

基于大核卷积残差块和坐标注意力的血管分割算法

张凯文¹, 龙俊¹, 袁品焄², 谢怡宁¹

¹东北林业大学机电工程学院, 黑龙江 哈尔滨

²东北林业大学计算机与控制工程学院, 黑龙江 哈尔滨

收稿日期: 2026年2月28日; 录用日期: 2026年3月27日; 发布日期: 2026年4月7日

摘要

随着机器学习和深度学习的发展, 研究者们应用各种算法和模型成功地从眼底图像中检测出相关疾病, 其中视网膜血管的精确分割是眼科疾病自动化辅助诊断的关键步骤。然而传统分割算法存在血管与噪声难以区分, 以及长距离血管拓扑结构难以保持连贯性等问题。为此, 本文提出了一种基于大核卷积残差块和坐标注意力的眼底血管分割算法TLadder-AMSF。该算法采用具有大感受野的大核卷积残差块以建模长距离血管依赖, 设计非对称多尺度融合模块AMSF以实现多尺度特征的精准聚合, 并引入坐标注意力以增强空间方向感知能力。在DRIVE和STARE数据集上的实验结果表明, 该方法在灵敏度上分别达到82.64%和83.56%, 在准确率上分别达到96.25%和98.22%, 各项指标与多个主流方法相比均达到先进水平。为了直观验证本文方法的分割性能, 进行了可视化讨论。此外, 消融实验证明大核卷积残差块、AMSF模块及坐标注意力提升了分割图像的拓扑连通性, 验证了各个模块的有效性。

关键词

血管分割, 非对称多尺度融合, 坐标注意力

Vessel Segmentation Algorithm Based on Large Kernel Convolution Residual Block and Coordinate Attention

Kaiwen Zhang¹, Jun Long¹, Pinhan Yuan², Yining Xie¹

¹College of Mechanical and Electrical Engineering, Northeast Forestry University, Harbin Heilongjiang

²College of Computer and Control Engineering, Northeast Forestry University, Harbin Heilongjiang

Received: February 28, 2026; accepted: March 27, 2026; published: April 7, 2026

文章引用: 张凯文, 龙俊, 袁品焄, 谢怡宁. 基于大核卷积残差块和坐标注意力的血管分割算法[J]. 计算机科学与应用, 2026, 16(4): 31-41. DOI: 10.12677/csa.2026.164107

Abstract

With the development of machine learning and deep learning, researchers have successfully applied various algorithms and models to detect related diseases from fundus images, among which the precise segmentation of retinal blood vessels is a key step in automated ophthalmic disease diagnosis. However, traditional segmentation algorithms face issues such as difficulty distinguishing vessels from noise and maintaining the continuity of long-distance vessel topology. To address this, this paper proposes a fundus vessel segmentation algorithm TLadder-AMSF based on a large kernel convolution residual block and coordinate attention. This algorithm uses a Large Kernel Convolution Residual Block with a large receptive field to model long-distance vessel dependencies, designs an asymmetric multi-scale fusion module AMSF to achieve accurate aggregation of multi-scale features, and introduces a coordinate attention to enhance spatial directional perception. Experimental results on the DRIVE and STARE datasets show that the method achieves sensitivities of 82.64% and 83.56%, and accuracies of 96.25% and 98.22%, respectively, with all metrics reaching advanced levels compared to several mainstream methods. To visually validate the segmentation performance of this method, visualization discussions were conducted. In addition, ablation experiments demonstrated that the Large Kernel Convolution Residual Block, AMSF module, and coordinate attention improved the topological connectivity of segmented images, confirming the effectiveness of each module.

Keywords

Vessel Segmentation, Asymmetric Multi-Scale Fusion, Coordinate Attention

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

视网膜血管是人体唯一能够以无创方式直接观察到的微循环系统。视网膜血管形态的变化(如直径、分支角度及曲率)是临床诊断糖尿病视网膜病变、高血压、青光眼及多种心脑血管疾病的重要生物标志物 [1]-[5]。通过对眼底图像进行精确的血管分割,能够量化血管形态指标,为临床医生提供客观的辅助诊断依据,实现疾病的早期筛查与精准诊断。自动化、高精度的血管分割是计算机辅助诊断系统的关键步骤。然而,目前临床上的视网膜分析多依赖专业医师的手动勾画,这不仅耗时费力,且受限于操作者的专业知识与主观经验,难以满足大规模筛查的需求。此外,由于视网膜图像存在光照不均、病变干扰(如渗出、出血)、以及血管自身尺度变化大(从粗大主干至纤细末梢)和结构复杂(交叉、缠绕)等挑战,实现鲁棒且精确的分割仍是一项艰巨任务。

随着计算机视觉与深度学习技术的发展,研究者们提出了多种自动化分割算法,涵盖了传统方法和深度学习方法,包括无监督和监督学习方法,显著提升了血管分割的效率与客观性。无监督血管分割方法不依赖于任何标记数据,这些方法通过提取眼底图像中的血管和背景特征,发现它们之间的相互关系,然后进行血管分割。Wang 等人 [6] 利用多小波核的匹配滤波器实现自动血管分割,在分割血管的同时特别去除了非血管边界部分。Yin 等人 [7] 提出了一种基于血管跟踪算法的概率跟踪方法。血管跟踪算法涉及手动选择初始点,然后进行迭代血管边缘检测。然而,血管跟踪方法常遇到血管间隙和不连续性的问

题。为解决细血管问题, Neto 等人[8]通过结合形态重建前一阶段获得的粗略分割轮廓, 创建了一个曲率图, 从而获得了更精细的分割结果, 并实现了一种由粗到细的血管分割方法。此外, 许多研究人员尝试通过传统方法解决血管问题。考虑到病灶对血管分割的影响, Lam 等人[9]提出了一种基于统计归一化的视网膜血管分割方法, 该方法利用病灶、血管和背景的凹凸特征实现分割。

