each of band. In the process of IEMD scheme, IMFs weight values are determined by means of Gaussian Firefly Algorithm (GFA). IMFs are employed as feature data vector for Support Vector Machine–Fuzzy Sigmoid Kernel (SVM-FSK) classification framework. The FSK is employed as a kernel function in an SVM classification framework in order to classify even the mixed pixels HSI images. The objective function results of the FSK-SVM are estimated based on the Hierarchical Dirichlet Process (HDP)–Hidden Markov Model (HMM). Proposed work representation of the entire system is illustrated in Fig. 1.

Dimensionality reduction using Kernel Fisher Discriminant Analysis (KFDA): In this study, proposed a Kernel Fisher Discriminant Analysis (KFDA) scheme (He *et al.*, 2006) for the purpose of removing and reducing the feature space for HSI is defined implicitly by means of a kernel. The effectiveness of KFDA scheme is completely based on the selection of the kernel. Consider $X = (x_1, \ldots x_n)$, $x \in \chi$ represents the HSI samples which is a random subset of R_n, for every image samples the spatial and spectral features of the samples is indicated as $x \in \mathcal{F} = \{fs_1, \ldots fs_m\}$ as the feature from 1 to m. A symmetric function $K: \mathcal{F} \times \mathcal{F} \rightarrow R$ is known as kernel function among two hyper spectral features when it meets the finitely positive semi-definite characteristic: For any hyper spectral features samples $fs_1, \ldots fs_m$ the gram matrix $G \in \mathbb{R}^{m \times m}$ is given as:

$$G_{ij} = K(fs_i, fs_j) \tag{1}$$

K implicitly maps the input HSI samples with features F with a high-dimensional Hilbert space $\mathcal{H}$ equipped with the inner product $\langle ., . \rangle_{\mathcal{H}}$ by means of a mapping feature space $\phi: \mathcal{F} \rightarrow \mathcal{H}$:

$$K(o, p) = \langle \phi(o), \phi(p) \rangle_{\mathcal{H}}, \forall o, p \in \mathcal{F} \tag{2}$$

Inner product $\langle \phi(o), \phi(p) \rangle_{\mathcal{H}}$ is known as Hilbert space. This space is also called as feature space, it is completely based on the kernel function K and will be indicated as ϕ_k & $\mathcal{H}_k$. The feature space of the HSI samples is given in the type of objective function h(fs), in order to learn hyper spectral feature samples decision is done between two dissimilar feature samples in the kernel function $\mathcal{H}_k$:

$$h(fs) = sgn(w^t \phi_k(fs) + b) \tag{3}$$

where, $w \in \mathcal{H}_k$ indicates the feature vector weight values $b \in R$ be the bias value for feature samples is the intercept and:

$$sgn(u) = \begin{cases} \text{positive feature if } u > 0 \\ \text{negative feature if } u < 0 \end{cases} \tag{4}$$
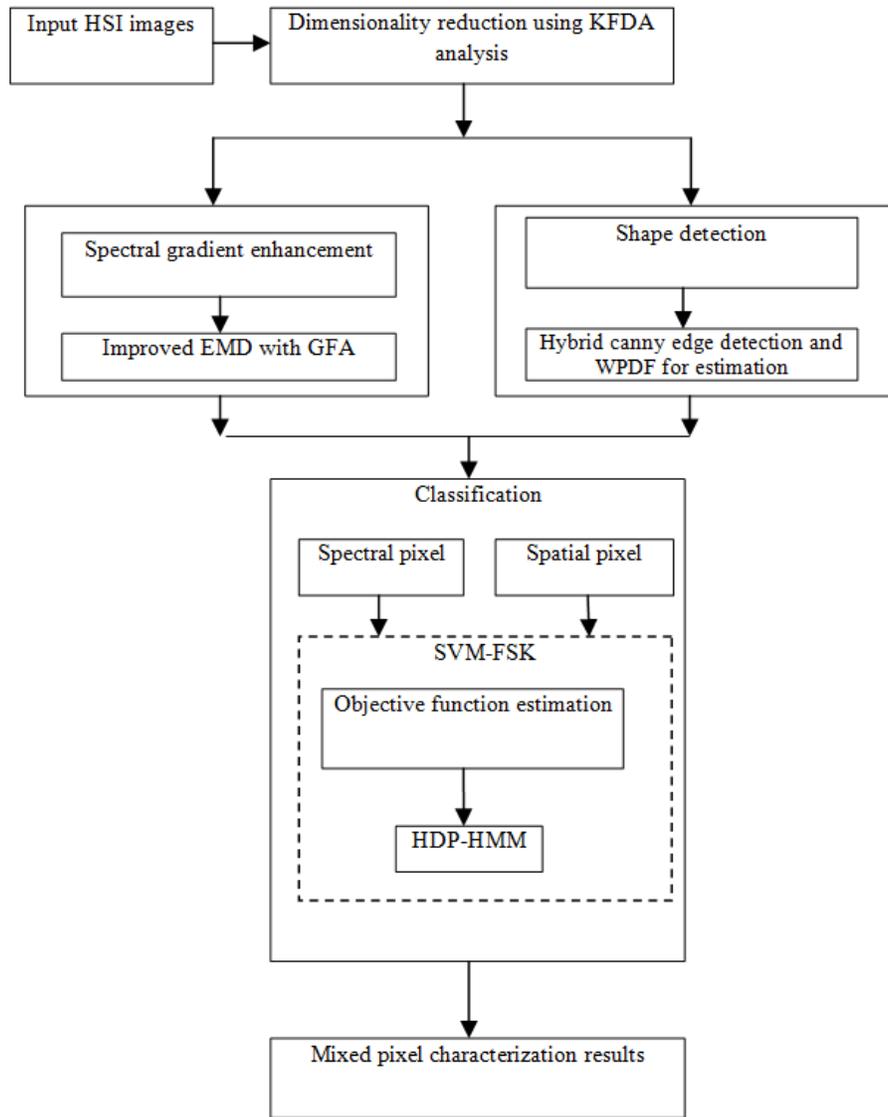
Fig. 1: Entire architecture of proposed work

