International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 1775-1791

https://doi.org/10.70517/ijhsa464155

Research on Automatic Identification Algorithm for Building Crack Defects Based on Convolutional Neural Networks

Xianyu Wang[1,*]
[1] International College, Chongqing University of Posts and Telecommunications, Chongqing, 400056, China
Corresponding authors: (e-mail: 13608120964@163.com).

Abstract The author uses the full convolutional neural network model for automatic identification of building crack defects. After processing the building crack defect data, the full convolutional neural network model is trained and optimized, and the automatic identification effect of the full convolutional neural network model in this paper is compared with other methods through experiments. In the calculation of building crack defects, image morphology is used to skeletonize the building cracks and calculate the crack length and width. The method is utilized to calculate the physical dimensions of building cracks and their static and dynamic distribution. The average training and validation accuracies of this paper's model at an initial learning rate of 1e-5 are 95.79% and 95.28%, respectively; the maximum training and validation accuracies are 81.15% and 77.77%, respectively; the maximum training and validation recalls are 80.81% and 80.48%, respectively; and the maximum training and validation F1 values are 84.19% and 79.11%, respectively. Its automatic identification effect is better than other methods. The relative error of maximum crack width between the model prediction results and real results in this paper is 2.1%~25.4%, and the relative error of crack length is 26.48%.The distribution ranges of static crack widths at three locations are between 1.3~3.4mm, 1.02~2.03mm, and 0.42~1.39mm, respectively. The dynamic crack width and area values showed an overall decreasing trend.

Index Terms fully convolutional neural network, image morphology, skeletonization, building cracks, automatic identification

I. Introduction

For the building, its surface is composed of concrete materials, and it is easy to form cracks in the constant changes of the external environment. The reasons for the formation of cracks in addition to floods, earthquakes and other natural disasters, mainly including construction technology and construction materials and temperature changes and other factors [1], [2]. Concrete is prone to structural deformation, and under the long-term erosion of rain and air, resulting in structural corrosion, reducing its load-bearing capacity, thus generating cracks [3], [4]. Prolonged overloading leads to uneven stress problems, which can also lead to cracks [5]. Small cracks can affect the aesthetics of the building structure and have a psychological impact on the occupants. Cracks with longer length and larger width will lead to the reduction of overall structural bearing capacity, deterioration of structural durability and seismic performance, shortening the service life of the building, and generating safety hazards [6]-[8]. Most of the cracks appear within 1-2 years of structural use, as analyzed by a large number of studies and observations [9], [10]. With the passage of time, these cracks will gradually accumulate, leading to a decline in concrete strength, the ability of the structure to resist damage from external loads is weakened, and in serious cases, it will lead to major injuries and deaths, resulting in unavoidable tragedies in order to safeguard the safety of people and property [11]-[14]. Therefore, the detection of crack defects on the surface of the building is of great significance in eliminating potential safety hazards.

As the focus of the building defects identification work, identification of its surface cracks can not only timely repair and fill defects to prevent the occurrence of construction hazards, to avoid causing aging or structural damage and other problems brought about by economic losses, but also test the quality of the building construction and construction materials, the overall quality of the project to monitor the results of the test of the building structure of the long-term dynamic safety assessment, and produce safety hazards of the building timely maintenance [15]-[18]. In this context, the problem of crack defect identification on building surfaces is studied. Crack defect identification can be said to be a very important part of crack detection, for the study of the problem of crack defect identification on the surface of the building, each country shows a very important attitude, which is due to the great significance of the problem for building maintenance, can effectively improve the overall level of building quality [19], [20]. Most of the traditional crack detection methods rely on manual visual inspection, which is not only inefficient, but also

easily affected by the subjective experience of the inspectors, resulting in poor accuracy of the detection results. In addition, manual inspection is usually accompanied by high labor and time costs, and working at high altitude or in hazardous environments increases the safety risk of inspectors. The identification of crack defects in buildings is currently evolving from traditional manual methods to intelligent identification methods.

Literature [21] developed an algorithm automatically used to identify exterior building cracks, mainly based on UAS (image collection) + deep learning (image processing and recognition) + photogrammetry (3D model construction to improve the clarity of the recognized image). Literature [22] describes an automatic building exterior crack inspection software, which is free and easy to use by applying UAV and smartphone camera to collect building projections, based on segmentation algorithms to realize concrete crack recognition, and has the function of generating quantitative and qualitative reports. Literature [23] provides a real-time building surface defect recognition application based on machine learning, 3D scanning techniques, Python programming language, PyTorch library, and Zed camera system, which not only recognizes crack defects and damages on the surface, but also distinguishes phenomena formed in natural environments such as shadows or stains. Literature [24] constructed an image preprocessing algorithm using improved bilateral filtering and created an image classification model based on the GhostNet algorithm to locate and identify building build crack defects in 2D pixel localization technique with 99% accuracy. Literature [25] automatically identifies the internal cracks of building components through ultrasonic sensors, combines GSM and GPS modules for crack localization, and can piece of localization information sent to the relevant personnel, and at the same time can realize the monitoring of cracks. Literature [26] uses supervised machine learning algorithms to automatically identify defects in granite masonry walls, realizing overall identification from individual defect identification, reducing the subjectivity and cost of manual identification. Literature [27] collected data from brick masonry walls with different colors, textures and sizes by hand, and performed automated brick segmentation and brick wall crack recognition with deep learning networks based on image processing techniques. Literature [28] integrated 2D automated phase expansion ultrasound technology and mobile machine vision to perform high-precision 2D spectral estimation of images of bridge surface cracks, respectively, via filtering and removing low-frequency crack information, to enhance the clarity of automated crack information recognition. Literature [29] for concrete surface crack recognition, designed an image segmentation method based on the Previtt operator and Osu threshold algorithm, image parameter de-contamination processing, combined with support vector machine for automatic image recognition. Literature [30] proposed an improved adaptive multi-head self-attention mechanism and coding module of U-Net to form a concrete surface crack recognition model, the recognition precision and accuracy were improved by 4.05% and 2.21%, respectively, and the crack width was analyzed in combination with orthogonal skeleton method. Literature [31] used UAV photogrammetry and digital twin technology for crack identification in cultural heritage buildings and inspected the structure, and achieved 83% recognition rate through crack data model construction and crack measurement accuracy analysis.

With the development of deep learning technology, Convolutional Neural Networks (CNN) are also increasingly used in the recognition of building cracks. In the field of image recognition, CNN can accurately identify various objects in an image by virtue of its powerful feature extraction ability, and can keenly capture the contour, texture and other key features of an object from a seemingly cluttered image, which plays a decisive role in deep learning to achieve accurate image classification [32]-[34]. Currently, scholars have established different types of CNN models for crack detection and achieved fruitful research results. Literature [35] formulated an automatic identification and classification process for concrete defects in aging buildings, and the whole process takes building images as input conditions and combines 3D modeling and CNN models for defect localization and identification, which improves the safety and effectiveness of identification. Literature [36] for self-service robotic inspector of the tunnel surface crack recognition for research and development, relying on the CNN model convolutional layer of the representation of the ability to deep CNN and heuristic image post-processing techniques were able to realize the crack recognition. Literature [37] constructed a deep CNN-based building crack analysis system, including classification and segmentation modules, and quantitatively analyzed the image segmentation results with image processing techniques, and the recognition accuracy was improved by nearly 6%. Literature [38] used an unmanned aerial vehicle to capture images of building facades, and combined the CNN model and the U-Net neural network model to classify and segment the captured images respectively, which reduced the image noise and thus improved the reliability and efficiency of crack recognition.

Similarly, literature [39] applied an unmanned aerial vehicle (UAV) to capture images of concrete cracks on building surfaces, and constructed multiple CNN models with the support of ResNet50 and YOLOv8, the former for classification (more than 99% on average) and the latter for target detection (85% on average), which have a high accuracy in identifying and localizing concrete cracks. Literature [40] designed a crack extraction algorithm to segment concrete bridge cracks and noise, combining CNN and plain Bayesian data fusion techniques to identify

the skeleton and continuous boundaries of cracks. Literature [41] applied a binary classification CNN model to automatically identify cracks in concrete structures and fused convolutional feature extraction layer, random forest and XGBoost to form a CNN model for automatic prediction of crack depth, which is an important reference for structural assessment and repair.

In this paper, convolutional neural network is applied to the automatic recognition of building cracks, firstly, the data of building surface cracks are collected and pre-processed, and the full convolutional neural network is selected to segment the building crack images end-to-end at pixel level. Secondly, the recognition effect of the automatic recognition model on building crack defects is improved by training and optimizing the full convolutional neural network. On the basis of image morphology calculation, the length and width of building cracks are calculated by skeletonizing the building crack defects. Finally, the method is applied to examples to calculate the physical dimensions of building cracks and analyze their crack static and dynamic distribution characteristics.

