

Research Article

Railway Subgrade Defect Automatic Recognition Method Based on Improved Faster R-CNN

Xinjun Xu ^{1,2}, Yang Lei ², and Feng Yang ¹

¹*School of Mechanical Electronic & Information Engineering, China University of Mining & Technology, Beijing 100083, China*

²*Infrastructure Inspection Research Institute, China Academy of Railway Sciences, Beijing 100081, China*

Correspondence should be addressed to Xinjun Xu; xinjun_xu@163.com

Received 9 February 2018; Revised 25 April 2018; Accepted 24 May 2018; Published 14 June 2018

Academic Editor: José M. Lanza-Gutiérrez

Copyright © 2018 Xinjun Xu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Railway subgrade defect is the serious threat to train safety. Vehicle-borne GPR method has become the main railway subgrade detection technology with its advantages of rapidness and nondestructiveness. However, due to the large amount of detection data and the variety in defect shape and size, defect recognition is a challenging task. In this work, the method based on deep learning is proposed to recognize defects from the ground penetrating radar (GPR) profile of subgrade detection data. Based on the Faster R-CNN framework, the improvement strategies of feature cascade, adversarial spatial dropout network (ASDN), Soft-NMS, and data augmentation have been integrated to improve recognition accuracy, according to the characteristics of subgrade defects. The experimental results indicates that compared with traditional SVM+HOG method and the baseline Faster R-CNN, the improved model can achieve better performance. The model robustness is demonstrated by a further comparison experiment of various defect types. In addition, the improvements to model performance of each improvement strategy are verified by an ablation experiment of improvement strategies. This paper tries to explore the new thinking for the application of deep learning method in the field of railway subgrade defect recognition.

1. Introduction

Railway subgrade defect is a serious threat to railway transport safety. At present, the main detection equipment of railway subgrade defects is the vehicle-mounted ground penetrating radar (GPR) [1, 2]. In spite of the large amounts of railway subgrade detection data on existing railway, the defect recognition method still primarily relies on the low efficient artificial recognition by experienced experts, which has influenced the further development of railway subgrade detection. Therefore, how to automatically recognize subgrade defects to improve the recognition efficiency and accuracy has been an urgent problem.

At present, the researches on railway subgrade defect recognition primarily focus on the extraction of hand-designed features and the application of traditional machine learning methods such as support vector machine (SVM) and shallow neural network in classifier training. Liao et al. [3] extracted features of mud pumping defects by analyzing layer drawing of GPR data and recognized defects by neural

network technology. Du et al. [4] extracted subband energy, variance, and layer position from subgrade GPR images as features to establish a vector neural network model and used the model to recognize subgrade defects, such as mud pumping, subgrade settlement, and ballast fouling. However, the blurred boundary between the features of subgrade settlement and ballast fouling affected the final recognition rate. Zou et al. [5] used instantaneous frequency and instantaneous amplitude as features and SVM as classifiers to recognize subgrade defects. However, the recognition range was limited to the training set. Du et al. [6] used the sparse representation method to express subgrade defect features and recognized subgrade defects by quarter-sphere SVM.

The essential features of railway subgrade defects cannot be obtained by the traditional way which designs defect features artificially. Therefore, ideal classification results are not achieved by traditional methods. In recent years, deep learning methods [7–9], especially the target detection algorithms [10–13] based on convolutional neural network, have achieved great success in many fields. By extracting hidden

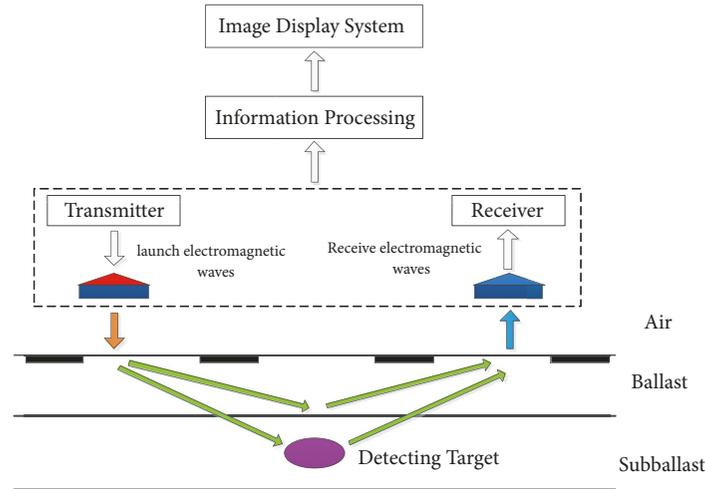


FIGURE 1: The schematic diagram of ground penetrating radar.

high-dimensional features from training data, deep learning methods can eliminate the drawbacks of artificial design features and tedious image processing steps in traditional approaches. Some researchers have tried the deep learning method to solve the problems in railway engineering. Yin et al. [14] proposed an automated diagnosis network of VOB for high-speed train via deep belief network (DBN) and improves the accuracy of fault diagnosis for VOBs to 90–95% in HSRs; Gibert et al. [15] combined multiple detectors within a multitask learning framework to improved accuracy for detecting defects on railway ties and fasteners; Pang et al. [16] recognized the faults of high-speed train bogie based on deep network and the recognition rate is 100% for nonwhole fault of key components of bogie at the different speed; Giben et al. [17] described a novel approach to visual track inspection using material classification and semantic segmentation with Deep Convolutional Neural Networks (DCNN).

In summary, the current researches on railway subgrade defect recognition mainly focus on combining artificial design features with traditional machine learning classification algorithms such as support vector machine and artificial neural network. However, the subgrade defects are various in shape and size, and the artificial design features are limited in the ability of feature representation. Therefore, the artificial design features are difficult to deal with the complicated subgrade environments. In addition, although the deep learning method has been introduced in the railway field, it has not been extended to the subgrade defects recognition. Therefore, different from the traditional methods, this paper uses the deep learning method to extract high-dimensional features and recognize defects from the GPR profile of subgrade detection data.

In this paper, a novel method of applying Faster R-CNN [13] to automatic recognition of railway subgrade defects was proposed. To solve the difficulties in subgrade defect recognition task, the improvement strategies of feature cascade, adversarial spatial dropout network (ASDN) [18], Soft-NMS [19], and data augmentation were integrated to

improve Faster R-CNN. In order to conduct experimental research, a dataset with a total of 2050 labelled subgrade defect profiles was made. In addition, based on the dataset, a large number of experiments including the ablation experiment of improvement strategies, the comparison experiment with traditional approaches and the comparison experiment with various defect types have been conducted to verify the effectiveness of the proposed method. This paper explores the possibility of applying deep learning methods to recognize subgrade defects, and provides a new thinking for the field of railway subgrade defect recognition. The recognition method based on Faster R-CNN improves the recognition accuracy compared with traditional methods.

