The Centerline Extraction Algorithm of Weld Line Structured Light Stripe Based on Pyramid Scene Parsing Network

WEIBO YU, YU LI, HONGTAO YANG, AND BAIZHU QIAN
School of Electrical and Electronic Engineering, Changchun University of Technology, Changchun 130012, China
Corresponding author: Hongtao Yang (hongtao_3110594@126.com)

This work was supported by the Jilin Province Science and Technology Development Plan Project “Research and Development of Active Vision Guidance System for Automatic Welding Robots” under Grant 20200401118GX.

ABSTRACT Based on the good feature learning ability of the pyramid scene parsing network, a method for extracting the centerline of structured light stripes of weld lines based on the pyramid scene parsing network and Steger algorithm is proposed. This method avoids the traditional complex weld image preprocessing technology, and simplifies the operation steps of extracting the centerline of the structured light stripe of the weld image line. In this paper, the pyramid scene parsing network is used to predict the pixels containing weld feature information. Through the pyramid pooling module, the local and global context feature information is fused to supplement the feature information of the weld edge, and then the Steger algorithm is used to extract the weld feature centerline. The results show that the method in this paper can accurately extract the centerline position of the structured light stripe of the weld line under the interference of reflection, and the average value reaches 86.8 % on the accuracy evaluation index mean intersection over union, the 18.93 pixels on the weld extraction accuracy index root mean square error, and the average time of extracting the center line of structured light stripe of weld line is 0.188 s.

INDEX TERMS Welding seam feature extraction, pyramid scene parsing network, Steger algorithm, image segmentation.

I. INTRODUCTION
In recent years, with the rapid development of automation process in many industries such as industrial manufacturing, and the rapid improvement of related technologies, relying on robots for metal welding operation has gradually become an indispensable link in the production process. In order to meet the requirements of welding process in modern automatic production, it is necessary to obtain real-time information of weld centerline and weld feature points in the welding process. Welding seam tracking technology based on active vision has become one of the research hotspots in the field of automatic welding [1]. In the production process of welding in the factory, it is always affected by many interference conditions such as splash, dust and strong arc light. Therefore, in the process of automatic welding, it is particularly important to use line structured light technology to accurately obtain weld position information [2].

At present, there are many methods for feature extraction of line structured light stripe center line in weld image. Weiming et al. [4] proposed a fast line laser stripe center extraction algorithm, which has certain anti-noise performance. However, when the background is complex, the light quality of line structured light is not good, and the stripe is non-uniform and robust, the algorithm needs to be combined with more algorithms to operate. Keqi et al. [5] proposed a weld detection method based on image preprocessing and post-processing, but the detection method is extremely complex and has poor anti-noise performance. Jie and Yawen [6] proposed a variable threshold segmentation algorithm based on Otsu threshold (Otus). This method effectively removes the influence of background noise, but it takes too long and affects efficiency. However, these methods for processing weld images need to cooperate with complex image preprocessing techniques due to the smoke and dust conditions and...
light conditions during welding, which requires more practical experience of practitioners to select suitable operation methods for different scenes and complex conditions. And the practical process of operation steps are too cumbersome, not conducive to universal application. Yongshuai et al. [7] proposed a method of weld feature extraction based on full convolution neural network, but the accuracy is not high. It can only extract the edge contour of the weld, and cannot extract the center line of the structured light stripe of the weld line. Therefore, the feature extraction algorithm of weld image needs to be developed to the intelligent learning algorithm with wider learning ability [8]. The feature extraction algorithm of weld image based on deep learning began to appear in relevant research fields to improve the adaptability and anti-interference ability of weld feature extraction [9].

Traditional neural network image segmentation algorithm in image segmentation, always for some not obvious small size object image segmentation accuracy is insufficient, or even misclassification. PSPNet [10] for scene parsing and semantic segmentation tasks, by introducing the Pyramid Pooling Module, the local and global information are fused together, and the optimization strategy of moderate supervision loss is used. The segmentation accuracy on multiple datasets exceeds FCN, DeepLab-v2 [11] and other neural network models. In this paper, a weld laser line feature extraction method based on PSPnet and Steger [12] algorithm is proposed. This method can accurately obtain the position information of the weld center line in the weld image, which can ensure the extraction accuracy of the weld center line and greatly simplify the operation steps.

