International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-2, July 2019

Unfolding the Restrained Encountered in Hyperspectral Images

Kriti, Urvashi Garg

Abstract*: The content of the natural scenes needs to be interpreted which is the primary concern in computer vision. In the advancement of the systems focusing on the intelligent image understanding, the most powerful key is the degree to which meaningful information is extracted by the computer. Moreover with this advancement in the field of image processing, precise and huge information capturing images are desired. The hyperspectral images find its place in such fields of applications. For a single scene, the hyperspectral images (HSI) are composed of hundreds of channels of spectral data. For different materials, with the availability of detailed spectral information, hundreds of contracted bands are collected by hyperspectral sensors. However, with the dimensional complexity, its impact varies from field to field. We reiterate our main focus in this article on providing the various challenges existing relating to HSI and a case study of the current solutions provided for each. A clear depiction of the current issues and approaches in the field of compression as well as some general issues are also discussed towards the end section.

Index Terms*: hyperspectral imaging (HSI), hyperspectral sensors, multispectral imaging.

I. OUTLINE TO HYPERSPECTRAL IMAGES

A. *Difference of HSI from usual images*

The image produced by usual digital camera contains either intensity or selected color representation (e.g. RGB) while the images that are multispectral in nature ease to deliver spectral evidence for each and every pixel in the wavelength range for a specified spectral resolution. The processing of the visual information is involved in certain applications where using the spectral data is of great prominence in performing tasks e.g. in remote sensing, medical imaging, fine arts, assessment of product quality. The increasing drift for these schemes to practice spectral and spatial information leads to hyperspectral image representation that finds numerous applications in estimating and analyzing the existence of chemical compounds, pathologies and other statistics that provides qualitative and at the same time quantitative evaluation of these features [1].

Revised Manuscript Received on 30 July 2019.
* Correspondence Author
Kriti*, Department of Computer Science and Engineering, Chandigarh University, India.
Urvashi Garg, Department of Computer Science and Engineering, Chandigarh University, India.

B. *Need for HIS*

In remote sensing, hyperspectral imaging has become a core technology and is also known as image spectrometry. A very high resolution is exhibited by hyperspectral imaging that provides diagnostic capabilities that are better for detection of objects, classification, and discrimination as compared to the multispectral imaging [2]. Moreover, in the traditional imaging sensors considering the human visual system, only three spectral measurements (red, green, and blue: RGB) are captured that covers the visible electromagnetic spectrum. On the contrary, there is a large quantity of information beyond the visible range being carried in other series of electromagnetic spectrum embracing near-infrared and infrared ranges. For the first time, new doors have been opened by HSI techniques allowing the provision of exploiting the information beyond the visible bands of the electromagnetic spectrum [3].

C. *Sensors in HSI*

To monitor the various environmental changes in our planet hyperspectral sensors have been developed in the latest years. For the wide variety of instances applicable to the classification of ground materials and recognition of objects, the hyperspectral sensors are employed to work for a narrow continual spectral band as they find applications in military surveillance, analysis of urban-growth, monitoring of agriculture, finding of minerals, detection of material defects and others [4]. From the visible to infrared spectra, the hyperspectral image (HSI) data set spans between hundreds of electromagnetic spectral bands [5]. The rich informative HSI data poses challenges in the expansion of effective and resourceful algorithms for the HSI classification [6]. From satellite, airborne as well as ground-based sensors, an enormous quantity of data is obtained with developing remote sensing technology. According to the spectral and spatial facts, to satisfy the purpose of remote sensing, the fundamental goal of classification is assigning the label to distinctive objects for an image. For the collection of hyperspectral image, only spectral information concerning each pixel is acquired whereas manual acquirement of label information is often done by experts [7].

D. *Classification in HSI*

The HSI classification remained a vibrant zone of investigation in the latest years. The classification task for HSI is developing rapidly due to enormous applications in mineralogy, agriculture as well as in surveillance and therefore a huge quantity of procedures are anticipated tackling this delinquent [8].

The purpose of classification approach is in allocating unique label to each and every pixel vector for a given set of observations in HSI. In figure 1, a general framework of HSI is given since the classification itself is a vast field to be incorporated in a single diagram, hence the major terminology associated with HSI classification is given enabling the reader to get the basic idea relating to the classification field. For more than three decades, the dimensional complexity has been known, and so its impact that varies from field to field. Nowadays, a resort to the model of statistical learning is a usual practice aimed at the users of remote sensing data. The satellite imagery of new generation is used for the classification of land-use where it is considered as the state of art algorithm that includes models such as neural networks, SVMs. A very high resolution is achieved under the applications of such models that prove their efficiency in handling the data remotely sensed [10]. One of the prominent topics in HSI analysis is instrument and measurement in remote sensing, having wide applications in skin imaging [11], and identification of ground elements [12] and exploration of minerals [13].

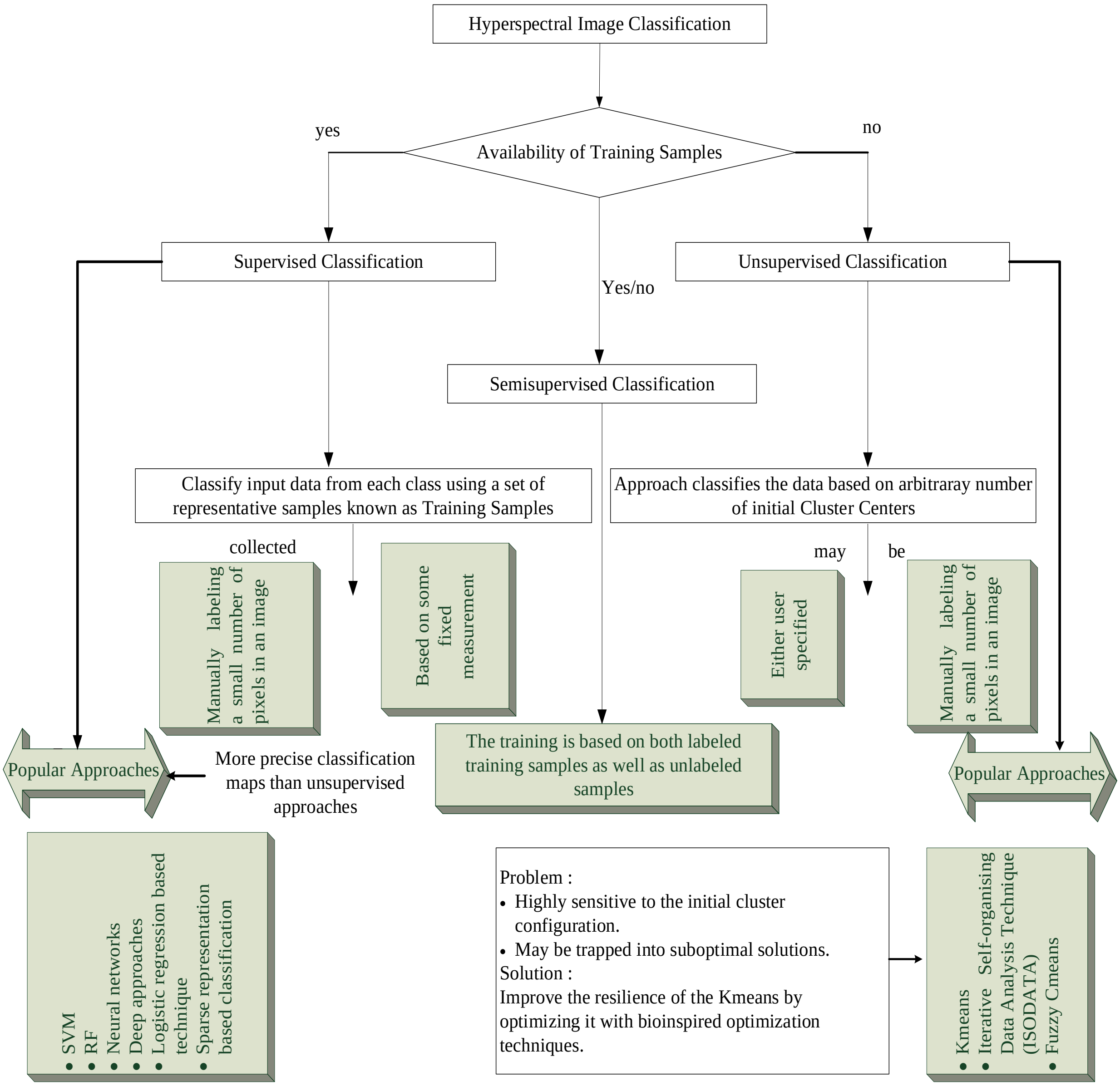

Fig. 1.General framework of Classification in HSI with an elaborated focus on supervised and unsupervised classification

Retrieval Number: B1763078219/19©BEIESP
DOI: 10.35940/ijrte.B1763.078219
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

The collection of HSI is done by captured and recorded spanning from visible to infrared spectrum [14]. Aimed at a single pixel, the spectrum information is signified by a vector whose entities corresponds to standards of dissimilar spectral band [15]. For the very first type of classification i.e. supervised algorithms, its performance however strongly be contingent on data representation used to train the classifier [16]. The second type of classification i.e. semisupervised classification is known to outperform the supervised classification approaches in accuracy. In the literature of hyperspectral scenes, benchmarkimages avail a minor amount of training illustrations [14]. With a great dimension of data as in case of HSI, a quantity of training models limit the classification performance. The unsupervised classification is yet another type in the arena of hyperspectral image classification and remains in great demand focusing on the domain where the existence of supervised and semisupervised classification approaches fails to give the required performance. With the increase in dimensionality, the classification algorithm accuracy decreases, as denoted by well-known Hughes effect [15]. An important role is played by images avail a minor amount of training illustrations [14]. With a great dimension of data as in case of HSI, a quantity of training models limit the classification performance. The unsupervised classification is yet another type in the arena of hyperspectral image classification and remains in great demand focusing on the domain where the existence of supervised and semisupervised classification approaches fails to give the required performance. With the increase in dimensionality, the classification algorithm accuracy decreases, as denoted by well-known Hughes effect [15]. An important role is played by quality and quantity to achieve accurate results for classification [16]. Moreover, in the process of HSI classification, factors like presence of redundant features, limited training samples availability and imbalance among them plus high dimensional data leads to challenges. The field of HSI classification itself is a large domain for the researchers to find a problem and a required solution to it. Most of the challenges discussed below are primarily related to the classification phase of HSI. The researchers can benefit from the article as the current or most popular remedies to the existing problems in the classification domain has been discussed with a clarity of the future work. Being the most difficult phase of research i.e. to identify the problems, this article benefits the researcher with multiple challenges associated with the hyperspectral images.

