

Research on Track Fastener Service Status Detection Based on Improved Yolov4 Model

Jing He¹, Weiqi Wang¹, Nengpu Yang²

¹College of Electrical and Information Engineering, Hunan University of Technology, Zhuzhou, Hunan

²College of Railway Transportation, Hunan University of Technology, Zhuzhou, Hunan

Email: hejing@263.net, 15219038373@163.com, yangnengpu@163.com

How to cite this paper: He, J., Wang, W.Q. and Yang, N.P. (2024) Research on Track Fastener Service Status Detection Based on Improved Yolov4 Model. *Journal of Transportation Technologies*, 14, 212-223. <https://doi.org/10.4236/jtts.2024.142013>

Received: April 1, 2024

Accepted: April 12, 2024

Published: April 15, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

As an important part of railway lines, the healthy service status of track fasteners was very important to ensure the safety of trains. The application of deep learning algorithms was becoming an important method to realize its state detection. However, there was often a deficiency that the detection accuracy and calculation speed of model were difficult to balance, when the traditional deep learning model is used to detect the service state of track fasteners. Targeting this issue, an improved Yolov4 model for detecting the service status of track fasteners was proposed. Firstly, the Mixup data augmentation technology was introduced into Yolov4 model to enhance the generalization ability of model. Secondly, the MobileNet-V2 lightweight network was employed in lieu of the CSPDarknet53 network as the backbone, thereby reducing the number of algorithm parameters and improving the model's computational efficiency. Finally, the SE attention mechanism was incorporated to boost the importance of rail fastener identification by emphasizing relevant image features, ensuring that the network's focus was primarily on the fasteners being inspected. The algorithm achieved both high precision and high speed operation of the rail fastener service state detection, while realizing the lightweight of model. The experimental results revealed that, the MAP value of the rail fastener service state detection algorithm based on the improved Yolov4 model reaches 83.2%, which is 2.83% higher than that of the traditional Yolov4 model, and the calculation speed was improved by 67.39%. Compared with the traditional Yolov4 model, the proposed method achieved the collaborative optimization of detection accuracy and calculation speed.

Keywords

Yolov4 Model, Service Status of Track Fasteners, Detection and Recognition, Data Augmentation, Lightweight Network, Attention Mechanism

1. Introduction

During the manufacturing, transportation and installation of sleepers, it is easy for the pre-embedded rail shoulder to be damaged, resulting in issues such as looseness and falling of track fasteners after a period of operation; Also due to the long-term repeated load imposed on them, the track fasteners can be in abnormal service states such as fracture, displacement and insufficient fastening pressure. Such abnormal service states of these fasteners will seriously affect the operation safety of the train and easily bring potential safety hazards to railway transportation. It is of great significance to detect and identify whether the track fasteners are healthy in service to ensure the safety of railway transportation.

Therefore, establishing advanced detection methods to identify the service status of track fasteners has always been a research hot spot for scholars at home and abroad. For example, Zhao Shanshan *et al.* [1] [2] proposed a fastener detection algorithm based on scale-invariant feature transform (SIFT). It classified the track fasteners by normalizing SIFT features into Fisher vectors of the same length, so that the track fasteners were detected and identified. Wang Qiang *et al.* [3] provided a fastener identification method with improved characteristics of LBP (local binary pattern) operator. It could detection and identify the track fasteners under different weather conditions and achieve better detection stability to a certain extent. Kalal *et al.* [4] carried out a study on the feature recognition of track fasteners by using support vector machine for classification. It extracted the features by integrating the pyramidal gradient direction histogram with macroscopic local binary pattern features. Gibert *et al.* [5] [6] adopted a track fastener detection algorithm based on the Bayesian framework, which effectively improved the accuracy of fastener detection and identification. It is worth mentioning that these methods require a large amount of prior knowledge, their performance of feature extraction and expression for track fasteners is just general, and with to-be-improved identification accuracy.

