

## Article

# CycleGAN-Based Data Augmentation for Subgrade Disease Detection in GPR Images with YOLOv5

Yang Yang <sup>1</sup>, Limin Huang <sup>1</sup>, Zhihou Zhang <sup>1,\*</sup>, Jian Zhang <sup>1</sup> and Guangmao Zhao <sup>2</sup>

<sup>1</sup> Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu 611756, China; doubley@my.swjtu.edu.cn (Y.Y.); limin@my.swjtu.edu.cn (L.H.); zhangjian93@swjtu.edu.cn (J.Z.)

<sup>2</sup> China Railway Design Corporation, Tianjin 300142, China; zhaoguanguangmao@crdc.com

\* Correspondence: logicprimer@swjtu.edu.cn

**Abstract:** Vehicle-mounted ground-penetrating radar (GPR) technology is an effective means of detecting railway subgrade diseases. However, existing methods of GPR data interpretation largely rely on manual identification, which is not only inefficient but also highly subjective. This paper proposes a semi-supervised deep learning method to identify railway subgrade diseases. This method addresses the sample imbalance problem in the defect dataset by utilizing a data augmentation method based on a generative adversarial network model. An initial network model for disease identification is obtained by training the YOLOv5 network with a small number of existing samples. The intelligently extended samples are then labeled to achieve a balance in the disease samples. The network is trained to improve the recognition accuracy of the intelligent model using a more complete dataset. The experimental results show that the accuracy of the proposed method can reach up to 94.53%, which is 23.85% higher than that of the supervised learning model without an extended dataset. This has strong industrial application value for railway subgrade disease detection as the potential learning ability of the model can be explored to a greater extent, thereby improving the recognition accuracy of subgrade diseases.

**Keywords:** disease detection; ground-penetrating radar; YOLO; GAN; semi-supervised learning



**Citation:** Yang, Y.; Huang, L.; Zhang, Z.; Zhang, J.; Zhao, G. CycleGAN-Based Data Augmentation for Subgrade Disease Detection in GPR Images with YOLOv5. *Electronics* **2024**, *13*, 830. <https://doi.org/10.3390/electronics13050830>

Academic Editor: Antoni Morell

Received: 7 December 2023

Revised: 12 February 2024

Accepted: 15 February 2024

Published: 21 February 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

As the foundation of the comprehensive transportation system, railways serve a crucial role in connecting the world, facilitating communication, driving development, and promoting equality. The subgrade, as the infrastructure of railways, directly impacts the safety of trains and the comfort of passengers. Diseases of the railway subgrade are diverse and have complex causes. The most common diseases are mud pumping, settlement, and water abnormalities [1].

Mud pumping is a prevalent disease in railway subgrades, particularly in regions with substantial rainfall. The clayey soil of the subgrade bed transforms into a slurry state after water absorption. Under the dynamic influence of the train, this slurry infiltrates the ballast voids. The slurry generates a suction effect under the repeated load, leading to an upward extrusion [2]. The settlement process involves an initial consolidation due to the railway's weight, followed by subsequent compression deformation under the train's load. Various factors, including embankment slope, embankment height, ground temperature, and local roadbed soil type, can influence this process [3]. Water abnormalities primarily occur when the water content of the subgrade material exceeds its standard range. This condition is typically caused by inadequate drainage, rainwater infiltration, rising groundwater levels, or improper control of material water content during construction. When the water content of the subgrade material is excessively high, its physical properties such as strength decrease and the compressibility increases, leading to subgrade deformation [4]. This deformation affects the railway's stability and safety. These diseases significantly impact

the regular operation of railway lines, necessitating the implementation of scientific and rational detection methods.

Ground-penetrating radar (GPR) is an effective tool for detecting diseases in railway subgrades. It operates by emitting pulses of electromagnetic radiation into the surface under survey from a transmitter. When these waves encounter any subsurface changes, a portion of the electromagnetic energy reflects back and is received by an antenna [5]. The time taken for the reflected signals to return is measured, providing an indication of the depth and location of the disruption [6]. GPR can penetrate various materials such as soil, debris, water, and concrete, each having different dielectric and conductive properties. After data processing, information is revealed about the internal structure, shape, and even the presence of abnormal objects in the subterranean space [7]. GPR has found extensive use in numerous fields, including groundwater detection, underground facility positioning, and geological structure identification. In addition, it has become a popular tool for detecting railway subgrade diseases due to its light weight, high efficiency, non-destructive nature, high resolution, and strong anti-interference capabilities [8]. However, the substantial amount of data obtained by ground-penetrating radar for railway subgrade detection are primarily interpreted manually. This leads to issues such as a low recognition efficiency, a lack of unified objective standards, and difficulty in ensuring accuracy, thereby limiting the widespread application of this technology in the field of subgrade detection.

Deep learning (DL) [9] is instrumental in the new industrial revolution. It uses a series of data transformation layers to map input data to an expected output. Despite the simple structure and function of each individual component, complex nonlinear mapping can be achieved when dozens or even hundreds of layers are combined. In the past decade, deep learning has made revolutionary progress in many historically challenging areas, such as image segmentation, speech recognition, machine translation, emotion classification, target detection, and Go games [10–15]. With the vigorous development of DL, many researchers have attempted to apply deep learning to the field of ground-penetrating radar and have achieved significant success.

According to Sun et al. [16], an improved EfficientNetv2-s classification method was proposed for more precise classification and recognition of radar radiation source signals. Compared with other deep learning image classification methods, this method demonstrated a better classification accuracy on the test set, with a top-1 accuracy of 98.12%, which is 0.17–3.12% higher than other methods. Liu et al. [17] discussed a deep neural network (DNN) structure referred to as GPRInvNet, which was developed to tackle the challenges of mapping ground-penetrating radar (GPR) B-Scan data to complex permittivity maps of subsurface structures. The results demonstrated that GPRInvNet is capable of effectively reconstructing complex tunnel lining defects with clear boundaries. Yue et al. [18] proposed an improved model called a least square generative adversarial network (LSGAN) to address the scarcity of labeled GPR data. This model incorporates the loss functions of LSGANs and convolutional neural networks (CNNs) to generate high-precision GPR images. It has been empirically demonstrated that including LSGAN-generated images in the training GPR dataset enhances target diversity and improves detection precision.

The field of CNNs has seen rapid growth and development in recent years, with researchers continually pushing the boundaries of what is possible with these powerful models. With the rapid growth in data amounts and the increasing power of graphic processor units, many researchers have improved CNNs and achieved state-of-the-art results in railway subgrade disease detection tasks. Zhang et al. [19] extracted an intelligent identification method for railway subgrade diseases based on unsupervised learning and utilized the k-means clustering algorithm and LandTrendr algorithm to achieve automatic classification of railway subgrade GPR signals. The test results demonstrated that this method is capable of distinguishing three types of signals, with the classification accuracy rate exceeding 95%. Xu et al. [20] integrated several improvement strategies, such as feature cascade, an adversarial spatial dropout network (ASDN), Soft-NMS, and data augmentation, into the Faster R-CNN framework to improve the recognition accuracy

according to the characteristics of subgrade defects. The experimental results indicated that compared with the traditional SVM+HOG method and the baseline Faster R-CNN, the improved model can achieve a better performance. Liu [21] developed a novel deep learning method to process GPR data, the CRNN, which combines a convolutional neural network (CNN) and a recurrent neural network (RNN). The authors proposed using the CNN to process raw GPR waveform data from signal channels and the RNN to process features from multiple channels. The results showed that the CRNN achieves a higher precision at 83.4%, with a recall of 77.3%.

