Object Detection in High resolution using Satellite Imagery with Deep Learning

Niharika Goswami[1], Keyurkumar Kathiriya[2], Santosh Yadav[3], Janki Bhatt[4], Dr. Sheshang Degadwala[5]

[1,2,3,4] U.G. Scholar, Sigma Institute of Engineering, Vadodara, Gujarat, India

[5] Associate Professor, Sigma Institute of Engineering, Vadodara, Gujarat, India

ABSTRACT

For several years, detecting objects in satellite imagery has been a difficult task. Higher accuracies in the identification of different artefacts from very high-resolution satellite images have been achieved thanks to the creation of successful deep learning algorithms and advancements in hardware systems. Satellite imaging has been successfully used for weather forecasts, spatial and geological purposes over the last few decades. For these types of applications, low-resolution satellite images are adequate. However, technical advancements in satellite imaging include high resolution cameras, which broadens its area of use. As a result, High-Resolution Satellite Imagery (HRSI) proved to be a viable complement to aerial photogrammetric data in terms of providing a new data base for object detection. Vehicle tracking from satellite data would be a safer option for automating such networks when traffic volumes in developed countries skyrocket. This study summarises recent advances in the field of deep learning-based object recognition from satellite imagery. In this article, a U-net model was used to classify satellite images into five distinct categories.

Keywords : Satellite Image, Single Shot Detector, You Only Look Once (YOLO), Conditional Random Fields (CRF) and Fully Convolutional Networks (FCN)

I. INTRODUCTION

One of the most effective and significant instruments used by meteorologists is satellite imagery. They are literally the sky's eyes. These photographs reassure forecasters about atmospheric activity because they offer a simple, succinct, and precise depiction of how conditions are unfolding. Without satellites, forecasting the weather and doing experiments will be incredibly difficult. The representations of atmospheric motion provided by data collected at stations around the country are small. It is still possible to get a decent interpretation from the results, but since the stations are hundreds of miles apart, major features could be overlooked. Satellite photographs tend to illustrate what cannot be measured or observed. Furthermore, satellite images are recognised as evidence. There is no space for mistake. Satellite imagery provides evidence that can be viewed "on the spot." Satellite images provide a

clear representation of what is happening around the world, particularly over oceans where data gaps exist. Data can only be collected at specific locations around the planet, but forecasting would be almost as impossible without it as it would be without satellites. It is important to have both. The combination of the two provides a much greater view of how the weather is operating and vastly increases forecasting accuracy. Satellite photographs come in a number of formats. The visible, ultraviolet, and water vapour photographs are the most useful to meteorologists. Many of these images can be captured by a single satellite in orbit.

Satellite imagery has also been used in international development and other fields. Examples include the study of Land Use and Land Cover (LULC) shift identification to locate illicit deforestation or to warn agricultural projects; the use of satellite imaging to study refugee settlements; and its analysis to assist in relief operations during natural disasters such as floods or fires. Another one of such impressive applications was published two years ago in a paper by authors from Stanford's Sustainability and Artificial Intelligence Lab. The research incorporated deep learning, day satellite imagery, night satellite imagery, and socioeconomic conventional surveys to provide a solution for predicting poverty estimates in various locations around Africa. Even with its shortcomings, new approaches to poverty calculation are still needed in the battle for the first UN Sustainable Development Goal, the abolition of poverty in all its forms, anywhere. The analysis is a fantastic example of how satellite imagery and artificial intelligence should work together.

II. RELATED WORKS

Nyan Linn Tun, Alexander Gavrilov, and Naing Min Tun used FCNN to differentiate satellite images in [1]. The proposed model achieves high precision separation in very short periods of time without using a lot of computing power. They have a 0.70 percent accuracy breakdown. Many image classification technologies have been developed, and classification effects will achieve substantial changes in both classification categories. They built an uncertainty matrix in a comparison database (with 21 classes) and achieved an accuracy of 70.48 percent. Using in-depth learning technologies, the satellite image classification has been created to compare different main regions of the globe.

Yohei Koga, Hiroyuki Miyazaki, and Ryosuke Shibasaki suggested an unregulated domain modification (DA) solution to the performance degradation caused by image aspect discrepancies between data domains in [3]. They increased the accuracy of receiving more than 10% of the desired domain by using the correlation of Correlation (CORAL) DA and the opposition DA on a regional-based vehicle detector. Their proposed approach outperformed the precision obtained from labelling data with the goal domain name. The DA approach is useful even when a well-labelled database does not reach the optimal optimum precision due to the difficulty of the target region, such as the target area, such as the complexity of the image element.

