

PAPER • OPEN ACCESS

Ocular Rectus Muscle Segmentation Based on Improved U-net

To cite this article: Cong Wu *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **533** 012058

View the [article online](#) for updates and enhancements.

Ocular Rectus Muscle Segmentation Based on Improved U-net

Cong Wu, Jinhao Zhan, Yixuan Zou, Fagang Jiang, Junjie Yang

Hubei University of technology, No.28, Nanli Road, Hong-shan District, Wuchang, Wuhan, Hubei China

zhanjinhao5@gmail.com

Abstract. In recent years, deep learning has made great progress in computer vision, and shown a good application prospect in reading medical images. TAO (thyroid-associated ophthalmopathy) is one of the most common orbital diseases in adults is autoimmune disease, the exact pathogenesis is not clear. In hospitals, CT images are usually used for disease analysis and the distortion of ocular rectus muscles is one of the main causes of TAO, thus the separation of the ocular rectus muscle is of great practical significance. However, the artificial segmentation of the ocular rectus muscle is a time-consuming task and relies heavily on the experience of the operator. In this paper, we improved the traditional U-net with GoogLeNet inception module and apply it on the segmentation of ocular rectus muscles. The experimental results show that the segmentation results of our method are more accurate and efficient than the traditional algorithm, and it is helpful for doctors to diagnose patients' diseases.

1. Introduction

TAO (thyroid-associated ophthalmopathy) is one of the most common orbital diseases in adults is autoimmune disease, the exact pathogenesis is not clear. [1] The severity of TAO is divided into three grades: mild, moderate and severe. In mild grade of TAO, It had little effect on the quality of life of the patients. Although mild TAO did not significantly threaten the vision of the patients, it had a great impact on their life and work. When it comes to severe TAO, the patient developed thyroid disease-related optic neuropathy and/or corneal lesions, and should take immediate intervention measures such as glucocorticoid intravenous shock [2]. The disease of TAO will cause swelling of the extraocular muscles and stimulate the differentiation of some pre-adipocytes into adipocytes [3]. There are three types of eye CT scanning in hospitals: transverse position, sagittal position and coronary position. Among them, the CT images of the transverse position scan shows the stripe structure of the left and right rectus muscle of the patient's eyes. The contour of the ocular rectus muscle was most clearly observed and was diagnosed by its morphology, such as the degree of obesity of ocular rectus muscle. See figure 1 for details.



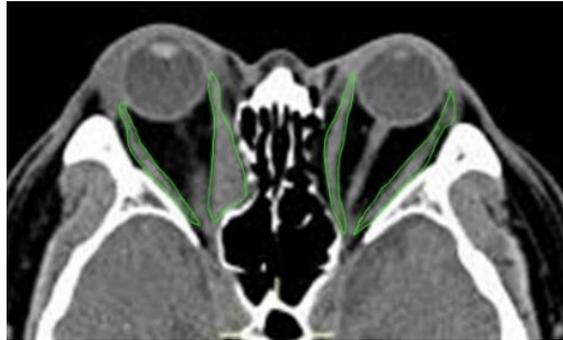


Figure 1. As shown in the figure, this is the image of the transverse position of the human eye CT image. The green contour delineates the left and right rectus muscle of the eye, which is the target area to be segmented.

2. Related Work

At present, in the field of computer vision, convolution neural network is the most effective depth model of deep models. The ImageNet Large Scale Visual Recognition Challenge (ILSVRC) witnessed the development of CNN model in the field of computer vision. From AlexNet[4] in 2012, VGGNet[5] and GoogLeNet[6] in 2014 to the Residual Network in 2015, the size and depth of CNN increased sharply. At the same time, the error rate of objects recognition task is also decreasing rapidly. At present, in 1 000 categories of object recognition task, CNN model achieved less error rate than human do. In addition, CNN has been successfully applied in a lot of general recognition tasks, such as object detection, semantic segmentation and contour detection.

Our approach draws on recent successes of deep nets for image segmentation. The most straightforward way of improving the performance of deep neural networks is by increasing their size. This includes both increasing the depth, the number of levels of the network and its width: the number of units at each level[6].The progress of Image segmentation give the credit to the development of image classification. Image segmentation changes the final full connection layer into the full convolution layer so as to achieve the purpose of segmentation. FCN [7] proposed the full convolution network and achieved remarkable results in different deep learning tasks.

