Article

Water Leakage and Crack Identification in Tunnels Based on Transfer-Learning and Convolutional Neural Networks

Ke Man [1], Ruilin Liu [1], Xiaoli Liu [2,*], Zhifei Song [1,*], Zongxu Liu [1], Zixiang Cao [1] and Liwen Wu [1]

1   School of Civil Engineering, North China University of Technology, Beijing 100144, China; man_ke@sina.cn (K.M.); liuruilin_ncut@163.com (R.L.); liuzongxu0127@163.com (Z.L.); cao_zixiang@163.com (Z.C.); wuliwen_ncut@163.com (L.W.)
2   State Key Laboratory of Hydroscience and Hydraulic Engineering, Tsinghua University, Beijing 100084, China
*   Correspondence: xiaoli.liu@tsinghua.edu.cn (X.L.); songzf@ncut.edu.cn (Z.S.); Tel.: +86-138-1180-4855 (X.L.); +86-186-1836-6770 (Z.S.)

Abstract: In order to solve the problems of long artificial time consumption, the inability to standardize the degree of damage, and the difficulty of maintaining data in traditional tunnel disease detection methods, this paper proposes the use of Residual Network (ResNet) models for tunnel water leakage and crack detection. ResNet proposes a residual learning framework to ease the training of networks that are deeper than those previously used. Furthermore, ResNet explicitly reformulates the layers as learning the residual functions of the reference layer inputs, rather than learning the unreferenced functions. The ResNet model is built on the Tensorflow Deep Learning (DL) framework and transfer-learning is used to optimize the model. The ResNet-V1 can be obtained by pre-training in ImageNet. The fully connected layers of the ResNet-V1 were modified to four classifications of tunnel disease. Then, the SoftMax function is used to recognize the tunnel diseases. Four network structures have been chosen, i.e., ResNet34 and ResNet50, with and without Transfer-learning, respectively. Those models were selected for testing and training on the sample dataset, and these four network structures were compared and analyzed using five types of evaluation indicators, which are the confusion matrix, accuracy, precision, recall ratio and F1. In identifying tunnel cracks and water leakage, the accuracy of ResNet50 and ResNet34 using the transfer-learning were 96.30% and 91.29%, and the accuracy of ResNet50 was 5.01% higher than that of ResNet34; for the network structure without the transfer-learning, the accuracy of ResNet50 was 90.36% and ResNet34's accuracy was 87.87%. These data show that the accuracy of ResNet50 is higher than that of ResNet34 with or without the transfer-learning, and the deep structure framework is superior in the identification of tunnel diseases; secondly, comparing the network structures with and without the transfer-learning, it can be found that using the Transfer-Learning can improve the ResNet network's accuracy for tunnel disease identification. The experiments and reliability analysis demonstrate the intelligent tunnel disease identification method proposed in this paper, and its good robustness and generalization performance. This method can be used for the rapid identification of cracks and water leakage in a tunnel survey, construction and maintenance, which has practical engineering implications for tunnel disease detection.

Keywords: tunnel diseases; deep learning; ResNet; transfer-learning; image augmentation

1. Introduction

With the rapid development of China's economy, the infrastructure construction of railways continues to advance rapidly, and a series of complex geological, oversized, buried, and extra-long tunnel projects are being or will be implemented [1]. Tunnel disease is a phenomenon in a tunnel that may prevent normal use. This includes water leakage, cracks and tunnel frost damage. These not only cause damage to the infrastructure in the tunnel but also significantly shorten the service life of the tunnel, so the accurate identification

of the tunnel disease is crucial to ensure tunnel safety [2,3]. Currently, the two most used approaches for detecting tunnel diseases are artificial identification and machine learning [4]. Traditional tunnel disease detection methods rely on artificial statistics, but artificial statistics is costly and slow, and the detection results are inevitably affected by human factors, making it difficult to obtain a uniform standard. Traditional artificial detection includes threshold segmentation [5], edge detection [6,7] and area growth [8], but this method is often constrained by the difference between the fracture pixels and the background pixels, and can only achieve better recognition results when the difference is obvious, and the choice of threshold value when using threshold segmentation methods can directly affect the recognition effect with large uncontrollable factors. In the field of machine learning, a method for crack edge detection based on the Gabor filter was presented by Roberto Medina [9], which is rotation-independent and can detect cracks independent of the growth direction; integral channel features were used to redefine the markers that constitute the cracks, which can achieve complex crack recognition [10]. While statistical machine learning can achieve better recognition results, it often requires the pre-processing of images to obtain abstract features of the tunnel disease, and the subsequent need for constant verification and modification to ensure that the validity of the features is a heavy workload [11–14]. The existing means of detecting tunnel engineering defects lags far behind current developmental needs. As artificial intelligence is an inevitable trend in the future of tunnel construction, the development of tunnel disease monitoring-technology based on intelligent fields has great practical engineering significance.

The deep learning method uses convolutional neural networks (CNNs) to extract features from tunnel diseases images, so the model learns to abstract tunnel diseases. This approach enables the automatic identification of tunnel cracks and tunnel leakage, making up for the shortcomings of traditional detection methods. DL is a branch of machine learning of which CNNs are an important method, and in recent years, CNNs have been used with great success in computer vision, natural language processing and speech [15–18]. With the continuous advancement of research in the field of geotechnical engineering, The deep supervised target detection network (DSOD) and ResNet model was combined to achieve intelligent recognition of rock properties, which has good robustness and generalization performance [19]. Full convolutional networks were used to learn the features of weak targets in a complex background and use residual learning in the network with acceleration network optimization, model parameters, etc., but the filter results generated by this method can cause the target signal increase in intensity and size. The expanded phenomenon causes the size and shape information of the original target to become lost [20]. Full convolutional network (FCN) was used to achieve the identification of water leakage in shield tunnels, and the method has superior robustness in overcoming pipeline occlusion [21]; in addition, the FCN was also used in the recognition algorithm, which not only achieves the identification of rock-concrete fractures, but also provides real-time statistics on their geometric length, width, and area [22]; the residual network structure was used to construct an over-prediction network model for depth feature extraction of water-rich fracture zones, which provides a basis for decision making in practical engineering [23]. At present, with the continuous development of DL, different model structures have been evolved and generated in CNNs, such as VGG, ResNet, FCN, R-CNN, etc. [24].

Different model structures can be used for tunnel disease identification in different fields to achieve better results, such as using Mask R-CNN to identify bridge cracks, using SC-Net to identify corrosion cracks in reinforced concrete [25,26], etc. These latest studies have achieved a good recognition effect. Considering the actual situation, tunnel cracks and water leakage account for a small proportion of the sample images, especially the crack images which do not obviously belong to the fine feature tunnel disease which requires deeper neural networks to extract more advanced abstract features. In this paper, the ResNet model structure is introduced, which cannot only use the deep network to extract small features but also solve the problem that the deep network is difficult to train.

