Partially-Supervised Learning for Vessel Segmentation in Ocular Images

Yanyu Xu[1], Xinxing Xu[1(✉)], Lei Jin[2], Shenghua Gao[2], Rick Siow Mong Goh[1], Daniel S. W. Ting[3,4], and Yong Liu[1]

[1] Institute of High Performance Computing, A*STAR, Singapore, Singapore
{xu_yanyu,xuxinx,gohsm,liuyong}@ihpc.a-star.edu.sg
[2] ShanghaiTech University, Pudong, China
{jinlei,gaoshh}@shanghaitech.edu.cn
[3] Singapore National Eye Centre, Singapore Eye Research Institute, Singapore, Singapore
daniel.ting.s.w@singhealth.com.sg
[4] Duke-NUS Medical School, Singapore, Singapore

Abstract. The vessel segmentation in ocular images is a fundamental and important step in the diagnosis of eye-related diseases. Existing vessel segmentation methods require a large-scale ocular images with pixel-level annotations. However, manually annotating masks is a laborious and tedious process. Compared with the traditional pipelines which either annotate the complete training set or several images in full, in this paper, we propose a novel supervision manner, named Partially-Supervised Learning (PSL), which only relies on partial annotations in the form of one patch from each of the few images. Targeting it, we propose an active learning framework with latent MixUp. The active learning strategy is employed to select the most informative patch for further annotation, while the latent MixUp is proposed to learn a proper visual representation of both the annotated and unannotated patches. The experimental results on two types of vessel segmentation datasets (Rose-1 (SVC) dataset for OCTA image, and DRIVE dataset for fundus image) validate the effectiveness of our model. With only 5% annotations on Rose-1 (SVC) and DRIVE dataset, our performance is comparable with the previous methods trained on the whole fully annotated dataset.

Keywords: Partially-supervised learning · Vessel segmentation

1 Introduction

The retinal blood vessel segmentation from color fundus images [20] or OCTA images [12] is a fundamental and important step in the diagnosis of eye-related

Electronic supplementary material The online version of this chapter (https://doi.org/10.1007/978-3-030-87193-2_26) contains supplementary material, which is available to authorized users.

M. de Bruijne et al. (Eds.): MICCAI 2021, LNCS 12901, pp. 271–281, 2021.
https://doi.org/10.1007/978-3-030-87193-2_26

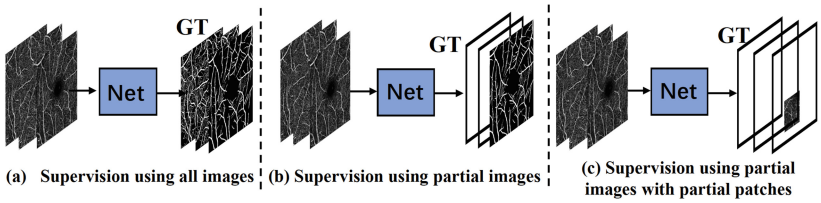


Fig. 1. An illustration of segmentation with different supervised way.

diseases, including systemic, metabolic, and hematologic diseases [13]. In the past decades, researchers have proposed many automatic segmentation methods, such as tracking [3] or filtering [2] based models, as well as thresholding based models [5,23]. Recently, deep learning based methods [8,12,16] have achieved promising results for the retinal blood vessel segmentation task. As shown in Fig. 1 (a), learning such segmentation models usually requires enough fully annotated pixel-wise masks, which is a laborious and tedious process. Besides, the annotations of retina blood vessels also require professional medical knowledge and proper training.

Naturally, a key question arises. Can we design a model that could still produce satisfactory performance but use as few annotations as possible? Along this direction, some researchers use a part of the dataset with fully annotated images in Fig. 1 (b). For example, some research works propose to use semi-supervised learning for cell segmentation [27], tumor segmentation [4], 3D abdominal CT [22], left atrium segmentation [10] or active learning for skin lesion segmentation [25]. Others propose an interactive annotation process for brain segmentation [21] or heart, aorta segmentation [9] or a full supervision directly on patches [6]. For such a labeling and training strategy, there are limited scenes and image conditions, which might reduce the generalization ability of the trained model. Based on the fact that the thick or thin vessels are usually consistent or the same in one image, it might be redundant to annotate all the vessels.

