



Autonomous pavement distress detection using ground penetrating radar and region-based deep learning

Jie Gao ^{a,b,c,*}, Dongdong Yuan ^b, Zheng Tong ^{d,*}, Jiangang Yang ^a, Di Yu ^a

^a School of Transportation and Logistics, East China Jiaotong University, Nanchang 330013, China

^b School of Highway, Chang'an University, Xi'an 710064, China

^c Chanba Ecological District Management Committee, Xi'an 710024, China

^d Sorbonne Université, Université de Technologie de Compiègne, CNRS, UMR 7253 Heudiasyc, CS 60319-60203 Compiègne cedex, France



ARTICLE INFO

Article history:

Received 2 March 2020

Received in revised form 30 April 2020

Accepted 4 June 2020

Available online 12 June 2020

Keywords:

Pavement distress

Faster region convolutional neural network

Ground penetrating radar

Image processing

ABSTRACT

Many data processing technologies have been utilized for pavement distress detection (e.g., reflection cracks, water-damage pits, and uneven settlements) using ground penetrating radar (GPR). However, the various real-world conditions have resulted in challenges of the accuracy and generalization ability of these techniques. To overcome these challenges, we proposed a deep-learning method, called faster region convolutional neural network (Faster R-ConvNet), to complete the task. The 30 Faster R-ConvNets were trained, validated, and tested using 2,557, 614, and 614 GPR images, respectively. The optimal anchor size and ratio were determined based on the validation results. The stability, superiority, real-time of the optimal Faster R-ConvNet were verified based on the test results. The results demonstrated that the optimal Faster R-ConvNet achieved 89.13% precision and 86.24% IoU. The stability of the model in different pavement structures was desirable. The comparative study indicated that the optimal Faster R-ConvNet outperformed other supervised learning methods in distress detection. Additionally, a real-time detection using optimal Faster R-ConvNet was conducted with acceptable accuracy.

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

Asphalt pavement distresses are an inevitable issue in pavement structure inspection and condition assessment. The detection of some distresses has proved to be challenging owing to their location in the pavement, such as reflection crack and water-damage pit [1]. For example, reflection cracks are the most common cracks in semi-rigid bases in China, and they always occur in either the semi-rigid base itself or in sub-bases before propagating to the pavement [2]. These distresses are not visual during their developing stages. However, it is essential to detect them before they propagate to the pavement surface to guarantee the highway performance.

In the past decade, several innovative technologies have been introduced to detect pavement distresses, mainly including ground penetrating radar (GPR) [2,3], ultrasonic testing [4,5], radio examination [6,7], and thermal image/infrared [8,9]. Compared with other technologies, GPR has superiority in high efficiency and pre-

cision, reasonable anti-interference capacity, and wide detection ranges. For example, Fernandes et al. [10] proposed a method to detect cracks with different widths using GPR. Ahn et al. [2] inspected the pavement subsurface distress severity and its distribution using GPR. However, there were still two problems remaining using GPR to detect pavement distresses: (a) complicated data preprocessing technologies should be utilized to extract pavement distress features; and (b) the human assistance is required to handle the complex background and foreground in the GPR data. For example, Solla et al. [11] presented a GPR-based method to detect pavement concealed crack and their features. The main disadvantage of this method was that an infrared thermometer was required for auxiliary analysis. Szymczyk et al. [12] proposed a signal transformation method to reconstruct GPR data to detect underground defects, but the conversion processes were complex. Xu et al. [13] presented a GPR detection method to locate pavement voids and cracks, but the distress classification should be realized by humans. In summary, the applications of GPR for pavement distress detection have been limited by the two problems. It is essential to develop a method to process GPR data autonomously or semi-autonomously, which has the capacity of handling the complex background and foreground in the GPR data.

* Corresponding authors at: School of Transportation and Logistics, East China Jiaotong University, Nanchang 330013, China (J. Gao).

E-mail addresses: gaojie@ecjtu.edu.cn (J. Gao), zheng.tong@hds.utc.fr (Z. Tong).

Many data processing technologies have been implemented for completing the tasks to partially replace human-assistance inspections, such as support vector machine (SVM) and S-transform. For example, Zhou et al. [14] used SVMs to classify different pavement distresses using GPR data. The imported data for SVMs were features acquired by clutter suppression and area-of-interest extraction. 3D S-transform [12] was proposed to process GPR data to look for sinkholes, but it required human assistance during the processes. Xue et al. [15] proposed an extension evaluation method by combining GPR and falling weight deflectometer to detect pavement voids. Tosti and Benedetto [16] recognize pavement pumping based on the GPR signal, although this method was only suitable for the certain object. Generally, these data-processing technologies detected or recognized some pavement distresses using GPR. However, there was still a problem remaining: complex distress and background features still should be extracted by humans because the shapes, grayscale, and other features of distresses were various in different GPR images. The problem led to the precisions of existing methods were varied from one real-world condition to another. Therefore, it is necessary to develop a detection method to extract distress features automatically, whose precision is stable in different real-world conditions.

With the explosive development of deep learning and its application, it has shown the significant advantages in pattern recognition [17,18], 3D object retrieval [19], object location [20,21], and regressive calculation [22,23]. Notably, deep-learning methods have been used in pavement distress detection. For example, Cha et al. [24] used convolutional neural networks (ConvNet) to recognize pavement cracks in real-world conditions with high precision. Sha et al. [25] proposed a ConvNet-based method to recognize different pavement distresses. Tong and colleagues [26,27] developed a series of deep-learning models to recognize, locate, three-dimension reconstruct pavement distresses. Although the recognition results in the two studies [26,27] were reasonable, the pave-

ment detection results were not accurate owing to the deep-learning structures. Ali and Cha [28] proposed a novel combination of deep learning and uncooled micro-bolometer for the detection of a steel bridge, achieving 96% accuracy and 97.79% specificity. Beckman et al. [29] designed an advanced deep learning approach integrated with a structural surface fitting algorithm for automatic volumetric damage quantification using a depth camera. The reported average precision and mean precision error were 90.79% and 9.45%, respectively. In summary, there are two attractive properties for the deep-learning method and its application in civil engineering fields. First one is stability, which indicates that the model has a great tolerance for translation and distortion of the detection objects [29,30]. In this study, stability means that the model is stable to detect distresses, though the features of distresses and background are various. The second one is automation, which indicates that the model can learn low-, medium-, high-features related to detection objects automatically [31–33]. In this study, automation means that the model has the capacity for learning a mass of features related to pavement distresses without human assistance. Therefore, it is feasible to utilize deep learning to detect pavement distress using GPR data.

Motivated by the drawback of the pavement distress detection in the previous study [26,27], we proposed Faster Region ConvNets (Faster R-ConvNets), which was a type of modified deep-learning models, to detect pavement distresses using GPR image. The main advantage of the method was that it required no human assistance to detect pavement distresses with high precision and recall. Additionally, the performance of the model was stable in different pavement structures and materials. Thus, it can be regarded as an automatic nondestructive testing method to replace traditional pavement inspection technologies, such as some unsupervised algorithms (e.g., S-transform and Sobel edge detection) and supervised algorithms proposed before the wide adoption of deep learning (e.g., artificial neural network and support vector machine).

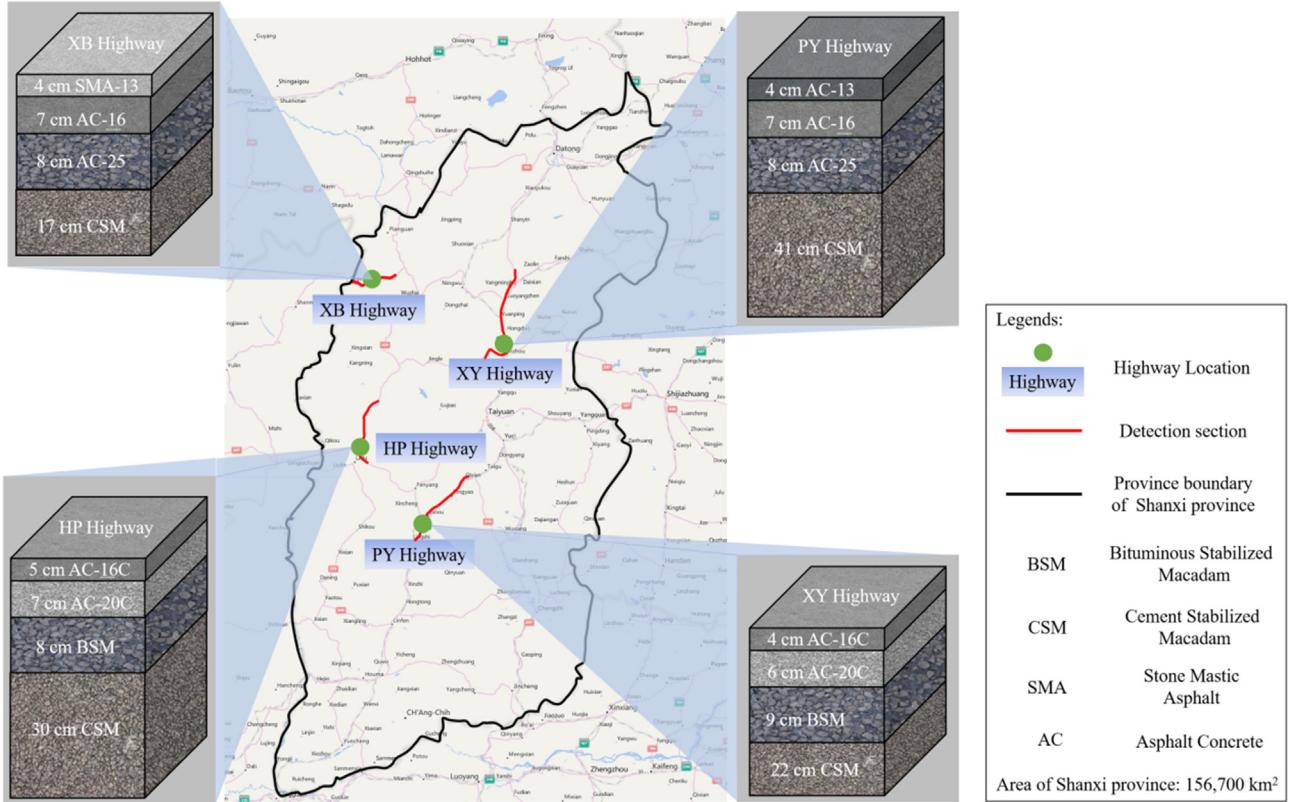


Fig. 1. Different Structures of four highways.

