

A Neural Network Approach to Identify Hyperspectral Image Content

Puttaswamy M. R.¹, Balamurugan P²

¹Research and Development Center, Bharathiar University, India

²Department of Computer Science, Government Arts College, India

Article Info

Article history:

Received Nov 26, 2017

Revised Feb 16, 2018

Accepted Mar 30, 2018

Keyword:

Band selection algorithm (BSA)

Hyperspectral image (HSI)

Linear projection (LP)

Monogenetic binary feature (MBF)

Neural network (NN)

ABSTRACT

A Hyperspectral is the imaging technique that contains very large dimension data with the hundreds of channels. Meanwhile, the Hyperspectral Images (HSIs) delivers the complete knowledge of imaging; therefore applying a classification algorithm is very important tool for practical uses. The HSIs are always having a large number of correlated and redundant feature, which causes the decrement in the classification accuracy; moreover, the features redundancy come up with some extra burden of computation that without adding any beneficial information to the classification accuracy. In this study, an unsupervised based Band Selection Algorithm (BSA) is considered with the Linear Projection (LP) that depends upon the metric-band similarities. Afterwards Monogenetic Binary Feature (MBF) has consider to perform the 'texture analysis' of the HSI, where three operational component represents the monogenetic signal such as; phase, amplitude and orientation. In post processing classification stage, feature-mapping function can provide important information, which help to adopt the Kernel based Neural Network (KNN) to optimize the generalization ability. However, an alternative method of multiclass application can be adopt through KNN, if we consider the multi-output nodes instead of taking single-output node.

Copyright © 2018 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Puttaswamy M. R.,
Research and Development Center,
Bharathiar University,
Coimbatore, Tamil Nadu 641046, India.
Email: mrp.gowda@gmail.com

1. INTRODUCTION

The HSI consist of very large dimension data with the hundreds of channels that ranging from the short infrared wave to the visible region at 'electro-magnetic (EM) spectrum' [1]. A hyperspectral is the imaging technique, which acquires the objects information based upon their EM-spectrum with the wavelength of 400nm to 2500nm. Meanwhile, the HIS delivers the complete knowledge of imaging, thus the several of application like as; material identification, target detection and object discovering has reported in [2]. Information extraction is very significant process in HSI; therefore applying a classification algorithm is very important tool for practical uses.

However, the HSI are always having a large number of correlated and redundant feature, which causes the decrement in the classification accuracy [3]; moreover, the features redundancy come up with some extra burden of computation that without adding any beneficial information to the classification accuracy. Hence, HSI data processing (that contain high volume data) become somewhat difficult, especially with the supervised learning method because the accuracy of classification decreases with specific set of training as increasing features number, this called as 'Hughes Occurrence'. In order to achieve higher classification accuracy, dimensionality reduction approach becoming very beneficial and it also reduces the

requirement of data storage and computational time [4]. In [5], they concluded that reducing in number of features tends to achieve better classification accuracy.

In dimensionality reduction, the most commonly applied technique is subjected to this paper is feature selection (FS), which chooses an important sub-set feature from the main feature set and dispose the remaining features. Another approach is feature extraction (FE), which generally extracts the important properties from a feature set and then transforms the 'main data' to create that more separable. Considering both the feature approach, a feature selection is most suitable for the 'real dataset' with respect to feature extraction. However, the HSI provides a comprehensive separation of analogous surface materials, this kind of spectral feature are specially correlate with the adjacent bands, hence it provide a redundant information [6]. A Hughes effect [7], is related to low generalization ability of the 'classifier', which frequently has encountered in several 'pattern recognition' applications such as; object recognition, text categorization, computer vision and gene expression data [8], [9]. There are several feature techniques such as 'MNF' (minimum-noise fraction [10]), 'PCA' (principal component analysis [11]), 'SPP' (sparsity preserving projection [12]), 'LPP' (local preserving projection [13]), 'MSME' (multi-structure manifold embedding [14]), and 'SPA' (sparsity preserving analysis [15]), etc. Though the some important information at feature approach has not obtained properly, therefore it causes the performance degradation in HSI classification. Therefore, it is imperious to develop an efficient and new feature selection technique, which integrates the spectral band selection and classification to remotely sense the HSI.

The dimensionality can be achieve through 'band selection approach' that is BSA; there are two types of BSA (i.e., supervised and unsupervised BSA). Here, we are performing unsupervised BSA, which is used to get the superlative information bands without having any information of objectives. In this study, an unsupervised based BSA is considered with the LP that depends upon the metric-band similarities. Afterwards MBF has consider to perform the 'texture analysis' of the HSI, where three operational component represents the monogenetic signal such as; phase, amplitude and orientation. In post processing classification stage, Neural Network (NN) performs the classification process after the feature extraction from MBF.

The NN comprises one hidden layer and one output layer; the weight assignment is perform randomly in between 'Input of NN' and hidden layer. Feature mapping function can provide important information, which help to adopt the kernel based NN (KNN) to optimize the generalization ability [16]. However, an alternative method of multiclass application can be provide through KNN, if we consider the multi-output nodes instead of taking single-output node. The two HSI data set namely as Salinas-scene [17] and Pavia University [17], we are going to use throughout this paper to validate our classification results with respect other state-of-art methods. This paper is organized as follows; Section 2 provide the detailed survey of state-of-the-art, Section 3 gives the proposed methodologies knowledge, Section 4 provides the experimental result and analysis of classification. Finally, Section 5 concludes this paper.

2. LITRATURE SURVEY

Analysis of HSIs task is difficult because the data sets which are having extremely large dimensionality, so choose some essential features it is most significant, and helpful for learning. In order to select essential features of heterogeneous, in [18] they has proposed a method called sparse feature selection method which is built on regularized regression model. However, in this presentation with the noise information are contaminated because of imaging devices. That noise will affects the process of learning for example hyperspectral images of high dimensional data analysis. Redundancy reducing as well as preserving data these two critical problems, which are necessarily to be handled: at the feature selection. In [19], based on a recently designed memetic procedure they proposed a feature selection technique for hyperspectral image classification. They designed a suitable objective task inside the proposed technique, that able be measure the enclosed redundancy info as well as essential info on the selected feature subsets. In [20], they selected no redundant information plus Gabor features for hyperspectral image classification and proposed a Markov blanket based symmetrical uncertainty approach.

