

基于注意力机制的多尺度残差 U-Net 眼底血管分割

赵凤¹, 钟蓓蓓^{1*}, 刘汉强²¹西安邮电大学通信与信息工程学院(人工智能学院), 陕西 西安 710121;²陕西师范大学计算机科学学院, 陕西 西安 710119

摘要 针对现有眼底血管分割方法难以辨别细小血管及交叉处血管分割断裂的问题,提出了一种基于注意力机制的多尺度 U 型网络。在编码阶段使用改进的残差块结构提取血管深度特征的同时有效解决过拟合问题,接着依次采用多尺度卷积模块和多尺度注意力模块进一步获取深度特征的多尺度特征信息。然后,使用 MaxBlurPool 进行池化,对数据进行降维并保证平移不变性。此外,在最后一个编码层引入混合注意力机制和并行空洞卷积,前者从通道和空间维度强调需要重点关注的信息,抑制背景区域的干扰;后者用来获取不同大小感受野的特征信息,且不会引入多余参数而导致计算负担。在解码部分,改进跳跃连接方式以抑制噪声的干扰并获得更加丰富的上下文信息。所提算法在公开的眼底数据集上取得了优于其他算法的分割效果。

关键词 图像处理; 图像分割; 视网膜血管; 注意力机制; 多尺度卷积; 空洞卷积

中图分类号 TP391

文献标志码 A

DOI: 10.3788/LOP202259.1810002

Multi-Scale Residual U-Net Fundus Blood Vessel Segmentation Based on Attention Mechanism

Zhao Feng¹, Zhong Beibei^{1*}, Liu Hanqiang²¹School of Communication and Information Engineering & School of Artificial Intelligence, Xi'an University of Posts and Telecommunications, Xi'an 710121, Shaanxi, China;²School of Computer Science, Shaanxi Normal University, Xi'an, Shaanxi 710119, China

Abstract Some existing retinal vessel segmentation methods have been unsuccessful in distinguishing weak blood vessels and have suffered from blood vessel segmentation disconnections at intersections. To solve this problem, a multi-scale U-shaped network based on attention mechanism was proposed in this paper. In the encoding part, the proposed algorithm employed the improved residual block structure to extract the depth features of blood vessels while effectively solving the overfitting problem. In turn, the multi-scale convolution module and multi-scale attention module were used to obtain multi-scale feature information of the depth features. Then, MaxBlurPool was used as the pooling method to reduce dimensions of data and ensure the translation invariance. In addition, hybrid attention module and parallel dilated convolution were presented in the last encoding layer, where the former emphasized the information that needs to be focused from the channel and space dimensions to suppress the interference of the background area and the latter was used to obtain the characteristic information of receptive fields with different sizes while not introducing redundant parameters to cause computational burden. In the decoding part, skip connection was improved to suppress noise and obtain more abundant context information. The proposed algorithm achieved better segmentation effect than other methods on public fundus datasets.

Key words image processing; image segmentation; retinal vessel; attention mechanism; multi-scale convolution; dilated convolution

1 引言

视网膜血管的结构信息对糖尿病、高血压、动脉硬化

化等疾病的临床诊断具有重要的指示作用^[1-3],临床经验丰富的医生能够通过观察眼底图像中血管的外观表现对相关疾病进行初步判断,有助于早期发现隐藏疾

收稿日期: 2021-06-07; 修回日期: 2021-07-05; 录用日期: 2021-07-20

基金项目: 国家自然科学基金(62071379, 62071378, 61571361)、陕西省自然科学基金基础研究计划(2021JM-461, 2020JM-299)、西安邮电大学西邮新星团队(xyt2016-01)

通信作者: *2871188907@qq.com

病并提醒患者及时治疗。然而,人工分割方法需要对大量眼底视网膜图像进行准确标注,这对标注人员的专业能力及熟练度都有着很高的要求,所以在实际临床诊断中存在一定的局限性。因此,有必要研发一种端到端的眼底血管自动分割方法,来协助医生客观高效地对疾病进行筛查、诊断和治疗。

眼底血管分割任务的目的是区分目标血管和背景区域,属于典型的二分类问题。专家学者将计算机辅助视网膜疾病诊断方法主要分为无监督学习分割方法和监督学习分割方法^[4]。无监督学习分割方法不需要数据对应的标签,通过学习数据自身的信息确定内部特性并找出数据之间的规律,是一种自主学习的过程。通常使用的无监督学习方法包括聚类方法、匹配滤波方法、几何形态学方法等。文献[5]提出了一种基于可能性模糊聚类的分割方法,并使用布谷鸟搜索对目标函数进行优化。文献[6]提出了一种结合匹配滤波器和高斯的一阶导数的视网膜血管提取方法,该方法能够有效区分血管结构和非血管结构。文献[7]提出了一种基于Kirsch模板的形态学操作来提取视网膜静脉,使得血管的连通性得以保持。文献[8]使用多结构元素重建的形态学操作来增强图像,消除伪边缘,保留了真正的细血管边缘,提高了血管分割准确性。无监督学习方法的优点是分割速度快,但根据样本数据的固有特征而设计的模型往往不易定位和提取细小血管的特征,因此分割效果通常并不理想。

不同于无监督方法根据样本数据进行自主学习的思想,监督学习分割方法的实现总体分为两步:首先使用特征提取器获得目标血管的特征信息;然后利用分类器判别每个像素点属于目标还是背景。通常情况下,监督学习分割方法分为传统监督学习方法和基于深度学习的分割方法。在传统监督学习方法中,特征提取器和分类器多选用经典的机器学习方法。文献[9]使用多尺度Gabor滤波器识别血管,并利用主成分分析方法提取特征,最后结合支持向量机和高斯混合模型进行血管分割。文献[10]使用滤波器组确定视网膜病变区域的位置,然后使用集成分类器完成分类。然而,传统的机器学习方法容易受到主观因素的影响而导致丢失部分信息,进而产生分割准确性不稳定现象。近年来,基于深度学习的分割方法占据主导地位,使用卷积神经网络自动提取图像的特征,可以学习到更加丰富的特征信息。文献[11]提出的基于改进HED网络的视网膜血管分割算法使用残差可变形卷积增强模型获取血管形状和尺寸的能力,能够有效解决在细小血管和病变区域血管分割能力不足的问题。文献[12]提出的R2U-Net将递归卷积神经网络和残差网络整合到U-Net结构,通过结合这两种网络的优势,在参数量相同的前提下得到更好的分割效果,其中残差网络可以用来训练深层网络,递归残差卷积层用来累积特征以获得目标图像更好的特征表示。文献[13]提出的

MS-NFN将图像输入两个具有相同U-Net结构的子模型中,通过提取视网膜图像的多尺度特征来提高对毛细血管的分割准确度。文献[14]提出了一个新的轻量级网络Vessel-Net,Vessel-Net将inception-residual模块嵌入U型编码-解码结构中,以提高血管特征的代表能力。文献[15]设计了一种“注意力引导滤波器”的注意力机制,用来传递来自前面特征图的结构信息,以更好地保留图像的结构信息。文献[16]提出了一种基于双流网络的视网膜血管分割方法,一个网络用来对血管图像进行整体分割,另一个网络用来对细血管进行精细分割。总体而言,基于深度学习的分割方法解决了无监督学习方法和传统监督学习方法存在的不足,使得视网膜血管的分割精度得到显著提升。

