International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 12 | Dec - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Crowd Sense: A Real-Time System for Crowd Analytics

1st Dr. Poornima Raikar
Department of Computer Science and Engineering (AIML)
KLS Vishwanathrao Deshpande Institute of Technology Haliyal, India
pnr@klsvdit.edu.in

2nd Dr. Venkatesh Shankar
Department of Computer Science and Engineering
KLS Vishwanathrao Deshpande Institute of Technology Haliyal, India
vss@klsvdit.edu.in

3rd Mr. Gopal Muri
Department of Computer Science and Engineering
KLS Vishwanathrao Deshpande Institute of Technology Haliyal, India
gopalmuri1919@gmail.com

4th Mr. Pradeep Pai
Department of Computer Science and Engineering
KLS Vishwanathrao Deshpande Institute of Technology Haliyal, India
paipradeep4123@gmail.com

5th Mr. Aditya Wandakar
Department of Computer Science and Engineering
KLS Vishwanathrao Deshpande Institute of Technology Haliyal, India
adityagw20@gmail.com

Abstract -Real-time crowd monitoring has become increasingly important in public and commercial environments such as shopping malls, campuses, transportation hubs, and event venues. Traditional surveillance systems rely on manual observation of video footage and do not provide meaningful real-time analytics. This paper presents Crowd Sense, a real-time crowd analytics system that detects and counts people from live video streams and performs gender classification using lightweight deep-learning models. The system employs MobileNet-SSD for person detection and a gender classification model for identifying male and female individuals. A Flask-based web dashboard integrated with SocketIO is developed to display live video, people count, gender distribution, and analytics updates without page refresh. The proposed system is designed to run efficiently on standard CPU-based hardware, eliminating the need for expensive GPU resources. Experimental results demonstrate reliable real-time performance under normal lighting conditions, making the system cost-effective, scalable, and suitable for practical deployment in real-world environments.
Keywords -- Crowd Analytics, Real-Time Monitoring, People Detection, Gender Classification, Computer Vision, MobileNet-SSD, Flask Dashboard

I. INTRODUCTION:

Real-time crowd monitoring has become increasingly important in modern public environments such as shopping malls, educational campuses, transportation hubs, offices, stadiums, and large events. With growing population density and frequent public gatherings, ensuring safety, efficient management, and optimized resource allocation has become a major challenge. Traditional CCTV systems, although widely deployed, merely record video footage and depend entirely on human operators for observation. This manual approach is inefficient, inconsistent, and susceptible to fatigue-related errors, especially when continuous monitoring is required.
Moreover, conventional surveillance systems lack analytical capabilities. They cannot automatically provide meaningful insights such as the total number of people in a scene, gender distribution, or real-time variations in crowd flow. During peak hours or emergencies, such limitations can delay decision-making and result in poor situational management.
Although automated crowd detection systems exist, most rely on computationally heavy deep-learning models like YOLO or Faster R-CNN, which demand high-performance GPU-based hardware. This makes them expensive and unsuitable for small businesses, educational institutions, and public offices that operate within limited budgets. Many solutions are also proprietary, restrict customization, or require cloud-based subscriptions, increasing cost and dependency.
In addition, several existing systems lack integrated dashboards to present analytics in an intuitive and real-time manner, making them less usable for non-technical administrators. Therefore, there is an urgent need for a lightweight, efficient, low-cost real-time crowd analytics system that can detect people, classify gender, and present analytics instantly — all using standard consumer-grade computer hardware and a simple web-based interface.

II. METHODS:

2.1 Materials Used

The **Crowd Sense** system was developed using a combination of standard hardware components, open-source software tools, and pre-trained deep-learning models to enable real-time crowd analytics on CPU-based systems. A webcam serves as the primary input device to capture live video streams from the monitored environment. All processing is performed on a personal computer or laptop equipped with a standard CPU and sufficient system memory, eliminating the need for GPU acceleration.
The software implementation is carried out using **Python**, selected for its extensive support for computer vision and machine learning applications. **OpenCV** is utilized for video acquisition, frame preprocessing, face detection, and visualization of detection results. For person detection, a pre-trained **MobileNet-SSD** model is employed due to its lightweight architecture and suitability for real-time execution on CPU hardware. Gender classification is performed using a pre-trained deep-learning model that analyzes detected facial regions. The system's web-based interface is developed using the **Flask** framework, while **Flask-SocketIO** enables real-time communication between the backend processing module and the dashboard.

