COMPUTER VISION INTERNSHIP AT OMNIPRESENT ROBOT TECHNOLOGIES

AN INDUSTRIAL INTERNSHIP REPORT

submitted by

SOUVIK DATTA

(Reg. No. 19BEE1213)

in partial fulfillment for the award of the degree of

BACHELOR OF TECHNOLOGY

in

ELECTRICAL AND ELECTRONICS ENGINEERING

VIT

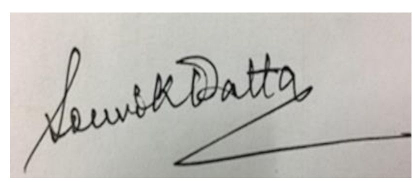
Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

AUGUST 2022

DECLARATION BY THE CANDIDATE

I hereby declare that the in-plant training report entitled "**COMPUTER VISION INTERNSHIP AT OMNIPRESENT ROBOT TECHNOLOGIES"** submitted by me to VIT University - Chennai Campus, in partial fulfillment of the requirement for the award of the degree of **Bachelor of Technology** in **Electrical and Electronics Engineering** is a record of bonafide industrial training undertaken by me under the supervision of **Mr. Shivankit Arun.** I further declare that the work reported in this report has not been submitted and will not be submitted, either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university.



Signature of the Candidate

Place: Chennai

Date: 20th July, 2022

SCHOOL OF ELECTRICAL ENGINEERING

BONAFIDE CERTIFICATE

This is to certify that the in-plant training report entitled "**COMPUTER VISION INTERNSHIP AT OMNIPRESENT ROBOT TECHNOLOGIES"** submitted by **SOUVIK DATTA (Reg. No. 19BEE1213)** to VIT University - Chennai Campus, in partial fulfillment of the requirement for the award of the degree of **Bachelor of Technology** in **Electrical and Electronics Engineering** is a record of bonafide in-plant training undertaken by him/her under my supervision. The training fulfills the requirements as per the regulations of this Institute and in my opinion meets the necessary standards for submission. The contents of this report have not been submitted and will not be submitted either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university.

<Signature of the HoD>

Head of the Department **(B.Tech EEE)**

Date:

Examiner (s) Signature

1. **2.**

OMNIPRESENT
ROBOT TECH

Date: 1 July 2022

CERTIFICATE BY THE TRAINING OFFICER

This is to certify that the project report entitled "Ground Surface Mapping with GPR" submitted by SOUVIK DATTA (Reg. No. 19BEE1213) to VIT University -Chennai Campus, in partial fulfillment of the requirement for the award of the degree of Bachelor of Technology in Electrical and Electronics Engineering is a record of bonafide in-plant training undertaken by him/her under my supervision. The training fulfills the requirements as per the regulations of this Institute and in my opinion meets the necessary standards for submission. The contents of this report have not been submitted and will not be submitted either in part or in full, for the award of any other degree or diploma in this institute or any other institutor's university.

Signature of the External Supervisor
Mr. Shashankit Arun
TRAINING OFFICER
Lead Software Engineer

Reg. Office - 20 Supreme Enclave, Mayur Vihar Phase 1, Delhi-110091
Business Office - Plot No.-11, Basement, DDA Complex, Okhla Phase I, Delhi-110020
+91 11 41327323 * www.omnipresenttech.com
CIN - U72900DL2010PTC210996

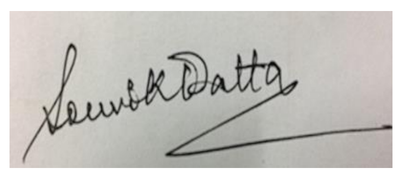
ACKNOWLEDGEMENT

I would like to express my heartfelt appreciation to Vellore Institute of Technology, Chennai for providing such an opportunity for students to broaden their perception of how the real world in the field of civil engineering looks like, as well as for organising the entire internship programme and making every effort to ensure that the entire internship programme achieves its desired goals.

I would like to thank Mr. Shivankit Arun, Lead Software Engineer and Mr. Manish Piplani, Chief Operations Officer at Omnipresent Robot Technologies, for giving me the opportunity to contribute in solving real world problems and for assisting me with my day-to-day tasks throughout my internship.

I would want to express my gratitude to the employees who assisted me with moral and material needs, as well as sharing their knowledge with me in order to better my theoretical understanding to the practical tasks.

Finally, I'd want to convey my heartfelt gratitude to my family and friends for their support in all aspects and for allowing me to spend my entire internship period at work.



Place : Chennai

(Souvik Datta)

Date : 20th July, 2022

TABLE OF CONTENTS

ABSTRACT

Ground-Penetrating Radar is a geophysical tool that uses radar pulses to photograph the subsurface. It is a non-invasive method of surveying the subsurface for underground utilities such as concrete, asphalt, metals, pipes, cables, or masonry. The project involved the use of MALA GPR to study underground utilities and process them later on using various AI and image processing methods.

Object Detection is a computer vision approach that enables us to recognise and find things in images or videos. Object detection can be used to count objects in a scene, determine and track their precise locations, and precisely label them using this type of identification and localization. Specifically, object detection generates bounding boxes around observed items, allowing us to determine where these objects are in (or how they move through) a given scene. The project involved the use of YOLOv5 to detect objects from UAVs over six classes.

Remote Sensing Change Detection (RSCD) is the process of identifying changes between scenes of the same location acquired at different times. This is an active research area with a broad range of applications. Satellite images taken on the earth's surface are analysed to identify the spatial and temporal changes that have occurred naturally or manmade. The project involved the use of Remote Sensing Models to detect changes between two images.

LIST OF TABLES –

LIST OF FIGURES –

OMNIPRESENT ROBOT TECHNOLOGIES

Omnipresent Robot Tech is India's leading robotics, industrial UAV/drone and video analytics startup. Founded by Mr. Aakash Sinha in 2010, they build drones and robots for industrial inspections and defence and the software that drives them. Using Computer Vision, Machine Learning and Virtual Reality, they are able to provide their clients with unsurpassed visual analytics and actionable insights. They are the leading Drone and Camera based Artificial Intelligence services provider in India with founders from Carnegie Melon University under the stewardship of Dr Raj Reddy (Father of Robotics across the globe).

Omnipresent Robot Technologies have 6 state government emplacements for providing AI services from Andhra Pradesh, Maharashtra, Karnataka, Gujarat, Punjab and West Bengal. Apart from Government bodies, the organization is currently pandering to the requirements of several leading private industry houses of India such as HINDALCO, Reliance, JSW, JK Cement, etc. by providing drone services for the betterment of surveillance, administration and engineering aspects of an operational plant.

Industry: Aviation and Aerospace Component Manufacturing
Company Size: ~ 50 employees

Current Projects –

Industrial Drones & Robots	Video Analytics Software
• Drones built for industrial inspections • River cleaning robots • Robots for logistics • Emergency Response drones • Defense Drones	• Software for Drones • 3D Modelling • Nerve Center – surveillance using machine learning, computer vision & video analytics

Website: http://www.omnipresenttech.com
Specialties: Robotics, Unmanned Aerial Vehicles, Unmanned Ground Vehicles, and Perception systems
News and Media –

1. Ro-Boat by Channel One (USA) – [YouTube]
2. Zerodha co-founder Nikhil Kamath invests in drone start-up Omnipresent Robot Tech – [Read More: Financial Express]
3. India's longest drone flight conducted in Haryana for HPCL – [Read More: New Indian Express]

CHAPTER - I: GROUND SURFACE MAPPING WITH GPR

INTRODUCTION

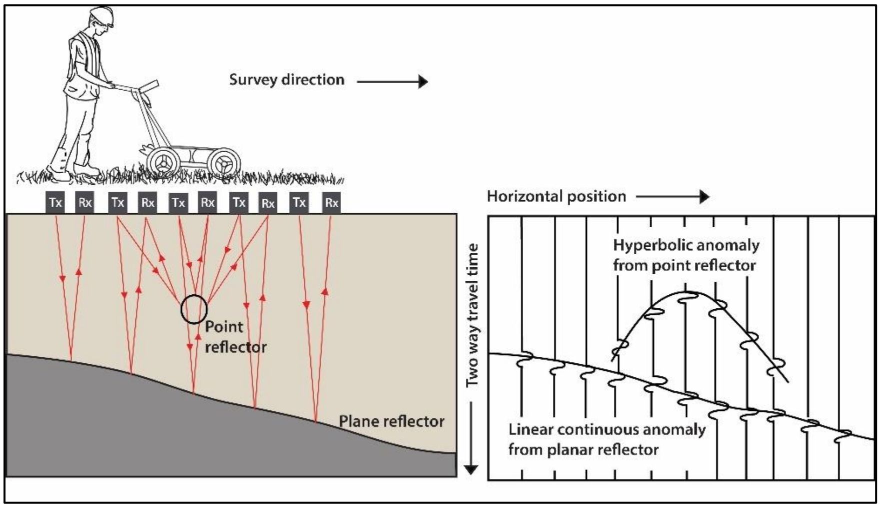


Figure 1. Ground Penetrating Radar Working Principle

Ground-Penetrating Radar is a geophysical tool that uses radar pulses to photograph the subsurface. It is a non-invasive method of surveying the subsurface for underground utilities such as concrete, asphalt, metals, pipes, cables, or masonry. The analysis of these GPR graphs will be influenced by a variety of factors. For example, a substance's dielectric constant, soil type, and so forth.

