Clustering of Hyperspectral Images using Entropy based Multiple Features (Bands) Set Selection

Hitenkumar Motiyani, Quazi Sameed, Prashant Kumar Mali, Anand Mehta
Department of Civil Engineering
Institute of Infrastructure, Technology, Research And Management, Ahmedabad, Gujarat, 380026.
hitenkumar.motiyani.19c@iitram.ac.in

***Abstract*—This study suggests a novel segmentation-based clustering algorithm that applies entropy based local feature selection to choose the top bands for each cluster. A framework with numerous steps constitutes the proposed methodology. *k*-means is initially used as an image segmentation technique. Then, Shannon Entropy is utilized to determine the significant clusters from the generated segments. Finally, cluster map is obtained after merging insignificant clusters into significant clusters. Further, multiple feature set obtained through entropy are also utilized while performing clustering. The performance of the proposed methodology is examined using three sets of hyperspectral images. Adjusted Mutual Information and Overall Accuracy are used as evaluation criteria. The results of the study demonstrate that the proposed methodology performs better than the other segmentation methodologies that were evaluated. The results indicate that accuracy is much enhanced by selecting multiple feature set while performing clustering.**

***Index Terms*—Segmentation, feature selection, hyperspectral image, entropy, clustering.**

I. Introduction

Hyperspectral imaging technique has gained popularity in recent years as an intelligent and promising analytical tool for analyses carried out in research, control, and industry. The installation of imaging spectrometers in various locations is necessary for the acquisition of hyperspectral images. This spectrometers are used to capture images of electromagnetic waves in the ultraviolet, visible, near-infrared, and mid-infrared ranges [1]. Each pixel in the employed wavelength range can obtain a fully reflected or emitted spectrum since the imaging spectrometer can capture in a number of continuous and extremely tiny bands. Hence, great spectral resolution, numerous bands, and a wealth of information are hallmarks of hyperspectral images [2]. Hyperspectral remote sensing images are processed using a variety of techniques, the most common of which are image segmentation, feature selection, clustering, dimensionality reduction, and classification.

Cluster analysis or clustering is a challenging process as it needs determine which group a pattern will belong to in the absence of a label [3]. Several variables or traits may be considered appropriate for clustering. The situation might be made worse by the curse of dimensionality. High computational cost and high dimensionality both have an impact on the consistency of algorithms [4].

The advantages of improving learning performance, increasing computing efficiency, reducing memory storage, and creating stronger generalisation models are shared by feature extraction and feature selection [5]. They are both considered to be efficient dimensionality reduction approaches as a result. On the one hand, feature extraction is favoured in many instances where the input data lacks any properties that a particular learning algorithm can recognise. Nevertheless, because feature extraction generates a new set of characteristics, subsequent analysis is difficult because we are unable to preserve the physical meanings of these features. In contrast, feature selection improves the readability and interpretability of models by maintaining the physical meanings of some of the original features [6].

A specific image processing approach known as image segmentation is used to separate an image into two or more useful regions. The process of making boundaries between various semantic components in an image is known as image segmentation [7]. Image segmentation, from a more technical point of view, is the process of giving each pixel in the image a label so that pixels with the same label are related with respect to some visual or semantic attribute [8].

Based on the measurement, the image might have certain properties like that grey level, colour intensity, texture information, depth, or motion. The clustering method is used in image segmentation. Identifying groups of related images is the process of clustering in image segmentation [9]. There are numerous clustering algorithms that can be categorised in order to obtain the super pixel information. To achieve the right outcome with great efficiency, which affects storage picture, clustering approach is used [10].

This research's main contribution is entropy based local feature selection technique and clustering method. The proposed local feature selection technique takes into account both relevancy and redundancy concept while identifying top features (bands). Local feature selection in this instance refers to selecting a reduced set of band for each cluster. The rest of the paper is organised as follows. The methodology is described in Section II of the paper. Section III includes a list of the datasets and operations used. Section IV presents the study's findings. The paper is concluded in Section V.