Compared to existing Faster R-CNN models, the proposed models can be considered as its modification. There are two parts of the modification: hyperparameter optimization to adapt the detection task and algorithm improvement to improve its accuracy. The first one is the optimization of anchor ratios and sizes. The aim is to make the bounding boxes desirable to detect pavement distresses using GPR data. The second one mainly includes the adoption of stochastic pooling and layer-sequential unit-variance initialization, both of which have been verified to improve the algorithm accuracy but not adopted in the existing Faster R-CNN. The rest of this paper is organized as follows. We present our research approaches, including the method for acquiring GPR images, the design of the Faster R-ConvNets, and implementation details in [Section 2](#). This is followed by a discussion of the performance of the Faster R-ConvNets based on the validation and testing results in [Section 3](#). The performance discussion includes the training and testing results, a stability analysis, a comparative study, and real-time detection. Our conclusions are summarized in [Section 4](#).

2. Research approaches

In this section, we first present the method for inspecting pavement distress using a GPR in [Section 2.1](#). Then we provide a detailed description of the Faster R-ConvNets in [Section 2.2](#). At last, the database and implementation details are shown in [Section 2.3](#).

2.1. Acquisition GPR images

A high-quality GPR image database was required to develop Faster R-ConvNets for detecting pavement distresses autonomously. The quality of GPR images was influenced by transmitting frequencies of GPR and pavement structures. Thus, GPR images under different conditions (transmitting frequencies and pavement structures) were collected to guarantee the integrality of the database.

In this study, an LTD-2000 air-coupled GPR was used to capture pavement distress images. The parameters of the LTD-2000 GPR has been introduced in a similar work Ref. [26]. The sweep speed

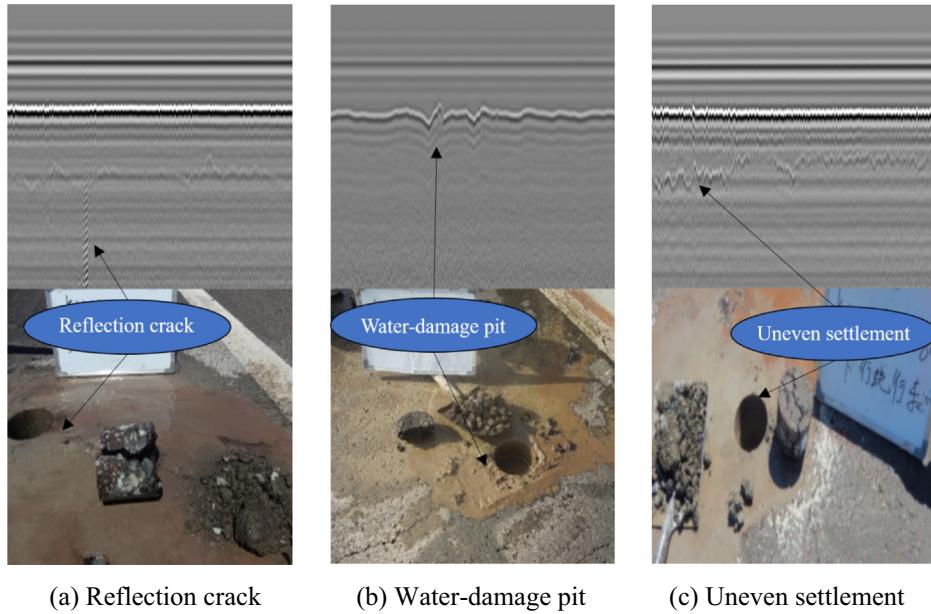


Fig. 2. Typical GPR images of pavement distresses.

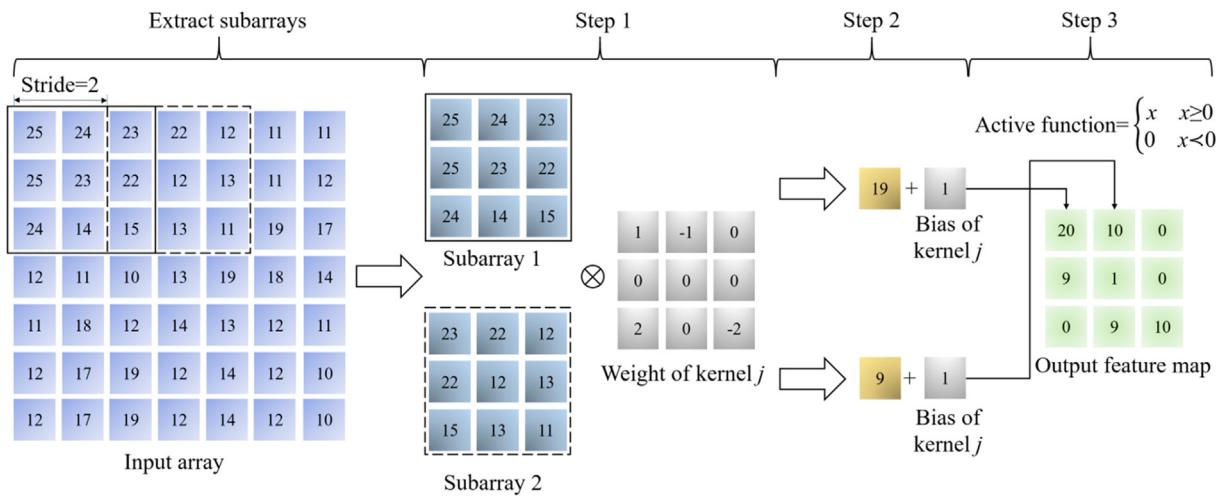


Fig. 3. An example of convolution operation.

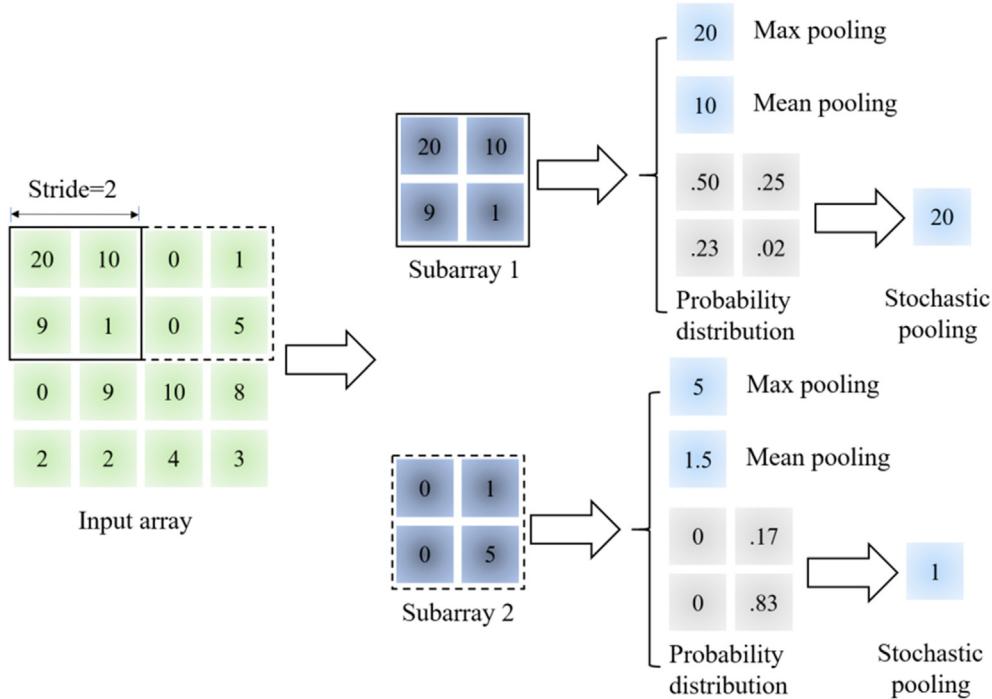


Fig. 4. Examples of three pooling operations.

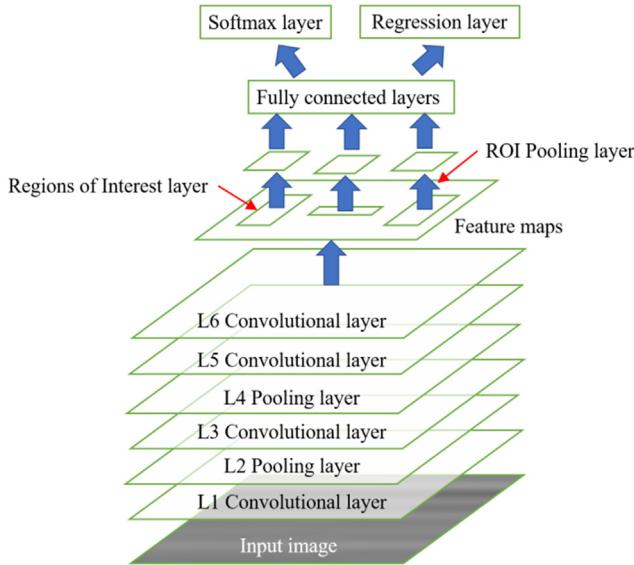


Fig. 5. Flow chart of a Fast R-ConvNet.

and time window during the pavement GPR image collection were 128 Hz and 5 ns, respectively. To guarantee the integrality of the GPR database, three different transmitting frequencies (300, 500, 900 MHz) were utilized to collect GPR images in the same pavements. The vertical resolution ranges of the different transmitting frequencies were 0.30–0.47 cm, 0.15–0.27 cm, and 0.09–0.13 cm, while the horizontal ranges were 3.0–5.2 cm, 1.3–2.9 cm, and 1.0–1.4 cm. Notably, the vertical resolution of the transmitting frequencies could meet the engineering demands, but the horizontal resolution could not meet the detection demand of reflection cracks [34], which led that the crack widths in GPR images were not their real width. However, as shown in the study of Guo et al. [35], the peak location of a reflection crack in a GPR image was the same as its real location in the pavement. Therefore, the horizontal resolution had limited influence on the detection of reflection cracks. Additionally, GPR images from four highways in Shanxi Province, China were acquired using the three transmitting frequencies to guarantee the integrality of the GPR database. Fig. 1 presents the structures of the four pavements.

