Hybrid Subspace Detection based on Spectral and Spatial Information for Effective Hyperspectral Image Classification

Sadia Zaman Mishu Assistant Professor Department of Computer Science & Engineering Rajshahi University of Engineering & Technology, Rajshahi, Bangladesh

ABSTRACT

Subspace detection of remote sensing hyperspectral image data cube has become an important area of research because of the challenges of dealing with high dimensional feature space for efficient identification of ground objects. Standard feature extraction method such as Principal Component Analysis (PCA) has several shortcomings as it depends solely on global variance of the data set generated ignoring the low variant components. In this paper these limitations are addressed and alternatively Folded-PCA (FPCA) is used for feature extraction. FPCA has some advantages over PCA as it utilizes both local and global structures of the image and requires comparatively less computational cost and memory. These properties make it suitable for feature extraction therefore our proposed method combines it with Quadratic Mutual Information (OMI) for the task of feature reduction. In this research, QMI is utilized as a means of feature selection over the new features generated from FPCA to obtain an informative subspace. The proposed method is named as (FPCA-QMI). It is tested on two hyperspectral datasets one is real mixed agricultural land and another one is an university area. Finally Kernel Support Vector Machine (KSVM) technique is applied to measure the classification accuracy of these two datasets. From the experimental analysis it is observed that the proposed method can detect effective subspace and obtains the highest accuracy of 98.0328% and 99.0431% on two real hyperspectral images which is better than the baseline approaches.

General Terms

Remote Sensing; Feature Extraction; Feature Selection; Principal Component Analysis; Classification

Keywords

Subspace; Folded Principal Component Analysis; Quadratic Mutual Information; Kernel Support Vector Machine

1. INTRODUCTION

Hyperspectral remote sensors capture hundreds of contiguous spectral bands in optical wavelength range [1]. These bands provide high spectral resultations that are used in many applications. several number of Among imaging spectrometers in remote sensing technology Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) and Reflective Optics Sysyetm Imaging Spectrometer (ROSIS) are commonly used imaging sensors. Nevertheless, these sensors provide high volume of data. For instance, AVIRIS sensor captures data in 220 different wavelengths of the bands with a spectral resolution of $0.01 \mu m$ that covers from visible light to (near) infrared region of electromagnetic wave spectrum to

Boshir Ahmed Professor Department of Computer Science & Engineering Rajshahi University of Engineering & Technology, Rajshahi, Bangladesh

denote narrow absorption features which provide strong discrimination capability in data inspection [2, 3]. This high dimensional data provides many research chanllenges as it requires more processing time. It also suffer from the curse of dimensionality problem which refers to that the classification accuracy increases with increase of the no of features and starts to decrese after a certain value of the no of bands. This effect occurs due to the imbalance of the ratio between the no of bands and the no of training samples which reduces the overall classification accuracy [2, 4]. Conversely, the close and contiguous bands are extremely correlated and some of them are redundant which are not important for classification [2]. These challanging research problems motivate us to extract relevant features with a view to improve classification accuracy. Different feature extraction and feature selection methos are available for feature reduction [5,6]. In current literature, Principal Component Analysis (PCA) is utilized for the task of feature extraction [1, 7]. But PCA has some limitations as it depends on global variance overlooking low variant components which may cause loss of information [1]. Again it requires large memory and more computational time for data prepocessing. As a result FPCA is introduced in [3] which has been theoretically proven to be able to reduce the size of the covariance matrices thus requires less memory and also gives better classification accuracy than PCA. Moreover FPCA extracts both global and local structure from the hypercube [8, 9]. So for all these specialities of FPCA, it is used for feature extraction in our proposed method.

But only FPCA is not enough to detect subspace of better features because it extracts features based on the high spectral variance but high variance doesn't always give guarantee of having most useful features. Low variance components may cover more spatial information which is important for classification. So feature selection is needed to rank the features having more spatial information generated from the feature extraction method. The mutual information(MI) is a widely used feature selection method. It is non parametric so it does not rely on the assumption about the shape of the distribution of the input parameters [2]. It has been widely used as a similarity measure for different types of imaging applications such as feature reduction [10], registration of image [11] and image clustering [12] etc. In this paper, an improved variant of MI called quadratic mutual information (QMI) is used for selecting features generated from FPCA. In [13] the QMI has been determined from the input samples and this QMI is useful for not only discrete but also for continuous features. In this paper, QMI is used over the FPCA images as a feature selection criterion inorder to maximize the relevance of the selected features. QMI is applied between FPCA

images and class labels. In this hybrid approach (FPCA-QMI), first FPCA is performed on input hypespectral image to generate the new features and then QMI is applied to rank the features having more spatial information generated from FPCA images. Finally the performance of the proposed method is evaluated using kernel support vector machine classifier for the two datasets.

