

Applied Mathematics and Nonlinear Sciences

<https://www.sciendo.com>

Deep Convolutional Neural Networks for Image Reconstruction and Damage Recognition in UAV Bridge Inspection

Shun Wang^{1,†}, Mingwei Sun²

1. Shanghai Municipal Highway Engineering Inspection Co., Ltd., Shanghai, 201108, China.

2. Jiangbei District Highway and Transportation Management Center, Ningbo, Zhejiang, 315000, China.

Submission Info

Communicated by Z. Sabir

Received October 27, 2024

Accepted February 23, 2025

Available online March 26, 2025

Abstract

In this paper, bridge image data are collected using a UAV, and the collected images are denoised and enhanced using wavelet analysis methods and nonlinear variations. A deep convolutional neural network (DCNN) is used to construct a model for the reconstruction task of the acquired bridge images and to reduce the impact of blurring and other problems generated in the process of image compression on the bridge damage identification. An improved deep convolutional neural network with step-by-step input capability is proposed, and after feature extraction by MobileNet-v2 lightweight network through deep separable convolution operation, the spine neural network is utilized to construct a decision module so that the decision-making information is fed into the fully connected layer to obtain the bridge damage recognition results. In this paper, the mean value of the peak signal-to-noise ratio of the image obtained after denoising the image using wavelet analysis is 51.27, and the mean value of the structural similarity is 0.974, which is significantly better than other denoising methods. The image reconstruction model constructed based on DCNN improves by more than 19% compared with all other algorithms, showing high accuracy and high efficiency. The bridge damage recognition model has an accuracy of more than 90% for the damage recognition of different working condition locations of the bridge, which can effectively realize the bridge damage recognition based on the images collected by UAV.

Keywords: Deep convolutional neural network; Wavelet analysis; Nonlinear transformation; Spine neural network; Damage recognition.

AMS 2010 codes: 68T45

[†]Corresponding author.

Email address: wangshun-1984@163.com

1 Introduction

Bridges are an important part of highways and also an important part of highway maintenance and management [1]. With the increase in traffic flow and automobile tonnage, highway bridges and municipal bridges with disease work are more common, which can easily cause traffic accidents [2-3]. Due to the lack of maintenance units for high-altitude operation, ships, and other related professional testing equipment, resulting in large, extra-large bridges, high-altitude tower columns, cables, steel pipe vaults, and other special important components of the detection work is difficult to carry out normally, resulting in special structural parts of the bridge disease can not be found in a timely manner and dealt with, in the long run for the management and maintenance of the important bridges is extremely detrimental to the maintenance of the bridge [4-8]. In addition to bridge damage also includes natural disasters and other factors [9-10]. Therefore, how to utilize scientific and technological means to monitor and inspect the safety of bridges during the service period to ensure that bridges are used for a long period has become a hot topic of research in the field of bridges at home and abroad [11-12].

At present, the mainstream bridge disease detection method is still using artificial bridge cars, telescopes, and other traditional means [13-14]. It has been proved by practice that with the traditional means of detection, there are many problems, and it is not easy to ensure an all-round detection of the bridge [15]. In recent years, due to the rapid development of drone camera technology, the technology has been utilized in the field of bridge appearance detection [16-17]. That is, the use of drones carrying image acquisition system to collect information, and then the use of image processing technology, noise removal, compression, and reduction of the image so that you can have a more clear intuitive understanding of the bridge's appearance of the disease condition, so as to improve the detection efficiency and reduce the detection cost [18-22]. Therefore, the application of digital image processing technology in bridge detection, with the ground control equipment and all kinds of powerful information processing software, effectively solves the problem of difficulty in observing the special structural parts of the bridge so that the bridge management and maintenance work into a new height [23-27].

In this paper, the DJI Mavic Ar2i quadcopter UAV is selected as the bridge image acquisition device, and the ideal bridge image is successfully acquired by setting the shooting distance. For the problems of redundancy and useless information in the original bridge images, the wavelet transform is utilized to process the original bridge images to eliminate the repetition of the image information and, at the same time, remove the noise in the images. In order to improve the visual effect of the image, dynamic compression is performed on the bridge rail image after wavelet analysis, and nonlinear transformation is used to make the bridge image softer under different lighting environments to enhance the quality of the captured bridge image, and to improve the efficiency and accuracy of image reconstruction. A bridge image reconstruction model is constructed based on the local perception and other capabilities of the deep convolutional neural network, and the enhanced image is input into the deep convolutional neural network for training. A stochastic method is used to initialize the relevant parameters of the deep convolutional neural network, and the gradient descent method is used for updating. Bridge image reconstruction experiments are carried out using the established model to analyze the reconstruction effect of the model in this paper. The MobileNet-v2 lightweight network improves the traditional deep convolutional neural network, and the improved deep convolutional neural network is utilized to extract damaged features from the reconstructed bridge images. Based on the unique local inference ability of the spine neural network, the decision layer of the spine neural network is introduced, and the thought features after spreading are split into two equal parts and input into it, combined with the accelerating effect of activation function on the model training, the identification results of the model on the bridge damage are finally obtained. Different working

condition positions and bridge damage indicators are selected to verify the performance of the bridge damage recognition model constructed in this paper in the training set and test set, respectively.

2 Bridge image acquisition

2.1 Drone equipment

The bridge structure is complex and varied, so the UAV needs to be flexible enough when collecting bridge images. DJI Mavic Ar2i quadcopter is selected for image acquisition, which has a compact body, is equipped with a seven-eye vision system, can hover and fly stably outdoors, and is equipped with a three-axis gimbal camera that can stably take 48-megapixel photos. However, the angle of view of the camera is $-90^{\circ}\sim+24^{\circ}$ (the angle of view is considered as 0° , downward is negative, and upward is positive). Therefore, a light conversion device and a corresponding stabilization device are designed in front of the camera in order to capture the image of the top of the UAV. The main component of the light conversion device is a trapezoidal reflector in front of the lens, which is at an angle of 45° to the vertical plane. When collecting images of the side of the bridge, the original UAV equipment can be used, but when collecting images of the bottom of the bridge, the light conversion device is installed under the fuselage, and the lens faces forward to capture images of the top of the UAV.

2.2 Shooting distance

The shooting distance is the key factor that affects the pixel size and the actual size. When shooting, the camera is parallel to the shooting plane so that the distance from the objects in the acquisition range to the lens is as consistent as possible. In order to obtain higher quality images as much as possible under the premise of ensuring the safety of the UAV, the shooting distance should be determined according to the pixel accuracy and the requirements of bridge disease detection. According to the relevant technical specifications, the crack width of the concrete bridge and prestressed concrete bridge superstructure is limited to 0.3~0.8mm, so the shooting distance needs to be satisfied with the observation of cracks of 0.1mm wide. According to the pixel accuracy of 5000pixel $\times$ 5000pixel, the shooting area of a single image is judged to be 50mm $\times$ 100mm. The orthogonal straight lines and borders of 50mm $\times$ 100mm are drawn on the white paper. The shooting lens is moved so that the straight lines are all over the lens. The shooting distance is recorded by repeating many times, and the average value of the shooting distance is 560mm, i.e., the shooting distance is required to be less than 560mm. The following formula can express the pixel size and shooting distance:

$$l/l' = P/P' \quad (1)$$

Where l is the maximum shooting distance of 560 mm. l' is the shooting distance determined according to the actual situation ($l' \leq 560\text{mm}$). P is the shooting pixel edge length of 0.1 mm. P' is the actual pixel edge length.

