Signal, Image and Video Processing
https://doi.org/10.1007/s11760-024-03300-3

ORIGINAL PAPER

A real-time traffic sign detection in intelligent transportation system using YOLOv8-based deep learning approach

Mingdeng Tang[1]

Received: 28 January 2024 / Revised: 7 May 2024 / Accepted: 16 May 2024

Abstract
Intelligent transportation systems rely heavily on accurate traffic sign detection (TSD) to enhance road safety and traffic management. Various methods have been explored in the literature for this purpose, with deep learning methods consistently demonstrating superior accuracy. However, existing research highlights the persistent challenge of achieving high accuracy rates while maintaining non-destructive and real-time requirements. In this study, we propose a deep learning model based on the YOLOv8 architecture to address this challenge. The model is trained and evaluated using a custom dataset, and extensive experiments and performance analysis demonstrate its ability to achieve precise results, thus offering a promising solution to the current research challenge in deep learning-based TSD.

Keywords Traffic sign detection · Deep learning · YOLOv8 model · Real-time · Intelligent transportation

1 Introduction

The advent of smart cities has ushered in a new era of urban development, where the integration of advanced technologies plays a pivotal role in enhancing the quality of life for citizens [1, 2]. One of the critical aspects of this transformation is the deployment of video-based traffic surveillance systems [3]. These systems utilize state-of-the-art cameras and computer vision techniques to monitor and manage traffic flow, ensuring a safer and more efficient urban environment. In this case, traffic sign detection (TSD) emerges as a key component in ensuring the smooth operation of these surveillance systems [4].

Traffic sign detection within video-based traffic surveillance systems [5] is of paramount importance. These signs serve as the language of the road, communicating vital information to drivers, such as speed limits, warnings, and directions. Accurate and real-time detection of these signs is essential to assist drivers, enforce traffic rules, and improve overall road safety. In the context of modern smart cities, the demand for precise and efficient traffic sign detection has never been higher [4, 6].

The field of traffic sign detection has witnessed significant advancements in recent years. Traditional computer vision techniques have paved the way for more sophisticated and robust methods. These include the incorporation of deep learning, which has demonstrated remarkable performance improvements. This paper aims to review the latest advances in traffic sign detection methods and delve into the reasons behind the increasing adoption of deep learning-based approaches [7].

Deep learning-based approaches have garnered substantial attention from researchers due to their ability to handle complex tasks. Deep neural networks have shown promising results in various computer vision applications, making them a natural choice for traffic sign detection. Their adaptability and capacity to learn intricate patterns have encouraged further exploration in this domain [8–11].

While deep learning methods offer remarkable potential, they also pose unique challenges, especially when it comes to achieving high accuracy and real-time performance. These challenges are crucial in the case of traffic sign detection, where misinterpretations or delays can have serious consequences. Addressing these research challenges is essential to ensure that deep learning-based approaches can meet the stringent requirements of modern video-based traffic surveillance [12–14].

✉ Mingdeng Tang
tang18680906057@126.com

1 Department of Network and Information Security, Chongqing Vocational Institute of Safety & Technical, Wanzhou, Chongqing 404120, China

Published online: 28 May 2024

In this study, we propose a deep learning approach using Convolutional Neural Networks (CNNs) to address the challenges of traffic sign detection. We will demonstrate that adopting a deep learning approach can effectively overcome the identified research challenges by leveraging the model's capacity to capture intricate features and patterns in traffic signs. Our approach is supported by a custom dataset and a comprehensive training, validation, and testing process, ensuring robustness and accuracy.

This research contributes to the field of traffic sign detection in several ways. First, we generate a custom dataset tailored to the challenges of traffic sign detection, aiding in the development and evaluation of future methods. Second, we propose an efficient deep-learning method that enhances the accuracy and real-time capabilities of traffic sign detection. Third, extensive experiments and performance evaluations are conducted to validate the effectiveness of our method, providing insights and benchmarks for further research and applications in the domain of smart city traffic surveillance.

The remainder of the paper is organized as follows. Related Work reviews state-of-the-art techniques in traffic sign detection, noting the advances and limitations of various methods, including enhancements to YOLOv5 for real-time, multi-scale detection. The Methodology section details the data collection process, emphasizing the use of a diverse dataset and data augmentation techniques essential for training the YOLOv8 model. It also explains the model training and evaluation procedures. In Results and Discussion, the paper presents experimental results and performance evaluations, discussing key metrics such as precision, recall, and mean Average Precision (mAP), and includes visual aids like confusion matrices and precision-recall curves. The paper concludes with a Conclusion and Future Works summarizing the findings and contributions of the research, discussing limitations, and suggesting future areas for improvement, such as enhancing model robustness and computational efficiency.

2 Related work

This section provides an overview of the state-of-the-art techniques utilized in the field of traffic sign detection for road safety and traffic management.

This paper [2] introduced an enhanced YOLOv5 network for real-time multi-scale traffic sign detection. The method employs deep learning techniques to achieve highly accurate and efficient detection of traffic signs across varying scales. However, there's a limitation regarding its adaptability to environmental factors like adverse weather or lighting conditions, which can affect accuracy. Despite this limitation, the proposed method presents a significant advancement in the field of intelligent transportation, offering a practical and effective solution for real-time traffic sign detection in diverse urban settings.

