Hyperspectral Image Content Identification Using Kernel Based Neural Network

Puttaswamy M R, Balamurugan P
Research Scholar, Research and Development Centre, Bharthiyar University, India
Asst. Professor, Computer Science Department, Govt. Arts College, India

***Abstract*—**Large dimension data of Hyperspectral Image (HSI) leads to high computation cost, more execution time and increase the memory demand; therefore, difficulty arises during the classification of HSI. Unsupervised-BSA (band selection algorithm) using linear projection (LP) dependent band metric similarity has considered for informatics band selection of HSI. However, space complexity and time constrain is very big challenge for many feature algorithm. In this paper, we will use monogenetic binary feature (MBF) to get the local and monogenic feature from each pixel. The extracted feature from the combined monogenetic parameter will be used for effective classification process. In post processing, the classification of extracted feature from the MBF is perform by the Neural Network (NN) and to provide an enhanced generalization ability we introduce a kernel based NN, and considering the unknown feature mapping. To validate the performance of proposed hyperspectral classification algorithm, we will compared with several state-of-the-art.

***Keywords*—**Hyperspectral Image (HSI), Linear Projection (LP), Neural Network (NN), Band Selection Algorithm (BSA) and Monogenetic Binary Feature (MBF)

I. Introduction

THE HSIs are very high-resolution images with the more number of continuous spectrum, which causes the large spectral information to keep and form a difficulty at processing of HSI [1]. Large dimension data of HSI leads to high computation cost, more execution time and increase the memory demand; therefore, difficulty arises during the classification of HSI. To reduce the dimension of HSI is the pre-processing step that apply significance role at the classification of image; both feature selection (FS) and the feature extraction (FE) are majorly used approach for the 'dimension-reduction'. FS is an approach of selecting appropriate feature subset for the construction of model; FE-Approach follow different technique from a FS-Approach, in this FE produces newly features from the original features function. However, FSA returns a feature subset and the selection technique also used in HSI domains (i.e., for many features).

The FS can be get through a band transformation method such as Band Selection Algorithm (BSA). This BSA separated in two categories; supervised-BSA has used in terms to preserve the preferred objective information and unsupervised BSA has used to find the best informative bands that without knowing of any objective info. There are many approaches has proposed for the BSA; for example, divergence [2] and canonical analysis [3] has performed for the selection of band, JM-distance [4] and the B-distance [5] has used for the selection function of prior bands. However, these techniques mainly aim towards obtaining important information of selected bands, also to provide better FS results than of unsupervised-BSA. Nevertheless, the obtain information was not enough to get reliability like unsupervised-BSA; therefore it is necessary to implement unsupervised-BSA. There has been many unsupervised-BSA has proposed by researchers, information-theory-based BS has found in [6] [7], distance-based-measurement has investigated in paper [8], noise-adjusted PCA (principal component analysis) were adopted for the unsupervised BS in [3] and, several methods were compared by [9]. Du et al. [10] proposed an unsupervised BS-algorithm thru linear prediction based similarity metric for band, they shown in result section that their approach provided a better result w.r.t several other algorithms in class-reparability and the information conservation. Feature selection is an operational methodology to reduce the input dimensionality of features thru preserving several important information for the HSI classification that will be advantageous to decrease 'computational complexity' and reducing the redundancy [11] [12]. There are some popular feature extraction techniques are available such as, minimum-noise fraction (MNF [13]), principal component analysis (PCA [14]), sparsity preserving projection (SPP [15]), local preserving projection (LPP [16]), multi-structure manifold embedding (MSME [17]), and sparsity preserving analysis (SPA [18]), etc. However, several important information for the HSI classification has cannot obtain properly, therefore the classification performance degrade.

In this paper, unsupervised-BSA using linear projection dependent band metric similarity has considered for band selection, this method not demand any class prior information to classify the data analysis phase. Afterwards, we presents an efficient and multiresolution approach i.e., monogenetic binary feature (MBF), monogenetic signal representation can be given by three operational component; phase, amplitude and orientation. MBF can be used as individually 'or' combined to optimize the feature performance. In post

processing, the classification of extracted feature from the MBF is perform by the Neural Network (NN), which contain a hidden layer and a output layer, assignment of weight in between hidden layer and NN input can be done randomly. NN-classifier has applied to get the probability-classification output through the obtained extracted features; the selection of NN tends to effective computation in less time and gives better-classified results. To provide an enhanced generalization ability we introduce a kernel based NN, and considering the unknown feature mapping. The cross-validation that based on the available training-samples has used to adjust the kernel of NN and the weights of output layer can be estimate thru Linear Regression Analysis (LRA). To validate the performance of proposed hyperspectral classification algorithm, we will compared with several state-of-the-art.