A fast forward (FS) approach [21], which based upon Gaussian Mixture (GM) prototype classifier. The GM-classifier has used for classifying the hyperspectral images, this approach choose the spectral feature to increase the prediction rate of classification. This can execute through 'k-fold' cross validation in order to achieve efficient implementation and fast computing time. Initially, the GM can upgraded with computed classification rate, instead of re-estimate overall model. Finally, GM marginalization can be sub-model from the full learning model of spectral features. The investigation of the unsupervised-BSA [22] for the spectral FS and the classification process based upon SVM (Support Vector Machine) [23], are mostly used in morphological profiles ('MP'). The difficulty at high dimensionality can minimize by resulting features and it is most usual activity where MP has extracted from PCA.

The main challenge of classification problems is selection of a subset, which has efficient and minimum from the enormous amount of features. The entire feature selection methods consist of a search algorithm used for selection of high time consuming finest candidate from the possible answers. In [24], proposed dual feature selection approaches without requirement of search algorithm. With using mean values plus simple calculation of standard deviation effective features of subset are selected from the method. In [25], [26], they expressed two class classification issue, where modified (KNN) k nearest neighbors classifier method has been applied to create validity rating, maximal coherence, plus categorize the test samples by k-fold validation. In [27], they emphasis on the HIS classification data, which captured from the Environmental Program (EnP). Moreover, they considered the dataset that presented with EnP contest, a standard has recently initialized with several classifying object and the land use classes depends upon hyperspectral data. In this [28], they has considered 'Radial Basis Function' (RBF), where full bandwidth-RBF kernel has used as feature weights, whenever the feature values become rescaled through the z-scores value.

The methodology of HIS allows recognizing materials through using photo-thermal 'infrared spectroscopy' [29]-[30]. In order to obtain the infrared spectroscopy, an infrared camera capture the hundreds of images of an object at different channels wavelength simultaneously while a 'QCL' (quantum cascade laser) causes the subjected material to be brightened. The main problems that effecting the classification arrives from the disparaging ratio of high dimension HIS and the small size 'training data' [31]. These kind of problems causes advancement in machine learning approaches such as; multiple learning systems, transductive learning, semi-supervised learning and active learning [32]-[34] and several methods of learning depend upon the unlabeled samples [35], [36]. To overcome from such kind of issues, some strategy has considered enhancing new similarity measurement for the band reduction and dimension reduction of HIS [37], [38]. However, the high correlation between the bands has misidentify for designing new techniques to reduce the data dimensionality, which including some of the methods that has great acceptance such as; PCA [39], 'linear discriminant analysis' (LDA) [40] or MNF [41]. It also has been investigated in preceding works that HIS classification after the preprocessing process dimensionality reduction 'or' band selection are usually outperforms better classification with respect to traditional HSI data classification [42], [43], which in order to reduce the computational complexity.

3. PROPOSED HYPERSPECTRAL CLASSIFICATION ALGORITHM

The presence of high spectral correlation in the HSI causes difficulty in feature selection process that is why it's necessary to perform dimensional reduction at spectral feature in order to obtain a useful 'compact set' spectral feature. The dimensionality can be achieve through 'band selection approach' that is BSA; there are two types of BSA (i.e., supervised and unsupervised BSA). Here, we are performing unsupervised BSA, which is used to get the superlative information bands without having any information of objectives. In this study, an unsupervised based BSA is considered with the linear projection that depends upon the metric-band similarities. This methodology does not require any information regarding prior-class to categorize the level of data analysis. Here, we select the two different bands (i.e., τ_1, τ_2) from LP-BSA [44] and the subset of bands is defined as;

$$\emptyset = \{\tau_1, \tau_2\} \quad (1)$$

The new selected subset of band is

$$\emptyset = (\emptyset) \cup (\tau_3) \quad (2)$$

where, τ_3 is a third band and above Equation (2) has repeated until third band is more than the subset of first two bands, when this criteria has satisfy then only the Equation (2) is updated.

$$\tau' = t_0 + t_1 \times (\tau_1) + t_2 \times (\tau_2) \quad (3)$$

Here, τ' denotes the prediction band of LP by selecting first two bands (i.e., τ_1, τ_2) and the error in LP can be minimalize through selecting t_0, t_1 and t_2 parameters,

$$\mathbb{E}_{LP} = \|\tau - \tau'\| \quad (4)$$

In order to calculate the minimal square solution,

$$t = (t_0 \times t_1 \times t_2)^Y \quad (5)$$

Furthermore,

$$t = (1/(\mathbb{x}^Y \times \mathbb{x})) \times \mathbb{x}^Y \times \delta \quad (6)$$

Where, \mathbb{x} denotes the 'input data' in matrix format (i.e., $V \times 3$ matrix form). Here, first column contain single value, second column contains all pixels values of band τ_1 and, third column contains all pixels values of band τ_2 . Additionally, the δ represents a vector matrix $V \times 1$ that include the all pixel values of τ , the bands τ_1 and τ_2 are compared with τ_3 band, if the value of third band found more than the subset of band (\emptyset) is updated.

