ISSN: 2366-1313

DETECTION AND CLASSIFICATION OF OBJECTS IN SATELLITE IMAGES USING CUSTOM CNN

[1]peddi Srinitha, [2]Dr.B. SATEESH KUMAR

[1] M. Tech Scholar, [2] Professor, Department Of CSE

JNTUH UNIVERSITY COLLEGE OF ENGINEERING, JAGITAL, T.S., INDIA

Abstract:

An rising number of uses for satellite image analysis include surveillance, military operations, geospatial research, and monitoring environmental effects and climate change. The ability to automatically identify and categorize objects is a key component of satellite image analysis. It becomes difficult to identify and categorize things in aerial pictures due to their nature, size, and variety of visual attributes. Due to the structure of these photographs and the data they contain, manually detecting objects in them takes a lot of time. Automating the identification of different characteristics or objects from these satellite photos is desired. The traditional techniques for classifying things comprise two steps: (i) locating the areas in the picture where items are present, and (ii) classifying the objects in those regions. The complexity of the backdrop, size, noise, and distance characteristics makes it more difficult to recognize things. This study suggests using a specifically designed convolutional neural network to identify and categorize three distinct things in the photos, including trees, buildings, and autos. Additionally, it seeks to comprehend and succinctly describe the performance characteristics of the hypothetical custom CNN.

Keywords— Satellite imagery, object detection, classification, custom, convolutional neural networks, image processing

I. INTRODUCTION

In various applications, including remote surveillance, monitoring of the environment, aerial survey, etc., satellite imagery is becoming more important. All of these apps use satellite pictures to search for items, events of interest, facilities, etc. In most applications, manually identifying objects and classifying them becomes exceedingly challenging, particularly when processing enormous amounts of data and a high number of satellite photos at once. Although the recognition and categorization of objects in pictures has been extensively explored in the field of image processing, the detection of things in satellite (aerial) photographs is more difficult due to the tiny size of the objects and the difficulty in tracking and capturing their visual characteristics. To this goal, a number of automated detection and classification algorithms have been suggested and are currently being developed. For object identification and categorization in satellite photos, several algorithms have been put forward, ranging from traditional machine learning (ML) to modern deep learning. Machine learning techniques for identifying and categorizing data have received the greatest attention of these during the last several decades. These techniques include taking different attributes from the photos and applying ML classifiers to categorize them. Because the size, direction, and backdrop of the target item might vary, automated object recognition is still difficult. The problems with automated object recognition cannot be solved by using traditional machine learning classifiers that manually choose features like HOG, Gabor, Hough transform, wavelet coefficients, etc. As a consequence, an effective strategy is required, and Deep Learning has demonstrated promising results when used to accomplish the goal of detection and classification utilizing CNN. Deep learning classification techniques that have recently been presented for highly accurate automated object recognition are able to automatically learn features from the photos rather than relying on human feature selection. Convolutional neural network (CNN)-based deep learning models are widely advocated for the identification and categorization of objects in satellite imagery. There are two stages in these models. The zones of item presence in the picture are found in the first stage. Using a convolutional neural network, the items are categorised in the second stage.

This study suggests using a tailored convolutional neural network to identify and categorize items in satellite photos. The algorithm used is trained to identify three things in satellite images: trees, buildings, and automobiles. The effectiveness of the YOLO V3 algorithm's real execution on the same dataset as well as other common benchmark information for different algorithms without actual execution are compared to the performance of detection and classification. Instead of doing so in two passes as is the case with traditional CNN models, the

ISSN: 2366-1313

YOLOV3 algorithm integrates the detection and categorization of objects in one step.

II. RELATED WORKS

This section surveys the current deep learning models for object identification and categorization. The authors of [1] have suggested a variant that use a CNN to detect objects. They proposed the idea of a rotation-invariant region-based convolutional neural network particularly for satellite photos. Prior to the categorization of the objects, the feature representation is normalized in order to accomplish and concentrate on the idea of an area (rotation invariant). Following that, classification is performed, which is now dependent on the fact that each picture has a more sophisticated calculation for patching results over rotation-invariant areas.

