Global Journal of Pure and Applied Mathematics.
ISSN 0973-1768 Volume 11, Number 3 (2015), pp. 1603-1626

http://www.ripublication.com

Dimensionality Reduction of Hyperspectral Image Using RBF-PCA and Mixed Pixel Wise HSMM Characterization Using SVM-FSK Forclassification

D.Regan[1], Dr.S.K.Srivatsa[2]
[1]ResearchScholar, St.Peter's University, Chennai, India
Email:reganocto@gmail.com
[2]Senior Professor, Prathyusha Institute of Technology & Management, Chennai, India

Abstract

The classification of hyperspectral remote sensing images is a tricky job, owing to numerous points of view, for instance, grouping of various pixels of present data; merely limited information is presented a priori, data dimension is substantial for conventional classification approaches; characteristically more than a few hundreds of spectral bands are obtained for every image. With the intention of solving all of these complications, in this paper formulated a novel classification method with a novel Radial Basis Function-Principal Component Analysis (RBF-PCA)-based prior model for dimensionality reduction of hyperspectral images. A prior prospect of the PCA is planned depending on the energy lying exterior to the span of principal components recognized in a higher-dimensional hyperspectral image feature space. This approach makes use of both mixed spatial and spectral information of the image which leads to more and enhanced class separability and consequently to improved classification accuracy. To advance the gradient level of spatial information, Improved Empirical Mode Decomposition (IEMD) with ABC (IEMD-ABC) is employed to increase the mixed pixel wise Support Vector Machine (SVM)-Fuzzy Sigmoid Kernel (FSK) classification accuracy. In EMD method the identifiable of Intrinsic Mode Functions (IMFs) of spectral band, weight values of IMFs are computed with the help of Artificial Bee Colony (ABC). The obtained both spectral and spatial information learning probability value from SVM -FSK are estimated by using HSMM. The proposed SVM-FSK is performed 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 value.

Keywords: Hyperspectral Image Classification, Dimensionality Reduction, Empirical Mode Decomposition (EMD), Support Vector Machine (SVM), Radial Basis Function (RBF), Fuzzy Sigmoid Kernel (FSK), Principal Component Analysis (PCA), Pixel Wise characterization, Hyper Spectral Images (HSI), Gradient Enhancement, Artificial Bee Colony (ABC), Hidden Semi Markov Model (HSMM)

Nomenclature

$\Re^{N}$	-	Hyperspectral features training samples
$\Re^{SSF}$	-	Hyperspectral spectral and spatial features samples
Φ	-	Nonlinear function
$k(fs, fs')$	-	RBF kernel feature space
$\lVert\Phi(fs)-\Phi(fs')\rVert_2^2$	-	Euclidean square distance between two different Hyperspectral feature spaces
σ	-	free parameter
$i=1,...M, j=1,...M$	-	Number of the features Hyperspectral image samples
A	-	Covariance matrix of the Radial Bias Kernel -principal component analysis (RBF-PCA)
λ_i	-	Eigen-values of the RBF-PCA
e [illegible]	-	normalized Eigen-vectors
FSN	-	Unobserved Feature space
V_i	-	Eigen vector
ψ_{FS_o}	-	Projection of FS on the principal subspace
FS_0	-	Feature space initial state
$P(FS)$	-	Gibb’s prior to derive the principal component analysis
ς	-	Sigma variant
FSO	-	Feature space reduced output from RBF-PCA
V_{FS}	-	Gradient of RBF-PCA
FSO_0	-	Final reduced Feature space results
$f(IMFSOW_a)$	-	Fitness value of the IMFs
$(IMFSOW_a)$	-	Intrinsic mode weight function
$A \& B$	-	A and B are the spatial dimensions
S	-	number of spectral bands
D	-	Dimension of IMF weight values for particles
p_a	-	Probability value p_a for $(IMFSOW_a)$
V_{ab}	-	IMFs weight value position
c And b	-	Randomly selected IMFs weight values
$\phi_{ab} \in [-1,1]$	-	Random vector
SN	-	Size of bees that is number of weight values in IMF

MNC	-	Maximum number of the cycles
$cycle = 1$	-	Termination condition
$RHIB$	-	The updated weight values of IMFs
R	-	Total number of IMFs used in the reconstruction,
q And r	-	Be the number of the different states in HSMM with classes
$SGFSII_i$	-	Spectral gradient reduced feature space input sample
$SGFSP$	-	Spectral gradient pixel
$SGFSOI$	-	Spectral gradient reduced feature space output
$SGFSOI_{ic} = \{+1, -1, middleclass\}$	-	Number of the classes for input sample
$svmb$	-	SVM Bias values for pixels
β	-	Regularization parameters
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 (fuzzytanh) kernel
l	-	labeled training samples
$SGFSOI_{qc} = \{+1, -1, mc\}$	-	Spectral gradient output image pixels classes
π_i	-	Initial probabilities for HSMM
$D_r(u)$	-	sojourn of the unobserved SVM-FSK objective function,
$SGFSS_t$	-	Semi-Markov chain spectral gradient feature space state at a time t
$L_c(sgfss_0^{\tau-1+u}, sgfsi_0^{\tau-1} \mid \theta)$	-	Complete-date likelihood results for SVM-FSK result
$L(\theta)$	-	Completed state sequence complicates the likelihood function
M^N	-	Possible sequences results
M	-	Hidden Semi Markov Model (HSMM) states
N	-	Number of input spectral information data

Introduction

The introduction of high spectral resolution airborne and satellite sensors enhances the capability for gathering of ground targets in fields as varied as farming, geology, geography or defense [1]. The technological development of optical sensors over the previous few decades has offered remote sensing analysts with wealthy spatial, spectral, and temporal data. Especially, the development in spectral resolution of hyperspectral images (HSIs) and infrared sounders unlocks the doors to new application fields and creates new methodological confronts in data investigation. HSIs permit the categorization of objects of attention with unprecedented accuracy, and maintain inventories the latest. HSI images comprise both spectral and spatial information data with high dimensional data space for characteristics and certain spatial information not straightforwardly noticeable by humans. These HSI methods and its applications have been concerned in extensive-ranging of research attempts [2].

The classification of hyperspectral remote sensing images is a demanding job, because of numerous complications, for instance, the mixed pixel wise categorization, dimensionality reduction, data dimension is substantial for conventional classification approaches; characteristically numerous hundreds of spectral bands are obtained for each image. These spectral bands can offer extremely rich spectral information of each pixel with the intention of recognizing the material of the objects. On the other hand, the spectral information alone occasionally does not permit the parting of structures. Consequently, contextual information, geometrical characteristics for instance, is essential for classification task of hyper spectral images.

The major complexity to compute the contextual information of hyper spectral images is the high dimension of the data. To diminish the data dimension, both supervised and non supervised approaches are formulated. The supervised approaches, for instance, band selection [3], decision boundary feature extraction and non-weighted feature extraction [3], convert the data in proportion to the training set with the aim of improving the separability of the data. On the other hand, the supervised approaches depend on the superiority of training set. The unsupervised approaches, for instance, PCA (Principle Component Analysis) or ICA (Independent Component Analysis), optimize certain statistical criterion (for example, the most uncorrelated elements or the mainly independent elements) to project the data onto a sub space with inferior dimension. The application of these approaches on hyperspectral data can be found in [4-5]. On the other hand, the elements obtained by optimizing the statistical condition do not essentially have physical meanings. Numerous numbers of the dimensionality reduction researches have been carried out in the previous works without any concern of the mixed pixel wise characterization classification construction.

