iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** V **Month of publication:** May 2024

DOI: https://doi.org/10.22214/ijraset.2024.62073

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

Performance Analysis: Object Detection System for Low Illumination

Meher Tanushri Balasaheb[1], Mhaske Sakshi Sanjay[2], More Shubham Somnath[3], Ugale Sakshi Bhaurao[4], Dr. N. R. Kale[5]
[1, 2, 3, 4]*Students,* [5]*Professor, Department of Information Technology, MET Institute of Engineering, Nashik, India.*

Abstract: ***Abnormal activity detection plays a very important role in surveillance applications. The existing research on surveillance for daytime has achieved better performance by detecting and tracking objects using deep learning algorithms. However, it is difficult to achieve the same performance for night vision mainly due to low illumination. Deep learning is a powerful machine learning technique in which the object detector automatically learns image features required for detection tasks. It is required to generate a model that detects objects under low illumination. The approach is to use thermal infrared images and detect external objects, if any, and classify whether it is human or animal in an isolation area. With the rapid growth of deep learning, more efficient techniques will be implemented to solve the problems of object detection using neural networks and deep learning.***

Keywords: Object Detection, RetinaNet, SSD, Thermal Infrared Images, YOLOv8.

I. INTRODUCTION

Object detection is an important aspect of computer vision technology that involves identifying objects in images for various applications such as surveillance, security, safety and military operations. Object detection in low-illumination environments is a challenging task due to the low quality of the images. With the advancements in deep learning algorithms, object detection has become more precise and accurate. However, for nighttime, the traditional methods for object detection using visible light cameras are limited as visible light cameras often struggle to capture images in low light illumination, leading to limited visibility and increasing difficulties in object detection. On the other hand, thermal imaging can detect infrared radiation emitted by objects, providing clear images in the dark. Thermal imaging can be used for night vision detection because it captures the heat signature of objects instead of visible light, making it an effective solution for low light illuminations. Unlike visible light cameras, thermal imaging cameras are not affected by darkness, fog or other atmospheric conditions that can obstruct the view. Instead, they can provide a clear image of the environment and the objects within it, even in complete darkness. In this research paper, we propose a method that can use deep learning algorithms for object detection in night vision using thermal infrared images. We suggest developing a model that can accurately detect objects in low light illuminations, making it useful for a wide range of applications [1]. We will implement deep learning algorithms such as YOLO (You Only Look Once), SSD (Single Shot Detector) and RetinaNet, particularly for night vision using thermal infrared images and compare them using various performance metrics, including accuracy, precision and recall. This proposal will provide insights into the effectiveness of using deep learning algorithms for object detection in night vision using thermal infrared images.

II. PROBLEM STATEMENT

Design and implement a real-time object detection system optimized for night vision scenarios using deep learning techniques is crucial for enhancing surveillance and safety measures. The system must accurately detect and classify three main categories of objects: humans, animals, vehicles. To construct a powerful deep learning framework capable of accurately detecting objects in low-light or night vision scenarios. To find the most accurate algorithm that can detect using thermal infrared images. To develop an SMS alert system that provide category-wise counts of detected objects.

The following are the goals:

1) To develop an object detection system for night vision using deep learning algorithms.
2) To implement two popular deep learning algorithms - YOLOv8 and SSD for object detection in low light conditions.
3) To compare the performance of YOLOv8 and SSD on the FLIR dataset and evaluate their performance using mAP, precision, recall, and F1 score.
4) To identify the best-performing algorithm and recommend it for object detection in night vision applications.

©IJRASET: All Rights are Reserved

III. SYSTEM ARCHITECTURE

1) Input Image: The project starts with an input image that is captured using a night vision camera.
2) Preprocessing: The input image is annotated into .txt and .xml using bounding boxes which is the required format for training YOLOv8 and SSD.
3) Feature Extraction: Two deep learning algorithms, YOLOv8 and SSD, are implemented to extract features from the preprocessed image. YOLOv8 and SSD are popular object detection algorithms that are widely used in computer vision applications.
4) Performance Comparison: The performance of YOLOv8 and SSD is compared based on several evaluation metrics, such as mean average precision (mAP), precision, recall, and F1 score. YOLOv8 is found to be better than SSD for object detection.
5) Object Detection: Once the object detection algorithm is selected, it is used to detect objects in the input image. This involves identifying the category of object present in the image and its probability.
6) Alert SMS: If a specific category of object is detected, the system sends an alert SMS using Twilio. The SMS contains the category-wise count of the detected objects. This feature allows users to receive real-time notifications about the presence of specific objects in the environment.
7) User Interface: The output image is displayed on a user interface that allows users to interact with the system. The user interface is a web application.

