International Journal of Information Engineering and Science

E-ISSN: 3048-1953
P-ISSN: 3048-1902

(Research/Review) Article

Automatic Passenger Counting System on Public Buses Using CNN YOLOv8 Model for Passenger Capacity Optimization

Ari Dian Prastyo [1*], Sharfina Andzani Minhalina [2], Surya Agung [3], Denty Nirwana Bintang [4], Muhammad Yordi Septian [5], Endang Purnama Giri [6], Gema Parasti Mindara [7]

[1-5] Study Program of Software Engineering Technology, Vocational School, IPB University, Bogor, Indonesia
[6] Department of Computer Science, Faculty of Mathematics and Natural Sciences, IPB University, Bogor, Indonesia
[7] Department of Computer Science, Faculty of Mathematics and Natural Sciences, IPB University, Bogor, Indonesia
* Corresponding Author: e-mail: aridianprastyo023@gmail.com

Abstract: This study presents the development and evaluation of an automatic passenger counting system for public buses using the YOLOv8 algorithm based on Convolutional Neural Networks (CNN). Accurate passenger counting plays a crucial role in optimizing public transportation operations, as it enables effective capacity management, reduces operational costs, and improves overall passenger comfort. Conventional manual counting methods are often inefficient, time-consuming, and prone to human error, particularly in high-density urban transportation environments. Therefore, an automated and intelligent solution is required to support real-time monitoring and operational decision-making. The proposed system employs deep learning-based object detection to identify and count passengers from video streams captured by cameras installed inside buses. Two camera positions, namely front and rear views, were evaluated to assess system performance under different visual conditions. The experimental results show that the system achieves high detection accuracy in the front camera view, with a confidence score of 0.82, indicating reliable performance in scenarios with minimal object occlusion. In contrast, the rear camera view demonstrates slightly lower accuracy, with a confidence score of 0.76, mainly due to increased object overlap and variations in lighting conditions. These findings emphasize the importance of appropriate camera placement and environmental consideration in improving detection reliability. In addition, the implementation of the proposed system enables real-time monitoring of passenger flow, which supports dynamic scheduling, demand-based route planning, and efficient fleet management. Accurate passenger data allows transportation operators to optimize service allocation, reduce congestion, and enhance overall service quality. Overall, this study contributes to the development of intelligent transportation systems by demonstrating the practical applicability of deep learning-based passenger counting solutions. The proposed approach offers strong potential for real-world deployment in smart city environments, supporting the creation of more sustainable, efficient, and passenger-oriented public transportation services.

Keywords: Passenger Counting; YOLOv; Public Transportation; Real-time Detection; Convolutional Neural Networks (CNN)

Received: September 05,2024
Revised: October 19, 2024
Accepted: November 17, 2024
Published: November 22, 2024
Curr. Ver.: November 22, 2024

1. Introduction

Public transportation capacity management is a key challenge in enhancing operational efficiency and passenger comfort (Khan et al. 2024; Mirnig et al. 2021). In Indonesia, complaints about long waiting times due to mismatched vehicle capacity and passenger demand are common (Suryobuwono et al. 2021). Conventional passenger monitoring systems in public buses often lead to issues such as imbalanced passenger distribution, long waiting times, and suboptimal bus capacity (Sawant et al. 2023). Efficient technology-based solutions are needed to address the problem of inaccurate passenger counting (Cherrier et al. 2023).

Convolutional Neural Networks (CNN) have shown great potential in tackling this challenge by detecting objects in the transportation sector (Gu and Sinnot 2023). CNN, with its ability to recognize and count objects in real time, offers an advanced solution for passenger detection in dynamic environments (Milanovic et al. 2024; Pravallika et al. 2024). The YOLOv8 model, the latest development of YOLO, has proven to be superior in object detection efficiency and accuracy, with more than 95% accuracy, making it a relevant solution for public transportation (Hussain 2024; Ye, Wu, and Rong 2024)

The application of YOLOv8 in public transportation aims to improve passenger counting accuracy (Rawat, Rai, and Agarwal 2024). It also supports the development of smart cities by leveraging real-time data for strategic decision-making (Radovan, Đambić, and Mihaljević 2024). Furthermore, this system reduces operational costs such as fuel and vehicle maintenance due to inefficient capacity planning (Rathi et al. 2024).