The rest of this paper is structured as follows. Section 2 introduces the foundation of railway subgrade detection and the establishment of defect dataset. Section 3 describes the method of railway subgrade defect recognition based on Faster R-CNN and the improvement strategies. The experiment and analysis of subgrade defect recognition are discussed in Section 4. Finally, Section 5 presents the conclusions and future work of the paper.

2. Railway Subgrade Detection and Defect Dataset Construction

2.1. Principle of Ground Penetrating Radar. Ground penetrating radar (GPR) [20] is a kind of electronic device that uses high-frequency electromagnetic waves to detect the distribution of subsurface media. The principle of GPR is shown in Figure 1. First, the transmitter emits high-frequency electromagnetic waves through the transmitting antenna, and the electromagnetic wave is reflected when encountering mediums with dielectric differences during the propagation process. Then, the reflected wave is input to the receiver through the receiving antenna. Finally, the host radar records the motion characteristics of the electromagnetic wave and saves it into a file. The cross-sectional scanning image of the underground media is generated by the image display system. According to the arrival time of the reflected signal and the



FIGURE 2: The appearance photo of subgrade status inspection vehicle.



FIGURE 3: Three-channel radar antenna group.



FIGURE 4: FastWave host radar.

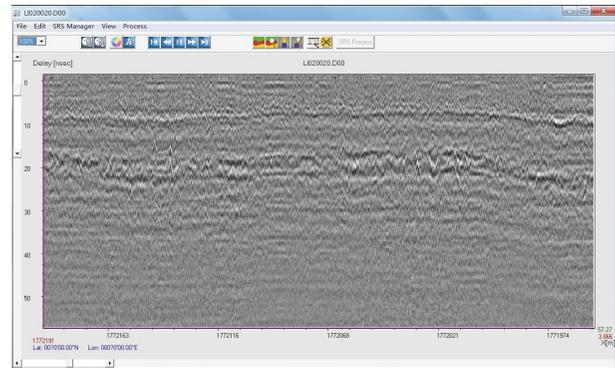


FIGURE 5: The interface of SRS-DPA postprocessing software.

average reflection velocity of target, the distance of detection target can be calculated.

2.2. Vehicle-Borne GPR Detection Technology. The subgrade data used in this paper is collected by the subgrade status inspection vehicle as shown in Figure 2, in which a fast and nondestructive detection technique of vehicle-borne GPR is adopted to detect subgrade. The detection vehicle is equipped with the RIS ground penetrating radar of Italy IDS company to build an acquisition system of railway subgrade data, which consists of three-channel radar antenna group (as shown in Figure 3), FastWave host radar (as shown in Figure 4), central control system, signal display instrument, range finder, signal transmission cable, etc.

When the system is started, the central control system sends a data acquisition instruction to the host radar. Then, the host radar transmits pulse signal through the air-coupled antenna in the radar antenna group and receives the reflected signal of subgrade structure layer simultaneously. The range finder can synchronously records the driving distance of train. Finally, all the data is transmitted to the central control system through the transmission cable, and the subgrade data profile can be viewed in real time through the signal display instrument.

A total of three survey lines are arranged during the detection, which are distributed on the centre of railway and the both sides of rails. Three 400 MHz air-coupled antennas are placed in parallel to collect subgrade data. The time

window of 400MHz antenna is set to 60ns, the number of sampling points is set to 512 points, and the sampling interval is set to 0.115m.

2.3. Radar Data Processing. The railway subgrade data collected by vehicle-borne GPR is unavoidably incorporated into noise signals, due to the interference caused by catenary, high-voltage wire pole, railway signal equipment, locomotives, and so on. In addition, the rails and reinforced concrete sleepers have strong reflections of radar electromagnetic waves which reduce downward radiated energy and affect the quality of inspection data. Therefore, the processing of the original signal can suppress the noise signal and highlight the interlayer interface or the effective characteristic information such as the amplitude, phase, and waveform of the reflected wave of the target defect.

The SRS-DPA postprocessing software is used for data processing and the software interface is shown in Figure 5. The radar signal processing includes preprocessing and further processing analysis. The preprocessing includes data sampling, zero-line correction, and mileage correction. The data processing includes horizontal filtering, vertical band-pass filter, gain setting, speed analysis, and contrast setting.

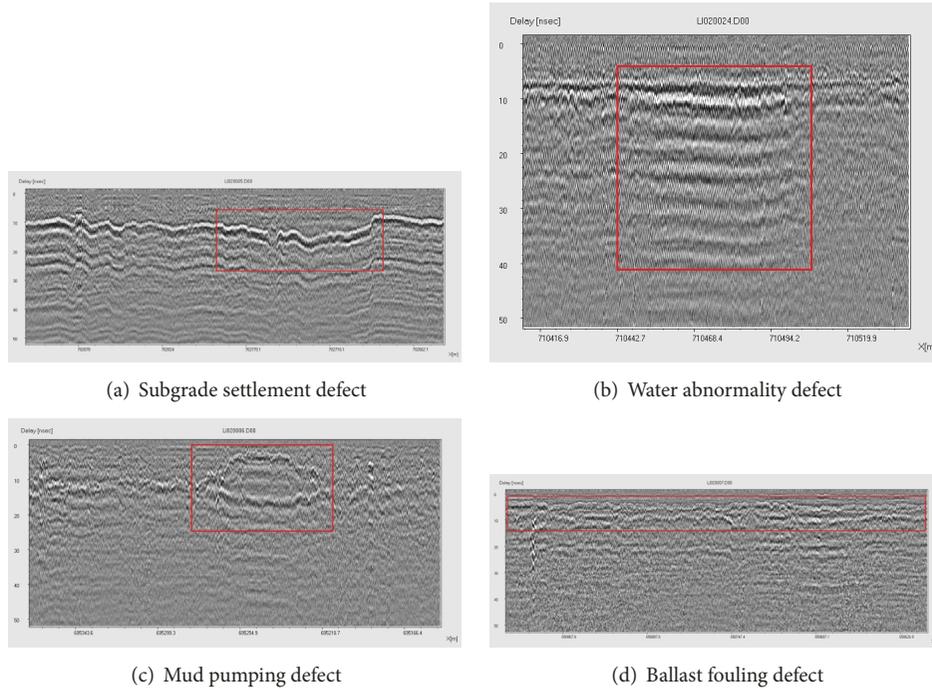


FIGURE 6: The ground penetrating profile of railway subgrade defects.