II. ALGORITHM PRINCIPLE
A. PSPNet ALGORITHM PRINCIPLE
PSPNet is a network proposed for relatively complex scene analysis problems. Among them, its network structure is based on the residual network (ResNet) [13] as the main skeleton, and some unique designs are added after the end of ResNet. This network is mainly aimed at the traditional convolutional neural network that does not make good use of the context information in the scene, and cannot highlight and predict some unobvious category features (details that are easily overlooked), so the pyramid structure is introduced to solve this problem. In deep learning, the size of the receptive field indirectly determines the extent to which image context information is used. ResNet effectively expands the receptive field through-hole convolution and cross-layer connection, but as the depth of the network increases, the actual receptive field is still more theoretical. Feel the wild to be small. The pyramid pooling module in PSPNet uses pooling of different scales to effectively alleviate the situation that the actual receptive field in the network is less than the theoretical receptive field. Among them, the image data input size:

\[ W_1 \times H_1 \times D_1 \]  

\[ W_1 \] is image width, \( H_1 \) is image height, \( D_1 \) is image depth (number of channels). The output image size after convolution is:

\[ W_n = (W_{n-1} - F + 2P) / S + 1 \]
\[ H_n = (H_{n-1} - F + 2P) / S + 1 \]
\[ D_n = K \]  

\( W \) is the width of the input picture; \( H \) is the length of the input picture; \( D \) is the image depth (number of channels); \( K \) is the specific value; \( F \) is the dimension of the convolution kernel; \( S \) is the step size; \( P \) is the number of pixels to be filled. \( n \) represents the order. Using the same function to ensure that the output is positive and prevent overfitting, the mathematical expression of the \( \text{relu} \) function is:

\[ f(x) = \max(0, x) \]  

Convolution and activation function operation:

\[ x_j^l = f\left( \sum_{l' < M_j} x_{l'}^{l-1} \ast k_j^{l'} + b_j^l \right) \]  

The part in brackets is the main operation for extracting features of the convolutional layer: \( l \) is the current convolutional layer of the network, \( j \) represents one of the feature maps of the current convolutional layer of the network. \( x_j^l \) is the feature map, \( x_{l-1}^{l-1} \) is the value of the \( i \)-th pixel in the receptive field, \( M_j \) represents the set of selected input maps, and \( k_j^{l'} \) is the connection between the \( i \)-th pixel in the \( x_{l-1}^{l-1} \) layer and the convolution kernel of the \( j \)-th map in the \( l \) layer, \( b_j^l \) is the bias term, and \( f() \) is the activation function. The \( q \)-th layer is the sub-sampling layer (pooling layer), and the \( h \)-th feature map is:

\[ x_j^l = f\left( w_j^l \downarrow (x_j^{l-1}) \right) + b_j^l \]  

\( \downarrow() \) is the down-sampling function (downsampling using the maximum method), \( w_j^l \) is the weight, and \( b_j^l \) is the offset. When the network propagates forward, the weight parameters of the network are updated by minimizing the training error divided by the loss. Losses during training are calculated as follows:

\[ \text{Loss} = \frac{1}{n} \sum_{i} \ln(Q(x = m)), \quad i = 0, 1, \ldots, n - 1 \]  

\[ Q(x = m) = \frac{\exp(z_m)}{\sum_{j} \exp(z_j)}, \quad j = 0, 1, \ldots, M - 1 \]  

In the formula, \( \text{Loss} \) is the network training \( \text{Loss}, Q(x = m) \) is the probability that the pixel \( x \) belongs to the category \( m \), \( n \) is the number of pixels in the current training image, \( z_j \) is the eigenvalue of the \( j \)-th category, \( M \) is the total number of categories, there are two categories in this paper, so \( M = 2 \). In the phase of backpropagation updating network weight parameters [14], this paper adopts the stochastic gradient descent method (SGD) to update the weight by the linear
combination of the negative gradient $\nabla L(W_t)$ and the updated value of the previous weight. The calculation is as follows:

$$V_{t+1} = \mu V_t - \alpha \nabla L(W_t)$$

(8)

$$W_{t+1} = W_t + V_{t+1}$$

(9)

In the formula, $W_t$ is the weight matrix of the tth iteration, $V_t$ is the weight update value of the tth iteration, $\alpha$ is the basic learning rate of the negative gradient, $\mu$ is the weight of the weight update value $V_t$, used to weight the influence of the previous gradient direction on the current gradient descent direction.

