



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 7, July 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com



Classification and Object Detection on Satellite Images Using Custom CNN Architecture

Sunitha A N¹, Prof. Susmitha M N²

M.Tech Student, Department of Computer Science and Engineering, MVJ College of Engineering, Bengaluru, India ¹

Assistance Professor, Department of Electronics and Communication Engineering, MVJ College of Engineering, Bengaluru, India²

ABSTRACT: As the variety of applications for satellite image analysis growing, some of these applications include surveillance, military operations, geospatial research, and environmental impact and climate change monitoring. One of the most important aspects of satellite image analysis is the capacity to automatically recognise and classify objects. Recognizing and classifying items in aviation images may be challenging due to the type and size of the objects as well as the shifting visual features. Because of the type and information included in these photographs, manual article location is quite laborious. It is appealing to automate the finding of various highlights or articles from these satellite photos. The conventional techniques for categorising things include two steps: I identifying the areas of the image where objects are present, and (ii) categorising the items in those areas. It's harder to recognise items because of the complicated background, size, noise, and distance features. According to this research, a specially created convolutional neural network may be used to recognise and classify three distinct objects in the photographs, including trees, buildings, and cars. Also, it aims to grasp and briefly sum up the performance characteristics of the fictitious custom CNN.

KEYWORDS: Image processing techniques include convolutional neural networks, satellite imaging, and object recognition and categorization.

I.INTRODUCTION

Applications for satellite photography are expanding in importance, including environmental monitoring, remote surveillance, aerial survey, etc. All of these programmes search via satellite photos for objects, noteworthy activities, structures, etc. While processing several satellite images concurrently and processing enormous amounts of data, the bulk of applications make it very challenging to manually classify and identify objects. Despite extensive study on object detection and classification in the field of image processing, it is more difficult to recognise objects in satellite (aerial) photographs due to their small size and the challenges in monitoring and collecting their visual characteristics. To accomplish this goal, a variety of automated detection and classification methods have been proposed and are currently being developed. Many approaches, from conventional machine learning (ML) to contemporary deep learning, have been proposed for object detection and categorization in satellite pictures. Machine learning methods for detection and classification have received the greatest attention during the last several decades. These techniques involve using ML classifiers to categorise different aspects of the images. Due to the possibility of the target item's size, direction, and background changing, automated object detection is still difficult. Automated object identification problems cannot be solved with conventional machine learning classifiers that explicitly choose features like HOG, Gabor, the Hough transform, wavelet coefficients, etc. Hence, a robust strategy is required, and Deep Learning has shown promising results when applied to the goal of detection and classification using CNN. Deep learning classification algorithms that have recently been presented for very accurate automated classification systems do not rely on human feature selection.

Object identification software may automatically deduce characteristics from the images. Convolutional neural networks (CNNs)-based deep learning techniques are widely suggested for classifying and identifying objects in satellite photos. There are two stages to these models. The zones of item presence in the image are found in the first stage. Using a convolutional neural network, the items are categorized in the second step.

The performance of detection and classification is compared to the efficacy of the YOLO V3 algorithm's real execution on the same dataset as well as other typical benchmark data for other algorithms without actual execution. The YOLOV3 technique incorporates object detection and classification in a single step as opposed to conventional CNN models, which do it in two stages.

II.LITERATURE REVIEW

This section surveys the current deep learning models for object detection and classification.

The option suggested by the authors in [1] suggests utilising CNN to discover items. They advised using a rotation-invariant region-based convolutional neural network, particularly for satellite photos. In order to create and concentrate on the concept of an area, the process of normalising feature representation is carried out prior to categorising the objects (rotation invariant). This is followed by the classification procedure, which now relies on the fact that each image has a more intricate computation for patching results across rotation-invariant regions.

In [2], the authors provide a technique for identifying notable river crossings and bridges that may be visible in satellite pictures. Recursive scanning, which employs geometric constraints to recognise aspects like the fact that the water body in question is a river, must first be used to detect or identify the water bodies in the picture, which are mostly rivers. In order to further detect or identify the pixels that may belong to a bridge, a scan is carried out over the area of the rivers that were identified after the use of the river systems that were identified in the first step and based on the utilisation of information relating to the spatial dimensions of various bridges. After the recognised pixels have been discovered, their connection or connectivity is examined to see whether a bridge segment exists in the picture for the pixels that were first spotted. The drawback of this strategy is that it requires prior information and a grasp of the spatial dimensions of the objects we are looking for, in this case, the bridges.

