

采用融合 ResNet 和 Transformer 的 U-Net 进行疟疾感染红细胞分割

刘潇霜, 张伟

甘肃中医药大学信息工程学院, 甘肃 兰州 730000

【摘要】针对疟疾感染红细胞图像分割模型分割性能不高的问题,提出一种改进的 U-Net 网络模型,融合 ResNet 和 Transformer。首先编码器部分使用 ResNet,加深特征提取网络,以提取更深层次的特征;然后将 ResNet 输出传入 Transformer 模块进行目标区域特征的加强;最后通过解码器模块进行特征融合并输出结果。在疟疾显微图像数据集上,本文方法的 Dice 相似系数、平均交并比、类别平均像素准确率均优于 U-Net 网络,分别达到了 87.40%、76.85%、85.28%。本文方法可以提高疟疾感染红细胞图像的分割精度,为疟疾诊断提供更有效和准确的解决方案。

【关键词】疟疾; U-Net; Transformer; 语义分割

【中图分类号】R318; TP391

【文献标志码】A

【文章编号】1005-202X(2024)02-0191-07

Segmentation of malaria-infected erythrocytes using U-Net incorporating Transformer and ResNet

LIU Xiaoshuang, ZHANG Wei

School of Information Engineering, Gansu University of Chinese Medicine, Lanzhou 730000, China

Abstract: A novel U-Net network model which integrates ResNet and Transformer is proposed to address the problem of poor malaria-infected erythrocyte performance of the existing models. ResNet is used in the encoder to deepen the feature extraction network for extracting deeper features, and the ResNet output is inputted into Transformer module for the feature enhancement in the target area, and finally the decoder module is used to perform feature fusion and output the results. The experiment on the malaria microscopy image dataset shows that the proposed method outperforms U-Net in Dice similarity coefficient, mean intersection over union, and mean pixel accuracy, reaching 87.40%, 76.85%, and 85.28%, respectively. The proposed method can improve the accuracy of malaria-infected erythrocyte segmentation and provide a more effective and accurate solution for malaria diagnosis.

Keywords: malaria; U-Net; Transformer; semantic segmentation

前言

疟疾是一种由寄生虫引起的传染性疾病,寄生虫通过雌性按蚊叮咬传播^[1]。世界卫生组织报告称 2021 年全球有 2.47 亿病例,约有 60 万人因感染疟疾死亡,非洲是疟疾感染最严重的地区,2021 年该区域疟疾病例总数占全球的 95%^[2]。疟疾的早期诊断和

治疗可以减少疾病感染,防止死亡,并有助于减少传播^[3]。尽管有各种新技术可用于疟疾诊断,但常规的手工显微镜检查是疟疾最普遍的诊断技术^[4]。疟疾寄生虫主要影响人类血液中的红细胞,可通过查看染色薄血涂片显微图像中红细胞的颜色、形状、大小等特征来判断感染情况,确定疟疾感染阶段,即环状体、滋养体、裂殖体和配子体 4 个感染阶段^[5]。诊断结果的有效性和准确性依赖于病理学家的经验,这一检测方法需要专业人员的专业经验,费时费力,成本高^[6]。计算机辅助疟疾诊断可以为该问题提供有效的解决方案,薄血涂片显微图像中红细胞的精确分割有助于提高专家诊断疟疾的效率。

在血涂片显微图像中细胞分割主要的方法可以分为传统方法和深度学习方法。传统分割方法包含

【收稿日期】2023-10-26

【基金项目】甘肃省教育厅创新基金(2022B-113)