2.2 Key Procedures and Techniques

The Crowd Sense system follows a well-defined and systematic sequence of procedures to perform real-time crowd analytics efficiently. The process begins with continuous video acquisition using a webcam, which captures live video streams from the surrounding environment. These video frames are obtained in real time using the OpenCV library, ensuring smooth and uninterrupted input to the system.
Once a frame is captured, it undergoes a preprocessing stage to improve detection efficiency. During this stage, the frame is resized and converted into a suitable format required by the detection model. Preprocessing helps reduce computational load and ensures faster execution, which is essential for maintaining real-time performance on CPU-based systems.
After preprocessing, person detection is performed on each frame

using the MobileNet-SSD model. This model identifies all individuals present in the scene and generates bounding boxes around detected persons. The bounding boxes provide spatial information that helps isolate individual regions for further analysis. For every detected person, the system extracts the corresponding facial region using a face detection technique. This step ensures that only relevant facial features are considered for demographic analysis.

The extracted face regions are then passed to the gender classification model, which predicts whether the detected individual is male or female. The classification results are combined with the detection data to compute real-time statistics such as total number of people, male count, and female count. These values are updated continuously as new frames are processed. Finally, the computed analytics are transmitted to a web-based dashboard using real-time communication techniques, allowing users to view live results instantly without manual page refresh.

2.3 Algorithms Used

The Crowd Sense system employs a combination of lightweight deep-learning models and classical computer vision algorithms to achieve efficient real-time crowd detection and demographic analysis. Each algorithm is selected with the objective of maintaining high processing speed and acceptable accuracy while operating on standard CPU-based hardware. By integrating these algorithms into a structured pipeline, the system ensures smooth real-time performance without relying on GPU acceleration. The major algorithms used in the system and their respective roles are described below.

- **MobileNet-SSD Algorithm (Person Detection)**

MobileNet-SSD (Single Shot Detector) is used as the primary algorithm for person detection in the Crowd Sense system. This deep-learning model is specifically designed for efficient object detection on devices with limited computational resources. Unlike traditional detection models that require multiple processing stages, MobileNet-SSD performs detection in a single forward pass, making it suitable for real-time applications.

The model processes each incoming video frame and identifies individuals present in the scene by generating bounding boxes around detected persons. One of the key advantages of MobileNet-SSD is its use of depthwise separable convolutions, which significantly reduce the number of computations required during inference. This architectural design enables fast execution on CPU-based systems while maintaining reliable detection accuracy. Along with bounding box coordinates, the model also outputs confidence scores for each detection. To improve overall accuracy and reduce false positives, detections with low confidence scores are filtered out before further processing.

- **Haar Cascade Algorithm (Face Detection)**

For face detection, the Crowd Sense system utilizes the Haar Cascade algorithm, a classical computer vision technique known for its simplicity and speed. Haar Cascade works by scanning the image using a set of rectangular filters and comparing pixel intensity differences with pre-trained Haar features. This approach allows the algorithm to quickly identify facial patterns within an image.

In the proposed system, face detection is applied only within the regions identified by the person detection stage. By restricting face detection to detected person bounding boxes, unnecessary computation is avoided, thereby improving processing efficiency. Haar Cascades are lightweight and require minimal computational resources, making them well-suited for real-time face detection on standard hardware. Their fast execution speed ensures that face detection does not introduce delays in the overall processing pipeline.

- **GenderNet Classification Algorithm (Gender Prediction)**

Gender classification in the Crowd Sense system is performed using GenderNet, a pre-trained deep-learning model designed to classify gender from facial images. Once a face is detected using the Haar Cascade algorithm, the cropped facial region is passed as input to the GenderNet model. The model analyzes facial features and produces probability values corresponding to male and female classes.

