



Randomized Principal Component Analysis for Hyperspectral Image Classification

Mustafa Ustuner

► To cite this version:

Mustafa Ustuner. Randomized Principal Component Analysis for Hyperspectral Image Classification. 2024. hal-04568194

HAL Id: hal-04568194

<https://hal.science/hal-04568194v1>

Preprint submitted on 3 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

RANDOMIZED PRINCIPAL COMPONENT ANALYSIS FOR HYPERSPECTRAL IMAGE CLASSIFICATION

Mustafa Ustuner

Department of Geomatics Engineering, Artvin Coruh University, 08100 Artvin, Turkey

ABSTRACT

The high-dimensional feature space of the hyperspectral imagery poses major challenges to the processing and analysis of the hyperspectral data sets. In such a case, dimensionality reduction is necessary to decrease the computational complexity. The random projections open up new ways of dimensionality reduction, especially for large data sets. In this paper, the principal component analysis (PCA) and randomized principal component analysis (R-PCA) for the classification of hyperspectral images using support vector machines (SVM) and light gradient boosting machines (LightGBM) have been investigated. In this experimental research, the number of features was reduced to 20 and 30 for classification of two hyperspectral datasets (Indian Pines and Pavia University). The experimental results demonstrated that PCA outperformed R-PCA for SVM for both datasets, but received close accuracy values for LightGBM. The highest classification accuracies were obtained as 0.9925 and 0.9639 by LightGBM with original features for the Pavia University and Indian Pines, respectively.

Index Terms— Hyperspectral, PCA, R-PCA, LightGBM, SVM

1. INTRODUCTION

In recent years, there has been a notable increase in the use of hyperspectral remote sensing images, acquired from airborne or spaceborne sensors, with rapid development in the Earth observation and imaging technology [1, 2, 3]. Hyperspectral images provide rich information about the land objects through narrow spectral bands [1, 4]. Though this increased dimensionality of data provides detailed information, it poses challenges (curse of dimensionality) to conventional techniques, such as classification, for the accurate analysis of hyperspectral images [1, 3, 5].

Hyperspectral images are used in different kinds of applications such as environmental monitoring, crop classification, and target detection. Though hyperspectral images provide essential and rich information regarding the land objects, there are some noisy and redundant spectral bands among the hundreds of bands that might pose challenges in their classification [1, 6, 7]. Applying dimension (feature) reduction

techniques is necessary to remove these redundancies in the hyperspectral data. The main objective of dimension (feature) reduction is to identify the best subset of original or transformed features while retaining as many spectral details as possible [1, 5, 6]. Dimensionality reduction of the hyperspectral data can be implemented in two ways: (1) feature selection or (2) feature extraction. Feature selection aims to identify the most suitable subset of original features based on selection criteria, while feature extraction transforms the data into a much lower-dimensional subspace. Dimension reduction techniques can be classified into three categories: unsupervised, supervised, and semi-supervised approaches [5].

Unsupervised dimensionality reduction techniques can be further grouped into four categories, based on the operation principle: (1) based on information content, (2) projection based, (3) similarity measures and (4) frequency analysis [6]. Additionally, numerous methods, such as PCA, independent component analysis (ICA), kernel PCA, linear discriminant analysis (LDA), R-PCA and local linear feature extraction, have been employed for reducing the dimensionality of hyperspectral images [1, 5]. PCA, a projection-based technique, is the most frequently utilized in hyperspectral image analysis among these methods. It aims to produce a smaller set of orthogonal variables in lower-dimensional subspace from the original features with maximum variance [8, 3]. However, the implementation of PCA can be computationally intensive and time-consuming. Therefore, the use of random projections is considered as an alternative approach for dimensionality reduction in terms of computation time. Random projections are computationally cheap and require less memory compared to classical dimensionality reduction techniques [9, 10]. The randomized PCA, which is a random projection of PCA, is computed via randomized (approximated) singular value decomposition (SVD) of the original data [9, 10, 11].

