

Retinal Vessel Segmentation with Skeletal Prior and Contrastive Loss

Yubo Tan, Kai-Fu Yang, Shi-Xuan Zhao, and Yong-Jie Li, *Senior Member, IEEE*

Abstract—The morphology of retinal vessels is closely associated with many kinds of ophthalmic diseases. Although huge progress in retinal vessel segmentation has been achieved with the advancement of deep learning, some challenging issues remain. For example, vessels can be disturbed or covered by other components presented in the retina (such as optic disc or lesions). Moreover, some thin vessels are also easily missed by current methods. In addition, existing fundus image datasets are generally tiny, due to the difficulty of vessel labeling. In this work, a new network called SkelCon is proposed to deal with these problems by introducing skeletal prior and contrastive loss. A skeleton fitting module is developed to preserve the morphology of the vessels and improve the completeness and continuity of thin vessels. A contrastive loss is employed to enhance the discrimination between vessels and background. In addition, a new data augmentation method is proposed to enrich the training samples and improve the robustness of the proposed model. Extensive validations were performed on several popular datasets (DRIVE, STARE, CHASE, and HRF), recently developed datasets (UoA-DR, IOSTAR, and RC-SLO), and some challenging clinical images (from RFMiD and JSIEC39 datasets). In addition, some specially designed metrics for vessel segmentation, including connectivity, overlapping area, consistency of vessel length, revised sensitivity, specificity, and accuracy were used for quantitative evaluation. The experimental results show that, the proposed model achieves state-of-the-art performance and significantly outperforms compared methods when extracting thin vessels in the regions of lesions or optic disc. Source code is available at <https://www.github.com/tyb311/SkelCon>.

Index Terms—fundus image, vessel segmentation, skeletal prior, data augmentation, contrastive loss

I. INTRODUCTION

Retinal vessel morphology is a key signal source for ophthalmic diagnosis, such as diabetic retinopathy, pigment epithelium, glaucoma, cardiovascular disease, hypertension, arteriosclerosis and age-related macula degeneration [6]–[8], or choroid correlation analysis [3], [4]. Structural changes in the retina may also indicate cardiovascular diseases [5]. Therefore, segmentation of retinal vessels is an important pre-processing step of fundus image analysis. However, manual segmentation of retinal vessels is very time-consuming and laborious, so

This work was supported by Key Area R&D Program of Guangdong Province (#2018B030338001); National Natural Science Foundation of China (#62076055). (Corresponding author: Kai-Fu Yang, email:yangkf@uestc.edu.cn).

Yubo Tan, Kai-Fu Yang, Shi-Xuan Zhao, and Yong-Jie Li are with the MOE Key Laboratory for Neuroinformation, University of Electronic Science and Technology of China, Chengdu 610054, China.

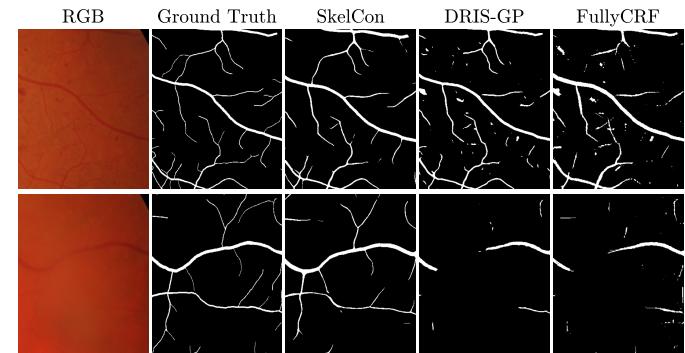


Fig. 1. Challenging cases of retinal vessel segmentation, including vessels in the region of lesions (top) and thin vessels (bottom). The proposed method (SkelCon) can reduce false alarms in lesion areas and improve the performance of thin vessel segmentation compared with the recent state-of-the-art methods, e.g., DRIS-GP [1] and FullyCRF [2].

highly precise segmentation algorithms are urgently required to facilitate diagnoses for both ophthalmologists and artificial intelligence systems.

Traditional vessel segmentation methods mainly include thresholding methods [9], tracking methods [10] [11], and filtering methods [4], [12], [13]. These methods are performed directly based on the pixel intensity or morphological features of vessels. Machine learning methods have been introduced to classify the extracted features of pixels [2], [14]–[17]. However, traditional classifiers generally require manually designed features, which rely heavily on domain knowledge. Recently, with the rise of deep learning, great progress has been made in retinal vessel segmentation [18], [19]. Deep networks can automatically learn features with numerous learnable parameters, and they can segment retinal vessels more accurately.

However, existing vessel segmentation methods still face several major problems. First, many diseases or lesions can seriously affect image quality, resulting in difficulties of retinal vessel segmentation. Other tissues in the retina, such as the optic disc, also interfere with vessel segmentation (shown in Fig. 1 (top row)). Second, thin vessels are easily missed by the current methods, leading to incomplete vessel segmentation (shown in Fig. 1 (bottom row)). In addition, deep learning methods strongly depend on large amounts of labeled data, but fundus datasets for vessel segmentation are usually small because of the difficulty of labeling [20]–[23]. This makes the segmentation of fine vessels and lesion areas more difficult.

In this work, a new deep learning network is proposed

to tackle the aforementioned problems. Accordingly, a contrastive loss function is built to distinguish vessels from lesions or other tissues. In addition, a skeletal prior is introduced to balance the segmentation of thick and thin vessels and improve their continuity. Finally, a new method is proposed to improve the effectiveness of data augmentation. Specifically, the main contributions of this paper can be summarized as follows:

- To improve the completeness and continuity of vessel segmentation, a vessel skeleton fitting module (SFM) and the corresponding loss functions are designed to guide the network to focus on thin vessel segmentation and improve the continuousness.
- To improve the discrimination of features between vessels and background pixels (especially lesions), a contrastive loss (CL) is introduced to reduce the intra-class distance and increase the inter-class variance. In addition, a vessel-background sampling strategy is proposed to build samples in the feature space.
- Additionally, a new data augmentation method, called color-space channel mixing (CSM), is proposed to expand the number of training samples and improve the generalization ability of the model.
- Experiments on nine classical datasets show that the proposed method achieves state-of-the-art performance compared with recent methods.

The paper is organized as follows. In Section II, we describe the traditional methods and deep learning methods designed for retinal vessel segmentation. Section III presents the proposed vessel segmentation method (i.e., SkelCon) in detail. The experimental results and the evaluations are presented in Section IV. Finally, the discussion and conclusion are respectively provided in Sections V and VI.

II. RELATED WORK

Much research has been conducted on retinal vessel segmentation, involving both traditional methods and deep learning methods. This section will introduce the existing research as well as some specific technologies related to the proposed method.

A. Traditional Methods for Vessel Segmentation

Traditional methods are performed directly based on the pixel intensity or morphological features of the vessels, such as the adaptive local threshold method [9], matched filtering algorithm [12], pattern detection algorithm [13], and multi-scale line detectors [4]. Sheng *et al.* used an efficient minimum generated superpixel tree combined with the geometric structure, texture, color, and spatial information to detect vessel structure [24]. Lam *et al.* [25] proposed a differentiable concave measure to deal with bright spots and dark lesions.

To extract the vessel tree, Mendonca *et al.* extracted the vessel centerline, selected the connected set of candidate points, and finally segmented vessels with an iterative region growth method [10]. Roychowdhury *et al.* proposed an iterative vessel segmentation algorithm [11] in which a global threshold is used to extract the initial value and a new stop criterion is used to terminate the iterative process. In general, these methods

can avoid the false segmentation of lesions to a certain extent, but with poor accuracy. Moreover, traditional methods usually depend heavily on hand-crafted filters, which are not flexible enough.

Machine learning typically requires the input of hand-crafted features and classifies these features using appropriate classifiers. Ricci *et al.* used the gray level of pixels to construct feature vectors and classified them with a support vector machine (SVM) classifier [14]. Similarly, Marin *et al.* calculated seven-dimensional moment-invariant features to represent pixels [15]. Annunziata *et al.* classified the curve structure with a random forest classifier [17]. Soares *et al.* used a multi-scale two-dimensional (2D) Gabor wavelet transform to construct feature vectors and then applied a Bayesian classifier to classify the pixels [16]. Orlando *et al.* combined conditional random field models and SVM for vessel segmentation [2]. Generally, machine learning approaches require manually designed features, which usually rely on domain knowledge.

The accuracy of traditional segmentation methods needs improvement, especially the anti-interference ability when facing lesions and shadows. Therefore, the emerging methods in recent years have mainly been deep learning methods.

B. Deep Learning Methods for Vessel Segmentation

Data-driven deep learning methods have exhibited higher performance than traditional methods in the field of retinal vessel segmentation. For example, Fu *et al.* took retinal vessel segmentation as a boundary detection task [26]. They used a multi-scale and multi-level network with a lateral output layer to learn rich hierarchical representation and model the remote interactions between pixels with conditional random fields. In the method developed by Maninis *et al.*, feature maps from a side output layer assisted vessel and optic disc segmentation [27]. By combining the multi-scale analysis provided by stationary wavelet transform with the multi-scale full convolutional neural network, Oliveira *et al.* produced a method that can deal well with changes in the width and direction of retinal vessel structure [28]. Gu *et al.* proposed CE-NET to obtain more advanced information and reserve spatial information for 2D medical image segmentation [29].

To improve the connectivity of retinal vessel segmentation, Mou *et al.* used the dense dilated network for initial retinal vessel detection and adopted the probability regularization walking algorithm for the fracture issue [30]. Although existing deep learning methods are very good at the segmentation of thick vessels, it is still challenging for them to obtain more complete vessels and avoid mis-segmentation of lesions.

Generally, several aspects can be considered when improving the deep-learning-based methods for vessel segmentation, including data augmentation, model designing, training strategies, and loss functions. Some existing solutions and improvements for the above challenges of deep vessel segmentation are as follows.

1) Data Augmentation: Recently, many studies have used generative methods to augment data. For example, Son *et al.* believed that existing detection methods were prone to

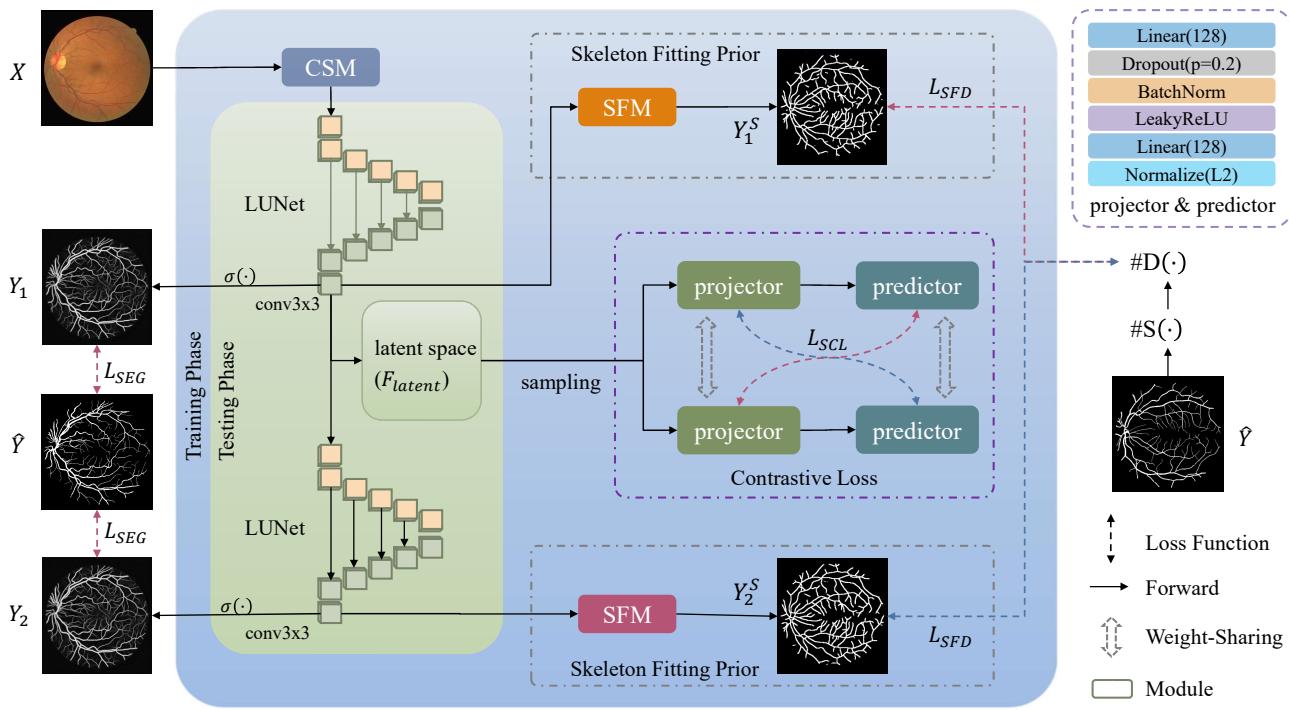


Fig. 2. The network architecture of the SkelCon. \mathbf{X} is the input fundus image, $\mathbf{Y}_1, \mathbf{Y}_2$ are outputs of two cascaded little UNets, $\mathbf{Y}_1^S, \mathbf{Y}_2^S$ are skeleton fitting vessels, and $\hat{\mathbf{Y}}$ is ground truth. In addition, $\#S(\cdot)$ and $\#D(\cdot)$ are morphological skeleton and dilation operators respectively, $\sigma(\cdot)$ indicates the sigmoid function. CSM: Color Space Mixing, LUNet: Little UNet, and SFM: Skeleton Fitting Module.

missing small vessels or generating false positives in terminal branches, so they proposed a generative adversarial network (GAN) for generating accurate retinal angiography [31]. Zhao *et al.* proposed R-sGAN technology to construct a synthetic fundus image dataset [32]. GANs can generate a large number of samples, but it is difficult for them to ensure the data distribution consistency between the new data and the real fundus images. For example, the changes in the depth of vessels and lesions cannot be well simulated. In contrast, to generate more authentic and diverse samples, the proposed method presented a novel data augmentation strategy that generates more realistic and diverse samples by randomly mixing color components from different color spaces.

2) Unbalanced Vessels: Due to the imbalance of pixel numbers between thick and thin vessels in fundus images, pixel-wise loss functions limit the accuracy of thin vessel segmentation by deep learning models. Yan *et al.* combined a segment loss function and a pixel loss function to better balance both thin and thick vessels [33]. Further, Yan *et al.* proposed a deep learning model divided into three stages: thick vessel segmentation, thin vessel segmentation, and vessel fusion [34]. This model minimized the negative effects of the imbalance between thick and thin retinal vessels. In addition, based on vascular geometric characteristics, Cherukuri *et al.* designed a representation layer with strong extraction ability and a residual network to segment vessels [1]. The improvement based on weighted loss function or prior knowledge based on vessel geometry can enhance the performance of vessel segmentation. Therefore, a comprehensive balance strategy between thick and thin vessels through the skeleton of the vascular tree is expected to further improve the accuracy of thin vessel

segmentation. Previous works usually balance the thin and thick vessels with weighted loss functions or divide them into different sub-tasks. In this work, we build a new skeleton fitting module and loss functions are correspondingly designed to guide the network to focus on thin vessel segmentation and improve the continuousness.

3) Contrastive Learning: Contrastive learning can be used to reduce the intra-class distance between samples and increase the inter-class distance. The general paradigm is to map the sample features to a hidden space (usually a hypersphere) and then construct positive and negative sample pairs. At present, contrastive learning is widely used for self-supervised representation learning [35]. This approach can achieve remarkable results with a small number of annotations. Grill *et al.* indicated that negative sample pairs may not be necessary, as their BYOL network directly used a predictor to map the hypersphere vectors to new views. Then a loss was calculated between the feature outputs of a pair of encoders, and the stop gradient technique was applied to avoid the collapse solution [36]. Chen *et al.* pointed out that momentum updating encoders or big batch sizes are not required [37]. In addition, contrastive learning is complementary to supervised learning [38]. Recently, contrastive learning has been shown to perform well in natural images and has been involved in the field of medical images [39], [40]. Instead of directly employing contrastive learning scheme, this work constructs a contrastive loss to better deal with retinal vessel detection in lesion areas. As far as we know, our work presented in this paper is the first attempt to improve vessel segmentation in lesion areas by introducing contrastive loss.

III. METHODOLOGY

The flowchart of the proposed SkelCon method is shown in Fig. 2. The backbone model consists of two cascaded little UNets (LUNets). In the training phase, a data augmentation module (color space mixing, CSM) is first used to expand training samples and improve the generalization of the proposed network. In addition, extra loss functions are provided via the skeleton fitting module (SFM) to enhance the thin vessel segmentation. Finally, the features of the latent space between two LUNets are fed into the contrastive learning inspired module, in which a contrastive loss is designed to further help the model distinguish vessel and background pixels in the feature space. In the testing phase, only the backbone network consisting of two cascaded LUNets is used to predict vessels and achieve lightweight vessel segmentation.

A. Data Augmentation with Color Space Mixing

To train a robust neural network, data augmentation is a widely used step to generate more differentiated training samples. It is particularly important to retinal vessel segmentation considering that existing labeled datasets usually only contain dozens of images. Popular data augmentation methods (e.g., scaling, flipping, and rotation) in the field of natural image recognition are not suitable for such small datasets, so augmentation methods specifically for retinal vessel segmentation are needed [41]. For example, Sun *et al.* proposed two new data augmentation modules for retinal image augmentation and improved the performance of vessel segmentation [41]. However, their method requires additional computational overhead. Therefore, a new data augmentation module for retinal vessel segmentation is proposed in this work. The basic idea of the proposed module is randomly mixing color components from different color spaces since different color spaces usually highlight different details.

Specifically, the original RGB image is first transformed into two different color spaces, which are randomly selected from popular color spaces, including RGB, YUV, YCrCb, LAB, XYZ, HLS, and LUV. Next, augmented samples are generated by combining two random color components from the two transformed color spaces. Let X denote the original fundus image, and let T_1 and T_2 denote any two color space transformations. The generated sample can be obtained with

$$I_{csm} = \alpha_{mix} \cdot T_1(X)^m + (1 - \alpha_{mix}) \cdot T_2(X)^n, \quad (1)$$

where m and n indicate the index of the component in the corresponding color space and α_{mix} is a weighting factor. They are generated randomly with $m, n \in \{1, 2, 3\}$ and $\alpha_{mix} \in [0, 1]$. Additional operations, such as adding random noises and flipping, are applied to I_{csm} .

B. The Cascaded Network

The basic UNet [42] adds a multi-scale skip-connection between the encoder and symmetric decoder, which can better fuse deep features and shallow details. Inspired by the lightweight W-NET proposed by Galdran *et al.* [43], a lightweight cascade model is built based on the UNet-like

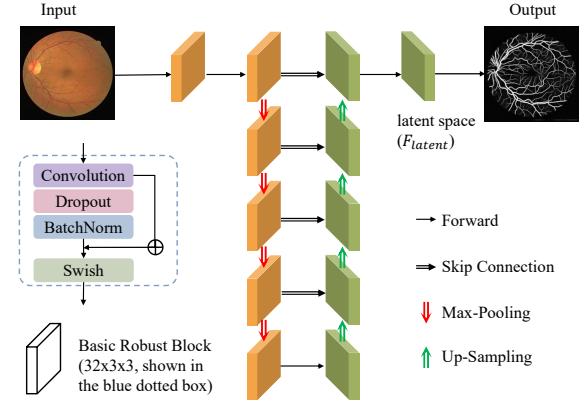


Fig. 3. The architecture of the little UNet (LUNet).

networks. First, a basic robust block (denoted as BRB) is designed, which includes the convolution, dropout, batchnorm, and swish activation, and is followed by a residual connection [44]. The original UNet's blocks are replaced with five layers of BRB, each containing 32 channels, to form a new little UNet (i.e., LUNet), as shown in Fig. 3. Finally, two UNet-like networks are cascaded to build the proposed model. Let Y_1 and Y_2 denote the outputs of the two LUNets, each of which is followed by a Sigmoid layer.

To train the proposed network, both LUNets receive the supervision of the ground truth with the focal loss, proposed by Lin *et al.* [45].

$$\mathcal{L}_{\text{FL}}(Y) = - \sum_i^N (1 - p_i)^\gamma \log(p_i), p_i \in Y. \quad (2)$$

Finally, the combined loss can be expressed as

$$\mathcal{L}_{\text{SEG}} = \mathcal{L}_{\text{FL}}(Y_2) + \alpha_{ds1} \cdot \mathcal{L}_{\text{FL}}(Y_1), \quad (3)$$

where α_{ds1} is a weighting coefficient, which is experimentally set to 0.9 in this work.

C. Skeleton Fitting Prior

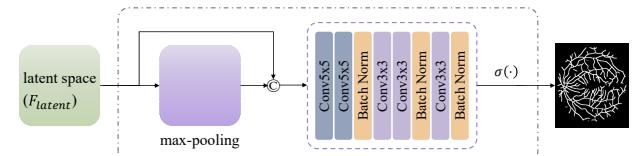


Fig. 4. The architecture of the skeleton fitting module. The © symbol denotes the concatenation operator.

Thin vessels occupy about 23% of the vessels in the fundus images [46], but they are likely to be missed by vessel segmentation methods based on pixel-matching metrics. Such methods have a bias towards thick vessels because thick vessels occupy more pixels. Therefore, a new module is proposed herein to balance the thick and thin vessel pixels with skeleton fitting and improve the performance of thin vessel segmentation.

First, the skeletons of the vessels are obtained and dilated to the same width, so the new ground truth of vessels can be obtained with

$$\hat{Y}^s = \#D(\#S(\hat{Y})), \quad (4)$$

where $\#D(\cdot)$ denotes the morphological dilation operator and $\#S(\cdot)$ denotes the morphological skeleton operator. The size of the template in the dilation operator is three pixels, which is used to expand vessels to the approximately maximum width of vessels.

Additionally, a new skeleton fitting module (shown in Fig. 4) is proposed to fit the vascular skeleton from the outputs of LUNets. A max-pooling operator is used to achieve the similar effects of morphological dilation, and then the output is concatenated to input features. The features are fed into a small sub-network (in Fig. 4). Finally, the output of the skeleton fitting module (Y^s) is obtained with a Sigmoid layer, and the Dice loss function is employed to constrain the distance between Y^s (including Y_1^s and Y_2^s from the two LUNets in the backbone network) and the dilated ground truth (\hat{Y}^s). i.e.,

$$\mathcal{L}_{SFD} = 2 - \frac{2|Y_2^s \cap \hat{Y}^s|}{|Y_2^s| + |\hat{Y}^s|} - \lambda_{ds2} \cdot \frac{2|Y_1^s \cap \hat{Y}^s|}{|Y_1^s| + |\hat{Y}^s|}, \quad (5)$$

where the weighting coefficient λ_{ds2} is set to 0.5.

D. Sample-Wise Contrastive Loss

One of the challenges in vessel segmentation is the disturbance from lesions or other tissues like the optic disc. To overcome this challenge, the neural network needs to improve the feature representation of vessels around lesions and other tissues. In this work, a contrastive loss is employed to improve the discrimination between vessels and the surrounding background. Inspired by contrastive learning methods, the goal of using contrastive loss is to enhance the similarity of the same category pixels in the feature spaces extracted from the neural network. The contrastive loss in this paper is built by sampling the features based on the labels by doctors (ground truth).

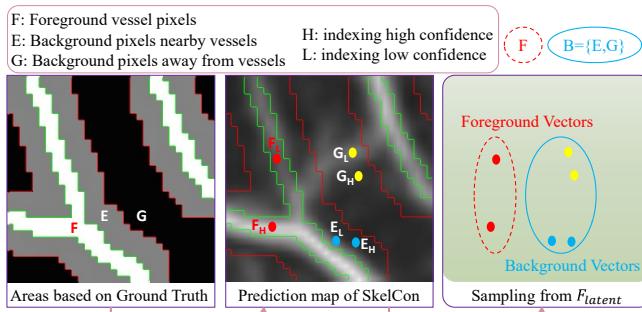


Fig. 5. The sampling of three categories of pixels, including vessel pixels, background pixels nearby, and background pixels further away from vessels.

Fig. 5 illustrates the supervised sampling process. According to the ground truth, the pixels in the retina are divided into three categories, including vessel foreground (denoted by F), background pixels next to vessels (denoted by E), and the other background pixels (denoted by G). For each category, the pixels are further divided into two classes, i.e., high confidence

pixels (indexed by H) and low confidence pixels (indexed by L), according to the final predicted probability of the cascaded LUNets. Thus, we obtain the vessel feature set from the latent feature space F_{latent} , i.e., $\{F_H, F_L\}$, which contains 1024 samples. Meanwhile, 512 background samples nearby vessels and 512 other background samples are obtained from the latent feature space to form the background feature set, i.e., $\{B_H, B_L\}$, where $B_H = \{E_H, G_H\}$ and $B_L = \{E_L, G_L\}$.

Subsequently, F_H , F_L , B_H , and B_L are fed into the projector (P_j) and predictor (P_d), both of which are simple multi-layer perceptrons including nonlinear layers and an $L2-norm$ layer (shown in Fig. 2). In addition, a cosine similarity loss is used between high and low confidence features in the same feature set. In detail, the cosine loss is defined as

$$\mathcal{D}(\vec{p}, \vec{q}) = -\frac{\vec{p}}{\|\vec{p}\|_2} \cdot \frac{\vec{q}}{\|\vec{q}\|_2}, \quad (6)$$

where $\|\cdot\|_2$ is the $L2-norm$ operator. Meanwhile, the contrastive loss is computed between the result of the predictor and the projector with high and low confidence samples.

$$\begin{aligned} \mathcal{L}_{SCL} = \frac{1}{2} \sum_{S \in \{F, B\}} & \left\{ \mathcal{D}(\Gamma\{P_j(S_H)\}, P_d(P_j(S_L))) \right. \\ & \left. + \mathcal{D}(\Gamma\{P_j(S_L)\}, P_d(P_j(S_H))) \right\}, \end{aligned} \quad (7)$$

where Γ denotes the stop gradient operator, which means that all tensors before this node do not compute gradients at the backward phase. This operator is suggested in [37] to avoid getting a collapse solution.

Finally, the overall loss function is

$$\mathcal{L}_{ALL} = \lambda_{seg} \cdot \mathcal{L}_{SEG} + \lambda_{sfd} \cdot \mathcal{L}_{SFD} + \lambda_{scl} \cdot \mathcal{L}_{SCL}, \quad (8)$$

where λ_{seg} , λ_{sfd} , and λ_{scl} are the weighting coefficients of the vessel segmentation loss function, vessel fitting loss function, and supervised contrastive loss function, respectively. Based on experiments, these coefficients are set as $\lambda_{seg} = 1$, $\lambda_{sfd} = 0.9$, and $\lambda_{scl} = 0.1$ in this work.

IV. EXPERIMENTS

A. Fundus Datasets

As previously mentioned, existing labeled datasets for retinal vessel segmentation are excessively small. This work employs four widely-used datasets, including DRIVE [21], STARE [20], CHASEDB1 [22], and HRF [23], as well as other recently developed datasets including RFMiD [49], JSIIC39 [50], UoA-DR [48], IOSTAR and RC-SLO [47].

DRIVE Dataset¹: The DRIVE dataset contains 40 pairs of fundus images and corresponding labels of vessel segmentation. The size of each image in the dataset is 565×584. Moreover, the images have been divided into training and test sets, each containing 20 pairs of images and labels. In particular, each image in the test set has been labeled by two doctors. Usually, the first label is regarded as the ground truth,

¹<http://www.isi.uu.nl/Research/Databases/DRIVE/>

and the second label is regarded as a human observation to evaluate the accuracy.

STARE Dataset²: The size of each image in the STARE dataset is 700×605 . There are 20 color fundus images in the STARE dataset and no allocation of training and test sets. To evaluate the performance of methods, there are two common schemes for test set allocation in previous works. The first is to put 10 images in the training set and the other 10 images in the test set. The second is the method of Leave-One-Out.

CHASEDB1 Dataset³: The CHASEDB1 dataset contains 28 images from both left and right eyes, and the size of each image is 999×960 . According to previous studies, the last 8 images are selected as testing samples and others as training samples.

HRF Dataset⁴: There are 45 fundus images in HRF dataset with a resolution of 3504×2336 . Fifteen images from each group of healthy children, diabetic retinopathy, and glaucoma patients are taken as the training set, and the other 30 images are taken as the test set. Images and labels are downsampled twice to reduce the computational cost [1], [2].

The DRIVE and HRF datasets provide the fields of view (FoVs) for all images, which can be used directly. However, the FoVs are not available in the STARE and CHASEDB1 datasets. Therefore, in this work, the FoVs generated by Marin *et al.* [15] were employed, and the pixels inside the FoV mask were used to calculate related metrics.

Other Vessel Segmentation Datasets: The IOSTAR dataset includes 30 images with a resolution of 1024×1024 and the RC-SLO dataset contains 40 image patches with a resolution of 360×320 [47]. There are 200 optic disk centered images, with a resolution of 2124×2056 in UoA-DR [48]. All the vessels in these datasets are annotated.

Challenging Clinical Datasets: RFMiD [49] consists of 3200 fundus images captured using three different fundus cameras. It constitutes a wide variety of diseases. JSIEC39 dataset [50] includes multiple common fundus diseases of 39 classes. There are no vascular labels in the both datasets.

B. Evaluation Metrics

To quantitatively evaluate the performance of segmentation methods, some metrics have been specially designed for vessel segmentation and widely used in previous works. For example, a set of metrics was proposed by Gegundez *et al.* [56] to evaluate the connectivity (C), overlapping area (A), and consistency of vessel length (L) of predicted vessels.

The overall metric (F) was defined as

$$F(C, A, L) = C \times A \times L. \quad (9)$$

Recently, another group of metrics has been used to evaluate the skeletal similarity of vessels. Yan *et al.* proposed curve similarity (cs_i) and thickness similarity (ts_i) for each skeleton

²<http://www.ces.clemson.edu/ahoover/stare/probing/index.html>

³<http://blogs.kingston.ac.uk/retinal/chasedb1/>

⁴<http://www5.informatik.uni-erlangen.de/research/data/fundus-images>

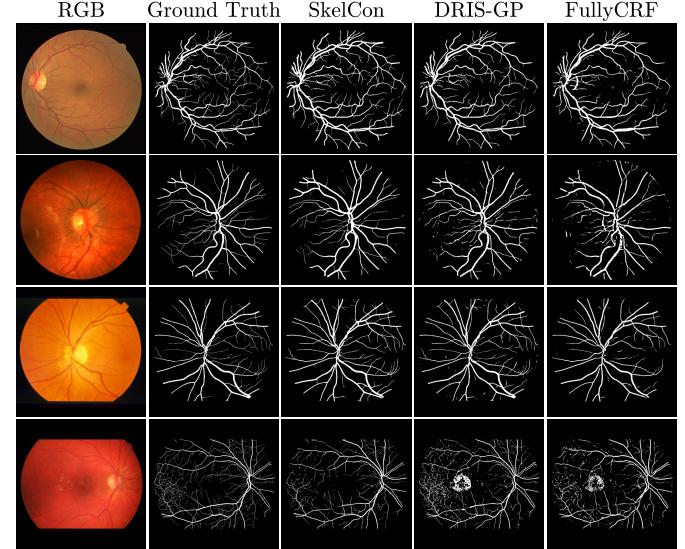


Fig. 6. Comparison of visual results on four images from the DRIVE, CHASEDB1, STARE, and HRF datasets (from top to bottom rows).

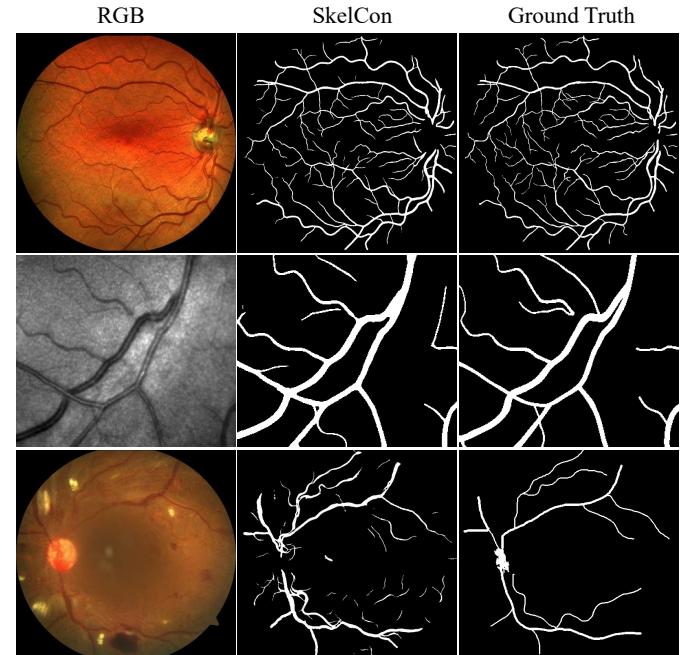


Fig. 7. Visual results on images from IOSTAR, RC-SLO and UoA-DR datasets.

segment. The final skeletal similarity of the i^{th} segment was defined as

$$ss_i = (1 - \alpha) \cdot cs_i + \alpha \cdot ts_i, \quad (10)$$

where the parameter α was set to 0.2. The overall skeletal similarity was defined as SS , which is a weighted combination of cs_i and ts_i of all segments of the total vessel skeleton.

Based on SS , Yan *et al.* redefined true positive (TP), false negative (FN), true negative (TN), and false positive (FP) based on skeleton instead of pixels. The revised metrics for evaluating the performance of vessel segmentation include Sensitivity (rSe), Specificity (rSp), and Accuracy (rAcc). More details about the computation of new metrics can be found

TABLE 1

COMPARISON WITH STATE-OF-THE-ART METHODS ON DRIVE. **RED**, **GREEN**, AND **BLUE** ARE THE TOP THREE (UNIT: %). THE P-VALUE OF THE PAIRED T-TEST BETWEEN EACH METHOD WITH THE PROPOSED SKELCON IS IN PARENTHESES.

Method	F	C	A	L	rSe	rSp	rAcc	Mcc
traditional methods								
Wavelets [16]	43.76(<0.001)	99.77(<0.001)	71.22(<0.001)	60.85(<0.001)	42.27(<0.001)	99.53(<0.001)	85.05(<0.001)	64.40(<0.001)
LineDet [14]	04.14(<0.001)	99.83(=0.584)	17.53(<0.001)	15.63(<0.001)	07.43(<0.001)	99.99 (<0.001)	76.55(<0.001)	23.68(<0.001)
Gray-Level [15]	60.52(<0.001)	99.90 (=0.002)	83.27(<0.001)	72.54(<0.001)	57.57(<0.001)	99.57(<0.001)	88.87(<0.001)	73.63(<0.001)
MSLineDet [4]	69.87(<0.001)	98.73(<0.001)	87.03(<0.001)	81.21(<0.001)	74.57(<0.001)	98.22(=0.041)	92.11(<0.001)	72.38(<0.001)
KernelBoost [51]	71.14(<0.001)	98.45(<0.001)	87.97(<0.001)	81.89(<0.001)	73.22(<0.001)	97.51(=0.007)	91.28(<0.001)	70.67(<0.001)
SE [52]	41.95(<0.001)	99.93 (<0.001)	71.95(<0.001)	57.78(<0.001)	32.09(<0.001)	98.06(=0.002)	81.30(<0.001)	61.02(<0.001)
N4-Fields [53]	72.24(<0.001)	99.77(<0.001)	89.20(<0.001)	80.90(<0.001)	64.85(<0.001)	99.69 (<0.001)	90.75(<0.001)	76.83(<0.001)
CRFs [3]	72.83(<0.001)	99.56(<0.001)	88.97(<0.001)	82.11(<0.001)	69.60(<0.001)	98.77(=0.156)	91.29(<0.001)	74.82(<0.001)
FullyCRF [2]	73.11(<0.001)	99.49(<0.001)	89.57(<0.001)	81.93(<0.001)	70.72(<0.001)	98.86(=0.046)	91.60(<0.001)	75.56(<0.001)
deep learning methods								
HED [54]	80.09(<0.001)	99.75(<0.001)	90.06(<0.001)	89.11(<0.001)	71.57(<0.001)	95.11(<0.001)	89.08(<0.001)	66.00(<0.001)
DRIU [27]	80.43(<0.001)	99.56(<0.001)	91.52(<0.001)	88.23(<0.001)	82.36 (=0.181)	96.85(<0.001)	93.13(<0.001)	71.61(<0.001)
DeepVessel [26]	61.74(<0.001)	99.60(<0.001)	84.23(<0.001)	73.38(<0.001)	54.93(<0.001)	99.78 (<0.001)	88.32(<0.001)	73.34(<0.001)
V-GAN [31]	84.82 (=0.287)	99.64(<0.001)	94.69 (<0.001)	89.84 (<0.001)	80.77(<0.001)	99.63(<0.001)	94.76 (=0.319)	80.24(=0.760)
JL-UNet [33]	81.06(<0.001)	99.61(<0.001)	93.08(=0.215)	87.35(<0.001)	76.11(<0.001)	99.57(<0.001)	93.53(<0.001)	78.98(<0.001)
SWT-FCN [28]	83.92(=0.034)	99.73(<0.001)	94.36(=0.007)	89.11(<0.001)	79.63(<0.001)	99.64(<0.001)	94.48(=0.387)	80.53 (=0.236)
DeepDyn [55]	84.53(=0.001)	99.70(<0.001)	94.58 (<0.001)	89.61(<0.001)	81.52 (<0.001)	99.44(<0.001)	94.82 (=0.134)	80.02(=0.062)
DRIS-GP [1]	84.94 (=0.503)	99.68(<0.001)	94.91 (<0.001)	89.74 (<0.001)	80.22(<0.001)	99.64(<0.001)	94.66 (=0.771)	81.84 (<0.001)
DAP [41]	82.55(<0.001)	99.72(<0.001)	93.74(=0.541)	88.24(<0.001)	78.57(<0.001)	99.57(<0.001)	94.15(=0.004)	79.00(<0.001)
SkelCon(Ours)	85.30	99.85	93.58	91.26	83.23	98.59	94.61	80.30

TABLE 2

COMPARISON WITH STATE-OF-THE-ART METHODS ON STARE. **RED**, **GREEN**, AND **BLUE** ARE THE TOP THREE (UNIT: %). THE P-VALUE OF THE PAIRED T-TEST BETWEEN EACH METHOD WITH THE PROPOSED SKELCON IS IN PARENTHESES.

Method	F	C	A	L	rSe	rSp	rAcc	Mcc
training/test=10/10								
Wavelets [16]	71.03(<0.001)	99.47(<0.001)	85.68(<0.001)	83.13(<0.001)	69.43(<0.001)	99.41 (<0.001)	94.18(<0.001)	73.77(<0.001)
LineDet [14]	18.83(<0.001)	99.78(<0.001)	41.42(<0.001)	42.94(<0.001)	26.28(<0.001)	99.98 (<0.001)	87.42(<0.001)	44.73(<0.001)
HED [54]	81.43 (<0.001)	99.85 (<0.001)	89.46(<0.001)	91.11 (<0.001)	75.36(<0.001)	95.53(<0.001)	92.10(<0.001)	69.04(<0.001)
DRIU [27]	80.75(<0.001)	99.84 (<0.001)	90.25 (<0.001)	89.49 (<0.001)	83.77 (<0.001)	97.27(<0.001)	95.04 (<0.001)	75.15 (<0.001)
V-GAN [31]	81.70 (<0.001)	99.75(<0.001)	91.91 (<0.001)	88.98(<0.001)	78.94 (<0.001)	99.59 (<0.001)	96.03 (<0.001)	81.39 (=0.562)
SkelCon(Ours)	83.59	99.92	91.25	91.58	85.14	98.48	96.22	81.95
twenty folds(Leave-One-Out)								
Gray-Level [15]	61.69(<0.001)	99.91 (<0.001)	79.97(<0.001)	76.37(<0.001)	60.24(<0.001)	98.99(<0.001)	92.57(<0.001)	69.74(<0.001)
MSLineDet [4]	63.15(<0.001)	98.86(<0.001)	80.19(<0.001)	78.01(<0.001)	85.07(<0.001)	96.65(<0.001)	94.63(<0.001)	67.92(=0.002)
DeepVessel [26]	66.25(<0.001)	99.65(<0.001)	83.18(<0.001)	79.58(<0.001)	62.43(<0.001)	99.60 (<0.001)	93.46(<0.001)	73.68(<0.001)
No-Pool [19]	05.86(<0.001)	99.88 (=0.172)	24.18(<0.001)	20.89(<0.001)	10.80(<0.001)	100.0 (<0.001)	85.10(<0.001)	29.39(<0.001)
FullyCRF [2]	65.72(<0.001)	99.10(=0.003)	82.68(<0.001)	78.82(<0.001)	72.85(<0.001)	98.15(=0.064)	93.84(<0.001)	72.57(<0.001)
JL-UNet [33]	76.92(=0.004)	99.60(<0.001)	88.79(=0.017)	86.52(=0.002)	80.96(=0.010)	99.00(=0.723)	96.18(<0.001)	78.23(=0.067)
SWT-FCN [28]	85.31 (=0.017)	99.81(<0.001)	93.62 (=0.039)	91.19 (=0.007)	85.72 (=0.657)	99.48 (<0.001)	97.22 (=0.019)	82.94 (=0.024)
DRIS-GP [1]	87.67 (<0.001)	99.76(<0.001)	95.16 (<0.001)	92.30 (<0.001)	90.72 (<0.001)	99.32(<0.001)	97.94 (<0.001)	85.45 (<0.001)
SkelCon(Ours)	82.73	99.86	92.28	89.65	86.10	99.05	96.91	80.92

TABLE 3

COMPARISON WITH STATE-OF-THE-ART METHODS ON CHASEDB1. **RED**, **GREEN**, AND **BLUE** ARE THE TOP THREE (UNIT: %). THE P-VALUE OF THE PAIRED T-TEST BETWEEN EACH METHOD WITH THE PROPOSED SKELCON IS IN PARENTHESES.

Method	F	C	A	L	rSe	rSp	rAcc	Mcc
DeepVessel [26]	64.88(<0.001)	99.73(<0.001)	80.03(<0.001)	81.21(<0.001)	55.00(<0.001)	98.22(=0.006)	93.15(<0.001)	67.52(<0.001)
No-Pool [19]	72.25(=0.002)	99.80(<0.001)	86.25(=0.007)	83.82(=0.002)	66.24(<0.001)	98.50 (=0.010)	94.69(<0.001)	76.65(=0.015)
FullyCRF [2]	63.70(<0.001)	99.73(<0.001)	80.60(<0.001)	79.09(<0.001)	63.44(<0.001)	97.33(=0.003)	93.33(<0.001)	70.10(<0.001)
JL-UNet [33]	74.59(=0.004)	99.77(<0.001)	87.56(=0.014)	85.32(=0.004)	69.81(=0.001)	97.86(=0.435)	94.56(<0.001)	76.64(=0.007)
SWT-FCN [28]	78.03 (=0.074)	99.89 (=0.007)	89.84 (=0.539)	86.88 (=0.015)	71.35 (=0.001)	98.72 (<0.001)	95.50 (=0.459)	79.92 (=0.726)
DRIS-GP [1]	80.82 (=0.825)	99.81 (<0.001)	91.71 (=0.149)	88.25 (=0.291)	76.06 (=0.314)	99.07 (<0.001)	96.36 (=0.032)	81.46 (=0.002)
SkelCon(Ours)	81.26	99.91	90.44	89.86	78.17	97.94	95.61	78.97

TABLE 4

COMPARISON WITH STATE-OF-THE-ART METHODS ON HRF. **RED**, **GREEN**, AND **BLUE** ARE THE TOP THREE (UNIT: %). THE P-VALUE OF THE PAIRED T-TEST BETWEEN EACH METHOD WITH THE PROPOSED SKELCON IS IN PARENTHESES.

Method	F	C	A	L	rSe	rSp	rAcc	Mcc
FullyCRF [2]	64.90 (=0.209)	99.46 (<0.001)	83.13 (<0.001)	78.24 (<0.001)	59.65 (=0.001)	97.09 (=0.098)	91.64 (=0.004)	69.17 (<0.001)
DRIS-GP [1]	71.20 (<0.001)	99.85 (<0.001)	87.16 (<0.001)	81.44 (=0.758)	66.11 (<0.001)	98.98 (<0.001)	94.25 (<0.001)	75.68 (<0.001)
SkelCon(Ours)	83.33	99.95	93.10	89.49	78.53	98.86	95.90	79.15

TABLE 5
PERFORMANCE ON RC-SLO, IOSTAR AND UoA-DR (UNIT: %). N.A. MEANS THE DATA IS NOT AVAILABLE.

Dataset	Method	F	C	A	L	rSe	rSp	rAcc	Se	Sp	Acc	Auc	Dice	Mcc
IOSTAR	LAD-OS [47]		N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	75.45	97.40	95.14	96.15	N.A.	73.18
	SkelCon(trained on DRIVE)	74.05	99.85	87.32	84.83	79.32	99.00	95.11	80.57	97.03	95.48	97.53	76.50	74.29
RC-SLO	LAD-OS [47]		N.A.	N.A.	N.A.	N.A.	N.A.	N.A.	77.87	97.10	95.12	96.26	N.A.	73.27
	SkelCon(trained on DRIVE)	82.15	99.89	90.55	90.60	76.35	98.27	95.36	79.32	97.56	95.48	97.91	78.52	76.51
UoA-DR	SkelCon(trained on DRIVE)	33.50	99.69	55.19	58.73	51.11	95.20	89.81	47.49	93.25	90.47	82.37	36.91	33.20

in [46]. In the present work, the MATLAB code⁵ provided by Yan *et al.* [46] was used to calculate these metrics. In addition, the Matthews Correlation Coefficient (Mcc) was also included [57]. Paired T test was used to calculate P values between the proposed SkelCon and existing methods. Statistical significance was defined as P value < 0.05.

C. Implementation Details

The model is trained on a TITAN RTX(12GB) GPU running the Pytorch framework. The RAdam optimizer is used, and the initial learning rate is set at 0.01 with a weight decay of 2e-4. The ReduceLROnPlateau mechanism is used with a coefficient of 0.7. The patience and cooldown parameters are both 2, and the minimum learning rate is set to 1e-5.

For data pre-processing, patches with a size of 128×128 are randomly cropped from the original images due to the large size of fundus images. All patches are normalized before being fed into the neural network. In addition, the batch size is set to 32 in the training phase. A total of 160 epochs are executed in each experiment, and all patches are iterated eight times in each training epoch. Finally, the probability maps predicted by the model are binarized with the threshold of 0.5 to obtain the final segmentation.

TABLE 6
SUBJECTIVE EVALUATION BY THREE OPHTHALMOLOGISTS ON CLINICAL IMAGES FROM JSIEC39 AND RFMiD DATASETS.

Best performance	SkelCon(Ours)	DAP [41]	DRIS-GP [1]
Ophthalmologist #1	76.25%	21.25%	2.50%
Ophthalmologist #2	71.25%	28.75%	0.00%
Ophthalmologist #3	80.00%	18.75%	1.25%

D. Comparison with State-of-the-Art Methods

1) General Comparison: To comprehensively evaluate the proposed method, experiments were first conducted on the four datasets, and the quantitative metrics are listed in Tables 1, 2, 3, and 4. Overall, the proposed method, SkelCon, obtains the best scores on the most considered metrics, especially the *F* metric, which reflects the connectivity and accuracy of vessel segmentation. On the STARE dataset, two schemes of testing were conducted, the 10/10 dataset division and the leave-one-out method. SkelCon achieves the best performance on almost all metrics when tested with 10/10 data division, and ranks the third overall when tested with the leave-one-out method (Table 2). It should be noted that the proposed

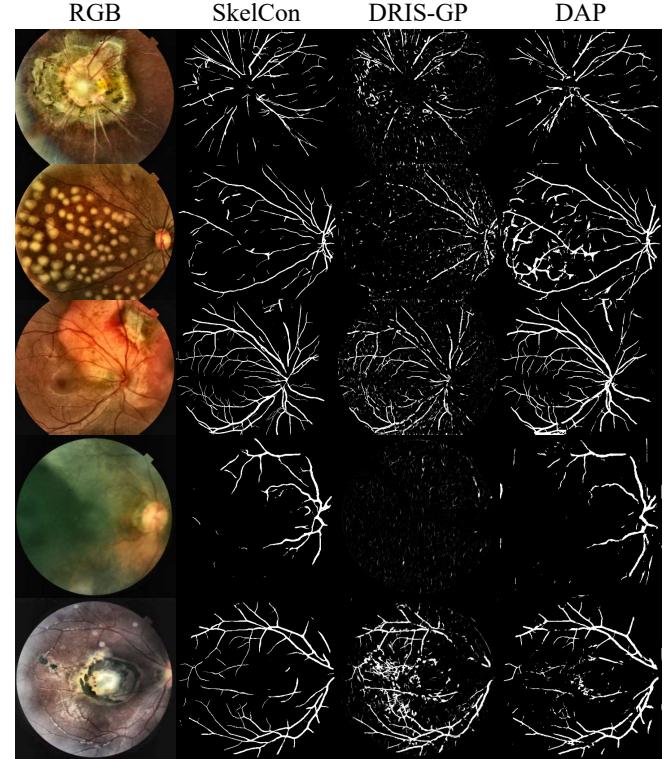


Fig. 8. Visual results on images from JSIEC39 and RFMiD datasets.

method significantly outperforms recent methods on the HRF dataset, as shown in Table 4. The HRF dataset contains some challenging images from diabetic retinopathy and glaucoma patients. From the experiments, the proposed method achieves state-of-the-art performance and performs better in challenging cases compared with the considered methods, including DRIS-GP [1], SWT-FCN [28], JL-UNet [33], and FullyCRF [2].

Fig. 6 presents several examples, showing that SkelCon obtains more complete segmentation of thin vessels and fewer false alarms. The last row of Fig. 6 indicates SkelCon can accurately extract vessels from regions with serious soft exudation. We have conducted experiments on the recently developed datasets, including UoA-DR [48], IOSTAR and RC-SLO [47]. In addition, some clinical fundus images with multiple disease disturbances were also considered, e.g., RFMiD [49] and JSIEC39 [50] datasets. The experiments produce desired results on these datasets, and we draw a similar conclusion with the experiments on the four popular datasets. In addition, even some clinical fundus images from RFMiD [49] and JSIEC39 [50] datasets contain serious pathological lesions,

⁵<https://github.com/ZengqiangYan/skeletalSimilarityMetric>

TABLE 7
COMPARING SKELCON WITH TWO OBSERVERS AND OTHER METHODS ON THE DRIVE DATASET (UNIT: %).

Ground Truth	Year	Method	F	C	A	L	rSe	rSp	rAcc	Mcc
1st observer	2021	SkelCon(Ours)	85.30	99.85	93.58	91.26	83.23	98.59	94.61	80.30
	2004	2nd observer	83.75	100	93.98	89.06	85.84	99.19	95.74	76.00
2nd observer	2021	SkelCon(Ours)	87.31	99.84	94.50	92.52	84.80	98.59	95.07	82.67
	2004	1st observer	83.75	100	93.98	89.06	87.02	98.95	95.87	76.00

TABLE 8

THE IMPROVED RANKING ON DRIVE, WITHOUT OUTLIERS (ONLY THE TOP 32 ARE SHOWN DUE TO SPACE CONSTRAINTS).

Key	Rank	Original rank	Rank diff.	Adjusted acc.	Published acc.	Category	Operation
Cherukuri2020 [1]	1	27	26	.9600	.9600	FoV	deep learning
Wu2020 [58]	2	31	29	.9582	.9582	FoV	deep learning
Jiang2019 [59]	3	7	4	.9578	.9706	all pixels	deep learning
Park2020 [60]	4	6	2	.9577	.9706	all pixels	deep learning
Oliveira2019 [28]	5	33	28	.9576	.9576	FoV	deep learning
Noh2019 [61]	6	36	30	.9569	.9569	FoV	deep learning
Son2017 [31]	7	37	30	.9568	.9568	FoV	deep learning
SkelCon(Proposed)	8	40	32	.9555	.9555	FoV	deep learning
Khanal2020 [55]	9	41	32	.9551	.9551	FoV	deep learning
Yan2018 [33]	10	44	34	.9542	.9542	FoV	deep learning
Sun2021 [41]	11	47	36	.9540	.9540	FoV	deep learning
Liskowski2016 [19]	12	50	38	.9535	.9535	FoV	deep learning
Ngo2017 [62]	13	52	39	.9533	.9533	FoV	deep learning
Dasgupta2017 [63]	14	53	39	.9533	.9533	FoV	deep learning
Li2016 [18]	15	58	43	.9527	.9527	FoV	deep learning
Mo2017 [64]	16	61	45	.9521	.9521	FoV	deep learning
Ganin2015 [53]	17	67	50	.9506	.9506	FoV	other
Zhang2018 [65]	18	68	50	.9504	.9504	FoV	deep learning
Song2017 [66]	19	72	53	.9499	.9499	FoV	deep learning
XiuQin2019 [67]	20	10	-10	.9495	.9650	all pixels	deep learning
Kovacs2016 [68]	21	74	53	.9494	.9494	FoV	classical
Roy2015 [11]	22	73	51	.9494	.9494	FoV	classical
Brancale2018 [69]	23	75	52	.9490	.9490	FoV	deep learning
Palanivel2020 [70]	24	77	53	.9480	.9480	FoV	supervised
Shah2017 [71]	25	80	55	.9479	.9479	FoV	supervised
Shukla2020 [72]	26	81	55	.9476	.9476	FoV	classical
Zhang2016 [47]	27	82	55	.9476	.9476	FoV	classical
Cheng2014 [73]	28	83	55	.9474	.9474	FoV	supervised
Sreejini2015 [74]	29	12	-17	.9470	.9633	all pixels	classical
Zhou2017 [75]	30	84	54	.9469	.9469	FoV	other
Melinscak2015 [76]	31	85	54	.9466	.9466	FoV	deep learning
Mendonca2006 [10]	32	86	54	.9463	.9463	FoV	classical

the proposed method still outperforms the compared methods. Quantitative metrics and some examples are listed in the Table 5, Fig. 7, and Fig. 8.

Note that some of the fundus datasets were not collected for vessel segmentation, and hence no vessel labels are available. For example, RFMiD [49] and JSIEC39 [50] datasets are for Multi-Disease Classification. Therefore, we show some qualitative results when evaluating the proposed methods in Fig. 8. It is easy to find that the proposed detects more complete and continuous vessels compared with previous methods on clinical images. In addition, we performed an experiment with subjective evaluation by three ophthalmologists from West China Hospital of Sichuan University. Eighty images were randomly selected from the RFMiD [49] and JSIEC39 [50]

datasets. For each image, the original and processed ones of three different methods were presented on the screen synchronously in random order, and three ophthalmologists were required to independently vote the one with the best result based on their knowledge and experiences. The percentage of the processed images with the best performance voted by ophthalmologists are listed in Table 6, which shows that the proposed method obtains higher performance than the other two compared methods. In addition, the proposed method and DAP outperform the DRIS-GP on clinical images, which may benefit from the efficient data augmentation methods in the proposed method and DAP.

We introduced the ranking with the ranking system in [77] on the DRIVE dataset. For convenient comparison, we recalculated the metric of ACC (accuracy) using the code of [77]. Our method won the eighth place (among 80 considered methods), which means the proposed method obtains good performance for vessel segmentation, considering that the metrics for the top-rank methods are just slightly different. The ranking is shown in the following Table 8.

2) Vessel Segmentation in Optic Disc and Lesion Areas:

More detailed comparisons are shown in Fig. 9, which presents zoomed-in views of optic disc and lesion areas. The first and second rows in Fig. 9 show that with state-of-the-art methods, vessels in the optic disc area may be missed due to low contrast, and the border of the optic disc may be incorrectly identified as vessels. In comparison, SkelCon usually extracts more complete vessels and fewer false alarms in the optic disc area. Lesions often make vessel segmentation difficult because many lesions occur near vessels, such as in the case of hemorrhage. The third and fourth rows in Fig. 9 show that the state-of-the-art methods may incorrectly segment lesions as vessels due to the similar local features. In contrast, SkelCon can accurately segment vessels in lesion areas. This might be attributed to SkelCon's contrastive loss to improve the model's discrimination of vessels and background pixels.

3) Segmentation of Thin Vessels: Accurate thin vessel segmentation is another contribution of the proposed method. Fig. 10 presents several images to verify the capability of thin vessel detection. Compared with state-of-the-art methods, SkelCon extracts more complete thin vessels. In addition, according to the two doctors' labels provided in the DRIVE dataset, some doctors may also miss tiny vessels. As shown in Fig. 11, the difference of labels between the two observers is mainly reflected in the thin vessels. SkelCon shows good performance when detecting thin vessels. The quantitative metrics listed in Table 1 further confirm this conclusion. SkelCon achieves comparable performance to human observers and obtains higher scores of *F* and *L* in Table 7.

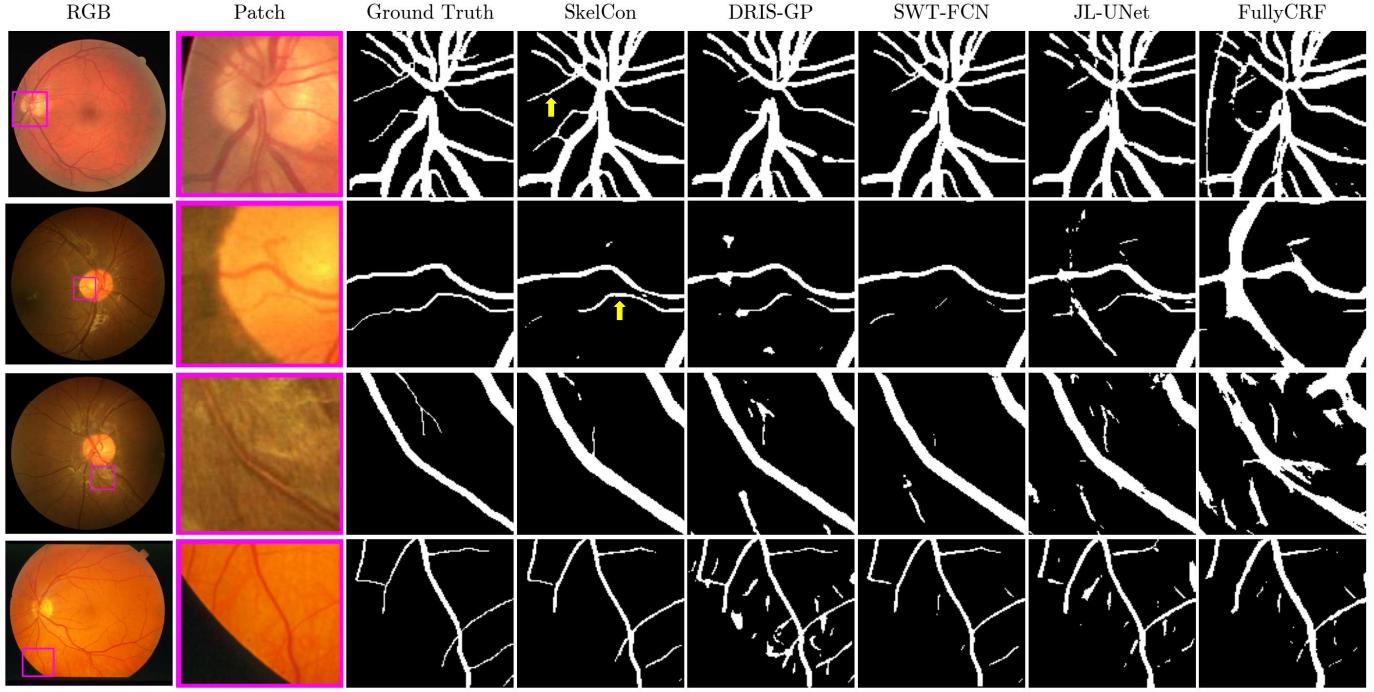


Fig. 9. Vessel segmentation in the optic disc area (the first two rows) and lesion areas (the last two rows). Yellow arrows in the SkelCon results in the first and second rows indicate that the proposed method usually extracts more complete vessels than compared methods.

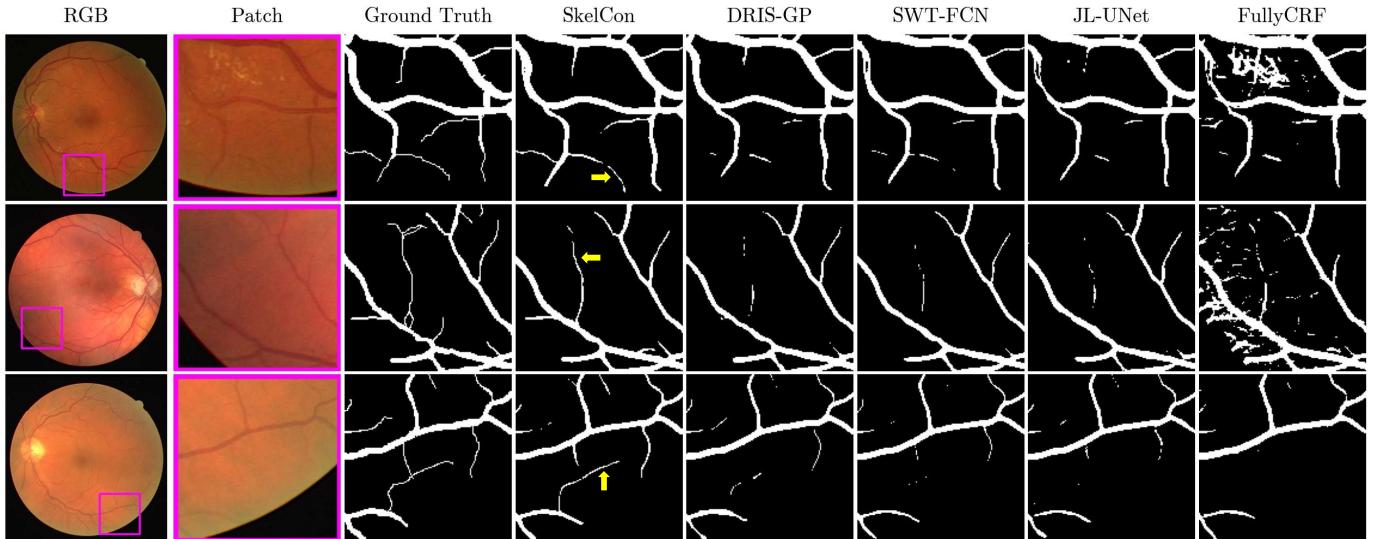


Fig. 10. Thin vessel segmentation indicated by yellow arrows in the SkelCon results.

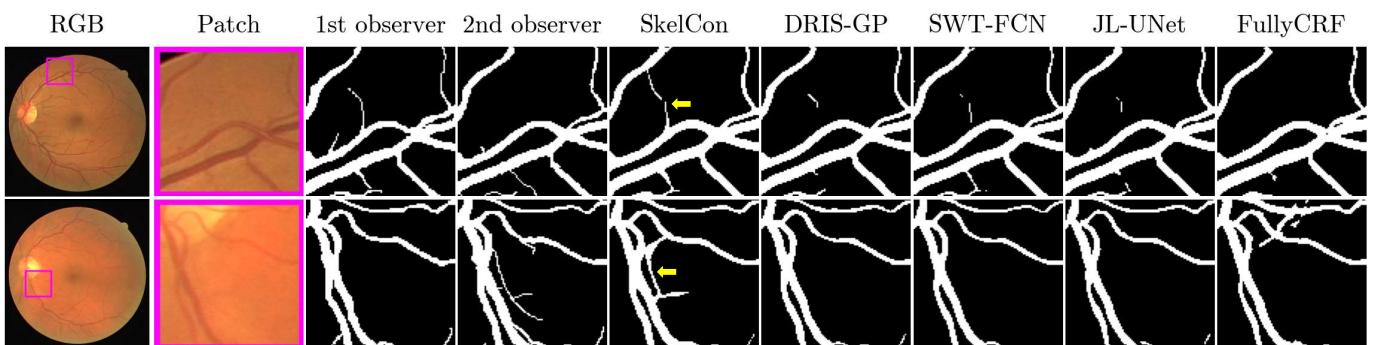


Fig. 11. Thin vessel segmentation comparison among SkelCon and two observers on the DRIVE dataset. Yellow arrows in the SkelCon results indicate thin vessels may be missed by doctors.

TABLE 9
CROSS-VALIDATION ON THE DRIVE AND STARE DATASETS (UNIT: %).

Training Set	Test Set	Method	F	C	A	L	rSe	rSp	rAcc	Mcc
DRIVE	STARE	JL-UNet [33]	73.07	99.50	86.37	84.46	82.18	98.84	96.29	75.26
		SkelCon(w/o CSM)	72.12	99.81	89.18	80.89	67.11	99.75	91.37	74.93
		SkelCon(Ours)	79.04	99.86	90.54	87.32	86.20	98.44	96.49	77.82
STARE	DRIVE	JL-UNet [33]	69.59	99.65	87.82	79.35	64.94	99.60	90.70	75.07
		SkelCon(w/o CSM)	73.15	99.80	89.89	81.40	67.79	99.78	91.70	75.96
		SkelCon(Ours)	74.69	99.88	90.38	82.61	70.86	99.73	92.30	76.10

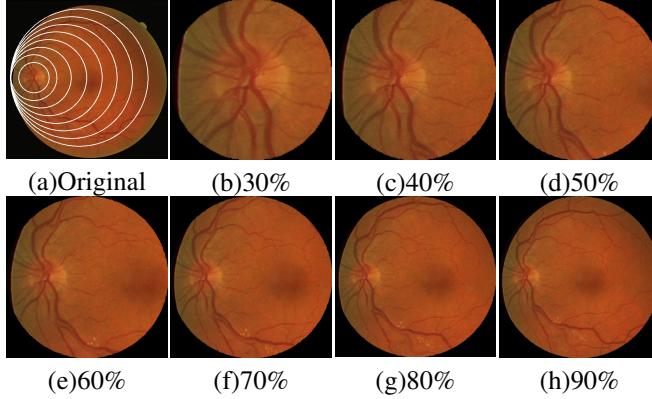


Fig. 12. Synthetic images in FoV of different percentages. The solid white lines in original image (a) represent 30 ~ 90% FoV, respectively.

TABLE 10

PERFORMANCE OF SMARTPHONE-BASED SYNTHETIC IMAGES ON DRIVE, INCLUDING SYNTHETIC IMAGES FROM PANOPTIC(PO), D-EYE(DE), PEEK RETINA(PR), AND INVIEW(IN) SYSTEMS, WITH 32%, 40%, 45%, AND 94% FOVS(UNIT:%).

Method	F	C	A	L	rSe	rSp	rAcc	Mcc
Original	85.30	99.85	93.58	91.26	83.23	98.59	94.61	80.30
94%(IN)	85.12	99.84	94.69	90.00	81.33	99.33	94.55	78.86
90%	84.34	99.86	94.31	89.50	80.42	99.24	94.27	78.81
80%	83.81	99.88	94.00	89.20	79.72	99.12	94.01	78.85
70%	83.49	99.90	93.78	89.05	78.95	98.96	93.86	79.18
60%	83.06	99.91	93.42	88.92	78.09	98.77	93.61	79.41
50%	83.20	99.92	93.33	89.13	77.06	98.58	93.33	80.54
45%(PR)	82.85	99.93	92.95	89.13	76.06	98.37	93.00	80.07
40%(DE)	82.38	99.94	92.46	89.07	75.01	98.15	92.63	80.05
32%(PO)	82.09	99.94	92.02	89.18	73.31	97.91	92.03	80.93
30%	81.51	99.95	91.49	89.07	71.66	97.67	91.41	80.28

E. Generalization

1) **Cross-validation:** Cross-validation between datasets means that a model is trained on one dataset, and then tested on a new dataset, so as to evaluate the generalization performance of the model.

In this experiment, the DRIVE and STARE datasets are selected for further cross-validation and the results are listed in Table 9. Compared with JL-UNet [33], SkelCon obtains significantly higher scores on almost all considered metrics. The results of SkelCon without the data augmentation module, i.e., CSM, are also listed in Table 9. These results indicate the improvement of generalization mainly benefits from the effective data augmentation provided by CSM.

2) **Performance of smartphone-based synthetic images:** To explore the possibility to detect vessels from the images acquired using smartphone-based imaging techniques, synthetic

images were first obtained according to the method reported in [78]. The main steps of synthetic image generation with the method provided in [78] are image cropping and resizing (shown in Fig. 12). We developed an extra experiment by testing the proposed method on smartphone-based synthetic images generated based on the DRIVE dataset.

We first cropped the original fundus images with 30~90% of the full FoV, and then resized each patch to 512x512, shown in Fig. 12. Experimental results are listed in Table 10. With the reducing of FoV, the performance decreases slowly, but acceptable results are still obtained. Therefore, we can safely conclude that the proposed method has great potential in applying to retinal images acquired using smartphone-based imaging techniques.

F. Ablation Study

TABLE 11

ABLATION STUDY ON THE DRIVE DATASET (UNIT: %). BASELINE IS THE CASCDED LUNETS WITH NOISE & FLIP, +CL IN THE LAST ROW IS THE PROPOSED SKELCON.

Method	F	C	A	L	rSe	rSp	rAcc	Mcc
Baseline	84.83	99.80	93.47	90.89	82.15	98.80	94.48	80.73
+CSM	85.07	99.78	93.57	91.08	82.59	98.74	94.56	80.75
+SFM	85.11	99.85	93.48	91.15	82.85	98.65	94.55	80.58
+CL	85.30	99.85	93.58	91.26	83.23	98.59	94.61	80.30

In the ablation study, the benefits of each module in the proposed model are shown in Table 11. The baseline is the cascaded LUNets without the data augmentation module (CSM), skeleton fitting module (SFM), or contrastive loss (CL). By adding CSM, SFM, and CL, almost all considered metrics increase correspondingly, except rSp (Table 11). For example, the final model obtains the highest F at 85.30% and rSe at 83.23%.

In addition, the metric of rSp, which reflects the specificity (i.e., true negative rate) of vessel segmentation, is carefully analyzed. SkelCon decreases the rSp compared with the baseline since it usually improves the connectivity of vessel segmentation. Fig. 13 shows two examples, illustrating that the proposed method can connect vessels missed by doctors (vessels indicated in yellow arrows in the first row). Of course, there may also be some wrong connections, like multiple vessels in the optic disc (see the bottom row in Fig. 13). Increasing connectivity with SkelCon is beneficial (see the results of compared methods in Fig. 13). It can extract more complete vessels, especially thin vessels, and it can further suppress disturbances from lesions that usually lack connectivity.

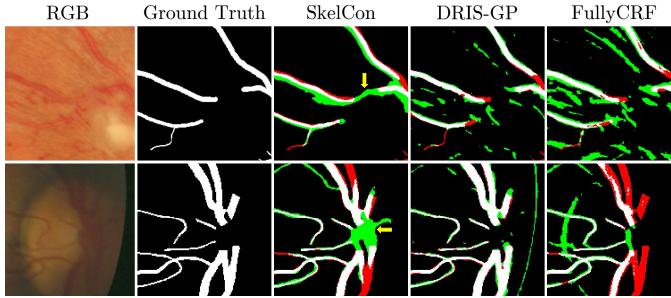


Fig. 13. The improved connectivity of vessel segmentation. Yellow arrows indicate the profitable connection missed by doctors (top) and wrong connection in the optic disc (bottom). White pixels: true positive detection; green pixels: false positive detection; red pixels: false negative detection; black pixels: true negative detection.

G. Comparison of Parameters

As Fig. 14 shows, SkelCon is compared with other methods in terms of the number of learnable parameters and main metrics on the DRIVE dataset (i.e., SS and F). It should be noted that CSM, SFM, and CL are only used in the training phase. Therefore, the cascaded LUNets have 0.42MB of parameters only. In contrast, JL-UNET [33], DeepDyn [55], and DAP [41] have about 33.0MB of parameters. There are also several lightweight networks in the current mainstream approaches, such as SWT-FCN (about 0.7MB of parameters), and DRIS-GP (only 0.13MB of parameters). However, the proposed achieves better performance with a lightweight network.

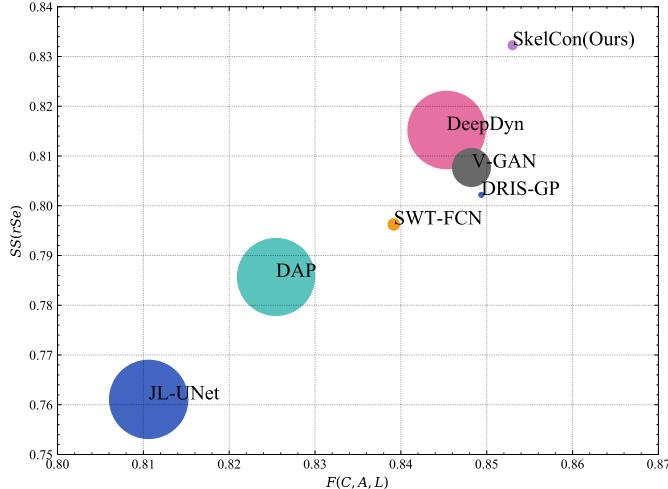


Fig. 14. Comparison of parameters and performance between SkelCon and state-of-the-art methods on the DRIVE dataset. The size of each circle in the figure represents the number of parameters.

V. DISCUSSION

Recently, great progress has been made in retinal vessel segmentation with the rise of deep learning. The main challenges of vessel detection that remain are the segmentation of fine vessels [33], [46] and tackling of lesion interference [1], [31]. Along with this line, our study also concentrates on improving the segmentation for the fundus images containing lesions and fine vessels by designing new methods for data augmentation and sample balancing. The proposed method can

effectively preserve the morphology of the vessels and improve the completeness and continuity of thin vessels.

In addition, in order to optimize the boundary of vessels and tackle the problem of false-positive samples caused by lesions, this study employs a contrastive loss inspired by the contrastive learning scheme [35]. As a self-supervised pre-training method, contrastive learning can have a better effect on a small amount of labeled data, with lots of unlabeled data [37], [79]. It is especially suitable for medical image processing tasks, which are facing difficulties in data acquisition and annotation [39], [40]. Moreover, contrastive learning can be combined with supervised learning to further improve the performance of models.

However, for some clinical fundus images in which vessels are seriously destroyed by lesions, e.g., images from the RFMiD [49] and JSIEC39 [50] datasets, the performances of almost all vessel segmentation methods decline significantly. To address this challenge, we believe that the performances of various deep-learning-based methods are expected to be further improved by collecting more diverse challenging images and labeled datasets. In addition, improving the quality of the fundus images in the process of image acquisition is a fantastic way. For example, it is expected to build a system that can provide feedback to the clinician to help him/her acquire and select informative images for processing. Then, an image enhancement system could be used to further improve the quality of the images. We believe such a system can facilitate not only clinic diagnosis but also pattern recognition algorithms.

VI. CONCLUSION

In this work, a new training paradigm for retinal vessel segmentation is proposed, which includes a new data augmentation method, skeletal prior guiding mechanism, and contrastive loss. Using data augmentation with color-space channel mixing, the proposed method can produce more diverse vessel samples to improve its generalization ability. In addition, prior knowledge of the vascular skeleton provided by the skeletal prior guiding mechanism can greatly avoid the fracture of vessel segmentation and improve the performance of thin vessel segmentation. Finally, the contrastive loss further strengthens the vessel segmentation, especially in lesion areas. The proposed method is shown to achieve state-of-the-art performance in various experiments. Future work includes building a more explicit representation of vascular structures with neural networks, and extending the method to other line segmentation problems in medicine, such as rib detection and cell boundary extraction.

REFERENCES

- [1] V. Cherukuri, V. K. Bg, R. Bala, and V. Monga, "Deep retinal image segmentation with regularization under geometric priors," *IEEE Transactions on Image Processing*, vol. 29, pp. 2552–2567, 2019.
- [2] J. I. Orlando, E. Prokofyeva, and M. B. Blaschko, "A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images," *IEEE transactions on Biomedical Engineering*, vol. 64, no. 1, pp. 16–27, 2016.

- [3] J. I. Orlando and M. Blaschko, "Learning fully-connected crfs for blood vessel segmentation in retinal images," in *international conference on medical image computing and computer-assisted intervention*, 2014, pp. 634–641.
- [4] U. T. Nguyen, A. Bhuiyan, L. A. Park, and K. Ramamohanarao, "An effective retinal blood vessel segmentation method using multi-scale line detection," *Pattern recognition*, vol. 46, no. 3, pp. 703–715, 2013.
- [5] R. Poplin, A. V. Varadarajan, K. Blumer, Y. Liu, M. V. McConnell, G. S. Corrado, L. Peng, and D. R. Webster, "Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning," *Nature Biomedical Engineering*, vol. 2, no. 3, pp. 158–164, 2018.
- [6] K. B. Khan, A. A. Khaliq, A. Jalil, M. A. Iftikhar, N. Ullah, M. W. Aziz, K. Ullah, and M. Shahid, "A review of retinal blood vessels extraction techniques: challenges, taxonomy, and future trends," *Pattern Analysis and Applications*, vol. 22, no. 3, pp. 767–802, 2019.
- [7] T. Li, W. Bo, C. Hu, H. Kang, H. Liu, K. Wang, and H. Fu, "Applications of deep learning in fundus images: A review," *Medical Image Analysis*, vol. 69, p. 101971, 2021.
- [8] D. S. Ting, L. Peng, A. V. Varadarajan, P. A. Keane, P. M. Burlina, M. F. Chiang, L. Schmetterer, L. R. Pasquale, N. M. Bressler, D. R. Webster *et al.*, "Deep learning in ophthalmology: the technical and clinical considerations," *Progress in retinal and eye research*, vol. 72, p. 100759, 2019.
- [9] X. Jiang and D. Mojon, "Adaptive local thresholding by verification-based multithreshold probing with application to vessel detection in retinal images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 1, pp. 131–137, 2003.
- [10] A. M. Mendonca and A. Campilho, "Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction," *IEEE transactions on medical imaging*, vol. 25, no. 9, pp. 1200–1213, 2006.
- [11] S. Roychowdhury, D. D. Koozekanani, and K. K. Parhi, "Iterative vessel segmentation of fundus images," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 7, pp. 1738–1749, 2015.
- [12] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, "Detection of blood vessels in retinal images using two-dimensional matched filters," *IEEE Transactions on medical imaging*, vol. 8, no. 3, pp. 263–269, 1989.
- [13] F. Zana and J.-C. Klein, "Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation," *IEEE transactions on image processing*, vol. 10, no. 7, pp. 1010–1019, 2001.
- [14] E. Ricci and R. Perfetti, "Retinal blood vessel segmentation using line operators and support vector classification," *IEEE transactions on medical imaging*, vol. 26, no. 10, pp. 1357–1365, 2007.
- [15] D. Marín, A. Aquino, M. E. Gegúndez-Arias, and J. M. Bravo, "A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features," *IEEE Transactions on medical imaging*, vol. 30, no. 1, pp. 146–158, 2010.
- [16] J. V. Soares, J. J. Leandro, R. M. Cesar, H. F. Jelinek, and M. J. Cree, "Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification," *IEEE Transactions on medical Imaging*, vol. 25, no. 9, pp. 1214–1222, 2006.
- [17] R. Annunziata and E. Trucco, "Accelerating convolutional sparse coding for curvilinear structures segmentation by refining scird-ts filter banks," *IEEE transactions on medical imaging*, vol. 35, no. 11, pp. 2381–2392, 2016.
- [18] Q. Li, B. Feng, L. Xie, P. Liang, H. Zhang, and T. Wang, "A cross-modality learning approach for vessel segmentation in retinal images," *IEEE transactions on medical imaging*, vol. 35, no. 1, pp. 109–118, 2015.
- [19] P. Liskowski and K. Krawiec, "Segmenting retinal blood vessels with deep neural networks," *IEEE transactions on medical imaging*, vol. 35, no. 11, pp. 2369–2380, 2016.
- [20] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Transactions on Medical imaging*, vol. 19, no. 3, pp. 203–210, 2000.
- [21] J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. Van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE transactions on medical imaging*, vol. 23, no. 4, pp. 501–509, 2004.
- [22] M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. R. Rudnicka, C. G. Owen, and S. A. Barman, "An ensemble classification-based approach applied to retinal blood vessel segmentation," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 9, pp. 2538–2548, 2012.
- [23] J. Odstrcilík, R. Kolar, A. Budai, J. Hornegger, J. Jan, J. Gazarek, T. Kubena, P. Černosek, O. Svoboda, and E. Angelopoulou, "Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database," *IET Image Processing*, vol. 7, no. 4, pp. 373–383, 2013.
- [24] B. Sheng, P. Li, S. Mo, H. Li, X. Hou, Q. Wu, J. Qin, R. Fang, and D. D. Feng, "Retinal vessel segmentation using minimum spanning superpixel tree detector," *IEEE transactions on cybernetics*, vol. 49, no. 7, pp. 2707–2719, 2018.
- [25] B. S. Lam, Y. Gao, and A. W.-C. Liew, "General retinal vessel segmentation using regularization-based multiconcavity modeling," *IEEE Transactions on Medical Imaging*, vol. 29, no. 7, pp. 1369–1381, 2010.
- [26] H. Fu, Y. Xu, S. Lin, D. W. Kee Wong, and J. Liu, "Deepvessel: Retinal vessel segmentation via deep learning and conditional random field," in *International conference on medical image computing and computer-assisted intervention*, 2016, pp. 132–139.
- [27] K.-K. Maninis, J. Pont-Tuset, P. Arbeláez, and L. V. Gool, "Deep retinal image understanding," in *International conference on medical image computing and computer-assisted intervention*, 2016, pp. 140–148.
- [28] A. Oliveira, S. Pereira, and C. A. Silva, "Retinal vessel segmentation based on fully convolutional neural networks," *Expert Systems with Applications*, vol. 112, pp. 229–242, 2018.
- [29] Z. Gu, J. Cheng, H. Fu, K. Zhou, H. Hao, Y. Zhao, T. Zhang, S. Gao, and J. Liu, "Ce-net: Context encoder network for 2d medical image segmentation," *IEEE transactions on medical imaging*, vol. 38, no. 10, pp. 2281–2292, 2019.
- [30] L. Mou, L. Chen, J. Cheng, Z. Gu, Y. Zhao, and J. Liu, "Dense dilated network with probability regularized walk for vessel detection," *IEEE transactions on medical imaging*, vol. 39, no. 5, pp. 1392–1403, 2019.
- [31] J. Son, S. J. Park, and K.-H. Jung, "Retinal vessel segmentation in fundoscopic images with generative adversarial networks," *arXiv preprint arXiv:1706.09318*, 2017.
- [32] H. Zhao, H. Li, S. Maurer-Stroh, Y. Guo, Q. Deng, and L. Cheng, "Supervised segmentation of un-annotated retinal fundus images by synthesis," *IEEE transactions on medical imaging*, vol. 38, no. 1, pp. 46–56, 2018.
- [33] Z. Yan, X. Yang, and K.-T. Cheng, "Joint segment-level and pixel-wise losses for deep learning based retinal vessel segmentation," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 9, pp. 1912–1923, 2018.
- [34] Z. Yan, X. Yang, and K.-T. Cheng, "A three-stage deep learning model for accurate retinal vessel segmentation," *IEEE journal of Biomedical and Health Informatics*, vol. 23, no. 4, pp. 1427–1436, 2018.
- [35] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, "Momentum contrast for unsupervised visual representation learning," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 9729–9738.
- [36] J.-B. Grill, F. Strub, F. Altché, C. Tallec, P. Richemond, E. Buchatskaya, C. Doersch, B. Avila Pires, Z. Guo, M. Gheshlaghi Azar *et al.*, "Bootstrap your own latent-a new approach to self-supervised learning," vol. 33, 2020, pp. 21271–21284.
- [37] X. Chen and K. He, "Exploring simple siamese representation learning," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 15750–15758.
- [38] P. Khosla, P. Teterwak, C. Wang, A. Sarna, Y. Tian, P. Isola, A. Maschinot, C. Liu, and D. Krishnan, "Supervised contrastive learning," vol. 33, 2020, pp. 18661–18673.
- [39] H. Li, X. Yang, J. Liang, W. Shi, C. Chen, H. Dou, R. Li, R. Gao, G. Zhou, J. Fang *et al.*, "Contrastive rendering for ultrasound image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2020, pp. 563–572.
- [40] P. Pandey, A. Pai, N. Bhatt, P. Das, G. Makharia, P. AP *et al.*, "Contrastive semi-supervised learning for 2d medical image segmentation," *arXiv preprint arXiv:2106.06801*, 2021.
- [41] X. Sun, H. Fang, Y. Yang, D. Zhu, L. Wang, J. Liu, and Y. Xu, "Robust retinal vessel segmentation from a data augmentation perspective," in *International Workshop on Ophthalmic Medical Image Analysis*, 2021, pp. 189–198.
- [42] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, 2015, pp. 234–241.
- [43] A. Galdran, A. Anjos, J. Dolz, H. Chakor, H. Lombaert, and I. B. Ayed, "The little w-net that could: state-of-the-art retinal vessel segmentation with minimalist models," *arXiv preprint arXiv:2009.01907*, 2020.
- [44] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

- [45] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," pp. 2980–2988, 2017.
- [46] Z. Yan, X. Yang, and K.-T. Cheng, "A skeletal similarity metric for quality evaluation of retinal vessel segmentation," *IEEE transactions on medical imaging*, vol. 37, no. 4, pp. 1045–1057, 2017.
- [47] J. Zhang, B. Dashtbozorg, E. Bekkers, J. P. Pluim, R. Duits, and B. M. ter Haar Romeny, "Robust retinal vessel segmentation via locally adaptive derivative frames in orientation scores," *IEEE transactions on medical imaging*, vol. 35, no. 12, pp. 2631–2644, 2016.
- [48] R. J. Chalakkal, W. H. Abdulla, and S. Sinumol, "Comparative analysis of university of auckland diabetic retinopathy database," in *Proceedings of the 9th International Conference on Signal Processing Systems*, 2017, pp. 235–239.
- [49] S. Pachade, P. Porwal, D. Thulkar, M. Kokare, G. Deshmukh, V. Sahasrabuddhe, L. Giancardo, G. Quellec, and F. Mériudeau, "Retinal fundus multi-disease image dataset (rfmid): a dataset for multi-disease detection research," *Data*, vol. 6, no. 2, p. 14, 2021.
- [50] L.-P. Cen, J. Ji, J.-W. Lin, S.-T. Ju, H.-J. Lin, T.-P. Li, Y. Wang, J.-F. Yang, Y.-F. Liu, S. Tan *et al.*, "Automatic detection of 39 fundus diseases and conditions in retinal photographs using deep neural networks," *Nature communications*, vol. 12, no. 1, pp. 1–13, 2021.
- [51] C. Becker, R. Rigamonti, V. Lepetit, and P. Fua, "Supervised feature learning for curvilinear structure segmentation," in *International conference on medical image computing and computer-assisted intervention*, 2013, pp. 526–533.
- [52] P. Dollár and C. L. Zitnick, "Structured forests for fast edge detection," in *Proceedings of the IEEE international conference on computer vision*, 2013, pp. 1841–1848.
- [53] Y. Ganin and V. Lempitsky, "N4-fields: Neural network nearest neighbor fields for image transforms," in *Asian conference on computer vision*, 2014, pp. 536–551.
- [54] S. Xie and Z. Tu, "Holistically-nested edge detection," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1395–1403.
- [55] A. Khanal and R. Estrada, "Dynamic deep networks for retinal vessel segmentation," *Frontiers in Computer Science*, p. 35, 2020.
- [56] M. E. Gegúndez-Arias, A. Aquino, J. M. Bravo, and D. Marín, "A function for quality evaluation of retinal vessel segmentations," *IEEE transactions on medical imaging*, vol. 31, no. 2, pp. 231–239, 2011.
- [57] K. B. Khan, M. S. Siddique, M. Ahmad, and M. Mazzara, "A hybrid unsupervised approach for retinal vessel segmentation," *BioMed Research International*, vol. 2020, 2020.
- [58] Y. Wu, Y. Xia, Y. Song, Y. Zhang, and W. Cai, "Nfn+: a novel network followed network for retinal vessel segmentation," *Neural Networks*, vol. 126, pp. 153–162, 2020.
- [59] Y. Jiang, H. Zhang, N. Tan, and L. Chen, "Automatic retinal blood vessel segmentation based on fully convolutional neural networks," *Symmetry*, vol. 11, no. 9, p. 1112, 2019.
- [60] K.-B. Park, S. H. Choi, and J. Y. Lee, "M-gan: Retinal blood vessel segmentation by balancing losses through stacked deep fully convolutional networks," *IEEE Access*, vol. 8, pp. 146 308–146 322, 2020.
- [61] K. J. Noh, S. J. Park, and S. Lee, "Scale-space approximated convolutional neural networks for retinal vessel segmentation," *Computer methods and programs in biomedicine*, vol. 178, pp. 237–246, 2019.
- [62] L. Ngo and J.-H. Han, "Multi-level deep neural network for efficient segmentation of blood vessels in fundus images," *Electronics Letters*, vol. 53, no. 16, pp. 1096–1098, 2017.
- [63] A. Dasgupta and S. Singh, "A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation," in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. IEEE, 2017, pp. 248–251.
- [64] J. Mo and L. Zhang, "Multi-level deep supervised networks for retinal vessel segmentation," *International journal of computer assisted radiology and surgery*, vol. 12, no. 12, pp. 2181–2193, 2017.
- [65] Y. Zhang and A. Chung, "Deep supervision with additional labels for retinal vessel segmentation task," in *International conference on medical image computing and computer-assisted intervention*, 2018, pp. 83–91.
- [66] J. Song and B. Lee, "Development of automatic retinal vessel segmentation method in fundus images via convolutional neural networks," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2017, pp. 681–684.
- [67] P. Xiuqin, Q. Zhang, H. Zhang, and S. Li, "A fundus retinal vessels segmentation scheme based on the improved deep learning u-net model," *IEEE Access*, vol. 7, pp. 122 634–122 643, 2019.
- [68] G. Kovács and A. Hajdu, "A self-calibrating approach for the segmentation of retinal vessels by template matching and contour reconstruction," *Medical image analysis*, vol. 29, pp. 24–46, 2016.
- [69] N. Brancati, M. Frucci, D. Gragnaniello, and D. Riccio, "Retinal vessels segmentation based on a convolutional neural network," in *Iberoamerican Congress on Pattern Recognition*, 2017, pp. 119–126.
- [70] D. A. Palanivel, S. Natarajan, and S. Gopalakrishnan, "Retinal vessel segmentation using multifractal characterization," *Applied Soft Computing*, vol. 94, p. 106439, 2020.
- [71] S. A. A. Shah, T. B. Tang, I. Faye, and A. Laude, "Blood vessel segmentation in color fundus images based on regional and hessian features," *Graefe's Archive for Clinical and Experimental Ophthalmology*, vol. 255, no. 8, pp. 1525–1533, 2017.
- [72] A. K. Shukla, R. K. Pandey, and R. B. Pachori, "A fractional filter based efficient algorithm for retinal blood vessel segmentation," *Biomedical Signal Processing and Control*, vol. 59, p. 101883, 2020.
- [73] E. Cheng, L. Du, Y. Wu, Y. J. Zhu, V. Megalooikonomou, and H. Ling, "Discriminative vessel segmentation in retinal images by fusing context-aware hybrid features," *Machine vision and applications*, vol. 25, no. 7, pp. 1779–1792, 2014.
- [74] K. Sreejini and V. Govindan, "Improved multiscale matched filter for retina vessel segmentation using pso algorithm," *Egyptian Informatics Journal*, vol. 16, no. 3, pp. 253–260, 2015.
- [75] L. Zhou, Q. Yu, X. Xu, Y. Gu, and J. Yang, "Improving dense conditional random field for retinal vessel segmentation by discriminative feature learning and thin-vessel enhancement," *Computer methods and programs in biomedicine*, vol. 148, pp. 13–25, 2017.
- [76] M. Melinscak, P. Prentasic, and S. Loncaric, "Retinal vessel segmentation using deep neural networks," in *VISAPP (1)*, 2015, pp. 577–582.
- [77] G. Kovács and A. Fazekas, "A new baseline for retinal vessel segmentation: Numerical identification and correction of methodological inconsistencies affecting 100+ papers," *Medical Image Analysis*, vol. 75, p. 102300, 2022.
- [78] R. E. Hacisoftaoğlu, M. Karakaya, and A. B. Sallam, "Deep learning frameworks for diabetic retinopathy detection with smartphone-based retinal imaging systems," *Pattern recognition letters*, vol. 135, pp. 409–417, 2020.
- [79] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," pp. 1597–1607, 2020.