Therefore, different from the above attempts that fully annotate a few training images, we propose a novel supervision manner, named Partially-Supervised Learning, which only requires one patch from each of the few images. In particular, each annotated image consists of both one annotated patch, *e.g.* 10%, and the unannotated patches, *e.g.* 90% in Fig. 1 (c). In this way, the model could preserve diversity in the scenes and reduce the costs of manual annotation. Besides, only annotating one patch in each of the few images could also reduce the redundancy of annotating similar texture or repeated patterns. In [15], they propose a method using partial point annotation to segment the cell regions. However, their model is specific for the cell data, since they use the prior knowledge, such as the locations of the cell nucleus, while our setting is more general for several kinds of vessel images, such as fundus and OCTA.

There are two challenges of our partially-supervised learning setting. The first one is how to select the most useful patch to be annotated, since different patches include different level of information. The second one is how to leverage

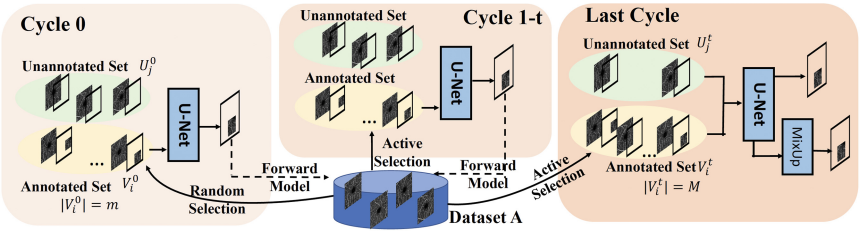


Fig. 2. The pipeline of our proposed model.

those unannotated patches and alleviate the need for labeled data, given the quite limited annotations. Targeting it, we propose an active learning framework with a latent MixUp for vessel segmentation in the partially-supervised learning setting. Given a labeling budget, we employ an active labeling strategy to iteratively annotate the most informative and uncertain patch in the unlabeled set and learn the segmentation model on them. Besides, all the rest unannotated patches in the annotated images and the rest unannotated images are used to train the network. Inspired by MixUp [1,24], we propose a latent MixUp to learn a proper visual representation of both the annotated and unannotated patches. Different from the combination between the inputs in image level in the existing MixUp, our latent MixUp combines the latent representations in patch level, such as the annotated patches and unannotated patches.

The contributions of this work are summarized as follows: To reduce the annotation cost and produce competitive performance, we study the segmentation task under a new partially-supervised learning setting, where we only annotate one patch from each of the few images. Targeting it, we propose an active learning framework with a latent MixUp. The active learning strategy is used to select the most uncertain patch based on the probability predicted by the segmentation model and the latent MixUp is designed to learn a proper visual representation of both the annotated and unannotated patches. The experimental results on two types of vessel segmentation datasets show the effectiveness of our proposed model. With only 5% annotated data, our proposed model could achieve comparable performance with those methods based on the fully annotated training set on the ROSE-1 (SVC) and DRIVE dataset.

2 Method

2.1 Partially-Supervised Learning

In this work, we propose a partially-supervised learning setting for the ocular image vessel segmentation. We only annotate one patch from each of the few images. All annotated and unannotated patches are used in the training process.

Firstly, we select M images from dataset A ($M < |A|$), using a random or other specific selection strategy, such as selecting the most uncertain one. For

each selected image $v_i \in \mathbb{R}^{3\times H\times W}$, we only annotate a patch about 10% of $H\times W$ area. The annotated patch is also selected by a random or other specific selection strategy, such as selecting the most uncertain one. Besides, the annotated patch is fully connected not disperse, since it is convenient to annotate one patch than several ones and several small annotated patches might disappear or be mixed with the neighboring unannotated patches after a successive convolutional layers.

Given the quite limited annotations, this partially-supervised learning setting has two challenges: (1) how to select the most useful patch to be annotated; (2) how to leverage those unannotated patches and alleviate the need for annotated data. Targeting at it, we propose an active learning framework with a latent MixUp, as illustrated in Fig. 2. The active learning framework is designed to iteratively select the most informative and uncertain patch to be annotated. Inspired by MixUp [1,24], the proposed latent MixUp is to learn a proper visual representation of both the annotated and unannotated patches. In our implementation, we use UNet [16] as backbone. The U-Net includes the contracting path (8 convolution layers following 4 max pooling layers) and expansive path (8 convolution layers and 4 up-sample layers). In the following, we will describe the proposed active learning framework with the latent MixUp in details.

