

BAND SELECTION USING COEFFICIENT OF VARIATION AND BAND DENSITY RANKING IN HYPERSPECTRAL IMAGE

O. SUBHASH CHANDER GOUD^{1,3}, T. HITENDRA SARMA², C. SHOBA BINDU¹

¹Research Scholar, Department of CSE, JNTUA College of Engineering Anantapur, Anantapuramu, A.P., India.

²Department of IT, Vasavi College of Engineering, Hyderabad. Telangana, India.

³Department of Data Science, Anurag University, Hyderabad, Telangana, India

E-mail: ^{1,3}organtsubhash@gmail.com, ¹shobabindhu@gmail.com, ²t.hitendrasarma@gmail.com

ABSTRACT

In Hyperspectral Image (HSI) processing, one of the critical challenges is addressing Dimensionality Reduction through Feature Selection, especially given the high volume of spectral bands and often limited labeled data. This study introduces an innovative Band Subset Selection (BSS) technique that employs a Ranking-Based Approach to tackle this problem efficiently. The proposed approach is distinguished by its unsupervised nature, leveraging the fusion of two essential statistical measures: Coefficient of Variation (CV) and Band Density (BD). By synergistically combining these metrics, each band in the HSI dataset is analyzed, ranked, and subsequently filtered, allowing the model to identify an optimal subset of bands with the most relevant spectral information. This curated Band Subset (BS) method, termed CV-BDS-BS, is meticulously compared against an existing ranking procedure called SAM-SC (Spectral Angle Mapper with Spatial Coherence). Both methods undergo rigorous evaluation using state-of-the-art machine learning algorithms to ensure the efficiency, robustness, and reliability of the dimensionality reduction process. This integrated CV-BDS-BS methodology streamlines HSI data by reducing dimensionality and preserving essential spectral and spatial information.

Keywords: *Dimensionality Reduction, Feature Selection, Band Subset Selection, Coefficient of Variation, Band Ranking.*

1. INTRODUCTION

Hyperspectral Imaging (HSI) technology has revolutionized numerous fields including environmental monitoring, agriculture, mineral exploration, healthcare, cultural heritage preservation, and forensic studies [1]. Its applications covers a wide spectrum, addressing diverse challenges and offering invaluable insights. However, working with HSI data presents several inherent challenges which includes managing the high volume of data generated, handling the complexity of high dimensionality, addressing cost constraints associated with acquisition and processing, and effectively interpreting the vast amount of information captured. Despite these problems, the richness in spatial and spectral information of HSI data is potentially useful to identify or classify the objects with utmost precision by leveraging its unique spectral signatures [2],

which is constantly driving its applications in defense [2], agriculture [3], medical[4], environmental monitoring[5], industrial automation[5] and land use land cover classification [6].

A hyperspectral dataset comprises D spectral bands and N spatial pixels, representing each pixel as a D -dimensional vector. The *peaking phenomenon* [7] highlights that as D increases to a large number, the performance efficiency of classification models tends to decline. This emphasizes the importance of *dimensionality reduction* (DR) which can be achieved through two primary approaches: *feature (or band) selection* and *feature extraction*. *Feature extraction* involves transforming the original data into a new reduced-dimensional space (e.g., using PCA, LLE, or autoencoders) [8][9], often improving model performance but at the cost of losing interpretability. On the other hand, *feature selection* focuses on

identifying the most relevant spectral bands while retaining the original feature space, offering better interpretability. When coupled with machine learning models, this approach has been shown to be highly effective in managing high-dimensional data, leveraging the spatial correlation among pixels and spectral features, and improving classification accuracy significantly, especially for hyperspectral imagery.

Feature selection can be undertaken through either supervised or unsupervised techniques. Unsupervised feature selection, in particular, emphasizes the identification of relevant features without relying on labeled data. Among these methods, clustering emerges as a powerful approach, frequently applied to hyperspectral imaging (HSI) data. By grouping spectral bands into clusters with high internal cohesion and clear separation from others, clustering not only reduces the dataset's dimensionality but also reveals significant patterns within the data. Beyond its role in feature selection, clustering proves invaluable in the analysis of satellite and aerial imagery. By leveraging attributes such as spectral reflectance, texture, and other pixel-level characteristics, clustering algorithms effectively segment and analyze image data, enabling applications like land cover classification, object detection, and anomaly detection. [10].

Unsupervised learning-based approaches are quite suitable for hyperspectral data analysis due to the limited availability of ground truth label information. Clustering-based band selection is a prominent technique in hyperspectral image analysis aimed at reducing data dimensionality while preserving essential spectral information. This method involves grouping similar spectral bands into clusters and selecting representative bands from each cluster, thereby minimizing redundancy. One of the notable contributions in this direction uses clustering of spectral bands such that intracluster variance is minimized where as the intercluster variance is maximized, in terms of some dissimilarity measures and based on band information content, such as mutual information and Kullback–Leibler divergence. [11]

Another approach uses the DBSCAN algorithm to cluster bands based on spectral similarity. A representative band is then chosen from each cluster, and the bands are ranked accordingly. This method effectively reduces dimensionality and enhances classification performance [12]. Apart from density based clustering methods, hierarchical methods also been explored for band selection. The divisive hierarchical clustering (DHC) method

effectively captures the intrinsic relationships among spectral bands, enabling the formation of any desired number of band groups. This DHC approach effectively identifies and excludes noise bands. This method is robust as it clusters and selects low-correlation bands using information entropy and mutual information. [13]

The Global Optimal Clustering (GOC) algorithm has been proposed, which assumes that all bands within a cluster are highly similar. By selecting representative bands from each cluster, GOC effectively reduces redundancy and maintains the integrity of the original data [14]. Ranking-based methods are highly advantageous due to their independence from prior knowledge about specific bands, making them adaptable to a wide range of applications. Although these methods often involve significant computational effort owing to their exhaustive evaluation processes, the trade-off is worthwhile as they consistently produce more accurate and refined band selections. [15].

Focusing on the similarity measure between two bands, a new approach for similarity between bands, called SSIM, has been introduced and it is integrated with similarity-based ranking (SR) procedures SSIM is observed to be more suitable for HSI than other conventional distance measures by exploring the spatial information contained in the spectral bands. [16].

Recently, another new clustering approach for band ranking method based on information similarity measures from the concepts of information theory viz., Normalized Mutual Information for similarity and Variation of Information for dissimilarity to create a ranking scheme based on the affinity of a band to become a cluster center. [17]. A detailed review of diverse band selection approaches is presented in [18],[19] Ranking-based methods, in particular, offer significant advantages, primarily due to their independence from prior knowledge of specific bands, making them remarkably versatile. While these methods can be computationally demanding owing to their thorough evaluation processes, their ability to achieve more accurate and refined band selections often outweighs the associated time complexity [20]. k-means is one of the simple unsupervised algorithms that is widely used in many applications. k-means with different similarity measures like Spectral Angle Mapper, Spectral Information Divergence, and Jeffereys Matusita are also studied for band selection in HSI [21],[22],[23].

