

Research



Cite this article: Pan S, Yan H, Liu Z, Chen N, Miao Y, Hou Y. 2023 Automatic pavement texture recognition using lightweight few-shot learning. *Phil. Trans. R. Soc. A* **381**: 20220166.
<https://doi.org/10.1098/rsta.2022.0166>

Received: 26 February 2023

Accepted: 8 May 2023

One contribution of 15 to a theme issue 'Artificial intelligence in failure analysis of transportation infrastructure and materials'.

Subject Areas:

civil engineering

Keywords:

pavement detection, deep learning, convolutional neural network, Siamese network, one-dimensional convolution

Authors for correspondence:

Zhuo Liu

e-mail: liuzhuo66@bjut.edu.cn

Yue Hou

e-mail: yue.hou@swansea.ac.uk

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.6725581>.

Automatic pavement texture recognition using lightweight few-shot learning

Shuo Pan¹, Hai Yan¹, Zhuo Liu¹, Ning Chen¹,
Yinghao Miao² and Yue Hou³

¹Beijing Key Laboratory of Traffic Engineering, Beijing University of Technology, Beijing 100124, People's Republic of China

²National Center for Materials Service Safety, University of Science and Technology Beijing, Beijing 100083, People's Republic of China

³Department of Civil Engineering, Faculty of Science and Engineering, Swansea University, Swansea SA1 8EN, UK

SP, 0000-0001-9452-9738; HY, 0000-0003-0660-7228; ZL, 0000-0001-9356-8989; YH, 0000-0002-4334-2620

Texture is a crucial characteristic of roads, closely related to their performance. The recognition of pavement texture is of great significance for road maintenance professionals to detect potential safety hazards and carry out necessary countermeasures. Although deep learning models have been applied for recognition, the scarcity of data has always been a limitation. To address this issue, this paper proposes a few-shot learning model based on the Siamese network for pavement texture recognition with a limited dataset. The model achieved 89.8% accuracy in a four-way five-shot task classifying the pavement textures of dense asphalt concrete, micro surface, open-graded friction course and stone matrix asphalt. To align with engineering practice, global average pooling (GAP) and one-dimensional convolution are implemented, creating lightweight models that save storage and training time. Comparative experiments show that the lightweight model with GAP implemented on dense layers and one-dimensional convolution on convolutional layers reduced storage volume by 94% and training time by 99%, despite a 2.9% decrease in classification accuracy. Moreover, the model with only GAP

implemented on dense layers achieved the highest accuracy at 93.5%, while reducing storage volume and training time by 83% and 6%, respectively.

This article is part of the theme issue 'Artificial intelligence in failure analysis of transportation infrastructure and materials'.

1. Introduction

Pavement texture is a vital road characteristic with notable effects on skid resistance [1,2], tire-pavement noise [3,4] and road maintenance [5,6]. Changes in pavement texture during road use may lead to potential traffic safety risks and affect the quality of road service. Thus, accurate recognition of pavement texture is an essential prerequisite for performing road maintenance. However, the complexity of factors affecting the appearance of pavement texture may pose challenges to identifying pavement texture types promptly and accurately in many situations. As a result, large-scale pavement texture recognition tasks typically require substantial human and financial resources.

Deep learning has gained remarkable attention due to its ability to extract target features and learn iteratively using convolutional neural networks (CNNs) in various fields, including autonomous driving [7], materials [8], medicine [9] and road detection [10]. Specifically, for pavement texture recognition, deep learning has shown excellent performance. Chen *et al.* [11] applied the Dense Convolutional Network to classify pavement texture by generating adversarial networks to expand pavement texture, demonstrating the superiority of deep learning over manual classification and traditional machine learning methods. Liu *et al.* [12] proposed a precise and stable framework for pavement texture measurement and reconstruction, using a depth CNN-based encoder to extract features from pavement images and converting them into models using feature mapping units and decoders. Dong *et al.* [13] proposed an effective encoder–decoder CNN architecture based on the residual network for the reconstruction of pavement texture, which proved to be effective for three-dimensional feature analysis. Yoo *et al.* [14] introduced a novel method of pavement classification based on deep learning by using the tire-pavement interaction noise, which reflects the surface contour and texture properties of the road, as the data source. Chhay *et al.* [15] proposed a faster region-based CNN based on deep learning to count the exposed aggregate number representing texture wavelengths on digital images of exposed aggregate concrete pavement.

Although deep learning methods have demonstrated impressive classification accuracy in various tasks, the requirement for substantial amounts of training data often poses a challenge in real-world engineering applications. Few-shot learning has emerged as a promising alternative to overcome this obstacle, as it enables effective classification tasks with limited data availability. Several network architectures, such as Siamese networks [16,17], matching networks [18], prototype networks [19] and relation networks [20], have been designed and implemented specifically for few-shot learning applications. Numerous studies have demonstrated satisfactory results in classification tasks using few-shot learning techniques, even with fewer than 10 samples available for training each classification category. This highlights the potential of few-shot learning as a viable solution for tackling data scarcity issues in various domains. For example, Wang *et al.* [21] proposed a multi-attention mutual information distributed framework that combines distributed learning with few-shot learning to reduce time consumption and achieve desired results in the evaluation. Hu *et al.* [22] integrated unsupervised descriptor selection into few-shot learning, which employs a descriptor selection module to select and localize semantically meaningful regions in feature maps without supervision, thereby improving the efficiency of task adaptation and classification performance. Singh *et al.* [23] introduced dual class information, base class label and self-supervised class label for the base class training images to help the model learn generic features and improve the novel class performance in few-shot learning. Zeng *et al.* [24] developed a self-attention and mutual-attention module to learn feature

correlations and reduce background interference in few-shot learning classification of remote-sensing images, which led to higher accuracy. Dong *et al.* [25] used the back-end network to extract multilevel feature information from the base class and improved the loss function to maximize the distance between different categories and minimize the distance between the same categories in few-shot learning. Guo *et al.* [26] introduced the pairwise similarity module into few-shot learning and generated the calibrated class centres suitable for the query sample by extracting the semantic correlations between the support and query sample and enhancing the discriminant regions. It is clear that few-shot learning has great potential to address pavement texture recognition issues.

