Line segment detection based on improved U-Net

Yunpeng Zhi[a], Baolong Guo[b, *], Zhijie Chen[c] and Cheng Li[d]

School of Aerospace Science and Technology, Xidian University; Xi' an 710071, China;

[a]zhiyunpeng_ok@163.com, [b]blguo@xidian.edu.cn, [c]ck102722@163.com

[d]licheng812@stu.xidian.edu.cn

*Corresponding author

Keywords: LSD, U-Net network, Low-Level feature

Abstract: Line segment detection measurement technology is an important and challenging task in computer vision, and it is also a hot research direction at home and abroad. This paper uses the improved U-Net network to extract line segments from simple textured images. The simple textured images no longer require high-level features, and pay more attention to the extraction of low-level features. The improved network enhances the low-level information in the image. Feature extraction has significantly improved the performance of straight line detection on images with simple textures.

1. Introduction

The concept of deep learning comes from artificial neural networks, and neural network models generally include perceptrons with different parameters.Traditional machine learning uses low-level features to express high-level features, and then the network learns the distribution performance of the features. Convolutional neural networks omit this complicated processing and directly learn the original data distribution. In the past few years, convolutional neural networks have become more and more widely used in image processing. Line segment detection is an important and important and challenging low-level tasks in computer vision. The generated line segment maps provides compact structural information and helps complete many advanced visual tasks such as 3D reconstruction, image segmentation, stereo matching, scene analysis, and image stitching[3]. In 2018,Nan xue, Song Bai, and Gui-song Xia, etc. proposed an improved LSD algorithm, which used dual representation dual representation of attraction field mapping based on line segment mapping and region division, and pose the problem of extracting line segment as a region coloring problem, which was performed by a convolutional neural network in deep learning as Image Semantic Segmentation[3].In this paper, the network architecture in this method is improved, which makes the extraction of Low-Level information features more prominent in the network structure. Through this algorithm, the line segments detection rate of simple texture images is improved.

2. U-Net network and LSD algorithm

2.1 U-Net network analysis

The U-Net network was proposed in 2015 and is mainly used for medical image segmentation [1]. When it was proposed, it was mainly used for lung nodule detection after cell wall segmentation and blood vessel extraction on the fundus retina. The network has some excellent segmentation effects in these areas.The U-Net network structure is like the capital "U". It is mainly composed of a convolution layer, a maximum pooling layer (down sampling), a deconvolution layer (up sampling), and a ReLU non-linear activation function. The entire network is divided into contracting path and expansion path, the contracting path is composed of 3 × 3 convolution and 2 × 2 pooling, which is repeated four times, the main purpose is to obtain the context information. The expansion path consists of copying and cropping the image before the corresponding maximum pooling layer, and stitching the manipulated images, performing 2*2 deconvolution and 3*3 convolution operations, and repeating the same four times, the last time Perform a 1*1 convolution operation.

2.2 LSD algorithm

The LSD algorithm is different from the Hough transform algorithm.It does not rely on edge detection.It can directly extract straight line segments from grayscale images, and it has less calculation than the Hough transform .Because of its fast and accurate characteristics, the algorithm is widely used in line detection and image recognition. The main idea of the algorithm is that the important concept of the algorithm is the gradient of each pixel, calculating the gradient value and the gradient direction of each pixel [2]. The image is calculated using a 2*2 template block, as shown in Figure 1.

$\ddots$	$\vdots$	$\vdots$	$\cdot^{\cdot^{\cdot}}$
$\cdots$	$i(x, y)$	$i(x+1, y)$	$\cdots$
$\cdots$	$i(x, y+1)$	$i(x+1, y+1)$	$\cdots$
$\cdot^{\cdot^{\cdot}}$	$\vdots$	$\vdots$	$\ddots$

Fig.1 Gradient calculation template

It's here $i(x, y)$ is the gray value of the pixel (x, y). the gradient value for this point ,The calculations $G(x, y)$ are shown in formulas (1)-(3):

$$g_x = \frac{i(x+1, y) + i(x+1, y+1) - i(x, y) - i(x, y+1)}{2} \tag{1}$$

$$g_y = \frac{i(x, y+1) + i(x+1, y+1) - i(x, y) - i(x+1, y)}{2} \tag{2}$$

$$G(x, y) = \sqrt{g_x^2(x, y) + g_y^2(x, y)} \tag{3}$$

The duality between the area representation and the boundary contour representation of the object or surface,which is a well-known fact in computer vision[3].In the literature, the line segment is mapped to the dual area, and the line segment detection problem is regarded as the coloring problem of the area, Through an effective and straightforward method to calculate the dual region representation, by re-formulating the LSD into the equivalent region coloring problem, we solve the above-mentioned challenges of dealing with local ambiguity and class imbalance in a principled way.