The MBF [45] has considered performing the 'texture analyses of the HSI, where three operational components represents the monogenetic signal such as; phase, amplitude and orientation. The operators of MBF can be given as 'MBF-P' (MBF-Phase), 'MBF-A' (MBF-Amplitude) and 'MBF-O' (MBF-Orientation). This can be computed through [45], where parameters of local variable and the image intensity are integrated to provide individual operator of monogenetic signal.

$$\text{MBF-P}(A_b) = [B'_o(A_b), B'_f(A_b), B'_p(A_b)],_{Bi} \quad (7)$$

$$\text{MBF-A}(A_b) = [B'_o(A_b), B'_f(A_b), B'_A(A_b)],_{Bi} \quad (8)$$

$$\text{MBF-O}(A_b) = [B'_o(A_b), B'_f(A_b), B'_O(A_b)],_{Bi} \quad (9)$$

The MBF-X represents the pattern number of MBF ($\text{MBF-X} \in \{P, A, O\}$), here, A_b is pixel of monogenetic signal that can be encoded with 2-bits ($B'_o(A_b), B'_f(A_b)$) $_{Bi}$. Each of the MBF operator feature can be used as individually for the classification process, otherwise the combination of this can be used for further classification process. In order to provide the optimized HSI, the single-vector histogram for the MBF-P, MBF-A and, MBF-O is;

$$F_h = fm_t(x, y) \begin{cases} x = 1, \dots, l \\ y = 1, 2, 3 \\ h = \{\text{MBF-P}, \text{MBF-A}, \text{MBF-O}\} \end{cases} \quad (10)$$

where, $fm_t(x, y)$ represents the (F_h) histogram vector, h represents is feature map on the x th scale and l denotes a level of sub-region. Neural Network (NN) performs the classification process after the feature extraction from MBF. The NN comprises one hidden layer and one output layer; the weight assignment is perform randomly in between 'Input of NN' and hidden layer. Weights of the linear output layer can be determined through LRA (Linear Regression Analysis) [46], which will reduce the computation time/cost. Feature mapping function can provide important information, which help to adopt the kernel based NN (KNN) to optimize the generalization ability [16].

An alternative method of multiclass application can be provide through KNN, if we consider the multi-output nodes instead of taking single-output node. The classifier of b -class have b -number of output nodes, if the original label of class is d , then the output vector with b -output nodes is;

$$c_i = [0, \dots, 0, 1^d, 0, \dots, 0]^C \quad (11)$$

If we consider only d th element of output vector (c_i) is one, then output vector is;

$$c_i = [c_{i,1}, \dots, c_{i,b}]^C \quad (12)$$

where, rests of the components are adjusting to zero and the multi-output nodes for the classification process of NN can be providing as;

$$\text{Minimize : } Q_{P_{NN}} = 0.5\|G\|^2 + 0.5 \times H \sum_{i=1}^K \|L_i\|^2 \quad (13)$$

where user-specified parameter (H) provides an adjustment distance between the training error and the separating margin. The output ‘weight vector’ between output layer and hidden layer is denote as G . Hidden layer of output that corresponds to input sample (x) is $M(x_i)$,

$$\text{Subject to : } M(x_i)G = c_i^C - c_i^C, \text{ where } i = 1, \dots, K \quad (14)$$

The ‘training error vector’ (L_i) of b -output nodes w.r.t the (x_i) training sample is,

$$L_i = [L_{i,1}, \dots, L_{i,b}]^C \quad (15)$$

‘Karush–Kuhn–Tucker’ approach (KKT) [47] is considered for training the KNN, which consider the dual optimization problem;

$$Q_{D_{NN}} = 0.5\|G\|^2 + 0.5 \times H \sum_{i=1}^K \|L_i\|^2 - \sum_{i=1}^K \sum_{j=1}^b S_{i,j} (M(x_i)G_j - c_{i,j} + L_{i,j}) \quad (16)$$

where, G_j is weight vector that linking the hidden layer to corresponding j th output node.

$$G = [G_1, \dots, G_b] \quad (17)$$

The optimize solution can be provide through corresponding KKT-Approach, such as;

$$\frac{\partial Q_{D_{NN}}}{\partial G_j} = 0 \rightarrow G_j = \sum_{i=1}^K S_{i,j} (M(x_i))^C \rightarrow G = R^C S \quad (18)$$

where, S_i is Lagrange multiplier [48] that corresponds to the (c_i, x_i) training samples.

$$\frac{\partial Q_{D_{NN}}}{\partial L_i} = 0 \rightarrow S_i = H L_i, \quad i = 1, \dots, K \quad (19)$$

$$\frac{\partial Q_{D_{NN}}}{\partial S_i} = 0 \rightarrow M(x_i)G + L_i^C - c_i^C = 0, \quad i = 1, \dots, K \quad (20)$$

In Equations (14) to (18) shows for the specific case of the multiple-output nodes and it also can be used for a single-output node via setting b value as one. Where, the Lagrange multiplier [48] for nonlinearity function can be given as;

$$S_i = [S_{i,1}, \dots, S_{i,b}]^C \quad (21)$$

$$S = [S_1, \dots, S_N]^C \quad (22)$$

Therefore, we will consider the multiple-output nodes for the multi-class classifier; in that the hidden layer matrix (H) size can be decide through the K -number of training samples. Q is number of the hidden nodes that are unrelated to output nodes number (i.e., number of classes). Here, we consider the circumstance where training samples number is very more than the feature space dimensionality. The number of training samples $K \gg Q$, we have the alternative solution from Equations (18) and (19);

$$G = H[R]^C L \quad (23)$$

$$L = H^{-1}[R^C]^{\dagger} G \quad (24)$$

From Equation (20) we can write it as;

$$\begin{aligned} R G - C + H^{-1}[R^C]^{\dagger} G &= 0 \\ R^C C &= R^C (R + H^{-1}[R^C]^{\dagger}) G \end{aligned} \quad (25)$$

$$G = R^C C \left[R^C C + \frac{1}{H} \right]^{-1} \quad (26)$$

For this circumstances, a decision function of linear feature space ‘or’ the output functionality of Kernel based NN-classifier is,

$$U(x) = M(x)G = M(x) \times R^{CC} \times \left[R^{CC} + \frac{1}{H} \right]^{-1} \quad (27)$$

Where, R^\dagger is the ‘Moore Penrose’ generalize inverse matrix [49], [50] and matrix $I = [R^C] \times [R^C]^\dagger$. This proposed model can be used for different size of applications, however, these approaches have various computational costs, and their efficiency may vary for the different applications. It found though the KNN implementation, that the KNN performance generalization is not much sensitive towards the feature space dimensionality (Q) and the better performance can be achieved as long as ‘ L ’ is more enough. Therefore, if the data-set of training are very more (i.e., $K \gg Q$) then a preferable choice is to apply Equation (27) in order to decrease the cost of computation.