【作者简介】刘潇霜,硕士研究生,研究方向:深度学习、医学图像处理, E-mail: 284385351@qq.com

【通信作者】张伟,副教授,硕士生导师,研究方向:医学信号处理、计算机仿真, E-mail: 4865354@qq.com

Ostu算法、分水岭算法和标记控制的分水岭算法等。Lee等^[7]为分离重叠红细胞,使用canny边缘算法增强细胞的重叠边缘,确定重叠顺序,再结合Ostu算法实现重叠细胞的自动提取。Das等^[8]使用标记控制分水岭算法在血涂片图像上分割红细胞。Savkare等^[9]将全局阈值和Ostu阈值相加生成细胞的二值图像,再使用分水岭变换算法来分离重叠细胞。Jalari等^[10]最大化原始彩色图像的亮度分量的类间方差,得到初始阈值图像,再用迭代阈值选择算法找到最佳阈值,用于分割显微血液图像中的疟原虫。Savkare等^[11]采用K均值聚类 and 全局阈值对血细胞图像进行增强和分割,重叠部分采用Sobel边缘检测器和分水岭变换进行分离。传统图像处理方法存在一定的局限性,无法处理来自不同领域的图像固有的可变性,并且在分割任务中的泛化性较差。近年来,随着神经网络越来越多地应用于图像的分割和分类,深度学习方法成功地解决了传统图像处理方法的局限性,并得到广泛的应用^[12]。与基于传统算法的分割方式相比,基于深度学习的语义分割算法可以更准确地提取语义特征信息,实现对每个类别的准确预测^[13]。朱世祺等^[14]运用DeepLabV3+神经网络对肠息肉内镜图像进行分割,其平均交并比达81.9%。Ronneberger等^[15]提出U-Net网络,该网络由编码器和解码器组成,可以从很少的图像中进行端到端训练,在生物医学分割应用中实现非常好的性能。U-Net下采样的过程中会丢失很多特征信息,从而导致分割精度不足,因此出现了多种基于U-Net的改进算法。Li等^[16]结合神经微分方程与U-Net网络来进行血涂片图像分割,以提高图像分割的准确性。Zhang等^[17]将可变形卷积层应用于经典的U-Net结构,提出dU-Net,具有更好的分离细胞、区分背景噪声并预测正确的细胞形状的能力。Li等^[18]提出Residual-Attention UNet++网络,能以最小的性能下降为代价实现更快的医学图像分割。Liu等^[19]在U-Net结构的基础上引入残差连接,提出一种改进的Res-Unet肺动脉分割方法。Nurçin等^[20]引入CNN-Transformer混合架构,提高重叠红细胞的分离准确率,网络的Jaccard相似性指数达94.5%。胡晓阳等^[21]将多头自注意力机制引入U-Net网络的跳跃连接,提高对肝脏的分割精度,在边缘轮廓处理上有更好表现。现有的分割方法存在分割精度不足,无法分割出清晰的细胞边缘等缺点,尚存在提升空间。

本研究对U-Net结构进行改进和优化,采用ResNet作为主干网络,减少网络丢失的特征信息,同时引入Transformer模块,用于疟疾感染红细胞显微图像,以提升细胞的分割效果。本研究主要可分为

下述3点:(1)数据集预处理,在已有的显微镜成像图中筛选可靠图像,采用旋转、裁剪、镜像等方法进行扩充,并对每张图像进行尺寸统一处理。(2)网络训练,将经过预处理的数据集传入本研究提出的融合ResNet和Transformer的U-Net网络模型中进行训练。(3)分割指标评估,使用模型对测试集数据进行预测,并与Ground Truth进行多项指标的计算。评价指标包含Dice相似系数(Dice Similarity Coefficient, DSC)、平均交并比(Mean Intersection over Union, mIoU)、类别平均像素准确率(Mean Pixel Accuracy, MPA)、准确率、精确率和召回率。

1 融合ResNet和Transformer的U-Net网络

1.1 总体设计

本研究对U-Net结构进行改进,提出融合ResNet和Transformer的U-Net网络对疟疾感染红细胞进行分割。当需要增加模型的层数来增强特征提取的能力时,ResNet的引入能很好地避免因为网络层数增加导致的缺点^[22]。Transformer的提出是用于机器翻译,但是后续很多学者将其运用到图像分割领域,凭借它具有的提供自注意力信息特征的特点,Transformer的变体在医学图像分割领域取得了不错的成绩^[23]。网络的总体结构如图1所示。本研究所使用的神经网络总体采用编码器和解码器结构。编码器网络主干部分使用ResNet并且加入自带全局自注意力机制的Transformer,二者结合使用增强特征提取和特征信息在网络层间的传播能力。解码器部分采用4次上采样和特征融合,从而细化解码器对浅层细节特征的恢复。