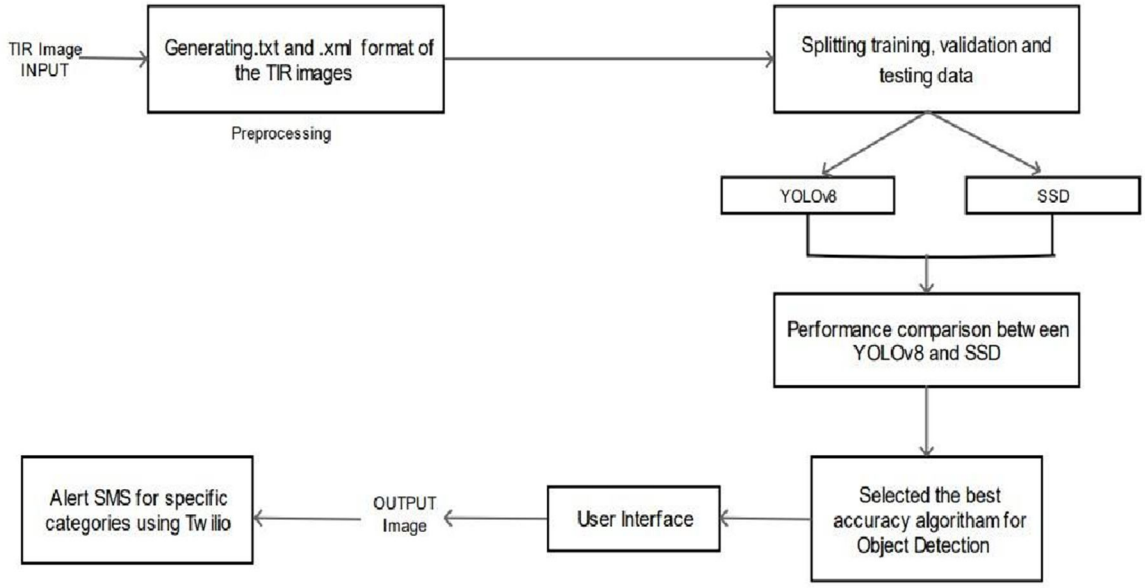


Figure 1: System Architecture

IV. METHODOLOGY

The idea proposed in this paper focuses on building robust models using deep learning algorithms for object detection in night vision. To find an accurate algorithm that can detect objects using thermal infrared images, we analysed different object detectors based on their performance parameters. Faster R-CNN, MRCNN[4], HOG[8] , YOLO V3[9] , YOLO V8, SSD[10], RetinaNet[7] achieved best results for object detection. We selected one stage detectors, namely YOLO, SSD and RetinaNet, as they have high inference speeds[3].

A. Data Collection and Pre-processing

We will gather different TIR(Thermal Infrared) videos, and further, they will be converted into image frames. Then these image frames can be resized according to different object detectors. YOLO keeps the aspect ratio safe without the need for explicit image resizing. Further data augmentation can be done on the resized images. Data augmentation is a technique where the size of the training dataset is increased by creating modified versions of existing data samples.

The purpose of data augmentation is to reduce overfitting and to increase the robustness and generalisation of the trained model by exposing it to a wider range of variations in the input data. Common methods for data augmentation include flipping, rotation, scaling, cropping and adding noise to images

B. YOLOv8

YOLO (You Only Look Once) is a popular object Detection algorithm that uses a single convolutional neural network to perform object detection and classification. Many versions of YOLO have been proposed, and in this paper, we have considered YOLOv8, the latest version of YOLO. YOLOv8 was developed by Ultralytics on 10 January 2023. It performs three tasks, namely object detection, image classification and instance segmentation. According to Ultralytics, YOLOv8 is fast, accurate, and easy to use, and it can be trained on large datasets. It has a new backbone network, a new anchor- free detection head, and a new loss function. The YOLOv8 detector can be tested in isolated areas using thermal infrared images in this experiment. As mentioned above, data augmentation is used for increasing the accuracy of the detector, but in YOLOv8, there is no need for data augmentation as it augments images during training automatically. Then we will annotate the images in the dataset using the LabelImg tool. LabelImg is open-source software that allows users to annotate images by drawing bounding boxes around objects in the image and then labelling those objects. After annotating the images, the annotations can be saved in YOLO format. We will obtain .txt files for every image. Each txt file will contain the class of object, its height, weight, x coordinate and y coordinate. Once the images are annotated, create a .yaml file which consists of the paths of the training and validation folder, the number of classes in the dataset and their names. Split the dataset along with its corresponding .txt files into 80%, 10%, and 10% for training, validation and testing, respectively. For implementing YOLOv8, the prerequisite is installing and importing Ultralytics, as YOLOv8 can be imported from it. Once all necessary things are done, the model can be trained. For training, it is essential to specify the hyperparameters, such as the number of epochs, image size, task, etc. After the model is trained, we can also obtain the time required for training per image which can be used as a parameter for comparing the object detectors. The validation can be performed on the 10% of the dataset, and confusion matrix, f1 score, recall curve and other evaluation parameters can also be obtained. These evaluation results can be saved for further analysis. Later, the model can be tested by providing some thermal infrared images, and the output can be saved. We can also provide a video for testing. The output will be an image or video with bounding boxes to the objects along with its class probability.