Several studies have examined the potential for use of R-PCA for feature reduction and classification of the hyperspectral images. Makantasis et al. introduced the R-PCA for dimensionality reduction and combined it with a convolutional neural network (CNN) for encoding spatial and spectral information (denoted as R-PCA CNN) [12]. Most recent studies that utilize R-PCA explore deep learning models for the classification of hyperspectral images. Mei et al. proposed a five-layer CNN (C-CNN) and feature-learning CNN (FL-

CNN) for hyperspectral image classification and compared the performances with state-of-the-art CNN-based methods, R-PCA CNN and CNN based on pixel-pair features (CNN-PPF) [13]. Zhang et al. introduce a framework, called diverse region-based deep CNN model (denoted as DR-CNN) for the classification of hyperspectral images and compare the performance of the DR-CNN with other deep learning and state-of-the-art classifiers such as SVM, CNN, R-PCA CNN, CNN-PPF and SS-CNN [14]. Liu et al. proposed a new method called deep few-shot learning method to address the small sample size problem of hyperspectral image classification. In their experimental research, they compared the performance of their proposed method with several CNN-based methods including R-PCA CNN [15]. Only a few papers have investigated the R-PCA with conventional machine learning algorithms, excluding deep learning. In [12], the classification performance of R-PCA CNN was compared with SVM-based methods (RBF-SVM and Linear-SVM). Damodaran et al. compared the R-PCA and kernel PCA for the dimensionality reduction of hyperspectral images using SVM classification [16]. So, the impacts of R-PCA features on hyperspectral image classification accuracy on supervised machine learning algorithms have not been fully explored yet.

Thus, the main objective of this research is to investigate the impacts of R-PCA on hyperspectral image classification accuracy and compare it with PCA on two public available hyperspectral scenes (Indian Pines (Fig 1) and Pavia University (Fig 2)) [17] using machine learning algorithms.

2. DATASET

In this experimental research, two publicly available and benchmark hyperspectral datasets (Indian Pines and Pavia University) were used and freely downloaded from [17].

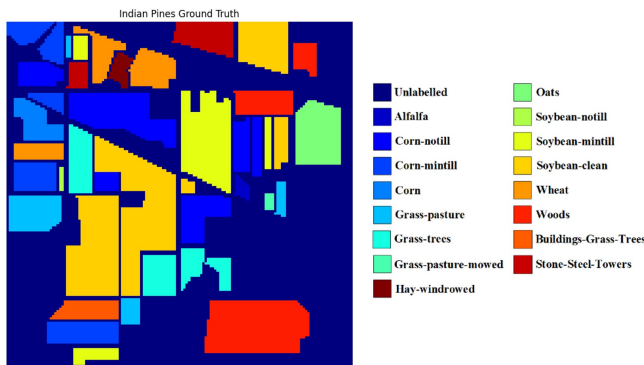


Fig. 1. Ground truth image for Indian Pines scene

The spatial resolutions are 20 m and 1.3 m for Indian Pines and Pavia University scene, respectively. The Indian Pines dataset has originally 220 spectral bands from AVIRIS sensor however 20 spectral bands were removed to exclude the water absorption region [17, 18]. Therefore the remaining 200 spectral bands were used in this experimental research.

There are 16 ground-truth classes in Indian Pines dataset (Table 1). The landscape of the Indian Pines scene comprises mostly agricultural land, approximately two-thirds, with the remaining one-third consisting of forest or other types of perennial natural vegetation [17, 18].

Table 1. Ground truth classes for the Indian Pines scene

	Class	Samples
1	Alfalfa	46
2	Corn-notill	1428
3	Corn-mintill	830
4	Corn	237
5	Grass-pasture	483
6	Grass-trees	730
7	Grass-pasture-mowed	28
8	Hay-windrowed	478
9	Oats	20
10	Soybean-notill	972
11	Soybean-mintill	2455
12	Soybean-clean	593
13	Wheat	205
14	Woods	1265
15	Buildings-Grass-Trees-Drives	386
16	Stone-Steel-Towers	93

The Pavia University dataset contains 103 bands acquired from the airborne sensor ROSIS-03 (Reflective Optics Systems Imaging Spectrometer) and has nine ground-truth classes (Table 2). There are originally 115 spectral bands obtained from the sensor however 12 bands were removed due to noise [17, 19].

