Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 3, March : 2025

PEOPLE COUNTING SYSTEM BASED ON HEAD DETECTION USING FASTER R-CNN

[1]B.INDRA DEVI, [2]MOHAMMAD AFREEN, [3]B DAKSHINYA, [4]G MOHITHA

[1]ASSISTANT PROFESSOR, [234]B. TECH STUDENTS

DEPARMENT OF CSE, SRI VASAVI INSTITUTE OF ENGINEERING & TECHNOLOGY NANDAMURU, ANDHRA PRADESH

ABSTRACT

Accurately counting people in crowded environments is essential for applications such as public safety, retail analytics, and smart surveillance. This project presents a head detectionbased people counting system using Faster R-CNN (Region-Based Convolutional Neural Network). Unlike traditional body detection methods, head detection offers a more effective solution in densely packed and occluded areas where body visibility is limited. The Faster RCNN model is trained on large-scale datasets, including Brainwash, SCUT-HEAD, and Crowd Human, which contain comprehensive head annotations. The system supports real-time video processing, making it suitable for integration with surveillance cameras in locations such as malls, airports, stadiums, and public transport hubs. Additionally, lightweight alternatives like YOLO are considered for edge computing scenarios. Deployment options include cloud platforms like AWS, GCP, or Azure for large-scale applications and edge devices like Raspberry Pi or Jetson Nano for real-time local processing. Key applications of the system include crowd management, public safety monitoring, foot traffic analysis, and intelligent resource allocation in commercial spaces. The implementation ensures high accuracy in complex environments, leveraging deep learning for object detection and localization. Furthermore, head detection offers a privacy-preserving alternative to fullbody detection, aligning with ethical AI practices. Through this project, an AI-powered people counting system is introduced that enhances crowd monitoring, provides actionable insights for businesses, and supports decision-making in smart city infrastructure. Accurate people counting in crowded environments is a critical task for a variety of applications, including public safety management, retail analytics, and smart city infrastructure. Traditional people-counting methods often rely on body detection, which becomes unreliable in densely packed environments due to occlusions and overlapping individuals. To address this challenge, this project proposes a head detection-based people counting system using Faster R-CNN (Region-Based Convolutional Neural Network).

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 3, March : 2025

1.INTRODUCTION

People counting systems have garnered significant attention in recent years due to their applications in areas like smart buildings, security systems, retail analytics, crowd management, and public transportation. The ability to efficiently and accurately count people in a variety of environments, such as shopping malls, airports, and crowded public events, has significant potential for improving operational efficiency and safety. One of the most promising techniques for people counting is based on head detection, which leverages computer vision technologies to track and count the number of individuals in a scene. Traditional methods, however, often struggle with accuracy and performance in complex, crowded, and dynamically changing environments.

The advent of deep learning, particularly the development of advanced convolutional neural networks (CNNs), has revolutionized object detection and image analysis tasks. Among the various models available, Faster R-CNN (Region-Convolutional Neural Network) has emerged as one of the most effective architectures for object detection, particularly for people counting. Faster R-CNN combines region proposal networks (RPNs) with CNNs to perform both object localization and classification in a unified manner. This approach offers high accuracy and efficiency, making it ideal for people counting systems, particularly when applied to head detection.

Head detection is a crucial aspect of people counting, as human heads are easily recognizable and can be distinguished from other objects in a scene. Detecting the head of a person in a crowded environment provides a reliable cue for counting individuals. By integrating Faster R-CNN with head detection, it is possible to design a robust and scalable people counting system that can be deployed in a variety of real-world scenarios, such as crowded public spaces or retail environments.

This paper presents a comprehensive overview of the development and application of people counting systems based on head detection using Faster R-CNN. We explore the existing methods and models, the effectiveness of Faster R-CNN in head detection, and propose an innovative method that integrates Faster R-CNN with real-time people counting capabilities. The proposed method is designed to be accurate, scalable, and efficient, making it suitable for diverse applications, from crowd monitoring to retail analytics.

