DOI:10.16016/j.2097-0927.202412075

基于DAEUnet和ConvNeXt网络的食管癌智能分割与分期诊断模型构建

熊铃焰[1],王润媛[2],张方红[1,3],杨 有[1,3],吴 毅[4],吴 蔚[5],武书磊[3] [1]重庆国家应用数学中心,重庆;[2]山西省肿瘤医院胸外科,山西太原;[3]重庆师范大学计算机与信息科学学院,重庆;陆军军医大学(第三军医大学):[4]生物医学工程与影像医学系数字医学教研室;[5]第一附属医院心外科,重庆

[**摘要**] **目的** 基于DAEUnet和ConvNeXt网络,通过迁移学习策略构建食管癌智能分割和T分期诊断模型。**方法** 收集2018年1月至2022年4月共126例确诊为食管癌患者Dicom原始数据,其中100例来源于陆军军医大学第一附属医院胸外科,26例来自山西省肿瘤医院胸外科,数据增强后共计60 275张,搭建DAEUnet食管癌智能分割网络,并在此基础上构建ConvNeXt、Swin Transfoemr以及ResNet 3个分类网络进行食管癌T分期诊断。**结果** 基于DAEUnet网络食管癌智能分割Dice相似系数(Dice similarity coefficient,DSC)为0.82,食管癌、主动脉、正常食管、纵隔淋巴结、心脏各个部位DSC分别为72.4%、87.5%、79.3%、60.5%、96.8%;在3个食管癌T分期诊断模型中,ConvNeXt模型效果最好,T1~T4期的精确度分别为0.65、0.727、0.889、0.92,曲线下面积(AUC)=0.892,优于ResNet和Swin Transformer网络。**结论** 基于DAEUnet和ConvNeXt网络的食管癌智能分割与T分期诊断模型,能够提高食管癌T分期的准确性与治疗效率。

[**关键词**] 食管癌;增强CT;智能分割;ConvNeXt;T分期诊断

[**中图法分类号**] R735.1;R445;TP391.7 [**文献标志码**] A

Intelligent segmentation and staging system for esophageal cancer based on DAEUnet and ConvNeXt networks

XIONG Lingyan[1], WANG Runyuan[2], ZHANG Fanghong[1,3], YANG You[1,3], WU Yi[4], WU Wei[5], WU Shulei[3] [1]Chongqing National Center for Applied Mathematics, Chongqing; [2]Department of Thoracic Surgery, Shanxi Cancer Hospital, Taiyuan, Shanxi; [3]School of Computer and Information Science, Chongqing Normal University, Chongqing; [4]Department of Digital Medicine, Faculty of Biomedical Engineering and Imaging Medicine; [5]Department of Cardiovascular Surgery, First Affiliated Hospital, Army Medical University (Third Military Medical University), Chongqing, China

[**Abstract**] Objective To construct an intelligent segmentation and T-stage diagnostic model for esophageal cancer based on the DAEUnet and ConvNeXt networks using transfer learning. Methods Dicom raw data from 126 patients diagnosed with esophageal cancer between January 2018 and April 2022 were collected, including 100 cases from Department of Thoracic Surgery at the First Affiliated Hospital of Army Medical University and 26 cases from the Department of Thoracic Surgery at Shanxi Cancer Hospital. After

[**基金项目**] 重庆市教委科学技术研究项目(KJZD202200504)

[**通信作者**] 张方红,E-mail:fanghong.zhang@cqnu.edu.cn

吴 毅,E-mail:wuy1979@tmmu.edu.cn

2097-0927/Copyright © 2025,陆军军医大学(第三军医大学)。这篇开放获取文章遵循CC BY许可协议(https://creativecommons.org/licenses/by/4.0/)

引用格式:熊铃焰,王润媛,张方红,等. 基于DAEUnet和ConvNeXt网络的食管癌智能分割与分期诊断模型构建[J]. 陆军军医大学学报,2025,47(10):1135-1144. DOI:10.16016/j.2097-0927.202412075.

