MODEL-BASED LABEL-TO-IMAGE DIFFUSION FOR SEMI-SUPERVISED CHOROIDAL VESSEL SEGMENTATION

Kun Huang[1], Xiao Ma[1], Na Su[2], Songtao Yuan[2], Qiang Chen[1]

1. Nanjing University of Science and Technology, China

2. The First Affiliated Hospital with Nanjing Medical University, China

1 – Background

Current successful choroidal vessel segmentation methods rely on large amounts of voxel-level annotations, which are hard and time-consuming.

Semi-supervised learning solves this issue by enabling model learning from both unlabeled data and a limited amount of labeled data.

The suboptimal performance of current semi-supervised methods due to three main factors:

1) Incomplete pseudo labels only provides limited supervision.
2) Pseudo labels are limited by distribution of labeled data.
3) Increasing domain shift between labeled data and unlabeled data with the amount of labeled data decreasing.

2 – Method

Overview

We propose a model-based label-to-image diffusion (MLD) framework. Rather than keep the consistency between the pseudo labels and the original images, we generate the corresponding images of each pseudo label for unlabeled data.

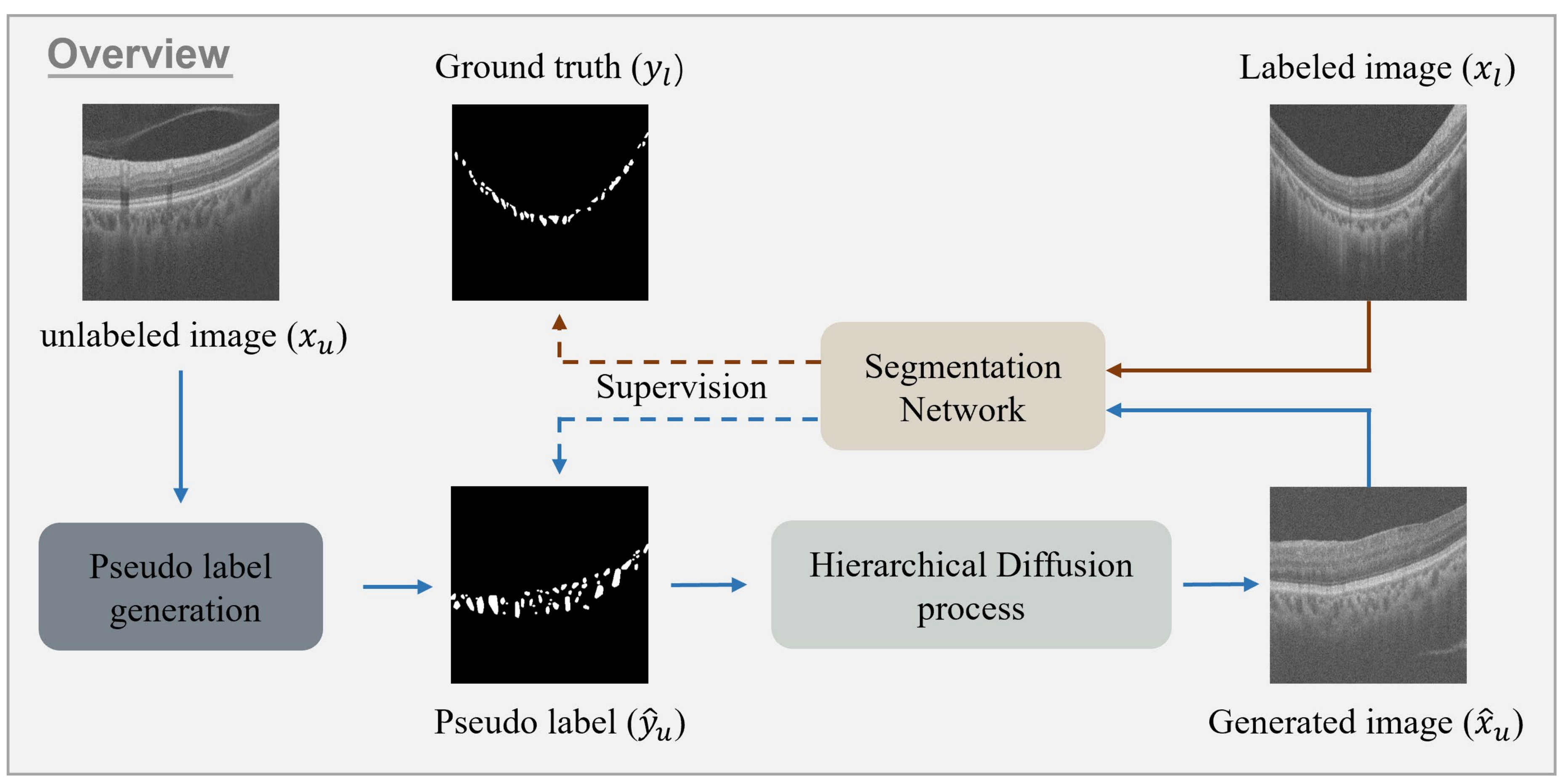

Fig. 1. Overview of the MLD framework

Pseudo Label Generation

Our framework does not require the consistency the original image and the pseudo label. Thus, we propose a general pseudo label generation strategy consisting of a series of unsupervised operations, which does not bias on the labeled data distribution.