In recent years, computer vision methods based on deep learning have shown good performance in the field of surface defect recognition [7], therefore attracted scholars' attention and been more and more widely used for the service status detection and identification of track fasteners. For example, Xu Guiyang *et al.* [8] improved the region proposal network in Mask R-CNN with the K-means clustering algorithm, and further proposed a track fastener status detection method based on this Mask R-CNN, which is proved to be with higher detection accuracy; Liu Yuting *et al.* [9] proposed a track fastener service status detection method based on Faster R-CNN with optimized original network operator. It achieved a fastener status identification of high accuracy. However, both of the above methods are two-stage algorithms. They have low false detection rate and high accuracy, however, they are slow in detection and not able to detect in real-time. In order to speed up the training, Gao Jialin *et al.* [10] proposed a railway fastener detection method based on an improved Yolov4 model. By adding an output end and a head structure, it was more effective in detecting

the fasteners, and with increased accuracy and detection rate; Zhang Zening *et al.* [11] applied a fast clip method for track fasteners service status detection based on the SSD convolutional neural network. By using the transfer learning method to pre-train the data set, the training speed of the model was effectively improved. As one-stage algorithms for target detection, the above discussed methods are fast in detection, however, they are easy to be disturbed by irrelevant image information, resulting in reduced detection accuracy. To sum up, both one-stage and two-stage target detection algorithms have achieved initial success in the detection and identification of track fastener service status, but none of them can balance the detection accuracy and calculation speed well.

In response to this problem, this paper conducts an in-depth study on the Yolov4 model [12] and proposes a track fastener service status detection method based on an improved Yolov4 model. In this method, the Mixup data augmentation method is introduced to enrich the data information of the model, as well as to improve the generalization ability of the model; The MobilenetV2 network is used to replace the backbone part of the Yolov4 model, so that model is lightweighted, and the computing ability can be greatly improved; Finally, the SE attention mechanism is introduced to enable the network model to automatically learn the feature weight according to the loss function. In this way, the identification weight of track fasteners is improved and an accurate detection can be realized. By these three alternations, the target detection capability of the traditional Yolov4 model is improved, which is conducive to for the collaborative optimization of the detection accuracy and the calculation speed of track fasteners.

2. Algorithm Principle

2.1. Basic Architecture of Yolov4 Model

As the advanced version of Yolov3 [13], the target detection algorithm Yolov4 mainly consists of three parts: backbone, neck network and Yolo head network. Among them, the backbone adopts the CSP Darknet53 network. The latter is improved from Darknet53 by referencing CSPNet to achieve greater network input resolution, deeper network layers, and more parameters. The neck network takes the SPP module as the additional module, while the PANet as the feature fusion module. The input arbitrary-size feature map is pooled with fixed size, and the features obtained from each pooling are combined to obtain the number of fixed-length features, which are finally input into the fully connected layer for training the network. The PANet is applied to replace FPN for parameter aggregation, so as for the target detection at different levels. Meanwhile, the fusion method is changed from addition to concatenation. The Yolo head network still uses the detection head of the Yolov3 algorithm. Hence, the architecture diagram of Yolov4 is as shown in **Figure 1**.

2.2. Loss Function

In order to guide the Yolov4 network model to autonomously learn relevant

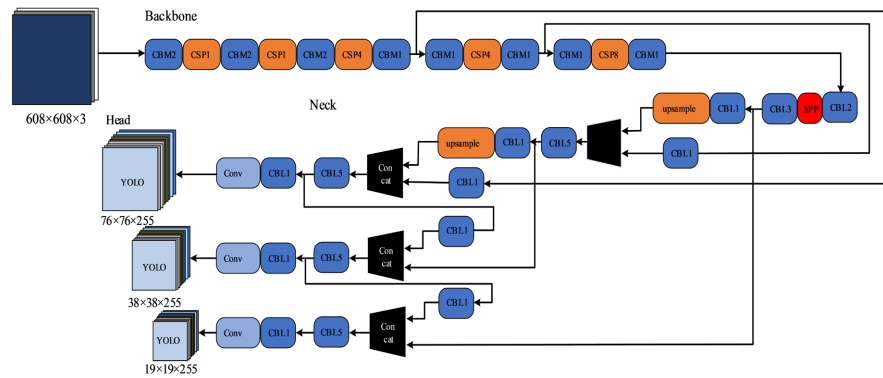


Figure 1. Yolov4 architecture diagram.

features, a joint loss function [14] composed of three parts is designed, *i.e.* bounding box regression prediction error L_{loc} confidence prediction error L_{conf} and classification prediction error L_{cls} as shown in Equation (1).