Rest of the paper has organized as follows; section-II represents the several state-of-the-art, section-III provides the knowledge of our proposed methodologies, section-IV provides the brief knowledge of selected dataset and accuracy of classification process and, afterwards we makes some conclusion-remarks.

.

II. Literature Survey

HSI leads to high computation cost, more execution time and increase the memory demand; therefore, difficulty arises during the classification of HSI. To reduce the dimension of HSI is the pre-processing step that apply significance role at the classification of image. Due to the benefits of in-depth knowledge, in this paper [19], a standardized in-depth feature extraction (FE) method has obtained for (hyperspectral image) HSI classification with the help of a convolutional-neural-network (CNN). The method which is introduced by employs few of the pooling layers and the convolutional is to abstract in-depth features from HSIs, wherein they are invariant, discriminant and nonlinear. These few important features are beneficial for target detection and classification of an image. Moreover, to represent the basic problems of disproportion between less availability and higher dimensionality of training sections for the ordering or grouping of the HSI, to overcome the overfitting in modelling of the data class, the several approaches such as dropout and the L2-regularization has examined. More significantly, we introduce a 3-D CNN-based model known as feature extraction including the regularization, which is been combined such that from hyperspectral-imagery an essential spectral spatial features can be extracted.

In paper [20], a NILE (near-isometric-linear-embedding's) method is been introduced so that low-dimensional features can be extracted from HSI (hyperspectral-imagery). In the data of HIS, the near-isometric-linear-embedding's tries to conserve local nearest neighbour's isometric ally for all the different points with the linear deterministic projection matrix, which has fewer rows. For the construction of the matrix the problems is similarly transformed as a problem called as affine rank minimization where in it stores the norms of all difference spectral vectors among the pairwise pixels data of the HIS(hyperspectral-imagery). In convex program to reduce the problem of rank minimization, a scheme of nuclear norm minimization has introduced and used. To resolve the major problems alternating direction method of multipliers (ADMM) framework is been implemented and as per the requirement the projection matrix is acquired.

In paper [21], introduce an improved version of volumetric directional pattern such it can be used proficiently for extracting information with high spatial context in HSI (hyperspectral imagery). In the inputs of HSI, the newly introduced method combines the texture information from 3 sequential bands. For assigning the object category the local image texture features has extracted for each pixel of attention are then provided into an extreme-learning classifier. The information, which is Spatial, has presented major support for classification of HSI (hyperspectral image). For abstracting the features of spatial texture, the Local Binary Pattern (LBP) [22] can be used, although it is unable to capture the features of the structural and textural of images at different resolutions. Therefore, to achieve the better features of spatial from HSI (hyperspectral imagery) a scheme called as multiscale is presented on CLBP (Complete LBP), also on the LBP. On two different HSI data sets the experiments where performed which proved that the scheme which is introduced can increase the accuracy in classification of both CLBP and LBP.

In paper [23], because of the redundant spectral information and high dimensionality in a HSI (hyperspectral image), the LDA (linear discriminant analysis) and PCA (principal component analysis) has normally used for the extraction of features. By transforming LDA (linear discriminant analysis) and PCA (principal component analysis) to the problems of the regression and on the regression coefficients the I_1-norm constraint are imposed, for the better and improved version for extraction of features a SDA (sparse discriminant analysis) and SPCA (sparse principal component analysis) is been developed. This paper [24], Deep-learning based has presented the better and favourable performance for the classification of the HSI (hyperspectral image), due to their ability of abstracting deep features from HIS. Although, a high number of training samples are required in these types of techniques. Normally, to deliver illustrative feature expression for HSI (hyperspectral image) it is quite hard for deep-learning model when there is a limited number of samples. In this document a RGF (rolling guidance filter), a basic and easy deep-learning model novel and R-VCANet (vertex component analysis network) is been introduced, which provides better accuracy whenever the amount of training samples is not available in large quantities.