1.2 网络结构

1.2.1 以ResNet-50为主干的编码器结构 本研究选择ResNet-50作为主干特征提取网络。ResNet的残差结构在增加网络深度同时提升性能,从而提取低级特征和空间细节特征^[24]。选择ResNet-50作为编码器能提升整体模型对特征提取的能力,并满足医学分割对边缘等细节信息的要求^[25]。

ResNet-50的结构如图2所示,总共有5层。第一层包含一个 7×7 的卷积层和最大池化操作。剩下的4层可以看作是4个不同的残差块,每个残差块中包含了多个残差单元,残差单元由3个卷积层组成,这4个残差块分别由3、4、6和3个残差单元组成,共包含48层卷积层。残差单元由两个 1×1 卷积层、一个 3×3 卷积层、BN层和ReLU激活函数构成。最后通过1个平均池化、全连接层和softmax组成的输出层,得到最后的 1×1 卷积输出。

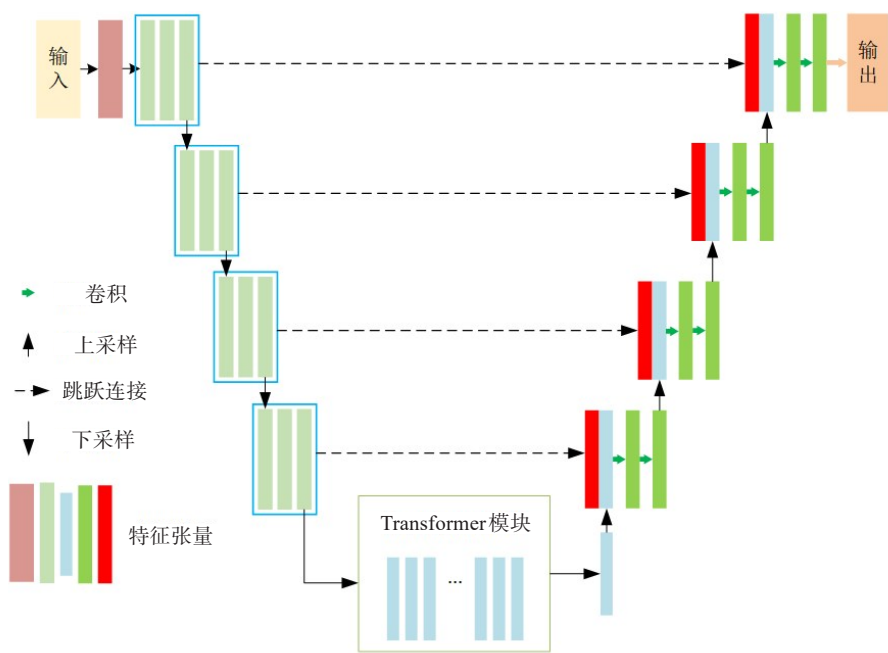


图1 本文网络结构

Figure 1 Structure of the proposed network

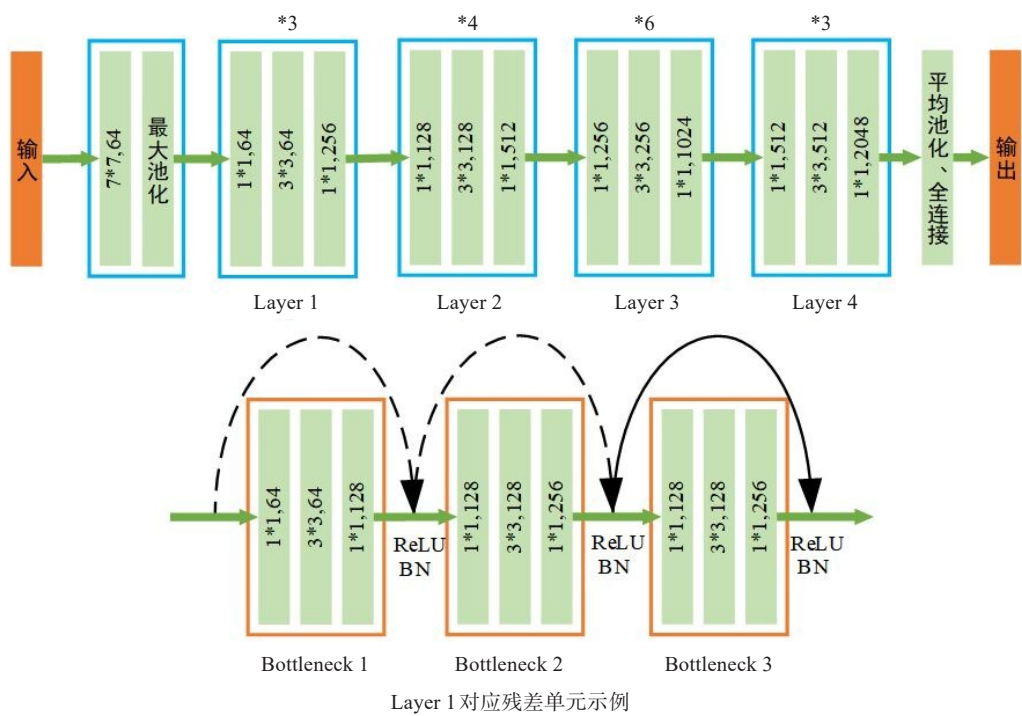


图2 ResNet-50 网络结构

Figure 2 ResNet-50 network architecture

1.2.2 Transfome 模块 残差网络在提取细节信息和高分辨率特征方面表现出色,但由于连续卷积操作,上下文信息的关联性降低^[26]。为解决这个问题,引入注意力机制,使网络能更集中地关注目标区域的信息。Transformer作为一种自注意力机制,能进一步增强网络对边缘信息的关注,建立像素点之间的

长距离特征依赖关系,从而提升分割效果^[27]。这种改进能更好地处理分割结果中边缘细节的不足。

