第 62 卷　第 5 期
2026 年 9 月

南京大学学报(自然科学)
JOURNAL OF NANJING UNIVERSITY
(NATURAL SCIENCE)

Vol. 62, No. 5
Sept., 2026

DOI:10.13232/j.cnki.jnju.2026.05.006

基于模糊引导与双分支Transformer解码的视网膜血管分割方法FGDB-TransUnet

吕力玮,丁卫平*,耿　胜,刘鸿升,陈　龙,沈暄词
(南通大学人工智能与计算机学院,南通,226019)

摘　要:视网膜血管分割是眼底图像分析中的关键任务,对相关眼底疾病的辅助诊断与早期筛查具有重要意义,然而,血管细长连续、分支密集且边界模糊,现有模型在细小血管和弱边界区域容易出现漏检与断裂.针对现有卷积神经网络与Transformer方法对模糊区域和不确定区域显式建模不足的问题,提出一种模糊引导的双分支Transformer-UNet视网膜血管分割框架,其核心思想是将模糊区域显式建模并贯穿结构设计与优化约束.首先,从输入眼底图像中构造无需额外标注参与的模糊隶属图,用以刻画局部纹理不稳定与边界模糊区域,形成可用于训练与推理的一致性先验.其次,在保持Transformer编码器全局表征能力的基础上,引入标准分支与模糊分支并行解码,并利用模糊门控机制对多尺度跳跃连接进行调制,使模糊分支聚焦细小血管与弱边界的精细重建.接着,设计像素级自适应融合模块动态整合两分支输出,以实现结构恢复与边界精修的互补协同.最后,在损失函数中引入面向模糊区域的加权监督与分支一致性约束,从优化层面增强不确定区域的判别强度并抑制边界抖动.在包含800幅多疾病高分辨率眼底图像的FIVES数据集上开展大样本验证,并通过六组递进消融实验和固定网络权重的单因素参数扰动实验,分别分析双分支、模糊门控、自适应融合、模糊感知损失及先验参数设置的影响.实验结果表明,该方法能够在不同数据分布下保持稳定的分割质量,其准确率与$F1$分数等指标表现良好,尤其在血管模糊边界与复杂分支结构的分割场景中具有更强的鲁棒性,验证了所提框架用于视网膜血管分割任务的有效性.

关键词:模糊隶属度,双分支解码,血管分割,Transformer架构
中图分类号:TP18　　**文献标志码**:A

FGDB-TransUnet: A retinal vessel segmentation method using fuzzy guidance and dual-branch transformer decoding architecture

Lü Liwei, Ding Weiping*, Geng Sheng, Liu Hongsheng, Chen Long, Shen Xuanci
(School of Artificial Intelligence and Computer Science, Nantong University, Nantong, 226019)

Abstract: Retinal vessel segmentation is a fundamental task in fundus image analysis, playing a vital role in computer-aided diagnosis and early screening for fundus disea ses. However, due to the slender, continuous structure, dense branching, and blurred boundaries of blood vessels, existing models - ranging from conventional convolutional neural net works to Transformer-based approaches-tend to suffer from missed detection and fragmentation, especially in regions with tiny vessels and ambiguous boundaries. This is because thes e methods lack sufficient capability to explicitly model such uncertain and fuzzy areas. To address these challenges, this paper proposes a fuzzy-guided dual-branch Transformer-UNet framework for retinal vessel segmentation. The core idea of our approach is to explicitly model fuzzy regions and integrate this concept into

基金项目:国家重点研发计划(2024YFE0202700),国家自然科学基金(62576178,U2433216),江苏省自然科学基金(BK20231337),海南省卫生健康科技创新联合项目(WSJK2026ZD285)
收稿日期:2026－07－29
* 通信联系人,E-mail: dwp9988@163.com

both the structural design and optim ization constraints throughout the framework. First, we construct a fuzzy membership map from the input fundus images without requiring additional annotations. This map characterizes regions with unstable local textures and blurry boundaries, providing a consistent prior that can be applied consistently across both the training and inference stages. Second, while preserving the global representation capability of the Transformer encoder, we introduce a standard branch and a fuzzy branch for parallel decoding. A fuzzy gating mechanism is employed to modulate multi-scale skip connections, enabling the fuzzy branch to focus on the fine reconstruction of micro-vessels and weak boundaries. Next, we design a pixel-wise adaptive fusion module that dynamically integrates the outputs of the two branches, achieving complementary collaboration between structural restoration and boundary refinement. Finally, we incorporate region-weighted supervision for fuzzy areas and inter-branch consistency constraints into the loss function. These additions enhance the discriminative power of uncertain regions from an optimization perspective and help suppress boundary jitter. In addition, we conduct a large-scale evaluation on the FIVES dataset, which contains 800 high-resolution fundus images from four different disease categories. We further perform six progressive ablation configurations and a frozen-weight one-factor-at-a-time sensitivity analysis to examine the contributions of the dual-branch decoder, fuzzy gating, adaptive fusion, fuzzy-aware losses, and prior-parameter settings. Experimental results demonstrate that the proposed method maintains stable segmentation performance across various data distributions and achieves favorable results in metrics such as accuracy and F1-score. In particular, it exhibits stronger robustness in segmenting vessel boundaries that are blurry or characterized by complex branching structures, thereby validating the effectiveness of our proposed framework for retinal vessel segmentation tasks.

Keywords: fuzzy membership degree, dual-branch decoding, vessel segmentation, transformer architecture

眼底彩照具有非侵入、成本较低、采集便捷等特点,是眼科疾病筛查与慢病管理中常用的医学影像形式. 视网膜血管的形态变化与糖尿病视网膜病变、高血压性视网膜病变、动脉硬化及心脑血管风险等多类疾病密切相关,血管口径、分叉角度、迂曲度和血管密度等结构性指标能够为临床诊断、疾病进展评估和风险分层提供重要参考[1−3]. 因此,准确提取视网膜血管结构不仅是眼底图像分析中的基础任务,也是后续病灶定位、血管定量测量和智能辅助诊断的重要前提. DRIVE,STARE 和 CHASE_DB1 等公开数据集的建立进一步推动了视网膜血管分割方法的标准化比较与模型评估[4−6].

然而,视网膜血管分割不是简单的二值分割问题. 与一般医学图像中的器官或病灶相比,视网膜血管具有细长、稀疏、分支复杂和拓扑连续性强等特点. 主干血管通常较易识别,而细小血管、末端分支、交叉点和低对比边界区域容易受到照明不均、视盘高亮、病灶遮挡和成像噪声等因素影响,表现出边界模糊、灰度响应弱和结构不连续等问题,使模型在实际分割过程中容易出现细血管漏检、分支断裂和边界收缩等现象. 由此可见,视网膜血管分割的关键不仅在于提高整体像素分类准确率,更在于增强模型对弱边界、细小分支和复杂交叉结构的稳定识别能力.

早期视网膜血管分割方法主要依赖滤波增强、形态学处理、区域生长、匹配滤波和传统分类器等人工设计特征[1−3],这类方法具有一定可解释性,但对图像质量、参数选择和预处理流程较为敏感,在病灶干扰、低对比度和血管尺度变化明显的场景下泛化能力有限. 随着深度学习的发展,卷积神经网络逐渐成为医学图像分割的主流方法. U-Net 通过编码器-解码器结构和跳跃连接实现了高层语义信息与低层空间细节的融合,在医学图像分割任务中得到了广泛应用[7]. 在此基础上,残差连接、循环卷积、密集连接、多尺度特征融合和注意力机制等结构被不断引入,以增强模型对局部纹理、边界细节和多尺度血管结构的表达能力[8−10].

尽管卷积网络在视网膜血管分割中取得了较好效果,但其局部感受野特性限制了模型对长程依赖和血管整体拓扑结构的建模能力. 近年来,Transformer 及其视觉变体通过自注意力机制提升了全局上下文建模能力. Vision Transformer 将图像划分为序列化 patch,并利用 Transformer 编码器学习全局表征[11];TransUNet 进一步将

Transformer与U-Net结构结合，使模型能够同时利用全局语义信息和局部空间细节，为医学图像分割提供局部-全局互补的建模方式[12]. 然而，在视网膜血管分割任务中，Transformer对全局结构的建模优势不能完全解决弱边界和细血管丢失的问题，当局部血管响应较弱或边界不确定时，模型仍可能在细小分支、交叉点和背景相似区域产生不稳定预测.

近年来，针对视网膜血管分割任务，研究者进一步提出多种改进网络. SA-UNet通过空间注意力增强血管区域响应[13]，IterNet利用血管网络结构冗余进行迭代式细化[14]，DUNet采用可变形卷积适应血管形态变化[10]，TiM-Net将Transformer引入多尺度分割结构以增强上下文表达[15]，ESD-MR-Net通过扩展-压缩和双尺度残差结构提取多尺度特征[16]. 此外，粗糙注意力、模糊注意力和多尺度特征融合等方法也被用于增强模型对复杂血管结构和模糊边界区域的表达能力[17−19]. 这些方法从不同角度提升了分割性能，但多数研究仍主要通过增强特征提取能力或扩大感受野来改善结果，对弱边界、低对比区域和细小分支的不确定性缺乏更加显式的建模.

从本质上看，视网膜血管分割中的困难区域往往具有明显的不确定性. 细小血管末端、血管交叉点、病灶邻近区域和低对比边界不总是呈现明确的前景或背景特征，而是表现为连续过渡和模糊归属状态. 传统二值监督将所有像素视为确定样本，容易使模型在训练过程中被大量背景像素主导，导致整体准确率较高，但细血管召回不足. 模糊集理论通过隶属度函数将样本对某一集合的归属程度表示为连续值，为弱边界和不确定区域建模提供了理论基础[20−21]. 因此，将模糊不确定性引入视网膜血管分割，并进一步作用于先验构造、特征调制、解码恢复和损失约束过程，是提升复杂血管结构分割稳定性的有效思路.

针对上述问题，本文提出FGDB-TransUnet模型. 该模型以TransUNet的局部-全局混合建模思想为基础，从输入的眼底图像中构造无需额外标注参与的模糊先验图，用于刻画局部纹理不稳定、边界过渡模糊和尺度响应不一致区域. 在解码阶段，设计结构分支与模糊引导分支并行的双分支解码框架，结构分支侧重主体血管连续恢复，模糊分支在模糊先验调制下强化细小血管与弱边界区域的表达，同时，通过像素级自适应融合和面向模糊区域的优化约束，实现结构恢复、边界精修和分支一致性的协同优化. 与单纯堆叠注意力模块或扩大网络容量不同，本文方法将血管分割中的困难区域转化为可计算、可调制和可监督的模糊区域建模问题，从而提升模型在复杂分支、低对比血管和边界过渡区域中的分割稳定性.

本文主要贡献如下.

(1)提出一种无需额外标注参与的模糊隶属图构造策略，以局部对比度与多尺度差分响应为依据来刻画不确定区域，作为训练与推理阶段一致可用的不确定性先验.

(2)设计一种模糊引导的双分支解码Trans-UNet框架，通过模糊门控机制调制多尺度跳跃连接，使模糊分支在结构上聚焦细血管与弱边界的精细恢复，标准分支保持主体结构的稳定重建.

(3)提出像素级自适应融合与面向模糊区域的优化约束，包括模糊区域加权监督与分支一致性正则化，以增强不确定区域的判别强度并降低边界预测振荡.

(4)在实验层面引入FIVES大样本验证，并构建A0～A5六组严格递进的受控消融配置，进一步通过固定网络权重的单因素扰动分析绿色通道、局部统计窗口、多尺度组合与融合系数对模型输出的影响，从样本规模、模块贡献和参数稳定性三个角度验证所提方法.

1　基础知识

1.1　模糊不确定性与隶属度建模　在传统二值分割任务中，图像中的每个像素通常被划分为血管或背景两类. 该方式在主干血管和均匀背景区域具有较好的适用性，但对于视网膜血管边缘、细小分支末端以及病灶邻近区域，像素灰度往往呈现连续过渡状态，不总是具有明确的类别边界. 若仍将这类像素强制视为完全确定的前景或背景，模型在训练过程中容易受到硬标签约束的影响，从而在弱边界区域产生过度收缩，在细血管区域产生漏检.

