

基于多尺度注意力的视网膜血管分割方法研究

朱思凡 朱国胜

湖北大学计算机与信息工程学院 武汉 430062

(1049120243@qq.com)

摘要 在医学图像分割中,视网膜血管分割对于眼科疾病的早期诊断与治疗是很重要的。视网膜血管分割不仅有助于诊断糖尿病视网膜病变、青光眼、动脉硬化等疾病,还在分析眼部血管形态、血流动力学等方面具有广泛的应用。但是现有方法在处理视网膜细小血管和血管边缘时还无法精确分割,在类别不平衡、血管形态复杂性和有限训练样本等方面仍然受到限制。为了提高血管分割精度并降低误判率,提出了一种基于多尺度注意力的视网膜血管分割模型(MDAF-Net)。该模型通过引入多尺度动态卷积来自适应地调整对不同尺度血管的关注度,缓解了细小血管提取不足的问题,结合通道和空间注意力机制优化特征融合,增强了模型对细节特征的提取能力,采用多尺度特征融合策略,提升了在血管形态复杂性下的分割效果。MDAF-Net在DRIVE和CHASE_DB1数据集上验证模型效果,得到Dice系数为0.764、MIoU为78.3%(DRIVE)和Dice系数为0.820、MIoU为82.5%(CHASE_DB1)。实验结果表明,MDAF-Net在分割精度和假阳性率控制方面具有显著优势,解决了传统方法在细小血管分割、类别不平衡和假阳性等方面的局限。

关键词: 多尺度;动态卷积;注意力融合;特征提取;视网膜血管分割

中图分类号 TP391

Retinal Vessel Segmentation Based on Multi-scale Attention

ZHU Sifan and ZHU Guosheng

School of Computer and Information Engineering, Hubei University, Wuhan 430062, China

Abstract In medical image segmentation, retinal vessel segmentation is very important for the early diagnosis and treatment of ophthalmic diseases. Retinal vessel segmentation is not only helpful for the diagnosis of diseases such as diabetic retinopathy, glaucoma, and arteriosclerosis, but also has wide applications in analyzing ocular vascular morphology and hemodynamics. However, existing methods cannot accurately segment small retinal blood vessels and blood vessel edges, and are still limited in terms of class imbalance, complexity of blood vessel morphology, and limited training samples. In order to improve the accuracy of blood vessel segmentation and reduce the false positive rate, this paper proposes a retinal vessel segmentation model based on multi-scale attention (MDAF-Net). The model introduces multi-scale dynamic convolution to adaptively adjust the attention to blood vessels of different scales, alleviates the problem of insufficient extraction of small blood vessels, combines channel and spatial attention mechanisms to optimize feature fusion, enhances the model's ability to extract detailed features, and adopts a multi-scale feature fusion strategy to improve the segmentation effect under the complexity of blood vessel morphology. MDAF-Net verifies the model effect on the DRIVE and CHASE_DB1 datasets, and obtains a Dice coefficient of 0.764 and an MIoU of 78.3% (DRIVE) and a Dice coefficient of 0.820 and an MIoU of 82.5% (CHASE_DB1). The experimental results show that MDAF-Net has significant advantages in segmentation accuracy and false positive rate control, and solves the limitations of traditional methods in small blood vessel segmentation, category imbalance and false positives.

Keywords Multi-scale, Dynamic convolution, Attention fusion, Feature extraction, Retinal vessel segmentation

1 引言

随着全球人口老龄化趋势的加剧,糖尿病和高血压等慢性疾病的发病率逐年增加^[1]。在临床诊断中,视网膜血管的形态学变化是多种眼科疾病早期的重要诊断依据^[2],例如糖尿病视网膜病变、青光眼和黄斑变性等。

视网膜血管在维持眼内稳态和保证视觉功能方面起着关键作用^[3],眼科医生通过血管分割,可以更加高效、准确地对患者的视网膜图像进行分析,发现潜在的病变区域,从而进行针对性的治疗和干预,来辅助进行疾病诊断。视网膜的眼底图像如图1所示。

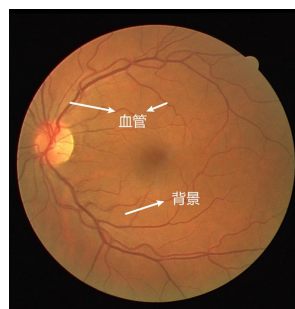


图1 视网膜眼底图像

Fig.1 Retinal fundus image

由图 1 可以看出,图像中分布着视网膜血管树,其主要由动脉和静脉等构成,还可大致看出视网膜血管的直径、分支角度以及长度等信息。

视网膜血管分割是医学图像处理中的一个重要分支,具体任务是视眼底图像中精确分割血管。

1.1 传统分割方法

早期的视网膜血管分割是由专业的医生对视网膜图像进行分析和处理,手动分割血管。但受到视网膜血管图片的尺度变化复杂、对比度低以及数据量不足的限制等,视网膜血管分割面临着巨大的挑战^[4]。这也导致了不同医生的分割结果不一致,并且手动分割的过程也相当耗时耗力。

随着技术的发展,研究人员将眼底图像进行技术上的处理,主要依赖于图像处理,使其能更加高效直观准确地分割出血管。例如 Xu 等^[5]提出了一种多层阈值技术进行精确的血管分割,通过将图像灰度值分成多个层次,采用多个阈值来区分不同的血管区域。然而,这类方法过于依赖手工设计的规则,无法有效处理视网膜图像中复杂的血管结构、不同尺度的血管、噪声和光照不均等问题。

近年来,一些改进的图像处理方法,如自适应阈值技术和多尺度图像增强方法相继被提出,以应对上述挑战,但依旧无法与基于深度学习的方法相抗衡。

1.2 基于机器学习的分割方法

基于机器学习的视网膜血管分割包括无监督方法和有监督方法。无监督方法不依赖于人工分割的结果,而有监督方法则需要依赖标注数据进行训练。无监督方法包括聚类分析^[6],匹配滤波^[7]等,通过自动提取图像中的特征进行分割。然而,由于无法从手工标注数据中获取有效的信息,其应用于视网膜血管分割的效果并不理想。而有监督方法包括支持向量机^[8]和随机森林^[9]等,能有效地利用标注数据进行训练,但一旦特征较为复杂,提取的效果就会较差,因此在分割任务中的泛化能力不足。尤其是将其应用在细小血管和低对比度图像的处理上,这类方法的表现依然有限。

近几年来,随着深度学习技术的发展,机器学习方法在视网膜血管分割中的应用已逐渐被深度学习方法所替代。因为深度学习能够利用神经网络自动从数据中学习特征,并在大规模标注数据的支持下取得更好的分割效果。

1.3 基于深度学习的分割方法

近年来,利用人工智能(AI)技术进行医学图像分割的方法受到了广泛关注^[10],深度学习方法通过使用大量的数据来自动学习特征,不再依赖于人工推理^[11]。Krizhevsky 等^[12]提出的 AlexNet 首次用于图像处理领域并取得突破,标志着深度学习开始在图像处理领域应用。Long 等^[13]提出 FCN 全卷积网络,将深度学习首次应用于语义分割,推动了医学图像分割的发展。

Ronneberger 等^[14]提出 U-Net 结构,在医疗图像分割领域获得了广泛应用,并被证明在视网膜血管分割任务中也取得了良好效果。Oktay^[15]提出了 Attention U-Net,将注意力融入 U-Net 中,增强了对有效特征的关注,抑制了冗余信息的干扰。

尽管 U-Net 及其改进网络在视网膜血管分割中取得了显著进展,但由于视网膜血管结构的复杂性,例如细小血管和

低对比度区域,现存的 U-Net 架构仍然无法对血管进行准确分割。简单的采样操作会导致上下文信息缺乏,而跳跃连接并没有将丢失的信息补充,因此单一的 U-Net 架构无法有效对视网膜血管的特征进行提取,容易产生误分割。

