

Research Article

Convolutional Neural Networks for Structural Damage Identification in Assembled Buildings

Chunhua You , Wenxiang Liu, and Lei Hou

School of Civil and Constructional Engineering, Hunan Institute of Technology, Hengyang 421002, China

Correspondence should be addressed to Chunhua You; 2005001486@hnit.edu.cn

Received 21 February 2022; Accepted 4 March 2022; Published 11 April 2022

Academic Editor: Hangjun Che

Copyright © 2022 Chunhua You et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper investigates the migration learning AlexNet-based algorithm for the recognition of assembly building structures and the recognition based on an improved algorithm, and gives an analysis of the results. The structure of AlexNet convolutional neural network is introduced and the basic principles of migration learning are analysed. The optimal model for the ceiling damage recognition task was obtained through parameter adjustment, with a test accuracy of 96.6%. The maximum improvement in test accuracy is about 4%, with 82.6% and 79.7% for beam and column damage recognition and infill wall damage recognition respectively.

1. Introduction

In the aftermath of a disaster, rapid assessment of the extent of damage has become an important basis for the allocation of emergency search and rescue resources following an earthquake [1]. The traditional method of assessment mainly involves a group of trained professional engineers and academic researchers visiting the site to survey the extent of the damage. While this approach is accurate, the safety of the assessment team personnel is not guaranteed given the occurrence of aftershocks, and the unsafe elements within the building, and the whole process lasts longer and is less efficient. But with another development in artificial intelligence, computer vision is starting to come into the picture. Intelligent classification and labelling of image data using deep learning methods has become the new craze [2, 3].

The rapid assessment system for postearthquake building damage proposed in this paper has no restrictions on the professionalism of the photographers and can even use drones to enter the interior to take pictures, which greatly accelerates the efficiency of postearthquake assessment and provides suggestions for further investigation [4].

In order to enable people in the earthquake zone to work safely and urgently after the earthquake, a rapid assessment of postearthquake building damage is needed, and relatively

safe houses can be used to prevent people from having nowhere to live and sleeping on the streets. This will provide the basic basis for the government to make scientific and effective regulation [5]. The assessment is divided into structural and nonstructural elements, with structural elements including beams and columns, and nonstructural elements including infill walls and ceilings [6]. The damage to structural elements is directly related to the safety of the overall building. From the photographic data recorded in the Wenchuan and Lushan earthquakes, the damage to beams and columns in houses that did not collapse was mainly as follows: slight cracks in the beams and columns, crushed external concrete, crushed concrete with yielding steel reinforcement, example pictures are shown in Figure 1.

Since the reform and opening up, with the development of the economic level, the quality of China's urban and rural buildings has improved significantly, and in the medium earthquake, there is rarely any damage to the load-bearing elements of houses. Through a large number of post-earthquake disaster picture data the analysis revealed that the infill walls, as the first line of seismic protection for reinforced concrete frame structures, were very seriously damaged, as shown in Figure 2.

Convolutional neural networks are trained on a large amount of data. Theoretically, the richer the data, the better



FIGURE 1: Damage to structural elements in the Wenchuan earthquake.



FIGURE 2: Damage to infill wall in the Lushan earthquake.

the learning ability of the trained convolutional neural network and the better the classification results. CNN have been known for their excellent performance in image classification since their inception. In recent years, the development of CNN has become more and more rapid and has gradually surpassed the human's own ability to classify images, which lays the foundation for using CNN in this paper to perform damage recognition of buildings after earthquakes [7]. In deep learning, training a deep neural network usually requires a large amount of labelled data as the training set, however, in many practical situations, obtaining sufficient data samples is very difficult. Training a model with an inadequate training set runs the risk of overfitting the model, resulting in poor generalisation of the trained model and failure to achieve the expected accuracy on the test data set [8]. The idea of migration learning, which has emerged in recent years, better solves the problem of inadequately large datasets in deep learning. In this paper, we collected a large amount of image data from the Institute of Engineering Mechanics of the China Earthquake Administration in the Wenchuan and Lushan earthquakes, which laid a solid foundation for the establishment of the dataset.

Compared with traditional recognition methods, we do not need to personally extract the damaged areas of the components for feature input, and can avoid the disadvantages of manual feature extraction, for example: (a) it is generally difficult to express complex high-level semantics of images based on some lifting layer feature information of the graph (e.g. colour, texture, etc.), so the generalisation ability is generally weak. (b) These methods are generally designed for specific applications in specific domains, and their generalisation and migration capabilities are mostly weak.

The convolutional neural network-based recognition method only requires us to build a convolutional neural network with completed training and upload the image in the visual operation interface to recognize the result, ensuring the intelligence, simplicity and practicality of the system. In the future, there is still a broad prospect for the development of postearthquake building damage recognition algorithms based on convolutional neural networks.