Citation:XIONG L Y,WANG R Y,ZHANG F H,*et al*. Intelligent segmentation and staging system for esophageal cancer based on DAEUnet and ConvNeXt networks[J]. J Army Med Univ,2025,47(10):1135-1144. DOI:10.16016/j.2097-0927.202412075.

data augmentation, a total of 60 275 images were obtained. The DAEUnet esophageal cancer intelligent segmentation network was built, and on this basis, 3 classification networks, ConvNeXt, Swin Transformer, and ResNet were constructed for T-stage diagnosis of esophageal cancer. **Results** The Dice similarity coefficient (DSC) for esophageal cancer intelligent segmentation using the DAEUnet network was 0.82, and the DSC value of the esophagus, aorta, normal esophagus, mediastinal lymph nodes, and heart was 72.4%, 87.5%, 79.3%, 60.5% and 96.8%, respectively. Among the 3 T-stage diagnosis models for esophageal cancer, the ConvNeXt model performed the best, with a precision value for T1~T4 stages of 0.65, 0.727, 0.889 and 0.92, respectively, and an AUC value of 0.892, which were superior to the ResNet and Swin Transformer networks. **Conclusion** The proposed DAEUnet and ConvNeXt-based intelligent segmentation and T-stage diagnosis model for esophageal cancer improves T-stage accuracy and treatment efficiency.

[**Key words**] esophageal cancer; enhanced CT; intelligent segmentation; ConvNeXt; T-stage diagnosis

Supported by the Scientific Research Project of Chongqing Municipal Education Commission (KJZD202200504). Corresponding author: ZHANG Fanghong, E-mail: fanghong.zhang@cqnu.edu.cn; WU Yi, E-mail: wuy1979@tmmu.edu.cn

食管癌(esophageal cancers,EC)是一种常见的消化系统恶性肿瘤,发生率和死亡率高,严重危害人民的生存和生活质量,其早期诊断困难,病情发展快,患者就诊时大多已至晚期[1-2]。根据2020年世界卫生组织国际癌症研究机构发布的癌症负担数据,2020年全球食管癌死亡病例约为54.41万[3]。目前,食管癌主要的治疗手段为外科手术切除[4-5],患者早期直接进行食管癌根治性手术,将显著改善其预后和生存率,提高生活质量[6-8]。因此,应高度重视食管癌早期快速精准识别与诊断,以缩短诊疗时间,提升诊疗效率,有助于后期治疗决策和预后评估。

目前,食管癌T分期诊断主要依赖于医师通过影像学检查的经验判断,具有一定的主观性和局限性[9-10]。近年来,深度学习在医学图像分析领域的快速发展为食管癌智能诊断提供了新的技术路径。例如LIU等[11]构建了Inception-ResNet模型,对食管恶性肿瘤进行了分型诊断,模型的总体准确率可达到85.83%。EBIGBO等[12]创建了ResNet网络改进模型(基于斑块的BE分析的分类网络),在模型中加入了"一出一进"模型交叉验证,提高了模型的分类精度。RAMDHAN ABDALLA等[13]提出了一个在量子框架中使用深度学习和监督学习技术的混合量子-经典神经网络,以学习给定数据集标签并优化电路参数,最大限度减少误差,在食管癌分类方面显示出较高精准性。GUO等[14]设计了SegNet模型,可以实时识别食管癌病变部位,该模型加入了最大池化指数以提取食管癌的特征图,该模型有效提高了食管癌以及毗邻结构的分割精度。此外,WAN等[15]构建了Gastro Net模型,采用了Bootstrapping和集合学习算法的半监督学习方法,对巴雷特食管进行定位识别和诊断,该模型准确率可达到87.5%。尽管已有研究取得显著进展,但在实际临床应用中仍面临病灶边缘模糊导致的细分误差,以及微观浸润特征识别困难造成的分期偏差的挑战。