This study introduces a ranking-based approach for band subset selection (BSS), while there are existing methods, such as covariance-based band selection by Kim et al. [54], where the covariance-based method proved to be worthy in detecting the target successfully. Due to its unsupervised nature, the CV-BDS performs BSS with lower computational complexity compared with clustering algorithm characteristics like k-Means, Mutual Information for Ranking of bands, or any other ranking methods for BSS. The proposed method can be implemented without global optima or any tedious transformations like PCA, ICA, etc. The proposed method evaluates bands based on relevance and significance, ranking them according to predetermined criteria. By prioritizing bands that contribute most effectively to the desired analysis outcomes, the ranking-based approach assures that the subset maintains the integrity of spectral information while effectively reducing dimensionality. Consequently, this method enables precise and efficient HSI analysis tailored to the specific requirements of the application at hand. The proposed method is applied to real-time HSI datasets, and the validation indices, like **Overall Accuracy (OA)**, **Kappa**, are utilized to check the possibility of correct predictions.

2. LITERATURE SURVEY

This section briefly reviews some unsupervised ranking-based methods for band selection. An unsupervised band selection algorithm using DBSCAN clustering extracts and clusters attributes from bands, followed by ranking based on non-Gaussianity to prioritize bands by their discriminatory power, as proposed by Datta et al. [12]. A clustering approach that identifies cluster centers as points of highest density and large distance from higher-density points was utilized by Rodriguez et al. [24] to define a method called Fast Density Peak Clustering (FDPC). Jia et al. [25] extended this approach in the Enhanced-FDPC method, which identifies cluster centers using local density and within-cluster distance, and employs exponential-based learning for Band Subset, a ranking-based clustering method for band selection.

Feng et al. [26] proposed a semi-supervised band discrimination method using Maximum Discrimination and Information (MDI) to select bands without redundancy by ensuring high discrimination, high information, and low redundancy (DIR). Later, Feng et al. [27] refined this semi-supervised method by incorporating a non-

negative low-rank representation to reduce redundancy while preserving the original bands.

The Squaring Weighted Low-Rank Subspace Clustering (SWLRSC) method, by Zhai et al. [28], captures global structural data on a hyperspectral image band set through a strongly connected adjacency matrix and dynamically adjusts the size of the band subset from the original bands in HSI. The band selection via rank minimization, employing spatial discrete gradient filtering and block statistics to characterize bands, followed by a low-rank model to determine an affinity matrix for clustering and choosing the most representative bands according to representation residuals. Zhu et al. [29] introduced method for hyperspectral. Sawant et al. [30] proposed a method that leverages spectral, texture, shape, and statistical attributes for pixel categorization in hyperspectral imagery. It combines clustering and ranking to select the top feature groups, with potential improvements through advanced feature selection and alternative clustering algorithms.

The Optimal Clustering Framework (OCF), utilizes Dynamic Programming and Continuous Band Indexes Constraint (CBIC) to form and rank band subsets, facilitating the selection of an optimal subset by clustering similar bands and ensuring the chosen bands are representative as proposed by Wang [31].

Various hyperspectral band selection methods, focusing on ranking-based techniques, have been analyzed by Sun et al. [18] and categorized into supervised and unsupervised. Unsupervised methods prioritize bands using criteria such as variance, spectral derivatives, and information entropy, while techniques like Fast Density Peak Clustering (FDPC) rank bands based on dissimilarity and exemplar component analysis. Another contribution by Patro et al. [32] describing the various unsupervised band selection approaches validated with 2 real HSI datasets, the study concluded that clustering approaches such as Dual clustering, and Superpixel ML-AP approaches perform better than the ranking-based band selection method. The study by Vaddi et al. [33] states that ranking methods consume less computation than clustering-based procedures while ranking procedures decrease accuracy in classifying the HSI data sets.

Another approach by Zhang et al. [34] called “data gravitation and weak correlation-based ranking (DGWCR)” clusters the bands based on connection center evolution (CCE) implying noisy band elimination, later the bands of each cluster are ranked using data gravitation, and entropy-

containing similarity. The Shared Nearest Neighbour Correlation Analysis (SNNCA) method, leveraging shared nearest neighbors to assess local band density and rank bands, enhances the identification of crucial bands as introduced by Yang et al. [35]. A Low ranking-based Band Selection (LRBS), that utilizes a low-rank coefficient matrix for spectral clustering and K-Means to select optimal bands, effectively combining low-rank structures with clustering algorithms proposed by Yu et al. [36].

Band Attention Module (BAM) is a deep learning framework for optimal band selection using a neural network, capturing nonlinear interdependencies among spectral bands and employing a reconstruction framework for unsupervised band reduction proposed by Cai et al. [37]. An image-denoising approach to rank bands employs a reference image from the first principal component, and parameters like mutual information, correlation coefficient, and structural similarity are utilized for ranking the bands Varade et al. [38]. The Spatial Entropy an extension of Shannon Entropy is defined for calculating Mutual Information called (SEMI) that is utilized for identifying optimal Bands, by Wang [39] for band selection. A new clustering method called "Fast and latent low-rank subspace clustering (FLLRSC)", incorporates Hadamard random projections to manage high-dimensional data, transforming HSI into a low-rank structure, and using correntropy measure similarity and spectral clustering to identify a band subset that is both computationally efficient and informative, as introduced by Sun et al. [40].

A band selection method proposed by Su et al. [41] called "Band Ranking via Extended Coefficient of Variation (BRECV)", which ranks hyperspectral bands using the Coefficient of Variation (CV) across all bands, focusing on bands with smaller averages and greater standard deviations relative to their neighbors. A modified CV matrix for neighboring band trios is utilized to rank and select an optimal subset. The "Spatial Residual Clustering and Entropy-based Ranking," method combines spatial residual filtering accompanying K-Means clustering of spatial features, and entropy-based ranking strategy to enhance band selection as defined by Kishore et al. [42]. Sun et al. [43] applied a new clustering mechanism called "hyperbolic clustering-based band hierarchy (HCBH)", that applies an adaptive hyperbolic distance which can get the similarity between bands both geometrically and informative. The new ranking procedure proposed by Li et al. [44] "Band selection for heterogeneity classification of

hyperspectral transmission images based on multi-criteria ranking." Based on a multi-criteria-based ranking (MCR) of bands which divides the data into subintervals the correlation coefficient is based on variance with information entropy. The top 'Q' bands are selected from each subinterval, and the subintervals are combined with the SVM classifier for classification.

A novel method for band selection proposed by Zhang et al. [45] that makes use of multi-objective functionality constructed with information entropy and Structural Similarity Measure (SSIM) to optimize this functionality the model utilizes particle swarm optimization called Adaptive Particle Swarm Optimization with self-repair mechanism, which selects the band subset. The Structural Similarity Index (SSIM) that measures the image quality as proposed by Wang et al. [46] is enhanced for band selection by Xu [47] to develop a "Structural Similarity based Ranking for Band Selection (SR-SSIM)," adapting the Structural Similarity Index (SSIM) to construct a similarity matrix between bands, later ranking them based on the average similarity and dissimilarity then choosing the top 'k' bands by normalizing these metric values.

