Journal of Theoretical and Applied Information Technology

31st October 2023. Vol.101. No 20
© 2023 Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

CLASSIFICATION OF MULTISPECTRAL IMAGES FOR LAND USE AND LAND COVER USING TRANSFORMER MODEL

ADLENE EBENEZER P [1] S. MANOHAR [2]

[1,2]Department of Computer Science and Engineering, College of Engineering and Technology, SRM Institute of Science and Technology, Vadapalani Campus, Vadapalani, Chennai, India.

E-mail: [1]ae6494@srmist.edu.in, [2]manohars@srmist.edu.in

ABSTRACT

Land cover is the term used to describe the geographical coverage that encompasses the planet's surface. It consists of trees, greenery, concrete, barren ground, and water bodies. Remote imaging is generally used to obtain data on the land covering. The information collected is then classified according to different land covers using an approach called land cover classification. In order to predict future changes in land utilisation and land cover for the forest, and non-forest covered regions of the hilly region in the region of Tiruvanamallai, India, this research develops the Vision Transformer framework. With an overall accuracy of 94.01% in 2015 and 94.19% in 2018, the Land Use/Land Cover (LU/LC) classification map during the classification yields positive validation results.

Keywords: *Land Use/Land Cover, Transformer model, Forest, Non-forest, Classification*

1. INTRODUCTION

Understanding the different driving forces behind changes in Land Use/Land Cover has recently become a prime study theme in the field of human aspects of worldwide change. LU/LC change, which are significantly associated with economic prosperity and natural alterations, have been fundamental areas of globally environmental development. For both global and territorial land resource exploration, a significant number of Land Use research initiatives have offered information and support assistance. The foundation for conducting study on Land Cover (LC) change is accurate classification [1]. Researchers have created multispectral approaches that can concurrently gather information on multiple wavelength intensities due to the advancement of remote sensing technology, providing detailed feature data on the target object [2]. The amount of accessible satellite imagery is constantly increasing as there are more satellites for Earth observation in orbit. Due to the abundance of data, statistical learning techniques that take advantage of big datasets can be used to solve a variety of Earth observation issues [3].

The satellite data are explained by remote sensing technology, which also facilitates efficient LU/LC prediction study. The data was collected by researchers using a variety of remote sensing satellite devices. The most important step in the

LULC change evaluation is choosing an appropriate dataset. Aerial photos, satellite images, auxiliary information, Google Maps were used to gather the images for the National Forest Service, Land Resource Management, and Urban developers. Meticulous collection is necessary in order to process the datasets from various time intervals in a remote sensing surroundings. And a few of the datasets for satellite systems are LISS-III (Linear Imaging Self-Scanning Sensor III) [4], LISS-IV (Linear Imaging Self-Scanning Sensor IV) from bhuvan portal [5], Landsat Series from earth explorer [6], Sentinel-2A and -2B from European Space Agency [7], MODIS (Moderate Resolution Imaging Spectroradiometer) [8], REIS (Rapid Eye Earth Imaging System), DEM (Digital Elevation Model) for identifying Independent Variable [9].

Pre-processing was the first action taken following data collection. This procedure was carried out in order to improve the satellite image's clarity and quality. One of the most crucial pre-processing stages is atmospheric correction as it can have an impact on the outcome. The primary goal of

atmospheric correction is to correct impacts on satellite images by determining the optical properties [10]. Direct terrain mapping cannot be done with satellite imagery because it frequently has geometrical distortions, especially if it comes to spatio-temporal analysis. Several variables, including distortion, terrain curvature, and landscape topography impact, and territory correction, can lead to such errors [11].

Remote sensing imagery must be improved through radiometric correction in order to give each pixel a measurable physical unit. Any mapping application reliant on remote sensing depends on selecting a relevant image for radiometric and atmospheric correction level. [12]. While using remote sensing images to explore forest resources, topographic correction can minimise the impact of geography while enhancing the accuracy of various kinds of forest trees classification [13]. The extraction of the area of interest, filtering [15], pan sharpening [16], equalisation of histograms [17], layer stacking, edge enhancement, and mosaicking [14] of satellite images are additional spatial enhancement techniques.