The project began with figuring out how to retrieve the data (.rd3/.rd7) files from the MALA device in the form of an image (.jpg or.png) before proceeding with the image processing. This was accomplished using the RGPR package in R-Studio. To acquire the optimal format for the GPR plot, some background removal and noise removal filters were utilised.

Following this the images were post-processed using two methods for detecting utility parabolas in GPR plots, which indicate the existence of utilities beneath the ground. To detect utilities, AI methodologies such as the MASK R-CNN and YOLOv3 models, as well as classification-based methods to determine the presence or absence of utilities in an image. Aside from that, image processing-based methods such as Gabor filters and Edge Detection methods to automate the utility detection process.

Lastly, a visualisation software for the entire process flow was developed. By selecting a.rd3/.rd7 file from the system, one can get a final output containing detected utilities in the image as well as four other files containing ***Coordinates***, ***Folium Map***, ***Google Earth kml file***, and ***iteration kml*** file showing the number of iterations the operator did while field testing.

WORK ASSIGNED

I. AI-BASED METHOD – 1

YOLOv3 (You Only Look Once, Version 3) is a real-time object detection algorithm that identifies specific objects in videos, live feeds, or images. The YOLO machine learning algorithm uses features learned by a deep convolutional neural network to detect an object. Object classification systems are used by Artificial Intelligence (AI) programs to perceive specific objects in a class as subjects of interest. The systems sort objects in images into groups where objects with similar characteristics are placed together, while others are neglected unless programmed to do otherwise.

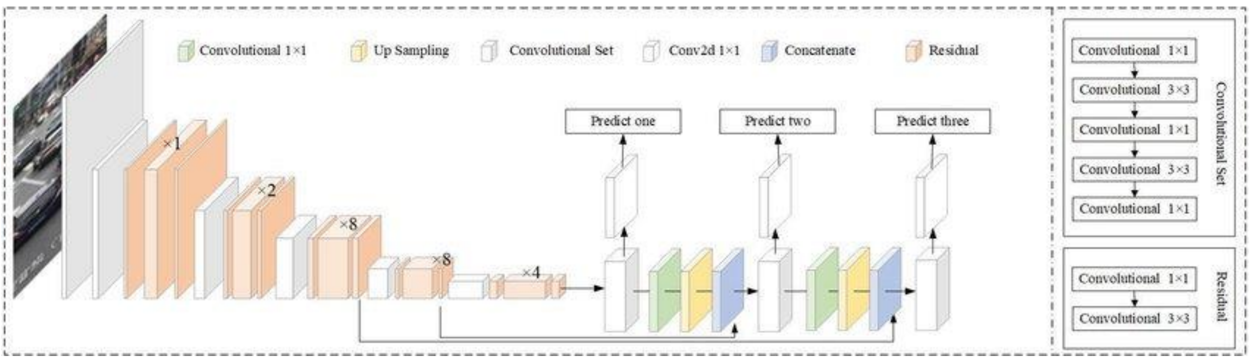


Figure 2. Neural Network Architecture of YOLOv3 [1].

To ensure lower false cases, a classification strategy is implemented in which two classes of datasets were used, "Parabolas" and "Negatives," to differentiate between a plot of GPR data including parabolas and thus utilities and a plot of GPR excluding parabolas.

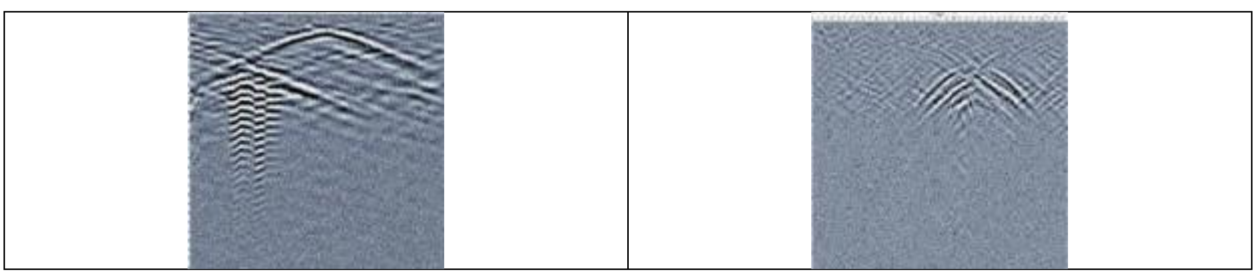


Figure 3. YOLOv3 Results depicting parabola's presence.

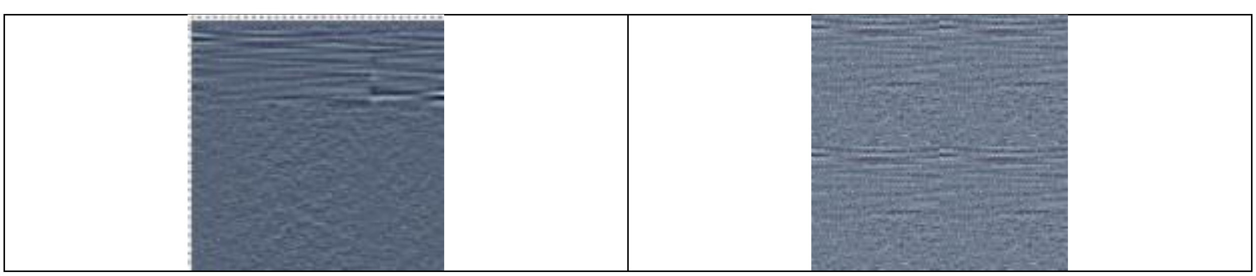


Figure 4. YOLOv3 Results depicting parabola's absence, thus Negatives.

II. AI-BASED METHOD – 2

Mask R-CNN (Mask Regional Convolutional Neural Network) is a deep neural network aimed to solve instance segmentation problems in machine learning or computer vision, it was developed by Facebook AI and Research (FAIR) in April 2017.

There are two stages of Mask R-CNN. First, it generates proposals about the regions where there might be an object based on the input image. Second, it predicts the class of the object, refines the bounding box and generates a mask at the pixel level of the object based on the first stage proposal. Both stages are connected to the backbone structure [2].

Fig. 5. depicts the concept of region-based CNN (R-CNN). This approach utilizes bounding boxes across the object regions, which then evaluates convolutional networks independently on all the Regions of Interest (ROI) to classify multiple image regions into the proposed class. The R-CNN architecture was designed to solve image detection tasks. Also, R-CNN architecture forms the basis of Mask R-CNN and it was improved into Faster R-CNN.

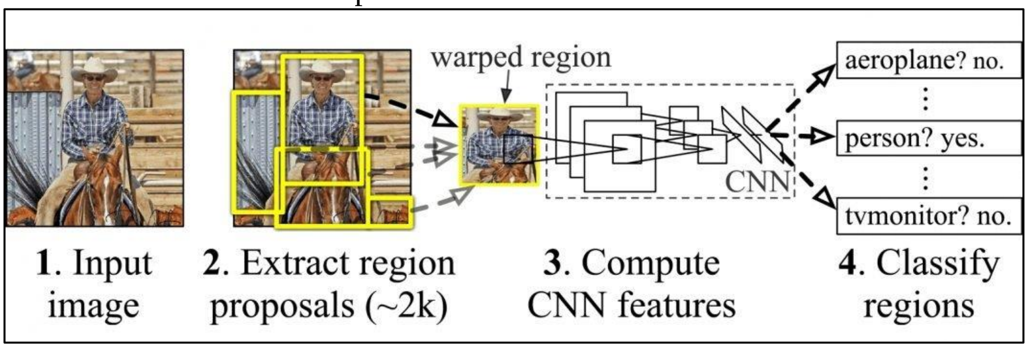


Figure 5. Concept of Regional Convolutional Neural Network (R-CNN) [3]

Mask R-CNN can automatically segment and construct pixel-wise masks for every object in an image.

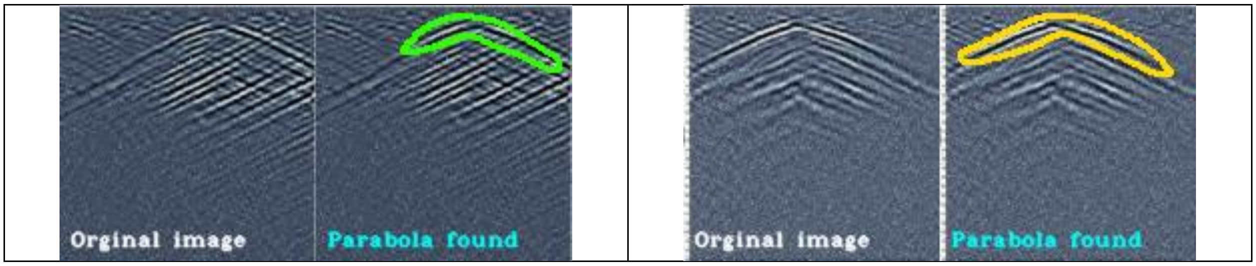


Figure 6. Mask-RCNN Results depicting parabola's presence

Figure 7. Mask-RCNN Results depicting parabola's absence

When the two models work together, they can produce findings with a very low number of false cases. It is a reliable way of distinguishing true positives from false negatives.

