Neural Computing and Applications (2022) 34:3151–3164
https://doi.org/10.1007/s00521-021-06578-3

ORIGINAL ARTICLE

# Hierarchical deep network with uncertainty-aware semi-supervised learning for vessel segmentation

Chenxin Li[1] · Wenao Ma[1] · Liyan Sun[1] · Xinghao Ding[1] · Yue Huang[1] · Guisheng Wang[2] · Yizhou Yu[3]

Received: 29 March 2021 / Accepted: 21 September 2021 / Published online: 12 October 2021


## Abstract

The analysis of organ vessels is essential for computer-aided diagnosis and surgical planning. But it is not an easy task since the fine-detailed connected regions of organ vessel bring a lot of ambiguity in vessel segmentation and sub-type recognition, especially for the low-contrast capillary regions. Furthermore, recent two-staged approaches would accumulate and even amplify these inaccuracies from the first-stage whole vessel segmentation into the second-stage sub-type vessel pixel-wise classification. Moreover, the scarcity of manual annotation in organ vessels poses another challenge. In this paper, to address the above issues, we propose a hierarchical deep network where an attention mechanism localizes the low-contrast capillary regions guided by the whole vessels, and enhance the spatial activation in those areas for the sub-type vessels. In addition, we propose an uncertainty-aware semi-supervised training framework to alleviate the annotation-hungry limitation of deep models. The proposed method achieves the state-of-the-art performance in the benchmarks of both retinal artery/vein segmentation in fundus images and liver portal/hepatic vessel segmentation in CT images. Our implementation is publicly available at https://github.com/XGGNet/Vessel-Seg.

**Keywords** Vessel segmentation · Hierarchical deep network · Attention mechanism · Semi-supervised learning

## 1 Introduction

The analysis and visualization of vessels in human organs are essential for diagnosis and treatment of various diseases. For example, automatic retinal artery and vein (A/V) segmentation assists doctors diagnose systemic and cardiovascular diseases [1, 24]. Specifically, the asymmetrical change of retinal A/V measured by arteriolar-venular ratio (AVR) is closely associated with several cardiovascular diseases [22, 53]. Decrease in AVR is a biomarker with increased risk of stroke [24]. Meanwhile, the narrowing of retinal arteriolar caliber has been reported to be related to the risk of hypertension and diabetes development [5, 38]. Similarly, accurate segmentation and visualization of liver vessels are key prerequisites for safe and efficient surgery in the liver regions [8, 46]. The analysis of hepatic and portal vessels can also provide valuable diagnostic information with regards to chronic liver disease and cirrhosis [41]. In a word, these clinical requirements are summarized as the accurate segmentation of organ vessels as well as their sub-types, e.g., the retinal vessels with arteries and veins in fundus images, and the liver vessels with hepatic and portal sub-types in liver CTs.

Although this field has received considerable research attention, several limitations hinder the routine application of vessel segmentation and sub-type classification in the clinical practice. First, the low contrast between capillaries and background and between different sub-type vessels poses a challenge in segmentation models. An example of retinal fundus image is shown in Fig. 1. Some capillary

✉ Xinghao Ding
dxh@xmu.edu.cn

Guisheng Wang
wanggs1996@tom.com

Yizhou Yu
yizhouy@acm.org

[1] School of Informatics, Xiamen University, Xiamen 361005, China

[2] Department of Radiology, The Third Medical Centre, Chinese PLA General Hospital, Beijing, China

[3] Deepwise AI Laboratory, Beijing 100125, China

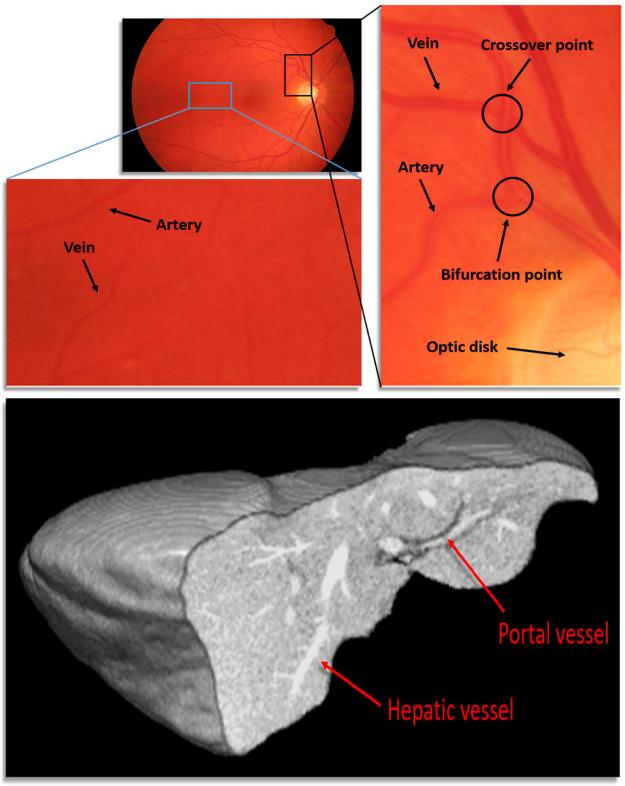


**Fig. 1** An example of 3-channel color retinal fundus image and of 1-channel grayscale anatomical CT volume of a liver. In the retinal image, the regions near and far from the optic disk regions are zoomed in for better visualization. In the liver image, the regions of portal vessels and hepatic vessels are marked

vessels in the areas far away from an optic disk are too indiscernible to detect. The contrast of A/V there also tends to be lower. The situation is similar when observing portal and hepatic vessels in liver anatomical CT images, as shown in Fig. 1. To address this challenge, we think the model should focus more on those low-contrast areas, which are usually capillaries, vessel edges, etc. We discover that the unconfident predictions of whole vessel segmentation are often generated from the low-contrast regions, as shown in Fig. 6. With this location information as a guide, we can spatially activate the feature representation of sub-type vessel segmentation, which enhances the spatial attention of the model on those 'hard' regions for the task of sub-type vessel segmentation.

Besides the above limitation, the previous methods that segment vessels in a two-stage regime, which first segment the whole vessels, then classify the results into sub-types, may propagate and even amplify the first-stage errors mainly caused by the low-contrast regions to the subsequent stage [2, 34, 36]. Instead, the multi-task fashion has attracted long-standing attention and proved to be effective in other medical practices [3, 51]. For example, the task of liver lesion segmentation is usually coupled with the liver segmentation in a multi-task manner. Similar in spirit, we integrate the segmentation of whole and sub-type vessels as a one-stage multi-task framework, to improve the feature sharing between the two tasks, alleviate the accumulated errors and further overcome the challenge caused by low-contrast regions.

Another challenge is the scarcity of the pixel-wise annotation for vessels and their sub-types. Usually, the labeling process for pixel-wise vessels requires radiology expertise and is laborious as well as time-consuming, which causes the acquiring labels for every image in a training dataset to be a burden. Thus, a more practical situation in practical clinical applications is that only some training data are labeled. To address this, we leverage the information from massive unlabeled datasets via semi-supervised learning methods. A common approach is self-labeling, where the model is trained with labeled data by supervised methods and then allocates pseudo-labels for unannotated data. However, the pseudo-generated labels are not always reliable due to the distribution shift between labeled and unlabeled data, which initiates the requirement of filtering unreliable pseudo-labels. Our intuition is that if an unlabeled sample is allocated a more consistent pseudo-label, i.e., maintains the same class, we think its pseudo-label is high-quality and reliable. Thus, we introduce the uncertainty estimation for the generated labels and retain only the ones with enough certainty for the subsequent optimization.

Overall, in this paper, we propose a hierarchical deep network with uncertainty-aware semi-supervised learning for organ vessel segmentation. First of all, we propose the hierarchical capillary-enhanced network for joint whole and sub-type vessel segmentation, where the multi-task design completes both sub-tasks, emphasizes on the capillary-like low-contrast regions based on the guide information from the sub-task of whole vessels segmentation, and regularizes the model to avoid the error propagation occurring in prior two-stage works. Moreover, we develop a uncertainty-aware semi-supervised learning method to efficiently leverage information from massive unlabeled datasets. We allow the model to self-label the unlabeled data and filter the low-quality pseudo-labels. To further model the reliability of the generated labels, we introduce the uncertainty-aware estimation and filter the pseudo-labeled data with high uncertainty.

Our initial work has been published on MICCAI 2019 [34]. We extend the conference version by incorporating a novel semi-supervised learning framework by uncertainty modeling. We also enrich the experimental section by validating the proposed model on liver vessel segmentation.

## 2 Related work

In this section, the previous models for organ vessel and sub-type vessel segmentation are reviewed. The semi-supervised learning methods that have been employed in the case of shortage of labeled data are then reviewed.

### 2.1 Organ vessel segmentation

In graph-based approaches, a vascular graph is built from the extracted vessel centerlines and then each individual graph tree is classified into arteries and veins. The model proposed by Niemeijer et al. [40] first extracts handcrafted features from the vessel centerlines. Then the soft labels are calculated, based on which the classification of vessels can be performed. Dashtbozorg et al. [9] designed a model for retinal A/V segmentation through a pipeline of whole vessel segmentation, centerline extraction and graph reconstruction. Estrada et al. [13] proposed a graph-theoretic framework for A/V segmentation by estimating the vascular topology using a global likelihood model and domain specific features. Xu et al. [55] extracted novel texture features from vessel centerlines and employed KNN as the classifier. In the recent work from Zhao et al. [59], the dominant set theory was used to identify the complicated branches and crossover junctions, building a vascular graph and classifying the centerlines into artery and vein. In terms of liver vessel segmentation, a contextual information extraction was used in a hierarchical voting mechanism in [6] to segment liver vessel. Incorporating knowledge of liver vascular structures, Jerman et al. [23] employed vascular filters based on Hessian filters with a loss constraint ratio to the vessel shape. Merveille et al. [35] utilized morphological filters by ranking the orientation responses of path operators. Lebre et al. [28] further proposed a 'Couinaud' scheme for the portal/hepatic vessel segmentation. Guo et al. [17] modeled a vascular network based on graph cut, thinning and vessel combination, based on which the vessels can be identified. Nevertheless, the above two-step methods suffer from the limitation that the performance of sub-type vessel classification affects the accuracy of the whole vessel segmentation, especially for the graph-based methods. Additionally, feature-engineering methods heavily rely on the fine design of handcrafted feature extraction, which presents difficulty in capturing complex vessel patterns.

With the rapid development of deep neural networks, the number of studies using fully convolutional networks (FCN) to detect and classify organ vessels is increasing. AlBadawi and Fraz [2] adopted the FCN with an encoder-decoder structure for the pixel-wise prediction of retinal arteries and veins. Meyer et al. [36] also used FCN for the same task and reported the performance on major vessels thicker than three pixels. Bruno et al. [4] further introduce the paradigm in the vessel segmentation on fluoroscopy images. Huang et al. [20] utilized FCN 3D-UNet for the vessel segmentation in liver CT images. Deep learning-based methods have demonstrated their potential to segment sub-type vessels in an end-to-end manner. However, these models applied direct segmentation without exploiting contextual image information, thus showing uncertainty in prediction and often classifying pixels on sub-type vessels especially in capillary as background. Thus, there is still room for improvements if the specific model design could be proposed to account for the ambiguity in prediction and incorporating proper contextual information.

### 2.2 Semi-supervised learning

With the aim at exploiting a large number of unlabeled data, plenty of researches raised the focus on semi-supervised learning strategy [18, 21, 31, 43, 52]. Hong et al. [18] proposed decoupled networks including classification and segmentation networks. The class-specific information across layers is manifpulated by the carefully designed module. Hung et al. [21] utilized an adversarial learning mechanism and adopted a logit (probability maps) output from network as a confidence map in semi-supervised learning. Moreover, some approaches designed unsupervised loss by some heuristic regularizations, e.g., minimum entropy, and incroporated it in the overall semi-supervised loss function [27, 50, 57]. Sajjadi et al. [45] proposed a consistency loss on the outputs of the model conditioned on random perturbations of one image to impose transform invariance. In some prior literature, researchers assigned pseudo-labels to the massive unlabeled datasets [29, 47]. Lee et al. [29] utilized a pre-trained network to generate pseudo-labels of unlabeled data by itself, which were then used for fine-tuning the model to achieve better robustness.

The model uncertainty exists in nature for learning systems and can be used to estimate the uncertainty and quality in turn for pseudo-labels. There are two possible sources of uncertainty [10, 25]: epistemic uncertainty and aleatoric uncertainty. The former claims that the inherent uncertainty of model comes from the ignorance during data generation [10]. While the later focuses more on the noise inevitably brought by the observation [25]. In medical imaging practice, uncertainty was used to display probable errors [7] and often aided medical image segmentation [11, 30]. In this paper, we explore the possibility of introducing uncertainty modeling in the semi-supervised learning scheme.

## 3 Method

A data pre-processing pipeline and a hierarchical network are described in this section. Then, the design of spatial activation to enhance low responses in sub-type vessels in capillary regions is detailed. Finally, the uncertainty-based semi-supervised framework built on a probabilistic model is presented. Please note that a retinal fundus image is used to illustrate our approach.

### 3.1 Data pre-processing

We use specific pre-processing on different organs and modalities. For retinal fundus images, we apply a series of pre-processing on raw images shown in the green box in Fig. 2. To correct the bias generated by non-uniform illumination, a illumination correction is applied. Then, a line detector [39] and Gabor filter [48] are utilized on corrected images in parallel for edge extraction. Raw images, illumination-corrected images and two edge maps are further cropped into pairs of patches and are further concatenated in the channel dimension to form an 8-channel input (i.e., 3 channels for a raw/corrected image, 1 channel for either of the two kinds of edge map). In the ablation study, we validate the effectiveness of this proposed multiple-inputs design. In the testing phase, a sliding window strategy is applied on the test image for the patch-wise predictions, as shown in Fig. 2; then, the patches are further merged to reconstruct the image prediction via averaging on logits of overlapping patches as post-processing.

For liver vessel segmentation validated on 3D liver CT scans, we randomly crop volumetric cubes from normalized liver CT scans as model input without other pre-processing operations in the training phase. A similar sliding-window approach is applied in testing phase.

### 3.2 Hierarchical capillary-enhanced network

Here, we present the architecture of our hierarchical capillary-enhanced network in Fig. 3 for joint whole and sub-type vessel segmentation. A U-Net like architecture is adopted as backbone, since U-Net [44] has proven its potency in segmenting medical images accurately and efficiently. Similar to the regular U-Net, we concatenate the feature maps from the previous decoding layer with the corresponding feature maps from the encoding layers to be used as input in each decoding layer.

To predict the whole vessel and sub-type vessels (i.e., arteries/veins in retinal image and portal/hepatic vessels in liver image) simultaneously, we design an ad hoc output block where contextual information from whole vessel segmentation helps guide the sub-type vessel segmentation, as shown in Fig. 3a. The motivation of this design is building a hierarchical model to aid a difficult task (sub-structure segmentation) with information from an easy task (whole vessel segmentation) and avoid accumulative segmentation error.

We apply two different sets of convolution on the input features of the output block to generate two chunks of feature maps. For whole vessel segmentation, one chunk of features is squeezed into a 1-channel representation via $1 \times 1$ convolution + BN; then, a sigmoid function is utilized to produce a logit ranging from 0 to 1, where a binary cross-entropy loss function $\mathcal{L}_v$ is imposed. Therefore, the output block is able to predict whole vessel segmentation while the resulting logit describes the contextual information of the confidence level of the model in making a

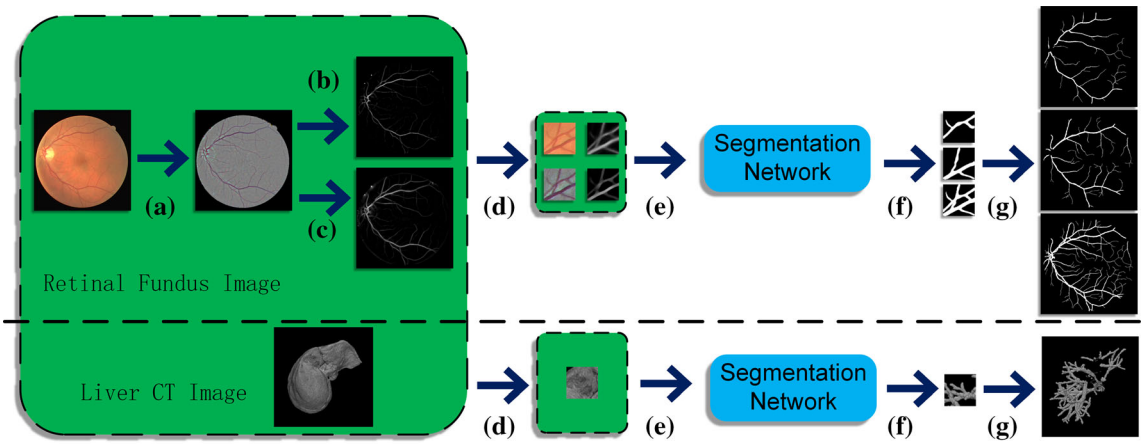


**Fig. 2** The basic pipeline of the segmentation for retinal A/V and liver vessels. A multi-input module is applied for retinal A/V, which leverages the results from three common image pre-processing techniques [34]. **a** Illumination correction; **b** Gabor filtering; **c** Line detector; **d** Patch extraction; **e** Feeding patches into segmentation network; **f** Logits of patch vessel segmentation; **g** Segmentation on whole image

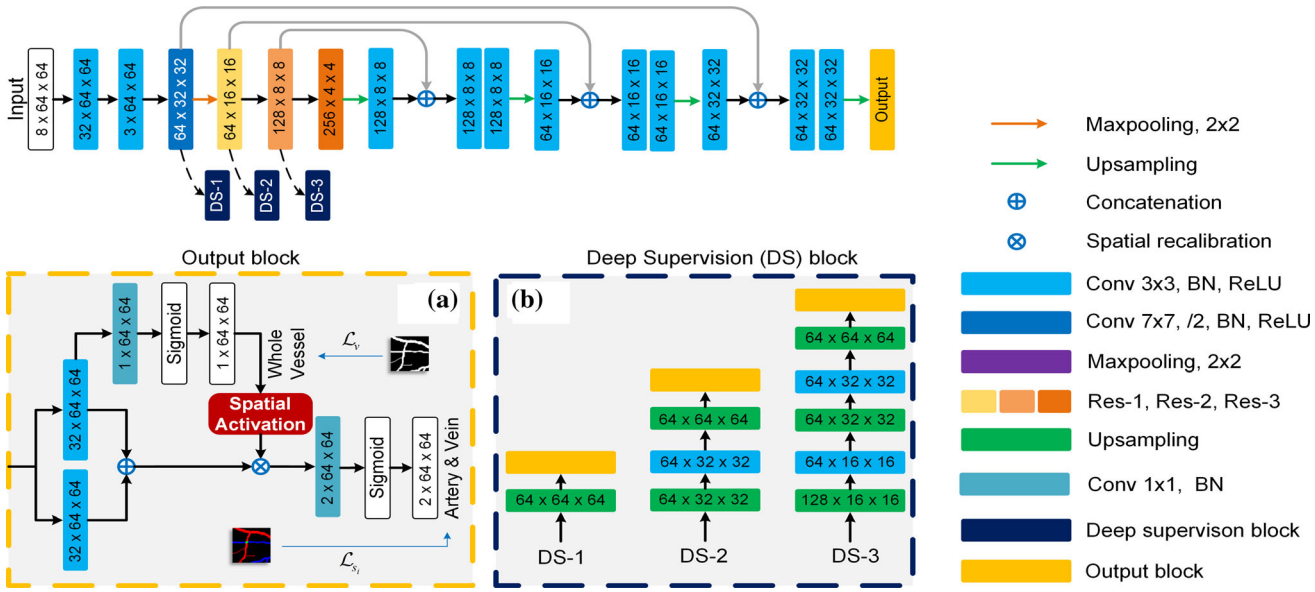


**Fig. 3** The architecture of our proposed hierarchical deep network. **a** Output block; **b** Deep supervision branches. In output block, the branch with one output channel is to generate the segmentation maps of whole vessels and the branch with two output channels is for the segmentation maps of arteries and veins, respectively

prediction about a certain pixel belonging to a vessel instead of background. In this stream, the common characteristics of sub-type vessels are extracted as the whole vessel regions to be distinguished from backgrounds.

As shown in the output block in Fig. 3a, two chunks of features obtained by convolving input features are concatenated together in another stream and weighted by the spatial activation which is discussed in the next subsection. Then, the weighted features are further forwarded into a $1 \times 1$ convolution + BN + Sigmoid layer, squeezing them into a logit of N channels with N denoting total number of sub-type vessel classes. Each channel in the logit represents the likelihood of this sub-type class being presented in the prediction. In this paper, N equals 2 for both retinal and liver vessel segmentation. Then, each channel of the logit is used for computing the binary cross-entropy loss to train the segmentation of sub-type vessels $\mathcal{L}_{s_i}$ where subscription $i$ represents a certain sub-type vessel. Note that the segmentation of sub-type vessels is not a one-hot problem since a pixel is likely to be classified into both artery and vein if it lies in the junction of vessels. Thus, either channel in the two-channel logits represents the possibility of one substructure of vessel. In this stream, the model learns to differentiate sub-type vessels (artery/vein in the retinal image or portal/hepatic vessel in the liver image). To designing such a two-stream multi-task output block, we exploit both the homogenous and heterogenous characteristics of all sub-type vessels.

### 3.3 Spatial activation

Here, a detailed motivation and description of the spatial activation of our capillary-enhanced network are presented. From Fig. 1, we observe the low-contrast intensity of capillary vessels compared with the background, which leads to uncertainty from the segmentation model in classifying them. To validate this hypothesis, we extract the probability heatmap of predicting the whole vessel and visualize it in Figure 4. In Fig. 4a, the golden standard vessel segmentation is shown. In Fig. 4b–d, the probability map is divided into three ranges: [0,0.3], (0.3,0,7] and (0.7,1.0]. The logits of vessel pixels tend to 1, and those of

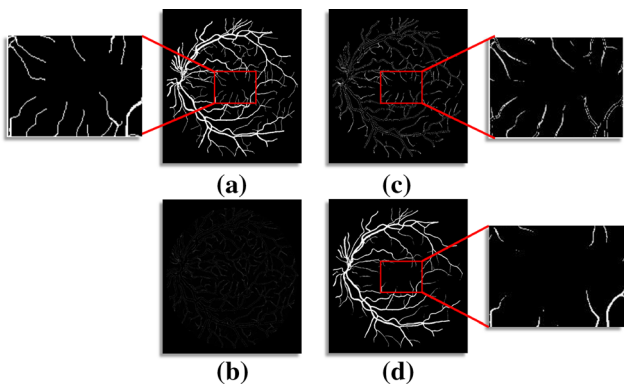


**Fig. 4** The visualization of probability maps of a retinal image in different ranges. **a** An ground truth vessel segmentation from the AV-DRIVE dataset; **b** Probability maps where pixels ranging between [0,0.3]; **c** Probability maps where pixels ranging between (0.3,0.7]; **d** Probability maps where pixels ranging between (0.7,1.0]

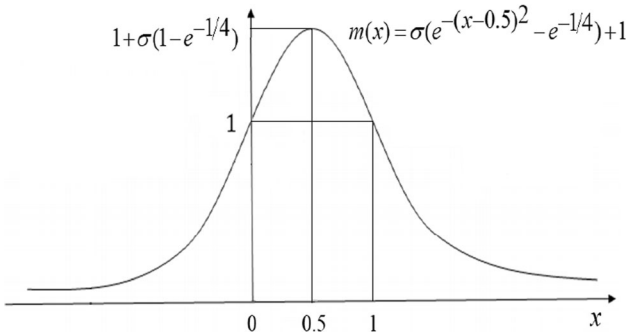


**Fig. 5** The activation function $m(x)$ in the proposed spatial activation module

backgrounds tend to 0. The majority of the pixels corresponding to thick vessels range between (0.7,1.0]. However, the pixels in the areas of capillary vessels range between (0.3,0.7], indicating that the network is highly uncertain in predicting these capillary vessels, decreasing the network performance in sensitivity.

To further utilize the contextual information of vessel segmentation to promote the difficult sub-type vessel segmentation, an ad hoc activation function is proposed, as in Eq. 1:

$$m(x) = \sigma\left(e^{-\left(x-\frac{1}{2}\right)^2} - e^{-\frac{1}{4}}\right) + 1, \tag{1}$$

where $\sigma$ denotes the activation factor and is set as the identity mapping in our design, $x$ denotes the value of each pixel on the probability map with regards to whole vessel segmentation. In Fig. 5, we show the graph of this activation function. It can be seen that as the value of $x$ approaches 0.5, the maximum $1+\sigma(1-e^{-\frac{1}{4}})$ of the function is approximated. As the value of $x$ approaches 0 or 1, then $m(x)$ collapses to the identity mapping. We can pass the logit (probability map) of the whole vessel segmentation to this activation function and receive an attention weighting map. We then apply the attention score map to weight each feature maps from sub-type segmentation via point-wise multiplication.

In such a design, capillary vessels with pixel value around 0.5 will be enhanced with higher weights close to $1+\sigma(1-e^{-\frac{1}{4}})$. Meanwhile, background or thick vessels that are easier to segment will remain unchanged with weights close to 1. An example of the activation map is presented in Fig. 6.

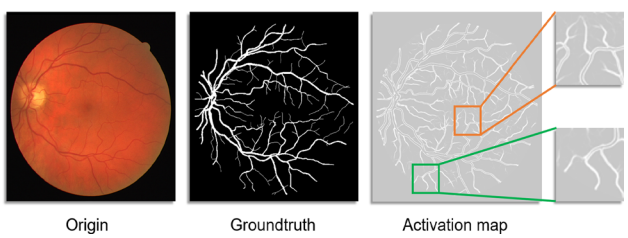


**Fig. 6** An example of the proposed spatial activation. The attention maps are obtained by the activation function $m(x)$ referred in Fig. 5

A similar weighting approach is the focal loss [32] where hard examples are assigned larger weights automatically. In our case, the hard example is the accurate segmentation of the capillary vessels. However, what distinguishes our work from the focal loss is we regularize the hard examples based on another related but easier task, i.e., the vessel segmentation in an hierarchical fashion. Additionally, focal loss is only used during training while our hierarchical spatial activation is applied on both training and testing phase.

### 3.4 Deep supervision

Lacking sufficient amount of labeled images often leads to over-fitting of a model. To deal with this problem, we adopt a deep supervision [12] method, as shown in Fig. 3b. By adding extra branches, the shallow layers can be guided closer to supervision information, which enables low-level layers to extract more semantics [58].

We design the loss function including binary cross-entropy loss of the main branch, deep supervision branches and a weight decay regularization term, as in Eqs. 2 and 3:

$$\begin{aligned}\mathcal{L}(\text{pred}, \text{GT}) = &\text{BCE}(\text{pred}_o, \text{GT}) \\ &+ \frac{1}{3}\sum_{i=1}^{3}\text{BCE}(\text{pred}_i, \text{GT}) \\ &+ \frac{\lambda}{2}||\Theta||_2^2,\end{aligned} \tag{2}$$

$$\begin{aligned}\text{BCE}(\text{pred}, \text{target}) =& \mathcal{L}_v + \mathcal{L}_{v1} + \mathcal{L}_{v2} \\ =& -\sum_{c=1}^{3}\mu_c \cdot \text{target}_c \cdot \log(\text{pred}_c),\end{aligned} \tag{3}$$

where $\Theta$ denotes the network parameters; $\text{pred}_o$ denotes the output of the segmentation network; $i$ denotes the $i$th deep supervision block; $c$ denotes the $c$th class; the weight of each class is denoted as $\mu_c$ with $\frac{3}{7}$, $\frac{2}{7}$ and $\frac{2}{7}$ for vessel, artery and vein in retinal image and portal/hepatic vessel segmentation. The binary cross-entropy is applied on every pixel in cropped patches.

### 3.5 Uncertainty-aware semi-supervised segmentation

Labeling vessel images is laborious and requires expert knowledge. To further leverage massive unlabeled data for better generalization, we propose a semi-supervised learning framework. Generating pseudo-labels for the unlabeled data by an automatic annotation algorithm is a popular approach in semi-supervised learning. However, simply applying a

deterministic model on the unlabeled dataset for pseudo-label acquisition ignores the common distribution shift between labeled data and unlabeled data. Once trained, the model based on labeled data may generate low-quality pseudo-labels for unlabeled data. Another major drawback of this labeling strategy lies in the fact that the annotation is performed by hard thresholding on the logits without handling the uncertainty in making such predictions, which leads to bias in pseudo-labels. To tackle these problems, we improve the pseudo-label generation by introducing the scheme of uncertainty estimation, which has been investigated in medical imaging practice [11, 16, 30, 56]. Inspired by the probabilistic U-Net which attempts to describe the ambious mode for image segmentation [26], we propose an uncertainty-aware architecture based on Bayesian networks [25], which is shown to be more efficient than Monte Carlo dropout [10].

The overview of the training and sampling process of our proposed uncertainty-aware semi-supervised segmentation method is shown in Fig. 7. Besides the proposed hierarchical segmentation network, we further introduce the prior network and the posterior network, to capture the prior and posterior latent representations of the input image, respectively. At a high level, the prior network is trained to embed similar representations to those by the posterior network with a Kullback–Leibler (KL) loss. Then in the sampling process, the latent representations captured by the trained prior network are expected to be close to the posterior representations and carry the label information. Following the spirit of conditional variational autoencoder (CVAE), the latent space is modeled as an N-dimensional axis-aligned Gaussian distribution with parameters as mean $\mu$ and covariance $\sigma^2$.

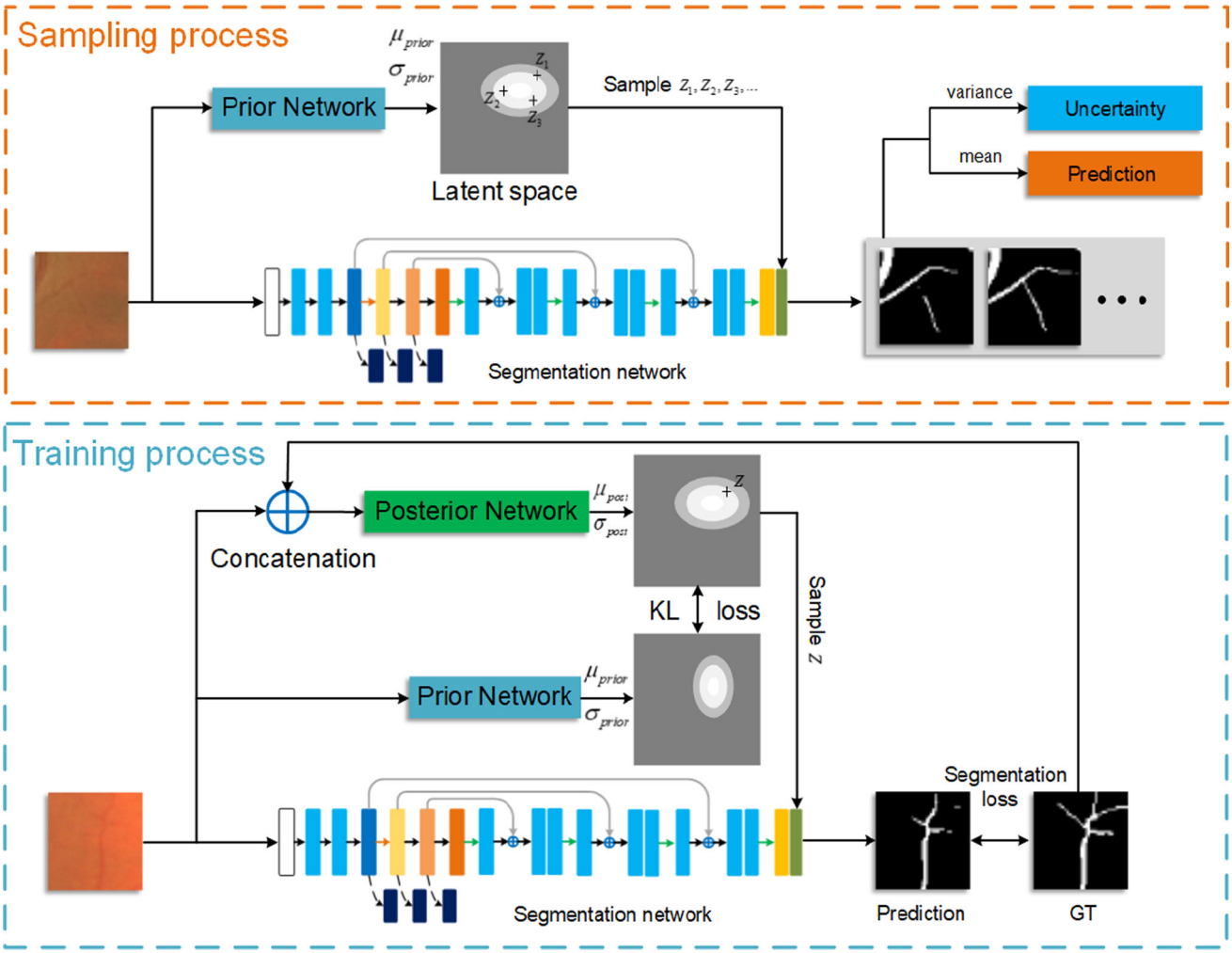


**Fig. 7** The overall architecture of the proposed uncertainty-aware semi-supervised learning framework. The sampling process is designed to generate the pseudo-annotation as well as its uncertainty estimation for an uncertainty-aware self-labeling program. The training process allows the model to capture the semantic-correlated pattern for the input unlabeled image. These two processes are alternative as well as progressive and can completes each other. Besides the proposed hierarchical segmentation network, we introduce a prior network and posterior network to implement the above processes

More precisely, in the training process, we train the prior network with the assist of a posterior network. The posterior network is fed with the input image and its ground truth and it is trained to generate the distribution that carries the information of semantic labels, which we call as posterior distribution. Then, we utilize a Kullback–Leibler (KL) divergence to constrain the prior distribution generated from the prior network to approach the posterior distribution. Once trained, the prior network is regarded to be able to capture a similar pattern to the posterior network, and the prior distribution is approximate to the posterior distribution.

Then in the sampling process, we can draw a sample $z_i$ from the approximate posterior distribution from the trained prior network. We broadcast it to an N-channel feature map and concatenate it with the output of the hierarchical segmentation network. By repeating the distribution sampling $M$ times, $M$ corresponding model outputs can be obtained. Then, the pixel-wise mean values and variance values among them are calculated as the pseudo-annotation and the uncertainty estimation. Please note that for an input image, the output of the segmentor (hierarchical capillary-enhanced network) is invariant during different samples of latent space, so the computation of these forward passes is cheap. Both the prior and posterior nets have the same architecture, which follows the previous experience in the referred paper [26], *P-UNet*. Concretely, the cores of the models feature 4 down-sampling operations, at each scale the blocks comprise three convolutional layers with 3 × 3-kernels, each followed by a ReLU-activation.

With the pseudo-labels with uncertainty estimation, we can improve the hierarchical capillary-enhanced model by retraining only the pseudo-labels with high confidence. Assume that the model uncertainty of input image is $U$, which is expressed as the variance computed by several forward passes of the sampling. The notation $u_v$ represents the uncertainty at the $v$-th pixel in the pseudo-label. The Loss function of the retraining can be formulated in Eq. 4 as:

$$\mathcal{L}_u(\text{pred}, \text{pseudo}) = \frac{\sum_v (\mathbb{I}(u_v < H) \cdot \text{BCE}(\text{pred}, \text{pseudo}))}{\sum_v \mathbb{I}(u_v < H)}, \tag{4}$$

where $\mathbb{I}(\cdot)$ is the indicator function outputting 1 if the condition is satisfied and 0 otherwise; pred and pseudo are the segmentation prediction and the pseudo-label at the $v$-th pixel; $H$ is a threshold to filter out uncertain label empirically set as 0.7. With this uncertainty-aware loss function, the segmentation network can learn from the reliable pseudo-labels as well as avoid the possible misleading from unreliable ones. The overall algorithm for the proposed semi-supervised learning framework is list in Algorithm 1.

**Algorithm 1** Overall flow of the proposed semi-supervised learning framework

**Input:** labeled dataset $\mathcal{D}_l$, unlabeled dataset $\mathcal{D}_u$; pretraining iterations $T_{pre}$, retraining iterations $T_{re}$; initialized parameter of segmentation network $\theta$, initialized parameter of prior and posterior networks $\theta'$.

**Output:** trained parameter $\theta$ of segmentation network.

1: **for** $loop = 1$ to $T_{pre}$ **do**
2: Optimize $\theta$ and $\theta'$ on labeled dataset $\mathcal{D}_l$ by the training process in Figure 7.
3: **end for**
4: **for** $loop = 1$ to $T_{re}$ **do**
5: Take multiple forward passes of pretrained model to generate outputs on unlabeled dataset $\mathcal{D}_u$ by the sampling process in Figure 7.
6: Calculate the mean values of the multiple outputs as pseudo annotation.
7: Calculate the variance values of the multiple outputs as uncertainty estimation.
8: Optimize $\theta$ on dataset $\mathcal{D}_u$ with the pseudo annotation and uncertainty estimation using Equation 4;
9: **end for**

## 4 Experimental results

### 4.1 Dataset

For retinal artery/vein segmentation, we trained and evaluated our model on the publicly available AV-DRIVE database [19]. The AV-DRIVE database contains 20 training and 20 testing colorful retinal fundus images with dimension of 584 × 565 pixels, with pixel-wise labeling of vessel segmentation and A/V classification provided. To validate our semi-supervised learning approach, we pretrained segmentation model on AV-DRIVE dataset and generate pseudo-annotations on HRF [42] dataset which contains 45 images of the dimension 800 × 1200. We further retrained our model on the combination of AV-DRIVE and HRF datasets and tested it on the INSPIRE-AVR dataset containing 40 fundus images with dimension of 2048 × 2392. Besides, with only the centerline-annotated labels available for INSPIRE-AVR, we evaluated the model performance only on the annotated centerline pixels.

Some examples of the public database can be seen in Fig. 8.

For the liver portal/hepatic vessel segmentation, we adopted the IRCAD datasets containing 3D CT-scans of 20

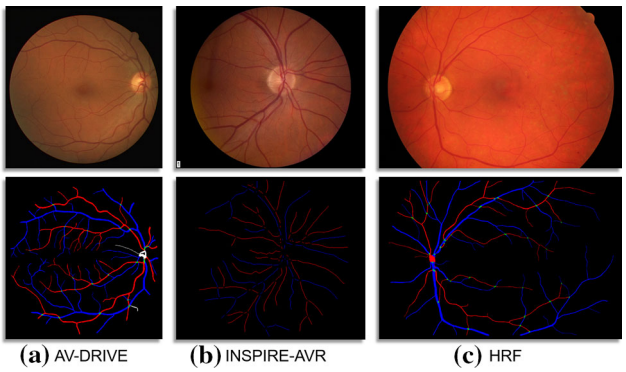


**Fig. 8** Some examples of the public database (first row) and their corresponding manual segmentation (second row). Please note, there are only centerline-annotated labels in INSPIRE-AVR, which is different from AV-DRIVE and HRF

different imaging subjects. All the 20 scans are labeled with the pixels of portal and hepatic vessel. We conducted K-fold cross-evaluation and report the mean results of them following prior study [17, 20, 23, 28, 35].

### 4.2 Implementation details

In the training stage, the stochastic gradient descent algorithm with momentum was adopted to optimize our model. The training iterations for all our experiments were 60,000. The initial learning rate was set as 0.05 and halved every 5000 iterations. The size of the training mini-batch was 16. Multi-scale patches with size of $64 \times 64, 96 \times 96, 128 \times 128$ were randomly cropped from the retinal images during training and were further resized to $64 \times 64$ to exploit multi-scale information. To alleviate the overfitting risk, we utilized some common data augmentation strategies as random flip and scaling. In the testing stage, patches with size of $64 \times 64$ were extracted at the stride of 10. The segmentation of sub-type vessels and full vessel of a image were obtained by aggregating the patches, with the overlapping regions being averaged as post-processing. The experiments were conducted on a single NVIDIA GPU (Tesla P40) using Pytorch platform. As for the computation time, the training process took about 3 h. It took about 1.5 s to perform segmentation on an image in the testing stage. All the results are obtained as the average over 5 random runs.

### 4.3 Evaluation metrics

We adopt common evaluation metrics: Accuracy(Acc), Sensitivity(Sen), Specificity(Sp), Area Under Curve(AUC), F1 score (F1). The Accuracy, Sensitivity and Specificity are defined as

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \tag{5}$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{6}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{FP} + \text{TN}}, \tag{7}$$

The metric AUC is defined as the receiver operating characteristics (ROC) area under the curve, which represents the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one.

The F1 score is defined as

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \tag{8}$$

where $\text{Precision} = \frac{\text{TP}}{\text{FP}+\text{FP}}$, $\text{Recall} = \frac{\text{TP}}{\text{FP}+\text{FN}}$

### 4.4 Ablation studies

On the retinal fundus image dataset AV-DRIVE, three groups of ablation studies are performed to evaluate the contribution of each design in our approach: (1) Joint prediction of vessel segmentation and artery/vein segmentation in a multi-task manner (denoted as MTs); (2) Multiple inputs using pre-processed images (denoted as MIs); (3) Spatial activation (denoted as AC). Baseline is built by removing the MTs, MIs and AC designs.

As indicated in Table 1, a multi-task way of performing vessel segmentation and sub-type segmentation simultaneously can improve both tasks. Also, by inputting multiple pre-processed images into the model, the A/V classification accuracy is improved by 0.3%. When the spatial activation is applied, the A/V segmentation can be further increased by 0.6%. Our proposed hierarchical capillary-enhanced network achieves a pixel-wise accuracy of 95.70% for vessel segmentation and 92.58% for sub-type vessel segmentation. We also compare our spatial activation with the design of focal loss. We observe no significant improvement when removing the activation block and equipping the cross-entropy loss with focal loss. In Fig. 9, an example is shown to compare our model with baseline. We zoom in multiple local areas, and our proposed method has remarkably improved the segmentation result of retinal artery and vein over the baseline by the MTs, MIs and AC design.

To evaluate the semi-supervised learning framework, a segmentation model is pretrained using the labeled AV-DRIVE dataset and generates pseudo-labels on the unlabeled HRF dataset. The evaluation is conducted on the INSPIRE database. We first build the lower bound of semi-supervised learning as applying supervised learning scheme on only labeled AV-DRIVE, which is denoted as

**Table 1** The ablation study results of vessel segmentation and A/V classification (%)

| Combination | | | Vessel Segmentation | | | | | A/V Classification | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| MTs | MIs | AC | Acc | Sen | Sp | AUC | F1 | Acc | Sen | Sp | F1 |
| | | | 94.98 | 68.86 | **98.79** | 97.60 | 93.86 | 91.25 | 89.68 | 92.55 | 90.30 |
| ✓ | | | 95.61 | 78.50 | 98.10 | 98.01 | 94.72 | 91.63 | 90.46 | 92.63 | 90.65 |
| ✓ | ✓ | | 95.66 | 78.30 | 98.19 | 98.08 | 94.91 | 91.98 | 90.36 | **93.42** | 91.01 |
| ✓ | ✓ | ✓ | **95.70** | **79.16** | 98.11 | **98.10** | **95.15** | **92.58** | **92.18** | 92.98 | **91.98** |

Bold indicates best results

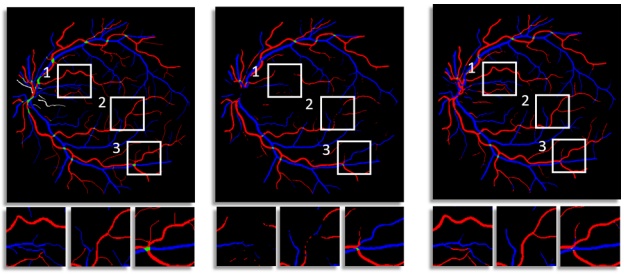


**Fig. 9** Comparison of model performance for baseline and proposed method. **a** Ground truth; **b** Baseline model performance; **c** Proposed method

**Supervised on only labeled**. We also build the upper bound as applying supervised learning scheme on all data, including both labeled AV-DRIVE and HRF, which is denoted as **Supervised on all**. For self-label semi-supervised setting, we first evaluate the benefits of additional datasets by using a trained segmentor on AV-DRIVE to directly annotate images on the HRF dataset without uncertainty modeling, denoted as **Vanilla Annotation**. In order to validate our uncertainty estimation approach, we conduct a comparative experiment to replace it with the Monte Carlo Dropout [15], denoted as **MC Dropout**.

The comparative experimental results on A/V segmentation can be seen in Table 2. Compared to lower bound model, Vanilla annotation in such a semi-supervised framework brings little improvements. This is related to the unreliable pseudo-labels, which may mislead the model. Experimentally, we count the error rate of pseudo-labels generated by the models on HRF, and we find that about 23% pseudo-labels are inconsistent with the ground truth ones. Besides, when applying our uncertainty estimation model, semi-supervised method produces more accurate segmentation than the one using MC Dropout. This result is also approaching the semi-supervised upper bound.

We also conduct the ablation experiments on the uncertainty threshold H in Eq. 4, as shown in Fig. 10. We observe the Acc performance achieved by our model with respect to H ranging from 0.1 to 0.9. At last, we find H=0.7 corresponds to the best Acc performance, which is 90.1%.

### 4.5 Comparison to existing methods

We compare our hierarchical capillary-enhanced network with other state-of-the-art approaches in the performances of whole retinal fundus vessel and artery/vein segmentation, as shown in Tables 3 and 4.

In Table 3, we compare the capillary-enhanced network with other state-of-the-art models for whole retinal vessel segmentation on the testing dataset in AV-DRIVE. Our approach achieves the better performance among the compared models, owing to the joint optimization for both whole and sub-type vessel segmentation.

In Table 4, we compare our capillary-enhanced network with other state-of-the-art models for retinal arteries/veins segmentation on the testing dataset in both AV-DRIVE and INSPIRE datasets. When performing under a fair comparison, the proposed model achieves a pixel-wise accuracy of 94.50%, which outperforms the current state-of-the-art A/V segmentation method by a considerable margin.

We also have validated our algorithm on liver vessel segmentation in IRCAD [49]. Because 3D liver computed tomography (CT) images are provided in this dataset, we adopt a 3D U-Net, incorporating the proposed designed

**Table 2** Performance comparison of A/V segmentation of semi-supervised learning

| Methods | Note | Acc (%) | Sp (%) | Sen (%) | F1 (%) |
|---|---|---|---|---|---|
| Supervised on only labeled | Lower bound | 84.7 | 92.5 | 77.1 | 83.6 |
| Vanilla Annotation | | 85.4 | **95.8** | 64.5 | 84.7 |
| MC Dropout | Semi-supervised | 88.3 | 88.9 | 88.2 | 87.8 |
| **Proposed** | | **90.1** | 88.2 | **92.4** | **89.5** |
| Supervised on all | Upper bound | 91.6 | 92.4 | 91.0 | 90.8 |

Bold indicates best results

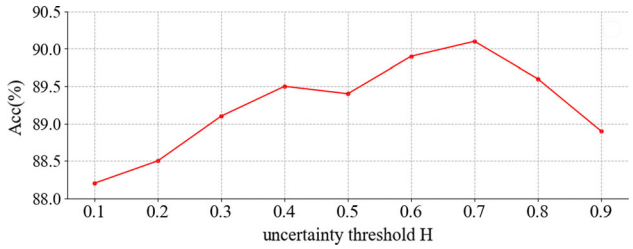


**Fig. 10** Ablation study on uncertainty threshold H

**Table 3** Performance comparison of whole vessel segmentation on the AV-DRIVE dataset

| Methods | Acc (%) | Sp (%) | Sen (%) | AUC (%) |
|---|---|---|---|---|
| Fu et al. [14] | 94.70 | – | 72.94 | – |
| Liskowski et al. [33] | 95.35 | 98.07 | 78.11 | 97.90 |
| Mo et al. [37] | 95.21 | 97.80 | 77.79 | 97.82 |
| Wu et al. [54] | 95.67 | **98.19** | 78.44 | 98.07 |
| **Proposed** | **95.70** | 98.11 | **79.16** | **98.10** |

Bold indicates best results

spatial activation and removing our multi-input module. In this database, we use our model to segment the portal and hepatic vessels in a multi-task manner. Then, we combine them to form the segmentation of whole veins in a liver. The 3D reconstructions of the segmented portal and hepatic vessels are shown in Fig. 11. We have compared our model with extensive prior approaches, as shown in Table 5. Note that we only report performances on whole vessel segmentation of some compared methods [17, 20, 23], because they were proposed for targeting on whole vessel regions. At last, the proposed method achieves superior performance against other state-of-the-art methods. The common improvements in the joint task and the two sub-tasks also demonstrate the extensive effectiveness of the proposed multi-task mechanism

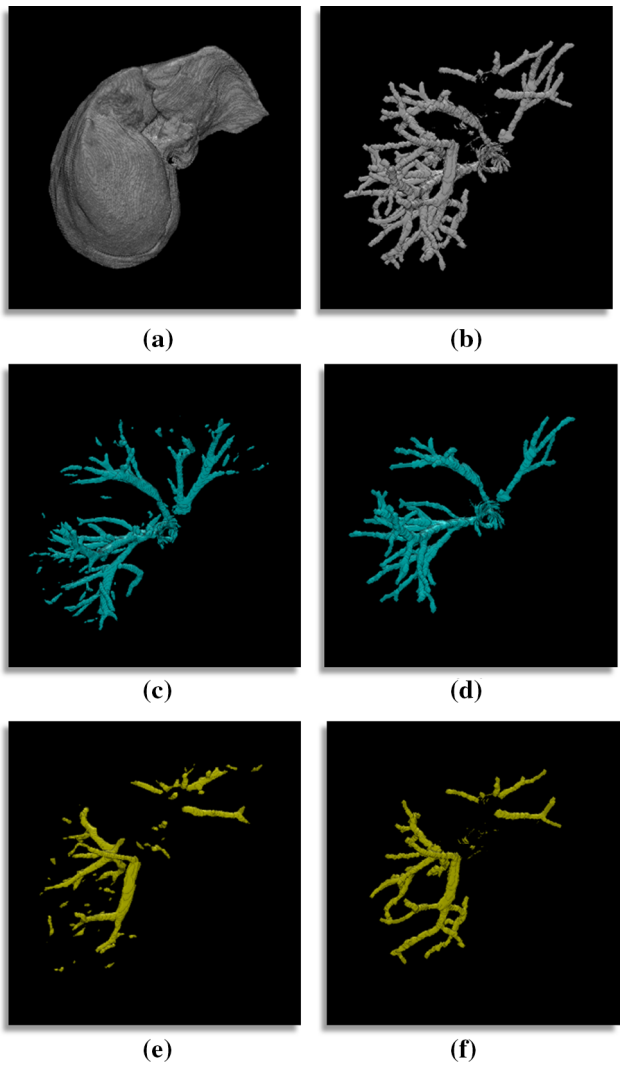


**Fig. 11** The visualization of segmentation of portal vessel and hepatic vessel. **a** CT liver volume from IRCAD dataset; **b** manual label of portal and hepatic vessel; **c** segmentation prediction of hepatic vessel; **d** manual label of hepatic vessel; **e** segmentation prediction of portal vessel; **f** manual label of portal vessel

**Table 4** Performance comparison of A/V classification on AV-DRIVE and INSPIRE datasets

| Methods | AV-DRIVE | | | INSPIRE | | |
|---|---|---|---|---|---|---|
| | Acc (%) | Sen (%) | Sp (%) | Acc (%) | Sen (%) | Sp (%) |
| Dashtbozorg et al. [9] | 87.4 | 90.0 | 84.0 | 84.9 | 91.0 | 86.0 |
| Estrada et al. [13] | 93.5 | 93.0 | 94.1 | 90.9 | 91.5 | 90.2 |
| Xu et al. [55] | 92.3 | 92.9 | 91.5 | – | – | – |
| Zhao et al. [59] | – | 91.9 | 91.5 | 91.0 | 91.8 | 90.2 |
| **Proposed** | **94.5** | **93.4** | **95.5** | **91.6** | **92.4** | **91.3** |

Bold indicates best results

**Table 5** Performance comparison of hepatic and portal vessel segmentation on the IRCAD dataset

| Methods | Vessel Segmentation | | | Portal Vessel Segmentation | | | Hepatic Vessel Segmentation | | |
|---|---|---|---|---|---|---|---|---|---|
| | Acc (%) | Sen (%) | Sp (%) | Acc (%) | Sen (%) | Sp (%) | Acc (%) | Sen (%) | Sp (%) |
| Merveille et al. [35] | 90.0 | 20.0 | 97.0 | 98.0 | 56.0 | **99.0** | 96.0 | 23.0 | 98.0 |
| Lebre et al. [28] | 97.0 | 69.0 | 98.0 | 97.0 | **70.0** | 98.0 | 98.0 | 54.0 | 98.0 |
| Huang et al. [20] | 97.1 | 74.3 | 98.3 | – | – | – | – | – | – |
| Jerman et al. [23] | 97.2 | **76.7** | 96.5 | – | – | – | – | – | – |
| Guo et al. [17] | 97.8 | 66.2 | **98.7** | – | – | – | – | – | – |
| **Proposed** | **97.9** | 76.4 | 98.6 | **98.9** | 61.5 | 98.3 | **99.7** | **62.4** | **99.4** |

Bold indicates best results

## 5 Conclusion

In this paper, we build the learning model that faces several serious practical challenges in organ vessel segmentation. First of all, we propose the hierarchical capillary-enhanced network to enhance the performance of vessel segmentation, especially for capillary regions. The proposed method enables the end-to-end segmentation of whole vessels and their sub-types simultaneously in a multi-task manner and allows the sub-tasks can benefit from the joint task. Moreover, we develop an uncertainty-aware semi-supervised learning framework to alleviate the practical issue of annotation scarcity. We evaluate the proposed overall method on some publicly available databases and compare it to the prior methods extensively in both the retinal fundus datasets and liver CT set. The experimental results show that our hierarchical capillary-enhanced model outperforms the state-of-the-art supervised methods under the fully annotated setting and other semi-supervised methods under the annotation-scarcity setting, which demonstrates the proposed model is effective extensively in the medical practice.

To summary, the proposed method significantly improves the accuracy and efficiency of A/V segmentation as well as alleviates the rely upon vessel annotation, which are both key requirements in real-world clinical applications. In the future work, we will go on working based on this research and develop a fully automatic and annotation-efficient machine for vascular parameter modeling to facilitate the clinical vascular biomarker study.

**Acknowledgements** This work is supported in part by National Key Research and Development Program of China (No. 2019YFC0118100), in part of ZheJiang Province Key Research Development Program (No. 2020-C03073), in part by National Natural Science Foundation of China under Grants 81671766, 61971369, U19B2031, U1605252, 61671309, in part by Open Fund of Science and Technology on Automatic Target Recognition Laboratory 6142503190202, in part by Fundamental Research Funds for the Central Universities 20720180059, 20720190116, 20720200003, and in part by Tencent Open Fund.

## Declarations

**Conflict of interest** Authors declare that there is no conflict of interests regarding the publication of this paper.

## References

1. Abràmoff MD, Garvin MK, Sonka M (2010) Retinal imaging and image analysis. IEEE Rev Biomed Eng 3:169–208
2. AlBadawi S, Fraz MM (2018) Arterioles and venules classification in retinal images using fully convolutional deep neural network. In: Int. Conf. on Img. Anal. and Recog., Springer, pp. 659–668
3. Bi L, Kim J, Kumar A, Feng D (2017) Automatic liver lesion detection using cascaded deep residual networks. arXiv preprint arXiv:1704.02703
4. Bruno P, Zaffino P, Scaramuzzino S, De Rosa S, Indolfi C, Calimeri F, Spadea MF (2018) Using cnns for designing and implementing an automatic vascular segmentation method of biomedical images. In: International Conference of the Italian Association for Artificial Intelligence. pp. 60–70. Springer
5. Chew SK, Xie J, Wang JJ (2012) Retinal arteriolar diameter and the prevalence and incidence of hypertension: a systematic review and meta-analysis of their association. Curr Hypertens Rep 14(2):144–151
6. Chi Y, Liu J, Venkatesh SK, Huang S, Zhou J, Tian Q, Nowinski WL (2010) Segmentation of liver vasculature from contrast enhanced ct images using context-based voting. IEEE Trans Biomed Eng 58(8):2144–2153
7. Ching T, Himmelstein DS, Beaulieu-Jones BK, Kalinin AA, Do BT, Way GP, Ferrero E, Agapow PM, Zietz M, Hoffman MM et al (2018) Opportunities and obstacles for deep learning in biology and medicine. J R Soc Interface 15(141):20170387
8. Clavien PA, Petrowsky H, DeOliveira ML, Graf R (2007) Strategies for safer liver surgery and partial liver transplantation. New Engl J Med 356(15):1545–1559
9. Dashtbozorg B, Mendonça AM, Campilho A (2014) An automatic graph-based approach for artery/vein classification in retinal images. IEEE Trans Image Process 23(3):1073–1083
10. Der Kiureghian A, Ditlevsen O (2009) Aleatory or epistemic? does it matter? Struct Saf 31(2):105–112
11. DeVries T, Taylor GW (2018) Leveraging uncertainty estimates for predicting segmentation quality. arXiv preprint arXiv:1807.00502
12. Dou Q, Chen H, Jin Y, Yu L, Qin J, Heng P (2016) 3d deeply supervised network for automatic liver segmentation from CT volumes pp. 149–157

13. Estrada R, Allingham MJ, Mettu PS, Cousins SW, Tomasi C, Farsiu S (2015) Retinal artery-vein classification via topology estimation. IEEE Trans Med Imaging 34(12):2518–2534
14. Fu H, Xu Y, Wong DWK, Liu J (2016) Retinal vessel segmentation via deep learning network and fully-connected conditional random fields. In: 2016 IEEE 13th international symposium on biomedical imaging, IEEE, pp. 698–701
15. Gal Y, Ghahramani Z (2016) Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In: International Conference on Machine Learning, pp. 1050–1059
16. Galdran A, Meyer M, Costa P, Campilho A, et al. (2019) Uncertainty-aware artery/vein classification on retinal images. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), IEEE, pp. 556–560
17. Guo X, Xiao R, Zhang T, Chen C, Wang J, Wang Z (2020) A novel method to model hepatic vascular network using vessel segmentation, thinning, and completion. Med Biol Eng Comput pp. 1–16
18. Hong S, Noh H, Han B (2015) Decoupled deep neural network for semi-supervised semantic segmentation. Adv Neural Inf Process Syst pp. 1495–1503
19. Hu Q, Abramoff MD, Garvin MK (2013) Automated separation of binary overlapping trees in low-contrast color retinal images. In: MICCAI, pp. 436–443
20. Huang Q, Sun J, Ding H, Wang X, Wang G (2018) Robust liver vessel extraction using 3d u-net with variant dice loss function. Comput Biol Med 101:153–162
21. Hung WC, Tsai YH, Liou YT, Lin YY, Yang MH (2018) Adversarial learning for semi-supervised semantic segmentation. arXiv preprint arXiv:1802.07934
22. Ikram MK, de Jong FJ, Vingerling JR, Witteman JC, Hofman A, Breteler MM, de Jong PT (2004) Are retinal arteriolar or venular diameters associated with markers for cardiovascular disorders? The Rotterdam Study. Investig Ophthamol Vis Sci 45(7):2129–2134
23. Jerman T, Pernuš F, Likar B, Špiclin Ž (2016) Enhancement of vascular structures in 3d and 2d angiographic images. IEEE Trans Med Imaging 35(9):2107–2118
24. Kawasaki R, Xie J, Cheung N et al (2012) Retinal microvascular signs and risk of stroke: the Multi-Ethnic Study of Atherosclerosis (MESA). Stroke 43(12):3245–3251
25. Kendall A, Gal Y (2017) What uncertainties do we need in bayesian deep learning for computer vision?. Adv Neural Inf Process Syst pp. 5574–5584
26. Kohl S, Romeraparedes B, Meyer C, De Fauw J, Ledsam JR, Maierhein KH, Eslami SMA, Rezende DJ, Ronneberger O (2018) A probabilistic U-Net for segmentation of ambiguous images pp. 6965–6975
27. Laine S, Aila T (2016) Temporal ensembling for semi-supervised learning. arXiv preprint arXiv:1610.02242
28. Lebre MA, Vacavant A, Grand-Brochier M, Rositi H, Abergel A, Chabrot P, Magnin B (2019) Automatic segmentation methods for liver and hepatic vessels from CT and MRI volumes, applied to the Couinaud scheme. Comput Biol Med 110:42–51
29. Lee DH (2013) Pseudo-label : the simple and efficient semi-supervised learning method for deep neural networks. In: ICMLW, pp. 2,3,7
30. Leibig C, Allken V, Ayhan MS, Berens P, Wahl S (2017) Leveraging uncertainty information from deep neural networks for disease detection. Sci Rep 7(1):1–14
31. Li Q, Arnab A, Torr PH (2018) Weakly-and semi-supervised panoptic segmentation. In: Proceedings of the European Conference on Computer Vision, pp. 102–118
32. Lin TY, Goyal P, Girshick R, He K, Dollar P (2017) Focal loss for dense object detection. In: ICCV. pp. 2980–2988
33. Liskowski P, Krawiec K (2016) Segmenting retinal blood vessels with deep neural networks. IEEE Trans Med Imaging 35(11):2369–2380
34. Ma W, Yu S, Ma K, Wang J, Ding X, Zheng Y (2019) Multi-task neural networks with spatial activation for retinal vessel segmentation and artery/vein classification. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, pp. 769–778
35. Merveille O, Talbot H, Najman L, Passat N (2017) Curvilinear structure analysis by ranking the orientation responses of path operators. IEEE Trans Pattern Anal Mach Intell 40(2):304–317
36. Meyer MI, Galdran A, Costa P, Mendonça AM, Campilho A (2018) Deep convolutional artery/vein classification of retinal vessels. In: Int. Conf. on Img. Anal. and Recog, pp. 622–630
37. Mo J, Zhang L (2017) Multi-level deep supervised networks for retinal vessel segmentation. Int J Comput Assist Radiol Surg 12(12):2181–2193
38. Nguyen TT, Wang JJ, Islam FA, Mitchell P, Tapp RJ, Zimmet PZ, Simpson R, Shaw J, Wong TY (2008) Retinal arteriolar narrowing predicts incidence of diabetes: the australian diabetes, obesity and lifestyle (AusDiab) study. Diabetes 57(3):536–539
39. Nguyen UT, Bhuiyan A, Park LA, Ramamohanarao K (2013) An effective retinal blood vessel segmentation method using multi-scale line detection. Pattern Recognit 46(3):703–715
40. Niemeijer M, van Ginneken B, Abràmoff M (2009) Automatic classification of retinal vessels into arteries and veins. In: Med. Img. 2009: Comp-Aided Diag 7260: 72601F
41. O'Donohue J, Ng C, Catnach S, Farrant P, Williams R (2004) Diagnostic value of doppler assessment of the hepatic and portal vessels and ultrasound of the spleen in liver disease. Eur J Gastroenterol Hepatol 16(2):147–155
42. Odstrcilik J, Kolar R, Budai A, Hornegger J, Jan J, Gazarek J, Kubena T, Cernosek P, Svoboda O, Angelopoulou E (2013) Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database. IET Image Process 7(4):373–383
43. Papandreou G, Chen LC, Murphy KP, Yuille AL (2015) Weakly-and semi-supervised learning of a deep convolutional network for semantic image segmentation. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1742–1750
44. Ronneberger O, Fischer P, Brox T (2015) U-Net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-assisted Intervention, Springer, pp. 234–241
45. Sajjadi M, Javanmardi M, Tasdizen T (2016) Mutual exclusivity loss for semi-supervised deep learning. In: 2016 IEEE International Conference on Image Processing
46. Selle D, Preim B, Schenk A, Peitgen HO (2002) Analysis of vasculature for liver surgical planning. IEEE Trans Med Imaging 21(11):1344–1357
47. Shi W, Gong Y, Ding C, Ma Z, Tao X, Zheng N (2018) Transductive semi-supervised deep learning using min-max features. In: European Conference on Computer Vision
48. Soares JV, Leandro JJ, Cesar RM, Jelinek HF, Cree MJ (2006) Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification. IEEE Trans Med Imaging 25(9):1214–1222
49. Soler L, Hostettler A, Agnus V, Charnoz A, Fasquel J, Moreau J, Osswald A, Bouhadjar M, Marescaux J (2010) 3d image reconstruction for comparison of algorithm database: a patient specific anatomical and medical image database. Tech, Rep, IRCAD, Strasbourg, France
50. Tarvainen A, Valpola H (2017) Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. Adv Neural Inf Process Syst pp. 1195–1204

51. Vorontsov E, Tang A, Pal C, Kadoury S (2018) Liver lesion segmentation informed by joint liver segmentation. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), IEEE, pp. 1332–1335
52. Wei Y, Xiao H, Shi H, Jie Z, Feng J, Huang TS (2018) Revisiting dilated convolution: A simple approach for weakly-and semi-supervised semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7268–7277
53. Wong TY, Klein R, Sharrett AR, Duncan BB, Couper DJ, Tielsch JM, Klein BE, Hubbard LD (2002) Retinal arteriolar narrowing and risk of coronary heart disease in men and women: the atherosclerosis risk in communities study. JAMA 287(9):1153–1159
54. Wu Y, Xia Y, Song Y, Zhang Y, Cai W (2018) Multiscale network followed network model for retinal vessel segmentation. In: MICCAI, pp. 119–126
55. Xu X, Ding W, Abràmoff MD, Cao R (2017) An improved arteriovenous classification method for the early diagnostics of various diseases in retinal image. Comp. Methods Prog. Biomed. 141:3–9
56. Yu , Wang S, Li X, Fu CW, Heng PA (2019) Uncertainty-aware self-ensembling model for semi-supervised 3d left atrium segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, pp. 605–613
57. Yves Grandvalet YB (2005) Semi-supervised learning by entropy minimization. Adv Neural Inf Process Syst 17
58. Zhang Z, Zhang X, Peng C, Xue X, Sun J (2018) Exfuse: Enhancing feature fusion for semantic segmentation. In: ECCV, pp. 269–284
59. Zhao Y, Xie J, Su P, Zheng Y, Liu Y, Cheng J, Liu J (2018) Retinal artery and vein classification via dominant sets clustering-based vascular topology estimation. In: MICCAI, pp. 56–64