

# Dimensionality Reduction Techniques for Hyperspectral Images

Shraddha P. Lodha<sup>1</sup> and Prof. S. M. Kamlapur<sup>2</sup>

<sup>1</sup>Student, M.E. Computer Engineering, KKWIEER, Savitribai Phule University of Pune, India

<sup>2</sup>Department of Computer Engineering, KKWIEER, Savitribai Phule University of Pune, India

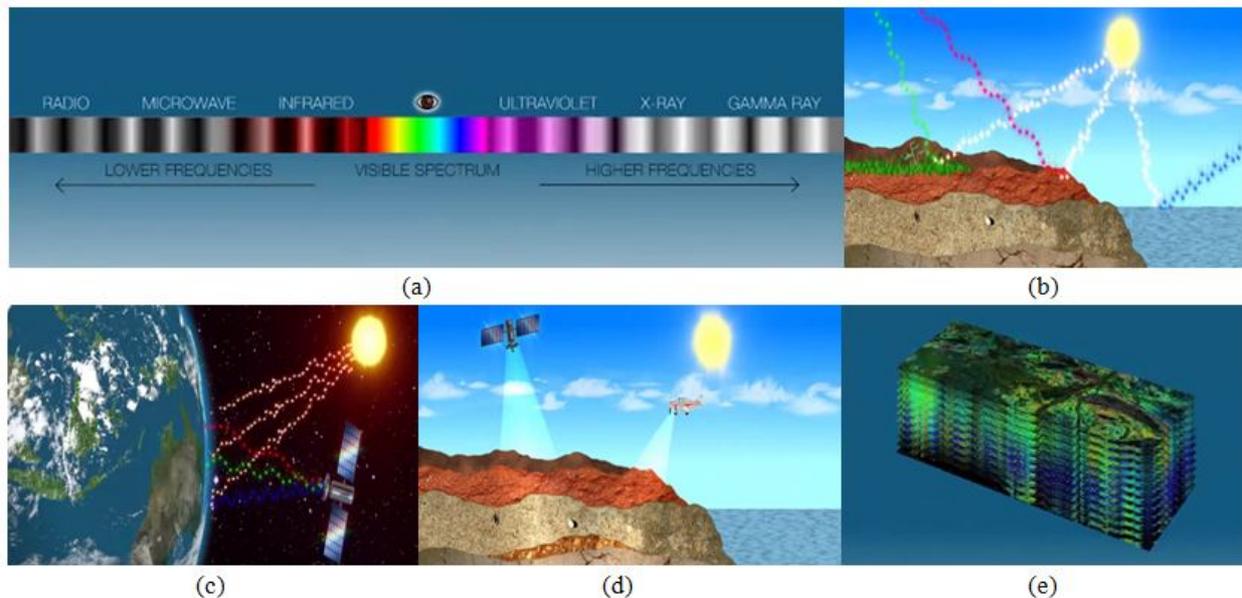
## ABSTRACT

*Hyperspectral Imaging produces an image where each pixel is having narrow spectral bands with plentiful spectral information. Spectral bands refer to the large number of measured wavelengths bands of Electromagnetic Spectrum. The large number of spectral bands in hyperspectral data increases the computational burden. So, dimensionality reduction through spectral feature selection thoroughly affects the accuracy of the given task. A fuzzy rough set is an approach that deals with the concepts of vagueness and indiscernibility. It finds feature subsets preserving the semantics of the given datasets. Therefore, this paper proposes the applicability of Fuzzy-Rough Set Approach to select the most significant spectral features from the hyperspectral data. Selected features are employed to build a more easy and understandable learning model in order to improve the classification quality of hyperspectral images.*

