IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 6, November 2024

Real-Time People Counting System

Devesh Kumar[1] and Dr. Chitra K[2]

Student MCA, IVth Semester[1]

Assistant Professor, Department of MCA[2]

Dayananda Sagar Academy of Technology and Management, Udaypura, Bangalore, Karnataka, India

deveshkumar10.dk@gmail.com and chitra-mca@dsatm.edu.in

Abstract*: This paper introduces an advanced system for real-time people counting, utilizing live video streams or IP cameras. The core of this system is the Single Shot Multibox Detector (SSD) framework, enhanced with MobileNet architecture to ensure precise object detection. To continuously track detected individuals, we employ a centroid tracking algorithm. The system's design prioritizes scalability, making it ideal for various business applications, such as monitoring foot traffic in stores, buildings, and shopping malls. A key feature of our system is its ability to send immediate alerts when the number of people exceeds a set threshold. This functionality is crucial for managing crowd control and adhering to safety protocols, particularly in scenarios like the COVID-19 pandemic. We have optimized the system's performance with threading, which improves its responsiveness and efficiency. By integrating advanced machine learning techniques with practical business applications, our real-time people counting system offers a dependable and scalable solution for contemporary crowd management challenges. Features like real-time alerts, automatic scheduling, and comprehensive data logging highlight its effectiveness in enhancing operational efficiency and safety across various settings.*

Keywords: Real-time people counting, Single Shot Multibox Detector (SSD), MobileNet architecture, centroidtracking, machine learning, Efficiency, Scalability

I. INTRODUCTION

Efficiently managing human traffic in various environments is crucial for optimizing resource allocation, enhancing customer experiences, and ensuring public safety. Traditional methods like manual counting and basic sensor systems often struggle with accuracy and adaptability, leading to inefficiencies in crowd management.

Recent advancements in computer vision and machine learning offer promising solutions to these challenges. This research aims to develop a robust, real-time people-counting system that leverages visual data processing for precise identification and counting of individuals. By harnessing these technologies, the system aims to provide accurate counts in dynamic and crowded environments.

The demand for accurate people-counting systems has been underscored by events like the COVID-19 pandemic, emphasizing the critical importance of maintaining safe distancing measures. A reliable counting system can play a crucial role in enforcing these protocols and supporting public health efforts. Through the integration of advanced technologies, our system seeks to offer scalable solutions to modern challenges in crowd management.

II. LITERATURE REVIEW

The evolution of people counting methods shows a growing need for precise and effective ways to manage crowds. Early methods, like manual counting and turnstile devices, offered basic functionality but were impractical for dynamic environments, often leading to inaccuracies and operational limitations. [1]

A major improvement came with digital sensors and surveillance cameras. Early versions using infrared sensors or pressure mats were more efficient but still struggled to accurately distinguish between individuals in crowded settings and adapt to different environmental conditions. [2]

Recent advancements in computer vision and machine learning have greatly changed the field of people counting. Techniques like convolutional neural networks (CNNs) have been very successful in tasks like object detection and

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 6, November 2024

recognition. For example, a model that divided images into regions and predicted bounding boxes and probabilities for each region set new standards for real-time object detection, especially in crowded environments. [3]

Another study introduced a simple and effective solution for tracking multiple objects in real-time. This straightforward design makes it ideal for dynamic environments where people counting is necessary. Additionally, integrating region proposal networks with object detection networks enabled near-real-time object detection. This significantly improved the ability of automated systems to count people accurately in complex environments by making object detection faster and more precise. [4]

Maintaining accurate trajectories of detected individuals over time is also crucial. A study on global data association for multi-object tracking using network flows emphasized this, highlighting its importance for reliable people counting in surveillance systems. Further advancements came with deep residual learning, which made it easier to train very deep neural networks. This opened new possibilities for improving the accuracy of people counting systems by using more complex models. [5]

Another approach redesigned computer vision architecture, leading to significant improvements in object detection accuracy and efficiency. This modular design allows for efficient scaling, making it suitable for real-time applications like people counting. A model using a single deep neural network pass to detect objects in images combined with speed and accuracy, making it a strong candidate for real-time people counting systems. The integration of machine learning algorithms with computer vision systems has further improved the accuracy and reliability of people counting. Fine-tuning these models on specific datasets has been effective in enhancing detection accuracy across different environments. Advanced tracking algorithms, such as the centroid tracker, enhance performance by consistently identifying individuals across consecutive frames. [6]

Scalability is important for modern people counting systems to handle large datasets efficiently without compromising performance. Implementing multi-threading and parallel processing techniques ensures rapid data processing and analysis, supporting real-time decision-making and resource allocation. [7]

Another model introduced Feature Pyramid Networks for object detection, which improved the ability to detect objects at various scales. This advancement is particularly beneficial for people counting in environments with different distances and perspectives. A deep network, despite its simplicity, has been very successful in various image recognition tasks. Its deep architecture provides a strong framework for feature extraction, which is crucial for detecting and counting individuals in diverse and complex environments. [8]

Finally, the concept of Histograms of Oriented Gradients (HOG) for human detection has been foundational in the development of many subsequent object detection algorithms. HOG's ability to capture edge and gradient structure information has been essential for accurate human detection and counting. [9]

The shift from manual and basic sensor-based methods to advanced computer vision and machine learning techniques represents a major technological advancement. State-of-the-art models combined with effective tracking algorithms provide reliable solutions for accurate and efficient people counting. These advancements promise to enhance crowd management, optimize resource allocation, and improve safety in public and commercial spaces. [10]

III. METHODOLOGY

The proposed system uses a multi-stage approach: video acquisition, pre-processing, object detection, tracking, and counting.