II. Automatic identification of crack defects on building surface based on convolutional neural network

II. A. Convolutional Neural Networks

Convolutional neural network (CNN) is one of the typical supervised deep learning algorithms [42], and its research originated in image processing. In the image domain, it is difficult to achieve good performance with ordinary fully connected neural networks. In order to reduce the number of free parameters, CNNs have been proposed and successfully applied.LeNet-5 is one of the earliest CNN models, which achieved superior classification scores in the MNIST image dataset and greatly contributed to the vigorous development of deep neural networks. Its general construction can be summarized as follows:

(1) In constructing the network, a convolutional layer, a pooling layer, and a fully connected network layer are used as the hidden layer.

(2) The convolutional layer extracts local information through convolution operation.

(3) The pooling layer improves the spatial robustness of the features and reduces the network complexity by downsampling (removing some unimportant features).

(4) Nonlinear mappings in the network generally use tanh, sigmoid, or Relu functions.

(5) The last layer (or layers) of the network often uses an ordinary fully connected neural network, which serves the purpose of classification.

Most current convolutional neural networks still use a similar network structure.

II. A. 1) Convolutional layers

The core idea behind a convolutional layer is that instead of a large, dense linear layer of neurons where there are connections between each input and each output, it consists of a number of very small linear layers, each of which typically has fewer than 25 inputs (related to the size of the convolutional kernel) and a single output, and which can be used in any position. Each small neuron layer is called a convolution kernel, but it is actually a very small linear layer that accepts a small number of inputs and serves as a single output.

Convolutional operations are an efficient strategy for extracting local features from an image. Different convolution kernels (templates for convolution operations, also called filters) can extract different local features, and the features extracted become more and more abstract (high-level features) as the number of network layers deepens.

The convolution operation contains three different modes: valid mode, same mode and full mode. There is no substantial difference between the different modes of convolution, except that the size of the output image after convolution is different, because the same mode and full mode supplement the pixel matrix of the input image with different levels of “0” in order to make the output image fit the expected size.

II. A. 2) Pooling layer

In convolutional neural networks, the most commonly used model is one or more convolutional computation layers followed by a pooling layer, which serves to reduce the size of the model, increase the inference speed of the network, and effectively prevent overfitting of the network structure. The pooling layer is also commonly referred to as downsampling, which is the process of sampling the sequence of features extracted from the convolution at regular intervals. In practice, it produces a set of hyperparameters, but the network does not need to learn any parameters as it is a fixed set of operations, which is why it brings high efficiency. In addition to this, such an operation creates some perturbation in the input image, which helps in improving the robustness of feature extraction.

In early convolutional neural networks, average pooling was more common and as the research progressed, it was found that maximum pooling brought better experimental results and hence, it became more popular to use

maximum pooling in current convolutional neural networks. An intuitive understanding of maximum pooling is that if a feature of the input data is extracted in the pooling window, its maximum value is retained as the output of the pooling layer, because the larger the value is, it means that the neural network has extracted some specific feature. If this feature is not extracted, it may not actually exist, then the maximum value in the pooling window is still very small, so using maximum pooling has a good feature differentiation effect.

II. A. 3) Activation function

Activation functions, also called transfer functions, are nonlinear functions that are being applied between convolution and pooling operations. Such a function is very important as it enhances the representation and classification ability of the network by accomplishing a nonlinear transformation of the input data. Early perceptron models lacked a nonlinear activation function, and although the number of network layers was increasing, the output values were nothing more than linear combinations of the input data, and thus their data representation ability was limited and they were unable to simulate nonlinear mapping relationships of complex features. In addition, as the number of network layers deepens, the linear transformation causes the input data to expand, the network cannot achieve convergence, and the output values become meaningless. Therefore the direct effect of the activation function is to compress the input data from an infinite space to a very small space, such as between 0 and 1, at which point the output of a single neuron can be interpreted as a probability, thus accomplishing the prediction of the input data.

For a nonlinear function to be used as an activation function, several constraints need to be met:

(1) The domain of definition is infinite and the function must be continuous, because the activation function must produce a definite output for any number, otherwise the neurons in the network will “die” (fail badly).

(2) The function is monotonic as far as possible and cannot change direction, meaning that the activation function cannot output the same value for different inputs. This is because when a neural network learns, it is actually finding the correct configuration of weights for a particular output. If there is more than one correct answer, the output becomes very difficult.

(3) The activation function and its derivatives can be computed efficiently. Because the activation function needs to be called repeatedly, it is not desired to compute it too slowly when faced with hundreds of millions of network parameters. Currently, the common activation functions that can satisfy these constraints are Sigmoid, Relu, and Tanh. Among them, the Relu function is the most frequently used one, and its expression is shown in (1):

$$f(x) = \max(0, x) \tag{1}$$

Compared to Sigmoid and Tanh, Relu eschews complex exponential operations and is more efficient in its calculations.

II. A. 4) Full connectivity layer

The fully connected layer is located at the end of the convolutional neural network, the actual operation, it will be the output of the network's feature map spread into a number of one-dimensional vector.

This operation causes each feature map to be highly condensed into each number, and finally these numbers are provided to a classifier, which weights and sums them according to the importance of the features in each dimension, and the resulting output is the predicted probability for the same target (e.g., cats and dogs). This classifier is the common Softmax as shown in equation (2):

$$P_j = \frac{e^{x_j}}{\sum_{j=1}^{n} e^{x_j}} \tag{2}$$

where P_j is the regression function for class j in the model, representing the probability of occurrence of class j, and n is the number of classes of training samples in the neural network. In this way the model completes the classification of the input image.

II. B.Data processing of crack defects on building surfaces

II. B. 1) Building surface crack image data acquisition

Building surface crack defect recognition through deep learning first requires the acquisition of image data of crack defects on the building surface, and then the sampling and quantification of the acquired images are carried out in order to realize the digitization of the defect images.

The acquisition of image data of crack defects is mainly carried out in three aspects, the collection of crack defect

images by manual field acquisition, the collection of crack defect images in the laboratory, and the collection of crack defect images by drones.

Among them, the manual field acquisition method is simpler, and the images are mainly acquired manually, using a PA line scan camera as the tool.

When image acquisition is carried out in the laboratory, the main purpose is to collect image data from experiments related to the flexural resistance of concrete beams in buildings, and the acquisition tool is a SLR digital camera, model Canon EOS T08.

When the images of crack defects were acquired by UAV, the model of the vehicle used was RTK M012, on which an aerial camera was paired with the specific model XS1-4.

In order to make the training model generalization ability and robust ability stronger, the image diversity needs to be fully considered when the defect images are acquired. Some of the images are acquired at a distance of 0.5 m or less, and the remaining images are acquired at a distance of 0.5-1 m. Shadows, light and other influencing factors are taken into account during acquisition.

A total of 300 images with a pixel resolution of 4806 × 3654 were acquired by manual field acquisition, 120 images with a resolution of 2253 × 1865 were acquired in the laboratory, and 360 images with a pixel resolution of 4 468 × 3846 were acquired by drone acquisition.700 of the acquired images were used in model testing, validation, and training process, and the remaining images were used to implement secondary testing. In order to increase the number of data in the overall dataset and to achieve higher recognition accuracy of the model, image cropping was performed on the collected images to achieve a pixel resolution of 150×150, and the defective images were automatically cropped by the window sliding segmentation algorithm in Python. Then some of the background images that do not contain crack defects and concrete are eliminated, and finally the images are labeled with crack defect-free images and crack defect images as a database for testing, verification, and training. A total of 20000 images are available in the constructed crack defect image database. Random selection is performed among them to generate the test set, validation set, and training set. This step of cropping ensures that the database is richer in data and contains a wider range of image types, such as large crack defects, small crack defects, edge cracks, distorted cracks, and crack-free defect images such as bumps, textures, and scratches.

In order to better understand the building surface crack image data, the dimensionality reduction visualization of the dataset is implemented through t-distributed neighborhood random embedding. Sampling and quantization of the acquired images, the specific process of digitization processing is shown in Figure 1.