The most important aspect of the remote sensing study to be determined is the LU/LC change classification. Various classification algorithms are used to categorize the material on the surface of the earth. The LU/LC analysis outcomes are contrasted with the variables of the ground truth using the evaluation of attainment. According to the training approach, the LU/LC change algorithms for classification are divided into supervised and unsupervised classifiers. K-means, Iterative Self-Organizing Data Analysis (ISODATA) clustering, Fuzzy C-means clustering, and the SOM (Self Organizing Map) Neural Network approach are the Unsupervised techniques.

The unsupervised learning model's drawback is that it requires a lot of processing time when the data are unorganized. MLC (Maximum Likelihood Classifier) [18], kNN (k-Nearest Neighbour) [19], Minimum Distance Classifier, Mahalanobis distance, SVM (Support Vector Machines), Parallelepiped, and RFC (Random Forest Classifiers) are a few examples of frequently used Supervised techniques [20].

Post-classification change identification was used to analyse the LU/LC variations in the time series data for the study region. The basic LU/LC Machine-Learning model has the problem of not having accurate information of the predicted map, which makes it challenging for urban planners to analyse the information further.

Accuracy evaluates how precisely the results resemble the real values and provides a comprehensive compilation of data from the obtained satellite imagery. A confusion matrix of data that distinguishes between user and producer accuracy has been established to evaluate the accuracy of supervised image classification. [21].

The extraction of LU/LC features has become a crucial aspect of research, and therefore, there is a need to establish a standard and accurate methodology for LU/LC classification. After conducting a brief survey on various classification algorithms, it has been established that the sustainable growth of the LU/LC environment for time-series data necessitates an accurate classification map. The Vision Transformers proposed for classifying LU/LC classes successfully performed on images from 2012 and 2015.

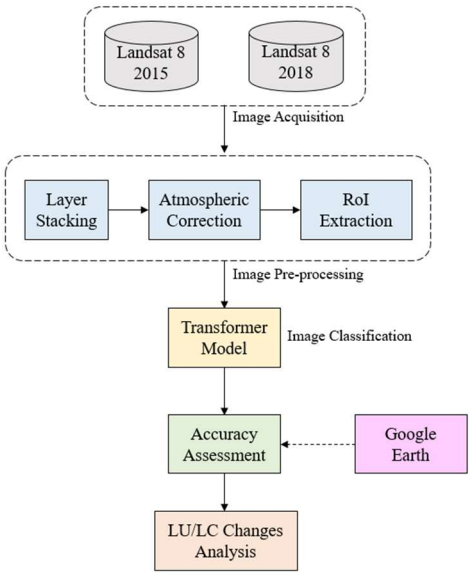
The following part of the article is provided as follows: The purpose of this motivation and its contribution are discussed in Section 2. The suggested approach for this research objective is illustrated in Section 3. Section 4 describes the outcomes and conclusions put forth in our research work, and Section 5 presents the study's conclusion.

2. MOTIVATION

The primary contribution of researchers worldwide is to offer new, cutting-edge information to the public, the government, and various institutions of learning in their corresponding fields. A detailed classification map is necessary for the time-series data used in the LU/LC environment, which is why our research was strongly motivated by the results of our brief survey of various classification algorithms. For the classification of Land Use/Land Cover changes in our study region, a Vision Transformer model is proposed. The proposed model, based on Vision Transformer, classifies LU/LC changes in our study area.

2.1 The LULC Classification Flow

Figure 1 shows the progressive work of LU/LC classification is in the context of the remote sensing environment that we have developed. The comprehensive Land Use/Land Cover classification workflow is defined as follows:



- Image Acquisition: The initial data acquisition procedure, which involved downloading and analyzing a time series of satellite photos for a particular area.
- Image Pre-processing: Enhancing the original remote sensing image using various pre-processing methods.
- Classification: Pre-processed image is classified utilising Transformer techniques under the LULC scheme.
- Validation: The comparison of a reference map to the classed map for evaluation.

