ACAT 2022

21st International Workshop on Advanced Computing and Analysis Techniques in Physics Research

Flow-Unet for High Dimensional Image Semantic Segmentation

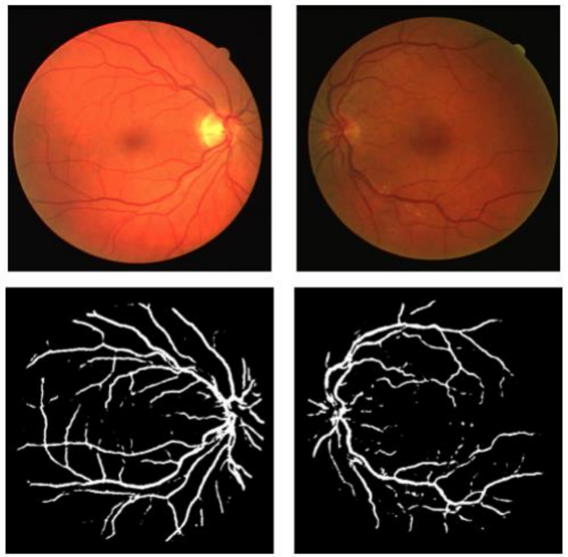
Xiaomeng Qiu[1], Yu Hu[2], Lin Wang[3], Wenjun Gu[1], Wei Song[1], Haolai Tian[2]

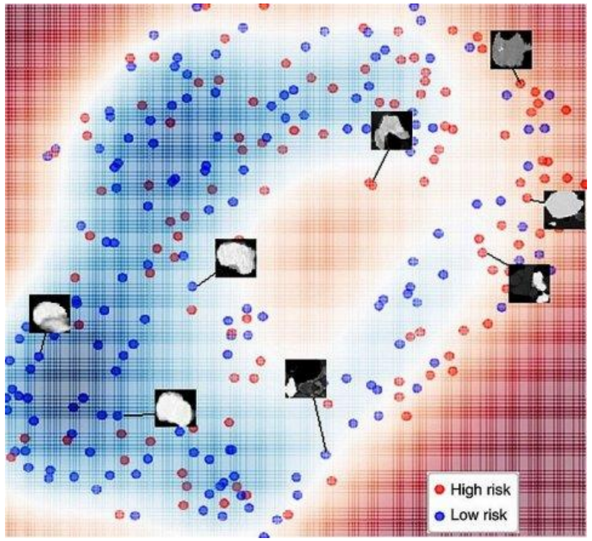
[1]Zhengzhou university, [2]Institute of High Energy Physics CAS, [3]Beijing WeMed Medical Equipment Co. Ltd.

October, 26 , 2022

Medical image processing

- Medical image processing is highly applicable. In which, the segmentation of organs, lesions and regions of interest in medical images is an important auxiliary means for medical diagnosis and surgical.
- The value of medical image processing is mainly reflected in:
  - Extract the region of interesting, to ignore the interference from other regions.
  - Measure the size of human organs or lesions, help the doctors to diagnose or modify patients' treatment plans
  - Obtain anatomical map information and provide original data for 3D reconstruction and visualization of medical images.

