

江玉彬, 张强. 基于深度学习的角膜溃疡分割方法研究[J]. 智能计算机与应用, 2025, 15(3): 145-151. DOI: 10.20169/j.issn.2095-2163.250321

基于深度学习的角膜溃疡分割方法研究

江玉彬, 张强

(河北工程大学 信息与电气工程学院, 河北 邯郸 056038)

摘要: 角膜溃疡是角膜疾病中最常见的症状, 可能会造成不可逆的视力损伤或失明。不同角膜溃疡的病理形态差异较大, 并且缺乏足够的具有真实标签的裂隙灯图像, 提出了一种半监督学习与改进的 Swin-UNet 结合的网络模型用于角膜溃疡病灶区域的自动语义分割。首先, 在 Swin-UNet 的网络模型的瓶颈引入 ResNeXt50 的第 5 层网络结构, 并使用 CNN 与 Transformer 交叉教学的半监督学习方法训练改进的网络模型。将方法与之前的分割方法进行比较, 实验结果表明, 所提方法可以进一步提高视角角膜溃疡的整体语义分割精度。

关键词: 角膜溃疡; Swin-UNet; 语义分割; 半监督学习; 聚合残差网络

中图分类号: TP 391.41

文献标志码: A

文章编号: 2095-2163(2025)03-0145-07

Research on corneal ulcer segmentation method based on deep learning

JIANG Yubin, ZHANG Qiang

(School of Information and Electrical Engineering, Hebei University of Engineering, Handan 056038, Hebei, China)

Abstract: Corneal ulcers are the most common symptom of corneal disease, which can cause irreversible visual impairment or blindness. The pathological morphology of different corneal ulcers is quite different, and there is a lack of sufficient slit lamp images with real labels, and a network model combining semi-supervised learning with improved Swin-UNet is proposed for automatic semantic segmentation of corneal ulcer lesions. Firstly, the fifth-layer network structure of ResNeXt50 is introduced in the bottleneck of the Swin-UNet network model, and the improved network model is trained using the semi-supervised learning method of cross-teaching CNN and Transformer. Comparing the method with the previous segmentation method, the experimental results show that the proposed method can further improve the overall semantic segmentation accuracy of the viewing angle membrane ulcer.

Key words: corneal ulcers; Swin-UNet; semantic segmentation; semi-supervised learning; aggregate residual networks

0 引言

角膜溃疡是角膜疾病中最常见的症状, 是世界范围内眼部发病率的重要原因之一, 涉及到角膜基质上皮层的破坏。太晚或不适当的治疗可能对视力造成不可逆的损害。在许多情况下, 由于疾病的进程往往缓慢和不受关注, 视力障碍的识别就容易发生推迟。在传统医学诊断中, 专家的判断几乎是一个决定性的考量因素。即使在深度学习和人工智能快速发展的今天, 医学诊断问题上, 依然严重依赖于医生的判断。然而, 早期干预可以极大地增加愈合几率, 降低恶化程度, 这使得具有成本效益的公共卫生预防方案对工业化和发展中经济体非常重要。显

而易见, 预防有助于有效治疗角膜致盲, 虽然有经验的眼科医生可以通过一些专业软件手动标记溃疡区域, 但这项任务既耗时又主观。为了量化角膜溃疡的严重程度, 需要对其进行分割, 尤其是片状溃疡、角膜溃疡的病症表象并不相同。溃疡区域具有不规则性, 并不仅限于某一处, 尤其对形状多维不规则的区域的估算则存在难度。因此, 探索一种能够自动、准确分割角膜溃疡区域的方法具有重要意义。

准确分割溃疡区域对于评估角膜溃疡的严重程度和制定治疗方案至关重要。3 种类型角膜溃疡的比较如图 1 所示。由图 1 可知, 根据病理特征和分布, 角膜溃疡可分为上面一排的点状角膜溃疡、中间一排的点片状混合角膜溃疡和下面一排的片状角膜

作者简介: 江玉彬(1998—), 女, 硕士研究生, 主要研究方向: 计算机视觉, 图像处理。Email: jiangyubin2020@163.com; 张强(1977—), 男, 副教授, 主要研究方向: 人工智能。

