End Face Detection of Rebar Based on Improved PP YOLO

Kunning Lai[1], Sibo Huang[1], Han Cui[2], Yong Cheng[3] and Wei Luo[1, *]

[1]Modern Education Technology Center, Huizhou University, Huizhou, China
[2]School of Electronic Information and Electrical Engineering, Huizhou University, Huizhou, China
[3]Beijng Qingtai Data Technology Company Limited, Beijing, China
[*]Corresponding author: luowei@hzu.edu.cn

Abstract: This paper introduces an improved PP-YOLO network method to enhance the accuracy of rebar end identification and counting in construction engineering. The network is optimized for the specific characteristics of rebar images, enhancing its recognition capabilities for rebars of various sizes and shapes. Notably, the network structure introduces new pathways in the 4th and 5th layers, enabling more effective learning and identification of rebar features from low-level characteristics, thus improving overall recognition and counting accuracy. Additionally, an optimized data augmentation strategy, tailored to the unique features of rebar images, replaces the traditional Mixup method. Specific algorithms are introduced to enhance the network's efficiency in learning rebar image characteristics. These improvements led to excellent performance in rebar end face recognition tests, achieving an average precision (AP) of 96.23%, a 1.07% increase compared to the original model. This significant performance improvement confirms the effectiveness of our proposed improvements in practical applications, offering new perspectives for the development of rebar detection technology in the construction industry.

Keywords: Deep Learning; Object Detection; PP-YOLO; Rebar.

1. Introduction

In construction engineering, rebar is a critical supporting material, and its precise use is crucial for efficiency and structural integrity [1][2][3]. Accurate identification and counting of rebar ends are vital for the stability and safety of structures. The varied sizes, shapes, and configurations of rebar in images complicate automated identification and counting. Developing a method for rapid and precise identification and counting of rebar ends is therefore a significant research focus in construction engineering. [4] This paper enhances the PP-YOLO network [5], renowned for its effectiveness in natural image object detection, to better suit rebar end identification and counting. Traditional rebar identification methods, mostly based on manual inspection, are not only inefficient but also require substantial expertise. The incorporation of deep learning, particularly in object detection, has demonstrated exceptional effectiveness and practicality globally. This research aims to improve the accuracy and efficiency of automated rebar end face identification and counting by refining the PP-YOLO network. Major advancements in image processing within deep learning began with the 2012 computer vision competition, where Krizhevsky[6] and others' modification of the AlexNet[7] network marked a significant milestone, achieving unparalleled accuracy in image classification tasks and advancing the use of convolutional neural networks (CNNs) in image recognition. With rapid advancements in hardware technology, more robust and intricate deep learning network architectures have evolved, greatly expanding the scope of computer vision. These developments offer robust technical support for image recognition tasks in specialized areas like rebar end detection and counting, enabling precise identification in complex scenarios. For example, the Region-based Convolutional Neural Network (R-CNN) algorithm [8], proposed by Girshick et al. [9] in 2014 and applied in rebar end face recognition, markedly improved recognition accuracy by segmenting the task into classification and localization. Later algorithms, such as Fast R-CNN [10] and Faster R-CNN, showed more efficient rebar image processing. The introduction of the YOLO series algorithms in 2017 marked a new era for quick and accurate rebar end counting, merging target bounding box localization and classification, thereby significantly boosting detection speed and accuracy. That same year, Lin [11] and others introduced a bottom-up and top-down structure in the Feature Pyramid Network (FPN), enhancing small rebar end face recognition accuracy. In 2019, the introduction of the CornerNet [12] algorithm addressed the inefficiencies of traditional anchor-based methods, contributing to progress in efficient recognition and counting of rebar ends.[13] [14] [15] The PP-YOLO network, developed on Baidu's deep learning framework PaddlePaddle, is particularly apt for complex image object detection.

2. Improved PP-YOLO Network for Rebar End Face Image Detection

The PP-YOLO network is a typical application of a single-stage object detection algorithm, which includes a backbone network, a detection neck (usually a feature pyramid network), and a detection head (for classification and localization). In this study, we made improvements to these parts, especially optimizing them for the identification and counting of rebar ends. Through these improvements, we aim to enhance the efficiency and accuracy of rebar image target detection.