II. METHODOLOGY

The proposed methodology is stated as follows: segmenting a hyperspectral image using *k*-means to obtain new segments. The following stage is identifying significant clusters using entropy. Combine the remaining clusters with the significant ones according to the minimum distance from significant clusters, taking into account significant segments with bands having least redundancy from the local band selection technique in the subsequent stage. Because HSI has a large number of bands or high dimensions. These bands have a lot of correlation. In other words, there is a lot of redundant information in bands. This leads us to believe that applying band selection method may improve accuracy. Fig. 1 provides a visual representation of the proposed methodology.

Shannon entropy is employed in methodology which basically is a measure of the uncertainty associated with a random variable. Specifically, Shannon entropy quantifies the expected value of the information contained in a message.

Basically, Shannon Entropy is represented by the equation:

$$\text{Shannon Entropy}(p) = -\sum_{i=1}^{J} p_i \log_2(p_i) \tag{1}$$

where, p_i is the pixel of *i*th cluster and *J* is the total number pixels in a cluster.

In this study, shannon entropy is employed to identify significant clusters by organising all segments in ascending order. This is because our hypothesis is that the lower the entropy, the more identical the pixel values are for a cluster. Additionally, bands with higher entropy values are arranged in descending order to identify the top bands in each cluster. This is done because we believe that higher entropy values indicate more band variability, which correlates to higher information content.

A. *Image Segmentation*

Image segmentation is the process of splitting up a digital image into several homogeneous subregions [4]. The goal of image segmentation is to collect pixels into meaningful image regions, or regions that correspond to things, or naturally occurring parts of objects, with high affinity and specificity [11]. By organising the dataset into a matrix using *k*-means, the process for segmenting an image is demonstrated. Whereas *k*-means is a clustering-based method. This approach looks for *k* divisions that satisfy a specific requirement while attempting to optimise the result [12]. First, choose a few pixels to serve as the first cluster centroid. Then, group the other pixels to their cluster centroids in accordance with the criterion of the minimal distance. Finally, we will acquire the initial classification. If the classification is inappropriate, however, we will alter it and repeat the process until we have a classification that is satisfactory [13]. Moreover, additional segments are produced following *k*-means segmentation that are organised using shannon entropy and are then used in band selection approach.

B. *Multiple feature (band) set selection*

The number of images (bands) in hyperspectral data gives it a very high dimensionality [14]. It's crucial to acquire the bands in a way that retains as much information as is feasible as an outcome. The band selection approach uses some informative measurements like entropy to determine the relative relevance of each spectral band. It then uses a ranking criteria to choose the top-ranked bands from a sorted sequence of more relevant and fewer redundant bands [5]. Moreover, bands with similar entropy values result in redundancy. In order to avoid this, a distance-weighted parameter score (γ) is introduced. The proposed band selection strategy is inspired by the methodology discussed in [6] and [15].

Formally, score (γ) is formulated in equation as:

$$\beta = \max_{ET_j > ET_i} d \tag{2}$$

$$\gamma = (\beta) \times (ET_i) \tag{3}$$

where, β is a function that store maximum distance for a test band from all bands whose entropy is higher than that of test band, *d* is distance between bands in a cluster to remaining bands of the same cluster. *ET* is the entropy for corresponding bands, *i* and *j* are band indexes. In this research, *ET* is used to determine the relevant bands while β is used to decrease redundancy and γ comprises both relevance and redundancy criteria. The following algorithm shows the steps for the proposed methodology.

Algorithm 1 Segmentation with multiple features (bands) set selection

```
Input: Hyperspectral Image.
Output: Clustered Image.
  • Using k-means clustering for segmentation. Generate the segmentation map from the cluster map. It will have multiple segments in it (group of pixels)
  • Every segment is now regarded as a cluster.
  • Sort every cluster according to increasing entropy.
  • Select the top segments that have more pixels than five because we assume that clusters with fewer than five pixels do not contain much information.
  • Calculate the shannon entropy and $\beta$ for all bands for each segment in a same manner. Higher entropy indicates more band variation, making it more significant.
  • Find the maximum distance between bands whose entropy is greater than the $j^{th}$ band and use that information to calculate the $\beta$.
  • Compute the $\gamma$ and arrange the segments according to the descending order of $\gamma$ values.
  • Evaluate the pairwise distance between the most significant and remaining segments while taking the bands that have been found for each significant cluster into account.
  • Labels should be distributed among clusters based on the shortest distance from significant clusters.
  • Assemble these segments into a cluster map now.
```