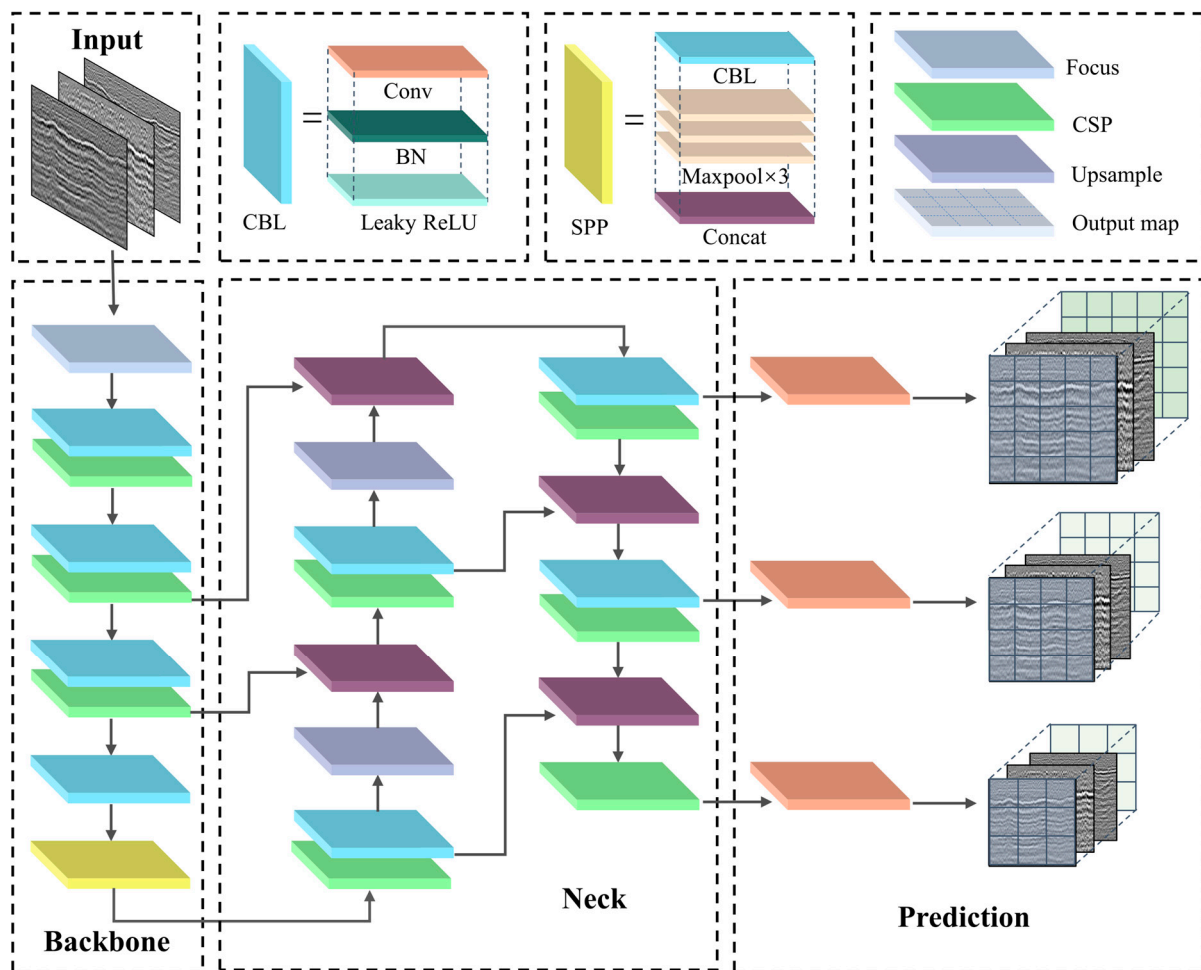
Although the above methods are feasible for intelligent recognition of ground-penetrating radar spectra, there are still some challenges that have not been overcome. These methods either rely on a large amount of annotated radar data or suffer from problems such as a weak model identification accuracy due to sample data imbalance, which cannot meet the requirements of real-time detection. Hence, it has become an urgent technical challenge to explore innovative DL technologies for accurately identifying railway subgrade diseases while reducing costs and enhancing efficiency.

In light of the above, this paper proposes a semi-supervised railway subgrade disease identification method based on YOLOv5. The method is divided into four steps: First, the original subgrade disease dataset is intelligently expanded, with the generative adversarial network [22] as the backbone network to obtain a certain amount of unlabeled disease data. Following this, the pretrained network is used to predict the expanded data and generate pseudo-labels, resulting in the acquisition of a pseudo-labeled disease sample dataset. Finally, the original dataset and the pseudo-labeled dataset are input into the YOLOv5 neural network for semi-supervised training.

## 2. Image Recognition for Ground-Penetrating Radar Based on Deep Learning

### 2.1. Radar Data Recognition Method Based on YOLOv5

This paper employs the YOLOv5 neural network for radar data prediction. YOLOv5, a convolutional-neural-network-based object detection model, outperforms its predecessors in the YOLO series [23–25] in terms of detection performance and speed. It comprises four network models, YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x, each increasing in complexity and recognition accuracy. The paper utilizes the YOLOv5x model for training, which is structured into four parts: Input, Backbone, Neck, and Prediction. As depicted in Figure 1, the radar data are initially input into the network for preprocessing. The Backbone segment then extracts image feature information, followed by multi-scale feature fusion in the Neck network. Finally, the image is output to complete the prediction.



**Figure 1.** The network structure diagram of YOLOv5.

### 2.1.1. Disease Image Preprocessing

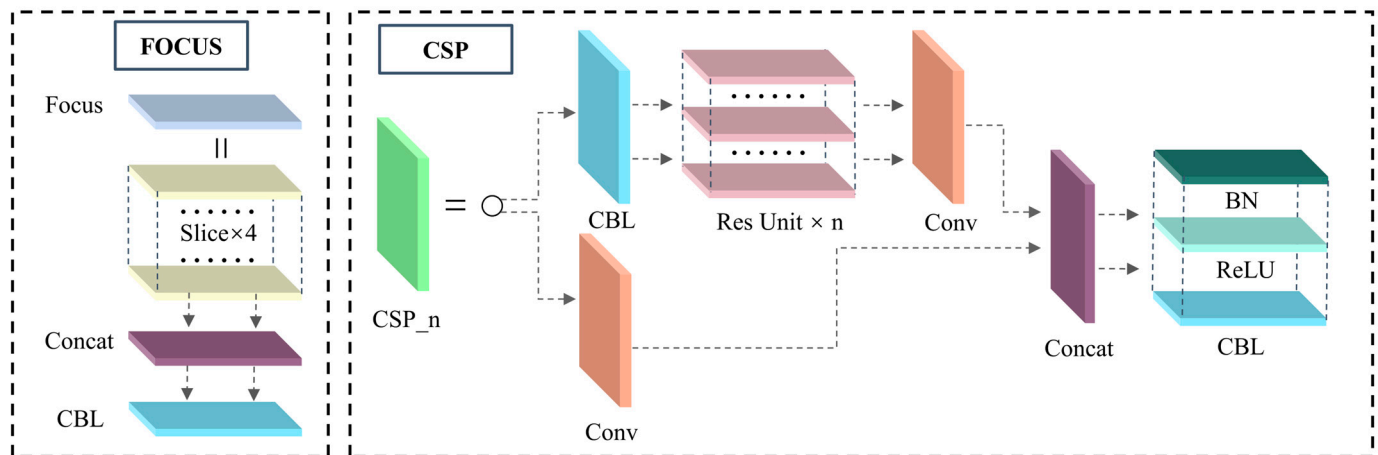
In the Input segment, the YOLOv5 neural network employs Mosaic data enhancement [26]. Specifically, four disease radar images are selected for random scaling, cropping, and arrangement for splicing. This method effectively enriches the dataset, particularly through the use of random scaling to process images, which increases the proportion of small targets in the dataset, thereby enhancing the network's robustness. Additionally, the YOLOv5 neural network utilizes an adaptive anchor box calculation method to compute the difference between the predicted and actual boxes, subsequently updating the network inversely. Through iterative updates, the optimal anchor box value is ultimately calculated. Interference information in the image is eliminated by relying on adaptive image scaling to reduce errors and further improve accuracy and detection speed, making it more suitable for identifying disease locations in radar data.

### 2.1.2. Extraction of Disease Feature Information

The Backbone segment primarily comprises the Focus and CSP structures, with their network structure diagrams depicted in Figure 2. The Focus structure emphasizes slicing operations. Following the preprocessing of the disease radar image, the original image is uniformly enlarged to a size of  $608 \times 608 \times 3$ . It is then input into the Focus structure for slicing operations, transforming it into a  $304 \times 304 \times 12$  feature image. After convolution operations, a  $304 \times 304 \times 32$  feature image is generated. This method effectively reduces the network model's computational load and enhances the network training speed. Upon extraction of the disease atlas features, the Convolutional Sparse Coding (CSP) [27] algorithm initially partitions the disease features from the base layer into two subsets. These



subsets are subsequently integrated using a cross-stage hierarchical structure, thereby ensuring the precision of the deep fusion process applied to the disease features while simultaneously minimizing computational requirements.



**Figure 2.** The network structure diagram of Focus and CSP.

### 2.1.3. Multi-Scale Feature Fusion of Diseases

The Neck component of the network, positioned between the Backbone and Output, serves to further extract and integrate the image feature information produced by the Backbone, subsequently passing it to the output layer. The YOLOv5 neural network employs a feature pyramid network (FPN) [28] structure within the Neck component. Through a series of upsampling and downsampling operations, feature images at various levels are amalgamated to generate the final disease feature image for target detection. This approach is particularly effective in handling disease entities of varying scales and sizes within radar data, enhancing the detection and recognition of disease entities across different proportions and sizes, thereby improving prediction accuracy.