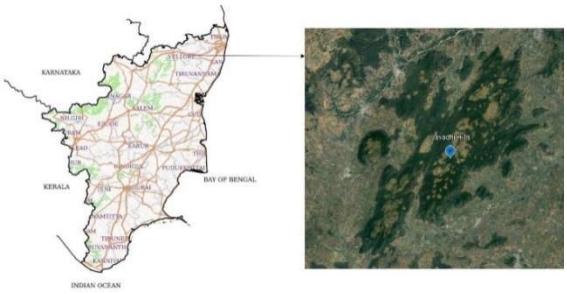


Figure 1: Outline for Land Cover Classification

3. MATERIALS AND METHODS

3.1 Study Region and Data Collection

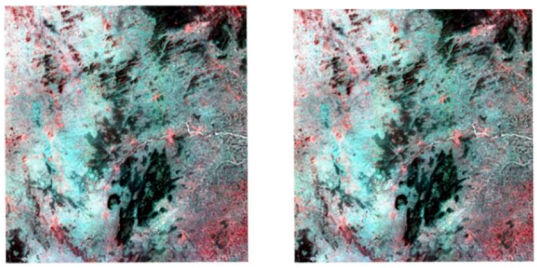
The Javadi Hills' forest- and non-forest-covered areas, whose geographic coordinates range from 78.74 E 12.4 N to 79.1 E 12.65 N, serve as the subject region for our research project. Our study region is situated between the Eastern Ghats in Tamil Nadu, India's district of Vellore and Tiruvannamalai.

Figure 2: Location of Study area

The images from Landsat 8 are taken for the purpose of our research. The USGS (United States Geological Survey) provided the Landsat 8. Figure 2 displays the study area map of the study region, which was taken from GEE. The characteristics of the remote sensing imagery from satellites are displayed in Table 1.

Table 1. Details of the satellite images

Satellite	Year	Source	Format
Landsat 8 OLI/TI	2015 2018	United States Geographical Survey	GeoTiff



3.2 Image Pre-processing

In our study, atmospheric corrections were applied to ensure that the derived spectral remote sensing image of the study region had adequate visibility. In order to obtain the Area of Interest (AOI) coordinates of our study area that lie in 78.74 E. 12.4 N and 79.1 E. 12.65 N, a geometric correction was applied. Figure 3 represents the extracted image for our study area and pre-processed multispectral image that has been shown in Figure 4.

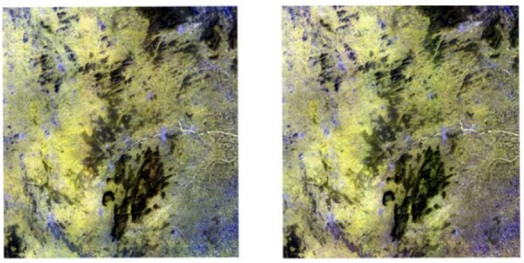
The technique of layer stacking is used to combine numerous multi spectral bands into a single image. This requires resampling additional bands whose spatial resolution is different from the intended resolution in order to make the images' extents (number of rows and columns) equal.

(a) *(b)*

Figure 3: Extracted Image for the year 2015 & 2018

By removing air effects from satellite images, an atmospheric correction is intended to ascertain real surface reflectance values and to extract physical properties of the Earth's surface,

including surface reflectance [22]. Perhaps the most crucial step in the pre-processing of data from satellites that have been remotely sensed is atmospheric correction. When comparing and analysing multi-temporal images, such a correction is essential.



(a) (b)

Figure 4: Pre-processed Image for the year 2015 & 2018

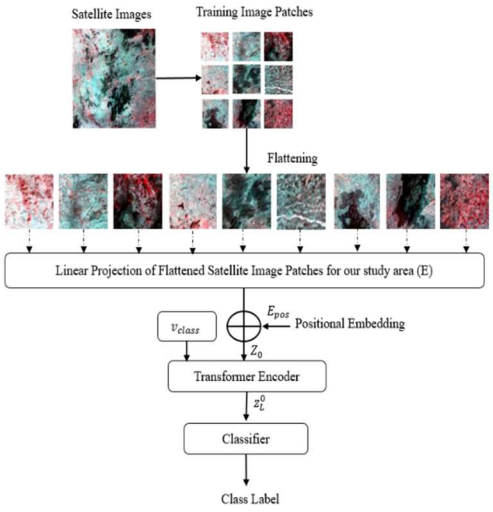
3.3 Transformer Model for LU/LC Classification

The self-attention method was used to create a deep learning algorithm called a transformer. The transformer analyses the sequential data simultaneously without relying on a recurrent network, adhering to the encoder-decoder architecture [23-27]. The Transformer builds an encoder module using the series of remote sensing image patches and the classified label to accomplish image classification. In these studies using the preprocessed Landsat satellite images of Javadi Hills from 2015 and 2018, the Transformer classification model was used.