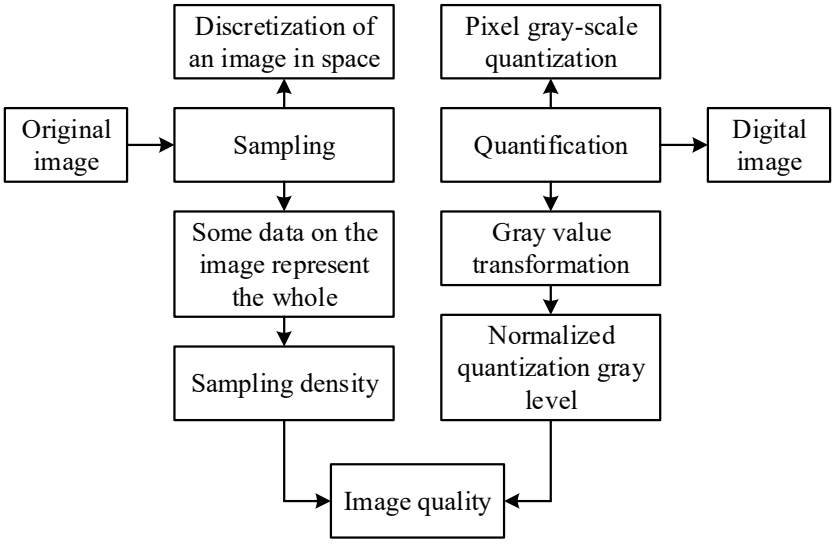


Figure 1: Specific process of digital processing

II. B. 2) Pre-processing of building surface crack image data

For the image data obtained after the digitization process, it is also necessary to pre-process it, and the specific steps include threshold segmentation processing, filtering processing and enhancement processing.

(1) Threshold segmentation processing

Threshold segmentation processing uses the OSTU automatic segmentation method [43], which belongs to the parameter automatic selection of global segmentation, mainly using the image gray scale statistics to obtain the gray scale histogram, and to maximize the interclass variance objective of the threshold to solve the threshold, the

threshold as a criterion to classify the image pixels, which are divided into two types: background and foreground.

Firstly, the actual gray level expectation of the first type of pixel point is calculated, i.e.:

$$P_1 = \sum_{i=0}^{k} p_i \tag{3}$$

where, P_i represents the actual gray level expectation of the pixel point of the 1st kind, p_i represents the actual frequency of occurrence of the pixel whose gray level is i in the image, k represents a certain gray level threshold.

And the actual gray level expectation of the 2nd kind of pixel point is calculated, i.e:

$$P_2 = \sum_{i=k+1}^{L-1} p_i \tag{4}$$

where P_2 denotes the actual grayscale expectation of the 2nd kind of pixel point, and L denotes the corresponding number of gray levels in the grayscale space.

Then the actual weighted expectation of the pixel point of the 1st kind is calculated:

$$m_k = \sum_{i=0}^{k} i \cdot p_i \tag{5}$$

and calculate the corresponding weighted expectation of all pixel points in the image:

$$m_G = \sum_{i=0}^{L-1} i \cdot p_i \tag{6}$$

Finally, the actual interclass variance of the two pixels is calculated, i.e:

$$\sigma_B^2 = P_1 \cdot P_2 \cdot (2m_k - m_G)^2 \tag{7}$$

where σ_B^2 denotes the actual interclass variance of the two pixels.

(2) Filtering process

The filter selected for filtering processing is LPF filter, and the specific filtering process is as follows:

The expression of the filter function is specified as follows:

$$H(u,v) = e^{-\frac{D(x,y)^2}{2D_o^2}} \tag{8}$$

where $H(u,v)$ represents the filter function, D_o represents the cutoff frequency, and $D(x,y)$ represents the frequency domain coordinates.

The expression for the frequency domain function is specified as follows:

$$D(u,v) = u^2 + v^2 \tag{9}$$

where u,v represents the left and right coordinates of the frequency domain pixels.

The post-frequency domain filter function is computed according to the above equation:

$$F'(u,v) = H(u,v) \cdot F(u,v) \tag{10}$$

where $F'(u,v)$ represents the post-frequency domain filtering function.

Finally the corresponding function after the original spatial domain filtering is calculated:

$$f'(x,y) = \frac{1}{M \cdot N} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F'(u,v) e^{i2\pi(\frac{ux}{M} + \frac{vy}{N})} \tag{11}$$

where $f'(x,y)$ represents the corresponding function after filtering in the original spatial domain and M,N represents the filtering threshold.

(3) Enhancement Processing

The enhancement processing uses the adaptive enhancement method, and the specific processing steps are as follows:

Firstly, the total value of the expected gray level of the crack image is calculated, and the specific calculation formula is as follows:

$$\mu = \sum_{i=-\infty}^{\infty} p_i \sum_{L-1=0}^{L-1} p_{L-1} \tag{12}$$

where, μ represents the total value of the desired gray level of the crack image.

Finally the adaptive enhancement of the image is implemented as follows:

$$g(x,y) = \frac{L-1}{f'_{\max} - f'_{\min}} \cdot (f'(x,y) - f'_{\min}) \tag{13}$$

where $g(x,y)$ denotes the coordinates of the image after adaptive enhancement, $f'_{\max}$ denotes the maximum value of the corresponding function after filtering in the original spatial domain, and $f'_{\min}$ denotes the minimum value of the corresponding function after filtering in the original spatial domain.

II. C.Fully Convolutional Neural Networks

In 2015, Long et al. proposed a pixel-level segmentation method that can realize end-to-end segmentation, called full convolutional neural network (FCN), compared with the traditional semantic segmentation method of convolutional neural network, the input of full convolutional neural network can be an image of arbitrary size, which can reduce the complex process of image preprocessing, and is more conducive to the transformation of deep learning results and application in real life.

II. C. 1) FCN Network Architecture

FCN network improves the original CNN network [44], the first 5 convolutional layers after the image input are the same as CNN, but the fully connected layers of the model are discarded in favor of a convolutional layer, after the convolutional layer the output is still outputted as a two-dimensional digital image matrix, and then the featured image is restored to the input image size by the bilinear interpolation method, and finally the classification of individual pixels is computed to obtain the segmentation results.

The three most significant features of FCN are: full convolutionalization, upsampling and jump structure.

(1) Full Convolutionalization

Fully convolutionalization means that all fully connected layers in the network structure are replaced by convolutional layers, and the output is changed from a one-dimensional vector to a two-dimensional matrix, and the feature image is reduced to the size of the input image through the inverse convolutional layer.

(2) Up-sampling

The upsampling operation is also known as deconvolution. Inverse convolution is not a simple inverse operation of convolution, but more accurately, it should be said that the upsampling is a transposed convolution. In order to achieve end-to-end pixel-level prediction, the size of the feature mapping is restored to the original input image size by up-sampling. Inverse convolution is a simple and reliable up-sampling method to convert small sparse matrices into large dense matrices.FCN uses bilinear interpolation, which is easy to implement and computationally small but tends to ignore some details. The output size of upsampling can be calculated by the following formula. The specific formula is shown in (14):

$$o = s(i-1) + k - 2p \tag{14}$$

where, s denotes the step size, k is the convolution kernel size and p is the image filling method.

(3) Jump structure

In the convolutional neural network, the network output resolution of the gradual layer is higher, retaining more detailed information, but with the increase of the number of layers, the deeper layers of the network learning features become more abstract, the resolution of the output feature maps is low, and the details are lost every time the downsampling operation is performed, which will result in the consequences of poor segmentation effect and low accuracy. The jump structure is to fuse the feature maps of different layers, effectively combining the information of the shallow network with that of the deep network, and then carry out the anti-convolution operation for optimization, which solves the problem of information loss to a certain extent.

II. C. 2) UNet network structure

UNet model is a full convolutional neural network improved on the basis of FCN, which has the features of small samples, fast detection and good segmentation, and was widely used in medicine at the beginning of its invention.The UNet shape is clean and refreshing, and happens to resemble the shape of a U, so it is called UNet network structure, which consists of left and right halves, with compression channel in the left half, and expansion channel in the right half. The compression channel is the most common convolutional neural network structure, which utilizes a structure consisting of two convolutional layers and a maximum pooling layer for downsampling four times, and after maximum pooling, the feature map doubles in dimension. In the expansion channel, one upsampling operation is performed first, where the dimension of the feature map is reduced by a factor of two, and then it is spliced with the cropped feature map of the corresponding compression channel, which is completed to obtain a feature map that is twice as large as the previous one, and then it continues to use the two convolutional layers for feature extraction and repeats the operation of the above structure. Finally, for the output layer, two convolutional layers are used to transform the 64-dimensional feature map into a 2-dimensional output map.

The structural model of UNet network follows the idea of FCN semantic segmentation, reading features through convolutional and pooling layers, and then reducing the image to the input size through the up-sampling layer. Moreover, UNet skillfully integrates the advantages of encoding-decoding structure and hopping network, which makes the model structure more perfect.