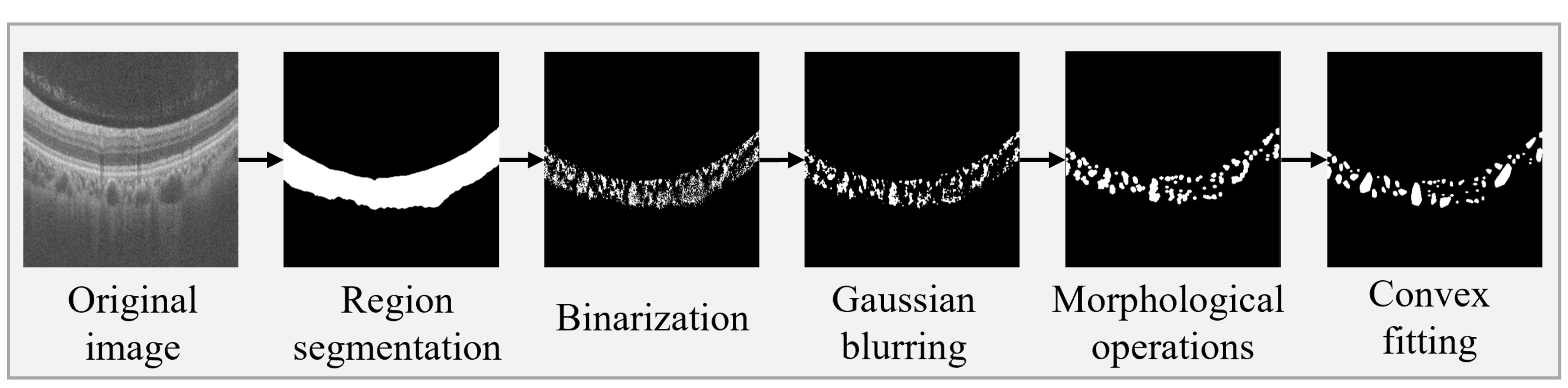

Fig. 2. Process of the pseudo label generation

Label-to-image Diffusion

We propose a hierarchical denoising diffusion probabilistic model that can learn from both labeled and unlabeled data with two denoising networks $\{f_u, f_l\}$.

- Forward process: For a given real image x_0, we add noises T steps as

$$x^t = \sqrt{1-\beta^t}x^{t-1} + \sqrt{\beta^t}z^t$$

- Training f_u: f_u takes the current noisy image x_u^t and a region mask y_r as input and estimates the added noise of the current step z^t:

$$L_u = \|z_u^t - f_u(x_u^t, y_r)\|_2$$

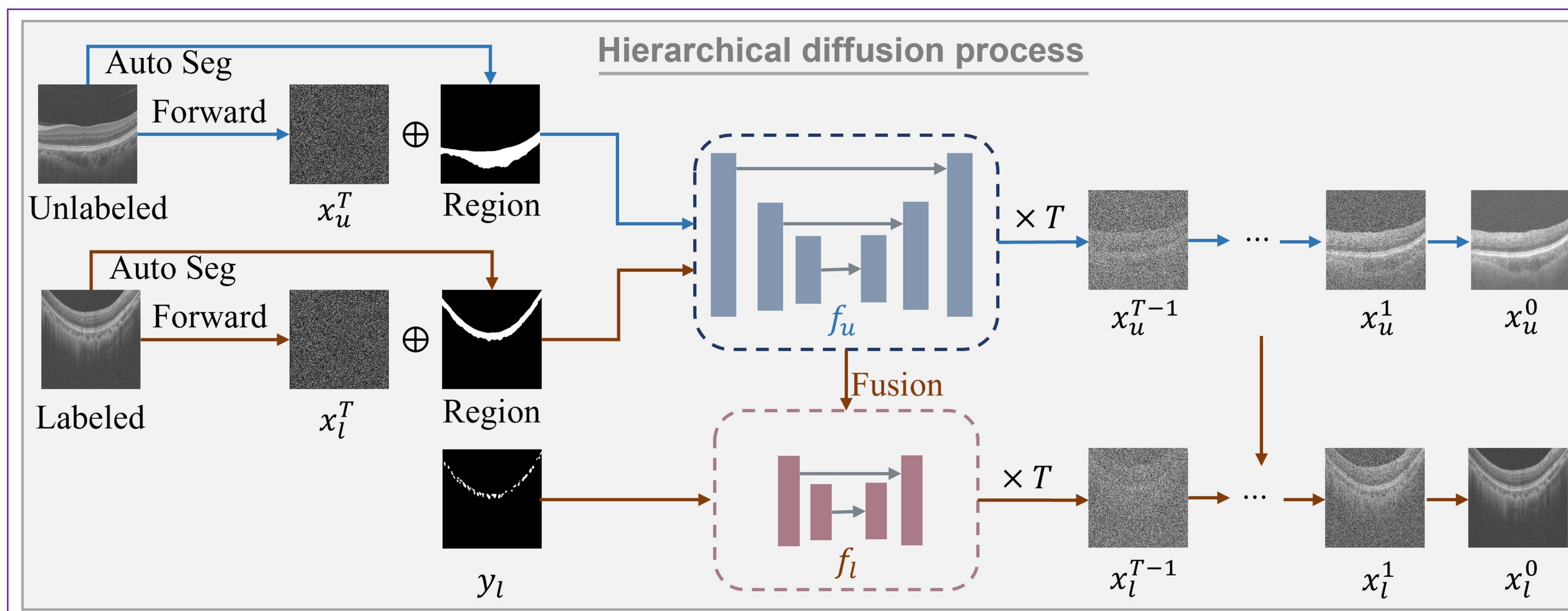

Fig. 3. Process of hierarchical diffusion probabilistic model

- Training f_l: f_l takes the last inferenced features of f_u and the label y_l as input:

$$L_l = \|z_l^t - f_l(f_u(x_l^t, y_r), y_l)\|_2$$

- Diffusion process:

$$x_l^{t-1} = \frac{1}{\sqrt{\alpha^t}}(x_l^t - \frac{1-\alpha^t}{\sqrt{1-\bar{\alpha}^t}} f_l(f_u(x_l^t, y_r), y_l) + \sigma^t z$$

3 – Topological connectivity metric

Pixel-based metrics are difficult to reflect the topological connectivity of choroidal vessels in the three-dimensional space. Thus, we propose a new metric named connectivity change ratio (CCR).

Wcalculate the change of the number of connected components brought by under-segmented regions as $|f_c(Y_P\&Y_l) - f_c(Y_l)|$ and over-segmented regions as $|f_c(Y_P|Y_l) - f_c(Y_l)|$. The CCR considers both the influence of under-segmented and over-segmented regions on topological connectivity as

$$CCR = \frac{|f_c(Y_p\&Y_l) - f_c(Y_l)|}{f_c(Y_l)} + \frac{|f_c(Y_p|Y_l) - f_c(Y_l)|}{f_c(Y_l)}$$

4 – Experimental Evaluation

Setting

1L+21U: 21 unlabeled volumes and 1 labeled volume

5L+21U: 21 unlabeled volumes and 5 labeled volumes

(400 images per volume)

Comparison with SSL methods

Method	1L+21U			5L+21U		
	Dice	95HD	CCR	Dice	95HD	CCR
SL	70.7 ± 7.8	19.4 ± 11.9	7.9 ± 2.8	79.9 ± 6.0	13.9 ± 17.1	3.7 ± 1.2
UA-MT [11]	70.2 ± 7.2	18.2 ± 12.3	9.6 ± 6.0	80.1 ± 5.7	9.5 ± 9.6	3.6 ± 1.2
DTC [6]	72.1 ± 7.4	19.3 ± 12.7	6.5 ± 2.7	80.9 ± 6.2	14.4 ± 19.6	3.7 ± 1.2
CPS [10]	75.6 ± 6.1	15.9 ± 15.5	6.8 ± 3.1	81.8 ± 5.7	12.8 ± 15.0	4.5 ± 1.6
R-Drop [5]	75.7 ± 5.5	17.8 ± 19.2	9.3 ± 4.5	81.6 ± 5.7	13.2 ± 16.5	3.9 ± 1.4
GCS [7]	70.5 ± 8.5	19.3 ± 10.9	8.2 ± 5.2	80.8 ± 5.9	8.4 ± 8.9	3.4 ± 1.2
MLD (Ours)	**78.2 ± 6.0**	**10.1 ± 9.5**	**5.6 ± 1.8**	**82.4 ± 5.6**	**7.1 ± 7.6**	**3.1 ± 1.2**

Downstream enhancement

Method		1L+21U			5L+21U		
		Dice	95HD	CCR	Dice	95HD	CCR
ChoroidNet [4]	w/o MLD	72.5	16.3	6.0	79.1	8.7	4.2
	w MLD	**77.6**	**10.1**	**5.5**	**80.8**	**8.0**	**3.7**
RefineNet [3]	w/o MLD	73.1	16.2	6.9	79.3	10.4	4.8
	w MLD	**77.8**	**9.8**	**5.8**	**81.7**	**7.7**	**4.2**

Generation quality