Research on real-time cloth edge extraction method based on ENet semantic segmentation

Xiaofei Gao, Qi Lizhe, Ma Chuangjia and Yunquan Sun

Abstract
A vision-based sewing trajectory extraction method aims at the difficulty of adaptive generation of sewing trajectory in garment automatic processing. Firstly, the ENet model is improved and the I-ENet semantic segmentation algorithm is proposed; Then, based on the semantic segmentation results, a dynamic edge extraction method based on the GaussMod fitting method is proposed. Based on segmented images, curve fitting, median filtering, and edge-preserving filtering are used. Finally, the Canny operator is used to find the sewing edge trajectory curve. At last, the sewing trajectory curve of the robot is generated. Through experimental verification, the MIoU and PA of the semantic segmentation algorithm I-ENet proposed in this paper reach 95.06% and 97.80% respectively. Compared with the MIoU of the original model of ENet, the MIoU is improved by 2.78%, the pixel accuracy is improved by 1.18%, and the frame rate is 30 FPS. It can realize cloth segmentation and extraction in an unstructured environment. The maximum error between the sewing edge trajectory and the actual edge is 1.35 mm, the mean error is 0.61 mm, and the mean value of dynamic variance is 0.57 mm. This method meets the practical application requirements.

Keywords
Semantic segmentation, image processing, edge detection, cloth

Introduction
Clothing is one of the essential items in people’s life. With the increase in labor cost year by year, garment manufacturing automation has become a hot issue in this industry. In upgrading garment manufacturing automation, simple automatic sewing equipment is challenging to meet the actual needs. Therefore, the mode of closed-loop operation based on the interaction between visual perception information and automated equipment has increasingly become a research hotspot. However, in the actual sewing scene, affected by the light, background, and external light, the traditional image processing algorithms lack certain adaptability. For example, the first-order edge detection operators such as the Sobel operator and the second-order edge detection operators such as the Laplacian operator and the Canny operator. But all of these cannot accurately extract the edge information of cloth. Because these detection algorithms only consider the sharp local changes of the image, especially the sharp changes in color, brightness, and other features. These features are detected to find cloth edges. However, these
features are difficult to highlight in more complex scenes, so the traditional algorithms lack certain generalizations and stability.

Based on the limitations of traditional algorithms, many scholars try to extract edge features by deep learning. The most significant difference is that deep learning can learn edge feature information from image data by itself because the depth model can extract feature information from pixel-level data. It provides a new idea and method for solving computer edge extraction. Shen et al.\(^5\) proposed the Deep Contour edge detection algorithm, which can highlight the image features through clustering to form labels. It can fit different edge contours through different model parameters. Through experimental verification, this method has a simple structure and easy implementation but lacks self-adaptability. Liu and Lew\(^6\) proposed an RDS algorithm, which uses different depth monitoring, integrates the diversity of the network, and has high detection accuracy. Liu et al.\(^7\) proposed a full convolution neural network model, which uses the idea of FPN to carry out edge detection combined with deep and shallow feature mapping. It optimizes the loss function at the same time, which is more universal. In the same year, Xu et al.\(^8\) introduced the Attention-Gated Conditional Random Fields and defined a new multiscale hierarchical depth network model based on attention guidance for edge detection. The network can learn deeper multiscale feature information. Deng et al.\(^9\) proposed an LPCB detection method to solve the problem of thick edges in edge detection. This method is an end-to-end network architecture of a bidirectional system. The network can effectively use the network hierarchy information to generate a boundary mask with high pixel accuracy, to obtain more accurate edge information: He et al.\(^10\) proposed a multi-scale network model BDCN and introduced the scale enhancement module. The edge detection effect has been greatly improved through the combination of the two. However, the above algorithms have the problem of a large amount of calculation and slow recognition speed. Orujov et al.\(^11\) proposed an image processing for contour detection algorithm based on Mamdani (Type-2) fuzzy rules, which mainly uses Mamdani (Type-2) fuzzy rules to detect the edge of image gradient values. The overall architecture has certain universality. Li et al.\(^12\) proposed a new multiscale edge detection method based on the multi-stream structure and image pyramid principle, which can solve the problems of fast sample training and accurate testing. This method constructs a down sampling pyramid network and lightweight up sampling pyramid network, respectively, to enrich the multiscale representation of the encoder and decoder. Experiments show that this method can double the training speed without a great change in accuracy.