模糊集理论通过隶属度描述样本属于某一集

合的程度,使像素状态能够由离散的二值判断扩展为连续表达[20-21]. 设图像定义域为Ω,血管区域对应的模糊集合为A,则任一像素$x_i\in\Omega$对血管集合的隶属程度可表示为:

$$\mu_A(x_i):\Omega\to[0,1] \tag{1}$$

其中,$\mu_A(x_i)$越接近1,表示该像素越可能属于血管结构;越接近0,其越可能属于背景;当取值处于中间区间时,通常意味着该像素可能位于边界过渡、低对比细血管或结构不稳定区域. 与二值标签相比,隶属度表达能保留更多关于局部不确定性的连续信息,更适合描述视网膜血管图像中模糊边界和弱响应区域.

在眼底图像中,不确定区域往往与局部灰度波动和多尺度结构差异有关. 主干血管通常具有较强灰度对比和较稳定结构,而细小血管、交叉点和边界过渡区域则更容易在邻域内产生灰度变化. 若用G表示输入图像的灰度表示,其局部均值和局部二阶矩可写为:

$$\begin{aligned}\mu&=\mathrm{AvgPool}_k(G)\\ \mu_2&=\mathrm{AvgPool}_k(G^2)\end{aligned} \tag{2}$$

则局部标准差响应为:

$$S=\sqrt{\mu_2-\mu^2+\varepsilon} \tag{3}$$

其中,$\mathrm{AvgPool}_k(\cdot)$表示窗口大小为$k$的平均池化操作,$\varepsilon$是防止数值不稳定的极小常数. 该响应能够刻画局部区域内灰度分布的不稳定程度. 当某一位置位于血管边缘或细小分支附近时,其邻域内灰度通常会发生较明显的变化,因此局部标准差的响应相对增强.

除局部统计波动外,视网膜血管还具有明显的尺度差异,不同宽度的血管在不同平滑尺度下会呈现不同响应,因此多尺度差分可用于刻画结构尺度变化. 设k_1和k_2分别表示小尺度与大尺度平滑窗口且$k_1<k_2$,则多尺度差分响应可写为:

$$\begin{aligned}B_s&=\mathrm{AvgPool}_{k_s}(G)\\ B_l&=\mathrm{AvgPool}_{kl}(G)\\ D&=|B_s-B_l|\end{aligned} \tag{4}$$

该响应能够突出不同尺度下表现不一致的区域,对细血管、弱边界和交叉结构具有较强敏感性. 由此可见,模糊不确定性并不是对血管标签的简单替代,而是对图像中“难以确定区域”的连续描述,为后续特征调制、分支解码和损失约束提供更细粒度的空间提示.

1.2 Transformer全局依赖建模 视网膜血管具有明显的树状结构和长程连续性,局部卷积能够较好地提取边缘、纹理和局部形态特征,但卷积核感受野有限,难以直接建立远距离血管片段之间的依赖关系. 当图像中存在低对比度血管或局部断裂时,仅依赖局部特征容易导致模型无法判断远端血管是否属于同一连续结构. Transformer的自注意力机制能够在全局范围内建模不同位置之间的相关性,因此被广泛用于增强医学图像分割中的上下文表达能力[11-12].

设输入特征序列为:

$$X\in\mathbb{R}^{N\times d} \tag{5}$$

通过线性映射得到查询、键、值:

$$Q=XW_Q,K=XW_K,V=XW_V \tag{6}$$

其中,W_Q, W_K, W_V为可学习参数矩阵. 随后,通过查询与键之间的相似性计算注意力权重,并对值矩阵进行加权聚合:

$$\mathrm{Attention}(Q,K,V)=\mathrm{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{7}$$

其中,d_k表示键向量维度,用于缩放点积结果,避免维度增大导致梯度不稳定. 该机制使任意两个空间位置之间都可以直接建立联系,从而增强模型对全局结构和远距离依赖的表达能力.

在视觉任务中,Vision Transformer将图像划分为固定大小的图像块,并将其映射为token序列进行全局建模[11]. 对于视网膜血管分割,这种建模方式有助于捕获主干血管走向、远端分支联系以及跨区域结构一致性. TransUNet进一步将Transformer编码器与U-Net式解码结构结合,一方面利用Transformer建模全局上下文,另一方面通过解码器和跳跃连接恢复局部空间细节[12]. 这种局部-全局结合的结构为视网膜血管分割提供了重要基础.

但是,Transformer的全局建模能力并不等同于对所有细节区域都具有稳定的恢复能力. 由于视网膜血管中的细小分支和弱边界区域本身响应较弱,当这些局部特征在编码或下采样过程中被削弱时,全局注意力可能更倾向于强化主干血管和显著结构,难以充分恢复低对比细节. 因此,在

利用Transformer建模全局依赖的同时，还需要引入面向弱边界和不确定区域的显式引导机制，使全局语义信息与局部细节补偿能协同发挥作用.

1.3　选择性特征融合与不平衡监督约束　编码-解码结构中的跳跃连接能够将浅层空间细节传递到解码阶段，是医学图像分割中保持边界信息的重要机制. 然而，浅层特征除了包含有用的血管边缘，也可能包含背景纹理、光照变化、视盘高亮和噪声响应，若直接将所有浅层特征与高层语义特征拼接，容易使无关背景信息进入解码器，影响弱边界区域的判别. 注意力门控和特征重标定机制通过学习空间或通道权重，对特征响应进行选择性增强，使网络更加关注与目标结构相关的区域[8,13].

一般形式上，设某尺度细节特征为S，门控/权重图为$A\in[0,1]^{H\times W}$，则选择性调制可写为：

$$\tilde{S}=S\odot A \tag{8}$$

若将A取为不确定性隶属度图（或其尺度对齐形式），则可在空间上抑制确定区域的噪声注入，同时，在弱边界与细结构区域保留或强化有效细节. 这类“以不确定性为权重的选择性调制”能在稳定性与细节性之间提供更柔性的折中空间.

当预测由多路信息组合得到时，固定权重融合往往难以适配跨区域差异，更一般的做法是学习像素级融合权重，使融合策略由数据驱动形成. 设两路特征为Φ_1,Φ_2，可学习权重图：

$$w=\sigma\big(g([\Phi_1;\Phi_2])\big)\in[0,1]^{H\times W} \tag{9}$$

并以加权方式得到融合输出：

$$Z=w\odot Z_1+(1-w)\odot Z_2 \tag{10}$$

其中，Z_1,Z_2为对应预测（或logits）. 该形式允许模型在清晰区域偏向稳定预测，在模糊区域偏向细节修复，从而在跨区域条件下获得更好的适配性.

除了结构层面的选择性融合外，不确定区域的优化通常还需要通过风险加权显式实现. 以逐像素损失ℓ_{ij}为例，引入权重图$W_{ij}\geqslant 0$后，加权经验风险可写为：

$$L_w=\frac{\sum \ell_{ij}W_{ij}}{\sum W_{ij}+\varepsilon} \tag{11}$$

当权重与不确定性隶属度相关（例如，$W=F^{\gamma}$）时，高不确定区域对梯度的贡献被放大，从而改变优化过程中风险分配的空间结构，使学习重心从“数量占优的确定区域”转向“决定质量的困难区域”. 进一步，在多路预测或多分支表达下，不确定区域的分歧往往更显著，若缺少约束易导致局部振荡. 此时，可引入加权一致性正则，设两路概率输出为P_1,P_2，则：

$$L_{cons}=\frac{\sum|P_1-P_2|\cdot F}{\sum F+\varepsilon} \tag{12}$$

该类约束将一致性要求集中于模糊区域，有助于降低弱边界处的预测波动并提高结构连续性. 总体上，不确定性隶属度提供了统一的空间权重接口，使选择性融合与加权优化能够在同一表征之上协同工作，为后续的细结构恢复与边界稳定提供理论层面的支撑.

2　基于模糊引导与双分支Transformer解码的视网膜血管分割方法（FGDB-TransUnet）

视网膜血管分割的主要误差集中于弱边界、低对比细分支与交叉点区域，在下采样与全局建模过程中，这类结构更易被背景吸收而产生细血管漏检、局部断裂与边界收缩. 为此，本文在TransUNet的混合编码-解码框架上引入面向“不确定区域”的结构化改动，将血管分割中的弱边界/低对比区域由被动噪声来源转化为可计算、可注入的不确定性引导信号. 具体地，首先从输入图像无监督构造模糊先验图F，以局部统计波动与多尺度差分响应共同刻画弱边界与细结构附近的空间不确定性. 随后，在解码端引入双分支协同机制，其中结构分支侧重拓扑连通与稳态恢复，模糊分支在多尺度跳跃连接回注处以$S_k\odot F_k$的形式对高分辨率细节进行门控调制，使细分支与过渡带获得更充分的细节补偿而确定区域的冗余纹理回注被抑制. 进一步，设计像素级融合权重w对两分支输出进行空间自适应组合，使主体结构与难区细节在同一输出中实现区域化、可控的权衡. 训练阶段结合模糊感知的加权监督与难区一致性约束，将优化压力持续聚焦于不确定区域并抑制分支冲突，从结构路径与优化路径两方面共同指向细血管检出、交叉点连接与边界过渡带

刻画的稳定提升. 模型的具体架构如图1所示.

符号约定:输入patch记为$X\in\mathbb{R}^{B\times C\times H\times W}$(默认$H=W=256$),标注为$Y\in\{0,1\}^{B\times H\times W}$,有效视野掩膜为$M\in\{0,1\}^{B\times H\times W}$. 网络输出二类logits $Z\in\mathbb{R}^{B\times 2\times H\times W}$,前景概率为:

$$P=\mathrm{Softmax}(Z)_{[:,1,:,:]}\in[0,1]^{B\times H\times W} \tag{13}$$

2.1 无监督不确定性先验F的构造 视网膜血管分割中的难点主要来源于细小血管、弱边界、交叉点以及局部低对比区域,这些区域在眼底图像中往往不会呈现稳定、清晰的二值结构,而是表现出灰度变化缓慢、边界过渡模糊、局部纹理不均以及尺度响应不一致等特征. 传统编码-解码模型在下采样过程中容易削弱这类细粒度结构,导致细血管漏检、分支断裂和边界收缩. 为了增强模型对上述不确定区域的感知能力,本文在不引入额外人工标注的条件下,根据输入图像自身的局部统计响应和多尺度差分响应来构造模糊先验图F,可为后续模糊分支解码提供空间引导.

设输入眼底图像块为$X\in\mathbb{R}^{C\times H\times W}$. 考虑到眼底图像绿色通道通常能够提供更清晰的血管-背景对比度,若输入图像为多通道表示则提取绿色通道作为灰度表示,若输入图像为单通道表示则直接将该通道作为灰度表示,由此得到灰度图$G\in[0,1]^{H\times W}$. 随后,对G进行逐样本归一化,使其像素值分布处于统一范围内,为后续局部统计计算和多尺度响应计算提供稳定的输入.

