

# Improved Target Detection Algorithm for Foggy Conditions in YOLOv8

Hanwei Mao \*, Peng Wang, Liming Zhou, Zhiren Zhu, Xiangmeng Ren

College of Electrical Engineering, Tianjin University of Technology and Education, Tianjin, China

\* Corresponding author: Hanwei Mao (Email: 15022471661@163.com)

**Abstract:** In a foggy environment, due to the significant decrease in visibility, the target image captured by the vehicle camera becomes blurred and the information is incomplete, which exacerbates the problem of misdetection and omission in the process of target detection, and poses a serious challenge to driving safety and navigation accuracy. In this regard, a foggy target detection algorithm based on improved YOLOv8 is proposed. First, by introducing a bidirectional feature pyramid BiFPN structure in the backbone network, the algorithm is able to better capture target detection features through bidirectional connections. Secondly, the Shuffle Attention mechanism is added to the neck network to enhance the diversity of the input sequences by randomly disrupting and grouping them, thus improving the performance of the self-attention mechanism, and thus improving the detection accuracy of the network model. The experimental results show that the average accuracy of the improved model on the RTTS dataset is increased by 1.3% mAP@0.5 and 1.2% by mAP@0.5:0.95. In summary, the improved model YOLO\_BIS can show good performance when dealing with target detection tasks in foggy scenarios.

**Keywords:** Target Detection; YOLOv8; BiFPN; Shuffle Attention.

## 1. Introduction

With the rapid development of science and technology, automatic driving [1] technology, as a core component of intelligent transportation system, is gradually moving from concept to reality, leading a profound change in the future travel mode. The realization of automatic driving technology relies on the in-depth integration of multiple advanced technologies, among which, target detection technology, as a key part of environment perception, is crucial for ensuring vehicle driving safety and enhancing the decision-making ability of automatic driving system. At present, the target detection technology based on YOLO algorithm is widely used in foggy scenes. Su Tong et al [2] based on DehazeNet defogging algorithm combined with improved YOLOv5 algorithm to detect vehicles and people in foggy days, using lightweight GSConv instead of the standard convolution, which improves the network feature fusion ability. Lirens et al[3]proposed an improved method based on the Double-Head framework, which optimizes the feature extraction part of the image and the prediction head, thus improving the ability of the network to focus on the target. Zhang et al [4] used YOLOv7 as an augmented network and introduced the MSFFA structure to improve visibility, which improves the effectiveness of the network in multi-class object target detection. Although the target detection technology for vehicles and pedestrians in foggy scenarios has made some progress in recent years, which significantly enhances the detection performance, however, the current mainstream foggy target detection methods are still facing many challenges, and their limitations cannot be ignored. For this reason, this paper proposes an improved target detection algorithm for YOLOv8 under foggy conditions.

## 2. YOLOv8 Algorithm

YOLOv8 [5] is an advanced target detection algorithm, which inherits the advantages of its predecessor YOLO series

algorithms and makes several important improvements and optimizations. YOLOv8 uses BCE Loss (Binary Cross Entropy Loss) as classification loss, DFL Loss (Distribution Focal Loss) and CIoU Loss (Complete Intersection over Union) as the regression loss, which helps to measure the similarity of the target frames more accurately and further improves the convergence speed and performance of the model. Compared with the original model, YOLOv8 is able to better balance the relationship between accuracy and speed, and is more suitable for target detection tasks in different scenarios.

## 3. Model Optimisation

### 3.1. BiFPN Module

Unlike traditional FPNs, BiFPN[6]introduces bi-directional connections between neighboring levels of the feature pyramid. This means that information can flow from higher level features to lower level features (top-down path) or from lower level features to higher level features (bottom-up path). Bidirectional connectivity [7] allows integration of information from different levels of the feature pyramid in both directions. This integration helps to capture multi-scale features efficiently. BiFPN uses a weighted feature fusion mechanism to combine features from different levels. The fused weights are learned during the training process, ensuring optimal feature integration. Bidirectional connectivity in BiFPN helps to better capture feature representations at different scales, improving the network's ability to handle objects of different sizes and complexity. As shown in Figure 1, this is particularly important in target detection tasks, where the size of objects in an image may vary significantly.