收稿日期: 2023-09-27

溃疡。

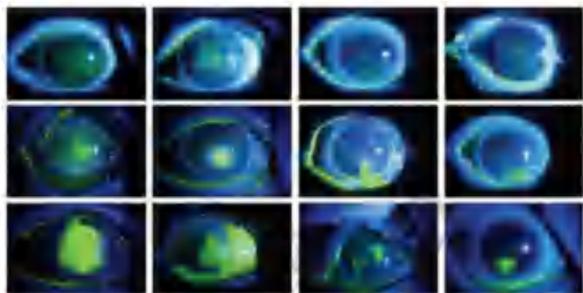


图1 3种类型角膜溃疡的比较

Fig. 1 Comparison of three types of corneal ulcers

医学图像分割旨在从医学影像中准确识别病变器官的像素点,以提取这些病变区域的信息特征,这是医学图像分析领域中一项具有挑战性的任务^[1]。随着时间的推移,医学图像分割方法逐渐演化出多种不同的算法。传统的医学图像分割方法包括基于灰度阈值的分割算法、基于边缘检测的分割算法、基于区域的分水岭分割算法,以及融合特定理论的分割算法等^[2]。

2015年,Ronneberger等学者^[3]提出U-Net网络架构,该网络结构由编码阶段和解码阶段组成。在编码过程中,下采样图像提取图片特征;解码过程中,对图片进行上采样,以便逐步恢复图片的大小。2018年,Xiao等学者^[4]提出Res-UNet,将UNet的每一个子模块分别替换为具有残差连接的形式,解决了视网膜血管的有效分割。2021年,Chen等学者^[5]提出了TransUNet,同时具有Transformers和U-Net的优点,是医学图像分割的强大替代方案。2017年,Sun等学者^[6]采用了基于补丁的方法,该方法基于深度卷积神经网络。2021年,Wang等学者^[7]提出了一种角膜溃疡分割网络CU-SegNet来分割荧光素染色图像中的角膜溃疡。2022年,Wang等学者^[8]提出了一种新型的半监督多尺度自转换生成对抗网络Semi-MsST-GAN,可以利用未标记的图像来提高裂隙灯图像荧光素染色中角膜溃疡分割的性能。

目前,虽然Transformer模型具有良好的全局捕获能力,但在深度学习中进行图像分割和病灶检测时,只考虑了图像的全局特征表示,局部细节学习表示仍有待加强。本次研究的主要贡献总结如下:

(1) 通过引入Swin-UNet模型,并使用ResNext50的第5层网络作为Swin-UNet模型的瓶颈,增加了子特征图的提取量,提高了病灶区域分割的准确性。

(2) 为了进一步提高无标记样本的分割性能,采用CNN和Transformer之间的交叉教学作为角膜溃疡图像分割的半监督学习框架。

(3) 基于SUSTech-SYSU数据集的综合实验证明了本文提出的方法的有效性。结果表明,与其他最先进的算法相比,本文提出的方法不仅获得了更高的分割精度,而且可以利用无标记数据进一步提高分割性能。

1 网络结构

1.1 Swin-UNet 网络结构

Swin-UNet是第一个纯粹的基于Transformer的U型架构,由编码器、瓶颈、解码器和跳跃连接组成^[9]。编码器、瓶颈和解码器都是基于Swin-Transformer Block构建的。

使用带有偏移窗口的分层Swin Transformer作为编码器来提取上下文特征,并设计了具有补丁扩展层的基于对称Swin Transformer的解码器来执行上采样操作,通过跳跃连接与编码器的多尺度特征融合,恢复特征图的空间分辨率,接下去则进行分割预测。

