

基于 RAA-UNet 的虹膜块状特征分割^{*}

陈宇 唐云祁

(中国人民公安大学侦查学院 北京 100038)

摘要: 目前虹膜识别结果尚不能应用到司法审判当中,法庭科学领域开始关注以虹膜可解释特征统计规律为基础的量化鉴定方法,为此需要实现虹膜纹理特征的自动分割提取。针对近红外虹膜图像中块状特征的提取问题,提出一种结合残差网络、注意力机制和空洞空间金字塔池化的虹膜块状特征分割网络。为此,首先构建了虹膜块状特征标注数据集,用于模型的训练、验证和测试。其次,以 UNet 为基础框架进行改进,将 UNet 的卷积模块替换为残差模块,促进梯度的传播并提高特征的保留和传递能力;在跳跃连接中加入注意力门模块以提高模型对块状特征的感知能力;在模型中部和末端采用空洞空间金字塔池化模块,扩大感受野并进行多尺度特征提取和融合。最后,提出了结合交叉熵和 Dice 系数的混合损失函数,并采用归一化和直方图均衡化等预处理技术以突出虹膜块状特征。实验结果表明,RAA-UNet 在测试集上的表现优于其他对比模型, F1 分数、mIoU 和 Dice 系数分别达到了 82.64%、84.21%、81.66%,较 UNet 提升 4.42%、3.37% 和 3.87%。损失函数实验确定了最佳权重因子,直方图均衡化处理显著提升了分割效果,消融实验验证了模型改进的有效性。提出的 RAA-UNet 语义分割模型能够实现虹膜块状特征的准确分割,可为虹膜鉴定的研究提供技术支撑。

关键词: 虹膜块状特征;UNet;虹膜鉴定;残差网络;金字塔池化;注意力机制

中图分类号: TP391.4; TN911.7 **文献标识码:** A **国家标准学科分类代码:** 520.60

Block-shaped iris feature segmentation based on RAA-UNet

Chen Yu Tang Yunqi

(School of Criminal Investigation, People's Public Security University of China, Beijing 100038, China)

Abstract: Currently, the results of iris recognition cannot be applied to judicial trials. The forensic science community has begun to focus on quantitative identification method based on the statistical rules of interpretable iris features. For this purpose, it is necessary to achieve automatic segmentation and extraction of iris texture features. A segmentation network for block-shaped iris features in near-infrared iris images is proposed, which combines residual networks, attention mechanisms, and atrous spatial pyramid pooling. First, a block-shaped iris feature annotation dataset was constructed for model training, validation, and testing. Secondly, improvements were made to the UNet framework as follows: the convolutional modules were replaced with residual modules to promote gradient propagation and enhance feature retention and transmission capabilities; attention gate modules were added to the skip connections to improve the model's perception of block-shaped features; and atrous spatial pyramid pooling modules were employed in the middle and end parts of the model to expand the receptive field and perform multi-scale feature extraction and fusion. Finally, a hybrid loss function combining cross-entropy and Dice coefficient was proposed, and preprocessing techniques such as normalization and histogram equalization were used to highlight block-shaped iris features. Experimental results show that the RAA-UNet outperforms other comparison models on the test set, with F1 score, mIoU, and Dice coefficient reaching 82.64%, 84.21%, and 81.66%, respectively, representing improvements of 4.42%, 3.37%, and 3.87% over UNet. The loss function experiments determined the optimal weight factor, and histogram equalization significantly improved segmentation performance. Ablation experiments verified the effectiveness of the model improvements. The proposed RAA-UNet semantic segmentation model can accurately segment block-shaped iris features, providing technical support for iris identification research.

Keywords: block-shaped iris features;UNet;iris identification;residual network;pyramid pooling;attention mechanism

0 引言

虹膜是人体唯一性极强的生物特征,在近红外光下呈

现出丰富的纹理特征^[1]。虹膜识别技术由于其精度高^[2]、比对速度快、非接触性等特点,已广泛应用于公共安全^[3]、国民身份系统^[4]等领域。虹膜识别技术需要从虹膜纹理中

提取到用于身份识别的区分性信息,为此诸多学者探讨了多种多样的特征提取方法,主要可以分成基于手工特征的传统方法和基于深度学习的方法^[5]。其中,基于手工特征的方法主要从图像的频域或空域中提取特征,从虹膜图像频域中提取特征的典型方法有 Gabor 滤波、小波变换等;从虹膜图像空域中提取特征的典型方法有局部二值模式^[6]、尺度不变特征变换^[7]等。

随着深度学习技术的发展,深度网络开始应用于虹膜识别,通过数据驱动的方式自动学习虹膜图像的特征表征。基于深度学习的虹膜识别可分为预训练网络搭配分类器、分类网络和相似度网络等 3 种策略^[8]。许多学者使用在大规模数据集上训练的深度学习网络搭配分类器进行虹膜图像的特征提取和匹配,如 Nguyen 等^[9]证明了 Inception、Resnet、DenseNet 等预训练的卷积神经网络能够有效地提取虹膜图像中的判别性视觉特征。分类网络使用具有 softmax 损失的深度架构将输入的虹膜图像分类为已知身份列表,典型网络有 Gangwar 等^[10]提出的基于卷积神经网络的 DeepIrisNet, Gao 等^[11]提出的基于 Transformer 的 SwinIris。相似度网络使用具有成对损失的深度架构,学习一种度量来表示两张虹膜图像的相似或不相似程度,而无需预先知道图像的身份,典型网络有 Zhao 等^[12]提出的 Uninet, Nguyen 等^[13]提出的 ComplexIrisNet。

但是,虹膜识别技术,无论是基于手工特征的方法还是深度学习的方法,所提取的特征往往难以被人所直观理解,不适用于法庭科学领域的物证辨析。基于手工特征的特征提取方法无法获得虹膜中特定类型纹理的大小、形状、位置等信息^[14],所提取到的特征通常既不易于解释,也不对应于任何物理可见的特征^[15]。基于深度学习的特征提取方法,所使用的深度神经网络具有高度非线性的复杂结构^[16],从虹膜图像中所提取到的特征通常是高维的、抽象的向量或矩阵,缺乏直观的物理意义。目前,在司法审判过程中,虹膜识别结果还无法直接成为司法证据来使用。

近年来,法庭科学领域开始关注以虹膜可解释特征的统计规律为基础的量化鉴定问题,尝试通过两张虹膜图像中的人类可理解特征的匹配数量来对虹膜进行检验鉴定。为此,需要实现虹膜图像中可解释纹理特征的分类和准确分割提取,以研究其在大规模数据集上的分布规律。陈子龙等^[17]提出将虹膜上的特征分成放射状沟线、向心沟、卷缩轮、隐窝和色素点五类。张翌阳等^[18]提出依据形态特点将虹膜上的纹理特征归纳成块状特征、线状特征和环状特征 3 类。块状特征是该 3 类特征中最为显著、分布最广的特征。

一些学者探讨了以可解释特征为基础的虹膜识别算法,为虹膜块状特征的分割提供了有价值的参考。Shen 等^[19]尝试通过虹膜隐窝进行虹膜匹配,通过实验证实了虹膜隐窝特征用于身份认定的有效性^[20],并提出一种虹膜隐窝的检测方法^[21]。该方法包括使用灰度形态学变换作为

初始检测器,结合多尺度金字塔架构来提取虹膜隐窝的候选区域,然后通过优化过程,根据像素对平均强度和平均梯度的贡献,调整虹膜隐窝的边界,以获得准确的检测结果。

同时,也有学者专注于虹膜中各种物理纹理的检测。刘笑楠等^[22]通过组合窗口搜索的方法实现了对虹膜块状纹理的精准检测。苑玮琦等^[23]采用线性模板和模糊规则实现了虹膜坑洞纹理的检测。朱立军等^[24]提出采用双线性模板和分块策略实现了虹膜色素斑的检测,朱立军等^[25]也提出通过线性模板和 BAB 策略结合 SVM 的方法实现了块状纹理的检测。此外,苑玮琦等^[26]提出运用形态学的方法提取出虹膜图像中的所有目标纹理,并通过 SVM 和定义的约束条件,筛选出最终的坑洞纹理。

然而,上述基于传统图像处理技术和统计学习的提取方法存在泛化性不佳、需要人工设计特征和调整参数、速度慢、精度低的不足,不能满足研究虹膜块状特征规律的需要。自 2012 年, Krizhevsky 等^[27]提出的 AlexNet 在 ImageNet 图像分类比赛中取得了突破性的成果以来,卷积神经网络被引入语义分割任务,带来了显著的性能提升。Long 等^[28]提出的 FCN 是首个端到端的语义分割深度学习模型。FCN 通过将全连接层替换为卷积层,实现了任意尺寸输入图像的语义分割,并通过上采样恢复原始分辨率。Ronneberger 等^[29]提出的 UNet 最初用于生物医学图像分割。UNet 采用了对称的编码-解码结构,通过跳跃连接融合不同尺度的特征信息,显著提升了分割精度。

