Innovative method for recognizing subgrade defects based on a convolutional neural network

Zheng Tong [a,b], Jie Gao [a], Haitao Zhang [b,*]

[a] School of Highway, Chang'an University, Xi'an 710061, China
[b] College of Civil Engineering, Northeast Forestry University, Harbin 150040, China

HIGHLIGHTS

- Accuracies of cascade and multi-stage CNNs classified subgrade defects according to training and validation results satisfy the requirements for subgrade assessment.
- Strategy using classifier 2 in cascade CNN improved robustness of object recognition in images obtained at different transmitting frequencies.
- CNN-based method using cascade CNN exhibited highly robust subgrade defect detection performance under various conditions compared with Sobel edge detection and K-value clustering analysis.

ARTICLE INFO

ABSTRACT

Subgrade defects originate below the base of an asphalt pavement and they contribute significantly to pavement damage. The detection of subgrade defects is considered challenging because the recognition of defects is difficult. Therefore, the utilization of ground penetrating radar (GPR) to detect subgrade defects has attracted significant interest in recent years. However, the use of manually processed GPR images for classifying defects is inefficient and inaccurate. Thus, in this study, we applied convolutional neural networks (CNNs) to GPR images for automatically classifying subgrade defects (e.g., uneven settlement, sinkholes, and subgrade cracks). Two CNNs called multi-stage CNN and cascade CNN with different structures were established to accomplish the tasks automatically. The main difference between the two CNNs is that the cascade CNN is a classifier 2, which is for recognition and trained only using hard samples. Each CNN was developed in training, validation, and testing processes. Based on the training and testing results, sensitivity analysis was performed to verify the stability of the CNNs. We compared state-of-the-art methods for defect detection and the CNN-based method in order to verify the superior performance of the CNNs. Finally, we tested an application of the CNN-based method to show that it is transferrable to other asphalt pavements. The training results indicated that the cascade CNN classified subgrade defects with 97.35% accuracy during training and 96.80% in validation, while the multi-stage CNN classified subgrade defects with 91.35% accuracy during training and 90.45% in validation. The sensitivity analysis results showed that the cascade CNN exhibited the expected stability in terms of the transmitting frequency, i.e., the frequency of a high-frequency electromagnet wave from the transmitting antenna of the GPR, and different highway structures, whereas the multi-stage CNN did not. In addition, compared with Sobel edge detection and K-value clustering analysis, the CNN-based method obtained more robust performance at subgrade defect detection under various conditions using raw images. These results indicate that the CNN-based method performs well and it can classify subgrade defects in realistic situations.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

The subgrade of an asphalt highway is an important component and the main part responsible for bearing loads on the pavement, including the pavement's weight and the vehicle load [1]. Due to the stress state of the subgrade, some defects are inevitable such

* Corresponding author.
E-mail address: zht6781@163.com (H. Zhang).

as uneven settlement, sinkholes, and subgrade cracks [2]. Subgrade cracks always occur in the subgrade before propagating to the surface but these subgrade defects cannot be observed directly. However, these defects always lead to pavement stresses and they affect the performance of highways. If effective measures are taken early to detect these defects in the subgrade, then the formation of pavement distress may be prevented. However, the detection of these defects is very challenging, mostly because they are located below the pavement.

In recent years, several innovative technologies have been used to detect subgrade defects, such as ground penetrating radar (GPR) technology [3–5], X-rays [6], and ultrasonic flaw detection [7]. GPR technology has some advantages such as its high efficiency, safe operation, nondestructive operation, and high anti-interference level. Therefore, GPR has been used widely to detect and evaluate highway structures. Significant progress has been made in the utilization of GPRs for detecting defects in subgrade structures, but this technology has several obvious disadvantages, i.e., its dependency on other auxiliary instruments for damage recognition, complicated data preprocessing requirements, and the difficulty of automatic defect recognition [8]. For example, Dong et al. [9] used a novel vehicle-mounted GPR to estimate the thicknesses of highway structure layers and to detect defects. The estimated results showed that the system and methods obtained satisfactory accuracy. This method must be assisted by a global positioning system and filtering algorithms. Khamzin et al. [10] used an air-launched GPR system comprising two truck-mounted GSSI antennae driven at near-highway speeds to assess highway structures and defects, but defects must be classified manually based on GPR data in this process. Szymczyk et al. [11] utilized an innovative S-transformation and successfully re-constructed GPR signals into three-dimensional models by using a complex conversion process. Yuri et al. [12] used the common offset and common midpoint method based on GPR data to analyze the composition of soil. Tosti and Benedetto [13] predicted pavement pumping using GPR, although this method can only recognize pavement pumping. In general, highway detection, especially subgrade defect detection, demands the development of an automatic defect recognition system that can use GPR data or images directly.

The methods described above partially address the aim of using GPR data or images but the following problems remain: (1) complex foreground, background, and feature information related to subgrade defects in GPR data cannot be handled easily; (2) the robustness of the algorithms can be affected by variable real-world situations (e.g., transmitting frequency of the GPR and the highway structure); and (3) human assistance is required to recognize the defects in images or data. Therefore, it is important to develop an automatic subgrade defect recognition system with sufficiently robust performance in variable real-world situations and under the influence of GPR noise.

Developments in machine learning have led to deep learning, especially convolutional neural networks (CNNs), which obtain good performance in the field of object recognition [14,15]. Lecun et al. [16,17] proposed the CNN as a type of artificial neural network with a structure where shared weights reduce the complexity of the network model. The structure is similar to that of biological neural networks [18–20]. Detailed information about CNNs can be found in previous studies [21–27]. In general, robustness is an attractive property of CNNs in civil engineering, where this robustness is evident in terms of their high stability at recognizing different objects such as humans and animals in different conditions. This property is important for defect recognition in GPR images where it is necessary to handle complex foreground, background, and defect features. Therefore, CNN can be utilized to classify different subgrade defects. In the last two years, CNNs have been used to detect highway defects, where Tong et al. [28] and Cha et al. [29]

employed CNNs to detect pavement cracks, while Cha et al. [30] used a fast CNN to detect multiple visual damage types in structures. The results of these three studies demonstrate that CNNs exhibit good stability in this field compared with conventional detection algorithms. Tong et al. [8] also used CNNs for the recognition, location, measurement, and three-dimensional reconstruction of concealed cracks based on GPR images. However, this CNN system could only distinguish concealed cracks from other types of distress, and thus it cannot be regarded strictly as defect recognition. Therefore, in the present study, we developed an application where we combined CNN with GPR images for defect detection in order to obtain highly accurate detection results based on an efficient detection procedure.

In this study, we aimed to employ CNNs to obtain appropriate models for automatically classifying subgrade defects using GPR images of asphalt pavements. The main advantages of the proposed CNN-based method are that it exhibits good stability with different transmitting frequency (300, 500, and 900 MHz), highway structures (five highways with different structures in Zhejiang Province, China), and various type of GPR noise (e.g., foreground and background). The remainder of this paper is organized as follows. In Section 2, we summarize the main approaches employed in this study, including the methods used for generating datasets for CNNs, building the CNN structures, training and testing the CNNs, and analyzing the stability and performance of the CNNs. The performance of the CNNs is discussed in Section 3, including the training and testing performance, stability analysis, and comparative studies. In Section 4, we present an application of a well-developed CNN in Zhejiang, China. We give our conclusions in Section 5.

2. Research approaches

Fig. 1 shows the main research procedure followed in this study. A collection method that employed an air-coupled GPR and a method for confirming different types of defects using core samples were utilized to prepare the datasets for our CNNs. Two CNNs were developed with different architectures based on GPR datasets to recognize subgrade defects. Fig. 2 shows the general developmental flow for the two CNNs. The two CNNs were implemented based on Caffe with an Intel (R) Core (TM) i7-6700 CPU, 8.00 GB random access memory (RAM), and an Nvidia GeForce GTX 1060 6 GB GPU.

2.1. Generating datasets for CNNs

The first step in the development of CNNs for recognizing subgrade defects was the generation of datasets. The complete dataset included training samples, testing samples, and their corresponding target samples.

A high-quality GPR dataset of subgrade defects was required to establish CNN models. The quality of a GPR dataset is often influenced by the acquisition equipment and pavement structures. An air-coupled GPR called OKO GPR was employed to capture GPR images. Three shielding antennae were utilized with 300, 500, and 900 MHz high-frequency electromagnet waves, where their vertical resolution ranges were 0.30–0.47 cm, 0.15–0.27 cm, and 0.09–0.13 cm, respectively. The vertical resolutions of the three shielding antennae met the engineering requirements [31]. The time sampling rate was 1024 scans/s and the distance sampling rate was 10 samples/cm. The height of the air-coupled antenna to the ground is 0.2 m.

