

Lightweight Fatigue Driving Detection Based on Improved YOLOv8

Shidong Sun, and Yang Xu*

Abstract—To address the challenges of computational complexity and parameter redundancy in existing fatigue-driving detection methods, we propose a novel approach named SL-YOLOv8 (StarNet-LWDH-Head-YOLOv8). This method integrates the YOLOv8n model with the StarNet backbone network, improving the capacity for driver fatigue detection and feature extraction effectiveness while reducing computational complexity. Furthermore, the Star module is incorporated into the neck network to form the C2f-star, improving feature fusion efficiency while maintaining a low computational cost. 3×3 depthwise separable convolutions are applied to the YOLOv8 detection head to optimize the model further, significantly reducing the number of model parameters. The experimental results demonstrate that the proposed SL-YOLOv8 achieves an average accuracy of 98.2%. Compared to YOLOv8n, it achieves a 53.7% reduction in parameters and a 45.7% reduction in computational cost. These findings highlight that SL-YOLOv8 effectively mitigates computational and parameter burdens while preserving high detection accuracy.

Index Terms—Image processing; Deep learning; Lightweight; Fatigue detection; YOLOv8

I. INTRODUCTION

WITH the rapid advancement of the economy and the transportation industry, the global number of traffic accidents is increasing rapidly. Studies show that fatigued driving is a crucial factor contributing to traffic accidents[1]. In a state of fatigue, a driver's reaction speed and judgment capacities are remarkably diminished, making it difficult to respond promptly to unexpected circumstances. Statistical data reveals that approximately half of traffic accidents are associated with excessive driver fatigue. In our transportation system, fatigue driving is a major causal factor of accidents. Consequently, the identification of the fatigue state of drivers holds substantial significance in curtailing the occurrence rate of traffic accidents and enhancing road safety [2].

Researchers have made remarkable progress in computer vision and target detection using deep learning in recent years, applying it to fatigue driving detection. Traditional target detection methods, such as HOG [3] (Histogram of Orientation Gradients) and DPM [4] (Part Modeling), encounter difficulties in efficiently accomplishing the detection task, exhibiting relatively low overall efficiency and accuracy. For instance, Zhang Zhiwei et al. use the VGG19[5] network to monitor driver distraction behavior; however, the deficiency in localization accuracy under

specific conditions restricts its practical application. In contrast, Xiu-Li Lu et al. combine facial feature point detection with an improved SSD[6] algorithm to determine fatigue status, yet the model is large and challenging to deploy on standard devices. Additionally, two-stage detection algorithms (e.g., R-CNN [7], Fast R-CNN [8], and Faster R-CNN [9]), although capable of enhancing accuracy, necessitate the generation of candidate frames, leading to a limited operation speed. Although each of these one- and two-stage algorithms possesses merits and demerits in fatigue driving detection, they fail to address the core issue comprehensively. The DETR [10] (Detection Transformer) algorithm detects targets through direct prediction; nevertheless, it remains inadequate when handling fatigue driving detection. For this reason, Joseph et al. propose the single-stage YOLO (You-Only-Look-Once) algorithm[11], which can simultaneously predict multiple bounding boxes and their corresponding categories. Compared with traditional two-stage algorithms, YOLO exhibits a significant superiority in processing speed and synchronizes classification and position prediction by transforming the target detection task into a regression problem. Consequently, researchers have extensively applied the YOLO algorithm in driving behavior detection. Murthy et al. utilized YOLOv5s [12] as a detection framework for driver assistance systems, demonstrating better performance than YOLOv3 and YOLOv4. However, it fails to adequately consider the significance of lightweight design in detecting fatigue driving. In contrast, YOLOv7 incorporates the E-ELAN structure to augment the learning capacity of the network and preserve the integrity of the gradient paths[13]. YOLOv8 adopts an anchorless design and introduces the C2f module, further refining the model's feature extraction ability and detection accuracy.

Although the aforementioned algorithms have addressed specific issues in fatigue driving detection, there remains scope for enhancement in optimizing the network structure, especially under diverse lighting conditions, reducing the number of parameters, and alleviating the computational burden. This paper proposes a new approach to overcome the limitations of existing algorithms under such circumstances. It proposes a lightweight improved algorithm, SL-YOLOv8, based on the YOLOv8n algorithm to further optimize performance and augment efficiency.

II. RELATED WORK

A. YOLOv8n Model

The YOLO series, a classic single-stage target detection approach, has been extensively employed in vision. YOLOv8 has undergone optimization in multiple aspects based on YOLOv5, particularly in detection accuracy and speed,

Manuscript received Dec 26, 2024, revised May 2, 2025. This research was funded by the project of the Liaoning Provincial Department of Education (Project Number: LJKZ0310).

Shidong Sun is a Postgraduate of the University of Science and Technology Liaoning, Anshan, Liaoning, China. (e-mail: ssdong_2001@outlook.com).

Yang Xu* is a Professor at the University of Science and Technology Liaoning, Anshan, Liaoning, China. (corresponding author to provide phone: +086-138-8978-5726; e-mail: xuyang_1981@aliyun.com).

achieving remarkable progress. One of the core novelties of YOLOv8 compared to its predecessors is the incorporation of anchor-free detection, which substantially accelerates the post-processing step of non-maximum suppression (NMS). With its outstanding detection speed and accuracy, YOLOv8 precisely recognizes and localizes targets within images and videos and supports tasks such as image classification and instance segmentation.

In contrast to other YOLO models, YOLOv8 has garnered widespread attention due to its efficient detection performance and rapid response capabilities. YOLOv8n divides its network structure into three main components: the input layer (Input), the backbone network (Backbone), and the neck and head (Neck and head). Moreover, YOLOv8 can furnish five distinct scale versions by diverse practical task requirements, namely N/S/M/L/X [14, 15]. By eliminating anchor frames and directly locating the target centroid, YOLOv8 reduces the utilization of a priori frames and optimizes the processing speed of NMS.

