Automatic Focus Fusion Method of Concrete Crack Image Based on Deep Learning

Chuang Wang(✉), Jiawei Pang, Xiaolu Deng, Yangjie Xia, Ruiyang Li, and Caihui Wu

China Construction Third Engineering Bureau Group Co., Ltd., Zhanjiang 524000, China
229247432@qq.com

Abstract. The algorithm used in the traditional image auto focus fusion method is easy to fall into local iteration, resulting in poor quality of image fusion results. A depth learning based concrete crack image auto focus fusion method is designed. Extract the features in the digital image to obtain the template feature set of the concrete crack image, match, and use the filter function to reduce noise, optimize the image auto focus fusion algorithm, and improve the quality of the fused image after transformation. In order to verify the effectiveness of the design method, comparative experiments are designed to compare the results of the design method and the traditional methods. In terms of fusion focusing results, output signal to noise ratio and image histogram, the results of the design method are better than the traditional methods. The image quality and output signal to noise ratio obtained are higher, and the histogram distribution is more uniform, indicating that the image quality is better.

Keywords: Deep learning · Artificial intelligence · Image autofocus · Image fusion

1 Introduction

Self focusing image fusion is an important part of image fusion technology, and the key is image processing. When forming images in a scene, due to the limited focusing range of the optical system, it is usually not easy for the optical imaging system to generate clear images of objects at different distances in the scene [1]. When the focus of the imaging system is focused on an object, it can form a clear image on the image plane. At this point, the images generated on the surface of objects located elsewhere will exhibit varying levels of blur. Therefore, Principles of Optical Lens Imaging enables the imaging system to continuously increase its resolution, but it is impossible to avoid the impact of the specified focus range on the overall effect of the imaging image, that is, It is not easy to obtain clear images of all objects in the same scene only through the imaging system [2]. To more realistically and comprehensively reflect the information of the scene, it is hoped to obtain, to clear images of all objects in the scene. The way to deal with this trouble is to focus on different objects within the scene to obtain multiple auto-Next, fuse these autofocus images to improve their respective clear areas, to Calculate the usage area of each, convenient obtain all the objects in the scene. Fusion image of all objects is clear, that is, auto-focus image fusion.

B. Wang et al. (Eds.): ICMTEL 2023, LNICST 533, pp. 200–211, 2024.
https://doi.org/10.1007/978-3-031-50574-4_14

Reference [3] ask a full-focus image fusion method of light field based on edge enhancement guided filtering, which can obtain full-focus image of light field, denoise through guided filtering technology, classify image characteristics using fuzzy clustering method, and fuse full-focus image of light field through edge enhancement. This method can improve the fusion effect, but the fusion efficiency is poor. Reference [4] proposes a multi-focus image fusion method using Laplacian and convolutional neural network. The multi-focus image is obtained according to the sensor, and the convolutional neural network is used to build a multi-focus image fusion trainer for image fusion training. The Laplacian method is used to increase the features and improve the multi-focus image fusion effect. However, this method is affected by the light intensity and has poor reliability.

To solve the above problems, this paper proposes an concrete crack image auto-focus fusion method based on depth learning. The auto-focus image fusion technology enables objects with different imaging distances to be clearly presented in an image. The artificial intelligence image auto-focus fusion method based on deep learning has laid a good foundation for feature extraction, image recognition and other processing, thus effectively improving the utilization of image information and the reliability of the system for target detection and recognition.

2 Automatic Focusing Fusion Method of Concrete Crack Image Based on Depth Learning

2.1 Concrete Crack Image Feature Extraction and Matching

In order to realize automatic focusing and fusion of concrete crack image, a remote concrete crack image filtering and detection model is construct, and the feature binding feature matching of concrete crack image is put into effect out under the information enhancement skill. The edge wheel thickness of the local sound array of concrete crack image is determined. The feature quantity is described by $p^* = \left(X^{(CS2)}, \theta, \rho\right)$, where $X^{(CS2)}$ is the fuzzy blocking feature quantity of concrete crack image, θ is the edge feature quantity of concrete crack image, and ρ is the contour feature quantity of concrete crack image. In the spatial coordinate system, the pixel data set of concrete crack image is as follows:

$$M = \left(\theta^e, \rho^e\right) = EFA(\theta, \rho) \tag{1}$$

Among them, *EFA* is factor analysis, h is the pixel set of concrete crack image, and the pixel characteristic structure and binary respectively of concrete crack image are performed using the positive diffusion method [5, 6]. The separation guide line is (θ^e, ρ^e), and the vector features combined with fuzzy blocks concrete crack image is obtained through the local quantitative The specific feature analysis is as follows:

$$\begin{cases} EX^{(CS2)} = \{x | x \in [0, h]\} \\ EY^{(CS2)} = \rho^e cos\theta^e \\ EZ^{(CS2)} = \rho^e sin\theta^e \end{cases} \tag{2}$$