2.2 Active Learning Framework

To select and annotate the most informative patch, we employ an active learning framework. Suppose there is a dataset A with labeling budget M, the maximum number of annotated images [11]. In initial cycle, we randomly select m images from the dataset A, randomly annotate one 10% patch in each selected images, then use them to train the segmentation model R^0. The annotated and unannotated set are represented as $V^0 = \{v_i, p_i\}$ and $U^0 = \{u_j\}$. After the training, we propose an active sample selection to choose m images with the most informative and uncertain patch from the unannotated set U^0, based on the predicted probabilities by model R^0. In the following cycle t, the updated annotated set V^t are used to train the segmentation model R^t. The process goes on until meeting the budget M. Only in the last cycle T ($T = \lceil M/m \rceil$), the unannotated set U is used to train the segmentation model R, using our latent MixUp.

Active Sample Selection. At the ending of cycle t, we want to select and annotate the most informative and uncertain patch and add them into the annotated set $V^t + 1$ to train the segmentation network R^{t+1}. These most uncertain patch usually is too complex or confused to make the right classification decision for the segmentation model with high confidence. The uncertainty of samples is directly indicated by the predicted probabilities of the learned segmentation model R. In particular, we forward the learned model R^t on the unannotated set U^t and get the probability map p_j for each images in U^t. In [19], Simple Margin learns an SVM on the existing labeled data and chooses as the next instance to query the instance that comes closest to the hyperplane. Similarly, a model is trained on labeled data and chooses the next most uncertain patches, which are closest to the decision boundary. In particular, it computes the distance between the

predicted probability and the median probability ($w = 0.5$). We calculate the smallest sum values through the following:

$$r_j = \arg\min_j \sum_l^L \sum_k^K (p_{j,lk} - w)^2, \tag{1}$$

where $[l : L, k : K]$ are the coordinate ranges of the unannotated candidate patches. When $p_{j,lk}$ equals w, the sum value is near 0, while $p_{j,lk}$ equals 0 or 1, the sum value is much larger. Thus, the smallest sum value represents this patch has the most uncertainty.

2.3 Latent MixUp

To learn a proper visual representation of the annotated and unannotated patches, we propose a latent MixUp, inspired by MixUp [1,24]. Mixup randomly combine the inputs and their corresponding labels to encourage model to behave linearly between training samples to reduce the oscillations during inference.

Different from the various existing MixUp versions in image level, our latent MixUp are used in patch level. Suppose we randomly select one partial annotated image I_1 and one unannotated image I_2. The proposed latent MixUp includes two kinds of combinations: 1) the combination between the partial annotated patch I_1^v in I_1 and the same coordinate unannotated patch I_2^v in I_2, as well as their corresponding ground truth label p_1^v and guessing label $\hat{p}_2^v$; 2) the combination between the rest unannotated patches I_1^u in I_1 and the same coordinate unannotated patches I_2^u in I_2, as well as their corresponding guessing label $\hat{p}_1^u$ and guessing label $\hat{p}_2^u$. For the guessing labels $\hat{p}$ of the unannotated patches, we use the probability predicted by the segmentation model R.

For a pair of one partial annotated image I_1 and one unannotated image I_2, we compute the new augmented latent representations from their last convolutional layer $\phi(I_1)$ and $\phi(I_2)$. The latent MixUp with a weight λ' is:

$$I^{v'} = \lambda'\phi(I_1^v) + (1-\lambda')\phi(I_2^v), p^{v'} = p_1^v \tag{2}$$

$$I^{u'} = \lambda'\phi(I_1^u) + (1-\lambda')\phi(I_2^u), p^{u'} = \lambda'\hat{p}_1^u + (1-\lambda')\hat{p}_2^u. \tag{3}$$

Following [1], λ' is generated by $\lambda \sim (\alpha, \alpha)$, $\lambda' = \max(\lambda, 1-\lambda)$, where α is a hyper-parameter. The latent MixUp enriches the distribution in-between training samples. To note that, the combination of existing MixUp occurs at inputs in image level, while ours occurs at latent representations in patch level.

