Classification of EO-1 Hyperion Data using Supervised Minimum Distance Algorithm and Spectral Angle Mapper

Vaibhav A. Didore [1], Renuka S. Vaidya [1], Dhananjay B. Nalawade [1], Karbhari V. Kale[1]

Department of Computer Science and Information Technology, Dr Babasaheb Ambedkar Marathwada University, Aurangabad, Maharashtra, India, 431001.

Abstract- In remote sensing, hyperspectral images contain most of the information that can be processed to identify objects. Satellite imagery classification plays a key role in various object recognition systems. Satellite imagery classification requires a selection of suitable classification techniques based on the user's requirement. In this paper, two supervised classification technique namely minimum distance classification (MDC) and Spectral angle mapper (SAM) has been used for the analysis of the remotely sensed data. In this experimental analysis, EO-1 Hyperion sensor data downloaded from USGS earth explorer has been used. The MDC technique estimates the Euclidean distance of each pixel to the mean vector of the class and assigns each pixel to the closest class. Mean vector and covariance metrics are the key component of MDC that can be retrieved from training data. The classification accuracy has been achieved for the MDC in terms of overall accuracy 79.31% and overall kappa accuracy is 73.99%. similarly overall accuracy calculated for SAM is 84.12% and kappa accuracy is 79.54%. From the experimental analysis it is evident that Minimum Distance classification technique is useful for small size data and it takes less time for classification, howevermore classification accuracy is achieved by using SAM.

Keywords- Image Classification; Minimum Distance; Spectral Angle Mapper; USGS EO-1; Hyperion data.

INTRODUCTION

A remote Sensing Satellite is the human's third eye that presents in the sky. With that artificial eye, we can closely observe ecology and other material present on theearth's surface. Remote sensing provides the ability to observe an object without making contact with the object[1]. Finding meaningful data from hyperspectral images is an art and science, which can be possible from classification. Image classification is considered an important aspect of remote sensing, image analysis, and pattern recognition [2] In remote sensing, the classification of the hyperspectral image is the process ofgrouping pixels in a finite set of individual classes based on their data values. The pixel is assigned to a particular class if it satisfies a certain set of rules to fit in a particular class [3]. The human eye and brain can easily recognize spectral and spatially enhanced images of certain colors and patterns. Merely for a computer system, pattern recognition is a mathematical processand pixels are classified and grouped in categories according to mathematical criteria.

These calculations divide into supervised classification and unsupervised classification. Unsupervised classification is a more computer automated process and it grouped the similar spectral characteristic pixel into the clusters. Unsupervised classification is used whendata are very less to be classified an image. The main advantage of this technique is time-saving because it doesn't require any analyst to the fed information before classification, but after classification, labeling should be done by the analyst [4]. Supervised classification is the process of grouping pixels into classes based on training data. Training data are groups of pixels that represent areas for which the class of information is already known (land cover, geological type, etc.). Supervised classification is used when a few classes are required for analysis. Some prior knowledge of pixels is also required that represent classes that you want to extract from the image after that analyst creates an appropriate signature form image for classification [5].

LITERATURE REVIEW

EL-Rahman et. al. used Spectral Angle Mapper (SAM) and Parallelepiped algorithms implemented by a software tool used to analyses and process geospatial images that is an Environment of Visualizing Images (ENVI). They are applied to Al-Kharj, Saudi Arabia, as a study area. The author achieved overall accuracy for SAM classification was 66.67%, and 33.33% for Parallelepiped classification. Therefore, the author finds that the SAM algorithm has provided a better imageclassification [6].

Kuching S. et. al used Spectral Angle Mapper (SAM), Maximum likelihood classification (MLC), Artificial Neural Network (ANN), and Decision tree (DT) Supervised classifier for classification of hyperspectral data of Forest research institute (FRI) in Kepong, Selangor, Malaysia. The author acquired data using AISA airborne imaging spectrometer onboard the NOMAD GAF-27 Aircraft. Author achievedclassification accuracy for MLC, SAM, ANN, & DT is 85.56%, 48.83%, 83.61%, & 50.67% respectively. The author suggests that MLC is the best classifier [7].