4. EXPERIMENTAL RESULTS AND ANALYSIS

4.1. Data-set description

In this paper, we will use the real HSI dataset, which is available for general-purpose use. All the experiments has performed using MATLAB simulation on an Intel-i5 2.80GHz machine, 64bit operating system with 8GB RAM. The two HIS data-set namely as Salinas-scene [17] and Pavia University [17], we are going to use throughout this paper to validate our classification results with respect other state-of-art methods. At first Salinas scene dataset has collected of Salinas valley (which is situated at California) through the ‘AVIRIS-sensor’ with 224-bands. Moreover, it can be characterize by the high ‘spatial level resolution’ that is 3.7mp; this dataset consist of many radiance data such as vineyard, bare soils, fields and vegetables. There are total sixteen classes present in the ground truth of Salinas dataset.

The second consider dataset is PaviaU, which dataset has collected of Pavia University at Northern Italy (which is during flight campaign) through the ‘ROSIS-sensor’ with 103-bands. Moreover, it can be characterize by the high ‘geometric resolution that is 1.3m with 610×610 pixels; there are total nine classes present in the ground truth of PaviaU dataset. Table 1 show the total number of samples at each class of Salinas Scene dataset [17] and it comprises of 54129 total samples. Similarly, Table 2 shows the total number of samples at each class of PaviaU dataset [17] and it comprises of 42776 total samples.

Table 1. Total Samples at each Class of Salinas Scene dataset [17]

CLASSES	TOTAL SAMPLES
Class1 ‘Brocoli-weeds_1’	2009.0
Class2 ‘Brocoli-weeds_2’	3726.0
Class3 ‘Fallow’	1976.0
Class4 ‘Rough-Fallow’	1394.0
Class5 ‘Smooth-Fallow’	2678.0
Class6 ‘Stubble’	3959.0
Class7 ‘Celery’	3579.0
Class8 ‘Grapes-untrained’	11271.0
Class9 ‘Soil-Vinyard’	6203.0
Class10 ‘Corn-green-weeds’	3278.0
Class11 ‘Lettuce_romaine-4wk’	10683.0
Class12 ‘Lettuce_romaine-5wk’	1927.0
Class13 ‘Lettuce_romaine-6wk’	916.0
Class14 ‘Lettuce_romaine-7wk’	1070.0
Class15 ‘Vinyard-untrained’	7268.0
Class16 ‘Vinyard-vertical’	1807.0
TOTAL SAMPLES	54129.0

Table 2. Total Samples at each Class of Pavia University dataset [17]

CLASSES	TOTAL SAMPLES
Class1 ‘Asphalt’	6631.0
Class2 ‘Meadows’	18649.0
Class2 ‘Gravel’	2099.0
Class2 ‘Trees’	3064.0
Class2 ‘Metal sheets’	1345.0
Class2 ‘Bare Soil’	5029.0
Class2 ‘Bitumen’	1330.0
Class2 ‘Self-Blocking-Bricks’	3682.0
Class2 ‘Shadows’	947.0
TOTAL SAMPLES	42776.0

4.2. Comparative study

Here we compare the classification performance of several state of art with respect to our proposed model. In this section, a proposed preprocessing MBF-Algorithm will compare with some existing popular feature extraction methods such as MNF [10], PCA [11], SPP [12], LPP [13], MSME [14], and SPA [15]. The kappa coefficient of classification (Kappa) and overall accuracy (OA) has used to show the accuracy after classification.

Taking a SalinasS dataset, the classification result of different method has compared with proposed MBF has shown in Table 3. The MBF (PA) is compared with respect to MNF [10], PCA [11], SPP [12], LPP [13], MSME [14], and SPA [15] method, based upon the Kappa (%), OA (%) and their individual class. In Table 3, five classes got accuracy of 100% at MBF algorithm (classes number are 3, 4, 9, 11 and 16). The overall accuracy of proposed MBF model has compared with respect to MNF [10], PCA [11], SPP [12], LPP [13], MSME [14], and SPA [15] and we got improved classification results of 6.23%, 7.8%, 9.4%, 7.9%, 4.9%, and 9.4%. The Kappa accuracy of proposed MBF model has compared with respect to MNF [10], PCA [11], SPP [12], LPP [13], MSME [14], SPA [15] and we got improved classification results of 7.03%, 8.75%, 10.57%, 8.8%, 5.5%, and 10.57%.

Table 3. Class wise Comparison on the SalinasS –dataset (%)

Class Number	MNF [10]	PCA [11]	SPP [12]	LPP [13]	MSME [14]	SPA [15]	MBF (PA)
1	99.80	99.30	98.90	99.40	99.90	88.70	99.74
2	100.0	99.60	100.0	100.0	100.0	98.60	99.89
3	100.0	99.30	97.20	100.0	100.0	98.40	100.0
4	99.80	99.70	99.30	99.60	99.90	97.50	100.0
5	97.70	96.70	92.70	97.80	98.40	94.60	96.60
6	99.80	99.70	99.90	99.90	99.90	97.00	96.70
7	99.60	98.40	99.20	99.60	99.70	98.50	99.97
8	85.00	77.40	78.90	83.10	86.30	75.20	98.73
9	99.40	98.10	96.20	99.60	99.90	96.40	100.0
10	95.60	91.10	91.60	93.00	97.30	82.80	98.17
11	99.00	98.10	96.90	98.60	99.00	93.80	100.0
12	100.0	100.0	99.90	100.0	99.90	92.60	99.67
13	98.20	99.50	98.00	98.20	100.0	87.40	99.50
14	95.30	93.10	93.90	94.60	98.50	96.80	98.97
15	74.80	81.90	69.10	66.80	80.50	70.40	98.60
16	99.00	98.70	98.50	98.90	97.60	96.30	100.0
OA (%)	92.80	91.20	89.60	91.10	94.10	89.60	98.97
Kappa (%)	91.90	90.20	88.40	90.10	93.40	88.40	98.85

Table 4 shows the class wise classification results on a PaviaU dataset. Out of nine classes, our proposed model MBF has more accuracy in five classes (class number 2,3,6,7 and 8). The overall accuracy of proposed MBF model has compared with respect to MNF [10], PCA [11], SPP [12], LPP [13], MSME [14], SPA [15] and we got improved classification results of 4.89%, 1.8%, 14.2%, 8.4%, 0.3%, and 15.8%. The Kappa accuracy of proposed MBF model has compared with respect to MNF [10], PCA [11], SPP [12], LPP [13], MSME [14], SPA [15] and we got improved classification results of 6.4%, 2.38%, 18.38%, 10.8%, 0.42%, and 20.7%.

