International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

ENHANCED OBJECT DETECTION WITH DEEP LEARNING

YARAMTI GOPALA KRISHNA
PG SCHOLAR ,Dept of CSE
SWARNANDHRA COLLEGE OF ENGINEERING AND TECHNOLOGY
Dr P.SRINIVASULU
HOD And Professor, Dept of CSE
SWARNANDHRA COLLEGE OF ENGINEERING AND TECHNOLOGY

ABSTRACT

Object detection is a critical problem for advanced driving assistance systems (ADAS). Recently, convolutional neural networks (CNN) achieved large successes on object detection, with performance improvement over traditional approaches, which use hand-engineered features. However, due to the challenging driving environment (e.g., large object scale variation, object occlusion, and bad light conditions), popular CNN detectors do not achieve very good object detection accuracy over the KITTI autonomous driving benchmark dataset. In this paper, we propose three enhancements for CNN-based visual object detection for ADAS. To address the large object scale variation challenge, deconvolution and fusion of CNN feature maps are proposed to add context and deeper features for better object detection at low feature map scales. In addition, soft nonmaximal suppression (NMS) is applied across object proposals at different feature scales to address the object occlusion challenge. As the cars and pedestrians have distinct aspect ratio features, we measure their aspect ratio statistics and exploit them to set anchor boxes properly for better object matching and localization. The proposed CNN enhancements are evaluated with various image input sizes by experiments over KITTI dataset. The experimental results demonstrate the effectiveness of the proposed enhancements with good detection performance over KITTI test set

.

INTRODUCTION

Since AlexNet has stormed the research world in 2012 ImageNet on a large scale visual recognition challenge, for detection in-depth learning, far exceeding the most traditional methods of artificial vision used in literature. In artificial vision, the neural convolution networks are distinguished in the classification of images. Object detection [9] and location in digital images has become one of the most important applications for industries to ease user, save time and to achieve parallelism. This is not a new technique but improvement in object detection is still required in order to achieve the targeted objective more efficiently and accurately. The main aim of studying and researching computer vision is to simulate the behavior and manner of human eyes directly by using a computer and later on develop a system that reduces human efforts shows the basic block diagram of detection and tracking. In this paper, an SSD and MobileNets based algorithms are implemented for detection and tracking in python environment. Object detection involves detecting region of interest of object from given class of image. Different methods are –Frame differencing, Optical flow, Background subtraction. This is a method of detecting and locating an object which is in motion with the help of a camera. Detection and tracking algorithms are described by extracting the features of image and video for security applications [3] [7] [8]. Features are extracted using CNN and deep learning [9]. Classifiers are used for image classification and counting [6]. YOLO based algorithm with GMM model by using the concepts of deep learning will give good accuracy for feature extraction and classification [10]. Section II describes SSD and MobileNets algorithm, section III explains method of implementation, and section IV describes simulation results and analysis.

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.

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".