III. IMAGE PROCESSING METHOD

The goal was to acquire single spots in the image for parabolas. Since the peaks of parabolas were the brightest in the image, an attempt was made using the **`Scikit-Image`** function **`peak_local_max`** to obtain local maxima points in the image based on pixel intensity. **`Scikit-Image`** is a Python library [5]. **`Scikit-Image`** is an open-source image processing library for the Python programming language. It includes algorithms for segmentation, geometric transformations, colour space manipulation, analysis, filtering, morphology, feature detection, and more.
The **`peak_local_max`** function returns the coordinates of local peaks (maxima) in an image. Internally, a maximum filter is used for finding local maxima. This operation dilates the original image and merges neighbouring local maxima closer than the size of the dilation. Locations, where the original image is equal to the dilated image, are returned as local maxima.

To reduce noise before applying the function, Gaussian Median Filter was utilised. Peaks in the image were detected by the function as shown in Fig. 8 and Fig. 9.

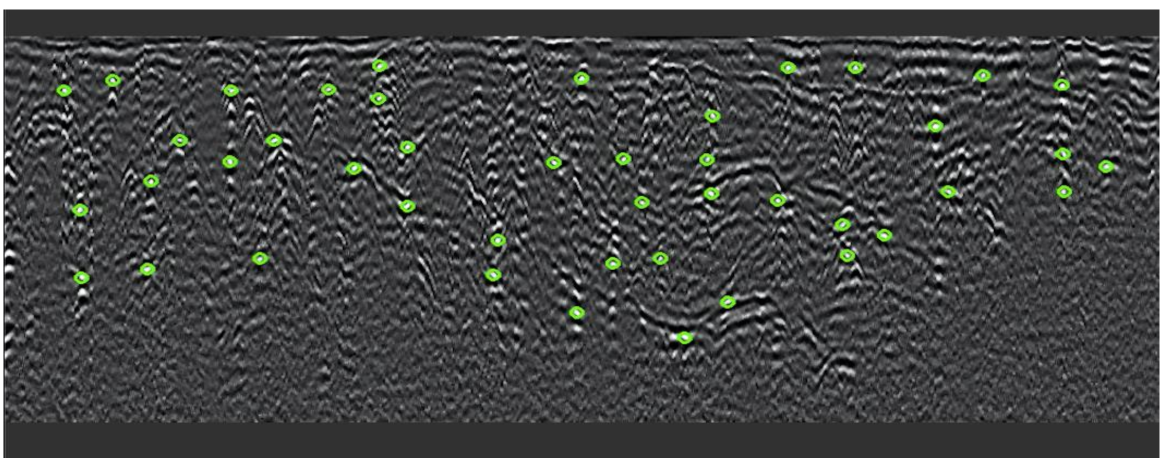


Figure 8. Detection of peaks

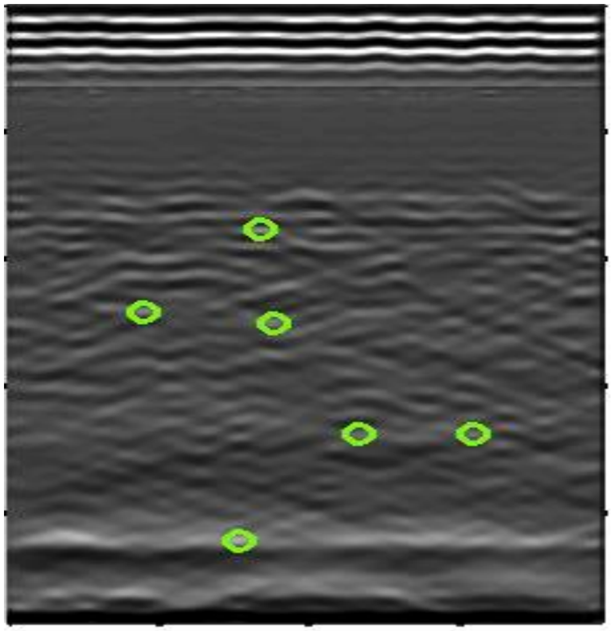


Figure 9. Detection of peaks on a .rd7 file

IV. PLOTTING UTILITIES

`Folium` is a powerful Python library that helps you create several types of Leaflet maps. By default, Folium creates a map in a separate HTML file. Since Folium results are interactive, this library is very useful for dashboard building. You can also create inline Jupyter maps in Folium [6]. Folium builds on the data wrangling strengths of the Python ecosystem and the mapping strengths of the Leaflet.js library. Using Folium, you can manipulate your data in Python, and then visualize it in a Leaflet map. Folium enables you to generate a base map of specified width and height with either default tilesets (i.e., map styles) or a custom tileset URL. The following tilesets are available by default with Folium:

1. OpenStreetMap
2. Mapbox Bright
3. Mapbox Control Room
4. Stamen (incl. Terrain, Toner, and Watercolor)
5. Cloudmade
6. Mapbox

The coordinates of utilities found had to be entered into the database and plotted on maps. This task was completed using Folium and the Mapbox API.

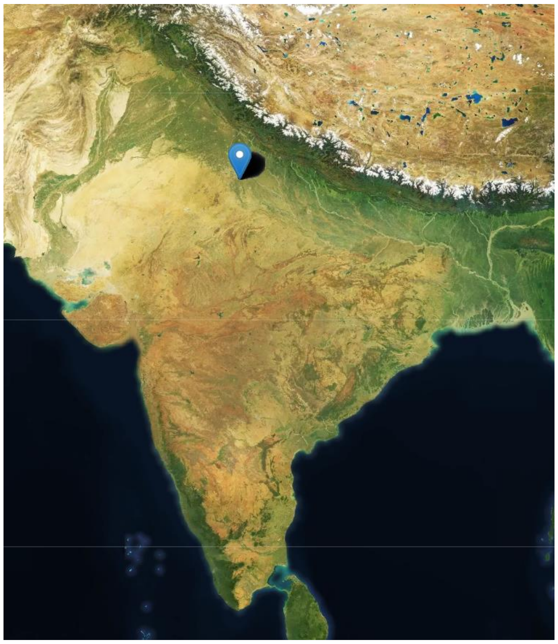


Figure 10. Folium plot on Physical Map of India

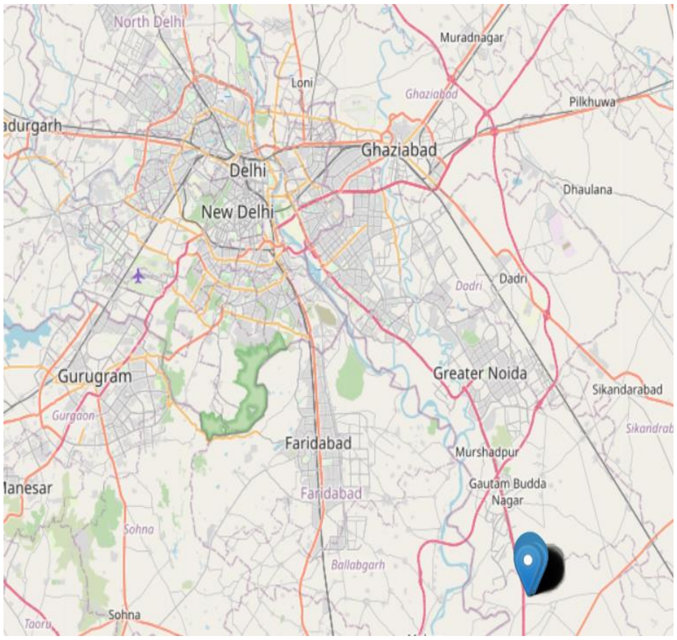


Figure 11. Folium plot depicting detected utilities

Figure 12. Representation of coordinates of utilities using Folium

V. SOFTWARE VISUALIZATION

Images from the GPR device (MALA) were transmitted to the system for additional processing via a USB drive. The photos have a.rd7 /.rd3 file extension and were converted to a readable format with the software. The image was further processed by using a median filter to eliminate noise and mathematical functions with AI models to detect parabolic peaks in the image.

To get the pixel coordinates of detected peaks in the image at the end of the operation. These pixel values were then used to correlate the GPS coordinates collected from the device in the .cor file, and final coordinates (*Latitude* and *Longitude*) describing the position of the device were retrieved.

