

引用格式: HUANG Hong, WANG Tao, LI Yuan, et al. Cancer Pathological Segmentation Network Based on Depth Feature Fusion[J]. Acta Photonica Sinica, 2022, 51(3):0310001

黄鸿,王涛,李远,等. 基于深度特征融合的癌症病理图像分割网络[J]. 光子学报, 2022, 51(3):0310001

※封面论文※

基于深度特征融合的癌症病理图像分割网络

黄鸿¹, 王涛¹, 李远¹, 周凡琳², 李昱²

(1 重庆大学 光电技术与系统教育部重点实验室, 重庆 400044)

(2 重庆大学附属肿瘤医院 病理科, 重庆 400030)

摘要:卷积神经网络在癌症病理图像分割中具有突出表现,但在临床应用上依然面临着切片染色多样、分辨率差异大等挑战。针对上述问题,提出了一种病灶分割网络 HU-Net,提高了癌症病理图像的分割精度。HU-Net 使用 U-Net 网络作为基本结构,利用经过预训练的 EfficientNet-B4 作为网络特征编码器,解码器部分在 U-Net 网络上进行改进,将不同深度特征重新进行组合进行特征融合,提升了深层特征在预测中的作用。在此基础上,利用各深度层融合特征预测输出,构建多损失函数共同训练,使深层语义信息更具鉴别力。最后,采用改进的通道注意力模块对融合特征进行选择,使网络对不同分辨率图像的适应性增强,提升了模型筛选重要特征的能力。在 BOT 数据集和 SEED 数据集上分别进行癌症病灶分割实验,所提方法的 DICE 系数得分在两个数据集上分别达到 77.99% 和 82.94%, 准确度得分分别达到 88.52% 和 87.42%。该方法相较于 U-Net 和 DeepLabv3+ 等网络有效提升了癌症病理图像病灶分割精度,实现了更准确的癌症病灶定位和分割。

关键词:计算机图像处理;分割算法;特征融合;癌症;病理图像;深度学习

中图分类号: TP391

文献标识码: A

doi:10.3788/gzxb20225103.0310001

0 引言

在癌症病理图像的临床诊断中,基于病理组织切片的显微分析可用于观察癌组织的发展状态从而判断癌症的变化、转移情况,因此病理图像又被称为癌症临床诊断中的金标准^[1]。近年来,随着全视野数字切片技术的发展,针对病理图像的定量分析研究逐渐兴起。病理图像智能分析对癌症患者的诊断和预后具有重要作用,其中病理切片的病灶分割可用于计算癌症病变的相关量化指标、定位癌组织区域以及参与制定手术方案等^[2-3]。精确的计算机辅助病理分割不仅能帮助临床医生提高诊断效率,还可以提升诊断的准确性,如何针对癌症病理图像特点设计高效的分割算法已成为医学图像分析领域的研究热点^[4]。

传统的图像分割方法主要包括阈值分割法、区域生长法和图割法等。典型的阈值分割法有大津法^[5],其主要原理是通过最大化类间方差来计算分割阈值,因其简单和计算量较小而被广泛应用^[6]。典型的区域生长法有分水岭算法^[7],其主要原理为通过将像素灰度模拟为海拔高度来模拟地理环境,谷底和峰顶的边界即为分水岭,其分割结果具有空间连续性,在分割连续小目标时效果较好,但该算法受噪声影响大且不适用于大区域分割。典型的图割法有 Graphcut 法^[8],其原理为先建立一张加权图,通过边权值的大小来筛选边,使划分的子图彼此尽量相互分离,从而达到分割的目的^[9]。该方法具有较好的泛化性被广泛应用于医学图像分割,但其时间复杂度较高对计算资源的消耗大。上述传统图像分割方法共有的缺点是依赖于手工特征,分割中不能充分利用图像中的语义信息,导致性能受限。

深度学习方法具有特征提取能力强和便于训练优化的特点,开始应用于医用图像分割中。这些方法主要为全卷积模型,代表性方法包括全卷积网络(Fully Convolutional Networks, FCN)^[10]、U-Net 网络^[11]和 U-

基金项目:国家自然科学基金(No. 42071302),重庆市留学人员回国创业创新支持计划(No. cx2019144)

第一作者:黄鸿(1980—),男,教授,博士,主要研究方向为模式分类、医学影像处理、遥感影像处理等。Email: hhuang@cqu.edu.cn

收稿日期:2021-07-21;录用日期:2021-10-19

<http://www.photon.ac.cn>

Net++网络^[12]等。早期,CIRESAN D C等^[13]基于卷积神经网络(Convolutional Neural Networks, CNN)利用滑窗进行像素级目标分割,但存在计算量大和效率低的问题。LONG J等提出了FCN网络作为替代方法,将经典卷积神经网络的全连接层替代为上采样卷积层,从而建立了一个只包含卷积层的网络,取消了输入尺寸固定的限制,可灵活应用于不同大小的图像上。在FCN网络的基础上,U-Net网络进一步发展了编码、解码的思想,其上采样和下采样结构呈U形对称,其解码层增加了相应的上采样层和通过跳跃连接进行特征融合使提取特征更加有效,提高了分割精度。ZHOU Z等在U-Net网络基础上提出了U-Net++网络,通过将长连接改用密集连接、在训练时自动剪枝的方式,使网络可以自适应学习合适的下采样层数以获得最优的对应U-Net网络深度,在提升分割性能的同时没有明显增大计算资源的消耗。黄鸿等^[14]基于融合自适应加权聚合策略提出了*i*U-Net++方法,该方法可以自适应学习各层特征权重,使包含不同信息的特征根据贡献程度聚合起来得到分割结果,在浅层特征融合中突出了关键特征,提升了分割性能。此外,还有学者将病理图像研究拓展到了高光谱病理图像的研究中,在利用空间信息的基础上增加了更丰富的病理组织光谱信息^[15]。郑少佳等^[16]基于傅里叶变换多频率通道注意力机制结合Inception网络构建了一种胆管癌高光谱图像分割网络,在一定程度上解决了数据不平衡问题并在胆管癌上取得了较高的分割精度。WEI X等^[17]设计了一个端到端网络,通过在原始光谱数据和其主成分之间建立逐像素映射实现了无监督学习,有效解决了监督卷积神经网络模型训练样本有限的问题。

上述方法针对病理图像通过融合不同深度U-Net网络特征,在CT、MRI等医学图像上取得了不错的分割效果,但直接在浅层进行预测,在一定程度上削弱了高层语义特征的指导作用,导致没有充分利用高层语义特征在精确定位病灶区域上的重要作用,尤其是病理图像存在切片染色多样、分辨率差异大等问题,限制了病理图像分割的性能。

