Edge tracking method of damaged mural images based on deep learning

ZHANG Li-wei[1],FENG Huan-ping[2]
{lihaixia09@sina.com[1],aini2125@tom.com[2]}
(1. Heibei Institute Of Communication College,shijiazhuang 051430,China;
2.Heibei Institute Of Communication College,shijiazhuang 051430,China)

Abstract: The traditional edge tracking method of damaged mural images has the phenomenon of high noise variance in the process of tracking, which affects the actual performance of tracking. Therefore, an edge tracking method based on deep learning is proposed.Through image preprocessing, the image gray mean value was unified and image quality was enhanced. Roberts edge detection operator was used to detect image edge features. RCF model in deep learning was used to fuse edge features and output them to find the starting point of fusion features.The test results show that the noise variance of the edge tracking method for damaged mural images designed based on deep learning is between 0.010 and 0.015, which is lower than the noise variance generated in the traditional tracking method, indicating that this method is better than the traditional tracking method.

Keywords: Deep learning; Image edge; Edge tracking;

1 Introduction

As the mural has a long history, the wall is not conducive to the long-term preservation of the mural, but also has the influence of natural disasters and human factors, which has caused various damages, such as fading, discoloration, falling off, etc., so it is important and necessary to protect the mural. When using different methods to repair murals, the edge tracking of damaged murals is a very important step[1-2].

From the perspective of the image boundary tracking implementation method, the traditional boundary tracking methods are mainly divided into two categories: one is based on run length, and the other is based on chain code. The run-length tracking method is used to divide multiple target areas into run-length representations in a row, and analyze the relationship between adjacent runs to obtain the area outline. The chain code-based tracking method is relatively intuitive. It only needs to track on the edges of the image, which is more efficient. However, for multi-region and multi-connected images, the above two methods are prone to excessive noise variance. In reference [3], a target tracking scheduling method based on wireless sensor networks is proposed. Considering the characteristics of wireless sensor

EAI MOBIMEDIA 2020, August 27-28, Harbin, People's Republic of China

DOI 10.4108/eai.27-8-2020.2295946

networks and the requirements of extended clarity of monitoring system functions, the original LCFs algorithm is improved, and a simple and effective target detection and tracking scheme is designed and implemented, A scheduling algorithm (dolb) for the definition optimal load balancing is proposed. However, the noise variance of this method is high in the process of image tracking.

In view of the problems of the above methods, this paper proposes an edge tracking method of damaged mural image based on deep learning. By using the hidden nodes in deep learning and feature learning, the purpose of feature fusion is achieved, and then the purpose of edge tracking of damaged image is solved.

2 Edge tracking method of broken mural image

2.1 Raw image preprocessing

Low-quality images will inevitably be produced due to the effects of damaged mural images due to insufficient light sources during the shooting process,In order to improve the visual effect of such images and the quality of the images, a series of enhancement processes are performed on such images.

Convert the original image into a grayscale histogram and set an initial label for each gray level of the grayscale image histogram in order to find a local maximum,to set the initial label, establish constraints:

$$y(\alpha)=\begin{cases}-1 & x(\alpha)<x(\alpha-1)\\ 1 & x(\alpha)\geq x(\alpha-1)\end{cases} \quad (1)$$

Where $y(\alpha)$ is the label on the α gray value, and $y(\alpha)$ is the number of pixels on the α gray value.

The above process is that if the total number of pixels on the α -th gray value is greater than or equal to the total number of pixels on the $\alpha-1$ gray value, the a-th label is set to 1, otherwise it is set to -1. Set all possible values of α satisfying $y(\alpha-9)=1,\ldots,y(\alpha+1)=1$ and $y(\alpha+1)=-1,\ldots,y(\alpha+9)=-1$ conditions as the local maximum we are looking for[4]. As shown in Figure 1.

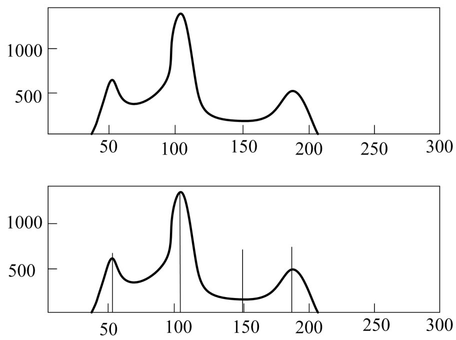

Figure 1 Histogram selection segmentation value

It can be seen from the figure that there are 4 extreme points that meet the conditions in the histogram, but the difference between the third extreme point and the minimum points on both sides is very small. In order to better enhance the image, these relative points need to be removed. Inconspicuous extreme points, so we set a condition, that is, the extreme point that meets our set conditions is the final optimal segmentation point,The conditions are as follows:

$$\begin{cases} \dfrac{\left(x(\alpha)\text{-}u(\alpha-1)\right)}{T} \geq \eta \\ \dfrac{\left(x(\alpha)\text{-}u(\alpha+1)\right)}{T} \geq \eta \end{cases} \quad (2)$$