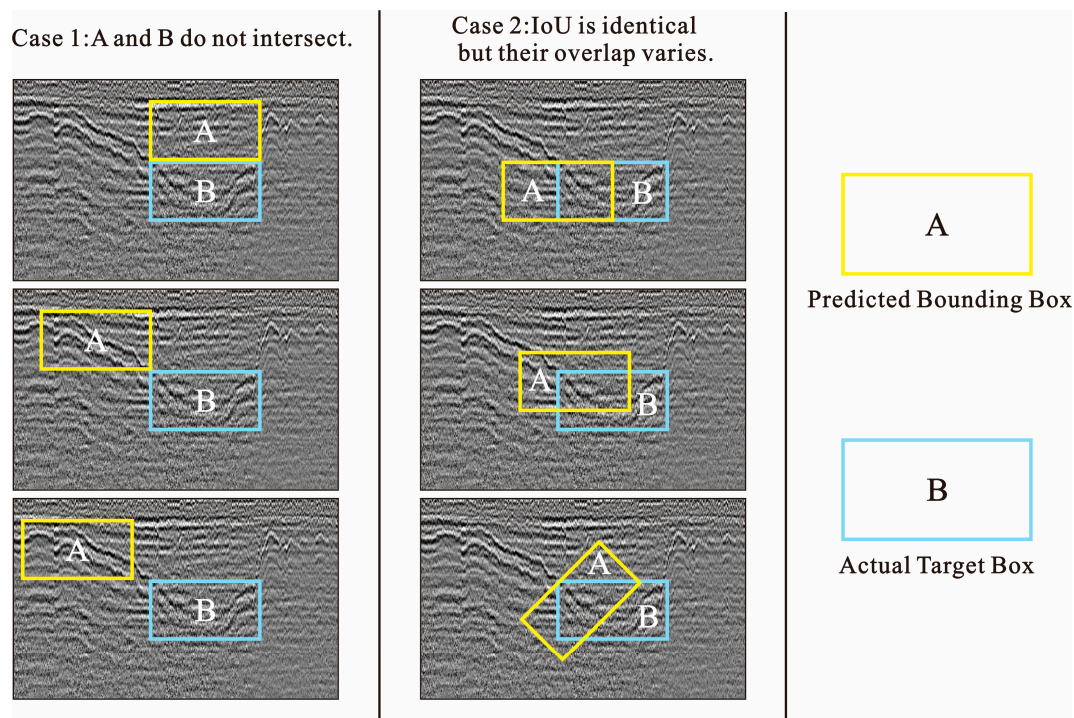
### 2.1.4. Radar Graph Output and Prediction

In the prediction segment, the YOLOv5 neural network employs GIoULoss [29] as the bounding box loss function. The calculation of GIoULoss, as depicted in Equation (1), involves several elements: A signifies the predicted bounding box by the algorithm, B denotes the actual target box, and C represents the smallest enclosing rectangle of A and B.

$$\text{GIoU} = \text{IoU} - \frac{|C - (A \cup B)|}{|C|} \quad (1)$$

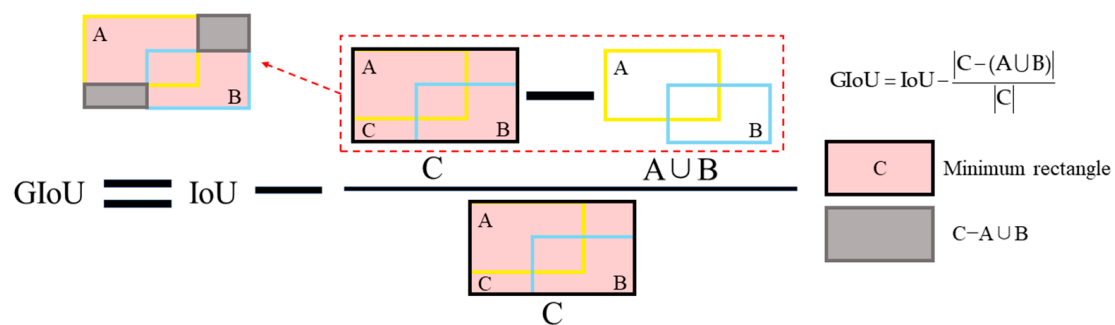
If the IoULoss [30] from the preceding version of YOLO is employed, specifically the computation method in Equation (2), the difference between A and B will not be accurately represented when the predicted and actual bounding boxes do not intersect. In these instances, the loss function is non-differentiable, suggesting that it cannot optimize scenarios where the predicted and actual bounding boxes do not intersect. Furthermore, once the intersection value of the two boxes is established, the IoU value remains unchanged and fails to depict the manner in which the two boxes intersect. Figure 3 illustrates two cases of IoU failure

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|} \quad (2)$$



**Figure 3.** Two cases of IoU failure.

GIoU was developed to address the limitations of IoU while fully leveraging its benefits. The computation of GIoU, depicted in Figure 4, introduces the concept of the minimum bounding rectangle and resolves the issues with IoU. When the IoU is 0, implying that box A is far from box B, the ratio of  $A \cup B / C$  approaches 0, and GIoU tends towards  $-1$ . Conversely, when the IoU is 1, indicating that the two boxes overlap, the ratio of  $A \cup B / C$  is 1. Thus, the value of GIoU lies within the interval  $(-1, 1]$ . This is how GIoULoss addresses these inherent issues in IoULoss, thereby enhancing network optimization.



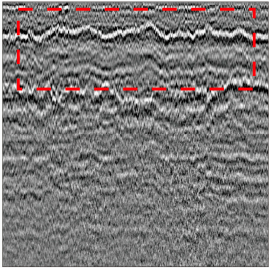
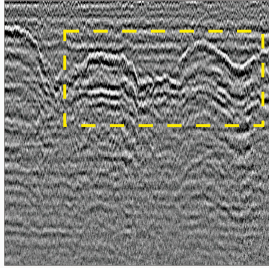
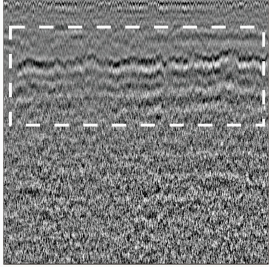
**Figure 4.** The calculation process of GIoU.

## 2.2. Dataset Construction

The GPR system used in our study is the RIS GPR, developed by the Italian company Ingegneria dei Sistemi (IDS). This system is utilized to scan and detect the railway subgrade, producing a grayscale image of its profile. The radar system comprises a three-channel antenna group, a central control system, a host radar unit, signal display meters, ranging devices, and transmission cables. The antenna group, which transmits at 400 MHz, is installed beneath a detection train developed by the China Academy of Railway Sciences, maintaining a clearance of 30 cm from the track surface. The data collection time window is set to 60 ns, with 512 sampling points and a tracking interval of 11.25 cm. The disease images we obtained were then manually interpreted to identify and mark typical diseases

such as mud pumping, subgrade settlement, and water abnormalities. The dataset contains a total of 625 images, including 500 mud pumping images, 68 subgrade settlement images, and 57 water abnormality images. Table 1 displays these three typical disease radar images and their distribution in the dataset. Since one railway subgrade disease image may contain multiple diseases, the actual number of diseases is greater than the total number of images.