针对上述问题,本文提出了一种改进的层级特征融合分割方法(Hierarchical Feature-fused U-Net, HU-Net)。该方法基于U-Net网络,利用EfficientNet-B4网络作为特征编码器,在解码器部分改进了特征融合方法,利用改进的通道注意力模块对融合特征进行筛选,训练时不同深度预测下对应的损失函数都进行反向传播,获得了更有效的深层语义特征,从而提升了模型获取和利用全局语义特征的能力,改善了癌症病理图像病灶分割精度。

1 本文方法

1.1 整体网络结构

U-Net是一种具有对称性的分割网络,它通过添加长连接的方式在每一层中都进行高层全局特征和局部细节特征的特征融合,有效利用了不同重要性特征进行分割预测^[18]。因此采用U-Net网络作为分割模型的基本网络结构,其整体网络结构如图1。

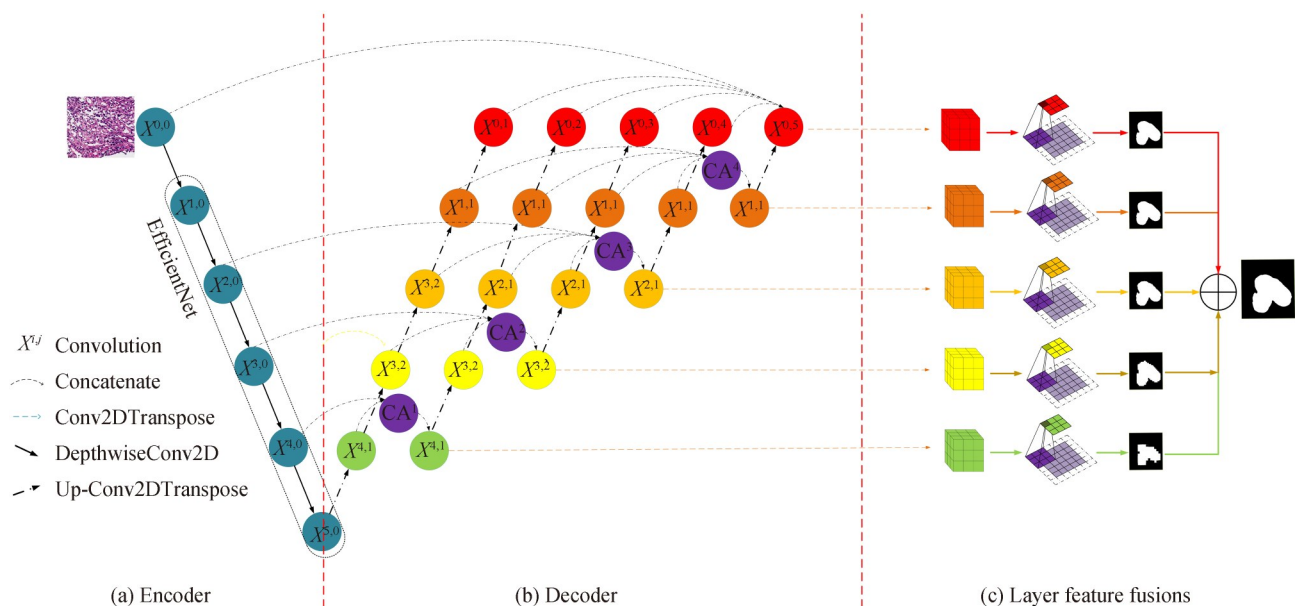


图1 HU-Net网络整体结构

Fig.1 The overall structure of the proposed HU-Net algorithm

该网络主要包括编码器、解码器和特征融合预测三个部分。图1(a)为基于 EfficientNet-B4 网络的特征编码器。图1(b)为解码器,在解码器中改进了 U-Net 原有特征融合方式,将不同深度特征重新进行了融合,从深层逐渐到浅层,每一层融合特征都不断上采样到最浅层,每层特征融合时都融合之前深层中的高层语义信息。图1(c)为特征融合预测模块,训练时将不同深度的层融合特征都进行预测输出,都计算损失函数并反向传播。同时,在网络解码器中应用了通道注意力模块(CA¹-CA⁴)对融合后的特征进行通道特征选择,如图1(b)中紫色部分。

1.2 编码器模块

EfficientNets 是 Google 利用神经网络架构搜索获得的一系列网络模型,分别为 EfficientNet-B0~B7^[19]。通过考虑运算量与计算资源,本文网络选择在 Imagenet 数据集上预训练的 EfficientNet-B4 网络作为特征编码器,通过迁移学习使编码器在病理图像上更快学习到关键特征,在实际使用时去掉了全连接层,具体结构如图2。该网络主要由 32 个移动翻转瓶颈卷积模块(Mobile Inverted Bottleneck Convolution, MBConv)和 2 个卷积层构成,各个 MBConv 块中卷积核的大小分别如图2所示。通过将步长设为 2,使输出特征图在长和宽两个方向上分别缩小为原来的一半,实现了压缩特征图以进行特征编码,将 EfficientNet-B4 中不同下采样特征通过图1中的跳跃连接和解码器特征进行融合从而代替了原 U-Net 中的编码器进行特征编码。同时, EfficientNet-B4 网络还引入了通道注意力思想,将特征图在特征通道上进行压缩和激发,使得在编码过程中进一步突出了关键特征。

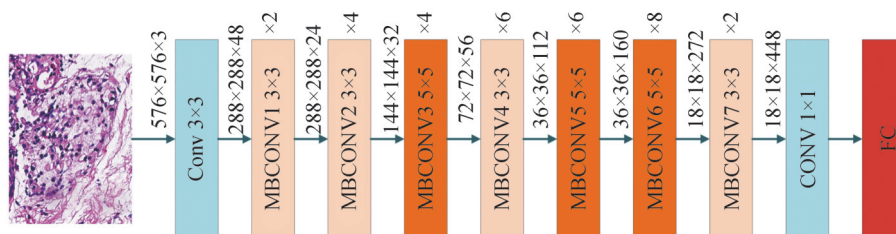


图2 EfficientNet-B4 整体结构
Fig.2 The structure of EfficientNet-B4

1.3 解码器模块

U-Net 网络只使用最浅层融合特征进行预测, U-Net++ 通过增加短连接和剪枝的方式实现了自动学习 U-Net 网络的合适层数,丰富了模型表示复杂特征的能力,但这种特征聚合方式并不能充分利用深层特征的全局信息以适用于癌症病理图像分割。针对病理图像内容复杂和相似性高的特点,需要进一步增加全局特征的预测作用以增强模型对病灶的区分能力。因此,在保留跳跃连接的基础上设计了如图1所示的解码器,对每一层要上采样的特征都分别进行若干次上采样,直至最浅层。