Rodrigo F. Berriel, Andre Teixeira Lopes, Alberto F. de Souza, and Thiago Oliveira-Santos suggested a classification scheme for big zebra satellite satellites in [4]. Using the Google Static Maps API, Google Maps Directions API, and OpenStreetMap, the app senses photographs of intersected and unpaved roads all over the world. The experiment was carried out in this novel dataset, which included 245,768 images from three continents, nine countries, and more than 20 towns. The test results verified the proposed system's power, with 97.11 percent accuracy in the global survey.

In [7] Mark Pritt and Gary Chern introduced an in-depth reading curriculum that identifies artefacts and

properties in high-definition satellite imagery. CNN's integration of post-processing neural networks, which provide forecasts from CNN and satellite metadata, is included in the software. The machine achieves an accuracy of 0.83 and an F1 rating of 0.797 in the IARPA fMoW database of one million images in 63 groups, including the false detection process. In the fMoW TopCoder challenge, it separates 15 classes with 95% or better accuracy and outperforms the John Hopkins APL model by 4.3 percent. Monitoring a satellite imagery store may assist law enforcement in detecting illicit mining activity or illegal fishing vessels, assist disaster response teams by mapping mud slides or storm damage, and empowering investors to monitor crop growth or oil resource development more effectively.

The method developed by [9] Vladimir Iglovikov, Sergey Mushinskiy, and Vladimir Osin is focused on an adaptation of a completely convolutional neural network for multispectral data processing. Their method consists of several stages, including the application of a completely convolutional network to multispectral satellite images with a shared training goal and the study of boundary effects and reflectance indices. Its precision is comparable to the first two solutions, but unlike those, it does not rely on sophisticated assembly techniques and therefore can be conveniently scaled for manufacturing implementation as part of automated feature labelling schemes for satellite imagery research.

In [11], Qiling Jiang, Liujuan Cao, Ming Cheng, Cheng Wang, and Jonathan Li suggest a Deep Convolutional Neural Network-based vehicle detection system in satellite images (DNN). When working with videos, DNN is a deep learning model with a high learning ability. A Deep Neuro Network (DNN)-based classifier is used to assess whether or not the target super pixel is a car. Their experiment reveals that DNN outperforms other methods in terms of efficiency. They will look at the possibilities of

transferring the qualified detector through resolutions, which is a common-sense problem in vehicle detection in satellite images.

In [12] presents a method for detecting buildings in multispectral satellite images developed by Tomohiro Ishii, Edgar Simo-Serra, Satoshi Iizuka, Yoshihiko Mochizuki, Akihiro Sugimoto, Hiroshi Ishikawa, and Ryosuke Nakamura. They adapt the classification network to identification by translating the network's fully-connected layers to convolutional layers, allowing the network to process images of any resolution. They offer an in-depth analysis of the seven separate spectral bands given by satellite images and demonstrate how important it is to combine them to achieve successful results. In [13], Ahmad Mansour, Ahmed Hassan, Wissam Hussein, and Ehab Said used two cutting-edge algorithms for target detection (Faster RCNN and SSD). They create vehicle datasets from Google Earth and other satellite samples including the JF-2 and WORLD-VIEW satellites. Mean average precision (MAP) is an output metric. According to the findings, faster R-CNN Inception-V2 outperforms SSD Inception-V2. However, the SSD Inception-V2 detects images in less time. The research will be expanded to include general vehicle identification (bicycle, motorcycle, bus, truck).

In [14] Y Harold Robinson, S Vimal, Manju Khari, Fernando Carlos Lopez Hernandez, and Ruben Gonzalez Crespo explain and compare various object classifiers in terms of accuracy and execution time. During the segmentation and classification measures, the image spectrum is used to minimise processing costs. To begin, objects in satellite images are removed to aid in the feature extraction process. They analysed and commented on the feasibility of the suggested strategies. Experiments show that our proposed tree-based CNN outperforms the other strategies in terms of classification efficiency and execution time. The enhanced knowledge can make use of deeper neural networks and improve

