Research Journal of Applied Sciences, Engineering and Technology 10(5): 504-513, 2015
DOI:10.19026/rjaset.10.2457
ISSN: 2040-7459; e-ISSN: 2040-7467

Submitted: August 22, 2014 Accepted: September 13, 2014 Published: June 15, 2015

Research Article

Restricted Bipartite Graphs Based Target Detection for Hyperspectral Image Classification with GFA-LFDA Multi Feature Selection

[1]T. Karthikeyan and [2]S. Venkatesh Kumar
[1]P.S.G. College of Arts and Science,
[2]Department of MCA, Dr. SNS Rajalakshmi College of Arts and Science, Coimbatore, India

Abstract: Hyper spectral imaging has recently become one of the most active research areas in remote sensing. Hyper spectral imagery possesses more spectral information than multispectral imagery because the number of spectral bands in hyper spectral imagery is in the hundreds rather than in the tens. However, the high dimensions of hyper spectral images cause redundancy in spatial-spectral feature domain and consider only spectral and spatial features only and ability of the classifier to excel even as training HSI images are limited. However, unless develop suitable algorithms for target detection or classification of the hyper spectral images data becomes difficult. Therefore, it is becomes essential to consider different features and find exact target detection rate to improve classification rate. In order to overcome this problem in this study presents a novel classification framework for hyper spectral data. Proposed system uses a graph based representation, Restricted Bipartite Graphs (RBG) for exact detection of the class values. Before that the feature of the HSI images are selected using the Gaussian Firefly Algorithm (GFA) for multiple feature selection and Local-Fisher's Discriminant Analysis (LFDA) based feature projection are performed in a raw spectral-spatial feature space for effective dimensionality reduction. Then RBG is proposed to represent the reduced feature results into graphical manner to solve exact target class matching problem, in hyper spectral imaginary. Classification is performed using the Hybrid Genetic Fuzzy Neural Network (HGFNN), Genetic algorithm is used to optimize the weights of the fuzzifier and the defuzzifier for labeled and unlabeled data samples. Experimentation results show that the proposed GFA-LFDA-RBG-HGFNN method outperforms in terms of the classification accuracy and less misclassification results than traditional methods.

Keywords: Gaussian Firefly Algorithm (GFA), Hybrid Genetic Fuzzy Neural Network (HGFNN), Hyper Spectral Imagery (HIS), Local-Fisher's Discriminant Analysis (LFDA), Restricted Bipartite Graphs (RBG), Spatial Gray Level Dependency (SGLD)

INTRODUCTION

The hyper spectral imaging is limited to the visible spectrum range; it facilitates what is in-discriminable to the human visual system to become discriminable. The typical Human Visual System (HVS) projects the light spectrum, the determinant of color, into a three-dimensional subspace. Recent improvement in remote sensing technologies has prepared Hyper Spectral Imagery (HSI) with good grace available to detect and classify objects on the earth by pattern recognition approaches. Hyper spectral signatures are composed of densely sampled reflectance values over a wide range of the electro-magnetic spectrum. Even though most of the conventional approaches for HSI analysis involve per-pixel spectral classification, spatial-spectral exploitation of HSI has the potential to further improve the classification performance chiefly when there is unique class-specific textual information in the scene. In recent work, incorporating spatial context (Fauvel *et al*., 2008; Tarabalka *et al*., 2009; Dell'Acqua *et al*., 2004) into per-pixel spectral classification has shown to considerably improve classification performance for HSI. One open challenge with deriving spatial features from HSI is the further deterioration in the over-dimensionality problem.

There are two common approaches for dimensionality reduction, that is, selection-based and projection-based approach. The goal of this work is to reduce the dimensionality of hyper spectral images, so that they can be simply displayed and interpreted by the human. Some researchers have also proposed nonlinear DR algorithms for remote sensing data algorithms include Supervised Local Tangent Space Alignment (SLTSA) (Ma *et al*., 2010). The primary disadvantage of these methods is that often, algorithms such as LFDA and LDA necessitate a reasonably large training sample size to effectively learn the projection.

Dimensionality is reduced still the selection of the feature subset is not achieved in earlier works, it is also

Corresponding Author: T. Karthikeyan, P.S.G. College of Arts and Science, Coimbatore, India

considered as an important in the classification task. The purpose of feature selection is to find a subset of all available features without a projection based on some criterion function. Most importantly, the subset of selected features should not significantly degrade the performance of the classifier. In previous work, GA has been shown to work very well for a variety of feature selection tasks (Ma *et al*., 2003). Every method considers multiple features that the different features are distributed in a unified feature space, because they have different physical meanings and statistical properties. It doesn't support multiple feature selection. To solve this problem a firefly algorithm is proposed in this research for multiple feature selection.

In nature Fireflies are able to produce light thanks to particular photogenic organs located very nearer to the body exterior behind a window of transparent cuticle (Babu and Kannan, 2002). Firefly algorithm has certain drawbacks such as trapping into several local optimums. Firefly algorithm do local search as well and occasionally can't get relieve of them. Firefly algorithm parameters are set constant and they do not vary by time. In each iteration, the problem of convergence speed is solved by Gaussian distribution to move all fireflies to global best.

In order to solve the target detection or classification problem with multiple feature selection methods for hyper spectral images samples, in this study proposed a restricted bipartite graphs based methods to improve the classification results of the multiple feature selection and dimensionality reduction methods. Initially the important multiple features of the hyper spectral image samples are selected using the Gaussian Firefly Algorithm (GFA) and dimensionality reduction is performed for selected features using the LFDA. After the over-dimensionality problem is solved then perform a restricted bipartite graphs for reduced dimensional feature sample to easily match the class variable of the samples to improve classification or learning results. Determine all the maximum matching class variables uniquely restricted or not are equivalent to find no more than two paths between two vertices particularly for reduced feature samples from hyper spectral images. Finally perform the hybrid genetic fuzzy neural classifier for labeled and unlabeled samples in addition to separating labeled samples in different classes from each other. The proposed HGFNN is computationally simple and the rule-bases of which have a direct interpretation. The key feature of our approach is genetic adaptation of membership functions to new data, i.e., learning is reflected in the shape of the membership functions using evolutionary techniques combined with fuzzy neural computation (Ishibuchi *et al*., 1994), it is successfully applied to the problem of learning spectral and spatial information in the pixels.