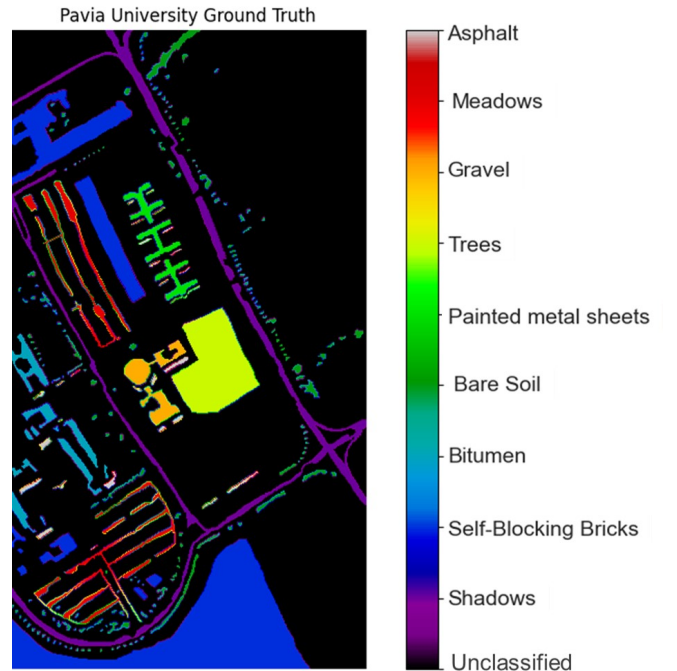


Fig. 2. Ground truth image for Pavia University scene

Table 2. Ground truth classes for the Pavia University scene

#	Class	Samples
1	Asphalt	6631
2	Meadows	18649
3	Gravel	2099
4	Trees	3064
5	Painted metal sheets	1345
6	Bare Soil	5029
7	Bitumen	1330
8	Self-Blocking Bricks	3682
9	Shadows	947

3. DIMENSIONALITY REDUCTION

The high spectral resolution of the hyperspectral data provides advantages for the classification but also presents challenges for conventional signal-processing techniques [5, 20, 7]. For such cases, dimensionality reduction techniques are utilized to map high dimensional space to lower dimensional subspace [1, 7]. The PCA is the most commonly employed feature reduction technique in hyperspectral image classification.

As an alternative to PCA, R-PCA, a linear dimensionality reduction technique like PCA, has been proposed for very large datasets to find approximate principal components using randomized SVD [9, 10, 11]. In the experimental setup, the number of components were reduced to 20 and 30.

4. IMAGE CLASSIFICATION

SVM has been the most frequently used machine learning model for classifying hyperspectral images over the past two decades. This is because SVM which is a kernel-based learning algorithm can handle high dimensional data to distinguish classes when a limited number of training data is available [21] therefore has shown great success in hyperspectral image classification [22, 23, 24]. The optimum parameters ($C=600$ and $\gamma=0.5$) for radial basis function kernel of SVM were determined by grid search strategy using cross validation approach. SVM classification was performed using the open-source Scikit-learn module in Python [25].

To analyze the impacts of R-PCA on image classification accuracy, another machine learning algorithm called LightGBM has been utilized in this experimental research. LightGBM is a cutting-edge ensemble learning algorithm and based on decision trees. It applies the leaf-wise (i.e., vertically) strategy while growing the trees and accelerates the training process by its two new functionalities: Gradient-based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB) [26, 27].