针对上述技术瓶颈,本研究构建DAEUnet食管癌智能分割网络,并结合ConvNeXt分类网络,提高食管癌分割精度和T分期准确性。DAEUnet网络通过双层变压器架构及多尺度特征学习,能够更好地捕捉全局上下文信息,在复杂的食管癌病灶分割任务中表现出较传统CNN方法更高的精度。另外,ConvNeXt模型优化卷积神经网络的特征提取过程,通过深度可分离卷积和窗口注意力机制[16],进一步提高食管癌T分期分类任务中的分类性能。基于这两种创新网络,本研究旨在为医师提供高效、精准的食管癌T分期诊断工具,提升诊断效率,辅助精准T分期,优化治疗决策。

1 资料与方法

1.1 临床资料

收集2018年1月至2022年4月共126例食管癌患者薄层CT图,其中100例来源于陆军军医大第一附属医院胸外科,26例来自山西省肿瘤医院胸外科,T1期24例,T2期40例,T3期50例,T4期12例。本研究已获得患者知情同意,并通过陆军军医大学第一附属医院和山西省肿瘤医院伦理委员会审批(KY2021165,2021050)。

纳入标准:①术后经组织病理学证实为食管鳞状细胞癌和食管腺癌;②食管癌手术前CT扫描为薄层增强CT图像且其图像的层厚<2 mm;③食管癌肿瘤区

域>5 mm的患者。

排除标准:①临床病理信息不完整或关键数据缺失(n=112);②图像质量差,如模糊或病灶区域被遮挡,导致无法辨识(n=156);③术前2周薄层CT扫描图像信息不全,特别是主动脉或淋巴结等区域缺失数据(n=22)。

1.2　图像标注数据集的构建

本实验将患者原始数据薄层CT图像以DICOM的格式上传至Amira软件(版本号6.0.0),由医师进行食管癌图像分割。通过调节窗宽窗位,提高图像对比度,从而更清晰地辨别肿瘤、淋巴结等微小结构的边界。所有CT图像均由低年资医师(具有1~2年临床工作经验)进行手动勾画,并由高年资医师(具有10年以上临床经验)进行复核修改。分割结构包括食管癌肿瘤及其毗邻的4个结构:心脏、主动脉、纵隔淋巴结、正常食管。

1.3　数据集预处理

1.3.1　数据样本不平衡处理　　为了解决食管癌数据中类别不平衡的问题,采用基于GAN的Pix2Pix模型对T1和T4阶段的CT影像进行数据增强[17]。由于T1(24例)和T4(12例)阶段样本量较少,这可能导致模型在这2个阶段的预测能力较弱,进而影响整体效果。通过Pix2Pix模型,调整输入图像的纹理、光照和细节等特征,从而生成不同风格的增强样本[18]。增强后,T1阶段的样本数量从24例增至36例,T4阶段从12例增至32例,整体数据集的样本量从126例增加到180例,提升了约43%。为评估数据增强的效果,使用结构相似性指数(SSIM)、峰值信噪比(PSNR)[19]和DSC等多个评价指标。结果显示,增强后的图像在结构和质量上与原始图像保持较高一致性(SSIM值约为0.84,PSNR值为33.5 dB),但通过调整纹理、光照和对比度等特征,成功模拟了不同的影像采集条件,增强了数据的多样性,从而提升了模型的泛化能力。在肿瘤区域分割任务中,使用增强数据训练后的模型DSC达到了0.82,显示出较高的准确性和鲁棒性。

1.3.2　实验数据去噪　　为了减少噪声对实验数据的影响,提高数据的质量和可靠性,对实验数据进行去噪操作。首先,使用平滑滤波器对数据进行平滑处理,减少数据中的高频噪声成分。接着,进行中值滤波处理有效去除数据中的脉冲噪声。此外,应用小波去噪技术,通过小波变换和阈值处理去除噪声信号。经过以上去噪操作,在一定程度上提高了实验数据的质量,为后续的分割与分类操作提供更清晰的图像。