On the other hand, the backbone network comprises a convolutional layer, a C2f module, and a fast pyramid pooling (SPPF) module, which is mainly responsible for the hierarchical extraction of multi-scale features. This segment employs residual connectivity and a bottleneck structure to diminish model complexity and enhance performance [16]. Compared to the C3 module of YOLOv5, the C2f module attains more favorable feature extraction outcomes with fewer parameters. The neck structure encompasses convolution, the C2f module, concatenation (Concat), and an up-sampling module, which effectively integrates shallow and deep feature information through the feature pyramid method (FPN) and path aggregation network (PAN). The neck network of YOLOv8 incorporates the SPPF module, the probabilistic anchor-box assignment (PAA) module, and two PAN modules, and the model subsequently feeds the fused features into three different scales of detection heads. The detection head generates the final target detection results via convolution and deconvolution operations. In addition, the classification head performs the target classification task through global average pooling [17].

B. Improved SL-YOLOv8n Module and StarNet Module

In order to effectively address the issues of excessive parameters and high computational complexity during the process of driver fatigue detection, this paper proposes an improved YOLOv8n target detection algorithm, namely SL-YOLOv8. The network structure diagram of the SL-YOLOv8 target detection algorithm is illustrated in Figure 1. The specific improvements are detailed as follows:

(1)SL-YOLOv8 integrates the backbone network of StarNet with YOLOv8, thereby realizing the mapping of high-dimensional and nonlinear feature space through the star operation. This combination not only averts the escalation of computational complexity but also fortifies the model's capacity to perceive fatigue detection under varying lighting conditions, consequently enhancing the efficacy of feature extraction.

(2)The C2f module of YOLOv8 is enhanced by substituting its neck structure with the Star Block, giving rise to the C2f-Star structure. This modification optimizes

the model architecture and markedly boosts the model's operational efficiency.

(3)By incorporating a lightweight asymmetric detection head, LWDH-Head, which utilizes 3×3 depth-separable convolution, the number of parameters of the model is substantially diminished, and the computational efficiency is enhanced.

The backbone network of YOLOv8n is composed of multiple convolutional and pooling layers. Compared to YOLOv5, YOLOv8n incorporates the C2f module[18], which supplants the original C3 structure. The C2f module curtails the number of parameters via optimized design while simultaneously enhancing the feature extraction ability. This module mitigates redundant parameters and augments computational efficiency, enabling the model to maintain high performance while remaining lightweight.

To further augment the feature extraction capacity and reduce the parameter count of the model, this paper proposes using StarNet instead of the original network architecture of YOLOv8n. StarNet maps the high-dimensional and nonlinear feature space through star operations without augmenting computational complexity, which optimizes the feature extraction outcome. Additionally, incorporating StarNet augments the model's adaptability within complex environments and facilitates more precise fatigue detection. This enhancement empowers SL-YOLOv8 to strike a superior balance between performance and efficiency, furnishing robust support for the practical applications of driver fatigue monitoring. StarNet streamlines the model representation in a single-layer neural network by integrating the weight matrix and the bias term as a unified entity. We denote this integration as $W = \begin{bmatrix} W \\ B \end{bmatrix}$, W represents the weight component, and B designates the bias term. Therefore, the input vector X is extended into a new input matrix to incorporate a constant term (typically assuming the value of 1), denoted as $X = \begin{bmatrix} X \\ 1 \end{bmatrix}$. On this basis, StarNet specifically implements the star operation $W_1^T X * W_2^T X$. In the single-input single-output scenario, we define the weight matrices W_1 and W_2 and the inputs $X \in \mathbb{R}^{(d+1) \times 1}$.

Where d represents the number of input channels, this approach can be readily extended to multiple output channels by simply adjusting the dimensions of the weight matrices W_1 and W_2 accordingly. Through this construction, StarNet not only enhances the computational efficiency of the model but also streamlines the processing of the deviation term. To sum up, we can express the star operation as follows:

$$\begin{aligned} \omega_1^T x * \omega_2^T x &= \left(\sum_{i=1}^{d+1} \omega_1^i x^i \right) * \left(\sum_{j=1}^{d+1} \omega_2^j x^j \right) \\ &= \sum_{i=1}^{d+1} \sum_{j=1}^{d+1} \omega_1^i \omega_2^j x^i x^j \\ &= \sum_{(i,j)} \alpha_{(i,j)} x^i x^j \end{aligned} \quad (1)$$

$$\alpha_{(i,j)} = \begin{cases} \omega_1^i \omega_2^j & \text{if } i = j, \\ \omega_1^i \omega_2^j + \omega_1^j \omega_2^i & \text{if } i \neq j. \end{cases} \quad (2)$$

Here, i and j denote the different channels, while ω denotes the coefficient of each element. The star operation can eventually be generalized to $\frac{(d+2)(d+1)}{2}$ portfolios As shown in equation (1). By stacking the multilayer structure, the dimensions of the hidden layers can grow recursively, approaching infinity. Assuming the width of the initial network layer is d , once generated, a star operation operation is performed $\sum_{i=1}^{d+1} \sum_{j=1}^{d+1} \omega_1^i \omega_2^j x^i x^j$. This leads to an implicit representation of the feature space $R\left(\frac{d}{\sqrt{2}}\right)^{2^1}$. In this way, we denote the star operation of the n th iteration by S_n .