II. CHALLENGES IN HSI

A growing interest has been witnessed in the recent years in the HSI processing. Considering the hyperspectral channels, rich amount of information is provided by HSI as compared to gray-scale and RGB images hence facilitating with more knowledge to each observation. In fields like multimedia, computer vision and image processing, hyperspectral technology, as a consequence have become a powerful tool. HSIs however in practice suffers from various issues like huge data size, the mass of redundant information etc. hindering its application in many cases [27]. Some of these issues have been stated in this article with the respective solutions to each. Since the solutions are richer than the existing problem and vary from application-to-application, hence very concise discussion to some of the popular or new solutions have been provided to avoid confusion to the researchers. Moreover, restricting to an aim of the article where the primary focus is to introduce the researchers to the domain of HSI and its problems, a brief introduction to some of the popular challenges has been given in this section. The paper primary focus is to provide a research problem to the reader enhancing their clarity in this field.

A. *Non-linearity of Data*

Over the past decade, the attentiveness in manifold learning to represent the topology of huge and nonlinear high dimensional data set is inferior yet still significant, wherein the dimensions to visualization and classification are growing rapidly particularly in the hyperspectral imagery analysis.The data constituting a resolution of high spectral and continuous band of HSI allows for discrimination between targets of interests that are spectrally similar providing capabilities in estimating constituent abundance within pixels and hence allowing direct exploitation in the predictive model for absorption features. For the HSI data, the dense spectral sampling allows for the association of spectral information in many adjacent bands that are exceedingly correlated, resulting into considerable lesser inherent dimensional space traversed by data (figure 2). To exploit the non-linear characteristics of the data the development of specialized methods has been motivated to meet the amplified accessibility of HSIs with superior admittance to advance computing. In this context, significant attention has been received by a selection of features and its extraction. From the perspective of back-end classifiers, both feature extraction and feature selection are relatively flexible. From the perspective of classification meaningful information is retained by the feature selection while featuring extraction projects data to intrinsic spaces of lower dimensions. The algorithms involved in feature selection are computationally intensive while for feature extraction it is superior to the finest feature selection. The feature selection is not vigorous in multifaceted scenes while extraction of feature is more vigorous to variations in the spectral signatures across the scenes. Feature selection preserves the original spectral signature. While in feature extraction the interpretation relative to that of the original spectral signature is lost. Hence the difference between feature extraction and feature selection are stated clearly. One of the approach to withstand the non-linearity of data is the nonlinear manifold learning in a graph entrenching structure. The advantage of the approach lies in its ability to mitigate the impact that effects electromagnetic energy navigating the atmosphere and reflected from the target. However, it has various disadvantages. Nonlinearity is not always exhibited in data, therefore user's evaluation of inherent nonlinearity in the data becomes always necessary. It requires parameter tuning and higher computation is exhibited by large-scale data sets. Being data-driven, manifold learning results strongly depends on data characteristics, hence one method cannot provide consistent results. Its assumption of inherent smoothness violated by data sets containing classes whose spectra is distinctly different resulting in either multiple manifolds or submanifolds that cannot be unified with the representation of solitary manifold.

Hence opportunities exist where appropriate characterization is developed that exploits unique characterization of manifold keeping structure of hierarchical manifold a merit. The work can further be extended beyond feature extraction. For local and global embedding methods, joint exploitation is needed in dynamic, multitemporal environment integrating active and semisupervised learning. Numerous other approaches are available in the literature like the contribution of localized spatial information have been known to provide ominously advanced precisions and superior visualizations, even though the computational overhead of such methods would prerequisite to be considered for outsized remotely sensed data sets [33]-[36]. The HSIs availability has increased providing grander access to progressive computing, in turn, motivating the expansion of generalized approaches exploiting the nonlinear physiognomies of the data. In this context, significant attention is received by dimensionality reduction problem with solutions provided by feature selection and extraction approaches. Both these approaches are quite successful in the arena of classification. Though some loss of information is seen relative to the original data, yet effective results are shown by both in classification [54].For a given data set with training samples, $\ddot{X} = \{xi\}_{i=1}^{n}$ in R^m (feature space of *m*-dimensional) and *n* being total quantity of samples for training [54]. A graph embedding framework is adapted by nonlinear dimensionality reduction algorithms in which G = {$\ddot{\mathbf{X}}$, $\hat{\mathbf{W}}$}, is the undirected weighted graph and $\hat{W}$ is the *n* # *n affinity* matrix. The concept of affinity weights $\hat{W}$ *ij ε* [0, 1] is utilized by the algorithms for measuring the "distance" among observations of dual sample. The class label information is not used by the affinity function instead it characterizes the neighborhood associations concerning entire pairs of points on the basis of differences of the features. The heat-kernel is a popularly used approach for measuring the affinity between samples x*i* and x*j* to the existing problem in the domain of non-linearity in the hyperspectral images. There are various preprocessing techniques that can be applied to the images to enhance their visibility and extract useful information from them so that the wide range of applications of hyperspectral image can fetch where $\gamma i = ||x_i - x_i^{(k_{nn})}||$ denoting the native scaling of samples of data in the neighborhood of x_i and $x_i^{(k_{nn})}$ is the k_{nn}-nearest neighbor of x*i* .As stated this is just one of the proposed solution

$$W_{ij} = \exp\left(-\frac{||x_i - x_j||^2}{y_i y_j}\right) \qquad (1)$$

maximum benefit from it elaborating its capability to provide a vast amount of knowledge. The applications of hyperspectral images exist owing to the opulent spectral and spatial statistics delivered by it. Hence the challenge of non-linearity has to be resolved in the very beginning by using any of the image enhancement technique well suited for hyperspectral images so that detailed study of it can be done in an appropriate manner.

B. *Ill-conditioned Problems with Remote Sensing Images (Classification by means of Restricted quantity of training samples and bias in Sample selection)*

In the remote sensing, the hyperspectral sensors capture several images in the visual and nonvisual range for the frequency bands/channels with a wavelength of 5-20 nm between the channels/bands. A very important branch in the arena of remote sensing is the classification of HSI. The class of the object may be recognized on the basis of spectral measurements at a point leading to the major mission of classification where it assigns a label to every pixel generating a map of land cover. The division for the classification methods is done in two forms namely supervised and unsupervised classification. For the purpose of better evaluation focus is primarily on supervised classification. Due to the complicated ground features, finding a common sparse representation is a very difficult task of remote sensing images. An image of remote sensing is regarded as the combination of sub-image of edges, even and point-like components respectively. To represent only a particular kind of texture or ground object, this capability is shown by each domain method so for each sub-image in the sparse representation group of domain representation is used. The ill-conditioned problems like super resolution (SR) and classification are solved using MSR (mixed sparse representation) which is regarded as a prior to maximum a posteriori and for the remote sensing images the effects of the framework have been demonstrated. When compared to the other common SR methods, the new framework of MSR constructs high-resolution images in a much better way that adds to the improvement in the accuracy in classification as well. To solve the ill-conditioned problems related to remote sensing images, a competitive candidate is the proposed framework MSR [55]-[68].

Retrieval Number: B1763078219/19©BEIESP
DOI: 10.35940/ijrte.B1763.078219
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

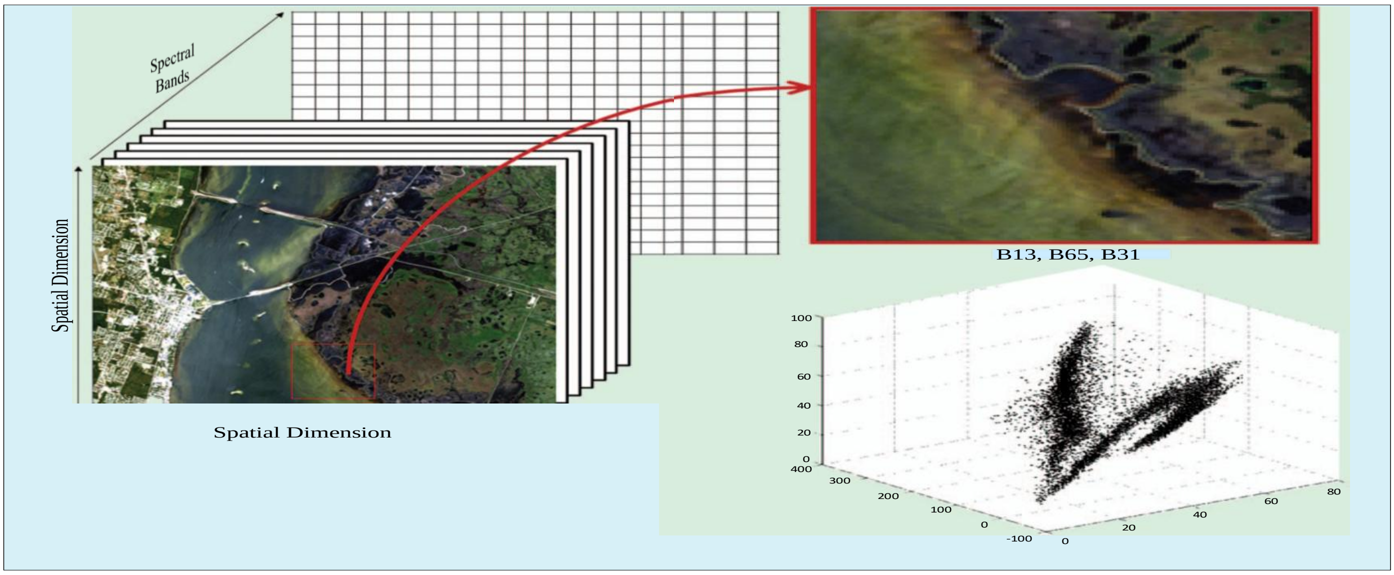

Fig . 2. An AVIRIS hyperspectral image of true color over the Kennedy Space Center (KSC), Florida. For spectral data the nonlinearity is exhibited by plot of bands: 13, 31 and 65.