本文提出了一种基于注意力机制的多尺度残差U型网络,该网络能够有效解决细小血管难以识别并且容易发生分割断裂的问题,进一步提高分割灵敏度和准确度。主要工作如下:1)使用改进的残差块结构和多尺度卷积模块进行特征提取,在提高血管多尺度深度特征提取能力的同时,能够有效解决层数加深带来的网络退化以及过拟合问题。接着,使用多尺度注意力模块为这些具有多尺度信息的关键特征分配更多的注意力,有效降低非血管像素造成的干扰;2)使用MaxBlurPool在下采样时尽可能地保持平移不变性,以保留更多的血管局部细节信息;3)在最后一个编码层的两个改进的残差块之间引入混合注意力模块,从通道和空间两个维度自适应细化特征的信息。在编码部分底部引入并行空洞卷积模块,在提取血管的多尺度特征信息的同时,有效降低额外参数带来的计算复杂度;4)改进解码部分的跳跃连接方式,首先引入注意力(AG)模块来抑制背景噪声的干扰,然后,使用多尺度注意力模块进一步关注AG模块获得的特征图的多尺度信息,为同级特征的拼接融合提供更加丰富的上下文信息,进而提高分割精度。

2 网络结构与算法原理

本实验组借鉴U-Net中的U型结构,如[图1(a)]所示,对U-Net中的普通卷积进行了改进,进而提出了一种基于注意力机制的多尺度残差U型网络,如[图1(b)]所示。在编码部分,使用两个改进的残差块结构,不仅能够解决由于网络层数加深而带来的梯度消失或梯度爆炸问题,而且还能对网络模型进行有效的正则化处理,缓解深度网络由于数据量较少而产生的过拟合现象,高效地提取眼底血管图像的深度特征。接着,使用MaxBlurPool作为池化方法来对特征信息进行降维,在帮助网络模型减小计算量的同时尽可能地避免特征信息的丢失。在最后一层结构中,引入混合注意力模块和并行空洞卷积。在解码部分对跳跃连接方式进行改进,首先使用AG模块调整血管和背景像素的权重因子,专注于目标血管的特征信息,有效抑

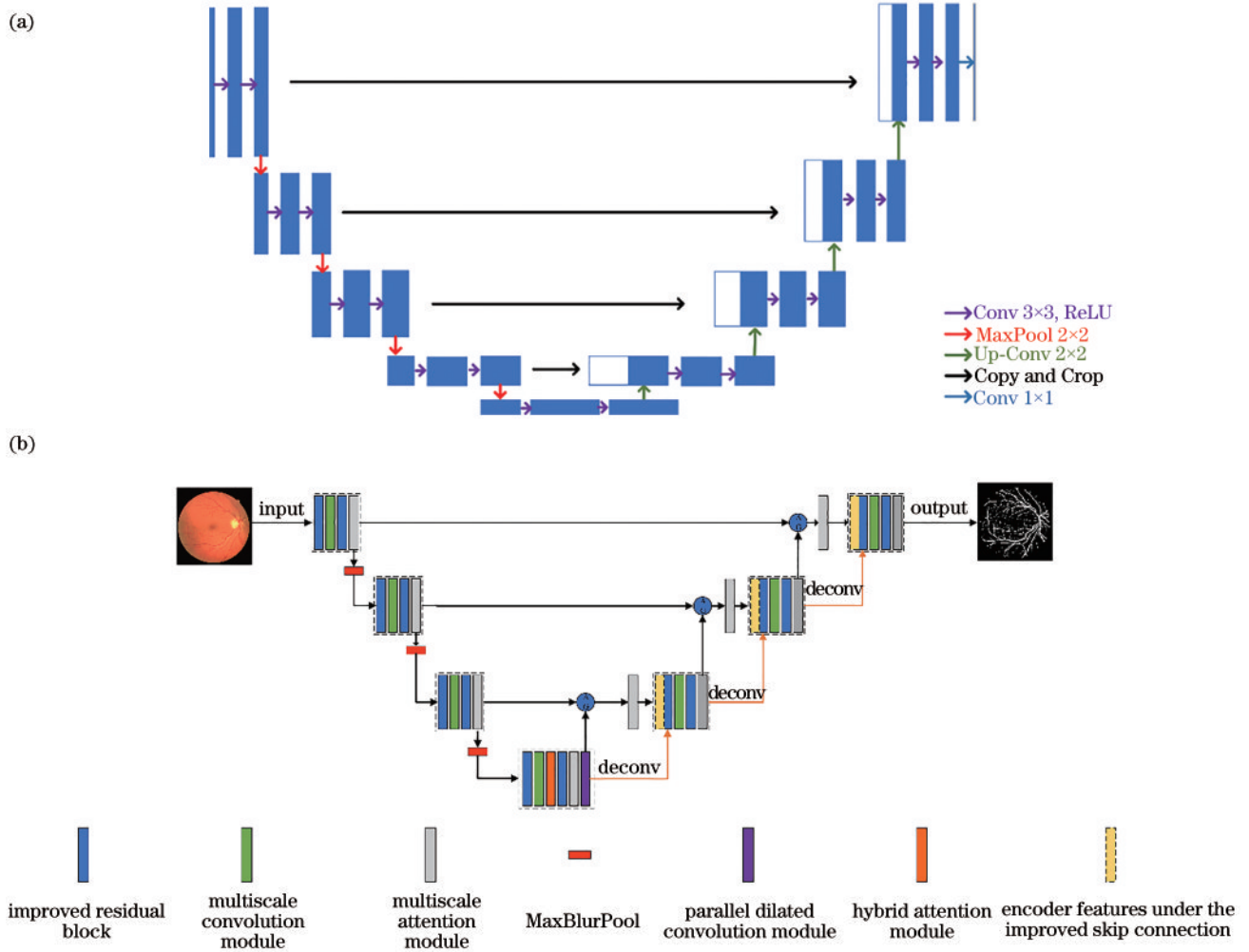


图 1 网络结构图对比。(a) U-Net; (b) 基于注意力机制的多尺度残差 U 型网络

Fig. 1 Network structure comparison. (a) U-Net; (b) multi-scale residual U-shaped network based on attention mechanism

制非血管区域带来的噪声干扰。然后,将获取到的关键特征输入多尺度注意力模块中,得到这些关键特征的多尺度信息。改进的跳跃连接方式能够更好地融合图像的上下文信息,帮助解码部分更好地恢复原始图像,进而提高分割精度。最后使用 softmax 激活函数对解码结果进行二分类,实现血管和背景的分割。

2.1 改进的残差块结构

为了提高图像分割性能,现有的经典卷积神经网络大多通过增加层数来更好地拟合特征^[17]。然而,一味地增加网络的深度不仅不会提高模型的分割精度,反而会出现梯度消失的问题,网络也不可避免地发生退化,从而阻碍网络模型的训练,此时层数少的网络在性能上反而优于层数多的网络。此外,网络模型层数的加深通常会使得模型的复杂度变高,因此容易导致网络发生过拟合。本实验组借鉴深度残差网络模型^[18-19]中的残差块结构,将结构化的 dropout 方法(即 DropBlock^[20])引入进来,设计了改进的残差块结构,具体如图 2 所示。在所提改进残差块结构中,首先使用批量归一化(BN)对传递来的数据进行标准化操作,使其符合正态分布,并用线性激活函数(ReLU)对输入

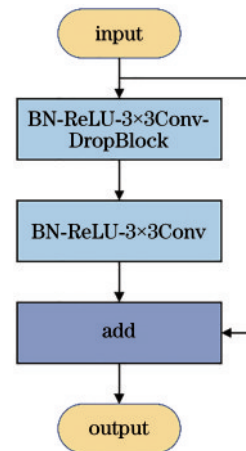


图 2 改进的残差块结构

Fig. 2 Improved residual block structure

数据进行激活,进而利用卷积和 DropBlock 操作提取有效的特征;然后再一次利用批量归一化、激活函数和卷积操作;最后对输入进行恒等映射,通过 1×1 卷积调整输入和输出的通道数量,作为改进的残差块结构的输出。新的残差结构在提高深层网络的分类性能的