Chen 等^[30]综合了膨胀卷积、金字塔池化等技术,提出了 DeepLabV3 模型。Chen 等^[31]在 DeepLabV3 基础上结合了编码器-解码器结构,通过上采样和跳跃连接将浅层特征与深层特征进行融合,提出了 DeepLabV3+ 模型。Transformer 模型在自然语言处理领域取得了巨大成功,受此启发,研究人员开始探索其在计算机视觉任务中的应用。Dosovitskiy 等^[32]提出的 ViT 首次将 Transformer 架构引入图像分类任务中,展示了其强大的特征提取能力。Xie 等^[33]将层次化结构的混合 Transformer (mix transformer, MiT) 编码器与轻量级的多层感知机 (multi-layer perceptron, MLP) 解码器相结合,提出了 SegFormer 语义分割模型,该模型通过多尺度特征提取和简化解码器设计实现了优越的性能和效率。

在上述的语义分割模型中,UNet 网络因其在医学图像分割中的优良性能而受到广泛关注。UNet 网络持续演化,衍生出如 Attention-UNet^[34], U2Net^[35] 等诸多变体网络模型。UNet 及其变体在血管、病变^[36] 等医学图像分割任务中取得了显著的成果。UNet 还具备数据量要求较低的优势,能够满足小样本训练的需求。因此本文考虑将 UNet 基础框架从医学图像分割领域迁移到虹膜鉴定领域,以精确分割虹膜块状特征。

由于虹膜图像中块状特征形状多变、大小不一,需要捕捉更多的细节特征,且图像背景复杂度高,存在大量的干扰

因素(如眼睑、睫毛等),对模型的抗干扰能力提出较高的要求。将 UNet 直接应用于虹膜块状特征分割不足以解决上述问题,故引入残差网络,促进梯度的传播,增强特征的传递,提高模型对细节信息的提取能力;添加注意力机制,在编码器与解码器之间加入注意力门模块,使模型能够聚焦于图像中的关键区域,减少噪声干扰,提高分割精度;加入空洞空间金字塔池化模块,提取多尺度特征信息,减轻因下采样带来的空间分辨率丢失问题,从而提升细节保留和特征提取能力。

本文针对虹膜鉴定中的特征提取问题,以虹膜块状特征为研究对象,提出使用改进的 RAA-UNet 语义分割模型实现虹膜图像中块状特征的分割提取。具体包括,建立虹膜块状特征数据集,改进和构建虹膜块状特征分割模型,比较不同模型在评价指标和分割图像效果上的性能差别。本文所提模型可为后续统计虹膜块状特征的分布规律提供技术支撑。

1 本文方法

1.1 RAA-UNet 网络模型结构

1) 残差模块

随着网络的深度增加,卷积神经网络在反向传播过程中可能会遭遇梯度消失或梯度爆炸的问题。这些问题会导致网络难以收敛,甚至导致性能的退化,限制了网络的深度和表达能力。为了解决梯度消失和梯度爆炸问题,He 等^[37]提出残差网络。其核心思想是引入残差连接,允许网络学习输入和输出之间的残差信息,而不是直接学习整个映射关系。

残差结构的设计使得网络更容易地传播梯度,加速了网络的训练过程,并提高了模型的性能和泛化能力。因此,为了更好地提取虹膜块状特征的信息,提高模型的特征提取能力,本文将 UNet 中编码器和解码器的卷积模块替换为残差模块。由于每个残差块的输入特征的维度和输出维度都不同,每个残差模块的残差连接均使用了 1×1 卷积调整通道维度。本文所用的残差结构表示方式为:

$$y = F(x_i, \{W_i\}) + W'x_i \quad (1)$$

式中: y 表示残差模块的输出, x_i 表示残差模块的输入, $F(x_i, \{W_i\})$ 表示特征提取部分,即要学习的残差映射, W' 表示用 1×1 卷积操作调整通道维度。

本文所用残差模块的具体结构如图 1 所示。图 1 展示了编码器中的第一个残差模块,其由主路径和残差连接组成。主路径包括两个卷积层,卷积核大小均为 3×3 ,输出通道数为 64。每个卷积层后面跟着批标准化处理(batch normalization, BN)和修正线性单元(rectified linear unit, ReLU)。BN 层通过规范化每一层的输入,加速了网络收敛,并提高了模型的泛化能力。ReLU 激活函数引入非线性变换,增强了网络的表达能力,同时还能产生稀疏激活,提高特征表示效率并减少过拟合风险。由于模块的输入和

输出维度不同,所以残差连接使用 1×1 卷积来调整维度,确保两个路径的特征图在维度上的尺寸一致。

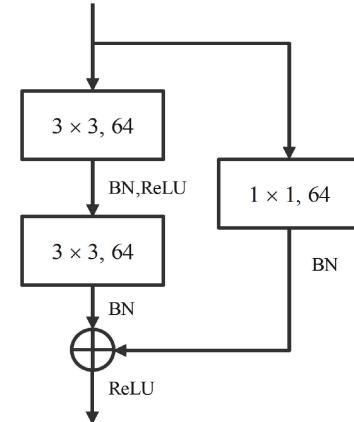


图 1 残差模块
Fig. 1 Residual module

残差模块的引入使得网络更容易地传播梯度,缓解了梯度消失和梯度爆炸问题。同时,残差连接还能够帮助网络更好地保留和传递低层次的特征信息,提高模型的特征提取能力和泛化能力。

2) 注意力门模块

虹膜图像中存在着眼睑、睫毛、背景噪声等干扰因素,为了提高模型对目标区域的关注程度,减少背景干扰,本文引入注意力机制。本文在编码器和解码器之间的跳跃连接中加入了注意力门模块,以帮助网络更好地聚焦于图像中重要的区域。这种机制允许网络根据任务的需求动态地分配权重,以调整对不同特征区域的关注力度。注意力门的具体结构如图 2 所示。

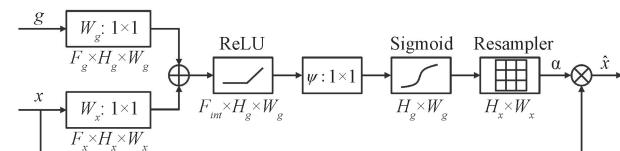


图 2 注意力门模块
Fig. 2 Attention gate module

注意力门模块接受来自解码器的特征图 g 和编码器的特征图 x 作为输入。两个特征图经过一系列线性变换和激活函数后,得到注意力权重 α 。这个权重图表征了每个像素点的重要性:权重接近 1 的像素点被认为是关键区域,而接近 0 的则被视为不重要区域。最后,原始的编码器特征图 x 与注意力权重 α 进行元素级的相乘得到加权后的特征图 \hat{x} 。这种特定区域的权重分配有助于提高网络对关键区域的感知能力,从而能够提高分割的准确性。

3) 空洞空间金字塔池化模块

在编码器提取特征的过程中,会伴随着连续的下采样操作,导致空间分辨率的降低。这种操作有助于增加感受野并减少计算量,但不可避免地丢失部分空间信息。膨胀

卷积(也称为空洞卷积)被提出用以扩展卷积神经网络的感受野,而不增加额外的参数或计算成本。膨胀卷积通过引入一个空洞率参数来扩大卷积核中元素的间隔,使得卷积核覆盖更广的输入区域。膨胀卷积通过增加感受野,可以有效缓解由于下采样而导致的空间分辨率问题。空洞空间金字塔池化(atrous spatial pyramid pooling, ASPP)是在膨胀卷积的基础上提出的多尺度特征提取模块,通过组合多个不同空洞率的膨胀卷积层并行工作,能够同时捕捉多尺度的上下文信息。其具体结构如图 3 所示。