**Keywords:** Dimensionality Reduction, Fuzzy-Rough Sets, Hyperspectral Imaging, Spectral features

## 1. INTRODUCTION

Hyperspectral Imaging is a foremost research area in Remote Sensing. It combines the potentiality of digital imaging and spectroscopy. Hyperspectral imaging collects information across the Electromagnetic Spectrum. This technology is used largely in the Minerals, Coal, Oil and Gas Industry. Hyperspectral images provide researchers with a picture of the surface of a surveyed area so they can deduce some conclusions. Our eyes only see visible band of the light spectrum. Many more bands of non-visible light spectrum with higher and lower frequencies exist to the left and right of visible spectrum. This is known as the Electromagnetic Spectrum and it includes the bands of ultra-violet, x-rays, radio waves, micro waves and infra-red radiations as shown in fig. 1(a) [1]. Hyperspectral Imaging is a technology which divides regions of those spectra which we can't see into many more bands and creates visible images from them. Earth's surface is overlaid with a variety of materials like soil, rocks, foliage and water, all composed of diversifying minerals and chemical compounds. These materials all reflect only a chunk of the Electromagnetic Spectrum and therefore can be identified by their unique Electromagnetic Signatures, see fig. 1(b). Hyperspectral sensors scan the sunlight bouncing off the Earth's surface as shown in fig. 1(c). By analyzing great portion of the Electromagnetic Spectrum contained within that sunlight, the composition of the region can be determined by comparing against known spectral signature of different minerals and compounds. Hyperspectral sensors take an advantage of passive sensors and hence, only scan the available light bouncing off the surface of the surveyed region. The sensors are usually mounted on either airplanes or low orbit satellites that scan and record area of the Earth's surface, refer fig. 1(d). Results from hyperspectral scans are compiled by computer software into data cubes. An example of Hyperspectral data cube (3-D cube) which is further known as hyperspectral image data cube is shown in fig. 1(e) [1]. Thus, every pixel in the hyperspectral image consists of a contiguous reflectance spectrum. That means each pixel contains contiguous bands of electromagnetic spectrum from visible to infrared regions and can be used to identify the objects in the scene since it provides much more detailed information about the scene than the normal color camera. In conjunction with other imaging technologies, hyperspectral spectral imaging can yield much more detailed information about the scene or the surveyed area. Thus, hyperspectral imaging leads to an extremely enhanced ability to classify the objects in the scene based on their spectral properties [1]. However, the classification of hyperspectral image is a very challenging task in recent days due to the huge number of band values for each pixel. Hence, dimensionality reduction becomes a crucial step for successfully implementing hyperspectral image classification.



**Figure 1** The concept of hyperspectral imaging technology which consists of (a) the electromagnetic spectrum, (b) reflectance of a chunk of the electromagnetic spectrum from the materials on the Earth’s surface, (c) hyperspectral sensors which scan the sunlight bouncing off the Earth’s surface, (d) sensors mounted on airplanes or low orbit satellites and (e) hyperspectral image cube.

## 2. DIMENSIONALITY REDUCTION IN HYPERSPECTRAL IMAGES

Hyperspectral image has very strong spectral correlation that means there is often a significant amount of redundant and misleading information present in the image and such information needs to be removed before any further processing can be carried out [2],[3],[4], and [6]. For this reason, it is often useful to reduce the data dimensionality. Dimensionality reduction techniques can be a transformation-based or a selection-based [5]. The key distinction between these two is whether a dimensionality reduction technique transforms or preserves the meaning of the original dataset in the process of reduction. The choice of dimensionality reduction technique is often guided by the particular problem at hand.

### 2.1 Transformation-based reduction

Transformation-based reduction approach is a semantics destroying dimensionality reduction technique (referred to as feature extraction) which in the process irreversibly transforms the original dataset features [5]. These methods are commonly employed in the situations where the semantics of the original datasets are not needed by further process. Many approaches such as principal component analysis [7], minimum noise fraction transform [10], discriminant analysis [11], [15], non-parametric weighted feature extraction [16], [17], wavelet transform [19], and spectral mixture analysis [18] have been used for feature extraction in hyperspectral images in order to reduce data redundancy or to extract specific information. This section briefly discusses several such popular techniques that have been used for hyperspectral images. Following Table 1 give a summary about popular feature extraction methods used for hyperspectral images.

