EVALUATION OF MACHINE LEARNING FOR HYPERSPECTRAL REMOTE SENSING IMAGE CLASSIFICATION

Jun-Wei Guo[1], Fuan Tsai*[2]
[1]Undergraduate Student, Department of Space Science & Engineering, National Central University, Taiwan.
No. 300, Jhongda Rd., Jhongli District, Taoyuan City 32001, Taiwan[1]
Email: 96007peter@gmail.com

[2]Professor, Center for Space and Remote Sensing Research, National Central University, Taiwan.
No. 300, Jhongda Rd., Jhongli District, Taoyuan City 32001, Taiwan[2]
Email: ftsai@csrsr.ncu.edu.tw

KEY WORDS: Hyperspectral Image, Machine Learning, Image Classification, Feature Extraction

ABSTRACT: As technology advances, hyperspectral remote sensing image has become an exciting area of research. With the advantages of hyperspectral images (HSI), such as detailed spectral information, there is a great potential for sophisticated applications across diversified fields, including enhancing image classification accuracy. Hyperspectral imagery may achieve more precise object differentiation. However, dealing with high-dimensional hyperspectral data presents challenges, including the curse of dimensionality and redundant information between bands. To overcome these obstacles, integrating machine learning techniques offers effective solutions. In this research, feature extraction is performed using the Minimum Noise Fraction (MNF) technique to obtain relevant and important information. Subsequently, three machine learning classification algorithms, including Support Vector Machine (SVM), Random Forest (RF), and eXtreme Gradient Boosting (XGBoost), are applied for classification. The research utilizes two airborne hyperspectral benchmark datasets, Indian Pines and Salinas, obtained from JPL's Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) project. The datasets consist of samples from various land cover categories, including forests, agricultural areas, bare soil, and vegetation, with samples representing different growth stages. Both datasets contain 16 classes of samples but cover different categories. Additionally, the study intends to incorporate EO-1 Hyperion hyperspectral satellite image, which provides a broader coverage area, serving as supplementary research data. Moreover, SPOT imagery data is utilized as a reference to evaluate the results of Hyperion classification. The objective of this research is to integrate the advantages of various algorithms and find a suitable and efficient classification process for hyperspectral image data, which can be successfully applied to both airborne and satellite images. Experimental results show that this process can achieve satisfactory classification results for both airborne and satellite hyperspectral images. All three classification algorithms produce acceptable results and the Random Forest algorithm produces the best classification accuracy. The best results show an overall accuracy of 0.86 and 0.96 for Indian Pines and Salinas airborne images, respectively. The best overall accuracy of the satellite image is 0.84 and Kappa is 0.79.

1. INTRODUCTION

Since the development of hyperspectral image (HSI), it has been frequently used in the remote sensing field for different applications. The advantages of the hyperspectral image lie in its wide spectral range, high spectral resolution, a large number of bands, and continuity, which can provide detailed and rich spectral information. Enables the identification of subtle changes in various targets. However, the information contents of a hyperspectral image lie in a lower dimensional subspace, due to the high correlations between the reflectance of the neighboring bands (Gewali et al., 2018). Dimension reduction is necessary before the algorithm can be used effectively for classification (Camps-Valls et al., 2013). Therefore, feature extraction is used as the dimension reduction method in this study. The main algorithm uses the Minimum Noise Fraction (MNF) technique.

Machine learning (ML) algorithms due to their outstanding predictive power have become a key tool for modern hyperspectral image analysis (Gewali et al., 2018). Classification is one of the hottest research topics in the hyperspectral field. In the past, a huge number of methods were proposed to deal with the hyperspectral data classification problem (Chen et al., 2014). With the availability of sensor data of various resolutions. Comparison and testing of different classification algorithms for various applications are also necessary (Lu and Weng, 2007). Therefore, this study proposes to utilize three different classification algorithms for the performance present of hyperspectral images and to establish a suitable processing procedure for such imagery. The three classification algorithms include Support Vector Machine (SVM), Random Forest (RF), and eXtreme Gradient Boosting (XGBoost).

The datasets utilize two airborne hyperspectral data. The datasets encompass forests, agricultural areas, bare ground, vegetation, etc. The samples have different growth periods, which are difficult to recognize by ordinary multispectral images (MSI). In addition, we also use EO-1 Hyperion hyperspectral satellite images to confirm the usefulness of the final complete process. In summary, this study aims to compare and analyze the classification results of three algorithms and improve the accuracy. Effectively find the appropriate spectral information by feature extraction. Identify the appropriate classification process for hyperspectral data and successfully apply it to different images.

2. METHODOLOGY

2.1 HYPERSPECTRAL IMAGE

Hyperspectral image has a higher spectral resolution and wider spectral range than multispectral image. The number of bands is large and continuous, generating tens to hundreds of bands. It extends from visible light to near-infrared (NIR) and short-wave infrared (SWIR), with each band as narrow as 10 nanometers (Shaw and Burke, 2003). This enhances the depth of spectral information and enables the discrimination of a wide range of substances. It is widely used in various fields, such as environmental monitoring, geological research, and vegetation research (Thenkabail et al., 2016).

2.2 FEATURE EXTRACTION

Despite hyperspectral images providing rich spectral data, increasing band numbers and data volume lead to heightened inter-band correlations, resulting in information redundancy. An abundance of information may result in the curse of dimensionality. To address this challenge, the typical solution involves employing feature extraction to reduce dimensionality. This approach revolves around converting data into informative and non-redundant features, enhancing the capacity to find the most important features in the feature space (Benediktsson et al., 2005). In the research conducted by (Lu and Weng, 2007), it is emphasized that selecting suitable variables is a critical step for successfully implementing image classification.