$$L = L_{loc} + L_{conf} + L_{cls} . \tag{1}$$

The bounding box regression prediction error L_{loc} adopts the CIOU loss function (as shown in Equation (2)). On the basis of IOU, it considers the scale information such as the boundary overlapping score, boundary center distance, and the boundary height ratio, so that the weakness of MSE function caused by using the boundary center distance and the boundary height as independent variables can be made up.

$$L_{loc} = 1 - \eta_{IOU} (M, N) + \frac{\rho^2 (M_{ctr}, N_{ctr})}{m^2} + \alpha v . \tag{2}$$

$$\alpha = \frac{v}{[1 - \eta_{IOU} (M, N)] + v} . \tag{3}$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{\omega^{gt}}{h^{gt}} - \arctan \frac{\omega}{h} \right)^2 . \tag{4}$$

In the formula, α is the trade-off factor; v is the rating factor for detecting the uniformity of the boundary aspect ratio; $\eta_{IOU} (M, N)$ indicates the intersection over union of the predicted box to the actual box; $\rho^2 (M_{ctr}, N_{ctr})$ is the Euclidean distance between the center point of the prediction box and that of the actual box; ω^{gt} and h^{gt} represent the width and height of the actual box, while ω and h are the width and height of the prediction box.

In order to independently evaluate the confidence of the prediction, the model proposed in this paper is constrained with the confidence prediction error shown in Equation (5).

$$L_{conf} = - \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{obj} \left[\overline{C}_i^j \log (C_i^j) + (1 - \overline{C}_i^j) \log (1 - C_i^j) \right] - \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{noobj} \left[\overline{C}_i^j \log (C_i^j) + (1 - \overline{C}_i^j) \log (1 - C_i^j) \right] . \tag{5}$$

where, B is the number of priori boxes in a single grid; S^2 is the number of grids;

I_{ij}^{obj} is an indication that the prediction bounding box contains the target; I_{ij}^{noobj} is an indication that the prediction bounding box does not contain a target; C_i^j is the prediction confidence; \overline{C}_i^j is the actual confidence; λ_{noobj} is the set parameter value.

The classification prediction error of the model is calculated by Equation (6).

$$L_{cls} = -\sum_{i=0}^{S^2} I_{ij}^{obj} \sum_{c \in C'} \left\{ \overline{P}_i^j(c) \log [P_i^j(c)] + [1 - \overline{P}_i^j(c)] \log [1 - P_i^j(c)] \right\} \quad (6)$$

where, c is the number of detection target types; C' represents the total number of defect categories; $P_i^j(c)$ is the prediction probability; $\overline{P}_i^j(c)$ is the actual probability.

3. Improvement of YOLOv4 Algorithm

In order to more effectively detect and identify the service status of track fasteners, this paper introduces data augmentation, network lightweight, and attention mechanism into the Yolov4 architecture for improvement, and a new fastener service status detection algorithm based on the improved Yolov4 model is established.

3.1. Mixup Data Augmentation

Data augmentation can effectively improve the generalization ability of the model, however, the augmentation process depends on the expansion of the data set. The conventional data augmentation does not model and analyze the domain relationship between different samples of different categories. As a simple and data-independent mixed-category augmentation method, the Mixup method uses the beta distribution shown in Equation (7) to (9) to calculate the mixed weight. Its mixing considers the domain relationship of different samples of different categories, so that the expansion of the training data set can be realized.

$$\lambda = \text{Beat}(\alpha, \beta). \quad (7)$$

$$\tilde{x} = \lambda x_i + (1 - \lambda) x_j. \quad (8)$$

$$\tilde{y} = \lambda y_i + (1 - \lambda) y_j. \quad (9)$$

where, *Beat* refers to the beta distribution; λ is the mixing weight calculated from the beta distribution with parameters α and β ; \tilde{x} and \tilde{y} are the mixed samples and the labels corresponding to them, respectively; x_i and y_i are the original sample and the corresponding labels, respectively.

Referencing the experimental settings in Reference [15], this paper conducts an experimental analysis of the value of the hyperparameter α . As shown in **Figure 2**, when α is 0.5, the test performance of the model is optimal; In addition, when adjusting parameters for multiple groups of experiments, the weight λ will be randomly generated in each batch of samples, and the expectation value of the weight in the N batches of samples throughout the training process is ap-

proximately 0.5, thus, no matter how the values of α and β are set, the expectation value of $\alpha/(\alpha + \beta)$ is always approximately 0.5. In consequence, the beta distribution hyperparameter of the mixing weight λ in this paper is set as $\alpha = \beta = 0.5$, so that the algorithm performance can be relatively optimal.

