

融合注意力机制与密集多尺度特征的异质噪声虹膜分割方法

尤轩昂^{1*}, 赵鹏^{1**}, 慕晓冬¹, 白坤¹, 练赛²

¹火箭军工程大学作战保障学院, 陕西 西安 710025;

²西安交通大学微电子学院, 陕西 西安 710049

摘要 针对当前分割方法不能兼顾可见光与近红外光谱的异质噪声虹膜图像的分割精度与效率问题, 提出一种基于注意力机制与密集多尺度特征融合的编-解码网络。首先, 引入基于深度可分离卷积的改进残差瓶颈单元, 降低参数量与计算量的同时防止信息丢失与梯度混淆; 其次, 改进密集空洞空间金字塔模块的空洞率组合并放置于编码器后, 以增强多尺度特征融合; 最后, 设计了一种高效并行空间-通道注意力模块, 将该模块添加至每个下采样层与解码器之后, 提升对噪声目标与虹膜像素的分辨能力。在 3 个公开虹膜数据集上的实验结果表明: 所设计网络的平均 F1 分数与平均交并比 (mIoU) 均优于已有算法; 空间占用、参数量、计算量分别比基准网络减少 41%、41.77%、65.35%; 有效改善了对多个光谱噪声虹膜的分割性能且更易部署于移动设备, 能够更加高效、精确地分辨噪声与虹膜目标。

关键词 图像处理; 虹膜分割; 可见光; 近红外; 特征融合

中图分类号 TP391

文献标志码 A

doi: 10.3788/LOP202259.0410006

Heterogeneous Noise Iris Segmentation Based on Attention Mechanism and Dense Multiscale Features

You Xuanang^{1*}, Zhao Peng^{1**}, Mu Xiaodong¹, Bai Kun¹, Lian Sai²

¹College of Operational Support, Rocket Force University of Engineering, Xi'an, Shaanxi 710025, China;

²College of Microelectronics, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China

Abstract Current segmentation methods do not account for the segmentation accuracy and efficiency of visible and near-infrared heterogeneous noisy iris images; thus, in this study, we propose a codec network based on the fusion of attention mechanism and dense multiscale features. First, an improved residual bottleneck element based on deep separable convolution was introduced to reduce the number of parameters and computation while preventing information loss and gradient confusion. Second, the dense void space pyramid module's void rate combination was improved and placed behind the encoder to improve multiscale feature fusion. Finally, to improve the resolution of noise targets and iris pixels, an efficient parallel space-channel attention module was designed and integrated into each down sampling layer and decoder. The experiments conducted on three open iris data sets show that both the average F1-score and mean intersection over union (mIoU) of the proposed network are superior to the existing algorithms. Compared with the benchmark network, the occupied space, number of parameters, and amount of computation are reduced by 41%, 41.77%, and 65.35%, respectively. It can effectively improve the segmentation

收稿日期: 2021-01-27; 修回日期: 2021-03-13; 录用日期: 2021-03-25

基金项目: 国家自然科学基金(61403399)、陕西省自然科学基金(2020JM358, 2016JM6061)

通信作者: *youxuanang@163.com; **zpxhh@163.com

performance of the network for multispectral noise iris and is easier to be deployed on mobile devices, which can more efficiently and accurately distinguish noise and iris targets.

Key words image processing; iris segmentation; visible light; near-infrared; feature fusion

1 引言

虹膜识别是一种安全可靠的非接触身份识别技术,广泛应用于金融支付、智能解锁、出入境管控等领域。虹膜区域的精确分割作为虹膜识别中的关键环节,对后续特征提取和识别的性能有着重要影响^[1]。低用户配合度的移动端虹膜识别场景下采集的图像尺度不一且存在模糊、睫毛遮挡、反射光斑等不良噪声。针对移动端场景下的异质光谱(可见光与近红外)虹膜图像,研究更精确鲁棒的噪声虹膜分割给虹膜识别带来了新的机遇与挑战。

目前虹膜分割可分为传统算法和基于深度学习的方法。传统算法最早可追溯至 20 世纪 90 年代,1993 年, Daugman^[2]提出一种基于积分-微分算子计算虹膜半径方向上灰度值梯度变化的边界定位算法。随后 1997 年, Wildes^[3]提出一种基于 Laplace 算子和 Hough 圆形变换的虹膜分割算法。之后的很多算法都是基于这两种方法进行改进和创新的,如基于聚类的边界定位^[4]、全变分模型(RTV-L)^[5]、测地线活动轮廓(GACs)模型^[6]等。然而以上传统算法在噪声虹膜图像上分割不够精细且规则复杂。

虹膜分割本质上属于二分类语义分割问题,近年来,得益于深度学习特别是卷积神经网络的迅速发展,许多网络被迁移或改进后用于虹膜分割。Shelhamer 等^[7]将分类网络 VGG 的全连接层替换为卷积层用于特征提取,提出全卷积网络(FCN)。输入图像经下采样得到尺寸较小的特征图,再使用反卷积恢复至原始尺寸,最终得到逐像素预测结果。由此, Tan 等^[8]提出多尺度全卷积网络(MFCN),该网络通过融合不同深度的特征来获取虹膜的细节纹理信息与全局结构约束,提升了分割精度,但在噪声较多的困难样本上泛化性能不佳。Hao 等^[9]采用贝叶斯近似改进 SegNet 并将其用于虹膜分割。该方法在预测阶段采用基于蒙特卡罗采样的 Dropout 操作,最终产生具有像素类别标签的后验分布与分割图,提升了可见光虹膜分割精度。Ronneberger 等^[10]设计了具有堆叠卷积层与最大池化层的医学图像分割网络 U-Net。Lian 等^[11]提出了

一种注意力机制虹膜分割网络 U-Net。该网络在编码阶段加入边界框回归子网络以生成掩模权重,并通过拼接来融合特征图使其更侧重学习虹膜特征,有效利用了先验几何信息来提高分割精度。综上,基于深度学习的分割方法改善了传统噪声虹膜分割中的错分、漏分现象,提升了精度和鲁棒性。然而,现有虹膜分割网络存在参数搜索空间大、计算冗余和噪声敏感等问题,训练好的模型大小动辄几百 MB 甚至 GB 级别,难以部署于低功耗、低算力的移动端边缘计算设备。

基于上述讨论,本文设计了一个新的分割网络来解决复杂识别场景下的虹膜精确定位。整体设计框架采用 U-Net 的编-解码结构,先在网络各层嵌入基于深度可分离卷积的改进残差瓶颈单元,在减少冗余计算量和参数数量的同时防止梯度消失,保证分割精度。对密集空洞空间金字塔池化模块进行改进,并添加在网络的瓶颈部分以融合多尺度上下文特征,使噪声目标分割更精细、平滑。针对通道注意力机制中小目标特征信息学习不足的问题,设计了高效空间-通道并行注意力机制,有效提高了网络的特征表示能力和训练速度。最后,在可见光与近红外光谱的 3 个虹膜数据集上对所提算法和现有算法进行对比,设计了未采用本文策略的基准网络与消融实验,以证明所提算法可改善分割精度并显著降低参数量与计算量。

2 融合注意力机制与密集多尺度特征的虹膜分割网络

2.1 模型分析与概述

基于语义分割的虹膜精确定位是基于计算机视觉技术的,采用边缘端的嵌入式设备拍摄采集眼部图像,对图像进行预处理与数据增强后输入到预先训练好的神经网络中,实现对虹膜区域像素级预测。低用户配合度场景复杂,如图 1 所示,其中包括注视偏移、虹膜缺失(闭眼)、毛发遮挡、虹膜旋转、模糊、镜面反光等干扰因素。