Typical GPR images of a reflection crack, a water-damage pit, and an uneven settlement are presented in Fig. 2. The distress annotations in all the GPR images were determined by a field

Table 1
Detailed parameters of a Fast R-ConvNet.

Layer number	Type	Input size	Filter size	Number	Stride
L0	Input	$512 \times 512 \times 1$	–	–	–
L1	Convolution	512×512	21×21	96	2
L2	Pooling	$246 \times 246 \times 96$	–	–	2
L3	Convolution	$123 \times 123 \times 96$	14×14	192	2
L4	Pooling	$55 \times 55 \times 192$	–	–	2
L5	Convolution	$27 \times 27 \times 192$	4×4	384	1
L6	Convolution	$24 \times 24 \times 384$	4×4	384	1
L7	ROI	Ran. \times Ran. \times 192	–	–	–
L8	ROI pooling	192	–	–	–
L9	Fully connected	1024	–	–	–
L10	Fully connected	256	–	–	–
L11	Fully connected	4	–	–	–
L12	Output	–	–	–	–

investigation to guarantee its reliability. As shown in Fig. 2 and the study of Tong [36], different pavement distresses showed different characteristics in GPR images, but these characteristics were various in the real-world conditions and not easy to summarize and extract by humans. For example, the shapes of reflection cracks were a scattered hyperbolic, but their symmetry of left and right sides was varied from one to another. Thus, the core sampling shown in Fig. 2 was used to determine the types of distresses that it occurred after GPR scanning once an abnormal wave was shown in a GPR image.

2.2. Region-based deep learning

To detect pavement distresses (e.g. reflection cracks, water-damage pits, and uneven settlements) using GPR images, the Faster R-ConvNet models were developed, which were a type of modified deep learning models. We start with some necessary fundamental concepts of deep learning briefly in Section 2.2.1. Then the Fast R-ConvNet and region proposal networks (RNP), as two key parts of the Faster R-ConvNets, are described in detailed in Section 2.2.2 and Section 2.2.3, respectively. At last, the combination model of

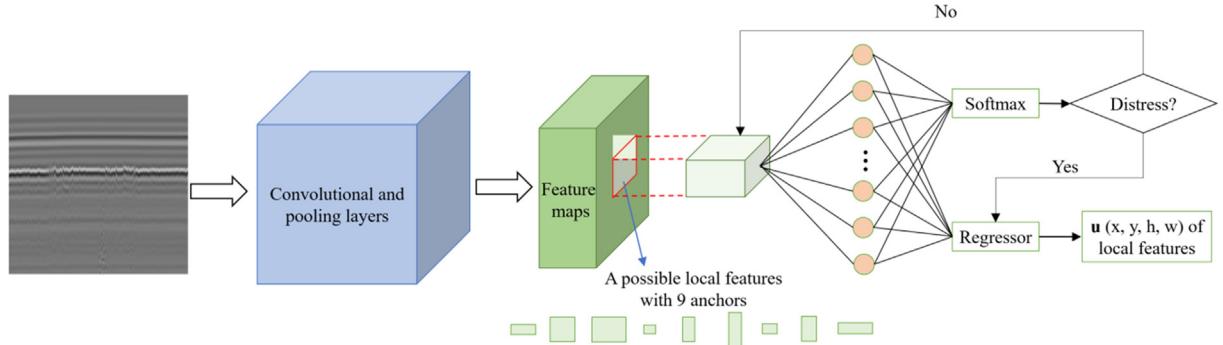


Fig. 6. Flow chart of an RNP.

Table 2
Detailed parameters of an RNP.

Layer number	Type	Input size	Filter size	Number	Stride
L0	Input	$512 \times 512 \times 1$	–	–	–
L1	Convolution	512×512	$21 \times 21 \times 1$	96	2
L2	Pooling	$246 \times 246 \times 96$	–	–	2
L3	Convolution	$123 \times 123 \times 96$	$14 \times 14 \times 96$	192	2
L4	Pooling	$55 \times 55 \times 192$	–	–	2
L5	Convolution	$27 \times 27 \times 192$	$4 \times 4 \times 192$	384	1
L6	Convolution	$24 \times 24 \times 384$	$4 \times 4 \times 384$	384	1
L7	Anchor	$21 \times 21 \times 384$	$4 \times 4 \times 384$	256	1
L8	Fully connected	256	–	–	–
L9	Output				–

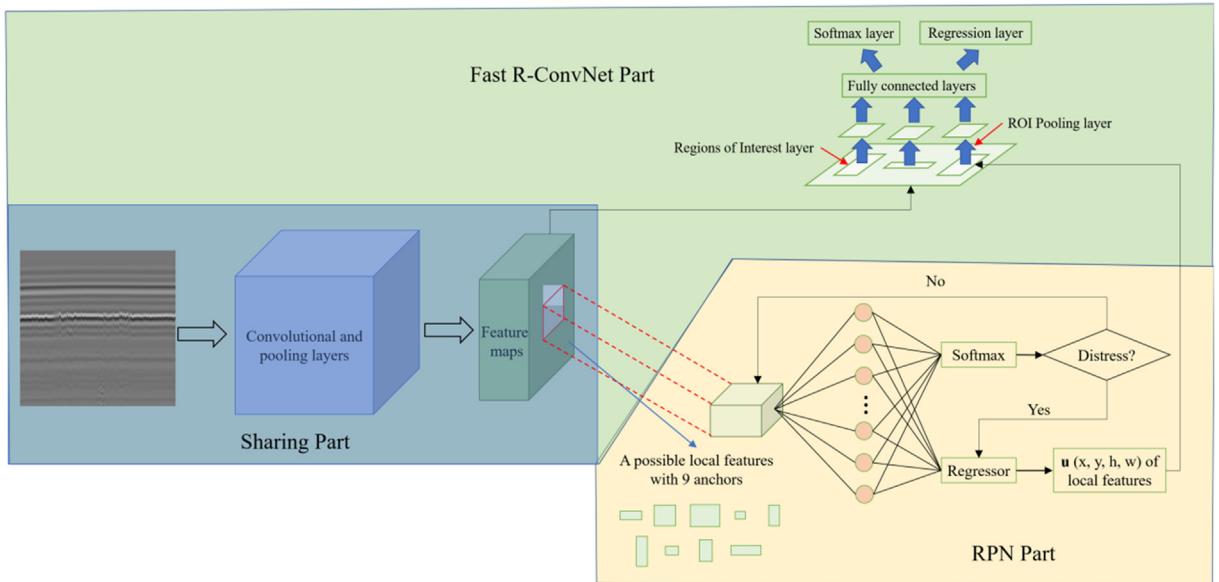


Fig. 7. Flow chart of a Faster R-ConvNet.

an RPN and a Fast R-ConvNet, called Faster R-ConvNet, are presented in [Section 2.2.4](#).

2.2.1. Fundamental concepts

Prior to the development of the Faster R-ConvNets, the fundamental concepts of deep learning are briefly recalled in this section as following:

- (1) *Convolutional layer.* The function of a convolutional layer is to extract low-, medium-, or high-features from input data using a large number of kernels, whose procedures are called convolution operation. There are three steps in a convolution operation as shown in [Fig. 3](#). Firstly, an element-by-element multiplication is conducted between a subarray of an input array and a kernel. Secondly, the multiplied values are summed, and bias is added to the summed values. Lastly,

the result in the second step is active by an active function, such as *ReLU*. The weights and bias of a kernel are given by training using gradient descent algorithms (e.g. stochastic gradient descent).

- (2) *Pooling layer.* The function of a pooling layer is to reduce the spatial size of an input array to avoid overfitting and reduce computation. There are mainly three different pooling options, max pooling, mean pooling, and stochastic pooling, as shown in [Fig. 4](#). The study of Zeiler & Fergus [37] showed that the stochastic pooling was a more effective method for regularizing ConvNets than the other two. Thus, all the pooling layers for this study were stochastic pooling layers. A stochastic pooling layer samples from one of an input array's subarrays based on a probability distribution.
- (3) *Fully connected layer.* The function of fully connected layers is to map the arrays acquired by convolutional layers and

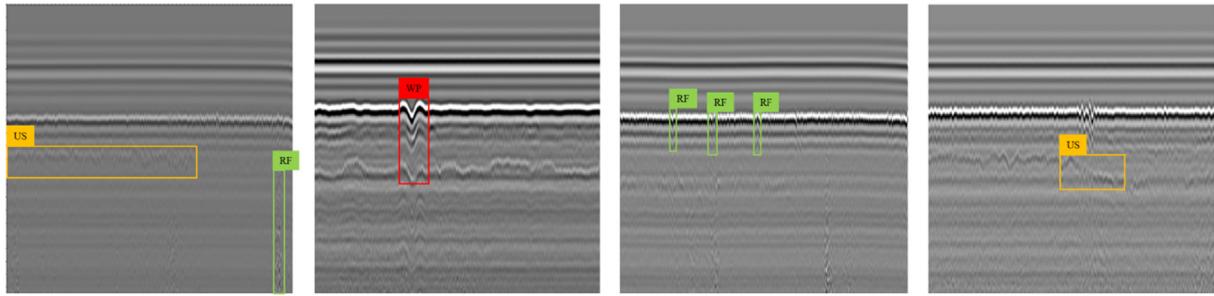


Fig. 8. Database examples with the labels and bounding boxes. RF, WP, and US stands for reflection crack, water-damage pit, uneven settlement, respectively.