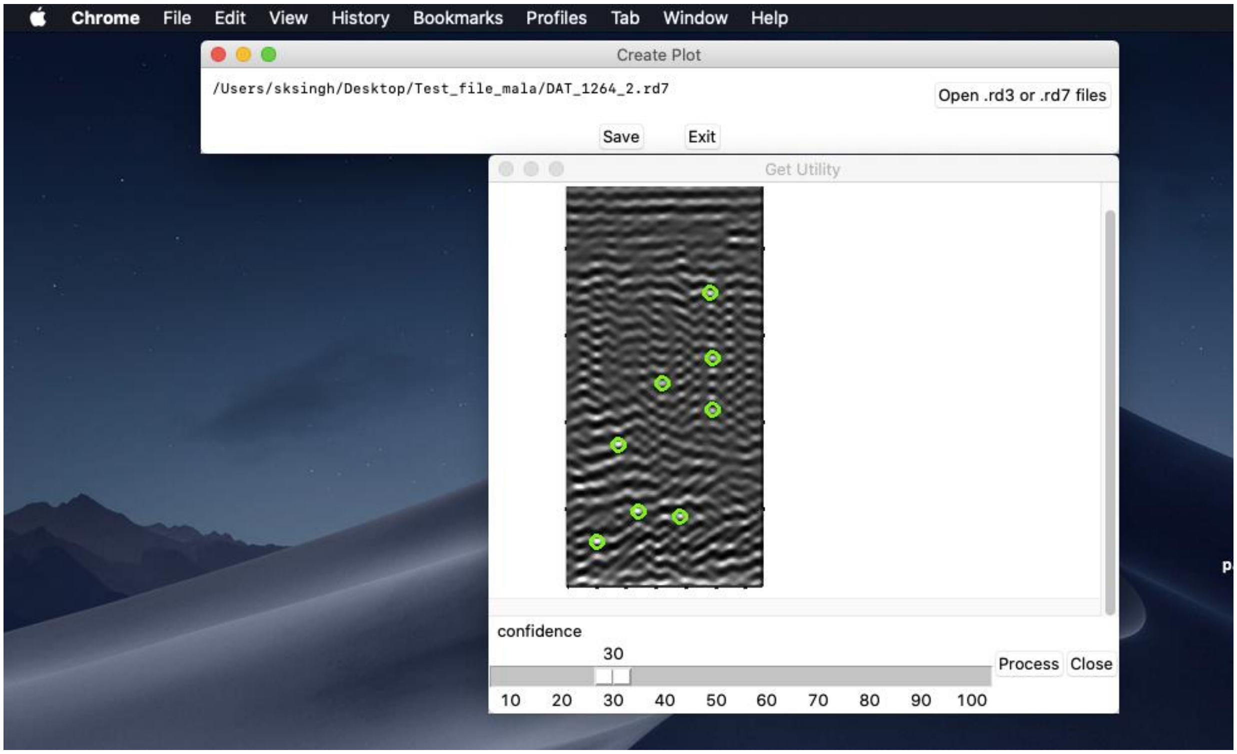


Figure 13. The ‘Tkinter’ application is designed to take in input images and process them.

CONCLUSION

To summarise the fundamental notion, ground-penetrating radar data is being used to automate the process of utility detection in a specific land's subsurface. The project's purpose is to generate useful parabola charts for utilities, as well as depth and feature information.

• Utilities in the image were detected with great accuracy.
• The user interface for the project provided a confidence threshold function that allowed the user to alter the threshold and compare the discovered utilities.
• This feature enables the operator to manually choose the ideal threshold to acquire the best detections in the image; this is simply a method of fine-tuning the results produced by our AI and statistical models.

HARDWARE MODEL of MALA GPR

Figure 14. MALA GPR with NVIDIA Jetson Nano

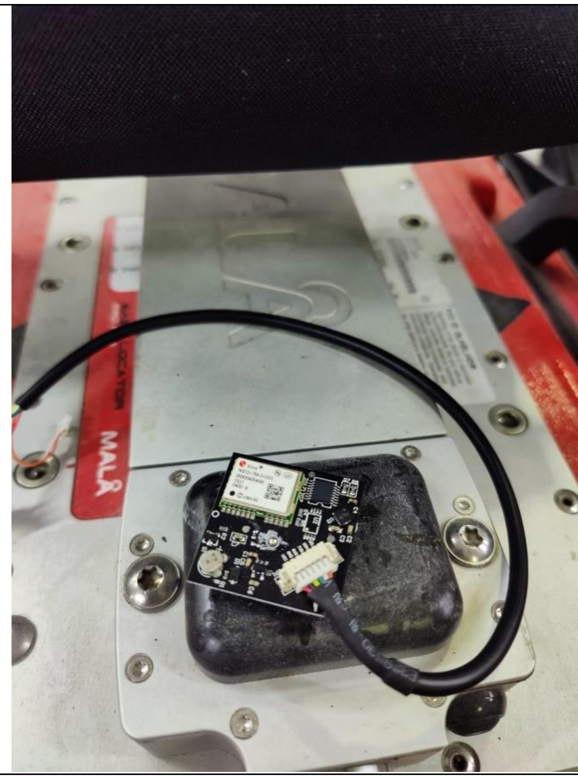


Figure 15. MALA GPR with Here 2 GPS Module

Figure 16. MALA GPR Device Overview - I

Figure 17. MALA GPR Device Overview - II

SUMMARY

Table 1. Summary of the Scope of Work and Corresponding Details associated with it

Stage1	**Scope of Work**	**Corresponding Details**
1	Development of Integration matrix for GPR equipment	Developed and integrated external GPS to achieve higher ground accuracies with the MALA GPR.
2	Utility Detection with depth (up to 3m) along with utility diameter	Depth is not a constraint with the developed algorithm. Depends on the capability of the GPR system.
3	Development of interpretation software	Rd3/Rd7 data is given as input to the application, which then processes the data to give the below-mentioned outputs. • Coordinates.csv • Folium Map • Coordinates.kml • Iterations.kml
4	Realtime Report Generation	1. Images from MALA GPR were transmitted to the system for additional processing using a USB drive. 2. .rd3 and .rd7 files were read by our application. 3. Images are then pushed to post-processing to eliminate noise and local maxima using a median filter to enhance the found parabolas. 4. Pixel coordinates of the detected peaks are collected which are then correlated with the GPS coordinates. 5. Final Latitudes and Longitudes of the detected utilities are presented in the application.
5	Data collection and training for the development of an automated utility depth detection module	Training modules for utility detection are generated.
6	Development of calibration module	A calibration option was also introduced to the application, which modifies the identified utilities by altering the confidence threshold.
7	AI platform development is based on a Neural network	1. The Mask-RCNN model is used to perform image segmentation and masking of

	developed from previously collected data to automate utility detection	portions of the image that have similar features to the dataset. 2. YOLOv3 working together with AI ensure high accuracies with a minimal number of false positives.
8	Output in 2D/3D Graphical format for interpretation	The following outputs are given from the application - 1.Coordinates.csv 2. Folium Map 3. Coordinates.kml 4. Iterations.kml

REFERENCES

[1] Mao, Q. C., Sun, H. M., Liu, Y. B., & Jia, R. S. (2019). Mini-YOLOv3: real-time object detector for embedded applications. *Ieee Access*, *7*, 133529-133538.

[2] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. In *Proceedings of the IEEE international conference on computer vision* (pp. 2961-2969).

[3] Odemakinde, E. (2022). Everything about Mask R-CNN: A Beginner's Guide - viso.ai. Retrieved 19 July 2022, from https://viso.ai/deep-learning/mask-r-cnn/

[4] Onyszko, K., & Fryśkowska-Skibniewska, A. (2021). A New Methodology for the Detection and Extraction of Hyperbolas in GPR Images. *Remote Sensing*, *13*(23), 4892.

[5] Boulogne, F., Warner, J. D., & Neil Yager, E. (2014). Scikit-image: Image processing in Python. *J. PeerJ*, *2*, 453.

[6] Wu, Q. (2021). Leafmap: A Python package for interactive mapping and geospatial analysis with minimal coding in a Jupyter environment. *Journal of Open Source Software*, *6*(63), 3414.

CHAPTER - II: OBJECT DETECTION USING UAVs

INTRODUCTION

Figure 18. Unmanned Aerial Vehicle used for Surveillance

Object Detection is a computer vision approach that enables us to recognise and find things in images or videos. Object detection can be used to count objects in a scene, determine and track their precise locations, and precisely label them using this type of identification and localization. Specifically, object detection generates bounding boxes around observed items, allowing us to determine where these objects are in (or how they move through) a given scene.

An image is labelled by image recognition. The term "dog" is applied to a photograph of a dog. A photo of two dogs is still labelled "dog." In contrast, object detection forms a box around each dog and labels it "dog." The model predicts where each object will be and what label will be assigned to it. Object detection, in this sense, gives more information about an image than recognition.

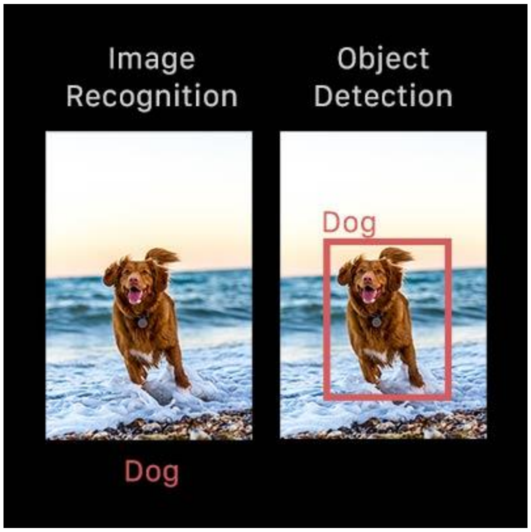


Figure 19. Object Recognition vs Object Detection

Given these key distinctions and object detection's unique capabilities, it can be applied in several ways -

- Crowd counting
- Self-driving cars
- Video surveillance
- Face detection
- Anomaly detection

Because state-of-the-art object detection techniques can accurately identify and track multiple instances of a given object in a scene, these techniques naturally lend themselves to automated video surveillance systems.
For instance, object detection models are capable of tracking multiple people at once, in real-time, as they move through a given scene or across video frames. From retail stores to industrial factory floors, this kind of granular tracking could provide invaluable insights into security, worker performance and safety, retail foot traffic, and more.