1.4 现有血管分割先进方法

1.4.1 动态可变形卷积

动态可变形卷积是一种具有创新性的卷积操作,利用偏移量动态调整卷积核形状,使其能够灵活地适应输入图像中的不同输入。与传统卷积不同,动态可变形卷积的卷积核可以在空间上进行位移,以适应图像中的不规则物体形态,将其应用在分割细小血管和复杂结构时效果更佳。

近年来,动态可变形卷积(DDConv)被应用于视网膜血管分割任务,研究表明,DDConv 能够有效增强对不规则血管形态的捕捉能力,提高细小血管分割的精度。例如, Yang 等^[16]提出了一种基于 DDConv 的网络架构,用于改善低对比度图像和细小血管区域的分割,实验结果表明,该方法在多个数据集上的分割效果均优于传统的 U-Net 模型。

1.4.2 多尺度特征提取

多尺度特征提取是指在图像的不同空间尺度下进行特征提取。其中视网膜血管的形态多样,血管的尺度差异大,有粗大的主血管和细小的分支血管。多尺度特征提取通过对不同尺度的血管特征进行处理,增强了模型捕捉不同尺寸血管的能力。结合多尺度信息的网络架构后,能够在保证局部细节的同时,提取出血管的全局结构。其中 Szegedy 等^[17-18]提出的 Inception 模块,使用不同大小的卷积核进行图像多尺度特征的提取。

许多改进的网络结构也引入了多尺度信息融合方法,这些网络通过在不同深度的卷积层中引入多尺度特征,进一步优化了血管分割的效果。例如, Zhou 等^[19]提出了一种基于多尺度卷积和自适应特征选择的 U-Net++ 模型,能够显著提高对不同尺度血管的分割精度。

此外,金字塔池化(Pyramid Pooling)和特征金字塔网络(FPN)等多尺度融合方法,也广泛应用于医学图像分割任务,能处理复杂背景和细小血管。例如, Liu 等^[20]提出了一种基于 FPN 和自注意力机制的血管分割模型,实验表明,该方法在细小血管的分割上明显优于传统 U-Net 模型。

在视网膜血管分割中,多尺度特征提取可以显著提升分割效果。例如, Yang 等^[21]提出了一种基于 U-Net 的分割方法,通过在编码器中引入特定的卷积模块,替换了传统的普通卷积操作。这些卷积模块能够在多个尺度上提取图像特征,相较于原始的卷积层,这种设计能够更有效捕捉不同尺度下的血管信息。Shi 等^[22]提出了 MD-Net 模型,该模型融合了密集连接和多尺度特征提取技术,不仅能同时捕捉图像中的全局特征和局部特征,还利用残差学习机制和多尺度感受野,进一步增强了网络对细节的提取能力和对复杂血管结构的识别精度,可更好地处理视网膜血管的细小部分和复杂交叉区域。以上研究表明,多尺度特征提取在视网膜血管分割中能够显著提高分割精度并增强模型的鲁棒性。

1.4.3 自适应特征融合与注意力机制

自适应特征融合方法将来自不同层级的特征进行加权融合,网络能够从不同层次的信息中提取特征。例如,

Li 等^[23]提出了一种具有多维注意和自适应特征融合的视网膜血管分割网络,在提升血管分割精度的同时,增强了模型的适应性和鲁棒性。

Guo 等^[24]引入 SA-UNet 模型进行视网膜血管分割,SA-UNet 通过合并 DropBlock^[25]和批处理归一化取代了 U-Net 框架中原始的卷积块。同时,在编码器和解码器之间集成了空间注意模块,使得网络能够自适应地关注重要的空间区域,进一步提升了血管分割的精度和鲁棒性。

自注意力机制(Self-Attention)通过计算输入特征图各部分之间的关系,强化了模型对重要区域的感知能力。例如, Golias̄ 等^[26]提出了一种具有多级自注意力的卷积网络架构应用于视网膜血管分割。Shen 等^[27]将自注意力机制与 U-Net 结合,通过整个血管图像生成全局注意力映射,将自注意力机制引入血管分割任务。

1.5 存在的挑战

1.5.1 细小血管与血管边缘的精确分割

视网膜血管的复杂性使得细小血管与血管边缘区域的分割困难,现有方法在处理这些细小区域时容易出现误差,无法提取到血管的细节,导致分割精度较低。

1.5.2 类别不平衡

血管与背景之间存在类别不平衡,细小血管在图像中所占的像素比例较小。训练过程中,模型容易忽视这些细小血管,从而影响最终分割的准确性。

1.5.3 血管形态复杂性

血管的粗细、弯曲和分支多样性增加了分割任务的难度。在复杂背景和噪声干扰下,精确分割变得更加困难。

1.5.4 有限的训练样本

现有数据集的样本数量有限,影响了模型对血管特征的充分学习,降低了其泛化能力。

1.5.5 细节分割与计算效率的平衡

视网膜血管分割对细节有着极高的要求。如何在不牺牲计算效率的前提下精确分割细小血管,仍然是一个重要的挑战。

1.5.6 多尺度特征提取的局限性

传统模型在多尺度特征提取方面存在局限,无法有效捕捉不同尺度下血管的多样特征,限制了分割效果的进一步提升。在假阳性率较高时,模型将背景区域误判为血管,增加了伪阳性的数量。

1.6 本文贡献

为了应对视网膜血管分割中的多个挑战,本文提出了一种基于多尺度注意力的视网膜血管分割模型 MDAF-Net,该模型基于 U-Net 架构设计,用于解决细小血管分割、类别不平衡和血管形态复杂性等问题。

对于细小血管与血管边缘的精确分割,MDAF-Net 引入了多尺度动态卷积模块,该模块能根据输入图像的特征自适应地调整模型对不同尺度血管的关注度。通过多分支结构结合动态感受野的方式,网络能够在捕捉粗大血管的全局信息和细小血管的局部特征时进行动态优化。这种方法可以提升细小血管和边缘区域的分割精度,降低现有方法处理细小区域的误差。

针对类别不平衡问题,MDAF-Net 在训练过程中通过

动态调整多尺度特征提取,增强模型对细小血管的关注,适应性地处理不同尺度的血管结构,提高模型对这些难以检测到的区域的敏感度,确保了在类别不平衡的情况下,细小血管分割的精确性。

对于血管形态的复杂性,MDAF-Net 采用了多尺度卷积和动态感受野调整策略,以捕捉血管的多样特征,例如应用在血管的粗细、弯曲和分支区域。多尺度特征融合技术将低层的细节特征与高层的语义特征相结合,在复杂背景和噪声干扰下也能保持较高的分割精度,从而有效应对血管形态的复杂性。

针对有限的训练样本问题,MDAF-Net 通过动态卷积和多尺度特征融合,利用现有样本的信息,增强特征提取能力,从而提高了在血管分割任务中的泛化能力。

为了解决细节分割与计算效率的平衡问题,MDAF-Net 结合了多尺度卷积和动态感受野调整,不仅确保了细小血管的精确分割,还通过优化计算流程,保持较高的计算效率,避免了计算资源的过度消耗。

针对多尺度特征提取的局限性,MDAF-Net 采用动态卷积模块,使得模型能够根据输入图像的特征自适应地调整感受野,有效避免了传统多尺度特征提取方法中的局限性,特别是在假阳性率较高的情况下,能更准确地区分血管与背景,降低伪阳性的发生。

