# 融合结构化卷积和双重注意力机制的轻量级眼底图像分割网络

汪华登[1,2)], 刘金[2)], 黎兵兵[3)], 潘细朋[1,2,4)], 刘振丙[1,2)], 蓝如师[1,2)*], 罗笑南[1,2)]

[1)] (广西图像图形与智能处理重点实验室 桂林 541004)
[2)] (桂林电子科技大学计算机与信息安全学院 桂林 541004)
[3)] (广东省人民医院赣州医院(赣州市立医院)病理科 赣州 341000)
[4)] (广东省医学科学院广东省人民医院放射科 广州 510080)
(rslan@guet.edu.cn)

**摘 要:** 眼底血管图像的自动分割对于多种眼科疾病的计算机辅助诊断具有重要作用. 针对血管的尺度差异和图像噪声导致眼底血管图像分割困难、使用单一尺度卷积运算的深度学习方法获取的特征感受野有限, 以及现有的方法复杂度过高的问题, 提出一个融合结构化卷积和双重注意力机制的轻量级眼底图像分割网络. 通过编码器增强、减少下采样次数和特征深度的编码-解码网络设计, 实现参数量只有 0.63M 的轻量化网络. 在编码阶段, 提出一种结构化卷积方法, 有效地避免了网络训练过拟合, 提高了网络捕获差异化血管特征的能力; 在解码阶段, 采用基于空间和通道的双重注意力机制, 使网络更加关注血管特征的上下文和几何空间信息, 抑制病变等噪声的干扰. 在 DRIVE, CHASE_DB1 和 STARE 数据集上进行实验的结果表明, 所提网络图像分割的准确率分别为 96.92%, 97.57%和 97.51%, 灵敏度分别为 83.68%, 84.99%和 84.87%, 受试者曲线下的面积(AUC)分别为 98.67%, 99.05%和 99.02%; 并通过在 DRIVE 和 STARE 数据集上的交叉训练, 验证了该网络的泛化能力.

**关键词:** 眼底图像分割; 编码-解码网络; 轻量级网络; 结构化卷积; 双重注意力机制
**中图法分类号:** TP391.41 **DOI:** 10.3724/SP.J.1089.2024.19843

# Lightweight Fundus Image Segmentation Network Combining Structured Convolution and Dual Attention Mechanism

Wang Huadeng[1,2)], Liu Jin[2)], Li Bingbing[3)], Pan Xipeng[1,2,4)], Liu Zhenbing[1,2)], Lan Rushi[1,2)*], and LuoXiaonan[1,2)]

[1)] (*Guangxi Key Laboratory of Image and Graphics Intelligent Processing, Guilin* 541004)
[2)] (*School of Computer Science and Information Security, Guilin University of Electronic Technology, Guilin* 541004)
[3)] (*Department of Pathology, Ganzhou Hospital of Guangdong Provincial People's Hospital (Ganzhou Municipal Hospital), Ganzhou* 341000)
[4)] (*Department of Radiology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou* 510080)

**Abstract:** Automatic segmentation of fundus blood vessel images plays an important role in computer-aided diagnosis of various ophthalmic diseases. In order to solve the problems of difficulty in segmenting fundus vascular images caused by vascular scale difference and image noise, limited feature field obtained by deep learning method using single-scale convolution operation, and high complexity of existing methods, a lightweight fundus image segmentation network combining structured convolution and dual attention mechanism is proposed. Through the design of coder-decoder network with encoder enhancement, downsampling

收稿日期: 2022-07-20; 修回日期: 2023-04-07. **基金项目**: 国家自然科学基金(62362014, 62172120, 62002082); 广西杰出青年基金(2019GXNSFFA245014); 广西图像图形与智能处理重点实验室(桂林电子科技大学)开放基金(GIIP2209). **汪华登**(1979—), 男, 博士, 正高级实验师, 硕士生导师, CCF 高级会员, 主要研究方向为医学图像处理、深度学习; **刘金**(1997—), 男, 硕士研究生, 主要研究方向为医学图像分割、深度学习; **黎兵兵**(1993—), 男, 硕士, 主治医师, 主要研究方向为肿瘤病理学与人工智能; **潘细朋**(1985—), 男, 博士, 副教授, 硕士生导师, 主要研究方向为机器学习、医学图像分析; **刘振丙**(1980—), 男, 博士, 教授, 博士生导师, CCF 会员, 主要研究方向为机器学习、图像处理; **蓝如师**(1986—), 男, 博士, 研究员, 博士生导师, CCF 会员, 论文通信作者, 主要研究方向为机器学习、医学信息处理; **罗笑南**(1963—), 男, 博士, 教授, 博士生导师, 主要研究方向为医学图像三维重建、机器学习.

times and feature depth reduction, the lightweight network with only 0.63M parameters is realized. In the coding stage, a structured convolution method is proposed, which effectively avoids overfitting of network training and improves the ability of network to capture differentiated vascular features. In the decoding stage, a dual attention mechanism based on spatial and channel is adopted to make the network pay more attention to the context and geometric spatial information of vascular features, and suppress the interference of lesion noise. Experiments on the DRIVE, CHASE_DB1 and STARE datasets show that the accuracy of the image segmentation is 96.92%, 97.57%, and 97.51%, the sensitivity is 83.68%, 84.99%, and 84.87%, and the area under curve of the receiver operating characteristic is 98.67%, 99.05%, and 99.02%, respectively. Cross-training on the DRIVE and STARE datasets validated the generalization ability of the network.

**Key words:** fundus image segmentation; encoder-decoder network; lightweight network; structured convolution; dual attention mechanism

眼底视网膜血管结构中包含大量眼部相关疾病诊断的信息. 在临床研究中, 如糖尿病视网膜病变[1]、糖尿病黄斑病变和高血压性视网膜病变等疾病, 都会导致视网膜血管形状、曲率发生不同程度的改变. 虽然可以依靠医生目视检查来获取不同形态的血管信息, 但这种方法耗时、费力且主观性强. 研究高效、准确的眼底图像自动分割方法, 形成计算机辅助诊断技术, 对于多种眼部疾病的早期诊断和精准治疗具有重要价值.

眼底图像的结构复杂, 还夹杂着不同程度的出血和渗出物等干扰噪声[2], 导致视网膜血管自动分割困难. 早期的分割方法大致可分为 3 类[3]: (1) 基于血管追踪的方法. 考虑血管结构的连续性, 通过设置初始种子节点, 沿血管方向追踪血管结构完成分割[4]; (2) 基于匹配滤波的方法. 通过计算对图像滤波后的最大响应值找到血管点, 常采用多尺度的高斯滤波器计算血管图像[5]; (3) 基于传统机器学习的方法. 通过构造特征向量训练分类器完成血管分割[6-7]. 虽然这些方法在一定程度上能够实现血管的自动分割, 但由于血管特征的差异性和图像噪声的干扰, 最终的分割性能依然不能令人满意. 近年来, 随着卷积神经网络[8]和深度学习在医学领域的蓬勃发展, 越来越多的学者将它们应用于眼底血管分割任务, 并取得了较好的效果. U-Net[9]是其中一个性能表现良好的分割网络, 因此许多学者基于该网络提出了不同的改进网络, 如 Att-UNet[10], UNet++[11]和 DDNet[12]等. 虽然这些网络改善了 U-Net 的分割性能, 但是过于追求提高分割的准确率, 忽略了网络的复杂度、参数量以及在现实应用中的算力支持, 例如, Att-UNet, UNet++和 DDNet 的网络参数量分别达到 8.21M, 9.04M 和 56.03M. 因此, 设计一个轻量且高效的眼底血管自动分割网络, 具有很大的研究意义和应用价值.

针对眼底血管分割存在的需求和挑战, 本文提出一个融合结构化卷积和双重注意力机制的轻量级网络. (1) 提出一种结构化卷积方法. 使用基于空洞卷积的结构化多尺度卷积模块(structured multi-scale convolutional module, SMCM)扩大网络的感受野, 增强网络对不同尺度的差异化血管的特征提取能力; 使用结构化 DropBlock[13]卷积模块(structured dropblock convolutional module, SDCM)替换常规的 3×3 卷积层, 防止网络过拟合. (2) 设计基于空间和通道的双重注意力机制. 将空间注意力模块(spatial attention module, SAM)嵌入编码器底层, 强调特征空间位置上的联系, 更加关注特征的上下文信息, 突出重要的血管特征; 将通道注意力模块(efficient channel attention module, ECAM)嵌入原始的跳跃连接中, 通过跨通道信息交互自适应地细化血管特征, SAM 和 ECAM 形成一种双重注意力机制, 抑制无关的背景噪声, 提高网络精准分割的能力. (3) 设计轻量级网络结构. 减少编码-解码网络下采样次数和特征深度, 并使用轻量化的设计增强编码器的结构, 改进 U 型编码-解码网络. 实验结果表明, 在保证网络轻量级的情况下, 本文网络实现了比主流眼底血管分割网络更好的分割性能.

## 1　相关工作

近年来, 随着计算机硬件性能的提高, 深度学习算法逐渐成为血管分割任务中的主流方法. Yan 等[14-15]改进了损失函数, 提出三阶段的深度学习

网络，分别实现粗细血管的分割以及血管的融合；但多阶段的设计大大增加了网络的复杂度. Deng 等[16]提出一个基于多尺度注意力的分割网络，结合具有残差机制的可变形卷积网络和改进的脉冲耦合神经网络提高血管分割的准确性和连通性；但是，多模块和多网络的结合增加了网络的复杂度和计算开销.