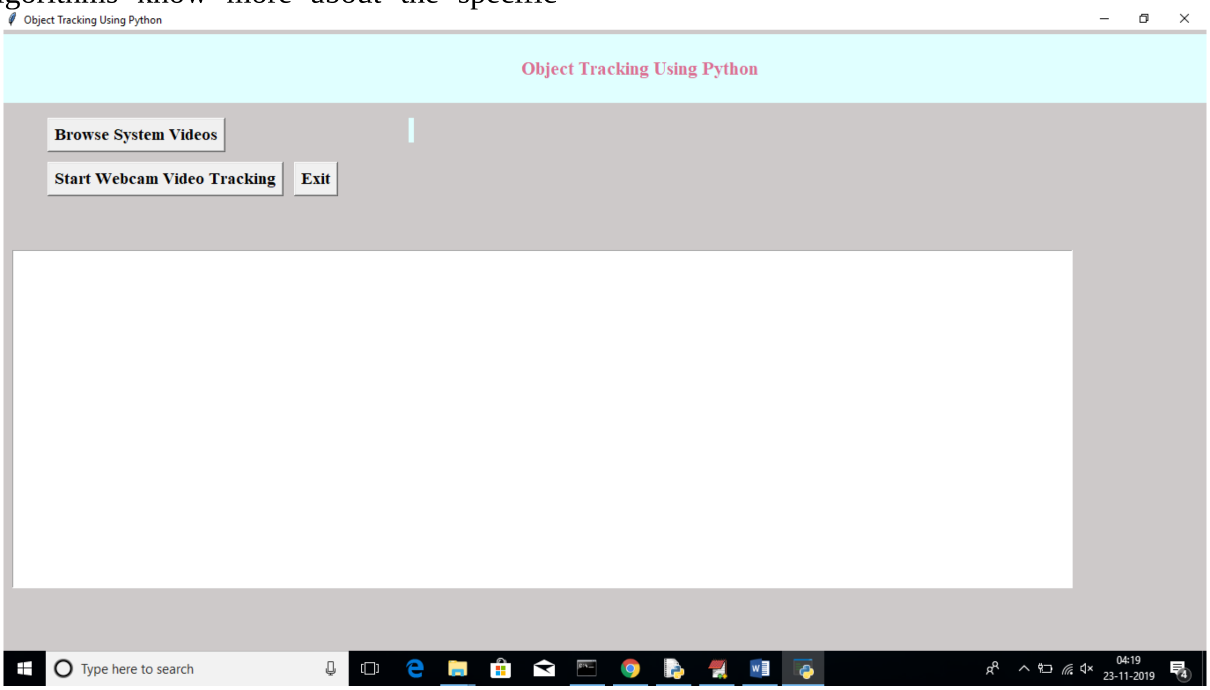
Meanshift and Camshift: These are algorithms for locating the maxima of a density function. They are also used for tracking.

Single object trackers: In this class of trackers, 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. In most real life applications, these trackers are used in conjunction with an object detector.

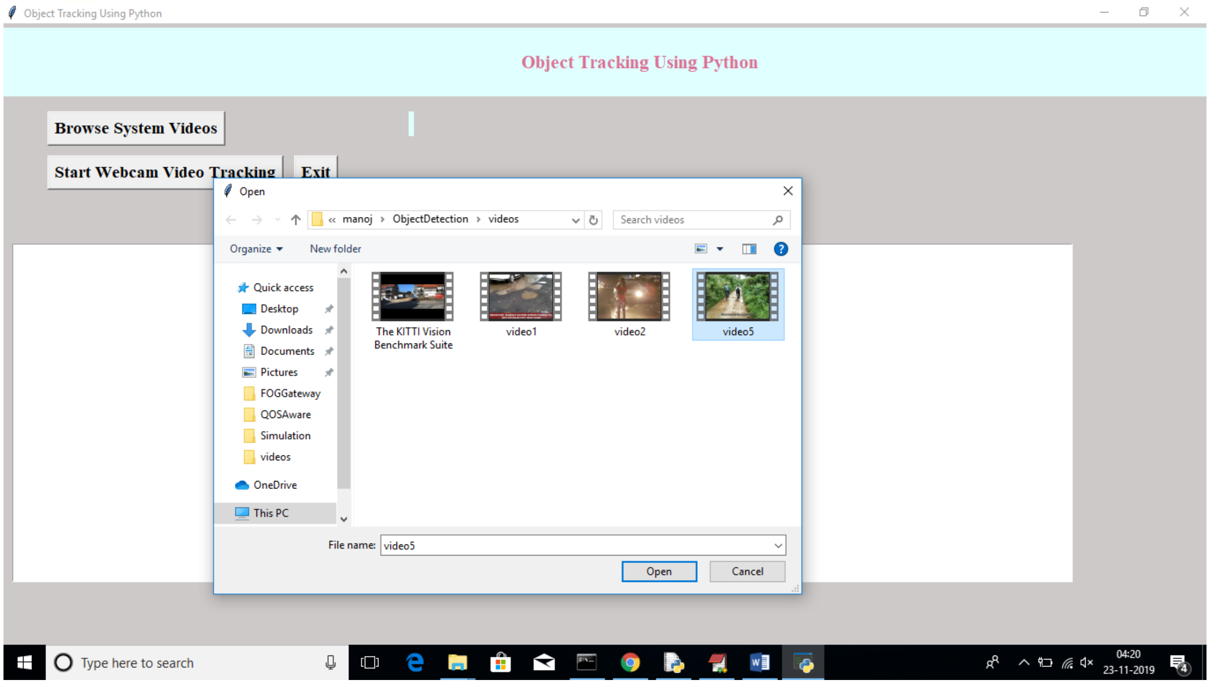
Multiple object track finding algorithms: In cases when we have a fast object detector, it makes sense to detect multiple objects in each frame and then run a track finding algorithm that identifies which rectangle in one frame corresponds to a rectangle in the next frame.