Even though numerous schemes have been introduced in the area of hyperspectral image classification methods, such as, and Support Vector Machines (SVMs) [3], Bayesian schemes like Relevance Vector Machines [6] and Gaussian processes classification [7]. However, the SVM has certainly turned out to be the most extensively employed method in HSI classification study [8]. One of the foremost objectives in investigative studies in HSI image classification is to classify pixels into diverse classes by considering thousands and hundreds of hyperspectral cluster information and the human labeling is enormously exclusive. The foremost complication take place in HSI images classification is to manage little number of training samples with huge data dimensionality. The entire existing HSI image classification approaches majorly concentrate on either feature selection or dimensionality reduction. With the intention of solving all of these complications and lessen the dimensionality of the features space with varied pixel wise spectral and spatial data pixel, in this research presented a novel classification framework.

This paper concentrates on the challenging setback of hyperspectral mixed pixelwise characterization framework together with dimensionality reduction complication in classification, which has lately achieved in recognition and drawn the attention of other scientific disciplines such as machine learning, image processing and computer vision [9]. In case of remote sensing research, the term classification is employed to indicate the process that allocates mixed pixels to a collection of classes and then allocated to a class. Introducing novel RBF-PCA based dimensionality reduction field for hyperspectral imaging for hyperspectral images with the intention of increasing the classification accuracy with a reduced amount of volume of data. Following the dimensional are reduced to amplify the spectral gradient value, this paper employs Empirical Mode Decomposition (EMD) which is formulated by Amir Eftekhar et al [10] to examine nonlinear and nonstationary data. EMD decomposes a signal into a fixed amount of Intrinsic Mode Functions (IMFs) and a residue.

In this paper, IEMD is implemented to each spectral band to acquire a fixed number of IMFs. Subsequently, ABC, with a spectral gradient objective function, is employed to decide the optimum weights of the obtained IMFs for rebuilding. The IMFs are summed by means of these weights to rebuild the characteristics that are employed in classification with SVM-FSK. Consequently, the proposed approach integrates spatial and spectral processing. Mixed pixel wise characterization approach based on a probabilistic HSMM framework for SVM-FSK. It is employed as a pointer to locally discover the overall number of mixed components that play a part in each pixel. It is also revealed that the proposed approach offers a significant increase in accuracy for hyperspectral image classification.

Related Work

In recent past, several approaches and schemes have been formulated to achieve enhanced hyperspectral image categorization accuracy. Numerous amounts of the dimensionality reduction approaches also carried out to lessen the dimensionality of the feature space. Hyperspectral data dimensionality reduction and end member extraction has been carried out by Muhammad et al [11]. This author formulated a novel approach to overcome the computational complexity of hyperspectral (HS) image data to discover multiple targets/end members precisely and competently by reducing time and complexity. With the intention of overcoming the computational complexity, standard deviation and chi square distance metric approaches are taken into account. The quantity of end members is approximated by unbiased iterative correlation scheme. In order to provide an algorithm for surpassing the computational complexities of hyperspectral data to discover the multiple targets and end members efficiently with fewer computational times was the main objective. The end member judgment was carried out by unbiased iterative correlation technique.

Bruce et al [12] dyadic discrete wavelet transform is developed for feature extraction from a high-dimensional data space. The wavelet's inbuilt multi resolutional characteristics are discussed based on multispectral and hyperspectral remote sensing. In addition, several wavelet-based features are employed to the complication of automatic classification of precise ground vegetations from hyperspectral signatures. The wavelet transform features are assessed by means of an automated statistical classifier. The classification is examined using hyperspectral data for several agricultural applications.

Burgers et al [13] have done a comparative investigation of dimensionality reduction approaches aiming to assess the performance of the dimensionality reduction schemes. Eight dissimilar approaches namely, Principal Component Analysis, Laplacian Eigen maps, Isomap, Kernel Principal Component Analysis, Independent Component Analysis, LMVU, Diffusion maps and LTSA have been assessed for their performance on the dimensionality reduction and determination of the inherent dimensionality of the hyperspectral images. Nonlinear approaches had provided comparably enhanced results however had a foremost complication of taking extremely long runtimes. Several hyperspectral data sets were employed in the experiment and the performance evaluation has been carried out by comparing both the classification accuracy and runtime of the algorithm. In this progression, PCA was detected to be best in running and provide the most precise results. Subsequent to investigation has been carried out, of all the odds PCA had surpassed and has been found as the most excellent dimensionality reduction approach providing the best results. Target detection was contentedly carried out by PCA, KPCA and ICA. PCA had the slightest error rates in the processes and surpassed in all the tasks evaluated to other approaches.

Investigation of hyper dimensional feature spaces, to assess their efficiency in categorizes complex forest regions in a multisource framework. The initial technique is a parametric regularized Gaussian Maximum Likelihood (GML) [14] classifier. Especially, in case of the number of training samples is lesser than the quantity of features, the covariance matrix employed in the decision rule become singular, and as a result, the GML cannot be exploited. With the intention of solving this complication that employs the Mixed- Leave One- Out-Covariance (LOOC) [15] to avoiding extreme covariance estimator error. Mixed-LOOC2 has improvements over LOOC and BLOOC and requires less computation than those two. Depending on Mixed-LOOC2, new DAFE and mixture classifier approaches are formulated. Current feature extraction algorithms, at the same time effective in certain circumstances, have noteworthy limitations. Discriminate Analysis Feature Extraction (DAFE) is quick however does not execute well with classes whose mean values are alike, and it creates only N-1 consistent features where N represents the number of classes.

The second approach is a distribution-free machine learning classifier depending on the SVM [16-17]. SVM have been extensively employed in previous investigations on classification of hyperspectral data (e.g., [16-17]), establishing their efficiency in hyper dimensional feature spaces; 2) both approaches are basically capable of solving ill-posed classification complications, where the ratio among the amount of available training samples and the amount of features is comparatively small (this is a characteristic condition with hyperspectral data). The SVM classifier is a distribution free complex classifier, which depends on machine learning and, therefore, on an entirely different theoretical background concerning GML-LOOC. SVM established to be extremely efficient for classification of hyperspectral data (e.g., [16] and [17]).

An assessment of SVM in remote sensing was provided by Mountrakis et al [18]. An investigation was made on SVM and its significance in remote sensing. SVM has an ability to simplify the data with extremely less training samples and provide enhanced accuracies compared against other training techniques. Since SVM is non-parametric, for classification it does not presume a statistical distribution. It requires and constantly bonds to global minima since it takes care of quadratic setbacks. At the same time, the remote sensing data have unidentified distributions, this property of SVM is extremely much helpful permitting to outperform than the other type for classification approaches.