**Table 1:** A summary of feature extraction methods for hyperspectral images

Feature Extraction Methods	Characteristics
Principal Component Analysis (PCA)	<p>In hyperspectral imaging, PCA is based on the fact that neighboring bands are highly correlated which often convey almost the same info about the object. Hence, the analysis is used to transfer higher dimensional hyperspectral data to lower dimensional hyperspectral data removing the correlation among the bands. It employs statistic properties of hyperspectral image [7].</p> <p><b>Merits:</b></p> <ul style="list-style-type: none"> <li>• Most of the information may exclusively be contained in the first few PCA bands</li> <li>• As number of PCA bands increases, the information content of the PCA bands reduces</li> <li>• Remaining bands contain noisy information</li> </ul> <p><b>Demerits:</b></p> <ul style="list-style-type: none"> <li>• Only the subsets of the original bands are used for the analysis</li> </ul>

Minimum Noise Fraction (MNF)	<p>and PCA is performed on these subsets</p> <ul style="list-style-type: none"><li>• PCA ca not eliminate anomalies that can be seen in one arbitrary band</li><li>• PCA transforms an image into a new coordinate system without considering factors such as noise</li></ul> <p>MNF is a data transformation technique [10]. In hyperspectral image processing, MNF is commonly used to align the data along axes of decreasing signal to noise ratio (SNR). Since, hyperspectral images tend to contain more noise, MNF is a technique to address the noise present in the image. MNF maximizes the estimated SNR. Few noise filters are used like Adaptive Filters (AF), etc.</p> <p><b>Merits:</b></p> <ul style="list-style-type: none"><li>• MNF of an image (i.e. higher order bands in hyperspectral image) can be truncated incorporating low SNR that eliminates much of the noise without losing much information</li></ul> <p><b>Demerits:</b></p> <ul style="list-style-type: none"><li>• If truncation consists of too many bands, too much noise is left in the image</li><li>• If truncation consists of too few bands, useful signal may be excluded from the resulting image</li></ul>
Discriminant Analysis	<p>Linear Discriminant Analysis (LDA): Fisher's linear discriminant analysis is well suited when data is normally distributed. It is a parametric feature extraction method. LDA is based on following three scatter matrices: within-class, between class and mixture (or total) scatter matrices which can be used to formulate criteria of class separability. Transforms the data from a high dimensional space into low dimensional subspace with maximized class separability. In hyperspectral images, transformation is done using the ratio of number of training samples and spectral features [10], [15].</p> <p><b>Merits:</b></p> <ul style="list-style-type: none"><li>• It is distribution-free</li><li>• Fast and easy to apply</li><li>• Gives better within class scatter matrices</li></ul> <p><b>Demerits:</b></p> <ul style="list-style-type: none"><li>• The rank of between-class scatter matrix is the number of classes minus one. Therefore, the maximum number of features that can be extracted remains the same which may not be sufficient for achieving better accuracy in practical application</li><li>• Cannot perform well for non normally distributed data</li><li>• Sensitive to outliers</li><li>• In hyperspectral image processing, the ratio of the number of training samples and features is small due to the fact that labeling is expensive and that the feature space has high dimensionality. Hence, it suffers from the singularity problem since within class covariance is singular</li></ul>
Non-parametric feature extraction	<p>It is commonly used for high dimensional multiclass pattern recognition problem. Based on non-parametric extension of scatter matrices. To overcome the limitations of LDA non-parametric feature extraction algorithms are developed like nonparametric discriminant analysis (NDA) based on nonparametric between class scatter matrix, nonparametric weighted feature extraction (NWFE) with fully nonparametric within-class and between-class scatter matrices, nonparametric weighted feature extraction (NWFE), some regularization techniques to solve singularity problem, and nonparametric fuzzy feature extraction (NFFE) which introduces fuzzy membership grades to design its within class and between-class scatter matrices [16], [17].</p> <p><b>Merits:</b></p> <ul style="list-style-type: none"><li>• Suitable for handling linear problems</li><li>• employment of a suitable regularization</li></ul>

