

文章编号 1004-924X(2025)20-3299-16

基于 CNN 和 Transformer 的睑板腺图像 多粒度分割

杨 松¹, 夏振平^{1*}, 李 笠², 吴彦澍¹

(1. 苏州科技大学 电子与信息工程学院, 江苏 苏州 215009;

2. 福建省立医院眼科, 福建 福州 350001)

摘要:针对睑板腺图像分割中的多阶段处理和边缘模糊问题,本文设计了一种端到端的多粒度分割算法。在编码阶段,采用 TransUNet 编码器架构,能够高效提取眼睑和腺体区域的共享特征;在解码阶段,采用双解码路径针对眼睑和腺体区域的不同特征分别设置不同的解码器分支。同时,针对腺体区域,设计了多尺度特征融合模块,并在跳跃连接中加入通道注意力机制。这些优化提升了边缘精度、纹理清晰度和形状轮廓,从而有效解决了边缘模糊和腺体粘连的问题。对于眼睑区域,则采用标准解码器结构进行分割预测。通过与现有先进分割方法的实验对比,所提算法在上下睑板腺的精度均值上表现优异,尤其在关键指标平均交并比和 Dice 相似性系数上,分别达到了 79.9% 和 76.5%,较 TransUNet 分别提高了 3.2% 和 5.3%。本文算法能够精准地分割睑板腺图像的目标区域,可以为睑板腺功能障碍的辅助诊断提供依据。

关键词:睑板腺图像分割;多粒度分割;CNN;Transformer;医学图像处理

中图分类号:TP391.41 **文献标识码:**A

doi:10.37188/OPE.20253320.3299

CSTR:32169.14.OPE.20253320.3299

Multi-granularity segmentation of meibomian gland images based on CNN and Transformer

YANG Song¹, XIA Zhenping^{1*}, LI Li², WU Yanshu¹

(1. School of Electronics and Information Engineering, Suzhou University of Science and Technology,
Suzhou 215009, China;

2. Department of Ophthalmology, Fujian Provincial Hospital, Fuzhou 350001, China)

* Corresponding author, E-mail: xzp@usts.edu.cn

Abstract: To address the multi-stage processing and edge blurring issues in meibomian gland image segmentation, this paper designed an end-to-end multi-granularity segmentation algorithm. During the encoding phase, the TransUNet encoder architecture was adopted to efficiently extract shared features of the eyelid and glandular regions. In the decoding phase, a dual decoding path was employed to set up different decoder branches for the unique features of the eyelid and glandular regions. Meanwhile, for the glandular region, a multi-scale feature fusion module was designed, and a channel attention mechanism was incorpo-

收稿日期:2025-05-14;修订日期:2025-07-18.

基金项目:国家自然科学基金资助项目(No. 62002254);江苏省自然科学基金资助项目(No. BK20200988);江苏省研究生科研与实践创新计划项目(No. SJCX24_1866)

rated into the skip connections. These optimizations improved edge accuracy, texture clarity, and shape contour, thereby effectively solving the problems of edge blurring and glandular adhesion. For the eyelid region, a standard decoder structure was used for segmentation prediction. Through experimental comparison with existing advanced segmentation methods, the proposed algorithm exhibits excellent performance in terms of the average accuracy for the upper and lower meibomian glands. Especially on the key indicators of mean Intersection over Union (IoU) and Dice Similarity Coefficient, it reaches 79.9% and 76.5% respectively, which are 3.2% and 5.3% higher than those of TransUNet. The algorithm in this paper can accurately segment the target regions of meibomian gland images, which can provide a basis for the auxiliary diagnosis of meibomian gland dysfunction.

Key words: meibomian gland image segmentation; multi-granularity segmentation; CNN; Transformer; medical image processing

1 引言

医学图像分割作为计算机辅助检测流程的第一步,旨在识别感兴趣对象的轮廓或区域信息^[1],其核心目标是将医学图像中的目标病变组织与无关组织分离^[2],精准勾画出感兴趣的区域,例如患者体内的不同组织结构、异常病变或解剖结构^[3]。这一技术为病情的定性、定量分析及治疗规划提供了重要依据,是现代医疗诊断与治疗不可或缺的一环^[4]。

睑板腺位于眼睑的睑板内,属于皮脂腺,其主要功能是合成和分泌脂质,形成泪液的脂质层,从而有效防止泪液的过度蒸发,保护眼部的生理环境^[5]。研究表明,睑板腺的功能与其形态密切相关,而睑板腺的堵塞、炎症以及激素水平的变化都会影响其形态,导致油脂分泌功能下降,从而影响患者的生活质量^[6],最终引发睑板腺功能障碍 (Meibomian Gland Dysfunction, MGD)^[7-8]。我国干眼的患病率高达 31.4%^[9]。Lemp 等人^[10]发现,超过 85% 的干眼患者同时患有睑板腺功能障碍。因此,准确诊断睑板腺功能障碍对于干眼症的诊断具有重要意义。睑板腺形态缺失作为诊断指标之一,具有重要临床价值。腺体萎缩百分比,即腺体损失面积与总眼睑面积的比值,是评估睑板腺功能障碍严重程度的重要临床因素^[11]。

非接触式红外成像是一种重要的诊断工具,通过红外光成像技术能够获取睑板腺图像。基于这些图像,医生可以评估睑板腺的形态特征,包括腺体的脱落面积、宽度、长度和弯曲度^[12]。

其中,腺体损失的测量对诊断睑板腺功能障碍具有重要临床意义。腺体萎缩百分比作为评估睑板腺功能障碍严重程度的关键指标,使用标准化的分级量表能够进行有效分类^[13]。睑板腺红外成像能够帮助医生定量分析患者的睑板腺图像,从而实现早期疾病严重程度的判断,制定更精准的治疗方案。

在众多分割方法中,传统的阈值分割因其原理简单、效率高而被广泛采用,二维熵分割法作为其中的代表,常被用于医学图像中组织与背景的自动区分^[14]。随着深度学习的发展,卷积神经网络 (Convolutional Neural Networks, CNN) 已广泛应用于各类分割任务,越来越多的研究将其应用于医学图像处理中,如乳腺癌分割^[15],皮肤病变分割^[16],直肠息肉分割^[17],肺分割^[18],视网膜图像分割^[19]等。近年来,深度学习方法在眼科领域的应用也取得了进展,例如在睑板腺图像分割任务中张贺童等人^[20]采用深度学习方法对眼睑区域进行分割,通过限制眼睑区域来增强腺体特征,从而提升眼科医生对睑板腺病情的辨识能力。此外,还有研究直接对腺体进行分割,其中包括将睑板腺图像划分为更小的子区域, CNN 对每个子区域进行独立分析,从而实现更精确的睑板腺诊断^[21-22]。张祖辉等人^[23]则通过引入预训练的 Mask R-CNN 模型来增强腺体的分割效果。然而,对于临床诊断而言,仍需经过两次训练来分别提取腺体区域和眼睑区域,这一过程需要进一步优化以提升效率和准确性。

这些研究成果为本研究奠定了理论基础,并为睑板腺功能障碍的诊断提供了明确的方向。

然而,目前关于睑板腺的分割大多将腺体分割和眼睑区域分割视为两个独立任务,未充分考虑二者之间的相关性。这一不足不仅浪费了计算资源,还可能影响诊断和治疗决策的效率。

在医学图像处理领域,基于U型结构和跳跃连接的深度神经网络模型,已被广泛应用于各种分割任务^[24]。U型网络结构解码层能够重用相应编码层特征图,更好地融合高低级语义特征^[25],在医学图像分割中展现了较强的鲁棒性和适应性。然而,CNN在捕获全局上下文信息方面存在固有局限性,尤其在处理复杂且小型病变区域时更为明显^[26]。由于腺体形态不规则、结构复杂且容易粘连,CNN可能忽略全局语义信息,导致分割结果的边缘模糊和形态丢失。

尽管DDUNet^[27]和UNet++^[28]通过多层次特征融合缓解了这一问题,但其在全局特征建模能力上的不足依然存在。特别是在同时处理多个腺体和复杂背景的需求下,这一问题更加突出。因此,设计能够平衡全局信息捕获与局部细节保留的分割模型成为解决这一难题的关键。