SVM has been employed as a tool for mapping mineral potentially by Renzuang [19]. The research confirmed that SVM is the most excellent geo-computational tool for spatial investigation. SVM was subjected to many variables for mineral prospective mapping. SVM approach with dissimilar kernel functions was assessed with the mineral region. The results obtained were satisfactory and pointed out that it is a helpful tool for integrating several evidence layers in mineral wealth mapping. These results promoted the practice of SVM since the study area is occupied with several mineral mines.

A Recursive Support Vector Machine (RSVM) for dimensionality reduction was formulated by Tao et al [20]. A multidimensional maximum margin feature extraction scheme was discussed broadly which is employed for building an orthogonal dependent dimensionality reduction. The analysis demonstrates that as the number of recursive constituents raises, the objective function of the SVM decreases. RSVM confirms improved accuracy than the standard SVM and Linear Discriminant Analysis (LDA) and have no singularity complications. The examination was performed on standard benchmark non-spatial data sets. The major idea of taking this literature is to analyze the same on the spatial high dimensional hyperspectral dataset.

Proposed Methodology

In this work, 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 hyperspectral image cube without losing any significant information and as a result circumvent 'the curse of dimensionality'. In this paper, at first carry out dimensionality reduction approaches with the help of Radial Basis Function (RBF)-Principal Component Analysis (RBF-PCA) scheme. This method maps the entire hyperspectral data from the original feature space and plots the efficient features accompanied by helpful information to a lower-dimensional subspace.

RBF-PCA is a nonlinear extension of conventional PCA for obtaining higher-order correlations in a hyperspectral data. Subsequently, Improved Empirical Mode Decomposition (IEMD) method is employed to improve gradient level of spatial data to separately specialize of Intrinsic Mode Functions (IMFs) of every one of band. In EMD approach, IMFs weight values are decided by means of Artificial Bee Colony (ABC) optimization algorithm. Quantity of IMFs is exploited as feature data vector for Support Vector Machine–Fuzzy Sigmoid Kernel (SVM-FSK) classification framework. Subsequently obtained results are implemented to mixed pixel wise probabilistic classification framework, in which SVM-FSK is employed as classifier. The fuzzy sigmoid function is exploited as a kernel function in an SVM classification framework to categorize mixed pixels HSI images. The objective function results of the FSK-SVM are approximated depending on the Hidden Semi Markov Model (HSMM). Proposed work representation of the complete system is given in Fig.1.

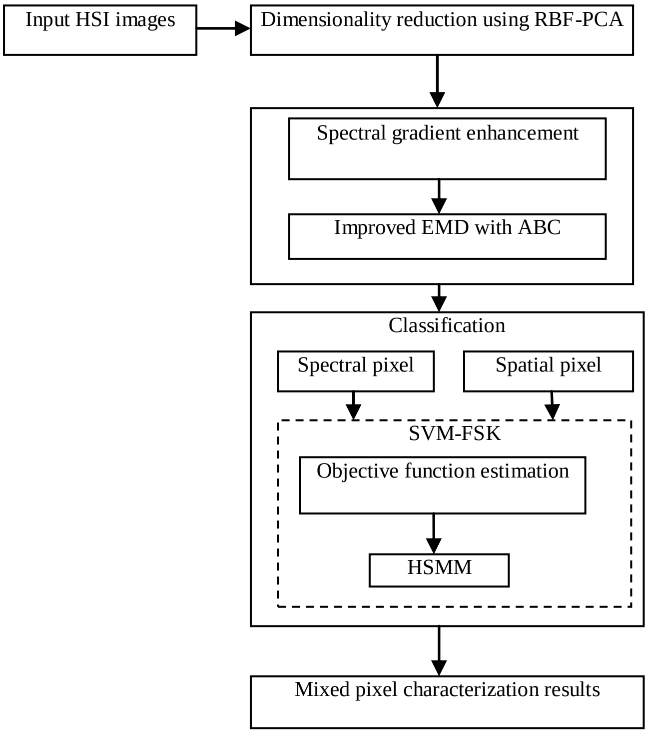

Figure 1: Entire Architecture of Proposed Work

Dimensionality Reduction Using Radial Basis Function-Principal Component Analysis (RBF-PCA)

PCA is one of the most extensively utilized dimensionality reduction approaches. It is a commonly accepted approach for eradicating redundancy in the data which depends on decorrelation. Even though it accomplishes redundancy reduction by removal of low variance constituents, this rotational transform is time-consuming as a result of its global nature. Furthermore, since it is a global transformation, it does not maintain local spectral signatures and consequently might not maintain all information required to obtain a better classification [21]. Principal component analysis is described as a transform of a given set of n input feature samples with the same size formed in the n-dimensional vector $x=(x_1,...x_n)$ into a vector y according to ,

$$y = A(x - m_x) \tag{1}$$

This point of view enables to form a simple formula (1) ,each row of the feature vector x of hyperspectral image contains of K values belonging to one input. The vector m_x is the vector of mean values for hyperspectral image samples defined by relation,

$$m_x = E(x) = \frac{1}{k}\sum_{k=1}^{K} x_k \tag{2}$$

Matrix A in Equation (1) is determined by the covariance matrix C_X. Rows in the A matrix are formed from the eigenvectors e of C_X ordered according to corresponding eigenvalues in descending order. The evaluation of the C_X matrix is possible according to relation

$$C_X = E\{(x-m_x)(x-m_x)^T\} = \frac{1}{K}\sum_{k=1}^{K} x_k x_k^T - m_x m_x^T \tag{3}$$

As the vector x of input variables is n-dimensional it is obvious that the size of C_X is n x n. The elements $C_X(i,i)$ lying in its main diagonal are the variances

$$C_X(i,i) = E(x_i - m_i)^2 \tag{4}$$

of x and the other values $C_X(i,i)$ determine the covariance between input variables x_i, x_j.

$$C_X(i,j) = E\{(x_i - m_i)(x_j - m_j)\} \tag{5}$$

between input feature vectors x_i, x_j, . The rows of A in Equation (1) are orthonormal so the inversion ofPCA is possible according to relation

$$X = A^T y + m_k \tag{6}$$

Extend PCA to consider higher order correlations among the spectral and spatial hyperspectral data pixels for accomplishing improved super-resolution potential. Depending on the schemes from [22], characterize a nonlinear function Φ on the hyperspectral image data sample with feature space $\Re^N$ as $\Phi : \Re^N \rightarrow \Re^{SSF}$ in which typically characterize the hyperspectral spectral and spatial features as $SSF >> NF$, $\Re^{SSF}$ is called as the spectral and spatial feature space. The function Φ is typically selected in order that $\Re^{SSF}$ includes higher-order product terms from number of hyperspectral features training samples $\Re^N$. Employing PCA on $\Re^{SSF}$ will then acquire higher-order correlations from number of hyperspectral features training samples $\Re^N$. On the other hand, openly computing the nonlinear map Φ is computationally expensive, particularly when the space is high-dimensional. With the intention of surpassing these complications in PCA, this paper employed RBF –PCA approach for each number of the hyperspectral data with non linear function. Characterize a kernel function k in order that,

$$K(fs, fs') = \exp\left(\frac{-\|\Phi(fs) - \Phi(fs')\|_2^2}{2\sigma^2}\right) \tag{7}$$