The Tensorflow DL framework is used to build the ResNet model, a transfer-learning mechanism is introduced, and the ImageNet dataset is used to pre-train the ResNet model to obtain the pre-trained ResNet-v1 model, keeping the front-end parameters and weights of the model unchanged and modifying the fully connected layer (FC) of the model to adapt to the four divisions of tunnel damage in this paper, followed by a 4:1 ratio. The new prediction model was then tested using the divided test set. Finally, the test results were analyzed using five evaluation metrics: confusion matrix, accuracy, precision, recall, and F1.

2. DL Based Tunnel Disease Identification

CNNs are very widely used models in DL; they are feedforward neural networks with local connectivity and shared weights in which a large number of neurons are organized in a certain way to produce responses to overlapping regions in the visual field [24]. The ResNet model is one of the cross-connected models of convolutional neural networks, proposed by Kaiming He [27]. Deeper levels of the network are required to continuously extract more abstract features since water leakage and cracks in tunnels represent a small proportion of pixels in the picture. The ResNet model was chosen for the implementation of the detection of tunnel diseases to prevent the problem of vanishing gradients or exploding gradients in deeper network layers. The recognition of tunnel diseases by the model consists of two main steps: one is the establishment of the sample dataset, and another is the construction of the model structure.

2.1. Dataset Construction

Through the statistical analysis of tunnel diseases, cracks and water leakage are the most important diseases, representing more than 86% of the tunnel diseases. This paper therefore focuses on the analysis of cracks and water leakage and considers the fact that other equipment in the tunnel, such as lining joints and pipelines, may interfere with the detection of diseases. This paper divides the sample dataset into four categories, namely Label 1, Label 2, Label 3, and Label 4, as shown in Figure 1.

Label 1: Tube Sheet Splices   Label 2: Pipes

Label 3: Water Leakage   Label 4: Cracks

Figure 1. Classification of the sample dataset.

Adequate sample data optimizes the model in two ways: first, it improves the abstract representation of the features of the network model, and second, it improves the network's

ability to generalize data and avoid overfitting the model [28]. The majority of the tunnel disease image dataset used in this paper was obtained from random photographs taken by field personnel on the tunnel surface, and a small portion of the images were retrieved from the network, for a total of 600 sample data. Since the number of obtained samples is limited, image augmentation is used to increase the size of the dataset by making a series of random changes to the training images to produce similar but different training samples. For example, we can crop the image in different ways so that the objects of interest appear in different locations, thus reducing the model's dependence on where the objects appear: we can also adjust factors such as brightness and color to reduce the model's sensitivity to color, etc. The training images were rotated once every 30°, 12 times in total. Additionally, a total of 7600 image samples were then obtained after image augmentation, among which the image samples were divided into a training set (6080) and a test set (1520) by a category in a 4:1 ratio [29]. The classification and number of tunnel diseases is shown in Table 1.

Table 1. Classification and number of tunnel diseases.

Types of Tunnel Diseases	Sample Sets	Training Sets	Test Sets
Label 1	2100	1680	420
Label 2	2020	1616	404
Label 3	1780	1424	356
Label 4	1700	1360	340
Total	7600	6080	1520

2.2. Construction of ResNet Models in Tunnel Disease Identification

2.2.1. ResNet

In 1962, Hubel and Wiesel proposed the concept of receptive fields by studying the visual cortical cells of cats [30,31]; in 1979, Fukushima proposed a neurocognitive machine model based on the concept of receptive fields which is considered to be the first convolutional neural network implemented [32]; in 1998, LeCun et al., combined convolutional layers and in 2012, Krizhevsky et al. proposed the AlexNet convolutional neural network using modified linear units as the activation function and achieved excellent results in the ImageNet image evaluation [33,34], which is an important milestone in the history of DL development. The feature extraction process from input to output of a standard convolutional neural network is shown in Figure 2.
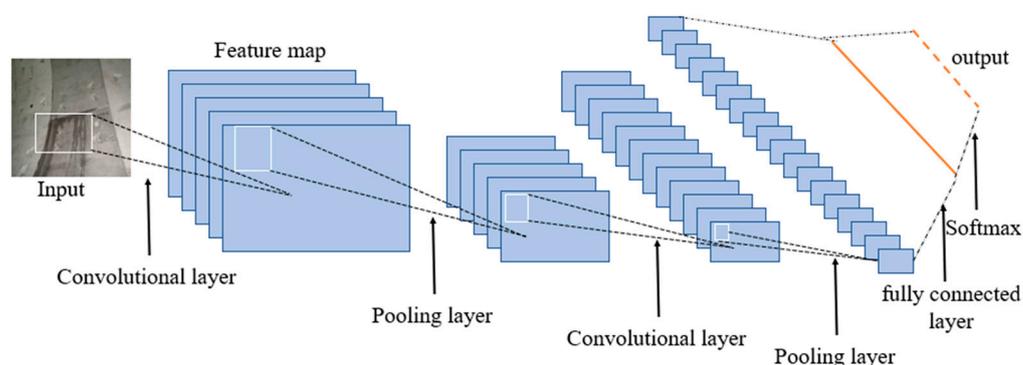

Figure 2. Standard convolutional neural network convolution process.

As shown in Figure 2, the convolutional neural network identifies the tunnel lesions in four main steps. Firstly, the input to the CNN is actually a three-dimensional neuron (one matrix for each of the color channels in RGB). The output of the convolutional layer is then obtained by multiplying the convolutional kernel with the digital matrix and then summing it, which achieves the initial extraction of the tunnel disease features. As the

feature maps obtained after convolution are still large, a pooling layer must be used to reduce the dimensionality of each feature map in order to retain the most useful image information. Finally, the extracted features are aggregated in a fully connected layer, and the score for each category is obtained by a classifier based on which the type of tunnel disease is predicted.