Furthermore, to the problem of ill-posed classification where the availability of the samples is too few with respect to the features number, the two different approaches considered are AL and SSL, where the estimation of the underlying distribution classes is not done correctly by any one. In the classification problems of real RS, ill-posed difficulties remain very expected to befall in particular for RS images classification of the past generation e.g. HSI and images of very high resolution (VHR), wherein the existing training set to the quantity of features is usually very trivial. The high number of spectral bands in the classification of HSI leads to defining the problem of classification in the feature space of high dimension. The poor spectral resolution is provided by the available spectral bands in the VHR image classification, therefore several features of texture and geometry are extracted to characterize. A novel SSL algorithm [72] has been proposed for improving the progressive semisupervised SVM as it integrates concepts that are considered in the AL methods. The three major iterative steps include initialization, the SSL, and the convergence. In the initialization the original training samples are used by a standard SVM as it solves the following constrained optimization problem:

$$\min_{\hat{w},\xi,b} \frac{1}{2}\|\hat{w}\|^2 + \check{C}\sum_{i=1}^{n}\xi i$$

$$\text{subjected to : } yi\,[\hat{w}\cdot\varphi(xi) + b] \geq 1 - \xi i \quad i = 1, \ldots, n$$

$$\xi i \geq 0 \tag{2}$$

where $\hat{w}$ is a vector orthogonal to the unraveling hyperplane, b is the term of bias such that $b/|\ \hat{w}\ |$ is representing the distance of the hyperplane at the origin, $\check{C}$ is the parameter of regularization, φ is the mapping function of the data in the feature space, ξi is the variable in slack, and n is the quantity of training samples. Rendering to the emblem of the subsequent conclusion function $f(\mathrm{x}) = \hat{w}\cdot\varphi(\mathrm{x}) + b$, pseudo labels that are assumed to be the unlabeled trials. For any iteration, after the initialization until convergence, a pool is formed by the set of samples µ of unlabeled samples that are iteratively designated and added together to the training set in addition to their pseudo labels and are then aloof from the pool. The following sample set id is defined that lies in both higher and lower sides of the margin:

$$Hup = \{\mathbf{x}|\mathbf{x} \in \mu,\ 0 \leq f(\mathbf{x}) \leq 1)\} \tag{3}$$

$$Hdown = \{\mathbf{x}|\mathbf{x} \in \mu, -1 \leq f(\mathbf{x}) \leq 0)\} \tag{4}$$

ρ samples are designated at every iteration from every side of the boundary. The ρ samples with $f(\mathrm{x})$ in particular, nearer to 1 are selected from *Hup*, and the ρ samples with $f(\mathrm{x})$ nearer to -1 are taken from *Hdown*. This grades in the assortment of an entire of 2ρ samples, that are namely semi-labeled. Therein the procedure of iteration is stationary when the quantity of mislabeled samples of training as well as the pseudo-labeled patterns numbers are lesser or equal than $\beta\cdot m$, where β is a parameter which is user-defined. After reaching the junction the SVM is trained for the final time, in accordance to the subsequent problem of minimization:

$$\min_{w,\xi,\xi^*,b} \frac{1}{2}\|\ \hat{w}\ \|^2 + \check{C}\sum_{i=1}^{n}\xi i + \check{C}^{*}_{max}\sum_{j=1}^{m}\xi^{*}_{j}$$

$$\text{subject to : } yi\,[\hat{\mathbf{w}}\cdot\varphi(\mathbf{x}i) + b] \geq 1 - \xi i \quad i = 1, \ldots, n$$

$$y^{*}_{j}[\hat{w}.\ \varphi(x^{*}_{j}) + \mathrm{b}] \geq 1\text{-}\xi^{*}_{j} \quad j = 1, \ldots, m$$

$$\xi i,\ \xi^{*}_{j} \geq 0 \tag{5}$$

where the complete set of semi-labeled samples is related with the similar parameter $\check{C}^{*}_{max}$ for regularization. To obtain the results of accurate classification, it is important to consider not only the quantity but the quality as well. To model the fundamental dispersal of real classes, this capability is directed by the training samples quality. From the underlying distribution, designing is the mutual supposition in the algorithms of learning to train the data consisting of instances strained autonomously. The problem of ill-conditioning can be solved by an approach that uses a new framework of mixed sparse representation (MSRs). The advantage of the approach is the improved classification accuracy. Moreover, high-resolution images are reconstructed. The competitive candidate solves the problem for image related MAP. Though this approach poses challenges like MSR-SR is slower in computation and involvement of extra procedure like warping, interpolation, and blurring as they are included in every iteration. It provides the opportunity for being extended to unsupervised classification.

The experiment is further characterized by few training samples and inturns affected by bias due to the selection of the sample. Moreover, the concepts of integration are investigated to find similarity between dual paradigms of learning to improve state-of-art procedures therein uniting AL and SSL towards mutually leverage recompenses of approaches mutually. The novel algorithm of SSL proposed is improving the liberal semisupervised SVM as it integrates concepts usually considered in AL methods [69]-[72]. Table 1 provides alternative solutions to the challenge of the inadequate volume of samples for training and the bias associated with the selected samples. The table illustrates the bifurcation of the approach that is semisupervised learning where each approach has its own importance in the field of classification from the use of separate classifiers to the use of joint probability and many more. An additional brief of active learning is also given and towards the end, these two approaches can be very well integrated providing a satisfactory solution to the resolve the problem of sample bias as well as a few labeling samples. The opportunities are also given in the finest manner that takes the charge of incrementing the diversity in the field of research that the researcher can very well implement in its new era of research.

The use of supervised classification technique is done to automate the classification of RS images on the accessibility of categorized trials to train the supervised algorithm. In the learning, an expensive and time-consuming process is the collection of labeled samples. For accurate classification maps, the central importance is given to the superiority and amount of available training samples. The quantity of training trials is not sufficient neither the quality to train the classifier properly foremost to a major delinquent in real-world scenario. To enrich the supervised learning algorithm with information to be given in input and improve the accuracy of classification, SSL techniques are adopted as mentioned in the table 1 and an alternative approach improving classifier learning is Active learning that

TABLE I. Brief Summary Of Ssl And Al To Provide Relief To The Challenge Of Restricted Quantity Of Training Samples And Bias In Sample Selection (2014)

Approach	Methodology	Opportunities
Self –training [SSL]	Repeatedly used supervised learning method. The labeled samples are trained in the beginning, followed by labeling the unlabeled samples in accordance with the current decision function and then adding training set to it	For a development of hybrid solutions, integration of AL and SSL can be investigated further.
Co-training [SSL]	For labeled data, training of two separate classifiers is done for two subfeature set respectively. Each classifier classifies labeled data inturn providing other classifier labeled sample which is most confident with predicted labels. With the additional training examples, each classifier is retained when given to supplementary classifier and procedure is recapitulated.	Using various strategies, validation methods for SSL can be deeply investigated. For training set, after inclusion of semi-labeled samples, consistency of model assessed by novel validation strategies.
Generative Probabilistic Models [SSL]	It is assumed for approximation of joint probability $P(\mathbf{x}, y\|\theta)$, given θ is limitation vector of the model to be projected from annotations. Joint exploitation of labeled and unlabeled samples is benefitted by estimation of parameter vector θ. Bayes rule forms a basis for final classification.	Convergence of learning algorithm detected forming a good solution or by moving towards inconsistency in conditions that are critical with either few or biased labeled data where a reliable way of using cross-validation is not possible.
Semisupervised SVM [SSL]	In-terms of TSVM: Given a set of training points, traditional classifiers induces decision function to incorporate a learning task, wherein the minimization of classification error is done.	
Graph-Based SSL	A graph is defined by graph-based SSL method where nodes are categorized and unlabeled samples and edges reproduce the resemblance. The label smoothness is assumed by these methods over graph (cluster/manifold regularization).	
Active Learning	AL does not lead to training set that is unbiased, as the strategy of completely random selection is obtained. To diminish the quantity of training samples to be labeled is an aim of AL for classification accuracy to be satisfactory.	

is based on the assumption that few new labeled samples are to labeled and at the same time to add them to the training customary. The iterative expansion of the original training set is done that includes supervisor who usually is a human expert that assigns a accurate label to any inquired sample. The method shows its effectiveness to optimize the training sample collection effectively concerning different application domains that include RS image classification [8]-[11].

C. *Pixel-wise Representation*

In HSI, classification of an image is one of the most crucial procedures labeling the pixels to unique classes on the basis of spectral features. In the fields of mineralogy as well as agriculture and surveillance, HSI is in great demand leading to the rapid development of HSI classification task with a huge amount of methods being projected tackling this delinquent [73]. For SRC, a assessment sample $\mathbf{y} \in \mathrm{R}^P$, where P defines the quantity of spectral bands that can be inscribed as a sparse linear amalgamation where $x \in \mathrm{R}^N$ and $\|x\|1 = \sum_{i=1}^{N} |xi|$ is l_1–norm. A = [a1, a2, . . . , aN] is a organized dictionary designed from the concatenation of numerous classwise subdictionaries, $\{ai\}i$=1,..., N signifies the columns of A, N defines the entire number of training samples from all the K classes, and λ is a scalar regularization parameter. The class label aimed at the test pixel y is resoluted through a minutest residual among y and its estimation from every classwise subdictionary

$$\text{class}(\mathbf{y}) = \arg \min_{g} \|\mathrm{y} - A\delta g(\mathrm{x})\|_2^2 \qquad (6)$$

A highly effective approach when compared with previous approaches on both classification results and computational efficiency is SVM. To improve its performance an extensive diversity of modifications to SVM's have been projected.