2.2.1 MINIMUM NOISE FRACTION

The concept of Minimum Noise Fraction (MNF) transformation (Green et al., 1988). The reason lies in the fact that the image quality of principal component images obtained after Principal Component Analysis (PCA) transformation often does not align with their sorting order. MNF transformation aims to enhance this by sorting images based on their quality. The principle is based on PCA and incorporates the concept of maximizing the Signal-to-Noise Ratio (SNR). During the transformation, images are sorted based on their SNR values (Uddin et al., 2021). In summary, this approach involves performing two consecutive layers of PCA transformation. The first step involves applying a high-pass filter template to the image, resulting in a new space projection of the original image. In the generated image, the noise possesses unit variance and is uncorrelated across bands. The second step is to perform PCA on the noise-whitened data. By implementing these two transformations, the MNF transformation matrix is obtained. Through MNF transformation, it can reduce the dimension of the hyperspectral images and separate the noise from the hyperspectral images. Furthermore, the results from the research by (Mundt et al., 2005) indicate that MNF transformation exhibits superior performance compared to PCA transformation.

2.3 MACHINE LEARNING CLASSIFICATION ALGORITHM

Machine learning has become the primary tool for the analysis of hyperspectral images (Gewali et al., 2018). In scenarios where direct information extraction from data is not always feasible, the prevalent approach involves applying machine learning methods for practical utilization.

2.3.1 SUPPORT VECTOR MACHINE

Support Vector Machine (SVM) is a machine learning algorithm constructed from statistical learning theory (Cortes and Vapnik, 1995). Research has indicated that SVM's ability to handle high-dimensional data is often leveraged for hyperspectral data classification (Ghamisi et al., 2017). Furthermore, it is widely utilized in various machine learning algorithms (Mahesh, 2020). The primary core concept of SVM is to generate a decision boundary with the maximum margin of separation between data samples belonging to different classes (Gewali et al., 2018). In cases of linearly separable data, direct classification is feasible. However, for non-linear data, the utilization of kernel trick offers a solution. Applying kernel transformation involves mapping the data's input feature space into a higher-dimensional space (Melgani et al., 2004). The kernel trick involves identifying a function in the original space that yields the outcome of the mapped inner product, referred to as the kernel function. In SVM, the fundamental kernel function is defined as (Hsu et al., 2003):

Radial Basis Function (RBF): $K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0$ (1)

Polynomial: $K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0$ (2)

Linear: $K(x_i, x_j) = x_i^T x_j$ (3)

where (x_i, x_j) represents a pair of data labels, γ, r, and d are kernel parameters. The choice of the kernel is an important issue in the SVM algorithm, and its performance largely depends on the kernel (Patle and Chouhan, 2013).

2.3.2 RANDOM FOREST

Random Forest (RF) (Breiman, 2001) is an algorithm that builds upon Decision Trees (DT) to create an enhanced method. It is an ensemble learning, which merges the results from multiple base predictors to produce a more accurate result (Gewali et al., 2018). This classifier has been widely used in conjunction with hyperspectral data (Ghamisi et al., 2017). One of its characteristics is the ability to reduce data features. Uses a reduced feature set to learn each ensemble member, which makes them less prone to overfitting (Gewali et al., 2018). Furthermore, even in scenarios with high dimensionality, imbalanced training samples, or missing values, RF still can provide satisfactory classification results. Another advantage of the RF classifier is that it is insensitive to noise in the training labels (Ghamisi et al., 2017), which contributes to the effective solution of issues in hyperspectral data.

2.3.3 EXTREME GRADIENT BOOSTING

eXtreme Gradient Boosting (XGBoost) (Chen et al., 2016). It is an algorithm derived from Decision Trees (DT). Similar to RF, XGBoost lies in the construction of DT. In RF, each DT generates independently, whereas XGBoost generates DTs sequentially, establishing correlations between them. Moreover, it continuously improves upon the errors generated by previous classifications. The core concept involves the combination of multiple weak classifiers into a strong classifier, thereby achieving better classification results.

2.4 ACCURACY ASSESSMENT

After applying different classification algorithms, assessing the quality of the classification results is necessary. The confusion matrix is a powerful tool for performance assessment by quantifying the classification overlap (Heydarian et al., 2022). The Confusion Matrix encompasses various metrics such as precision, overall accuracy, kappa, etc., which are used to assess the performance of individual classes and overall performance. The relevant formulas are shown as:

$$precision_c = \frac{TP}{TP+FP} \quad (4)$$

$$overall\ accuracy = \frac{TP+TN}{TP+FN+TN+FP} \quad (5)$$

$$kappa = \frac{2\cdot(TP\cdot TN - FP\cdot FN)}{(TP+FP)\cdot(FP+TN)+(TP+FN)\cdot(FN+TN)} \quad (6)$$

where c represents each class, T represents the true, F represents the false, P represents the positive, and N represents the negative. This study primarily employs overall accuracy and kappa as the evaluation criterion.

3. EXPERIMENTS

3.1 DATASETS AND STUDY AREA

Setup

For all the data, we split the datasets into 30% for testing and 70% for training, ensuring that each generated dataset was randomized. As this research employed Python for classification and aimed to address the issue of imbalanced samples, the training parameters were set with the option: class_weight = ' balanced '. Each class weight is calculated as: number of samples / (total number of classes * number of samples for this class). A weight is assigned to each class within the training set. Classes with a higher number of samples receive lower weights, while those with fewer samples are

assigned higher weights. Applying this method will enhance result accuracy and prevent the occurrence of sample imbalance issues.