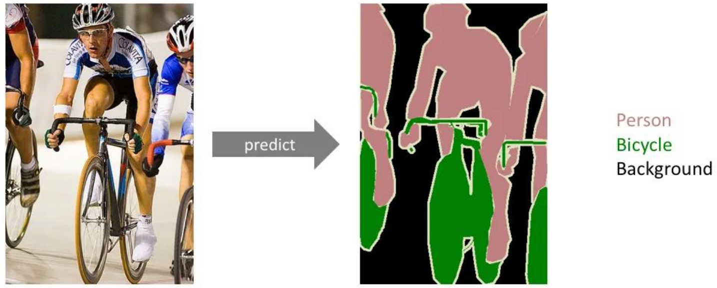




Semantic Segmentation

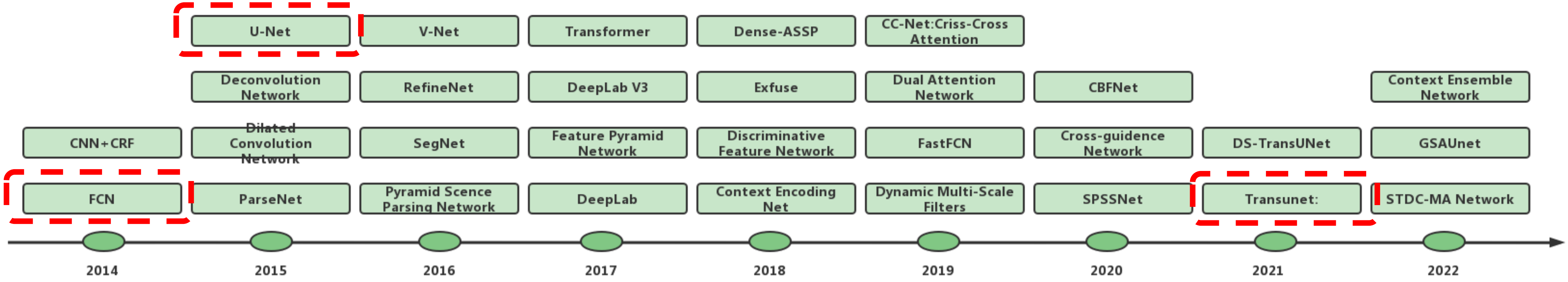
Image Tasks

- Image classification: use computer model to predict what objects are in the picture
- Object detection: detect what is in the picture (mark with a box, point …)
- Semantic segmentation: distinguish each pixel in the graph
- Instance segmentation: It is a combination of semantic segmentation and object detection, and different individuals of the same object need to be marked



Semantic Segmentation

- It is a typical computer vision problem, which involves taking some raw data as input and converting them into masks with highlighted regions of interest.
- Medical image processing often use the full pixel semantic segmentation, where each pixel in an image is assigned a category ID according to the object of interest to which it belongs to.



High Dimensional Image

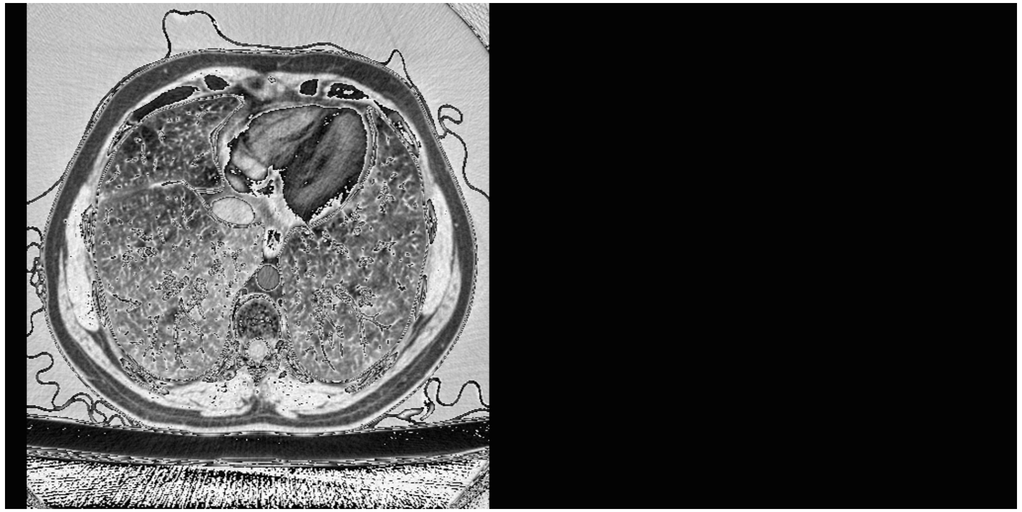
2D image data

- Flat image without depth information
- Only area, no volume

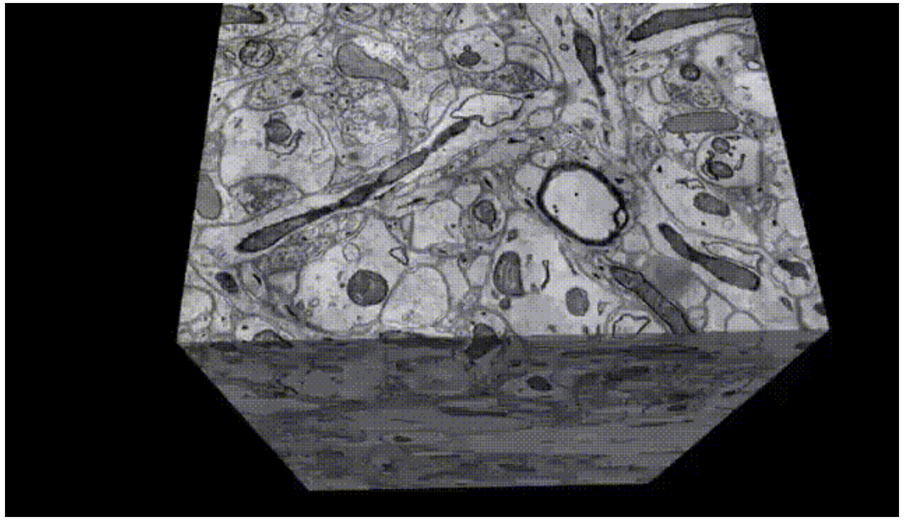
High dimensional image data

- Additional information of time or volume are attached

The accuracy of segmentation can be enhanced through the correlation between time series or depth information