与无监督方法相比, 监督方法通过添加标记数据(通常由专家提供)来训练模型。Wu 等人[10]提出的 U-GAN 使用视网膜血管自动分割模型作为 GAN 的生成器, 同时将分割结果和真实值发送到判别网络。通过对抗训练不断优化分割结果。此外, 一些研究人员通过重复 U 形结构来提高最终分割结果。Zhuang 等人[11]采用了两个 U 形结构, 并将原来的通道数拼接替换为 U 形网络内的像素值相加。第一个 U 形网络的解码器和第二个 U 形网络的编码器通过像素值相加连接, 形成了 LadderNet。鉴于血管在分割难度上存在差异, 一些血管仍然难以准确分割。Wang 等人[12]提出了 HAnet, 该网络由一个编码器和解码器组成, 编码器用于提取特征, 其中一个解码器获得粗略分割结果并将特征分为易于分割或难于分割, 然后另外两个解码器专门用于进一步分割视网膜血管。然而, 视网膜血管分割任务仍面临严峻挑战。首先, 血管末梢血管极易被背景噪声淹没导致漏检。其次, 传统卷积核的感受野有限, 难以捕捉长距离的血管拓扑信息, 导致在分割结果中频繁出现血管断裂现象。单独依靠传统的卷积块或对称的特征融合方式, 往往无法在复杂背景下精确平衡局部细节与全局连通性, 导致模型在处理细小血管时的鲁棒性受限。针对这些问题, 本文提出一种基于大核卷积残差块和坐标注意力的迭代细化分割算法。通过 LadderNet 循环细化提取的特征图。通过引入大尺寸深度可分离卷积, 显著提升了模型的有效感受野, 增强了模型对血管长距离拓扑连通性的感知能力。此外, 提出了一种非对称多尺度特征融合(AMSF)策略, 根据模型不同深度特征层级的分布差异, 在编解码路径的瓶颈处设计了两个功能互补的融合模块。其中, 在 AMSF-2 模块中嵌入了坐标注意力(Coordinate attention)[13]机制, 增强了模型的空间感知能力, 提升了血管分割精度。

2. 实验设置

2.1. 数据集

为了验证本文提出的模型在视网膜血管分割任务中的有效性, 我们在两个公开的临床标准数据集 DRIVE 和 STARE 上进行了验证。DRIVE 数据集由一台 CanonCR5 非显微镜相机拍摄, 包含 40 张分辨率为 565×584 的彩色眼底图像, 划分为 20 张训练集和 20 张测试集。每张图像提供由第一位专家标注的分割掩模作为金标准。STARE 数据集由美国佛罗里达大学的研究团队构建, 包含 20 幅 700×605 的彩色眼底图像, 图像这些图像由 TopconTRV-50 眼底相机采集得到涵盖了不同年龄、性别和眼部状况的患者群体, 每幅图像由专家进行像素级标注。

2.2. 数据预处理

眼底图像往往存在光照不均、对比度低以及噪声干扰等问题, 因此进行预处理。提取彩色图像的绿色通道(该通道血管与背景对比度最高), 并进行标准差归一化。采用限制对比度自适应直方图均衡化(CLAHE)算法提升血管末梢与背景的辨识度。将所有图像及其对应金标准统一中心裁剪或填充至固定尺寸 512×512 。通过 Gamma 变换进一步修正图像的亮度分布。为缓解过拟合并增强模型对不同血管形态的适应性, 在训练阶段对训练样本实时施加随机增强, 包括: 水平/垂直翻转、随机旋转。

2.3. 实验环境设置

在实验中, 使用 AdamW 作为优化器, 权重衰减为 0.05, 动量为 0.99, 初始学习率为 0.0001, 采用

余弦退火学习率调度。实验训练批大小为 8，训练轮次为 50 轮，使用早停策略(early stopping = 8)。实验在 12GB 内存的 NVIDIA GeForce RTX 3080 Ti 上进行。

3. 模型结构

本文提出了一种基于大核卷积残差块和坐标注意力的迭代细化分割网络 TLadder-AMSF。首先，该方法通过第一个 U 形网络进行基础特征提取，该阶段采用编码器 - 解码器结构，利用大核卷积残差块 LKCRB 逐层提取空间特征并进行下采样，在瓶颈处通过 AMSF-1 模块捕获多尺度信息，随后通过上采样路径初步生成血管特征图。随后，特征进入第二个 U 形网络进行特征迭代精炼。在该阶段，编码器不仅接收内部的传递信息，还通过横向连接融合第一个 U 形网络对应层级的特征，实现跨路径的信息复用。在 U 形网络的最底层，通过集成坐标注意力机制的 AMSF-2 模块对深层特征进行空间校准与增强。最后，经过解码器的逐层还原，输出与原图尺寸一致的高精度血管分割结果。整体模型架构如图 1 所示。