For finding road networks that may be visible or present in aerial photographs captured by a satellite or unmanned aerial vehicle, the authors of [3] provide a two-stage technique (UAV). Using the trimming down or pruning step after the detection stage allows for automatic detection. Next, after initially identifying these areas or shapes, a Bayesian model is utilised to categorise regions that have similar or homogeneous qualities. In the second stage, conditional probability is used to determine the likelihood that any certain chunk or segment is a road. This detection technique has a high computational complexity, much as many others.

The authors of [4] propose an object-based image analysis technique that is primarily used to classify the topography of distinct lengths of land in satellite images and may be used to determine the land cover. Texture properties must be retrieved before an SVM classifier can be trained, which necessitates segmenting each item in the picture. When this is complete, the segmented objects are used to carry out the extraction procedure. The segmented objects' textural characteristics are then recovered and used to train an SVM classifier, which categorises the land according to its usage and cover. As feature extraction is often incomplete, imperfect, and limited in scope, the accuracy of this method is constrained.

The authors of [5] focused their attention on the related problem of identifying objects in satellite pictures and provided a description of a shoddy solution. It makes use of the Boltzmann machine, a generative model that does both encoding and decoding, which are fundamental activities. A Bayesian framework uses the properties of any picture as input for the first function of encode to build the capability for object identification. To be employed during training in a semi-supervised model, the Bayesian framework should be trained using data sets that comprise a sample of pictures with objects that are classified using different labels.

III.METHODOLOGY OF PROPOSED SURVEY

Figure 1 shows the design for these deep learning (DL) model proposed of object finding then object organization. The suggested remedy has the following crucial features:

- Acquisition of Image
- Preprocessing and Preparation data
- Training and Construction CNN
- Results Analysis and CNN Validation
- Bounding Box Drawing

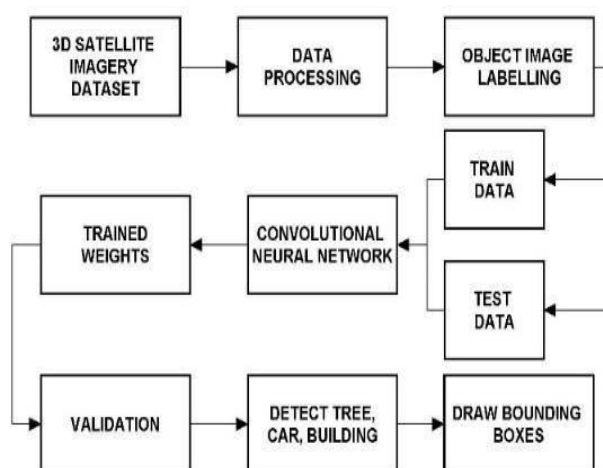


Figure 1 – System Architecture

the position of the features is determined by where they are in the input pictures. This problem restricts the method's ability to provide reliable results since it is more specifically fitted to the training data set and generates results that are specific rather than being able to generate data that is more broad, which aggravates the problem. Overfitting is a possible explanation for the issue, which leads to the unsatisfactory outcomes. The pooling layers, which are a crucial component of the CNN, may be used to generalise the existence of features over the whole picture. The average pooling and max pooling layer both have their advantages and disadvantages, with the former providing a clearer image due to its emphasis on maximum values, such as the intensity of the light in this scenario. Although average pooling aims to preserve the essential aspects of the picture while providing a smoother appearance. Avg-pooling was selected for this use-case because it can easily extract specified characteristics from an input picture. network receives its input from the annotated images. As shown in Figure 2, the customised convolutional neural network's architecture and layer details are both customised. Because the required training requires a lot of resources. A Nvidia GPU is used for the training together with Google. In comparison to a network with normal convolutions of the same depth, the suggested customised CNN has the benefit of a large decrease in the number of parameters to be evaluated.