II. C. 3) Improved UNet network architecture

Semantic segmentation is one of the important research directions in the field of target detection, which enables pixel-level target detection. Semantic segmentation consists of encoder and decoder. The encoder utilizes a convolutional layer to input the features of the image, and the pooling layer approach reduces the size and speed of the feature reading and reduces the burden of network computation. The decoder, on the other hand, utilizes the inverse convolution method to restore the feature image to the input image size and predict the result.

In this paper, Res UNet is applied to bridge crack recognition for the first time. Compared with traditional neural networks, the biggest difference of ResNet is that the input is added to the output, which has the same effect as fusing the feature information of the lowest layer to the upper layer. When the network training layer is deeper, it also means that more features can be extracted and better express the image semantics, but at the same time, it will face serious gradient disappearance and network degradation problems. The addition of the residual block perfectly solves this problem network is difficult to optimize. The traditional module and the residual module are shown in Fig. 2, compared with the traditional module, the residual module adds a constant shortcut link. Where $H(x)$ is the optimal mapping, $F(x)$ is the residual mapping, and $H(x)=F(x)+x$. This makes it easier to learn $F(x)$ by switching from learning of $H(x)$ to learning of $F(x)$.

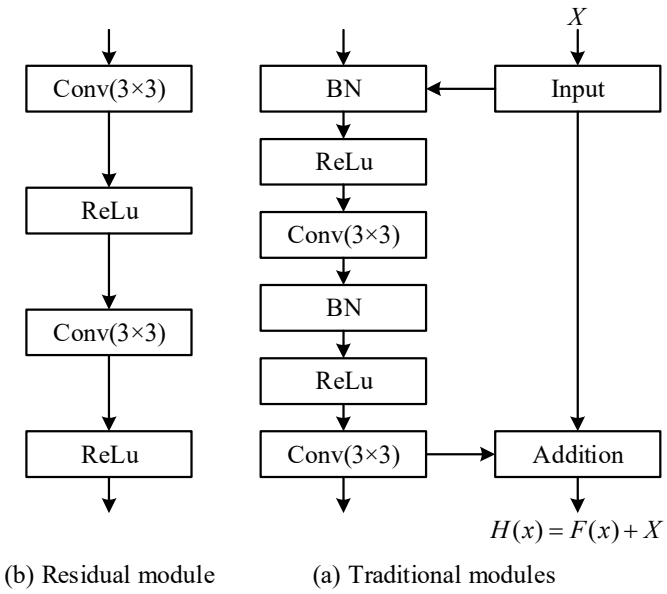


Figure 2: Traditional module and residual module

In this paper, we have utilized to incorporate Residual module (residual module) into UNet network and to identify

and detect bridge cracks. The crack pixels and background pixels are segmented from the input image.Res UNet network is a semantic segmentation network model based on Res Net and UNet, and its structure type is largely similar to UNet.Res UNet network consists of three parts: encoder, connector and decoder. The decoder is composed of 3 residual modules and its main role is to decode the input image and read the features, which is the process of downsampling. Connector is applied to connect the information propagation path between encoder and decoder. Decoder is applied to the process of image recovery, the principle is to classify each pixel in the image, mainly consists of 3 decoding modules, each decoding module includes batch normalization, ReLu layer and inverse convolution layer. The inclusion of residual networks into the UNet model effectively overcomes the problems of overfitting, parameter redundancy, and depth model degradation due to the excessive number of network layers. Moreover, due to the inclusion of residual blocks, the training speed is greatly improved, which also enables the network to ensure high loss accuracy while preserving fewer parameters. The network structure is shown in Fig. 3.

