International Journal in IT & Engineering (IJITE)
Volume 10 Issue 10, Oct 2022 ISSN: 2321-1776 Impact Factor: 7.256
Journal Homepage: http://ijmr.net.in, Email: irjmss@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal

DEEP LEARNING TECHNIQUES FOR EFFECTIVE FEATURE RECOGNITION, SELECTION, AND EXTRACTION FROM COMPLICATED REMOTE SENSING DATASETS

Juvvala Bala Ambedkar
(Computer Science & Engineering)
Dr. Ajay Agarwal (Professor)
Glocal School of Technology and Computer Science

ABSTRACT

The development of trustworthy systems that enable many options for "Internet of Things" (IoT) and remote sensing photos has been made possible through the application of machine learning models and middleware characteristics. The usage of remote sensing is particularly necessary to collect geographical data on huge proportions. The primary objective of this research is to conduct a study on deep learning techniques for effective feature recognition, selection, and extraction from complicated remote sensing datasets. According to the findings of the study, the area calculated for the class using the ESRI LULC (Land use/ Land Cover) dataset and the RF (Random Forest) classifier were practically identical. Additionally, the RF classifier has the highest accuracy at 92.17 percent.

Keywords: *Deep Learning Techniques; Remote Sensing Datasets; Internet of Things; Feature Recognition.*

INTRODUCTION

The field of remote sensing is a developing area of research that has a global influence and contributes to an improved understanding of the complex relationship that exists between human activities and shifts in the global environment. Recent decades have seen the development of "remote sensed" (RS) capabilities including hyper spectral scanning and "synthetic apertures radio" (SAR), electro-optical, thermal, "detecting and locating light sources", and other remote sensing (RS) devices have been accumulating enormous volumes of data, which has made it possible for humanity to investigate the cosmos (Ball et al., 2018). In this context, the use Classification and detection of changes are two applications of telemetry of pictures that have had promise. (Ma et al., 2019).

There are many uses for the use of computer vision and deep learning, including picture categorization, the detection of objects during industrial production, the study of medical images, the recognition of actions, and remote sensing (Shi et al., 2022). The process of image classification takes place in stages and begins with the creation of a scheme for the classification of the images that are desired. After that, the images go through a process called pre-processing, which includes image grouping, image enhancement, scaling, and other similar operations. The next part of the procedure involves isolating the sought-after portions of the photographs and creating preliminary clusters. The machine learning technique is next applied to the photos to provide the required categorization. (Karimi Jafarbigloo and Danyali, 2021). The previous literatures that have been published on this subject will be elaborated on in the next part.

LITERATURE REVIEW

AUTHORS AND YEAR	METHODOLOGY	FINDINGS
Huang et al. (2018)	The "five-layer-fifteen-level" (FLFL) satellite remotely sensed data administration framework has been detailed and altered to build a better "four-layer-twelve-level" (FLTL) form for farm far-removed data administration and applications. Big data for agriculture uses cameras on high-resolution (HR) sensors on unmanned helicopters and ground-based infrastructure.	Predicts the potential synchronisation of remotely sensing big data management as well as applications at the neighbourhood, county, and field scale.
Kalantar et al., (2020)	The "Generalised Logical Modelling" (GLM), "Aided Recurrent Designs" (BRT or GBM), and "Random Forests" (RF) are trained shortly after the "Flexible Discriminant Assessment" (FDA) supervised instruction algorithms has trained the LSM methods can be compared against additional frequently utilised computations for an identical for a reason	In comparison to GBM, GLM, and FDA, RF appears to be the most capable when it comes to coping with all of the conditioning elements.
Nasiri et al. (2022)	Surveys recent developments in technology pertaining to big data. Its purpose is to provide assistance in selecting and implementing the appropriate combination of various Big Data technologies in accordance with the technological needs of the organisation and the requirements of the particular applications being used	Classifies several technologies and analyses their features, advantages, limitations, and applications
Tzenios, Reddy, and Bharadiya (2023)	This study uncovers holes in understanding of techniques for deep learning and data from aerial imagery within a particular location, which aids in understanding how vegetation indices and environmental factors affect agricultural productivity.	The usage of the Moderate-ResolutionMulticolor Spectroscopic radiometer in satellite communication, is the most often used type of remote sensing. The results showed that vegetation indicators were the most often used factor for predicting crop yield. This evaluation contrasts all of these methods and discusses their benefits and drawbacks.
Hao et al. (2023)	Using deep learning has become more popular, and to solve this problem, techniques for enhancing data are also being developed. Various solutions have also been proposed. By presenting and reviewing the current state of the science on information enhancement for remotely sensed object recognition, the present piece seeks to close that hole in the literature.	this study addresses the shortcomings of the current approaches and identifies potential lines of inquiry for data augmentation methods in the future.