WORK ASSIGNED

YOLOv5 - YOLOv5 is the latest object detection model developed by **`ultralytics`**, the same company that developed the PyTorch version of YOLOv3, and was released in June 2020 [1].
YOLO proposes the use of an end-to-end neural network that makes predictions of bounding boxes and class probabilities all at once.

Following a fundamentally different approach to object detection, YOLO achieves state-of-the-art results beating other real-time object detection algorithms by a large margin.
In addition to increased accuracy in predictions and a better Intersection over Union in bounding boxes (compared to real-time object detectors), YOLO has the inherent advantage of speed. YOLO is a much faster algorithm than its counterparts, running at as high as 45 FPS.

YOLO Workflow –

The YOLO algorithm works by dividing the image into *N* grids, each having an equal dimensional region of SxS. Each of these *N* grids is responsible for the detection and localization of the object it contains.
Correspondingly, these grids predict bounding box coordinates relative to their cell coordinates, along with the object label and probability of the object being present in the cell. This process greatly lowers the computation as both detection and recognition are handled by cells from the image.

YOLO makes use of Non-Maximal Suppression to deal with this issue. In Non-Maximal Suppression, YOLO suppresses all bounding boxes that have lower probability scores. YOLO achieves this by first looking at the probability scores associated with each decision and taking the largest one. Following this, it suppresses the bounding boxes having the largest Intersection over Union with the current high probability bounding box. This step is repeated till the final bounding boxes are obtained.

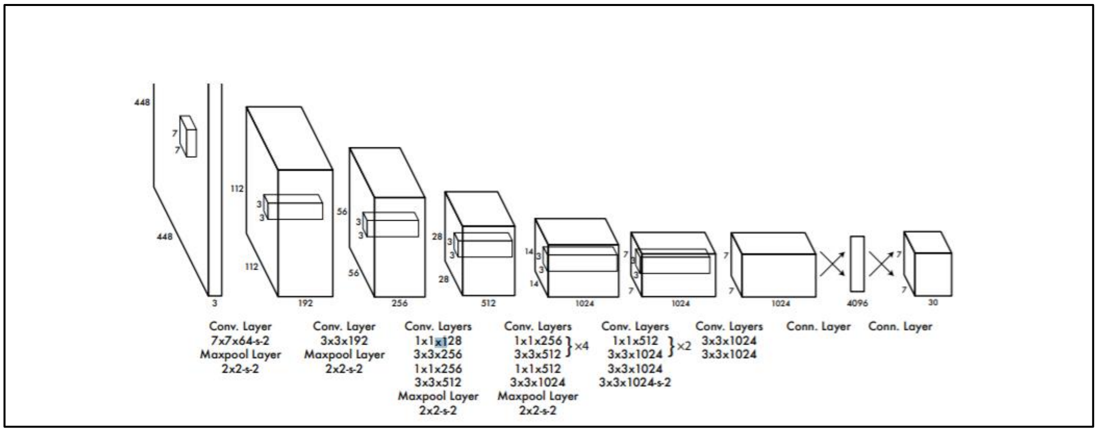


Figure 20. Figure depicting YOLO Neural Network Architecture

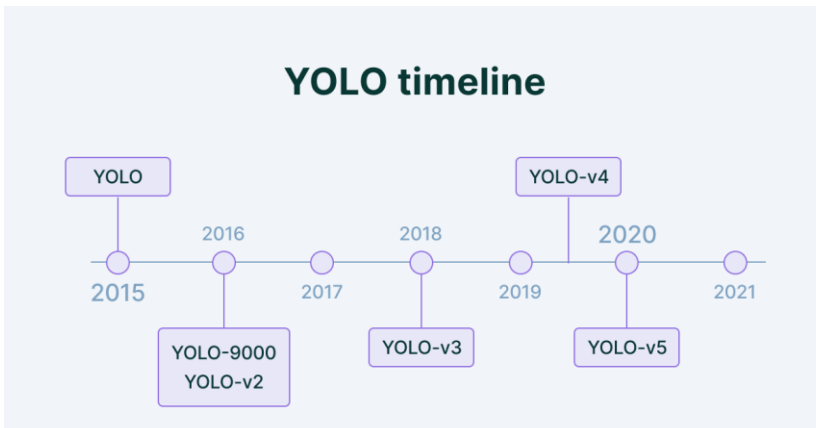


Figure 21. Figure depicting the timeline of all YOLO versions

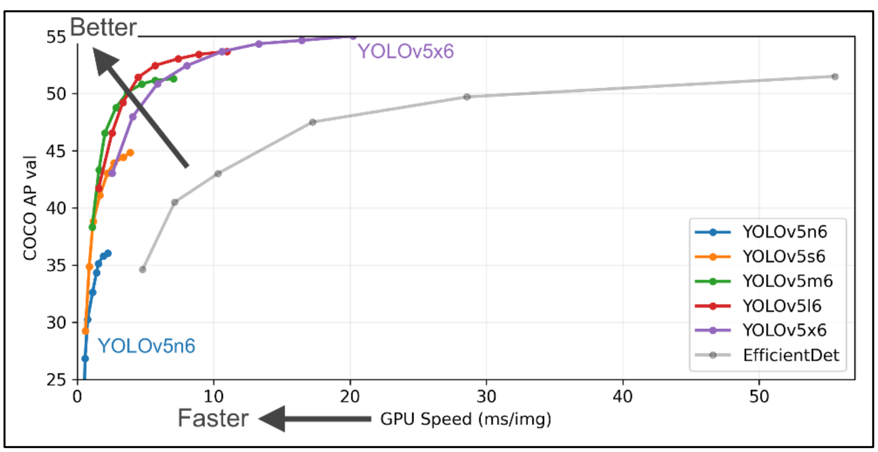


Figure 22. Figure depicting various YOLOv5 versions

AU-AIR Dataset - AU-AIR dataset is the first multi-modal UAV dataset for object detection [2]. It meets vision and robotics for UAVs having the multi-modal data from different onboard sensors and pushes forward the development of computer vision and robotic algorithms targeted at autonomous aerial surveillance. AU-AIR has several features:

Object detection in aerial images

- 32,823 labelled frames
- 132,034 object instances
- 8 object categories related to traffic surveillance
- Frames are also labelled with time, GPS, IMU, altitude, and linear velocities of the UAV

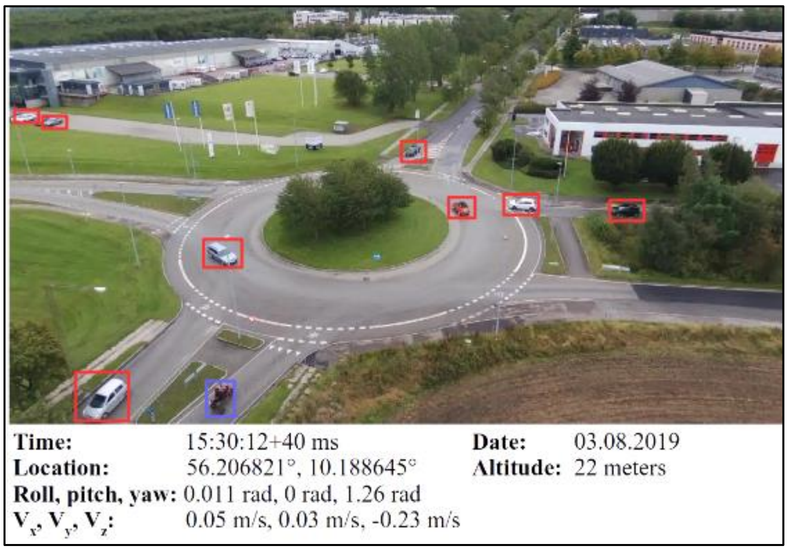


Figure 23. AU-AIR Dataset and corresponding label details.

VisDrone Dataset – Drones, or general unmanned aerial vehicles (UAVs), outfitted with cameras, have been rapidly used in a variety of applications, including agriculture, aerial photography, fast delivery, and surveillance. As a result, automatic understanding of visual data collected from these platforms has become increasingly important, bringing computer vision closer to drones. VisDrone, a large-scale benchmark with extensively annotated ground-truth for different major computer vision applications, to bring vision and drones together.

The VisDrone2019 dataset is collected by the AISKYEYE team at the Lab of Machine Learning and Data Mining, Tianjin University, China [3]. The benchmark dataset consists of 288 video clips formed by 261,908 frames and 10,209 static images, captured by various drone-mounted cameras, covering a wide range of aspects including location (taken from 14 different cities separated by thousands of kilometres in China), environment (urban and country), objects (pedestrian, vehicles, bicycles, etc.), and density (sparse and crowded scenes).

The dataset was gathered utilising several drone platforms (i.e., different models of drones), in diverse settings, and under varying weather and illumination circumstances. These frames are manually annotated with over 2.6 million bounding boxes representing popular targets like walkers, vehicles, bicycles, and tricycles.

Train a custom YOLOv5 Model to detect the following classes –

Table 2. The following six classes have been used in the model

Person	Car	Truck
Van	Bus	Negative

In YOLO, a bounding box is represented by four values **`[x_center, y_center, width, height]`**.