**B. PRINCIPLE OF STEGER ALGORITHM**

Steger algorithm is a method for extracting the centerline of structured light stripes of weld line image based on Hessian matrix [15]. The direction where the absolute value of the second-order directional derivative in the image $I(x, y)$ takes the maximum represents the normal direction of the linear structured light stripe and this direction can be determined by calculating the eigenvalues and eigenvectors of the Hessian matrix. In the online structured light stripe image (weld image), the Hessian matrix is expressed as follows:

$$H(x, y) = \begin{pmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{pmatrix}$$

(10)

In the formula:

$$I_{xx} = \frac{\partial^2 G(x,y)}{\partial^2 x} \otimes I(x, y)$$

(11)

$$I_{yy} = \frac{\partial^2 G(x,y)}{\partial^2 y} \otimes I(x, y)$$

(12)

$$I_{xy} = \frac{\partial^2 G(x,y)}{\partial x \partial y} \otimes I(x, y)$$

(13)

The formula $G(x, y)$ is a two-dimensional Gaussian convolution template, which highlights the gray distribution characteristics of the stripe. $I(x, y)$ is an image matrix centered on an image point $(x, y)$ and equal in size to a two-dimensional Gaussian convolution template.

The normal direction of the image point $(x_0, y_0)$ is the direction where the absolute value of the second-order derivative of the point takes the maximum value [16], which is given by the eigenvector $(n_x, n_y)$ corresponding to the absolute value of the maximum eigenvalue of the Hessian matrix of the image point $(x_0, y_0)$ . Its tangent direction is perpendicular to the normal direction. The gray distribution function in the normal direction of the point $(x_0, y_0)$ is expanded by the second-order Taylor expansion, which is represented by $I(x_0 + m_x, y_0 + m_y)$. Since the gray value of the light stripe is Gaussian distribution in the normal direction of the light stripe, and the greater the light intensity is, the closer it is to the center of the light stripe. Thus the point where the first derivative of $I(x_0 + m_x, y_0 + m_y)$ is zero that the sub-pixel stripe center point $(x_0 + m_x, y_0 + m_y)$ in the normal direction of point $(x_0, y_0)$. The value of $t$ is shown below:

$$t = \frac{n_x I_x + n_y I_y}{n_x^2 I_{xx} + 2n_x n_y I_{xy} + n_y^2 I_{yy}}$$

(14)

In the formula:

$$I_x = \frac{\partial G(x, y)}{\partial x} \otimes I(x, y)$$

(15)

$$I_y = \frac{\partial G(x, y)}{\partial y} \otimes I(x, y)$$

(16)

**III. THE CENTERLINE EXTRACTION ALGORITHM OF WELD LINE STRUCTURED LIGHT STRIPE BASED ON PSPNet**

**A. OVERALL DESIGN OF ALGORITHM**

In this paper, the extraction method of weld line structured light stripe centerline is mainly based on PSPNet for image segmentation of line structured light stripe centerline in weld image and then combined with Steger algorithm for line structured light stripe centerline extraction. The algorithm consists of two parts, the specific process is shown in Figure 1. Among them, in order to reduce the time required for training the model and accelerate the convergence rate, the pre-trained ResNet-50 is used for parameter fine-tuning. Among them, PSPNet extracts the features of weld image from ResNet-50 network, and extracts the deep and shallow features of the image through the pyramid pooling module composed of pooling layers of different sizes of pooling cores, and fuses them at multiple levels to reduce the probability of false segmentation. PSPNet network for image feature extraction classification stage and the data after processing up-sampling stage, and finally through a convolution layer and up-sampling layer, to achieve end-to-end output. The data set is input into the network and trained by fine-tuning the network parameters. When the network Loss converges to a certain extent, the network training is stopped. The trained neural network model is used to predict the corresponding image segmentation, and then the Steger algorithm is used to extract the centerline of the structured light stripe. Finally, the segmentation effect and the completed parameters are checked.