C. Single Shot Detector(SSD)

A Single Shot Detector is an object detector based on VGG16 architecture. Single Shot Detector uses bounding boxes of various aspect ratios, compares them with ground truth boxes to obtain confidence scores and extracts feature maps. SSD uses VGG16 for the extraction of the feature map. The input image is modified by adding the ground truth boxes around all the objects present in the image. These images are fed into the VGG16. In VGG16, there are thirteen convolutional layers, five Max Pooling layers, and three Dense layers though it has only sixteen weight layers. The convolution layers are of a 3x3 filter with stride 1 and always use the same padding and maxpool layer of a 2x2 filter. Max pooling reduces the feature map dimension. VGG16 performs strongly and does image classification tasks. The architecture of the SSD consists of 6 convolutional layers preceded by VGG16, as shown in Fig.1. This set of auxiliary convolution layers helps feature extraction at multiple scales. It ultimately decreases the input size to the next corresponding layers[10]. For each object, we get 8732 bounding boxes. From these 8732 bounding boxes top 200 predictions are made based on the calculated confidence score of each box. The goal is to find the box that best fits the ground truth box. Intersection over Union (IoU) is calculated for the predicted bounding boxes and the ground truth boxes.[27]. Non-max suppression is used to filter these boxes and remove the duplicates. [12] SSD takes input images with the ground truth boxes. SSD achieves better accuracy using various aspect ratios than other object detectors like YOLO. SSD has various applications like video forensics, legal investigations, landmark detections, and many more. In order to start implementation, we will use the LabelImg tool for annotation to obtain the images with ground truth boxes and the corresponding .xml files for each image. The collected thermal infrared images will be split into train, validation and test set into ratios of 80%, 10% and 10%, respectively— Import Tensorflow object detection API. Create a .csv file storing the details of the .xml files in comma-separated format for each training and validation set. Label_map.pbtxt file will be automatically generated, which stores the classes using the .xml files. Upload the generate_tfrecord.py file and generate TFrecords for each train and validation set. TFRecord format is a simple format a sequence of binary records. Load the tensorboard and train the model. Then validation is done using 10% data and finds confusion matrix, f1 score, recall curve and other evaluation parameters. Save the results of the evaluation. After validation, test the trained object detection model based on the test images, record the outputs and save them along with the bounding box and its class probability.

D. RetinaNet

RetinaNet is a single-shot object detection model developed by Facebook AI Research (FAIR). A deep neural network architecture extracts features from an image using a backbone network.

Then it uses two separate branches to predict the class probabilities and the bounding box locations for each object in the image. RetinaNet made improvements over existing single-stage object detection models, i.e. Feature Pyramid Networks(FPN)[13] and Focal Loss. The various object detectors evaluate only a few locations of an image, and background objects are left, which leads to class imbalance problems. RetinaNet object detector uses Focal Loss to fill in for the class imbalances and inconsistencies, thus improving the speed of RetinaNet. As shown in Fig.2, the architecture of RetinaNet breaks down into three components[14]

Backbone Network(Bottom-up pathway + Top down the pathway with lateral connections) Bottom-up pathway used in RetinaNet is Resnet. It is used for feature extraction, where the network starts from lower-level features and builds up to more complex features through multiple layers. This allows the network to learn the hierarchy of features.