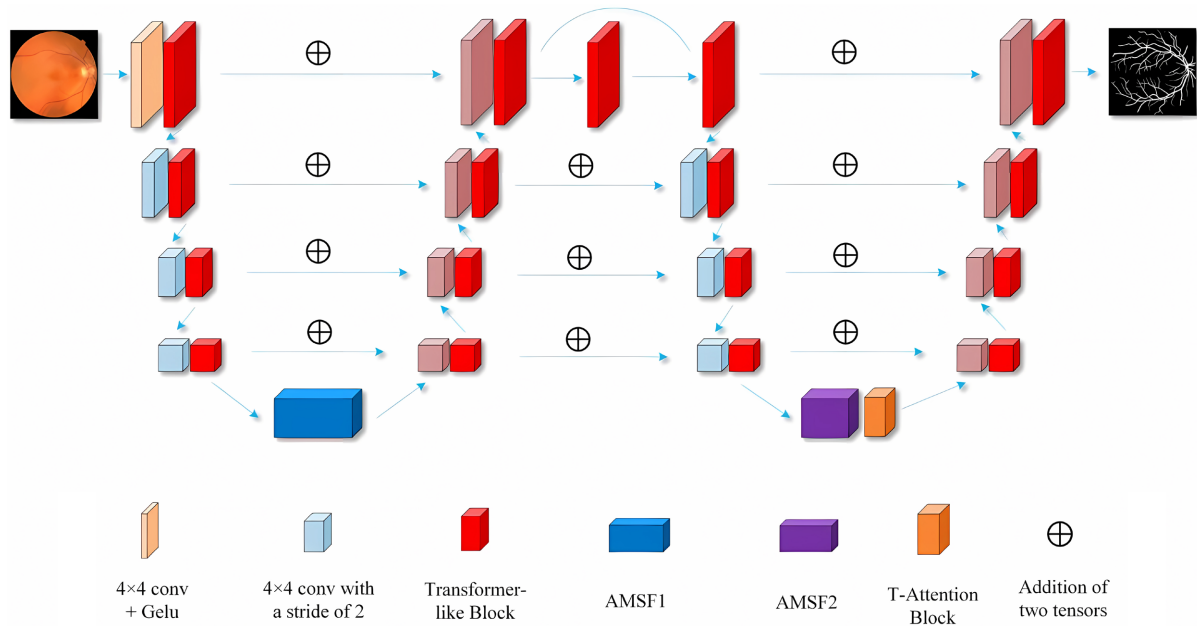


Figure 1. TLadder-AMSF structural diagram
图 1. TLadder-AMSF 结构图

3.1. 大核卷积残差块 LKCRB

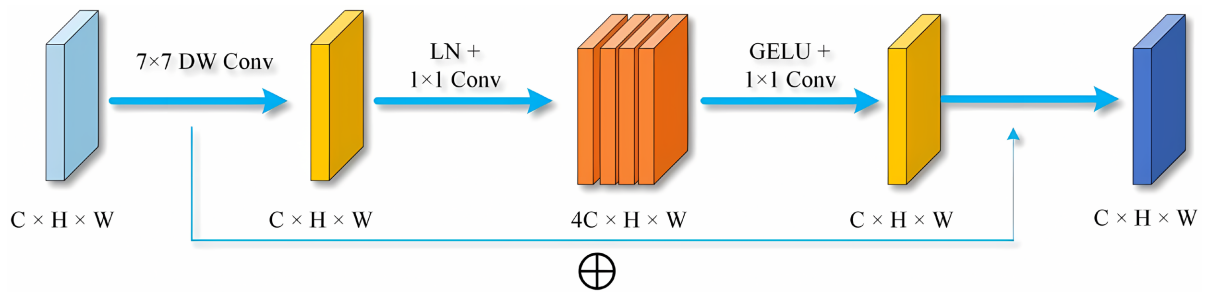


Figure 2. Large kernel convolution residual block structure diagram
图 2. 大核卷积残差块结构图

为了在保持卷积神经网络高效计算优势的同时,引入类似 Transformer 的长程建模能力,本文将网络的基础提取单元设计为大核卷积残差块 LKCRB。该模块通过微观结构的优化,显著提升了模型对血管拓扑连通性的感知。传统的卷积神经网络如 ResNet 及其变体,通常受限于较小的卷积核(如 3×3),导致其有效感受野不足。在视网膜血管分割中,这种局限性使得模型难以捕捉长距离的血管拓扑特征,极易导致细长血管在视觉连通性上的断裂。虽然 Vision Transformer 具有强大的全局建模能力,但其计算开销巨大。因此,本文引入大核卷积残差块,旨在利用大尺寸深度可分离卷积模拟 Transformer 的长程依赖获取能力,同时保留卷积网络的归纳偏置和计算效率,从微观层面增强对血管连通性的保护。模块结构图如图 2 所示。

对于第 1 层的输入特征 $x_l \in \mathbb{R}^{C \times H \times W}$, 其计算流程如下。

模块起始处采用 7×7 的深度可分离卷积(Depthwise Convolution)。相比于传统的 3×3 卷积核,大尺寸卷积核有效扩大了感受野,使模型能捕捉更长跨度的血管特征,减少了由于局部信息不足导致的血管断裂。该过程如公式(1)所示:

$$x_{dw} = \text{DWConv}_{7 \times 7}(x_l) \quad (1)$$

其后,遵循“窄-宽-窄”的设计原则,首先通过层归一化(LayerNorm)稳定分布,随后利用 1×1 卷积将特征通道扩张 4 倍,经 GELU 激活函数处理后,再由另一个 1×1 卷积压缩回原始维度。该过程如公式(2)所示:

$$x_{up} = \text{GELU}(\text{PWConv}_1(\text{LN}(x_{dw}))) \quad (2)$$

其中 PWConv_1 表示升维的点卷积操作。

最后,在末端采用残差连接将输入与输出相加,确保深层网络训练时的梯度流稳定性,有效防止特征退化。该过程如公式(3)所示:

$$x_{l+1} = x_l + \text{PWConv}_2(x_{up}) \quad (3)$$

通过这种结构,模型在保持卷积归纳偏置的同时,获得了类似于 Transformer 的全局表征能力。

3.2. 对称多尺度融合模块 AMSF

视网膜血管具有极高的尺度动态性,主干血管与毛细血管末梢在图像中的像素表现跨度显著。单一感受野的卷积核难以同时有效捕捉所有尺度的血管。常规的 U 型网络在瓶颈处通常采用对称或单一尺度的结构,这导致模型在第一阶段容易丢失微小血管特征,而在第二阶段又缺乏对全局背景的有效抑制。虽然多尺度特征融合是常见解决方案,但在网络不同深度(或不同处理阶段),对多尺度信息的需求应有侧重。为此,我们设计了非对称多尺度融合模块,采用非对称策略,分别嵌入在两个阶段的瓶颈处。在第一个阶段,利用小扩张率组合“广泛探测”所有可能的血管候选特征,避免早期丢失微小血管信号;在第二个阶段,在更丰富的上下文和初步分割线索基础上,利用大扩张率与全局路径进行“精准聚焦”和融合,以区分真实血管与噪声,并修正细节。这种设计逻辑确保了模型在不同迭代深度下能够自适应地平衡局部细节与全局语义。两个 AMSF 模块具有不同的结构,如图 3 所示。

设输入特征图为 $F_m \in \mathbb{R}^{H \times W \times C}$, 其计算流程如下。

首先,利用 AMSF-1 (即第一阶段底部),该模块采用并行的扩张卷积路径,扩张率(dilation rate)设置为较小的 $r \in \{1, 2, 4\}$ 。这种设计侧重于在相对局部的范围内,以不同感受野捕获血管特征,确保基础特征中包含了从细小到中等粗细血管的响应。该过程如公式(4)所示:

$$F_{\text{branch}}^r = \text{Conv}_{3 \times 3, \text{dilation}=r}(F_m), r \in \{1, 2, 4\} \quad (4)$$

随后, 各分支特征通过拼接(Concatenation)和 1×1 卷积进行融合。该过程如公式(5)所示。

$$F_{\text{fusion1}} = \text{Conv}_{1 \times 1} \left(\text{Concat} \left(F_{\text{branch}}^1, F_{\text{branch}}^2, F_{\text{branch}}^4 \right) \right) \quad (5)$$

其次, 利用 AMSF-2 (即第二阶段底部)融合更广阔的上下文和全局信息。它包含三个并行分支, 使用更大的扩张率 $r \in \{6, 12\}$, 以覆盖图像中更广阔的区域, 建模血管的远程空间关系。通过全局平均池化(GAP)和 1×1 卷积获取图像级的语义上下文信息 F_{gap} 。通过等映射分支, 保留原始特征 F_{in} , 确保信息流的稳定性。各分支特征经适当调整后相加融合, 该过程如公式(6)所示。

$$F_{\text{fusion2}} = \text{Conv}_{1 \times 1} (F_{\text{in}}) + \sum_{r \in \{6, 12\}} \text{Conv}_{3 \times 3, \text{dilation}=r} (F_{\text{in}}) + \text{Upsample} \left(\text{Conv}_{1 \times 1} (F_{\text{gap}}) \right) \quad (6)$$

关键的是, 在 AMSF-2 中, 融合后的特征 F_{fusion2} 将被送入一个坐标注意力模块(CA)进行进一步的空间感知增强, 即 $F_{\text{out}} = \text{CA}(F_{\text{fusion2}})$ 。这种非对称设计——AMSF-1 “广撒网”、AMSF-2 “精加工”——构成了网络性能跃升的核心枢纽。

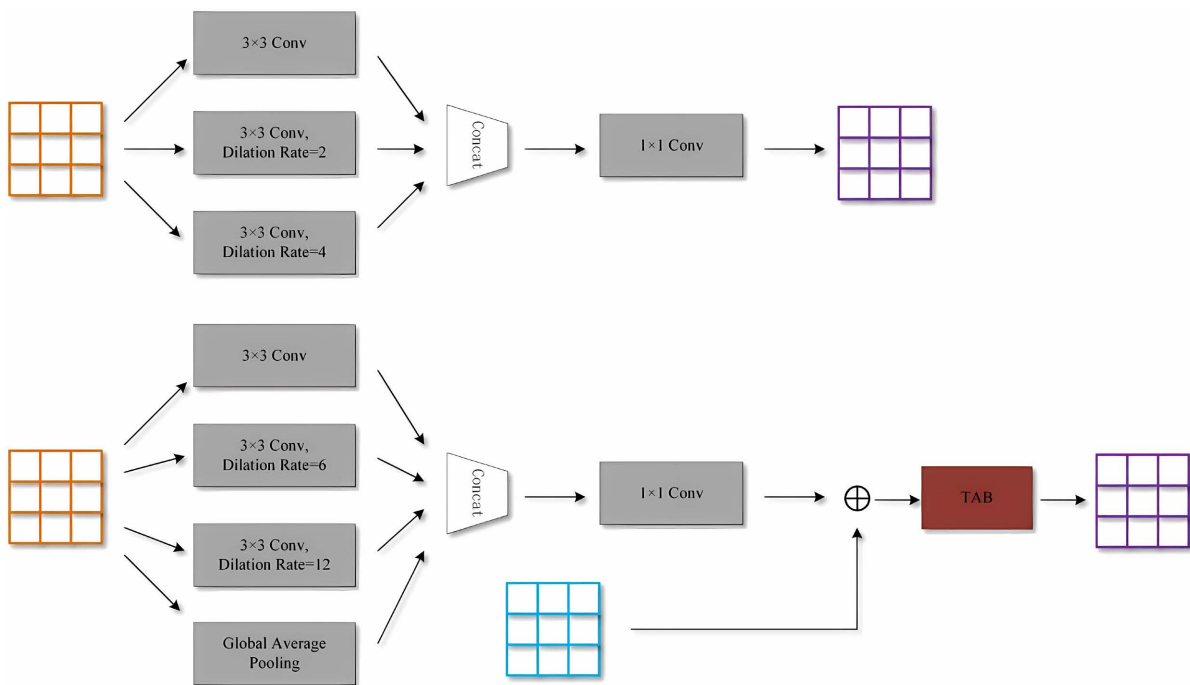


Figure 3. Asymmetric multi-scale fusion module structure diagram
图 3. 非对称多尺度融合模块结构图

