

基于对抗学习和引导机制的视盘和视杯联合分割

马晓月, 陈媛媛*

(四川大学计算机学院, 四川 成都 610065)

摘要: 准确的视盘(OD)和视杯(OC)分割能够有效地辅助青光眼的诊断和监测,从而进一步提高治疗效果。然而,现有方法没有考虑到眼底图像不同通道之间的差异,并且难以实现对 OC 边界的精确分割。针对这个问题,提出一种基于对抗学习和引导机制的网络框架 ALG-Net,旨在提高 OD 和 OC 的分割性能。ALG-Net 由分割网络和鉴别器两部分组成。在分割网络中,构建引导融合模块(GFM),该模块将单通道特征信息与 RGB 图像特征融合,使网络充分学习眼底图像不同通道之间的差异信息,引导分割网络聚焦于关键区域。ALG-Net 网络框架还采用了鉴别器,通过对抗学习的方式促进分割网络生成更真实的分割结果。在 REFUGE 和 Drishti-GS 数据集上进行广泛的实验评估,实验结果表明,ALG-Net 在 REFUGE 数据集上 OD 和 OC 分割的平衡精度分别达到了 98.6% 和 95.9%,在 Drishti-GS 数据集上也表现出优异的性能。此外,ALG-Net 的分割结果应用于青光眼分类任务,在 REFUGE 数据集上 ROC 曲线下面积(AUC)为 0.983,相较于经典 UNet 算法提高了 0.015,为青光眼的早期诊断和监测提供了有力的支持。

关键词: 青光眼诊断;视盘分割;视杯分割;UNet 模型;注意力机制;引导机制;对抗学习

中图分类号: TP391

文献标志码: A

DOI: 10.19678/j.issn.1000-3428.0068734

Joint Optic Disc and Optic Cup Segmentation Based on Adversarial Learning and Guidance Mechanism

MA Xiaoyue, CHEN Yuanyuan*

(School of Computer Science, Sichuan University, Chengdu 610065, Sichuan, China)

【Abstract】 Accurate segmentation of the Optic Disc(OD) and Optic Cup(OC) effectively assists in diagnosing and monitoring glaucoma, thereby improving treatment outcomes. However, existing methods do not consider the differences between the various channels of fundus images, making it challenging to achieve accurate segmentation of the OC boundary. To address this problem, a network framework based on adversarial learning and a guidance mechanism, termed ALG-Net, is proposed to improve OD and OC segmentation performance. ALG-Net comprises two main components: a segmentation network and a discriminator. The segmentation network includes a Guidance Fusion Module(GFM) designed to merge single-channel feature information with RGB image features. This allows the network to learn the differences among the various channels of the fundus image, guiding the segmentation network to focus on key regions. The framework also incorporates a discriminator, which encourages the segmentation network to generate more realistic results through adversarial learning. Extensive experimental evaluations are conducted on the REFUGE and Drishti-GS datasets. The results show that ALG-Net achieved balanced accuracies of 98.6% and 95.9% for OD and OC segmentation, respectively, on the REFUGE dataset and demonstrated better performance on the Drishti-GS dataset. In addition, applying ALG-Net's segmentation results to glaucoma classification tasks yielded an Area Under the ROC Curve(AUC) of 0.983 on the REFUGE dataset, surpassing the classic UNet algorithm by 0.015. This demonstrates ALG-Net's strong support for early diagnosis and monitoring of glaucoma.

【Key words】 glaucoma diagnosis; Optic Disc(OD) segmentation; Optic Cup(OC) segmentation; UNet model; attention mechanism; guidance mechanism; adversarial learning

0 引言

青光眼是导致失明的主要病症,也是不可逆性失明的首要因素。根据最新发布的数据,预计到

2040 年,全球青光眼患者数量将达到 1.1 亿人^[1]。面对这样的情况,青光眼的防治绝对不可忽视。及时筛查青光眼,对于早期治疗是至关重要的。在青光眼的诊断中,视盘比(CDR)是一个非常重要的判

收稿日期: 2023-10-31 修回日期: 2024-01-18

基金项目: 国家自然科学基金(62376173)。

通信作者 E-mail: *chenyuanyuan@scu.edu.cn

断依据。CDR 为垂直杯直径与垂直盘直径的比率^[2]。一般来说,CDR 越大,青光眼风险越高。因此,视盘(OD)和视杯(OC)的准确分割是 CDR 测量的关键。OD 相对于背景区域有比较清晰的边界,而 OC 区域没有明显的分界线,但亮度通常更大,通常医生只能根据经验来分割 OC 区域。然而,人工分割是非常耗时的,而且会携带主观判断,经验不足的医生很难做到精准分割。

目前,数据驱动的深度神经网络方法在 OD 和 OC 分割上取得了很好的效果。相较于传统方法基于手工特征得到分割结果,深度神经网络方法还充分考虑到图像中的语义信息。其中,全卷积网络(FCN)^[3]和 UNet 网络^[4]等网络模型彻底改变了传统的图像分割领域,提供了优异的 OD 和 OC 分割结果。然而这些方法没有考虑到眼底图像的不同通道信息之间的差异。在眼底图像中,相对于其他通道,绿色通道具有更加清晰的 OC 区域信息,更适合用于 OC 的分割^[5]。为了充分考虑到不同通道之间的信息差异,学习更丰富的特征信息,本文提出了一种基于对抗学习和引导机制的网络框架 ALG-Net 来实现更精细的分割,主要贡献如下:

1) ALG-Net 框架由分割网络和鉴别器组成。在分割网络中,为了充分学习眼底图像通道信息之间的差异,提出了引导融合模块(GFM)。GFM 将绿色通道提取到的语义信息与 RGB 图像的特征图进行融合,从而引导网络更好地关注和定位感兴趣的区域和特征。

2) 在分割网络中构建了解码器融合输出模块(DFOM),进一步促进了浅层语义与深层语义的融合,同时,采用侧边输出来构建多尺度输出层,以此提高分割性能。

3) ALG-Net 框架通过对抗学习的方式促使分割网络生成更逼真的分割结果,利用对抗学习中的判别误差和分割网络中的分割误差一起来训练网络,促使分割结果更加真实。

4) 在 REFUGE^[6]和 Drishti-GS^[7]数据集上对所提方法的性能进行了验证,结果表明 ALG-Net 框架取得了极佳的分割性能:在 REFUGE 数据集上分割 OD 和 OC 的重叠误差分别为 0.047 和 0.187;在 Drishti-GS 数据集上分割 OD 和 OC 的重叠误差分别为 0.041 和 0.149。

1 相关工作

2014 年,SHELHAMER 等^[3]提出了 FCN 模型,开启了深度神经网络方法在分割领域的应用研

究^[8-9]。该网络自提出后就成为语义分割的基本框架,后续算法基本都是由该网络框架发展而来。例如目前医学图像分割中最重要的网络模型——UNet 模型,就是基于 FCN 发展而来。

UNet 模型基于编码器-解码器结构,利用跳跃链接融合语义信息,形成更为强大的语义特征。自该模型提出以来,基于 UNet 模型的医学图像分割研究层出不穷^[10-24]。例如,一些方法^[10,15,20]将多尺度输入融入到了 UNet 模型中。多尺度输入避免参数大幅度增长的同时,增加了网络的宽度。