The authors in [15] presented two innovative models for traffic sign detection utilizing YOLOv5s. The method leverages YOLOv5s architecture to detect traffic signs in real-time efficiently, demonstrating high accuracy. However, it faces a limitation concerning the detection of damaged or obscured signs, impacting its robustness in challenging conditions. Nevertheless, these novel models contribute to enhancing the field of intelligent transportation by providing effective solutions for real-time traffic sign detection.

The paper [16] introduced Attention-YOLOV4, a real-time and highly accurate traffic sign detection algorithm. The method incorporates attention mechanisms within the YOLOV4 framework to enhance detection precision. However, a potential limitation could be its computational complexity, which may require powerful hardware for real-time applications. Nonetheless, Attention-YOLOV4 significantly advances traffic sign detection, offering a compelling solution for improved accuracy in intelligent transportation systems.

The authors in [17] focused on Indian traffic sign detection and recognition using deep learning techniques. The method employs deep neural networks to achieve accurate detection and classification of Indian traffic signs. However, a limitation may arise from the diversity of traffic signs and varying road conditions, potentially affecting recognition accuracy. Nonetheless, this research contributes to enhancing road safety and traffic management in India by providing a foundation for intelligent transportation systems to effectively identify and interpret Indian traffic signs, even with the challenge of diversity and variability.

The paper [18] introduced MTSDet, a multi-scale traffic sign detection method that incorporates attention and path aggregation techniques. This innovative approach enhances the detection of traffic signs across varying scales and complex urban environments. A potential limitation of MTSDet may be the computational resources required for real-time deployment, which can pose practical challenges. Nevertheless, MTSDet represents a significant advancement in intelligent transportation systems, offering a comprehensive solution for multi-scale traffic sign detection with the potential to contribute to enhanced road safety and traffic management.

The five papers explore various aspects of traffic sign detection, with a common emphasis on real-time and accurate detection. The [2] and [15] both build on the YOLOv5 architecture, with the former proposing improvements and the latter introducing novel models. They share a foundation in YOLOv5 but differ in the specific enhancements, potentially aimed at better real-time multi-scale detection. The [16] paper introduces attention mechanisms within the YOLOv4 framework to achieve high-accuracy, real-time traffic sign detection. On the other hand, [17] focuses on a specific geographical context (India), using deep learning methods for traffic sign detection. Lastly, [18] stands out for its unique emphasis on path aggregation and multi-scale detection techniques, promising improved precision.

In summary, these papers demonstrate the versatility of deep learning-based approaches in traffic sign detection. The first two papers focus on enhancing YOLOv5 for real-time multi-scale detection, contributing variations within a common framework. The [16] and [18] both incorporate attention mechanisms, but the latter introduces path aggregation as well, promising higher precision. The [17] takes a regional perspective, which is both a strength and limitation, given its context-specific focus. Each paper provides valuable insights into the challenges and solutions related to traffic sign detection within the realm of intelligent transportation systems.

3 Methodology

3.1 Data collection

In our research, we curated a diverse dataset for traffic sign detection by sourcing images from internet resources and leveraging the capabilities of Roboflow's extensive image database. This dataset compilation process involved collecting images of traffic signs captured under various conditions, such as different weather, lighting, and urban environments. The dataset comprises 6279 samples, collected after performing data augmentation to address the challenge of dataset diversity. The DA is employed to enhance the variety and robustness of a dataset by applying transformations such as rotation, translation, scaling, or flipping to the original samples. This process helps in better training machine learning models to recognize patterns and generalize well to unseen data. In this dataset, the classes encompass a total of 21 various traffic signs, signals, and road markings commonly encountered in urban and suburban environments. They are essential for training models to accurately detect and classify objects in images or video frames, contributing to the development of robust and reliable computer vision systems for traffic management, autonomous driving, and pedestrian safety applications.

The dataset used for traffic sign detection comprises 6,279 images, categorized into 21 distinct classes of traffic signs, signals, and road markings typical in urban and suburban settings. To prepare the dataset for robust machine learning applications, it has been divided into training (70%), validation (20%), and testing (10%) subsets. Data augmentation techniques such as rotation, translation, scaling, and flipping were employed to enhance the diversity and resilience of the dataset. These augmentations are crucial for training models to recognize and generalize from varied patterns, which is essential for effective traffic sign detection in real-world conditions.

By combining data from internet resources and Roboflow, we ensured a comprehensive representation of traffic sign variations and scenarios, which is crucial for training a robust model capable of detecting signs in real-world situations. To further enrich our dataset and increase its diversity, we employed data augmentation techniques. Data augmentation involves applying various transformations to the existing images, effectively generating new training samples with minimal manual effort. Some common data augmentation techniques relevant to traffic sign detection include image rotation, resizing, cropping, and flipping. Rotation helps the model adapt to traffic signs in different orientations while resizing and cropping simulate variations in the distance from which signs are observed. Flipping horizontally is particularly useful for capturing the symmetrical nature of many traffic signs. Additionally, techniques like brightness and contrast adjustment mimic changing lighting conditions, while adding noise can simulate real-world imperfections in images. These augmentations not only expand the dataset but also expose the model to a wide array of potential variations and challenges it may encounter in practice. This, in turn, enables the model to generalize better and perform accurately in diverse environmental settings, making it a crucial step in training a robust traffic sign detection model. traffic sign dataset classes involves: bus_stop, do_not_enter, do_not_stop, do_not_turn_l, do_not_turn_r, do_not_u_turn, enter_left_lane, green_light, left_right_lane, no_parking, parking, ped_crossing, ped_zebra_cross, railway_crossing, red_light, stop, t_intersection_l, traffic_light, u_turn, warning and yellow_light.

