



Comparison of Supervised Classification Algorithms Using a Hyperspectral Image for Land Use/Land Cover Classification [†]

Sonia Sharma Banjade ^{*}, Nitant Rai and Bipana Subedi

Department of Forest Resources and Environmental Conservation, Virginia Polytechnic Institute and State University, Blacksburg, VA 24060, USA; nitant3@vt.edu (N.R.); sbipana98@vt.edu (B.S.)

^{*} Correspondence: soniasharma07@vt.edu

[†] Presented at the 5th International Electronic Conference on Remote Sensing, 7–21 November 2023; Available online: <https://ecrs2023.sciforum.net/>.

Abstract: Hyperspectral imaging is becoming popular in land use/land cover classification because of its ability to capture detailed information through higher spatial resolution and contiguous spectral bands. Using the hyperspectral image from G-LiHT (Goddard's LiDAR, Hyperspectral, and Thermal) Airborne Imager covering a study area in Tennessee, Knoxville, we compared the performance of Spectral Angle Mapper (SAM), Spectral Information Divergence (SID), and Support Vector Machine (SVM) for land use/land cover classification. We used a confusion matrix for the accuracy assessment of the classifiers. Among the three classifiers, SVM showed the highest accuracy with 92.03%. Our results also show that some classes, such as water and forests, are consistently distinguishable across all classification methods, while others, such as built-up areas, vary depending on the technique used.

Keywords: supervised classification; hyperspectral image; land use; spatial resolution; classifiers

1. Introduction

The detailed mapping of land use/land cover change has advanced in recent decades with emerging satellite data based on multispectral or hyperspectral sensors. Since the 1960s, remote sensing data has been used in land cover mapping. The detailed mapping assists in analyzing changes over time in various land use/land cover classes and assessing risk at various scales [1]. This information plays a vital role in preserving ecologically sensitive areas and solving environmental issues. With the increasing urbanization and land degradation, the significance of land use classification has increased [2].

Multispectral images are extensively utilized in image classification, particularly for land use and land change classification [3]. However, with the rapid advancement in technology, hyperspectral images have also been used recently. The hyperspectral images provide spectral data for each pixel in numerous contiguous spectral bands often covering a wide range of wavelengths [4]. Also, these images have high spectral resolution, making them able to capture detailed information about the spectral characteristics of the observed objects or surface.

The development of reliable image classification depends on the performance of classification algorithms. Specifically dealing with hyperspectral images, the high dimensionality and spectral mixing are the major challenges [5], which can significantly impact the accuracy of classification results. Additionally, an inadequate number of ground-truth data, as well as potential redundancy in hyperspectral images, add complexity to the classification process [6]. Therefore, this study aims to investigate the robustness of classification algorithms in handling spectral unmixing and limited ground-truth information. It compares various image classification algorithms using a hyperspectral image.

Image classification involves a process where each individual pixel within the image is categorized into discrete land use classes [7]. The two most often used methods of



Citation: Sharma Banjade, S.; Rai, N.; Subedi, B. Comparison of Supervised Classification Algorithms Using a Hyperspectral Image for Land Use/Land Cover Classification. *Environ. Sci. Proc.* **2024**, *29*, 59. <https://doi.org/10.3390/ECRS2023-16702>

Academic Editor: Luca Lelli

Published: 11 December 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

classification are supervised classification and unsupervised classification. In supervised classification, the analyst knows about the labels of classes or has the training dataset [8] before applying the classification of the image.

Hence, this study aims to perform and compare supervised classifiers (SVM, SAM, and SID) for land use land cover classification using a hyperspectral image.

2. Materials and Methods

2.1. Data

A hyperspectral image from G-LiHT (Goddard's LiDAR, Hyperspectral, and Thermal) was used for this research. This data underwent processing to generate standardized data products, including 1 m at-sensor reflectance hyperspectral imagery. The flight was conducted on 7 May 2015, in Stanton of Knoxville, Tennessee. The image was acquired as UTK_7May2015_Stanton and was downloaded from the NASA G-LiHT website accessed on 3 September 2023 (URL: <https://glihtdata.gsfc.nasa.gov/>) with 119 spectral bands between 418 and 918 nm. The hyperspectral imaging spectrometer model was Hyperspec model 1002A-00451; Headwall Photonics [9].