2. RELATED WORKS

Combining the existing feature extraction and feature selection methods are extensively used at present to obtain a fusion approach which gives better result than only feature extraction or feature selection.

Hossain et al. in paper [13] have combined mutual information with PCA for feature extraction. They have achieved about 80% classification accuracy. But they have used an unsupervised approach which can be further improved if training samples are available. Here also mean channel has been used as the reference which may affect the selection.

Hossain et al. in paper [1] have combined PCA with normalized mutual information and achieved 96% accuracy. But hard thresholding has been utilized while assessing the input features.

In paper [2] PCA has been combined with QMI and 99% accuracy is obtained. But among 16 classes of AVIRIS dataset only 5 classes have been used which is not sufficient for analysis.

In all the current works, PCA is used for feature extraction. But PCA suffers from high computational cost and memory requirement and also rely on global variance disregarding the low variant components which may cause information loss. In this paper these limitations have been addressed and alternatively FPCA has been used which shows prominent result than PCA in terms of memory requirement, computational time and also dealing with both the local and global structure which makes our proposed method more robust than the existing benchmark approaches.

3. METHODOLOGY

3.1 Principal Component Analysis

Principal Component Analysis (PCA) is an unsupervised method that reduces the number of features by an orthogonal projection and truncation of the transformed data [14]. The new dataset is generated from the linear combination of the input features The resultant images are ordered depending on the high value of the variance [2,15]. Let X is an hyperspectral image and $X_n = [X_{n1}, X_{n2}, ..., X_{nF}]^T$ is the spectral vector of a pixel in the image. The rang of the values of n is from 1 to S, where S is number of pixels in the image. The mean-adjusted matrix M_n is generated by subtracting the average spectral vector \overline{I} from X_n as follows:

$$\mathbf{M}_{\mathbf{n}} = X_{\mathbf{n}} - \overline{I} \tag{1}$$

$$\bar{I} = \frac{1}{S} \sum_{n=1}^{S} X_n \tag{2}$$

Next, the covariance matrix **C** is calculated as follows:

$$\mathbf{C} = \frac{1}{S} \mathbf{M} \, \mathbf{M}^{\mathrm{T}} \tag{3}$$

The eigenvalues and eigenvectors are gained by the decomposition of ${\bf C}$ as

$$\mathbf{C} = \mathbf{V}\mathbf{E}\mathbf{V}^{\mathrm{T}} \tag{4}$$

Where, **E** is the diagonal matrix of eigenvalues of **C** and **V** is the orthonormal matrix of corresponding eigenvectors.

3.2 Folded Principal Component Analysis

If the spatial size S is very large, the calculation of covariance matrix is difficult using PCA due to memory management issue [7]. Furthermore, PCA be unsuccessful to catch the individual contribution of each of the F bands and considers all bands of hyperspectral image (HSI) equally in covariance matrix calculation [8]. Thus, FPCA is offered that decreases the covariance matrix size and extracts the local structure from the entire hypercube spectral domain effectually [9].

In implementation of FPCA [7] each mean-adjusted spectral vector M_n is transformed into a H×W matrix D_n which is defined as

$$\boldsymbol{D}_{\boldsymbol{n}} = [d_{n1}, d_{n2}, \dots, d_{nh}]^T$$
(5)

Finally, the overall covariance matrix for the whole dataset is obtained by accumulating all these partial matrices which is given by

$$\mathbf{D}_{\mathbf{F}} = \frac{1}{s} \sum_{n=1}^{S} \mathbf{C}_{\mathbf{n}} \tag{6}$$

The projection matrix is then computed after performing Eigen decomposition on D_F .