同时,也增强了模型的泛化能力,有效解决卷积层的过拟合问题。

2.2 多尺度卷积模块

深层网络通过加深网络层数增加模型复杂度来更好地拟合特征。而为了使每一层网络都能够学习到更丰富的特征,在深层网络的基础上进一步提升模型性能,需要增加网络的宽度,也就是增加通道数。受 Szegedy 等^[21]提出的 Inception 模块的启发,本实验组提出了一种多尺度卷积模块,每个并行的分支分别使用 1×1 、 3×3 和两个级联的 3×3 大小的卷积核对输入特征同时进行卷积,如图 3 所示。使用两个级联的 3×3 卷积替代 5×5 卷积,是因为这两种卷积方式获得的感受野相同,而小的卷积核能够减少模型训练中的参数。此外,不同于 Inception 模块将每个分支得到的特征图进行拼接融合,所设计的多尺度卷积模块则将每个分支得到的特征图按照对应位置像素进行相加聚合,能够增加每一维的特征信息。因此,将这些多尺度特征进行融合,来获取更加丰富全面的特征信息,提升网络模型的识别和分类能力。

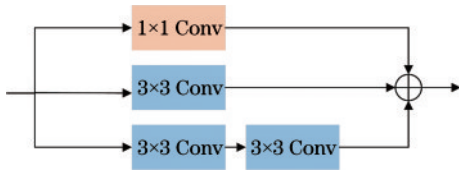


图 3 多尺度卷积模块

Fig. 3 Multi-scale convolution module

2.3 池化方法

经典的卷积神经网络在处理分类识别任务时不具备平移不变性^[22],文献^[23]指出:目标物体即使只发生微小位移变换,其对应的输出都可能产生剧烈抖动,原因在于特征提取过程未遵守采样定理,使得高频信号和低频信号发生混叠,进而导致平移不变性的丢失。因此,为了提高网络模型的识别和分类能力,有必要在特征提取时保留平移不变性。MaxBlurPool^[24]在最大池化的基础上引入低通滤波器,对进行区域最大值计算后的特征图使用模糊核进行加权平均,解决下采样带来的平移不变性丢失问题。本实验组使用 MaxBlurPool 作为网络模型的池化操作,能够有效缓解下采样造成的锯齿现象,减轻由于输入发生微小位移而导致输出发生剧烈波动的现象,从而尽可能地保证网络的平移不变性,增强了模型的鲁棒性。

2.4 并行空洞卷积模块

眼底血管的结构复杂多样,形态大小各不相同。在分割任务中,检测大血管需要大的感受野,而检测小血管则需要小的感受野,因此相同卷积操作得到的感受野并不适用于所有血管的分割。因此,本实验组提出了并行空洞卷积模块,如图 4 所示。该模块将网络上一层获取的特征图输入四个并行的空洞卷积核^[25]

中,然后将每个空洞卷积得到的特征图进行相加整合,其中每个空洞卷积都使用 1×1 大小的卷积核,不但可以随时调整特征图的维度,而且可以在不改变特征图大小的前提下增加非线性特性。此外,由上到下依次将空洞卷积的膨胀率(dr)设置为 2、4、8、16,获得不同大小的感受野,进而提取血管图像的多尺度特征信息。最后,特征图的相加整合使得描述图像的每个特征的信息增多。并行空洞卷积模块通过设置不同的膨胀率,在保持分辨率的同时任意增大感受野,进而获得图像的多尺度特征,捕获更加丰富的上下文信息,并且空洞卷积不会引入多余参数来增加计算量。

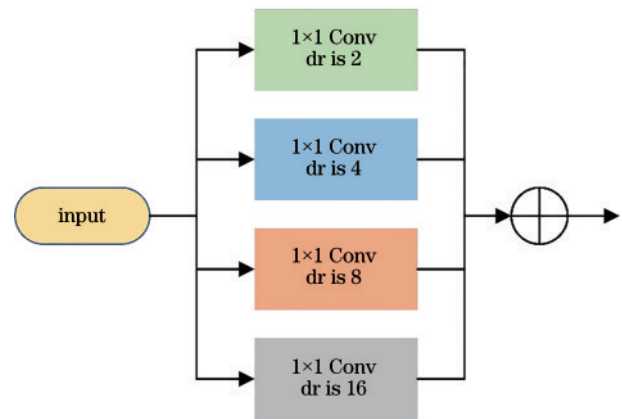


图 4 并行空洞卷积模块

Fig. 4 Parallel dilated convolution module

2.5 注意力模块

注意力机制模拟人类的视觉感知系统,将有限的注意力分配给感兴趣区域,帮助网络模型快速地获取图像的关键信息,进而提升分割任务的工作效率。注意力机制具有以下优点:引入较少的额外参数,减少计算量;操作简单快捷,即插即用;提高分割精度,视觉效果好。

2.5.1 多尺度注意力模块

眼底图像中细小、微弱的血管难以检测和识别,因此在分割过程中容易出现微小血管断裂的现象。所提多尺度注意力模块使用不同大小的卷积核提取血管图像的多尺度特征,从而获取血管在多种形态下的特征信息,为网络的上采样提供更加丰富的底层细节特征,提高网络对不同尺寸血管的适应能力,学习眼底图像中血管复杂多变的结构信息,更好地实现准确分割。多尺度注意力模块的结构如图 5 所示,首先将特征图输入三个并行分支中,每个分支分别采用 1×1 、 3×3 、两个级联 3×3 的卷积操作提取多尺度的特征信息。其次,在每次卷积提取特征后,每个分支都会进行一次 1×1 的卷积操作,用来调整特征的通道数目,降低额外参数带来的计算复杂度。然后,将每个分支提取到的相关性强的特征通过相加进行聚合,弱化不相关特征带来的影响,同时加快网络的收敛速度。最后,利用

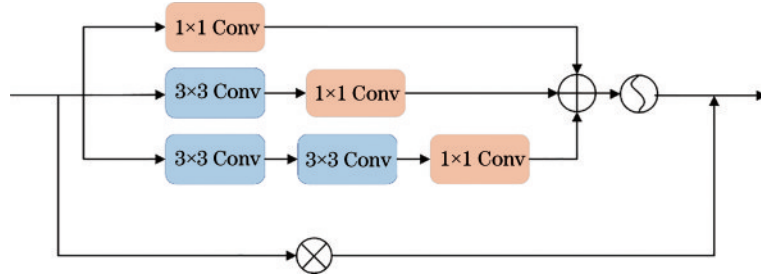


图 5 多尺度注意力模块

Fig. 5 Multi-scale attention module

sigmoid 激活函数计算权重系数,并将其与输入特征相乘得到多尺度注意力图。

2.5.2 混合注意力模块

除了在这一层利用多尺度注意力模块,本实验组还在编码部分的最后一层设计了一种混合注意力模块,该模块依次沿着通道维度和空间维度获得特征的注意力图,如图 6 所示。该模块首先将特征送入有效通道注意力(ECA)模块,得到通道注意力图,然后将其作为空间注意力模块(SA)模块的输入,提取空间维度的注意力信息。ECA 模块在对输入特征进行全局

平均池化后,使用一维卷积替换全连接层进行权值共享,并且将每一组的权重设置相等,有效降低了参数量^[26]。SA 模块^[27]中,输入特征图沿通道轴同时进行最大池化和平均池化,每个操作都会得到一幅一维的特征图。接着,把这两幅特征图按照通道进行拼接,然后进行一个 7×7 的卷积操作并利用 sigmoid 函数计算权重因子。最后将输入特征图和该权重因子相乘得到注意力图。混合注意力模块结合通道注意力和空间注意力的优势,获取更加丰富的注意力信息,同时抑制背景区域的干扰,提升模型的分割能力。