Indian Pines dataset and Salinas dataset

Both airborne datasets were acquired using the AVIRIS sensor, encompassing 224 bands, covering a wavelength range of 0.4-2 μm. The Indian Pines test site in north-western Indiana, with a size of 145×145 pixels and a spatial resolution of 20 meters. The Salinas test site in Salinas Valley, California, with a size of 512×217 pixels and a spatial resolution of 3.7 meters. Two datasets also contain 16 classes, with some specific details shown in Figures 1 and 2 and Table 1.

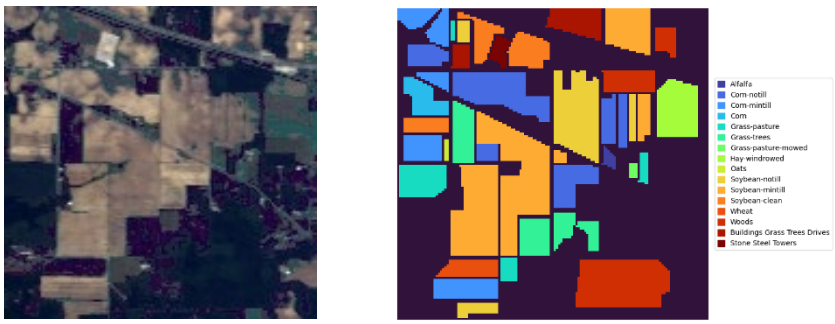


Figure 1. (left) Indian Pines false-color composite image (bands 29, 19, and 9). (right) Ground truth.

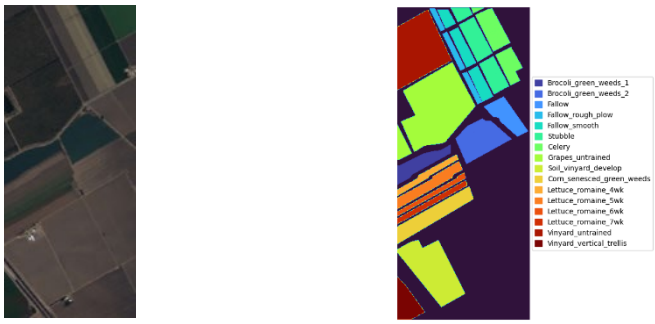


Figure 2. (left) Salinas false-color composite image (bands 29, 19, and 9). (right) Ground truth.

Table 1. Original sample details for the Indian Pines and Salinas

Class No.	Class (Indian Pines)	Number of Samples	Class (Salinas)	Number of Samples
1	Alfalfa	46	Brocoli_green_weeds_1	2009
2	Corn-notill	1428	Brocoli_green_weeds_2	3726
3	Corn-mintill	830	Fallow	1976
4	Corn	237	Fallow_rough_plow	1394
5	Grass-pasture	483	Fallow_smooth	2678
6	Grass-trees	730	Stubble	3959
7	Grass-pasture-mowed	28	Celery	3579
8	Hay-windrowed	478	Grapes_untrained	11271
9	Oats	20	Soil_vinyard_develop	6203
10	Soybean-notill	972	Corn_senesced_green_weeds	3278
11	Soybean-mintill	2455	Lettuce_romaine_4wk	1068
12	Soybean-clean	593	Lettuce_romaine_5wk	1927
13	Wheat	205	Lettuce_romaine_6wk	916
14	Woods	1265	Lettuce_romaine_7wk	1070
15	Buildings-Grass-Trees Drives	386	Vinyard_untrained	7268
16	Stone-Steel-Towers	93	Vinyard_vertical_trellis	1807

EO-1 Hyperion

The EO-1 Hyperion hyperspectral satellite image consists of 242 bands, covering a wavelength range of approximately 0.35 to 2.58 μm (Tsai et al., 2006). The study area is located in Taipei, Taiwan. The image size is 373×532 pixels, with a

spatial resolution of 30 meters. This image was acquired on July 6, 2002. The sample classes include building, vegetation, crop, river, and bridges & highways. Moreover, the SPOT-4 image is utilized as a reference to evaluate the results. The SPOT-4 image was acquired on January 8, 2002. The spatial resolution is multispectral: 20 m and panchromatic: 10 m. For more information about EO1-Hyperion image data and SPOT-4 image in this study are shown in Figure 3, Figure 4, and Table 2.

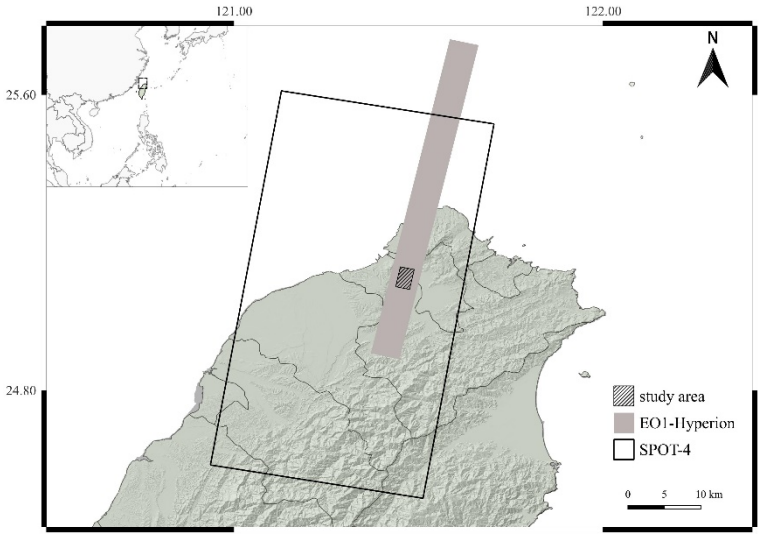
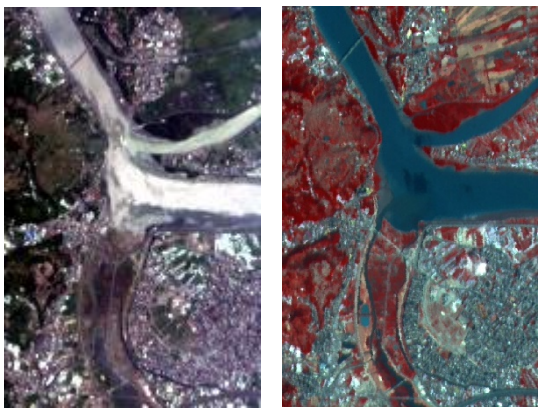