Liu X. et. al used SAM, and K-SAM classification algorithm to classify Beijing-1 micro-satellite multispectral images with 32m spatial resolution. The author achieved an overall classification accuracy of 88.17% for SAM and 89.83% for K-SAM. The author suggests that the polynomial K-SAM algorithm issuperior to the SAM algorithm in remote sensing image classification [8].

Sisodia P. S. et. al used Minimum distance (MD), Parallelepiped (PP), Mahalanobis Distance (MhD), and Maximum likelihood classification to classify different Landsat images such as Landsat Multispectral Scanner (MSS), Thematic Mapper (TM), & Enhanced Thematic mapper + (ETM+) images of Jaipur District, Rajasthan. Author achieved overall accuracy for MhD, MLC, MD, and PP is 85.46%, 83.21%, 86.52%, and 83.11%

respectively. The author suggests that Minimum Distance classification produced better accuracy for Landsat MSS images, whereas MLC produce better accuracy with Landsat TM & ETM+ images [9].

Niknejad M. et.al used Support Vector Machine (SVM), Maximum Likelihood Classification (MLC), Minimum Distance (MD) and Parallelepiped (PP) for Landsat ETM+ data & achieve highest 88.65% accuracy for SVM [10].

Classification Technique	Dataset	Best Classifier	Author
SAM, & PP	EO-1 Hyperion	SAM	EL-Rahman et al[6]
MD, MLC, MhD, & PP	Landsat MSS, Landsat TM, Landsat ETM+	MD & MLC	Sisodia P.S. et.al[9]
SVM, MLC, MD, & PP	Landsat ETM+	SVM	Niknejad M. et.al[10]
SAM, & K-SAM	Beijing-1 micro-satellite	K-SAM	Liu X.et. al[8]
SAM, MLC, ANN & DT	AISA imaging spectrometer	MLC	Kuching S. et. al[7]
PP, MD, MLC,ANN, & SVM	Quick bird	SVM	Doma M. I.et al.[11]

Table 1: Review of literature

MINIMUM DISTANCE CLASSIFIER (MDC) is the most commonly used supervised classification and used in a variety of applications. In MDC, training data is used only to determine class means; classification is then performed by placing a pixel in the class of the nearest mean [12] [13]. The minimum distance algorithm is also attractive since it is a faster technique than maximum likelihood classification [14].

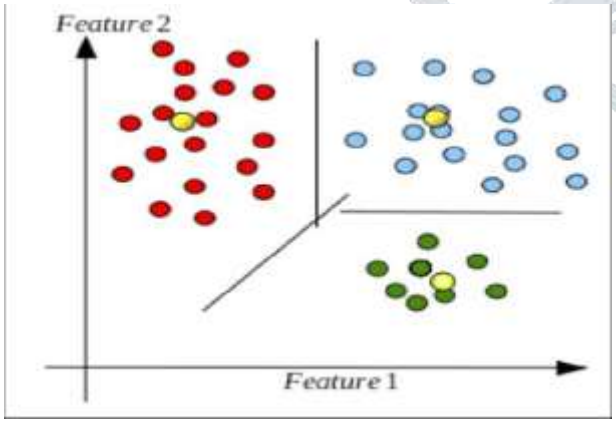

Fig. 1. BASIC CONSEPT OF MDC [14].

The flow of MDC is [15]:
1) Select training samples.
2) Calculate class training sample mean.
3) Calculate the distance from the test sample to the class mean vector.
4) The minimum distance is the category to which the best sample belongs.