Provided hyperspectral image samples feature vectors in the dimensional space; the RBF kernel function computes their $-\|\Phi(fs) - \Phi(fs')\|_2^2$ Euclidean square distance among two different hyperspectral feature space vectors in the feature space. σ represents a free parameter, select k suitably to stimulate a desired Φ. In the last step, required to carry out the PCA to each hyperspectral feature space, fine-tune the $\Re^{SSF}$ parameters. For that function, required to transform the Φ and k to obtain $\tilde{\Phi}, \tilde{k}$ values to each hyperspectral feature space in the following approach,

$$\tilde{\Phi}(fs) = \Phi(fs) - \frac{1}{M}\sum_{i=1}^{M}\Phi(Tfs_i) \tag{8}$$

$$\tilde{k}(fs, fs') = \tilde{\Phi}(fs, fs') = k(fs, fs') - \frac{1}{M}\sum_{i=1}^{M}k(fs, TFS_i) - \frac{1}{M}\sum_{i=1}^{M}k(TFS_i, fs') + \frac{1}{M^2}\sum_{i=1}^{M}\sum_{j=1}^{M}k(TFS_i, TFS_j) \tag{9}$$

In order to work out the covariance matrix for each hyperspectral feature data space fs is computed as AA^T where

$$A = \left(\frac{1}{\sqrt{M}}\right)\{\tilde{\Phi}(TFS_1), \ldots .\tilde{\Phi}(TFS_m)\} \tag{10}$$

$$\lambda_i = (AA^T)_{ij} = \frac{1}{M}\tilde{\Phi}(TFS_i)^T\tilde{\Phi}(TFS_j) = \frac{1}{M}\tilde{k}(TFS_i, TFS_j) \tag{11}$$

Subsequently, fragment the covariance matrix results into Eigen-decompose $A^T A$ to obtain the Eigen-values $\{\lambda_i\}$ and normalized Eigen-vectors $\{e_i\}$. Depending on the values of $\{\lambda_i\}$ select the initial m eigen-vectors as the principal components. For particular new hyperspectral feature sample dataspace FSN, compute the projection of $\tilde{\Phi}(FSN)$ on Eigen-vector V_i as

$$\tilde{\Phi}(FSN)^T V_i = \tilde{\Phi}(FSN)^T Ae_i = \sum_{j=1}^{M} e_{i_j}\tilde{k}(FSN, TFS_j) \tag{12}$$

Characterize the projection of FS on the principal subspace as given below,

$$\psi_{FS_0} = \sum_{i=1}^{M}\tilde{\Phi}(FS_0)^T, V_i)V_i \tag{13}$$

Execute a Gibb's earlier than deriving the PCA term to each feature dimensional space term as given below,

$$P(FS)=\left(\frac{1}{C}\right)\exp\{\frac{-\|\psi_{FS_0}-\tilde{\Phi}(FS)\|^2}{\varsigma}\} \tag{14}$$

$$\hat{FS}=\arg\min_{Fs}\{\|FSO-HFS\|^2 + \Gamma_{FSN}\|\Psi_{FS_0}-\tilde{\Phi}(FS)\|^2\} \tag{15}$$

$\tilde{\Phi}(FS_0)^T V_i$ is computed in equation (12). From (15) the expression for the gradient of $\|\Psi_{FS_0}-\tilde{\Phi}(FS)\|^2$ with regard to FS can be computed as,

$$\nabla_{FS}\|\Psi_{FS_0}-\tilde{\Phi}(FS)\|^2=\nabla_{FS}\tilde{k}(FS,FS_0) - 2\sum_{i=1}^{m}(\tilde{\Phi}(FS_0)^T V_i\sum_{j=1}^{M}e_{i_j}\nabla_{FS}\tilde{k}(FS,TFS_j)) \tag{16}$$

$$\nabla_{FS}\tilde{k}(FS,TFS_j)=2n(1+Fs^T FS)^{n-1}FS - -\frac{2}{m}\sum_{i=1}^{n}n(1+Fs^T TFS_i)^{n-1}TFS_i \tag{17}$$

$$\nabla_{FS}\tilde{k}(FS,TFS_j)=n(1+Fs^T TFS_j)^{n-1}TFS_j - -\frac{1}{m}\sum_{i=1}^{n}n(1+Fs^T TFS_i)^{n-1}TFS_i \tag{18}$$

$P(FS)$ increases as FS meets results to Ψ_{FS} and get maximum feature space for hyperspectral data for preliminary feature input space FS_0

$$FS_0=\arg\min_{Fs}\|FSO-HFS\|^2+\gamma\|FS\|^2 \tag{19}$$

Improved EMD for Spectral Gradient

Subsequent to feature dimensional space results from the RBF-PCA approach, carries 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 paper, 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 $IMFSOW_i$ are computed using Artificial Bee Colony (ABC) algorithm, 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.

$$f(IMFSOW_a) = specgrad = \frac{1}{A \times B \times S} \sum_{m=1}^{A} \sum_{n=1}^{B} \sum_{b=1}^{S} \left| \frac{\partial FSO(m,n,b)}{\partial b} \right| \quad (20)$$

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 some of the ABC with bee initial population. IMFs weight values are computed through the searching manner of the honey with dancing manner with arbitrarily selection of the $IMFSOW_i$ weight value and updated by considering previous maximum probability values P_i to each IMFs weight values up to now found by honey bee. The weight values of IMF are revised in accordance with food searching manner of bee in every searching space dimensional space D with some absolute spectral gradient values in every spectral band of the hyperspectral data. 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 one of the bee to compute the weight values of hyperspectral reduced feature dimensional space IMF is given as

$IMFSOW_a = (imfsow_{a1}, ...imfsow_D)$ in which D represents the dimension of IMFs weight values for investigation search space to hyperspectral image sample.

ABC imitates the foraging manner of honey bees. When bee looks for best $IMFSOW_a$ they typically deposit a unique dancing manner of the bees. ABC has been productively employed for numerous optimization complications [23] since it is simple to build up and solve several optimization complications with only a few controls of parameters [24]. In ABC optimization, employed $IMFSOW$ bees visit the food source location. In ABC each one of the employee bee collects information regarding weight values of the diminished feature dimensional space results from RBF-PCA. Employed bees carry out the local examination to discover the best $IMFSOW$ and attempt to use the nearest best $IMFSOW$ neighboring locations results for each one of reduced feature dimensional space. The bees waiting in the nest region to discover most excellent $IMFSOW$ for hyperspectral feature space are regarded as onlooker bees. Onlooker bees carry out the global examination to discover the most excellent $IMFSOW$ and revise global optimum $IMFSOW$ results in position update phase.

