2021 Sixth International Conference on Image Information Processing (ICIIP) | 978-1-6654-3361-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICIIP53038.2021.9702568

# Quantitative Analysis of Different Land-use and Land-cover Classifiers using Hyperspectral Dataset

Neelam Dahiya
Chitkara University Institute of Engineering and Technology,
Chitkara University, Punjab, India.
neelam.ran@chitkara.edu.in
ORCID: 0000-0003-1839-9040

Sheifali Gupta
Chitkara University Institute of Engineering and Technology,
Chitkara University, Punjab, India.
sheifali.gupta@chitkara.edu.in
ORCID: 0000-0001-5692-418X

Sartajvir Singh
Chitkara University School of Engineering and Technology, Chitkara University,
Himachal Pradesh, India.
sartajvir.singh@chitkarauniversity.edu.in
ORCID: 0000-0002-4451-4949

***Abstract*— Remote sensing plays a vital role in analyzing land-use and land-cover (LULC) areas using hyperspectral imagery which was not feasible with multispectral imagery. The analysis of LULC is an important factor for environmental application. Classification is one of the methods used for the categorization of different types of class cateogries. The performance evaluation of various classifiers using hyperspectral dataset is essential to be performed for the categorization of various classes. Therefore, the main aim of this paper is to analyze and implement the various supervised classifiers such as random forest (RF), neural network (NN), and minimum distance classifier (MDC). The hyperspectral dataset has been used for these classifiers over a part of Haryana and Uttar Pradesh states, India. The results have shown that NN (88%) algorithm has achieved higher accuracy than other algorithms .These experimental outcomes highlights the potential of NN in handling complex problems and useful in the mapping of changes over LULC using hyperspectral imagery. This study is beneficial in applications such as crop monitoring, mapping seasonal variations, and snow detection.**

***Keywords—Hyperspectral, Land-use (LU) random forest (RF), neural network (NN), minimum distance classifier (MDC).***

## I. Introduction

The term Land-use and Land-cover (LULC) provides information about the area covered by wetland, agriculture ecosystem and water bodies, etc. [1]. The LULC changes are vigorous and difficult to acquire correct information with traditional methods. The main reasons found behind the changes in the LULC region are human interference like urbanization, industrialization, environmental changes, etc. [2]. Therefore, accurate monitoring of such regions is required which can be achieved with the help of remote sensing. Remote sensing and geographical information system (GIS) play an important role in mapping and monitoring LULC changes. Remote sensing plays a vital role in analyzing LULC areas using hyperspectral imagery which was not feasible with multispectral imagery [3].

Hyperspectral imagery is used to capture the fine and accurate information which is present in different narrow bands. In contrast to multispectral imaging, hyperspectral imaging provides various advantages such as crop identification, agriculture, identification of water bodies, etc [4]. In recent years, studies on different change detection algorithms and classification procedures were addressed to monitor or detect the changes over different land cover regions using multispectral datasets but rarely done with hyperspectral datasets [5].

Classification is the most common method which is used to categorize the given data into various classes based on training data. It can be done on both structured and unstructured data. The classification process is further divided into supervised and unsupervised. Supervision is required to train the model for supervised learning, whereas no supervision is required to train the model for unsupervised learning. There are various classifiers used for classification such as random forest (RF), neural network (NN) [6], minimum distance classifier (MDC) as supervised, and k-means [7] as unsupervised. These classifiers are used to identify the changes in the LULC with satellite imaging.

During the past decade, various methodologies have been used for HSI classification. The common methods for HSI classification are pixel-based such as support vector machine (SVM) [8]. In addition, some other approaches used are principal component analysis [9] and independent component analysis(ICA) [10].The most popular classifier used for image classification and to utilize spatial and spectral information from hyperspectral data is Convolutional neural Network (CNN) [11].On the other hand, some researchers focused on KNN [12] for small datasets for spectral-spatial hyperspectral image classification .

In the present paper, a comparative analysis of various supervised classifiers namely RF, NN, and MDC has been performed. The accuracy assessment is used to validate the applicability of each classifier. This paper also describes the advantages and disadvantages of each classifier. The paper is arranged into four main sections. Following this introduction part, the second section provides the details about the study area and satellite dataset. The third section consists of the description of various classifiers which are used for the present study. The fourth section includes the results and discussion of classifiers based on accuracy assessment, followed by a conclusion.