2.4 Loss Function

For vessel segmentation task, we adopt the commonly used pixel-wise cross entropy loss $L_{seg} = -\sum_i p^v \log(R(I^v))$, where $R(I^v)$ is the segmentation mask prediction and p^v is the partial annotated ground truth of image i. Given an image pair $< (I^{v'}, p^{v'}), (I^{u'}, p^{u'}) >$, the latent MixUp loss function is computed

on their mixed representations. Following [1], the loss item for the annotated regions is a cross entropy loss $L_V = -\sum p^{v'} \log(R(I^{v'}))$, and the loss item for unannotated regions is MSE loss $L_U = \sum \|p^{u'} - R(I^{u'})\|_2^2$. In the last cycle, the total loss function is $L = L_{seg} + L_V + \lambda_U L_U$. λ_U is a hyper-parameter.

3 Experiment

3.1 Experiment Setting

Implementation Details. We use the PyTorch [14] platform to implement our model with the parameter settings: mini-batch size (4), learning rate (1.0e−4), momentum (0.95), weight decay (0.0005), and 100 epochs. We employ the default initialization to initialize the model. In the latent MixUp, the hyper-parameter α, λ_U are 0.75, 0.01. The labeling budget M is 10 and 15, and m equals 5.

Table 1. The performance comparison **Left** on the ROSE-1 (SVC) dataset and **Right** on the DRIVE dataset. FSL: Fully-Supervised Learning. SSL: Semi-Supervised Learning. PSL: Partially-Supervised Learning. LO: Label-Only. In our model, $m = 5$.

Method	Type	Ratio	ACC ↑	G-mean ↑	Kappa ↑	Dice ↑	FDR ↓	Method	Type	Ratio	ACC ↑	SPE ↑	SEN ↑	Dice ↑
U-Net [16]	FSL	100%	0.8955	0.8068	0.6476	0.7116	0.2627	UNet [16]	FSL	100%	0.9686	0.9861	0.7887	0.8140
CE-Net [8]	FSL	100%	0.9121	0.8256	0.6978	0.7511	0.1997	DEU-Net [20]	FSL	100%	0.9567	0.9816	0.7940	0.8270
OCTA-Net [12]	FSL	100%	0.9182	0.8361	0.7201	0.7697	0.1775	BEFD-UNet [26]	FSL	100%	0.9701	0.9845	0.8215	0.8267
MT [18]	SSL	3.3%	0.8827	0.7696	0.5951	0.6658	0.2896	MT [18]	SSL	5.0%	0.9519	0.9768	0.6403	0.6988
MixMatch [1]	SSL	3.3%	0.9079	0.8046	0.6750	0.7299	0.2313	MixMatch [1]	SSL	5.0%	0.9568	0.9741	0.6519	0.7245
LO (M = 10)	PSL	3.3%	0.9086	0.8263	0.6883	0.7425	0.2260	LO (M = 10)	PSL	5.0%	0.9605	0.9799	0.7584	0.7708
LO (M = 15)	PSL	5.0%	0.9102	0.8291	0.6935	0.7480	0.2228	LO (M = 15)	PSL	7.5%	0.9609	0.9794	0.7688	0.7752
Ours (M = 10)	PSL	3.3%	0.9119	0.8260	0.6948	0.7483	0.2008	Ours (M = 10)	PSL	5.0%	0.9618	0.9768	0.8020	0.7875
Ours (M = 15)	PSL	5.0%	0.9134	0.8273	0.6985	0.7510	0.1794	Ours (M = 15)	PSL	7.5%	0.9630	0.9788	0.8088	0.7919

Dataset. We use two types of eye images datasets: *The DRIVE dataset* [17] consists of 20 training and 20 testing fundus images centered on the macula. Each image includes a circular field-of-view mask; *The ROSE-1 dataset* [12] consists of 117 OCTA images. In this work, we use pixel-level annotation.

Metrics. Following the existing works [8,12,17], we use the following metrics: Dice Coefficient (Dice), Accuracy (ACC), Specificity (SPE), Sensitivity (SEN), Kappa score, False Discovery Rate (FDR) and G-mean score.

3.2 Performance Comparison

We evaluate our model with the following methods on the datasets on the commonly used metrics. Table 1 and Fig. 3 show the quantitative and qualitative results on the ROSE-1 (SVC) and DRIVE datasets.