`x_center and y_center` are the *normalized coordinates* of the centre of the bounding box. To make coordinates normalized, take pixel values of x and y, which marks the centre of the bounding box on the x- and y-axis. Then divide the value of x by the width of the image and the value of y by the height of the image.
`width` and **`height`** represent the width and the height of the bounding box and they are normalized as well.

`An Example of a YOLO B-Box Format from the custom dataset used for this project -`

`0 0.561811 0.657244 0.185432 0.586572`

- **0** - Represents the number of classes. One might argue as to why the number of classes is 0 instead of 1. The format works in such a way that if there are **2** classes then the numbers would be represented as **0 & 1**. Similarly, if the number of classes was **8** then it would be represented from 0~7.
- **`0.561811 - x_center`**
- **`0.657244 - y_center`**
- **`0.185432 - width`**
- **`0.586572 – height`**

CONCLUSION

The training takes about 8 GPU hours utilizing about 17,985 training images and 500 epochs.

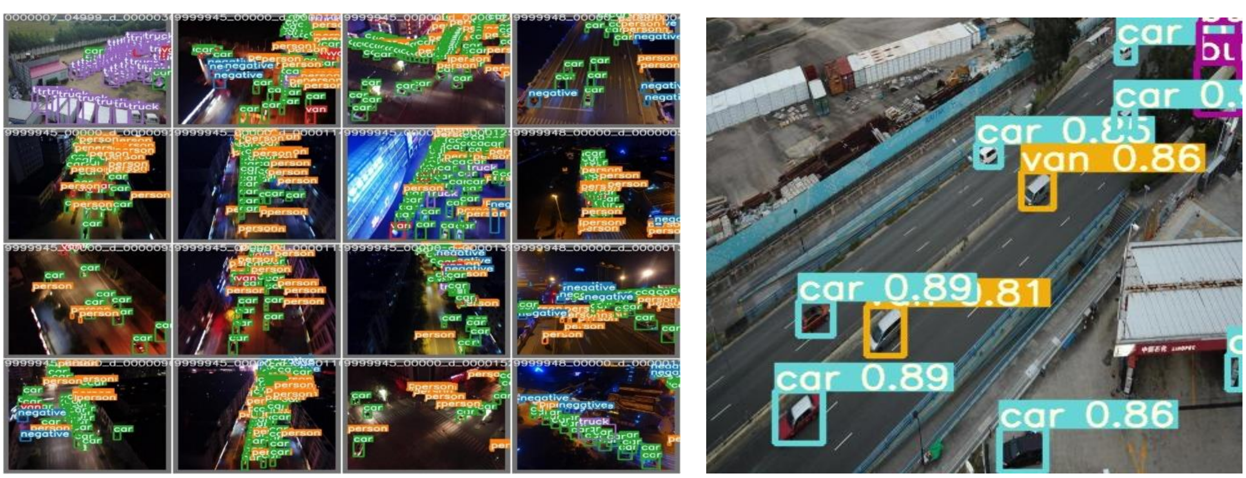


Figure 24. Final Test Results

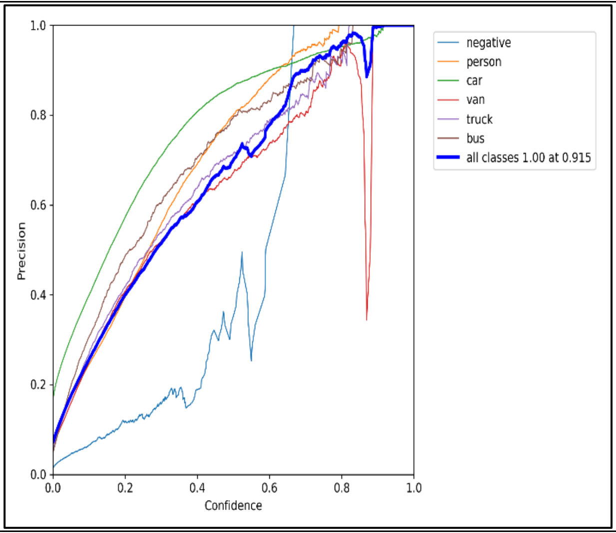


Figure 25. Precision vs Confidence Score Graph

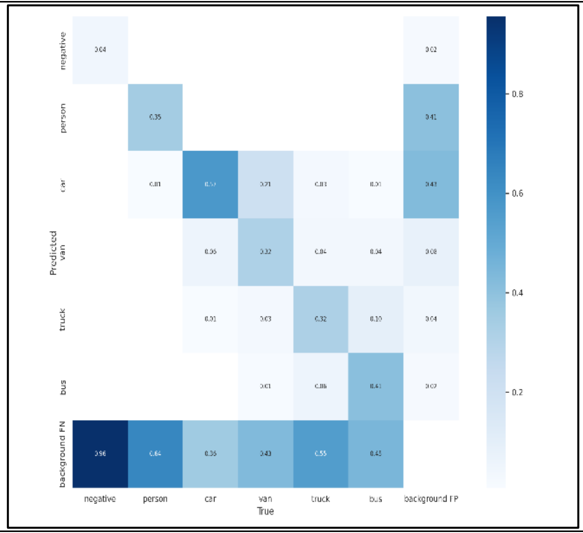


Figure 26. Confusion Matrix of the Model

REFERENCES

[1] Jocher, G., Chaurasia, A., Stoken, A., Borovec, J., NanoCode012, Kwon, Y., ... & Michael, K. (2022). ultralytics/yolov5: v6. 1-TensorRT, TensorFlow Edge TPU and OpenVINO Export and Inference. Zenodo, Feb, 22.

[2] Bozcan, Ilker, and Erdal Kayaan. "AU-AIR: A Multi-modal Unmanned Aerial Vehicle Dataset for Low Altitude Traffic Surveillance." IEEE International Conference on Robotics and Automation (ICRA), 2020, to appear.

[3] Du, D., Zhu, P., Wen, L., Bian, X., Lin, H., Hu, Q., ... & Zhang, L. (2019). VisDrone-DET2019: The vision meets drone object detection in image challenge results. In Proceedings of the IEEE/CVF international conference on computer vision workshops (pp. 0-0).

CHAPTER - III: CHANGE DETECTION USING UAVs

INTRODUCTION

Remote Sensing Change Detection (RSCD) is the process of identifying changes between scenes of the same location acquired at different times. This is an active research area with a broad range of applications. Satellite images taken on the earth's surface are analysed to identify the spatial and temporal changes that have occurred naturally or manmade.

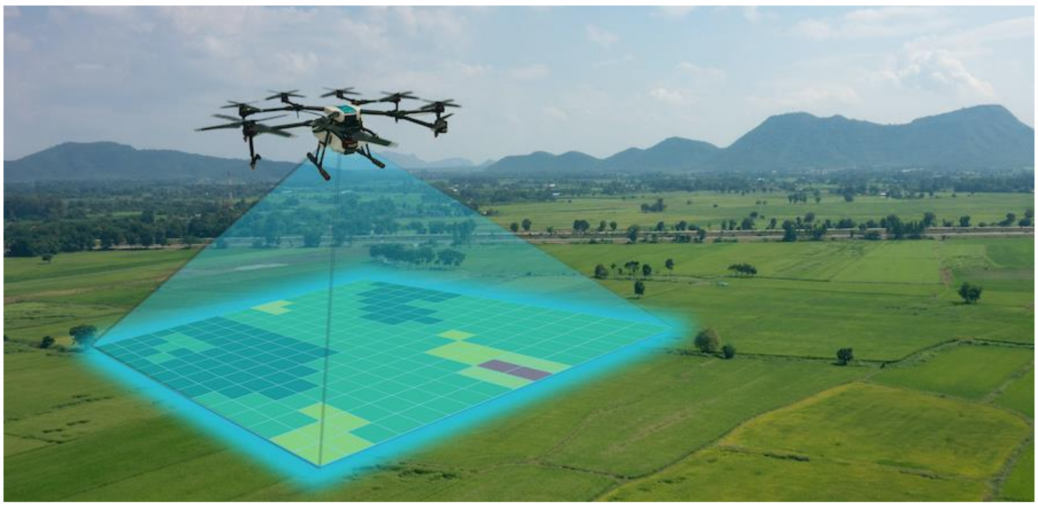


Figure 27. UAVs engaged in Remote Sensing in farmlands

Real-time prediction of change provides an understanding related to the land cover, environmental changes, habitat fragmentation, coastal alteration, urban sprawl, etc. Change Detection (CD) is at the heart of many impactful applications of remote sensing. Studying differences in land cover and land use over time with remote sensing imagery can shed light on urbanization trends, ecosystem dynamics, surface water and sea ice trends and damages through natural disasters.

Some specific uses of remotely sensed images of the Earth include

1. Large forest fires can be mapped from space, allowing rangers to see a much larger area than from the ground.
2. Tracking clouds to help predict the weather or watching erupting volcanoes, and help watch for dust storms.
3. Tracking the growth of a city and changes in farmland or forests over several years or decades.
4. Discovery and mapping of the rugged topography of the ocean floor (e.g., huge mountain ranges, deep canyons, and the “magnetic striping” on the ocean floor).