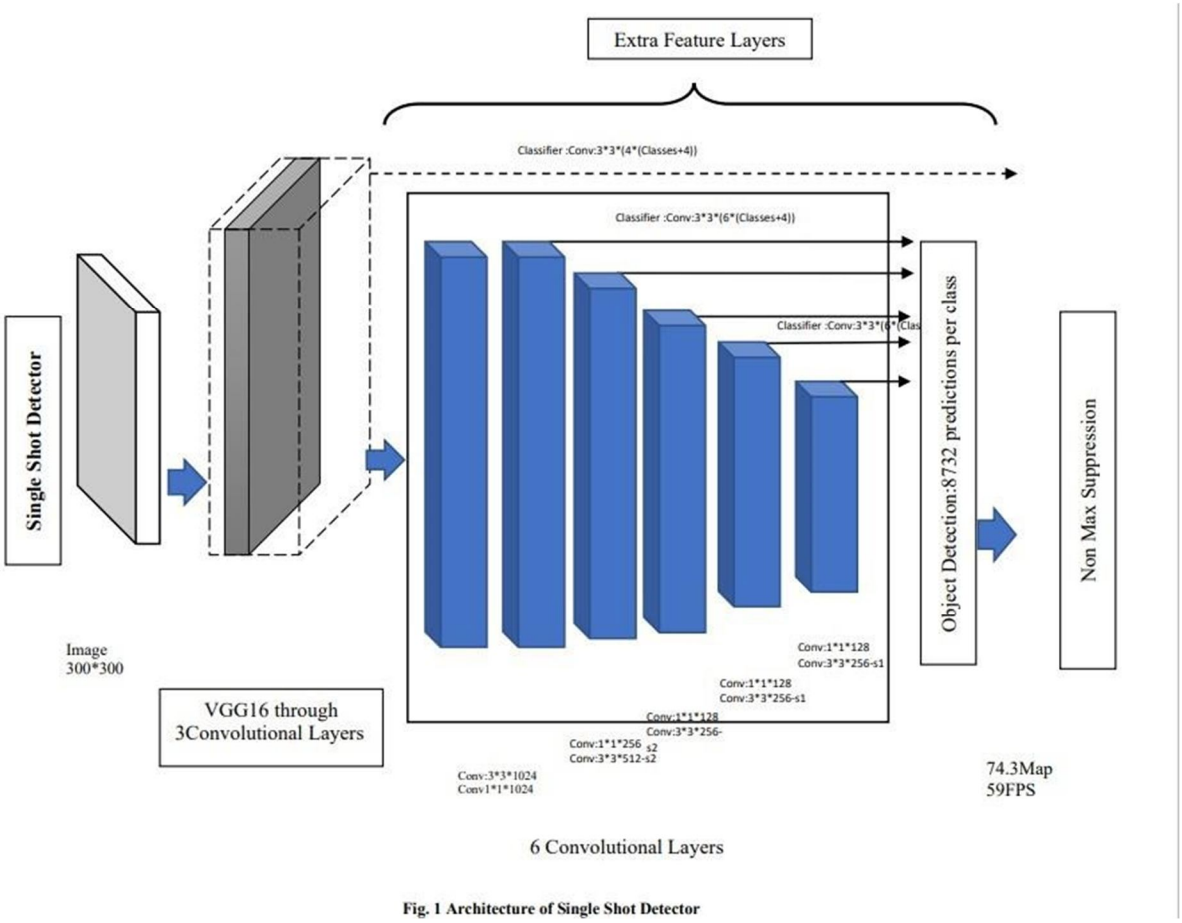


Fig. 1 Architecture of Single Shot Detector

E. Top-down Pathway with Lateral Connections

The top-down pathway means that the network starts with high-level features and refines them through multiple layers to obtain the final object detection. The lateral connections allow information to flow between different layers of the network, enabling the network to make predictions based on both high-level and low-level features. This topdown and lateral connections combination helps RetinaNet balance the trade-off between accuracy and computational efficiency. For this, RetinaNet uses Feature Pyramid Network(FPN). 3.4.2. Sub-network for Object Classification In RetinaNet, the sub-network responsible for object classification is called the "classification subnet". The classification subnet takes the feature maps generated by the backbone network. It produces a set of class scores for each anchor box, indicating the likelihood of each anchor box containing a specific object class. The classification subnet typically consists of several fully connected (FC) layers, sometimes preceded by a few convolutional (Conv) layers to reduce the number of feature channels and increase the spatial resolution.

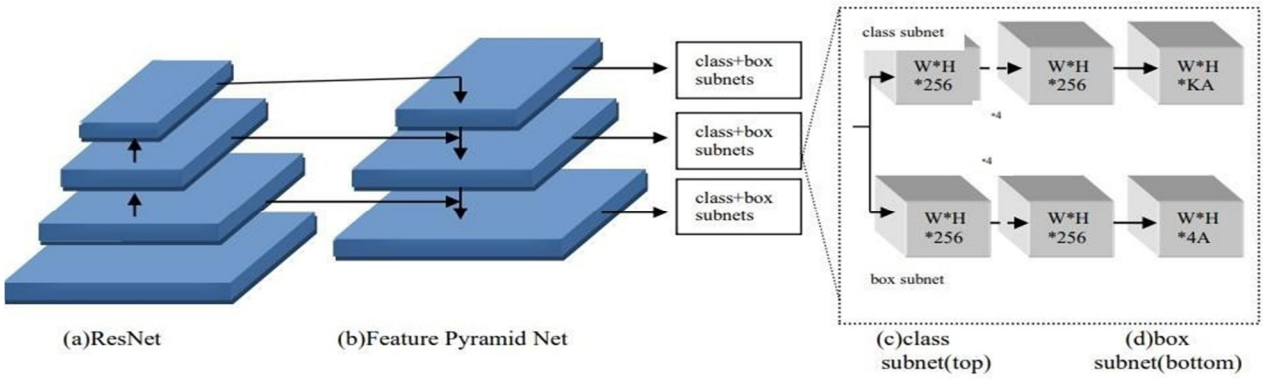


Fig. 2 Architecture of RetinaNet

The final FC layer produces the class scores, which are used along with the anchor boxes and the bounding box regression subnet to produce the final object detection results.

