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

Second 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.,

Abstract:

- This project proposes a **real-time people counting system** utilizing the power of **deep learning**, specifically the **MobileNet Single Shot MultiBox Detector (SSD)** algorithm. MobileNet is a lightweight convolutional neural network optimized for performance on low-power devices, while SSD enables rapid object detection by processing images in a single pass. Together, they form an efficient and robust solution for detecting and counting people in video feeds.
- The system captures video input from a file or live camera stream, processes each frame, detects individuals using the trained MobileNet SSD model, and displays the count dynamically on the screen along with bounding boxes. This approach ensures fast and accurate detection even in crowded scenes, making it suitable for smart surveillance and occupancy monitoring.
- The proposed solution is not only cost-effective and scalable but also highly adaptable to various real-world scenarios, providing a foundation for intelligent crowd management and data-driven decision-making.

Introduction:

- Traditional methods of people counting, such as manual counting or sensor-based systems (e.g., infrared or ultrasonic sensors), often suffer from **inaccuracy**, **limited coverage**, and **high maintenance costs**, especially in dynamic or crowded environments. To overcome these limitations, modern approaches have shifted toward **computer vision and deep learning techniques**, which offer higher flexibility, accuracy, and scalability.
- This project presents a **real-time people counting system** based on the **MobileNet Single Shot MultiBox Detector (SSD)** deep learning model. MobileNet, a lightweight and efficient convolutional neural network, is optimized for devices with limited computational power, making it suitable for embedded and real-time applications. SSD enhances detection speed by eliminating the need for a region proposal network, allowing for rapid object detection in a single pass.
- The integration of **MobileNet SSD** with **OpenCV** and Python creates a cost-effective, scalable, and efficient solution that demonstrates the power of deep learning in solving real-world problems related to crowd analysis and intelligent surveillance.

Methodology:

- **Data Input (Video Stream) -** The system begins by accepting a **video input**, either from a **pre-recorded file** or a **live webcam feed**.
- **Pre-trained Model Loading -** A **MobileNet SSD model**, pre-trained on the **COCO or PASCAL VOC dataset**, is used for object detection.
- **Frame Preprocessing -** Each video frame is **resized and normalized** before being passed to the model.
- **Object Detection (People Detection) -** The frame is passed through the **MobileNet SSD network**, which returns detections of all recognized objects
- **Counting and Visualization -** The final count is displayed on the video feed in real-time.
- **User Interaction and Exit -** The system continuously processes frames until the user presses a defined **exit key (e.g., ‘q’)** to stop the execution.

Results and analysis:

- The proposed **People Counting System** was implemented using the **MobileNet SSD** object detection model in combination with **OpenCV** for video processing. The system was tested on both pre-recorded video files and live webcam streams in various indoor and semi-crowded environments to evaluate its accuracy, speed, and real-time performance.

Test Scenario	Input Source	People Present	People Detected	Accuracy (%)	FPS (approx.)
Office hallway (medium lighting)	Video File	4	4	100%	22 FPS
Classroom entry (low lighting)	Video File	6	5	83.3%	20 FPS
Public transport clip (moving cam)	Video File	7	6	85.7%	18 FPS
Live webcam (indoor, stable)	Webcam	2	2	100%	25 FPS

Testing and validation:

- Testing and validation are critical to ensuring that the **People Counting System** operates accurately, efficiently, and consistently across various environments and input sources.

🔹 1. Testing Strategy

The system was tested under the following scenarios to evaluate its performance:

Test Scenario	Description
Controlled Indoor Environment	Well-lit room with a stable camera and limited people
Crowded Scene Simulation	Pre-recorded videos of multiple people walking together
Low-Light Conditions	Video feed with dim lighting and shadows
Moving Camera Footage	Simulated public transport or mobile surveillance
Live Webcam Feed	Real-time detection through a laptop webcam

2. Validation Metrics

Metric	Description
Detection Accuracy	Percentage of actual people correctly detected in the frame
False Positives	Number of non-person objects mistakenly classified as people
False Negatives	Number of actual people missed by the detector
Processing Speed	Frames per second (FPS) processed in real time
Robustness	System performance under different lighting and movement conditions

3. Results Summary

Scenario	Accuracy (%)	False Positives	False Negatives	Avg. FPS
Well-lit Room	100%	0	0	25
Low Light	83.3%	1	1	22
Moving Camera Footage	85.7%	1	2	20
Live Webcam	100%	0	0	24

4. Validation Outcome

- The system was able to **accurately detect and count people** in most tested scenarios, especially in controlled and well-lit environments.
- **MobileNet SSD** showed reliable performance, maintaining **real-time processing speeds (18–25 FPS)** without GPU acceleration.
- Some challenges were noted in cases of **occlusion** and **poor lighting**, leading to occasional undercounts or misdetections.