On the respect of medical images segmentation, U-net [8] has been unanimously recognized, and has been greatly ahead in the IBIS cell segmentation task with only 30 training images. And that, the related scholar in The field of deep learning put forward a few variants of U-net, literature [9] discusses The utilize of resnet in U-net, It deepen the U-net without causing vanishing gradient problem and enhance the feature extraction capacity of the network. Literature [10] introduced dense block, which can deliver low-level features in the previous layer, and work parallel with higher-level features in higher layers so as to realize the efficient reuse of features. This paper explores the performance improvement of inception module in U-net.

3. Network Achitecture Design

Inspired by NIN [11] and GoogLeNet V3[12], we found that 1×1 convolution layer can fuse information of network layers in different channels, in other words, enlarging the width of the network helps to extract image features. In GoogLeNet v3, 3 inception module_a, 5 inception module_b and 2 inception module_c are inserted into the network structure after several convolution layers.

It has been proved that dimensional reduction and parallel structures of the Inception modules which allows for mitigating the impact of structural changes on nearby components[12] and the utilise of inception modules can reduce parameters of the whole network: GoogleNet employed only 5 million parameters, which represented a $12\times$ reduction with respect to its predecessor AlexNet, which used 60 million parameters. Furthermore, VGGNet employed about $3\times$ more parameters than AlexNet[12].

The structure of the three inception modules are shown in figure 2.

The fact that different inception module with different amount being placed in different parts of the network illustrates their different functions: module_a can do feature extraction in a relatively shallow network depth, accordingly, inception module_b and inception module_c can perform better in middle depth and a deeper depth in network.

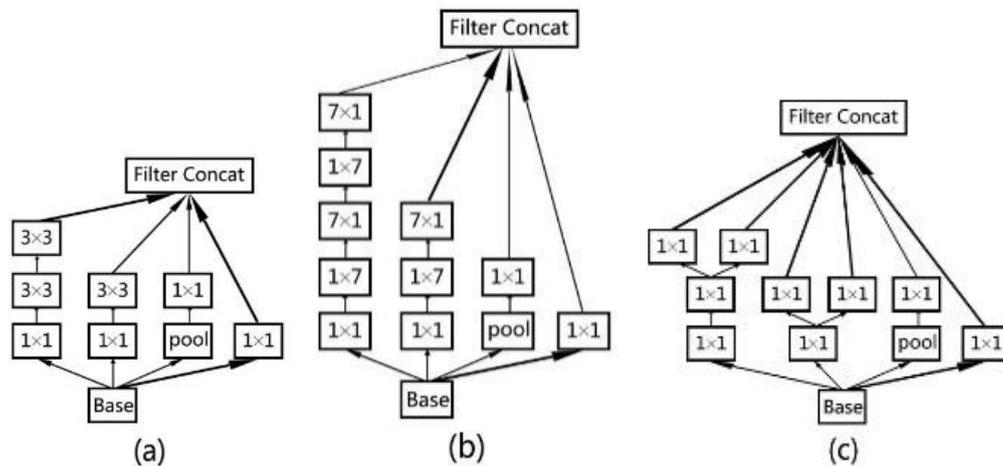


Figure 2. GoogLeNet inception module. (a).inception module_a appears after a few convolutional layers in the network,(b).inception module_b appears after inception module_a,(c).inception module_c appears after inception module_b.

Among various image segmentation algorithms, U-net can train without a large amount of training samples and perform well in segmentation tasks. In ISBI cell tracking challenge 2014 and 2015, U-net achieved an average IOU (intersection over union) of 92%, which is significantly better than the second best algorithm with 83% [8]. So it's a segmentation algorithm for medical image segmentation.

In this paper, we design several networks based on U-net to explore how to insert inception modules into U-net to get an ideal network architecture thus to increase segmentation accuracy. U-net network belongs to an encoder-decoder framework structure. In the encoding phase, the network extracts image features through several convolutional layers. After getting the feature map, the network does a few transpose convolutions and meanwhile merges layers in the encoder phase in the same shape to get the segmentation result. U-net is the most effective segmentation network in the field of medical image segmentation. The several network structures designed in this paper are as shown in Table 1:

Table 1. The configuration of seven improved U-net.

net	A			B			C			D			E			F			G		
Smallest image size	8			8			8			8			8			32			32		
Insert when picture size is	128	64	32	64	32	16	32	16	8	32	16	8	32	16	8	128	64	32	128	64	32
Insert phase	Encoding phase			Encoding phase			Encoding phase			Encoding phase			Encoding phase			Encoding-decoding phase			Encoding phase		
Module amounts	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c
	1	1	1	1	1	1	1	1	1	2	3	1	3	5	2	1	1	1	1	1	1

Through the comparison between A and G networks, we can know the influence of network depth on network training. Through comparative experiments of A, B and C networks, it can be seen that the inception module is best placed in that location. Comparison experiments of network F and G can tell us whether it is better to insert the inception module in the encoding-decoding stage or not. Through the

comparison experiment of C, D and E networks, we can know the influence of different amount of inception module on the network, and finally obtain the optimal network in current task.