2.1. Improvement of the Detection Neck

Improvement of the Detection Neck: This paper improves the PP-YOLO network, a typical application of a single-stage object detection algorithm, focusing on enhancing performance for rebar end identification and counting. We made targeted improvements to the network's three main parts

- the backbone network, the detection neck (usually a feature pyramid network), and the detection head (for classification and localization) - to enhance the efficiency and accuracy of rebar image target detection. In terms of improvements to the detection neck, for rebar end identification applications, we reconstructed the feature pyramid network, adding lateral connection structures. Moreover, we introduced dilated convolutions to expand the receptive field and capture more extensive contextual information, thus improving the recognition ability of rebar end features. We also integrated a self-attention mechanism[16], especially in the 4th and 5th layer feature mappings (C4 and C5), allowing the network to focus more on the critical parts of the rebar ends and improve feature discrimination. In terms of network architecture adjustment, we retained the basic framework of the last three layers of feature mappings (C3, C4, and C5) in the backbone network and strengthened the feature fusion and information transmission between these layers. Especially between the C4 and C5 layers, the combination of self-attention mechanisms and dilated convolutions effectively transmitted more abundant semantic information from lower to higher layers. We defined the feature mapping of the l-th layer output of the feature pyramid as Pl, and in this paper's experiments, l = 3, 4, 5. When the input image size is W × H, the output feature mapping Pl is:

$$P_l = \frac{W}{2^l} \times \frac{H}{2^l},\ l = 3,4,5 \tag{1}$$

In key applications for rebar end face recognition, we made systematic improvements to the detection neck of the PP-YOLO network, especially by adding dilated convolutions and attention mechanisms to enhance the model's performance. These improvements were based on an analysis of the original PP-YOLO network, identifying its limitations in information transmission: although the network could transmit low-level semantic information to higher levels, it failed to fully utilize the rich spatial features of the lower levels. To address this issue, we introduced dilated convolutions in the C3, C4, and C5 layers of the feature pyramid. This convolution form expands the receptive field by introducing spatial intervals in standard convolution kernels, maintaining a balance between parameter count and computational complexity, while also enhancing the model's ability to capture spatial information. This is particularly beneficial before the upsampling block, as dilated convolutions help maintain details that may be lost during downsampling. Furthermore, we integrated a self-attention mechanism in the P4 and P5 feature mappings, allowing the network to focus on key areas containing more target information. At the output end of the convolution block, especially at the nodes where upsampling and downsampling information are merged, we also applied attention modules. This design further enhanced feature discrimination and promoted the fine fusion of features at different levels. With these structural improvements, especially through the Downsample Block module, we established a new pathway for effective information transmission from lower to higher layers. This improvement not only enhanced the efficiency of converting low-level spatial features to high-level semantic information but also significantly improved the model's accuracy in rebar end face recognition tasks. Experimental results validated the effectiveness of this design, as shown in Figure 2, where the improved detection neck structure exhibited exceptional performance in rebar end face recognition, indeed improving recognition accuracy.