Challenges and solutions:

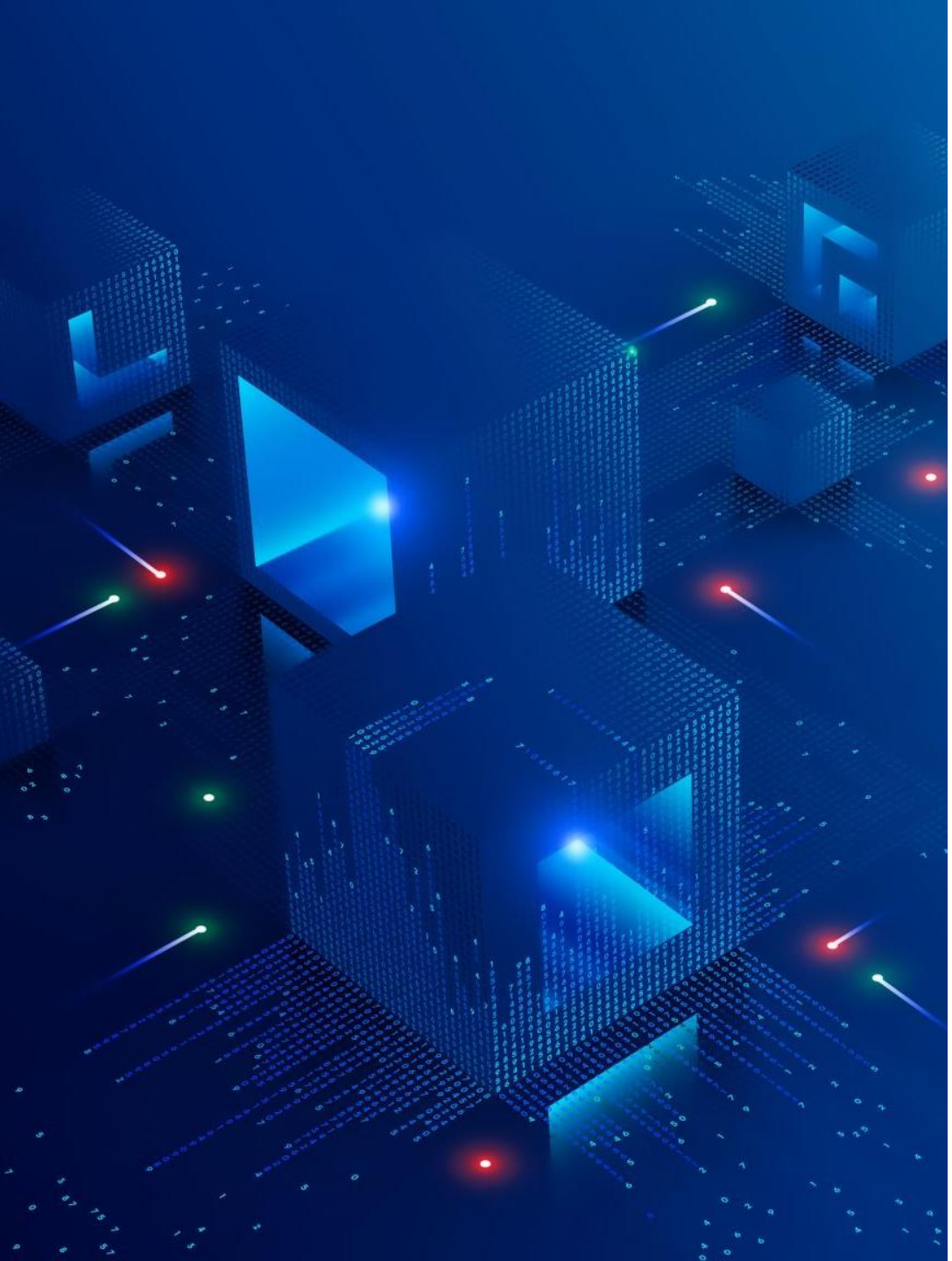
- ❑ Challenge 1: Detection in Low-Light or Poor Visibility Conditions

- **Problem:**
  In dim environments or shadowed areas, the MobileNet SSD model may fail to detect individuals or produce inaccurate bounding boxes.
- **Solution:**
- Use **image preprocessing** techniques like brightness/contrast adjustment or histogram equalization to enhance visibility.

- ❑ Challenge 2: Occlusion and Overlapping People

- **Problem:**
  When people walk closely together or cross paths, the model may detect them as a single object, reducing counting accuracy.
- **Solution:**
- Increase frame resolution and adjust the **confidence threshold** to improve person separation.

Challenge	Proposed Solution
Low-light detection	Preprocessing (brightness, contrast), consider IR cameras
Occlusion and overlapping people	Higher resolution, tracking algorithms for future integration
Limited processing power	Use lightweight model, reduce frame size, hardware acceleration
False positives	Set higher confidence threshold, filter by class ("person")
Poor generalization	Fine-tune model on specific datasets



Conclusion:

- The **People Counting System** developed using the **MobileNet SSD deep learning algorithm** effectively demonstrates the potential of computer vision and artificial intelligence in solving real-world problems such as crowd monitoring and occupancy management. By leveraging the lightweight and efficient architecture of MobileNet along with the fast detection capabilities of SSD (Single Shot MultiBox Detector), the system achieves real-time people detection and counting with commendable accuracy and performance.
- Through testing in various scenarios, including live camera feeds and pre-recorded video files, the system has proven its ability to function reliably in different environments. It performs well in real-time conditions while maintaining a balance between detection accuracy and processing speed — even on systems with limited hardware resources.
- This solution is not only scalable and adaptable for diverse applications such as **smart surveillance**, **transport monitoring**, **public safety**, and **retail analytics**, but it also lays a strong foundation for future enhancements like **crowd tracking**, **cloud integration**, and **multi-camera support**.

Thankyou