Tracking vs Detection

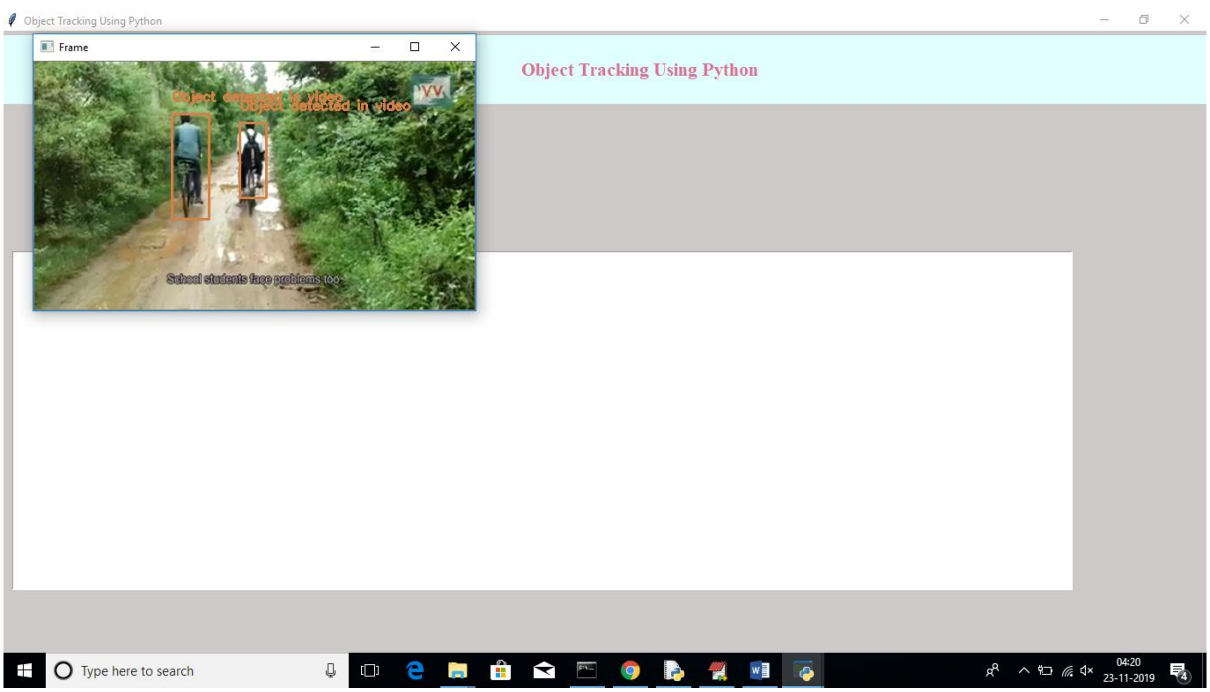
If you have ever played with OpenCV face detection, you know that it works in real time and you can easily detect the face in every frame. So, why do you need tracking in the first place? Let's explore the different reasons you may want to track objects in a video and not just do repeated detections. Tracking is faster than Detection: Usually tracking algorithms are faster than detection algorithms. The reason is simple. When you are tracking an object that was detected in the previous frame, you know a lot about the appearance of the object. You also know the location in the previous frame and the direction and speed of its motion. So in the next frame, you can use all this information to predict the location of the object in the next frame and do a small search around the expected location of the object to accurately locate the object. A good tracking algorithm will use all information it has about the object up to that point while a detection algorithm always starts from scratch. Therefore, while designing an efficient system usually an object detection is run on every nth frame while the tracking algorithm is employed in the n-1 frames in between. Why don't we simply detect the object in the first frame and track subsequently? It is true that tracking benefits from the extra information it has, but you can also lose track of an object when they go behind an obstacle for an extended period of time or if they move so fast that the tracking algorithm cannot catch up. It is also common for tracking algorithms to accumulate errors and the bounding box tracking the object slowly drifts away from the object it is tracking. To fix these problems with tracking algorithms, a detection algorithm is run every so often. Detection algorithms are trained on a large number of examples of the object. They, therefore, have more knowledge about the general class of the object. On the other hand, tracking algorithms know more about the specific instance of the class they are tracking. Tracking can help when detection fails: If you are running a face detector on a video and the person's face get's occluded by an object, the face detector will most likely fail. A good tracking algorithm, on the other hand, will handle some level of occlusion. Double click on 'run.bat' file to get below screen