III. Experimental Setup

The stated approaches was programmed using Python through a personal machine with sixteen gigabytes of Random Access Memory (3200 MHz) and Intel i5 10th Generation

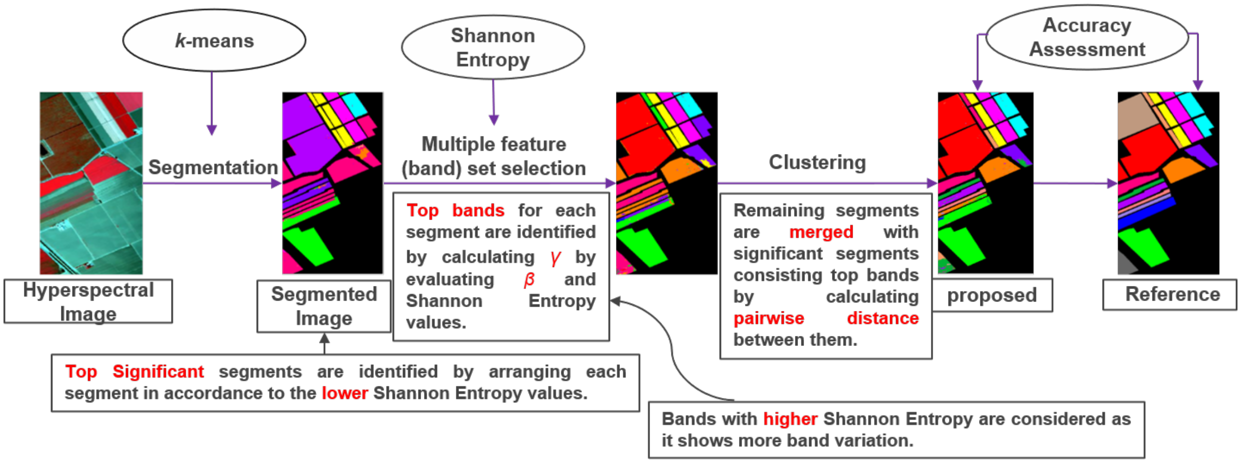
k-means
Shannon Entropy
Accuracy Assessment
Segmentation
Multiple feature (band) set selection
Clustering
Top bands for each segment are identified by calculating γ by evaluating β and Shannon Entropy values.
Remaining segments are merged with significant segments consisting top bands by calculating pairwise distance between them.
Hyperspectral Image
Segmented Image
proposed
Reference
Top Significant segments are identified by arranging each segment in accordance to the lower Shannon Entropy values.
Bands with higher Shannon Entropy are considered as it shows more band variation.

Fig. 1: Proposed Methodology

processor (4.30 GHz Max Turbo Frequency), with a four gigabytes of dedicated graphics memory of NVIDIA GTX 1650, operated on windows 11 [16].

This approach was evaluated using various hyperspectral data, and it was then compared at various phases of the proposed methodology. Firstly using k-means for image segmentation to generate new clusters, then utilising the multiple feature (band) set selection method while taking into account significant clusters with bands having minimal redundancy, combine the remaining segments with the significant clusters according to the minimum distance from significant clusters.

Nevertheless, a few additional approaches were used for comparison, as illustrated. Principal Component Analysis (PCA) with k-means was implemented where the information were normalised as well as to ascertain their associated covariance was discovered. Moreover, Simple Linear Iterative Clustering (SLIC) [17] is performed which suggests the similarity clustering of pixel space position and pixel colour feature to achieve a segmented image. Firstly, segments are created in SLIC and then, gathered together by defining entropy of each segment. Further, according to increasing entropy, each segment is arranged, as the segment having least entropy value shows more information. Likely, best segments are selected and rest of them are accumulated with the selected ones by obtaining mean of individual segments and calculating distances between all of them, respectively. Additionally, the proposed approach is also compared with two other hyperspectral clustering approaches, H2NMF [18] and CFSFDP [19]. Similar to the initial phase of the proposed method, segmentation comes before CFSFDP.