综上所述,MDAF-Net 通过多尺度动态卷积、动态感受野调整、特征融合以及注意力机制,解决了视网膜血管分割中的多个难点,提高了细小血管分割精度。在 DRIVE 和 CHASE_DB1 数据集上的实验结果及分割结果的可视化分析表明,与其他经典分割模型相比,本文 MDAF-Net 在性能和分割效果上具有显著优势,特别是在血管分割精度和假阳性率控制方面表现优异。

2 算法原理

2.1 模型框架

图 2 展示了 MDAF-Net 的框架图。MDAF-Net 是一种基于多尺度注意力的视网膜分割模型,主干网络基于 U-Net 的编码器-解码器架构,主要针对视网膜血管分割任务进行优化。该模型主要分为多尺度特征提取模块(MDDF)和高低层注意力融合模块(HLAF)。

MDDF 模块被设计用于替代原始 U-Net 结构中的跳跃连接部分。通过结合多分支的动态可变形卷积(DDConv)、池化操作以及通道级特征融合,MDDF 模块能够高效地提取多尺度的特征。其多分支结构使得模型能更好地捕捉血管的多样性特征,在处理细小血管和复杂结构时,也能有效表达上下文信息,展现出明显的优势。为了进一步提升特征表达能力,MDDF 模块的输出通过残差连接与原始输入特征融合,增强了网络在不同尺度血管上的适应性,提升了对细节特征的捕捉能力。

HLAF 模块位于 MDDF 模块的残差连接之后,用于整合来自不同网络层的特征信息,实现高低层特征的协同优化。具体来说,HLAF 模块将解码器中上采样处理后的特征图作为低层特征,编码器经过 MDDF 模块处理后的特征图作为高层特征,并在此基础上进行全局和局部的上下文建模。该模

块分别增强低层和高层特征,使网络能够更精确地捕捉血管边缘和细节区域。在此过程中,HLAF 模块结合了通道注意力机制和特征融合操作,使高、低层特征的结合更加高效,同时关注血管的关键区域。通过高、低层特征的加权融合,HLAF 模块不仅实现了跨层信息的高效传递,还显著强化了对细小血管、模糊区域以及复杂血管结构的分割效果。

通过 MDDF 和 HLAF 模块的协同作用,MDAF-Net 在

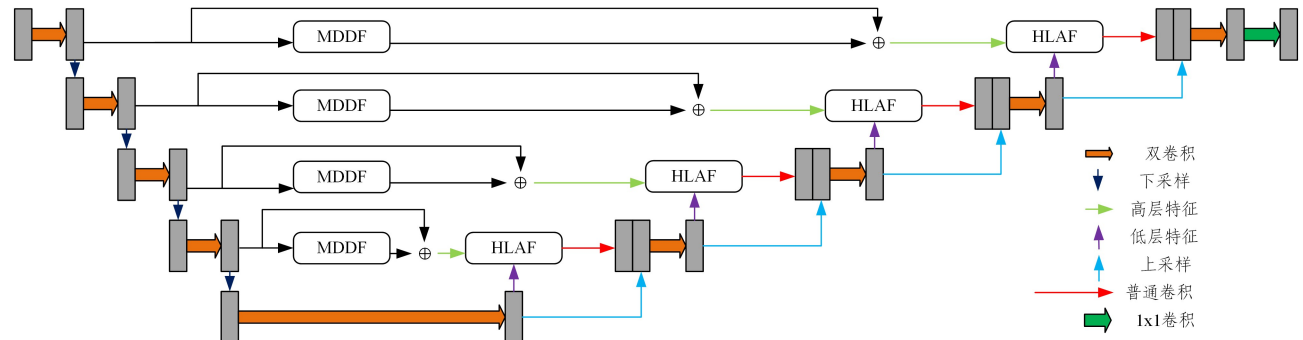


图 2 MDAF-Net 框架

Fig. 2 Framework of MDAF-Net

2.2 多尺度特征提取模块

多尺度特征提取的核心在于能够捕获特征图中不同尺度的语义信息,提取更为丰富且多样化的特征表达。在视网膜血管分割任务中,不同尺度的特征能够帮助网络同时关注细小血管的精细结构和较大血管的整体形态,提升分割的全面

性和精准性。因此,本文提出了一种多尺度动态可变形卷积特征融合模块(Multi-scale Dynamic Deformable Convolution Feature Fusion, MDDF),如图 3 所示。该模块利用多分支结构设计,结合动态感受野和多尺度特征提取策略,增强了网络对复杂形态血管的建模能力。

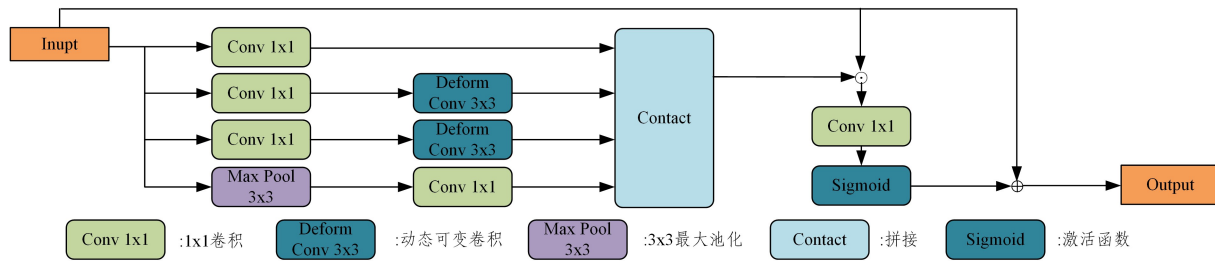


图 3 多尺度动态可变形卷积特征融合模块(MDDF)

Fig. 3 Multi-scale Dynamic Deformable Convolutional Feature Fusion Module(MDDF)

MDDF 模块的多分支结构是其实现多尺度特征提取的关键。该模块在每个分支中采用不同大小的动态感受野,以适应血管形态的多样性。例如,小感受野的分支主要用于捕获细小血管和边缘区域的特征,大感受野的分支则更关注于较大血管和全局结构的特征。这种设计通过动态调整感受野的范围,使网络能够针对不同尺度的血管进行特征提取。

分支,分别负责不同感受野的特征提取。分支 1 通过 1×1 卷积提取全局上下文信息并压缩特征维度。分支 2 采用 3×3 动态可变形卷积,结合偏移量学习机制增强局部特征建模能力。分支 3 利用 5×5 动态可变形卷积扩展感受野,以捕获大范围上下文信息。分支 4 通过 3×3 最大池化提取全局特征并通过 1×1 卷积进行压缩。将这些分支的特征图沿通道维度拼接后,与原始卷积进行点乘,再通过 1×1 卷积和 Sigmoid 激活生成注意力图,并与原始输入进行残差连接融合,以进一步强化特征表达。MDDF 模块通过动态感受野捕捉细粒度细节,同时结合多尺度特征提升对微小血管和边缘区域的分割能力。MDDF 模块中各部分的输入与输出维度变化如表 1 所列。

动态可变形卷积(DDConv)增强了 MDDF 模块的特征提取能力。由于传统卷积操作的感受野是固定的,无法适应血管的复杂形态。而 DDConv 通过动态学习位移偏移量,根据血管形态灵活调整感受野的形状和位置,捕捉血管区域的细节特征。在多分支结构中融合 DDConv,使 MDDF 模块在捕捉细小血管、分叉点和模糊区域特征方面表现更优。

传统卷积操作的卷积核以固定的感受野进行局部特征提取,其几何结构固定且对位置信息不敏感,导致难以处理具有复杂形态如视网膜血管。例如,细小血管和分叉点处的几何变化无法被固定感受野的卷积核充分捕获,限制了模型性能。为解决这一问题,利用动态可变形卷积对卷积核采样位置的动态调整,使卷积能够自适应地聚焦于目标区域的关键特征。