In the above formula, $x(\alpha)$ indicates the number of pixels of the maximum point, and $u(\alpha+1)$ indicates the number of pixels of the interval minimum point.T represents the size of the gray interval, and the parameter η is the set threshold. The intensity of the false extreme points can be removed by setting the value of η [5].According to the obtained optimal segmentation value, the gray interval of the original image is divided, and a certain enhanced gray interval is reassigned for each gray interval. The interval mapping function is as follows:

$$\begin{cases} Q_i = h_i - l_i \\ t_i = \dfrac{Q_i * \log_{10}\left(\dfrac{S_i}{Q_i}\right)}{T} \\ T = \sum_{i=0}^{n} Q_i * \log_{10}\left(\dfrac{S_i}{Q_i}\right) \end{cases} \quad (3)$$

In the above formula, Q_i represents the range of the gray interval of the i -th sub-histogram, and h_i represents the maximum gray value of the gray interval. l_i represents the minimum grayscale value of the i -th interval, S_i is the total number of pixels in the grayscale interval, and t_i is the size of the redistributed grayscale interval of the i th sub-histogram[6].Assume that the gray interval of the sub-histogram of the first output image is $[0, t_i]$, then the range of the i th gray interval can be obtained by the following function:

$$\begin{cases} start_i = \sum_{\alpha=0}^{i=1} t_\alpha + 1 \\ end_i = \sum_{\alpha=0}^{i} t_\alpha \end{cases} \quad (4)$$

$start_i$ represents the starting gray value of the i th gray interval, that is, the minimum gray value; end_i represents the maximum gray value of the i th interval.If $i = 1$ is the first interval, then $start_i = 0$; if i is the last interval, end_i is 255.

In order to further ensure the stability of the gray value of the gray image after the enhancement process, the histogram of each interval is cut to prevent the pixels in the gray image from being too concentrated in a small gray interval Causes excessive enhancement in the gray interval that takes up too much in the mapping process, and processes the histogram of each interval.According to the segmented histogram and segmentation value obtained in the previous step, the shear closure value of each interval is calculated respectively. The shear threshold of each interval can be calculated according to the following formula:

$$C_i = \frac{1}{\alpha_i - \alpha_{i-1}} \times \sum_{s=\alpha_{i-1}}^{\alpha_i} x(s) \quad (5)$$

In the formula, C_i is the calculated threshold for each interval, where α_i represents the maximum gray value of the i th interval, α_{i-1} represents the minimum gray value of the gray interval, and $x(s)$ represents the total number of pixels in the gray interval.The process of clipping the input gray image histogram is to perform the following processing on the number of pixels at each gray level according to the clipping threshold of each interval. If the number of pixels is less than the clipping threshold, the gray level Value unchanged, otherwise change the value of this gray level to the clipping threshold[7].

According to the histogram processed by the segmentation interval and the peak clipping operation, an independent histogram grayscale unification operation is performed on the sub-images of each interval separately to maintain the stability of the gray average value before and after the gray image processing,Assume that the gray value of the input gray image $P(u,v)$ is ω_i. For each pixel of the output image, multiply its gray value by an offset coefficient, that is, do the following processing:

$$P^{'}(x,y)=\omega_i P(u,v) \quad (6)$$

In the formula, $P^{'}(x,y)$ represents the final enhanced image. After the above formula is processed, the gray average value of the output image can be further adjusted to be close to the average gray value of the input image to achieve the purpose of enhancing image quality.

2.2 Detect image edge features

In the process of image edge feature detection, the processed image is regarded as a two-dimensional matrix, and the points in the matrix are regarded as sample points of the continuous change function of the gray intensity of the image[8].The edge of the image is the set of points where the gray value changes significantly. Therefore, the edge of the image can be determined by obtaining the first derivative of the gray intensity function.The magnitude of the gradient is expressed as the intensity of the edge, and the direction of the gradient is perpendicular to the direction of the edge. In general, the gradient magnitude of a continuous image function $f(u,v)$ is expressed as:

$$grad\left(f(u,v)\right)=\left[\frac{\partial f(u,v)}{\partial x},\frac{\partial f(u,v)}{\partial y}\right] \quad (7)$$