Baselines. Since this is the first work to study partially-supervised learning setting ('PSL'), we compare our model with the following recent methods, divided into three groups. The first group is related to the fully-supervised learning methods, denoted as 'FSL'. We list some recent state-of-the-art methods using all samples as training, such as CE-Net [8], OCTA-Net [12] on the ROSE-1

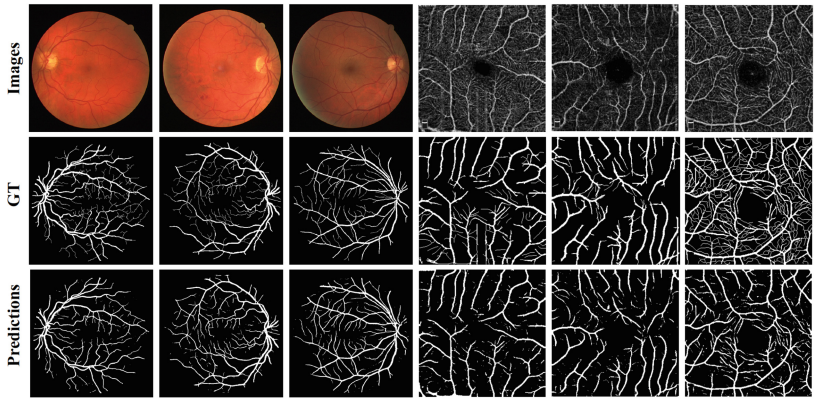


Fig. 3. Some examples of our model on the DRIVE and ROSE-1 (SVC) dataset.

dataset as well as DEU-Net [20] and BEFD-Net [26] on the DRIVE dataset. Secondly, we compare our model with the following related semi-supervised learning methods, denoted as 'SSL'. Mean teacher (MT) [18] is a classic consistency-based semi-supervised learning approach for classification. We modify it to segmentation task, by replacing the classifier by the segmentation mask generation and using the color augmentation in [7]. MixMatch [1] is a recently work for a semi-supervised learning, based on the mixup data augmentation [24] but performed on unlabeled data. The similar modification is applied for segmentation task. Finally, we also design a simple baseline, under partially-supervised learning 'PSL'. Label only ('Label-Only') uses the partial annotation (10% regions) with M images to train the U-Net model using masks on the loss function. To note that, we run five times the SSL and PSL baselines and our models and report the average values in the Table 1 and 2.

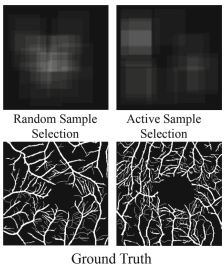
The 'Ratio' column represents how many annotated regions percentage the method uses as training samples. The PSL methods use 100% annotated data. The SSL methods use the 5% fully annotated images as training samples. For the PSL methods on the ROSE-1 (SVC) dataset, 5% means they use 15 (15/30 = 50%) partial annotated images, and each of them has 10% annotated region.

The ROSE-1 (SVC) Dataset. We conduct experiments on the OCTA images. The experimental results are shown on the Table 1 Left. With a small amount of annotated regions in the training process, our network can still converge well in this dataset. Under the same ratio of annotated regions, such as 5%, the PSL methods in the third block always perform better than the SSL methods in the second block on the all metrics, which indicates our proposed novel supervision could capture more variance and reduce the redundancy of the dataset. Even using only 5% annotated data, our method improves slightly better performance than fully-supervised method CE-Net [8] in all metrics. We also show some quantitative examples predicted by our model ($M = 15, m = 5$). For the thick vessels, our model could estimate well, while might fall to predict these

Table 2. The ablation studies on the ROSE-1 (SVC) dataset. **The Left**: The effect of Active Sample Selection. RS: Random Selection. AC: Active Selection. **The Middle**: The effect of the Latent MixUp (LM). **The Right**: The visualization of the distribution of the selected patches by random selection and active selection.

Method	Type	Dice ↑	FDR ↓
M=5, m=5	AC/RS	0.7335	0.2343
M=10, m=5	RS	0.7444	0.2318
M=10, m=5	AC	**0.7471**	**0.2145**
M=15, m=5	RS	0.7468	0.2266
M=15, m=5	AC	**0.7494**	**0.2172**
M=20, m=5	RS	0.7521	0.2156
M=20, m=5	AC	**0.7525**	**0.1962**

Method	LM	Dice ↑	FDR ↓
M=5, m=5, AC	×	0.7335	0.2343
M=5, m=5, AC	√	**0.7378**	**0.2189**
M=10, m=5, AC	×	0.7471	0.2145
M=10, m=5, AC	√	**0.7483**	**0.2008**
M=15, m=5, AC	×	0.7494	0.2172
M=15, m=5, AC	√	**0.7510**	**0.1794**
M=20, m=5, AC	×	0.7525	0.1962
M=20, m=5, AC	√	**0.7546**	**0.1943**



thin vessels, owning to the lack of enough annotated data. More quantitative comparison with different M could be found in supplementary materials.