3.2 Model training

The YOLOv8-based deep learning model is trained on the dataset. The model architecture consists of convolutional layers, followed by detection layers responsible for predicting bounding boxes and class probabilities. During training, the model learns to predict the presence and location of traffic signs within the input images. The training process involves optimizing the model's parameters using gradient descent optimization algorithms, such as Adam or SGD, to minimize the detection loss function.

The training process typically takes several epochs, with each epoch consisting of multiple iterations (or batches) of forward and backward passes through the network. The model's performance is evaluated using metrics such as mean Average Precision (mAP) and Intersection over Union (IoU) on a separate validation dataset. Hyperparameters such as learning rate, batch size, and optimizer settings are tuned to optimize the model's performance.

3.3 Model evaluation

The trained model is evaluated on a separate validation dataset to assess its performance metrics, including precision, recall, and mAP at different IoU thresholds. Precision measures the fraction of true positive detections among all predicted detections, while recall measures the fraction of true positive detections among all ground truth traffic signs. mAP calculates the average precision across all classes and is a commonly used metric for object detection tasks. Evaluation metrics are computed by comparing the model's predictions with ground truth annotations. Precision, recall, and mAP scores are calculated for each class (traffic sign type) and averaged across all classes to obtain overall performance metrics. The evaluation results provide insights into the model's ability to detect traffic signs accurately and efficiently.

3.4 Model optimization

Based on the evaluation results, the model may undergo further optimization to improve its accuracy and efficiency. Techniques such as transfer learning, model pruning, and architecture modifications may be applied to enhance the model's performance. Transfer learning involves fine-tuning the pre-trained YOLOv8 model on the specific task of traffic sign detection using the collected dataset. Model pruning techniques may be used to remove redundant or unnecessary parameters from the model to reduce its size and computational complexity. Architecture modifications, such as adjusting the number of layers or adding additional convolutional filters, may be explored to improve the model's ability to detect small or occluded traffic signs.

3.5 Deployment and real-time inference

Once the model is optimized, it is deployed in real-time traffic sign detection systems. The deployed model is capable of processing live video streams or images from traffic surveillance cameras in real-time. Inference results, including detected traffic signs and their corresponding bounding boxes, are displayed on the user interface for real-time monitoring and analysis by transportation authorities or autonomous vehicles.

The deployed model is optimized for efficient inference on hardware platforms commonly used in intelligent transportation systems, such as CPUs, GPUs, or specialized accelerators like NVIDIA Jetson. The model's inference speed is measured in frames per second (FPS), indicating the number of frames processed per second during real-time operation. The model's accuracy and efficiency are validated through extensive testing on real-world traffic sign detection scenarios, ensuring reliable performance under varying environmental conditions.

4 Results and discussion

In this section, we present the experimental results and performance evaluation of our real-time traffic sign detection system using the YOLOv8-based deep learning approach. The performance of the model is evaluated based on several key metrics, including precision, recall, and mAP. Figure 1 shows visual representation of experimental resuly result for the proposed YOLOv8 model.

Evaluating the performance of a YOLOv8 model for traffic sign detection involves assessing its precision, recall, and mAP metrics.

Precision measures the accuracy of the positive predictions made by the model. In the case of traffic sign detection, it tells us how many of the detected signs are correct.

- For each detected traffic sign, determine whether it is a true positive (correctly detected) or a false positive (incorrectly detected).
- Calculate the ratio of true positives to the total number of positive predictions (true positives + false positives).

Recall, also known as sensitivity or true positive rate, quantifies the model's ability to capture all the actual positive instances (traffic signs). In other words, it measures how many of the real signs were detected.

By evaluating precision, recall, and mAP, we gain a comprehensive understanding of our YOLOv8 model's performance in traffic sign detection. High precision suggests that the model makes accurate detections, while high recall indicates that it captures a significant portion of actual signs. mAP provides an aggregate measure of the model's performance, considering different confidence levels, and is particularly useful for comparing models or fine-tuning them to meet specific requirements.

4.1 Confusion matrix

The confusion matrix is a performance evaluation tool for classification tasks, including traffic sign detection using the YOLOv8 model. It provides a detailed summary of the model's predictions, showing how well it classifies instances into different classes, as well as the instances that were misclassified. The confusion matrix is particularly useful for understanding the model's strengths and weaknesses in classifying specific traffic sign classes.

In our case, with multiple traffic sign classes such as "bus_stop," "do_not_enter," "red_light," "stop," and others, the confusion matrix will be a square matrix where each

Fig. 1 Result of the YOLOv8 model

row corresponds to the actual class, and each column corresponds to the predicted class. Each cell in the matrix contains the number of instances that were classified as a particular class. The confusion matrix and these rates are essential for assessing the YOLOv8 model's performance in detecting different traffic sign classes, pinpointing areas where it excels and where it may need improvement. Figure 2 shows the confusion matrix of the YOLOv8 model.