A. *Data Set*

The 204-band Salinas dataset was captured over the Salinas Valley in California using a 224-band Airborne Visible Infrared Imaging Spectrometer sensor. It distinguishes itself with a high spatial resolution of 3.7-m pixels. There are 512 lines and 217 samples in the covered region. Salinas ground truth includes 16 distinct forms of land use and cover [20].

The other datasets used in the tests are from Pavia Center and Pavia University. During a flight campaign over the northern Italian city of Pavia, the ROSIS (Reflective Optics System Imaging Spectrometer) sensor obtained these two images. Pavia University has 103 spectral information and nine classes, compared to Pavia Center's 102 spectral and nine classes, and its subset's eight. The original datasets from Pavia Centre and Pavia University are 1096 × 1096 and 610 × 610 pixels, respectively., respectively, are reduced here into 400 × 400 pixels and 610 × 340 pixels. Both datasets have 1.3 metres of geometric resolution.

B. *Accuracy Assessment*

The functions Adjusted Mutual Information (AMI) and Overall Accuracy, respectively, are used to evaluate the clustering [21]. For comparing clusterings in this context, adjusted mutual information, a kind of mutual information, may be employed. When the two partitions are identical, the AMI is 1, and when the MI between the two partitions is equal to the value that would be anticipated by chance alone, it is 0 [22].

Let $\mathbb{C}$ be the set of classes produced from the ground reference data, and let Ω be the set of clusters derived by the algorithm. Their mutual information, $\mathrm{MI}(\Omega, \mathbb{C})$ can be obtained as follows:

$$\mathrm{MI}(\Omega, \mathbb{C}) = \sum_{k} \sum_{j} p(\omega_k \cap c_j) \log_2 \frac{p(\omega_k \cap c_j)}{p(\omega_k) \cdot p(c_j)} \qquad (4)$$

where $p(\omega_k)$, $p(c_j)$ and $p(\omega_k \cap c_j)$ are probabilities that a randomly chosen pixel belongs to the cluster ω_k, class c_j and cluster ω_k as well as class c_j , respectively. Also, here n_{kj} denotes the number of objects that are common to these clusters.

When the two partitions have more clusters (with a fixed number of set elements N), the baseline value of mutual information between them tends to be higher and does not take on a constant value. By adopting a hypergeometric model of randomness, it can be shown that the expected mutual information between two random clusterings is:

$$\text{EMI}(\Omega,\mathbb{C}) = \sum_{k=1}\sum_{j=1}\sum_{n_{kj}=(a_k+b_j-N)^+}^{\min(a_k,b_j)} \frac{n_{kj}}{N}\log(\frac{N.n_{kj}}{a_k b_j}) \times \frac{a_k!b_j!(N-a_k)!(N-b_j)!}{N!n_{kj}!(a_k-n_{kj})!(b_j-n_{kj})!(N-a_k-b_j+n_{kj})!} \quad (5)$$

where, $(a_k$ + b_j - $N)^+$ denotes max(0, a_k + b_j - N). The variables a_k and b_j are partial sums of the contingency table; that is,

$$a_k = \sum_{j=1} n_{kj}$$

and

$$b_j = \sum_{k=1} n_{kj}$$

The adjusted measure for the mutual information may then be defined to be:

$$\text{AMI}(\Omega,\mathbb{C}) = \frac{\text{MI}(\Omega,\mathbb{C}) - \text{EMI}(\Omega,\mathbb{C})\}}{\max\{\text{H}(\Omega),\text{H}(\mathbb{C})\} - \text{EMI}(\Omega,\mathbb{C})} \quad (6)$$

where $\text{H}(\Omega) = -\sum_k p(\omega_k)\log_2 p(\omega_k)$ and $\text{H}(\mathbb{C}) = -\sum_j p(c_j)\log_2 p(c_j)$ are the entropies of Ω and $\mathbb{C}$, respectively.