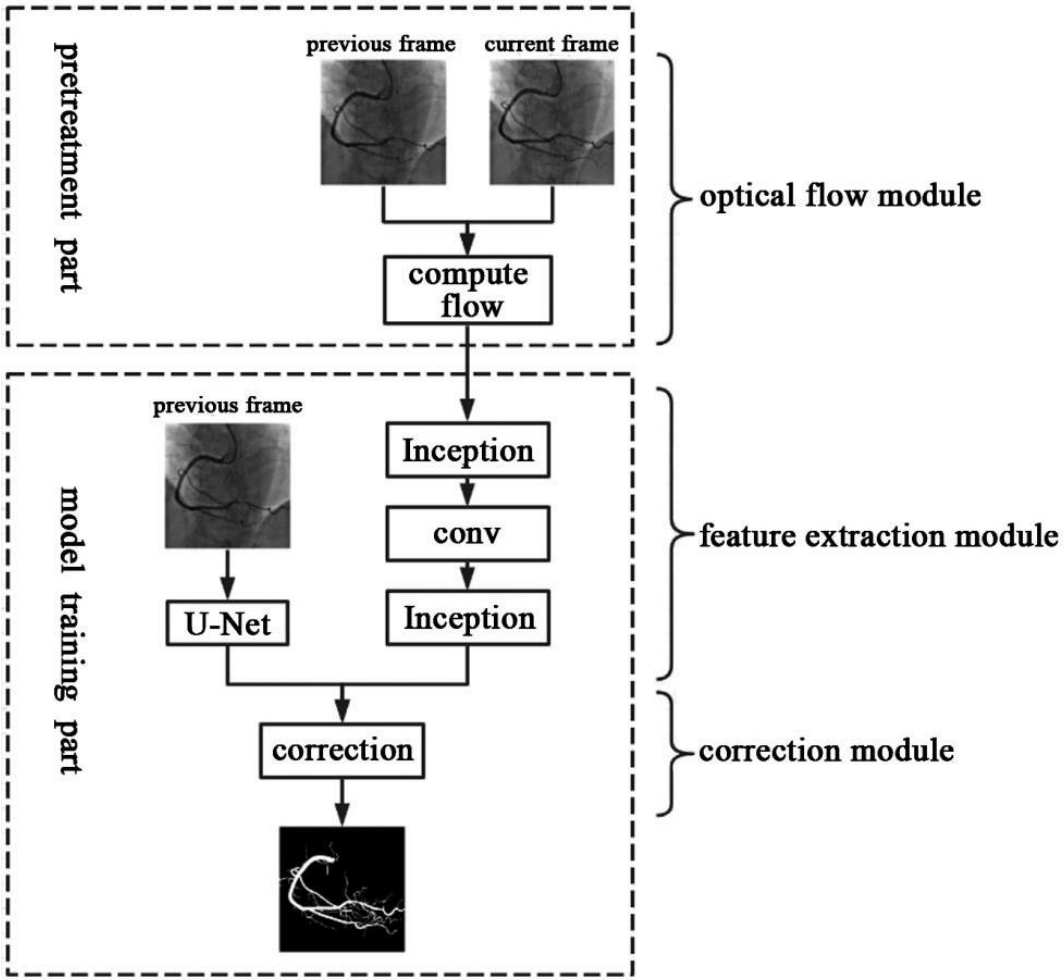
Temporal segmentation



Stereo segmentation

Our approach in Deep Learning —— Flow-Unet

- For images with timing sequence characteristic, the relationship between adjacent frames can be fully used to transmit the segmentation results of the previous frame to the next frame as prior knowledge.
- Optical flow can transmit timing sequence characteristic, obtain more segmentation information.