Table 4. Class wise Comparison on the PaviaU –dataset (%)

Class Number	MNF [10]	PCA [11]	SPP [12]	LPP [13]	MSME [14]	SPA [15]	MBF (PA)
1	87.30	85.50	76.80	84.10	92.30	69.20	90.47
2	86.70	93.20	76.70	81.30	94.40	82.40	99.58
3	84.30	89.00	79.60	80.80	78.10	71.10	93.88
4	92.70	93.40	93.20	92.60	94.00	88.70	66.68
5	100.0	99.80	99.80	100.0	100.0	92.50	90.80
6	92.80	91.70	83.90	92.50	93.90	78.70	99.90
7	97.40	97.50	87.10	96.90	96.90	82.50	99.80
8	91.90	93.90	77.90	89.50	94.30	65.50	96.28
9	99.90	100.0	100.0	100.0	99.90	96.50	48.97
OA (%)	89.30	92.20	80.50	86.00	93.60	79.00	93.90
Kappa (%)	86.00	89.70	75.00	81.90	91.50	72.80	91.89

Classification results at 10-sample size of SalinasS-dataset has shown in Figure 1, where top line mark at each bar represent the error bar. Proposed model is compared in terms of OA and Kappa (KA) with MNF [10], PCA [11], SPP [12], LPP [13], MSME [14], and SPA [15]. As per Figure 1, MBF is compared in terms OA (%) and we got 11.43%, 9.9%, 15.89%, 15.69%, 13.16% and 11.13% improvement w.r.t considered state of art. Similarly, MBF is compared in terms KA (%) and we got 12.7%, 10.97%, 17.68%, 17.47%, 14.3% and 12.39% improvement w.r.t considered state of art. Figure 2 shows the classification results at 20-sample size (SalinasS-dataset), here MBF is compared in terms OA (%) and we got 11.61%, 7.96%, 12.11%, 13%, 10% and 7.9% improvement w.r.t considered state of art. Similarly, MBF is compared in terms KA (%) and

we got 12.8%, 8.8%, 13.41%, 14.4%, 11.2% and 8.65% improvement w.r.t considered state of art. In Figure 3, a proposed MBF algorithm performs 10.74%, 6.7%, 9.9%, 10.4%, 9%, 7.3% better overall accuracy than the existing feature algorithm MNF [10], PCA [11], SPP [12], LPP [13], MSME [14], and SPA [15]; in terms KA (%), we got 11.9%, 7.5%, 11%, 11.5%, 10% and 8.1% improvement. At 40-sampling size that shows in Figure 4, MBF performs 10.4%, 5.3 %, 8%, 9.5%, 6.4%, 6.6% better overall accuracy than the considered feature algorithm and in terms KA (%), we got 11.6%, 6.7%, 10%, 11.38%, 8.3% and 8.14% improvement.

As given in Figure 1 to Figure 4 for SalinasS-dataset, similarly in Figure 5 to Figure 8 has given for Pavia U-dataset. Figure 5, shows the classification results at 10-sample size (PaviaU-dataset), here MBF is compared in terms OA (%) and we got 25%, 14%, 41%, 34%, 19.5% and 33.6% improvement with respect to considered state of art methods. Similarly, MBF is compared in terms KA (%) and we got 31%, 18%, 50%, 41.4%, 24% and 41.6% improvement w.r.t considered state of art. In Figure 6 and 7, the kappa coefficient of MBF is 91.52 and 91.8, and the overall accuracy is 93.8 and 93.9. In Figure 8, a proposed MBF algorithm performs 17%, 4.4%, 18.7%, 22%, 5.8%, 14.36% better KA (%) than the existing feature algorithm MNF [10], PCA [11], SPP [12], LPP [13], MSME [14], and SPA [15]. The analysis of above result provides the performance evaluation of MBF model, where the performance in every sampling size is increasing as per increment in the training-sample-size.

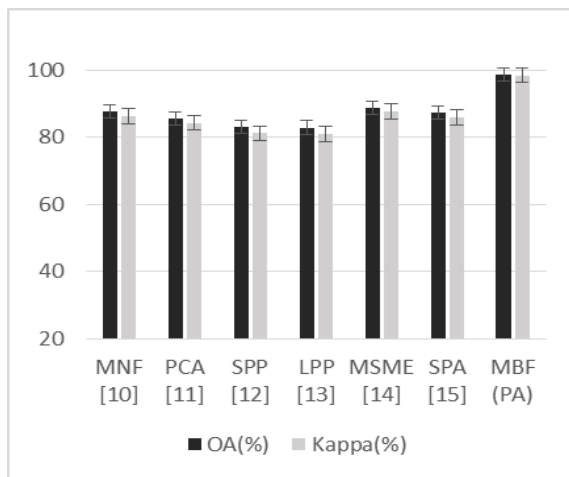


Figure 1. Classification results at 10-sample size (SalinasS-dataset)