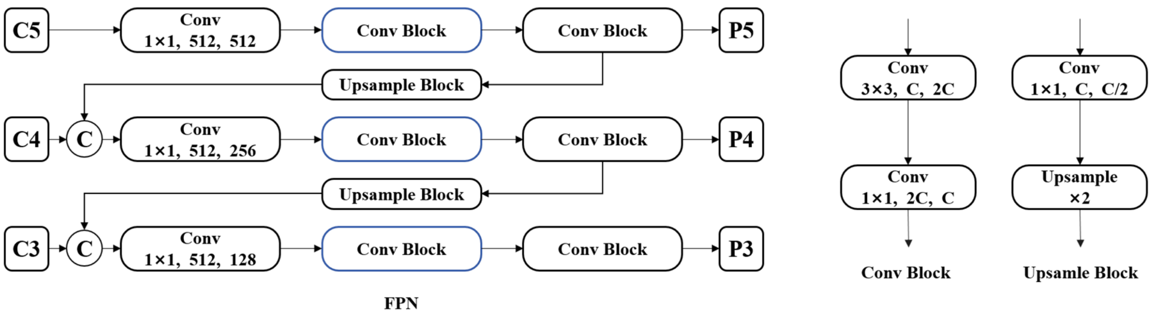

Figure 1. Detailed Structure of the PPYOLO Feature Pyramid

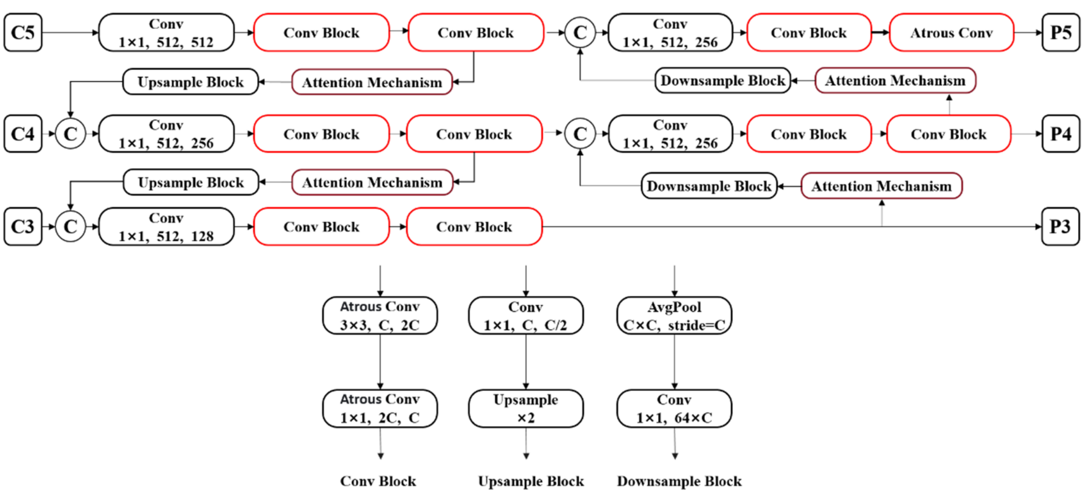

Figure 2. Block diagram of detection neck optimisation strategy

2.2. Detection Head

In the PP-YOLO network for rebar end face image identification, the detection head uses 3×3 and 1×1 convolution layers for final predictions. These convolution layers are meticulously designed to include necessary batch normalization and activation functions, effectively processing the features of rebar end faces. For each input image of rebar ends, the prediction results are associated with specially designed anchor boxes, which are adjusted to fit the unique shape and size of rebars, optimizing the accuracy of predicted probabilities and location coordinates.

To further optimize the accurate identification and precise localization of rebar ends, we employed an improved loss function strategy. The classification task uses cross-entropy loss, but we introduced label smoothing technology to reduce the model's overconfidence in a single category, thereby improving its generalization ability. For localization tasks, we used CIoU loss [18], an IoU-based loss function that considers not only the overlap area between the prediction box and the true box but also the distance and shape of their centers. This is very effective in improving the accuracy of bounding box localization. The calculation formula for CIoU loss is as follows:

$$CIOU\ LOSS = 1 - IOU + \frac{p^2(b,b_{gt})}{c^2} + \alpha \times v \tag{2}$$

Where IoU denotes the intersection over union between the predicted and actual boxes. The term 'pbbgt' refers to the Euclidean distance between the centers of these boxes. 'c' represents the diagonal length of the smallest area enclosing both boxes. The parameter 'α' serves as a weight, while 'v' quantifies the aspect ratio consistency between the predicted and actual boxes.