3.3 Mutual Information

Mutual Information is a measure that is used to calculate the similarity between two random variables which is called the mutual dependence [2]. The entropy is the amount of information that a random variable comprises and this is used to determine mutual information. According to Shannons information theory [16] the entropy of a discrete random variable A can be represented by the following equation

$$H(\mathbf{A}) = \sum_{a \in \mathbf{A}} p(a) \log p(a) \tag{7}$$

Where p(a) denotes the probability density function of A. So the mutual information I(A, B) of random variable A and random variable B is given by [2]

$$I(\boldsymbol{A},\boldsymbol{B}) = H(\boldsymbol{A}) + H(\boldsymbol{B}) - H(\boldsymbol{A},\boldsymbol{B})$$
(8)

H(A,B) indicates joint entropy of A and B and finally the mutual information between these two random variables can be defined by

$$I(\boldsymbol{A}, \boldsymbol{B}) = \sum_{b \in B} \sum_{a \in A} p(a, b) \log \frac{P(a, b)}{p(a)p(b)}$$
(9)

Where p(a) and p(b) indicates the marginal probability distribution function of random variable A and B respectively. If these two random variables do not comprise any information about each other, the value of MI will be zero and it will be maximum if all the information of A is shared by B and vice versa [2].

3.4 Quadratic Mutual Information

Quadratic Mutual Information (QMI) is the improved variant of mutual information based feature selection technique which is used to rank the feature subsets based on the information they provide about the output. The main perception of QMI is based on Cauchy–Schwarz divergence and quadratic Renyi entropy [2]. Renyi entropy is expressed by:

$$(\mathbf{A}) = \frac{1}{1-q} \log \sum_{i=1}^{S} p(A_i)^q$$
(10)

If the value of the parameter i intended to become 1 in the limit then it converges to Shannon entropy and if q=2 then it is termed as quadratic Renyi entropy [2, 17]. One of the

measures of mutual information with respect to Renyi entropy is known as the Cauchy–Schwarz divergence [2, 18] which is given by

$$D_{CS}(p_1, p_2) = 2R_{E_2}(\boldsymbol{A}; \boldsymbol{B}) - R_{E_2}(\boldsymbol{A}) - R_{E_2}(\boldsymbol{B}) \quad (11)$$

Where $R_{E_2}(\mathbf{A}), R_{E_2}(\mathbf{B})$ are quadratic Renyi entropy, $R_{E_2}(\mathbf{A}; \mathbf{B})$ represents the quadratic Renyi-cross entropy and finally the QMI [2] is defined by

$$I_{CS}(\boldsymbol{A}, \boldsymbol{B}) = D_{CS}(p_{12}(a, b), p_1(a)p_2(b))$$
(12)

Here the value of $I_{CS}(A, B)$ becomes zero if A and B are independent of each other and positive otherwise. So the features are selected based on the QMI values [2]. The features having highest QMI values are selected as informative features for the classification task.

4. PROPOSED METHOD

The proposed method mainly focuses on the hybrid approach for detecting the effective subspace of relevant features for better classification. It combines FPCA and QMI for feature reduction. The flowchart of the proposed method is given in figure-1





At first FPCA has been applied on the original data for the task of feature extraction and QMI has been calculated between class label data and generated FPCA features. Features are then ordered based on the QMI value and the features having QMI value greater than threshold 0.3 are selected as most informative features. Then these selected features are evaluated with kernel support vector machine classifier.

5. EXPERIMENTAL RESULT ANALYSIS

5.1 Dataset Description

The experiment of this work is assessed by two hyperspectral image datasets one is Indian Pine dataset and another one is Pavia University dataset. First one was collected over the agriculture area of Indian Pine which is located in the north part of Indiana [18]. The spectral resolution of this dataset is $0.4 \ \mu m$ to $2.5 \ \mu m$. It has 220 bands and 16 classes. Because of the inadequate training samples some classes were not used.

The second one was collected from the Pavia University, north Itally. It was captured by the ROSIS sensor [20]. It has a spectral resolution of 0.43 μm to 0.86 μm , 103 bands and 9 classes. Again due to insufficient training sample class "Shadows" was not used. Training and test samples of these two datasets are given in Table-1 and Table-2.

The training and test samples of these two datasets are collected based on the ground truth image.