基于编码-解码网络的深度学习算法已被证明在视网膜血管分割等医学分割方面具有良好的效果[9]，常用的分割网络 U-Net 就是典型的编码-解码网络. 然而，直接将 U-Net 应用在血管分割任务中还存在着一些缺陷：(1) 网络结构的深度增加导致编码器的特征积累达到极限，很难进一步提取有效特征；(2) 可获得的眼底图像数据量较少，容易使 U-Net 训练过拟合；(3) 单一尺度的卷积操作的感受野是固定的，很难捕捉到尺度变化巨大的血管特征信息，尤其是一些微细血管；(4) 虽然 U-Net 解码阶段的跳跃连接可以将局部信息进行特征融合，但这会引入太多无关的背景噪声，使网络难以区分真实的血管与周围的病变噪声. 针对这些问题，研究人员提出许多基于 U-Net 的改进算法. Schlemper 等[10]提出一种基于 U-Net 的注意力门控结构——Att-UNet，能够减少图像中噪声带来的负面影响，但对于微细血管的分割能力较差；Zhou 等[11]提出基于 U-Net 的改进网络 UNet++，其内嵌了不同深度的 U-Net，并将长连接改成密集的短连接，使用剪枝策略加快推理速度，虽然相比 U-Net 实现了更高的分割性能，但是对于低对比度的图像效果仍旧不好，并且嵌套 U-Net 的方式导致网络参数量较大；受到 DropBlock 的启发，Guo 等[17]基于 U-Net 设计了 SD-UNet，有效地缓解了因眼底图像数据量小造成的训练过拟合现象，但网络分割准确率和灵敏度较差. 除了基于 U-Net 改进的算法之外，使用 ResNet[18]和 DenseNet[19]等主体网络结构对编码器进行构造，也衍生出了许多基于编码-解码网络的分割网络. Gu 等[20]提出 CE-Net，通过密集空洞卷积块和残差多核池化块增强编码器结构；Wang 等[21]提出一个上下文空间网络用于血管分割，设计双通道编码器进行特征提取，提升了网络对于分割血管和检测血管边界的能力，但双编码的结构增加了网络的复杂度；Yang 等[22]基于编码-解码网络提出综合上下文网络 CC-Net 用于血管分割，一方面弥补了池化操作带来的空间信息损失，另一方面更加关注特征的通道信息，达到了不错的分割性能，但对于微细血管的分割还存在不足. 除了对编码器的改进，文献[23]研究也表明，解码器中的上采样操作和反卷积操作有利于某些分割任务.

虽然以上视网膜血管分割网络在一定程度上改善了分割性能，但也存在着诸多不足：(1) 使用过多的辅助输出和密集、残差网络连接[12,14-15,20-21]来追求网络性能的提高，大大增加了网络的复杂度和计算量，忽略了网络的轻便性以及临床应用中的算力支持；(2) 在网络输出后增加后处理步骤[12,16]破坏了网络端到端的特性；(3) 使用固定的特征感受野提取血管信息[9-11,22]不利于差异化的血管分割.

## 2 本文网络

### 2.1 结构搭建

本文提出一个端到端的轻量级神经网络，输入数据(或信息)是彩色视网膜血管图像，输出是视网膜血管分割图像. 如图 1 所示，本文网络是具有编码器(左侧)-解码器(右侧)的 U 型网络. 在编码器阶段，原始数据首先经过一个 SMCM、一个 SDCM 和一个最大池化(max pooling, MP)层完成第 1 阶段编码，特征图的通道数由 3 变为 16；第 2~3 编码阶段保持与第 1 阶段相同的操作，特征图通道数从 32 变为 64；第 4 阶段嵌入 SAM 以突出血管特征，输出的特征图尺寸为 $(H/8\times W/8)\times 128$. 编码器和解码器的结构是对称的，在解码器阶段，首先通过转置卷积将编码阶段得到的特征图进行放大，将结果与经过注意力处理后的相应编码阶段的特征图进行通道维度的拼接，得到的特征通道数变为 $128(64+64)$；第 2~4 解码阶段的操作保持一致，每个阶段包括 2 个 SDCM 和一个转置卷积操作，输出特征图尺寸为 $H\times W\times 16$；在每个解码阶段的跳跃连接中嵌入 ECAM，将长连接变为短连接，调整通道响应，抑制干扰噪声，细化血管特征. 在网络的最后一层，使用 1×1 卷积和 Sigmoid 激活函数输出血管分割图.

### 2.2 编码-解码网络中的结构化卷积方法

在视网膜血管图像中，不仅血管的尺度差异较大，而且一些病变噪声的颜色、亮度与视网膜血管相近，使用固定感受野的常规卷积运算难以充分提取血管特征，容易出现错误分割. 尽管本文使用了一些数据增强手段来扩充眼底数据集，但训练样本的数量仍然较少，导致本文网络容易出现

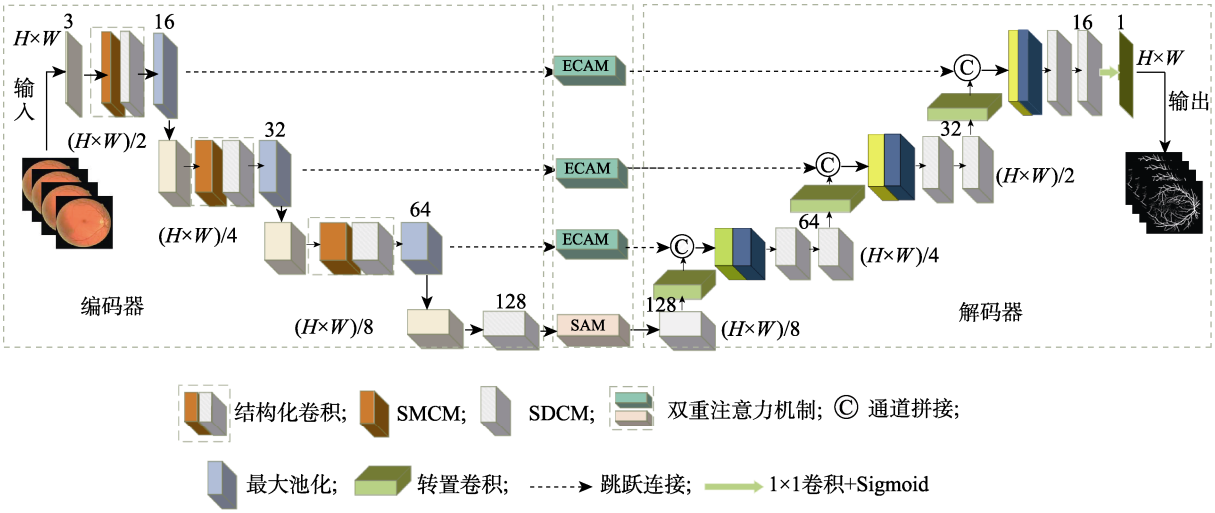



图 1　本文网络结构

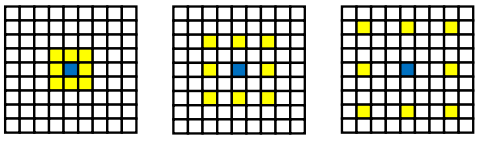
过拟合的现象. 为了应对这些问题，本文在编码-解码网络中设计一种结构化卷积方法，有效地解决网络训练过拟合问题，提高了网络捕获差异化血管特征的能力. 结构化卷积方法由 SMCM 和 SDCM 这 2 部分组成.

(1) SMCM

与传统的卷积运算相比，空洞卷积能够在不增加参数量的前提下扩大感受野，并且不会降低特征图的分辨率[24]. 对于大小为 $n\times n$ 的卷积核使用空洞卷积之后，卷积核的大小被扩大为 $N_r\times N_r$，公式为

$$N_r = n+(n-1)\times(r-1) \tag{1}$$

其中，$r$ 表示空洞率，当 $r=1$ 时，空洞卷积简化为标准卷积. 如图 2 所示，其中，黄色像素格是计算蓝色像素格所涉及的像素位置，以 3×3 的标准卷积核为例，$r=1$ 的卷积核大小 $N_r$ 依然是 3，$r=2$ 的卷积核大小 $N_r=5$，$r=3$ 的卷积核大小 $N_r=7$. 所以，拥有大的空洞率的卷积核会获得大的感受野.



a. $r=1$　　b. $r=2$　　c. $r=3$

图 2　使用不同空洞率的空洞卷积

本文使用空洞卷积代替常规的卷积操作，以便更好地提取血管特征. 考虑视网膜血管结构的差异性，单一尺度的卷积核无法充分地兼顾不同尺度的血管特征，本文提出 SMCM 捕获尺度差异化的血管特征，实现不同尺度的特征信息融合，提高网络对于复杂血管结构的分割能力. 如图 3a 所示，在 SMCM 中，使用 4 个具有不同空洞率的空洞卷积层和 1 个标准 1×1 卷积层捕获血管特征图上的特征信息. 4 个并行的空洞卷积层的空洞率分别为 1，2，4 和 8. 其中，大的空洞率提取的特征感受野比较大，语义信息表征能力较强，有助于较大血管的识别；小的空洞率提取的特征感受野较小，获取血管几何细节信息的能力较强，有利于微细分叉血管的捕获. SMCM 中，一方面，组合 4 个不同空洞率的空洞卷积，以充分地提取不同尺度的血管特征信息；另一方面，使用标准的 1×1 卷积层保留当前尺度的信息特征；最后，将所有分支所得到的特征映射进行通道维度的拼接，得到大小为 $H\times W\times 5C$ 的多尺度特征信息，再使用 1×1 卷积运算融合得到的多尺度特征信息，以获得最终的输出特征图 $\boldsymbol{O}^l$. 整体过程可以描述为

$$\boldsymbol{O}^l(\boldsymbol{x}_i)=\begin{cases} W_{1\times1}(\boldsymbol{X}), & i=1 \\ D_r(\boldsymbol{X}), & (i,r)=(2,1),(3,2),(4,4),(5,5) \end{cases} \tag{2}$$

$$\boldsymbol{O}^l = W_{1\times1}\left(\tau\left(\boldsymbol{O}^l(\boldsymbol{x}_1),\boldsymbol{O}^l(\boldsymbol{x}_2),\cdots,\boldsymbol{O}^l(\boldsymbol{x}_i)\right)\right),\quad i=1,2,\cdots,5 \tag{3}$$

其中，$\boldsymbol{O}^l(\boldsymbol{x}_i)$ 表示在编码器的第 $l$ ($l=1,2,3$)阶段 SMCM 中第 $i$ 支路上的输出；$W_{1\times1}(\cdot)$ 表示 1×1 的卷积运算；$D_r(\cdot)$ 表示 3×3 卷积核空洞率为 $r$ 的空洞卷积运算；$\tau(\cdot)$ 表示通道拼接操作.