In the structural design of the detection head, we paid special attention to the final output configuration of each output channel. Considering that rebar end face identification usually involves more specific categories, we adjusted the final output of each output channel to 3×(1 + 4), that is, one rebar end face category plus four positional coordinate parameters. This configuration ensures that the network can effectively classify and locate rebar end faces. We also adjusted the weights of the cross-entropy loss function. This adjustment is based on the characteristics of rebar features, such as the size, shape, and frequency of appearance in images. In this way, we can optimize the classification training effect for rebar ends and improve the model's recognition ability for different categories of rebars. Figure 3 in this paper details the structure of the detection head and its coordinated use with the loss function. Specifically, the introduction of CIoU loss is explained in detail, showing how it optimizes localization tasks. The weight adjustment of the cross-entropy loss function is also elucidated, including its definition and calculation method, as well as examples of how to adjust weights based on rebar features. With these improvements, our PP-YOLO network not only excels in the detection of rebar end faces but also demonstrates high accuracy and efficiency in processing end face images in complex environments. The definition of the cross-entropy loss function is as shown in formula (3).Where weight is the specified weight for each category, and input is the input for the specified category.

$$loss_j = weight[class](-input[class] + \log(\textstyle\sum_{i=0}^{K} \exp(input_i))\ j = 1,2,\ldots,K \tag{3}$$

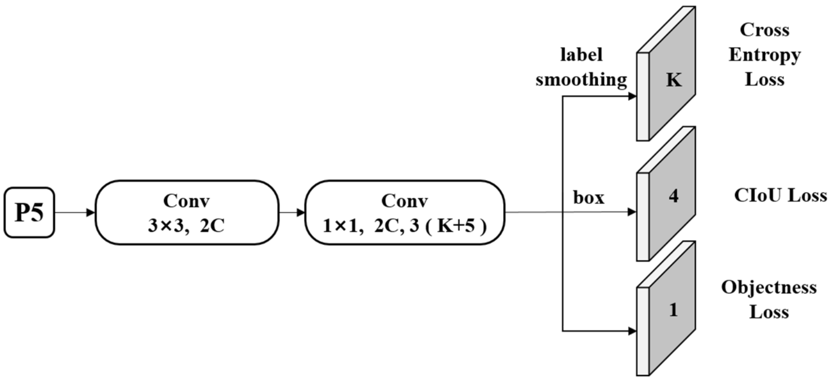

Figure 3. Block diagram of the improved data enhancement strategy

2.3. Improvements in Data Enhancement Optimisation Strategies

In order to further improve the generalisation performance of the rebar end face recognition model and to avoid overfitting phenomena, we optimised our data enhancement strategy. Although initially, our PP-YOLO model used the Image Mixup algorithm in the expectation of enhancing the dataset through weighted mixing between images, we observed that this method may produce undesirable mixing effects on rebar images, which in turn may adversely affect the training efficiency of the model.

To overcome these limitations, we turned to the CutMix algorithm, which generates training samples that are better suited to the characteristics of the rebar by replacing certain portions of the image in order to introduce a more diverse set of image features.CutMix creates a new combination of images by randomly cropping blocks of images and stitching them together with other blocks of images.This combination preserves the key feature information of the rebar while increasing the diversity of backgrounds and environments.

In addition, we introduced a series of direct data enhancement techniques, including basic geometric transformations (e.g., rotation and scaling), lighting and contrast adjustments, noise injection, and blurring and sharpening treatments, all designed to simulate a wide range of conditions likely to be encountered in the real world. With these combined data enhancements, the model is able to recognise rebar end faces in a variety of changing environments, thus improving the robustness and accuracy of the model in real-world application scenarios.

The improved data augmentation process is demonstrated in Fig. 4, which not only effectively enhances the quality and diversity of the training samples, but also ensures that the additional data processing does not significantly increase the computational burden during training and inference. With this well-designed data augmentation scheme, our model is significantly improved in both recognition accuracy and efficiency.