Road maintenance professionals tend to use mobile devices in the field due to their portability, which limits the storage capacity of embedded models. Additionally, rapid feedback of classification results is often needed in engineering practice. Therefore, it is important to develop lightweight few-shot learning models applicable to practical engineering scenarios. Reducing the time and computational cost of deep learning models can bridge the gap between theoretical models and practical applications in many fields, and a number of lightweight methods have been proposed to this end. Sun *et al.* [27] proposed a lightweight network architecture for image super-resolution, using depthwise separable convolution to enhance the efficiency of the convolution operation and stacked weighted multi-scale residual blocks to improve the representation capability. Jiang *et al.* [28] developed a method for network compression in super-resolution, which employed weight pruning and a multi-slicing network of information to extract and integrate multi-scale features. Grassucci *et al.* [29] introduced a family of parameterized hypercomplex neural networks that directly captured convolution rules and filter organization from data, making them flexible in any user-defined or tuned domain. Cheng *et al.* [30] proposed a lightweight unified fusion network for multi-focus image fusion, which used guided filtering to separate the source image into basic and detailed layers and applied a gradient perception strategy to handle the fusion problem. Liu *et al.* [31] developed a lightweight generative adversarial network that used squeeze and expand, multi-scale convolution, and depthwise separable convolution to reduce network parameters and training time. Gong *et al.* [32] introduced a global contextually guided lightweight network that employed secondary cross-modal integration to remove redundant information and added a hybrid feature-cascaded aggregation module to emphasize the global context, along with calibration. However, despite the potential of few-shot learning for pavement texture classification, the development of lightweight few-shot learning models specifically for this application has not been extensively investigated, presenting an important research gap.

This study proposes a pavement texture classification model based on few-shot learning, aiming to improve its suitability for engineering practice by making it lightweight. The dataset used was constructed with three-dimensional laser scanning pavement texture data. The Siamese network is chosen as the basic framework to construct the original few-shot learning model, considering the specific characteristics of the pavement texture data. To reduce the parameter quantity and multiply-accumulate operations, lightweight methods such as global average pooling (GAP) and one-dimensional convolution are employed in different parts of the few-shot learning model. Comparative experiments were conducted to evaluate the performances of the lightweight models and the original few-shot learning model. The applicability of the proposed models in actual engineering tasks was then discussed based on the results. This study contributes twofold: (i) verifying the feasibility of few-shot learning on four-way five-shot pavement texture classification and (ii) proposing specific lightweight pavement texture classification models, balancing accuracy with storage volume and training time. By focusing on the practical requirements of pavement texture recognition in the field of road maintenance and addressing the limitations of current deep learning models, this study aims to provide valuable insights for practitioners and researchers alike. The proposed lightweight few-shot learning models have the potential to significantly impact the efficiency and effectiveness of pavement texture recognition, ultimately contributing to safer and more sustainable road infrastructure.

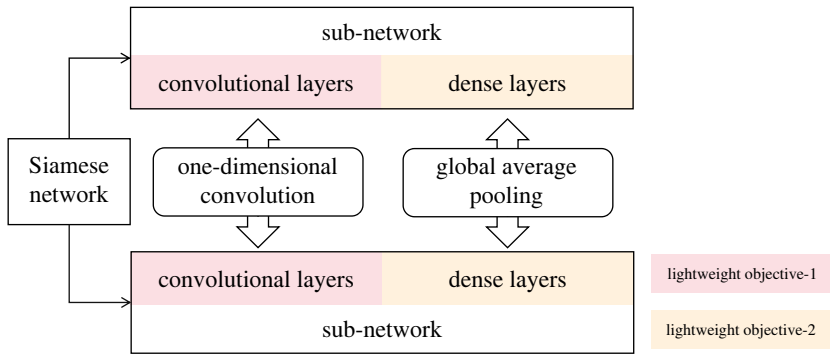


Figure 1. The flow chart of methodology.

2. Methodology

This section presents the methodology employed in this paper for pavement texture recognition. First, the Siamese network is introduced, which serves as the few-shot learning network framework, highlighting the potential targets to lightweight in the Siamese network. Subsequently, alternative lightweight methods are discussed, including GAP, one-dimensional convolution and the integration of both. The figure 1 presents the basic structure and the lightweight routes of the Siamese network.

(a) Siamese network architecture and lightweight objectives

In this study, the Siamese network was chosen for pavement texture classification, primarily due to its exceptional performance across various domains and its relatively simple architecture compared to other few-shot learning frameworks. The simplicity of the Siamese network makes it an appealing choice for addressing complex classification tasks while maintaining efficiency and ease of implementation. Furthermore, its demonstrated success in diverse applications [33–35] offers a strong foundation for adapting it to the specific requirements of pavement texture classification, ultimately contributing to the creation of more accurate and robust models. Although the Siamese network may entail higher computational demands when dealing with a larger number of categories, it remains applicable and well suited for the four-type pavement texture classification task addressed in this paper.

The Siamese network consists of two sub-networks with the same structure that share the same weights. These sub-networks generate embeddings for the input data, which are then fed to the energy function. The similarity of the input data is calculated based on the output of the energy function. For image classification, the Visual Geometry Group 16 (VGG16) was chosen as the sub-networks, which has been proven effective for extracting feature vectors in many studies based on Siamese networks [36–38].

Specifically, a pair of pavement texture images is input into the original Siamese network, with each image being processed by a separate sub-network. The sub-networks consist of a convolution layer, a maximum pooling layer and a dense connection layer, as shown in figure 2. The convolutional layers have a kernel size of 3×3 and channels of 64, 128, 256, 512 and 512, with the Relu activation function connecting each convolutional layer. The maximum pooling layers have kernels with a 2×2 size and strides with a 2×2 size. The first two dense connection layers have 4096 output nodes connected to the Relu activation function, while the last dense connection layer has 1000 output nodes connected to the Sigmoid activation function. The feature vectors output by the sub-networks are then used to calculate the energy function, which ultimately produces the similarity between the input images.

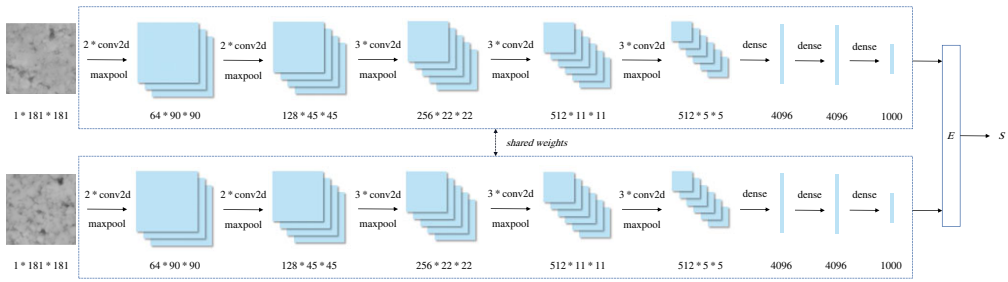


Figure 2. Original Siamese network.