Figure 3. Study area



(a) (b)

Table 2. Sample details for the EO-1 Hyperion

Class No.	Class	Number of Samples
1	building	346
2	vegetation	304
3	crop	200
4	river	442
5	bridges & highways	135

Figure 4. (a) Hyperspectral image of the study area (bands 29, 19, and 9) (EO1-Hyperion image courtesy of the U.S. Geological Survey). (b) Reference pansharpened image. SPOT-4 © CNES 2002, Distribution Airbus DS.

The pan-sharpening technique was applied to the SPOT-4 image in this study to enhance object recognition and spatial resolution. Pan-sharpening is the process of merging a high-resolution panchromatic image with a lower-resolution multispectral image to create a single high-resolution color image, as illustrated in Figure 5.

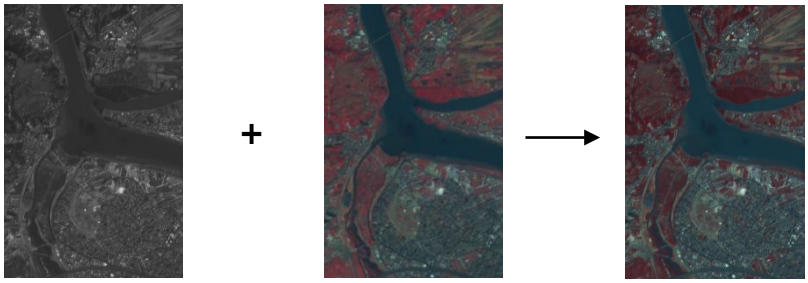


Figure 5. Image pan-sharpening with SPOT-4 images. (left) panchromatic image. (middle) multispectral image. (right) pansharpened image.

3.2 OUTLINE OF THE METHOD

The research flowchart is illustrated in Figure 8. Most of the processes in this study were carried out using Python, while some pre-processing tasks were performed using the Environment for Visualizing Images (ENVI). Firstly, both satellite and airborne images undergo pre-processing. After the MNF transformation, a suitable number of images are selected. Following this, three machine learning classification methods are applied. Finally, the classification results are evaluated.