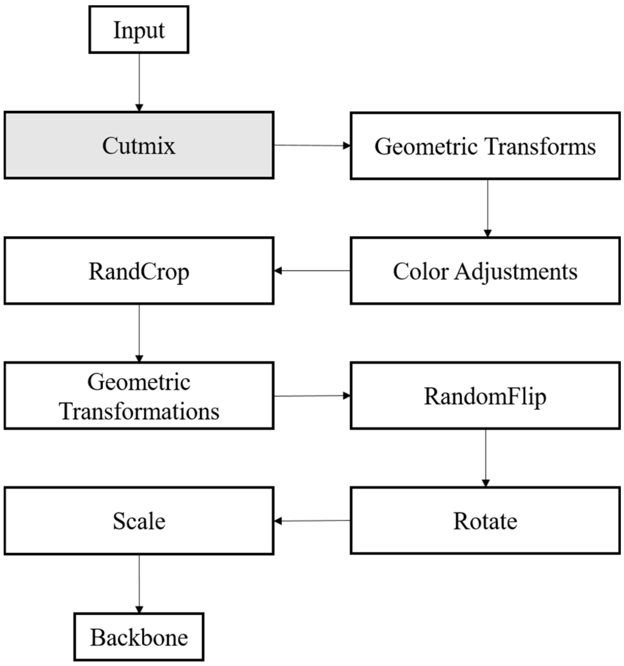

Figure 4. Block diagram of the improved data enhancement strategy

2.4. Optimisation of training strategies

The experimental datasets in this paper focus on rebar endface recognition, and therefore specialised datasets containing a wide range of rebar types and configurations were chosen. Due to the uneven distribution of rebar types in these datasets (e.g., some types have far more rebar images than others), this may result in the model recognising some common types with higher accuracy than less common types during model training. In addition, there may be visual similarities between different types of rebar, which may lead to mutual inhibition of the models during learning. To address these challenges and optimise the training process, we performed specific processing on the dataset. First, we analysed the rebar dataset to pick out difficult-to-recognise rebar types and mixed these with data of common types to achieve a more balanced training set. In addition, we developed a training strategy optimisation procedure to improve the performance of the model in identifying various types of rebar. We used a method of dynamically adjusting the training set by adjusting the composition of the dataset based on the model's performance at different stages, focusing on training those rebar types that the model has poorer recognition performance. In this way, we ensure that the model is able to learn various types of rebar effectively, while avoiding the problem of uneven recognition accuracy due to dataset bias.

3. Experimental Results and Analysis

3.1. Dataset and Experimental Environment

In this study, we primarily utilized a publicly available rebar end face dataset provided by Baidu for our experiments in rebar end face recognition and counting. To enhance the efficacy and accuracy of the experiments, we normalized the dataset and applied standardized sizing to minimize the impact of factors such as lighting and background, ensuring that the images fed into the network had uniform dimensions and proportions. Additionally, to augment the model's ability to recognize rebar at various angles and scales, we employed data augmentation techniques like rotation, scaling, and flipping during the data preprocessing stage. Furthermore, to improve the model's adaptability to variations in image quality in actual engineering environments, we incorporated random noise and blurring processes. In terms of dataset division, we used 80% of the images as the training set for model learning and optimization, while the remaining 20% served as the test set for evaluating the model's performance. The deep learning architecture used for the experiment is Baidu PaddlePaddle2.2.2, Python version 3.7, paddlex version 2.0.0. The running environment is GPU，NVIDIA Tesla V100, with 32GB of video memory, a 4-core CPU, 32GB of CPU RAM, and a 100G hard drive.

3.2. Experimental Parameter Setup and Evaluation Metrics

In the experimental part of this study, we meticulously set the experimental parameters to ensure accuracy and repeatability. The primary objective of the experiment was to compare the performance differences between the improved PP-YOLO network and the classic YOLOv3, as well as the latest YOLOv5 algorithms in the task of rebar end face recognition. To ensure a fair comparison, all models were trained and evaluated in the same hardware and software environments. The main steps of the experimental process can be summarized as follows: initially setting the training model path and evaluation model path, sequentially setting the model's backbone, detection neck, detection head, loss function (learning rate set to 0.0003); then proceeding with training data augmentation and optimization; and finally, setting the parameters for the prediction model.

In the rebar end face recognition experiments for a single category, Average Precision (AP) is used as an evaluation metric since we are only concerned with the recognition accuracy of a specific category.AP is an important performance metric to measure the performance of a model on a single category. In this case, mAP (Mean Average Precision) is not applicable because we only focus on the recognition results of a single category.The formula for AP is:

$$AP = \sum_{k=1}^{N} P(k)\Delta r(k) \quad (4)$$

In this equation, P(k) is the percent detection at the kth threshold, Δr(k) is the amount of change in percent detection at the kth threshold, and N is the number of thresholds. In essence, this formula calculates the area under the entire P-R curve by integrating the check-accuracy and check-completeness at different thresholds. By calculating the AP, we are able to fully evaluate the performance of the model on the rebar end face recognition task. A high AP value indicates that the model performs well in terms of both recognition accuracy and coverage, while a low AP value suggests that the model needs to be further optimised.