**Table 1.** Radar images of three typical diseases and their distribution in the dataset.

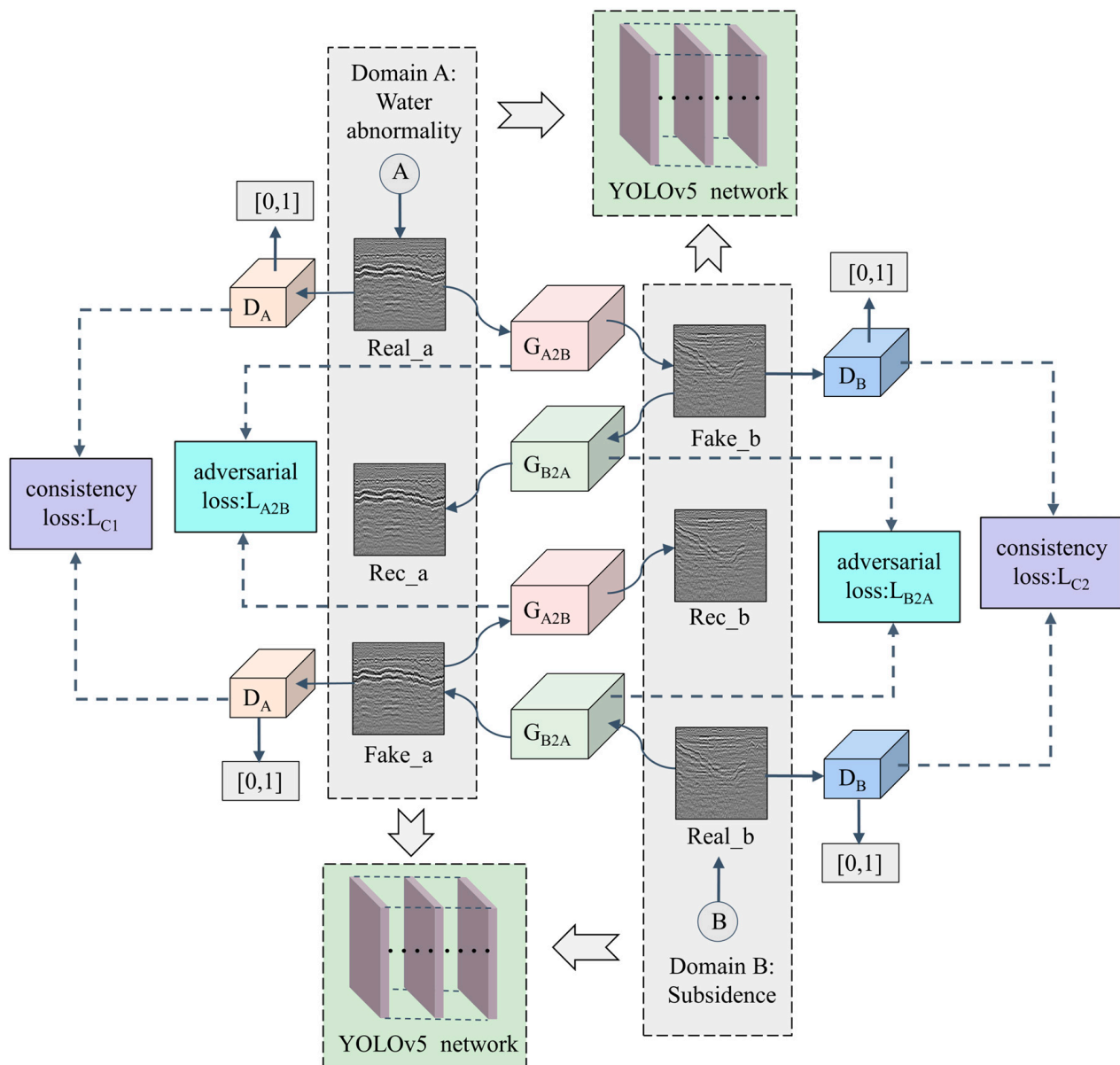
| Railway Subgrade Defect Types | Typical Defects Image                                                               | Number of Images | Image Proportion |
|-------------------------------|-------------------------------------------------------------------------------------|------------------|------------------|
| Mud pumping                   |    | 500              | 80%              |
| Settlement                    |   | 68               | 10.88%           |
| Water abnormality             |  | 57               | 9.12%            |

As indicated in Table 1, the distribution of these three diseases in the original dataset is significantly unbalanced. Specifically, the number of images related to mud pumping far exceeds those of the other two diseases. Mud pumping is closely associated with rainfall, while settlement is linked to the upper load and the compactness of the subgrade. The occurrence of water abnormalities is tied to the roadbed material. In the environment where data collection was conducted, rainfall is frequent, leading to a high incidence of mud pumping. Consequently, a model trained on this dataset identifies all disease images as mud pumping, achieving an accuracy of up to 80%. However, such a prediction is evidently meaningless. For deep learning, the balance and adequacy of sample data directly influence the quality of the trained model and determine its generalizability and practicality. Ample sample data can also effectively circumvent the long-tail effect and prevent overfitting [31,32]. Therefore, to enhance the quality of the sample dataset effectively, this paper proposes a radar data augmentation method based on the CycleGAN neural network.

### 3. Intelligent Augmentation of Image Data

#### 3.1. Data Expansion Based on CycleGAN

CycleGAN [33] is a deep learning model derived from traditional generative adversarial networks (GANs) [22] that can use unsupervised learning to achieve image conversion. It can convert images from one set (source images) to another (target images) without providing any label or annotation information about the target image. CycleGAN uses two GANs to learn the mapping from the source image to the target image and vice versa. The output between them is coordinated by a set of cyclic constraints to ensure that the generated results are as close to the real data as possible. Figure 5 shows the model structure of CycleGAN.



**Figure 5.** The network structure diagram of CycleGAN.

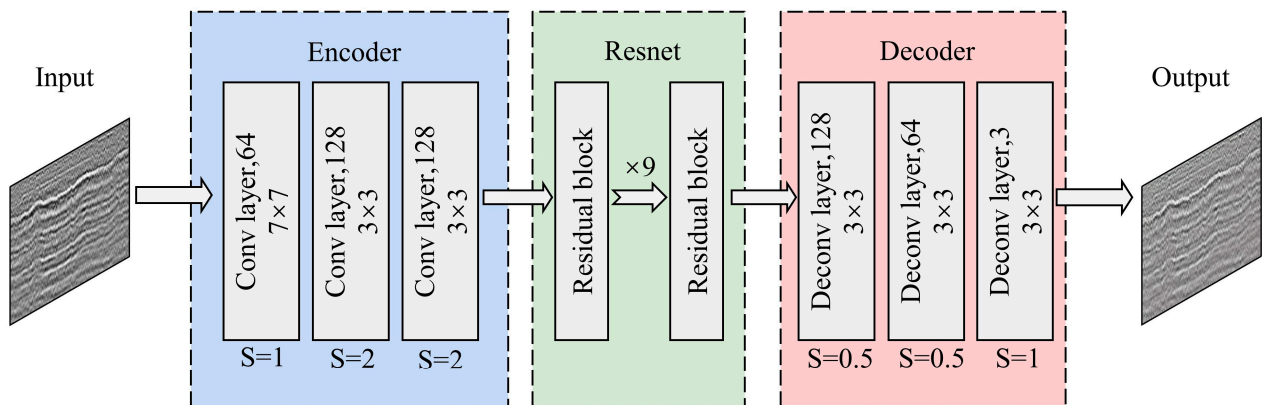
In Figure 5, Real\_a represents an image belonging to the water abnormality disease class, and Real\_b represents an image belonging to the subgrade settlement disease class. First, Real\_a is input into the generator  $G_{A2B}$  to generate a B-style image called Fake\_b, which is then input into the generator  $G_{B2A}$  to generate a reconstructed image named Rec\_a. Similarly, Real\_b is processed in the same way as Real\_a to obtain Fake\_a and Rec\_b. The



discriminator  $D_A$  calculates the probability of Real\_a and Fake\_a belonging to the water abnormality disease class, while the discriminator  $D_B$  calculates the probability of Real\_b and Fake\_b belonging to the subgrade settlement disease class. During this process, the reconstructed images Rec\_a and Rec\_b are used to construct cycle consistency loss, and the probability values calculated by discriminators  $D_A$  and  $D_B$  are utilized to construct adversarial loss.

### 3.1.1. Network Structure

The generator structure of CycleGAN, as shown in Figure 6, is mainly divided into an encoding–decoding structure. The encoder uses three convolutional layers for downsampling to extract image features. A residual block [34] structure is then introduced between the encoder and the decoder to increase the network depth, ensure the smooth transmission of gradient information in the deep network, and achieve style conversion. The decoder is the opposite of the encoder, using three deconvolutional layers for upsampling to restore the changed features step by step. After each convolution and deconvolution operation, instance normalization [35] is performed to make the network more inclined to use all features and reduce the difficulty of training. The normalized features are then activated with the ReLU activation function to increase the nonlinearity of the neural network.



**Figure 6.** The network structure diagram of the generator.

The discriminator of CycleGAN adopts the Markov structure [36], which is composed entirely of convolutional layers and outputs an  $N \times N$  matrix. Each point in the matrix represents a receptive field in the original image, corresponding to a region of the original image. Finally, the average value of the output matrix is taken as the basis for judging the generated image and the real image. The Markov discriminator takes into account the influence of different parts of the image and can effectively maintain the high resolution and high detail of the image.

