Object Detetcion using SSD-MobileNet

B.MathuraBai[1], Vishnu Priya.Maddali[2], Chaithya.Devineni[3], Iswarya.Bhukya[4], Shirisha.Bandari[5]

[1]Associate Professor, Dept.of Information Technology, Vnr Vjiet, Telangana, India
[2-5]Under Gradute, Dept.of Information Technology, Vnr Vjiet, Telangana, India

---***---

Abstract - *Object detection is employed in practically every real-world application, including autonomous traversal, visual systems, and face identification. Real-time object detection is a vast, colourful, and intricate domain of computer vision. Object detection is critical in the realm of computer vision. There have been several improved object detection algorithms published in literature; nonetheless, the majority of them are aimed to enhance detection accuracy. As a result, the necessity to reduce computational complexity is frequently overlooked. These upgraded object detectors require a high-end GPU to attain real-time performance. In this article, we offer a lightweight object detection model built on Mobilenet-v2. The real-time object detector developed here can be used in embedded systems with limited processing resources. This is a critical design component of current autonomous driving assistance systems (ADAS). The suggested lightweight object detector offers a wide range of applications.*

Key Words: **Computer Vision, Object Detection, MobileNet, Single Shot Multi-Box Detector, OpenCv.**

1. INTRODUCTION

Object detection is one of the critical regions of studies in cpmputer vision and prescient today. This is an image classification technique that aims to detect one or more types of elements in an image and use a bounding box to determine their presence.The formatter must build these components while taking into account the following requirements. Object detection is the process of identifying each object in a photograph or image using computer/software. Object recognition is one of the most pressing issues in wireless network computer vision. It is often used in wireless networks and serves as the muse for complex imaginative and prescient obligations which includes goal monitoring and scene interpretation. The aim of object detection is to identify whether there are any objects in the image that belong to the defined category. If it exists, the next job is to determine its category and location. Traditional object detection algorithms are mostly focused on detecting a few types of objects, such as pedestrian detection and infrared target detection. Object identification algorithms have made significant progress as a result of recent advances in deep learning technology, particularly the introduction of deep convolution neural network (CNN) technology. Three primary approaches extensively used in this sector are You Only Look Once (YOLO), single shot multi-box detector (SSD), and faster region CNN (F-RCNN).This study adds to the current literature by enhancing the accuracy of the SSD method for recognising tiny items. The SSD algorithm detects large things effectively but is less reliable when recognising tiny objects. As a result, we alter the SSD method to obtain acceptable detection accuracy for tiny items. The photos or sceneries were captured using web cams, and we conducted tests using common objects in context (COCO) datasets. We collect object detection (OD) datasets from our facility for use in our image processing lab. We create a network using several libraries including tensorflow-GPU 1.5. Tensorflow directory, SSD Mobilenet FPN Feature Extractor, object detection API of TensorFlow, and jupyter notebook are used for the experimental setup. This overall arrangement allows us to provide superior real-time object detection. However, MobileNet with the powerful SSD framework has been a warm research factor in latest times, because of the purposeful barriers of running robust neural nets on low-stop gadgets like mobileular phones/laptops to moreover amplify the horde of achievable effects with admire to real-time applications.

2. LITERATURE SURVEY

Celiet. Al presents an object detector based on small sample learning. The proposed model makes use of object semantic relevance to improve the accuracy of weak feature objects in complex scenarios. Tanget. Al concentrates on the framework design and model working principles, as well as the model's real-time performance and detection accuracy. Christian szegedy and colleagues provide a simple but effective formulation of object detection as a regression problem on object bounding box masks. It describes a multiscale inference process that, when used by a few network applications, produces high-resolution object detections at a cheap cost. Xiaogang wanget offers an overview of deep learning with an emphasis on applications in object identification, detection, and segmentation, which are fundamental difficulties for computer vision with various applications to photos and videos. A novel object detection technique is introduced, and objects are further differentiated by mean shift (MS) segmentation. There is vision with the assistance of depth information produced from stereo fixed number of sliding window templates. It is also possible to use supervised learning to solve the trouble through manner of approach of imposing it with Decision trees or, SVM in deep learning, as stated in Malay Shahet. ZhongQiuZhaoet provides a complete overview of deep learning-based object identification frameworks that address various sub-issues such as confusion and low resolution due to varying degrees of RCNN changes. Sandeep Kumaret works with the easynet model, which allows for detecting predictions with a single network. At testing time, the easynet version examines the complete image, so predictions are knowledgeable through global context.