Flow-Unet

- **Optical Flow Module**

  Obtain optical flow information between adjacent frames

- **Feature Extraction Module**

  Previous frame: split with U-Net network

  Optical Flow: feature extraction using the inception module

- **Correction Module**

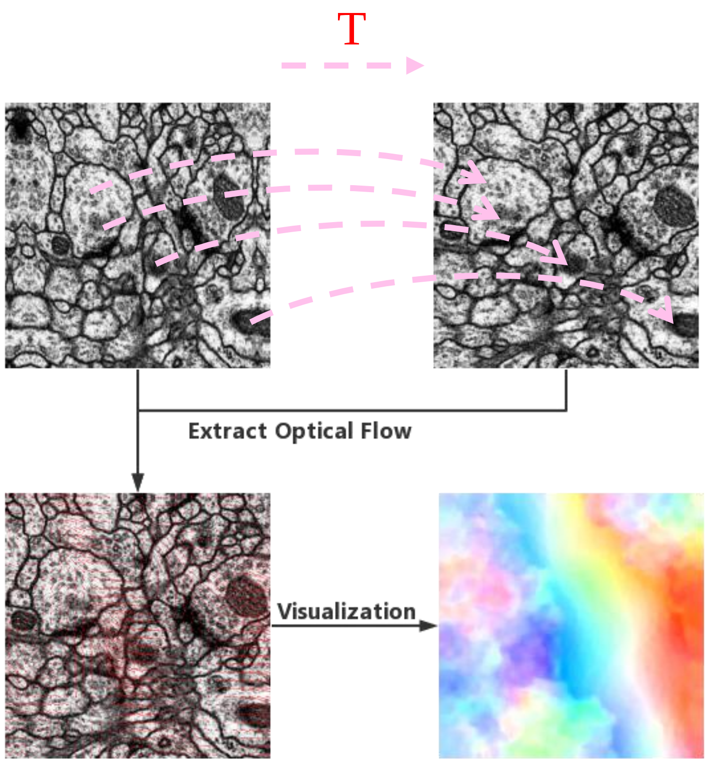
  Assign weight to the two parts obtained from the feature extraction module.

Considering both medical image features and temporal information.

Method——Optical Flow

- Describe the object in a scene that changes dynamically between two consecutive frames due to motion.
- Essence: a two-dimensional vector field.
- Each vector represents the displacement of the point in the scene from the previous frame to the next frame.

Optical Flow Equation



Hue-Saturation-Value graphic

The optical flow method is usually based on three assumptions:

① The light energy of the front and rear frames remains unchanged

② The motion of the same pixel between adjacent frames is small

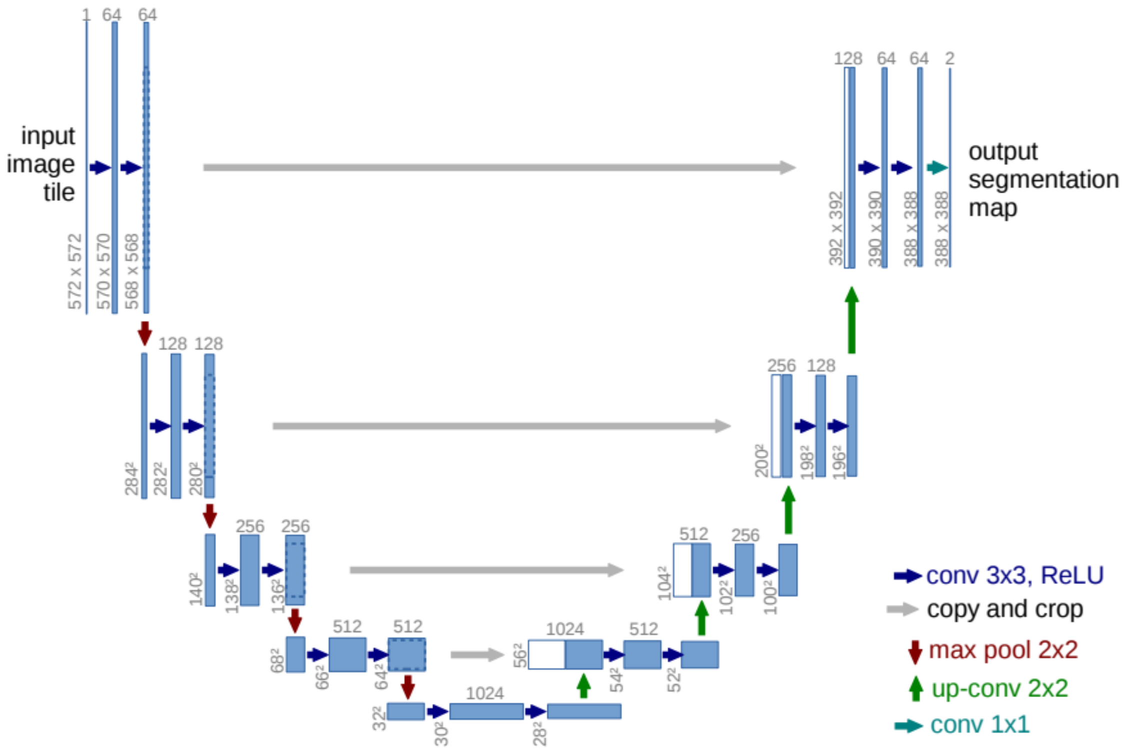
③ Motion of adjacent pixels is similar

$$f_x u + f_y v + f_t = 0$$

$$f_x = \frac{\partial f}{\partial x} \quad f_y = \frac{\partial f}{\partial y} \quad u = \frac{dx}{dt} \quad v = \frac{dy}{dt}$$

