# A Coarse-to-Fine Model Structure with Vessel Direction Enhancement for Retinal Vessel Segmentation

***Abstract*—Vessel segmentation plays an essential role in many clinical applications. Among them retinal vessel segmentation is a significant task for developing the computer-aided diagnosis system for retinal diseases.Although supervised methods have achieved state-of-art performance, acquiring expert annotation is laborious and mostly limited for two-dimensional datasets with a small sample size. Many methods have been proposed to make the data augmentation and these methods have made great contributions to the accurate segmentation of retinal images. Our method is based on the Study Group Learning(SGL) method, trying to use the erased labels in SGL model. In the enhancement part, we add a coarse vessel segmentation task and a vessel direction prediction task as the new auxiliary enhancement. In the segmentation part, all the enhancement outputs will be sent into the segmentation module for fine results. So it is a coarse-to-fine model structure. Our changes of the model can improve the vessel segmentation results.**

***Index Terms*—Retinal vessel segmentation, Image enhancement.**

## I. Introduction

Retinal inspection is an effective approach for the diagnose of multiple retinal diseases. Most disorders can be detected and dealt with correctly if we make the accurate segmentation of the retinal image segmentation. What's more, the same as other vessel disorders such as coronary artery stenosis and abdominal aortic aneurysm, the accurate segmentation of the DSA images will lead to the successful interventional surgical treatment. So it is quite important to develop the advanced vessel segmentation algorithms.

Plenty of previous efforts have been made in automatic retinal vessel segmentation. Since the boost development of Deep Learning, the end-to-end methods have paved the way for accurate segmentation of retinal vessel segmentation. However, the data-driven methods suffer from the number of samples and the models will easily over-fit the small-scale training data. In order to make the data augmentation and improve the model robustness. The Study Group Learning method [1] augmented the training dataset DRIVE and CHASE in order to make the model more robust and less sensitive to the noise. However, as mentioned in SGL method, they cannot make full use of the thick vessel data generated by using the skeleton-tracing approach [2]. So, we proposed a method based on the SGL model, and our idea about the vessel direction prediction as a enhancement method came from the work by Jena $et\,al.$ [3].

We improved the enhancement part of the SGL model, still using the U-net structure as the enhancement module,but we added the prediction of thick vessel $I_t$ and the prediction of vessel direction vectors $I_v$. The prediction of thick vessel $I_t$ and vessel direction vectors $I_v$ will be the input of the segmentation module, which is still the U-net structure, as the same of SGL model. Experiments show that our model can provide accurate segmentation of retinal images as well as SGL model but with less training epoches.

## II. Related works

Recently, data-driven based methods utilize UNet-based model [4] or its variants to achieve significant performance compared with traditional methods. The Study Group Learning method [1] is still the state-of-art work in the dataset of DRIVE and CHASE segmentation.

Nowadays attention have been paid not only to the segmentation task but also to the vessel image enhancement and vessel width estimation tasks. The work by Jena $et\,al.$ [3] shows a novel Self-supervised Vessel Enhancement Using Flow-Based Consistencies, using the transformation and the template tricks to enhance the vessel image. Such enhancement method is a self-supervised method, which do not need the annotation. The vessel width estimation method by Li R Q$et\,al.$ [5] treat the problem of vessel width estimation as the pixel-level regression. Such idea also inspires us in our work.

## III. Methodology

### A. Model Structure

The proposed model structure is illustrated in Fig. 1. Similar to the SGL model, we utilize the concatenated UNet consisting of the enhancement and segmentation modules to learn both the enhancement and the segmentation map. We also use the group study labels in SGL model in order to improve the robustness of our model. What is different from SGL model is that we add two more tasks for the enhancement module: the thick vessel prediction and vessel direction prediction. The thick vessel prediction and vessel direction prediction will be sent with the enhancement together to the segmentation module for the final segmentation.

### B. Thick vessel prediction

In order to make full use of the thick vessel data generated by using the skeleton-tracing approach [2]. We set one channel of the output of the enhancement module as the prediction of the thick vessel $I_t$ and it will be trained under the supervision of the thick part generated from the train labels. So the loss function of this part is set as:

$$L_t = CE(I_t, I'_t) \tag{1}$$

where $I_t$ is the output prediction of the thick vessel and $I'_t$ is the label generated by using the skeleton-tracing approach of the annotation label $I'$, CE is the cross entropy loss. The prediction of the thick vessel can provide the model with a coarse view of the vessel branch, and then the segmentation module can make the fine segmentation of the details of the vessel. It can be regarded as a coarse-to-fine process.

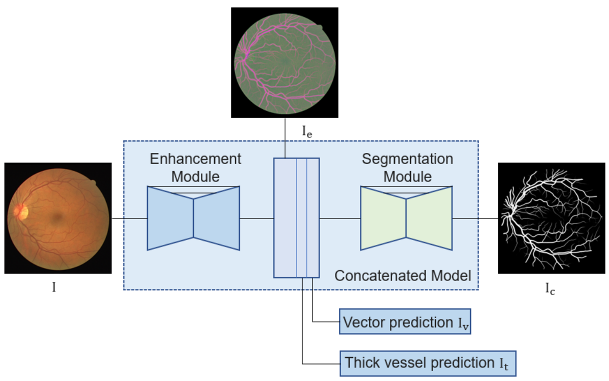


Fig. 1. The model structure.