1.3.3　实验数据增强　　为了增加训练数据的多样性,提高T分期模型的泛化能力和鲁棒性,对图像进行了数据增强[20],共生成82 750张增强图像。首先,通过垂直和水平翻转操作,增加图像的对称性变化,帮助模型识别不同角度的物体。其次,加入高斯噪声和高斯模糊,模拟真实场景中的噪声和模糊效果,提升模型对复杂环境的适应能力。最后,利用仿射变换和分段仿射变换,对图像进行了多样化的几何变换(如缩放、旋转、剪切和平移),增强了数据的多样性,提升了模型对不同尺度和形态的适应性,从而提高了模型的泛化能力和稳健性。

1.3.4　数据和标签归一化处理　　为提高T分期模型的训练稳定性和收敛速度,对实验图像数据进行了灰度归一化,将像素值缩放至[0, 1]范围内,确保图像间的可比性,利于模型学习图像特征。此外,采用one-hot编码将每个像素的标签转化为1个类别向量,向量长度等于类别数,标签位置为1,其余为0,避免数据分布不一致对模型性能造成的影响。经过数据样本不平衡处理、去噪、增强和归一化等步骤,完成数据预处

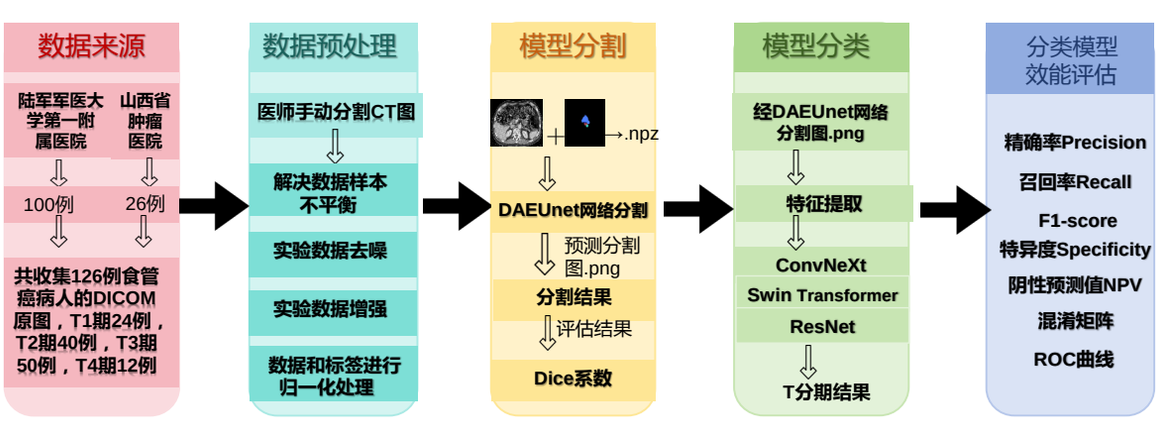


图1　基于CT图像的食管癌T分期智能分割与诊断模型构建流程

理。为便于理解整个模型构建过程,图1展示了基于CT图像的食管癌T分期智能分割与诊断模型的整体流程框架,包含从数据集收集到模型评估的各个核心步骤。

1.4 食管癌智能T分期模型的构建

由于数据集样本较小,本研究采用迁移学习与5折分层交叉验证相结合的方法。首先,通过迁移学习,将DAEUnet网络分割后的预测图与预训练模型的初始权重迁移到3个分类模型中进行训练,最终输出食管癌T分期诊断结果。图2为食管癌T分期诊断网络的构建框架,图中展示了基于食管癌CT数据集的分割网络编码器与解码器提取特征的过程、关键技术的处理方式、食管癌智能分割网络与分类模型的结合,以及T分期预测的核心流程。

1.4.1 5折分层交叉验证 针对数据小样本特点,对数据进行了5折分层交叉验证。在进行交叉验证之前,首先将数据集按8:2划分为训练集和测试集,其中训练集包含128例样本,测试集包含32例样本。测试集仅用于最终的模型评估,不参与训练过程,确保评估结果的有效性和公平性。