While, the Overall Accuracy (OA), on the other hand, is determined by a contingency matrix that reports the intersection cardinality for every true/predicted cluster pair. When the samples are independent and have an identical distribution, it gives enough data for all clustering metrics, eliminating the need to take into consideration instances that are not clustered in all cases. OA presents it's results in percentage.

IV. Results And Discussion

Hyperspectral images are classified using image segmentation and band selection approaches while accounting for relevance and redundancy. The results of the proposed methodology (hereinafter referred as SegETMFS) were first examined with following clustering strategies: Seg_*k* and Seg_*k* + MFS_R.

However, in addition to the proposed approach, techniques like the multiple feature (band) set selection technique with redundancy are discussed. Let's say Seg_*k* denotes for segmentation using *k*-means, MFS_R represents multiple feature (band) set selection technique with redundancy factor not included and that *k*M stands for *k*-means clustering. Nonetheless, *k*_seg denotes the quantity of segments, *k*_cl denotes the quantity of significant clusters, and *L* denotes the quantity of top bands taken into account.

Furthermore, PCA + *k*M is employed, with the percentage of principle components (PC) chosen to guarantee that PCs can account for at least 99% of the variance in the dataset. The best segment values for the Salinas, Pavia Center, and Pavia University datasets in PCA + SLIC + *k*M are 120, 80, and 150, respectively.

A. *Salinas Dataset*

For proposed strategy, many initialised *k*_seg values are implemented in the range of 2 to 15 (as shown in Fig. 2(a)) and after fixing the value of *k*_seg i.e 7, according to the maximum AMI values, further, *k*_cl values were altered in a set of 16 to 30 and generated as 24 (as shown in Fig. 2(b)). Finally, after fixing *k*_seg and *k*_cl values, *L* was changed from 11 to 20 and was found out to be 13 (as shown in Fig. 2(c)). Thus, the accuracy attained is 84.31%.

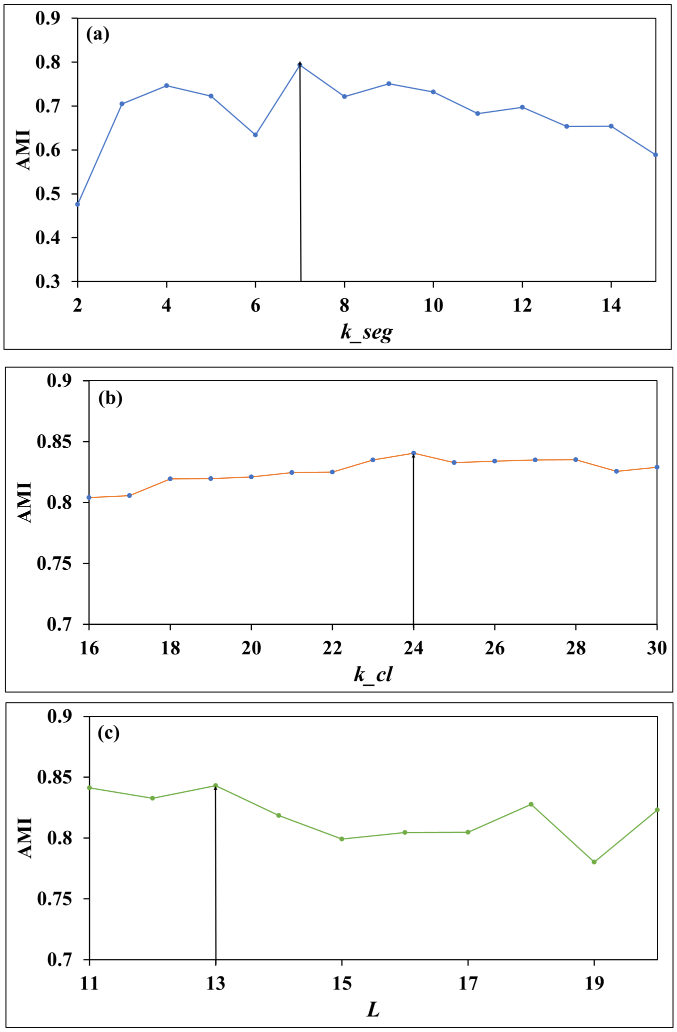

Fig. 2: Proposed strategy of Salinas Dataset

TABLE I: Salinas accuracy values.