为了实现不同分支特征的有效整合,MDDF 模块采用特征级的融合策略。在各分支的特征提取完成后,通过通道级融合操作将不同尺度的特征进行统一处理,获得更全面的语义信息。这样,MDDF 模块不仅保留了不同分支特征的多样性,还增强了特征间的协同性,使网络在处理复杂血管结构时具有更强的表达能力和鲁棒性。MDDF 模块包含 4 个并行

动态可变形卷积通过在常规卷积的基础上引入偏移量和变形操作,实现对卷积核采样位置的动态调整,如图4所示。将输入特征图经过一个平方卷积核 K ,生成每个位置对应的偏移量。图4中 3×3 的卷积核可以表示为:

$$K = (0,0), (0,1), (0,2), \dots, (2,1), (2,2) \quad (1)$$

其中, K 表示卷积核的几何位置,这里表示为一个 3×3 的卷积核,每对括号表示卷积核的相对位置坐标。

则在 P 处的输出特征映射 Y 可以表示为:

$$Y(P_0) = \sum_{P_n \in K} K(P) \cdot X(P_0 + P_n) \quad (2)$$

其中, $Y(P_0)$ 是在位置 P_0 处的输出特征图,表示卷积操作的结果;

$K(P)$ 表示卷积核中位置 P 的权重; $X(P_0 + P_n)$ 表示输入特征图中位置 $P_0 + P_n$ 的像素值, P_n 是卷积核的位置偏移。

然后对每个输出位置,动态可变形卷积通过偏移量调整采样点位置,当偏移量为 ΔP ,则在 P 的输出特征映射 Y 变为:

$$Y(P_0) = \sum_{P_n \in K} K(P) \cdot X(P_0 + P_n + \Delta P) \quad (3)$$

对于非整数采样位置,采用双线性插值从特征图中获取特征值。根据调整后的采样点位置,动态卷积核权重与采样特征进行加权求和,最后得到一个与输入特征图大小和维度相同的特征图。

表1 MDDF 模块维度变化
Table 1 MDDF module dimension changes

| 模块 | 输入维度 | 输出维度 | 备注 |
|----------------|------------------|------------------|--------------------------------------|
| 整体输入 | (N, C, H, W) | — | 输入到整个 MDDF 模块的特征图 |
| branch1(1×1) | (N, C, H, W) | $(N, C/4, H, W)$ | 使用 1×1 卷积将通道数从 C 降维到 C/4 |
| branch2_1(1×1) | (N, C, H, W) | $(N, C/4, H, W)$ | 使用 1×1 卷积将通道数从 C 降维到 C/4 |
| branch2_offset | $(N, C/4, H, W)$ | $(N, 18, H, W)$ | 用于计算偏移量的 3×3 卷积,偏移量通道数为 18 |
| branch2_ddc | $(N, C/4, H, W)$ | $(N, C/4, H, W)$ | 动态可变形卷积(Deformable Conv)输出与输入维度一致 |
| branch3_1(1×1) | (N, C, H, W) | $(N, C/4, H, W)$ | 使用 1×1 卷积将通道数从 C 降维到 C/4 |
| branch3_offset | $(N, C/4, H, W)$ | $(N, 50, H, W)$ | 用于计算偏移量的 5×5 卷积,偏移量通道数为 50 |
| branch3_ddc | $(N, C/4, H, W)$ | $(N, C/4, H, W)$ | 动态可变形卷积(Deformable Conv)输出与输入维度一致 |
| branch4_pool | (N, C, H, W) | (N, C, H, W) | 3×3 最大池化,不改变通道数 |
| branch4_conv | (N, C, H, W) | $(N, C/4, H, W)$ | 使用 1×1 卷积将通道数从 C 降维到 C/4 |
| 特征融合(concat) | (N, C, H, W) | (N, C, H, W) | 将 4 个分支的输出在通道维度拼接后,使用 1×1 卷积恢复到 C 通道 |
| 激活(sigmoid) | (N, C, H, W) | (N, C, H, W) | 对融合后的特征图应用 Sigmoid 激活函数 |

动态可变形卷积的自适应性强,通过动态调整采样位置,卷积核能够灵活适应血管的几何形状和大小。与直接增加卷积核的数量和大小相比,其能聚焦于目标区域的关键特征,这不仅解决了固定大小卷积核自适应特征提取能力差的问题,

而且克服了不同输入共享相同卷积核参数的缺陷。对于边缘和复杂形态的细节信息,也能有效提升分割性能。偏移量的引入有效扩展了卷积操作的感受野,使模型能够捕获更丰富的上下文信息。

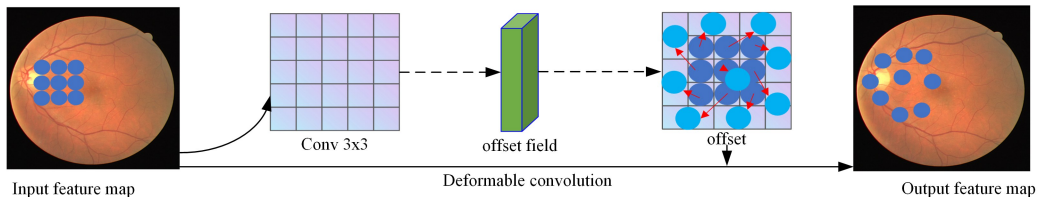


图4 动态可变形卷积的偏移量

Fig. 4 Offset of dynamic deformable convolution

2.3 高底层注意力融合模块

2.3.1 HLAFF

为了进一步提升视网膜血管分割任务中的特征融合效果,本文提出了高低层注意力融合(High Low Attention Fusion, HLAFF)模块。该模块的结构如图5所示,可以整合网络中低级特征与高级特征的信息,从而兼顾血管分割任务中的细节信息与全局语义特征。

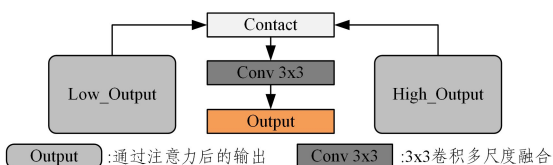


图5 注意力多尺度融合

Fig. 5 Attention multi-scale fusion

的空间细节信息,如血管边缘和纹理特征。而高级特征来源于编码器的深层特征,主要包括血管的全局结构和语义信息。HLAFF将两者进行融合,保留了空间分辨率高的细节特征,增强了对语义层次高的全局信息的理解能力。

HLAFF模块的实现包含3个关键机制,即通道注意力、空间注意力和多尺度特征融合。通道注意力机制用于动态调整特征图中不同通道的权重,根据输入特征的重要性有选择地强化血管相关的通道特征,精确聚焦于血管区域。空间注意力机制通过分析特征图在空间维度上的分布,捕捉空间位置信息,增强模型对血管边缘和模糊区域的感知能力。通过多尺度特征融合,HLAFF模块整合了低级特征与高级特征信息。