### 3.1.2. Loss Function

As demonstrated in the above analysis, CycleGAN comprises a ring network structure with two sets of generators and discriminators. Its loss function consists of two components: adversarial loss and cycle consistency loss. The two sets of generators facilitate the conversion between water abnormalities and subgrade settlement and vice versa. Consequently, the adversarial loss of CycleGAN comprises two parts, denoted as  $L_{A2B}$  and  $L_{B2A}$ . The loss function of the generator is defined as follows:

$$L_{A2B} = E_{b \sim p_{data}(b)} [\log D_B(b)] + E_{a \sim p_{data}(a)} [\log(1 - D_B(G_{A2B}(a)))] \quad (3)$$

$$L_{B2A} = E_{a \sim p_{data}(a)} [\log D_A(a)] + E_{b \sim p_{data}(b)} [\log(1 - D_A(G_{B2A}(b)))] \quad (4)$$



In this formula,  $E_x$  denotes the computation of the expected value, while  $P_{data(a)}$  and  $P_{data(b)}$  represent the probability distribution of the A and B image domains, respectively. As such, the adversarial loss,  $L_{GAN}$ , is defined as follows:

$$L_{GAN} = L_{A2B} + L_{B2A} \quad (5)$$

The least squares generative adversarial network (LSGAN) [37] employs a least squares loss function as a substitute for the cross-entropy loss function used in traditional GANs. This effectively enhances the quality of image generation, accelerates the convergence rate of the model, and improves the stability of training. By adopting the LSGAN method, the cross-entropy loss is replaced with a least squares loss, and the logarithmic operation is optimized to a square operation. The resulting loss function is as follows:

$$L_{A2B} = E_{b|data(b)}[D_B(b)]^2 + E_{a|data(a)}[1 - D_B(G_{A2B}(a))]^2 \quad (6)$$

$$L_{B2A} = E_{a|data(a)}[D_A(a)]^2 + E_{b|data(b)}[1 - D_A(G_{B2A}(b))]^2 \quad (7)$$

The adversarial loss function alone cannot guarantee that a single input can be mapped. It is possible for the generator to convert all water abnormality images into the same subgrade settlement image, rendering the adversarial loss function ineffective. To prevent this, a cycle consistency loss function,  $L_C$ , is introduced. By converting a to b and then converting b back to a to obtain  $a'$ , the loss value between a and  $a'$  is calculated, ensuring the effectiveness of training.

$$L_C = E_{a|data(a)}[\|a' - a\|_1] + E_{b|data(b)}[\|b' - b\|_1] \quad (8)$$

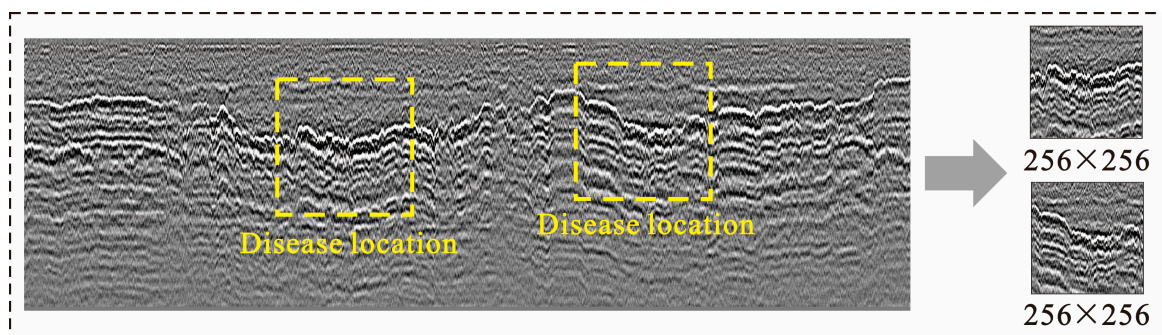
Using Formulas (6)–(8), the total loss function  $L_{CycleGAN}$  of CycleGAN can be derived:

$$L_{cyclegan} = L_{A2B} + L_{B2A} + \lambda L_C \quad (9)$$

In this formula, the parameter  $\lambda$  represents the weight of the cycle consistency loss.

### 3.2. Experimental Result

The previously described method is employed to augment two types of diseases with limited image data: settlement and water abnormalities. Initially, grayscale images of these diseases are processed according to the method depicted in Figure 7. The regions where disease is present are selected and cropped into  $256 \times 256$ -sized images for inputting into CycleGAN for training. Since a single disease image may contain multiple diseases, the actual number of cropped images exceeds the number of original disease images.



**Figure 7.** Cropping an image of subgrade settlement disease.

After cropping, a total of 121 images of subgrade settlement and water abnormality diseases was obtained. These 242 images were input into the CycleGAN model for intelligent expansion, and the expansion results are shown in Figure 8. In the end, a total of

510 images were obtained, with 255 images of subgrade settlement and water abnormality diseases each. The changes in the amount of image data before and after expansion are shown in Table 2.

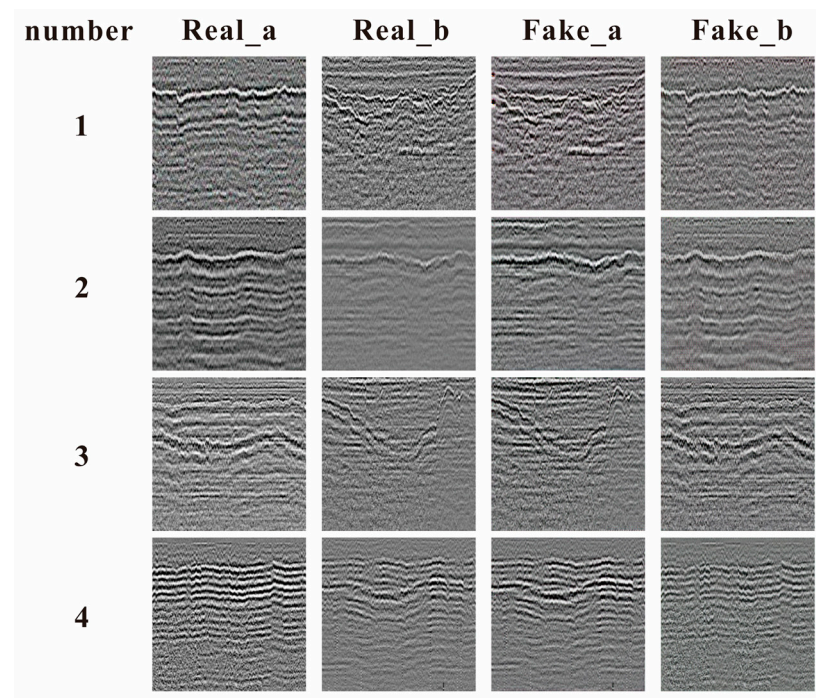


Figure 8. Extension result example.

Table 2. Statistical table of image data volume before and after expansion.

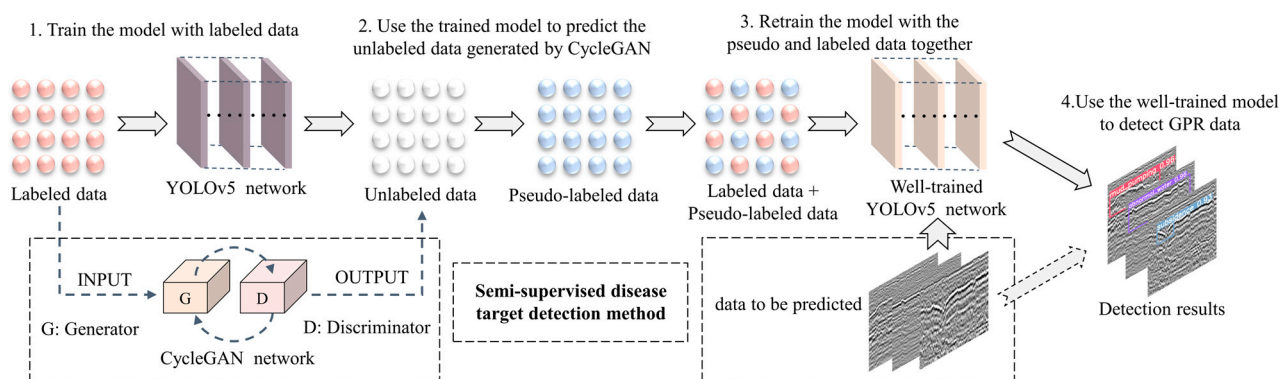
| Disease Type      | Original Dataset | After Cropping | Full Dataset | Increase Quantity |
|-------------------|------------------|----------------|--------------|-------------------|
| Mud pumping       | 500              | 500            | 500          | 0                 |
| Settlement        | 68               | 121            | 255          | 187               |
| Water abnormality | 57               | 121            | 255          | 198               |
| Total             | 625              | 742            | 1010         | 385               |

#### 4. Semi-Supervised Disease Target Detection Method

##### 4.1. Semi-Supervised Disease Target Detection Method Based on Radar Data

Semi-supervised learning, a significant methodology in deep learning, amalgamates the discriminative power of supervised learning with the generalization capabilities of unsupervised learning. This approach has emerged as a focal point of research in the machine learning domain and has been employed to address a multitude of practical challenges, including, but not limited to, natural language processing [38], digital image processing [39–42], video tagging [43–45], and biometric recognition [46]. Semi-supervised learning leverages labeled data to generate pseudo-labels for unlabeled data, thereby enabling learners to optimize the generalization performance on the unlabeled data. Yue et al. [47] noted that the performance of a CNN significantly deteriorates when the labeled samples are insufficient. To effectively utilize the unlabeled samples, they presented a novel semi-supervised CNN method. It has been proven that their method can effectively improve the SAR ATR accuracy when the labeled samples are insufficient. According to Gao et al. [48], a semi-supervised learning method for object recognition in synthetic aperture radar (SAR) images was discussed. This method achieved state-of-the-art results on the Moving and Stationary Target Acquisition and Recognition (MSTAR) dataset, and it was shown that using generated images to train the networks can improve the recognition accuracy with a small number of labeled samples.