A normalized confusion matrix is a variation of the standard confusion matrix that provides a more interpretable view of the classification performance by representing values as proportions or percentages. Instead of displaying raw counts of instances, it presents the rates of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) in relation to the total instances for each class. This allows for a better comparison of the model's performance across different traffic sign classes, especially when the classes have varying numbers of instances.

In our scenario, with numerous traffic sign classes like "bus_stop," "do_not_enter," "red_light," "stop," and others, the normalized confusion matrix would show the following:

- Each row in the matrix would represent the actual class (ground truth).
- Each column would represent the predicted class by the YOLOv8 model.

The values in the cells would be the rates of instances correctly or incorrectly classified relative to the total number of instances of that actual class.

This normalized confusion matrix offers a clearer perspective on how well the model performs for each class. It allows us to assess the model's accuracy and misclassification rates for different traffic sign classes, aiding in identifying areas that require improvement and understanding the relative difficulty of classifying specific signs in traffic sign detection tasks. The confusion matrix normalized of the YOLOv8 model is conducted in Fig. 3.

4.2 Precision-confidence

The Precision-Confidence Curve is a valuable tool for evaluating the efficiency and effectiveness of a model in multi-class classification tasks like traffic sign detection. This curve illustrates how the precision of the model varies at

Fig. 2 Confusion Matrix of YOLOv8 model

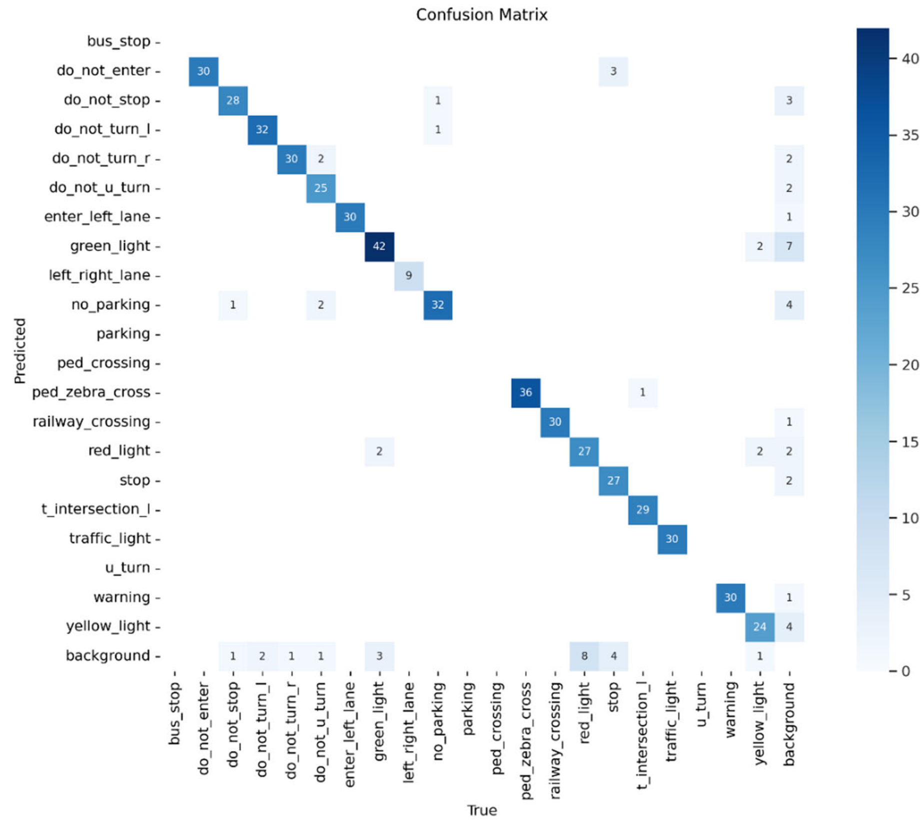
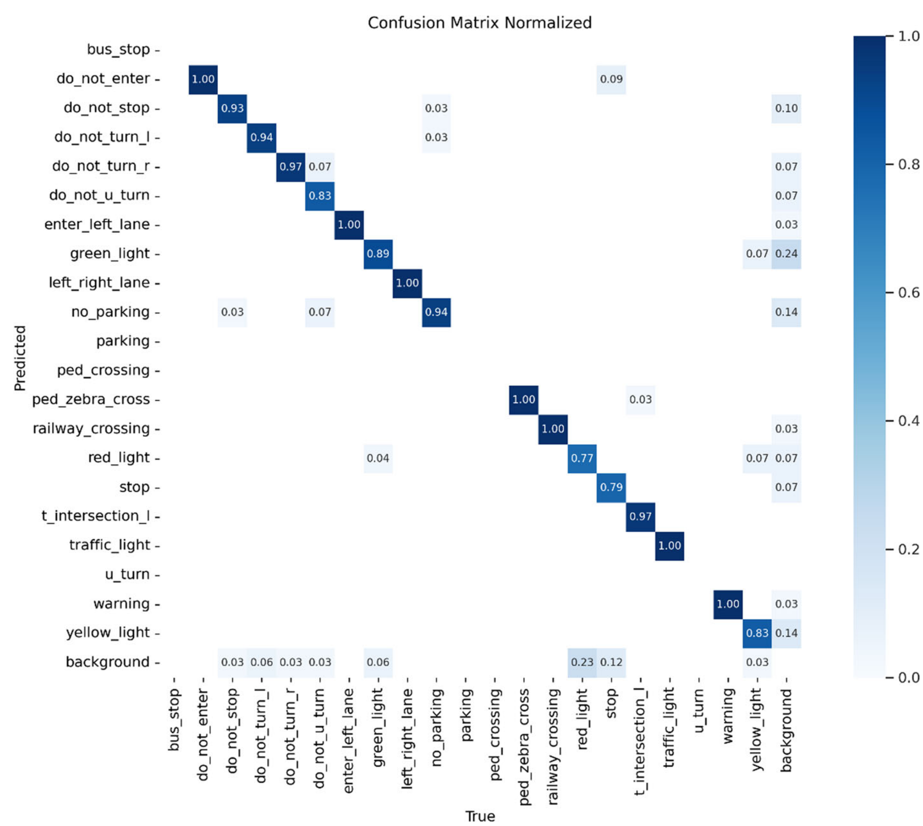