为了克服现有睑板腺图像分割算法在多阶段处理和边缘模糊方面的不足,本文提出了一种基于TransUNet^[29]编码器的创新网络架构。该架构通过共享编码器高效提取共性特征,并采用双解码器策略,分别应对眼睑区域的粗粒度分割和腺体区域的细粒度分割需求。受到SCU-Net++^[30]的启发,通过在腺体区域的解码器中结合了多尺度特征融合和通道注意力模块,有效增强了高频信息的捕获能力,而眼睑区域的解码器则保留传统结构,用于完成分割任务。该方法提高了两类区域的同步分割效率和精度,通过使用预训练参数加速模型收敛,缓解了小样本问题。实验结果表明,该方法在睑板腺图像分割任务中具有较好的性能表现,为相关研究提供了新思路。

2 方法

2.1 图像预处理

在医学图像分割任务中,图像的质量和清晰度直接影响分割效果。然而,受拍摄角度、光线条件及设备性能等因素的限制,实际采集的图像往往存在模糊、噪声和反光等问题。这些缺陷对

后续图像分析和处理带来显著挑战,可能导致分割精度降低,进而干扰医学诊断的准确性。在睑板腺图像分割任务中,泪液反光和腺体区域纹理复杂性尤为突出,这不仅使得腺体边界变得模糊,还可能引发腺体区域与背景的混淆,从而增加了分割的难度,并进一步导致腺体面积测量和功能评估的偏差。这种图像质量问题如果得不到有效改善,可能会影响下游任务,如腺体萎缩分析、干眼症诊断及疗效评估的可靠性。

为了解决这一问题,本文提出了一种结合(Hue Saturation Value, HSV)颜色空间分解和快速行进方法(Fast Marching Method, FMM)^[31]以及灰度直方图均衡化的预处理策略。首先,将睑板腺图像转换到HSV色彩空间,以亮度作为主要参考维度。HSV色彩空间通过将颜色信息分为色相、饱和度和亮度三部分,能够有效地区分图像中的颜色信息与亮度信息。在睑板腺图像中,泪液反光区域通常表现为高亮区域,且往往伴随有强烈的亮度变化,因此,亮度分量在此任务中具有决定性的作用。

为了精确提取泪液反光区域,本文使用了基于动态亮度阈值计算的方法。首先,通过对图像的亮度通道进行统计分析,计算出每张图像的亮度均值和标准差。然后,动态设置亮度阈值,这一阈值依据每张图像的亮度特征进行调整,以适应不同图像的光照变化。在实际操作中,利用此动态阈值,能够更加精准地提取出图像中的反光区域,同时有效避开由于不同光照条件所带来的亮度差异。

接着,引入FMM算法对提取的反光区域进行修复。在修复过程中,FMM基于波前传播的原理,利用已知区域的像素灰度和纹理信息,通过迭代更新像素到达时间,逐步推测并填充反光区域的未知像素值。FMM是一种高效的数值计算方法,能够描述波前传播的动态过程。在图像处理领域,FMM被广泛应用于图像修复、距离变换和形状提取等任务。通过该方法,可以有效地修复图像模糊区域,恢复腺体边界的清晰度,从而提高图像分割的准确性。FMM算法过程如式(1)所示:

$$P(p) = C(p) \cdot D(p), \quad (1)$$

其中: $C(p)$ 反映当前点的已知信息信任程度, D

(p) 反映当前点的修复方向是否与纹理边界一致,其中 $C(p)$ 的定义如式(2)所示:

$$C(p) = \frac{\sum_{q \in \Psi_p} C(q)}{\text{Area}(U_p)}, \quad (2)$$

其中: U_p 表示点 p 周围的补丁区域, $C(q)$ 表示补丁内已知像素点的置信度值, $D(p)$ 的定义如式(3)所示:

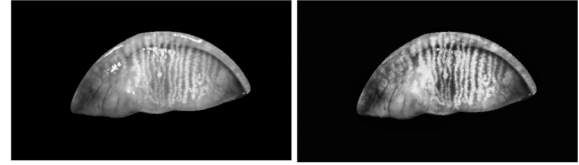
$$D(p) = \frac{|\nabla I^\perp \cdot n|}{\alpha}, \quad (3)$$

其中: $\nabla I^\perp$ 表示图像梯度的垂直分量,用于描述纹理变化。 n 表示边界法向量。 α 归一化因子,用于控制优先级范围。

从修复边界中选取优先级最高的点 p ,在其邻近的已知区域中寻找最相似的补丁,并将补丁的像素值复制到待修复补丁。随后,更新修复边界和置信度,当待修复区域的所有像素均被填充时,修复完成。在修复完成后,对每张睑板腺图像进行灰度直方图均衡化,加大腺体和背景的分度。

图1展示了睑板腺图像预处理前后的效果对比图,在睑板腺图像修复的具体应用中,FMM和

灰度直方图均衡化有效缓解了泪液反光导致的腺体边界模糊问题,使腺体区域的轮廓更加清晰。



(a) 处理前 (b) 处理后
(a) Image before preprocessing (b) Image after preprocessing

图1 睑板腺图像预处理前后对比图

Fig. 1 Meibomian gland images before and after preprocessing

2.2 模型结构

本文所提出的网络结构如图2所示。该架构通过共享编码器和双解码器设计,实现了高效的特征提取和精确的分割。具体而言,该模型包括四个关键模块:编码器模块、多尺度特征融合模块、双解码器模块和通道关注模块。在编码阶段,首先利用CNN从输入图像中提取特征,通过降采样减少图像分辨率,同时增加通道数,以便

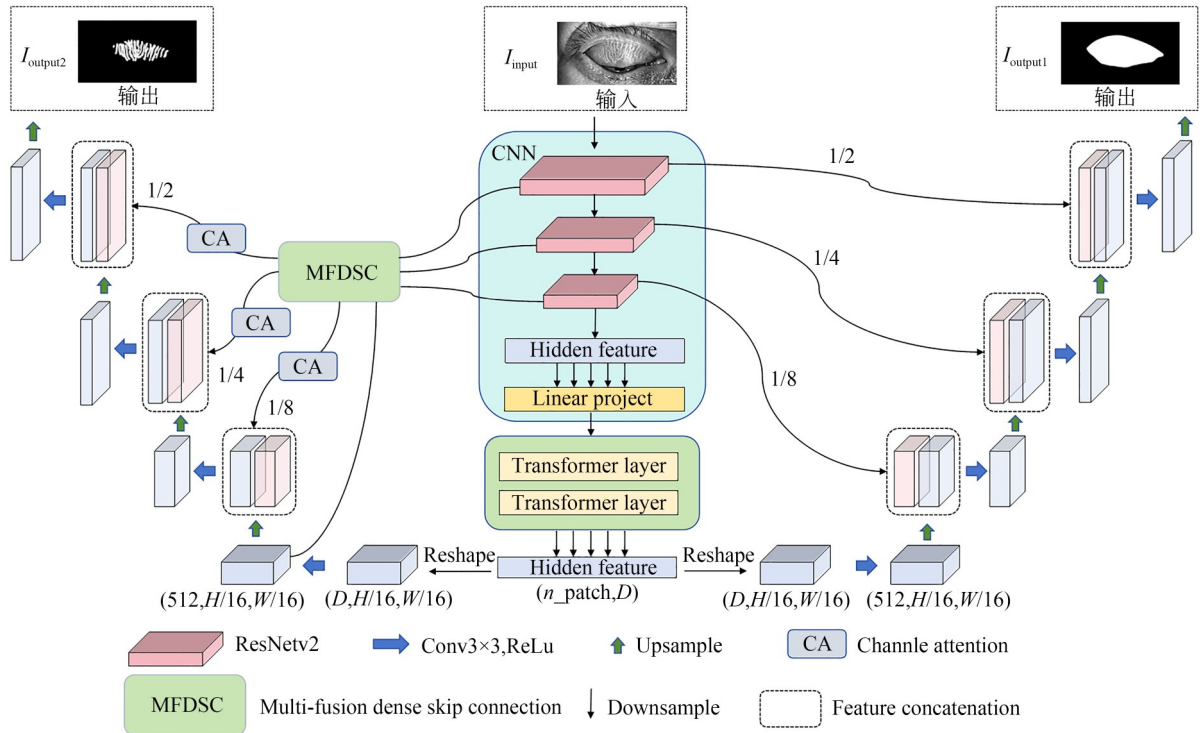


图2 整体网络结构

Fig. 2 Overall network structure

捕获局部特征。接着,将大小为 $W \times H$ 的图像划分为多个 16×16 的不重叠块,并通过线性嵌入层将这些块的特征映射为 C 维。经过线性变换后,图像块被输入至一系列 Vision Transformer^[32] 模块,用于捕获全局图像特征。