III. Proposed Hyperspectral classification Algorithm

However, the HSI contains a very high spectral correlation, due to that; the dimension reduction by spectral feature is an important pre-processing parameter to get a compact set but

useful spectral features. The reduction in dimensionality can be get through a band transformation method such as Band Selection Algorithm (BSA). This BSA separated in two categories; supervised-BSA has used in terms to preserve the preferred objective information and unsupervised BSA has used to find the best informative bands that without knowing of any objective info. In this paper, unsupervised-BSA using linear projection (LP) dependent band metric similarity has considered, this method not demand any class prior information to classify the data analysis phase. In LP-BSA [10], the two bands τ_1 and τ_2 has chosen; there selected subset can be given thru,

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

In order to obtain the third band τ_3, the subset of selected band should be update as

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

The equation 2 has to repeat until the τ_3 bands in $\emptyset$ is become more enough and the current subset of band contains τ_1 and τ_2 to select the 3rd band, which is unlike of bands τ_1 and τ_2, considering the τ band can be predicted as;

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

The LP of band τ can be denote by τ' through bands τ_1 and τ_2, the minimizing parameter t_0, t_1 and t_2 can minimize the LPE (i.e. Linear Prediction Error: $\mathbb{E}_{LP} = \|\tau - \tau'\|$). The minimal square solution $t = (t_0 \times t_1 \times t_2)^Y$ can be given as;

$$t = (1/(\mathbb{x}^Y \times \mathbb{x})) \times \mathbb{x}^Y \times \delta \quad (4)$$

Where, the matrix form of input data has denoted thru $\mathbb{x}$, which is in $V \times 3$ matrix and the 1st column include one value, 2nd-column having all pixels in band τ_1 and, the 3rd-column include all pixels in band τ_2 and δ represents the $V \times 1$ vector matrix including all pixels in τ. The most unlike band compared to τ_1 and τ_2 has maximize the error and it will chose as the band τ_3 for the $\emptyset$.

In order to obtain the texture analysis of HSI the MBF [25] is come in to picture and, monogenetic signal representation can be given by three operational component; phase, amplitude and orientation. The MBF can be compute through; MBF-P, MBF-A and MBF-O, this are the code words of phase, amplitude and orientation. This operator can be form through combining local variation parameters and local image-intensity parameters given thru [25];

$$\mathrm{MBFP}(A_b) = \left[B_o^t(A_b), B_f^t(A_b), B_p^t(A_b),\right]_{Bt} \quad (5)$$

$$\mathrm{MBFA}(A_b) = \left[B_o^t(A_b), B_f^t(A_b), B_A^t(A_b),\right]_{Bt} \quad (6)$$

$$\mathrm{MBFO}(A_b) = \left[B_o^t(A_b), B_f^t(A_b), B_O^t(A_b),\right]_{Bt} \quad (7)$$

The number of pattern in $\mathrm{MBF-X}$ can be given by

$$\mathrm{MBF-X} \in \{\mathrm{P}, \mathrm{A}, \mathrm{O}\} \quad (8)$$

The imaginary parameter of monogenetic signal can be represents at A_b pixel that encoded in 2-bits $(B_o'(A_b), B_f'(A_b))_{Bt}$. This above three MBF can be used as individually feature for further classification, also they can be used combined to optimize the feature performance.

Here, all local-histograms at several regions and scale will concentrate as a single vector of histogram to present the enhanced HSI image and the histogram of MBF-P, MBF-A and MBF-O can be given as;

$$F_h = fm_t(x,y) \quad (9)$$

$$\begin{cases} x = 1,..l \\ y = 1,2,3 \\ h = \{\mathrm{MBF-P}, \mathrm{MBF-A}, \mathrm{MBF-O}\} \end{cases} \quad (1)$$

Where, number of 'sub-region' at each scale given by l, h is feature map on a xth scale y and $fm_t(x,y)$ denotes the histogram vector of F_h. (2)

Here, the classification of extracted feature from the HBF is perform by the Neural Network (NN), which contain a hidden layer and a output layer, assignment of weight in between hidden layer and NN input can be done randomly. Moreover, the linear output layer weights can be calculated thru Linear Regression Analysis (i.e., LRA [26]) therefore, the computed cost will be much lesser than other NN-based approach. At given number of classes M, the labels of class has given through,

$$D_n \in (1 \le n \le D), \{1, -1\} \quad (10)$$

Where, D is a constructed row-vector and can be given as,

$$D = (D_1, \ldots, D_n, \ldots, D_M) \quad (11)$$

Equation (11) denotes the class number from that an individual sample belongs, as for example, if the 1st element of D_n is 1 and the other value is -1, then we can say that sample is belong from the nth classes. The training samples and their corresponding labels can be represent as;

$$\{C_t, D_t\}_{t=1}^k, \text{ Where } \begin{cases} C_t \in \forall^{vd} \\ D_t \in \forall^M \end{cases} \quad (12)$$