Now click on 'Browse System Videos' button to upload videos from system

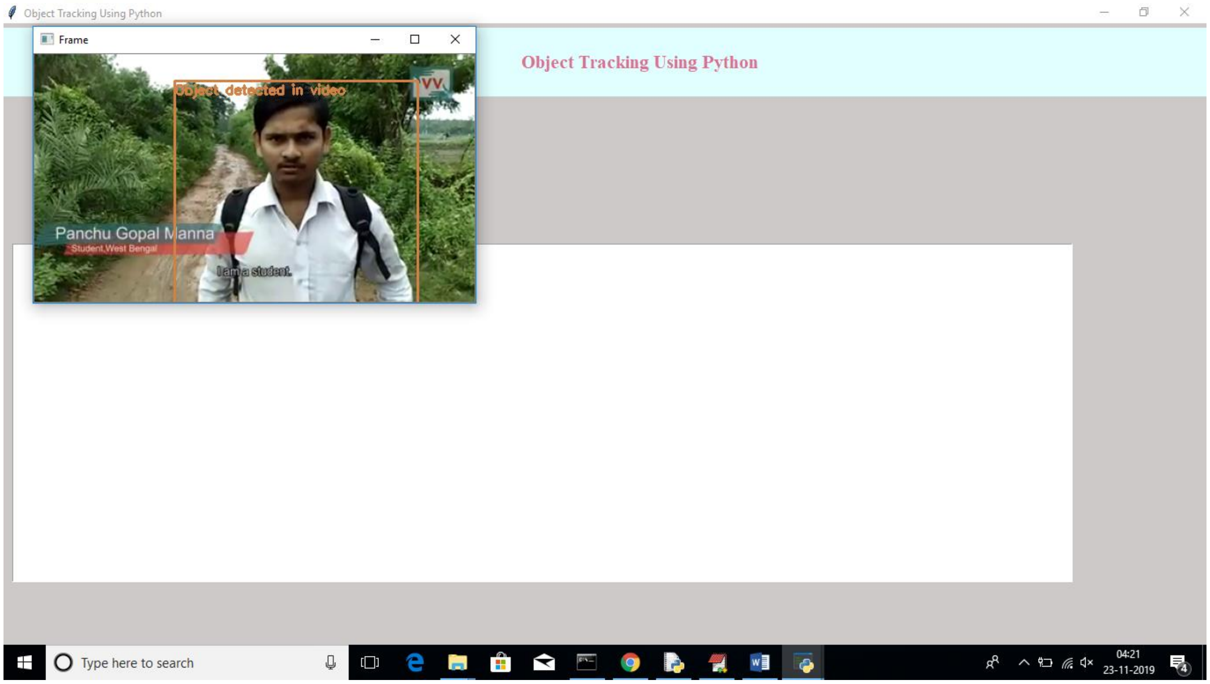


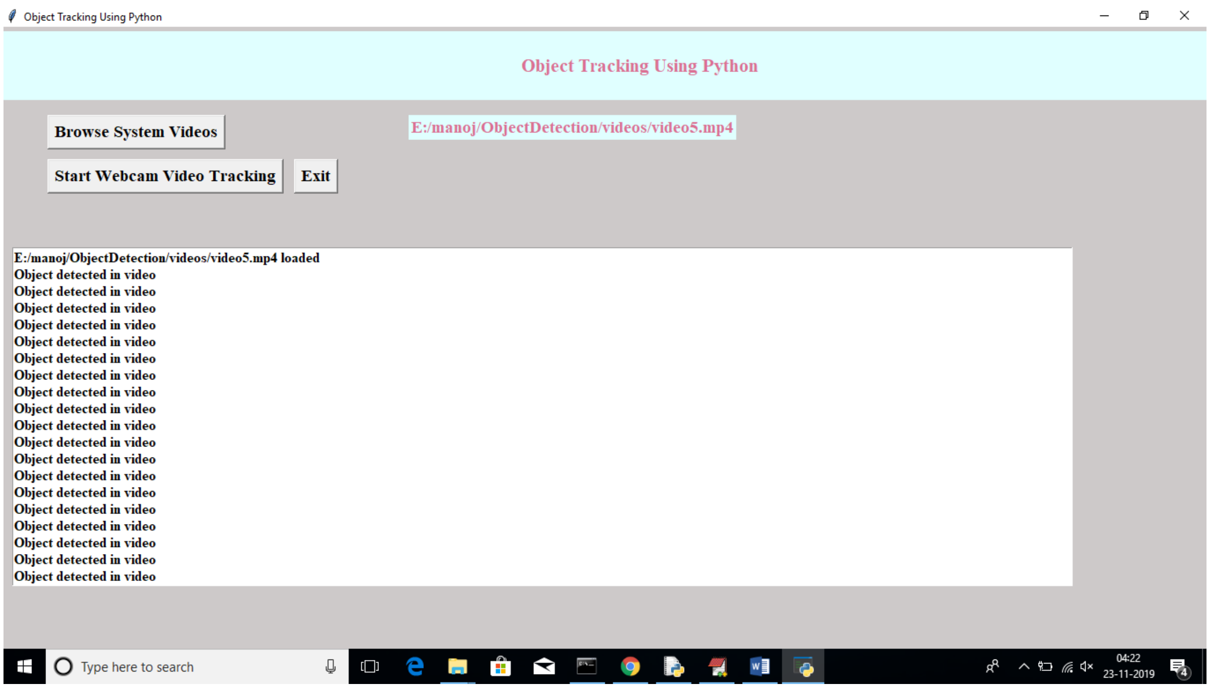
In above screen I am uploading one video, after upload will get below screen



In above video we can see application start tracking objects from video and mark them with bounding boxes. Similarly we can upload any video and track objects from video