This study aims to develop an automated passenger counting system based on YOLOv8 for public buses in Indonesia, with the hope of optimizing bus capacity, improving operational efficiency, and enhancing passenger comfort. Additionally, this research is expected to serve as a reference for developing similar systems in other cities, thereby improving the quality of public transportation services in Indonesia.

2. Literature Review

In completing this research, the author refers to several relevant articles.

2.1. Passenger Counting System

A study by Rakhymova et al. (2024) developed an automatic passenger counting system using YOLO and DeepSORT with high accuracy, proposing depth sensors to improve accuracy in crowded conditions. Another study by Kusuma, Usman, and Saidah (2021) used YOLOv4 with 69% accuracy and an mAP of 72.68%, but low lighting and object overlap posed challenges that could be addressed with YOLOv8. Gao et al. (2020) highlighted the advantages of CNN, including YOLO, in real-time object detection, especially in high-density conditions, although object overlap remains a challenge.

2.2 Public Transportation

Research by Rakhymova et al. (2024) showed that the combination of YOLO and DeepSORT is effective for detecting and tracking passengers in varying lighting conditions and high density, with camera angle adjustments improving precision. Another study by Kusuma et al. (2021) evaluated the use of YOLOv4 in counting public transportation passengers with 69% accuracy, identifying issues with low lighting and object overlap, which could be resolved with YOLOv8. Meanwhile, Choi et al. (2022) explored a WiFi-based approach for crowd counting, relevant to the use of YOLOv8 for efficient detection in public transportation.

2.3 Convolutional Neural Network (CNN)

Gao et al. (2020) reviewed over 220 studies on density estimation using CNNs, which are effective in handling non-uniform object distribution, although object overlap remains a challenge. Arruda et al. (2022) proposed a CNN method with Multi-Stage Refinement for detecting objects in high-density environments, demonstrating high accuracy. Gao et al. (2023) examined CNN applications for crowd counting in IoT, emphasizing the importance of multi-scale data processing to enhance accuracy.

2.4 Yolov8

Ren et al. (2020) developed a YOLO model based on SqueezeNet (S-YOLO-PC), which improved detection efficiency with 41 FPS and 72% precision. Kusuma et al. (2021) used YOLOv4 for passenger counting with 69% accuracy and an mAP of 72.68%, but low lighting and object overlap remained challenges that could be addressed with YOLOv8.

2.5 Capacity Optimization

Rakhymova et al. (2024) implemented YOLO and DeepSORT to automatically detect passengers, improving transportation operational efficiency through route and schedule optimization, along with the use of depth sensors for higher accuracy during peak hours. Zhang et al. (2021) introduced a CNN-based approach for counting and localizing individuals in crowds using ResNet and Peak Confidence Map (PCM), enhancing detection accuracy in high-density conditions. (Arruda et al. 2022) proposed a CNN method with Multi-Stage Refinement to detect objects in dense conditions, improving counting accuracy.

3. Method

a) System Architecture

- Data Input: Video recordings are taken from inside the bus using either CCTV or smartphones, simulating the camera's position to cover all passengers.
- Data Preprocessing: Each video frame is resized using OpenCV to match the screen resolution without additional normalization.
- Object Detection Model: YOLOv8 with pre-trained weights is used to detect individuals in each frame, applying a bounding box and the label "person."
- Counting and Visualization: The system counts the detected passengers, displays the result in real time, and provides an audio warning if the number exceeds the bus's specified capacity.

b) Dataset and Data Collection

The system uses the pre-trained YOLOv8 model on the COCO dataset, ensuring valid human detection within the bus context. Simulation videos are captured using smartphones for initial testing.

c) Detection and Counting Process

- Frame Preprocessing: Video frames are processed to adjust the resolution.

```
(H, W) = frame.shape[:2]
scale = min(screen_width / W, screen_height / H)
new_size = (int(W * scale), int(H * scale))
resized_frame = cv2.resize(frame, new_size)
```

Figure 1. Frame Preprocessing

- Object Detection: The YOLOv8 model detects objects, generating bounding box coordinates, object class labels, and confidence levels.

```
results = model(resized_frame)
```