Few of them integrate the contextual information in the classifiers [74] [75]. In order to pursue the rule of sparse decision, others design sparse SVM by using l1-norm as the regularizer [76]. To resolve numerous computer visualization errands SRC has been proposed recently [77], leading to state-of-the-art recital sparsity being used as a prior. To rely on the scrutiny that hyperspectral pixels belong to the identical class, SRC is being applied to HSI classification that approximately lies in the equivalent low-dimensional subspace. In the analysis of HSI, one of the most important procedure is the pixel-wise classification that assigns every pixel to a predefined class. An SRC gives rather plausible results as it represents an assessment pixel as a linear amalgamation of a trivial subset of labeled pixels as compared to traditional classifiers like SVM. The second-generation SRCs appeared in the literature when additionally structured scarcity priors are incorporated that repots to expand the performance of HSI further. The exploitation of the spatial dependencies forms the basis of these priors amongst the neighboring pixels, the intrinsic configuration of the dictionary, or both. The author also proposes a novel arranged prior baptized as the low-rank (LR) assembly prior, that can be deliberated as a alteration of the LR prior. Besides, differently regulated priors to advance the outcome for the HSI classification is also being investigated [67]. Again as table 2 depicts one of the approaches to support the pixel-wise classification based on the sparse representation classifier where it outperforms the traditional classifiers.

TABLE II. Differentiation Between The Spatial-Spectral Techniques Based On Pixel-Wise Classifiers

<table>
<tr><td colspan="4">Pixel-wise classifiers
↓
Spatial-spectral techniques</td></tr>
<tr><td rowspan="2">Extract spatial features
(incorporate both)
↓
Spatial and spectral features
(techniques incorporated)
↓
• Mean and median filters.
• Gabor wavelets.
• Morphological profiles.
• Mean shift (MS).
(after extracting spatial features)
↓
• Composite kernals (CK's).
• Generalized CK's.
• Multiple kernals.
(considering spatial features in addition to spectral features)
↓
Used to perform Final Classification</td><td colspan="3">Integrate spatial and spectral information separately.</td></tr>
<tr><td>Probabilistic SVM

First estimate the class conditional probability density function.
↓
Then the context-based class priors is estimated using Markov Random Fields (MRF).</td><td>Bayesian Framework

Uses MLR to learn posterior probability distributions from spectral information contained in the data and uses MRF to include spatial contextual information in the classifiers.</td><td>Original HSI

First classified by a pixel-wise classifier and simultaneously segmented into some adaptive neighbors by using some segmentation techniques
(such as):
• Partitional clustering.
• Morphological body-sheds.
• Minimum spanning forests.

Majority voting adopted integrating pixelwise classification results and the obtained segmentation map.</td></tr>
</table>

D. *The Outperformance of Classifier in both Quality and Quantity*

In remote sensing instrument and measurement, analysis of HSI is an important topic having wide applications in fields such as skin imaging, identification of ground elements and exploration of minerals. The electromagnetic wave reflected from materials is used to capture and record HSI spanning from visible to the infrared spectrum. For a single pixel, the spectral information for HSI is epitomized by a vector whose entities corresponds to those values of spectral bands that are different. For the pixels, the class labels are determined by spectral information of the undeviating variances of pixels. The spectral evidence in the previous work is utilized by the previous work for HIS classification without considering spatial contexts hence resulting in the classification maps with noisy appearance. For HSI, classification performance is enhanced attempting to jointly exploit spatial-spectral information. A region-based sparse representation is proposed that assumes that pixels within local region belong to the same class or the same material. For each test pixel, a fixed-size region is first defined in the region-based sparse representation, then inside this expanse, the pixels are mutually disintegrated on the identical dictionary atoms. An approach is highlighted in table 3 regarding the outperformance of the classifier in both quality and quantity.Sparse representation classification (SRC) [35] outline undertakes the pixel ẍ of a particular class that can be embodied by a lined amalgamation of atoms selected from a dictionary $Đ = [Đ1, \ldots, Đ\,c, \ldots, Đ\,C] \in R^{NXM}$. C is defined as the quantity of classes and $M = \sum_{c=1}^{C} M_c$ is the total number of samples for training. The sub-dictionary $Đ_c \in R^{NXM_c}$ is created by unswervingly extracting pixels of the cth class in unique HSI. Then, subsequently directing the identical process, one can acquire solitary subdictionary for every class, besides all these subdictionaries are established as the concluding dictionary $Đ = [Đ\,1, \ldots, Đ\,c, \ldots, Đ\,C]$. Then ẍ can be sporadically represented as follows:

$$\ddot{\mathbf{x}} = Đ\,\alpha + \varepsilon \tag{7}$$

where $\alpha \in R^{MX1}$ defines the sparse coefficients for the ẍ and ε is an inaccurate residual item. The value of α is attained by resolving the subsequent optimization problem:

$$\hat{a} = \arg \min_{\alpha} \|\ddot{x} - Đ\boldsymbol{\alpha}\|_2 \quad \text{s.t. } \|\alpha\|_0 \le K_0 \tag{8}$$

where $K0$ is the level of sparsity, which is equivalent to the superior bound quantity of nonzero rows in ậ. Once ậ is acquired, the class label of the assessment pixel can be resolute by associating the error of reconstruction of every class

$$\text{class}(\ddot{\mathbf{x}}) = \arg \min_{c=1,\ldots,C} r^c(\ddot{\mathbf{x}}) = \arg \min_{c=1,\ldots,C} \|\ddot{x} - Đc\hat{a}_c\|_2 \tag{9}$$

where $\hat{a}_c$ is the sparse coefficient subset of ậ that belongs to c-th class. For HSI, the complex spatial information cannot be sufficiently exploited ever since the magnitude of the designated region is stable. Also for the HSI classification, some recent work combines extracted features with the classifiers. In [24] the author mined linear as well as non-linear features developing a mechanism for multiple feature learning. Though an outstanding performance is achieved by the multiple-feature-based classifier, still few issues need to be addressed. One of the issues is multiple-feature-based classifiers are applied only to the fixed window on the feature and the exploitation on the HSI features for spatial correlation cannot be sufficient [23]. In the classifiers based on multiple-feature, the correlations among different features are not well considered [25]. In this paper, MFASR process is projected to address the above two issues [32]. This method utilizes the adaptive sparse representation exploiting the correlations effectively midst four features that are extracted. For the HSI features to utilize the spectral-spatial information fully a shape-adaptive region is adopted for the pixel in every feature. The proposed MFASR demonstrated the superiority over several well-known classifiers demonstrated by the experimental results on three real HSIs both in terms of quantitative and qualitative measures. For the HSI four features were empirically selected in this paper. For the HSI an automatic strategy for selection of feature is to be considered in the future work to enrich the performance of HSI classification for MFASR [52]-[55].

E. *Small Number of Labelled Samples*

For classification purposes, to augment and increase the size of training samples, semi-supervised [24] and domain adaptation/transductive learning [56], approaches are proposed in the literature of remote sensing. To improve the classifier systems performance active learning methods is proposed that uses fewer training samples [57]. Using the labeled samples that are available possibilities are offered by the new samples for training, probably semi-supervised learning (SSL) [56]. The window structure is created using a blend of both spatial as well as spectral statistics i.e. proposed via the integration of contextual information from the neighborhood pixels. Better performance and promising results are shown by the proposed framework even when the quantity of initially labeled samples are trivial [62].

F. *Exploitation of Spatial Features*

In the process of classification, the fundamental role is played by the exploitation of spatial features. To the surrounding pixels are generally related spatial characteristics of a pixel. The sliding window helps in defining fixed spatial neighborhood for the extraction techniques of spatial features [47] or these techniques are known for creating adaptive three-dimensional district to the relating parameter value of the considered filter [39], therefore computing spatial measures within the regions obtained. To extract the spatial information relating the hyperspectral images RMSR (random multiscale representation) technique is used that includes a spatial-spectral method of classification. The dense scales are used by the RMSR technique that represents spatial characteristics of HSI, so that complementary information can be benefited collecting various scales capturing spatial evidence nearby the pixel. The spatial features are mined by the computing principles on the quadrilateral scales that are dense. In the process of keeping all the information being extracted spatially, the spatial features obtained on dense scales are concatenated as rich-dimensional multiscale features. To compress the spatial features of high dimension without any loss of salient features into the lower dimensional features, actual sparse random matrix of measurement is familiarized. With the help of computing criteria, a complete process of RMSR is accomplished originally at random scales of the random bands according to nonzero entries of identical sparse measurement matrix.