A key factor that affects the image quality is the collection method employed, where the optimal collection method should consider both efficiency and precision. Detailed information about
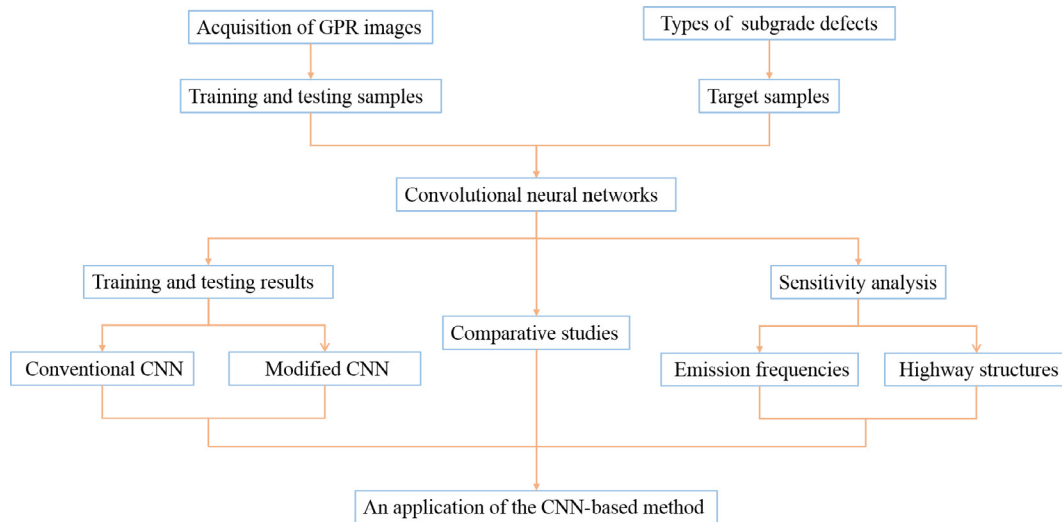
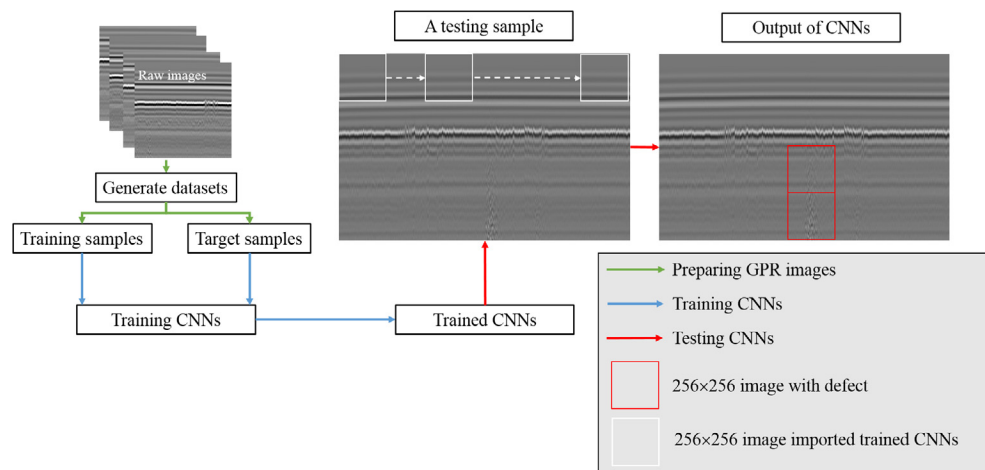**Fig. 1.** Flow chart illustrating the CNN development process.

Fig. 2. Flow chart illustrating the detection of subgrade defects.

the scanning process can be found in a previous study [8]. The antenna was activated to record an abnormal subgrade defect while its position is moved from location 1 to location 2. The distance between locations 1 and 2 was 4 cm, thereby ensuring that the GPR scanned and recorded more than four echo signals from the topmost point of a subgrade defect in order to obtain precise GPR images. The speed of the air-coupled GPR connected to a vehicle was 60 km/h to ensure the efficiency of GPR image acquisition. Denoising, wavelet analysis, and the detection of GPR images were performed automatically using LTD-2000 systems based on the GPR data collected. In total, 312 raw GPR images were acquired with different transmitting frequency (i.e., 267 images at 5120 × 3072 pixel resolution for training and 45 images at 5120 × 3072 pixel resolution for testing). Four GPR images are shown in Fig. 3. Clearly, the collection method could acquire GPR images containing subgrade defects with high efficiency and precision. In each of the 312 GPR images, the distance and the depth covered in a single GPR image were 25 m and 1 m, respectively.

The 267 images were cropped into small images each with a pixel resolution of 256 × 256. The smaller images were manually annotated as containing no defect, uneven settlement, sinkhole, or subgrade crack. The basis of the annotations for the core samples is shown in Fig. 3. Sinkholes were confirmed based on the

pavement conditions and core samples. Uneven settlement and subgrade cracks were confirmed based on core samples. Excavation was performed to confirm detection if the defects could not be determined based only on the core samples.

A target dataset was also used for developing the CNNs, which contained different types of defects with their corresponding GPR images. In this study, the four conditions present in the GPR images comprised subgrade with no defect, uneven settlement, sinkhole, and subgrade crack. Considering that quadrature encoding had advantages in terms of its high precision and fast response speed [28], these types of defects were quadrature encoded as $[1\,0\,0\,0]'$, $[0\,1\,0\,0]'$, $[0\,0\,1\,0]'$, and $[0\,0\,0\,1]'$ to represent "subgrade with no defect," "uneven settlement," "sinkhole," and "subgrade crack," respectively.

2.2. CNN

In this section, we explain the overall architectures of the two CNNs, including the details of each layer. The two CNNs were created using multiple layers, which mainly comprised the input, convolution, pooling, and output layers, and we explain the differences between the two CNNs in the following. Other important functions
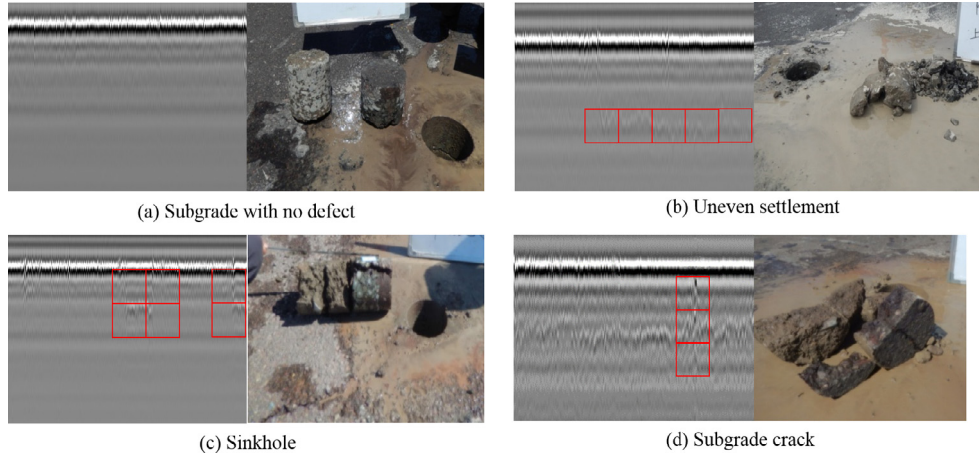
(a) Subgrade with no defect

(b) Uneven settlement

(c) Sinkhole

(d) Subgrade crack

Fig. 3. Typical GPR images and core samples.

and parameters used for network learning are described in Section 2.2.2.

2.2.1. Architecture description

The architectures of the two CNNs called the multi-stage CNN and cascade CNN are shown in Fig. 4. In general, a neural network is a mechanism for machine learning based on an input layer, multiple hidden layers, and an output layer. Each layer comprises several neuron nodes. A CNN is a type of the neural network with two special layers called the convolution layer and pooling layer. In this study, two CNNs were employed to classify GPR images.

(1) Architecture of the multi-stage CNN

The sizes of the sub-images and sequence of layers were based on previous studies [21,29]. The convolutional layer and pooling layer allow GPR images to be used directly where the input and feature maps are acquired in the convolutional layers. Several feature maps comprised the convolutional layers, as shown in Fig. 4 (a). We used specific sizes and numbers of kernels in the convolutional layers, as explained in our previous studies [28,32,33]. A feature map represented a special type of feature of subgrade defects based on the convolution operation with different kernels. A kernel can be regarded as a convolution filter that represents a small set of connection weights in a general neural network. As shown in Fig. 4(a), the multi-stage CNN had five kernels in C1 and 10 kernels in C2. The 15 kernels were initialized with random values and given the final values by network training. Each convolution result formed an element of the feature map in the next layer. The connections between the two layers were established in this manner. The procedure described above can be represented by Eq. (1):

$$I_j = sigmoid\left(\sum_{i=0}^{n} x_i \cdot w_{ij} + b_j\right) \tag{1}$$

where x_i denotes the ith feature map from the previous layer or the input image from the input layer, w_{ij} and b_j denote the weights and bias of the jth kernel, respectively, and $sigmoid(x)$ is a function given by the following equation.

$$sigmoid(x) = \frac{1}{1 + e^{-x}} \tag{2}$$

Pooling layers are another key feature of CNN where they reduce the spatial size of feature maps in a process known as downsampling or pooling. Some studies [29,34] have shown that the max pooling performance with image data sets is better than

that of other pooling methods. Thus, max pooling was utilized in the CNN architecture in this study. Max pooling takes the max values from a feature map, and an example of max pooling as shown in Fig. 5.