performance by incorporating neural network principles.
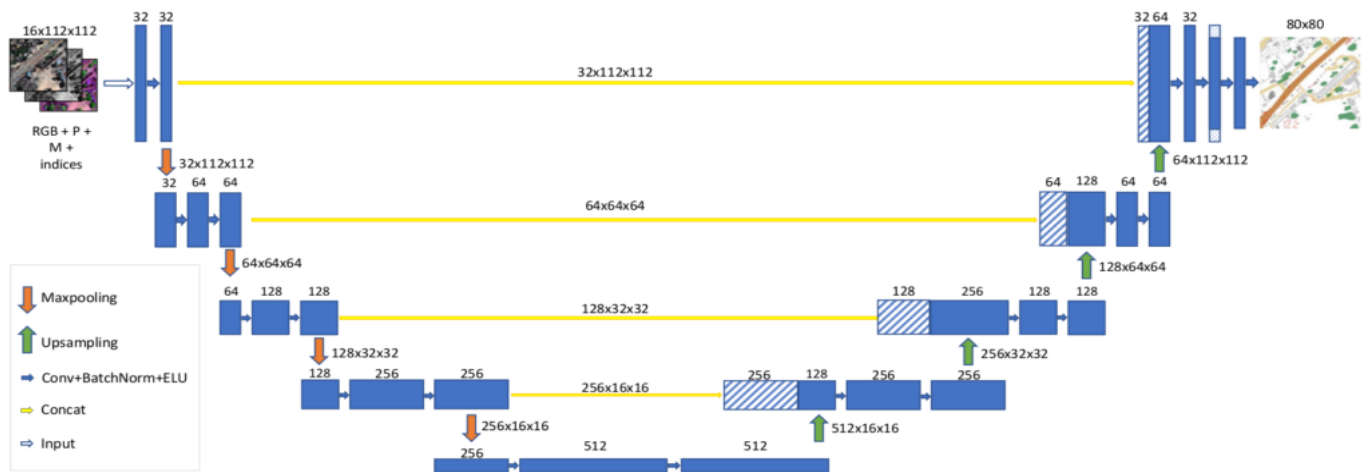
III. PROPOSED METHODOLOGY

FIGURE 1: U-NET ARCHITECTURE

The architecture resembles a letter 'U,' thereby justifying its name. This architecture is divided into three sections: contraction, bottleneck, and expansion. The contraction part is comprised of several contraction blocks. Each block starts with an input and then applies two 3X3 convolution layers followed by a 2X2 max pooling. After each block, the number of kernels or function maps doubles, allowing architecture to learn complex structures effectively. The bottommost layer acts as a bridge between the contraction and expansion layers. It employs two 3X3 CNN layers, which are joined by a 2X2 up convolution layer. The heart of this design, however, is in the expansion section. It, like the contraction sheet, is made up of multiple expansion blocks. Each block sends the input to two 3X3 CNN layers, which are then accompanied by a 2X2 up sampling layer. In order to preserve symmetry, the number of feature maps used by the convolutional layer is reduced by half after each block. However, each time the input is appended, function maps from the corresponding contraction layer are also appended. This intervention ensures that the features discovered when contracting the image are used to restore it. The number of contraction blocks equals the number of expansion blocks. The resulting mapping is then passed into another 3X3 CNN layer with the number of function maps equal to the number of Segment wanted.

IV. RESULTS AND ANALYSIS

In this part, we will look at a Keras-based implementation of a deep U-Net for satellite image segmentation. The dataset is made up of 8-band commercial grade satellite imagery from the SpaceNet dataset. Each of the 24 positions is represented by a few tiff files in the Train Array. Each position has an 8-channel image with spectral information from several wavelength channels (red, red edge, coastal, blue, green, yellow, near-IR1 and near-IR2). These files can be found in the data/mband/ directory. Mask photographs, which are accurately segmented images of each training area, are also available. These files contain data on five distinct types of objects: houses, bridges, plants, crops, and water (note that original Kaggle contest had 10 classes). Satellite photographs

have a resolution of 16 bits. Mask-files, on the other hand, are 8-bit.
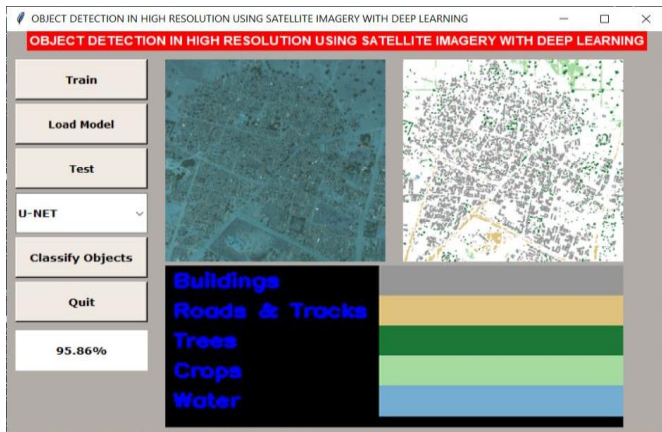