**B. PSPNet MODEL CONSTRUCTION**

PSPNet network extracts features from the input image through a convolution neural network (CNN) to obtain the feature map. The size of the feature map is 1/8 of the input, and the obtained feature map is combined with the pyramid adaptive average pooling module to capture the features of different sub-regions, so as to obtain the overall context information. The small graph of sub-region features is upsampled to the size of the feature graph and combined with the
previous global features. so that the current features include
global features and local features. Finally, the final prediction
results are obtained by convolution and up-sampling. The
final output result is obtained after the last convolution layer.
PSPNet network structure is shown in Figure 2.

The black dotted line in Figure 2 is the core part of the
PSPNet network, which is the pyramid pooling module. The
resulting feature maps are pooled adaptively on four scales
of \(1 \times 1\) (red), \(2 \times 2\) (yellow), \(3 \times 3\) (blue), and \(6 \times 6\) (green),
which helps the network integrate global context information.
Taking the uppermost layer as an example, the uppermost
layer is the roughest \(1 \times 1\) global pooling to generate a single-
pixel multi-channel output. If \(N\) levels are set in the pyramid,
\(1 \times 1\) convolution is used to reduce the number of channels
to the original \(1/N\) after pooling. After convolution, each
level is up-sampled by bilinear interpolation, and fused with
the original feature map as a priori information. The feature
maps of different scales are up-sampled and fused with the
original feature map before the pyramid pooling structure is
applied, and further convolution is performed in turn. Finally,
a complete output image is obtained.

C. EXTRACTION OF STRIPE CENTER LINE OF LINE
STRUCTURED LIGHT IN WELD IMAGE

After obtaining the processed weld image by the above
algorithm, the Steger algorithm with high accuracy is used
to extract the center line of the weld line structured light
stripe. In order to reduce the processing time, according to
the light stripe area of the weld image, the corresponding
image coordinates are set by OPENCV (computer vision and
machine learning open source software library) Rect class
operation, and the line structured light stripe line in the image
is divided into a separate region of interest (ROI). In order to
ensure the real-time tracking of weld seam, reduce the
time of image processing as far as possible, and reduce the
interference of unnecessary weld image area on the extraction
accuracy of line structured light image stripe center line. For
the collected weld image, the specific steps of extracting the
stripe center line of the line structured light image are as
follows:

Step1: Get the original weld image and input it into the
PSPNet model that has been trained by the data set to segment
the image.

Step2: According to the preset image size, the ROI image
region is obtained with the weld image as the center.

Step3: Steger algorithm is used to extract the centerline of
the weld line structured light stripe. And output the image
extraction results.

Through the above steps to optimize the original image of
the weld, you can see the final results as shown in Figure 3,
the algorithm effectively removes the interference such as
reflection, and retain the integrity of the centerline of the line
structured light weld image stripe.

IV. EXPERIMENTAL PROCESS AND RESULTS ANALYSIS

A. DATA SET PREPARATION

In this paper, the PSPNet-50 (one of PSPnet) with high accu-
racy is selected. They are all neural networks for supervised
learning, so it is necessary to prepare the relevant sample
label data set for network training, generate the corresponding
network model, and realize the segmentation of weld images.
The original image data set used in the experiment is 800 origi-
nal image training data of welding seam taken by high resolu-
tion industrial digital camera, 200 validation set and 100 test
data. In order to reduce the hardware resource cost and time
cost of network model training, the size of the collected
images is adjusted to \([320 \times 320]\) size. The Label-Me image
annotation tool is used to annotate weld line structured light
images, where _background_ represents the background and
the corresponding visual RGB value is \((0,0,0)\).1001 repre-
sents the weld, and the corresponding visual RGB value is
\((128,0,0)\). The details are shown in Figure 4.

In order to ensure the final segmentation effect of the
image, improve the generalization performance of the trained
network model, and prevent the underfitting problem caused
by too little sample data. The original image data set and
label image data set are expanded, and the original image and
label image are mirrored and rotated. Finally, 3200 images
are obtained as the training set, 800 images as the validation
set and 400 images as the test set, as shown in Figure 5.