The 'Dual-Constrained Low-Rank Representation BS (DCLRRBS)' method, incorporating super-pixel and imbalanced class-wise constraints to enhance hyperspectral image (HSI) classification, is proposed by Yu et al. [48]. Band selection methods also leveraged the Improved Affinity Propagation (IAP), starting with the computation of information entropy for each band, followed by constructing a self-similarity matrix. This matrix is constructed based on 'k' that is limited between 1 and the half of the square root of the total number of bands ($\sqrt{L} \div 2$, with L denoting the bands in hyperspectral imagery), dividing it into 'k' blocks, and using the Affinity Propagation (AP) algorithm to generate exemplars that represent Band Subset as defined by Zhu et al. [49].

The Sequential Band Selection Ranking (SBSR) method, computes the entropy of each band, arranging them in descending order, and iteratively selecting bands based on their entropy values and correlation with already selected bands to optimize the number of bands for the band subset proposed by Laveria et al. [50]. Li et al. [51] introduced the Difference between Intergroups (DIG) method for band selection, which merges a Grouping Strategy driven by Intragroup Similarity (GSIS) with a Ranking Strategy emphasizing Intergroup Differences (RSDI). GSIS minimizes intra-group similarity, while RSDI identifies bands with maximum local density representation. The method

addresses redundant bands and evaluates multiple band subsets to determine the optimal one.

3. METHODOLOGY

The proposed methodology for band selection is applied in two phases i) Band Subset Selection (BSS) with Coefficient of Variation (CV) and Band Density (BD) ii) Classification of HSI. In the first phase, the band is calculated with two parameters CV and BD without supervised knowledge of each band. Consider σ_i is the variance of band i and μ_i is the mean of the band i . m_i denotes the coefficient of variation of band i . Matrix M denotes the coefficient of variation of the band i . CV is calculated for every band as per equation (2). CV is a dimensionless statistic which is also known as relative standard deviation as defined by Su et al.[42]. The CV is utilized to calculate has relatively smaller mean and relatively larger standard deviation, which tells that the band represents more information than the adjacent bands. A CV with a more realtive mean and lower standard deviation has less information than the adjacent bands, and the window for calculating the CV is reduced to 3 adjacent bands to lower the computational capacity and also to know the band's representativeness with the local neighbourhood.

$$m_i = \sigma_i / \mu_i \quad (1)$$

$$M = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1/\mu_1 & \sigma_2/\mu_1 & \sigma_3/\mu_1 \\ \sigma_1/\mu_2 & \sigma_2/\mu_2 & \sigma_3/\mu_2 \\ \sigma_1/\mu_3 & \sigma_2/\mu_3 & \sigma_3/\mu_3 \end{bmatrix}$$

$$CV(i) = m_{22} - (m_{21} + m_{11} + m_{12}) \\ + m_{22} - (m_{23} + m_{32} + m_{33}) \quad [41] \quad (2)$$

The CV inspired by Su et al. [41] of each band calculates the standard deviation concerning the adjacent bands, in the matrix M , the m is the coefficient of variance of the bands with their adjacent bands. Each row in matrix M calculates the Root Mean Square Deviation of a band concerning the mean of the neighbouring bands like m_{12} which refers to the band's standard deviation of band 1 to the mean of band 2, and so on

$$S(i,j) = \sum_{i,j}^L (R(i) - R(j))^2 \quad (3)$$

The $S(i,j)$ represents the similarity matrix constructed as $S \in R^{L \times L}$ where R is the HSI dataset with L number of bands in the dataset. The S is calculated for the HSI data set with the $R(i)$ representing the band 'i' pixel intensities and $R(j)$ representing the band 'j' pixel intensities, whereas the similarity matrix is computed for the complete set of bands in the HSI band dataset $i < j < L$ fro the complete HSI dataset.

$$D(i,j) = \frac{\sqrt{S(i,j)}}{L} \quad (4)$$

$D(i,j)$ is the matrix derived from similarity matrix S representing the band 'i' tendency with the complete set of bands of the HSI dataset calculated by equation (4). $D(i,j)$ is calculated in FDPC to determine the local density of the points in a cluster; here, it's calculated as a parameter for deciding the band density.

$$\rho = \sum_j X(D(i,j) - d_c) \quad (5)$$

$$d_c = \mu_i \quad (6)$$

The ' ρ ' calculated by equation (5) represents the sum of the band 'i' tendency and its deviation with the mean of the pixel reflectance intensities of band 'i', which makes the band density represent the band intensity varying with the complete set of bands in the data set. The maximum the ' ρ ' value the more the band represents the ground elements.

$$X(x) = \begin{cases} 1 & \text{if } x < 0 \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

$$BDS(i) = \rho_i \quad (8)$$

$$BS(i) = CV(i) * BDS(i) \quad (9)$$

Equation (9) which is a combination of CV and Band Density combines the two parameters of a band, among which one is the local representativeness of a band within the neighbourhood, and the latter with the representation of the band with more information compared to total bands in HSI. The Band Density or local density as referred to by Jia et al. [25] where the local density of the band is represented by equation (8) where the equation (3) calculates the similarity of the bands with one another based on the reflectance of each

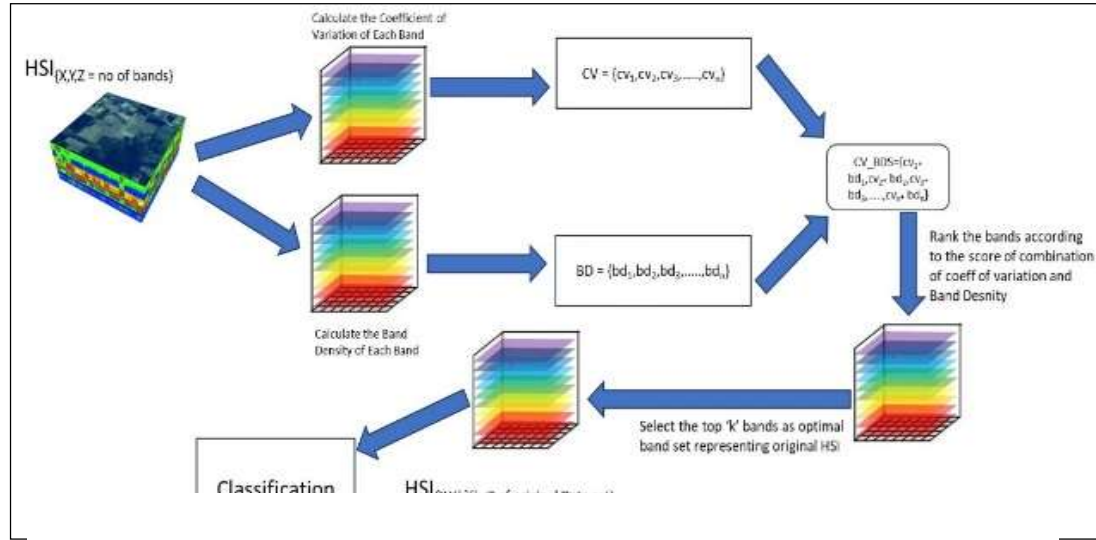


Figure 1: Block Diagram of the Proposed Approach

pixel in the band. Later equation (5) calculates the band's original local density, which points to the similarity of the bands from one another based on the reflectance of the ground elements in the bands. The dc in equation (6) is the threshold cutoff for calculating a band's local density, which here is considered as the average value of band i similarity values depending on the matrix generated with equation (3).