假设输入的图像为 $p_i \in R^{H \times W \times C}$, H 、 W 和 C 分别表示输入图像的高、宽和通道数,Transformer的高级特征表示过程如下:

(1)首先进行输入Transformer的格式初始化,将

p_i 转为2D向量序列表示的图像块,并且尺寸设为 $P \times P$,得到 N 个图像块的输入。

(2)使用可训练的线性映射方式将图像预处理过程得到的序列化图像映射到潜在的高维空间,将每个图像块表示为一个高维向量。映射过程表示为:

$$Z = [p_1E, p_2E, \cdots, p_nE]$$

(1)

其中, Z 表示映射得到的高维向量结果, p_1E, p_2E, \cdots, p_nE 表示不同图像块的类别编码。

(3)本研究采用的编码器模块遵循原始的Transformer结构设计方法,由多个堆叠的Transformer层组成。每个Transformer层的结构如图3所示,包含两个标准化层、一个多头自注意力模块(Multi-Headed Self-Attention, MSA)和一个多层感知机模块(Multilayer Perceptron, MLP)。该模块对嵌入向量进行处理。

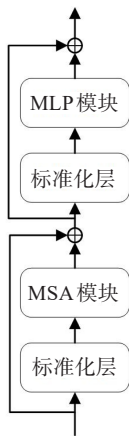


图3 Transformer层结构
Figure 3 Transformer layer structure

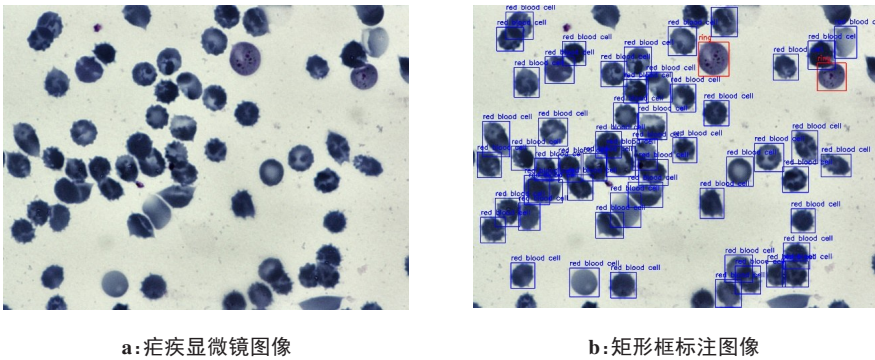


图4 疟疾显微镜图像和其对应矩形框标注图像示例
Figure 4 Example of a malaria microscope image and its corresponding rectangular box labeled image