2. Related Work

Nowadays, more people have become aware of the importance of deep learning, especially in the field of machine learning, which has become more effective. The reason for this is that deep learning has advanced and developed very significantly in several areas, such as sound and text, and most crucially, the use of Internet technology, which has led to an artificial intelligence revolution. Reference [9] proposes a method capable of recovering the properties of cracks. In this method, crack points are first located by means of state-of-the-art crack detection techniques. Then, the skeletal structure of each point is identified using image refinement methods. These structures are integrated into the distance field of the crack point by means of a distance transformation. In this way, crack width, length and direction can be automatically recovered. Reference [10] proposes a new method for detecting spalling regions on the surface of reinforced concrete columns. The properties of the spalling regions are obtained from the image data, according to which the spalling regions are first separated using a threshold-holding algorithm based on local entropy [11]. Based on this, a new global adaptive thresholding algorithm was combined to measure the longitudinal reinforcement

(spalling depth in the column) and the spalling length in the column. A new detection method for postdisaster image classification is proposed in [12], which has four steps: firstly raw data screening, secondly scene classification, then target detection and finally damage assessment. The method was validated in the classification of specific examples. Distinguishing themselves from previous work, [13] et al. propose a new method for structural damage detection using deep CNNs that automatically obtains information from low-level waveform signals rather than relying on manual labelling. The method implements structural damage detection based on data alone without relying on human expert knowledge, and numerical simulations are performed to obtain the response data of the structure, followed by data preprocessing, data augmentation, and finally training of the augmented dataset by a deep CNN to estimate its classification performance for damage localization, and demonstrates superiority compared to another damage extraction method [14].

3. Data Set Creation and Expansion

3.1. Data Set Creation. Most of the data set in this paper comes from the picture data of the Wenchuan and Lushan earthquakes provided by the Institute of Engineering Mechanics of the China Earthquake Administration, and a small part of the data comes from the Internet, as shown in Figure 3.

In this paper, according to the damage characteristics of the building, the frame building is divided into two parts: structural elements, beams and columns, and nonstructural elements, infill walls and nonstructural ceiling elements. The specific classification levels and condition descriptions are shown in Table 1:

3.2. Data Preprocessing. (a) Samples from the seismic images were selected to meet the experimental requirements and the target region of interest in the images was made central by manual processing.

The histogram of a grey-scale image in the grey-scale range $[0, L-1]$ is defined as shown in (1):

$$h(r_k) = n_k. \quad (1)$$

Here r_k — kth level of grey in the image; n_k — Number of pixels in the image with a grey level of r_k .

$$p(r_k) = h(r_k) = \frac{n_k}{MN}, \quad (2)$$

where M — Number of rows of grayscale charts; N — Number of columns in the grey-scale chart; MN — Total number of pixels in the image.

$1/MN$ is the histogram defined in the normalisation (2), which after normalisation $p(r_k)$ can be interpreted as an estimate of the probability that grey level r_k will occur in the image. The histogram is the basis for a variety of other image processing techniques, and the spatial domain processing of images can be expressed as equation (3)

$$g(x, y) = T[f(x, y)], \quad (3)$$

where $f(x, y)$ is the input image. $g(x, y)$ is the transformed image. T is a transform function defined on the neighbourhood of the point (x, y) .

The smallest domain of this equation is 1×1 , in which case g depends only on the value of f at point (x, y) , while T in equation (3) becomes a grey-scale transform function shaped as in equation (4):

$$s = T(r). \quad (4)$$

The immediate effect of equation (4) is to convert the input grey-scale values to other values for the purposes of contrast stretching, binarisation, etc. Histogram equalisation is a special conversion method in the form of equation (4). There exists a very important transformation function in the image domain of the following form:

$$s = T(r) = (L-1) \int_0^r p_r(w)dw, \quad (5)$$

where w — Integral dummy variables; $p_r(w)$ — The probability density function of a random variable r .

The integral part of the right-hand side of the formula is the cumulative distribution function of the random variable r . Because the probability density function of the random variable r is always positive, the cumulative distribution function is increasing and has a maximum value of 1. Therefore, the transformation $T(r)$ is an increasing function and has a value range of $[0, L-1]$. When the random variable is a continuous random variable, the random variable s obtained by the above transformation will be a uniform probability density variable. The pixel values in an image can be treated as if the interval were $[0, L-1]$ a discrete variable, and according to equation (5) the histogram equalisation is obtained, which changes in the following form:

$$\begin{aligned} s_k = T(r_k) &= (L-1) \sum_{j=0}^k p_r(r_j) \\ &= \frac{(L-1)}{MN} \sum_{j=0}^k n_j, k = 0, 1, \dots, L-1. \end{aligned} \quad (6)$$

In equation (6), r_k , MN , n_k , and $p_r(r_k)$ all correspond to the converted values as described previously, and L is the number of grey levels in the image. The histogram equalisation conversion formula ensures two important conditions: the conversion function $T(r)$ is a monotonically increasing function on the interval $0 \leq r \leq L-1$; on the interval $0 \leq r \leq L-1$, there is $0 \leq T(r) \leq L-1$.

3.3. Data Enhancement. In the case of insufficient data samples, data augmentation is often used to expand the sample. A common means of data augmentation is geometric transformation, and in this paper the sample is expanded by means of rotation and mirroring. In the field of image recognition, the rotation, translation and mirroring of an image do not actually change the essential information of



FIGURE 3: Damage to structural elements—beams and columns. (a) in good condition or slightly damaged. (b) Moderate damage. (c) Severe damage.