is calculated with equation (9). where each band is ordered in descending order. The top-ranked subset of k bands is chosen as the optimal band selection in the CV and BDS-based (CV-BDS-BS) method.

$$SAM(S_i, S_j) = \cos^{-1} \theta$$

$$\theta = \frac{\sum_{i,j=1}^L S_i S_j}{\sqrt{\sum_{i=1}^L S_i^2} \sqrt{\sum_{j=1}^L S_j^2}} \quad (10)$$

$$SC(i, j) = \sum_{i,j=1}^L S_i * S_j \quad (11)$$

$$SAM - SC(S_i, S_j) = \frac{SC(S_i, S_j)}{SAM(S_i, S_j)} \quad (12)$$

Algorithm 1 Coefficient of Variation with Band Density for Band Selection in HSI ($D_{N \times M}, k$) (CV-BDS-BS)

Initialize: Select all the bands for calculating the coefficient of variation using equation (2)

1. For each Band Calculate the Band Density as equation (8).
2. Compute the Band Score(i) as equation (9).
3. Arrange the Band Score or all the bands in Descending order.
4. Select the top 'k' band as Bands with high priority band subset.

Output: The final optimal spectral bands $\{b_1^{(o)}, b_2^{(o)}, \dots, b_k^{(o)}\}$. =0.

After the calculation of CV and BDS from equations (2) and (8) the product of the two Band Score (BS)

In the comparison, the band's selection is made through another hybrid method called 'Spectral Angle Mapper with Spatial Coherence (SAM-SC)' where the spectral angle mapper gives the similarity between two bands of an HSI by using the equation (10) and Spatial Coherence by equation (11). The SAM-SC proposed by Wang [53] is calculated with equation (12).

The band subset selection is applied using equation (12) between all the bands and the ranking of the bands is ordered according to the descending order of the SAM-SC score of all the bands. The first 'k' bands are chosen to form the optimal band subset.

4. RESULTS AND DISCUSSION

This section presents the details of the datasets used and the analysis of the results to validate the proposed methodology for the band section. Indian

Pines (IP), Salinas (SA), Pavia Center (PC), WHU-Hi-LongKou, and WHU-Hi-HongHu are the data sets used in the current experimental study and the details are provided in Table 1.

4.1 DATA SETS

Table 1: Data sets utilized for the proposed CV-BDS-BS Method.

Dataset	Bands	Pixels	Classes	Sensor
Indian Pines (IP) [37]	224	145 x 145	16	AVIRIS
Salinas (SA) [37]	224	512 x 217	16	AVIRIS
Pavia Centre (PC) [37]	110	610 x 340	9	ROSIS
WHU-Hi-LongKou [54]	270	550 x 400	9	Headwall Nano-Hyperspec
WHU-Hi-HongHu [54]	270	940 x 475	22	Headwall Nano-Hyperspec

4.2 RESULTS AND ANALYSIS

Table 2 presents the results obtained by applying Algorithm 1, where band subsets for all datasets are represented. The band subsets are chosen using both the SAM-SC and CV-BDS-BS methods for all datasets under consideration. For SAM-SC, equation (12) is used to select the chosen k bands as the BSS, while in CV-BDS-BS, Algorithm 1 is applied to identify the selected k bands as the optimal BSS.

The band subsets selected using SAM-SC and CV-BDS-BS are then used to classify the HSI dataset with various classifiers. The train-test split is set at 70% and 30% on the newly reduced datasets, as shown in Table 2. Table 3 demonstrates that the proposed method yields higher overall accuracy, with KNN (for $K=3$) outperforming other band selection methods.

Figure 2 shows that CV-BDS-BS achieves better results than SAM-SC across all datasets. Likewise, Tables 3 and 4 highlight that KNN consistently produces higher Recall and F1 Scores than other techniques for classifying all datasets. Figure 3 displays the Ground Truth and

classification maps generated by each classifier for the Indian Pines dataset.

Figure 4 shows the GT and results generated by different classifiers for the Pavia Centre dataset. Figure 5 presents the GT and classification maps for the Salinas dataset, also produced by DT, KNN, and RF. Figures 6 and 7 illustrate the GT and results generated by these classifiers for the Whi Hi Long Kou and Whi Hi Hong Hu datasets.

The overall accuracy (OA) improvement shown in Figure 2 reveals a significant increase of over 5% to 70% when comparing the SAM-SC method to the CV-BDS-BS method. Specifically, in the PC dataset, there is a modest 5% improvement when utilizing the RF classifier. In contrast, the most substantial gain is observed in the LongKou dataset, where the DT classifier achieves a remarkable 70% increase in accuracy.

Using the RF classifier with the CV-BDS-BS method results in lower accuracy improvements across all datasets, showing just a 5% increase in the PC dataset and a 50% improvement in the LongKou dataset. On the other hand, the DT classifier consistently outperforms RF, achieving at least a 22% improvement in the PC dataset and a maximum of 70% in the LongKou dataset.

In Table 3, the IP dataset shows a Recall score as low as 56% with the RF classifier, while the KNN classifier improves this to 74%. In the PC dataset, SVM achieves the highest Recall score at 96%, with RF recording the lowest at 89%. The LongKou dataset demonstrates a notable improvement over the previous SAM-SC method, with KNN reaching 87% and DT scoring 82%. For the HongHu dataset, SVM reaches 66%, while DT has the lowest score of 59%. Finally, in the Salinas dataset, Recall is comparable to the PC dataset, with KNN achieving the highest score of 92% and DT the lowest at 60%.

Table 4 indicates that the PC dataset achieves the highest F1 Score of 96% with the SVM method, while the lowest score recorded is 86% with the Decision Tree (DT). The Salinas dataset follows with an F1 Score of 91% using KNN, also showing a lower score with DT. The LongKou dataset maintains consistent accuracy across all measures, scoring 86% with KNN and 80% with DT. In comparison, the IP and HongHu datasets have the lowest scores at 47% and 51% with DT, while KNN yields the highest scores of 73% and 59% for these datasets, respectively.

Figure 8 and 9 reflects the proposed method CV-BDS accuracy improvement compared with other methods proposed by Su et al.,[41]. In both the

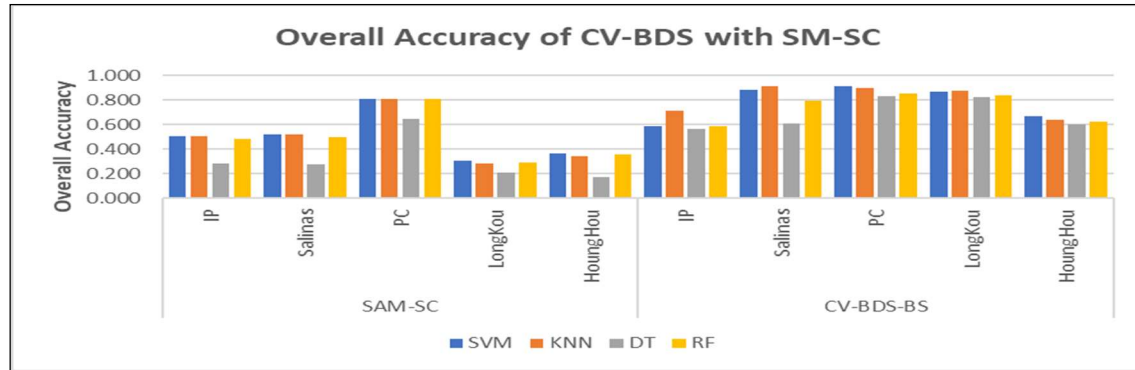


Figure 2: Overall Accuracy of SAM-SC and CV-BDS-BS.

figures where Indian Pines, Salinas data set we can observe KNN based CV-BDS has improved in terms of Overall Accuracy compared with other methods like SVM based CV-BDS and also the methods like Optimal clustering framework (OCF), Optimal neighbourhood reconstruction (ONR), Band Ranking via CV with Dropping Adjacent Bands (BRECVD).