在训练集上,采用5折分层交叉验证。具体步骤如下:首先,将训练集中的数据根据T分期的类别分布划分为5个大致相等的子集,确保每个子集中的类别比例与训练集中的总类别分布相似。接着,在每次交叉验证迭代中,将其中的4个折用于模型训练,剩余的1个折用于模型验证。该过程重复5次,确保每个折至少1次作为验证集。最终,利用交叉验证的结果来评估模型在训练集上的性能,并通过独立的测试集评估模型的泛化能力。

1.4.2 食管癌智能分割网络搭建 本研究提出了一种创新性的食管癌智能分割网络——DAEUnet。为了验证该模型在食管癌智能分割任务中的有效性,将其与近年来较新的医学图像分割网络,如TransUnet、Swin-Unet和UNet等进行比较。DAEUnet食管癌智能分割网络采用双层变压器架构,能够高效提取目标区域特征。与传统卷积神经网络相比,DAEUnet通过引入变压器结构,使得网络能够更好地捕捉全局上下文信息,从而克服了传统方法在长距离依赖特征提取上的局限性。与传统卷积神经网络相比,DAEUnet分割网络通过引入变压器结构,使得网络能够更好地捕捉全局上下文信息,克服传统方法在长距离依赖特征提取上的局限性。

输入图像首先通过补丁嵌入模块进行处理,生成重叠补丁标记,经过4倍下采样后得到标记的输入图像。随后,图像通过包含3个编码器块的编码器模块进行处理,每个编码器块由2个连续双重变压器层和一个补丁合并层组成。在补丁合并阶段,2×2的补丁合并操作不仅有效减少了空间维度,还通过加倍通道维度,使网络能够以层次化方式学习多尺度特征,极大提升了对食管癌不同形态和尺度病灶的分割能力。

为了进一步提升分割精度,DAEUnet在特征融合方面引入了空间和通道注意力机制,使网络能够动态地选择最相关的特征进行融合,从而避免了信息冗余和特征丢失。解码器阶段通过扩展特征并利用跳跃连接的特征融合,有效恢复了高分辨率的细节信息。最后,经过2层双层变压器进一步特征融合,网络通过线性投影层输出最终的分割图,从而精准地完成食管癌的智能分割。

DAEUnet食管癌智能分割网络不仅结合了变压器架构和多尺度特征学习,还能够通过空间和通道注意力机制灵活地调节特征融合策略,在一定程度上提升了对食管癌及其毗邻结构的精准分割能力。这种精确的分割方法能够为后续的临床决策、手术规划和三维重建提供更高效的支持(图3)。