在每一层融合时,除了融合上一解码层和对应编码层的输出,还包括之前解码层上采样到本层的特征,融合方式均为按特征通道进行叠加。设从网络最深层到最浅层分别为 0, 1, 2, ..., 当前层为第 i 层, $\text{Conc}(\cdot)$ 表示特征叠加融合操作,则每一层中特征融合表示为

$$X = \text{Conc}(X_0, X_1, \dots, X_i) \quad (1)$$

通过式(1)中的融合方式, HU-Net 网络增强了深层特征在模型预测中的重要性,最深层的语义特征在浅层进行预测时依然发挥作用,但融合特征数量的增加带来了融合后特征通道数量的增加。为了确定不同通道特征的重要性并筛选有效特征,使用了通道注意力模块对融合特征进行选择。为了使注意力模块更加有效以及方便使用,降低对网络结构的依赖,对通道注意力模块的使用采用如图3所示流程。输入特征在经过通道注意力模块校正后首先使用卷积操作将特征通道压缩为原来的一半,接着对特征进行批归一化和激活,最后将通道压缩的特征进行叠加。这样可以进一步突出经通道注意力校正得到的关键特征,同时使注意力模块在网络各个地方可以灵活使用。具体的融合方式如图1(b)。

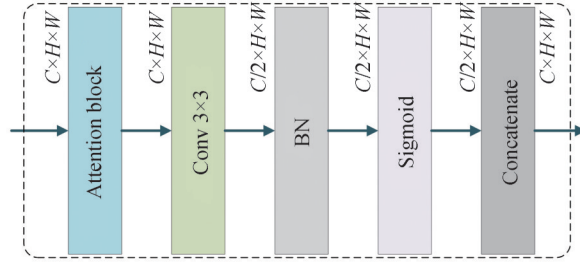


图3 注意力模块的使用方法
Fig.3 The method of using attention block

1.4 层级融合模块

为了加速模型训练收敛以及更有效地进行模型微调,同时使网络在训练过程中在不同深度都获得更有效特征,设计了如图1(c)所示的特征融合预测模块。在解码器的不同层,都将该层融合筛选后的特征通过转置卷积直接上采样到图像原尺寸,并利用Sigmoid函数进行激活,输出之后得到每层对应的分割预测结果。然后,将不同层的预测进行叠加融合,再经过卷积并激活后得到融合的预测结果作为模型最终预测。训练时,模型对最终的融合预测结果和各层单独的预测结果都进行输出,分别依次计算损失函数和进行误差反向传播。在模型验证和测试时,只采用融合预测结果作为模型输出。通过这种方式,可以直接在深层和浅层同时得到更有效的特征,通过观察不同层的预测结果可以分析不同层的训练学习情况以及评估模型的特征表达能力,增强了模型的可解释性。

上述融合和训练方式,使网络在深层中将更加关注病灶位置,可以获得更有效的高层语义信息进行病灶区域定位,利用浅层特征进行边界细节预测,使模型整体分割预测更加精确。

1.5 通道注意力模块

在U-Net网络中,使用了改进的通道注意力模块(CA¹⁻⁴),具体网络结构如图4。输入为编码特征和解码特征叠加得到的融合特征 $X \in \mathbb{R}^{C \times H \times W}$ 。在该注意力模块中为了提升模型的收敛速度首先针对输入进行了卷积,具体可表示为

$$X_i^{k+1} = R \left[\sum_{j=1}^C (X_j^k * H_i^{k+1} + b_i^{k+1}) \right] \quad (2)$$

式中,*为卷积操作, $X^k \in \mathbb{R}^{C \times H \times W}$ 表示卷积层输入,为模型第 k 层的输出, X_j^k 表示第 j 通道特征。 $X^{k+1} \in \mathbb{R}^{C \times H \times W}$ 表示经过卷积层后得到的第 $k+1$ 层特征, $\sum(\cdot)$ 操作代表卷积层中第 i 个卷积核对 C 通道输入特征的卷积操作,其中 H_i 和 b_i 分别表示对应的可学习权重参数和偏置项。 $R[\cdot]$ 表示修正线性单元激活函数(Rectified Linear Unit, Relu)。在此基础上,进行批归一化操作

$$\hat{X}^{k+1} = \gamma \frac{X^{k+1} - E(X^{k+1})}{\sqrt{\text{Var}(X^{k+1})}} + \beta \quad (3)$$

式中, X^{k+1} 代表卷积操作的输出结果作为输入, \hat{X}^{k+1} 代表批归一化的输出结果, $E(\cdot)$ 和 $\text{Var}(\cdot)$ 操作分别表示期望和方差操作, γ 和 β 表示批归一化层中的可学习参数。

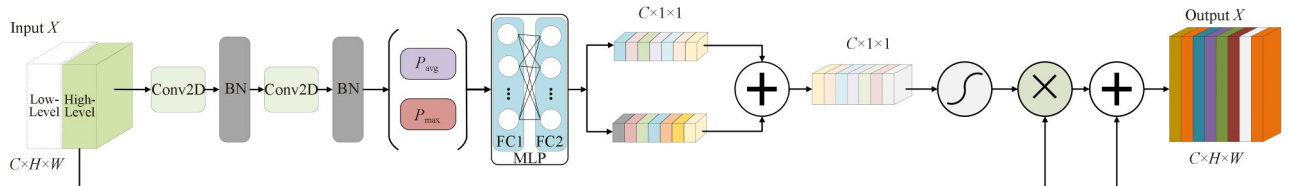


图4 通道注意力模块结构
Fig.4 The structure of attention block

进行两次卷积和批归一化操作后,需利用全局池化学习全局参数。目前的压缩激励注意力网络(Squeeze-and-Excitation Networks, SENet)^[20]只利用了全局平均池化(Global Average Pooling, GAP)来提

取对应通道权重特征,但容易导致细节信息丢失。因此,本文增加了通过全局最大池化(Global Max Pooling,GMP)来提取通道特征,通过保留更多信息从而增强注意力模块的学习和表示能力。

$$P_{\text{avg}} = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X^k(i, j) \quad (4)$$

$$P_{\text{max}} = \text{Max}(X^k) \quad (5)$$

式(4)表示对池化层输入特征 $X^k \in \mathbb{R}^{C \times H \times W}$ 的每一个通道求均值,式(5)中 $\text{Max}(\cdot)$ 表示针对特征 $X^k \in \mathbb{R}^{C \times H \times W}$ 的每一个通道求最大值,全局平均池化和全局最大池化输出分别为 $P_{\text{avg}} \in \mathbb{R}^{C \times 1 \times 1}$ 和 $P_{\text{max}} \in \mathbb{R}^{C \times 1 \times 1}$ 。