Table 1. Training and Test Samples of Indian Pines Dataset

Indian Pines Dataset			
Class Name	Training Samples	Test Sample	
Hay-widrowed	165	135	
Soybeans-notill	109	85	
Woods	279	248	
Wheat	21	10	
Grass/Trees	15	14	
Soybeans-min	48	44	
Corn-notill	20	15	
Alfalfa	25	20	
Corn-min	108	72	
Corn	25	20	
Soybean-clean	15	15	

Table 2. Training and Test Samples of Pavia University Dataset

Pavia University Dataset			
Class Name	Training Samples	Test Samples	
Asphalt	19889	1826	
Meadows	5595	4730	
Gravel	630	520	
Trees	919	813	
Paint Metal Sheet	403	369	
Bare soil	1509	1406	
Bitumen	399	366	
Self block bricks	1105	936	

(%)

The folding options for FPCA for Indian Pines and Pavia University datasets are given in table 3 and 4. The folding options are chosen in such a way so that the multiplication of height L and width W must be equal to the no of bands. That's why the band 63, 102 and 163 are removed as noisy band for Pavia University dataset to make the no of bands even.

Fable 3. Folding	Options	for	Indian	Pines	Dataset
------------------	---------	-----	--------	-------	---------

Option	L	W
FPCA-1	2	110
FPCA-2	4	55
FPCA-3	5	44
FPCA-4	10	22
FPCA-5	11	20

Table 4. Folding Options for Pavia University Dataset

Option	L	W
FPCA-1	2	50
FPCA-2	4	25
FPCA-3	5	20

In this paper the dimension of the Indian Pines image has been reduced to 10 from 220 bands using FPCA because 10 folded principal components cover almost 100% cumulative variance which is shown in figure 2 for Indian Pines dataset.

Again the dimension of the Pavia University image has been reduced to 10 from 103 bands respectively using FPCA because 10 folded principal components cover almost 100% cumulative variance for this dataset which is shown in figure 3.

From the result of feature extraction using FPCA, it can be seen that image of principal components extracted by FPCA are sharper than the principal components extracted by PCA which is illustrated in figure 4



Fig 2: Cumulative Variance of Indian Pines dataset



Fig 4: Principal components extracted by FPCA are sharper than principal components extracted by PCA

From this figure it is revealed that the folded principal components (FPCs) extracted by FPCA are visibly clear and sharper than principal components (PCs) extracted by PCA. The reason of this is that PCA treats all bands equally considering global variance and ignoring low variant components. Conversely FPCA extracts the local structure from the hypercube by folding the spectral vectors and these components are less prone to noise than the PCA components. Same result is also achieved for Pavia University dataset. Another advantage of FPCA is that it reduces the size of the covariance matrix so it requires less computational time and memory compared to PCA. Let L and W be the height and width of the data matrix of FPCA and S is the spatial size of the dataset. The comparison of the memory requirement for PCA and FPCA is given in table-5.

Table 5. Comparison of Memory Requirement

Method	Data matrix size	Covariance matrix size
PCA	S x LW	LW x LW
FPCA	L x W	W x W

So for all these specialties of FPCA motivate us to use FPCA for feature extraction instead of PCA.

5.3 Result of Feature Selection



Fig 3: Cumulative Variance of Pavia University dataset

Higher variance does not give guarantee of consuming required information. Some features with lower variance may be further informative [19]. So only feature extraction is not adequate for informative subspace detection. As shown in figure 2 for Indian Pines dataset, FPC-11 contains more spatial information than FPC-10. But in FPCA, FPC-10 is ranked higher than FPC-11 based on variance. To address this issue, QMI between the class labels and FPCA principal components is calculated. The features are then ordered based on the value of QMI. The QMI value of first 20 folded principal components is shown in figure 4 for Indian Pines dataset.



Fig 5: QMI value of FPCs for Indian Pines Dataset

The value of threshold value T is set as 0.3. The features having QMI value greater than 0.3 are selected as most informative features. The top 10 ranked features for Indian Pines dataset is given in table 6.

Similarly for Pavia University dataset FPC-3 contains more spatial information than FPC-1 which is ranked higher in FPCA based on variance. So after applying QMI on FPCA image, the most spatially informative features are selected. The QMI value of the 1st 20 folded principal components for Pavia University dataset is shown in figure 6.

 Table 6. Top Ranked Features Selected by FPCA and PFCA-QMI for Indian Pines dataset

Method	Fold	Order Of The Selected Features
FPCA	_	FPC:1,2,3,4,5,6,7,8,9,10
FPCA-QMI1	2 x 110	FPC:2,3,7,1,11,5,13,8,4,16
FPCA-QMI2	4 x 55	FPC:1,5,2,3,6,11,4,15,13,7
FPCA-QMI3	5 x 44	FPC:1,7,6,12,3,4,11,2,5,15
FPCA-QMI4	10 x 22	FPC:1,3,7,11,6,5,13,16,2,9
FPCA-QMI5	11 x 20	FPC:2,1,3,14,6,11,4,7,10,5



Fig 6: QMI value of FPCs for Pavia University Dataset

In case of Pavia University dataset the threshold value is also set as 0.3 which is selected through the inspection. Top 7 ranked features having threshold greater than 0.3 are given in table 7.