On the surface, the more layers of convolution there are in a CNN, the more likely the model is to acquire higher-level abstract features. However, overly deep CNNs are generally difficult to train, due to the exponential decay or growth of the cumulative backpropagation error signal as the number of layers of the neural network increases, whereas ResNet introduce the cross-layer connections and constructs the residual modules that solve the degradation problem compared to ordinary neural networks [35]. The following is a brief description of its principle in the general network back propagation process—the loss function and gradient values are shown in Equation (1) [36,37]:

$$
\begin{cases}
Loss = F(X, W) \\
\frac{\partial Loss}{\partial X} = \frac{\partial F(X,W)}{\partial X}
\end{cases}
\tag{1}
$$

The $F(X, W)$ is the loss function of the network and is used to estimate the inconsistency between the predicted value F of the model and the real value Y, where W is the parameter of the neural network and X is our given input. Taking the loss function as the optimization goal of the entire network, the optimized network parameter W is obtained by the multiple derivation of input X, and the error can reach the convergence state; finally, the model has a good prediction effect. Similarly, it extended the loss function of the network to a multi-layer neural network and the value of the gradient was introduced according to the chain rule, as shown in Equation (2) [36,37].

$$
\begin{cases}
Loss = F_n(X_n, W_n); L_n = F_{n-1}(X_{n-1}, W_{n-1}); \dots L_2 = F_1(X_1, W_1) \\
\frac{\partial Loss}{\partial X_i} = \frac{\partial F_n(X_n, W_n)}{\partial X_n} * \dots * \frac{\partial F_n(X_{i+1}, W_{i+1})}{\partial X_{i+1}}
\end{cases}
\tag{2}
$$

We can see that the gradient of the front layer network becomes smaller and smaller as the error is back propagated. ResNet cleverly introduces a residual structure, as shown in Figure 3.
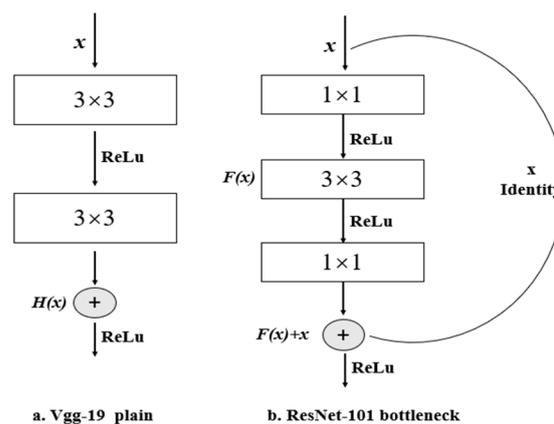

Figure 3. Comparison of general and residual networks.

From Figure 3, the ResNet modifies the output layer $H(X) = F(X)$ to $H(X) = F(X) + X$, at which point the gradient value of the multi-layer neural network changes from Equation (2) to Equation (3) [27]:

$$
\frac{\partial X_{i+1}}{\partial X_i} = \frac{\partial X_i + \partial F(X_i, W_i)}{\partial X_i} = 1 + \frac{\partial F(X_{i+1}, W_{i+1})}{\partial X_i}
\tag{3}
$$

We thus learn that the gradient does not disappear from Equation (3), even if the network is very deep.

2.2.2. Pre-Training Model ResNet_V1 Based on Transfer-Learning

Transfer-learning is a new type of machine learning approach that uses existing knowledge to solve problems in related domains, concentrating on the target task rather than learning all source and target tasks simultaneously [38]. In transfer-learning, the roles of the source and target tasks are no longer symmetrical, and unlike traditional machine learning, which has to be retrained for different target tasks, transfer-learning does not [39]. The difference between traditional machine learning and transfer-learning can be seen in Figure 4.
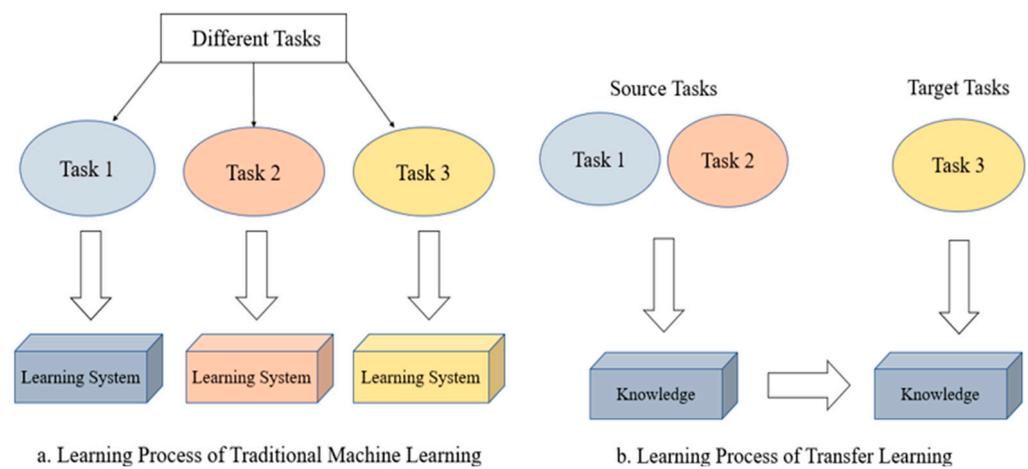

Figure 4. Different learning processes between (**a**) traditional machine learning and (**b**) transfer learning.

Transfer-learning can find the same parameter information between the source and target domains, allowing the data of the source domain and the data of the target domain share some model parameters. This can be applied to a new and different dataset through feature extraction and parameter tuning, thus avoiding costly data labelling efforts and greatly improving the learning performance. There are four main applications of transfer-learning: firstly, the new (target) dataset is small and similar to the original training dataset; secondly, the new (target) dataset is small and different from the original training dataset; thirdly, the new (target) dataset is large and similar to the original training dataset; and fourthly, the new (target) dataset is large and different from the original training dataset. We chose the second scenario in transfer-learning since the number of new training samples cannot reach our desired value. We added a new fully connected layer to the remaining pre-training layer with the number of classifications matching the tunnel disease category we set for the target dataset, then randomized the weights of the new FC layer and froze all the pre-trained network weights, and finally only the parameters of the new fully connected layer was trained after feature extraction. This approach avoids the problem of overfitting caused by the small size of the new dataset; this process is shown in Figure 5.
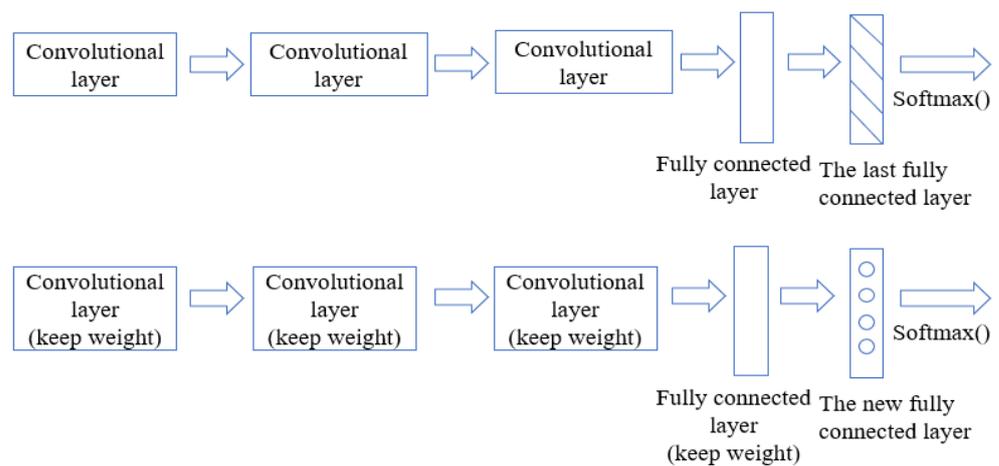
Figure 5. Transfer-learning with a small dataset and different data from the original.