## II. Study Area And Satellite Dataset

### A. *Study Area*

The study area is a part of Haryana and Uttar Pradesh states (India) acquired from Hyperion E0-1 having geographical coordinates from 29°4' N to 77°1" N in latitude and 22°2' E to 18°4" E in longitude as shown in Fig. 1. The classes which are selected for the study include deciduous vegetation, dense vegetation, built-up, barren, and water area. Vegetation plays an important role in the ecosystem and includes various parts such as crops, shrubs, forests, wetlands, etc. So, its continuous monitoring is necessary for the management of environment and earth resources which can be achieved with satellite datasets [13]. The Hyperion EO-1 dataset is used for the present study and is useful for monitoring various changes over LULC. Different supervised and unsupervised classifiers are used to detect changes over the study area. The dataset found applications in the field of snow detection, weather monitoring, crop identification, agriculture, etc.

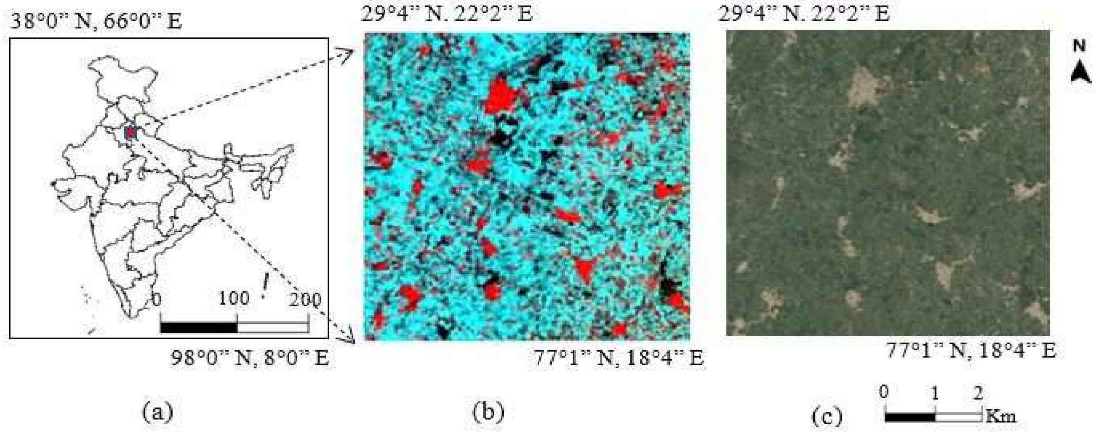

Fig. 1. Location of study site (a) Image of India (highlighted area representing selected study area) (b) False color image of the study area (07, March 2017) (c) Ground image of the study area

### *B. Satellite Dataset*

In the present study, a cloud-free hyperspectral dataset namely Hyperion EO-1 is used which was acquired on 07 March 2017 with a spatial resolution of 30-m and found various applications in the field of agriculture, wetland, forest, etc [14]. It consists of 242 spectral bands which cover the wavelength in the range of 356-2577 nm. The dataset was downloaded from Unites States Geological Survey (USGS) earth explorer's online web platform (https://earthexplorer.usgs.gov/). The hyperspectral dataset provides more information due to a large number of spectral bands as compared to multispectral datasets. It is also helpful in the extraction of various earth surface features.

The technical specifications of the hyperspectral imaging dataset are shown in Table I.

TABLE I. TECHNICAL SPECIFICATIONS OF THE HYPERSPECTRAL DATASET.

| S. No. | Parameters | Values/Name |
|---|---|---|
| 1. | Dataset | Hyperion EO-1 |
| 2. | Acquisition Date | 07/03/2017 |
| 2. | Number of Bands | 242 |
| 3. | Spectral range | 0.4 - 2.5 µm |
| 4. | Spatial resolution | 30 (m) |
| 5. | Swath width | 7.5 km |
| 6. | Spectral resolution | 10 (nm) |
| 7. | Spectral coverage | continuous |
| 8. | Pan-band resolution | N/A |
| 9. | File Type | .TIF |
| 10. | Image dimension | 1000 × 3271 |

## III. CLASSIFICATION APPROACHES

The three supervised classification algorithms named (a) RF (b) NN and (c) MDC have been implemented using the Hyperion EO-1 dataset. The supervised classification approaches have been used to dispense a precise evaluation of classes over pixels. In these approaches, training is given to the model on labeled data which helps predict future outcomes. Each classifier is outlined below with its various advantages and disadvantages and the taxonomy of the classifier is shown as in Fig. 2.