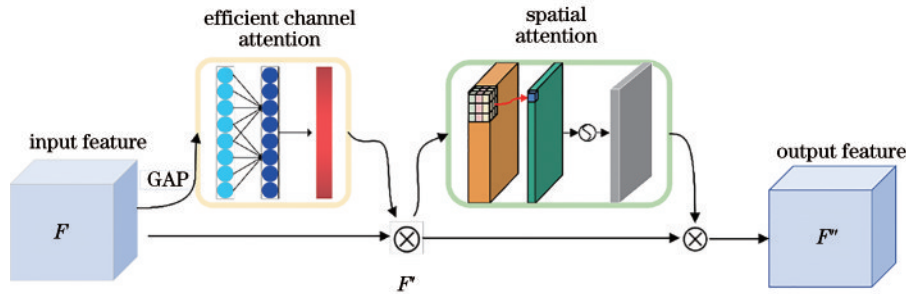


图 6 混合注意力模块

Fig. 6 Hybrid attention module

2.6 改进的跳跃连接

在解码阶段,本实验组改进了跳跃连接方式。首先,AG 模块^[28]能够更好地结合眼底血管图像的底层细节信息和高层语义信息,获取图像更丰富的特征信息。同时,AG 模块能够通过抑制背景噪声增强网络模型对目标血管的敏感度。然后,将 AG 模块得到的特征图输入多尺度注意力模块,对特征信息进行多尺度变换,以获取不同感受野下的特征信息,便于进一步提取目标血管的形态、大小等外观特征。由 AG 模块和多尺度注意力模块共同作用的跳跃连接方式,能够得到眼底血管图像更为完整的上下文信息,提高网络模型的像素分类能力,进而提升分割准确度。

3 实验过程与结果分析

3.1 数据集与实验环境

实验选用 digital retinal images for vessel extraction (DRIVE)^[29] 和 child heart and health study in England

(CHASE DB1)^[30] 两个公开的数据集。DRIVE 数据集来自荷兰的一个糖尿病视网膜病变筛查项目,该项目对 400 多名年龄分布在 25 岁~90 岁的糖尿病患者进行筛查。DRIVE 数据集由 40 幅彩色眼底图像组成,并且有 7 幅图像存在早期糖尿病性的视网膜病变,每幅图像的分辨率大小为 $565 \text{ pixel} \times 584 \text{ pixel}$ 。此外,将这 40 幅图像划分为 20 幅训练图像和 20 幅测试图像,同时每幅彩色眼底图像都有与之相对应的专家手动分割的结果。

CHASE DB1 数据集共有 28 幅彩色眼底图像,分别从 14 个儿童的左眼和右眼获取数据,每幅图像的分辨率大小为 $999 \text{ pixel} \times 960 \text{ pixel}$,每幅彩色眼底图像都有与其对应的专家分割结果作为参考。此外,由于官方并未给出具体的划分情况,根据 DRIVE 数据集的划分方式,将 CHASE DB1 数据集分为 20 幅训练图像和 8 幅测试图像。

实验使用以 TensorFlow 为端口的 Keras 深度学习

框架进行模型的搭建。实验硬件设备为 Intel E5-2640 v3 CPU 和 NVIDIA GeForce RTX 2080 Ti GPU, 软件环境为 Windows 10 操作系统。

3.2 数据预处理

彩色眼底图像中血管的颜色及形态和背景区域相似, 对比度不明显, 难以进行区分, 因此极易造成分类错误。同时, 除了自身存在的内部噪声之外, 图像的成像装置及采集数据的设备也不可避免地引入外部噪声。因此, 在进行视网膜血管分割之前, 有必要对血管进行增强处理, 突出血管的特征信息。

使用对比度受限自适应直方图均衡化 (CLAHE) 算法和伽马变换对目标血管图像进行增强处理。首先, 对彩色视网膜图像进行灰度变化, 然后利用 CLAHE 算法将图像划分成多个小区域, 并对每个小区域的对比度进行限制, 该操作不仅可以提高目标血管和背景之间的对比度, 还能够限制噪声同幅度增强。经过 CLAHE 算法处理的眼底图像整体得到初步增强, 但仍存在灰度过高或过低的不均衡区域, 于是使用伽马变换进行修正, 再次突显目标血管区域。最后, 对图像进行归一化处理, 统一数据样本的分布情况, 加速网络的收敛。图 7 展示了原始图像和预处理后的图像。

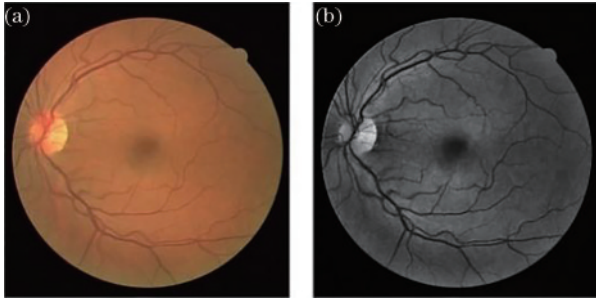


图 7 图像预处理。(a)DRIVE 数据集原始图像;(b)预处理后的图像

Fig. 7 Image preprocessing. (a) Original image of DRIVE dataset; (b) pre-processed image

3.3 评价指标

视网膜血管分割是一项像素级别的语义分割任务, 目的是将血管和背景进行分类, 本质上属于二分类问题。评价视网膜血管的分割效果, 除了通过人眼观察获得主观感受之外, 还需要评价指标对模型进行客观评判。

评估模型时, 需要将模型分割结果与其对应的真实标签进行比较, 像素的比较结果分为如下情况: 真阳性 (TP) 代表血管像素被正确判定为血管; 真阴性 (TN) 代表背景像素被正确判定为背景; 假阴性 (FN) 代表将血管像素错误判定为背景; 假阳性 (FP) 代表将背景像素错误判定为血管。选用的评价指标有灵敏度 (SEN)、特异性 (SPE)、准确率 (ACC)、F1 分数 (F1), 具体计算公式如表 1 所示。灵敏度计算实际血管像素

被判断为血管所占的比例, 用来衡量血管像素的识别能力; 特异性计算实际背景像素被判断为背景所占的比例, 用来衡量背景像素的识别能力; 准确率计算正确分类的样本占所有样本的比例, 用来衡量模型的分割精度; F1 用来平衡灵敏度和精确度的表现, 能够更全面准确地评价模型的性能。为了进一步衡量模型的性能, 也使用受试者工作特征 (ROC) 曲线下方的面积 (AUC) 作为评价指标, 用来直观衡量模型分类能力。当 AUC 值为 1 时, 就意味着实现了完美分割。

表 1 评价指标的公式

Table 1 Formula of evaluation index

Index	Formula
SEN	$R_{SEN} = \frac{N_{TP}}{N_{TP} + N_{FN}}$
SPE	$R_{SPE} = \frac{N_{TN}}{N_{TN} + N_{FP}}$
F1-score	$S_{F1-score} = \frac{2N_{TP}}{2N_{TP} + N_{FN} + N_{FP}}$
ACC	$R_{ACC} = \frac{N_{TN} + N_{TP}}{N_{TP} + N_{FP} + N_{TN} + N_{FN}}$

3.4 数据扩充与训练细节

实验所用数据集的原始训练样本较少, 因此需要进行数据扩充来避免模型出现过拟合现象。使用的扩充方式包括加入高斯噪声、进行翻转变换、色彩转换, 然后从扩充后的数据中挑选 10% 作为验证集, 用来随时监测模型的训练情况。接着, 采用正态分布对网络的权重进行随机初始化, 训练时使用二值交叉熵作为损失函数, 并选用自适应矩阵估计算法 (Adam) 对模型进行优化。将模型训练次数设置为 50 时能够达到最优的训练效果, 具体操作是将前 40 轮训练的学习率设置为 0.001, 后 10 轮训练的学习率设置为 0.0001, 批归一化操作的动量设置为 0.9。此外, 为了更好地实现正则化效果, 使用 DropBlock 方法, 并且将丢弃块的大小设置为 7。不同之处在于, 训练 DRIVE 数据集时将丢包率设置为 0.20, 而训练 CHASE DB1 数据集时将丢包率设置为 0.12, 以此达到最佳的分割效果。