	Type	Filters	Size	Output
	Convolutional	32	3 × 3	256 × 256
	Convolutional	64	3 × 3 / 2	128 × 128
1x	Convolutional	32	1 × 1	128 × 128
	Convolutional	64	3 × 3	
	Residual			
	Convolutional	128	3 × 3 / 2	64 × 64
2x	Convolutional	64	1 × 1	64 × 64
	Convolutional	128	3 × 3	
	Residual			
	Convolutional	256	3 × 3 / 2	32 × 32
8x	Convolutional	128	1 × 1	32 × 32
	Convolutional	256	3 × 3	
	Residual			
	Convolutional	512	3 × 3 / 2	16 × 16
8x	Convolutional	256	1 × 1	16 × 16
	Convolutional	512	3 × 3	
	Residual			
	Convolutional	1024	3 × 3 / 2	8 × 8
4x	Convolutional	512	1 × 1	8 × 8
	Convolutional	1024	3 × 3	
	Residual			
	Avgpool		Global	
	Connected		1000	
	Softmax			

Figure 2: Architecture of custom CNN

Convolution is a basic mathematical method that is used to several different aspects of image processing. It is a simple mathematical operation, yet it is essential to the study of image processing. It is feasible to merge two arrays with the same number of dimensions but different sizes using the convolution function to create a third array with the same number of dimensions. This is typically a necessary and important element that may be used in image processing in order to develop operators that receive certain pixel values as input and then produce as output pixels that are essentially a simple linear combination of the input pixels. In the context of image, a grayscale picture is often represented as an input array of integers, which only represents one of the input arrays. The kernel, a second array with two dimensions and sometimes of smaller magnitude or size, may have the same thickness as a single pixel.

Convolution simply involves sliding a 2x2 or else comparable seed crossways there image after upper leftward together bottom right though making sure there seed goes finished altogether possible motions other locations anywhere its see to non-cross there pictures border then instead refuges there entire area popular the thing motion.

The pooling layers, an essential component of CNNs, are what distinguish them from conventional neural networks and give them the capacity to handle and analyse massive quantities of input. The convolutional layers indicate the presence of features in an input picture. Nevertheless, the layers themselves present a challenge in that the sensitivity of the position of the features is determined by where they are in the input pictures. This problem restricts the method's ability to provide reliable results since it is more specifically fitted to the training data set and generates results that are specific rather than being able to generate data that is more broad, which aggravates the problem. Overfitting is a possible explanation for the issue, which leads to the unsatisfactory outcomes. The pooling layers, which are a crucial component of the CNN, may be used to generalise the existence of features over the whole picture. The average pooling and max pooling layer both have their advantages and disadvantages, with the former providing a clearer image due to its emphasis on maximum values, such as the intensity of the light in this scenario. Although average pooling aims to preserve the essential aspects of the picture while providing a smoother appearance. Avg-pooling was selected for this use-case because it can easily extract specified characteristics from an input picture.

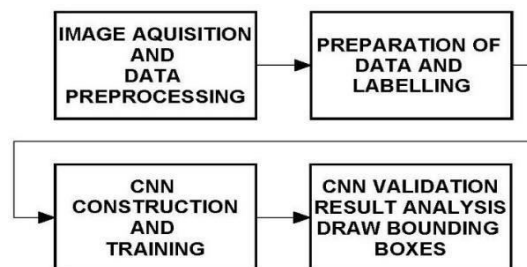


Figure 3 – System Modules

System architecture serves as the foundation for creating a strategy or design for the software that may communicate with several other subsystems in order to complete the implementation process. In order to perform the mission, it works along with other subsystems, and a framework was created to allow for communication and control amongst them. Figure 3 lists the primary components of the system.

A. Image Acquisition and Data Preprocessing

Acquiring the image dataset is essential for any object detection-based system development. While collecting datasets, there are a few rules to follow that might help with prediction and effective training of the convolutional neural network model. A dataset with more data will often provide better results. The data received from the VALID website has never been processed in any way and uncooked. Also, since real-world satellite imagery data is not accessible to the general public, a 3D model of the same was employed in this study as the proof of concept.