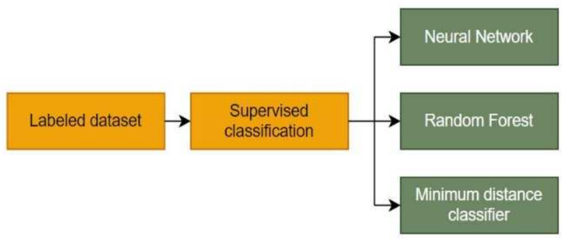


Fig. 2. Taxonomy of different classifiers is considered in the present work.

### *A. NN*

It is also known as an artificial neural network (ANN) which is comprised of the input layer, output layers, and various hidden layers. The number of hidden layers can be chosen according to requirement of applications [15]. The main functionality of hidden layers is to perform non-linear transformation of the inputs entered into the network . Here, input is fed forward through the neurons to get the appropriate result. It depends upon the past training data to learn and can improve accuracy over time. It serves various applications in the field of agriculture, object detection, weather forecasting, etc. Some of the drawbacks of NN are (a) Expensive (b) Needs more data (c) Hardware dependence. Instead of such limitations, it performs very well for LULC classification and is also used for various complex problems [16].

### *B. RF*

It is a supervised algorithm that is used to resolve classification problems. It is based on the ensemble learning method which consists of numerous decision trees on various subsets of the given dataset [17]. Instead of relying on one decision tree, the average of all is considered for output prediction which improves the accuracy of the algorithm. This algorithm takes lesser time as compared to others and also prevents overfitting issues. It is also capable to handle the value of the missing dataset automatically. The various application areas of RF are agriculture, healthcare, and banking sectors. One of the major limitations of this algorithm that can hamper its performance is more training time as compared to others algorithm [18]. It is also complex due to the involvement of various trees to generate the output.

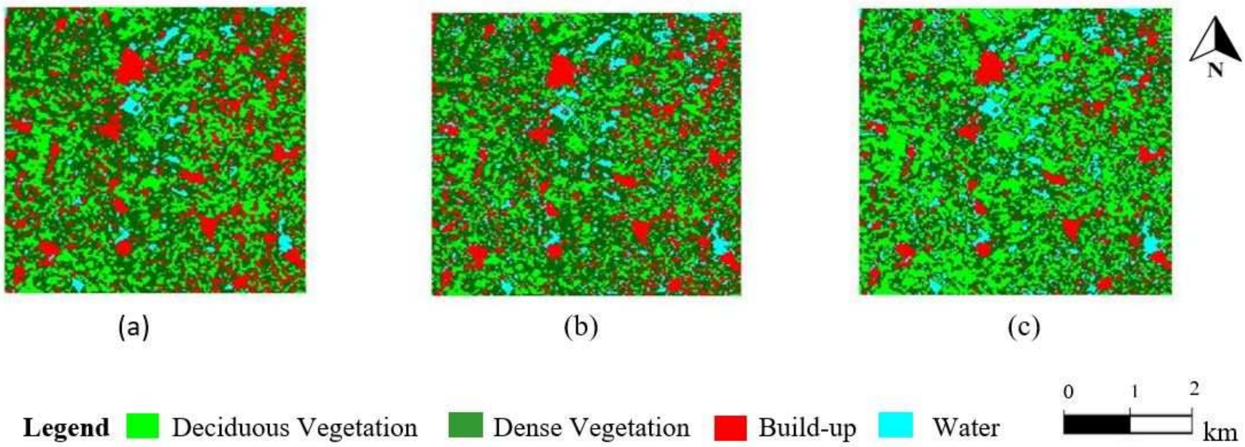


Fig. 3. Classified output using different classifiers (a) random forest (RF) (b) neural network (NN) (c) maximum-likelihood classifier (MLC)

TABLE II. ACCURACY ASSESSMENT OF DIFFERENT CLASSIFIERS.

| Classifier | Classified Data | Reference Total (%) | Classified Total (%) | Correct Number (%) | Producer Accuracy (%) | User Accuracy (%) | Kappa value | Overall Accuracy (%) | Overall Kappa |
|---|---|---|---|---|---|---|---|---|---|
| RF | Deciduous Vegetation | 36 | 22 | 22 | 61.11 | 100 | 1.000 | 85 | 0.7713 |
| | Dense Vegetation | 49 | 49 | 49 | 100 | 100 | 1.000 | | |
| | Built-up | 11 | 19 | 10 | 90.91 | 52.63 | 0.4678 | | |
| | Water | 04 | 10 | 04 | 100 | 40 | 0.3750 | | |
| NN | Deciduous Vegetation | 32 | 27 | 27 | 84.38 | 100 | 1.000 | 88 | 0.8224 |
| | Dense Vegetation | 47 | 43 | 40 | 85.11 | 93 | 0.8684 | | |
| | Built-up | 15 | 20 | 15 | 100 | 75 | 0.7059 | | |
| | Water | 06 | 10 | 06 | 100 | 60 | 0.5745 | | |
| MDC | Deciduous Vegetation | 34 | 33 | 31 | 91.18 | 93.94 | 0.9082 | 83 | 0.7277 |
| | Dense Vegetation | 56 | 45 | 42 | 75.00 | 93.33 | 0.8485 | | |
| | Built-up | 08 | 12 | 08 | 100 | 66.67 | 0.6377 | | |
| | Water | 02 | 10 | 02 | 100 | 20 | 0.1837 | | |