Among them, $X^{(CS2)}$, $Y^{(CS2)}$, $Z^{(CS2)}$ are respectively the fuzzy block feature matching vector on the three-dimensional point, x is the pixel on the $X^{(CS2)}$ plane, and the three-dimensional scattered point of the concrete crack image is used about information fusion, and $\rho^e - R$ is the displacement center [7, 8]. Among them, R is the displacement distance. Based on The sparse method was used to cut the features of the concrete crack image, resulting in the length of the edges of the concrete crack image:

$$Team(z) = argmax \underset{k=1,2,\cdots,R}{(y_k.z + e_k)} \tag{3}$$

Among them, z is the pixel of the concrete crack image on the $Z^{(CS2)}$ plane in the spatial rectangular coordinate system, and k is the pixel. Then, the local structure feature decomposition method is used to obtain the template feature set of concrete crack images [9, 10]. Establish the analytical model with the same characteristics as the concrete crack image uses the Gradient descent method to track the pixels. Under the super-resolution visual imaging, the filter function of the concrete crack image is:

$$sim(x_c, x_d) = \frac{\sum_{m=1}^{p} y_{cm} \times \sum_{m=1}^{p} y_{dm}}{\sqrt{\sum_{m=1}^{p} y_{cm}^2 \times \sum_{m=1}^{p} y_{dm}^2}} \tag{4}$$

Among them, x_c and x_d are the concrete crack image pixel abscissa under the super-resolution visual imaging condition, y_c, y_d are the concrete crack image pixel ordinate under the super-resolution visual imaging condition, and m is the amplitude-frequency characteristic of the filter function.

In the sub-blocks $G_{a,b}$ with $M \times N$ and 2×2 sub-blocks, the automatic focusing and information fusion processing of the image are performed by the method of guided filter detection. Figure 1 shows the block diagram of the digital image auto-focus fusion route in this article.

It can be seen from the analysis of Fig. 1 that in this paper, the concrete crack image filtering is extracted, then the image features are compared, and the image edge scale function is obtained through pixel feature reconstruction. After calculation, the concrete crack image focusing automatic fusion is achieved by removing the image clutter, and after guided filtering detection.

According to the sparse prior visual feature extraction method, the feature segmentation mold of the concrete crack image is established, and the Iterative reconstruction model of the concrete crack image is created by combining the features of the concrete crack image with the information enhancement technology. The specific grayscale pixels are as follows:

$$pixel\_C = max\left(\sum_{i=1}^{S} (Q - P)\right) \tag{5}$$

Among them, P is the concrete crack image feature, and Q is the artificial intelligence feature after three-dimensional reconstruction. Based on this, a distributed combination

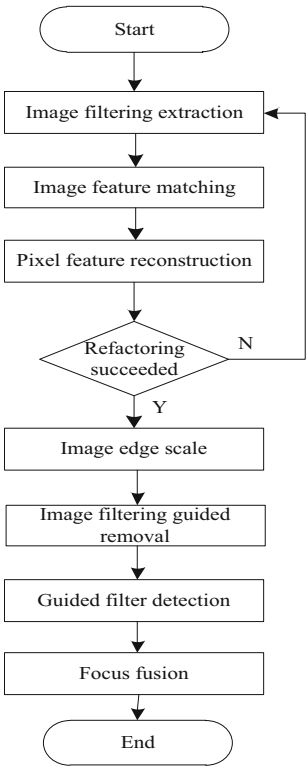


Fig. 1. Road map of concrete crack image auto focusing fusion

model of low resolution and high resolution templates concrete crack image is established, and the gray feature decomposition of concrete crack image execute through the scheme and fuzzy relationship degree of spatial block fusion of the concrete crack image is $K(x_0, y_0)$, and $K(x_0, y_0)$ is In the center, remote fusion is performed, and the grayscale marginal set of concrete crack images is acquire as:

$$pixel\_D = max\left(\sum_{i=1}^{S}(R-K)\right) \tag{6}$$

If $pixel\_C < pixel\_D$, through the difference fusion and scheduling method, the template matching coefficient of the concrete crack image is obtained as:

$$\begin{cases} \beta = \theta / \pi \\ s = 1 - \frac{\lambda_1}{\lambda_2} \\ v = \lambda_1 + \lambda_2 \end{cases} \tag{7}$$