VIDEO ACQUISITION

High-resolution IP cameras are strategically placed to maximize coverage. These cameras continuously capture video streams, which are transmitted to a central processing unit.

PRE- PROCESSING

Pre-processing enhances video data quality through resizing, noise reduction, and frame differencing to highlight moving objects. These steps ensure efficient and accurate detection and tracking.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, November 2024

SINGLE SHOT DETECTOR (SSD)

The SSD model, combined with MobileNet, is used for object detection. SSD is a single-shot detector that detects objects in an image in one forward pass, making it faster compared to two-shot detectors like R-CNN.

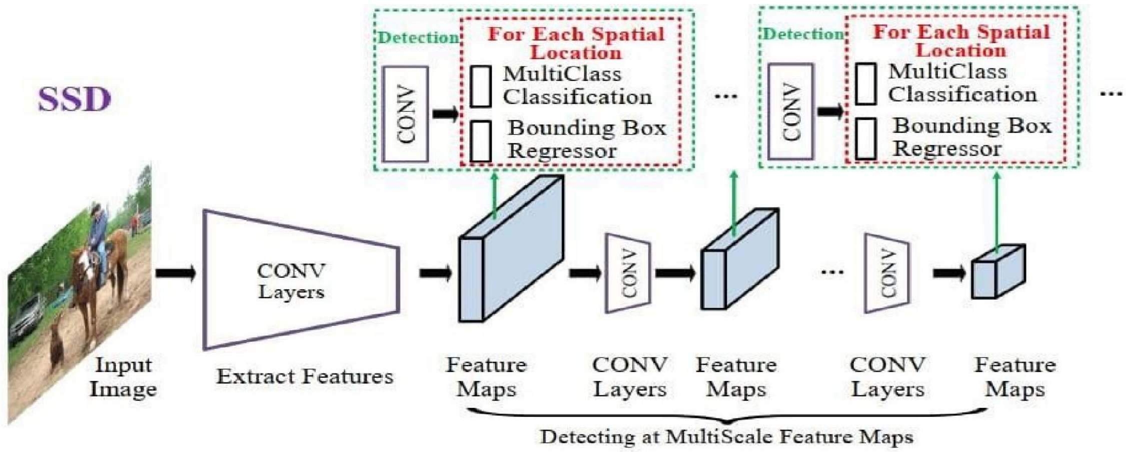

Figure 1: Single Shot Detector (SSD)

MOBILENET

MobileNet is a lightweight, efficient deep neural network designed for mobile and embedded vision applications. It reduces the computational cost while maintaining accuracy, making it suitable for real- time applications.

CENTROID TRACKER

The centroid tracker assigns a unique ID to each detected object and tracks its centroid over subsequent frames. This approach is computationally efficient and reliable for tracking multiple objects.

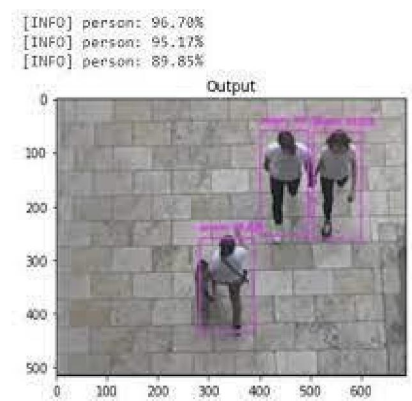

Figure 2: Detection of People in Open Space along with bounding box coordinates

COUNTING

The counting module increments the count when an individual crosses a predefined region of interest (ROI), such as an entry or exit point. The system ensures accurate counting by considering movement direction and avoiding multiple counts for the same individual.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, November 2024

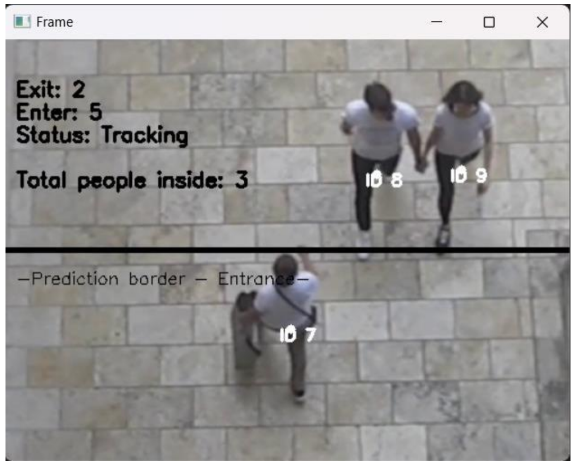

Figure 3: GUI Application

IV. FEATURES

REAL-TIME ALERTS

When the number of people exceeds the threshold, an alert is sent via email. This feature is controlled by the ALERT and Threshold settings in the configuration file.

