A New Method for Automated Monitoring of Road Pavement Aging Conditions Based on Recurrent Neural Network

Xiao Chen, Xianfeng Zhang, Jonathan Li, *Senior Member, IEEE*, Miao Ren, and Bo Zhou

***Abstract*— The automated monitoring of road pavement conditions is a challenging subject in intelligent transportation. However, the existing studies mostly focus on extracting pavement damages such as cracks, while the pavement aging conditions are still less investigated. In this paper, a novel method based on a modified recurrent neural network is designed for automated monitoring of asphalt pavement aging phenomena from fine-resolution satellite imagery. A spectral augmentation method is proposed to enhance the spectral details of the road pavements. A novel loss function is also proposed to improve the bi-directional gated recurrent unit (Bi-GRU) network in order to better classify different degrees of road pavement aging and non-pavement objects. In order to demonstrate the outperformance of the modified network Bi-GRU+, the Worldview-2 satellite image (16360 *7728) covering 16 asphalt roads in the southwestern suburb of Beijing City is used. The results show that the proposed approach has better performance than existing machine learning methods, with an overall accuracy of 98.16% and a Kappa coefficient of 0.97. The overall processing time of the proposed method is 7836 seconds in our case study. The proposed method is efficient for large-scale monitoring of road health conditions from fine-resolution satellite imagery. It can become a part of intelligent transportation and provide a new foundation for large-range automated monitoring of road pavement aging conditions.**

***Index Terms*— Remote sensing, recurrent neural network, gated recurrent unit (GRU), asphalt pavement, aging conditions, multispectral imagery.**

I. Introduction

ROAD networks connecting buildings, villages, cities, and countries are the most important transportation infrastructure in modern life. The conditions of road pavements are directly related to the driving experience, traveling comfort, and driving safety [1].

The road pavement conditions can be simply divided into two components: pavement aging [2], [3], [4] and pavement damages [5], [6], [7], [8], [9]. For flexible pavement systems, pavement aging denotes the quality of asphalt pavement degradation over time caused by weathering, loads, and moisture [2], [10]. The damages such as cracks, alligator cracks, and potholes are likely to appear on the heavily aged pavement. Therefore, road maintenance engineers should regularly assess the road conditions and timely conduct maintenance to extend the road life [11].

A common method for road pavement condition assessment is field inspection by engineers in order to evaluate the physical and chemical parameters of the pavement using indices such as the pavement condition index, structural index, and overall pavement condition index [3], [12]. In recent years, the advanced pavement management system (PMS) mounted on a vehicle has been used for road pavement inspection. The PMS includes several sensors and equipment, including GIS, GNSS, laser scanners, image acquisition systems, odometers, and ground-penetrating radar [13], [14], [15], in order to instantaneously evaluate the road health and provide information for maintenance strategies and decision-making [15], [16]. However, the field investigation and the PMS are labor-intensive and have technical limitations, such as single-lane detection, traffic obstruction, road surface destruction, costly large-scale monitoring, and time-consuming and laborious [2], [17].

In recent years, with the advancement of remote sensing technology and computer algorithms, high-resolution satellite imagery and deep learning approaches have been used for pavement conditions mapping [2], [17], [18], [19]. However, most of the existing deep learning algorithms focus on pavement damage monitoring and are suitable for fine-scale monitoring of pavement quality in a small range, but not suitable for monitoring large-scale pavement aging processes. In addition, the conventional shallow machine learning algorithms rely on the manual selection of features or thresholds, while having a low degree of automation and relatively poor accuracy. Therefore, the existing methods fail to combine deep learning and high-resolution satellite imagery well. A method that can detect road aging conditions over a large range with a high automation and generalization ability is not yet developed.

A novel method for monitoring asphalt pavement aging conditions based on a deep learning architecture is proposed to solve this problem. The proposed method does not require manual selection of features and thresholds and can be used for large-range asphalt pavement aging monitoring. The proposed deep learning network model is improved on the basis of the gated recurrent unit (GRU) to be suitable for monitoring pavement aging conditions. To the best of our knowledge, it is the first time developing an RNN-based classification model

Manuscript received 31 December 2021; revised 15 April 2022 and 4 June 2022; accepted 24 August 2022. This work was supported by the National Foundation of Natural Science of China under Grant 42171327. The Associate Editor for this article was S. S. Nedevschi. *(Corresponding author: Xianfeng Zhang.)*

Xiao Chen, Xianfeng Zhang, Miao Ren, and Bo Zhou are with the Institute of Remote Sensing and Geographic Information System, Peking University, Beijing 100871, China (e-mail: xfzhang@pku.edu.cn).

Jonathan Li is with the Department of Geography and Environmental Management, University of Waterloo, Waterloo, ON N2L 3G1, Canada.

Digital Object Identifier 10.1109/TITS.2022.3204334

used to monitor asphalt pavement aging conditions from submeter resolution satellite imagery.

The contributions of this paper are summarized as follows:

(1) We propose a biGRU-based road pavement aging classification method that does not require the manual selection of features or thresholds and is highly automated. To our knowledge, this is the first approach that uses RNN network architecture to develop an algorithm for monitoring asphalt pavement aging conditions.

(2) Combined with the characteristics of the pavement aging process, a novel loss function is proposed based on the cross-entropy loss. It can allow the model to focus on learning "hard samples" and "similar samples" and improve the model's ability to classify different degrees of asphalt pavement aging.

(3) Using satellite imagery as the data source, an automated pavement aging monitoring method including spectral augmentation and statistical analysis is developed. It allows the monitoring of large-range pavement aging conditions.

The remainder of this paper is organized as follows. The Related works are reviewed in Section II. Section III addresses the details of the proposed methods. The experiments and results of this study are presented in Section IV, and the discussion is provided in Section V. Finally, the conclusions are drawn in Section VI.

II. Related Work

A. Pavement Aging Monitoring

The asphalt pavement aging conditions can be monitored using spaceborne and aerial remote sensing imagery [19]. Multispectral and hyperspectral image data cover a broad spectral range and have a relatively fine spectral resolution, allowing the monitoring of the pavement aging process [4], [11], [20], [21], [22], [23], [24].

The pavement quality changes with the composition and content of asphalt concrete, which results in changes of the pavement spectral reflectance. For instance, the volatilization and oxidation of asphalt can be characterized by the absorption of iron oxides at 520, 670, and 870 nm, while the slope of the spectral curve between the visible and near-infrared wavelengths increases as the asphalt pavement ages. Based on these characteristics, Herold *et al.* [3], [4] construct three spectral indices to characterize the slope in the visible to near-infrared (VNIR) bands, in order to represent the degree of pavement aging. However, these indices are less efficient for monitoring seriously aged pavements and are only suitable for relatively healthy pavements.

The field investigation demonstrated that the asphalt content of pavement decreased with the increase of the service life, which results in increasing pavement reflectivity and brightness in the images [20]. Therefore, asphalt pavement with different aging conditions can be monitored based on the pavement brightness [20], [25]. Mettas *et al.* [23] divide the aged pavements into three levels and explore the possibility of using Landsat7 ETM+ data to monitor the pavement conditions. They deduced that the three aging levels of asphalt pavements had significant spectral differences, which demonstrated the potential of using satellite imagery to monitor the pavement aging conditions. Mohammadi [26] and Andreou [27] use hyperspectral images as data sources. They deduce that the Spectral Angle Mapper(SAM) is relatively efficient and more suitable for pavement aging conditions mapping. Pan *et al.* [28] extracted the aging classes of asphalt pavement using multiple endmember spectral mixture analysis (MESMA) and Worldview-2 images. Their method minimizes the influence of mixed pixels on the classification result.

However, these approaches for monitoring pavement aging conditions based on a small sample size have low automation, require threshold selection and have poor generalization ability when the regions of interest or image data change. These problems have limited the application of remote sensing in practical road pavement conditions monitoring.

B. Deep Learning

Deep learning is a data-driven approach which provides high accuracy, automation, and high generalization performance for massive samples [29]. Researchers have discovered the potential of deep learning in road pavement conditions monitoring. Most of the existing deep learning architectures for monitoring road pavement conditions focus on extracting cracks and other pavement damages. They can be grouped into three categories.

The first category is the image classification into healthy and damaged road sections [30], [31], [32]. This type of method has a high classification accuracy and demonstrates the potential of deep learning for pavement monitoring. It is often combined with other methods to detect road damage. For instance, Pan *et al.* [33] perform multi-scale segmentation of road images and use the AlexNet to obtain different damage categories.

The second category is the object detection, which determines the road damage in very fine resolution imagery using a detection box [34], [35], [36]. It includes the one-step methods that directly detect the damage [9], [37], [38], and the two-step methods that use a detection step following classification [39], [40], [41], [42]. These methods generally used rectangular boxes to locate the detected damaged objects. Due to the fact that road pavement damages have irregular shapes, these methods have limited identification accuracy and cannot obtain the shape of the damaged sections for subsequent road health assessment.