特征提取完成后,得到的特征向量被重塑为 $W \times H$ 的特征图。然后,特征图被送入两个独立的卷积模块,以提取眼睑区域和腺体区域的局部特征并进行上采样处理。

对于眼睑区域,采用 CNN 减少通道数后进行上采样,提升图像分辨率,并通过跳跃连接补充上采样过程中丢失的低级特征。对于腺体区域,特别是边缘模糊且与背景对比度低的情况,引入了多尺度特征融合模块和通道关注模块,以增强对细节的识别能力。最后,采用扩展层进行两倍上采样,将特征图恢复到原始输入分辨率,并通过线性投影层进行像素级分割预测。

为了提高分割效率并降低计算资源需求,该架构采用了预训练参数锁定共享编码器,以提取腺体和眼睑区域的共同特征。针对不同的分割任务,设计了两个独立优化的解码器路径,从而高效地解决眼睑区域与腺体区域的分割问题。

2.3 编码器

在编码阶段,本研究采用了一种结合 CNN 与 Vision Transformer 优势的复杂编码器结构,用于高效提取眼睑和腺体区域的特征。CNN 通过卷积层捕获图像的局部细节特征,在腺体区域的高精度细节提取中表现突出;同时,Vision Transformer 借助自注意力机制建模全局依赖关系,有效增强全局特征理解能力,尤其适用于眼睑区域的分割任务。眼睑区域因其复杂背景和全局轮廓特征的依赖性,需要更多上下文信息来确保分割的准确性,而 Vision Transformer 的全局建模能力恰好满足这一需求。

其中对于 CNN 部分使用 ResNetv2 模型结构,ResNetv2 是 ResNet 的优化版本,通过调整批归一化和激活函数位置,并采用预激活策略,加速了训练过程。

Vision Transformer 的结构如图 3 所示,其结构由嵌入层、Transformer 编码器和多层感知头组成。嵌入层将输入图像分割成固定大小的图

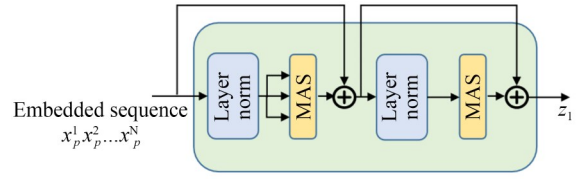


图 3 Vision Transformer 结构

Fig. 3 Vision Transformer structure

块,并通过展平和线性映射转化为特征向量,再通过位置信息嵌入形成完整的序列表示。在 Transformer 编码器中,这些序列经过多层自注意力机制和前馈神经网络处理,提取全局上下文特征,从而增强模型的分割能力。这种模块化结构实现了全局与局部特征的互补,显著提升了模型的分割性能,其中 patch 嵌入的过程如式(4)所示:

$$z_0 = [x_p^1 E; x_p^2 E; \dots; x_p^n E] + E_{\text{pos}}, \quad (4)$$

其中: E 为补丁嵌入投影, E_{pos} 表示位置嵌入。编码器由 L 层多头自注意力机制(Multi-Head Self-Attention, MSA)组成和多层感知机(Multi-Layer Perceptron, MLP)因此每一层的输出的表达如式(5)所示:

$$\begin{aligned} z_{\lambda'} &= \text{MSA}(\text{LN}(z_{\lambda-1})) + z_{\lambda-1}, \\ z_{\lambda} &= \text{MLP}(\text{LN}(z_{\lambda'})) + z_{\lambda'} \end{aligned} \quad (5)$$

其中: $\text{LN}(\cdot)$ 表示层归一化操作符, z_{λ} 为编码后的图像表示。

本文的模型架构采用两层 Transformer 模块,充分利用图像的全局上下文信息,增强模型表示能力。首先,通过 ResNetV2 提取图像的高层语义特征图,然后将特征图划分为固定大小的图像块。接着,使用线性变换将每个图像块映射到低维特征空间,生成嵌入表示。这些嵌入向量结合位置编码和可学习的嵌入向量,提供丰富的空间位置信息,使模型能够同时捕捉局部特征和全局上下文,从而提升分割精度。

2.4 多重融合密集跳跃连接

为了有效应对腺体分割任务中常见的边缘模糊问题,本文提出的模型借鉴了密集连接的理念,采用了一种多尺度融合的密集跳跃连接方法。通过引入多尺度语义信息融合机制,模型能够在不同尺度上全面提取关键特征,兼顾解码阶段的灵活性与编码器性能的充分发挥。密集跳跃连接通过嵌套的多融合结构,充分整合来自不

同尺度和层次的特征信息,显著提升了特征表达的全面性和准确性。该设计不仅优化了模型的上下文理解能力,还能在腺体区域中精确捕捉微观结构和复杂纹理特征。

多重融合密集跳跃连接模块的结构如图 4 所

示。在该框架中,融合多尺度语义信息为模型提供了全局视野与局部细节的结合。多层次、多尺度的特征捕捉机制通过分层递进的方式,逐步积累丰富的语义信息,缓解了编码器和解码器模块间可能存在的语义鸿沟。

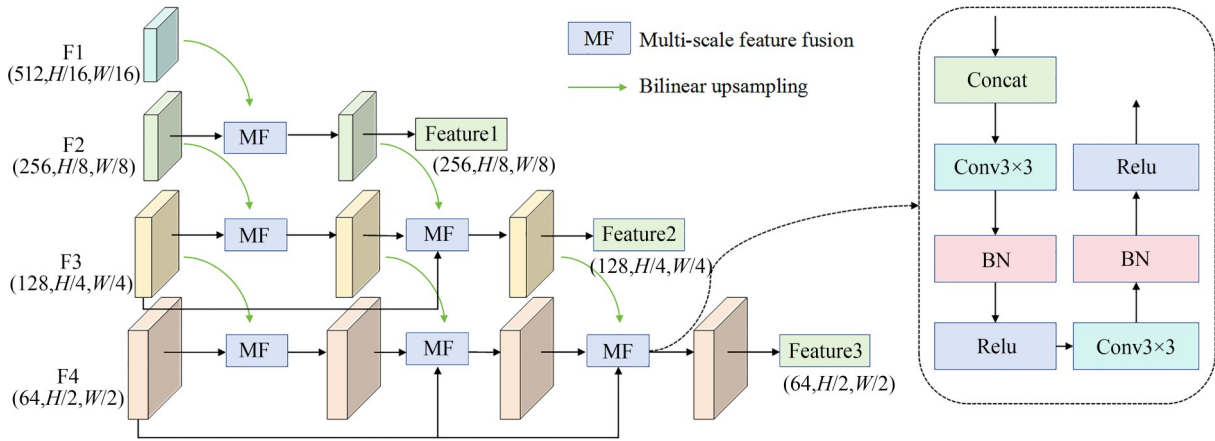


图 4 多重融合密集跳跃连接模块

Fig. 4 Multiple fusion dense jump connection module

2.5 通道注意力

通道注意力模块的结构如图 5 所示。该模块的核心设计理念在于通过选择性增强输入特征，

对与任务相关的特征通道进行突出,同时抑制无用或冗余通道,从而提高网络对关键信息的识别、提取和利用能力。

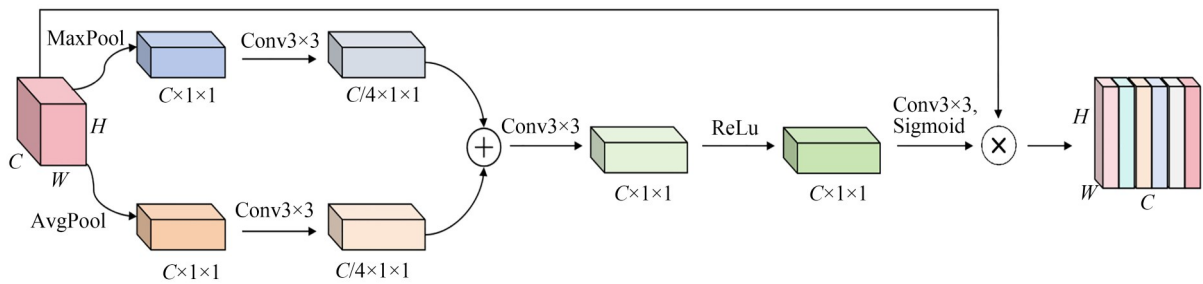


图 5 通道注意力模块

Fig. 5 Channel attention module

模块的处理流程分为多个关键阶段,旨在实现高效的特征融合和动态优化。首先,模块对输入特征图执行全局平均池化和最大池化操作,分别提取特征图的全局统计信息和显著性信息。全局平均池化能够捕获特征的整体分布,而最大池化则聚焦于提取显著性特征。两个池化操作从不同角度为特征融合提供了丰富的先验信息。