Table 1: Literature review

According to the past literatures, when it comes to distant sensing, where accurate and transparent reasoning is crucial, developing techniques for interpreting and explaining the decisions made by deep learning models is a research gap that needs attention. So, the main aim of this research is to conduct a study on deep learning techniques for effective feature recognition, selection, and extraction from complicated remote sensing datasets.

METHODOLOGY

Themethodology relies heavily on satellite technologies in which this paper examined. Their cooperation has helped future large-scale real-time mapping and monitoring at multiple spatiotemporal scales. Satellite-related technologies include cloud-computing platforms, machine learning, and deep learning. The categorization methods and methodologies used for mapping and monitoring depend on the satellite sensors, the region of interest with different spatial scales, and the computing systems. Satellite-related technologies used in research applications were discussed in the following sections. Over the past five years, several cloud computing platforms have evolved and advanced, including Google Earth Engine (GEE), which is widely used. This allowed Big Data paradigms to be used for research and management, focusing on data-driven, fast, and cost-effective data access, massive computational resources, and high-end visualisation.

The system used machine learning, deep learning, and cloud technology to incorporate complex remote sensing datasets. Several real-time applications analyse and display classifier data. Google Earth Engine's web-based platform shows its ability to manage large amounts of data and makes it easier for the author to choose tools and classifiers for each application due to its cloud-based nature.

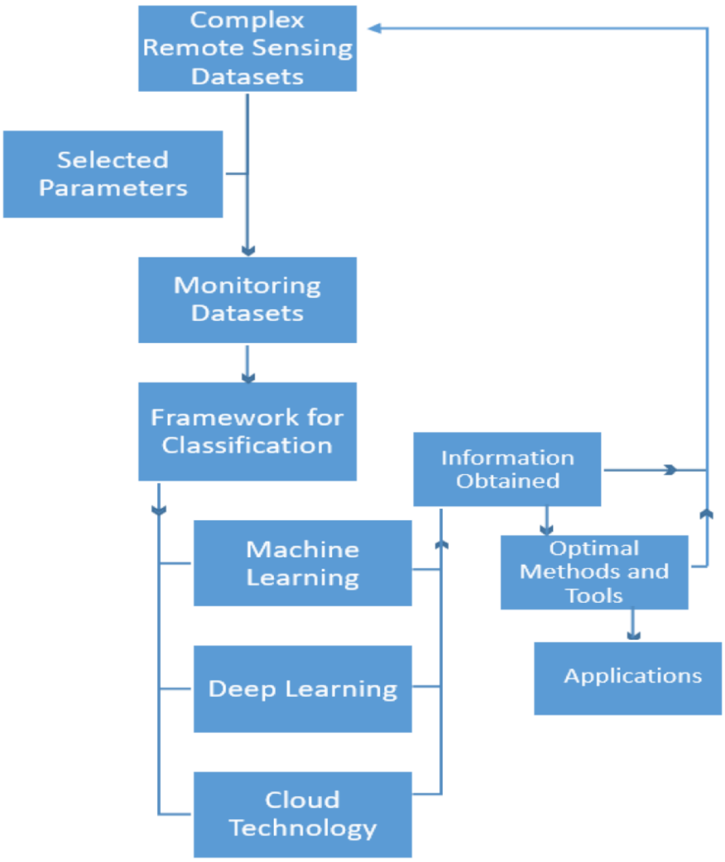

Figure 1: Overall Methodology adopted.

RESULTS AND DISCUSSIONS

This research seeks the best classifiers for each application. This final strategy produced the best classifiers and achieved the goal of comparing machine learning and deep learning. The detection and recognition steps made up machine learning. End-to-end deep learning is common. The time needed to train and test learning algorithms in a physical device. Training and exams are separate because of this. Interpretability is the ability to forecast accurately by following a known sequence. Deep learning hides the prediction process. The study examined processing stages, suggesting a logical approach to reaching the desired result. Computing needs storage and space. The criteria and elemental analysis include computers since it requires operational procedures and hardware components to complete. Table (2) shows how parameters differed between deep learning and machine learning.