Let Ti and Ti = {Yi, xi}i=1r denotes a series of remote sensing images, r, where Yi is a remote sensing image; xi denotes the class labels {xi ϵ 1, 2, . . . , n} correlated with the Yi, and n indicates the count of defined LU/LC classes. Figure 5, describes the architecture model for Classification.

A multispectral image of Yi from the training set is used as the input for the initial stage of the Transformer model. The image is segmented into a set of patches in the same size that do not overlap. The Transformer considers each patch as a unique token. Then retrieved the patches of proportion cp2 where p is the patch size from an image of size h*w*d in which h is the height, c is the number of channels, and w is the width. The collected patches are flattened to create a series of images (y1, y2, y3,..., ym) of length m, where m is equal to hw/p2.

Figure 5: Architecture of the model



3.3.1 Linear Embedding Layer

With the standard embedding matrix, E, a vector configuration with model dimension d is created by linearly projecting the image segments. Along with the trained classification token vclass, embedding representations are integrated to complete the classification task. The patch representation is coded with the positional data, or Epos. Positional embedding was used to process the acquired image patches' spatial configurations. Equation (1) gives the set of spectral image patches that arise from positional embedding with token z0.

$$z_0 = [vclass; y1E; y2E; \dots ymE] + E_{pos}, E \epsilon \mathrm{R}^{(p^2c)*d},\ E_{pos} \epsilon \mathrm{R}^{(n+1)*d} \quad (1)$$

3.3.2 Transformer Encoder

The encoder for the transformer receives the series of embedded image patches, z_0, with L identical layers. The GeLU activation function is enclosed between a MSA and a fully connected feed-forward MLP block. The normalisation layer (LN),

which connects the two encoder subcomponents, facilitates communication with the remaining skip connections. Equations (2) and (3) provide a depiction of the encoder's two primary parts. The primary classifier receives the first element of the sequence z0L from the final layer of the encoder in order to produce the LU/LC categorised classes. In figure 6 the transformer encoder module has been explained.

$$Z_l^1 = MSA\left(LN\left(Z_{l-1}\right)\right) + Z_{l-1}, l = 1 \ldots L \quad (2)$$

$$Z_l = MLP\left(LN(Z_L^1)\right) + Z_l^1, l = 1 \ldots L \quad (3)$$

$$x_i = LN(Z_L^0) \quad (4)$$

The encoder- MSA block is regarded as the transformer's core element. The significance of a single patch image embedding in relation to the other embeddings in the sequence is determined by the MSA block. This block has four layers: a linear layer, a layer for self-attention, a layer for integration, and a final layer for linearity. The weighted total of each value in the sequence is used to calculate the attention weight. The SA head computes the Q_K_V scaling dot product using the attention weights. The element was multiplied against three learnt matrices, UQKV, to provide the query (Q), key (K), and value (V) (Equation (5)). The dot product is utilised among the one element of Q vectors and other elements of K vectors to determine the importance of the elements on the sequence. The result highlights how crucial the image patches are to the sequence. In Equation (6), the outcome of the dot product was performed and sent into the Activation function, Softmax.

$$[Q, K, V] = zU_{QKV}, U_{QKV} \epsilon \, \mathrm{R}^{d*3D_k} \quad (5)$$

$$A = softmax\left(\frac{QK^T}{\sqrt{D_K}}\right) A \epsilon \, \mathrm{R}^{n*n} \quad (6)$$

The SA (Self-Attention) block's scaling-dot-product procedure relates to the common dot product but also takes into account the key DK's dimension as a scaling factor. The outputs that are obtained through Softmax are products of the values of each patch embedding vector to process the patches with the highest attention scores. In order to obtain the feature map along with pixel value representation, the output of all attention heads is combined and given to the Multi-Layer Perceptron classification. Resampling was done to alter the size of the feature map by ensuring that the obtained classed image is displayed in a uniform manner when accuracy was being evaluated. Through the accuracy assessment, the classified map evaluation is accomplished. For our study region, the classified map has been examined for the LU/LC changes analysis's additional information in light of the good accuracy results.