一个Swin Transformer block由一个W-MSA和一个SW-MSA组成,如图2所示。

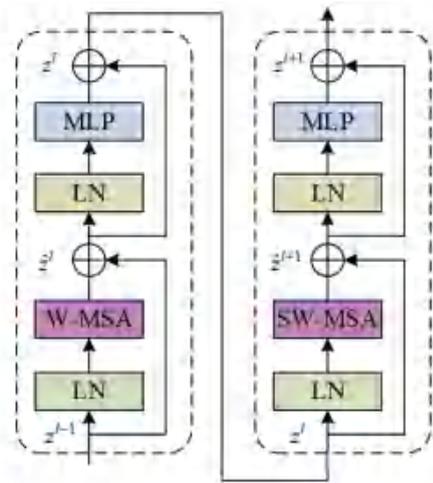


图2 2个连续的Swin Transformer块

Fig. 2 Two consecutive Swin Transformer blocks

2个连续Swin Transformer块可表示为:

$$\hat{z}^l = W - MSA(LN(z^{l-1})) + z^{l-1} \quad (1)$$

$$z^l = MLP(LN(\hat{z}^l)) + \hat{z}^l \quad (2)$$

$$\hat{z}^{l+1} = SW - MSA(LN(z^l)) + z^l \quad (3)$$

$$z^{l+1} = MLP(LN(\hat{z}^{l+1})) + \hat{z}^{l+1} \quad (4)$$

在 \hat{z}^l 和 z^l 表示(S)W-MSA 模块和 l 的 MLP 模块的输出与 Vision Transformer 相似,自我注意的计算方法如下:

$$Attention(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d}} + B\right)V \quad (5)$$

其中, Q, K, V 分别表示查询向量矩阵、键向量矩阵、值向量矩阵,这 3 个矩阵中每一行分别表示一个对应的向量; B 表示相对位置偏移。

1.2 ResNeXt 的网络层

Swin-UNet 在提取上下文信息和恢复空间分辨率方面更加有效,但在深度瓶颈中,由于 Transformer 太深,Transformer 模块在图像特征计算的收敛上无法获得很好的效果。为了更有效地提取角膜溃疡区特征,提高整体分割效果,对 Swin-UNet 中的瓶颈进行了改进。由于残差网络中残差块的设计不会导致特征提取的计算能力随着网络的深化而下降,可以在瓶颈处实现子特征图的分辨率和特征维数,所以用 ResNeXt 残差块来替换 Swin-UNet 瓶颈位置是合适的。ResNext50 的第 5 层网络结构如图 3(a) 所示。ResNext50 由多个残差模块组成,解决了网络层次深化导致的网络退化问题,实现了数千层的网络

参数计算。

传统的方法中要提高模型的准确率,都是加深或加宽网络,但是随着超参数数量的增加,网络设计的难度和计算开销也会加大。因此 ResNeXt 结构可以在不增加参数复杂度的前提下提高准确率,同时还减少了超参数的数量^[10]。

ResNeXt 同时采用 VGG 堆叠的思想和 Inception 的 split-transform-merge 思想。VGG 主要采用堆叠网络来实现,之前的 ResNet 也借用了这样的思想。此后提到的 Inception 系列网络,简单讲就是 split-transform-merge 的策略,即 Inception 是将多个卷积或池化操作放在一起组装成一个网络模块,设计神经网络时,以模块为单位去组装整个网络结构。但是 Inception 系列网络的问题是:网络的超参数设定的针对性较强,当应用在别的数据集上时需要修改许多参数,因此可扩展性一般。

ResNeXt 关键点在于基数 (*cardinality*), 参见图 3(b) 中 *cardinality* = 32 的设计,这里每个被聚合的拓扑结构都是一样的,增加 *cardinality* 比增加深度和宽度要更加有效。