Furthermore, in the rebar endface recognition experiments of this study, we pay special attention to the detection time of the model, i.e., the time required for the model to process a single image. This is because in real-world applications, such as construction sites or manufacturing environments, fast and accurate identification of rebar end faces is crucial for improving efficiency and ensuring quality control. The detection time has a direct impact on the responsiveness and real-time performance of the overall system.

3.3. Experimental Results and Analysis

In the experiments on rebar end-face detection, we conducted separate tests for each improvement point of the improved PP-YOLO network proposed in this paper to ensure the accuracy and reliability of the experimental results. These tests were conducted in a separate experimental environment to avoid interference from non-relevant factors. Specifically, the experiments involved a number of different network configurations, including the use of different backbone networks and the detection of improvements to the neck structure.

One of the key configurations in the experiment was to base the improved network on the overall architecture of the ResNet101-vd-dcn backbone network, and this experiment was designated as Method 1. In addition, we explored the optimisation of the training strategy based on the ResNet50-vd-dcn and ResNet101-vd-dcn backbone networks, and these were designated as Method 2 and Method 3, respectively. in particular, the experiment for the optimised detection of neck structures in the ResNet50-vd-dcn backbone network was designated as Method 4. the experiment for detecting neck structures optimised in the ResNet50-vd-dcn backbone network is designated as Method 4. The results of the YOLOV3 ,YOLOV5,PP YOLO model training for this experiment are shown in Figures 8-11. The results of the comparison experiment are shown in Table 1.

The experimental results show that the optimised training strategy significantly improves the performance of the network in rebar end face detection. In particular, when using Method 4, we observed a 1.07% improvement in the AP value compared to Method 1, which demonstrates the effectiveness of the optimised detection neck structure in improving the model accuracy. Although this modification may have slightly increased the number of parameters of the network and may have led to a reduction in the speed of inference, the significant effect in terms of accuracy improvement makes it a worthwhile improvement to adopt. Through comparative analysis, we found that the network, while performing well in learning features for small and very small rebar targets, requires further optimisation when dealing with medium or larger rebar targets. This finding provides valuable guidance on the direction of future improvements to the rebar end-face detection network.

In this rebar end-face detection experiment, a detailed loss function convergence analysis is performed in order to evaluate the training efficiency and stability of different models. This analysis compares the convergence performance of the classic YOLOv3, the latest YOLOv5, and our improved PP-YOLO network on the loss function. As the main objective of optimisation during model training, the convergence speed and stability of the loss function are key indicators to assess the effectiveness of model training. During the experiments, we monitored the change of the loss value of each model throughout the training cycle. By recording the loss values at each iteration step, we were able to observe the performance changes during model learning. The fast and steady decrease of the loss function indicates that the model is able to learn effectively from the training data and has better generalisation ability. The loss function convergence plot for the experimental YOLOV3 ,YOLOV5, PP YOLO model is shown in Figure 5-7.

The experimental results show that the improved PP-YOLO network outperforms YOLOv3 and YOLOv5 in terms of the convergence speed and stability of the loss function.Specifically, PP-YOLO shows a faster loss decline in the early stages of training and maintains low loss fluctuations throughout the training process, which demonstrates the effectiveness of the optimised algorithms and network structure. In contrast, YOLOv3 and YOLOv5, while also demonstrating a decrease in loss values, showed large fluctuations during certain training phases, which may indicate that the training stability of these models in specific scenarios could be improved.