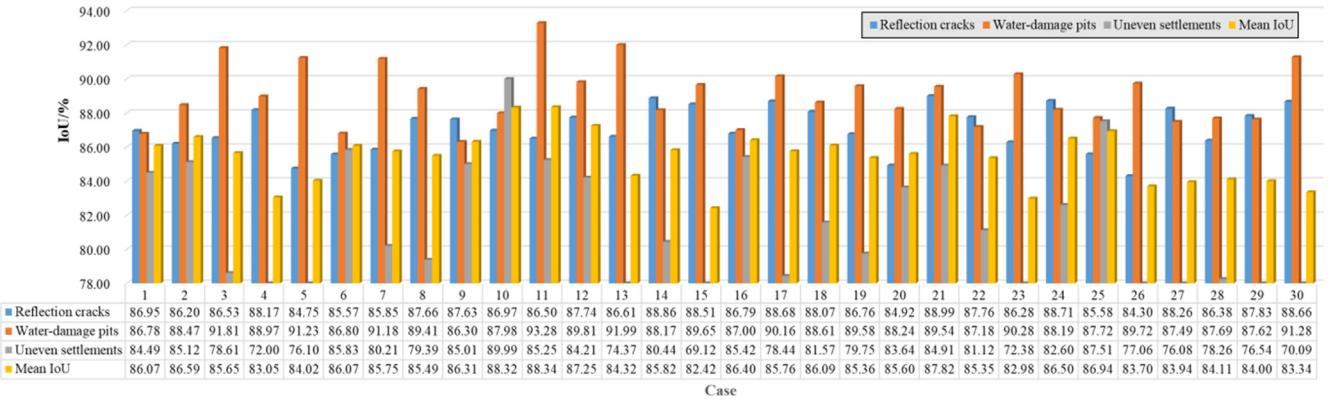


Fig. 9. Validation IoUs of the 30 Faster R-ConvNets.

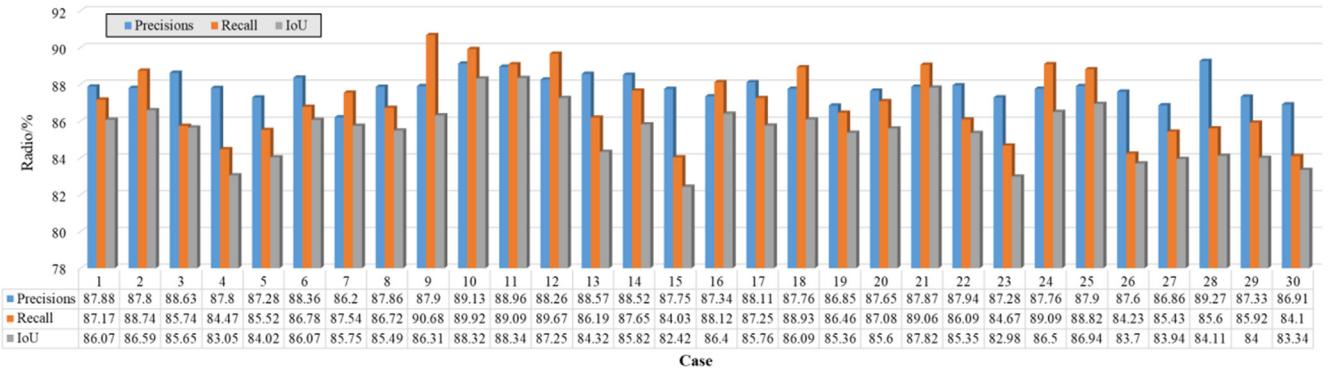


Fig. 10. Validation precisions, recalls and IoUs of the 30 Faster R-ConvNets.

pooling layers to a target space. The outputs of the final fully connected layer are used to classify and regress in output layers.

(4) Output layers. There are two types of output layers for detecting pavement distresses using GPR images. The first one is

Table 3
Precisions, recalls, and IoUs of No. 10 Faster R-ConvNet (Unit: %).

Object	Precision	Recall	IoU
Reflection cracks	88.31	89.04	86.53
Water-damage pits	90.56	89.68	87.26
Uneven settlements	88.51	91.04	84.93
Overall	87.13	89.92	86.24

a softmax layer, which is used to predict the distress class, and the second one is a regression layer, which is used to compute the location of the distress in a GPR image. Detail information of the two layers is introduced in the next Section.

2.2.2. Fast R-ConvNet

Fig. 5 and **Table 1** present the flow chart and parameters of the Fast R-ConvNet in this study. The Fast R-ConvNet consisted of convolutional layers, pooling layers, a region of interest (ROI) layer, an ROI pooling layer, and two output layers. The ROI and ROI pooling layers were the key to detect pavement distresses, whereas a conventional ConvNet could only recognize distresses in images but could not locate them [27]. The function of the ROI layer was used to randomly select local features from feature maps generated by convolutional layers and pooling layers. These local features were

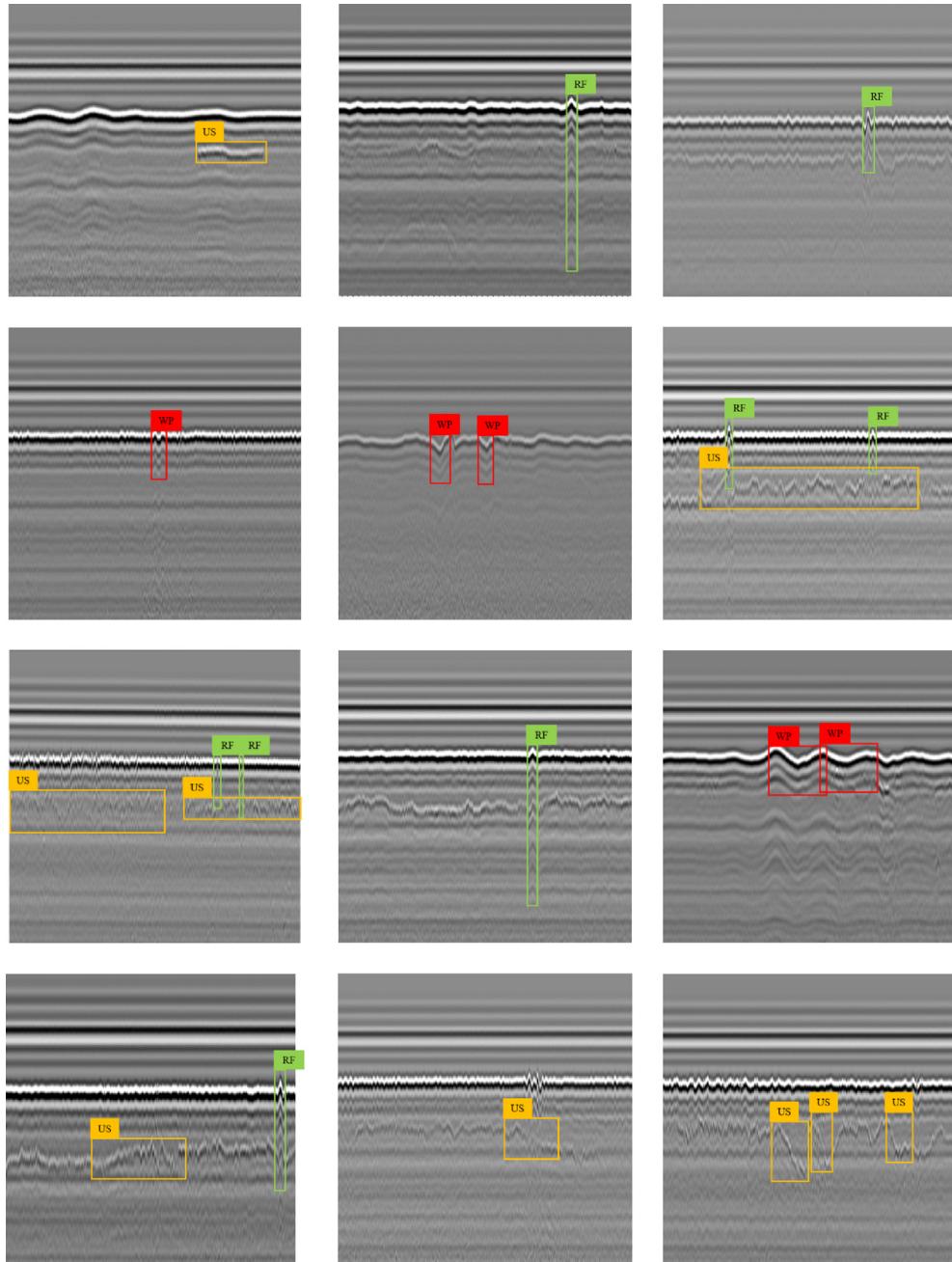


Fig. 11. Examples of testing results.

imported into fully connected layers and output layers in sequence to determine whether a distress was shown in their corresponding parts in the GPR image or not. However, there was a problem that the sizes of local features are random. Thus, an ROI pooling layer was utilized to normalize these local features to fit the input size of the first fully connected layer.