In general, MD classifies each pixel in the nearest class (regardless of its distance). The distance between the "image pixel" and the "layer mean vector" can be calculated in different ways (e.g. Euclidean distance, Mahalanobis distance, etc.) In the MDC classifier, to find an average of the pixels of the training sets in the n-direction space, all pixels in the image are sorted by the mean of the layer to which they are closest. This is one of the simple algorithms, using the training set of a class represented as a centre point based on information about the mean of all pixels of the sample class.

Remote sensing data are capable to understand the city morphology like shape, density, patch, patch density, and spread of built-up area [16]. Any unknown pixel will definitely be assigned to one of any classes, there will be no unclassified pixel [17].

In the Minimum Distance Classification:
- Each class is represented by its own mean vector.
- Learning is done using objects (pixels) of a known class.
- Calculate the mean of feature vectors of an object in a class.
- New objects are ordered by determining the nearest mean vector (Image Analysis Center).

This classification requires mean vectors for each class in each band k from the training data. The Euclidean distance is computed for all pixels with all signature means.

$$D = \sqrt{(BV_{ijk} - \mu_{ck})^2 + (BV_{ijl} - \mu_{cl})^2}$$

Where,
μ_{ck} and μ_{cl} represent the mean vectors for class c measured in bands k and l.

SPECTRAL ANGLE MAPPER (SAM) is an algorithm is based on an ideal assumption that a single pixel of remote sensing images represents one certain ground cover material, and can be uniquely assigned to only one ground cover class. The SAM is simply based on measuring the spectral similarity between two spectra. The spectral similarity can be obtained by considering each spectrum as a vector in q-dimensional space [18]. where q is the number of bands. The SAM algorithm determines the spectral similarity between two spectra by calculating the angle between the two spectra, treating them as vectors in a space with dimensionality equal to the number of bands [19]. The SAM algorithm, in ENVI, assigns angles to the output channels, then each pixel is assigned to the layer defined by the reference spectrum. The class assigned to each pixel is stored in the output channel [20]. The main limitation of this method is the problem of spectral mixing, which often occurs in medium resolution images.
In this paper, supervised minimum distance classification (MDC) and spectral angle mapper classification (SAM) is used to classify remotely sensed images for classification accuracy and identify different ground covers like water body, vegetation, buildup area, and barren land etc.