接着,池化结果被输入到共享权重的卷积层进行处理。通过卷积操作,模块能够捕获特征的

局部细节并实现初步的特征融合。为了增强模型的泛化能力,融合过程中引入了正则化技术,以避免因特定特征的过度学习而导致模型过拟合。此外,ReLU 激活函数被引入卷积层之间,增加了非线性建模能力,提升了特征表达的多样性和有效性。

在特征融合阶段,模块将全局平均池化和最大池化得到的特征表示进行合并,并通过两个连续的 3×3 卷积层进一步细化特征。

最后,经过特征融合的结果通过 Sigmoid 激活函数进行归一化,通道权重被限制在 0 到 1 的范围内,从而实现对各通道特征的动态选择。这些权重反映了每个通道在当前任务中的重要性,能够突出关键特征并抑制噪声与冗余信息。最终,生成的通道权重被逐元素应用到原始输入特征图中,得到加权后的特征图。通过这种加权方式,任务相关特征的表达能力得到了显著增强,同时对模型预测无关或干扰信息的影响也得到了有效抑制。

2.6 解码器

在模型架构中,解码器是关键组件,负责将编码器提取的高级语义特征逐步还原到原始图像尺寸,同时生成精确的分割掩码。在编码器通过卷积和池化层对图像进行下采样时,尽管获得了丰富的高级语义信息,但空间分辨率显著降低。为了解决这一问题,解码器通过上采样操作逐步恢复特征图的分辨率,使其接近输入图像的尺寸。

为了实现眼睑和腺体的精准分割,解码器被设计为两个独立分支,这些分支共享编码器提取的特征,但分别使用不同的标签和损失函数进行训练,从而独立生成针对眼睑和腺体的分割结果。

在解码的每个阶段,上采样后的特征图通过跳跃连接与编码器对应阶段的特征图融合。这种融合方法将浅层的精确位置信息与深层的语义特征有机结合,生成更加丰富的特征表示。融合后的特征图进一步通过卷积层处理,以增强类别之间的区分能力。通过这种迭代处理,特征图的分辨率逐步恢复到与输入图像一致。

在解码器的最终阶段,使用 1×1 卷积层将高维特征图转换为类别得分图,每个像素的得分代表其属于不同类别的概率,最终生成精确的分割结果。双分支解码器的设计大幅提升了眼睑和腺体的分割效率。

2.7 损失函数

在睑板腺图像分割任务中,类别不平衡问题尤为显著,主要表现为分割区域的组织区域与背景区域之间像素数量的巨大差异。这种不平衡容易导致机器学习模型在训练过程中偏向多数类别,进而忽视占比较小但关键的类别,例如分

割目标区域。为应对此问题,本文设计了一种复合损失函数,将加权交叉熵损失^[33]与 Dice 损失^[34]相结合,从而有效增强模型对不平衡类别的识别能力,并提升整体分割精度。

Dice 损失的引入进一步优化了模型在分割任务中的表现。它通过直接度量预测分割区域与实际分割区域的重叠程度,显著改善了边界处理能力。其中加权交叉熵损失如式(6)所示:

$$L_{\text{weighted}} = -\frac{1}{N} \sum_{i=1}^N [\alpha y_i \log(p_i) + \beta (1 - y_i) \log(1 - p_i)], \quad (6)$$

其中: N 代表样本数量。 y_i 第 i 个像素的真实标签, P_i 第 i 个像素的预测概率,表示模型预测该像素属于目标区域的概率, α 代表目标类的加权系数。该参数用于调整目标类的重要性。当目标类样本较少时,增加 α 的值可以使模型更加关注目标区域的分割,减少背景区域对损失的影响。 β 表示背景类的加权系数。用于调节背景类的重要性。

此外,通过为睑板腺和腺体设置不同的权重因子,本文的方法确保模型在训练过程中能够充分关注这些关键的少数类别,从而显著提升了它们的识别和分割效果。这种策略不仅有效提高了分割精度,还增强了模型在复杂医疗图像处理中的鲁棒性,展现出在实际医疗图像分析中的高度实用性和可靠性。

除了传统方法,DiceLoss 作为一种基于区域的评估指标,能够直接衡量预测分割区域与目标区域之间的重叠程度。与传统的逐像素损失函数不同,DiceLoss 能够更全面地反映分割结果在空间上的一致性和形状的准确性。DiceLoss 计算如式(7)所示:

$$\text{DiceLoss}(X, Y) = 1 - \frac{2|X \cap Y|}{|X| + |Y|}, \quad (7)$$

其中: X 表示预测结果的二值化图像, Y 表示真实标签的二值化图像。

在睑板腺分割任务中,结合交叉熵损失与 DiceLoss 能有效缓解类别不平衡问题,特别是在目标与背景区域比例失衡的情况下。这种组合不仅增强了模型对关键区域的识别能力,还改善了预测区域与实际区域的重叠精度,减少误分类。同时,交叉熵与 DiceLoss 的结合在边缘区域

表现出更强的鲁棒性,能够精准捕捉细节,尤其是在边缘模糊或部分遮挡的情况下。这使得该方法在复杂医学图像分割任务中提供了更加稳定和可靠的结果,为临床诊断和图像分析提供有力支持。

3 实验结果和分析

3.1 实验数据与评价指标

在本研究中,本文采用 MGD-1k^[35]数据集对模型性能进行了全面评估。MGD-1k 是首个针对睑板腺的图像数据集。该数据集包含 1 000 张睑板腺图像,其中包括 467 张上眼睑图像和 533 张下眼睑图像,数据集分为 800 张训练集和 200 张测试集。在数据集中,正常睑板腺图像有 180 张,一级病变图像有 612 张,二级病变图像有 167 张,三级病变图像有 40 张。所有图像均由专业医师进行了严格标注,包含眼睑区域和腺体区域的分割标签,为模型训练和验证提供了精确的基准。

在模型性能评估方面,本文采用了学术界广泛认可的两种评价指标: Dice 相似系数、(mean Intersection over Union, mIoU) 平均交并比。这些指标反映了分割质量的不同维度,是衡量图像分割性能的重要手段。

Dice 用于测量预测分割掩膜与实际分割标签之间的重叠程度。它通过计算预测与真实标签交集的两倍大小与两者总大小之比来得出。Dice 的计算如式(8)所示:

$$Dice = \frac{2|X \cap Y|}{|X| + |Y|}, \quad (8)$$

其中: X 表示预测结果的二值化图像, Y 表示真实标签的二值化图像。该指标为重叠程度提供了一种稳健的度量方式,尤其适用于边界精确分割至关重要的任务。通过引入这些指标,能够全面评估模型在图像分割任务中的有效性和准确性。

mIoU 用于评估模型预测的准确性,定义为预测区域与实际分割区域的交集与并集之比。mIoU 的计算如式(9)所示:

$$mIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{TP}{FN + FP + TP}, \quad (9)$$

其中: TP 表示正确分割的目标像素数量,即实际为正例且预测为正例的像素数量; FP 表示错误

分割的目标像素数量,即实际为负例但预测为正例的像素数量, FN 表示错误分割的目标像素数量,即实际为正例但预测为负例的像素数量; k 表示分割目标的数量。

通过严格应用这些评估指标,能够有效地评估模型在图像分割任务中的表现。通过系统地分析模型在不同维度上的表现,能够明确其适用范围和局限性,为实际应用提供可靠的理论支持。

此外,采用这些标准化的评估指标确保了研究的公平性和客观性,使得本文的研究能够与其他研究工作进行比较。

3.2 实验设置

在本研究中,本文采用 PyTorch 框架构建并训练模型,并使用 NVIDIA GeForce RTX 3090 GPU 以实现高效训练。在训练过程中,本文将批量大小设置为 4,并使用 Adam 优化算法更新模型梯度,一阶矩和二阶矩指数衰减率分别设置为 $\beta_1=0.9$, $\beta_2=0.999$, 为确保充分学习,本文将训练配置为运行 20 个 epoch,并设定初始学习率为 0.01。