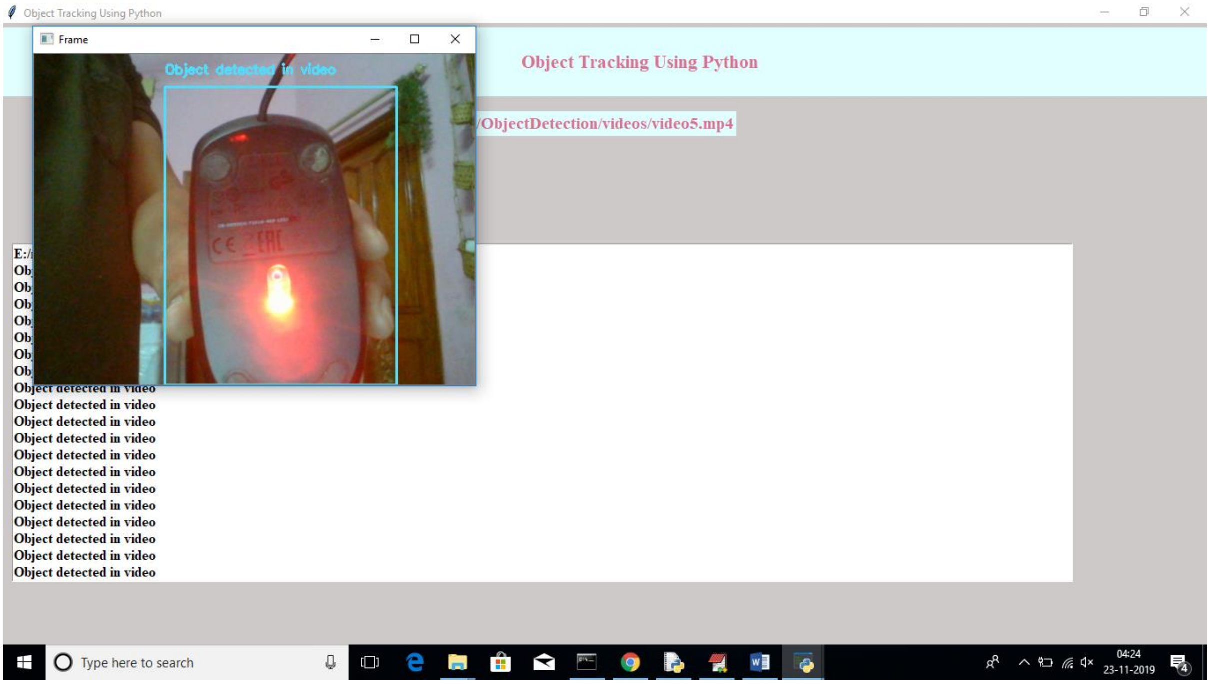
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362



