

基于DSE-Net的甲状腺相关眼病患眼区域轻量化分割算法

陈家毓¹, 何宏¹, 朱海鹏¹, 宋雪霏²

(1. 上海理工大学健康科学与工程学院, 上海 200093; 2. 上海交通大学医学院附属第九人民医院眼科, 上海 200011)

摘要: 临床活动性评分(Clinical activity score, CAS)是临床诊断甲状腺相关眼病(Thyroid associated ophthalmopathy, TAO)的重要评估方法之一。由于TAO症状的多样性和非患病区域的影响,人工诊断TAO容易受医生的主观经验影响。精准获取TAO患者脸部关键区域是早期诊断TAO的重要前提之一。因此,本文提出了一种基于DSE-Net的TAO患病区域自动分割的轻量化算法。DSE-Net采用U-Net作为主干模型,设计的密集型挤压-激励(Dense squeeze-and-excitation, DSE)通道注意力模块逐层提取编码结构的低级特征并融合解码结构的高级特征,进一步增强模型的特征提取能力。在巩膜、眼睑和泪阜数据集上的测试证明了DSE-Net的有效性,其中Dice系数分别达到了84.8%、84.7%和92.7%,IoU分别达到了74.0%、74.7%和86.5%。同时经过大量的对比实验证明了DSE-Net的优越性。提出的模型具有参数少、结构简单和特征提取能力强等特点,为TAO的早期诊断和预后治疗提供了重要信息。

关键词: 甲状腺相关眼病;DSE-Net;通道注意力;图像分割;轻量化模型

中图分类号: R771.3;R581;TP391.4 **文献标志码:** A

Lightweight Segmentation Algorithm for TAO Diseased Areas Based on DSE-Net

CHEN Jiayu¹, HE Hong¹, ZHU Haipeng¹, SONG Xuefei²

(1. School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China;
2. Department of Ophthalmology, Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200011, China)

Abstract: The clinical activity score (CAS) is one of the important assessment methods for clinical diagnosis of thyroid associated ophthalmopathy (TAO) disease. Manual diagnosis of TAO is susceptible to the subjective experience of ophthalmologists due to the diversity of TAO symptoms and the influence of non-diseased areas. The accurate acquisition of key facial areas of TAO patients is one of the significant prerequisites for early diagnosis of TAO. Therefore, this paper proposes a lightweight algorithm for automatic segmentation of TAO diseased areas based on DSE-Net. The DSE-Net adopts U-Net as the backbone model, and the dense squeeze-and-excitation (DSE) channel attention module, which is designed to extract low-level features of the encoding structure layer by layer and fuse high-level features of the decoding structure layer, further enhances the feature extraction capability of the model. Tests on the sclera, eyelid, and lacrimal caruncle datasets demonstrate the effectiveness of DSE-Net, with Dice

基金项目: 国家科技部项目(G2021013008);上海市科学技术委员会项目(18070503000);上海理工大学医工交叉重点项目(1020308405,1022308502)。

收稿日期: 2022-09-28; **修订日期:** 2023-02-27

coefficients reaching 84.8%, 84.7%, and 92.7%, and IoUs reaching 74.0%, 74.7%, and 86.5%, respectively. The superiority of DSE-Net is also proved by a large number of comparative experiments. The proposed model has fewer parameters, simple structure and strong feature extraction ability, providing significant information for the early diagnosis and prognosis treatment of TAO.

Key words: thyroid associated ophthalmopathy(TAO); DSE-Net; channel attention; image segmentation; lightweight model

引 言

甲状腺相关眼病(Thyroid associated ophthalmopathy, TAO)是一种累及眼框内和框周组织的严重致盲性眼眶病,也是成人最常见的眼眶病,属于自身免疫性疾病^[1]。TAO严重损害患者视功能,80%的患者眼球突出、眼睑退缩、眼部畸形,50%的患者出现斜视、复视、视力下降等视功能障碍,更有5%的患者致盲^[2]。TAO患者的日常活动、心理健康和视力综合评分低于白内障、青光眼和视网膜病变等其他眼病。甲状腺相关眼病活动度分期对患者的治疗时机选择和预后估计有重要意义,欧洲格里夫斯眼病协会(European group on graves' orbitopathy, EUGOGO)^[3]推荐临床活动性评分(Clinical activity score, CAS)对患者活动性评价。CAS根据是否出现眼睑充血、结膜充血、泪阜或皱襞肿胀等7个症状进行评价,评分达到3分及以上定义为活动性TAO,在此期间使用糖皮质激素治疗可以有效减轻软组织水肿、眼球突出等症状。

临床上,眼科专家无需借助专业仪器即可为患者眼部相关部位进行检查,评价是否有相关症状及症状的严重程度,给出CAS值。这项工作首先准确定位患者眼部的巩膜(位于球结膜上层)、眼睑、泪阜等区域,再进一步判断症状,整个过程耗时耗力,且要求医生具备丰富经验。设计机器学习方法自动精准分割患者面部图像中眼睑、巩膜、泪阜等部位,分割结果可以有效辅助医生查看TAO患者病情、提高医生的诊断效率,也为自动判断病症的识别算法模型提供精准的感兴趣区域。模型拥有高分割性能的同时应尽可能降低复杂度、快速分割,符合实际应用需求,因此面向TAO患者患病区域的轻量型分割算法对辅助临床医生诊断具有重要意义。