B. Data preparation and labelling of object images

Data preparation is the process of gathering and applying domain expertise to the dataset's images in order to sort and iterate through them. This section and the ones that follow it go into great depth about the steps that go into processing raw data. During data preprocessing, the enormous quantity for twin statistics gathered after there 3D satellite imaging collection is converted and cleaned to create a trainable dataset. According to the objects seen in the image, each image is opened and given a label. In this case, a separate XML file is used to store the labels.

Since there employment for this effort is built taking place administered knowledge, labelling of items in the photos is necessary. It belongs to the AI subcategory and is also popularly referred to as supervised deep learning. Labeling of As the system implementation for this work relies on supervised learning, it is necessary to identify objects in the photos. It also goes by the name "supervised deep learning," and it falls within the AI subcategory in common parlance. Its primary method for predicting results or for the task of categorizing data is to train algorithms using labelled datasets. Up until the model is regarded to be well fitted, the method, which is a model or algorithm that continuously modifies its weights depending on incoming data. Cross-validation involves a learning and changing process. It provides precise location information to the neural network during the training phase for objects like trees, cars, and buildings. Because CNN is the basis for this work's system implementation, labelling of objects in the photos is necessary.



C. CNN Training and Construction

The subset used to test the trained model is known as the Test set, while the subset used to train a model is known as the Training phase. Designing a neural network architecture designed for this purpose involves use-case is known as CNN creation. In this manner, we can 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 work. The coding for the custom CNN has been implemented in Fig 6. 90% of the 3D satellite imagery dataset was used for training, and the remaining 10% was used for testing. This increases the amount of training data available to the CNN, improving prediction accuracy. The ultimate anticipated value is determined iteratively by calculating the weighted average of this expected value. This particular system falls under the custom object detection sub-class because it must identify objects like automobiles, buildings, and trees. Training in this situation requires a lot of resources and cannot be completed on a laptop or local PC. To solve this issue, a virtual machine running on Google Collab was equipped with an Nvidia Tesla T4 Tensor Core GPU and 16GB of virtual Memory. Training was carried out twice, with 50 and 100 epochs, respectively, being used for each training iteration. Figure 4 screenshot depicts training that was conducted over approximately 100 epochs. The total number of objects taught during that era will be listed after each epoch.

```

1 from torch.autograd import Variable
2 import torch.nn.functional as F
3
4
5
6
7
8 class customCNN(torch.nn.Module):
9     def __init__(self):
10        super(customCNN, self).__init__()
11        self.conv1 = torch.nn.Conv2d(1, 16, kernel_size=3, stride=1, padding=1)
12        self.pool = torch.nn.MaxPool2d(kernel_size=2, stride=2, padding=0)
13        self.conv2 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
14        self.conv3 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
15        self.conv4 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
16        self.conv5 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
17        self.conv6 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
18        self.conv7 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
19        self.conv8 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
20        self.conv9 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
21        self.conv10 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
22        self.conv11 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
23        self.conv12 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
24        self.conv13 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
25        self.conv14 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
26        self.conv15 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
27        self.conv16 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
28        self.conv17 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
29        self.conv18 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
30        self.conv19 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
31        self.conv20 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
32        self.conv21 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
33        self.conv22 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
34        self.conv23 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
35        self.conv24 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
36        self.conv25 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
37        self.conv26 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
38        self.conv27 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
39        self.conv28 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
40        self.conv29 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
41        self.conv30 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
42        self.conv31 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
43        self.conv32 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
44        self.conv33 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
45        self.conv34 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
46        self.conv35 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
47        self.conv36 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
48        self.conv37 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
49        self.conv38 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
50        self.conv39 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
51        self.conv40 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
52        self.conv41 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
53        self.conv42 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
54        self.conv43 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
55        self.conv44 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
56        self.conv45 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
57        self.conv46 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
58        self.conv47 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
59        self.conv48 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
60        self.conv49 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
61        self.conv50 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
62        self.conv51 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
63        self.conv52 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
64        self.conv53 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
65        self.conv54 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
66        self.conv55 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
67        self.conv56 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
68        self.conv57 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
69        self.conv58 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
70        self.conv59 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
71        self.conv60 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
72        self.conv61 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
73        self.conv62 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
74        self.conv63 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
75        self.conv64 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
76        self.conv65 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
77        self.conv66 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
78        self.conv67 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
79        self.conv68 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
80        self.conv69 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
81        self.conv70 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
82        self.conv71 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
83        self.conv72 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
84        self.conv73 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
85        self.conv74 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
86        self.conv75 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
87        self.conv76 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
88        self.conv77 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
89        self.conv78 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
90        self.conv79 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
91        self.conv80 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
92        self.conv81 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
93        self.conv82 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
94        self.conv83 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
95        self.conv84 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
96        self.conv85 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
97        self.conv86 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
98        self.conv87 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
99        self.conv88 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
100       self.conv89 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
101       self.conv90 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
102       self.conv91 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
103       self.conv92 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
104       self.conv93 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
105       self.conv94 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
106       self.conv95 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
107       self.conv96 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
108       self.conv97 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
109       self.conv98 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
110       self.conv99 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
111       self.conv100 = torch.nn.Conv2d(16, 16, kernel_size=3, stride=1, padding=1)
112       self.fc1 = torch.nn.Linear(16 * 16 * 16, 10)
113       self.fc2 = torch.nn.Linear(10, 10)
114