The data which are necessary to perform KFDA are the means and covariances for dimensionality reduced features in the feature space. In actual fact, to carry out dimensionality reduction process for HSI samples KFDA with the sample means:

$$\mu_K = \frac{1}{m}\sum_{i=1}^{m} \phi_k(fs_i) \tag{5}$$

$$\Sigma_K = \frac{1}{m}\sum_{i=1}^{m} \phi_k(fs_i) - \mu_K \tag{6}$$

The fundamental concept of KFDA is to find a track in the HSI feature space $\mathcal{H}_k$ onto which the projections of the two sets between two different hyper spectral feature samples are well separated. Specifically, the separation between two different features of hyper spectral features is measured by the ratio of the variance and mean. Consequently, KFDA maximizes the FDR:

$$FD_\lambda(w, K) = \frac{\left(w^T(\mu_K)\right)^2}{w^T(\Sigma_K + \lambda I)w} \tag{7}$$

where, λ indicates a positive regularization parameter and I indicates the identity operator in HSI feature space $\mathcal{H}_k$. The weight value of the every feature space is decided in accordance with the Gaussian Firefly Algorithm (GFA), can show that the weight vector for feature space:

$$w^* = (\Sigma_K + \lambda I)^{-1} \cdot (\mu_K) \tag{8}$$

increases the FDR. The maximum FDR is realized by w^* is given by:

$$FD_\lambda^*(K) = \max_{w \in \mathcal{H}_k\{0\}} FD_\lambda(w, K) = (\mu_K)^T (\Sigma_K + \lambda I)^{-1} (\mu_K) \qquad (9)$$

Shape detection using hybrid canny edge detection and WPDF: During this phase, the shape is detected from HSI samples by applying Hybrid Canny Edge Detector to obtain edge pixels as Primitives. HSI samples are divided into sub-images with every sub-image then processed using a revised canny operator (Li *et al.*, 2014). The output image is constructed by means of assembling results from these sub-images. A typical canny operator makes use of the following phases to extract edges:

- Eliminate white Gaussian noise by means of smoothing the image using a Gaussian filter.
- Compute the magnitude and direction of the gradient at every pixel. Consider f(x, y) represents raw HSI Laplacian of Gaussian operator (LoG) has been found that makes use of the second-order image derivative, also known as Laplacian magnitude, for the purpose of extracting edges. The Laplacian magnitude is given as:

$$\nabla^2 f(x, y) = \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2} \qquad (10)$$

- When the gradient's magnitude at the processing point is bigger than the two neighbours' gradients, when placed in the gradient route, the processing pixel is spotted as the edge. If not, it is regarded as the background.
- Make use of hysteresis thresholding to eliminate weak edges.

Applying this scheme to real HSI samples observed that inaccuracies in segmentation outcome principally from step 3, because it is a process that is extremely sensitive to noise. This is the cause why an substitute scheme is formulated depends based on Fuzzy Entropy is initiated. Fuzzy entropy is a kind of entropy of a fuzzy set indicating the information of ambiguity for the entire categories of images and applications. It is normally employed to quantify the value of input image and P(X) as the probability mass function for HSI samples, subsequently the equivalent entropy H is given as follows:

$$H(x) = \sum_i P(x_i) HI(x_i) = -\sum_i P(x_i) \log_b P(x_i) \qquad (11)$$

where, HI indicates the information content of HSI samples x and b is the base of the logarithm.

In case of a HSI, there are typically two classes, objects and background. When a membership function is characterized, the degrees of spectral and spatial pixels belonging to the diverse sets can be computed. In accordance with the obtained memberships, spectral and spatial pixels can be divided into correct groups; the fundamental concept behind pixel clustering. Here, the Fuzzy Entropy theory is employed to differentiate gradients for edges from ones generated by noise. The gradients of the HSI produced with the canny operator are a set of values which can be categorized into two groups, edges and noise. At this point, the membership functions, μ, are given as:

$$\mu_x = \begin{cases} 0, x \\ \frac{b-x}{b-a}, a \le x \le b \\ 1, x > c \end{cases} \qquad (12)$$

As a result, the gradient increases the joint entropy of this fuzzy set is fixed as the threshold. The conception of joint entropy was initially formulated. Consider the entropy as H, in order that the entropy for set of edges from hyper spectral image A can be given as follows:

$$H(A) = -\sum_{j=1}^{L} \frac{\mu_x(j)h_j}{P(A)} \log \frac{\mu_x(j)h_j}{P(A)} \qquad (13)$$

where, P(A) represents the probability value of edges corresponding to the hyper spectral image A, $h_j = \frac{N_j}{N_{total}}$, N_j represents the number of points whose gradient is equivalent to j and N_{total} indicates the number of total points. Thus, the joint entropy is given as:

$$H(AB) = H(A|B) + H(B) \le H(A) + H(B) \qquad (14)$$

They happen to be equivalent if and only if these two sets of data are independent. Based on (14), it is clear that the process of thresholding is converted to that of discovering the maximum of the total entropy of gradients. The detected edge pixels from HSI samples are provided as two directions (x, y) ∈ HI. Edge pixels obtained from hybrid canny edge detection schemes to discover the edges of the HSI samples to determine the shape of the object. For this purpose, Weibull Probability Density Function (WPDF) probabilistic framework is employed. Hence, characterize possible edge pixels as continuous random PDF variables. Subsequently, estimate the shape edge pixels in accordance with their PDF by considering the extracted edge pixels as observations. WPDF with continuous random variables is characterized for extracted edge pixels. The Weibull distribution that functions in accordance with the shape parameter value is a versatile most extensively lifetime distributions in consistency engineering, consequently constituting the features of other categories of distributions. The 3-parameter WPDF is given as:

Table 1: Parameters of Gaussian Firefly Algorithm (GFA)