*Note RF: random forest; NN: neural network, MDC: minimum distance classifier*

### *C. MDC*

Conventionally, the MDC model is used to classify unknown images and for pattern recognition. The main idea is to calculate the minimum distance from the mean value of each class of the training data to the digital value of each pixel in the imagery [19]. The Euclidean distance method is used to find the minimum distance. The main advantage of this algorithm is that every pixel is designated to a class and is very easy to compute. Its training is simple and easy as compared to others. However, the major drawback of this approach is that it does not account for class variability. Moreover, it also leads to poor classification due to insufficient training samples [20].

## IV. RESULTS AND DISCUSSION

The Hyperion EO-1 dataset is used as input data, acquired on 07 March 2017. After data acquisition, three well-known supervised classification algorithms such as RF, NN, and MDC are implemented to explore the impact of the Hyperion EO-1 dataset over the LULC region. During the classification, various categories are explored as deciduous vegetation, dense vegetation, built-up, and water. The final classified output of different classifiers is shown in Fig. 3. To classify the dataset, more than 20 samples are selected for each category to process the data. The evaluation of various classifiers is done with the help of accuracy assessment. The accuracy assessment includes the following terms such as producer accuracy, user accuracy, overall accuracy, Kappa coefficient which are calculated using an error matrix [21]. The accuracy assessment is shown in Table 2. The accuracy assessment table shows that NN (88%) outperforms as compared to RF (85%), MDC (83%) for the LULC area. The main reason behind these results is that NN can learn and improve over time by itself as compared to other algorithms. Moreover, NN performs better for both small as well as large datasets. However, the user's and producer's accuracy can be improved by making algorithms fully automatic.

## V. Conclusion

In this paper, different supervised algorithms i.e Neural Network (NN), Random Forest (RF), and Minimum Distance Classifier (MDC) have been implemented to evaluate the best classification technique for LULC. The comparative analysis of these classifiers is done with the help of a hyperspectral dataset i.e., Hyperion EO-1 over the part of Haryana state, India. This study is done to evaluate the Hyperion EO-1 over LULC accurately. The present study shows that the NN supervised algorithm achieved better and highest accuracy for LULC monitoring as compared to another algorithm. Further, it can also be used for other study areas.

## VI. FUTURE SCOPE

The present paper is helpful for the researcher to collect basic knowledge about various latest classifers.This study can also be enhanced for snow detection, forest analysis, plants diseases monitoring, etc.The researcher will be able to use this technique on various different field with more ease.

## Acknowledgment

The authors would like to thank United States Geological Survey (USGS) for providing the Hyperion EO-1 (Earth Observation) dataset for research purposes.

## References

[1] A. M. Dewan and Y. Yamaguchi, "Land use and land cover change in Greater Dhaka, Bangladesh: Using remote sensing to promote sustainable urbanization," *Appl. Geogr.*, vol. 29, no. 3, pp. 390–401, 2009, doi: 10.1016/j.apgeog.2008.12.005.

[2] S. Twisa and M. F. Buchroithner, "Land-use and land-cover (LULC) change detection in Wami river basin, Tanzania," *Land*, vol. 8, no. 9, 2019, doi: 10.3390/land8090136.

[3] P. Mishra, M. S. M. Asaari, A. Herrero-Langreo, S. Lohumi, B. Diezma, and P. Scheunders, "Close range hyperspectral imaging of plants: A review," *Biosyst. Eng.*, vol. 164, pp. 49–67, 2017, doi: 10.1016/j.biosystemseng.2017.09.009.

[4] H. Okamoto and W. S. Lee, "Green citrus detection using hyperspectral imaging," *Comput. Electron. Agric.*, vol. 66, no. 2, pp. 201–208, 2009, doi: 10.1016/j.compag.2009.02.004.

[5] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *Int. J. Remote Sens.*, vol. 28, no. 5, pp. 823–870, 2007, doi: 10.1080/01431160600746456.