STUDY AREA AND DATA USED

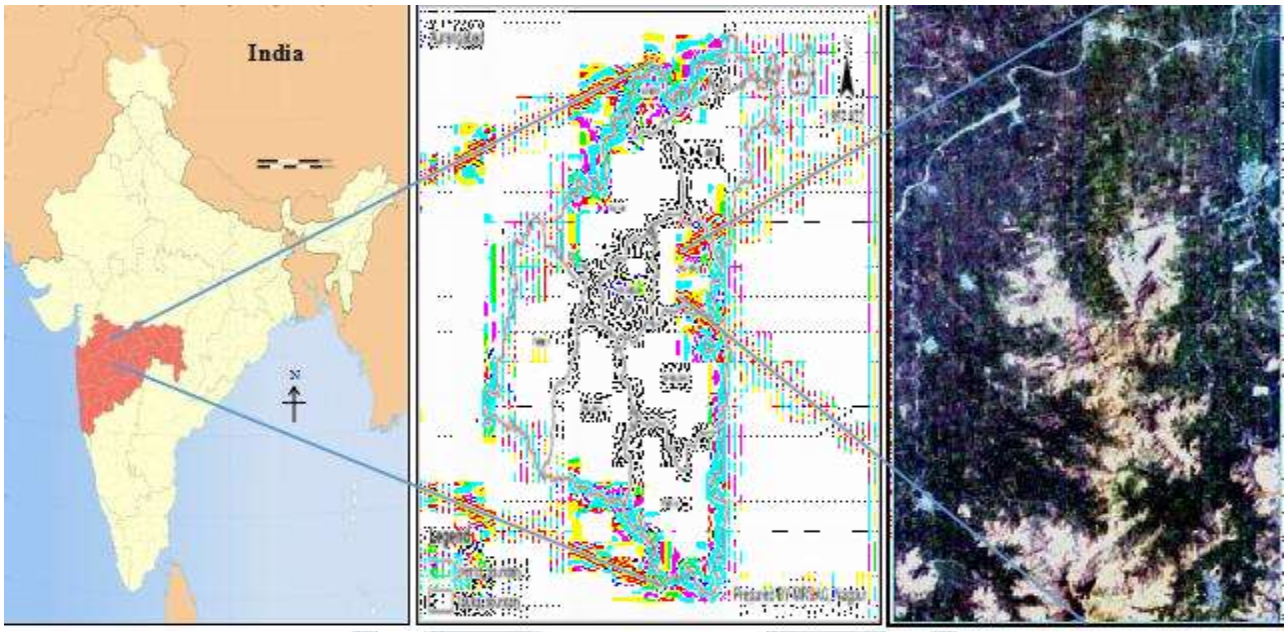

Figure 1: LOCATION OF STUDY AREA ON MAP OF INDIA.

For the classification of the remotely sensed image, Phulambri Tahsil of Aurangabad district of Maharashtra has been chosen. The study area extends on latitude 20° 1' 30" N to 20° 8' 0" N and longitude 75° 20' 20" E to 75° 24' 30"E. Aurangabad district is located in the central of Maharashtra, the world-famous heritage site is situated in Aurangabad districts like Ajanta caves andEllora caves. Out of 10,100 Sq.km area, 141.1 Sq.km and 9959 Sq.km area are categories under urban and rural [21]. The Average rainfall of Aurangabad district is 734 mm and the minimum temperature is 5.6 °C and maximum temperature is 45.9 °C. Aurangabad have near about 135.75Sq.km forest area and gives 9% contribution in total forest area of whole Maharashtra [22].

Remote sensing data used in this study has been acquired from the public domain United States Geological Survey (USGS) [23]. USGS is a scientific agency governed by the United States. Scientists from the U.S. Geological Survey study the landscape, natural resources, and natural disasters that threaten the United States. The work of the organization covers the fields of biology, geography, geology and hydrology. The US Geological Survey is a fact-finding research organization with no regulatory responsibility. For image accuracy assessment 290 for Minimum distance and 315 for Spectral angle mapper random sample points were collected.

The EO-1 Hyperion image of Phulambri Tahsil was acquired from the public domain U.S. Geological Survey. EO-1 images were rectified on Geographic (lat/long) projection. After capturing the images, we applied an image enhancement with histogram alignment for better interpretation and accuracy. The study area was separated from the image using sub-setting the area of interest. Supervised Minimum Distance Classification (MDC) and Spectra Angle Mapper (SAM) algorithm have been used for classification. Five signature classes were selected for classification like water bodies, vegetation, buildup area, barren land, hilly area. For the buildup area, we consider residential buildings, highways, industries, commercial buildings etc. Vegetation represents the agricultural land, forest, reserve centuries, etc. Barren land represents the non-vegetation area, and the hilly area represents land slop, rocks, deep valleys and forests present in the hilly region etc. Random sampling is used to evaluate accuracy. Random location is used for the ground truth data collection. For doubt about classification, we have generated a confusion matrix. A field survey using google earth of the study area has been done for classification accuracy. We have calculated overall accuracy, producer accuracy, user accuracy and overall kappa for the classified images. EO-1 Hyperion image false color composite, the classified image of Minimum Distance, and the classified image of Spectral Angle Mapper have shown by fig. 5, 6, and 7 respectively. Environment for Visualizing Image (ENVI 5.2) software was used for experimental work [24].