3.3. 坐标注意力 CA

在复杂的眼底图像中, 血管与背景(如视盘、病变区域)的纹理有时可能相似, 仅依赖通道注意力或局部上下文容易导致误判。虽然通道注意力机制(如 SE 模块)能增强重要通道的响应, 但它通过全局平均池化压缩了空间信息, 导致模型在分割细小血管时无法感知“血管在哪里”。此外, 血管的本质是具有强烈方向性和空间位置规律的管状结构。例如, 主血管通常从视盘放射状发出。因此, 我们引入坐标注意力机制 CA [13], 其通过沿水平和垂直两个坐标轴进行空间编码, 将精确的空间位置信息(坐标信息)编码到通道注意力中, 使模型不仅知道“关注哪些特征通道”, 还能知道“在空间的哪个位置加强或抑制”, 从而更精准地定位走向各异的血管, 抑制形态不规则的非血管噪声。该模块结构如图 4 所示。

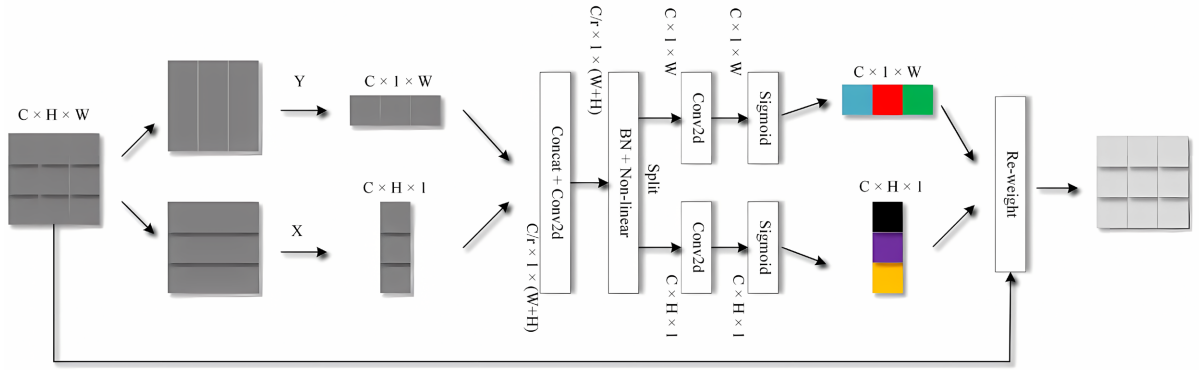


Figure 4. Coordinate attention structure diagram
图 4. 坐标注意力结构图

给定 AMSF-2 输出的特征 $X \in \mathbb{R}^{H \times W \times C}$ ，TCA 模块通过以下流程生成空间感知的注意力权重。

首先，进行方向感知的特征编码，即分别沿水平(X)和垂直(Y)方向进行一维全局平均池化，生成一对方向感知的特征向量，其计算流程如下：

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} x_c(h, i) \quad (7)$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w) \quad (8)$$

其中， $z^h \in \mathbb{R}^{C \times H}$ ， $z^w \in \mathbb{R}^{C \times W}$ 。

其次，进行坐标信息生成与融合：将两个方向的特征拼接后，通过一个共享的 1×1 卷积 f 和激活函数(如 Sigmoid)进行非线性变换，生成中间特征图，再分解为两个独立的方向注意力权重张量，其计算流程如下：

$$f = \delta(\text{Conv}_{1 \times 1}(\text{Concat}(z^h, z^w))) \quad (9)$$

$$g^h = \sigma(\text{Conv}_{1 \times 1}^h(f)), g^w = \sigma(\text{Conv}_{1 \times 1}^w(f)) \quad (10)$$

其中， $g^h \in \mathbb{R}^{C \times H \times 1}$ ， $g^w \in \mathbb{R}^{C \times 1 \times W}$ ， δ 和 σ 分别为非线性激活函数。

最后，进行特征重校准，将生成的水平与垂直注意力权重与原始输入特征逐元素相乘，实现对特征图的空间位置自适应校准，其计算流程如下：

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (11)$$

输出 $Y \in \mathbb{R}^{H \times W \times C}$ 即为经过空间坐标信息增强的特征。

通过这一机制，CA 模块使网络能够捕捉“在特定行和列上，哪些通道的特征是重要的”，从而显著增强了对血管走向的敏感性，在血管交叉、边界模糊及视盘附近等挑战区域表现出更精确的分割能力。

4. 实验结果

4.1. 对比试验

为了验证本文方法的优越性，我们将本文提出的完整模型与近年来在视网膜血管分割领域具有代表性的先进方法[11] [14]-[18]在 DRIVE 和 STARE 数据集上进行了全面对比。

由表 1 可知, 在 STARE 数据集上, 本文提出的方法在多项关键评价指标上均表现优异。本文方法的 Se 达到了 0.8264, 较原始 LadderNet 提升了约 4.48%。这主要归功于大核卷积残差块 LKCRB 的大感受野特征提取以及 AMSF 模块对多尺度血管特征的充分捕获, 使得模型能够有效识别对比度极低的微细血管末梢。此外, 在准确率(Acc)和 Auc 上, 本文方法分别达到了 0.9625 和 0.9848, 均位居所有对比算法之首。尽管灵敏度大幅提升, 本文方法的特异性(Sp)仍保持在 0.9812 的高位, 证明了坐标注意力(CA)的引入能够有效利用空间坐标信息校准特征权重, 在增强血管响应的同时, 精准抑制了视盘及病变区域的类血管伪影。