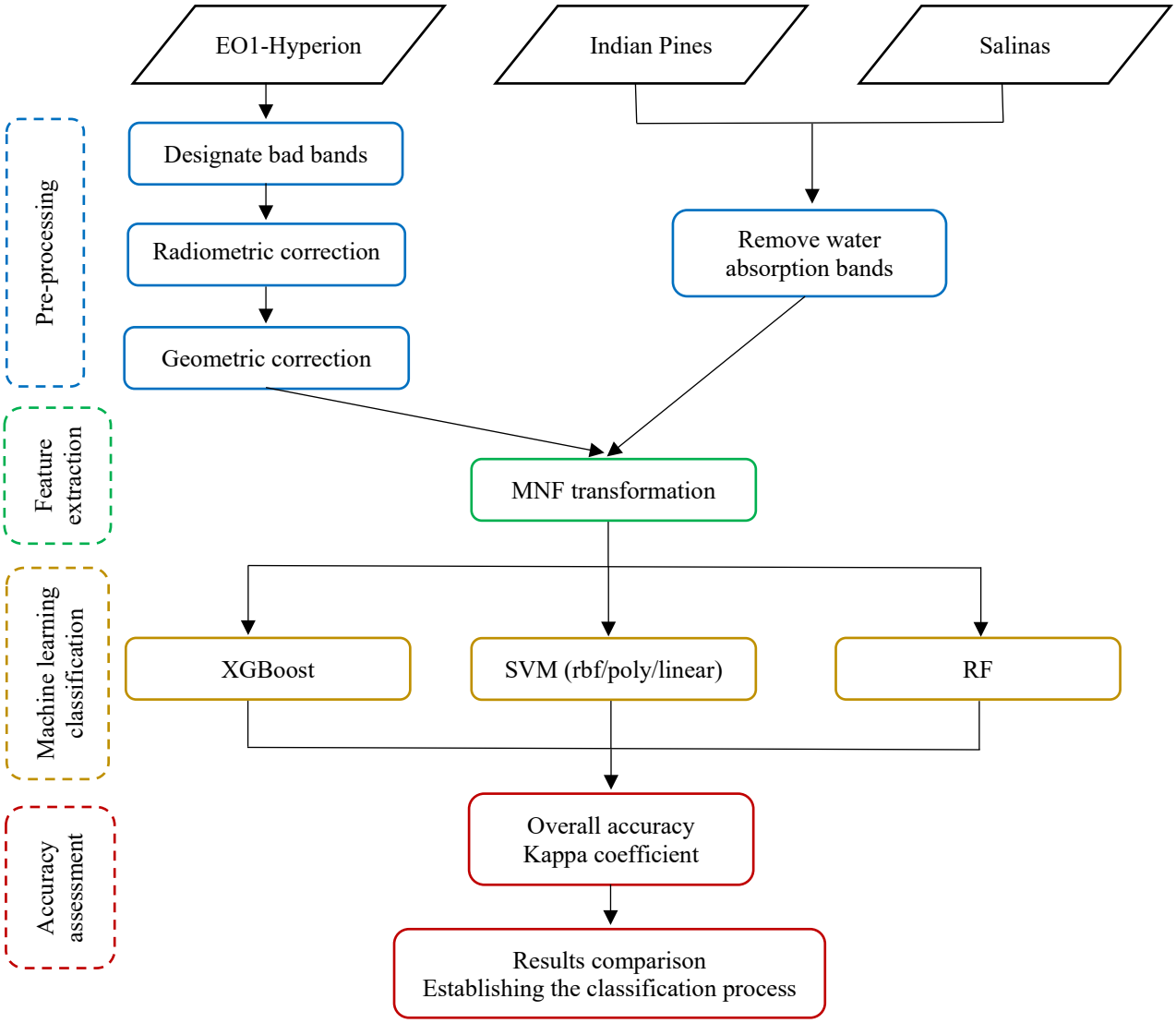


Figure 6. Research flowchart

3.3 PREPROCESSING

The Indian Pines and Salinas hyperspectral images have undergone pre-processing steps such as removing water absorption bands. As a result, the initial set of 224 bands has been reduced. The Indian Pines hyperspectral image now has 200 bands, while the Salinas hyperspectral image has been reduced to 204 bands. The Hyperion image also required the removal of bad bands, including not illuminated (1-7, 222-242), overlap regions (58-76), and water absorption bands (121-126, 167-180). Following the reduction from 242 to 175 bands, subsequent steps included radiometric correction, atmospheric correction, and geometric correction.

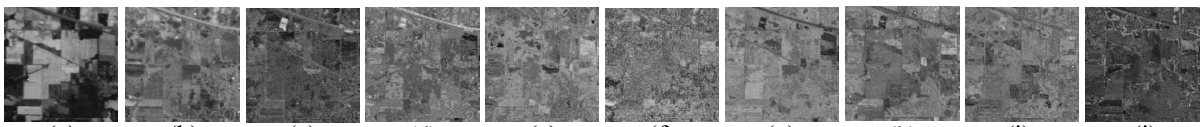
3.4 ALGORITHM SETUP

In this study, as mentioned earlier, Python was predominantly used for the feature extraction and classification process. Algorithm and parameter tuning are necessary in this regard. The data will be adjusted initially (train_size = 0.7, test_size = 0.3). Furthermore, to address data imbalance, the class_weight = ' balanced ' parameter is employed in each algorithm, and the stratify = y approach is used. Some slight parameter adjustments were also used for the three machine learning classification algorithms. The most noticeable adjustments were made to the SVM classifier, where three different kernel functions were employed. The three kernel functions are the radial basis function (kernel = ' rbf '), polynomial (kernel = ' poly '), and linear (kernel = ' linear ').

4. RESULTS AND DISCUSSIONS

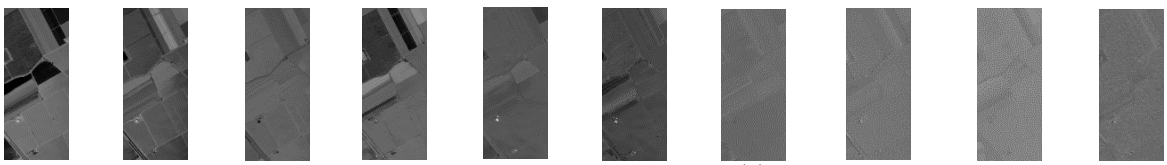
4.1 MINIMUM NOISE FRACTION RESULT

To reduce the dimensionality, this study uses minimum noise fraction transformation. After MNF transformation, we choose to retain a minimum of 0.999 of the image cumulative explained variance. The Indian Pines image is left with 17 components. The Salinas image is left with 25 components. The EO-1 Hyperion image is left with 31 components. The result in Figure 7, Figure 8, and Figure 9 shows the first ten components extracted by MNF.



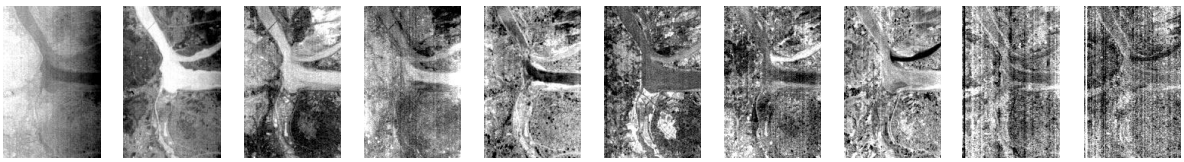
(a) (b) (c) (d) (e) (f) (g) (h) (i) (j)

Figure 7. (a) ~ (j) the first ten components extracted by the MNF from the Indian Pines image.



(a) (b) (c) (d) (e) (f) (g) (h) (i) (j)

Figure 8. (a) ~ (j) the first ten components extracted by the MNF from the Salinas image.