Moreover, with the H hidden nodes, the NN output function can be given through

$$fn_H(C_t) = \sum_{j=1}^{H} \varepsilon_j . NL_f(\varphi_j + C_t . \mu_j) = D_t, \quad \text{where } t = 1, \ldots, k \quad (13)$$

Where, μ_j denotes the 'weight-vector' that connecting the input nodes to jth hidden node, ε_j denotes the 'weight-vector' that connecting the output nodes to jth hidden node, NL_f denotes the function of non-linear activation and φ_j is the bias-vector of jth 'hidden-node'. Here, linear product of function C_t and μ_f has computed thru $C_t . \mu_j$ and some value (e.g., 1) is padded with C_t, then a dimension (1)

(1)

vector vd is ($vd+1$). Though the bias vector can be taken as the weight-vector element that can be assign randomly and above (13) equations for k can be given as,

$$D = \lambda \varepsilon \quad (14)$$

Where, λ represents the output matrix of hidden layer in neural network.

$$D = [D_1; D_2; D_3; \ldots\ldots; D_k] \in \forall^{k \times M} \quad (15)$$

$$\varepsilon = [\varepsilon_1; \varepsilon_2; \varepsilon_3; \ldots\ldots; \varepsilon_k] \in \forall^{H \times M} \quad (16)$$

The hidden-layer output matrix can be written as,

$$\lambda = \begin{bmatrix} \tilde{\lambda}(C_1) \\ \vdots \\ \tilde{\lambda}(C_k) \end{bmatrix} = \quad (17)$$

Subm

$$\begin{bmatrix} \tilde{\lambda}(\varphi_1 + C_1.\mu_1) \ldots\ldots & \tilde{\lambda}(\varphi_H + C_1.\mu_H) \\ \vdots & \vdots \\ \tilde{\lambda}(\varphi_1 + C_k.\mu_1) \ldots\ldots & \tilde{\lambda}(\varphi_H + C_k.\mu_H) \end{bmatrix}$$

$$\tilde{\lambda}(C_i) = \quad (18)$$

$$[\tilde{\lambda}(\varphi_1 + C_i.\mu_i), \ldots\ldots, \tilde{\lambda}(\varphi_H + C_i.\mu_H)]$$

Where, $\tilde{\lambda}(C_i)$ is a matrix function of λ and equation (18) is hidden nodes output with respect to C_i input function that maps the from H-dimension hidden nodes to the vd-dimension vector. Generally, the hidden neurons number (H) is very lesser than training samples number (k), and equation (14) can be rewritten for the LRA as,

$$\varepsilon' = D\lambda^{Inv} \quad (19)$$

Where, λ^{Inv} is the inverse generalized function matrix λ [27] and it can be shown as,

$$\lambda^{Inv} = \left(1/_{\lambda\lambda^T}\right).(\lambda^T) \quad (20)$$

The above equation (20) can provide the generalization and stability to the NN, a +ve parameter (ψ^{-1}) is added to individual diagonal element (i.e., $\lambda\lambda^T$), therefore, the NN classification output is given thru,

$$fn_H(C_i) = \quad (21)$$

$$D.\tilde{\lambda}(C_i).\lambda^T\left(\frac{1}{I.\psi^{-1} + \lambda\lambda^T}\right) = \tilde{\lambda}(C_i).\varepsilon$$

In equation (21), the function of feature mapping $\tilde{\lambda}(C_i)$ is knowledgeable to us. Here to provide an enhanced generalization ability we introduce a kernel based NN [28], and considering the unknown feature mapping, the K-NN matrix can be written as,

$$N_{KM} = \lambda(\lambda^{Inv}) \quad (22)$$

$$N_{KM(i,j)} = [\tilde{\lambda}(C_i).\tilde{\lambda}(C_j)] \quad (23)$$

Where, $[\tilde{\lambda}(C_i).\tilde{\lambda}(C_j)]$ can be write as $N(C_i, C_j)$ and, K-NN output function can be written as,

$$fn_H(C_i) = D.\begin{bmatrix} N(C_i, C_1) \\ \vdots \\ N(C_i, C_k) \end{bmatrix}^T \frac{1}{\psi^{-1}I + N_{KM}} \quad (24)$$

The label of input data can be determine accordance to output node-indexing w.r.t the higher values. Here, we consider the cross-validation that has based on the available training-samples to adjust the kernel of NN.