Linear Spectral Mixture Analysis (LSMA)	<ul style="list-style-type: none"><li>• Adjustment of the extracted features can be done</li></ul> <p><b>Demerits:</b></p> <ul style="list-style-type: none"><li>• NDA becomes hardy on unequal covariance or complexly distributed data</li><li>• May not be suitable for non-linear problems. To extract the non-linear features kernel version of the algorithms have been used</li></ul>
Wavelet-based Methods	<p>Linear Spectral Mixture Analysis is widely used to address the mixed pixel problem. In hyperspectral image analysis, LSMA is commonly used for subpixel detection and mixed pixel classification. It is based on the assumption that the image pixel is linearly mixed by materials (referred to as targets). It determines the component parts of mixed pixels and identifies the proportion of a pixel which belongs to a particular class or feature based on the spectral characteristics of the endmembers where endmembers are the 'pure' spectra corresponding to each of the classes [19].</p> <p><b>Demerits:</b></p> <ul style="list-style-type: none"><li>• Sensitive to target knowledge and noise</li><li>• Abundance fraction of targets present in an image are assumed to be unknown that needs to be estimated</li></ul> <p>To address these issues some algorithms are modified and developed like Independent component analysis (ICA) based linear spectral random mixture analysis (LSRMA) assuming abundance fraction as random.</p> <p>Wavelet-based reduction for hyperspectral imagery is based on the fact that the spectral data from original feature space is transformed to a reduced feature space. The main principle of this method is to apply Discrete Wavelet Transform (DWT) to hyperspectral data in spectral domain and at each pixel location. Here, wavelet reduced data represent a spectral distribution similar to the original distribution but in compressed form [18].</p> <p><b>Demerits:</b></p> <ul style="list-style-type: none"><li>• Lossy compression</li><li>• Removed high frequency signal may contain useful information for class separation and identification</li></ul>

Thus, two major issues arise from such transformation-based dimensionality reduction techniques. One is the number of data dimensions required for dimensionality reduction to avoid significant loss of information. The other is that, since the data are transformed, some crucial and critical information may have been compromised and distorted. Hence, the transformation-based dimensionality reduction technique usually changes the physical meaning of the original data [4], [2].

## 2.2 Selection-based reduction

Selection-based reduction approach is a semantics preserving technique (referred to as feature selection or band selection in hyperspectral images) which attempt to retain the meaning of the original dataset. The main aim of selection-based dimensionality reduction technique is to find a minimal subset of the original data without losing their physical meaning [4], [2]. In hyperspectral images, to reduce the dimensionality of hyperspectral data, band selection technique is commonly used. Band selection in hyperspectral images refer to select a subset of original bands containing most useful information about the object, removing redundant, irrelevant or misleading spectral features from the hyperspectral data. According to the availability of class information, band selection is categorized into two types: unsupervised band selection method and supervised band selection method [6].

### 2.2.1 Unsupervised band selection

When class information is not known in advance, unsupervised band selection has to be adopted to find the most informative and distinctive bands. Unsupervised band selection methods can be categorized based on the following classification as shown in Table 2.

**Table 2:** Unsupervised band selection methods for hyperspectral images

Categories	Characteristics
Methods based on Information Theory	It includes the use of information theory like mutual information. Provides a methodology to find generalized correlation among image bands. Divergence measures can be used to compare different probability distributions [14].
Methods based on Constrained Band Selection (CBS)	It is different from variance-based methods or information theoretic criteria-based methods, where this approach linearly constrains a band to minimize the correlation or dependency of this particular band from the other bands in a hyperspectral image [13], [14].
Methods based on eigenanalysis and decorrelation of the input data	Maximum Variance Principal Component Analysis (MVP) is one of the technique based on eigenanalysis and decorrelation of input data. In this, a joint band- prioritization and band-decorrelation approach is adopted for hyperspectral image classification. The band prioritization depends on an eigenanalysis, decomposing a matrix into an eigen form matrix from which a loading factors matrix has been constructed and used to prioritize-bands [12], [14].
Methods based on clustering strategy	Clustering process strategy refers to group bands that minimizes the intra-cluster variance and maximizes the inter-cluster variance in terms of dissimilarity measures based on band information content such as mutual information, etc. [9].

### 2.2.2 Supervised band selection

When class information is already known supervised band selection is applied to preserve the desired object information. Supervised feature selection methods involve two major steps: *subset generation and subset evaluation* [23]. Subset generation is nothing but a process of heuristic search, where each point in the search space, specifies a candidate subset for evaluation. Band selection in hyperspectral images requires an efficient searching strategy [21], [22]. Optimal search techniques such as Exhaustive Search or Branch and Bound are not applicable for hyperspectral images due to the huge amount of complexity, since in these techniques all possible band combinations are tested and then optimal one will be selected. Hence, to avoid this, different suboptimal techniques, such as SFS, SBS, SFFS and SBFS are used [21]. Various stochastic optimization based feature (band) selection techniques for hyperspectral images have been used such as Genetic Algorithms (GA) [25], Particle Swarm Optimization (PSO) [24], [26] and Ant Colony Optimization (ACO) [20]. In subset evaluation, the fitness function (evaluation criteria) of each subset of feature is calculated. Based on different evaluation criteria, supervised band selection methods regardless of their searching strategy stage, can be classified into two approaches: filter approach and wrapper approach [3].