2.2. Data Pre-Processing, Training, and Testing Dataset

We preprocessed the hyperspectral image to reduce redundancy and noise. We used Minimum Noise Fraction (MNF) transformation, a widely used technique that serves to decorrelate spectral bands and reduce noise, effectively isolating the signal from undesirable variations [10]. Following this transformation, an eigenvalue analysis was conducted to determine the importance of each MNF component. MNF components with higher eigenvalues were prioritized as they capture more detailed information about the land cover classes being analyzed. We subsetted the original image with 30 bands, only reducing the band with low eigenvalues.

We used the spectral library (Figure 1) created through an in-field survey using a spectrometer as the ground-truth data or training dataset for image classification. We used a random sampling method to choose our samples in the field. A total of 60 samples of each land cover type (vegetation, grassland, built-up, bare soil, water) were recorded using a spectrometer. Then, 40 samples were chosen for the training dataset randomly. The remaining 20 samples were chosen for testing.

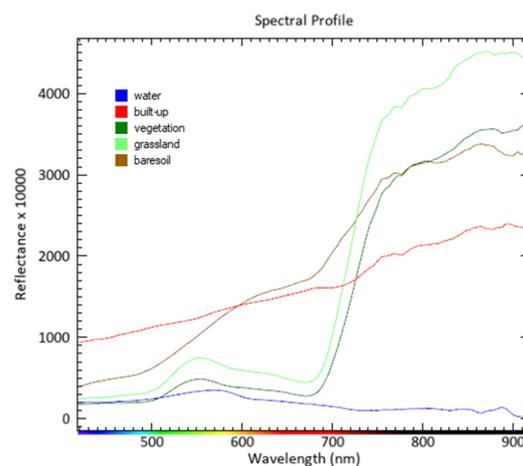


Figure 1. Spectral profile of water, built-up, vegetation, grassland, and bare soil.

2.3. Image Classification

2.3.1. Spectral Angle Mapper (SAM)

SAM is particularly valuable for identifying and characterizing materials or objects within a scene based on their spectral signatures. SAM operates on the principle that the similarity between two spectra can be quantified by measuring the angle between them in a high-dimensional space, where each dimension corresponds to a spectral band or

wavelength [11]. For this classification method, we tried various values of maximum angle radians, and the best result was obtained when a value of 0.3 was used.

2.3.2. Spectral Information Divergence (SID)

SID is commonly used in hyperspectral image processing for tasks like anomaly detection, target detection, and classification. It helps identify areas or objects in an image that deviate significantly from the expected spectral distribution, which can be useful in image classification [12]. For the SID algorithm, we used the maximum divergence threshold value of 0.5 to obtain the best result.

2.3.3. Support Vector Machine (SVM)

SVM is a supervised machine learning algorithm used for classification and is effective in high-dimensional images [13]. We used a radial-based kernel function using a gamma value of 0.009 for SVM.

2.4. Accuracy Assessment

The accuracy assessment is a crucial step in land use/land cover classification for the validation of the classified image. We used a confusion matrix, which summarizes the class labels against the predicted labels to evaluate the performance of supervised classification algorithms. The total accuracy was calculated as

$$\text{Overall accuracy} = (\text{Number of correctly classified pixels} \div \text{Total number of pixels}) \times 100$$

3. Results

3.1. Image Classification

The image classification results are shown in Figure 2. SAM appears to give no data value surrounding the water bodies and built-up areas. SID can remove the no data value from the image. It can be seen that the forest and built-up areas are classified, although there seems to be some noise over the water bodies. Meanwhile, SVM performed well in detecting the land cover types, removing the no data value over the water bodies and surrounding built-up areas.

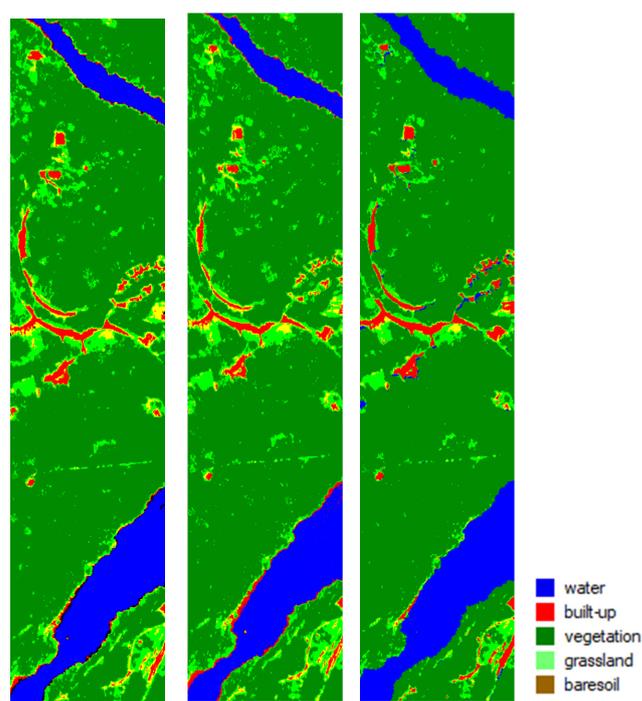


Figure 2. Supervised classification results using SAM, SID, and SVM (left to right) showing five (water, built-up, vegetation, grassland, bare soil) classes and unclassified labels.