另一方面,有些研究人员在跳跃链接上做了改进。例如 YUAN 等^[15]将跳跃链接改进为一个上下文语义提取模块,减少了深层和浅层特征之间的语义鸿沟,促进了特征的融合。同样为了补偿由于在编码器中进行池化而丢失的空间信息,最小化特征退化带来的影响,HASAN 等^[18]和 ZHOU 等^[20]采用残差跳跃连接融合编码器和解码器的语义特征来保证分割质量。

为了充分学习医学图像中的语义特征,除了对模型的改进外,一些方法还增添一些预处理措施来辅助 UNet 的分割。如一些方法^[10,17,21,23-26]将图像转换到极坐标系下进行训练,充分利用了 OD 和 OC 之间的约束关系。XIE 等^[13]为了充分利用已有数据集,将图像转化为以边界轮廓为中心。在转化过程中通过设置不同的拟合椭圆参数,得到了多张转化后的图像,实现了基于几何参数的数据增强。同样地,为了充分利用 OD 和 OC 内部像素之间的关系,YIN 等^[14]不做预处理操作,而是将 UNet 模型与水平集结合,将网络模型的输出视为一个水平集,最终 OD/OC 分割是预测水平集函数的零水平集。这些方法针对 OD 和 OC 之间的约束关系,利用极坐标转换、水平集等方法在 UNet 模型上做了不同的改进,充分验证了 UNet 模型在医学图像分割领域的优势和潜力。

随着深度神经网络的广泛应用,人们结合不同网络模型的优势来设计分割框架。例如,文献[2,23,25]方法采用 VGG-16 作为网络主干来实现特征提取,获取图像语义信息。还有一些方法^[26-27]结合 ResNet 来优化模型,缓解网络退化问题。这些研究结合不同模型优势,使 OD 和 OC 分割更加精确。

近几年,许多研究将生成对抗网络(GAN)^[28]应用到 OD 和 OC 分割中,利用 GAN 模型对抗学习的机制,促使分割网络进一步细化分割结果。例如,BIAN 等^[29]将分割网络作为生成器,通过生成器产

生伪造标签欺骗鉴别器,而鉴别器又通过不断学习来辨别真伪,从而缓解了分割网络在低学习率下的过拟合问题,并且改善了局部分割细节,无需再做后处理。

除此之外,一些研究也做出了其他的探索。例如:刘丽霞等^[30]采用多个专家的标签,增加关于分割网络不确定性估计的研究,提升分割网络的性能;SHANKARANARAYANA 等^[9]用深度估计图像引导分割任务。与 OC 区域附近的血管弯曲类似,OC 区域附近视网膜表面的深度不连续被认为是青光眼检测的重要线索^[9]。尽管该方法需要光学相干断层成像的支持,但进一步展示出利用 OC 区域的相关信息来引导分割任务的适用性。

受深度估计图像引导分割方法^[9]启发,为了充分考虑到不同通道之间的信息差异,学习更丰富的特征信息,本文提出了一种基于对抗学习和引导机制

制的网络框架 ALG-Net。ALG-Net 在 UNet 模型基础上利用引导融合将绿色通道的特征信息与 RGB 图像的特征信息进行充分融合,再通过对抗学习机制促进分割网络生成更逼真的分割结果,实现更精细的分割。

2 基于对抗学习和引导机制的网络框架

2.1 整体流程

ALG-Net 网络架构训练流程如图 1 所示(彩色效果见《计算机工程》官网 HTML 版,下同),其中, X 表示眼底图像, Y 表示真实标签, Y' 表示分割网络输出的预测分割。首先,对眼底图像进行预处理。相对于整个眼底图像,OD 和 OC 区域所占面积较小,背景区域的面积相对较大,在网络训练的过程中容易出现过拟合现象。因此,先对 OD 和 OC 区域进行裁剪,得到感兴趣区域(ROI)。

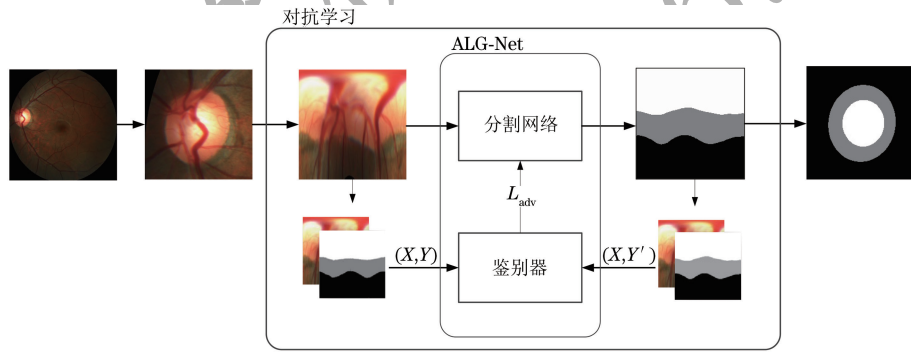


图 1 ALG-Net 训练流程

Fig. 1 The training process of ALG-Net

此外,为了更好地利用背景包含 OD、OD 包含 OC 的空间约束关系,本文将眼底图像从笛卡尔坐标系转换至极坐标系。在笛卡尔坐标系下,令 P 表示眼底图像上的一点,假设 OD 的中心是 O 。极坐标系下对应点表示为 $P'(\theta, r)$, θ 和 r 分别表示点 P 到中心 O 的方向角和半径。笛卡尔坐标系下 $P(u, v)$ 与极坐标系下 $P'(\theta, r)$ 的转换关系如式(1)所示:

$$\begin{cases} u = r \cos \theta \\ v = r \sin \theta \end{cases} \Leftrightarrow \begin{cases} r = \sqrt{u^2 + v^2} \\ \theta = \tan^{-1}(v/u) \end{cases} \quad (1)$$

2.2 ALG-Net 中的分割网络

2.2.1 网络框架

ALG-Net 中所采用的分割网络在传统 UNet 模型的基础上采用引导机制和多尺度输出加强网络对分割任务中关键信息的聚焦,缓解信息过载问题,提高 OD 和 OC 分割的效率和准确性。该网络模型结构如图 2 所示,其采用编码器-解码器模型。

在编码器阶段,分割网络采用多次卷积和下采样操作来提取绿色通道图像和 RGB 图像的语义特征,并利用 GFM 来促进特征融合,获取更丰富的语义信息,引导分割网络关注不同通道之间的差异,进一步提高对 ROI 的分割性能。

在解码器阶段,本文构建了 DFOM 和多尺度输出层。DFOM 在逐层向上恢复高级语义特征尺寸的同时,进一步减少深层和浅层特征之间的语义鸿沟,促进特征的融合。对于多尺度输出层,本文对 3 个 DFOM 的侧边输出和解码器最后一层的输出进行上采样操作得到多尺度输出结果。之后构建一个平均层来组合所有的多尺度输出预测图作为最终的预测图,得到 K 个通道概率图输出,其中 K 为类别数(在本文工作中,因为是针对 OD 和 OC 两个类别,所以 $K=2$)。多尺度输出层可以及时地将输出损失反向传播到解码器的早期层,缓解梯度消失问题,帮助早期卷积层的训练。