眼部图像包含皮肤、毛发、巩膜、虹膜、瞳孔 5 种不同的区域。巩膜区域颜色较浅,瞳孔区域与毛发颜色较深,皮肤部分的颜色介于中间,虹膜呈现出

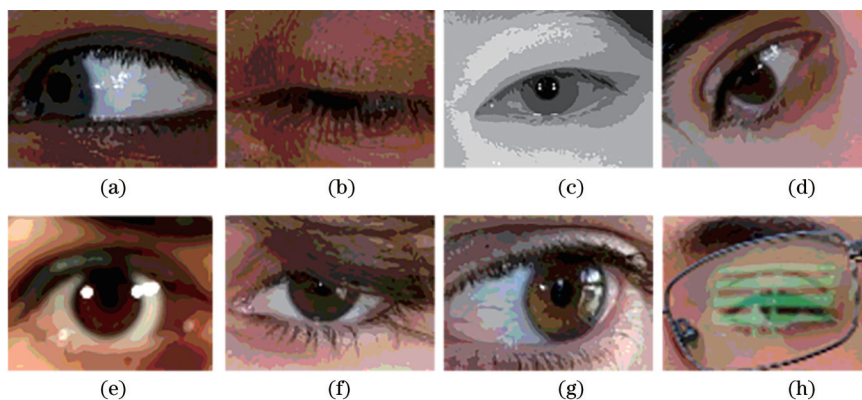


图 1 复杂场景下采集的眼部图像示例。(a) 注视偏移; (b) 虹膜缺失; (c) 睫毛遮挡; (d) 虹膜旋转; (e) 模糊; (f) 头发阴影; (g) 镜面反射; (h) 眼镜遮挡

Fig. 1 Examples of eye image acquisition in complex scenes. (a) Gaze deviation; (b) absence of iris; (c) eyelash occlusion; (d) iris rotation; (e) blur; (f) hair shade; (g) specular reflection; (h) glasses occlusion

独特的颜色与纹理特征。若为用户高度配合下采集的眼部图像,其虹膜区域较为干净,不同目标界限分明,亮度适中,不存在噪声干扰的情况。由于亮光下受试者佩戴的镜片产生了面积较大反射光斑并覆盖在了虹膜上,容易将该处像素误识别为其他区域,难以准确分割。虹膜部分缺失或遮挡情况也对算法的精度提出了挑战。

为精确分割出虹膜,采用卷积神经网络(CNN)提取虹膜特征,并提出一种新的分割网络 MFFIris-Unet。该网络参考 U-Net 语义分割模型的端到端编-解码结构,用于特征提取的编码器部分,采用基于深度可分离卷积(DSC)设计的改进残差瓶颈单元(RBU)构建的残差网络,以降低训练参数量。通过在网络中间的瓶颈处放置所设计的基于空洞卷积的改进密集多尺度特征融合模块,提升不同尺度特征的表达能力。在上采样阶段,采用双线性插值与反卷积相结合的方式恢复特征图,该方法相比仅使用插值或反卷积的方式性能更均衡。在每层的跳跃连接的通道拼接操作前加入注意力机制,提升细节信息获取的有效性,提升分割精度与训练速度。通过以上操作构建的模型在兼顾分割精度的前提下降低了模型的参数量与计算量,提升了对复杂虹膜图像的噪声目标分割的鲁棒性。

2.2 MFFIris-Unet 算法

根据上述分析,借鉴 U-Net 的网络架构,重新设计了一种基于注意力机制与密集多尺度特征融合的轻量级异质噪声虹膜分割深度神经网络模型,并将其命名为 Multiscale Feature Fusion Iris segmentation U-Net(MFFIris-Unet),结构如图 2 所

示,上方标注了每一层特征图的深度,下方虚线框内的图例注明了不同符号所表示的含义。

首先,裁剪后的输入图像尺寸为 288×288 ,为提升网络对图像中不同尺度分割目标的特征表示与提取能力,以改善边缘处的分割细节,将编码阶段的下采样(down sampling)操作增加至 5 次。进行编码阶段的 5 次下采样特征提取操作后,最终得到尺寸缩小 $1/32$ 、深度扩大 10 倍的特征图。在编码网络与解码网络之间的瓶颈处放置了改进的特征金字塔模块(FPM),融合前面各层不同尺度特征并提高大感受野下的训练精度。网络结构可分为 10 层,编码阶段的 1~5 次下采样与解码部分的 6~10 次上采样通过“跳跃连接”联系,其中第 6、7 次上采样(up sampling)采用双线性插值(bilinear interpolation),第 8、9、10 次上采样采用反卷积(transpose convolution)。经实验证明,深层采用双线性插值操作与浅层采用反卷积操作结合的上采样方式能较好地平衡解码阶段的速度与计算量。其次,利用深度可分离卷积替换传统卷积降低计算量的策略,在编码和解码阶段的每一层下采样和上采样操作后均嵌入改进的残差瓶颈单元,从而能更好地利用卷积层间的残差连通性,促进高层语义信息在网络中的传播,以较浅的网络深度达到更高的精度。随后,在解码阶段的跳跃连接之前与模型上采样的最后加入高效并行空间-通道注意力模块(EPSCAM),提升不同层次与维度特征图语义特征表示的有效性,再通过一个 1×1 卷积对同层次的特征图与上一次上采样的特征图进行拼接操作,以实现同尺度信息的相互融合。最后,通过一个