LITERATURE REVIEW

Leaf Conventional classification approaches are not used without including dimension reduction as a preprocessing step. This is owing to the 'curse' of dimensionality. Several approaches have been proposed to alleviate the effects of dimensionality on information extraction from hyper spectral data, such as Principal Component Analysis (PCA) (Jolliffe, 2002). They all based on linear projection and can result in loss of nonlinear properties of the original data.

Wang and Chang (2006) introduced three ICA-based dimensionality reduction methods for hyper spectral data. Belkin and Niyogi (2003) developed the Laplacian Eigen map for dimensionality reduction and it conserve the relative distance between data points. Chang and Yeung (2006) presented a healthy locally linear embedding for nonlinear dimensionality reduction and they confirmed that the method is suitable for outlier problem. Chen and Qian (2007) proposed an improved LLE by introducing a spatial neighbourhood window for hyper spectral dimensionality reduction.

Cai *et al*. (2007) presented a Locality Sensitive Discriminant Analysis (LSDA) algorithm to discover a projection which maximizes the margin between data points from various classes at each local area. Purposely, the data points are mapped into a subspace in which the near points with the same label are close to each other whereas the nearby points with different labels are distant.

Gomes *et al*. (2008) proposed simple voting combinations of individual classifier decisions. Always, the aforesaid methods employ educated heuristics in combining decisions from multiple decision engines whereas advocating the choice of a fixed set of features.

In Sun *et al*. (2007), the best set of features is adaptively learned from a collection of two different types of features. In recent times, a two-stage meta-classification framework (Srinivas *et al*., 2011) is proposed, in which the vector of 'soft' outputs from multiple classifiers is infer as a meta-feature vector and feed to a second classification stage to attain the final class decision. These approaches expose the presence of balancing yet correlated information present in distinct feature sets, which is subjugated to a first order by fusing classifier outputs that use these features. The approaches mentioned above don't select multiple feature based selection and target detection results become less for classification of the hyper spectral image samples.

PROPOSED METHODOLOGY

The major aim of the proposed study is to improve the classification accuracy of the systems by

Table 1: Spatial features

Feature	Formula
Angular second moment	$f_1 = \sum_i \sum_j \{p(i,j)\}^2$
Contrast	$f_2 = \sum_{n=0}^{N_s-1} n^2 \{ \sum_{i=1}^{N_\varphi} \sum_{j=1}^{N_\varphi} p(i,j)$ $\lvert i-j \rvert = n$
Correlation	$f_3 = \frac{\sum_i \sum_j (ij) p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$, $\mu = \sum_i \sum_j P_{i,j}$,and σ^2 is the variance
Sum of squares	$f_4 = \sum_i \sum_j (i-\mu)^2 p(i,j)$
Inverse difference moment	$f_5 = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j)$
Sum average	$f_6 = \sum_{i=2}^{2N_s} i p_{x+y}(i)$
Sum variance	$f_7 = \sum_{i=2}^{2N_s} (i-f_s)^2 p_{x+y}(i)$
Sum entropy	$f_8 = -\sum_{i=2}^{2N_s} p_{x+y}(i) \log \{p_x + y(i)\}$
Entropy	$f_9 = -\sum_i \sum_j p(i,j) \log(p(i,j))$
Difference variance	$f_{10} =$ variance of p_{x-y}
Difference entropy	$f_{11} = -\sum_{i=0}^{N_s-1} p_{x-y}(i) \log \{p_{x-y}(i)\}$
Information measure correlation I	$f_{12} = \frac{H_{XY} - H_{XY1}}{\max\{H_X, H_Y\}}$ Where $H_{XY} = -\sum_i \sum_j p(i,j) \log(p(i,j)$ $H_{XY1} = -\sum_i \sum_j p(i,j) \log(p_x(i) p_y(j)$ $H_{XY2} = -\sum_i \sum_j p_x(i) p_y(j) \log(p_x(i) p_y(j)$
Information measure correlation II	$f_{13} = (1 - \exp[2.0(H_{XY2} - H_{XY})]^{1/2}$
Maximal correlation coefficient	$f_{14} =$ (second largest eigenvalue ofQ)$^{1/2}$ Where $Q(i,j) = \sum_i \frac{p(i,k)p(j,k)}{p_x(i) p_y(k)}$

$p(i,j)$ $(i,j)^{th}$ entry in a normalized gray tone spatial dependence matrix $= P(i,j)/R$

$p_x(i)$ i^{th} entry in the marginal probability matrix obtained by summing the rows of $P(i,j) = \sum_{j=1}^{N_s} P(i,j)$

N_s number of distinct gray level in the quantized image. $\sum_i$ and $\sum_j$ $\sum_{i=1}^{N_s}$ & $\Sigma_{j=1}^{N_s}$ respectively

$$P_y(j) = \sum_{i=1}^{N_s} p(i,j)$$

$$P_{x+y}(k) = \sum_{i=1}^{N_s} \sum_{j=1}^{N_s} p(i,j) \quad k = 2,3, \ldots .2N_s$$

$$P_{x-y}(k) = \sum_{i=1}^{N_s} \sum_{j=1}^{N_s} p(i,j) \quad k = 0,1, \ldots . N_s - 1, |i-j| = k$$

introducing the graph based methods for reduced dimensionality features from LFDA. Since in all the existing work the identification of the class inference (class assignment) that can exploit class-conditional correlations between the aforementioned feature sets. This is reasonably a hard task as HSI images that are typically high dimensional and the number of training images corresponding to a target class is limited. In order to overcome these problems in proposed study before the classification of the reduced dimensionality features in the HSI images apply the bipartite graphs based restricted matching to solve the above mentioned problems after the completion of the dimensionality reduction using LFDA. Before that initially the features