**Filter Approach:** Band selection method based on filter approach is dependent on general characteristics of the data to evaluate and select feature (band) subsets without involving any classification algorithm so that undesirable features are omitted before the classification process begins. It uses some statistical methods (i.e. discrimination measures) such as the Mahalanobis Distance [26], Bhattacharyya Distance [26], Jeffreys-Matusita Distance [27] to identify an optimal subset of bands in hyperspectral image.

**Wrapper Approach:** Band selection method based on wrapper approach is dependent on one predetermined classification algorithm and uses its performance as an evaluation criterion. It determines features (bands) which better suit the classification algorithm in terms of performance. Feature Shaving is an example of this type which removes the whole band at once including extra parameters such as maximum size of smoothing window, a few runs of whole procedure, an expert knowledge, etc. [8].

## 3. PROPOSED SYSTEM

Rough Set Theory (RST) introduced by Pawlack has been widely used as a selection tool to discover data dependencies and to reduce number of attributes contained in a dataset using the data alone and no additional information. It deals with incomplete, imperfect data, and data vagueness. A Rough Set is defined by two approximation sets called as lower

approximation sets and upper approximation sets. Lower approximation sets consists of the set of objects that are certainly belong to the subset of interest, while upper approximation sets consists of the objects that are possibly belong to the subset of interest. Given a dataset with discretized attribute values, then using RST it is possible to find a subset (termed a reduct) of the original attributes that are most informative and all other attributes can be removed from the dataset with minimal information loss. However, the dataset containing that the values of attributes may be both crisp and real-valued, and this is where traditional RST encounters a problem. Therefore, it is desirable to develop techniques to provide the means of data reduction for crisp and real valued attribute datasets and this can be achieved through the use of fuzzy-rough sets [5]. Fuzzy rough set approach combines the rough set and fuzzy set to tackle the weakness of the classical rough set in order to deal with real value data without discretization. Fuzzy-rough sets abridge the related but distinct concepts of vagueness and indiscernibility both of which occur as a result of uncertainty in knowledge. Thus, it preserves the meaning of the data knowledge. In addition, it does not require additional information for the process of feature selection. Feature selection based on fuzzy rough sets particularly depends on defining a fuzzy-similarity relation between objects within a subset of features. Using fuzzy similarity relation, fuzzy concepts can be approximated by defining the fuzzy lower approximation sets and fuzzy upper approximation sets [5], [28]. Let, a universe  $U$  as a finite non empty set of objects each defined by a set of attributes, denoted by  $A$ . Here, the pair  $(U, A)$  is known to be an information system (IS) where for every subset  $P \subseteq A$  there exist an associated similarity relation.  $\mu_{R_p(x,y)}$  denotes the similarity between objects  $x$  and  $y$  induced by the subset of features  $p$ . Given  $X \in U$ ,  $X$  can be approximated by the information contained in  $P$  constructing the  $P$ -lower and  $P$ -upper approximations of  $X$  as defined in Equations (1) and (2) respectively [28].

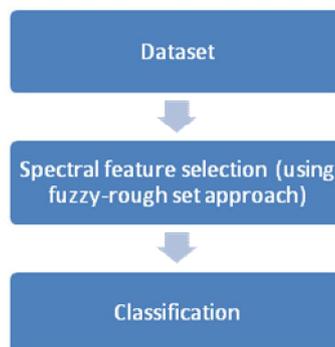
$$\underline{\mu}_{R_p} X(x) = \inf_{y \in U} I(\mu_{R_p}(x,y), \mu_x(y)) \tag{1}$$

$$\overline{\mu}_{R_p} X(x) = \sup_{y \in U} T(\mu_{R_p}(x,y), \mu_x(y)) \tag{2}$$