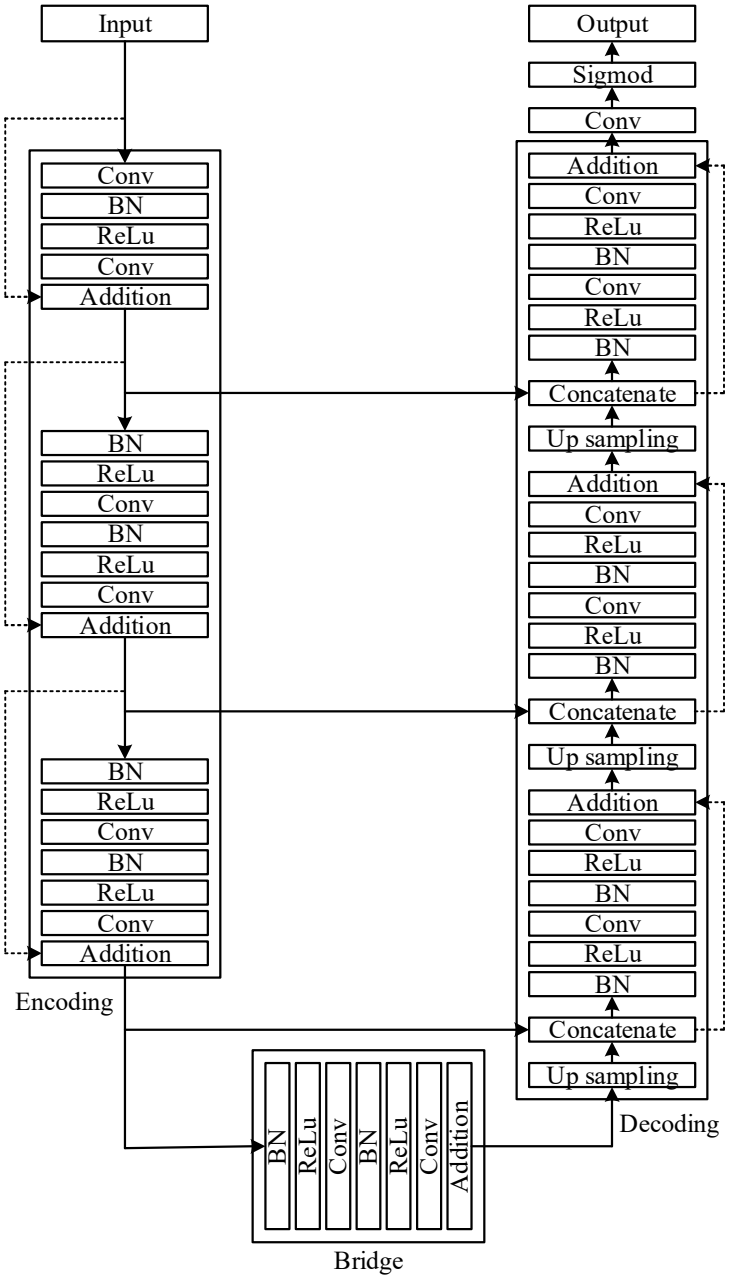


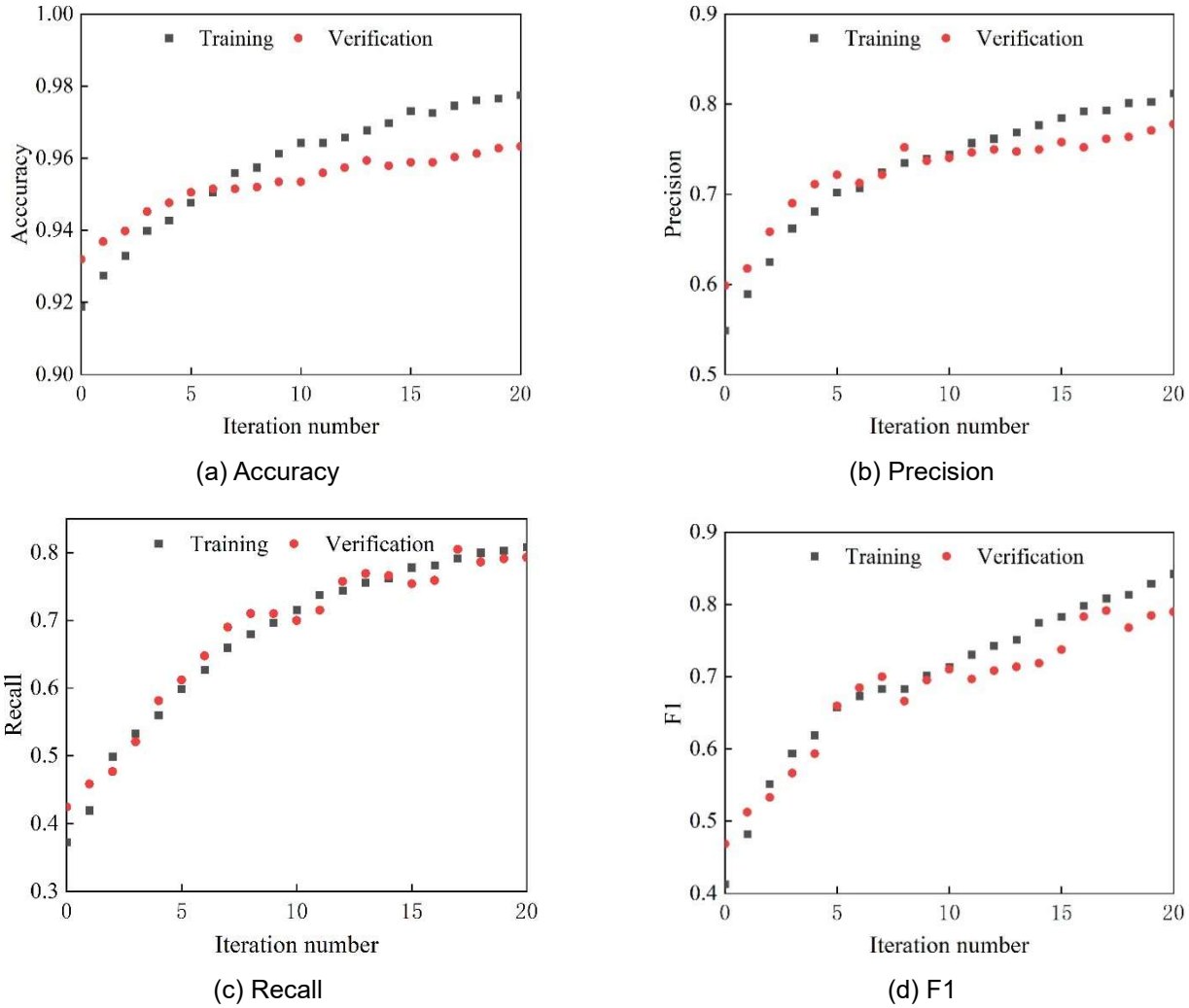
Figure 3: Res Net structure

III. Improved UNet network for automatic identification of building crack defects

The training process was performed on a workstation with a high-performance GPU (Geforce GTX 1080TI) and CPU (Intel Core i7, 3.60 GHz 8), where the GPU is capable of massively parallel computation and the CPU is used for network model updating. The code was written in Python 3.5 and the virtual environment was built with TensorFlow 1.4 and Numpy 1.14, a configuration that ensures that the training process can be executed on both Windows and Linux systems.

III. A. Full Convolutional Neural Network Model Training

The full convolutional neural network model designed in this paper, when the initial learning rate is 1e-5, the network is shown in Fig. 4(a) for the training and validation sets, the variation of accuracy in each time period, the average of accuracy in the training phase is 95.79% and the average of accuracy in the validation phase is 95.28%. When the initial learning rate is 1e-5, in Figs. 4(b), (c), and (d), the maximum precision, recall, and F1 score in the training phase are 81.15%, 80.81%, and 84.19%, respectively. The maximum precision, recall and F1 in the validation phase are 77.77%, 80.48% and 79.11%, respectively. Therefore, in practical applications, the initial learning rate can be set to 10^{-4} for crack identification, and although the full convolutional neural network algorithm in this paper performs better than random structure forest and convolutional neural network modeling, its performance is not as good as that of CrackNet with precision, recall, and F1 scores of 90.56%, 87.96%, and 88.95%, respectively The results are good. The gap between the training and validation data can be seen in Fig. 4 and it is found that overfitting occurs after 15 times of training with all the samples in the training set, which can be caused by fewer types and number of cracks in the dataset.



(a) Accuracy (b) Precision

(c) Recall (d) F1

Figure 4: Evaluation index

Although the accuracy of the full convolutional neural network algorithm in this paper did not reach the expected maximum, the training time was significantly shorter. Compared to CrackNet's 9-day training time, after 600 iterations, the full convolutional neural network in this paper completed 1200 iterations in less than an hour.

III. B. Comparison of experimental results for crack identification

For the images of building cracks collected as training samples in equally complex environments, the accuracy of detection can reach 98.26%, and the experimental structure of the precision, recall and F1 score of its trained network model on the validation set is shown in Table 1. The evaluation index of this paper's full convolutional neural network algorithm is similar to the results of other three classical algorithms Random Structure Forest, CNN and CrackNet for crack detection, but the advantage of this paper's full convolutional neural network algorithm model on the one hand does not have to require all the training images and test images to have the same size. On the other hand it is more efficient and does not need to classify each image block, which can significantly reduce the training time.

Table 1: Crack identification comparison

Identification algorithm	Precision	Recall	F1
Canny	0.452	0.734	0.482
Random structure forest	0.763	0.864	0.853
CNN	0.806	0.857	0.833
CrackNet	0.911	0.889	0.878
Ours	0.857	0.809	0.813

IV. Calculation of crack characteristics for images of building walls

IV. A. Image morphological operations

A common processing method used in crack recognition is the mathematical morphology operation. Mathematical morphology is similar to convolutional neural network and also has a filter in convolutional neural network but in mathematical morphology it is called structural element which is the basic unit of traversal scanning image processing. In mathematical morphology, the point is able to obtain specific image information by moving through the image and performing pixel calculations. Depending on the different purposes of our image operations, we can choose structural elements of different colors, sizes and shapes for detection.

In preliminary image processing, we most often use expansion and erosion operations. With the research scholars it is found that the combination of two operations in sequential order will be more effective when used in combination, i.e., open and closed operations.

(1) Erosion operation

The definition of corrosion operation is to use the structure element to move and judge in the target image, if it meets the requirements, the point will be included in the output results, and then slide scanning the whole map to find out all the pixel points to meet the requirements of the collection of pixel points that is the result obtained by the corrosion process. The essence is to find out the pixel position points of the original image with the same sliding structure.

The specific process of erosion operation is to move the structure element B on X along axes x and y respectively, which is similar to the convolution process. When the structure element B is moved to pixel (x, y) in the image X corresponding to the origin, the pixel (x, y) is assigned a value of 1 if B is contained in the image X at point (x, y), otherwise it is 0. The specific expression is as follows:

$$X \Theta B = \{x, y \mid (B)_{xy} \in X\} \tag{15}$$

where, Θ represents the corrosion operation.

In the crack recognition study, the corrosion operation will reduce the range boundary of the crack region, and the tiny points in the image will be reduced to none, so as to eliminate the tiny points. In practice, we can find that after the corrosion operation, although the fine objects and points are eliminated, the area of the crack itself is also reduced, so the expansion operation is also needed to restore the crack area.

(2) Expansion operation

The definition of the expansion operation is similar to that of the corrosion operation [45], the difference is that the judgment criterion of the expansion operation is that as long as a region in the original image contains a structural element, then the origin position of the structural element will be marked, and the sliding scanning of the whole image to find out all the points that meet the requirements of the set of points that is the result of the expansion

operation. The essence is to find out the points that contain structural elements in the original image.

The specific process of the expansion operation is to move the structure element B on X along axes x and y, respectively, and again, this moving process is similar to the convolution process. When the structure element B is moved to pixel element (x, y) in the image X corresponding to the origin, if B overlaps with image A at point (x, y), then the pixel element (x, y) is assigned a value of 1, otherwise it is 0. The specific expression is as follows:

$$X \oplus B = \{x, y \mid (B)_{xy} \cap X \neq \varnothing\} \tag{16}$$

where $\oplus$ represents the expansion operation.

It can be seen that the erosion operation is equivalent to a contraction of the boundary of the crack region, while the expansion operation is equivalent to an expansion of the boundary of the crack region. The expansion operation is similar to the erosion operation in that it expands the crack region by doing a sliding traversal of each pixel in the image using a numerical matrix. The expansion operation can reconnect the disconnected regions caused by the prediction error of the neural network, so that the whole region has more continuity, which can be used to repair the fracture error caused by the crack segmentation process. Similar to the corrosion operation, the crack region will be enlarged and distorted after the expansion operation, so it needs to be used together with the corrosion operation.

(3) Open and close operation

The open operation is to rot the candle first, and then expand.

Open operation can make the image contour becomes smooth, does not change the crack area and can remove the crack surface protrusions and burrs, but also can eliminate the image of small dots and disconnect the narrow connectivity domain, its specific mathematical expressions are as follows:

$$A \circ B = (A \Theta B) \oplus B \tag{17}$$

The closed operation is to expand and then erode.