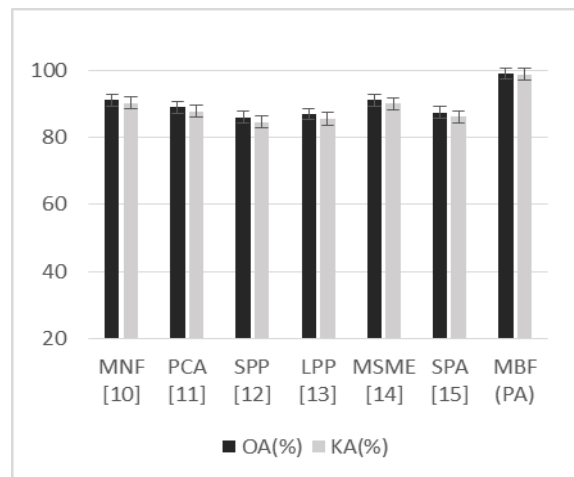


Figure 2. Classification results at 20-sample size (SalinasS-dataset)

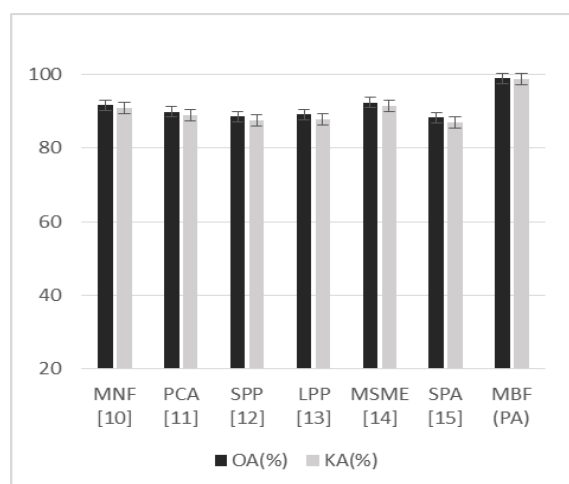


Figure 3. Classification results at 30-sample size (SalinasS-dataset)

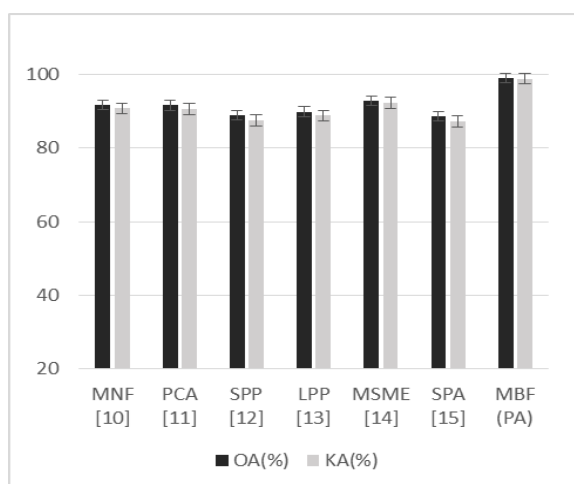


Figure 4. Classification results at 40-sample size (SalinasS-dataset)

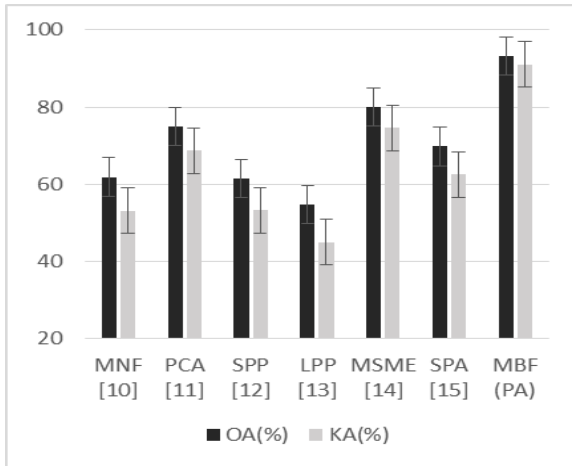


Figure 5. Classification results at 10-sample size (PaviaU-dataset)

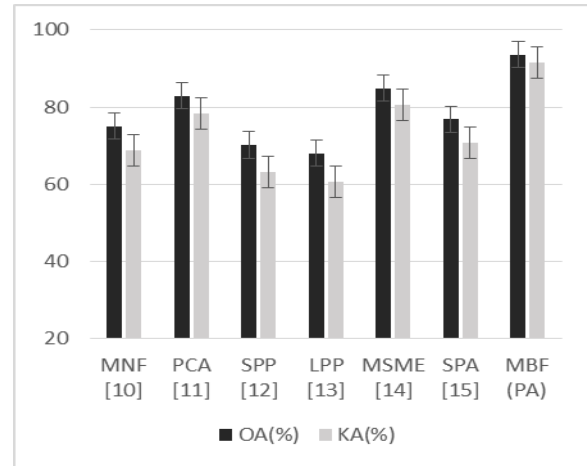


Figure 6. Classification results at 20-sample size (PaviaU-dataset)

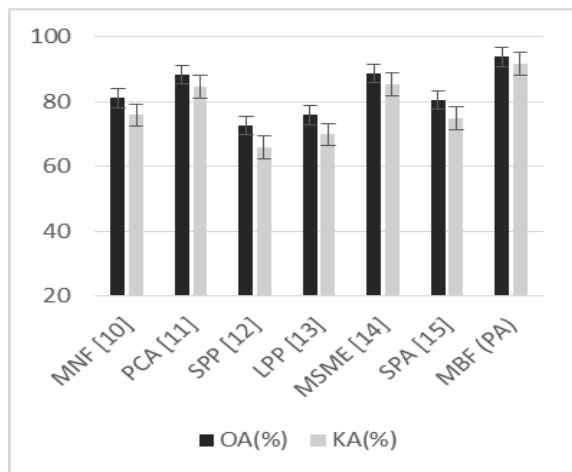


Figure 7. Classification results at 30-sample size (PaviaU-dataset)

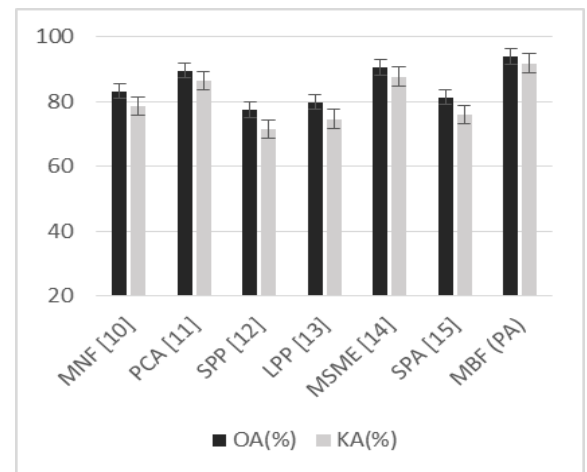


Figure 8. Classification results at 40-sample size (PaviaU-dataset)

