SKP ENGINEERING COLLEGE

Approved by AICTE New Delhi | Affiliated to Anna University – Chennai
Tiruvannamalai, Tamil Nadu | Phone: +91-4175-252633 | +91-9443105139

DEPARTMENT OF INFORMATION TECHNOLOGY

IT3811 – Project work

Zeroth Review

PEOPLE COUNTING SYSTEM USING MOBILENET, SSD ALGORITHM BY DEEP LEARNING

Submitted by:

Manikandan. S	**(512221205010)**
Ramanan. S	**(512221205015)**
Ramkumar. A	**(512221205016)**

Project Co-Ordinator:

MS.M.SAMHITHA ME.,

Project Guide:

Dr.V.Raji ME.,Ph.D.,

Introduction:

In recent years, the demand for intelligent surveillance and crowd management solutions has significantly increased across various sectors such as transportation, retail, public safety, and smart cities. One such application is **automated people counting**, which enables systems to monitor and analyze human presence in real-time without manual intervention.

This project presents a **People Counting System** built using the **MobileNet Single Shot MultiBox Detector (SSD)** deep learning algorithm. The system utilizes a pre-trained deep neural network to detect and count the number of people in a given video feed or live camera stream. MobileNet SSD is known for its **lightweight architecture**, which makes it highly efficient and suitable for **real-time applications**, even on devices with limited computational power.

By leveraging **OpenCV** and **deep learning models**, each frame of the video is analyzed to detect the presence of individuals. The model identifies objects (in this case, people), draws bounding boxes around them, and keeps a running count of the detected persons. The result is then displayed in real time on the video feed, offering a seamless and interactive way to monitor crowd density.

This system can be deployed in real-world environments such as **public transport systems (buses, trains)**, **shopping malls**, **events**, or **offices** where knowing the number of occupants is crucial for **resource optimization**, **safety**, and **data analytics**.

Literature review:

- Deep Learning-Based Object Detection - One of the most effective approaches in modern object detection is the use of **Single Shot MultiBox Detector (SSD)**, introduced by Liu et al. in 2016. SSD performs object detection in a single forward pass, making it faster than region-based algorithms like R-CNN and its variants. It balances speed and accuracy by detecting multiple objects of different sizes and aspect ratios through default bounding boxes at multiple feature map scales.
- People Detection and Counting - People detection has been a core challenge in computer vision, especially in crowded scenes. Earlier methods used **background subtraction** and **motion-based tracking**, which suffered from occlusions and lighting variations. More recent solutions employ pretrained models like **YOLO (You Only Look Once)**, **Faster R-CNN**, and **MobileNet SSD**, which detect people directly from raw image data.
- Research Gap - While existing models like YOLO and Faster R-CNN offer high accuracy, they often demand significant computational resources. MobileNet SSD, although slightly less accurate, presents a balanced solution for real-time people counting in constrained environments. However, improvements in handling **occlusion**, **variable poses**, and **low-light conditions** remain open challenges.

Base paper:

- Developing a people counting system using MobileNet and the Single Shot MultiBox Detector (SSD) algorithm leverages advancements in deep learning to accurately and efficiently detect and count individuals in various environments. This approach is particularly beneficial for applications requiring real-time analytics, such as crowd management, retail analytics, and public safety monitoring.
- **MobileNet** is a class of efficient convolutional neural networks designed for mobile and embedded vision applications. It utilizes depth-wise separable convolutions to build lightweight deep neural networks, enabling the deployment of complex models on resource-constrained devices without significant loss of accuracy.
- **SSD (Single Shot MultiBox Detector)** is an object detection algorithm that discretizes the output space of bounding boxes into a set of default boxes over different aspect ratios and scales per feature map location. At prediction time, the network generates scores for the presence of each object category in each default box and adjusts the boxes to better match object shapes. This method eliminates the need for a separate proposal generation stage, making it faster and more efficient for real-time applications.
- Integrating MobileNet with SSD combines the strengths of both: the efficiency of MobileNet's lightweight architecture and SSD's capability for real-time object detection. This integration facilitates the development of people counting systems that can operate effectively on edge devices, providing timely data without the latency associated with cloud processing

Problem Statement:

- In environments such as public transportation, shopping malls, offices, and public events, the ability to **accurately count the number of people** in real time is essential for effective **crowd management, resource allocation, security, and data analytics**.
- With the advancement of **deep learning and computer vision**, there is an opportunity to develop a system that can **automatically detect and count people** using video input with high accuracy and real-time performance.
- The problem this project addresses is the **design and implementation of a lightweight, real-time, and accurate people counting system** using the **MobileNet SSD (Single Shot MultiBox Detector)** algorithm, which can be effectively deployed in practical scenarios for **real-time crowd analysis** without the need for high-end hardware.

Objective of the problem:

- **To design and implement an object detection-based system** capable of detecting and counting people in video streams using deep learning.
- **To utilize the MobileNet SSD model** for real-time object detection, providing a balance between speed, accuracy, and computational efficiency.
- **To process input from video files or live camera feeds** and continuously update the count of detected individuals on each frame.
- **To visualize detections by drawing bounding boxes and labels** around people detected in the video frames.
- **To create a lightweight and portable solution** that can run on low-power or embedded devices (such as Raspberry Pi or Jetson Nano) without requiring high-end GPUs.
- **To enable future scalability**, such as integrating cloud-based databases (e.g., Firebase) for data storage or analytics.

Thankyou