Since the pixels in the image are discrete, the partial differential in the formula can be approximated by difference. The simple gradient approximation is:

$$\begin{cases} f_u = f[u,v+1] - f[u,v] \\ f_v = f[u,v] - f[u+1,v] \end{cases} \quad (8)$$

To quote Roberts edge detection operator, calculate the difference of pixels in the diagonal direction in the field of 2 * 2:

$$\begin{cases} f_u = f[u,v] - f[u+1,v+1] \\ f_v = f[u,v+1] - f[u+1,v] \end{cases} \quad (9)$$

The gradient of Roberts edge detection operator in the image $f(u,v)$ is:

$$g(f(u,v)) = \max\{f(u,v) - f(c,s)\} \quad (10)$$

In the formula, $f(c,s)$ represents the four domain points of the pixel (u,v). Its convolution template in the u and v directions is shown below.

1	0
0	-1

0	1
-1	0

Figure 2 Convolution template of edge detection operator

According to these two convolution templates, the Roberts gradient amplitude $r(u,v)$ can be calculated. An appropriate threshold is set to classify the image pixels. When the amplitude is greater than T, it is an edge feature point[9].Roberts edge detection operator uses the gray difference between two adjacent pixels on the diagonal to approximate the edge features of the image. There are some breakpoints in the detected edge features. Deep learning technology is used to fuse the edge features of the image.

2.3 Design of image edge feature fusion based on deep learning

The RCF model in deep learning is used to output fusion features. The internal structure of the first stage of the model is shown in the figure below.

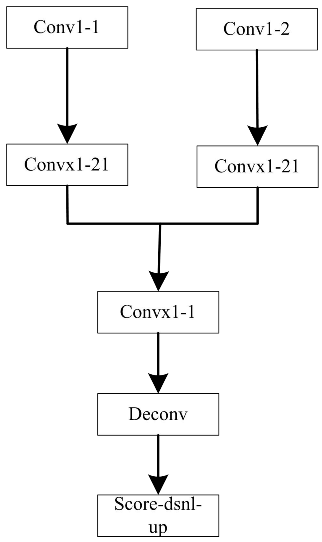

Figure 3 internal structure of RCF model

The figure shows the feature fusion mode of the RCF model in phase 1. In the figure, Convl 1 and Convl 2 represent the first two convolutional layers of the first stage. Convlxl-1 and convlxl-21 respectively represent convolutional layers with convolution kernel size of 1x1 and output channel of 1 and 21[10]. Deconv represents a deconvolution layer, also called a transposed convolution layer, whose role is to upsample features to the original picture size.

Using this model, using a multi-level feature fusion strategy, feature extraction is performed as a level feature for each stage, and it is finally used for the detailed feature compensation of the final prediction output. The fusion of hierarchical features and preliminary classification features, the fusion results are shown in the following figure.

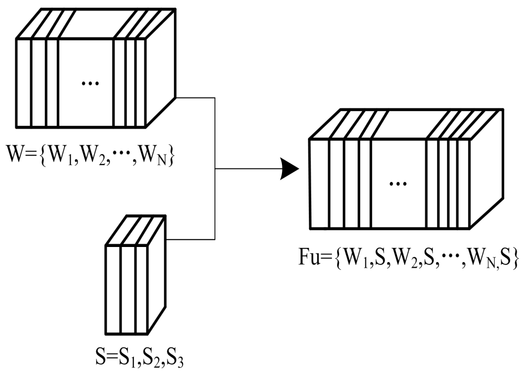

Figure 4 Multi-level feature fusion

The result of the first stage fusion shown in the figure above, the stacked feature fusion is used in the second stage, so that the fused features can contain more effective detailed information, and the features between the stages are further fused. The stacked structure is shown in Figure 4.

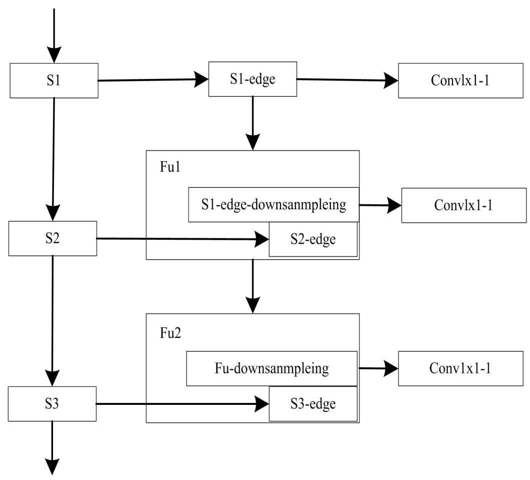

Figure 5 Phased feature fusion of the stacked structure

In Figure 5, S1 represents the first stage of feature extraction in the model, and S1-edge represents the hierarchical features extracted at this stage. The specific stacking method is: the next stage feature is first stitched all the above stage features, and then a Convlxl-1 convolution layer is used for feature fusion.Because it is necessary to ensure that all feature sizes are consistent when stitching feature channels, the features of the upper layer must be down-sampled to the same size as the stitched layer before stitching.In order to save the important information of the subject in the original feature during the downsampling, the downsampling method uses the maximum pooling operation to obtain the final fusion feature.