Fig. 3 The confusion matrix normalized of YOLOv8 model



different confidence thresholds for each class. Each class is represented on the x-axis, and the y-axis shows the precision values at different confidence thresholds. It helps in understanding the trade-off between precision and confidence for different classes. In our case, with a maximum precision rate of almost 92.6% for all classes, this signifies that the YOLOv8 model is highly efficient in recognizing the specified traffic sign classes. A precision rate of 92.6% indicates that when the model predicts a traffic sign for a given class, it is correct approximately 92.6% of the time. This is a robust performance, as it indicates that the model has a low rate of false positives for most classes, making it reliable for traffic sign detection.

The consistency of high precision across all classes further underscores the model's effectiveness. It suggests that the model performs well for a diverse set of traffic sign types, from warning signs to traffic lights and various road instructions. The 92.6% precision rate implies that the model can confidently recognize and classify these signs with a high degree of accuracy.

In summary, the YOLOv8 model's excellent precision across the wide range of traffic sign classes, with a maximum of 92.6%, indicates its strong performance in efficiently recognizing these signs. This is a positive sign for its practical use in intelligent transportation systems and road safety applications, as it ensures accurate and reliable detection of traffic signs for various purposes. Figure 4 shows the precision-confidence curve of the YOLOv8 model.

4.3 Recall-confidence

The Recall-Confidence Curve is a tool used to evaluate the efficiency and effectiveness of a model, such as the YOLOv8 model, in multi-class classification tasks like traffic sign detection. This curve shows how the recall rate of the model changes at different confidence thresholds for each class. The recall rate, often called sensitivity, is a measure of the model's ability to capture all relevant instances. In our case, achieving a maximum recall rate of 98% for all classes indicates that the YOLOv8 model is highly efficient in recognizing the specified traffic sign classes. This means that when the model predicts a traffic sign for a given class, it correctly detects approximately 98% of the actual instances of that class, demonstrating the model's strong ability to identify and capture these signs.

The high recall rate across all classes suggests that the YOLOv8 model can reliably recognize a wide range of traffic sign types, making it effective for traffic sign detection. This is particularly important for road safety and traffic management applications, as it indicates that the model can successfully capture and identify the majority of relevant traffic signs in various situations, contributing to improved safety and efficiency on the roads. Figure 5 shows the recall-confidence curve of the YOLOv8 model.

Table 1 The comparison results of the proposed model and others

Model	Precision (%)	Recall (%)	mAP @0.5 (%)
YOLO5s	93.20	94.50	87.50
YOLO8s	96.70	92.60	98
Faster R-CNN	91.80	90.50	89.70
SSD	92.30	89.90	91.20
RetinaNet	94.20	91.70	93.80
Mask R-CNN	92.90	93.50	92.30

4.4 mAP metric

The mAP_0.5 curve is a valuable tool for evaluating the YOLOv8 model's efficiency and effectiveness in traffic sign detection. It provides a visual representation of how mAPchanges at various Intersections over Union (IoU) thresholds. The mAP metrics, including mAP_0.5, offer a comprehensive assessment of the model's precision and recall trade-offs, class-specific performance, and overall model effectiveness. Achieving a high mAP of 96.7% demonstrates that the YOLOv8 model excels in accurately and efficiently detecting traffic signs, making it a promising solution for road safety and traffic management applications. Figure 6 shows the precision-recall curve of the YOLOv8 model.

4.5 Models comparison

In this section, we present a model comparison of six different object detection models: YOLO5s, YOLO8s, Faster R-CNN, SSD, RetinaNet, and Mask R-CNN. We implemented each of these models and collected experimental results to compare their performance for our study. Table 1 presntes the comparison results of the proposed model and others.

As shown in Table 1, the performance of various object detection models is evaluated based on three key metrics: mAP @0.5, recall, and precision. Among the models compared, YOLO8s achieves the highest mAP @0.5 score of 98%, indicating its superior ability to accurately detect and localize objects in images compared to other models. This high mAP @0.5 score suggests that YOLO8s can achieve consistently high performance across a wide range of object detection tasks, making it a robust choice for real-world applications where accuracy is paramount.