After the convolution and pooling operations, a layer is required to classify the input data based on the feature maps. Softmax is the main method used to classify input data. The procedure followed by the softmax layer in our multi-stage CNN can be represented by Eq. (3):

$$P(y_i = 4|x_i; W) = \begin{bmatrix} p(y_i = 1|x_i; W) \\ p(y_i = 2|x_i; W) \\ p(y_i = 3|x_i; W) \\ p(y_i = 4|x_i; W) \end{bmatrix} = \frac{1}{\sum_{j=1}^{n} e^{W_j^T x^{(i)}}} \begin{bmatrix} ce^{W_1^T x^{(i)}} \\ e^{W_2^T x^{(i)}} \\ e^{W_3^T x^{(i)}} \\ e^{W_4^T x^{(i)}} \end{bmatrix} \tag{3}$$

where $P(y_i = 4|x_i; W)$ denotes the probabilistic expression for the ith training sample and $W_{jx}^{T(i)}$ are the inputs for the softmax layer. The inputs for the softmax layer are the feature maps obtained from the pooling layers. The sum of the right-hand side is always 1 because the function always normalizes the distribution. Thus, Eq. (3) returns the probabilities of each input's individual classes.

(2) Architecture of cascade CNN

Hard samples have always been a problem for deep learning [34]. In the field of subgrade detection using GPR images, hard samples comprise those samples with low resolution. Samples with low resolution are collected using a low transmitting frequency and Tong et al. [32] showed that the accuracy of CNNs decreases with the transmitting frequency. However, a low transmitting frequency is sometimes used to collect subgrade information from a great depth. It is difficult for the network to converge when these types of samples are present. Thus, identifying these samples and training them individually is a good solution. Therefore, as described in a similar study [33], the multi-stage CNNs were cascaded as a new CNN called a cascade CNN, as shown in Fig. 4(b).

As shown in Fig. 4(b), a hard sample is selected based on its transmitting frequencies. In this study, a hard sample was defined as an image collected using a transmitting frequency below 500 MHz. Low transmitting frequencies yield images with low resolution so it is difficult to determine whether the defects in GPR images are sinkholes or subgrade cracks with only one classifier. These images were regarded as hard samples. After identifying the hard samples, two classifiers with the same structure as the multi-stage CNN were trained individually until the recognition accuracy exceeded a threshold. The main difference between the
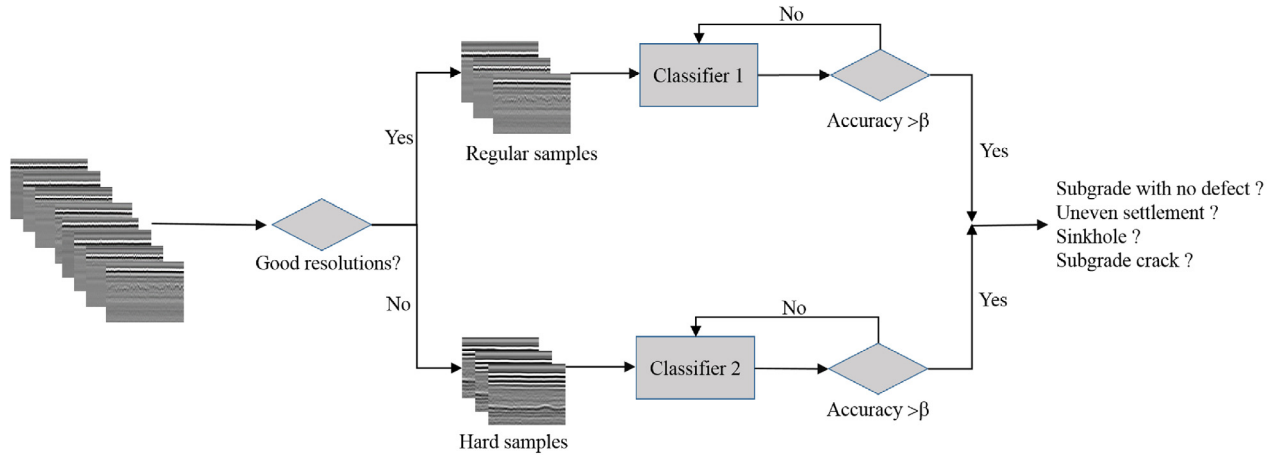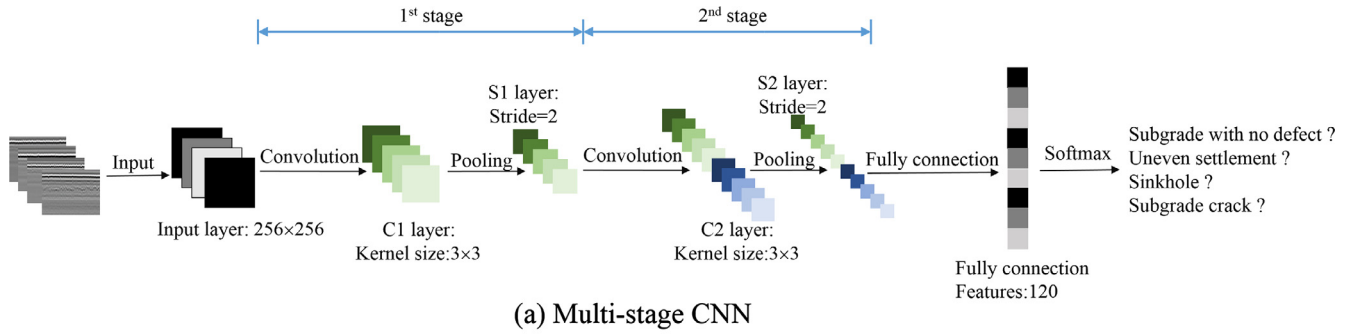
(a) Multi-stage CNN

Notes: Classifier 1 and classifier 2 have the same structure as the Multi-stage CNN

(b) Cascade CNN

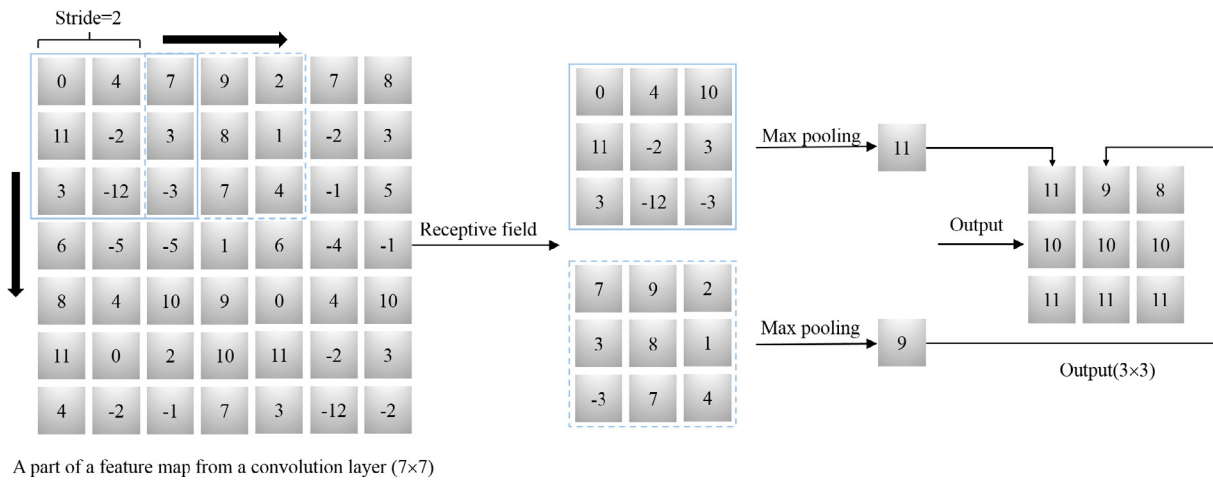**Fig. 4.** Structures of the two CNNs.

Fig. 5. Example of max pooling.

multi-stage CNN and the cascade CNN is classifier 2, which has the same structure as the multi-stage CNN but it is only trained using hard samples.

2.2.2. Network learning

The aim of network learning is to minimize the distance between the outputs of the CNNs and the target samples by adjust-ing the connecting weights and bias. As mentioned above, the ini-tial values of the weights and bias were randomly assigned before training. The outputs of the CNNs and the target classes did not coincide, and the difference was calculated with Eq. (4):

$$L = \frac{1}{m} \left[\sum_{i=1}^{m} \sum_{j=1}^{n} 1\{y_i = j\} \log \frac{e^{W_j^T x_i}}{\sum_{l=1}^{n} e^{W_l^T x_i}} \right] + \frac{\lambda}{2} \sum_{j=1}^{n} W_j^2 \quad (4)$$

where $1\{y_i = j\}$ is a logical expression that always returns either zeros or ones; in particular, if an output for the ith GPR image is true for j class, then the term returns ones; otherwise, zeros are returned; and $\frac{\lambda}{2}\sum_{j=1}^{n}W_j^2$ is a regularization term to prevent overfitting [36].

To narrow the deviations, the backpropagation algorithm was utilized for training. The connecting weights and bias were optimized using Eqs. (5) and (6), respectively,

$$W_{ij}(p+1) = W_{ij}(p) - \alpha_z \frac{\partial L}{\partial W_{ij}(p+1)} + \beta\left[W_{ij}(p) - W_{ij}(p-1)\right] \quad (5)$$

$$b_j(p+1) = b_j(p) - \alpha_j \frac{\partial E}{\partial b_j(p+1)} + \beta\left[b_j(p) - b_j(p-1)\right] \quad (6)$$

where $p-1$, p, and $p+1$ denote the p-1th, pth, and $p+1$th iterations during training, respectively, α_z and α_j are the learning rates with a range of $[0, +\infty]$, and β is a damping coefficient.