$$\begin{cases} S_1 = \sum_{i=1}^{d+1} \sum_{j=1}^{d+1} \omega_{(1,1)}^i \omega_{(1,2)}^j x^i x^j & \in R\left(\frac{d}{\sqrt{2}}\right)^{2^1} \\ S_2 = \mathbf{W}_{2,1}^T S_1 * \mathbf{W}_{2,2}^T S_1 & \in R\left(\frac{d}{\sqrt{2}}\right)^{2^2} \\ S_3 = \mathbf{W}_{3,1}^T S_2 * \mathbf{W}_{3,2}^T S_2 & \in R\left(\frac{d}{\sqrt{2}}\right)^{2^3} \\ \vdots & \\ S_n = \mathbf{W}_{n,1}^T S_{n-1} * \mathbf{W}_{n,2}^T S_{n-1} & \in R\left(\frac{d}{\sqrt{2}}\right)^{2^n} \end{cases} \quad (3)$$

As the number of model iterations increases, the model becomes increasingly adept at capturing complex and abstract features, leading to a significant improvement in overall performance. By stacking multiple layers of star operations, the potential feature dimensions grow exponentially. This paper, building upon this principle, proposes a new method., this paper introduces a Star Block, which is further developed into a Star Network (StarNet),

as illustrated in Figure 2. The proposed model features a simple yet efficient structure that leverages the advantages of star operations to enhance real-time performance while maintaining high accuracy, showcasing considerable potential for practical applications. As the depth of the neural network augments and the number of feature map channels escalates. However, the network performance is enhanced. It concomitantly leads to the accumulation of redundant information and substantially elevates the computational cost. This issue is also conspicuously prominent in the YOLO series of models. This paper proposes a new method to enhance the efficiency of fatigue driving detection. This paper proposes StarNet, a lightweight backbone network that is designed to effectively diminish the computational and parameter magnitudes. The overall model performance is further optimized by substituting the backbone network of YOLOv8n with StarNet.

C. C2f-star module

The C2f module represents a crucial constituent in YOLOv8, and its design stems from enhancements made to the C3 module in YOLOv5, intending to maintain a lightweight architecture while augmenting the richness of gradient flow information. In the YOLOv8n model, the C2f module augments the overall network performance by integrating multiple bottleneck structures. Nevertheless, redundant or irrelevant feature information may be incorporated due to the excessive accumulation of bottleneck structures within this module. Such redundancy not only augments the computational burden of the model but may also exert a negative impact on the recognition accuracy.