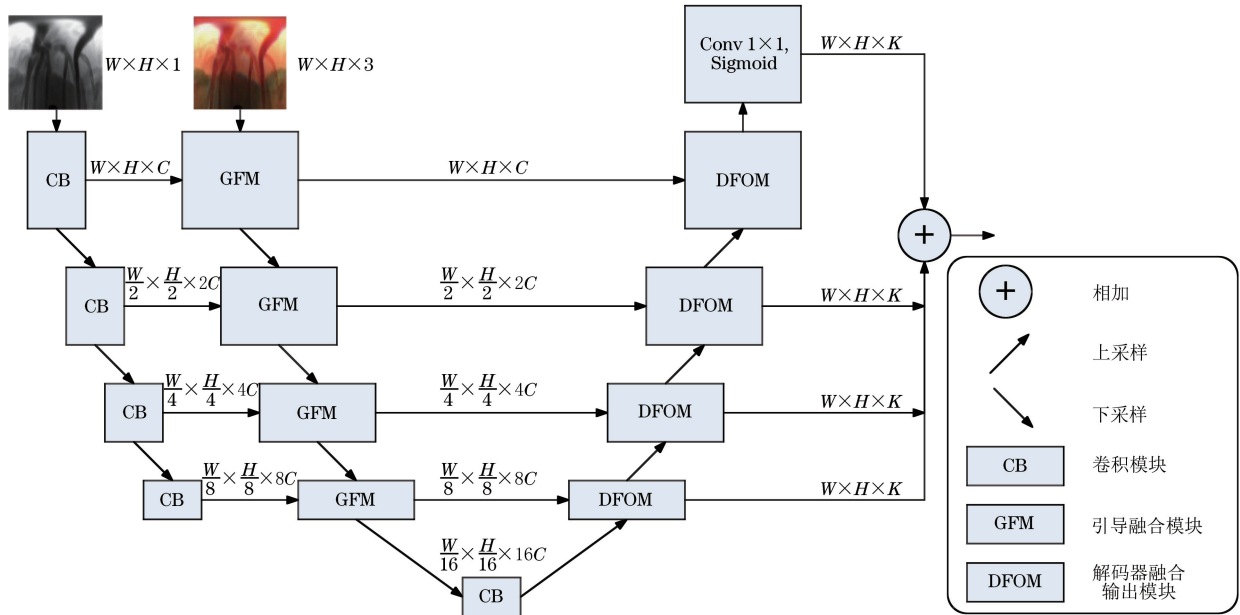


图 2 ALG-Net 中的分割网络
Fig.2 Segmentation network in ALG-Net

2.2.2 引导融合模块

已有的研究已经发现绿色通道具有更加清晰的 OC 区域信息,更适合于 OC 的分割^[5]。RGB 图像虽然已经包含绿色通道,但仍然能够深入挖掘独属于绿色通道的特征信息,并利用提取到的特征信息引导网络从 RGB 图像的众多信息中提取特定区域信息,使网络更有针对性地学习到绿色通道的差异性特征,有助于提高网络对 OC 区域的敏感度。

如图3所示,为了充分利用绿色通道中的 OC

区域信息,GFM将 U 型编码器提取到的绿色通道特征作为侧边输入,再将侧边输入经过 SimAM 注意力机制^[31]操作加强对 ROI 关键信息的学习,之后采用跳跃链接将学习到的特征信息与 RGB 图像的特征信息进行融合。为了充分利用不同通道、不同感受野之间的信息差异,GFM 将融合得到的特征信息经过两次 3×3 卷积后,再采用 1×1 、 3×3 和 5×5 三种不同的卷积核来提取特征。最后,GFM 中引入了残差链接的思想,缓解过拟合问题。

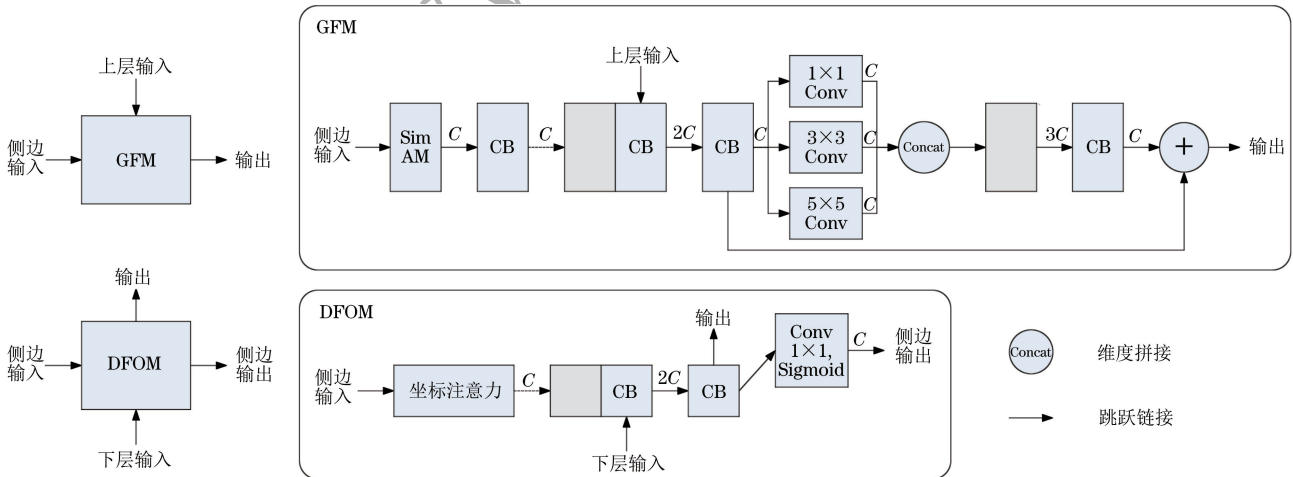


图 3 GFM 和 DFOM 结构
Fig.3 Structure of GFM and DFOM

在 GFM 中,分割网络通过学习权重对不同通道之间的信息进行调整,使得模型能够更灵活地处理不同通道之间的信息差异,保证融合后的特征图中尽可能多地包含 ROI 的特征信息,引导分割网络关注 OC 区域特征,提高 OC 分割性能。

2.2.3 解码器融合输出模块

如图3所示,DFOM将编码器的浅层语义特征作为侧边输入,侧边输入经过坐标注意力(CA)^[32]操作,获取到跨通道的特征信息以及方向感知和位置感知的特征信息,以此来增强对 ROI 的表示,提

高分割网络对 OC 区域的关注。

之后,DFOM 再采用跳跃链接与解码器的深层语义特征进行融合,进一步减少深层和浅层特征之间的语义鸿沟,促进特征的融合。融合得到的特征信息经过卷积操作得到传递给上层的输出,再通过上采样将特征信息的尺寸恢复至 $W \times H \times C$,得到本层的侧边输出。

2.2.4 注意力机制

在 GFM 和 DFOM 中,本文分别使用了 SimAM 注意力机制^[31]和坐标注意力机制^[32]让网络更好地关注和定位感兴趣的区域和特征。

1) SimAM 注意力机制。

该注意力机制无需添加额外的参数,直接为特征图推断三维注意力权重。具体来说,其通过求解能量函数解析解的方式来推断每个神经元的重要性。能量函数定义如式(2)所示:

$$e_t(\omega_t, b_t, y, x_i) = (y_T - \hat{t})^2 + \frac{1}{N-1} \sum_{i=1}^{N-1} (y_0 - \hat{x}_i)^2 \quad (2)$$