Table 2 Bands Selected using the proposed CV-BDS-BS and SAM-SC Methods.

Method	Dataset	Bands Selected
CV-BDS	IP	15,16,18,19,20,21,23,17,24,22,25,10,26,30,14,72,27,9,54,11
	Salinas	104,38,29,28,27,58,32,30,31,151,26,37,200,25,24,23,20,21,22,77
	PC	14,17,19,15,18,16,21,13,20,12,22,11,10,24,23,25,26,27,9,28
	L-Kou	121, 123, 122, 119, 120, 237, 235, 221, 236, 222, 124, 219, 220, 227, 232, 233, 215, 218, 226, 216
	H-Hou	121, 119, 123, 122, 120, 118, 117, 124, 125, 116, 115, 114, 113, 126, 112, 111, 127, 110, 109, 108
SAM-SC	IP	96,97,76,75,74,73,51,52,53,54,55,56,57,58,59,60,198,61,63,64,65,66,67,68,69,70,71,72,62,199
	Salinas	98,99,89,202,76,74,52,53,54,55,56,57,58,59,60,61,75,62,64,65,66,67,68,69,70,71,72,73,63,203
	PC	40,39,25,38,36,35,34,33,32,31,30,29,28,27,37,101

L-Kou	81,82,100,83,85,86,87,88,89,90,91,92,93,94,95,96,97,98, 84, 269
H-Hou	86,87,88,81,82,100,83,85,89,90,91,92,93,94,95,96,97,98,84, 249

Table 3: Recall values of various classification methods.

Method/Data Set	Indian Pines	Salinas	Pavia Center	Long Kou	Hong Hu
SVM	0.650	0.900	0.967	0.865	0.666
KNN	0.742	0.920	0.952	0.874	0.641
DT	0.565	0.606	0.867	0.826	0.599
RF	0.612	0.820	0.899	0.840	0.628

Table 4: F1 Scores of various classification methods

Method / Data Set	Indian Pines	Salinas	Pavia Center	Long Kou	Hong Hu

SVM	0.598	0.894	0.967	0.839	0.590
KNN	0.732	0.919	0.952	0.866	0.593
DT	0.477	0.563	0.861	0.808	0.517
RF	0.555	0.806	0.891	0.822	0.538

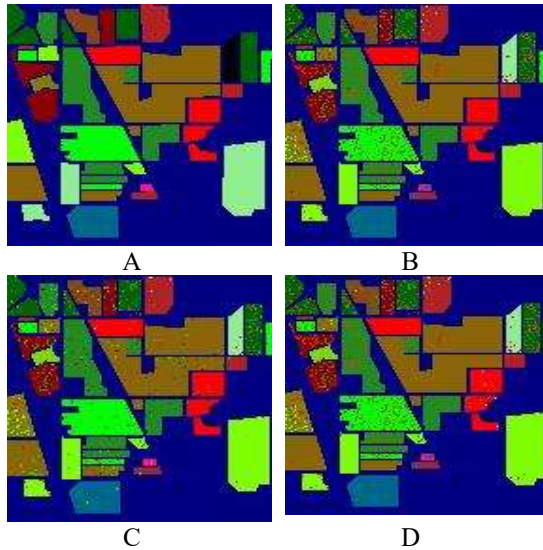


Figure 3: A) Ground Truth of Indian Pines. Classification Maps by B) Decision Tree C) K-Nearest Neighbor D) Random Forest.

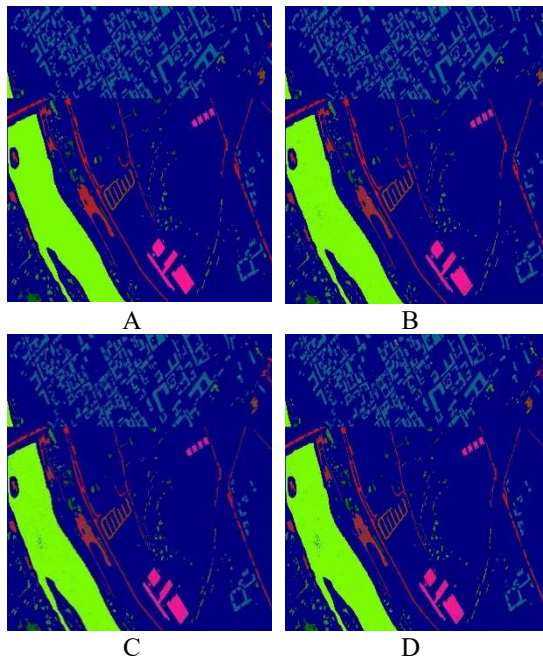


Figure 4: A) Ground Truth of Pavia Center. Classification Maps by B) Decision Tree C) K-Nearest Neighbor D) Random Forest.

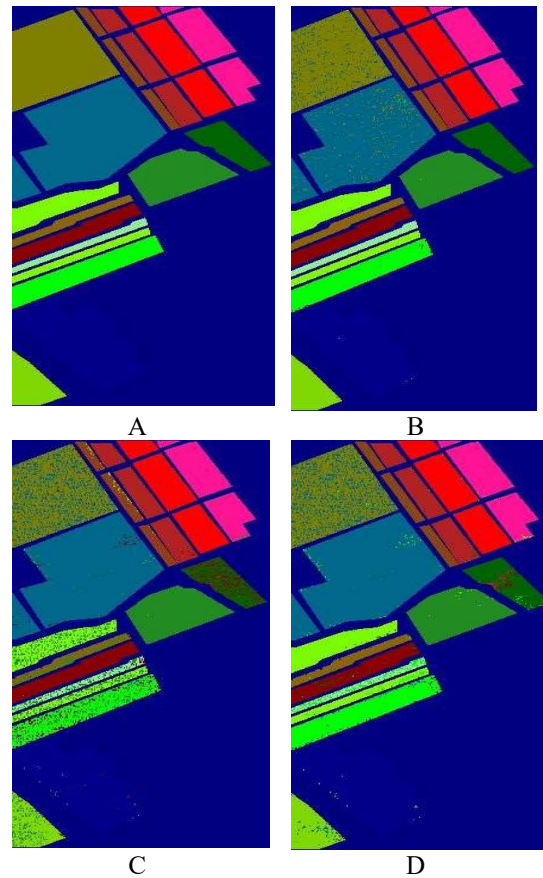


Figure 5: A) Ground Truth of Salinas. Classification Maps by B) Decision Tree C) K-Nearest Neighbor D) Random Forest.