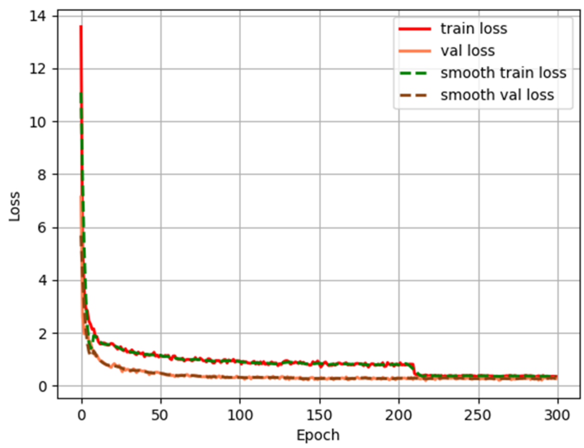

Figure 5. Loss Function Variation Curve for PP-YOLO

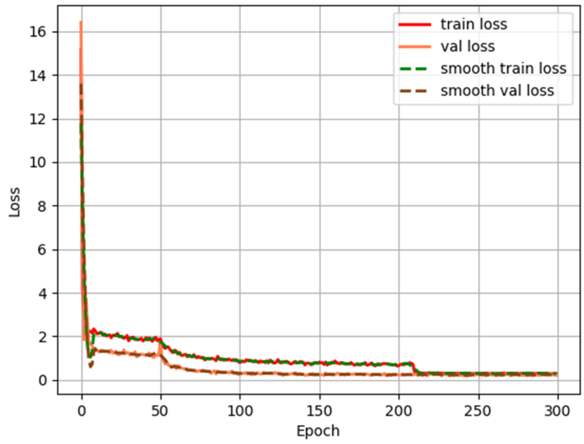

Figure 6. Convergence Curve of the Loss Function for YOLOv3

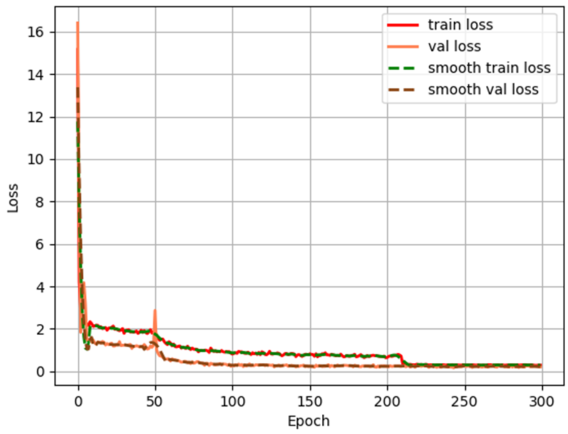

Figure 7. Convergence Curve of the Loss Function for YOLOv5

Figure 8. Visualization of Training Strategy Optimization Results for PP-YOLO

Figure 9. Visualization of Bottleneck Optimization Results for PP-YOLO Detection

Figure 10. Visualization of Unoptimized Results for PP-YOLO

Figure 11. Visualization of Optimization Strategy Results for PP-YOLO

Table 1. Detection Performance of Various Improvement Methods on Public Datasets

Model	Optimization	AP	Detect Time
Yolov3	unoptimized	93.20%	45.4ms
Yolov5	unoptimized	94.23%	44.0ms
PP-YOLO	unoptimized	95.16%	23.3ms
PP-YOLO	Training strategy optimization	96.17%	23.1ms
PP-YOLO	Data augmentation optimization	96.20%	21.2ms
PP-YOLO	Detecting bottleneck optimization	96.23%	22.3ms

4. Conclusions

In this paper, we propose an improved PP-YOLO network method dedicated to the identification and counting of rebar end face images. We perform targeted optimization of the training strategy and the detection of the neck part. The optimization of the training method is based on the characteristics of rebar images, we adopt the CutMix data enhancement algorithm, which is suitable for the characteristics of rebar images, to replace the traditional Mixup data enhancement method, and introduce the GridMask algorithm to enhance the network's ability to learn rebar features. These optimizations enable the model to more fully and evenly learn the various features of rebar end faces, which effectively improves the recognition accuracy of various types of rebars. Our training strategy optimization solves the problems of large size variation and inter-class similarity in rebar images, and achieves a more balanced recognition accuracy of different types of rebars, with a maximum of 96.23% AP. In the improvement of the detection neck part, we retain the basic architecture of the original network, and at the same time, we add the information transfer pathway from the low-layer to the high-layer network in layers 4 and 5. This improvement not only allows the low-level network to learn the feature information of the high-level network, but also enhances the ability of the high-level network to learn the features of the rebar. This optimized method of detecting necks significantly improves the accuracy of rebar end face recognition and ensures that the model has excellent generalization ability in the face of different datasets, and its AP is improved by 1.07% compared to the unimproved PP-YOLO network.