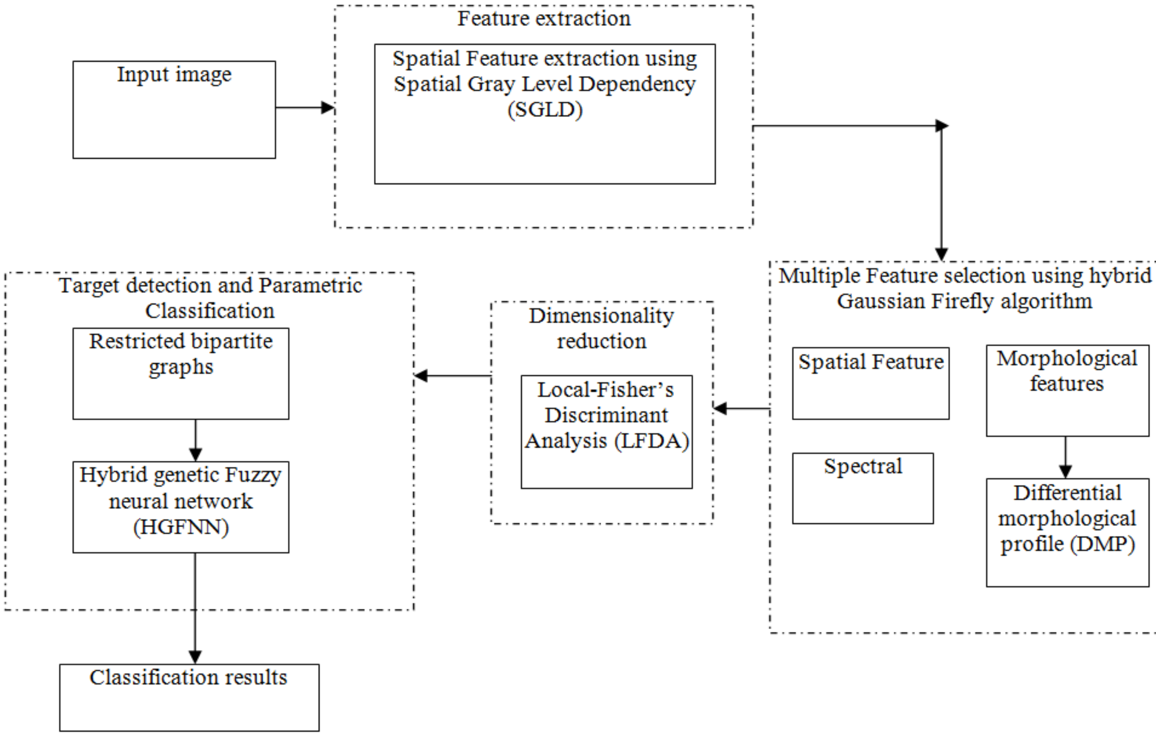

Fig. 1: Block diagram representation of the proposed study

of the hyper spectral images are extracted using the Spatial Gray Level Dependency (SGLD) is specified in Table 1. Then perform multiple feature selection using Gaussian Firefly Algorithm (GFA) which improves the Firefly behavior based on the Gaussian distribution function and dimensionality reduction of features using Local-Fisher's Discriminant Analysis (LFDA). Next apply bipartite graphs based restricted matching methods to exact matching of the class values for each features in the hyper spectral data to improve the Hybrid Genetic Fuzzy Neural Network (HGFNN) that incorporates labeled and unlabeled data in the target detection framework, it is provided with some available labeled information in addition to the unlabeled information, thus allowing encoding some knowledge about the geometry and the shape of the dataset. The block diagram representation of the proposed study is illustrated in Fig. 1.

Spatial feature extraction for HSI images: The SGLD describes spatial context of images based on how frequently two grey levels appear according to a position operator within an image. Statistical methods make use of second order statistics to form the relationships between pixels in the region by constructing Spatial Gray Level Dependency (SGLD) matrices (Kaizer, 1955). A SGLD matrix is the joint probability rate of gray levels i and j for two pixels of the spatial context of the HSI images with a definite spatial relationship in an image. The spatial relationship is defined in terms of distance d and angle θ. From each matrix, 14 statistical measures are extracted together with angular second moment, contrast, correlation, variance, inverse different moment, sum average, sum variance, sum entropy, difference variance, difference entropy, information measure of correlation I, information measure of correlation II and maximal correlation coefficient is mentioned in Table 1.

Multiple features selection using FA-KLD for HSI images: Multiple features feature selection method is performed to improve the classification accuracy of the Hybrid Genetic Fuzzy Neural Network (HGFNN) methods. The input image is represented as $x_i \in \mathbb{R}^m$ of input multiple features $\{f^k \in R^{L_k}\}_{k=1}^{m}$ in which m is the number of features and k is a specific feature within a population of m features $(k = 1, \ldots, m)$ and L_k is the length of the k^{th} feature vector. Three kinds of features are employed as a case study, i.e., the spectral feature, the morphological feature and the shape feature.

The spectral feature: The spectral feature of the each and every hyper spectral image samples is obtained by position its experimental surface reflectance in each and every one of the l bands.

The morphological feature: The morphological feature of the Hyper spectral images are selected based

Table 2: Parameters of Firefly Algorithm (FA)

Parameters	Values	Description
α	0.2	Alpha
β_0	0.3	Beta_0
γ	0.2	Gamma
Iterations	20	Generations

on the DMP (Pesaresi and Benediktsson, 2001). The proposed DMP methods stores the information of the about image structure. In hyper spectral images the two types of letters are used to store the image structure, where capital letters are used to represent the binary set values of the images and lowercase letters are used to represent the gray value of the pixels for feature extracted image samples. These two letters denotes the operations of the DMP such as opening and closing. The binary set values are represented as $G = \{p_n\}$ contains of information about number of the foreground points $p_n \in E$, where $n \in [1, N]$ and $E \subset \mathbb{Z}^2$, for hyperspectral image data samples with extracted feature results whilst $g : E \rightarrow \mathbb{R}$ is a grayscale function from grid-space E to height-values from $\mathbb{R}$. In generally the morphological operators are specified through ω (or) ω_S, where ω is defined as the square-shaped function and their corresponding scale value of the square-shaped function is represented as ω_S with attribute function Λ. This work focus a segmentation method based on the multi-scale grid for DMP concept.

