International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN: 2147-6799 www.ijisae.org Original Research Paper

A Review of Statistical Approaches for Band Reduction Techniques for Hyperspectral Image Classification

[1]Biju Kumar S. P., [2]Meera Nair

Submitted: 05/06/2024 **Revised:** 17/07/2024 **Accepted:** 25/07/2024

Abstract

Hyperspectral imaging (HSI) captures a wide spectrum of light, divided into countless, closely spaced bands. This is very useful to record more information in a data pixel, enabling precise identification of materials and objects. The vast number of spectral bands in HSI data can introduce challenges like redundant information and higher computational requirements. Numerous methods exist to reduce HSI dimensionality by selecting informative bands. Band selection is a critical step, focusing on identifying meaningful bands and eliminating redundant or irrelevant ones. This study compares major statistical approaches in band selection methods: Information-Theoretic approaches, PCA (Principal Component Analysis), ICA (Independent Component Analysis), and Clustering-Based Method and evaluate the performance and methodology of these approaches in the field of hyperspectral image classification.

Keywords- *Hyperspectral Imaging, Band Selection, Statistical Methods, Feature Selection, Classification, Remote Sensing*

Introduction

A Hyperspectral image includes hundreds of narrow spectral bands [1] with in the visible to infrared regions of electromagnetic spectrum. This high spectral resolution helps to analyse a detailed material composition and properties of objects in the image that can be utilized to perform accurate material identification & classification. Some of applications of HSI are food quality checking [2], urban planning [3], [4], agriculture[5], Biotechnology [6], Environmental Monitoring [7], Forensic Science [8], Oil and Gases [9], and Medical field [10]. Conversely, HSI have a very large volume of data[11]. The common image processing algorithms are not efficient in managing the dimensionality of HSI. In the case of hyper spectral image classification this high dimensionality badly effects the computational performance and classification accuracy. So, Band selection methods play a crucial role in HSI processing Numerous approaches have been developed to reduce the number of spectral bands while preserving essential information. Statistical methods play a major role in these approaches. This literature review focuses on some of the frequently used statistical methods for hyperspectral band selection along with their mathematical formulations and a comparative analysis.

[1]*Assistant Professor, Department of Computer Applications, NSS College, Rajakumari, Idukki, Kerala Email: bijukumarsp@gmail.com*

[2]*Assistant Professor, Department of Computer Applications, NSS College, Rajakumari, Idukki, Kerala. Email: meeranairmail@gmail.com*

Statistical approaches for band selection

Information-Theoretic Methods

Mutual Information (MI) is a popular method for evaluating the relevance of bands in HSI. MI is a measure that assesses the information shared between two variables, such as a spectral band and the target class label. The mathematical formulation of mutual information given by:

$$MI(Q,\theta) = \sum_{q\in Q}\sum_{i\in\theta} P_{(Q,\theta)}(q,i)\log\left(\frac{P_{(Q,\theta)}(q,i)}{P_Q(q)P_\theta(i)}\right)$$

Where $P_{(Q,\ \theta)}$ - joint probability mass function
P_Q - marginal probability mass functions of Q
P_θ -marginal probability mass functions of θ .

In the case of hyperspectral images, random variable(Q) represents the pixels' value in the HSI and random variable (θ) represents the corresponding label in the ground truth, mutual information between Q and θ means that the dependency between HSI and the ground truth value.

Information Entropy is another method closely related to Mutual information [2] , ie Mutual information can be expressed by joint entropy or conditional entropy.

$$MI(Q,\theta) = E(Q) + E(\theta) - E(Q,\theta) \text{ or}$$
$$MI(Q,\theta) = E(Q) - E(Q/\theta)$$

Where E(Q) and E(θ) are the entropy of Q and θ, E(Q, θ) is the joint entropy and E(Q/ θ) is the conditional entropy of Q given θ.