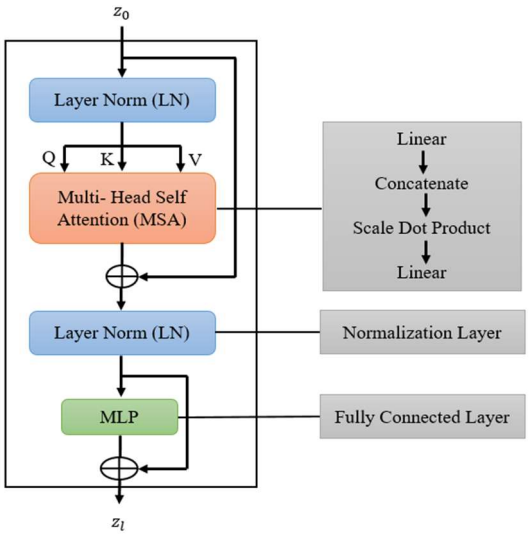


Figure 6: Transformer Encoder Module

4. RESULT AND DISCUSSION

In the present paper, the LU/LC classification in our study area between 2015 and 2018 was analyzed. Using imagery from satellites between 2015 and 2018, the LU/LC analysis in the study region's non-forested and forest-covered areas had been studied.

On an Intel Core i7 CPU running Windows 11 (64-bit), all experiments were performed using Visual Studio and geospatial processing tools like ArcGIS and Google Earth Engine.

4.1 Accuracy Evaluation

An accuracy evaluation needs to be carried out in order for the classified images to be acknowledged more precisely, by contrasting the classification outcomes with the ground truth value, the accuracy was determined [28-30]. From the input training image, 9 patches (size = 64 * 4) were recovered, and each patch comprises the trained classes.

The Transformer model was used to do the LU/LC classification for the input image from our

© 2023 Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

study area. Table 2 displays the hyper-parameters applied during the training of the classification model.

Table 2. Hyperparameters of Transformer Model

Hyper_parameters	Values
Learning_Rates	0.001
Weight_Decay	0.0001
Batch_Size	15
Number_of_Epochs	50
Image_Size	256*256
Patch_Size	64
Patchs_per_Images	9
Number_of_Heads	4
Transformer_Layers	8
Activation_Function	GeLU
Optimizer	Adam

In order to evaluate the accuracy of the suggested approach, 551 random sampling points were gathered for each identified LU/LC class. Google Earth images that were captured during time-consistent intervals were used to observe the reference points that were obtained. Along with the kappa statistics result, a precision, recall, and overall accuracy report is produced. The accuracy evaluation for 2015 and 2018 is shown in Table 3 and 4.

Table 3: Description of LULC Class

2015		
LU/LC Class/Year	Producer's Accuracy	User's Accuracy
Forest	97.12	95.62
Non- Forest	90.46	96.57
2018		
LU/LC Class/Year	Producer's Accuracy	User's Accuracy
Forest	98.72	94.85
Non- Forest	91.35	93.24

Table 4: Accuracy Assessment for years 2015 & 2018

LU/LC Classification	2015	2018
Overall_Accuracy	94.01	94.19
Kappa_Coefficient	0.901	0.904
Precision	0.94	0.95
Recall	0.95	0.96
F1 Score	0.95	0.95

4.1 Comparative Analysis

We have the Transformer Model for classification in this research paper. For the Javadi Hills region, we compared our model to KNN, DT, SAM, DCNN, and traditional LU/LC classification methods. In terms of precision and computing efficiency, our model performs better than the other conventional classification techniques.

Table 5: Comparative Analysis for various Classification Algorithm

S.No	Algorithm	Accuracy (%)
1	Ours	94.12
2	KNN [31]	85.67
3	DT [32]	89.00
4	SAM [33]	90.45
5	DCNN [34]	91.03

The typical LU/LC algorithms have been executed, and Table 5 provides a comparison. Figure 7 shows the classification model's relative accuracy in comparison.