过去几年,传统图像分割算法在辅助医学影像分析任务取得了令人较为满意的结果,但传统分割算法严重依靠先验知识,准确率较低。传统分割方法主要有阈值法、区域生长法、边缘检测法、聚类法、图分割和活动轮廓模型等方法^[4-11]。阈值分割是传统分割模型最基础的方法,Qian等^[5]提出一种采用级联Haar特征和Otsu动态阈值分割方法提高现有瞳孔检测算法的性能;Dong等^[7]基于改进的OTSU阈值分割方法实现了巩膜分割;梁挺^[4]为了解决眼睑闭合不全程度诊断中准确率低和诊断效率低的问题,提出瞳孔中心定位法和改进的区域生长法,对眼睑闭合不全程度的诊断在准确率上有较大提升;Chatterjee等^[9]使用Kirsch滤波器、Canny、Prewitt、Sobel和Fuzzy-C等基于边缘的方法分割视网膜血管;Jin等^[6]基于K-Means聚类和改进的Hough变换实现了虹膜分割;Banerjee等^[10]使用测地线活动轮廓估计瞳孔和虹膜的外部轮廓,并基于GrabCut图分割的椭圆拟合算法实现了对虹膜区域的准确分割;Das等^[8]提出了基于主动轮廓的区域生长分割方法分割巩膜,以虹膜中心为参考点分别在两边选取种子点;Zhao等^[11]将区域生长法和基于区域的水平集方法提取结果结合,准确分割视网膜血管。传统方法多为无监督学习,依靠人工设计特征提取、设定分类策略实现分割图像,适合小型数据集,但面对复杂问题建立模型难度大,且模型泛化性有限。

基于深度学习的分割方法主要为端到端的监督学习,根据含有标签的数据集自主提取特征。Naqvi等^[14]使用SegNet作为主干网络,融合残差跳跃连接传输高频信息,对巩膜和虹膜进行分割;Chaud-

hary等^[15]结合U-Net和DenseNet深度神经网络,并提出边界感知损失函数和损失调度策略训练深度语义分割模型,分割巩膜、虹膜和瞳孔等区域,实现了实时视线追踪;Radu等^[16]将颜色、形状和边缘相关的三种类型特征结合,输入多分类器结构再经过具有一层隐藏层的前馈神经网络对巩膜进行分割;Lucio等^[17]使用生成对抗网络和全卷积神经网络在MICHE-I和UBIRIS.v2两个数据集测试了巩膜分割效果。上述方法都是面向正常人巩膜、虹膜和瞳孔的分割,未应用于TAO患病区域巩膜、眼睑和泪阜的分割。Zhu等^[18]提出SV-net模型,在疑似TAO患者的眼眶CT图像中精准分割全部眼外肌和视神经协助对TAO患者的临床决策。当前其他关于TAO患者的图像分析大多基于MRI和CT图像^[19]。TAO眼病患者的巩膜、眼睑和泪阜较正常人会出现不同严重程度的充血、肿胀,呈多样性与复杂性。

因此,本文通过设计密集型挤压-激励(Dense squeeze-and-excitation, DSE)通道注意力模块,将多个卷积块和挤压-激励模块通过密集连接结合,增强有效特征信息的同时避免输入信息的丢失,进一步提升特征通道方向的敏感性,并基于此提出了DSE-Net,对TAO患病区域轻量型分割,克服患者面部图像复杂、多样的形态对分割模型的影响。最后,在上海交通大学医学院附属第九人民医院眼科TAO患者真实面部图像数据进行实验,实现TAO患者巩膜、眼睑和泪阜区域精准分割,从而验证本文方法的有效性 with 准确性。

1 本文方法

1.1 模型框架

TAO患病区域特征如图1所示,其中图1(a)为患病区域细节,红色和蓝色区域分别为内、外眼角放大图,绿色区域为巩膜,黄色区域为泪阜,橙色区域为眼睑。巩膜与外眼角皮肤交界区域纹理相近,边界模糊不易界定;巩膜形态有正常的乳白色、遍布血丝的充血状和完全充血肿胀,样本间差异巨大,如图1(b)所示。泪阜区域受采集环境的影响易有反光点;样本的泪阜大小、形态、颜色饱和度各异,部分泪阜隐藏在内眼角,面积小不易捕捉,如图1(c)所示。眼睑指上眼皮隆起区域,位于眉毛与眼睛之间,左右两侧与面部皮肤没有清晰、明确的边缘特征。为了应对上述复杂情况,实现精准、快速的分割,本文设计了DSE通道注意力机制模块,融合编码器-解码器结构,逐层对编码器低级特征不同通道分配权重并融合解码器的高级特征,提升模型特征提取能力并控制模型参数量。

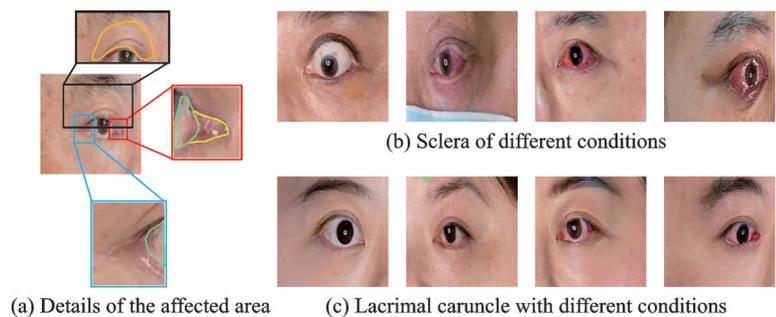


图1 TAO患者患病区域特征

Fig.1 Characteristics of diseased regions in TAO patients

网络模型结构如图2所示,主要分为编码器、解码器、通道注意力模块3部分。编码器经过4次下采样提取带有语义信息的特征图,解码器则通过4次上采样将提取的高维特征图经反卷积逐步还原为原始尺寸的分割结果图像,通道注意力模块连接编码器和解码器提升模型分割性能。模型输入 $256 \times 256 \times 3$ 的眼部感兴趣区域彩色图像,经过两个 3×3 卷积块(Convolution)提取图像特征,通道数增加至32;再使用 2×2 最大池化操作(Max pooling)进行下采样,将 $W \times H \times C$ 尺寸的特征图汇聚为 $W/2 \times H/2 \times C$ 输入下一层编码器,压缩了特征图尺度并保留有效信息,一定程度上避免过拟合。经过4次下采样,编码器不断压缩特征图尺寸、增加特征图通道,最终提取了 $16 \times 16 \times 512$ 的特征图。