ImageNet is a large visual database for visual object recognition software research which contains more than 14 million images covering more than 20,000 categories, among which more than a million images have explicit category annotations and annotations of object locations in the images [34]. Transfer-learning is generally performed in two ways: either fine-tuning or fixing weights [38]. Fixing weights is the way we choose to: use ResNet_V1 as the pre-training model obtained after training ResNet on the ImageNet dataset; fix the ResNet_V1 correlation layer and modify the network output layer for the type of tunnel disease recognition; use the method of fixed weights to avoid the update of the ResNet model parameters after each round of training by the structural layer after fine-tuning, enabling us to accelerate the network training. Tunnel disease identification is a multi-classification problem ($c > 2$) for which the SoftMax function was chosen as the classification layer of the convolutional neural network which takes the output of the fully connected layer as input and exponentiates it, the SoftMax function is defined as Equation (4) [34].

$$\begin{cases} V_i = x * w + b \\ S_i = \dfrac{e^{V_i}}{\sum_j^c e^{V_j}} \end{cases} \tag{4}$$

x is the input value of the previous layer of neurons; w is the weight; b is the bias; V_i represents the output of the classifier's pre-stage output unit; i indicates the category index; the total number of categories is C; S_i is the ratio of the index of the current element to the sum of the indices of all elements; SoftMax converts the output of multiclassification into a relative probability with a positive value and a sum of 1. In summary, the process of tunnel disease identification with the transfer-learning-based ResNet model in this paper is given in Figure 6.
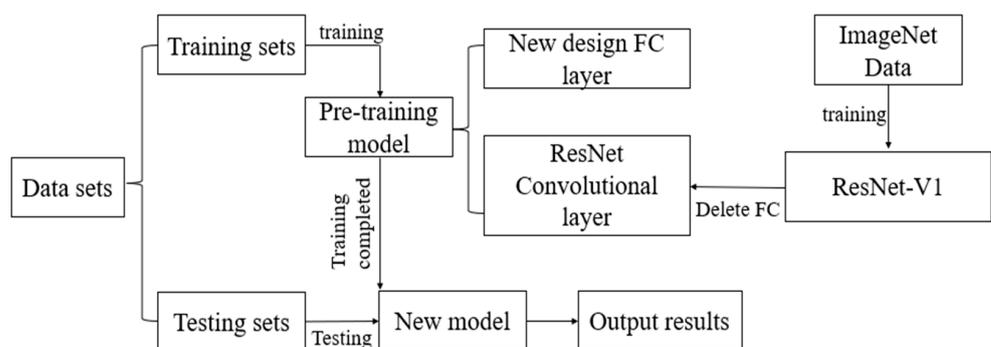

Figure 6. Transfer learning process for tunnel disease identification methods.

2.3. Model Evaluation Indicators

The confusion matrix directly reflects the model's classification of tunnel defects. In the sample dataset, we already know the labels of the images in the real case, and we can also use the known labels to determine which predictions the trained model makes on the test set are correct and which are wrong. For this, we can introduce four basic indicators: one is the number of true positives (TPs), where the true value is positive and the model considers it to be positive; second, the number of false negatives (FNs), for which the true value is positive and the model considers it to be negative; third, the number of false positives (FPs), for which the true value is negative and the model considers it to be positive; fourth, the number of true negatives (TNs), for which the true value is negative and the model considers it to be a negative quantity true negative (TN). Combining the above four indicators in a table gives the confusion matrix for the dichotomous problem, as in Table 2.

Table 2. Confusion matrix for binary classification problems.

Confusion Matrix		True Value	
		Positive	Negative
Predicted value	Positive	TP	FP
	Negative	FN	TN

We can learn that the larger the number of TPs and TNs and the smaller the number of FPs and FNs, the better the model predicts the classification. However, the statistics in the confusion matrix are individual numbers, and sometimes it is difficult to measure the recognition effect of the model just by counting the number of individuals in the face of a large amount of data. We therefore introduced four secondary indicators, namely accuracy (*ACC*), precision (*P*), recall (*R*) and F_1, to analyze the overall data to value the model [40].

ACC is expressed as the proportion of samples correctly classified by the model as a percentage of the total number of samples, for all categories of measure, as in Equation (5) [41]:

$$ACC = \frac{TP + TN}{TP + FP + FN + TN} \tag{5}$$

P is the number of samples for which the model predicts positive tunnel disease, as a proportion of the actual samples that are also positive. The precision rate measures the ability of the model to discriminate between negative samples; the higher the precision rate, the better the model is at discriminating between negative samples, as in Equation (6) [41]:

$$P = \frac{TP}{TP + FP} \tag{6}$$

R is the proportion of all samples that are actually positive cases that are predicted to be positive. The higher the recall, the better the model's ability to discriminate between positive samples, as in Equation (7) [41]:

$$R = \frac{TP}{TP + FN} \tag{7}$$

F_1 is the summed average of the precision and recall in the range [0, 1], and the higher the value is, then the more robust the model, as in Equation (8) [41]:

$$F_1 = \frac{2PR}{P + R} \tag{8}$$

3. Experimental Analysis of the Model

To verify the effectiveness of transfer-learning and model selection in this paper, two models, ResNet34 and ResNet50, were selected using a dataset of 7600 samples with and without transfer-learning base on the training set, and a test set was used to compare the recognition effectiveness of the four models. The superiority of the ResNet50 model for tunnel disease recognition was confirmed using confusion matrix, *ACC*, *P*, *R*, and F_1 score.