3.2. Lightweighting of MobileNet-V2 Network

In order to improve the computational efficiency of the model, the backbone network structure of the Yolov4 model is chosen to be MobileNet-V2 network instead of the original CSPDarknet53 (as shown in Figure 3) to lightweight the model network.

The CSPDarknet53 network structure used in the traditional Yolov4 model unfolds the input image data into a 3×3 ordinary convolution calculation, performs

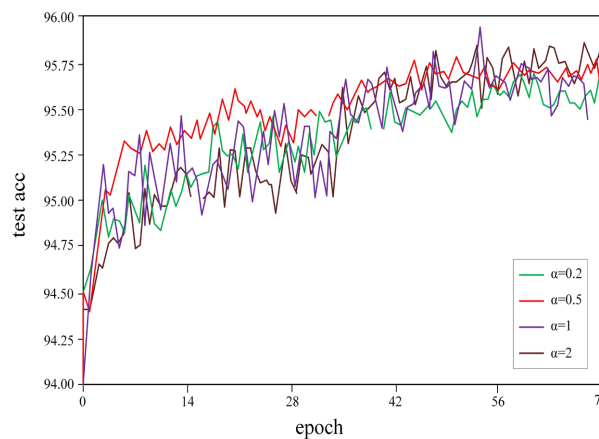


Figure 2. Test accuracy with different α .

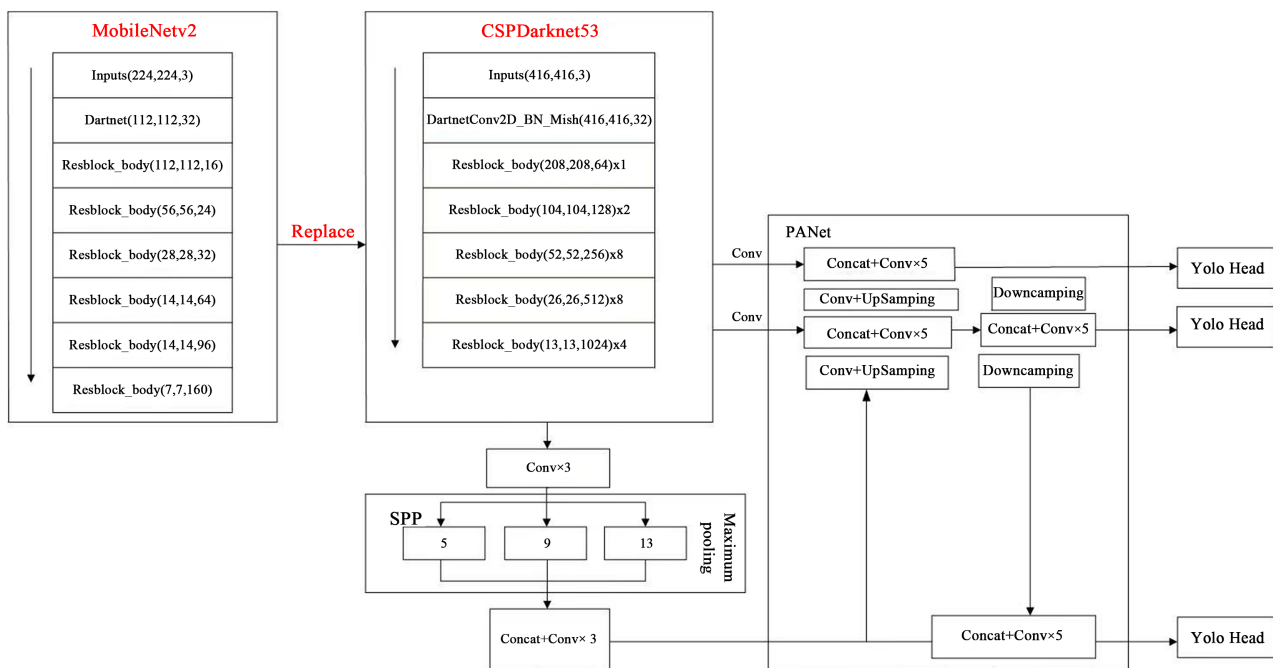


Figure 3. Lightweight network structure.

channel compression, and then inputs it into the Resblock1 module. The ResBlock1 module downsamples the input feature map first, and then conducts 1×1 ordinary convolution to divide the output feature map into two branches. The feature map of one branch is input to the residual convolution block for further unfolding and calculation, after which a 1×1 ordinary convolution is performed to integrate the channel features. Finally, the two output feature maps are stacked on the channel dimension. After stacking, an ordinary convolution is performed to fuse the channel information to obtain the final feature map output.