THREADING

Threading is used to improve performance by reducing the lag between frame capture and processing. This is particularly useful for real-time applications where high frame rates are essential.

SCHEDULER

The scheduler feature allows the system to run at specific times, reducing the computational load when the system is not in use.

TIMER

The timer stops the system after a predefined duration, ensuring that resources are not wasted by running the system longer than necessary.

DATA LOGGING

The system logs the counting data at the end of each session, which can be used for footfall analysis and other business insights.

EXPERIMENTAL SETUP

The system was deployed in a retail store environment to evaluate its performance. Cameras were positioned at entry and exit points, and video data was collected for a particular period. The system's performance was assessed based on accuracy, processing time, and scalability.

ACCURACY

Accuracy was measured by comparing the system's counts with manual counts obtained from reviewing video footage. The system achieved an accuracy of 95%, demonstrating its effectiveness in real-world scenarios.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, November 2024

PROCESSING TIME

The processing time per frame was measured to ensure the system operates in real time. On average, the system processed each frame in under 50 milliseconds, which is suitable for real-time applications.

SCALABILITY

Scalability was evaluated by testing the system in different environments with varying numbers of people and different camera setups. The system maintained high accuracy and performance across all tested scenarios.

V. RESULTS AND DISCUSSIONS

During testing, the system performed robustly across a range of video sources, including webcams and IP cameras, demonstrating reliable real-time performance with minimal latency, thanks to efficient threading implementation. The alert system effectively notified staff whenever the crowd count surpassed predefined thresholds, highlighting its practical utility in managing public spaces.

PERFORMANCE ANALYSIS

The decision to employ SSD with MobileNet architecture proved advantageous, offering a compelling blend of speed and accuracy in object detection. This setup enabled the system to handle live video streams adeptly while ensuring precise identification and counting of individuals. Additionally, the centroid tracker algorithm played a pivotal role in maintaining continuity by effectively tracking multiple objects, even in busy and dynamic environments.

LIMITATIONS

Despite its overall effectiveness, the system's performance is contingent upon the quality and consistency of the video feed. Variations in video resolution, lighting conditions, and camera positioning can impact the accuracy of object detection and tracking. Particularly in environments with substantial background movement or occlusions, such as crowded public areas, the system may encounter challenges in distinguishing between genuine individuals and false positives—instances where inanimate objects or transient obstructions are mistakenly identified.

FUTURE DIRECTIONS

Looking ahead, there are several avenues for enhancing the system's capabilities and versatility. Implementing advanced pre-processing techniques, such as sophisticated background subtraction algorithms and image enhancement methods, could mitigate the influence of environmental factors on detection accuracy. Moreover, further training deep learning models on comprehensive datasets encompassing diverse scenarios could bolster the system's ability to discern individuals in complex settings.

VI. CONCLUSION

This study presents a robust real-time people-counting system designed to address the challenges of accurate counting in dynamic environments. By leveraging advanced computer vision techniques, specifically the SSD with MobileNet architecture and centroid tracking, our system has demonstrated high accuracy and efficiency across various video sources, including webcams and IP cameras. These capabilities make it well-suited for a wide range of applications, from optimizing resource allocation in public spaces to enhancing customer service in retail environments.

The system's performance analysis underscores its ability to operate seamlessly in real-time with minimal latency, thanks to efficient threading implementation. This ensures prompt data processing and timely alerts when the number of people exceeds predefined thresholds. However, the system is not without limitations. Environments with variable video feed quality or significant background movement may pose challenges, potentially leading to false positives. Addressing these issues will be crucial for enhancing reliability and ensuring consistent performance in diverse settings.

In conclusion, while our developed system represents a significant leap forward in real-time people counting technology, ongoing research and development efforts are essential to unlock its full potential. By addressing current limitations and exploring new avenues for improvement, we aim to contribute towards more effective crowd management solutions and enhance overall societal benefits.

REFERENCES

[1] Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. arXiv preprint arXiv:1804.02767.

[2] Bewley, A., Ge, Z., Ott, L., Ramos, F., & Upcroft, B. (2016). Simple Online and real-time tracking. 2016 IEEE International Conference on Image Processing (ICIP), 3464-3468.

[3] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. Advances in Neural Information Processing Systems, 91-99.

[4] Zhang, L., Li, Y., & Nevatia, R. (2015). Global Data Association for Multi-Object Tracking Using Network Flows. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1-8.

[5] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770-778.

[6] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2961-2969.

[7] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real- Time Object Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 779-788.

[8] Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature Pyramid Networks for Object Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 936-944.

[9] Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv preprint arXiv:1409.1556.

[10] Dalal, N., & Triggs, B. (2005). Histograms of Oriented Gradients for Human Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 886-893.