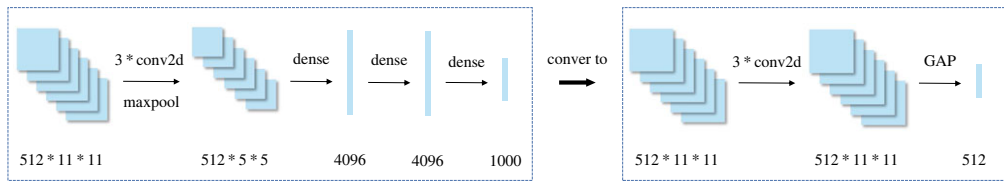


Figure 3. Lightweight for dense layers based on GAP.

Regarding the lightweighting of the original Siamese network, the main targets are reducing storage volume and training time. The reduction of the number of parameters is the primary consideration for reducing the storage volume occupied by the classification model, which can also decrease the likelihood of overfitting. In terms of training time, the reduction of multiply-accumulate operations (MACC) in the network, a significant portion of computational overhead, is crucial. Table 1 shows the volume of parameters and MACC for the sub-networks in the original Siamese network. It is evident that the dense layers contain the majority of the parameters in the network, while the convolutional layers are the primary source of computational cost. Consequently, the lightweighting of the dense layers is crucial for reducing the storage volume of the network, and the lightweighting of the convolutional layers is essential for saving the training time.

(b) Lightweight method based on global average pooling

GAP is a lightweight method used to reduce the number of parameters in dense layers [39]. It involves taking the average of all the pixels in the feature map of each channel, resulting in a single value per channel. The feature map is then transformed into a feature vector with a length equal to the number of channels. This method is useful in avoiding the high input dimensions that are usually required by dense layers. In this study, GAP is applied in place of dense layers, as illustrated in figure 3. This approach helps to reduce the number of parameters, storage volume and potential for overfitting.

(c) Lightweight method based on one-dimensional convolution

One-dimensional CNNs are commonly used for processing text data [40]. However, given the simple characteristics of pavement texture, it is possible to perform pavement texture classification using one-dimensional convolution. Unlike two-dimensional convolution, one-dimensional convolution operates on feature vectors with only one dimension, and its convolution kernels also have only one dimension, significantly reducing the computational

Table 1. Parameters and MACC of the sub-networks of the original Siamese network.

no.	layer	output shape	parameters ($\times 10^6$)	MACC ($\times 10^9$)
1	conv2D	(64, 181, 181)	0.00064	0.019
2	conv2D	(64, 181, 181)	0.037	1.2
3	pooling	(64, 90, 90)	0	/
4	conv2D	(128, 90, 90)	0.074	0.60
5	conv2D	(128, 90, 90)	0.15	1.2
6	pooling	(128, 45, 45)	0	/
7	conv2D	(256, 45, 45)	0.30	0.60
8	conv2D	(256, 45, 45)	0.59	1.2
9	conv2D	(256, 45, 45)	0.59	1.2
10	pooling	(256, 22, 22)	0	/
11	conv2D	(512, 22, 22)	1.2	0.57
12	conv2D	(512, 22, 22)	2.4	1.1
13	conv2D	(512, 22, 22)	2.4	1.1
14	pooling	(512, 11, 11)	0	/
15	conv2D	(512, 11, 11)	2.4	0.29
16	conv2D	(512, 11, 11)	2.4	0.29
17	conv2D	(512, 11, 11)	2.4	0.29
18	pooling	(512, 5, 5)	0	/
19	flatten	(12 800)	0	/
20	dense	(4096)	52	0.052
21	dense	(4096)	17	0.017
22	dense	(1000)	4.1	0.00041
total parameters: $88 (\times 10^6)$.				
total MACC: $9.8 (\times 10^9)$				

overhead and time cost. To accommodate one-dimensional convolution, the input of the sub-network is a 181-length vector with 181 channels, containing the same information as the original 181×181 single-channel image. Consequently, all two-dimensional convolution operations are replaced by one-dimensional convolution, and the pooling layers are updated to one-dimensional pooling, as illustrated in figure 4.

(d) Lightweight method based on the integration of global average pooling and one-dimensional convolution

The combination of GAP and one-dimensional convolution can yield an optimal lightweight effect. In this approach, the feature vector output of the last one-dimensional convolution layer is subjected to a GAP operation, where each dimension of the feature vector is averaged to obtain a scalar value. The resulting feature vector, which has a length equal to the number of channels of feature vectors generated in the last one-dimensional convolution layer, serves as the embedding that is fed into the energy function by the sub-network. The lightweight Siamese

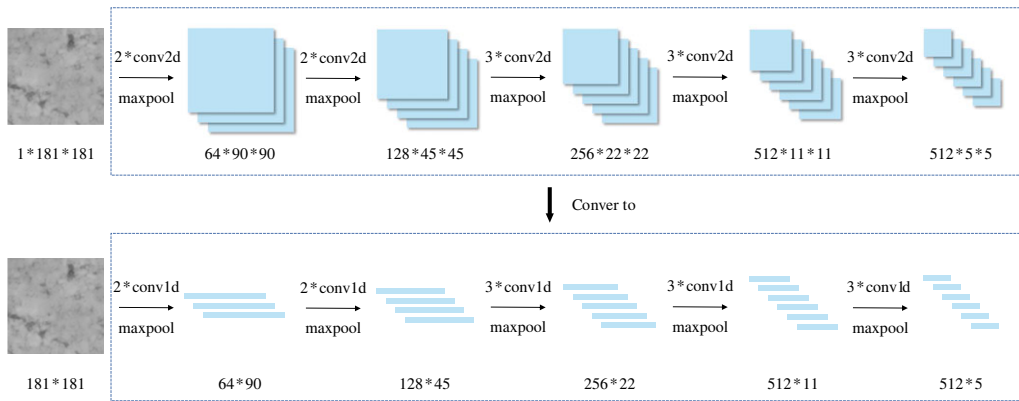


Figure 4. Lightweight for convolutional layers based on one-dimensional convolution.

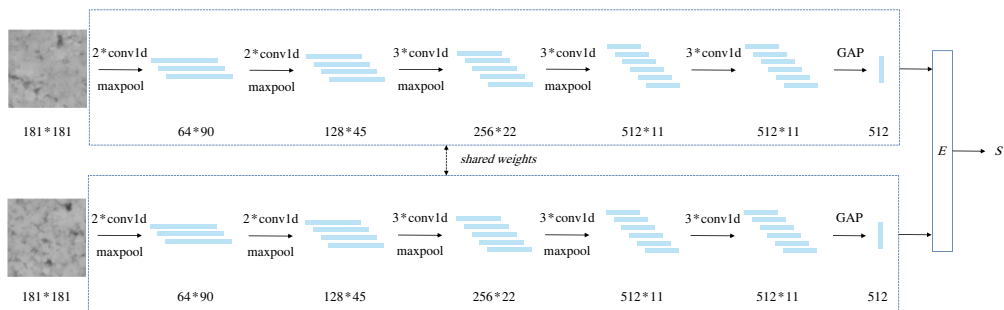


Figure 5. Lightweight Siamese network based on GAP and one-dimensional convolution.

network structure that incorporates both GAP and one-dimensional convolution is illustrated in figure 5.