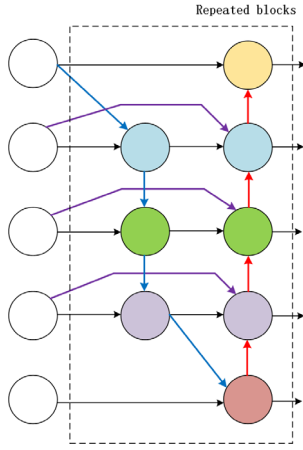


Figure 1. BiFPN module structure diagram

### 3.2. ShuffleAttention Module

Shuffle Attention Attention Mechanism [8] uses a self-attention mechanism to compute the attention weights between each vector and the other vectors. The vectors within each group are weighted and summarized according to the attention weights to obtain the representation within the group. Reorganization can recombine the representations within each group into a new sequence. Reflective shooting maps the new sequence back to the original high-dimensional vector space via an inverse mapping function. The output yields a new representation that can be used for subsequent tasks such as classification or generation. As shown in Figure 2, the core idea is to enhance the diversity of input sequences by randomly disrupting and grouping them, thus improving the performance of the self-attention mechanism.

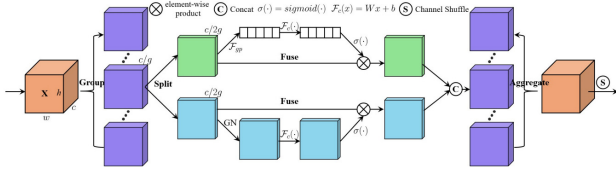


Figure 2. ShuffleAttention module structure

In this paper, based on the YOLOv8n model, a new foggy target detection algorithm, YOLO\_BIS, is proposed to address the misdetection and omission problems that occur when it performs target detection in foggy scenarios. Through the application of this algorithm, the probability of misdetection and omission is reduced, and the robustness and reliability of the model are improved.

A weighted bidirectional feature pyramid network, BiFPN, is added to the Neck section, which introduces learnable weights to learn the importance of different input features, while top-down and bottom-up multiscale feature fusion is applied iteratively. One of the core innovations of BiFPN is the bidirectional cross-scale connectivity, which allows for features to be more comprehensively informative between different layers through top-down and bottom-up paths transfer and fusion. This is different from traditional FPNs and PANs, which mainly use top-down feature propagation.

The ShuffleAttention mechanism is added to layers 16 and 20 of the Neck part, which can enhance the diversity of input sequences by randomly disrupting and grouping to improve the performance of the self-attention mechanism. The structure of the YOLO\_BIS network is shown in Figure 3.

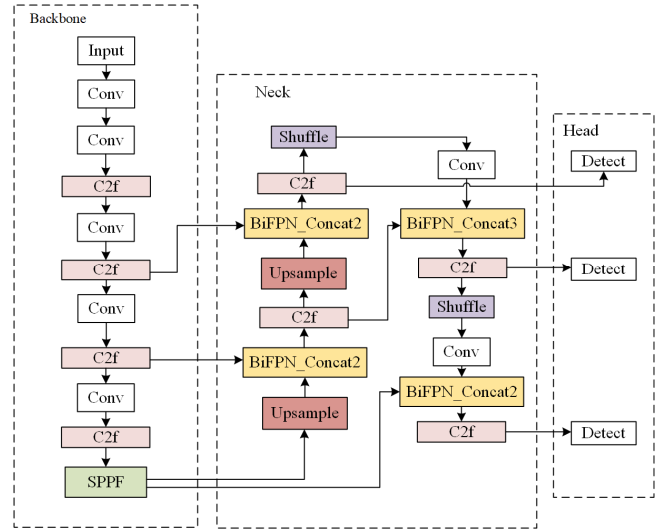


Figure 3. Improved YOLOv8 network structure

## 4. Experimental Result and Analysis

### 4.1. Datasets

The operating system for this experiment is Windows 11 Professional, 64-bit OS, running on 128G of RAM, Intel(R) Xeon(R) Gold 6226R CPU @ 2.90GHz 2.89 GHz (2 processors). Python programming language was used to complete the training of the model in Pycharm version 2023.1. batch-size was set to 4 and the number of iterations epoch was set to 200.