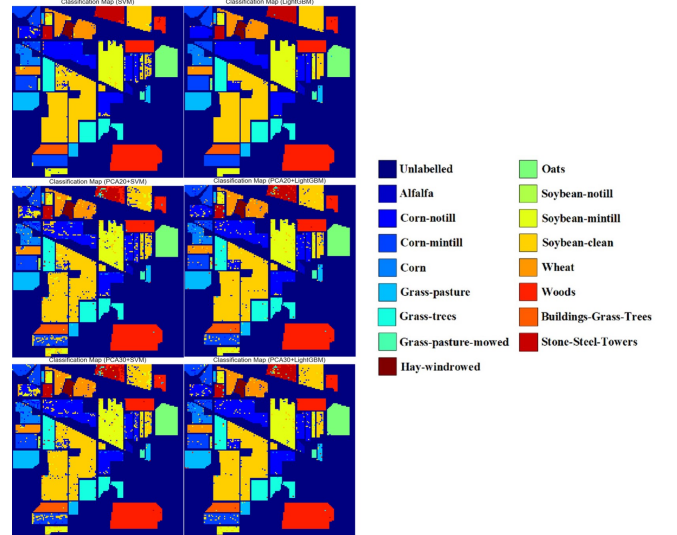
In the experimental setup, the 70% of ground truth dataset selected as training set (randomly) and the remaining parts were selected as test set.

5. EXPERIMENTAL RESULTS AND DISCUSSION

In this experimental research, the impacts of R-PCA on hyperspectral image classification accuracy with two supervised machine learning methods (SVM and LightGBM) have been investigated and the results were compared with PCA on two hyperspectral datasets. The accuracy scores and classification maps were presented on Table 3, Table 4 and Figure 3 and Figure 4 for Indian Pines and Pavia University scenes, respectively.

Table 3. Accuracy Scores for Indian Pines

Indian Pines					
Methods	Original	PCA-20	RPCA-20	PCA-30	RPCA-30
SVM	0.9473	0.8657	0.8471	0.8803	0.8354
LightGBM	0.9639	0.8481	0.3148	0.8530	0.8546

**Fig. 3.** Classification Maps for the Indian Pines scene

The highest classification accuracies (overall accuracies) were received as 0.9925 and 0.9639 by LightGBM with original features for Pavia University and Indian Pines, respectively. For both datasets, the original features gives better results than the dimensional reduced features in terms of overall accuracies by SVM and LightGBM. For the Indian Pines scene, the PCA with SVM for both first 20 and 30 principal components received higher accuracies than R-PCA however R-PCA with LightGBM slightly outperformed PCA for the first 30 principal components (0.8546 vs 0.8530).

Table 4. Accuracy Scores for Pavia University

Pavia University					
Methods	Original	PCA-20	RPCA-20	PCA-30	RPCA-30
SVM	0.9760	0.9565	0.9327	0.9577	0.9342
LightGBM	0.9925	0.9553	0.9550	0.9540	0.9539

For the Pavia University scene, the PCA with SVM for both first 20 and 30 principal components received higher

accuracies than R-PCA however for the R-PCA with LightGBM, the classification accuracies are very close (equal) to those received with PCA. The statistical significance between the accuracies were evaluated by McNemar's test (with the 95% confidence interval) and it was found that there is no significantly difference between R-PCA and PCA for the classification of Pavia University scene with LightGBM. For the classification of the original data, the LightGBM received higher accuracy than SVM for Pavia University.