(2) SDCM

在深度学习中经常会遇到网络过拟合的问题. 由于本文网络是一个全卷积网络，因此卷积输出特征图的各个元素之间有很强的相关性，即使使用 Dropout 机制[25]随机丢弃一些相关单元，其邻域的相似特征信息仍然会传递到网络的下一层，

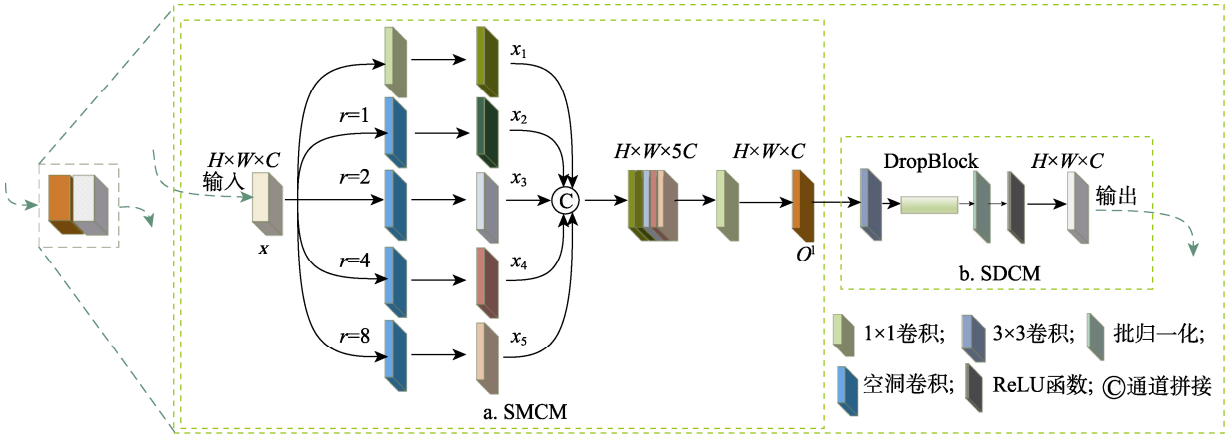



图 3 结构化卷积方法示意图

从而导致 Dropout 无效. 与 Dropout 不同, DropBlock 随机丢弃的是一块连续的区域, 并非独立的单元. DropBlock 中有 $\alpha$ 和 $\gamma$ 这 2 个主要的参数, 其中, $\alpha$ 表示要丢弃的连续区域块大小, 当 $\alpha$=1 时, DropBlock 等同于普通的 Dropout, 本文的实验中 $\alpha$=7; $\gamma$ 表示随机丢弃过程中的概率, 公式为

$$\gamma=\frac{(W\times H)\times(1-P_{\mathrm{R}})}{(W-\alpha+1)\times(H-\alpha+1)\times\alpha^2} \tag{4}$$

其中, $W$ 和 $H$ 分别表示特征图的宽度和高度; $P_{\mathrm{R}}$ 表示神经单元被保留的概率.

文献[17,26-27]的工作表明, 将 DropBlock 集成在编码-解码网络中能够有效地缓解网络过拟合问题. 因此, 本文将 DropBlock 正则化方法集成在 SDCM 中防止网络过拟合, SDCM 的结构如图 3b 所示. 其中, 在编码器的第 1~3 阶段, 将 SDCM 使用在结构化卷积方法中; 在编码器的第 4 阶段和解码器中, 使用 SDCM 替换双层的 3×3 卷积块.

### 2.3 基于空间和通道的双重注意力机制

ECAM 对于实现特征图不同维度信息交互具有良好效果. Wang 等[28]发现, 适当的跨通道信息对于网络学习有效的特征信息非常重要. ECAM 通过权重共享的一维卷积运算实现跨通道的信息交互, 避免通道压缩操作, 实现了较好的网络分割性能. 在编码-解码网络中, 虽然长连接的跳跃连接能够实现浅层和深层语义信息的融合, 但它们之间的语义差别较大, 直接粗糙地融合会加大网络的学习难度. 因此, 本文将 ECAM 嵌入在跳跃连接中, 将长连接变为短连接进行噪声过滤, 细化突出浅层语义信息中的重要的血管特征信息, 减少与深层语义信息的差别, 降低网络的学习难度, 提高网络分割性能. 本文嵌入的 ECAM 只含有 9 个参数, 因此具有绝对的轻量性, 其结构如图 4a 所示. ECAM 设计一种自适应调节卷积核的方法, 卷积核的大小 $k$ 和特征通道数 $C$ 之间存在一种特殊的非线性特征映射 $\kappa(\cdot)$, 给定 $C$, $k$ 可以自适应地确定大小. 整体过程可以描述为

$$\boldsymbol{E}=\boldsymbol{X}'\otimes\sigma\left(W_{1\mathrm{D}}\left(\mathrm{GAP}\left(\boldsymbol{X}'\right)\right)\right) \tag{5}$$

$$k=\kappa(C)=\phi\left(\frac{\mathrm{lb}C}{\theta}+\frac{\lambda}{\theta}\right) \tag{6}$$

其中, $\boldsymbol{X}'$ 表示输入特征图; $W_{1\mathrm{D}}(\cdot)$ 表示一维卷积运算; $\mathrm{GAP}(\cdot)$ 表示全局平均池化(global average

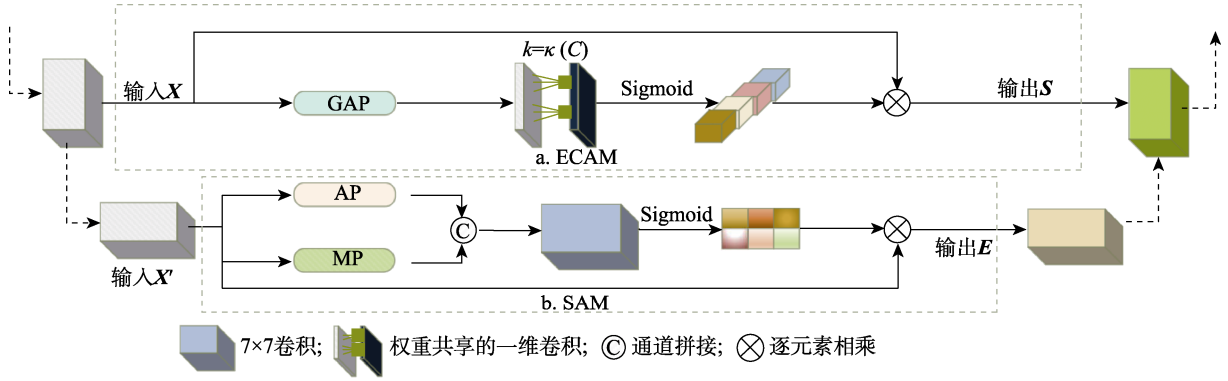



图 4 ECAM 和 SAM 双重注意力结构图

pooling, GAP)操作；$\phi(\cdot)$ 函数表示取相邻的奇数；$\boldsymbol{E}$ 表示通道注意力特征图；$\theta$=2，$\lambda$=1.

SAM 是卷积注意力模块[29]的一部分，广泛应用于分类和检测任务. 在编码器的底部，网络特征积累达到顶峰，将 SAM 嵌入其中，能够充分利用特征像素之间的空间关系来生成空间注意力特征图，使网络更加关注血管的上下文信息，突出重要的血管特征，抑制背景噪声. 此外，SAM 只有 98 个参数，是一种轻量化的结构，能够在不给网络带来计算负担的前提下改善分割性能. SAM 的结构如图 4b 所示，整体过程可以描述为

$$\boldsymbol{S}=\otimes\sigma\left(W_{7\times7}\left(\tau\left(\mathrm{AP}\left(\boldsymbol{X}\right),\mathrm{MP}\left(\boldsymbol{X}\right)\right)\right)\right) \tag{7}$$

其中，$\boldsymbol{X}$ 表示输入特征图；$\otimes$ 表示逐元素相乘；$W_{7\times7}(\cdot)$ 表示 7×7 的卷积运算；$\sigma(\cdot)$ 表示 Sigmoid 激活函数；$\mathrm{AP}(\cdot)$ 和 $\mathrm{MP}(\cdot)$ 分别表示平均池化(average pooling, AP)和最大池化操作；$\tau(\cdot)$ 表示通道拼接操作；$\boldsymbol{S}$ 表示空间注意力特征图.

### 2.4　网络轻量化方法

已有的许多分割网络忽视了网络的复杂度和现实应用中的算力支持. 如何在保证分割性能的前提下降低网络的复杂度，实现网络轻量化，是一个非常重要的问题. 通常，网络的参数量决定网络的复杂度，参数量越少，网络的复杂度越低，网络对于算力的需求越低，也更加容易在实践中部署. 但是，过于追求减少网络的参数量，也会造成网络的性能降低.

编码-解码网络遵循的设计原则是：每次下采样时，特征图的长度和高度减半，但深度会增加一倍. 深度的增加直接导致卷积核数量的增加，网络的参数量会逐渐增加. 因此，本文网络减少了下采样次数和特征深度，使得网络中最深的特征映射只有 128. 本节讨论 SMCM 和双重注意力机制对于网络参数总量的影响. 图 1 中，SMCM 取代了原始编码器中的卷积层，注意力模块被嵌入在编码-解码网络中. 为了在不增加网络计算量的前提下扩大卷积核从而扩大感受野，本文在 SMCM 中使用空洞卷积. 如在 SMCM 的 $x_3$ 分支，其生成的参数为 $C\times3\times3\times C=9C^2$，若使用普通卷积获得相同的感受野，则生成的参数为 $C\times5\times5\times C=25C^2$，其中，$C$ 表示特征图的深度，这样做相当于减少了 64%的参数量. 另外，本文计算 SAM 和 ECAM 的双重注意力模块的参数量只有 107，而这对于整个网络 638 732 的参数量来说是可以忽略不计的.

## 3　实验与结果分析

### 3.1　数据集

使用 DRIVE[30]，CHASE_DB1[31]和 STARE[32]这 3 个公开的彩色眼底图像数据集评估本文提出的轻量级分割网络. 对于 3 个数据集，测试集中的每幅图像都由 2 位专家进行手动标注，整个实验过程中，本文都选择第 1 位专家标注作为实验的金标准，同时与第 2 位专家标注结果进行比较，3 个数据集的信息如表 1 所示.

**表 1　3 个公开数据集的信息**

| 数据集 | 数量 | 图像大小/像素 | 训练-测试 |
|---|---|---|---|
| DRIVE | 40 | 584×565 | 20-20 |
| CHASE_DB1 | 28 | 999×960 | 20-8 |
| STARE | 20 | 700×603 | 4 折交叉验证 |

DRIVE 数据集包含 40 幅彩色眼底图像，其中，33 幅图像来自非糖尿病患者，7 幅图像来自轻度糖尿病患者，每幅图像的分辨率为 584×565 像素. 本文遵循官方的数据集划分方式，使用 20 幅图像进行训练，20 幅图像进行测试.