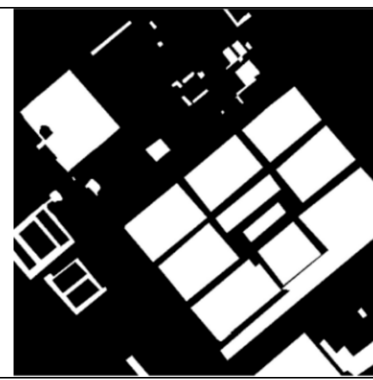
WORK ASSIGNED

Change detection in remote sensing images is a critical and challenging task, and its specific work refers to the quantitative analysis of multiple temporal remote sensing images for the same target area, determining the features and scope of surface changes and detecting the changed and unchanged parts. Remote sensing image change detection is utilised to detect illegal buildings, water area supervision, natural disaster assessment, urban planning expansion research, and military reconnaissance. Because of the increasing amount of data from remote sensing images and the increasing demand in this direction, manual comparison and analysis of the changing area appear time-consuming and laborious.
Due to factors such as -

- Seasons
- Solar illumination

imaging styles of different phases have huge differences, which make it difficult to solve the change detection task by computer vision.

Change detection is a unique task for remote sensing image processing, which can be regarded as a dichotomy problem of a region changing or not, as shown in Figure 28. Figure 28(a) shows the remote sensing images of a certain region of Yinchuan City in 2015, Figure 28(b) shows the remote sensing images of this region in 2017, and Figure 28(c) shows the change label of this region, where **black indicates that the location has not changed and white indicates that the location has changed**. The task of change detection is to identify the changing areas in different phases.

		
Figure 14. 28. (a) Area in 2015	Figure 15. 28. (b) Area in 2017	Figure 16. 28. (c) Change Map

DESCRIPTION OF THE DATASET

Onera Change Detection Dataset (OSCD): OSCD is a benchmark for bitemporal urban CD based on multispectral Sentinel-2 images [3]. It contains manual annotations of binary changes for 24 cities across the globe where 14 are used for training and 10 for testing. The labels focus on urban changes such as newly constructed buildings and natural changes such as a sea-level rise or differences in vegetation are not annotated. The two images per city are selected to be cloud-free and are generally taken about 1-3 years apart.

		
Figure 17. 29 (a) City: Dubai, UAE - Original Image	Figure 18. 29. (b) City: Dubai, UAE - Altered Image	Figure 19. 29. (c) Corresponding Change Map of Dubai

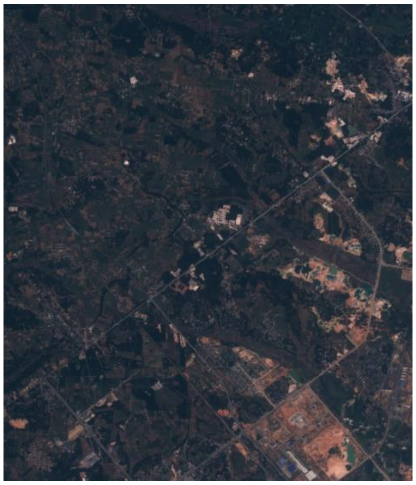
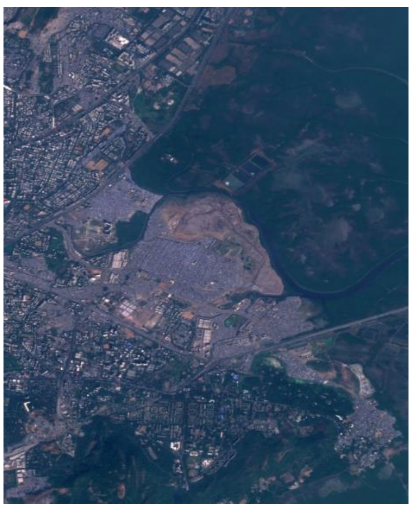
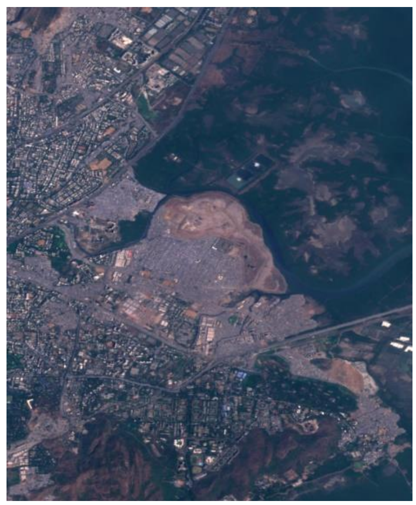
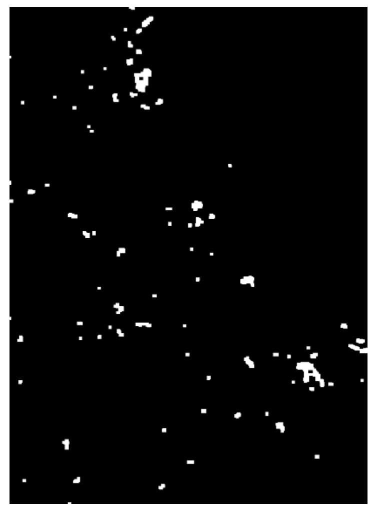
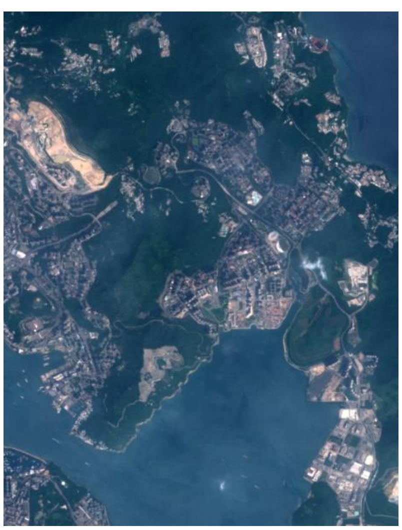
Our model has the following parameters [1] [2]:

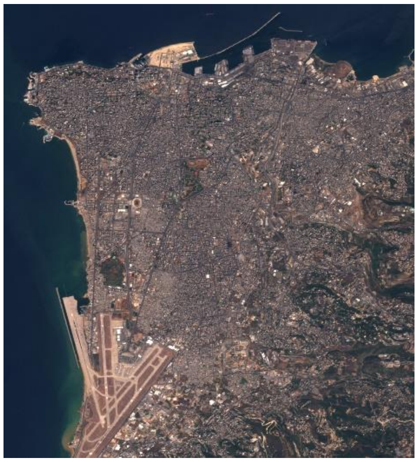
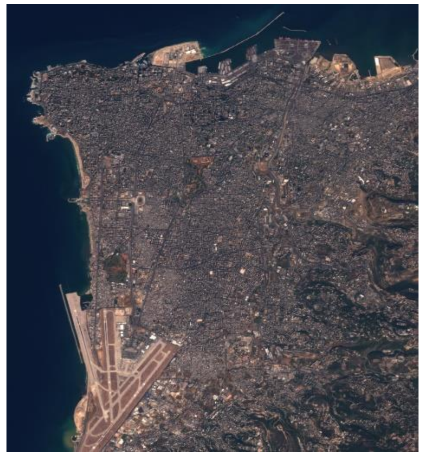
1. Maximum neighbourhood size: n max
2. Step-Size of the ensemble: s
3. Filter Size of Morphological Operations: p
4. Voting Threshold: $0 \leq v \leq 1$ (Impacts Probability of the Change Threshold)
5. Otsu Factor: $0 < v < 5$ (Impacts Sensitivity of the model to change)

Initial Results of our Model on Satellite Images from the Dataset with parameter **voting_threshold = 0.95** and **otsu_factor = 1.9** –

Table 3. Table showing Results from SiROC Model

Original Image	Altered Image	Change Map
Las Vegas		
		
Beihai		

		
Mumbai		
		
Hong Kong		
		

Beirut		
		

Principal Component Analysis (PCA) and K-means clustering techniques over different images to detect changes in multi-temporal images satellite imagery.

Automatic change detection in images of a region acquired at different times is one of the most interesting topics of image processing. Such images are known as multi-temporal images. Change detection involves the analysis of two multi-temporal satellite images to find any changes that might have occurred between the two-time stamps.

It is one of the major utilization of remote sensing and finds application in a wide range of tasks like defence inspections, deforestation assessment, land use analysis, disaster assessment and monitoring of many other environmental/man-made changes.

Methodology - An ***unsupervised method*** for change detection is used in this code. It involves the automatic analysis of the change data, i.e. the difference image, constructed using the multi-temporal images. A different image is the pixel-by-pixel subtraction of the 2 images. Eigenvectors of pixel blocks from the difference image will then be extracted by Principal Component Analysis (PCA).

