ISSN: 2664-9640

Research on the Optimization of Traffic Sign Recognition Technology based on Deep Learning in Higher Vocational Computer Application Majors

Dan Lian

Zhejiang DongFang polytechnic, Wenzhou, Zhejiang, 325000, China

Abstract

This paper takes the computer application major in higher vocational and technical colleges as the research background and focuses on the application of deep learning - related knowledge in the field of traffic sign recognition. The paper selects images from traffic scene environments, constructs relevant object detection models such as YOLO and Faster R - CNN, and trains and optimizes these models. By drawing charts of the model convergence process and comparing the performance indicators of different models, it elaborates on the performance of different models in traffic sign recognition. The research process and the idea of comparative analysis are applied to the computer application major in higher vocational education, hoping to improve students' practical application ability of deep - learning object detection technology and provide references for industry practitioners and researchers.

Keywords

Higher Vocational Computer Application Major; Object Detection Model; Deep Learning; Traffic Sign Recognition.

1. Introduction

With the rapid development of science and technology, intelligent transportation systems are being integrated into people's lives and actual production at an unprecedented speed, bringing great convenience to daily life and work. Intelligent transportation systems can serve intelligent traffic recognition well[1]. Among them, traffic sign recognition is a core element in ensuring road safety and smoothness and is also one of the key links in the entire intelligent transportation system. Accurately recognizing traffic signs can ensure that drivers obtain relevant information in a timely manner and can effectively assist drivers in making correct decisions, thereby reducing traffic accidents and improving traffic efficiency. Although the traffic recognition system is of great importance, traditional traffic recognition methods mainly rely on manual judgment and feature information extraction, combined with machine - learning algorithms for judgment. Affected by various adverse factors such as light changes and weather interference, it is often difficult to make accurate recognition and judgment, and errors may even occur. Obviously, the traditional methods have obvious limitations and innate disadvantages and are difficult to meet the current high - efficiency requirements. Therefore, exploring new technologies to promote intelligent traffic recognition systems is of great significance. In recent years, deep - learning technology has developed rapidly. Many experts and scholars have actively explored its applications, creating favorable conditions for intelligent traffic sign recognition. Deep learning can automatically extract deep - level features from massive image data and has a powerful model fitting ability, which greatly improves the accuracy and stability of the intelligent traffic recognition system[2].

For students majoring in computer applications in higher vocational colleges, actively engaging in research on deep - learning - based traffic sign recognition is of great significance. On the one hand, it can create conditions for students to keep up with cutting - edge industry technologies. On the other hand, through the experiment and operation of actual projects, it can effectively improve students' thinking and practical abilities, enhance their problem - solving skills and experience, and lay a solid foundation for their future participation in related industries and jobs.

2. Introduction to Deep - Learning Object Detection Models

2.1. YOLO Model

The YOLO (You Only Look Once) series of models is an end - to - end object detection algorithm. It regards the object detection task as a regression problem and directly predicts the category and location of objects on the image. The YOLO model has the advantage of fast detection speed and can be applied in scenarios with high real - time requirements. For example, YOLOv5 uses technologies such as CSPNet (Cross Stage Partial Network) in its network structure, which reduces the amount of calculation while enhancing the feature extraction ability. Through multi - scale feature fusion, it can achieve good detection results for traffic signs of different sizes[3].

2.2. Faster R - CNN Model

The Faster R - CNN belongs to the region - based object detection model. Currently, this model mainly consists of components such as the Region Proposal Network (RPN) and the Fast R - CNN detection network. In practical work, the RPN plays a crucial role. It can screen out candidate regions where objects may exist from numerous images, pointing out the direction for subsequent detections[4]. The Fast R - CNN then accurately classifies the candidate regions indicated by the former, determines which specific object category they belong to, and also performs regression on the object's position to determine the precise location of the object in the image. The greatest advantage of this model is its extremely high detection accuracy. It can excel in scenarios with extremely high requirements for detection accuracy, such as high - precision security monitoring and industrial part defect detection. However, it also has certain limitations[5].

3. Collection and Preprocessing of Traffic Sign Image Data

3.1. Data Collection

Students used high - definition cameras to collect images in different traffic scenarios, covering various environments such as urban streets, highways, and rural roads to ensure the diversity of the collected data. The types of traffic signs collected include common types such as prohibition signs, instruction signs, and warning signs, with a total of 5,000 images collected. To ensure data quality, the collected images were screened, and images with blurring, severe occlusion, and other issues that made it impossible to accurately identify traffic signs were removed.