4. Experiments

The experimental environment is as follows:

Hardware: Inter(R) Core(TM) i7-6700, 12G, NVIDIA GeForce GTX TITIAN X

Software: Windows 64bit, tensorflow1.7.0, keras2.0.8, CUDA8.0, cunn5.1

4.1. Dataset

Since there is no way to get the label of the test set of IBSI data set, 30 training sets are divided into 20 training sets and 10 test sets. The data expansion method is shown in table 2:

Table 2. Configuration of data augmentation.

Data expansion configuration	
Rotation	0.20
Width adjustment	0.05
Height adjustment	0.05
Shear	0.05
Scale	0.05

The TAO data set comes from the partner hospital. There are 89 CT image data sets of the transecting position of TAO. The distribution of data sets is training set: test set =1:1. All images in the dataset have been desensitized to patient information. To preprocess data sets, we first segment the ROI area which contains the ocular rectus muscle, and then resize the images to 512×512. After that the labels of the dataset are to be made. The black and white mask tags are made with the image labeling tool – labelme.

4.2. Experiment Result

We designed seven network based on U-net and inception module, aiming to explore ideal network architecture of improved U-net under different network configuration. Through experiment, all experiment result in IBSI task are as shown in table 3:

Table 3. Configuration of data augmentation.

net	Training accuracy	Training loss	Test PA	Test SN	Test SP
A	0.9637	0.0833	0.9106	0.9446	0.6858
B	0.9577	0.1826	0.9068	0.9487	0.6693
C	0.9586	0.1816	0.9214	0.9307	0.7312
D	0.9623	0.5041	0.8851	0.9624	0.5768
E	0.9603	0.5067	0.8785	0.9651	0.5481
F	0.9486	0.1948	0.8937	0.9547	0.6154
G	0.9621	0.5096	0.8830	0.9588	0.5695

The definitions of evaluation indexes PA, SN and SP are as shown in table 4:

Table 4. Defination of Segmentation Indexes

Index	PA	SN	SP
defination	$\frac{S_{pred} \cap S_{label}}{S_{pred}}$	$\frac{S_{pred} \cap S_{label}}{S_{label}}$	$\frac{S_{pred_bg} \cap S_{label_bg}}{S_{label_bg}}$

S_{pred} represents the target area predicted result, S_{label} represents the target area of label image, S_{pre_bg} represents the background area of the predicted result, S_{label_bg} represents background area of the label image.

According to the data in the table, it can be seen that: 1) when the smallest image size in the network is 8, the network performs better. 2) from the perspective of insertion location of inception module, insert module_a when image size is 128×128 and module_b and module_c when image size is 64×64 and 32×32 respectively can make the network perform better; 3) To insert 3 module_a, 5 module_b and 2 module_c is the better; 4) It's better to insert inception module in encoding-decoding phase. In conclusion, in the current task, the ideal network structure is to insert 3 module_a when image size is 128×128 , 3 module_b when image size is 64×64 and 2 module_c when image size is 32×32 in encoding-decoding phase of the network symmetrically. With this ideal improved U-net, the segmentation accuracy is significantly improved.

The comparison between loss and accuracy in the training process of original u-net and optimal improved u-net is shown in figure 3:

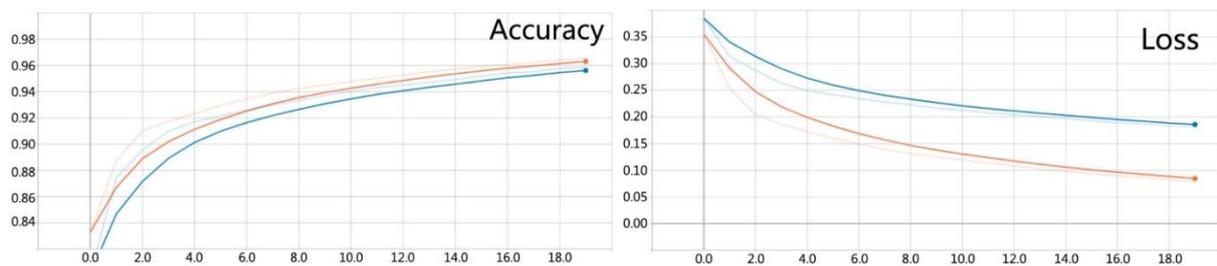


Figure 3. The upper figure shows comparison of accuracy convergence curve and the figure below shows the loss reduction curves of original U-net and improved U-net on the same task