x,y: pixel coordinate $\quad$ f : gradient $\quad$ t : time

Method —— Feature Extraction



Main contributions: U-shaped structure network

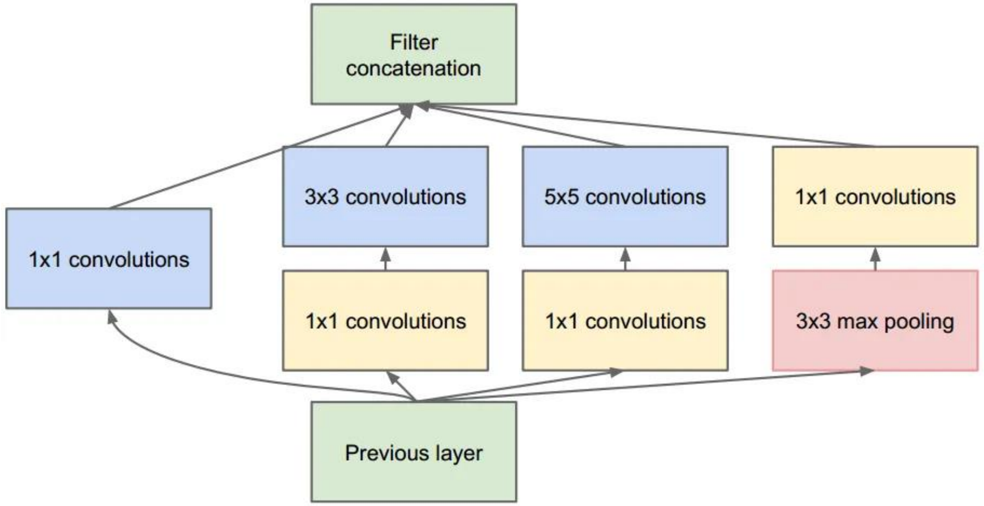
Advantage:

- The output result can locate the position of the target category
- Since the input training data is patches, it is equivalent to data enhancement, thus solving the problem of small number of biomedical images

Why U-net performs well in medical image segmentation?

- All features of medical images are important, so low-level features and high-level semantic features are important, so the U-shaped skip connection structure is more useful
- In the case of small order of magnitude, the segmentation SOTA model is not very different from the lightweight Unet
- In medical imaging tasks, we often need to design our own network to extract different modal features, so lightweight and simple Unet can have more operation space

Method——Feature Extraction



The main idea is: how to use a dense component to approximate or replace the optimal local sparse structure

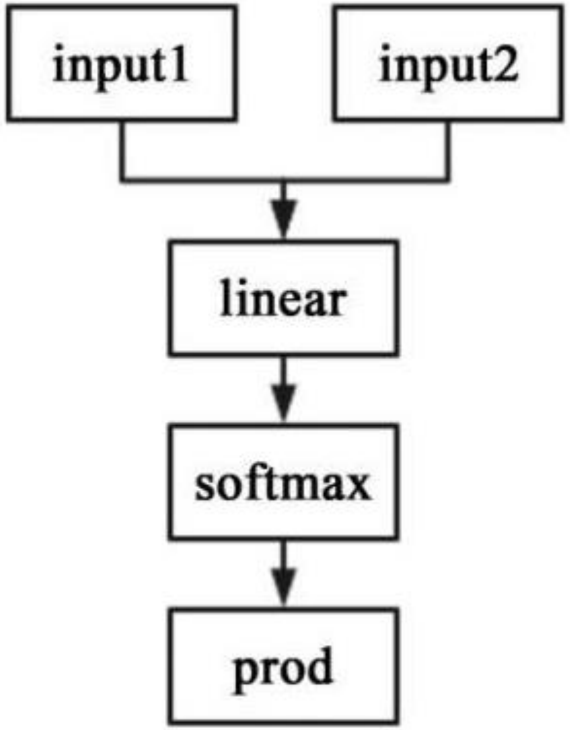
The advantages of Inception model :

- 1x1 convolution kernel is employed to add a layer of characteristic transformation and nonlinear transformation with less computation
- Batch Normalization is proposed to avoid gradient disappearance and accelerate convergence
- The last full connection layer is removed and replaced by the global average pooling layer, which greatly reduces the computation.

Two inception model are used and connected by a layer with two 1x1 convolution kernel , The role of 1x1 convolutional kernel:

- Cross channel feature integration
- Dimension raising and dimension reducing of feature channel
- Reduce parameter amount

Method——Correction



Correction module

- In order to modify the segmentation results by optical flow, the weights of the two segmentation results are allocated.
- Firstly, segmentation result of the current frame and the optical flow information are regarded as two separate parts, and input into the linear layer to obtain the weight information of the two parts.
- Then, Softmax operation will normalize the weights of those two parts. The real number field of the output of the linear model will be mapped to [0,1] to represent the effective real number space of the probability distribution, so as to obtain the weight occupied by the two parts respectively.
- Finally, the two segmentation results are respectively multiplied by their own weight matrix and concatenated to get the final segmentation result.