The third category is the image segmentation, which classifies each pixel into healthy and damaged pavement [29], [33], [43], [44], [45], [46]. In general, pavement damages comprise a relatively small area in road pavement images. Thus, these methods suffer from the sample imbalance problem. In addition, some studies propose improvements in the network structure, such as the feature pyramid and hierarchical boosting network (FPHBN) [43], in which the samples are weighted layer by layer. The model focuses on learning difficult samples (i.e., hard samples), minimizing the sample imbalance problem. Thus the approach performs well on multiple sample sets. Kang *et al.* [18], [47], [48] perform crack segmentation in complex environments and different lighting conditions by integrating three independent computer

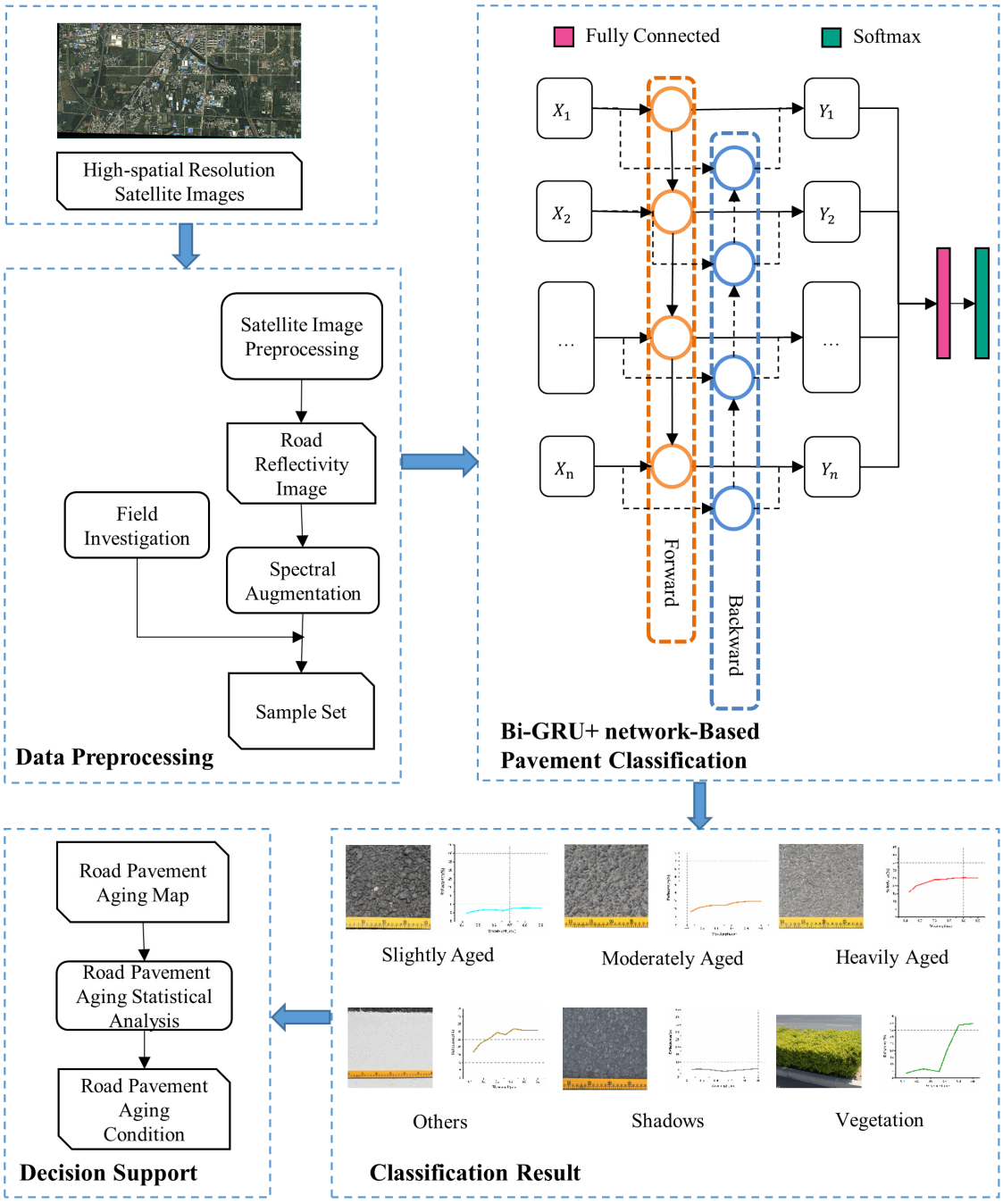

Fig. 1. Framework for asphalt pavement aging monitoring from fine-resolution satellite imagery using the Bi-GRU+ network.

vision algorithms and developed a new encoder with an attention module. Choi *et al.* [49] propose a real-time crack segmentation DL architecture, referred to as SDDNet-V1, which can greatly improve the time efficiency and identify relatively vague cracks.

These deep learning models are mainly based on two-dimensional convolutional neural networks (2DCNNs). Their high image processing ability ensures the detection of pavement damages (e.g. cracks) with distinguishing spatial characteristics. However, different pavement aging degrees are characterized by spectral characteristics, and 2DCNNs are not well suited for fully using the abundant spectral information in multispectral or hyperspectral data. In addition, the spatial resolution of the remotely sensed data required for pavement aging monitoring is often not sufficient for damage monitoring, and the spatial information is less, which further limits the application of 2DCNN in pavement aging monitoring. Therefore, these existing deep learning methods cannot be applied to road pavement aging mapping. Therefore, new deep learning architectures are required to monitor pavement aging conditions from fine-resolution satellite imagery.

III. Framework and Methods

A. Framework Architecture

The proposed framework (Fig. 1) is composed of three parts: data preprocessing, Bi-GRU+ network-based pavement

classification and decision support. The satellite image is processed to produce a road reflectance image in the data preprocessing part. In addition, a nonlinear function is proposed for spectral augmentation and the field investigation is performed within three months of the satellite image acquisition from which the true aging conditions of the road pavements are collected. The Bi-GRU+ network is developed and used to classify the augmented reflectance image pixel by pixel. After the statistical analysis of the classification results, the pavement aging conditions of each road are derived. They can be used as decision support for the road maintenance department

B. Data Preprocessing

In order to meet the requirements of large-scale road pavement aging monitoring, the proposed framework uses fine spatial resolution multispectral satellite images as the data source. The spatial resolution should be less than one meter, and the spectrum should cover at least visual and near-infrared (VNIR) bands. A radiometric correction of the image is performed using the gain and offset data provided in the image metadata. In addition, an atmospheric correction is performed using the Fast Line-of-sight Atmospheric Analysis of Hypercubes (FLAASH). The resultant reflectance images are then clipped to contain only the road pavements, in order to reduce the influence of non-pavement areas on the classification process. A field investigation within three months of the image acquisition date is performed to collect the true conditions of road pavement aging in the study area.

The aging of asphalt pavement is a gradual process. The field investigations demonstrated that the aging of asphalt pavement can be divided into three categories [2], i.e., slightly aged, moderately aged, and heavily aged. The slightly aged pavement refers to a small reduction in the asphalt content due to volatilization, oxidation, absorption, and photochemical reactions. The asphalt oil film has mostly disappeared. However, the gravel aggregate has not been exposed (Fig. 2b). In moderately aged pavement, friction damage occurs due to the traffic load and physical weathering, causing the exposure of gravel aggregate in the asphalt mixture (Fig. 2c). When the exposed gravel aggregate is subjected to ongoing friction and pressure of vehicle loads and weathering, the coarse aggregate breaks and drops, which results in heavily aged pavement (Fig. 2d). During the aging processes, the deformability and structural strength of the road pavement is degraded. Eventually, a wide range of road pavement damage occurs in different forms, such as cracks, potholes, etc.

The Munsell neutral value scale card (MNVSC) is used in the field investigation as a quantitative reference to describe asphalt pavement aging conditions. The MNVSC divides the gray color from pure black to pure white into 37 levels according to the ISCC-NBS international standard [50]. Each level has a value, which is a dimensionless number. Field comparisons and visual discriminations are performed on the color of the asphalt pavement of different aging degrees, in order to obtain the gray values of the underlying asphalt pavement. The three categories of aging conditions (Table I) correspond to the grayscale values: slightly aged asphalt pavement ([N0.5/-N4.25/]), moderately aged asphalt pavement ([N4.25-N6.75/]), and heavily aged asphalt pavement ([N6.75/-N8.75/]). The level [N8.75/N9.5/] is white, and it is ignored because it does not correspond to any asphalt pavement.

TABLE I

ISCC-NBS COLOR NAMES, VALUES AND SPECTRAL REFLECTANCE OF THE MNVSC FOR THREE ASPHALT AGING CONDITIONS

Color Name	Munsell Value	Reflectance (%)	Road Aging Condition
Black	[N0.5/, N2.25/]	[0.6,3.8]	
Dark Gray to Black	(N2.25/, N2.75/]	(3.8,5.5]	Slightly aged
Dark Gray	(N2.75/, N4.25/]	(5.5,13.7]	
Medium to Dark Gray	(N4.25/, N4.75/]	(13.7,17.6]	Moderately aged
Medium Gray	(N4.75/, N6.25/]	(17.6,33.0]	
Medium to Light Gray	(N6.25/, N6.75/]	(33.0,39.5]	
Light Gray	(N6.75/, N8.25/]	(39.5,63.6]	Heavily aged
White to Light Gray	(N8.25/, N8.75/]	(63.6,73.4]	
White	(N8.75/, N9.5/]	(73.4,90.0]	Ignored

An analytical spectral devices full range (ASD Field Spec-FR, ASD Corporation) is used to record spectral data. The instrument has three detectors covering the RGB, near-infrared (VNIR) and a short-wave infrared (SWIR1 and SWIR2) band, with a spectral sampling interval of 1.4 nm for the VNIR detector and 2.0 nm for the SWIR. A field spectral measurement is performed between local time 11:00 and 13:00 under a clear sky, and the dark current is removed at every beginning of a measurement. The optical fiber bundle collects the reflected radiation with a 25-deg conical field of view at 50 cm above the pavement, which corresponds to a 22 cm$\times$22 cm region. Besides the ground targets, a white reference is measured with the Spectral on Panel (Lab sphere Inc., North Sutton, New Hampshire) to standardize and calculate the spectral reflectance of all the ground targets. The spectrum of every target, including the white reference, is recorded five times per measurement, and then the mean of five spectra was used in further processing. In addition to the asphalt pavement, the spectral information of other related objects including vegetation, cement sidewalks, traffic lines, and bare soil was also collected. Based on the field investigation, the pavement features are categorized into three aging conditions (slightly aged, moderately aged and heavily aged), vegetation, shadows and others six categories. The others category includes traffic lines, skyways and so on.