完成上述操作后,利用一个共享多层感知器模块(Multiple Layer Perception, MLP)来获取每一流的全局特征系数。MLP主要包含两个全连接层,第一个全连接层的输出结点数为 C/r ,后面接Relu激活函数,第二个全连接层的输出结点数为 C 。本实验中设定 $r=1$,可根据计算资源情况将 r 设置为 2、3 等,从而平衡算法性能和计算资源消耗。在此基础上,将两组通道权值进行相加,利用 Sigmoid 函数进行激活得到通道注意力系数 $\beta \in [0, 1]^{C \times 1 \times 1}$ 。将获得的通道注意力系数乘以注意力模块输入特征得到经过注意力模块校正的层融合特征。为了在计算时加速运算速度以及提升模型稳定性,使用了残差结构得到经过校正后的输出,可表示为

$$Y_{\text{CA}} = X \cdot \beta + X \quad (6)$$

2 实验结果及分析

在整个实验中,实验的硬件环境为 PANYAO 7048GR,CPU 为正式版 E5-2096 2.3G-3.6G 45M Cache 22NM,内存为 256 G,显卡为 NVIDIA GeForce GTX TITAN RTX,应用的深度学习框架为基于 Ubuntu 系统的 Tensorflow2.0 框架。CUDA 以及 CUDNN 的版本为 CUDA10.1.243 和 CUDNN7.6.5,采用的编程语言为 Python3.8。

2.1 实验数据集

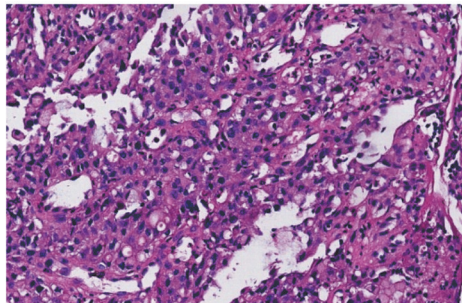
本文使用了两个癌症病理数据集来评估本文算法模型,第一个为 2020 年“华录杯”江苏大数据开发与应用大赛(“SEED”大赛)—医疗卫生赛道癌症病理数据集^[21],第二个数据集为 2017 年中国大数据人工智能创新创业大赛(Brain Of Things, BOT)—病理切片识别 AI 挑战赛中的病理数据集^[22]。上述数据集均为胃部癌变组织,为正常、癌变组织分割问题。

1) SEED 病理数据集

该数据集在江苏省人民医院中收集,共包含 732 个带胃癌病灶精标注的数字病理切片样本,其中有 17 个样本病理图像为空白无法使用,1 个样本出现错误标注,在使用中均被剔除,实验中实际使用样本数为 714。病理图像以及对应的标注都为 RGB 图像。该数据集中图像尺度差异大,如图 5(c)所示。不同图像的尺度从 372×489 到 $12\,047 \times 18\,257$ 变化,尺度上的巨大变化为病灶精确分割带来了挑战。

2) BOT 病理数据集

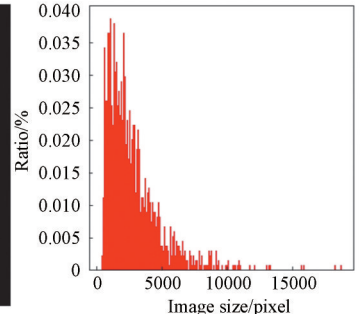
BOT 数据集包含 560 个样例,每个样本包含病理图像和对应的病灶区域标注,图像大小全部为 $2\,048 \times 2\,048$ 。该数据集的病理图像中大部分病灶区域都相对较小,是分割中的难点问题。病理图像和对应专家标注的掩膜如图 5(d)~(g)。标注掩膜图像中,白色区域表示癌组织对应区域,黑色区域表示正常组织。



(a) Original image in SEED dataset



(b) Annotation image in SEED dataset



(c) Size of images in the dataset

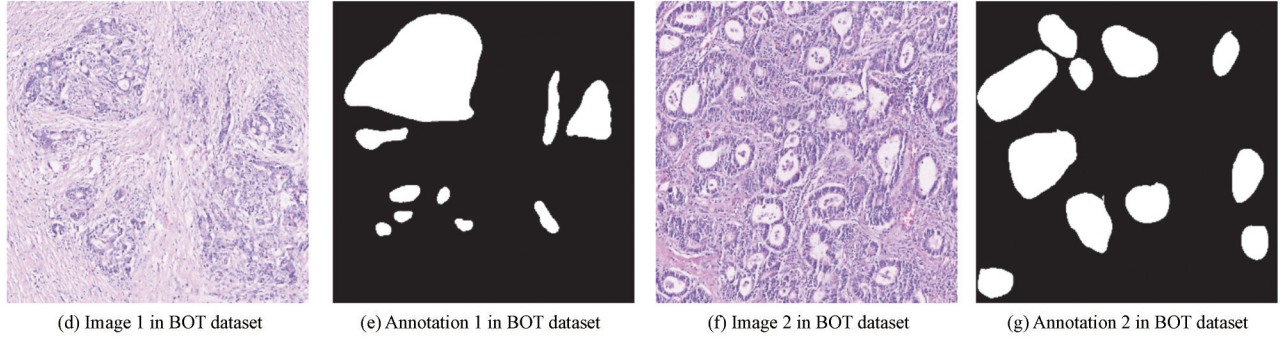


图5 SEED和BOT数据集

Fig.5 The SEED and BOT data sets used in the experiment

2.2 实验设置

为了弥补训练数据相对不足,实验中对训练集数据进行了数据增广,增广比例设置为1.2倍,以减小增广数据可能引入的额外干扰。具体增广方式包括旋转、镜像、翻转操作,其中旋转操作固定为 30° 、 60° 的顺时针和逆时针旋转,镜像操作为水平镜像,翻转操作为上下翻转。每个训练样本的增广概率为20%,对每次数据增广只随机选取一种变换方式处理图像。每种增广方式下获得的图像数量,其比例相近且在增广数据中均匀分布。

出于尽量保留训练图像细节和减少模型计算量,训练HU-Net网络时两个数据集中图像均缩放为 576×576 大小,训练采用随机梯度下降(Stochastic Gradient Descent, SGD)作为模型优化器。实验采用DICE系数差异函数作为训练损失函数,训练时利用模型在不同深度下对应的预测结果共同进行训练。预测时,所有预测结果均以0.5为概率阈值对模型输出进行二值化后得到。训练初始学习率设置为0.01,训练epoch总数设置为150。随着训练模型逐步拟合,当验证精度超过10个epoch没有增加时,则减小学习率,每次学习率减小0.005,最小学习率为0.0001,当经过连续20个epoch验证精度没有提高时结束训练,批次大小(batch size)为5。本实验所有参数为在多次实验条件下确定的最优值。

