Real-Time Object Detection for Night Vision Scenarios Using Deep Learning

Tanushri Meher[1], Sakshi Mhaske[2], Shubham More[3], Sakshi Ugale[4]
[1,2,3,4]*Department of Information Technology Engineering, MET's Institute of Engineering, Nashik, Maharashtra, India.*

---***---

Abstract - Abnormal activity detection is crucial for effective surveillance applications. While daytime surveillance has benefited from deep learning algorithms for object detection and tracking, achieving similar performance in night vision is challenging due to low illumination. Deep learning, a powerful machine learning technique, automatically learns image features for detection tasks. This project aims to develop a model capable of detecting objects under low illumination, specifically using thermal infrared images. The proposed approach focuses on utilizing thermal infrared images to detect external objects. The performance of the developed model is evaluated through experimentation and analysis, including comparative studies with existing methods. This assessment provides insights into the advantages and limitations of using deep learning for object detection in low-light conditions. With the continuous growth of deep learning, more efficient techniques have been implemented to address the challenges associated with object detection using neural networks and deep learning.

***Keywords* -** Object Detection, Thermal Infrared Images, Neural Network.

1. INTRODUCTION

The project aims to tackle a significant challenge or address a specific problem in Object detection for night vision using deep learning algorithms. This project endeavors to explore innovative solutions, develop practical applications, or conduct in-depth research to contribute to the advancement of knowledge and bring about tangible outcomes. The purpose of this project is to compare different deep-learning algorithms for Night time and evaluate the performance of the models. Through rigorous analysis, experimentation, and implementation, the project aims to provide valuable insights, novel approaches, or practical solutions to address the identified problem or fulfil a specific need in the field. Object detection is a computer vision technique that in volves identifying and localizing multiple objects within an image or a video. It combines the tasks of object recognition (classification) and object localization (determining the position and extent of the objects). Object detection algorithms typically leverage deep learning to produce meaningful results. When humans look at images or videos, we can recognize and locate objects of interest within a matter of moments. The goal of object detection is to replicate this intelligence using a computer. Surveillance has become a critical task in recent times mainly due to the increase in crime rates. To develop a real-time object detection system using deep learning algorithms, specifically YOLOv8 and SSD, for night vision scenarios. The system is designed to detect specific categories of objects, and once these objects are detected, it will utilize the Twilio integration to send SMS alerts, providing a category-wise count of the detected objects. The project will involve preprocessing the image data, implementing the object detection algorithms, comparing their performance, and integrating the Twilio API for SMS alerts. The system can be applied in various scenarios such as surveillance, security, and monitoring in low-light conditions.

A. Objective

Objective Project Objective.

To build the Deep learning Model using YOLOv8 and SSD. To evaluate the models using metrics such as mean average precision (MAP),precision, recall, and F1 score. To compare the performance of YOLOv8 and SSD and identify the best-performing algorithm for object detection in low-light conditions. To enhance the system, the highest-performing algorithm will be integrated with the Twilio API. This integration will enable the system to send alert messages, providing a category-wise count, when specific categories of objects are detected in the images.

2.PROBLEM STATEMENT

The aim of this project is 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 provides category-wise counts of detected objects.

3.LITERATURE SURVEY

The paper [1] Human Detection in Aerial Thermal Images Using Faster R-CNN and SSD Algorithms was published in April 2022. In this paper, the performance of Faster RCNN and SSD was analysed. The Faster R-CNN ResNet50 model achieved the highest MAP while SSD achieved the highest detection speed. The limitation was that it required high computational power for preprocessing.

The paper [2] Object detection from UAV thermal infrared images and videos using YOLO models was published in year August2022. In this paper, the YOLO models were designed to extract features from groundbased TIR images and videos. Inaccurate detection due to the effects of observation angles in UAVs were the major drawback mentioned in the paper.

The paper [3] Thermal Object detection in Difficult weather conditions using YOLO was published in July 2020. This research paper undertook a comprehensive analysis, comparing the effectiveness of various object detectors including Faster R-CNN, Cascade R-CNN, SSD and YOLOv3, inaccurately detecting objects under challenging weather conditions. And it was found that YOLOv3 is faster than other detectors. The drawback was that it faced difficulty in detecting a variety of objects.

The paper [4] Night Vision Surveillance: Object Detection using Thermal and Visible Images was published in July 2020. The idea proposed an efficient object detection module using the fusion of thermal and visible images. Faster R-CNN was used for object detection and MR-CNN for object classification. The proposed method per forms better on real-time nightvision images.

4. PROPOSED METHODOLOGY

The following figure represents the steps involved in the proposed approach in a sequence manner:

YOLOv8:

You Only Look Once (YOLOv8) is a real-time object detection algorithm that predicts classes and bounding boxes for objects in an image in a single run. It identifies the center (bx, by), width (b_w), height (b_h), class value (c), and object probability (pc) for bounding boxes. Developed by Ultralytics, YOLOv8 features a new backbone network (CSPDarknet53) and anchor-free detection head, claiming speed, accuracy, and ease of use.