Table 2: Assessment of Machine learning techniques and "Deep Neural Network Models"

Parameters	Machine Learning		Deep Neural Network	
	Local Extent	**Regional Extent**	**National Extent**	**Global Extent**
Data Dependence	Works well with medium-resolution datasets	Works well with high resolution datasets with detailed information	Works well by increasing the data characteristics and data processing	Performance depends on enhancing the characteristics of datasets
Processing	Can operate with CPU and software packages	Cannot operate using software packages	GPU is required for managing big data	GPU required for mapping a large area
Computing Platform	ArcGIS, Arc MAP, GEE	GEE, Google Earth	GEE, Google Colab	Keras, TensorFlow
Algorithms Used	CART, SVM, RF	CART, SVM, RF, Otsu, Change Detection	Convolutional neural network (ResNet)	Convolutional Neural Network (ResNet)
Results	Results from SVM and RF had the highest accuracy with minimum training samples.	Remarkable results and accuracy achieved	Quick results are provided with rare training samples as well as obtaining high accuracies.	Results are satisfied as validated using readily available datasets
Interpretability	Easy and quick to understand and operate as well as a user-friendly interface	Can predict and easily applied and combined with other datasets to conclude with immediate decision making	Need to learn linear and non-linear features for deep analysis and prediction	Deep analysis to a global extent is only possible by the easy working of the neurons

Image processing works on this scale are tough to accomplish and demand a powerful computer environment. Classification studies (Hao et al., 2023; Huang et al., 2018; Ball et al., 2018) on vast areas are usually necessary, and in order to complete these jobs, the environment must be powerful. GEE made it easy to process multi-temporal Landsat-8 image series and classify vast study regions. GEE provided a variety of machine learning classifiers such as RF, CART, and SVM for its customers. The sole drawback that can be attributed to GEE in relation to the present investigation is the absence of instruments that can be used to carry out geostatistical sampling processes and statistical calculations.

The GEE built-in machine classifiers classify the study area's LULC using machine-learning training samples. Classifiers are trained using these samples. The same research region uses CART, SVM, and RF for LULC classification, all of which are interesting classifiers. These methodologies are integrated into GEE for data processing. The best classifier for localised mapping is found using these classifications. The region is divided between built-up regions, grasslands, croplands, and woods and trees. Final category: barren ground. This reclassification combines forests and grasslands into one vegetation category. The LULC is built using CART, RF, and SVM supervised classification. The figure below shows that CART performed moderately compared to the other two classifiers.

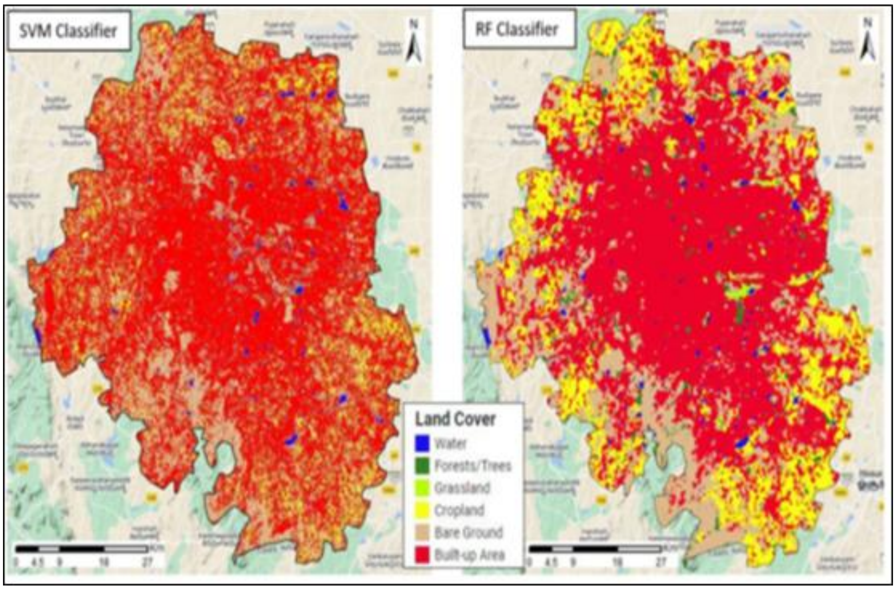

Figure 2: LULC map of Bangalore generated using SVM, and RF classifier.