3.2. Accuracy Assessment

The confusion matrix table for each of the classification algorithms (SVM, SID, and SAM) is represented in Tables 1–3.

SVM achieved an exceptional accuracy of 92.03%, SID had 89.60%, and SAM had 91.23%. The confusion matrices provide further insights into the classification performance, detailing the distribution of true positives (correctly classified pixels), true negatives, false positives, and false negatives for each classifier. The high values in the diagonal of the confusion matrices indicate strong agreement between the predicted and actual class labels.

Table 1. Accuracy assessment of SVM.

Class	Bare Soil	Grassland	Water	Built-Up	Vegetation	Total
Unclassified	0	0	0	0	0	0
Baresoil	77.62	2.98	0	2.46	0	1.31
Grassland	14.69	94.47	0	7.91	1.09	3.71
Water	0	0	99.25	4.61	0	23.68
Built-up	6.99	0	0.75	41.55	0.08	5.06
Vegetation	0.7	2.55	0	43.47	98.83	66.24
Total	100	100	100	100	100	100

Overall accuracy = 92.03%.

Table 2. Accuracy assessment of SID.

Class	Bare Soil	Grassland	Water	Built-Up	Vegetation	Total
Unclassified	0	0	0	0	0	0
Baresoil	79.02	3.4	0.3	10.91	0.03	2.39
Grassland	20.98	96.6	0	13.52	2.98	5.64
Water	0	0	96.85	0	0	22.6
Built-up	0	0	2.85	34.79	0.03	4.65
Vegetation	0	0	0	40.78	96.96	64.72
Total	100	100	100	100	100	100

Overall accuracy = 89.60%.

Table 3. Accuracy assessment of SAM.

Class	Bare Soil	Grassland	Water	Built-Up	Vegetation	Total
Unclassified	0	0	1.43	0	0	0.33
Baresoil	92.31	2.98	0.23	8.76	0.08	2.32
Grassland	7.69	97.02	0	10.6	2.08	4.59
Water	0	0	97.52	0	0	22.75
Built-up	0	0	0.83	41.55	0.03	4.95
Vegetation	0	0	0	39.09	97.81	65.05
Total	100	100	100	100	100	100

Overall accuracy = 91.23%.

4. Discussion

The accuracy assessment results of the land use/land cover classification, employing SVM, SAM, and SID classifiers, reveal promising outcomes for the hyperspectral image. The achieved accuracies for all three classes indicate they performed very well for detailed classification, while SVM stands as a top-performing classifier with the highest accuracy of 92.03% among the three of them. The result was consistent with a comparative study on the effectiveness of image classification algorithms, including SVM, SAM, and SID, conducted by [14,15], which also concluded that SVM performs better than other methods. Despite showing the highest accuracy, SVM is computationally intensive, especially with large datasets. Though SAM had negligible differences with SVM, SAM has proven to be best in capturing spectral similarity based on spectral angles [16].

The notable outcome of this research is the consistency in the distinguishability of forest and water across all employed classification schemes. This implies the spectral

signatures of these classes are distinct and easily discernible by the selected classifiers. In contrast, variability is seen in built-up areas. Also, the challenges seen in this research are shadows, particularly tall structures, and trees. In many cases, these shadows create dark pixels within the image and can be incorrectly classified as water bodies.

5. Conclusions

Following an analysis of different supervised image classifiers, we discovered that SVM outperforms other classifiers in accurately identifying land cover/land use classes and is also effective at handling high-dimensional data. Following SVM, SAM can also serve as a suitable method for detecting land cover/land use classes, as there was negligible difference between SAM and SVM. The detection in built-up areas and water bodies is slightly mislabeled as a shadow by SID, whereas the SVM demonstrated its effectiveness in handling such scenarios. Hence, these three supervised classifiers were demonstrated to be effective in classifying remotely sensed data.

Author Contributions: Conceptualization, S.S.B.; data curation, S.S.B. and N.R.; formal analysis, S.S.B., N.R. and B.S.; methodology, S.S.B.; writing—original draft, S.S.B. and B.S.; writing—review and editing, S.S.B. and N.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data will be available on request.