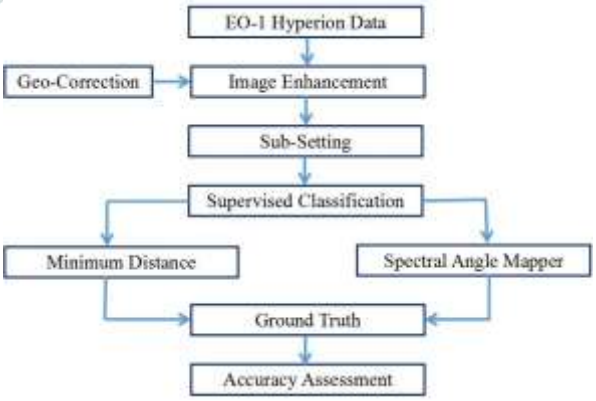

FIG. 3: FLOW CHART OF METHODO

RESULTS

FIG. 4: FALSE COLOUR COMPOSITE

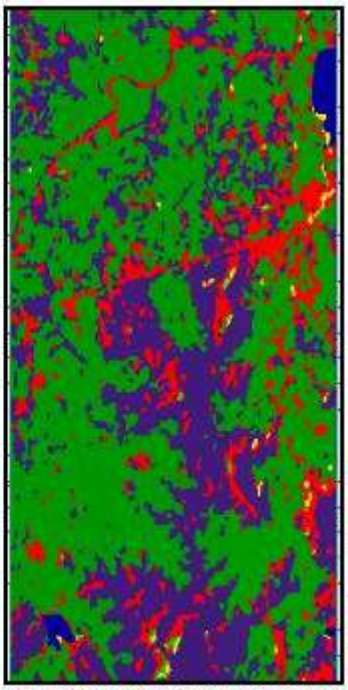

FIG. 5: MINIMUM DISTANCE ALGORITHM

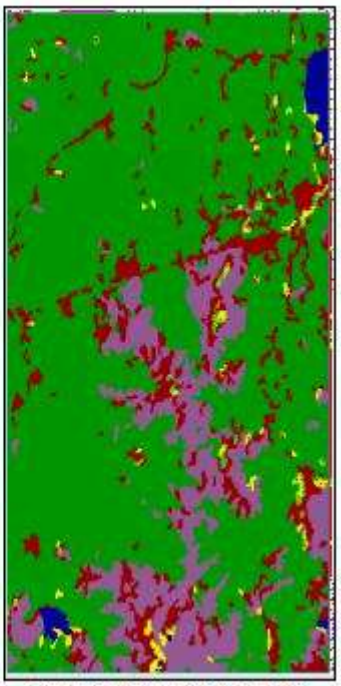

FIG. 6: SPECTRAL ANGLE MAPPER

Ground Truthing and validation was made out for classification using a field survey of 290 locations for MD and 315 locations for SAM that cover the entire work region, Confusion metrics were generated where we had doubts about classification as shown in Table 2 and Table 3. The classified image of study area for MD (Fig. 5) was having an Overall Accuracy (OA) of 79.31%, Producer Accuracy (PA) 80.83%, User Accuracy (UA) 78.83% and Overall Kappa coefficient (K) of 73.99%. and for Spectral Angle Mapper (Fig. 6) was having an Overall Accuracy (OA) of 84.12 %, Producer Accuracy (PA) 80.16 %, User Accuracy (UA) 78.48 % and Overall Kappa coefficient (K) of 79.54 %.

$$O.A. = \frac{\text{Total no. of correctly classified pixel}}{\text{Total no. of pixel used for accuracy assessment}}$$

$$\text{Overall Accuracy} = \frac{23}{29} = 79.31\% \text{ for MD}$$

Error matrix of ENVI classification with MD						
Classes\ Reference Classes	Water Body	Vegetation Area	Buildup Area	Barren Land	Hilly Area	Total
Water Body	40	0	0	0	10	50
Vegetation Area	0	70	0	10	0	80
Buildup Area	0	0	50	0	0	50
Barren Land	0	0	10	30	10	50
Hilly Area	0	0	20	0	40	60
Total	40	70	80	40	60	290
User Accuracy (UA)	80%	87.5%	100%	60%	66.33%	78.83%
Producer Accuracy (PA)	100%	100%	62.5%	75%	66.66%	80.83%
Overall Accuracy (OA) of MD= 79.31 %, Kappa= 73.99 %						

Table 2: Confusion Matrix of Study Area for MD.