如图2所示,模糊先验生成过程由局部标准差响应和多尺度差分响应两条路径组成. 第一条路径用于刻画像素邻域内的局部灰度波动. 对于背景平坦区域或血管内部较稳定区域,其局部灰度变化较小,响应值相对较低;在血管边缘、细小分支、交叉点以及局部纹理变化明显的位置,邻域

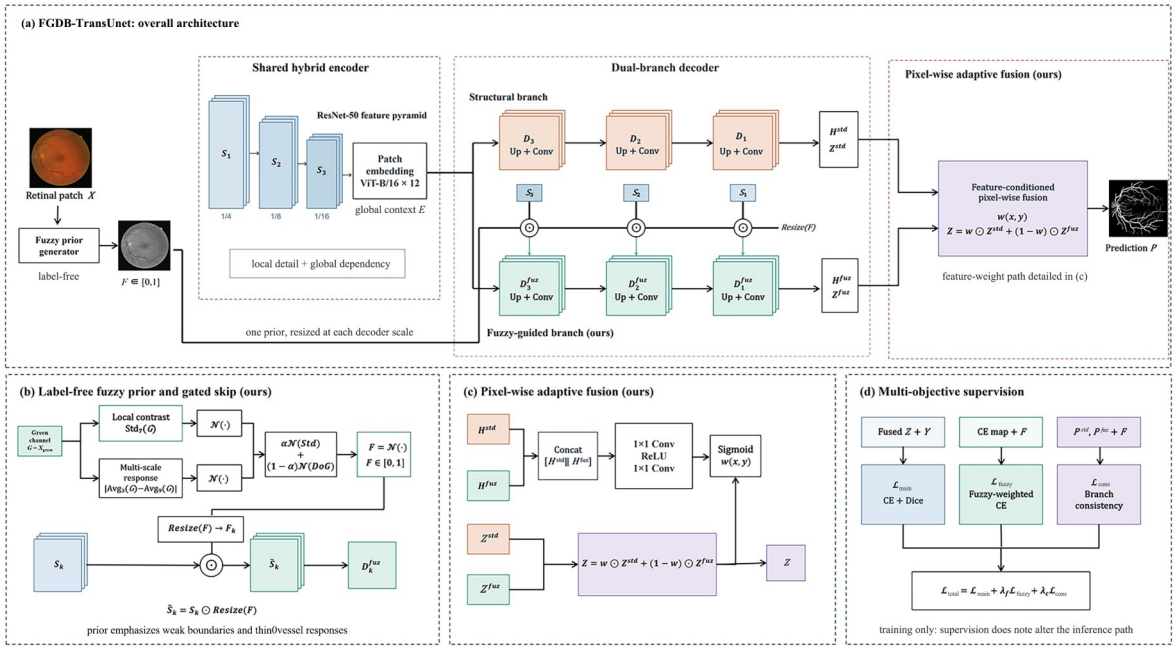


图1 FGDB-TransUnet模型

Fig. 1 FGDB-TransUnet model

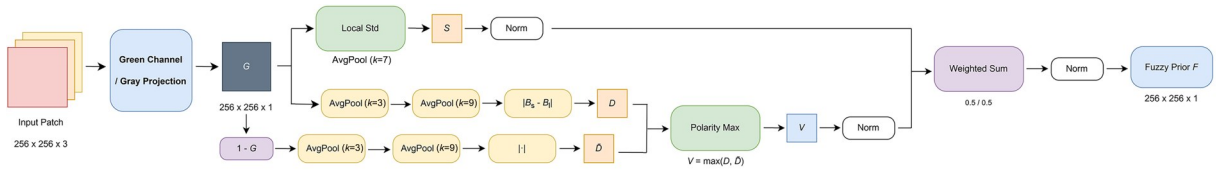


图2 模糊先验生成模块

Fig. 2 Fuzzy prior generation module

内灰度分布更加不稳定，局部标准差响应会相应增强．具体地，首先通过平均池化计算灰度图 G 的局部均值 μ：

$$\mu = \mathrm{AvgPool}_{k_s}(G) \tag{14}$$

同时计算局部二阶矩 μ_2：

$$\mu_2 = \mathrm{AvgPool}_{k_s}(G^2) \tag{15}$$

在此基础上，局部标准差响应 S 可表示为：

$$S = \sqrt{\mu_2 - \mu^2 + \varepsilon} \tag{16}$$

其中，$\mathrm{AvgPool}_{k_s}(\cdot)$ 表示窗口大小为 k_s 的平均池化操作，ε 为防止数值不稳定的极小常数．采用局部窗口 $k_s = 7$ 计算标准差响应，使其能够在一定邻域范围内捕获灰度变化较明显的区域．该响应并不直接判断某一像素是否属于血管，而是用于刻画该位置附近是否存在边界、纹理波动或结构不确定性．

第二条路径用于提取多尺度差分响应．视网膜血管具有明显的尺度差异，主干血管在较大尺度平滑后仍能保持结构连续，而细小血管、弱边界和局部交叉点则更容易在不同尺度平滑结果之间产生响应差异．为此，分别采用小尺度平均池化和大尺度平均池化对灰度图 G 进行平滑，得到：

$$\begin{aligned} G_{k_1} &= \mathrm{AvgPool}_{k_1}(G) \\ G_{k_2} &= \mathrm{AvgPool}_{k_2}(G) \end{aligned} \tag{17}$$

其中，k_1 和 k_2 分别表示小尺度和大尺度平均池化窗口且满足 $k_1 < k_2$．对应采用 $k_1 = 3, k_2 = 9$．通过计算两种尺度平滑结果之间的绝对差值，可以得到原始灰度图上的多尺度差分响应 D：

$$D = \left| G_{k_1} - G_{k_2} \right| \tag{18}$$

该响应能够突出不同尺度下结构表现不一致的区域，尤其对细小血管、弱边界以及局部纹理变化区域较为敏感．

考虑到眼底图像中血管通常表现为暗线结构，但在局部反光、亮度不均或图像增强影响下，部分区域可能出现响应极性变化，仅在原始灰度图上计算差分响应可能无法充分覆盖所有不确定区域．进一步对反相灰度图 $1 - G$ 进行相同的多尺度差分计算，首先得到反相图在两种尺度下的平滑结果：

$$\begin{aligned} \overline{G_{k_1}} &= \mathrm{AvgPool}_{k_1}(1 - G) \\ \overline{G_{k_2}} &= \mathrm{AvgPool}_{k_2}(1 - G) \end{aligned} \tag{19}$$

随后计算反相多尺度差分响应 $\overline{D}$：

$$\overline{D} = \left| \overline{G_{k_1}} - \overline{G_{k_2}} \right| \tag{20}$$

为了同时保留原始灰度方向和反相灰度方向中的显著响应，对 D 和 $\overline{D}$ 进行逐像素最大选择，得到极性鲁棒的多尺度响应 V：

$$V = \max\left(D, \overline{D}\right) \tag{21}$$

通过该操作，模糊先验生成过程能够避免过度依赖单一灰度极性，使模型在面对暗血管、局部亮度变化和复杂背景纹理时仍能获得较稳定的不确定区域提示．此时，S 更侧重于刻画局部统计波动，V 更侧重于刻画多尺度结构差异，两者从不同角度描述了视网膜血管图像中的模糊区域．

在得到局部标准差响应 S 和极性鲁棒多尺度响应 V 后，分别对二者进行逐样本归一化处理，使其映射到统一的数值范围内：

$$\begin{aligned} \hat{S} &= \mathrm{Norm}(S) \\ \hat{V} &= \mathrm{Norm}(V) \end{aligned} \tag{22}$$

其中，$\mathrm{Norm}(\cdot)$ 表示逐样本归一化操作．随后，将两类响应进行加权融合，并再次归一化得到最终模糊先验图 F：

$$F = \mathrm{Norm}\left(\lambda_s \hat{S} + \lambda_v \hat{V}\right) \tag{23}$$

其中，λ_s 和 λ_v 分别表示局部标准差响应和多尺度差分响应的权重系数．设置 $\lambda_s = \lambda_v = 0.5$，使局部统计不确定性和多尺度结构不确定性以相同权重参与先验构造．最终得到 $F \in [0,1]^{H \times W}$，其数值越大，表示该位置越可能处于细小血管、弱边界、交叉点或局部结构不稳定区域．

需要说明的是，模糊先验图 F 并不是血管标签的替代，也不是对人工标注的近似复制，而是由输入图像自身计算得到的连续不确定性提示．与二值标签相比，F 能保留更多关于局部边界模糊性、尺度变化和纹理波动的信息．在后续解码阶段，模糊分支利用该先验对多尺度跳跃特征进行门控，使模型在保持主体血管结构恢复能力的同时，对弱边界和细小分支区域分配更高关注，从而缓解传统编码-解码结构在视网膜血管分割中容易出现的细节丢失和边界收缩问题．

为了更清晰地描述模糊先验的构造过程,算法 1 给出了从输入 patch 到模糊先验图的计算流程. 该过程仅依赖输入图像自身的局部统计响应与多尺度差分响应,不需要额外人工标注,因此能够在训练阶段和推理阶段均可直接使用.

算法 1　模糊先验生成

输入:输入图像块 $X\in\mathbb{R}^{B\times C\times H\times W}$;局部统计窗口大小 k_s;多尺度平滑窗口大小 k_1, k_2;

输出:模糊先验图 $F\in[0,1]^{B\times 1\times H\times W}$.

1. 若 $C\geqslant 2$, 提取输入图像的绿色通道作为灰度表示 G,否则,直接将单通道输入作为灰度表示 G;

2. 将 G 归一化到 $[0,1]$ 区间,得到标准化灰度图;

3. 利用平均池化计算局部均值 μ 与局部二阶矩 μ_2;

4. 根据局部均值与局部二阶矩计算局部标准差响应:

$R_{std}=\sqrt{\mu_2-\mu^2+\varepsilon}$;

5. 分别采用小尺度窗口 k_1 与大尺度窗口 k_2 对 G 进行平滑,得到 G_{k_1} 与 G_{k_2};

6. 计算多尺度差分响应:$R_{dog}=\left|G_{k_1}-G_{k_2}\right|$;

7. 对反相图 $1-G$ 重复步骤 5~6,得到反相差分响应 R_{dog}^{inv};

8. 取两种极性响应的逐像素最大值,得到极性鲁棒的尺度响应:$R_{ms}=\max\left(R_{dog},R_{dog}^{inv}\right)$;

9. 分别对 R_{std} 与 R_{ms} 进行逐样本归一化,得到 $\hat{R}_{std}$ 与 $\hat{R}_{ms}$;

10. 融合两类响应并再次归一化,得到模糊先验图 $F=\mathrm{Norm}\left(\lambda_s\hat{R}_{std}+\lambda_m\hat{R}_{ms}\right)$;

11. 返回 F.

2.2　先验门控的双分支解码机制　跳跃连接是恢复细节的关键通道,但它对“哪里该回注,哪里不该回注”缺少选择性. 在确定区域,过度回注浅层纹理会放大噪声并诱发误报;在弱边界与细分支区域,回注强度不足就容易出现断裂与边界收缩. 为了在不破坏全局拓扑建模的前提下重塑细节入口的信息流,本方法在解码端引入双分支. 结构分支维持稳态结构恢复,模糊分支在跳跃回注处以先验 F 进行门控,形成“难区细节补偿”路径. 两个分支共享同一个编码输出并保持相近的解码形态,使该输出可被后续融合与一致性约束有效利用.

设编码端输出的多尺度跳跃特征为 $\{S_k\}$,其中,k 表示尺度层级. 对输入尺寸 $H=W=256$,所用的三条 skip 通常对应 128×128, 64×64, 32×32 三个分辨率(无需展开通道与 token 细节). 在每个尺度,门控先将 F 插值到该尺度得到 F_k,再对 skip 做逐像素乘法调制:

$$F_k=Interp(F;H_k\times W_k),\widetilde{S_k}=S_k\odot F_k \tag{24}$$

采用乘法门控的关键在于,它没有“凭空增强特征”,而是对 skip 的有效注入区域施加空间选择性. 当 F_k 较小时,模糊分支对该位置的浅层纹理回注被抑制;当 F_k 较大时,门控退化为近似常规回注,从而把容量集中到不确定区域. 门控施加于 skip 而非主干语义特征的原因在于主干语义负责全局拓扑一致性,而误差更多地发生在高分辨率细节回注阶段,因此,在 skip 入口处重塑信息流能够更直接地改善细血管与边界过渡带,而不会削弱全局连通约束.