The spectral characteristics of the categories are summarized as follows:

1) As the pavement ages, the reflectance of the asphalt pavement increases. The reflectance values of the slightly aged pavement range from 5% to 10%, those of the moderately aged pavement range from 12% to 20%, and that of the heavily aged pavement range from 14% to 35%.

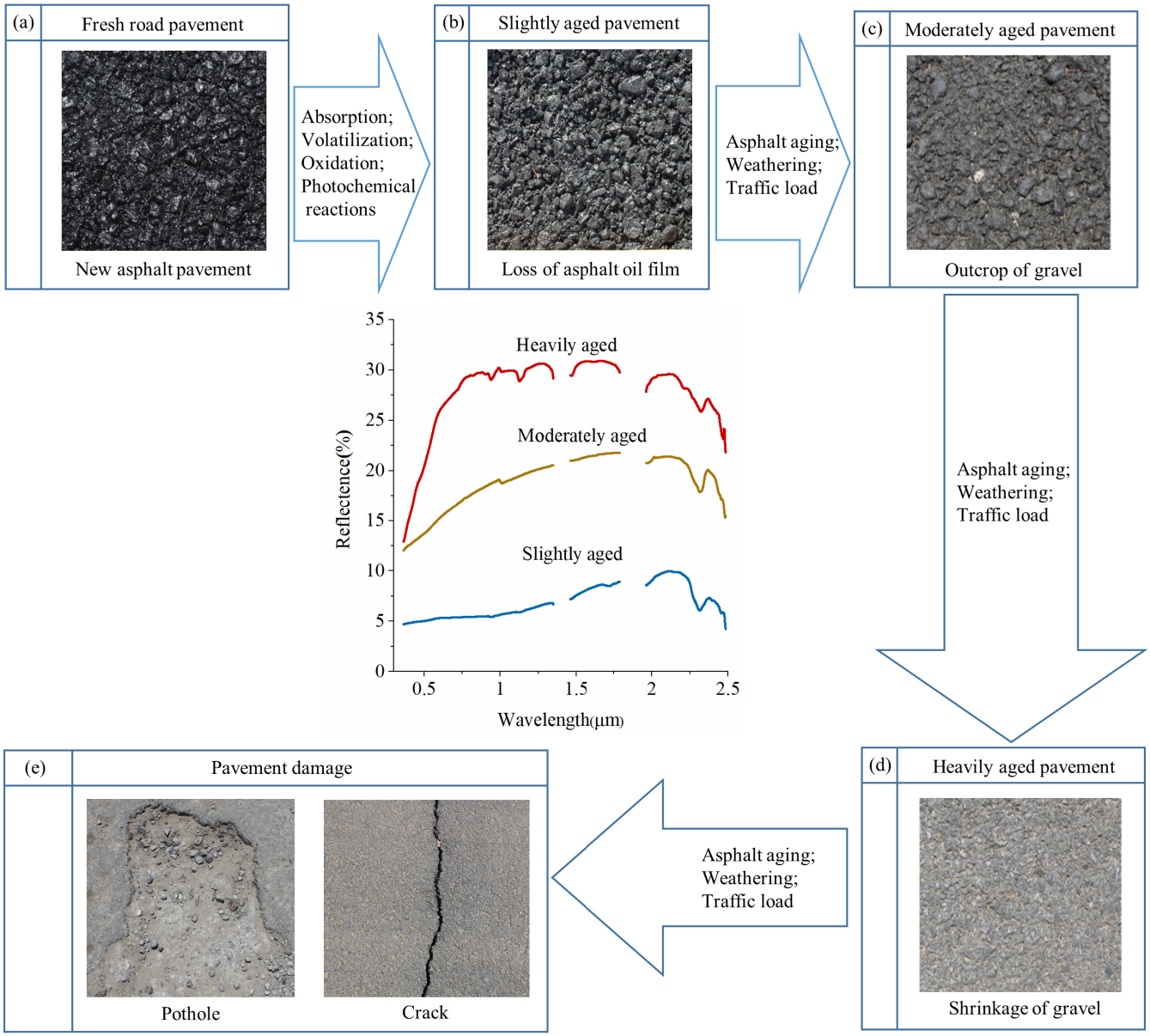

Fig. 2. Illustration of the aging processes of asphalt pavement over time. (a) Fresh pavement, (b) Slightly aged, (c) Moderately aged, (d) Heavily aged, and (e) Pavement damages (e.g. crack, pothole, etc.). The pavement photos were captured by Canon 500, and the spectral curves were made from the in-situ measurements of pavement spectra using the portable ASD device.

2) When the pavement ages, the slope of the reflectance curve in the VNIR bands increases.

3) The reflectance values of the pavement surface and other related objects range from 0% to 40%.

In general, the spectral characteristics of different aging conditions are quite clear. The specific environment may affect them. For example, the spectral characteristics of shadows and heavily aged pavement are relatively similar. Therefore, a spectral augmentation method is proposed in this paper. Equation (1) is used for spectral augmentation.

$$x = -(\rho - 1)^2 + 1 \tag{1}$$

where $\rho \epsilon [0, 1]$ is the original reflectance and $x \epsilon [0, 1]$ is the augmented spectral reflectance.

Equation (1) can enlarge the value of $\rho \epsilon [0, 0.4]$ to $x \epsilon [0, 0.64]$ without changing the value range of ρ, which can enlarge the spectral characteristics of different categories.

C. Bi-GRU+ Network-Based Pavement Classification

In contrast to the traditional feed-forward neural networks such as CNNs, the hidden layers of an RNN are connected between nodes and form a directed graph along a sequence. Therefore, an RNN has a high performance for sequential data analysis, and it has been widely used in natural language processing [51]. However, the RNN is problematic for spectral data, which results in gradient disappearance and the inability to obtain spectral contextual information [52]. To address these problems, a Bi-GRU+ network model is proposed.

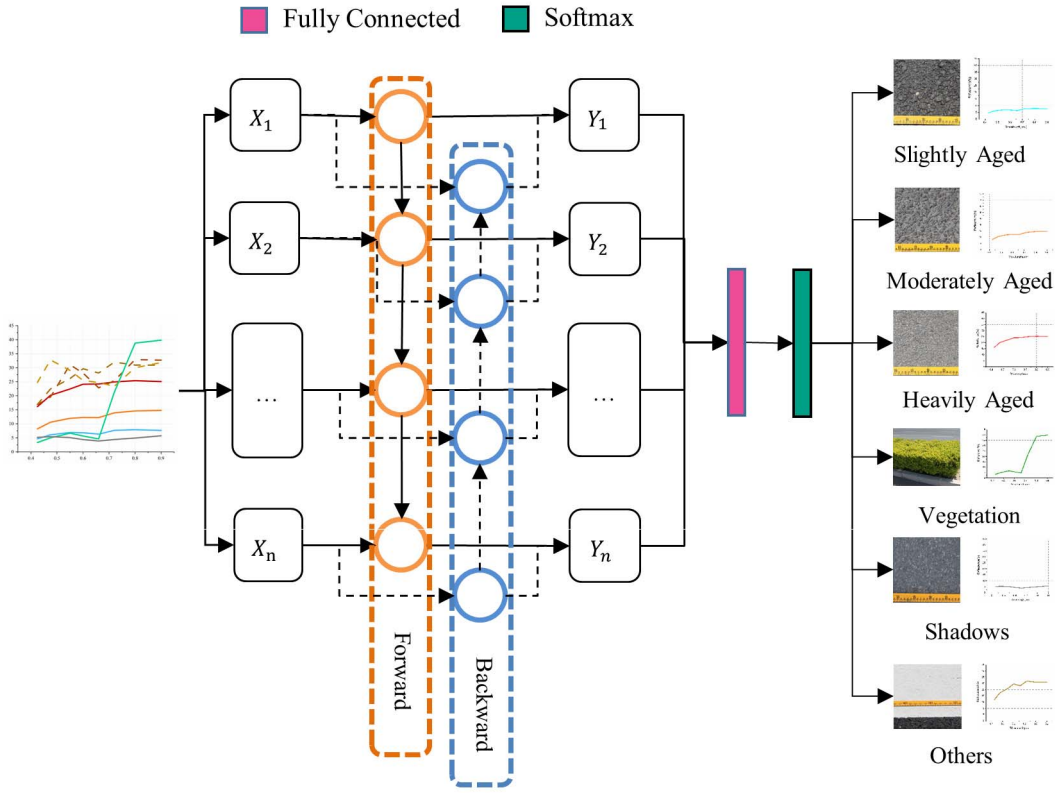

Fig. 3. Diagram of the proposed Bi-GRU+ network.

The proposed network is less affected by gradient disappearance, has fewer parameters and higher efficiency than the long short-term memory network (LSTM) [53]. It can also extract contextual information from spectral sequence data to learn spectral characteristics, which makes it suitable for multispectral or hyperspectral image data. In addition, the proposed loss function is improved using cross-entropy, in order to be suitable for pavement aging monitoring.

The architecture of the proposed network is shown in Fig. 3. Each pixel of a multispectral or hyperspectral image is considered as a spectral sequence $x = (x_1, x_2, \ldots\ldots, x_n)$, where x_t represents the pixel value in band t. Consequently, the remote sensing images can be transformed into spectral sequence data as input into the network. The input spectral sequence data contains contextual information between the spectra. This information is learned using the forward and backward hidden layers and is input into a fully connected layer to obtain the spectral features. The Softmax classifier is then used to predict the classes.