Table 1. Comparative experimental results on the DRIVE dataset

表 1. 在 DRIVE 数据集上的对比实验结果

方法	Se	Sp	Acc	Auc
U-Net	0.7991	0.9792	0.9562	0.9792
AAU-Net	0.8034	0.9708	0.9495	0.9688
LadderNet	0.7816	0.9789	0.9538	0.9766
IterMiUNet	0.7854	0.9784	0.9538	0.9755
DenseUNet	0.7564	0.9845	0.9555	0.9775
AttUNet	0.8037	0.9778	0.9556	0.9792
本文方法	0.8264	0.9812	0.9625	0.9848

如表 2 所示, 在更具挑战性的 STARE 数据集上, 本文方法同样展现出了卓越的分割性能。本文方法的 Se 指标达到了 0.8356, 显著优于 AttUNet (0.8123)和原始 LadderNet (0.7958)。这表明通过集成大尺寸卷积核的 LKCRB 与 AMSF 结构, 模型能够更有效地应对 STARE 数据集中细小血管浓密且对比度低的问题。STARE 数据集中包含较多的病理噪声, 本文方法通过坐标注意力(CA)对水平与垂直空间信息进行编码, 使 Sp 指标保持在 0.9805 的高水平, 能够精准区分血管与类血管病变。Acc (0.9662)与 Auc (0.9885)均为全表最高值, 证明了本文方法在处理复杂病理背景下的眼底图像时, 具有极高的准确性与泛化能力。

Table 2. Comparative experimental results on the STARE dataset

表 2. 在 STARE 数据集上的对比实验结果

方法	Se	Sp	Acc	Auc
U-Net	0.7845	0.9754	0.9542	0.9782
AAU-Net	0.7921	0.9712	0.9510	0.9782
LadderNet	0.7958	0.9734	0.9545	0.9791
IterMiUNet	0.8012	0.9721	0.9552	0.9802
DenseUNet	0.7756	0.9821	0.9584	0.9815
AttUNet	0.8123	0.9765	0.9591	0.9824
本文方法	0.8356	0.9805	0.9662	0.9885

为了进一步直观验证本文方法的分割性能，图5展示了本文方法与U-Net、LadderNet的可视化对比结果及局部放大效果，我们选取了具有不同血管形态与背景复杂度的代表性样本进行展示。观察第一组和第二组局部放大图可以发现，由于视网膜末梢血管与其背景的对比度极低，U-Net和LadderNet在处理纤细分支时均出现了明显的特征丢失与拓扑断裂现象。而本文方法得益于大核卷积残差块LKCRB提供的大尺寸感受野以及AMSF模块的多尺度特征融合，能够精准捕捉到对比度极低的微细血管，并保持了血管树结构的逻辑连通性，分割结果最接近专家标注(标签)。在血管密集分布及靠近视盘的区域，传统算法极易受到背景亮度不均或组织干扰，导致分割边缘模糊。本文方法生成的分割图边缘更加锐利且平滑，这证明了坐标注意力(CA)通过水平与垂直方向的空间编码，实现了对血管位置的精准校准，从而有效抑制了非血管区域的误判噪声。通过对比观察，本文方法在处理血管交叉与分支点处展现出更强的形态学还原能力，避免了LadderNet中常见的像素粘连问题。这种高保真度的分割效果，为后续临床上进行血管直径测量和扭曲度分析提供了可靠的数据支撑。整体而言，可视化结果与前述量化指标相互印证，表明本文方法在提升血管检出敏感性的同时，亦保持了较好的结构一致性与细节还原能力。