Parameters	Values	Description
α	0.2	Alpha
β_0	0.3	Beta$_0$
γ_F	0.2	Gamma
Iterations	20	Generations

$$N(x,y) = \frac{\beta}{\eta}\left(\frac{x-y}{\eta}\right)^{\beta-1} e^{-\left(\frac{x-\gamma}{\eta}\right)^{\beta}} \qquad (15)$$

where,

$$t \geq 0, \beta > 0, \eta > 0, -\infty < \gamma < +\infty$$

η = Scale parameter
β = Slope parameter
γ = Location parameter values for HSI

This WPDF meets the two constraints as:

$$\sum_x \sum_y N(x,y) = 1 \, \& \, N(x,y) \geq 0, \forall(x,y)$$

Improved EMD for spectral gradient: Subsequent to feature dimensional space results from the KFDA scheme, carried out mixed pixel wise characterization probabilistic classification framework to enhance the classification rate of HSI images. For that function, initially required to approximate gradient level of the HSI images in inside lessens feature space hyperspectral data for both spectral and spatial data. The gradient level of spectral data is approximated by means of Empirical Mode Decomposition (EMD) methods and spectral information data are reorganized with the help of IMFs. In this study, IMFs takes part considerable role to enhance the gradient level of spatial data, as a result the computation of weight values for IMF becomes also significant. The weight values of the IMF IMFSOW$_1$ are computed using Gaussian Firefly Algorithm (GFA) (Farahani *et al*., 2011), it enhances the classification accuracy rate for mixed pixel wise SVM-FSK hyperspectral data. As a result, the mean value of restructuring spectral gradient data is employed as the target purpose of the ABC:

$$f(IMFSOW_a) = SpecGrad =$$
$$\frac{1}{A \times B \times S}\sum_{m=1}^{A}\sum_{n=1}^{B}\sum_{b=1}^{S}\left|\frac{\partial FSO(m,n,b)}{\partial b}\right| \qquad (16)$$

where, FSO indicates the hyperspectral reduced feature dimensional space image sample, A and B represents the spatial dimensions and S represents the number of spectral bands. The weight values of the IMFs match to somebody of the FA with firefly initial population. IMFs weight values are computed through the searching flashing behavior of firefly. The weight values of IMF are revised in accordance with flashing behavior of the each firefly for weight values of IMF in the searching space. The present weight values of the IMFs for local HSI images and complete HSI image

samples weight values (IMFs) are compared against each others to choose best IMFs weight values to the overall gradient process. The current location for each firefly is to calculate the weight values of hyperspectral reduced feature dimensional space IMF is given as:

$$IMFSOW_a = (imfsow_{a1}, imfsow_{a2}, \ldots, imfsow_D\}$$

where, D denotes the dimension of IMFs weight values. In GFA, each IMF weight values are regarded as Fireflies. Yang employed this behavior of fireflies and launched Firefly Algorithm in Yang (2010). Subsequently, most significant three characteristics as mentioned above to carry out to calculate weight values for IMF function:

- The entire fireflies are unisex. As a result, IMF weight value will be attracted to other IMF weight value not considering their sex;
- Attractiveness β_o is relative in accordance with their brightness. As a result, for any two flashing IMF weight values, the less bright IMF weight value will shift in the direction of the brighter IMF weight value. When there is no brighter one IMF weight value is found than a specific weight value of IMF (firefly), it will travel arbitrarily;
- The brightness one weight value for IMF (firefly) is influenced as a result of fitness function from f(IMFSOW$_a$) (16). For a maximization setback, the brightness weight value can simply chosen in accordance with the fitness function from f(IMFSOW$_a$) (16). The parameter values of the GFA are given in Table 1.

Social behavior: Random walk is a executed for the purpose of selecting each IMF weight value arbitrarily based on Brownian motion. With the intention of moving the entire IMF weight values, random walk concept is used based on a Gaussian distribution that is given as:

$$p = f(IMFSOW_a|\mu, \delta) = \frac{1}{\delta\sqrt{2\pi}}e^{-(IMFSOW_a - \mu)^2/2\delta^2} \qquad (17)$$

where, er indicates an error between best IMF weight value solution and fitness value of IMF weight values (firefly):

$$er = f(g_{best}) - f(IMFSOW_a) \qquad (18)$$

$\mu = 0$ and $\delta = 1$, represents the mean and standard deviation function correspondingly. Social behavior of IMF, weight values (fireflies) is introduced by:

$$IMFSOW_{ai} = IMFSOW_{ai} + \alpha * (1 - p) * rand() \qquad (19)$$

where, α represents firefly parameter that is regulated through adaptive parameter scheme. In this scheme, GFA depends on the random walk each IMF weight

firefly IMFSOW$_a$ movement is pulled towards the best objective function is provided in Eq. (16).

Algorithm 1: Gaussian firefly algorithm (GFA)
Initialize algorithm parameters: Objective function of f(IMFSOW$_a$) from, where IMFSOW = (IMFSOW$_1$, … …, IMFSOW$_a$)^T
Create initial population of fireflies as number of weight values in IMF
Describe light intensity of I$_a$ at IMFSOW$_a$ by means of f(IMFSOW$_a$), from Eq. (16)
While (t < MaxGen)
For a = 1 to n (all n fireflies);
For b = 1 to n (all n fireflies);
If (I$_b$>I$_a$), move firefly a in the direction of b; end if
Attractiveness varies with distance r by means of Exp [-r2 and γF
Calculate new IMF, weight values solutions and revise light intensity;
End for b
End for a
Rank the current weight value for IMF and discover the current best weight values with the help of (16)
Define normal distribution
For c = 1, …, n all c fireflies
Draw a random IMF weight value and execute Gaussian distribution
Evaluate new solution (new solution(c))
Id new (new solution (c)<solution (a)&&(new solution (c) < last solution(c))
Move chosen IMF weight in the direction of spectral gradient enhancement
End if
End for c
End while;
Post process results and visualization;
End procedure;
Return the best weight values for IMF and its fitness value IMFSOW$_1$

The updated weight values of IMFs by means of their respective weights to get hold of the new hyperspectral data representation that will be employed for classification is given as:

$$RHIB = \sum_{d=1}^{R} IMFSOw_a \times IMFSO_a \qquad (20)$$

where, IMFSOW$_a$ demonstrates the equivalent weight of the IMF, R is the overall number of IMFs employed in the reconstruction process and RHIB indicates the reconstructed HSI band.
Then execute classification technique.