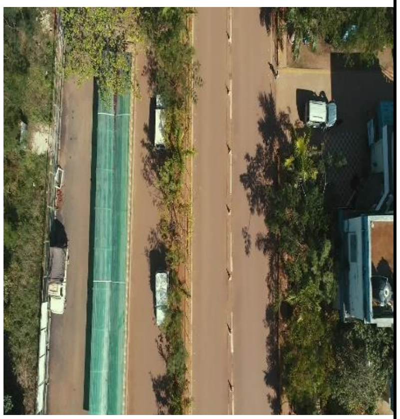
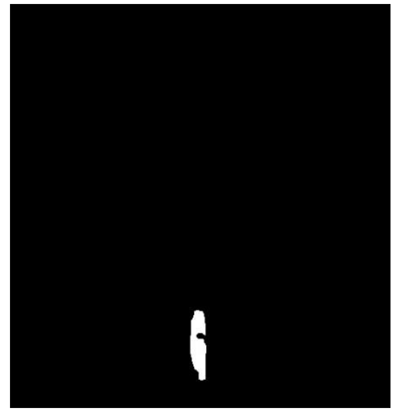
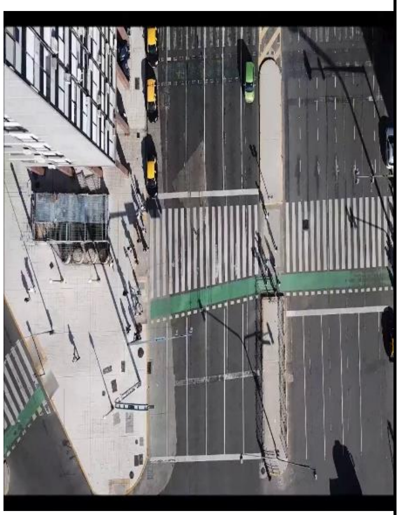
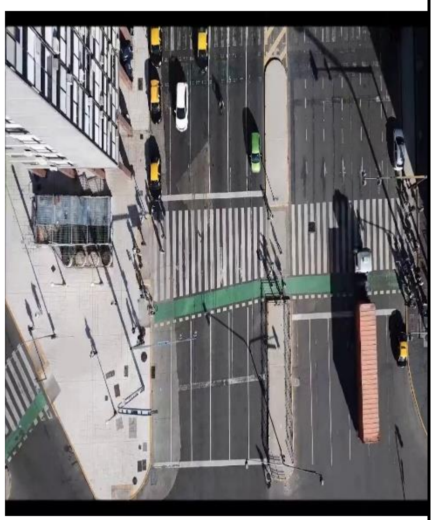
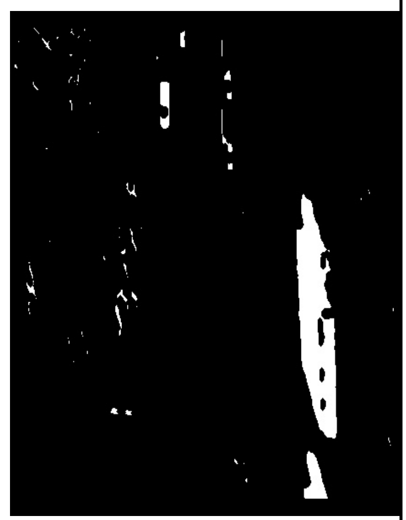
Key improvements over previous methods -

1. Significant reduction in image noise yielding much more accurate change maps.
2. Reduced processing time to achieve results.

Requirements to run this code -

1. Needs **`SciPy==1.2.0`**, which isn't available on Windows (`code was executed on Google Colab`) [4] [5].
2. The length and width of image pairs must be the same.
3. Image pairs must be grayscale rather than RGB [6].

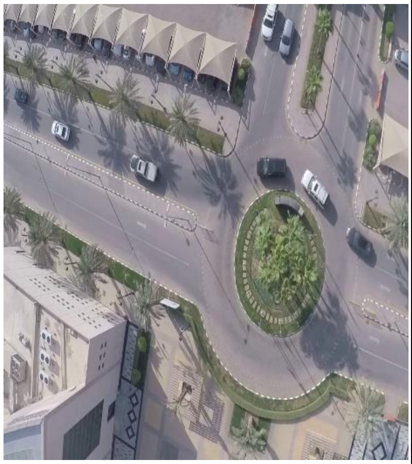
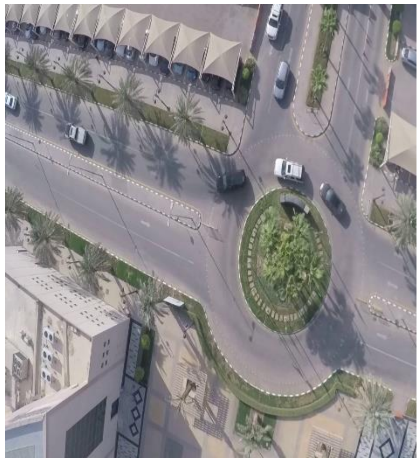
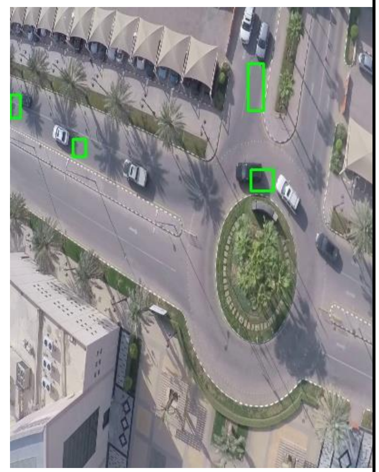
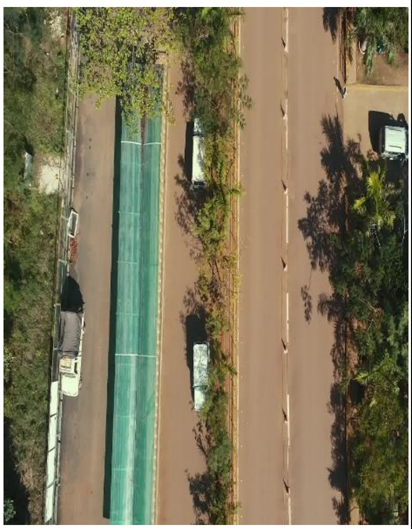
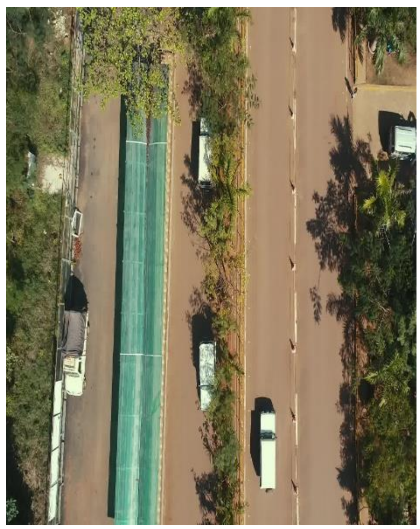
Table 4. K-Means PCA Results

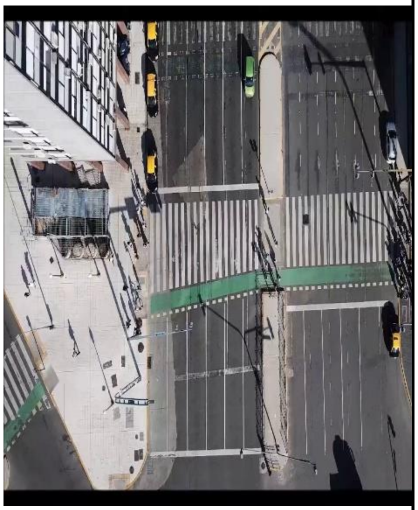
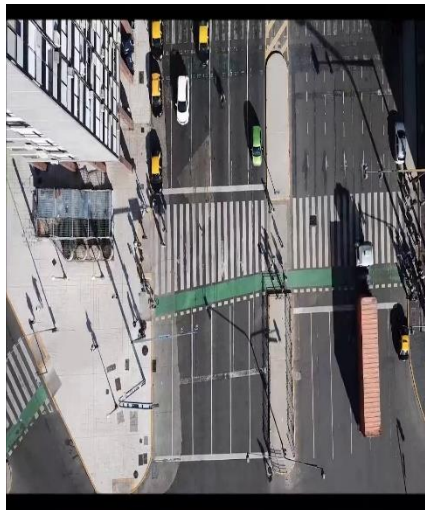
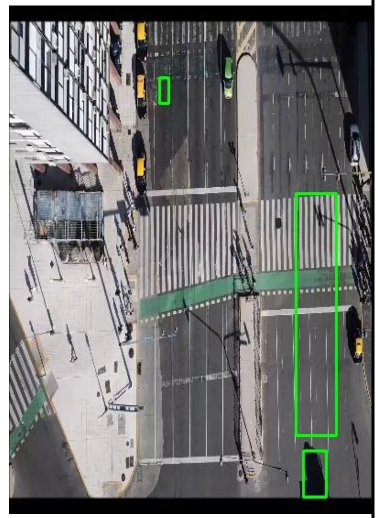
	Lane Traffic	
		
	Street Crossing	
		

Addition of Bounding Box

- The bounding box is an imaginary rectangle drawn around a given object and it serves as the region of interest.
- To draw a bounding box around an object in the given image, make use of a function called the `findContours( )` and `rectangle( )` in OpenCV.

Table 5. Change Maps with Bounding Boxes

Roundabout Traffic		
		
	Lane Traffic, Threshold = 3	
		
	Street Crossing, Threshold = 3	

		
	Dataset, Threshold = 2	
		

REFERENCES

[1] SiROC - Spatial Context Awareness for Unsupervised Change Detection in Optical Satellite Images - [GitHub]

[2] L. Kondmann, A. Toker, S. Saha, B. Schölkopf, L. Leal-Taixé and X. X. Zhu, "Spatial Context Awareness for Unsupervised Change Detection in Optical Satellite Images," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-15, 2022 - IEEE, PDF

[3] OSCD - Onera Satellite Change Detection Dataset - [IEEE Dataport]