$E(Q) = -\sum_Q p(Q)log\ p(Q)$ where P(Q) is the probability distribution function of Q

$$E(Q,\theta) = -\sum_{q\in Q}\sum_{i\in\theta} p(q,i)logp(q,i)$$

$$E(Q/\theta) = -\sum_{q\in Q}\sum_{i\in\theta} p(q/i\ )logp(q/i)$$

Steps for band selection

1. Calculate Mutual Information: For each spectral band, calculate the mutual information relationship between the bands and the class labels. Bands that share more information with the class labels are more relevant for classification.
2. Rank the Bands: Based on the computed mutual information, rank the spectral bands from most informative to least informative.
3. Select Top Bands: Select all bands with high ranks that reduces the dimensionality of the data while preserving spectral information.
4. Handle Redundancy: Some of the selected band provide similar information (redundant bands), to avoid this use some advanced techniques like conditional mutual information or joint mutual information.

MI is good approach to capture non-linear dependencies between spectral bands and class labels but linear methods like correlation fails to measure nonlinear dependencies.

Table 1 Band selection using Information-Theoretic approach

Reference	Approach	Classifier, Data set, Accuracy
[12]	Utilize a two-phase feature selection strategy. First, apply minimal-redundancy-maximal-relevance criterion to generate a candidate feature pool. Then, use wrappers (e.g., naive Bayes classifier) to select a compact subset from this pool	NB, HDR-MultiFeature, 96.8 SVM, HDR-MultiFeature, 96.5 LDA, HDR-MultiFeature,95.9 NB, Arrhythmia, 82.14 SVM, Arrhythmia, 80.48 LDA, Arrhythmia, 81.67 NB, NCI, 85 SVM, NCI, 86.67 LDA, NCI, 75 NB, Lymphoma, 94.79 SVM, Lymphoma, 96.87 LDA, Lymphoma, 96.87

[13]	Use mutual information (MI) values and correlation with adjacent bands are used to select informative bands.	Classifier: SVM Data Set: AVIRIS 92AV3C Accuracy: 80.67 (for 20 bands) ,90.62 (for 80 bands)
[14]	Apply a differentiable mutual information for selecting informative bands. It helps to generate an automatic solution by gradient searching	Classifier: SVM Data Set: AVIRIS 92AV3C Accuracy: 85-90% (Depending on the spectral window width and position)
[15]	Symmetric Uncertainty and Mutual Information for Dimensionality Reduction	Classifier: SVM Data Set: AVIRIS 92AV3C Accuracy: 84.16 (for 42 bands)
[16]	This method selects some of redundant bands with high mutual information values, which helps to proper mapping to ground truth values. To manage redundancy, a complementary threshold is applied to the final mutual information value. The error probability of classification is calculated using Fano's inequality.	Classifier: SVM Data Set: AVIRIS 92AV3C Accuracy: 88.14 (for 18 selected bands and threshold of 0.03 to control redundancy)
[17]	Combines mutual information with a steepest ascent strategy to optimize feature selection dynamically. This algorithm uses a wrapper approach for feature selection	Classifier: SVM Data Set: AVIRIS 92AV3C Accuracy: 84.16 (for 42 selected bands and threshold of 0.56)
[18]	This method uses different band discrimination methods such as, and SIDAM creating the band channel, it helps to select relevant bands that are not highly intercorrelated. This approach effectively addresses the challenge of inter-band correlation and redundant data	Classifier: SVM Datasets: 1. Indiana Indian Pines 2. Salinas 3. University of Pavia
[19]	It uses Sequential Forward Search (SFS) to select informative, non-redundant bands based on mutual information. Selects bands with high information content and low redundancy by calculating the average mutual information between each band.	KNN, Botswana Dataset, 97.35 SVM, Botswana Dataset, 92.25 KNN, Indian Pine, 89.56 SVM, Indian Pine, 75.22

Principal Component Analysis (PCA)

PCA is a dimensionality reduction a technique simplifies complex data without losing key information. It identifies uncorrelated principal components with maximum variance in the data. The principal components are mutually independent, meaning they are not correlated with each other. They are also arranged in descending order of variance. Covariance formula for a sample is

$$cov_{x,y} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

Band Selection method

1. Reshaping the Hyperspectral Data (3D) to 2D form
2. Calculate the covariance matrix of HSI.
3. Find the eigen values and eigen vectors of the covariance matrix.
4. Arrange the eigenvalues in

descending order and retrieve the associated eigenvectors.