Among them, λ_1 and λ_2 the detection results are for the long and short edges, respectively. Based on edge contour testing, the template matching output is obtained:

$$I_{GSM} = \sum_{i=1}^{N} (\beta(C_i + v_i|s_i) - \beta(v_i)) \tag{8}$$

Among them, N is a combination of natural numbers and i the number of test point. Using adaptive feature detection and parameter estimation algorithms, guided filtering is applied to concrete crack images, resulting in the following filtering function:

$$\hat{x}(r/r) = \sum_{j}^{\delta} \hat{x}^i(r/r)u_j(k) \tag{9}$$

Among them, r represents the filter function guide value, k matching coefficient for concrete crack image points, δ break down the scale for Eigen decomposition of a matrix, $u_j(k)$ pixel intensity, and j Is the energy coefficient.

Foundation on the combination of digital feature matching and communication tracking methods,

the concrete crack image guided filtering and noise reduction processing are performed.

2.2 Introduce Deep Learning Neural Network

Within the pulse coupled neural network put forward akhorn], a pulse coupled neuron N_{ij} is composed of three parts: receiving domain, modulation part, and pulse generator. Figure 2 shows the structure of pulse coupled neurons.

There are two passageway in the receive domain. One passageway is called channel F, which is used to receive feedback input F_{ij}; The other channel is called channel L and is used to receive link input L_{ij}. The external stimulation signal is transmitted to neuron N_{ij} from channel F, while the output signal of neurons in other adjacent positions is transmitted to neuron N_{ij} from channel L. A pulse coupled neuron can usually be described by the following mathematical formula:

$$\begin{cases} F_{ij}(n) = \exp\left(-1/\alpha_F\right)F_{ij}(n-1) + S_{ij} + V_F \sum_{kl} M_{ijkl}Y_{kl}(n-1) \\ L_{ij}(n) = \exp\left(-1/\alpha_L\right)L_{ij}(n-1) + V_L \sum_{kl} W_{ijkl}Y_{kl}(n-1) \\ U_{ij}(n) = F_{ij}(n)\left[1 + \beta L_{ij}(n)\right] \\ \theta_{ij}(n) = \begin{cases} \exp\left(-1/\alpha_L\right)\theta_{ij}(n-1), & if\ Y_{ij}(n-1) = 0 \\ V_T, & otherwise \end{cases} \\ Y_{ij}(n) = \begin{cases} 1, & if\ U_{ij}(n) > \theta_{ij}(n) \\ 0, & otherwise \end{cases} \end{cases} \tag{10}$$

It can be seen from Eq. (10) that the internal activity u is composed of the inputs of channel F and channel L. After obtaining the internal activity U, U is compared with the

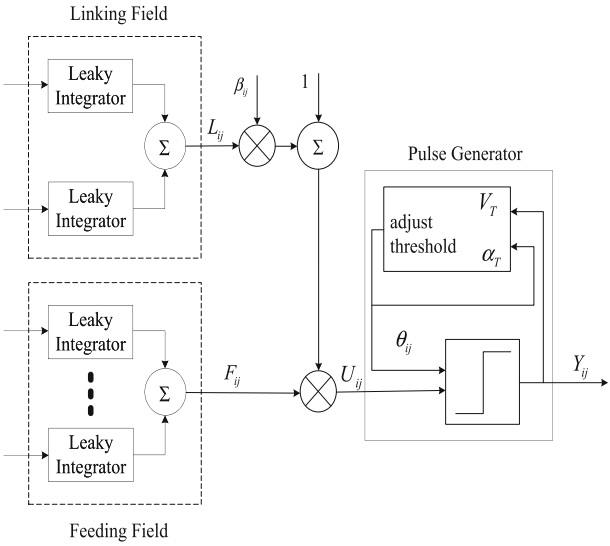


Fig. 2. Structure of pulse coupled neurons

dynamic threshold B of neurons, and the output Y is obtained by comparison. If there is significant internal activity U, neurons will produce pulse signals.