Where  $I$  denotes the fuzzy similarity implicator,  $T$  is the t-norm, and  $\mu_p$  denotes the fuzzy similarity relation induced by the subset of features  $P$ . The main issue of fuzzy rough set approach is to select a subset of features called reduct  $R$  that represents the original set of features. So feature selection based on fuzzy rough set approach search for the optimal subset of features that approximately has either the same certainty or uncertainty of the original set of features. It depends on finding the degree of dependency between the objects which can be computed using the lower approximations. Let degree of dependency of object  $Q$  on a subset of features  $P$  denoted as  $\nu'_P(Q)$  and can be computed using Equation (3) [28].

$$\nu'_P(Q) = \frac{\sum_{x \in U} \mu_{pos_{R_p}}(Q)(x)}{|U|} \tag{3}$$

Thus, in the proposed system, fuzzy-rough set approach is used for spectral feature selection to reduce the dimensionality of the original hyperspectral data. The objective of the proposed system is twofold. First, select spectral features from the original hyperspectral data which in turn will reduce the number of bands in hyperspectral image. Second, use selected features for hyperspectral image classification. Fig. 2 shows the flow chart for the proposed system.



**Figure 2** Flow chart of the proposed system

#### 4. CONCLUSIONS

This paper provides a survey of various dimensionality reduction techniques in terms of feature extraction and band selection methods for hyperspectral images. In feature extraction methods, original data are transformed and compact set of original features are generated. Commonly used feature extraction methods used such as principal component

analysis, minimum noise fraction transform, discriminant analysis, non-parametric weighted feature extraction, wavelet transform, and spectral mixture analysis have been discussed for hyperspectral images. In band selection, a subset of significant features is selected for preserving the meaning of the original dataset. Supervised and unsupervised approaches for band selection have been carefully studied. A new framework for dimensionality reduction have been proposed based on fuzzy rough sets which selects important features with no additional information.

## References

- [1] [http://www.youtube.com/watch?v=17cRTkG31\\_w](http://www.youtube.com/watch?v=17cRTkG31_w)
- [2] Jun Li, Prashanth Reddy Marpu, Antonio Plaza, Jose M. Bioucas-Dias, and Jon Atli Benediktsson, "Generalized Composite Kernel Framework for Hyperspectral Image Classification," *IEEE transactions on geoscience and remote sensing*, vol. 51, no. 9, September 2013.
- [3] B. Mojaradia, h.emamib, m.varshosazc and s. Jamali, "A Novel Band -Selection Method for Hyperspectral Data Analysis," *The international archives of the photogrammetry, remote sensing and spatial information sciences*. Vol. Xxxvii. Part b7. Beijing 2008.
- [4] Kun tan, Erzhu Li, Giandu, and Peijun Du, "Hyperspectral Image Classification Using Band Selection And Morphological Profiles," *IEEE journal of selected topics in applied earth observations and remote sensing*, vol. 7, no. 1, January 2014.
- [5] Richard Jensen and Giang Shen, *Computational Intelligence and Feature Selection: Rough and Fuzzy Approaches*.
- [6] Qian du and He Yang, "Similarity-based Unsupervised Band Selection for Hyperspectral Image Analysis," *IEEE geoscience and remote sensing letters*, vol. 5, no. 4, October 2008.
- [7] Craig Rodarmel and Jieshan, "Principal Component Analysis for Hyperspectral Image Classification," *Surveying and Land Information Systems*, vol. 62, no. 2, 2002.
- [8] Sergueiverzakov, Pavelpačlik, and Robert P.W. Duin, "Feature Shaving for Spectroscopic Data," Springer-Verlag Berlin Heidelberg 2004.
- [9] Adolfo Martinez-Uso, Filibertopla, Jose MartinezSotoca and Pedro Garcia-Sevilla, "Clustering-based Hyperspectral Band Selection Using Information Measures," *IEEE transactions on geoscience and remote sensing*, vol. 45, no. 12, December 2007.
- [10] Rhonda D. Phillips and Layne T. Watson, Christine e. Blinn and Randolph H. Wynne, "An Adaptive Noise Reduction Technique for Improving The Utility of Hyperspectral Data," *Pecora17 – The future of land imaging. Going operational November 18 – 20, 2008*.
- [11] Qian Du, Chein-I Chang, "A Linear Constrained Distance-based Discriminant Analysis for Hyperspectral Image Classification" *Pattern Recognition* 34 (2001).
- [12] Chein-I chang, Qian Du, Tzu-Lung Sun, and Mark I. G. Althouse, "A Joint Band-Prioritization And Band-Decorrelation Approach To Band Selection for Hyperspectral Image Classification" *IEEE transactions on geoscience and remote sensing*, vol. 37, no. 6, November 1999.
- [13] Chein-I Chang, and Su Wang, "Constrained Band Selection for Hyperspectral Imagery," *IEEE transactions on geoscience and remote sensing*, vol. 44, no. 6, June 2006.
- [14] Adolfo Martinez-Uso, Filibertopla, Jose M. Sotoca, and Pedro Garcsevilla, "Comparison of Unsupervised Band-Selection Methods for Hyperspectral Imaging," Springer-Verlag Berlin Heidelberg, pp. 30–38, 2007.
- [15] Tatyana V. Bandos, Lorenzo Bruzzone and Gustavo Camps-Valls, "Classification of Hyperspectral Images With Regularized Linear Discriminant Analysis," *IEEE transactions on geoscience and remote sensing*, vol. 47, no. 3, March 2009.
- [16] Jinn-Min Yang, Pao-Ta Yu, Bor-Chen Kuo, and Ming-Hsiang Su, "Nonparametric Fuzzy Feature Extraction for Hyperspectral Image Classification," *International Journal of Fuzzy Systems*, vol. 12, no. 3, September 2010.
- [17] Bor-Chen Kuo and David A. Landgrebe, "Hyperspectral Data Classification Using Nonparametric Weighted Feature Extraction".
- [18] Chein-I Chang, Shao-Shan Chiang, James A. Smith and Irving W. Ginsberg, "Linear Spectral Random Mixture Analysis for Hyperspectral Imagery," *IEEE transactions on geoscience and remote sensing*, vol. 40, no. 2, February 2002.
- [19] B.Salehia and M.J.Valadanzoelj, "Wavele-based Reduction of Hyperspectral Imagery".
- [20] Jianwei Gao, Gian Du, Lianru Gao, Xu Sun and Bing Zhang, "Ant colony optimization-based supervised and unsupervised band selections for hyperspectral urban data classification," *Journal of Applied Remote Sensing* 085094-1 vol. 8, 2014.
- [21] Aloke Datta, Susmita Ghosh and Asish Ghosh, "Wrapper-based Feature Selection In Hyperspectral Image Data Using Self-adaptive Differential Evolution," 2011 International Conference On Image Information Processing (iciip 2011).
- [22] P. A. Devijver and J. Kittler, "Pattern Recognition: A Statistical Approach," Prentice-Hall int., first ed., 1982.