In addition, the loss function and data set in deep learning are also studied to obtain higher detection accuracy, generalization, and robustness.\(^13\)\(^15\) Although the edge detection method based on deep learning has made remarkable achievements, it has not been studied in garment sewing, especially image processing in the sewing process. This process has certain requirements for edge detection’s processing speed and accuracy. The processing speed of VGGNet, Google Net, ResNet, and other network models studied is currently slow. Large amounts of calculation cannot solve the problem of image processing under high-speed sewing. Therefore, this paper analyzes fabric image characteristics and the actual sewing processing needs. A real-time fabric edge extraction method based on ENet image semantic segmentation is proposed to complete the deviation between the fabric edge curve and fabric in the sewing process. Keep the edge information of fabric as much as possible to help fabric sewing track planning and automatic deviation correction.

Method

ENet semantic segmentation network\(^16\) is a commonly used segmentation network with simple network architecture, fast running time, and few parameter variables. So it can be applied to real-time image segmentation and mobile devices. Compared with other conventional semantic segmentation networks, such as FCN and SegNet, the running time and memory of the ENet are reduced by dozens of times, respectively. Because the application scenario of this paper is image segmentation and detection in the state of high-speed sewing, the network model is selected as the basic neural network architecture. The model is optimized and improved to meet the needs of real-time and accuracy. To extract the edge of the cloth and the deviation of double-layer fabric, a new image processing algorithm combining the neural network and the traditional algorithms is proposed to obtain the edge curve and deviation value of the cloth. The technical flow chart of this paper is shown in Figure 1.

**Improvement of the semantic segmentation model**

As a typical lightweight semantic segmentation model, the ENet network has excellent real-time performance. Still, the segmentation accuracy is not high, especially when the cloth is similar to the background color. To improve the accuracy of ENet network segmentation to ensure real-time performance, we widen the last stage of the encoder in the original model of ENet, that is, double the number of channels. The improved model is pruned to ensure real-time performance while improving the segmentation accuracy. Width is one factor affecting the neural network’s expression ability. In contrast, the network layer with larger width can learn more rich texture, color, and other characteristics. It can also improve the prediction performance of the network and better distribute the material and background.
The improved model architecture is shown in Figure 2. The transparent module represents the clipped layer. The green module indicates the added layer. Multiple identical convolution operations on the same scale feature map will repeatedly extract feature information, resulting in certain information redundancy. Therefore, we cut stage 1 (the yellow part of Figure 2) of the original model and deleted two modules ($64 \times 128 \times 128$), which reduces the number of parameters of the original model – assuming that the size of the network input data is $3 \times H \times W$. After passing through the original encoder of ENet, the output size is $128 \times (H/8) \times (W/8)$. After the original encoder, we added a downsampling module to stage 3 (the light blue part of Figure 2) following stage 2 (the dark blue part of Figure 2) to further reduce the resolution of the feature map of the coding layer and increase the number of channels. The final output size is $256 \times (H/16) \times (W/16)$, increasing its width. The network has a larger receptive field in the convolution process of the feature map, combined with more context information, and learns richer, specific, and effective texture features. To maintain the symmetry of the network structure, we add an upsampling stage to the decoder. The improved network structure is named I-ENet (Improved ENet).

### Design of edge extraction algorithm

This section mainly processes the segmented image and extracts the cloth’s edge and deviation information. If the segmented image is directly edge detected using the traditional edge detection operator, there is more noise interference in the results, and it isn’t easy to filter. It is easy to hurt the automatic sewing operation using the edge information.

The traditional edge detection methods are divided into three categories. The first category is the classical edge detection methods, such as the differential operator...
method, optimal operator method, and suitable method. The second is the global extraction method based on energy minimization, which is characterized by using rigorous mathematical methods to analyze this problem. It gives a one-dimensional cost function as the optimal extraction basis and extracts edges from the perspective of global optimization, such as the relaxation method. The third category is image edge extraction methods developed recently, such as wavelet transform, mathematical morphology, fuzzy mathematics, and fractal theory. Especially the method of image edge extraction based on multiscale wavelet transformation is a more research topic at present. Among them, each of the above categories contains a variety of different edge detection algorithms.

To reduce the influence of the overall noise of the segmented image on the edge feature extraction, a dynamic ROI segmentation method is proposed according to the obtained pixel list of the outer contour of the cloth. The result is shown in Figure 4. The image read by OpenCV is shown in Figure 4(a).