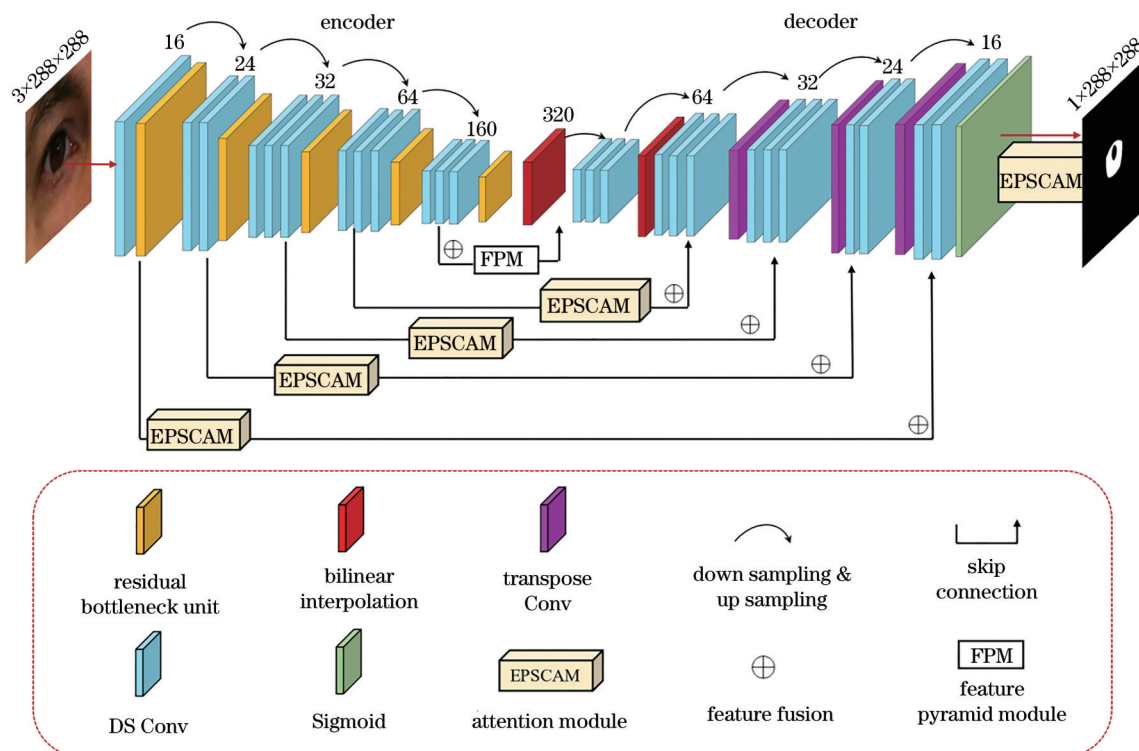


图 2 MFFIris-Unet 架构

Fig. 2 MFFIris-Unet architecture

Sigmoid 函数对图像进行像素级分类,得到与输入图像相同尺寸的预测图。

2.3 改进的残差瓶颈单元

深度可分离卷积^[12]的核心是利用卷积层通道与空间相关的可退耦特性,将常规卷积分解为深度卷积(DW)和一个1×1的点卷积(PW)。DW单卷积核对应单个通道,输出的特征图通道数与输入相同,即不改变深度。PW可实现升维或降维,有效编码通道信息。将DSC替换传统卷积的网络可在保证精度的同时显著降低计算量与参数量。如采用DSC的MFFIris-Unet的计算成本大约仅为标准卷积的1/8且精度损失甚微^[12]。构建残差网络,可有效改善网络过深导致的梯度消失与梯度爆炸问题。倒残差块(inverted residual block)结构如图3所示,由PW+DW+PW的方式组合而成,通过将低维特征采用线性表示,去除第2个PW批规范化(BN)层后的ReLU6激活函数来构造线性瓶颈层,学习表示空间特征的上下文信息。另外,首尾相加的快捷连接结构帮助网络避免出现随着深度增加而退化的情况。

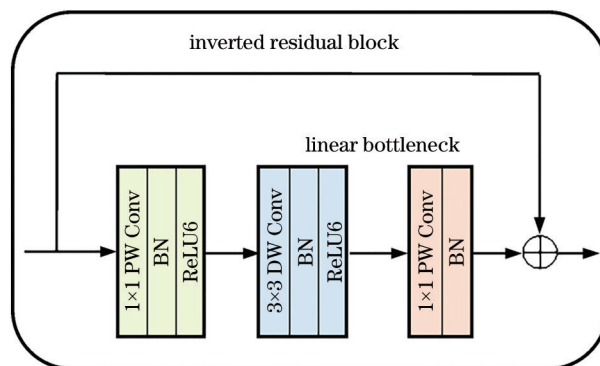


图 3 倒残差块结构

Fig. 3 Structure of inverted residual block

针对倒残差块存在的信息丢失与梯度混淆问题,提出改进的残差瓶颈单元(RBU)^[13],其结构如图4所示。RBU由3×3的DW+PW+PW+3×3

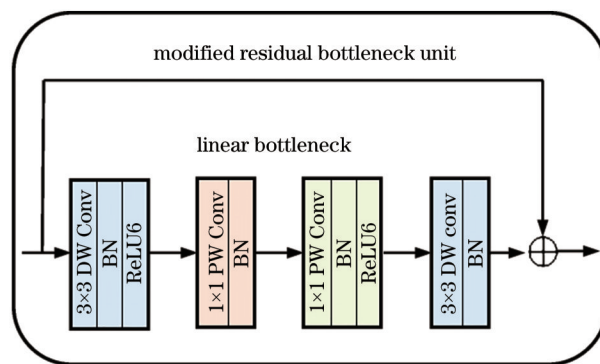


图 4 改进残差瓶颈单元结构

Fig. 4 Structure of the modified residual bottleneck unit

的 DW 四部分组成,其中在第 2 层与第 4 层后不添加激活函数,在特征的高维表示间建立快捷连接,促进了梯度信息的层间传播。实验证明,相比采用倒残差块的模型,采用改进残差瓶颈单元的模型在 ImageNet 分类任务和 VOC2007 目标检测数据集上的精度分别提升 1.7% 和 0.9%^[13]。因此,所提网络采用基于 DSC 的 RBU。

2.4 高效并行空间-通道注意力机制

注意力机制在图像识别、自然语言处理等智能信息处理领域应用广泛^[14]。压缩激励(SE)模块是一种通过学习特征图通道间的维度映射关系来提升网络性能的注意力机制,先通过压缩(squeeze)操

作的全局平均池化获取特征图的空间关系和通道特征。随后由全连接层-ReLU-全连接层-Sigmoid组成的激励(excitation)操作获取不同通道的权重^[13]。其中两个全连接层的作用是获取非线性的跨通道交互和降维,从而降低模型复杂性,但降维的同时也阻碍了通道注意预测,故对小目标分割效果不佳。

为提高模型对有效特征的代表能力,受文献[15]中高效通道注意力(ECA)模块启发,以 scSE^[15]的并行注意力架构为基础,结合局部信息跨通道交互策略,设计出更适用于所设计网络的注意力机制,结构如图 5 所示。

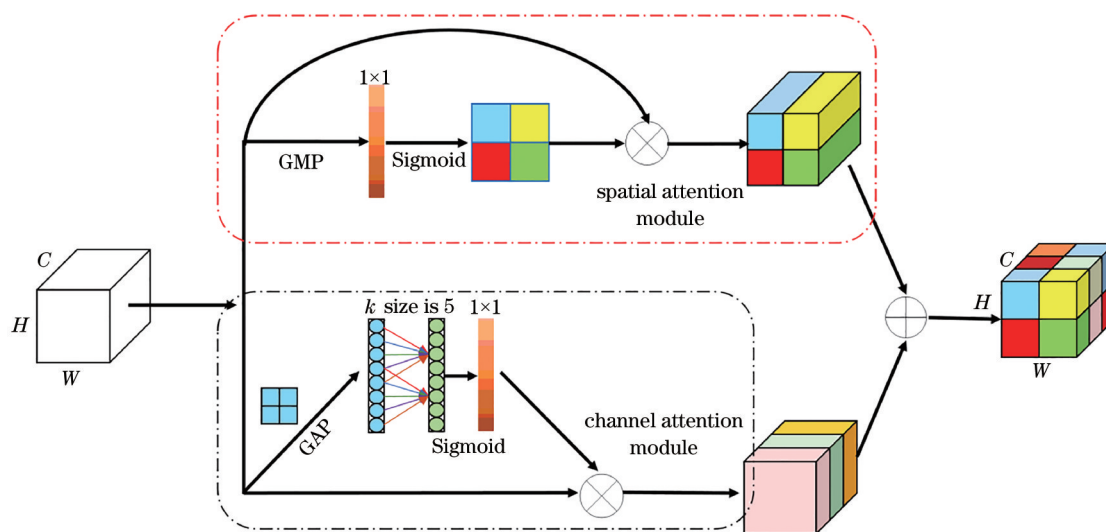


图 5 空间-通道并行注意力模块结构

Fig. 5 Spatial-channel parallel attention module architecture

通道注意力模块(CAM)中,切片后的特征图先经不降低维数的全局平均池化,随后使用核大小 k 为 5 的快速一维卷积实现局部通道信息交互,其中 k 表示参与每个通道注意力预测的其他邻近通道数目。最后经 Sigmoid 操作后和输入特征图进行点乘,实现权重概率的分配。空间注意力模块(SAM)首先经全局最大池化(GMP)操作获取当前感受野下的细节信息,随后经 1×1 卷积、Sigmoid 操作降维与低层特征相乘后得到加权特征图。该过程使网络在空间维度上学习到特征的重要性。通道注意力与空间注意力机制对通道和空间两个维度的信息重要性进行权重分配,以此来提升对有价值的信息利用,抑制无关信息的干扰,在提高网络的特征表示能力与效率的同时也有助于防止梯度消失和网络退化。