The MobileNet-V2 network introduces an inverted residual bottleneck structure based on a depthwise separable convolution “DW-Conv” into the backbone part (as shown in Figure 4). For the input, 1×1 ordinary convolution is first conducted to increase the dimensionality, then a DW convolution in a 3×3 depthwise separable convolution is used to extract the image features. Finally, a 1×1 ordinary convolution is performed to reduce dimensionality to obtain feature map output, so as to lightweight the model network and effectively reduce model calculation.

In these convolution operations, the calculation amount of ordinary convolution FLOPs_N is shown in Equation (10), and the calculation amount of depthwise separable convolution FLOPs_{DW} is shown in Equation (11).

$$FLOPs_N = \{[C_i \times K_w \times K_h] + [C_i \times K_w \times K_h - 1] + 1\} \times W \times H \times C_o . \quad (10)$$

$$FLOPs_{DW} = K_w \times K_h \times C_i \times W \times H . \quad (11)$$

where, C_i and C_o are the number of input and output channels respectively, K_w and K_h are the width and height of the input image data, while W and H are the width and height of the convolution kernel.

From the above equations, it can be known that when CSPDarknet53 and MobileNet-V2 are used as the backbone network for unfolding respectively to process the same image of size $(224 \times 224 \times 3)$, the model calculation amount is 716.37 M and 121.26 M accordingly. It means that with the same input, the calculation amount of the model lightweighted by MobileNet-V2 network is only about 1/6 of that of the original Yolov4 model. The model computational efficiency is significantly improved.

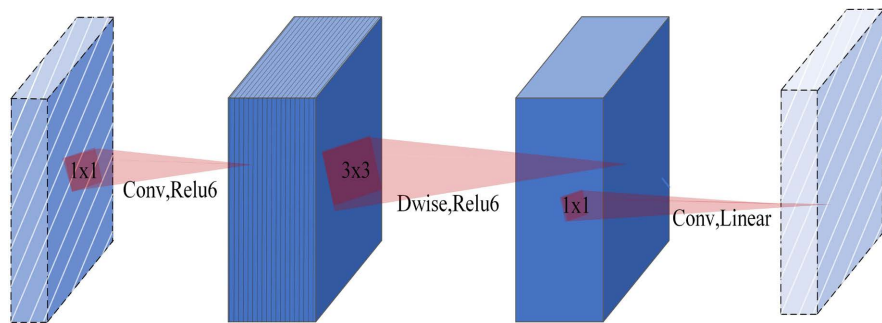


Figure 4. Inverted residual bottleneck structure.

3.3. SE Attention Mechanism

While the model network is lightweighted by MobileNet-V2, a certain degree of reduction in detection accuracy is inevitably caused. In order to compensate the descend of accuracy, the SE Block module is also introduced into the proposed model. By exploring the weighted mapping relationship between the feature channel and the feature map in the convolutional layer, different weights are given to different positions of the image from the perspective of the channel domain, so as to highlight the feature information of track fasteners, and to realize the adaptive attention weighting of the model. In this way, the model training can further focus on valid information such as the status of the track fastener, leading to accuracy improvement.

The SE Block module is mainly composed of Squeeze, Excitation and Scale. Its calculation process is shown in Figure 5.

In this calculation process, the convolution operation F_{tr} that organizes the input features is first carried out on the input feature map “X” to generate the feature map “U”; At the same time, the Squeeze (F_{sq}) operation is used to perform global average pooling on the feature map “U” to generate a $1 \times 1 \times C$ vector to obtain the weight feature z indicating the channel importance information; Each element z_c in z is expressed as Equation (12).

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j). \tag{12}$$

where, F_{sq} represents the global average pooling operation.