The cloth sewing edge area concerned by this task is always below the image. The dynamic ROI setting method can change according to the shape and angle of the cloth and effectively reduce the interference of outside areas in the process of extracting edge information. The process of setting dynamic ROI is as follows:

(a) Binarization

Because the segmented image features are relatively simple, the commonly used fixed threshold method is used for binarization. The basic idea is shown in Formula (1):

$$g(x,y) = \begin{cases} 255, & f(x,y) \geq T \\ 0, & \text{others} \end{cases}$$  

Where $T$ is the global threshold.

The result of binarization is shown in Figure 4(b).

(b) Nonlinear fitting

The coordinates of the image in (a) are extracted. It can be seen from Figure 3(a) that the edge of the segmented image is similar to a non-equilateral closed quadrilateral, and there exists “one-to-many” data. So effective linear fitting cannot be carried out. Since the target object of this paper is the sewing edge, the data is arranged in clockwise order. Half of the total number of pixels in the pixel list is selected as the scattered point data for curve fitting, as shown in Figure 3(b). The basic idea of the nonlinear fitting through GaussMod is shown in formula (2):

$$f(x) = y_0 + (f_1 * f_2)(x)$$

$$= y_0 + \frac{A}{t_0} e^{-\left(\frac{1}{2}\right) \left(\frac{x-x_e}{t}\right)^2} \int_{-\infty}^{\infty} e^{-\frac{1}{2\pi} \left(\frac{y-z}{2}\right)^2} dy$$  

Where $f_1(x) = \frac{A}{t_0} e^{-\frac{x-x_e}{t_0}}$, $f_2(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-x_c)^2}{2\sigma^2}}$, $z = x - x_c - \frac{w}{t_0}$, $A, x_c, y_0, w$, and $t_0$ are constants.

The fitted curve is shown in Figure 3(c).

(c) Determining the tangent centerline of the cloth edge

The derivative of the curve fitted in (b) is obtained to get the position coordinate where the curvature of the curve changes the most. It can form a linear relationship with the endpoint of the curve to form the center line $g(x)$ of the dynamic ROI, as shown in Figure 4(c):

$$g(x) = \frac{f(x) - f'(x_{\text{max}})(x-x_{\text{max}})}{x_e - x_{\text{max}}} + f'(x)_{\text{max}}$$  

Where $x_e$ is the largest point in the scatter point and $x_{\text{max}}$ is the coordinate point where the curvature change is the largest.
(d) Take the straight line obtained in (c) as the center line, offset upward with an offset distance of 100 pixels, and then bias downward. The offset straight line is the straight line passing through the farthest point of the image pixel coordinate to ensure that the cut image can include all data.

(e) A binary operation is performed on each corresponding pixel in Figure 4(a) and (e) to obtain the dynamic ROI of the image.

After obtaining the image shown in Figure 3(f), it is necessary to extract the edge of the image. Through the edge-preserving filter method provided by OpenCV, it was carried out to avoid serious loss of edge information of the two pieces of cloth after noise filtering. And then, it is processed by a median filter with a kernel size of 3. Then the edge information is extracted by using adaptive threshold binarization. Finally, the median filter is used to filter the scattered noise, and the outer edge of the fabric and the boundary between the two fabrics are detected. The overall process is shown in Figure 5.

**Experiment**

**Construction of experimental platform**

The experiment is carried out on the same hardware configuration and software framework. The experimental device is shown in Figure 6. The camera used in the experiment is Mindvision (Model: MV-SUA501GC/M-T), a 5 million pixels color camera. The resolution and frame rate is 2448 × 2048 and 40 fps. The experimental environment is Windows 10 operating system (Intel i5-10400f processor, GeForce GTX 1660 SUPER 6GB graphics card). Python version is 3.8, and the deep learning framework is pytorch1.7.0.
We set the optimizer to use the Adam algorithm to train the network model, use a single GPU, set the learning rate to 5e-4, L2 weight attenuation to 2e-4, batch size to 4, and select the cross-entropy loss function as the loss function to help the network converge faster. Limited by the size of the video memory, the adopted image resolution is 640 × 480. The training set and verification sets are sent to the I-ENet network for training, and the training effect is compared with that of the ENet network.

**Data set**

In this paper, the data set is collected in the laboratory and collected by an industrial camera. To make the trained network model have better generalization ability and reduce the influence of different colors, different illumination, and shooting angles on the segmentation structure, the collected image should include all kinds of clothes and backgrounds as much as possible. In this paper, the pure color cloth is taken as the segmentation target of the experiment. Firstly, pictures of the sewing process were collected in a laboratory environment. Some of the images contained distractors similar in color to the fabric. Secondly, images in different background environments are collected. This is shown in Figure 7. The resolution size of the collected images is 2448 × 2048, and the saving format is JPEG.