CHASE_DB1 数据集包含 28 幅彩色眼底图像(拍摄收集于 14 名儿童的左右眼)，每幅图像的分辨率为 999×960 像素. 使用 20 幅图像进行训练，8 幅图像进行测试. 该数据集中的图像背景照明不均，血管较薄，血管对比度较差，小动脉分布广泛.

STARE 数据集包含 20 幅彩色眼底图像，其中，10 幅属于病理性图像，10 幅属于非病理性图像，每幅图像的分辨率为 700×603 像素. 由于该数据集没有官方的数据集划分方法，因此使用 4 折交叉验证方式进行训练(每折包含 15 幅训练图像，5 幅测试图像).

### 3.2　实验细节与参数设置

本文实验的设备环境为 Windows10 操作系统，开发测试软件环境为 PyCharm；远程服务器环境为 Ubuntu 18.04.5 LTS，NVIDIA Tesla P100-16 GB GPU. 使用的深度学习框架为 Tensorflow 和 Keras. 将 3 个数据集分别分为训练集和测试集，使用训练集进行训练. 为了适应本文网络，分别将 DRIVE，CHASE_DB1 和 STARE 数据集的每幅图像用零填充为 592×592 像素，1 008×1 008 像素和 704×704 像素；使用数据增强(随机角度旋转、水平翻转、垂直翻转和对角翻转)扩大训练集，训练集数量从 20，20 和 15 扩充到 200，200 和 150，并且在网络训练阶段随机选择 10%的增强图像作为验证集. 为了

与现实情况尽可能一致, 在网络评估阶段, 本文将分割结果裁剪为数据图像的原始大小. 为了达到最好的网络分割性能, 对于 DRIVE 数据集, batch size= 2, $P_R$ = 0.81; 对于 CHASE_DB1 数据集, batch size= 1, $P_R$ = 0.84; 对于 STARE 数据集, batch size= 3, $P_R$ = 0.85; 所有的数据集都训练 150 次迭代, 使用Adam优化器和交叉熵损失函数优化本文网络. 交叉熵损失函数的公式为

$$L_{CE}(\hat{y},y)=-\sum_i y_i\ln\hat{y}_i+(1-y_i)\ln(1-\hat{y}_i) \tag{8}$$

其中, $\hat{y}$ 和 $y$ 分别表示预测值和真实值; $i$ 表示第 $i$ 个训练样本. 本文采用自适应学习率的策略, 在训练的前 100 次迭代, 学习率保持在$10^{-3}$; 后 50 次迭代, 学习率保持在$10^{-4}$.

### 3.3 评价指标

视网膜血管分割任务可以看作一个二元分割任务. 因此, 分割结果可以使用常见的准确率 $A$, 灵敏度 $S_{sen}$, 特异性 $S_{spe}$ 和 $F_1$ 进行衡量; 此外, 本文还评测了马修斯相关系数 $J_M$ 和科恩卡帕系数 $J_K$ [33]. 这些评价指标的计算公式分别为

$$A=\frac{TP+TN}{TP+TN+FP+FN} \tag{9}$$

$$S_{sen}=\frac{TP}{TP+FN} \tag{10}$$

$$S_{spe}=\frac{TN}{TN+FP} \tag{11}$$

$$F_1=\frac{2TP}{2TP+FP+FN} \tag{12}$$

$$J_M=\frac{TP\times TN-FP\times FN}{\sqrt{(TP+FP)\times(TP+FN)\times(TN+FP)\times(TN+FN)}} \tag{13}$$

$$J_K=\frac{P_O-P_E}{1-P_E} \tag{14}$$

$$P_O=\frac{TP+TN}{TP+TN+FP+FN} \tag{15}$$

$$P_E=\frac{(TP+FP)(TP+FN)+(FN+TN)(FP+TN)}{(TP+FP+FN+TN)^2} \tag{16}$$

其中, TP, TN, FP 和 FN 分别表示分割结果中每个像素所产生的真阳性(true positive, TP), 真阴性(true negative, TN), 假阳性(false positive, FP)和假阴性(false negative, FN); $P_O$ 和 $P_E$ 分别表示预测值和真实值的一致性比例和整体随机一致性; $S_{sen}$ 和 $S_{spe}$ 分别表示网络对于血管和非血管像素的分割性能; $A$ 用于衡量网络的整体性能; $F_1$ 表示网络分割结果和真实标签之间的相似程度; $J_M$ 能够在数据类别不平衡的情况下衡量网络的整体性能; $J_K$ 用于衡量本文网络的分割结果与医学专家注释之间的一致性. 为了与其他方法进行比较, 本文还采用受试者工作特性(receiver operating characteristic, ROC)曲线下的面积(area under curve, AUC)进行评价, AUC 越高, 表示网络分割性能越好, 鲁棒性越强.

### 3.4 超参数敏感性比较

(1) 空洞率 $r$ 的验证

本文在 SMCM 中集成空洞卷积来提高网络特征提取能力, 设置不同的空洞率 $r$ 会对网络性能造成不同的影响. 为此, 本文在 DRIVE 数据集上选取 6 组不同空洞率的 SMCM 对其敏感性进行实验, 并对实验结果进行评估, 结果如表 2 所示. 可以看出, 本文使用 $r$ = (1, 2, 4, 8) 组合的 SMCM 获得的 $A$ 和 $S_{spe}$ 虽然不是最高的, 但是 $S_{sen}$, $F_1$, AUC, $J_M$ 和 $J_K$ 都是最高的, 表明其综合分割性能是最好的, 也说明本文所选空洞率的合理性.

(2) 随机丢弃概率 $P_R$ 的验证

本文在 SDCM 中集成 DropBlock 防止网络训练过拟合, 改善网络的分割性能. $P_R$ 决定 DropBlock 过程中随机丢弃神经元的概率, 其值不同也会影响网络性能. 因此, 本文在 3 个数据集上分别选取 6 个不同 $P_R$ 值的 SDCM 对其进行敏感性实验, 并对实验结果进行评估, 结果如表3所示. 可以看出, 对于 DRIVE 数据集, 当 $P_R$ = 0.81 时, 虽然 $A$ 和 $S_{spe}$ 略低于最高值, 但是其 $S_{sen}$, $F_1$, AUC, $J_M$ 和

**表 2 本文网络在 DRIVE 上对 SMCM 中 *r* 的敏感性对比** %

| $r$ | $A$ | $S_{sen}$ | $S_{spe}$ | $F_1$ | AUC | $J_M$ | $J_K$ |
|---|---|---|---|---|---|---|---|
| (1, 2, 3, 4) | 96.94 | 81.83 | 98.34 | 82.41 | 98.59 | 80.74 | 80.73 |
| (1, 2, 3, 5) | 96.89 | 83.02 | 98.22 | 82.39 | 98.55 | 80.69 | 80.69 |
| (1, 2, 4, 6) | 96.93 | 81.96 | 98.38 | 82.43 | 98.55 | 80.75 | 80.74 |
| (1, 2, 4, 8) | 96.92 | **83.68** | 98.20 | **82.66** | **98.67** | **80.98** | **80.97** |
| (1, 3, 5, 7) | 96.93 | 82.57 | 98.31 | 82.49 | 98.54 | 80.80 | 80.80 |
| (1, 4, 6, 8) | **96.96** | 82.38 | **98.49** | 82.38 | 98.56 | 80.73 | 80.71 |

注. 加粗表示最优值.

**表 3　本文网络在 3 个数据集上对 $P_R$ 的敏感性实验**　　%

| 数据集 | $P_R$ | $A$ | $S_{sen}$ | $S_{spe}$ | $F_1$ | AUC | $J_M$ | $J_K$ |
|---|---|---|---|---|---|---|---|---|
| DRIVE | 0.80 | 96.92 | 82.07 | 98.35 | 82.37 | 98.54 | 80.69 | 80.69 |
| | 0.81 | 96.92 | **83.68** | 98.20 | **82.66** | **98.67** | **80.98** | **80.97** |
| | 0.82 | 96.97 | 80.74 | 98.53 | 82.36 | 98.51 | 80.72 | 80.70 |
| | 0.83 | 96.96 | 82.11 | 98.38 | 82.53 | 98.64 | 80.86 | 80.86 |
| | 0.84 | 96.91 | 82.94 | 98.26 | 82.48 | 98.58 | 80.79 | 80.78 |
| | 0.85 | **97.02** | 80.54 | **98.60** | 82.55 | 98.67 | 80.95 | 80.92 |
| CHASE_DB1 | 0.80 | 97.46 | **86.12** | 98.23 | 81.06 | 99.00 | 79.87 | 79.71 |
| | 0.81 | 97.52 | 84.93 | 98.36 | 81.18 | 98.95 | 79.94 | 79.85 |
| | 0.82 | **97.58** | 83.79 | **98.51** | 81.36 | 98.83 | 80.10 | 80.06 |
| | 0.83 | 97.42 | 86.10 | 98.18 | 80.80 | 98.94 | 79.60 | 79.43 |
| | 0.84 | 97.57 | 84.99 | 98.42 | **81.51** | **99.05** | **80.29** | **80.21** |
| | 0.85 | 97.56 | 83.85 | 98.48 | 81.21 | 98.96 | 79.96 | 79.91 |
| STARE | 0.80 | 97.47 | 83.62 | **98.61** | 83.38 | 99.01 | 82.08 | 82.02 |
| | 0.81 | 97.46 | 83.31 | 98.60 | 83.17 | 98.93 | 81.92 | 81.80 |
| | 0.82 | 97.48 | 84.28 | 98.56 | 83.53 | 99.00 | 82.23 | 82.17 |
| | 0.83 | 97.47 | 83.68 | 98.59 | 83.35 | 98.96 | 82.07 | 81.98 |
| | 0.84 | 97.50 | 84.07 | 98.61 | 83.63 | **99.04** | 82.36 | 82.28 |
| | 0.85 | **97.51** | **84.87** | 98.55 | **83.81** | 99.02 | **82.50** | **82.46** |

注. 加粗表示最优值.

