Proceedings of CONF-CDS 2025 Symposium: Application of Machine Learning in Engineering
DOI: 10.54254/2755-2721/2025.AST26309

A Review of YOLO Algorithm Application in Object Detection in Complex Traffic Environments

Zebin Luo

Faculty of Science and Technology, Middlesex University, London, United Kingdom
ZL248@live.mdx.ac.uk

Abstract. Complex traffic environments object detection is a key area in intelligent transportation and autonomous driving technology research. It has a direct influence on target recognition accuracy and provides credible data support to environmental understanding and decision-making. This paper reviews recent advances in YOLO-based object detection methods in traffic scenarios, including challenges of long-range detection, occlusion, truncation, illumination, and weather. Over the last few years, many of these models have been optimized and trained for the above purposes with methods such as attention mechanisms, feature fusion, and improved loss functions. Detection accuracy, robustness, and real-time response under some conditions have all been enhanced using these methods. This paper adopts literature review and analysis techniques to conclude the performance of general YOLO models after some optimizations in some complex environments, and discusses future possibilities for model improvement. The findings are expected to help with intelligent traffic system development and serve as a suitable reference for future practical use.

Keywords: YOLO, machine learning, complex traffic environments, object detection

1. Introduction

In recent years, with the continuous development of modern society, the complexity of traffic environments has been increasing. This growing complexity places higher demands on autonomous driving technologies and makes object detection in complex traffic scenes a significant research focus. Complex traffic environments typically involve a wide variety of road participants, such as pedestrians, vehicles, dense and overlapping object distributions, as well as diverse weather conditions and varying lighting across different times of day [1]. YOLO-based object detection methods offer promising solutions for this task. You Only Look Once (YOLO) is a viral and widely used algorithm [2]. The YOLO algorithm is known for its simple architecture and fast detection speed, making it suitable for real-time applications. However, its accuracy may be lower than that of region proposal-based methods. To address this limitation, multiple versions of the YOLO algorithm have been developed, providing feasible solutions for object detection in increasingly complex traffic scenarios. This paper focuses on recent developments of YOLO-based object detection models in complex traffic environments. The study is conducted through a literature review and analysis to investigate how the models handle specific complex cases. Conclusions derive effective

DOI: 10.54254/2755-2721/2025.AST26309

insights for intelligent transportation system development and offer good references for subsequent research and model optimization in complex and dynamic environments.

2. Application of YOLO algorithms for object detection in complex traffic environments

2.1. Vehicle detection

Vehicle detection in traffic scenes provides three typical challenges: far-distance detection, object truncation, and occlusion. These are also central issues for vehicle detection in advanced traffic scenarios, in that they tend to affect the consistency of decision-making for intelligent driving systems and thereby have a direct influence on traffic safety overall. A robust model specially designed for harsh traffic scenes was proposed in VP-YOLO. A novel Visual Attention Module(VAM), inspired by visual cross-connection, was introduced to blend high-quality horizontal and vertical spatial features. The module seeks to capture channel-wise information between heterogeneous spatial locations to mimic cross-view dependency mechanisms [3]. For addressing the issue of truncated objects, the model contains a Feature Reconstruction Module (FRM). This module, combined with attention mechanisms, helps achieve more accurate object localization and detection of partially occluded objects. Moreover, the model enhances the representation of small targets at the augmentation stage by enriching contextual information and taking advantage of the intrinsic features of the objects.

Compared to traditional YOLOv7, the method handles long-range, truncated, and occluded conditions with stronger capability. Furthermore, a VP-specific dataset was proposed a VP-specific dataset containing a set of difficult traffic scenarios, which remedies the weaknesses of traditional datasets to study detection algorithms in real-world scenarios. This improvement proves effective in enhancing the model's perception ability in adverse environments.

2.2. Pedestrian detection

In complex traffic scenes, pedestrian detection is crucial to public safety as it directly impacts pedestrian path prediction and driving behavior in intelligent driving systems. Pedestrian detection in real-world environments, however, faces many challenges such as diverse postures, mutual occlusion, complex weather conditions, and long-distance recognition, all of which make model development challenging. A lightweight real-time pedestrian detection model, named LP-YOLO, was proposed to address these challenges [4]. It is composed of multiple crucial modules: an enhanced Cross Stage Partial (CSP) bottleneck with a fast implementation of spatial-channel convolution (C3k2SC), and a spatial-channel decoupled downsampling module (SCDown). The model also integrates an optimized feature fusion receptive field attention mechanism for enhancing multi-scale feature integration. Compared to YOLOv11n on dataset CityPerson, the model reduces parameters by 24.8% and computation cost by 7.94% while increasing mAP@0.5 by 1.6%. It is able to achieve this in lower complexity without losing important feature information. LP-YOLO also shows strong generalization ability, performing comparably to the baseline model, especially under extreme weather conditions and at ultra-long distances.