3.5 实验结果与分析

3.5.1 分割效果对比与细节展示

将所提基于注意力机制的多尺度残差 U 型网络用于公共眼底图像数据的分割实验中, 并使用数据集中的标准图来对比分析网络的分割效果。图 8 给出了不同算法在 DRIVE 数据集和 CHASE DB1 数据集 (第 1、2 行图像来源于 DRIVE 数据集, 第 3、4 行图像来源于 CHASE DB1 数据集) 上的整体分割效果图, 并且这些图像相对应的细节分割效果如图 9 所示。

针对第 1 幅图像, 从图 8、图 9 可以看出: 文献 [28] 的算法出现交叉处血管错误分割的现象, 同时部分细小血管的信息没有分割完整; 文献 [12] 的算法均出现

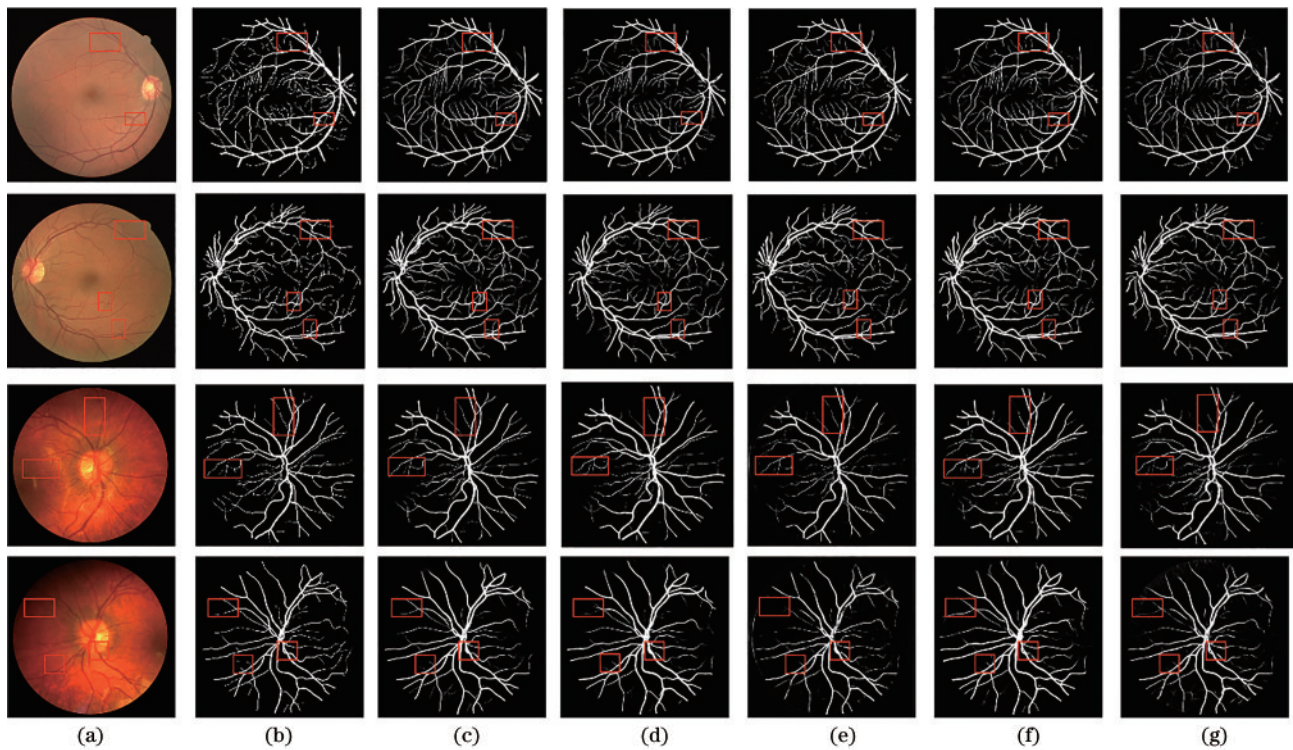


图 8 不同算法视网膜血管分割结果。(a)原始图像;(b)标准图;(c)所提算法;(d) Residual U-Net^[12];(e) Recurrent U-Net^[12];(f) R2U-Net^[12];(g)文献[28]中的算法

Fig. 8 Retinal vessel segmentation results of different algorithms. (a) Original images; (b) ground truth; (c) proposed algorithm; (d) Residual U-Net^[12]; (e) Recurrent U-Net^[12]; (f) R2U-Net^[12]; (g) algorithm in reference [28]

交叉处血管分割断裂的现象,并且缺失较多细节信息;所提算法能够较好地分割交叉处血管,并较为完整地保留细小血管。

针对第 2 幅图像,从图 8、图 9 可以看出:Residual U-Net^[12]难以保留分支处细小血管的全部信息,且会在交叉连接处发生断裂;其余对比算法则会出现无法分出或错误分割分支处细小血管的现象;所提算法能够实现分支处细小血管的连续分割。

针对第 3 幅图像,从图 8、图 9 可以看出:文献[12]除 R2U-Net 外的其余算法均在血管交叉连接处出现隔断分割现象,并且对细小血管分割不足;文献[28]的算法则在交叉连接处发生断裂的同时,会出现错误分割细小血管的现象;所提算法和 R2U-Net^[12]都能够较为准确地分割交叉处血管及细小血管。

针对第 4 幅图像,从图 8、图 9 可以看出:Residual U-Net^[12]能较好地分割细小血管,但会出现交叉处隔断的现象;Recurrent U-Net^[12]在血管交叉处分割不连续的同时,会发生细小血管的错分;R2U-Net^[12]和文献[28]的算法在部分交叉点处出现细微断裂,同时有待提高细小血管完整信息的保留能力;所提算法在完整保留细小血管的同时,有效避免了视盘区域附近由于光照不均匀造成的粗血管连接处分割断裂的问题。

在交叉分支处血管出现分割断裂现象,这是因为特征提取能力不够强,并且在网络学习过程中随着层数的增加,会面临梯度消失的问题,同时池化操作会导

致局部细节信息丢失。相较而言,所提算法使用“改进的残差块结构+多尺度卷积模块”,在提高特征提取能力的同时,有效避免了网络加深带来的梯度消失的问题,并且在网络的编码部分引入多尺度注意力模块和混合注意力模块来降低光照不均、噪声等外界因素带来的干扰。同时,改进的池化方法尽可能地保留血管局部细节信息,并增强了模型的鲁棒性。因此,所提算法能够较好解决交叉分支处的血管断裂问题。

在细小血管的分割表现出局限性,这是因为提取的细节信息较少,并且网络细节之间的融合能力不够强。所提算法在编码部分使用并行空洞卷积模块,在解码部分引入带有注意力机制的改进的跳跃连接方式,在捕捉多尺度血管信息的同时能够有效恢复和融合编码部分的细节信息,所以对细节信息的分割效果较好。

综上所述,所提算法不仅能够将血管和背景分开,还可以实现粗血管的完整分割。特别地,所提算法能够有效解决交叉处血管分割断裂的问题,并且能够更好地保留细小血管的完整信息,相较于对比算法具有一定的稳定性优势。

3.5.2 单一模块的作用验证实验

所提算法是在原始 U-Net^[31]上进行一系列改进而来的,可对每个模块进行验证实验,以此来展示所提的每一个模块给 U-Net 带来的影响。1)原始 U-Net 网络,记为 M1; 2)将原始 U-Net 中所有层的第 1 个卷积