3. Experiment design

This section introduces the experimental design. First, it presents the pavement texture dataset used in this study. Subsequently, the design of the comparative experiments is discussed, including the specific network configurations and hyperparameters in the experiments designed to test the performance of the recognition models employing different lightweight methods. Finally, the experimental environment is described. The flowchart of the experiment is presented in figure 6.

(a) Dataset of pavement texture

Three-dimensional pavement texture data were acquired using a laser scanner in Beijing, China, with a size of 90×90 mm and sampling intervals of 0.5 mm in two horizontal directions [41]. Each data was recorded as a 181×181 two-dimensional array with the corresponding height value. The height of the pavement texture was normalized prior to further processing, and the data were converted into grayscale images suitable for the application of image classification methods and visualization.

The classification task involves identifying four of the most common types of pavement texture: dense asphalt concrete (DAC), micro surface (MS), open-graded friction course (OGFC) and stone matrix asphalt (SMA). Typical images of each type of pavement texture refer to [11]. In

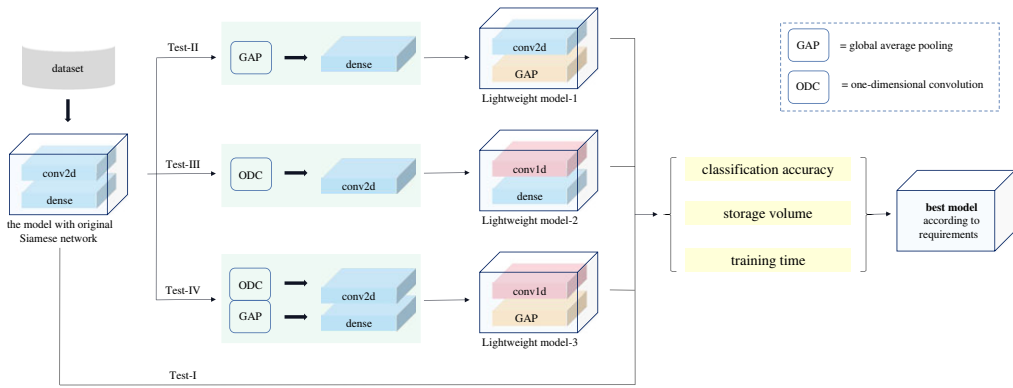


Figure 6. The flowchart of experiment design.

Table 2. Sample composition of the dataset.

	DAC	MS	OGFC	SMA
sample size of training set	5	5	5	5
sample size of test set	45	4	14	45

this study, the training of the pavement texture recognition model was performed with a limited dataset of only five images per category. This constraint was motivated by the fact that collecting diverse and extensive samples of pavement texture images is a challenging and laborious process in road detection work. The main objective of this study was to propose a reliable model capable of achieving high accuracy in recognizing pavement textures despite the limited availability of training data, ultimately contributing to the overall efficiency and effectiveness of road surface analysis and maintenance. The number of images used for model training is presented in table 2. The proposed few-shot learning models were evaluated using a four-way five-shot classification task, which is typical in the field of few-shot learning [42–45].

(b) Comparative experiments for evaluating the performances of the classification models

To reduce the storage volume and training time of the Siamese network and investigate how this affects the accuracy of the network in classifying pavement texture, a comparative experiment was designed. Various methods were applied to obtain lightweight models from the original Siamese network, which were then trained under the same conditions. The performance of the models with the original Siamese network and lightweight Siamese networks for pavement texture classification was compared comprehensively, considering their storage volume, training time and classification accuracy.

Specifically, four experiments were designed to evaluate the performance of the original Siamese network model and three lightweight models, as shown in table 3. The original Siamese network consisted of convolutional layers and dense layers, while the lightweight models were obtained by applying three different lightweight methods to the original Siamese network. Lightweight model-1 was obtained by replacing the dense layers of the original Siamese network with GAP. Lightweight model-2 used one-dimensional convolution for the convolutional layers in the original Siamese network. Lightweight model-3 applied GAP and one-dimensional convolution to the dense layers and the convolutional layers, respectively.

Prior to the training, sample pairs were constructed by pairing two pavement texture images and inputting them into the Siamese network. A sample pair consisting of two images belonging to the same category was defined as a positive sample pair, while a sample pair consisting of two

Table 3. Experiments using different lightweight methods of Siamese network.

Experiment	model	lightweight method for dense layers	lightweight method for convolutional layers
Test-I	model with original Siamese network	/	/
Test-II	Lightweight model-1	GAP	/
Test-III	Lightweight model-2	/	one-dimensional convolution
Test-IV	Lightweight model-3	GAP	one-dimensional convolution

Table 4. Sample pairs for each experiment.

experiment	model	sample pairs		sample pairs partition	
		positive sample pairs	negative sample pairs	sample pairs of training (%)	sample pairs of validation (%)
Test-I	model with original Siamese network	40	150	50	50
Test-II	Lightweight model-1	40	150	50	50
Test-III	Lightweight model-2	40	150	50	50
Test-IV	Lightweight model-3	40	150	50	50

images belonging to different categories was defined as a negative sample pair. All images in the training set were used to construct sample pairs. The number of positive and negative sample pairs was calculated using equations (1) and (2), respectively.

$$NUM_p = N \cdot \frac{K!}{2(K-2)!} \quad (1)$$

$$NUM_n = K^2 \cdot \frac{N!}{2(N-2)!} \quad (2)$$

where N is the number of categories and K is the sample size of each category. In the pavement texture classification task of this study, the numbers of positive and negative sample pairs were 40 and 150, respectively. All the sample pairs were randomly divided into two equal parts for training and validation of the Siamese network. The statistics of sample pairs used for each experiment are presented in table 4.

The Siamese networks used the Euclidean distance function as the energy function, which was defined as follows:

$$E = \|x_n - z_n\|_2, \quad (3)$$

where x and z represent the feature vectors output by the sub-networks and n is the dimension of the feature vector. The Energy function serves as a measure of the difference between samples. In particular, for a positive sample pair, the value of E will be very small, while for a negative sample pair, the value of E will be large.