3.1. Training Model

The ResNet shallow network structures are ResNet18 and ResNet34, while the deeper ones are ResNet50, ResNet101, ResNet152, etc. The number after ResNet represents the number of layers that need to be updated with parameters through training. The residual module is introduced in both structures, the shallow network structure uses the BasicBlock residual module, and the deeper network structure uses the BottleNeck residual module, as shown in Figure 7.
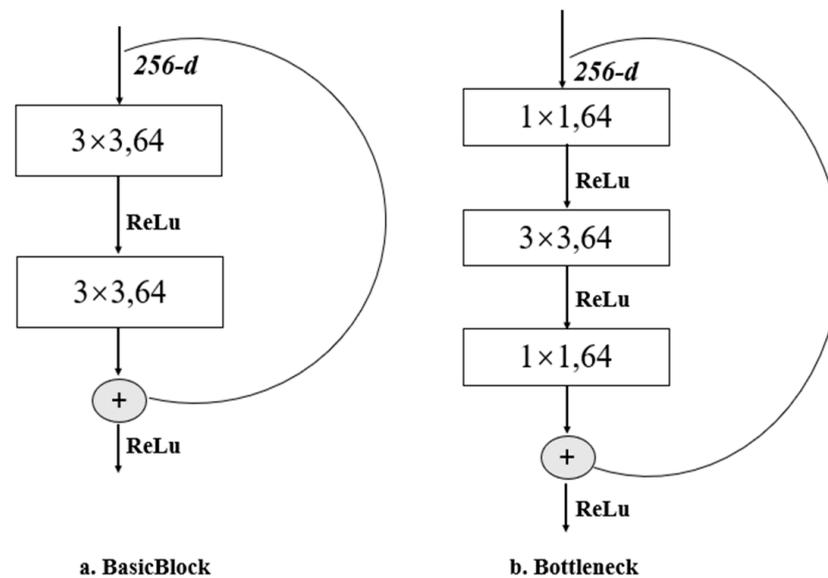

Figure 7. BasicBlock residual module and Bottleneck residual module.

ResNet34 and ResNet50 were selected as network structures for comparison: the number of training iterations was 5000, the learning rate was 0.0002, and the momentum decay parameter was 0.9. As the number of iterations increases, the accuracy rate shows an increasing trend, indicating that the ResNet network can be selected for the identification of tunnel diseases. Within a certain range, the number of iteration steps is positively correlated with the model recognition accuracy, and the test set image is tested once every iteration and its accuracy is recorded. The maximum accuracy of the transfer-learning based ResNet50 on the test set can be seen to be 96.30%, and the maximum accuracy of the transfer-learning-based ResNet34 on the test set is 91.29%; the accuracy of ResNet50 without transfer-learning-based learning was up to 90.36%, and the accuracy of ResNet34 without transfer-learning based learning was up to 87.87%, as shown in Figure 8. Indicating that the transfer-learning model of ResNet50 was better than ResNet34 at identifying tunnel diseases, and the transfer-learning improved the accuracy of the model in extracting the tunnel disease features, and the model robustness is better.
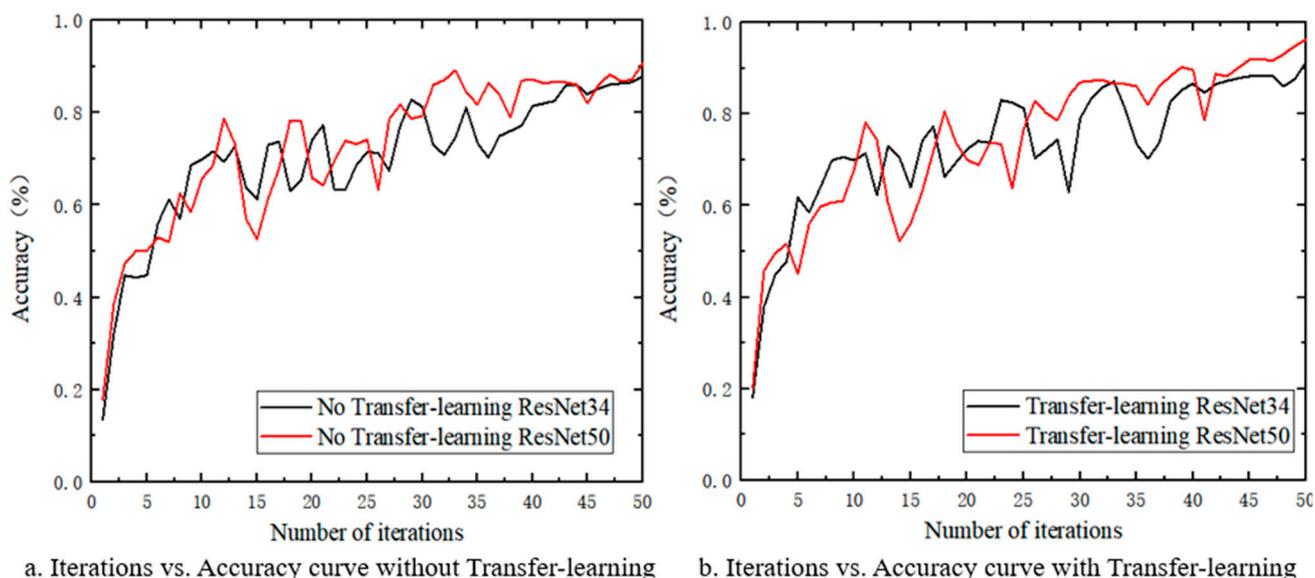
a. Iterations vs. Accuracy curve without Transfer-learning   b. Iterations vs. Accuracy curve with Transfer-learning

Figure 8. Different models with and without transfer-learning.

3.2. Results and Analysis

The transfer-learning-based ResNet34 and ResNet50 models were compared and the sample data identified on the tunnel distress test set were represented as a confusion matrix. As shown in Figure 9 ResNet34 identified water leakage with 27 samples identified as tunnel sheet splices and 13 samples identified as tunnel lines, while ResNet50 identified water leakage with 2 samples identified as tunnel sheet splices and 10 samples identified as tunnel lines; ResNet34 identified cracks with 27 samples identified as tunnel sheet splices and 3 samples identified as tunnel lines; ResNet34 identified cracks with 27 samples and 3 samples with tunnel lines, and ResNet50 identified cracks with 9 samples and 2 samples with tunnel lines. For sample 1 and sample 2, these were not tunnel diseases. When ResNet34 identified sample 1, 13 samples were incorrectly identified as tunnel lines, and 40 samples were incorrectly identified as tunnel cracks, while ResNet50 identified that in sample 1, 8 samples were incorrectly identified as tunnel lines, and 13 samples were incorrectly identified as tunnel cracks. Similarly for sample 2, when using ResNet34 for identification, 29 samples were incorrectly identified as the tunnel sheet splices, 4 samples were incorrectly identified as tunnel water leakage, and 20 samples were incorrectly identified as tunnel cracks. When using ResNet50, the number of wrong samples for identifying tunnel sheet splices and tunnel cracks is greatly reduced, and the number of wrong samples for tunnel leakage is 0. It can be seen that the transfer-learning-based ResNet50 is more stable when classifying tunnel deceases.