3. Improved U-Net mode

3.1 shallow U-Net model

The improved U-Net model proposes a shallow layer based on the original model.The For unnecessary deep information, no information feature extraction is performed.It is mentioned in the related literature that shallow structures can extract more low-level information from the image. The deep network structure will expand the receptive field.The role of the convolution operation is largely
to expand the receptive field again to extract image high-level information.The improved model is shown in Figure 2.This is a redundant feature extraction for the simple image that obscures the image description. Therefore, in order to extract the shallow information of the image, the reduction of the convolution operation and the reduction of the network depth were originally performed. The network model still presents a U-shape.

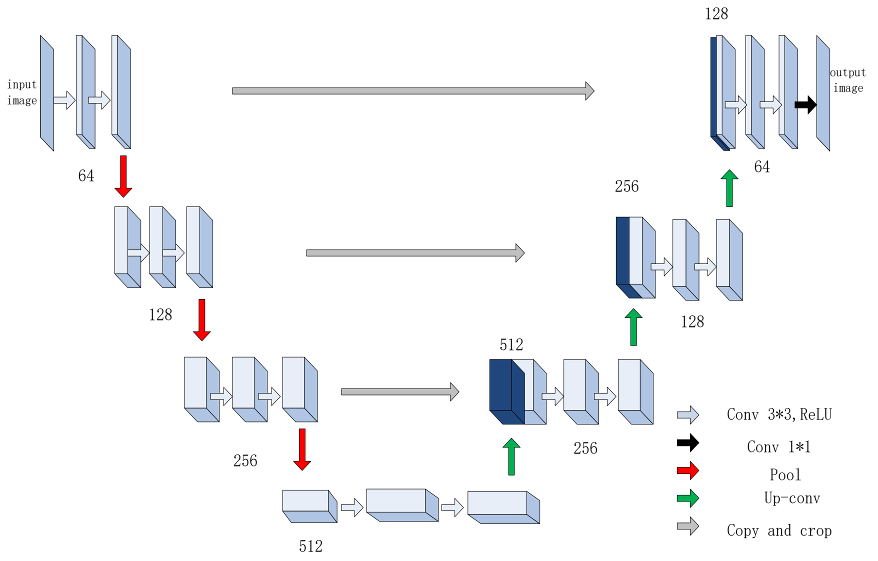

Fig.2 Improved U-Net network model

3.2 Enhanced low-level information extraction

In order to strengthen the low-dimensional information, in the improved shallow network, the extraction of shallow information is enhanced. In the contracting path, the input image undergoes multiple convolution and pooling operations to achieve a sublimation from low-level feature information to high-level feature information. so compared to the expansion path of the right half of the network, the biological contracting path of the left half of the network contains more low-level information.In order to increase the low-level information and improve the resolution of the extended path generation map.It is easier to lose the low-dimensional features of the image by making a single skip architecture connection.In order to connect the low-level features to the extended path of the network through skip architecture, the number of skip architectures between the contraction path and the extended path is increased. Better optimization to extract shallow information. This can better make up for lost image textures and contours, and can improve the resolution of the image. The schematic diagram of the skip architecture connection is shown in Figure 3

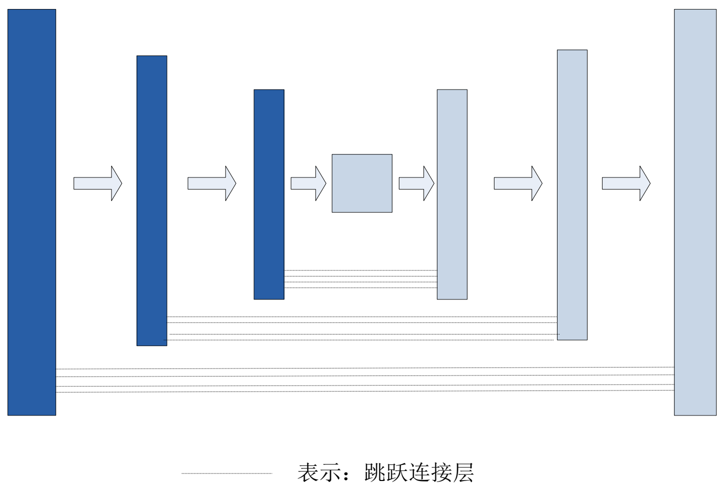

Fig.3 Schematic diagram of adding skip architecture

Increased the information transmission channel between the convolutional layer and deconvolution, so that for the colorless and simple texture images in this article, the features can be better extracted, and the shallow features are more connected to the deep features. Weakened the existence of deep features and ensured maximum extraction of low-level information.