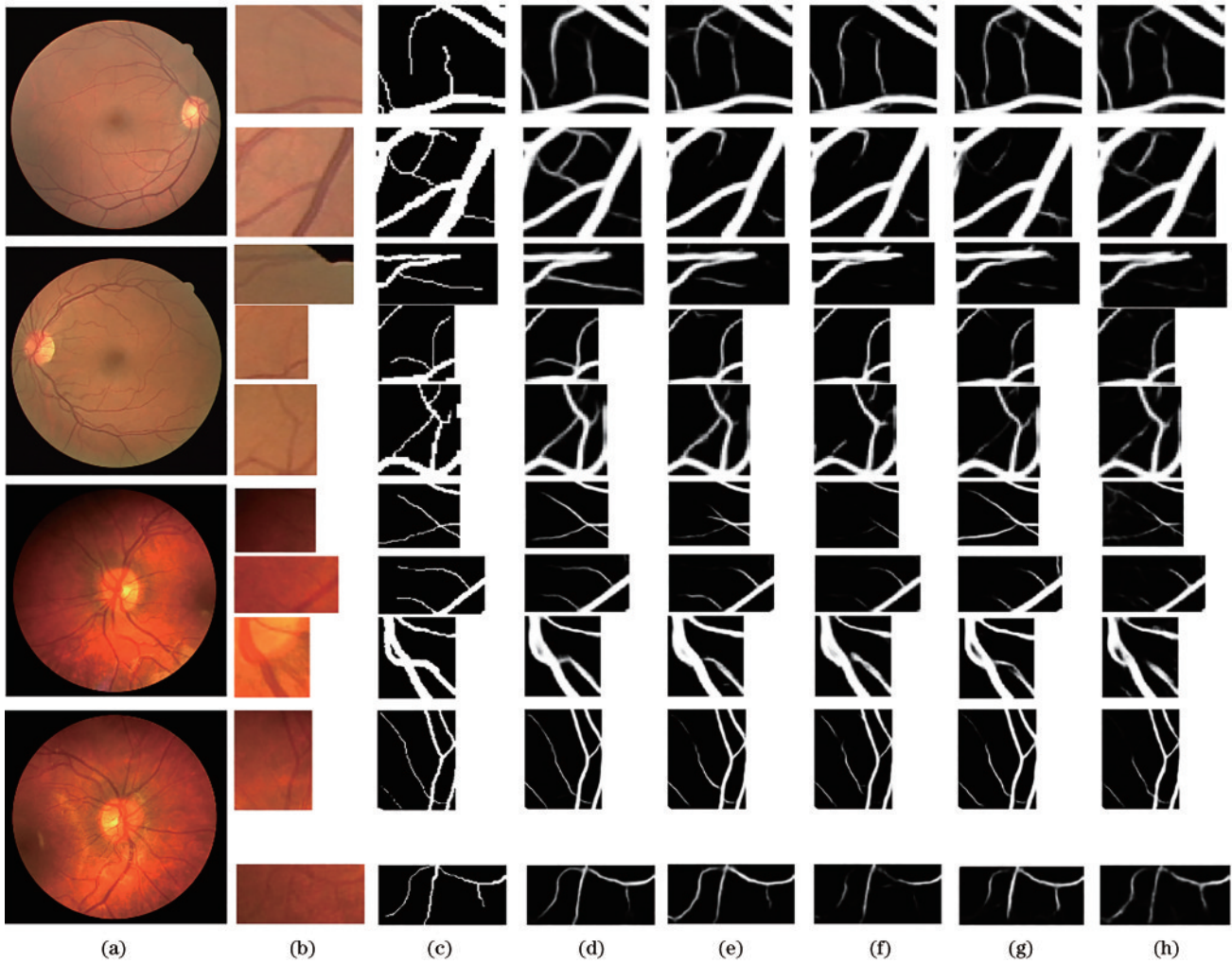


图 9 分割结果细节对比。(a)原始图像;(b)原始图像细节;(c)标准图细节;(d)本文算法细节;(e)Residual U-Net^[12]细节;
(f)Recurrent U-Net^[12]细节;(g)R2U-Net^[12]细节;(h)文献 [28] 算法细节

Fig. 9 Detail comparison of segmentation results. (a) Original image; (b) details of original images; (c) details of ground truth; (d) details of proposed algorithm; (e) details of Residual U-Net^[12]; (f) details of Recurrent U-Net^[12]; (g) details of R2U-Net^[12]; (h) details of algorithm reference [28]

操作替换为“改进的残差块结构+多尺度卷积模块”，把第 2 个卷积操作替换为“改进的残差块结构+多尺度注意力模块”，记为 M2;3)在原始 U-Net 中使用改进的跳跃方式，记为 M3;4)在原始 U-Net 的下采样操作的最后一个编码层引入混合注意力模块和并行空洞卷积模块，记为 M4。

图 10 给出单一模块分别作用在 U-Net 上的视觉分割效果,表 2 给出相应的评价指标。从图 10 的局部区域分割效果可以看出,每个模块都能够帮助 U-Net 提高交叉处血管或细小血管的分割能力,进而提升眼底血管图像的分割性能。表 2 中的结果同样表明,每个模块都能够提高 U-Net 的评价指标。

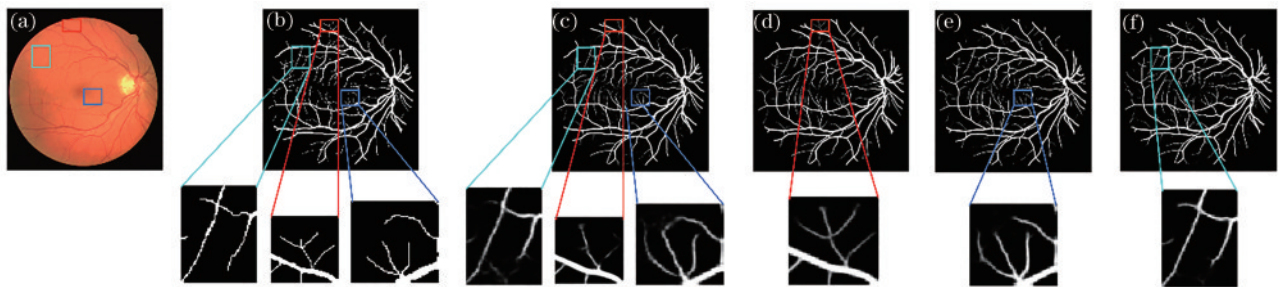


图 10 单一模块的作用验证。(a)原始图像;(b)标准图;(c)M1;(d)M2;(e)M3;(f)M4

Fig. 10 Verification of role of a single module. (a) Original images; (b) Ground truth; (c) M1; (d) M2; (e) M3; (f) M4

表 2 单一模块的作用验证实验

Table 2 Verification experiment of role of a single module

Method	SEN	SPE	F1	ACC	AUC
M1	0.7736	0.9822	0.7901	0.9640	0.9754
M2	0.7638	0.9891	0.8141	0.9694	0.9849
M3	0.7830	0.9870	0.8158	0.9690	0.9845
M4	0.7757	0.9873	0.8114	0.9684	0.9829

3.5.3 多模块累积作用验证实验

单一模块的作用验证实验表明,所提算法所使用的每一个模块都能够提高原始 U-Net 的分割性能。为了直观感受各个模块共同作用的效果,设计了多模块累积作用验证实验。1)原始 U-Net 网络,记为 N1; 2)将原始 U-Net 的第 1 个卷积层替换为“改进的残差块结构+多尺度卷积模块”,第 2 个卷积层替换为“改进的残差块结构+多尺度注意力模块”,记为 N2;3)基于 N2,采用改进的跳跃连接方式进行同级特征融合,记为 N3;4)基于 N3,在编码部分的最后一层的多尺度卷积模块之后引入混合注意力模块,记为 N4;5)基于 N4,在编码部分的最后一层的多尺度注意力模块之后引入并行空洞卷积,记为 N5;6)基于 N5,使用 MaxBlurPool 作为池化方法,也就是所提算法。