Spectral and spatial characterization using SVM-FSK: Then the gradient level of spectral and spatial information for HSI images are determined from improved EMD approaches and then execute probabilistic mixed pixel-wise SVM-FSK classification framework. The objective function outcome of SVM-FSK is estimated by means of HDP-HMM. In order to learn mixed pixel-wise SVM-FSK based classification results, approximate the HDP-HMM probability value to every spectral gradient information from improved EMD to enhance the classification outcome. Fuzzy sigmoid kernel function is employed as kernel function to SVM classification approaches with the intention of enhancing classification outcome of HSI images. SVM-FSK approach potentially recognizes appropriate and inappropriate features vectors through maximization of margin size among feature vectors. The exploitation of Fuzzy sigmoid kernel function discovers the maximization of margin hyperplane is converted in spatial domain version. In view of the fact that maximum margin classifiers are well standardized techniques and it doesn't corrupts the performance of classification for infinite dimensional data. Shorten the operation of mixed pixel wise SVM-FSK classification structure and determining the similarity between variables, it employs inner product as metric. In these classification approaches, when any dependent variables exist, those variables information might be lodged through supplementary dimensions and consequently can identify by a mapping (Samuelson and Brown, 2011). In this study, let SGFSII$_i$ = {SGFSII$_1$, SGFSII$_2$, ….SGFSII$_n$} represent the spectral gradient outcome from IEMD based GFA approach with diminished feature dimensional space. Also let SGFSII$_i$ = [SGFSII$_{i1}$, …., SGFSII$_{id}$]^T which represents the spectral gradient outcomes that related with reduced feature dimensional images pixel SGFSII$_i$ ∈ SGFSP. The SGFSP is characterized as a spectral gradient pixel SGFSP ∈ {1, 2,…., n} with indexing of n pixels of SGFSII$_i$ & SGFSB in which SGFSB indicates the amount of spectral gradient feature space bands. SGFSOI = [SGFSOI$_1$, SGFSOI$_2$, …, SGFSOI$_{1c}$, …, SGFSOI$_{1k}$]^T represent the classification results of SVM-FSK, in which C indicates the quantity of classes SGFSOI$_{ic}$ = {+1, -1, middle class} for c = 1, …C and $\sum_c$ SGFSOI$_{ic}$ = 1 ϕ (.) is indicated as non linear mapping function of gradient function, it is carried out in accordance with the Cover's theorem (Zemzami *et al*., 2013), which promises elevated classification accuracy rate for linearly separated feature vector samples and it is commonly higher dimensional feature space fs:

$$\min_{W,\xi_i,svmb} \left\{ \begin{array}{c} \frac{1}{2}||w||^2 + \beta_{cl} \\ \sum_i \xi_i \end{array} \right\} \qquad (21)$$

Constrained to the:

$$SGFSOI_i(\phi^T(j)W + ba) \geq 1 - \xi_q \qquad (22)$$

$$\forall i = 1, ... n$$
$$\xi_i \geq 0 \; \forall i = 1, ... n \qquad (23)$$

where, W&*svmb* represents a linear classifier for spectral gradient HSIs. Classification outcome of SVM approaches are managed by regularization constraint β and it is automatically selected by user, the error values of feature vectors are indicated by the parameter ξ_i. The result of mixed pixel wise SVM-FSK classification framework is approximated depending on probabilistic method HDP-HMM is given below:

$$SGFSOI_{ic} = \begin{cases} 1 \text{ if } p(SGFSOI_{ic} = 1|SGFSII_i, SGFSTR_l) \\ > p(SGFSOI_{ic_t} = 1|SGFSII_i, SGFSTR_l) \\ \forall c_t \neq c \\ 0, \text{otherwise} \end{cases} \qquad (24)$$

In order to enhance the classification accuracy, in this attempt employs a kernel function K:

$$K(SGFSOI_i, SGFSOI_i) = \phi(SGFSOI_i).\phi(SGFSOI_j) \qquad (25)$$

This kernel function result not enhances classification accuracy rate for certain data, to overcome these complication, in this study kernel function are approximated depending on fuzzy sigmiod function is defined in Eq. (24) given below:

$$f(SGFSOI) = sgn \sum_{i,j=1}^{n} SGFSOI_i SGFSOI_i \alpha_i \alpha_j$$
$$K(SGFSII_i, SGFSII_j) + svmb \qquad (26)$$

where, the SVM biases value (SVMb) of fuzzy kernel can be effortlessly computed from the α_q, it happens to be neither 0 nor C. This study extends the fundamental ideas of hyberbolic tangent function from (Soria-Olivas *et al.*, 2003) and it is given as follows (25):

$$\begin{aligned} &(SGFSII_i, SGFSII_j) = \\ &\begin{cases} -1 \; SGFSII_i.SGFSII_j \text{ is low} \\ +1 \; SGFSII_i.SGFSII_j \text{ is high} \\ m.SGFSOI_i.SGFSOI_i \; SGFSII_i.SGFSII_j \\ \quad\quad \text{ is medium} \end{cases} \end{aligned} \qquad (27)$$

where, m represents a constant value indicating the effectiveness of the sigmoid tract. In the statement of fuzzy logic idea, the sigmoid kernel function is defined as collection of fuzzy membership functions. Several fuzzy membership functions presents; however in this study only concentrated on three triangular function, owing to their simplicity. Subsequent to FSK function be continuous, as a result the expression (26) can be readily re-written as a function of a and γ, as follows:

$$K(SGFSII_i, SGFSII_i) = \begin{cases} -1 \; SGFSII_i.SGFSII_j \leq \gamma - \left(\frac{1}{a}\right) \\ +1 \; SGFSII_i.SGFSII_j \geq \gamma + \left(\frac{1}{a}\right) \\ 2(SGFSOI_i.SGFSOI_j - \gamma) \\ -a^2(SGFSOI_i.SGFSOI_j - \gamma) \\ |(SGFSII_i.SGFSII_j - \gamma)| \end{cases} \qquad (28)$$

which is the absolute form of the proposed fuzzy sigmoid (fuzzy tanh) kernel. In addition, measure the outcome of fuzzy sigmoid kernel function through objective function for I labeled training samples that is:

$$TS_l = (SGFSII_1, SGFSOI_1), ... (SGFSII_l, SGFSOI_l)$$

Based on the above discussed steps, the major function is to approximate the probabilistic value for mixed pixels $SGII_1$ with class label vector $SGFSOI_1$. This vector results can be obtained from HDP-HMM through the computing probability function:

$$SGFSOI_{ic} = \begin{cases} 1 \text{ if } p(c_{SGFSII} = k|SGFSOI, \theta) \\ > p(ct_{SGFSII} = k|SGFSOI, \theta) \\ \forall c_t \neq c \\ 0 \text{ otherwise} \end{cases} \qquad (29)$$

Probability estimation using HDP-HMM: In this research study, a concept of the Hidden Markov Model (HMM) with Hierarchical Dirichlet Process (HDP) is used and it is named as HDP-HMM, formalism for each spectral gradient input image with number of spectral gradient diminished feature space input sample states $SGFSII_q = \{SGFSII_1, SGFSII_2, ...SGFSII_n\}$, which includes number of output classes $SGFSOI_{qc} = \{+1, -1,$ middle class$\}$ for c = 1, ..., C of spectral gradient diminished feature space. In order to re-estimate the error found by SVM function, value for spectral gradient input image computation of backwards process from spectral gradient t input samples to t+1 input samples. The (F_t) variables point out that there is a Markov changeover at spectral gradient t input samples to t+1 input samples with diminished feature sample space and $F_t = 1$, as:

$$\beta_t(i) \triangleq p(SGFSOI_{t+1}: T|SGFSII_t = i, F_t = 1)$$
$$= \sum_j \beta_t^*(j)p(SGFSII_{t+1} = j|SGFSII_t = i)$$
$$\beta_t^*(i) \triangleq p(SGFSOI_{t+1}: T|SGFSII_{t+1} = i, F_t = 1)$$

$$= \sum_{d=1}^{T-t} \beta_{t+d}(i) \overbrace{P(D_{t+1} = d|SGFSII_{t+1} = i)}^{\text{duration of prior term}} \qquad (30)$$
$$\underbrace{P(SGFSOI_{t+1:t+d}|SGFSII_{t+1} = i, D = d)}_{\text{Likelihood term}}$$
$$\beta_t(i) \triangleq 1$$

In order carry out the above step in perfect way, divide the number of states into β, β* and make use of $SGFSOI_{t+1}$ to indicate SVM-FSK function estimate

Table 2: Notation used in proposed methodology

Notation	Explanation
x	Hyperspectral features training samples
$\{fs_1, \ldots fs_m\}$	Hyperspectral spectral and spatial features samples
$K: \mathcal{F} \times \mathcal{F} \to R$	Kernel function
$G \in \mathbb{R}^{m \times m}$	Gram matrix
ϕ	Kernel linear map function
H_k	Hillbert kernel space for feature samples
w	Feature vector weight values
$b \in R$	Be the bias value for feature samples
$Sgn(u)$	Sigmoid function for features
μK	Mean value for features
	Covariance matrix value for features
$\displaystyle\sum_K$	
$i = 1, \ldots n, j = 1, \ldots m$	Feature samples for HSI
$h(fs)$	Objective function for dimensionality reduction of features
$FD_\lambda(w, K)$	Fisher discriminant analysis
P(FS)	Gibb's prior to derive the principal component analysis
λ	is a positive regularization parameter
I	is the identity operator in HSI feature space H_k
F(x, y)	Hyperspectral raw image
$\nabla^2 f(x, y)$	Laplacian magnitude
$H(x)$	Entropy
H1	Information content of HSI
μx	Fuzzy membership function
N_j	is the number of points whose gradient is equal to j
N_{total}	Denotes the number of total points
N(x, y)	3-parameter Weibull pdf
η	Scale parameter
β	Slope parameter
γ	Location parameter values for HSI
A and B	are the spatial dimensions,
S	is the number of spectral bands
$IMFSOW_a$	Intrinsic mode function weight values
D	Denotes the dimension of IMFs weight values
FSo	hyperspectral reduced feature dimensional space image sample
$f(IMFSOW_a|\mu, \delta)$	Gaussian distribution
er	Error between best IMF weight value solution and fitness value of IMF Weight values
$\mu = 0$ and δ	be the mean and standard deviation function for firefly
α	Firefly parameter
I_b	Intensity value for IMF weight values
RHIB	The updated weight values of IMFs
R	Total number of IMFs used in the reconstruction,
$SGFSII_1$	Spectral gradient reduced feature space input sample
SGFSP	Spectral gradient pixel
SGFSOI	Spectral gradient reduced feature space output
$SGFSOI_{ic} = \{+1, -1, \text{middle class}\}$	Number of the classes for input sample
svmb	SVM Bias values for pixels
Bc_1	Regularization parameters for classification
ξ_i	Error values of feature vectors
m	Constant value representing the effectiveness of the sigmoid tract
$\gamma - \frac{1}{a}$ where $\gamma + \frac{1}{a}$	Fuzzy membership limits
$K(SGFSII_i, SGFSII_i)$	Fuzzy sigmoid (fuzzy $^{tanh \square}$) kernel.
l	Labeled training samples
$SGFSOI_{qc} = \{+1, -1, mc\}$	Spectral gradient output image pixels classes
t to t+1	Reduced feature HSI samples
F_t	Markov transition at spectral gradient t input samples
β, β^*	States of HMM for each hyper spectral reduced feature samples
$P(D_{t+1} = d|SGFSII_{t+1} = i)$	Duration of prior term for hyper spectral reduced feature samples
$P(SGFSII_{t+1; t+d}|SGFSII_{t+1} = i, D = d)$	Likelihood term
π_j	Initial probability matrix for hyper spectral reduced feature samples
z_s	As a super-state s for spectral gradient input samples with reduced dimensional space
$\{\omega_j\}_{j=1}^{\infty}$	To represent the parameters value for each one of the D_{t+1}
D	Denotes the $\{+1, -1, \text{middle class}\}$ class