5. Identify the top 'n' eigenvectors having top 'n' eigenvalues to get the 'n' principal components.

The selected principal components can then be used to find the informative bands form the given data set.

PCA is a widely adopted and powerful technique for dimensionality reduction problems that helps to find more informative bands that helps to increase the efficiency of the classifier. Minimum Noise Fraction (MNF), Folded-PCA(FPCA), Spectrally Segmented PCA(SSPCA) and Segmented-PCA(SPCA) are some linear extensions of PCA. Kernel Entropy Component Analysis(KECA) and Kernel-PCA(KPCA) are the nonlinear extensions [20].

Table 2 Band selection using PCA approach

Reference	Approach	Classifier, Data set, Accuracy
[21]	Reduce dimensionality of hyperspectral image by extracting its principal components with high variance.	Classifier: Maximum Likelihood Classifier (MLC) Data set: HYDICE, AVIRIS Accuracy: Up to 90% accuracy using all PCA bands.
[22]	This approach uses Semi-supervised PCA and hypergraph models for dimensionality reduction.	Classifier: Sparse Representation Classifier (SRC) Data set: APHI, Indian Pines, Washington DCMall.
[23]	Different feature extraction techniques (e.g., PCA, ICA) are applied to reduce the dimensionality and improve classification.	Classifier: SVM with RBF kernel Data set: Indian Pines. Accuracy: PCA using 46 bands – 93.74 SPCA using 50 bands -93.74 FPCA using 40 bands -95.12 KPCA using 107 bands – 95.92 KECA using 147 bands -95.63
[24]	Edge-preserving filtering technique used for removing textures and noise then uses PCA for selecting bands.	Classifier: SVM Data set: Indian Pines Accuracy: 95
[25]	PCA used to extract principal components and BEMD (Bi-dimensional empirical mode decomposition) decomposes the image into three BIMFs (bi-dimensional intrinsic mode functions) and residue image. Finally, BIMFs and residue image used for classification.	Classifier: SVM Data set: Indian Pines - 96.7 Pavia University - 97.6
[26]	Randomized PCA (R-PCA) used for dimensionality reduction.	SVM, Pavia University, 93.42 SVM, Indian Pines, 83.54 LightGBM, Pavia University, 95.39 LightGBM, Indian Pines, 85.46

Independent Component Analysis (ICA)

ICA is used to separate statistically independent components from a mixture of signals. In PCA, which maximizes variance, ICA maximizes the statistical independence between components. In the case of hyperspectral images (HSI) spectral bands can be treated as independent sources of information. ICA aims to represent the observed multivariate signal X as a linear combination of statistically independent components.

$$X=AS$$

where:

- **X** is the matrix of observed data (with rows representing different spectral bands and columns representing different observations).
- **A** is an unknown mixing matrix that mixes the independent components.
- **S** is the matrix of independent components (the underlying sources).

The ICA algorithm works to estimate W (unmixing matrix), such that

$$S=WX$$

The elements within each row of S should be statistically independent. The independence is typically measured using criteria like kurtosis (to identify non-Gaussian signals).

Steps in ICA for Band Selection

- ICA Estimation: Algorithms such as FastICA are applied to find the unmixing matrix W. FastICA uses an iterative approach to maximize the non-Gaussianity of the estimated components, which helps in ensuring statistical independence.
- Selection of Bands: After finding the independent components, select bands with the highest absolute values in the mixing matrix A.