The contrastive loss function was used to evaluate the match between E and the corresponding label, which can be defined as follows:

$$\text{Loss} = \frac{1}{2N_s} \sum_{i=1}^{N_s} yE^2 + (1-y) \max(m-E, 0)^2, \quad (4)$$

where y represents the label of sample pairs, 1 for positive sample pairs and 0 for negative sample pairs, N_s is the number of samples, and m is a threshold set to 1 in this paper. If the Siamese network is effective, the value of E for positive sample pairs tends to approach 0, while the value of E for negative sample pairs is significantly larger, resulting in a low value of *Loss*. Conversely,

Table 5. Comparison of classification accuracies.

experiment	model	classification accuracy on test set (%)				
		DAC	MS	OGFC	SMA	average
Test-I	model with original Siamese network	95.6	100	80.0	83.6	89.8
Test-II	Lightweight model-1	96.9	100	91.4	85.8	93.5
Test-III	Lightweight model-2	91.1	100	68.6	69.3	82.3
Test-IV	Lightweight model-3	96.4	100	78.6	72.4	86.9

if the value of E for positive sample pairs is large and that for negative sample pairs is close to 0, the $Loss$ value will be high.

The RMSprop optimizer was used with a learning rate of 10^{-5} , and the training epochs were set to 20 with a batch size of 16. The other parameters were set to their default values; each experiment was repeated five times under the same conditions.

(c) Experimental environment and dependent equipment

The experiments were performed on a laptop with an Intel Core i9-9880H CPU@2.30GHz, 16 cores and 32GB RAM. All models were implemented in Python using the TensorFlow deep learning library.

4. Result and discussion

(a) Comparison of pavement texture classification accuracy among the lightweight models

The Siamese network can only infer similarity between two input images and cannot directly classify them into specific categories. Therefore, the comparison verification method was used to classify pavement texture images in the test set. Specifically, for each image in the test set, sample pairs were constructed using five images of each category in the training set and input into the trained Siamese networks. The energy function values of the test image and each category of images in the training set were averaged to obtain the differences between the test image and various pavement textures. The category with the least difference from the test image was considered as the classification result. The classification confusion matrix of each experiment is presented in figure 7. The high classification accuracy for each category of pavement texture in each experiment indicates that the original Siamese network and each lightweight Siamese network were effective.

The comparison of classification accuracies is presented in table 5, where the average accuracy was defined as the arithmetic average of accuracy in each category. It was observed that Lightweight model-1 achieved the highest average accuracy, which could be attributed to the GAP method that helped alleviate the overfitting problem during the training process. On the other hand, the average accuracies of Lightweight model-2 and Lightweight model-3 were lower than that of the Original Siamese network, indicating that one-dimensional convolution was less effective for classification than two-dimensional convolution. Nevertheless, all models proposed in this study achieved an average accuracy higher than 80%, thus indicating their effectiveness for pavement texture classification. Notably, the MS classification accuracy was 100% due to its significant differences from other categories, while the OGFC and SMA classification accuracies were low as they were similar to each other.

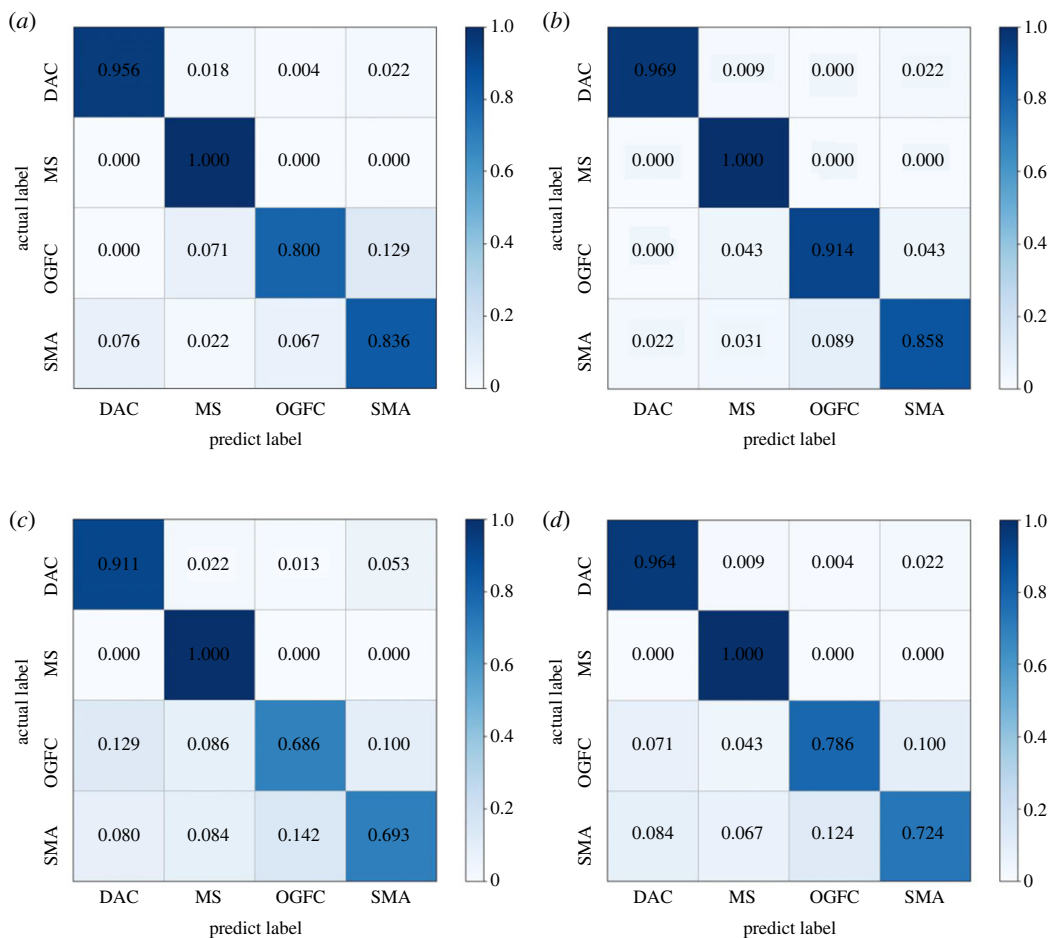


Figure 7. Classification confusion matrix of each experiment. (a) Test-I, (b) Test-II, (c) Test-III and (d) Test-IV.

(b) Comparison of lightweight degree among the lightweight models

The degree of lightweight of a network is a critical factor in practical engineering. As the major indicator of network lightness, storage volume was compressed by reducing parameters and training time was saved by reducing MACC in this study. The degree of lightness of the proposed models is summarized in table 6. Compared to the original Siamese network-based model, the storage volume of Lightweight model-1 and Lightweight model-3 was substantially reduced due to GAP, which decreased dense layer parameters. Lightweight model-2 and Lightweight model-3 were superior in terms of training time, as one-dimensional convolution actively reduced the MACC of the convolutional layer. Overall, Lightweight model-3 achieved the most desirable lightweight effect, with the lowest storage volume and shortest training time.