3.2. Data Preprocessing

The collected original images had problems such as uneven illumination and large size differences and needed to be preprocessed. First, image normalization was performed. The pixel values of the images were mapped to the interval [0, 1], and the formula is:

$$x_{norm} = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$$

ISSN: 2664-9640

where x is the original pixel value, and xmin and xmax are the minimum and maximum pixel values in the image, respectively. Second, the images were cropped and scaled to uniformly adjust them to the size required for model input. In addition, to enhance the generalization ability of the model, data augmentation techniques such as rotation, translation, and flipping were used, expanding the dataset to 10,000 images.

4. Construction and Training of Traffic Sign Recognition Models based on Deep Learning

4.1. Model Construction

According to the research requirements, YOLOv5 and Faster R - CNN models were respectively built. When building the YOLOv5 model, based on the characteristics of traffic signs and the scale of the dataset, an appropriate version of the network structure (such as different scales like s, m, l, x) was selected, and parameters were adjusted, such as the number of convolutional kernels in the backbone network and the connection method of the feature fusion layer. For the Faster R - CNN model, a suitable basic convolutional neural network (such as ResNet50) was selected as the feature extraction network, and parameters such as the anchor box size and ratio of the RPN network were set.

4.2. Model Training

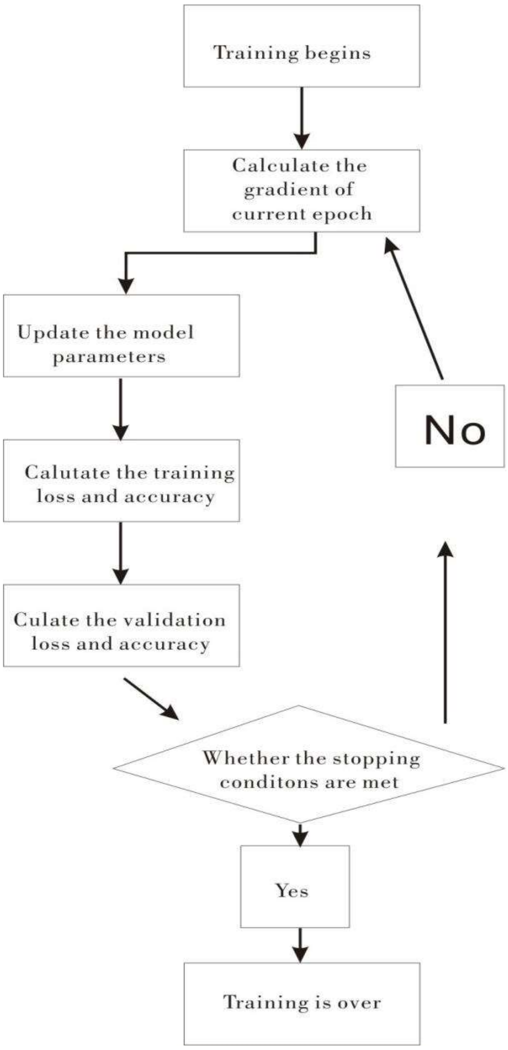

Figure 1. Comparison of Gradient Changes during the Training of YOLOv5 and Faster R-CNN Models

The preprocessed dataset was used to train the built models. During the training process, the cross - entropy loss function was used to measure the difference between the model's prediction results and the true labels. The Adam optimizer was selected, and the learning rate was set to 0.001. The dataset was divided into a training set, a validation set, and a test set in the ratio of 70%:15%:15%. During the training process, after each epoch of training, validation was performed on the validation set. The model parameters were adjusted according to the loss and accuracy on the validation set to prevent overfitting. At the same time, the training loss, validation loss, training accuracy, and validation accuracy of each epoch were recorded, and relevant charts of the model convergence process were drawn, such as the gradient change graph shown in Figure 1.

The training process usually lasts for dozens or even hundreds of epochs. The specific duration depends on the complexity of the model, the scale of the dataset, and the performance of the hardware device. In this study, the training was carried out on a workstation equipped with an NVIDIA GPU. After 50 epochs of training, the accuracy of the validation set of the YOLOv5 model gradually stabilized. Due to its high computational complexity, the Faster R - CNN model took 80 epochs of training to reach a relatively stable state. During the training process, it was found that the training loss of the YOLOv5 model decreased rapidly, while the Faster R - CNN model had a slower decrease in training loss in the early stage but could achieve higher accuracy in the later stage.