3 DCNN-based bridge image reconstruction method

3.1 Denoising of bridge images

The original bridge image has a three-dimensional structure, including a large amount of information. There is a certain degree of redundancy between the information, and it also contains a lot of useless

information, which negatively affects the efficiency of the image reconstruction and, at the same time, makes the image transmission speed slower and the storage space of the bridge image large. Therefore, the wavelet transform is used for preprocessing of the original bridge image to eliminate the repetition of the image information and at the same time, to remove the image Noise. Let the original bridge image signal be s_j , and by introducing the wavelet analysis technique to decompose s_j [28], a finer bridge image signal is obtained as follows:

$$s_{1,l}^0 = \text{even}(s_j) = s_{2l} \quad (2)$$

$$d_{1,l}^0 = \text{odd}(s_j) = s_{2l+1} \quad (3)$$

where l represents the acquisition time of the bridge image signal.

The estimation of $d_{1,l}^0$ using $s_{1,l}^0$ of the bridge image signal yields an estimation error of:

$$d_{1,l}^i = d_{1,l}^{i-1} - \sum_{\kappa} p^i K S_{1,l-k}^{i-1} \quad (4)$$

Where, i denotes the bridge image signal sequence and P denotes the bridge image information quantity.

The bridge image signal is obtained after M times wavelet analysis:

$$s_{1,l} = s_{1,l}^M / K \quad (5)$$

$$d_{1,l} = K d_{1,l}^M \quad (6)$$

After the above process, the bridge image information processing can be realized, and then wavelet analysis is used to denoise the bridge image operation.

3.2 Enhancement operations for bridge images

In order to make the information in the bridge image better displayed and improve the visual effect of the image, the dynamic compression of the bridge image after wavelet analysis using logarithmic transformation is adopted so that the dark pixels in the image amplitude expanded, making the image closer to the human visual system. Let the bridge image light component be $v(x, y)$, then the logarithmic transformation can be described as:

$$R_1(x, y) = \log(1 + v(x, y)) \quad (7)$$

After the bridge image is logarithmically transformed, it may produce some high-light areas, which makes the bridge image too bright. For this reason, the bridge image is then nonlinearly transformed to make the bridge image softer in different lighting environments:

$$I_2(x, y) = \left(\frac{2}{1 + \exp(-m)} - 1 \right) \times 255 - v(x, y) \quad (8)$$

$$m = k(v) \times \frac{v}{v + a\bar{v}} \quad (9)$$

$$k(v) = \begin{cases} 7 & 0 \leq v \leq 60 \\ (v - 600) / 70 + 7 & 60 < v \leq 200 \\ (v - 200) / 55 + 9 & 200 < v \leq 255 \end{cases} \quad (10)$$

Where, v is the mean gray value of the bridge image and a denotes the adjustment factor of the illumination analysis amount.

After realizing dynamic compression for bridge images, the probability of causing bridge images to be more blurred is quite high, therefore, it is necessary to carry out inverse sharpening processing for bridge images to further enrich the detailed information of bridge images. Let the Gaussian blurring of bridge images [29] can be described as:

$$U(x, y) = I(x, y) * G(x, y) \quad (11)$$

Where $I(x, y)$ denotes the obtained bridge image and $G(x, y)$ denotes the Gaussian function which is defined as:

$$G(x, y) = \frac{1}{2\pi\sigma^2} \cdot \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (12)$$

Anti-sharpening can be described:

$$I_{\text{En}}(x, y) = \frac{c}{2c-1} I(x, y) - \frac{1-c}{2c-1} U(x, y) \quad (13)$$

The inverse sharpening process can remove the low-frequency information in the bridge image and retain the high-frequency information of the bridge image so that the detailed information of the bridge image is significantly enhanced.

3.3 Bridge Image Reconstruction with DCNN

3.3.1 Deep Convolutional Neural Networks

Deep Convolutional Neural Network (DCNN) belongs to the type of feedforward neural network [30], but deep convolutional neural networks have superior learning performance compared to traditional feedforward neural networks. Deep convolutional neural networks have the following advantages:

- 1) Local perception ability: each neuron can perceive the connection between the bridge image pixels, and these local connections can be composed together to get the global connection of the bridge image, laying the foundation for the bridge image reconstruction.
- 2) Weight sharing ability, different bridge image regions, can be feature extracted by convolutional layer to extract bridge image reconstruction features from different perspectives, more comprehensively describing the connection between bridge image regions.

- 3) Pooling layer: A deep convolutional neural network can increase the pooling layer in order to reduce the dimension of bridge image reconstruction features, and at the same time, it has strong robustness to the bridge image rotation and other transformations, which reduces the bridge image reconstruction workload and reconstruction time.

DCNN includes many convolutional layers and pooling layers, mapping the input and output layers through the intermediate layers, and its basic structure is shown in Figure 1.

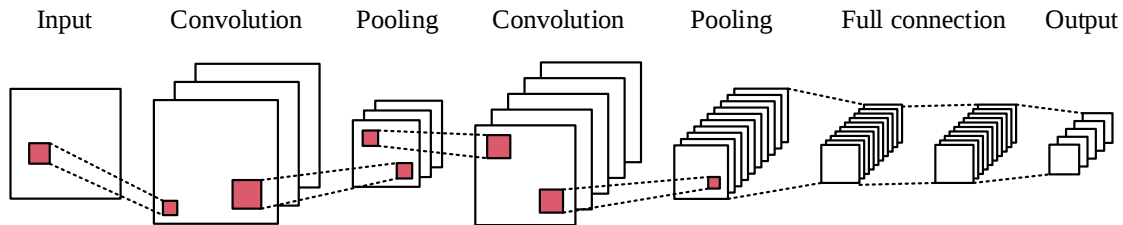


Figure 1. Basic structure of deep convolutional neural network

3.3.2 Steps for DCNN-based bridge image reconstruction

Step 1: Acquire bridge images by UAV equipment, normalize the bridge images to be reconstructed so that they are of the same size and dimension, and also perform horizontal flipping and panning operations.

Step 2: A denoising method is designed using wavelet transform, which is adapted to perform the denoising operation of bridge images to reduce the negative impact of noise on bridge images.

Step 3: Enhance the denoised bridge image by using the nonlinear transform method to make the bridge image brightness and contrast more reasonable, improve the bridge image clarity, and remove ambiguity.

Step 4: Input the enhanced processed bridge image into the deep convolutional neural network for training, initialize the relevant parameters of the deep convolutional neural network by random method through the convolution, pooling, and other operations, and calculate the error of the output layer.

Step 5: When the output layer error of the deep convolutional neural network is larger than the practical application requirements, it enters the error feedback operation and adopts the gradient descent method to update the relevant parameters of the deep convolutional neural network so as to reduce the output error continuously.

Step 6: When the output layer error of the deep convolutional neural network is less than the actual application requirements, establish the optimal bridge image reconstruction model.

Step 7: Conduct bridge image reconstruction experiments using the established model and analyze its reconstruction effect.