4层编码器通过密集型挤压-激励通道注意力模块(DSE module)筛选有用的通道信息,与解码器上层输出特征使用捷径连接(Shortcut connection)叠加融合,共同输入对应层的解码器,恢复空间信息,补充边缘等高频信息,提升分割效果。解码器与编码器结构对应,首先经过两个 3×3 的卷积块解码,再对特征图使用 3×3 卷积核反卷积(Transposed convolution), $W/2 \times H/2 \times C/2$ 的特征图上采样至 $W \times H \times C$ 。在四层解码器计算后,特征图尺寸还原为 $256 \times 256 \times 32$ 。最后通过 3×3 卷积块压缩通道数,得到 $256 \times 256 \times 1$ 尺寸的分割结果图。

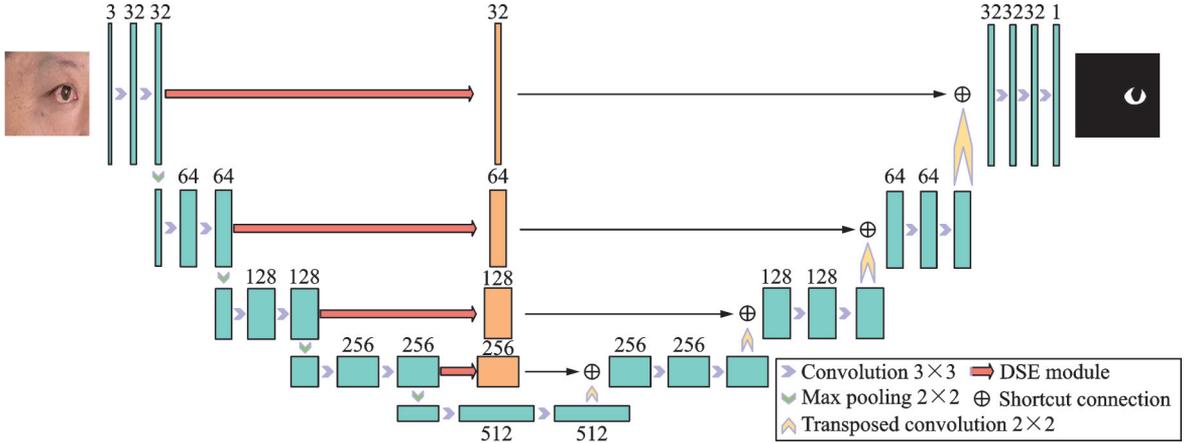


图2 网络模型结构

Fig.2 Network model structure

1.2 DSE 模块

注意力机制受人类视觉注意力的启发,在视觉任务中为特征不同部分分配权重,突出特征图中的重要信息,削弱低价值信息,以此提升模型性能。Hu等^[20]提出了挤压-激励网络(Squeeze-and-excitation net, SE-Net)通道注意力机制模型,首先通过全局池化将特征图压缩为,进行全连接,得到 $1 \times 1 \times C$ 维的向量,其中 r 为通道压缩比例,经实验取16,再连接ReLU激活,经过全连接将 C/r 维向量还原为 C 维向量,然后经过Sigmoid激活函数得到数值区间在 $[0, 1]$ 的通道权重向量。最后将权重向量沿通道赋予原特征图,实现通道注意力机制。通道权重向量 s 为

$$s = F(z, W) = \sigma(g(z, W)) = \sigma(\text{FC}_2 \delta(\text{FC}_1 z)) \quad (1)$$

式中: FC_1 为输出降维至 C/r 个通道的全连接层; δ 为ReLU激活函数; FC_2 为输出为 C 个通道的全连接层; σ 为Sigmoid函数。此模块结构简单,但丢失了部分输入信息,特征提取的能力有限。

本文提出了密集型挤压-激励(DSE)模块,模型结构如图3所示。SE子模块通过全局平均池化对输入张量 X 从通道维度特征压缩,获得特征图 $z \in \mathbf{R}^{1 \times 1 \times C}$ 。 z 相当具备了全域感受野,表示沿特征通道方向的全局分布。再学习通道间的非线性关系,通过 W 参数生成每个特征通道的权重,生成的权重向量为

$$s = F_{\text{ex}}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)) \quad (2)$$

式中: W_1 为输出压缩至 C/r 个通道的 1×1 卷积层,取16; W_2 为输出还原为 C 个通道的 1×1 卷积层。最后式(2)提取的权重向量与沿特征图 X 对应的通道相乘,对原始特征图进行缩放,得到输出特征图 $\tilde{X} \in \mathbf{R}^{H \times W \times C}$ 为 $\tilde{X} = s \cdot X$ 。

DSE模块的输入为 $X_{\text{in}} \in \mathbf{R}^{H \times W \times C}$,先通过两个 3×3 的卷积块进一步提取特征,传入SE子模块,