式中: $\hat{t} = \omega_t t + b_t$ 和 $\hat{x}_i = \omega_i x_i + b_i$ 分别是输入特征的单个通道中的目标神经元和其他神经元; i 是空间维度上的一个通道; $N = H \times W$ 表示该通道上的神经元数量; ω_t 和 b_t 是权重和偏置。当 $t = y_T$, $x_i = y_0$ 时,方程式(2)取得最小值。 y_T 和 y_0 是两个不同的值,为了简单起见,对 y_T 和 y_0 使用了二进制标签(即 1 和 -1),通过极小化该方程来寻找目标神经元 t 与同一通道内所有其他神经元 x_i 之间的线性可分性。每个神经元的重要性也可以通过计算能量函数求解得到。

2) 坐标注意力机制。

坐标注意力将位置信息嵌入到通道中,通过分解为两个一维特征的编码过程,得到分别沿两个不同空间方向的聚合特征^[33]。通过这种方法,可以在一个空间方向上捕获远程依赖关系,同时在另一个空间方向上保存精确的位置信息。具体网络框架如图 4 所示。整体结构可以分为坐标信息嵌入和坐标注意力生成两个模块。坐标信息嵌入模块将输入使用尺寸 $(H, 1)$ 和 $(1, W)$ 的池化核沿着水平坐标方向和竖直坐标方向对每个通道进行编码变换,之后再将这两个变换进行特征聚合,返回一对方向感知注意力图。坐标注意力生成模块将坐标信息嵌入模块生成的两个特征图进行级联,然后使用卷积进行变换,在本文的工作中,卷积操作降低倍数 $r=32$ 。最后,将卷积后的特征信息沿着空间维度切分为两个单独的张量完成卷积操作,得到最终输出。

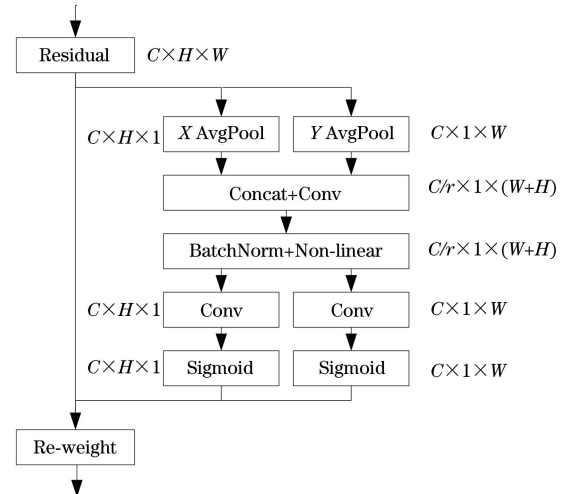


图 4 坐标注意力机制

Fig. 4 Coordinate attention mechanism

2.2.5 损失函数

本文采用的是 OD 和 OC 联合分割的方法,将 OD 和 OC 的分割任务转化为一个多标签问题来解决。由于 OC 位于 OD 内部的特殊性,OC 区域的像素点也属于 OD 区域,因此在计算输出损失时,本文将 OD 和 OC 作为两个独立的二分类器分别计算输出损失,再乘以权重求和。在本文的方法中,引入了一个基于 Dice 系数的多标签损失函数,具体如式(3)所示:

$$L_{\text{dice}} = 1 - \sum_k \frac{2\alpha_k \sum_i p_{(k,i)} g_{(k,i)}}{\sum_i p_{(k,i)}^2 + \sum_i g_{(k,i)}^2} \quad (3)$$

式中: $N = H \times W$ 代表像素数; K 代表类别数; $p_{(k,i)}$ 和 $g_{(k,i)}$ 分别代表预测分割的像素点和真实标签的像素点; α_k 代表每一类别的权重。由于 OD 和 OC 分割在任务中同等重要,因此本文设置 $\alpha_k = 0.5$ 。

ALG-Net 框架通过对抗学习的方式促使分割网络生成更逼真的分割结果,因此分割网络在训练时引入了对抗损失 L_{adv} ,具体定义如式(4)所示:

$$L_{\text{adv}} = E_{x \sim P_g} [D(x)] \quad (4)$$

式中: P_g 表示分割网络生成的预测分割; $D(x)$ 表示判别器的输出。

对于分割网络,整体损失函数由分割损失 L_{dice} 和对抗损失 L_{adv} 两部分组成,具体如式(5)所示:

$$L_{\text{seg}} = \beta_1 L_{\text{adv}} + \beta_2 L_{\text{dice}} \quad (5)$$

式中: β_1 和 β_2 表示不同损失的权重。由于在初始训练时对抗损失数值远大于分割损失,并且在训练后期对抗损失快速收敛,容易使得网络陷入局部最优,基于对任务需求和实验结果的综合考虑,设置 $\beta_1 = 0.01$, $\beta_2 = 1$ 。

2.3 ALG-Net 中的对抗学习架构

2.3.1 网络框架

在 GAN 模型中,生成器产生伪造数据欺骗鉴别器,而鉴别器又不断学习来辨别数据真伪,两者在博弈的过程中不断提升自己的技术。本文利用 GAN 模型的对抗学习机制,让鉴别器区分输入的图像是来自

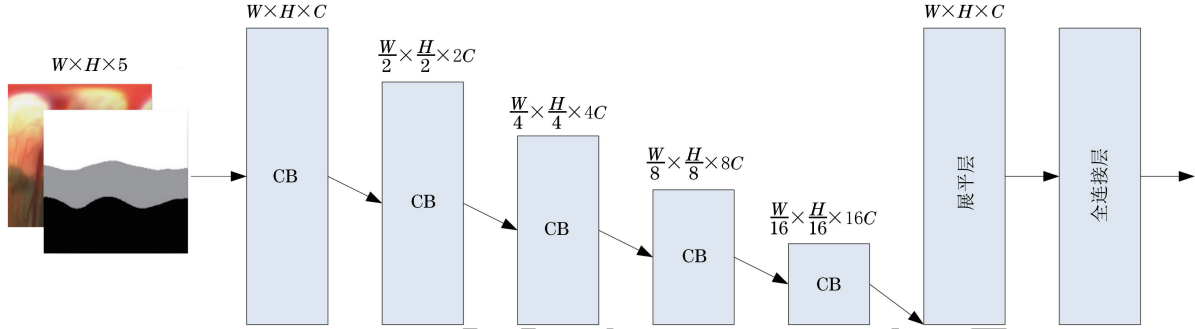


图 5 ALG-Net 中的鉴别器

Fig.5 Discriminator in ALG-Net

然而原始 GAN 模型训练不稳定,很难平衡生成器和鉴别器的训练程度,容易出现梯度消失问题。因此,本文采用 Wasserstein GAN^[34] 来进行训练,该模型用 Wasserstein 距离代替 JS 散度。Wasserstein 距离与 KL 散度、JS 散度相比,即便两个数据分布没有重叠,Wasserstein 距离仍然能够反映它们的远近。所以利用 Wasserstein 距离计算得到的损失曲线不会出现突变,而是平滑的,可以提供有意义的梯度。