4 Bridge damage identification technology based on DCNN

4.1 Model Architecture

In this section, a deep neural network is first utilized to achieve fast extraction of bridge damage features after image reconstruction. Then, the spine neural network with step-by-step input is constructed to accurately recognize the damage, and the overall framework of the model is shown in Fig. 2. Based on the model parameter migration and the characteristics of the spine neural network module, this paper proposes an improved lightweight deep convolutional neural network with step-by-step input capability. The network obtains $7*7*128$ high-dimensional feature maps after deep separable convolution operation, and the redundant information in the shallow features is constructed through the maximum pooling layer to obtain $1*1*1280$ highly robust feature maps as the input of the spine neural network. The input part is split into two parts and gradually fed into the spine neural network. Finally, in this paper, the outputs of multiple spine neural networks with local inference ability are fused and the features are input to the fully connected layer to get the final classification results.

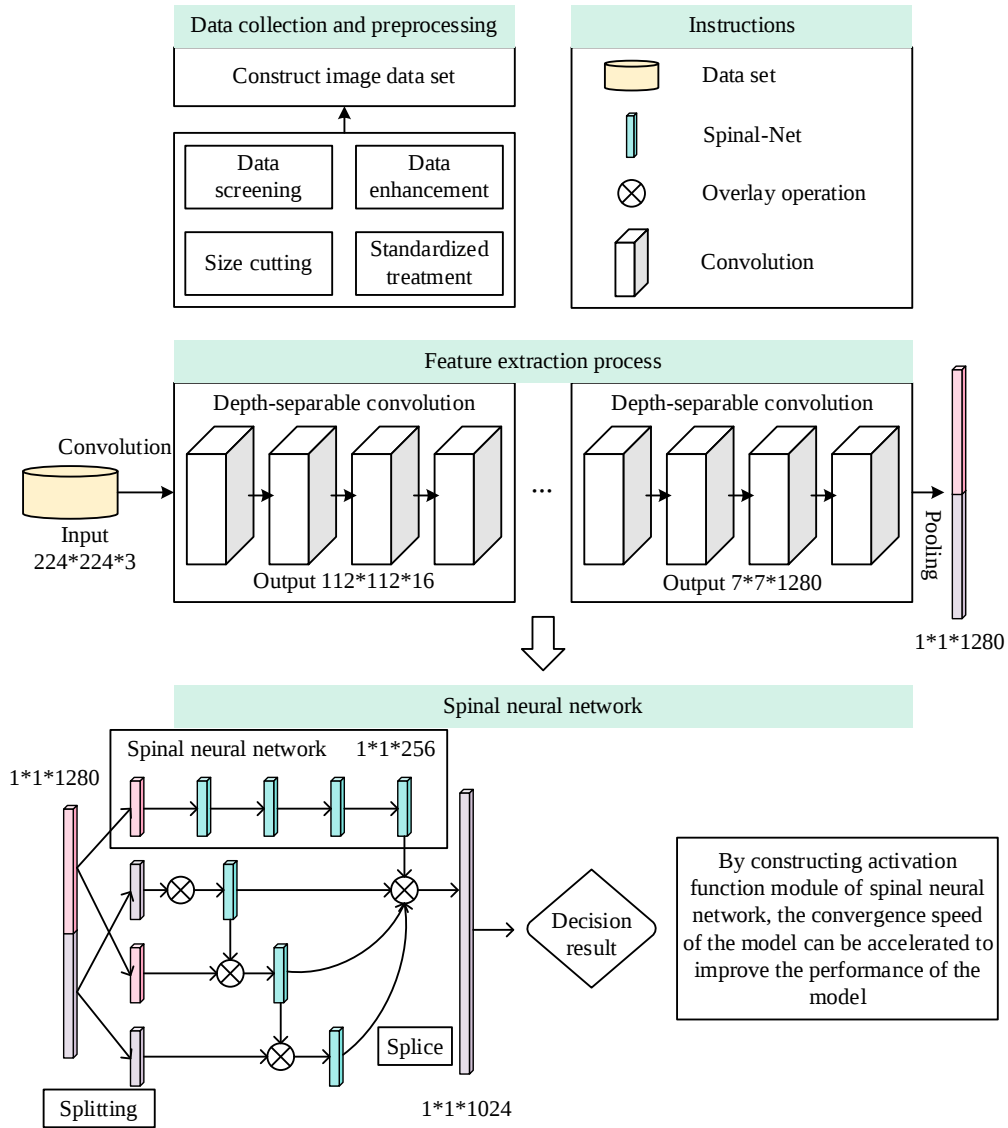


Figure 2. Overall network framework

4.2 Feature extraction process

4.2.1 Migration learning

The disadvantage of deep convolutional neural network models is that they are data-dependent and require a large amount of data available for training. In the bridge damage recognition task under UAV inspection, there is a problem with small data samples. Migration learning reduces the sample dependence on data and improves the performance of the model by transferring the a priori knowledge learned by the model in a large number of datasets to the learning process of the target dataset. In this paper, we choose the migration learning strategy based on network parameters. Firstly, the ImageNet dataset is utilized to train the network in the source domain. Secondly, the feature extraction part of the pre-trained network is used for model design to train the target dataset. Finally, the parameters of the classifiers are fine-tuned by freezing the parameter updates of the feature extraction module. This results in a shared model structure and parameters, which improves accuracy while reducing configuration requirements and decreasing training time.

4.2.2 Light-weighted networks

MobileNet-v2, a typical representative of lightweight networks [31], is widely used for image classification tasks. The model tends to have a smaller model space complexity compared to traditional deep convolutional neural networks, allowing it to work on mobile devices and computers with lower computing power. The depth-separable convolution in the model creatively separates the traditional convolution process into depth (DW) convolution and point-by-point (PW) convolution, effectively reducing the computational effort of the model.

One of the basic modules consists of two DW convolutions as well as one PW convolution. The computational procedure is shown below. Suppose the input sample is x . After the DW convolution, the feature mapping of the hidden layer is as in Equation (14):

$$D_w(x) = \sigma_1(B_N(x \otimes w_3^k + b^k)) \quad (14)$$

In Equation (14), σ_1 denotes the activation function, B_N refers to the normalization process, x refers to the original input, $\otimes$ denotes the convolution process, w denotes the convolution kernel, 3 denotes the dimension of the convolution kernel as 3*3, and k is the number of convolution kernels. b represents the bias of the network, k represents the dimension of the bias.

And P_w convolution chooses whether to use a nonlinear activation function or not according to the different positions it is in. After P_w convolution, the feature mapping of the hidden layer is as in Eq:

$$P_{w1}(x) = \sigma_1(B_N(x \otimes w_1^k + b^k)) \quad (15)$$

$$P_{w2}(x) = B_N(x \otimes w_1^k + b^k) \quad (16)$$

In Eqs. (15), (16), σ_1 denotes the activation function, B_N refers to the normalization process, x refers to the original input, $\otimes$ denotes the convolution process, w represents the convolution kernel, 3 denotes the convolution kernel with the dimension of 3*3, and k is the number of convolution kernels. k denotes the bias of the network and k denotes the dimension of the bias.

A short join section is added at the beginning and end of each base module to weaken the problem of gradient vanishing by superimposing the input x directly onto the final output when the sliding step of the convolution kernel is 1, and the input feature size is equal to the output feature size. The mapping of the output behaves as in Eq. (17):

$$S(x) \begin{cases} = P_{w2}(D_w(P_{w1}(x))) \\ = P_{w2}(D_w(P_{w1}(x))) + x \end{cases} \quad (17)$$

Where P_{w1} and P_{w2} are shown in Eqs. (15), (16), D_w convolution is shown in Eq. (14) and x refers to the original input.