Scout bees determine arbitrarily select new IMFSOW weight value of the diminished hyperspectral feature dimensional space which is not looked by the employed bees, these three phases are continued until a highest number of the iterations termination criterion is met. The fitness value $f(IMFSOW_i)$ is computed depending on the equation (20). An artificial onlooker bee chooses the $IMFSOW_i$ in accordance with the calculation of the probability value p_a by using expression given below,

$$p_a = \frac{f(IMFSOW_a)}{\sum_{w=1}^{SN} f(IMFSOW_a)} \quad (21)$$

Where $f(IMFSOW_i)$ indicates the fitness value IMF weight to every employee i in the location and SN represents the size of the population. The chosen IMF weight position is revised by using the following equation (22)

$$v_{ab} = IMFSOW_{ab} + \theta_{ab}(IMFSO_{ab} - \chi_{cb}) \quad (22)$$

Where c and b are arbitrarily chosen IMFs weight values $c \in \{1,....SN\}$ & $b \in \{1,....D\}$, $\phi_{ab} \in [-1,1]$ based on this result, the parameter value of $IMFSO_{ab}$ surpasses its threshold value, the result of the bees for IMFs weight value is satisfactory otherwise it is not satisfactory, it is also substituted by the scouts bees, In case of ABC, when a bee present IMFs weight value position does not enhance the result within a pre-specified number of iterations, subsequently the current IMFs weight values to be assumed as abandoned and it is updates as given below,

$$IMFSOW_a^b = IMFSOW_{\min} + rand(0,1) \\ (IMFSO_{\max}^b - IMFSO_{\min}^b) \quad (23)$$

All the above discussed steps majorly based on following parameters which limits the operation SN, maximum number of the cycles (MCN).

Algorithm 1: Weight estimation using ABC

Input: SN size of bees that is number of weight values in IMF
D dimensions of search space for weight values
Output: Update weight values for hyperspectral images
Method

1. Initialize bee's weight values in IMF with arbitrary positions and velocities $VIMFW_i$ in the search space D, and then select the most excellent IMF weight values
2. Fix $cycle = 1$
3. Repeat
4. For the entire weight values for IMF in $IMFW_i$ in bee do
5. Estimate the objective function of each IMF weight values by means of fitness function from $f(IMFSOW_a)$ (20)
6. Construct new IMF weight value calculation for each bee position v_{ab} with the help of (22)
7. Execute the greedy selection process for the employed bees
8. Compute the probability values p_a for chosen $IMFSOW_a$ with the help of (21)
9. Construct current weight values position solutions v_a for the onlookers bees depending on P_i and assess them
10. Implement the greedy selection method for the onlookers

11. Decide the discarded tasks results in the scout, when exists, and substitute it with a new arbitrarily produced solution $IMFSOW_a^b$ by (23)

12. Remember the best solution accomplished to this point

13. $cycle = cycle + 1$

14. Until $cycle = MCN$ is satisfied (typically an adequately better IMF weight value fitness)

15. Return the best weight values for IMF and its fitness value $IMFSOW_i$.

16. The updated weight values of IMFs by means of their corresponding weights to acquire the new hyperspectral data representation that will be employed for classification

$$RHIB = \sum_{d=1}^{R} IMFSOw_a \times IMFSO_a \quad (24)$$

Where $IMFSOw_a$ shows the equivalent weight of the IMF, R is the overall number of IMFs employed in the reconstruction, and $RHIB$ indicates the reconstructed hyperspectral image band.

17. Execute classification using following 3.3 and 3.4 sections.

Spectral and Spatial Characterization

In Soon after 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 HSMM. In order to learn mixed pixel-wise SVM-FSK based classification results, approximate the HSMM 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 competently 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 [25]. In this paper, let $SGFSII_i = \{SGFSII_1, ..., SGFSII_n\}$ represent the spectral gradient outcome from IEMD based ABC approach with diminished feature dimensional space. Also let $SGFSII_i = \{SGFSII_1, ..., SGFSII_{id}\}^T$ which represents the spectral gradient outcomes that related with reduced feature dimensional images pixel $SGFSII_i \in SGFSP$. The $SGFSP$ is characterized as a spectral gradient pixel $SGFSP \in \{1, ...n\}$ with indexing of n

pixels of $SGFSII_i$ & $SGFSB$ in which SGFSB indicates the amount of spectral gradient feature space bands $SGFSOI_i = \{SGFSOI_1,..., SGFSOI_{ik}\}^T$ represent the classification results of SVM-FSK, in which C indicates the quantity of classes $SGFSOI_{ic} = \{+1,-1, middleclass\}$ for $c = 1,...C$ and $\sum_c SGFSOI_{ic} = 1$, $\phi(.)$ is indicated as non linear mapping function of gradient function, it is carried out in accordance with the Cover's theorem [26], which promises elevated classification accuracy rate for linearly separated feature vector samples and it is commonly higher dimensional feature space fs.

$$\min_{W,\xi,svmb} \left\{ \frac{1}{2} \| w \|^2 + \beta \sum_i \xi_i \right\} \tag{25}$$

Constrained to the,

$$SGFSOI_i(\phi^T(j)w + svmb) \geq 1 - \xi_q, \forall i = 1,..n \tag{26}$$

$$\xi_i \geq 0, \forall i = 1,..n \tag{27}$$

Where W & $svmb$ represents a linear classifier for spectral gradient hyperspectral images. 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 HSMM is given below,

$$SGFSOI_{ic} = \begin{cases} 1 if p(SGFSOI_{ic} = 1 \mid SGFSTR_i) > p(SGFSOI_{ic_t} = 1 \mid SGFSTR_i) \\ 0 otherwise \end{cases} \tag{27}$$

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

$$K(SGFSOI_i, SGFSOI_j) = \phi(SGFSOI_i)\phi(SGFSOI_j) \tag{28}$$

This kernel function result not enhances classification accuracy rate for certain data, to overcome these complication, in this paper kernel function are approximated depending on fuzzy sigmiod function is defined in equation (28) given below,

$$f(SGFSOI) = \text{sgn} \sum_{i,j=1}^{n} SGFSOI_i SGFSOI_j \alpha_i \alpha_j$$
$$K(SGFSOI_i, SGFSOI_j) + svmb \tag{29}$$

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 paper extends the fundamental ideas of hyberbolic tangent function from [27] and it is given as follows (30)

$$K(SGFSII_i, SGFSII_j) = \begin{cases} -1 SGFSII_i.SGFSII_j islow \\ +1 SGFSII_i.SGFSII_j ishigh \\ m.SGFSOI_i.SGFSOI_j. \\ SGFSII_i.SGFSII_j ismedium \end{cases} \quad (30)$$

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 paper only concentrated on three triangular function, owing to their straightforwardness as illustrated in Fig.2. The fuzzy sigmoid kernel function be continuous, the membership limits are specified by $\gamma - \frac{1}{a} to \gamma + \frac{1}{a}$