In terms of recall, YOLO5s exhibits the highest value of 94.50%, indicating its ability to correctly identify a large proportion of true positive detections among all ground truth objects. However, despite its high recall rate, YOLO5s lags

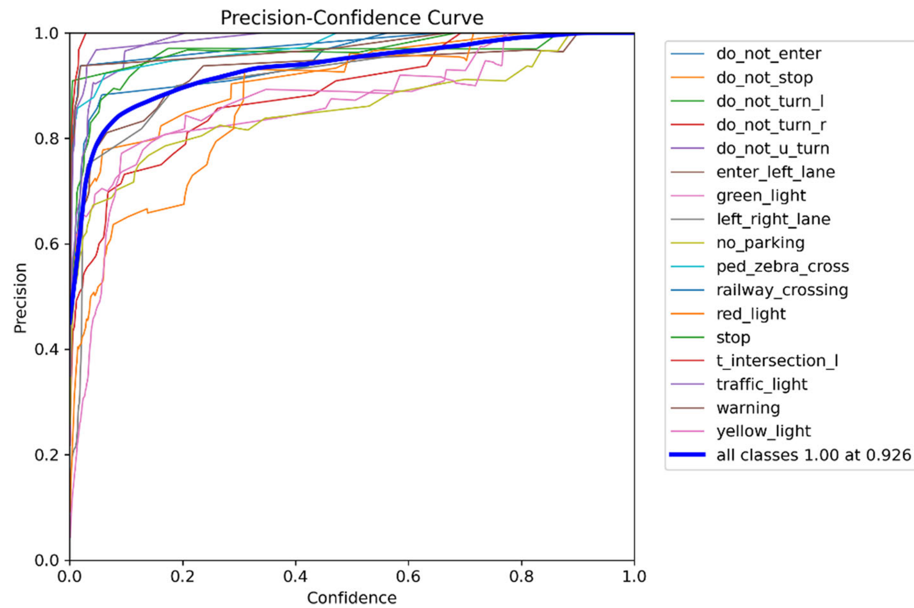


Fig. 4 Precision-Confidence Curve of the YOLOv8 model

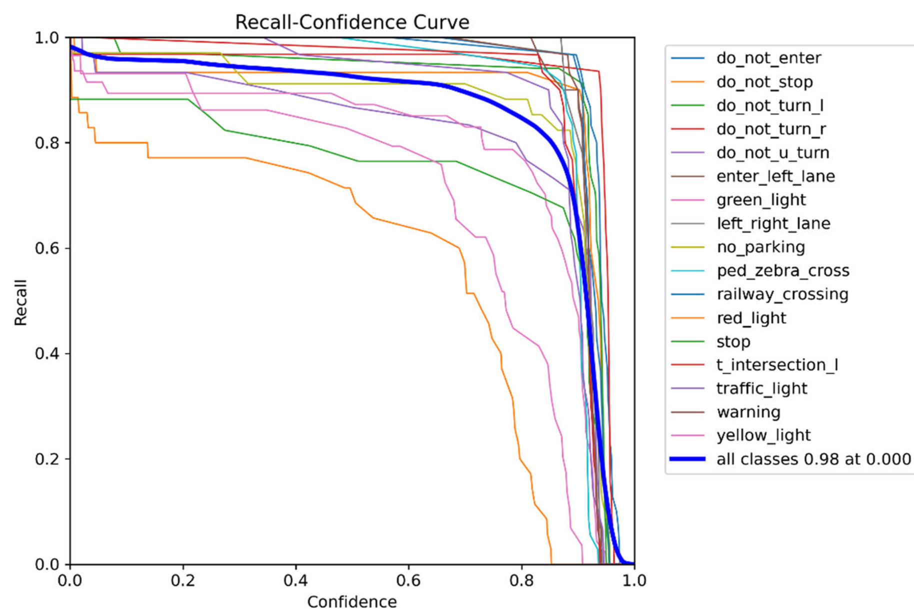


Fig. 5 Recall-Confidence Curve of the YOLOv8 model

behind YOLO8s in terms of mAP @0.5, suggesting that while it may identify a higher number of true positive detections, it may also generate more false positives, leading to a lower overall accuracy.

When considering precision, YOLO8s outperforms all other models with a precision score of 96.70%. This indicates that YOLO8s is highly accurate in correctly identifying true positive detections among all predicted detections, resulting in fewer false positive detections compared to other models. The high precision of YOLO8s makes it particularly well-suited for applications where minimizing false alarms is critical, such as in autonomous driving systems or medical image analysis.

Among the other models evaluated, RetinaNet demonstrates competitive performance with a precision score of 94.20% and an mAP @0.5 score of 93.80%. While RetinaNet falls slightly behind YOLO8s in terms of precision and mAP @0.5, it still outperforms other models such as Faster R-CNN, SSD, and Mask R-CNN. This highlights the effectiveness of RetinaNet in accurately detecting objects in images and its potential utility in various computer vision tasks.