```

Figure 4 : Code Implementation of CNN Layers

The custom CNN provides a (.pt) weights file after successfully completing training, whose container been utilised by there inference way together conduct finding on chance objects. A parameter like weight is crucial in the operation of a neural network in order to modify input data from within the network's various hidden layers. In other words, the network is nothing more than a gathering of bumps that convey then process info or information, similar to how neurons are connected in the brain. There are three internal factors that each node uses to define itself: heaviness, a collection of ideas, then a partiality value. A node processes the input in a sequential manner similar to an assembly line process, multiplying the input by the weight to produce an output. The output is then evaluated to determine if it should be tested further or sent on to the CNN's next layer. The network's hidden layers are frequently where the weights are found. The weights are set appropriately such that the system can recognise the designated objects (cars, buildings, and trees) with accuracy from satellite pictures that are given as input.

D. CNN Validation, Analysis of Results

There "visualize detections" method is used to deploy the bounding box function, which can be written in plain Python. This particular function accepts a few arguments as input, including the images, boxes, classes, predictions' scores, figure size, line width, and bounding box colour as RGB channels (Red-Green-Blue). In essence, these findings are metrics that make it simpler to analyse the outcomes of predictions. In the findings, the names of the items that it was able to recognise are also printed. The results of this experiment include the system's ability to recognise the things (cars, trees, and buildings) as well as how effectively it can predict them.

IV .RESULTS

At this stage, a small sample of randomly selected images from the test folder files may be used to assess the proposed system. As a result, it will be able to recognise and identify objects in the photographs like automobiles, trees, and buildings as well as create bounding boxes around them to make identification simpler. In a Python virtual environment, detection is performed via Google Collab or CLI.

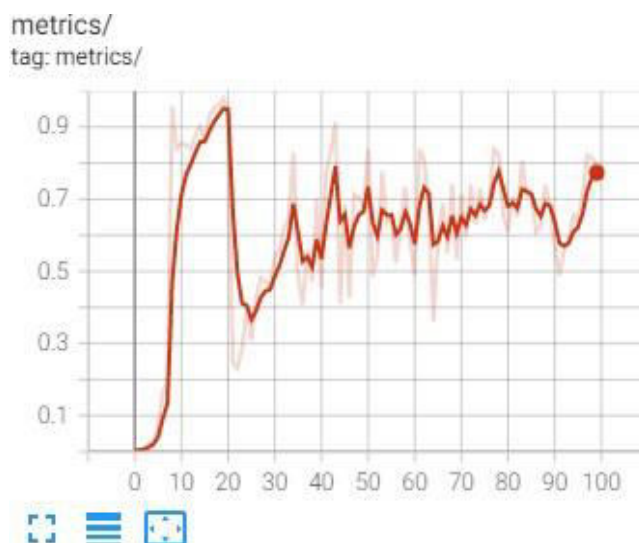


Figure 5 Accuracy of system with 100 epochs

The accuracy of Figure 5's representation of the final assessment metrics is 94.65%.