为了在性能与计算复杂度之间取得平衡,我们将 Transformer 编码器的层数设置为 2 层。该参数设置基于初步的消融实验结果,实验中发现继续增加层数对性能提升有限,甚至在部分样本上出现过拟合现象。因此,本文采用两层结构进行实验。

为确保模型比较的公平性,本文针对上眼睑和下眼睑分别进行独立训练。上眼睑图像分为 345 张训练集和 40 张验证集以及 82 张测试集,下眼睑图像分为 375 张训练集和 40 张验证集以及 118 张测试集。这种划分方式有效避免了数据分布的不平衡对模型性能评估的影响,确保了实验结果的客观性和可靠性。

3.3 对比实验

在对比实验中,本文将所提出的模型与其他先进的分割模型进行性能比较,包括 TransUNet, Rolling-UNet^[36], 以及经典的分割模型 U-Net^[37] 和 Deeplabv3+^[38]。本文对提出的模型与传统分割模型的分割效果进行了详细的对比分析。

其结果如表 1 和表 2 所示,实验显示,所提出

表 1 不同算法在上睑板腺分割定量对比

Tab. 1 Quantitative comparison of different algorithms for upper meibomian gland segmentation

算法	mIoU	Dice
Deeplabv3+	77.9%	73.4%
Rolling-UNet	76.6%	71.2%
TransUNet	74.8%	68.2%
UNet	77.1%	71.2%
Dual-TransUNet (Ours)	78.8%	74.7%

的模型在睑板腺图像分割任务中表现卓越,优于现有主流模型,取得了领先的性能指标。

分割结果如图 6 所示,通过与其他方法的对比,模型在面对复杂背景、形状不规则或含有噪

表 2 不同算法在下睑板腺分割定量对比

Tab. 2 Quantitative comparison of different algorithms for lower meibomian gland segmentation

算法	mIoU	Dice
Deeplabv3+	78.3%	73.1%
Rolling-UNet	77.6%	72.1%
TransUNet	78.7%	73.5%
UNet	79.5%	74.4%
Dual-TransUNet (Ours)	81.0%	77.4%

声的图像时,依然能够保持较高的精度,这表明该模型具备很好的鲁棒性。与传统方法相比,本文提出的模型在处理这些高难度的医学图像时展现出了更加稳定的性能。

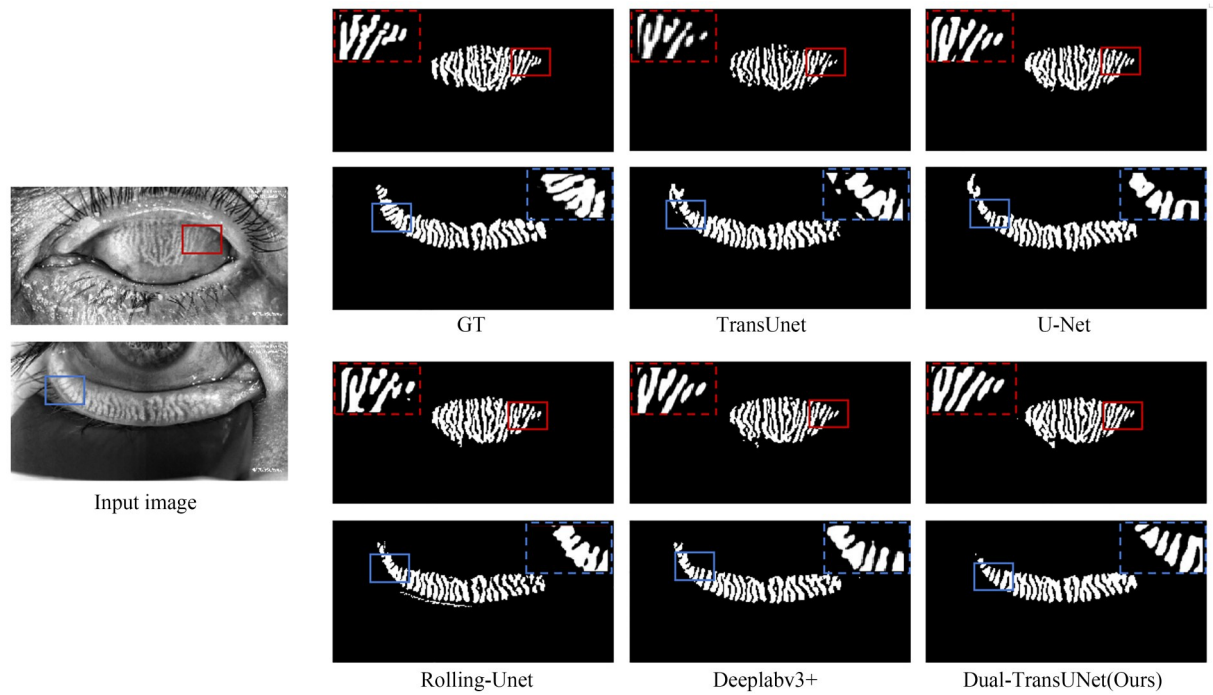


图 6 不同模型在上眼睑和下眼睑下腺体分割效果对比

Fig. 6 Comparison of segmentation performance of different models on the upper and lower eyelid meibomian glands

研究中引入的多尺度融合与密集跳跃连接策略极大地增强了模型的特征表达能力,使其能够充分融合来自不同尺度和层次的信息。这一创新设计提高了分割的全面性和准确性,尤其在处理复杂细节和微小结构时表现出了强大的优势。通过这种方式,模型能够在多个尺度上有效提取关键信息,并在解码阶段保持高精度的特征

表达,从而提升了分割效果。多尺度融合方法帮助模型从不同的视角理解图像,提升了模型对细节的感知能力,并确保了分割结果的精确度。

表 1 比较了多种分割算法在上睑板腺分割任务中的定量性能,采用 mIoU 和 Dice 两个指标进行评估。结果表明,传统分割模型 UNet 和 Deeplabv3+ 在 mIoU 和 Dice 指标上表现较为突出,其

中 Deeplabv3+ 综合性能更为均衡。相比之下, Rolling-UNet 和 TransUNet 的整体性能相对较低。

本文提出模型在 mIoU 和 Dice 指标上分别取得了 78.8% 和 74.7% 的最高分, 超越了其他分割算法。特别是在区域重叠度和分割一致性方面, 改进模型展现出明显优势。

表 2 中对比了多种分割算法在下睑板腺分割任务中的性能, 使用 mIoU 和 Dice 作为评价指标。结果显示, 传统模型 UNet 和 Deeplabv3+ 整体表现较好, 特别是 UNet, 在 mIoU 和 Dice 上分别达到 79.5% 和 74.4%。Rolling-UNet 的性能略低, mIoU 和 Dice 分别为 77.6% 和 73.1%。另

外, 本文展示了不同模型进行分割预测的可视化对比, 如图 6 所示, 其中 GT (Ground Truth) 代表原始的目标图像。

3.4 消融实验

文的消融研究重点分析了三个关键模块的作用: 锁定参数训练、通道注意力模块和多尺度特征融合模块。图 7 展示了消融实验中不同模型的分割结果, 其中 GT 代表标签图像, TU 为基准模型, TU-LP 是在基准模型基础上锁定编码器参数的版本, TU-LP-CA 是在锁定参数的基础上引入了通道注意力机制, TU-LP-MS 则在此基础上进一步集成了多尺度密集连接结构, 而 TU-LP-MS-CA 是本文提出的最终模型。