共选取了DICE系数、交叉比系数(Intersection Over Union, IOU)、敏感度(Sensitivity, Sen)、准确率(Mean Accuracy, MA)和精确率(Precision, Pre)五个评价指标进行定量评价,具体表示为

$$\text{DICE} = \frac{2 |M_a \cap P_{re}|}{|M_a \cup P_{re}|} \quad (7)$$

$$\text{IOU} = \frac{|M_a \cap P_{re}|}{|M_a \cup P_{re}|} \quad (8)$$

$$\text{Sen} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (9)$$

$$\text{Pre} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (10)$$

$$\text{MA} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \quad (11)$$

式中, M_a 表示预测矩阵, P_{re} 表示标签矩阵。TP、FP、TN、FN分别表示为真阳性、假阳性、真阴性、假阴性的对应像素预测结果,其中最终的预测矩阵为将模型预测结果以0.5为阈值进行二值化操作后得出。五个评价指标分别从多个角度定量地度量了预测结果的分割效果,使评价结果较为客观和全面。

2.3 BOT数据集实验结果

2.3.1 HU-Net网络和经典网络对比

为验证HU-Net网络有效性,选取了FCN-VGG16^[10]、U-Net^[11]、SegNet^[23]、DeepLabv3+^[24]、CA-Net^[25]和TransUnet^[26]进行对比。训练时,为了减少参数对于各方法分割性能的影响,各个模型的输入大小都设置为 576×576 大小,批处理大小(Batch size)设置为5,且采用和HU-Net训练相同的数据增广方式。初始学习率和学习率设置如下:最小学习率都调整为0.0001,初始学习率分别设置为0.01、0.005、0.005、0.001、0.005、

0.005,每超过 10 个 epoch 验证精度未提高时触发学习率变化,每次变化学习率下降为原来的 0.8 倍,学习率的设置为多次实验取得的最优值。所有算法的实验结果都为五折交叉验证结果的平均值。

实验前,按照 4:1 的比例划分训练集和测试集,不同算法在 BOT 数据集上的分割结果如表 1。

表 1 不同算法在 BOT 数据集上的结果
Table 1 The result of different algorithms on the BOT dataset

Algorithm	DICE /%	MA /%	Sen /%	Pre /%	IOU /%
U-Net	70.06	86.04	68.57	77.93	58.74
FCN-VGG16	68.01	85.05	66.72	73.67	56.91
SegNet	72.05	85.86	70.81	75.56	60.30
DeepLabv3+	76.52	87.87	75.98	78.81	65.34
TransUnet	75.02	87.49	78.40	75.14	65.09
CA-Net	74.33	83.47	77.53	72.12	62.00
HU-Net	77.99	88.52	79.22	78.78	67.01

*Note: The test result of the model is the average value of the 5-fold cross-validation experiment, the best results are marked in bold

由表 1 可得知,FCN 网络取得了较差的分割精度,这是因为 BOT 病理图像中的病灶区域普遍较小,FCN 网络在进行预测时只利用了最深层语义信息,导致细节信息利用不足,而 U-Net 网络因为融合了细节信息因此取得了更高的预测分割精度。TransUnet、CA-Net 和 SegNet 在保留细节信息的基础上分别采用了更有效的特征提取器、利用注意力模块进行特征筛选和改进上采样的方式,进一步改善了病灶分割效果,但这些方法共同的缺点是没有突出全局特征在病理图像分割中的作用。本文提出的 HU-Net 算法改进了特征融合方式,突出了深层语义信息在预测中的作用,因此在 BOT 数据集上取得了最好的分割效果。DeepLabv3+ 相较于 FCN 网络减少了下采样深度,调节了深层特征在预测中作用,在精确度上取得了不错效果,但是减少下采样会导致计算量显著增加。

为了直观对比不同方法的预测,将标注图像(GT)同所有方法在测试集上的预测结果分别进行叠加显示,其中标签图像用红色表示,模型预测结果使用绿色进行表示,预测结果和标注部分重合区域表示为黄色,结果如图 6。

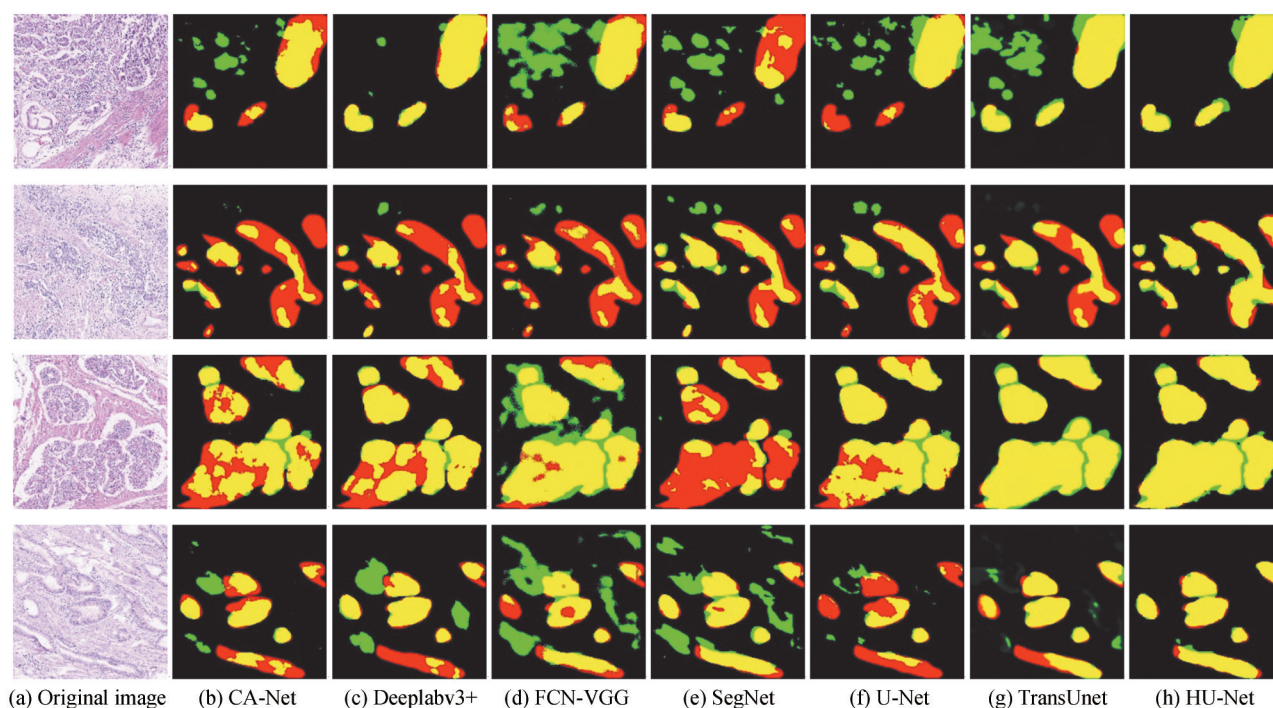


图 6 不同算法在 BOT 数据集上的分割结果
Fig.6 Segmentation masks of different algorithms on BOT dataset