α_F, α_L and α_T are the attenuation ordinary courtes of PCNN neurons, and V_F, V_L and V_T are amplification factors. M and W are the weights of the feedback part and the link part of the receiving domain, respectively. The constant β is the connection strength between neurons.

In the system put in this paper, PCNN is used to process image directly. The structure of PCNN is single-layer, similar to a two-dimensional matrix. One neuron is connected with other surrounding neurons. The structure and function of all neurons in PCNN are the same. The number of neurons in PCNN is consistent with the number of pixels of the input image. Moreover, the pixels of the input image correspond to neurons one by one.

For neuron N_{ij}, the input of channel F is the gray price of pixel S_{ij} relevant; to the neuron.

$$F_{ij}(n) = S_{ij} \tag{11}$$

Generally, the output of an ordinary neuron in PCNN can be expressed by formula (10). However, in this chapter, the output of neuronal hunger is indicate by under; formula (12).

$$Y_{ij}(n) = \begin{cases} U_{ij}(n), & if\ U_{ij}(n) > \theta_{ij}(n) \\ 0, & otherwise \end{cases} \tag{12}$$

The output of PCNN used in this paper is also an image, that is, image a, with O size of $P \times Q$ picture element. The picture element O_{ij} value is the sum of the outputs of its

corresponding neuron N_{ij}, as shown in formula (13):

$$O_{ij} = \sum_{n} Y_{ij}(n) \tag{13}$$

We use formulas (10) to (12) to describe the pulse coupled neuron and PCNN used in multi focus image fusion in this paper.

Before describing the execution steps of PCNN used in this paper, we need to introduce the symbols of variables and constants used.

External stimulus (feedback input) F is a two-dimensional matrix with size $P \times Q$, as shown in formula (11). The value of each element in F is the gray value of the pixel of its corresponding input image. L is the link input matrix with size $P \times Q$. $L = \exp\left(-1/\alpha_L\right) * L + V_L * work$, where $work = (Yconv2K)$. Y is the output matrix, which records the output of all current neurons. K is the kernel matrix, and its size is $(2 \times r + 1) \times (2 \times r + 1)$. Θ is the threshold matrix. U is the matrix recording the internal activity of each neuron. O is the output matrix of PCNN. F, L, $work$, Y, Θ, U and O are two-dimensional matrices with the same size Represents dot product, '$conv2$' represents two-dimensional convolution.

The specific execution steps of PCNN for multi focus image fusion proposed in this paper are as follows:

1) Initialize the relevant parameters and matrices.

$$\begin{cases} F = S \\ L = U = Y = 0 \\ \Theta = 1 \end{cases} \tag{14}$$

The initialization of K is different from other array. The value of the elementary substance at the center of the K matrix is 1. The value of other elements in the K matrix is $1/d$, and d is the range from the element to the pivot, and $d \leq r$.

2) Select the number of iterations np of PCNN. Initialize n, $n = 1$.

$$\begin{cases} L = exp\left(-1/\alpha_L\right) * L + V_L * work \\ U = F * (1 + \beta * L) \end{cases} \tag{15}$$

3) *If* $n > np$, goto step 4), else go back to step 2).
4) Matrix O is used as the output of PCNN.

2.3 Optimized Image Auto Focus Fusion Algorithm

In this paper, (x, y) is used to represent the coordinates of any pixel, $g_i^A(x, y)$ and $f_N^A(x, y)$, $g_i^B(x, y)$ and $f_N^B(x, y)(i = 1, 2, \cdots, N)$ are used to represent the high-frequency and low-frequency subband coefficients of the original figure A and B after NWF transformation and decomposition respectively; $g_i^F(x, y)$ and $f_N^F(x, y)$ represent the combined NWF transform high-frequency and low-frequency incidentally factor The high-frequency take actor of the original image after NWF transform represent the detail communication of the original presentation. The more detail communication, the higher the definition.