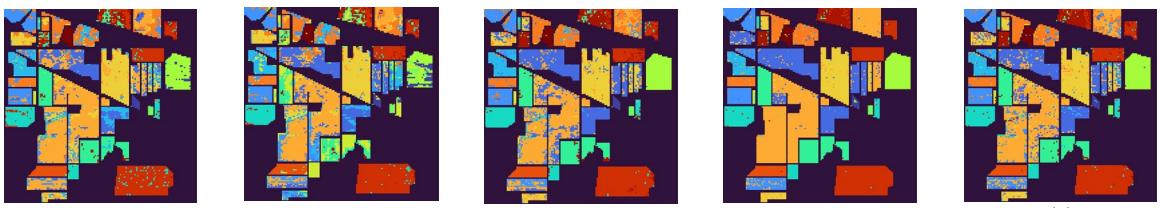
(a) (b) (c) (d) (e) (f) (g) (h) (i) (j)

Figure 9. (a) ~ (j) the first ten components extracted by the MNF from the EO-1 Hyperion image.

From Fig 7 to Fig 9, it can be seen that the quality of the image decreases by (a) ~ (f). The quality decrement is most noticeable in EO-1 Hyperion images.

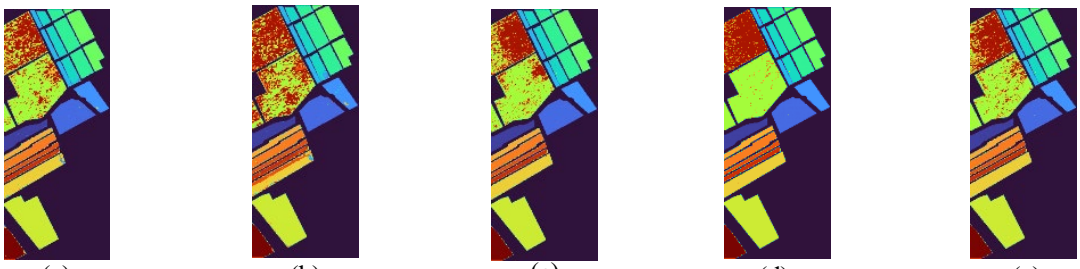
4.2 CLASSIFICATION RESULT

After obtaining the MNF transformation results, we select appropriate components for classification. Some of the data are subjected to sample selection before being classified. The three datasets are classified using SVM (rbf / poly / linear), RF, and XGBoost respectively. The classification results are shown in Figure 10, Figure 11, and Figure 12. At the same time, the overall accuracy (OA) and kappa coefficient (kappa) were calculated and are presented in Tables 3, 4, and 5.



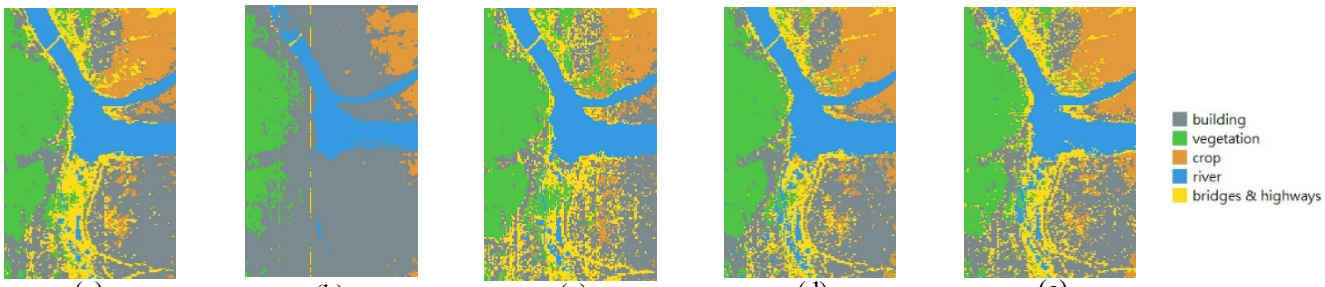
(a) (b) (c) (d) (e)

Figure 10. Classification results for Indian Pines data. (a) SVM_rbf, (b) SVM_poly, (c) SVM_linear, (d) RF, and (e) XGBoost



(a) (b) (c) (d) (e)

Figure 11. Classification results for Salinas data. (a) SVM_rbf, (b) SVM_poly, (c) SVM_linear, (d) RF, and (e) XGBoost



(a) (b) (c) (d) (e)

Figure 12. Classification results for EO1-Hyperion data. (a) SVM_rbf, (b) SVM_poly, (c) SVM_linear, (d) RF, and (e) XGBoost

Table 3. Accuracy values obtained for the Indian Pines

	SVM_rbf	SVM_poly	SVM_linear	RF	XGBoost
Alfalfa	0.42	0.20	0.93	1.00	0.85
Corn-notill	0.77	0.81	0.72	0.82	0.77
Corn-mintill	0.62	0.51	0.68	0.89	0.69
Corn	0.31	0.26	0.54	0.78	0.48
Grass-pasture	0.75	0.86	0.83	0.96	0.86
Grass-trees	0.88	0.81	0.95	0.90	0.93
Grass-pasture-mowed	0.54	0.13	1.00	1.00	0.86
Hay-windrowed	0.99	1.00	0.99	0.95	0.98
Oats	0.60	0.02	0.83	1.00	0.50
Soybean-notill	0.60	0.62	0.70	0.85	0.65
Soybean-mintill	0.71	0.85	0.82	0.79	0.81
Soybean-clean	0.51	0.46	0.61	0.88	0.65
Wheat	0.92	0.90	0.94	0.98	0.98
Woods	0.89	0.94	0.96	0.91	0.92
Buildings-Grass-Trees Drives	0.67	0.68	0.65	0.90	0.68
Stone-Steel-Towers	0.97	0.90	0.96	0.96	0.96
OA	**0.7057**	**0.6449**	**0.7860**	**0.8576**	**0.7854**
kappa	**0.6682**	**0.6051**	**0.7588**	**0.8361**	**0.7572**