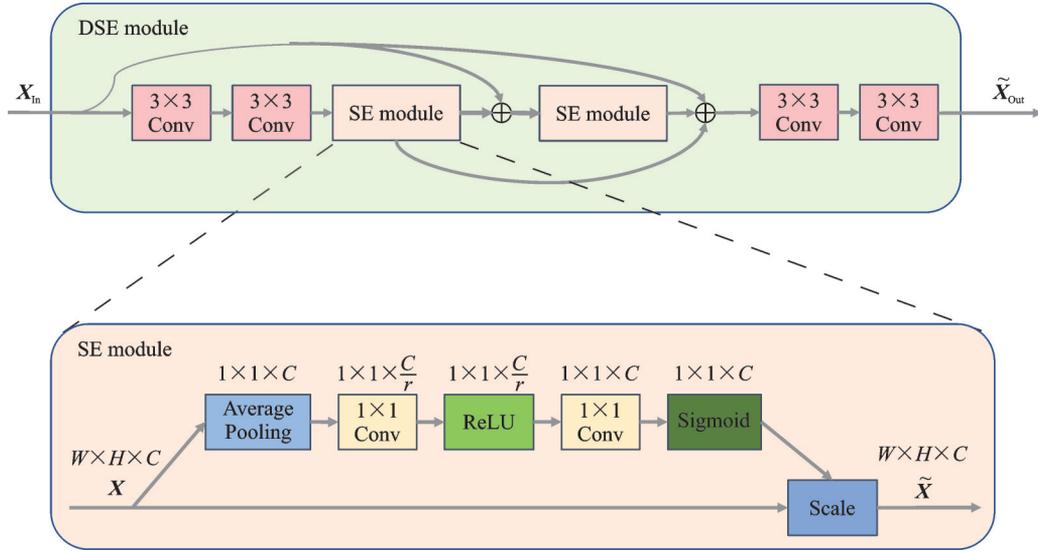


图3 DSE 模块结构

Fig.3 DSE module structure

获取初估计的通道注意力特征图 \tilde{X}_{SE1}

$$\tilde{X}_{SE1} = SE_1(\text{Conv}_2(\text{Conv}_1(X_{in}))) \tag{3}$$

式中 Conv 为 3×3 卷积层。再将 \tilde{X}_{SE1} 与 X_{in} 融合输入 SE 子模块进一步提升通道注意力的估计能力,得到精估计的特征图 \tilde{X}_{SE2} 为

$$\tilde{X}_{SE2} = SE_2(\tilde{X}_{SE1} + X_{in}) \tag{4}$$

为了保留输入特征图 X_{in} 的原始信息并放大重要通道的信息权重,合并特征图 \tilde{X}_{SE2} 、 \tilde{X}_{SE1} 与 X_{in} ,再传入两个 3×3 的卷积块整合包含通道注意力信息的特征图,最终提取到 $\tilde{X}_{Out} \in \mathbf{R}^{H \times W \times C}$ 为

$$\tilde{X}_{Out} = \text{Conv}_4(\text{Conv}_3(\tilde{X}_{SE2} + \tilde{X}_{SE1} + X_{in})) \tag{5}$$

本文提出的 DSE 模块密集连接多个特征图、结合多个卷积块,强调特征有效信息又避免了输入信息的丢失,提升了特征通道方向的敏感性,从而增强了模型的特征提取能力,提高了模型分割性能;且模块保持了相对简单的结构,参数量增加较少。

1.3 训练设置

网络采用二分类交叉熵损失函数与 Dice 损失函数的和作为总损失函数进行优化。预测结果的精确度越高,损失值就越低,较好地衡量了两个概率分布之间的差异性。二分类交叉熵损失函数为

$$\text{BCELoss}(x_n, y_n) = -w_n [y_n \times \lg x_n + (1 - y_n) \times \lg(1 - x_n)] \tag{6}$$

式中 x_n 为预测图像,范围为 $0 \sim 1$ 间; y_n 为真实标签,值为 0 或 1; w_n 为损失值缩放因子,用于调整样本间的权重,本文取 1。

Dice 系数是一种集合相似度度量函数,通常用于表示两个样本的相似度

$$\text{Dice Coefficient} = \frac{2|X \cap Y|}{|X| + |Y|} \tag{7}$$

式中 $|X \cap Y|$ 表示 X 与 Y 两个集合的交集; $|X|$ 、 $|Y|$ 表示元素个数。由于需要不断优化损失值减小, Dice 损失函数为

$$\text{Dice Loss} = 1 - \frac{2|X \cap Y| + \text{smooth}}{|X| + |Y| + \text{smooth}} \quad (8)$$

式中 smooth 为极小的常数取 1×10^{-6} , 防止分母为 0。

模型采用 RMSProp 优化器, 其中衰减速率为 0.9, 保持数值稳定的小常数为 1×10^{-8} , 初始学习率为 1×10^{-4} 。防止模型过拟合, 每当验证集 Dice 系数停止提升两次后学习率将减少 1×10^{-8} 。训练过程中批处理大小 (Batch size) 设置为 32, 共训练 200 轮。

2 实验评估与分析

2.1 实验环境与数据集

2.1.1 实验环境

本文实验的硬件平台为 CPU: Intel(R) Core(TM) i7-9700 CPU @ 3.00 GHz, 内存: 16 GB, 软件环境: PyTorch 1.11、OpenCV 4.5.5。

2.1.2 数据集

本文数据全部采集自 TAO 患者, 采集设备为 Bellus3D Face Camera Pro 相机, 由上海交通大学医学院附属第九人民医院眼科提供。数据经预处理裁剪为 $256 \text{像素} \times 256 \text{像素}$ 大小的图像样本, 每只眼睛为一个样本, 各分割任务样本如图 4 所示。经眼科医生指导对眼睑、巩膜、泪阜样本分别进行手动标注, 其中泪阜过小, 标签还包括临近泪阜的一部分巩膜区域。各分割任务样本集数量如表 1 所示, 所有样本皆为 TAO 患者。深度学习模型的训练集与验证集比重为 9:1, 在训练模型时随机划分。

2.2 方法比较

为了验证本模型所具有的优势, 对比了 4 种应用于眼部图像传统分割方法, 包括阈值分割^[7]、K-means 聚类^[6]、水平集^[11]和 GrabCut^[10], 同时引入在自然图像、医学图像等领域常用的深度学习模型, 比较了 U-Net^[21]、SE-UNet^[20]、SegNet^[22]、DeepLab v3^[23]、DeepLab v3+^[24]、TransUNet^[25]、TransFuse^[26] 七种模型。