Complex environments can cause problems such as sample imbalance or mixed pixels in actual scenes. For example, the proportion of different aging conditions on the road pavement is often different, which causes the problem of sample imbalance. These problems make some pixels, denoted by "hard samples", difficult to classify. On the other hand, the aging of asphalt pavement is a gradual process. Thus some pixels are difficult to classify between two different aging conditions, denoted by "similar samples".

To solve these problems, the cross-entropy function is improved as (2).

$$L_{loss} = \sum_{i=0}^{C-1} \left(\alpha * \left(1 - p_i\right) * p_i - \left(1 - p_i\right)^2 * y_i * \log p_i\right) \quad (2)$$

where L_{loss} is the improved cross-entropy loss function, C is the number of categories, y_i is the one-hot encoding of the i-th category, and p_i is the predicted probability of the i-th category. Compared with the original cross-entropy loss function L_{cross}, the modulating factor $\left(1 - p_i\right)^2$ is used to make the model focus on training the hard samples. As $p_i \to 1$, the factor tends to 0 and the loss for well-classified samples is down-weighted. Moreover, an additional item $\alpha * \left(1 - p_i\right) * p_i$ is added to improve the classification result of similar samples, where $\alpha \geq 0$ is the tunable hyperparameter. As $p_i \to 0.5$, the item increases and the loss for similar samples is up-weighted.

The advantage of the GRU is that hidden states can be selectively reset and updated by the reset gate and update gate. The advantage of the proposed network is that it can focus on the feature bands to distinguish between different objects. The network hidden layer is computed as:

$$\begin{cases} \overrightarrow{R_t} = \sigma\left(x_t \overrightarrow{W_{xr}} + H_{t-1} \overrightarrow{W_{hr}} + \overrightarrow{b_r}\right) \\ \overrightarrow{Z_t} = \sigma\left(x_t \overrightarrow{W_{xz}} + H_{t-1} \overrightarrow{W_{hz}} + \overrightarrow{b_z}\right) \end{cases} \quad (3)$$

$$\overrightarrow{\tilde{H}_t} = \tanh\left(x_t \overrightarrow{W_{xh}} + \left(\overrightarrow{R_t} \odot H_{t-1}\right) \overrightarrow{W_{hh}} + \overrightarrow{b_n}\right) \quad (4)$$

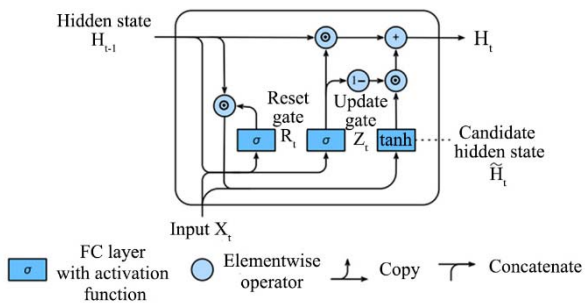

Fig. 4. Calculation of the forward hidden layers.

$$\overrightarrow{H_t} = \overrightarrow{Z_t} \odot H_{t-1} + \left(1 - \overrightarrow{Z_t}\right) \odot \overrightarrow{\widetilde{H_t}} \tag{5}$$

Equations (3)-(5) are used for the forward hidden layers, where $\boldsymbol{x}_t$ is the input data, and $\overrightarrow{\boldsymbol{H}_t}$ is the output at step t (Fig. 4).

In (3), $\overrightarrow{\boldsymbol{R}_t}$ and $\overrightarrow{\boldsymbol{Z}_t}$ represent the reset gate and update gate of the forward hidden layer, $\boldsymbol{H}_{t-1}$ is the previous hidden state, $\overrightarrow{\boldsymbol{W}_{xr}}$, $\overrightarrow{\boldsymbol{W}_{hr}}$, $\overrightarrow{\boldsymbol{W}_{xz}}$, and $\overrightarrow{\boldsymbol{W}_{hz}}$ are the weight parameters, $\overrightarrow{\boldsymbol{b}_r}$ and $\overrightarrow{\boldsymbol{b}_n}$ are the biases, and σ represents the sigmoid activation function. More precisely, the reset gate $\overrightarrow{\boldsymbol{R}_t}$ controls how the hidden state of the previous step flows into the candidate hidden state of the current step. In addition, the hidden state of the previous step contains all the historical information of the sequence data up to the previous step. Therefore, the reset gate can be used to drop prediction-irrelevant historical information, as expressed in (4). The update gate $\overrightarrow{\boldsymbol{Z}_t}$ helps to capture partial long-term dependencies in the sequence data, as shown in (5). Equation (4) is used to calculate the candidate hidden state. $\overrightarrow{\boldsymbol{W}_{xh}}$ and $\overrightarrow{\boldsymbol{W}_{hh}}$ are the weight parameters, $\overrightarrow{\boldsymbol{b}_n}$ is the biases, $\odot$ denotes the multiplication by the element, and $\left(\overrightarrow{\boldsymbol{R}_t} \odot \boldsymbol{H}_{t-1}\right)$ represents the output of the reset gate of the current step and the hidden state of the previous step. If the element value in the reset gate is close to 0, the reset gate corresponds to the hidden state element of 0. That is, the hidden state of the previous step is discarded. If the element value is close to 1, the hidden state of the previous step is retained. Subsequently, the element multiplication of the result is used to link to the input of the current step. The candidate hidden state is calculated by the fully connected layer with the tanh activation function.

Equation (5) is used to calculate the hidden state in the current step. Since (4) uses the sigmoid activation function, the update gate $\overrightarrow{\boldsymbol{Z}_t}$ ranges between 0 and 1. $\overrightarrow{\boldsymbol{Z}_t} \odot \boldsymbol{H}_{t-1}$ and $\left(1 - \overrightarrow{\boldsymbol{Z}_t}\right) \odot \overrightarrow{\widetilde{\boldsymbol{H}_t}}$ represent the network with the memory and the forgotten sequences of the above and current information, respectively.

The backward hidden layer structure is similar to the forward hidden layer structure. It extracts the following information of the sequence data using (6) $\sim$ (8).

$$\begin{cases} \overleftarrow{R}_t = \sigma\left(x_t \overleftarrow{W_{xr}} + H_{t+1} \overleftarrow{W_{hr}} + \overleftarrow{b_r}\right) \\ \overleftarrow{Z_t} = \sigma\left(x_t \overleftarrow{W_{xz}} + H_{t+1} \overleftarrow{W_{hz}} + \overleftarrow{b_z}\right) \end{cases} \tag{6}$$

$$\overleftarrow{\widetilde{H}_t} = \tanh\left(x_t \overleftarrow{W_{xh}} + \left(\overleftarrow{R_t} \odot H_{t+1}\right) \overleftarrow{W_{hh}} + \overleftarrow{b_n}\right) \tag{7}$$

$$\overleftarrow{H_t} = \overleftarrow{Z_t} \odot H_{t+1} + \left(1 - \overleftarrow{Z_t}\right) \odot \overleftarrow{\widetilde{H_t}} \tag{8}$$

The forward state $\overrightarrow{\boldsymbol{H}_t}$ and back hidden state $\overleftarrow{\boldsymbol{H}_t}$ are merged into $\boldsymbol{H}_t$ and sent $\boldsymbol{H}_t$ to the fully connected layer. The final classification results are obtained using the Softmax classifier.

D. Decision Support

The classification map is achieved using the abovementioned approach. Some statistical analysis can be extracted from the classification map so as to obtain the pavement aging conditions in the study area. The aging conditions and their distribution on each road pavement will serve as further decision support for the road maintenance department. The heavily aged road pavements are often accompanied by cracks and other damages, which seriously affect driving safety and should be maintained as soon as possible. There is a risk of further deterioration of moderately aged road pavements, which also requires attention. The proportion of pixels with different aging degrees on each road is counted, and they are weighted as 0.05, 0.3, and 0.65 corresponding to slightly, moderately, and heavily aged, so as to calculate the pavement aging index of each road. The larger the index value, the more serious the aging condition of the road pavement. If the index value is greater than 0.5, the road surface is heavily aged and needs maintenance.

E. Accuracy Assessment

The overall accuracy (OA), average accuracy (AA), class accuracy, Kappa coefficient K, Macro Precision, Macro Recall, and Macro F1 are used to evaluate the classification performance of the proposed network on the test dataset. The OA is the ratio between the number of correctly classified pixels and the total number of pixels in the test set, while the AA is the average of the class accuracies. The K is an index used to determine whether the model prediction results and actual classification results are consistent, as defined in (9).

$$\mathrm{K} = \frac{OA - \frac{\sum_k n_{k1} * n_{k2}}{N^2}}{OA + \frac{\sum_k n_{k1} * n_{k2}}{N^2}} \tag{9}$$

where N is the sum of the confusion matrix elements, n_{k1} is the sum of elements in row k, n_{k2} is the sum of elements in col k.

Macro Precision, Macro Recall, and Macro F1 are the averages of the Precision, Recall, and F1 of all the classes, respectively. The larger the values of the metrics, the better the performance.

IV. Experiments and Results

All the experiments are implemented on a Windows Server with Python3.6, Pytorch1.8.1 [54] and scikit-learn [55] (CPU: Intel Xeon Silver 4116@2.1GHz; RAM: 128GB; GPU: NVIDIA RTX 1080TI).