The processed data are saved as images for the subsequent image annotation.

2.4. Data Interpretation and Defect Labelling. The basis of defect determination in this paper mainly includes the GPR features of railway subgrade defects, on-site excavation verification of defects, and expert interpreting experience.

Firstly, different types of subgrade defects have their own features in GPR profiles. Therefore, different defects such as mud pumping and subgrade settlement can be distinguished and discriminated based on their radar image features, as shown in Figure 6. The radar image features of subgrade settlement are that the reflection events of the interface between the ballast and subgrade bed and the one between surface layer and bottom layer of subgrade bed are obvious bend sinking and depth-down offset, and the events are discontinuous and uneven near the same depth position, as shown in Figure 6(a). The radar image features of water abnormality are that the medium interface is low-frequency and strong reflection, large amplitude, contrary phase, and multiple reflections, as shown in Figure 6(b). The radar image features of mud pumping are that the wave group of defects is disorderly and discontinuous, and the strong low-frequency reflection resembles a mountain tip or straw hat, as shown in Figure 6(c). The radar image features of ballast fouling are that there is a snow shaped or flocculent reflection inside ballast bed, and the reflection events on the interface between ballast bed and subgrade bed are not clear or discontinuous, as shown in Figure 6(d).

Secondly, some of the defects detected have been excavated on site to verify the correctness of defect radar features. The on-site verification images are shown in Figure 7.

Subgrade defects will directly affect traffic safety. For example, mud pumping defect will cause problems such as track offset, uneven rail surface, and hard bending of rails, which will have a serious impact on the entire track structure; subgrade settlement defect could cause recess to make the line uneven; ballast fouling defect will lead to poor drainage and water accumulation, which will infiltrate into the subgrade and cause defects like mud pumping under the influence of train loads; water abnormality defect could cause certain settlement or partial collapse. If the subgrade lacks adequate protection and reinforcement equipment, the subgrade stability will be affected or destroyed.

Thirdly, due to the complex geological structure of railway subgrade, the subgrade defects are various in shape and size. Therefore, the identification of specific defects still needs expert experience. The defect images are labelled with the open source software Labeling [21]. The annotated images are shown in Figure 6. Different defect types are labelled with rectangle frames in different colours. Then, the software will automatically generate the XML index file needed for model training phase, including the position and length-width information of defects in the image. Image annotation is carried out by experienced subgrade detection experts. The labelling rules are that the defects verified by on-site excavation will be preferentially selected; then, the defects with obvious defect GPR features will be labelled; finally, the fuzzy defects that even the experienced experts cannot confirm will be preferentially eliminated.

2.5. Building Defect Dataset. The defect dataset is constructed from the labelled defect images and its data distribution is shown in Table 1. The dataset includes a total of 2050 images,

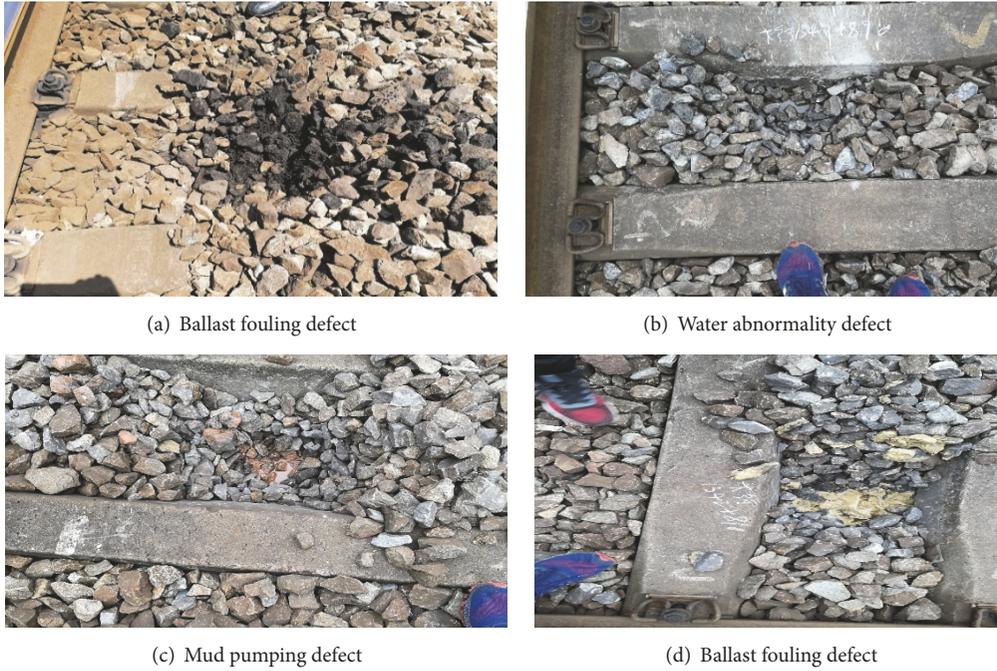


FIGURE 7: The site excavation photos of subgrade defects.

TABLE 1: Data distribution in dataset.

Dataset	Mud pumping	Subgrade settlement	Water abnormality	Ballast fouling	Normal subgrade	Total
Training set	360	280	280	320	400	1640
Validation set	45	35	35	40	50	205
Test set	45	35	35	40	50	205
Total	450	350	350	400	500	2050

in which 500 are normal subgrade images and 1550 are defect images. Among the 1,550 defect images, 450 are mud pumping defects, 350 are subgrade settlement defects, 350 are water abnormality defects, and 400 are ballast fouling defects. The data set is divided into training set, validation set, and test set according to the ratio of 8:1:1. Each set randomly selects images in proportion from the four defect types and normal subgrade to ensure that each set is in the same distribution. Therefore, the training set includes 1640 images, of which 400 are normal subgrade, 360 are mud pumping, 280 are subgrade settlement, 280 are water abnormality, and 320 are ballast fouling; the verification set and the test set, respectively, include 205 images, of which 50 are normal subgrade, 45 are mud pumping, 35 are subgrade settlement, 35 are water abnormality, and 40 are ballast fouling.

3. Subgrade Defect Recognition with Improved Faster R-CNN

3.1. Faster R-CNN. With the accumulation of R-CNN [10] and Fast R-CNN [12], Ross B. Girshick presented a new detection algorithm, Faster R-CNN [13]. On the structure, Faster R-CNN has integrated the four basic steps of target detection into a single deep network, which were feature extraction, proposal generation, bounding box regression,

and classification. Therefore, a significant performance has been achieved, especially on the detection speed. The traditional Selective Search [22] was replaced with the Region Proposal Network (RPN) in Faster R-CNN, a fully convolutional network [23]. Therefore, the Faster R-CNN was composed of two components: RPN and Fast R-CNN. The RPN generated a series of rectangular proposal regions which might include objects and transferred them to the Fast R-CNN; the Fast R-CNN performed target classification and position correction of bounding boxes. The architecture of Faster R-CNN is presented in Figure 8.