TABLE 1: Criteria for classifying the damage state of beams and columns.

Failure state	State description
Basically intact and slightly damaged	Basically, there is no crack or slight crack, which does not affect the structural safety
Moderate damage	Large cracks appear on the concrete surface, a small part of the concrete is crushed, and the reinforcement is not bent
Severe damage	Large areas of concrete are crushed, steel bars are bent, and even tilt or collapse

the image. Moreover, by rotating, panning and mirroring the image, the sample can be obtained from multiple angles and directions with different sizes [15]. This transformation can improve the recognition accuracy of the convolutional

neural network, because it can effectively avoid the possibility that the convolutional neural network is not well trained due to the different shooting equipment or shooting angles, thus causing errors in recognition. In this paper, the



(a)



(b)



(c)



(d)



(e)



(f)

FIGURE 4: Continued.

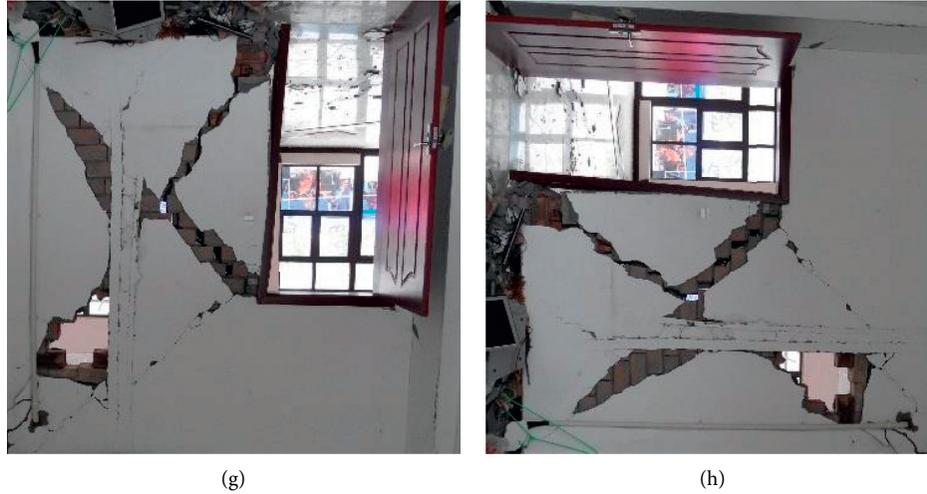


FIGURE 4: Geometric transformation diagram. (a) Original. (b) Original rotated by 90° . (c) Original rotated by 180° . (d) Original rotated by 270° . (e) Mirror image. (f) Mirror rotation 90° . (g) Mirror rotation 180° . (h) Mirror rotation 270° .

original image is rotated and mirrored to increase the number of images to 8 times the original size, as shown in Figure 4.

4. Based on the Improved AlexNet

In this paper, AlexNet has a low accuracy for the task of identifying the extent of damage to beams and columns and the extent of damage to infill walls, and cannot be significantly improved by simply adjusting the hyper-parameters, so improvements to the original AlexNet are considered. In recent years, many scholars have modified and enhanced AlexNet to improve neural network recognition accuracy. In [16], the WN (Weight Normalization) was proposed to replace the LRN layer with the LRN (Local Response Normalization), and the WN was placed after all pooling layers to improve the accuracy of AlexNet model training. In [17], a fusion segmentation algorithm of ReLU6 and Swish was proposed to address the problem of partial failure to update the weights and gradient explosion of the ReLU activation function in AlexNet, which improved the convergence speed and accuracy of AlexNet model training and also alleviated the occurrence of overfitting [18].

4.1. An Improved AlexNet Network Structure. However, as the number of layers increases the disadvantages of a small data set become apparent, as not only does the test error become higher when the network is deeper, but its training error is surprisingly higher as well. Deeper networks are associated with gradient disappearance and explosion problems, which hinder the convergence of the network, and this phenomenon of deeper networks with reduced performance is generally referred to as the degradation problem.

However, in order to prevent the model from being trained poorly due to the increase in parameters after the

addition of the convolutional layer, we choose to combine the residual structure of ResNet with AlexNet to achieve the goal of increasing the convolutional layer without a surge in parameters, so as to improve the problem of low accuracy and loss of convergence in multitask recognition [19, 20].

Assuming that the input to a particular segment of the neural network is x and the desired output is $H(x)$, i.e., $H(x)$ is the desired complex potential mapping, it would be more difficult to train such a model if it were to be learnt; recalling the previous assumption, if a more saturated accuracy has been learnt (or when the error in the lower layers is found to become larger), then the next learning goal shifts to the learning of a constant mapping, i.e. making the input x approximate the output $H(x)$ to keep from causing a drop in accuracy in the later layers.