The procedure used for network learning with the backpropagation algorithm in our study is as follows.

Step 1. Give the stochastic connecting weights and bias to the initial CNNs.
Step 2. Select 500 training samples randomly for each iteration and then input them into the CNNs.
Step 3. Calculate L for 500 training samples.
Step 4. Adjust parameters W_{ij} and b_j with Eqs. (5) and (6).
Step 5. Select another 500 stochastic training samples and repeat steps 2–4 until L meets the requirement.
Step 6. Test the trained CNN with the testing dataset.

2.3. Stability and performance analysis

2.3.1. Stability analysis

The L results from training can only reflect the average accuracies of the two CNNs. As mentioned above, sensitivity analysis was required to verify the stability of the CNNs in different real-world conditions, including different transmitting frequency and highway structures.

(1) Sensitivity analysis with different transmitting frequencies.

The effects of different GPR transmitting frequencies were considered. In this study, three different transmitting frequencies, i.e., 300, 500, and 900 MHz, were used to verify the sensitivity of the CNNs. In the testing dataset, 45 GPR images were divided based on their transmitting frequency. There were 15 GPR images for each transmitting frequency. These images were cropped to obtain small images with 256×256 pixel resolution. These small images were then inputted into the two CNNs and the results were used to verify the sensitivity of the CNNs to the transmitting frequency.

(2) Sensitivity analysis with different structures

The effects of different highway structures were also considered. We collected 45 GPR images from five highways in Zhejiang Province, China, and their structures are shown in Table 1. There were 15 GPR images for each type of highway structures. These images were inputted into the two CNNs and the results were used to verify the sensitivity of the CNNs to the highway structures.

2.3.2. Comparisons with state-of-the-art methods for defect detection

We compared the performance of the CNN-based method with the state-of-the-art methods for detecting subgrade defects, where three different images from the testing dataset collected with different highway structures and transmitting frequencies were used.

We selected two well-known methods comprising Sobel edge detection and K-value clustering analysis for the comparisons. The results obtained using the three methods were employed to verify the superior performance of the CNN-based method. A previous study showed that Sobel edge detection obtained unsatisfactory results with digital images [29] but GPR images have different characteristics compared with digital images of highways. For example, GPR images contain a large amount of background noise information due to the properties of GPR, which may have negative effects on object recognition. Thus, Sobel edge detection was employed to show that CNN performed better as dealing with background noise information.

3. Results and discussion

3.1. Training and testing results

Unlike other image-based methods for detecting subgrade defects, we performed automatic feature extraction using network learning by adjusting the weights and bias of the kernels. Fig. 6 shows the training and validation results. The ratio of the number of training images relative to validation images was 4:1, and a total of 10,000 features were acquired from the 267 images with 5120 × 3072 pixel resolution. Thus, the training and validation accuracies were calculated as 8000 and 2000, respectively. As shown in Fig. 6, the network learning performance was exceptional. The final training and validation accuracies of the cascade CNN were 97.35% at the 144th iteration and 96.80% at the 145th iteration, respectively. The final training and validation accuracies of the multi-stage CNN were 91.35% at the 134th iteration and 90.45% at the 132nd iteration, respectively. The training and validation accuracies of the two CNNs satisfied the demands for subgrade assessment in China [31].

Fig. 7 shows images of the 1st field visualization (C1) for the multi-stage CNN and classifier 1 in the cascade CNN for a GPR image. As mentioned above, the multi-stage CNN and classifier 1 in the cascade CNN had the same structure and they were trained using the same dataset. Thus, they had the same kernels in the convolutional layers. The visualizations of each receptive field were regarded as feature maps of subgrade defects, where they could indicate whether the network needed more training and the types of features recognized by the well-trained CNNs [29]. In particular, the 1st field visualization (C1) for the multi-stage CNN and classifier 1 in the cascade CNN had five receptive fields, as shown in Fig. 4. Considering that sinkholes should produce a hyperbola in GPR images and that uneven settlement should produce no horizontal reflections, then features 2 and 3 were probably sinkholes and uneven settlement, respectively. Cracks (in the pavement or subgrade) should not affect the GPR image greatly, as shown by Lu et al. [35,37], so the shapes of concealed cracks in GPR images comprised a scattered hyperbola with symmetrical left and right sides. Feature 1 in Fig. 7 was probably a subgrade crack. In general, the two CNNs learned the features of defects via automatic training. Thus, the two CNNs had the capacity for automatic feature learning. In addition, the feature maps contained smooth patterns with some noisy features due to the complex and arbitrary patters of the asphalt highway structures. These noisy features may explain the erroneous assessments.

Training the multi-stage CNN with a GPU (Nvidia GeForce GTX 1060 6 GB) and a CPU (Intel (R) Core (TM) i7-6700 CPU) required about 10 s for one iteration and the approximate estimated running time on only one CPU was about 144 s. Similar times were required for training the cascade CNN. Therefore, the use of a GPU boosted the training speed. The two well-trained CNNs were used for sensitivity analysis to verify their stability.

Table 1
Highway structures.

Highway 1	Highway 2

Highway 1:
- 5 cm Asphalt concrete (AC-13)
- 6 cm Asphalt concrete (AC-16)
- 8 cm Asphalt concrete (AC-16)
- 8 cm Bituminous Stabilized Macadam
- 35 cm Cement stabilized gavels

Highway 2:
- 5 cm SBS Asphalt concrete (AC-16C)
- 7 cm SBS Asphalt concrete (AC-20C)
- 8 cm Bituminous Stabilized Macadam
- 30 cm Cement stabilized gavels

Highway 3	Highway 4

Highway 3:
- 4 cm SBS Asphalt concrete (AC-16C)
- 6 cm SBS Asphalt concrete (AC-20C)
- 9cm Bituminous Stabilized Macadam
- 32 cm Cement stabilized gavels

Highway 4:
- 4 cm SBS Asphalt Concrete (AC-13)
- 7 cm Asphalt concrete (AC-16)
- 8 cm Asphalt concrete (AC-25)
- 41 cm Cement stabilized gavels
- 20 cm Natural gravel

Highway 5

Highway 5:
- 4 cm Stone mastic asphalt (SMA-13)
- 7 cm Asphalt concrete (AC-16)
- 8 cm Asphalt concrete (AC-25)
- 17 cm Cement stabilized gavels
- 12cm Natural gravel

3.2. Sensitivity analysis results

The average accuracies of the two CNN reflected their overall performance. As mentioned above, sensitivity analysis was performed to verify the stability of the CNNs in different conditions, including different transmitting frequency and highway structures.

3.2.1. Recognition accuracy with different transmitting frequency

The 45 GPR images in the testing dataset were divided based on their transmitting frequency (i.e., 300, 500, or 900 MHz). There were 15 GPR images for each transmitting frequency. These images were cropped to obtain small images at 256 × 256 pixel resolution. The 10,800 small images were inputted into the two CNN. The classification accuracy rates are shown in Figs. 8 and 9.
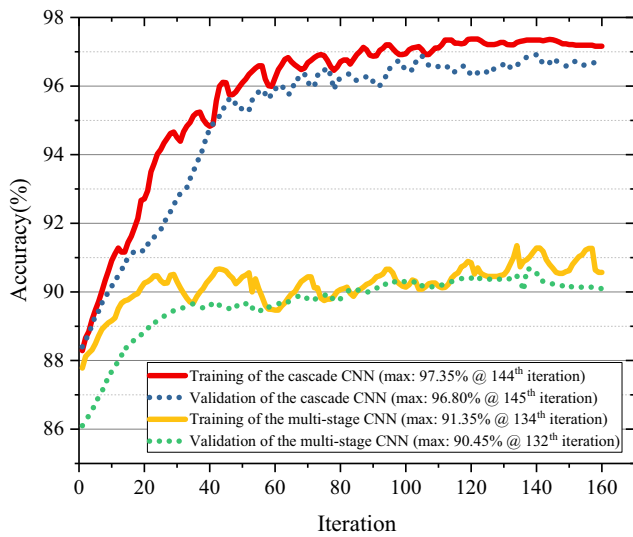
Fig. 6. Accuracy for each iteration.

Fig. 8 shows that the classification accuracies obtained by the multi-stage CNN with transmitting frequencies of 300, 500, and 900 MHz were 78.87%, 92.33%, and 95.28%, respectively. Detailed information about the recognition rates is shown in the confusion matrices. The confusion matrices demonstrate that the recognition rates varied significantly, thereby showing that the classification results were influenced by transmitting frequencies using the multi-stage CNN. Thus, we can conclude that the multi-stage CNN exhibited instability that depended on the transmitting frequencies. The low transmitting frequencies had a negative influence on the performance of the multi-stage CNN because less detail of the subgrade defects was present in the GPR images as the transmitting frequencies decreased. Thus, some defect features could not be acquired after the convolution operation and negative changes also occurred at the edges of defects. Due to this loss of important feature information, the results obtained by the softmax layer deviated from the actual results. Super-resolution solutions such as regularized deconvolution [38] may be used to address this loss of important feature information. If these methods are utilized in GPR with a low transmitting frequency, such as 300 MHz, then the accuracy of the multi-stage CNN may be improved. However, the improvement might not be significant if the transmitting frequency is high. The utilization of super-resolution solutions to improve the accuracy of the multi-stage CNN should be addressed in future research.