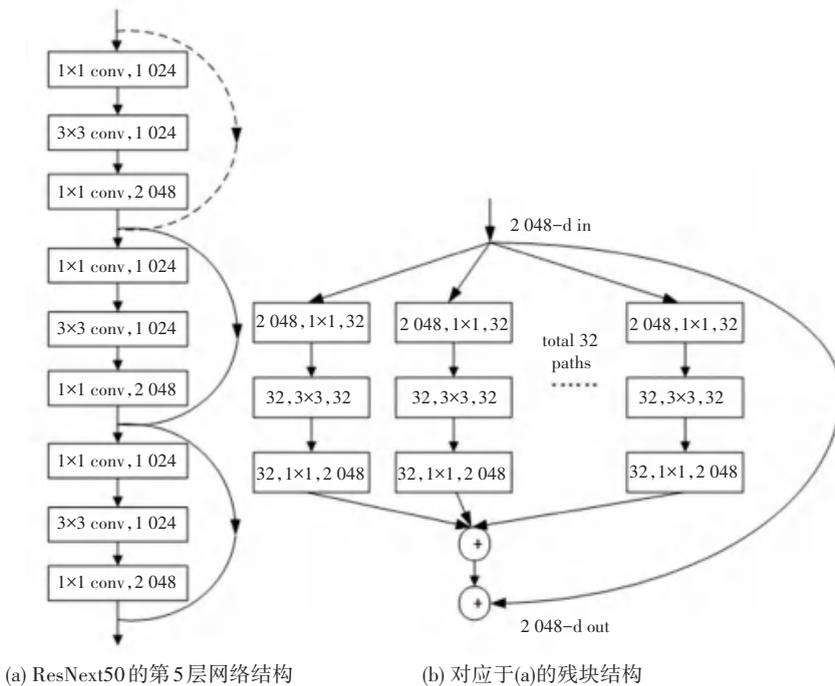


图 3 ResNeXt 的结构

Fig. 3 Structure of ResNeXt

1.3 改进的 Swin-UNet 模型

本文将 Swin-UNet 网络瓶颈部分由原来的 2 个连续 Swin Transformer 块替换为 ResNext50 的第 5 层

网络结构。图 4 显示了本文提出的改进的 Swin-UNet 网络模型。



图4 改进的 Swin-UNet 网络模型

Fig. 4 Improved Swin-UNet network model

在瓶颈处,考虑到 ResNeXt 中的残差网络块不会随着层的加深而退化的特点,本文利用 ResNeXt50 中的第 5 层网络结构,弥补了 Transformer 模块在深度网络计算中无法收敛的不足。因此,对角膜图像的病变区域进行深度特征学习和提取。在深度瓶颈处使用卷积模块可以增加标记数据上的特征学习,其中特征尺寸保持 $H/32 \times W/32 \times 8C$ 不变。

1.4 半监督学习

SSL 是监督学习和无监督学习的结合,通过使用一小部分标记示例和大量未标记数据,模型可从

中学习并对新示例进行预测。基本过程涉及使用现有的标记数据来标记剩余的未标记数据,从而有效地帮助增加训练数据。

交叉教学将经典的深度协同训练从一致性正则化加以简化,其中一个网络的预测被用作伪标签,直接端到端监督另一个网络。交叉教学是隐式一致性正则化,与显式一致性正则化相比,可以产生更稳定、更准确的伪标签。显式一致性正则化强制最小化不同网络预测的差异,同时对其进行优化,这可能导致不同网络的预测相同但预测错误。

采用 CNN 和 Transformer 之间的交叉教学^[11]为

半监督角膜溃疡图像分割提供一个非常简单而高效的框架, CNN 与 Transformer 交叉教学如图 5 所示。如前所述, CNN 和 Transformer 是视觉识别的不同学习范式, 其中 CNN 依赖于局部卷积运算, 而 Transformer 基于远程自关注, CNN 和 Transformer 在

学习范式上存在差异, 并且这些属性可以在训练期间相互补充, 因此交叉教学可以有助于同时学习具有这 2 种属性的统一分割器。交叉教学的灵感来自于 3 个成果: 深度协同训练^[12]、协同教学^[13] 和交叉伪监督^[14]。