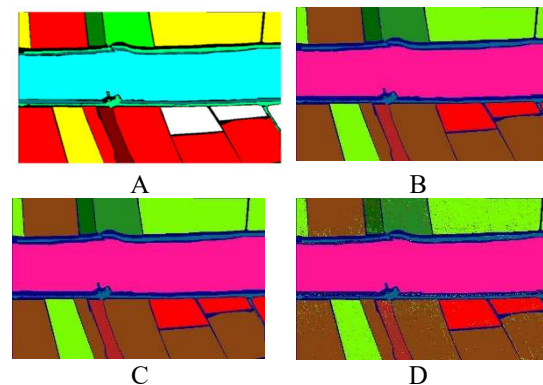


Figure 6: A) Ground Truth of Whu-Hi-LongKou. Classification Maps by B) Decision Tree C) K-Nearest Neighbor D) Random Forest.

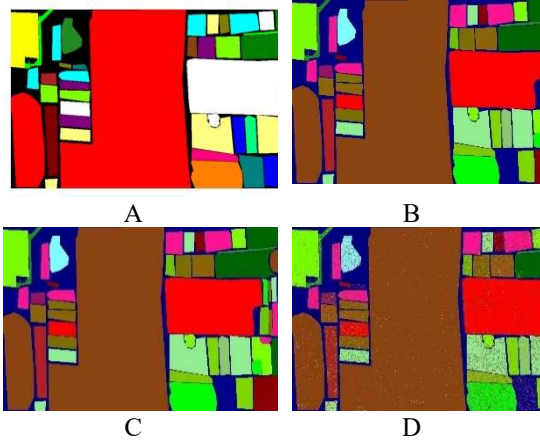


Figure 7: A) Ground Truth of Whu-Hi-Hong Hu. Classification Maps by B) Decision Tree C) K-Nearest Neighbor D) Random Forest.

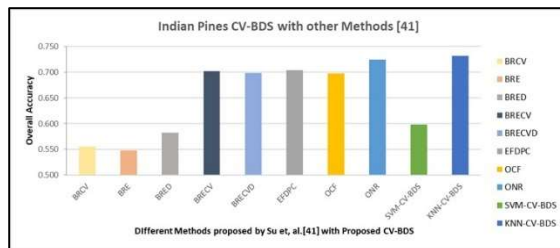


Figure 8 Comparison of Proposed CV-BDS method with Su, et al.,[41].

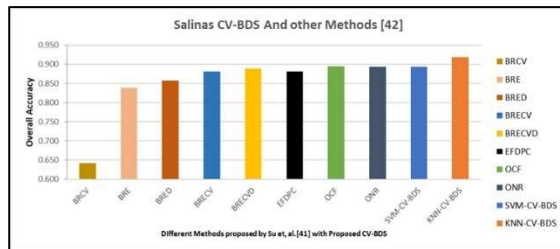


Figure 9 Comparison of Proposed CV-BDS method for Salinas dataset with Su, et al.,[41].

5. CONCLUSION

The present study demonstrates that utilizing the Coefficient of Variation with Band Density hybrid model for band selection yields better results in classifying Hyperspectral Imagery (HSI) datasets. The proposed approach shows better accuracy than other band selection approaches with the KNN algorithm on all the datasets and better band subset selection than the SAM-SC approach. Although SAM-SC also demonstrates improved accuracy, it falls short of the performance achieved by CV-BDS. Classification accuracy and Kappa improvement of more than 50% in some datasets with the proposed CV-BDS band selection, the classification

techniques among which the better one observed is KNN with K as '5'. The computation of CV-BDS-BS is less than other band selection strategies as it is more of statistical parameter extraction from the data set rather than reducing the dimensions with feature extraction, the CV-BDS-BS shows better classification accuracy than SAM-SC. Based on the study the proposed method CV-BDS has the effects of adjacent bands selection which brings redundancy in band selection, which may decrease the accuracy of the classification of HSI. The same method with improvement of removal of adjacent bands will lead to optimal BSS, in turn reflecting the better classification of HSI.

REFERENCES:

- [1] Huang SY, Mukundan A, Tsao YM, Kim Y, Lin FC, Wang HC. Recent advances in counterfeit art, document, photo, hologram, and currency detection using hyperspectral imaging. *Sensors*. 2022 Sep 26;22(19):7308. doi: 10.3390/s22197308.
- [2] Yuen PW, Richardson M. An introduction to hyperspectral imaging and its application for security, surveillance and target acquisition. *The Imaging Science Journal*. 2010 Oct 1;58(5):241-53. doi: 10.1179/174313110X12771950995716.
- [3] Wang B, Sun J, Xia L, Liu J, Wang Z, Li P, Guo Y, Sun X. The applications of hyperspectral imaging technology for agricultural products quality analysis: a review. *Food Reviews International*. 2023 Feb 17;39(2):1043-62. doi: 10.1080/87559129.2021.1929297.
- [4] Fei B. Hyperspectral imaging in medical applications. *InData handling in science and technology 2019 Jan 1 (Vol. 32, pp. 523-565)*. Elsevier. doi: 10.1016/B978-0-444-63977-6.00021-3.
- [5] Khan MJ, Khan HS, Yousaf A, Khurshid K, Abbas A. Modern trends in hyperspectral image analysis: A review. *Ieee Access*. 2018 Mar 12;6:14118-29. doi: 10.1109/ACCESS.2018.2812999.
- [6] Lou C, Al-qaness MA, AL-Alimi D, Dahou A, Abd Elaziz M, Abualigah L, Ewees AA. Land use/land cover (LULC) classification using hyperspectral images: a review. *Geo-spatial information Science*. 2024 Apr 17:1-42. doi: 10.1080/10095020.2024.2332638.
- [7] Hughes G. On the mean accuracy of statistical pattern recognizers. *IEEE transactions on information theory*. 1968 Jan;14(1):55-63. doi: 10.1109/TIT.1968.1054102.
- [8] Firat H, Asker ME, Hanbay D. Classification of hyperspectral remote sensing images using