The shape feature of the hyper spectral image data samples are described based on the Pixel Shape Index (PSI) (Shih, 2009), it consists of three major steps:

- For input image samples expand the direction of the each lines through the calculation of the gray-level similarity values
- Calculate the length of the each direction through the direction line results
- Lastly the shape feature can be characterized as:

$$Shape = [d_1, d_2, \ldots . d_l]^T \tag{1}$$

The extracted features then select those above mentioned features for hyper spectral image data samples are selected by using Gaussian Firefly optimization Algorithm (GFA). Fireflies are the special creatures in nature. Most of fireflies formed short and rhythmic flashes and have different flashing behavior. Fireflies employ these flashes for communication and attracting the possible prey. By seeing this behavior of fireflies Yang introduced Firefly Algorithm in 2010 (Zhang *et al*., 2006). In this study each and every firefly is well thought-out as a multiple features matrix. Then most important three features mentioned above is selected for reduced dimensionality matrix. There three idealized rules to select multiple features in the image samples:

- All fireflies are unisex. So, one feature matrix (firefly) will be paying attention to other feature images samples (fireflies) regardless of their sex.
- Attractiveness is proportional to their brightness. Therefore, for any two flashing features matrix samples (fireflies), the one which is less bright will go towards the brighter one. The attractiveness is proportional to the brightness. If there is no brighter one than a firefly will move randomly.
- The brightness of a feature matrix (firefly) is affected by the landscape of the local Fisher's ratio is considered as objective function. For a maximization problem, the brightness can simply be proportional to the value of the objective function. The parameter values of the firefly algorithm is represented in Table 2.

In this FA the selection of the multiple features is based on the movements of fireflies' moves the position of the current fireflies to overcome the problem of general FA, also this works uses a random walk for multiple feature selection meant for dimensionality reduction using LFDA.

Social behavior: Random walk is a random procedure which comprises of captivating a consecutive random step series of consecutive random steps. Here the step size or length in a random walk can be fixed or changing. If the step length complies with the Gaussian distribution, the random walk becomes the Brownian motion (Yazdani and Meybodi, 2010). In order to move all fireflies from hyper spectral image data samples from SGLD in a same manner, it is used random walk model to move all of the agents based on a Gaussian distribution. In proposed algorithm, at the end of each iteration, it is introduced normal Gaussian distribution determined by:

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

where, x is an error between best solution and fitness value of multiple features (firefly) i:

$$x = f(g_{best}) - f(fm_i) \tag{3}$$

μ is mean and δ is standard deviation. Because of use of standard normal distribution, it is set to $\mu = 0$ and $\delta = 1$. Then a random number will be drained from this Gaussian distribution that is associated to each firefly (Multiple Features matrix from SGLD) probability (p). Social behavior of fireflies is introduced by:

$$fm_i = fm_i + \alpha * (1 - \rho) * rand() \tag{4}$$

where, α is a firefly parameter to adjust adaptive parameter in Yang (2010). But if the new position of firefly i gives better fitness for that special firefly, it will move to that new position. The proposed GFA based feature selection method is based on the random walk each firefly (features) from single image feature matrix i movement is attracted to best solution that is more attractive (brighter) based on the position for each firefly for next iteration and they get more near to global best specified in Eq. (3).

Algorithm 1: Gaussian Firefly Algorithm (GFA)
Initialize algorithm parameters:
Objective function of f(x), from LFDA in Eq. (10), where $x = (x_1, \ldots\ldots, x_d)T$
Generate initial population of fireflies or x_i $(i = 1, 2, \ldots, n)$ assigning values from a feature index set $(i = 1, 2, \ldots, n)$
Define light intensity of I_i at x_i via f (x_i)
While $(t < MaxGen)$
For $i = 1$ to n $(all\ n\ fireflies)$;
For $j = 1$ to n $(all\ n\ fireflies)$
If $(I_j > I_i)$, move firefly i towards j; end if
Attractiveness varies with distance r via Exp $[-$ r2$]$
Evaluate new solutions and update light intensity;
End for j;
End for i;
Rank the fireflies and find the current best features
Define normal distribution
For $k = 1, \ldots n$ all n fireflies
Draw a random feature matrix from Gaussian distribution apply (2) for selected feature matrix value for each hyper spectral data
Evaluate new solution (new solution (k))
Id new
$(new\ solution(k)$
$< solution(i))\&\&(new\ solution(k)$
$< last\ solution(k))$
Move firefly (features) towards to the best features
End if
End for k
End while;
Post process results and visualization;
End procedure;

Dimensionality reduction of multiple features using LFDA for HSI images: Assume the training samples $\{(y_i, z_i)|Y_i \in \mathbb{R}^m, z_i \in \{1,2,\ldots,c\}, i = 1,2,\ldots,n\}$, where X_i is the training sample consists of multiple features, $\{f^k \in R^{L_k}\}_{k=1}^{m}$ in which the m is number of features (k = 1,…m) and L_k is the length of the k vector z_i is the corresponding label of Y_i, c is the number of classes and is the total number of training samples. Let n_l be the number of training samples in class ω_l and $n = \sum_{l=1}^{c} n_l$.