5. CONCLUSION

The HSI are always having a large number of correlated and redundant features, which causes the decrement in the classification accuracy. Hence, HIS data processing becomes somewhat difficult, especially with the supervised learning method because the accuracy of classification decreases with specific set of training as increasing features number. In order to achieve higher classification accuracy, dimensionality reduction approach becoming very beneficial and it also reduces the requirement of data storage and computational time. In this paper, an unsupervised based BSA is considered with the LP that depends upon the metric-band similarities. Subsequently MBF has considered performing the ‘texture analyses of the HSI, where three operational components represents the monogenetic signal such as; phase, amplitude and orientation. In post processing classification stage, the KNN has applied to optimize the generalization ability by an alternative method of multiclass application through considering the multi-output nodes instead of taking single-output node.

The proposed model classification accuracy has compared with the accuracy of several state-of-art methods and from the above analysis of results our proposed model has perform considerably well in every training set. Proposed model OA and Kappa values at 40-training samples is 98.97% and 98.85% at Salinas-S dataset; similarly at Pavia-U dataset, we got 93.9% and 91.9% of overall and Kappa classification accuracy. In future work, experiments should be implement on very large datasets in order to calculate the scalability of several approaches.

REFERENCES

- [1] A. Muntasa, I. A. Sirajudin, M. H. Purnomo, "Appearance global and local structure fusion for face image recognition", *TELKOMNIKA (Telecommunication, Computing, Electronics and Control)*, 2011, vol. 9, no. 1, pp. 125-132.
- [2] H. Grahn and P. Geladi, "Techniques and applications of hyperspectral image analysis," Chichester, England; Hoboken, J. Wiley, 2007.
- [3] K. Fukunaga and R. R. Hayes, "Effects of sample size in classifier design," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 8, pp. 873-885, 1989.
- [4] M. Pal and G. M. Foody, "Feature Selection for Classification of Hyperspectral Data by SVM," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 5, pp. 2297-2307, 2010.
- [5] A. Janecek, *et al.*, "On the Relationship between Feature Selection and Classification Accuracy," *Journal of Machine Learning Research - Proceedings Track*, vol. 4, pp. 90-105, 2008.
- [6] B. M. Shahshahani and D. A. Landgrebe, "The effect of unlabeled samples in reducing the small sample size problem and mitigating the Hughes phenomenon," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 5, pp. 1087-1095, 1994.
- [7] G. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Trans. Inf. Theory*, vol. 14, no. 1, pp. 55-63, 1968.
- [8] Z. Sun, *et al.*, "Object detection using feature subset selection," *Pattern Recognit.*, vol. 37, no. 11, pp. 2165-2176, 2004.
- [9] I. Guyon, *et al.*, "Gene selection for cancer classification using support vector machines," *Mach. Learn.*, vol. 46, pp. 1-3, pp. 389-422, 2002.
- [10] G. Lixin, *et al.*, "Segmented minimum noise fraction transformation for efficient feature extraction of hyperspectral images," *Pattern Recognit.*, vol. 48, no. 10, pp. 3216-3226, 2015.
- [11] X. M. Cheng, *et al.*, "A novel integrated PCA and FLD method on hyperspectral image feature extraction for cucumber chilling damage inspection," *Trans. ASAE*, vol. 47, no. 4, pp. 1313-1320, 2004.
- [12] L. Qiao, *et al.*, "Sparsity preserving projections with applications to face recognition," *Pattern Recognit.*, vol. 43, no. 1, pp. 331-341, 2010.
- [13] Z. Wang and B. He, "Locality perserving projections algorithm for hyperspectral image dimensionality reduction," in *Proc. 19th Int. Conf. Geoinform., Shanghai, China*, pp. 1-4, 2011.
- [14] Y. Gan, *et al.*, "Feature Extraction Based Multi-Structure Manifold Embedding for Hyperspectral Remote Sensing Image Classification," in *IEEE Access*, vol. 5, pp. 25069-25080, 2017.
- [15] F. Luo, *et al.*, "Fusion of graph embedding and sparse representation for feature extraction and classification of hyperspectral imagery," *Photogram. Eng. Remote Sens.*, vol. 83, no. 1, pp. 37-46, 2017.
- [16] K. R. Muller, *et al.*, "An introduction to kernel-based learning algorithms," in *IEEE Transactions on Neural Networks*, vol. 12, no. 2, pp. 181-201, 2001.
- [17] http://www.ehu.es/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes
- [18] Q. Zhang, *et al.*, "Automatic spatial-spectral feature selection for hyperspectral image via discriminative sparse multimodal learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 1, pp. 261-279, 2015.
- [19] M. Zhang, *et al.*, "Memetic algorithm based feature selection for hyperspectral images classification," *2017 IEEE Congress on Evolutionary Computation (CEC), San Sebastian*, pp. 495-502, 2017.
- [20] L. Shen, *et al.*, "Discriminative Gabor Feature Selection for Hyperspectral Image Classification," in *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 1, pp. 29-33, 2013.
- [21] M. Fauvel, *et al.*, "Fast Forward Feature Selection of Hyperspectral Images for Classification With Gaussian Mixture Models," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 6, pp. 2824-2831, 2015.
- [22] K. Tan, *et al.*, "Hyperspectral Image Classification Using Band Selection and Morphological Profiles," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 1, pp. 40-48, 2014.
- [23] M. Pal and G. M. Foody, "Feature Selection for Classification of Hyperspectral Data by SVM," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 5, pp. 2297-2307, 2010.
- [24] M. Imani and H. Ghassemian, "Fast feature selection methods for classification of hyperspectral images," *Telecommunications (IST), 2014 7th International Symposium on, Tehran*, pp. 78-83, 2014.
- [25] N. Zhang, *et al.*, "An Enhanced K-Nearest Neighbor Classification Method Based on Maximal Coherence and Validity Ratings," *14th International Symposium, Japan*, 2017.
- [26] A. Woodley, *et al.*, "Efficient Feature Selection and Nearest Neighbour Search for Hyperspectral Image Classification," *2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA), Gold Coast, QLD*, pp. 1-8, 2016.
- [27] S. Keller, *et al.*, "Investigation of the impact of dimensionality reduction and feature selection on the classification of hyperspectral EnMAP data," *2016 8th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Los Angeles, CA*, pp. 1-5, 2016.
- [28] P. J. Hsieh, *et al.*, "A nonlinear feature selection method based on kernel separability measure for hyperspectral image classification," *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan*, pp. 461-464, 2015.
- [29] C. A. Kendziora, *et al.*, "Advances in standoff detection of trace explosives by infrared photo-thermal imaging," *Proceedings SPIE 7664, Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XV*, 2010.