Independent Component Analysis (ICA) helps to select hyperspectral bands by focusing on statistical independence. But computational complexity is very high than PCA. Infomax, FastICA, and JADE are popular ICA algorithms used in signal processing [27].

Table 3 Band selection using ICA approach

Reference	Approach	Classifier, Data set, Accuracy
[27]	Compares different ICA methods like Infomax, FastICA and JADE for dimensionality reduction and their performance in hyperspectral image classification.	Classifier: SVM, Random Forest Data set: Indian Pines, Pavia University Accuracy: Some ICA methods outperform PCA for certain datasets
[28]	In this approach, choose a transform matrix using Independent Component Analysis (ICA) then estimate the probability distributions of the independent components using a non-parametric method. Finally, classify the data using Bayes' rule.	Classifier: ICDA Data Set: Hekla, Indian Pine, ROSIS Accuracy: ROSIS -82.14, Indian Pine-6.58, Hekla -93.91

[29]	Use a nonparametric kernel density estimator to approximate the probability distributions of the independent components	Classifier: ICDA Data Set: Hekla, Indian Pine, ROSIS, HYDICE DC Mall Accuracy: ROSIS -82.14, Indian Pine- 67.08, Hekla -94.53 HYDICE DC Mall-98.69
[30]	In this method ICA is used for initial dimensionality reduction and then genetic algorithm is used to optimize the selection. Reduced morphological attributes used for spatial feature selection. This integrated approach enhances both spectral and spatial classification.	Classifier: SVM with Radial Basis Function (RBF) kernel Data set: Pavia University, Pavia Center, Salinas, and Hekla Accuracy: Pavia University -95.6 Pavia Center -99.12 Salinas -99.51 Hekla -98.87
[31]	This paper introduces a Random Fourier Feature (RFF)-based ICA method is used for reducing the dimensionality of HSI. This approach addresses the computational complexity issues associated with Kernel ICA	Classifier: SVM with RBF kernel. Data set: Pavia University, Salinas Scene Accuracy (35 selected bands): Salinas- 90.29 Pavia University -88.18

Clustering Methods

Clustering methods are commonly used to group spectral bands together subject to some properties of the band. Using this method easily identify bands with similar information and then select a band from that group containing all the properties of that group that helps to reduce redundancy.

K-means clustering is a popular method for hyperspectral band selection. It's a Euclidean distance-based algorithm that partitions spectral bands into k clusters based on their similarity. The objective function for k-means is:

$$\min_{b1,\ldots,bk} \sum_{i=1}^{k} \sum_{x \in Bi} \|x - \mu_i\|^2$$

where Bi represent the clusters, μi represent the mean of cluster Bi.

K-means algorithm use different characteristics of data for clustering, i.e., mean, median, mode, interquartile range and geometric mean etc... After clustering, one band from each cluster is selected as a representative.

Affinity propagation

Unlike k-means, affinity propagation [32] is a clustering algorithm that determines the number of clusters automatically. The cluster quality of k-means depends on initial values. In the affinity propagation-based clustering approach, representative bands are identified by searching for a matching set of exemplars. This method takes into account both the inter-band relationships and the individual discriminative power of each band.

Key-steps of clustering by affinity propagation

1. Compute similarity: Euclidean distance, negative squared Euclidean distance etc... are some popular methods used to compute a similarity matrix
2. Responsibility Update: Sets the initial responsibility matrix, where the (i, k)th element represents the responsibility of data point i to be the exemplar for data point k.
3. Update availability matrix: The availability matrix (A) is created, with A(i, k) representing the suitability of data point i as an exemplar for data

point k. Data points with high availability from others are more likely to become exemplars.