LFDA: LFDA is a recent extension to LDA (Sugeno and Park, 1993) that can effectively handle the multi-modal/non-Gaussian problem. It is a supervised feature projection technique that effectively combines the properties of LDA and an unsupervised manifold learning technique-Locality Preserving Projection (LPP). For more information about LPP, readers. The overall idea of LFDA is to obtain a good separation of samples from different classes while preserving the local structure of point-clouds of each class. The local within-class scatter matrix $S^{(lw)}$ and the local between class scatter matrix $S^{(lb)}$ used in LFDA are defined as follows:

$$S^{(lb)} = \frac{1}{2}\sum_{i,j=1}^{n} W_{i,j}^{(lb)} \left(X_i - X_j\right)\left(X_i - X_j\right)^T \quad (5)$$

$$S^{(lw)} = \frac{1}{2}\sum_{i,j=1}^{n} W_{i,j}^{(lw)} \left(X_i - X_j\right)\left(X_i - X_j\right)^T \quad (6)$$

where, $W^{(lb)}$ and $W^{(lw)}$ are $n \times n$ matrices defined as:

$$W_{i,j}^{(lb)} = \begin{cases} A_{i,j}\left(\frac{1}{n} - \frac{1}{n_l}\right), \text{if } y_i = y_j = l, \\ \frac{1}{n} \text{ if } y_i \neq y_j, \end{cases} \quad (7)$$

$$W_{i,j}^{(lw)} = \begin{cases} \frac{A_{i,j}}{n_l} \text{ if } y_i = y_j = l, \\ 0, \text{if } y_i \neq y_j, \end{cases} \quad (8)$$

The affinity matrix $A_{i,j}$ used in this study is defined as:

$$A_{i,j} = \exp\left(-\frac{\|x_i - x_j\|^2}{Y_i Y_j}\right) \quad (9)$$

where, $Y_i = \left\|X_i - X_j^{(k_{nn})}\right\|$ represents the local scaling of data samples in the neighborhood of X_i and X_i^k is the th-nearest neighbor of X_i. The transformation matrix $\widehat{T}_{LFDA}$ of LFDA can then be computed by maximizing the local Fisher's ratio:

$$\widehat{T}_{LFDA} = \underset{T}{\arg\max}\left\{\text{trace}\left[\left(T^T S^{(lw)} T\right)^{-1} T^T S^{(lb)} T\right]\right\} \quad (10)$$

which can be solved as a generalized Eigen value problem involving $S^{(lb)}$ and $S^{(lw)}$.

Restricted matching in bipartite graphs for target detection: From the reduced multiple feature matrix, employed for learning the class values for each and every HSI image concern several classification problems. If the number of samples becomes more since the identification of exact class becomes difficult in feature reduced matrix. In order to overcome this problem and improve the classification learning results for reduced feature matrix of multiples feature selection in this the reduced multiple feature matrix is converted into graphical form for target class identification in exact manner.

In order to perform the exact matching of the class variables for hyper spectral image data samples or reduced features matrix in the input samples, in this study use as graph based representation for reduce features results from LFDA. The samples results from $\widehat{T}_{LFDA}$ is represented as graph $G = (V, E)$ is defined by a set of nodes $V = (v_1, \dots v_n)$ and a set of (undirected) edges $E \subset V \times V$, i.e., the set of unordered pairs of nodes. A graph varies in structural complexity from sparse tree graphs to fully-connected dense graphs. In this study refer an bipartite graph is represented as $G = (X, Y, E)$, X, Y represents the different input reduced feature matrix from LFDA, corresponds to set of edges $M \subset (X, Y)$ is a matching of the class variables which don't share the common feature reduced matrix from LFDA represented in the vertex V.

A matching class variables results from the LFDA feature reduce matrix for HSI image samples M is uniquely restricted if its saturated vertices induce a subgraph which has a unique perfect matching and denotes as M_{ur}. A subset edges $S \subset M$ is a forcing set for a matching M if S is in no other ideal matching of G. If the edges in the graph of reduced feature matrix results from LFDA is matched to similar class variables in the graph is denoted as M as G (M) and name all of the vertices not saturated by M as free vertex set V_f. If the particular feature reduced matrix samples is matched to target class variables then the matching results of the feature need to satisfy the following conditions:

$$V = \{x | (x, y) \in M \tag{11}$$

$$A = \{< x_1, x_2 > | (x_1, x_1') \in M \wedge (x_2, x_2') \in M\ (x_1, x_2') \in E - M\} \tag{12}$$

In uniquely matching the class variable for each reduced feature matrix will include vertex with degree 1. In order to efficiently perform the target detection results for reduced feature matrix from LFDA extend the existing BD mapping methods. The extension of the BD method as follows: $V = V_f \cup V_1$ and $A = (A_1 \cup A_f)$, where $D(V_1, A_1)$ is a RZ-mapping digraph and A_f is a pair of arcs between $(v_i, v_j) \in G(v_i \in V_f)$ and all of $v_i \in M$ if $d(v_i) = 1$. Let D response to the extends BD-mapping digraph and V_t, Vs are the set of terminal nodes (number of the reduced feature matrix image samples from LFDA) in $D(G, M, V_f)$, when satisfies following one of three conditions:

All of $v_i \in V_f$, there exists only one path from v_i to $v_j \in V_t$.
All of $v_i \in V_f$, there exits only one path from $v_j \in V_s$ to v_i.

For any two $v_i, v_j \in V_f$, if there exists at most one $v_k \in D(G, M)$ have the path from v_i to v_k and v_j to v_k, then all maximum matching of G are uniquely restricted, where V_f is set of free nodes of $D(G, M)$.