$$Output = C_{3 \times 3}(contact(O_L, O_H)) \quad (4)$$

其中, O_L 为低层特征通过通道注意力后的输出, O_H 为高层特征通过空间注意力后的输出。

HLAFF模块中各部分的输入输出维度变化如表2所列。

低级特征通常来源于解码器的浅层特征图,包含了丰富

表 2 HLAF 模块维度变化

Table 2 HLAF module dimension changes

| 模块 | 输入维度 | 输出维度 | 备注 |
|---------------------|----------------------|----------------------|--|
| 整体输入 | (N, C, H, W) | — | HLAF 模块的输入特征图, 通常来自上一层网络 |
| global_avg_pool | (N, C, H, W) | $(N, C, 1, 1)$ | 自适应全局平均池化, 将空间维 $H \times W$ 压缩为 1×1 |
| fc1 | (N, C) | $(N, C/4)$ | 全连接层, 通道数从 C 减少到 $C/4$ |
| fc2 | $(N, C/4)$ | (N, C) | 全连接层, 将通道数从 $C/4$ 恢复到 C |
| sigmoid | (N, C) | (N, C) | 对通道注意力分数应用 Sigmoid 激活函数 |
| context_block | $(N, C/2, H, W)$ | $(N, C/2, H, W)$ | 通过上下文注意力模块生成增强的特征 |
| conv_mask | $(N, C/2, H, W)$ | $(N, 1, H, W)$ | 使用 1×1 卷积生成注意力权重掩码 |
| softmax | $(N, 1, H \times W)$ | $(N, 1, H \times W)$ | 对注意力权重掩码进行归一化处理 |
| channel_mul_conv | $(N, C/2, H, W)$ | $(N, C/2, H, W)$ | 使用两层 1×1 卷积和 ReLU 激活函数生成通道增强特征 |
| channel_mul_conv(1) | $(N, C/2, H, W)$ | $(N, C/8, H, W)$ | 第一层 1×1 卷积, 通道数从 $C/2$ 降低到 $C/8$ |
| LayerNorm | $(N, C/8, 1, 1)$ | $(N, C/8, 1, 1)$ | 层归一化, 保持输出维度不变 |
| ReLU | $(N, C/8, 1, 1)$ | $(N, C/8, 1, 1)$ | 应用 ReLU 激活函数 |
| channel_mul_conv(2) | $(N, C/8, H, W)$ | $(N, C/2, H, W)$ | 第二层 1×1 卷积, 通道数从 $C/8$ 恢复到 $C/2$ |
| fusion_conv | $(N, 2C, H, W)$ | (N, C, H, W) | 使用 3×3 卷积对输入特征进行融合 |
| 整体输出 | (N, C, H, W) | — | HLAF 模块的最终输出特征图, 通道数为 C , 与空间分辨率 $H \times W$ 保持一致 |

2.3.2 通道注意力

低级特征中包含丰富的空间细节信息, 对于捕获血管细小结构尤为重要。然而, 由于低级特征缺乏显著性信息, 直接使用可能会干扰分割任务的整体性能。为此, HLAF 模块通过通道注意力机制对低级特征进行增强, 如图 6 所示。

$$O_L = F_L \otimes (\text{sigmoid}(FC(\text{ReLU}(FC(\text{GAP}(F_L)))))) \quad (5)$$

通过全局平均池化操作, 将输入的低级特征压缩为全局通道描述向量。通过两个全连接层及 ReLU 激活函数实现通道压缩和非线性变换, 并计算通道注意力权重。将通道注意力权重与低级特征逐元素相乘, 以增强重要通道, 抑制无关通道, 从而生成增强后的低级特征。

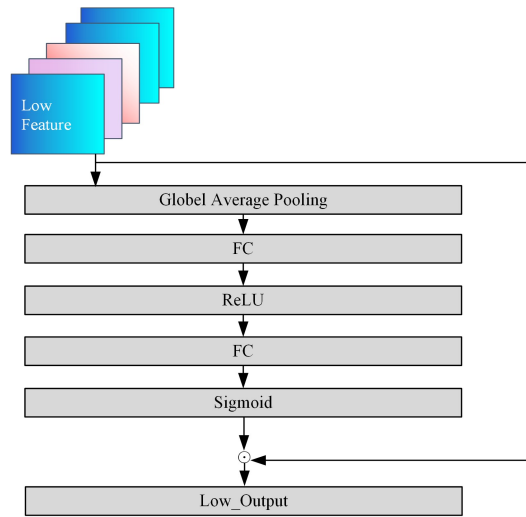


图 6 通道注意力

Fig. 6 Channel attention

2.3.3 空间注意力

高级特征通常包含丰富的语义信息, 有助于模型理解全局上下文。然而其空间细节往往不足, 直接使用可能会导致边缘模糊和细小血管的分割不准确。为解决此问题, HLAF 模块通过 ContextBlock 空间注意力机制对高级特征进行处理, 如图 7 所示。首先进行尺寸对齐, 将输入高级特征通过双线性插值上采样至与低级特征一致的空间分辨率; 再通过上下文建模, 使用 ContextBlock 模块提取高级特征的空间上下文, 生成增强后的高级特征。

ContextBlock 包含通道乘法注意力和通道加法注意力。

前者通过通道间乘法建模空间信息, 后者通过通道间加法整合上下文信息。通过组合两种方式, 模块可自适应地增强关键特征。

具体地, 高层注意力融合操作先使用一个 1×1 的卷积层和 softmax 激活函数来生成注意力权重 F_A 。

$$F_A = \text{softmax}(C_{1 \times 1}(F_L)) \quad (6)$$

再与输入进行矩阵乘法, 以获得上下文, 通过一个 1×1 的卷积层、归一化层和 ReLU 激活函数, 最后通过一个 1×1 的卷积恢复维度, 返回融合后的输出。

$$O_H = F_L \otimes (C_{1 \times 1}(\text{ReLU}(\text{LN}(C_{1 \times 1}(F_L \otimes F_A)))) \quad (7)$$

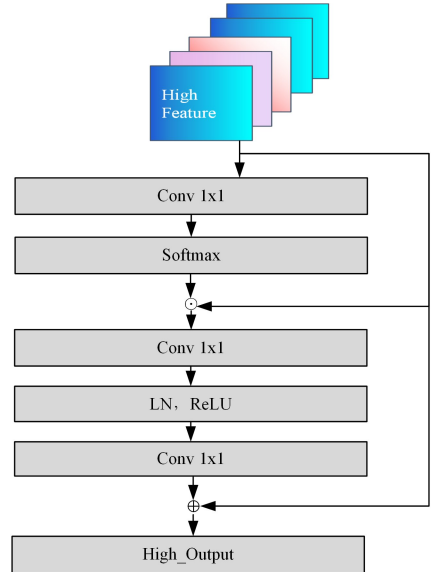


图 7 空间注意力

Fig. 7 Spatial attention

2.4 损失函数

为了提高视网膜血管分割任务中模型的表现, 实验采用了联合优化的损失机制, 结合了交叉熵损失(见式(8))和 Dice 损失(见式(9))。

$$L_{CE} = - \frac{1}{|\Omega_{roi}|} \sum_{i \in \Omega_{roi}} \log(\hat{p}_{i, y_i}) \quad (8)$$

其中, \hat{p}_{i, y_i} 是第 i 个像素对应的真实类别 y_i 的预测概率, Ω_{roi} 是感兴趣区域的像素索引集合。这样函数可通过真实标签 y 直接索引到对应类别的预测概率。

$$L_{Dice} = 1 - \frac{2 \sum_{i \in \Omega_{roi}} \hat{p}_{i,y_i} y_i + \epsilon}{\sum_{i \in \Omega} \hat{p}_{i,y_i}^2 + \epsilon \sum_{i \in \Omega_{roi}} y_i^2 + \epsilon} \quad (9)$$

其中, $\hat{p}_{i,c}$ 是预测的第 i 个像素属于类别 c 的概率; y_i 是第 i 个像素的真实标签; Ω_{roi} 是感兴趣区域, 即有效计算损失的区域, 通常是去除填充区域; ϵ 是很小的正数, 用于避免除零错误。