Acknowledgment

This work was supported in part by the ‘QingTai Digital Intelligence Integration’ Collaborative Innovation Project of the Science and Technology Development Center of the Ministry of Education under Grant 2020QT16, in part by the Featured Innovative Projects of Department of Education of Guangdong Province (No.2019KTSCX175, No.2023KTSCX144), and part by the Guangdong Basic and Applied Basic Research Foundation (No.2022A1515140120).

References

[1] Zhao, W. Q., Kong, Z. X., Zhou, Z. D., et al. (2021). Enhanced feature detection of small targets in aerial remote sensing. Journal of China Image and Graphics, 26(3), 644-653.

[2] Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). ImageNet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems (pp. 1097-1105). New York, USA.

[3] Girshick, R., Donahue, J., Darrell, T., et al. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 580-587). Piscataway, USA.

[4] Girshick, R. B. (2015). Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (pp. 1440-1448). Piscataway, USA.

[5] Ren, S., He, K., Girshiek, R. B., et al. (2017). Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(6), 91-99.

[6] He, K., Gkioxari, G., Dollar, P., et al. (2017). Mask R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (pp. 2961-2969). Piscataway, USA.

[7] Joseph, R., Santosh, D., Ross, G., et al. (2016). You only look once: Unified, real-time object detection. In IEEE Conference on Computer Vision and Pattern Recognition (pp. 779-788). Las Vegas, USA.

[8] Wei, Y. H., & Wen, Y. L. (2021). Ratio-and-scale-aware YOLO for pedestrian detection. IEEE Transactions on Image Processing, 30, 934-947.

[9] Venkat, A. A., Huan, C. C., Chi, C. H., et al. (2021). Applying deep learning to defect detection in printed circuit boards via a newest model of you-only-look-once. Mathematical Biosciences and Engineering, 18(4), 4411-4428.

[10] Lin, T. Y., Dollar, P., Girshick, R., et al. (2017). Feature pyramid networks for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2117-2125). Piscataway, USA.

[11] Law, H., & Deng, J. (2019). Cornernet: Detecting objects as paired keypoints [EB/OL]. Available at https://arxiv.org/abs/1808.01244. arXiv preprint arXiv:1808.01244.

[12] Long, X., Deng, K., Wang, G., et al. (2020). PP-YOLO: An effective and efficient implementation of object detector [EB/OL]. Available at https://arxiv.org/abs/2007.12099. arXiv preprint arXiv:2007.12099.

[13] Ma, J., Yao, Z., Xu, C. F., et al. (2021). A real-time tracking algorithm for multiple drones based on improved PP-YOLO and Deep SORT. Computer Applications. doi:10.11772/j.issn.1001-9081.2021071146.

[14] Liu, R., Lehman, J., Molino, P., et al. (2018). An intriguing failing of convolutional neural networks and the coordconv solution [EB/OL]. Available at https://arxiv.org/abs/1807.03247. arXiv preprint arXiv:1807.03247.

[15] Wang, Y., Zhen, P. B., Hou, J. H., et al. (2021). Convolutional neural networks with dynamic regularization. IEEE Transactions on Neural Networks and Learning Systems, 32(5), 2299-2304.

[16] Liu, Z. H., Zhang, Y. D., Chen, Y. Z., et al. (2020). Detection of algorithmically generated domain names using the recurrent convolutional neural network with spatial pyramid pooling. Entropy, 22(9), 1-20.

[17] Yun, S., Han, D., Oh, J. S., et al. (2019). CutMix: Regularization strategy to train strong classifiers with localizable features. In 2019 IEEE/CVF International Conference on Computer Vision (ICCV) (pp. 6022-6031). Piscataway, USA.

[18] Li, K. Q., Chen, Y., Liu, J. C., et al. A review of object detection algorithms based on deep learning. Computer Engineering. doi: 10.19678/j.issn.1000-3428.0062.