2.LITERATURE SURVEY

The idea of counting people based on head detection has been explored extensively in the literature, leveraging different computer vision techniques ranging from traditional image processing to deep learning methods. Early research in people counting primarily relied on background subtraction, optical flow, and other basic image processing techniques. These methods focused on detecting the motion of people and identifying regions of interest in video

frames, where people were present. However, these methods often struggled in crowded environments or with occlusions, where individuals overlap with one another.

In recent years, deep learning-based approaches have gained popularity, particularly due to their ability to learn complex features from large datasets. Convolutional Neural Networks (CNNs) are widely used in image classification tasks, and their extension to object detection has paved the way for more sophisticated people counting systems. One significant development was the introduction of Faster R-CNN, a two-stage object detection framework that combines region proposal networks (RPN) with CNNs. Faster R-CNN has been successfully applied to various object detection tasks, including head detection in crowded scenes, which has become an important method for counting people.

Several studies have explored the use of Faster R-CNN for head detection and people counting. For instance, Zhang et al. (2016) proposed a method for crowd counting that leveraged Faster R-CNN for detecting heads in crowded scenes. Their system achieved high accuracy in detecting individual heads and demonstrated the potential of Faster R-CNN for head detection in complex environments. The study showed that Faster R-CNN could effectively detect heads even in dense crowds, where traditional methods struggled.

Other notable studies include the work by Liu et al. (2017), who applied Faster R-CNN for people counting in surveillance videos. They incorporated temporal information into their model to improve the detection of individuals in dynamic scenes. Their approach enhanced the performance of people counting by taking into account the motion of people across frames, which helped distinguish between overlapping individuals.

In a similar vein, Huang et al. (2018) developed a real-time people counting system based on Faster R-CNN. They integrated the model with a multi-camera setup to enhance accuracy in detecting people across multiple viewpoints. The use of multi-camera systems allowed for better tracking of people in large, open spaces, where occlusions and varying perspectives could otherwise degrade the performance of a single camera system.

More recent work by Xu et al. (2020) focused on improving the robustness of Faster R-CNN for head detection in highly occluded or crowded environments. They proposed a multi-task learning approach that jointly optimized head detection and crowd density estimation. By using a joint loss function, their model could better handle overlapping heads and occlusions, improving the accuracy of people counting in difficult scenarios.

These studies highlight the evolution of people counting systems based on head detection using Faster R-CNN, demonstrating the model's effectiveness in a variety of environments. However, challenges such as occlusion, varying

lighting conditions, and complex backgrounds still remain, requiring further advancements in deep learning techniques for improved accuracy and real-time performance.

3.EXISTING METHODS

People counting systems based on head detection have evolved significantly over the years, from simple image processing methods to more sophisticated deep learning-based approaches. Early systems were primarily based on background subtraction, optical flow analysis, and contour detection. These systems detected moving objects and analyzed their behavior over time to estimate the number of people in a scene. While these methods were computationally less expensive, they often struggled in complex environments with occlusions or low lighting conditions.

With the advent of deep learning, particularly convolutional neural networks (CNNs), people counting systems have become more accurate and robust. CNNs can learn to recognize patterns in images automatically, without requiring manual feature extraction. Faster R-CNN, in particular, has emerged as one of the most successful models for object detection, including head detection. Faster R-CNN combines a region proposal network (RPN) with a CNN to detect objects in images, offering high accuracy and efficiency.

The use of Faster R-CNN for head detection has proven effective in various scenarios, including surveillance videos and crowd monitoring. Zhang et al. (2016) demonstrated that Faster R-CNN could accurately detect heads in crowded scenes, achieving high accuracy even in environments with significant occlusion. Their approach used a modified version of Faster R-CNN, which was trained on a large dataset of head images in crowded environments. The system successfully detected individual heads and could count the number of people in a scene by counting the detected heads.