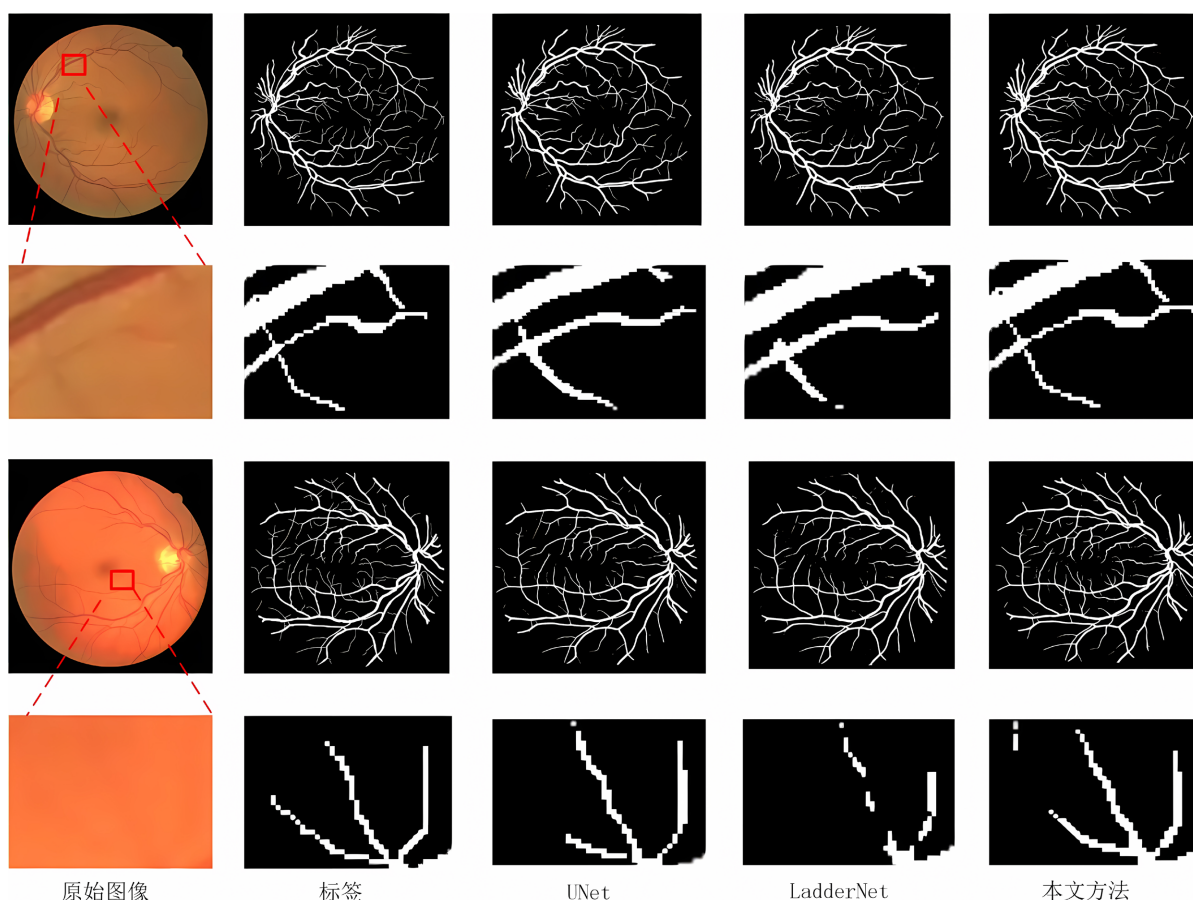


Figure 5. Segmentation result visualization
图5. 分割结果可视化图

4.2. 消融实验

为了验证本文提出的各个模块对视网膜血管分割性能的具体贡献，我们在DRIVE数据集上进行了消融实验，结果如表3所示。

Table 3. Ablation experiment results on the DRIVE dataset
表 3. 在 DRIVE 数据集上的消融实验结果

方法设计				血管分割			
Baseline	LKCRB	AMSF	CA	Se	Sp	Acc	Auc
✓				0.7816	0.9789	0.9538	0.9766
✓	✓			0.8012	0.9794	0.9572	0.9798
✓	✓	✓		0.8155	0.9798	0.9598	0.9821
✓	✓	✓	✓	0.8264	0.9812	0.9625	0.9848

在 Baseline 的基础上替换为大核卷积残差块 LKCRB 后, 灵敏度(Se)从 0.7816 显著提升至 0.8012。传统的残差块结构通常受限于 3×3 卷积核的局部感受野, 在处理细长且扭曲的血管路径时, 难以捕捉跨区域的全局拓扑依赖, 容易导致血管在视觉特征微弱处发生断裂。而大核卷积残差块 LKCRB 凭借其 7×7 深度可分离卷积架构, 模拟了类似 Transformer 的长程建模能力, 在不显著增加计算开销的前提下大幅扩展了有效感受野。这一机制强化了模型对血管连续性的感知, 显著缓解了微小血管的断裂问题。

进一步叠加 AMSF 模块后, Se 指标攀升至 0.8155, AUC 增长至 0.9821。常规的 U 型架构在瓶颈处往往采用单一尺度的特征映射, 这使得模型在面对视网膜血管极端的尺度动态分布(从粗大的中心主干到极其微弱的毛细末梢)时, 难以平衡不同粒度的特征表征。AMSF 模块通过并行的非对称扩张卷积支路, 实现了对多尺度空间信息的动态采集。这种“由粗及精”的特征融合策略, 有效填补了单一感受野在特征提取时的盲区, 使得模型对细小血管分支的查全率得到了进一步提升。

在最终引入坐标注意力后, 特异性(Sp)提升至 0.9812, 准确率(Acc)达到 0.9625 的峰值。此前的方法(如 SE 注意力)往往通过全局平均池化将空间信息压缩为单一的通道向量, 这导致模型失去了对血管像素精确坐标的感知能力, 在面对视盘边缘、病变渗出等高对比度类血管伪影时, 容易产生误判。TAB 通过将空间坐标信息分别沿水平与垂直方向进行编码, 实现了精准的空间位置权重分配。该机制不仅增强了血管区域的特征响应, 更赋予了模型强大的背景抑制能力, 从而在保持高灵敏度的同时, 显著降低了误诊率, 实现了分割性能的全面优化。

5. 结论

本文提出了一种名为 TLadder-UMSF 的高精度视网膜血管分割方法。为了全面提取眼底图像中复杂的血管形态特征, 我们构建了基于双阶段迭代细化的阶梯式网络架构。通过引入集成了大感受野特征提取能力的大核卷积残差块, 有效增强了模型对微细血管分支的捕捉及拓扑连通性的保持。同时, 本文设计了非对称多尺度融合(AMSF)策略与坐标注意力机制, 旨在实现对不同尺度血管特征的深度挖掘与精准空间定位, 为血管分割提供了更具辨识度的多维视角。经过实验验证, 结果表明本文方法在 DRIVE 和 STARE 数据集上均取得了优于经典算法及近年来主流算法的效果, 具有更强的鲁棒性与可解释性。未来的工作将偏向于对异质性临床数据集的进一步补充, 增强模型的泛化能力, 并在模型轻量化方面进一步探索, 尝试扩展到更广泛的医学图像分割任务。