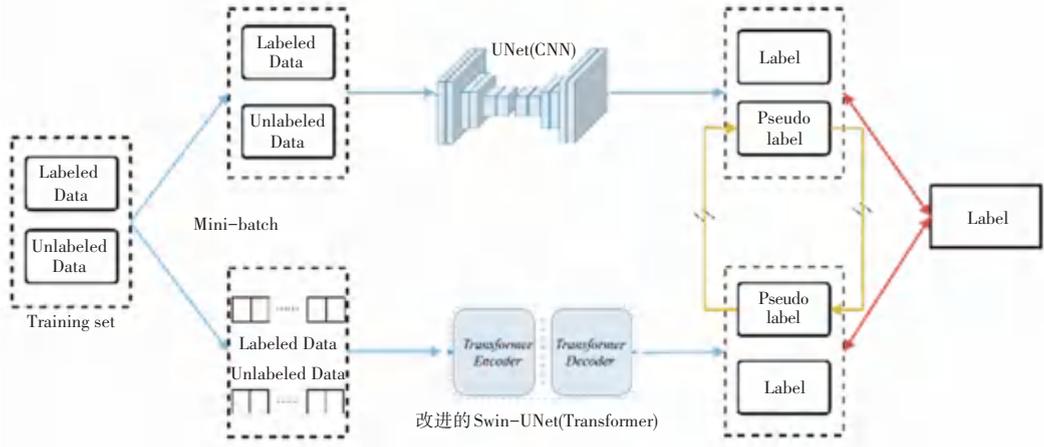


图 5 Cross Teaching between CNN and Transformer 结构图

Fig. 5 Structure diagram of Cross Teaching between CNN and Transformer

对于监督学习使用的损失函数为有标签预测数据的 *Dice* 损失和交叉熵损失的综合, 表达式如下:

$$Loss_{sup} = Loss_{ce} + Loss_{dice} \quad (6)$$

交叉教学策略的伪标签通过函数 *argmax* 生成, p_{ic} 、 p_{it} 分别为 CNN 网络和 Transformer 网络训练生成的伪标签, p_{ic} 和 p_{it} 分别表示 2 个网络在同一无标签数据上的预测结果。交叉教学策略的伪标签是这样生成的:

$$p_{ic} = \text{argmax}(p_{it}); p_{it} = \text{argmax}(p_{ic}) \quad (7)$$

交叉教学损失是一个双向损失函数。一个流是从 CNN 到 Transformer, 另一个流是从 Transformer 到 CNN。未标记数据的交叉教学损失定义为:

$$Loss_{cross} = Loss_{dice}(p_{ic}, p_{ic}) + Loss_{dice}(p_{it}, p_{it}) \quad (8)$$

整体训练目标函数是由标记数据上的监督损失和未标记数据的无监督损失组成的联合损失。总体损失定义为:

$$Loss_{total} = Loss_{sup} + \lambda Loss_{cross} \quad (9)$$

其中, λ 是一个权重因子, 通常由时间依赖性高斯预热函数定义。

本文的 CNN 网络采用 UNet 网络, Transformer 网络则变更为 Transformer 和 CNN 相结合的网络, 即采用改进的 Swin-UNet 网络。

2 实验分析

2.1 数据集

提出的方法在 SUSTech-SYSU 公共裂隙灯荧光素染色图像数据集^[15] 上进行了综合实验, 该数据集用于开发和评估角膜溃疡自动分割算法。其中有 354 张点片状混合片状角膜溃疡裂隙灯荧光素染色图像、由眼科医生标注了真实标签, 还有 358 张无标签的点状角膜溃疡图像。每张分辨率为 2 592 × 1 728 像素的 RGB 图像只包含一个角膜区域, 位于视场的中间位置, 将输入图像大小设置为 224 × 224 像素。

2.2 实验配置

在实验中, 开发环境为 Pytorch1. 10. 0, GPU 为 NVIDIA RTX 3090(24 GB 显存)。随机梯度下降优化器的脉冲为 0. 9, 权值衰减为 0. 000 1, 初始学习率为 0. 01。首先, 对图像数据进行增广, 包括从 -20° ~ 20° 的旋转、水平翻转、垂直翻转等操作, 然后选取 70% 作为训练集, 20% 作为验证集, 10% 作为测试集。采用 SGD 优化器, *epoch* 为 150, *batch size* 为 24, 每个 *epoch* 输出验证集的评价指标值。

2.3 评价指标

通过评价指标来评估训练完成的网络模型, 主要从 *Dice*、*IoU*、*Recall* 和 *Precision* 四个方面来评价

模型的优劣,计算定义如下所示:

$$Dice = \frac{2 \times TP}{2 \times TP + TN + FP} \quad (10)$$

$$IoU = \frac{TP}{TP + FN + FP} \quad (11)$$

$$Recall = \frac{TP}{TP + FN} \quad (12)$$

$$Precision = \frac{TP}{TP + FP} \quad (13)$$

其中, TN 、 TP 、 FN 和 FP 分别表示真阴性、真阳性、假阴性和假阳性。根据评价指标的反馈信息,调整模型的参数,选出效果更好的模型为后续系统的搭建做准备,评价价值越高,模型分割效果越好。