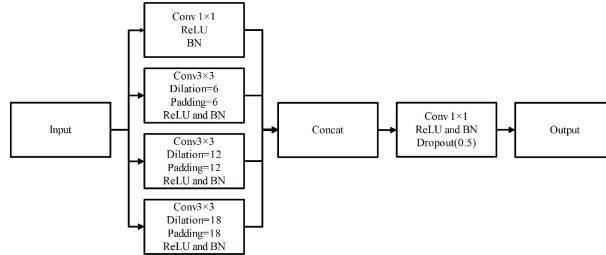


图 3 ASPP 模块

Fig. 3 ASPP module

ASPP 模块包括 4 个并行分支,包括 1 个 1×1 标准卷积分支和 3 个 3×3 膨胀卷积分支。3 个膨胀卷积分支采用不同空洞率来获取不同尺度的上下文信息,并通过相应的填充确保了卷积操作后特征图的尺寸保持不变。所有分支并行处理输入特征图,各分支的输出再在通道维度上拼接整合,使模型能够同时利用不同尺度的特征信息。拼接后的特征图经 1×1 卷积层整合不同分支的信息,调整通道的维度,并通过 dropout 层减少过拟合,提高模型的泛化能力。

4) 双线性插值

双线性插值是一种常用的图像重采样技术,其基本原理是对 4 个最近邻的像素值进行加权平均,从而计算出新

像素点的值,如图 4 所示。该方法的插值公式可以表示为:设 f 是图像的像素强度函数,需要在 X 和 Y 方向上对像素点进行插值。首先,在 X 方向上对两个邻近点进行插值:

$$R_1 = \frac{x_2 - x}{x_2 - x_1} \cdot f(x_1, y_1) + \frac{x - x_1}{x_2 - x_1} \cdot f(x_2, y_1) \quad (2)$$

$$R_2 = \frac{x_2 - x}{x_2 - x_1} \cdot f(x_1, y_2) + \frac{x - x_1}{x_2 - x_1} \cdot f(x_2, y_2) \quad (3)$$

之后,在 Y 方向上进行最终的线性插值得到目标点 P 的像素值:

$$P = \frac{y_2 - y}{y_2 - y_1} \cdot R_1 + \frac{y - y_1}{y_2 - y_1} \cdot R_2 \quad (4)$$

通过这种方式,双线性插值能够在上采样过程中平滑地计算像素值,从而得到更加连续和自然的图像过渡。双线性插值还能够避免转置卷积可能引入的棋盘效应,并且其不涉及任何学习参数,计算效率更高。

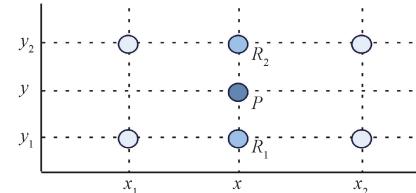


图 4 双线性插值示意图

Fig. 4 Bilinear interpolation schematic diagram

5) RAA-UNet 结构

本文提出的模型主要由对称的编码器和解码器两部分组成。编码器负责提取特征,解码器负责恢复空间细节。整个模型由 8 个残差模块,2 个 ASPP 模块,4 个注意力门模块,3 个池化层,3 个上采样层组成。输入的图像尺寸为 $540 \times 70 \times 1$,输出图像尺寸为 $540 \times 70 \times 2$ 。将该模型命名为 RAA-UNet (residual attention and atrous spatial pyramid pooling UNet, RAA-UNet),示意图如图 5 所示。

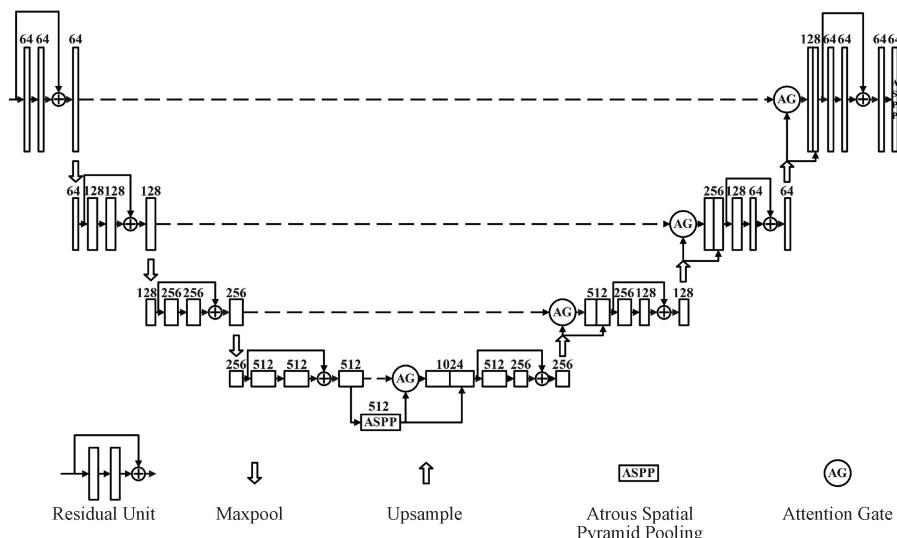


图 5 RAA-UNet 模型网络结构示意图

Fig. 5 RAA-UNet network structure diagram

在整个模型中,编码器和解码器原本的卷积模块均被残差模块取代。编码器和解码器各有 4 个残差模块,卷积核的大小均为 3×3 。相较于原来的卷积模块,残差模块引入了残差连接,能够避免模型深度过深导致的退化问题,并加强特征的传递,从而提高模型的特征提取和分割能力。在编码器中,采用 2×2 最大池化进行下采样。在解码器中,采用双线性插值进行上采样。

为了提高模型对目标的关注度,抑制无关区域,本文在编码器和解码器中间的跳跃连接上加入注意力门模块。上采样后的解码器特征图与经过跳跃连接的编码器特征图同时输入到注意力门中,生成注意力权重 α ,最后得到经过注意力调整后的编码器特征图。调整后的编码器特征图与上采样后的解码器特征图在通道维度上拼接,并输入

到解码器中继续处理。在跳跃连接上加入注意力门,可以引导模型聚焦于更有意义的区域,增强模型的分割性能。

为了减少空间信息的损失,更好地提取虹膜图像中的特征,提高分割效果,本文将 UNet 的第 4 次下采样和对应的卷积模块替换成 ASPP 模块,在多个尺度上提取图像的特征信息,并将多尺度特征信息进行融合。同时在解码器的输出处也增加了 ASPP 模块,以提高模型的分割能力。考虑到输入到 ASPP 图像的大小,将第 1 个 ASPP 的空洞率设为(1,2,3),第 2 个 ASPP 的空洞率设为(6,12,18)。

本文所提出的网络模型的具体结构如表 1 所示。表 1 第 2 列和第 5 列表示输出的图像大小和通道数。第 3 列和第 6 列表示模型的内部细节,如“ $3 \times 3, 64$ ”表示卷积核为 3×3 ,输出通道为 64 的卷积层。

表 1 RAA-UNet 模型网络结构
Table 1 RAA-UNet network constructure

模块	特征大小	RAA-UNet	模块	特征大小	RAA-UNet
Input	540×70	—	AG2	$135 \times 17 \times 256$	—
Conv1	$540 \times 70 \times 64$	$3 \times 3, 64$, BN, ReLU	Concat2	$135 \times 17 \times 512$	AG2+Upsample_1
	$540 \times 70 \times 64$	$3 \times 3, 64$, BN, ReLU	Conv6	$135 \times 17 \times 256$	$3 \times 3, 256$, BN, ReLU
Maxpool	$270 \times 35 \times 64$	2×2		$135 \times 17 \times 128$	$3 \times 3, 128$, BN, ReLU
Conv2	$270 \times 35 \times 128$	$3 \times 3, 128$, BN, ReLU	Upsample2	$270 \times 35 \times 128$	2×2 , bilinear
	$270 \times 35 \times 128$	$3 \times 3, 128$, BN, ReLU	AG3	$270 \times 35 \times 128$	—
Maxpool	$135 \times 17 \times 128$	2×2	Concat3	$270 \times 35 \times 256$	AG3+Upsample_2
Conv3	$135 \times 17 \times 256$	$3 \times 3, 256$, BN, ReLU	Conv7	$270 \times 35 \times 128$	$3 \times 3, 128$, BN, ReLU
	$135 \times 17 \times 256$	$3 \times 3, 256$, BN, ReLU		$270 \times 35 \times 64$	$3 \times 3, 64$, BN, ReLU
Maxpool	$67 \times 8 \times 256$	2×2	Upsample3	$540 \times 70 \times 64$	2×2 , bilinear
Conv4	$67 \times 8 \times 512$	$3 \times 3, 512$, BN, ReLU	AG4	$540 \times 70 \times 64$	—
	$67 \times 8 \times 512$	$3 \times 3, 512$, BN, ReLU	Concat4	$540 \times 70 \times 128$	AG4+Upsample_3
ASPP1	$67 \times 8 \times 512$	Rate=(1, 2, 3)		$540 \times 70 \times 64$	$3 \times 3, 64$, BN, ReLU
AG1	$67 \times 8 \times 512$	—	Conv8	$540 \times 70 \times 64$	$3 \times 3, 64$, BN, ReLU
Concat1	$67 \times 8 \times 1024$	AG1+ASPP1	ASPP2	$540 \times 70 \times 64$	Rate=(6, 12, 18)
Conv5	$67 \times 8 \times 512$	$3 \times 3, 512$, BN, ReLU	Conv9	$540 \times 70 \times 2$	$1 \times 1, 2$
	$67 \times 8 \times 256$	$3 \times 3, 256$, BN, ReLU			
Upsample1	$135 \times 17 \times 256$	2×2 , bilinear			