用 $D_k(\cdot)$ 表示尺度 k 的解码块操作(上采样、拼接、卷积更新的组合),则结构分支与模糊分支的逐尺度更新分别为:

$$\begin{aligned}U_k^{std}&=D_k\left(U_{k-1}^{std},S_k\right)\\U_k^{fuz}&=D_k\left(U_{k-1}^{fuz},\widetilde{S_k}\right)=D_k\left(U_{k-1}^{fuz},S_k\odot F_k\right)\end{aligned} \tag{25}$$

三尺度门控不是冗余,粗尺度更利于约束主干走向与拓扑一致性,中尺度对交叉点与中等血管分叉的连接更敏感,细尺度直接决定边界过渡带与细血管末端的可辨性. 将门控在多尺度重复,相当于在不同空间频段上同时对“难区细节入口”进行重分配,避免仅在单尺度补偿导致的局部改善与全局不稳.

门控同样影响反向传播的梯度分配. 由于 $\widetilde{S_k}$ 对 S_k 的敏感性在像素层面被 F_k 调制,弱边界/细分支位置(F_k 较大)对应的 skip 通道将获得更大的有效更新;确定区域(F_k 较小)则对浅层纹理的依赖被抑制,降低背景纹理捷径对收敛方向的主导性. 由此,结构分支更偏向稳态恢复,模糊分支更偏向难区补偿,两者形成互补输出.

两分支末端特征分别为 F^{std} 与 F^{fuz},经各自分割头得到 logits:

$$Z^{std}=H\left(F^{std}\right),Z^{fuz}=H\left(F^{fuz}\right) \tag{26}$$

上述过程本质上是在解码阶段对多尺度跳跃

连接进行差异化处理. 结构分支保持常规跳跃连接以恢复主体拓扑，模糊分支利用尺度对齐后的先验图对跳跃特征进行门控，使弱边界和细分支区域获得更强的细节回注.

算法2给出了该解码机制的具体流程.

算法2　先验门控双分支解码机制

输入：编码器输出特征E；多尺度跳跃特征集合$\{S_l\}_{l=1}^{L}$；模糊先验图F；插值算子Resize$(\cdot)$；结构分支解码块$D_s(\cdot)$；模糊分支解码块$D_f(\cdot)$；分割头Head$(\cdot)$；

输出：结构分支logits Z_s；模糊分支logits Z_f；结构分支末端特征H_s；模糊分支末端特征H_f.

1. 将编码器输出特征E同时送入结构分支和模糊分支，初始化：$H_s^0=E, H_f^0=E$；

2. 按照由低分辨率到高分辨率的顺序遍历跳跃特征$\{S_l\}_{l=1}^{L}$；

3. 在第l个尺度，将结构分支特征H_s^{l-1}上采样至当前尺度；

4. 将结构分支上采样特征与原始跳跃特征S_l拼接，并通过结构解码块更新：$H_s^l=D_s\left(H_s^{l-1},S_l\right)$；

5. 将模糊分支特征H_f^{l-1}上采样至当前尺度；

6. 将模糊先验图F插值到当前尺度，得到尺度对齐先验：$F_l=\mathrm{Resize}\left(F,S_l\right)$；

7. 利用F_l对跳跃特征进行逐像素门控，得到模糊门控跳跃特征：$\tilde{S}_l=F_l\odot S_l$；

8. 将模糊分支上采样特征与$\tilde{S}_l$拼接，并通过模糊解码块更新：$H_f^l=D_f\left(H_f^{l-1},\tilde{S}_l\right)$；

9. 重复步骤3～8，直至所有尺度的跳跃特征均完成解码更新；

10. 得到两分支末端特征：$H_s=H_s^L, H_f=H_f^L$；

11. 分别通过分割头得到两分支预测logits：$Z_s=\mathrm{Head}\left(H_s\right), Z_f=\mathrm{Head}\left(H_f\right)$；

12. 返回Z_s, Z_f, H_s, H_f.

2.3　像素级自适应融合　双分支输出的可靠性具有显著空间异质性，主体血管与大面积背景更依赖稳态预测以抑制误报，而弱边界与细分支更依赖难区补偿以减少漏检与断裂. 若采用固定系数对两分支输出进行全局加权，往往会把这种异质性抹平，造成“细血管收益不足或误报上升”的权衡. 为此，引入像素级自适应融合头来学习权重图$w\in(0,1)^{B\times1\times H\times W}$，使融合成为逐像素的“分支选择”.

融合权重从两分支末端特征学习而来，而非直接根据logits幅值决定. 将末端特征在通道维拼接后，通过两层1×1卷积生成权重（中间ReLU、末端Sigmoid）：

$$w=\sigma\left(\mathrm{Conv}_{1\times1}^{(2)}\left(\delta\left(\mathrm{Conv}_{1\times1}^{(1)}\left(\left[F^{std};F^{fuz}\right]\right)\right)\right)\right) \tag{27}$$

其中，$[;]$表示通道拼接，$\delta(\cdot)$为ReLU，$\sigma(\cdot)$为Sigmoid. 最终在logits层完成逐像素融合：

$$Z=w\odot Z^{std}+(1-w)\odot Z^{fuz} \tag{28}$$

该设计的关键点是w由特征驱动学习，可利用边界纹理、局部上下文一致性与语义置信等信息形成更稳健的空间决策. 在主体区域融合倾向于稳态结构分支，在弱边界与细分支区域融合倾向于模糊补偿分支，从而把双分支互补性转化为单一输出中的区域化优势.

与固定系数融合不同，像素级自适应融合需要先从两分支末端特征中学习空间权重，再在logits层完成逐像素组合，这样可以避免仅依据概率幅值进行硬选择，使融合过程同时利用语义上下文、边界纹理与分支置信信息. 算法3给出了自适应融合的具体计算过程.

算法3　像素级自适应融合

输入：结构分支末端特征H_s；模糊分支末端特征H_f；结构分支logits Z_s；模糊分支logits Z_f；融合权重头$\Phi(\cdot)$；

输出：像素级融合权重w；融合logits Z；最终血管概率图P.

1. 对两分支末端特征进行空间尺寸对齐，保证H_s与H_f具有相同的高度和宽度；

2. 在通道维度拼接两分支末端特征，得到融合特征：$H_c=\mathrm{Concat}\left(H_s,H_f\right)$；

3. 将H_c输入融合权重头，通过1×1卷积压缩通道并提取分支互补信息；

4. 经ReLU激活后，再通过1×1卷积生成单通道融合权重图；

5. 采用Sigmoid函数将权重映射到$[0,1]$区间：$A=\sigma\left(\Phi\left(H_c\right)\right)$；

6. 在logits层对两分支输出进行逐像素融合：$Z=w\odot Z^{std}+(1-w)\odot Z^{fuz}$；

7. 对融合logits进行Softmax或Sigmoid变换，得到最终血管概率图P，并返回w, Z, P.

2.4　模糊感知优化约束　结构上的门控与分工并不必然等价于“训练会聚焦难区”. 血管分割具

有显著的类别不平衡,若仅依赖常规损失,梯度往往由大量背景像素主导,难区信号被稀释. 为了使先验引导形成闭环,本方法在有效视野内采用基础监督并叠加两类模糊感知约束. 难区加权监督把优化压力持续压向 F 指示区域,难区一致性约束抑制两分支在难区的强分歧,避免融合阶段输出的不稳定.

基础监督在掩膜 M 内计算. 记逐像素交叉熵图为 $CE(Z,Y)\in\mathbb{R}^{B\times H\times W}$,则掩膜交叉熵定义为:

$$L_{ce}=\frac{\sum\left(CE(Z,Y)\odot M\right)}{\sum M+\varepsilon} \tag{29}$$

Dice损失采用前景概率 P 计算软Dice:

$$\begin{aligned}&\mathrm{Dice}(P,Y)=\frac{2\sum\left(P\odot Y\odot M\right)+\varepsilon}{\sum\left(P\odot M\right)+\sum\left(Y\odot M\right)+\varepsilon}\\&L_{dice}=1-\mathrm{Dice}(P,Y)\end{aligned} \tag{30}$$

基础主损失为 $L_{\mathrm{main}}=L_{ce}+L_{\mathrm{dice}}$,用于保证全局可分性与稳定收敛.

难区加权监督通过先验权重来重分配交叉熵代价. 为了保持先验的统计语义并防止网络通过反向传播“规避权重”,对 F 采用停止梯度 $\mathrm{stopgrad}(\cdot)$. 设强调指数为 γ(默认 $\gamma=2$),则权重与损失定义为:

$$\begin{aligned}&W_f=\left(\mathrm{stopgrad}(F)\right)^{\gamma}\odot M\\&L_{\mathrm{fuzzy}}=\frac{\sum\left(CE(Z,Y)\odot W_f\right)}{\sum W_f+\varepsilon}\end{aligned} \tag{31}$$

该项的作用不是简单地提高损失值,而是改变梯度在空间上的分布. 弱边界与细分支区域承担更高代价,缓解了“背景安全预测”主导收敛导致的细血管漏检.

同时,引入难区一致性约束以协调双分支. 记两分支前景概率为:

$$\begin{aligned}&P^{std}=\mathrm{Softmax}\left(Z^{std}\right)_{[:,1,:,:]}\\&P^{fuz}=\mathrm{Softmax}\left(Z^{fuz}\right)_{[:,1,:,:]}\end{aligned} \tag{32}$$

一致性项在先验指示区域聚焦,采用加权 L_1:

$$\begin{aligned}&W_c=\mathrm{stopgrad}(F)\odot M\\&L_{cons}=\frac{\sum\left(\left|P^{std}-P^{fuz}\right|\odot W_c\right)}{\sum W_c+\varepsilon}\end{aligned} \tag{33}$$

该约束并不要求两分支全图趋同,而是抑制它们在困难区域的强分歧,使“稳态结构恢复”与“难区细节补偿”在空间上更可协调,减少融合后局部伪影与振荡风险.

最终总损失为:

$$L_{\mathrm{total}}=L_{\mathrm{main}}+\lambda_f L_{\mathrm{fuzzy}}+\lambda_c L_{\mathrm{cons}} \tag{34}$$

其中,默认 $\lambda_f=0.35$, $\lambda_c=0.1$, $\gamma=2$.

2.5 基于模糊引导与双分支 Transformer 解码的视网膜血管分割方法 将上述模块嵌入骨干网络后,整体形成“先验指示难区-门控改变细节入口-像素级选择融合-损失再压向难区”的闭环链路. 输入首先生成模糊先验 F,其在推理阶段同样可获得,因此空间含义稳定且不引入训练-测试偏移. 解码阶段中,结构分支沿用常规跳跃回注来维持主体血管连通性与低误报;模糊分支通过 $\tilde{S}_k=S_k\odot F_k$ 将回注集中在弱边界与细分支处,并在多尺度重复该操作,使难区补偿同时覆盖拓扑约束、中尺度分叉连接与细尺度边界过渡带刻画. 随后,融合头根据末端特征预测逐像素权重 w,并以 $Z=w\odot Z^{std}+(1-w)\odot Z^{fuz}$ 实现空间自适应选择,从而在主体区域更偏向稳态预测,在不确定区域更偏向细节补偿. 在训练阶段,L_{fuzzy} 将监督代价持续压向 F 指示区域,使细血管末端与弱边界获得更强梯度;L_{cons} 抑制两分支在难区的强分歧,避免融合阶段产生局部冲突. 由此,细血管漏检、交叉点断裂与边界收缩等典型误差在结构路径与优化路径上均具有对应的抑制机制,模型对不同数据集的成像差异与对比度波动也更不敏感,表现为更一致的泛化收益.

为了进一步概括所提方法在训练与推理阶段的整体执行逻辑,算法4给出了完整的计算流程,将模糊先验生成、混合编码、先验门控双分支解码、像素级自适应融合以及模糊感知优化约束统一到同一框架中,形成“先验构造-结构解码-细节补偿-自适应融合-损失约束”的闭环.

算法4 FGDB-TransUnet模型

输入: 眼底图像patch X;人工标注 Y;有效视野掩膜 M;混合编码器 $Encoder(\cdot)$;双分支解码器;融合权重头;损失权重 λ_f,λ_c;

输出: 最终血管概率图 P;训练阶段总损失 $\mathcal{L}_{\mathrm{total}}$.

