Customized Smart Object Detection Using Yolo and R-Cnn in Machine Learning

[1] Mohammed Zabi Uddin, [2] Mohammed Abeel Ahmed Mohiuddin, [3] Mohd Abdullah Ansari, [4] Dr. Syed Asadullah Hussaini

[1] BE Student, Dept. of Computer Science Engineering, ISL Engineering College

[2] BE Student, Dept. of Computer Science Engineering, ISL Engineering College

[3] BE Student, Dept. of Computer Science Engineering, ISL Engineering College

[4] Associate Professor, Dept. of Computer Science Engineering, ISL Engineering College

ABSTRACT
In this paper work, using python and OPENCV module we are detecting objects from videos and webcam. This application consists of two modules such as 'Browse System Videos' and 'Start Webcam Video Tracking'. Object detection is an important task in computer vision and has numerous applications in fields such as surveillance, robotics, and autonomous driving. In this project, we aim to develop an object tracking system using Python and the OpenCV module. The system consists of two modules: "Browse System Videos" and "Start Webcam Video Tracking." The first module allows the user to select a video file from their system to track objects in, while the second module tracks objects in real-time using the user's webcam. Our system uses a combination of computer vision techniques, such as color thresholding and blob detection, to detect and track objects in the video or webcam feed. By developing this system, we hope to demonstrate the potential of Python and OpenCV for object tracking applications and inspire further development in the field.
Keywords: Object Detection, YOLO, R-CNN, Machine Learning, Video Tracking.

I. Introduction

Object Detection is an emerging technology in the field of Computer Vision and Image Processing that deals with detecting objects of a particular class in digital images. It has considered being one of the complicated and challenging tasks in computer vision [1]. It works with the methods of automatic extraction, analysis and understanding of useful information [9]. Object detection is a key concept in the field of AI, permitting computer frameworks to "see" their situations by recognizing objects in visual pictures or recordings.. Object detection is commonly known as object recognition, object identification, and image detection. It has wide range of applications such as object detection, traffic control, human-computer interaction, gesture recognition, video surveillance, security, defence, healthcare [5][7]. Object detection is one of the most important research directions for computer vision. Object detection algorithms are mainly divided into the traditional methods and the deep learning methods [3][4]. It is the system of detecting a goal item in a photo or a selected part of the video. Object detection will work only if the image that is needed, is visible on the given input. If the target

object is covered up by any obstruction, it'll not be able to distinguish it. In existing system, many problems in computer vision were saturating on their accuracy before a decade. However, with the rise of deep learning techniques, the accuracy of these problems drastically improved. One of the major problem was that of image classification, which is defined as predicting the class of the image. So, in this work, we focused on customized object detection by using OpenCV and R-CNN. It works by uploading videos from system and start tracking the objects the from video and then mark them with bounding boxes. It also connects the application to web cam and start streaming to track the objects from the web cam.

II. Literature Survey

Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi presented You Only Look Once: Unified, Real-Time Object Detection [3], which is a fast and simple approach to detecting real time images was introduced in this paper as You Only Look Once. The model was built to detect images accurately, fast and to differentiate between art and real images. In comparison with Object detection techniques that came before YOLO, like R-CNN, YOLO introduced a single unified architecture for regression go image into bounding boxesand finding class probabilities for each box. This meant that YOLO performed much faster and also provided more accuracy. It could also predict artwork correctly .

Chengji Liu, Yufan Tao, JiaweiLiang, Kai Li1, YihangChen presented Object Detection Based on YOLO Network [4], in which a generalized object detection network was developed by applying complex Degradation processes on training sets like noise, blurring, rotating and cropping of images.

The model was trained with the degraded training sets which resulted in better generalizing ability and higher robustness. The experiment showed that the model trained with the standard sets does not have good generalization ability for the degraded images and has poor robustness. Then the model was trained using degraded images which resulted in improved average precision. It was proved that the average precision for degraded images was better in general degenerative model compared to the standard model.