1.2 损失函数

损失函数在语义分割模型的训练中十分重要,它直接影响着模型的训练效果和最终的性能表现。本节将介绍本文所采用的主要损失函数,包括交叉熵损失函数和 Dice 损失函数,以及将它们结合起来的混合损失函数。

二元交叉熵损失是在二分类任务中常用的损失函数,用于衡量模型预测值与真实标签之间的差异。其定义为:

$$BCEloss = -\frac{1}{N} \sum_{i=1}^N [y_i \log(p_i) + (1-y_i) \log(1-p_i)] \quad (5)$$

式中: N 是图像中像素的个数; y_i 是第 i 个像素的真实标签,背景为 0,目标为 1; p_i 是第 i 个像素预测为目标的概率值,在 $[0,1]$ 区间。

二元交叉熵损失函数对所有像素的预测结果进行评估,对所有对于类别不平衡的情况非常敏感,会偏向于像素数量较多的类别,陷入局部最优解^[38],从而导致模型在少数类别上的表现不佳。在虹膜块状特征的分割中,由于前景(虹膜块状特征)和背景(非虹膜块状特征)的比例不平衡,很可能导致对虹膜块状特征区域的分割的准确性较低。为平衡模型对前景和背景的关注度,增加正样本在计

算损失时的贡献,引入加权二元交叉熵损失函数,其定义为:

$$\text{Weighted BCEloss} = -\frac{1}{N} \sum_{i=1}^N [w_{pos} y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \quad (6)$$

式中: w_{pos} 为正样本权重,取值为 2。

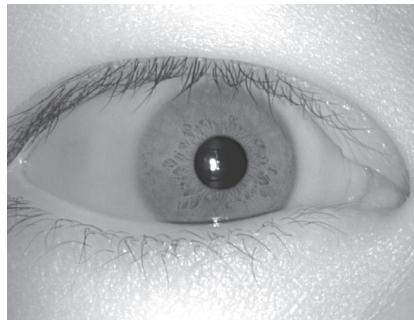
在类别不平衡的语义分割任务中,通常使用 Dice 损失函数作为模型的损失函数,其定义为:

$$\text{DiceLoss} = 1 - \frac{2 \times \sum_{i=1}^N y_i p_i + \epsilon}{\sum_{i=1}^N y_i + \sum_{i=1}^N p_i + \epsilon} \quad (7)$$

式中: ϵ 为防止分母为 0 而设置的调节值,取值 10^{-6} 。

Dice 损失关注预测结果与真实标签的重叠相似程度,通过 Dice 系数来优化网络,能够较好的平衡背景和前景不平衡的情况。但 Dice 损失在本质上是不稳定的,在梯度计算时容易出现剧烈抖动,且对边界分割效果不佳^[39]。因此,本文综合考虑两个损失函数的特点,提出一种混合损失函数,其定义为:

$$\text{Loss} = (1 - \lambda) \cdot \text{WeightedBCEloss} + \lambda \cdot \text{DiceLoss} \quad (8)$$



(a) 虹膜原始图像
(a) Original iris image



(b) 归一化后的图像(a)
(b) Normalized version of image (a)

图 6 虹膜图像的归一化

Fig. 6 Iris image normalization

归一化后的虹膜图像的灰度值分布集中在一个狭窄的范围内,导致图像的对比度较低,细节信息不明显,不利于虹膜块状特征的分割。为了克服这一问题,采用了直方图均衡化的方法来增强虹膜图像。直方图均衡化是一种广泛应用的图像增强技术,它通过调整图像的灰度值分布,使其较为均匀地分布在整个灰度范围内,从而增强图像的对比度^[41]。

在进行直方图均衡化时,首先计算图像的直方图,然后通过累计直方图计算出每个灰度级的累计分布函数,接下来将累计分布函数映射到新的灰度范围内,最后根据映射结果更新图像的灰度值。经过直方图均衡化处理后,虹膜图像的对比度得到了显著的提升,图像的细节变得更为清晰,突出了虹膜块状特征,如图 7 所示。

式中: λ 为权重因子,取值范围为 $[0, 1]$,用于调节加权二元交叉熵损失函数和 Dice 损失函数的比例。

1.3 图像预处理

本文所用数据是使用中科虹霸接触式虹膜采集仪 IKUSB-E30 采集的近红外光虹膜图像,大小为 $640 \text{ pixel} \times 480 \text{ pixel}$ 。在提取虹膜块状特征之前,需要对原始虹膜图像进行预处理操作,包括虹膜图像的归一化和直方图均衡化。

在采集虹膜的过程中,由于光照条件的不同,瞳孔会随着光照强度的变化而缩放,导致虹膜上的块状特征发生形变。同时,采集到的虹膜图像上还存在着大量如瞳孔、巩膜等非虹膜的干扰因素。因此,为了减弱纹理的形变、去除无关信息,需要对虹膜图像进行归一化处理。虹膜归一化是指通过对虹膜内外圆的定位,将分割后的圆环状虹膜展开成统一大小矩形图像的操作^[40]。

经归一化处理后,所有虹膜均统一到相同尺寸,去除了如瞳孔、巩膜等非虹膜区域,消除了由于瞳孔缩放引起的纹理变化,有助于虹膜块状特征的进一步分割提取和规律统计。原始虹膜图像分辨率为 $640 \text{ pixel} \times 480 \text{ pixel}$,归一化后的虹膜图像的分辨率为 $540 \text{ pixel} \times 70 \text{ pixel}$ 。虹膜图像的归一化操作如图 6 所示。



图 7 直方图均衡化效果

Fig. 7 Illustration of the impact of histogram equalization

2 实验

2.1 数据集的构建

本文采用监督学习的方法训练 RAA-UNet 语义分割模型,为此构建一个虹膜块状特征的标注数据集。本文使用 labelme 软件创建分割标签,并在预处理后的虹膜图像上手动标注虹膜块状特征,总计标注了 200 张虹膜图像,标注示例如图 8 所示。

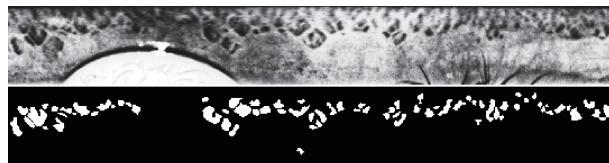


图 8 人工标注块状特征的图例

Fig. 8 An example of manually labeled block-shaped iris features

为了增加样本的多样性并降低人工标注成本,本文采取了以下数据增强策略。首先,对每张图像进行水平翻转、垂直翻转和水平垂直翻转,将样本数量扩增到原来的 4 倍。接着,将每张图像进行四等分切片,并重新排序以模拟虹膜原图旋转 90°、180° 和 270° 的效果,再次扩增 4 倍,如图 9 所示。因此,最终共得到了 3 200 张图像。为了进行训练、验证和测试,按照 8:1:1 的比例划分了训练集、验证集和测试集。

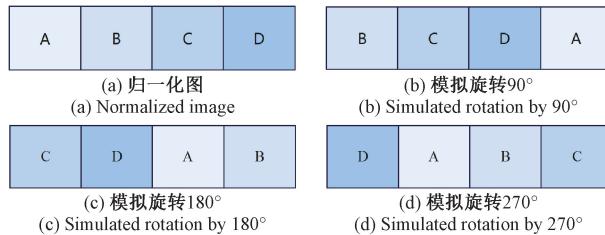


图 9 数据增强示意图

Fig. 9 Data enhancement diagram

2.2 实验环境和设置

实验平台的硬件设备为 Intel(R) Xeon(R) Gold 6346 CPU @ 3.10GHz, 内存 156 GB, GPU 是 NVIDIA A100, 内存 40 GB。软件配置为 Windows 10, CUDA11.4, 深度学习框架为 Pytorch, 版本 1.13.1+cu117。

本实验用于训练和测试的图像是预处理后的虹膜图像,大小为 540 pixel×70 pixel。虹膜图像是 BMP 格式,标注图像是 PNG 格式,训练期间通过随机裁剪、随机缩放、随机旋转进行数据增强。Batch size 设置为 16, 整个训练过程包括 50 个 epoch。使用 Adam 优化器,初始学习率设为 0.001, betas 设为 (0.9, 0.999)。