The overall feature extraction network input is shown in equation (18):

$$M(x) = pool(C_{3 \times 3}(S_6(x \otimes w_3^k + b^k))) \quad (18)$$

Where $pool$ refers to the maximum pooling operation, C denotes the 3×3 convolution operation, S_6 denotes the superposition of 6 base modules, x refers to the original input, w represents the convolution kernel, k is the number of convolution kernels, and 3 denotes that the convolution kernel has a dimension of 3×3 . b denotes the bias of the network and k denotes the dimension of the bias.

4.3 Spinal Neural Networks

4.3.1 Decision-making module

Spinal neural networks are nonlinear decision-making layers capable of accepting partial information, with local inference capabilities not generally found in fully connected layers. It can reduce the model error rate and improve the model performance. Usually, the one-dimensional features after flattening are split into two equal parts, which are fed into the spine neural network decision layer, and the decision information of the bottom layer and the split inputs are spliced together as the higher spine neural network decision layer. Finally, all the decision information is spliced into the last fully connected layer to get the output.

After the feature extraction process, the feature M_x , split into two parts A_1 and A_2 shown below:

$$A_1 = \{x_{M1}, x_{M2} \cdots x_{M640}\} \quad (19)$$

$$A_2 = \{x_{M640}, x_{M641} \cdots x_{M1280}\} \quad (20)$$

The spinal neural network used in this section consists of four ganglia, with the first layer of spinal ganglia having input A_1 and output $S_p(B_1)$, as shown below:

$$S_p(B_1) = \sigma_2(Linear(Droop(A_1))) \quad (21)$$

where $Droop$ is a random zeroing operation and $Linear$ is a linear layer.

The second spinal ganglion layer accepts the outputs $S_p(B_2)$ as well as A_2 from the previous one and outputs a $1*1*256$ feature vector $S_p(B_2)$ as follows:

$$B_2 = A_2 + S_p(B_1) \quad (22)$$

$$S_p(B_2) = \sigma_2(\text{Linear}(\text{Droop}(B_2))) \quad (23)$$

The third spinal ganglion layer accepts the outputs $S_p(B_2)$ as well as A_2 from the second layer and outputs a $1*1*256$ feature vector $S_p(B_2)$ as shown below:

$$B_3 = S_p(B_2) \oplus A_1 \quad (24)$$

$$S_p(B_3) = \sigma_2(\text{Linear}(\text{Droop}(B_3))) \quad (25)$$

The last layer of the spinal ganglia accepts the outputs $S_p(B_3)$ as well as A_1 from the third layer and outputs the $1*1*256$ feature vector $S_p(B_4)$ as shown below:

$$B_4 = S_p(B_3) \oplus A_2 \quad (26)$$

$$S_p(A_4) = \sigma_2(\text{Linear}(\text{Droop}(A_4))) \quad (27)$$

The fully connected layer accepts the outputs from all the ganglia and self-learns to assign different weights and outputs the classification result f_x as shown below:

$$B_{all} = \alpha S_p(B_1) \beta \oplus S_p(B_2) \oplus \chi S_p(B_3) \oplus \delta S_p(B_4) \quad (28)$$

$$f(x) = \text{linear}(B_{all}) \quad (29)$$

4.3.2 Activation of modules

The essence of a deep convolutional neural network is a multilayer composite function. The role of the activation function in the neural network is to use the information in the neural network for nonlinear conversion and increase the expression ability of the neural network so as to better fit the nonlinear complex model. The activation module of the spinal neural network can directly affect the model training effect. The appropriate activation function can play a multiplier effect, while the inappropriate activation function will affect the model fitting effect. In the local decision-making model proposed in this section, the activation function, as an important part of the decision-making module, plays an important role in the response correction of the neural network, which is conducive to speeding up the training process and improving the classification performance. In general, Leaky-ReLU and its variants (PReLU, RReLU) outperform the ReLU activation function in general bridge datasets. In this paper, the UAV bridge inspection dataset is not suitable for the ReLU activation function used by conventional spine neural networks due to its characteristics, so the activation function with superior performance is selected by comparing Leaky-ReLU and its variants (PReLU, RReLU).

5 Experimentation and analysis

5.1 Experimental data set

After determining the structure of the algorithm, it is necessary to construct a suitable dataset as the training dataset of the model. Concrete material has the characteristics of easy access, good fire resistance, not easy to weather, and is often used as a material for bridge construction. Therefore, this experiment utilizes UAV equipment to collect multiple concrete bridge images (including damaged and undamaged images) on a bridge, and by rotating panning cropping, and other operations on the acquired images, a dataset of more than 5000 concrete bridge images is obtained, which contains a variety of concrete bridge images. The bridge damage images in the dataset include different states of damage, such as transverse cracks, longitudinal cracks, bridge corrosion, and so on. After obtaining the dataset, the dataset is labeled using the Labeling image annotation tool with the type of Crack, and the training and test sets are divided according to the ratio of 7:3.

5.2 Image Reconstruction Analysis

5.2.1 Analysis of bridge Image denoising effect

In order to verify the effectiveness of the wavelet analysis method for bridge image denoising, the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) are used to objectively evaluate the merits of the denoised image results, which are calculated as follows:

$$PSNR = 10 \lg \frac{255^2 \times N_1 \times N_2}{\|\hat{x} - x\|^2} \quad (30)$$

$$SSIM = \frac{(2\mu_{\hat{x}}\mu_x + C1)(2\sigma_{\hat{x}\hat{x}} + C2)}{(\mu_{\hat{x}}^2 + \mu_x^2 + C1)(\sigma_{\hat{x}}^2 + \sigma_x^2 + C2)} \quad (31)$$

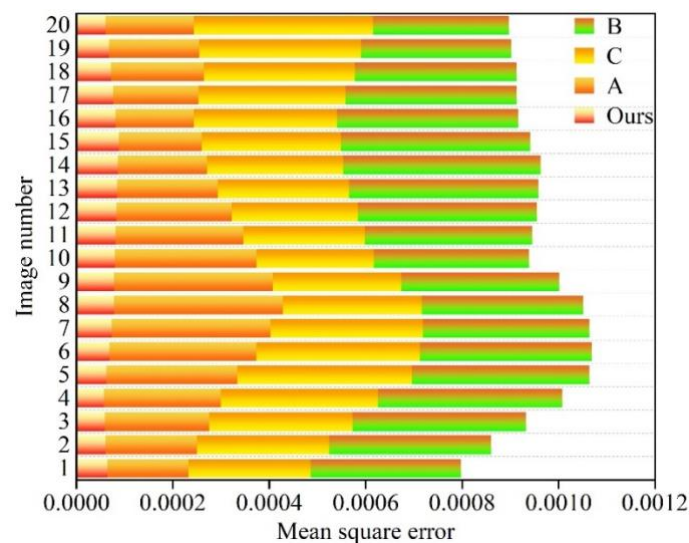
Where, x and $\hat{x}$ are the original image and the denoised image respectively, $\mu_{\hat{x}}, \mu_x$ and $\sigma_{\hat{x}}, \sigma_x$ are the mean and variance, and $\sigma_{\hat{x}\hat{x}}$ is the covariance.