	*k*_seg	*k*_cl	*L*	AMI	OA(%)
Seg_*k*	8	16	All bands	0.7408	67.95
Seg_*k* + MFS_R	6	25	10	0.8210	69.05
SegETMFS	7	24	13	0.8431	72.85

According to the Table I, carrying out similar procedure for other experiments as carried for proposed strategy, it is examined that accuracy is improved by 9.76% and 2.61%,

TABLE II: Assessment of Salinas, Pavia Center, and Pavia University's accuracy using several methods.

Dataset	Salinas			Pavia Center			Pavia University		
	*k*_cl	AMI	OA(%)	*k*_cl	AMI	OA(%)	*k*_cl	AMI	OA(%)
*k*M	16	0.7248	67.04	8	0.7694	79.28	9	0.5330	53.40
PCA + *k*M	16	0.7235	66.87	7	0.7791	79.75	8	0.5481	46.29
Seg_*k*	16	0.7408	67.95	8	0.7306	79.26	9	0.4604	67.85
Seg_*k* + MFS_R	25	0.8210	69.05	16	0.6785	73.89	11	0.4614	59.87
PCA + SLIC + *k*M	16	0.7506	65.46	8	0.6329	59.81	9	0.3100	36.51
H2NMF [19]	26	0.7041	70.41	8	0.7715	85.28	9	0.4721	43.75
CFSFDP [18]	26	0.7903	75.48	14	0.7829	83.20	9	0.5392	53.41
SegETMFS	24	0.8431	72.85	20	0.7903	90.25	13	0.5408	68.74