In order to deeply investigate the target detection effectiveness of YOLOv8 in foggy environment, the RTTS dataset [9] is chosen for empirical analysis. The RTTS dataset is derived from the foggy day subset of RESIDE dataset (RESIDE-Foggy Day Dataset), which is a publicly available dataset specifically designed for foggy day image processing and computer vision research. Specifically, the RTTS dataset covers 4322 real foggy day images that contain diverse target types as well as various occlusion situations. In the experiments in this paper, we divide the dataset into a test set (433 images), a training set (2161 images), and a validation set (1728 images). These datasets are widely used in computer vision tasks such as target detection, image enhancement and defogging. In order to improve the performance and robustness of the model under foggy conditions, we especially use rectangular labeling frames to specifically filter out a variety of targets such as bicycles, buses, cars, motorcycles and pedestrians for training.



Figure 4. Images of selected datasets

### 4.2. Evaluation Indicators

In this paper, Average Precision (AP), Precision (P), Recall (R) and Mean Average Precision (mAP) [10] are used as the evaluation metrics for target detection, and the corresponding calculation formulas are shown in Eqs. (1) to (4).

$$AP = \int_0^1 P(R)dR \quad (1)$$

$$precision = \frac{TP}{TP + FP} \quad (2)$$

$$recall = \frac{TP}{TP + FN} \quad (3)$$

$$mAP = \frac{\sum AP}{N} \quad (4)$$

Where TP denotes the number of correct positive samples, FP denotes the number of incorrect positive samples, i.e., the actual category is a negative sample but the model predicts it as a positive sample. fn denotes the number of incorrect negative samples, i.e., the actual category is a positive sample and the model predicts it as a negative sample. n is the number of all categories.

### 4.3. Analysis of the Results

In order to assess the effectiveness of the improved model, the original algorithm was compared with the improved algorithm YOLO\_BIS. mAP@0.5 The curve comparison graph is shown in Figure 6. From Fig. 5, it can be seen that the detection effect of the improved algorithm is significantly improved compared with the original algorithm, which proves that the improved algorithm is effective for detection.

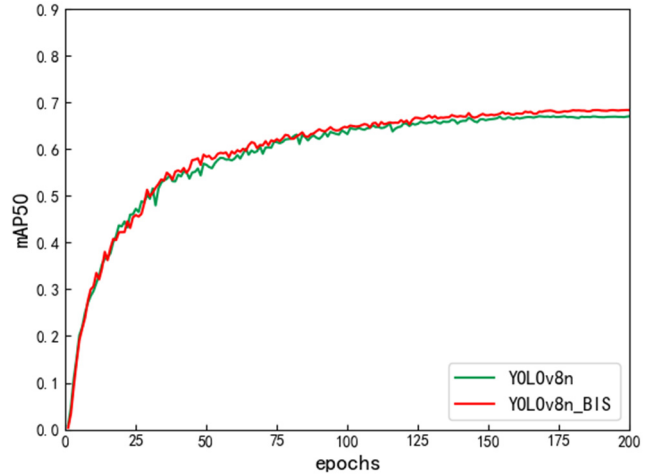


Figure 5. Comparison chart of mAP@0.5 curve

### 4.4. Ablation Experiments

In order to verify the effectiveness of the model improvement methods on the experiments, the different improvement methods are evaluated by doing ablation experiments. Table 1 shows that the YOLOv8 base model has an average accuracy of 67% on the dataset. After adding BiFPN, the YOLOv8\_B model has a 0.5% increase on mAP0.5 compared to the original base model. It shows that BiFPN has some effect on model accuracy improvement. After adding ShuffleAttention, the YOLOv8\_S model shows a growth of 0.3% on mAP0.5 compared to the original base model. In terms of detection accuracy, the improved model improves by 1.3% compared to the traditional model, and the YOLO\_BIS model has improved target detection accuracy for the foggy dataset, which verifies the applicability to target detection in foggy scenes.

Table 1. Ablation experiments

Algorithm	BiFPN	Shuffle	mAP@0.5/%	mAP@0.5:0.95/%	FPS
YOLOv8n			67	43.4	62.11
YOLOv8n_B	√		67.5	44	47.39
YOLOv8n_S		√	67.3	43.7	45.87
YOLO_BIS	√	√	<b>68.3</b>	44.6	51.28