Four complex image denoising models, including wavelet analysis, full variational method (A), partial differential method (B), and median filtering method (C) in this paper, are respectively utilized to process the noisy bridge images collected by UAV in real-time. In order to reduce the error value of the bridge image denoising results, the signal-to-noise ratio PSNR of the processed bridge images is calculated after processing 20 different bridge images by the four denoising models, and the results are shown in Table 1. Analysis of the data in the table reveals that the wavelet analysis method used in this paper denoises the bridge images with the highest signal-to-noise ratio and the cleanest noise removal to obtain the ideal bridge image.

Table 1. The bridge image peak signal-to-noise ratio of different methods

Image number	Ours	A	B	C
1	45.42	15.91	21.53	13.17
2	55.67	27.47	26.06	16.7
3	46.94	23.74	12.98	29.82
4	49.74	26.93	16.75	24.65
5	47.12	15.44	16.27	12.55
6	51.04	26.50	13.21	27.25
7	58.94	27.31	26.79	27.97
8	48.83	28.95	27.93	13.89
9	46.40	17.68	26.87	26.13
10	48.62	26.79	26.10	16.86
11	53.62	25.11	19.46	12.93
12	52.84	24.73	27.26	14.15
13	56.60	20.35	14.08	18.59
14	59.39	16.52	22.49	17.32
15	48.46	26.57	11.67	22.86
16	51.31	18.35	26.02	26.10
17	55.70	23.85	20.18	27.9
18	55.58	20.57	25.56	28.78
19	46.51	15.25	17.84	19.74
20	46.65	16.01	26.57	28.13

On the basis of the above experiments, further testing of different methods in the bridge image denoising after the mean square error, this indicator represents the size of the error between the processed image and the original image, the smaller the value, indicating that the processed image is closer to the original image, i.e., the better the denoising effect, the results are shown in Figure 3. Analysis can be obtained in this paper wavelet analysis method of bridge image denoising. The mean square error is always maintained within 0.0001, has superior performance, and can achieve a clear display of the bridge image.

**Figure 3.** Comparison of the mean square error results

Structural similarity (SSIM) is another evaluation index of bridge image denoising processing. This index can evaluate the degree of similarity between the denoised image grayscale and detailed texture features and the original image. The closer the value is to 1 indicates that the closer the image is to the original image, i.e., the better the denoising effect, the results are shown in Table 2. Analyzing the data in the table, the structural similarity of the bridge image after denoising by the wavelet analysis method in this paper is 0.974 on average, which is closer to 1 than other methods, indicating that the wavelet analysis method has a better denoising effect, and it has certain application value.

Table 2. Comparison of SSIM results

Image number	Ours	A	B	C
1	0.987	0.777	0.683	0.750
2	0.984	0.770	0.678	0.754
3	0.958	0.788	0.682	0.745
4	0.975	0.753	0.656	0.747
5	0.987	0.764	0.669	0.751
6	0.987	0.766	0.661	0.755
7	0.984	0.790	0.664	0.737
8	0.988	0.782	0.662	0.760
9	0.969	0.780	0.667	0.732
10	0.987	0.769	0.666	0.760
11	0.963	0.773	0.679	0.743
12	0.964	0.774	0.645	0.730
13	0.966	0.764	0.683	0.723
14	0.975	0.786	0.657	0.730
15	0.967	0.763	0.641	0.728
16	0.983	0.775	0.660	0.749
17	0.952	0.777	0.644	0.742
18	0.973	0.769	0.681	0.746
19	0.969	0.755	0.645	0.741
20	0.956	0.759	0.686	0.764

5.2.2 Analysis of Bridge Image Reconstruction Effect

In order to verify that the deep convolutional neural network model with completed training has good reconstruction effect and generalization ability, i.e., it can reconstruct untrained bridge images. From the test set, 20 bridge images that have completed the denoising process and contain both damaged and undamaged bridges are selected for image reconstruction using the DCNN algorithm. BP neural network (BP), convolutional neural network (CNN), and feedforward neural network (FNN) are selected as the comparison models of this paper's algorithm to reconstruct the same images under the same experimental environment in order to show the superior performance of this paper's algorithm.

Fig. 4 shows the loss rate and iteration process changes of each model trained on the dataset, from which it can be seen that when the number of iterations is more than 6000, the loss rate of the DCNN model in this paper tends to be stable and maintains at about 0.02. While BP, CNN and FNN have higher values of loss rate than this paper's model when the loss rate tends to stabilize.

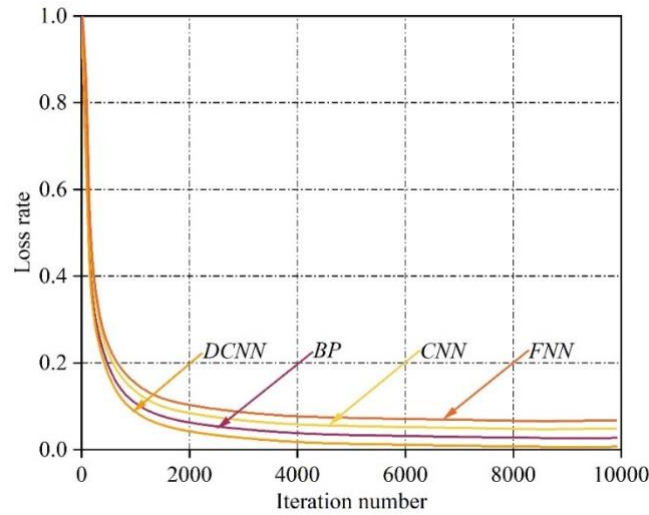


Figure 4. Model training loss rate

Figure 5 shows the reconstruction accuracy and iterative process variation of each model during training. The image reconstruction accuracy is calculated as:

$$Acc = \frac{N_r}{N_a} \times 100\% \quad (32)$$

Where N_r is the number of cells calculated correctly. N_a is the total number of cells.

From the figure, it can be seen that when the number of model training times is more than 6000 times, the image reconstruction accuracy under the model of this paper tends to stabilize, which is about 0.995. At this time, the reconstruction accuracy of the BP neural network, convolutional neural network, and feed-forward neural network are 0.798, 0.682, and 0.603, respectively. Compared with the reconstruction accuracy of this paper's model is improved by 24.69%, 45.89%, and 65.01%, respectively. It shows that the neural convolutional neural network in this paper performs better in the training set and has higher reconstruction accuracy for the bridge images in the training set.

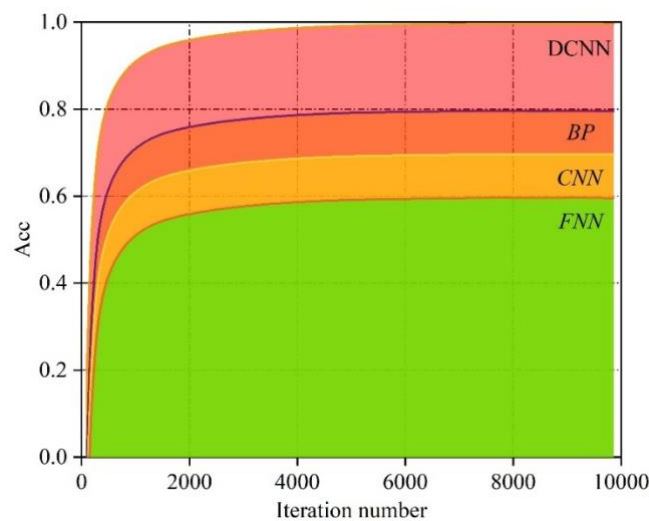


Figure 5. Model training accuracy

The reconstruction accuracy and reconstruction time of each model on 20 bridge images selected from the test set are shown in Table 3. From the table, it can be clearly seen that the mean value of reconstruction accuracy of deep convolutional neural network in this paper on 20 images is 0.981, while the mean values of reconstruction accuracy of the other three algorithms are 0.823, 0.765, and 0.689, respectively. The reconstruction accuracy of this paper's model for the bridge images is improved by 19.20%, 28.24%, and 42.38%, respectively. At the same time, the average time required for this model to complete the reconstruction task of 20 bridge images is 0.459s, while the average reconstruction time of the other three algorithms is 6.789s, 13.788s, and 20.570s, respectively, compared with the reduction of the reconstruction time of the bridge images in this paper's model is 93.24%, 96.67%, and 97.77%, respectively. Combined with the experimental results of the training set, it can be concluded that the deep convolutional neural network used in this paper shows excellent performance of high reconstruction accuracy and short reconstruction time in the bridge image reconstruction task collected by UAVs and is able to efficiently complete the bridge image reconstruction task.