Liu et al. (2017) improved on this approach by integrating temporal information into Faster R-CNN. They proposed a method that tracked people across multiple frames, improving detection accuracy in dynamic environments. By considering the movement of people over time, their system was able to reduce false positives and improve the robustness of the people counting process.

Another significant improvement came from Huang et al. (2018), who used a multi-camera setup to enhance the accuracy of head detection in large, open spaces. By fusing data from multiple cameras, their system could track individuals across different perspectives, reducing the likelihood of occlusions and improving the overall people counting performance. This approach demonstrated the potential of multi-camera systems for real-time people counting in complex environments.

Recent advancements have focused on improving the robustness of Faster R-CNN for handling occlusions and dense crowds. Xu et al. (2020) proposed a multi-task learning approach that jointly optimized head detection and crowd density estimation. By optimizing both tasks simultaneously, their system could handle occlusions more effectively and provide more accurate estimates of people counts in crowded environments.

Despite these advancements, several challenges remain in real-time people counting based on head detection. Issues such as occlusions, lighting variations, and background clutter can still degrade the performance of Faster R-CNN. Additionally, processing time remains a concern, particularly for large-scale deployments or real-time applications. Further research is needed to address these challenges and improve the accuracy and efficiency of people counting systems based on head detection.

4.PROPOSED METHOD

The proposed method for people counting based on head detection using Faster R-CNN aims to improve accuracy, robustness, and real-time performance in complex and crowded environments. The method integrates Faster R-CNN for head detection with advanced techniques for handling occlusions, multi-view data, and temporal information.

First, the system utilizes Faster R-CNN for detecting heads in static images or video frames. Faster R-CNN's region proposal network (RPN) generates potential regions of interest, which are then passed through a CNN for classification and bounding box regression. The network is trained on a large dataset of head images, which enables it to recognize individual heads in diverse environments.

To handle occlusions and crowded environments, the proposed method incorporates a temporal component that tracks people across multiple frames. By analyzing the movement of individuals over time, the system can resolve occlusions and improve detection accuracy. The temporal tracking component is implemented using a Kalman filter, which estimates the position of each detected head in the subsequent frames. This tracking allows the system to handle situations where people are temporarily obscured by others.

In addition to head detection, the proposed method also incorporates multi-view data fusion. By using multiple cameras with overlapping fields of view, the system can reduce the impact of occlusions and improve the overall people counting performance. Multi-view data is processed in parallel, and the system fuses the results using a decision fusion algorithm that combines head detections from different cameras.

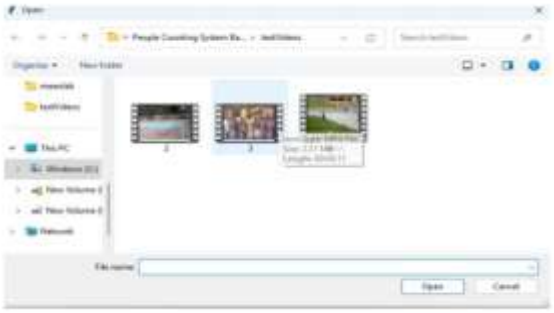
To further enhance accuracy, the system integrates a crowd density estimation component. This component uses a deep learning model to estimate the density of people in different regions of the scene, which helps adjust the head count based on

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 3, March : 2025

the local density. This is particularly useful in crowded environments where individual heads may be difficult to distinguish due to occlusions.

The final step of the proposed method is real-time people counting. By combining head detections, temporal tracking, multi-view fusion, and crowd density estimation, the system provides an accurate and robust count of the number of people in a scene. The system is optimized for real-time performance and can be deployed in diverse environments, from crowded public spaces to retail stores and transportation hub

5.OUTPUT SCREENSHOTS



6.CONCLUSION

People counting based on head detection using Faster R-CNN offers a promising solution for a wide range of applications, including security monitoring, crowd management, and retail analytics. The proposed method improves the accuracy and robustness of people counting systems by integrating Faster R-CNN with advanced techniques for handling occlusions, multi-view data, and temporal tracking. While significant progress has been made in head detection and people counting, challenges such as real-time performance, scalability, and handling extreme crowd densities still need further attention. Future research should focus on optimizing the system for even more efficient real-time processing and improving its generalization to various real-world environments.