The DRIVE Dataset. We also conduct experiments on the fundus images. Table 1 Right is the experimental results. With less 10% annotated data, our network still can perform well on this dataset. Similarly, compared with the SSL methods, the PSL baseline and our PSL model could always achieve higher performance on the all metrics under the same ratio of the annotated regions, such as 5%. Besides, compared with the fully-supervised learning method DEU-Net [20], our method improves comparable performance, even slightly better performance on two metrics. Figure 3 shows some prediction of our model ($M = 15, m = 5$). We can see that our model still could perform well on these thick vessels, while might fall on the very thin vessels. If using more annotated thin regions as training samples, the performance could be further improved.

3.3 Ablation Studies

We conduct ablation studies on the ROSE-1 (SVC) dataset.

The Effect of the Budget M. We use different budgets M to annotated the training samples, such as $M = 5, 10, 15, 20$. The results are shown on the Table 2. We can see that with the increasing M, the model could achieve the better performance. It is obvious that using more annotated regions as training samples, the segmentation model could perform much better. We also train the models using less or more annotations, such as 5%, 25% and 50% in the selected images, to investigate how the annotated ratios effect the performance. Similarly, we find that the bigger the annotated ratio is, the higher performance the model could achieve. More comparisons could be found in supplementary materials.

The Effect of Active Sample Selection. To validate the effect of the active sample selection (AC), we train the baselines, replacing it with random selection (RS). The results are shown in the Table 2 Left part. We can see that our proposed active sample selection always outperforms the random selection baselines

on the metrics. Besides, we also show the distribution of the annotated patches, selected by two strategies, in the Fig. 2 Right. The proposed active sample selection could avoid selecting the central regions, where there almost exist no vessels and it is much easier to predict, as shown in the Fig. 2 Right. Both of them could indicate that our proposed active sample selection could select the most informative and uncertain patch to be annotated. Besides, we also use other sample selection ways, such as using $p = 0.25$, 0.75, entropy using output probability $p \log(1-p)$, or Monte-Carlo Dropout. The results on the Dice metric are 0.750 (MC Dropout), 0.750 ($p \log(1-p)$), 0.747 ($p = 0.25$), 0.750 ($p = 0.75$) and 0.751 (Ours).

The Effect of the Latent MixUp. To evaluate the effect of the latent MixUp, we train the baselines, removing the proposed latent MixUp with different M. The experimental results are shown on the Table 2 Middle part. We can see that our proposed latent MixUp could always outperform the baselines on the metrics, which indicates that our proposed latent MixUp could make use of these large unannotated images to learn more useful visual features.

4 Conclusion

We propose a novel partially-supervised learning setting to estimate the vessel segmentation using as few annotations as possible and producing competitive performance. Compared with semi-supervised learning, it could bring the various challenging scenes using the same even less annotation costs. Besides, we also design an active learning framework with latent MixUp. Further, our model only uses 5% annotated data and the results also indicate there is a further improving space in the vessel segmentation under the partially-supervised learning setting.

Acknowledgements. This research is supported by A*STAR under its PROGRAMMATIC FUND (Grant No. A20H4g2141).

References

1. Berthelot, D., Carlini, N., Goodfellow, I., Papernot, N., Oliver, A., Raffel, C.A.: MixMatch: a holistic approach to semi-supervised learning. In: Advances in Neural Information Processing Systems, pp. 5049–5059 (2019)
2. Chaudhuri, S., Chatterjee, S., Katz, N., Nelson, M., Goldbaum, M.: Detection of blood vessels in retinal images using two-dimensional matched filters. IEEE Trans. Med. Imaging **8**(3), 263–269 (1989)
3. Chutatape, O., Zheng, L., Krishnan, S.M.: Retinal blood vessel detection and tracking by matched Gaussian and Kalman filters. In: Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 20 Biomedical Engineering Towards the Year 2000 and Beyond (Cat. No. 98CH36286), vol. 6, pp. 3144–3149. IEEE (1998)
4. Fang, K., Li, W.-J.: DMNet: difference minimization network for semi-supervised segmentation in medical images. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12261, pp. 532–541. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59710-8_52