$J_K$ 都是 6 组数据中的最高值，说明选取 0.81 的 $P_R$ 分割 DRIVE 的有效性；同样，对于 CHASE_DB1 和 STARE 数据集，本文选取多项指标表现优异的 0.84 的 $P_R$ 和 0.85 的 $P_R$ 分别作为分割这 2 个数据集的最终参数设置. 实验结果表明，本文选取的 $P_R$ 值能够更好地发挥网络的性能，实现最优的分割效果.

### 3.5　与第 2 位专家分割结果的比较

(1) 定量比较. 将本文网络与第 2 位专家在 3 个数据集上的分割结果进行比较，结果如表 4 所示. 可以看出，本文网络大部分的评价指标都优于第 2 位专家，并且在复杂血管图像(CHASE_DB1 和 STARE 数据集)上表现出良好的分割性能. 虽然本文网络是通过第 1 位专家的标注信息作为金标准训练的，但是在衡量网络综合性能的指标 $F_1$ 上，与第 2 位专家的结果相比，本文网络有了明显的提高. 通常，$0 \leqslant J_K \leqslant 1$，当 $J_K \geqslant 0.8$ 时表示网络分割结果和专家结果一致性非常高. 表 4 中，本文网络在 3 个数据集上的 $J_K$ 分别为 DRIVE(0.809 7), CHASE_DB1(0.802 1)和 STARE(0.824 6)，表明该网络的分割结果与专家标注结果的高度一致性. 实验结果表明，本文网络具有一定的优势和良好的鲁棒性，并且对于血管分割任务中不同类型的血管图像表现良好.

(2) 可视化比较. 为了更加直观地验证本文网络的优越性，一方面，对分割结果进行可视化，并与第 2 位专家的标注结果进行比较；另一方面，对网络在 3 个数据集上的 ROC 曲线进行可视化. 表 5 所示为本文网络在 3 个测试集中分割结果的可视

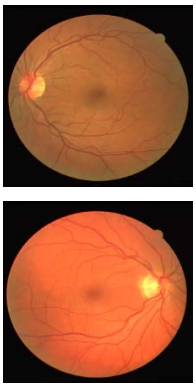
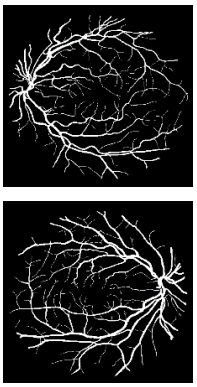
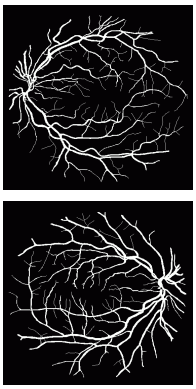
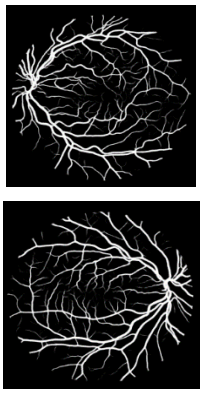
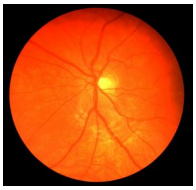
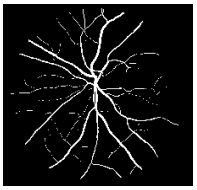
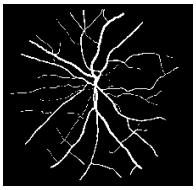
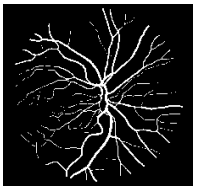
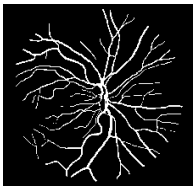
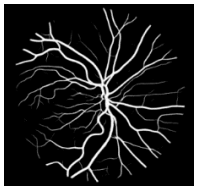
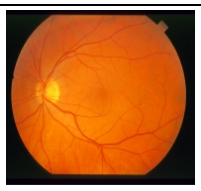
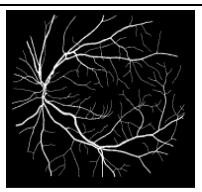
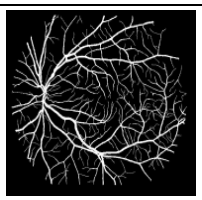
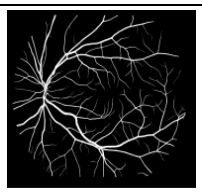
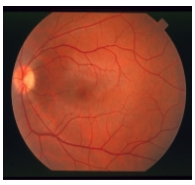
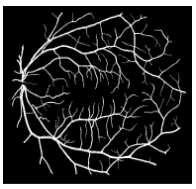
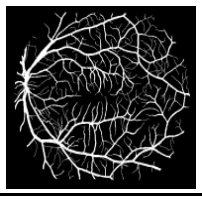
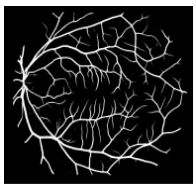
**表 4　本文网络与第 2 位专家在 3 个数据集上的分割结果比较**　　%

| 数据集 | 方法 | $A$ | $S_{sen}$ | $S_{spe}$ | $F_1$ | AUC | $J_M$ | $J_K$ |
|---|---|---|---|---|---|---|---|---|
| DRIVE | 第 2 位专家标注结果[32] | 96.37 | 77.43 | 98.19 | 78.89 | 87.81 | | |
| | 本文 | **96.92** | **83.68** | **98.20** | **82.66** | **98.67** | **80.98** | **80.97** |
| CHASE_DB1 | 第 2 位专家标注结果[32] | 97.33 | 83.13 | 98.29 | 79.69 | 90.71 | | |
| | 本文 | **97.57** | **84.99** | **98.42** | **81.51** | **99.05** | **80.29** | **80.21** |
| STARE | 第 2 位专家标注结果[32] | 95.22 | **90.71** | 95.64 | 74.17 | 92.91 | | |
| | 本文 | **97.51** | 84.87 | **98.55** | **83.81** | **99.02** | **82.50** | **82.46** |

注. 加粗表示最优值.

化效果图，图 5 所示为本文网络在 3 个数据集中的 ROC 曲线．从表 5 可以看出，与第 2 位专家的标注结果相比，本文网络的分割效果更好，血管的完整性和连通性也相对更强．

**表 5　本文网络在 3 个数据集上的可视化分割结果比较**

| 数据集 | 原始图像 | 金标准 | 第 2 位专家标注结果 | 本文网络 |
|---|---|---|---|---|
| DRIVE |  |  |  |  |
| |  |  |  |  |
| CHASE_DB1 |  |  |  |  |
| |  |  |  |  |
| STARE |  |  |  |  |
| |  |  |  |  |

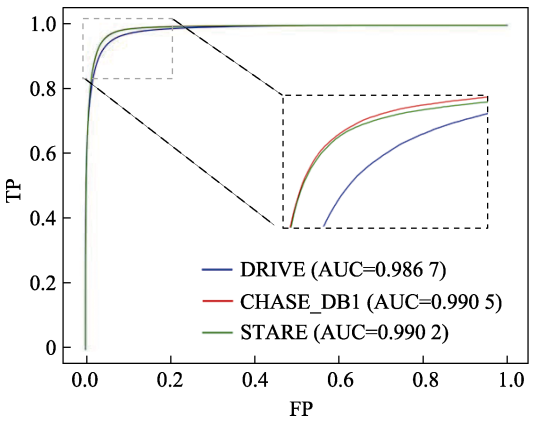



图 5　本文网络在 3 个数据集中的 ROC 曲线

### 3.6　与其他网络在 3 个数据集上的性能比较

将本文网络与 9 个目前主流的视网膜血管图像分割网络在 3 个数据集上的 $A$，$S_{sen}$，$S_{spe}$ 和 AUC 的性能指标进行对比．为了比较的公平性，本文复现了 U-Net, Att-UNet 和 UNet++, 并使用与本文网络相同的实验设置和训练策略．

表 6 所示为 10 个分割网络在 DRIVE 数据集上的性能比较．可以看出，本文网络的 $A$，$S_{sen}$，$S_{spe}$ 和 AUC 值分别为 96.92%，83.68%，98.20% 和 98.67%，其中，$A$，$S_{sen}$ 和 AUC 取得最高值；虽然 SD-UNet 取得最高的 $S_{spe}$，但是其他 3 个指标均低

**表 6　10 个网络在 DRIVE 数据集上的性能比较 %**

| 网络 | 出版年 | $A$ | $S_{sen}$ | $S_{spe}$ | AUC |
|---|---|---|---|---|---|
| U-Net[9] | 2015 | 96.41 | 76.95 | 98.28 | 97.34 |
| SD-UNet[17] | 2019 | 96.74 | 78.91 | **98.48** | 98.36 |
| CE-Net[20] | 2019 | 95.45 | 83.09 | 97.47 | 97.79 |
| Att-UNet[10] | 2019 | 96.51 | 77.46 | 98.34 | 97.78 |
| DDNet[12] | 2020 | 95.94 | 81.26 | 97.88 | 97.96 |
| UNet++[11] | 2020 | 96.71 | 76.30 | 98.67 | 97.90 |
| CSU-Net[21] | 2021 | 95.65 | 80.71 | 97.82 | 98.01 |
| DMNet[16] | 2022 | 96.83 | 83.63 | 98.11 | |
| CC-Net[22] | 2022 | 96.70 | 81.33 | 98.18 | 98.46 |
| 本文 | | **96.92** | **83.68** | 98.20 | **98.67** |

注. 加粗表示最优值.

于本文网络，尤其是体现网络能否正确识别血管的 $S_{sen}$ 指标上，本文网络比 SD-UNet 高 4.77 个百分点.

表 7 所示为 10 个分割网络在 CHASE_DB1 数据集上的性能比较. 可以看出，本文网络取得的 $A$，$S_{sen}$，$S_{spe}$ 和 AUC 值分别为 97.57%，84.99%，98.42%和 99.05%，其中，$A$ 和 AUC 达到最高值，$S_{sen}$ 达到次高值；虽然 DMNet 在 $S_{sen}$ 上略高于本文网络，但是在 $A$ 和 $S_{spe}$ 上都低于本文网络，说明本文网络在对血管像素分割的准确度上和对背景像素分割的性能上都优于 DMNet；与拥有最高特异性的 SD-UNet 相比，本文网络的 $S_{sen}$ 提高 9.4 个百分点.