Table 7. Top Ranked Features Selected by FPCA and PFCA-QMI for Pavia University dataset

Method	Fold	Order of the Selected Features
FPCA	-	FPC:1,2,3,4,5,6,7,8,9,10
FPCA-QMI1	2 x 50	FPC: 3,1,5,2,7,9,11
FPCA-QMI2	4 x 25	FPC: 2,3,1,5,10,7,6
FPCA-QMI3	5 x 20	FPC: 3,1,2,7,5,9,10

The subspace of these selected features are then assed to KSVM classifier for measuring classification accuracy.

5.4 Result of Classification

Kernel Support Vector Machine classifier has been used for the task of classification. To train the classifier RBF kernel is utilized. The parameters of kernel such as best cost parameter C=10 and kernel width γ =2.4 for Indian Pines and C=8 and γ =0.3 for Pavia University are defined by 10-fold cross validation. If the classification accuracy is measured without using any feature extraction or feature selection, then it gives 69.03% accuracy for Indian Pines dataset and 71.83% accuracy for Pavia University dataset which are very inadequate and it highly encourages us to apply feature reduction technique. The classification results of PCA, FPCA and FPCA-QMI with different folding options are given in Table 8 for Indian Pines dataset and Table 9 for Pavia University dataset.

The overall classification accuracies using different no of features are shown in figure 7 for Indian Pines dataset. From this figure it can be seen that FPCA-QMI performs better than PCA and FPCA in all cases for top ten ranked features. For Indian Pines dataset PCA gives 92.8211% accuracy which is lower than the other methods. Again the highest accuracy of FPCA is 95.8567% for FPCA-4 which is greater than PCA. But FPCA selects features based on the highest value of variance which does not give reassurance of having high informative features. That's why QMI is applied to these extracted features for selecting appropriate features with local characteristics competently and the accuracy achieved by FPCA-QMI4 using only 10 features is 98.0328% which is higher than the accuracy given by PCA and FPCA for all other folding options.

Table 8. Classification Accuracy for Indian Pines Dataset

Method	Accuracy
PCA	92.8211
FPCA-1	95.0321
FPCA-2	95.5194
FPCA-3	94.9804
FPCA-4	95.8567
FPCA-5	93.6390
FPCA-QMI1	97.9801
FPCA-QMI2	96.8575
FPCA-QMI3	95.9876
FPCA-QMI4	98.0328
FPCA-QMI5	94.4103

Table 9. Classification Accuracy for Pavia University Dataset

Method	Accuracy
PCA	93.0987
FPCA-1	96.5762
FPCA-2	96.3680
FPCA-3	97.5490
FPCA-QMI1	97.0356
FPCA-QMI2	98.5042
FPCA-QMI3	99.0431



Fig 7: Classification Accuracy of Indian Pines Dataset

Similarly for Pavia University dataset, from figure 8 it can be observed that the accuracy achieved by PCA is 93.0987% which is not satisfactory. Furthermore the combination of FPCA-QMI3 gives highest accuracy of 99.0431% with only 7 features ranked with the proposed method.



Fig 8: Classification Accuracy of Pavia University Dataset

So from all these analysis it is observed that our proposed method performs better in terms of memory requirement and classification accuracy than the methods studied in this paper.

6. CONCLUSION

An improved subspace detection technique has been proposed here to obtain better classification result which has been justified by using kernel SVM classifier. A comparative study has been presented in Table 8 and 9. It can be seen that the proposed method outperforms the standard baseline approaches for the task of memory requirement and classification accuracy. This is because the selected features are uncorrelated and able to form subspace of most informative features. By applying QMI as a feature selection criterion the proposed method becomes more effective as it can handle both linear and non-linear relationship between the classes. Finally the processing speed of feature selection based on QMI can be further improved by applying adaptive thresholding.