To estimate the target class variables results in the graph define two different class distributions by p and q respectively. Let $\{y_i^p\}$trc represent the set of training feature vectors corresponding to projection P_i from LFDA and class p similarly define $\{y_i^q\}$trc. Further, let the class distributions corresponding to p and q for the i-th set of features be denoted by $f_p(y_i)$ and $f_q(y_i)$, respectively. A pair of m-node discriminative tree graphs G_i^p and G_i^q is learnt for each wavelet basis projection P_i, $i = (1, 2, .., M)$, by solving (13)-(14):

$$\hat{p} = \arg\min_{\hat{p}\ tree} D(\tilde{p}\,||\hat{p}) - D(\tilde{q}||\hat{p}) \tag{13}$$

$$\hat{q} = \arg\min_{q\ tree} D(\tilde{q}\,||\hat{q}) - D(\tilde{p}||\hat{q}) \tag{14}$$

Algorithm 2: Restricted Matching in Bipartite Graphs for target detection
1. Feature extraction (training): Obtain feature vectors results $y_i\ (i = 1, \dots M)$in $\mathbb{R}^m$ from LFDA $P_i, i = 1, \dots M$ for each input samples
2. Initial disjoint graphs
Begin
(1) Generate a BD-mapping graph $D(V, A) = f(G[M])$.
(2) If M is not perfect then
(3) Return non unique perfect matching based on the
(4) Else
(5) If D is acyclic then return unique perfect matching
(6) Else return non unique perfect matching
end if
end
3. Boosting the algorithm using the hybrid FNN

Hybrid Genetic Fuzzy Neural Network (HGFNN) methods for classification: The Genetic Fuzzy Neural Network (GFNN) algorithm is same as like Sugeno controller. Initially, the input vector results from the Restricted Matching in Bipartite Graphs $bgy_p \in \mathbb{R}^{m\times 1}$ is fuzzified by the vector of membership functions in the m distinct property dimensions, $\tilde{f} = (\mu_1, \dots \mu_m)$, $\tilde{f}$: bgx $\rightarrow [0, e]$, where e is the m-dimensional unit vector. The fuzzy vector $\tilde{f}$(bgx) is then aggregated to a fuzzy signal by a T-norm, i.e., $\propto_p = T(\tilde{f}(bgx_p))$. If needed, the input features bgx_j of bgx can be expressed in terms of memberships in each of the linguistic property sets: poor, medium and good. Input m-dimensional pattern $bgx = (bgx_1, \dots, bgx_m)$ would be represented by a with a three fuzzifier parameters:

$$\tilde{f} = (\mu_{1,poor}, \mu_{1,medium,}, \dots \mu_{m,good}) \tag{15}$$

In this condition, the memberships may possibly be represented continently by the π-function. The key features of the proposed approach are the defuzzification of the combined signal. The signal $\propto_p$ is initially mapped to each fuzzy output group by a vector $h = (\mu_{G_1}, \ldots . \mu_{G_k})$ and lastly, aggregated to a crisp group indicator by a appropriate T-conorm:

$$S\,(a, b) \coloneqq 1 - T\,(1 - a, 1 - b), a, b \in [0, 1] \quad (16)$$

This approach allows specification of part membership functions for each crisp input and for each fuzzy output group. Therefore, the final group membership (gfmp) for an observation bgxp is obtained by the following operations:

$$g_p = \mathrm{iarg}_{G_i}(S(h(\propto_p))) = \mathrm{iarg}_{G_i}(S(h(T(\tilde{\tilde{f}}(x_p))))) \quad (17)$$

where, if required $\tilde{f}$ is suitably aggregated over the linguistic properties of the each Restricted Matching in Bipartite Graphs results for each features in bgx:

$$\mathbb{R}^{m\times 1} \ni \tilde{f} = (S\left(\mu_{1,poor}, \mu_{1,medium}, \mu_{1,good}\right), \ldots . S\,(\mu_{m,poor}, \mu_{m,medium}, \mu_{m,good}) \quad (18)$$

In this case, the components of the fuzzifier $\tilde{f}$ (first hidden layer) point out the degree of membership in the fuzzy set a best economic indicators for the resultant inputs bgx_j; $j = 1, \ldots, m$. The mapping function $\tilde{h}$, once more, symbolize the degree of membership in each fuzzy output groups, as activated by the aggregate signal $\propto_p$. In FNN algorithm the misclassification takes place due to the firing level of the fuzzy signal goes beyond the point of intersection for the defuzzifier functions. On misclassification, the membership functions in the fuzzifier and defuzzifier are adjusted analytically by the GFNN-algorithm, as explained. The adjustment process consists of learning element in the algorithm. To facilitate this problem of the FNN for classification of the matched features matrix results from Bipartite Graphs, a genetic algorithm is used to adjust the parameter values of the FNN. Let us assume $L = 1$ to L_{MAX} be the number of the predefined runs and T is the maximum number of the iterations from step 1 to 4.

Step 0: State the fuzzy membership function in the fuzzifier $\tilde{f}$ and the defuzzifier $(\tilde{h})$. Give the adjustment parameters, initialize the bounds values for each weight value in the Fuzzy Neural Network (FNN),$w = \left(\tilde{f}^{\,T}, \tilde{h}^{\,T}\right)$, Let t = 0 specify the $f\,(w, C_t)$, C_t be the class and initialize the population:

$$POP(t) = \begin{bmatrix} w_{1t} \\ \vdots \\ w_{Nt} \end{bmatrix} = \begin{bmatrix} (bgx_{1t}, y_{1t}) \\ \vdots \\ (bgx_{Nt}, y_{Nt}) \end{bmatrix} \quad (19)$$

where, w_{1t} is the initial (weight) value and $w_{it} \in [\mathrm{left}, \mathrm{right}]$:

$$w_{ik} = \begin{cases} w_{ij} + \Delta(t, \mathrm{right}(j) - w_{ik}) \text{ if } r = 0 \\ w_{ij} + \Delta\left(t, w_{ik} - \mathrm{left}(j)\right) \text{ if } r = 1 \end{cases} \quad (20)$$

With,

$$\Delta(t, \mu) = \mu\alpha\left(1 - \frac{t}{T}\right)^b \quad (21)$$

$\alpha \in [0,1]$ is random variable and b is the predefined system parameter influential the degree of the mutation operator. IJ $j \in [k + 1, k + s]$ the w_{ij} is correctly rounded (upwards if r = 0 and downwards if r = 1).