具体来说, 为了减少分割图像中不感兴趣区域对训练过程的影响, 交叉熵损失在计算时忽略了这些区域。Dice 损失用于衡量预测结果与真实标签之间的重叠程度。

为了进一步增强模型捕捉复杂特征的稳定性, 该损失机制引入了辅助输出损失, 如式(10)所示, 辅助输出是从模型中间层生成的预测, 通过额外的监督信号提升主分支的学习能力。最终通过主输出(见式(11)), 与辅助输出的加权和进行优化, 如式(12)所示。这一方法通过多层次的监督, 促进了模型的快速收敛, 改善了分割效果。

$$L_{main} = L_{CE}(X_{out}, y, \Omega_{roi}) + L_{Dice}(X_{out}, y, \Omega_{roi}) \quad (10)$$

其中, X_{out} 为主输出预测, y 为真实标签, Ω_{roi} 为感兴趣区域。

$$L_{aux} = L_{CE}(X_{aux}, y, \Omega_{roi}) + L_{Dice}(X_{aux}, y, \Omega_{roi}) \quad (11)$$

其中, X_{aux} 为辅助输出预测。

$$L_{total} = L_{main} + \lambda_{aux} L_{aux} \quad (12)$$

其中, λ_{aux} 为辅助损失权重。

3 实验

3.1 软硬件平台

实验采用的软硬件配置如表 3 所列, 包括操作系统、Python 版本、深度学习框架以及硬件资源等。

表 3 软硬件平台

Table 3 Hardware and Software Platform

| 类别 | 配置详情 |
|-----------|-------------------------------------|
| 操作系统 | Ubuntu 20.04 |
| Python 版本 | Python 3.8 |
| 深度学习框架 | PyTorch 1.11.0 |
| CUDA 版本 | CUDA 11.3 |
| 内存 | 80 GB |
| GPU | NVIDIA RTX 4090D(24 GB) × 1 |
| CPU | 15 vCPU Intel® Xeon® Platinum 8474C |

3.2 数据集

实验使用了 DRIVE 数据集和 CHASE_DB1 数据集进行评估, 表 4 列出了每个数据集的详细信息。考虑到视网膜血管分割任务中数据集的样本量较少, 数据量的不足会限制模型的泛化能力, 导致过拟合。为此, 在数据预处理阶段, 实验采用了一种综合的增强策略, 结合了随机增强和归一化处理, 对视网膜图像进行旋转、缩放、翻转和平移等操作, 提高了模型训练的稳定性。这些数据增强方法有效地帮助模型学习不同方向和形状的血管特征。

表 4 视网膜血管数据集

Table 4 Retinal blood vessel dataset

| DataSet | Total Images | Training Set | Testing Set |
|-----------|--------------|--------------|-------------|
| DRIVE | 40 | 20 | 20 |
| CHASE_DB1 | 28 | 22 | 6 |

实验中 DRIVE 已经划分好了训练集和测试集, 因此再对 CHASE_DB1 数据集进行了训练集和测试集的划分。考

虑到这两个数据集的样本量较小, 进一步划分验证集会致训练数据不足, 影响模型的学习效果, 因此未引入独立的验证集或交叉验证。为减轻过拟合风险, 实验采用了数据增强、正则化等技术, 并通过在测试集上的最终评估来验证模型的泛化能力。

3.2.1 DRIVE

DRIVE 数据集是一个广泛使用的视网膜血管分割数据集, 包含 40 张视网膜图像, 其中 20 张用于训练, 另外 20 张用于测试, 分辨率为 584×565 像素。每张图像均由专家进行手动标注, 提供了血管的精确分割结果。其图像包含不同的视网膜血管结构和背景噪声, 都对算法提出了较高要求。

3.2.2 CHASE_DB1

CHASE_DB1 是一个用于视网膜血管分割的数据集, 广泛应用于医学图像处理领域的研究。该数据集包含了 28 张高分辨率视网膜彩色图像, 分辨率为 1280×960 , 每张图像由两位专家手动标注, 提供了精确的血管分割标准。

3.3 实验配置

3.3.1 数据预处理

针对视网膜血管数据的特殊性, 实验对视网膜血管图片数据进行了特定优化。首先加载原始图像和手动标注图, 并生成感兴趣区域的掩码。为了消除不感兴趣区域的影响, 将这些区域的像素值设为 255, 并将其叠加到标注图上, 得到最终的分割掩码图, 其中血管前景为 1, 背景为 0, 不感兴趣区域为 255。将预处理后的图像和掩码均转换为 PIL 格式, 便于后续的数据增强。为提高训练效率并增强模型的泛化能力, 设计了批次数据打包方式, 填充较小图像, 使批次内的图像和掩码尺寸保持一致。

在数据增强方面, 通过随机尺寸调整、水平与垂直翻转以及随机裁剪等方法, 增加样本的多样性, 模拟真实场景中不同分辨率和视角的特性。对输入图像进行标准化处理, 使得数据分布具有零均值和单位方差, 降低不同通道间的特征差异, 加快梯度收敛速度。在验证阶段, 移除随机增强, 仅保留固定裁剪与归一化, 确保模型评估的结果具有一致性。

3.3.2 评价指标

对于分割性能的评估, 引入常用的评价指标, 如 Dice 系数、MIoU、敏感性 (SE)、特异性 (SP)、F1 分数和准确性 (ACC), 对模型进行性能分析, 多维度地评估模型在血管分割任务上的表现。

Dice 系数适合测量小血管结构的分割精度, 主要衡量模型分割结果与真实标签之间的重叠程度, 取值范围在 $0 \sim 1$, 值越接近 1, 说明分割结果越好, 如式(13)所示。

$$Dice = \frac{2 \cdot |A \cap B|}{|A| + |B|} \quad (13)$$

其中, A 表示模型预测的分割结果(如二值分割图中像素为 1 的区域), B 表示真实的分割标注, $|A \cap B|$ 表示预测结果和真实标签的交集区域中像素为 1 的个数。MIoU 也是重要的分割评价指标, 用于计算模型预测结果与真实标注之间的重叠率, 格外关注小区域的重叠。MIoU 是在 IoU 的基础上计算的平均值, 对于单个 IoU 的计算式如式(14)所示。

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (14)$$

其中, A 表示模型预测的分割结果区域, B 表示真实的标注区

域。MIoU 的计算式如式(15)所示。

$$MIoU = \frac{1}{C} \sum_{i=1}^C \frac{|A_i \cap B_i|}{|A_i \cup B_i|} \quad (15)$$

其中, C 表示类别总数。

SE 为灵敏度,也称为真阳性率或召回率,用于衡量血管被正确分割的比例,如式(16)所示。

$$SE = \frac{TP}{TP + FN} \quad (16)$$

SP 为特异性,用于衡量非血管区域即背景被正确分割的比例,如式(17)所示。

$$SP = \frac{TN}{TN + FP} \quad (17)$$

F1 分数是评估模型性能的综合指标,用于衡量分割结果和标准结果之间相似性,如式(18)所示。

$$F1 = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (18)$$

ACC 用于衡量视网膜血管总体分割的准确性,在像素级别中,表示整体正确预测的像素占总像素的比例,如式(19)所示。

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} \quad (19)$$

其中, TP 表示被正确分割的血管像素数量, FP 表示错误分割的血管像素数量, TN 表示正确分割的背景像素数量, FN 表示错误分割的背景像素数量。

3.4 结果分析

3.4.1 分割精度评估指标

本文选取 Dice 和 MIoU 分别用于定量评估模型输出的分割结果与真实标注之间的相似度,以及衡量分割区域的重叠程度。选取了 U-Net, U-Net++ 和 U-Net_att 作为基础模型,将 MDAF-Net 与这些基础模型进行对比。表 5 列出了在 DRIVE 数据集上,各个模型的评价指标。可以看到,与基础的模型相比,本文模型的 Dice 和 MIoU 均达到了最优。