**表 7　10 个网络在 CHASE_DB1 数据集上的性能比较 %**

| 网络 | 出版年 | $A$ | $S_{sen}$ | $S_{spe}$ | AUC |
|---|---|---|---|---|---|
| U-Net[9] | 2015 | 97.07 | 78.08 | 98.35 | 97.43 |
| SD-UNet[17] | 2019 | 97.38 | 75.59 | **99.00** | 98.72 |
| Att-UNet[10] | 2019 | 97.28 | 77.67 | 98.76 | 98.09 |
| DDNet[12] | 2020 | 96.37 | 82.68 | 97.73 | 98.12 |
| UNet++[11] | 2020 | 97.34 | 79.65 | 98.52 | 98.24 |
| CAR-UNet[27] | 2021 | 97.51 | 84.39 | 98.39 | 98.98 |
| CSU-Net[21] | 2021 | 97.06 | 84.27 | 98.36 | 98.24 |
| DMNet[16] | 2022 | 97.14 | **85.41** | 97.94 | |
| CC-Net[22] | 2022 | 97.18 | 82.09 | 98.16 | 98.56 |
| 本文 | | **97.57** | 84.99 | 98.42 | **99.05** |

注. 加粗表示最优值.

表 8 所示为 10 个分割网络在 STARE 数据集上的性能比较. 可以看出，本文网络取得的 $A$，$S_{sen}$，$S_{spe}$ 和 AUC 值分别为 97.51%，84.87%，98.55%和 99.02%，其中，$A$ 和 $S_{sen}$ 达到最高值，AUC 达到次高值；虽然 CAR-UNet 取得了最高的 AUC 值，但是本文网络在其他 3 个指标($A$，$S_{sen}$ 和 $S_{spe}$ )上均高于 CAR-UNet.

**表 8　10 个网络在 STARE 数据集上的性能比较 %**

| 网络 | 出版年 | $A$ | $S_{sen}$ | $S_{spe}$ | AUC |
|---|---|---|---|---|---|
| U-Net[9] | 2015 | 96.78 | 76.11 | 98.49 | 97.75 |
| SD-UNet[17] | 2019 | 97.25 | 75.48 | **98.99** | 98.50 |
| Att-UNet[10] | 2019 | 97.09 | 81.00 | 98.43 | 98.29 |
| DDNet[12] | 2020 | 96.98 | 83.98 | 97.69 | 98.58 |
| UNet++[11] | 2020 | 97.31 | 78.33 | 98.87 | 98.55 |
| CAR-UNet[27] | 2021 | 97.43 | 84.45 | 98.50 | **99.11** |
| CSU-Net[21] | 2021 | 97.02 | 84.32 | 98.45 | 98.25 |
| DMNet[16] | 2022 | 97.32 | 82.72 | 98.47 | |
| CC-Net[22] | 2022 | 97.24 | 82.01 | 98.51 | 98.79 |
| 本文 | | **97.51** | **84.87** | 98.55 | 99.02 |

注. 加粗表示最优值.

综上所述，从表 6~表 8 可以看出，本文网络在视网膜血管分割方面的性能优于大部分现有主流网络，表明本文网络能够很好地完成视网膜血管分割任务. 表 9~表 11 所示为本文网络在 3 个数据集上单幅图像的测试结果比较.

进一步，对相同实验条件下 4 个网络的分割结果进行可视化，结果如表 12 所示，其中，不同网络分割结果的局部细节被放大到每个子图下方的矩形框中展示. 可以看出，本文网络能够分割出眼底图像中大部分血管，包括一些低对比度的尺度差异的粗血管和微细血管；在放大的局部细节中，与其他网络相比，本文网络更好地分割出分叉和边缘血管，避免了噪声干扰，保留了血管的完整性和连通性. 从图 5b~5d 可以看出，本文网络获得了最好的 ROC 曲线. 综合考虑表 6~表 12 和图 6，本文网络在定量和可视化的比较中都获得较为优秀的结果，证明了本文网络在视网膜血管分割任务中具有的优越性.

### 3.7　在 DRIVE 数据集上的消融实验

在 DRIVE 数据集上进行消融实验，验证本文网络中组件的有效性，进一步证明本文网络的有效性，结果如表 13 所示. 其中，Base 表示不含任何改进组件，基于编码-解码网络的轻量级 U 型网络. 训练-测试集依据 DRIVE 数据集官方的数据划分方法，避免因数据划分不同带来的随机性.

从表 13 可以看出：

(1) 比较 Base 和 Base_SAM. 在轻量级的 Base 基础上加入 SAM 后，6 项指标全部得到提升，在参

**表 9 本文网络在 DRIVE 上单幅测试图像分割结果** %

| 图像 | $A$ | $S_{sen}$ | $S_{spe}$ | AUC | 图像 | $A$ | $S_{sen}$ | $S_{spe}$ | AUC |
|---|---|---|---|---|---|---|---|---|---|
| 01 | 96.78 | 89.33 | 97.51 | 99.06 | 12 | 97.04 | 85.42 | 98.14 | 98.92 |
| 02 | 97.02 | 86.92 | 98.17 | 99.05 | 13 | 96.68 | 80.35 | 98.45 | 98.43 |
| 03 | 96.45 | 77.77 | 98.52 | 98.50 | 14 | 96.94 | 88.01 | 97.72 | 98.90 |
| 04 | 97.10 | 81.45 | 98.69 | 98.41 | 15 | 97.32 | 87.35 | 98.09 | 99.04 |
| 05 | 98.81 | 79.31 | 98.62 | 98.32 | 16 | 97.09 | 85.78 | 98.21 | 99.05 |
| 06 | 96.59 | 78.23 | 98.57 | 98.19 | 17 | 96.85 | 83.25 | 98.10 | 98.82 |
| 07 | 96.98 | 77.03 | 98.98 | 98.45 | 18 | 96.76 | 90.10 | 97.34 | 99.09 |
| 08 | 96.87 | 79.38 | 98.51 | 98.40 | 19 | 97.54 | 92.07 | 98.04 | 99.31 |
| 09 | 97.09 | 80.89 | 98.52 | 98.48 | 20 | 96.97 | 90.31 | 97.50 | 99.12 |
| 10 | 96.96 | 82.50 | 98.26 | 98.49 | 平均值 | 96.92 | 83.68 | 98.20 | 98.67 |
| 11 | 96.64 | 82.74 | 98.01 | 98.43 | | | | | |

**表 10 本文网络在 CHASE_DB1 上单幅测试图像分割结果** %

| 图像 | $A$ | $S_{sen}$ | $S_{spe}$ | AUC | 图像 | $A$ | $S_{sen}$ | $S_{spe}$ | AUC |
|---|---|---|---|---|---|---|---|---|---|
| 11L | 98.17 | 86.95 | 98.8 | 99.23 | 13R | 97.08 | 83.33 | 98.01 | 98.78 |
| 11R | 98.13 | 84.97 | 98.87 | 99.25 | 14L | 97.88 | 82.48 | 99.02 | 99.31 |
| 12L | 96.94 | 86.39 | 97.54 | 98.78 | 14R | 97.63 | 82.57 | 98.56 | 98.95 |
| 12R | 97.32 | 87.18 | 98.13 | 99.12 | 平均值 | 97.57 | 84.99 | 98.42 | 99.05 |
| 13L | 97.41 | 85.9 | 98.16 | 98.97 | | | | | |

**表 11 本文网络在 STARE 上单幅测试图像分割结果** %

| 图像 | $A$ | $S_{sen}$ | $S_{spe}$ | AUC | 图像 | $A$ | $S_{sen}$ | $S_{spe}$ | AUC |
|---|---|---|---|---|---|---|---|---|---|
| 01 | 97.31 | 76.00 | 99.16 | 98.89 | 163 | 98.16 | 91.45 | 98.73 | 99.65 |
| 02 | 97.53 | 70.15 | 99.48 | 98.64 | 235 | 97.71 | 90.45 | 98.42 | 99.52 |
| 03 | 98.01 | 84.34 | 98.88 | 99.05 | 236 | 98.07 | 87.47 | 99.12 | 99.58 |
| 04 | 96.67 | 65.82 | 99.14 | 98.52 | 239 | 97.34 | 88.96 | 98.13 | 98.97 |
| 05 | 96.18 | 82.25 | 97.56 | 98.56 | 240 | 96.69 | 85.04 | 98.02 | 98.29 |
| 44 | 96.68 | 92.38 | 97.01 | 99.09 | 255 | 97.72 | 87.92 | 98.68 | 99.30 |
| 77 | 97.86 | 92.62 | 98.32 | 99.57 | 291 | 98.05 | 85.90 | 98.70 | 98.46 |
| 81 | 97.99 | 91.04 | 98.55 | 99.52 | 319 | 98.17 | 87.17 | 98.67 | 99.39 |
| 82 | 98.23 | 87.92 | 99.11 | 99.61 | 324 | 96.98 | 70.37 | 98.89 | 97.96 |
| 139 | 96.72 | 91.85 | 97.14 | 99.01 | 平均值 | 97.51 | 84.87 | 98.55 | 99.02 |
| 162 | 98.03 | 83.37 | 99.15 | 99.48 | | | | | |

数量仅增加了 98 的情况下, 性能指标从 96.13%, 74.89%, 98.16%, 77.22%, 96.47%和 75.15%提高到 96.22%, 75.65%, 98.20%, 77.82%, 96.83%和 75.79%, 说明加入 SAM 组件策略的有效性.

(2) 比较 Base_SAM 和 Base_SAM_ECAM. 在仅仅增加 9 个参数的情况下, Base_SAM_ECAM 的指标全部得到提升, 与 Base_SAM 相比, $A$ 提高 0.50 个百分点, $S_{sen}$ 提高 1.97 个百分点, $S_{spe}$ 提高 0.36 个百分点, $F_1$ 提高 2.78 个百分点, AUC 提高 0.89 个百分点, $J_M$ 提高 3.10 个百分点, 说明本文在跳跃连接阶段使用 ECAM 处理视网膜血管特征的有效性.

(3) 比较 Base_SAM_ECAM 和 Base_SAM_ECAM_SDCM. 使用 SDCM 之后, 网络的 $A$, $S_{sen}$, $F_1$, AUC 和 $J_M$ 都得到提高, 虽然 $S_{spe}$ 有所下降, 但是体现网络整体分割性能的 $F_1$ 和 AUC 分别提高 1.51 个百分点和 0.52 个百分点, 说明 SDCM 对提高网络的分割性能是有效的.