IV. Results and Analysis

In this section, we will compare the performance of classification with the help of pre-processing MBF proposed algorithm, which will compared with the several popular feature extraction techniques are available such as, MNF [13], PCA [14], SPP [15], LPP [16], MSME [17], and SPA [18]. Our proposed algorithm uses the real available HSI dataset such as Salinas-scene [29] and Pavia University [29]. The Salinas-scene was collected through the AVIRIS sensor (224-bands) over the Salinas valley (at California), also it characterized thru the high-level spatial resolution (i.e., 3.7-meter-pixels), this HSI data available for a sensor-radiance data, which includes bare soils, vineyard fields, and vegetables. Salinas scene ground truth comprises '16-classes'. A Pavia University HSI data has acquired thru a ROSIS sensor (during flight campaign at Pavia in Northern Italy), 103 spectral bands with 610*610 pixels and their geometric resolution is about 1.3 m. Pavia University ground truth comprises '9-classes'. The ground-truth comprises 16-classes in Salinas dataset, 54129 total sample and, per class total samples has given in Table 1(a). In Pavia university dataset, ground-truth comprises 16-classes, 54129 total sample and, per class total samples has given in Table 1(b). The overall accuracy (OA), kappa coefficient of the agreement (KA) has been used to acquire the classification accuracy. All experiments are perform using MATLAB 2016b on an Intel i5 3.00-GHz machine with the 8 GB RAM.

Table1: Total number of samples for each class
1(a) Sixteen classes of Salinas Scene dataset [29]

Classes	Total Samples
C1-Brocoli_green_weeds_1	2009
C2-Brocoli_green_weeds_2	3726
C3-Fallow	1976
C4-Rough_Plow_Fallow	1394
C5-Smooth-Fallow	2678
C6-Stubble	3959
C7-Celery	3579
C8-Grapes_untrained	11271

C9-Soil_vinyard_develop	6203
C10-Corn_senesced_green_weeds	3278
C11-Lettuce_romaine_4wk	1068
C12-Lettuce_romaine_5wk	1927
C13-Lettuce_romaine_6wk	916
C14-Lettuce_romaine_7wk	1070
C15-Vinyard_untrained	7268
C16-Vinyard_vertical_trellis	1807
Salinas Scene Total samples	54129

1(b) Sixteen classes of Pavia university dataset [29]

Classes	Total Samples
C1-Asphalt	6631
C2-Meadows	18649
C2-Gravel	2099
C2-Trees	3064
C2-Metal sheets	1345
C2-Bare Soil	5029
C2-Bitumen	1330
C2-Self-Blocking-Bricks	3682
C2-Shadows	947
Pavia U Total Samples	42776

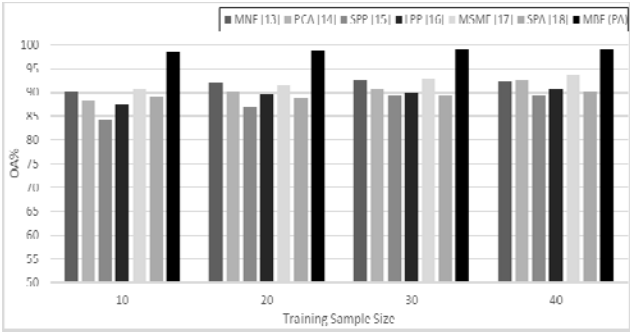

Figure 1: Classification results at different training-sample-sizes (10 to 40) on the SalinasS-dataset

Figure 1 shows the Classification results at different training-sample-sizes on the SalinasS-dataset, which varies 10 to 40 samples size. Our proposed model compares with the several existing technique such as, MNF [13], PCA [14], SPP [15], LPP [16], MSME [17], and SPA [18]. Considering 10 sampling size (SS), proposed algorithm MBF performs 8.59%, 10.4%, 14.37%, 11.3%, 8.2%, 9.7% better overall accuracy than the existing feature algorithm MNF [13], PCA [14], SPP [15], LPP [16], MSME [17], and SPA [18]. Now at 20 training-SS, we got 6.9%, 8.8%, 11.8%, 9.3%, 7.38%, 10.12% improved overall accuracy than the corresponding feature algorithm. Moreover, we computed for the 30 and 40 training-SS, where we can see the considerable improvement at MBF algorithm. At 40 sampling size, MBF performs 6.6%, 6.4%, 9.56%, 8.34%, 5.3%, 8.85% better overall accuracy than the considered feature algorithm.