2.3 评价指标

在对语义分割的结果进行验证和评价时,不仅需要视觉感知上的定性评价,还需要客观的定量评价^[42]。为了准确评价模型的分割性能,本文从准确性和相关性两方面,使用了 6 个常用评价指标。其中准确性指标包括像素精度、召回率、精确率、F1 分数,相关性指标包括平均交并比和 Dice 系数。

准确性指标可以使用混淆矩阵来定义,如表 2 所示。对于虹膜块状特征分割任务,本文将虹膜块状特征设为正类,将非虹膜块状特征区域(背景区域)设为负类,从而分出真负例(true negative, TN),假正例(false positive, FP),假负例(false negative, FN),真正例(true positive, TP)。

表 2 混淆矩阵

Table 2 Confusion matrix

真实标签	模型预测	
	预测为负类 (背景)	预测为正类 (块状特征)
标签为负类(背景)	TN	FP
标签为正类(块状特征)	FN	TP

通过混淆矩阵,可以计算以下评价指标。其中,像素精度(pixel accuracy, PA)衡量了模型在整个图像上的预测准确性,它表示从预测正确的像素数量占总体的比例,计算公式如式(9)所示。召回率(Recall)表示所有真正为正类的样本中,被正确预测为正类的样本所占的比例,如式(10)所示。精确率(Precision)表示在预测为正类的结果中,真正例所占的比例,如式(11)所示。精确率和召回率都有一定的片面性,不能单独作为评价指标。F1 分数(F1 Score)是精确率和召回率的调和平均数,可用于综合评价模型的性能,如式(12)所示。

$$\text{Pixel Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (10)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (11)$$

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} = \frac{2TP}{2TP + FP + FN} \quad (12)$$

相关性指标指的是模型预测的分割区域与标签区域之间的相似程度,主要包括交并比(intersection over union, IoU)和 Dice 系数(Dice similarity coefficient, DSC)。交并比表示预测分割区域与真实标签区域之间的重叠程度,定义为预测分割区域与标签区域的交集面积除以它们的并集面积,如式(13)所示。平均交并比(mean intersection over union, mIoU)计算每个类别的 IoU 值的算术平均值,如式(14)所示。Dice 系数衡量预测分割区域与真实标签区域之间的相似度,被定义为预测分割区域与标签区域的两倍交集面积除以它们的总面积之和,如式(15)所示。

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|} \quad (13)$$

$$\text{mIoU} = \frac{1}{k+1} \sum_{k=0}^k \frac{|A_k \cap B_k|}{|A_k \cup B_k|} \quad (14)$$

$$\text{Dice similarity coefficient} = \frac{2 |A \cap B| + \epsilon}{|A| + |B| + \epsilon} \quad (15)$$

式中: k 表示前景种类数量; A 、 B 分别为标签图像和预测图像; ϵ 防止分母为 0 而设置的调节值,取值为 10^{-6} 。

2.4 实验结果与分析

1) 损失函数对分割的影响

为了探究损失函数对模型分割效果的影响,选取最佳损失函数,本文对混合损失函数中的权重因子 λ 进行了一系列实验。在实验过程中,将 λ 从 0~1 以 0.25 为步长进行变化,共得到 5 个不同的权重因子。针对 RAA-UNet 模型,进行了五组实验,每组实验都采用不同权重因子的损失函数进行训练,每组实验都重复训练 4 次,保存在验证集上取得最佳效果的模型权重。随后,在测试集上计算评价指标,并计算其均值,最终绘制得到了 Dice 系数随 λ 变化的折线图,如图 10 所示。

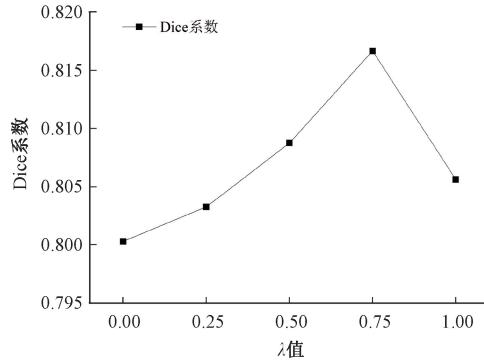


图 10 Dice 系数随 λ 变化趋势

Fig. 10 Trend of Dice coefficient with λ

可以看出,Dice 系数随 λ 的增大先增加再减少,在 $\lambda=0$ 时最低,仅为 0.8003,在 $\lambda=0.75$ 时最高,为 0.8166。这说明在虹膜块状特征分割任务中,单独使用加权二元交叉熵损失函数或者单独使用 Dice 损失效果均不佳,而结合加权二元交叉熵损失和 Dice 损失的混合损失函数能有效提高模型的分割性能。因此,本文将混合损失函数的 λ 定为 0.75,并以此进行后续的实验。

2) 直方图均衡化对分割的影响

为了探究直方图均衡化对模型分割效果的影响,使用未经直方图均衡化的标注数据集和均衡化后的标注数据集分别进行 RAA-UNet 模型的训练,两组实验都重复训练 4 次,保存在验证集上取得最佳效果的模型权重。随后,在测试集上计算评价指标,结果如表 3 所示。

表 3 直方图均衡化对分割的影响

Table 3 The impact of histogram equalization on segmentation

预处理操作	像素精度	召回率	精确率	F1 分数	平均交并比	Dice 系数
未处理	97.66	66.55	85.49	74.83	78.68	74.26
预处理	98.09	86.79	78.99	82.64	84.21	81.66

注:加粗字体为每列最优值。

由表 3 可知,与使用直方图均衡化图像进行训练相比,使用原图进行训练的结果精确率略高,但召回率很低。

这说明此时模型只能识别出最明显的正类样本,虽然识别出的样本大多数是正确的,但很多正类样本被漏掉了。精确率和召回率只能反映模型的某一方面,具有片面性,所以实际分割效果需要更多地参考能体现综合性能的 F1 分数以及相关性指标(mIoU 和 Dice 系数)。使用直方图均衡化后的图像进行训练的模型,其 F1 分数、mIoU 和 Dice 系数都有了明显的提升。

实验结果表明,均衡化后的图像增强了模型的召回率和综合性能,提高了模型对目标区域的识别能力和分割效果。因此,直方图均衡化作为一种预处理方法,能够有效提高 RAA-UNet 模型分割虹膜块状特征的性能表现。

3) 对比实验

为了验证本文提出的 RAA-UNet 模型在虹膜块状特征分割任务中的有效性和优势,本文选择了 FCN、LRASPP、DeepLabV3、DeepLabV3+、SegFormer、UNet 和 U2Net 作为对比模型。这些模型代表了当前语义分割领域中不同架构和技术的主要进展,涵盖了经典的卷积神经网络及基于 Transformer 的网络,并且都在各类分割任务中取得了良好的表现。FCN、DeepLabV3 和 DeepLabV3+ 均采用了 ResNet-50 作为特征提取的主干网络,LRASPP 采用了 MobileNetV3 作为主干网络,并均使用了 PyTorch 官方提供的预训练权重进行初始化。SegFormer 采用了 MiT-b0 和 MiT-b5 两种主干网络,并使用了预训练权重进行初始化。

对所有模型分别进行了 4 次重复训练,每次都在验证集上选择效果最佳的权重,并在测试集上进行评估,最后取四次测试的均值,结果如表 4 所示。从表 4 可以看出,基于 U 形结构的 Unet、U2Net 和 RAA-UNet 在各项指标中均明显优于其他模型。这说明编码器-解码器的基础架构在虹膜分割任务中具有明显的优势。并且,本文提出的 RAA-UNet 模型在所有指标上都优于 UNet 和 U2Net。以具备代表性的 Dice 为例,UNet 为 77.79%,U2Net 为 80.82%,而 RAA-UNet 则为 81.66%,较 UNet 提升了 3.87%,比 U2Net 高 0.84%。其他模型的 Dice 系数均低于 70%。

为了更全面地比较各个模型,本文还使用了 ROC 曲线和 PR 曲线来可视化其性能,如图 11 和 12 所示。从图 11 可知,在 ROC 曲线的分布中,RAA-UNet 最靠近左上角。RAA-UNet 模型 AUC 值为 0.9945,高于其他模型。这表明 RAA-UNet 在平衡真阳性率和假阳性率方面的整体性能最佳。从图 12 可知,RAA-UNet 的 PR 曲线最靠近右上角,AUPRC 为 0.9172,高于其他模型,在所有模型中表现最佳。