As can be seen from figure 3, the improved U-net has a faster convergence speed, a higher segmentation accuracy and a faster loss reduction speed than the original U-net training. The segmentation results of the improved U-net on the ocular rectus muscle in the transverse position are shown in figure 4:

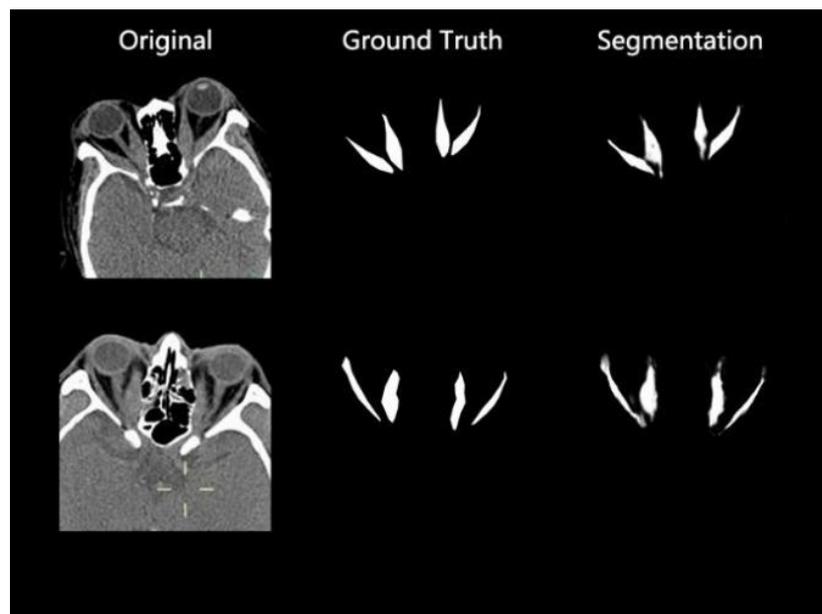


Figure 4. segmentation result of TAO ocular rectus muscle

Compared with original U-net, the accuracy of the improved u-net in the segmentation task of the ocular rectus muscle is shown in table 5:

Table 5. Configuration of data augmentation

	PA	SN	SP
U-net	0.8799	0.8017	0.9955
Our method	0.9537	0.8388	0.9984

5. Conclusion

This paper presents an Ocular rectus muscle segmentation algorithm based on U - Net, the CT images of the ROI area is extracted firstly, then after preprocessing the image as input of network, we trained the ideal network do segmentation of the CT image and finally obtain the ocular rectus muscle. The results of finalization and quantitative experiments show that compared with the traditional segmentation algorithm, the segmentation effect of our method is more accurate and efficient, and it is an auxiliary for doctors to do segmentation of ocular rectus muscle.

Acknowledgments

Thank you for the valuable data provided by Wuhan Union Medical College Hospital, Hubei province, China and relevant technical and knowledge guidance provided by professor Fa-Gang Jiang and doctor Jun-Jie Yang.

References

- [1] Guoqiang li. The application of CT and MRI in thyroid - related eye diseases diagnosis [J] . Journal of hubei university of traditional Chinese medicine, 2013, 15 (3) : 103-104.
- [2] Fuping liu. Evaluation and treatment strategy of thyroid - related eye diseases [J]. Chinese journal of clinical physicians: electronic edition, 2010, 4(1):4-7.
- [3] Yanjiao Zhang. Current status of thyroid - related immunosuppressive therapy for eye diseases [J]. Medical review, 2004, 10(9):553-555.
- [4] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]// International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012:1097-1105.
- [5] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. Computer Science, 2014.
- [6] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2015:1-9.
- [7] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]// Computer Vision and Pattern Recognition. IEEE, 2015:3431-3440.
- [8] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation[J]. 2015, 9351:234-241.
- [9] Drozdal M, Vorontsov E, Chartrand G, et al. The Importance of Skip Connections in Biomedical Image Segmentation[J]. 2016:179-187.
- [10] Jégou S, Drozdal M, Vazquez D, et al. The One Hundred Layers Tiramisu: Fully Convolutional DenseNets for Semantic Segmentation[C]// Computer Vision and Pattern Recognition Workshops. IEEE, 2017:1175-1183.
- [11] Lin M, Chen Q, Yan S. Network In Network[J]. Computer Science, 2013.
- [12] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the Inception Architecture for Computer Vision[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2016:2818-2826