- [30] F. Fuchs, *et al.*, "Infrared hyperspectral standoff detection of explosives," *Proceedings SPIE 8710, Chemical, Biological, Radiological, Nuclear, and Explosives (CBRNE) Sensing XIV*, 2013.
- [31] C. I. Chang, "Hyperspectral Imaging: Techniques for Spectral Detection and Classification," 2003.
- [32] J. Li, *et al.*, "Semisupervised hyperspectral image classification using soft sparse multinomial logistic regression," *Geoscience and Remote Sensing Letters, IEEE*, vol. 10, no. 2, pp. 318-322, 2013.
- [33] M. M. Crawford, *et al.*, "Active learning: Any value for classification of remotely sensed data?" *Proceedings of the IEEE*, vol. 101, no. 3, pp. 593-608, 2013.
- [34] L. Bruzzone, *et al.*, "A novel transductive SVM for semisupervised classification of remote-sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 11, pp. 3363-3373, 2006.
- [35] J. Li, *et al.*, "Generalized composite kernel framework for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 9, pp. 4816-4829, 2013.
- [36] D. Tuia, *et al.*, "Active learning methods for remote sensing image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 7, pp. 2218-2232, 2009.
- [37] C. M. Bachmann, *et al.*, "Exploiting manifold geometry in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 441-454, 2005.
- [38] C. I. Chang, "Hyperspectral Data Processing: Algorithm Design and Analysis," John Wiley & Sons, 2013.
- [39] Q. Du and J. E. Fowler, "Hyperspectral image compression using jpeg2000 and principal component analysis," *Geoscience and Remote Sensing Letters, IEEE*, vol. 4, no. 2, pp. 201-205, 2007.
- [40] G. Licciardi, *et al.*, "Decision fusion for the classification of hyperspectral data: Outcome of the 2008 grs-s data fusion contest," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 11, pp. 3857-3865, 2009.
- [41] J. Li and J. M. B. Dias, "Minimum volume simplex analysis: a fast algorithm to unmix hyperspectral data," in *Geoscience and Remote Sensing Symposium, IGARSS 2008. IEEE International*, vol. 3, pp. III250-III253, 2008.
- [42] A. Plaza, *et al.*, "Recent advances in techniques for hyperspectral image processing," *Remote Sensing of Environment*, vol. 113, pp. S110-S122, 2009.
- [43] J. M. B. Dias, *et al.*, "Hyperspectral remote sensing data analysis and future challenges," *Geoscience and Remote Sensing Magazine, IEEE*, vol. 1, no. 2, pp. 6-36, 2013.
- [44] Q. Du and H. Yang, "Similarity-Based Unsupervised Band Selection for Hyperspectral Image Analysis," in *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 4, pp. 564-568, 2008.
- [45] M. Yang, *et al.*, "Monogenic Binary Coding: An Efficient Local Feature Extraction Approach to Face Recognition," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 6, pp. 1738-1751, 2012.
- [46] P. M. Arsal, *et al.*, "Neural Network and Linear Regression methods for prediction of students' academic achievement," *2014 IEEE Global Engineering Education Conference (EDUCON), Istanbul*, pp. 916-921, 2014.
- [47] X. Hu and J. Wang, "A Recurrent Neural Network for Solving Nonconvex Optimization Problems," *The 2006 IEEE International Joint Conference on Neural Network Proceedings, Vancouver, BC*, pp. 4522-4528, 2006.
- [48] S. Zhang and A. G. Constantinides, "Lagrange programming neural networks," in *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 39, no. 7, pp. 441-452, 1992.
- [49] D. Serre, "Matrices: Theory and Applications," New York, Springer-Verlag, 2002.
- [50] C. R. Rao and S. K. Mitra, "Generalized Inverse of Matrices and its Applications," New York, Wiley, 1971.

BIOGRAPHIES OF AUTHORS



Puttaswamy M. R. received the BSc degree from the University of Mysore, India, in 2001 and the MCA degree in Computer Applications from Visvesvaraya Technological University, India in 2005 and the M.Tech in Information Technology from Karnataka State Open University, India in 2010 and the PgCert in Higher Education Professional Practice from Coventry University, UK in 2014. He is currently working toward Ph.D degree in the Research and Development Center, Bharthiyar University, India. Puttaswamy is a Fellow of Higher Education Academy (UK) and also he is a Senior Fellow of Higher Education Academy (UK). He is currently member of the Society of Photographic Instrumentation Engineers (SPIE) and also member of IEEE. His current research interests include image processing, hyperspectral Image Classification/ Extraction and automatic target recognition.



Dr. P. Balamurugan is currently an Assistant Professor in the Department of Computer Science Govt. Arts College, Coimbatore, India where teaches digital image processing. His research interests are Digital Image Processing and Pattern recognition. He is the author (or coauthor) of reputed journal papers, several book chapters. He is a referee of several international journals and has served on the Scientific Committees of several national and international conferences. He is a member of the IEEE.