3.2. Overview of the Proposed Defect Recognition Method.

In this paper, the Faster R-CNN is applied to the automatic recognition of railway subgrade defects. In view of the difficulties in the detection task of railway subgrade defects, multiple improvements have been made to improve the traditional Faster R-CNN. The specific implementation process is presented in Figure 9.

As shown in Figure 9, using deep learning method to conduct the research of subgrade defect recognition, the development environment needs to be built firstly. In this paper, we build the development environment based on Caffe, which has two model training modes that are CPU and GPU. After the key parameters are set up, we need to choose

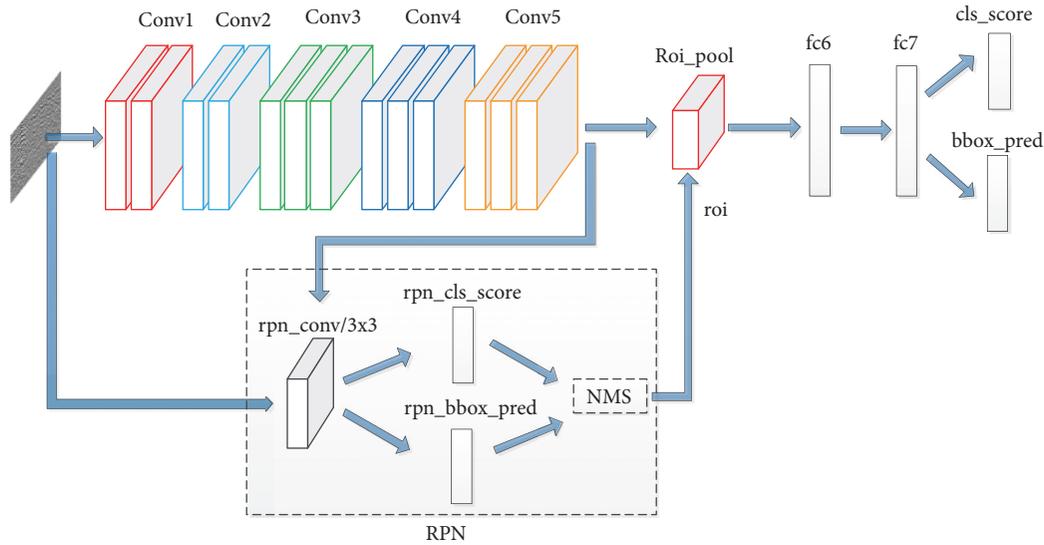


FIGURE 8: The network structure diagram of Faster R-CNN.

whether to use pretraining model and the improvement strategies of feature cascade, ASDN, Soft-NMS, and data augmentation to improve Faster R-CNN. Then, we use our self-made subgrade defect dataset to fine tune the model. Finally, we analyze the detecting results and conduct ablation experiments of improvement strategies and a comparison experiment with traditional methods.

3.3. Improvement Strategies. The method of directly applying traditional Faster R-CNN to railway subgrade defect recognition cannot achieve ideal recognition effect. The main reasons are that different from general target detection task of natural images, in railway subgrade defect detection task, the boundaries between target and background are blurred, the dimension and shape of detecting targets are diverse, the proportion of positive and negative samples is highly imbalanced, and the number of samples is limited. Therefore, in view of the characteristics and difficulties in railway subgrade defect detection task, four improvement strategies of feature cascade, ASDN, Soft-NMS, and data augmentation have been integrated to improve the traditional Faster R-CNN. The network structure of the improved model is shown in Figure 10 and implementation details will be described in following parts.

3.3.1. Feature Cascade. The traditional Faster R-CNN uses the features pooled by the Roi-pooling layer to perform image classification and regression, which have wide acceptance region and coarse feature grain. Therefore, small targets are easy to be ignored. However, in the railway subgrade defect detection task, the dimension and shape of target defects are changeable, so traditional Faster R-CNN does not achieve expected detection performance. Inspired by the studies in [24, 25], an improvement strategy called feature cascade has been used to improve the traditional Faster R-CNN. This improvement strategy combines the output features of shallow convolutional layers with those of

deep convolutional layers to form a new multisized feature, which can detect multiscale targets and improve model performance.

The specific implementation method is shown in a blue dashed box in Figure 10. Firstly, the output features of conv3_3, conv4_3, and conv5_3 layers are, respectively, pooled and normalized by their own Roi-Polling layer and L2_Norm layer. Then, the three output features are combined into a single feature via the Concat layer. Finally, the Conv1×1 layer integrates the multichannel features and reduces output feature dimension to match the input dimension of the fc6 layer in traditional Faster R-CNN. In addition, it is notable that the L2 Normalization layer is necessary, because the feature dimensions in shallow convolutional layers and deep convolutional layers are different. By means of the L2_Norm layer, all features from the multiple convolutional layers in different depth are unified into one dimension.

3.3.2. Adversarial Spatial Dropout Network. Compared with the target detection task of natural images, the number of samples which can be collected in the railway subgrade defect recognition task is limited, and the shape and dimension of defects are diverse. These factors will lead to the incomplete coverage of defect dataset. Therefore, a kind of adversarial nets called adversarial spatial dropout network (ASDN) [18] is introduced to solve the problem. The ASDN generates hard positive samples which are infrequent and hard to classify by confrontational learning. Then, those hard samples are input to the detector to enhance their influence so that the detecting network and its adversary, the ASDN, could learn together to improve model performance. The specific structure of the ASDN is shown in a green dotted box in Figure 10. A mask with the same size of the output of feature cascade is generated to modify its partial feature, which makes the detector difficult to distinguish. Through the improvement of the ASDN, the detection precision of the model is improved.