In addition, hardware devices have a significant impact on the training efficiency and effect of the model. Using a high - performance GPU can accelerate the calculation and shorten the training time. If the hardware resources are limited, the training time will be greatly extended, and the model may even fail to converge to the optimal solution.

5. Model Evaluation and Result Analysis

5.1. Selection of Evaluation Indicators

Indicators such as accuracy, recall, and mean average precision (mAP) were used to evaluate the trained models. Accuracy reflects the proportion of correctly predicted samples in the total number of samples. Recall measures the proportion of correctly detected target samples in the actual number of target samples. mAP comprehensively considers the detection accuracy of different types of targets and can more comprehensively evaluate the performance of the model.

5.2. Testing in Complex Scenarios

The models were tested in various complex traffic scenarios, including situations such as strong or dim lighting, partial occlusion of traffic signs, and blurred images. For example, in the low - light scenario at dusk, the recall rate of the YOLOv5 model decreased, and some traffic signs with darker colors were likely to be missed. Although the accuracy of the Faster R - CNN model was less affected, its detection speed decreased significantly. When traffic signs were partially occluded by branches, the mAP values of both models decreased. However, the Faster R - CNN model, with its fine - grained feature extraction ability, had a relatively better recognition effect for partially occluded signs.

5.3. Comparative Analysis of Different Models

The performance of the YOLOv5 and Faster R - CNN models on the same test set was compared. In normal traffic scenarios, the YOLOv5 model had a significant advantage in detection speed, being able to process dozens of frames per second. Its accuracy could reach about 85%, the recall rate was 80%, and the mAP value was 0.82. The Faster R - CNN model had a slower detection speed but a high accuracy of 90%, a recall rate of 85%, and an mAP value of 0.88. In complex scenarios, the accuracy advantage of the Faster R - CNN model became more

ISSN: 2664-9640

prominent, but its disadvantage in detection speed was further magnified. Although the YOLOv5 model still had a fast detection speed, its accuracy loss was relatively large.

6. Conclusion

In this study, two deep - learning object detection models, YOLOv5 and Faster R - CNN, for traffic sign recognition were successfully constructed and trained. It was found that both models had their own advantages and disadvantages. In practical applications, the selection should be based on specific requirements. For example, the YOLOv5 model is suitable for real - time recognition scenarios in autonomous driving, while the Faster R - CNN model is more applicable to post - analysis of traffic surveillance videos. Future research can be carried out from the following three aspects: optimizing the existing model structure, combining new architectures and technologies to improve the comprehensive performance in complex scenarios; expanding the scale of the dataset to cover special scenarios and rare traffic signs to enhance the generalization ability; and exploring multi - model fusion to combine the advantages of different models to achieve more efficient and accurate recognition. In addition, the research results can be integrated into the curriculum practice of higher vocational computer application majors to help cultivate high - quality technical and skilled talents who meet industry requirements.

References

[1] Yang Guangzhu, Long Zelian, Li Yi, et al. Research Progress on Traffic Sign Recognition Technology Based on Deep Learning [J]. Western China Communication Science & Technology, 2023,(12):194 - 197. DOI:10.13282/j.cnki.wccst.2023.12.059.

[2] Dong Xiaohua, Wei Yuke. Review of Traffic Sign Recognition Technology Methods [J]. Computer Knowledge and Technology, 2020,16(25):193 - 194 + 199. DOI:10.14004/j.cnki.ckt.2020.3029.

[3] Gao Mingyue, Dong Quande. Research on Traffic Sign Recognition Technology Based on Deep Learning [J]. Journal of Lanzhou University of Arts and Science (Natural Science Edition), 2020,34(03):93 - 97. DOI:10.13804/j.cnki.2095 - 6991.2020.03.019.

[4] Zhong Huijuan. Application Research on Traffic Sign Recognition Technology Based on Feature Reuse [J]. Digital Technology and Application, 2020,38(02):24 - 26. DOI:10.19695/j.cnki.cn12 - 1369.2020.02.14.

[5] Yu Jincheng, Xie Guanghan, Luo Fang. Exploration of Digital Recognition Technology for Road Traffic Signs Based on Deep Learning [J]. Digital Technology and Application, 2013,(12):62 - 63 + 66. DOI:10.19695/j.cnki.cn12 - 1369.2013.12.038.