参考文献

- [1] Liew, G., Wang, J.J., Mitchell, P. and Wong, T.Y. (2008) Retinal Vascular Imaging: A New Tool in Microvascular Disease Research. *Circulation: Cardiovascular Imaging*, 1, 156-161. <https://doi.org/10.1161/circimaging.108.784876>

- [2] Cheung, C.Y., Xu, D., Cheng, C., Sabanayagam, C., Tham, Y., Yu, M., *et al.* (2021) A Deep-Learning System for the Assessment of Cardiovascular Disease Risk via the Measurement of Retinal-Vessel Calibre. *Nature Biomedical Engineering*, **5**, 498-508. <https://doi.org/10.1038/s41551-020-00626-4>
- [3] Hughes, A.D., Falaschetti, E., Witt, N., Wijetunge, S., Thom, S.A.M., Tillin, T., *et al.* (2016) Association of Retinopathy and Retinal Microvascular Abnormalities with Stroke and Cerebrovascular Disease. *Stroke*, **47**, 2862-2864. <https://doi.org/10.1161/strokeaha.116.014998>
- [4] Zhang, K., Liu, X., Xu, J., Yuan, J., Cai, W., Chen, T., *et al.* (2021) Deep-Learning Models for the Detection and Incidence Prediction of Chronic Kidney Disease and Type 2 Diabetes from Retinal Fundus Images. *Nature Biomedical Engineering*, **5**, 533-545. <https://doi.org/10.1038/s41551-021-00745-6>
- [5] De Fauw, J., Ledsam, J.R., Romera-Paredes, B., Nikolov, S., Tomasev, N., Blackwell, S., *et al.* (2018) Clinically Applicable Deep Learning for Diagnosis and Referral in Retinal Disease. *Nature Medicine*, **24**, 1342-1350. <https://doi.org/10.1038/s41591-018-0107-6>
- [6] Wang, Y., Ji, G., Lin, P. and Trucco, E. (2013) Retinal Vessel Segmentation Using Multiwavelet Kernels and Multiscale Hierarchical Decomposition. *Pattern Recognition*, **46**, 2117-2133. <https://doi.org/10.1016/j.patcog.2012.12.014>
- [7] Yin, Y., Adel, M. and Bourennane, S. (2012) Retinal Vessel Segmentation Using a Probabilistic Tracking Method. *Pattern Recognition*, **45**, 1235-1244. <https://doi.org/10.1016/j.patcog.2011.09.019>
- [8] Câmara Neto, L., Ramalho, G.L.B., Rocha Neto, J.F.S., Veras, R.M.S. and Medeiros, F.N.S. (2017) An Unsupervised Coarse-to-Fine Algorithm for Blood Vessel Segmentation in Fundus Images. *Expert Systems with Applications*, **78**, 182-192. <https://doi.org/10.1016/j.eswa.2017.02.015>
- [9] Lam, B.S.Y., Yongsheng Gao, and Liew, A.W. (2010) General Retinal Vessel Segmentation Using Regularization-Based Multiconcavity Modeling. *IEEE Transactions on Medical Imaging*, **29**, 1369-1381. <https://doi.org/10.1109/tmi.2010.2043259>
- [10] Wu, C., Zou, Y. and Yang, Z. (2019) U-GAN: Generative Adversarial Networks with U-Net for Retinal Vessel Segmentation. 2019 14th International Conference on Computer Science & Education (ICCSE), Toronto, 19-21 August 2019, 642-646. <https://doi.org/10.1109/iccse.2019.8845397>
- [11] Zhuang, J. (2018) LadderNet: Multi-Path Networks Based on U-Net for Medical Image Segmentation. arXiv:1810.07810.
- [12] Wang, D., Haytham, A., Pottenburgh, J., Saeedi, O. and Tao, Y. (2020) Hard Attention Net for Automatic Retinal Vessel Segmentation. *IEEE Journal of Biomedical and Health Informatics*, **24**, 3384-3396. <https://doi.org/10.1109/jbhi.2020.3002985>
- [13] Hou, Q., Zhou, D. and Feng, J. (2021) Coordinate Attention for Efficient Mobile Network Design. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 13713-13722. <https://doi.org/10.1109/cvpr46437.2021.01350>
- [14] Ronneberger, O., Fischer, P. and Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W. and Frangi, A., Eds., *Lecture Notes in Computer Science*, Springer International Publishing, 234-241. https://doi.org/10.1007/978-3-319-24574-4_28
- [15] Chen, G., Li, L., Dai, Y., Zhang, J. and Yap, M.H. (2023) AAU-Net: An Adaptive Attention U-Net for Breast Lesions Segmentation in Ultrasound Images. *IEEE Transactions on Medical Imaging*, **42**, 1289-1300. <https://doi.org/10.1109/tmi.2022.3226268>
- [16] Li, X., Chen, H., Qi, X., Dou, Q., Fu, C. and Heng, P. (2018) H-Denseunet: Hybrid Densely Connected UNet for Liver and Tumor Segmentation from CT Volumes. *IEEE Transactions on Medical Imaging*, **37**, 2663-2674. <https://doi.org/10.1109/tmi.2018.2845918>
- [17] Wang, S., Li, L. and Zhuang, X. (2022) Attu-Net: Attention U-Net for Brain Tumor Segmentation. In: Crimi, A. and Bakas, S. Eds., *Lecture Notes in Computer Science*, Springer International Publishing, 302-311. https://doi.org/10.1007/978-3-031-09002-8_27
- [18] Kumar, A., Agrawal, R.K. and Joseph, L. (2023) Itermiunet: A Lightweight Architecture for Automatic Blood Vessel Segmentation. *Multimedia Tools and Applications*, **82**, 43207-43231. <https://doi.org/10.1007/s11042-023-15433-7>