In addition, we expand the collected data set through data enhancement methods, such as random erase and flip et al., to improve the generalization ability and robustness of the model and achieve a better cloth segmentation effect. Figure 8 is an example of an enhanced image. In the experiment, we use labelme for data annotation. The contents of the image are uniformly divided into cloth and background. The cloth is marked in red, and the background is marked in black.

**Evaluation criterion**

In this experiment, there is only one segmentation object of cloth, so the actual number of categories for semantic segmentation is 2 (cloth and background). We select pixel accuracy (PA) and mean intersection over union (MIoU) as the model performance evaluation indicators.

PA is the most direct measure to measure the effect of semantic segmentation. Its meaning refers to the proportion of the sum of the number of pixels correctly predicted and inferred by the network in all pixels. The calculation formula of PA is shown in formula (4):

$$PA = \frac{\sum_{i=0}^{k} P_{ii}}{\sum_{i=0}^{k} \sum_{j=0}^{k} P_{ij}}$$

Where $k$ is the number of categories, $p_{ii}$ indicates the number of the label category is $i$, and classified as $i$, that is, the number of pixels correctly classified. $p_{ij}$ indicates the number of the label category is $i$, but classified as $j$, that is, the number of pixels incorrectly classified. The denominator represents the total number of pixels.

MIoU is the most commonly used measure of semantic segmentation algorithm, which means the average value of
the sum of the ratio of the predicted result of each classification to the intersection and union of real categories.\textsuperscript{24} The calculation formula of MIoU is shown in formula (5):

\[
MIoU = \frac{1}{k + 1} \sum_{i=0}^{k} \sum_{j=0}^{k} \frac{P_{ij}}{1 + \sum_{j=0}^{k} P_{ij} - P_{ji}}
\]

(5)

Where \(P_{ij}\) and \(P_{ji}\) have the same meaning of formula (4). \(P_{ij}\) and \(P_{ji}\) represent false positive and false negative respectively.

To verify the accuracy of edge extraction, this paper analyzes the average error of single frame image \(x\), and determines the accuracy of single image detection in the sewing process. We define the unit pixel area of the label image as \(y\), and the unit pixel area of the segmented image as \(\hat{y}\). If the unit pixel area of the label image is the same as the pixel area of the segmented image, the XOR value is denoted as 0. If not, it’s 1. Then the area of the classified pixel region is calculated, and the error value of a single frame image is:

\[
x_i = \sum_{j=1}^{k} \frac{S_j(r_j)}{N_k}
\]

(6)

Where \(r_j \subseteq y_j \oplus \hat{y}_j\), \(S_j\) is the area of the \(j\)th pixel area. \(N_k\) is the number of pixels in the \(j\)th pixel area. \(k\) is the number of areas.

The stability of detection accuracy in the sewing process is determined by analyzing the dynamic variance of the first \(n\) frames of images. The formula of dynamic variance is:

\[
f(n) = \sum_{i=1}^{n} \left( \frac{x_i - \overline{x}}{n} \right)^2
\]

(7)

Where \(x_i\) is the detection error in the frame image. \(\overline{x}\) is the average value of the image error of the first \(n\) frames. \(n\) is the number of samples.

**Experimental verification**

To verify the effectiveness of this method, we compare different neural network models for the same segmentation task. The data sets shown in Figures 7 and 8 are used for training, and the prediction results of the test set are compared through the network. The evaluation results are shown in Table 1.

| Algorithm     | MIoU (%) | PA (%) | Time (ms) |
|---------------|----------|--------|-----------|
| ENet          | 92.28    | 96.62  | 17.0      |
| DDRNet23      | 80.26    | 91.94  | 18.1      |
| DDRNet23-slim | 87.96    | 94.99  | 10.6      |
| regseg        | 92.23    | 96.90  | 18.6      |
| Deeplabv3+    | 95.74    | 97.32  | 38.5      |
| PSPNet        | 94.09    | 96.30  | 22.7      |
| I-ENet        | 95.06    | 97.80  | 20.0      |