4. Update availability and responsibility matrix: Repletely update the availability and responsibility matrices until convergence.
5. Calculate net responsibility: To find net responsibility, calculates the sum of responsibility and availability for each point.
6. Select exemplar: This crucial step identifies data points with high net responsibility as exemplars, which are then used as cluster canters.
7. Cluster formation: Lastly, data points are assigned to their nearest exemplars to form clusters based on similarity

Table 4 Band selection using clustering approach

Reference	Approach	Classifier, Data Set, Accuracy
[33]	Denoise the spatial images using wavelet shrinkage and then select informative bands from the denoised data using affinity propagation	KNN, Indian Pines, 85.04 SVM, Indian Pines, 94.65 KNN, Kennedy Space Center,95.22 SVM, Kennedy Space Center,98.49 KNN, Botswana, 93.33 SVM, Botswana, 97.77
[34]	Use FDPC (Fast density-peak clustering) approach to select bands. Combine normalized local density and intracluster distance to calculate the ranking score for each band. Using an exponential learning rule, define the cutoff threshold for selected bands.	KNN, Indian Pine (10 bands), 51.47 SVM, Indian Pine (10 bands), 57.91 KNN, KSC (15 bands), 79.15 SVM, KSC (15 bands), 83.27 KNN, Pavia Center (14 bands), 91.9 SVM, Pavia Center (14 bands), 94.66
[35]	Clustering with k-means is performed twice because in each time k-means select different bands due to random behaviour of k-means.	SVM, Indian Pines(20 bands),78.75 kNN, Indian Pines(20 bands),67.33 CART, Indian Pines(20 bands), 63.01 Naive Bayes, Indian Pines(20 bands),52.92 SVM, Salinas scene(20 bands),89.94 kNN, Salinas scene(20 bands),85.02 CART, Salinas scene(20 bands),83.89 Naive Bayes, Salinas scene(20 bands),78.73 SVM, Pavia University(20 bands),89.92 kNN, Pavia University(20 bands),79.42 CART, Pavia University(20 bands),74.06 Naive Bayes, Pavia University (20 bands),67.76
[36]	SWLRSC (Squaring weighted low-rank subspace clustering) method is used for dimensionality reduction	Classifier: SVM with RBF kernel Data Set: Pavia University (15 bands), 68.06 Salinas, (15 bands), 84.45
[37]	Each intrinsic cluster is divided into multiple subsets using the Improved Affinity Propagation algorithm. The initial availability matrix is modified using information entropy to identify a suitable number of clusters	Classifier: SVM with RBF kernel Data Set: Salinas scene Accuracy: 91.45

Advantages of Statistical Approaches in Band Selection:

1. Data-Driven Decisions: Statistical methods rely on real-world data, enabling good decisions that reflect the specific characteristics of the dataset.
2. Redundancy Reduction: These methods help to identify and eliminate redundant bands without loss important information.
3. Enhanced Classification Accuracy: Choosing the most relevant bands can enhance classifier accuracy.
4. Flexibility: Statistical methods can be applied in both supervised and unsupervised manner. For example, MI can be used in a supervised manner to maximize class separability, or in an unsupervised way to focus on intrinsic data properties.
5. Combinability: These methods can be combined with other approaches, such as both clustering or machine learning algorithms are combined to create a hybrid method.

Table 5 Comparative study

Method	Advantages	Disadvantages
Mutual Information	Captures non-linear dependencies, robust to noise	Computationally expensive, estimation challenges
PCA	Reduces dimensionality, maximizes variance, noise reduction	Loss of interpretability, linear assumption, unsupervised
ICA	Identifies independent sources, effective in noise reduction	Sensitive to noise, requires good initialization
Clustering-based	Reduces redundancy, easy implementation	May need predefined number of clusters

Conclusion

All the approaches for band selection directly or indirectly use statistical concepts. Statistical approaches are not only good but often essential in band selection, particularly in hyperspectral imaging, where the goal is to extract the most relevant information from a vast amount of data. These approaches are very simple and effective in the area of hyper spectral image dimensionality reduction. This comparative study mainly focuses the important approaches used in hyperspectral image classification. Information-Theoretic methods are good for balancing informativeness and redundancy. PCA is an efficient approach, but it has a linear nature and may not capture complex spectral relationships. ICA excels at separating mixed signals into their individual components. Clustering-Based methods is a traditional approach to maintaining diversity in the selected bands. To ensure optimal performance, band selection method should be based on the specific characteristics of the hyperspectral data and desired outcomes of the analysis.