Closed operation is able to fill the holes in the cracked regions of the original image, and also has the function of smoothing the junction of cracked and non-cracked regions as well as connecting two non-connected regions. The mathematical expression for the closure operation is:

$$A \cdot B = (A \oplus B) \Theta B \tag{18}$$

In this study of crack recognition, the open and close operation can effectively improve the task of analyzing the crack segmentation map with the comparison between the original segmentation image of the crack and the open and close operation.

IV. B. Skeletonization for crack length and width

IV. B. 1) Morphological skeletonization of crack images

The purpose of skeletonizing cracks is to transform the cracks at the pixel level to provide a clear visualization of single pixel width cracks and their topology, and the acquired crack skeleton can be used as an important indicator for structural health monitoring of bridges. Skeletonization methods for images include the mid-axis method, the Hilditch algorithm, and the axis refinement algorithm. In this study, the central axis method can accurately remove the boundary of each crack, and the inspectors can use the central axis algorithm to quickly and accurately extract the crack skeleton.

Skeletonization is a commonly used method in image processing [46], which can be used as an algorithm for describing the key structure of a crack image and to realize the skeleton extraction of cracks by turning multi-pixel-width cracks into cracks with only a unit pixel width. The skeletonization of crack images can highlight the crack features more effectively as well as facilitate the calculation of information such as the length, width, morphology, and direction of the cracks.

IV. B. 2) Calculation of crack length and width

Full Convolutional Neural Network Prediction Image After the skeletonization is completed, the essence of the image is actually a set of two-dimensional arrays in which the numerical matrix pixel points are able to express the information about the horizontal and vertical coordinates of the skeleton of the crack image, $i = 0, 1, 2, ..., n$.

On this basis, the specific calculation process of the length, maximum width, and average width of the crack in the image crack recognition analysis is shown below:

(1) Length of crack

In the calculation of crack length, the length direction of the crack skeleton pixel in the continuous region is summed up, and the final result is the length of the crack. Assuming that the length of the crack is l, the formula

for calculating the crack length is shown in equation (19):

$$l = \sum_{i=1}^{n} \sqrt[2]{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} \tag{19}$$

(2) Maximum width of the crack

And the width of the crack is the vertical distance between the two boundaries of the predicted crack image before skeletonization.

Any point in a crack image can be represented as a matrix pixel coordinate point, $i = 0,1,2,...,n$, $k = 0,1,2,...,n$. Let us assume that the boundary pixel coordinates of the lower edge of the crack in column k are $(i, Z(i,k))$, then the width of the vertical crack pixel located in that column is $P(k)$. The formula for $P(k)$ is as follows:

$$P(k) = Z(i,k) - Z(j,k) \tag{20}$$

At the same time we set $M(k) = \frac{z(i,k) + z(j,k)}{2}$ and the crack inclination is $\theta(k)$, then $\theta(k)$ is calculated as follows:

$$\theta(k) = \arctan(\frac{M(k+1) - M(k-1)}{2}) \tag{21}$$

Then the pixel width $W(k)$ of the crack is:

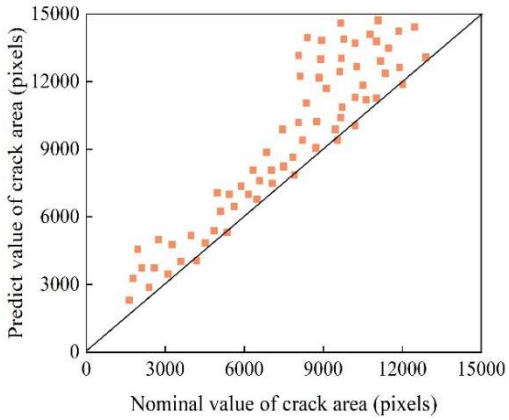
$$W(k) = P(k) \times \cos(\theta(k)) \tag{22}$$

Finally, just take the maximum value among all the obtained crack pixel widths.

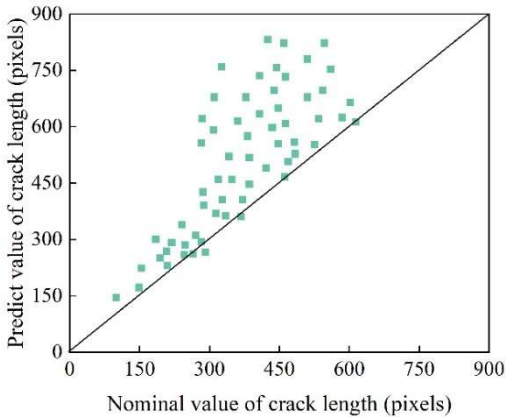
V. Quantitative crack analysis

V. A. Calculation of physical dimensions of cracks

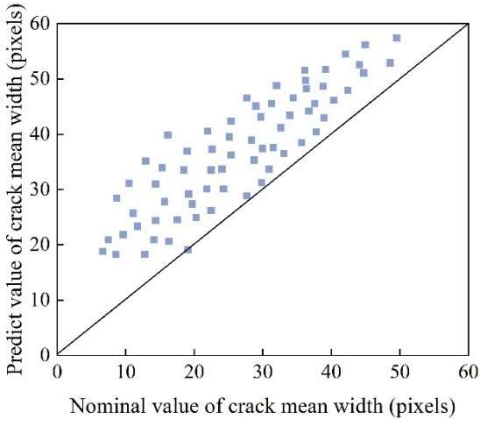
In order to test the performance of the algorithm in this paper, a skeleton is generated based on the framework proposed in this paper and four metrics, namely, crack length, area, average width and maximum width, are quantified using 100 detected images. The comparison between the predicted and real results is shown in Fig. 5. It is evident from the distribution of points along the centerline. In Fig. 5(d), it can be seen that the crack maximum width measurements are roughly distributed around the centerline, and the corresponding relative errors are in the range of 2.1% to 25.4%, indicating that the test results can meet the expectations. From Fig. 5(a)(b), it can be seen that the accuracy of crack area and length is generally large, 68% of the measured crack area is higher than the area of the labeled cracks, 55% of the measured crack length is higher than the length of the labeled cracks, and the relative error of the crack length is 26.48%. In Fig. 5(c), the distribution of crack mean width is more dispersed relative to the other parameters. Since the crack mean width is calculated by dividing the crack area by the crack length, this situation proves the measurement bias of the crack mean width, i.e., a certain crack width may be calculated from different area and length values, resulting in an uneven distribution in the end.



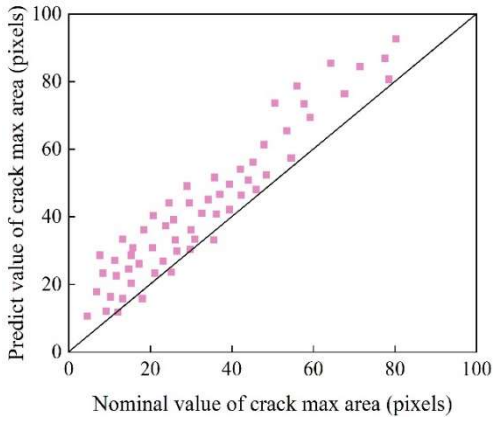
(a) Crack area



(b) Crack length



(c) Crack mean width



(d) Crack max area

Figure 5: Comparison of crack quantifications at a pixel level

V. B. Analysis of experimental results for crack quantification

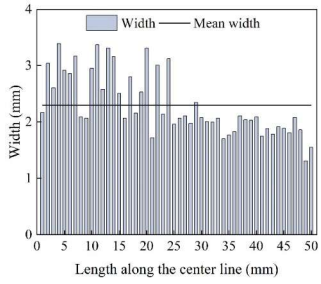
V. B. 1) Static distribution of cracks

The pixel size of the image for this experiment is 4112×3008 pixels, and the camera is used to take a picture of the main crack area of the final damage pattern of the concrete beam. However, since the original crack pixels are too large for deep learning to recognize accurately, they need to be cropped before quantitative analysis. In this experiment, Photoshop software was used to crop the photographed images one by one according to the crack region using a 500×500 pixel frame, so that most of the crack regions were included. The final damage surface of the main crack region A was cropped into 4 parts labeled as A-1, A-2, and A-3.

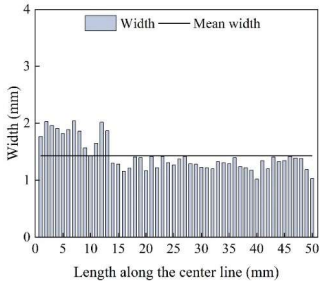
In order to verify the accuracy of the trained full convolutional neural network model of this paper to output the crack quantization data, since the pixels of the cropped image are 500×500 pixels and the area of the crack on the surface of the close-up beam is 50mm×50mm. therefore, the correlation between the pixel value and the actual size is established, which yields that a single pixel represents the value of the actual length of 0.1mm, so as to establish the link between the pixel value and the actual size.A-1~A -3 The quantization results are shown in Fig. 6, and the results of the quantized and measured values of the characteristic parameters are shown in Table 2.