2.4 Implement image edge tracking

The edge tracking of the damaged mural image first extracts the starting point in the fused feature. This starting point must be the end point of the extracted edges. The set edge tracking strategy is tracking from small scale to large scale, so the starting point is searched for the fused features at the smallest scale.

Let e_i be an edge point in the fusion feature at the minimum scale. If there is only one edge point in the neighborhood of e_i, then e_i is considered as a starting point.Suppose e_i is a point in the fused feature. Let $E(e_i)$ be the parent node corresponding to e_i in the gradient image. Define a detection operator $\gamma(e_i)$, if e_i is a local maximum point, $\gamma(e_i)=1$; otherwise, $\gamma(e_i)=0$.

Define the tracking indicator operator $r(e_i,e_j)$:

$$\begin{aligned} r(e_i,e_j) &= \gamma(e_i)\wedge\gamma(E(e_i)) \\ &\wedge\gamma(e_j)\wedge\gamma(E(e_j)) \end{aligned} \quad (11)$$

Among them, $\wedge$ represents a logical AND operator. If $r(e_i,e_j)=1$, it indicates that the edge point e_j after the tracking algorithm is selected to be connected to e_i. At this time, there is a path between the edge points e_i and e_j;Otherwise, point e_j is treated as a non-edge point and is set to 0. In $r(e_i,e_j)=1$, the path between edge points e_i and e_j is shown in the figure below.

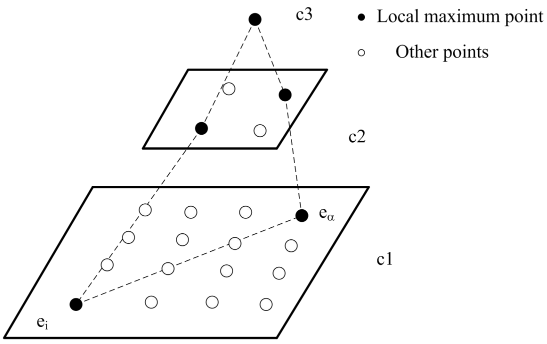

Figure 6 Image edge point tracking

If the path of two non-adjacent edge points is finally obtained from the above process, the purpose of image edge tracking is achieved. At this time, the edge tracking method of the damaged mural image based on deep learning is designed.

3 Simulation test of edge tracking method of broken mural image

3.1 test environment and data set

The test hardware environment is Intel Xeon CPU E5-2620 2.10GHz, 256DDR4 memory, 1TB solid-state hard disk; the software environment is MATLAB software, Caffe deep learning framework.

In order to verify the actual performance of the designed edge tracking method based on deep learning for broken mural images and the traditional tracking method. A unified SBD data set is adopted, which is an open source data set. This dataset contains more than 20,000 image data, which belong to 20 categories. The original image data in the SBD dataset is from the VOC2011PASCAL dataset.

3.2 Test data training hyperparameter settings

Because the image data in the SBD data set is too large, before the parameter setting, the data set required for testing is reasonably divided, and 8498 of them are used as the training set and 2857 are used as the test set.At the same time, the data was enhanced and the data in the SBD data set was scaled down and scaled to achieve the effect of increasing data samples.At the same time, the above data is divided into ten groups in an increasing relationship for future testing.

The mean reduction operation is performed on the data in the data input layer to ensure that the network speeds up the convergence of each layer's weight when backpropagating.At the same time, in order to enhance the generalization ability of the data, the input is also subjected to random mirror flip processing. After random flip, each input is not a fixed value, so that the tracking method based on deep learning has a different input situation. The better the learning, the more robust the data obtained after training.

The data was randomly cropped with a fixed size of 472x472. This size can be adjusted according to the consumption of resources, but it must be a multiple of 8, because the input is continuously down-sampled during the test. Each stage will down-sample to half of the previous stage, so a multiple of 8 is good for data processing.