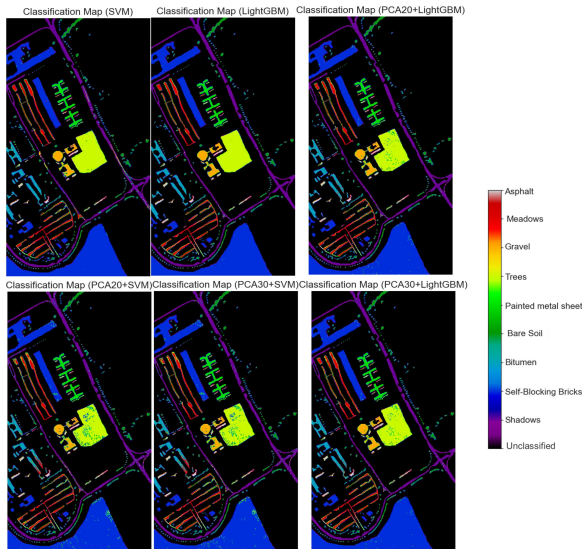


Fig. 4. Classification Maps for Pavia University scene

6. CONCLUSION

In most of the previous studies, R-PCA was investigated along with CNN-based methods however has not been explored yet with supervised learning algorithms rather than SVM for hyperspectral image classification. With this research, the impacts of R-PCA was investigated for the classification of hyperspectral images using LightGBM and SVM and the classification performance was compared with PCA. This research has also tested the sensitivity of the kernel-based learning and ensemble learning algorithms to the same dataset for dimensionality reduction. This experimental results demonstrated that the original features yield better results than PCA and RPCA. The highest classification accuracies for the Indian Pines and Pavia University were received as 0.9639 and 0.9925 by LightGBM, respectively.

7. ACKNOWLEDGEMENT

The author would like to thank Prof. D. A. Landgrebe for providing the Indian Pines data set and Prof. P. Gamba for providing the Pavia University data set. Also, thanks to Computational Intelligence Group of the University of the Basque Country for making the datasets freely available at [17]

8. REFERENCES

- [1] X. Jia, B.-C. Kuo, and M. M. Crawford, "Feature mining for hyperspectral image classification," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 676–697, 2013.
- [2] L. He, J. Li, C. Liu, and S. Li, "Recent advances on spectral-spatial hyperspectral image classification: An overview and new guidelines," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 3, pp. 1579–1597, 2018.
- [3] B. Rasti, D. Hong, R. Hang, P. Ghamisi, X. Kang, J. Chanussot, and J. A. Benediktsson, "Feature extraction for hyperspectral imagery: The evolution from shallow to deep: Overview and toolbox," *IEEE Geoscience and Remote Sensing Magazine*, vol. 8, no. 4, pp. 60–88, 2020.
- [4] Md. A. Mamun Md. P. Uddin and Md. A. Hossain, "Pca-based feature reduction for hyperspectral remote sensing image classification," *IETE Technical Review*, vol. 38, no. 4, pp. 377–396, 2021.
- [5] P. Ghamisi, N. Yokoya, J. Li, W. Liao, S. Liu, J. Plaza, B. Rasti, and A. Plaza, "Advances in hyperspectral image and signal processing: A comprehensive overview of the state of the art," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 4, pp. 37–78, 2017.
- [6] D. R. Hidalgo, B. B. Cortés, and E. C. Bravo, "Dimensionality reduction of hyperspectral images of vegetation and crops based on self-organized maps," *Information Processing in Agriculture*, vol. 8, no. 2, pp. 310–327, 2021.
- [7] Gustavo Camps-Valls, Devis Tuia, Lorenzo Bruzzone, and Jon Atli Benediktsson, "Advances in hyperspectral image classification: Earth monitoring with statistical learning methods," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 45–54, 2014.
- [8] J. Senthilnath, S. N. Omkar, V. Mani, N. Karnwal, and Shreyas P. B., "Crop stage classification of hyperspectral data using unsupervised techniques," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 2, pp. 861–866, 2013.
- [9] N. Halko, P. G. Martinsson, and J. A. Tropp, "Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions," *SIAM Review*, vol. 53, no. 2, pp. 217–288, 2011.
- [10] M. Wojnowicz, D. Zhang, G. Chisholm, X. Zhao, and M. Wolff, "Projecting "better than randomly": How to reduce the dimensionality of very large datasets in a way