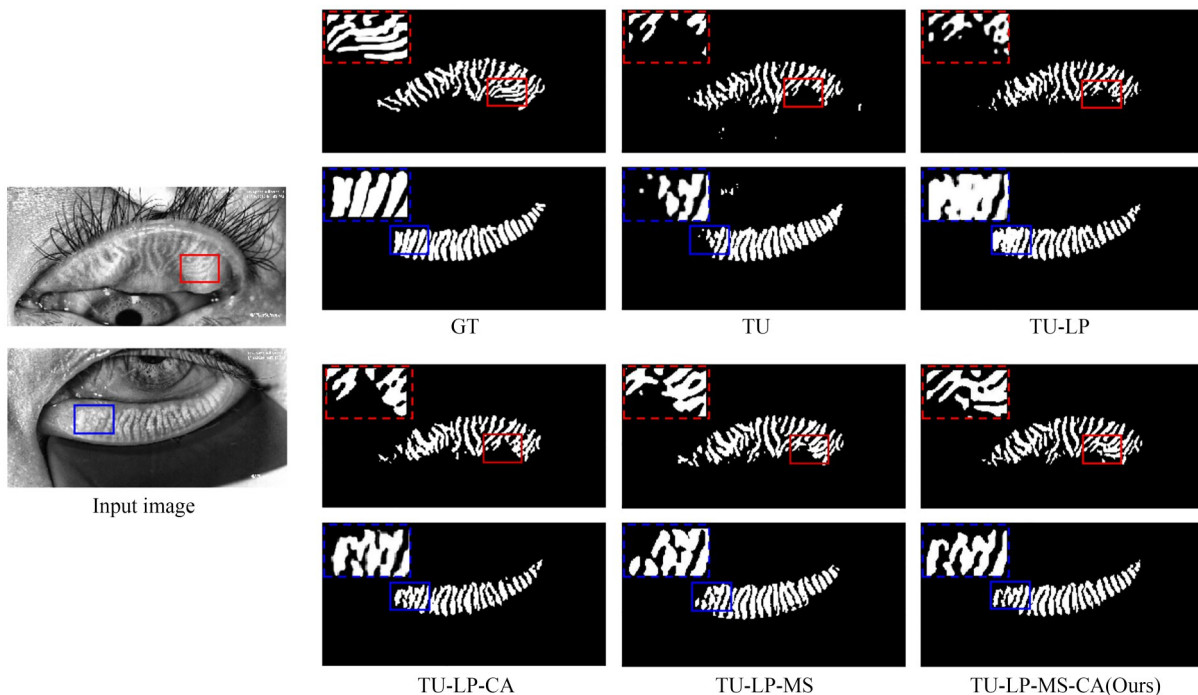


图 7 上眼睑与下眼睑下腺体分割的消融实验对比

Fig. 7 Ablation study comparison for upper and lower eyelid meibomian gland segmentation

从消融实验的结果可以明显看出, 单独使用基准模型时, 尽管模型的总体预测精度较高, 但依然存在背景区域被误预测为目标区域的问题。通过在基准模型中引入参数锁定, 模型的预测精度有所提高, 特别是在目标区域的定位上更为准确。然而, 锁定参数后的模型仍然容易受到背景干扰, 造成目标区域的误预测。

当进一步引入通道注意力机制时, 模型对目标区域的关注得到了增强, 使得细节部分的分割

效果有所改进, 尤其是在复杂边缘和小目标的分割中表现更加突出。结合多尺度密集连接结构后, 模型进一步提升了对不同尺度特征的捕捉能力, 使得分割结果更加全面, 对细节的处理更为精确。

通过结合锁定参数、通道注意力机制和多尺度密集连接模块的 TU-LP-MS-CA 模型取得了最佳效果。此模型能够在保证准确分割目标区域的同时, 减少背景干扰。通过多个模块的有效

融合,模型在精度、细节处理及边缘分割等方面都表现出了较强的适应性,预测结果接近实际标签。

表 3 和表 4 展示了上眼睑和下眼睑任务中消融实验的定量对比结果,分析了各模块对模型性能的影响。在上眼睑任务中,逐步引入不同模块后,模型的 mIoU 和 Dice 系数逐步提高。例如,添加多尺度特征融合(TU-MS)和通道注意力机制(TU-CA)后,模型的 mIoU 和 Dice 系数均有所

提升。接着,加入锁定参数(TU-LP)进一步提升了模型性能。最终,结合多尺度特征融合、通道注意力机制和参数锁定(TU-MS-CA-LP(Ours))的模型,达到了最佳性能。

下眼睑任务的实验结果呈现类似的趋势。随着不同模块的加入,模型的 mIoU 和 Dice 系数逐渐提升,最终通过结合多尺度特征融合、通道注意力机制和参数锁定(TU-MS-CA-LP(Ours))获得了最佳性能。

表 3 不同模块在上眼睑上的消融实验定量对比

Tab. 3 Quantitative comparison of ablation experiments on the upper eyelid for different modules

模型	MFDSC	CA	LP	mIoU	Dice
TU				74.8%	68.2%
TU-MS	✓			77.5%	71.9%
TU-CA		✓		76.4%	69.9%
TU-LP			✓	77.8%	73.1%
TU-CA-LP		✓	✓	78.1%	73.5%
TU-MS-LP	✓		✓	78.7%	74.5%
TU-MS-CA	✓	✓		77.6%	72.4%
TU-MS-CA-LP (Ours)	✓	✓	✓	78.8%	74.7%

表 4 不同模块在下眼睑上的消融实验定量对比

Tab. 4 Quantitative comparison of ablation experiments on the lower eyelid for different modules

模型	MFDSC	CA	LP	mIoU	Dice
TU				78.7%	73.5%
TU-MS	✓			80.2%	75.9%
TU-CA		✓		79.7%	75.1%
TU-LP			✓	78.9%	74.2%
TU-CA-LP		✓	✓	79.2%	74.6%
TU-MS-LP	✓		✓	80.5%	76.8%
TU-MS-CA	✓	✓		79.9%	75.2%
TU-MS-CA-LP (Ours)	✓	✓	✓	81.0%	77.4%

此外,从图中的测试样本中可以看出基础模型在边界模糊或腺体结构不清晰的图像中易出现过分割或漏分问题,而引入多尺度特征融合后,模型在不同尺寸腺体结构的识别上更加稳健;通道注意力机制增强了关键特征通道的响应,有效减少了背景干扰造成的误判。参数锁定机制则在一定程度上提升了模型对不同样本分

布的适应性,降低了性能波动。这些结果进一步验证了各模块对提升模型稳定性与泛化能力的积极作用。

综合表 3 和表 4 的结果可以看出,通过多尺度特征融合和通道注意力机制以及锁参数的结合,模型在上眼睑腺体分割任务中 mIoU 和 Dice 系数分别提升了 4% 和 6.5%,模型在下眼睑腺

体分割任务中 mIoU 和 Dice 系数分别提升了 2.3% 和 3.9%。

接下来,本文探讨了冻结模型编码器参数对分割效果的影响。实验结果表明,冻结编码器参数后,模型分割上眼睑腺体的 Dice 提升至 73.1%,分割下眼睑腺体的 Dice 提升至 74.2%。这一现象表明,在数据量有限的情况下,适当冻结预训练网络层的参数不仅能够有效防止过拟合,还能促进模型在小型数据集上的更好收敛,从而提高了模型的稳定性和性能。

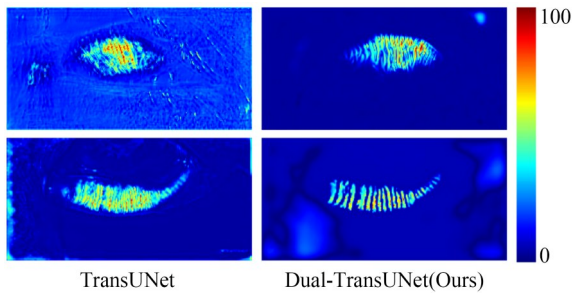


图 8 上眼睑与下眼睑腺体分割特征图对比可视化

Fig. 8 Comparison visualization of upper and lower eyelid gland segmentation feature maps

为了更全面地展示本文所添加模块对分割效果的贡献,本文对模型在分割过程中的特征图进行了可视化,具体结果如图 8 所示。为了便于对比,本文将每个特征图的权重归一化至 0~100 的范围。通过观察特征图,本文可以明显看到,相较于基准模型,本文的模型在分割过程中更加专注于腺体区域,表现出更高的特征权重集中在腺体的边缘和内部区域。而对于无关的背景或区域,特征权重则较低,表明模型能够有效抑制干扰信息,提升分割精度。这些结果验证了本文所添加模块在引导模型聚焦于重要特征方面的有效性,从而增强了分割效果。