2.2 实验设置

将数据经过预处理和扩充后传入网络中训练,

(4)经过多个Transformer编码器模块的处理后,最后一个编码器的输出被视为图像的高级特征表示。

2 实验分析

2.1 数据集及预处理

2.1.1 数据集 在MIT-Malaria数据集上进行实验,该数据集由Broad研究所的Dr.Anne Carpenter团队收集,包括1 328张疟疾显微镜图像以及对应每张图像的json格式的标注文件。数据集包含红细胞、环状体、滋养体、裂殖体、配子体、白细胞以及困难细胞7类细胞,其中红细胞和白细胞属于未受感染的细胞,红细胞约占总细胞数的97%。无法区分类别的细胞定义为困难细胞,困难细胞和白细胞数据数量太少,在训练中被忽略。本实验分割的类别分为5类:健康红细胞、环状体、滋养体、裂殖体、配子体。图4为经过手动矩形边界框标注的疟疾显微镜图像示例。

2.1.2 数据预处理 实验首先对数据集进行手动掩膜标注,结合专家标注的矩形框标注图像对数据进行掩膜标注,给不同类型的细胞标注不同颜色的掩膜,健康红细胞、环状体、滋养体、裂殖体和配子体分别用深蓝色、粉色、浅蓝色、红色和绿色来表示。原始疟疾显微镜图像和其对应的掩膜标注图像示例如图5所示。

其次,由于数据集中红细胞占比太大,导致网络对不同类型的特征的学习不均衡,占比较少的细胞无法预测出结果。因此本研究根据标注信息,对类别数量较少的细胞进行数据增强处理,经过旋转、镜像、裁剪等方法进行数据扩充。最后将图像尺寸调整为256×256,并进行归一化处理,再将数据集转换为VOC格式进行训练。

使用Adam优化器训练网络参数。最大学习率设置为0.000 1,最小学习率为0.000 001。Adam优化器适

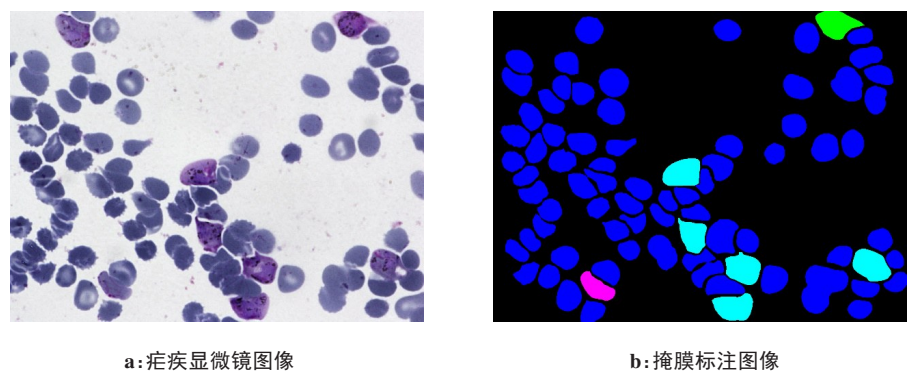


图5 疟疾显微镜图像和其对应的掩膜标注图像示例

Figure 5 Example of a malaria microscopy image and its corresponding mask labeled image

用于快速拟合高维度数据并使参数平稳。为平衡速度和效果,将批量大小设置为4,即每次将4张图像传入网络进行训练。预处理后的数据集按照9:1的比例随机划分为训练集和测试集。模型训练时,最大epoch设为100。

本文方法使用PyTorch框架进行实现,依赖库版本包括PyTorch-GPU 1.10、PIL和Numpy等。模型学习设备的配置为CPU: Intel Core i9-13900K、内存: 16 GB、操作系统: Windows 11 x64、GPU: NVIDIA GeForce RTX 4060。