3. EXISTING SYSTEM

Prior work was offered to accelerate the spatial pyramid pooling networks approach. This helped speed up feature extraction, but it was effectively a forward pass caching approach. Fast RCNN is a spatially-localized and oriented method to classify and localize objects in images. It works at an unprecedented speed, with high accuracy, and is powered by TensorFlow. The model is provided as a single model rather than a pipeline for immediate training and output of regions and classifications. The architecture takes an image as input, processing it to generate a collection of range recommendations. This collection will be processed by a deep convolutional neural network with more than 1 billion parameters and 100 million neurons, in order to give you the most up-to-date styles. This procedure will be performed on top of a pre-trained CNN model. The RoI Pooling layer, close to the belief of the deep CNN, retrieves traits unique to a particular input candidate region. The CNN model is first trained with a pre-trained network, which we provided. The CNN output is then processed via means of a fully linked layer, and then the version splits into outputs, one for class label prediction through a softmax layer and some other with a linear output for the bounding box. After each region of interest has been extracted and detected, the next step is to perform a comparison between each of these regions. Fast RCNN is significantly faster than RCNN in training and test sessions. When looking at the performance of Fast R-CNN during testing, incorporating region suggestions greatly slows down the algorithm. Previously, all object detection techniques hired areas to discover the object within the picture. The network does now no longer study the complete picture. Instead, regions of the photo with a excessive probability of containing the object are used. YOLO, or You Only Look Once, is an object identifying method that differs notably from the region-primarily based totally algorithms mentioned above. In YOLO, a single neural network predicts the bounding containers in addition to their class probabilities.

4. PROPOSED SYSTEM

The proposed system uses the Mobilenet SSD architecture to quickly and efficiently identify objects in real time. A Python script is written using OpenCV 3.4 that uses a deep neural network to discover objects. The system works as follows: The input is real-time video from a camera or webcam with a simplified MobileNet architecture that builds a lightweight deep neural network with depth-separable convolution. The input video is split into frames before being sent to the MobileNet layer. Each feature value is calculated as the difference between the amount of pixel intensity in the bright areas and the amount of pixel intensity in the dark areas. These components are computed using all of the image's available sizes and areas. Images can contain both irrelevant and related elements that can be used to identify items. The task of the MobileNet layer is to convert the pixels of the input image into highlights that characterize the image content. The bounding boxes and related class(label) of objects are then determined using the MobileNet-SSD model. The only remaining step is to display or view the output.

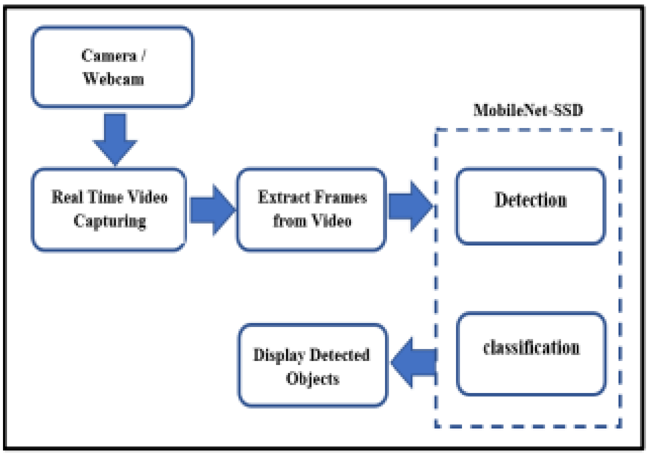

Fig -1: Proposed Methodology Architecture Design

5. RESULTS

This object detection algorithm achieves good results with any fps low quality camera and can detect objects in real time with decent accuracy. In our experiment we gave different images as input and the model has identified then with a good accuracy. And then we used the webcam to detect objects in the real-time which also produced the desired results.