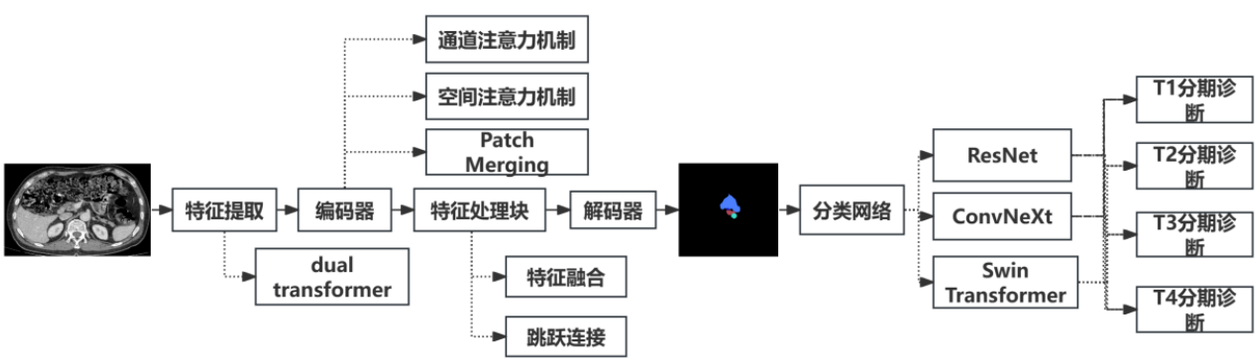


图2 食管癌T分期诊断网络的构建技术方案

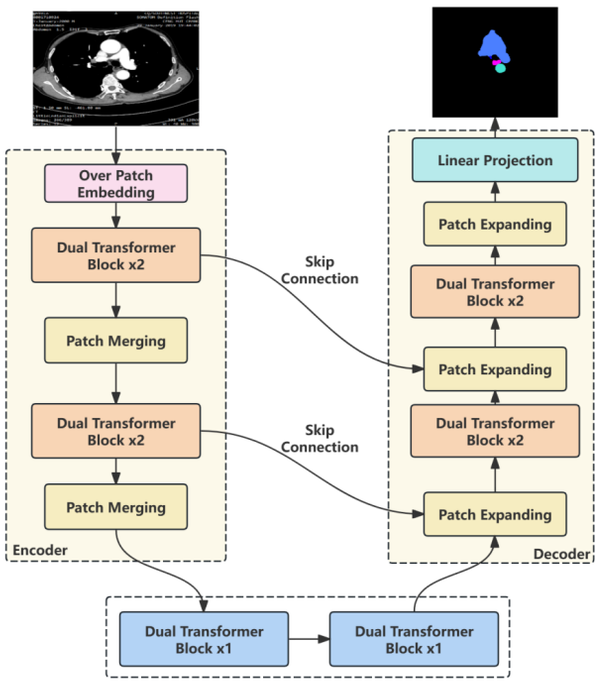


图3 DAEUnet食管癌智能分割网络模型结构

1.4.3 食管癌智能分期诊断网络搭建 在经DAEUnet网络智能分割后,分割结果与原始图像一同输入至Swin Transformer、ConvNeXt和ResNet 3个分类网络进行训练处理。数据首先进行预处理和数据增强等操作。此外,分割网络还自动计算食管癌的分割体积,并将其与分割图像一起输入3分类网络,用于食管癌T分期的精准诊断。

ConvNeXt是一种经过优化的卷积神经网络,其在食管癌分类任务中展现出独特的优势。与传统的CNN相比,ConvNeXt通过深度可分离卷积来优化特征提取过程,显著提升了计算效率。此外,卷积层分解为多个小型并行卷积层,使模型能够在不同维度上更精确地学习图像特征,同时减少了参数量,提高了速度与性能。通过引入组卷积,ConvNeXt进一步降低了计算复杂度,同时增强了网络处理高维度数据的能力。

ConvNeXt网络在食管癌的图像分类任务中具备强大的细节捕捉能力,通过增大卷积核和调整步长,ConvNeXt能够更好地捕捉图像中的微小细节与上下文信息,这对食管癌这种复杂病变的精准诊断至关重要。ConvNeXt还采用了独特的窗口注意力机制和偏移窗口注意力,能够有效地捕捉图像中的空间和上下文信息,提高分类准确性。

ConvNeXt网络的处理流程如下:输入的512×512×3的图像首先经过初始卷积层进行压缩,然后通过4个Stage逐步处理。在每个Stage中,包含下采样模块和多个Block模块。下采样模块通过2×2卷积核和步幅为2的操作减小特征图尺寸并提取更高级的特征。在每个Block中,深度卷积提取空间特征,并通过1×1卷积融合通道信息,再通过DELU激活函数进行非线性变换,最后进行分类以得到最终结果。ConvNeXt网络的具体结构见图4。

1.5 模型效能评估

为评估模型的泛化能力,对数据增强后的160例数据集进行了训练集(128例)与测试集(32例)的划分。测试集仅用于最终的独立评估,确保在训练过程中完全未参与任何形式的训练或验证,避免数据泄漏问题。在模型训练过程中,使用5折分层交叉验证法,将训练集分为5个折,每次交叉验证时,4个折用于训练,剩余的1个折用于验证。这个过程重复5次,确保每个折都至少1次作为验证集,最终将各次验证结果的分类结果取平均值,评估模型在训练集上的性能。最终的测试集仅用于评估模型的泛化能力,从而确保模型效能评估的有效性和公平性。