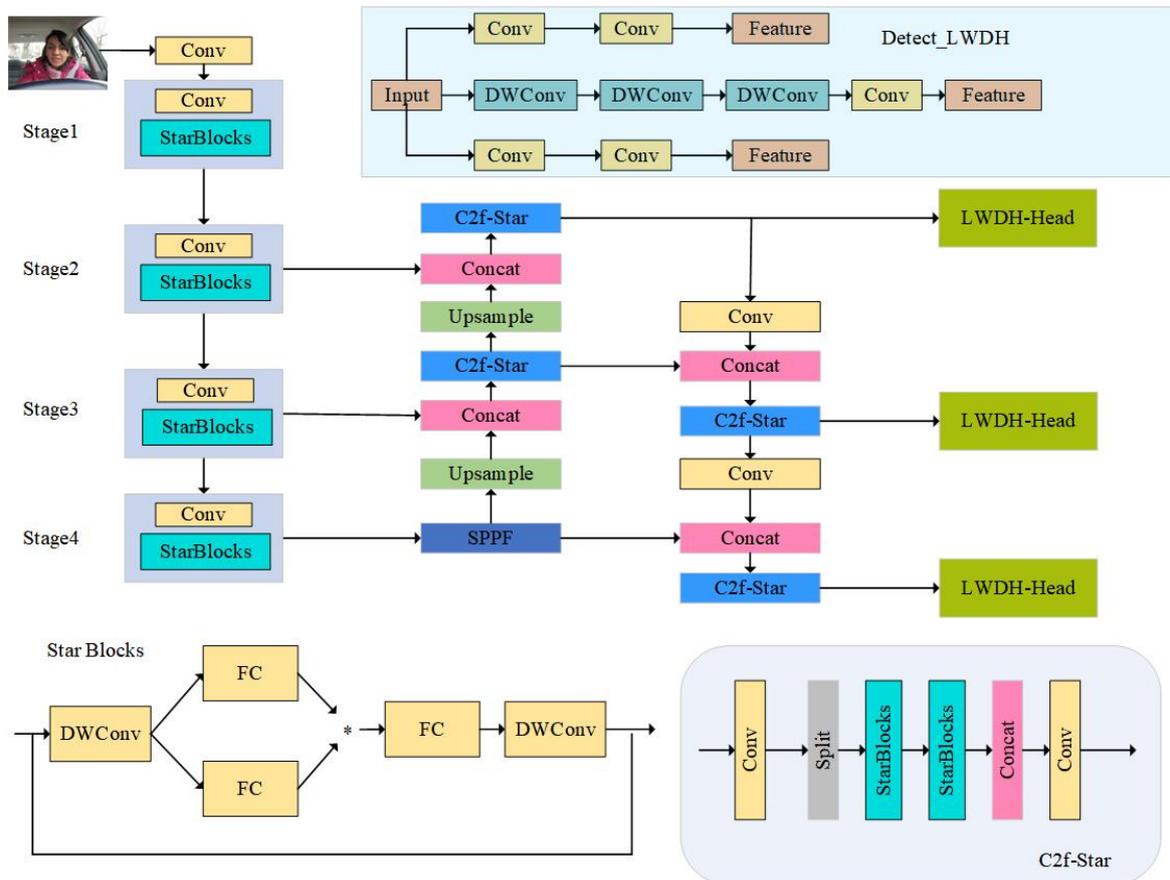


Fig. 1. SL-YOLOv8 Structure Diagram

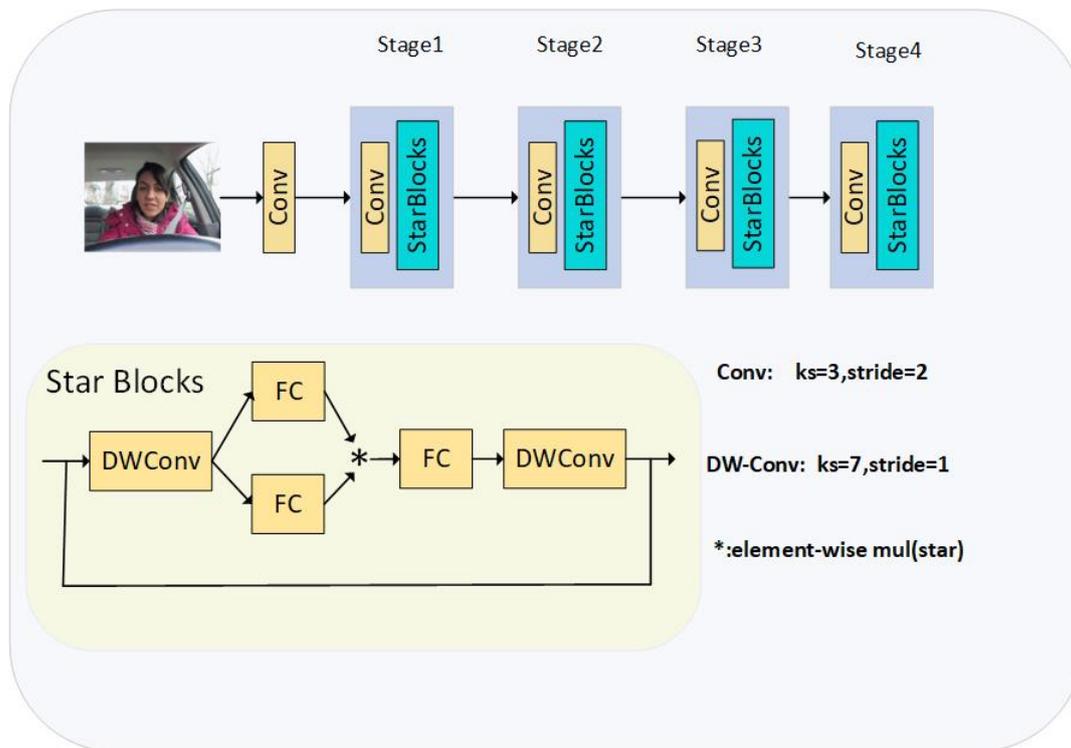


Fig. 2. StarNet Structure Diagram

For this reason, this paper introduces the star-shaped block as a substitute for the bottleneck structure. The star-shaped block exhibits functionality comparable to that of the bottleneck structure at a reduced computational cost and can attain equivalent or superior recognition outcomes. Consequently, replacing the traditional bottleneck structure in the C2f module with star-shaped blocks optimizes the performance and computational efficiency of the model. Figure 3 illustrates the C2f-star structure. Based on this, by introducing the star-shaped block, the C2f module not only effectively diminishes the computational complexity but also enhances the model's efficiency in exploiting the feature information. This improvement renders the model more adaptable in diverse application scenarios, particularly in resource-constrained environments. Additionally, the structural design of star-shaped blocks streamlines the network architecture. It facilitates a smoother gradient flow during the backpropagation process, accelerating the training process and enhancing the convergence speed. These advantages endow SL-YOLOv8 with enhanced competitiveness and practicality in practical applications regarding performance and computational efficiency. Figure 3 depicts the structure of C2f-star.

The star block's processing flow is initiated by conducting a convolution operation on the input data. Subsequently, the data is processed through two fully connected layers, one of which applies a ReLU activation function. Thereafter, the outputs of these layers are integrated through the star operation. Eventually, the merged results are combined with the initial input data, and an additional convolution operation is carried out to produce the final output.

The C2f-Star module utilizes star operations in the star block to remove the traditional bottleneck structure, thus effectively reducing redundant computation and model size. The star block can extract high-dimensional features from

low-dimensional input space, which significantly improves the extraction efficiency of fatigue driving features. This feature makes the model show higher efficiency and practicability in practical applications.

D. Lightweight Detection Head LWDH Module

In YOLO series models, the detection head generally comprises three branches, each of which processes information from different scales of the same object. However, conventionally, these branches operate independently, which may result in inefficient utilization of model parameters and an elevated risk of overfitting. To tackle this problem, this paper substitutes the standard 3×3 convolution in each branch with 3×3 depthwise separable convolutions (DWConv) [19]. In contrast to traditional convolutions, DWConv remarkably reduces the number of parameters by disassembling the convolution operation into depthwise convolution and pointwise convolution. During the depthwise convolution stage, the kernel acts on each channel independently, ensuring that the number of feature maps remains equal to the number of input channels without expanding the dimensionality of the feature maps. Subsequently, pointwise convolution, via 1×1 convolution, integrates information across different channels, modifying the number of channels while maintaining the spatial dimensions of the feature maps. By decreasing the number of parameters, depthwise separable convolutions enhance computational efficiency, and pointwise convolutions effectively integrate cross-channel information, retaining the model's feature extraction capabilities. Figure 4 illustrates the structure of the LWDH-Head module.

This method further diminishes the model parameters and augments the detection speed. Excessive coupling between tasks is effectively circumvented by introducing three 3×3 depth-separable convolutional layers for the classification

task and the bounding box regression task, respectively. After substituting the head structure in the YOLOv8 network with the LWDH-Head, the number of parameters of the model is markedly reduced while the stability of the detection accuracy is maintained, signifying that the optimization strategy curtails the model complexity without impairing the detection performance.