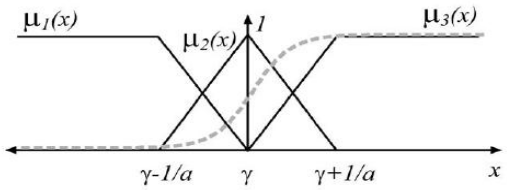

Figure 2: Schematic of The Three Membership Functions

Subsequent to fuzzy sigmoid kernel function be continuous, as a result the expression (30) can be readily re-written as a function of a and γ as follows:

$$K(SGFSII_i, SGFSII_j) = \begin{cases} -1 SGFSII_i.SGFSII_j \le \gamma - \frac{1}{a} \\ +1 SGFSII_i.SGFSII_j \ge \gamma - \frac{1}{a} \\ 2.(SGFSOI_i.SGFSOI_j - \gamma) - a^2 (SGFSOI_i.SGFSOI_j - \gamma) \\ |(SGFSII_i.SGFSII_j - \gamma)| \end{cases} \quad (31)$$

which is the absolute form of the proposed fuzzy sigmoid (fuzzy tanh) kernel. The major benefits of this function are (1) it carries out quicker than, since the final results of classification purpose are conveyed in a series of saturated samples (Eq. (30)), (2) it allows to select diverse levels of non-linearity by choosing the quantity and authority of the membership functions. In addition, measure the outcome of fuzzy sigmoid kernel function through objective function for l labeled training samples that is $TS_l = (SGFSII_1, SGFSOI_1)...(SGFSII_l, SGFSOI_L)$ Based on the above discussed steps is referred from [28], the major function is to approximate the probabilistic value for mixed pixels $SGII_i$ with class label vector $SGFSOI_i$. This vector results can be obtained from Hidden Semi Markov Model (HSMM) through the computing probability function.

$$SGFSOI_{ic}, = \begin{cases} -1\, p(c_{SGFSII} = k \mid SGFSOI, \theta) > \\ p(ct_{SGFSII} = k \mid SGFSOI, \theta) \\ 0 otherwise \end{cases} \quad (32)$$

Probability Estimation Using HSMM

A Hidden Semi Markov Models with markov chain is initiated in [29], it includes quantity of classes $SGFSOI_{ic} = \{+1, -1, middleclass\}$ for $c = 1,..C$ with several number of spectral gradient reduced feature space input sample states $SGFSII_q = \{SGFSII_1, ..., SGFSII_n\}$. The observed process $SGFSII_t$ is connected to the hidden, is described for each one of the states $SGFSII_t$ with classes.

Initial probabilities

$$\prod_i := P(SGFSS_0 = j), \sum_i \prod_i = 1 \quad (33)$$

The transition probabilities for the state r. For each $r \neq q$

$$P_{ij} = P(SGFSS_{t+1} = j \mid SGFSS_{t+1} \neq i, SGFSS_t = i) \quad (34)$$

$$\sum_{j \neq i} p_{ij} = 1 \,\&\, p_{ij} = 0 \quad (35)$$

The occupancy of definite classes $d_j(u)$ together with the entity has to be allocated to each of the states by,

$$\begin{aligned} d_j(u) &= P(SGFSS_{t+u+1} \neq j \mid SGFSS_{t+u-v} = j, \\ v &= 0,.., u-2 \mid SGFS_{t+1} = j, SGFS_t \neq j \end{aligned} \quad (35)$$

for $u \in \{1,...M_r\}$. Sojourn time is described as occupancy of each and every state in SHMM to finish the estimation of the SVM-FSK objective function development with several number of the classes $SGFSOI_{qc} = \{+1, -1, middleclass\}$. The sojourn of the unobserved SVM-FSK objective function result is indicated as $D_r(u)$ .It is between the boundaries $t+1 tot + u$ in the state r. Before and following the sojourn time, each $SGFSOI_{rc}$ state r with several classes is indicated as

$$D_j(u) := \sum_{v > u} d_j(u) \quad (36)$$

When the $SGFSOI_{qc}$ estimation results for hyper spectral reduced feature space results in state r at time $t = 0$, the subsequent relation can be confirmed,

$$P(SGFSS_t \neq j \mid SGFSS_{t-v} = j, v = 1,..t) = d_j(t)\Pi_j) \quad (37)$$

The last observed estimation results for mixed pixel wise characterization of $SGFSII_t$ is associated to the Semi-Markov chain $SGFSS_t$ by the observation probabilities,

$$b_j(SGFSII_t) = P(SGFSII_t = x_t \mid SGFSS_t = j) \,\&\, \sum_{x_t} b_j(SGFSII_t) = 1 \quad (38)$$

The observation of the particular estimated SVM-FSK objective function is exemplified by the conditional independence property,

$$\begin{aligned} &P(SGFSII_t = sgfsii_t \mid SGFSII_0^{\tau-1} = sgfsii_0^{\tau-1}) \\ &= P(SGFSII_t = sgfsii_t \mid SGFSS_t = sgfss_t) \end{aligned} \quad (39)$$

Complete-date likelihood,

$$L_C(sgfss_0^{\tau-1+u}, sgfsii_0^{\tau-1} \mid \theta) = P(SGFSS_0^{\tau-1} = sgfss_0^{\tau-1}, SGFSS_{\tau-1+v} = sgfss_{\tau-1,v} \quad (40)$$
$$= 1....u-1. SGFSS_{\tau-1+v} \neq sgfss_{\tau-1}, SGFSII_0^{\tau-1} = sgfsii_0^{\tau-1}(\theta)$$

The completed state sequence cause difficulties to the likelihood function by a supplementary sum over all probable prolongations of the overload detection sequence $s_0,..., s_{\tau-1}$.It is specified as,

$$L(\theta) = \sum_{s_0,...,s_{\tau-1}} \sum_{u_{\tau+}} L_c(sgfss_0^{\tau-1}, sgfsii_0^{\tau-1} \mid \theta) \quad (41)$$

The major inspiration for the Hidden Semi Markov Model (HSMM) work is to build up a proficient probabilistic method to compute the probability values of $\mathrm{SGFSOI_i}$. Consider there is M -state Markov chain of length N happens to be the number of input spectral information data from spectral gradient IEMD-ABC, and there are M^N probable sequences results are employed to acquire complete $SGFSOI_{ic}$ spectral gradient output image class. On the other hand, it is obvious that, when the state space of HSMM is huge, numerous of other sequences might be also interested to carry out mixed pixel wise probabilistic estimation.

Experimental Results

In this paper, for the purpose of assessing the results of proposed approach and existing work, experimentation results are done with a single data set specifically, the Indian Pine data in Table .1that belongs to hyperspectral data. It is one of the primarily significant dataset generally used at present. Since certain spectral data classes are extremely comparable and some pixels in the images include combined pixel data values. The Indian Pine data set, which was formulated in 1992, it includes pixel range of image that corresponds to 145 by 145 pixels and totally there are 220 spectral bands. However in the spectral band data there are certain noises exist and water absorption also exist, in order to surpass these complications, noises and water absorption are eradicated, the absolute 220 spectral bands are reduced to 200 spectral bands data is 4–2.5 μm, and the spatial resolution is 20 m. Although the original hyperspectral data from Indian pine dataset includes 16 classes. From the complete training samples when certain classes include only little data samples those classes are generally not considered, and the remaining classes training samples is taken as input to training process.