从图6可得知,FCN因为没有利用细节信息所以出现了严重的过分割,U-Net和SegNet网络因为过于强化细节信息在预测中的作用,导致弱化了模型定位病灶的能力,出现欠分割现象。而CA-Net和TansUnet等通过筛选有效特征和增强特征提取能力,在一定程度上改善了分割效果。本文提出的HU-Net算法突出了全局特征同时保留了细节信息,使分割结果更加精准。

2.3.2 HU-Net算法消融实验

在此基础上,进行了消融实验,具体设置为:1)在U-Net的基础上使用在ImageNet数据集上预训练的EfficientNet-B4网络作为特征提取器,实验中称为EU-Net。2)在EU-Net的基础上使用多深度特征融合策略并使用多输出进行多损失函数共同训练,实验中命名为FU-Net。3)在EU-Net的基础上增加使用注意力模块,注意力模块的使用位置同样在各解码层对应的特征融合位置,实验中命名为AU-Net。4)本文所提方法HU-Net网络。消融实验中的模型训练设置与HU-Net训练相同,取五折交叉验证实验结果均值作为各模型得分,实验结果如表2。

表2 不同模型在BOT数据集上的结果
Table 2 The results of different models on the BOT dataset

Algorithm	DICE/%	MA/%	Sen/%	Pre/%	IOU/%
EU-Net	75.26	87.00	77.59	74.59	63.73
FU-Net	77.45	88.52	77.46	78.78	66.53
AU-Net	76.98	87.64	76.78	78.36	65.56
HU-Net	77.99	88.65	79.22	78.95	67.01

*Note: The test result of the model is the average value of the 5-fold cross-validation experiment, the best results are marked in bold

从表2可看出,EU-Net在BOT数据集上提升了分割精度,说明更换有效的编码器可以获得更具鉴别力的特征。同时,添加注意力模块(AU-Net)和深度特征融合(FU-Net)两种改进方式相较于EU-Net又在一定程度上提升了病灶的预测分割精度,FU-Net得分更高说明深层信息对于预测而言有着重要作用,而HU-Net模型融合了两种方式的优点获得了最高DICE系数。

2.4 SEED数据集实验结果

2.4.1 HU-Net网络和经典网络对比

为了测试HU-Net模型在不同数据集上的泛化性,在SEED数据集上进行了对比实验,实验设置同2.3.1节,实验结果如表3。

表3 不同方法在SEED数据集上的结果
Table 3 The result of different algorithms on the SEED dataset

Algorithm	Dice/%	MA/%	Sen/%	Pre/%	IOU/%
U-Net	66.36	78.90	65.61	70.15	53.65
FCN-VGG16	71.47	75.37	78.52	71.36	56.87
SegNet	75.01	80.23	78.82	73.75	61.55
DeepLabv3+	81.86	86.24	82.59	80.66	69.63
TransUnet	80.25	85.16	82.95	79.79	68.94
CA-Net	76.74	81.13	80.68	74.35	63.70
HU-Net	82.94	87.42	84.01	82.56	72.08

*Note: The test result of the model is the average value of the 5-fold cross-validation experiment, the best results are marked in bold

在表3中,FCN比U-Net获得了更好的分割效果,主要是因为SEED病理图像中病灶区域较大,因此突出全局特征的FCN分割精度更高。SegNet、TransUnet和CA-Net分别通过改进上采样方式、替换编码器和利用注意力筛选特征,在不同程度上提升了分割精度。不同于FCN网络进行了多倍特征下采样和U-Net直接利用浅层特征预测,DeepLabv3+减少了下采样倍数并在深度层中进行预测,调节了深层特征在预测中的作用,提升了分割精度。本文提出的HU-Net直接在融合结构上突出了深层特征,并利用注意力模块进行

筛选,因此不同评价指标下均取得了最高得分,相较于U-Net网络DICE系数提升了16.58%。

各算法在测试集上的预测结果如图7。在图7中,FCN能大致分割出病灶区域,但由于没有利用细节特征进行预测,因此在病灶边缘的分割误差较大。U-Net网络因为弱化了全局特征在预测中的作用,定位病灶区域能力受限,出现欠分割现象。相较于对比算法,HU-Net取得了最好的分割效果。