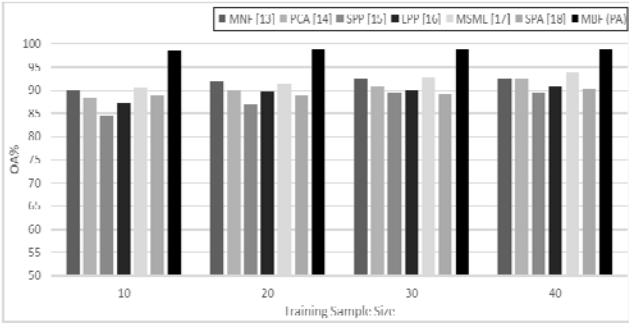

Figure 1: Classification results at different training-sample-sizes (10 to 40) on the PaviaU-dataset

Figure 2 shows the Classification results at different training-sample-sizes on the PaviaU-dataset, which also varies 10 to 40 samples size. Considering 40 sampling size (SS), proposed algorithm MBF performs 9.26%, 1.49%, 14.4%, 11.9%, 1.49%, 7.98% better overall accuracy than the existing feature algorithm MNF [13], PCA [14], SPP [15], LPP [16], MSME [17], and SPA [18].

At different samples 10, 20, 30 and 40, the proposed algorithm got 93.24, 93.6, 93.84 and 93.9 percent of OA in considered Pavia-U-Dataset, similarly while considering Salinas-S-Dataset, the obtained OA are 98.57%, 98.8%, 98.93% and 98.96% (i.e., 10, 20, 30 and 40 SS).

Table2: Classification results of several methodologies on the SalinasS –dataset (%).

Class No.	MNF [13]	PCA [14]	SPP [15]	LPP [16]	MSME [17]	SPA [18]	MBF (PA)
1	99.8	99.3	98.9	99.4	99.9	88.7	99.74
2	100	99.6	100	100	100	98.6	99.89
3	100	99.3	97.2	100	100	98.4	100
4	99.8	99.7	99.3	99.6	99.9	97.5	100
5	97.7	96.7	92.7	97.8	98.4	94.6	96.6
6	99.8	99.7	99.9	99.9	99.9	97	96.7
7	99.6	98.4	99.2	99.6	99.7	98.5	99.97
8	85	77.4	78.9	83.1	86.3	75.2	98.73
9	99.4	98.1	96.2	99.6	99.9	96.4	100
10	95.6	91.1	91.6	93	97.3	82.8	98.17
11	99	98.1	96.9	98.6	99	93.8	100
12	100	100	99.9	100	99.9	92.6	99.67
13	98.2	99.5	98	98.2	100	87.4	99.5
14	95.3	93.1	93.9	94.6	98.5	96.8	98.97
15	74.8	81.9	69.1	66.8	80.5	70.4	98.6
16	99	98.7	98.5	98.9	97.6	96.3	100
OA(%)	**92.8**	**91.2**	**89.6**	**91.1**	**94.1**	**89.6**	**98.97**
KA(%)	**91.9**	**90.2**	**88.4**	**90.1**	**93.4**	**88.4**	**98.85**

The classification results of several methodologies on the SalinasS-dataset is given in above table 2, where our proposed algorithm MBF has compared with MNF [13], PCA [14], SPP [15], LPP [16], MSME [17], and SPA [18] on the basis of

individual class, overall accuracy (OA%) and KA (%). From the table 2, there are five classes, which got 100% accuracy through MBF algorithm (classes 3, 4, 9, 11 and 16). Through computed OA, MBF has compared with MNF [13], PCA [14], SPP [15], LPP [16], MSME [17], SPA [18] and we got 6.23%, 7.8%, 9.4%, 7.9%, 4.9%, 9.4% improvement in classification results. Moreover by computed KA, proposed model MBF has compared with MNF [13], PCA [14], SPP [15], LPP [16], MSME [17], SPA [18] and we got 7.03%, 8.75%, 10.57%, 8.8%, 5.5%, 10.57% improvement in the classification results.

Table3: Classification results of several methodologies on the PaviaU –dataset (%).