7. REFERENCES

- Hossain, M. A., X. Jia, and M. Pickering, 2014. "Subspace detection using a mutual information measure for hyperspectral image classification," IEEE Geoscience and Remote Sensing Letters, vol. 11, no. 2, pp. 424–428.
- [2] Mishu, S. Z., M. A. Hossain, and B. Ahmed. 2018. "Hybrid Sub-space Detection Technique for Effective Hyperspectral Image Classification." Paper presented at the IEEE 4th International Conference on Computer, Communication, Chemical, Material and Electronic Engineering, Bangladesh
- [3] Zabalza, J., J. Ren, M. Yang, Y. Zhang, J. Wang, S. Marshall, and J. Han. 2014. "Novel Folded-PCA for Improved Feature Extraction and Data Reduction with Hyperspectral Imaging and SAR in Remote Sensing." *ELSEVIER ISPRS*
- [4] Hughes, G. 1968. "On the mean accuracy of statistical pattern recognizers", *IEEE Transactions on Information Theory*, IT-14(1): 55-63. doi:10.1109/TIT.1968.1054102.
- [5] Jia, X., B. Kuo, and M. M. Crawford, 2013. "Feature mining for hyperspectral image classification," Proceedings of the IEEE Journals and Magazines, vol. 101, pp. 676-697,
- [6] Richards, J. A. and X. Jia, 2006. "Remote Sensing Digital Image Analysis," 4th Edition: Springer-Verlag Berlin Heidelberg, Germany
- [7] Deepa, P.,and K. Thilagavathi, 2015. "Feature Extraction of Hyperspectral Image Using Principal Component Analysis and Folded-principal Component Analysis", Paper presented at 2nd International Conference on Electronics and Communication Systems (ICECS): 656-660.
- [8] Uddin, M. P., M. A. Mamun, and M. A. Hossain, 2017. "Feature Extraction for Hyperspectral Image Classification", Paper presented at the IEEE 5th Region 10 Humanitarian Technology Conference (R10-HTC), Bangladesh.
- [9] Uddin, M. P., M. A. Mamun, and M. A. Hossain, 2017. "Segmented FPCA for hHyperspectral Image Classification", Paper presented at the IEEE 3rd International Conference on Electrical Information and Communication Technology (EICT), Bangladesh.
- [10] Torkkola, K, 2003. "Feature extraction by nonparametric mutual information maximization," Journal of Machine Learning Research, vol. 3, pp. 1415–1438.

- [11] Viola, P. and W. M. Wells III, 1997. "Alignment by maximization of mutual information," International Journal of Computer Vision, vol. 24, no. 2, pp. 137 – 154.
- [12] Vinh, N. X. and J. Bailey, 2010. "Information theoretic measures for clusterings comparison: variants, properties, normalization and correction for chance," Journal of Machine Learning Research, vol. 11, pp. 2837–2854.
- [13] Hossain, M. A., M. Pickering, X. Jia, 2011. "Unsupervised feature extraction based on mutual information measure for hyperspectral image classification", *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 978-1-4577-1005-6.
- [14] Deolindo, C. S., A. C. B. Kunicki, F. L. Brasil, and R. C. Moioli, 2014. "Limitations of Principal Component Analysis as a Method to Detect Neuronal Assemblies" Paper presented at IEEE 16th International Conference on e-health Networking, Applications and Services (Healthcom).
- [15] Rodarmel, C. and J. Shan, 2002."Principal component analysis for hyperspectral image classification". In: Surveying and Land Information Systems, 62.2, pp. 115– 122.
- [16] Shannon, C. E. and W. Weaver, 1949. "The mathematical theory of communication," Urbana, IL: University of Illinois Press.
- [17] Sluga, D. and U. Lotri, 2017. "Quadratic mutual information feature selection,"Journal of Entropy, 19, 157.
- [18] Sluga, D. and U. Lotric, 2013. "Generalized Informationtheoretic Measures for Feature Selection," In Proceedings of the International Conference on Adaptive and Natural Computing Algorithms, Lausanne, Switzerland, Springer: Berlin/Heidelberg, Germany, pp. 189–197. 4-6.
- [19] Hossain, M.A., H. E. Jannat, B. Ahmed, and M. A. Mamun,"Feature mining for effective subpace detection and classification of hyperspectral images," International Conference on Electrical, Computer and Communication Engineering (ECCE), Cox's Bazar, Bangladesh, 16-18 Februar, 2017.
- [20] Kavitha, K., S. Arivazhagan and B. Suriya, 2014, "Classification of Pavia University Hyperspectral Image using Gabor and SVM Classifier", International Journal of New Trends in Electronics and Communication, vol.2 issue. 3.