(4) 比较 Base_SAM_ECAM_SDCM 和 Base_SAM_ECAM_SDCM_SMCM(即本文网络). 使用 SMCM 之后, 网络的 $A$, $S_{sen}$, $F_1$, AUC 和 $J_M$ 都得到提高; 最明显的改进体现在 $A$, 从 81.72%提高到 83.68%, 说明 SMCM 可以有效地捕获不同尺度的血管特征, 在避免血管信息丢失的同时, 有效

**表 12　4 个网络在 3 个数据集分割结果图细节比较**

| 数据集 | 原始图像 | 金标准 | U-Net[9] | Att-UNet[10] | UNet++[11] | 本文网络 |
|---|---|---|---|---|---|---|
| DRIVE | | | | | | |
| CHASE_DB1 | | | | | | |
| STARE | | | | | | |

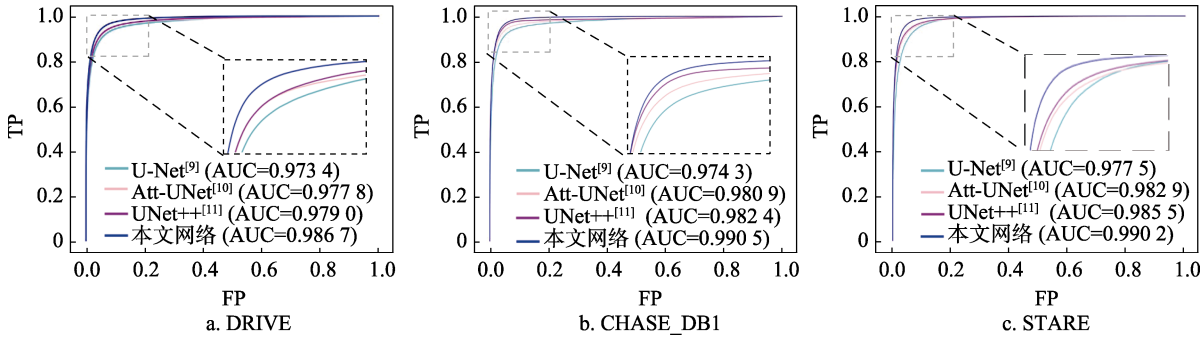



图 6　4 个网络在 3 个数据集上的 ROC 曲线图比较

**表 13　本文网络在 DRIVE 数据集上的消融实验结果比较**　　%

| 网络 | 参数量 | $A$ | $S_{sen}$ | $S_{spe}$ | $F_1$ | AUC | $J_M$ |
|---|---|---|---|---|---|---|---|
| Base | 535 793 | 96.13 | 74.89 | 98.16 | 77.22 | 96.47 | 75.15 |
| Base_SAM | 535 891 | 96.22 | 75.65 | 98.20 | 77.82 | 96.83 | 75.79 |
| Base_SAM_ECAM | 535 900 | 96.72 | 77.62 | **98.56** | 80.60 | 97.72 | 78.89 |
| Base_SAM_ECAM_SDCM | 538 716 | 96.88 | 81.72 | 98.33 | 82.11 | 98.24 | 80.41 |
| Base_SAM_ECAM_SDCM_SMCM | 638 732 | **96.92** | **83.68** | 98.20 | **82.66** | **98.67** | **80.98** |

注．加粗表示最优值．

地兼顾边缘血管特征信息的提取．虽然参数量有所增加，但是本文网络最终的参数量只有 638 732．因此，对于视网膜血管分割任务，本文网络是一个轻量且高效的神经网络．

综上，消融实验结果验证了本文提出的方法和模块的有效性．

图 7 定性地展示了不同实验网络在 DRIVE 数据集上的分割结果，其中，每幅分割图像下包含 2 个被放大的局部细节图．可以看出，本文网络的分割效果优于其他实验网络，尤其是对于微细血管

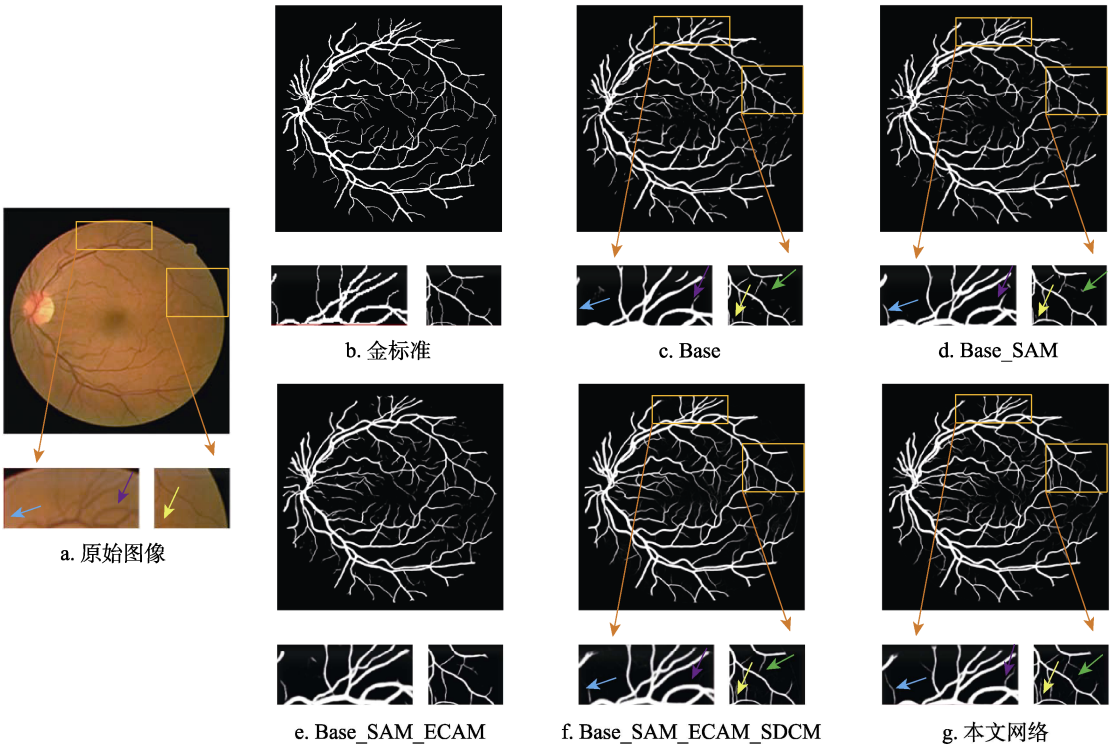



图 7 不同实验网络在 DRIVE 数据集分割结果的局部细节对比

的分割方面优势更明显.

### 3.8 泛化能力

在 DRIVE 和 STARE 数据集上使用交叉训练策略证明本文网络的泛化能力，选择这 2 个数据集是因为它们的原始图像具有相似大小的分辨率. 本文将在一个数据集上训练出的网络直接用于另一个数据集的评估而无需进行重新训练，这样做将更具有实际意义. 此外，对于已有的可复现的文献[14-15,34-35]也基于这 2 个数据集进行了交叉训练，定量的统计结果如表 14 所示. 可以看出，当在 DRIVE 上训练得到本文网络、在 STARE 上进行测试时，与其他的网络相比，本文网络在 $A$, $S_{sen}$ 和 AUC 上都取得了最高值；当本文网络在 STARE 上训练、在 DRIVE 上进行测试时，尽管 STARE 数据集中微细血管的标注信息较少，但是与其他网络相比，本文网络仍然取得了最高的 $A$, $S_{spe}$ $S_{spe}$ 和 AUC. 实验结果表明，本文网络比其他网络具有更好的泛化能力.

**表 14 验证网络泛化性的交叉训练实验结果比较**

| 数据集 | | 网络 | 指标/% | | | |
|---|---|---|---|---|---|---|
| 训练 | 测试 | | $A$ | $S_{sen}$ | $S_{spe}$ | AUC |
| DRIVE | STARE | Fraz 等[34] | 94.95 | 70.10 | 97.70 | 96.60 |
| | | Yan 等[15] | 95.80 | 73.19 | 98.40 | 96.78 |
| | | Yan 等[14] | 95.69 | 72.11 | **98.40** | 97.08 |
| | | CcNet[35] | 95.63 | 74.99 | 97.98 | 96.21 |
| | | 本文 | **96.58** | **75.47** | 98.31 | **97.35** |
| STARE | DRIVE | Fraz 等[34] | 94.56 | 72.42 | 97.92 | 96.97 |
| | | Yan 等[15] | 94.44 | 70.14 | 98.02 | 95.68 |
| | | Yan 等[14] | 94.94 | **72.92** | 98.15 | 95.99 |
| | | CcNet[35] | 94.86 | 72.17 | 98.20 | 93.27 |
| | | 本文 | **96.41** | 72.44 | **98.71** | **97.26** |

注. 加粗表示最优值.

### 3.9 讨 论

#### 3.9.1 计算复杂度

在医疗临床应用环境中，通常不一定有足够的硬件计算资源来部署和运行高复杂度的网络，因此，深度学习网络的计算复杂度是评估其应用前景的重要指标. 网络的参数量 $Q$ 和浮点运算次数(floating-point operations per second, FLOPs)[36] 是适用于衡量网络复杂度的评价指标，公式分别为

$$Q=\left(C_{in}\times K^2+1\right)\times C_{out} \tag{17}$$

$$\text{FLOPs}=2\times H\times W\left(C_{in}\times K^2+1\right)\times C_{out} \tag{18}$$

其中，$C_{in}$ 和 $C_{out}$ 分别表示输入和输出特征图的通道数；$K$ 表示卷积核大小；$H$ 和 $W$ 分别表示特征图的高和宽. 表 15 所示为本文网络和其他相同实验条件下的视网膜血管分割网络的复杂度比较. 可以看出，与其他网络相比，本文网络的 $Q$ 和 FLOPs 分别只有 0.63M 和 1.27M，复杂度是最低

**表 15　STARE 上不同网络的复杂度比较**

| 网络 | $Q$ | FLOPs | $A$/% | $S_{sen}$/% | $F_1$/% | $J_M$/% |
|---|---|---|---|---|---|---|
| U-Net[9] | 7.76M | 15.51M | 96.78 | 76.11 | 78.32 | 76.70 |
| Att-UNet[10] | 8.21M | 16.42M | 97.09 | 81.00 | 81.06 | 79.49 |
| UNet++[11] | 9.04M | 18.07M | 97.31 | 78.33 | 81.51 | 80.15 |
| CAR-UNet[27] | 1.05M | 2.10M | 97.43 | 84.45 | 83.45 | 82.16 |
| 本文 | **0.63M** | **1.27M** | **97.51** | **84.87** | **83.81** | **82.50** |

注. 加粗表示最优值.