Our study employs the CycleGAN neural network to intelligently augment existing radar data, generating more representative radar data of diseases. This approach effectively addresses the issue of insufficient and unbalanced original radar datasets. However, the augmented dataset lacks annotation information. Training the network model using a supervised learning method would require manual annotation of the generated data. To circumvent this, we propose using semi-supervised learning in place of supervised learning, as illustrated in Figure 9. This method eliminates the need for manual annotation of the intelligently augmented radar data. Instead, both labeled and unlabeled data are concurrently input into the network for training. This approach not only establishes a model that satisfies accuracy requirements but also resolves the issue of insufficient annotation data.



**Figure 9.** Railway subgrade disease detection based on a semi-supervised learning method.

#### 4.2. Comparison of Supervised and Semi-Supervised Training Results

Initially, the radar data of manually annotated diseases (original dataset) are input into the YOLOv5 neural network for supervised training. The training, testing, and validation sets are partitioned in an 8:1:1 ratio, utilizing 500 radar data for training and allocating 62 each for testing and validation.

Table 3 displays the confusion matrix of the identification results for three types of railway subgrade diseases using the supervised training method and provides an analysis of the model's disease recognition capabilities. It can be observed that when employing the network model obtained through supervised learning for prediction, its accuracy rate is approximately 70.68%. This is due to the uneven distribution of data volume among the three diseases, leading to a low recognition accuracy of the trained model.

**Table 3.** Confusion matrix of three disease identification results in supervised learning.

| Confusion Matrix |                   | Reference   |            |                   |      |
|------------------|-------------------|-------------|------------|-------------------|------|
|                  |                   | Mud Pumping | Settlement | Water Abnormality | FP   |
| prediction       | Mud pumping       | 0.89        | 0.00       | 0.00              | 0.67 |
|                  | Settlement        | 0.00        | 0.73       | 0.00              | 0.33 |
|                  | Water abnormality | 0.00        | 0.00       | 0.51              | 0.00 |
|                  | TP                | 0.11        | 0.27       | 0.49              | 0.00 |

The original dataset is expanded using the CycleGAN neural network, obtaining a total of 1010 annotated and unannotated data (full dataset). The training, test, and validation sets are also divided in a ratio of 8:1:1, with 808 radar data used for training and 101 each for testing and validation. The annotated data and unannotated data are jointly input into the YOLOv5 network for semi-supervised training.

Table 4 presents the confusion matrix of the identification results for three types of railway subgrade diseases using the semi-supervised training method. It can be seen that

the recognition accuracy of the network model has significantly improved at this time. The recognition accuracy rate for mud pumping is 98%, for settlement is 94%, and for water abnormality is 91%. The recognition accuracy rate for the entire expanded dataset can reach 94.53%.

**Table 4.** Confusion matrix of three disease identification results in semi-supervised learning.

| Confusion Matrix |                   | Reference   |            |                   |      |
|------------------|-------------------|-------------|------------|-------------------|------|
|                  |                   | Mud Pumping | Settlement | Water Abnormality | FP   |
| prediction       | Mud pumping       | 0.98        | 0.00       | 0.00              | 0.88 |
|                  | Settlement        | 0.00        | 0.94       | 0.00              | 0.08 |
|                  | Water abnormality | 0.00        | 0.00       | 0.91              | 0.04 |
|                  | TP                | 0.02        | 0.06       | 0.09              | 0.00 |

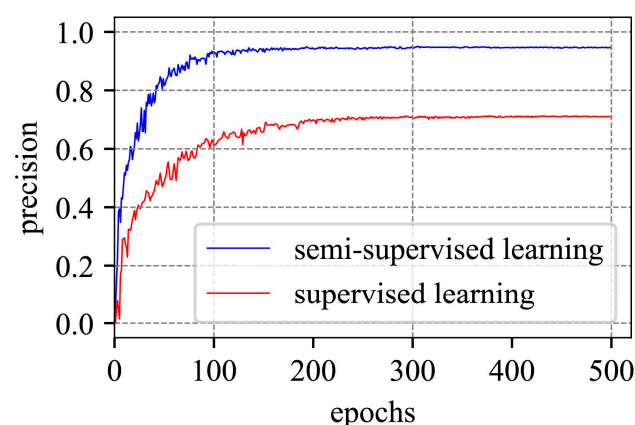
Precision and recall are used as evaluation indicators for training results. Precision represents the ratio of the number of correctly predicted samples in the dataset model to the total number of samples, while recall represents the proportion of correctly identified samples among all samples that need to be identified. The calculation formulas for these two indicators are shown in Equations (10) and (11), respectively.

$$\text{precision} = \frac{TP}{TP + FP} \times 100\% \quad (10)$$

$$\text{recall} = \frac{TP}{TP + FN} \times 100\% \quad (11)$$

In Equations (10) and (11), TP represents the number of samples correctly recognized by the model; FP represents the number of samples incorrectly predicted as positive, i.e., the number of false positives; and FN represents the number of samples incorrectly predicted as negative, i.e., the number of false negatives.

The performance of the two methods in terms of precision and recall is shown in Figures 10 and 11. From the training results, it can be seen that after intelligently expanding the radar data and using a semi-supervised deep learning training method, the established model has better performance in precision and recall, with an overall accuracy improvement of about 23%.

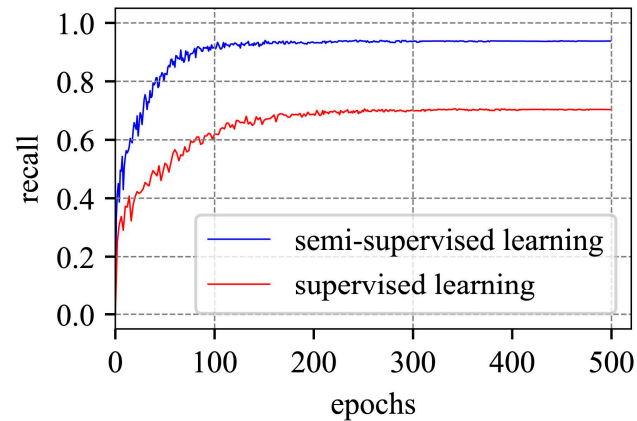


**Figure 10.** Comparison of precision between supervised learning and semi-supervised learning.

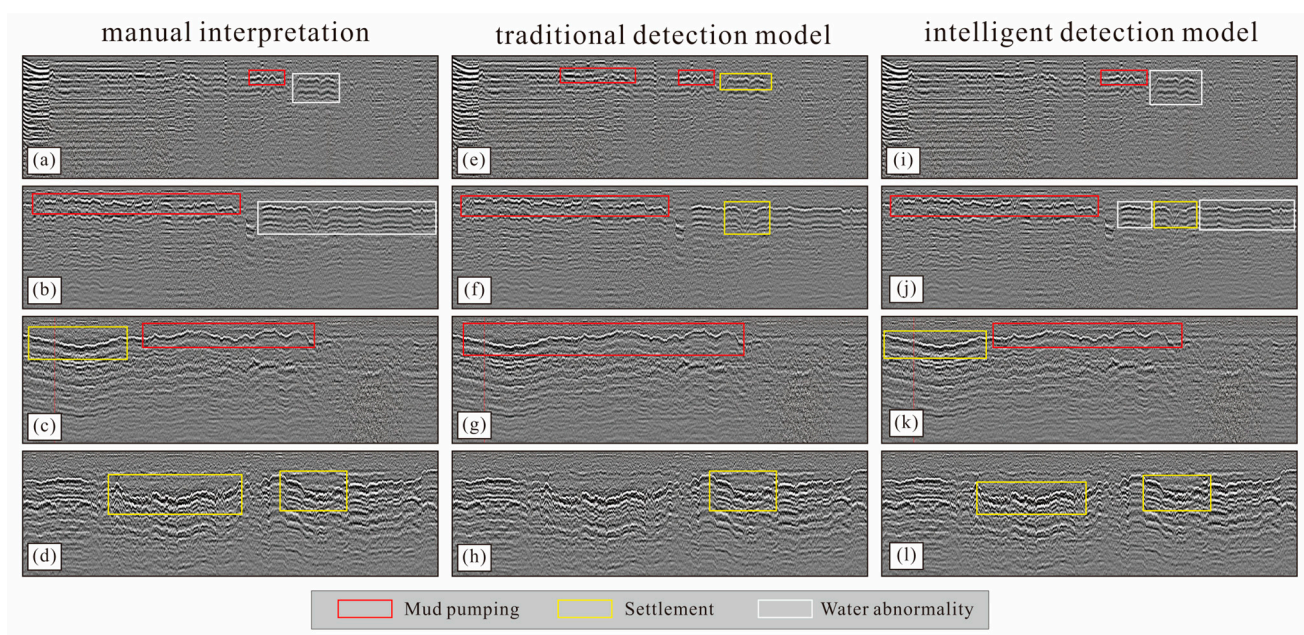
Figure 12 provides a comparative analysis of the detection efficacy between the conventional method and the intelligent method proposed in this study. The first column displays the results of manual interpretation, the second column presents the outcomes from the traditional model interpretation (the model is obtained by directly training the YOLOv5 network using a supervised learning method without data augmentation), and



the third column showcases the detection results of the proposed model. The intelligent model introduced in this study exhibits a superior consistency, with excellent interpretation results in terms of disease classification and localization. Furthermore, it demonstrates an enhanced accuracy and robustness compared to the traditional detection model.