Table 3. Image reconstruction comparison

Number	DCNN		BP		CNN		FNN	
	Acc	Time/s	Acc	Time/s	Acc	Time/s	Acc	Time/s
1	0.976	0.288	0.837	7.772	0.778	13.771	0.754	19.265
2	0.965	0.066	0.747	6.232	0.819	14.257	0.632	24.307
3	0.970	0.071	0.909	7.430	0.779	13.28	0.726	23.820
4	0.985	0.670	0.832	7.237	0.792	15.718	0.639	19.348
5	0.998	0.793	0.789	6.208	0.823	13.643	0.641	18.086
6	0.993	0.318	0.877	7.765	0.795	11.642	0.642	15.660
7	0.998	0.309	0.822	5.572	0.795	12.924	0.703	20.258
8	0.979	0.059	0.768	5.949	0.769	14.637	0.696	24.699
9	0.99	0.590	0.760	7.991	0.790	15.454	0.753	19.772
10	0.985	0.404	0.890	5.503	0.738	14.562	0.625	16.051
11	0.977	0.876	0.745	7.476	0.771	11.004	0.665	16.611
12	0.980	0.161	0.838	6.495	0.742	12.908	0.680	20.618
13	0.999	0.702	0.762	5.825	0.762	15.696	0.735	24.489
14	0.985	0.779	0.812	6.284	0.780	15.526	0.706	17.989
15	0.981	0.456	0.844	5.138	0.702	14.072	0.686	23.111
16	0.963	0.314	0.842	6.920	0.729	14.944	0.678	20.063
17	0.986	0.779	0.906	7.680	0.756	13.074	0.756	20.746
18	0.964	0.203	0.784	7.806	0.766	11.540	0.712	18.323
19	0.971	0.733	0.816	6.618	0.707	12.782	0.714	23.987
20	0.982	0.599	0.873	7.869	0.710	14.328	0.644	24.188
Mean	0.981	0.459	0.823	6.789	0.765	13.788	0.689	20.570

5.3 Bridge Damage Identification Analysis

5.3.1 Selection of Injury Indicators

The selection of damage indicators is a difficult point in the work of bridge damage identification, the ideal selection of bridge damage indicators should have the following characteristics: sensitive to

structural damage. Easy to obtain in the actual engineering. Strong robustness and low sensitivity to testing errors.

At present, the acceleration response signal is used as the damage indicator in damage identification research. In order to study the actual use of different characteristic parameters of bridges for damage identification work, a bridge monitoring parameter is selected as a damage indicator for comparative study.

Deflection is one of the most intuitive indicators to reflect the health condition of the bridge. When the bridge component damage will lead to a decrease in the main frequency of the vibration characteristics of the change, the deflection will become larger. Deflection is also easily accessible data in daily bridge monitoring work.

Therefore, this paper selects deflection and acceleration as damage indicators for damage identification.

5.3.2 DCNN damage identification

Build MobileNet-v2 lightweight network and spinal neural network based on Matlab deep network designer. Bridge Damage Recognition Model Building Based on DCNN During the training process, the denoising processed sample set is re-divided into a training set and test set in the ratio of 7:3. Where the training set samples are used to train the model in order to make the model learn the feature parameters of the data. The test set sample data is used to test the generalization ability of the model.

Fig. 6 shows the model training plot, and (a) and (b) show the sample set model training plot based on deflection data and the sample set model training plot based on acceleration data, respectively. The network extracts the information features of the learning training dataset through three convolution operations and two pooling operations, minimizes the loss rate after several iterations, and then outputs the damage recognition results through the fully connected layer and the classification layer. The model is trained after 200 rounds of iterations to reach the optimal training results.

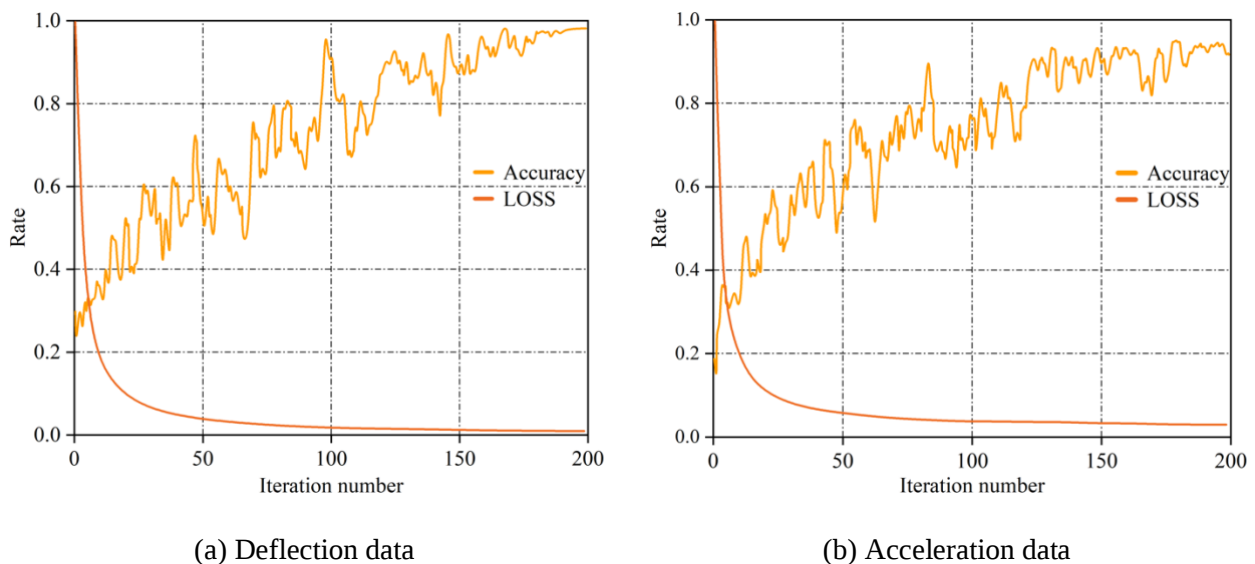


Figure 6. Model training results

5.3.3 Damage Identification Results

At the end of training the model, Matlab was used to write the code to apply the model, and the test set data samples with the same dimension as the model training were input to verify the model's damage recognition accuracy. Table 4 shows the results of model damage identification based on the deflection damage index, and Table 5 shows the results of model damage identification based on the acceleration damage index. 1 in the table indicates that there is no damage to the bridge, and 2-10 indicate the working condition positions of the north side beam end of one span, the north side span middle of one span, the south side span middle of one span, the south side beam end of one span, the north side beam end of the second span, the middle of two spans, the south side beam end of two spans, the middle of the main tower, and the bottom of the main tower, etc., respectively. From the table, it can be seen that although the bridge damage identification model constructed in this paper has the situation of misplaced damage identification, the overall model can reach more than 90% of the bridge damage identification rate, which fully demonstrates the good performance of this paper's model in the bridge damage identification task.