Class No.	MNF [13]	PCA [14]	SPP [15]	LPP [16]	MSME [17]	SPA [18]	MBF (PA)
1	87.3	85.5	76.8	84.1	92.3	69.2	90.47
2	86.7	93.2	76.7	81.3	94.4	82.4	99.58
3	84.3	89	79.6	80.8	78.1	71.1	93.88
4	92.7	93.4	93.2	92.6	94	88.7	66.68
5	100	99.8	99.8	100	100	92.5	90.8
6	92.8	91.7	83.9	92.5	93.9	78.7	99.9
7	97.4	97.5	87.1	96.9	96.9	82.5	99.8
8	91.9	93.9	77.9	89.5	94.3	65.5	96.28
9	99.9	100	100	100	99.9	96.5	48.97
OA(%)	**89.3**	**92.2**	**80.5**	**86**	**93.6**	**79**	**93.9**
KA(%)	**86**	**89.7**	**75**	**81.9**	**91.5**	**72.8**	**91.89**

Table 3 shows the classification results of several methodologies on the PaviaU-dataset. Through computed OA, MBF has compared with MNF [13], PCA [14], SPP [15], LPP [16], MSME [17], SPA [18] and we got 4.89%, 1.8%, 14.2%, 8.4%, 0.3%, 15.8% improvement in classification results. Moreover by computed KA, proposed model MBF has compared with MNF [13], PCA [14], SPP [15], LPP [16], MSME [17], SPA [18] and we got 6.4%, 2.38%, 18.38%, 10.8%, 0.42%, 20.7% improvement in the classification results. From the above result analysis we can say that the MBF model has perform very well in every sampling size and as per increasing in training-sample-size the overall accuracy and KA is also increasing.

V. Conclusion

In this paper, framework based on the MBF has proposed after the U-BSA pre-processing step of band selection and the extracted feature used for further classification process. The monogenetic signal representation from MBF algorithm can be given by three operational component; phase, amplitude and orientation and that can be used as individually ‘or’ combined to optimize the feature performance. Moreover, in classification process, the considered cross-validation that based on the available training-samples has used to adjust the kernel of NN. The proposed algorithm MBF has compared with the several popular feature extraction techniques, which has shown in result section and from the analysis, we can say that the MBF model has perform very well in every considered situations. In Salinas-S and Pavia-U dataset, the obtained average value of OA w.r.t considered feature process is 7.64% and 7.59% through our proposed MBF algorithm. In the future work, more sophisticated classification methodologies will be explore.

References

[1] M. He, F. M. Imran, B. Belkacem and S. Mei, "Improving hyperspectral image classification accuracy using Iterative SVM with spatial-spectral information," 2013 IEEE China Summit and International Conference on Signal and Information Processing, Beijing, 2013, pp. 471-475.

[2] R. Huang and M. He, “Band selection based on feature weighting for classification of hyperspectral imagery,” IEEE Geosci. Remote Sens. Lett., vol. 2, no. 2, pp. 156–159, Apr. 2005.

[3] C.-I Chang, Q. Du, T.-L. Sun, and M. L. G. Althouse, “A joint band prioritization and band decorrelation approach to band selection for hyperspectral image classification,” IEEE Trans. Geosci. Remote Sens., vol. 37, no. 6, pp. 2631–2641, Jun. 1999.

[4] Ifarraguerri, “Visual method for spectral band selection,” IEEE Geosci. Remote Sens. Lett., vol. 1, no. 2, pp. 101–106, Apr. 2004.

[5] S. D. Backer, P. Kempeneers,W. Debruyn, and P. Scheunders, “A band selection technique for spectral classification,” IEEE Geosci. Remote Sens. Lett., vol. 2, no. 3, pp. 319–323, Jul. 2005.

[6] C. Conese and F. Maselli, “Selection of optimum bands from TM scenes through mutual information analysis,” ISPRS J. Photogramm. Remote Sens., vol. 48, no. 3, pp. 2–11, 1993.

[7] S. S. Shen and E. M. Bassett, “Information theory based band selection and utility evaluation for reflective spectray systems,” Proc. SPIE, vol. 4725, pp. 18–29, 2002.

[8] N. Keshava, “Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries,” IEEE Trans. Geosci. Remote Sens., vol. 42, no. 7, pp. 1552–1565, Jul. 2004.

[9] P. Bajcsy and P. Groves, “Methodology for hyperspectral band selection,” Photogramm. Eng. Remote Sens., vol. 70, no. 7, pp. 793–802, 2004.