Rumin Zhang, Yifeng Yang [6] presented An Algorithm for Obstacle Detection based on YOLO and Light Filed Camera. An obstacle detection algorithm in the indoor environment is proposed which combines the YOLO object detection algorithm and the light field camera and will classify objects into categories and mark them in the image. The images of the common obstacles were labelled and used for training YOLO. The object filter is applied to remove the unconcern obstacle. Different types of scene, including pedestrian, chairs, books and so on, are demonstrated to prove the effectiveness of this obstacle detection algorithm.

III. Proposed System

In this work, using python and OPENCV module, we are detecting objects from videos and webcam. The goal of object tracking is to locate and track an object's movement over time in a video.

This is typically done by selecting an object in the first frame of the video and tracking it as it moves throughout the subsequent frames. Object tracking can be challenging due to changes in lighting, occlusions, and object appearance changes.

This application consists of two modules such as –

'Browse System Videos' and 'Start Webcam Video Tracking'.

• One of the major problem was that of image classification, which is defined as predicting the class of the image.

• There is no object detection in existing system by using OpenCV.

This system can be applied in the areas of surveillance system, face recognition, fault detection, character recognition, etc.

It is extremely fast compared to other real time detectors which came before it as it uses a Unified Model where the detection is seen as a single regression problem and there is no complex pipeline, just a neural network run on the image.

IV. Implementation

OpenCV will use following algorithms to track object in videos Dense Optical flow: These algorithms help estimate the motion vector of every pixel in a video frame.

• Sparse optical flow: These algorithms, like the Kanade-Lucas-Tomashi (KLT) feature tracker, track the location of a few feature points in an image [11].

• Kalman Filtering: A very popular signal processing algorithm used to predict the location of a moving object based on prior motion information. One of the early applications of this algorithm was missile guidance! Also as mentioned here, "the on-board computer that guided the descent of the Apollo 11 lunar module to the moon had a Kalman filter" [11].

• Meanshift and Camshift: These are algorithms for locating the maxima of a density function. They are also used for tracking.

• Single object trackers: In this, the first frame is marked using a rectangle to indicate the location of the object we want to track. The object is then tracked in subsequent frames using the tracking algorithm. [2].

• Multiple object track finding algorithms: It makes sense to detect multiple objects in each frame and then run a track finding algorithm that identifies which rectangle in one frame corresponds to a rectangle in the next frame [5][10].

YOLO – You Only Look Once

It is a single-shot detector model that uses a fully convolutional neural network (CNN) to process an image. The YOLO algorithm takes an image as input and then uses a simple deep convolutional neural network to detect objects in the image.
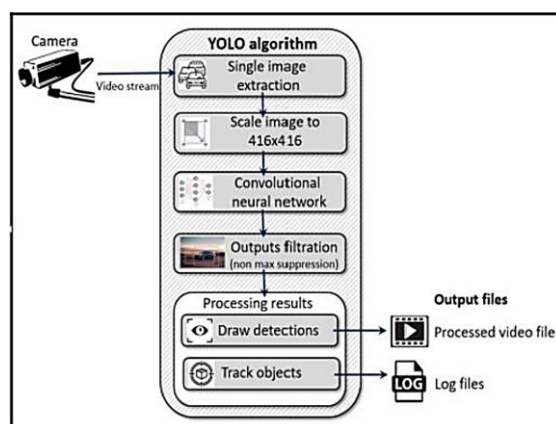
Fig. 1: Object detection with YOLO

R-CNN – Region-Based Convolutional Neural Network. The CNN algorithm have been extensively used to classify images. But to detect in an image and to draw bounding boxes around them is a tough problem to solve. So, to solve this problem R-CNN algorithm is considered, which is a technique for addressing object detection tasks, designed for model performance.
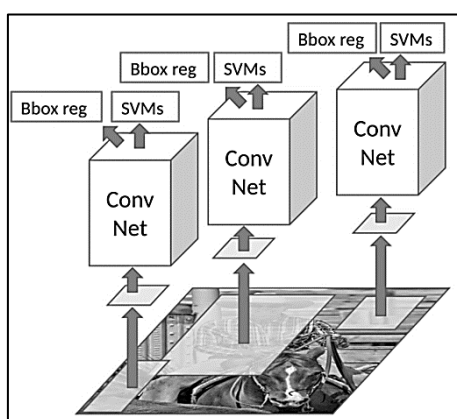

Fig. 2: Object detection with R-CNN

Modules description:

In this work, using python and OPENCV module we are detecting objects from videos and webcam.