[6] S. Singh, R. K. Tiwari, V. Sood, and H. S. Gusain, "Detection and validation of spatiotemporal snow cover variability in the Himalayas using Ku-band (13.5 GHz) SCATSAT-1 data," *Int. J. Remote Sens.*, vol. 42, no. 3, pp. 805–815, 2021, doi: 10.1080/2150704X.2020.1825866.

[7] B. Simhachalam and G. Ganesan, "Performance comparison of fuzzy and non-fuzzy classification methods," *Egypt. INFORMATICS J.*, 2015, doi: 10.1016/j.eij.2015.10.004.

[8] V. Sood and S. Gupta, "A Comparative Review on Different Sub-Pixel Classification Algorithms," no. February, pp. 7–9, 2013, [Online]. Available: https://www.researchgate.net/publication/323185559_A_Comparative_Review_on_Different_Sub-Pixel_Classification_Algorithms.

[9] M. P. Uddin, M. Al Mamun, and M. A. Hossain, "Effective feature extraction through segmentation-based folded-PCA for hyperspectral image classification," *Int. J. Remote Sens.*, vol. 40, no. 18, pp. 7190–7220, 2019, doi: 10.1080/01431161.2019.1601284.

[10] E. Ibarrola-Ulzurrun, J. Marcello, and C. Gonzalo-Martin, "Assessment of component selection strategies in hyperspectral imagery," *Entropy*, vol. 19, no. 12, 2017, doi: 10.3390/e19120666.

[11] R. Vaddi and P. Manoharan, "Hyperspectral image classification using CNN with spectral and spatial features integration," *Infrared Phys. Technol.*, vol. 107, no. April, p. 103296, 2020, doi: 10.1016/j.infrared.2020.103296.

[12] C. Bo, H. Lu, and D. Wang, "Spectral-spatial K-Nearest Neighbor approach for hyperspectral image classification," *Multimed. Tools Appl.*, pp. 1–18, 2017, doi: 10.1007/s11042-017-4403-9.

[13] J. Im and J. R. Jensen, "Hyperspectral Remote Sensing of Vegetation," *Hyperspectral Remote Sens. Veg.*, vol. 6, pp. 1943–1961, 2016, doi: 10.1201/b11222.

[14] K. S. Khurshid *et al.*, "Preprocessing of EO-1 Hyperion data," *Can. J. Remote Sens.*, vol. 32, no. 2, pp. 84–97, 2006, doi: 10.5589/m06-014.

[15] M. Awad, "Sea water chlorophyll-a estimation using hyperspectral images and supervised Artificial Neural Network," *Ecol. Inform.*, vol. 24, pp. 60–68, 2014, doi: 10.1016/j.ecoinf.2014.07.004.

[16] T. Zhan *et al.*, "TDSSC: A Three Directions Spectral-Spatial Convolution Neural Networks for Hyperspectral Image Change Detection," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 1680, pp. 1–1, 2020, doi: 10.1109/jstars.2020.3037070.

[17] G. Krishna, R. N. Sahoo, S. Pradhan, T. Ahmad, and P. M. Sahoo, "Hyperspectral satellite data analysis for pure pixels extraction and evaluation of advanced classifier algorithms for LULC classification," *Earth Sci. Informatics*, vol. 11, no. 2, pp. 159–170, 2018, doi: 10.1007/s12145-017-0324-4.

[18] N. Puletti, N. Camarretta, and P. Corona, "Evaluating EO1-Hyperion capability for mapping conifer and broadleaved forests," *Eur. J. Remote*

*Sens.*, vol. 49, no. June, pp. 157–169, 2016, doi: 10.5721/EuJRS20164909.

[19] A. Sharma, K. K. Paliwal, and G. C. Onwubolu, “Class-dependent PCA, MDC and LDA: A combined classifier for pattern classification,” *Pattern Recognit.*, vol. 39, no. 7, pp. 1215–1229, 2006, doi: 10.1016/j.patcog.2006.02.001.

[20] A. K. Shakya, A. Ramola, A. Kandwal, and R. Prakash, “Comparison of Supervised Classification Techniques With Alos Palsar Sensor for Roorkee Region of Uttarakhand, India,” *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. XLII–5, no. November, pp. 693–701, 2018, doi: 10.5194/isprs-archives-xlii-5-693-2018.

[21] R. G. Congalton, “A review of assessing the accuracy of classifications of remotely sensed data,” *Remote Sens. Environ.*, vol. 37, no. 1, pp. 35–46, 1991, doi: 10.1016/0034-4257(91)90048-B.