**Table 1.** Performance of different networks on the test set in this paper.
Figure 9. Extract results of the actual sewing process: (a) original image, (b) label, (c) ENet, (d) I-ENet, and (e) I-ENet segmentation diagram.
As can be seen from Table 1, the MIoU and PA values of the I-ENet model reached 95.06% and 97.80%, which increased by 2.78% and 1.18% compared with the original model of it. The improved model adds an upper sampling module and network parameters in the decoding layer, which increases the reasoning time of the model by 3 ms. But it can still meet the speed requirements in practical work. Although the MIoU value of I-ENet is 0.68% lower than Deeplabv3, the time is saved by 18.5 ms. Although the time is a little longer than that of the first four models in the table, its accuracy has been improved accordingly. Therefore, under comprehensive consideration, the I-ENet model is the optimal network model, which provides support for cloth edge detection and helps the manipulator claw drives the cloth to complete automatic sewing. To more intuitively show the effectiveness of the method proposed in this paper, we list and compare the segmentation effects of some test images in ENet and I-ENet networks in Figure 9. Figure 9(a) is the original image collected by the camera. Figure 9(b) is the data label image. Figure 9(c) is the segmentation effect image of the ENet original model. Figure 9(d) is the segmentation effect image of the I-ENet model proposed in this paper. Figure 9(e) is the gray image after segmentation.

It can be seen from the figure that both models can filter out the background, including the characteristics of interferences and shadow parts with similar colors. However, the ENet model has obvious prediction errors under the same training conditions, mainly reflected in the wrong identification of the sewing machine nameplate and cloth edge details. However, the prediction effect in I-ENet has been significantly improved, and the nameplate can be...
filtered out. The segmentation effect of cloth edge detail is also better. In addition, in the sewing process, the illumination with different intensities is changed, which is reflected in lines 3–9 respectively. The original model will be disturbed to varying degrees, while the I-ENet model can better adapt to the influence of different illumination and achieve a good cloth segmentation effect.

After segmenting the original image through the neural network, the segmented gray image, as shown in Figure 9(e) is obtained. The find contours method provided by OpenCV is used to extract the list of cloth contours, which contains all pixel positions constituting the cloth contour. Use the method in the previous section to determine the required ROI, as shown in the third and fourth lines of Figure 10.

Based on selecting ROI, the image is processed successively through an edge-preserving filter and median filter to filter out the noise and retain edge information as much as possible. Finally, the edge information is extracted through adaptive threshold binarization and the median filter, especially the deviation value between two fabrics, is proposed, as shown in the last two lines of Figure 10.

A total of 84 experimental images are collected and screened and brought into formula (6) and formula (7) to obtain the analysis diagram of edge detection error in the sewing process, as shown in Figure 11. From the diagram, it can be concluded that the maximum error in the sewing process is 1.35 mm, the mean error is 0.61 mm, the variance is 0.22 mm, and the mean dynamic variance of the first $n$ images is 0.57 mm, and the variance is 0.11 mm.

Finally, the extracted edge image is directly written to the list of coordinate values through the find contour function in OpenCV and transmitted to the robot controller for use.

**Conclusion**

In this paper, the improved ENet neural network model is trained on the self-built data set. Compared with the original model, it has higher segmentation accuracy, can adapt to the cloth edge segmentation and detection in an unstructured environment, and better retain the cloth edge information. The MIoU is 95.06%, and the PA value is 97.80%. A dynamic edge extraction method based on GaussMod fitting is proposed, which can effectively reduce the noise interference when extracting the cloth edge curve in the sewing process and then effectively filters the noise points through an edge-preserving filter and median filter. Finally, the effective cloth edge information and deviation information are extracted, and the effective sewing track curve is obtained. The maximum error between the sewing edge trajectory and the actual edge calculated by the experiment is 1.35 mm, the mean error is 0.61 mm, and the mean value of dynamic variance is 0.57 mm. The detection accuracy can meet the actual needs, provide a theoretical data model for the subsequent sewing trajectory, and lay a foundation for the realization of automatic sewing.