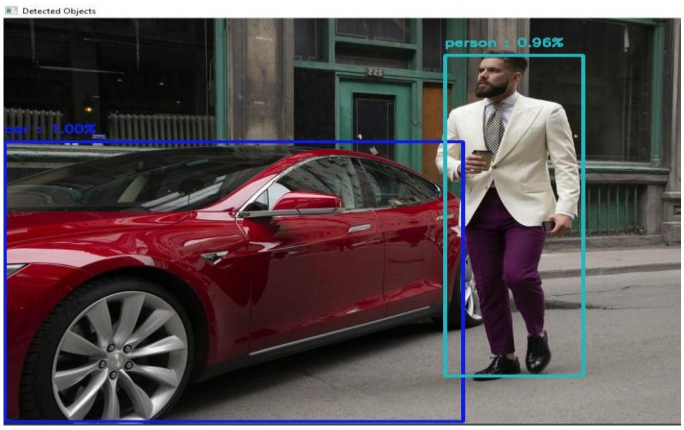

Fig -2: Objects(Person, Car) detected in image

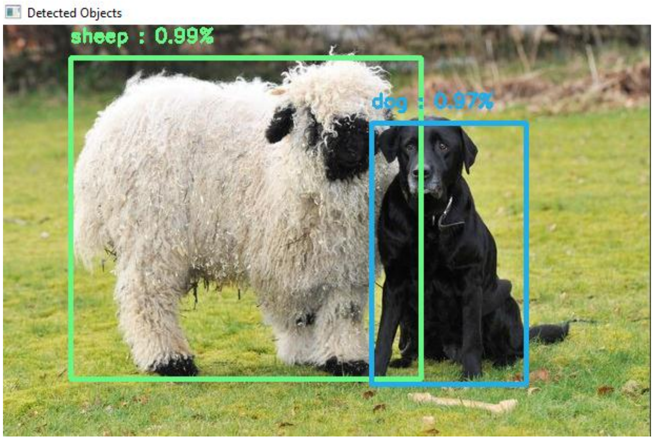

Fig -3: Objects (Animals) detected in image

Fig -4: Objects detected using webcam

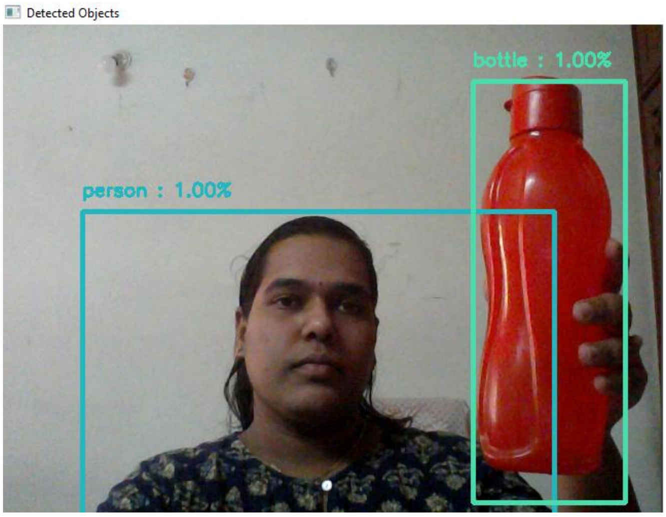

Fig -5: Objects detected using webcam

6. CONCLUSION

In this study, we developed a deep learning model for step-by-step identification of the position of objects in an image. The framework, like other best-in-class frameworks, could recognise the item with a good accuracy. In this manner, we employ an object detection module capable of recognising what's within the real-time video stream. It uses MobileNet and SSD frameworks to run modules to provide fast and productive object detection techniques based on deep learning. In the future, we can hold to enhance our detection model , which includes lowering reminiscence use and growing performance, in addition to including new classes of objects.

REFERENCES

[1].Y. Zhong, Y. Yang, X. Zhu, E. Dutkiewicz, Z. Zhou, T. Jiang, Device-free sensing for personnel detection in a foliage environment.

[2].S.Z. Su, S.Z. Li, S.Y. Chen, G.R. Cai, Y.D. Wu, A survey on pedestrian detection.

[3].M. Zeng, J. Li, Z. Peng, The design of top-hat morphological filter and application to infrared target detection. Infrared Physics & Technology

[4].L. Deng, D. Yu, Deep learning: methods and applications.Foundations and Trends in Signal Processing

[5].J. Redmon, S.Divvala, R.Girshick, A. Farhadi, You onlylook once: Unified, real-time object detection. In

[6].A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks. In

[7].Advances in neural information processing systemsH. Jiang, E. Learned-Miller, Face detection with the faster

[8].R-CNN. In 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition

[9].R. Girshick, J. Donahue, T. Darrell, J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition

[10].X. Peng, C. Schmid, Multi-region two-stream R-CNN for action detection. In European conference on computer vision (pp. 744-759). Springer, Cham. (2016, October).

[11].J. Redmon, A.Angelova, Real-time grasp detection using convolutional neural networks. In 2015 IEEE International Conference on Robotics and Automation (ICRA)

[12].R. Girshick, Fast r-cnn. In Proceedings of the IEEE international conference on computer vision

[13].S. Ren, K. He, R. Girshick, J. Sun, Faster r-cnn: Towards real-time object detection with region proposal networks. In Advances in neural information processing systems

[14].Y.J. Cao, G.M. Xu, G.C. Shi, Low altitude armored target detection based on rotation invariant faster R-CNN[J]. Laser & Optoelectronics Progress