B. PSPNet NETWORK MODEL PARAMETERS AND
TRAINING

The hardware experiment environment in this article is:
CPU is i7–9750 processor; RAM is 16G; GPU graphics
W. Yu et al.: Centerline Extraction Algorithm of Weld Line Structured Light Stripe

The Keras framework is used to implement the model. The software experiment environment adopts win−10 system; keras−2.2.4; Tenorflow−1.9; CUDA−9.0; Python−3.6; opencv−4.2. Among them, Python is Anaconda3 (64bit) installed. Using Adam as the optimizer of the model, the momentum parameter is set to 0.9, and the learning rate is 0.0001. The batch size is set to 1, and the epochs is set to 300. Use BatchNorm for normalization. ReLU is used as the activation function. Among them, Conv and Conv2d are both convolutional layers. In order to prevent over−fitting, this article uses Dropout after conv2d_5 with a ratio of 0.1. MaxPooling2D represents the maximum pooling operation. Res_branch represents the residual module, lambda represents the linear interpolation module, and concatenate_1 represents the fusion operation between the layers. average_pooling2d_1 represents average pooling. The specific parameter information of each inter−layer model is shown in Table 1.

C. EXPERIMENTAL RESULT ANALYSIS

In order to verify whether the algorithm in this paper has some progress in image segmentation accuracy, this paper selects the Unet network and SegNet network to compare the segmentation results and uses the same data and Resnet−50 model weight parameters for training. Use test data sets to validate the trained model. Among them, to reduce the chance of algorithm optimization, in this paper, in the 3200 images of the data set (800 images for each weld image), four original images are randomly selected from four weld images, and three algorithms are used for image segmentation prediction. The comparison results of image segmentation accuracy of PSPnet algorithm, UNet algorithm and SegNet algorithm are shown in Figure 6. To better compare the performance of this algorithm, several important evaluation indexes commonly used in the field of neural network image segmentation algorithm are used to measure the accuracy of image segmentation: pixel accuracy (PA), mean pixel accuracy (MPA) and mean intersection over union (MIoU). The comparison results of evaluation indicators are shown in Tables 2, 3, and 4.

As can be seen from the experimental results, the PSPNet algorithm used in this paper, four weld image segmentation mean, of which PA is 99.8 %. MPA was 96.7 %. The MIoU value is 86.8 %. From the comparison of Tables 2, 3, and 4, it can be seen that PSPNet algorithm is 0.9 % ahead of UNet algorithm in MPA and 1 % ahead of UNet algorithm in MIoU. On PA they are the same. The gap between PSPNet algorithm and SegNet algorithm is even larger. PSPNet algorithm leads SegNet algorithm by 8.8 % in MPA and 3.1% in MIoU. On PA, the difference is only 0.2 %. The main reason for this phenomenon is related to the data set used in this paper. The background in the weld segmentation image accounts for the main body of the whole image, so the PA value is high. And for lap weld, butt weld segmentation, the performance of three algorithms are good, so in the pixel accuracy PA, the three algorithms are not much different. Based on the above data, the PSPNet algorithm adopted in this paper has certain advantages in accuracy compared with the traditional UNet algorithm and the SegNet algorithm.