The concept of contrast is applied to multi-sensor image fusion, and the image contrast C is defined as:

$$C = \frac{L_P - L_B}{L_H} \times L_B \tag{16}$$

where, L_P is the part gray level of the image. L_B is the indigenous background gray level of the image (equivalent to the low-frequency component product after NWF transformation); $L_H = L_P - L_B$. The high-frequency components after the equivalent NWF transformation and the measurement accuracy of image comparison are as follows:

$$(x, y) = \sum_{i=1}^{N} \sum_{(m,n)\in N(x,y)} \frac{g_i^I(x, y)^2}{f_N^I(x, y)} \tag{17}$$

where: $I = A, B$; $N(x, y)$ is a window defined with pixel (x, y) as the center, which is taken as 10 in this paper × 10 (pixels). The larger $C^I(x, y)$, the greater the contrast and the higher the definition of the local area where pixel (x, y) is located in the original image I. According to the contrast measurement of the image, the clear and blurred regions of the pixels can be approximately determined. The fusion process is to combine the NWF transform coefficients of the pixels belonging to the clear region in the multi focus image. The high-frequency subband coefficients of multifocal images are combined as follows:

$$g_i^F(x, y) = \begin{cases} g_i^A(x, y), C^A(x, y) \geq C^B(x, y) \\ g_N^B(x, y), C^A(x, y) < C^B(x, y) \end{cases} \tag{18}$$

The low-frequency subband fusion process of multi focus image is similar, that is:

$$f_N^F(x, y) = \begin{cases} f_N^A(x, y), C^A(x, y) \geq C^B(x, y) \\ f_N^B(x, y), C^A(x, y) < C^B(x, y) \end{cases} \tag{19}$$

In multi focus image, considering the correlation of adjacent pixels, the possibility of the following situations is generally small: a NWF coefficient after combination comes from the coefficient after image A transformation, while most of the coefficients after combination of adjacent pixels come from the coefficient after image B transformation; And vice versa. Therefore, consistency test should be conducted on the combined NWF transform coefficients to ensure that most of the combined NWF coefficients of a pixel and its adjacent pixels come from the same image. After consistency test, the combined NWF coefficients can be inversely transformed to have the blend image.

3 Experiment

3.1 Experimental Parameters

In order to verify the application performance of this method in digital image dynamic focusing fusion, experimental tests and analysis are carried out. Relevant parameters are shown in Table 1.

By setting the parameters above, take two groups of image for experiments.

Table 1. Experimental Data

Serial number	Parameter	Size
1	Image window size	25 * 25 pixels
2	Image filtering coefficient	0.26
3	Image edgeretention	0.8
4	Guided filterscale	4
5	Numbero fiterations	200times

3.2 Comparison of Fusion Focusing Results

The imitate tentative of digital image focusing fusion method is completed by using PC equipped with MATLABR2018b software. The computer configuration of the simulation experiment PC is Intel (R) Pentium CPU of 2.42 GHz, the memory RAM of the computer is 4.00 GB, and the operating system of the PC is the 64 bit system of windows 8.1.

In this paper, the VITOVision database is selected as the experimental data set to obtain the concrete crack image fusion focusing results and comparison, as shown in Fig. 3.