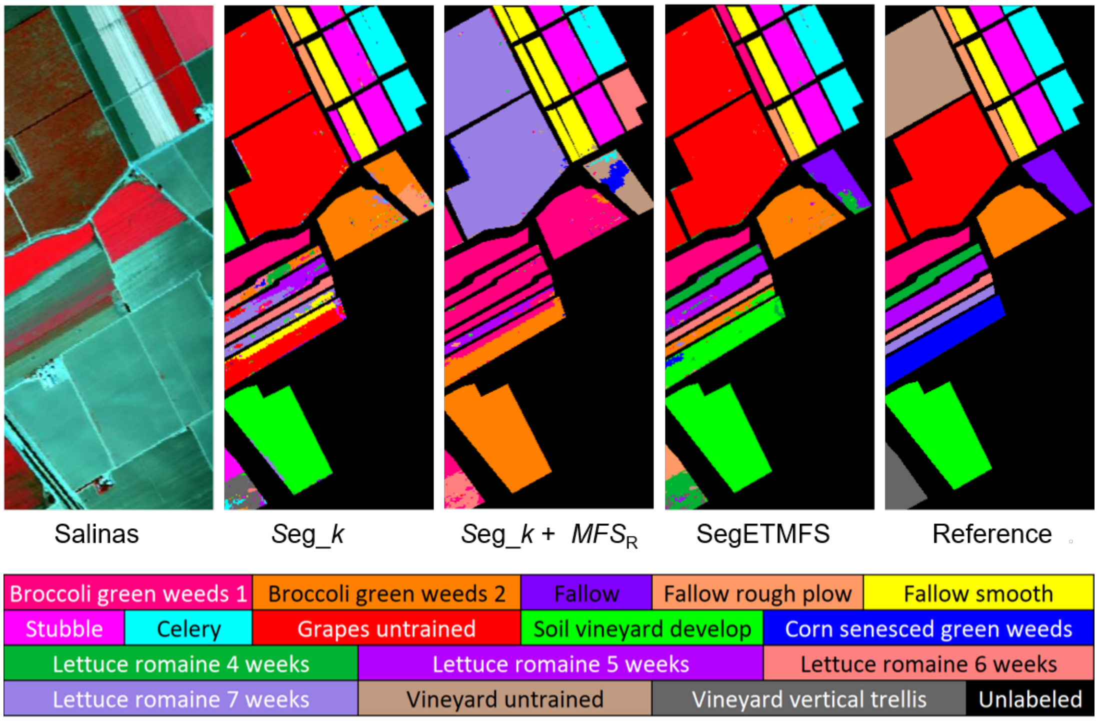

Fig. 3: Salinas Dataset's clustering maps.

respectively, using the multiple feature (band) set selection technique with the redundancy removal score on bands.

B. Pavia Center and Pavia University Dataset

TABLE III: Pavia Center's accuracy values.

	*k*_seg	*k*_cl	*L*	AMI	OA(%)
Seg_*k*	5	8	All bands	0.7306	79.26
Seg_*k* + MFS_R	4	16	16	0.6785	73.89
SegETMFS	5	20	10	0.7603	90.25

k-means segmentation-based clustering for the Pavia Center dataset makes use of a lot of initial random *k*_seg values in the range of 2 to 15 along with discovering number of suitable AMI values and indicating maximum AMI value. Five (as specified in Table III) was found to be the optimum number for Pavia Center for proposed strategy.

The same way, a set of *k*_cl values is utilised for clustering, and *L* values are modified for band selection. The relevant AMI values are observed from *k*_cl = 8 to 20 and *L* altered from 11 to 20, and the highest AMI value is taken into account. The best *k*_cl and *L* values are used to calculate the AMI, which is then investigated to find that it has significantly increased from 67.85% to 76.03%. Using the multiple feature (band) set selection technique and the redundancy reduction criteria increases accuracy by 10.75%, respectively.

TABLE IV: Pavia University's Accuracy values.

	*k*_seg	*k*_cl	*L*	AMI	OA(%)
Seg_*k*	3	9	All bands	0.4604	67.85
Seg_*k* + MFS_R	3	14	11	0.4614	59.87
SegETMFS	2	13	12	0.5408	68.74

k-means segmentation-based clustering for the Pavia University dataset makes use of a lot of initial random *k*_seg values in the range of 2 to 15 along with discovering number of suitable AMI values and indicating maximum AMI value. Two (as specified in Table IV) was found to be the optimum number for Pavia University for proposed strategy.

The same way, a set of k_cl values is utilised for clustering, and L values are modified for band selection. The relevant AMI values are observed from k_cl = 8 to 20 and L altered from 11 to 20, and the highest AMI value is taken into account. The best k_cl and L values are used to calculate the AMI, which is then investigated to find that it has significantly increased from 46.04% to 54.08%. Using the multiple feature (band) set selection technique and the redundancy reduction criteria increases accuracy by 0.22% and 14.68%, respectively.

Table II shows accuracy values corresponding to various parameters and Fig. 3 displays a distinction between the original image and the classified images from the Salinas dataset. The proposed methodology outperformed the compared segmentation-based clustering methodologies. Also, it is noted that the higher accuracy scores are generated from proposed strategy.

V. Conclusion

A novel clustering method and local feature selection is proposed in this paper. After performing image segmentation, the segments are subsequently clustered according to Entropy. Also, a multiple feature (band) set selection based on entropy is suggested and used in the clustering strategy's final stage. Multiple feature (band) set selection takes into consideration both relevance and redundancy criteria. Following that, the proposed technique (SegETMFS) was compared to several clustering strategies. The experiments revealed that using the multiple feature (band) set selection method produced higher accuracy. Therefore, it can be said that the proposed method has achieved greater performance than other clustering algorithms that were compared.

VI. Acknowledgement

The Hyperspectral Images used in this study were provided by Prof. P. Gamba and Prof. David Landgrebe, for which the authors are grateful.