results, subsequently HDP-HMM utilizes a prior suggestions to approximate SVM-FSK function for every spectral gradient reduced feature dimensional space sample:

$$\beta|\gamma \sim GEM(\gamma), \pi_j|\beta, \alpha \sim DP(\alpha, \beta),$$
$$\theta_j|H, \lambda \sim H(\lambda) \ \omega_j|\Omega \sim \Omega$$
$$\tau := 0, s := 1, \text{while } \tau < T \ do:$$
$$z_s|\{\pi_j\}_{j=1}^\infty, z_{s-1} \sim \tilde{\pi}_{z_{s-1}}$$
$$D_s|\omega \sim D(\omega_{z_s})$$
$$SGFSOI_s = SGFSOI_{\tau+1:\tau} + D_s + 1 \ |\{\theta_j\}_{j=1}^\infty, z_s, D_s \overset{iid}{\sim} f(\theta_{z_s})$$
$$\tau := \tau + D_s \ s := s + 1 \tag{31}$$

π_j represents the initial probability matrix for hyper spectral reduced feature samples, z_s indicates a super-state s for spectral gradient input samples with reduced dimensional space and $\{\omega_j\}_{j=1}^\infty$ to indicate the constraint values for each one of the D_{t+1}, with D indicates the {+1, -1, middle class} class. To estimate SVM-FSK, in HDP-HMM, calculate backward values for β and β^* in (30), subsequently posterior probability of input sample states $SGFSII_q = \{SGFSII_1, SGFSII_2,$ …., $SGFSII_n\}$ in the HMM state in accordance with:

$$p(SGFSII_1 = i|SGFSOI_{1:T})$$
$$\propto p(SGFSII_1 = i)p(SGFSOI_{1:T}|SGFSII_1 = i, F_0 = 1)$$
$$= p(SGFSII_1 = i)\beta_0^*(i) \tag{32}$$

The posterior probability value for input sample states $SGFSII_q = \{SGFSII_1, SGFSII_2, ….SGFSII_n\}$ in HMM state is represented via:

$$p(D_1 = d|SGFSOI_{1:T}, SGFSII_1 = \overline{SGFSII}_1, F_0 = 1)$$
$$= \frac{p(D_1=d)p(SGFSOI_{1:d}|D_1=d,SGFSII_1=\overline{SGFSII}_1,F_0=1)\beta_d(\overline{SGFSII}_1)}{\beta_0^*(\overline{SGFSII}_1)} \tag{33}$$

Reiterate the steps by considering $SGFSII_{D_1+1}$ with initial distribution given by $p(SGFSII_{D_1+1} = i|SGFSII_1 = \overline{SGFSII}_1)$. On the other hand, it is obvious that, when the state space of HMM is huge, some of other sequences might be also interested to carry out mixed pixel wise probabilistic estimation. The notation in the proposed work and their explanation is specified in Table 2,

EXPERIMENTAL RESULTS

Experimental results are provided for three hyperspectral data sets which are commonly used to evaluate the performance of hyperspectral classification algorithms, namely, the Indian Pine data in Table 3. The Indian Pine data set, which was taken over Indiana in 1992, consists of 145×145 pixels and 220 spectral bands. The spectral bands containing atmospheric noise and water absorption are removed, resulting in 200 bands. The spectral range of the data is 0.4-2.5μm and the spatial resolution is 20 m. Although the original ground-truth data contain 16 classes, the classes with a small number of samples are usually ignored and the

Table 3: Indian pine data

Class	No of samples
Corn-No till	1434
Corn-Min till	834
Grass/Pasture	497
Grass/Trees	747
Hay-Windrowed	489
Soybean-No till	968
Soybean-Min till	2468
Soybean-Clean	614
Woods	1294
Alfalfa	54
Building /Grass/Trees/Drives	380
Corn	234
Grass/Pasture-Mowed	26
Oats	20
Stone-steel-towers	95
Wheat	212

Fig. 2: The input image samples from Indian pine data se

Fig. 3: Noise incorporated input samples from Indian pine data set

Fig. 4: Shape detection using hybrid canny edge detection and WPDF from Indian pine data set

nine classes with a high number of samples. The Indian Pine data set is one of the most commonly used HSI. It is regarded to be a challenging data set, which is difficult to classify, because the spectral signatures of some classes are very similar and the pixels in the image are heavily mixed. Figure 2 Illustrates the input image samples from Indian Pine data set. After the noise values are added to input image samples is also illustrated in Fig. 3.

The shape detection results using hybrid canny edge detection and WPDF for Indian Pine data set samples are shown in Fig. 4. This Table 4 shows the

Table 4: Indian pine data results for spectral gradient

Training data	EMD	EMD-SEGA	EMD-SEPSO	EMD-SEGFA
10	93.7	94.5	95.42	96.56
20	93.81	94.65	95.51	96.38
30	93.98	94.78	95.64	96. 69
40	94.21	94.89	95.87	96.38
50	94.35	95.12	96.22	97.24