The predicted gender is determined by selecting the class with the higher probability value. GenderNet is optimized for fast inference and operates effectively on CPU-based systems, making it suitable for real-time deployment. The integration of GenderNet enables the system to extract meaningful demographic information from live video streams, supporting applications such as crowd analysis and targeted advertisement display.

- **Real-Time Streaming Algorithm (Flask-SocketIO)**

To enable real-time visualization of detection and analytics results, the Crowd Sense system uses Flask-SocketIO for data streaming between the backend and the dashboard. Unlike traditional web applications that require frequent page refreshes, Flask-SocketIO uses WebSocket-based communication to push updates instantly to the client interface.

This real-time streaming mechanism allows the dashboard to display live video frames, bounding boxes, gender labels, and analytic values such as total people count, male count, and female count without delay. By maintaining continuous communication between the processing module and the user interface, the system ensures smooth interaction and immediate feedback. This approach significantly enhances usability and makes the system suitable for real-time monitoring scenarios.

2.4 System Architecture

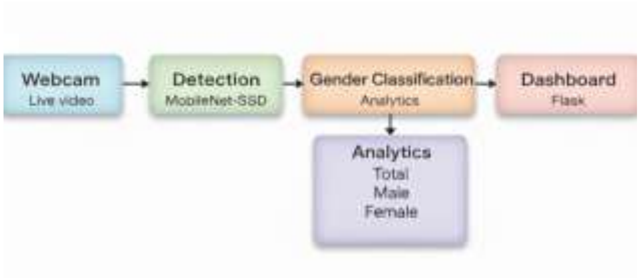

Fig 2.1: System Architecture of the Crowd Sense Real-Time Analytics

The system architecture of Crowd Sense is designed as a simple yet efficient real-time processing pipeline to ensure smooth data flow from video capture to visualization. The architecture begins with a webcam that continuously captures live video frames from the environment. These frames are forwarded to the processing module, which serves as the core component of the system.

Within the processing module, person detection, face detection, and gender classification are performed sequentially on each frame. The detected individuals are marked with bounding boxes, and gender labels are assigned based on facial analysis. The results obtained from this stage are then passed to the analytics module, where real-time statistical metrics such as total people count and gender distribution are computed.

The analytics module continuously updates these values based on incoming frames and forwards both the processed video frames and computed statistics to the backend server. A Flask-based backend, integrated with SocketIO, enables real-time data transmission to the web dashboard. The dashboard displays the live video feed along with dynamically updated analytics, ensuring responsive interaction and clear visualization. This architecture allows the system to operate efficiently on standard hardware while delivering real-time crowd insights in a user-friendly manner.

2.5 Statistical Analysis Methods

Basic statistical methods are applied to analyze and present crowd analytics results. The system computes descriptive statistics, including total count, male count, and female count, from each processed frame. These values are updated continuously to reflect real-time changes in crowd composition.

To minimize sudden fluctuations caused by temporary detection errors, frame-wise smoothing is applied by averaging values across consecutive frames. The final results are displayed on the dashboard using numerical indicators and visual charts, enabling easy interpretation of crowd trends over time. These statistical methods ensure stable and meaningful representation of real-time crowd data without introducing additional computational overhead.

III. RESULTS

The performance of the proposed **Crowd Sense** system was evaluated using live video streams captured through a webcam in indoor environments. The system was tested under different lighting conditions and varying crowd sizes to analyze its real-time detection capability and reliability.

The MobileNet-SSD model successfully detected multiple individuals in real time. Bounding boxes were accurately drawn around detected persons, demonstrating stable and consistent detection in moderately crowded scenes. The system maintained continuous detection without frame drops during normal operation.

Face detection using the Haar Cascade algorithm performed reliably when faces were clearly visible and frontal. Based on the detected facial regions, the GenderNet model classified individuals as male or female in real time. The system continuously updated the total people count along with male and female counts as new frames were processed.

The system achieved an average processing speed of approximately **12–18 frames per second (FPS)** on a standard CPU-based computer without GPU acceleration. This confirms that the proposed system is capable of real-time crowd analytics using lightweight deep-learning models. The system operated continuously without crashes or significant delays during testing.