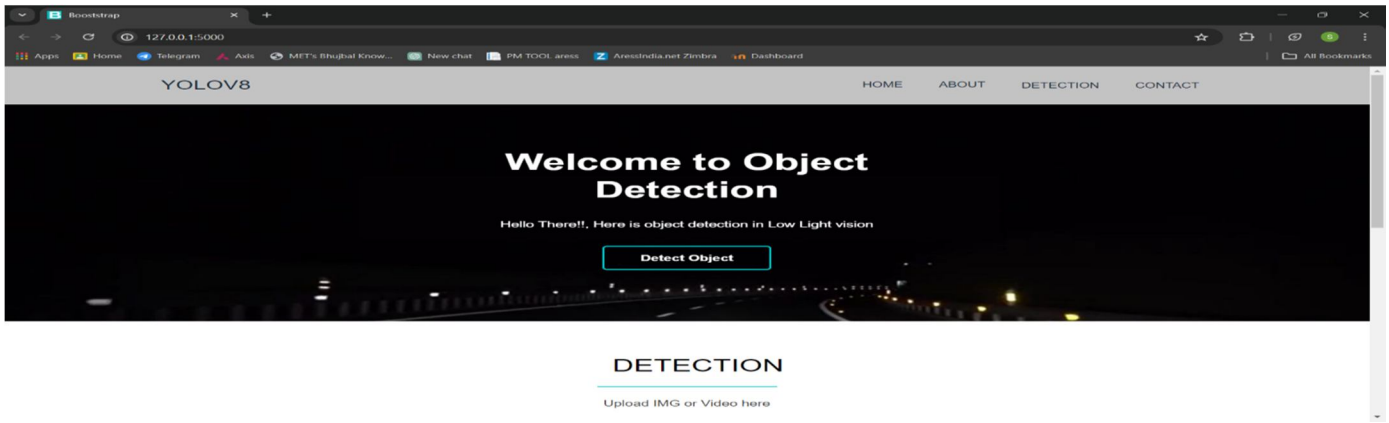
F. 5zSub-network for Object Regression

In RetinaNet, the regression sub-network is parallel Network (FPN) in a parallel manner. The sub-network for object regression is a fully connected layer that inputs features extracted by the Feature Pyramid Network (FPN). It outputs the bounding box coordinates for each object in the image. The sub-network uses a combination of anchor boxes and regression to predict the object-bounding boxes. The anchor boxes serve as prior knowledge, and the regression values refine the anchor boxes to be more accurate. The loss function used to train the sub-network penalises predicted bounding boxes far from the ground-truth boxes, encouraging the network to learn to generate accurate predictions. For implementing RetinaNet, the first step that will be carried out is annotating the dataset using the LabelImg tool. Along with the YOLO format, the annotations can also be saved as .xml files in PASCAL VOC format, which RetinaNet needs. Once the images are annotated, we will create a pascal_voc.py file and define the dataset classes in the file. Later, we will split the dataset into training, validation and testing in an 80:10:10 ratio. Then we will produce train.txt and test.txt files containing lists of train and test files, respectively. After this step, we will download the pretrained weights, and the model can be trained. The argument we can pass while training the model can be batchsize, steps, epochs, weights, etc. Once the model is trained, we can get a weight file. Load it and build an inference model. Afterwards, the model will detect and evaluate objects in the images.

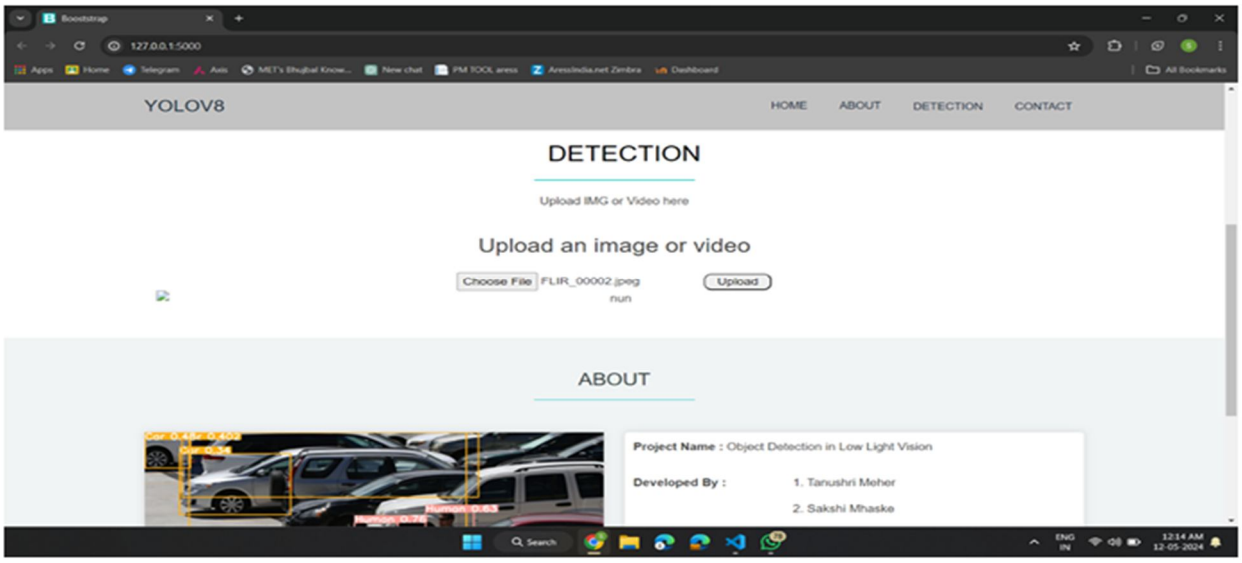
V. RESULT ANALYSIS

The proposed system was successfully tested to show its effectiveness and achievability. It reduces the manpower, time of college administrators and paper work. It also reduces the efforts of the students to travel all the way to college for enquiry purposes. In this paper we have a developed a chatbot which will interact with the users and provide all the college related information. The student/parent and the college admin are interacted through a chatbot. The questions which are not answered by the chatbot will be updated by the college admin.

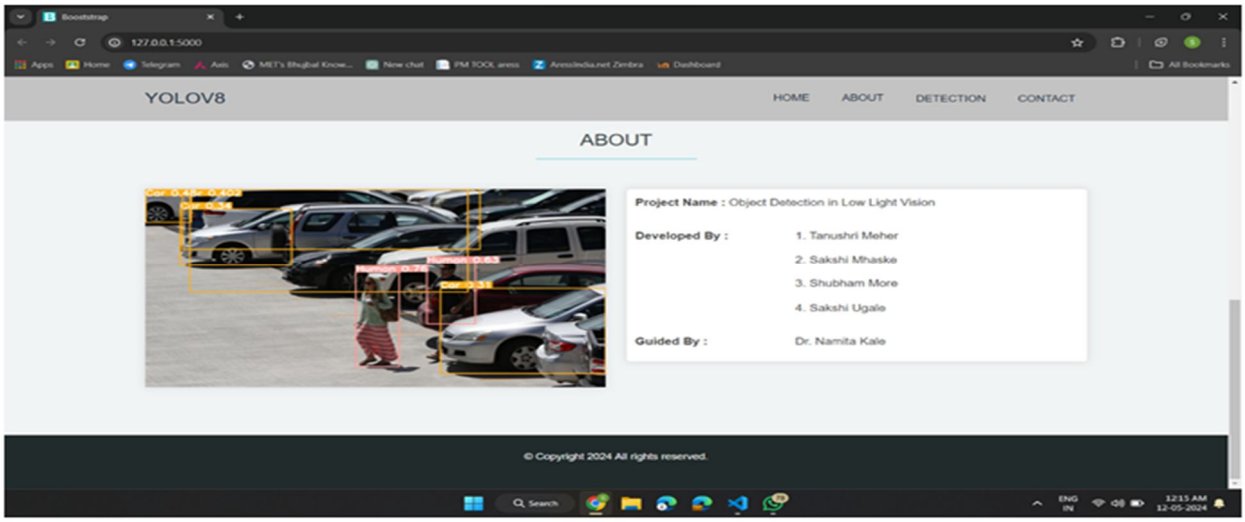
1) Home page of object detection.