This application consists of two modules such as 'Browse System Videos' and 'Start Webcam Video Tracking'.

Module 1 –

Browse System Videos: Using this module application allow user to upload any video from his system and application will connect to that video and start playing it, while playing if application detect any object then it will mark that object with bounding boxes, while playing

video if user wants to stop tracking then he need to press 'q' key from keyboard to stop video playing.

Module 2 –

Start Webcam Video Tracking: Using this module application connect itself with inbuilt system webcam and start video streaming, while streaming if application detect any object then it will surround that object with bounding boxes, while playing press 'q' to stop web cam streaming.
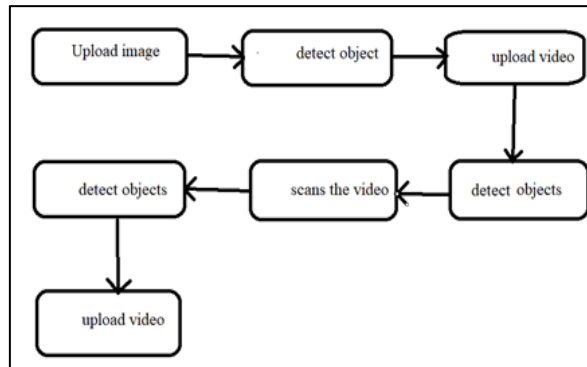
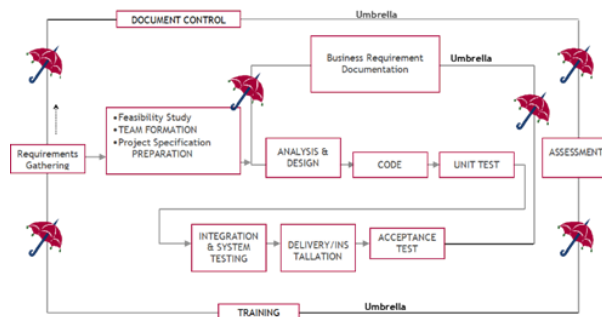**Architecture diagrams:**

Fig. 3: Architecture diagram

Fig 4. Umbrella Model

Screenshots:

The two modules of the application are shown below.

Fig 5. Application Modules

When we click on the first module, 'Browse System Videos' button to upload videos from system, we can see the application starts tracking the objects from the video and mark them with bounding boxes, as shown in fig. 6.
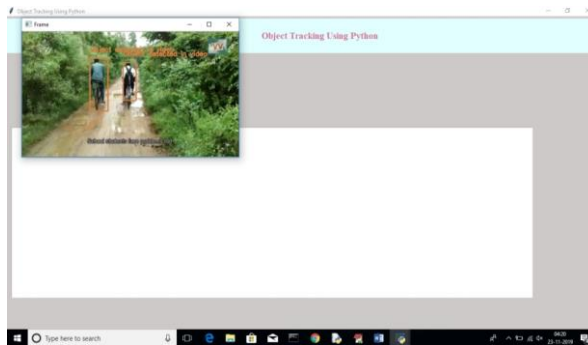