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 opensource 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.

CSPDarknet53, with 53 convolutional layers, employs cross-stage connections to enhance information flow and maintain accuracy with a relatively compact model size.

SSD (Single Shot Detector):

Single Shot Detector (SSD) uses a single neural network for object detection and classification in images. It employs a CNN for feature extraction and predicts bounding boxes for object classes, refining them through non-maximum suppression (NMS). SSD's advantages include speed, accuracy, and flexibility in detecting diverse object shapes and sizes. Its architecture involves 6 convolutional layers preceded by VGG16, aiding in multi-scale feature extraction.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 01 | January - 2024 SJIF Rating: 8.176 ISSN: 2582-3930

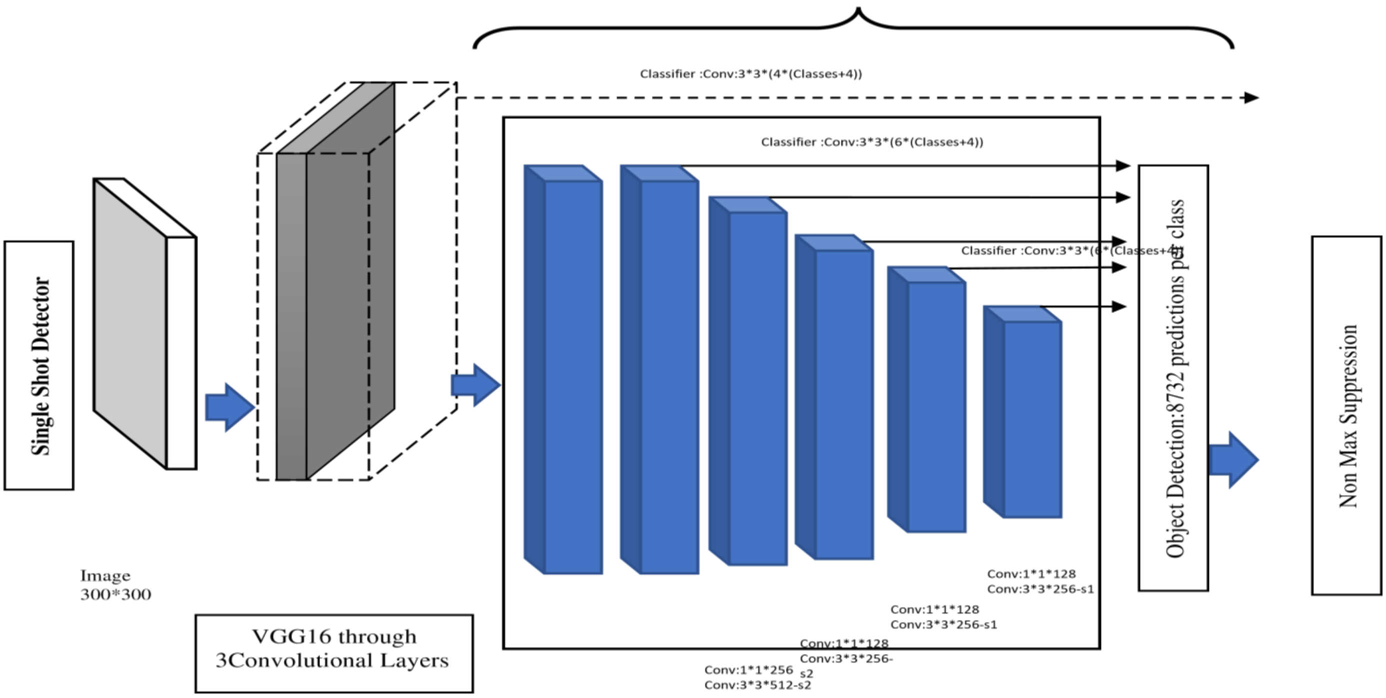

Fig -1: **Architecture of Single Shot Detector**

5.SYSTEMS PROPOSED ARCHITECTURE

The project commences by capturing an input image through a night vision camera, essential for subsequent stages. Preprocessing of this image involves annotating it into .txt and .xml formats using bounding boxes, which aligns with the necessary format for training YOLOv8 and SSD, two prevalent deep learning algorithms utilized in computer vision tasks. Feature extraction is then conducted employing these algorithms to derive valuable information from the pre-processed image.