Step 1: Estimate the pop (t) in the ascending order by means of $f\,(w, C_t)$. Load input feature mapped matching graph results from the bgx and categorize the samples through FNN. If $\hat{g}_p \neq g_p$ then go to step 2 or else go to step 3.

Step 2: Let $\mu_{\hat{g}_p}$ and μ_{g_p} corresponds to the fuzzy membership function result of the false and true results. If $\hat{g}_p \neq g_p$ go to step 3 else adjust the membership function in the fuzzifier $\tilde{f}$ and aggregate the mapping vector $\tilde{h}$.

Step 3: Select the $\frac{N}{2}$ best individuals from pop (t) and generate n new individuals in pop (t) by the mixed-integer crossover and non uniform mixed integer mutation.

Step 4: Repeat the steps 1-3 until convergence take place, $L = L_{MAX}$, then define the termination condition as follows: Let $\overline{F_t}$ be the moving average of the best objective function values get hold of over a predefined number of the iterations, if $|F_t - \overline{F_t}| < \varepsilon$ then stop.

EXPERIMENTAL RESULTS

In this section, the experimental hyper spectral datasets used in this study is presented with experimental setup as shown in Fig. 2. Rio Hondo image with region of interests used to validate and quantify the efficacy of the proposed approach, as measured by classification accuracies and classification maps.

Eleven different classes were defined in this image, which are given in Table 3.

Experimental hyperspectral datasets: The foremost experimental HSI dataset in use was obtained from NASA's AVIRIS sensor and was collected over northwest Indian Pine test location in June 1992. The image correspond to a vegetation-classification state with 145×145 pixels and 220 bands in the 0.4- to

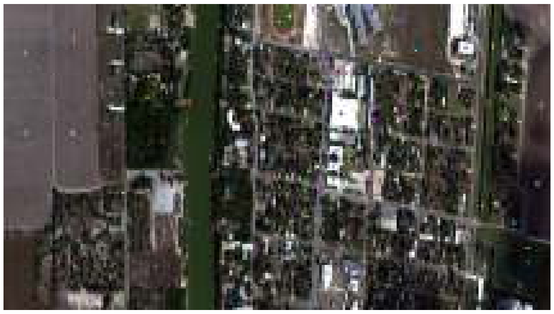

Fig. 2: Rio Hondo image with region of interests

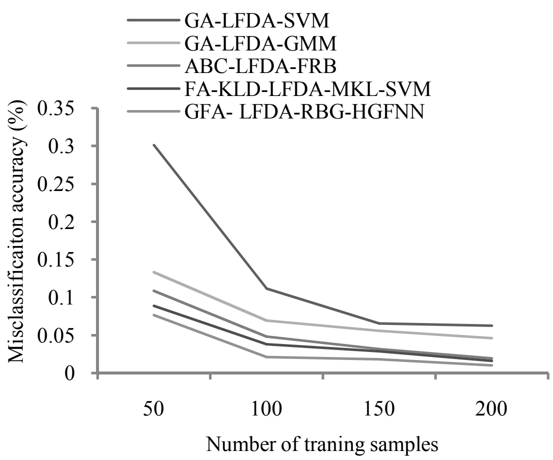

Fig. 3: Misclassification results comparison for methods

Table 3: Class cover types for the CASI data and size of the set of available labeled samples

Class name	Sample size
Tree	209
House	425
Building	204
Grass	225
Ground	206
Road	269
Parking lot	226
Water	234
Sand	202
Running lane	191
Agriculture field	246

Table 4: Overall accuracies (%) and standard deviation obtained as a function of number of training samples per class

Number of training samples	GA-LFDA-SVM	GA-LFDA-GMM	ABC-LFDA-FRB	FA-KLD-LFDA-MKL-SVM	GFA- LFDA-RBG-HGFNN
20	69.92 (2.50)	86.70 (0.68)	89.12 (0.53)	91.12 (0.48)	92.34 (0.32)
40	88.82 (0.81)	93.07 (1.38)	95.19 (1.24)	96.19 (1.04)	97.89 (0.89)
60	93.45 (0.42)	94.41 (0.63)	96.84 (0.55)	97.14 (0.45)	98.17 (0.38)
80	93.74 (0.27)	95.38 (0.53)	98.04 (0.48)	98. 40 (0.32)	98.97 (0.32)

2.45-μm region of the evident and infrared spectrum with a spatial resolution of 20 m. From the 16 different land-cover classes in the image, seven classes are redundant due to their inadequate number of training samples. Twenty noisy bands are taken away in the sight covering the region of water absorption and 200 spectral bands are used in the experiments.

Table 5: Number of features selected by FA, ABC and extracted by LFDA as a function of number of training samples per class

Number of training samples per class	20	40	60	80
Number of features selected by ABC	21	67	171	185
Number of features extracted by LFDA	9	10	18	42
Number of features selected by FA-KLD	18	64	165	169
Number of features extracted by LFDA	7	8	13	39
Number of features selected by GFA	15	57	157	161
Number of features extracted by LFDA	5	6	9	32

Performance evaluation: Performance of the proposed method is measured using experimental implementation representing challenging real-world state of affairs. In the initial setting, the classification accuracy is considered as a function of changing number of training samples. Due to the large number of spectral and spatial features, the preliminary feature selection using ABC would be beneficial, especially when the number of training sample is small. Table 4 show the overall mean classification accuracy and standard deviation around the mean of this GFA-LFDA-RBG-HGFNN algorithm, comparing it against mean accuracies with FA-KLD-LFDA-MKL-SVM, ABC-LFDA-FRB, GA-LFDA-GMM and GA-LFDA followed by SVM (GA-LFDA-SVM).

Note the high classification performance of the GFA-LFDA-RBG-HGFNN algorithm approach when spatial-spectral-morphological features are employed. It is also clear that the FA-KLD-LFDA-KLM-SVM algorithm approach is very robust to the amount of training samples employed for performing the FA-KLD search and for training the training classifier. Hence infer that GFA- LFDA-RBG-HGFNN algorithm is very effective at exploiting spatial-spectral features extracted from hyper spectral imagery. Also report the optimal number of features selected by FA-KLD and extracted by LFDA for University of Pavia data in Table 5.