Table 4. Recognition results based on deflection damage index

Position	Damage degree/%	Recognition result	Recognition rate/%	Position	Damage degree/%	Recognition result	Recognition rate
1	0	1	100	6	20	6	100
	0	1	100		40	6	100
	0	1	100		60	6	100
	0	1	100		80	6	100
2	20	2	100	7	20	1	90.82
	40	2	100		40	7	100
	60	2	100		60	5	91.48
	80	2	100		80	8	93.56
3	20	3	100	8	20	1	92.38
	40	3	100		40	7	95.66
	60	3	100		60	8	100
	80	3	100		80	8	100
4	20	1	92.17	9	20	9	100
	40	4	100		40	9	100
	60	4	100		60	9	100
	80	4	100		80	9	100
5	20	1	92.36	10	20	10	100
	40	5	100		40	10	100
	60	5	100		60	10	100
	80	5	100		80	10	100

Table 5. Recognition results based on acceleration damage indicators

Position	Damage degree/%	Recognition result	Recognition rate/%	Position	Damage degree/%	Recognition result	Recognition rate
1	0	1	100	6	20	6	100
	0	1	100		40	6	100
	0	1	100		60	6	100
	0	2	99.38		80	2	91.64
2	20	1	93.74	7	20	7	100
	40	2	100		40	1	90.39
	60	6	90.62		60	7	100
	80	3	95.84		80	7	100
3	20	7	90.71	8	20	8	100
	40	3	100		40	2	91.26
	60	3	100		60	8	100
	80	3	100		80	8	100
4	20	4	100	9	20	9	100
	40	5	94.76		40	9	100
	60	4	100		60	9	100
	80	4	100		80	9	100
5	20	2	92.83	10	20	9	98.47
	40	3	93.46		40	10	100
	60	5	100		60	10	100
	80	1	90.17		80	10	100

Fig. 7 shows the comparison between the real label and the predicted label of the model, and (a) and (b) show the predicted recognition results of the model with deflection data and acceleration as the dataset, respectively. From the figure, it can be seen that the accuracy of damage recognition of the model trained using the two damage metrics is high. Combined with the data in Table 4 and Table 5, the credibility of the model for damage identification generally improves with the aggravation of the damage degree, and the differences between the real label and the predicted label are concentrated in the middle of the second span, the middle of the south span of the first span and the south beam end of the first span, while the error identification of the model damage recognition with acceleration as the damage index is mainly concentrated in the north beam end of the first span and the south beam end of the first span.

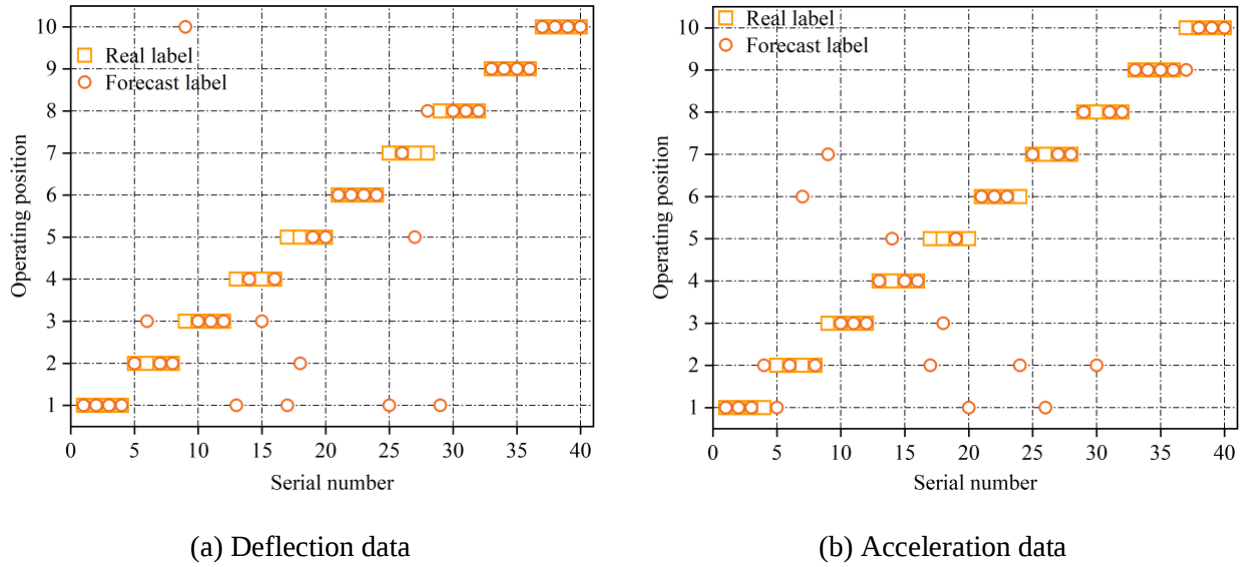


Figure 7. Comparison of the real label and forecast label

6 Conclusion

In this paper, the reconstruction and damage recognition task of bridge images collected by UAV based on a deep convolutional neural network is performed, and wavelet analysis and nonlinear transformation are used to enhance the image quality and improve the accuracy of image reconstruction and damage recognition during the task.

The reconstruction accuracy of the bridge image tends to stabilize at about 0.995 after the model has been trained more than 6000 times, while the reconstruction accuracies of the BP neural network, convolutional neural network, and feedforward neural network at this time are 0.798, 0.682, and 0.603, respectively. Compared with the reconstruction accuracies of this paper's model, the reconstruction accuracies of this paper's model are improved by 24.69%, 45.89%, and 65.01%, respectively. In the test set, the average reconstruction accuracy of this paper's model is improved by 19.20%, 28.24%, and 42.38% compared with the other three models, and the reconstruction time is reduced by 93.24%, 96.67%, and 97.77% compared with the other three algorithms. It shows that the bridge image reconstruction model constructed based on deep neural networks in this paper can significantly reduce the image reconstruction time and efficiently complete the bridge image reconstruction task under the premise of ensuring high reconstruction accuracy. At the same time, the bridge damage recognition model maintains the accuracy of damage recognition in 9 working condition locations, such as the girder end on the north side of the first span of the bridge and the north side of the first span. Through the identification of bridge damage by the model, it can accurately and quickly locate the bridge damage position, significantly improve the efficiency of UAV bridge inspection, shorten the time of damage investigation in bridge maintenance, and extend the service life of the bridge.

References

- [1] Ayele, Y. Z., Aliyari, M., Griffiths, D., & Droguett, E. L. (2020). Automatic crack segmentation for UAV-assisted bridge inspection. *Energies*, 13(23), 6250.
- [2] Peng, X., Zhong, X., Zhao, C., Chen, A., & Zhang, T. (2021). A UAV-based machine vision method for bridge crack recognition and width quantification through hybrid feature learning. *Construction and Building Materials*, 299, 123896.