Overall, the results of the evaluation demonstrate that YOLO8s stands out as the superior model among the ones compared, achieving the highest scores in both precision and mAP @0.5 metrics. While YOLO5s exhibits a higher

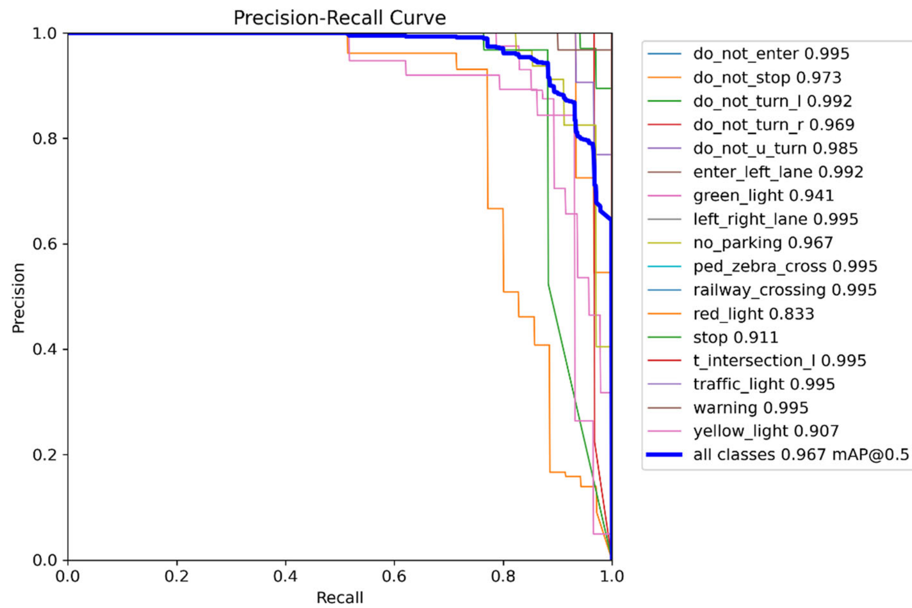


Fig. 6 Precision-Recall Curve of the YOLOv8 model

recall rate, its lower precision and mAP @0.5 scores suggest that YOLO8s offers a better balance between accuracy and recall, making it the preferred choice for object detection tasks requiring high precision and robust performance.

4.6 Evaluation of using the proposed model in real-time

When assessing the application of a YOLOv8-based deep learning model for real-time traffic sign detection, several critical factors must be considered. These encompass performance metrics, computational efficiency, and the challenges of practical deployment.

Performance metrics YOLOv8 is celebrated for its superior performance in object detection, especially noted for its accuracy. Key metrics such as precision, recall, and mean Average Precision (mAP) are fundamental, ensuring minimal false positives and maximum detection of relevant traffic signs, crucial for maintaining safety. The model's real-time responsiveness, vital for traffic sign detection, is achieved through an optimized architecture that enables quick predictions without compromising precision. The frame rate, measured in frames per second (FPS), is essential as it indicates the model's capability to process images swiftly, a necessity in dynamic traffic environments where prompt and accurate detection is critical.

Computational efficiency The practical application of YOLOv8 heavily depends on its ability to operate efficiently across commonly available hardware platforms in intelligent transportation systems, such as CPUs, GPUs, and specialized accelerators like NVIDIA Jetson. To enhance the model's efficiency, optimization techniques such as model pruning, quantization, and the utilization of lighter convolutional layers are employed. These methods help in reducing the model size and accelerating inference times, which are indispensable for deployments in environments with limited computational resources.

Practical deployment challenges Deploying YOLOv8 involves addressing issues like environmental variability and scalability. The model must perform robustly across a range of conditions, including varying lighting and weather, challenges that demand extensive testing and advanced training techniques like domain adaptation. Moreover, the model needs to remain adaptable to accommodate evolving traffic rules and sign designs, necessitating mechanisms such as continuous learning or periodic re-training with updated datasets.

Cost-effectiveness Implementing YOLOv8-based systems entails evaluating initial setup costs—including hardware procurement and system integration—and ongoing operational costs, such as maintenance and upgrades. These expenses must be balanced against the benefits of enhanced traffic safety and efficiency to assess the overall cost-effectiveness.

Safety and reliability In traffic management systems utilizing YOLOv8, safety and reliability are paramount. Even minor errors in traffic sign detection can lead to significant consequences, especially in high-speed or densely populated urban settings. Therefore, ensuring the system's reliability across various scenarios is crucial. Integrating robust fail-safe mechanisms to manage potential system failures effectively is also essential, underscoring the importance of prioritizing safety and reliability in the system's design and implementation.

In conclusion, the deployment of a YOLOv8-based model for real-time traffic sign detection requires a holistic approach, considering not just technical specifications and performance but also the practical aspects of deployment, cost management, and safety protocols.