1. 对输入图像 X 进行通道选择、归一化和尺寸整理,得到标准化输入;

2. 根据局部标准差响应与多尺度差分响应构造模糊先验图 F;

3. 将输入X送入混合编码器，获得编码特征E与多尺度跳跃特征集合$\{S_l\}_{l=1}^{L}$；

4. 初始化结构分支和模糊分支，使两分支共享相同的编码输出E；

5. 结构分支按照常规跳跃连接方式逐尺度解码，用于恢复主体血管结构与全局连通性；

6. 在每个解码尺度，将模糊先验图F插值到当前尺度，得到F_l；

7. 利用F_l对对应尺度的跳跃特征S_l进行门控，得到模糊门控跳跃特征$\tilde{S}_l$；

8. 模糊分支利用$\tilde{S}_l$逐尺度解码，用于强化弱边界、细小血管和交叉点区域的细节恢复；

9. 两分支分别通过分割头得到结构分支logits Z_s与模糊分支logits Z_f；

10. 根据两分支末端特征学习像素级融合权重图w；

11. 在logits层完成逐像素自适应融合，得到融合预测：$Z=w\odot Z^{std}+(1-w)\odot Z^{fuz}$；

12. 由融合logits计算最终血管概率图P；

13. 若处于训练阶段，则在有效视野掩膜M内计算基础分割损失；

14. 结合模糊先验图F计算模糊区域加权监督损失与两分支一致性约束；

15. 得到总损失：$L_{\text{total}}=L_{\text{main}}+\lambda_f L_{\text{fuzzy}}+\lambda_c L_{\text{cons}}$.

3　实验结果与分析

为了验证提出的基于模糊引导与双分支Transformer解码的视网膜血管分割方法在不同成像条件、不同样本规模及复杂血管结构下的有效性，在DRIVE，STARE，CHASE_DB1和FIVES四个公开眼底视网膜血管数据集上进行实验．前三个数据集用来保持与既有研究的可比性；FIVES包含800幅2048×2048多疾病眼底图像，用来补充大样本验证．所有评价指标均在视野掩膜（Field of View Mask，FOV）范围内计算，以避免视野外无效背景像素对分割结果造成偏置．

实验平台为Windows 11操作系统，开发环境为PyCharm，模型基于PyTorch框架实现．训练阶段采用Adam优化器进行参数更新，输入图像统一裁剪为256×256的patch形式．相比整图训练，patch训练能够在有限样本条件下增加训练样本数量，并使模型在局部区域内更充分地学习细血管、边界过渡带和局部交叉结构．模型测试阶段，将测试图像在FOV区域内进行预测，并将逐像素预测结果与人工标注进行比较，得到准确率、灵敏度、特异性、AUC（Area under Curve）和$F1$分数等评价指标．

3.1　实验数据集　选取DRIVE，STARE，CHASE_DB1和FIVES四个公开视网膜血管分割数据集作为实验对象．DRIVE数据集的图像分辨率为565×584，包含20幅训练图像和20幅测试图像，图像质量较稳定，常用于验证视网膜血管分割模型的基本性能．STARE数据集图像的分辨率为700×605，包含更多病灶、亮度不均和局部噪声干扰，能够更好地检验模型在复杂病理背景下的鲁棒性．CHASE_DB1数据集图像的分辨率为999×960，血管分支更加密集，局部交叉和末梢细血管更复杂，对模型的细结构恢复能力提出了更高的要求．FIVES数据集包含800幅2048×2048高分辨率彩色眼底图像以及像素级血管标注，覆盖年龄相关性黄斑变性、糖尿病视网膜病变、青光眼和正常四类，每类200幅．本文保留官方600/200训练-测试划分，并从官方训练集按疾病类别分层划分540幅训练图像和60幅验证图像，官方200幅测试图像在训练和检查点选择阶段保持不可见．所有数据集的详细信息如表1所示．

表1　数据集描述

Table 1　Dataset descriptions

数据集	分辨率	图像格式	实验划分与用途
DRIVE	565×584	RGB	原对比实验20/20 递进消融16/4/20
STARE	700×605	RGB	原对比实验10/10
CHASE_DB1	999×960	RGB	原对比实验20/8
FIVES	2048×2048	RGB	540/60/200 用于大样本验证

为了直观地展示不同经典数据集的成像差异，图3给出DRIVE，STARE和CHASE_DB1三个经典小样本数据集的原始眼底图像、人工标注和FOV掩膜示例．FIVES数据集用于后续大样本定量比较，其数据规模和划分亦见表1.

在预处理阶段，保留眼底图像中与血管结构相关性更强的绿色通道，并对像素值进行归一化

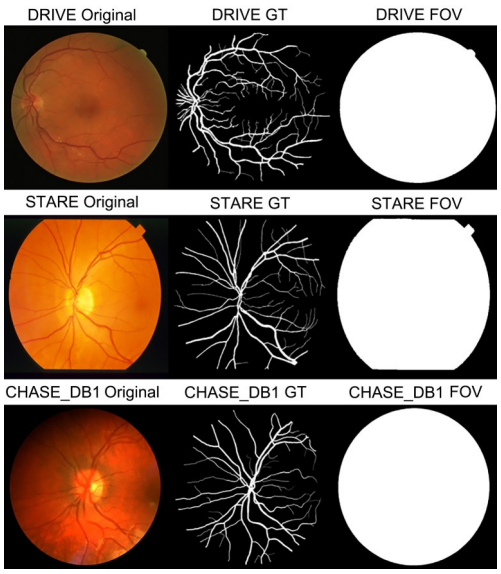


图 3　三个数据集的眼底图像、人工标注和 FOV 掩膜示例

Fig. 3　Examples of fundus images, ground truth and FOV masks on three datasets

处理. 视网膜血管边界通常表现为灰度渐变而非清晰跳变,单纯依赖原始灰度信息容易导致细血管被背景纹理淹没. 为此,进一步根据局部标准差响应和多尺度差分响应构造模糊先验图,用于刻画图像中潜在的弱边界、低对比细血管和局部纹理不稳定区域. 图 4 展示了原始图像经过绿色通道提取、局部响应计算和模糊先验生成后的可视化结果. 由图可见,模糊先验图在主干血管边缘、细小分支和局部交叉区域具有较强响应,能够为后续模糊分支解码和损失加权提供空间提示.

考虑到视网膜血管像素在整幅眼底图像中占比较低,若直接采用整图训练,模型容易受大量背景像素主导,削弱对细小血管的学习. 采用 256×256 patch 训练策略,并使 patch 采样区域尽可能落在 FOV 有效视野范围内. 图 5 展示了 patch 采样过程,其中,原始图像被划分为局部图像块,模型在局部区域内学习血管与背景之间的细粒度差异. 该策略一方面能够扩充训练样本数量,另一方面能够增加细小血管、弱边界和交叉区域在训练过程中的出现频率,提升模型对困难区域的关注能力.

3.2　实验结果评估指标　视网膜血管分割任务的目标是从眼底图像中分离出血管区域. 为了便于评价,将该任务视为像素级二分类问题,血管像素记为正类(Positive),非血管像素记为负类(Negative). 对模型输出的分割结果与人工标注的真实标签进行逐像素比较,将预测结果划分为四类:真阳性(TP),指血管像素被正确预测为血管;假阳性(FP),指非血管像素被错误预测为血管;真阴性(TN),指非血管像素被正确预测为非血管;假阴性(FN),指血管像素被错误预测为非血管. 所有指标均在 FOV 限定的有效区域内统计,避免视野外无效像素对评价产生偏置.

基于上述四类像素统计,可计算准确率(Accuracy,Acc)、灵敏度(Sensitivity,Sen)、特异性(Specificity,Spe)以及 $F1$ 分数($F1$ Score),其具体定义如下:

$$\begin{aligned} Acc &= \frac{TP+TN}{TP+FP+TN+FN} \\ Sen &= \frac{TP}{TP+FN} \\ Spe &= \frac{TN}{TN+FP} \\ F1 &= \frac{2TP}{2TP+FP+FN} \end{aligned} \tag{35}$$

其中,Acc 反映整体像素分类正确率;Sen 描述对血管正类的检出能力,通常对细小血管漏检更敏感;Spe 描述对背景负类的抑制能力,用于衡量误检控制水平;$F1$ 综合考虑漏检与误检,对类别不均衡的血管分割任务更具代表性.

3.3　实验结果　为了验证本文模型的有效性,首先将其与现有视网膜血管分割方法在 DRIVE 数据集上进行定量比较. DRIVE 数据集图像的质量相对稳定,血管主干结构清晰,但细小血管和边界过渡区域仍然容易出现漏检,因此该数据集能反映模型在常规眼底图像条件下的基础分割能力.

实验结果如表 2 所示. 由表可见,在 DRIVE 数据集上,本文模型的 $Acc=0.9602$,$Sen=0.8388$,$Spe=0.9775$,$F1=0.8365$. 与现有方法相比,本文模型的 Sen 和 $F1$ 指标的表现较好,说明模糊先验与双分支解码机制能增强模型对细小血管和弱边界区域的识别能力. 虽然部分方法的 Spe 指标略高,但过高的 Spe 往往意味着模型对背景区域更保守,可能伴随细血管漏检. 相比之下,本文模型能保持较高背景抑制能力,还进一步提高了血管检出能力,使 $F1$ 分数获得稳定提升.

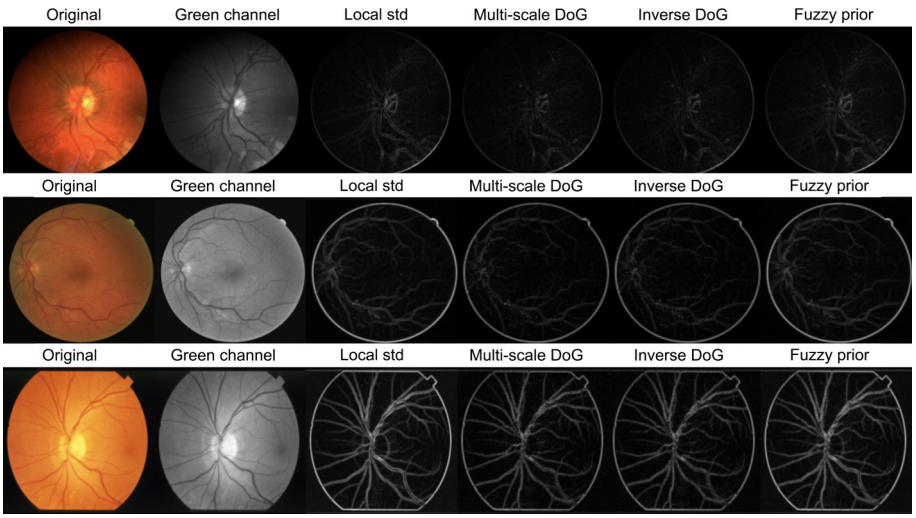


图4　预处理与模糊先验生成示例

Fig.4　Examples of preprocessing and fuzzy prior generation

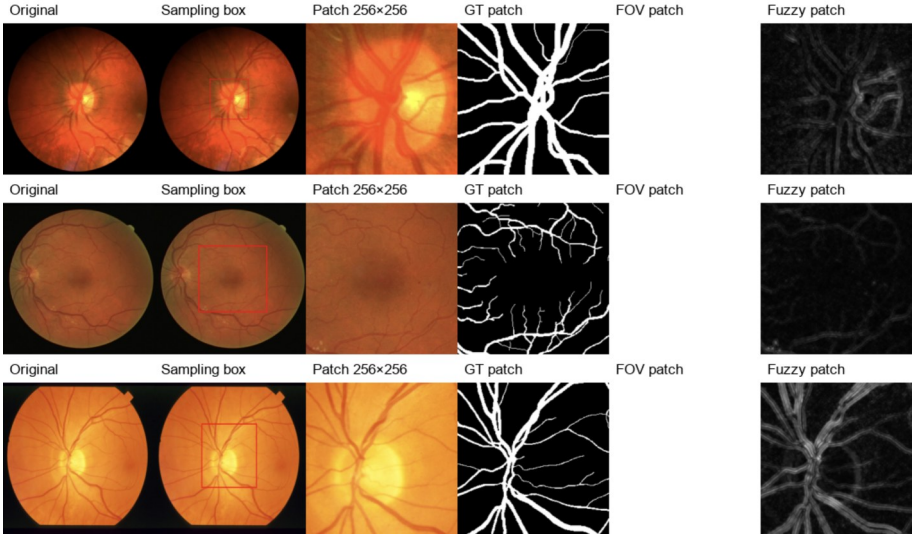


图5　视网膜血管图像的patch采样示例

Fig.5　Patch sampling examples of retinal vessel images

为了更直观地展示不同模型各指标的差异，根据表2绘制DRIVE数据集上的指标柱状对比图，如图6所示. 由图可见，本文模型的灵敏度和$F1$分数的优势较明显，这与本文方法针对模糊区域和细血管结构进行显式建模的设计目标一致. 说明模糊先验不是简单提高整体像素准确率，而是主要作用于传统模型容易忽略的低对比血管区域.