Before proceeding further, it is important to note that the aforementioned data is reliable. The parameters quantity and MACC are considered to be theoretical values, whereas storage volume, training time and accuracy are experimental values that require rational analysis. Based on the results obtained after each experiment was repeated five times, it was observed that the storage volume remained constant, and the accuracy and training time remained mostly stable, as depicted in figure 8. Furthermore, calculations revealed that the accuracy and training time for each experiment fell within twice the s.d. range with no outliers detected.

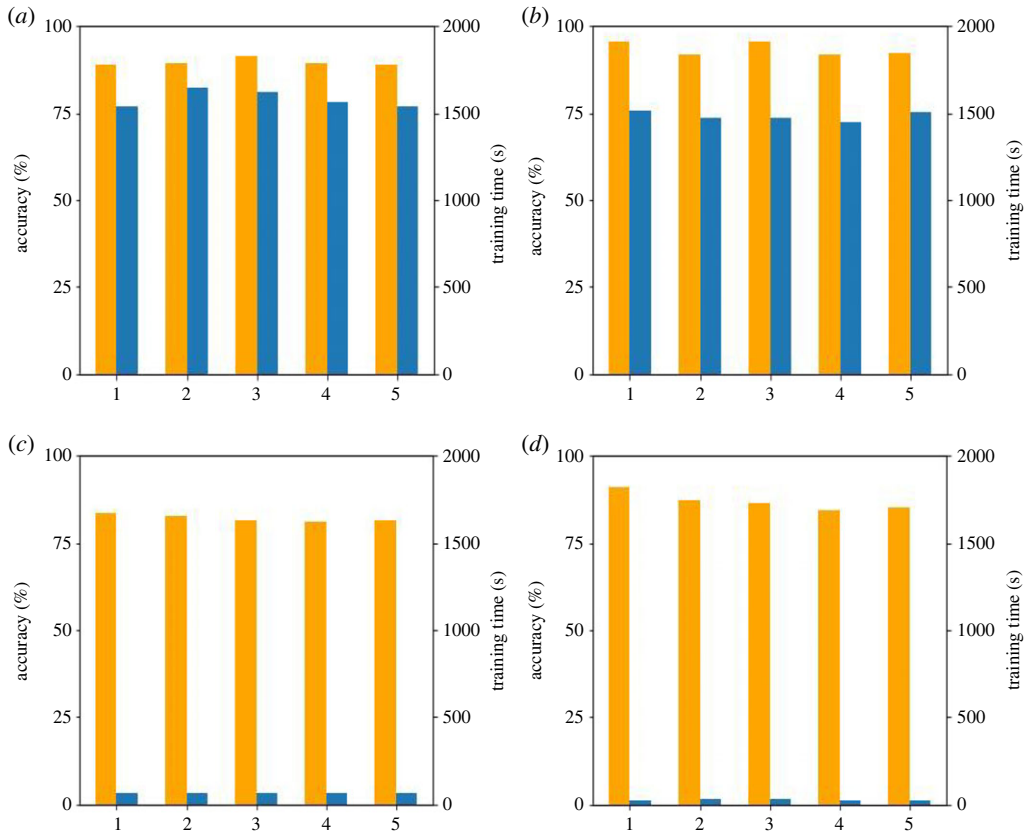


Figure 8. Accuracy and training time of each experiment repeated five times. (a) Test-I, (b) Test-II, (c) Test-III and (d) Test-IV.

Table 6. Indicators of network lightness of the proposed models.

experiment	model	parameter quantity ($\times 10^6$)	storage volume (MB)	MACC ($\times 10^9$)	training time (s)
Test-I	model with original Siamese network	88	336	9.8	1583
Test-II	Lightweight model-1	15	56	9.7	1484
Test-III	Lightweight model-2	36	139	0.18	65
Test-IV	Lightweight model-3	4.9	19	0.11	27

(c) Comprehensive analysis of classification accuracy and lightweight degree

A comprehensive comparison of model performances is presented in table 7. The experimental results of the Original Siamese network were used as the benchmark for comparison. Lightweight model-1 achieved the highest classification accuracy of 93.5%, which was 3.7% higher than the Original Siamese network. Importantly, the accuracy increase referred to the absolute increase in this paper. In addition, Lightweight model-1 significantly compressed the storage volume by 83% and saved 6% training time. For the lightweight effect, Lightweight model-3 performed the best, compressing 94% of the storage volume compared to the model with the original Siamese

Table 7. Comprehensive comparison of Siamese networks.

model	parameter quantity		storage volume		MACC		training time		test accuracy	
	($\times 10^6$)	reduction (%)	(MB)	reduction (%)	($\times 10^9$)	reduction (%)	(s)	reduction (%)	(%)	increase (%)
model with original Siamese network	88	/	336	/	9.8	/	1583	/	89.8	/
Lightweight model-1	15	83	56	83	9.7	1	1484	6	93.5	3.7
Lightweight model-2	36	59	139	59	0.18	98	65	96	82.3	-7.5
Lightweight model-3	4.9	94	19	94	0.11	99	27	98	86.9	-2.9

network and taking only 27 s to complete the training. However, the classification accuracy of pavement texture based on Lightweight model-3 was reduced by 2.9%.

In essence, GAP significantly reduced the storage volume by replacing dense layers, while one-dimensional convolution also achieved some effect by reducing the number of input nodes in dense layers. Moreover, one-dimensional convolution drastically cut the training time of the Siamese network by reducing the MACC of the convolutional layers. However, GAP had little effect since the dense layers only account for a small fraction of the MACC of the network. It should be noted that the reduction rate of storage volume and parameter quantity was the same since the network saved only parameters when storing. However, the reduction rate of training time was different from the reduction rate of MACC due to reasons such as the amount of computation of the activation function not being counted. Additionally, GAP improved accuracy slightly, while one-dimensional convolution had a slight effect on reducing accuracy.

5. Conclusion

This study proposes a few-shot learning approach based on a Siamese network for classifying pavement texture with a limited dataset. To enhance the practical application of the models, lightweight methods are incorporated into the Siamese network. Comparative experiments are conducted to evaluate the performance of the original Siamese network and the lightweighted Siamese networks in a four-way five-shot pavement texture classification task. The findings indicate that (i) all models achieve over 80% accuracy, confirming the feasibility and effectiveness of few-shot learning for pavement texture classification. Moreover, implementing lightweight methods reduces the storage volume and training time of the models without compromising accuracy. (ii) The model that applies GAP on the dense layers of the Siamese network attains a classification accuracy of 93.5% while reducing storage volume by 83% and saving 6% of training time, thus serving as a suitable choice for high-accuracy classification tasks. (iii) The lightweight model that uses one-dimensional convolution on convolutional layers and GAP on dense layers may be more appropriate for tasks that prioritize classification speed and storage volume, with only 19 MB of storage requirement and less training time.