Table 1: Indian Pine Data

CLASS	NUMBER 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

Table.2shows the results of overall accuracy for the original data representation as well as the EMD-GA, EMD with spectral enhancement (denoted as IEMD-PSO), and the dimensionality reduced features RBF-PCA method for IEMD-ABC, it can be found that the overall accuracy results of the proposed RBF-PCA IEMD-ABC have higher accuracy than the existing methods since it reduces the dimensionality of the features by using the RBF-PCA methods.

Table 2: Accuracy In Indian Pine Data For Spectral Gradient

Images	**Accuracy (%)**			
	EMD	**EMD-GA**	**IEMD-PSO**	**RBF-PCA- IEMD-ABC**
1	91.5	92.23	95.42	96.8
2	90.5	93.84	95.51	97.15
3	90.25	93.95	95.64	97.45
4	90.8	94.12	95.87	97.89
5	90.5	94.25	96.12	97.98

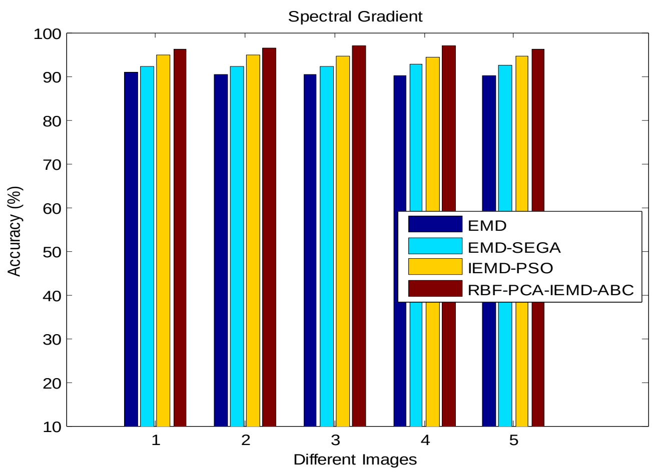

Figure 3: Accuracy In Indian Pine Data For Spectral Gradient

Fig.3 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

ABC with reduced features results from the RBF-PCA methods with the existing methods EMD, EMD-GA, it can be compared with existing EMD-GA. It shows that proposed RBF-PCA-IEMD-ABC achieves higher accuracy than earlier methods.

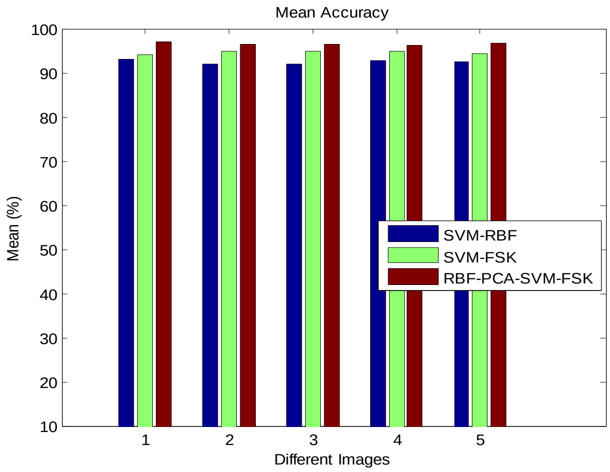

Figure 4: Mean Value Results For Classification Methods

Fig.4 shows the evaluation of classification methods performance in terms of mean function. Mean values evaluation is carried out among classification approaches such as, SVM with RBF kernel and SVM with Fuzzy sigmoid function, feature reduced dimensionality reduction method (RBF-PCA) with SVM-FSK. In case when the results of mean values are higher for particular classification approach, similarly the classification accuracy of those approaches is also higher. When comparing against SVM-RBF method, SVM-FSK methods the mean value of the proposed RBF-PCA-SVM-FSK accomplishes higher mean values as shown in Fig.5 and values are tabulated in Table.3.

Table 3: Mean And Standard Deviation In Indian Pine Data Classification Results

Images	**SVM-RBF**	**SVM-FSK**	**RBF-PCA-SVM-FSK**	**SVM-RBF**	**SVM-FSK**	**RBF-PCA-SVM-FSK**
	Mean (μ)			**Standard deviation (σ)**		
1	92.56	93.24	96.95	0.97	0.78	0.608
2	92.13	94.41	96.89	0.964	0.77	0.61
3	93.56	94.64	96.63	0.958	0.71	0.617
4	93.39	95. 22	97.23	0.95	0.72	0.604
5	91.9	95.8	97.58	0.978	0.74	0.608

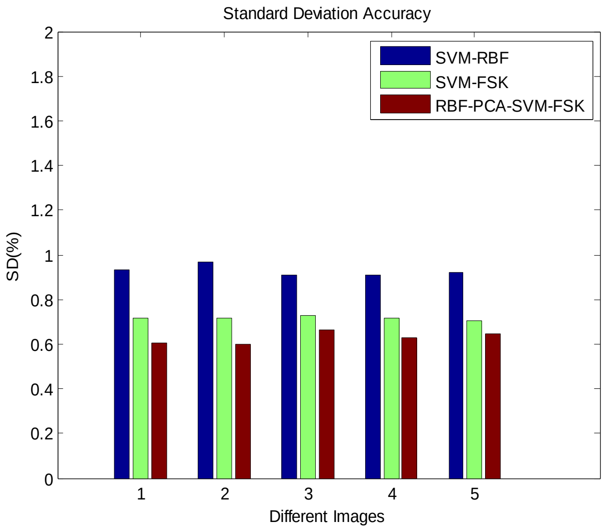

Figure 5: Standard deviation for classification results

Fig.5 shows the results of standard deviation estimation between classification methods such as SVM-RBF, SVM–FSK and proposed RBF-PCA-SVM-FSK method. In case when the results of standard deviation are lesser, it shows better classification accuracy results. It shows that proposed RBF-PCA-SVM-FSK (Radial Basis Function- Principal Component Analysis -Support vector machine-Fuzzy Sigmoid Kernel) have less standard deviation than SVM-RBF, SVM-FSK.

Conclusion and Future Work

In this paper, a novel RBF-PCA based dimensionality reduction method for hyperspectral data is formulated for IEMD-ABC. This model is employed to standardize the reconstruction of high values of the feature vector pixels for each hyperspectral image data samples and both spatial and spectral domain features are employed for generating helpful features. By nonlinearly mapping the hyperspectral images to a higher-dimensional feature space and carrying out PCA on the feature space and obtain higher-order correlations exist in hyperspectral image data. Improved EMD together with ABC (IEMD-ABC) for the purpose of spectral gradient enhancement is used to enhance the classification accuracy. In IMED approach, the hyperspectral bands information are transformed into IMFs and weight values of IMFs are acquired from ABC approach to optimize the spectral gradient. Here proposed a mixed pixel-wise characterization (SVM-FSK) method for classification of spectral-spatial information. SVM-FSK probability values are computed with the help of the probabilistic HSMM method; it discovers the misclassification results and a collection of previous resultant classes, in the same way. In addition, classification is typically enhanced by HSMM. It also permits the capability to design sparse techniques that are capable of working in applicable feature subspaces, where compact and computationally well-organized methods can be run. At last, classifiers can now integrate mixed spatial and spectral data that separate uncertainties exist in land-cover classification. The proposed RBF-PCA approach is very effective in noisy images when less training data and more features are employed. This method provides considerable performance enhancement in terms of classification accuracy when less features are utilized compared to other frequency based dimensionality reduction methods.