A probability \mathbf{P} (P_0, P_1, P_2, P_4) and four parameters (x, y, h, w) were generated by the softmax and regression layers based on the normalized local features, respectively. P_0, P_1, P_2, P_4 were the probabilities of no distress, reflection crack, water-damage pit, and uneven settlement. x and y were the center coordinates of a distress, while h and w were the height and width of the distress.

Considering the initial weights and bias in each layer were given randomly, there were gaps between the outputs of the two output layers and the ground truths. A training strategy, named mini-batch gradient descent, was employed to reduce the gaps. In a mini-batch gradient descent epoch, 3 GPR images and 192 local features were chosen randomly. The gaps, also called the loss of the Fast R-ConvNet, were computed by the combination of the deviation from the two output layers as Equation (1). Then weights and bias in each layer were adjusted using Equation (2) until the outputs were closed to the ground truths,

$$\text{Loss_output} = L_{\text{softmax}} + 0\{P_0 = 0\} \cdot L_{\text{regression}} \quad (1a)$$

$$L_{\text{softmax}} = \frac{1}{1 + e^{-|P_{\text{output}} - P_{\text{ground_truth}}|}} \quad (1b)$$

$$L_{\text{regression}} = \frac{1}{2N} \sum_{i=1}^{192} \| u_{\text{output}} - u_{\text{ground_truth}} \|_2^2 \quad (1c)$$

$$\text{weight}^{i+1} = \text{weight}^i + \alpha \cdot \frac{\partial \text{Loss_output}}{\partial \text{weight}^i} \quad (2a)$$

$$\text{bias}^{i+1} = \text{bias}^i + \alpha \cdot \frac{\partial \text{Loss_output}}{\partial \text{bias}^i} \quad (2b)$$

where $0\{P_0 = 0\}$ was a logical expression, which returned 0 if a predicted class was no distress, returning 1 otherwise. \mathbf{u} was a vector of (x, y, h, w) . α was learning rate during the training. α was 0.002 in this study.

2.2.3. Region proposal network

In the Fast R-ConvNet, the ROI layer generated local features randomly. It led to unnecessary computation during training and testing. Additionally, it also led to the loss of some important local features. In this study, region proposal networks (RPNs) were used to generate local features to replace the random selection. The flow chart and parameters of an RPN used in this study are presented in Fig. 6 and Table 2. There were three steps for an RNP to generate local features as following:

- (1) Feature maps were acquired by convolutional and pooling layers;

	RC	WP	US		RC	WP	US		RC	WP	US
RC	86.57	0.02	0.36	WP	88.76	0.06	0.47	US	89.60	0.42	0.49
WP	0.21	89.24	0.41	US	0.14	90.42	0.32	BG	0.41	92.02	0.30
US	0.24	0.06	87.99	BG	0.13	0.21	88.24	US	0.26	0.20	89.30
BG	12.98	10.68	11.24		10.97	9.31	10.97	BG	9.73	7.36	9.91

(a) 300 MHz

(b) 500 MHz

(c) 900 MHz

Fig. 12. Confused matrices of precision results in different transmitting frequencies. RC, WP, US, and BG are reflection cracks, water-damage pits, uneven settlement, and background.

	RC	WP	US		RC	WP	US		RC	WP	US
RC	86.17	0.65	0.57	WP	88.41	0.78	1.34	US	92.54	0.41	0.21
WP	1.29	86.82	0.52	US	0.52	88.97	1.31	BG	0.11	93.25	0.69
US	1.16	0.74	87.93	BG	0.73	0.93	90.51	US	0.14	0.17	95.68
BG	11.38	11.79	10.98		10.34	9.32	6.84	BG	7.21	6.17	4.42

(a) 300 MHz

(b) 500 MHz

(c) 900 MHz

Fig. 13. Confused matrices of recall results in different transmitting frequencies. RC, WP, US, and BG are reflection cracks, water-damage pits, uneven settlement, and background.

- (2) Nine rectangular boxes, called anchors, with three different widths and heights were used to generate possible local features. Possible local features were feature maps might include a pavement distress in this study.
- (3) Possible local features were imported into the softmax layer and regression layer. The softmax layer was used to determine whether a possible local feature included a partial or whole pavement distress. The regression layer was used to reduce the deviation between an anchor and the ground truth. The detailed procedures for reducing the deviation have been introduced in the study of Ren et al. [38].

As shown in the three steps, the key issue was to design suitable anchors to generate possible local features. We expressed an anchor as $(x_i, y_i, h_{ij}, w_{ij})$, $i = 1, 2, \dots, 9$, $j = 1, 2, 3$. x_i and y_i were the center of an anchor, while h_{ij} and w_{ij} were the height and width of the anchor. To guarantee the rationality of the sizes and ratios of the nine anchors, 15 combinations of seven different ratios (0.3, 0.4, 0.7, 1.0, 1.3, 1.6, and 1.9) were utilized in this study, while 2 combinations of sizes (16, 32, 48, 64, 128, and 192) were selected. Therefore, we designed 30 RPNs to produce possible local features. The performance of the 30 RPNs were evaluated by the validation performance of the 30 corresponding Faster R-ConvNet. Finally, the optimal RNP and its Faster R-ConvNet were chosen to detect pavement distresses for the testing data set.

2.2.4. Region proposal network

As shown in Tables 1 and 2, the 30 RPNs had the same convolution layers and pooling layers (L1-L4) as the Fast R-ConvNet. Thus, the two model were combined by sharing the L1-L4, which saved time in the both training and testing. The combination model, as shown in Fig. 7, was named Faster R-ConvNet. Referring a study

of transfer learning [39], there were four steps to train the 30 Faster R-ConvNets as follows.

Step 1. An RNP was given the weights with layer-sequential unit-variance initialization [40] and trained until the RNP could distinguish GPR images with pavement distresses from normal GPR images.

Step 2. The learned weights in L1-L4 were given to the Fast R-ConvNet.

Step 3. The RPN was used to generate a large number of local features as the training data set for the Fast R-ConvNet.

Step 4. The local features were used to fine tune the Fast R-ConvNet, whose initial weights were given in the Step2.

Each RPN and the Fast R-CNN network were trained with a learning rate of 0.002, a momentum of 1.0, and a weight decay of 0.0010 for 60,000 and 30,000 iterations, respectively.

2.3. Database and implementation details

A database including 3785 GPR images with a pixel resolution of 512×512 was generated by the method introduced in Section 2.1. The ratios of three transmitting frequencies (300, 500, 900 MHz) and four highways were 1:1:1 and 1:1:1:1, respectively. Three pavement distresses (reflection crack, water-damage pit, and uneven settlement) were included in the database. We annotated the labels and bounding boxes of pavement distresses in a Python environment as shown in Fig. 8.

A training and validation data set were generated by randomly selecting GPR images from the database. The training and validation data set included 2557 and 614 GPR images, respectively. The training and validation data set had GPR images under three transmitting frequencies from four highways. The rest images were used as a testing data set.

	RC	WP	US
RC	87.14	0.29	0.37
WP	0.28	90.04	0.31
US	0.86	0.50	88.41
BG	11.72	9.17	10.91

(a) XB Highway

	RC	WP	US
RC	89.21	0.61	0.17
WP	0.20	89.97	0.61
US	0.26	0.14	89.02
BG	10.23	9.28	10.20

(b) XY Highway

	RC	WP	US
RC	89.71	0.31	0.10
WP	0.37	89.74	0.33
US	0.49	0.22	91.43
BG	9.43	9.73	10.98

(a) XB Highway

	RC	WP	US
RC	88.21	0.24	0.27
WP	1.32	89.97	0.28
US	0.56	0.43	90.72
BG	9.89	9.36	8.73

(b) XY Highway

	RC	WP	US
RC	87.95	0.60	0.36
WP	0.13	91.78	0.35
US	0.42	0.15	88.97
BG	11.50	7.47	10.32

(c) HP Highway

	RC	WP	US
RC	88.94	0.12	0.67
WP	0.25	90.45	0.24
US	0.06	0.17	87.64
BG	10.75	9.26	11.45

(d) PY Highway

	RC	WP	US
RC	89.63	0.17	0.08
WP	0.24	88.76	0
US	0.35	0.83	90.32
BG	9.78	10.24	9.67

(c) HP Highway

	RC	WP	US
RC	88.61	0.24	0.43
WP	0.90	90.25	0.76
US	0.25	0.50	91.69
BG	10.24	9.01	7.42

(d) PY Highway

Fig. 14. Confused matrices of precisions in different highways. RC, WP, US, and BG are reflection cracks, water-damage pits, uneven settlement, and background.

Fig. 15. Confused matrices of recalls in different highways. RC, WP, US, and BG are reflection cracks, water-damage pits, uneven settlement, and background.

The processes in [Section 2.2](#) were conducted in a Python environment on a computer with a Core i7 8750H @ 3.4 GHz CPU, 32 GB of DDR4 memory, and an 8 GB NVIDIA 1080 GPU.

3. Results and discussion

3.1. Training and testing results

The four steps in [Section 2.2](#) were conducted to train the 30 Faster R-ConvNets, whose aim was to find the optimal one with reasonable anchor sizes and ratios. The intersection over unions (IoUs) of the three distresses and mean IoU in the validation were used to evaluate the performance of the 30 Faster R-ConvNet during the training. The IoUs of the three distresses were computed by Equation (3) as

$$\text{IoU} = \frac{\text{Intersection}}{\text{Distress} + \text{Bounding box} - \text{Intersection}} \quad (3)$$

In this study, the mean IoU was the average of the IoUs of the reflection cracks, water-damage pits, and uneven settlements.