2.5 多尺度特征融合模块

充分利用多尺度信息可提高网络对不同尺度

目标的分割性能。常用的多尺度特征提取结构有空间金字塔池化(PSP)模块和空洞空间金字塔池化(ASPP)模块^[16]。PSP 模块以并行拼接的方式聚合多尺度信息,参数量较大。ASPP 模块对不同扩张率的空洞卷积生成的特征图进行级联,使输出结合多个感受野大小的信息,避免了多次池化导致的信息丢失。感受野增加,扩张率也要随之增加,随着空洞率的增加,分割效率逐渐下降,另外 DSC 的引入和网络层数的增加可能导致精度损失。

为更好地利用虹膜图像中噪声小目标与虹膜区域特征存在的差异,分割出具有更细粒度的多尺度目标,提升噪声虹膜的分割精细度,对文献[16]提出的密集空洞空间金字塔模块(Dense-ASPP)进行改进,得到的 FPM 结构如图 6 所示。先采用 4 个空洞率组合为 (3, 6, 12, 18) 的 3×3 空洞卷积,并加入 BN 操作^[14],使结构更适合虹膜图像的尺寸,增加

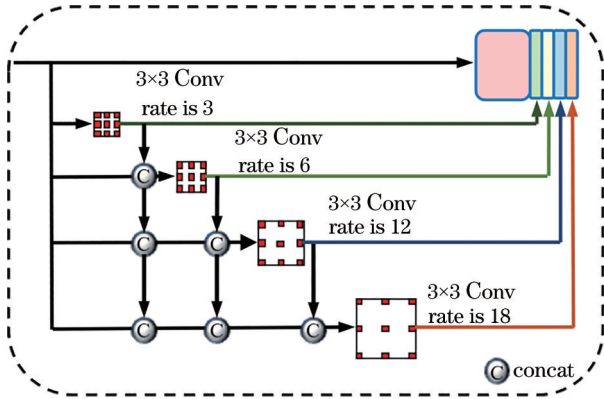


图 6 改进的 Dense-ASPP 结构

Fig. 6 Improved Dense-ASPP structure

了训练的稳定性。不同空洞率卷积支路的特征图通过全局平均池化后输入到 256 个滤波器的 1×1 卷积中,随后使用双线性插值上采样至统一尺寸。最后通过一个 1×1 卷积拼接所有支路的特征。

在编码部分后特征融合前加入改进的 FPM,在保持特征图分辨率不变的情况下增加感受野,以逐像素相加的方式将特征融合至解码部分的对应层中,从而提升了网络提取多个尺度下的上下文特征信息的能力。

2.6 损失函数

合适的损失函数对提升分割性能至关重要。虹膜图像常存在背景像素占比远大于虹膜区域像素的情况,即正样本占全图比例过低的类别不平衡问题。二分类语义分割网络中常用的二值交叉熵(BCE)损失函数对类别不平衡图像中正样本像素类别特征学习困难且易陷入局部最优,导致训练出的模型鲁棒性不佳。为解决该问题,所提网络的损失函数采用文献[17]提出的 Dice 损失,实际使用中为减少过拟合,将分子分母同时加 1,改进后损失函数定义式为

$$D_{Loss}(X, Y) = 1 - \frac{2|X \cap Y| + 1}{|X| + |Y| + 1}, \quad (1)$$

式中: D_{Loss} 为 Dice 损失函数值; X 、 Y 分别表示 ground truth 标签和预测结果。

相比于 BCE 损失, Dice 损失将衡量预测结果与 ground truth 标签之间的重叠像素作为损失函数值。这一思想本质上与语义分割评价指标交并比(IoU)一致,训练过程中, Dice 损失值随着模型的收敛而降低,同时 IoU 逐渐变大,模型的预测精度不断上升。因此,对于存在较多类别不平衡图像的分割任务, Dice 损失对网络训练的优化效果比 BCE 损失更好。

3 实验与结果分析

3.1 数据集与数据增强

为验证模型性能,选择 CASIA. v4-distance (CASIA)^[18]、UBIRIS. v2 (UBIRIS)^[19]、MICHE-I (MICHE)^[20] 3 个具有挑战性的公开虹膜数据集对所提算法进行评价。CASIA 包含远距离近红外相机对 142 名受试者采集的 2567 幅包括左右眼的上半人脸图像,本文选择文献[21]中裁剪的前 40 名受试者的 400 幅 640×480 像素的虹膜图像及提供的逐像素标注的精细标签图。UBIRIS 由移动和远距离状态下拍摄的 261 名对象共 11102 张彩色虹膜图像组成,本文采用噪声虹膜分割比赛 NICE-I^[19] 的 1000 幅带标注样本进行训练和测试。MICHE 由多种型号移动设备拍摄的共 3000 张可见光图像组成,本文使用文献[22]提供的 2 个子集共 871 个标注图像对进行实验。上述数据集涵盖了不同时间段和采集距离的可见光与近红外多光谱噪声虹膜图像,包含遮挡、注视偏移、镜面反射、模糊等丰富噪声且尺度不一。因此,采用上述数据对所提算法进行多光谱噪声虹膜分割性能评价是科学、合理和有效的。

按照 7:1:2,将原始数据集划分为训练集、验证集、测试集三部分。因可获取的已标注数据较少,故先对训练集进行随机组合旋转、平移、缩放、裁剪等几何变换操作并随机加噪(运动模糊、高斯模糊)进行数据增强,以提升训练效果。具体操作分三个阶段:首先,以 0.5、0.75、1.25、1.5、1.75 五个不同尺度随机调整原图;其次,对图像与对应的标签同时作上下左右的随机翻转、 $-15^\circ \sim 15^\circ$ 的平移操作、 30° 以内的随机旋转,增强后得到的 CASIA、MICHE、UBIRIS 训练总样本数分别为 2224、5560、5960;最后,在输入网络时将图像统一裁剪成 288×288 的固定大小。3 个数据集经数据增强后的训练样本和标签实例如图 7 所示。