The system's object detection phase is illustrated with 50 and 100 epochs, respectively. As can be seen, the system underwent 100 epochs of training, which resulted in significantly higher accuracy, which also served as the final forecast accuracy obtained. The most accurate and precise model with the most current training weights was evaluated using metrics from the Tensor-Board library, and the ultimate accuracy reached was 94.65%. Here, the tallest mountain will be considered.

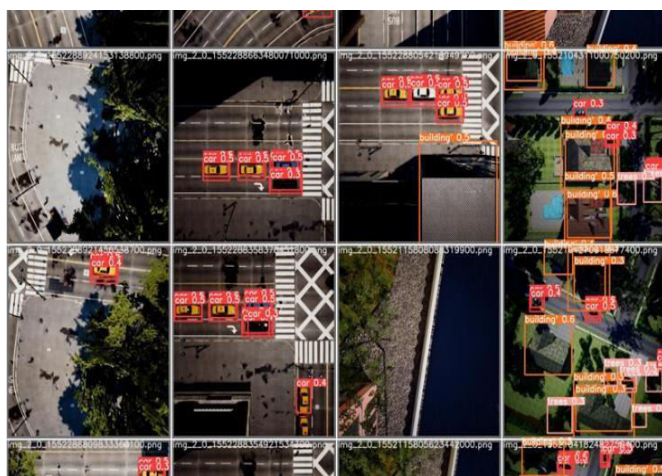


Figure 6 : System during run-time (with 100 epochs)



Figure 7: System during run-time (with 50 epochs)