2.3.2 损失函数

通过对抗学习,将模型的优化问题转化为博弈问题,交替更新分割网络和鉴别器的权重。鉴别器 D 判断输入是来自真实标签还是预测分割,其损失函数定义如式(6)所示:

$$L_D = E_{x \sim P_r} [D(x)] - E_{x \sim P_g} [D(x)] \quad (6)$$

式中: P_r 表示真实标签。

3 实验

3.1 数据集

为了验证所提方法的有效性,本文在 REFUGE^[6] 和 Drishti-GS^[7] 两个数据集上进行了训练和测试,并基于 REFUGE^[6]、Drishti-GS^[7] 和 RIGA^[35] 三个数据集来验证模型迁移学习的能力。

REFUGE 数据集是 MICCAI 2018 视网膜眼底青光眼挑战赛所发布的大型数据集,包含了 1 200 张眼底图像,其中青光眼图像 120 张,正常眼底图像 1 080 张,为 OD、OC 分割和青光眼检测提供了充分的数据支撑。

Drishti-GS 数据集包含了 101 张图像,其中青

光眼图像 70 张,正常眼底图像 31 张。每张图像的标签分别由具有 3、5、9 和 20 年经验的青光眼专家来标注,并将至少有 3 位专家达成一致的区域作为 OD 和 OC 的真实标签。

RIGA 数据集的图像来自 3 个不同的渠道 (MESSIDOR 数据集、Bin Rushed 眼科中心和 Magrabi 眼科中心),图像以 TIFF 的格式进行保存。该数据集包含了 750 张图像,每张图像都有手工标注的 OD 和 OC。

RIGA 数据集的图像来自 3 个不同的渠道 (MESSIDOR 数据集、Bin Rushed 眼科中心和 Magrabi 眼科中心),图像以 TIFF 的格式进行保存。该数据集包含了 750 张图像,每张图像都有手工标注的 OD 和 OC。

3.2 评价指标

为了评估分割性能,本文采用交叉误差 E 、平衡准确率 A 作为评价指标。具体公式如下:

$$E = 1 - \frac{|l_A \cap l_B|}{|l_A \cup l_B|} \quad (7)$$

$$A = (R + S) / 2 \quad (8)$$

$$R = \frac{n_{TP}}{n_{TP} + n_{FN}} \quad (9)$$

$$S = \frac{n_{TN}}{n_{TN} + n_{FP}} \quad (10)$$

式中: l_A 和 l_B 分别表示预测分割和真实标签; n_{TP} 和 n_{TN} 分别表示真阳性和真阴性的个数; n_{FP} 和 n_{FN} 分别表示假阳性和假阴性的个数。

3.3 实验设置

将 REFUGE 数据集中有标签的 800 张图像按照 7:3 比例进行随机划分,560 张作为训练图像,240 张作为测试图像。对 Drishti-GS 和 RIGA 数据集图像也进行同样的操作。此外,由于 RIGA 数据集中的图像只包含 ROI,无需再做图像裁剪操作,因此将图像统一调整成 500×500 像素大小。

在实验中, GPU 为 NVIDIA GTX 1650, 内存为 8 GB, 编码框架采用的是 PyTorch 1. 13. 1, CUDA 11. 7。采用 RMSprop 优化器, 学习率(lr)为 0. 000 05, Batch Size 设置为 2, Epoch 设置为 1 000。

3. 4 实验结果分析

3. 4. 1 OD 和 OC 分割实验

1) 消融实验。

为了验证 ALG-Net 框架的有效性, 在 REFUGE 数据集上对网络框架的各模块进行了消融实验, 结果如表 1 所示, 其中加粗表示最优值, 下同。将 UNet 作为基线网络, 依次将 GFM、DFOM 以及对抗学习 WGAN 融入进行验证。

表 1 模块消融实验结果

Table 1 Module ablation experiment result				
算法	E_{cup}	E_{dish}	$A_{cup}/\%$	$A_{dish}/\%$
UNet(baseline)	0. 201	0. 070	94. 3	97. 1
UNet+WGAN	0. 189	0. 069	95. 0	98. 0
UNet+GFM	0. 188	0. 062	95. 4	97. 8
UNet+DFOM	0. 190	0. 062	94. 9	97. 8
UNet+GFM+DFOM	0. 185	0. 058	95. 6	98. 2
ALG-Net	0. 187	0. 047	95. 9	98. 6

从表 1 可以看出: 使用对抗学习机制构建 WGAN 模型可以促使分割结果更加真实, 有效提升网络的性能, 在 A_{cup} 和 A_{dish} 指标上加入对抗学习机制比原始 UNet 模型提高了 0. 7、0. 9 个百分点; 使用 GFM 改进骨干网络可以有效提升 OC 分割的性能, 比原始 UNet 模型在 E_{cup} 指标上降低了 0. 013, 在 A_{cup} 指标上提高了 1. 1 个百分点; 再引入 DFOM 后 E_{cup} 降低到了 0. 185, 在评价指标上具有比较明显的优势, 表明了 GFM 与 DFOM 在分割网络中联合使用的有效性。表 1 中的结果验证了各个模块对于提高 OD 和 OC 分割性能均有一定的有效性。最终, 将以上模块融合得到本文提出的框架 ALG-Net, 相较于原始 UNet 模型, ALG-Net 在 E_{cup} 和 E_{dish} 指标上降低了 0. 014、0. 023, 在 A_{cup} 和 A_{dish} 指标上提升了 1. 6、1. 5 个百分点, 充分验证了本文提出的方法对改进 UNet 模型的有效性。

之后, 为了更全面地探讨通道融合对模型性能的影响, 本文基于 REFUGE 数据集将不同通道特征与 RGB 图像进行融合, 并对比效果, 实验结果如表 2 所示。

根据表 2 的数据可以发现, 绿色通道的融合显著提高了网络对 OC 区域的敏感性, 实现了最优的性能。此外, 红色通道的融合也对 OD 区域的分割

表 2 融合通道消融实验结果

Table 2 Fusion channel ablation experiment result				
算法	E_{cup}	E_{dish}	$A_{cup}/\%$	$A_{dish}/\%$
UNet(baseline)	0. 201	0. 070	94. 3	97. 1
融合所有通道	0. 191	0. 055	93. 7	98. 4
融合红色通道	0. 206	0. 048	94. 6	98. 5
融合蓝色通道	0. 191	0. 050	95. 4	98. 2
融合绿色通道	0. 187	0. 047	95. 9	98. 6

产生了正面效果, E_{dish} 和 A_{dish} 分别达到了 0. 048 和 98. 5%。尽管性能略逊于融合绿色通道的实验, 但再次验证了绿色通道能够更清晰地呈现 OC 区域信息, 更适用于 OC 的分割, 而红色通道则能够更明显地展示 OD 区域信息, 更适用于 OD 的分割^[5]。然而, 融合蓝色通道和融合所有通道的实验未显示出显著优势, 特别是融合所有通道, 其包含的信息繁多, 与原有 RGB 图像提取的信息大部分重叠, 无法充分学习到不同通道间的信息差异, A_{dish} 达到了 98. 4%, 但 A_{cup} 仅为 93. 7%。