的; 另外, 该网络并未因为复杂度的降低而损失分割性能, 因为其在 $A$, $S_{sen}$, $F_1$ 和 $J_M$ 得分都是最高的, 表明本文网络具有更好的实际应用价值.

3.9.2　额外的评估和未来的工作

目前, 所有的视网膜血管分割指标都是基于像素的测量, 由于微细血管的所占像素有限, 因此对于指标的贡献不大. 同时, 由于视网膜血管图像中只有9%~14%的像素属于血管像素, 因此上述基于像素的指标很大可能会遇到数据类别不平衡问题. 显然, 在视网膜血管分割任务中, 微细血管的分割比粗血管更具有挑战性. 为了证明本文的网络在微细血管方面的分割优势, 本文进一步测试了本文网络与其他 4 种主流分割网络在 STARE 数据集上的 $J_M$ 值(因为 $J_M$ 更加适合数据类别不平衡性问题), 统计结果如表 15 所示. 可以看出, 本文网络具有最高的 $J_M$, 进一步说明本文网络在微细血管分割性能方面的优越性.

但是, 本文中用于实验的数据类型较为有限, 除了常用的彩色眼底照相图片, 还有扫描激光检眼镜(scanning laser ophthalmoscope, SLO)视网膜图像和光学相干断层扫描(optical coherence tomography, OCT)视网膜图像. 与常用的彩色眼底图像相比, SLO 和 OCT 图像包含更多的信息, 能够更有效地反映眼底的精细结构; 尤其是 OCT 图像, 能更清晰地显示出视网膜的 10 层结构, 其也是临床眼科检查的重要诊断依据[37]. 未来将专注于 SLO 和 OCT 图像, 使本文网络可以更加有效地应用于临床实践.

## 4　结　语

本文提出一个融合结构化卷积和双重注意力机制的轻量级眼底图像分割网络. 提出一种结构化卷积方法, 使用 SMCM 提取并融合多尺度的特征信息, 弥补了单一尺度卷积感受野有限, 特征提取不足的缺点; 使用 SDCM 有效地防止网络过拟合, 提高网络分割性能; 通过添加少量的参数, 将基于空间和通道的双重注意力模块嵌入到编码-解码网络中, 重新调整空间和通道的特征映射, 突出重要的血管特征, 抑制干扰噪声, 增强网络对于差异化血管的精准分割. 在 DRIVE, CHASE_DB1 和 STARE 数据集上的实验结果表明, 本文网络对于尺度差异化血管的分割较为精确, 更好地避免了干扰噪声的影响; 与其他主流的血管分割网络相比, 本文网络在参数量仅有 0.63M 的情况下, 在 3 个数据集上都实现了最优的分割性能. 此外, 在 DRIVE 上的消融实验验证了网络中相关方法和模块的有效性; 分析讨论结果表明, 本文网络具有最优的网络复杂度和良好的泛化能力. 今后, 将在保证网络轻量化的基础上进一步提高分割方法的性能, 并使其能应用于其他类型的眼底图像数据集.

## 参考文献(References):


[1] Pang Hao, Wang Cong. deep learning model for diabetic retinopathy detection[J]. Journal of Software, 2017, 28(11): 3018-3029(in Chinese)
(庞浩, 王枞. 用于糖尿病视网膜病变检测的深度学习模型[J]. 软件学报, 2017, 28(11): 3018-3029)

[2] Guo Song, Li Tao, Li Ning, *et al*. Red lesion segmentation of fundus image with multi-task learning[J]. Journal of Software, 2021, 32(11): 3646-3658(in Chinese)
(郭松, 李涛, 李宁, 等. 基于多任务学习的眼底图像红色病变点分割[J]. 软件学报, 2021, 32(11): 3646-3658)

[3] Zhu Chengzhang, Zou Beiji, Xiang Yao, *et al*. A survey of retinal vessel segmentation in fundus images[J]. Journal of Computer-Aided Design & Computer Graphics, 2015, 27(11): 2046-2057(in Chinese)
(朱承璋, 邹北骥, 向遥, 等. 彩色眼底图像视网膜血管分割方法研究进展[J]. 计算机辅助设计与图形学学报, 2015, 27(11): 2046-2057)

[4] Yin Y, Adel M, Bourennane S. Retinal vessel segmentation using a probabilistic tracking method[J]. Pattern Recognition, 2012, 45(4): 1235-1244

[5] Chaudhuri S, Chatterjee S, Katz N, *et al*. Detection of blood vessels in retinal images using two-dimensional matched filters[J]. IEEE Transactions on Medical Imaging, 1989, 8(3): 263-269

[6] Zhu Chengzhang, Cui Jinkai, Zou Beiji, *et al*. Retinal vessel segmentation based on multiple feature fusion and random forest[J]. Journal of Computer-Aided Design & Computer Graphics, 2017, 29(4): 584-592(in Chinese)
(朱承璋, 崔锦恺, 邹北骥, 等. 基于多特征融合和随机森林的视网膜血管分割[J]. 计算机辅助设计与图形学学报, 2017, 29(4): 584-592)

[7] Zhu Chengzhang, Xiang Yao, Zou Beiji, *et al*. Retinal vessel segmentation in fundus images using CART and AdaBoost[J]. Journal of Computer-Aided Design & Computer Graphics,

2014, 26(3): 445-451(in Chinese)
(朱承璋, 向遥, 邹北骥, 等. 基于分类回归树和 AdaBoost 的眼底图像视网膜血管分割[J]. 计算机辅助设计与图形学学报, 2014, 26(3): 445-451)

[8] Reda I, Ayinde B O, Elmogy M, *et al*. A new CNN-based system for early diagnosis of prostate cancer[C] //Proceedings of the 15th IEEE International Symposium on Biomedical Imaging. Los Alamitos: IEEE Computer Society Press, 2018: 207-210

[9] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[C] //Proceedings of 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Heidelberg: Springer, 2015: 234-241

[10] Schlemper J, Oktay O, Schaap M, *et al*. Attention gated networks: Learning to leverage salient regions in medical images[J]. Medical Image Analysis, 2019, 53: 197-207

[11] Zhou Z W, Siddiquee M M R, Tajbakhsh N, *et al*. UNet++: redesigning skip connections to exploit multiscale features in image segmentation[J]. IEEE Transactions on Medical Imaging, 2020, 39(6): 1856-1867

[12] Mou L, Chen L, Cheng J, *et al*. Dense dilated network with probability regularized walk for vessel detection[J]. IEEE Transactions on Medical Imaging, 2020, 39(5): 1392-1403

[13] Ghiasi G, Lin T Y, Le Q V. DropBlock: a regularization method for convolutional networks[C] //Proceedings of the 32nd International Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2018: 10750-10760

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

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

[16] Deng X Y, Ye J H. A retinal blood vessel segmentation based on improved D-MNet and pulse-coupled neural network[J]. Biomedical Signal Processing and Control, 2022, 73: Article No.103467

[17] Guo C L, Szemenyei M, Pei Y, *et al*. SD-Unet: a structured dropout U-Net for retinal vessel segmentation[C] //Proceedings of the 19th IEEE International Conference on Bioinformatics and Bioengineering. Los Alamitos: IEEE Computer Society Press, 2019: 439-444

[18] He K M, Zhang X Y, Ren S Q, *et al*. Deep residual learning for image recognition[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2016: 770-778

[19] Huang G, Liu Z, Van Der Maaten L, *et al*. Densely connected convolutional networks[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2017: 2261-2269

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

[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[J]. IEEE Journal of Biomedical and Health Informatics, 2021, 25(4): 1128-1138

[22] Yang D X, Zhao H D, Han T C. Learning feature-rich integrated comprehensive context networks for automated fundus retinal vessel analysis[J]. Neurocomputing, 2022, 491: 132-143

[23] Noh H, Hong S, Han B. Learning deconvolution network for semantic segmentation[C] //Proceedings of the IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2015: 1520-1528

[24] Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions[OL]. [2022-07-20]. https://arxiv.org/abs/1511.07122

[25] Tompson J, Goroshin R, Jain A, *et al*. Efficient object localization using convolutional networks[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2015: 648-656

[26] Guo C L, Szemenyei M, Yi Y G, *et al*. SA-UNet: spatial attention U-Net for retinal vessel segmentation[C] //Proceedings of the 25th International Conference on Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2021: 1236-1242

[27] Guo C L, Szemenyei M, Hu Y T, *et al*. Channel attention residual U-Net for retinal vessel segmentation[C] //Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Los Alamitos: IEEE Computer Society Press, 2021: 1185-1189

[28] Wang Q L, Wu B G, Zhu P F, *et al*. ECA-Net: efficient channel attention for deep convolutional neural networks[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2020: 11531-11539

[29] Woo S, Park J, Lee J Y, *et al*. CBAM: convolutional block attention module[C] //Proceedings of the 15th European Conference on Computer Vision. Heidelberg: Springer, 2018: 3-19

[30] Staal J, Abràmoff 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

[31] Fraz M M, Rudnicka A R, Owen C G, *et al*. Delineation of blood vessels in pediatric retinal images using decision trees-based ensemble classification[J]. International Journal of Computer Assisted Radiology and Surgery, 2014, 9(5): 795-811

[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] Cohen J. A coefficient of agreement of nominal scales[J]. Educational and Psychological Measurement, 1960, 20(1): 37-46

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

[35] Feng S T, Zhuo Z S, Pan D R, *et al*. CcNet: a cross-connected convolutional network for segmenting retinal vessels using multi-scale features[J]. Neurocomputing, 2020, 392: 268-276

[36] Molchanov P, Tyree S, Karras T, *et al*. Pruning convolutional neural networks for resource efficient inference[OL]. [2022-07-20]. https://arxiv.org/abs/1611.06440

[37] Xu Jun, Chen Qiang, Niu Sijie. Correlation between optic nerve head parameters and retinal nerve fiber layer in glaucoma[J]. Journal of Computer-Aided Design & Computer Graphics, 2017, 29(6): 977-983(in Chinese)
(徐军, 陈强, 牛四杰. 青光眼视神经头参数与视网膜神经纤维层的相关性分析[J]. 计算机辅助设计与图形学学报, 2017, 29(6): 977-983)