In Figure 5, the residual network structure is shown by means of "shortcut connections", where the input x is passed directly to the output as the initial result, and the output is $H(x) = F(x) + x$. If $F(x) = 0$, then $H(x) = x$, which is the above mentioned constant mapping. Thus, ResNet is equivalent to changing the learning goal from learning a complete output to the difference between the target value $H(x)$ and x , also known as the residual $F(x) = H(x) - x$. Thus, the training goal later is to approximate the residual result to 0, so that the accuracy does not decrease as the network deepens.

The two structures in Figure 6 are for ResNet34 and ResNet50/101/152 respectively, and their purpose is mainly to reduce the number of parameters. The left figure shows two $3 \times 3 \times 256$ convolutions with the number of parameters: $3 \times 3 \times 256 \times 256 = 1179648$, while the right figure first reduces the 256-dimensional channels to 64 dimensions by 1×1 convolution and finally recovers them by 1×1 convolution, using the overall number of parameters: $1 \times 1 \times 256 \times 64 + 3 \times 3 \times 64 \times 64 + 1 \times 1 \times 64 \times 256 = 69632$. The number of parameters is reduced by a factor of 16.94 compared to the left graph, so the main purpose of the right graph is to reduce the number of parameters and hence the computational effort [21].

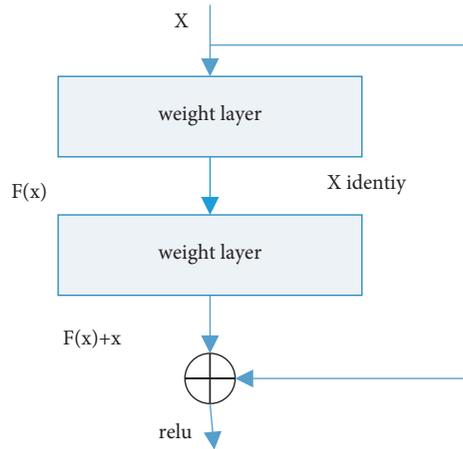


FIGURE 5: Structure of the residuals.

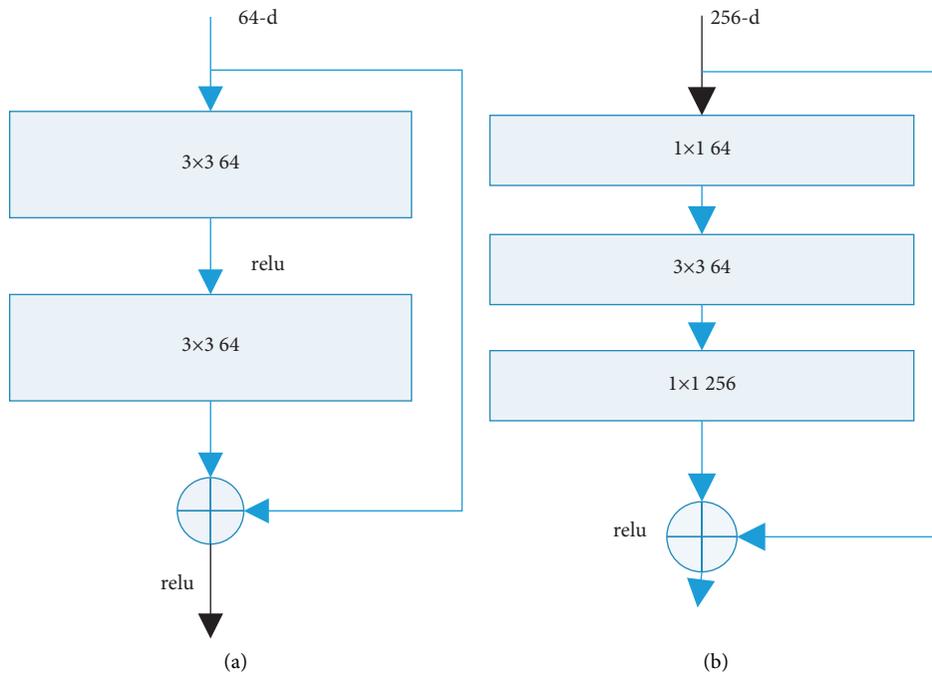


FIGURE 6: Two forms of residual structure. (a) Basic block. (b) Bottleneck.

For a regular ResNet (Figure 6(a)), basic block can be used for networks with 34 layers or less. (Figure 6(b)) Bottleneck is used for deeper networks (e.g. 101 layers) with the aim of reducing the number of computations and parameters. This paper therefore considers adding the Basic block module after the convolutional layer of AlexNet, replacing the three fully-connected layers with one. The AlexNet convolutional layer is used to extract the feature information of the images in the dataset, and a comparison experiment is set up to add one Basic block block, two Basic block blocks and three Basic block blocks to the classification layer, and then the accuracy of the network is tested after the training is completed. The network improvement scheme is shown in Figure 7.

5. Damage Recognition Experiments

5.1. Vibration Recognition. The time-frequency response of the sensor was tested at different frequencies and accelerations. The time and frequency response of the sensor is shown in Figure 8 for external excitation frequencies of 90 Hz and 350 Hz and accelerations of 1.0 g and 3.0 g respectively. The output waveform of the FBG vibration sensor is a sinusoidal waveform, as can be seen from the time domain signal; Figure 8(b) and (d) show the frequency spectrum obtained from the Fast Fourier Transform (FFT) of the time domain signal. The results are consistent with the experimentally set frequency parameters.