**Figure 11.** Comparison of recall between supervised learning and semi-supervised learning.

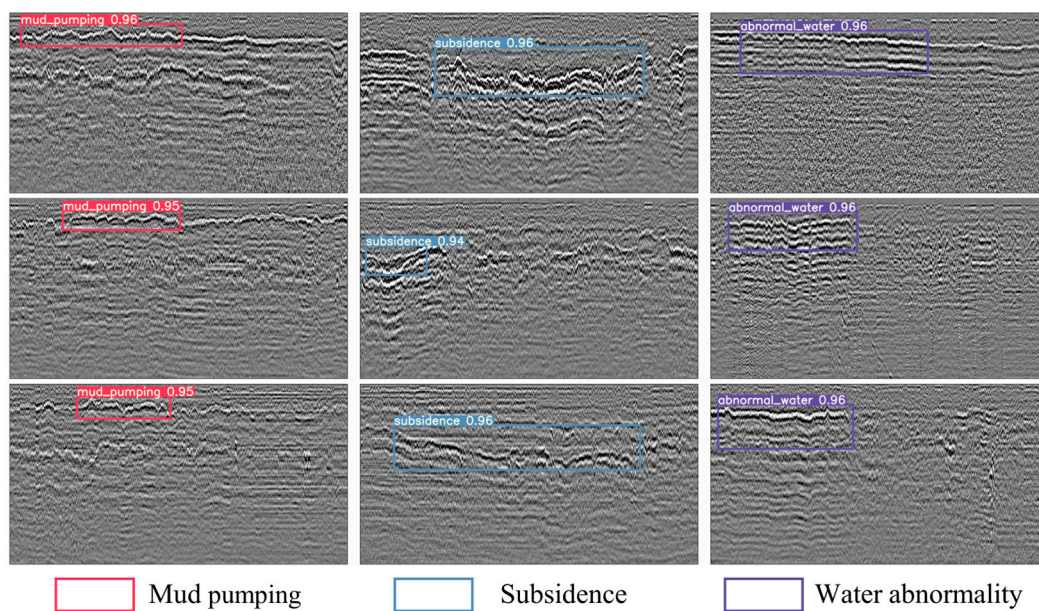


**Figure 12.** Comparison of the detection effects between supervised learning and semi-supervised learning. (a–d) Manual interpretation results; (e–h) supervised learning detection results; (i–l) semi-supervised learning detection results.

#### 4.3. Case Study

The model presented in this article was used to detect diseases in grayscale images of ground-penetrating radar profiles of a railway. Figure 13 illustrates the detection results, with the first column representing mud pumping, the second column representing settlement, and the third column representing water abnormalities. A disease-free radar spectrum exhibits a clear layered structure, a straight and continuous inphase axis, and a uniform reflected energy. When scanning sections afflicted with diseases, the waveforms in the graph display distinct characteristics.





**Figure 13.** The detection effect of the model on measured data. The red box represents mud pumping, the blue box represents settlement, and the purple box represents water abnormalities.

For instance, mud pumping disease usually occurs within the range of 0.05 to 0.5 m underground. Its radar image waveform is clearly different from the surrounding area from top to bottom, showing a chaotic, discontinuous, low-frequency strong reflection shape. This happens because there are tiny particles in the track bed and soil. These particles can become sticky under certain conditions. Under the action of water and the repeated load of trains, it softens and deforms, forming mud and causing changes in dielectric constant.

In scenarios where the rainfall surges due to climate change and the drainage capacity of the line is inadequate, subgrade settlement may occur due to excessive local water content. This manifests in radar images as a noticeably curved reflection of the same phase axis, with a downward shift in depth, interruption or discontinuity near the same depth position, and uneven ups and downs.

Water abnormality refers to the phenomenon in which the water content in a certain length of roadbed section is relatively large compared to adjacent sections. Due to signal attenuation when radar waves pass through water bodies, radar waves are strongly reflected on the surface of water-containing layers. The spectrum will show obvious brightening or phase inversion, and there may be multiple reflections.

Figure 13 demonstrates that the model proposed in this article can accurately identify and annotate the characteristics and locations of three types of diseases in radar images, exhibiting good robustness.

## 5. Conclusions and Perspective

This article presents a radar data augmentation method based on CycleGAN, and the network is trained using semi-supervised learning. The trained network model is applied to actual radar data. The CycleGAN neural network is first used to increase the quantity of images of two types of diseases, subgrade settlement and water abnormalities, to address the issue of poor training results caused by an insufficient data volume. A semi-supervised method is then employed to train the network, inputting annotated raw data and the expanded unlabeled data for training. This addresses the problem of the low recognition accuracy in the GPR field due to the difficulty of obtaining annotated radar data. Finally, by inputting the radar data to be identified into the trained network model, the specific locations of diseases in the radar image can be determined. During experimentation, the highest recognition accuracy of the model was 94.53%, demonstrating the effectiveness and practicality of the proposed method in practical engineering applications.

We utilize object detection networks to interpret radar data, with the goal of overcoming the subjectivity and inefficiency inherent in traditional interpretation methods. In practical applications, we frequently encounter challenges in creating datasets or managing imbalanced datasets for a variety of reasons. Our experimental validation demonstrates that the use of the CycleGAN neural network for data augmentation effectively mitigates data imbalances. Furthermore, data annotation during network training is a laborious and time-intensive task. By employing a semi-supervised training approach, we can construct a high-accuracy model with only a minimal amount of annotated data.

In future work, we plan to use more semi-supervised networks for model training, explore other networks that are more suitable for radar data detection than YOLOv5, and identify more suitable networks through comparative experiments. Additionally, due to the complexity of the actual working environment and the difficulty of obtaining annotated data, we have only expanded the data for two types of diseases: settlement and water abnormalities. Determining whether CycleGAN can be used to expand the data of more types of diseases and whether the detection coverage of the model can be increased are key issues for our future work.

**Author Contributions:** Conceptualization, Y.Y.; methodology, Y.Y. and L.H.; validation, Z.Z. and J.Z.; formal analysis, J.Z.; investigation, Y.Y. and Z.Z.; data curation, L.H.; writing—original draft preparation, Y.Y. and G.Z.; writing—review and editing, Z.Z. and J.Z.; visualization, G.Z.; supervision, Z.Z.; funding acquisition, Z.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by Sichuan Provincial Department of Science and Technology Project (2021YJ0031), Chengdu City Technology Innovation R&D Project (2022-YFO5-O0004-SN), and Open Topic of the National Engineering Research Center for Digital Construction and Evaluation Technology of Urban Rail Transit (2023KC01).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** No new data were created or analyzed in this study. Data sharing is not applicable to this article.

**Conflicts of Interest:** Author Guangmao Zhao was employed by the company China Railway Design Corporation. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. Li, Y.; Liu, H.; Wang, S.; Jiang, B.; Fischer, S. Method of Railway Subgrade Diseases (defects) Inspection, based on Ground Penetrating Radar. *Acta Polytech. Hung.* **2023**, *20*, 199–211. [\[CrossRef\]](#)
2. Wilk, S.T.; Li, D. A deep investigation into the mechanisms and factors producing mud pumping of railway track. *Transp. Geotech.* **2023**, *38*, 100908. [\[CrossRef\]](#)
3. Niu, F.; Lin, Z.; Lu, J.; Liu, H. Study of the influencing factors of roadbed settlement in embankment-bridge transition section along Qinghai-Tibet Railway. *Rock Soil Mech.* **2023**, *32* (Suppl. S2), 372–377.
4. Liu, S.; Lu, Q.; Li, H.; Wang, Y. Estimation of Moisture Content in Railway Subgrade by Ground Penetrating Radar. *Remote Sens.* **2020**, *12*, 2912. [\[CrossRef\]](#)
5. Feng, D.; Liu, Y.; Zhang, B.; Wang, X. Special Issue on Ground Penetrating Radar: Theory, Methods, and Applications. *Appl. Sci.* **2023**, *13*, 9847. [\[CrossRef\]](#)
6. Motevalli, Z.; Zakeri, B. Time-Domain Spectral Inversion Method for Characterization of Subsurface Layers in Ground-Penetrating-Radar (GPR) Applications. *Appl. Comput. Electromagn. Soc. J. (ACES)* **2019**, *34*, 93–99.
7. Dinh, K.; Gucunski, N.; Duong, T.H. An algorithm for automatic localization and detection of rebars from GPR data of concrete bridge decks. *Autom. Constr.* **2018**, *89*, 292–298. [\[CrossRef\]](#)
8. Artagan, S.S.; Bianchini Ciampoli, L.; D’Amico, F.; Calvi, A.; Tosti, F. Non-destructive Assessment and Health Monitoring of Railway Infrastructures. *Surv. Geophys.* **2019**, *41*, 447–483. [\[CrossRef\]](#)
9. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [\[CrossRef\]](#)
10. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [\[CrossRef\]](#)