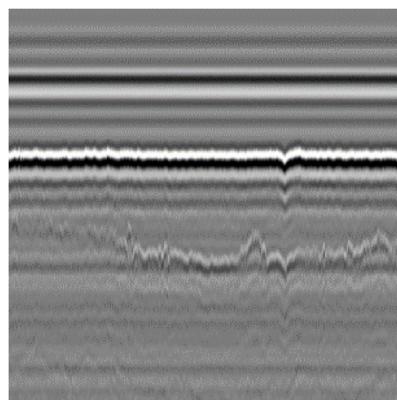
[Fig. 9](#) presents the IoUs of the three pavement distresses and mean IoU. The maximum IoU in the 30 Faster R-ConvNets was 88.34% in No. 11 case. However, the 85.24% IoU of uneven settle-

ments in No. 11 case was not acceptable, though the IoUs of reflection cracks and water-damage pits were 86.49% and 93.27%, respectively, which were desirable. Considering the task of the Faster R-ConvNet was to detect all the three distresses, it was necessary to select a Faster R-ConvNet whose all IoUs of the three distresses were acceptable and balanceable. Compared with No. 11 case, the average IoUs of the three distresses in the No. 10 Faster R-ConvNet were balanceable, which were 86.97%, 87.98%, and 89.99%. Therefore, the No. 10 case was considered as the optimal one, whose ratios were 0.3, 0.7, and 1.6 and sizes were 16, 64, and 128.

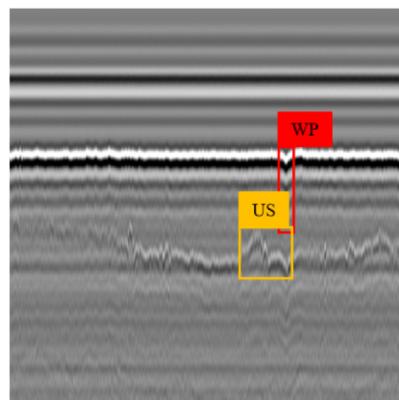
Recall and precision were also used to further evaluate the trained Faster R-ConvNets as

$$\text{Recall} = \frac{\text{Intersection}}{\text{Distress}} \quad (4)$$

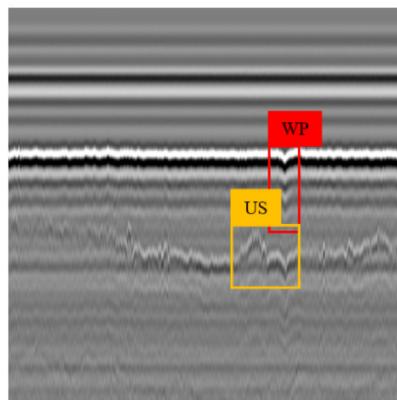
$$\text{Precision} = \frac{\text{Intersection}}{\text{Bounding box}} \quad (5)$$



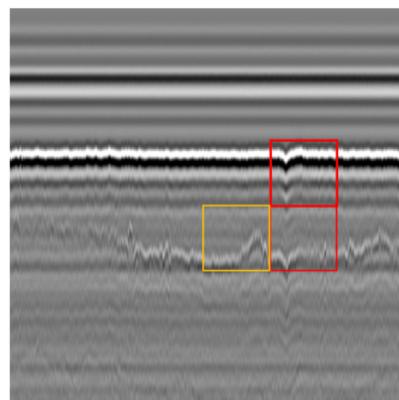
(a) Original GPR image



(b) Result of the Faster R-ConvNet



(c) Result of the cascade ConvNet



(d) Result of the region ConvNet

Fig. 16. Detection results of the three methods in PY Highway and 900 MHz.

The validation recalls and IoUs were computed by Eqs. (4) and (5). Fig. 10 presents mean average precisions, recalls, and IoU of the 30 Faster R-ConvNets. The recall and precision of the No 10 case were 89.13% and 89.09%, respectively.. It indicated that the No. 10 Faster R-ConvNet was a reasonable and balanceable one, though its precision, recall, and IoU were not the best among the 30 cases. Therefore, the No. 10 case was regarded as the optimal one.

The testing data set mentioned in Section 2.3 was used to test the trained No. 10 Faster R-ConvNet. Table 3 presents the precisions, recalls, and IoUs of the No. 10 Faster R-ConvNet in the testing. Fig. 11 shows some testing results. We could find that the testing performance of the No. 10 Faster R-ConvNet was closed to the validation performance. It indicated that the No. 10 Faster R-ConvNet had a great generalization and out-of-sample ability. We could also find that the testing performance in different distresses was similar. It indicated that the No. 10 Faster R-ConvNet could detect reflection cracks, water-damage pits, and uneven settlements well. Thus, we concluded that the No. 10 Faster R-ConvNet could detect asphalt pavement distresses in various real-world condition. The No. 10 Faster R-ConvNet was well-trained in general.

3.2. Stability analysis

In the pavement inspection engineering, it is important to guarantee the model to work well in pavements with different struc-

tures. Thus, a stability analysis based on the testing results was conducted to verify the stability of the model in different real-world conditions, including different transmitting frequencies and highway structures.

(1) Different transmitting frequencies

The testing data set was divided based on the transmitting frequencies including 300, 500, and 900 MHz. The ratio of GPR images in the three transmitting frequencies was 1:1:1. The confused matrices shown in Fig. 12 and Fig. 13 present the precisions and recalls of the No. 10 Faster R-ConvNet in different transmitting frequencies. We found that the precisions and recalls of the No. 10 Faster R-ConvNet decreased with the decrease of the transmitting frequencies, though the precisions and recalls in 300 MHz was still acceptable. For example, the recalls of the water-damage pits in 900 MHz and 300 MHz were 93.25% and 86.82%, respectively. Further, the IoUs in 300, 500, and 900 MHz were 84.35%, 86.27%, and 88.10%, respectively. It indicated that a low transmitting frequency had a negative influence on the detection accuracy of the model because less detail of the pavement distresses was collected and shown in the GPR images in a low transmitting frequency condition. Therefore, some necessary features related to the distresses were not acquired by the convolutional and pooling layers, which led to a deviation of the outputs of the softmax and regression layer. The cascade ConvNet system [41] may be a solution to the problem.

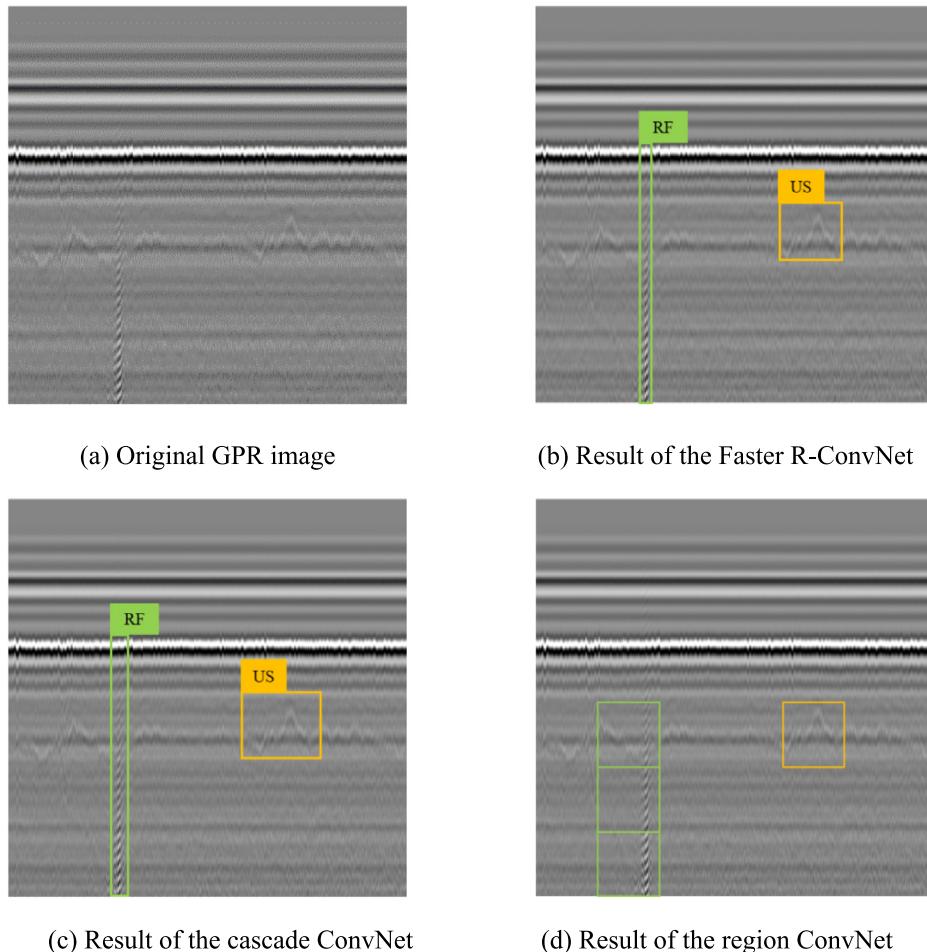


Fig. 17. Detection results of the three methods in PY Highway and 600 MHz.

(2) Different highway structures

The testing data set was divided based on the four highways shown in Fig. 1. The ratio of GPR images in the four highways was 1:1:1:1. The confused matrices shown in Figs. 14 and 15 present the precisions and recalls of the No. 10 Faster R-ConvNet in different highways. The precisions and recalls of the No. 10 Faster R-ConvNet in four highways were similar. For example, the precisions of reflection cracks in the four highways were 87.14%, 89.21%, 87.95%, and 88.94%, respectively. Additionally, the IoUs of the four highways were 86.03%, 86.27%, 86.34%, and 86.32%, respectively. It indicated that the pavement structures and materials had limited influence on the precision and recall of the model. There were three reasons why the performance of the No. 10 Faster R-ConvNet was not influenced by the pavement structures and materials: (1) the difference of the pavement structures and materials had limited influence on the performance of the GPR, which guaranteed the quality of the GPR images; (2) the kernels of convolutional layers had learned low-, medium-, and high-features related to pavement distresses, which could extract distresses from GPR images under different conditions during the testing; (3) the pooling layers improved the tolerance of the Faster R-ConvNet to the limited influence of the pavement structures and materials [42].