In RMSR, criteria are investigated that includes first two moments, which are simple and at the same time can be meaningfully enhanced by the integral technique of the image. A load of computation becomes light because of these effective methods. The CK-based approach conducts the final classification that weights appropriately the spatial features with regard to spectral features. The two widely used hyperspectral datasets are tested for the proposed method [65]. The actual matrix of sparse random measurement [35] is given below that will be utilized in RMSR. In random projection, a rich-dimensional feature vector $a \in R^m$ is anticipated to a inferior dimensional feature vector $v \in R^n$ by the use of a random matrix $R \in R^{nxm}$ as follows:

$$v = Ra \qquad (10)$$

wherein the columns of **R** is having entity length, and n is very less than m. Precisely, for an integer d and let $n \geq n0 = O(\varepsilon^{-2}\ln(d))$ with ε is greater than 0, for whichever dual vectors $\mathbf{a}_i$, $\mathbf{a}_j$ in a restricted collection X of d vectors in R^m, we have

$$(1-\varepsilon)\,\|a_i - a_j\|_2^2 \leq \|Ra_i - Ra_j\|_2^2 \leq (1+\varepsilon)\,\|a_i - a_j\|_2^2 \qquad (11)$$

Since the random Gaussian matrix **R** is dense, involving substantial reminiscence in addition to calculation necessities wherein m is hefty. For incapacitating this difficult, a sparse random matrix **R** needs to be presented with its entries given as follows:

$$\mathbf{R}_{ij} = \sqrt{p}\ \mathrm{X} \begin{Bmatrix} 1 & \text{with probability } \frac{1}{2\rho} \\ 0 & \text{with probability } 1-\frac{1}{\rho} \\ -1 & \text{with probability } \frac{1}{2\rho} \end{Bmatrix} \qquad (12)$$

TABLE III. Approach To Show The Outperformance Of The Classifier In Both Quality And Quantity (2017)

Approach	Advantages	Challenges	Future Work
A multiple feature-based adaptive sparse representation (MFASR) method is proposed.	1) MFASR approach is better than classification methods it is being compared to in terms of both visual quality on classification maps and quantitative metrics OA, AA and Kappa and is highly competitive with MASR on OA. 2) Excellence in classification accuracy for only 1% training samples, being superior to other classifiers. 3) On multiple features, the performance of the method is better than one based on the spectral feature. 4) From feature images, dictionaries are constructed directly from feature images where the size of the dictionary is same as a number of training samples. 5) Feature dictionary directly extracting pixels and hence the size of the dictionary being large leads to high computational cost.	Performance deteriorated by very large sparsity level due to the fact that final classification is misled by some dictionary atoms selected from other class when sparsity level is very large.	To examine effective dictionary learning algorithm, future work concentrates on more representative feature dictionaries.

wherein $\rho = 3,^2$ and verified that if $n \geq (4+2\beta)(\varepsilon^2/2 - \varepsilon^3/3)^{-1}\ln(d)$ with $\beta > 0$, the proclamation (12) embraces accurateness through likelihood of at least $1 - d^{-\beta}$. The matrix R is easy to compute by the use of a unvarying random originator. The adaptive spatial vicinity is used to advance approaches for classification of HSI. The adaptive neighborhood area is used to achieve segmentation results. After surveying, as innovation these fundamental methods of segmentation (unsupervised) can be then widespread to HSI remote sensing. Before getting into the details of the proposed solution it is important to define the three basic classes of image segmentation which is discussed in figure 4. As depicted in the figure below, the criterion to define image segmentation is spectral and spatial domain dependent which is then further defined based on margin, group and threshold with marked or grouped. Then a new approach is developed merging spatial and spectral statistics of Hyperspectral classification. The segmentation map is used as spatial information while classification map as spectral information in this novel approach as shown in figure 5. There are different segmentation method that is generalized to HSI analysis which includes Robust Color Morphological Gradient (RCMG), Expectation Maximization (EM) etc. while their extensions were applied in the empirical implementation that includes Hyperspectral Robust Color Morphological Gradient (HRCMG), Adequate Expectation Maximization (AEM) etc. Indian Pines and Hekla are the two available hyperspectral data sets on which experiments were applied. The three analysis measurements, their classification maps with pixelwise approaches plus spatial-spectral approaches previously defined that includes EMP and ECHO were used to compare the experimental results. The results clearly indicate better performance in quantitate quality measures than the other reviewed approaches but indicating that for the proposed approach classification map is so artificial in some cases.

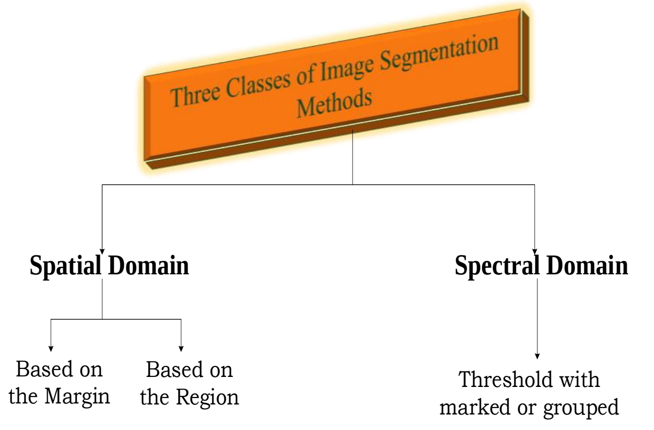

Fig. 3. Classes of Image Segmentation

Compared to the elder schemes improved accuracy is shown by the novel segmentation methods (HRCMG, AEM, and HRHSEG), when employing median voting scheme. The proposed approach is found to be fit and it preserves altogether spectral and spatial statistics appropriately. The projected approach is fast, accurate and simple. The future work includes introducing the three major approaches for processing of spectral-spatial simultaneously: the fusion of spatial-spectral statistics at pixel feature vector, Object-oriented classification and pixel classification and fusion of segmentation maps [66].

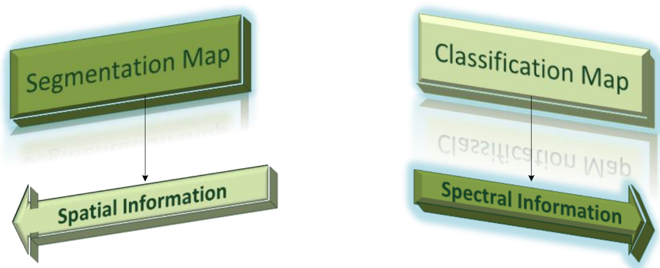

Fig. 4. Approach used to excavate the Spectral Information (2016)

G. Operative Feature Abstraction and Denoising of HSI data (supervised classification)

Intended for many applications, applying the techniques of denoising directly, probably removes fine features when removing noise which is undesirable. Therefore, multiscale wavelet transform-based methods for image denoising are widely used avoiding this problem, where at different decomposition levels the image is putrefied into a set of wavelet quantities using the transform and hence removing noise existing in the low-energy channels of the transformed domain. The analysis of the data of the hyperspectral images is of prime importance and the technique efficient for its analysis is given in figure 9, where both the previous and proposed method has been given with the interest of differentiation among them according to various properties. Though there exist a wide variety of approaches to analyzing the hyperspectral data but the widely used is the wavelet transform. It is one of the oldest methods for the analysis of the hyperspectral data. Since it is not possible to discuss in detail all the reformed approaches, hence a general approach has been discussed in the form of comparative structure between the wavelet transform and curvelet transform briefing the researcher with its effective analysis indexing.

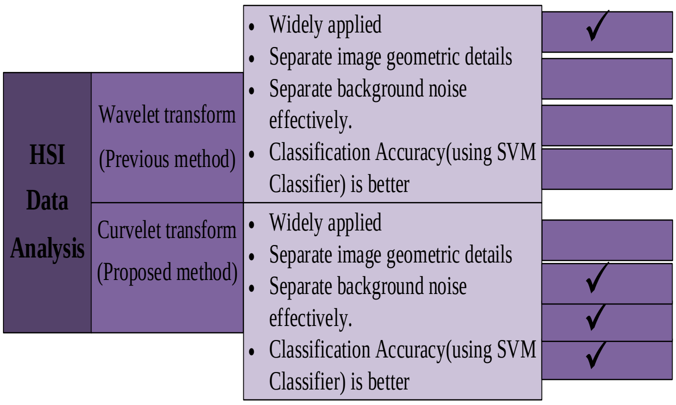

Fig. 5. Methods to analyze the data of hyperspectral images

A popular research topic in the recent years is the HSI classification, and before this classification task, an important task is the extraction of features effectively. In other words, feature extraction can be defined as the process of transforming original features into a well-defined new set of features. Traditionally, to the HSI data cube, the extraction techniques for spectral features is applied directly. Feature extraction is an important phenomenon associated with hyperspectral images in the field of image processing. In the procedure of feature extraction, from the original bands, new features need to be learned, wherein losing the physical significance of the bands but in turn containing information that is more discriminant for the process of classification. While utilizing the label information the methods of feature extraction is divided into three methods namely: unsupervised, supervised, and semi-supervised. During the process of learning, both the supervised and semi-supervised methods are in need of labeled samples labeled pixel. On the other hand, the unsupervised approaches have been attracted in the recent years since they entertain no requirement of label information. Being a vast field, the involving some amount of artificial work to be done on the complete illustration of the process of feature extraction cannot be defined, hence the popular and most recent work done in accordance with the same is given diagrammatically below considering one of the most prominent criteria i.e. reduced dimensionality, further bifurcating the process. Hence, if reduced dimensionality is provided in either supervised or unsupervised classification, then the various approaches satisfying the functionality are listed below some of which includes approaches like principal component analysis and many more. The other is based on the supervised feature extraction excluding the involvement of dimensionality reduction. An innovative algorithm is recommended in the paper [67] for the extraction of HSI features that exploit the curvelet transformed domain by means of a comparatively new spectral feature dispensation technique, i.e. singular spectrum analysis (SSA). On the spectral dimension in order to extract features, the proposed approach relies mainly on SSA, since it belongs to the category of spectral feature extraction. To show its efficacy, the method proposed is associated to some state-of-the-art technique used for spectral feature extraction.

The proposed method, in addition, is known to prove that it is capable to eliminate the unwanted relics that are presented through the process of data attainment. Also using the proposed approach, an addition of an extra step of spatial preprocessing is achieved to the classified map, where the performance of classification is comparable to several methods of recent spectral–spatial classification [57].