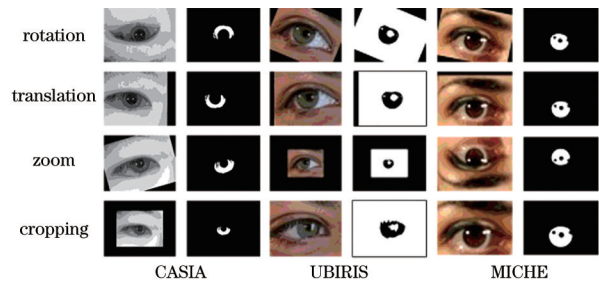


图 7 经数据增强的训练样本示例

Fig. 7 Examples of data enhanced training samples

3.2 网络训练

使用的云计算平台软硬件配置: Intel(R) Xeon(R) Gold 6132 CPU @ 2.60 GHz 处理器、内存 192.0 GB, 操作系统为 Ubuntu 18.04.5 LTS 64 bit、CUDA 版本 10.1、OpenCV 4.4、python 3.6, 整个模型的训练全程使用 PyTorch1.3.1 深度学习框架, 训练过程中采用 1 块显存 16 GB 的 NVIDIA® Tesla T4 GPU 进行加速。在训练过程中, 超参数和优化器设置如下: 选择 Adam 优化器

在反向传播阶段更新网络的权值; 固定初始学习率为 0.01 并采用自适应的策略, 随着网络训练进行不断更新学习率; batch size 为 16, 初始估计指数衰减率 beta1 设置为 0.9, 第 2 次估计指数衰减率设置为 0.999, 模型训练的 epoch 次数设置为 300, 根据训练情况可手动调整。图 8 为所设计模型分别在 CASIA、UBIRIS、MICHE 数据集上训练的损失下降曲线与平均交并比 (mIoU) 精度变化曲线。

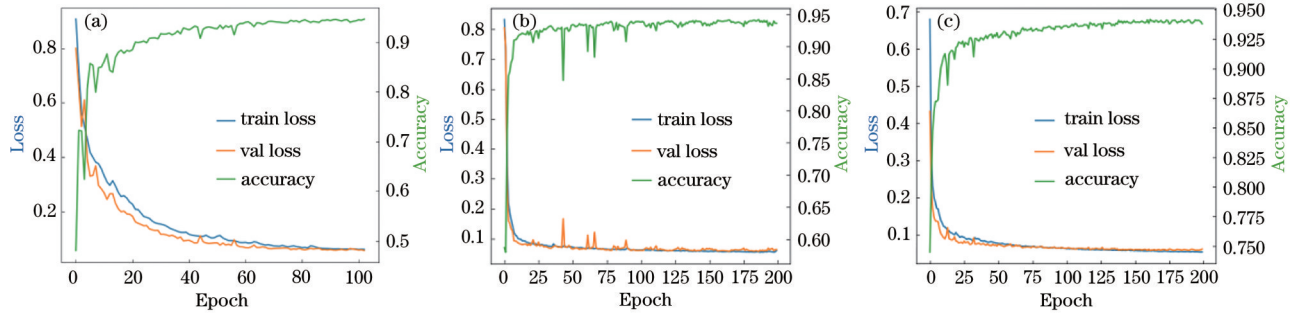


图 8 不同数据集上的训练损失函数与精度变化曲线。(a) CASIA; (b) UBIRIS; (c) MICHE

Fig. 8 Curves of training loss function and precision change at different datasets. (a) CASIA; (b) UBIRIS; (c) MICHE

3.3 实验结果分析

3.3.1 评价指标

由于移动端虹膜分割更偏重关注预测的正确率, 故选择 mIoU 和平均 F1 分数两个指标评价分割精度。交并比即预测结果与标签图的交集与并集之比^[23]。平均交并比是所有测试图像分割交并比的均值, 其值介于 [0, 1], 越接近 1, 分割结果越精确。mIoU 的公式为

$$\alpha_{mIoU} = \frac{1}{n} \sum_{i=1}^n \left[\frac{N_{tp}}{N_{tp} + N_{fn} + N_{fp}} \right], \quad (2)$$

式中: N_{tp} 为真阳性率 (true positive), 表示标注与预测相同的像素数; N_{fp} 为假阳性率 (false positive), 表示非虹膜像素预测为虹膜的像素数; N_{fn} 为假阴性率 (false negative), 表示属于虹膜像素但误识别为其他像素的个数。F1 分数的定义为准确率 (P) 和召回率 (R) 的调和平均值, 即

$$F_{Score} = \frac{2R \times P}{R + P} = \frac{2N_{tp}}{2N_{tp} + N_{fp} + N_{fn}}, \quad (3)$$

式中: $P = \frac{N_{tp}}{N_{tp} + N_{fp}}$, 表示预测结果与真实像素之比, $R = \frac{N_{tp}}{N_{tp} + N_{fn}}$, 表示在标注图中虹膜像素分割正确的比例。 P, R, F_{Score} 也均介于 [0, 1], 越接近 1, 表示结果越好。为评价测试集的整体分割性能, 分

别计算了 3 个指标的均值 μ 和方差 σ 。均值越高, 方差越小, 表示算法性能越好。另外, 还从参数量 (Params)、每秒浮点运算次数 (FLOPs)、占用运行内存空间 (Storage space)、平均预测时间 (Average time) 4 个指标评价算法空间复杂度和时间复杂度。

3.3.2 不同算法对比

选择当前效果较好的传统算法 (RTV-L) 和基于深度学习的算法 (Deeplab V3、U-Net、PI-Unet^[24]) 与所提算法进行对比, 分别在可见光与近红外光谱的 3 个虹膜数据集上进行训练和测试 (所有算法均采用与原论文相同的参数)。并随机选取了包括睫毛遮挡、镜面反射、闭眼、斜视等困难分割样本可视化分析预测结果。部分困难样本的分割实例如图 9~11 所示, 从左至右依次为原图、ground truth 标注图、Deeplab V3 分割结果、U-Net 分割结果、RTV-L 分割结果、PI-Unet 分割结果、所提 MFFIris-Unet 分割结果。红色标注为错误拒绝像素 (标记为虹膜而算法预测为其他像素), 绿色标注为错误接受像素 (标记为其他像素但算法预测为虹膜像素)。

实验结果显示: 由于近红外图像受环境光线变化小, 虹膜轮廓相对可见光图像更显著, RTV-L 算法在近红外的 CASIA 数据集上分割轮廓清晰可见, 但对包含睫毛遮挡、眯眼、镜面反射等噪声的样本分割效果较差; 与传统算法 RTV-L 相比, 采用编-解

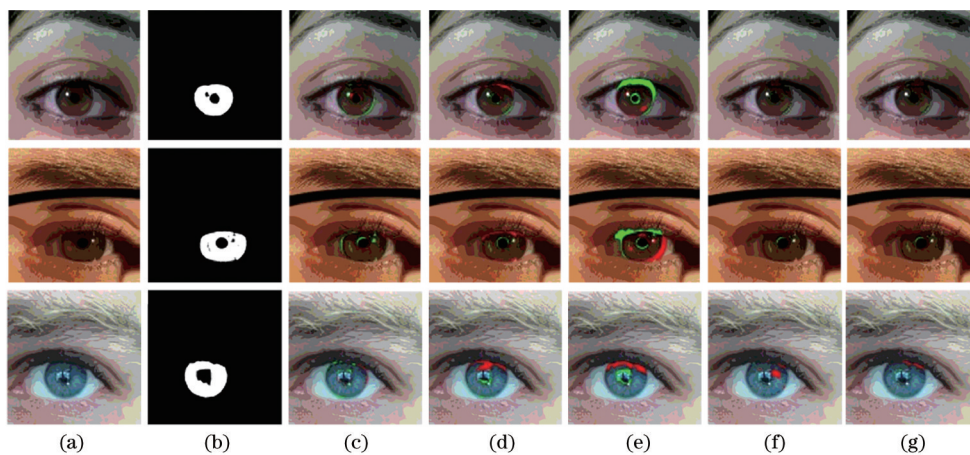


图 9 不同方法在 MICHE 数据集上的分割结果。(a)原图;(b)标注图;(c) Deeplab V3 分割结果;(d) U-Net 分割结果;
(e) RTV-L 分割结果;(f) PI-Unet 分割结果;(g) MFFIris-Unet 分割结果