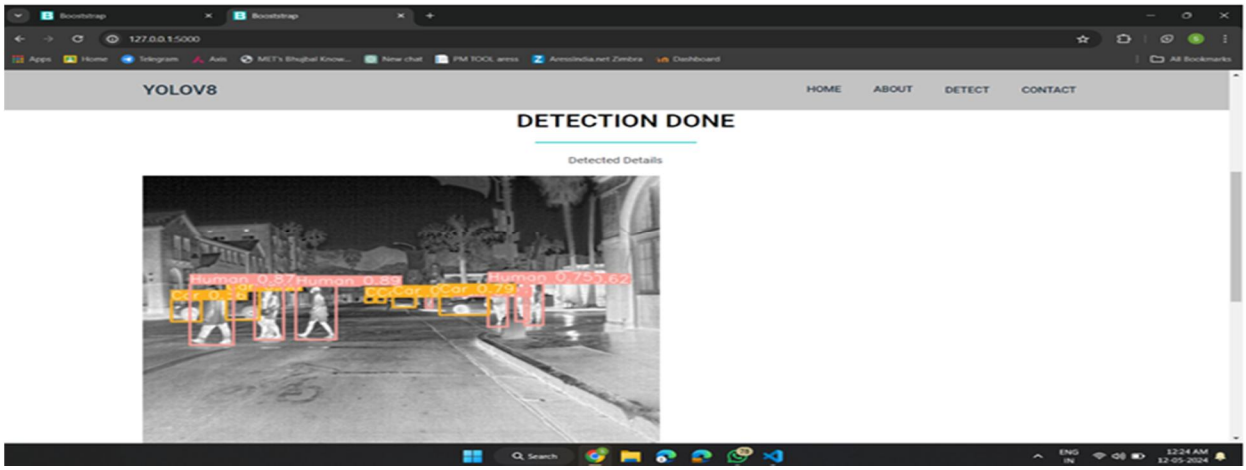
2) *Upload an image or video.*



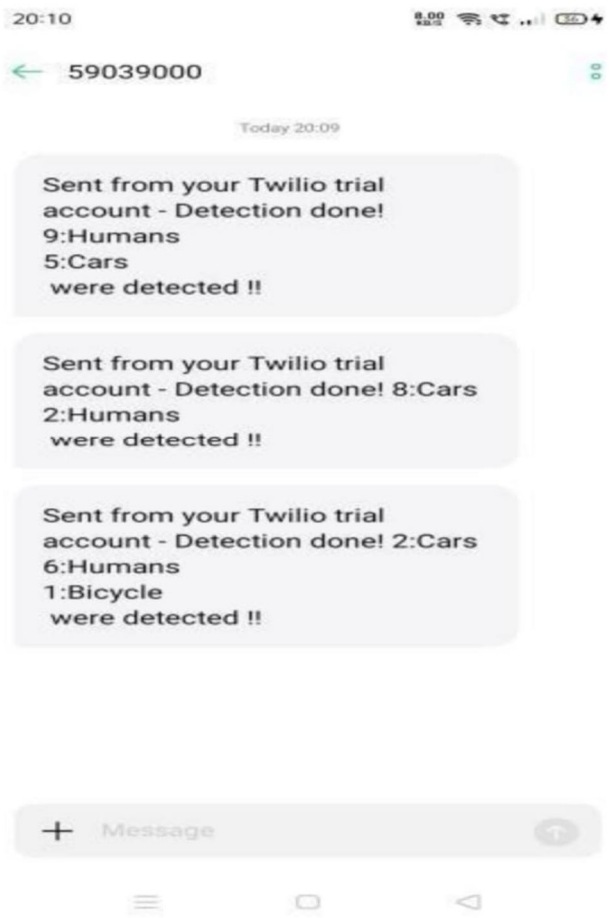
3) *YOLOV8 Detection*



4) *Result of uploaded image or video.*



5) .SMS Alert.