III. EXPERIMENT

A. Experimental environment

The operating system used for the experiments in this paper is Win11, using Python3.8, Cuda11.8, and Pytorch2.0 as the development environment and deep learning framework. The graphics card is NVIDIA GeForce RTX 3090(24GB), and the CPU is 15 vCPU Intel(R) Xeon(R) Platinum 8362 CPU @ 2.80GHz. YOLOv8n was used as the baseline model, the input image size was 640×640, the batch size (BatchSize) was 32, the initial learning rate was 0.01, and 300 rounds of iterative training were performed (Epoch).

B. Dataset

In order to validate the feasibility of this algorithm, it is essential to acquire an appropriate dataset. During the dataset collection process, this paper organizes the datasets under factors such as age, gender, and lighting. Therefore, this experiment will gather images from the following datasets:

(1) YAWDD dataset[20]: YawDD is a dataset specifically designed for driver fatigue detection. It captures images of male and female drivers by installing a camera beneath the rearview mirror of a vehicle.

(2) CEW dataset: The CEW dataset was collected and produced by Nanjing University of Aeronautics and Astronautics (NUAA), focusing on closed-eye detection. It encompasses facial images of individuals of diverse genders, ages, and races when their eyes are closed.

(3) Self-constructed dataset: This dataset comprises images of drivers in the eye-open, eye-closed, and yawning states under varying lighting conditions.

In this paper, the three datasets are integrated and named FDD (Fatigue Driving Detection), amounting to 8019 images. These images are annotated using a labeling tool to categorize them into four classes: eyes open, eyes closed, mouth open, and mouth shut. The processed files are saved in JPG format, annotated in XML format, and converted to TXT format before the experiment. Figure 5 presents some of the images within the dataset.

C. Evaluation Metrics

The model evaluation indexes in this paper are Precision P, Recall R, and Mean Average Precision mAP. The main calculation formula is as follows:

$$P = \frac{TP}{TP + FP} \quad (4)$$

$$R = \frac{TP}{TP + FN} \quad (5)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (6)$$

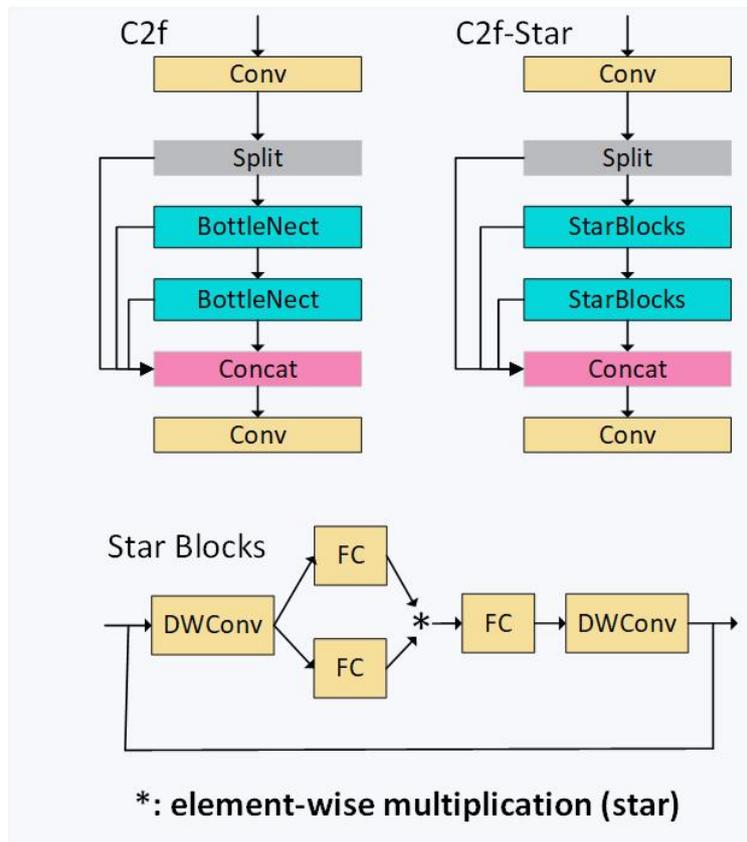


Fig. 3. C2f and C2f-star module

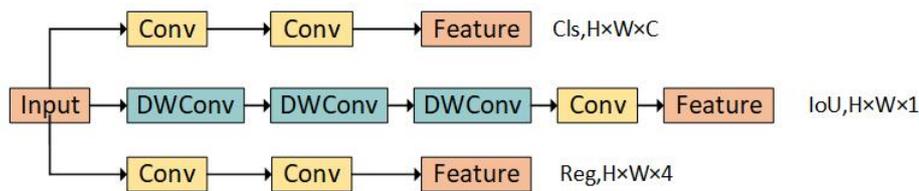


Fig. 4. LWDH-Head module



Fig. 5. Partial images in the dataset

TP is the correctly identified target in the picture, FP refers to the category that recognizes the target's location but incorrectly identifies the target, FN represents an error in forecasting results. N is the category of representation, AP_i is the area under each class PR curve.

In addition, this paper employs other metrics, such as the number of parameters, computational volume, and model weight size, to compare with other mainstream target detection models, thereby comprehensively demonstrating the improved model's superior detection performance. The number of parameters represents the total count of all parameters within the model and reflects the complexity of the network. Generally, a more significant number of parameters endows the model with greater expressiveness, yet it also entails higher computational and storage demands.

The computational volume is typically quantified in FLOPs (Floating Point Operations Counts), which signifies the computational cost the model requires during inference. A smaller computational volume usually implies a faster model inference. The weight size, on the other hand, pertains to the storage capacity of the model file. Models with smaller weights exhibit higher efficiency during training and streamline the deployment process, primarily when utilized on resource-constrained devices[21].

D. Backbone Network Comparison Experiment

In order to more intuitively assess the performance superiority of substituting the YOLOv8 backbone network with StarNet, this paper undertakes a series of backbone network comparison experiments. We compare several networks that have gained popularity in recent years as performance benchmarks. We conduct these experiments within the same environment and on the FDD dataset, and present the experimental results in Table I.