Then the Excitation (F_{ex}) is operated. By parameter learning through two fully connected layer, the corresponding weight s_c of each feature channel is updated, and a weight array s with a dimension of $1 \times 1 \times C$ is obtained, as shown in Equation (13)

$$s = F_{ex}(z, W) = \sigma(W_2 \delta(W_1 z)). \tag{13}$$

where, F_{ex} represents the excitation operation; δ and σ indicate Relu and Sigmoid activation functions, respectively; $W_1 \in \mathbb{R}^{\frac{C}{r} \times C}$, $W_2 \in \mathbb{R}^{C \times \frac{C}{r}}$, and r is the scaling factor.

Finally, the dot product operation Scale (F_{scale}) is performed for the one-dimensional weight array s obtained after the excitation operation by combining and the feature map “U”, so that the feature channel is weighted:

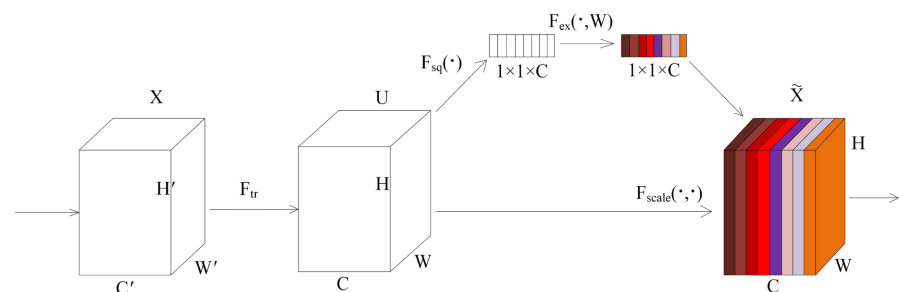


Figure 5. SE Block calculation flow chart.

$$\tilde{X} = F_{scale}(U, s) = U \cdot s, \tilde{X} \in \mathbb{R}^{H \times W \times C}. \quad (14)$$

where, F_{scale} represents the dot product operation; \tilde{X} is output of SE Block module with weighted attention.

Based on this attention mechanism of attention weighted output, the network training of the proposed model can focus on the state characteristics of track fasteners. Other information regarding such as tracks and sleepers can be effectively filtered out. In this way, the detection accuracy of the algorithm is improved.

4. Experiment Results and Analysis

This paper proposes a track fastener service state detection method based on an improved Yolov4 model by introducing in Mixup data augmentation, MobileNet-V2 network lightweight, and SE attention mechanism. In order to verify the effectiveness of these three alternations, training and tests were carried out on the proposed method with the data collected from the track fasteners. The results were compared with those of traditional Yolov4 model.

4.1. Test Description

4.1.1. Experiment Setup

The operating environment of the experiment simulation in this paper was Pytorch, with a configuration as: 1) CPU: Intel I7-8700, 3.2 GHz; 2) operating memory: 16 GB; 3) GPU: NVIDIA GeForce RTX3080, video memory 12 GB, computing power 8.6; 4) code operating environment: Torch = 1.9, Python = 3.7; 5) CUDA version: CUDA11.1. The optimizer adopted SGD, with a weight decay 0.0005, a momentum coefficient 0.9. The learning rate decay mode was cosine decay, the batch-size was 10, and the number of training iterations (Epoch) was 70.

4.1.2. Brief Introduction of Data Set

The data sets of track fasteners used in this experiment were from the track inspection car of Zhuzhou CRRC Times Electric Co., Ltd.. They were collected by CCD camera on a metro line in a city with a unified standard. A total 3334 images were included. The image data format was PNG, and the resolution of each image was 3072×1024 . The data is divided into a training set, a validation set and a test set according to the ratio 8:1:1. The training set was used to train the model and fit the weight parameters; The validation set was used for adjusting the model parameters to obtain the optimal model; While the test set was used to test the final model and evaluate the final output.

The open source software Labelmg was utilized for data annotation to labeling the status and position information of track fasteners in the image data. **Figure 6** demonstrates the service status data of some track fasteners.

4.2. Experiment Results and Analysis

First, model training and validation were carried out with the training and validation data sets. This is to verify the detection effectiveness of the improved Yolov4

model proposed in this paper for the service state of track fasteners. The calculation curve of the model is shown in **Figure 7**. It can be seen that the loss values of both the training set and the validation set converge to 0.36 in advance and the final convergence values are basically the same. This indicates that the parameter settings and prediction results of the track fastening detection model proposed in this paper are basically reasonable.