In order to verify whether the centerline extraction algorithm in this paper has made some progress in the accuracy of the centerline extraction of the line structured light stripes in the weld image, and whether it has good stability and extraction speed. In this paper, the segmentation results of the above algorithms are used as experimental objects, and the traditional gray weighted centroid algorithm (GWCA) and extreme value algorithm (EVA) are selected to compare with the proposed algorithm. The weld line structured light stripe center lines extracted by each algorithm are shown in Figure 7, 8, 9. The centerline of the structured light stripe of the weld line extracted by the algorithm in this paper is...
| Model | Layer | Kernel_Size | Filters | Output Shape |
|-------|-------|-------------|---------|--------------|
| Input | – | – | (320,320,3) |
| Conv1 | 3x3 | 64 | (160,160,64) |
| Conv2 | 3x3 | 64 | (160,160,64) |
| Conv3 | 3x3 | 64 | (160,160,128) |
| MaxPooling2D | 3x3 | 128 | (80,80,128) |
| Res2_branch | 1x1, 3x3, 64 | 64 | (80,80,256) |
| Res3_branch | 3x3, 2x2, 128 | 128 | (40,40,512) |
| Res4_branch | 3x3, 2x2, 256 | 512 | (40,40,1024) |
| Res5_branch | 3x3, 2x2, 512 | 512 | (40,40,2048) |
| average_pooling2d_1 | 30x30 | 2048 | – |
| average_pooling2d_2 | 15x15 | 2048 | – |
| average_pooling2d_3 | 10x10 | 2048 | – |
| average_pooling2d_4 | 5x5 | 2048 | – |
| conv2d_1 | 1x1 | 512 | – |
| conv2d_2 | 1x1 | 512 | – |
| conv2d_3 | 1x1 | 512 | – |
| conv2d_4 | 1x1 | 512 | – |
| lambda_1 | – | – | (40,40,512) |
| lambda_2 | – | – | (40,40,512) |
| lambda_3 | – | – | (40,40,512) |
| lambda_4 | – | – | (40,40,512) |
| concatenate_1 | – | – | (40,40,4096) |
| conv2d_5 | 3x3 | 512 | (40,40,512) |
| dropout_1 | – | – | (40,40,512) |
| conv2d_6 | 1x1 | 2 | (40,40,2) |
| lambda_5 | – | – | (320,320,2) |
| Output | – | – | (320,320,2) |

The centerline of the stripe extracted by the other algorithms, there are obviously different degrees of position offsets. In particular, the centerline of the light stripe extracted by the Extreme value algorithm in several groups of weld images can obviously see that the centerline offset of the light stripe is very serious. In order to accurately compare the accuracy of each algorithm, the root mean square error (RMSE) is used to quantitatively analyze the effect of extraction. The results are shown in tables 5, 6, and 7. It can be seen from Table 5, 6, and 7 that the overall RMSE of the PSPnet-Steger algorithm extraction results is 18.93 pixel according to the extraction results of the center line of the structured light stripe of the weld line of the T-type weld, lap weld and butt weld. This result is smaller than that of Segnet-Steger method and UNet-Steger algorithm, and is much smaller.
than that of PSPNet-extreme value algorithm and PSPNet-gray weighted centroid algorithm. The extraction of the center line of the structured light stripe of the weld line of the V-type weld is slightly inferior to that of the UNet-Steger algorithm. And from the average results of each algorithm, the overall RMSE of Steger algorithm is less than gray center of gravity algorithm and extreme value algorithm. It shows that this algorithm still has high accuracy.

In order to accurately compare the time loss between algorithms, the time comparison results of PSPNet—Steger algorithm and each algorithm are shown in tables 8, 9 and 10. In terms of the time comparison results of each algorithm, the PSPNet-Steger algorithm takes slightly more time than the PSPNet—gray gravity center algorithm and the PSPNet—extremum value algorithm. Among them, the PSPNet—extremum value algorithm with the fastest speed takes 0.062s. PSPNet—Steger algorithm takes 0.188s. However, the three algorithms have little difference in time, and all meet the real—time requirements in the welding process. In summary, from the overall results, the performance of

### TABLE 2. UNet algorithm evaluation index results.

| Standard      | PA (%) | MPA(%) | MIOU(%) |
|---------------|--------|--------|---------|
| T-type        | 99.7   | 96.4   | 84.5    |
| V-type        | 99.9   | 94.2   | 88.7    |
| Lap type weld | 99.7   | 95.7   | 85.3    |
| Butt type weld| 99.8   | 97.0   | 84.6    |
| Percentage    | 99.8   | 95.8   | 85.8    |

### TABLE 3. SegNet algorithm evaluation index results.

| Standard      | PA (%) | MPA(%) | MIOU(%) |
|---------------|--------|--------|---------|
| T-type        | 99.8   | 85.9   | 84.2    |
| V-type        | 99.6   | 84.8   | 81.0    |
| Lap type weld | 99.4   | 86.2   | 82.2    |
| Butt type weld| 99.5   | 94.8   | 87.3    |
| Percentage    | 99.6   | 87.9   | 83.7    |