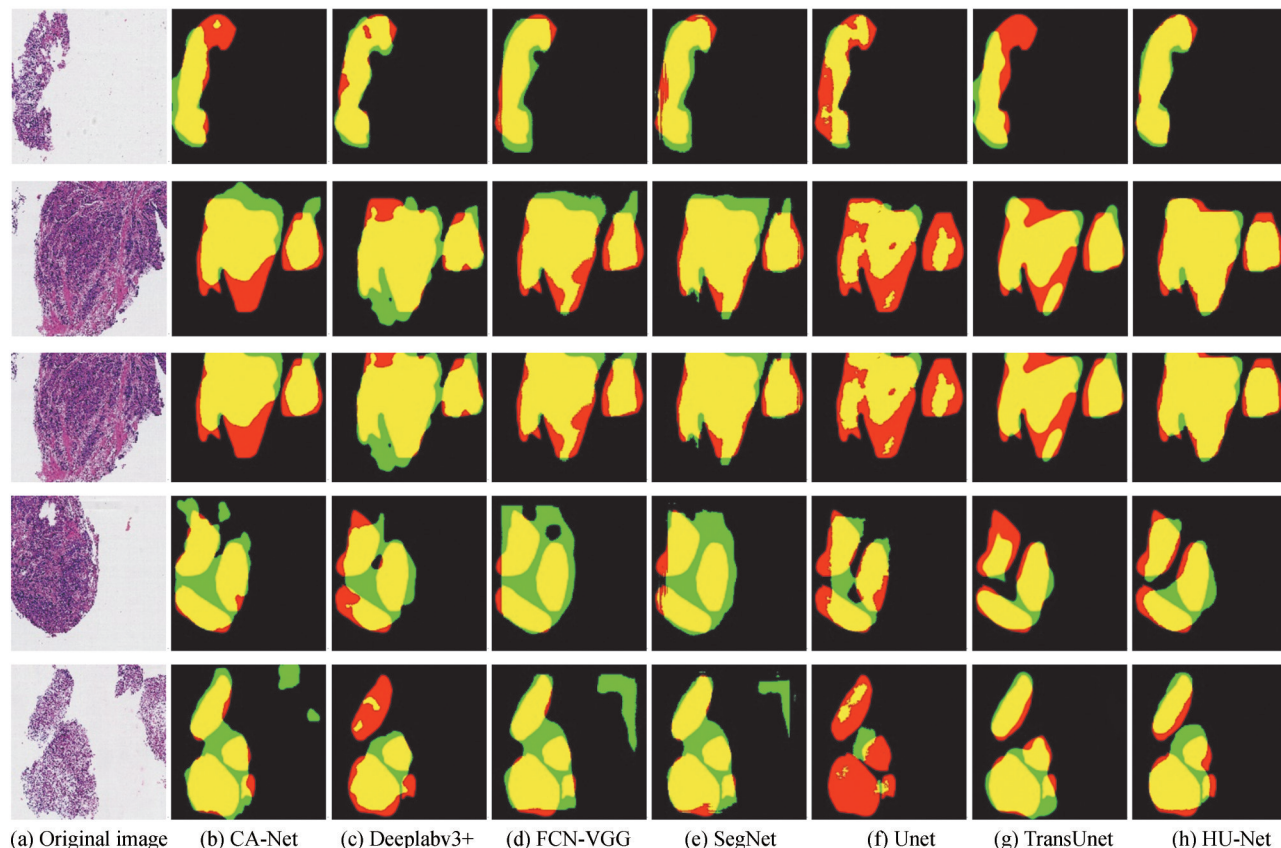


图7 不同算法在SEED数据集上的分割结果

Fig. 7 Segmentation masks of different algorithms on SEED dataset

2.4.2 算法消融实验

在SEED数据集上消融实验设置与2.3.2节一致,实验结果如表4。

表4 不同模型在SEED数据集上的结果
Table 4 The results of different models on the SEED dataset

Algorithm	Dice/%	MA/%	Sen/%	Pre/%	IOU/%
EU-Net	78.95	83.42	80.85	77.56	66.10
AU-Net	79.76	85.02	80.26	79.87	67.74
FU-Net	81.76	86.84	82.23	81.92	70.48
HU-Net	82.94	87.42	84.0	82.56	72.08

*Note: The test result of the model is the average value of the 5-fold cross-validation experiment, the best results are marked in bold

从表4可以看出,AU-Net和FU-Net网络相较于EU-Net在DICE系数上有明显提升,说明采用通道注意力模块和改进的深层特征融合方式都能提高病理图像的病灶分割性能。HU-Net网络在病灶分割上取得了最高的分割精度,验证了该方法对于提升模型分割性能的有效性。结合表2中的实验结果,可以发现深层特征融合方式的改进对模型的分割性能提升尤为明显,这说明在癌症病理图像分割任务中,深层语义特征相较于细节特征具有更重要作用。

对实验进行综合对比分析:1)观察图5、7中所示图像可以看出实验用BOT、SEED数据集具有明显染色

差异,但对比分析表2、4可以发现在添加注意力模块之后AU-Net相较于EU-Net分别提升了1.72%和0.81%,说明本文注意力模块在不同深度层中筛选得到了更有效的通道特征,明显增强了网络应对图像颜色变化的能力。2)FU-Net相较于EU-Net在两个数据集上分别提升了2.19%和2.81%,说明改进特征融合方式和增加不同层的反向传播训练能更好地适应病理图像分割需要,而FU-Net在SEED数据集上取得了更高分割精度的提升,进一步说明这种改进方法在应对尺度变化的图像上表现出了明显作用。3)同时,分析表1、4可以看出HU-Net相较经典网络以及改进对比网络在两个数据集上均取得了最优效果,说明了本文所提方法较好解决了图像染色差异和图像尺度变化的问题。

3 结论

本文提出了一种改进的病理图像语义分割算法HU-Net。该算法首先增强了编码器提取特征的能力,然后利用自然图像到病理图像的迁移学习来缓解病理图像相对不足的问题,在解码器中通过改进深层和浅层特征的融合方式增强了全局特征在预测中的作用,并采用注意力模块筛选了重要特征,最后通过网络多输出训练在各个层中获得了更具鉴别力的特征。在BOT和SEED病理图像数据集上的结果表明,本文所提方法获得了更好的分割效果,其DICE分割精度分别可达77.99%和82.94%,从而为医生临床诊断提供辅助支持。

病理图像往往具有大小不一、尺度变化大的特点,本文所提方法直接将所有图像统一变换为了576×576大小以方便训练,可能会导致病理图像细节信息丢失,因此下一步研究工作将关注如何更有效利用图像中的细节信息进行网络模型训练,改善分割性能。

参考文献

- [1] LITJENS G, KOOI T, BEJNORDI B E, et al. A survey on deep learning in medical image analysis[J]. *Medical Image Analysis*, 2017, 42:60-88.
- [2] COUDRAY N, OCAMPO P S, SAKELLAROPOULOS T, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning[J]. *Nature Medicine*, 2018, 10(24):1559-1567.
- [3] GHAZNAVI F, EVANS A, MADABHUSHI A, et al. Digital imaging in pathology: whole-slide imaging and beyond [J]. *Annual Review of Pathology-Mechanisms of Disease*, 2013, 8(1):331-359.
- [4] JIN Xu, WEN Ke, LV Guofeng, et al. Survey on the applications of deep learning to histopathology[J]. *Journal of Image and Graphics*, 2020, 25(10):1982-1993.
金旭,文可,吕国锋,等.深度学习在组织病理学中的应用综述[J].*中国图象图形学报*,2020,25(10):1982-1993.
- [5] OTSU N. A threshold selection method from gray-level histograms [J]. *IEEE Transactions on Systems Man & Cybernetics*, 2007, 9(1):62-66.
- [6] CHEN X, LIAN C F, WANG L, et al. Diverse data augmentation for learning image segmentation with cross-modality annotations[J]. *Medical Image Analysis*, 2021, 71:102060.
- [7] VINCENT L, SOILLE P. Watersheds in digital spaces: an efficient algorithm based on immersion simulations[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1991, 13(6):583-598.
- [8] BOYKOV Y Y, JOLLY M P. Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images [C]. *Proceedings Eighth IEEE International Conference on Computer Vision*, 2001:5-112.
- [9] WANG Weixing, TIAN Liping, WANG Yue. Segmentation of cell images based on improved graph MST and skeleton distance mapping[J]. *Optics and Precision Engineering*, 2013, 21(9):2464-2472.
王卫星,田利平,王悦.基于改进的图论最小生成树及骨架距离直方图分割细胞图像[J].*光学精密工程*,2013,21(9):2464-2472.
- [10] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 39(4):640-651.
- [11] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation [C]. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Munich: MICCAI, 2015: 234-241.
- [12] ZHOU Z, SIDDIQUEE M M R, TAJBAKSHI N, et al. UNet++: a nested U-Net architecture for medical image segmentation [C]. *4th Deep Learning in Medical Image Analysis :DLMIA*, 2018:3-11.
- [13] CIRESAN D C, GIUSTI A, GAMBARDELLA L M, et al. Deep neural networks segment neuronal membranes in electron microscopy images [C]. *Proceedings of the 25th International Conference on Neural Information Processing Systems*, 2012, 2:2843-2851.