3.5 结果分析

睑板腺缺失按 Arita^[39]等人提出的四级划分,0 级:无腺体脱失,1 级:缺失面积小于总腺体区域面积的三分之一,2 级:缺失面积介于三分之一至三分之二之间,3 级:缺失面积超过三分之二。在对眼睑和腺体区域进行分割后,可以计算腺体密度,其计算过程如式(10)所示:

$$Density = \frac{S_{gland}}{S_{lid}}, \quad (10)$$

其中: S_{gland} 代表眼睑区域的面积, S_{lid} 代表腺面积。

在对腺体和眼睑区域进行分割后,本文将数据分为 MGD 组和正常组,其中正常组的分级为 1 级或 0 级,而 MGD 组的分级为 2 级或 3 级。本文选取了 273 张上眼睑图像、273 张下眼睑图像以及 261 对全眼睑图像计算平均腺体密度,并将数据集按 80% 用于训练,20% 用于测试。利用腺体密度对睑板腺分级进行了基于支持向量回归^[40]的拟合,并对结果进行了 Spearman 相关性分析。分析结果如表 5 所示。

表 5 睑板腺密度预测与睑板腺分级的相关性分析

Tab. 5 Correlation analysis between meibomian gland density prediction and meibomian gland grading

类别	正常组	MGD 组	相关系数
上眼睑睑板腺体密度	0.271	0.191	0.736**
下眼睑睑板腺体密度	0.331	0.228	0.812**
全眼睑睑板腺体密度	0.301	0.234	0.768**

注:** $P < 0.01$

表 5 展示了上眼睑、下眼睑以及全眼睑图像中腺体密度的均值,以及预测结果与睑板腺缺失分级之间的 Spearman 相关性分析结果。上眼睑的相关系数表明,腺体密度与人工分级结果之间存在较强的正相关关系;而在下眼睑中,该相关系数进一步提高,说明下眼睑区域的预测结果与人工分级之间具有更高的一致性。此外,全眼睑的相关系数也显示出良好的预测能力。这一结果验证了腺体密度拟合在不同眼睑区域均具备较高的预测能力,且在下眼睑区域表现出更稳定的预测效果。

依托双分支网络结构的优势,系统能够同步提取眼睑区域与腺体区域,实现腺体密度的直接计算与非线性拟合。拟合结果为睑板腺缺失严重程度的自动评估提供了可靠支持。通过该自动化流程,系统可生成一个更加精准且具有可解释性的缺失严重程度评价指标,为临床辅助诊断提供有力依据。

4 结 论

本文提出了一种端到端的多粒度分割算

法,针对睑板腺图像分割中的多阶段处理和边缘模糊问题进行优化。算法采用TransUNet编码器架构,能够高效提取眼睑和腺体区域的共享特征;在解码阶段,设计了双解码路径,分别为眼睑和腺体区域设置了不同的解码器分支。此外,引入多尺度特征融合模块和通道注意力机制,进一步提升了腺体区域的边缘精度、纹理清晰度和形状轮廓,有效解决了传统方法中的边缘模糊和腺体粘连问题。实验结果表明,该算法上下睑板腺的精度均值上表现优异,在mIoU和Dice相似性系数上分别达到了79.9%和76.5%,相比基准模型TransUNet,分别提高了3.2%和5.3%。为医学图像分析领域的自动化和精度提升提供了有效技术支持。该方法能

够精准地提取腺体区域并计算腺体密度,为临床上睑板腺缺失的判断提供了科学依据,特别是在自动化评估睑板腺缺失严重程度方面具有重要应用前景。

作者贡献声明:

杨松:论文构思,实验验证,结果可视化展示,论文初稿撰写;

夏振平:论文构思,方法设计,项目经费获取,论文审核与撰写指导;

李笠:研究指导,理论与形式化分析,资源支持与平台保障;

吴彦澍:调查研究,数据处理与管理,项目执行与统筹。

参考文献:

- [1] 吴瞳,胡浩基,冯洋,等. 分割一切模型(SAM)在医学图像分割中的应用[J]. 中国激光, 2024, 51(21): 27-42.
WU T, HU H J, FENG Y, *et al.* Application of segment anything model in medical image segmentation[J]. *Chinese Journal of Lasers*, 2024, 51(21): 27-42. (in Chinese)
- [2] 高威,蒋慧,焦一平,等. 基于多任务和注意力的胰腺癌全切片图像多组织分割模型[J]. 生物医学工程学杂志, 2023, 40(1): 70-78.
GAO W, JIANG H, JIAO Y P, *et al.* Multi-tissue segmentation model of whole slide image of pancreatic cancer based on multi task and attention mechanism [J]. *Journal of Biomedical Engineering*, 2023, 40(1): 70-78. (in Chinese)
- [3] 徐昌佳,易见兵,曹锋,等. 采用DoubleUNet网络的结直肠息肉分割算法[J]. 光学精密工程, 2022, 30(8): 970-983.
XU C J, YI J B, CAO F, *et al.* Colorectal polyp segmentation algorithm using DoubleUNet network [J]. *Opt. Precision Eng.*, 2022, 30(8): 970-983. (in Chinese)
- [4] 言方荣. 人工智能在生物医药领域中的应用和进展[J]. 中国药科大学学报, 2023, 54(3): 263-268.
YAN F R. Application and advance of artificial intelligence in biomedical field [J]. *Journal of China Pharmaceutical University*, 2023, 54 (3) : 263-268. (in Chinese)
- [5] 蒋冬冬,新荷. 睑板腺功能障碍相关干眼的诊疗进展[J]. 国际眼科杂志, 2021, 21(7): 1209-1212.
JIANG D D, JIN H. Progress in diagnosis and treatment of meibomian gland dysfunction associated with dry eye [J]. *International Eye Science*, 2021, 21(7): 1209-1212. (in Chinese)
- [6] VU C H V, KAWASHIMA M, YAMADA M, *et al.* Influence of meibomian gland dysfunction and friction-related disease on the severity of dry eye [J]. *Ophthalmology*, 2018, 125(8): 1181-1188.
- [7] KNOP N, KNOP E. Meibomian glands. Part I: anatomy, embryology and histology of the Meibomian glands [J]. *Der Ophthalmologe*, 2009, 106(10): 872-883.
- [8] WISE R J, SOBEL R K, ALLEN R C. Meibography: a review of techniques and technologies [J]. *Saudi Journal of Ophthalmology*, 2012, 26(4): 349-356.
- [9] SONG P G, XIA W, WANG M L, *et al.* Variations of dry eye disease prevalence by age, sex and geographic characteristics in China: a systematic review and meta-analysis [J]. *Journal of Global Health*, 2018, 8(2): 020503.
- [10] LEMP M A, CREWS L A, BRON A J, *et al.* Distribution of aqueous-deficient and evaporative dry eye in a clinic-based patient cohort: a retrospective study [J]. *Cornea*, 2012, 31(5): 472-478.