Table 4. Accuracy values obtained for the Salinas

	SVM_rbf	SVM_poly	SVM_linear	RF	XGBoost
Brocoli_green_weeds_1	1.00	1.00	1.00	1.00	1.00
Brocoli_green_weeds_2	0.99	1.00	1.00	1.00	1.00
Fallow	0.93	0.86	1.00	1.00	1.00
Fallow_rough_plow	0.99	0.99	1.00	1.00	1.00
Fallow_smooth	0.99	0.98	0.99	1.00	0.99
Stubble	1.00	1.00	1.00	1.00	1.00
Celery	1.00	1.00	1.00	1.00	1.00
Grapes_untrained	0.81	0.82	0.84	0.87	0.87
Soil_vinyard_develop	0.99	0.99	1.00	1.00	1.00
Corn_senesced_green_weeds	0.84	0.82	0.99	0.99	0.97
Lettuce_romaine_4wk	0.93	0.90	1.00	1.00	0.99
Lettuce_romaine_5wk	0.97	0.95	1.00	1.00	0.99
Lettuce_romaine_6wk	0.95	0.38	1.00	1.00	1.00
Lettuce_romaine_7wk	0.96	0.93	0.99	0.98	0.98
Vinyard_untrained	0.17	0.55	0.72	1.90	0.76
Vinyard_vertical_trellis	0.99	0.99	1.00	1.00	1.00
OA	**0.9018**	**0.8387**	**0.9286**	**0.9563**	**0.9353**
kappa	**0.8909**	**0.8221**	**0.9175**	**0.9513**	**0.9281**

Table 5. Accuracy values obtained for the EO1-Hyperion

	SVM_rbf	SVM_poly	SVM_linear	RF	XGBoost
building	0.90	0.50	0.70	1.00	0.94
vegetation	0.86	1.00	0.80	0.90	0.86
crop	0.39	0.50	0.50	0.56	0.63
river	1.00	1.00	1.00	1.00	1.00
bridges & highways	0.64	0.20	0.52	0.46	0.56
OA	**0.78**	**0.61**	**0.72**	**0.84**	**0.82**
kappa	**0.70**	**0.43**	**0.64**	**0.79**	**0.76**

As shown in Figure 13, we compared the results of three datasets using different algorithms.

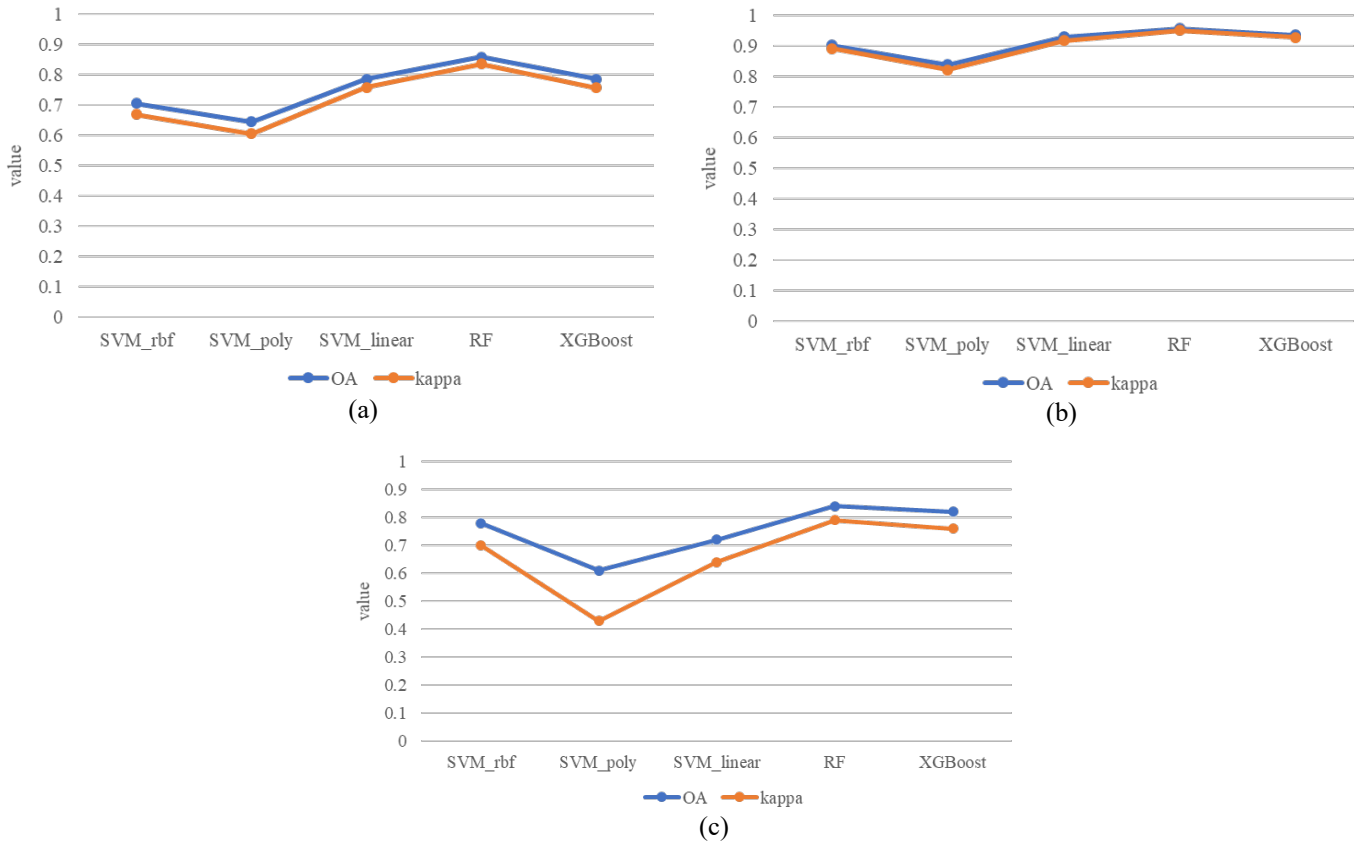


Figure 13. Result comparison. (a) Indian Pines, (b) Salinas, and (c) EO1-Hyperion