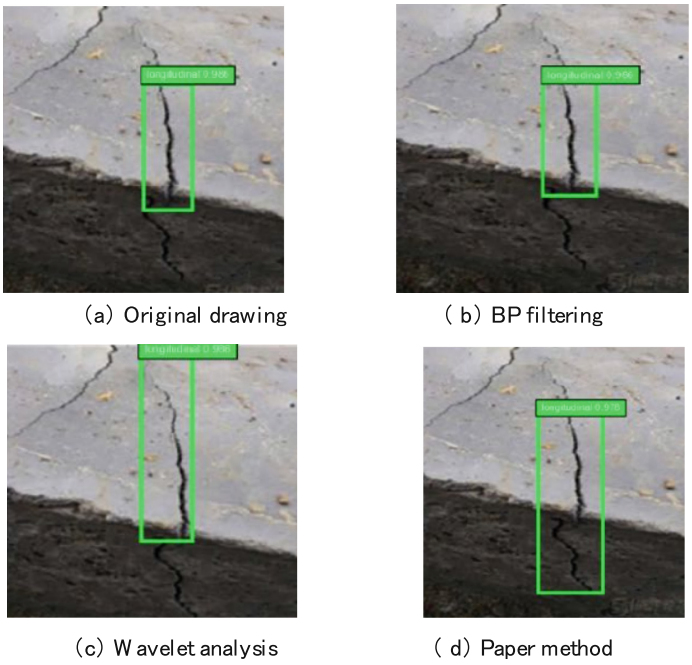
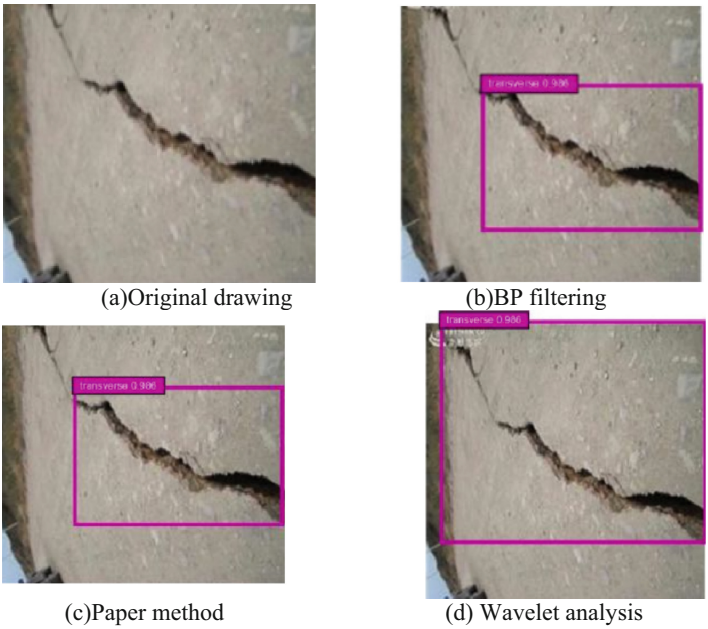


Fig. 3. Automatic focus fusion result of concrete longitudinal crack image

The concrete crack image fusion focusing results and comparison are shown in Fig. 4.



(a)Original drawing (b)BP filtering

(c)Paper method (d) Wavelet analysis

Fig. 4. Concrete transverse crack image auto focus fusion results

Based on the analysis of the above simulation results, it can be concluded that this method can effectively achieve the automatic focusing combination of concrete crack images and increase the digital imaging quality.

3.3 Output SNR Comparison

The conclusion of comparing the signal-to-noise ratio of the output is shown in Table 2.

Table 2. Signal noise ratio comparison test

Iterations/time	Paper method	Edge enhancement guided filtering/dB	Laplace method/dB
50	32.12	12.33	23.36
100	45.45	15.43	29.11
150	52.34	19.21	35.46
200	54.34	21.15	42.12

According to the comparison test results of signal-to-noise ratio, the maximum signal-to-noise ratio of the edge enhanced guided filtering method and the Laplace's method is 21.15 dB and 42.12 dB, respectively, while the maximum signal-to-noise ratio of this method can reach 54.34 dB, indicating that the imaging quality of this method is better.

3.4 Image Histogram Comparison

So as to further verify the image focus fusion effect of the method in this paper, the image histograms processed by different methods are obtained, and the results are shown in Fig. 5.