2.3 评价指标

实验采用交并比 (Intersection over union, IoU)、Dice 系数、精确率 (Precision)、召回率 (Recall) 和像素准确度 (Pixel accuracy, PA) 五个评价指标对分割效果进行评价, 如表 2 所示, 其中 i, j 分别表示目标类别与背景, p_{ii} 表示正确分割的像素数量即真正, p_{ij}, p_{ji} 分别表示假正与假负。IoU 与 Dice 系数是主要的分割效果评价。

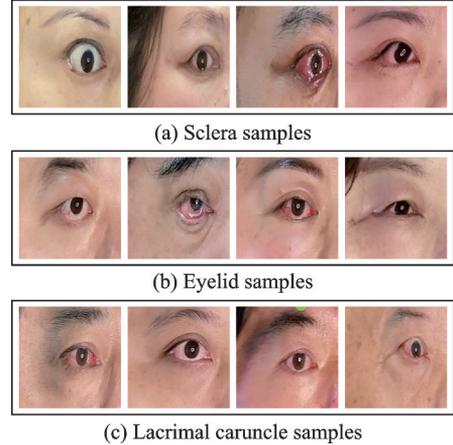


图 4 数据集样本示例

Fig.4 Examples of dataset samples

表 1 各分割任务样本集数量

Table 1 Number of sample sets for each segmentation task

任务集合	巩膜	眼睑	泪阜
训练集	538	611	590
验证集	59	67	65
测试集	60	60	60

表 2 评价指标

Table 2 Evaluation metrics

指标	公式
交并比	$\text{IoU} = \frac{p_{ii}}{p_{ij} + p_{ji} + p_{ii}}$
Dice 系数	$\text{Dice} = \frac{2 \cdot p_{ii}}{p_{ij} + p_{ji} + 2 \cdot p_{ii}}$
精确率	$\text{Precision} = \frac{p_{ii}}{p_{ji} + p_{ii}}$
召回率	$\text{Recall} = \frac{p_{ii}}{p_{ij} + p_{ii}}$
像素准确度	$\text{PA} = \frac{p_{ii} + p_{jj}}{p_{ii} + p_{ij} + p_{ji} + p_{ii}}$

2.4 对比实验

2.4.1 分割结果比较

本文模型在巩膜分割的训练损失曲线如图5所示,结果表明DSE-Net可以有效收敛。在巩膜分割与眼睑分割两个任务中,本文方法对比了4种传统分割方法和7种深度学习方法;由于传统方法的局限,泪阜分割任务比较了7种深度学习方法。各方法在3个数据集上的分割效果评价指标如表3所示。巩膜和眼睑分割测试集的分割结果表明,手动提取特征的4种传统方法分割效果与深度学习方法差距大。在3个分割任务中,DSE-Net模型IoU和Dice系数均为最佳。巩膜分割任务中,本文方法IoU值达到了74.0%,Dice系数为84.8%,较DeepLab v3+分别提高了

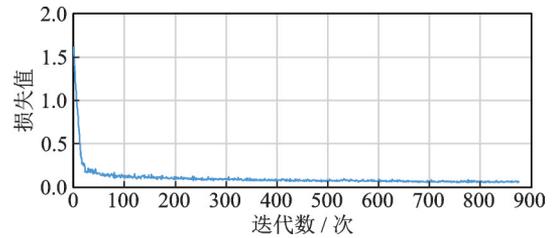


图5 DSE-Net巩膜分割任务损失收敛曲线

Fig.5 Loss convergence curve of DSE-Net for scleral segmentation task

表3 各模型分割结果

Table 3 Segmentation results of each model

数据集	方法	IoU	Dice	Precision	Recall	Accuracy
巩膜	阈值分割	34.2	49.7	62.4	43.2	98.6
	K-means聚类	36.8	52.7	77.0	41.8	98.1
	水平集	25.1	38.9	73.7	28.1	98.0
	Grab Cut	44.2	60.0	59.7	67.0	98.4
	U-Net	71.7	83.0	95.0	74.5	99.2
	SE-UNet	73.4	84.3	93.9	77.2	99.3
	SegNet	71.8	83.1	92.7	76.4	99.2
	DeepLab v3	71.3	82.9	91.1	76.8	99.2
	DeepLab v3+	73.7	84.5	93.1	77.9	99.3
	TransUNet	72.7	83.9	92.4	77.4	99.3
	TransFuse	61.3	74.4	94.8	63.5	99.0
DSE-Net	74.0	84.8	93.6	78.0	99.3	
眼睑	阈值分割	50.3	66.6	61.8	75.6	91.9
	K-means聚类	50.2	66.6	61.4	76.1	92.9
	水平集	40.8	57.0	62.7	60.0	92.0
	Grab Cut	46.7	63.3	59.1	71.8	92.9
	U-Net	74.6	84.7	92.2	80.1	96.9
	SE-UNet	73.7	83.9	90.8	80.2	96.7
	SegNet	73.7	84.0	90.0	80.8	96.7
	DeepLab v3	74.0	84.5	90.4	81.2	96.8
	DeepLab v3+	73.3	83.7	88.7	81.1	96.6
	TransUNet	73.5	84.1	88.2	82.3	96.7
	TransFuse	74.5	84.5	89.8	81.5	96.7
DSE-Net	74.7	84.7	89.4	82.4	96.8	
泪阜	U-Net	86.3	92.5	92.8	92.7	99.8
	SE-UNet	85.0	91.7	91.2	92.8	99.8
	SegNet	83.6	90.8	90.7	91.7	99.8
	DeepLab v3	83.7	91.0	90.0	92.5	99.8
	DeepLab v3+	83.8	90.9	91.7	91.2	99.8
	TransUNet	84.0	91.1	89.6	93.3	99.8
	TransFuse	80.0	87.1	90.6	86.6	99.8
	DSE-Net	86.5	92.7	92.1	93.6	99.8