In the experiments, this paper systematically compares the performance of multiple network architectures within YOLOv8. The experimental results indicate that the incorporation of EfficientViT[22] and MobileNetV4[23] fails to enhance the detection accuracy of the model; instead, it leads to a substantial increase in the number of parameters and computational load, augmenting the complexity of the model without yielding the anticipated performance improvement. The average accuracy of Fasternet [24] decreases by 0.4%, accompanied by a significant elevation in its number of parameters and computational complexity. In contrast, HGNetV2 [25] does exhibit a notable reduction in the number of parameters and computational cost, yet its performance advantage over StarNet is not pronounced. Although GhostHGNetV2 [26] performs better in reducing the number of parameters and computational complexity, its average accuracy declines by 0.3%, which fails to enhance the overall model performance effectively. In contrast, StarNet not only surpasses YOLOv8n in terms of the number of parameters and computational complexity but also attains an average accuracy comparable to that of the baseline model. When compared with other network architectures, StarNet manifests significant performance advantages. Therefore, the selection of StarNet as the backbone network of YOLOv8n remarkably improves the perception of fatigue detection and augments the effect of feature extraction, validating its superiority in complex detection tasks.

E. Ablation experiment

In order to verify the effectiveness of each module on the algorithm's performance, this paper conducts ablation experiments under the premise that the configuration of the self-constructed dataset FDD and the environment remains

TABLE I
COMPARISON EXPERIMENT RESULTS OF BACKBONE NETWORK

Model	Precision/%	Recall/%	mAP@0.5/%	Params/M	GFLOPs
YOLOv8n	96.3	96.1	98.3	3.0	8.1
YOLOv8n + EfficientViT	96	96	98.3	4.0	9.4
YOLOv8n + Fasternet	95.8	94.8	97.9	4.2	10.7
YOLOv8n + Mobilenetv4	95.8	96.3	98.3	5.7	22.5
YOLOv8n + HGNetV2	96.7	95.1	98.3	2.4	6.9
YOLOv8n + GhostHGNetV2	96.3	95.1	98	2.3	6.8
YOLOv8n + StarNet	96.6	94.9	98.3	2.2	6.5

TABLE II
RESULTS OF THE ABLATION EXPERIMENT

StarNet	C2f-Star	LWDH	mAP@0.5/%	Params/M	GFLOPs	Size/M
	Baseline		98.3	3.0	8.1	6.0
✓			98.3	2.2	6.5	4.5
	✓		98.1	2.4	6.9	5.1
		✓	98.3	2.4	5.7	4.8
✓	✓		98.1	2.0	6.1	4.1
✓		✓	98.2	1.6	4.5	3.4
	✓	✓	98.0	1.9	4.5	3.9
✓	✓	✓	98.2	1.4	3.7	2.9

consistent. Through these experiments, we observe the trends of different model performance indicators and present the experimental results in Table II.

Following optimization, SL-YOLOv8 significantly reduced both the complexity of the model and the computational load. Specifically, YOLOv8n sequentially incorporated three modules: StarNet, C2f-Star, and LWDH, resulting in reductions of 26.3%, 15.7%, and 20.7% in the number of parameters, respectively. Gleichzeitig, the computational load was reduced by 19.8%, 14.8% and 29.6%. These optimizations effectively reduced the complexity of the model while the detection accuracy remained largely unchanged. More specifically, when we replaced the original YOLOv8 with the optimized model that included the StarNet and LWDH modules, the number of parameters dropped from 3.0M to 1.6M, reducing 45.7%, while the computational load decreased from 8.1G to 4.5G, reducing by 44.4%. After incorporating StarNet, LWDH, and C2f-Star modules, the parameter count of SL-YOLOv8 was reduced to 1.4M, about 46.3% of the parameters of YOLOv8n, while the computational load was reduced to 3.7G, representing 45.6% of the baseline model load. Despite these reductions, SL-YOLOv8 maintained a high detection accuracy, proving that the optimization did not compromise performance. These results demonstrate that SL-YOLOv8, through effective structural optimization, reduced model complexity and enhanced computational efficiency. Furthermore, despite the substantial reduction in both the parameters and the computational load, SL-YOLOv8 still provided a high detection accuracy, further confirming the effectiveness of the proposed optimization strategy. This study provides valuable information for deploying deep learning models in resource-constrained environments and offers an effective solution for reducing model complexity without sacrificing

accuracy.

F. Results and Analysis

To assess the performance of the SL-YOLOv8 model, the aforementioned indicators were used as measures of model performance, and comparisons were made with current standard target detection models under the same experimental environmental conditions. The experimental results are presented in Table 3. First, this study compared the improved SL-YOLOv8 with other models in the YOLO family, such as the newly introduced YOLOv9, YOLOv10, and YOLOv11. Although the detection accuracy of SL-YOLOv8 is slightly lower than that of some lightweight mainstream YOLO models, such as YOLOv6n, YOLOv7-tiny, and YOLOv9t, it exhibits significant advantages in terms of parameter count and computational complexity. Specifically, in resource-constrained environments, SL-YOLOv8, with its reduced computational demands, can effectively enhance real-time processing capabilities and meet the requirements of practical applications. Compared with YOLOv12n, which maintains similar detection accuracy, SL-YOLOv8 offers a reduced parameter count and lower computational complexity, further demonstrating the superior performance of SL-YOLOv8. Compared to the latest RT-DETR-R18 model, SL-YOLOv8 has a precise parameter count and computational complexity advantage and exhibits superior detection accuracy. Compared to Faster-RCNN, SL-YOLOv8 still maintains certain advantages in terms of parameter quantity and computational cost while also achieving better detection performance. In conclusion, although the detection accuracy of SL-YOLOv8 is comparable to that of some mainstream algorithms, its lower parameter count and computational complexity are significantly reduced compared to other models, thereby demonstrating