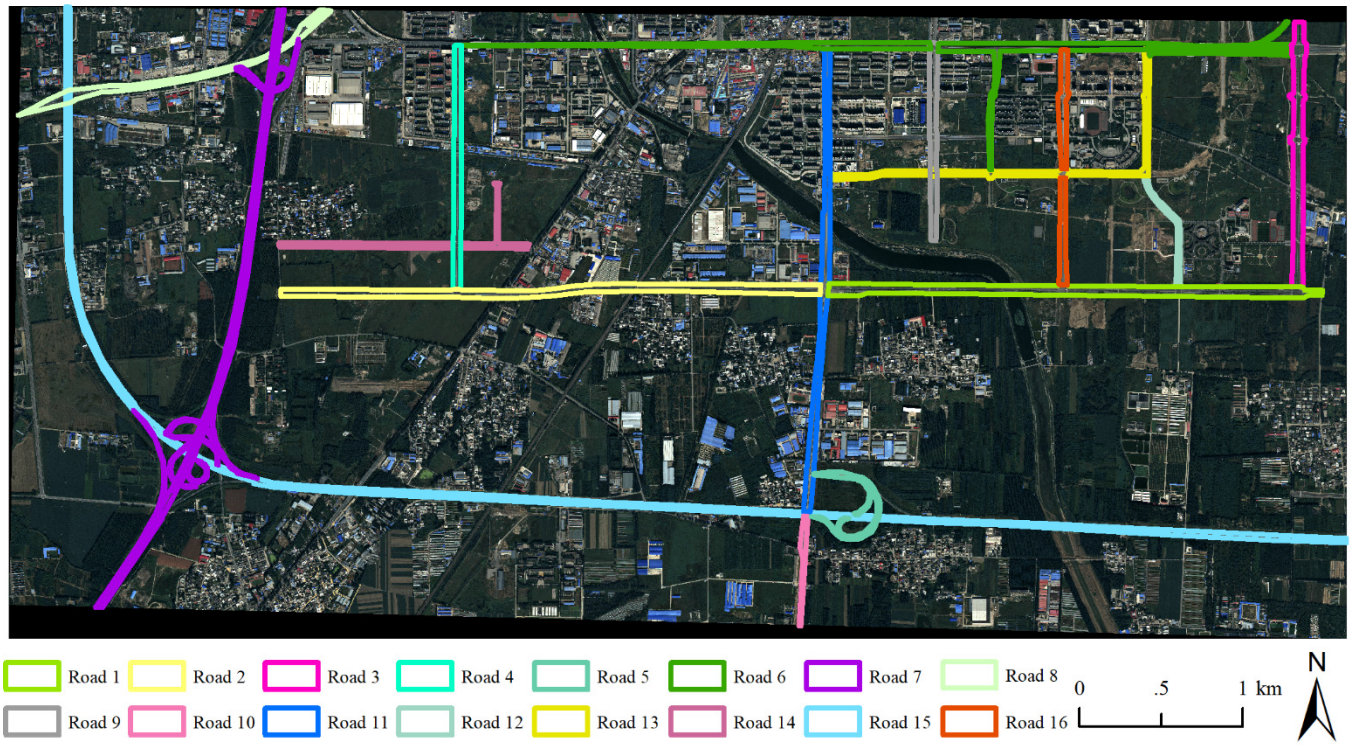

Fig. 5. Worldview-2 satellite image and roads in the study area.

A. Study Area and Sample Set

Part of Fangshan district, southwest of Beijing City, China, is considered as the study area. This area is relatively spacious, with light traffic, no tall buildings and little shadow on roads. Thus, it is suitable for monitoring road aging conditions from satellite remote sensing data. There are 16 asphalt-paved roads located in the study area, covering high-grade roads and urban secondary roads (Fig. 5). For the convenience of description, the roads accordingly are coded from No.1 to No.16. The field investigation shows that these road pavements are in different aging conditions.

Worldview-2 satellite image data covering the study area are used to evaluate the proposed framework. The Worldview-2 sensor has eight bands with wavelengths covering visible and near-infrared bands (0.4–1.04μm). The image was acquired on 21 September 2013. The Gram-Schmidt algorithm is used to pan sharpen the panchromatic and multispectral image. The resultant reflectance images have a spatial resolution of 0.5 m.

Combined with the field investigation results, 17000 pixels in total are selected on the five road images for the categories of different aging degrees (Road 1, Road 3, Road 12, Road 13, and Road 16) by visual interpretation and field investigation. The train and validation sets are randomly generated using a ratio of 7:3, and 49000 pixels on the remaining roads are extracted to generate a test set (Table II).

TABLE II
NUMBER OF TRAIN/VALIDATION/TEST SAMPLES

No.	Category	Train	Validation	Test
1	Slightly aged	2800	1200	8000
2	Moderately aged	4200	1800	20000
3	Heavily aged	2800	1200	16000
4	Others	700	300	1000
5	Vegetation	1050	450	3500
6	Shadows	350	150	500
	Total	11900	5100	49000

B. Hyperparameter Setting

The three main hyperparameters (the learning rate, the factor α in the loss function and the hidden state unit) significantly affect the results. The impact of different hyperparameters on the model performance is evaluated. The adaptive moment estimation (Adam) [56] optimizer is used. In extreme cases, a large learning rate may lead to fluctuations in the model accuracy during training, and a small learning rate may result in an inability of model fitting. In this study, three initial learning rates of 0.000001, 0.0001, and 0.001 are assessed, and the optimal initial learning rate is 0.001.

The factor α can control how much the model focuses on the similar samples. The larger the value of α is, the more the model will focus on the learning of similar samples, but it may affect the learning effect of ordinary samples. It is deduced that $\alpha = 0.1$ leads to the best results in the experiments.

The hidden unit determines the number of neurons in the implied hidden layer. The larger the number, the more difficult the training, while a small number of hidden units may also reduce the model performance. The hidden unit variable is determined as the exponent of two (i.e. 32, 64,

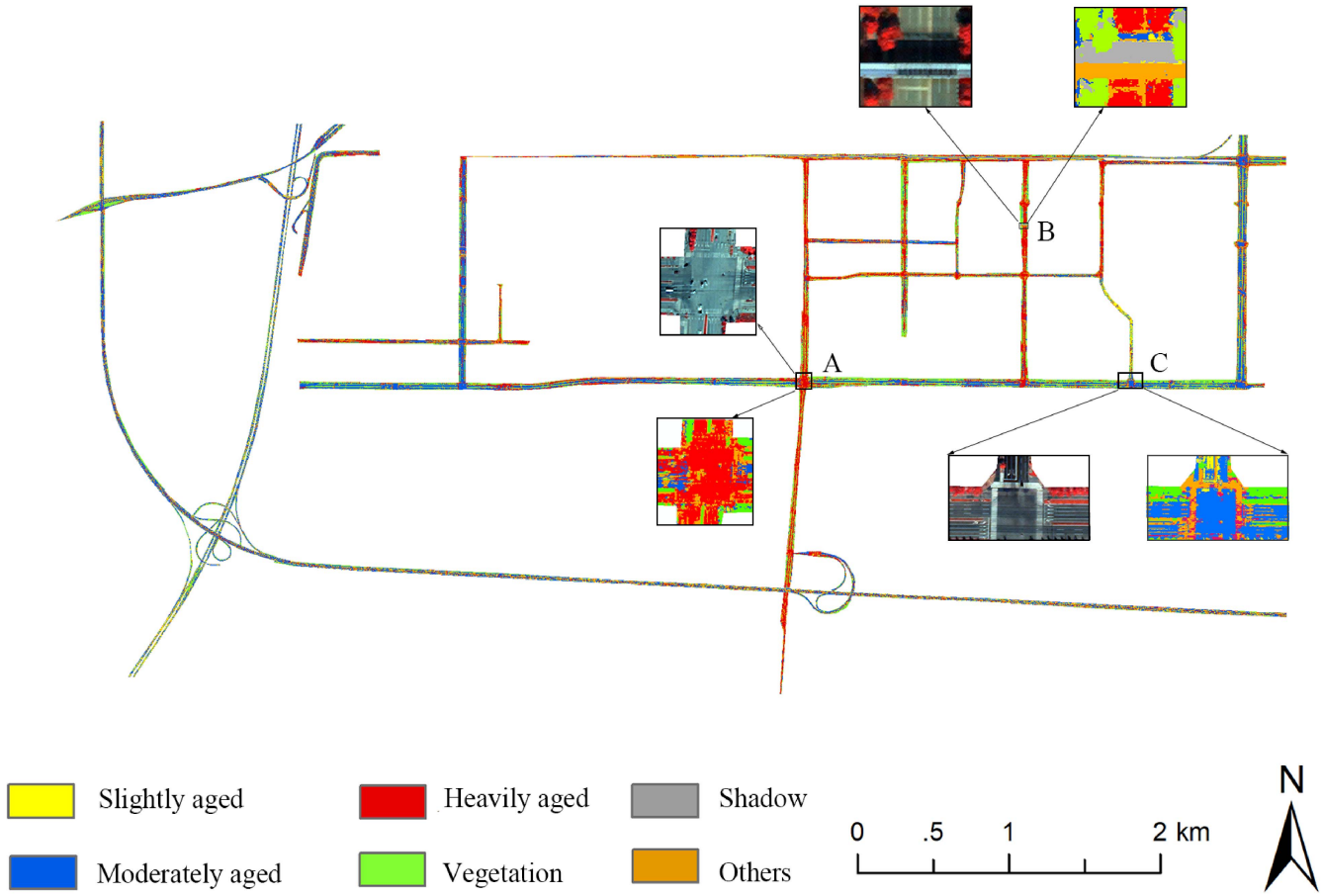

Fig. 6. Classification of road pavement conditions from the Worldview-2 images using the Bi-GRU network. The zoom-in images are satellite images and classification results in the A, B and C locations.

TABLE III
CONFUSION MATRIX OF THE CLASSIFICATION ON THE TEST SAMPLE SET

	Slightly aged	Moderately aged	Heavily aged	Others	Vegetation	Shadows	**User's accuracy (%)**
Slightly aged	7727	250	23	0	0	0	**96.56**
Moderately aged	177	19593	30	0	0	0	**98.95**
Heavily aged	19	183	15798	200	0	0	**97.52**
Others	0	6	6	986	2	0	**98.6**
Vegetation	0	1	0	3	3496	0	**99.89**
Shadows	3	0	0	0	0	497	**99.4**
Producer's accuracy (%)	**97.49**	**97.8**	**99.62**	**82.93**	**99.94**	**100**	

128, 256, 512,1024, and 2048), in order to find the optimal hyperparameters. The results show that 512 hidden units lead to the best performance in the study area.