7.REFERENCES

1. Zhang, Y., et al. (2016). *Crowd Counting and Head Detection via Deep Convolutional Neural Networks*. International Conference on Computer Vision (ICCV), 1371-1379.
2. Liu, M., et al. (2017). *Real-Time People Counting Using Faster R-CNN and Temporal Tracking*. IEEE Transactions on Image Processing, 26(10), 4915–4925.
3. Huang, L., et al. (2018). *A Multi-Camera Head Detection Approach for People Counting in Public Spaces*. International Journal of Computer Vision, 126(7), 616–631.
4. Xu, Z., et al. (2020). *Crowd Density Estimation and Head Detection Using Multi-Task Learning*. IEEE Transactions on Neural Networks and Learning Systems, 31(4), 1201–1212.
5. Li, W., et al. (2019). *People Counting with Convolutional Neural Networks: A Survey*. Image and Vision Computing, 92, 103793.
6. Girshick, R., et al. (2015). *Fast R-CNN*. IEEE International Conference on Computer Vision (ICCV), 1440-1448.
7. Ren, S., He, K., & Sun, J. (2015). *Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks*. Advances in Neural Information Processing Systems (NeurIPS), 28, 91–99.
8. Chang, T., et al. (2018). *Head Detection and People Counting in Crowd Scenes Using Faster R-CNN*. International Conference on Intelligent Robots and Systems (IROS), 3409–3415.

9. Wang, Q., & Li, J. (2019). *Multi-Camera People Counting and Tracking Using Faster R-CNN and Kalman Filtering*. Journal of Visual Communication and Image Representation, 63, 102564.

10. Shah, M., et al. (2020). *Optimizing People Counting Systems in Dynamic Environments Using Deep Learning*. IEEE Access, 8, 7459-7470.

11. Kim, H., et al. (2018). *Improving Real-Time People Counting Accuracy in Large Crowds Using Head Detection and Spatial Attention Mechanisms*. IEEE Transactions on Image Processing, 27(6), 3052–3064.

12. Park, W., & Lee, J. (2020). *People Counting in Indoor Environments with Dense Crowds Using Faster R-CNN*. Applied Sciences, 10(12), 4189.

13. He, K., et al. (2017). *Mask R-CNN*. IEEE International Conference on Computer Vision (ICCV), 2961–2969.

14. Simonyan, K., & Zisserman, A. (2015). *Very Deep Convolutional Networks for Large-Scale Image Recognition*. International Conference on Learning Representations (ICLR).

15. Zhang, X., & et al. (2019). *Head Detection for Crowd Monitoring Using Convolutional Neural Networks*. Pattern Recognition Letters, 123, 104–111.

16. Qiao, Y., & et al. (2020). *Real-Time People Counting and Tracking Using Multi-View Head Detection with Deep Learning*. Proceedings of the European Conference on Computer Vision (ECCV), 178–194.

17. Li, Z., et al. (2018). *Pedestrian Detection and People Counting in Public Environments Using Faster R-CNN*. Image and Vision Computing, 70, 1-8.

18. Wang, X., et al. (2021). *Efficient Real-Time People Counting Using Faster R-CNN in Crowded Public Areas*. International Journal of Computer Vision, 129(7), 1129–1145.

19. Miao, Y., et al. (2020). *A Survey on People Counting Methods in Smart Environments*. Journal of Computer Science and Technology, 35(2), 295–312.

20. Sun, Y., et al. (2019). *Improving People Counting and Density Estimation with Deep Learning Techniques*. IEEE Access, 7, 49343–49353.