5 Conclusion and future work

Traffic sign detection is a fundamental component of intelligent transportation systems, contributing to road safety and efficient traffic management. Numerous methods have been explored in the literature to address this challenge, with deep learning approaches emerging as the most promising due to their superior accuracy. However, current research confronts the challenge of achieving high accuracy rates while meeting the non-destructive and real-time requirements of intelligent transportation systems. In this study, we propose a deep learning model based on the YOLOv8 architecture to address this challenge. Leveraging a custom dataset and rigorous training, validation, and testing procedures, we demonstrate the efficacy of our method through extensive experiments and performance analysis, achieving precise and reliable results, thus providing a solution to the existing research challenge in deep learning-based traffic sign detection. While our proposed deep learning-based approach for traffic sign detection shows promising results, there are two notable limitations that need to be addressed in future research. First, the method's performance may still be influenced by environmental factors, such as adverse weather conditions or lighting variations, which can hinder the accuracy of detection. Secondly, our approach, while efficient, may require significant computational resources, making it less practical for deployment in resource-constrained environments. In light of these limitations, future work should focus on developing robust models that are less susceptible to environmental variability and on optimizing the computational efficiency of the deep learning method. Additionally, exploring the integration of real-time data feeds and further enhancing the model's adaptability to different urban settings could be areas of future research to improve the practicality and effectiveness of traffic sign detection in intelligent transportation systems.

Acknowledgements I would like to take this opportunity to acknowledge that there are no individuals or organizations that require acknowledgment for their contributions to this work.

Author contributions Mingdeng Tang: Writing-Original draft preparation, Conceptualization, Supervision, Project administration.

Funding This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data availability Data can be shared upon request.

Declarations

Competing interests The author declare no competing interests.

Ethical approval The research paper has received ethical approval from the institutional review board, ensuring the protection of participants' rights and compliance with the relevant ethical guidelines.

References

1. Zhang, J., Zou, X., Kuang, L.-D., Wang, J., Sherratt, R.S., Yu, X.: CCTSDB 2021: a more comprehensive traffic sign detection benchmark. Human Centric Comput. Inf. Sci. **12**, 2022 (2021)
2. Wang, J., Chen, Y., Dong, Z., Gao, M.: Improved YOLOv5 network for real-time multi-scale traffic sign detection. Neural Comput. Appl. **35**(10), 7853–7865 (2023)
3. Gu, Y., Si, B.: A novel lightweight real-time traffic sign detection integration framework based on YOLOv4. Entropy **24**(4), 487 (2022)
4. Liang, T., Bao, H., Pan, W., Pan, F.: Traffic sign detection via improved sparse R-CNN for autonomous vehicles. J. Adv. Transp. **2022**, 1–16 (2022)
5. Ang, M.C., Sundararajan, E., Ng, K.W., Aghamohammadi, A., Lim, T.L.: Investigation of threading building blocks framework on real time visual object tracking algorithm. Appl. Mech. Mater. **666**, 240–244 (2014)
6. Li, X., Xie, Z., Deng, X., Wu, Y., Pi, Y.: Traffic sign detection based on improved faster R-CNN for autonomous driving. J. Supercomput. **78**, 1–21 (2022)
7. Liu, Y., Shi, G., Li, Y., Zhao, Z.: M-YOLO: Traffic sign detection algorithm applicable to complex scenarios. Symmetry (Basel) **14**(5), 952 (2022)
8. Jiang, L., Liu, H., Zhu, H., Zhang, G.: Improved YOLO v5 with balanced feature pyramid and attention module for traffic sign detection. In: MATEC Web of Conferences, EDP Sciences, p. 03023 (2022)
9. Lahmyed, R., El Ansari, M., Kerkaou, Z.: Automatic road sign detection and recognition based on neural network. Soft. Comput. **26**, 1–22 (2022)
10. Ang, M.E.I.C., Aghamohammadi, A., Ng, K.O.K.W., Sundararajan, E., Mogharrebi, M., Lim, T.L.: Multi-core frameworks investigation on a real-time object tracking application. J. Theor. Appl. Inf. Technol. **70**(1), 163–171 (2014)
11. Aghamohammadi, A., Ang, M.C., Sundararajan, E.A., Weng, N.K., Mogharrebi, M., Banihashem, S.Y.: A parallel spatiotemporal saliency and discriminative online learning method for visual target tracking in aerial videos. PLoS ONE **13**(2), e0192246 (2018)
12. Karthika, R., Parameswaran, L.: A novel convolutional neural network based architecture for object detection and recognition with an application to traffic sign recognition from road scenes. Pattern Recognit Image Anal. **32**(2), 351–362 (2022)
13. Min, W., Liu, R., He, D., Han, Q., Wei, Q., Wang, Q.: Traffic sign recognition based on semantic scene understanding and structural traffic sign location. IEEE Trans. Intell. Transp. Syst. **23**(9), 15794–15807 (2022)
14. Zhu, Y., Yan, W.Q.: Traffic sign recognition based on deep learning. Multimed. Tools Appl. **81**(13), 17779–17791 (2022)
15. Bai, W., et al.: Two novel models for traffic sign detection based on YOLOv5s. Axioms **12**(2), 160 (2023)
16. Li, Y., Li, J., Meng, P.: Attention-YOLOV4: a real-time and high-accurate traffic sign detection algorithm. Multimed Tools Appl. **82**(5), 7567–7582 (2023)

17. Megalingam, R.K., Thanigundala, K., Musani, S.R., Nidamanuru, H., Gadde, L.: Indian traffic sign detection and recognition using deep learning. Int. J. Transp. Sci. Technol. **12**(3), 683–699 (2023)
18. Wei, H., Zhang, Q., Qian, Y., Xu, Z., Han, J.: MTSDet: multi-scale traffic sign detection with attention and path aggregation. Appl. Intell. **53**(1), 238–250 (2023)