表 5 在 DRIVE 上比较 Dice 和 MIoU

Table 5 Comparison of Dice and MIoU on DRIVE

| Model | Dice | MIoU |
|-----------|--------------|-------------|
| U-Net | 0.749 | 77.6 |
| U-Net++ | 0.752 | 77.4 |
| U-Net_att | 0.759 | 77.9 |
| Ours | 0.764 | 78.3 |

表 6 列出了在 CHASE_DB1 数据集上,本文模型与基础模型的对比。可以看到,与基础模型相比,本文模型在 Dice 指标中能达到较好的效果。

表 6 在 CHASE_DB1 上比较 Dice 和 MIoU

Table 6 Comparison of Dice and MIoU on CHASE_DB1

| Model | Dice | MIoU |
|-----------|--------------|-------------|
| U-Net | 0.814 | 82.6 |
| U-Net++ | 0.803 | 82.1 |
| U-Net_att | 0.809 | 82.2 |
| Ours | 0.820 | 82.5 |

3.4.2 模型判别能力评估指标

本文选取 SE, SP, F1 和 ACC 指标用于衡量模型在区分血管与非血管区域方面的能力,全面评估了模型的分类性能,反映了模型对血管和背景的判别能力。同样选取 U-Net, U-Net++ 和 U-Net_att 为基础模型,将 MDAF-Net 与这些基础

模型作进行对比。表 7 列出了在 DRIVE 数据集上,各个模型的评价指标。可以看到,与基础模型相比,本文模型在 SE, F1 分数和 ACC 均达到了最优。

表 7 在 DRIVE 上比较 4 个评价指标

Table 7 Comparison of four evaluation metrics on DRIVE

| Model | SE | SP | F1 | ACC |
|-----------|-------------|-------------|-------------|-------------|
| U-Net | 69.7 | 98.0 | 76.1 | 94.4 |
| U-Net++ | 72.9 | 97.2 | 75.9 | 94.1 |
| U-Net_att | 71.7 | 97.7 | 76.6 | 94.3 |
| Ours | 74.8 | 97.2 | 77.2 | 94.4 |

表 8 列出了在 CHASE_DB1 数据集上,本文模型与基础模型的对比。可以看到,与基础模型相比,本文模型在 SE, SP 和 ACC 这 3 个指标中能达到较好的效果。

表 8 在 CHASE_DB1 上比较 4 个评价指标

Table 8 Comparison of four evaluation metrics on CHASE_DB1

| Model | SE | SP | F1 | ACC |
|-----------|-------------|-------------|-------------|-------------|
| U-Net | 77.5 | 98.5 | 81.7 | 96.4 |
| U-Net++ | 76.1 | 98.2 | 81.1 | 94.4 |
| U-Net_att | 78.2 | 98.4 | 81.3 | 96.3 |
| Ours | 78.9 | 98.5 | 81.6 | 96.4 |

3.4.3 收敛性分析

训练过程中的损失函数变化如图 8 所示。从图 8 中可以看到,模型的收敛性在前期迅速下降,表明模型能够快速学习有效特征,而后期则趋于平稳,说明模型达到了最优解。

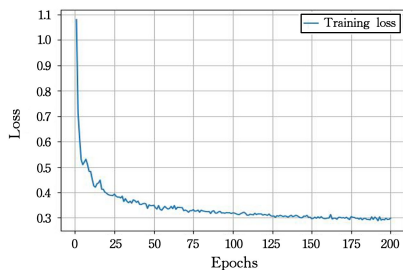


图 8 损失函数变化

Fig. 8 Loss function changes

3.4.4 学习率调度策略

在深度学习模型的训练过程中,固定的学习率往往难以适应训练不同阶段的需求,初期过大的学习率会导致训练不稳定,而在训练后期固定的较高学习率会导致错过更优解。

为了确保模型训练的有效性并促进其良好的收敛性,实验对超参数进行了优化,在学习率调度上采用了结合 Warmup 机制和非线性衰减的策略,在训练过程中逐步优化学习率。即在训练初期,学习率从一个较小的初始值线性增加至设定的学习率,Warmup 的作用在于缓解训练初期梯度爆炸的问题,确保模型能够稳定地学习有效特征。设定 Warmup $epochs=5$,即在前 5 个 epoch 中逐步增加学习率,帮助模型适应初期训练。在完成 Warmup 后,学习率逐渐减少,模型进行更精细的参数调整。在衰减过程采用指数衰减策略,非线性衰减帮助模型在接近收敛时进行更精细的调整,从而提高模型的泛化能力,降低过拟合的风险。

实验设置了不同学习率参数,以评估其对模型性能的影响。将批次大小默认设置为 1,使得模型可以更加频繁地更新参数,有助于视网膜血管分割更好地捕捉细节特征,避免过早的过拟合,增加模型的泛化能力。

表 9 列出了在 DRIVE 数据集上,使用模型 MDAF-Net,不同超参数设置下模型的性能对比。结果表明,结合 Warmup 机制和非线性衰减的学习率调度策略,不仅在训练过程中保持了良好的收敛性,而且在分割精度评估指标上取得了最佳的表现。

表 9 学习率调度策略

Table 9 Learning rate scheduling strategy

| 超参数设置 | 学习率策略 | Dice | MIoU |
|---------------------|--------------|--------------|-------------|
| 固定学习率 0.001 | 无调度 | 0.752 | 77.8 |
| 固定学习率 0.001+ Warmup | Warmup+固定学习率 | 0.758 | 77.9 |
| Warmup+非线性衰减 | Warmup+指数衰减 | 0.760 | 78.1 |
| Warmup+非线性衰减 (长周期) | Warmup+指数衰减 | 0.764 | 78.3 |

3.4.5 可视化分析

在 DRIVE 数据集中选取一个测试样本图像进行测试分割,其分割结果如图 9 所示。从得到的分割图来看,本文模型对于 DRIVE 数据集有着较好的分割效果,而且对细小血管和非血管区域的识别,相较于其他模型也有较好的效果。

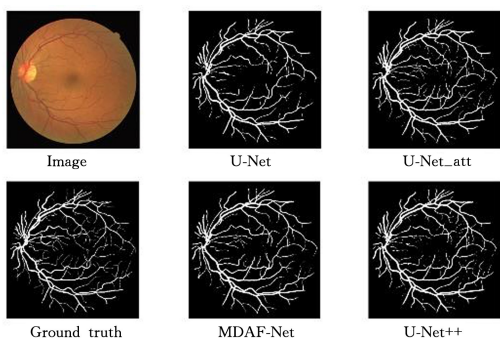


图 9 在 DRIVE 上的可视化分析

Fig. 9 Visual analytics on DRIVE

在 CHASE_DB1 数据集中选取一个测试样本图像进行测试分割,其分割结果如图 10 所示。从得到的分割图来看,本文模型对于 CHASE_DB1 数据集有着较好的分割效果,而且对细小血管和非血管区域的识别,相较于其他模型也有较好的效果。

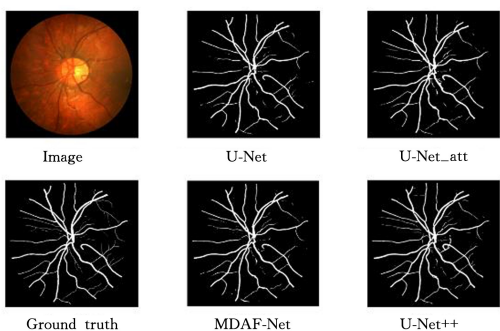


图 10 在 CHASE_DB1 上的可视化分析

Fig. 10 Visual analysis on CHASE_DB1

从模型的分割结果可以看出,假阳性主要出现在视盘周围和背景噪声较强的区域,图像表现为孤立的虚假血管线条或小片状结构。通过对训练样本的深入分析发现,这些区域的特征与血管结构(如亮度、纹理)存在一定相似性,导致了模型误判。从图 11 中可以看出,在 DRIVE 数据集中,MDAF-Net 能较好地识别出不属于血管的区域。