2) 不同算法对比。

为了进一步验证所提方法的性能, 本文将 ALG-Net 与几种分割算法进行比较, 包括本文所采用的基线算法 UNet^[4]、同样采用极坐标转换与多尺度输出的 MNet^[8], 以及与本文密切相关的方法, 如本文参考的方法 Depth guided Net^[9]、Cascade Net^[29]等。此外, 本文还将 ALG-Net 与近三年内针对 OD 和 OC 分割的最新算法进行比较, 以确保实验比较更全面、更具说服力。这些算法包括 DDSC-Net^[36]和 EARDS^[37], 以及与 Transformer 结合的先进算法 SwinUNet^[38]和 TUNet^[39]。关于经过再训练的比较算法, 在表 3 中详细说明了训练时所使用的参数。

表 3 再训练算法参数对比

Table 3 Re-training algorithm parameter comparison	
算法	参数
UNet, MNet	lr 为 0. 000 05, Batch Size 为 2, Epoch 为 1 000
Cascade Net ^[29]	lr 为 0. 000 1, Batch Size 为 2, Epoch 为 1 000
SwinUNet, TUNet	lr 为 0. 01, Batch Size 为 3, Epoch 为 1 000, vit_patches_size 为 16

各算法的实验结果见表 4, 其中“—”表示无数据。结果表明本文所提出的 ALG-Net 在 REFUGE 和 Drishti-GS 数据集上表现出极佳的效果, 比其他算法更具优势。可以看出, 本文所提出的方法在 E_{dish} 、 A_{dish} 和 A_{cup} 上达到了最优性

能。在 REFUGE 数据集上,相较于经典的 MNet 算法,ALG-Net 在 E_{cup} 和 E_{dish} 指标上降低了 0.014、0.023,在 A_{cup} 和 A_{dish} 指标上提升了 1.5、1.0 个百分点,进一步证明了对抗学习和引导机制的有效性。此外,相较于与本文密切相关的方

法以及近期的先进算法,ALG-Net 在 E_{dish} 、 A_{dish} 和 A_{cup} 指标上也表现出了更优性能。根据实验结果可知,在 Drishti-GS 数据集上,ALG-Net 各项评价指标都表现出最优性能,验证了本文方法的优越性。

表 4 不同方法在 REFUGE 和 Drishti-GS 数据集上的结果比较

Table 4 Result comparison of different methods on the REFUGE and Drishti-GS datasets

算法	Drishti-GS				REFUGE			
	E_{cup}	E_{dish}	$A_{cup} / \%$	$A_{dish} / \%$	E_{cup}	E_{dish}	$A_{cup} / \%$	$A_{dish} / \%$
UNet ^[4]	0.169	0.046	95.8	98.1	0.198	0.070	94.3	97.5
MNet ^[8]	0.165	0.045	95.5	98.5	0.201	0.070	94.4	97.6
Depth guided Net ^[9]	0.240	0.071	94.1	97.2	—	—	—	—
Cascade Net ^[29]	0.197	0.073	94.5	97.4	0.208	0.125	93.1	96.9
DDSC-Net ^[36]	0.156	0.043	—	—	0.194	0.076	—	—
SwinUNet ^[38]	0.182	0.085	94.7	96.8	0.187	0.055	94.5	98.1
TUNet ^[39]	0.152	0.051	95.6	98.3	0.179	0.052	94.8	98.3
EARDS ^[37]	0.151	0.050	92.1	97.5	0.198	0.085	89.6	95.6
ALG-Net	0.149	0.041	96.6	98.5	0.187	0.047	95.9	98.6

为了深入展示不同算法之间的差异,将本文所提出的 ALG-Net 与 UNet、MNet、SwinUNet、TUNet 以及 Cascade Net 在参数量和收敛性两方面进行了比较。具体而言,每个算法所使用的参数量见表 5,本文所提出的 ALG-Net 包含 2.279×10^7 个参数,其中分割网络包含 1.178×10^7 个参数,鉴别器包含 1.101×10^7 个参数。从表 4 和表 5 可知 ALG-Net 在性能方面表现出与 SwinUNet、TUNet 相媲美,甚至更出色的性能,而其参数量远低于 SwinUNet 和 TUNet。

表 5 不同算法的参数量

Table 5 Number of parameters for different algorithms

算法	参数量/ 10^6
UNet	17.27
MNet	10.22
SwinUNet	41.34
TUNet	105.28
Cascade Net	80.96
ALG-Net	22.79

随后,在 Drishti-GS 数据集上进行了收敛性比较。为确保公平比较,为每个算法采用相同的参数进行了 1 000 个 Epoch 的训练。图 6 展示了各算法在训练过程中 L_{dice} 的变化趋势。从图中可以观察到,本文提出的 ALG-Net 误差曲线快速收敛,并趋向平缓,在 200 轮次内完成了模型的收敛。

最后,本文展示了部分 OD 和 OC 分割结果,如

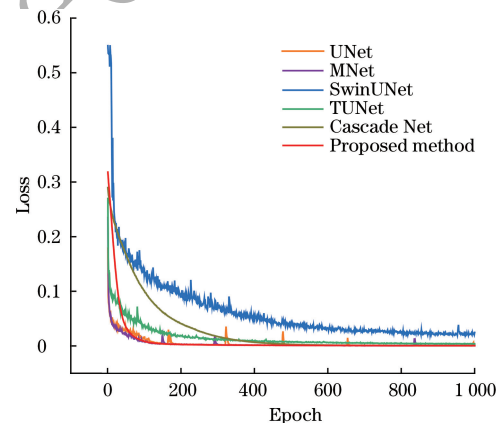


图 6 不同算法在 Drishti-GS 数据集上的 Dice 误差曲线

Fig. 6 Dice loss curves of different algorithms on the Drishti-GS dataset

图 7 所示,其中灰色和白色的区域分别表示 OD 和 OC。从图中可以看出,所提出的 ALG-Net 在分割 OD 和 OC 时表现出更加逼真的效果,对边界的处理也要优于其他方法,并且对于一些特殊图像的处理也更加真实,如图 7 的第 4 行示例,该图像相较于数据集上的其他图像具有明显不同的数据分布,所对比的几个算法皆无法正确定位 OD 和 OC 边界,在分割结果上呈现出明显异常。本文所提出的 ALG-Net 则表现出更好的适应能力和鲁棒性,分割结果的数据分布更加贴合真实标签,再次验证了相较于其他算法,ALG-Net 利用引导机制,加强了对关键信息的关注和定位,对抗学习机制的引入也促进了分割网络生成更逼真的 OD 和 OC 分割结果。