2.3 性能评价指标

为评估本文方法与现有的一些方法之间的性能差异,使用DSC、mIoU、MPA、准确率(Accuracy, Acc)、精确率(Precision)、召回率(Recall)综合评估模型性能。

DSC主要用于衡量分割结果与真实标签之间的相似性,对两者的重叠部分进行评估。计算公式为:

$$DSC = \frac{2TP}{2TP + FP + FN}$$

(2)

其中,TP表示一个样本被预测为正类,并且真实标签为正类;FN表示一个样本被预测为反例,但是真实标签为正例;FP表示一个样本被预测为正例,但是真实标签为反例;TN表示一个样本为预测为反例,并且真实标签为反例。

对于多目标类别的语义分割任务,mIoU这一评价指标被广泛使用,它的定义为目标类的预测值与真实值的交集与并集之比。具有N个类别的语义分割任务的mIoU计算公式为:

$$mIoU = \frac{1}{N} \sum_{i=0}^N \frac{TP_i}{TP_i + FP_i + FN_i}, i = 0, 1, \dots, N$$

(3)

MPA是一种用于衡量图像分割算法在像素级别上的准确性的常用指标。具有N个类别的语义分割任务的MPA计算公式见式(4):

$$MPA = \sum_{i=0}^N \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i}, i = 0, 1, \dots, N$$

(4)

准确率衡量的是模型预测正确的像素数量占总像素数量的比例,即正确分类的像素数除以总像素数。

$$Acc = \frac{TP + TN}{TP + TN + FP + FN}$$

(5)

精确率衡量的是在模型预测为某一类别的像素中,有多少是真正属于该类别的像素。它可以用来评估模型的分类准确性。

$$Precision = \frac{TP}{TP + FP}$$

(6)

召回率描述的是模型能正确识别出正样本的能力,即模型能覆盖多少实际的正样本,不考虑模型预测的准确性。召回率和精确率是两个互补的指标。精确率关注的是模型预测为正样本的准确性,而召回率关注的是模型对实际正样本的覆盖率。实际应用中,需要综合考虑召回率和精确率,以全面评估模型性能。

$$Recall = \frac{TP}{TP + FN}$$

(7)

2.4 结果分析

2.4.1 不同分割算法评价指标对比 选择几种广泛使用的分割算法,包括DeepLabv3+、U-Net和ResUNet,并将它们与本文方法进行比较(表1)。本文方法相比U-Net,DSC得到5.43%的提升,mIoU得到4.86%的提升,MPA得到3.98%的提升;与ResUNet相比,本文方法的性能也有所提升。综合看来本文方法在疟疾感染红细胞的分割上表现较优越,能提供相对较高的分割准确度。

2.4.2 分割效果对比 数据的对比不能直观说明网络的性能,因此,为评估不同分割算法的性能,本研究进行分割结果的可视化对比分析。在测试集上,使

表 1 分割结果对比
Table 1 Comparison of segmentation results

模型	DSC	mIoU	MPA	准确率	精准率	召回率
DeepLabv3+	0.8177	0.7006	0.8250	0.9337	0.8085	0.8300
U-Net	0.8197	0.7199	0.8130	0.9400	0.8241	0.8173
ResUNet	0.8468	0.7619	0.8388	0.9450	0.8559	0.8388
本文方法	0.8740	0.7685	0.8528	0.9655	0.8795	0.8641

用 DeepLabv3+、U-Net、ResUNet 和本文方法进行实际图像分割任务,并取分割后可视化结果。通过可视化分割结果可以发现 U-Net 会出现错误分割的情况并且分割出的细胞边界模糊,ResUNet 分割出的感染细胞边缘上也存在一定误差。本文方法相较于其它的方法,分割效果更佳,细胞边缘分割更精确(图6)。