Fig. 9 Segmentation results of different methods on MICHE dataset. (a) Original image; (b) ground truth; (c) results of Deeplab V3; (d) results of U-Net; (e) results of RTV-L; (f) results of PI-Unet; (g) results of MFFIris-Unet

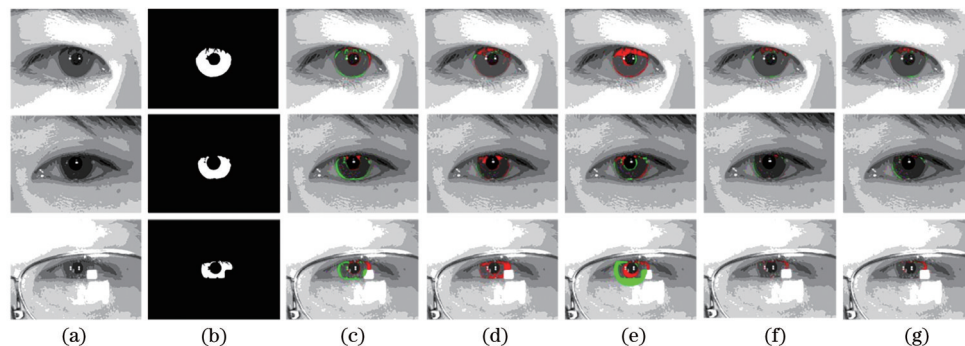


图 10 不同方法在 CASIA 数据集上的分割结果。(a)原图;(b)标注图;(c)Deeplab V3 分割结果;(d)U-Net 分割结果;
(e)RTV-L 分割结果;(f)PI-Unet 分割结果;(g)MFFIris-Unet 分割结果

Fig. 10 Segmentation results of different methods on CASIA dataset. (a) Original image; (b) ground truth; (c) results of Deeplab V3; (d) results of U-Net; (e) results of RTV-L; (f) results of PI-Unet; (g) results of MFFIris-Unet

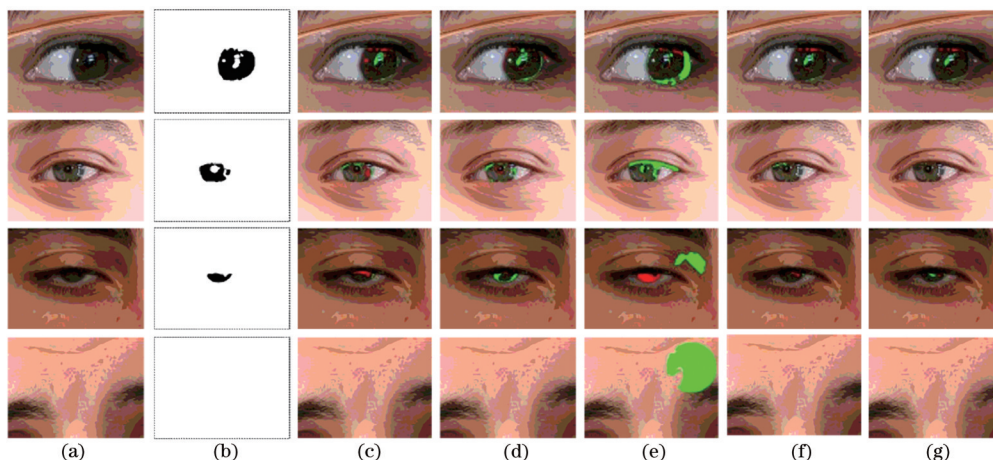


图 11 不同方法在 UBIRIS 数据集上的分割结果。(a)原图;(b)标注图;(c) Deeplab V3 分割结果;(d) U-Net 分割结果;
(e) RTV-L 分割结果;(f) PI-Unet 分割结果;(g) MFFIris-Unet 分割结果

Fig. 11 Segmentation results of different methods on UBIRIS dataset. (a) Original image; (b) ground truth; (c) results of Deeplab V3; (d) results of U-Net; (e) results of RTV-L; (f) results of PI-Unet; (g) results of MFFIris-Unet

码器架构的 U-Net 模型预测结果中误分辨像素明显减少,且在不同数据集上表现更均衡;由于采用了串并连接结合的空洞空间金字塔池化模块, Deeplab V3 模型分割结果在噪声区域的细节比 U-Net 有所优化;PI-Unet 的分割结果较精细,最接近所提算法;所提 MFFIris-Unet 模型在编-解码器对应层中加入改进 RBU 和下采样末端使用密集连接的多尺度特征融合,提高了网络对不同层次特征的提取能力,在可见光与近红外光谱的虹膜数据集上的分割性能较其他几种算法更为均衡,分割结果更接近真实的标签图,虹膜内外边缘更精细平滑、像素误判更少、分割精度更高。另外,与传统的虹膜分割算法相比,深度虹膜分割神经网络的准确率明显更高,对困难分割样本的鲁棒性更好。特别是当

图像中不存在虹膜区域时,传统方法易出现误判现象,而深度学习方法则能准确地分辨出是否存在虹膜区域。

表 1 展示了 5 种虹膜分割算法在 3 种虹膜数据集上的定量指标结果,加粗项为对应数据集中指标的最好值。与其他模型相比,MFFIris-Unet 分割精度更高,在 CASIA、UBIRIS、MICHE 三个虹膜数据集上的平均 F1 分数分别达到 97.14%、96.59%、96.54%; mIoU 为 94.61%、94.28%、93.63%;另外,所提算法训练时间更短,模型预测速度更快,优势明显。

图 12 为 5 种不同方法在 CASIA、UBIRIS、MICHE 3 个数据集上的 mIoU 与平均 F1 分数的直方图对比。

表 1 不同方法在 3 个虹膜数据集上的评价指标结果

Table 1 Evaluation index results of different methods on three iris datasets

Method	Dataset	R		P		F1-Score		mIoU / %	Average time / s
		μ / %	σ / %	μ / %	σ / %	μ / %	σ / %		
Deeplab V3	CASIA	90.13	6.62	92.80	4.12	93.21	3.67	88.21	0.56
	UBIRIS	85.17	9.53	90.92	4.01	87.55	6.32	79.24	0.44
	MICHE	89.84	10.93	91.66	8.12	91.18	8.89	84.69	0.41
U-Net	CASIA	91.77	7.62	95.23	3.51	91.78	5.58	87.34	0.93
	UBIRIS	91.96	7.82	90.29	4.63	90.81	4.92	81.92	0.67
	MICHE	88.86	13.13	90.75	8.56	88.25	10.52	81.20	0.66
RTV-L	CASIA	80.95	6.59	95.83	3.91	87.55	4.58	78.11	2.68
	UBIRIS	88.23	9.66	85.16	10.58	85.97	8.72	74.01	1.15
	MICHE	84.56	17.61	74.27	16.82	77.10	14.71	64.21	1.57
PI-Unet	CASIA	93.11	8.69	95.22	5.31	96.53	5.41	94.21	0.18
	UBIRIS	91.87	7.43	91.98	4.55	95.25	6.25	92.31	0.26
	MICHE	93.52	10.11	93.65	9.16	94.02	8.52	93.53	0.33
MFFIris-Unet	CASIA	92.62	7.65	96.56	3.69	97.14	4.36	94.61	0.11
	UBIRIS	92.87	6.87	92.96	3.68	96.59	4.11	94.28	0.10
	MICHE	94.05	10.21	93.14	8.69	96.54	8.62	93.63	0.07