References

1. A. F. H. Goetz, "Three decades of hyperspectral remote sensing of the Earth: A personal view," Remote Sens. Environ., vol. 113, pp. S5–S16, Sep. 2009, doi: 10.1016/j.rse.2007.12.014.
2. W. Lan et al., "A method using near infrared hyperspectral imaging to highlight the internal quality of apple fruit slices," Postharvest Biol. Technol., vol. 175, p. 111497, May 2021, doi:

10.1016/j.postharvbio.2021.111497.
3. M. S. Navin and L. Agilandeeswari, “Multispectral and hyperspectral images based land use / land cover change prediction analysis: an extensive review,” Multimed. Tools Appl., vol. 79, no. 39–40, pp. 29751–29774, Oct. 2020, doi: 10.1007/s11042-020-09531-z.
4. X. Tong, H. Xie, and Q. Weng, “Urban Land Cover Classification With Airborne Hyperspectral Data: What Features to Use?,” IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 7, no. 10, pp. 3998–4009, Oct. 2014, doi: 10.1109/JSTARS.2013.2272212.
5. C. M. Gevaert, J. Suomalainen, J. Tang, and L. Kooistra, “Generation of Spectral– Temporal Response Surfaces by Combining Multispectral Satellite and Hyperspectral UAV Imagery for Precision Agriculture Applications,” IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 8, no. 6, pp. 3140–3146, Jun. 2015, doi: 10.1109/JSTARS.2015.2406339.
6. I. Baianu, “Applications Of Microspectroscopy, Hyperspectral Chemical Imaging And Fluorescence Microscopy In Chemistry, Biochemistry, Biotechnology, Molecular And Cell Biology,” Nat. Preced., Nov. 2011, doi: 10.1038/npre.2011.6593.1.
7. M. B. Stuart, M. Davies, M. J. Hobbs, T. D. Pering, A. J. S. McGonigle, and J. R. Willmott, “High-Resolution Hyperspectral Imaging Using Low-Cost Components: Application within Environmental Monitoring Scenarios,” Sensors, vol. 22, no. 12, p. 4652, Jun. 2022, doi: 10.3390/s22124652.
8. B. Melit Devassy and S. George, “Forensic analysis of beverage stains using hyperspectral imaging,” Sci. Rep., vol. 11, no. 1, p. 6512, Mar. 2021, doi: 10.1038/s41598-021-85737-x.
9. M. T. Kuska, J. Behmann, and A.-K. Mahlein, “Potential of hyperspectral imaging to detect and identify the impact of chemical warfare compounds on plant tissue,” Pure Appl. Chem., vol. 90, no. 10, pp. 1615–1624, Oct. 2018, doi: 10.1515/pac-2018-0102.
10. G. Lu and B. Fei, “Medical hyperspectral imaging: a review,” J. Biomed. Opt., vol. 19, no. 1, p. 010901, Jan. 2014, doi: 10.1117/1.JBO.19.1.010901.
11. S. J. Hook, “NASA 2014 The Hyperspectral InfraredImager (HyspIRI) – Science Impact ofDeploying Instruments on SeparatePlatforms.pdf.” JPL CA Inst. Technol., Pasadena, White Paper, 2014.
12. Hanchuan Peng, Fuhui Long, and C. Ding, “Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 8, pp. 1226–1238, Aug. 2005, doi: 10.1109/TPAMI.2005.159.
13. B. Guo, S. R. Gunn, R. I. Damper, and J. D. B. Nelson, “Band Selection for Hyperspectral Image Classification Using Mutual Information,” IEEE Geosci. Remote Sens. Lett., vol. 3, no. 4, pp. 522–526, Oct. 2006, doi: 10.1109/LGRS.2006.878240.
14. B. Guo, Yuesong Lin, Dongliang Peng, and Anke Xue, “Applying differentiable mutual information to hyperspectral band selection,” in 2011 4th International Congress on Image and Signal Processing, Shanghai, China: IEEE, Oct. 2011, pp. 1609–1613. doi: 10.1109/CISP.2011.6100406.
15. El. Sarhrouni, A. Hammouch, and D. Aboutajdine, “Application of Symmetric Uncertainty and Mutual Information to Dimensionality Reduction and Classification of Hyperspectral Images,” Dec. 17, 2012, arXiv: arXiv:1211.0613. Accessed: Sep. 04, 2024. [Online]. Available: http://arxiv.org/abs/1211.0613
16. E. Sarhrouni, A. Hammouch, and D. Aboutajdine, “Band selection and