FIGURE 2: METRO CITY VIEW

As shown in the figure 2 satellite image of Metro city view is segmented by proposed U-net architecture and give the segmented view for easy understand.
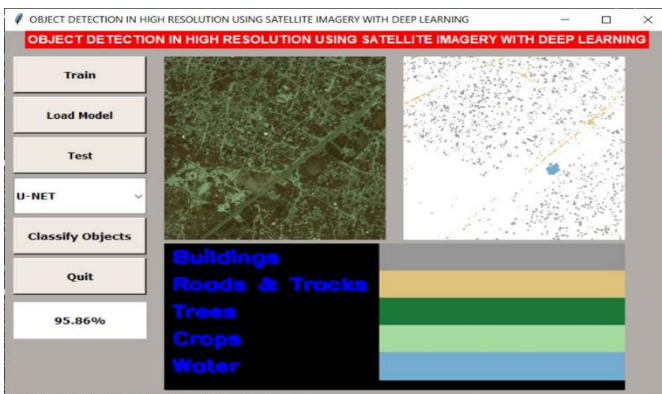

FIGURE 3: WATER RESOURCE IN CITY AREA

As shown in the figure 3 satellite image of Water Resource in City Area view is segmented by proposed U-net architecture and give the segmented view for easy understand.
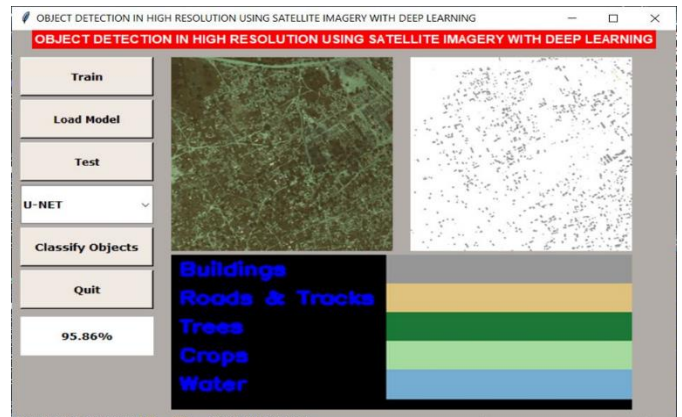

FIGURE 4: BUILDING AREA

As shown in the figure 4 satellite image of building view is segmented by proposed U-net architecture and give the segmented view for easy understand.
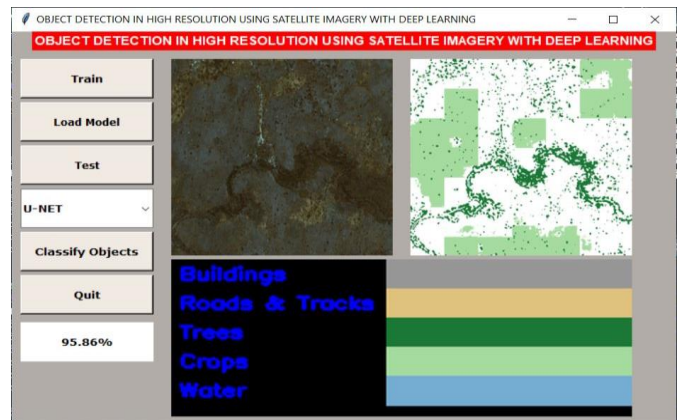

FIGURE 5: FOREST LAND

As shown in the figure 4 satellite image of Forest view is segmented by proposed U-net architecture and give the segmented view for easy understand.
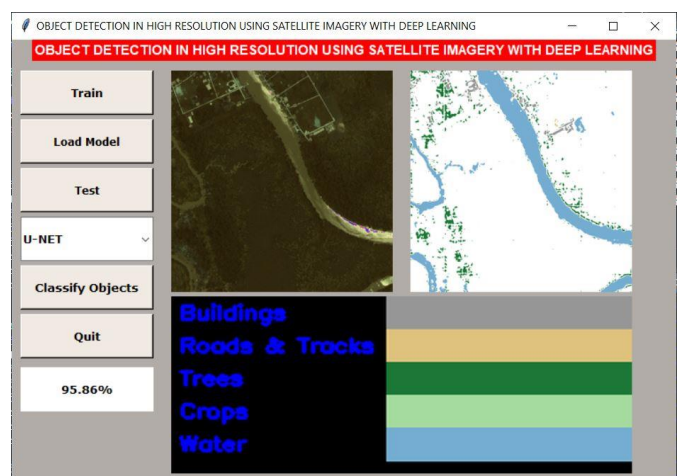