From Fig. 6 and Table 2, it can be seen that when the concrete beam reaches the ultimate load damage, the width distribution of Fig. A-1, which is nearest to the crack initiation point, ranges from 1.3 to 3.4 mm, and the widths of A-2 and A-3 cracks are distributed between 1.02 and 2.03 mm and 0.42 and 1.39 mm, respectively. The surface cracks in the main crack region decreased in parameters such as area, average width, and maximum width from the crack opening to the crack boundary. Although the internal width distribution of each crack is complex, there is an overall decreasing trend, and the crack width gradually decreases along the crack extension direction, indicating that the more complex the crack distribution is along the crack extension direction.

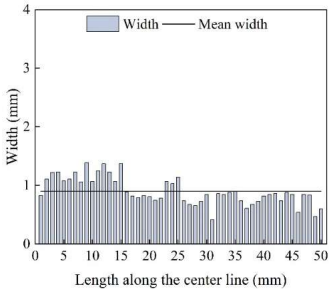
When the experiment is finished, the characteristic parameters of the above crack areas are measured by the crack measuring device respectively. By comparing the results with those obtained after deep learning segmentation, it is found that the error values of crack area and length are gradually reduced, mainly because the orientation angle of the cracks along the extended area is reduced and the crack extension tends to be stabilized. For the average and maximum widths, the accuracy is high.



(a) A-1



(b) A-2



(c) A-3

Figure 6: Length along centerline

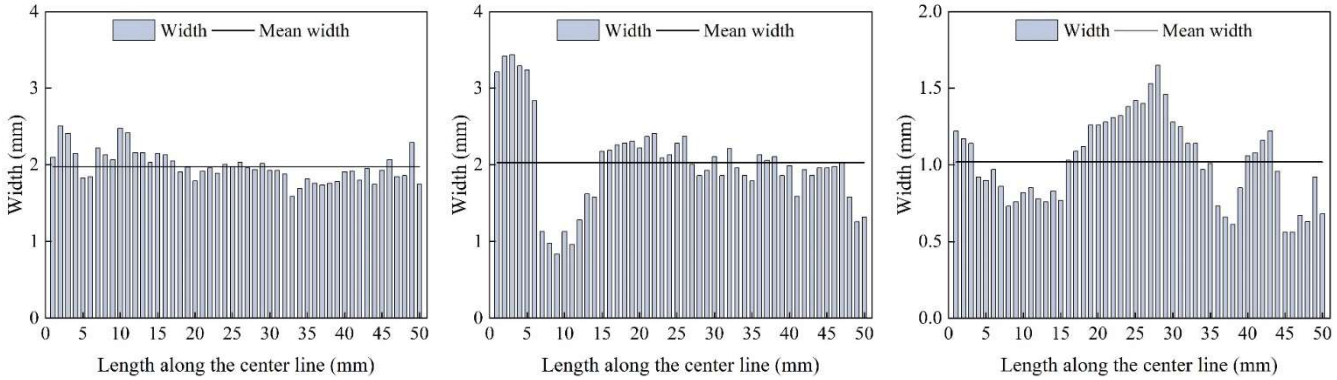
Table 2: Comparison of quantitative parameters and measured values of fractures

Quantitative index		Fracture number		
		A-1	A-2	A-3
Area (mm^2)	Model segmentation	97.8	77.4	43.2
	Tool measurement	105.6	79.5	44.1
	Error (%)	-7.4	-2.6	-2.0
Length (mm)	Model segmentation	45.3	42.2	39.6
	Tool measurement	43.2	40.8	38.5
	Error (%)	4.9	3.4	2.9
Orientation Angle (°)	Model segmentation	14.6	9.5	8.5
	Tool measurement	13.7	9.1	8.9
	Error (%)	6.6	4.4	-4.5
Mean width (mm)	Model segmentation	3.3	2.1	0.7
	Tool measurement	3.1	2.1	0.7
	Error (%)	6.5	0	0
Max width (mm)	Model segmentation	2.4	1.6	0.6
	Tool measurement	2.2	1.6	0.6
	Error (%)	9.1	0	0

V. B. 2) Dynamic distribution of cracks

When the load gradually reaches the ultimate load this process is actually the dynamic process of cracks from nothing to something, for the cracks from cracking to cracks gradually become wider to the concrete surface spalling, the experimental beams completely destroyed. For the B main crack region, it is also cropped into three sheets of 500×500 pixels, named B-1, B-2 and B-3, respectively, and the quantitative result map is shown in Fig. 7, and the statistical values of each characteristic parameter value are shown in Table 3.

The crack width and area values show an overall decreasing trend, and when the length range along the centerline is between 40 and 50 mm, both the area and width values are significantly higher than in other areas. It is only due to the occurrence of localized spalling phenomenon that the area and width values appear to increase. With the direction of crack expansion, its parameters also appear to gradually reduce the trend, when the length reaches 100mm, at this time the cracks are fine cracks, the width are less than 1mm.



(a) B-1 (b) B-2 (c) B-3

Figure 7: B crack quantification map of the main fracture area

Table 3: B main fracture area fracture quantitative parameters

Number	Area/mm^2	Length/mm	Angle/°	Max width/mm	Mean width/mm
B-1	82.4	40.3	13.6	2.5	2.0
B-2	68.2	53.2	58.4	3.4	2.0
B-3	27.5	55.7	4.5	1.7	1.0

VI. Conclusion

In this paper, in order to achieve automatic identification of building crack defects, a full convolutional neural network is selected for end-to-end pixel-level segmentation of building cracks, and the full convolutional neural network model is trained and optimized. The image morphology calculation is used to skeletonize the building crack defects so as to calculate the crack features.

(1) The training average accuracy, maximum precision, recall and F1 value of the full convolutional neural network model in this paper are 95.79%, 81.15%, 80.81% and 84.19% when the initial learning rate is 1e-5. The validation average accuracy, maximum precision, recall and F1 value are 95.28%, 77.77%, 80.48% and 79.11%, respectively. Taken together, the model in this paper outperforms the other models.

(2) The relative errors of the maximum width of cracks between the model prediction results and the real results in this paper range from 2.1% to 25.4%. The overall accuracy of crack area and length is large, 68% of the measured crack area is higher than the area of the labeled crack, and the relative error of crack length is 26.48%. Taking concrete beam cracks as an example, the distribution of static crack widths ranged from 1.3 to 3.4 mm, 1.02 to 2.03 mm, and 0.42 to 1.39 mm. The dynamic crack width and area values decreased gradually as a whole.

References

[1] Kohns, J., Zahs, V., Klonner, C., Höfle, B., Stempniewski, L., & Stark, A. (2025). Building damage assessment in natural disasters: A trans- and interdisciplinary approach combining domain knowledge, 3D machine learning, and crowdsourcing. Progress in Disaster Science, 100427.

[2] Golewski, G. L. (2023). The phenomenon of cracking in cement concretes and reinforced concrete structures: the mechanism of cracks formation, causes of their initiation, types and places of occurrence, and methods of detection—a review. Buildings, 13(3), 765.

[3] Le, D. B., Tran, S. D., Dao, V. T., & Torero, J. (2017). Deformation capturing of concrete structures at elevated temperatures. Procedia engineering, 210, 613-621.

[4] Angst, U. M. (2019, June). Durable concrete structures: Cracks & corrosion and corrosion & cracks. In 10th International Conference on Fracture Mechanics of Concrete and Concrete Structures FraMCoS-X (Vol. 233307).

[5] Yang, C., He, P., Wang, S., Wang, J., & Zhu, Z. (2025). Study on the Causes of Cracking in Concrete Components of a High-Pile Beam Plate Wharf. Buildings, 15(8), 1352.

[6] Foster, S., Hooper, P., & Easthope, H. (2022). Cracking up? Associations between building defects and mental health in new Australian apartment buildings. Cities & health, 6(6), 1152-1163.

[7] Prasanth, S., & Ghosh, G. (2021, December). Effect of cracked section properties on the resilience based seismic performance evaluation of a building. In Structures (Vol. 34, pp. 1021-1033). Elsevier.

[8] Chen, F., Baji, H., & Li, C. Q. (2018). A comparative study on factors affecting time to cover cracking as a service life indicator. Construction and Building Materials, 163, 681-694.