为了进一步分析九个模型的实际分割效果,本文对比了其在测试集上的实际分割结果,如图 13 所示。图 13 展示了典型样本的分割结果,包括输入图像、人工标注图以及九个模型的分割结果。图 13(a)和(b)分别为输入图像、

人工标注图,图13(c)~(k)分别为FCN、LRASPP、DeepLabV3、DeepLabV3+、SegFormer_b0、SegFormer_

b5、UNet、U2Net、RAA-UNet等模型对图13(a)的预测结果分割图。

表4 不同模型评价指标对比结果

Table 4 Comparison results of evaluation indicators for different models

模型	像素精度	召回率	精确率	F1分数	平均交并比	Dice系数
FCN ^[28]	94.81	70.23	50.54	58.68	68.07	58.10
LRASPP ^[43]	94.77	65.05	50.15	56.60	67.03	55.81
DeepLabV3 ^[30]	95.12	64.37	53.20	58.00	67.90	57.57
DeepLabV3+ ^[31]	96.34	70.63	63.65	66.89	73.22	66.29
SegFormer_b0 ^[33]	96.93	61.15	75.69	67.50	73.90	67.18
SegFormer_b5 ^[33]	95.98	76.09	59.04	66.46	72.79	65.58
UNet ^[29]	97.54	84.07	73.26	78.22	80.84	77.79
U2Net ^[35]	98.00	85.50	78.42	81.75	83.52	80.82
RAA-UNet	98.09	86.79	78.99	82.64	84.21	81.66

注:加粗字体为每列最优值。

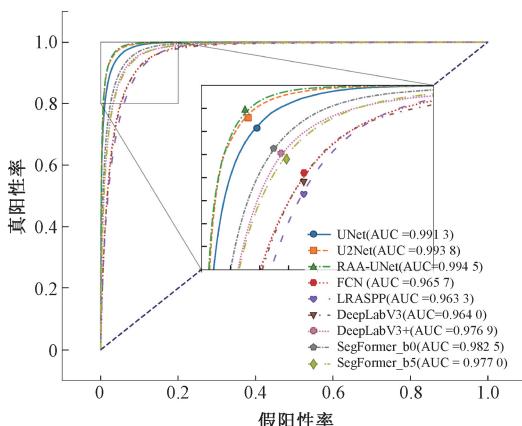


图11 不同模型的ROC曲线对比

Fig. 11 Comparison of ROC curves of different models

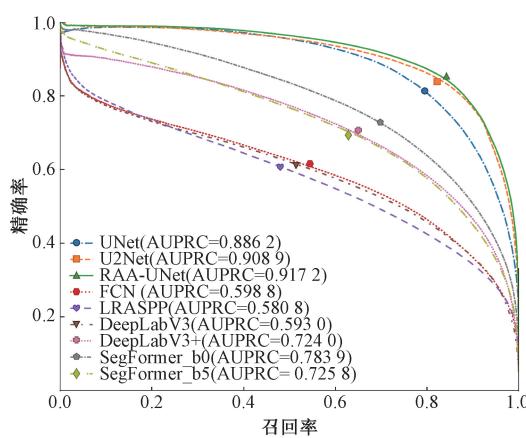
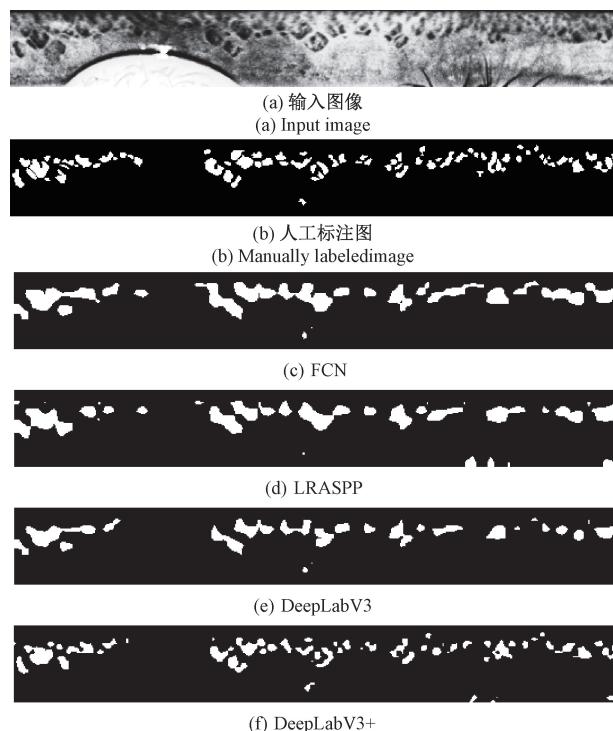


图12 不同模型的PR曲线对比

Fig. 12 Comparison of PR curves of different models

从图中可以看出,FCN、LRASPP和DeepLabV3在分

割虹膜块状特征时表现得十分粗糙,只能大致描绘出特征的附近区域,而未能精确捕捉到虹膜块状特征的细致结构。虽然FCN和DeepLabV3使用了Resnet50作为特征提取的主干网络,LRASPP使用MobileNetV3作为主干网络,均具备了较为强大的特征提取能力,但其在提取图像的深层特征时进行了多次下采样,缩小了特征图的尺寸,会导致丧失部分细节信息。在上采样的过程中,3种模型均直接采用了8倍上采样,将图像尺寸直接扩大8倍,这可能导致模型无法准确地恢复目标的细节信息。因此这3种模型不适合处理形状较小,数量较多的目标的分割任务。



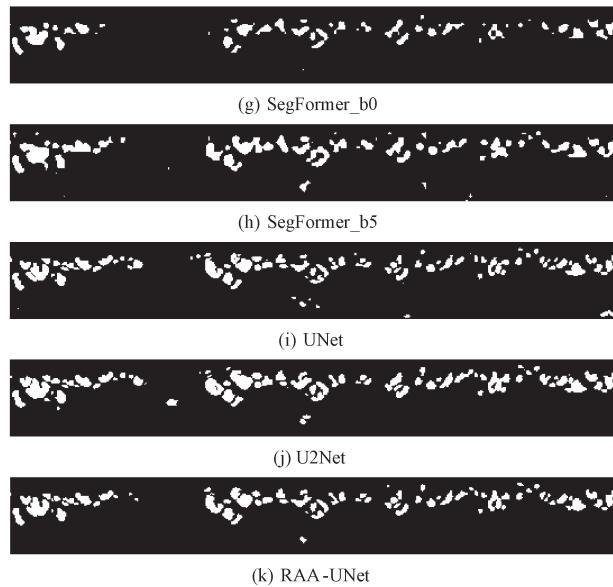


图 13 不同模型的分割效果对比

Fig. 13 Comparison of segmentation effects of different models

DeepLabV3+的分割效果较 DeepLabV3 有了一定的提升。这可能是 DeepLabV3+在 DeepLabV3 的基础上增加了解码器结构,通过跳跃连接将浅层特征与上采样后的深层特征进行拼接,然后再进行四倍上采样恢复到原始尺寸,较直接八倍上采样能恢复更多的空间细节。SegFormer 在虹膜块状特征的实际分割效果不佳,能够大致分割出目标,但难以捕捉到精细的边缘信息。在预测过程中,SegFormer 将 MiT 主干网络中的 32、16、8 倍和 4 倍下采样特征通过 MLP 处理,并上采样到 1/4 原尺寸后进行拼接,再经两个 MLP 层分别融合特征和预测掩膜,最终通过四倍上采样恢复到原有尺寸。四倍上采样会导致部分信息的损失。实际上,本文也尝试通过更改模型下采样倍率,将 SegFormer_b0 的下采样倍率调整为 1、2、4 倍和 8 倍进行训练,分割效果显著提升,Dice 系数达到了 80.50%。然而,这一结果仍低于本文所提出模型的 81.66%。

本文所提出的模型采用了特殊的“U”型网络结构,在编码器和解码器之间添加了跳跃连接,将同一尺度的下采样特征与上采样特征进行拼接。这种设计既保留了全局

的语义信息,又能够恢复细节丰富的空间信息,适合处理形状较小,数量较多的物体。进一步对比 RAA-UNet、UNet 和 U2Net 的分割图像,发现三者的整体分割效果整体接近,但在细节方面仍存在差异。UNet 存在易分割出噪点等干扰因素、边缘不清晰的不足,U2Net 也存在边缘不清晰、彼此粘连的情况。而 RAA-UNet 的分割结果与人工标注标准最为接近,对块状特征的分割最为准确,边界也最为平滑和精细。