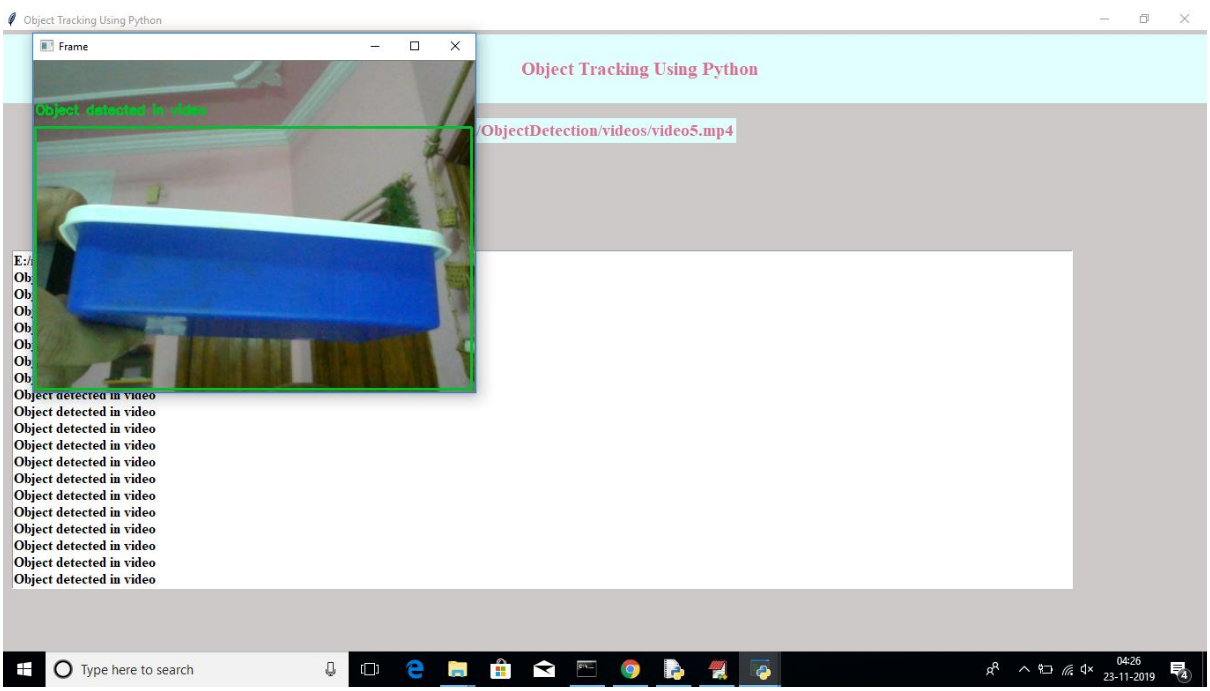


In above screen now click on another button called 'Start Webcam Video Tracking' to connect application to web cam and start streaming. After connecting to webcam will get below screen

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362



In above screen we can see objects is getting tracked from webcam also. In above screen it track computer mouse from web cam video



International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

CONCLUSION

Real time accurate object detection is one of the most critical problems for advanced driving assistance systems (ADAS) and autonomous driving. Recently convolutional neural networks (CNN) achieved huge successes on visual object detection over traditional object detectors, which use handengineered features. However, due to the challenging driving environment (e.g., large object scale variation, object occlusion and bad light conditions), popular CNN detectors including Faster-RCNN and SSD do not produce good detection performance over the KITTI driving benchmark dataset. In this paper we proposed three enhancements on a multiple scale CNN network model for ADAS object detection. Firstly, CNN feature maps deconvolution and fusion was proposed to add context and deeper features for better object detection at lower scale of feature maps, to address the large object scale variation challenge. Then, soft non-maximal suppression (NMS) was applied across object proposals at different image scales to address the object occlusion challenge. As the cars and pedestrians have distinct aspect ratio features, we measured their aspect ratio statistics and exploited them to set anchor boxes properly for better object matching and localization. The proposed CNN enhancements with various input image sizes were individually and jointly evaluated by extensive experiments over KITTI dataset. The effectiveness of the proposed enhancements was verified by experiment results with improved or comparable detection performance