The real-time dashboard successfully displayed live video, bounding boxes, and crowd analytics. Using Flask-SocketIO, the dashboard updated automatically without page refresh, enabling smooth and uninterrupted visualization of real-time crowd statistics.

Fig 3.1: Live Detection Output

This **Fig 3.1** shows the live detection output of the Crowd Sense system during execution. The webcam feed is processed in real time to detect individuals, and bounding boxes are drawn around each detected person. Gender labels are displayed based on facial analysis, and the corresponding analytics such as total people count and gender distribution are updated instantly on the dashboard. This image confirms the system's ability to perform real-time detection and visualization effectively.

Fig 3.2: Real-Time Analytics Display

This **Fig 3.2** shows the real-time analytics panel of the Crowd Sense system. It presents statistical information such as crowd size trend over time, current people count, gender-wise distribution, and advertisement display statistics. The analytics update dynamically as the system processes live video input. This display helps administrators monitor crowd behavior, demographic patterns, and advertisement performance in an intuitive and organized manner.

IV. DISCUSSIONS

The results obtained from the experimental evaluation demonstrate that the proposed Crowd Sense system effectively performs real-time crowd monitoring and demographic analysis using lightweight models on standard CPU-based hardware. The successful detection of multiple individuals in live video streams confirms the suitability of the MobileNet-SSD model for real-time person detection in moderately crowded indoor environments. The ability of the system to maintain stable detection without significant frame drops highlights the efficiency of the chosen detection approach.

The integration of the Haar Cascade algorithm for face detection and the GenderNet model for gender classification enabled the extraction of meaningful demographic information from live video input. Gender classification results were reliable when facial features were clearly visible and lighting conditions were adequate. However, reduced accuracy was observed in scenarios involving partial occlusion, side-facing individuals, and low-light conditions. These limitations are inherent to vision-based systems that rely on facial features and indicate the need for improved robustness in challenging environments.

The achieved processing speed of approximately 12–18 frames per second demonstrates that real-time analytics can be achieved without GPU acceleration. This performance is particularly significant for deployment in cost-sensitive environments such as educational institutions, small retail spaces, and public offices, where access to high-performance hardware is limited. The results confirm that lightweight deep-learning architectures can provide a practical balance between accuracy and computational efficiency.

The real-time dashboard played a crucial role in enhancing system usability. By utilizing Flask-SocketIO for continuous data transmission, the dashboard provided instantaneous updates of live video, people count, and gender distribution without requiring manual page refresh. This real-time visualization improves situational awareness and supports timely decision-making for administrators and security personnel.

Despite its effectiveness, the system has certain limitations. Performance decreases in highly crowded scenes where individuals overlap, and gender classification accuracy depends heavily on clear facial visibility. Additionally, the current implementation supports only a single camera feed, which limits scalability for large environments. Nevertheless, these limitations present opportunities for future enhancements such as multi-camera integration, improved face detection models, and advanced analytics.

Overall, the discussion confirms that the Crowd Sense system offers a reliable, low-cost, and deployable solution for real-time crowd analytics. By prioritizing efficiency, simplicity, and real-time performance, the system addresses practical challenges associated with traditional surveillance methods and provides a strong foundation for future extensions and improvements.

V. CONCLUSIONS

The *Crowd Sense* system successfully demonstrates an efficient and real-time solution for crowd analytics using lightweight computer vision models. By integrating MobileNet-SSD for person detection and GenderNet for gender classification, the system provides accurate demographic insights without requiring high-performance GPU hardware. The use of OpenCV for frame processing and Flask-SocketIO for continuous dashboard updates ensures smooth, real-time visualization of analytics, making the system practical for real-world environments such as malls, campuses, and public gathering spaces.

The implementation confirms that reliable crowd monitoring can be achieved on standard computing devices through optimized models and an effective processing pipeline. While the system performs well under normal conditions, some limitations exist, particularly related to lighting, face visibility, and single-camera support. However, these challenges open potential areas for future enhancement.