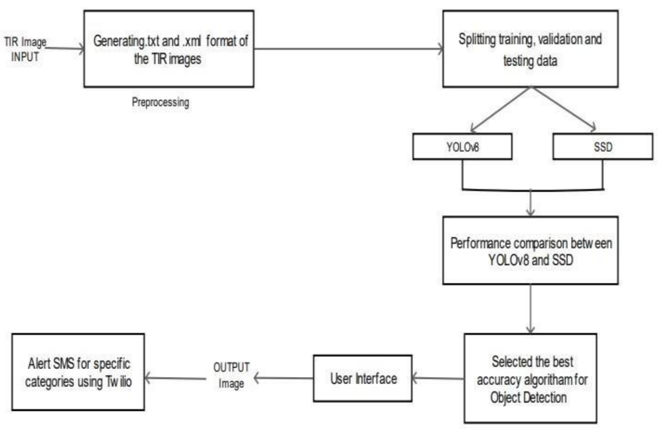

Fig -2: System Design

Following this, a meticulous comparison of YOLOv8 and SSD performance is executed using varied evaluation metrics like mean average precision (mAP), precision, recall, and F1 score. The findings indicate YOLOv8's superiority over SSD in object detection tasks, affirming its efficacy for the project's objectives.

Once the optimal detection algorithm, YOLOv8, is identified, it is deployed to discern objects within the input image. This process involves categorizing the identified objects and determining their respective probabilities. In the event a specific object category is detected, an alert SMS is generated using Twilio, promptly notifying users in real-time. Notably, this SMS includes a detailed count of the detected objects based on their categories, enhancing users' situational awareness.

To facilitate user interaction and engagement, the output image and detection results are displayed via a user-friendly web application user interface. This interface serves as an accessible platform through which users can observe, interact, and potentially further analyze the detected objects, ensuring a seamless and comprehensive user experience.

6. CONCLUSION

It will detect the abnormal activity in banks, roads and many more crowded areas, this can be implemented through YOLOv8 to automate and simplify manual work. It will detect variety of objects. The project will successfully implement both algorithms and compare their performance using evaluation metrics such as MAP score, precision, recall, and F2 score. The project will also integrate the Twilio API to send SMS alerts for a specific category of detected objects.

7. REFERENCES

[1] Chenchen Jiang, Huazhong Ren, Xin Ye, Jinshun Zhu, Hui Zeng, Yang Nan , Min Sun, Xiang Ren, Hongtao Huo, "Object detection from UAV thermal infrared images and videos using YOLOv8 models",2022.

[2] 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

[3] Mate Kristo, Marina Ivasic-Kos, Miran Pobar, "Thermal Object Detection in Difficult Weather Conditions Using YOLOv8", in Department of Informatics University of Rijeka, Rijeka, Croatia, IEEE 2020.

[4] Heena Patel and Kishor P. Upla, "Night Vision Surveillance: Object Detection using Thermal and Visible Images", in 2020 International Conference for Emerging Technology (INCET) Belgaum, India. Jun 5-7, 2020.

[5] Yuxuan Xiao, Aiwen Jiang, (Member, Ieee), Jihua Ye, And Ming-Wenwang, "Making of Night Vision: Object Detection Under Low-Illumination", School of Computer and Information Engineering, Jiangxi Normal University, Nanchang 330022, China, July 7, 2020.

[6] Aashish Bhandari, Aayush Kafle, Pranjal Dhakal, Prateek Raj Joshi and Dinesh Baniya Kshatri, "Image Enhancement and Object Recognition for Night Vision Surveillance", Department of Electronics and Computer Engineering IoE, Tribhuvan University Lalitpur, Nepal, 2020.

[7] Determining detection, recognition, and identification ranges of thermal cameras on the basis of laboratory measurements and TTP model." Infrared Imaging Systems: Design, Analysis, Modelling, and Testing XXIII. Vol. 8355.

[8] International Society for Optics and Photonics, 2012.

[9] Z. Wu, N. Fuller, D. Theriault, and M. Betke, "A thermal infrared video benchmark for visual analysis," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2014, pp. 201–208.

[10] B. Blecha, "Forward Looking Infrared (FLIR),"in Encyclopaedia of optical engineering, 2003, pp. 560–568.Thermal camera characteristics, URL:https://www.termogram.com/pdf/therma_cam_p_series/thermacam_p10/P10_datasheet.pdf

[11] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-CNN: Towards real-time object detection with region proposal networks," in Advances in neural information processing systems, 2015, pp. 91–99.

[12] Z. Cai and N. Vasconcelos, "Cascade R-CNN: Delving into High-Quality Object Detection," arXiv:1712.00726 [cs], Dec. 2017, Accessed: Jun. 08, 2020. [Online]. Available: http://arxiv.org/abs/1712.00726.

[13] Z. Tian, C. Shen, H. Chen, and T. He, "FCOS: Fully Convolutional One-Stage Object Detection," in 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South), Oct. 2019, pp. 9626–9635, DOI: 10.1109/ICCV.2019.00972.

[14] M. Buric, M. Pobar, M. Ivasic-Kos. An overview of action recognition in videos, 2017 40th International Convention on Information and Communication Technology, Electronics, and Microelectronics (MIPRO). IEEE, 2019.