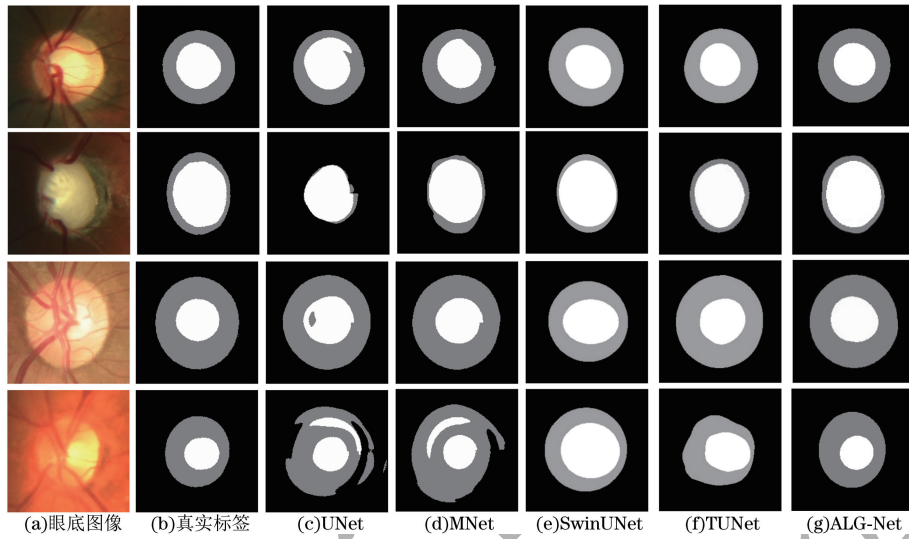


图 7 视盘和视杯分割的可视化示例

Fig.7 Visualization example of optic disc and optic cup segmentation

3) 迁移学习。

为了评估所提方法在不同数据集上迁移学习的性能,本文进行了一系列实验。首先在 RIGA 数据集上进行训练,然后将经过训练的模型分别在 Drishti-GS 和 REFUGE 数据集上进行微调,并进行测试。同样地,在 REFUGE 数据集上进行训练,然后将训练好的模型分别在 RIGA 和 REFUGE 数据集上进行微调,并

进行测试。在微调过程中,设置学习率为 0.000 05,训练 Epoch 为 10。调整过程仅涉及分割网络,因此保留了 L_{dice} 作为唯一的损失函数进行计算。

迁移学习的实验结果如表 6 所示,通过这一系列实验,旨在验证所提方法在不同数据集上的泛化性能和迁移学习效果,对模型在实际应用中的适用性进行更全面的评估。

表 6 迁移学习实验结果

Table 6 Transfer learning experiment result

Source	Target	E_{cup}	E_{dish}	$A_{\text{cup}}/\%$	$A_{\text{dish}}/\%$
RIGA	DRISHTI	0.156	0.054	96.6	98.1
	REFUGE	0.182	0.063	94.9	97.8
REFUGE	DRISHTI	0.152	0.046	96.3	98.4
	RIGA	0.160	0.035	96.2	99.0

3.4.2 青光眼分类实验

本文根据 ALG-Net 对 OD 和 OC 的分割结果,进一步计算 CDR。利用椭圆拟合的方法,寻找分割结果的最小外接椭圆,得到垂直 OD 和垂直 OC 的直径,计算得到 CDR,进而判断是否患有青光眼。在 REFUGE 的测试集中随机选取了 124 张标签,其中 18 张青光眼,其受试者工作特征(ROC)曲线如图 8 所示。

本文利用 ROC 曲线下面积(AUC)作为青光眼诊断的衡量指标。从图 8 中可以看出,ALG-Net 在青光眼分类中获得了最高的性能指标(AUC 为 0.983),相较于经典 UNet 算法提高了 0.015,相较于 TUNet 提高了 0.011。本文所提出的方法在青光眼分类中具有很强的竞争力,可以为青光眼诊断提供有力的支持。

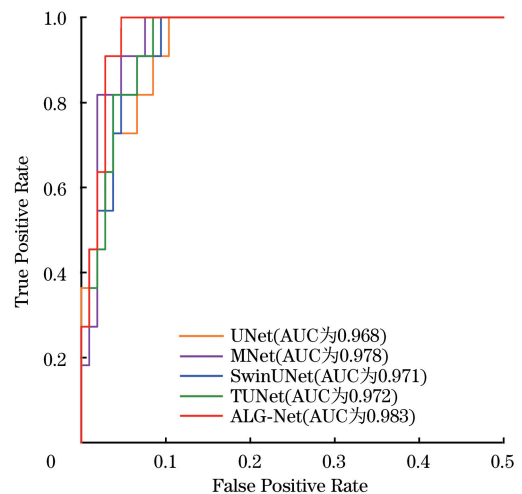


图 8 基于 CDR 的青光眼筛查 ROC 曲线及 AUC 值

Fig.8 ROC curve and AUC value of glaucoma screening based on CDR

4 结束语

针对现有方法通道信息利用不充分、OC 分割不精确的问题,本文提出了一种基于对抗学习和引导机制的 OD 和 OC 联合分割框架 ALG-Net。通过引导分割网络加强对 ROI 的关注,提取不同通道之间的关键语义信息。此外,利用对抗学习的机制促使分割网络生成更逼真的分割结果,提升图像分割精度。在 REFUGE 和 Drishti-GS 数据集上的实验结果表明,本文方法相较于其他方法更具竞争力,可以为青光眼的诊断提供更为可靠的判断依据。