H. Allocate an Exclusive Label to each pixel vector

An active area of research for HSI classification is the application of hyperspectral data analysis, in the last two decades [38] [39]. The goal of classification for a given a set of observations is to dispense an exclusive label to every pixel vector. The supervised classifiers along with many spectral-based techniques have been employed to tackle this problem, including spectral band analysis, k-nearest neighbors, SVM, and SRC. To the HSI, set of 2-D Gabor filter is applied first extracting discriminative features. Gabor filters [43] are applied band by band to hyperspectral images imitating the associations amongst surface materials in the spatial domain. Precisely, a series of 2-D Gabor filters through altered scales and alignments is premeditated to excerpt Gabor features

$$\Psi_{f_u, \theta_v}(x,y) = \exp(-\Pi(a^2x'^2 + b^2y'^2)) \,.\, \exp(j2\,\Pi f_u x'^2)$$
$$x' = x\cos\theta_v + y\sin\theta_v,\ y' = -x\sin\theta_v + y\cos\theta_v$$
$$a = 0.9589 f_u,\ b = 1.1866 f_u \qquad (18)$$

where *fu* and *θv* denote the central frequency and orientation of the Gabor filter, respectively. Subsequently the Gabor filters have been gained, denoted by $\{ \Psi_t, t = 1, 2, \ldots, T \}$ for expediency (T is the quantity of the Gabor filters), Gabor features are mined by convoluting each spectral band of Approach followed to assign a unique label to every pixel vector (2017) HSI with every Gabor filter, namely

$$\mathbf{G}t\,(x, y, b) = |(R_b \otimes \Psi_{f_t})|(x,y)| \qquad (19)$$

where $\mathbf{R}b \in R^{XxY}$ is the bth band, $1 \le b \le B$, and $\otimes$ and $|\cdot|$ defined as convolution and absolute operators, correspondingly. Apparently, $\mathbf{G}t\,(x, y) = [\mathbf{G}t\,(x, y, 1); \mathbf{G}t(x, y, 2); \ldots; \mathbf{G}t\,(x, y, B)] \,\varepsilon\, R^B$ is the retort of the tth Gabor filter at all bands. Through applying Ψ_t on all pixels of the hyperspectral images, a Gabor cube $\mathbf{G}t \,\varepsilon\, R^{XxYxB}$ can be obtained, which has the same size as the original hyperspectral data. Further, after each Gabor filter Ψ_t $t = 1, 2, \ldots, T$ has been convolved with the HSI , a total of T Gabor cubes $\mathbf{G}t$, $t = 1, 2, \ldots, T$ was extracted.

From the HSI a superpixel map is then generated. Then the super-pixel based spatial-spectral Schroedinger eigenmaps (S4E) method is then adopted reducing the dimensions of each Gabor cube being extracted. Using the SVM- based multitasking learning framework, the classification is finally carried out [10].

I. Compression

In Geoscience and Remote Sensing, the advancement in data acquisition technologies leads to the growth of spatial and spectral resolutions with the frequencies of obtaining data. Up to hundreds of several bands are covered by modern hyperspectral sensors having both high spatial as well as spectral resolutions. A comparative analysis of the two different type of compression i.e. lossy and lossless compression difference is clearly stated with a plot of comparison on various parameters below in table 4.

The analysis is done on various parameters to provide an overview to the researcher giving clarity on the parameters to differentiate the well-known compression i.e. lossy as well as lossless compression. The parameters includes the efficiency in compression and the effectiveness in terms of the technology applied. Furthermore, the selective loss in information is also the well-known parameter considered followed by the reduction in the data reduction, then the application where the loss of information is caused by certain types or levels of distortion and is insensitive to the system of human visual and hearing etc.

TABLE IV. Comparison Between Lossy And Lossless Compression

Comparison Factor	Lossy Compression	Lossless Compression
Improves the efficiency of compression and is an effective technology.	✓	
Selective information loss.	✓	
Application where the loss of information is caused by certain types or levels of distortion and is insensitive to the system of human visual and hearing.	✓	
Greater Data Reduction.	✓	
For applications using HSI demanding accuracy.		✓
For computers to analyze images automatically.		✓
No loss of the data to be reconstructed.		✓

An essential feature of the system that incorporates hyperspectral sensors is data compression owing to the huge difference amongst lossy and lossless compression volume of data collected by the sensors that need to transmitted or stored. The heart of the problem is to find an appropriate model in the hyperspectral data compression [43]. In remote-sensing, current sensors collect information in large amounts to be readily transmitted onto the ground having a limitation of limited memory capacity [47]. A large amount of in the spatial, spectral, and temporal domain of the earth surface is provided by the remote sensors. It must meet the requirement of a wide range of important applications providing frequent and at the same time fine coverage of large areas. For current storage and transmission systems, a tough challenge was the increment in the number of high-resolution sensors. For instance, images with upwelling spectral radiance in around 224 contiguous spectral channels having a wavelength between 400 and 2500 nm are delivered by the NASA Airborne Visible Infrared Imaging Spectrometer (AVIRIS) [41]. This clearly predicts that for remote sensing data the need for efficient coding techniques has become more and more imperative improving the capabilities of both storage and transmission. To this, a wide variety of solutions have been proposed solving the problem of large volume data, among which the inpainting finds its application in the field of HIS

TABLE V. A Comparative Analysis Of Popular Approaches For Compression In HSI

Approach	Assumption/ criterion	Advantages	Opportunities
PDE(partial differential equation) -based inpainting algorithm [12]	The separate inpainting of known data is done in spatial and spectral dimensions. Based on the fusion of both combination and predictive based method.	• Performance better than CCSDS-123.0. • As a spectral transform, it is better than JPEG 2000 Part 2 with DWT 9/7 at all bit rates. • Competitive to JPEG 2000 with PCA.	For compression ratios from moderate to low, these lag behind prevalent behavior compared to prediction-based and transform-based approaches.
Lossy hyperspectral data compression framework based on sparse representation [13]	Online dictionary learning algorithm, orthogonal matching pursuit (OMP): input data representation, final bit stream is formulated by applying quantization and entropy coding.	• The online dictionary helps to represent hyperspectral data sparsely incorporating correlation both spectrally and spatially. • The compression is offered with the benefit of rate-distortion performance competitive to Wavelet transform-based methods.	Beside compaction feature for energy, other features are embodied by sparse coefficient, e.g., discriminative, energy compaction feature, exploited in other hyperspectral data processing tasks. In dictionary learning, some challenges need to be addressed.
Patch-based low-rank tensor decomposition (PLTD) [HSI compression and reconstruction algorithm][27]	Each local patch represented: Third-order tensor (for HSI). For nonlocal similarity, similar tensor patches grouped by cluster forming fourth-order tensor per cluster. Grouped tensor assumed to be redundant.	• Redundancy in spectral and spatial domain removed and outperforms traditional image compression approaches; other tensor-based methods. • Fully Reserves neighborhood relationship; global correlation. • Simplification in obtaining reconstructed HSI.	The algorithm can be improved further by finding the optimal dimensions automatically making the algorithm more practical and at the same time effective for the compression of HSI.
Rate control algorithm [integrated into lossy extension to the CCSDS(Consultative Committee for Space Data Systems) -123 lossless compression recommendation] [14]	Lossy predictive coding operated in a so-called near-lossless mode, bounding maximum absolute error by constant on pixels being reconstructed. Gives variable output rate becoming drawback of near-lossless compression that depends on image content	• Fast and simple to implement providing the same accuracy for output rate and quality of images. • Overcomes drawback of limited downlink bandwidth. • Most popular solution meeting requirements of low-complexity, high-throughput and excellent performance for rate-distortion. • Overcomes complexity limitations that onboard compression algorithms face due to constraints on hardware.	Some flexibility is lost in terms of spatial modulation of quantization step sizes.

where it focuses on filling missing information and dead pixels occurred because of sensor failures. To it, a novel PDE-based inpainting algorithm is proposed compressing the HSI [42]. Being rich in the number of solutions proposed for the terminology associated with HSI compression, the literature remains wide open to be discussed in one single article, hence a summary of different and current compression approaches has been prepared to benefit the reader as discussed in table 5.

III. ADDITIVE CHALLENGES WITH A BRIEF DEFINE PROBLEM AND AN APPROPRIATE SOLUTION

As mentioned earlier in the article, the field of image processing is wide and so is the hyperspectral imagery. The article tends to elaborate the various challenges associated with the hyperspectral images. Moreover, to each challenge an appropriate solution is given with a condition that it may not be depending on the application desired or the required functionality desired by the end user. This, in turn, is a motivation to the researcher to work on one of the problem domain defined. In the article and open new area of research with a well-defined solution to them. Being such wide area not all the problem from the very beginning can be mentioned. Hence to wind up the article in favor of the researcher a short summary of few other problem is defined in table 6.