[9] Velumani, P., Mukilan, K., Varun, G., Divakar, S., Doss, R. M., & Ganeshkumar, P. (2020, December). Analysis of cracks in structures and buildings. In Journal of Physics: Conference Series (Vol. 1706, No. 1, p. 012116). IOP Publishing.

[10] Safiuddin, M., Kaish, A. A., Woon, C. O., & Raman, S. N. (2018). Early-age cracking in concrete: Causes, consequences, remedial measures, and recommendations. Applied Sciences, 8(10), 1730.

[11] Kueh, A. B. H., Wang, X. H., Chen, Y., & Gui, S. J. (2020). Contesting crack modes modeling of reinforced concrete structure threatened by the progressive rust expansion in rebars in the presence of external load. Construction and Building Materials, 263, 120127.

[12] Miura, T., Sato, K., & Nakamura, H. (2020). The role of microcracking on the compressive strength and stiffness of cracked concrete with different crack widths and angles evaluated by DIC. Cement and Concrete Composites, 114, 103768.

[13] Zhou, Z., Wang, C., & Han, X. (2022). Safety evaluation of cracked concrete structures with crack length index. Theoretical and Applied Fracture Mechanics, 122, 103662.

[14] Garrido-Torres, G., & Brescia-Norambuena, L. (2022). Impact of concrete cracking on building projects: case study in Santiago, chile. Practice Periodical on Structural Design and Construction, 27(4), 05022004.

[15] Silva, C. D., Coelho, F., de Brito, J., Silvestre, J., & Pereira, C. (2017). Inspection, diagnosis, and repair system for architectural concrete surfaces. Journal of Performance of Constructed Facilities, 31(5), 04017035.

[16] Munawar, H. S., Hammad, A. W., Waller, S. T., & Islam, M. R. (2022). Modern crack detection for bridge infrastructure maintenance using machine learning. Human-Centric Intelligent Systems, 2(3), 95-112.

[17] Mo, D. H., Wu, Y. C., & Lin, C. S. (2022). The dynamic image analysis of retaining wall crack detection and gap hazard evaluation method with deep learning. Applied Sciences, 12(18), 9289.

[18] Chen, J., Chan, I., & Brilakis, I. (2024). Shifting research from defect detection to defect modeling in computer vision-based structural health monitoring. Automation in Construction, 164, 105481.

[19] El-Din Hemdan, E., & Al-Atroush, M. E. (2025). A review study of intelligent road crack detection: Algorithms and systems. International Journal of Pavement Research and Technology, 1-31.

[20] Inam, H., Islam, N. U., Akram, M. U., & Ullah, F. (2023). Smart and automated infrastructure management: A deep learning approach for crack detection in bridge images. Sustainability, 15(3), 1866.

[21] Ko, P., Prieto, S. A., & de Soto, B. G. (2021). ABECIS: An automated building exterior crack inspection system using UAVs, open-source deep learning and photogrammetry. In ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction (Vol. 38, pp. 637-644). IAARC Publications.

[22] Ko, P., Prieto, S. A., & De Soto, B. G. (2023). Developing a free and open-source semi-automated building exterior crack inspection software for construction and facility managers. IEEe Access, 11, 77099-77116.

[23] Mariniuc, A. M., Cojocaru, D., & Abagiu, M. M. (2024). Building surface defect detection using machine learning and 3d scanning techniques in the construction domain. Buildings, 14(3), 669.

[24] Huo, Z., Wu, X., & Cheng, T. (2024). Crack recognition and defect detection of assembly building constructions for intelligent construction. Journal of Measurements in Engineering, 12(3), 485-501.
[25] Babu, J. C., Kumar, M. S., Jayagopal, P., Sathishkumar, V. E., Rajendran, S., Kumar, S., ... & Mahseena, A. M. (2022). IoT-Based Intelligent System for Internal Crack Detection in Building Blocks. Journal of Nanomaterials, 2022(1), 3947760.
[26] Valero, E., Forster, A., Bosché, F., Hyslop, E., Wilson, L., & Turmel, A. (2019). Automated defect detection and classification in ashlar masonry walls using machine learning. Automation in construction, 106, 102846.
[27] Loverdos, D., & Sarhosis, V. (2022). Automatic image-based brick segmentation and crack detection of masonry walls using machine learning. Automation in Construction, 140, 104389.
[28] Dan, D., & Dan, Q. (2021). Automatic recognition of surface cracks in bridges based on 2D-APES and mobile machine vision. Measurement, 168, 108429.
[29] Wang, R., Chen, R. Q., Guo, X. X., Liu, J. X., & Yu, H. Y. (2024). Automatic recognition system for concrete cracks with support vector machine based on crack features. Scientific Reports, 14(1), 20057.
[30] Dong, X., Liu, Y., & Dai, J. (2025). Recognition of Concrete Surface Cracks Based on Improved TransUNet. Buildings (2075-5309), 15(4).
[31] Yiğit, A. Y., & Uysal, M. (2024). Automatic crack detection and structural inspection of cultural heritage buildings using UAV photogrammetry and digital twin technology. Journal of Building Engineering, 94, 109952.
[32] Miao, P., & Srimahachota, T. (2021). Cost-effective system for detection and quantification of concrete surface cracks by combination of convolutional neural network and image processing techniques. Construction and Building Materials, 293, 123549.
[33] Tripathi, M. (2021). Analysis of convolutional neural network based image classification techniques. Journal of Innovative Image Processing (JIIP), 3(02), 100-117.
[34] Tian, Y. (2020). Artificial intelligence image recognition method based on convolutional neural network algorithm. Ieee Access, 8, 125731-125744.
[35] Kang, S., Kim, S., & Kim, S. (2024). Automatic detection and classification process for concrete defects in deteriorating buildings based on image data. Journal of Asian Architecture and Building Engineering, 1-15.
[36] Protopapadakis, E., Voulodimos, A., Doulamis, A., Doulamis, N., & Stathaki, T. (2019). Automatic crack detection for tunnel inspection using deep learning and heuristic image post-processing. Applied intelligence, 49, 2793-2806.
[37] Wu, X., & Liu, X. (2021). Building crack identification and total quality management method based on deep learning. Pattern Recognition Letters, 145, 225-231.
[38] Chen, K., Reichard, G., Xu, X., & Akanmu, A. (2021). Automated crack segmentation in close-range building façade inspection images using deep learning techniques. Journal of Building Engineering, 43, 102913.
[39] Wang, J., Wang, P., Qu, L., Pei, Z., & Ueda, T. (2024). Automatic detection of building surface cracks using UAV and deep learning-combined approach. Structural Concrete, 25(4), 2302-2322.
[40] Li, G., Liu, Q., Zhao, S., Qiao, W., & Ren, X. (2020). Automatic crack recognition for concrete bridges using a fully convolutional neural network and naive Bayes data fusion based on a visual detection system. Measurement Science and Technology, 31(7), 075403.
[41] Laxman, K. C., Tabassum, N., Ai, L., Cole, C., & Ziehl, P. (2023). Automated crack detection and crack depth prediction for reinforced concrete structures using deep learning. Construction and Building Materials, 370, 130709.
[42] Cristina Saldivia Siracusa,Eduardo Santos Carlos de Souza,Arnaldo Vitor Barros da Silva,Anna Luíza Damaceno Araújo,Caíque Mariano Pedroso,Tarcília Aparecida da Silva... & Alan Roger Santos Silva. (2025). Automated classification of oral potentially malignant disorders and oral squamous cell carcinoma using a convolutional neural network framework: a cross-sectional study. Lancet regional health. Americas,47,101138.
[43] Jian Wu,Chenlin Liu,Aiguo Ouyang,Bin Li,Nan Chen,Jing Wang & Yande Liu. (2024). Early Detection of Slight Bruises in Yellow Peaches (Amygdalus persica) Using Multispectral Structured-Illumination Reflectance Imaging and an Improved Ostu Method. Foods,13(23),3843-3843.
[44] Cheng Chen,Binquan Li,Huiming Zhang,Maihuan Zhao,Zhongmin Liang,Kuang Li & Xindai An. (2025). Performance enhancement of deep learning model with attention mechanism and FCN model in flood forecasting. Journal of Hydrology,658,133221-133221.
[45] Wenshu Zha,Shu Yan,Daolun Li & Detang Lu. (2017). A study of correlation between permeability and pore space based on dilation operation. Advances in Geo-Energy Research,1(2),93-99.
[46] Yuyang Li,Bin Shu,Chuanxiong Wu,Zhengwei Liu,Jian Deng,Zhiqin Zeng... & Zhenghua Huang. (2025). Deep learning-based multi-scale crack image segmentation and improved skeletonization measurement method. Materials Today Communications,46,112727-112727.