Method —— Loss function

Loss function:

$$Loss_{total} = 0.5 \times Loss_{dice} + 0.5 \times Loss_{bce}$$

Dice Loss is one of the most commonly used loss functions in medical image segmentation, used to evaluate the similarity between a predicted image and a labeled image:

$$Loss_{dice} = 1 - \frac{2|X \cap Y|}{|X| + |Y|}$$

Focuses on the similarity

Dichotomous cross-entropy loss is used to evaluate the loss of data when classifying each pixel in the segmentation process, and can measure the degree of difference between two different probability distributions in the same random variable:

$$Loss_{bce} = -\sum_{i=1}^{n} y_i \log(x_i)$$

Keep the pixels with smooth gradients

Experiment

- Dataset
  - Coronary angiogram: This data set is the real data of the medical company, which is the coronary angiography. There are 1200 pictures in the training data, which are divided into training set and validation set according to the proportion of 8:2, and 38 pictures in the test set.
  - Electron microscope of Drosophila melanogaster: Open data set provided by ISBI Challenge. The data are 30 groups of images from the ventral nerve cord of Drosophila melanogaster, each group is composed of several images.
  - Comprehensive Organogram of Healthy Abdomen(CHAOS): This experiment only uses CT images, and the data format is DICOM. The training set contains 2050 data, the validation set contains 266 data and the test set contains 558 data.
- Model
  - We compared the proposed model with other popular models.
- Training
  - Run 600 epochs of mini-batch gradient descent(Adam) with an initial learning rate of 10e−4 (decay 1% per 10 epochs) and a batch size of 4.
- Test
  - Calculation Pixel Accuracy, dice coefficient and Intersection over Union to evaluate.

Metric

Dice Similarity

$$\text{Dice} = \frac{2 \times \text{TP}}{2 \times \text{TP} + \text{FP} + \text{FN}}$$