### TABLE 4. PSPNet algorithm evaluation index results.

| Standard      | PA (%) | MPA(%) | MIOU(%) |
|---------------|--------|--------|---------|
| T-type        | 99.8   | 98.9   | 86.1    |
| V-type        | 99.8   | 93.5   | 87.7    |
| Lap type weld | 99.9   | 96.1   | 87.9    |
| Butt type weld| 99.8   | 98.3   | 85.3    |
| Percentage    | 99.8   | 96.7   | 86.8    |

### TABLE 5. PSPNet-Steger algorithm and evaluation index (RMSE) comparison results of each algorithm.

| Standard      | PSPNet-Steger | SegNet-Steger | Unet-Steger |
|---------------|---------------|---------------|-------------|
| T-type        | 22.86         | 25.19         | 24.33       |
| V-type        | 20.52         | 21.45         | 18.77       |
| Lap type weld | 15.71         | 16.07         | 18.94       |
| Butt type weld| 16.61         | 20.41         | 23.80       |
| Percentage    | 18.93         | 20.78         | 21.46       |
TABLE 6. PSPnet-Gray weighted centroid algorithm and evaluation index (RMSE) comparison results of each algorithm.

| Standard     | PSPnet-GWCA | SegNet-GWCA | Unet-GWCA |
|--------------|-------------|-------------|-----------|
| T-type       | 58.56       | 55.03       | 67.50     |
| V-type       | 52.51       | 55.76       | 58.29     |
| Lap type weld image | 65.20       | 57.85      | 54.89     |
| Butt type weld image | 50.89       | 49.40      | 61.33     |
| Percentage   | 56.79       | 54.51       | 60.50     |

TABLE 7. PSPnet-Extreme value algorithm and evaluation index (RMSE) comparison results of each algorithm.

| Standard     | PSPnet-EVA | SegNet-EVA | Unet-EVA |
|--------------|------------|------------|----------|
| T-type       | 49.89      | 51.98      | 64.20    |
| V-type       | 52.68      | 55.19      | 55.25    |
| Lap type weld image | 66.27       | 59.85      | 55.76    |
| Butt type weld image | 54.61       | 52.23      | 61.62    |
| Percentage   | 55.86      | 54.81      | 59.21    |

TABLE 8. The time-use comparison results of PSPnet-Steger algorithm and each algorithm.

| Standard     | TIME (s) |
|--------------|----------|
| PSPnet-Steger | 0.216   |
| SegNet-Steger | 0.236   |
| Unet-Steger  | 0.235    |

TABLE 9. The time-use comparison results of PSPnet-Gray weighted centroid algorithm and each algorithm.

| Standard     | TIME (s) |
|--------------|----------|
| PSPnet-GWCA  | 0.127    |
| SegNet-GWCA  | 0.147    |
| Unet-GWCA    | 0.157    |

TABLE 10. The time-use comparison results of PSPnet-Extreme value algorithm and each algorithm.

| Standard     | TIME (s) |
|--------------|----------|
| PSPnet-EVA   | 0.078    |
| SegNet-EVA   | 0.088    |
| Unet-EVA     | 0.114    |

V. CONCLUSION

In this paper, a central line extraction algorithm of a weld line structured light stripe based on PSPNet is proposed. At present, the linear structured light welding robot is susceptible to environmental factors, which leads to welding deviation, and the image preprocessing is too complex. Combined with deep learning technology, PSPNet is applied to the extraction process of weld line structured light stripe centerline. Compared with the SegNet algorithm and the Unet algorithm, PSPNet has good anti-interference ability. The extraction accuracy of the centerline of the structured light stripe of the weld line is high, and the image preprocessing process is simplified. The manual complex operation is omitted, which provides a new idea for the weld image processing method.