VP-YOLO is a model that excels at detecting pedestrians and cars in real-world traffic conditions. It uses a dedicated dataset called VP, which is based on publicly available datasets but re-annotated for focusing the task on occluded or truncated cars and pedestrians. At the enhancement level, it incorporates VAM and FRM modules, which increase the accuracy, robustness, and

DOI: 10.54254/2755-2721/2025.AST26309

explainability of the model to a large degree. This renders it comparatively more suitable for use in intricate and dynamic traffic conditions.

2.3. Traffic sign detection

Traffic signs play a significant role in traffic management. They ensure road safety and direct the actors in traffic situations. However, existing detection solutions largely lack high accuracy and poor real-time performance in sophisticated and dynamic traffic environments. Furthermore, traffic signs always appear as miniature and distant targets, significantly increasing the detection difficulty.

DP-YOLO, based on YOLOv8s, proposes an approach in which the baseline model's large-object detection head is discarded and a new small-object detection head is introduced [5]. It includes the DBBNCSPELAN4 module to improve feature extraction ability and designs the PTCSP module, which embeds Transformer mechanisms into the feature processing network. These modifications save both computational and parameter expenses. DP-YOLO, as compared to YOLOv8, has a parameter saving of 77.0% with improvement in mAP@0.5 by 5.8%, 2.7%, and 1.3% on TT100K, GTSDB, and CCTSDB datasets respectively. The model significantly improves small traffic sign detection and is adequate for edge deployment applications.

YOLO-BS, built upon the YOLOv8 framework, introduced a detection method with a small-object detection head and a Bidirectional Feature Pyramid Network (BiFPN) [6]. The detection achieved 90.1% mAP@0.5 and 78 FPS, outperforming mainstream models. The method achieves higher accuracy and more stable detection, especially for far and small traffic signs.

2.4. Road marking detection

Road marking detection is an active area of computer vision research that has a particular application in autonomous driving and instance segmentation. Effective detection of road markings makes them visible and improves the safety of drivers, pedestrians, and the overall traffic environment.

Gupta et al. proposed an optimized version of YOLOv7 to improve road marking detection [7]. The model combines the enhanced AdamW optimizer and a novel network structure called the Multi-Stage Cross Convolution Bottleneck Network (MSCCBN). The MSCCBN is designed to learn the different patterns and shapes of road markings effectively. Their model achieved a mAP@0.5 score of 0.889 on the Ceymo dataset with high accuracy and lower false positives and false negatives.

However, the model requires high-resolution input images to accurately identify road markings. While it shows robustness to low-light and partial overexposure, it may not be affected in more inclement weather, such as snow or fog, which were not included in the training data.

3. Application of YOLO algorithm in object detection in common usage scenarios

3.1. Detection under complex weather conditions

The YOLO algorithm is now a versatile and popular object detection model. However, its performance in adverse weather conditions has been of relatively less interest so far. Under tough-to-manage environments such as heavy rain, fog, snow, or low-light, the detection faults in the model would be more severe. Li et al. proposed an improved object detection algorithm after YOLOv11 as MFA-YOLO [8]. They introduced the wavelet transform-based C3k2-WT module, SimSPPF+, and C2PSA-HAFM modules, and developed a new classification loss function called

DOI: 10.54254/2755-2721/2025.AST26309

BE-ATFL. This method achieved an mAP of 61.7% and accuracy of 70.6% on the EWD, GAD, and UA_DETRAC datasets, providing its reliability and generalization ability and thus making it suitable for object detection under adverse weather in autonomous driving.

Liu et al. presented RFCS-YOLO, a new version of YOLOv7 with stronger receptive field and multi-scale features [9]. RFCS-YOLO uses an Efficient Feature Extraction Module(EFEM) and a Cross-Scale Fusion module (CSF), and adopts a new loss function, Focaler-Minimum Point Distance Intersection over Union (F-MPDloU). Compared to the standard YOLOv7, it works much better on the DWAN dataset with a mAP of 86.5%. On a generalization test on the autonomous driving dataset SODA10M, it scored an mAP@0.5 of 56.7%, showcasing its capacity to reduce missed and incorrect detections of proceeding further in augmenting the dataset with scene images captured automatically under different weather and environmental conditions to further improve the model's generalization under dynamic traffic conditions.