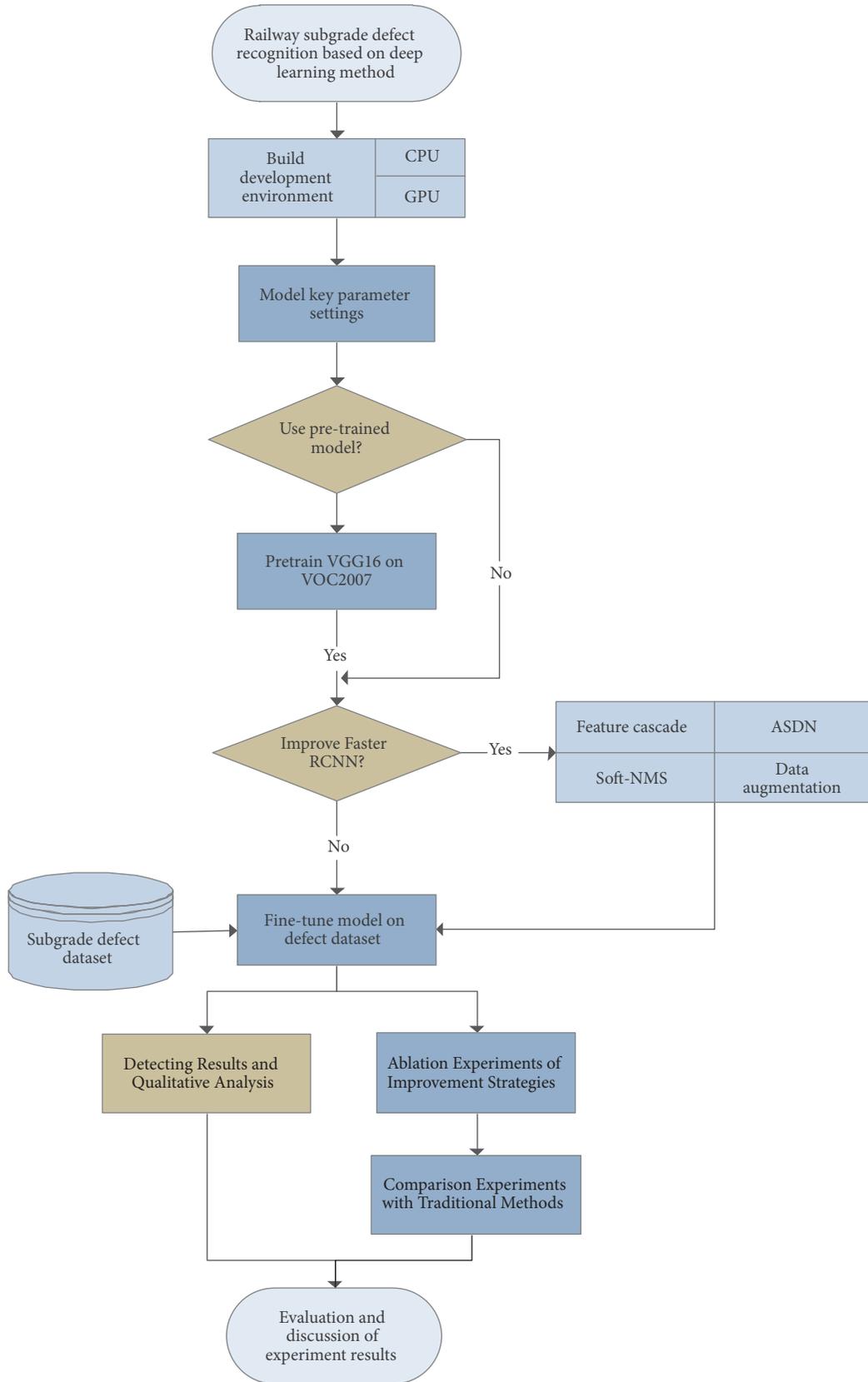


FIGURE 9: The flow diagram of the proposed method for subgrade defect recognition.

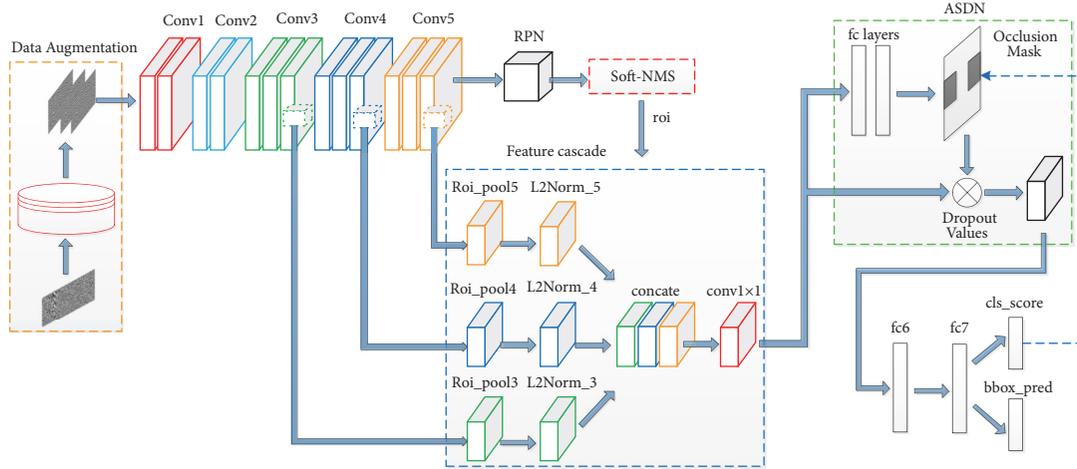


FIGURE 10: The network structure of the improved model.

3.3.3. Soft-NMS. Nonmaximum suppression (NMS) [26] is a vital part of target detection process, which can effectively suppress redundant detection boxes. However, in traditional NMS, only the detection box with the highest confidence level will be left and all the adjacent detection boxes whose overlap with the highest one is greater than the set threshold will be suppressed. Therefore, the traditional NMS cannot detect two adjacent and similar targets. Additionally, in railway subgrade defect recognition task, two or more similar defects often occur in adjacent regions, when the traditional NMS will lead to leakage judgment. Therefore, Soft-NMS [19], an improved NMS method, is adopted in this paper, which is marked with a red dotted box in Figure 10.

Soft-NMS does not directly suppress the detection box whose overlapping degree is greater than the threshold, while reducing the confidence level according to the overlapping area of detection boxes. It effectively reduces the number of false positive results and improves model performance. Soft-NMS uses the attenuation function to adjust the confidence level of detection box, which includes a linear weighted and a Gaussian weighted function forms. In this paper, the Gaussian weighted form is used, as given by

$$s_i = s_i e^{-iou(M, b_i)^2 / \sigma}, \quad \forall b_i \notin D \quad (1)$$

where M is the detection box with the greatest confidence level; b_i is the pending detection box; $iou(M, b_i)$ is the overlap degree of M and b_i ; s_i is the corresponding confidence level of b_i ; D is the set of detection boxes reserved.