0.3%。眼睑分割任务中,本文方法 IoU 取得 74.7%,较 U-Net 提升 0.1%;Dice 系数与 U-Net 相当。泪阜分割任务中,本文方法 IoU 达到了 86.5%,Dice 系数达到了 92.7%,较 U-Net 均提升 0.2%。结果表明 DSE-Net 性能优秀且稳定。

DSE-Net 在 3 个分割任务中召回率均最高,由表 2 中的召回率公式可知同一任务 ($p_{ij} + p_{ii}$) 不变,表明 DSE-Net 模型正确分割像素数量 p_{ii} 最大。但 DSE-Net 精确率指标均未取得最佳,不及 U-Net。由表 2 中的精确率公式可得 DSE-Net 模型全部预测数量 ($p_{ji} + p_{ii}$) 偏大,表明 DSE-Net 分割结果面积广于 U-Net,有相对轻微的分割现象。由于背景区域面积过大,各模型像素准确度表现相近。

各算法的指标均值如图 6 所示,深度学习方法之间差距较小,整体优于传统方法,DSE-Net 的 IoU、Dice 系数、召回率和像素准确度 4 项指标均值皆为最佳,精准率均值达到了 91.7%。

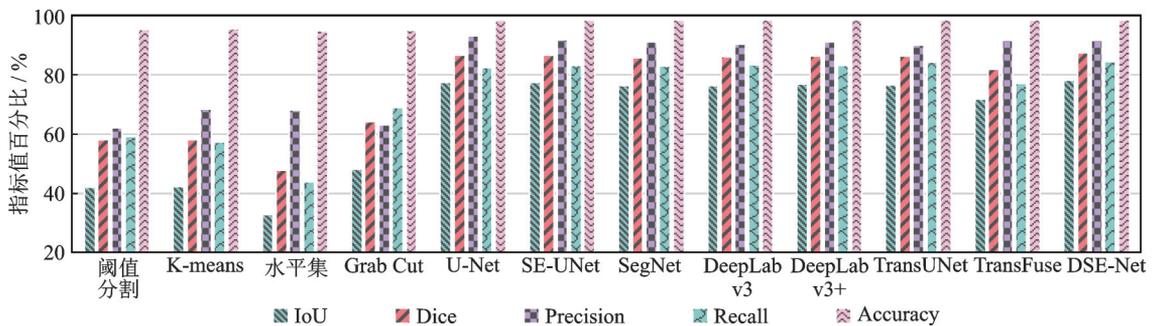


图 6 算法指标均值

Fig.6 Mean of algorithmic metrics

本文针对在各数据集选取两个样本进行可视化分析,样本输入如图 7(a)所示。泪阜分割结果如图 7(b)所示,SE-UNet 错误分割了样本 1,DeepLab v3、DeepLab v3+、TransUNet 和 TransFuse 对样本 1 存在欠分割,DSE-Net 较好地分割出目标区域。巩膜分割如图 7(c)所示,各深度学习模型面对上眼睑退缩样本 1 存在不同程度的欠分割,DSE-Net 相比其他分割方法更完整。眼睑分割如图 7(d)所示,DSE-Net 相比其他深度学习方法分割的边缘更加圆滑、自然,但各方法在眼睑红肿的样本 2 均有欠分割。

各实验结果表明,以 DSE 模块为核心的 DSE-Net 融合了编码器和解码器的特征,通过提升特征通道方向的敏感性进一步增强了模型的特征提取能力,模型提高了分割性能且表现稳定。

2.4.2 模型复杂度比较

为了验证模型的复杂度,测试了各算法在同一硬件上的运行时间,分割 60 张测试集图像求得平均每张图像消耗的时间,如表 4 所示。传统算法中阈值分割运算速度极快,K-means 聚类其次,而水平集和 Grab Cut 的计算时间远远高于其他方法,尤其水平集需要耗费数 10 s。深度学习方法各算法计算时间接近,DSE-Net 虽然推理速度次于 TransFuse,但分割性能较其更佳。且 DSE-Net 模型参数量仅为 11.02×10^6 ,为 SE-UNet 参数量的 53.3%,比 U-Net 参数量减少 20.02×10^6 ,如图 8 所示,远小于其他深度学习模型,属于轻量级网络。

2.5 消融实验

为了验证本文提出的 DSE 模块的有效性,在编码器-解码器结构上分别应用 SE 模块、包含一个 SE 模块的 DSE(DSE-1)和连接两个 SE 模块的 DSE(DSE-2),设置训练 50 轮,测试 3 个数据集的分割性能,取均值比较模型综合表现,结果如表 5 所示。各模块均能提升网络的分割性能,SE 通道注意力模块使网络的 IoU 提升了 1.2%、Dice 系数提高了 0.8%。DSE-1 模块相比 SE 模块在 IoU 和 Dice 系数上提升