- that outperforms random projections,” in *2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, 2016, pp. 184–193.
- [11] H. Firat, M. E. Asker, and D. Hanbay, “Classification of hyperspectral remote sensing images using different dimension reduction methods with 3d/2d cnn,” *Remote Sensing Applications: Society and Environment*, vol. 25, pp. 100694, 2022.
 - [12] K. Makantasis, K. Karantzalos, A. Doulamis, and N. Doulamis, “Deep supervised learning for hyperspectral data classification through convolutional neural networks,” in *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2015, pp. 4959–4962.
 - [13] S. Mei, J. Ji, J. Hou, X. Li, and Q. Du, “Learning sensor-specific spatial-spectral features of hyperspectral images via convolutional neural networks,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 8, pp. 4520–4533, 2017.
 - [14] M. Zhang, W. Li, and Q. Du, “Diverse region-based cnn for hyperspectral image classification,” *IEEE Transactions on Image Processing*, vol. 27, no. 6, pp. 2623–2634, 2018.
 - [15] B. Liu, X. Yu, A. Yu, P. Zhang, G. Wan, and R. Wang, “Deep few-shot learning for hyperspectral image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 4, pp. 2290–2304, 2019.
 - [16] B. B. Damodaran, N. Courty, and R. Tavenard, “Randomized nonlinear component analysis for dimensionality reduction of hyperspectral images,” in *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2017, pp. 5–8.
 - [17] Grupo de Inteligencia Computacional, “Hyperspectral remote sensing scenes,” http://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes/, 2021, Accessed: 2023-11-10.
 - [18] Ping Ma, Jinchang Ren, Huimin Zhao, Genyun Sun, Paul Murray, and Jiangbin Zheng, “Multiscale 2-d singular spectrum analysis and principal component analysis for spatial-spectral noise-robust feature extraction and classification of hyperspectral images,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 1233–1245, 2021.
 - [19] Ying Li, Haokui Zhang, and Qiang Shen, “Spectral-spatial classification of hyperspectral imagery with 3d convolutional neural network,” *Remote Sensing*, vol. 9, no. 1, 2017.
 - [20] Pedram Ghamisi, Javier Plaza, Yushi Chen, Jun Li, and Antonio J Plaza, “Advanced spectral classifiers for hyperspectral images: A review,” *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 1, pp. 8–32, 2017.
 - [21] J.A. Gualtieri and S. Chettri, “Support vector machines for classification of hyperspectral data,” in *IGARSS 2000. IEEE 2000 International Geoscience and Remote Sensing Symposium. Taking the Pulse of the Planet: The Role of Remote Sensing in Managing the Environment. Proceedings (Cat. No.00CH37120)*, 2000, vol. 2, pp. 813–815 vol.2.
 - [22] G. Camps-Valls and L. Bruzzone, “Kernel-based methods for hyperspectral image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 6, pp. 1351–1362, 2005.
 - [23] M. Pal and P. M. Mather, “Support vector machines for classification in remote sensing,” *International Journal of Remote Sensing*, vol. 26, no. 5, pp. 1007–1011, 2005.
 - [24] Mahdi Hasanlou, Farhad Samadzadegan, and Saeid Homayouni, “Svm-based hyperspectral image classification using intrinsic dimension,” *Arabian Journal of Geosciences*, vol. 8, pp. 477–487, 2015.
 - [25] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay, “Scikit-learn: Machine learning in python,” *Journal of Machine Learning Research*, vol. 12, no. 85, pp. 2825–2830, 2011.
 - [26] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T. Liu, “Lightgbm: A highly efficient gradient boosting decision tree,” in *Advances in Neural Information Processing Systems*, I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds. 2017, vol. 30, Curran Associates, Inc.
 - [27] H. Jafarzadeh, M. Mahdianpari, E. Gill, F. Mohammadianesh, and S. Homayouni, “Bagging and boosting ensemble classifiers for classification of multispectral, hyperspectral and polsar data: A comparative evaluation,” *Remote Sensing*, vol. 13, no. 21, 2021.