Fig. 9 shows that the classification accuracies of cascade CNN with transmitting frequencies of 300, 500, and 900 MHz were 95.36%, 95.78%, and 95.89%, respectively. Detailed information about the recognition rates is shown in the confusion matrices. According to the confusion matrices in Fig. 9, the recognition rates did not differ greatly, thereby demonstrating that the classification results obtained by cascade CNN were not influenced significantly

by the transmitting frequency. Thus, we can conclude that cascade CNN exhibited stability with the transmitting frequency, and classifier 2 in cascade CNN, which was only trained using hard samples, could improve the robustness of object recognition with various transmitting frequencies.

In general, cascade CNN performed better in terms of both the accuracy and its robustness to different transmitting frequencies compared with multi-stage CNN. Therefore, cascade CNN was used for the sensitivity analysis in order to verify its stability with different highway structures.

3.2.2. Recognition accuracy with different highway structures

We collected 45 GPR images from five highways in Zhejiang Province, China, and their structures are shown in Table 1. There were 15 GPR images for each highway structure. The testing images obtained at 900 MHz were utilized to verify the stability of cascade CNN with different highway structures. The results of the classification accuracy rates are shown in Fig. 10.

Fig. 10 shows that the classification accuracy rates using cascade CNN with highway structures 1–5 were 96.61%, 95.69%, 96.25%, 95.67%, and 95.23%, respectively. Detailed information about the recognition rates is shown in the confusion matrices. According to the confusion matrices in Fig. 10, the recognition rates did not vary greatly, so the classification results obtained by cascade CNN were not influenced by the different highway structures. Thus, we conclude that cascade CNN exhibited stability with different highway structures. Therefore, cascade CNN was used to demonstrate its superior performance compared with the state-of-the-art methods for defect detection.

3.3. Comparative studies

In order to compare the performance of the CNN-based method using cascade CNN with the state-of-the-art methods for detecting subgrade defects, we used three images from the testing dataset with different highway structures and transmitting frequencies. We selected two well-known methods comprising Sobel edge detection [39] and K-value clustering analysis [40] for the comparison. The results obtained using these three methods were employed to verify the superior performance of the CNN-based method.

The first image from Highway 1 had a resolution of 5120 × 3072 pixels and a transmitting frequency of 900 MHz, as shown in Fig. 11(a). The CNN-based method achieved accurate subgrade crack recognition and localization, as shown in Fig. 11(b). Sobel edge detection and K-value clustering analysis obtained some information about the subgrade crack, as shown in Fig. 11(c) and (d), but they could only recognize part of the information about the subgrade defects.

The second image from Highway 4 had a resolution of 5120 × 3072 pixels and a transmitting frequency of 300 MHz, as shown in Fig. 12(a). Similar to the first image, the CNN-based method achieved accurate uneven settlement recognition and localization, as shown in Fig. 12(b). However, Sobel edge detection and K-value clustering analysis obtained little information about the subgrade
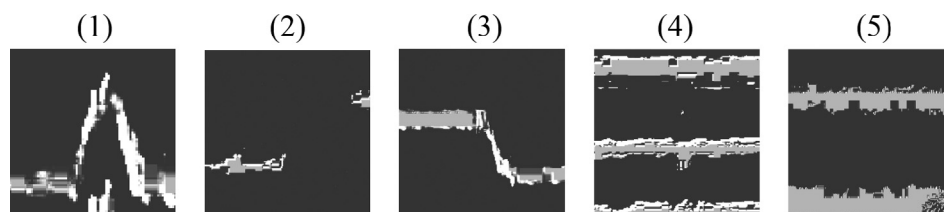

Fig. 7. Learned features.
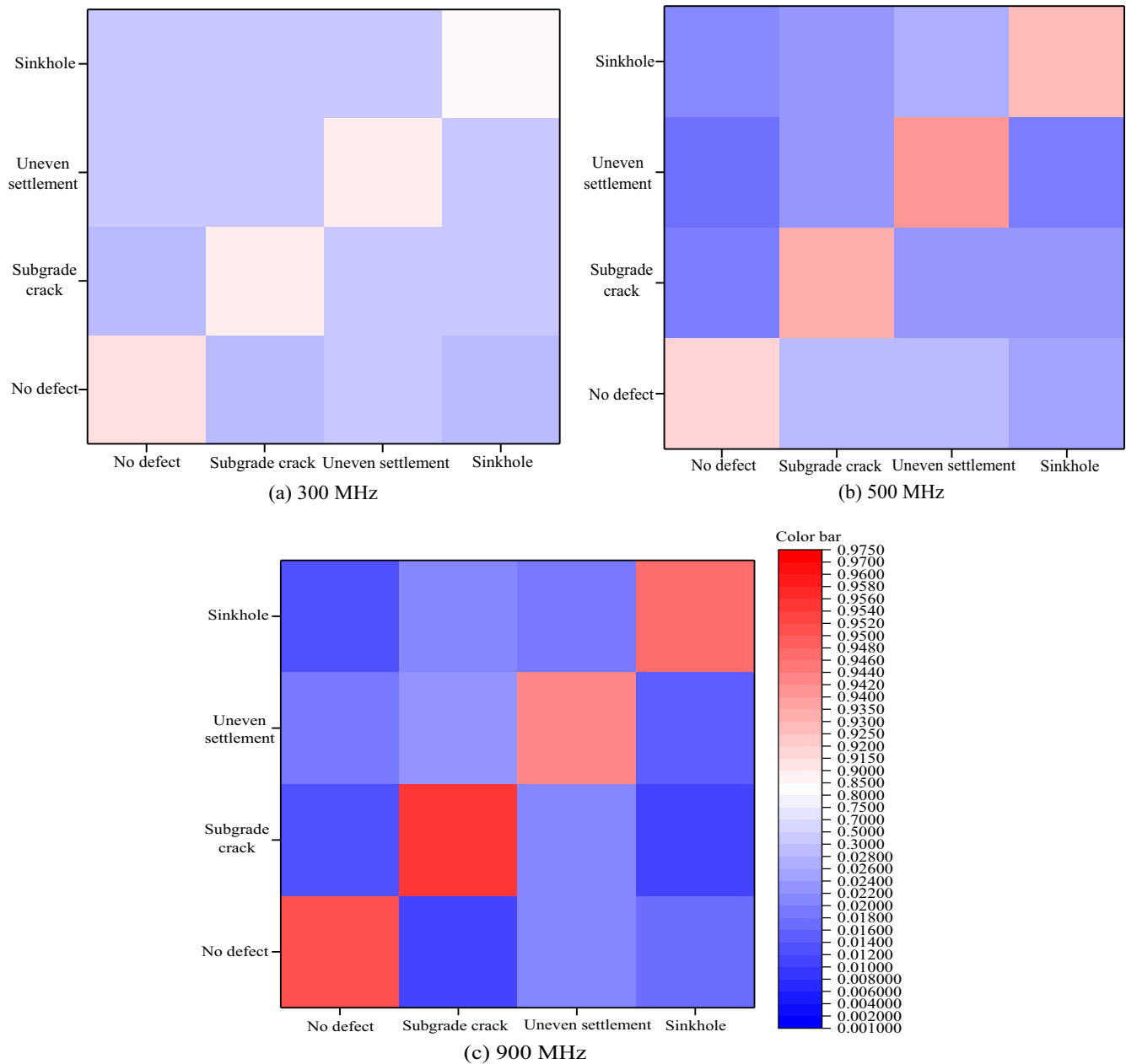
(a) 300 MHz

(b) 500 MHz

(c) 900 MHz

Fig. 8. Confusion matrices based on the testing dataset for the multi-stage CNN with different transmitting frequency, i.e., 300, 500, and 900 MHz. The X-axis labels are the ground truth labels and the Y-axis labels are the predicted labels.

defect because the performance of these methods is highly dependent on the image quality, whereas the CNN-based method was affected little by the quality of the image.

The third image from Highway 3 had a resolution of 5120 × 3072 pixels and a transmitting frequency of 500 MHz. This image contained two defects comprising uneven settlement and a sinkhole, as shown in Fig. 13(a). The CNN-based method achieved accurate recognition and localization for the two defects, as shown in Fig. 13(b), but Sobel edge detection and K-value clustering analysis only obtained some information about the defect types. Thus, the CNN-based method could detect different defects, whereas Sobel edge detection and K-value clustering analysis could not. The CNN-based method with different kernels could detect different defect features, as shown in Fig. 7, but Sobel edge detection and K-value clustering analysis could only detect one feature. In addition, the CNN-based method could detect the defects in each small image with a resolution of 256 × 256 pixels, as shown in Fig. 2,

whereas Sobel edge detection and K-value clustering analysis could only detect global defects in an image.