- [23] A. Jain and D. Zongker, "Feature Selection: Evaluation, Application, and Small Sample Performance," IEEE transaction on pattern analysis and machine intelligence, vol. 19, pp. 153--189, 1997.
- [24] S. Ding and L. Chan, "Classification of Hyperspectral Remote Sensing Images With Support Vector Machines And Particle Swarm Optimization," in proc. Of the int. Conf. on information engineering and computer science (iciecs'09), pp. 1--5, 2009.
- [25] J. P. Ma, Z. B. Zheng, Q. X. Tong, and L. F. Zheng, "An Application of Genetic Algorithms On Band Selection for Hyperspectral Image Classification," in proc. of the 2nd int. Conf. on machine learning and cybernetics, pp. 2810--2813, 2003.
- [26] P. A. Devijver and J. Kittler, Pattern Recognition: A Statistical Approach. Prentice-Hall int., first ed., 1982.
- [27] L. Bruzzone and S. B. Serpico, "A Technique for Feature Selection in Multiclass Problems," int. Journal of remote sensing, vol. 21, no.3, pp. 549--563, 2000.
- [28] Basem O. Alijla, Ahamad Tajudin Khader, Lim Chee Peng, Mohammed Azmi Al-Betar and Wong Li Pei, "Fuzzy- rough set approach for selecting the most significant texture features in mammogram images," 2013 Palestinian International Conference on Information and Communication Technology.

#### **AUTHOR**

**Shraddha P. Lodha** received the B.E. degree in Computer Engineering from the K. K. Wagh Institute of Engineering Education and Research, Savitribai Phule University of Pune, India in 2013 and she is currently pursuing the M.E. degree in Computer Engineering at K. K. Wagh Institute of Engineering Education and Research, Savitribai Phule University of Pune, India.