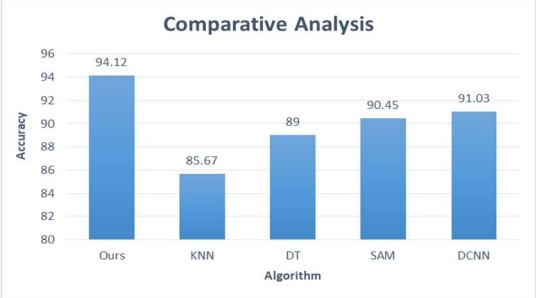


Figure 7: Performance Analysis for various Classification Algorithm

5. CONCLUSION

The significant impacts of human activity on the environment and natural resources are revealed by the LU/LC change. The major changes that took place in the study region assist policymakers and other urban planners in taking the appropriate decisions regarding land resource management. The multispectral Landsat satellite images of the study region from the years 2015 and 2018 were downloaded for this research. It was proposed to use the Transformer model to carry out the LU/LC classification, and Google Earth Images were used to evaluate the accuracy of the model. Our Transformer model had a 94.01% in 2015 and 94.19% in 2018 as classification accuracy. This examination of the change in our study region was done to pave the way for future research into the LU/LC change prediction for our study region.

REFERENCES

[1] Li, Zhengtao, Guokun Chen, and Tianxu Zhang. "A CNN-transformer hybrid approach for crop classification using multitemporal multisensor images." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 13 (2020): 847-858.

[2] Zhang, Zhiwen, et al. "Introducing Improved Transformer to Land Cover Classification Using Multispectral LiDAR Point Clouds." *Remote Sensing* 14.15 (2022): 3808.

[3] Scheibenreif, Linus, et al. "Self-Supervised Vision Transformers for Land-Cover Segmentation and Classification." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022.

[4] Hasim, Sheak, and Kalyan Kumar Bhar. "Seasonal Cropping Pattern Extraction Using NDVI from IRS LISS-III Image of Kangsabati Commanded Area." Procedia Computer Science 167 (2020): 900-906.

[5] Kumar, Ajay, and Amit Kumar Gorai. "Design of an optimized deep learning algorithm for automatic classification of high-resolution satellite dataset (LISS IV) for studying land-use patterns in a mining region." *Computers & Geosciences* 170 (2023): 105251.

[6] Hemati, MohammadAli, et al. "A systematic review of landsat data for change detection applications: 50 years of monitoring the earth." Remote sensing 13.15 (2021): 2869.

[7] Phiri, Darius, et al. "Sentinel-2 data for land cover/use mapping: a review." Remote Sensing 12.14 (2020): 2291.

[8] Palchaudhuri, Moumita, and Sujata Biswas. "Application of LISS III and MODIS-derived vegetation indices for assessment of micro-level agricultural drought." *The Egyptian Journal of Remote Sensing and Space Science* 23.2 (2020): 221-229

[9] Mesa-Mingorance, José L., and Francisco J. Ariza-López. "Accuracy assessment of digital elevation models (DEMs): A critical review of practices of the past three decades." Remote Sensing 12.16 (2020): 2630.

[10] Rumora, Luka, Mario Miler, and Damir Medak. "Impact of various atmospheric corrections on sentinel-2 land cover classification accuracy using machine learning classifiers." ISPRS International Journal of Geo-Information 9.4 (2020): 277.

[11] Gašparović, Mateo. "Urban growth pattern detection and analysis." Urban Ecology. Elsevier, 2020. 35-48.

[12] Kamal, Muhammad, Faaris H. Muhammad, and Shifa A. Mahardhika. "Effect of image radiometric correction levels of Landsat images to the land cover maps resulted from maximum likelihood classification." E3S Web of Conferences. Vol. 153. EDP Sciences, 2020.

[13] Dong, Chao, et al. "The effect of topographic correction on forest tree species classification accuracy." Remote Sensing 12.5 (2020): 787.

[14] Carvalho, Osmar Luiz Ferreira de, et al. "Instance segmentation for large, multi-channel remote sensing imagery using mask-RCNN and a mosaicking approach." Remote Sensing 13.1 (2020): 39.

[15] Bhatti, Uzair Aslam, et al. "Local similarity-based spatial–spectral fusion hyperspectral image classification with deep CNN and Gabor filtering." IEEE Transactions on Geoscience and Remote Sensing 60 (2021): 1-15.