在进行多模块累积验证实验时,当前测试使用的网

络都是在它前一个网络的基础上通过加入不同的模块进行构造的,测试结果如表 3 所示。从表 3 可以看出,在原始 U-Net 的基础上,每叠加一个单一模块,评价指标整体上都会有相应的提高。在所有模块的共同作用下,所提算法的分割效果明显,取得了理想的分割结果。

表 3 多模块累积作用验证实验

Table 3 Multi-module cumulative effect verification experiment

Method	SEN	SPE	F1	ACC	AUC
N1	0.7736	0.9822	0.7901	0.9640	0.9754
N2	0.7638	0.9891	0.8141	0.9694	0.9849
N3	0.8059	0.9861	0.8265	0.9703	0.9865
N4	0.8188	0.9863	0.8285	0.9704	0.9869
N5	0.8017	0.9868	0.8268	0.9706	0.9872
Proposed method	0.8267	0.9851	0.8308	0.9707	0.9876

3.5.4 混合注意力模块的验证实验

在编码部分的最后一层引入混合注意力模块,包含空间注意力模块和通道注意力模块,为了确定分割效果更好的组合方式,进行了 2 次验证实验:1)先使用空间注意力模块,再使用通道注意力模块;2)先使用通道注意力模块,再使用空间注意力模块。实验结果如表 4 所示。从表 4 可以看出,先通道后空间的混合方式对应的评价指标整体上更高,分割效果更好。

表 4 混合注意力模块的验证实验

Table 4 Validation experiment of hybrid attention module

Model	SEN	SPE	F1	ACC	AUC
Spatial-channel attention	0.8245	0.9844	0.8299	0.9704	0.9871
Channel-spatial attention	0.8267	0.9851	0.8308	0.9707	0.9876

3.5.5 不同分割算法对比分析

为了验证所提算法对彩色眼底血管图像的分割性能,表 5、表 6 给出了所提算法与对比算法在 DRIVE 和 CHASE DB1 数据集上的性能比较。在 DRIVE 数据

集分割结果中,所提算法的 SEN 值为 0.8267,SPE 值为 0.9851,F1 分数为 0.8308,ACC 值为 0.9707,AUC 值为 0.9876;在 CHASE DB1 数据集的结果中,所提算法的 SEN 值为 0.8520,SPE 值为 0.9850,F1 分数为

表 5 DRIVE 数据集眼底血管分割结果

Table 5 DRIVE dataset fundus blood vessel segmentation results

Type	Method	Year	SEN	SPE	F1	ACC	AUC
Unsupervised method	Reference [6]	2010	0.7120	0.9724		0.9382	
	Reference [5]	2014	0.6280	0.9840		0.9380	
	Reference [7]	2019	0.7030	0.9850		0.9510	
Supervised method	Residual U-Net ^[12]	2018	0.7726	0.9820	0.8149	0.9553	0.9779
	Recurrent U-Net ^[12]	2018	0.7751	0.9816	0.8155	0.9556	0.9782
	R2U-Net ^[12]	2018	0.7792	0.9813	0.8171	0.9556	0.9784
	Reference [28]	2018	0.7730	0.9823	0.8148	0.9676	0.9725
	Reference [13]	2018	0.7844	0.9819		0.9567	0.9807
	Reference [14]	2019	0.8038	0.9802		0.9578	0.9821
	Reference [15]	2019	0.8100	0.9848		0.9692	0.9856
	Reference [16]	2020	0.8062	0.9769		0.9547	0.9739
	Reference [32]	2020	0.7651	0.9818		0.9547	0.9750
	Proposed method	2021	0.8267	0.9851	0.8308	0.9707	0.9876

表 6 CHASE DB1 数据集眼底血管分割结果
Table 6 CHASE DB1 dataset fundus blood vessel segmentation results

Type	Method	Year	SEN	SPE	F1	ACC	AUC
Unsupervised method	Reference[33]	2015	0.7201	0.9824		0.9530	0.9532
	Reference[34]	2018	0.7555	0.9807		0.9521	
Supervised method	Residual U-Net ^[12]	2018	0.7726	0.9820	0.7800	0.9553	0.9779
	Recurrent U-Net ^[12]	2018	0.7459	0.9836	0.7810	0.9622	0.9803
	R2U-Net ^[12]	2018	0.7756	0.9820	0.7928	0.9634	0.9815
	Reference[28]	2018	0.7820	0.9850	0.8012	0.9680	0.9819
	Reference[13]	2018	0.7538	0.9847		0.9637	0.9825
	Reference[14]	2019	0.8132	0.9814		0.9661	0.9860
	Reference[15]	2019	0.8186	0.9848		0.9743	0.9863
	Reference[16]	2020	0.8135	0.9762		0.9617	0.9782
	Reference[35]	2020	0.8477	0.9825	0.8652	0.9643	0.9448
Proposed method	2021	0.8520	0.9850	0.8201	0.9765	0.9911	

0.8201, ACC 值为 0.9765, AUC 值为 0.9911。从表 5、表 6 可以直观感受到,所提算法在 DRIVE 和 CHASE DB1 数据集上的所有评价指标值都明显高于基于无监督学习的分割方法,即监督学习方法解决眼底血管图像问题的有效性。同时,将所提算法与现有的监督学习方法进行了比较,仅在 CHASE DB1 数据集上的 F1 的值低于文献[35]提出的方法,这是因为文献[35]使用的动态数据扩充方式在训练时进行数据增强,相较于所提算法在训练前进行数据扩充的方式,能够在一定程度降低血管的误检率。所提算法改善了 ACC 值和 AUC 值,在整体上具有最佳的分割效果,这意味着所提算法是一种有效的眼底血管图像分割方法。

4 结 论

提出了一种基于注意力机制的多尺度 U 型网络,对眼底血管图像实现了端到端的分割,该网络分为编码部分和解码部分。在编码部分,改进的残差块结构能够提取图像的深度特征并有效解决因数据量少而导致的过拟合问题。在残差块结构后接着引入多尺度卷积模块用来进一步获取多尺度特征,引入多尺度注意力模块有助于网络关注不同感受野下的特征信息,从而给血管区域分配更多的注意力,弱化背景区域的干扰。接着,使用 MaxBlurPool 进行池化,使得下采样操作在完成数据降维的同时保持平移不变性,增强平移鲁棒性。然后,在最后一层引入混合注意力模块分别从通道维度和空间维度获取关键特征信息,引入并行空洞卷积能够在提取多尺度特征的同时随时调整通道数目,并且不会带来额外参数增加计算量。在解码部分,改进原始的跳跃连接方式,将同一阶段编码部分的下采样结果和上一级的反卷积结果使用 AG 模块进行融合,在保留更多编码信息的同时能够抑制噪声的干扰,接着,将 AG 模块得到的特征图输入到多尺度注意力模块进一步获取关键特征的多尺度信息。改进的跳

跃连接方式不仅能够获取眼底血管图像丰富的上下文信息,而且能够有效抑制噪声的干扰。在 DRIVE 和 CHASE DB1 数据集的实验结果表明,所提算法具有一定的稳健性。不但可以对粗血管和大血管实现完整分割,而且对细小血管也同样可以取得较好的分割结果。特别地,所提算法重点关注眼底血管图像的多尺度特征信息,因此网络可以较好地分割细小血管和微弱血管,并且尽可能地保持血管的连通性,减少出现血管分割断裂的次数。下一步将提高网络的泛化能力,完成多种医学图像的分割任务。