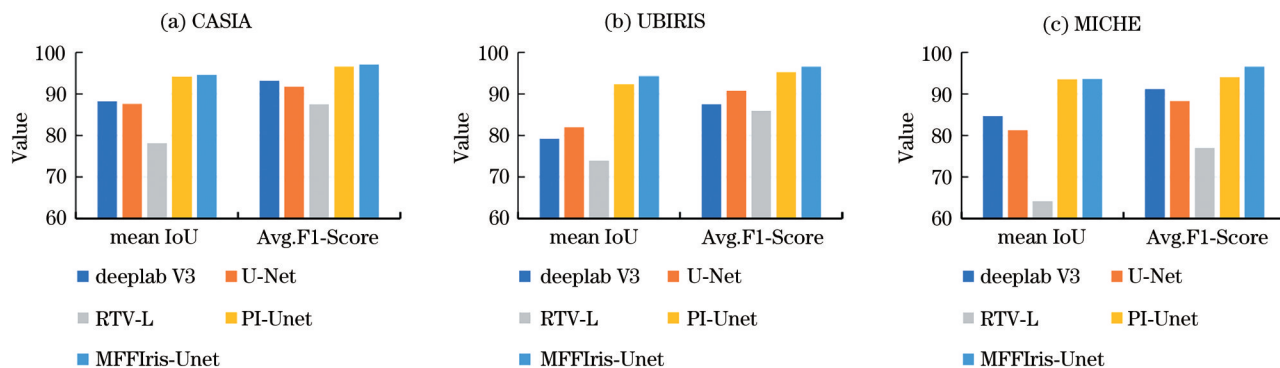


图 12 3 个测试集上的 mIoU 与平均 F1 分数直方图

Fig. 12 Histograms of mIoU and average F1 scores on three datasets

表 2 比较了不同深度学习分割网络在输入图片大小为 288×288 时的参数量、计算量(每秒浮点运算次数)、存储空间大小。所提网络模型参数量、计算量、存储空间远低于 FCN8s、U-Net、SegNet、Deeplab V3 四种经典语义分割网络与最新的轻量级虹膜分割模型 PI-Unet, 模型更轻量化, 应用优势明显。

3.3.3 消融实验

为验证所采用的数据增强策略、损失函数与多种模块组合对提高虹膜分割性能的有效性, 设计了 4 种不同网络进行消融研究。首先基准网络为未采用本文策略的 MobileV2+Unet 结构, 再在 MFFIris-Unet 基础上分别去除注意力机制与多尺度特征融合模块, 设计了 2 个不同网络。比较 4 种

表 2 不同方法的参数量、计算量、存储空间对比
Table 2 Comparison of the number of parameters, computation amount, and storage space of different methods

Method	Params / 10^6	FLOPs / 10^9	Storage space / GB
FCN8s	134.27	84.99	0.513
U-Net	26.36	62.61	0.121
SegNet	16.31	53.93	0.123
Deeplab V3	18.86	26.29	0.072
PI-Unet	2.86	1.56	0.012
MFFIris-Unet	1.45	0.35	0.005

网络在 3 个不同光谱的虹膜数据集上的平均 F1 分数、mIoU、平均时间、训练时间、模型大小 5 个评价指标, 结果如表 3 所示。

表 3 消融实验结果

Table 3 Results of ablation experiments

Method	Dataset	F1-Score / %	mIoU / %	Average time / s	Train time / h	Model size / MB
Base	CASIA	96.95	94.40	0.76	6	9.69
	UBIRIS	96.23	93.61	0.61	22	
	MICHE	96.47	93.51	0.55	23	
Base+RBU+Atten	CASIA	96.56	93.85	0.22	3	5.32
	UBIRIS	95.22	93.44	0.16	8	
	MICHE	95.62	93.21	0.13	5	
Base+RBU+FPM	CASIA	96.89	94.59	0.12	5.5	5.66
	UBIRIS	96.29	93.96	0.08	21	
	MICHE	96.54	93.69	0.10	22	
MFFIris-Unet	CASIA	97.14	94.61	0.11	3	5.66
	UBIRIS	96.59	94.28	0.07	8	
	MICHE	96.54	93.63	0.10	5	

由表 3 可知: 与采用倒残差模块的基准网络相比, 使用本方案设计的网络更轻量, 训练后模型大小仅为 5.66 MB 且测试时预测速度明显更快, 仅为基准网络的 1/6; 替换 RBU 结构后, 网络参数量与计算量减少明显; 而在加入注意力机制后模型收敛更快, 训练时间大幅缩短; 采用 FPM 与 RBU 的网络比基准网络精度更高、更轻量。网络之间分割精度相差较小, 但仍略好于基准网络, 在 CASIA、UBIRIS、MICHE 数据集上, MFFIris-Unet 的平均 F1 分数比基准网络分别提高了 0.19 百分点、0.36 百分点、0.07 百分点, mIoU 分别提高了 0.21 百分点、0.67 百分点、0.12 百分点, 体现了所提算法的优势。

图 13 为基准网络和 MFFIris-Unet 在三个数据

集上部分预测可视化结果。从左至右依次为原图、ground truth 标注图、基准网络分割结果、MFFIris-Unet 分割结果。从主观视觉上来看, 基准网络对噪声较多的近红外虹膜分割效果不佳, MFFIris-Unet 在异质虹膜图像上均表现良好且能分割出更多噪声信息, 对于小目标的细节处理更佳, 直接证明了所提多个模块组合对提升噪声虹膜图像分割精度的有效性。

另外经计算得, 基准网络的参数量和计算量分别为 2.49×10^6 、1.01 GMac, 而 MFFIris-Unet 参数量和计算量更低, 仅为 1.45×10^6 、 0.35×10^9 FLOPs, 分别减少了 41.77%、65.35%, 显著优于基准网络。综上, 本设计策略有助于提升分割效率, 实现更高效、鲁棒的噪声虹膜分割。

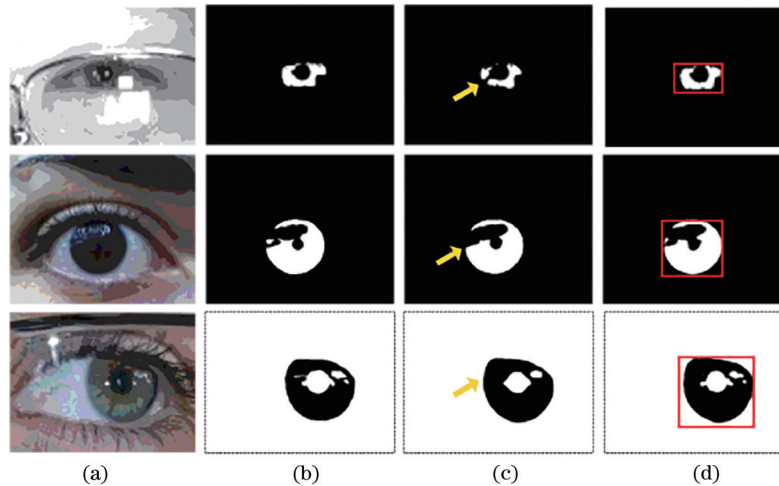


图 13 MFFIris-Unet 与基准网络的可视化预测结果。(a)原始图像;(b)标注图;(c)基准网络分割结果;(d)MFFIris-Unet 分割结果

Fig. 13 Visualized results predicted by the base model and MFFIris-Unet. (a) Original image; (b) ground truth; (c) results of base method; (d) results of MFFIris-Unet