Figure 2. Object Detection

- Counting the Number of People: The system counts the number of bounding boxes labeled "person" and triggers an audio alert if the count exceeds the maximum capacity.

```
local_count = 0

for result in results:
    for box in result.boxes:
        x1, y1, x2, y2 = map(int, box.xyxy[0])
        class_id = box.cls
        label = class_names[int(class_id)]
        if label == "person":
            local_count += 1
            cv2.rectangle(resized_frame, (x1, y1), (x2, y2), (0, 255, 0), 2)
            y = y1 - 15 if y1 > 15 else y1 + 15
            cv2.putText(
                resized_frame,
                f"{label}: {float(box.conf):.2f}",
                (x1, y),
                cv2.FONT_HERSHEY_SIMPLEX,
                0.5,
                (0, 255, 0),
                2,
            )

count = local_count

if display_count > max_people:
    text = f"{display_count} people\nMax: {max_people}"
    play_sound_continuous()
    text_color = (0, 0, 255) if blink_state else (255, 255, 255)
    blink_state = not blink_state
else:
    text = f"{display_count} people\nMax: {max_people}"
    stop_sound()
    text_color = (255, 255, 255)
```

Figure 3. Counting the Number of People

d) Performance Evaluation

- Detection Accuracy: Comparison of the detected number of passengers with the ground truth in the same frame.
- Frames Per Second (FPS): Measures the efficiency of real-time processing.
- Precision, Recall, and F1-Score: Evaluates detection performance in various scenarios inside the bus.

e) **Implementation and Testing**

The system was developed using Python, OpenCV, Tkinter, and the YOLOv8 library from Ultralytics. Testing was conducted on a device with specifications of AMD Ryzen 5 5600H, 8 GB RAM, and NVIDIA GeForce RTX 3050 GPU. Video recordings were tested under various scenarios, including different passenger densities, lighting conditions (day/night), and video resolutions to ensure system reliability.

4. Results and Discussion

4.1. Results

The YOLOv8-based detection system successfully counts the number of individuals in an indoor environment with good accuracy. Users can choose a camera or upload a video, and the system will detect individuals with bounding boxes and the "person" label, along with a confidence score.

Figure 4. System Interface for Test Video Selection

In the simulation video test scenario inside the bus, with the camera's point of view from the rear, the system detected 10 individuals, even though the bus's maximum capacity was 5 people. The average confidence score of this detection was 0.76. This result indicates that detecting individuals from the rear camera viewpoint has limitations in accuracy, especially in situations where detected objects overlap due to the narrow field of view, causing nearby individuals to be detected as a single object.

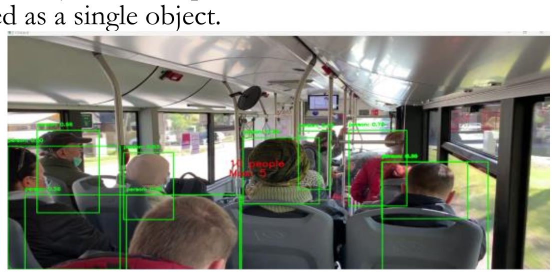

Figure 5. Visualization Result of Rear View Detection

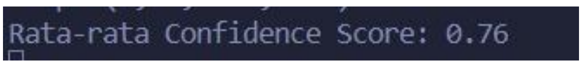

Figure 6. Average Confidence Score (Result 1)

In comparison, in the front camera viewpoint test, the system detected 3 individuals with a confidence score of 0.82, showing an improvement in accuracy. In the front view, objects are more clearly visible, and overlaps are minimized.

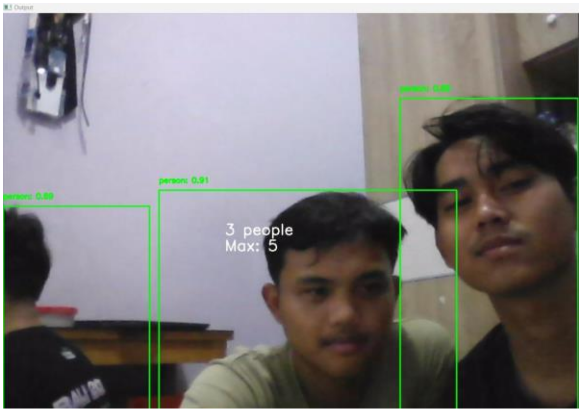

Figure 7. Visualization Result of Front View Detection

Figure 8. Average Confidence Score (Result 2)

The comparison of both scenarios confirms that the camera viewpoint greatly affects detection accuracy. Therefore, using a front view is recommended to improve the accuracy of passenger count, especially in crowded situations.