Acknowledgments: We would like to thank Santosh Rijal for having the lab sessions in the “Remote Sensing” course where we were able to collect field spectrometer data.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Petropoulos, G.P.; Arvanitis, K.; Sigrimis, N. Hyperion hyperspectral imagery analysis combined with machine learning classifiers for land use/cover mapping. *Expert Syst. Appl.* **2012**, *39*, 3800–3809. [[CrossRef](#)]
2. Shafri, H.Z.M.; Suhaili, A.; Mansor, S. The performance of maximum likelihood, spectral angle mapper, neural network and decision tree classifiers in hyperspectral image analysis. *J. Comput. Sci.* **2007**, *3*, 419–423. [[CrossRef](#)]
3. Joshi, S.; Rai, N.; Sharma, R.; Baral, N. Land Use/Land Cover (LULC) Change in Suburb of Central Himalayas: A Study from Chandragiri, Kathmandu. *J. For. Environ. Sci.* **2021**, *37*, 44–51. [[CrossRef](#)]
4. Raczko, E.; Zagajewski, B. Comparison of support vector machine, random forest and neural network classifiers for tree species classification on airborne hyperspectral APEX images. *EuJRS* **2017**, *50*, 144–154. [[CrossRef](#)]
5. Deilmai, B.R.; Ahmad, B.B.; Zabihi, H. Comparison of two Classification methods (MLC and SVM) to extract land use and land cover in Johor Malaysia. *IOP Conf. Ser. Earth Environ. Sci.* **2014**, *20*, 012052. [[CrossRef](#)]
6. Ballanti, L.; Blesius, L.; Hines, E.; Kruse, B. Tree species classification using hyperspectral imagery: A comparison of two classifiers. *Remote Sens.* **2016**, *8*, 445. [[CrossRef](#)]
7. Barnsley, M.J.; Barr, S.L. Inferring urban land use from satellite sensor images using kernel-based spatial reclassification. *Photogramm. Eng. Remote Sens.* **1996**, *62*, 949–958.
8. Bagirov, A.M.; Rubinov, A.M.; Soukhoroukova, N.V.; Yearwood, J. Unsupervised and supervised data classification via nonsmooth and global optimization. *TOP* **2003**, *11*, 1–75. [[CrossRef](#)]
9. Cook, B.D.; Corp, L.A.; Nelson, R.F.; Middleton, E.M.; Morton, D.C.; McCorkel, J.T.; Masek, J.G.; Ranson, K.J.; Ly, V.; Montesano, P.M. NASA Goddard’s Lidar, Hyperspectral and Thermal (G-LiHT) airborne imager. *Remote Sens.* **2013**, *5*, 4045–4066. [[CrossRef](#)]
10. Luo, G.; Chen, G.; Tian, L.; Qin, K.; Qian, S.-E. Minimum Noise Fraction versus Principal Component Analysis as a Preprocessing Step for Hyperspectral Imagery Denoising. *Can. J. Remote Sens.* **2016**, *42*, 106–116. [[CrossRef](#)]
11. De Carvalho, O.A.; Meneses, P.R. Spectral Correlation Mapper (SCM): An Improvement on the Spectral Angle Mapper (SAM). In *Summaries of the 9th JPL Airborne Earth Science Workshop*; JPL Publication: Pasadena, CA, USA, 2000.
12. Chang, C.I. Spectral information divergence for hyperspectral image analysis. In *Proceedings of the IEEE 1999 International Geoscience and Remote Sensing Symposium, IGARSS’99 (Cat. No.99CH36293)*, Hamburg, Germany, 28 June–2 July 1999; pp. 509–511.
13. Noble, W.S. What is a support vector machine? *Nat. Biotechnol.* **2006**, *24*, 1565–1567. [[CrossRef](#)] [[PubMed](#)]

14. Shakya, A.K.; Ramola, A.; Kandwal, A.; Prakash, R. Comparison of supervised classification techniques with alos palsar sensor for roorkee region of uttarakhand, india. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **2018**, *XLII-5*, 693–701. [[CrossRef](#)]
15. Wang, K.; Cheng, L.; Yong, B. Spectral-Similarity-Based Kernel of SVM for Hyperspectral Image Classification. *Remote Sens.* **2020**, *12*, 2154. [[CrossRef](#)]
16. Dennison, P.E.; Halligan, K.Q.; Roberts, D.A. A comparison of error metrics and constraints for multiple endmember spectral mixture analysis and spectral angle mapper. *Remote Sens. Environ.* **2004**, *93*, 359–367. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.