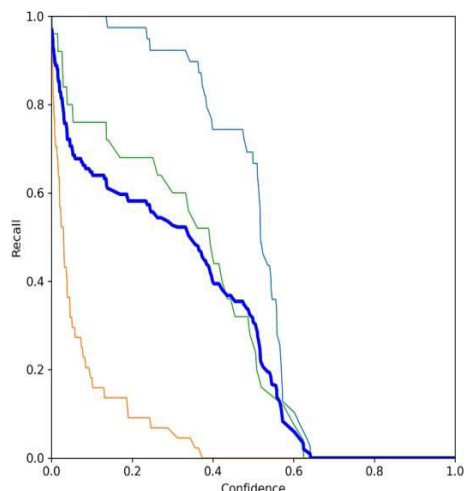


Figure :8.a

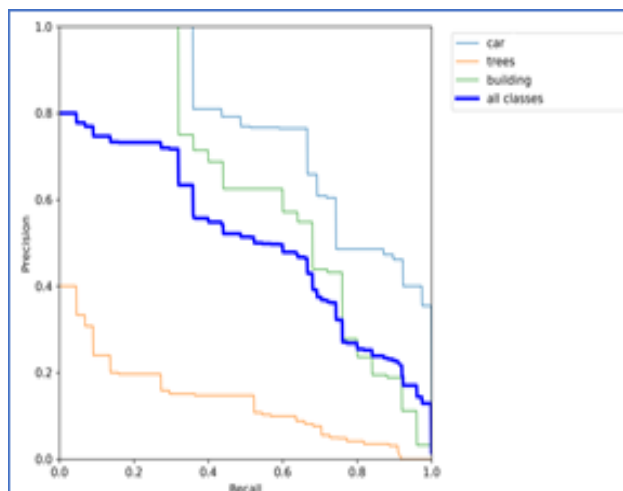


Figure :8.b

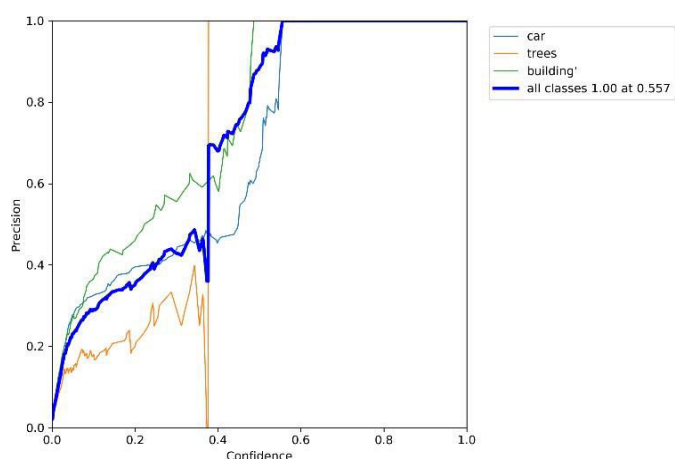


Figure :8.c

Figure 8 – Comparative Analysis of Performance Metrics

The accuracy map across various levels of For a number of object types, the confidence in the recommended custom CNN is shown below.

Under the suggested method, the automobile object achieves a superior balance between recall and precision, followed by the structure and the trees.

In the proposed custom CNN, car objects are categorized with higher confidence, followed by buildings and trees. Precision map across different confidence levels.

Precision to recall ratio: In the suggested custom CNN solution, the recall is highest for automobile construction and tree items come next.

F1 measure plot for various object classes - According to the findings, the suggested solution's classification of cars and buildings is superior than that of trees.

Following successful identification, the suggested Implementation using a proprietary convolutional neural network and the other implementation mentioned in the comparison were compared using the comparison graph shown in Figure 10 below. We can see from the comparison that the custom CNN is superior to the other solutions for object detection from satellite imagery. Nevertheless, comprehensive testing of YOLO v3 and v4 for the identical use-case scenario produced only about 90% net prediction accuracy. In comparison to YOLO v3, the proposed custom CNN method is 4.25 percent more accurate.

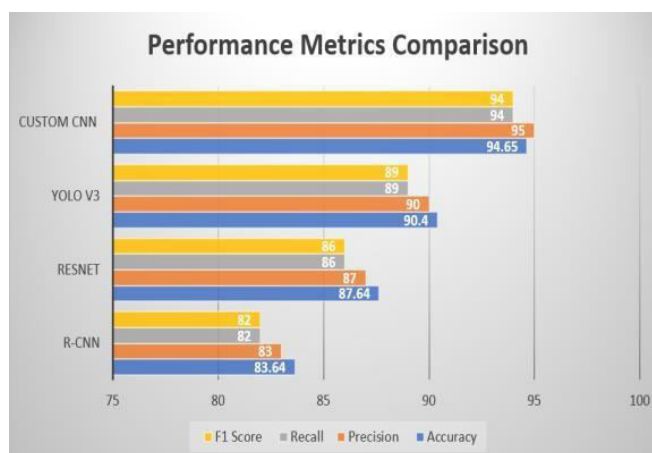


Figure 9: – Performance Metrics of Comparative Analysis

V.CONCLUSION AND FUTURE WORK

In the detection and categorization of chemicals in satellite imagery using a novel CNN model pictures were proposed. Three distinct objects—a car, a structure, and a tree—were used to assess the effectiveness of the suggested solution. The strategy was successful in achieving 94.65% accuracy. Although only a small number of training images were used in this study, we intend to test the model's effectiveness with many more datasets in the future. Real-time satellite images can be utilised to train the suggested neural network (with the necessary permits granted by the relevant space agency), which is a future advancement that can be applied to this study to further it.

REFERENCES

- [1] 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
- [2] 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.
- [3] 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.
- [4] 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
- [5] 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).
- [6] 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.
- [7] Arshitha Femin and Biju K.S., "Accurate Detection of Buildings from Satellite Images using CNN" IEEE, 2020.
- [8] 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.
- [9] F. Hu, G.-S. Xia, J. Hu, and L. Zhang, “Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery,” *Remote Sens.*, vol. 7, no. 11, pp. 14 680–14 707, 2015.
 - [10] 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
 - [11] Kadhim, Mohammed & Abed, Mohammed, Convolutional Neural Network for Satellite Image Classification, 2020.
 - [12] 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.
 - [13] 10.1016/j.procs.2018.10.325.
 - [14] Liang, Ming and Xiaolin Hu (2015) “Recurrent Convolutional Neural Network for Object Recognition” *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR’15)*.
 - [15] 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 scene-level 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



Impact Factor: 8.423



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details