11. Hinton, G.; Deng, L.; Yu, D.; Dahl, G.; Mohamed, A.-R.; Jaitly, N.; Senior, A.; Vanhoucke, V.; Nguyen, P.; Sainath, T.; et al. Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. *IEEE Signal Process. Mag.* **2012**, *29*, 82–97. [\[CrossRef\]](#)
12. Sutskever, I.; Vinyals, O.; Le, Q.V. Sequence to sequence learning with neural networks. *Adv. Neural Inf. Process. Syst.* **2014**, *27*.
13. Zhao, J.; Mao, X.; Chen, L. Speech emotion recognition using deep 1D & 2D CNN LSTM networks. *Biomed. Signal Process. Control* **2019**, *47*, 312–323.
14. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [\[CrossRef\]](#)
15. Silver, D.; Huang, A.; Maddison, C.J.; Guez, A.; Sifre, L.; Van Den Driessche, G.; Schrittwieser, J.; Antonoglou, I.; Panneershelvam, V.; Lanctot, M. Mastering the game of Go with deep neural networks and tree search. *Nature* **2016**, *529*, 484–489. [\[CrossRef\]](#)
16. Sun, Z.; Li, K.; Zheng, Y.; Li, X.; Mao, Y. Radar Spectrum Image Classification Based on Deep Learning. *Electronics* **2023**, *12*, 2110. [\[CrossRef\]](#)
17. Liu, B.; Ren, Y.; Liu, H.; Xu, H.; Wang, Z.; Cohn, A.G.; Jiang, P. GPRInvNet: Deep learning-based ground-penetrating radar data inversion for tunnel linings. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 8305–8325. [\[CrossRef\]](#)
18. Yue, Y.; Liu, H.; Meng, X.; Li, Y.; Du, Y. Generation of high-precision ground penetrating radar images using improved least square generative adversarial networks. *Remote Sens.* **2021**, *13*, 4590. [\[CrossRef\]](#)
19. Zhang, K.; Du, C. Intelligent identification of railway roadbed diseases based on unsupervised learning. In Proceedings of the 4th International Conference on Electronic Engineering and Informatics, Guiyang, China, 24–26 June 2022; EEI 2022; VDE: Berlin, Germany, 2022; pp. 1–4.
20. Xu, X.; Lei, Y.; Yang, F. Railway Subgrade Defect Automatic Recognition Method Based on Improved Faster R-CNN. *Sci. Program.* **2018**, *2018*, 4832972. [\[CrossRef\]](#)
21. Liu, H.; Wang, S.; Jing, G.; Yu, Z.; Yang, J.; Zhang, Y.; Guo, Y. Combined CNN and RNN Neural Networks for GPR Detection of Railway Subgrade Diseases. *Sensors* **2023**, *23*, 5383. [\[CrossRef\]](#)
22. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. *Commun. ACM* **2020**, *63*, 139–144. [\[CrossRef\]](#)
23. Redmon, J.; Farhadi, A. YOLO9000: Better, faster, stronger. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 7263–7271.
24. Redmon, J.; Farhadi, A. Yolov3: An incremental improvement. *arXiv* **2018**, arXiv:1804.02767.
25. Bochkovskiy, A.; Wang, C.-Y.; Liao, H.-Y.M. Yolov4: Optimal speed and accuracy of object detection. *arXiv* **2020**, arXiv:2004.10934.
26. Hao, W.; Zhili, S. Improved mosaic: Algorithms for more complex images. In *Journal of Physics: Conference Series*; IOP Publishing: Bristol, UK, 2020; p. 012094.
27. Wang, C.-Y.; Liao, H.-Y.M.; Wu, Y.-H.; Chen, P.-Y.; Hsieh, J.-W.; Yeh, I.-H. CSPNet: A new backbone that can enhance learning capability of CNN. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 14–19 June 2020; pp. 390–391.
28. Lin, T.-Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature pyramid networks for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2117–2125.
29. Rezaatofghi, H.; Tsoi, N.; Gwak, J.; Sadeghian, A.; Reid, I.; Savarese, S. Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 658–666.
30. Yu, J.; Jiang, Y.; Wang, Z.; Cao, Z.; Huang, T. UnitBox. In Proceedings of the 24th ACM International Conference on Multimedia, Amsterdam, The Netherlands, 15–19 October 2016; pp. 516–520.
31. Liu, L.; Ouyang, W.; Wang, X.; Fieguth, P.; Chen, J.; Liu, X.; Pietikäinen, M. Deep learning for generic object detection: A survey. *Int. J. Comput. Vis.* **2020**, *128*, 261–318. [\[CrossRef\]](#)
32. Zou, Z.; Shi, Z.; Guo, Y.; Ye, J. Object detection in 20 years: A survey. *arXiv* **2019**, arXiv:1905.05055. [\[CrossRef\]](#)
33. Zhu, J.-Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2242–2251.
34. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
35. Ulyanov, D.; Vedaldi, A.; Lempitsky, V. Instance normalization: The missing ingredient for fast stylization. *arXiv* **2016**, arXiv:1607.08022.
36. Isola, P.; Zhu, J.-Y.; Zhou, T.; Efros, A.A. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1125–1134.
37. Mao, X.; Li, Q.; Xie, H.; Lau, R.Y.K.; Wang, Z.; Smolley, S.P. Least Squares Generative Adversarial Networks. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2813–2821.
38. Nigam, K.; McCallum, A.K.; Thrun, S.; Mitchell, T. Text classification from labeled and unlabeled documents using EM. *Mach. Learn.* **2000**, *39*, 103–134. [\[CrossRef\]](#)
39. Tsuda, K.; Rätsch, G. Image reconstruction by linear programming. *Adv. Neural Inf. Process. Syst.* **2003**, *16*, 737–744. [\[CrossRef\]](#)

40. Zhou, Z.-H.; Chen, K.-J.; Jiang, Y. Exploiting unlabeled data in content-based image retrieval. In *Machine Learning: ECML 2004, Proceedings of the 15th European Conference on Machine Learning, Pisa, Italy, 20–24 September 2004*; Proceedings 15, 2004; Springer: Berlin/Heidelberg, Germany, 2004; pp. 525–536.
41. Zhou, Z.-H.; Chen, K.-J.; Dai, H.-B. Enhancing relevance feedback in image retrieval using unlabeled data. *ACM Trans. Inf. Syst. (TOIS)* **2006**, *24*, 219–244. [[CrossRef](#)]
42. Song, Y.; Zhang, C.; Lee, J.; Wang, F.; Xiang, S.; Zhang, D. Semi-supervised discriminative classification with application to tumorous tissues segmentation of MR brain images. *Pattern Anal. Appl.* **2009**, *12*, 99–115. [[CrossRef](#)]
43. He, J.; Li, M.; Zhang, H.-J.; Tong, H.; Zhang, C. Manifold-ranking based image retrieval. In *Proceedings of the 12th Annual ACM International Conference on Multimedia, New York, NY, USA, 10–16 October 2004*; pp. 9–16.
44. Yan, R.; Naphade, M. Semi-supervised cross feature learning for semantic concept detection in videos. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–26 June 2005*; IEEE: Piscataway, NJ, USA, 2005; pp. 657–663.
45. Tang, J.; Hua, X.-S.; Qi, G.-J.; Wang, M.; Mei, T.; Wu, X. Structure-sensitive manifold ranking for video concept detection. In *Proceedings of the 15th ACM International Conference on Multimedia, New York, NY, USA, 25–29 September 2007*; pp. 852–861.
46. Feng, W.; Xie, L.; Zeng, J.; Liu, Z.-Q. Audio-visual human recognition using semi-supervised spectral learning and hidden Markov models. *J. Vis. Lang. Comput.* **2009**, *20*, 188–195. [[CrossRef](#)]
47. Yue, Z.; Gao, F.; Xiong, Q.; Wang, J.; Huang, T.; Yang, E.; Zhou, H. A novel semi-supervised convolutional neural network method for synthetic aperture radar image recognition. *Cogn. Comput.* **2021**, *13*, 795–806. [[CrossRef](#)]
48. Gao, F.; Huang, Y.; Wang, J.; Sun, J.; Yang, E.; Zhou, H. A deep convolutional generative adversarial networks (DCGANs)-based semi-supervised method for object recognition in synthetic aperture radar (SAR) images. *Remote Sens.* **2018**, *10*, 846. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.