3.3.4. Data Augmentation. As the limited samples in railway subgrade defect dataset, a variety of data augmentation methods are comprehensively applied to expand the defect samples, which can effectively prevent the overfitting situation during model training process and improve model performance. Data augmentation allows the improved model to learn more image invariance features by geometric transformation, which means the pixel values of image samples are unchanged and only the pixel location is changed. The data augmentation methods used are presented in Table 2.

4. Experiments

4.1. Model Training and Key Parameters. The development environment is built on the deep learning infrastructure, Caffe [27]. Model training and testing are based on the Python implementation version [28] of Faster R-CNN. A large model, VGG16 [29], is chosen as the basic network model. Caused by the few samples in training set, the pretrained model of VOC2007 ImageNet [30] is used to initialize model parameters. Then, on this basis, self-made dataset is applied to further fine tune the model. Alternating training is used to train the model in GPU acceleration mode. The main configurations of development environment are shown in Table 3.

The maximum iterations are set up as [60000, 30000, 60000, 30000]. The global learning rate is set to 0.001, and in order to accelerate convergence process, the learning rate of new added layers in the improved model is increased. The spatial_scale parameter in Roi_pooling layer is set up according to feature map dimension, with 0.25 in roi_pool3, 0.125 in roi_pool4, and 0.0625 in roi_pool5. The num_output parameter in Conv1x1 layer is set up to 512 to match the input dimension of original fc6 layer.

4.2. Detecting Results and Qualitative Analysis. The mud pumping defects are taken as examples; two detecting results selected randomly in testing set are shown in Figure 11. It can be obtained from Figure 11(a) that the two mud pumping defects, separated by a short section of normal subgrade, are correctly recognized by the proposed model; they both get a high confidence level and a detection box which is basically consistent with actual defect region. From Figure 11(b), under the interference of the bridge in the middle, the mud pumping defect is still accurately detected with a more than 90% confidence level.

4.3. Ablation Experiments of Improvement Strategies. In order to quantitatively evaluate the effect of each improvement strategy proposed to the model performance, multitudes

TABLE 2: Data augmentation methods applied in the improved model.

Data Augmentation Method	Method description
Flip	Flipping the image horizontally or vertically.
Contrast	Changing the saturation S and V luminance component in the HSV color space of the image, and keeping the tone H unchanged.
Zoom	Enlarging or compressing the image according to a certain proportion.
Random Cropping	Clipping the image randomly, including the scale jittering method and the enhancement transform of dimension and length-width ratio.
Rotation	Rotating the image at arbitrary angle to change the orientation.

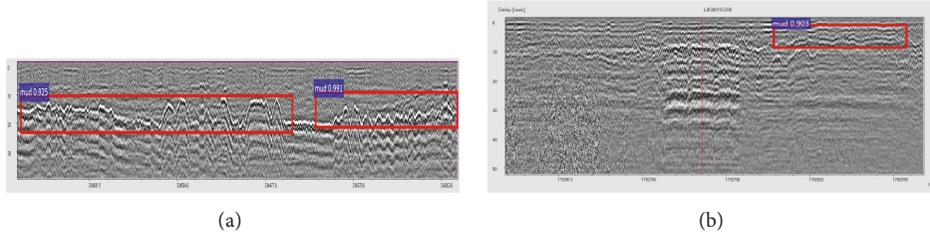


FIGURE 11: The defect detection results of the proposed method. (a) shows the result of two adjacent mud pumping detected. (b) shows the detected mud pumping close to a bridge.

TABLE 3: The configurations of development environment.

Equipment	Configuration
CPU	Intel Core i7-5930K
Graphics Card	NVIDIA 1070, 8 GB GDDR5
Memory	32 GB DDR4
Operating system	Linux (Ubuntu 14.04)

of ablation experiments are designed, which are shown in Table 4. The experimental results displayed in ROC curves are shown in Figure 12.

In Table 4, the symbols of check mark and cross indicate whether the model located in this line has applied the kind of improvement strategy. Therefore, none of the improvement strategies has been used to Model 1. Pretrained model has been applied to improve Model 2. Model 3 applied the improvement strategies of pretrained model and feature cascade. Model 4 applies pretrained model and ASDN. Model 5 applies three improvement strategies which are pretrained model, feature cascade, and ASDN. Model 6 uses all improvement strategies except data enhancement. Model 7 applies all five improvement strategies.

Firstly, the effect of pretrained model on model performance is examined by the ablation experiment (Model 1 versus Model 2). Model 1 is a traditional Faster R-CNN detection algorithm without any improvement strategy, while Model 2 uses the pretrained model on PASCAL VOC2007 for parameter initialization. It can be obtained from the ROC curves of the two models that the light blue one of Model 2 is closer to upper left corner than the black one of Model 1. Therefore, the performance of Model 2 is better.

Secondly, the improvement strategy of feature cascade is validated by the ablation experiment experiments (Model 2

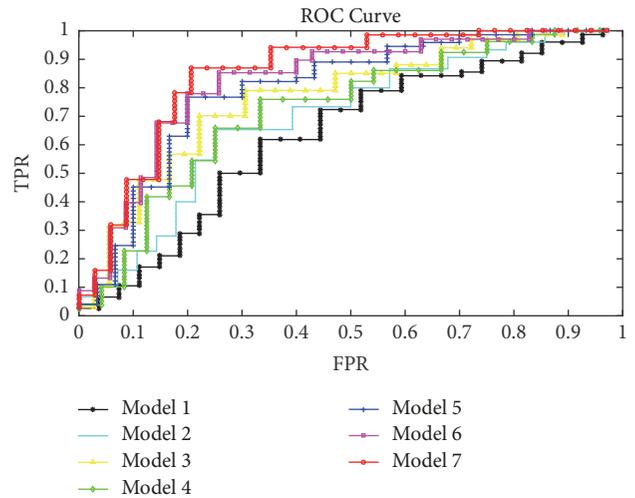


FIGURE 12: The ROC curve of ablation experiment results.

versus Model 3 and Model 4 versus Model 5). As shown in Figure 12, Model 3 gets the better model performance than Model 2. Compared with Model 4, Model 5 has been significantly improved on model performance. Therefore, it can be obtained that feature cascade is an effective improvement strategy.

Thirdly, from the ablation experiment (Model 2 versus Model 4), the ASDN improvement strategy indeed improves the model performance. However, it will generate more false positive samples at the same time. The use of Soft-NMS just makes up for the shortcoming of ASDN, and the promotion has been confirmed by a further experiment (Model 5 versus Model 6).

TABLE 4: The ablation experiment of the improvement strategies.