The authors of [2] provide a technique for identifying major crossings and bridges on waterways that are visible in satellite imagery. Recursive scanning, a method that applies geometric constraints to identify such features as the fact that the water body in question is a river, is a technique that must first be used to detect or identify the water bodies in the picture, which are mostly rivers. Following the use of the rivers that were identified in the first step, and based on the application of knowledge relating to the spatial dimensions of various bridges, a scan is carried out over the area of the rivers that were identified in order to further detect or identify the pixels that may belong to a bridge. After the pixels have been located, their connection and relationship are analyzed, and based on the results of that analysis, it is decided if the picture has a bridge segment for the pixels that were located initially. However, there is a drawback to this kind of method, namely the need that we have previous information and awareness of the three-dimensional nature of the objects or constructions, in this instance bridges, that we are seeking to detect.

In [3], authors describe a two-stage method for identifying road networks that may be seen or are present in aerial photographs such as those obtained by satellite or UAV. Through the use of the cutting down or pruning step after the detection stage, it results in automated detection. By identifying these areas or shapes at this step, a Bayesian model is then utilized to classify regions that have comparable or homogenous characteristics. The second step involves using a method known as conditional

ISSN: 2366-1313

probability to determine the chance that any specific chunk or segment is a road. This detection approach, like many others, has a very high computational complexity. The authors of [4] suggest an object-based image analysis method that may be used to identify the land cover and is mostly used to categorize the topography of various lengths of land in satellite pictures. Before an SVM classifier can be trained, texture characteristics must be extracted, which requires segmenting each item in the picture. These segmented objects are utilized to perform the extraction operation when this is finished. The textural properties of the segmented objects are then retrieved and utilized to train an SVM classifier in order to categorize the land based on its cover and use. The precision of this technique is controlled by feature extraction, which is often limited and not full, complete, or accurate in and of itself.

III. EVALUATION OF EXISTING WORK

In the fields of AI and Deep Learning, there are several implementations that may be used for object detection. But each of those approaches has pros and cons of its own. The most widely used implementations, including Faster R-CNN, ResNet, and YOLO V3, have been chosen for this section. They will also be judged according to performance measures, such as accuracy, precision, recall, and F1 score. The similarities stated above are shown in Table 1.

TABLE 1: PERFORMANCE EVALUATION METRICS OF EXISTING SYSTEMS

Implementation	**Accuracy**	**Precision**	**Recall**	**F1 score**
YOLO V3	91.36	0.92	0.91	0.90
ResNet	86.21	0.89	0.90	0.88
Faster R-CNN	84.26	0.84	0.87	0.87

IV. PROPOSED METHODOLOGY

A labeling technique is used to categorize the pictures of buildings, trees, and automobiles that were extracted from the 3D satellite imaging collection. The specialized convolutional neural network receives its input from the annotated pictures. Figure 1 depicts the architecture and layer breakdown of the customised convolutional neural network.

ISSN: 2366-1313

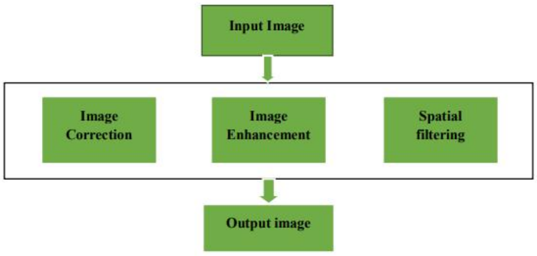

Fig 1 : Image Preprocessing

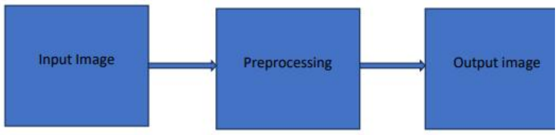

Fig 2 :Image Enhancement

Convolution is a basic mathematical technique that is used in many different image processing operations. It is a simple mathematical procedure, yet one that is crucial to the field of image processing. Through the process of convolution, it is possible to combine two arrays of equal dimensions but different sizes to produce a third array with the same number of dimensions. In order to create operators that accept certain pixel values as input and then deliver as output pixels that are basically a simple linear arrangement of these input pixels, this is often a required and significant component that may be employed in image processing. A grayscale picture is often represented in the setting of processing images as an array of inputs containing numbers, and thus represents only one of those input arrays. The kernel, a second array that is often lower in magnitude or size, has two dimensions and might very well be the same thickness as a single pixel.