[16] Ma, Jiayi, et al. "Pan-GAN: An unsupervised pan-sharpening method for remote sensing image fusion." Information Fusion 62 (2020): 110-120.
[17] Mousania, Younes, Salman Karimi, and Ali Farmani. "Optical remote sensing, brightness preserving and contrast enhancement of medical images using histogram equalization with minimum cross-entropy-Otsu algorithm." Optical and Quantum Electronics 55.2 (2023): 1-22.
[18] Alkaradaghi, Karwan, et al."Evaluation of land use & land cover change using multi-temporal landsat imagery: a case study Sulaimaniyah Governorate, Iraq." Journal of Geographic Information System 10.6 (2018): 247-260.
[19] Thakur, Reena, and Prashant Panse. "Classification Performance of Land Use from Multispectral Remote Sensing Images using Decision Tree, K-Nearest Neighbor, Random Forest and Support Vector Machine Using EuroSAT Data." International Journal of Intelligent Systems and Applications in Engineering 10.1s (2022): 67-77.
[20] Talukdar, Swapan, et al. "Land-use land-cover classification by machine learning classifiers for satellite observations—A review." Remote Sensing 12.7 (2020): 1135.
[21] Choudhury, Deblina, Kalikinkar Das, and Arijit Das. "Assessment of land use land cover changes and its impact on variations of land surface temperature in Asansol-Durgapur Development Region." The Egyptian Journal of Remote Sensing and Space Science 22.2 (2019): 203-218.
[22] Hadjimitsis, Diofandos G., et al. "Atmospheric correction for satellite remotely sensed data intended for agricultural applications: impact on vegetation indices." Natural Hazards and Earth System Sciences 10.1 (2010): 89-95.
[23] Kaselimi, Maria, et al. "A vision transformer model for convolution-free multilabel classification of satellite imagery in deforestation monitoring." IEEE Transactions on Neural Networks and Learning Systems (2022).
[24] Xue, Zhixiang, et al. "Deep hierarchical vision transformer for hyperspectral and lidar data classification." IEEE Transactions on Image Processing 31 (2022): 3095-3110.
[25] Lv, Pengyuan, et al. "SCViT: A spatial-channel feature preserving vision transformer for remote sensing image scene classification." IEEE Transactions on Geoscience and Remote Sensing 60 (2022): 1-12.
[26] Gao, Yunhao, et al. "Fusion classification of HSI and MSI using a spatial-spectral vision transformer for wetland biodiversity estimation." Remote Sensing 14.4 (2022): 850.
[27] Wang, Hongmiao, et al. "Land Cover Classification for Polarimetric SAR Images Based on Vision Transformer." Remote Sensing 14.18 (2022): 4656.
[28] Gashaw, Temesgen, et al. "Evaluation and prediction of land use/land cover changes in the Andassa watershed, Blue Nile Basin, Ethiopia." Environmental Systems Research 6.1 (2017): 1-15.
[29] Cheruto, Mercy C., et al. "Assessment of land use and land cover change using GIS and remote sensing techniques: a case study of Makueni County, Kenya." (2016).
[30] Rwanga, Sophia S., and Julius M. Ndambuki. "Accuracy assessment of land use/land cover classification using remote sensing and GIS." International Journal of Geosciences 8.04 (2017): 611.
[31] Mishra, P. K., Rai, A., & Rai, S. C. (2020). Land use and land cover change detection using geospatial techniques in the Sikkim Himalaya, India. The Egyptian Journal of Remote Sensing and Space Science, 23(2), 133-143.
[32] Das, S., & Sarkar, R. (2019). Predicting the land use and land cover change using Markov model: A catchment level analysis of the Bhagirathi-Hugli River. Spatial Information Research, 27(4), 439-452.
[33] Mohamed, M. M., & Elmahdy, S. (2018). Land use/land cover changes monitoring and analysis of Dubai Emirate, UAE using multi-temporal remote sensing data. EPiC Series in Engineering, 3, 1435-1443.
[34] Omeiza, D. (2019). Efficient machine learning for large-scale urban land-use forecasting in Sub-Saharan Africa. arXiv preprint arXiv:1908.00340.