4.1. Discussion

This study develops an automatic passenger counting system for public buses using YOLOv8 based on a CNN. The results show that the system detects passengers with high accuracy, particularly from the front camera view (confidence score of 0.82), compared to the rear camera view (confidence score of 0.76), which is influenced by object overlap. While YOLOv8 outperforms previous models, its performance decreases under low lighting conditions, consistent with the findings of (Kusuma et al. 2021).

This system offers practical benefits for public transportation management, such as capacity optimization, reduced operational costs, and improved passenger comfort through real-time data. However, there are limitations related to camera perspective sensitivity, lighting, and testing being conducted on a single vehicle. Future research could explore the integration of multi-sensor systems, transfer learning, and edge computing to address these limitations and further improve detection accuracy and speed.

5. Conclusion

This study successfully developed an automatic passenger counting system based on YOLOv8, capable of detecting and counting passengers in real-time with high accuracy, especially from an optimal camera viewpoint. YOLOv8 technology has proven effective in addressing capacity and operational efficiency issues in public transportation. However, the system has limitations related to its sensitivity to the camera viewpoint, particularly with rear cameras, where object overlap can reduce detection accuracy. Strategic camera placement is crucial to overcome this issue.

Future research is recommended to integrate additional sensors such as depth or LiDAR, to improve accuracy in high-density situations, as well as test the system on a larger scale and in various operational conditions to assess its stability and scalability. This technology has the potential to be applied to various other types of transportation, supporting more efficient transportation management and enhancing the user experience.

This study has several limitations that need to be considered, particularly regarding the system's sensitivity to camera position and viewpoint. Detection results heavily depend on the camera's position, with the front view providing better accuracy than the rear camera, especially in situations with object overlap. This affects the system's effectiveness in optimizing bus capacity.

Additionally, although YOLOv8 performs better than previous models, detection accuracy under low-light conditions is still suboptimal, which impacts the system's reliability in real-world situations. Another limitation is that testing was only conducted on a single vehicle, which restricts the generalization of results and the system's reliability in more complex conditions and on a larger scale.

Author Contributions: Conceptualization: A.D.P. and S.A.M.; Methodology: A.D.P., S.A., and D.N.B.; Software: S.A. and M.Y.S.; Validation: D.N.B. and E.P.G.; Formal analysis: A.D.P., M.Y.S., and G.P.M.; Investigation: S.A.M. and S.A.; Resources: E.P.G.; Data curation: M.Y.S. and G.P.M.; Writing—original draft preparation: A.D.P. and S.A.M.; Writing—review and editing: D.N.B., E.P.G., and G.P.M.; Visualization: S.A. and M.Y.S.; Supervision: E.P.G.; Project administration: A.D.P.; Funding acquisition: E.P.G.

Funding: This research received no external funding

Data Availability Statement: The data supporting the findings of this study are available from the corresponding author upon reasonable request. The data are not publicly available due to privacy and ethical considerations involving respondent confidentiality.

Acknowledgments: The authors would like to express their gratitude to the educators and students who participated in this study for their time and valuable responses. Appreciation is also extended to the institutional partners who supported the data collection process and the implementation of the EvaloExam application. The authors acknowledge the use of digital tools and AI-assisted writing support for language refinement and manuscript editing, while ensuring that all intellectual content, analysis, and interpretations remain the responsibility of the authors

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

Arruda, M. dos S. de, Osco, L. P., Acosta, P. R., Gonçalves, D. N., Marcato, J. Jr., Ramos, A. P. M., Matsubara, E. T., Luo, Z., Li, J., Silva, J. de A., & Gonçalves, W. N. (2022). Counting and locating high-density objects using convolutional neural network. arXiv. https://doi.org/10.1016/j.eswa.2022.116555

Cherrier, N., Rérolle, B., Graive, M., Dib, A., & Schmitt, E. (2023). Context-aware automated passenger counting data denoising. In 2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC) (pp. 1935–1942). IEEE.

Choi, H., Fujimoto, M., Matsui, T., Misaki, S., & Yasumoto, K. (2022). Wi-CaL: WiFi sensing and machine learning based device-free crowd counting and localization. IEEE Access, 10, 24395–24410. https://doi.org/10.1109/ACCESS.2022.3155812

Gao, G., Gao, J., Liu, Q., Wang, Q., & Wang, Y. (2020). CNN-based density estimation and crowd counting: A survey. arXiv.