Table 5: Indian pine data classification results

	Mean (μ)			Standard deviation (σ)		
Training data	SVM-RBF	SVM-FSK	SVM-FSK after edge detection	SVM-RBF	SVM-FSK	SVM-FSK after edge detection
10	93.5	95.23	96.89	1.4	1.21	1.02
20	93.85	95.41	96.92	1.25	0.81	0.56
30	94.24	95.64	96.38	1.2	0.738	0.65
40	94.39	95.92	96.79	1.18	0.73	0.66
50	94.9	96.32	97.81	1.08	0.71	0.63
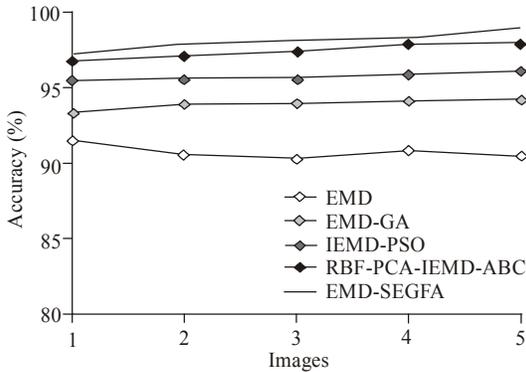
Fig. 5: Indian pine data results for spectral gradient
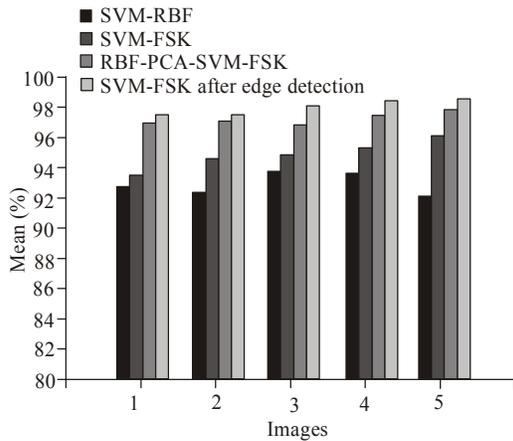
Fig. 6: Indian pine data classification results

overall accuracy of results for the original data representation as well as the EMD and the proposed EMD with spectral enhancement (denoted as EMD-SE) with PSO compare with existing EMD-SEGA.

Figure 5 shows the overall accuracy of results for the original Indian Pine data representation as well as the EMD and the proposed EMD with spectral enhancement and GFA is denoted as (EMD-SEGFA) compare with existing methods such as EMD-SEPSO,
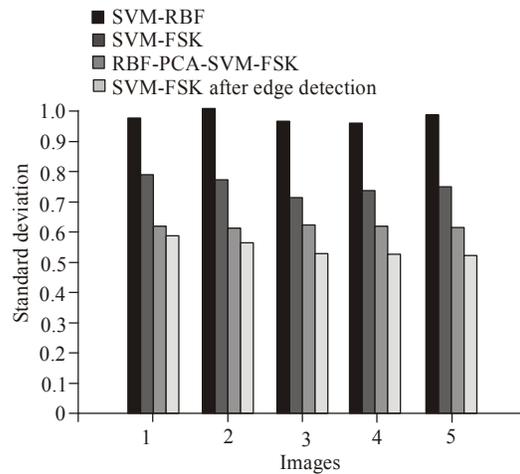
Fig. 7: Standard deviation for classification results

EMD-SEGA and EMD. It shows that proposed EMD-SEGFA spectral enhancement achieves higher 8.96% accuracy than earlier EMD-SEPSO methods, since the proposed work the shape detection is performed based on the Hybrid canny edge detection and estimated using WPDF, which produces exact classification results for all HSI images samples (Table 5).

Figure 6 classification performances are evaluated using mean function. The SVM classification with RBF kernel and SVM classification with Fuzzy sigmoid function of before and after edge detection. A classification result with respect to highest mean values shows the better accuracy results. It shows that proposed SVM-FSK (Support vector machine –Fuzzy Sigmoid Kernel) after the edge detection have achieves higher mean values when compare to before edge detection of the SVM-FSK and SVM-RBF (Support vector machine-Radial basis function) methods.

Figure 7 classification performances are evaluated using standard deviation function. The SVM classification with RBF kernel and SVM classification with Fuzzy sigmoid function of before and after edge detection. A classification result with respect to lowest standard deviation shows the better classification accuracy results. It shows that proposed SVM-FSK (Support Vector Machine-Fuzzy Sigmoid Kernel) after edge detection have less standard deviation than SVM-RBF (Support vector machine – Radial basis function) and before edge detection of SVM-FSK method, is illustrated in Fig. 7.

CONCLUSION AND RECOMMENDATIONS

In this study, the complication of dimensionality reduction and shape detection for HSI is taken into consideration in order to enhance classification accuracy. With this aim, this study presents Support Vetor Machine (SVM)-Fuzzy Sigmoid Kernel (FSK)

framework in the context of classification of hyperspectral data. Specifically, EMD-based scheme with spectral gradient enhancement has been formulated for HSI classification. EMD is employed for the purpose of decomposing hyperspectral bands into their IMFs. GFA-based optimization is then employed for the purpose of increasing the spectral gradient of the EMD-based representation by means of obtaining the weights of IMFs in an attempt to optimize the spectral gradient. At the same time, KFDA is used to reduce the data dimensionality, suppresses unwanted or interfering spectral signatures and identifies the occurrence of a spectral signature of interest. The fundamental idea is to project each pixel vector onto a fisher discriminant subspace with kernel values. In this study, hybrid Canny edge detection operator and the Fuzzy Entropy theory has been established to identify the shape of each object in HSI samples. The edge detection outcome is estimated based on the Weibull Probability Density Function (WPDF). SVM-FSK classifier is proposed for the purpose of classification of mixed pixels wise spectral spatial data. The SVM-FSK results is estimated in accordance with the HDP-HMM for mixed pixelwise characterization of complete image and a set of previously derived class combination maps, correspondingly. The proposed approach SVM-FSK with HDP-HMM, which intends to characterize mixed pixels in the scene and assumes that these pixels are normally mixed by only a few constituents, offers certain distinctive features with regards to other existing schemes. Experimental results confirm that the proposed SVM-FSK approach provides a significant increase in class separability for both synthetic and real hyperspectral scenes of various classification methods. The approach is applicable to both spectrally pure as well as mixed pixels. Further developments of this study include a comprehensive research of the influence of the KFDA algorithm used to enforce investigation of the possibility of including contextual spatial information within the SVM-FSK framework. The SVM is computationally less demanding for small training sets, so which is solved by using applying other classification methods such as Fuzzy Neural Network (FNN), Extreme Learning Machine (ELM) and their types, online dictionary learning schemes, etc to reduce the time complexity during classification process.