Overall, *Crowd Sense* provides a cost-effective, scalable, and user-friendly platform for real-time crowd monitoring and demographic analysis. It serves as a strong foundation for further advancements such as multi-camera deployment, improved child detection models, behavioural analytics, and deeper integration with smart systems and IoT technologies.

VI. REFERENCES

1] Y. Ghrban and N. K. El Abbadi, “Gender Classification from Face and Eyes Images Using Deep Learning Algorithm,” J. Computer Science, vol. 19, no. 3, pp. 345-362, 2023. thescipub.com

2] A. A. Lubis, A. Prasasta, and D. A. Sari, “Towards Efficient Crowd Counting and Behavior Analysis Using YOLOv11,” None, Apr. 2025, doi: 10.63876/ijtm.v4i1.128.

3] Y. Jiang, Y. Miao, B. Alzahrani, A. Barnawi, R. Alotaibi, and L. Hu, “Ultra Large-Scale Crowd Monitoring System Architecture and Design Issues,” Institute of Electrical and Electronics Engineers, Apr. 2021, doi:

4] Y. Ying, H. Zhu, Q. Jin, C. Pan, and D. Miao, “Survey on Deep Learning Based Crowd Counting,” J. Comput. Res. Dev., vol. 58, no. 12, pp. 2724-2747, 2021. crad.ict.ac.cn

5] H. Honmote, S. Gadekar, P. Katta, and M. Kulkarni, “Real Time Object Detection and Recognition using MobileNet-SSD with OpenCV,” Int. J. Eng. Res. Technol. (IJERT), vol. 11, no. 01, Jan. 2022. ijert.org

6] Z. S. A. Ghrban and N. K. E. Abbadi, “Gender Classification from Face and Eyes Images Using Deep Learning Algorithm,” J. Comput. Sci., vol. 19, no. 3, pp. 345-362, 2023. thescipub.com+1

7] N. R. Hamza, “Gender Classification from Human Face Images Using Deep Learning Based on MobileNetV2 Architecture,” J. Qadisiyah for Comput. Sci. & Math., 2025. jqcsm.qu.edu.iq

8] Y. Hao, H. Du, M. Mao, Y. Liu, and J. Fan, “A Survey on Regression-Based Crowd Counting Techniques,” Information Technology and Control, vol. 52, no. 3, 2023. itc.ktu.lt[9] Y. Wang and F. Luo, “A Survey of Crowd Counting Algorithm Based on Domain Adaptation,” Acad. J. Sci. Technol., vol. 5, no. 2, 2025. drpress.org

9] Y. Gu, M. Wu, Q. Wang, S. Chen and L. Yang, “A Deep Learning-Based Crowd Counting Method and System Implementation on Neural Processing Unit Platform,” Computers, Materials & Continua, vol. 75, no. 1, pp. 493-512, 2023. techscience.com

10] A. George, V. Vinothina and J. Beulah G., “Neural Network-Based Crowd Counting Systems: State of the Art, Challenges, and Perspectives,” J. Advances in Information Technology, vol. 14, no. 6, pp. 1450-1460, 2023. jait.us

11] “Texture features-based lightweight passive multi-state crowd counting algorithm,” EURASIP J. Wireless Communications & Networking, 2023, article 79. SpringerOpen

12] G. Tahyudin, M. Sulistiyo and M. Arzaki, “Classifying Gender Based on Face Images Using Vision Transformer,” Int. J. on Informatics Visualization, vol. 8, no. 1, 2024. joiv.org

13] A survey: “A survey of deep learning methods for density estimation and crowd counting,” 2025, Vicinagearth , vol. 2, article 2. SpringerLink

14] N. Ryadh Hamza, “Gender Classification from Human Face Images Using Deep Learning Based on MobileNetV2 Architecture,” J. Al-Qadisiyah for Computer Science & Mathematics, 2025. jqcsm.qu.edu.iq

15] Q. M. B. Aljelawy, E. K. Hanoun and G. M. Ali, “Real-time People Counting with Deep Learning: A Solution for Crowd Management,” Al-Mansour J., vol. 42, no. 1, pp. 120-127, 2025. journal.muc.edu.iq