- different dimension reduction methods with 3D/2D CNN. Remote Sensing Applications: Society and Environment. 2022 Jan 1;25:100694. doi: [10.1016/j.rsase.2022.100694](https://doi.org/10.1016/j.rsase.2022.100694)
- [9] Aloupogianni E, Ishikawa M, Ichimura T, Hamada M, Murakami T, Sasaki A, Nakamura K, Kobayashi N, Obi T. Effects of dimension reduction of hyperspectral images in skin gross pathology. Skin Research and Technology. 2023 Feb;29(2):e13270. doi: [10.1111/srt.13270](https://doi.org/10.1111/srt.13270)
- [10] Huang S, Zhang H, Pižurica A. A structural subspace clustering approach for hyperspectral band selection. IEEE Transactions on Geoscience and Remote Sensing. 2021 Aug 11;60:1-5. doi: [10.1109/TGRS.2021.3102422](https://doi.org/10.1109/TGRS.2021.3102422).
- [11] Martínez-Usó Martínez-Usó A, Pla F, Sotoca JM, García-Sevilla P. Clustering-based hyperspectral band selection using information measures. IEEE transactions on geoscience and remote sensing. 2007 Nov 19;45(12):4158-71. doi: [10.1109/TGRS.2007.904951](https://doi.org/10.1109/TGRS.2007.904951).
- [12] Datta A, Ghosh S, Ghosh A. Clustering based band selection for hyperspectral images. In 2012 international conference on communications, devices and intelligent systems (CODIS) 2012 Dec 28 (pp. 101-104). IEEE. doi: [10.1109/CODIS.2012.6422146](https://doi.org/10.1109/CODIS.2012.6422146).
- [13] Ji H, Zuo Z, Han QL. A divisive hierarchical clustering approach to hyperspectral band selection. IEEE Transactions on Instrumentation and Measurement. 2022 Jun 20;71:1-2. doi: [10.1109/TIM.2022.3184353](https://doi.org/10.1109/TIM.2022.3184353).
- [14] Zhang F, Wang Q, Li X. Hyperspectral image band selection via global optimal clustering. In 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) 2017 Jul 23 (pp. 1-4). IEEE. doi: [10.1109/IGARSS.2017.8126818](https://doi.org/10.1109/IGARSS.2017.8126818).
- [15] Datta A, Ghosh S, Ghosh A. Combination of clustering and ranking techniques for unsupervised band selection of hyperspectral images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2015 May 22;8(6):2814-23. doi: [10.1109/JSTARS.2015.2428276](https://doi.org/10.1109/JSTARS.2015.2428276)
- [16] Xu B, Li X, Hou W, Wang Y, Wei Y. A similarity-based ranking method for hyperspectral band selection. IEEE Transactions on Geoscience and Remote Sensing. 2021 Jan 14;59(11):9585-99. doi: [10.1109/TGRS.2020.3048138](https://doi.org/10.1109/TGRS.2020.3048138).
- [17] Sarma TH, Reddy RD, Mrudula K, Rao KR, Dhondi S, Kanthi M. Band selection in Hyperspectral Images using information similarity ranking. In 2024 International Conference on Machine Intelligence for GeoAnalytics and Remote Sensing (MIGARS) 2024 Apr 8 (pp. 1-4). IEEE. doi: [10.1109/MIGARS61408.2024.10544790](https://doi.org/10.1109/MIGARS61408.2024.10544790).
- [18] Sun W, Du Q. Hyperspectral band selection: A review. IEEE Geoscience and Remote Sensing Magazine. 2019 Jun 17;7(2):118-39. doi: [10.1109/MGRS.2019.2911100](https://doi.org/10.1109/MGRS.2019.2911100).
- [19] O. Subhash Chander Goud. Band Selection Methods for Hyperspectral Imagery Analysis – A Critical Comparison. Int J Intell Syst Appl Eng. 2024 Mar. 24;12(3):3093-109 [Link](#).
- [20] Sawant SS, Prabukumar M, Samiappan S. Ranking and grouping based feature selection for hyperspectral image classification. In 39th Asian Conf. Remote Sens.(ACRS 2018) 2018. Weblink: [Ranking and grouping](#).
- [21] Goud OS, Sarma TH, Bindu CS. Optimal Band Selection in Hyperspectral Images Using Improved K-Means Clustering with Spectral Similarity Measures. In 2023 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAET) 2023 Sep 12 (pp. 268-271). IEEE. doi: [10.1109/IICAET59451.2023.10291338](https://doi.org/10.1109/IICAET59451.2023.10291338).
- [22] Goud OS, Sarma TH, Bindu CS. Improved K-Means Clustering Algorithm for Band Selection in Hyperspectral Images. In 2023 International Conference on Electrical, Electronics, Communication and Computers (ELEXCOM) 2023 Aug 26 (pp. 1-6). IEEE. doi: [10.1109/ELEXCOM58812.2023.10370382](https://doi.org/10.1109/ELEXCOM58812.2023.10370382).
- [23] Varade D, Maurya AK, Dikshit O. Unsupervised band selection of hyperspectral data based on mutual information derived from weighted cluster entropy for snow classification. Geocarto International. 2021 Sep 14;36(15):1709-31. doi: [10.1080/10106049.2019.1665717](https://doi.org/10.1080/10106049.2019.1665717).
- [24] Rodriguez A, Laio A. Clustering by fast search and find of density peaks. science. 2014 Jun 27;344(6191):1492-6. doi: [10.1126/science.1242072](https://doi.org/10.1126/science.1242072).
- [25] Jia S, Tang G, Zhu J, Li Q. A novel ranking-based clustering approach for hyperspectral band selection. IEEE Transactions on Geoscience and Remote Sensing. 2015 Jul 17;54(1):88-102. doi: [10.1109/TGRS.2015.2450759](https://doi.org/10.1109/TGRS.2015.2450759).
- [26] Feng J, Jiao L, Liu F, Sun T, Zhang X. Mutual-information-based semi-supervised hyperspectral band selection with high discrimination, high information, and low