III. CONCLUSION

To find a research problem in HSI is a cumbersome task. To privilege the researcher in finding an appropriate research problem, this paper is a small contribution that not only defines the challenges associated with hyperspectral images but also an outlook to the approach or approaches that have been previously employed to solve that particular challenge. An addition to it is the analysis of that one or more approaches to determine its advantages, disadvantages, the challenges occurring in that particular approach and future work. The simplicity of the article is maintained by not getting into the detailing as discussing even a single challenge mentioned above can be turned into a complete review itself. Hence various challenges are defined by the classification based issues to the compression and foremost the last section of the

TABLE VI. Additive Issues In Hsi Along With The Proposed Approaches To Help Resolve Them

Challenges	Approach Followed	Advantage
Identify constituent spectra plus the estimation of its fractional abundance from mixed pixels.	Novel framework coupling sparse hyperspectral unmixing with abundance estimation error reduction [30].	• Suppress abundance estimation error. • Improves the unmixing accuracy.
Estimate number of end members (NOE).	A method proposed based on statistics of IDD (Indegree distribution dimensionality) of data nearest neighbor graph [31].	• IDD: High dependence on intrinsic dimensionality and skewed for increased dimensionality
HSI processing algorithms always based on the assumption of no spatial as well as spectral correlation in noise.	One effective solution is employing hyperspectral noise estimation which is based on regression residuals. Or use per-pixel noise estimator [32].	• Improvement in noise variance estimation compared to a classic residual method with the perk added in case of uncorrelated noise as well.
Expensive computation to sharpen HSI due to a large number of bands.	AATPRK (appropriate area-to-point regression kriging) transforming original HSI to new feature space [38].	• It expedites ATPRK inheriting its pros of maintaining similar performance in sharpening plus conserves spectral properties.
Hyperspectral anomaly detection	GSEAD (graphical scoring estimation based anomaly detector): The graphical data description are utilized achieving anomaly detection procedure based on data-adaptive analysis [39].	Achievements superior to other state-of-the-art methods for: ➢ Receiver operating characteristic curves. ➢ The area under ROC curves values. Background-anomaly separation
Hyperspectral face recognition is difficult due to factors: • Data acquisition. • Low signal-to-noise ratio. • Higher dimensionality.	For hyperspectral face recognition, the author initially compared five frequently used descriptors for 2D face recognition that already existed and then used CRC (collaborative representation classifier) with two voting techniques [40].	For PolyU-HSFD database: Gabor filter bank-based features are robust to both Gaussian white noise and shot noise achieving competitive classification results. For CMU-HSFD database: HOG (histogram of oriented gradients) yields good classification results for low noise level. High noise level: Raw facial images without feature extraction perform very well in term of correct classification rate. Facial image with no noise: The local binary pattern and HOG descriptor achieve good classification rates.

paper highlights the general challenges in the arena of HSI to endeavor the maximum of the benefit to the researcher. The novelty of the paper lies in easing the most crucial task in the era of research that is to find a research problem.

REFERENCES

1. Landgrebe, David. “Hyperspectral image data analysis.” IEEE Signal Processing Magazine 19, no. 1 (2002): 17-28.
2. Bryant, Jack. “On displaying multispectral imagery.” Photogrammetric Engineering and Remote Sensing 54 (1988): 1739-1743.
3. Di, Liping, and DONALDC RUNDQUIST. “Color-composite image generation on an eight-bit graphics workstation.” Photogrammetric Engineering and Remote Sensing 54 (1988): 1745-1748.
4. Durand, Jean M., and Yann H. Kerr. “An improved decorrelation method for the efficient display of multispectral data.” IEEE Transactions on Geoscience and Remote Sensing 27, no. 5 (1989): 611-619.
5. Lin, Qian, and Jan Allebach. “Displaying Multispectral Images On Video Ternunals In Rgb Color.” In Geoscience and Remote Sensing Symposium, 1990. IGARSS’90.’Remote Sensing Science for the Nineties’., 10th Annual International, pp. 357-360. IEEE, 1990.
6. Du, Qian, Nareenart Raksuntorn, Shangshu Cai, and Robert J. Moorhead. “Color display for hyperspectral imagery.” IEEE Transactions on Geoscience and Remote Sensing 46, no. 6 (2008): 1858-1866.
7. Jimenez, Luis O., and David A. Landgrebe. “Supervised classification in high-dimensional space: geometrical, statistical, and asymptotical properties of multivariate data.” IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 28, no. 1 (1998): 39-54.
8. Friedman, Jerome, Trevor Hastie, and Robert Tibshirani. The elements of statistical learning. Vol. 1. Springer, Berlin: Springer series in statistics, 2001.
9. Boser, Bernhard E., Isabelle M. Guyon, and Vladimir N. Vapnik. “A training algorithm for optimal margin classifiers.” In Proceedings of the fifth annual workshop on Computational learning theory, pp. 144-152. ACM, 1992.
10. Schölkopf, Bernhard, and Alexander J. Smola. Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT press, 2002.
11. S. Haykin, Neural Networks and Learning Machines, 3rd ed. Cambridge, MA, Prentice Hall, 2008.
12. Camps-Valls, Gustavo, and Lorenzo Bruzzone, eds. Kernel methods for remote sensing data analysis. John Wiley & Sons, 2009.
13. Bruzzone, Lorenzo, and Lorenzo Carlin. “A multilevel context-based system for classification of very high spatial resolution images.” IEEE transactions on Geoscience and Remote Sensing 44, no. 9 (2006): 2587-2600.
14. Pacifici, Fabio, Marco Chini, and William J. Emery. “A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification.” Remote Sensing of Environment 113, no. 6 (2009): 1276-1292.
15. Tuia, Devis, Fabio Pacifici, Mikhail Kanevski, and William J. Emery. “Classification of very high spatial resolution imagery using mathematical morphology and support vector machines.” IEEE Transactions on Geoscience and Remote Sensing 47, no. 11 (2009): 3866-3879.
16. Melgani, Farid, and Lorenzo Bruzzone. “Classification of hyperspectral remote sensing images with support vector machines.” IEEE Transactions on geoscience and remote sensing 42, no. 8 (2004): 1778-1790.
17. Camps-Valls, Gustavo, Luis Gómez-Chova, Javier Calpe-Maravilla, José David Martín-Guerrero, Emilio Soria-Olivas, Luis Alonso-Chordá, and José Moreno. “Robust support vector method for hyperspectral data classification and knowledge discovery.” IEEE Transactions on Geoscience and Remote sensing 42, no. 7 (2004): 1530-1542.
18. Fauvel, Mathieu, Jón Atli Benediktsson, Jocelyn Chanussot, and Johannes R. Sveinsson. “Spectral and spatial classification of hyperspectral data using SVMs and morphological profiles.” IEEE Transactions on Geoscience and Remote Sensing 46, no. 11 (2008): 3804-3814.
19. Foody, Giles M., and Ajay Mathur. “Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification.” Remote Sensing of Environment 93, no. 1 (2004): 107-117.
20. Y. Chen, N. M. Nasrabadi, and T. D. Tran, “Hyperspectral image classification using dictionary-based sparse representation,” IEEE Trans. Geosci. Remote. Sens., vol. 49, no. 10, pp. 3973–3985, Oct. 2011.
21. J. Li *et al.* , “Multiple feature learning for hyperspectral image classification,” IEEE Trans. Geosci. Remote Sens., vol. 53, no. 3, pp. 1592–1606, Mar. 2015.
22. E. Zhang, X. Zhang, H. Liu, and L. Jiao, “Fast multifeature joint sparse representation for hyperspectral image classification,” IEEE Geosci. Remote Sens. Lett., vol. 12, no. 7, pp. 1397–1401, Jul. 2015.
23. Du, Bo, Mengfei Zhang, Lefei Zhang, Ruimin Hu, and Dacheng Tao. “PLTD: Patch-based low-rank tensor decomposition for hyperspectral images.” IEEE Transactions on Multimedia 19, no. 1 (2017): 67-79.
24. A. Chakrabarti and T. Zickler, “Statistics of real-world hyperspectral images,” in Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, 2011, pp. 193–200.
25. J. Holloway, T. Priya, A. Veeraraghavan, and S. Prasad, “Image classification in natural scenes: Are a few selective spectral channels sufficient?” in Proceedings of IEEE International Conference on Image Processing, 2014, pp. 655–659.
26. Wang, Rui, Heng-Chao Li, Wenzhi Liao, Xin Huang, and Wilfried Philips. “Centralized Collaborative Sparse Unmixing for Hyperspectral Images.” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 10, no. 5 (2017): 1949-1962.
27. Bachmann, Charles M., Thomas L. Ainsworth, and Robert A. Fusina. “Exploiting manifold geometry in hyperspectral imagery.” IEEE transactions on Geoscience and Remote Sensing 43, no. 3 (2005): 441-454.
28. Tenenbaum, Joshua B., Vin De Silva, and John C. Langford. “A global geometric framework for nonlinear dimensionality reduction.” Science 290, no. 5500 (2000): 2319-2323.
29. Schölkopf, Bernhard, Alexander Smola, and Klaus-Robert Müller. “Nonlinear component analysis as a kernel eigenvalue problem.” Neural computation10, no. 5 (1998): 1299-1319.
30. Roweis, Sam T., and Lawrence K. Saul. “Nonlinear dimensionality reduction by locally linear embedding.” Science 290, no. 5500 (2000): 2323-2326.
31. Silva, Vin D., and Joshua B. Tenenbaum. “Global versus local methods in nonlinear dimensionality reduction.” In Advances in neural information processing systems, pp. 721-728. 2003.
32. Saul, Lawrence K., and Sam T. Roweis. “Think globally, fit locally: unsupervised learning of low dimensional manifolds.” Journal of Machine Learning Research 4, no. Jun (2003): 119-155.
33. Zhang, Zhenyue, and Hongyuan Zha. “Principal manifolds and nonlinear dimensionality reduction via tangent space alignment.” SIAM journal on scientific computing 26, no. 1 (2004): 313-338.
34. Belkin, Mikhail, and Partha Niyogi. “Laplacian eigenmaps for dimensionality reduction and data representation.” Neural computation 15, no. 6 (2003): 1373-1396.
35. Agrafiotis, Dimitris K. “Stochastic proximity embedding.” Journal of computational chemistry 24, no. 10 (2003): 1215-1221.
36. Yan, Shuicheng, Dong Xu, Benyu Zhang, Hong-Jiang Zhang, Qiang Yang, and Stephen Lin. “Graph embedding and extensions: A general framework for dimensionality reduction.” IEEE transactions on pattern analysis and machine intelligence 29, no. 1 (2007).
37. X. Jia and J. A. Richards, “Cluster-space representation for hyperspectral data classification,” IEEE Trans. Geosci. Remote Sens., vol. 40, no. 3,pp. 593–598, Mar. 2002.
38. J. A. Gualtieri and R. F. Cromp, “Support vector machines for hyperspectral remote sensing classification,” Proc. SPIE, vol. 3584, p. 221,Oct. 1998.
39. C.-H. Li, H.-S. Chu, B.-C. Kuo, and C.-T. Lin, “Hyperspectral image classification using spectral and spatial information based linear discriminant analysis,” in Proc. IEEE Int. Symp. Geosci. Remote Sens., Jul. 2011, pp. 1716–171 9.
40. J. Li, M. Bioucas-Dias, and A. Plaza, “Spectral–spatial classification of hyperspectral data using loopy belief propagation and active learning,” IEEE Trans. Geosci. Remote Sens., vol. 51, no. 2, pp. 844–856, Feb. 2013.