over KITTI test set. The average precision (AP) for pedestrian detection category "Easy" and the computation speed rank the first among the published works, the second for pedestrian category "Moderate" and "Hard", the third for car category "Moderate". And the network inference time for cars per 384×1280 image is only 0.08 second, much faster than the other top ranked published methods in KITTI leader board. In our future works we will investigate more CNN models and enhancements to improve object detection for safe and intelligent transport.

REFERENCES

[1] Wei Liu and Alexander C. Berg, "SSD: Single Shot MultiBox Detector",Google Inc., Dec 2016.
[2] Andrew G. Howard, and Hartwig Adam, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications", Google Inc., 17 Apr 2017.
[3] Justin Lai, Sydney Maples, "Ammunition Detection: Developing a Real- Time Gun Detection Classifier", Stanford University, Feb 2017
[4] Shreyamsh Kamate, "UAV: Application of Object Detection and Tracking Techniques for Unmanned Aerial Vehicles", Texas A&M University, 2015.
[5] Adrian Rosebrock, "Object detection with deep learning and OpenCV", pyimagesearch.
[6] Mohana and H. V. R. Aradhya, "Elegant and efficient algorithms for real time object detection, counting and classification for video surveillance applications from single fixed camera," 2016 International Conference on Circuits, Controls, Communications and Computing (I4C), Bangalore, 2016, pp. 1-7.
[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] 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.
[9] Manjunath Jogin, Mohana, "Feature extraction using Convolution Neural Networks (CNN) and Deep Learning" 2018 IEEE International Conference On Recent Trends In Electronics Information Communication Technology,(RTEICT) 2018, India.
[10] Arka Prava Jana, Abhiraj Biswas, Mohana, "YOLO based Detection and Classification of Objects in video records" 2018 IEEE International Conference On Recent Trends In Electronics Information Communication Technology,