- redundancy. IEEE transactions on geoscience and remote sensing. 2014 Dec 5;53(5):2956-69. doi: [10.1109/TGRS.2014.2367022](https://doi.org/10.1109/TGRS.2014.2367022).
- [27] Feng Y, Yuan Y, Lu X. A non-negative low-rank representation for hyperspectral band selection. International Journal of Remote Sensing. 2016 Oct 1;37(19):4590-609. doi: [10.1080/01431161.2016.1214299](https://doi.org/10.1080/01431161.2016.1214299).
- [28] Zhai H, Zhang H, Zhang L, Li P. Squaring weighted low-rank subspace clustering for hyperspectral image band selection. In 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS) 2016 Jul 10 (pp. 2434-2437). IEEE. doi: [10.1109/IGARSS.2016.7729628](https://doi.org/10.1109/IGARSS.2016.7729628).
- [29] Zhu G, Huang Y, Li S, Tang J, Liang D. Hyperspectral band selection via rank minimization. IEEE Geoscience and Remote Sensing Letters. 2017 Nov 14;14(12):2320-4. doi: [10.1109/LGRS.2017.2763183](https://doi.org/10.1109/LGRS.2017.2763183).
- [30] Sawant SS, Prabukumar M, Samiappan S. Ranking and grouping based feature selection for hyperspectral image classification. In 39th Asian Conf. Remote Sens. (ACRS 2018) 2018. [weblink](#).
- [31] Wang Q, Zhang F, Li X. Optimal clustering framework for hyperspectral band selection. IEEE Transactions on Geoscience and Remote Sensing. 2018 May 9;56(10):5910-22. doi: [10.1109/TGRS.2018.2828161](https://doi.org/10.1109/TGRS.2018.2828161).
- [32] Patro RN, Subudhi S, Biswal PK, Dell'acqua F. A review of unsupervised band selection techniques: Land cover classification for hyperspectral earth observation data. IEEE Geoscience and Remote Sensing Magazine. 2021 Feb 24;9(3):72-111. doi: [10.1109/MGRS.2021.3051979](https://doi.org/10.1109/MGRS.2021.3051979).
- [33] Vaddi R, Kumar BP, Manoharan P, Agilandeewari L, Sangeetha V. Strategies for dimensionality reduction in hyperspectral remote sensing: A comprehensive overview. The Egyptian Journal of Remote Sensing and Space Sciences. 2024 Mar 1;27(1):82-92. doi: [10.1016/j.ejrs.2024.01.005](https://doi.org/10.1016/j.ejrs.2024.01.005).
- [34] Zhang C, Zhang Z, Yu D, Cheng Q, Shan S, Li M, Mou L, Yang X, Ma X. Unsupervised band selection of medical hyperspectral images guided by data gravitation and weak correlation. Computer Methods and Programs in Biomedicine. 2023 Oct 1;240:107721. doi: [10.1016/j.cmpb.2023.107721](https://doi.org/10.1016/j.cmpb.2023.107721).
- [35] Yang R, Kan J. An unsupervised hyperspectral band selection method based on shared nearest neighbor and correlation analysis. IEEE Access. 2019 Dec 20;7:185532-42. doi: [10.1109/ACCESS.2019.2961256](https://doi.org/10.1109/ACCESS.2019.2961256).
- [36] Yu C, Cen K, Chang CI, Li F. Unsupervised hyperspectral band selection method based on low-rank representation. In Communications, Signal Processing, and Systems: Proceedings of the 2018 CSPA Volume I: Communications 7th 2019 (pp. 1053-1061). Springer Singapore. doi: [10.1007/978-981-13-6264-4_124](https://doi.org/10.1007/978-981-13-6264-4_124).
- [37] Cai Y, Liu X, Cai Z. BS-Nets: An end-to-end framework for band selection of hyperspectral image. IEEE transactions on geoscience and remote sensing. 2019 Nov 20;58(3):1969-84. doi: [10.1109/TGRS.2019.2951433](https://doi.org/10.1109/TGRS.2019.2951433).
- [38] Varade D, Maurya AK, Dikshit O. Unsupervised hyperspectral band selection using ranking based on a denoising error matching approach. International Journal of Remote Sensing. 2019 Oct 18;40(20):8031-53. doi: [10.1080/01431161.2019.1608388](https://doi.org/10.1080/01431161.2019.1608388).
- [39] Wang B, Wang X, Chen Z. SPATIAL ENTROPY BASED MUTUAL INFORMATION IN HYPERSPECTRAL BAND SELECTION FOR SUPERVISED CLASSIFICATION. International Journal of Numerical Analysis & Modeling. 2012 May 1;9(2). [Web link](#)
- [40] Sun W, Peng J, Yang G, Du Q. Fast and latent low-rank subspace clustering for hyperspectral band selection. IEEE Transactions on Geoscience and Remote Sensing. 2020 Jan 3;58(6):3906-15. doi: [10.1109/TGRS.2019.2959342](https://doi.org/10.1109/TGRS.2019.2959342).
- [41] Su P, Tarkoma S, Pellikka PK. Band ranking via extended coefficient of variation for hyperspectral band selection. Remote Sensing. 2020 Oct 12;12(20):3319. doi: [10.3390/rs12203319](https://doi.org/10.3390/rs12203319).
- [42] Kishore Raju K, Saradhi Varma GP, Rajya Lakshmi D. Spatial residual clustering and entropy based ranking for hyperspectral band selection. European Journal of Remote Sensing. 2020 Jun;53:82-92. doi: [10.1080/22797254.2019.1703559](https://doi.org/10.1080/22797254.2019.1703559).
- [43] Sun H, Zhang L, Ren J, Huang H. Novel hyperbolic clustering-based band hierarchy (HCBH) for effective unsupervised band selection of hyperspectral images. Pattern Recognition. 2022 Oct 1;130:108788. doi: [10.1016/j.patcog.2022.108788](https://doi.org/10.1016/j.patcog.2022.108788).
- [44] Li G, Ma S, Li K, Zhou M, Lin L. Band selection for heterogeneity classification of

- hyperspectral transmission images based on multi-criteria ranking. *Infrared Physics & Technology*. 2022 Sep 1;125:104317. doi: [10.1016/j.infrared.2022.104317](https://doi.org/10.1016/j.infrared.2022.104317).
- [45] Zhang Y, Lin Q, Li L, Xiao Z, Ming Z, Leung VC. Multiobjective band selection approach via an adaptive particle swarm optimizer for remote sensing hyperspectral images. *Swarm and Evolutionary Computation*. 2024 Aug 1;89:101614. doi: [10.1016/j.swevo.2024.101614](https://doi.org/10.1016/j.swevo.2024.101614).
- [46] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*. 2004 Apr 13;13(4):600-12. doi: [10.1109/TIP.2003.819861](https://doi.org/10.1109/TIP.2003.819861).
- [47] Xu B, Li X, Hou W, Wang Y, Wei Y. A similarity-based ranking method for hyperspectral band selection. *IEEE Transactions on Geoscience and Remote Sensing*. 2021 Jan 14;59(11):9585-99. doi: [10.1109/TGRS.2020.3048138](https://doi.org/10.1109/TGRS.2020.3048138).
- [48] Yu C, Zhou S, Song M, Chang CI. Semisupervised hyperspectral band selection based on dual-constrained low-rank representation. *IEEE Geoscience and Remote Sensing Letters*. 2021 Jan 20;19:1-5. doi: [10.1109/LGRS.2021.3049267](https://doi.org/10.1109/LGRS.2021.3049267).
- [49] Zhu Q, Wang Y, Wang F, Song M, Chang CI. Hyperspectral band selection based on improved affinity propagation. In *2021 11th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*. 2021 Mar 24 (pp. 1-4). IEEE. doi: [10.1109/WHISPERS52202.2021.9484004](https://doi.org/10.1109/WHISPERS52202.2021.9484004).
- [50] Llaveria D, Camps A, Park H, Narayan R. Ranking Methodology for Sequential Band Selection Combining Data Dispersion and Spectral Band Correlation. In *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium* 2022 Jul 17 (pp. 775-778). IEEE. doi: [10.1109/IGARSS46834.2022.9884380](https://doi.org/10.1109/IGARSS46834.2022.9884380).
- [51] Li S, Peng B, Fang L, Zhang Q, Cheng L, Li Q. Hyperspectral band selection via difference between intergroups. *IEEE Transactions on Geoscience and Remote Sensing*. 2023 Feb 3;61:1-0. doi: [10.1109/TGRS.2023.3242239](https://doi.org/10.1109/TGRS.2023.3242239).
- [52] Wang P, Wang L, Leung H, Zhang G. Super-resolution mapping based on spatial-spectral correlation for spectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*. 2020 Jul 1;59(3):2256-68. doi: <https://doi.org/10.1109/TGRS.2020.3004353>
- [53] Zhong Y, Hu X, Luo C, Wang X, Zhao J, Zhang L. WHU-Hi: UAV-borne hyperspectral with high spatial resolution (H2) benchmark datasets and classifier for precise crop identification based on deep convolutional neural network with CRF. *Remote Sensing of Environment*. 2020 Dec 1;250:112012. doi: [10.1016/j.rse.2020.112012](https://doi.org/10.1016/j.rse.2020.112012).
- [54] Kim JH, Kim J, Yang Y, Kim S, Kim HS. Covariance-based band selection and its application to near-real-time hyperspectral target detection. *Optical Engineering*. 2017 May 1;56(5):053101-. doi: [10.1117/1.OE.56.5.053101](https://doi.org/10.1117/1.OE.56.5.053101)