According to these comparative studies, the CNN-based method using the cascade CNN exhibited more robust performance at subgrade defect detection under various conditions based on the raw images compared with Sobel edge detection and K-value clustering analysis.

4. Application of the CNN-based method

To verify its transferability to other asphalt pavements, the CNN-based method using cascade CNN was employed to detect subgrade defects in practical cases according to the processes explained in Sections 2 and 3. Four highways in Zhejiang, China, were observed and 50 raw images were acquired from each road according to the collection method explained above. The 50 images
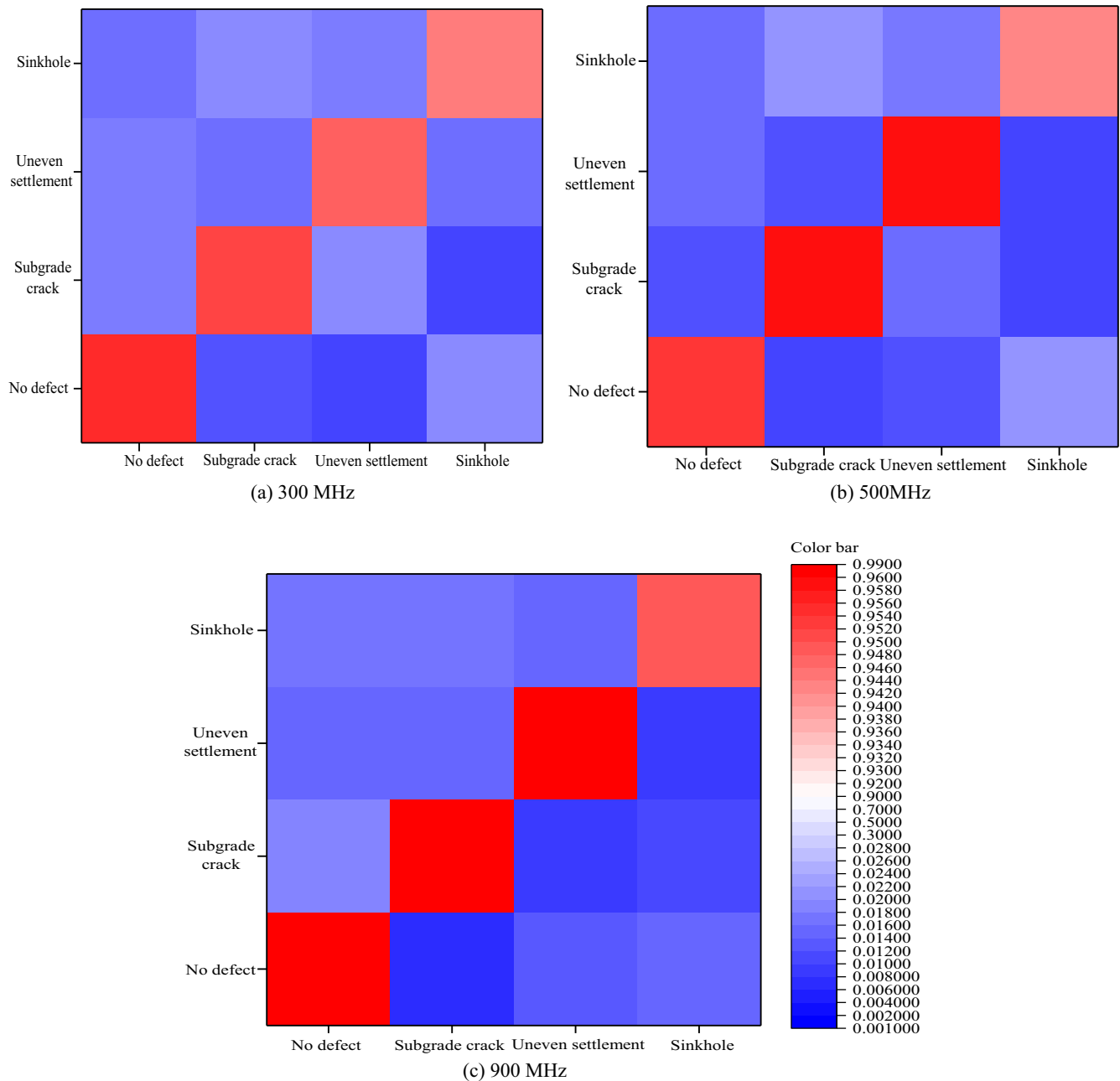
Fig. 9. Confusion matrices based on the testing dataset for cascade CNN with different transmitting frequency, i.e., 300, 500, and 900 MHz. The X-axis labels are the ground truth labels and the Y-axis labels are the predicted labels.

were imported into cascade CNN and defects were detected in small images at a resolution of 256 × 256 pixels. The results are shown in Table 2, where they demonstrate that the CNN-based method could classify different defects with 95.36% accuracy. Fig. 14 shows a GPR image after recognition using the CNN-based method. Table 2 and Fig. 14 demonstrate that the CNN-based method could detect subgrade defects with high accuracy in practical tests.

According to the defect recognition results obtained using the CNN-based method, suggestions can be made regarding the maintenance of the four highways. Thus, the efficiency of subgrade detection can be improved by saving the time taken to classify the types of defects in GPR images by using cascade CNN.

5. Conclusions

In this study, we described the application of CNN using GPR images for recognizing subgrade defects and we give the following conclusions.

(1) Cascade CNN classified subgrade defects with accuracies of 97.35% during training and 96.80% in the validation process. The multi-stage CNN classified subgrade defects with accuracies of 91.35% during training and 90.45% in the validation process. The accuracies of the two CNNs according to the training and validation results satisfy the requirements for subgrade assessment in China.
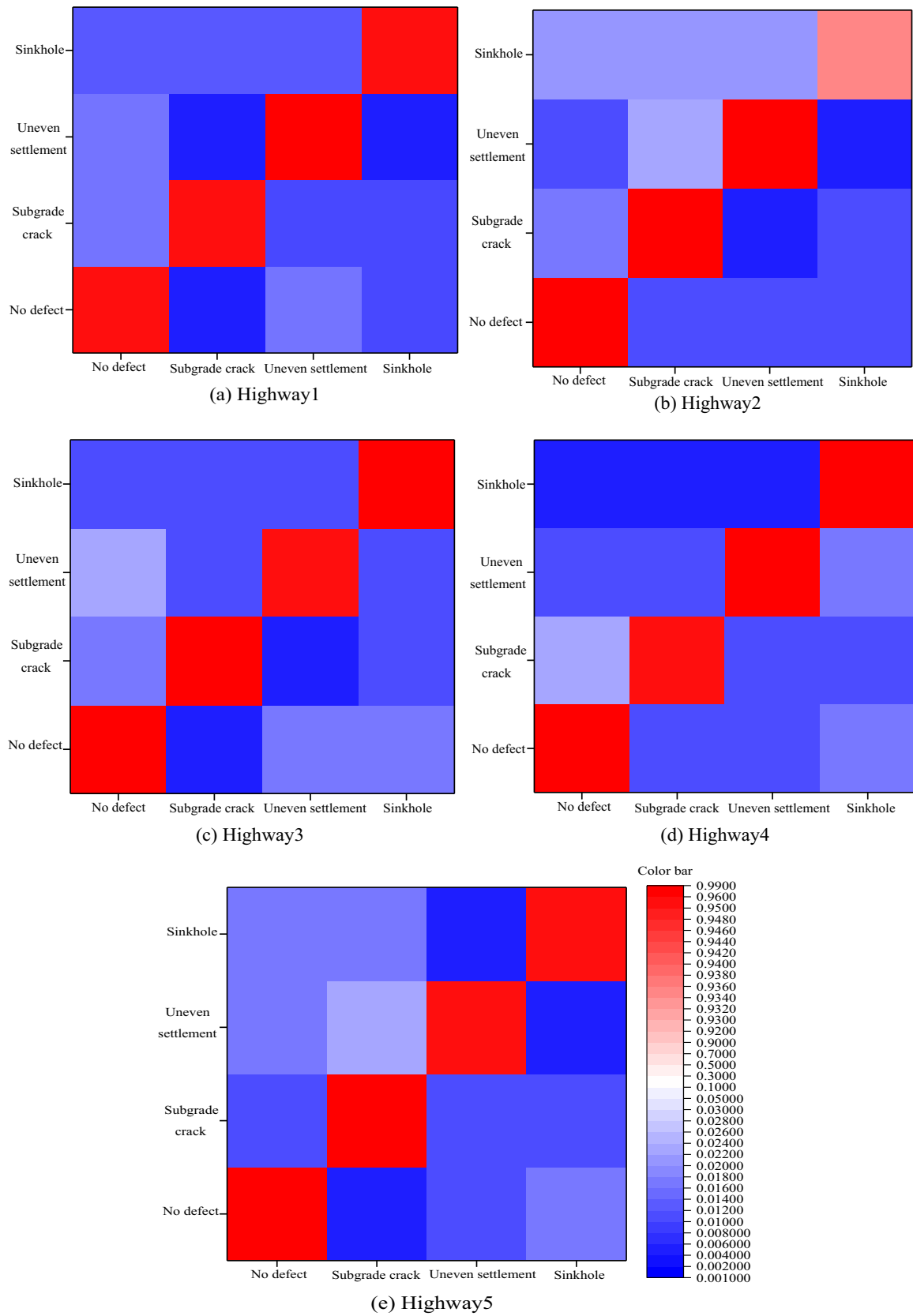
(a) Highway1