As an initial exploration of pavement texture classification using few-shot learning, this study has several limitations that warrant acknowledgement. First, the study only investigated the use of Siamese networks for constructing few-shot learning models. Other network architectures that implement few-shot learning may offer advantages over Siamese networks and should be considered for further investigation. Second, the dataset used in this study is limited due to the difficulty of obtaining pavement texture images, which constrained the types of pavement textures considered in the recognition model. Considering the diverse range of pavement textures and the potential variability within the same type of pavement surface, future research should focus on expanding the applicable types of textures and improving recognition accuracy with larger datasets. Finally, there may be additional approaches for creating lightweight pavement texture classification models beyond the two methods employed in this study. Future research should investigate the development of lightweight models that both improve accuracy and significantly reduce storage volume and training time.

Data accessibility. The test/train dataset used in this study is available as supplementary material [46].

Authors' contributions. S.P.: data curation, formal analysis, methodology and writing—original draft; H.Y.: conceptualization, methodology, project administration and writing—review and editing; Z.L.: methodology, supervision, writing—original draft and writing—review and editing; N.C.: conceptualization, investigation and writing—review and editing; Y.M.: data curation, resources and writing—review and editing; Y.H.: conceptualization, data curation, funding acquisition, methodology, project administration, resources, supervision and writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This work was supported by Key Science and Technology Projects in the Transportation Industry in 2021 (grant no. 2021-ZD2-047), Key Science and Technology Projects in the Transportation Industry in 2022

(grant no. 2022-ZD6-116), Plan Project of Shandong Transportation S&T (grant no. 2021B49), Opening Project Fund of Materials Service Safety Assessment Facilities (grant no. MSAF-2021-005) and National Natural Science Foundation of China (grant no. 51978048).

References

- Rith M, Kim YK, Lee SW. 2020 Characterization of long-term skid resistance in exposed aggregate concrete pavement. *Construct. Build. Materials* **256**, 119423. (doi:10.1016/j.conbuildmat.2020.119423)
- Li QJ, Zhan Y, Yang GW, Wang KCP. 2020 Pavement skid resistance as a function of pavement surface and aggregate texture properties. *Int. J. Pavement Eng.* **21**, 1159–1169. (doi:10.1080/10298436.2018.1525489)
- Del Pizzo LG, Bianco F, Moro A, Schiaffino G, Licitra G. 2021 Relationship between tyre cavity noise and road surface characteristics on low-noise pavements. *Transport. Res. Part D* **98**, 102971. (doi:10.1016/j.trd.2021.102971)
- Zhang ZW, Luan B, Liu X, Zhang MQ. 2020 Effects of surface texture on tire-pavement noise and skid resistance in long freeway tunnels: from field investigation to technical practice. *Appl. Acoust.* **160**, 107120. (doi:10.1016/j.apacoust.2019.107120)
- Gkyrtis K, Loizos A, Plati C. 2020 Integrating pavement sensing data for pavement condition evaluation. *Sensors* **21**, 3104. (doi:10.3390/s21093104)
- Vieira T, Sandberg U, Erlingsson S. 2021 Negative texture, positive for the environment: effects of horizontal grinding of asphalt pavements. *Road Materials Pavement Design* **22**, 1–22. (doi:10.1080/14680629.2019.1610476)
- Keberia PM, Khosravi A, Salaken SM, Nahavandi S. 2020 Deep imitation learning for autonomous vehicles based on convolutional neural networks. *IEEE-CAA J. Automat. Sin.* **7**, 82–95. (doi:10.1109/JAS.2019.1911825)
- Perelli A, Andersen MS. 2021 Regularization by denoising sub-sampled Newton method for spectral CT multi-material decomposition. *Phil. Trans. R. Soc. A* **379**, 20200191. (doi:10.1098/rsta.2020.0191)
- Duggento A, Conti A, Guerrisi M, Toschi N. 2021 A novel multi-branch architecture for state of the art robust detection of pathological phonocardiograms. *Phil. Trans. R. Soc. A* **379**, 20200264. (doi:10.1098/rsta.2020.0264)
- Hou Y, Li QH, Han Q, Peng B, Wang LB, Gu XY, Wang DW. 2021 MobileCrack: object classification in asphalt pavements using an adaptive lightweight deep learning. *J. Transport. Eng. Part B* **147**, 04020092. (doi:10.1061/JPEODX.0000245)
- Chen N, Xu ZJ, Liu Z, Chen YH, Miao YH, Li QH, Hou Y, Wang LB. 2022 Data augmentation and intelligent recognition in pavement texture using a deep learning. *IEEE Trans. Intell. Transp. Syst.* **23**, 25 427–25 436. (doi:10.1109/TITS.2022.3140586)
- Liu CQ, Li J, Gao J, Yuan DD, Gao ZQ, Chen ZJ. 2021 Three-dimensional texture measurement using deep learning and multi-view pavement images. *Measurement* **172**, 108828. (doi:10.1016/j.measurement.2020.108828)
- Dong SH, Han S, Wu C, Xu OM, Kong HY. 2022 Asphalt pavement macrotexture reconstruction from monocular image based on deep convolutional neural network. *Computer-Aided Civil Infrastructure Eng.* **37**, 1754–1768. (doi:10.1111/mice.12878)
- Yoo J, Lee CH, Jea HM, Lee SK, Yoon Y, Lee J, Yum K, Hwang SU. 2022 Classification of road surfaces based on CNN architecture and tire acoustical signals. *Appl. Sci. Basel* **12**, 9521. (doi:10.3390/app12199521)
- Chhay L, Park LJ, Kim YK, Lee SW. 2023 Counting of exposed aggregate number on pavement surface based on computer vision technique. *Road Materials Pavement Design* **24**, 86–102. (doi:10.1080/14680629.2021.2011381)
- Bromley J, Guyon I, LeCun Y, Säckinger E, Shah R. 1993 Signature verification using a Siamese time delay neural network. *Adv. Neural Information Process. Syst.* **6**, 737–744. (doi:10.1142/S0218001493000339)
- Chopra S, Hadsell R, LeCun Y. 2005 Learning a similarity metric discriminatively, with application to face verification. In *Conf. on Computer Vision and Pattern Recognition, San Diego, CA, 20–25 June 2005*, pp. 539–546. New York, NY: IEEE. (doi:10.1109/cvpr.2005.202)