REFERENCES

[1] Z. Yang and D. Yi, “Welding wires centerline detection method based on image gradient segmentation,” in Proc. IEEE 3rd Int. Conf. Inf. Commun. Signal Process. (ICICSP), Shanghai, China, Sep. 2020, pp. 290–294.
[2] H. Yu, C. Peng, Z. Zhao, L. Bai, and J. Han, “Visual texture-based 3-D roughness measurement for additive manufacturing surfaces,” IEEE Access, vol. 7, pp. 186646–186656, 2019.
[3] Z. Yuan, M. Xiangqun, J. Dengbiao, and T. Houjun, “Structural light stripe center extraction method under complex interference,” Chin. Laser, vol. 47, no. 12, pp. 172–180, 2020.
[4] L. Weiming, X. P. Guo, and D. Chang, “Fast extraction algorithm of line laser stripe center,” China Laser, vol. 47, no. 3, pp. 192–199, 2020.
[5] L. Keqi, L. Yizhong, X. Nan, and G. Zhengliang, “Research on weld recognition based on industrial robot arc welding,” Mod. Manuf. Technol. Equip., no. 10, pp. 13–17, 2019.
[6] H. Jie and L. Yawen, “Variable threshold segmentation algorithm of line structured light stripe center extraction,” Surv. Geograph. Inf., vol. 45, no. 3, pp. 20–23, 2020.
[7] Z. Yongshuai, Y. Guowei, W. QiQi, M. Lei, and W. Yizhong, “Welding seam feature extraction based on fully convolutional neural network,” China Laser, vol. 46, no. 3, pp. 36–43, 2019.
[8] E. Shelhamer, J. Long, and T. Darrell, “Fully convolutional networks for semantic segmentation,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 4, pp. 640–651, Apr. 2017.
[9] Y. Tian, H. Liu, L. Li, G. Yuan, J. Feng, Y. Chen, and W. Wang, “Automatic identification of multi-type weld seam based on vision sensor with silhouette-mapping,” IEEE Sensors J., vol. 21, no. 4, pp. 5402–5412, Feb. 2021.
[10] H. Zhao, J. Shi, and X. Qi. (2016). Pyramid scene parsing network. IEEE Computer Society. [Online]. Available: http://www.bookref.com
[11] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, “DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 4, pp. 834–848, Apr. 2018, doi: 10.1109/TPAMI.2017.2699184.
[12] C. Steger, “An unbiased detector of curvilinear structures,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 2, pp. 113–125, Feb. 1998.

[13] S. Kiliç, I. Askерзаде, and Y. Kaya, “Using ResNet transfer deep learning methods in person identification according to physical actions,” *IEEE Access*, vol. 8, pp. 220364–220373, 2020.

[14] E. K. Kumar, P. V. V. Kishore, M. T. K. Kumar, D. A. Kumar, and A. S. C. S. Sastry, “Three-dimensional sign language recognition with angular velocity maps and convolved feature resnet,” *IEEE Signal Process. Lett.*, vol. 25, no. 12, pp. 1860–1864, Dec. 2018.

[15] T. Sun and S. Yang, “An approach to formulate the hessian matrix for dynamic control of parallel robots,” *IEEE/ASME Trans. Mechatronics*, vol. 24, no. 1, pp. 271–281, Feb. 2019.

[16] O. van der Heide, A. Sbrizzi, and C. A. T. van den Berg, “Accelerated MR-STAT reconstructions using sparse hessian approximations,” *IEEE Trans. Med. Imag.*, vol. 39, no. 11, pp. 3737–3748, Nov. 2020.

WEIBO YU received the M.Sc. degree in electric drive and automation from the Jilin Institute of Technology. She is currently a Professor and a Master’s Tutor with the School of Electrical and Electronic Engineering, Changchun University of Technology. She has long-term engagement in intelligent instruments and intelligent control, digital image processing, and machine vision research. As the person in charge and main participants, she has completed several national natural science funds, national science and technology support plan project, and Jilin province science and technology development plan project. She has published more than 30 academic papers in domestic and foreign journals and conferences, of which 15 were retrieved by EI. As a major participant, she received second prize and third prize for provincial and ministerial scientific and technological progress.

YU LI received the B.E. degree from Jilin Engineering Normal University, in 2019. He is currently pursuing the M.A.Eng. degree with the Changchun University of Technology. His research interests include computer vision, digital image processing, and machine vision.

HONGTAO YANG received the Ph.D. degree in mechanical and electronic engineering from the University of Chinese Academy of Sciences, Jilin, China. His current research interests include the application of visual servo systems to robot kinematics to improve safety, stability, and efficiency of robots in conjunction with human operator. He is a member of the Society of Electrical Technology, Jilin.

BAIZHU QIAN graduated from Jilin Engineering Normal University. She is currently pursuing the degree with the Changchun University of Technology. Her current main research interests include machine vision and point cloud processing.