VI. APPLICATIONS

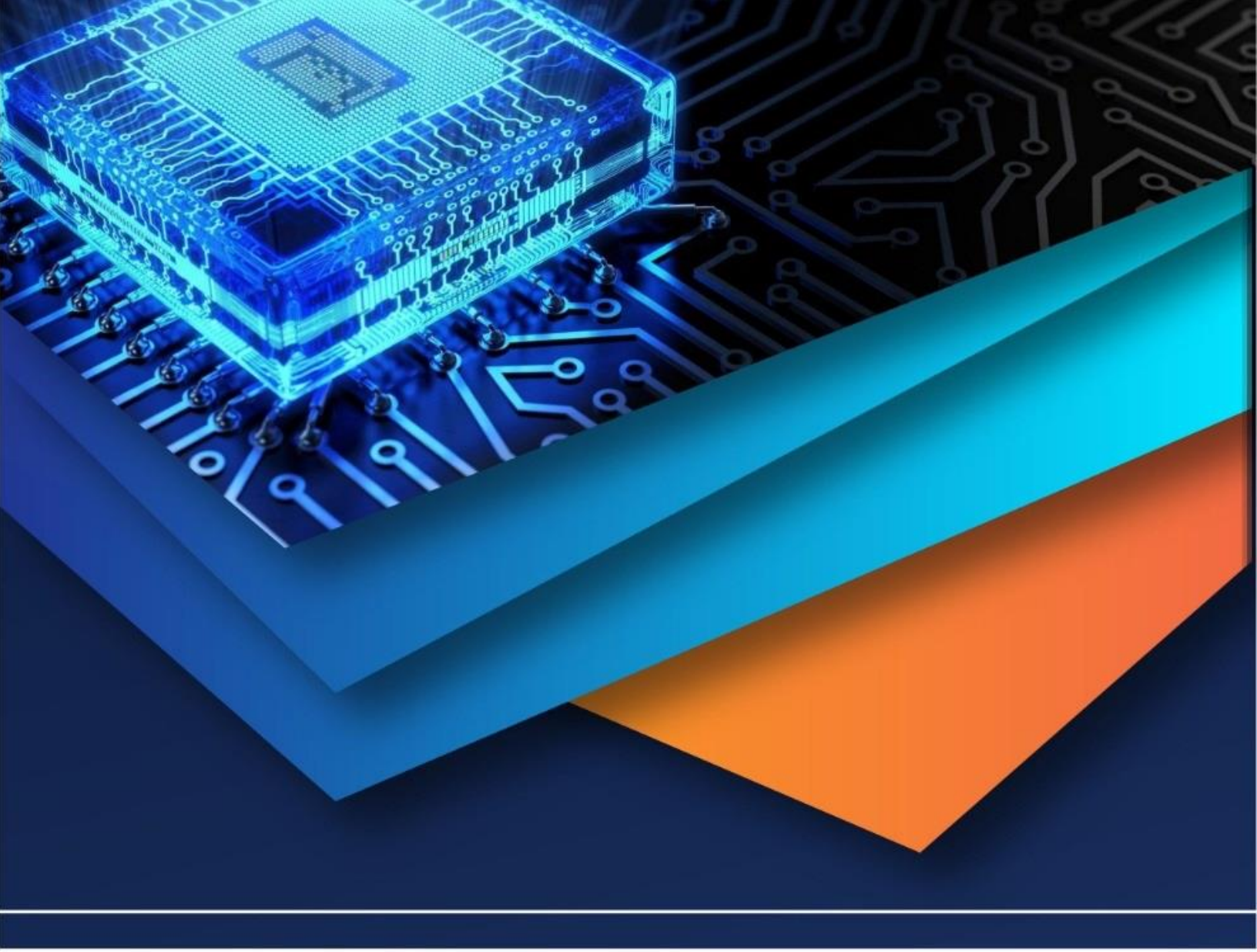
1) Surveillance and Security: The system can be used for surveillance and security purposes where immediate action is required upon detection of particular objects.
2) Automotive Industry: The system can be integrated into vehicles to detect objects on the road and provide alerts to drivers to avoid collisions.
3) Industrial Automation: The system can be used in manufacturing plants to detect objects on the assembly line and automate the production process.
4) Robotics: The system can be used in robotics for object recognition and detection, enabling robots to navigate and interact with their environment more effectively.
5) Medical Imaging: The system can be used in medical imaging to detect anomalies and abnormalities in medical images, aiding in the diagnosis of diseases.

VII. CONCLUSION

YOLOv8,SSD, and RetinaNet being one-stage detectors, may increase the speed of object detection. Furthermore, the models in this paper can be implemented, and the performance of these models can be compared to get the best object detector model for night vision. Also, the best model can be used for real-time object detection. Presently, we are working on the implementation of these models, and the results will soon be published in the paper.