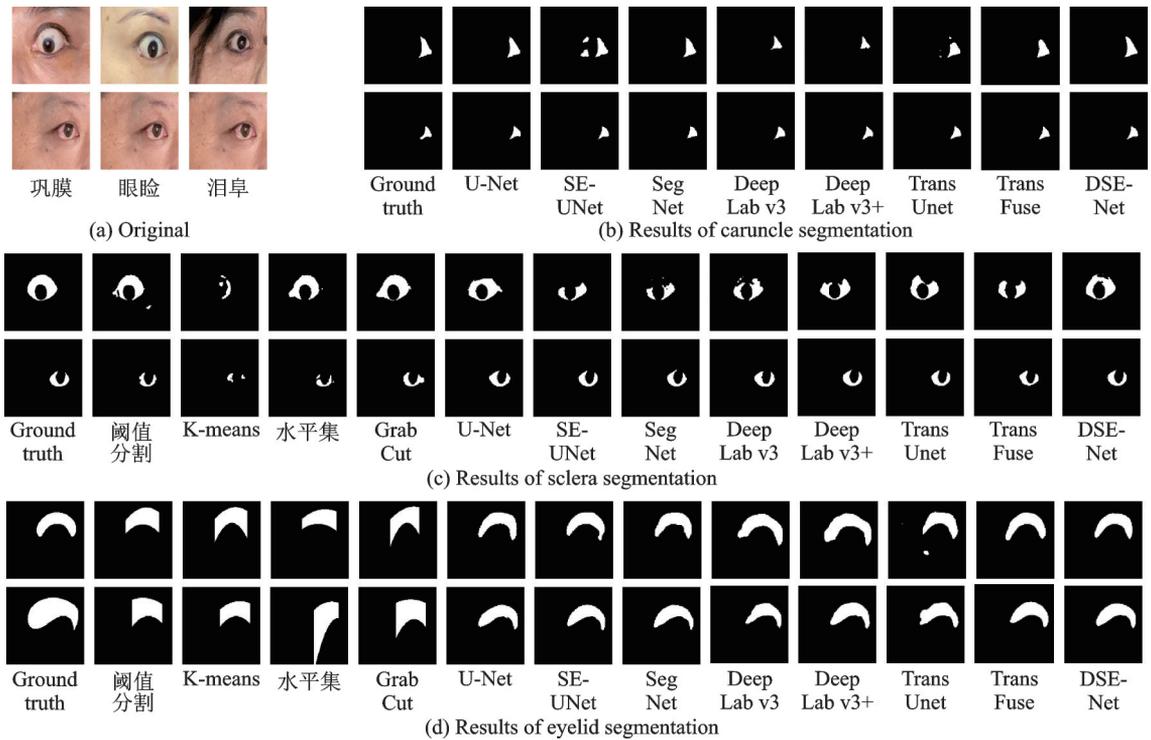


图7 分割结果可视化

Fig.7 Visualization of segmentation results

表4 各分割算法运行时间

Table 4 Running time of each segmentation algorithm

方法	阈值分割	K-means	水平集	Grab Cut	U-Net	SE-UNet	SegNet	DeepLab v3	DeepLab v3+	TransUNet	TransFuse	DSE-Net
巩膜	0.027	0.180	50.328	1.045	0.403	0.321	0.409	0.341	0.353	0.321	0.175	0.276
眼睑	0.062	0.243	17.143	1.700	0.400	0.313	0.331	0.340	0.389	0.348	0.169	0.273
泪阜	—	—	—	—	0.345	0.277	0.316	0.380	0.354	0.323	0.156	0.267

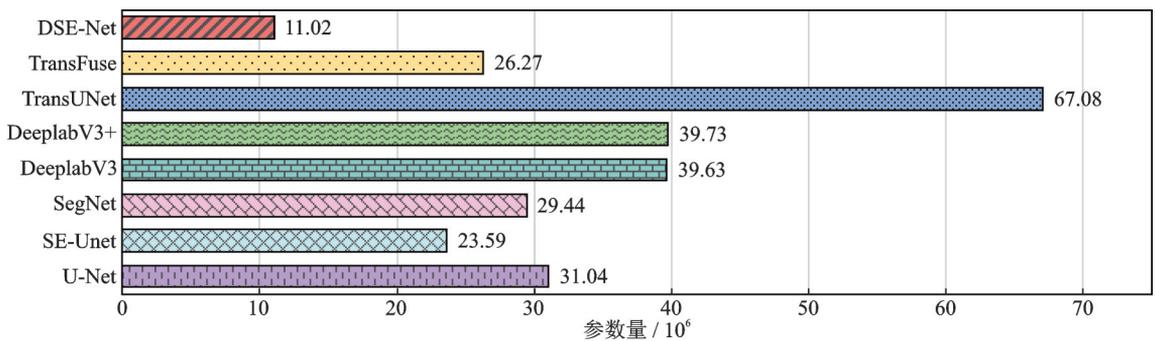


图8 各模型参数量

Fig.8 Number of parameters for each model

微小,而包含2个SE模块的DSE-2更充分地强调了重要通道的信息,相比基础网络模型IoU提升了1.8%、Dice系数提升了1.4%,精确率下降了0.9%,但召回率提高了2.6%。

表5 消融实验分割性能

Table 5 Segmentation performance of ablation experiment

%

SE	DSE-1	DSE-2	IoU	Dice	Precision	Recall	Accuracy
			76.8	86.2	93.0	81.9	98.6
			78.0	87.0	92.9	83.2	98.6
			78.1	87.2	91.9	84.1	98.6
			78.6	87.6	92.1	84.5	98.6

3 结束语

本文提出了一种面向TAO患者眼部区域分割的轻量分割算法。该算法在经典的编码器-解码器结构上引入DSE通道注意力模块,逐层提取编码结构的低级特征直接与解码结构的高级特征连接,增强了模型对特征的提取能力。基于医院门诊采集的TAO患者面部数据制作了巩膜、眼睑和泪阜3个数据集,经过实验验证了本文提出方法的有效性,分割结果Dice系数在3个分割任务上分别达到84.8%、84.7%和92.7%,在与多种主流模型对比中表现出稳定的性能,IoU、Dice和召回率均能达到最佳分割效果,同时参数量远远低于主流深度学习分割模型。DSE-Net可以精准分割不同病情的TAO患者巩膜、眼睑和泪阜区域,有较好的实时性,模型能为TAO疾病的早期诊断和预后治疗提供重要帮助。目前模型只能精准分割单个区域,无法实现一个模型对巩膜、眼睑、泪阜3个区域的同时分割,在实际部署算法时造成冗余,所以在未来的工作中将进一步优化模型实现多类别精准分割。

参考文献:

- [1] ŞAHLİE, GÜNDÜZ K. Thyroid-associated ophthalmopathy[J]. Turkish Journal of Ophthalmology, 2017, 47 (2): 94-105.
- [2] BAHN R S. Graves' ophthalmopathy[J]. New England Journal of Medicine, 2010, 362 (8): 726-738.
- [3] BARTALENA L, BALDESCHI L, BOBORIDIS K, et al. The 2016 European thyroid association/European group on graves' orbitopathy guidelines for the management of graves' orbitopathy[J]. European Thyroid Journal, 2016, 5 (1): 9-26.
- [4] 梁挺. 基于甲状腺相关眼病图像的虹膜定位和疾病分级方法研究[D]. 重庆:重庆邮电大学, 2021.
LIANG Ting. Research on iris location and disease classification method based on images of thyroid-related eye diseases[D]. Chongqing: Chongqing University of Posts and Telecommunications, 2021.
- [5] QIAN F, LEI H. A pupil location algorithm based on cascaded Haar feature and OTSU dynamic threshold segmentation[C]// Proceedings of Applied Optics and Photonics China 2021. USA: SPIE, 2021: 796-801.
- [6] JIN L, XIAO F, HAOPENG W. Iris image segmentation based on K-means cluster[C]// Proceedings of 2010 IEEE International Conference on Intelligent Computing and Intelligent Systems. [S.l.]: IEEE, 2010: 194-198.
- [7] DONG W, ZHOU H, XU D. A new sclera segmentation and vessels extraction method for sclera recognition[C]// Proceedings of 2018 10th International Conference on Communication Software and Networks (ICCSN). [S.l.]: [s.n.], 2018: 552-556.
- [8] DAS A, PAL U, BALLESTER M A F, et al. Sclera recognition using dense-SIFT[C]// Proceedings of 2013 13th International Conference on Intelligent Systems Design and Applications. [S.l.]: [s.n.], 2013: 74-79.
- [9] CHATTERJEE S, SUMAN A, GAURAV R, et al. Retinal blood vessel segmentation using edge detection method[C]// Proceedings of Journal of Physics: Conference Series, 2021: 012008.
- [10] BANERJEE S, MERY D. IRIS segmentation using geodesic active contours and grabcut[C]// Proceedings of Image and Video Technology-PSIVT 2015 Workshops. Auckland, New Zealand: [s.n.], 2016: 48-60.
- [11] ZHAO Y Q, WANG X H, WANG X F, et al. Retinal vessels segmentation based on level set and region growing[J]. Pattern Recognition, 2014, 47 (7): 2437-2446.
- [12] ALKASSAR S, WOO W L, DLAY S S, et al. Robust sclera recognition system with novel sclera segmentation and validation techniques[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2015, 47 (3): 474-486.
- [13] ALKASSAR S, WOO W L, DLAY S S, et al. Enhanced segmentation and complex-sclera features for human recognition

- with unconstrained visible-wavelength imaging[C]//Proceedings of 2016 International Conference on Biometrics (ICB). [S.l.]: [s.n.], 2016: 1-8.
- [14] NAQVI R A, LEE S W, LOH W K. Ocular-net: Lite-residual encoder decoder network for accurate ocular regions segmentation in various sensor images[C]//Proceedings of 2020 IEEE International Conference on Big Data and Smart Computing (BigComp). [S.l.]: IEEE, 2020: 121-124.
- [15] CHAUDHARY A K, KOTHARI R, ACHARYA M, et al. RITnet: Real-time semantic segmentation of the eye for gaze tracking[C]//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). [S.l.]: IEEE, 2019: 3698-3702.
- [16] RADU P, FERRYMAN J, WILD P. A robust sclera segmentation algorithm[C]//Proceedings of 2015 IEEE 7th International Conference on Biometrics Theory, Applications and Systems (BTAS). [S.l.]: IEEE, 2015: 1-6.
- [17] LUCIO D R, LAROCA R, SEVERO E, et al. Fully convolutional networks and generative adversarial networks applied to sclera segmentation[C]//Proceedings of 2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS). [S.l.]: IEEE, 2018: 1-7.
- [18] ZHU F, GAO Z, ZHAO C, et al. Semantic segmentation using deep learning to extract total extraocular muscles and optic nerve from orbital computed tomography images[J].Optik, 2021, 244: 167551.
- [19] JI Y, CHEN N, LIU S, et al. Research progress of artificial intelligence image analysis in systemic disease-related ophthalmopathy[J].Disease Markers, 2022, 2022: 3406890.
- [20] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2018: 7132-7141.
- [21] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]// Proceedings of Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015: 18th International Conference. Munich, Germany: [s.n.], 2015: 234-241.
- [22] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation[J].IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39 (12): 2481-2495.
- [23] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation[EB/OL]. (2017-06-17). <https://doi.org/10.48550/arXiv.1706.05587>.
- [24] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Proceedings of the European Conference on Computer Vision (ECCV). [S.l.]: [s.n.], 2018: 801-818.
- [25] CHEN J, LU Y, YU Q, et al. TransUNet: Transformers make strong encoders for medical image segmentation[EB/OL]. (2021-02-08). <https://doi.org/10.48550/arXiv.2102.04306>.
- [26] ZHANG Y, LIU H, HU Q. TransFuse: Fusing transformers and CNNs for medical image segmentation[C]//Proceedings of Medical Image Computing and Computer Assisted Intervention - MICCAI 2021. Strasbourg, France: [s.n.], 2021: 14-24.

作者简介:



陈家毓(1997-),男,硕士研究生,研究方向:医学图像处理,E-mail:202562383@st.usst.edu.cn。



何宏(1973-),通信作者,女,教授,研究方向:医学人工智能、人机交互系统、医疗大数据分析,E-mail:hehong@usst.edu.cn。



朱海鹏(1993-),男,博士研究生,研究方向:人工智能、医学图像、模式识别。



宋雪霏(1985-),男,助理研究员,研究方向:眼科学、数字医学、临床研究。