随后，在STARE数据集上进一步验证模型在复杂病理背景和光照不均条件下的分割性能，实验结果如表3所示. 与DRIVE数据集相比，STARE数据集中存在更多局部病灶、亮度变化和背景纹理干扰，血管与背景之间的灰度差异更不稳定，因此该数据集更能检验模型的鲁棒性.

由表可见，在STARE数据集上，本文模型的$Acc=0.9684$，$Sen=0.8042$，$Spe=0.9882$，$F1=0.8389$. 说明在更复杂的病灶背景和亮度变化条件下，模型仍能保持较高的整体分类能力和背景抑制能力. 尤其是$F1$指标保持在较高水平，表明本文方法不仅能够对明显血管区域进行稳定分割，也能在噪声干扰较强的情况下保持较好的血管与背景区分能力.

根据表3绘制的STARE数据集指标对比图

表 2 FGDB-TransUnet 与其他模型在 DRIVE 数据集上的比较

Table 2 Comparison of FGDB-TransUnet with other models on DRIVE dataset

模型	*Acc*	*Sen*	*Spe*	*F*1
Ronneberger et al[7]	0.9555	0.7822	0.9808	0.8174
Yang et al[22]	**0.9626**	0.7946	0.9821	0.8155
Wang et al[23]	0.9565	0.8071	0.9782	0.8251
Sun et al[24]	0.9537	0.7487	**0.9836**	0.8047
Zhang et al[25]	0.9565	0.7853	0.9818	0.8203
Li et al[26]	0.9578	0.8262	0.9783	0.8307
Yuan et al[27]	0.9581	0.8046	0.9805	0.8303
FGDB-TransUnet	0.9602	**0.8388**	0.9775	**0.8365**

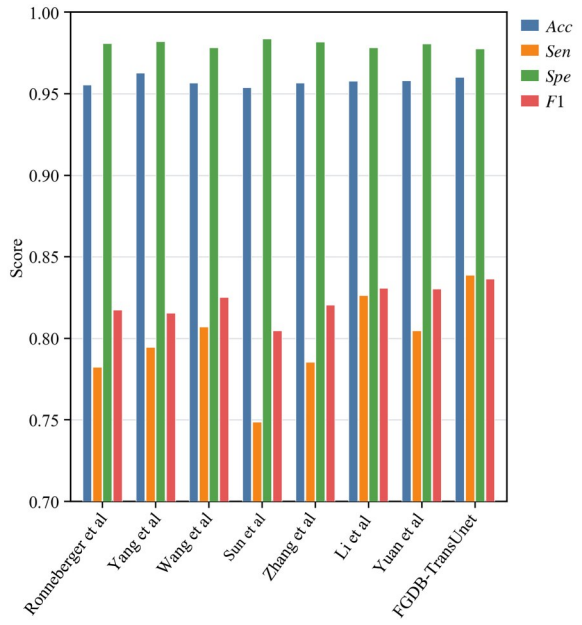


图 6 不同模型在 DRIVE 数据集上的指标对比

Fig. 6 Metric of different models on DRIVE dataset

表 3 FGDB-TransUnet 与其他模型在 STARE 数据集上的比较

Table 3 Comparison of FGDB-TransUnet with other models on STARE dataset

模型	*Acc*	*Sen*	*Spe*	*F*1
Ronneberger et al[7]	0.9526	0.6124	0.9929	0.7327
Sun et al[24]	0.9594	0.7487	0.9905	0.7850
Zhang et al[25]	0.9554	0.7791	0.9758	*
Yuan et al[27]	0.9612	0.6979	0.9846	*
FGDB-TransUnet	**0.9684**	**0.8042**	**0.9882**	**0.8389**

如图 7 所示. 由图可见,本文模型在整体性能上保持稳定,但 STARE 数据集中局部病灶和非均匀照明可能使部分低对比细血管的响应受到抑制. 因此,在分析 STARE 数据集结果时,不能只观察单一指标,而应结合 *Sen*,*Spe* 和 *F*1 共同判断模型性能. 本文模型中较高的 *Spe* 表明其对复杂背景具有较强抑制能力,而 *F*1 分数的提升进一步说,明其在减少误检的同时,仍然保持了较好的血管检出能力.

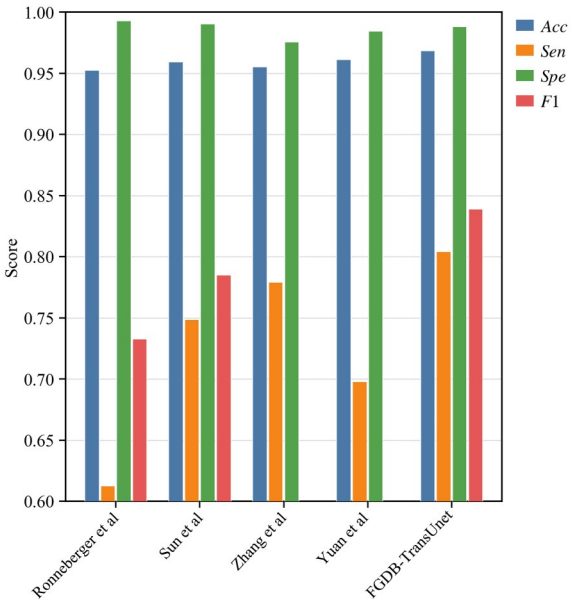


图 7 不同模型在 STARE 数据集上的指标对比

Fig. 7 Metric of different models on STARE dataset

最后,在 CHASE_DB1 数据集上对模型进行验证,实验结果如表 4 所示. CHASE_DB1 数据集分辨率较高,血管分支密集,交叉点和末梢血管数量较多,对模型的细粒度结构恢复能力提出了更高的要求.

由表可见,在 CHASE_DB1 数据集上,本文模型 $Acc=0.9661$, $Sen=0.8490$, $Spe=0.9778$. 其中,*Sen* 指标的表现较突出,说明模型在高分辨率、密集分支和复杂交叉结构条件下能够检出更多血管像素. CHASE_DB1 数据集中细血管数量较多,部分末梢血管边界较弱,传统模型容易将其误判为背景. 本文通过模糊先验对不确定区域进行引导,使模糊分支在解码阶段能够获得更多细节补偿,改善细小血管断裂和漏检现象.

表4　FGDB-TransUnet与其他模型在CHASE_DB1数据集上的比较

Table 4　Comparison of FGDB-TransUnet with other models on CHASE_DB1 dataset

模型	*Acc*	*Spe*	*Sen*
Zhang et al[28]	0.9452	0.9661	0.7626
Sun et al[24]	0.9566	0.9798	0.7475
Yan et al[29]	0.9610	**0.9809**	0.7633
Fan et al[30]	0.9510	0.9730	0.6570
Wang et al[31]	0.9603	0.9792	0.7730
Yan et al[32]	0.9607	0.9806	0.7641
FGDB-TransUnet	**0.9661**	0.9778	**0.8490**

图8展示了不同模型在CHASE_DB1数据集上的指标对比．由图可见，本文模型在血管检出相关指标上的表现较好．但由于CHASE_DB1数据集中背景纹理和局部反光较复杂，模型在提高*Sen*的同时也可能引入少量误检，因此*F*1的提升幅度需要结合误差图进一步分析．这也说明，在高分辨率复杂数据集上，血管分割性能不仅取决于主干血管是否完整，更取决于细小分支、交叉点和边界过渡区域能否被稳定恢复．

3.4　分割效果可视化分析　为了进一步观察模型在不同数据集上的实际分割效果，图9给出了本文模型在DRIVE，STARE和CHASE_DB1三个数据集上的可视化结果．每组图像依次展示原始眼底图像、人工标注、预测概率图、二值分割图、误差图和分割叠加图．与单纯的指标表相比，可视化结果能够更直观地反映模型在细血管、交叉点和边界过渡区域的表现．

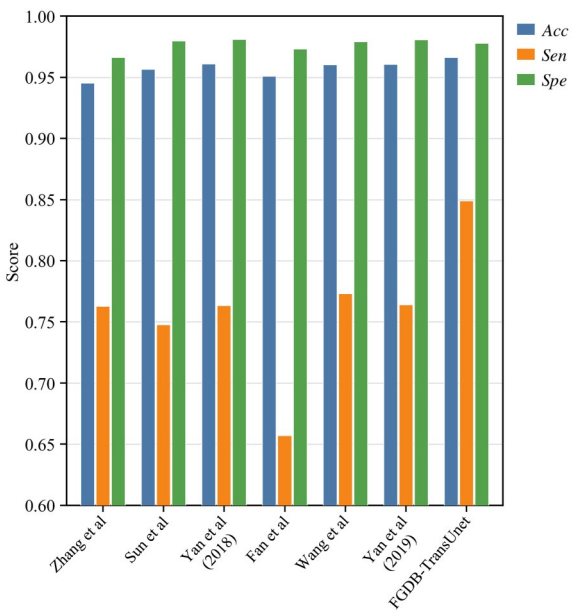


图8　不同模型在CHASE_DB1数据集上的指标对比

Fig. 8　Metric comparison of different models on CHASE_DB1 dataset

由图可见，本文模型能够较完整地恢复主干血管和多数分支结构，预测结果与人工标注整体一致．在DRIVE数据集中，模型对主干血管和中等尺度分支具有较稳定的分割效果，误差主要集中在血管末端和极细分支区域．在STARE数据集中，由于局部病灶和亮度不均影响，部分弱对比区域容易产生边界收缩，但模型整体上能够较好地抑制背景纹理误检．在CHASE_DB1数据集中，模型对密集分支和交叉点区域表现出较强的检出能力，但在局部反光和复杂背景区域仍可能出现少量假阳性．

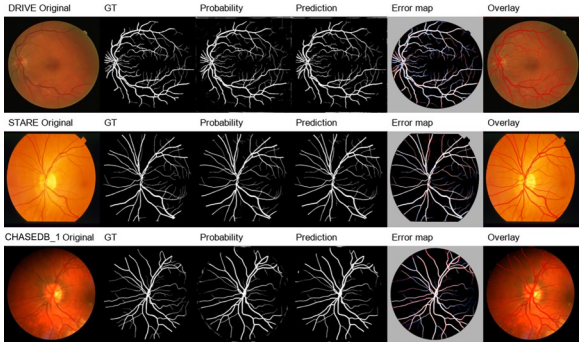


图9　FGDB-TransUnet在三个数据集上的分割效果示例

Fig. 9　Segmentation examples of the FGDB-TransUnet on three datasets

为了更加明确地分析模型对困难区域的处理能力，图10进一步展示了局部放大结果．选取区域主要包括细小血管末端、血管交叉点和弱边界过渡带．由图可见，本文方法在部分细小血管区域能够保持较好的连续性，交叉点附近的结构断裂现象有所减少．这说明模糊先验图对难分区域的空间提示能够有效引导解码分支关注局部细节，同时像素级融合机制能够在主体结构恢复和边界精修之间形成互补．

3.5　FIVES大样本验证与模糊理论基线比较　为了减轻仅依赖小样本数据集评价的局限，并回

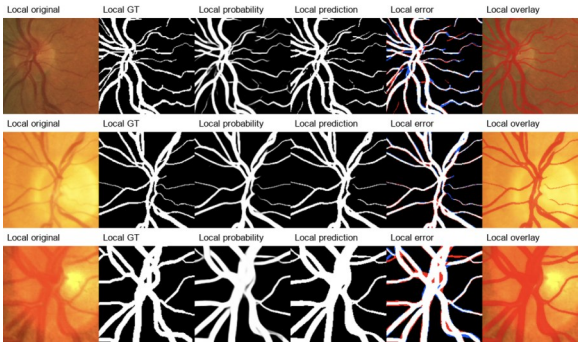


图 10　细小血管、交叉点和弱边界区域的局部分割效果

Fig. 10　Local segmentation results on thin vessels, bifurcations and weak boundaries

应模糊理论基线不足的问题,进一步在FIVES数据集上比较受控中实现的Fuzzy U-Net与FGDB-TransUnet. Fuzzy U-Net以标准U-Net为主体,并采用与本文相同的无标注模糊先验对多尺度跳跃特征进行门控. 其训练损失与普通U-Net保持一致,不使用双分支解码、自适应融合、模糊区域加权损失或分支一致性约束. 因此,该模型用于评价模糊先验嵌入卷积式编码-解码框架后的性能,并与所提完整框架的同协议模糊理论对照.