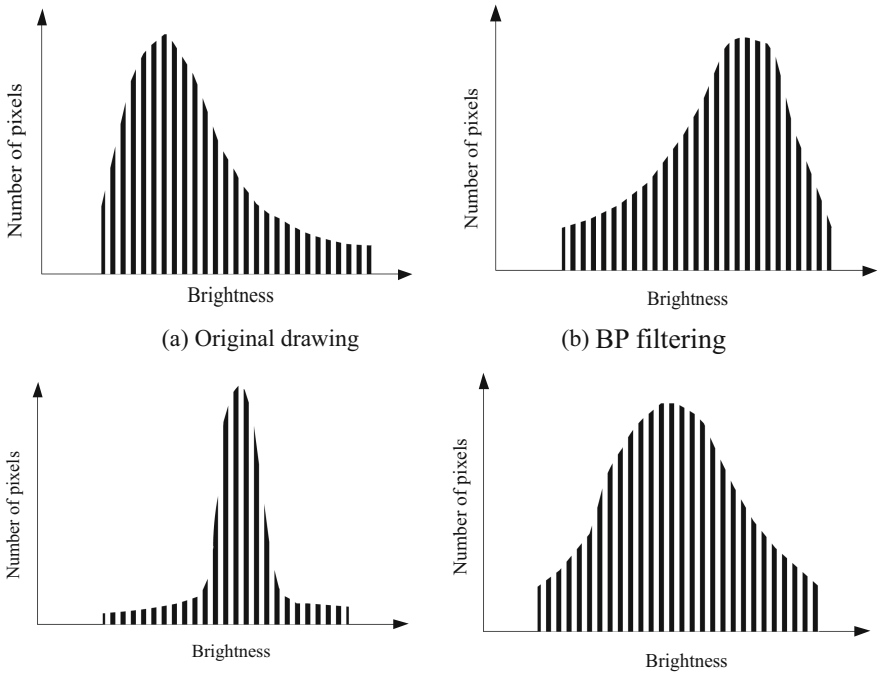


Fig. 5. Image histogram under different methods

From Fig. 5, it can be seen that the original image has low brightness and poor brightness; After Laplace's method processing, the brightness of the image is significantly improved, but the brightness is too bright; After using edge enhancement guided filtering method, the image brightness is too concentrated and the image processing effect is not good; After the processing of the method in this article, the histogram distribution of the image is relatively uniform, achieving the best image processing effect.

4 Conclusion

3D information features of joint concrete crack images, the concrete crack image processing is implement through filtering and testing methods. The article proposes an automatic focusing fusion method of aconcrete crack images based on depth learning. The feature matching of concrete crack images under information execution enhancement

ability, characteristic segmentation model of remote geophysical remote sensing concrete crack images is built, and the matching with features of concrete crack presentation is carried out under the message enhancement technique, The spatial matching function of concrete crack image under remote geophysical exploration and remote sensing is obtained. In shape distribution blocks of concrete crack image pixel clusters, combined with shape features, the focus fusion model of concrete crack image is obtained in the multi-scale input image to achieve automatic focus fusion of concrete crack image. The experimental results show that:

(1) The solution in the article can be well implemented the automatic focus fusion of remote geophysical remote sensing digital images, and improve the quality of digital imaging.
(2) The signal-to-noise ratio of the method in this article can reach 54.34 dB for automatic focus fusion of far digital images, effectively improving the image quality.
(3) After processing by this method, the image histogram distribution is relatively uniform, indicating that the image processing effect of this method is the best, which is neither bright nor dark.

Although the methods in the article include made good progress in image fusion, the image fusion process is relatively complex and the fusion time is too long. Therefore, how to shorten the integration time will be my research direction.

References

1. Zhou, H., Zhao, L.H., Liu, H.: Research on image restoration methods for global optimization of damaged areas. Comput. Simul. **37**(09), 469–473 (2020)
2. Mao, Y.P., Yu, L., Guan, Z.J.: Multi-focus image fusion based on fractional differential. Comput. Sci. **46**(S2), 315–319 (2019)
3. Wu, Y.C., Wang, Y.M., Wang, A.H.: Light field all-in-focus image fusion based on edge enhanced guided filtering. J. Electron. Inf. Technol. **42**(09), 2293–2301 (2020)
4. Zhai, H., Zhuang, Y.: Multi-focus image fusion method using energy of Laplacian and convolutional neural network. J. Harbin Inst. Technol. **52**(05), 137–147(2020)
5. Zhao, D.D., Ji, Y.Q.: Multi-focus image fusion combining regional variance and EAV. Chinese J. Liq. Cryst. Displays **34**(03), 278–282 (2019)
6. Zeng, Z.X., Liu, J.: Microscopic image segmentation method of C.elegans based on deep learning. J. Comput. Appl. **40**(05), 1453–1459 (2020)
7. Cao, J., Chen, H., Zhang, J.W.: Research on multi-focus image fusion algorithm based on super resolution. Comput. Eng. Appl. **56**(03), 180–186 (2020)
8. Chen, Q.J., Wang, Z.B., Chai, Y.Z.: Multi-focus image fusion method based on improved VGG network. J. Appl. Opt. **41**(03), 500–507 (2020)
9. Xie, Y.X., Wu, Y.C., Wang, Y.M.: Light field all-in-focus image fusion based on wavelet domain sharpness evaluation. J. Beijing Univ. Aeronaut. Astronaut. **45**(09), 1848–1854 (2019)
10. Liu, B., Han, G.L., Luo, H.Y.: Twin convolution neural network image fusion algorithm based on multi-scale details. Liq. Cryst. Disp. **36**(09), 1283–1293 (2021)