Then, with the test data set, the traditional Yolov4 model and the improved Yolov4 model proposed in this paper were respectively used to detection the fastener service state, and the results were compared. **Table 1** gives the result comparison of two models based on the test data set.

According to the experiment results, compared with the traditional Yolov4 model, the improved Yolov4 model with Mixup data augmentation, MobileNet-V2

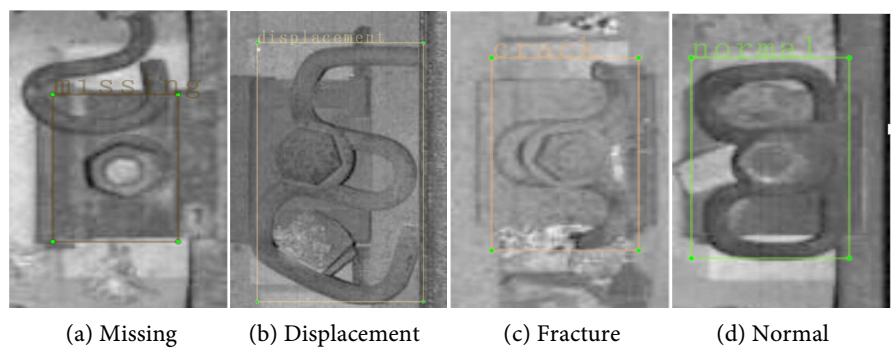


Figure 6. Presentation of partial data.

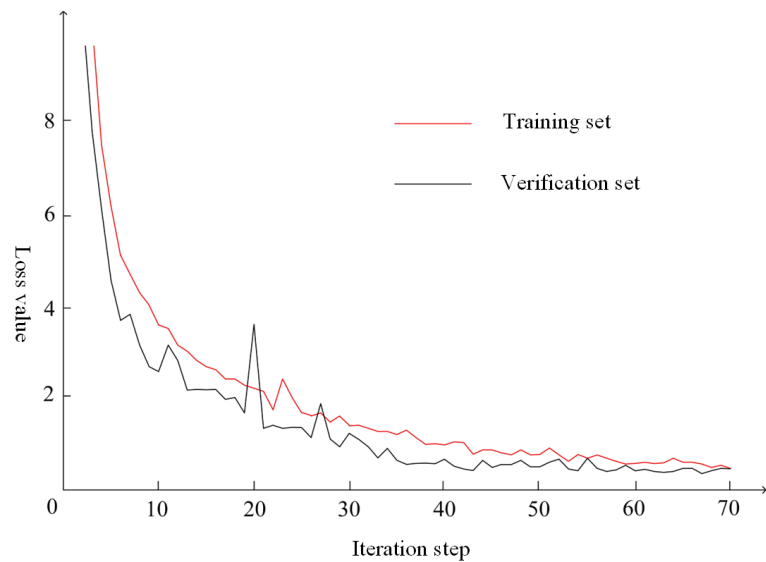


Figure 7. Loss ratio.

Table 1. Performance comparison of models.

Model	Mixup	Mobilenet-V2	SE Block	MAP (%)	FPS (Hz)
Yolov4				80.37	46
Proposed algorithm	√	√	√	83.20	77

network lightweight, and SE attention mechanism has a higher detection accuracy (MAP), which is increased by 2.83%, as well as has a faster processing speed. The processing frame rate (FPS) increased by 31 frames per second, *i.e.* the model calculation speed is increased by 67.39%.