Probability of misclassification accuracy results for the training samples of the learning algorithms are shown in Fig. 3. It shows that the misclassification results of the proposed GFA-LFDA-RBG-HGFNN algorithm, have less when compare to the existing FA-KLD-LFDA-MKL-SVM, ABC-LFDA-FRB, GA-LFDA-GMM and GA-LFDA followed by SVM (GA-LFDA-SVM).

CONCLUSION

The experimental results reported in this study are very promising, resulting in very high classification accuracies and demonstrate the efficacy of the proposed GFA-LFDA-RBG-HGFNN system for addressing the classification problem, multiple feature selection, target class matching problem, dimensionality reduction for small sample size as well as mixed pixel conditions. In this research, a novel Restricted Bipartite Graphs (RBG) method is introduced to solve target classification problem. Multiple features of the hyper spectral images are selected using Gaussian Firefly Algorithm (GFA) and dimension reduction algorithm is performed by using LFDA. The LFDA represents a meaningful low-dimensional features results pruning out the irrelevant features for classification tasks. In Restricted Bipartite Graphs (RBG) the reduced feature matrix are converted in the form of graph and perform exact matching of the class values, then classification learning is performed using the Hybrid Genetic Fuzzy Neural Network (HGFNN). It demonstrates the efficacy of the proposed GFA-LFDA-RBG-HGFNN achieves consistent improvements in classification performance. In future research, some other hybrid classification methods will choose and it is applied to other type of the sensor image datasets.

REFERENCES

Babu, B.G. and M. Kannan, 2002. Lightning bugs. Resonance, 7(9): 49-55.

Belkin, M. and P. Niyogi, 2003. Laplacian eigenmaps for dimensionality reduction and data representation. Neural Comput., 15: 1373-1396.

Cai, D., X.F. He and K. Zhou, 2007. Locality sensitive discriminant analysis. Proceeding of the International Conference on Artificial Intelligence. Hyderabad, India, pp: 708-713.

Chang, H. and D.Y. Yeung, 2006. Robust locally linear embedding. Pattern Recogn., 39(6): 1053-1065.

Chen, G.Y. and S.E. Qian, 2007. Dimensionality reduction of hyperspectral imagery using improved locally linear embedding. J. Appl. Remote Sens., 1(1).

Dell'Acqua, F., P. Gamba, A. Ferrari, J. Palmason, J. Benediktsson and K. Arnason, 2004. Exploiting spectral and spatial information in hyperspectral urban data with high resolution. IEEE Geosci. Remote S., 1(4): 322-326.

Fauvel, M., J. Benediktsson, J. Chanussot and J. Sveinsson, 2008. Spectral and spatial classification of hyperspectral data using SVMS and morphological profiles. IEEE T. Geosci. Remote, 46(11): 3804-3814.

Gomes, J.P.P., J.F.B. Brancalion and D. Fernandes, 2008. Automatic target recognition in synthetic aperture radar image using multiresolution analysis and classifiers combination. Proceeding of the IEEE Radar Conference (PADAR'08). Rome, pp: 1-5.

Ishibuchi, H., K. Nozaki, N. Yamamoto and H. Tanaka, 1994. Construction of fuzzy classification systems with rectangular fuzzy rules using genetic algorithms. Fuzzy Set. Syst., 65: 237-253.

Jolliffe, I.T., 2002. Principal Component Analysis. Springer, New York.

Kaizer, H., 1955. A quantification of textures on aerial photographs. Technical Note 221, Boston University Research Laboratories, Boston, MA.

Ma, L., M.M. Crawford and J. Tian, 2010. Generalised supervised local tangent space alignment for hyperspectral image classification. Electron. Lett., 46(7): 497-498.

Ma, J., Z. Zheng, Q. Tong and L. Zheng, 2003. An application of genetic algorithms on band selection for hyperspectral image classification. Proceeding of the IEEE International Conference on Machine Learning and Cybernetics, 5: 2810-2813.

Pesaresi, M. and J.A. Benediktsson, 2001. A new approach for the morphological segmentation of high-resolution satellite imagery. IEEE T. Geosci. Remote, 39(2): 309-320.

Shih, F.Y., 2009. Image Processing and Mathematical Morphology: Fundamentals and Applications. CRC Press, Boca Raton.

Srinivas, U., V. Monga and R.G. Raj, 2011. Meta-classifiers for exploiting feature dependencies in automatic target recognition. Proceeding of the IEEE Radar Conference, pp: 147-151.

Sugeno, M. and G.K. Park, 1993. An approach to linguistic instructions based learning. Int. J. Uncertain. Fuzz., 1: 19-56.

Sun, Y., Z. Liu, S. Todorovic and J. Li, 2007. Adaptive boosting for SAR automatic target recognition. IEEE T. Aero. Elec. Sys., 43(1): 112-125.

Tarabalka, Y., J. Benediktsson and J. Chanussot, 2009. Spectral-spatial classification of hyperspectral imagery based on partitional clustering techniques. IEEE T. Geosci. Remote, 47(8): 2973-2987.

Wang, J. and C.I. Chang, 2006. Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis. IEEE T. Geosci. Remote, 44(6): 1586-1600.

Yang, X.S., 2010. Firefly algorithm stochastic test functions and design optimisation. Int. J. Bioinspired Comput., 2(2): 78-84.

Yazdani, M. and M.R. Meybodi, 2010. AFSA-LA: A new model for optimization. Proceeding of the 15th annual CSI Computer Conference (CSICC'10), Tehran, Iran.

Zhang, L., X. Huang, B. Huang and P. Li, 2006. A pixel shape index coupled with spectral information for classification of high spatial resolution remotely sensed imagery. IEEE T. Geosci. Remote, 44(10): 2950-2961.