- [3] Mandirola, M., Casarotti, C., Peloso, S., Lanese, I., Brunesi, E., & Senaldi, I. (2022). Use of UAS for damage inspection and assessment of bridge infrastructures. *International Journal of Disaster Risk Reduction*, 72, 102824.
- [4] Dorafshan, S., Maguire, M., Hoffer, N. V., & Coopmans, C. (2017, June). Challenges in bridge inspection using small unmanned aerial systems: Results and lessons learned. In *2017 international conference on unmanned aircraft systems (ICUAS)* (pp. 1722-1730). IEEE.
- [5] Escobar-Wolf, R., Oommen, T., Brooks, C. N., Dobson, R. J., & Ahlborn, T. M. (2018). Unmanned aerial vehicle (UAV)-based assessment of concrete bridge deck delamination using thermal and visible camera sensors: A preliminary analysis. *Research in Nondestructive Evaluation*, 29(4), 183-198.
- [6] Khaloo, A., Lattanzi, D., Cunningham, K., Dell'Andrea, R., & Riley, M. (2018). Unmanned aerial vehicle inspection of the Placer River Trail Bridge through image-based 3D modelling. *Structure and Infrastructure Engineering*, 14(1), 124-136.
- [7] Liu, Y. F., Nie, X., Fan, J. S., & Liu, X. G. (2020). Image-based crack assessment of bridge piers using unmanned aerial vehicles and three-dimensional scene reconstruction. *Computer-Aided Civil and Infrastructure Engineering*, 35(5), 511-529.
- [8] Peng, X., Zhong, X., Zhao, C., Chen, Y. F., & Zhang, T. (2020). The feasibility assessment study of bridge crack width recognition in images based on special inspection UAV. *Advances in Civil Engineering*, 2020(1), 8811649.
- [9] Kim, I. H., Jeon, H., Baek, S. C., Hong, W. H., & Jung, H. J. (2018). Application of crack identification techniques for an aging concrete bridge inspection using an unmanned aerial vehicle. *Sensors*, 18(6), 1881.
- [10] Perry, B. J., Guo, Y., Atadero, R., & van de Lindt, J. W. (2020). Streamlined bridge inspection system utilizing unmanned aerial vehicles (UAVs) and machine learning. *Measurement*, 164, 108048.
- [11] Ameli, Z., Aremanda, Y., Friess, W. A., & Landis, E. N. (2022). Impact of UAV hardware options on bridge inspection mission capabilities. *Drones*, 6(3), 64.
- [12] Feroz, S., & Abu Dabous, S. (2021). Uav-based remote sensing applications for bridge condition assessment. *Remote Sensing*, 13(9), 1809.
- [13] Erdenebat, D., & Waldmann, D. (2020). Application of the DAD method for damage localisation on an existing bridge structure using close-range UAV photogrammetry. *Engineering Structures*, 218, 110727.
- [14] Duque, L., Seo, J., & Wacker, J. (2018). Synthesis of unmanned aerial vehicle applications for infrastructures. *Journal of Performance of Constructed Facilities*, 32(4), 04018046.
- [15] Zhang, C., Zou, Y., Wang, F., del Rey Castillo, E., Dimyadi, J., & Chen, L. (2022). Towards fully automated unmanned aerial vehicle-enabled bridge inspection: Where are we at?. *Construction and Building Materials*, 347, 128543.
- [16] Chen, S., Laefer, D. F., Mangina, E., Zolanvari, S. I., & Byrne, J. (2019). UAV bridge inspection through evaluated 3D reconstructions. *Journal of Bridge Engineering*, 24(4), 05019001.
- [17] Chen, S., Truong-Hong, L., Laefer, D., & Mangina, E. (2018, August). Automated bridge deck evaluation through UAV derived point cloud. In *CERI-ITRN2018* (pp. 735-740).
- [18] Lei, B., Wang, N., Xu, P., & Song, G. (2018). New crack detection method for bridge inspection using UAV incorporating image processing. *Journal of Aerospace Engineering*, 31(5), 04018058.
- [19] Zollini, S., Alicandro, M., Dominici, D., Quaresima, R., & Giallonardo, M. (2020). UAV photogrammetry for concrete bridge inspection using object-based image analysis (OBIA). *Remote Sensing*, 12(19), 3180.
- [20] Li, J., Li, X., Liu, K., & Yao, Z. (2022). Crack identification for bridge structures using an unmanned aerial vehicle (UAV) incorporating image geometric correction. *Buildings*, 12(11), 1869.
- [21] Jeong, E., Seo, J., & Wacker, J. P. (2022). UAV-aided bridge inspection protocol through machine learning with improved visibility images. *Expert Systems with Applications*, 197, 116791.

- [22] Hackl, J., Adey, B. T., Woźniak, M., & Schümperlin, O. (2018). Use of unmanned aerial vehicle photogrammetry to obtain topographical information to improve bridge risk assessment. *Journal of infrastructure systems*, 24(1), 04017041.
- [23] Omar, T., & Nehdi, M. L. (2017). Remote sensing of concrete bridge decks using unmanned aerial vehicle infrared thermography. *Automation in Construction*, 83, 360-371.
- [24] Reagan, D., Sabato, A., & Niezrecki, C. (2018). Feasibility of using digital image correlation for unmanned aerial vehicle structural health monitoring of bridges. *Structural Health Monitoring*, 17(5), 1056-1072.
- [25] Xu, Y., & Zhang, J. (2022). UAV-based bridge geometric shape measurement using automatic bridge component detection and distributed multi-view reconstruction. *Automation in Construction*, 140, 104376.
- [26] Dorafshan, S., & Maguire, M. (2018). Bridge inspection: Human performance, unmanned aerial systems and automation. *Journal of Civil Structural Health Monitoring*, 8, 443-476.
- [27] Saleem, M. R., Park, J. W., Lee, J. H., Jung, H. J., & Sarwar, M. Z. (2021). Instant bridge visual inspection using an unmanned aerial vehicle by image capturing and geo-tagging system and deep convolutional neural network. *Structural Health Monitoring*, 20(4), 1760-1777.
- [28] Claude Rospars, Pierre Argoul & Raphaël Carpine. (2024). Wavelet analysis to detect nonstationary and nonlinear behaviours in railway bridge. *Procedia Structural Integrity* 716-723.
- [29] Syed Muhammad Waqas, Sumra Khan, Kashif Talpur & Rizwan Ahmed Khan. (2024). Density peaks clustering based on Gaussian fuzzy neighborhood with noise parameter. *Expert Systems With Applications (PC)*, 124782-124782.
- [30] Ananthakrishnan Balasundaram, Prem Sundaresan, Aryan Bhavsar, Mishti Mattu, Muthu Subash Kavitha & Ayesha Shaik. (2025). Tea leaf disease detection using segment anything model and deep convolutional neural networks. *Results in Engineering* 103784-103784.
- [31] Chen Junde, Zhang Defu, Suzaiddola Md & Zeb Adnan. (2021). Identifying crop diseases using attention embedded MobileNet-V2 model. *Applied Soft Computing Journal (PA)*.

About the Authors

Shun Wang (1984-), male, Han nationality, Heze, Shandong, M.C.E, Senior Engineer, main research directions: Road and bridge detection.

Mingwei Sun (1983-), male, Han nationality, Heze, Shandong, M.C.E, Senior Engineer, main research directions: Road management.