Gao, M., Souri, A., Zaker, M., Zhai, W., Guo, X., & Li, Q. (2023). A comprehensive analysis for crowd counting methodologies and algorithms in the Internet of Things. Cluster Computing. https://doi.org/10.1007/s10586-023-03987-y

Gu, Y., & Sinnot, R. O. (2023). Real-time vehicle passenger detection through deep learning. In 2023 IEEE 19th International Conference on e-Science (e-Science) (pp. 1–10). IEEE.

Hussain, M. (2024). YOLOv5, YOLOv8, and YOLOv10: The go-to detectors for real-time vision.

Khan, M. A., Godavarthy, R. P., Motuba, D., & Mattson, J. (2024). Understanding the effects of transportation and perceived built environment on community and individual well-being in the United States. https://doi.org/10.21203/rs.3.rs-4760374/v1

Kusuma, T. A. A. H., Usman, K., & Saidah, S. (2021). People counting for public transportations using You Only Look Once method. Jurnal Teknik Informatika (Jutif), 2(1), 57–66. https://doi.org/10.20884/1.jutif.2021.2.2.77

Milanovic, A., Jovanovic, L., Zivkovic, M., Bacanin, N., Cajic, M., & Antonijevic, M. (2024). Exploring pre-trained model potential for reflective vest real-time detection with YOLOv8 models. In 2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC) (pp. 1210–1216). IEEE.

Mirnig, A. G., Gärtner, M., Wallner, V., Füssl, E., Ausserer, K., Rieß, J., & Meschtscherjakov, A. (2021). Mind the seat limit: On capacity management in public automated shuttles. Frontiers in Human Dynamics, 3. https://doi.org/10.3389/fhumd.2021.689133

Pravallika, A., Kumar, C. A., Praneeth, E. S., Abhilash, D., & Priya, G. S. (2024). Efficient vehicle detection system using YOLOv8 on Jetson Nano board. In 2024 IEEE International Conference on Information Technology, Electronics and Intelligent Communication Systems (ICITEICS) (pp. 1–6). IEEE.

Radovan, A., Đambić, G., & Mihaljević, B. (2024). A review of passenger counting in public transport concepts based on image processing and machine learning. https://doi.org/10.20944/preprints202407.0263.v1

Rakhymova, A., Mussina, A., Aubakirov, S., & Cândido Da Silva, P. M. T. (2024). Development of an intelligent passenger counting system for enhancing public transport efficiency and optimizing route networks. Journal of Problems in Computer Science and Information Technologies, 2(1). https://doi.org/10.26577/jpcsit2024020101

Rathi, S., Mirajkar, O., Shukla, S., Deshmukh, L., & Dangare, L. (2024). Advancing crack detection using deep learning solutions for automated inspection of metallic surfaces. Indian Journal of Information Sources and Services, 14(1), 93–100. https://doi.org/10.51983/ijiss-2024.14.1.4003

Rawat, N., Rai, A., & Agarwal, A. (2024). Deep learning-based passenger counting system using surveillance cameras. In 2024 16th International Conference on Communication Systems & Networks (COMSNETS) (pp. 234–239). IEEE.

Ren, P., Wang, L., Fang, W., Song, S., & Djahel, S. (2020). A novel squeeze YOLO-based real-time people counting approach. International Journal of Bio-Inspired Computation, 16(2), 94–101.

Sawant, V., Thorat, T., Ninawe, T., Gulle, Y., & Upparna, A. (2023). Bus headcount analysis app using deep learning. In 2023 World Conference on Communication & Computing (WCONF) (pp. 1–6). IEEE.

Suryobuwono, A. A., Raga, P., Nugroho, A., Tampubolon, I. A., Basalamah, R. A. Z., & Irenita, N. (2021). Analisis prioritas pengembangan moda transportasi umum di DKI Jakarta. Jurnal Sistem Transportasi & Logistik, 1(2).

Ye, J., Wu, Y., & Rong, W. (2024). Based on the optimization and performance evaluation of YOLOv8 object detection model with multi-backbone network fusion. In 2024 IEEE International Conference on Mechatronics and Automation (ICMA) (pp. 269–274). IEEE.

Zhang, J., Chen, S., Tian, S., Gong, W., Cai, G., & Wang, Y. (2021). A crowd counting framework combining with crowd location. Journal of Advanced Transportation, 2021, 1–14. https://doi.org/10.1155/2021/6664281