(b) Highway2

(c) Highway3

(d) Highway4

(e) Highway5

Fig. 10. Confusion matrices based on the testing dataset using cascade CNN with different highway structures. The X-axis labels are the ground truth labels and the Y-axis labels are the predicted labels.
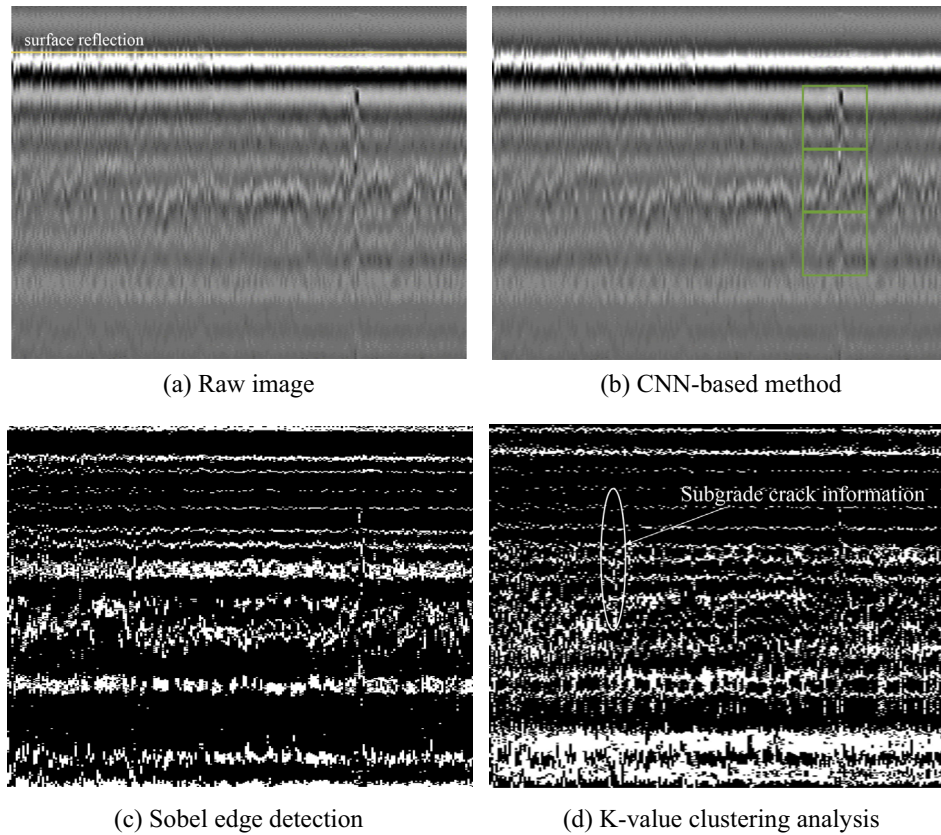
(a) Raw image

(b) CNN-based method

(c) Sobel edge detection

(d) K-value clustering analysis

Fig. 11. Subgrade crack with a transmitting frequency of 900 MHz. *Note:* The green box shows a small image at 256 × 256 pixel resolution including a subgrade crack. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
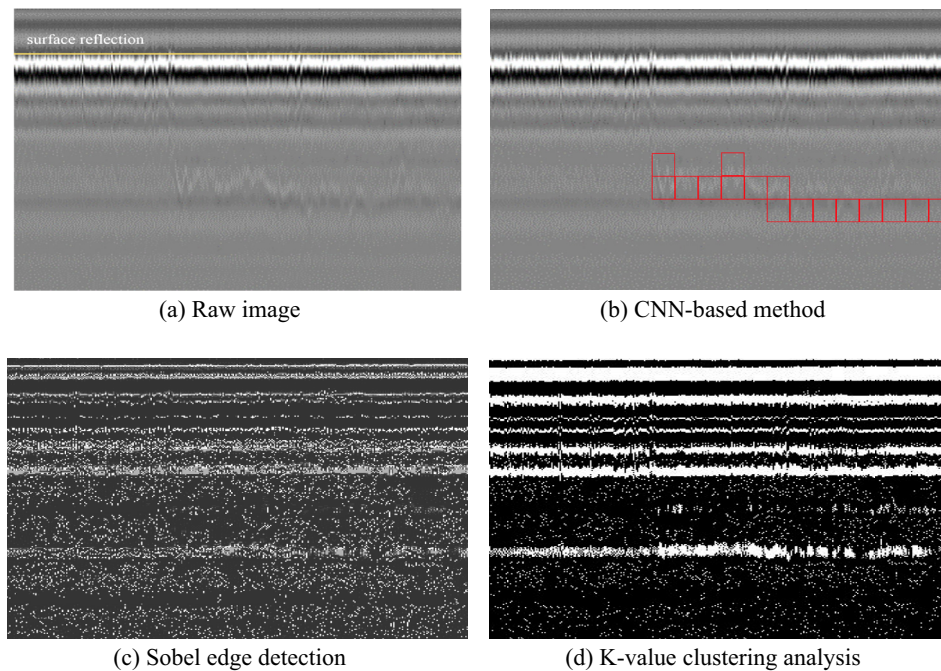

(a) Raw image

(b) CNN-based method

(c) Sobel edge detection

(d) K-value clustering analysis

Fig. 12. Uneven settlement with a transmitting frequency of 500 MHz. *Note:* The red box shows a small image at 256 × 256 pixel resolution including uneven settlement. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(2) Cascade CNN exhibited stability with respect to variations in both the transmitting frequency and highway structure, whereas the defect recognition performance of multi-stage CNN was unstable when using images obtained with different transmitting frequency. The strategy using classifier 2 in cascade CNN, which was trained only using hard samples, improved the robustness of object recognition in images obtained at different transmitting frequencies.
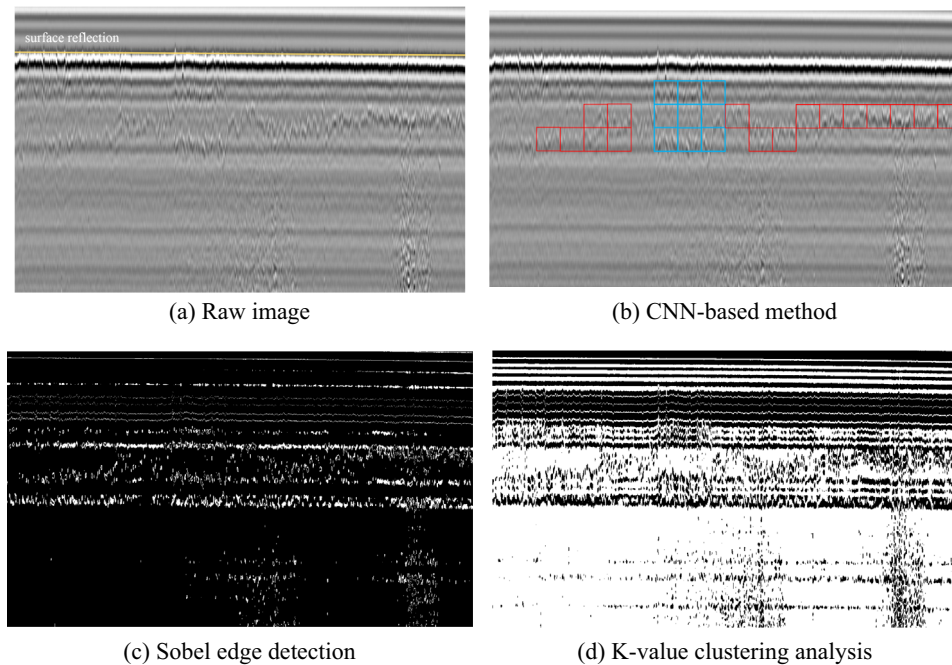
(a) Raw image

(b) CNN-based method

(c) Sobel edge detection

(d) K-value clustering analysis

Fig. 13. Two defects with a transmitting frequency of 500 MHz. *Note:* The red box shows a small image at 256 × 256 pixel resolution including uneven settlement. The blue box shows a small image at 256 × 256 pixel resolution including a sinkhole. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Confusion matrices obtained for the test application.

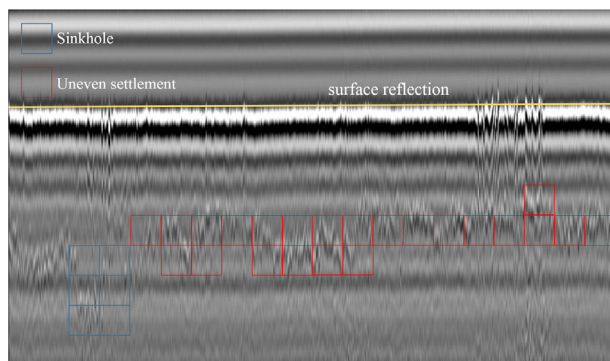| Accuracy	Sinkhole	Uneven settlement	Subgrade crack	No defect
Sinkhole	95.30%	1.33%	1.00%	1.70%
Uneven settlement	1.07%	94.93%	1.37%	2.17%
Subgrade crack	1.7%	1.13%	96.33%	1.23%
No defect	1.93%	2.60%	1.30%	94.90%

Fig. 14. Recognition results.