为了控制训练条件,两种模型使用完全相同的540/60/200训练,验证和测试划分,随机种子2345,训练迭代数,patch采样数,批大小,数据预处理和验证集选模规则. 模型训练完成后,均选择验证集$F1$最优的检查点,并在此前保持不可见的200幅测试图像上评估一次. 表5给出两种模型的图像级宏平均结果.

由表可见,Fuzzy U-Net的Acc,Sen,Spe,AUC,$F1$和clDice分别为0.9768,0.7882,0.9937,0.9846,0.8432和0.8362;FGDB-TransUnet的对应结果分别为0.9852,0.8943,0.9934,0.9932,0.9099,0.9112. 在相同实验协议下,FGDB-TransUnet相对Fuzzy U-Net的$F1$和clDice分别变化0.0667和0.0750. $F1$与clDice的同步提高表明,所提方法的收益不仅体现在像素级分类平衡上,也体现在血管中心线和分支连通性的保持上. 说明双分支协同、自适应融合及模糊感知优化在单分支模糊门控基础上提供了进一步补偿.

表 5　FGDB-TransUnet与Fuzzy U-Net在FIVES数据集上的同协议比较

Table 5　Same-protocol comparison between FGDB-TransUnet and Fuzzy U-Net on FIVES dataset

模型	*Acc*	*Spe*	*Sen*	*AUC*	*F1*	clDice
Fuzzy U-Net	0.9768	**0.9937**	0.7882	0.9846	0.8432	0.8362
FGDB-TransUnet	**0.9852**	0.9934	**0.8943**	**0.9932**	**0.9099**	**0.9112**

3.6　模块递进消融实验　为了分析各组成模块在相同训练条件下的增量作用,在DRIVE数据集的16/4/20训练、验证和测试划分上设置A0～A5六组递进配置,并统一随机种子1234、训练预算、数据预处理、验证集选模和测试流程. A0为基础TransUNet;A1在共享编码器后引入结构分支和模糊分支,并采用固定0.5等权融合;A2在A1基础上使用模糊先验门控多尺度跳跃特征;A3将固定融合替换为像素级自适应融合;A4进一步加入模糊区域加权损失;A5加入分支一致性约束,构成完整FGDB-TransUnet. 各配置仅相对前一配置增加一个设计因素,结果如表6所示.

由表可见引入双分支解码后,A1相较A0的$F1$,AUC和clDice分别提高0.0011,0.0004和0.0002,表现出轻微但较为一致的改善. 在固定

表 6　FGDB-TransUnet模块递进消融实验

Table 6　Progressive module ablation of FGDB-TransUnet

配置	双分支	模糊门控	融合方式	模糊加权损失	一致性约束	*AUC*	*F1*	clDice
A0	—	—	—	—	—	0.9801	0.8340	0.8296
A1	√	—	固定0.5	—	—	0.9805	0.8351	0.8298
A2	√	√	固定0.5	—	—	0.9736	0.8310	0.8176
A3	√	√	自适应	—	—	0.9799	0.8355	0.8276
A4	√	√	自适应	√	—	0.9799	0.8360	0.8306
A5	√	√	自适应	√	√	0.9784	0.8334	0.8240

0.5融合条件下加入模糊门控后，A2的$F1$，AUC和clDice分别下降0.0041，0.0069和0.0122，说明固定等权融合可能导致模糊门控对血管响应产生过度抑制．将固定融合替换为像素级自适应融合后，A3相较A2的$F1$，AUC和clDice分别回升0.0045，0.0063和0.0100，表明自适应融合能够缓解固定融合下的性能退化．进一步加入模糊区域加权损失后，A4的$F1$和clDice分别提高0.0005和0.0030，并取得六组配置中的最高值0.8360和0.8306．加入一致性约束后，A5的$F1$，AUC和clDice分别下降0.0026，0.0015和0.0066，提示当前一致性约束权重可能引入较强的正则化作用．总体上，递进消融支持双分支解码、自适应融合和模糊区域加权损失在当前实验设置下的有效性，同时说明模糊门控的作用依赖于融合策略，而一致性约束未在当前随机种子和固定超参数下带来额外增益．各模块并非在全部指标上均产生单调提升．

图11将表6中的$F1$，AUC和clDice变化进行可视化．三项指标从A0到A5呈现明显的局部波动，而非单调上升：A1相较A0略有改善，A2出现较明显下降，A3恢复至接近或超过基础配置的水平，A4取得最高的$F1$和clDice，而A5再次出现一定程度的回落．其中，变化最明显的恢复阶段为A2至A3，对应将固定0.5融合替换为像素级自适应融合，$F1$，AUC和clDice分别提高0.0045，0.0063和0.0100．这说明自适应融合能够根据空间位置协调结构分支与模糊分支的贡

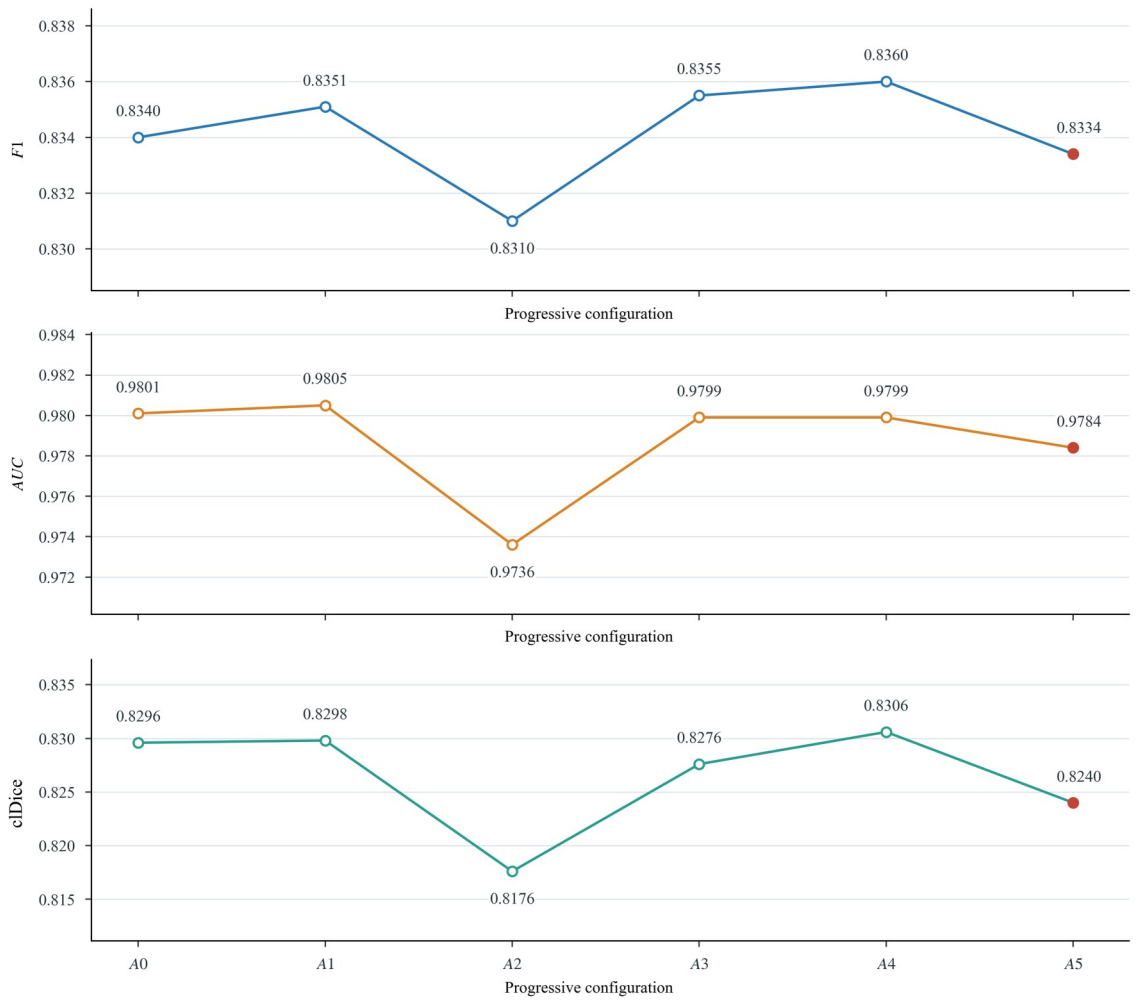


图11　FGDB-TransUnet模块递进消融结果

Fig. 11　Progressive ablation results of FGDB-TransUnet

献,缓解固定融合条件下模糊门控对血管响应的过度抑制,并同时改善像素级分类平衡、阈值无关判别能力和血管拓扑连通性. 图中的局部波动与表 6 的定量结果一致.

3.7 模糊先验参数敏感性分析 为了评估已训练模型对推理阶段模糊先验配置变化的敏感程度,固定 FIVES 数据集上随机种子 2345 训练得到的 FGDB-TransUnet 网络权重,并从官方测试集中按年龄相关性黄斑变性、糖尿病视网膜病变、青光眼和正常人四类各固定抽取 10 幅图像,构成包含 40 幅图像的疾病分层测试子集. 默认模糊先验参数设置为 Green 通道、局部标准差窗口 7、多尺度窗口组合(3,9)和融合系数 $\alpha=0.5$. 采用单因素控制方式,每次仅改变一个参数,网络权重、测试图像、滑动窗口推理策略、判定阈值和其余参数均保持不变.

需要说明,该实验用于评价冻结权重模型对推理阶段先验配置偏移的适应能力,而不是重新训练条件下的超参数搜索. 40 幅疾病分层图像仅用于降低重复推理的计算开销,FGDB-TransUnet 在 FIVES 完整测试集上的性能仍以表 5 中 200 幅测试图像的结果为准.

如表 7 所示,在默认参数设置下,模型在 40 幅疾病分层测试图像上的 $F1$ 和 clDice 分别为 0.8990 和 0.9030. 改变模糊先验输入通道后,$F1$ 和 clDice 相对默认设置的变化范围均为 $-0.0001\sim+0.0002$,最大绝对变化为 0.0002. 使用 Blue 通道时,两项指标分别增加 0.0002 和 0.0002;使用 Red 或 Gray 通道时,两项指标仅出现不超过 0.0001 的下降. 虽然 Blue 通道在当前子集上获得了略高结果,但该差异幅度较小,且测试子集不用于超参数选择,因此不据此调整默认通道设置.