FIGURE 6: RIVER VIEW

As shown in the figure 4 satellite image of River view is segmented by proposed U-net architecture and give the segmented view for easy understand.
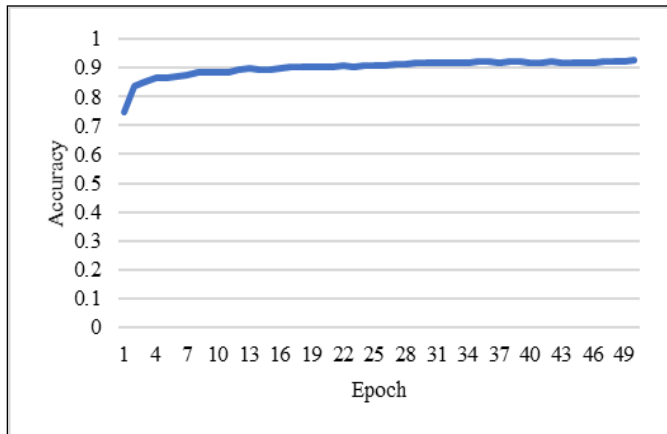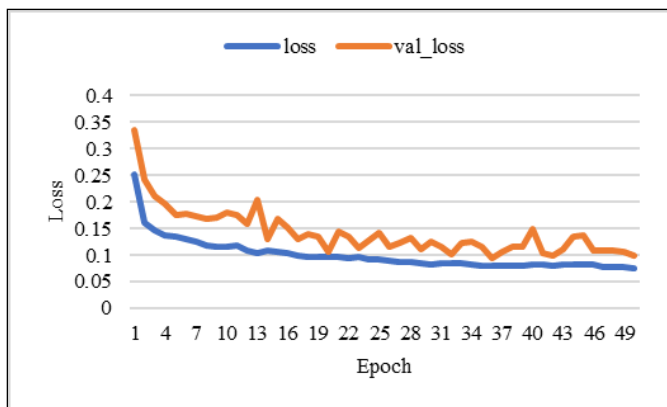

FIGURE 7: ACCURACY PLOT

FIGURE 7: MODEL LOSS

As shown in the figure 7 and 8 U-net model training parameters for 49 Epoches. As no of epoch increse the loss decrease and accuracy increses.

V. CONCLUSION

The satellite designation of imaginary objects is needed for traffic monitoring, land classification, and military applications. The previous approach was based on classifying only one or two of the three land types. It would not include all of the object's detail. In this study, the deep learning system U-net is used to classify objects. The key concept is to complement a traditional contracting network with successive layers that substitute pooling operators with up sampling operators. As a result, these layers maximise the output's resolution. It combines the position information from the down-sampling path with the contextual information from the up-sampling path to produce a general information mixing localization and context, which is needed to predict a good segmentation map based on object prediction in satellite images. The proposed U-net model seems to have an overall accuracy of 92.42 percent for five class objects.

VI. REFERENCES

[1]. N. L. Tun, A. Gavrilov, and N. M. Tun, "Multi-classification of satellite imagery using fully convolutional neural network," Proc. - 2020 Int. Conf. Ind. Eng. Appl. Manuf. ICIEAM 2020, pp. 7–11, 2020, doi: 10.1109/ICIEAM48468.2020.9111928.

[2]. A. Van Etten, "Satellite imagery multiscale rapid detection with windowed networks," Proc. - 2019 IEEE Winter Conf. Appl. Comput. Vision, WACV 2019, pp. 735–743, 2019,

[3]. A. Van Etten, "You Only Look Twice: Rapid Multi-Scale Object Detection in Satellite Imagery," 2018.

[4]. Y. Koga, H. Miyazaki, and R. Shibasaki, "Correction: A Method for Vehicle Detection in High-Resolution Satellite Images That Uses a Region-Based Object Detector and Unsupervised Domain Adaptation. [Remote Sensing 2020, 12, 575] doi: 10.3390/rs12030575," Remote Sens., vol. 12, no. 7,

[5]. R. F. Berriel, A. T. Lopes, A. F. De Souza, and T. Oliveira-Santos, "Deep Learning-Based Large-Scale Automatic Satellite Crosswalk Classification," IEEE Geosci. Remote Sens. Lett., vol. 14, no. 9, pp. 1513–1517, 2017, doi: 10.1109/LGRS.2017.2719863.