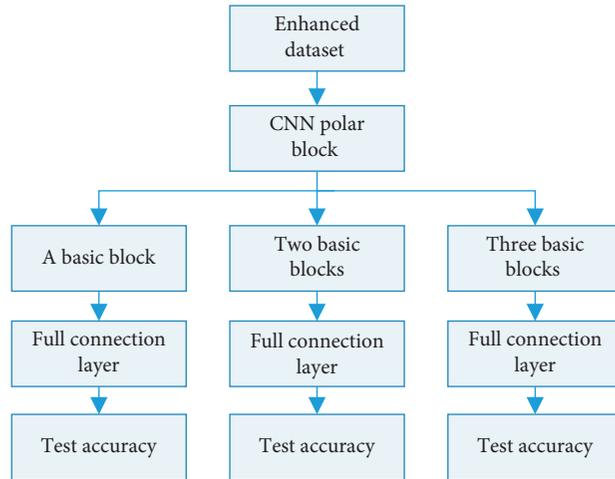


FIGURE 7: Design options for network improvements.

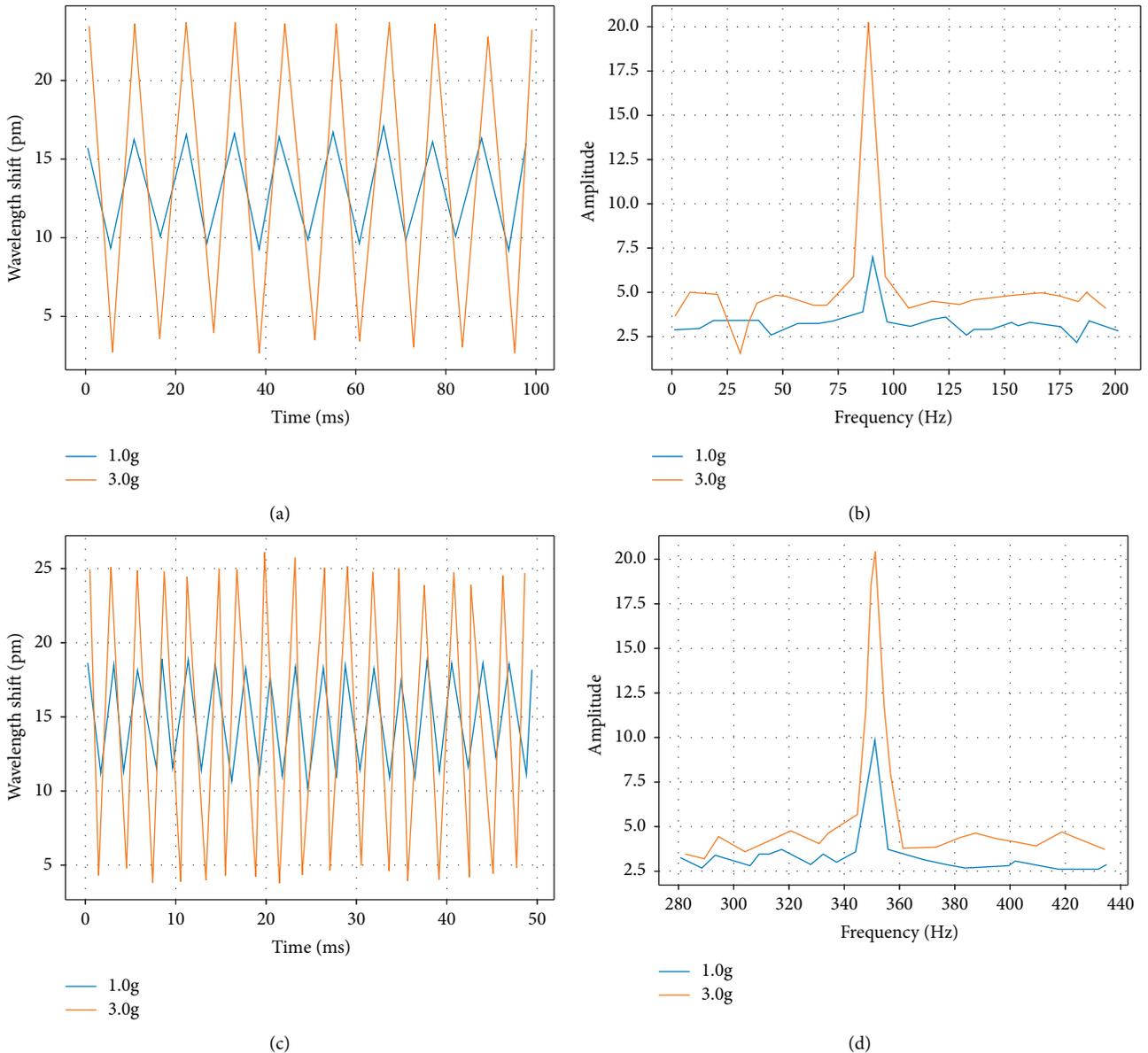


FIGURE 8: Response curves of the time-frequency characteristics of the sensor. (a) 90 Hz time domain; (b) 90 Hz frequency domain; (c) 350 Hz time domain; (d) 350 Hz frequency domain.