Error matrix of ENVI classification with SAM						
Classes\ Reference Classes	Water Body	Vegetation Area	Buildup Area	Barren Land	Hilly Area	Total
Water Body	71	0	4	0	0	75
Vegetation Area	0	95	0	3	0	98
Buildup Area	0	0	39	3	5	47
Barren Land	0	2	10	45	3	60
Hilly Area	0	0	15	5	15	35
Total	71	97	68	56	23	315
User Accuracy (UA)	94.66 %	96.93%	82.97%	75.00 %	42.85%	78.48%
Producer Accuracy (PA)	100%	97.93%	57.35%	80.35 %	65.21%	80.16%
Overall Accuracy (OA) of SAM = 84.12 %, Kappa= 79.54 %						

Table 3: Confusion Matrix of Study Area for MD.

EO-1 Data	Users Accuracy	Producers Accuracy	Overall Accuracy	Kappa Coefficient
MD	78.83%	80.83%	**79.31%**	73.99%
SAM	78.48%	80.16%	**84.12%**	79.54%

Table 4: Accuracy Result of EO-1 Data.

CONCLUSION

In this paper, supervised minimum distance classification (MDC) and spectral angle mapper (SAM) technique are compared by using EO-1 Hyperion remote sensing data. From the experimental analysis we have found that MDC perform well on small dataset, still there is a ambiguityin classification. Whereas this SAM technique classifies vegetation and water-body more correctly than MDC, however here the mountain area shows low classification accuracy due to shadows of plant and rock.

FUTURE WORK

In the future, we will use Maximum likelihood classification, Decision Tree, Parallelepiped, and SVM algorithm as a classifier to classify remotely sensed images analysis These image analyses can be used for land use/land cover mapping, forest and precision agricultural.

REFERENCES

[1] Kale, K. V., Solankar, M. M., & Nalawade, D. B. (2019). Hyperspectral Endmember Extraction Techniques. In Processing and Analysis of Hyperspectral Data. Intech Open.

[2] Abbas, K., & Rydh, M. (2012). Satellite image classification and segmentation by using JSEG segmentation algorithm. International Journal of Image, Graphics and Signal Processing, 4(10), 48.

[3] Kale, K.V., Solankar, M.M., Nalawade, D.B. et al. A Research Review on Hyperspectral Data Processing and Analysis Algorithms. Proc. Natl. Acad. Sci., India, Sect. A Phys. Sci. 87, 541–555(2017). https://doi.org/10.1007/s40010-017-0433-y

[4] Salah, M. (2017). A survey of modern classification techniques in remote sensing for improved image classification. Journal of Geomatics, 11(1), 1-21.

[5] Serpico, S. B., Bruzzone, L., & Roli, F. (1996). An experimental comparison of neural and statistical non-parametric algorithms for supervised classification of remote-sensing images. Pattern recognition letters, 17(13), 1331-1341.

[6] El-Rahman, S. A. (2016) Performance of spectral angle mapper and parallelepiped classifiers in agriculture hyperspectral image. International Journal of Advanced Computer Science and Applications (IJACSA), 7(5), 55.

[7] Kuching, S. (2007). The performance of maximum likelihood, spectral angle mapper, neural network and decision tree classifiers in hyperspectral image analysis. Journal of Computer Science, 3(6), 419-423.

[8] Liu, X., & Yang, C. (2013, December). A kernel spectral angle mapper algorithm for remote sensing image classification. In 2013 6th International Congress on Image and Signal Processing (CISP) (Vol. 2, pp. 814-818). IEEE.