(3) The CNN-based method using cascade CNN exhibited highly robust subgrade defect detection performance under various conditions using the raw images compared with Sobel edge detection and K-value clustering analysis. The CNN-based method recognized most of the information in the GPR images and its performance was robust irrespective of the image quality. In addition, the CNN-based method could detect different defects in raw images.

(4) A more advanced CNN needs to be developed to allow three-dimensional reconstruction in order to describe subgrade defects in greater detail, and this will be the focus of our future research.

References

[1] Bowen Tai, Jiankun Liu, Tengfei Wang, Yupeng Shen, Xu Li, Numerical modelling of anti-frost heave measures of high-speed railway subgrade in cold regions, Cold Regions Sci. Technol. 141 (2017) 28–35.
[2] Ahmed Mancy Mosa, Amer Hasan Taher, Layth A. Al-Jaberi. Improvement of poor subgrade soils using cement kiln dust, Case Stud. Constr. Mater. (Available online 23 June 2017).
[3] Wallace Wai-Lok Lai, Xavier Dérobert, Peter Annan. A review of ground penetrating radar application in civil engineering: a 30-year journey from locating and testing to imaging and diagnosis. NDT & E Int. (Available online 23 May 2017).
[4] Shan Zhao, Imad L. Al-Qadi, Development of regularization methods on simulated ground-penetrating radar signals to predict thin asphalt overlay thickness, Signal Process. 132 (2017) 261–271.
[5] María Varela-González, Mercedes Solla, Joaquín Martínez-Sánchez, Pedro Arias, A semi-automatic processing and visualisation tool for ground-penetrating radar pavement thickness data, Autom. Constr. 45 (2014) 42–49.
[6] Tao Liu, Xiao-ning Zhang, Zhi Li, Zhou-quan Chen, Research on the homogeneity of asphalt pavement quality using X-ray computed tomography (CT) and fractal theory, Constr. Build. Mater. 68 (2014) 587–598.
[7] Lulu Edwards, Haley P. Bell, Comparative evaluation of nondestructive devices for measuring pavement thickness in the field, Int. J. Pavement Res. Technol. 9 (2) (2016) 102–111.

[8] Zheng Tong, Jie Gao, Haitao Zhang, Recognition, location, measurement, and 3D reconstruction of concealed cracks using convolutional neural networks, Constr. Build. Mater. 146 (2017) 775–787.

[9] Zehua Dong, Shengbo Ye, Yunze Gao, Guangyou Fang, Xiaojuan Zhang, Zhongjun Xue, Tao Zhang, Rapid detection methods for asphalt pavement thicknesses and defects by a vehicle-mounted ground penetrating radar (GPR) system, Sensors 16 (12) (2017).

[10] Aleksey K. Khamzin, Aleksandra V. Varnavinaa, Evgeniy V. Torgashov, Neil L. Anderson, Lesley H. Sneed, Utilization of air-launched ground penetrating radar (GPR) for pavement condition assessment, Constr. Build. Mater. 141 (2017) 130–139.

[11] P. Szymczyk, M. Szymczyk, Non-destructive building investigation through analysis of GPR signal by S-transform, Autom. Constr. 55 (2015) 35–46.

[12] Yuri A. Sukhobok, Viktor V. Pupatenko, Gennady M. Stoyanovich, Yulia V. Ponomarchuk, Soil formation lithological profiling using ground penetrating radar, Proc. Eng. 143 (2016) 1236–1243.

[13] F. Tosti, A. Benedetto, Pavement pumping prediction using ground penetrating radar, Proc. Soc. Behav. Sci. 52 (2012) 1044–1053.

[14] Cécile Barat, Christophe Ducottet, String representations and distances in deep Convolutional Neural Networks for image classification, Pattern Recogn. 54 (2016) 104–115.

[15] Shi Baoguang, Bai Xiang, Yao Cong, Script identification in the wild via discriminative convolutional neural network, Pattern Recogn. 52 (2016) 448–458.

[16] Y. LeCun, Generalization and network design strategies, in: Connections in Perspective, North-Holland, Amsterdam, 1989, pp. 143–155.

[17] Y. LeCun, Y. Bengio, in: Convolutional Networks for Images, Speech, and Time Series, The Hand Book of Brain Theory and Neural Networks, MIT Press, 1995, p. 3361.

[18] Jun Xu, Xiaofei Luo, Guanhao Wang, Hannah Gilmore, Anant Madabhushi, A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images, Neurocomputing 191 (2016) 214–223.

[19] E.P. Ijjina, K.M. Chalavadi, Human action recognition using genetic algorithms and convolutional neural networks, Pattern Recogn. 59 (2016) 199–212.

[20] Amin Jalali, Rammohan Mallipeddi, Minho Lee, Sensitive deep convolutional neural network for face recognition at large standoffs with small dataset, Expert Syst. Appl. 97 (2017) 304–315.

[21] Zh.u. Yingying, Zhang Chengquan, Zhou Duoyou, Wang Xinggang, Bai Xiang, Liu Wenyu, Traffic sign detection and recognition using fully convolutional network guided proposals, Neurocomputing 214 (2016) 758–766.

[22] Yuan Dong, Yue Wu, Adaptive Cascade Deep Convolutional Neural Networks for face alignment, Comput. Stand. Interfaces 42 (2015) 105–112.

[23] D. Tomè, F. Monti, L. Baroffio, L. Bondi, M. Tagliasacchi, S. Tubaro, Deep Convolutional Neural Networks for pedestrian detection, Signal Process. Image Commun. 47 (2016) 482–489.

[24] Min Fu, Pei Xu, Xudong Li, Qihe Liu, Mao Ye, Ce Zhu, Fast crowd density estimation with convolutional neural networks, Eng. Appl. Artif. Intell. 43 (2015) 81–88.

[25] Biao Leng, Shuang Guo, Xiangyang Zhang, Zhang Xiong, 3D object retrieval with stacked local convolutional autoencoder, Signal Process. 112 (2015) 119–128.

[26] B. Leng, Z. Xiong, Modelseek: an effective 3D model retrieval system, Multimedia Tools Appl. 51 (3) (2011) 935–962.

[27] B. Leng, Z. Xiong, X. Fu, A 3D shape retrieval framework for 3D smart cities, Front. Comput. Sci. China 4 (3) (2010) 394–404.

[28] Zheng Tong, Jie Gao, Zhenqiang Han, Zhenjun Wang, Recognition of asphalt pavement crack length using deep convolutional neural networks, Road Mater. Pavement Des. (Published online).

[29] Yong-Jin Cha, Wooram Choi, Oral Buyukozturk, Deep learning-based crack damage detection using convolutional neural networks, Comput. Aided Civil Infrastruct. Eng. 32 (5) (2017) 361–378.

[30] Yong-Jin Cha, Choi Wooram, Suh Gahyun, Sadegh Mahmoudkhani, Oral Büyüköztürk, Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types, Comput. Aided Civil Infrastruct. Eng. (Published online).

[31] JTG H20-2007, Highway Performance Assessment Standards, Beijing, China.

[32] Zheng Tong, Aimin Sha, Jie Gao, Innovation for Recognition of Pavement Distresses by Using Convolutional Neural Network, World Transport Convention, Beijing, China, 2017.

[33] Zheng Tong, Jie Gao, Haitao Zhang, Innovation for evaluating aggregate angularity based upon 3D convolutional neural network, Constr. Build. Mater. 155 (2017) 919–929.

[34] Dominik Scherer, Andreas Muller, Sven Behnke, Evaluation of pooling operations in convolutional architectures for object recognition, Int. Conf. Artif. Neural Netw. (2010) 92–101.

[35] Fu. Min, Xu. Pei, Xudong Li, Qihe Liu, Mao Ye, Ce Zhu, Fast crowd density estimation with convolutional neural networks, Eng. Appl. Artif. Intell. 43 (2015) 81–88.

[36] Yoshua Bengio, Practical recommendations for gradient-based training of deep architectures, in: G. Montavon, G.B. Orr, K.-R. Muller (Eds.), Neural Networks: Tricks of the Trade, second ed., Springer, Berlin Heidelberg, 2012, pp. 437–478.

[37] Lu. Cheng-Ming, Qin Zhen, Zh.u. Hai-Long, Li Xiu-Zhong, Practical methods for detection of concealed cracks in highway pavement using ground penetration radar data, Chin. J. Geophys. 50 (5) (2007) 1558–1568.

[38] S. Zhao, P. Shangguan, I.L. Al-Qadi, Application of regularized deconvolution technique for predicting pavement thin layer thicknesses from ground penetrating radar data, NDT & E Int. 73 (2015) 1–7.

[39] A. Ayenu-Prah, N. Attoh-Okine, Evaluating pavement cracks with bidimensional empirical mode decomposition, EURASIP J. Adv. Signal Process. 2008 (2008) 861701, https://doi.org/10.1155/2008/861701.

[40] O. Yashon, M. Ouma, Hahn. Pothole detection on asphalt pavements from 2D-colour pothole images using fuzzy c-means clustering and morphological reconstruction, Autom. Constr. 83 (2017) 196–211.