4) 分割结果分析

进一步对 RAA-UNet 的分割结果进行分析。从评价指标上看,模型在测试集上的 F1 分数为 82.64%,说明其能较好地平衡精确率和召回率,有很强的分类能力;Dice 系数为 81.66%,说明模型预测结果与人工标注之间的重叠程度较高;观察 RAA-UNet 的实际分割图,可以看出其与人工标注图非常接近,且在边界上分割也非常精细。使用 RAA-UNet 进行虹膜块状特征的分割,其目的在于将块状特征有效的分割出来,辅助人工进行鉴定,以及作为研究块状特征分布规律的技术基础。因此,还需要从法庭科学的视角出发,在实例层面分析虹膜块状特征的漏检和错检。

对比测试集中的人工标注图和本文模型预测图像,可以发现人工标注的块状特征绝大部分都能被模型检测出来,漏检的块状特征的数量非常少。漏检的原因可能是这些块状特征与背景的对比度低,较难区分。从实例的角度分析错检,可以发现预测图像中错误检出的区域数量很少,且大多形状很小,对应区域在原始图像中属于颜色较深的噪点。这些噪点在鉴定时价值很小,一般不认为是虹膜块状特征,人工在进行标注时忽略,但会被模型分割出来。整体来看,RAA-UNet 可以将虹膜块状特征准确地分割出来。

5) 消融实验

RAA-UNet 模型在 UNet 的基础上将编码器和解码器的卷积模块改为残差模块,将解码器上采样方式由转置卷积更换为双线性插值,在跳跃连接中加入了注意力门模块,并在编码器和解码器的末端分别引入了 ASPP 模块。为了探究每种改进对模型性能的提升效果,本文以 UNet 为基准模型,进行了大量的消融实验。每种模型均训练 4 次,评价指标取均值,实验结果如表 5 所示。

表 5 消融实验对比结果

Table 5 Comparison results of ablation experiments

模型	像素精度	召回率	精确率	F1 分数	平均交并比	Dice 系数
UNet	97.54	84.07	73.26	78.22	80.84	77.79
UNet+残差模块	97.83	86.73	75.56	80.72	82.70	79.93
UNet+残差模块+双线性插值	97.86	86.76	75.86	80.94	82.87	80.17
UNet+残差模块+双线性插值+注意力门模块	97.96	86.61	77.32	81.67	83.44	80.89
UNet+残差模块+双线性插值+注意力门模块+ASPP 模块 (RAA-UNet)	98.09	86.79	78.99	82.64	84.21	81.66

注:加粗字体为每列最优值。

相对于基线模型UNet,加入残差模块后的UNet在像素准确率、精确率、召回率、特异率、F1分数、mIoU和Dice系数等多个指标上均有提升,表明残差模块能有效提升特征提取能力。采用双线性插值上采样后,在减少了参数量的同时,各项指标略有提高。在增添注意力机制后,虽然召回率略有下降,但精确率明显提升,其他指标也均有提高。使用ASPP模块后,各指标再次提升,各项指标分数均为最高,达到最优效果。

Dice系数是医学图像分割任务中最常用的评价指标。单从Dice系数来看,UNet的Dice系数仅为77.79%,而采用了“残差模块”,“残差模块+双线性插值”,“残差模块+双线性插值+注意力门模块”以及“残差模块+双线性插值+注意力门模块+ASPP模块”的改进UNet分别为79.93%,80.17%,80.89%和81.66%,均有了较为明显的提升。从整体趋势来看,随着模型结构不断完善,各指标的数值呈现逐步提升态势。

本文还使用了ROC曲线和PR曲线对采用不同模块的模型进行了评价,如图14和15所示。

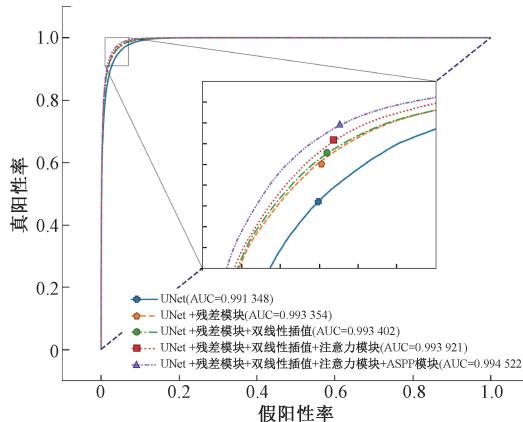


图14 消融实验ROC曲线对比

Fig.14 Comparison of ROC curves of ablation experiments

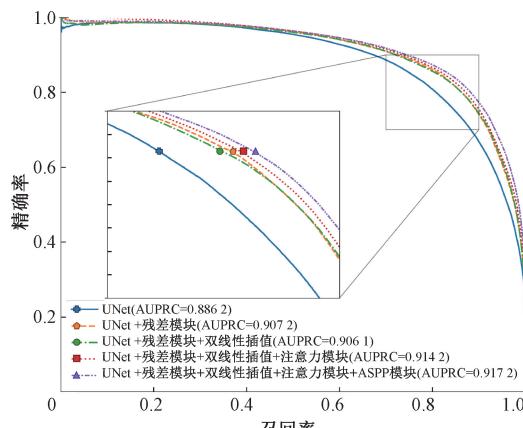


图15 消融实验PR曲线对比

Fig.15 Comparison of PR curves of ablation experiments

ROC和PR曲线提供了对模型性能的一种比较直观的视觉展示。从ROC曲线图中可以看出,随着模型模块的增加,AUC值逐步变大,且ROC曲线更向左上角凸起;从PR曲线中可以看出随着模型结构的改进,AUPRC值整体呈现增加趋势,且PR曲线更向右上角凸起。

可见,随着模型结构的改进,模型的性能逐渐提高,证明了每项改进都有其必要性。本文所提出的RAA-UNet模型优于UNet和其他模块组合模型。

3 结 论

针对虹膜鉴定中的虹膜块状特征提取问题,本文提出了基于RAA-UNet的虹膜块状特征分割方法并建立了虹膜块状特征标注数据集。RAA-UNet在UNet基础上更换了残差模块,引入残差连接以促进梯度传播并提高特征的传递能力;添加注意力机制,使模型专注于虹膜块状特征,抑制无关信息;在模型的中部和末端添加ASPP模块,提高模型的感受野并提取多尺度特征;选用双线性插值进行上采样以避免棋盘伪影并提高计算效率。在训练过程中,采用加权二元交叉熵和Dice系数构建混合损失函数,结合两种损失函数优势提高模型训练效果。

对比实验结果表明,RAA-UNet的各项评价指标均明显优于本文所对比的其他语义分割模型,并且相较于UNet在F1分数提升了4.42%,mIoU提升了3.37%,Dice系数提升了3.87%。相较于其他模型,RAA-UNet的分割效果更为准确和精细,更不易受到睫毛、背景噪声等干扰因素的影响,且分割图像的界限更为清晰。消融实验结果表明,每一步优化都有必要性,均提高了模型分割的性能表现。综合而言,RAA-UNet能够实现虹膜块状特征的准确分割,可为虹膜鉴定的进一步研究提供技术支撑。

参考文献

- [1] 孙哲南,赫然,王亮,等.生物特征识别学科发展报告[J].中国图象图形学报,2021,26(6): 1254-1329. SUN ZH N, HE R, WANG L, et al. Overview of biometrics research [J]. Journal of Image and Graphics, 2021, 26(6): 1254-1329.
- [2] DAUGMAN J. Information theory and the iriscode [J]. IEEE Transactions on Information Forensics and Security, 2016, 11(2): 400-409.
- [3] 周千里,王雅丽,徐雪婧,等.城市级大规模公安虹膜采集识别应用体系[J].计算机应用与软件,2021,38(8): 7-11,77. ZHOU Q L, WANG Y L, XU X J, et al. Application system of iris acquisition and recognition in large-scale public security at city level[J]. Computer Applications and Software, 2021, 38(8): 7-11,77.
- [4] BANERJEE S. Aadhaar: Digital inclusion and public services in india [R]. World Development Report, 2016.