- [11] 洪晶. 我国睑板腺功能障碍诊断与治疗专家共识 (2017 年) [J]. 中华眼科杂志, 2017, 53 (9): 657-661.
HONG J. Expert consensus on the diagnosis and treatment of meibomian gland dysfunction in China (2017) [J]. *Chinese Journal of Ophthalmology*, 2017, 53(9): 657-661. (in Chinese)
- [12] 骆仲舟, 邓宇晴, 王耿媛, 等. 基于红外成像原理的睑板腺图像量化分析系统 [J]. 眼科学报, 2021, 36(1): 30-37.
LUO Z Z, DENG Y Q, WANG G Y, *et al.* Meibomian gland image quantitative analysis system based on infrared imaging principle [J]. *Yan Ke Xue Bao*, 2021, 36(1): 30-37. (in Chinese)
- [13] 董芮岚, 周义霖, 原昊, 等. “精于诊断, 准其治疗”: 精准医疗背景下北京大学第三医院干眼特色诊疗平台建设的探索与分析 [J]. 眼科学报, 2021, 36(4): 306-318.
DONG R L, ZHOU Y L, YUAN H, *et al.* “Be exact in diagnosis and accurate in treatment”: Exploration and analysis of construction of characteristic diagnosis and treatment platform for dry eye in Peking University Third Hospital under the background of precision medicine [J]. *YanKe XueBao*, 2021, 36(4): 306-318. (in Chinese)
- [14] 柳长安, 冯雪菱, 孙长浩, 等. 基于改进麻雀算法的最大 2 维熵分割方法 [J]. 激光技术, 2022, 46 (2): 274-282.
LIU C A, FENG X L, SUN C H, *et al.* Maximum 2-D entropy image segmentation method based on improved sparrow algorithm [J]. *Laser Technology*, 2022, 46(2): 274-282. (in Chinese)
- [15] 秦传波, 宋子玉, 曾军英, 等. 联合多尺度和注意力-残差的深度监督乳腺癌分割 [J]. 光学精密工程, 2021, 29(4): 877-895.
QIN C B, SONG Z Y, ZENG J Y, *et al.* Deeply supervised breast cancer segmentation combined with multi-scale and attention-residuals [J]. *Opt. Precision Eng.*, 2021, 29(4): 877-895. (in Chinese)
- [16] 李大湘, 杨福杰, 刘颖, 等. 融入交叉注意力编码的皮肤病变分割网络 [J]. 光学精密工程, 2024, 32(4): 609-621.
LI D X, YANG F J, LIU Y, *et al.* Skin lesion segmentation network with cross-attention coding [J]. *Opt. Precision Eng.*, 2024, 32(4): 609-621. (in Chinese)
- [17] 谢斌, 刘阳倩, 李俞霖. 结合极化自注意力和 Transformer 的结直肠息肉分割方法 [J]. 光电工程, 2024, 51(10): 97-112.
XIE B, LIU Y Q, LI Y L. Colorectal polyp segmentation method combining polarized self-attention and Transformer [J]. *Opto-Electronic Engineering*, 2024, 51(10): 97-112. (in Chinese)
- [18] 雷雨婷, 张东, 杨双. 基于逻辑校准的多分类残差网络的肺分割算法 [J]. 半导体光电, 2021, 42 (4): 585-589, 595.
LEI Y T, ZHANG D, YANG S. An algorithm of lung segmentation based on logit adjustment in multi-class residual network [J]. *Semiconductor Optoelectronics*, 2021, 42(4): 585-589, 595. (in Chinese)
- [19] 陈明惠, 王腾, 袁媛, 等. 引入双编码器模型的 OCT 视网膜图像分割 [J]. 光电工程, 2023, 50 (10): 32-40.
CHEN M H, WANG T, YUAN Y, *et al.* Study on retinal OCT segmentation with dual-encoder [J]. *Opto-Electronic Engineering*, 2023, 50 (10): 32-40. (in Chinese)
- [20] 张贺童, 姚康, 丁上上, 等. 深度学习红外成像睑板腺分割处理系统 [J]. 中国医疗器械杂志, 2022, 46(4): 377-381, 427.
ZHANG H T, YAO K, DING S S, *et al.* Infra-red imaging meibomian gland segmentation system based on deep learning [J]. *Chinese Journal of Medical Instrumentation*, 2022, 46(4): 377-381, 427. (in Chinese)
- [21] URBIETA I, MUJICA A, PIÉROLA G, *et al.* WebLabel: OpenLABEL-compliant multi-sensor labelling [J]. *Multimedia Tools and Applications*, 2024, 83(9): 26505-26524.
- [22] 林嘉雯, 林智明, 赖泰辰, 等. 基于深度学习的睑板腺腺体分割方法研究 [J]. 国际眼科杂志, 2022, 22(7): 1191-1194.
LIN J W, LIN Z M, LAI T C, *et al.* Segmentation of meibomian glands based on deep learning [J]. *International Eye Science*, 2022, 22(7): 1191-1194. (in Chinese)
- [23] ZHANG Z H, LIN X L, YU X X, *et al.* Meibomian gland density: an effective evaluation index of

- meibomian gland dysfunction based on deep learning and transfer learning [J]. *Journal of Clinical Medicine*, 2022, 11(9): 2396.
- [24] 吴玉超,林岚,王婧璇,等. 基于卷积神经网络的语义分割在医学图像中的应用[J]. 生物医学工程学杂志, 2020, 37(3): 533-540.
- WU Y C, LIN L, WANG J X, *et al.* Application of semantic segmentation based on convolutional neural network in medical images [J]. *Journal of Biomedical Engineering*, 2020, 37(3): 533-540. (in Chinese)
- [25] 梁礼明,周琬颂,冯骏,等. 基于高分辨率复合网络的皮肤病变分割[J]. 光学精密工程, 2022, 30(16): 2021-2038.
- LIANG L M, ZHOU L S, FENG J, *et al.* Skin lesion segmentation based on high-resolution composite network[J]. *Opt. Precision Eng.*, 2022, 30(16): 2021-2038. (in Chinese)
- [26] 周腊珍,陈红池,李秋霞,等. 基于 Transformer 深度学习模型在医学图像分割中的研究进展[J]. 中国生物医学工程学报, 2024, 43(4): 467-476.
- ZHOU L Z, CHEN H C, LI Q X, *et al.* Research progress on transformer-based deep learning models for medical image segmentation [J]. *Chinese Journal of Biomedical Engineering*, 2024, 43(4): 467-476. (in Chinese)
- [27] JIA F, WONG W H, ZENG T Y. DDUNet: dense dense U-Net with applications in image denoising[C]. 2021 *IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*. October 11-17, 2021, Montreal, BC, Canada. IEEE, 2021: 354-364.
- [28] ZHOU Z W, RAHMAN SIDDIQUEE M M, TAJBAKHS N, *et al.* UNet++: A nested U-Net Architecture for Medical Image Segmentation [M]. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Cham: Springer International Publishing, 2018: 3-11.
- [29] CHEN J N, MEI J R, LI X H, *et al.* TransU-Net: Rethinking the U-Net architecture design for medical image segmentation through the lens of transformers [J]. *Medical Image Analysis*, 2024, 97: 103280.
- [30] CHEN Y F, ZOU B F, GUO Z X, *et al.* SCU-Net: swin-UNet and CNN bottleneck hybrid architecture with multi-fusion dense skip connection for pulmonary embolism CT image segmentation[C]. 2024 *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. January 3-8, 2024, Waikoloa, HI, USA. IEEE, 2024: 7744-7752.
- [31] 李雪涛,王耀雄,高放. 图像修复方法综述[J]. 激光与光电子学进展, 2023, 60(2): 29-44.
- LI X T, WANG Y X, GAO F. Review of image inpainting methods [J]. *Laser & Optoelectronics Progress*, 2023, 60(2): 29-44. (in Chinese)
- [32] HAN K, WANG Y, CHEN H, *et al.* A survey on vision transformer [J]. *IEEE Trans Pattern Anal Mach Intell*, 2023, 45(1): 87-110.
- [33] HE H B, GARCIA E A. Learning from imbalanced data[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2009, 21(9): 1263-1284.
- [34] DICE L R. Measures of the amount of ecologic association between species[J]. *Ecology*, 1945, 26(3): 297-302.
- [35] SAHA R K, MAHMUD CHOWDHURY A M, NA K S, *et al.* Automated quantification of meibomian gland dropout in infrared meibography using deep learning[J]. *The Ocular Surface*, 2022, 26: 283-294.
- [36] LIU Y T, ZHU H J, LIU M T, *et al.* Rolling-unet: revitalizing MLP's ability to efficiently extract long-distance dependencies for medical image segmentation [J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2024, 38(4): 3819-3827.
- [37] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional Networks for Biomedical Image Segmentation[M]. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015*. Cham: Springer International Publishing, 2015: 234-241.
- [38] CHEN L C, ZHU Y K, PAPANDREOU G, *et al.* Encoder-decoder with atrous separable convolution for semantic image segmentation[C]. *Computer Vision-ECCV 2018*. Cham: Springer International Publishing, 2018: 833-851.
- [39] ARITA R, ITOH K, INOUE K, *et al.* Noncontact infrared meibography to document age-related

changes of the meibomian glands in a normal population [J]. *Ophthalmology*, 2008, 115 (5) : 911-915.

[40] SMOLA A J, SCHÖLKOPF B. A tutorial on support vector regression [J]. *Statistics and Computing*, 2004, 14(3): 199-222.

作者简介:



杨 松(1999—),男,湖北孝感人,硕士研究生,2022年于徐州医科大学获得学士学位,主要从事医学图像处理、医学辅助诊断等方面的研究。E-mail: 15549516759@163.com

通讯作者:



夏振平(1985—),男,江苏兴化人,博士,副教授,硕士生导师,2014年于东南大学获得博士学位,主要从事视觉相关的医学图像处理及辅助诊断、显示图像质量测评和优化方面的研究。E-mail: xzp@usts.edu.cn