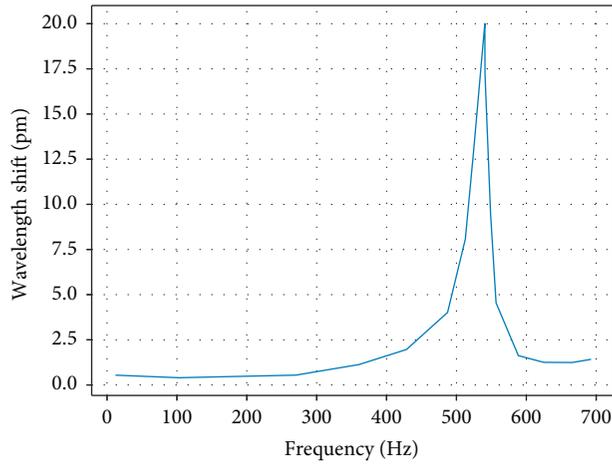


FIGURE 9: Amplitude frequency response curve of sensor.

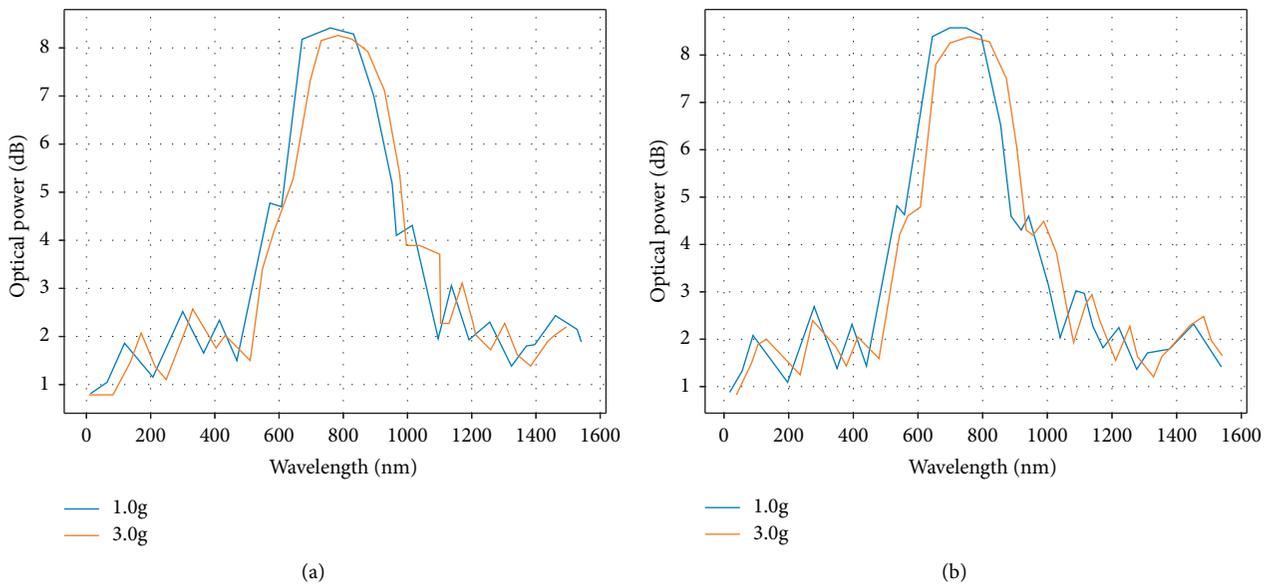


FIGURE 10: Spectra of short grating at different frequencies.(a) 90 Hz; (b) 350 Hz.

To verify the amplitude-frequency characteristics of the sensor, the maximum drift of the central wavelength of the sensor was investigated under certain conditions of acceleration, corresponding to different frequencies. The maximum value of the centre wavelength of the sensor was recorded and the curve is shown in Figure 9. It can be seen that the inherent frequency of the sensor at this acceleration is approximately 543.9 Hz, which is in good agreement with the simulated result of 568.6 Hz. The slight deviation in frequency is due to the sensor assembly error and other reasons.

The spectrograms of the sensor were measured at different frequencies and acceleration excitations. The spectra of the sensor are shown in Figure 10 for external excitation frequencies of 90 Hz and 350 Hz and accelerations of 1.0 g and 3.0 g respectively. It can be seen that the FBG spectrum did not broaden with increasing acceleration, indicating that the chirp phenomenon was effectively avoided.