5. Conclusion

Aiming to solve the issue in the current service state detection methods for track fasteners, *i.e.* the detection accuracy and the calculation speed are not well balanced, this paper proposes a new detection method based on improved the Yolov4 model for a collaborative optimization of detection accuracy and calculation speed. By introducing in the Mixup data augmentation, the generalization ability of the Yolov4 model is enhanced; By using the MobileNet-V2 network as the backbone network instead of CSPDarknet53, the model is lightweighted; Finally, the SE attention mechanism is introduced, so that the detection of the proposed model focuses on the service state characteristics of track fasteners, *i.e.* the detection accuracy of the model is further improved. The comparison of experiment results of the traditional and the improved Yolov4 models verifies that the detection algorithm proposed in this paper provides a certain substantial improvement in detection accuracy and detection speed, achieving a collaborative optimization of model detection accuracy and calculation speed.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Zhao, S.S., He, N. and Cao, S. (2018) Railway Fastener State Detection Algorithm Based on SIFT Feature. *Transducer and Microsystem Technologies*, **37**, 148-150+154. [https://doi.org/10.13873/j.1000-9787\(2018\)11-0148-03](https://doi.org/10.13873/j.1000-9787(2018)11-0148-03)
- [2] Lowe, D.G. (2004) Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, **60**, 91-110. <https://doi.org/10.1023/B:VISI.0000029664.99615.94>
- [3] Wang, Q., Li, B.L., Hou, Y., *et al.* (2018) An Improved LBP Feature for Rail Fastener Identification. *Journal of Southwest Jiaotong University*, **53**, 893-899. <https://doi.org/10.3969/j.issn.0258-2724.2018.05.003>
- [4] Kalal, Z., Mikolajczyk, K. and Matas, J. (2012) Tracking Learning Detection. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **24**, 1409-1422. <https://doi.org/10.1109/TPAMI.2011.239>
- [5] Gibert, X., Patel, V.M. and Chellappa, R. (2017) Deep Multitask Learning for Railway Track Inspection. *IEEE Transactions on Intelligent Transportation Systems*, **18**, 153-164. <https://doi.org/10.1109/TITS.2016.2568758>
- [6] Gibert, X., Patel, V.M. and Chellappa, R. (2015) Sequential Score Adaptation with Extreme Value Theory for Robust Railway Track Inspection. *IEEE International Conference on Computer Vision Workshop (ICCVW)*, Santiago, 7-13 December 2015, 131-138. <https://doi.org/10.1109/ICCVW.2015.27>

-
- [7] Niu, S.L., Li, B., Wang, X.G., *et al.* (2020) Defect Image Sample Generation with GAN for Improving Defect Recognition. *IEEE Transactions on Automation Science and Engineering*, **17**, 1611-1622. <https://doi.org/10.1109/TASE.2020.2967415>
- [8] Xu, G.Y., Li, J.Y., Bai, T.B., *et al.* (2022) Detection Method of Track Fastener State Based on Improved Mask R-CNN. *China Railway Science*, **43**, 44-51. <https://doi.org/10.3969/j.issn.1001-4632.2022.01.06>
- [9] Liu, Y.T., Zhang, T., Wang, X., *et al.* (2020) Research on Railway Fastener Status Detection Using Faster R-CNN. *Journal of Dalian Minzu University*, **22**, 202-207. <https://doi.org/10.13744/j.cnki.cn21-1431/g4.2020.03.003>
- [10] Gao, J.L., Bai, T.B., Yao, D.C., *et al.* (2022) Detection of Track Fastener Based on Improved YOLOv4 Algorithm. *Science Technology and Engineering*, **22**, 2872-2877. <https://doi.org/10.3969/j.issn.1671-1815.2022.07.042>
- [11] Zhang, Z.N., Zheng, S.B. and Li, L.M. (2021) SSD-Based Track Fastener Rapid Spring Status Detection. *Computer & Digital Engineering*, **49**, 1560-1565. <https://doi.org/10.3969/j.issn.1672-9722.2021.08.011>
- [12] Bochkovskiy, A., Wang, C.Y. and Liao, H.Y.M. (2020) YOLOv4: Optimal Speed and Accuracy of Object Detection. <https://doi.org/10.48550/arXiv.2004.10934>
- [13] Wang, J., Luo, L.F., Ye, W., *et al.* (2020) A Defect-Detection Method of Split Pins in the Catenary Fastening Devices of High-Speed Railway Based on Deep Learning. *IEEE Transactions on Instrumentation and Measurement*, **69**, 9517-9525. <https://doi.org/10.1109/TIM.2020.3006324>
- [14] Chen, Q. and Xiong, Q. (2020) Garbage Classification Detection Based on Improved YOLOV4. *Journal of Computer and Communications*, **8**, 285-294. <https://doi.org/10.4236/jcc.2020.812023>
- [15] Zhang, H., Cisse, M., Dauphin, Y.N., *et al.* (2017) Mixup: Beyond Empirical Risk Minimization. arXiv:1710.09412. <https://doi.org/10.48550/arXiv.1710.09412>