C. Results and Accuracy Validation

The classification results of the road surface aging conditions from the Worldview-2 images of the 16 roads in the study area are presented in Fig. 6. The different aging conditions in the road images and the different categories of image elements are accurately classified with OA of 98.16%, AA of 98.49%, Kappa coefficient of 0.97, Macro Precision of 96.17%, Macro Recall of 98.67%, and Macro F1 of 97.17% (Table III).

After the statistical analysis, the pavement ageing index of each road is calculated to quantitatively assess the aging conditions of the asphalt pavement in the study area (Fig. 7). Consistent with the field investigations, Road 12 has the best road condition with an aging index value of 0.1515. Roads 7 and 15 are also in good condition, with aging index values of 0.251 and 0.2965 respectively. On the contrary, Roads 14, 9, 13, 16, 11, and 10 are heavily aged, with aging index

TABLE IV
COMPARISON BETWEEN THE CLASSIFICATION ACCURACIES (%) OF DIFFERENT METHODS ON THE TEST SAMPLE SET

Category	SAM	MESMA	ResNet18	ResNeXt50	1DCNN	Bi-GRU+
Slightly aged	47.38	80.3	15.54	14.19	94.12	96.59
Moderately aged	33.99	75	47.52	37.2	93.64	98.95
Heavily aged	88.21	83.05	27.09	18.99	97.54	97.52
Others	24.8	78.2	8.38	14.61	74	98.6
Vegetation	100	-	17.78	23.23	99.94	99.89
Shadows	39.6	82.76	1.06	1.94	93.2	99.4
OA	58.69	81.71	34.49	27.03	98.12	**98.16**
AA	55.66	79.86	17.52	18.33	93.24	**98.49**
K	44.39	77	6.93	4.58	97.17	**97.35**
Macro Precision	53.15	78.61	17.5	19.5	90.33	**96.17**
Macro Recall	56.8	79.86	19	18.17	93.33	**98.67**
Macro F1	50.1	79.23	15.3	15.5	91.33	**97.17**

values greater than 0.5, and should be maintained as soon as possible. The information obtained by the proposed method is valuable for management departments to facilitate the efficient and timely road maintenance.

D. Comparative Study

The proposed Bi-GRU method is compared with several widely used state-of-the-art methods, including SAM, MESMA, ResNet18, ResNeXt50 and 1DCNN in order to demonstrate its efficiency.

SAM is a physically-based spectral classification that uses an n-D angle to match pixels to reference spectra. This method is often used to monitor pavement aging conditions with hyperspectral imagery [26], [27]. MESMA [2] is currently the state-of-the-art method for monitoring aging from satellite imagery. This method divides Vegetation into Others. ResNet [57] and has a high performance in image classification. ResNeXt [58] combined the advantages of Inception [59] and ResNet. It has been widely used for image classification. In this paper, ResNet18 and ResNeXt50 are used to test whether the traditional CNN network can be applied to aging monitoring. The sample set of ResNet18 and ResNeXt50 is the patches generated by the initial sample set, in which each pixel is converted into a $3*3$ patch. A one-dimensional convolutional neural network (1DCNN) is designed for comparison. Its architecture is shown in Fig. 8. The random seed is set to 0 in the experiments to avoid the effect of randomness on the model accuracy.

The classification accuracies obtained by different methods are presented in Table IV. The classification result obtained by the SAM method is poor, which proves that it is not suitable for multispectral satellite imagery. Although MESMA is the state-of-art method, it is far less accurate than the proposed method. The ResNet18 and ResNeXt50 have the worst performance and are mainly unable to distinguish each category. More precisely, they have a serious over-fitting phenomenon. This is mainly because satellite images cannot provide enough geometric information of pavement aging. This confirms that 2DCNN can't be directly applied to pavement aging monitoring, as previously mentioned. The 1DCNN model achieved accuracy second only to the proposed method Bi-GRU+. However, the classification accuracy for heavily aged and vegetation was poor. Compared with the proposed method, 1DCNN can't learn well the contextual information between spectra.

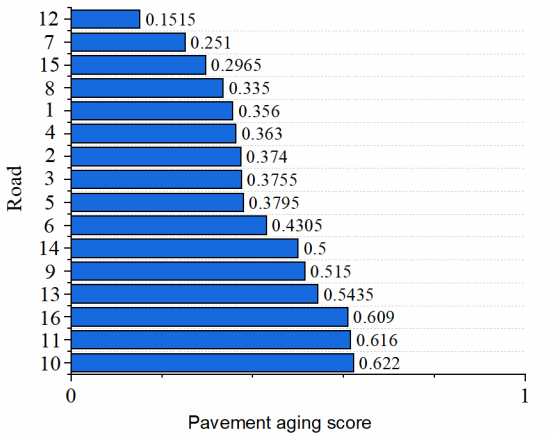

Fig. 7. Pavement aging index value for each road in the study area.

The proposed Bi-GRU+ model achieves better performance for learning spectral contextual information from different bands than the CNN methods, allowing it to distinguish classes with similar spectral features. Thus, the Bi-GRU+ achieves the highest OA, AA, Kappa coefficient, Macro Precision, Macro Recall, and Macro F1. Moreover, it is more important that the proposed model can identify severely aging target roads in practical applications. Therefore, although the accuracy of 1DCNN is relatively close to Bi-GRU+, our method is still the best one and most suitable for practical applications.

E. Ablation Study

The Bi-GRU+ learns the contextual features of the spectral bands for extracting spectral features. It can distinguish different aging conditions from satellite images. Tackling the problem of "hard samples" and "similar samples" in road condition monitoring, an improved loss function based on the cross-entropy function is proposed. The influence of each module of the proposed Bi-GRU+ Network on the classification performance on the test set is evaluated.

The detailed classification results with different modules are presented in Table V. Noted that GRU denotes the

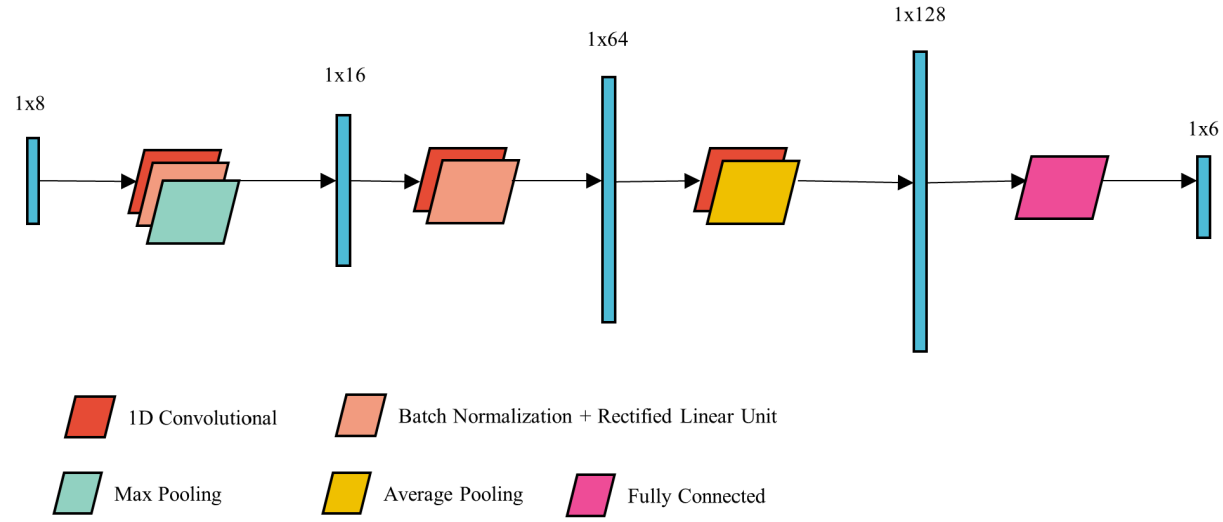

Fig. 8. The architecture of 1DCNN.

TABLE V
ABLATION EXPERIMENT ON THE SAMPLE SET

Category	GRU	GRU+	Bi-GRU	Bi-GRU+
Slightly aged	70.64	82.9	99.19	96.59
Moderately aged	98.07	99.04	97.14	98.95
Heavily aged	97.86	97.14	97.35	97.52
Others	98.4	98.1	97.8	98.6
Vegetation	99.89	99.91	100	99.89
Shadows	98.8	99.2	99.08	99.4
OA	94.43	96.39	97.79	**98.16**
AA	94.26	96.3	**98.55**	98.49
K	91.86	94.75	96.83	**97.35**
Macro Precision	95.83	96	95.5	**96.17**
Macro Recall	94.33	96.3	98.5	**98.67**
Macro F1	94.5	96	97	**97.17**

normal GRU, Bi-GRU means the bi-directional GRU, and "+" represents the improved Network with the proposed loss function L_{loss}.

It can be inferred from Table V that both bi-directional networks and L_{loss} improve the accuracy of GRU towards classification, and the proposed Bi-GRU+ achieved the best results in the most indicators. The high performance of the Bi-GRU+ on the moderately aging roads and shadow categories further proves the effectiveness of L_{loss}. Both the GRU+ and Bi-GRU+ models improve the classification accuracy with the moderately aged, which proves that L_{loss} can enhance the learning effect of the model for similar samples. The AA of the Bi-GRU+ model is slightly lower than that of Bi-GRU, which might be caused by the decrease of accuracy of the slightly aged. This actually demonstrated that the Bi-GRU+ model could perform better in the learning of similar samples of the moderately and heavily aged. In practical application, moderately aged and heavily aged are more noteworthy, thus we believe that a slight decrease in the accuracy of the slightly aged is acceptable.