We further compared the transfer-learning-based ResNet34 and ResNet50 models on the labeled data using P, R, and F_1. As shown in Tables 3 and 4, the precision and recall rates of the transfer-learning-based ResNet50 are higher than those of ResNet34, and it can be seen that ResNet50 is more capable of distinguishing between the positive and negative samples. From Tables 3 and 4, it can be seen that R is smaller when P is larger and P is smaller when R is larger; indicating that P and R are in fact contradictory measures. This cannot allow for a comprehensive judgement of the model with a single comparison of P and R, so we introduce the value of F_1 that takes into account both P and R to allow both P and R to be maximized and balanced. As shown in Figure 10, the F_1 of the transfer-learning-based ResNet50 model is over 94% for each label, and is higher than that of ResNet34 for each label.
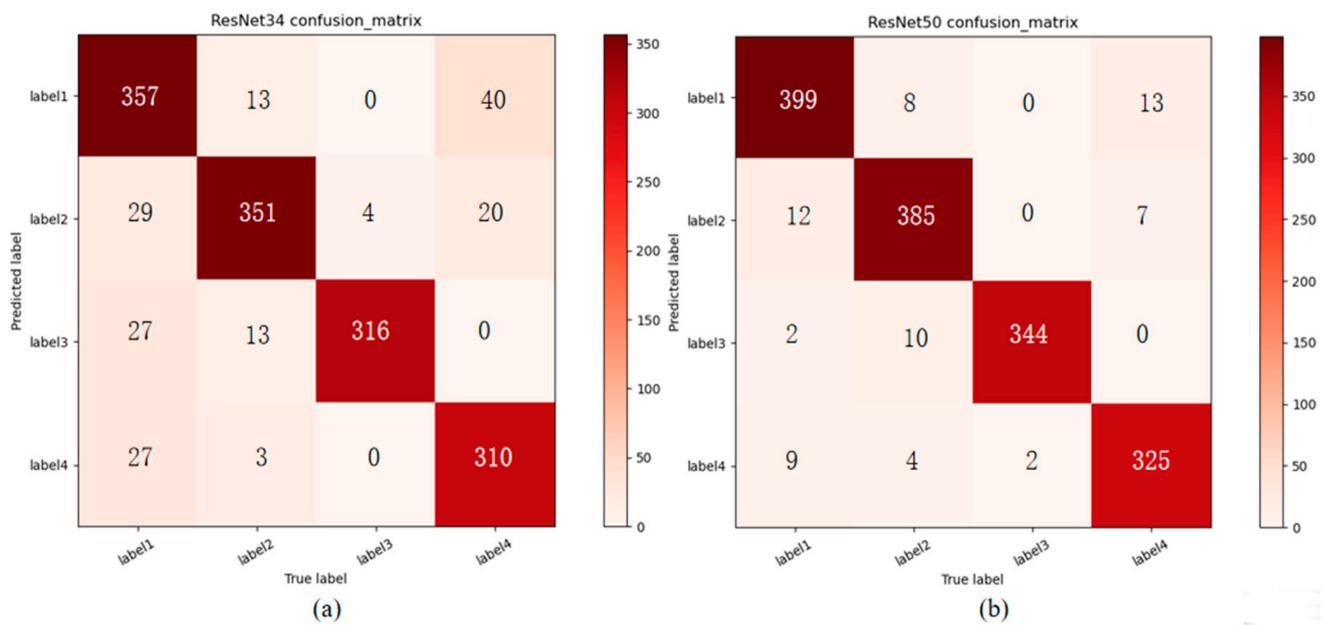
Figure 9. Confusion matrix of the transfer learning models of tunnel disease identification.

Table 3. Assessment indicators for each label transfer-learning ResNet34.

Labels	P/%	R/%	F_1
1	87.07	81.14	0.84
2	86.88	92.37	0.90
3	88.76	98.75	0.93
4	91.18	83.78	0.87

Table 4. Assessment indicators for each label transfer-learning ResNet50.

Labels	P/%	R/%	F_1
1	95.00	94.55	0.95
2	95.30	94.59	0.95
3	96.63	99.42	0.98
4	95.59	94.20	0.95

In summary, it can be seen that the transfer-learning-based ResNet50 has a better ability to identify tunnel diseases, and we can build a transfer-learning-based ResNet50 model and use convolutional neural networks to identify tunnel diseases.
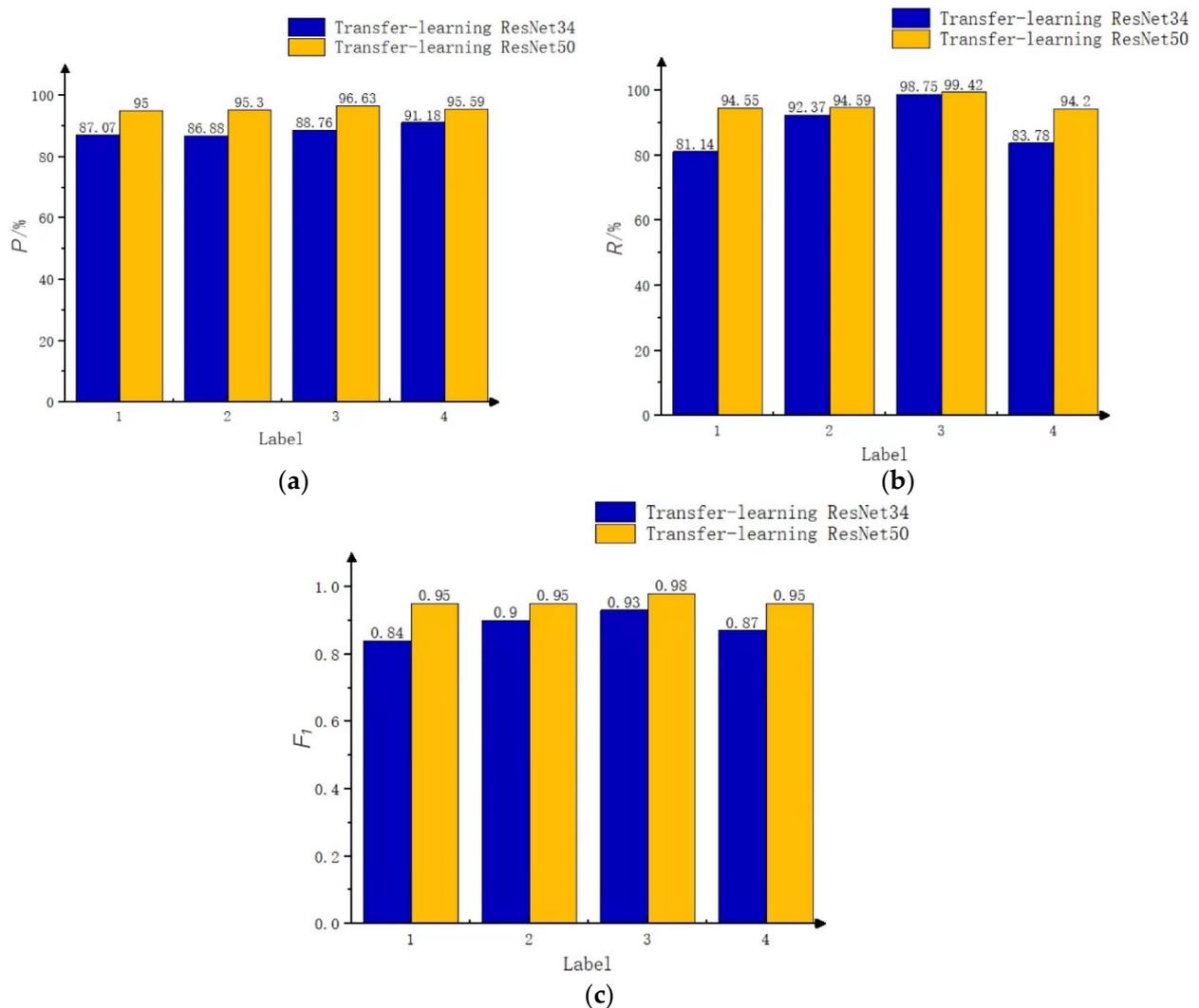
(a)