3.2. Detection in low-light environment

For nighttime object detection, many algorithms have already been set up to support traffic detection under low-light conditions. A representative solution for this task is Dark-YOLO, which improves detection performance under nighttime conditions [10]. However, it also fails in attempting to detect distant or small traffic lights under very low light. In order to counter this, Liu et al. designed Dark-YOLO, which is a low-light object detection model that employs an Adaptive Image Module, Dynamic Feature Extraction Module, and Inverted Dimensional Attention Module. All of these modules work in association with each other to consider various features in low-light images. Dark-YOLO on the ExDark dataset had a mAP@50 of 71.3%, which proves its capability in object detection in low-light conditions.

In the meantime, LNT-YOLO focuses on detecting small-sized traffic lights in nighttime scenes and is developed based on the YOLOv7-tiny model [11]. It is designed to recover small-object features discarded in the neck of YOLOv7-tiny with low-level feature knowledge. The model introduced a new attention module, SEAM, by which they use the benefit of SimAM and ECA to extract more channel and spatial features. They also introduced a loss function, HSM-EloU, which is aimed at promoting learning signal on drivers with tiny traffic lights. The authors also introduced the TN-TLD dataset, a well-annotated image dataset of nighttime driving scenes. Compared to the base YOLOv7-tiny model, LNT-YOLO boosted mAP by 13.7% to 26.2%, and improved performance on the LISA dataset by 9.5% to 25.5%. These results show that the model is incredibly viable and resilient and is thus well-suited for the detection of small traffic lights in low-light night environments.

4. Conclusion

This article introduces the latest development of YOLO-based object detection methods in complicated traffic situations. This paper finds that complicated traffic situations bring various challenges, such as long-range detection, occlusion, truncation, complex weather, and night vision, which are likely to affect the decision-making and reliability of intelligent driving systems significantly. Many models have been developed in the last few years to detect vehicles, pedestrians, and traffic signs in complicated traffic situations. Such models typically incorporate specialized modules, e.g., attention, feature merging enhanced, or new loss functions, to improve accuracy, robustness, and real-time performance. Observing that most of the earlier work has already carried out systematic reviews of the evolution of different versions of YOLO, this paper focuses more on

DOI: 10.54254/2755-2721/2025.AST26309

their actual applications in real-world traffic scenarios with complex traffic conditions and does not go into the evolutionary growth of conditions like heavy fog, snow, and complex occlusions. Subsequent research can focus on developing richer and dedicated datasets, combining data from multiple sources, and developing more accurate modules to further improve YOLO performance in complex and dynamic traffic situations.

References

[1] Zhao, L., Li, T., Zhang, H., Liu, Q., & Meng, R. (2025). MDSD-YOLO: A target detection method for complex road scenes. Engineering Technology II.

[2] Sultana, F., Sufian, A., & Dutta, P. (2020). A review of object detection models based on convolutional neural network. Intelligent Computing: Image Processing Based Applications, 1-16.

[3] Liu, W., Qiao, X., Zhao, C., Deng, T., & Yan, F. (2025). VP-YOLO: A human visual perception-inspired robust vehicle-pedestrian detection model for complex traffic scenarios. Expert Systems with Applications, 126837.

[4] Qu, Z., Liu, H., Kong, W., Gu, J., Wang, C., Deng, L., . . . Lin, F. (2025). LP-YOLO: An improved lightweight pedestrian detection algorithm based on YOLOv11. Digital Signal Processing, 105343.

[5] Qiu, J., Zhang, W., Xu, S., & Zhou, H. (2025). DP-YOLO: A lightweight traffic sign detection model for small object detection. Digital Signal Processing, 105311.

[6] Zhang, H., Liang, M., & Wang, Y. (2025). YOLO-BS: a traffic sign detection algorithm based on YOLOv8. Scientific Reports, 7558.

[7] Gupta, R. K., Pateriy, R., Swain, D., & Kumar, A. (2025). Road Marking Detection in Challenging Environments using Optimized YOLOv7. Procedia Computer Science, 258, 925-936.

[8] Li, Z., & Fan, H. (2025). MFA-YOLO: Multi-Scale Fusion and Attention-Based Object Detection for Autonomous Driving in Extreme Weather. Electronics, 14(5), 1014.

[9] Liu, G., Huang, Y., Yan, S., & Hou, E. (2025). RFCS-YOLO: Target Detection Algorithm in Adverse Weather Conditions via Receptive Field Enhancement and Cross-Scale Fusion. Sensors, 25(3), 912.

[10] Munir, S., & Lin, H.-Y. (2025). LNT-YOLO: A Lightweight Nighttime Traffic Light Detection Model. Smart Cities, 8(3), 95.

[11] Liu, Y., Li, S., Zhou, L., Liu, H., & Li, Z. (2025). Dark-YOLO: A Low-Light Object Detection Algorithm Integrating Multiple Attention Mechanisms. Applied Sciences, 15(9), 5170.