References

[1] M. Mateen, J. Wen, D. Nasrullah, and M. Azeem Akbar, "The role of hyperspectral imaging: A literature review," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 8, 09 2018.

[2] A. Mehta and O. Dikshit, "Comparative study on projected clustering methods for hyperspectral imagery classification," *Geocarto International*, vol. 31, pp. 296–307, 05 2015.

[3] A. F. Alkarkhi and W. A. Alqaraghuli, "Cluster analysis, chapter 11," in *Easy Statistics for Food Science with R*, A. F. Alkarkhi and W. A. Alqaraghuli, Eds. Academic Press, 2019, pp. 177–186.

[4] A. Mehta and O. Dikshit, "Projected clustering of hyperspectral imagery using region merging," *Remote Sensing Letters*, vol. 7, no. 8, pp. 721–730, 2016.

[5] J. A. Richards, *Remote Sensing Digital Image Analysis: An Introduction*. Springer-Verlag, Berlin, Heidelberg, 5th Edition, 2013, pp. 403–446.

[6] A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks," *Science*, vol. 344, no. 6191, pp. 1492–1496, 2014. [Online]. Available: https://www.science.org/doi/abs/10.1126/science.1242072

[7] A. Mehta and O. Dikshit, "Segmentation-based projected clustering of hyperspectral images using mutual nearest neighbour," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 12, 2017, pp. 5237–5244.

[8] H. Motiyani, P. K. Mali, and A. Mehta, "Hyperspectral image segmentation, feature reduction and clustering using k-means," in *2022 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*, 2022, pp. 389–393.

[9] S. L. Polk and J. M. Murphy, "Multiscale clustering of hyperspectral images through spectral-spatial diffusion geometry," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 2021, pp. 4688–4691.

[10] C. Hinojosa, E. Vera, and H. Arguello, "A fast and accurate similarity-constrained subspace clustering algorithm for hyperspectral image," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 10 773–10 783, 2021.

[11] H. Chokshi and A. Agarwal, "Image segmentation," in *Advanced Sensing in Image Processing and IoT 1st Edition*, 02 2022, pp. 43–62.

[12] S. Lloyd, "Least squares quantization in pcm," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 129–137, 1982.

[13] A. K. Jain, M. N. Murty, and P. J. Flynn., *Data Clustering: A Review*, 1999, pp. 264–323.

[14] A. Femenias and S. Marín, "Hyperspectral imaging, chapter 15," in *Electromagnetic Technologies in Food Science*. John Wiley Sons, Ltd, 2021, pp. 363–390.

[15] A. Mehta and O. Dikshit, "Segmentation-based clustering of hyperspectral images using local band selection," *Journal of Applied Remote Sensing*, vol. 11, no. 1, pp. 015 028–015 028, March 2017.

[16] S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, T. Yu, and the scikit-image contributors, "scikit-image: image processing in Python," *PeerJ*, vol. 2, p. e453, 6 2014. [Online]. Available: https://doi.org/10.7717/peerj.453

[17] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, no. 11, pp. 2274–2282, 2012.

[18] N. Gillis, D. Kuang, and H. Park, "Hierarchical clustering of hyperspectral images using rank-two nonnegative matrix factorization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 4, pp. 2066–2078, 2014.

[19] A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks," *Science*, vol. 344, no. 6191, pp. 1492–1496, 2014.

[20] M Graña, MA Veganzons, B Ayerdi, "Hyperspectral remote sensing scenes," accessed: 2021-12-26. [Online]. Available: https://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes

[21] N. X. Vinh, J. Epps, and J. Bailey, "Information theoretic measures for clusterings comparison: Is a correction for chance necessary?" in *Proceedings of the 26th Annual International Conference on Machine Learning*, ser. ICML '09. New York, NY, USA: Association for Computing Machinery, 2009, p. 1073–1080. [Online]. Available: https://doi.org/10.1145/1553374.1553511

[22] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge University Press, 2008, cambridge Books.