Pixel-wise Divide and Conquer for Federated Vessel Segmentation

Tian Chen*, **Wenke Huang***, **Zhihao Wang***, **Zekun Shi**, **He Li**, **Wenhui Dong**, **Mang Ye**, **Bo Du** and **Yongchao Xu**†

School of Computer Science, Wuhan University

{tian.chen, yongchao.xu}@whu.edu.cn

Abstract

Accurate vessel segmentation is essential for diagnosing and managing vascular and ophthalmic diseases. Traditional learning-based vessel segmentation methods heavily rely on high-quality, pixel-level annotated datasets. However, segmentation performance suffers significantly when applied in federated learning settings due to vessel morphology inconsistency and vessel-background imbalance. The former limits the ability of models to capture fine-grained vessels, while the latter overemphasizes background pixels and biases the model towards them. To address these challenges, we propose a novel method named Federated Vessel-Aware Calibration (*FVAC*), which leverages global uncertainty to provide differentiated guidance for clients, focusing on pixels of various morphologies that are difficult to distinguish. Furthermore, we introduce a foreground-background decoupling alignment strategy that utilizes more stable and balanced global features to mitigate semantic drift caused by vessel-background imbalance in local clients. Comprehensive experiments confirm the effectiveness of our method.

1 Introduction

Vessel segmentation plays a crucial role in the image registration of blood vessels and in the diagnosis and management of vascular and ophthalmic diseases, such as coronary artery disease, diabetic retinopathy, and glaucoma [Liu *et al.*, 2024]. Advances in deep learning, particularly through U-Net and its variants [Ronneberger *et al.*, 2015; Liu *et al.*, 2022], have significantly improved segmentation accuracy and robustness. However, progress in vessel segmentation is hindered by the lack of large-scale, high-quality, pixel-level annotated datasets, which are expensive and time-consuming to construct. In addition, data collected from different institutions often vary in imaging modalities, acquisition protocols, and labeling standards, resulting in heterogeneous vessel image distribution. Notably, privacy regulations further restrict multi-center data sharing, emphasizing the need for distributed approaches that can preserve privacy while leveraging diverse and heterogeneous datasets. Recently, Federated Learning (FL) has provided an ideal solution for vessel segmentation by addressing data privacy concerns and effectively handling data heterogeneity [McMahan *et al.*, 2016; Hyeon-Woo *et al.*, 2022; Elgabli *et al.*, 2022; Yang *et al.*, 2023a; Huang *et al.*, 2024b; Wang *et al.*, 2025]. By enabling collaborative training without requiring the exchange of raw data, FL offers a reliable framework to advance multiparty vessel segmentation collaboration, enhancing model generalization and scalability across diverse clinical environments.

Federated vessel segmentation faces greater performance challenges than general federated medical segmentation due to **vessel morphology inconsistency** and **vessel-background imbalance**. The former arises from significant variations in the distribution of thin and thick vessel pixels across clients, compounded by the complexity of intricate regions, such as vessel intersections. These inconsistencies bias the global model toward dominant features (*e.g.*, thick vessels or simpler regions) while neglecting subtle yet critical details, such as thin vessels and ambiguous boundaries. Existing methods commonly optimize local clients using cross-entropy loss, which treats all units (*e.g.*, images or pixels) uniformly without accounting for fine-grained challenges. Some studies have explored sample reweighting strategies to address data imbalance, enhance the representation of specific samples, and promote fairness across clients [Zhao and Joshi, 2022]. However, these strategies primarily focus on image-level reweighting, while the full potential of pixel-level reweighting remains largely unexplored. This gap is particularly critical for vessel segmentation, where fine-grained structures, such as thin and tortuous vessels, require more precise and adaptive guidance to ensure accurate, reliable, and faithful representation. Consequently, a critical question emerges: *"Q1: How can global information be leveraged to provide differentiated guidance for vessel pixels?"*

Regarding the **vessel-background imbalance**, current federated learning frameworks typically perform one-to-one model or parameter alignment, which involves either full-parameter alignment [Li *et al.*, 2020b] or direct point-to-point feature alignment [Mendieta *et al.*, 2022]. However, this approach significantly limits the applicability of federated learning in vessel segmentation, where the foreground-background disparity is particularly pronounced. Overemphasis on one-to-one alignment often leads to prioritizing background pixels, which weakens the semantic representation of sparse cat-

egories like vessels. Additionally, the aligned features are unstable and lack robustness when handling heterogeneous client data. Consequently, another key question arises: *"Q2: How to use robust feature alignment strategies to balance foreground and background global knowledge?"*

Based on the above issues, we propose our method. Concerning Question 1, we propose Federated Morphology Uncertainty Guidance (**FMUG**), inspired by the sample reweighting strategy in Federated Learning (FL). The first and most important step is to reveal the "hard kernel", which consists of challenging pixels that significantly impact the segmentation performance. Therefore, our method combines global and local uncertainty to guide the model in focusing on these key pixels. To be precise, we generate a global uncertainty map using the predictions of the global model and the ground truth alongside a local uncertainty map derived from the optimization process of the client's local model. These uncertainty maps facilitate a progressive learning process by dynamically identifying areas requiring focused attention at each stage. Leveraging these maps, we employ an uncertainty-based weighted cross-entropy loss to encourage the model to improve its accuracy while enhancing the uncertainty associated with each prediction. As for Question 2, we propose Federated Vessel Decoupling Alignment (**FVDA**), a foreground-background separation alignment approach. Existing methods primarily focus on aligning local and global directions for each pixel, treating this alignment as a holistic optimization term. In contrast, our work aims to decouple the foreground and background using the provided ground-truth labels. Specifically, we reconstruct foreground and background masks based on pixel labels and align each part independently. We argue that this approach is both compatible and flexible, ensuring semantic consistency while maintaining overall coherence with the global style.

In this paper, we aim to tackle the unique challenges in federated vessel segmentation, including vessel morphology inconsistency, foreground-background imbalance. To address these critical issues, we design innovative approaches that efficiently enhance segmentation accuracy, preserve semantic consistency, and improve generalization across diverse datasets. Our contributions are summarized as follows:

- We propose a novel pixel-level federated vessel segmentation method based on a divide-and-conquer approach, which leverages global uncertainty to guide local models in identifying critical pixels and effectively tackles challenges related to vessel morphology inconsistencies.
- We introduce a foreground-background separation alignment approach for vessel segmentation, effectively decoupling and aligning vessel and background features to ensure semantic consistency, thereby improving the generalization of the global model across diverse and heterogeneous client datasets.
- We conduct extensive experiments on two modalities of vessel datasets: one consisting of color fundus images (DRIVE [Staal *et al.*, 2004], STARE [Hoover *et al.*, 2000], CHASEDB1 [Owen *et al.*, 2009]) and the other composed of optical coherence tomography angiography (OCTA) images (ROSE-1 [Ma *et al.*, 2021] and OCT-500 [Li *et al.*, 2024]). Our method demonstrates superior performance compared to related approaches.

2 Related Work

2.1 Federated Learning

With the introduction of FedAvg by McMahan *et al.* [McMahan *et al.*, 2017], federated learning (FL) began to emerge as a promising and scalable approach for distributed learning. This groundbreaking work clearly showcased the potential of FL to enable collaborative model training without requiring raw data exchange, addressing privacy and security concerns that traditional centralized methods could not easily overcome. However, it experiences degraded performance when handling non-i.i.d. data (data heterogeneity). Some approaches focus on leveraging shared models as global signals [Li *et al.*, 2020b; Li *et al.*, 2021; Huang *et al.*, 2022; Huang *et al.*, 2023b; Huang *et al.*, 2023a], while others emphasize utilizing statistical distributions [Zhang *et al.*, 2022; Zhou and Konukoglu, 2023; Tang *et al.*, 2024]. Additionally, certain methods are built upon semantic class prototypes [Tan *et al.*, 2022; Mu *et al.*, 2023; Huang *et al.*, 2023c; Huang *et al.*, 2024a], and some rely on gradient aggregation [Karimireddy *et al.*, 2020; Gao *et al.*, 2022] to enhance collaboration in federated learning. These methods have achieved remarkable success, significantly advancing the development and adoption of federated learning. However, these methods are not specifically designed for pixel-level dense prediction tasks. Their granularity in federated learning is insufficient, especially for medical segmentation tasks, where they struggle to perform well in segmenting objects with blurry boundaries or thin and elongated structures.

2.2 Medical Segmentation Meet Federation

Fueled by the rapid development of deep learning, medical image segmentation has achieved remarkable progress [Ronneberger *et al.*, 2015; Isensee *et al.*, 2021; Ma *et al.*, 2024], greatly enhancing clinical diagnosis and treatment by enabling precise and efficient delineation of anatomical structures and pathological lesions, such as brain tumors [Yang *et al.*, 2023b; Zhou *et al.*, 2024], optic discs [Yii *et al.*, 2023; Chen *et al.*, 2024], and lung nodules [Wang *et al.*, 2023a]. However, training fully supervised models often demands costly and labor-intensive pixel-level annotations, particularly in the medical domain. Additionally, the sensitive nature of patient data and strict privacy regulations create significant legal and practical barriers to centralized data collection and sharing. In this context, federated learning (FL) has emerged as a natural, scalable, and promising solution for medical image segmentation [Wang *et al.*, 2022; Wang *et al.*, 2023b]. Jiang *et al.* [Jiang *et al.*, 2022] effectively mitigates local drift using amplitude normalization and addresses global drift through weight perturbation, thereby resolving style domain discrepancies in medical images.

In this paper, we focus on federated vessel segmentation, where the targets are characterized by variable thickness and tortuous structures. Additionally, the significant foreground-background disparity, with sparse vessel pixels compared to

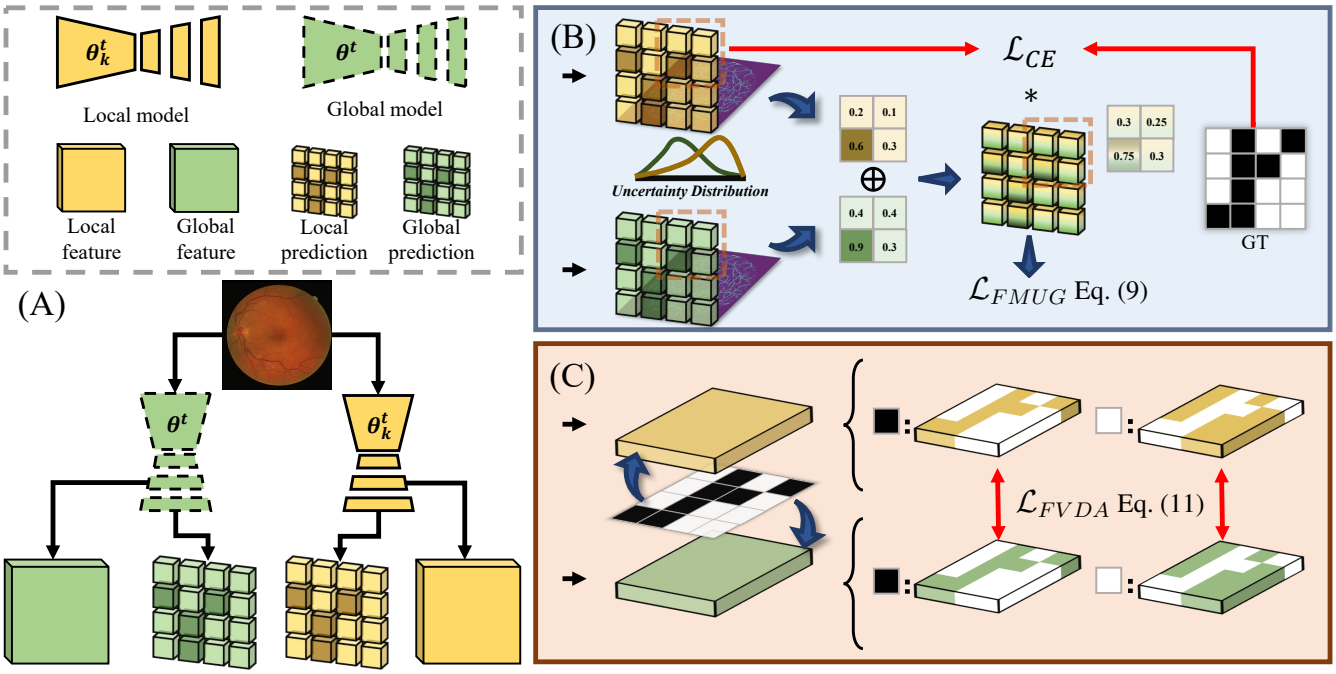


Figure 1: **Illustration of *FVAC***.(A) The workflow of our model, which generates both global and local predictions along with intermediate features. (B) *Federated Morphology Uncertainty Guidance*, as described in Sec. 3.2: The global and local predictions are used as inputs, from which uncertainty maps $\mathcal{U}^{\mathrm{g}}$ and $\mathcal{U}^{\mathrm{k}}$ are derived by projecting the prediction results. The weights $\boldsymbol{\alpha}$ are then obtained through the combination and linear mapping of both. (C) *Federated Vessel Decoupling Alignment*, as described in Sec. 3.2: Based on the labels, the feature map $\mathbf{F}$ is decoupled into the foreground and background features (*e.g.* $((\mathbf{F}^{\mathrm{k}}_{\mathrm{fg}}, \mathbf{F}^{\mathrm{g}}_{\mathrm{fg}}), (\mathbf{F}^{\mathrm{k}}_{\mathrm{bg}}, \mathbf{F}^{\mathrm{g}}_{\mathrm{bg}}))$, which are then used for group-wise alignment.

the dominant background, poses unique challenges for accurate segmentation in federated settings.

2.3 Federated Vessel Segmentation

Vessel segmentation has experienced rapid advancements in recent years. Accurately segmenting vessels with complex shapes and variable thicknesses remains a challenging and important problem. As a landmark network in medical image segmentation, U-Net [Ronneberger *et al.*, 2015] has demonstrated excellent performance in vessel segmentation, owing to its symmetric encoder-decoder and skip connection architecture. Inspired by U-Net, numerous variant networks for vessel segmentation have been developed [Li *et al.*, 2020a; Liu *et al.*, 2022]. Nevertheless, they all rely on detailed annotations, limiting applicability. Moreover, due to strict patient confidentiality and privacy policies, centralized data collection and sharing are highly challenging and often infeasible.

Federated learning (FL) offers a promising new paradigm for vessel segmentation by addressing the challenges of centralized data collection and the reliance on finely detailed annotations. Unfortunately, existing federated medical segmentation fails to fully address the unique challenges of vascular segmentation due to the high variability in vascular shapes and the uneven distribution of vessel thickness.

In this paper, we introduce the first federated learning-based framework designed explicitly for vessel segmentation. This framework enables clients to effectively focus on key pixels, enhance local model performance, and simultaneously learn more generalized, robust, and contextually aware semantic representations across heterogeneous datasets.

3 Methodology

3.1 Preliminaries

Following the setup of FedAvg [McMahan *et al.*, 2017], typical **Federated Learning (FL)** usually involves a centralized server and multiple clients. Each client k holds its private dataset $\mathcal{D}_k$, which can be represented at the sample level as:

$$\mathcal{D}_k = \{x_i, y_i\}_{i=1}^{n_k}, \tag{1}$$

where n_k represents the number of samples in the k-th client's local data. At the beginning of each communication round t, the global model parameter $\boldsymbol{\theta}^t$ is broadcasted by the server to all clients. Each client then performs local optimization on its dataset by minimizing a local objective (*e.g.*, cross-entropy loss), resulting in updated local parameters $\boldsymbol{\theta}_k^t$, which are sent back to the server. The server aggregates these local updates using a weighted averaging mechanism to produce the global model for the next round:

$$\boldsymbol{\theta}^{t+1} = \sum_{k=1}^{M} \alpha_k \boldsymbol{\theta}_k^t, \quad \alpha_k = \frac{n_k}{\sum_{k'=1}^{M} n_{k'}}, \tag{2}$$

where M denotes the total number of clients, and the weight α_k ensures that each client's contribution is proportional to its dataset size. This framework enables collaborative learning without sharing raw data, effectively preserving privacy while utilizing distributed datasets across multiple clients.

For segmentation-specific tasks, the client's dataset can also be expressed at the collection level as:

$$\mathcal{D}_k = \{X_i, Y_i\}_{i=1}^{n_k}, \tag{3}$$

where $\mathcal{P}(h, w)$ denotes the predicted probability for pixel (h, w), and $Y(h, w)$ is its ground truth label. The function $\text{argmax}(\mathcal{P}(h, w))$ returns the class with the highest predicted probability. For correctly classified pixels, uncertainty increases when non-predicted classes have higher probabilities, indicating weaker separation. For misclassified pixels, higher probability for the incorrect class leads to increased uncertainty, reflecting overconfidence. This behavior underscores the need to analyze the full probability distribution to assess confidence and detect weaknesses.

3.2 Proposed Method

Motivation. Federated Learning (FL) has emerged as a privacy-preserving collaborative learning framework, enabling multiple institutions to train global models without sharing sensitive data. In medical image segmentation tasks, existing FL methods primarily focus on addressing data heterogeneity caused by differences in imaging modalities (*e.g.*, MRI, CT) or acquisition protocols, emphasizing mitigating style shifts in global and local feature distributions.

However, segmenting thin and tortuous structures (*e.g.*, blood vessels) presents unique challenges in federated learning. Blood vessel segmentation suffers from **vessel morphology inconsistency**, where the distribution of thin and thick vessels varies significantly across clients. As a result, the global model often overlooks subtle yet important details, such as thin vessels and unclear boundaries, while focusing on dominant features like thick vessels or simpler regions. Moreover, the **vessel-background imbalance**, where vessels occupy only a small fraction of pixels, further weakens the representation of sparse categories and causes unstable feature alignment across heterogeneous clients.

Considering these challenges, this work focuses on:

- Leveraging the global uncertainty map to guide local clients in effectively prioritizing key regions (*e.g.*, thin vessels and blurred edges) during training.
- Foreground-background separation semantic alignment to help local clients learn more generalizable features, ensuring consistency across diverse clients while preserving fine-grained segmentation performance.

Overview of Framework. The framework of our method is illustrated in Fig. 1. Specifically, We leverage the uncertainty from the prediction results to assist local models in learning the hard kernels while utilizing decoupled features to mitigate the bias toward the dominant class. Next, we will introduce Federated Morphology Uncertainty Guidance and Federated Vessel Decoupling Alignment in detail.

I) Federated Morphology Uncertainty Guidance. Considering vessel morphology inconsistency, which represents pixel-level data heterogeneity (*e.g.*, varying vessel crossing degrees), this issue severely hampers the learning of fine-grained features. As a result, the global model becomes biased toward dominant features (*e.g.*, thick vessels or simpler regions), leaving it uncertain or even incapable of segmenting thin, elongated structures or ambiguous boundaries.

To address this, we leverage global uncertainty to guide local models in further refining their learning on hard-to-distinguish pixels, improving segmentation accuracy. Additionally, we treat the local client optimization process as a progressive learning approach, specifically focusing on uncertain pixels. By incorporating local uncertainty, we dynamically adjust the global uncertainty, deriving stable pixel-level learning weights to enhance model performance.

For a given input X, the model's predictions are represented as:

$$\mathcal{Z} = f(X; \boldsymbol{\theta}), \tag{4}$$

$$\mathcal{P} = \sigma(\mathcal{Z}), \tag{5}$$

where $\mathcal{Z}$ denotes the logits of the model, and σ represents the softmax function, which transforms the logits into a pixel-wise probability distribution $\mathcal{P}$.

Based on these predictions, uncertainty maps are derived to quantify the reliability of the predictions at the pixel level:

$$\mathcal{U} = \begin{cases} \min(\mathcal{P}(h, w)), & \text{if } \text{argmax}(\mathcal{P}(h, w)) = Y(h, w), \\ \max(\mathcal{P}(h, w)), & \text{otherwise.} \end{cases} \tag{6}$$

where $\mathcal{P}(h, w)$ represents the predicted probability for pixel (h, w), and $Y(h, w)$ denotes the ground truth label for that pixel. The function $\text{argmax}(\mathcal{P}(h, w))$ returns the index of the class with the highest predicted probability at pixel (h, w). For correctly classified pixels, uncertainty increases when the probabilities assigned to non-predicted classes are higher, indicating a weaker separation between the predicted class and others. Conversely, for misclassified pixels, uncertainty increases when the probability of the incorrect class is higher, reflecting the model's overconfidence in its prediction. This nuanced behavior highlights the importance of understanding the probability distribution across all classes to effectively assess confidence and identify weaknesses or inconsistencies.

Thus, based on the uncertainty maps $\mathcal{U}^{\text{g}} \in \mathbb{R}^{H \times W}$ and $\mathcal{U}^{\text{k}} \in \mathbb{R}^{H \times W}$, the pixel-wise weight map for the entire image is computed as a linear combination of the global and local uncertainty maps:

$$\boldsymbol{\alpha} = 0.5 \cdot \mathcal{U}^{\text{g}} + 0.5 \cdot \mathcal{U}^{\text{k}}. \tag{7}$$

where the superscripts g and k denote the global uncertainty (aggregated across clients) and the local uncertainty (specific to the current client), respectively. To ensure proper normalization, the weights are scaled as follows:

$$\boldsymbol{\alpha} = \frac{\boldsymbol{\alpha}}{\sum_{h=1}^{H} \sum_{w=1}^{W} \boldsymbol{\alpha}(h, w)}, \tag{8}$$

where the denominator ensures that the sum of all weights across the image equals 1. These normalized weights are then used to compute the weighted cross-entropy loss over the entire image:

$$\mathcal{L}_{\text{FMUG}} = -\sum_{h,w} \sum_{c=0}^{C-1} \boldsymbol{\alpha}(h, w) \cdot \mathbf{1}_{Y(h,w)=c} \cdot \log\left(\mathcal{P}_c(h, w)\right), \tag{9}$$

where $\mathbf{1}_{Y(h,w)=c}$ selects the predicted probability $\mathcal{P}_c(h, w)$ corresponding to the ground-truth class $c \in \{0, \ldots, C-1\}$.

Here, C denotes the total number of semantic classes (and prediction channels). $\boldsymbol{\alpha}(h, w)$ is a pixel-wise adaptive weight derived from both global and local uncertainty. In our binary vessel segmentation task, $C = 2$.

This approach adaptively penalizes ambiguity in correct predictions and overconfidence in incorrect ones. Leveraging uncertainty maps, the model focuses on difficult regions, enhancing robustness and performance.

II) Federated Vessel Decoupling Alignment. Vessel–background imbalance poses another significant challenge in federated vessel segmentation, weakening the semantic representation of sparse classes like vessels. Common one-to-one feature learning or parameter-sharing strategies in general federated learning often exacerbate this by overemphasizing background pixels, leading to unstable feature learning and reduced model robustness.

To address this, we extract foreground and background features from intermediate model features $\mathbf{F} \in \mathbb{R}^{C\times H\times W}$, where C, H, and W are the number of channels, height, and width. The ground truth Y is resized to $H \times W$ to obtain the mask $\mathbf{S}$, and we optimize foreground and background features separately. The optimization is as follows:

$$\begin{aligned}\mathbf{F}_{\text{fg}} &= \frac{1}{H\cdot W}\sum_{h=1}^{H}\sum_{w=1}^{W}\mathbf{F}[:,h,w]\cdot\mathbf{1}_{\mathbf{S}(h,w)=1},\\ \mathbf{F}_{\text{bg}} &= \frac{1}{H\cdot W}\sum_{h=1}^{H}\sum_{w=1}^{W}\mathbf{F}[:,h,w]\cdot\mathbf{1}_{\mathbf{S}(h,w)=0},\end{aligned} \tag{10}$$

where $\mathbf{F}[:, h, w]$ represents the feature vector at pixel (h, w), and $\mathbf{S}(h, w)$ is the mask value at pixel (h, w). The features $\mathbf{F}_{\text{fg}}$ and $\mathbf{F}_{\text{bg}}$ are averaged across all pixels, ultimately yielding a $C \times 1$ feature vector for each respective category.

To ensure consistency between the local and global feature representations, we calculate the Mean Squared Error (MSE) loss for both foreground and background features in the model. The loss is then computed as follows:

$$\mathcal{L}_{FVDA} = \frac{1}{C}\left\|\mathbf{F}^{\text{k}}_{\text{fg}} - \mathbf{F}^{\text{g}}_{\text{fg}}\right\|_2^2 + \frac{1}{C}\left\|\mathbf{F}^{\text{k}}_{\text{bg}} - \mathbf{F}^{\text{g}}_{\text{bg}}\right\|_2^2. \tag{11}$$

Through the combined use of the reweighted cross-entropy loss and the decoupled alignment loss, the final training objective is defined as:

$$\mathcal{L}_{FVAC} = \mathcal{L}_{FMUG} + \beta\cdot\mathcal{L}_{FVDA}, \tag{12}$$

where β controls the balance between the two loss terms. The reweighted cross-entropy loss $\mathcal{L}_{FMUG}$ improves segmentation accuracy and confidence, while the decoupled alignment loss $\mathcal{L}_{FVDA}$ ensures consistent features between local and global models. The overall algorithm is illustrated in Algorithm 1. Jointly optimizing both losses boosts accuracy and preserves foreground–background consistency.

Algorithm 1: Model training in *FVAC*

Input: Communication rounds T, local epochs E, number of participants K, the k^{th} participant private data $D_k(x, y)$, private model θ_k

Output: The final global model θ^T

for $t = 1, 2, ..., T$ **do**
  Participant Side;
  for $k = 1, 2, ..., K$ **in parallel do**
    $\theta_k^t \leftarrow$ **LocalUpdating**$(\theta^t, \mathcal{G})$
  Server Side;
  $\theta^{t+1} \leftarrow \frac{1}{K}\sum_{k=1}^{K}\theta_k^t$

LocalUpdating$(\theta^t, \mathcal{G})$:

$\theta_k^t \leftarrow \theta^t$; `// Distribute global parameter`
$\theta_{fixed} \leftarrow \theta^t$; `// Fix global parameter`
for $e = 1, 2, ..., E$ **do**
  for $(X_i, Y_i) \in D_k$ **do**
    $\mathcal{Z}_i = f(X_i, \theta_k^t)$, $\mathcal{Z}_i^g = f(X_i, \theta_{fixed})$
    $\mathcal{P}_i = \sigma(\mathcal{Z}_i)$, $\mathcal{P}_i^g = \sigma(\mathcal{Z}_i^g)$
    `/* Uncertainty Estimation */`
    $U_i \leftarrow (\mathcal{P}_i, Y_i)$, $U_i^g \leftarrow (\mathcal{P}_i^g, Y_i)$ in Eq. (6)
    `/* Weight Assignment */`
    $\alpha \leftarrow (U_i^g, U_i)$ in Eqs. (7) and (8)
    `/* Feature Decoupling */`
    $\mathbf{F_{fg_i}}, \mathbf{F_{bg_i}} \leftarrow (\mathbf{F_i}, \mathbf{S_i})$ in Eq. (10)
    $\mathbf{F^g_{fg_i}}, \mathbf{F^g_{bg_i}} \leftarrow (\mathbf{F^g_i}, \mathbf{S_i})$ in Eq. (10)
    $\mathcal{L}^k_{FMUG} \leftarrow (\alpha, \mathcal{P}_i, Y_i)$ in Eq. (9)
    $\mathcal{L}^k_{FVDA} \leftarrow (\mathbf{F_{fg_i}}, \mathbf{F_{bg_i}}, \mathbf{F^g_{fg_i}}, \mathbf{F^g_{bg_i}})$ in Eq. (11)
    $\mathcal{L}^k_{FVAC} \leftarrow (\mathcal{L}^k_{FMUG}, \mathcal{L}^k_{FVDA})$in Eq. (12)
    $\theta_k^e \leftarrow \theta_k^e - \eta\nabla\mathcal{L}^k$
return θ_k^e

3.3 Discussion

Conceptual Difference. Most FL methods use cross-entropy loss and sample reweighting but overlook pixel-level weighting and global context, limiting their ability to handle complex pixels. We first introduce global uncertainty in FL for vessel segmentation, guiding local models toward hard pixels and boosting accuracy. Moreover, conventional FL aligns models one-to-one, favoring dominant classes like the background and causing unstable feature learning. We address this by leveraging global uncertainty and decoupling foreground-background features for separate alignment, boosting segmentation precision and overall performance.

Related Uncertainty Learning. Uncertainty learning in semi-supervised settings helps models focus on informative regions and avoid noisy inputs, improving robustness and accuracy. Inspired by this, we incorporate uncertainty maps into FL to address vessel morphology inconsistencies, which often lead the global model to miss key pixels. Enhancing local learning on such regions mitigates global forgetting and boosts segmentation performance.

Limitation. *FVAC* achieves strong vessel segmentation via a pixel-wise divide-and-conquer strategy. However, it incurs extra training cost due to the additional forward pass for global uncertainty and overlooks potential attacks on the global model, which may mislead local clients. Future work will focus on enhancing robustness and security against adversarial threats while preserving efficiency and accuracy.

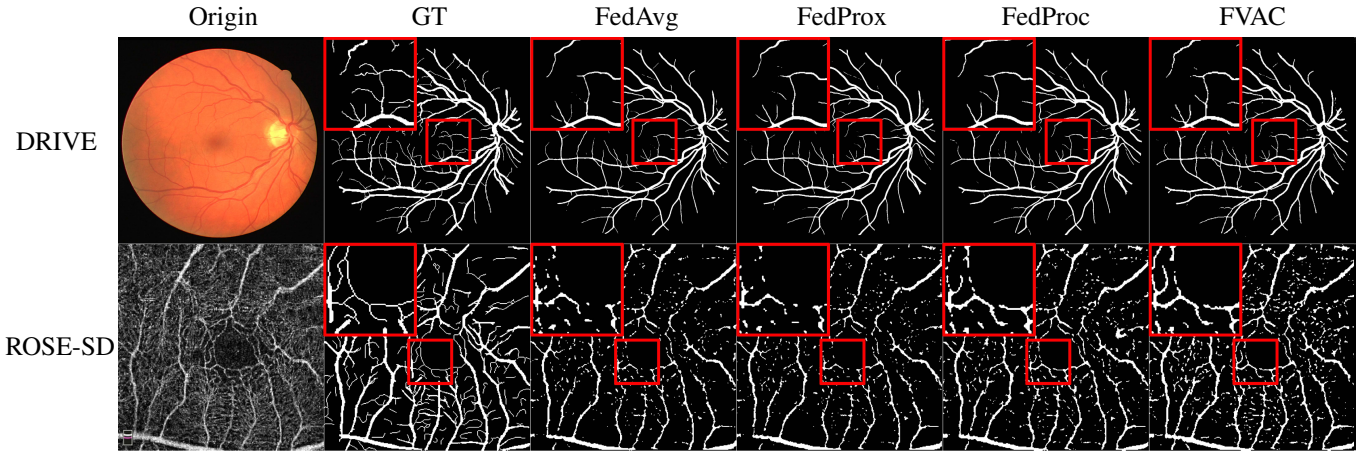


Figure 2: **Qualitative comparison results of exemplars**. The top-left corner highlights the zoomed-in details within the red box for each image. Our method demonstrates significantly more complete segmentation of vessel intersections and fine capillaries, effectively capturing intricate structures that are often challenging to distinguish, thereby improving overall segmentation performance. See detail in Sec. 4.2

Methods	Color Funds				OCTA				
	DRIVE	STARE	CHASEDB1	AVG	ROSE-SVC	ROSE-SD	OCTA-3MM	OCTA-6MM	AVG
FedAvg [ASTAT17]	77.23	79.74	74.00	76.99	64.27	57.98	59.34	41.76	55.84
FedProx [MLSys18]	77.60	79.95	<u>74.72</u>	<u>77.42</u>	63.46	57.24	59.56	41.14	55.35
MOON [CVPR21]	<u>78.10</u>	79.57	72.24	76.64	61.42	51.15	**60.50**	39.68	53.19
FedDyn [ICLR21]	77.24	79.74	74.00	76.99	64.27	<u>57.99</u>	59.35	41.77	55.85
FedOPT [ICLR21]	73.66	77.47	68.62	73.25	57.18	53.65	57.53	37.76	51.53
FedProto [AAAI22]	77.27	79.52	73.18	76.66	63.45	57.50	59.53	39.80	55.07
FedProc [FGCS23]	77.56	<u>79.96</u>	73.53	77.01	<u>64.70</u>	56.46	59.40	<u>43.07</u>	<u>55.90</u>
FVAC	**78.58** ↑ 1.35	**80.26** ↑ 0.52	**74.83** ↑ 0.83	**77.89** ↑ 0.90	**67.62** ↑ 3.35	**61.50** ↑ 3.52	<u>59.70</u> ↑ 0.36	**48.20** ↑ 6.44	**59.30** ↑ 3.46

Table 1: **Comparison with the sota methods** on the Color Fundus and OCTA datasets. AVG denotes the average Dice coefficient calculated across all domains. The best result is highlighted in bold, and the second-best is underlined. ↑ indicates an improvement in Dice compared to FedAvg, which serves as the baseline for our method. These notations apply to all other comparisons as well. See detail in Sec. 4.2

4 Experiments

4.1 Experimental Setup

Datasets and Evaluation Metric. We evaluate *FVAC* on both color fundus and OCTA retinal datasets. The fundus data include DRIVE [Staal *et al.*, 2004], STARE [Hoover *et al.*, 2000], and CHASEDB1 [Owen *et al.*, 2009], while the OCTA data consist of ROSE-1 [Ma *et al.*, 2021] and OCT-500 [Li *et al.*, 2024].

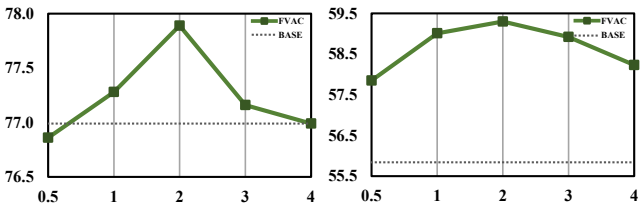
- DRIVE [Staal *et al.*, 2004] contains 40 high-resolution color retinal images (565×584), evenly split into 20 for training and 20 for testing.
- STARE [Hoover *et al.*, 2000] contains 20 manually annotated color retinal images (700×605), with 16 used for training and 4 for testing.
- CHASEDB1 [Owen *et al.*, 2009] contains 28 color retinal images with a resolution of 999×960 pixels, split into 20 images for training and 8 for testing.
- ROSE-1 [Ma *et al.*, 2021], a subset of ROSE, includes 117 OCTA images from 39 subjects, with 90 for training and 27 for testing. Each subject has enface angiograms of the superficial (SVC), deep (DVC), and combined (SVC+DVC) vascular plexus. We use ROSE-SVC and ROSE-SD (SVC+DVC) as two domain-specific subsets of our OCTA dataset.
- OCT-500 [Li *et al.*, 2024] provides OCTA images from 500 subjects across two fields of view, split into OCTA-6MM (300 images) and OCTA-3MM (200 images). We use 280/20 for training/testing in OCTA-6MM and 180/20 in OCTA-3MM. Both subsets are treated as separate domains in our OCTA dataset.

We evaluate all vessel datasets using the Dice coefficient as the sole performance metric, as it is widely adopted in medical image segmentation to reflect the overlap between predicted and ground truth regions.

Comparison Methods. We compare our method, *FVAC*, with several state-of-the-art approaches, including FedProx [Li *et al.*, 2020b], FedDyn [Durmus *et al.*, 2021], FedOPT [Reddi *et al.*, 2021], MOON [Li *et al.*, 2021], FedProto [Tan *et al.*, 2022], and FedProc [Mu *et al.*, 2023].Although experimental settings may vary slightly across methods, we ensure fairness by preserving key features of each approach for comparison.

FMUG	FVDA	Color Funds			
		DRIVE	STARE	CHASEDB1	AVG
		77.23	79.74	74.00	76.99
✓		77.27	79.90	74.51	77.23
	✓	77.34	79.85	74.50	77.23
✓	✓	78.58	80.26	74.83	**77.89**

Table 2: **Ablation study of key components** of our method on the Color Funds Dataset demonstrates each branch plays a distinct and complementary role. See detail in Sec. 4.2



(a) Color Funds (b) OCTA

Figure 3: **Analysis of *FVAC* with different beta (Eq. (12))** on Color Funds and OCTA datasets. "BASE" denotes FedAvg. $\beta = 2$ gives the best results on different datasets. See detail in Sec. 4.2

Implementation Details. For the experiment, to ensure simplicity and general applicability, we choose U-Net, a widely used and effective architecture, as the backbone for all Federated Learning methods. To ensure the reproducibility and consistency of our results, we fix the random seed across all experiments. In the federated learning process, models are trained using the AdamW optimizer [Loshchilov and Hutter, 2019] with a batch size of 4. The communication round is set to 100, and the local training epoch is 5 for all datasets.

We set up six participants per experiment with different combined datasets and randomly assign domains. For color fundus data, the split is DRIVE:3, STARE:1, CHASEDB1:1; for OCTA, ROSE-SVC:2, ROSE-SD:1, OCTA-3MM:1, OCTA-6MM:2. Each participant receives 1% of the original data from their assigned domains. We also apply data augmentations on each client, including color jitter, gamma correction, flipping, rotation, and random cropping.

We use a learning rate of 1e-4 for both collaborative and local updates on the color retinal datasets across all methods. In our method, all participants share the same hyperparameters (*i.e.*, $\beta = 2$). For the OCTA datasets, the learning rate is reduced to 1e-5 due to higher segmentation difficulty from varying modalities and denser vessel structures.

4.2 Results

Performance Comparison. For clarity and brevity, we present representative visualizations from the final model after federated learning, as shown in Fig. 2. Our method preserves more complete vascular structures, especially in regions with complex, twisted, or entangled connections. Though such areas constitute only a small part of the overall network, they are essential for downstream tasks like blood flow modeling or vascular assessment. Based on earlier qualitative results, the final Dice scores across popular methods are reported in Tab. 1, showing that our approach consistently outperforms others. On the OCTA dataset, it achieves a Dice improvement of nearly 3.5% over the baseline. Both qualitative and quantitative results confirm our method's strength in capturing complex and detailed vascular structures.

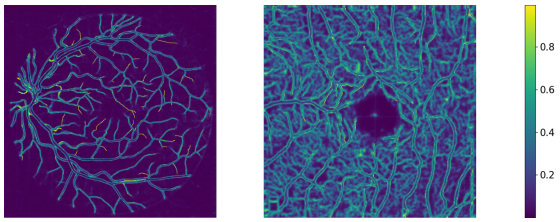


Figure 4: **Visualization of the pixel-wise weights (Eqs. (7) and (8)) in the FMUG**. As shown in the color bar, brighter regions distinctly highlight areas the model prioritizes more, emphasizing its focus on these regions. See detail in Sec. 4.2

Hyper-Parameter Setting. We evaluate *FVAC* under different β values as defined in Eq. (12), searching over $[0.5, 1, 2, 3, 4]$. As shown in Fig. 3, a small β overemphasizes the weighted cross-entropy loss, causing the local model to overfit difficult pixels and misalign with global features. A large β, in contrast, leads to plain feature alignment and fails to leverage challenging pixels. We choose $\beta = 2$ as the optimal setting and apply it consistently in all experiments.

Ablation Study. Compared to FedAvg, FMUG better captures challenging vessel structures, such as edges, capillaries, and crossing or adjacent vessels, as shown in Fig. 4. These hard-to-classify pixels, caused by morphology inconsistency across clients, often degrade performance. To address this, FMUG leverages uncertainty to guide the global model in helping local models focus on such regions, as shown in Tab. 2. Meanwhile, FVDA improves segmentation by separately aligning foreground and background features. When combined, *FVAC* maintains sensitivity to morphology variations while learning more robust and generalized representations, boosting segmentation performance.

5 Conclusion

In this paper, we are the first to explore federated vessel segmentation. To address two key challenges in federated vessel segmentation, namely vessel morphology inconsistency and vessel-background imbalance, we propose a pixel-wise divide-and-conquer approach called Federated Vessel-Aware Calibration (*FVAC*). By leveraging global uncertainty, our method directs the local models' attention to the hard kernel, which consists of pixels with high uncertainty caused by variable vessel morphology. Additionally, we use foreground-background separation alignment to mitigate the semantic bias induced by pixel-level class imbalance, ultimately leading to more robust semantic feature learning. Numerous experiments have clearly demonstrated the effectiveness and generalizability of the *FVAC* method, particularly in handling complex and detailed vascular structures. In the future, we will further explore innovative solutions to enhance the performance, robustness, and security of federated vessel segmentation, with a particular focus on reliably defending against potential adversarial threats to the global model.

Contribution Statement

Tian Chen, Wenke Huang, and Zhihao Wang contributed equally to this work and are marked with *. Yongchao Xu is the corresponding author and is marked with †.

Acknowledgments

This work was supported in part by NSFC 62222112, 62225113, 62325111, 62176186, 62361166629, 62176188, and 623B2080, the Innovative Research Group Project of Hubei Province under Grants (2024AFA017) and the National Key Research and Development Program of China (2024YFC3308400).
We thank Zekun Shi for his valuable collaboration. This work would not have been possible without him.

References

[Chen *et al.*, 2024] Ziyang Chen, Yongsheng Pan, Yiwen Ye, Zhiyong Wang, and Yong Xia. Trila: Triple-level alignment based unsupervised domain adaptation for joint segmentation of optic disc and optic cup. *IEEE Journal of Biomedical and Health Informatics*, 2024.

[Durmus *et al.*, 2021] Alp Emre Durmus, Zhao Yue, Matas Ramon, Mattina Matthew, Whatmough Paul, and Saligrama Venkatesh. Federated learning based on dynamic regularization. In *Proc. of Intl. Conf. on Learning Representations*, 2021.

[Elgabli *et al.*, 2022] Anis Elgabli, Chaouki Ben Issaid, Amrit Singh Bedi, Ketan Rajawat, Mehdi Bennis, and Vaneet Aggarwal. Fednew: A communication-efficient and privacy-preserving newton-type method for federated learning. In *Proc. of Intl. Conf. on Machine Learning*, 2022.

[Gao *et al.*, 2022] Liang Gao, Huazhu Fu, Li Li, Yingwen Chen, Ming Xu, and Cheng-Zhong Xu. Feddc: Federated learning with non-iid data via local drift decoupling and correction. In *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, 2022.

[Hoover *et al.*, 2000] AD Hoover, Valentina Kouznetsova, and Michael Goldbaum. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Trans. on Medical Imaging*, 2000.

[Huang *et al.*, 2022] Wenke Huang, Mang Ye, and Bo Du. Learn from others and be yourself in heterogeneous federated learning. In *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, 2022.

[Huang *et al.*, 2023a] Wenke Huang, Guancheng Wan, Mang Ye, and Bo Du. Federated graph semantic and structural learning. In *Proc. of Intl. Joint Conf. on Artificial Intelligence*, 2023.

[Huang *et al.*, 2023b] Wenke Huang, Mang Ye, Zekun Shi, and Bo Du. Generalizable heterogeneous federated cross-correlation and instance similarity learning. *IEEE Trans. on Pattern Anal. and Mach. Intell.*, 2023.

[Huang *et al.*, 2023c] Wenke Huang, Mang Ye, Zekun Shi, He Li, and Bo Du. Rethinking federated learning with domain shift: A prototype view. In *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, 2023.

[Huang *et al.*, 2024a] Wenke Huang, Yuxia Liu, Mang Ye, Jun Chen, and Bo Du. Federated learning with long-tailed data via representation unification and classifier rectification. *IEEE Transactions on Information Forensics and Security*, 2024.

[Huang *et al.*, 2024b] Wenke Huang, Mang Ye, Zekun Shi, Guancheng Wan, He Li, Bo Du, and Qiang Yang. Federated learning for generalization, robustness, fairness: A survey and benchmark. *IEEE Trans. on Pattern Anal. and Mach. Intell.*, 2024.

[Hyeon-Woo *et al.*, 2022] Nam Hyeon-Woo, Moon Ye-Bin, and Tae-Hyun Oh. Fedpara: Low-rank hadamard product for communication-efficient federated learning. In *Proc. of Intl. Conf. on Learning Representations*, 2022.

[Isensee *et al.*, 2021] Fabian Isensee, Paul F Jaeger, Simon AA Kohl, Jens Petersen, and Klaus H Maier-Hein. nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature methods*, 2021.

[Jiang *et al.*, 2022] Meirui Jiang, Zirui Wang, and Qi Dou. Harmofl: Harmonizing local and global drifts in federated learning on heterogeneous medical images. In *Proc. of the AAAI Conf. on Artificial Intelligence*, 2022.

[Karimireddy *et al.*, 2020] Sai Praneeth Karimireddy, Satyen Kale, Mehryar Mohri, Sashank J Reddi, Sebastian U Stich, and Ananda Theertha Suresh. Scaffold: Stochastic controlled averaging for on-device federated learning. In *Proc. of Intl. Conf. on Machine Learning*, 2020.

[Li *et al.*, 2020a] Liangzhi Li, Manisha Verma, Yuta Nakashima, Hajime Nagahara, and Ryo Kawasaki. Iternet: Retinal image segmentation utilizing structural redundancy in vessel networks. In *Proc. of IEEE Winter Conf. on Applications of Computer Vision*, 2020.

[Li *et al.*, 2020b] Tian Li, Anit Kumar Sahu, Manzil Zaheer, Maziar Sanjabi, Ameet Talwalkar, and Virginia Smith. Federated optimization in heterogeneous networks. *Proceedings of Machine learning and systems*, 2020.

[Li *et al.*, 2021] Qinbin Li, Bingsheng He, and Dawn Song. Model-contrastive federated learning. In *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, 2021.

[Li *et al.*, 2024] Mingchao Li, Kun Huang, Qiuzhuo Xu, Jiadong Yang, Yuhan Zhang, Zexuan Ji, Keren Xie, Songtao Yuan, Qinghuai Liu, and Qiang Chen. Octa-500: a retinal dataset for optical coherence tomography angiography study. *Medical image analysis*, 2024.

[Liu *et al.*, 2022] Wentao Liu, Huihua Yang, Tong Tian, Zhiwei Cao, Xipeng Pan, Weijin Xu, Yang Jin, and Feng Gao. Full-resolution network and dual-threshold iteration for retinal vessel and coronary angiograph segmentation. *IEEE Journal of Biomedical and Health Informatics*, 2022.

[Liu *et al.*, 2024] Yepeng Liu, Baosheng Yu, Tian Chen, Yuliang Gu, Bo Du, Yongchao Xu, and Jun Cheng. Progressive retinal image registration via global and local deformable transformations. In *IEEE Intl. Conf. on Bioinformatics and Biomedicine*, 2024.

[Loshchilov and Hutter, 2019] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *Proc. of Intl. Conf. on Learning Representations*, 2019.

[Ma *et al.*, 2021] Yuhui Ma, Huaying Hao, Jianyang Xie, Huazhu Fu, Jiong Zhang, Jianlong Yang, Zhen Wang, Jiang Liu, Yalin Zheng, and Yitian Zhao. Rose: a retinal oct-angiography vessel segmentation dataset and new model. *IEEE Trans. on Medical Imaging*, 2021.

[Ma *et al.*, 2024] Jun Ma, Yuting He, Feifei Li, Lin Han, Chenyu You, and Bo Wang. Segment anything in medical images. *Nature Communications*, 2024.

[McMahan *et al.*, 2016] H Brendan McMahan, FX Yu, P Richtarik, AT Suresh, D Bacon, et al. Federated learning: Strategies for improving communication efficiency. In *Proc. of Adv. in Neural Information Processing Systems*, 2016.

[McMahan *et al.*, 2017] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, 2017.

[Mendieta *et al.*, 2022] Matias Mendieta, Taojiannan Yang, Pu Wang, Minwoo Lee, Zhengming Ding, and Chen Chen. Local learning matters: Rethinking data heterogeneity in federated learning. In *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, 2022.

[Mu *et al.*, 2023] Xutong Mu, Yulong Shen, Ke Cheng, Xueli Geng, Jiaxuan Fu, Tao Zhang, and Zhiwei Zhang. Fedproc: Prototypical contrastive federated learning on non-iid data. *Future Gener. Comput. Syst.*, 2023.

[Owen *et al.*, 2009] Christopher G Owen, Alicja R Rudnicka, Robert Mullen, Sarah A Barman, Dorothy Monekosso, Peter H Whincup, Jeffrey Ng, and Carl Paterson. Measuring retinal vessel tortuosity in 10-year-old children: validation of the computer-assisted image analysis of the retina (caiar) program. *Investigative ophthalmology & visual science*, 2009.

[Reddi *et al.*, 2021] Sashank J Reddi, Zachary Charles, Manzil Zaheer, Zachary Garrett, Keith Rush, Jakub Konečnỳ, Sanjiv Kumar, and Hugh Brendan McMahan. Adaptive federated optimization. In *Proc. of Intl. Conf. on Learning Representations*, 2021.

[Ronneberger *et al.*, 2015] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Proc. of Intl. Conf. on Medical Image Computing and Computer Assisted Intervention*, 2015.

[Staal *et al.*, 2004] Joes Staal, Michael D Abràmoff, Meindert Niemeijer, Max A Viergever, and Bram Van Ginneken. Ridge-based vessel segmentation in color images of the retina. *IEEE Trans. on Medical Imaging*, 2004.

[Tan *et al.*, 2022] Yue Tan, Guodong Long, Lu Liu, Tianyi Zhou, Qinghua Lu, Jing Jiang, and Chengqi Zhang. Fedproto: Federated prototype learning across heterogeneous clients. In *Proc. of the AAAI Conf. on Artificial Intelligence*, 2022.

[Tang *et al.*, 2024] Zhenheng Tang, Yonggang Zhang, Shaohuai Shi, Xinmei Tian, Tongliang Liu, Bo Han, and Xiaowen Chu. Fedimpro: Measuring and improving client update in federated learning. In *Proc. of Intl. Conf. on Learning Representations*, 2024.

[Wang *et al.*, 2022] Jiacheng Wang, Yueming Jin, and Liansheng Wang. Personalizing federated medical image segmentation via local calibration. In *Proc. of European Conf. on Computer Vision*, 2022.

[Wang *et al.*, 2023a] Changwei Wang, Rongtao Xu, Shibiao Xu, Weiliang Meng, Jun Xiao, and Xiaopeng Zhang. Accurate lung nodule segmentation with detailed representation transfer and soft mask supervision. *IEEE Trans. on Neural Networks and Learning Systems*, 2023.

[Wang *et al.*, 2023b] Jiacheng Wang, Yueming Jin, Danail Stoyanov, and Liansheng Wang. Feddp: Dual personalization in federated medical image segmentation. *IEEE Transactions on Medical Imaging*, 2023.

[Wang *et al.*, 2025] Zhihao Wang, He Bai, Wenke Huang, Duantengchuan Li, Jian Wang, and Bing Li. Federated recommendation with explicitly encoding item bias. *Proc. of the AAAI Conf. on Artificial Intelligence*, 2025.

[Yang *et al.*, 2023a] Fu-En Yang, Chien-Yi Wang, and Yu-Chiang Frank Wang. Efficient model personalization in federated learning via client-specific prompt generation. In *Proc. of IEEE Intl. Conf. on Computer Vision*, 2023.

[Yang *et al.*, 2023b] Hengyi Yang, Tao Zhou, Yi Zhou, Yizhe Zhang, and Huazhu Fu. Flexible fusion network for multi-modal brain tumor segmentation. *IEEE Journal of Biomedical and Health Informatics*, 2023.

[Yii *et al.*, 2023] Fabian Yii, Tom MacGillivray, and Miguel O Bernabeu. Data efficiency of segment anything model for optic disc and cup segmentation. In *Proc. of Intl. Conf. on Medical Image Computing and Computer Assisted Intervention*. Springer, 2023.

[Zhang *et al.*, 2022] Jie Zhang, Zhiqi Li, Bo Li, Jianghe Xu, Shuang Wu, Shouhong Ding, and Chao Wu. Federated learning with label distribution skew via logits calibration. In *Proc. of Intl. Conf. on Machine Learning*, 2022.

[Zhao and Joshi, 2022] Zhiyuan Zhao and Gauri Joshi. A dynamic reweighting strategy for fair federated learning. In *IEEE Intl. Conf. on Acoustics, Speech and Signal Processing*, 2022.

[Zhou and Konukoglu, 2023] Tianfei Zhou and Ender Konukoglu. FedFA: Federated feature augmentation. In *Proc. of Intl. Conf. on Learning Representations*, 2023.

[Zhou *et al.*, 2024] Lifang Zhou, Yu Jiang, Weisheng Li, Jun Hu, and Shenhai Zheng. Shape-scale co-awareness network for 3d brain tumor segmentation. *IEEE Trans. on Medical Imaging*, 2024.