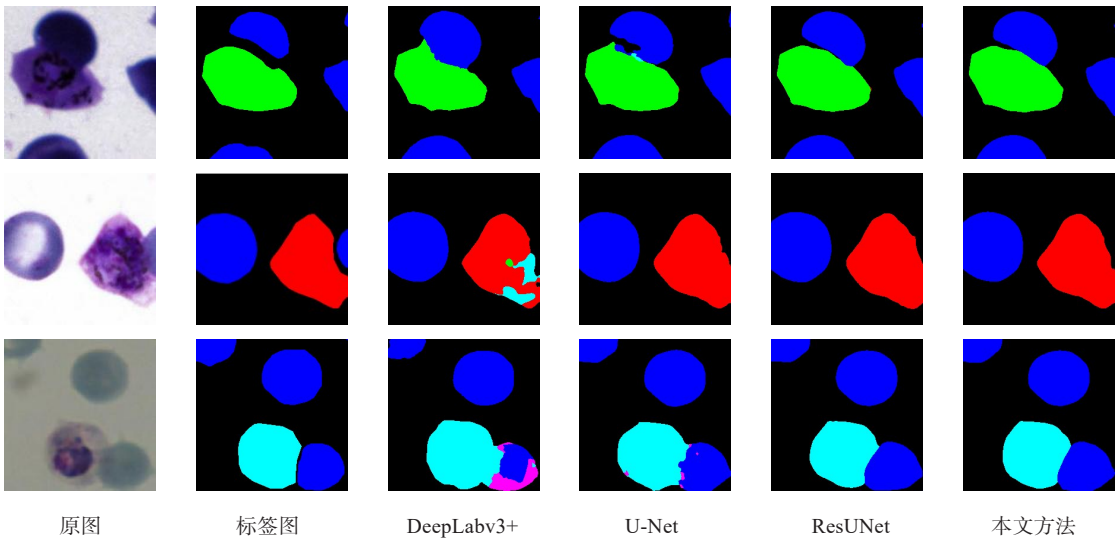


图6 分割结果的可视化对比
Figure 6 Visualization comparison of segmentation results

3 结 论

本研究以深度学习技术解决红细胞分割问题为方向,以提高分割精度和分割类别的准确性为目标,提出一种改进的 U-Net 网络,在 U-Net 中引入 ResNet 和 Transformer,提高网络对深层特征的提取能力,并使用 Transformer 的自注意力机制,使得网络更关注目标区域。经过数据扩充等预处理方法进行数据处理,并进行模型训练和评估。与 Deeplabv3+、U-Net 和 ResUNet 相比,本文方法获得更优的评价指标并且可视化结果显示本文方法的分割边缘精度更好,分割的类别准确度更高,识别错误的概率也更低。

【参考文献】

[1] Paul S, Batra S, Mohiuddin K, et al. A novel ensemble weight-assisted yolov5-based deep learning technique for the localization and detection of malaria parasites[J]. Electronics, 2022, 11(23): 3999.
[2] 刘巧, 刘珏. 全球疟疾流行病学研究进展[J]. 中国公共卫生, 2023, 39(4): 509-513.
Liu Q, Liu J. Global progress in malaria epidemiology[J]. Chinese Journal of Public Health, 2023, 39(4): 509-513.
[3] Çinar A, Yildirim M. Classification of malaria cell images with deep learning architectures[J]. Ing des Syst d'Information, 2020, 25(1): 35.
[4] Mbanefo A, Kumar N. Evaluation of malaria diagnostic methods as a key for successful control and elimination programs[J]. Trop Med