Set the maximum number of training iterations on the SBD dataset to 22000/08000 times. The basic learning rate is set to 2.5e-8, the weight update impulse is 0.9, and the weight attenuation is 5e-4.After the setup is completed, test the variation of noise variance in the process of tracking the edge of the damaged image with different tracking methods.

3.3 Noise Variance Test Results

Different edge mural tracking methods were used to test the variance of noise generated during the tracking process in the same environment.Using third-party software to statistically test the results, the results are shown below:

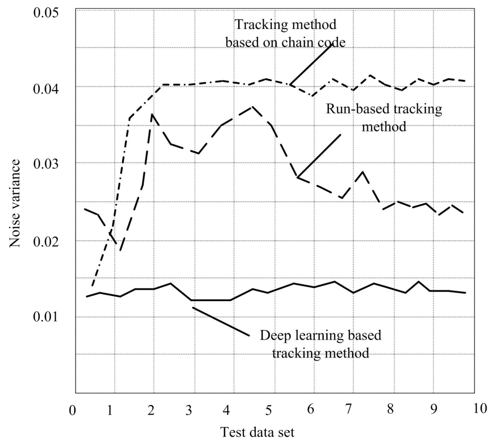

Figure 7 Noise variance test results of different tracking methods

Because the amount of data in the ten test groups has an increasing relationship, the data amount gradually increases during the test by default. Observing the results in the figure, it shows that using the chain code-based tracking method, as the amount of data increases, the noise variance gradually increases, reaching the peak in the second set of tests, and has been in a higher position since then, without a downward trend;The test results based on the run-length tracking method show that with the increase of the amount of data, the noise variance fluctuates greatly, there is no obvious law, the overall fluctuates between 0.01 ~ 0.04, and most of them are in a higher position; the tracking method based on deep learning shows that as the amount of data increases, the noise variance does not change significantly, always between 0.010 and 0.015, which is in a lower position.

In summary, the noise variance of the edge tracking method of the damaged mural image based on deep learning is lower than the other two methods, which indicates that the designed tracking method is more suitable for edge tracking of the damaged mural image.

4 concluding remarks

The generation of edge tracking method for broken mural images is of great significance for the repair and protection of broken murals.With the development of technology, traditional tracking methods can not meet the actual needs of today's edge tracking of broken murals. Therefore, deep learning technology is used to design a method for tracking broken mural images based on deep learning, and fuse images through network models in deep learning Edge features to reduce the variance of noise generated during the tracking process.The designed comparative test proves that the method in this paper can effectively solve the problems existing in the traditional method, and has certain practical significance for the subsequent development of image edge tracking methods.

References

[1] Chen ,S.W, Liu ,Y.J.Liu, D.: AlexNet Model and Adaptive Contrast Enhancement Based Ultrasound Imaging Classification . Computer Science.Vol.46,pp.146-152(2019)

[2] Qiu ,F. Hou ,F. Yuan,Y.: Blind image deblurring with reinforced use of edges . Journal of Image and Graphics.Vol.24,pp.847-858(2019)

[3] Mao, J.: A Video Target Tracking Algorithm Based on Wireless Sensor Networks. Control Engineering of China. Vol.25,pp.71-75(2018)

[4] Zheng, Y.P. Li, G.Y. Li ,Y.:Survey of Application of Deep Learning in Image Recognition. Computer Engineering and Applications.Vol.55,pp.20-36(2019)

[5] Tang ,C. Ling ,Y.S. Zheng, K.D.: Object detection method of multi-view SSD based on deep learning. Infrared and Laser Engineering.Vol.47,pp.302-310(2018)

[6] Lang ,X.T.: Frame Error Eliminating Technology in Moving Image Tracking Process . Science Technology and Engineering.Vol.19,pp.227-232(2019)

[7] Yang ,G.W. Jia ,Y.R.: Accurate Extraction and Simulation of Characteristic Target of Meteorological Satellite Remote Sensing Image . Computer Simulation.Vol.35,pp.372-376(2018)

[8] Li ,L.J, Liu ,X.H. Huang ,X.X.: Multispectral Image Registration Based on Symmetrical Edge. Semiconductor Optoelectronics.Vol. 40,pp.395-400(2019)

[9] Tian ,W. Hu ,Y.R.: Color image edge extraction algorithm based on co-occurrence matrix coupled with Otsu threshold. Journal of Electronic Measurement and Instrumentation.Vol. 32,pp.52-60(2018)

[10] Mao ,T.Q. Liu ,W. Huang ,J.: Change Detection of SAR Images Using Dyadic Wavelet Enhancement and Edge Local Information FCM . Journal of Signal Processing.Vol.34,pp.54-61(2018)