3.3. Comparative study

In order to compare the performance of the No. 10 Faster R-ConvNet with other state-of-the-art methods for detecting asphalt pavement distresses, we selected three GPR images from the testing dataset to conduct a comparative study, as shown in Fig. 16 (a)-Fig. 18 (a). Considering some studies [24,27] proved that traditional unsupervised methods were not suitable for distress

detection, only two surprised methods were employed in the comparison study. The first one was named cascade ConvNet-based method [43], where a cascade ConvNet was used to classify and detect pavement distresses using cropped GPR images. The second one was named region ConvNet using uncropped GPR images [26], whose work procedures were similar to the RNPs. The two models were also trained by the training data set presented in Section 2.3.

Figs. 16 and 17 present detection results of the three methods for GPR images from PY Highway in 900 MHz and 600 MHz. All the three methods detected distresses successfully. It indicated that the deep-learning methods had a good capacity for handling complexity background and foreground in GPR images. In this study, the foreground was pavement distresses in the GPR images, while the background was the part of GPR images with no distress. However, the precisions and recalls of the three methods were different due to their different working principles. The performance of the cascade ConvNet-based method was worst in the three models because it could only process cropped GPR images. It led its precision and recall were limited by the sizes of the cropped GPR images. Thus, there were many backgrounds included in the bounding box. The performance of the Faster R-CNN was better than the performance of the region ConvNet. This was because the Faster R-CNN detected distresses using local features proposed by the RNP, and a correction was then conducted in the Fast R-ConvNet part. Thus, the detection results were more precise than local features. However, the region ConvNet, whose working principle was similar to a RNP, outputted the local features directly as results. In a word, the Faster R-CNN had a superiority in the working principles for detecting distresses.

Fig. 18 presents detection results of the three methods for a GPR image from HP Highway in 300 MHz. It could be found that the

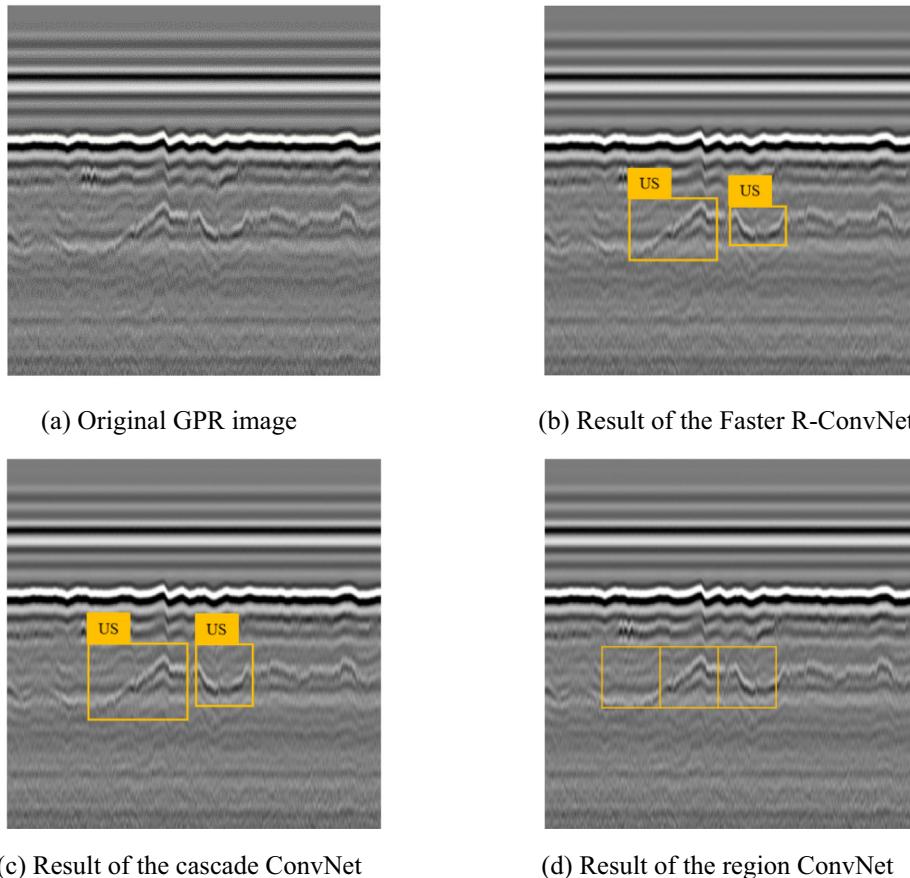


Fig. 18. Detection results of the three methods in HP Highway and 300 MHz.

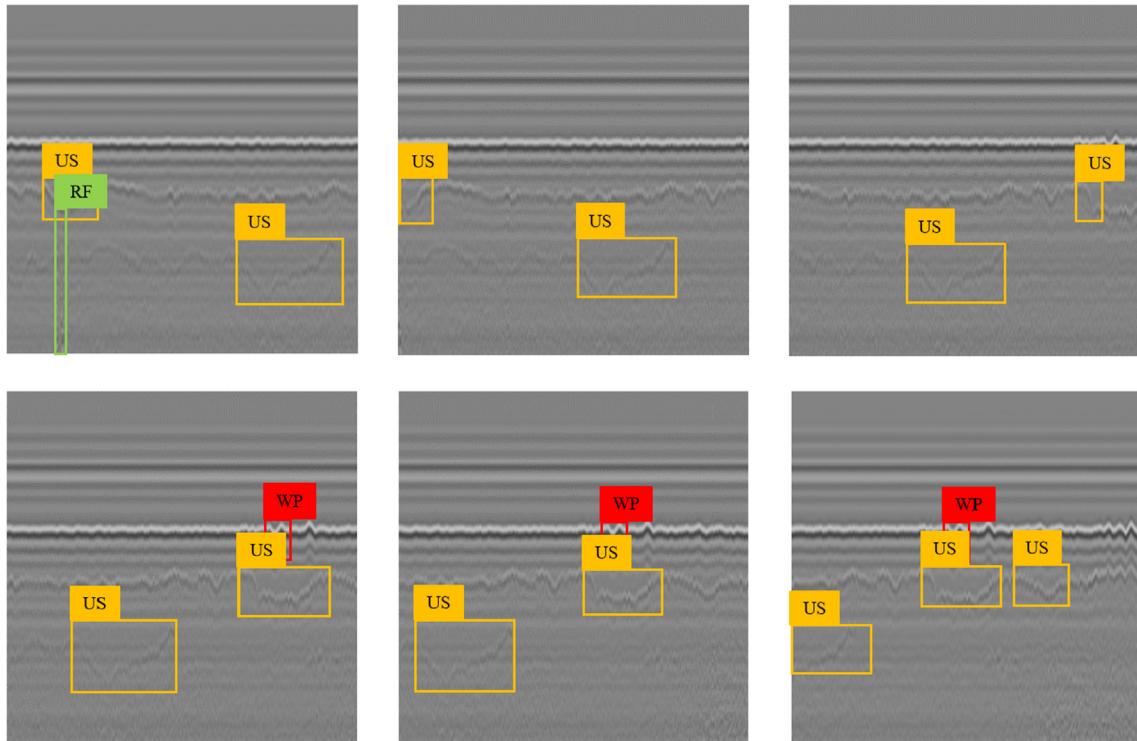


Fig. 19. Example of the real-time detection.

performance of the Faster R-CNN and region ConvNet decreased, while the performance of the cascade ConvNet-based method was stable. It indicated that the performance of the Faster R-CNN and region ConvNet were influenced by the transmitting frequencies in some degree, though the results were still acceptable. It also indicated that the cascade structure for deep learning was a way to solve the problem.

3.4. A real-time detection

As shown in Fig. 2, a GPR image could only reflect a part of the highway. To evaluate the entire conditions of highways, a real-time detection should be realized. The No. 10 Faster R-ConvNet realized the detection of a GPR image with a pixel resolution of 512×512 in GPU mode in approximately 0.03 s. A GPR image with a pixel resolution of 512×512 presented a 0.6 m highway. The speed of the testing vehicle was 20–25 m/s, which were less than a window increment speed of the GPR and the runtime of the Faster R-CNN. Therefore, the detection procedure did not introduce a delay.

Real-time detection results are shown in Fig. 19. As shown in Fig. 19, the method based on the Faster R-ConvNet had the capacity of achieving a real-time pavement distress detection in GPU mode, which could be used to replace a human-assisted pavement condition assessment based on real-time GPR data.

4. Conclusions

In this study, the Faster R-ConvNets combining RPNs and a Fast R-ConvNet were utilized to detect asphalt pavement distresses using GPR images, and following conclusion can be drawn:

- (1) The No. 10 Faster R-ConvNet, whose anchors ratios were 0.3, 0.7, and 1.6 and sizes were 16, 64, and 128, achieved the 86.53%, 87.26%, and 84.93 IoUs of reflection cracks, water-damage pits, and uneven settlements, respectively. Addi-

tionally, its precisions of the three distresses were 88.31%, 90.56%, and 88.51%. Thus, the No. 10 Faster R-ConvNet was regarded as the optimal one.

- (2) The performance of the No. 10 Faster R-ConvNet was not obviously influenced by the pavement structures and materials. However, its performance was influenced by the transmitting frequencies of the GPR due to the detail loss of the pavement distresses in the GPR images.
- (3) Compared with other supervised methods, the Faster R-ConvNet could detect pavement distresses more precisely under various real-world conditions in pavements. This was because the local features generated by the RPN provided a reasonable area for the Fast R-ConvNet part to regress the distress location. Additionally, nine anchors with different sizes and ratios had a better capacity to detect pavement distresses than cropped images with a fixed size.
- (4) A real-time detection was realized to locate distresses in asphalt pavements. The time for the Faster R-ConvNet to detect a GPR image with 512×512 pixels in GPU mode was approximately 0.3 s. The results provided the possibility for a real-time pavement detection using GPR images. The real-time detection results showed that the performance of the Faster R-ConvNet was stable and reasonable.