TABLE III
COMPARATIVE EXPERIMENTAL RESULTS

Model	mAP50/%	mAP50-95/%	Params/M	GFLOPs
yolov3-tiny	98.1	61.4	8.7	12.9
yolov5n	98.0	62.4	1.8	4.1
yolov6n	98.4	61.7	4.6	11.3
yolov7-tiny	98.3	61.2	6.0	13.0
yolov8n	98.3	64.7	3.0	8.1
yolov9t	98.4	64.1	2.6	11.0
yolov10n	98.3	64.6	2.3	6.5
yolov11n	98.3	64.5	2.6	6.3
yolov12n	98.3	64.4	2.6	6.3
Faster-RCNN	83.6	56.5	110.9	370.2
RT-DETR-r18	95.5	60.1	2.0	57.0
SL-YOLOv8 (ours)	98.2	63.6	1.4	3.7

its superiority and potential for practical applications, particularly in tasks requiring high efficiency and real-time performance.

Experimental comparisons reveal that the SL-YOLOv8 model proposed in this paper significantly enhances detection performance and model size compared to mainstream models. This renders SL-YOLOv8 more suitable for fatigue driving detection, especially for installation and deployment on small-scale devices. Figure 6 (a) depicts the detection plot of YOLOv8, and Figure 6 (b) presents the detection plot of the improved SL-YOLOv8 algorithm. As observable in Fig. 6, the improved SL-YOLOv8 algorithm demonstrates only a marginal reduction in detection accuracy while substantially decreasing the number of parameters under normal lighting conditions and in low-light or even dark night environments. This indicates that the improved algorithm can maintain a high level of detection performance while considerably enhancing the efficiency, thereby possessing a

robust practical application value.

IV. CONCLUSIONS

The paper proposes a new algorithm called SL-YOLOv8, which addresses key challenges such as computational complexity and parameter redundancy in fatigue driving detection. The method integrates the YOLOv8n model with the StarNet backbone network, enhancing driver fatigue detection capabilities and feature extraction effectiveness while reducing computational complexity. Furthermore, incorporating the Star module into the neck network to form C2f-star improves feature fusion efficiency without significantly increasing computational cost. Additionally, applying 3x3 depthwise separable convolutions to the YOLOv8 detection head reduces the model's parameters. Experimental results show that SL-YOLOv8 achieves an impressive detection accuracy of 98.2%. Compared to YOLOv8n, it reduces the number of parameters by 53.7%

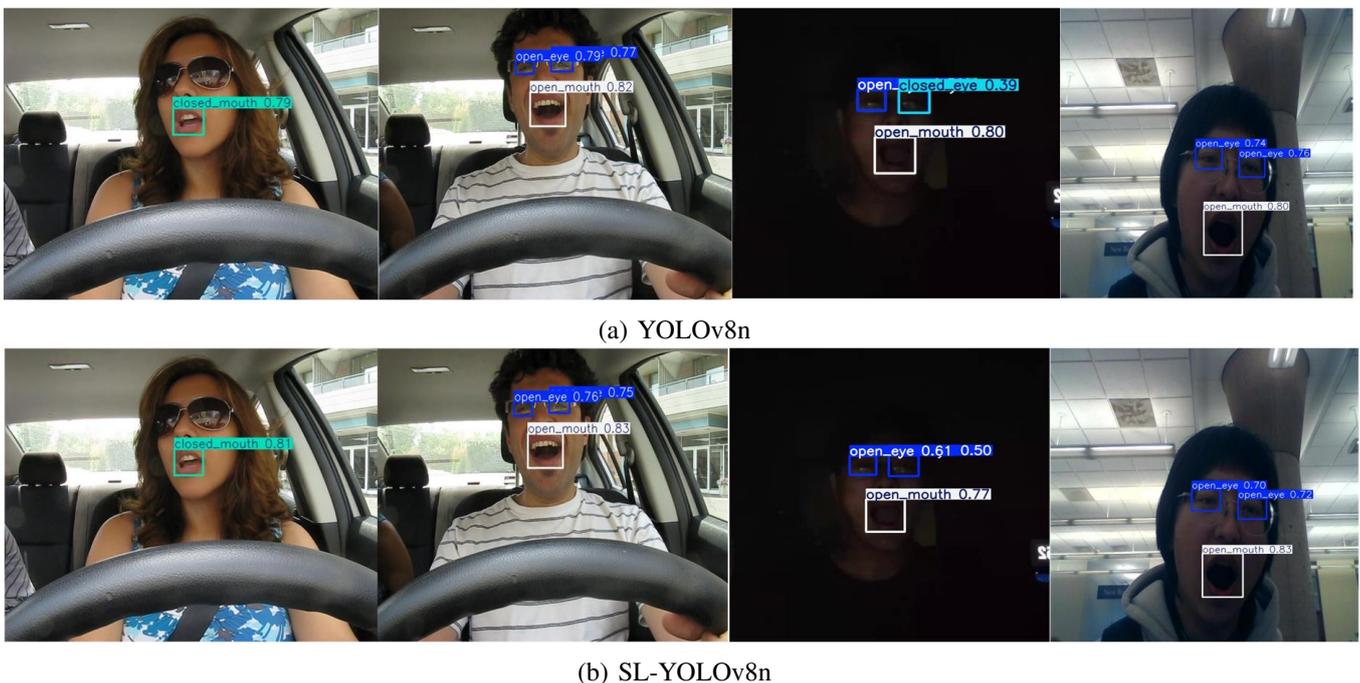


Fig. 6. The visual comparison chart between YOLOv8n and SL-YOLOv8

and the computational load by 45.7%, thus demonstrating significant performance improvements. These findings highlight that SL-YOLOv8 mitigates the computational and parameter burdens of existing models and maintains high detection accuracy, making it a strong candidate for fatigue detection applications. The conclusion suggests that while the SL-YOLOv8 method shows promising results, challenges remain, such as false positives and missed detections in specific environments (e.g., at night or when the driver wears sunglasses). Future research will further enhance the model by integrating multiple datasets to improve detection accuracy in diverse environmental conditions and optimize the network architecture to broaden its applicability.