- [5] LEE M B, KANG J K, YOON H S, et al. Enhanced iris recognition method by generative adversarial network-based image reconstruction [J]. IEEE Access, 2021, 9: 10120-10135.
- [6] HE Y Q, FENG G Q, HOU Y SH, et al. Iris feature extraction method based on LBP and chunked encoding [C]. 2011 Seventh International Conference on Natural Computation. Piscataway: IEEE, 2011: 1663-1667.
- [7] RATHGEB C, WAGNER J, BUSCH C. Sift-based iris recognition revisited: prerequisites, advantages and improvements [J]. Pattern Analysis and Applications, 2019, 22: 889-906.
- [8] NGUYEN K, PROENÇA H, ALONSO-FERNANDEZ F. Deep learning for iris recognition: A survey[J]. ACM Computing Surveys, 2024, 56(9): 1-35.
- [9] NGUYEN K, FOOKES C, ROSS A, et al. Iris recognition with off-the-shelf CNN features: A deep learning perspective [J]. IEEE Access, 2017, 6: 18848-18855.
- [10] GANGWAR A, JOSHI A. DeepIrisNet: Deep iris representation with applications in iris recognition and cross-sensor iris recognition [C]. 2016 IEEE International Conference on Image Processing (ICIP). Piscataway: IEEE, 2016: 2301-2305.
- [11] GAO R Q, BOURLAI T. On designing a swiniris transformer based iris recognition system[J]. IEEE Access, 2024, 12: 30723-30737.
- [12] ZHAO Z J, KUMAR A. Towards more accurate iris recognition using deeply learned spatially corresponding features[C]. 2017 IEEE International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2017: 3809-3818.
- [13] NGUYEN K, FOOKES C, SRIDHARAN S, et al. Complex-valued iris recognition network [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 45(1): 182-196.
- [14] 刘笑楠. 虹膜图像块状纹理检测方法研究[D]. 沈阳: 沈阳工业大学, 2014.
- LIU X N. Iris plaque-like textures detection [D]. Shenyang: Shenyang University of Technology, 2014.
- [15] CHEN J X, SHEN F, CHEN D Z, et al. Iris recognition based on human-interpretable features[J]. IEEE Transactions on Information Forensics and Security, 2016, 11(7): 1476-1485.
- [16] 陈冲, 陈杰, 张慧, 等. 深度学习可解释性综述[J]. 计算机科学, 2023, 50(5): 52-63.
- CHEN CH, CHEN J, ZHANG H, et al. Review on interpretability of deep learning [J]. Computer Science, 2023, 50(5): 52-63.
- [17] 陈子龙, 秦旗, 胡文锋, 等. 虹膜人工特征选择和标注方法的初步研究[J]. 刑事技术, 2021, 46(3): 221-227.
- CHEN Z L, QIN Q, HU W F, et al. Tentative exploration into artificial selection and labeling of iris feature[J]. Forensic Science and Technology, 2021, 46(3): 221-227.
- [18] 张翌阳, 唐云祁. 虹膜鉴定特征研究[J]. 中国公安大学学报(自然科学版), 2021, 27(2): 7-14.
- ZHANG Y Y, TANG Y Q. Research on iris identification features[J]. Journal of People's Public Security University of China (Natural Science Edition), 2021, 27(2): 7-14.
- [19] SHEN F, FLYNN P J. Iris matching by crypts and anti-crypts [C]. 2012 IEEE Conference on Technologies for Homeland Security (HST). Piscataway: IEEE, 2012: 208-213.
- [20] SHEN F, FLYNN P J. Are iris crypts useful in identity recognition? [C]. 2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS). Piscataway: IEEE, 2013: 1-6.
- [21] SHEN F, FLYNN P J. Iris crypts: Multi-scale detection and shape-based matching[C]. IEEE Winter Conference on Applications of Computer Vision. Piscataway: IEEE, 2014: 977-983.
- [22] 刘笑楠, 苑玮琦, 张波. 基于组合窗口搜索的虹膜块状纹理检测[J]. 仪器仪表学报, 2014, 35(8): 1900-1906.
- LIU X N, YUAN W Q, ZHANG B. Iris plaque-like texture detection based on combined window searching[J]. Chinese Journal of Scientific Instrument, 2014, 35(8): 1900-1906.
- [23] 苑玮琦, 朱立军. 基于线性模板和模糊规则的虹膜坑洞纹理检测[J]. 仪器仪表学报, 2015, 36(6): 1363-1371.
- YUAN W Q, ZHU L J. Iris crypt texture detection based on linear template and fuzzy rules[J]. Chinese Journal of Scientific Instrument, 2015, 36(6): 1363-1371.
- [24] 朱立军, 苑玮琦. 基于双线性模板和分块策略的虹膜色素斑检测[J]. 仪器仪表学报, 2015, 36(12): 2714-2721.
- ZHU L J, YUAN W Q. Iris pigmented spot detection based on bilinear templates and block partitioning strategy[J]. Chinese Journal of Scientific Instrument, 2015, 36(12): 2714-2721.
- [25] 朱立军, 苑玮琦. 基于 BAB 策略与 SVM 的虹膜块状纹理检测[J]. 仪器仪表学报, 2016, 37(11): 2438-2444.

- ZHU L J, YUAN W Q. Iris image lump-like texture detection based on BAB strategy and SVM [J]. Chinese Journal of Scientific Instrument, 2016, 37 (11): 2438-2444.
- [26] 苑玮琦, 朱立军, 张波. 基于形态学与支持向量机的虹膜坑洞纹理检测[J]. 仪器仪表学报, 2017, 38(3): 664-671.
- YUAN W Q, ZHU L J, ZHANG B. Iris crypt texture detection based on morphology and SVM[J]. Chinese Journal of Scientific Instrument, 2017, 38(3): 664-671.
- [27] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [C]. 25th Neural Information Processing Systems (NIPS). Cambridge: MIT Press, 2012: 1097-1105.
- [28] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2015: 3431-3440.
- [29] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation [C]. 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2015: 234-241.
- [30] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation[EB/OL]. (2017-12-05) [2024-09-27]. <https://arxiv.org/abs/1706.05587>.
- [31] CHEN L C, ZHU Y K, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]. 15th European Conference on Computer Vision (ECCV). Cham: Springer, 2018: 801-818.
- [32] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16×16 words: Transformers for image recognition at scale[EB/OL]. (2021-06-03) [2024-09-27]. <https://arxiv.org/abs/2010.11929>.
- [33] XIE EN Z, WANG W H, YU ZH D, et al. SegFormer: Simple and efficient design for semantic segmentation with transformers [J]. Advances in Neural Information Processing Systems, 2021, 34: 12077-12090.
- [34] OKTAY O, SCHLEMPER J, FOLGOC L L, et al. Attention U-Net: Learning Where to look for the pancreas [EB/OL]. (2018-5-20) [2024-09-27]. <https://arxiv.org/abs/1804.03999>.
- [35] QIN X B, ZHANG Z CH, HUANG CH Y, et al. U2-Net: Going deeper with nested U-structure for salient object detection [J]. Pattern Recognition, 2020, 106: 107404.
- [36] 刘拥民, 张毅, 欧阳凌轩, 等. 基于改进 U-Net 的轻量级眼底病变分割算法设计[J]. 电子测量技术, 2024, 47(3): 127-134.
- LIU Y M, ZHANG Y, OUYANG L X, et al. Design of light weight fundus lesion segmentation algorithm based on improved U-Net[J]. Electronic Measurement Technology, 2024, 47(3): 127-134.
- [37] HE K M, ZHANG X Y, REN SH Q, et al. Deep residual learning for image recognition[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2016: 770-778.
- [38] 陈英, 张伟, 林洪平, 等. 医学图像分割算法的损失函数综述[J]. 生物医学工程学杂志, 2023, 40(2): 392-400.
- CHEN Y, ZHANG W, LING H P, et al. A survey of loss function of medical image segmentation algorithms[J]. Journal of Biomedical Engineering, 2023, 40(2): 392-400.
- [39] ZHANG Y, LIU SH J, LI CH L, et al. Rethinking the dice loss for deep learning lesion segmentation in medical images [J]. Journal of Shanghai Jiaotong University(Science), 2021, 26(1): 93-102.
- [40] MALGHEET J R, MANSHOR N B, AFFENDEY L S. Iris recognition development techniques: A comprehensive review[J]. Complexity, 2021, 2021: 1-32.
- [41] AGRAWAL S, PANDA R, MISHRO P K, et al. A novel joint histogram equalization based image contrast enhancement [J]. Journal of King Saud University-Computer and Information Sciences, 2022, 34 (4): 1172-1182.
- [42] 于营, 王春平, 付强, 等. 语义分割评价指标和评价方法综述[J]. 计算机工程与应用, 2023, 59(6): 57-69.
- YU Y, WANG CH P, FU Q, et al. Survey of evaluation metrics and methods for semantic segmentation [J]. Computer Engineering and Applications, 2023, 59(6): 57-69.
- [43] HOWARD A, SANDLER M, CHU G, et al. Searching for mobilenetv3 [C]. 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2019: 1314-1324.

作者简介

陈宇, 硕士研究生, 主要研究方向为电子数据检验。

E-mail: 1019081653@qq.com

唐云祁(通信作者), 博士, 教授, 主要研究方向为电子数据检验和智能图像识别。

E-mail: tangyunqi@ppsuc.edu.cn