(b)

(c)

Figure 10. F_1 histogram based on transfer-learning.

4. Conclusions

With the development of artificial intelligence, deep learning has made great progress in tunnel engineering in recent years. Based on existing research results, this paper selects a network model that has not been previously used in the identification of tunnel diseases, which broadens the research direction for the identification of tunnel diseases. Using the unique cross-layer connection of ResNet, a residual module is constructed which realizes the fusion of local features and abstract features of cracks so as to retain detailed features of cracks and greatly improve the detection accuracy. Secondly, the method of transfer learning is introduced, combined with the model structure, and the main conclusions can be obtained through experimental verification as follows:

(1) In the model structure used in the past, the neural network can only identify a single tunnel disease, such as water leakage or cracks. The disadvantage of this method is that it needs to train two models which reduce the work efficiency and increase the time cost. In order to identify the two tunnel diseases of cracks and leaking water at the same time, a new fully connected layer with the same classification number as the tunnel diseases category was set up in this paper, which improves the efficiency of tunnel detection.

(2) The experimental data show that the accuracy rate of Res-Net50 without the transfer learning method is 90.36%, the accuracy rate of Res-Net50 with transfer learning method is 96.3%; the accuracy rate of Res-Net34 without transfer learning method is

87.87%; and the accuracy of Res-Net34 using the transfer learning method is 91.29%. It can be seen that using the transfer learning and fine-tuning training methods, a model with superior performance can be obtained in a shorter time.

(3) On the one hand, compared with the structure of the ResNet50 model and the ResNet34 model, the ResNet50 model is more complex, so it takes longer to train the ResNet50 network structure. On the other hand, according to the experimental data obtained from the test set, the five evaluation indicators of ResNet50 are higher than those of ResNet34, so the recognition accuracy of ResNet50 is the best. Considering that in practical engineering applications, due to the non-real-time change of cracks, engineering detection pays more attention to the model with the best crack recognition accuracy, so the improvement of the recognition rate has more practical use value.

(4) The research results of this paper can be applied to the automatic detection of tunnel cracks and water leakage in combination with photographic equipment which can not only reduce labor costs but also improve the efficiency of detection works. However, due to the complex tunnel environment, there are not only two types of tunnel diseases which are water leakage and cracks; in the next study, more types of tunnel diseases should be considered. Additionally, the way to obtain sample data in the article is too simple, mainly from on-site photography; in the future research, the sample data collected will come from different time and different places, and a mature database will be established to further improve the recognition effect of the model.

Author Contributions: K.M. contributed to the conception of the study; R.L. contributed to the write of the manuscript; X.L. helped perform the analysis with constructive discussions; Z.S. contributed significantly to analysis and manuscript preparation; Z.L., L.W. and Z.C. performed the experiment and the data analyses. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tian, S.; Wang, W.; Gong, J. Development and prospect of railway tunnels in China (including statistics of railway tunnels in China by the end of 2020). *Tunn. Constr.* **2021**, *41*, 308.
2. Liu, D.; Zhong, F.; Huang, H.; Zuo, J.; Xue, Y.; Zhang, D. Present Status and Development Trend of Diagnosis and Treatment of Tunnel Lining Diseases. *China J. Highw. Transp.* **2021**, *34*, 178–199.
3. Gong, X.; Guo, P. Prevention and Mitigation Methods for Water Leakage in Tunnels and Underground Structures. *China J. Highw. Transp.* **2021**, *34*, 1–30.
4. Zhou, F.; Jin, L.; Dong, J. Review of Convolutional Neural Network. *Chin. J. Comput.* **2017**, *40*, 1229–1251.
5. Wang, R.; Qi, T.; Lei, B.; Li, Y.; Zhu, X. Characteristic extraction of cracks of tunnel lining. *Chin. J. Rock Mech. Eng.* **2015**, *34*, 1211–1217.
6. Zhang, C.X.; Feng, C.; Chen, Z.; Hu, W.; Li, M. Parallel multiscale context-based edge-preserving optical flow estimation with occlusion detection. *Signal Process.-Image Commun.* **2022**, *101*, 14. [CrossRef]
7. Xuan, W.J.; Huang, S.L.; Liu, J.H.; Do, B. FCL-Net: Towards accurate edge detection via Fine-scale Corrective Learning. *Neural Netw.* **2022**, *145*, 248–259. [CrossRef]
8. Xue, Y.; Wang, P. Automatic recognition of cracks in tunnel lining based on characteristics of local grids in images. *Chin. J. Rock Mech. Eng.* **2012**, *31*, 991–999.
9. Medina, R.; Llamas, J.; Gómez-García-Bermejo, J.; Zalama, E.; Segarra, M.J. Crack Detection in Concrete Tunnels Using a Gabor Filter Invariant to Rotation. *Sensors* **2017**, *17*, 1670. [CrossRef]
10. Shi, Y.; Cui, L.; Qi, Z.; Meng, F.; Chen, Z. Automatic Road Crack Detection Using Random Structured Forests. *IEEE Trans. Intell. Transp. Syst.* **2016**, *17*, 3434–3445. [CrossRef]

11. Wang, W.; Tan, J. Rock joint image segmentation based on fractional differential and multi-grade combination in mathematical morphology. *J. Comput. Appl.* **2010**, *30*, 929–931, 942. [CrossRef]

12. Leng, B.; Qiu, W.; Wang, G.; Zhang, L. Research on Digital Image Processing Technology Used in Geological Analysis of Tunnel Engineering. *Railw. Stand. Des.* **2013**, *11*, 77–81.