Infect Dis, 2020, 5(2): 102.
[5] Hung J, Carpenter A. Applying faster R-CNN for object detection on malaria images[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2017: 56-61.
[6] Delgado-Ortet M, Molina A, Alf  rez S, et al. A deep learning approach for segmentation of red blood cell images and malaria detection[J]. Entropy, 2020, 22(6): 657.
[7] Lee H, Chen YP. Cell morphology based classification for red cells in blood smear images[J]. Pattern Recognit Lett, 2014, 49: 155-161.
[8] Das DK, Ghosh M, Pal M, et al. Machine learning approach for automated screening of malaria parasite using light microscopic images [J]. Micron, 2013, 45: 97-106.
[9] Savkare SS, Narote SP. Automatic system for classification of erythrocytes infected with malaria and identification of parasite's life stage[J]. Proc Technol, 2012, 6: 405-410.
[10] Jalari S, Reddy BE. A novel two-stage thresholding method for segmentation of malaria parasites in microscopic blood images[J]. J Biomed Eng Med Imaging, 2017, 4(2): 31.
[11] Savkare SS, Narote SP. Blood cell segmentation from microscopic blood images [C]//2015 International Conference on Information Processing. IEEE, 2015: 502-505.
[12] Jdey I, Hcini G, Ltifi H. Deep learning and machine learning for malaria detection: overview, challenges and future directions[J]. arXiv preprint. arXiv:2209.13292, 2022.
[13] Jan Z, Khan A, Sajjad M, et al. A review on automated diagnosis of malaria parasite in microscopic blood smears images[J]. Multimed Tools Appl, 2018, 77: 9801-9826.
[14] 朱世祺, 徐昶, 周鑫, 等. 基于 DeepLab V3+深度神经网络的结直肠息肉内镜图像分割[J]. 中国医学物理学杂志, 2023, 40(8): 944-949.
Zhu SQ, Xu C, Zhou X, et al. Endoscopic image segmentation of colorectal polyps based on DeepLab V3+ deep neural network[J]. Chinese Journal of Medical Physics, 2023, 40(8): 944-949.
[15] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for

- biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference. Springer International Publishing, 2015: 234-241.
- [16] Li D, Tang P, Zhang R, et al. Robust blood cell image segmentation method based on neural ordinary differential equations[J]. Comput Math Methods Med, 2021, 2021: 5590180.
- [17] Zhang M, Li X, Xu M, et al. Automated semantic segmentation of red blood cells for sickle cell disease[J]. IEEE J Biomed Health Inform, 2020, 24(11): 3095-3102.
- [18] Li Z, Zhang H, Li Z, et al. Residual-attention UNet++: a nested residual-attention U-Net for medical image segmentation[J]. Appl Sci, 2022, 12(14): 7149.
- [19] Liu Z, Yuan H. An Res-Unet method for pulmonary artery segmentation of CT images[J]. J Phys Conf Ser, 2021, 1924(1): 012018.
- [20] Nurçin FV. Improved segmentation of overlapping red blood cells on malaria blood smear images with TransUNet architecture[J]. Int J Imaging Syst Technol, 2022, 32(5): 1673-1680.
- [21] 胡晓阳, 李哲. 基于卷积神经网络和 Transformer 的肝脏 CT 图像分割方法[J]. 中国医学物理学杂志, 2023, 40(4): 423-428.
- Hu XY, Li Z. Liver CT image segmentation method based on convolutional neural network and Transformer[J]. Chinese Journal of Medical Physics, 2023, 40(4): 423-428.
- [22] Hou J, Yan C, Li R, et al. Lung nodule segmentation algorithm with SMR-UNet[J]. IEEE Access, 2023, 11: 34319-34331.
- [23] Chen J, Lu Y, Yu Q, et al. TransUNet: Transformers make strong encoders for medical image segmentation[J]. arXiv preprint, arXiv: 2102.04306, 2021.
- [24] 杨朝乾, 樊永生, 余红英. 一种改进的 U-Net 网络的图像语义分割算法[J]. 中北大学学报(自然科学版), 2023, 44(04): 397-402.
- Yang CQ, Fan YS, Yu HY. An improved image semantic segmentation algorithm for U-Net networks[J]. Journal of North University of China (Natural Science Edition), 2023, 44(4): 397-402.
- [25] Ali Y, Janabi-Sharifi F, Beheshti S. Echocardiographic image segmentation using deep Res-U network[J]. Biomed Signal Process Control, 2021, 64: 102248.
- [26] Ehab W, Li Y. Performance analysis of UNet and variants for medical image segmentation[J]. arXiv preprint, arXiv:2309.13013, 2023.
- [27] 侯向丹, 李紫宇, 牛敬钰, 等. 结合注意力机制和多路径 U-Net 的视网膜血管分割[J]. 计算机辅助设计与图形学学报, 2023, 35(1): 55-65.
- Hou XD, Li ZY, Niu JY, et al. Retinal vascular segmentation combining attention mechanism and multipath U-Net[J]. Journal of Computer Aided Design and Computer Graphics, 2023, 35(1): 55-65.

(编辑:谭斯允)