[10] 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, Oct. 2008. doi:10.1109/LGRS.2008.2000619

[11] F. Feng, W. Li, Q. Du, and B. Zhang, ``Dimensionality reduction of hyperspectral image with graph-based discriminant analysis considering spectral similarity," Remote Sens., vol. 9, no. 4, p. 323, Mar. 2017.

[12] B. Du, W. Xiong, J. Wu, L. Zhang, L. Zhang, and D. Tao, ``Stacked convolutional denoising auto-encoders for feature representation," IEEE Trans. Cybern., vol. 47, no. 4, pp. 1017_1027, Apr. 2017.

[13] G. Lixin, X. Weixin, and P. Jihong, ``Segmented minimum noise fraction transformation for ef_cient feature extraction of hyperspectral images," Pattern Recognit., vol. 48, no. 10, pp. 3216_3226, Mar. 2015.

[14] X. M. Cheng, Y. R. Chen, Y. Tao, C. Y. Wang, M. S. Kim, and A. M. Lefcourt, ``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, Jul. 2004.

[15] L. Qiao, S. Chen, and X. Tan, ``Sparsity preserving projections with applications to face recognition," Pattern Recognit., vol. 43, no. 1, pp. 331_341, 2010.

[16] Z. Wang and B. He, ``Locality perserving projections algorithm for hyperspectral image dimensionality reduction," in Proc. 19th Int. Conf. Geoinform., Shanghai, China, Jun. 2011, pp. 1_4.

[17] Y. Gan, F. Luo, J. Liu, B. Lei, T. Zhang and K. Liu, "Feature Extraction Based Multi-Structure Manifold Embedding for Hyperspectral Remote Sensing Image Classification," in IEEE Access, vol. 5, pp. 25069-25080, 2017.

[18] F. Luo, H. Huang, J. Liu, and Z. Ma, ``Fusion of graph embedding and sparse representation for feature extraction and classi_cation of

hyperspectral imagery," Photogram. Eng. Remote Sens., vol. 83, no. 1, pp. 37_46, Jan. 2017.
[19] Y. Chen, H. Jiang, C. Li, X. Jia and P. Ghamisi, "Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks," in IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 10, pp. 6232-6251, Oct. 2016.
[20] W. Sun, L. Zhang and B. Du, "Feature extraction using near-isometric linear embeddings for hyperspectral imagery classification," 2016 8th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Los Angeles, CA, 2016, pp. 1-4.
[21] Essa, P. Sidike and V. Asari, "Volumetric Directional Pattern for Spatial Feature Extraction in Hyperspectral Imagery," in IEEE Geoscience and Remote Sensing Letters, vol. 14, no. 7, pp. 1056-1060, July 2017.
[22] P. Sidike, C. Chen, V. Asari, Y. Xu and W. Li, "Classification of hyperspectral image using multiscale spatial texture features," 2016 8th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Los Angeles, CA, 2016, pp. 1-4.
[23] L. Wang, X. Xie, W. Li, Q. Du and G. Li, "Sparse feature extraction for hyperspectral image classification," 2015 IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP), Chengdu, 2015, pp. 1067-1070.
[24] B. Pan, Z. Shi and X. Xu, "R-VCANet: A New Deep-Learning-Based Hyperspectral Image Classification Method," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 10, no. 5, pp. 1975-1986, May 2017.
[25] M. Yang, L. Zhang, S. C. K. Shiu and D. Zhang, "Monogenic Binary Coding: An Efficient Local Feature Extraction Approach to Face Recognition," in IEEE Transactions on Information Forensics and Security, vol. 7, no. 6, pp. 1738-1751, Dec. 2012.
[26] P. Mohd Arsad, N. Buniyamin and J. l. Ab Manan, "Neural Network and Linear Regression methods for prediction of students' academic achievement," 2014 IEEE Global Engineering Education Conference (EDUCON), Istanbul, 2014, pp. 916-921.
[27] Rao, C. Radhakrishna and Mitra, Sujit Kumar, “Generalized inverse of a matrix and its applications” University of California Press, 1972, "https://projecteuclid.org/euclid.bsmsp/12005 14113",
[28] K. R. Muller, S. Mika, G. Ratsch, K. Tsuda and B. Scholkopf, "An introduction to kernel-based learning algorithms," in IEEE Transactions on Neural Networks, vol. 12, no. 2, pp. 181-201, Mar 2001.doi: 10.1109/72.914517
[29] http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes.

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 interest include image processing, hyperspectral imaging 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.