The classification results show that the proposed Bi-GRU+ network is efficient, provides high-level automation and has higher accuracy than the existing methods. Thus, the proposed framework is more suitable for large-scale road pavement quality monitoring.

V. DISCUSSION

A. Applicability of Aerospace Remote Sensing

Dense vehicles will interfere to a certain extent with satellites and other aerospace remote sensing platforms on traffic busy roads. However, most roads do not always have heavy traffic, and high-resolution satellite imagery is usually acquired between 10:00 am and 2:00 pm, which is usually not as dense as in the morning and evening rush hours. In addition, the UAV remote sensing can be used as an auxiliary, and it is more convenient to choose the time window to obtain pure road pavement images. In future work, we aim at combining the advantages of UAV and satellite remote sensing.

On the other hand, according to the field investigation, secondary roads in suburban areas and rural roads are more likely to be ignored by road managers than high-grade roads. There are relatively few vehicles on secondary roads and rural roads, and satellites can sufficiently obtain pure road pavement images.

Therefore, the proposed method has applicability in pavement aging conditions mapping.

B. Practical Contribution

Previous studies on the monitoring of road pavement conditions from remote sensing images mainly focus on road damages such as cracks, while the pavement aged conditions are currently less assessed. Therefore, we proposed a novel framework for the highly automated monitoring of asphalt pavement aging conditions. This framework can use fine-resolution satellite remote sensing images as data sources and has the potential

for large-scale monitoring applications. With the popularization of high-resolution multi-spectral satellite remote sensing images, we believe that the proposed framework can become an important part of intelligent transportation.

C. Limitations

This study still has some limitations. Firstly, non-pavement objects such as shadows of trees on both sides of the road may change the spectral characteristics of the road pavement, masking the road conditions. A comprehensive analysis of multi-seasonal remote sensing images may solve this problem. In addition, due to the fact that the proposed method is a pixel-based classification method, the "salt and pepper" phenomenon occurs. One possible solution consists in using images with the higher spatial resolution to allow the use of spatial information. Another solution consists in combining CNN and RNN to develop spatial-spectral neural network classification models. Furthermore, the mixed pixels may also affect the classification result. Further studies necessary on combining deep learning and mixed pixel decomposition for monitoring aging road pavements are crucial.

VI. Conclusion

This study presented a Bi-GRU+ method for monitoring asphalt pavement aging conditions from fine-resolution satellite imagery. To the best of our knowledge, this is the first designed RNN-based extraction and classification method, to evaluate the asphalt pavement health conditions. This method uses a nonlinear function for spectral augmentation and the Bi-GRU+ with the improved loss function in order to generate a classification map of road pavement aging conditions. The experiments were performed using a Worldview-2 satellite image (16360 *7728) of the study area covering 16 roads in southwest Beijing, and the processing time is 7836 seconds in the experimental environment. Since pavement aging is a relatively long-term process in months or even years, the mapping accuracy is more important than the processing time. Compared with different state-of-the-art algorithms, the results show that the proposed method achieves better performance than the other methods for distinguishing pavement conditions with similar spectral characteristics. The Bi-GRU+ model achieves the best classification of the Worldview-2 images with OA of 98.16% and the Kappa coefficient of 0.97. The resultant pavement aging index for each road in the study area was also consistent with the field investigation.

The proposed framework demonstrates the efficiency of the RNN for road aging monitoring from satellite multispectral imagery. With the continuous advancement of spaceborne remote sensing technology, it is expected that it will become possible to obtain more inexpensive high-spatial resolution multispectral satellite imagery and allow the practical application of the framework.

In future work, more experiments will be performed, such as monitoring road aging conditions and road damages by combining UAV-acquired multispectral/hyperspectral images. The spectral mixture of image pixels will also be investigated.

Acknowledgment

The authors would like to thank the anonymous reviewers for their valuable comments and questions, which help improve this manuscript.

References

[1] C. W. Yi, Y. T. Chuang, and C. S. Nian, "Toward crowdsourcing-based road pavement monitoring by mobile sensing technologies," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 4, pp. 1905–1917, Aug. 2015, doi: 10.1109/TITS.2014.2378511.

[2] Y. Pan, X. Zhang, J. Tian, X. Jin, L. Luo, and K. Yang, "Mapping asphalt pavement aging and condition using multiple endmember spectral mixture analysis in Beijing, China," *J. Appl. Remote Sens.*, vol. 11, no. 1, 2017, Art. no. 016003.

[3] M. Herold and D. Roberts, "Spectral characteristics of asphalt road aging and deterioration: Implications for remote-sensing applications," *Appl. Opt.*, vol. 44, no. 20, pp. 4327–4334, Jul. 2005, doi: 10.1364/AO.44.004327.

[4] M. Herold *et al.*, "Imaging spectrometry and asphalt road surveys," *Transp. Res. C, Emerg. Technol.*, vol. 16, no. 2, pp. 153–166, Apr. 2008, doi: 10.1016/j.trc.2007.07.001.

[5] H. D. Cheng, J. Wang, Y. G. Hu, C. Glazier, X. J. Shi, and X. W. Chen, "Novel approach to pavement cracking detection based on neural network," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1764, no. 1, pp. 119–127, Jan. 2001, doi: 10.3141/1764-13.

[6] Y. Fei *et al.*, "Pixel-level cracking detection on 3D asphalt pavement images through deep-learning- based CrackNet-V," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 1, pp. 273–284, Jan. 2020, doi: 10.1109/TITS.2019.2891167.

[7] H. Oliveira and P. L. Correia, "Automatic road crack detection and characterization," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 1, pp. 155–168, Mar. 2013, doi: 10.1109/TITS.2012.2208630.

[8] Y. Shi, L. Cui, Z. Qi, F. Meng, and Z. Chen, "Automatic road crack detection using random structured forests," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 12, pp. 3434–3445, Dec. 2016, doi: 10.1109/TITS.2016.2552248.

[9] X. Zhang, X. Xia, N. Li, M. Lin, J. Song, and N. Ding, "Exploring the tricks for road damage detection with a one-stage detector," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Atlanta, GA, USA, Dec. 2020, pp. 5616–5621.

[10] M. da Mota Lopes, M. D. Zhao, E. Chailleux, M. Kane, T. Gabet, and C. Petiteau, "Characterization of aging processes on the asphalt mixture surface," in *Proc. 2nd Int. Symp. Asphalt Pavements Environ., Transp. Res. Board, Nat. Academies*, 2012, pp. 1–10.

[11] E. Schnebele, B. Tanyu, G. Cervone, and N. Waters, "Review of remote sensing methodologies for pavement management and assessment," *Eur. Transp. Res. Rev.*, vol. 7, no. 2, pp. 1–19, Mar. 2015, doi: 10.1007/S12544-015-0156-6.

[12] Y. U. Shah *et al.*, "Development of overall pavement condition index for urban road network," *Proc.-Social Behav. Sci.*, vol. 104, pp. 332–341, Dec. 2013, doi: 10.1016/j.sbspro.2013.11.126.

[13] M. Montoya-Alcaraz, A. Mungaray-Moctezuma, and L. García, "Sustainable road maintenance planning in developing countries based on pavement management systems case study in Baja California México," *Sustainability*, vol. 12, no. 1, p. 36, 2020, doi: 10.3390/su12010036.

[14] A. Cord and S. Chambon, "Automatic road defect detection by textural pattern recognition based on AdaBoost," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 27, no. 4, pp. 244–259, Sep. 2012, doi: 10.1111/j.1467-8667.2011.00736.x.

[15] K. Themistocleous *et al.*, "Damage assessment using advanced non-intrusive inspection methods: Integration of space, UAV, GPR, and field spectroscopy," in *Proc. 2nd Int. Conf. Remote Sens. Geoinf. Environ.*, Paphos, Cyprus, 2014, Art. no. 92291O.

[16] K. G. Fendi, S. M. Adam, N. Kokkas, and M. Smith, "An approach to produce a GIS database for road surface monitoring," *APCBEE Proc.*, vol. 9, pp. 235–240, Dec. 2014, doi: 10.1016/j.apcbee.2014.01.042.

[17] N. Nappo, O. Mavrouli, F. Nex, C. van Westen, R. Gambillara, and A. M. Michetti, "Use of UAV-based photogrammetry products for semi-automatic detection and classification of asphalt road damage in landslide-affected areas," *Eng. Geol.*, vol. 294, Dec. 2021, Art. no. 106363, doi: 10.1016/j.enggeo.2021.106363.

[18] D. H. Kang, S. Benipal, and Y.-J. Cha, "Hybrid concrete crack segmentation and quantification across complex backgrounds without a large training dataset," in *Data Science in Engineering*, Cham, Switzerland: Springer, 2022, pp. 123–128.

[19] Y. Pan, X. Zhang, G. Cervone, and L. Yang, "Detection of asphalt pavement potholes and cracks based on the unmanned aerial vehicle multispectral imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 10, pp. 3701–3712, Oct. 2018, doi: 10.1109/JSTARS.2018.2865528.

[20] A. Mei, R. Salvatori, N. Fiore, A. Allegrini, and A. D'Andrea, "Integration of field and laboratory spectral data with multi-resolution remote sensed imagery for asphalt surface differentiation," *Remote Sens.*, vol. 6, no. 4, pp. 2765–2781, Mar. 2014, doi: 10.3390/rs6042765.