- [14] HUANG Hong, LV Rongfei, TAO Junli, et al. Segmentation of lung nodules in CT images using improved U-Net++ [J]. *Acta Photonica Sinica*, 2021, 50(2): 0210001.
黄鸿,吕容飞,陶俊利,等.基于改进U-Net++的CT影像肺结节分割算法[J].*光子学报*,2021,50(2): 0210001.
- [15] LI Q, HE X, WANG Y, et al. Review of spectral imaging technology in biomedical engineering: achievements and challenges[J]. *Journal of Biomedical Optics*, 2013, 18(10):100901.
- [16] ZHENG Shaojia, LI Qingli, ZHOU Mei, et al. Fourier transform channel attention network for cholangiocarcinoma hyperspectral image segmentation[J]. *Journal of Image and Graphics*, 2021, 26(8):1836-1846.
郑少佳,邱崧,李庆利,等.傅里叶变换通道注意力网络的胆管癌高光谱图像分割[J].*中国图象图形学报*,2021,26(8): 1836-1846.
- [17] WEI X, LI W, ZHANG M, et al. Medical hyperspectral image classification based on end-to-end fusion deep neural network[J]. *IEEE Transactions on Instrumentation & Measurement*, 2019, 68(11):4481-4492.
- [18] SONG Jie, XIAO Liang, LIAN Zhichao, et al. Overview and prospect of deep learning for image segmentation in digital pathology[J]. *Journal of Software*, 2021, 32(5): 1427-1460.
宋杰,肖亮,练智超,等.基于深度学习的数字病理图像分割综述与展望[J].*软件学报*,2021,32(5):1427-1460.
- [19] TAN M X, LE Q V. EfficientNet: rethinking model scaling for convolutional neural networks [C]. 36th International Conference on Machine Learning (ICML), California, 2019:691-700.
- [20] HU J, SHEN L, SUN G, et al. Squeeze-and-excitation networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(8): 2011-2023.
- [21] SEED: the challenge of intelligent diagnosis of cancer risk[EB/OL]. [2020-7-16]. <https://www.jseedata.com/>.
- [22] Brain of things: China big data & artificial intelligence innovation and entrepreneurship competition [EB/OL]. [2017-7-7]. <http://www.data dreams.org/>.
- [23] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017, 39(12):2481-2495.
- [24] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]. *European Conference on Computer Vision*. Springer, Cham,2018.
- [25] GU R, WANG G T, SONG T, et al. CA-Net: comprehensive attention convolutional neural networks for explainable medical image segmentation[J]. *IEEE Transactions on Medical Imaging*, 2020, 40(2):699-711.
- [26] CHEN J, LU Y, YU Q, et al. TransUNet: transformers make strong encoders for medical image segmentation [C]. *Computer Vision and Pattern Recognition (cs.CV)*, 2021.

Cancer Pathological Segmentation Network Based on Depth Feature Fusion

HUANG Hong¹, WANG Tao¹, LI Yuan¹, ZHOU Fanlin², LI Yu²

(1 Key Laboratory of Optoelectronic Technique System of the Ministry of Education, Chongqing University, Chongqing 400044, China)

(2 Department of Pathology, Chongqing University Cancer Hospital & Chongqing Cancer Institute & Chongqing Cancer Hospital, Chongqing 400030, China)

Abstract: At present, image segmentation technology is more and more widely used in the auxiliary diagnosis and treatment of cancer. However, mature and advanced image segmentation methods are mainly focused on natural images. Compared with the natural image, the content of pathological image is more complex, and there are great differences between different images. At the same time, the cancer cells and normal cells in the pathological image are mixed with each other, and there are great similarities between the two cells. These characteristics make many excellent natural image segmentation algorithms can not be directly applied to the segmentation task of pathological images and achieve good performance, so that artificial intelligence algorithms can not be quickly applied to medical auxiliary diagnosis and treatment. Therefore, more accurate segmentation of tumor pathological images is of great significance to promote clinical cancer diagnosis. Aiming at the problems of multiple slice staining, large resolution difference and complex image content in medical pathological images, an improved hierarchical feature fusion segmentation method is proposed. The method mainly includes four parts: encoder, decoder, channel attention module and loss function. Firstly, the method selects U-Net network as the basic network

structure. Efficientnet-b4 network is used to replace the original encoder of U-Net network for feature extraction, and the features are output from different layers as the output of the original encoder. The efficientnet-b4 network is transferred from natural images to pathological images for transfer learning, which effectively improves the ability of the network to extract effective features. In the decoder part, the feature fusion method is improved, and adds the hierarchical features of all layers before for feature fusion. Therefore, even the shallowest layer still contains the deepest global features. This method is used to gradually increase the role of global features in segmentation prediction from the deepest to the shallowest layer. It weakens the role of detail features in U-Net network. Therefore, the localization ability of the network on the main area of the lesion and the adaptability of the network to different resolution images are enhanced. At the same time, an improved channel attention module is used in each decoding layer, which is more suitable for pathological images than before. By adding global maximum pooling to extract channel features, more feature information is retained, so as to enhance the learning ability of the attention module and more effectively use the attention mechanism to filter the fused features. It highlights the effective features and suppresses the redundant features at the same time. In order to make the deep semantic information more distinguishable, the fusion characteristics of each depth layer are used to predict the output to construct a multi loss function. During training, the model performs prediction output at each decoding layer, and calculates the corresponding loss function for back-propagation training. Through this method, more effective deep semantic features are obtained for lesion location, which enhances the ability of the model to distinguish between normal and cancerous tissues, and improves the ability of the model to obtain and use global semantic features. Experiments are carried out on the BOT dataset and the seed dataset respectively. The Dice coefficient score of this method is 77.99% and 82.94% respectively, and the accuracy score is 88.52% and 87.42% respectively. Compared with U-Net and deeplabv3+, this method can effectively improve the segmentation accuracy and accuracy of tumor focus tissue, realize more accurate tumor location and segmentation in tumor pathological images, provide auxiliary support for doctors' clinical diagnosis more effectively, and improve the efficiency and accuracy of diagnosis and treatment. At the same time, ablation experiments are carried out on two data sets for the main improvements. The experimental results show the effectiveness of the improved method and can jointly promote the segmentation efficiency of HU-Net in pathological images from different aspects.

Key words: Computer image processing; Segmentation algorithm; Features fusion; Cancer; Pathology image; Deep learning

OCIS Codes: 100.4996; 100.4994; 170.4580; 100.2000