REFERENCES

Amira-Biad, S., T. Bouden, M. Nibouche and E. Elbasi, 2015. A bi-dimensional empirical mode decomposition based watermarking scheme. Int. Arab J. Inf. Techn. (IAJIT), 12(1).

Anthony, G., H. Gregg and M. Tshilidzi, 2007. Image classification using SVMs: One-against-one vs one-against-all. Proceeding of the 28th Asian Conference on Remote Sensing.

Bazi, Y. and F. Melgani, 2006. Toward an optimal SVM classification system for hyperspectral remote sensing images. IEEE T. Geosci. Remote, 44(11): 3374-3385.

Chen, J. and M. Jegen-Kulcsar, 2006. The Empirical Mode Decomposition (EMD) method in MT data processing. Proceeding of Kolloquium Elektromagnetische Tiefenforschung Conference, pp: 67-76.

Chen, Y., N.M. Nasrabadi and T.D. Tran, 2011. Hyperspectral image classification using dictionary-based sparse representation. IEEE T. Geosci. Remote, 49(10): 3973-3985.

Chen, Y., N.M. Nasrabadi and T.D. Tran, 2013. Hyperspectral image classification via kernel sparse representation. IEEE T. Geosci. Remote, 51(1): 217-231.

Demir, B. and S. Erturk, 2008. Spectral magnitude and spectral derivative feature fusion for improved classification of hyperspectral images. Proceeding of IEEE International Geoscience Remote Sensing Symposium, 3: III-1020-III-1023.

Du, Q.D., 2013. Foreword to the special issue on hyperspectral remote sensing: Theory, methods and applications. IEEE J. Sel. Top. Appl., 6(2): 459-465.

Farahani, S.M., A.A. Abshouri, B. Nasiri and M.R. Meybodi, 2011. A Gaussian firefly algorithm. Int. J. Mach. Learn. Comput., 1(5): 448-453.

Gao, Y. and T.S. Chua, 2013. Hyperspectral image classification by using pixel spatial correlation. In: Li, S. *et al.*, (Eds.): MMM, 2013. Part I, LNCS 7732, Springer-Verlag, Berlin, Heidelberg, pp: 141-151.

He, B., L. Tong, X. Han and W. Xu, 2006. SAR image texture classification based on kernel fisher discriminant analysis. Proceeding of the IEEE International Conference on Geoscience and Remote Sensing Symposium (IGARSS, 2006), pp: 3127-3129.

Jabari, S. and Y. Zhang, 2013. Very high resolution satellite image classification using fuzzy rule-based systems. Algorithms, 6: 762-781.

Ji, R., Y. Gao, R. Hong, Q. Liu, D. Tao and X. Li, 2014. Spectral-spatial constraint hyperspectral image classification. IEEE T. Geosci. Remote, 52(3): 1811-1824.

Kalluri, H.R., S. Prasad and L.M. Bruce, 2010. Decision-level fusion of spectral reflectance and derivative information for robust hyperspectral land cover classification. IEEE T. Geosci. Remote, 48(11): 4047-4058.

Kim, D. and H.S. Oh, 2009. EMD: A package for empirical mode decomposition and hilbert spectrum. R J., 1(1): 40-46.

Landgrebe, D.A., 2003. Signal Theory Methods in Multispectral Remote Sensing. John Wiley and Sons, Hoboken, NJ.

Li, J., P.R. Marpu, A. Plaza, J.M. Bioucas-Dias and J.A. Benediktsson, 2013. Generalized composite kernel framework for hyperspectral image classification. IEEE T. Geosci. Remote, 51(9): 4816-4829.

Li, Y., S.Y. Cho and J. Crowe, 2014. A hybrid edge detection method for cell images based on fuzzy entropy and the canny operator. J. Image Graph., 2(2).

Samuelson, F. and D.G. Brown, 2011. Application of Cover's theorem to the evaluation of the performance of CI observers. Proceeding of the 2011 International Joint Conference Neural Networks (IJCNN), pp: 1020-1026.

Shi, W., Y. Tian, Y. Huang, H. Mao and K. Liu, 2009. A two-dimensional empirical mode decomposition method with application for fusing panchromatic and multispectral satellite images. Int. J. Remote Sens., 30(10): 2637-2652.

Soria-Olivas, E., J.D. Martin-Guerrero, G. Camps-Valls, A.J. Serrano-Lopez, J. Calpe-Maravilla and L. Gomez-Chova, 2003. A low-complexity fuzzy activation function for artificial neural networks. IEEE T. Neural Networ., 14(6): 1576-1579.

Verzakov, S., P. Paclík and R.P. Duin, 2006. Edge detection in hyperspectral imaging: Multivariate statistical approaches. In: Yeung, D.Y. *et al.* (Eds.), SSPR&SPR, 2006. LNCS 4109, Springer, Berlin, Heidelberg, pp: 551-559.

Weng, B. and K.E. Barner, 2008. Optimal signal reconstruction using the empirical mode decomposition. EURASIP J. Adv. Sig. Pr., 2008: 845294, Doi:10.1155/2008/845294.

Yang, X.S., 2010. Firefly Algorithm. In: Engineering Optimization. John. Wiley & Sons, Inc., pp: 221-230.

Yang, X. and J. Li, 2012. Advances in Mapping from Remote Sensor Imagery: Techniques and Applications. Taylor and Francis, CRC Press, pp: 444, ISBN 978-1-43-987458-5.

Zemzami, O.A., H. Aksasse, M. Ouanan, B. Aksasse and A. Benkuider, 2013. Decomposition of 3D medical image based on fast and adaptive bidimensional empirical mode decomposition. Int. J. Comput. Netw. Commun. Secur., 7: 299-309.