41. J. Yin, Y. Wang, and J. Hu, "A new dimensionality reduction algorithm for hyperspectral image using evolutionary strategy," IEEE Trans. Ind. Informat., vol. 8, no. 4, pp. 935–943, Nov. 2012.
42. J. Yin, C. Gao, and X. Jia, "Using hurst and Lyapunov exponent for hyperspectral image feature extraction," IEEE Geosci. Remote Sens. Lett., vol. 9, no. 4, pp. 705–709, Jul. 2012.
43. K. Kavitha, S. Arivazhagan, and N. Kayalvizhi, "Wavelet based spatial—Spectral hyperspectral image classification technique using support vector machines," in Proc. Int. Conf. Comput. Commun. Netw. Technol. (ICCCNT), Jul. 2010, pp. 1–6.
44. C. Chen, W. Li, E. W. Tramel, M. Cui, S. Prasad, and J. E. Fowler, "Spectral–spatial preprocessing using multihypothesis prediction for noise-robust hyperspectral image classification," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 7, no. 4, pp. 1047–1059, Apr. 2014.
45. Li, Feng, Lei Xin, Yi Guo, Junbin Gao, and Xiuping Jia. "A framework of mixed sparse representations for remote sensing images." IEEE Transactions on Geoscience and Remote Sensing 55, no. 2 (2017): 1210-1221.
46. J. J. Heckman, "Sample selection bias as a specification error," Econometrica, vol. 47, no. 1, pp. 153–161, Jan. 1979.
47. B. Zadrozny, "Learning and evaluating classifiers under sample selection bias," in Proc. 21st Int. Conf. Mach. Learn., 2004, pp. 903–910.
48. R. G. Congalton and K. Green, Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. Boca Raton, FL, USA: CRC Press, 1999.
49. Persello, Claudio, and Lorenzo Bruzzone. "Active and semisupervised learning for the classification of remote sensing images." IEEE Transactions on Geoscience and Remote Sensing 52, no. 11 (2014): 6937-6956.
50. A. Plaza, J. Benediktsson, J. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, M. Marconcini, J. Tiltoni, and G. Trianni, "Recent advances in techniques for hyperspectral image processing," Remote Sens. Environ., vol. 113, no. S1, pp. S110–S122, Sep. 2009.
51. G. Camps-Valls, L. Gomez-Chova, J. Muñoz-Marì, J. Vila-Francés, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," IEEE Geosci. Remote Sens. Lett., vol. 3, no. 1, pp. 93–97, Jan. 2006.
52. L. Gomez-Chova, G. Camps-Valls, J. Muñoz-Marì, and J. Calpe-Maravilla, "Semi supervised image classification with Laplacian support vector machines," IEEE Geosci. Remote Sens. Lett., vol. 5, no. 3, pp. 336–340, Jul. 2008.
53. J. Zhu, S. Rosset, T. Hastie, and R. Tibshirani, "1-norm support vector machines," in Proc. NIPS, Dec. 2003, vol. 16, pp. 16–23.
54. J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 2, pp. 210–227, Feb. 2009.
55. J. Wright, J.Mairal, G. Sapiro, T. S. Huang, and S. Yan, "Sparse representation for computer vision and pattern recognition," Proc. IEEE, vol. 98, no. 6, pp. 1031–1044, Jun. 2010.
56. Y. Chen, M. Nasrabadi, and T. Tran, "Hyperspectral image classification using dictionary-based sparse representation," IEEE Trans. Geosci. Remote Sens., vol. 49, no. 10, pp. 3973–3985, Oct. 2011.
57. Sun, Xiaoxia, Qing Qu, Nasser M. Nasrabadi, and Trac D. Tran. "Structured priors for sparse-representation-based hyperspectral image classification."IEEE geoscience and remote sensing letters 11, no. 7 (2014): 1235-1239.
58. Z. He, Y. Shen, Q. Wang, and Y. Wang, "Optimized ensemble EMD-based spectral features for hyperspectral image classification," IEEE Trans. Instrum. Meas., vol. 63, no. 5, pp. 1041–1056, May 2014.
59. Z. He, Q. Wang, Y. Shen, J. Jin, and Y. Wang, "Multivariate Gray model-based BEMD for hyperspectral image classification," IEEE Trans. Instrum. Meas., vol. 62, no. 5, pp. 889–904, May 2013.
60. A. Prieto, F. Bellas, R. J. Duro, and F. López-Peña, "An adaptive approach for the progressive integration of spatial and spectral features when training ground-based hyperspectral imaging classifiers," IEEE Trans. Instrum. Meas., vol. 59, no. 8, pp. 2083–2093, Aug. 2010.
61. L. Fang, S. Li, X. Kang, and J. A. Benediktsson, "Spectral–spatial classification of hyperspectral images with a superpixel-based discriminative sparse model," IEEE Trans. Geosci. Remote Sens., vol. 53, no. 8, pp. 4186–4201, Aug. 2015.
62. Fang, Leyuan, Cheng Wang, Shutao Li, and Jón Atli Benediktsson. "Hyperspectral Image Classification via Multiple-Feature-Based Adaptive Sparse Representation." IEEE Transactions on Instrumentation and Measurement (2017).
63. J. Li, J. M. Bioucas-Dias, and A. Plaza, "Semisupervised hyperspectral image classification using soft sparse multinomial logistic regression," IEEE Geosci. Remote Sens. Lett., vol. 10, no. 2, pp. 318–322, Mar. 2013.
64. T. Oommen, D. Misra, N. K. C. Twarakavi, A. Prakash, B. Sahoo, and S. Bandopadhyay, "An objective analysis of support vector machine based classification for remote sensing," Math. Geosci., vol. 40, no. 4, pp. 409–424, 2008.
65. C. Persello and L. Bruzzone, "Active and semisupervised learning for the classification of remote sensing images," IEEE Trans. Geosci. Remote Sens., vol. 52, no. 11, pp. 6937–6956, Nov. 2014.
66. Q. Shi, B. Du, and L. Zhang, "Domain adaptation for remote sensing image classification: A low-rank reconstruction and instance weighting label propagation inspired algorithm," IEEE Trans. Geosci. Remote Sens., vol. 53, no. 10, pp. 5677–5689, Oct. 2015.
67. Z. Sun, C. Wang, D. Li, and J. Li, "Semisupervised classification for hyperspectral imagery with transductive multiple-kernel learning," IEEE Geosci. Remote Sens. Lett., vol. 11, no. 11, pp. 1991–1995,Nov. 2014.
68. D. Tuia, M. Volpi, L. Copa, M. Kanevski, and J. Muñoz-Marí, "A survey of active learning algorithms for supervised remote sensing image classification," IEEE J. Sel. Topics Signal Process., vol. 5, no. 3, pp. 606–617, Jun. 2011.
69. Aydemir, Muhammet Said, and Gokhan Bilgin. "Semisupervised Hyperspectral Image Classification Using Small Sample Sizes." IEEE Geoscience and Remote Sensing Letters 14, no. 5 (2017): 621-625.
70. G. Camps-Valls, L. Gomez-Chova, J. Mu˜noz-Mar´I, J. Vila-Franc´es, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," IEEE Geosci. Remote Sens. Lett., vol. 3, no. 1, pp. 93–97, Jan. 2006.
71. M. Fauvel, J. Chanussot, and J. A. Benediktsson, "A spatial–spectral kernel-based approach for the classification of remote-sensing images," Pattern Recog., vol. 45, no. 1, pp. 381–392, 2012.
72. Liu, Jianjun, Zebin Wu, Jun Li, Liang Xiao, Antonio Plaza, and Jon Atli Benediktsson. "Spatial–Spectral Hyperspectral Image Classification Using Random Multiscale Representation." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 9, no. 9 (2016): 4129-4141.
73. Borhani, Mostafa, and Hassan Ghassemian. "Hyperspectral spatial-spectral feature classification based on adequate adaptive segmentation." InIntelligent Systems (ICIS), 2014 Iranian Conference on, pp. 1-6. IEEE, 2014.
74. Qiao, Tong, Jinchang Ren, Zheng Wang, Jaime Zabalza, Meijun Sun, Huimin Zhao, Shutao Li, Jon Atli Benediktsson, Qingyun Dai, and Stephen Marshall. "Effective denoising and classification of hyperspectral images using curvelet transform and singular spectrum analysis." IEEE Transactions on Geoscience and Remote Sensing 55, no. 1 (2017): 119-133.
75. D. Manolakis, D. Mardon, and G. A. Shaw, "Hyperspectral image processing for automatic target detection applications," Lincoln Lab. J., vol. 14, no. 1, pp. 79–116, 2003.
76. J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. M. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," IEEE Geosci. Remote Sens. Mag., vol. 1, no. 2, pp. 6–36, Jun. 2013.
77. Jia, Sen, Bin Deng, Jiasong Zhu, Xiuping Jia, and Qingquan Li. "Superpixel-Based Multitask Learning Framework for Hyperspectral Image Classification." IEEE Transactions on Geoscience and Remote Sensing 55, no. 5 (2017): 2575-2588.

AUTHORS PROFILE

Kriti received B.E. degree (2012) from MBS College, Jammu and M.Tech degree (2015) from Chandigarh University, Punjab, India. She is currently pursuing Ph.D. Degree from Department of Computer Science and Engineering at Chandigarh University, Punjab, India. Her area of interest is hyperspectral image processing.

Dr. Urvashi Garg is working as an Associate Professor in the Department of Computer Science and Engineering at Chandigarh University, Mohali, India. She had authored several research papers, published at reputed conferences and journals.

Retrieval Number: B1763078219/19©BEIESP
DOI: 10.35940/ijrte.B1763.078219
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication