



图结构化辅助监督的视网膜血管分割网络

吕佳, 杨鸿强

引用本文:

吕佳, 杨鸿强. 图结构化辅助监督的视网膜血管分割网络[J]. 智能系统学报, 2026, 21(5): 1166–1179.

LÜ Jia, YANG Hongqiang. U-Net-based graph-structured auxiliary supervision network for retinal vessel segmentation[J]. CAAI transactions on intelligent systems, 2026, 21(5): 1166–1179.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202509041>

您可能感兴趣的其他文章

基于反馈注意力机制和上下文融合的非模式实例分割

Feedback attention mechanism and context fusion based amodal instance segmentation

智能系统学报. 2021, 16(4): 801–810 <https://dx.doi.org/10.11992/tis.202007042>

差异特征融合的无监督SAR图像变化检测

Unsupervised SAR image change detection based on difference feature fusion

智能系统学报. 2021, 16(3): 595–604 <https://dx.doi.org/10.11992/tis.202103011>

一种改进的深度学习道路交通标识识别算法

An improved deep learning algorithm for road traffic identification

智能系统学报. 2020, 15(6): 1121–1130 <https://dx.doi.org/10.11992/tis.201811009>

生成对抗网络辅助学习的舰船目标精细识别

Fine-grained inshore ship recognition assisted by deep-learning generative adversarial networks

智能系统学报. 2020, 15(2): 296–301 <https://dx.doi.org/10.11992/tis.201901004>

利用多模态U形网络的CT图像前列腺分割

Prostate segmentation in CT images with multimodal U-net

智能系统学报. 2018, 13(6): 981–988 <https://dx.doi.org/10.11992/tis.201806012>

深度学习在无人驾驶汽车领域应用的研究进展

Deep learning in driverless vehicles

智能系统学报. 2018, 13(1): 55–69 <https://dx.doi.org/10.11992/tis.201609029>

DOI: 10.11992/tis.202509041

网络出版地址: <https://link.cnki.net/urlid/23.1538.TP.20260609.1523.004>

图结构化辅助监督的视网膜血管分割网络

吕佳^{1,2}, 杨鸿强¹

(1. 重庆师范大学 计算机与信息科学学院, 重庆 401331; 2. 重庆国家应用数学中心, 重庆 401331)

摘要: 现有基于深度学习的视网膜血管分割方法因仅依赖逐像素掩码监督, 难以有效建模血管网络的全局连通性和复杂分支结构, 常导致分割结果中出现血管断裂等结构性误差。为此, 本文提出了一种基于 U-Net 的图结构化辅助监督视网膜血管分割网络。通过血管骨架引导的结构编码算法, 将关键分叉点与连通关系转换为结构化监督信号, 引导网络学习视网膜血管的拓扑结构特征。设计了多层次信息流注意力融合模块, 有效融合视觉语义特征和拓扑结构特征, 提升网络的结构表达能力。利用辅助监督机制, 引导网络同时学习拓扑结构特征和视觉语义特征, 以兼顾分割精度与血管形态一致性。在 CHASEDB1、DRIVE 和 STARE 数据集上的实验结果表明, 本文网络在精准分割视网膜血管的同时, 可有效减少血管断裂问题, 为保障分割结果的生物学一致性提供了参考。

关键词: 视网膜血管分割; 深度学习; U-Net; 图结构化辅助监督; 血管骨架; 拓扑结构; 多层次特征融合; 连通性
中图分类号: TP391.4 **文献标志码:** A **文章编号:** 1673-4785(2026)05-1166-14

中文引用格式: 吕佳, 杨鸿强. 图结构化辅助监督的视网膜血管分割网络 [J]. 智能系统学报, 2026, 21(5): 1166–1179.

英文引用格式: LÜ Jia, YANG Hongqiang. U-Net-based graph-structured auxiliary supervision network for retinal vessel segmentation[J]. CAAI transactions on intelligent systems, 2026, 21(5): 1166–1179.

U-Net-based graph-structured auxiliary supervision network for retinal vessel segmentation

LÜ Jia^{1,2}, YANG Hongqiang¹

(1. College of Computer and Information Sciences, Chongqing Normal University, Chongqing 401331, China; 2. National Center for Applied Mathematics in Chongqing, Chongqing 401331, China)

Abstract: Existing deep learning-based retinal vessel segmentation methods, relying exclusively on pixel-wise mask supervision, encounter significant challenges in effectively modeling intricate branching patterns and overall global connectivity features, which leads to structural anomalies such as vascular discontinuities in the segmentation results. Thus, a U-Net-based graph-structured auxiliary supervision network for retinal vessel segmentation is presented in this paper. Firstly, using a vessel skeleton-guided structural encoding algorithm, key bifurcation points and connectivity relationships are transformed into structural supervision cues, enabling the network to capture the global structural attributes of retinal vessels. Secondly, to achieve effective integration of visual-semantic and topological structural features, a multi-level information flow attention fusion module is designed, enhancing the network's structural representation capability. Finally, an auxiliary supervision mechanism is utilized to guide the network in jointly learning topological structural features and visual-semantic features, balancing segmentation accuracy with vessel morphological consistency. Evaluation on CHASEDB1, DRIVE and STARE datasets reveals that the proposed network achieves accurate retinal vessel segmentation, while effectively mitigates vascular discontinuities issues, thereby providing a reference for preserving biological consistency of segmentation outcomes.

Keywords: retinal vessel segmentation; deep learning; U-Net; graph-structured auxiliary supervision; vessel skeleton; topological structure; multi-level feature fusion; connectivity

视网膜血管作为全身微循环系统之一, 对全身生理代谢状态具有高敏感性^[1]。任何生理代谢

的异常或失衡都会在视网膜血管结构上引发相应的病理生理反应, 从而导致其形态特征发生变化^[2-3]。这些形态上的变化能够直接反映个体是否存在糖尿病视网膜病变、高血压和心血管疾病等^[4-5], 为疾病的早期干预提供了关键依据。作为眼底图像分析的关键预处理步骤, 准确且高效的

收稿日期: 2025-09-30. 网络出版日期: 2026-06-09.

基金项目: 国家自然科学基金重大项目 (11991024); 重庆市自然科学基金创新发展联合基金重点项目 (CSTB2025NSCQ-LZX0074); 重庆市研究生科研创新项目 (CYS260405).

通信作者: 吕佳. E-mail: lvjia@cqu.edu.cn.

视网膜血管分割关系到临床诊断的可靠性。然而,传统的视网膜血管分割方法高度依赖于临床医生的手动操作,耗时耗力,严重制约了大规模疾病筛查的实施。因此,研发高效、准确的自动化视网膜血管分割技术对于疾病筛查具有极其重要的作用^[6]。

得益于深度学习技术的飞速发展,特别是卷积神经网络(convolutional neural network, CNN)的广泛应用,视网膜血管分割任务的逐像素精度已取得较大的提升^[7-8]。基于CNN的U-Net^[9]因其跳跃连接设计,能够有效融合网络浅层的结构信息和深层的高级语义信息,在视网膜血管分割任务上取得了不错的效果。然而,受限与卷积核的形状,常规卷积核难以充分提取形态各异的血管特征。为此,缪林一等^[10]将可变形卷积引入U-Net替代常规卷积,自适应调整卷积核,以捕捉不同形态血管的特征,分割效果得到改善。同时,受限与感受野的大小,CNN难以捕获全局上下文信息,吴晨玥等^[11]将空洞卷积添加到网络当中,在不增加参数量的情况下捕获更大范围的上下文信息。然而,传统CNN缺乏跨通道或跨空间维度上的动态权重机制,难以区分重要特征与无关特征,为了使网络更加关注血管区域并抑制噪声,易三莉等^[12]将空间注意力机制引入到U-Net网络当中,使网络能够学习并关注图像中的重要区域,提高模型在复杂背景下的泛化能力。此外,U-Net通过跳跃连接将特征图进行直接拼接的方式难以有效保留语义信息。为此,Yi等^[13]通过引入双向ConvLSTM(convolutional long short-term memory)和非线性操作整合特征,替代U-Net中简单拼接,更高效地融合编码器的细节特征与解码器的高级语义特征,强化特征间的关联性。然而,CNN在捕获长距离依赖信息方面存在的局限却难以克服。具体而言,CNN在处理全局上下文信息时表现出难以整合较远像素之间的关联性,导致易出现误判和遗漏问题,为了缓解这个问题,Jia等^[14]采用Transformer与CNN串行结合的方式构建网络,在利用CNN提取局部特征同时,借助Transformer的全局建模优势来捕获长距离依赖信息,提升视网膜血管分割准确率。然而,传统Transformer参数量大,依赖大量数据进行训练且对硬件依赖性强,在医学影像等小数据集场景中易出现过拟合和部署困难的问题。为此,Tong等^[15]引入MobileViT(mobile vision Transformer)实现CNN与Transformer的轻量化深度融合,权衡高精度与实时性。此外,为应对Transformer在

处理长序列时存在计算效率较低的问题,Li等^[16]通过Mamba语义感知模块,在提高计算效率的同时,有效地融合全局与边缘特征,从而提升了低对比度细小血管的分割准确性。

为进一步提升精细化分割的性能,研究者们从多个维度展开深入研究。在特征提取与融合方面,研究者们设计了更精巧的网络结构,例如,张蓉蓉等^[17]提出的双路径编码网络通过多向差分残差模块增强了对细微血管的捕捉,并通过跨级特征解耦解决了特征混淆问题。在全局上下文建模方面,为克服传统CNN受限与局部感受野的问题,研究者们开始将Transformer和最新的状态空间模型引入分割网络。例如,彭圆圆等^[18]将Mamba结构与CNN结合,旨在实现计算效率与全局特征提取能力的更优平衡。此外,部分研究从训练数据出发,如王植炜等^[19]利用生成对抗网络对噪声标签进行优化,以减少低质量标注对模型训练的干扰。然而,上述方法强调视网膜血管高级语义信息的建模,重点关注低对比度、模糊边界以及血管尺度多变等常规的精细化分割难点,却忽视了视网膜血管结构信息的重要性,导致其连通性和生物学一致性无法保持。对于血管这类网状结构而言,其分支结构以及连通性尤为重要,连通性缺失会导致血流动力学改变,进而影响疾病判断的准确性^[20]。因此本文研究在保持视网膜血管分割精度的前提下,更关注网络结构建模能力,从而增强视网膜血管连通性与生物学一致性。

为增强网络对视网膜血管分割结果的连通性,现有研究从增强网络的结构建模能力出发,主要沿3条思路展开。

1) 通过引入图卷积神经网络(graph convolution network, GCN)来加强模型对视网膜血管结构的感知。如Wang等^[21]提出部分类激活映射(partial class activation mapping, PCAM)引导的图卷积级联U-Net,通过GCN构建动态图并引入类激活先验以抑制噪声传播并强化血管特征,提升细小血管的连通性。

2) 通过设计结构感知增强模块来加强网络对视网膜血管的结构建模能力。如Zeng等^[22]通过注意力门控(attention gates, AGs)模块直接筛选视网膜血管特征,使用柯尔莫哥洛夫-阿诺尔德网络(Kolmogorov-Arnold network, KAN)替换多层感知机(multilayer perceptron, MLP)来改进非线性建模能力,辅以损失对网络进行优化,提升微小血管、复杂分支血管的连通性。

3) 设计全新的损失函数来对网络优化的方向进行引导。深度学习网络会根据最终目标,利用反向传播来对网络内部参数进行更新。此类方法通过重新定义监督信号,将网络的优化目标从单纯的像素点分类转向关注其结构属性。例如,Shit 等^[23]提出 clDice(centerline dice)损失函数,通过骨架与掩码的重合度来强化中心线连续与连通性。Lux 等^[24]提出组件图损失,通过识别图中会造成拓扑错误的组件来计算损失,优化网络对拓扑关键区域的预测,以此增强预测结果的连通性。Schacht 等^[25]提出 Closs(contextloss),通过利用上下文信息来计算损失,优化网络的拓扑保持能力,以此实现更准确的拓扑结构分割。Zhou 等^[26]提出血管结构补全损失,通过在训练中随机遮挡部分血管并迫使网络预测完整结构来计算损失,优化网络对断裂区域的补全能力,以此增强最终分割结果的完整性和连通性。

前两类方法虽然在结构理解方面有所加强,但监督层面大多仍关注血管像素的类别是否正确。这种以像素类别为重点的监督方式,在处理复杂的血管拓扑结构、微小分支以及交叉区域时,往往难以确保预测结果的全局连通性,进而导致分割断裂或错误的连接。相比之下,第三类方法深入探究深度学习网络的内在优化机制,通过对监督信号进行隐式结构转换来提升连通性,然而,这类方法所转换的标注通常未能显式地表明视网膜血管的连接关系^[23-24]。具体而言,血管的端点、分叉点以及不同路径之间的可达性等关键的拓扑结构信息未作为明确的训练标注被直接提供给网络。这种间接的监督方式,使得网络的优化过程更容易倾向于局部区域的纹理与形状一致性,而非对血管分支连通性进行直接且精细的学习与建模,从而限制了其在全局连通性上的进一步提升。

基于上述分析,本文提出了一种基于 U-Net 的图结构化辅助监督视网膜血管分割网络(U-Net-based graph-structured auxiliary supervision network for retinal vessel segmentation, GASU-Net),使用显式的图结构化标注进行辅助监督,旨在引导网络在学习逐像素特征的同时,有效建模并保留血管的连通性与分支结构,从而提升网络分割结果的连通性和生物学一致性。本文的主要贡献如下:

1) 提出一个由血管骨架引导的结构编码(vessel skeleton-guided structural encoding, VSGSE)算法。该算法通过识别血管分叉点和端点,并以此构建图结构化标注,旨在显式表征血管拓扑分支结构信息,为网络提供显式的结构关系标注。

2) 设计多层次信息流注意力融合(multi-level information flow attention fusion, MIFAF)模块。该模块能够有效融合视觉语义特征和拓扑结构特征,增强网络对视网膜血管结构的建模和感知能力。

3) 采用图结构化辅助监督(graph-structured auxiliary supervision, GAS)机制。在保留逐像素掩码(pixel-wise mask)作为监督信号的同时,引入图结构化标注对网络进行辅助监督,以促进网络对血管分支结构、连通性等高层次拓扑结构特征的学习与表征。

1 基本原理

1.1 视网膜血管的拓扑结构

视网膜血管网络状的生物学结构决定了其在形态学分析中具有显著的拓扑特性。其本质上是一个由主血管、分支血管和毛细血管构成的复杂树状结构,其核心元素包括:明确的分叉点、端点以及连接这些关键节点的血管路径和连通路程。这些拓扑元素共同定义了血管的全局连通性及其在视网膜中的空间分布模式。图 1 给出了视网膜血管网络中的典型拓扑元素,图 1(b)用不同颜色的标记标示了分叉点和端点,白色像素表示视网膜血管骨架,而绿色线条则表示血管连通路程,图 1(c)则展示了原图血管的拓扑结构。精准识别并建模这些分叉点、端点及其连接的血管路径,对于评估视网膜健康状况、诊断如糖尿病视网膜病变和高血压性视网膜病变等疾病具有重要的临床价值。

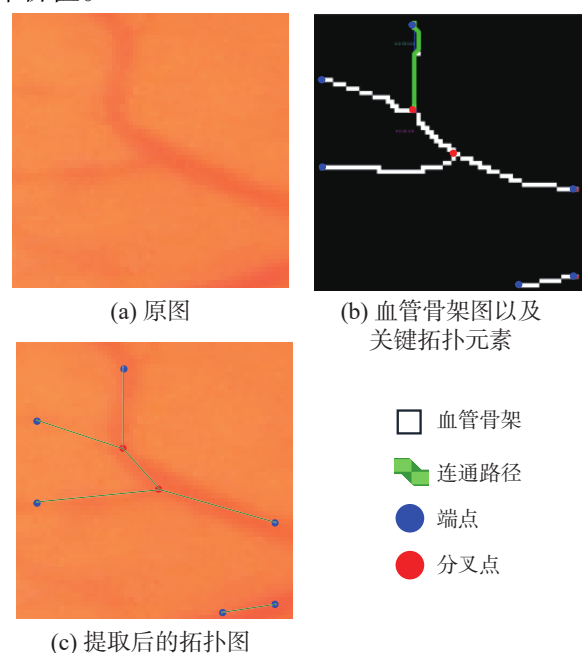


图 1 视网膜血管网络的拓扑结构

Fig. 1 Topological structure of retinal vascular network

2.1.1 关键节点识别

KNI 算法首先对输入的血管掩码图像 I 进行骨架化处理, 得到骨架图 I' 。本文采用 Zhang-Suen 细化算法^[27] 将血管区域简化为单像素宽的骨架线, 进而用于后续的关键点识别。

在骨架化之后, KNI 算法识别血管网络中的分叉点和端点加入集合 V_{topo} 。具体而言, KNI 算法将分析骨架上每个像素的局部连通性, 初步识别出连通性等于 3 的像素点作为潜在的分叉点, 以及连通性为 1 的像素点作为潜在的端点, 这些潜在点命名为 $P_i (i \in \{1, 2, \dots, n\})$ 。接下来 KNI 算法应用了一系列筛选策略, 对 P_i 进行筛选。这包括对空间上相近的多个 P_i 进行聚合处理, 模拟移除中心像素并观察连通分量数量变化, 并根据变化量为其指定一个“目标连接度”值, 以指示该节点在最终拓扑图中应具有的连接数量, 最后若“目标连接度”大于 2 则将 P_i 加入 V_{topo} 。图 3 给出了关键点识别步骤, 算法 1 为 KNI 算法的伪代码。

算法 1 关键节点识别算法

输入 血管掩码图像 I

输出 图 G (包含点集不包含边集)

- 1) 对 I 进行骨架化, 得到骨架图像 I'
- 2) 初始化空图 G
- 3) 从 I' 中识别关键点 P_i
- 4) **for** 每个 P_i
- 5) 确定 P_i 的目标度数
- 6) **if** P_i 目标度数大于 2
- 7) 将 P_i 加入 G 的节点集 V_{topo}
- 8) **End if**
- 9) **End for**

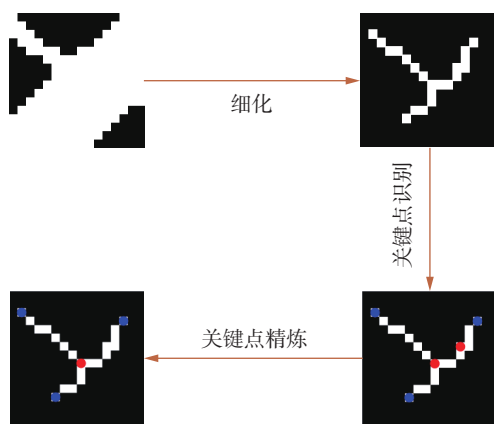


图 3 关键点识别步骤

Fig. 3 Key node identification steps

2.1.2 拓扑连接构建

在关键节点确定之后, TCC 算法开始进行节点的连接。图 4 给出了拓扑连接构建步骤。

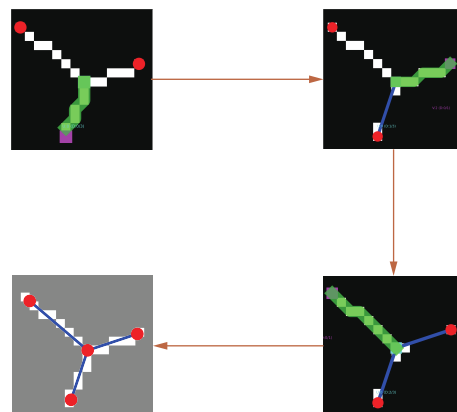


图 4 拓扑连接构建步骤

Fig. 4 Topological connection construction steps

TCC 算法通过在骨架像素构成的图像空间中执行广度优先搜索实现。TCC 算法首先会创建一个原骨架图 I' 的副本 I'_c , 并从关键点集 V_{topo} 中选择一个未满足其目标连接度的关键点 U , 沿着骨架线副本 I'_c 搜索可能的连接目标关键点 V' 。为确保所构建的边独立且不重叠, 一旦路径 $P_{UV'}$ 被确认为图的边, 该路径上的骨架像素都会从 I'_c 中删除, 从而引导后续搜索寻找新的连接路径。TCC 将重复上述过程, 直至所有节点的连接需求得到满足。最终, 血管网络被转换为一个拓扑图。算法 2 是 TCC 算法所对应的伪代码。

算法 2 拓扑连接构建算法

输入 图 G (包含点集不包含边集), 骨架图像 I'

输出 图 G (包含点集和边集)

- 1) 创建 I' 的副本 I'_c
- 2) **While** 图 G 中有节点未达目标度数时
- 3) 选择为未达标节点 U 作为起点
- 4) **if** U 为空
- 5) **break**
- 6) **End if**
- 7) 选择未达标节点 V'
- 8) 沿 I'_c 最短路径 $P_{UV'}$
- 9) **if** 找到路径
- 10) 在 G 中添加 (U, V')
- 11) 更新 G 和 V' 的度数
- 12) 从 I'_c 移除 $P_{UV'}$ 像素
- 13) **End if**
- 14) **End while**

2.1.3 拓扑图张量表示

为了使提取的拓扑结构信息能够被深度学习网络有效利用, VSGSE 算法需要将拓扑图转换为 3D 张量作为最终的标注。参考 He 等^[28] 提出的一种将图转换为 3D 张量的方法, 本文针对视网膜血管的特点对其进行了优化。

TGTR 算法将图 $G=(V,E)$ 转换为特定尺寸的三维特征图, 其中 H 和 W 分别是特征图的长和宽, 对应掩码的长和宽。通道数 C 设定为 13。

如图 5 所示, 图中每个节点 $v \in V$ 都被映射到唯一空间坐标 $(x, y) \in \{0, \dots, H-1 \times 0, \dots, W-1\}$ 上。每个位置的向量由 $1+4 \times 3=13$ 个值构成。特征图的第一个元素 $P_v \in [0,1]$ 用于编码当前位置存在节点的概率。之后由 4 组 3 元素向量组成, 每一组的第一个元素 $P_i \in [0,1]$ 用于编码存在相连节点的概率, 紧跟的 Δx_i 和 Δy_i 代表相连节点的位置, 使用相对位置来进行编码, $i \in \{1, 2, 3, 4\}$, i 越大节点存在边的节点距离当前节点越远。根据视网膜血管的生长特性, 每个分叉点最多会出现 3 个分支^[29], 但考虑到动静脉在掩码图上会出现重叠的情况, 编码时多添加一组冗余组, 以应对这种情况。

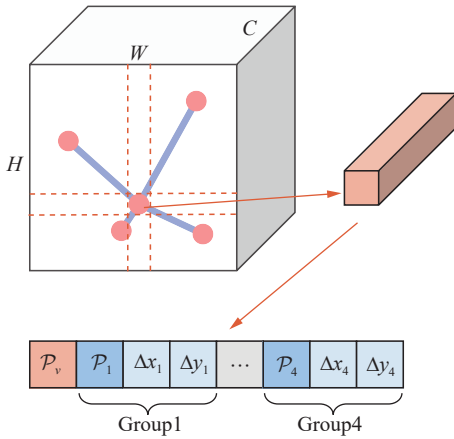


图 5 三维特征图

Fig. 5 3D feature map

在训练过程中, 将此三维拓扑张量作为 TSE 的学习目标, 通过最小化 TSE 的预测输出与此目标张量之间的损失, 迫使网络参数进行调整与优化, 将 TSE 从一个通用的视觉特征提取器, 特化为一个提取血管分叉、端点等拓扑元素拓扑特征的编码器。

2.2 多层次信息流注意力融合模块

MIFAF 是 GASU-Net 实现不同类型不同尺度特征融合的重要组成部分, 具体结构如图 6 所示。该模块将 TSE 提取的拓扑结构特征作为查询信号, 引导模块从 VSE 获取的视觉内容特征中提取与血管结构相关的语义信息。由于拓扑结构特征与语义特征在空间分布和语义层级上存在差异, 若不进行对齐而直接融合, 两类特征在空间上的不匹配会产生语义冲突, 导致网络无法正确利用拓扑先验来增强视觉特征。

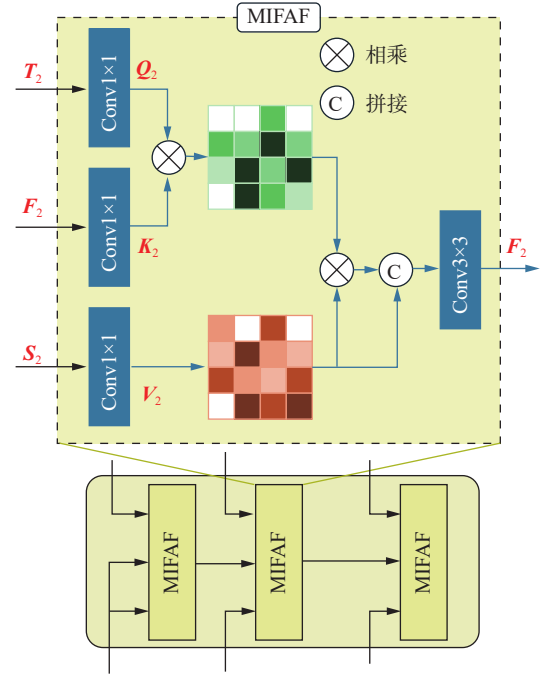


图 6 多层次信息流注意力融合模块

Fig. 6 Multi-level information-flow attention fusion module

TSE 输出的拓扑结构特征被转换为查询 (Query, Q) 信号, 而 VSE 提供的多尺度视觉特征则同时转换为键 (Key, K) 和值 (Value, V)。通过计算 Q 与 K 之间的注意力分数, 通过注意力机制, MIFAF 能够根据拓扑结构特征, 在视觉语义特征中寻找最匹配的模式。通过将注意力分数应用于 V , 使得视觉语义特征中与拓扑结构特征高度对齐的特征进行加权和筛选。

设在第 i 个融合阶段 ($i \in \{1, 2, 3\}$), VSE 提取的语义信息为 $F_{S,i} \in \mathbb{R}^{H \times W \times C}$, TSE 提取的拓扑结构特征为 $F_{T,i} \in \mathbb{R}^{H \times W \times C}$ 。MIFAF 模块首先使用 1×1 卷积将 $F_{T,i}$ 和 $F_{S,i}$ 投影为查询 (Q_i) 和键 (V_i)。生成方式为

$$Q_i = W_{q,i} F_{T,i}, V_i = W_{k,i} F_{S,i}$$

式中: $W_{q,i}$ 和 $W_{k,i}$ 为可学习的权重矩阵。而值 (K_i) 则根据融合阶段的不同有所区分, 当 $i=1$ 时, 融合模块对 $F_{S,1}$ 进行线性映射得到 K_1 。当 $i>1$ 时, 融合模块使用上一个融合的特征作为原始特征, 对其进行线性映射后作为 K_i 。计算公式为

$$K_i = \begin{cases} W_{v,1} F_{S,1}, & i = 1 \\ W_{v,i} \text{Concat}(F_{S,i}, F_{\text{fusion},i-1}), & i > 1 \end{cases}$$

式中: $W_{v,1}$ 和 $W_{v,i}$ 是各自阶段对应的投影矩阵, $\text{Concat}(\cdot)$ 表示特征拼接操作, $F_{\text{fusion},i}$ 表示第 i 阶段融合后的特征。随后, 模块通过点积计算注意力分数, 并进行归一化, 从而获得结构化加权的语义特征 F_A , 计算过程为

$$F_{A,i} = \text{Softmax} \left(\frac{Q_i K_i^T}{\sqrt{d_k}} \right) V_i$$

式中 d_k 是 Key 向量的维度, 用于缩放点积结果并稳定梯度。之后, $F_{A,i}$ 与 V_i 进行拼接, 并使用卷积操作进行深度融合, 得到最后的 $F_{\text{fusion},i}$ 。计算公式为

$$F_{\text{fusion},i} = \text{Conv}(\text{Concat}(F_{A,i}, V_i))$$

式中 $\text{Conv}(\bullet)$ 表示卷积操作。通过引入注意力机制, 网络能够根据拓扑结构特征的引导, 自适应地筛选和强化视觉语义特征中与血管相关的部分, 同时减少噪声的干扰。

2.3 特征解码器

FDecoder 用于逐步重建空间信息, 并生成最终血管分割掩码。整个解码过程分为 3 个阶段 ($j \in \{1, 2, 3\}$), 每个阶段均涉及两次不同类型特征图的拼接与卷积操作, 具体结构如图 7 所示。

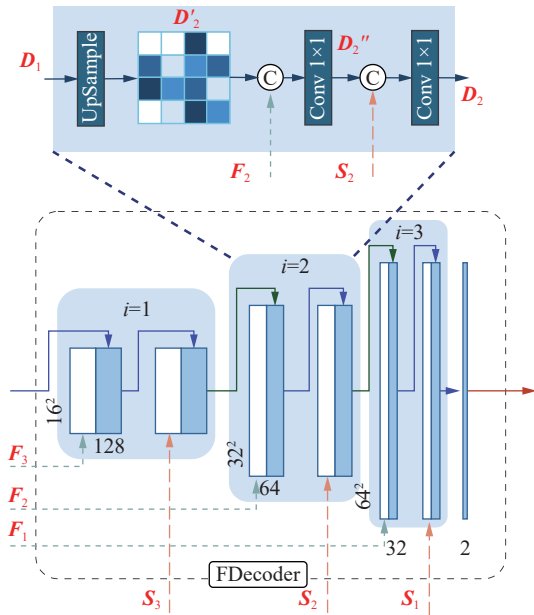


图 7 特征解码器

Fig. 7 Feature decoder

在每个解码阶段, 首先进行上采样操作, 将之前层的特征 D_j 提升至更高分辨率得到 D_{j+1} 。随后, D_{j+1} 和来自对应 MIFAF 阶段的输出 $F_{\text{fusion},4-j}$ 进行拼接, 这一步旨在加强筛选后的融合特征, 加强血管特征。接着进行一次卷积操作, 得到 D'_{j+1} , 旨在利用深度融合后的语义信息引导特征恢复。计算公式为

$$D'_{j+1} = \text{UpSample}(D_j)$$

$$D''_{j+1} = \text{Conv}(\text{Concat}(D'_{j+1}, F_{\text{fusion},4-j}))$$

得到的 D''_{j+1} 接着与来自 VSE 的对应层级的原始非融合视觉语义特征 S_{4-j} 进行再次拼接。对拼接后的特征再进行一次卷积操作, 将融合后的特征进行更精细的特征还原, 最终完成当前解码阶段的特征恢复, 得到该阶段的输出 D''_{j+1} 。计算公

式为

$$D_{j+1} = \text{Conv}(\text{Concat}(D''_{j+1}, S_{4-j}))$$

这种双重拼接策略旨在整合融合后的高级语义信息, 并补充 VSE 提供的原始、未经抽象的低级细节, 以有效恢复因上采样可能丢失的空间信息。再次拼接后的特征流会通过第二组基础卷积操作进行处理, 完成当前解码阶段的特征恢复。

FDecoder 从瓶颈层的集成特征开始, 逐层向上进行上采样和特征融合, 直至恢复到与原始输入图像相同的分辨率。最终, 这些高分辨率的特征图通过一个 1×1 卷积层进行最终的特征映射, 生成血管的预测概率图。该概率图随后经过 Softmax 激活函数处理, 输出每个像素属于血管或背景的概率分布, 从而形成最终的精细化血管分割掩码。

2.4 图结构化辅助监督

在 GAS 机制中, 监督信号通过损失函数被量化, 这些损失函数作为衡量网络性能和引导网络优化的关键途径。网络的总损失函数 L_{Total} 由血管分割损失函数 L_{Seg} 和拓扑结构监督损失函数 L_{Graph} 通过加权融合构成, 其计算公式为

$$L_{\text{Total}} = L_{\text{Seg}} + \lambda L_{\text{Graph}}$$

式中 λ 为可学习的自适应权重因子, 用于动态调控拓扑结构监督损失对整体训练目标的贡献, 从而在语义特征与拓扑结构二者之间保持平衡。

血管掩码分割任务采用 BCE Loss 量化监督。计算公式为

$$L_{\text{Seg}} = \text{BCE}(g_i, p_i)$$

式中: g_i 表示像素 i 的真实类别, p_i 则表示像素 i 预测为正类的概率。

为了引导网络学习血管网络的拓扑结构, 本文设计了一个专用的拓扑结构监督损失 L_{Graph} 。该损失函数用于 TSE 的输出与预构建的真实图张量之间, 并根据图张量中不同通道的语义信息, 采用不同的损失计算方式。

图张量包含 13 个通道, 可分为两类: 一类是概率通道, 即节点存在概率 (通道 0) 和连接概率 (通道 1、4、7、10)。另一类是位置通道表示相对位置偏移量 (通道 2、3、5、6、8、9、11、12)。

对于概率通道, 使用 BCE Loss 来衡量预测概率与真实标注之间的差异。其中, 通道 0 (节点存在概率) 的损失单独计算, 其他连接概率通道的损失则统一计算。计算公式为

$$L_{\text{node}} = \text{BCE}(P_{\text{pred},0}, P_{\text{gt},0})$$

$$L_c = \text{BCE}(P_{\text{pred},c}, P_{\text{gt},c})$$

式中: $P_{\text{pred},0}$ 、 $P_{\text{gt},0}$ 分别为预测和真实的节点存在概

率图。 $P_{\text{pred},c}$ 和 $P_{\text{gt},c}$ 分别为预测和真实的连接概率图, $c \in \{1, 4, 7, 10\}$ 表示连接概率通道。

对于位置通道, 使用均方误差损失 (mean squared error loss, MSE Loss) 来衡量预测偏移量与真实偏移量之间的平方误差。真实偏移量在此处会根据图像尺寸进行缩放归一化, 以匹配网络预测的范围。计算公式为

$$L_{\text{pos}} = \text{MSE}(\tanh(P_{\text{pred}}^{\text{os}}), P_{\text{gt}}^{\text{os}})$$

式中: $P_{\text{pred}}^{\text{os}}$ 和 $P_{\text{gt}}^{\text{os}}$ 分别为预测和真实的相对位置偏移量。 $P^{\text{os}} \in \{2, 3, 5, 6, 8, 9, 11, 12\}$ 表示相对位置偏移量所在通道。为确保预测图张量与真实图张量在空间分辨率上保持一致, 真实图张量会根据预测图张量的尺寸进行自适应调整。对于位置通道的预测值, 网络会施加 $\tanh$ 激活函数以限制其输出范围, 并对真实位置目标进行归一化处理。

最终, 拓扑结构监督损失 L_{Graph} 由上述各项损失相加而成, 计算公式为

$$L_{\text{Graph}} = L_{\text{node}} + L_{\text{conn}} + L_{\text{pos}}$$

3 实验与结果分析

3.1 数据集及预处理

为验证网络有效性, 本文选择在 CHASEDB1^[30]、DRIVE^[31] 和 STARE^[32] 3 个公开数据集上进行训练和测试。

为标准化输入数据并增强视网膜血管特征, 本文对所有数据集进行了预处理操作。本文对图像的绿色通道执行了对比度有限的自适应直方图均衡化处理。该方法能够有效增强图像的局部对比度, 同时抑制噪声的过度放大, 从而改善血管的视觉可辨识度。同时, 本文还对绿色通道进行了伽马校正, 进一步优化血管特征的提取。此外, 为保证实验结果的可比性和标准化, 训练集和测试集的划分比例按照数据集官方预定义的方式进行划分。详细划分比例见表 1 最后一列。

3.2 评估指标

为定量评估本文网络, 实验选择了两类指标进行实验, 一类是逐像素指标, 分别是平均交并比 (mean intersection over union, mIoU)、曲线下面积 (area under curve, AUC)、准确率 (accuracy, Acc)、

特异度 (specificity, Sp)、灵敏度 (sensitivity, Se)、F1 分数 (F1-score, F1) 和马修斯相关系数 (Matthews correlation coefficient, Mcc)。另一类是结构级指标, 专门针对管状结构分割设计的拓扑结构保持指标 cDice^[23]。此外还对比了各网络模型的参数数量 (parameters, params) 和每秒十亿次浮点运算次数 (giga floating-point operations per second, GFLOPs)。

表 1 数据集设置
Table 1 Setting of dataset

数据集	分辨率	图片数量	增强后数量	训练:测试
CHASEDB1	999×960	28	5 000	20:8
DRIVE	565×584	40	5 000	20:20
STARE	700×605	20	3 750	15:5

3.3 实验室环境及指标

实验采用 Python 语言编程, 基于 PyTorch 框架实现, 网络在配备 32 GB 内存的 Intel(R) Xeon(R) Gold 5218R CPU 和 16 GB NVIDIA Tesla T4 显卡的计算机上运行, 部署环境为 Ubuntu 20.04 操作系统。网络输入的图像尺寸为 64×64, 使用随机裁切进行数据增强, 每张图片被裁切为 250 张 64×64 尺寸的小图用于训练。网络训练轮次总数设置为 500 轮, 设置早停机制, batch size 设置为 16, 采用优化算法 Adam 训练整体网络, 初始学习率设置为 0.000 3, 并利用余弦退火策略辅助训练。自适应权重因子 λ 设置为 1。

3.4 实验结果与分析

3.4.1 消融实验

为验证各关键模块和监督信号对 GASU-Net 性能的有效性和实际贡献, 本文设计了 5 组消融实验, 实验结果如表 2 所示, 表中加粗数据代表同一指标下的最优结果, 下划线数据代表同一指标下的次优结果, $\uparrow$ 表示指标越大越好, 上述标注规则同样适用于本文后续所有表格。主要对比 5 组网络组合, 辅助监督方式分别考察了无辅助监督、PCAM 辅助监督和 GAS, 三者为互斥关系, 此外, 还探讨了有无 MIFAF 模块对分割性能的影响。

表 2 在 CHASEDB1 数据集上的不同模块组合的实验结果
Table 2 Experimental results for different combinations of modules on the CHASEDB1 dataset

辅助监督方式			MIFAF	mIoU($\uparrow$)	AUC($\uparrow$)	Acc($\uparrow$)	F1($\uparrow$)	Mcc($\uparrow$)	cDice($\uparrow$)
无辅助监督	PCAM	GAS							
$\checkmark$				0.662 8	0.940 4	0.961 9	<u>0.804 4</u>	<u>0.775 7</u>	0.801 5
		$\checkmark$		0.660 7	0.948 9	<u>0.962 8</u>	0.795 7	0.775 2	<u>0.809 7</u>
$\checkmark$			$\checkmark$	0.649 5	<u>0.951 2</u>	0.961 2	0.787 5	0.766 2	0.794 8

续表 2

辅助监督方式			MIFAF	mIoU(↑)	AUC(↑)	Acc(↑)	F1(↑)	Mcc(↑)	cIDice(↑)
无辅助监督	PCAM	GAS							
	√		√	0.604 1	<u>0.951 2</u>	0.953 8	0.753 2	0.728 2	0.763 9
		√	√	0.678 7	0.956 3	0.965 2	0.808 6	0.789 4	0.826 4

注: 加粗数据代表同一指标下的最优结果, 下划线数据代表同一指标下的次优结果, ↑表示指标越大越好。

第 1 组为基准网络, 在 GASU-Net 基础上去除了 GAS 监督和 MIFAF 模块, 保留其余结构。此时网络主要依赖像素标注或局部区域对比, 缺乏对全局血管结构编码能力, 因此在连通性和分支保持等结构相关指标上表现落后。

在第 2 组实验中, 网络单独使用 GAS 时, 与基准网络相比均有提升, cIDice 和 F1 指标分别提升 0.82 和 0.74 百分点, 说明网络的视网膜血管结构保持能力有所加强。

第 3 组实验中, 当 MIFAF 单独启用时, 网络效果较第 1 组基准网络有所下降, 表明若网络缺乏有效的监督指导, 模块可能会起反向作用, 过多相似的语义信息反而会对网络的准确判断构成干扰, 并导致其整体性能的下降。

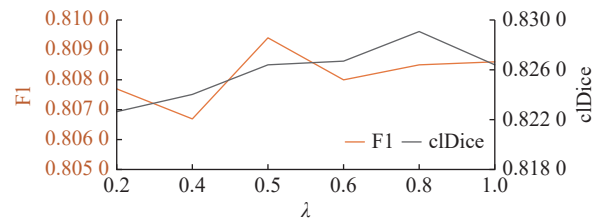
在第 4 组采用 PCAM 辅助监督时, 血管分割结果连续性提升有限, 原因在于 PCAM 主要以抑制噪声传递、强化血管特征为主要约束, 但在结构建模约束方面仍不及图结构化监督的直接引导。

在第 5 组采用 GAS 和 MIFAF 模块组合时, 网络性能达到最优。cIDice、mIoU、F1 等结构性和综合性指标上均取得最佳表现, 突出展示了网络在血管连通性和整体结构一致性方面的优势。证明图 GAS 能够显式捕捉血管分支、端点与连通关系, 为网络提供高阶拓扑结构引导, 而融合模块有助于视觉语义特征和拓扑结构特征的深度集成, 这使得网络兼具逐像素判别力和全局结构建模能力。

由第 2 组和第 3 组实验分析可知, GAS 机制在提供精确指导中的价值, 以及 MIFAF 在促进语义与结构特征有效融合中的作用, GAS 和 MIFAF 共同构成了提升网络结构感知能力、减少冗余信息并实现最佳分割性能的关键协同要素。

为探究不同权值的 λ 对网络性能的影响, 本文设计了 6 组关于 λ 的消融实验, 取 $\lambda \in \{0.2, 0.4, 0.5, 0.6, 0.8, 1.0\}$, 实验结果如图 8 所示。当 $\lambda=0.8$ 时, GASU-Net 在 cIDice 指标上取得最佳性能, 当 $\lambda=0.5$ 时, GASU-Net 在逐像素指标 F1 上取得最佳性能。这表明逐像素分割精度与拓扑结构保持之间存在相互影响的关系^[33]。 λ 过小, GAS 机制的作用被减弱, 血管断裂问题改善有限, 难

以保持血管的连通性。而 λ 过大, 网络会过度关注拓扑结构而忽略逐像素的细节, 导致分割精度下降。

图 8 F1 和 cIDice 指标随 λ 的变化Fig. 8 Variation of the F1 and cIDice metric with respect to λ

为定量评估 GASU-Net 对图结构化监督信号的鲁棒性与泛化性, 本文设计了有无监督标签噪声下的网络性能消融实验, 通过在 VSGSE 生成的监督标签加入人工噪声, 随机移除 10% 的节点, 来模拟现实世界中可能存在的标注误差。图 9 给出了使用无噪声标签与带噪声标签训练后, GASU-Net 在 CHASEDB1 测试集上的性能对比。当使用带噪声标签进行训练时, cIDice 指标表现出轻微下降的趋势。这表明, 监督信号的质量对最终的连通性存在影响, 但 GASU-Net 并非完全地依赖于单一监督信号, 所以其性能下降的幅度较小, 这表明 GASU-Net 在学习拓扑结构方面有一定鲁棒性。而 F1 指标相较于无噪声标签组, 呈现出小幅度的提升。此现象可能源于一种隐式的正则化效应, 即在监督标签中加入少量噪声, 能够有效防止网络对训练集中部分拓扑结构和像素细节产生过拟合, 从而增强 GASU-Net 在未知测试集上的泛化能力, 最终体现在逐像素分割性能的改善上。

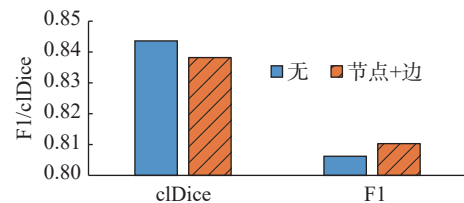


图 9 有无监督标签噪声下的 GASU-Net 性能对比

Fig. 9 Performance comparison with and without noisy supervision labels on GASU-Net

3.4.2 对比实验

为评估 GASU-Net 在视网膜血管分割任务中

的实际效果, 将其在 CHASEDB1、DRIVE 与 STARE 数据集上, 分别与其他主流网络进行了对比实验。实验结果见表 3~5。在 CHASEDB1 数据集上, GASU-Net 在 cIDice、mIoU、F1 和 Mcc 等指标上取得最优, 如表 3 所示。相较于 PGU-Net, GASU-Net 在 cIDice 和 mIoU 分别高出 1.06 和 0.58 百分

点, 这主要得益于 GASU-Net 采用的 GAS 机制, 在这种机制的作用下, 网络会学习视网膜血管的全局结构, 从而有效补全因低对比度、噪声或复杂的血管形态导致的潜在断裂, 并确保血管分支的准确连接, 进一步强化了整体的连通性和形态完整性。

表 3 在 CHASEDB1 数据集上与现有网络的实验结果对比

Table 3 Comparison of experimental results with existing networks on the CHASEDB1 dataset

网络	mIoU	AUC	Acc	Sp(↑)	Se(↑)	F1	Mcc	cIDice	Params/10 ⁶	GFLOPs
U-Net ^[9]	0.663 6	0.981 9	0.962 7	0.978 1	0.811 2	0.797 6	0.777 8	0.821 3	3.35	0.49
U-Net++ ^[34]	0.666 7	0.981 3	0.963 2	0.978 5	<u>0.811 8</u>	0.799 8	0.780 2	0.822 1	9.16	2.16
AttU-Net ^[35]	0.663 6	0.981 9	0.962 9	0.979 0	0.805 0	0.797 4	0.777 7	0.817 7	34.86	4.16
ResU-Net++ ^[36]	0.623 3	0.973 9	0.960 2	<u>0.983 4</u>	0.727 2	0.767 9	0.747 6	0.769 3	14.48	4.43
IterNet ^[37]	0.664 0	0.981 8	0.963 1	0.979 1	0.804 8	0.797 8	0.778 1	0.818 5	13.60	2.70
cldice ^[23]	0.653 1	0.943 4	0.961 5	0.977 5	0.800 5	0.790 2	0.769 1	0.818 9	8.64	1.03
PGU-Net ^[21]	<u>0.672 9</u>	0.985 3	0.965 7	0.984 4	0.778 0	<u>0.804 5</u>	<u>0.786 3</u>	0.815 8	2.17	2.55
AttUKAN ^[22]	0.670 9	<u>0.982 2</u>	0.964 1	0.980 0	0.807 2	0.802 8	0.783 9	<u>0.825 2</u>	24.95	0.62
GASU-Net	0.678 7	0.956 3	<u>0.965 2</u>	0.980 3	0.812 6	0.808 6	0.789 4	0.826 4	4.87	1.81

注: 加粗数据代表同一指标下的最优结果, 下划线数据代表同一指标下的次优结果, ↑表示指标越大越好。

表 4 在 DRIVE 数据集上与现有网络的实验结果对比

Table 4 Comparison of experimental results with existing networks on the DRIVE dataset

网络	mIoU	AUC	Acc	Sp	Se	F1	Mcc	cIDice
U-Net	0.672 9	0.979 5	0.954 8	<u>0.987 5</u>	0.730 8	0.804 5	0.784 3	0.790 9
U-Net++	0.675 9	0.978 1	0.954 4	0.984 8	0.746 4	0.806 6	0.784 3	0.791 5
AttU-Net	0.670 2	0.976 2	0.954 1	0.986 5	0.732 4	0.802 5	0.781 4	0.779 4
ResU-Net++	0.672 6	0.973 2	0.953 7	0.983 8	0.747 4	0.804 2	0.781 1	0.784 2
IterNet	0.689 5	0.975 2	0.955 4	0.981 4	0.777 3	0.816 2	0.792 2	0.802 7
cldice	0.678 1	0.939 1	0.952 7	0.977 3	0.783 8	0.808 2	0.781 7	0.820 8
PGU-Net	0.667 3	<u>0.978 3</u>	0.954 1	0.987 9	0.722 6	0.800 4	0.780 7	0.786 4
AttUKA	<u>0.698 3</u>	0.978 1	<u>0.955 6</u>	0.976 8	0.811 6	<u>0.822 2</u>	0.798 2	<u>0.821 1</u>
GASU-Net	0.699 0	0.933 3	0.956 0	0.978 4	<u>0.802 4</u>	0.822 9	<u>0.798 1</u>	0.822 2

注: 加粗数据代表同一指标下的最优结果, 下划线数据代表同一指标下的次优结果。

表 5 在 STARE 数据集上与现有网络的实验结果对比

Table 5 Comparison of experimental results with existing networks on the STARE dataset

网络	mIoU	AUC	Acc	Sp	Se	F1	Mcc	cIDice
U-Net	0.526 1	0.971 9	0.952 6	0.997 3	0.538 7	0.689 5	0.698 5	0.693 6
U-Net++	0.567 2	0.979 4	0.956 3	0.996 4	0.586 2	0.723 8	0.725 3	0.748 8
AttU-Net	0.536 9	0.972 1	0.953 4	0.996 7	0.553 5	0.698 7	0.704 0	0.726 6
ResU-Net++	0.377 2	0.942 8	0.938 3	0.998 5	0.382 4	0.547 8	0.586 5	0.555 4
IterNet	0.603 3	<u>0.973 6</u>	0.959 1	0.994 3	0.634 9	0.752 6	<u>0.746 5</u>	0.777 5
cldice	<u>0.609 0</u>	0.911 3	0.958 8	0.991 6	0.656 5	<u>0.757 0</u>	0.745 6	0.792 6
PGU-Net	0.477 2	0.965 0	0.947 7	<u>0.997 4</u>	0.488 5	0.646 1	0.661 7	0.635 5
AttUKAN	0.604 2	0.970 6	<u>0.959 2</u>	0.991 4	<u>0.658 1</u>	0.751 5	0.740 8	<u>0.809 6</u>
GASU-Net	0.643 7	0.910 5	0.962 7	0.992 4	0.689 1	0.783 3	0.771 9	0.836 5

注: 加粗数据代表同一指标下的最优结果, 下划线数据代表同一指标下的次优结果。

GASU-Net 的 Mcc 指标较 PGU-Net 高出 0.31 百分点,这是因为 GAS 通过显式地引导网络关注血管的全局结构和连接性,减少了假阳性(将背景误判为血管)和假阴性(将血管漏判为背景)的发生,从而在处理类别不平衡数据时能够达到较好的性能。而 AUC、Acc 和 Sp 指标上相较 PGU-Net 略有下降,这是在于 PGU-Net

更加注重噪声的抑制,更能区分被背景遮挡的细小血管,在 CHASEDB1 这个包含儿童精细视网膜血管的数据集上具有一定优势。从图 10 的第一行放大细节图可以看出,特别是红色方框细节, GASU-Net 在保持连通性上表现较好,其他网络均出现断裂情况难以保持视网膜血管的连通性。

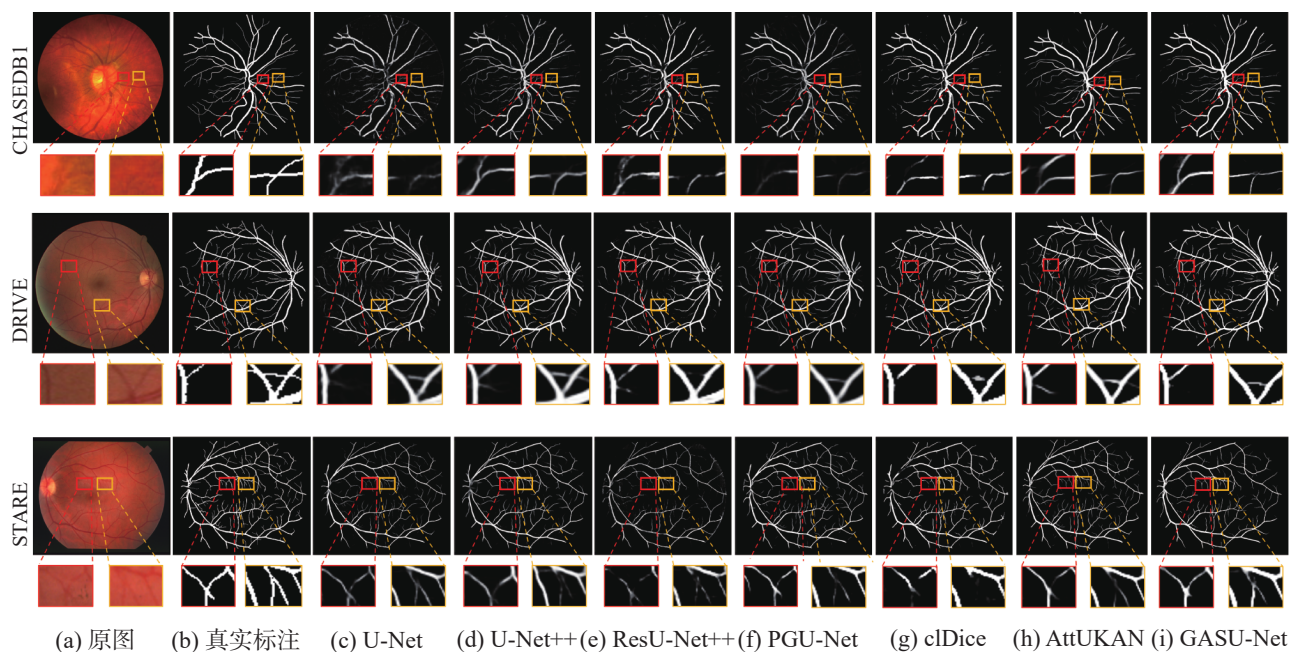


图 10 3 个数据集上不同网络的局部分割结果

Fig. 10 Local segmentation results of different networks on three datasets

相较于 CHASEDB1 数据集, DRIVE 数据集清晰度更高且细小血管相对较少,但包含的病变样本仍构成不小的分割挑战。在 DRIVE 数据集上的对比实验结果如表 4 所示。GASU-Net 的 Acc 指标取得最优,比次优 PGU-Net 高出 0.04 百分点,究其原因在于 DRIVE 数据集相较 CHASEDB1 有更高的清晰度,细小血管数也相较更少,且 MIFAF 在 TSE 特征的指导下从 VSE 中关注重要特征的方式也能够抑制一部分噪声,避免病变区域的干扰。GASU-Net 在 clDice、mIoU 和 F1 指标上依旧保持领先,相较于次优的 AttUKAN 分别提升 0.11、0.07 和 0.07 百分点。尽管 AttUKAN 设计了结构捕获模块在一定程度上提升了分割性能,但在对血管连通性和整体结构完整性至关重要的 clDice 和 F1 指标上, GASU-Net 仍处于领先地位。

这进一步证实了 GASU-Net 在血管分割中通过显式结构建模所带来的优势。从图 10 的第二行放大细节图的红色方框可以看出, GASU-

Net 能够有效避免虚假连接的情况出现,从黄色方框看, GASU-Net 能较好地避免血管断裂的情况出现。

表 5 为各网络在 STARE 数据集上的对比实验结果。STARE 数据当中的病变样本比例较 DRIVE 数据集有所增加,进一步加大了分割难度。尽管如此, GASU-Net 在 clDice、mIoU、Acc、Se、F1 和 Mcc 指标上达到最优。这表明 GASU-Net 的 GAS 在复杂图像背景和病变干扰下,仍能有效引导网络学习血管的连通性和分支结构,且效果相较于前两个数据集表现更好。这也表明了 GASU-Net 通过显式结构建模能够一定程度上提升网络的鲁棒性,尤其是在处理包含大量病变样本时,其能够有效避免单一逐像素监督方法易出现的结构性断裂或错误连接。和从图 10 的最后一行放大细节图的方框可以看出, GASU-Net 能够很好地维持分割出的视网膜血管的连通性,而其他网络都有不同程度的断裂情况出现。更多病变样本下的分割结果如图 11 所示,这些样本包

含视网膜色素上皮萎缩、脉络膜毛细血管层萎缩和硬性渗出物等病理病变。



图 11 GASU-Net 病变样本上的分割效果

Fig. 11 Local segmentation results of GASU-Net on lesion samples

为分析 GASU-Net 在视网膜血管分割这类类别不平衡数据集上的性能, 图 12 给出了 GASU-Net 在 CHASEDB1、DRIVE 和 STARE 3 个公开数据集的测试集上的混淆矩阵。该矩阵在直观展现了视网膜血管分割数据集的类别不平衡数据集的同时, 也展现了 GASU-Net 在识别血管与排除背景这两个对立目标之间的权衡能力。在类别严重不平衡的场景下, Mcc 是评估这种权衡能力的指标, GASU-Net 在两个数据集上均取得了最优的 Mcc 值, 这表明 GASU-Net 在处理类别不平衡数据时能够达到较好的性能。

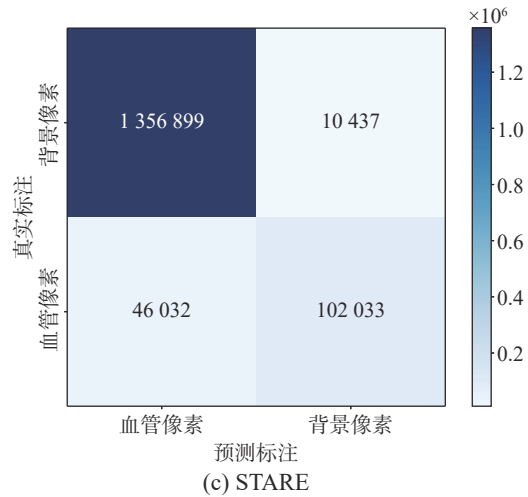
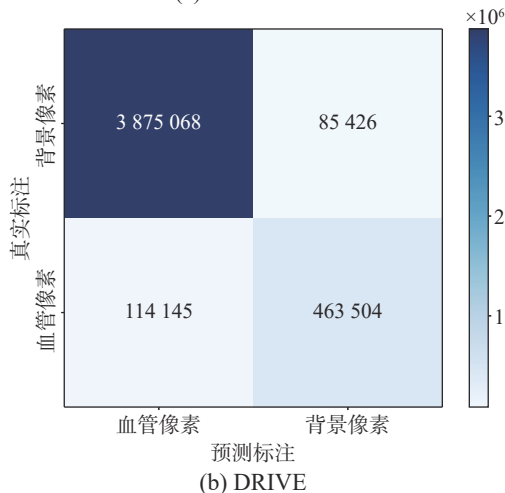
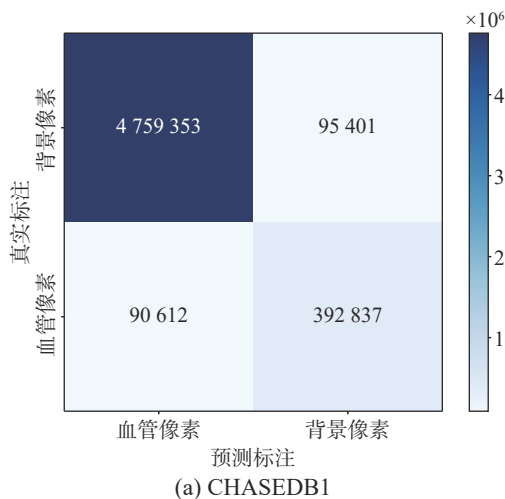


图 12 混淆矩阵对比

Fig. 12 Comparison of confusion matrix.

GASU-Net 在 3 个对比实验结果上均显示其在 cDice、mIoU 和 F1 指标上始终保持领先, 这表明 GASU-Net 具有更好的保持视网膜血管分割结果连通性的能力。mIoU 上的良好表现则进一步印证了该模型在分割区域的准确性和重叠度方面的优势。另外, 面向视网膜血管分割任务的数据集具有类别不平衡的特点, 而 GASU-Net 在 Mcc 指标上达到了两个最优和一个次优, 证明了该网络在精确率和召回率之间的良好平衡性。

值得注意的是, 在部分数据集的 AUC 和 Acc 等逐像素分类精度指标上, GASU-Net 未能取得最优值。究其原因在于 GASU-Net 的首要优化目标是确保分割结果的拓扑结构连通性, 即生成连续、无断裂且分支结构正确的血管树。为实现这一目标, 当 GASU-Net 面临因图像质量、对比度模糊或病变干扰导致的低置信度像素区域时, 它会策略性地倾向于保持拓扑结构连通性, 即会对某些孤立的、无法可靠连接的微弱血管像素进行抑制。这种对结构可靠性的优先级设定, 能够提升 cDice 指标, 但会在一定程度上影响到逐像素分类准确率的 AUC 和 Acc 指标。

4 结束语

本文针对现有分割网络难以有效保持视网膜血管连通性与分支结构的不足, 提出 GASU-Net。为增强网络对血管结构信息的建模能力, 设计了 VSGSE 算法, 并以图结构化张量作为高层次辅助监督信号, 从而显式引导网络学习和保留血管网络的整体结构特征。GASU-Net 通过双编码器单解码器架构, 联合 MIFAF 模块, 实现了语义特征与结构特征的深度融合, 提升了血管分割的连通性。本文在 3 个公开数据集上验证了其有效性。

GASU-Net 虽能有效提升视网膜血管分割的连通性和生物学一致性,但部分实验指标,如 AUC、Acc 指标未能实现持平,这表示网络在边界模糊区域的分割能力仍有不足。为精细化边界信息的建模,在未来工作中,将设计一个边界信息增强建模网络,与 GASU-Net 进行组合,构成一个级联式的双网络协同框架,在保留视网膜血管连通性的同时,对边界信息进行建模,提升视网膜血管的逐像素分割精度,最优化 AUC、Acc 等指标。

参考文献:

- [1] SINGH C. Metabolism and vascular retinopathies: current perspectives and future directions[J]. *Diagnostics*, 2022, 12(4): 903.
- [2] LIU Xinghui, TAN Hongwen, WANG Wu, et al. Deep learning based retinal vessel segmentation and hypertensive retinopathy quantification using heterogeneous features cross-attention neural network[J]. *Frontiers in medicine*, 2024, 11: 1377479.
- [3] CHEN Shuli, XU Yi, CHEN Bo, et al. Remnant cholesterol is correlated with retinal vascular morphology and diabetic retinopathy in type 2 diabetes mellitus: a cross-sectional study[J]. *Lipids in health and disease*, 2024, 23(1): 75.
- [4] SUNG J Y, LEE K H, JUN J H, et al. Changes in peripapillary microvasculature in patients with type 2 diabetes patients: effect of systemic hypertension[J]. *Scientific reports*, 2023, 13: 19459.
- [5] 张菡, 贾红艳, 凌赛广, 等. 70 岁以上冠心病患者基于人工智能的视网膜血管形态参数研究[J]. *眼科*, 2024, 33(4): 285–289.
ZHANG Han, JIA Hongyan, LING Saiguang, et al. Correlation analysis between retinal vascular morphology parameters and 70 years-old or over coronary heart disease based on artificial intelligence[J]. *Ophthalmology in China*, 2024, 33(4): 285–289.
- [6] 吕佳, 王泽宇. 基于深度学习的视网膜血管分割方法综述[J]. *重庆师范大学学报 (自然科学版)*, 2024, 41(4): 110–125.
LÜ Jia, WANG Zeyu. A survey of retinal vessel segmentation algorithms based on deep learning[J]. *Journal of Chongqing Normal University (natural science)*, 2024, 41(4): 110–125.
- [7] 汪有崧, 裴峻鹏, 李增辉, 等. 深度学习的视网膜血管分割研究综述[J]. *计算机科学与探索*, 2024, 18(8): 1960–1978.
WANG Yousong, PEI Junpeng, LI Zenghui, et al. Review of research on deep learning in retinal blood vessel segmentation[J]. *Journal of frontiers of computer science and technology*, 2024, 18(8): 1960–1978.
- [8] LIU Li, QIANG Ma, OUYANG Cheng, et al. Topology optimization in medical image segmentation with fast Euler characteristic[EB/OL]. (2025–08–05)[2025–09–27]. <https://doi.org/10.48550/arXiv.2507.23763>.
- [9] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]// *Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer, 2015: 234–241.
- [10] 缪林一, 李峰. 基于动态特征图卷积网络的视网膜血管分割方法[J]. *中国激光*, 2024, 51(15): 181–191.
MIAO Linyi, LI Feng. Retinal vessel segmentation based on dynamic feature graph convolutional network[J]. *Chinese journal of lasers*, 2024, 51(15): 181–191.
- [11] 吴晨玥, 易本顺, 章云港, 等. 基于改进卷积神经网络的视网膜血管图像分割[J]. *光学学报*, 2018, 38(11): 125–131.
WU Chenyue, YI Benshun, ZHANG Yungang, et al. Retinal vessel image segmentation based on improved convolutional neural network[J]. *Acta optica sinica*, 2018, 38(11): 125–131.
- [12] 易三莉, 陈建亭, 贺建峰. ASR-UNet: 一种基于注意力机制改进的视网膜血管分割算法[J]. *山东大学学报 (理学版)*, 2021, 56(9): 13–20.
YI Sanli, CHEN Jianting, HE Jianfeng. ASR-UNet: an improved retinal vessels segmentation algorithm based on attention mechanism[J]. *Journal of Shandong University (natural science)*, 2021, 56(9): 13–20.
- [13] YI Yugen, GUO Changlu, HU Yangtao, et al. BCR-UNet: bi-directional ConvLSTM residual U-Net for retinal blood vessel segmentation[J]. *Frontiers in public health*, 2022, 10: 1056226.
- [14] JIA Wenran, MA Simin, GENG Peng, et al. DT-net: joint dual-input transformer and CNN for retinal vessel segmentation[J]. *Computers, materials & continua*, 2023, 76(3): 3393–3411.
- [15] TONG Le, LI Tianjiu, ZHANG Qian, et al. LiViT-Net: a U-Net-like, lightweight Transformer network for retinal vessel segmentation[J]. *Computational and structural biotechnology journal*, 2024, 24: 213–224.
- [16] LI Daxiang, SU Miao, LIU Ying. MSPDD-net: Mamba semantic perception dual decoding network for retinal image vessel segmentation[J]. *Computers in biology and medicine*, 2025, 193: 110370.
- [17] 张蓉蓉, 吕晓琪, 李菁, 等. 双路径编码和跨级解耦的视网膜血管分割[J]. *光学精密工程*, 2025, 33(14): 2242–2261.
ZHANG Rongrong, LÜ Xiaoqi, LI Jing, et al. Dual-path encoding and cross-level decoupling for retinal vessel segmentation[J]. *Optics and precision engineering*, 2025, 33(14): 2242–2261.
- [18] 彭圆圆, 黎浩洋, 李文, 等. MFAL-Mamba: 多尺度特征聚合和注意力引导的眼底血管分割算法[J]. *光学学报*, 2025, 45(22): 214–224.
PENG Yuanyuan, LI Haoyang, LI Wen, et al. MFAL-

- mamba: fundus blood vessel segmentation algorithm based on multi-scale features and attention mechanism[J]. *Acta optica sinica*, 2025, 45(22): 214–224.
- [19] 王植炜, 黄梓航, 郭英杰, 等. 基于 GAN 的视网膜血管分割标签优化方法[J]. 华中科技大学学报 (自然科学版), 2025, 53(9): 108–113.
- WANG Zhiwei, HUANG Zihang, GUO Yingjie, et al. GAN-based label refinement method for retinal vessel segmentation[J]. *Journal of Huazhong University of Science and Technology (natural science edition)*, 2025, 53(9): 108–113.
- [20] HERNANDEZ R J, ROBERTS P A, EL-BOURI W K. Advancing treatment of retinal disease through in silico trials[J]. *Progress in biomedical engineering*, 2023, 5(2): 022002.
- [21] WANG Zeyu, JIA L V, LIANG Haocheng. Partial class activation mapping guided graph convolution cascaded U-Net for retinal vessel segmentation[J]. *Computers in biology and medicine*, 2024, 178: 108736.
- [22] ZENG Shuang, LEE C H, NNAMDI M C, et al. Novel extraction of discriminative fine-grained feature to improve retinal vessel segmentation[EB/OL]. (2025–05–06)[2025–09–27]. <https://doi.org/10.48550/arXiv.2505.03896>.
- [23] SHIT S, PAETZOLD J C, SEKUBOYINA A, et al. clDice - a novel topology-preserving loss function for tubular structure segmentation[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 16555–16564.
- [24] LUX L, BERGER A H, WEERS A, et al. Topograph: an efficient graph-based framework for strictly topology preserving image segmentation[EB/OL]. (2025–04–17)[2025–09–27]. <https://doi.org/10.48550/arXiv.2411.03228>.
- [25] SCHACHT B, GREVING I, FRINTROP S, et al. ContextLoss: context information for topology-preserving segmentation[EB/OL]. (2025–06–10)[2025–09–27]. <https://doi.org/10.48550/arXiv.2506.11134>.
- [26] ZHOU Yi, AHMED T S, WANG Meng, et al. Masked vascular structure segmentation and completion in retinal images[J]. *IEEE transactions on medical imaging*, 2025, 44(6): 2492–2503.
- [27] ZHANG T Y, SUEN C Y. A fast parallel algorithm for thinning digital patterns[J]. *Communications of the ACM*, 1984, 27(3): 236–239.
- [28] HE Songtao, BASTANI F, JAGWANI S, et al. Sat2Graph: road graph extraction through graph-tensor encoding[C]//Computer Vision–ECCV 2020. Cham: Springer, 2020: 51–67.
- [29] AGUIAR M, CASTANO F, TRUJILLO M. Segmentation and detection of vascular bifurcations and crossings in retinal images[C]//Advances in Computing. Cham: Springer, 2018: 432–443.
- [30] OWEN C G, RUDNICKA A R, NIGHTINGALE C M, et al. Retinal arteriolar tortuosity and cardiovascular risk factors in a multi-ethnic population study of 10-year-old children; the Child Heart and Health Study in England (CHASE)[J]. *Arteriosclerosis, thrombosis, and vascular biology*, 2011, 31(8): 1933–1938.
- [31] STAAL J, ABRAMOFF M D, NIEMEIJER M, et al. Ridge-based vessel segmentation in color images of the retina[J]. *IEEE transactions on medical imaging*, 2004, 23(4): 501–509.
- [32] HOOVER A D, KOUZNETSOVA V, GOLDBAUM M. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response[J]. *IEEE transactions on medical imaging*, 2000, 19(3): 203–210.
- [33] HU Xiaoling, LI Fuxin, SAMARAS D, et al. Topology-preserving deep image segmentation[C]//Proceedings of the 33rd international conference on neural information processing systems. Red Hook: Curran Associates Inc., 2019: 5657–5668.
- [34] ZHOU Zongwei, RAHMAN SIDDIQUEE M M, TABAKHSH N, et al. UNet++: a nested U-Net architecture for medical image segmentation[C]//Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham: Springer, 2018: 3–11.
- [35] OKTAY O, SCHLEMPER J, LE FOLGOC L, et al. Attention U-Net: learning where to look for the pancreas[EB/OL]. (2018–05–20)[2025–09–27]. <https://doi.org/10.48550/arXiv.1804.03999>.
- [36] JHA D, SMEDSRUD P H, RIEGLER M A, et al. ResUNet++: an advanced architecture for medical image segmentation[C]//2019 IEEE International Symposium on Multimedia. San Diego: IEEE, 2019: 225–2255.
- [37] LI Liangzhi, VERMA M, NAKASHIMA Y, et al. IterNet: retinal image segmentation utilizing structural redundancy in vessel networks[C]//2020 IEEE Winter Conference on Applications of Computer Vision. Snowmass: IEEE, 2020: 3645–3654.

作者简介:



吕佳, 教授, 博士, 主要研究方向为机器学习、数据挖掘及其在医学图像处理等方面的应用。主持或参与国家、省部级科研项目 20 余项, 发表学术论文 70 余篇。E-mail: lvjia@cqnu.edu.cn。



杨鸿强, 硕士研究生, 主要研究方向为深度学习及其在医学图像处理中的应用。E-mail: 2024110516027@stu.cqnu.edu.cn。