Model ID	Pre-trained model	Feature cascade	ASDN	Soft-NMS	Data augmentation
Model 1	×	×	×	×	×
Model 2	√	×	×	×	×
Model 3	√	√	×	×	×
Model 4	√	×	√	×	×
Model 5	√	√	√	×	×
Model 6	√	√	√	√	×
Model 7	√	√	√	√	√

TABLE 5: The results of the model performance comparison experiment.

Method	Detect speed/ (sec)		Precision	Recall	F-Score
	GPU	CPU			
HOG+SVM	-	0.136	54.7%	29.8%	38.6%
Faster-RCNN	0.091	0.367	68.3%	66.9%	67.6%
The proposed method	0.093	0.391	85.2%	82.1%	83.6%

Finally, combing all the improvement strategies, Model 7 achieves the best detection performance, which verifies the effectiveness of improvement strategies applied in this paper.

4.4. Comparison Experiment with Traditional Methods

4.4.1. Experiment Design. In order to comprehensively evaluate the performance of the proposed method, a comparison experiment with the traditional sliding method of SVM and histogram of oriented gradient (HOG) is conducted. Since SVM is a binary classifier, four linear SVM classifiers are trained separately for four kinds of subgrade defects which are the mud pumping, subgrade settlement, ballast fouling, and water abnormality. Firstly, the positive samples of each defect type are cropped from the self-made subgrade defect dataset, and the negative ones are randomly cropped from nontarget regions. Then, the training samples are uniformly rescaled to 224×224, and HOG features are extracted to train linear SVM classifiers. Finally, in detection stage, a sliding window with the dimension of 224×224 is used for defect detection, which is consistent with the input dimension cropped in Faster R-CNN.

4.4.2. Evaluation Method. In order to objectively evaluate the performance of three models in this experiment, three evaluation indexes which are the Precision, Recall, and F-Score are used. Precision is the proportion of predicted defect samples that are actual defects, as given by

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

where TP is the number of actual defect samples which are predicted to be defects; FP is the number of actual background samples which are predicted to be defects.

Recall is the proportion of actual defect samples that are correctly predicted to be defects, as given by

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

where FN is the number of the actual defect samples that are predicted to be backgrounds.

F-Score is the harmonic mean of Precision and Recall, as given by

$$F - Score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

4.4.3. Results and Analysis. The comparison experiment results are shown in Table 5. It can be obtained that the precision of the proposed method and traditional Faster R-CNN are greater than the HOG+SVM, because in deep learning approach, hidden features which are better than HOG can be extracted from training data. Meanwhile, the precision of the proposed method is greater than traditional Faster R-CNN, since the comprehensive application of improvement strategies could effectively overcome the difficulties in railway subgrade defect recognition. For the index of recall, the proposed method performs much better than the others. Furthermore, the recall index is the most vital index in the subgrade defect recognition task. If this index is quite low, a large number of actual defects will be omitted. Finally, from the comprehensive performance index F-Score, the proposed method achieves the greatest detection performance.

In addition, the detection speed of the three models is compared in CPU and GPU mode, respectively. The comparative results are displayed in Table 5. The index of detect speed means the average detection time per image in testing set. It can be obtained that although the detection speed of the proposed method is lower than the other models in CPU mode, it increases nearly 6 times in GPU mode and outperforms the traditional approach, which has basically achieved the real-time detection.

A sensitivity analysis has been conducted to test whether the model proposed is overfitting. The 10-fold cross-validation has been used to design experiments. Therefore,

TABLE 6: The results of 10-fold cross-validation experiment.

Method	Dataset										Average
	1	2	3	4	5	6	7	8	9	10	
HOG+SVM	37.7	38.6	38.9	39.5	36.9	37.9	39.3	38.3	38.8	39.4	38.5
Faster-RCNN	67.2	67.5	66.8	67.8	68.3	67.9	67.3	67.8	67.5	68.1	67.6
Proposed method	82.9	84.4	83.7	82.9	82.3	84.5	84.8	83.7	83.8	82.6	83.6

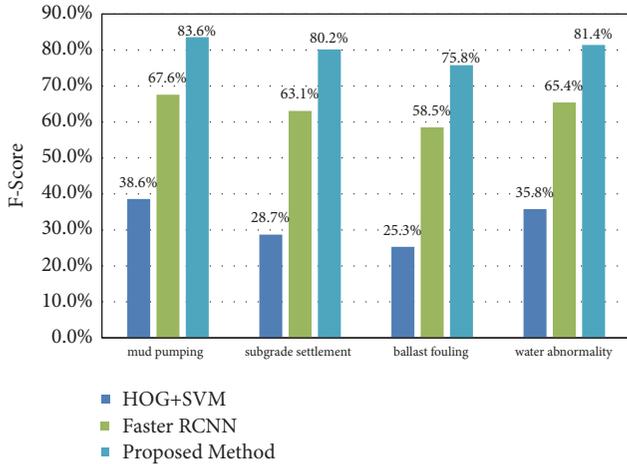


FIGURE 13: The detection accuracy comparison of the four kinds of railway subgrade defects.

the defect dataset is evenly and randomly divided into ten parts and the data of each defect type and normal subgrade in each part are kept in the same distribution. Then, one of them is selected as test set in turn, and the rest is used as training set and validation set to train and evaluate model. The cross-validation experimental results are shown in Table 6.

It can be seen from the experimental results in Table 6 that the F-Score of the 10 group dataset of the proposed method are all above 82% and relatively stable and achieves an average F-Score of 83.6%. Therefore, the model trained in this paper has no overfitting. In the same way, the overfitting phenomenon do not occur in the other two models of HOG+SVM and Faster R-CNN, which achieve the average F-Score of 38.5% and 67.6%, respectively.

Furthermore, in order to examine the model performance for various defect types, a comparison experiment is conducted with the same testing set consisting of mud pumping, subgrade settlement, ballast fouling, and water abnormality. The experimental results are shown in Figure 13. Compared with the other two methods, the proposed method achieves the greatest detection precision and stability for four kinds of defects. It is notable that the detecting precision of the ballast fouling defect is lower in all three models, because the features of ballast fouling in GPR profile are extremely complex and the ballast fouling is sensitive to signal interference from the structures such as turnouts and bridge guardrails.