The acceleration response of the sensor was tested at a certain frequency. The acceleration response of the sensor was tested at a frequency of 300 Hz with an external excitation of 0.2 g. The results showed that the acceleration and wavelength shift of the sensor were quasi-linear, with a linear fit of 99% or more and good linearity; the sensitivity of the sensor was 6.7 pm/g; the repeatability error of the sensor was about 1.7%, with good repeatability

5.2. Comparison of Whether Damage Is Identified or Not. In Scheme 1 for determining whether damage is present, as shown in Figure 11, the accuracy of both CNN1 and CNN2 is below 70% and the accuracy of LSTM is over 95%, so only LSTM can meet the requirements in this step of damage identification. The accuracy of CNN2 is improved by only 2.23% over CNN1, while LSTM is improved by 35.42% over CNN1, which indicates that in

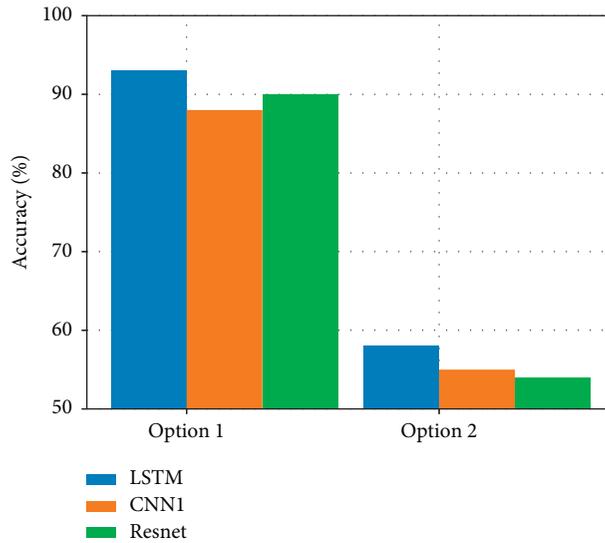


FIGURE 11: Accuracy of whether or not an injury was sustained.

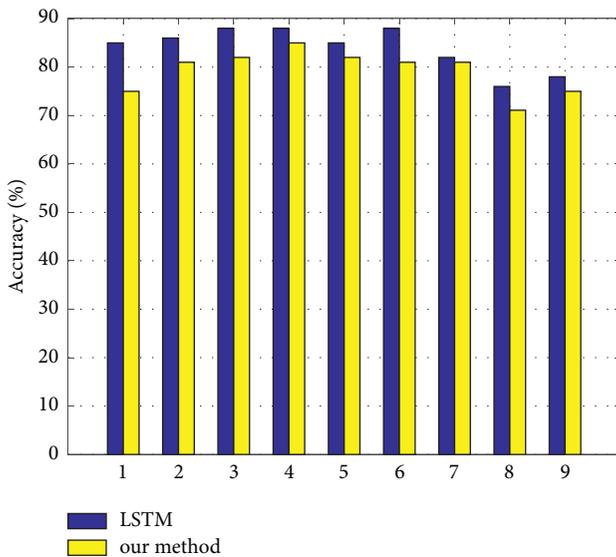


FIGURE 12: Accuracy of damage localization by location.

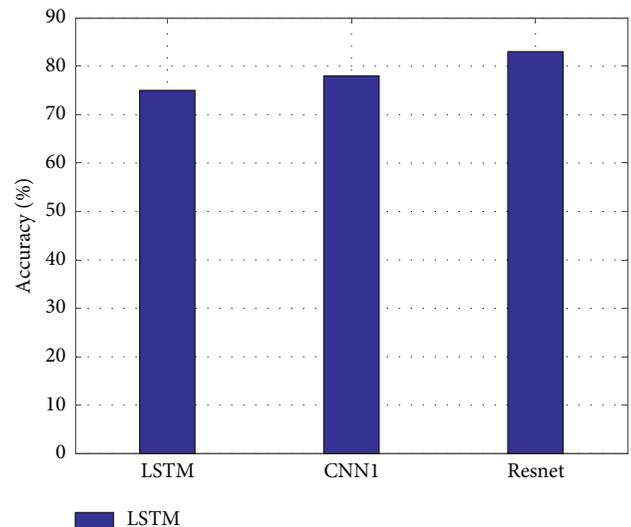


FIGURE 13: Overall average accuracy of injury localization.

Scheme 1, the The increase in sample data is not critical but the method of recognition, and therefore the LSTM method is more appropriate than the CNN method in Scenario 1.

In Scheme 2, as shown in Figure 11, the accuracy of all three methods is relatively high, so all three methods are possible. However, the recognition in Scheme 2 is more tedious and has to be divided into 10 to identify each one.

Combining the accuracy of the three methods of Option 1 and Option 2, it is more appropriate to adopt Option 1 and use the LSTM method when determining whether the damage is.

In the recognition of damage localization, as shown in Figures 12 and 13, the accuracy of damage localization at each location and the overall average accuracy are LSTM > CNN2 > CNN1, where the accuracy of both CNN1

and CNN2 is about 75, while the accuracy of LSTM can reach 91.76%, among which only LSTM meets the requirements.

6. Conclusions

Through the comparative experiments in this paper, it can be seen that these two methods effectively alleviate the AlexNet network’s overfitting problem. At this point, the improved convolutional neural network algorithm has effectively solved the problem of framed building vandalism. The evaluation of the entire postearthquake frame building will be carried out quickly. Experimentation with the number of parameters involved in training shows that with the feature extractor training approach, the model only needs to retrain the function of the last fully-connected layer.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declared that they have no conflicts of interest regarding this work.