18. Vinyals O, Blundell C, Lillicrap T, Kavukcuoglu K, Wierstra D. 2016 Matching networks for one shot learning. In *30th Conf. on Neural Information Processing Systems (NIPS), Barcelona, 5–10 December 2016*. Red Hook, NJ: Curran Associates. (doi:10.48550/arXiv.1606.04080)
19. Snell J, Swersky K, Zemel R. 2017 Prototypical networks for few-shot learning. In *31st Annu. Conf. on Neural Information Processing Systems (NIPS), Long Beach, CA, 4–19 December 2017*. Red Hook, NJ: Curran Associates. (doi:10.48550/arXiv.1703.05175)
20. Hu H, Gu JY, Zhang Z, Dai JF, Wei YC. 2018 *Relation networks for object detection*. In *31st IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, 18–27 June 2018*, pp. 3588–3597. New York, NY: IEEE. (doi:10.1109/CVPR.2018.00378)
21. Wang Z, Ma PC, Chi ZQ, Li DD, Yang H, Du WL. 2022 Multi-attention mutual information distributed framework for few-shot learning. *Expert Syst. Appl.* **202**, 117062. (doi:10.1016/j.eswa.2022.117062)
22. Hu ZP, Li ZJ, Wang XY, Zheng SY. 2022 Unsupervised descriptor selection based meta-learning networks for few-shot classification. *Pattern Recognit.* **122**, 108304. (doi:10.1016/j.patcog.2021.108304)
23. Singh P, Mazumder P. 2022 Dual class representation learning for few-shot image classification. *Knowledge-Based Syst.* **238**, 107840. (doi:10.1016/j.knosys.2021.107840)
24. Zeng QJ, Geng J. 2022 Task-specific contrastive learning for few-shot remote sensing image scene classification. *ISPRS J. Photogrammetry Remote Sens.* **191**, 143–154. (doi:10.1016/j.isprsjprs.2022.07.013)
25. Dong HW, Song KC, Wang Q, Yan YH, Jiang P. 2022 Deep metric learning-based for multi-target few-shot pavement distress classification. *IEEE Trans. Ind. Inf.* **18**, 1801–1810. (doi:10.1109/TII.2021.3090036)
26. Guo YR, Du RY, Li XX, Xie JY, Ma ZY, Dong Y. 2022 Learning calibrated class centers for few-shot classification by pair-wise similarity. *IEEE Trans. Image Process.* **31**, 4543–4555. (doi:10.1109/TIP.2022.3184813)
27. Sun L, Liu ZB, Sun XY, Liu LC, Lan RS, Luo XN. 2022 Lightweight image super-resolution via weighted multi-scale residual network. *IEEE-CAA J. Automatica Sin.* **8**, 1271–1280. (doi:10.1109/JAS.2021.1004009)
28. Jiang XR, Wang NN, Xin JW, Xia XB, Yang X, Gao XB. 2021 Learning lightweight super-resolution networks with weight pruning. *Neural Netw.* **144**, 21–32. (doi:10.1016/j.neunet.2021.08.002)
29. Grassucci E, Zhang A, Comminiello D. 2022 PHNNs: lightweight neural networks via parameterized hypercomplex convolutions. *IEEE Trans. Neural Networks Learn. Syst.* (doi:10.1109/TNNLS.2022.3226772)
30. Cheng J, Wu YL, Yuepeng E, You JL, Li T, Li H, Ge JG. 2021 MATEC: a lightweight neural network for online encrypted traffic classification. *Comput. Netw.* **199**, 108472. (doi:10.1016/j.comnet.2021.108472)
31. Liu Z, Pan S, Gao Z, Chen N, Li F, Wang L, Hou Y. 2023 Automatic intelligent recognition of pavement distresses with limited dataset using generative adversarial networks. *Automat. Construct.* **146**, 104674. (doi:10.1016/j.autcon.2022.104674)
32. Gong TT, Zhou WJ, Qian XH, Lei JS, Yu L. 2023 Global contextually guided lightweight network for RGB-thermal urban scene understanding. *Eng. Appl. Artif. Intell.* **117**, 105510. (doi:10.1016/j.engappai.2022.105510)
33. He Z, He D. 2020 Spatial-adaptive Siamese residual network for multi-/hyperspectral classification. *Remote Sensing* **12**, 1640. (doi:10.3390/rs12101640)
34. Jia S, Jiang SG, Lin ZJ, Xu M, Sun WW, Huang Q, Zhu JS, Jia XP. 2022 A semisupervised Siamese network for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **60**, 1–7. (doi:10.1109/TGRS.2021.3116138)
35. Wu YJ, Li J, Wu J, Chang J. 2020 Siamese capsule networks with global and local features for text classification. *Neurocomputing* **390**, 88–98. (doi:10.1016/j.neucom.2020.01.064)
36. Cheng YH, Wang AM, Wu L. 2022 A classification method for electronic components based on Siamese network. *Sensors* **22**, 6478. (doi:10.3390/s22176478)
37. Zhou W, Jia JY, Huang CX, Cheng YQ. 2020 Web3D learning framework for 3D shape retrieval based on hybrid convolutional neural networks. *Tsinghua Sci. Technol.* **25**, 93–102. (doi:10.26599/TST.2018.9010113)

38. Mehmood A, Maqsood M, Bashir M, Yang SY. 2020 A deep Siamese convolution neural network for multi-class classification of Alzheimer disease. *Brain Sci.* **10**, 84. (doi:10.3390/brainsci10020084)
39. Lin M, Chen Q, Yan S. 2013 Network in network. arXiv preprint arXiv:1312.4400.
40. Chen Y. 2015 Convolutional neural network for sentence classification. MS thesis, University of Waterloo.
41. Miao YH, Song PP, Gong XQ. 2014 Fractal and multifractal characteristics of 3D asphalt pavement macrotexture. *J. Mater. Civ. Eng.* **26**, 04014033. (doi:10.1061/(ASCE)MT.1943-5533.0000912)
42. Rahimian E, Zabihi S, Asif A, Farina D, Atashzar SF, Mohammadi A. 2021 FS-HGR: few-shot learning for hand gesture recognition via electromyography. *IEEE Trans. Neural Syst. Rehabil. Eng.* **29**, 1004–1015. (doi:10.1109/TNSRE.2021.3077413)
43. Gao FR, Cai LJ, Yang ZY, Song SJ, Wu C. 2022 Multi-distance metric network for few-shot learning. *Int. J. Mach. Learn. Cybernetics* **13**, 2495–2506. (doi:10.1007/s13042-022-01539-1)
44. Tian RL, Shi HM. 2023 Momentum memory contrastive learning for transfer-based few-shot classification. *Appl. Intelligence* **53**, 864–878. (doi:10.1007/s10489-022-03506-3)
45. Cheng J, Hao FS, He FX, Liu L, Zhang QS. 2023 Mixer-based semantic spread for few-shot learning. *IEEE Trans. Multimedia* **25**, 191–202. (doi:10.1109/TMM.2021.3123813)
46. Pan S, Yan H, Liu Z, Chen N, Miao Y, Hou Y. 2023 Automatic pavement texture recognition using lightweight few-shot learning. Figshare. (doi:10.6084/m9.figshare.c.6725581)