5. CONCLUSIONS

The results show that the three hyperspectral datasets, after MNF transformation, can effectively undergo dimensionality reduction while maintaining image quality ranking. This study successfully combines different algorithms to find an effective classification process applicable to both airborne and satellite hyperspectral images. The RF classifier consistently outperforms others, demonstrating reduced category confusion. The best results show an overall accuracy of 0.86 for Indian Pines and 0.96 for Salinas airborne images. The kappa values are 0.84 and 0.95, respectively. The best overall accuracy of the satellite image is 0.84 and Kappa is 0.79. Moreover, the SVM classifier using the rbf and linear kernel functions, as well as the XGBoost classifier, all demonstrate similar and satisfactory classification results. The SVM classifier using the poly kernel function, which is suitable for nonlinear data, tends to yield weaker classification results compared to the other classifiers. For all images, samples with similar characteristics among classes are prone to misclassification, leading to false negatives. Fragmented regions also appear in the classification results. Future research will utilize morphological techniques to address and improve the issue of fragmentation in classification. Furthermore, future research will incorporate more extensive open data for analysis and ensure the reliability of experimental results. Additionally, deep learning will be integrated into this study to align with emerging trends.

ACKNOWLEDGEMENTS

This research was supported by the National Science and Technology Council (NSTC) under project number: 112-2813-C-008-001-M.

REFERENCES

Breiman, L. (2001). Random forests. Machine learning, 45, 5-32.

Benediktsson, J. A., Palmason, J. A., & Sveinsson, J. R. (2005). Classification of hyperspectral data from urban areas based on extended morphological profiles. IEEE Transactions on Geoscience and Remote Sensing, 43(3), 480-491.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine learning, 20(3), 273-297.

Camps-Valls, G., Tuia, D., Bruzzone, L., & Benediktsson, J. A. (2013). Advances in hyperspectral image classification: Earth monitoring with statistical learning methods. IEEE signal processing magazine, 31(1), 45-54.

Chen, Y., Lin, Z., Zhao, X., Wang, G., & Gu, Y. (2014). Deep learning-based classification of hyperspectral data. IEEE Journal of Selected topics in applied earth observations and remote sensing, 7(6), 2094-2107.

Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining (pp. 785-794).

Green, A. A., Berman, M., Switzer, P., & Craig, M. D. (1988). A transformation for ordering multispectral data in terms of image quality with implications for noise removal. IEEE Transactions on geoscience and remote sensing, 26(1), 65-74.

Ghamisi, P., Plaza, J., Chen, Y., Li, J., & Plaza, A. J. (2017). Advanced spectral classifiers for hyperspectral images: A review. IEEE Geoscience and Remote Sensing Magazine, 5(1), 8-32.

Gewali, U. B., Monteiro, S. T., & Saber, E. (2018). Machine learning based hyperspectral image analysis: a survey. arXiv preprint arXiv:1802.08701.

Hsu, C. W., Chang, C. C., & Lin, C. J. (2003). A practical guide to support vector classification.

Heydarian, M., Doyle, T. E., & Samavi, R. (2022). MLCM: multi-label confusion matrix. IEEE Access, 10, 19083-19095.

Lu, D., & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. International journal of Remote sensing, 28(5), 823-870.

Melgani, F., & Bruzzone, L. (2004). Classification of hyperspectral remote sensing images with support vector machines. IEEE Transactions on geoscience and remote sensing, 42(8), 1778-1790.

Mundt, J. T., Glenn, N. F., Weber, K. T., Prather, T. S., Lass, L. W., & Pettingill, J. (2005). Discrimination of hoary cress and determination of its detection limits via hyperspectral image processing and accuracy assessment techniques. Remote Sensing of Environment, 96(3-4), 509-517.

Mahesh, B. (2020). Machine learning algorithms-a review. International Journal of Science and Research (IJSR).[Internet], 9, 381-386.

Patle, A., & Chouhan, D. S. (2013, January). SVM kernel functions for classification. In 2013 International Conference on Advances in Technology and Engineering (ICATE) (pp. 1-9). IEEE.

Shaw, G. A., & Burke, H. K. (2003). Spectral imaging for remote sensing. Lincoln laboratory journal, 14(1), 3-28.

Tsai, F., Chang, C. K., & Liu, G. R. (2006, October). Texture analysis for three dimensional remote sensing data by 3D GLCM. In Proc. Asian Conf. Remote Sens.(ACRS) (pp. 430-435).

Thenkabail, P. S., & Lyon, J. G. (Eds.). (2016). Hyperspectral remote sensing of vegetation. CRC press.

Uddin, M. P., Mamun, M. A., & Hossain, M. A. (2021). PCA-based feature reduction for hyperspectral remote sensing image classification. IETE Technical Review, 38(4), 377-396.