[6]. T. Ophoff, S. Puttemans, V. Kalogirou, J. P. Robin, and T. Goedemé, "Vehicle and vessel detection on satellite imagery: A comparative study on single-shot detectors," Remote Sens., vol. 12, no. 7, pp. 1–21, 2020, doi: 10.3390/rs12071217.

[7]. J. Yuan, "Automatic Building Extraction in Aerial Scenes Using Convolutional Networks," 2016, [Online]. Available: http://arxiv.org/abs/1602.06564.

[8]. M. Pritt and G. Chern, "Satellite image classification with deep learning," Proc. - Appl. Imag. Pattern Recognit. Work., vol. 2017-October, pp. 1–7, 2018, doi: 10.1109/AIPR.2017.8457969.

[9]. "Three Applications Of Deep Learning Algorithms For Object Detection Milena Napiorkowska (1), David Petit (1), Paula Martí (2) (1) Deimos Space UK Ltd ., (2) Deimos Engenharia," no. 1, pp. 4839–4842, 2018.

[10]. V. Iglovikov, S. Mushinskiy, and V. Osin, "Satellite Imagery Feature Detection using Deep Convolutional Neural Network: A Kaggle Competition," 2017, [Online]. Available: http://arxiv.org/abs/1706.06169.

[11]. E. Kalinicheva, J. Sublime, and M. Trocan, "Object-Based Change Detection in Satellite Images Combined with Neural Network Autoencoder Feature Extraction," 2019 9th Int. Conf. Image Process. Theory, Tools Appl. IPTA 2019, pp. 1–6, 2019, doi: 10.1109/IPTA.2019.8936085.

[12]. C. Wang, Q. Jiang, M. Cheng, J. Li, and L. Cao, "Deep Neural Networks-Based Vehicle Detection In Satellite Images Fujian Key Laboratory of Sensing and Computing for Smart City School of Information Science and Engineering, Xiamen University Xiamen, China," IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work., 2016.

[13]. T. Ishii et al., "Detection by classification of buildings in multispectral satellite imagery," Proc. - Int. Conf. Pattern Recognit., vol. 0, pp. 3344–3349, 2016, doi: 10.1109/ICPR.2016.7900150.

[14]. A. Mansour, A. Hassan, W. M. Hussein, and E. Said, "Automated vehicle detection in satellite images using deep learning," IOP Conf. Ser. Mater. Sci. Eng., vol. 610, no. 1, 2019, doi: 10.1088/1757-899X/610/1/012027.

[15]. G. Cheng, J. Han, and X. Lu, "Remote Sensing Image Scene Classification: Benchmark and State of the Art," Proc. IEEE, vol. 105, no. 10, pp. 1865–1883, 2017, doi: 10.1109/JPROC.2017.2675998.

[16]. Y. H. Robinson, S. Vimal, M. Khari, F. C. L. Hernández, and R. G. Crespo, "Tree-based convolutional neural networks for object classification in segmented satellite images," Int. J. High Perform. Comput. Appl., 2020, doi: 10.1177/1094342020945026.

[17]. N. Imamoglu, P. Martínez-Gómez, R. Hamaguchi, K. Sakurada, and R. Nakamura, "Exploring recurrent and feedback CNNs for multi-spectral satellite image classification," Procedia Comput. Sci., vol. 140, pp. 162–169, 2018, doi: 10.1016/j.procs.2018.10.325.

[18]. A. Hosny and A. Parziale, "A Study on Deep Learning," vol. 9, no. 4, pp. 21482–21483, 2019.

[19]. D. Dai and W. Yang, "Satellite image classification via two-layer sparse coding with biased image representation," IEEE Geosci. Remote Sens. Lett., vol. 8, no. 1, pp. 173–176, 2011, doi: 10.1109/LGRS.2010.2055033.

[20]. J. Han, D. Zhang, G. Cheng, L. Guo, and J. Ren, "Han_etal_IEEE_TGRS_2015_Object_detection _in_optical_remote_sensing_images_based_wea kly.pdf," pp. 1–26.

[21]. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect.

Notes Bioinformatics), vol. 9351, pp. 234–241, 2015, doi: 10.1007/978-3-319-24574-4_28.

[22]. P. Helber, B. Bischke, A. Dengel, and D. Borth, "Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 12, no. 7, pp. 2217–2226, 2019, doi: 10.1109/JSTARS.2019.2918242.

Cite this article as :