Extend the method proposed to hyperspectral images to estimate the characteristic scales of the structures. In future developments, will further explore the relationship between the parameters of our method and the spatial resolution, level of noise, and complexity of the analyzed scenes. Also planning on exploring the applications of the presented method for the analysis of multitemporal data sets.

References

[1] Chang, C. I. 2000. An information-theoretic approach to spectral variability, similarity, and discrimination for hyperspectral image analysis, *IEEE Transaction Information Theory*, Vol.46, No.5, pp. 1927–1932.

[2] Chang, C. I. 2013. Hyperspectral Data Processing: Algorithm Design and Analysis, John Wiley & Sons.

[3] Guo, B., Gunn, S. R., Damper, R. I., Nelson,J. D. B. 2006. Band selection for hyperspectral image classification using mutual information, IEEE Geoscience and Remote Sensing Letters, Vol.3,No.4, pp. 522-526.

[4] Fauvel, M., Benediktsson, J. A., Chanussot, J., Sveinsson, J. R. 2008. Spectral and spatial classification of hyperspectral data using SVMs and morphological profile, IEEE Transaction on Geoscience and Remote Sensing, Vol.46, No.11, pp. 3804-3814.

[5] Mathieu, F., Jocelyn, C., JónAtli, B., 2009. Kernel principal component analysis for the classification of hyperspectral remote-sensing data over urban areas, EURASIP Journal on Advances in Signal Processing.

[6] Poggi, G., Scarpa, G., Zerubia, J. B. 2005. Supervised segmentation of remote sensing images based on a tree structure MRF model, IEEE Transaction on Geoscience and Remote Sensing, Vol.43,No.8, pp. 1901-1911.

[7] Pesaresi, M., Benediktsson, J. A. 2001. A new approach for the morphological segmentation of high-resolution satellite imagery," IEEE Transaction on Geoscience and Remote Sensing, Vol.39,No.2, pp. 309-320.

[8] Luo, B., Aujol, J. F., Gousseau, Y., 2009. Local scale measure from the topographic map and application to remote sensing images, Multiscale modeling & simulation, Vol.8, No.1, pp.1-29.

[9] Cipolla, R., Battiato, S., Farinella, G. M. 2012. Machine Learning for Computer Vision , Springer Publishing Company, Incorporated.

[10] Eftekhar, A., Toumazou, C., Drakakis, E. M., 2013. Empirical Mode Decomposition: Real-Time Implementation and Applications, Journal of Signal Processing Systems, Vol.73,No. 1, pp 43-58.

[11] Bruce,L. M., Koger, C. H., Li, J., 2002. Dimensionality reduction of hyperspectral data using discrete wavelet transform feature extraction, IEEE Transactions on Geoscience and Remote Sensing, Vol.40 ,No.10, pp. 2331 – 2338.

[12] Ahmad, M., Lee, S., Haq, I. U., Mushtaq, Q. 2012. Hyperspectral Remote Sensing: Dimensional Reduction and End member Extraction, International Journal of Soft Computing and Engineering, Vol.2,No.2, pp.170 – 175.

[13] Burgers, K., Fessehatsion, Y., Rahmani, S., Seo, J., Wittman, T., 2009. A Comparative Analysis of Dimension Reduction Algorithms on Hyperspectral Data, LAMDA Research Group.

[14] Berge, A., Solberg, A. H. S. 2006. Structured Gaussian components for hyperspectral image classification, IEEE Transaction on Geoscience Remote Sensing, Vol.44, No.11, pp. 3386–3396.

[15] Kuo, B. C., Landgrebe, D., 2001. Improved Statistics Estimation and Feature Extraction for Hyperspectral Data Classification, Department Of Electrical And Computer Engineering Technical Reports, pp.1-91.

[16] Camps-Valls, G., Bruzzone, L. 2005. Kernel-based methods for hyperspectral image classification, IEEE Transaction on Geoscience Remote Sensing , Vol.43,No.6, pp. 1351–1362.

[17] Melgani, F., Bruzzone, L. 2004. Classification of hyperspectral remote sensing images with support vector machines, IEEE Transaction on Geosciences Remote Sensing , Vol.42, No.8,2004, pp. 1778–1790.

[18] Mountrakis, G. J., Ogole, C. 2011. Support Vector Machines in Remote Sensing: A Review, ISPRS Journal of Photogrammetry and Remote Sensing, Vol.66, No.3, pp.247–259.

[19] Renzuang, Z., Carranza, M. 2011. Support vector machine: A tool for mapping mineral prosperity, Computers and geosciences, Vol.37, pp.1967-1975.

[20] Tao,Q., Chu, D., Wang, J. 2008. Recursive Support Vector Machines for Dimensionality Reduction, IEEE Transactions on Neural Networks, Vol. 19, No.1, pp.189–193.

[21] Kaewpijit, S., Le Moigne, J., El-Ghazawi, T. 2003. Automatic reduction of hyperspectral imagery using wavelet spectral analysis," IEEE Transaction on Geoscience Remote Sensing , Vol.41,No.4, pp. 863–871.

[22] Liu, C., 2004. Gabor-based kernel PCA with fractional power polynomial models for face recognition, IEEE Transaction on Pattern Analysis Machine Intelligence, Vol.26,No.5, pp. 572–581.

[23] Suguna, N., Thanushkodi, K. G, 2011. An Independent Rough Set Approach Hybrid with Artificial Bee Colony Algorithm for Dimensionality Reduction, American Journal of Applied Science, Vol. 8 No.3, pp. 261 – 266.

[24] Bao, L., Zeng, J. C., 2009. Comparison and Analysis of the Selection Mechanism in the Artificial Bee Colony Algorithm, Proceedings of IEEE Ninth International Conference on Hybrid Intelligent Systems, Vol.1, pp.411-416.

[25] Bernhard,S., Alex,S., 2002. Learning with kernels, MIT Press, Cambridge, MA.

[26] Samuelson, F., Brown, D. G. 2011. Application of Cover's theorem to the evaluation of the performance of CI observers, International Joint Conference Neural Networks (IJCNN), pp.1020 – 1026.

[27] Soria-Olivas, E., Martín-Guerrero, J. D., Camps-Valls, G., Serrano-López, A. J., Calpe-Maravilla, J., Gómez-Chova, L. 2003. A low complexity fuzzy activation function for artificial neural networks , IEEE Transaction Neural Networks, Vol.14, No.6, pp.1576–1579.

[28] Regan, D., Srivatsa, S. K. 2014. Mixed Pixel Wise Characterization Based on HMM and Hyper spectral Image Gradient Enhancement for Classification Using SVM-FSK, International Review on Computers and Software (IRECOS), Vol.9, No.6,pp.1017-1026.

[29] Guedon, Y. 2003. Estimating hidden semi-Markov chains from discrete sequences, Journal of Computational and Graphical Statistics , Vol.12,No.3, pp.604–639.