Fig 6. Object tracking from video

In fig. 7 and fig. 8, when we click on another module button, called 'Start Webcam Video Tracking' to connect the application to the web cam and start streaming. It will track the computer mouse and the tiffin box, as the detected objects from the web cam.
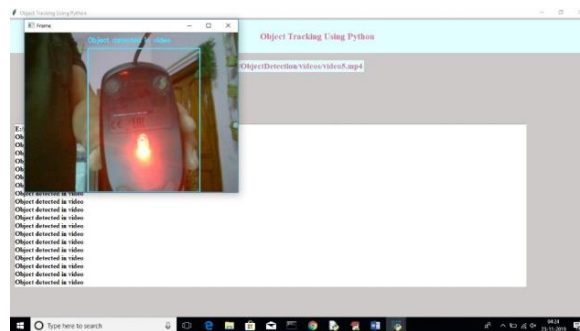
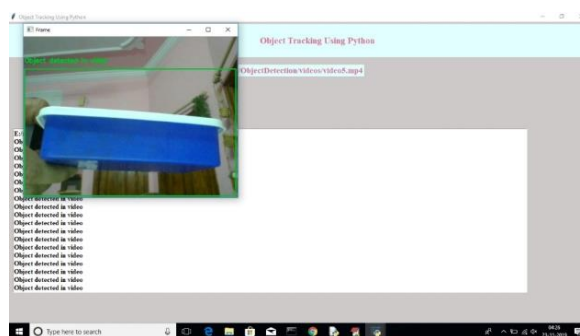

Fig 7. Object tracking from webcam (1)

Fig 8. Object tracking from webcam (2)

V. Conclusion & Future Work

In this work, an accurate and efficient object detection system has been developed which achieves comparable metrics with the existing state-of-the-art system. This work uses recent techniques in the field of computer vision and deep learning.

Features either the local or global used for recognition can be increased, to increase the efficiency.

Features either the local or global used for recognition can be increased, to increase the efficiency of the object recognition system. Geometric properties of the image can be included in the feature vector for recognition. Fully occluded object cannot be tracked and considered as a new object in the next frame.

VI. REFERENCES

1. Sumahassan, Udaya Kumar Addanki, Navya Irlapati, "Object Detection using Deep Learning Algorithm CNN", International Journal for Research in Applied Science and Engineering Technology, Vol 8, July 2020.

2. G Chandan, Ayush Jain, Harsh Jain, "Real Time Object Detection and Tracking Using Deep Learning and OpenCV",International Conference on Inventive Research in Computing Applications, July 2018.

3. Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, "You Only Look Once: Unified, Real-Time Object Detection", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), **ISSN:** 1063-6919, December 2016.

4. Chengji Liu, Yufan Tao, JiaweiLiang, Kai Li1, YihangChen, "Object Detection Based on YOLO Network", IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC), June 2019.

5. Apoorva Raghunandan, Mohana, Pakala Raghav and H. V. Ravish Aradhya, "Object Detection Algorithms for video surveillance applications" International conference on communication and signal processing (ICCSP), India, 2018, pp. 0570-0575.

6. R. Zhang, Y. Yang, W. Wang, L. Zeng, J. Chen and S. McGrath, "An Algorithm for Obstacle Detection based on YOLO and Light Filed Camera," 2018 12th International Conference on Sensing Technology (ICST)*,* Limerick, Ireland, 2018, pp. 223-226, doi: 10.1109/ICSensT.2018.8603600.

7. Akshay Mangawati, Mohana, Mohammed Leesan, H. V. Ravish Aradhya, "Object Tracking Algorithms for video surveillance applications" International conference on communication and signal processing (ICCSP), India, 2018, pp. 0676-0680.

8. D. Held, S. Thrun, and S. Savarese "Learning to Track at 100 FPS with Deep Regression Networks", ECCV 2016.

9. Tanvir Ahamed Nayeem, SM Motaharuzzaman, Anika Tabassum Hoque, "Computer Vision Based Object Detection and Recognition System for Image Searching", 12th International Conference on Electrical and Computer Engineering (ICECE), Electronic ISSN: 2771-7917, April 2023.

10. Ayoub Benali Amjoud, Mustapha Amrouch, "Object Detection Using Deep Learning, CNNs and Vision Transformers: A Review", IEEE Access, vol.11, pp.35479-35516, 2023.

11. Yaran Chen, Dongbin Zhao, Haoran Li, "Deep Kalman Filter with Optical Flow for Multiple Object Tracking", IEEE International Conference on Systems, Man and Cybernetics (SMC), Electronic ISSN: 2577-1655, November 2019.

12. M.A.Bari & Shahanawaj Ahamad, "Code Cloning: The Analysis, Detection and Removal", in International Journal of Computer Applications(IJCA),ISSN:0975-887,Vol:20, U.S.A.,April 2011.