参 考 文 献

- [1] Fleming A D, Philip S, Goatman K A, et al. Automated microaneurysm detection using local contrast normalization and local vessel detection[J]. IEEE Transactions on Medical Imaging, 2006, 25(9): 1223-1232.
- [2] Cheung C Y L, Zheng Y F, Hsu W, et al. Retinal vascular tortuosity, blood pressure, and cardiovascular risk factors[J]. Ophthalmology, 2011, 118(5): 812-818.
- [3] Mukaida M, Okami Y, Suetake N, et al. Contour line extraction of vein with use of special tracking strategy and fuzzy inference for arteriosclerosis diagnosis of retinal blood vessel[C]//ECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society, October 18-21, 2020, Singapore. New York: IEEE Press, 2020: 435-440.
- [4] 田丰, 李莹, 王静. 基于多尺度小波变换融合的视网膜血管分割[J]. 光学学报, 2021, 41(4): 0410001.
Tian F, Li Y, Wang J. Retinal blood vessel segmentation based on multi-scale wavelet transform fusion[J]. Acta Optica Sinica, 2021, 41(4): 0410001.
- [5] Emary E, Zawbaa H M, Hassanien A E, et al. Retinal vessel segmentation based on possibilistic fuzzy c-means clustering optimised with cuckoo search[C]//2014 International Joint Conference on Neural Networks (IJCNN), July 6-11, 2014, Beijing, China. New York: IEEE Press, 2014: 1792-1796.
- [6] Zhang B, Zhang L, Zhang L, et al. Retinal vessel

- extraction by matched filter with first-order derivative of Gaussian[J]. *Computers in Biology and Medicine*, 2010, 40(4): 438-445.
- [7] Dash J, Bhoi N. Retinal blood vessel extraction using morphological operators and Kirsch's template[M]// Wang J C, Reddy R M, Prasad V K, et al. *Soft computing and signal processing. Advances in intelligent systems and computing*. Singapore: Springer, 2019, 900: 603-611.
- [8] Karthika D, Marimuthu A. Retinal image analysis using contourlet transform and multistructure elements morphology by reconstruction[C]//2014 World Congress on Computing and Communication Technologies, February 27-March 1, 2014, Trichirappalli, India. New York: IEEE Press, 2014: 54-59.
- [9] Osareh A, Shadgar B. Automatic blood vessel segmentation in color images of retina[J]. *Iranian Journal of Science and Technology, Transaction B: Engineering*, 2009, 33(2): 191-206.
- [10] Akram M U, Khalid S, Tariq A, et al. Detection and classification of retinal lesions for grading of diabetic retinopathy[J]. *Computers in Biology and Medicine*, 2014, 45(1): 161-171.
- [11] 张赛, 李艳萍. 基于改进 HED 网络的视网膜血管图像分割[J]. *光学学报*, 2020, 40(6): 0610002.
Zhang S, Li Y P. Retinal vascular image segmentation based on improved HED network[J]. *Acta Optica Sinica*, 2020, 40(6): 0610002.
- [12] Alom M Z, Hasan M, Yakopcic C, et al. Recurrent residual convolutional neural network based on U-Net (R2U-Net) for medical image segmentation[EB/OL]. (2018-02-19)[2018-01-18]. <https://arxiv.org/abs/1802.06955>.
- [13] Wu Y C, Xia Y, Song Y, et al. Multiscale network followed network model for retinal vessel segmentation [M]//Frangi A F, Schnabel J A, Davatzikos C, et al. *Medical image computing and computer assisted intervention-MICCAI 2018. Lecture notes in computer science*. Cham: Springer, 2018, 11071: 119-126.
- [14] Wu Y C, Xia Y, Song Y, et al. Vessel-net: retinal vessel segmentation under multi-path supervision[M]// Shen D G, Liu T M, Peters T M, et al. *Medical image computing and computer assisted intervention-MICCAI 2019. Lecture notes in computer science*. Cham: Springer, 2019, 11764: 264-272.
- [15] Zhang S H, Fu H Z, Yan Y G, et al. Attention guided network for retinal image segmentation[M]//Shen D G, Liu T M, Peters T M, et al. *Medical image computing and computer assisted intervention-MICCAI 2019. Lecture notes in computer science*. Cham: Springer, 2019, 11764: 797-805.
- [16] 吕晓文, 邵枫, 熊义明, 等. 基于双流网络的视网膜血管分割方法[J]. *光学学报*, 2020, 40(4): 0410002.
Lü X W, Shao F, Xiong Y M, et al. Retinal vessel segmentation method based on two-stream networks[J]. *Acta Optica Sinica*, 2020, 40(4): 0410002.
- [17] Li D, Dharmawan D A, Ng B P, et al. Residual U-net for retinal vessel segmentation[C]//2019 IEEE International Conference on Image Processing, September 22-25, 2019, Taipei, Taiwan, China. New York: IEEE Press, 2019: 1425-1429.
- [18] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [19] He K M, Zhang X Y, Ren S Q, et al. Identity mappings in deep residual networks[M]//Leibe B, Matas J, Sebe N, et al. *Computer vision-ECCV 2016. Lecture notes in computer science*. Cham: Springer, 2016, 9908: 630-645.
- [20] Ghiasi G, Lin T Y, Le Q V. DropBlock: a regularization method for convolutional networks[EB/OL]. (2018-10-30)[2021-03-06]. <https://arxiv.org/abs/1810.12890>.
- [21] Szegedy C, Liu W, Jia Y Q, et al. Going deeper with convolutions[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition, June 7-12, 2015, Boston, MA. New York: IEEE Press, 2015: 15523970.
- [22] Kauderer-Abrams E. Quantifying translation-invariance in convolutional neural networks[EB/OL]. (2017-12-10)[2021-02-03]. <https://arxiv.org/abs/1801.01450>.
- [23] Aharon A, Yair W. Why do deep convolutional networks generalize so poorly to small image transformations? [EB/OL]. (2018-05-30)[2021-03-05]. <https://arxiv.org/abs/1805.12177>.
- [24] Zhang R. Making convolutional networks shift-invariant again[EB/OL]. (2019-04-25)[2021-02-03]. <https://arxiv.org/abs/1904.11486>.
- [25] Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions[J]. (2015-11-23)[2021-03-06]. <https://arxiv.org/abs/1511.07122>.
- [26] Wang Q L, Wu B G, Zhu P F, et al. ECA-net: efficient channel attention for deep convolutional neural networks [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 11531-11539.
- [27] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. *Computer vision-ECCV 2018. Lecture notes in computer science*. Cham: Springer, 2018, 11211: 3-19.
- [28] Oktay O, Schlemper J, Folgoc L L, et al. Attention U-Net: learning where to look for the pancreas[EB/OL]. (2018-08-20)[2020-01-02]. <https://arxiv.org/abs/1804.03999>.
- [29] Staal J, Abramoff M D, Niemeijer M, et al. Ridge-based vessel segmentation in color images of the retina[J]. *IEEE Transactions on Medical Imaging*, 2004, 23(4): 501-509.
- [30] 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.
- [31] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W M, et al. *Medical image*

- computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.
- [32] Kromm C, Rohr K. Inception capsule network for retinal blood vessel segmentation and centerline extraction[C]// 2020 IEEE 17th International Symposium on Biomedical Imaging, April 3-7, 2020, Iowa City, IA, USA. New York: IEEE Press, 2020: 1223-1226.
- [33] Roychowdhury S, Koozekanani D D, Parhi K K. Blood vessel segmentation of fundus images by major vessel extraction and subimage classification[J]. IEEE Journal of Biomedical and Health Informatics, 2015, 19(3): 1118-1128.
- [34] Dash J, Bhoi N. An unsupervised approach for extraction of blood vessels from fundus images[J]. Journal of Digital Imaging, 2018, 31(6): 857-868.
- [35] Mathews M R, Anzar S M, Kalesh Krishnan R, et al. EfficientNet for retinal blood vessel segmentation[C]// 2020 3rd International Conference on Signal Processing and Information Security (ICSPIS), November 25-26, 2020, DUBAI, United Arab Emirates. New York: IEEE Press, 2020: 20350335.