calculate the similarity of two samples

Pixel Accuracy

$$\text{PA} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}$$

Calculate the proportion of actual positive examples in positive examples

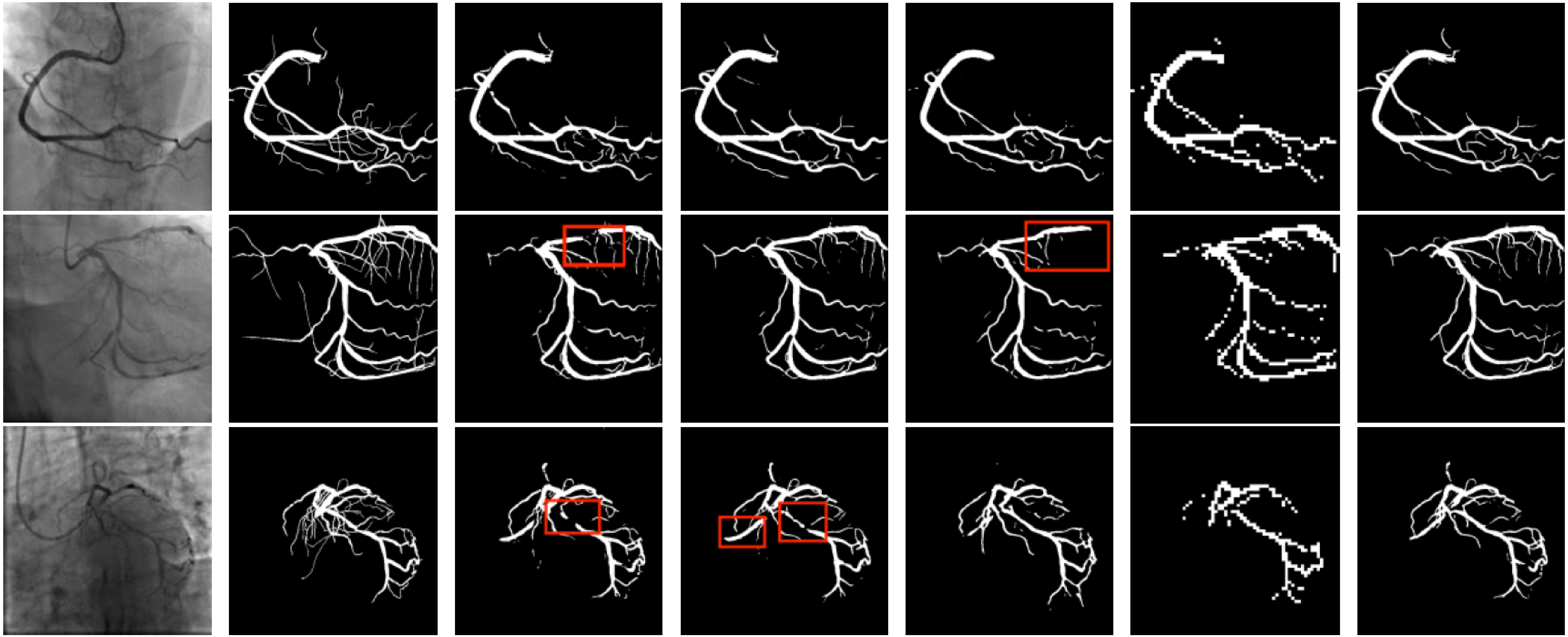
Intersection over Union

$$\text{IoU} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}}$$

Calculate the ratio between the intersection and union of a category prediction result and the real label

Result —— Coronary angiogram dataset

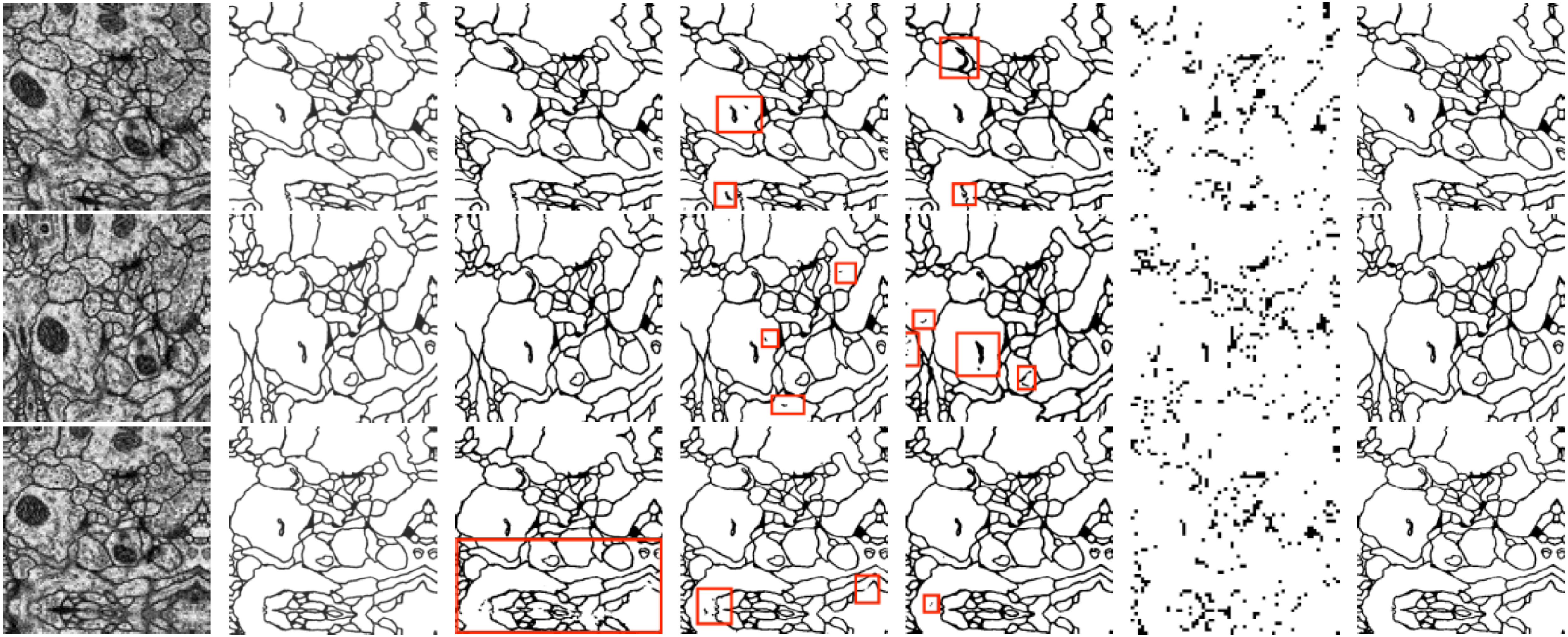
Get better results in the connectivity of blood vessels, and can obtain more detail on the small blood vessels.



