Object Detection Using SSD MobileNet

Abstract:

Implemented object detection using the SSD (Single Shot MultiBox Detector) MobileNet model on a Linux-based operating system. The SSD MobileNet model is a popular deep learning algorithm for real-time object detection, known for its speed and accuracy.

Trained the SSD MobileNet model on a custom dataset, fine-tuning the pre-trained weights on the COCO (Common Objects in Context) dataset. The custom dataset consisted of various object categories relevant to the application domain.

Utilized the TensorFlow framework to train the model, leveraging transfer learning techniques and data augmentation to improve performance. The model achieved high accuracy and detection performance on the custom dataset.

Performed real-time object detection on images and videos using the trained SSD MobileNet model. The model efficiently processed each frame, detecting and localizing objects with bounding box coordinates and corresponding class labels.

Incorporated additional features such as inference time calculations and frames per second (FPS) measurements to assess the model's efficiency in real-time applications. The model achieved fast inference times and high FPS rates, making it suitable for real-time object detection scenarios.

The SSD MobileNet model demonstrated high performance and accuracy in detecting objects in real-time. The low inference time and high FPS rate make the model suitable for real-time applications where speed is crucial.This project enhanced understanding of deep learning techniques for object detection, improved skills in model training and

deployment on Linux systems, and provided hands-on experience with state-of-the-art computer vision algorithms.

System Configuration:

- System OS: Ubuntu 20.04
- RAM: 16GB
- Disk Space: 500GB
- Processor: Intel i5-7300U
- Graphics: Mesa Intel HD Graphics 620
- Python IDE: Visual Studio Code

Packages and Libraries Used:

- OpenCV
- NumPy
- Time

FPS and Inference Time WIth Different Input Size :

Sl.No	Size	Inference Time(ms)	FPS
1	(128, 128)	19.8037	34.5773
2	(144, 144)	27.9483	35.12611
3	(198, 198)	40.91072	20.6074
4	(224, 224)	31.9969	25.18133
5	(272, 272)	42.22297	21.43034
6	(320, 320)	67.5342	12.52644
7	(416, 416)	104.9768	8.44682
8	(464, 464)	124.6509	7.29330
9	(512, 512)	139.7950	6.47224

The table above shows the performance metrics of the SSD MobileNet model for various video resolutions. The inference time indicates the time taken by the model to process each frame and make predictions. The FPS (Frames Per Second) represents the number of frames processed by the model in one second.

As the video resolution increases, the inference time tends to increase, resulting in a lower FPS. This is expected as higher-resolution videos require more computational resources for object detection. It's important to consider the trade-off between accuracy and real-time performance when choosing the appropriate video resolution for object detection tasks.

Overall, the SSD MobileNet model demonstrates efficient object detection capabilities while maintaining real-time performance on a Linux-based system. The implemented project serves as a valuable tool for various computer vision applications and deep learning research.