**Declaration of conflicting interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Thanks to the fund of the major general technology research project of Guangdong Jihua Laboratory (Y80311W180) and Shanghai artificial intelligence major project (2021SHZDZX0103).
References

1. Kittler J. On the accuracy of the Sobel edge detector. *Image Vis Comput* 1983; 1(1): 37–42.
2. Torre V and Poggio TA. On edge detection. *IEEE Trans Pattern Anal Mach Intell* 1986; 8(2): 147–163.
3. Wang BT and Yang SL. Edge detection of gray image based on Gauss Laplace operator. *Comput Eng Appl* 2003; 26: 132–134.
4. Canny J. A computational approach to edge detection. *IEEE Trans Pattern Anal Mach Intell* 1986; 8(6): 679–698.
5. Shen W, Wang X, Wang Y, et al. DeepContour: a deep convolutional feature learned by positive-sharing loss for contour detection. In: *Proceedings of the 2015 IEEE conference on computer vision and pattern recognition*, Boston, MA, 7–12 June 2015, pp.3982–3991. Piscataway, NJ: IEEE.
6. Liu Y and Lew MS. Learning relaxed deep supervision for better edge detection. In: *Proceedings of the 2016 IEEE conference on computer vision and pattern recognition*, Las Vegas, NV, 27–30 June 2016, pp.231–240. Piscataway, NJ: IEEE.
7. Liu Y, Cheng M, Hu X, et al. Richer convolutional features for edge detection. In: *Proceedings of the 2017 IEEE conference on computer vision and pattern recognition*, Honolulu, HI, 21–26 July 2017, pp.5872–5881. Piscataway, NJ: IEEE.
8. Xu D, Ouyang W, Alameda PX, et al. Learning deep structured multi-scale features using attention-gated CRFs for contour prediction. In: *Proceedings of the 31st international conference on neural information processing systems*, 2017, pp.3961–3970. Red Hook, NY: Curran Associates Inc.
9. Deng R, Shen C, Liu S, et al. Learning to predict crisp boundaries. In: *Proceedings of the 2018 European conference on computer vision*, LNCS 11210. Cham: Springer, 2018, pp.570–586.
10. He J, Zhang S, Yang M, et al. Bi-directional cascade network for perceptual edge detection. In: *Proceedings of the 2019 IEEE/CVF conference on computer vision and pattern recognition*, Long Beach, CA, 15–20 June 2019, pp.3823–3832. Piscataway, NJ: IEEE.
11. Orujov F, Maskeliūnas R, Damaševičius R, et al. Fuzzy based image edge detection algorithm for blood vessel detection in retinal images. *Appl Soft Comput* 2020; 94: 1568–1594.
12. Li K, Tian Y, Wang B, et al. Bi-directional pyramid network for edge detection. *Electronics* 2021; 10(3): 329.
13. Xie S and Tu Z. Holistically-nested edge detection. In: *Proceedings of the 2015 IEEE international conference on computer vision*, Santiago, Chile, 7–13 December 2015, pp.1395–1403. Piscataway, NJ: IEEE.
14. Liu Y, Cheng M, Hu X, et al. Richer convolutional features for edge detection. In: *Proceedings of the 2017 IEEE conference on computer vision and pattern recognition*, Honolulu, HI, 21–26 July 2017, pp.5872–5881. Piscataway, NJ: IEEE.
15. Everingham M, Eslami SMA, Van Gool L, et al. The PASCAL visual object classes challenge: a retrospective. *Int J Comput Vis* 2015; 111: 98–136.
16. Paszke A, Chaurasia A, Kim S, et al. ENet: a deep neural network architecture for real-time semantic segmentation. *Comput Vis Pattern Recognit* 2016; 2016: 1–10.
17. Badrinarayanan V, Kendall A and Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans Pattern Anal Mach Intell* 2017; 39(12): 2481–2495.
18. Duan RL, Li QX and Li YH. Overview of image edge detection methods. *Opt Technol* 2005; 31(3): 5.
19. Shu Rong P, Yaonan W, Guoci L, et al. A strategy improving registration accuracy progressively for INSAR complex image. *IEEE Int Conf Automation Sci Eng* 2006; 2006: 626–629.
20. Xu D, Dong WF, Pan ZK, et al. Edge detection algorithm based on mathematical morphology and wavelet transform. *Comput Appl* 2012; 32(A02): 165–167.
21. Perez L and Wang J. The effectiveness of data augmentation in image classification using deep learning. *IEEE Trans Image Process* 2018; 27(3): 547–554.
22. Zhong Z, Zheng L, Kang G, et al. Random erasing data augmentation. *Proc Conf AAAI Artif Intell* 2020; 34(7): 13001–13008.
23. Simonyan K and Zisserman A. Very deep convolutional networks for large-scale image recognition. *Int Conf Learn Representations* 2015; 5(3): 345–358.
24. Kingma D and Ba J. Adam: a method for stochastic optimization. *Comput Sci* 2014; 1: 1–15.