13. Hafsa Fatima, Shayesta Nazneen, Maryam Banu, Dr. Mohammed Abdul Bari, "Tensor flow-Based Automatic Personality Recognition Used in Asynchronous Video Interviews", Journal of Engineering Science (JES), ISSN NO:0377-9254, Vol 13, Issue 05, May 2022.

14. Ijteba Sultana, Mohd Abdul Bari and Sanjay, "Impact of Intermediate Bottleneck Nodes on the QoS Provision in Wireless Infrastructure less Networks", Journal of Physics: Conference Series, Conf. Ser. 1998 012029 , CONSILIO Aug 2021

15. Syed Shehriyar Ali, Mohammed Sarfaraz Shaikh, Syed Safi Uddin, Dr. Mohammed Abdul Bari, "Saas Product Comparison and Reviews Using Nlp", Journal of Engineering Science (JES), ISSN NO:0377-9254, Vol 13, Issue 05, MAY/2022.

16. Mohammed Shoeb, Mohammed Akram Ali, Mohammed Shadeel, Dr. Mohammed Abdul Bari, "Self-Driving Car: Using Opencv2 and Machine Learning", The International journal of analytical and experimental modal analysis (IJAEMA), ISSN NO: 0886-9367, Volume XIV, Issue V, May/2022.

17. Talukdar, V., Dhabliya, D., Kumar, B., Talukdar, S. B., Ahamad, S., & Gupta, A. (2022). Suspicious activity detection and classification in IoT environment using machine learning approach. Paper presented at the PDGC 2022 - 2022 7th International Conference on Parallel, Distributed and Grid Computing, 531-535. doi:10.1109/PDGC56933.2022.10053312 Retrieved from www.scopus.com

18. Timande, S., & Dhabliya, D. (2019). Designing multi-cloud server for scalable and secure sharing over web. International Journal of Psychosocial Rehabilitation, 23(5), 835-841. doi:10.37200/IJPR/V23I5/PR190698

19. Umbarkar, A. M., Sherie, N. P., Agrawal, S. A., Kharche, P. P., & Dhabliya, D. (2021). Robust design of optimal location analysis for piezoelectric sensor in a cantilever beam. Materials Today: Proceedings, doi:10.1016/j.matpr.2020.12.1058

20. Vadivu, N. S., Gupta, G., Naveed, Q. N., Rasheed, T., Singh, S. K., & Dhabliya, D. (2022). Correlation-based mutual information model for analysis of lung cancer CT image. BioMed Research International, 2022, 6451770. doi:10.1155/2022/6451770

21. Veeraiah, D., Mohanty, R., Kundu, S., Dhabliya, D., Tiwari, M., Jamal, S. S., & Halifa, A. (2022). Detection of malicious cloud bandwidth consumption in cloud computing using machine learning techniques. Computational Intelligence and Neuroscience, 2022 doi:10.1155/2022/4003403

22. Mr. Pathan Ahmed Khan, Dr. M.A Bari,: "Impact Of Emergence With Robotics At Educational Institution And Emerging Challenges", International Journal of Multidisciplinary Engineering in Current Research(IJMEC), ISSN: 2456-4265, Volume 6, Issue 12, December 2021,Page 43-46.

23. Mohammed Abdul Bari, Shahanawaj Ahamad, Mohammed Rahmat Ali, "Smartphone Security and Protection Practices", International Journal of Engineering and Applied Computer Science (IJEACS) ; ISBN: 9798799755577 Volume: 03, Issue: 01, December 2021.

24. Shahanawaj Ahamad, Mohammed Abdul Bari, "Big Data Processing Model for Smart City Design: A Systematic Review" , VOL 2021: ISSUE 08 IS SN : 0011-9342 ;Design Engineering (Toronto) Elsevier SCI Oct.

25. Mohammed Abdul Bari, Shahanawaj Ahamad, Mohammed Rahmat Ali, "Smartphone Security and Protection Practices", International Journal of Engineering and Applied Computer Science (IJEACS) ; ISBN: 9798799755577 Volume: 03, Issue: 01, December 2021 (International Journal,U K) Pages 1-6.