对局部标准差窗口进行扰动时,$F1$ 的最大绝对变化为 0.0001,clDice 在表 7 所示精度下基本保持不变. 改变多尺度窗口组合后,$F1$ 和 clDice 的变化幅度均小于 0.0001. 当融合系数 α 由 0.5 调整为 0.3 或 0.7 时,两项指标的最大绝对变化均不超过 0.0001. 总体上,在本实验覆盖的参数范围内,所有扰动设置引起的 $F1$ 和 clDice 最大绝对变化均不超过 0.0002,表明固定权重模型对输入通道、局部统计窗口、多尺度组合及融合系数的小范围偏移具有较好的数值稳定性. 由于该实验仅基于一个训练随机种子和 40 幅测试图像,上述结果反映的是当前固定训练条件下的推理稳健性,不能替代多随机种子重复训练或完整测试集上的参数稳定性分析.

3.8 实验结果 综合 DRIVE, STARE 和 CHASE_DB1 三个经典数据集以及 FIVES 大样实验可以看出,FGDB-TransUnet 在不同图像分

表 7 固定网络权重的模糊先验参数敏感性实验

Table 7 Frozen-weight sensitivity analysis of fuzzy-prior parameters

因素	设置	默认设置	$F1$	clDice	$\Delta F1$	ΔclDice	测试图像数
输入通道	Green	√	0.8990	0.9030	0	0	40
输入通道	Red		0.8989	0.9028	−0.0001	−0.0001	40
输入通道	Blue		0.8992	0.9031	0.0002	0.0002	40
输入通道	Gray		0.8989	0.9029	−0.0001	−0.0001	40
局部标准差窗口	7	√	0.8990	0.9030	0	0	40
局部标准差窗口	5		0.8990	0.9029	0	0	40
局部标准差窗口	9		0.8991	0.9030	0.0001	0	40
多尺度窗口	(3,9)	√	0.8990	0.9030	0	0	40
多尺度窗口	(3,7)		0.8990	0.9029	0	0	40
多尺度窗口	(5,11)		0.8991	0.9029	0	0	40
融合系数 α	0.5	√	0.8990	0.9030	0	0	40
融合系数 α	0.3		0.8989	0.9029	−0.0001	−0.0001	40
融合系数 α	0.7		0.8991	0.9030	0	0	40

辨率、病理背景和样本规模下完成了较系统的验证，并在像素级分类与血管拓扑保持指标上取得总体稳定的结果. DRIVE，STARE和CHASE_DB1实验用于保持与既有研究的可比性；FIVES实验将相互独立的训练、验证和测试图像数量扩展至540，60和200幅，为多疾病、高分辨率和较大样本规模场景下的评价提供了补充证据.

在完全相同的FIVES实验协议下，FGDB-TransUnet相较受控重实现的Fuzzy U-Net，$F1$和clDice分别提高0.0667和0.0750. 该结果表明，在单分支模糊先验门控的基础上，引入双分支协同解码、像素级自适应融合和模糊感知优化能够进一步改善像素级误检-漏检平衡以及血管中心线和分支连通性的保持. 对于Sen与Spe存在的变化，应结合AUC，$F1$和clDice共同判断，避免依据单一指标评价模型性能.

A0～A5递进消融结果表明，引入双分支解码后，模型的$F1$，AUC和clDice获得小幅改善；在固定0.5融合条件下直接加入模糊门控会降低血管灵敏度和拓扑指标，而像素级自适应融合能够明显恢复相应性能. 进一步加入模糊区域加权损失后，A4取得六组配置中最高的$F1$和clDice，说明该损失有助于改善困难区域的像素分类和血管拓扑保持. 加入一致性约束后的A5未获得额外增益，提示当前一致性约束权重可能产生较强的正则化作用. 表7中的冻结权重参数扰动结果进一步显示，不同输入通道、局部标准差窗口、多尺度窗口和融合系数引起的$F1$与clDice最大绝对变化均不超过0.0002，未发现对当前测试范围特别敏感的先验因素. 递进消融和参数敏感性实验分别回答了“性能变化来自哪些模块”以及“已训练模型对先验配置偏移是否敏感”的问题，与FIVES同协议比较共同构成对所提方法的补充验证.

4　结论

针对视网膜血管分割中细小血管易漏检、弱边界易收缩以及交叉区域易断裂等问题，本文提出一种基于模糊引导与双分支Transformer解码的视网膜血管分割方法. 该方法以TransUNet的混合编码结构为基础，从输入眼底图像中构造无需额外标注参与的模糊先验图，用于刻画局部纹理不稳定和边界过渡模糊区域；在解码阶段，引入结构分支与模糊分支的协同机制，使结构分支侧重主体血管的连续恢复，而模糊分支通过先验门控增强细小血管与弱边界区域的特征回注；同时，利用像素级自适应融合和面向模糊区域的优化约束，进一步协调结构恢复与边界精修之间的关系. 实验结果表明，本文方法在DRIVE，STARE和CHASE_DB1三个经典公开数据集上保持了较好的分割性能，并在包含800幅多疾病高分辨率图像的FIVES数据集上完成了进一步验证.

尽管本文方法在多个公开数据集上表现出较好的稳定性，但仍存在改进空间. 首先，当前模糊先验主要由局部统计响应和多尺度差分响应构造，对强病灶干扰、局部反光和极端低对比区域的适应性仍有提升空间，后续可引入更细粒度的图像质量评估或动态先验权重调节策略. 其次，本文主要围绕二值血管分割任务展开，没有进一步区分动静脉类别、血管宽度等级和病灶相关区域，未来可结合多任务学习方式，将血管分割与血管定量分析、病灶检测和疾病风险评估进行联合建模. 最后，本文实验主要基于公开眼底图像数据集完成，后续仍需在更多来源、更复杂成像条件和真实临床场景下进行验证，以进一步评估模型的泛化能力和实际应用价值.

参考文献

[1] Khandouzi A, Ariafar A, Mashayekhpour Z, et al. Retinal vessel segmentation: A review of classic and deep methods. Annals of Biomedical Engineering, 2022, 50(10): 1292－1314.

[2] Chen C H, Chuah J H, Ali R, et al. Retinal vessel segmentation using deep learning: A review. IEEE Access, 2021(9): 111985－112004.

[3] Qin Q, Chen Y Y. A review of retinal vessel segmentation for fundus image analysis. Engineering Applications of Artificial Intelligence, 2024, 128: 107454.

[4] Staal J, Abrämoff M D, Niemeijer M, et al. Ridge-based vessel segmentation in color images of the retina. IEEE Transactions on Medical Imaging, 2004, 23(4): 501－509.

[5] Hoover A, Kouznetsova V, Goldbaum M. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. IEEE Transactions on Medical Imaging, 2000, 19(3): 203－210.

[6] Fraz M M, Remagnino P, Hoppe A, et al. An ensemble classification-based approach applied to retinal blood vessel segmentation. IEEE Transactions on Biomedical Engineering, 2012, 59(9): 2538－2548.

[7] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation// The 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich, Germany: Springer, 2015: 234－241.

[8] Oktay O, Schlemper J, Folgoc L L, et al. Attention U-Net: Learning where to look for the pancreas. https://arxiv.org/abs/1804.03999, 2018－05－20.

[9] Gu Z W, Cheng J, Fu H Z, et al. CE-Net: Context encoder network for 2D medical image segmentation. IEEE Transactions on Medical Imaging, 2019, 38(10): 2281－2292.

[10] Jin Q G, Meng Z P, Pham T D, et al. DUNet: A deformable network for retinal vessel segmentation. Knowledge-Based Systems, 2019, 178: 149－162.

[11] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16×16 words: Transformers for image recognition at scale. https://arxiv. org/abs/2010.11929v2, 2021－06－03.

[12] Chen J N, Lu Y Y, Yu Q H, et al. Transunet: Transformers make strong encoders for medical image segmentation. https://arxiv. org/abs/2102.04306, 2021－02－08.

[13] Guo C L, Szemenyei M, Yi Y G, et al. SA-UNet: Spatial attention U-Net for retinal vessel segmentation// 2020 25th International Conference on Pattern Recognition. Milan, Italy: IEEE, 2021: 1236－1242.

[14] Cao H, Wang Y Y, Chen J, et al. Swin-unet: Unet-like pure transformer for medical image segmentation// Computer Vision - ECCV 2022 Workshops. Tel Aviv, Israel: Springer, 2023: 205－218.

[15] Milletari F, Navab N, Ahmadi S A. V-Net: Fully convolutional neural networks for volumetric medical image segmentation// 2016 4th International Conference on 3D Vision. Stanford, CA, USA: IEEE, 2016: 565－571.

[16] Salehi S S M, Erdogmus D, Gholipour A. Tversky loss function for image segmentation using 3D fully convolutional deep networks// The 8th International Workshop on Machine Learning in Medical Imaging. Quebec City, Canada: Springer, 2017: 379－387.

[17] Zadeh L A. Fuzzy sets. Information and Control, 1965, 8(3): 338－353.

[18] Bezdek J C. Pattern recognition with fuzzy objective function algorithms. Cham: Springer Science& Business Media, 2013.

[19] Yan Z Q, Yang X, Cheng K T. Joint segment-level and pixel-wise losses for deep learning based retinal vessel segmentation. IEEE Transactions on Biomedical Engineering, 2018, 65(9): 1912－1923.

[20] Wu Y C, Xia Y, Song Y, et al. NFN+: A novel network followed network for retinal vessel segmentation. Neural Networks, 2020, 126: 153－162.

[21] Wang B, Wang S P, Qiu S, et al. CSU-Net: A context spatial U-Net for accurate blood vessel segmentation in fundus images. IEEE Journal of Biomedical and Health Informatics, 2021, 25(4): 1128－1138.

[22] Wu H S, Wang W, Zhong J F, et al. SCS-Net: A scale and context sensitive network for retinal vessel segmentation. Medical Image Analysis, 2021, 70: 102025.

[23] Wang D Y, Haytham A, Pottenburgh J, et al. Hard attention net for automatic retinal vessel segmentation. IEEE Journal of Biomedical and Health Informatics, 2020, 24(12): 3384－3396.

[24] Yang L, Wang H X, Zeng Q S, et al. A hybrid deep segmentation network for fundus vessels via deep-learning framework. Neurocomputing, 2021, 448: 168－178.

[25] Ding W P, Sun Y, Huang J S, et al. RCAR-UNet: Retinal vessel segmentation network algorithm via novel rough attention mechanism. Information Sciences, 2024, 657: 120007.

[26] Zhang Y, He M, Chen Z N, et al. Bridge-Net: Context-involved U-net with patch-based loss weight mapping for retinal blood vessel segmentation. Expert Systems with Applications, 2022, 195: 116526.

[27] Li J Y, Gao G, Liu Y H, et al. MAGF-Net: A multiscale attention-guided fusion network for retinal vessel segmentation. Measurement, 2023, 206: 112316.

[28] Yuan Y C, Zhang L, Wang L T, et al. Multi-level attention network for retinal vessel segmentation. IEEE Journal of Biomedical and Health Informatics, 2022, 26(1): 312－323.

[29] Zhang J, Dashtbozorg B, Bekkers E, et al. Robust retinal vessel segmentation via locally adaptive derivative frames in orientation scores. IEEE Transactions on Medical Imaging, 2016, 35(12): 2631－2644.

[30] Fan Z, Lu J W, Wei C M, et al. A hierarchical image matting model for blood vessel segmentation in fundus images. IEEE Transactions on Image Processing, 2019, 28(5): 2367－2377.

[31] Wang X H, Jiang X D, Ren J F. Blood vessel segmentation from fundus image by a cascade classification framework. Pattern Recognition, 2019, 88: 331－341.

[32] Yan Z Q, Yang X, Cheng K T. A three-stage deep learning model for accurate retinal vessel segmentation. IEEE Journal of Biomedical and Health Informatics, 2019, 23(4): 1427－1436.

（责任编辑　杨可盛）