2.4 结果与讨论

不同方法的测试评价结果见表 1。由表 1 看出, Swin-UNet 比 U-Net、R2U-Net、Att-UNet、U-Net++ 效果都要好,改进的 Swin-UNet 的 $Dice$ 指数为 72.97%, 比 Swin-UNet 的 $Dice$ 指数高了 4.54%, $Recall$ 指数高了 5.58%。本文方法在这 4 个评价指标上取得了最好的分割效果, $Dice$ 高达 75.56%, $Precision$ 高达 81.65%, 说明本文方法具有良好的模型泛化性能和较强的算法鲁棒性,交叉教学的半监督学习实现了同时学习 CNN 和 Transformer, 分别专注于局部信息和全局信息。

表 1 不同方法的评价结果

Table 1 Evaluation results for different methods

方法	$Dice$	IoU	$Precision$	$Recall$
U-Net	0.668 1	0.556 4	0.759 2	0.669 9
R2U-Net	0.641 1	0.543 7	0.741 0	0.636 0
Att-UNet	0.664 3	0.552 5	0.783 2	0.653 3
U-Net++	0.646 5	0.528 4	0.769 9	0.643 2
Swin-UNet	0.684 3	0.585 4	0.758 7	0.673 1
改进的 Swin-UNet	0.729 7	0.621 8	0.798 5	0.728 9
Swin-UNet+交叉教学	0.740 4	0.632 8	0.810 5	0.739 2
改进的 Swin-UNet+交叉教学	0.755 6	0.652 4	0.816 5	0.751 7

角膜溃疡分割示例如图 6 所示。由图 6 可知,原始图像、真实标签和表 1 展示的 8 种不同方法的分割结果分别显示在 3 张图像上。从图 6 中可以看出,本文提出的方法获得了最好的分割结果。由于多尺度多区域、形状不规则、位置不固定等因素,角膜裂隙灯图像中的溃疡部分难以分割。改进的方法不仅能准确分割小尺寸的病灶,而且在分割大目标时保持良好的区域连续性。

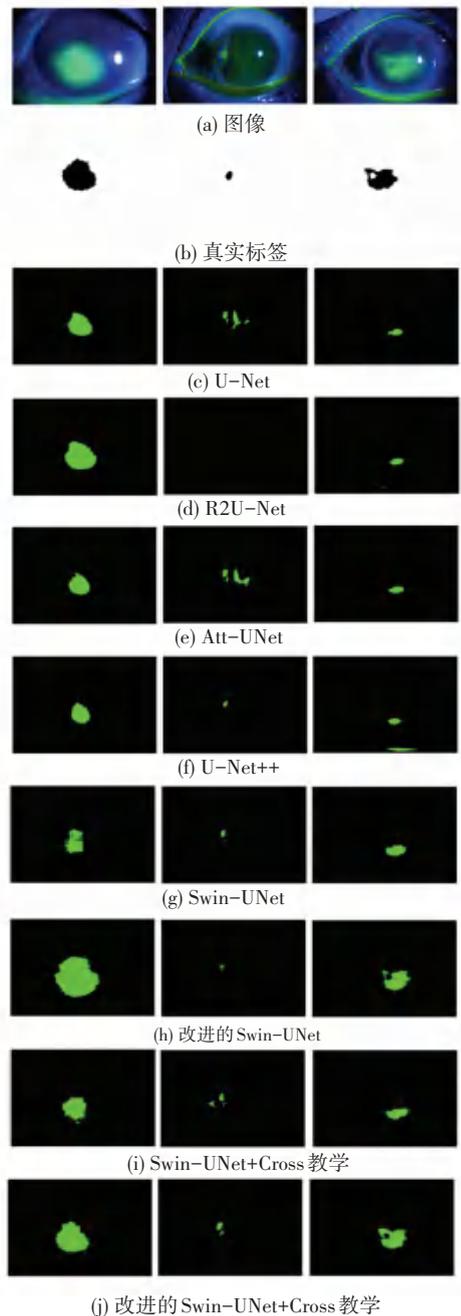


图 6 角膜溃疡分割示例

Fig. 6 Example of corneal ulcer segmentation

3 结束语

本文提出了一种改进的 Swin-UNet 网络模型,主要解决了 2 个问题:

(1) 为了提高模型的泛化能力,对数据集进行了扩充,并将 ResNext50 中的网络层作为 Swin-UNet 网络的瓶颈进行训练,提高了深度网络的计算收敛性。

(2) 选取无标签的数据,采用 CNN 和 Transformer 之间的交叉教学的半监督学习来提高网

络的分割精度,最后对结果进行预测。同时进行了综合实验,以评估所提方法的有效性和鲁棒性。实验证明,与其他最先进的算法相比,本文提出的基于半监督的方法的分割性能有了明显的提高。后续会收集更多的数据,尽量降低模型的时间复杂度,从而获取更好的分割结果。

参考文献

- [1] MOHAMMAD H H, JIA Wenjing, HE Xiangjian, et al. Deep learning techniques for medical image segmentation: Achievements and challenges [J]. *Journal of Digital Imaging*, 2019, 32(4): 582-596.
- [2] 刘宇,陈胜.医学图像分割方法综述[J]. *电子科技*, 2017, 30(8): 169-172.
- [3] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation [C]// *Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015)*. Cham: Springer, 2015: 234-241.
- [4] XIAO Xiao, LIAN Shen, LUO Zhiming, et al. Weighted RES-UNet for high-quality retina vessel segmentation [C]// *2018 9th International Conference on Information Technology in Medicine and Education (ITME)*. Piscataway, NJ: IEEE, 2018: 327-331.
- [5] CHEN Jieneng, LU Yongyi, YU Qihang, et al. Transunet: Transformers make strong encoders for medical image segmentation [J]. *arXiv preprint arXiv, 2102.04306*, 2021.
- [6] SUN Qichao, DENG Lijie, LIU Jianwei, et al. Patch-based deep convolutional neural network for corneal ulcer area segmentation [M]// *Fetal, Infant and Ophthalmic Medical Image Analysis (OMIA 2017)*. Lecture Notes in Computer Science. Cham: Springer, 2017, 10554: 101-108.
- [7] WANG Tingting, ZHU Weifang, WANG Meng, et al. Cu-Segnet: corneal ulcer segmentation network [C]// *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*. Piscataway, NJ: IEEE, 2021: 1518-1521.
- [8] WANG T, WANG M, ZHU W, et al. Semi-MsST-GAN: A semi-supervised segmentation method for corneal ulcer segmentation in slit-lamp images [J]. *Frontiers in Neuroscience*, 2022, 15: 1705.
- [9] CAO Hu, WANG Yueyue, CHEN J, et al. Swin-UNet: UNet-like pure transformer for medical image segmentation [C]// *Proceedings of the European Conference on Computer Vision*. Cham: Springer, 2022: 205-218.
- [10] XIE Saining, GIRSHICK R, DOLLÁR P, et al. Aggregated residual transformations for deep neural networks [C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Piscataway, NJ: IEEE, 2017: 1492-1500.
- [11] LUO Xiangde, HU Minhao, SONG Tao, et al. Semi-supervised medical image segmentation via cross teaching between CNN and transformer [J]. *arXiv preprint arXiv, 2112.04894*, 2021.
- [12] QIAO Siyuan, SHEN Wei, ZHANG Zhishuai, et al. Deep co-training for semi-supervised image recognition [C]// *Proceedings of the 15th European Conference on Computer Vision (ECCV 2018)*. Cham: Springer, 2018: 142-159.
- [13] COOK L, FRIEND M. Co-Teaching: Guidelines for creating effective practices [J]. *Focus on Exceptional Children*, 1995, 28(3): 1-16.
- [14] CHEN Xiaokang, YUAN Yuhui, ZENG Gang, et al. Semi-supervised semantic segmentation with cross pseudo supervision [J]. *arXiv preprint arXiv, 2106.01226*, 2021.
- [15] DENG Lijie, LYU Junyan, HUANG Haixiang, et al. The SUSTech-SYSU dataset for automatically segmenting and classifying corneal ulcers [J]. *Scientific Data*, 2020, 7(1): 23.