13. Tao, K.; Wei, H.; Liao, M.; Wang, X. A Dam Surface Crack Damage Scale Intelligent Recognition Based on SVM Method. *Value Eng.* **2017**, *36*, 124–125.

14. Rahimzad, M.; Nia, A.M.; Zolfonoon, H.; Soltani, J.; Mehr, A.D.; Kwon, H.-H. Performance Comparison of an LSTM-based Deep Learning Model versus Conventional Machine Learning Algorithms for Streamflow Forecasting. *Water Resour. Manag.* **2021**, *35*, 4167–4187. [CrossRef]

15. Amiri-Ardakani, Y.; Najafzadeh, M. Pipe Break Rate Assessment While Considering Physical and Operational Factors: A Methodology based on Global Positioning System and Data-Driven Techniques. *Water Resour. Manag.* **2021**, *35*, 3703–3720. [CrossRef]

16. Shu, X.; Ding, W.; Peng, Y.; Wang, Z.; Wu, J.; Li, M. Monthly Streamflow Forecasting Using Convolutional Neural Network. *Water Resour. Manag.* **2021**, *35*, 5089–5104. [CrossRef]

17. Li, X.; Jiao, L.; Cao, X. Application of Residual Network Algorithm in the Identification of Tunnel Lining Diseases. *Softw. Guide* **2021**, *20*, 168–173.

18. Zhang, B.; Han, B.; Li, N.; Gan, G. Research progress on unmanned inspection technology and disease identification methods for long distance hydraulic tunnels during operation. *J. Basic Sci. Eng.* **2021**, *29*, 1245–1264.

19. Xu, Z.; Ma, W.; Lin, P.; Shi, H.; Liu, T.; Pan, D. Intelligent recognition of lithology based on migration learning of rock images. *J. Basic Sci. Eng.* **2021**, *29*, 1075–1092.

20. Yang, Q. Dim and Small Target Detection Based on Full Convolutional Recursise network. *Acta Opt. Sin.* **2019**, *46*, 36–43.

21. Huang, H.; Li, Q. Image recognition for water leakage in shield tunnel based on deep learning. *Chin. J. Rock Mech. Eng.* **2017**, *36*, 2861–2871.

22. Xue, D.; Tang, Q.; Wang, A.; Zhang, L.; Zhou, H. FCN-based intelligent identification of crack geometry in rock or concrete. *Chin. J. Rock Mech. Eng.* **2019**, *38*, 3393–3403.

23. Chen, P.; Yuan, Q.; Zhang, Z.; Yang, L.; Chen, Z.; Wu, L. A convolutional neural network-based method for image interpretation of geological over-prediction of water-rich fracture zones in tunnels. *J. Basic Sci. Eng.* **2022**, *30*, 196–207.

24. Goodfellow, I.; Bengio, Y.; Courville, A. *DL*; MIT Press: Cambridge, MA, USA, 2016; pp. 1–775.

25. Yu, J.; Li, F.; Xue, X.; Zhu, P.; Wu, X.; Lu, P. Intelligent Identification of Bridge Structural Cracks Based on Unmanned Aerial Vehicle and Mask R-CNN. *China J. Highw. Transp.* **2021**, *34*, 80–90.

26. Xu, Y.; Zhang, T.; Jin, G. Identification of Corroded Cracks in Reinforced Concrete Based on Deep Learning SCNet Model. *J. Hunan Univ. (Nat. Sci.)* **2022**, *49*, 101–110.

27. 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.

28. Fan, L.; Zhao, H.; Zhao, H.; Hu, H.; Wang, Z. Survey of target detection based on deep convolutional neural networks. *Optics and Precis. Eng.* **2020**, *28*, 1152–1164.

29. Cha, Y.J.; Choi, W.; Buyukozturk, O. DL-Based Crack Damage Detection Using Convolutional Neural Networks. *Comput. Aided Civ. Infrastruct. Eng.* **2017**, *32*, 361–378. [CrossRef]

30. Hubel, D.H.; Wiesel, T.N. Receptive fields, Binocular interaction and functional architecture in cats visual cortex. *J. Physiol.-Lond.* **1962**, *160*, 106. [CrossRef]

31. Hubel, D.H.; Wiesel, T.N. Receptive fields of single neurons in the cats striate cortex. *J. Physiol.-Lond.* **1959**, *148*, 574–591. [CrossRef]

32. Fukushima, K. NEOCOGNITRON—A self-organizing neural network model for a mechanism of pattern-recognition unaffected by shift in position. *Biol. Cybern.* **1980**, *36*, 193–202. [CrossRef]

33. Lecun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [CrossRef]

34. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. *Commun. Acm* **2017**, *60*, 84–90. [CrossRef]

35. Muthumani, I.; Malmurugan, N.; Ganesan, L. ResNet CNN with LSTM Based Tamil Text Detection from Video Frames. *Intell. Autom. Soft Comput.* **2022**, *31*, 917–928. [CrossRef]

36. Jiang, Z.; Qin, J.; Zhang, S. Parameterized Pooling Convolution Neural Network for Image Classification. *Acta Electron. Sin.* **2020**, *48*, 1729–1734.

37. Zhu, Y.; Guo, C.; Li, K.; Wu, X. Weights upgrade in error back propagation convolutional neural network. *J. Inf. Eng. Univ.* **2021**, *22*, 537–544.

38. Pan, S.J.; Yang, Q.A. A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [CrossRef]

39. Liu, Y.; Lei, Y.; Fan, J.; Wang, F.; Yan, G. Survey on Image Classification Technology Based on Small Sample Learning. *Acta Autom. Sin.* **2021**, *47*, 297–315.

40. Liu, C.; Wang, W.; Wang, M.; Lv, F.; Konan, M. An efficient instance selection algorithm to reconstruct training set for support vector machine. *Knowl.-Based Syst.* **2017**, *116*, 58–73. [CrossRef]
41. Rezaei-Ravari, M.; Eftekhari, M.; Saberi-Movahed, F. Regularizing extreme learning machine by dual locally linear embedding manifold learning for training multi-label neural network classifiers. *Eng. Appl. Artif. Intell.* **2021**, *97*, 104062. [CrossRef]