CRediT authorship contribution statement

Jie Gao: Conceptualization, Methodology, Resources, Supervision. **Dongdong Yuan:** Investigation, Data curation. **Zheng Tong:** Formal analysis, Writing - review & editing, Writing - original draft, Software. **Jiangang Yang:** . **Di Yu:** .

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study is jointly funded by the Science and Technology Research Project of Jiangxi Provincial Department of Education (Grant No. GJJ190361), the Fundamental Research Funds for the Central Universities, CHD (Grant No. 300102210501). This work is also supported by Co-operation Program with the UTs and INSAs (France) funded by the China Scholarship Council (No. CSC201801810108).

References:

- [1] S. Li, C. Yuan, D. Liu, H. Cai, Integrated processing of image and GPR data for automated pothole detection, *J. Comput. Civil Eng.* 30 (6) (2016) 04016015.
- [2] H. Ahn, J. Lee, Subsurface Condition Evaluation of Asphalt Pavement for Pavement Preservation Treatments, *J. Test. Eval.* 44 (3) (2015) 1183–1193.
- [3] S.H. Dessouky, J. Oh, M. Ilias, S.I. Lee, D. Park, Investigation of various pavement repairs in low-volume roads over expansive soil, *J. Perform. Constr. Facil.* 29 (6) (2014) 04014146.
- [4] X. Cui, J. Zhang, N. Zhang, Y. Zhou, Z. Gao, W. Sui, Laboratory simulation tests of effect of mechanical damage on moisture damage evolution in hot-mix asphalt pavement, *Int. J. Pavement Eng.* 16 (8) (2015) 699–709.
- [5] Y.C. Wu, G. Cascante, M.D. Pandey, Condition assessment of longitudinal pavement joints using ultrasonic surface waves, *Can. J. Civ. Eng.* 41 (12) (2014) 1019–1028.
- [6] P. Hao, M. Zhang, Q. Zhang, J. Xu, S. Dong, Study on evaluation method of mud-pumping of cement concrete bridge deck pavement, *Constr. Build. Mater.* 167 (2018) 106–114.
- [7] Y. Chen, G. Tebaldi, R. Roque, G. Lopp, Y. Su, Effects of interface condition characteristics on open-graded friction course top-down cracking performance, *Road Materials and Pavement Design* 13 (sup1) (2012) 56–75.
- [8] M.G. Arab, R.A. Mousa, A.R. Gabr, A.M. Azam, S.M. El-Badawy, A.F. Hassan, Resilient Behavior of Sodium Alginate-Treated Cohesive Soils for Pavement Applications, *J. Mater. Civ. Eng.* 31 (1) (2018) 04018361.
- [9] Y.K. Cho, T. Bode, J. Song, J.H. Jeong, Thermography-driven distress prediction from hot mix asphalt road paving construction, *Journal of Construction Engineering and Management* 138 (2) (2011) 206–214.
- [10] F.M. Fernandes, J.C. Pais, Laboratory observation of cracks in road pavements with GPR, *Constr. Build. Mater.* 154 (2017) 1130–1138.
- [11] M. Solla, S. Lagüela, H. González-Jorge, P. Arias, Approach to identify cracking in asphalt pavement using GPR and infrared thermographic methods: Preliminary findings, *Ndt & E International* 62 (2014) 55–65.
- [12] P. Szymczyk, M. Szymczyk, Non-destructive building investigation through analysis of GPR signal by S-transform, *Autom. Constr.* 55 (2015) 35–46.
- [13] X. Xu, Q. Zeng, D. Li, J. Wu, X. Wu, J. Shen, GPR detection of several common subsurface voids inside dikes and dams, *Eng. Geol.* 111 (1–4) (2010) 31–42.
- [14] H.L. Zhou, Y.L. Jiang, L.H. Xu, G.Q. Liang, Automatic detection algorithm for expressway subgrade diseases based on SVM, *Zhongguo Gonglou Xuebao (China Journal of Highway and Transport)* 26 (2) (2013) 42–47.
- [15] Yanqing, X., Xiaoming, H., Wei, Z., and Shunzhi, Q. (2013). Improvement on Recognition Method of Void beneath Slab Based on Nondestructive Testing Technologies. In: Sustainable Construction Materials 2012 (pp. 77–91).
- [16] F. Tosti, A. Benedetto, Pavement pumping prediction using ground penetrating radar, *Procedia-Social and Behavioral Sciences* 53 (2012) 1044–1053.
- [17] C. Barat, C. Ducottet, String representations and distances in deep convolutional neural networks for image classification, *Pattern Recogn.* 54 (2016) 104–115.
- [18] B. Shi, X. Bai, C. Yao, Script identification in the wild via discriminative convolutional neural network, *Pattern Recogn.* 52 (2016) 448–458.
- [19] B. Leng, Z. Xiong, X. Fu, A 3D shape retrieval framework for 3D smart cities, *Frontiers of Computer Science in China* 4 (3) (2010) 394–404.
- [20] Y.J. Cha, W. Choi, G. Suh, S. Mahmoudkhani, O. Büyüköztürk, Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types, *Comput.-Aided Civ. Infrastruct. Eng.* 33 (9) (2018) 731–747.
- [21] H. Liu, A. Sha, Z. Tong, J. Gao, Autonomous microscopic bunch inspection using region-based deep learning for evaluating graphite powder dispersion, *Constr. Build. Mater.* 173 (2018) 525–539.
- [22] Z. Tong, J. Gao, A. Sha, L. Hu, S. Li, Convolutional neural network for asphalt pavement surface texture analysis, *Comput.-Aided Civ. Infrastruct. Eng.* 33 (12) (2018) 1056–1072.
- [23] Z. Tong, J. Gao, H. Zhang, Innovation for evaluating aggregate angularity based upon 3D convolutional neural network, *Constr. Build. Mater.* 155 (2017) 919–929.
- [24] Y.J. Cha, W. Choi, O. Büyüköztürk, Deep learning-based crack damage detection using convolutional neural networks, *Comput.-Aided Civ. Infrastruct. Eng.* 32 (5) (2017) 361–378.
- [25] Sh.a. Aimin, Tong Zheng, Gao Jie, Recognition and measurement of pavement disasters based on convolutional neural networks, *China journal of highway and transport* 31 (1) (2018) 1–10.
- [26] Z. Tong, J. Gao, H. Zhang, Recognition, location, measurement, and 3D reconstruction of concealed cracks using convolutional neural networks, *Constr. Build. Mater.* 146 (2017) 775–787.
- [27] Z. Tong, J. Gao, H. Zhang, Innovative method for recognizing subgrade defects based on a convolutional neural network, *Constr. Build. Mater.* 169 (2018) 69–82.
- [28] R. Ali, Y.J. Cha, Subsurface damage detection of a steel bridge using deep learning and uncooled micro-bolometer, *Constr. Build. Mater.* 226 (2019) 376–387.
- [29] G.H. Beckman, D. Polyzois, Y.J. Cha, Deep learning-based automatic volumetric damage quantification using depth camera, *Autom. Constr.* 99 (2019) 114–124.
- [30] G. Hinton, L. Deng, D. Yu, G. Dahl, A.R. Mohamed, N. Jaitly, T. Sainath, Deep neural networks for acoustic modeling in speech recognition, *IEEE Signal Processing magazine* (2012) 29.
- [31] Abdel-Hamid, O., Mohamed, A. R., Jiang, H., and Penn, G. (2012, March). Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition. In: 2012 IEEE international conference on Acoustics, speech and signal processing (ICASSP) (pp. 4277–4280). IEEE.
- [32] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems (pp. 1097–1105).
- [33] Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. R. (2012). Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint arXiv:1207.0580.
- [34] JTG H20-2007, Highway Performance Assessment Standards, Beijing, China.
- [35] G.U.O. Shili, Z.H.U. Peimin, S.H.I. Xinghua, L.I. Xiuzhong, Comparative analysis on response of ground penetrating radar wave field to crack width, *Chinese Journal Of Radio Science* 28 (1) (2013) 130–136 (in Chinese).
- [36] Z. Tong, D. Yuan, J. Gao, Y. Wei, H. Dou, Pavement-distress detection using ground-penetrating radar and network in networks, *Constr. Build. Mater.* 233 (2020) 117352.
- [37] Zeiler, M. D., and Fergus, R. (2013). Stochastic pooling for regularization of deep convolutional neural networks. arXiv preprint arXiv:1301.3557.
- [38] Ren, S., He, K., Girshick, R., and Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. In: Advances in neural information processing systems (pp. 91–99).
- [39] S.J. Pan, Q. Yang, A survey on transfer learning, *IEEE Transaction on Knowledge Discovery and Data Engineering* 22 (10) (2010).
- [40] Mishkin, Dmytro, and Jiri Matas. (2015). All you need is a good init. arXiv preprint arXiv:1511.06422.
- [41] W. Yang, Y. Chen, Y. Liu, L. Zhong, G. Qin, Z. Lu, W. Chen, Cascade of multi-scale convolutional neural networks for bone suppression of chest radiographs in gradient domain, *Med. Image Anal.* 35 (2017) 421–433.
- [42] Scherer, D., Müller, A., and Behnke, S. (2010, September). Evaluation of pooling operations in convolutional architectures for object recognition. In: International conference on artificial neural networks (pp. 92–101). Springer, Berlin, Heidelberg.
- [43] M. Fu, P. Xu, X. Li, Q. Liu, M. Ye, C. Zhu, Fast crowd density estimation with convolutional neural networks, *Eng. Appl. Artif. Intell.* 43 (2015) 81–88.