### C. Vessel direction prediction

We got the idea of this from the work of Jena *et al.* [3] and the work of Li R Q*et al.* [5]. We set two channels of the output of the enhancement module as the 2-D direction of the vessel, in the work of Jena *et al.* [3] the directions should follow the constraint of Path Continuity and Profile Consistency. However, the training of the enhancement module under these constraints costs too much computational time, which has brought difficulties for its implementation. So we only multiply the estimated direction of the point with the estimated direction of its right and upper points, under the hypothesis that the points in the neighborhood share the same directions. So the loss function of this part should be set as:

$$L_v = -SUM(I_v \cdot I'_v) \tag{2}$$

where $I_v$ is the output of the vectors, and $I'_v$ is the transformed $I_v$ in order to make the dot product of the point's vectors and its neighbors'.

So the total loss function of the model is set as:

$$L = CE(I_c, I'_c) + \lambda_0 CE(I_c, I_c t') + \lambda_1 L_t + \lambda_2 L_v \tag{3}$$

where $I_c$ is the output of the segmentation module, the final segmentation of the vessel. $I'_c$ and $I_c t'$ are the labels generated in the SGL method, where the $I_c t'$ is the augmentation. $\lambda_0, \lambda_1, \lambda_2$ are all set as 1.0.

TABLE I
COMPARISON WITH OTHER BASELINE METHODS ON DRIVE DATASET.

| Method | Sensitivity | Specificity | DICE | Accuracy | AUC |
|---|---|---|---|---|---|
| R2U-Net | 0.7792 | 0.9813 | 0.8171 | 0.9556 | 0.9784 |
| LadderNet | 0.7856 | 0.9810 | 0.8202 | 0.9561 | 0.9793 |
| IterNet | 0.7791 | 0.9831 | 0.8218 | 0.9574 | 0.9813 |
| SA-UNet | 0.8212 | 0.9840 | 0.8263 | 0.9698 | 0.9864 |
| BEFD-UNet | 0.8215 | 0.9845 | 0.8267 | 0.9701 | 0.9867 |
| SGL(K=8) | 0.8380 | 0.9834 | 0.8316 | 0.9705 | 0.9886 |
| Ours | **0.8295** | 0.9845 | **0.8317** | **0.9707** | 0.9883 |

TABLE II
COMPARISON WITH OTHER BASELINE METHODS ON CHASE DB1 DATASET.

| Method | Sensitivity | Specificity | DICE | Accuracy | AUC |
|---|---|---|---|---|---|
| UNet | 0.7840 | 0.9880 | 0.7994 | 0.9752 | 0.9870 |
| DUNet | 0.7858 | 0.9880 | 0.8000 | 0.9752 | 0.9887 |
| IterNet | 0.7969 | 0.9881 | 0.8072 | 0.9760 | 0.9899 |
| SA-UNet | 0.8573 | 0.9835 | 0.8153 | 0.9755 | 0.9905 |
| SGL(K=8) | 0.8690 | 0.9843 | 0.8271 | 0.9771 | 0.9920 |
| Ours | 0.8686 | 0.9838 | 0.8241 | 0.9766 | **0.9920** |

## IV. EXPERIMENTS

### A. Training

While training the model, we randomly crop the images into 256 × 256 patches, and apply data augmentation including horizontal and vertical flip, rotation, transpose, and random elastic warping [6].We only train for 10 epocheson each task. All training process are made on two TITANX GPUs.

### B. Results

The result of our model are shown in the two tables, our model performs as well as these state-of-art works, with less epoches to train. Compared with the SGL model with at least 20 epoches, we only need 10 epoches of training.

### C. Analysis

Because of the modification of enhancement in our work, our model needs less training epoches but performs as well as SGL model, that shows the effectiveness of our method.

## V. SUMMARY

We make full use of the group study labels in SGL model in order to improve the robustness of our model. What is different from SGL model is that we add two more tasks for the enhancement module: the thick vessel prediction and vessel direction prediction. The thick vessel prediction and vessel direction prediction will be sent with the enhancement together to the segmentation module for the final segmentation. Our work performs as well as the other SOTAs but takes less training epoches.

## REFERENCES


[1] Zhou Y, Yu H, Shi H. Study group learning: Improving retinal vessel segmentation trained with noisy labels[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2021: 57-67.

[2] Zhang T Y, Suen C Y. A fast parallel algorithm for thinning digital patterns[J]. Communications of the ACM, 1984, 27(3): 236-239.

[3] Jena R, Singla S, Batmanghelich K. Self-supervised vessel enhancement using flow-based consistencies[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2021: 242-251.

[4] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015: 234-241.

[5] Li R Q, Bian G B, Zhou X H, et al. Vessel Width Estimation via Convolutional Regression[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2021: 600-610.

[6] Simard P Y, Steinkraus D, Platt J C. Best practices for convolutional neural networks applied to visual document analysis[C]//Icdar. 2003, 3(2003).

[7] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.