4. Conclusions

The experimental data in this article are the public datasets Wireframe dataset and York Urban. The experimental environment is Ubuntu 16.04 system, deep learning framework Pyorch 1.0. The number of training iterations is 2000. In this paper, the improved network has a certain improvement in detection speed compared with the previous detection. The running speed before the improvement is 10.3 frames per second. After the improvement, the running speed is 13.6 frames per second. There is a certain increase in speed.The experimental results are shown below.

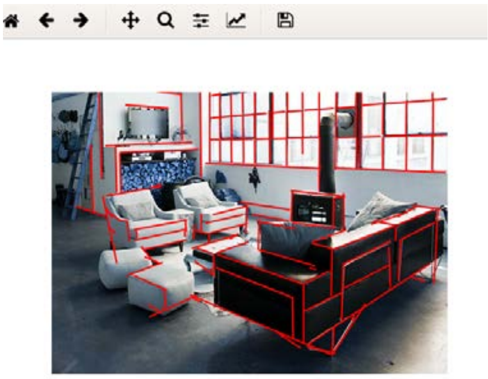

Fig.4 (a)

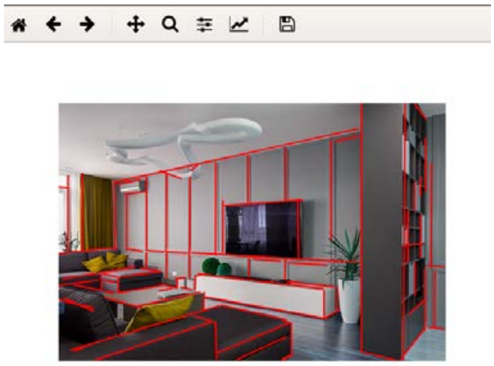

Fig.4 (b)

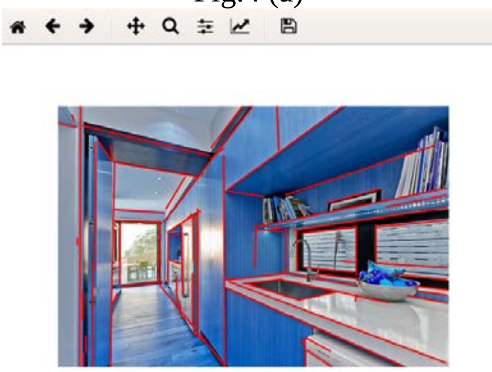

Fig.4 (c)

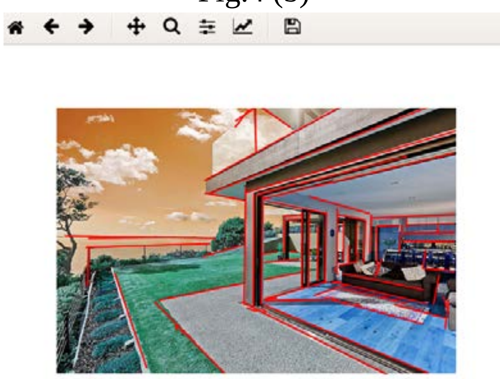

Fig.4 (d)

Fig.4 Experimental results

This paper proposes a line segment detection method based on an improved U-Net network for simple texture images.Based on the duality between the area representation of the object or surface and the boundary contour representation, the idea of converting line segment detection into region segmentation is proposed. Utilize the improved U-Net network to extract more low-level information of the image for line segment detection. From the experimental results, the improved U-Net network can effectively perform line detection, bu but there are also missed detections. In the future, we need to further study to overcome these shortcomings.

References

[1] Zhuang J.Ladder Net: multi-path networks based on U-Net for medical image segmentation [J].ar Xiv preprint ar Xiv: 1810.07810v2, 2018

[2] Von Gioi R G, Jakubowicz J, Morel J M, etal. LSD: a Line Segment Detector [J]. Image Processing on Line, 2012 (2): 35-55.

[3] Nan Xue, Song Bai , Fudong Wang, Gui-Song Xia .Learning Attraction Field Representation for Robust Line Segment Detection. IEEE Conference on Computer Vision and Pattern Recognition, 2019.

[4] Ronneberger O, Fischer P, Brox T.U-Net: convolutional networks for biomedical image segmentation [C] Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2015: 234-241.