classification of hyperspectral images using Mutual Information: An algorithm based on minimizing the error probability using the inequality of Fano," in 2012 International Conference on Multimedia Computing and Systems, Tangiers, Morocco: IEEE, May 2012, pp. 155–159. doi: 10.1109/ICMCS.2012.6320192.
17. E. Sarhrouni, A. Hammouch, and D. Aboutajdine, "Feature selection intelligent algorithm with mutual information and steepest ascent strategy," Int. J. Adv. Intell. Paradig., vol. 5, no. 4, p. 257, 2013, doi: 10.1504/IJAIP.2013.058300.
18. C.-I. Chang, Y.-M. Kuo, S. Chen, C.-C. Liang, K. Y. Ma, and P. F. Hu, "Self-Mutual Information-Based Band Selection for Hyperspectral Image Classification," IEEE Trans. Geosci. Remote Sens., vol. 59, no. 7, pp. 5979–5997, Jul. 2021, doi: 10.1109/TGRS.2020.3024602.
19. N. Agrawal and K. Verma, "Hyperspectral Band Selection using Mutual Information," in 2021 International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT), Bhilai, India: IEEE, Feb. 2021, pp. 1–4. doi: 10.1109/ICAECT49130.2021.9392504.
20. Md. P. Uddin, Md. A. Mamun, and Md. A. Hossain, "PCA-based Feature Reduction for Hyperspectral Remote Sensing Image Classification," IETE Tech. Rev., vol. 38, no. 4,pp. 377–396, Jul. 2021, doi: 10.1080/02564602.2020.1740615.
21. C. Rodarmel and J. Shan, "Principal Component Analysis for Hyperspectral Image Classification".
22. Z. Guo, H. Yang, X. Bai, Z. Zhang, and J. Zhou, "Semi-supervised hyperspectral band selection via sparse linear regression and hypergraph models," in 2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS, Melbourne, Australia: IEEE, Jul. 2013, pp. 1474–1477. doi: 10.1109/IGARSS.2013.6723064.
23. M. P. Uddin, M. A. Mamun, and M. A. Hossain, "Feature extraction for hyperspectral image classification," in 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), Dhaka: IEEE, Dec. 2017, pp. 379–382. doi: 10.1109/R10-HTC.2017.8288979.
24. X. Kang, X. Xiang, S. Li, and J. A. Benediktsson, "PCA-Based Edge-Preserving Features for Hyperspectral Image Classification," IEEE Trans. Geosci. Remote Sens., vol. 55, no. 12, pp. 7140–7151, Dec. 2017, doi: 10.1109/TGRS.2017.2743102.
25. J. Dr. J. H. Harikiran, "Hyperspectral image classification using support vector machines," IAES Int. J. Artif. Intell. IJ-AI, vol. 9, no. 4, p. 684, Dec. 2020, doi: 10.11591/ijai.v9.i4.pp684-690.
26. M. Ustuner, "Randomized Principal Component Analysis for Hyperspectral Image Classification," in 2024 IEEE Mediterranean and Middle-East Geoscience and Remote Sensing Symposium (M2GARSS), Apr. 2024, pp. 26–30. doi: 10.1109/M2GARSS57310.2024.10537329.
27. N. Falco, J. A. Benediktsson, and L. Bruzzone, "A Study on the Effectiveness of Different Independent Component Analysis Algorithms for Hyperspectral Image Classification," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 7, no. 6, pp. 2183– 2199, Jun. 2014, doi: 10.1109/JSTARS.2014.2329792.
28. A. Villa, J. A. Benediktsson, J. Chanussot, and C. Jutten, "Independent Component Discriminant Analysis for hyperspectral image classification," in 2010 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, Reykjavik: IEEE, Jun. 2010, pp. 1–4. doi: 10.1109/WHISPERS.2010.5594853.
29. A. Villa, J. A. Benediktsson, J.