References

- [1] O. Abdeljaber, O. Avci, S. Kiranyaz, M. Gabbouj, and D. J. Inman, "Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks," *Journal of Sound and Vibration*, vol. 388, pp. 154–170, 2017.
- [2] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, and D. J. Inman, "1D convolutional neural networks and applications: a survey," *Mechanical Systems and Signal Processing*, vol. 151, Article ID 107398, 2021.
- [3] Y.-J. Cha, W. Choi, and O. Büyüköztürk, "Deep learning-based crack damage detection using convolutional neural networks," *Computer-Aided Civil and Infrastructure Engineering*, vol. 32, no. 5, pp. 361–378, 2017.
- [4] C. Li and M. Wand, "Combining Markov random fields and convolutional neural networks for image synthesis," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2479–2486, San Juan, PR, USA, June, 2016.
- [5] X. Tao, D. Zhang, W. Ma, X. Liu, and D. Xu, "Automatic metallic surface defect detection and recognition with convolutional neural networks," *Applied Sciences*, vol. 8, no. 9, p. 1575, 2018.
- [6] O. Abdeljaber, O. Avci, M. S. Kiranyaz, B. Boashash, H. Sodano, and D. J. Inman, "1-D CNNs for structural damage detection: verification on a structural health monitoring benchmark data," *Neurocomputing*, vol. 275, pp. 1308–1317, 2018.
- [7] K. Wang, X. Wang, L. Lin, M. Wang, and W. Zuo, "3D human activity recognition with reconfigurable convolutional neural networks," in *Proceedings of the 22nd ACM International Conference on Multimedia*, pp. 97–106, New York, NY, USA, 2014, November.
- [8] D. Wu, C. Zhang, L. Ji, R. Ran, H. Wu, and Y. Xu, "Forest fire recognition based on feature extraction from multi-view images," *Traitement du Signal*, vol. 38, no. 3, pp. 775–783, 2021.
- [9] O. Avci, O. Abdeljaber, S. Kiranyaz, M. Hussein, M. Gabbouj, and D. J. Inman, "A review of vibration-based damage detection in civil structures: from traditional methods to Machine Learning and Deep Learning applications," *Mechanical Systems and Signal Processing*, vol. 147, Article ID 107077, 2021.
- [10] Z. Waszczyszyn and L. Ziemiański, "Neural networks in mechanics of structures and materials—new results and prospects of applications," *Computers & Structures*, vol. 79, no. 22–25, pp. 2261–2276, 2001.
- [11] M. Mahdianpari, B. Salehi, M. Rezaee, F. Mohammadimanesh, and Y. Zhang, "Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery," *Remote Sensing*, vol. 10, no. 7, p. 1119, 2018.
- [12] D. Cho, Y.-W. Tai, and I. Kweon, "Natural image matting using deep convolutional neural networks," in *European Conference on Computer Vision*, pp. 626–643, Springer, New York, NY, USA, 2016.
- [13] L. Jing, T. Wang, M. Zhao, and P. Wang, "An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox," *Sensors*, vol. 17, no. 2, p. 414, 2017.
- [14] C. V. Dung and L. D. Anh, "Autonomous concrete crack detection using deep fully convolutional neural network," *Automation in Construction*, vol. 99, pp. 52–58, 2019.
- [15] A. Payan and G. Montana, "Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks," 2015, <https://arxiv.org/abs/1502.02506>.
- [16] R. Oullette, M. Browne, and K. Hirasawa, "Genetic algorithm optimization of a convolutional neural network for autonomous crack detection," vol. 1, pp. 516–521, in *Proceedings of the 2004 congress on evolutionary computation (IEEE Cat. No. 04TH8753)*, vol. 1, pp. 516–521, IEEE, Portland, OR, USA, 2004, June.
- [17] Y. Xue and Y. Li, "A Fast detection method via region-based fully convolutional neural networks for shield tunnel lining defects," *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 8, pp. 638–654, 2018.
- [18] J. Kawahara, C. J. Brown, S. P. Miller et al., "BrainNetCNN: convolutional neural networks for brain networks; towards predicting neurodevelopment," *NeuroImage*, vol. 146, pp. 1038–1049, 2017.
- [19] M. Sahin and R. A. Sheno, "Quantification and localisation of damage in beam-like structures by using artificial neural networks with experimental validation," *Engineering Structures*, vol. 25, no. 14, pp. 1785–1802, 2003.
- [20] X. Yang, H. Sun, X. Sun, M. Yan, Z. Guo, and K. Fu, "Position detection and direction prediction for arbitrary-oriented ships via multitask rotation region convolutional neural network," *IEEE Access*, vol. 6, Article ID 50839, 2018.
- [21] S. H. Lee, C. S. Chan, P. Wilkin, and P. Remagnino, "Deep-plant: plant identification with convolutional neural networks," in *Proceedings of the 2015 IEEE International Conference on Image Processing (ICIP)*, pp. 452–456, IEEE, Quebec City, QC, Canada, 2015, September.