REFERENCES

- [1] A. Amodio, M. Ermidoro, D. Maggi, S. Formentin, and S. M. Savaresi, "Automatic detection of driver impairment based on pupillary light reflex," *IEEE transactions on intelligent transportation systems*, vol. 20, no. 8, pp. 3038–3048, 2018.
- [2] H. Zeng, X. Li, G. Borghini, Y. Zhao, P. Aricò, G. Di Flumeri, N. Sciaraffa, W. Zakaria, W. Kong, and F. Babiloni, "An eeg-based transfer learning method for cross-subject fatigue mental state prediction," *Sensors*, vol. 21, no. 7, p. 2369, 2021.
- [3] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR*, vol. 1. IEEE, 2001, pp. I–I.
- [4] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, vol. 1. IEEE, 2005, pp. 886–893.
- [5] Z. W. Zhang, "Research on abnormal driving behavior detection method based on machine vision," Master's thesis, Hunan University, Changsha, China, 2020.
- [6] L. X. L., L. X. F., and B. Y. Q., "Research on driving fatigue detection based on ssd multi factor fusion," *Electronic Measurement Technology*, vol. 45, no. 15, pp. 138–143, 2022.
- [7] P. Agrawal, R. Girshick, and J. Malik, "Analyzing the performance of multilayer neural networks for object recognition," *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part VII 13*. Springer, 2014, pp. 329–344.
- [8] X. Wang, A. Shrivastava, and A. Gupta, "A-fast-rcnn: Hard positive generation via adversary for object detection," *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2606–2615.
- [9] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 6, pp. 1137–1149, 2016.
- [10] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, "End-to-end object detection with transformers," *European conference on computer vision*. Springer, 2020, pp. 213–229.
- [11] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "Proceedings of the IEEE conference on computer vision and pattern recognition (cvpr)," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779–788.
- [12] J. S. Murthy, G. Siddesh, W.-C. Lai, B. Parameshachari, S. N. Patil, and K. Hemalatha, "Objectdetect: A real-time object detection framework for advanced driver assistant systems using yolov5," *Wireless Communications and Mobile Computing*, vol. 2022, no. 1, p. 9444360, 2022.
- [13] M. Sun and Y. Tian, "Research on traffic sign object detection algorithm based on deep learning," *Engineering Letters*, vol. 32, no. 8, pp. 1562–1568, 2024.
- [14] J. Terven, D.-M. Córdova-Esparza, and J.-A. Romero-González, "A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas," *Machine Learning and Knowledge Extraction*, vol. 5, no. 4, pp. 1680–1716, 2023.
- [15] G. Yang, J. Wang, Z. Nie, H. Yang, and S. Yu, "A lightweight yolov8 tomato detection algorithm combining feature enhancement and attention," *Agronomy*, vol. 13, no. 7, p. 1824, 2023.
- [16] S. Guo, N. Zhao, X. Ouyang, and Y. Ouyang, "Rbl-yolov8: A lightweight multi-scale detection and recognition method for traffic signs," *Engineering Letters*, vol. 32, no. 11, pp. 2180–2190, 2024.
- [17] H. Li and J. Wu, "Lsod-yolov8s: A lightweight small object detection model based on yolov8 for uav aerial images," *Engineering Letters*, vol. 32, no. 11, pp. 2073–2082, 2024.
- [18] K. Niu and Y. Yan, "A small-object-detection model based on improved yolov8 for uav aerial images," *2023 2nd International Conference on Artificial Intelligence and Intelligent Information Processing (AIIP)*. IEEE, 2023, pp. 57–60.
- [19] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *CoRR*, vol. abs/1704.04861, 2017. [Online]. Available: <http://arxiv.org/abs/1704.04861>
- [20] S. Abtahi, M. Omidyeganeh, S. Shirmohammadi, and B. Hariri, "Yawdd: A yawning detection dataset," *Proceedings of the 5th ACM multimedia systems conference*, 2014, pp. 24–28.
- [21] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*. Springer, 2014, pp. 740–755.
- [22] X. Liu, H. Peng, N. Zheng, Y. Yang, H. Hu, and Y. Yuan, "Efficientvit: Memory efficient vision transformer with cascaded group attention," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 14 420–14 430.
- [23] D. Qin, C. Leichner, M. Delakis, M. Fornoni, S. Luo, F. Yang, W. Wang, C. Banbury, C. Ye, B. Akin *et al.*, "Mobilenetv4: Universal models for the mobile ecosystem," *European Conference on Computer Vision*. Springer, 2025, pp. 78–96.
- [24] J. Chen, S.-h. Kao, H. He, W. Zhuo, S. Wen, C.-H. Lee, and S.-H. G. Chan, "Run, don't walk: chasing higher flops for faster neural networks," *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2023, pp. 12 021–12 031.
- [25] Y. Hong, B.-W. Min, S. Wang, and Y. Hu, "Adaptive spatial feature fusion-based sar ship detection algorithm," *2024 6th International Conference on Data-driven Optimization of Complex Systems (DOCS)*. IEEE, 2024, pp. 837–841.
- [26] X. Zhao, W. Zhang, H. Zhang, C. Zheng, J. Ma, and Z. Zhang, "Itd-yolov8: An infrared target detection model based on yolov8 for unmanned aerial vehicles," *Drones*, vol. 8, no. 4, p. 161, 2024.