5. Conclusions

An automatic recognition method of railway subgrade defects based on an improved Faster R-CNN was proposed. The

ablation experiments of improvement strategies are shown that the pretrained model, feature cascade, ASDN, Soft-NMS, and data augmentation can effectively improve the model performance. The result of the comparison experiment with traditional approach showed that the proposed method achieves the greatest mAP of 83.6%, much greater than the 38.6% of HOG+SVM, and is 16% greater than that of traditional Faster R-CNN. Furthermore, the detection speed of 0.093s in GPU mode outperforms the traditional approaches and basically meets the requirement of real-time detection. Additionally, the comparison experiment for various defect types further verifies the strong robustness of the proposed method. In the future, the following deficiencies still need to be further studied: (1) the original GPR signal data of GPR loses part of data accuracy during the process of converting to GPR profile, which affects the recognition accuracy. Therefore, how to recognize defects from raw data format of GPR may be a future research direction. (2) The current recognition algorithm can only identify four types of defects, so the promotion of recognition algorithms to other types of subgrade defects needs to be further studied.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

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

Acknowledgments

This work was supported by the Key Program of High-speed Railway Basic Research Joint Fund of China Railway (U1434211), the Science and Technology Development Program of China Railway (2017G003-H), and the China University of Mining & Technology (Beijing) "Yueqi Outstanding Scholar".

Supplementary Materials

Figure 000001~000115: original images of ballast fouling defect. Figure 000116~000332: original images of mud pumping defect. Figure 000333~000447: original images of subgrade settlement defect. Figure 000448~000584: original images of water abnormality defect. Figure 000585~000684: original images of normal subgrade. XML file 000001~

000115: annotated files of ballast fouling defect. XML file 000116~000332: annotated files of mud pumping defect. XML file 000333~000447: annotated files of subgrade settlement defect. XML file 000448~000584: annotated files of water abnormality defect. (*Supplementary Materials*)

References

- [1] J. P. Xiao, Y. Q. Wang, and L. B. Liu, "A multiband-pass filtering method to suppress sleeper noise in railway subgrade vehicle-mounted GPR data," in *Proceedings of the 16th International Conference of Ground Penetrating Radar, GPR 2016*, pp. 1–5, June 2016.
- [2] J. Xiao and L. Liu, "Multi-frequency GPR signal fusion using forward and inverse S-transform for detecting railway subgrade defects," in *Proceedings of the 8th International Workshop on Advanced Ground Penetrating Radar, IWAGPR 2015*, pp. 1–4, July 2015.
- [3] L. Liao, X. Yang, and C. Ding, "Technique for automatic interpretation of GPR images of railroad subgrades," *China Civil Engineering Journal*, vol. 42, no. 6, pp. 102–107, 2009.
- [4] P. Du, L. Liao, and X. Yang, "Intelligent recognition of defects in railway subgrade," *Journal of the China Railway Society*, vol. 32, no. 3, pp. 142–146, 2010.
- [5] H. Zou, "A study on roadbed disease recognition algorithm and application," *Progress in Geophysics*, vol. 24, no. 6, pp. 2302–2307, 2009.
- [6] Y. Du, Z. Hou, and W. Zhao, "An identification method for heavy-haul railway subgrade defects based on sparse representation," *China Civil Engineering Journal*, vol. 46, no. 11, pp. 138–144, 2013.
- [7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proceedings of the International Conference on Neural Information Processing Systems, ICONIP 2012*, pp. 1097–1105, November 2012.
- [8] C. Szegedy, W. Liu, Y. Jia et al., "Going deeper with convolutions," in *Proceedings of the 15th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015*, pp. 1–9, Boston, Mass, USA, June 2015.
- [9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the 16th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016*, pp. 770–778, Las Vegas, Nev, USA, June 2016.
- [10] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the 14th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2014*, pp. 580–587, June 2014.
- [11] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 9, pp. 1904–1916, 2015.
- [12] R. Girshick, "Fast R-CNN," in *Proceedings of the 15th IEEE International Conference on Computer Vision, ICCV 2015*, pp. 1440–1448, December 2015.
- [13] 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, 2017.
- [14] J. Yin and W. Zhao, "Fault diagnosis network design for vehicle on-board equipments of high-speed railway: A deep learning approach," *Engineering Applications of Artificial Intelligence*, vol. 56, pp. 250–259, 2016.
- [15] X. Gibert, V. M. Patel, and R. Chellappa, "Deep multitask learning for railway track inspection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 1, pp. 153–164, 2017.
- [16] R. Pang, Z. B. Yu, W. Y. Xiong, and H. Li, "Faults recognition of high-speed train bogie based on deep learning," *Journal of Railway Science and Engineering*, vol. 12, no. 6, pp. 1283–1288, 2015.
- [17] X. Giben, V. M. Patel, and R. Chellappa, "Material classification and semantic segmentation of railway track images with deep convolutional neural networks," in *Proceedings of the IEEE International Conference on Image Processing, ICIP 2015*, pp. 621–625, December 2015.
- [18] X. Wang, A. Shrivastava, and A. Gupta, "A-fast-RCNN: hard positive generation via adversary for object detection," in *Proceedings of the 17th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, February 2017.
- [19] N. Bodla, B. Singh, R. Chellappa, and L. S. Davis, "Improving object detection with one line of code," in *Proceedings of the IEEE International Conference on Computer Vision, ICCV 2017*, October 2017.
- [20] D. J. Daniels, "Ground penetrating radar," *Encyclopedia of RF and Microwave Engineering*, vol. 58, no. 12, pp. 183–194, 2005.
- [21] J. R. R. Uijlings, K. E. A. Van De Sande, T. Gevers, and A. W. M. Smeulders, "Selective search for object recognition," *International Journal of Computer Vision*, vol. 104, no. 2, pp. 154–171, 2013.
- [22] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE International Conference on Computer Vision, CVPR 2015*, pp. 3431–3440, June 2015.
- [23] S. Bell, C. L. Zitnick, K. Bala, and R. Girshick, "Inside-outside net: Detecting objects in context with skip pooling and recurrent neural networks," in *Proceedings of the IEEE Conference on 16th Computer Vision and Pattern Recognition, CVPR 2016*, pp. 2874–2883, June 2016.
- [24] T. Kong, A. Yao, Y. Chen, and F. Sun, "HyperNet: towards accurate region proposal generation and joint object detection," in *Proceedings of the 16th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016*, pp. 845–853, June 2016.
- [25] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1627–1645, 2010.
- [26] <https://github.com/tzutalin/labelImg>.
- [27] Y. Jia, E. Shelhamer, J. Donahue et al., "Caffe: convolutional architecture for fast feature embedding," in *Proceedings of the 22nd ACM International Conference on Multimedia, ACM 2014*, pp. 675–678, November 2014.
- [28] <https://github.com/rbgirshick/py-faster-rcnn>.
- [29] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," <https://arxiv.org/abs/1409.1556>.
- [30] O. Russakovsky, J. Deng, H. Su et al., "ImageNet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.



Hindawi

Submit your manuscripts at
www.hindawi.com