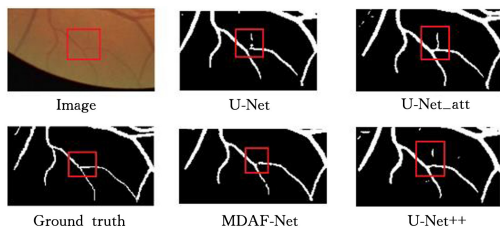


图 11 假阳性对比

Fig. 11 False positive comparison

3.5 消融实验

为了验证提出的多尺度特征提取模块(MDDF)和高低层注意力融合模块(HLAF)的有效性,本文设计了消融实验,分别针对 MDDF 模块和 HLAF 模块进行移除或替换,并分析模型在视网膜血管分割任务中的性能变化。

模型性能通过 Dice 系数和 MIoU 指标进行评估,并结合分割结果的可视化图像进行对比分析,如表 10 所列。结果表明,MDDF 模块和 HLAF 模块的结合能够进一步提升模型的分割性能。这些实验结果充分证明了所提出创新模块在视网膜血管分割任务中的有效性和适用性。

表 10 消融实验

Table 10 Ablation experiment

| 模型配置 | Dice | MIoU |
|------------------|--------------|-------------|
| 完整模型(MDAF-Net) | 0.764 | 78.3 |
| 移除 MDDF,仅使用跳跃连接 | 0.761 | 78.2 |
| MDDF 中去掉 DDCConv | 0.762 | 78.1 |
| 移除 HLAF 模块 | 0.756 | 77.9 |
| HLAF 模块使用简单拼接 | 0.759 | 77.8 |
| 同时移除 MDDF 和 HLAF | 0.755 | 77.7 |

结束语 本文算法结合 U-Net、多尺度特征提取和注意力机制,是一种新的基于 U-Net 的分割模型。使用 DRIVE 和 CHASE_DB1 数据集中的样本图像对该算法的性能进行分析。结果表明,MDAF-Net 分割算法相比与其他算法,在各方面均有较好的性能,因此能证明该算法的有效性。

本文算法在动态可变形卷积的实现中,对于偏移量的选取是固定的,可以尝试根据输入特征图的特点,例如在本实验中的血管分割中,将偏移量朝着适应血管方向进行偏移,这样可以增强卷积核对血管的特征提取能力。同时,还需要对分割后的血管进行进一步处理,从而更有利于辅助医生对疾病进行准确的判断。

参考文献

- [1] ODURO J K, OKYERE J, NYADOR J. Risky health behaviours and chronic conditions among aged persons: analysis of SAGE selected countries [J]. BMC Geriatr, 2023, 23(1): 145.
- [2] SALEH G A, BATOUTY N M, HAGGAG S, et al. The role of medical image modalities and AI in the early detection, diagnosis and grading of retinal diseases: a survey [J]. Bioengineering, 2022, 9(8): 366.
- [3] WANG S, CHEN Y, YI Z. A Multi-Scale Attention Fusion Network for Retinal Vessel Segmentation [J]. Applied Sciences, 2024, 14(7): 2955.
- [4] CHEN C, CHUAH J H, ALI R, et al. Retinal Vessel Segmentation Using Deep Learning: A Review [J]. IEEE Access, 2021, 9: 111985-112004.
- [5] XU M, CHEN S, GAO X, et al. Research on Fast Multi-Threshold

- old Image Segmentation Technique Using Histogram Analysis [J]. *Electronics*, 2023, 12(21):4446.
- [6] WISAENG K. Retinal blood vessel segmentation using density-based fuzzy C-means clustering and vessel neighborhood connected component[J]. *Measurement*, 2024, 242:116229.
- [7] KAR S S, MAITY S P. Blood vessel extraction and optic disc removal using curvelet transform and kernel fuzzy c-means [J]. *Computer in Biology and Medicine*, 2016, 70:174-189.
- [8] RICCI E, PERFETTI R. Retinal blood vessel segmentation using line operators and support vector classification [J]. *IEEE Transactions on Medical Imaging*, 2007, 26(10):1357-1365.
- [9] SKOUTA A, ELMOUFIDI A, JAI-ANDALOUSSI S, et al. Semantic Segmentation of Retinal Blood Vessels from Fundus Images by using CNN and the Random Forest Algorithm [C] // *SENSORNETS*. 2022:163-170.
- [10] KUFEL J, BARGIEL-LACZEK K, KOCOT S, et al. What is machine learning, artificial neural networks and deep learning? — Examples of practical applications in medicine[J]. *Diagnostics*, 2023, 13(15):2582.
- [11] COŞKUN M, YILDIRIM Ö, UÇAR A, et al. An overview of popular deep learning methods[J]. *European Journal of Technique*, 2017, 7(2):165-176.
- [12] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C]// *Advances in Neural Information Processing Systems*. 2012.
- [13] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015:3431-3440.
- [14] RONNEBERGERO, FISCHER P, BROXT. U-net: Convolutional networks for biomedical image segmentation[C]// *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015; 18th International Conference*. Springer, 2015:234-241.
- [15] OKTAY O. Attention u-net: Learning where to look for the pancreas[J]. *arXiv*:1804.03999, 2018.
- [16] YANG X, LI Z, GUO Y, et al. DCU-net: A deformable convolutional neural network based on cascade U-net for retinal vessel segmentation [J]. *Multimedia Tools and Applications*, 2022, 81(11):15593-15607.
- [17] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015:1-9.
- [18] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016:2818-2826.
- [19] ZHOU Z, RAHMAN SIDDIQUEE M M, TAJBAKHS N, et al. Unet++: A nested u-net architecture for medical image segmentation[C] // *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support; 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018*. Springer, 2018:3-11.
- [20] LIU Y P, RUI X, LI Z, et al. Feature pyramid U-Net for retinal vessel segmentation[J]. *IET Image Processing*, 2021, 15(8):1733-1744.
- [21] YANG D, LIU G, REN M, et al. A multi-scale feature fusion method based on u-net for retinal vessel segmentation[J]. *Entropy*, 2020, 22(8):811.
- [22] SHI Z, WANG T, HUANG Z, et al. MD-Net: A multi-scale dense network for retinal vessel segmentation [J]. *Biomedical Signal Processing and Control*, 2021, 70:102977.
- [23] LI J, GAO G, YANG L, et al. A retinal vessel segmentation network with multiple-dimension attention and adaptive feature fusion[J]. *Computers in Biology and Medicine*, 2024, 172:108315.
- [24] GUO C, SZEMENYEI M, YI Y, et al. Sa-unet: Spatial attention u-net for retinal vessel segmentation[C] // *2020 25th International Conference on Pattern Recognition (ICPR)*. IEEE, 2021:1236-1242.
- [25] GHIASI G, LIN T Y, LE Q V. Dropblock: A regularization method for convolutional networks[C]// *Advances in Neural Information Processing Systems*. 2018.
- [26] GOLIAŠ M, ŠIKUDO VÁ E. Retinal blood vessel segmentation and inpainting networks with multi-level self-attention[J]. *Biomedical Signal Processing and Control*, 2025, 102:107343.
- [27] SHEN X, XU J, JIA H, et al. Self-attentional microvessel segmentation via squeeze-excitation transformer Unet[J]. *Computerized Medical Imaging and Graphics*, 2022, 97:102055.



ZHU Sifan, born in 2000, postgraduate. His main research interests include deep learning and medical image segmentation.



ZHU Guosheng, born in 1972, Ph.D, professor, is a member of CCF (No. D4728M). His main research interests include next-generation and software-defined networks.