参考文献

- [1] PACHADE S, PORWAL P, KOKARE M, et al. NENet: nested EfficientNet and adversarial learning for joint optic disc and cup segmentation[J]. *Medical Image Analysis*, 2021, 74: 102253.
- [2] JIANG Y M, DUAN L X, CHENG J, et al. JointRCNN: a region-based convolutional neural network for optic disc and cup segmentation[J]. *IEEE Transactions on Bio-Medical Engineering*, 2020, 67(2): 335-343.
- [3] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 39(4): 640-651.
- [4] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]// *Proceedings of MICCAI 2015*. Berlin, Germany: Springer, 2015: 234-241.
- [5] ISSAC A, PARTHA SARATHI M, DUTTA M K. An adaptive threshold based image processing technique for improved glaucoma detection and classification[J]. *Computer Methods and Programs in Biomedicine*, 2015, 122(2): 229-244.
- [6] ORLANDO J I, FU H Z, BARBOSA B J, et al. REFUGE challenge: a unified framework for evaluating automated methods for glaucoma assessment from fundus photographs[J]. *Medical Image Analysis*, 2020, 59: 101570.
- [7] SIVASWAMY J, KRISHNADAS S R, DATT J G, et al. Drishti-GS: retinal image dataset for Optic Nerve Head (ONH) segmentation [C] // *Proceedings of the 11th International Symposium on Biomedical Imaging (ISBI)*. Washington D. C., USA: IEEE Press, 2014: 53-56.
- [8] EDUPUGANTI V G, CHAWLA A, KALE A. Automatic optic disc and cup segmentation of fundus images using deep learning[C]// *Proceedings of the 25th IEEE International Conference on Image Processing*. Washington D. C., USA: IEEE Press, 2018: 2227-2231.
- [9] SHANKARANARAYANA S M, RAM K, MITRA K, et al. Fully convolutional networks for monocular retinal depth estimation and optic disc-cup segmentation[J]. *IEEE Journal of Biomedical and Health Informatics*, 2019, 23(4): 1417-1426.
- [10] FU H Z, CHENG J, XU Y W, et al. Joint optic disc and cup segmentation based on multi-label deep network and polar transformation[J]. *IEEE Transactions on Medical Imaging*, 2018, 37(7): 1597-1605.
- [11] CHENG J, LI Z G, GU Z W, et al. Structure-preserving guided retinal image filtering and its application for optic disk analysis[J]. *IEEE Transactions on Medical Imaging*, 2018, 37(11): 2536-2546.
- [12] YU S, XIAO D, FROST S, et al. Robust optic disc and cup segmentation with deep learning for glaucoma detection[J]. *Computerized Medical Imaging and Graphics*, 2019, 74: 61-71.
- [13] XIE Z, LING T H, YANG Y Y, et al. Optic disc and cup image segmentation utilizing contour-based transformation and sequence labeling networks[J]. *Journal of Medical Systems*, 2020, 44(5): 96.
- [14] YIN P S, XU Y W, ZHU J H, et al. Deep level set learning for optic disc and cup segmentation[J]. *Neurocomputing*, 2021, 464: 330-341.
- [15] YUAN X, ZHOU L X, YU S Y, et al. A multi-scale convolutional neural network with context for joint segmentation of optic disc and cup[J]. *Artificial Intelligence in Medicine*, 2021, 113: 102035.
- [16] GUO W, ZHANG G D, GONG Z X, et al. Effective integration of object boundaries and regions for improving the performance of medical image segmentation by using two cascaded networks[J]. *Computer Methods and Programs in Biomedicine*, 2021, 212: 106423.
- [17] SHANMUGAM P, RAJA J, PITCHAI R. An automatic recognition of glaucoma in fundus images using deep learning and random forest classifier[J]. *Applied Soft Computing*, 2021, 109: 107512.
- [18] HASAN M K, ALAM M A, ELAHI M T E, et al. DRNet: segmentation and localization of optic disc and fovea from diabetic retinopathy image[J]. *Artificial Intelligence in Medicine*, 2021, 111: 102001.
- [19] HERVELLA Á S, ROUCO J, NOVO J, et al. End-to-end multi-task learning for simultaneous optic disc and cup segmentation and glaucoma classification in eye fundus images[J]. *Applied Soft Computing*, 2022, 116: 108347.
- [20] ZHOU W, PENG Y H, JI J H, et al. RMSDSC-Net: a robust multiscale feature extraction with depthwise separable convolution network for optic disc and cup segmentation[J]. *International Journal of Intelligent Systems*, 2022, 37(12): 11482-11505.
- [21] 蒋芸, 高静, 王发林. SAG-Net: 用于联合视盘和视杯分割的新型跳过注意力指导网络[J]. *计算机工程与科学*, 2021, 43(7): 1273-1282.
- JIANG Y, GAO J, WANG F L. SAG-Net: a new skip attention guided network for joint disc and cup segmentation[J]. *Computer Engineering & Science*, 2021, 43(7): 1273-1282. (in Chinese)
- [22] 徐蓬泉, 梁宇翔, 李英. 融合多尺度语义和剩余瓶颈注意力的医学图像分割[J]. *计算机工程*, 2023, 49(10): 162-170.
- XU P Q, LIANG Y X, LI Y. Medical image segmentation fusing multi-scale semantic and residual bottleneck attention[J]. *Computer Engineering*, 2023, 49(10): 162-170. (in Chinese)
- [23] 刘洪普, 赵一浩, 侯向丹, 等. 融合上下文和注意力的视盘视杯分割[J]. *中国图象图形学报*, 2021, 26(5): 1041-1057.
- LIU H P, ZHAO Y H, HOU X D, et al. Optic disc and cup segmentation by combining context and attention[J]. *Journal of Image and Graphics*, 2021, 26(5): 1041-1057. (in Chinese)
- [24] 刘然, 刘建霞, 王海翼. 改进 U-net++ 的青光眼视盘视杯分割方法[J]. *电子设计工程*, 2023, 31(1): 27-33.
- LIU R, LIU J X, WANG H Y. The improved U-net++ method for glaucoma optic disc and optic cup segmentation[J]. *Electronic Design Engineering*, 2023, 31(1): 27-33. (in Chinese)
- [25] IMTIAZ R, KHAN T M, NAQVI S S, et al. Screening of glaucoma disease from retinal vessel images using semantic segmentation[J]. *Computers & Electrical Engineering*, 2021, 91: 107036.

- [26] HAIDER A, ARSALAN M, LEE M B, et al. Artificial intelligence-based computer-aided diagnosis of glaucoma using retinal fundus images[J]. Expert Systems with Applications, 2022, 207: 117968.
- [27] WANG S J, YU L Q, YANG X, et al. Patch-based output space adversarial learning for joint optic disc and cup segmentation[J]. IEEE Transactions on Medical Imaging, 2019, 38(11): 2485-2495.
- [28] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[EB/OL]. [2023-09-20]. <http://arxiv.org/abs/1406.2661>.
- [29] BIAN X S, LUO X B, WANG C, et al. Optic disc and optic cup segmentation based on anatomy guided cascade network[J]. Computer Methods and Programs in Biomedicine, 2020, 197: 105717.
- [30] 刘丽霞, 宣士斌, 刘畅, 等. 多专家注释的视杯和视盘不确定性量化[J]. 计算机工程, 2023, 49(1): 250-257, 269.
LIU L X, XUAN S B, LIU C, et al. Quantification of optic cup and optic disc uncertainty with multi-expert annotations[J]. Computer Engineering, 2023, 49(1): 250-257, 269. (in Chinese)
- [31] YANG L, ZHANG R Y, LI L, et al. SimAM: a simple, parameter-free attention module for convolutional neural networks [C] // Proceedings of the 38th International Conference on Machine Learning. [S. l.]: PMLR, 2021: 11863-11874.
- [32] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2021: 13713-13722.
- [33] 崔博, 贾兆年, 姬鹏, 等. 基于改进 EfficientNetV2 网络的脑肿瘤分类方法[J]. 吉林大学学报(理学版), 2023, 61(5): 1169-1177.
- CUI B, JIA Z N, JI P, et al. Brain tumor classification method based on improved EfficientNetV2 network [J]. Journal of Jilin University(Science Edition), 2023, 61(5): 1169-1177. (in Chinese)
- [34] ARJOVSKY M, CHINTALA S, BOTTOU L. Wasserstein GAN[EB/OL]. [2023-09-20]. <http://arxiv.org/abs/1701.07875>.
- [35] ALMAZROA A, ALODHAYB S, OSMAN E, et al. Retinal fundus images for glaucoma analysis: the RIGA dataset[C]// Proceedings of Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications. [S. l.]: SPIE, 2018: 55-62.
- [36] LIU B Y, PAN D R, SONG H. Joint optic disc and cup segmentation based on densely connected depthwise separable convolution deep network[J]. BMC Medical Imaging, 2021, 21(1): 14.
- [37] ZHOU W, JI J H, JIANG Y, et al. EARDS: EfficientNet and attention-based residual depth-wise separable convolution for joint OD and OC segmentation [J]. Frontiers in Neuroscience, 2023, 17: 1139181.
- [38] DEWA J K, RACHMAWATIE, KOSALA G. Investigating self-attention in Swin-UNet model for disc and cup segmentation [C] // Proceedings of the 10th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE). Washington D. C., USA: IEEE Press, 2023: 1-10.
- [39] LI Z R, ZHAO C, HAN Z K, et al. TUNet and domain adaptation based learning for joint optic disc and cup segmentation[J]. Computers in Biology and Medicine, 2023, 163: 107209.

编辑 金胡考