4 结 论

针对移动端虹膜图像噪声干扰复杂,传统虹膜分割网络难以兼顾可见光与近红外光谱图像的分割精度问题,在深度编-解码网络的基础上,首先,用改进残差瓶颈单元替换残差块,在防止梯度消失的同时减少了网络参数量与计算量;其次,改进Dense-ASPP并将其放置于编-解码网络的瓶颈处,用于增强多尺度特征信息的提取;最后,在下采样每层之后和解码阶段的最后引入空间-通道并行注意力机制,增强网络对通道和空间信息重要性的判断能力,缩短了训练时间。在3个虹膜数据集上的实验结果表明,采用本设计策略,结合使用Dice损失函数及数据增强策略进行训练,各指标均优于传统深度学习分割模型,精度比未改进的基准网络略高且在多光谱上分割性能更均衡,在三个数据集上的平均F1分数最高分别为97.14%、96.59%、96.54%,mIoU为94.61%、94.28%、93.63%,而模型空间占用、参数量、计算量分别减少41%、41.77%、65.35%。结合可视化困难样本分割结果,表明所提算法能更高效、精细地分割包含遮挡、反光、模糊等复杂噪声的可见光与近红外光谱的虹膜图像,有效改善了对异质噪声虹膜图像的分割性能,在低性能的边缘计算设备上具有良好的应用价值。

参 考 文 献

[1] Wang C Y, Muhammad J, Wang Y L, et al.

Towards complete and accurate iris segmentation using deep multi-task attention network for non-cooperative iris recognition[J]. IEEE Transactions on Information Forensics and Security, 2020, 15: 2944-2959.

- [2] Daugman J G. High confidence visual recognition of persons by a test of statistical independence[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 15(11): 1148-1161.
- [3] Wildes R P. Iris recognition: an emerging biometric technology[J]. Proceedings of the IEEE, 1997, 85(9): 1348-1363.
- [4] Tan T N, He Z F, Sun Z N. Efficient and robust segmentation of noisy iris images for non-cooperative iris recognition[J]. Image and Vision Computing, 2010, 28(2): 223-230.
- [5] Zhao Z J, Kumar A. An accurate iris segmentation framework under relaxed imaging constraints using total variation model[C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 3828-3836.
- [6] Liu N F, Li H Q, Zhang M, et al. Accurate iris segmentation in non-cooperative environments using fully convolutional networks[C]//2016 International Conference on Biometrics (ICB), June 13-16, 2016, Halmstad, Sweden. New York: IEEE Press, 2016: 16252559.
- [7] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions

- on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640-651.
- [8] Tan C W, Kumar A. Unified framework for automated iris segmentation using distantly acquired face images[J]. IEEE Transactions on Image Processing, 2012, 21(9): 4068-4079.
- [9] Hao K L, Feng G R, Ren Y L, et al. Iris segmentation using feature channel optimization for noisy environments[J]. Cognitive Computation, 2020, 12(6): 1205-1216.
- [10] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[M] // Navab N, Hornegger J, Wells W M, et al. Medical image computing and computer-assisted intervention-MICCAI 2015. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.
- [11] Lian S, Luo Z M, Zhong Z, et al. Attention guided U-Net for accurate iris segmentation[J]. Journal of Visual Communication and Image Representation, 2018, 56: 296-304.
- [12] Sandler M, Howard A, Zhu M L, et al. MobileNetV2: inverted residuals and linear bottlenecks [C] //2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 4510-4520.
- [13] Roy A G, Navab N, Wachinger C. Concurrent spatial and channel 'squeeze & excitation' in fully convolutional networks[M]//Frangi A F, Schnabel J A, Davatzikos C, et al. Medical image computing and computer assisted intervention-MICCAI 2018. Lecture notes in computer science. Cham: Springer, 2018, 11070: 421-429.
- [14] Yang M K, Yu K, Zhang C, et al. DenseASPP for semantic segmentation in street scenes[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 3684-3692.
- [15] Zhang W X, Zhu Z C, Zhang Y H, et al. Cell image segmentation method based on residual block and attention mechanism[J]. Acta Optica Sinica, 2020, 40(17): 1710001.
张文秀, 朱振才, 张永合, 等. 基于残差块和注意力机制的细胞图像分割方法[J]. 光学学报, 2020, 40(17): 1710001.
- [16] Zhou D Q, Hou Q B, Chen Y P, et al. Rethinking bottleneck structure for efficient mobile network design[M]//Vedaldi A, Bischof H, Brox T, et al. Computer vision-ECCV 2020. Lecture notes in computer science. Cham: Springer, 2020, 12348: 680-697.
- [17] Li D X, Zhang Z. Improved U-Net segmentation algorithm for the retinal blood vessel images[J]. Acta Optica Sinica, 2020, 40(10): 1010001.
李大湘, 张振. 基于改进U-Net视网膜血管图像分割算法[J]. 光学学报, 2020, 40(10): 1010001.
- [18] Milletari F, Navab N, Ahmadi S A. V-net: Fully convolutional neural networks for volumetric medical image segmentation[C] //2016 Fourth International Conference on 3D Vision (3DV), October 25-28, 2016, Stanford, CA, USA. New York: IEEE Press, 2016: 565-571.
- [19] BIT. CASIA iris image database [EB/OL]. [2021-01-26]. <http://www.idealtest.org/#/datasetDetail/4>.
- [20] Proenca H, Alexandre L A. The NICE.I: noisy iris challenge evaluation-part I[C] //2007 First IEEE International Conference on Biometrics: Theory, Applications, and Systems, September 27-29, 2007, Crystal City, VA, USA. New York: IEEE Press, 2007: 9871330.
- [21] de Marsico M, Nappi M, Riccio D, et al. Mobile Iris Challenge Evaluation (MICHE) -I, biometric iris dataset and protocols[J]. Pattern Recognition Letters, 2015, 57: 17-23.
- [22] Wang C Y, Sun Z N. A benchmark for iris segmentation[J]. Journal of Computer Research and Development, 2020, 57(2): 395-412.
王财勇, 孙哲南. 虹膜分割算法评价基准[J]. 计算机研究与发展, 2020, 57(2): 395-412.
- [23] Wang Y, Liu L B. Bilinear residual attention networks for fine-grained image classification[J]. Laser & Optoelectronics Progress, 2020, 57(12): 121011.
王阳, 刘立波. 面向细粒度图像分类的双线性残差注意力网络[J]. 激光与光电子学进展, 2020, 57(12): 121011.
- [24] Zhou R Y, Shen W Z. PI-Unet: a neural network model for heterogeneous iris segmentation[J]. Computer Engineering and Application, 2021, 57(15): 223-229.
周锐焯, 沈文忠. PI-Unet:异质虹膜精确分割神经网络模型的研究[J]. 计算机工程与应用, 2021, 57(15): 223-229.