Image Acquisition and Data Preprocessing

Acquiring the picture dataset is essential for the growth of any sort of detection-based system. Some guidelines that may aid in prediction and proper training of the model of convolutional neural networks must be followed while gathering datasets. In general, a dataset with more data will provide better results. The information gathered from the VALID website is entirely unprocessed and uncooked. A 3D model of the same satellite imaging data was utilized in this study to show the proof of concept since real-world spacecraft imagery data is not available for public usage.

Preparation of Data and Object Image Labelling

Data preparation is the act of gathering and applying domain expertise to the dataset's pictures in order to sort and iterate over them. This section and the ones that follow it go into great depth on the processes that go into processing raw data. Following data preparation, the enormous quantity of picture data gathered from the 3D satellite imaging collection is converted and cleaned to create a trainable dataset.

CNN Construction and Training

The Test set is the subset used to test the trained model, while the Training phase is the subset used to train a model. The process of creating a neural network architecture specifically for this use case is known as CNN building. In this manner, we may provide the target attribute, or the right response, as the input to the CNN training data. A special neural network created for satellite pictures is used in this study.

CNN Validation, Analysis of Results

Validation is one of the most crucial steps in comparing the model to the input set of data. The test data is utilized to compare the actual results with those anticipated by CNN's expected output prediction after the training phase.

V. RESULTS

At this point, the proposed system may be evaluated using a small sample of random photographs from the test directory of files. This will allow it to recognize and detect items like cars, trees, and buildings in the images, as well as construct boundaries around them to make identification easier.

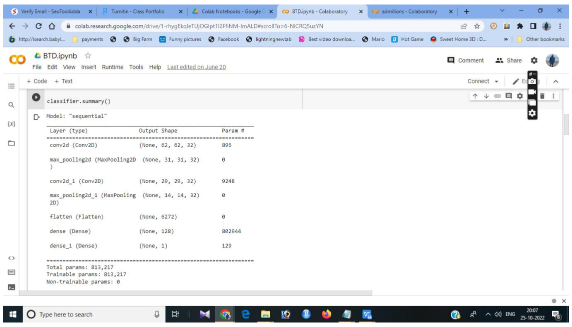

Fig 3 : Model Systematization

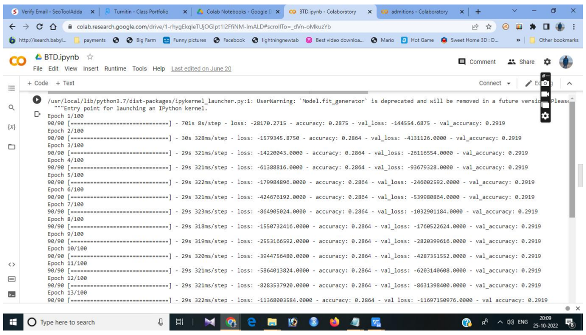

Fig 4 : Model training

The Google Collab or CLI is used for detection in a Python virtual environment. The system's object detection phase is shown with 50 and 100 epochs, respectively. We can see that the system was trained for a total of 100 epochs, which produced a much higher accuracy. This accuracy was also used to make the final prediction. The model with the most recent training weights with the best accuracy and precision was evaluated using measures from the Tensor-Board library, and the ultimate accuracy reached was 94.65%. Here, the tallest mountain will be taken into account.

VI. CONCLUSION

ISSN: 2366-1313

In this study, a unique CNN model for the identification and categorization of stuff in satellite pictures was suggested. Three distinct objects—a vehicle, a structure, and a tree—were used to assess the effectiveness of the suggested approach. The strategy was successful in achieving 94.65% accuracy. Although just a limited number of training images were utilized in this study, we want to evaluate the model's effectiveness with many more datasets in the future. Real-time satellite images may be utilized to train the suggested neural network (with the necessary permits granted by the relevant space agency), which is a future advancement that can be used to this study to further it. The prototype system will be prepared for deployment in a live environment in this fashion.

REFERENCES.

[1] Cheng, G.; Zhou, P.; Han, J. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. IEEE Trans. Geosci. Remote Sens. 2016.