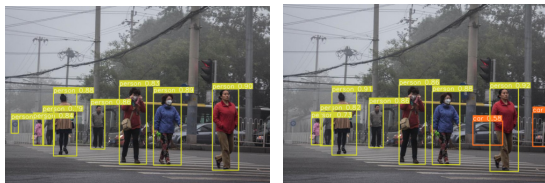
### 4.5. Detection Effect

Considering that the traffic situation in foggy scenes is complex and changeable, and there will be situations such as target occlusion and visible range changes, in order to verify the applicability of the improved model more intuitively, visual comparison experiments are carried out after constructing the target test set in foggy environments with the aim of comparing the performance differences between the traditional model and the improved model in target recognition. The comparison result graphs are shown in Fig. 6, with the target detection graph of the original model YOLOv8n on the left and the target detection graph of the improved model YOLO\_BIS on the right. Among them, Fig. 6(a) represents the comparison map for the target occlusion case; Fig.6(b) represents the comparison map for the dense fog environment; Fig.6(c) represents the comparison map for the close-range case. From Fig.6, it can be seen that the original model has target misdetection and omission in foggy scenarios, and the improved model reduces the omission of targets, improves the detection accuracy of the model, and has

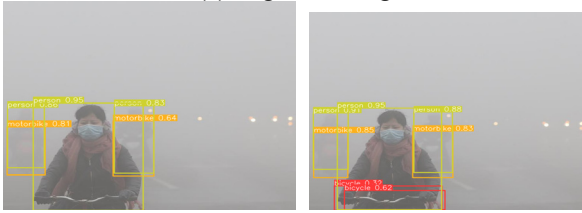
better applicability for the target detection task in foggy scenarios.

## 5. Conclusion

Aiming at the problems existing in the current target detection method on the lane in foggy scenarios, this paper proposes a foggy target detection algorithm based on improved YOLOv8: YOLO\_BIS introduced the ShuffleAttention attention mechanism into the neck network to enhance the diversity of input sequences so as to improve the detection performance of the model. The original Concat was replaced by the BiFPN weighted bidirectional feature pyramid network, allowing the model to adaptively adjust the fusion mode according to the importance of different features, so as to improve the detection performance. Compared with the traditional YOLO\_BIS V8 model, the performance of YOLO\_BIS model has improved in all aspects. mAP@0.5 has increased by 1.3% to 68.3%, and mAP@0.5-0.95 has increased by 1.2% to 44.6%. The improved model can more accurately identify and detect vehicles and people in foggy scenarios.



(a)Target masking



(b)Dense fog conditions



(c)Proximity

**Figure 6.** Comparison of visualization results

## References

- [1] YANG Lei, CHEN Yanfei, LI Haiming, et al. Target detection algorithm for autopilot scene based on improved YOLOv8 [J/OL]. Computer Engineering and Application,1-17[2024-09-02].
- [2] Su Tong, Wang Ying, Deng Qiyang, et al. The improved pedestrian and vehicle detection algorithm [J / OL]. Journal of System Simulation, 2024:1-11.
- [3] Li Rensi, Shi Yunyu, Liu Xiang, et al. Image object detection based on Double-Head [J]. Liquid Crystal crystal and Display, 2023, 38 (12): 1717-1727.
- [4] Qiang Zhang, Xiaojian Hu. MSFFA-YOLO Network: Multiclass Object Detection for Traffic Investigations in Foggy Weather. IEEE T. Instrumentation and Measurement, 2023, 72: 1-12.
- [5] WEI Liu-Mei, LUO Xue-Mei, KANG Jian. Improved small target detection algorithm for aerial images with YOLOv8 [J/OL]. Computer Engineering and Science,1-13[2024-09-02].
- [6] M. Tan, R. Pang and Q. V. Le, EfficientDet: Scalable and Efficient Object Detection, 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2020, pp. 10778-10787, doi: 10.1109/CVPR42600.2020.01079.
- [7] Liu, Lingzhi. Research on target detection algorithm based on multi-scale attention and dense connected network[D]. Guilin University of Electronic Science and Technology, 2023. DOI:10. 27049/d.cnki.ggldc.2023.001546.
- [8] Q. -L. Zhang and Y. -B. Yang, SA-Net: Shuffle Attention for Deep Convolutional Neural Networks, ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 2021, pp. 2235-2239, doi: 10.1109/ICASSP39728.2021.9414568.
- [9] Yuxi Cheng. Research on traffic sign recognition in foggy weather combining GAN and YOLOv7[D]. Liaoning University of Engineering and Technology, 2023. DOI: 10. 27210/ d.cnki.glnju.2023.000096.
- [10] Zhu Yan, Zhang Yuexia. SEP-YOLO: Improved road target detection algorithm based on YOLOv8[J/OL]. Computer Applications and Software,1-8[2024-09-02].