a. input b. GT c. U-Net d. U-Net++ e. PspNet f. CSANet g. ours

Result —— Drosophila electron microscope dataset

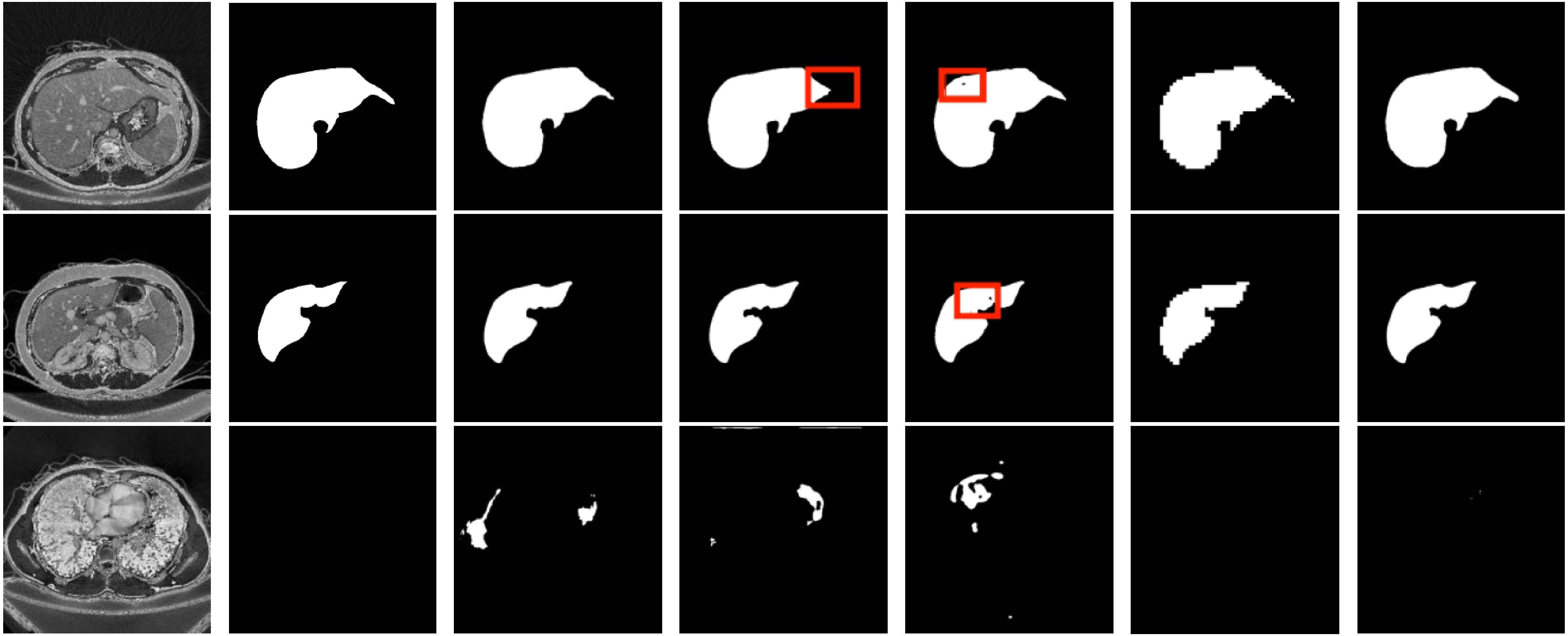
Overcome the breakage and noise. Furthermore , the segmentation information in the unlabeled area is also more accurate.



a. input b. GT c. U-Net d. U-Net++ e. PspNet f. CSANet g. ours

Result —— CHAOS dataset

The segmentation result is more accurate without noise and small black hole phenomenon.



a. input b. GT c. U-Net d. U-Net++ e. PspNet f. CSANet g. ours

Result

In quantitative analysis, Flow-Unet show best performance over other methods.

		Dice	PA	IoU
Coronary Angiogram	U-Net	0.7966	0.9915	0.8376
	U-Net++	0.8042	0.9736	0.6725
	PspNet	0.7504	0.9886	0.7725
	CSANet	0.6696	0.9569	0.5034
	Ours	**0.8102**	**0.9928**	**0.8589**
Electron Microscope of Drosophila Melanogaster	U-Net	0.9882	0.9811	0.9767
	U-Net++	0.9863	0.9780	0.9729
	PspNet	0.9830	0.9730	0.9665
	CSANet	0.9290	0.8837	0.6874
	Ours	**0.9924**	**0.9878**	**0.9850**
CHAOS	U-Net	0.9575	0.9953	0.9186
	U-Net++	0.9652	0.9961	0.9327
	PspNet	0.9504	0.9945	0.9055
	CSANet	0.9566	0.9952	0.9168
	Ours	**0.9760**	**0.9966**	**0.9532**

SUMMARY

- **We propose a method to segment high-dimensional images**
- **In the process of segmentation, not only the characteristics of medical images are considered, but also temporal information is used to enhance the segmentation effect**
- **The correction module is used to improve the segmentation effect**
- **We tested on several datasets and proved that our model has good generalization**

Next step

- Different optical flow extraction methods
- Expanding to data sets in different fields (not limited to biomedicine)

Thanks for your attention!