[21] R. B. Gomez, "Hyperspectral imaging: A useful technology for transportation analysis," *Opt. Eng.*, vol. 41, no. 9, pp. 2137–2143, Sep. 2002, doi: 10.1117/1.1497985.

[22] M. R. Resende, L. L. B. Bernucci, and J. A. Quintanilha, "Monitoring the condition of roads pavement surfaces: Proposal of methodology using hyperspectral images," *J. Transp. Literature*, vol. 8, no. 2, pp. 201–220, Apr. 2014.

[23] C. Mettas *et al.*, "Risk provision using field spectroscopy to identify spectral regions for the detection of defects in flexible pavements," *Natural Hazards*, vol. 83, no. 1, pp. 83–96, 2016.

[24] C. Mettas, K. Themistocleous, K. Neocleous, A. Christofe, K. Pilakoutas, and D. Hadjimitsis, "Monitoring asphalt pavement damages using remote sensing techniques," in *Proc. 3rd Int. Conf. Remote Sens. Geoinf. Environ.*, 2015, pp. 237–248.

[25] W. Emery, A. Yerasi, N. Longbotham, and F. Pacifici, "Assessing paved road surface condition with high resolution satellite imagery," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Quebec, QC, Canada, Jul. 2014, pp. 13–18.

[26] M. Mohammadi, "Road classification and condition determination using hyperspectral imagery," *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. 39, p. B7, Jan. 2012.

[27] C. Andreou, V. Karathanassi, and P. Kolokoussis, "Investigation of hyperspectral remote sensing for mapping asphalt road conditions," *Int. J. Remote Sens.*, vol. 32, no. 21, pp. 6315–6333, 2011.

[28] N. Carmon and E. Ben-Dor, "Mapping asphaltic roads' skid resistance using imaging spectroscopy," *Remote Sens.*, vol. 10, no. 3, p. 430, Jan. 2018, doi: 10.1117/1.JRS.11.016003.

[29] Z. Liu, Y. Cao, Y. Wang, and W. Wang, "Computer vision-based concrete crack detection using U-Net fully convolutional networks," *Autom. Construct.*, vol. 104, pp. 129–139, Dec. 2019, doi: 10.1016/j.autcon.2019.04.005.

[30] L. Pauly, D. Hogg, R. Fuentes, and H. Peel, "Deeper networks for pavement crack detection," in *Proc. 34th ISARC*, Taiwan, 2017, pp. 479–485.

[31] K. Gopalakrishnan, S. K. Khaitan, A. Choudhary, and A. Agrawal, "Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection," *Construct. Building Mater.*, vol. 157, pp. 322–330, Dec. 2017, doi: 10.1016/j.conbuildmat.2017.09.110.

[32] V. Pereira, S. Tamura, S. Hayamizu, and H. Fukai, "A deep learning-based approach for road pothole detection in Timor Leste," in *Proc. IEEE Int. Conf. Service Oper. Logistics, Informat. (SOLI)*, Aug. 2018, pp. 279–284.

[33] Y. Pan, X. Chen, Q. Sun, and X. Zhang, "Monitoring asphalt pavement aging and damage conditions from low-altitude UAV imagery based on a CNN approach," *Can. J. Remote Sens.*, vol. 47, no. 3, pp. 1–17, Jan. 2021, doi: 10.1080/07038992.2020.1870217.

[34] D. Arya *et al.*, "Transfer learning-based road damage detection for multiple countries," 2020, *arXiv:2008.13101*.

[35] Y. J. Cha, W. Choi, and O. Büyüköztürk, "Deep learning-based crack damage detection using convolutional neural networks," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 32, no. 5, pp. 361–378, Mar. 2017, doi: 10.1111/mice.12263.

[36] J. Li, X. Zhao, and H. Li, "Method for detecting road pavement damage based on deep learning," *Proc. SPIE*, vol. 10972, Apr. 2019, Art. no. 109722D.

[37] V. Hegde, D. Trivedi, A. Alfarrarjeh, A. Deepak, S. H. Kim, and C. Shahabi, "Yet another deep learning approach for road damage detection using ensemble learning," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Atlanta, GA, USA, Dec. 2020, pp. 5553–5558.

[38] H. Maeda, Y. Sekimoto, T. Seto, T. Kashiyama, and H. Omata, "Road damage detection and classification using deep neural networks with smartphone images," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 33, no. 12, pp. 1127–1141, Jun. 2018, doi: 10.1111/mice.12387.

[39] Y. J. Cha *et al.*, "Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 33, no. 9, pp. 731–747, Nov. 2018, doi: 10.1111/mice.12334.

[40] T. Hascoet, Y. Zhang, A. Persch, R. Takashima, T. Takiguchi, and Y. Ariki, "FasterRCNN monitoring of road damages: Competition and deployment," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Atlanta, GA, USA, Dec. 2020, pp. 5545–5552.

[41] Z. Pei, R. Lin, X. Zhang, H. Shen, J. Tang, and Y. Yang, "CFM: A consistency filtering mechanism for road damage detection," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Atlanta, GA, USA, Dec. 2020, pp. 5584–5591.

[42] L. N. H. Truong, O. E. Mora, W. Cheng, H. Tang, and M. Singh, "Deep learning to detect road distress from unmanned aerial system imagery," *Transp. Res. Rec.*, vol. 2675, no. 9, Apr. 2021, Art. no. 03611981211004973, doi: 10.1177/03611981211004973.

[43] F. Yang, L. Zhang, S. Yu, D. Prokhorov, X. Mei, and H. Ling, "Feature pyramid and hierarchical boosting network for pavement crack detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 4, pp. 1525–1535, Apr. 2020, doi: 10.1109/TITS.2019.2910595.

[44] A. Zhang *et al.*, "Automated pixel-level pavement crack detection on 3D asphalt surfaces using a deep-learning network," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 32, no. 10, pp. 805–819, Aug. 2019, doi: 10.1111/mice.12297.

[45] J. Huyan, W. Li, S. Tighe, Z. Xu, and J. Zhai, "CrackU-Net: A novel deep convolutional neural network for pixelwise pavement crack detection," *Struct. Control Health Monitor.*, vol. 27, no. 8, p. e2551, Mar. 2020, doi: 10.1002/stc.2551.

[46] X. Yang, H. Li, Y. Yu, X. Luo, T. Huang, and X. Yang, "Automatic pixel-level crack detection and measurement using fully convolutional network," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 33, no. 12, pp. 1090–1109, Aug. 2018, doi: 10.1111/mice.12412.

[47] D. H. Kang and Y.-J. Cha, "Efficient attention-based deep encoder and decoder for automatic crack segmentation," *Struct. Health Monitor.*, vol. 21, no. 5, 2021, Art. no. 14759217211053776.

[48] D. Kang *et al.*, "Hybrid pixel-level concrete crack segmentation and quantification across complex backgrounds using deep learning," *Autom. Construct.*, vol. 118, Oct. 2020, Art. no. 103291.

[49] W. Choi and Y.-J. Cha, "SDDNet: Real-time crack segmentation," *IEEE Trans. Ind. Electron.*, vol. 67, no. 9, pp. 8016–8025, Sep. 2020.

[50] A. Mei, C. Manzo, C. Bassani, R. Salvatori, and A. Allegrini, "Bitumen removal determination on asphalt pavement using digital imaging processing and spectral analysis," *Open J. Appl. Sci.*, vol. 2014, pp. 1–9, May 2014, doi: 10.4236/ojapps.2014.46034.

[51] M. Sundermeyer, H. Ney, and R. Schlüter, "From feedforward to recurrent LSTM neural networks for language modeling," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 23, no. 3, pp. 517–529, Mar. 2015.

[52] R. Hang, Q. Liu, D. Hong, and P. Ghamisi, "Cascaded recurrent neural networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5384–5394, Aug. 2019, doi: 10.1109/TGRS.2019.2899129.

[53] K. Cho, B. van Merrienboer, D. Bahdanau, and Y. Bengio, "On the properties of neural machine translation: Encoder–decoder approaches," 2014, *arXiv:1409.1259*.

[54] A. Paszke *et al.*, "PyTorch: An imperative style, high-performance deep learning library," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, pp. 8024–8035.

[55] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, no. 10, pp. 2825–2830, Jul. 2017.

[56] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[57] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[58] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2017, pp. 1492–1500.

[59] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.

Xiao Chen received the B.S. degree in geographic information science from China Agricultural University. He is currently pursuing the Ph.D. degree with the Institute of Remote Sensing and Geographic Information System, Peking University, China. His main research interests include remote sensing for disaster management, deep learning, and remotely sensed image classification.

Xianfeng Zhang received the Ph.D. degree from the University of Western Ontario, London, Canada, in 2005. He is currently a Full Professor and the Deputy Director of the Institute of Remote Sensing and GIS, Peking University. His research interests include remote sensing for natural hazard management, deep learning, remote sensing big data, and remote sensing for transportation. He is an Associate Editor of *International Journal of Applied Earth Observation and Geoinformation*. ORCID ID: 0000-0001-9486-82.

Jonathan Li (Senior Member, IEEE) is currently a Full Professor with the Department of Geography and Environmental Management, University of Waterloo, Canada. His research interests are in urban remote sensing and geospatial data science, especially in the intelligent extraction of geometric and semantic information for earth observation images, and LiDAR point clouds using machine learning and deep learning algorithms.

Miao Ren is currently pursuing the Ph.D. degree with the Institute of Remote Sensing and Geographic Information System, Peking University. His main research interests include pavement condition measurement, deep learning, and remote sensing for transportation.

Bo Zhou received the B.Eng. degree in science and technology of remote sensing from Wuhan University in 2021. He is currently pursuing the M.S. degree with the Institute of Remote Sensing and Geographic Information System, Peking University. His research interests include hyperspectral image pan-sharpening, deep learning, and remotely sensed image segmentation.