REFERENCES

[1] Yuxuan Xiao et al., "Making of Night Vision: Object Detection Under Low-Illumination," IEEE Access, vol. 8, pp. 123075-123086, 2020. Crossref, https://doi.org/10.1109/ACCESS.2020.3007610
[2] Muhammad Javed Iqbal et al., "Real-Time Surveillance Using Deep Learning," Security and Communications Networks, 2021.
[3] Crossref, https://doi.org/10.1155/2021/6184756
[4] Aditya Lohia et al., Bibliometric Analysis of One-stage and Two-stage Object Detection, University of Nebraska-Lincoln, 2021.
[5] Heena Patel, and Kishor P. Upla, "Night Vision Surveillance: Object Detection using Thermal and Visible Images," 2020 International Conference for Emerging Technology (INCET), 2020. Crossref, https://doi.org/10.1109/INCET49848.2020.9154066
[6] K. R. Akshatha et al., "Human Detection in Aerial Thermal Images Using Faster R-CNN and SSD Algorithms," Electronics, vol. 11, no. 7, p. 1151, 2022. Crossref, https://doi.org/10.3390/electronics11071151
[7] Mate Kristo, Marina Ivasic-Kos, and Miran Pobar, "Thermal Object Detection in Difficult Weather Conditions Using YOLO," IEEE Access, vol. 8, pp. 125459-125476, 2020. Crossref, https://doi.org/10.1109/ACCESS.2020.3007481 class+box subnets class+box subnets class+box subnets W*H *256 W*H *256 W*H *256 W*H *256 W*H *KA W*H *4A *4 *4 class subnet box subnet Dipali Bhabad et al. / IJCTT, 71(2), 87-92, 2023 92
[8] Mohanad Al-Hasanat et al., "RetinaNet-based Approach for Object Detection and Distance Estimation in an Image," International Journal on Communications Antenna and Propagation (IRECAP), vol. 11, no. 1, p. 19, 2021. Crossref, http://dx.doi.org/10.15866/irecap.v11i1.19341
[9] Navneet Dalal, and Bill Triggs, "Histograms of Oriented Gradients for Human Detection," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. Crossref, https://doi.org/10.1109/CVPR.2005.177
[10] Qasem Abu Al-Haija, Manaf Gharaibeh, and Ammar Odeh, "Detection in Adverse Weather Conditions for Autonomous Vehicles via Deep Learning," AI Journal, vol. 3, no. 2, pp. 303-317, 2022. Crossref, https://doi.org/10.3390/ai3020019
[11] Wei Liu et al., "SSD: Single Shot MultiBox Detector," Computer Vision and Pattern Recognition, 2015. Crossref, https://doi.org/10.48550/arXiv.1512.02325
[12] M. Maheswari, M. S. Josephine, and V. Jeyabalaraja, "YOLO Architecture-based Object Detection for Optimizing Performance in Video Streams," International Journal of Engineering Trends and Technology, vol. 70, no. 11, pp. 187-196, 2022. Crossref, https://doi.org/10.14445/22315381/IJETT-V70I11P220
[13] Jan Hosang, Rodrigo Benenson, and Bernt Schiele, "Learning Non-maximum Suppression," Computer Vision and Pattern Recognition, 2017. Crossref, https://doi.org/10.48550/arXiv.1705.02950
[14] Tsung-Yi Lin et al., "Feature Pyramid Networks for Object Detection," Computer Vision and Pattern Recognition, 2016. Crossref, https://doi.org/10.48550/arXiv.1612.03144
[15] Tsung-Yi Lin et al., "Focal Loss for Dense Object Detection," Computer Vision and Pattern Recognition, 2017. Crossref, https://doi.org/10.48550/arXiv.1708.02002
[16] Aashish Bhandari et al., "Image Enhancement and Object Recognition for Night Vision Surveillance," ICTRCET 18 at Bangaluru, 2018.
[17] Rupesh P.Raghatate et al, "Night Vision Techniques and Their Applications," International Journal of Modern Engineering Research (IJMER), vol. 3, no. 2, pp. 816-820, 2013.
[18] D. Malarvizhi, V. Lavanya, and M. Nivetha priya, "Night Vision Technology," International Journal for Scientific Research and Development, vol. 3, no. 8, 2017.
[19] Kai Hu et al., "A Marine Object Detection Algorithm Based on SSD and Feature Enhancement," Complexity, 2020. Crossref, https://doi.org/10.1155/2020/5476142
[20] Sreehari Patibandla, Maruthavanan Archana, and Rama Chaithanya Tanguturi, "Object Tracking using Multi Adaptive Feature Extraction Technique," International Journal of Engineering Trends and Technology, vol. 70, no. 6, pp. 279-286, 2022. Crossref, https://doi.org/10.14445/22315381/IJETT-V70I6P229
[21] Pavan Sai Vemulapalli et al., "Multi-object Detection in Night Time," Asian Journal of Convergence in Technology, vol. 5, no. 3, pp. 1- 7, 2019.
[22] Tanvir Ahmad et al., "Object Detection through Modified YOLO Neural Network," Scientific Programming, 2020. Crossref, https://doi.org/10.1155/2020/8403262
[23] Muskan Choudhary et al., "Object Detection Using YOLO Models," International Research Journal of Engineering and Technology, vol. 9, no. 5, pp. 3785-3789, 2022.
[24] Joaquın Royo Miquel et al., "RetinaNet Object Detector based on Analog-to-Spiking Neural Network Conversion," Image and Video Processing, 2021. Crossref, https://doi.org/10.48550/arXiv.2106.05624
[25] R. Manasa, K Karibasappa, and J. Rajeshwari, "Autonomous Path Finder and Object Detection using an Intelligent Edge Detection Approach," SSRG International Journal of Electrical and Electronics Engineering, vol. 9, no. 8, pp. 1-7, 2022. Crossref, https://doi.org/10.14445/23488379/IJEEE-V9I8P101
[26] Tsung-Yi Lin et al., "Feature Pyramid Networks for Object Detection," in Processing IEEE Conference on Computer Visison Pattern Recognition (CVPR), 2016. Crossref, https://doi.org/10.48550/arXiv.1612.03144
[27] Tsung-Yi Lin et al., "Feature Pyramid Networks for Object Detection," in Processing IEEE Conference on Computer Visison Pattern Recognition (CVPR), 2016. Crossref, https://doi.org/10.48550/arXiv.1612.03144
[28] Wu Zheng et al., "CIA-SSD: Confident IoU-Aware Single-Stage Object Detector from Point Cloud," in Processing IEEE Conference on Computer Visison Pattern Recognition (CVPR), 2020. Crossref, https://doi.org/10.48550/arXiv.2012.03015
[29] Amit Tiwari, and Jalaj Gupta, "A Simulation of Night Vision Technology Aided with AI," AKGEC International Journal of Technology, vol. 12, no. 1, 2022. [29] Pranav Adarsh, Pratibha Rathi, and Manoj Kumar, "YOLO v3-Tiny: Object Detection and Recognition using One Stage Improved Model," 2020 6th International Conference on Advanced Computing and Communication Systems, 2020. Crossref, https://doi.org/10.1109/ICACCS48705.2020.907431



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)