The maps illustrated that the majority of the territory is controlled by the urban class, while the countryside served as cropland on the region's periphery. The area that is covered by vegetation and crops is quite little within the main city, but as walk further out from the main city centre, the size of the area covered by vegetation and crops rises. The pattern repeats itself across the entirety of the territory; hence, it is obvious by looking at the maps that the urban area has expanded over the course of time. Taking this into consideration, the area of significant LULC classifications, such as urban, vegetation, and agricultural, is estimated. The area that was obtained using the RF classifier was almost an exact match for the area that was estimated for the class using the ESRI LULC dataset. up addition, the overall accuracy attained by the RF classifier is the best possible, coming up at 92.17 percent as given in table below.

Table 3: Results obtained from varied classifiers and accuracy assessment.

LULC area/ Classifier	Landsat-8			Sentinel-2
	CART	SVM	RF	ESRI Dataset
Urban (Km^2)	663	696	723	738
Vegetation (Km^2)	34	33	38	42
Cropland (Km^2)	1493	73	1283	1224
Kappa Coefficient	0.67	0.79	0.92	0.89
Overall Accuracy	84.66	90.14	92.17	94.18

CONCLUSION

The evaluation of accuracy is helpful in identifying the effectiveness of a variety of classifiers as well as the impact of the underlying training sampling schemes. The test samples are randomly selected from each specified subclass across the information set using an example map. By doing this, it is made guaranteed that the specimens being tested do not coincide with the initial data set, which was produced using a variety of sampling techniques. Similar test specimens were utilised in experiments utilising the stratified random selection methods for a specified number of samples of the preparation data. This allowed the various classifiers to be evaluated.

REFERENCES

1. Ball, J.E., Anderson, D.T. and Chan, C.S. (2018) 'Special section guest editorial: Feature and deep learning in remote sensing applications', *Journal of Applied Remote Sensing*, 11(04), p. 1. doi:10.1117/1.jrs.11.042601.
2. Bharadiya, J.P., Tzenios, N.T. and Reddy, M. (2023) 'Forecasting of crop yield using remote sensing data, agrarian factors and machine learning approaches', *Journal of Engineering Research and Reports*, 24(12), pp. 29–44. doi:10.9734/jerr/2023/v24i12858.
3. Hao, X. *et al.* (2023) 'A review of data augmentation methods of Remote Sensing Image Target Recognition', *Remote Sensing*, 15(3), p. 827. doi:10.3390/rs15030827.
4. Huang, Y. *et al.* (2018) 'Agricultural Remote Sensing Big Data: Management and Applications', *Journal of Integrative Agriculture*, 17(9), pp. 1915–1931. doi:10.1016/s2095-3119(17)61859-8.
5. Kalantar, B. *et al.* (2020) 'Landslide susceptibility mapping: Machine and ensemble learning based on Remote Sensing Big Data', *Remote Sensing*, 12(11), p. 1737. doi:10.3390/rs12111737.
6. Karimi Jafarbigloo, S. and Danyali, H. (2021) 'Nuclear atypia grading in breast cancer histopathological images based on CNN feature extraction and LSTM classification', *CAAI Transactions on Intelligence Technology*, 6(4), pp. 426–439. doi:10.1049/cit2.12061.
7. Ma, L. *et al.* (2019) 'Deep learning in remote sensing applications: A Meta-analysis and Review', *ISPRS Journal of Photogrammetry and Remote Sensing*, 152, pp. 166–177. doi:10.1016/j.isprsjprs.2019.04.015.
8. Nasiri, V. *et al.* (2022) 'Modeling forest canopy cover: A synergistic use of sentinel-2, aerial photogrammetry data, and machine learning', *Remote Sensing*, 14(6), p. 1453. doi:10.3390/rs14061453.
9. Shi, C. *et al.* (2022) 'Remote sensing scene image classification based on self-compensating convolution neural network', *Remote Sensing*, 14(3), p. 545. doi:10.3390/rs14030545.