5. Gao, S.S., et al.: Compensation for reflectance variation in vessel density quantification by optical coherence tomography angiography. Invest. Ophthalmol. Vis. Sci. **57**(10), 4485–4492 (2016)
6. Giarratano, Y., et al.: Automated segmentation of optical coherence tomography angiography images: benchmark data and clinically relevant metrics. Transl. Vis. Sci. Technol. **9**(13), 5 (2020)
7. Godard, C., Mac Aodha, O., Firman, M., Brostow, G.J.: Digging into self-supervised monocular depth prediction (October 2019)
8. Gu, Z., et al.: CE-Net: context encoder network for 2D medical image segmentation. IEEE Trans. Med. Imaging **38**(10), 2281–2292 (2019)
9. Khan, S., Shahin, A.H., Villafruela, J., Shen, J., Shao, L.: Extreme points derived confidence map as a cue for class-agnostic interactive segmentation using deep neural network. In: Shen, D., et al. (eds.) MICCAI 2019. LNCS, vol. 11765, pp. 66–73. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32245-8_8
10. Li, S., Zhang, C., He, X.: Shape-aware semi-supervised 3D semantic segmentation for medical images. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12261, pp. 552–561. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59710-8_54
11. Liu, X., Van De Weijer, J., Bagdanov, A.D.: Exploiting unlabeled data in CNNs by self-supervised learning to rank. IEEE Trans. Pattern Anal. Mach. Intell. **41**(8), 1862–1878 (2019)
12. Ma, Y., et al.: ROSE: a retinal OCT-angiography vessel segmentation dataset and new model. arXiv preprint arXiv:2007.05201 (2020)
13. Mou, L., et al.: CS-Net: channel and spatial attention network for curvilinear structure segmentation. In: Shen, D., et al. (eds.) MICCAI 2019. LNCS, vol. 11764, pp. 721–730. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32239-7_80
14. Paszke, A., et al.: PyTorch: an imperative style, high-performance deep learning library. In: Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E., Garnett, R. (eds.) Advances in Neural Information Processing Systems 32, pp. 8024–8035. Curran Associates, Inc. (2019)
15. Qu, H., et al.: Weakly supervised deep nuclei segmentation using partial points annotation in histopathology images. IEEE Trans. Med. Imaging **39**(11), 3655–3666 (2020)
16. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
17. Staal, J., Abràmoff, M.D., Niemeijer, M., Viergever, M.A., Van Ginneken, B.: Ridge-based vessel segmentation in color images of the retina. IEEE Trans. Med. Imaging **23**(4), 501–509 (2004)
18. Tarvainen, A., Valpola, H.: Mean teachers are better role models: weight-averaged consistency targets improve semi-supervised deep learning results. In: Advances in Neural Information Processing Systems, pp. 1195–1204 (2017)
19. Tong, S., Koller, D.: Support vector machine active learning with applications to text classification. J. Mach. Learn. Res. **2**, 45–66 (2001)
20. Wang, B., Qiu, S., He, H.: Dual encoding U-Net for retinal vessel segmentation. In: Shen, D., et al. (eds.) MICCAI 2019. LNCS, vol. 11764, pp. 84–92. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32239-7_10

21. Wang, G., Aertsen, M., Deprest, J., Ourselin, S., Vercauteren, T., Zhang, S.: Uncertainty-guided efficient interactive refinement of fetal brain segmentation from stacks of MRI slices. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12264, pp. 279–288. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59719-1_28
22. Wang, Y., et al.: Double-uncertainty weighted method for semi-supervised learning. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12261, pp. 542–551. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59710-8_53
23. Yousefi, S., Liu, T., Wang, R.K.: Segmentation and quantification of blood vessels for oct-based micro-angiograms using hybrid shape/intensity compounding. Microvasc. Res. **97**, 37–46 (2015)
24. Zhang, H., Cisse, M., Dauphin, Y.N., Lopez-Paz, D.: mixup: beyond empirical risk minimization. arXiv preprint arXiv:1710.09412 (2017)
25. Zhang, M., Dong, B., Li, Q.: Deep active contour network for medical image segmentation. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12264, pp. 321–331. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59719-1_32
26. Zhang, M., Yu, F., Zhao, J., Zhang, L., Li, Q.: BEFD: boundary enhancement and feature denoising for vessel segmentation. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12265, pp. 775–785. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59722-1_75
27. Zhou, Y., Chen, H., Lin, H., Heng, P.-A.: Deep semi-supervised knowledge distillation for overlapping cervical cell instance segmentation. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12261, pp. 521–531. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59710-8_51