[2] Chaudhuri, D., Samal, A., An automatic bridge detection technique for multispectral images. IEEE Trans. Geosci. Remote Sens. 2008

[3] Hu, J., Razdan, A., Femiani, J.C., Cui, M., Wonka, P., Road network extraction and intersection detection from aerial images by tracking road footprints. IEEE Trans. Geosci. Remote Sens., 2007

[4] Goodin, D.G., Anibas, K.L., Bezymennyi, M., 2015. Mapping land cover and land use from object-based classification: an example from a complex agricultural landscape. Int. J. Remote Sens. 36, 4702-4723.

[5] Han, J., Zhang, D., Cheng, G., Guo, L., Ren, J., Object detection in optical remote sensing images based on weakly supervised learning and high-level feature learning. IEEE Trans. Geosci. Remote Sens. 2015.

[6] Wang, J., Song, J., Chen, M., Yang, Z., Road network extraction: a neural-dynamic framework based on deep learning and a finite state machine. Int. J. Remote Sens.2015

[7] Duarte, D., et al. "Satellite image classification of building damages using airborne and satellite image samples in a deep learning approach." ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences 4.2 (2018).

[8] Zhong, Y.; Fei, F.; Liu, Y.; Zhao, B.; Jiao, H.; Zhang, L. SatCNN: Satellite image dataset classification using agile convolutional neural networks. Remote Sens. Lett. 2016, 2, 136–145.

[9] Arshitha Femin and Biju K.S., "Accurate Detection of Buildings from

Satellite Images using CNN" IEEE, 2020.

[10] G. Scott, M. England, W. Starms, R. Marcum, and C. Davis, "Training deep convolutional neural networks for land–cover classification of high-resolution imagery", IEEE Geoscience and Remote Sensing Letters, vol. 14, no. 4, 2017.

[11] F. Hu, G.-S. Xia, J. Hu, and L. Zhang, "Transferring deep convolutional neural networks for the scene classification of highresolution remote sensing imagery," Remote Sens., vol. 7, no. 11, pp. 14 680–14 707, 2015.

[12] F. Luus, B. Salmon, F. van den Bergh, and B. Maharaj, "Multiview deep learning for land-use classification," IEEE Geosci. Remote Sens. Lett., vol. 12, no. 12, pp. 2448–2452, Dec. 2015

[13] Kadhim, Mohammed & Abed, Mohammed, Convolutional Neural Network for Satellite Image Classification, 2020.

[14] Imamoğlu, Nevrez & Martinez-Gomez, Pascual & Hamaguchi, Ryuhei & Sakurada, Ken & Nakamura, Ryosuke. (2018). Exploring Recurrent and Feedback CNNs for Multi-Spectral Satellite Image Classification.Procedia Computer Science. 140. 162-169. 10.1016/j.procs.2018.10.325.

[15] Liang, Ming and Xiaolin Hu (2015) "Recurrent Convolutional Neural Network for Object Recognition" IEEE International Conference on Computer Vision and Pattern Recognition (CVPR'15).

[16] Tan, Y.; Xiong, S.; Li, Y. Automatic Extraction of Built-Up Areas from Panchromatic and Multispectral Remote Sensing Images Using Double-Stream Deep Convolutional Neural Networks. IEEE J. Sel.Top. Appl. Earth Obs. Remote Sens. 2018, 11, 3988–4004

[17] Li, Y.; Zhang, Y.; Huang, X.; Yuille, A.L. Deep networks under scenelevel supervision for multi-class geospatial object detection from remote sensing images. ISPRS J. Photogramm. Remote Sens. 2018

[18] L. Mou, P. Ghamisi and X. X. Zhu, "Deep Recurrent Neural Networks for Hyperspectral Image Classification," in IEEE Transactions on Geoscience and Remote Sensing, vol. 55, no. 7, July 2017

[19] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016

[20] Bochkovskiy, Alexey & Wang, Chien-Yao & Liao, Hong-yuan. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection., 2020

[21] Chien-Yao Wang, Hong-Yuan Mark Liao, Yueh-Hua Wu, Ping-Yang

ISSN: 2366-1313

Chen, Jun-Wei Hsieh, and I-Hau Yeh. CSPNet: A new backbone that can enhance learning capability of cnn. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshop (CVPR Workshop), 2020.