[9] Sisodia, P. S., Tiwari, V., & Kumar, A. (2014, September). A comparative analysis of remote sensing image classification techniques. In 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI) (pp. 1418-1421). IEEE.

[10] Niknejad, M., Zadeh, V. M., & Heydari, M. (2014). Comparing different classifications of satellite imagery in forest mapping (case study: Zagros forests in Iran). International Research Journal of Applied and Basic Sciences, 8(9), 1407-1415.

[11] Doma, M.L., M.S. Gomaa and R.A. Amer, R.A. (2015) Sensitivity of pixel-based classifiers to training sample size in case of high resolution satellite imagery. Journal of Geomatics, 9(2): 53-58.

[12] Lv, W., & Wang, X. (2020). Overview of hyperspectral image classification. Journal of Sensors, 2020.

[13] Nagne, A. D., Dhumal, R. K., Vibhute, A. D., Gaikwad, S., Kale, K., & Mehrotra, S. (2017, October). Land use land cover change detection by different supervised classifiers on LISS-III temporal datasets.In 2017 1st International Conference on Intelligent Systems and Information Management (ICISIM)

[14] El_Rahman, S. A., & Zolait, A. H. S. (2017). Oil spill hyperspectral data analysis: using minimum distance and binary encoding algorithms. International Journal of Computing and Network Technology, 5(01)

[15] Wacker, A. G., and D. A. Landgrebe. "Minimum distance classification in remote sensing." LARS Technical Reports (1972).

[16] Renuka Bhokarkar Vaidya, Dhananjay Nalawade, Dr KV Kale, Vaibhav A. Didore, "HYPERSPECTRAL IMAGERY FOR CROP YIELD ESTIMATION IN PRECISION AGRICULTURE USING MACHINE LEARNING APPROACHES: A REVIEW", International Journal of Creative Research Thoughts (IJCRT), ISSN:2320-2882, Volume.9, Issue 9, pp.a777a789, September 2021, http://www.ijcrt.org/papers/IJCRT2109095.pdf

[17] Silva, A. C. G., & Trevisan, G. D. M. (2019). Comparison of the classification of land use of the municipality of Frederico Westphalen-RS, using the ISODATA and Minimum Distance. Nativa: Pesquisas Agrárias e Ambientais, 7(6), 727-733.

[18] Workman Jr, J., & Springsteen, A. (1998). Applied spectroscopy: a compact reference for practitioners. Academic Press.

[19] Liu, X., & Yang, C. (2013, December). A kernel spectral angle mapper algorithm for remote sensing image classification. In 2013 6th International Congress on Image and Signal Processing (CISP) (Vol. 2, pp. 814-818). IEEE.

[20] Kruse, F. A., Lefkoff, A. B., Boardman, J. W., Heidebrecht, K. B., Shapiro, A. T., Barloon, P. J., & Goetz, A. F. H. (1993). The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data. Remote sensing of environment, 44(2-3), 145-163.

[21] Bhokarkar, Renuka P, Sandeep V. Gaikwad, and K. V. Kale. "Development of ATIS-web based system for Aurangabad city." (2018).

[22] Nalawade, D. B., Kashid, S. D., Dhumal, R. K., Nagne, A. D., & Kale, K. V. (2015). Analysis of present transport system of Aurangabad city using geographic information system. International Journal of Computer Sciences and Engineering, 3(6), 124-128.

[23] https://earthexplorer.usgs.gov/ access on 6 Sep 2021.

[24] Research Systems, Inc. (2003). ENVI user's guide. Research Systems.

[25] Fourati, H. T., Bouaziz, M., Benzina, M., & Bouaziz, S. (2017). Detection of terrain indices related to soil salinity and mapping salt-affected soils using remote sensing and geostatistical techniques. Environmental monitoring and assessment, 189(4), 177