Chanussot, and C. Jutten, "Hyperspectral Image Classification With Independent Component Discriminant Analysis," IEEE Trans. Geosci. Remote Sens., vol. 49, no. 12, pp. 4865–4876, Dec. 2011, doi: 10.1109/TGRS.2011.2153861.
30. N. Falco, J. A. Benediktsson, and L. Bruzzone, "Spectral and Spatial Classification of Hyperspectral Images Based on ICA and Reduced Morphological Attribute Profiles," IEEE Trans. Geosci. Remote Sens., vol. 53, no. 11, pp. 6223–6240, Nov. 2015, doi: 10.1109/TGRS.2015.2436335.
31. C. Jayaprakash, B. B. Damodaran, S. V., and K. P. Soman, "Dimensionality Reduction of Hyperspectral Images for Classification using Randomized Independent Component Analysis," in 2018 5th International Conference on Signal Processing and Integrated Networks (SPIN), Noida: IEEE, Feb. 2018, pp. 492–496. doi: 10.1109/SPIN.2018.8474266.
32. S. Jia, Y. Qian, and Z. Ji, "Band Selection for Hyperspectral Imagery Using Affinity Propagation," in 2008 Digital Image Computing: Techniques and Applications, Canberra, Australia: IEEE, 2008, pp. 137–141. doi: 10.1109/DICTA.2008.42.
33. S. Jia, Z. Ji, Y. Qian, and L. Shen, "Unsupervised Band Selection for Hyperspectral Imagery Classification Without Manual Band Removal," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 5, no. 2, pp. 531–543, Apr. 2012, doi: 10.1109/JSTARS.2012.2187434.
34. S. Jia, G. Tang, J. Zhu, and Q. Li, "A Novel Ranking-Based Clustering Approach for Hyperspectral Band Selection," IEEE Trans. Geosci. Remote Sens., vol. 54, no. 1, pp. 88–102, Jan. 2016, doi: 10.1109/TGRS.2015.2450759.
35. Y. Yuan, J. Lin, and Q. Wang, "Dual-Clustering-Based Hyperspectral Band Selection by Contextual Analysis," IEEE Trans. Geosci. Remote Sens., vol. 54, no. 3, pp. 1431– 1445, Mar. 2016, doi: 10.1109/TGRS.2015.2480866.
36. H. Zhai, H. Zhang, L. Zhang, and P. Li, "Squaring weighted low-rank subspace clustering for hyperspectral image band selection," in 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China: IEEE, Jul. 2016, pp. 2434–2437. doi: 10.1109/IGARSS.2016.7729628.
37. Q. Zhu, Y. Wang, F. Wang, M. Song, and C.-I. Chang, "Hyperspectral Band Selection Based on Improved Affinity Propagation," in 2021 11th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS), Amsterdam, Netherlands: IEEE, Mar. 2021, pp. 1–4. doi: 10.1109/WHISPERS52202.2021.9484004.