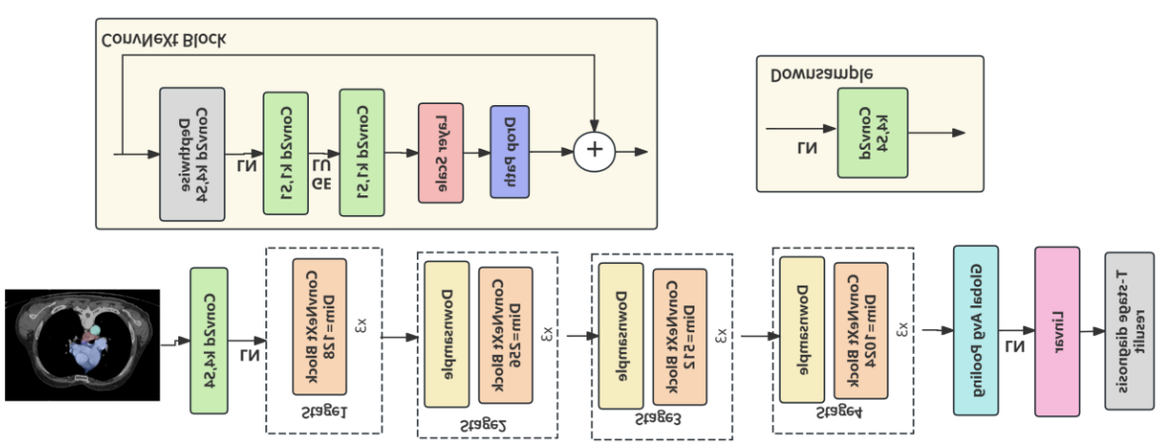


图4 ConvNeXt网络结构

表1 患者一般资料及其临床病理特征的Mann-Whitney检验结果

变量	中位数$M(P25,P75)$		Mann Whitney检验统计量		P
	训练集(n=101)	测试集(n=25)	U值	Z值	
粗糙程度评分	2.00(2.0,3.0)	2.00(2.0,3.0)	763.00	−0.787	0.431
肿瘤位置	4.00(3.0,5.0)	4.00(3.0,5.0)	845.50	−0.043	0.966
性别	1.00(1.0,1.0)	1.00(1.0,1.0)	733.50	−1.537	0.124
神经血管侵犯	1.00(1.0,1.0)	1.00(1.0,1.0)	834.00	−0.195	0.845
分化程度	2.00(2.0,2.0)	2.00(1.0,2.0)	694.50	−1.565	0.118
N分期	0.00(0.0,1.0)	0.00(0.0,1.0)	730.00	−1.168	0.243

模型智能分割评估采用Dice相似系数用于评估分割结果的质量。DSC通过比较预测分割和真实分割之间的重叠程度来衡量分割的准确性。

模型分类采用精确率[21]、召回率[22]、特异性[22]、召回率的调和均值(F1 score)[23]、阴性预测值(NPV)[23]以及曲线下面积(AUC)[24]来进行分类(分期诊断)效能评估。最后将模型分类与医师分类结果进行对比,由混淆矩阵[25]、ROC曲线图[25]等对模型的分类结果进行展示。

1.6 统计学分析

采用SPSS25.0软件进行统计学分析。组间差异采用Mann-Whitney检验,P<0.05表示差异具有统计学意义。

2 结果

2.1 患者基本信息与数据划分

本研究共纳入126例食管癌患者。数据集按8∶2的比例进行简单随机划分为训练集和测试集。结果显示,训练集和测试集的临床病理特征(包括性别、粗糙度评分等)没有统计学差异,说明两组具有均衡性。患者一般资料及其临床病理特征见表1。

2.2 基于DAEUnet和其他网络的食管癌及毗邻结构智能分割对比

本研究构建了DAEUnet、TransUnet、Swin-Unet和UNet 4种食管癌及其毗邻结构智能分割网络(表2)。其中,DAEUnet在所有结构的平均DSC(0.82)优于TransUnet、Swin-Unet和UNet,尤其在分割食管癌及其毗邻结构方面效果最佳。与其他方法相比,DAEUnet的DSC值分别提高了4.6%、1.95%和11.27%。食管癌、正常食管和纵隔淋巴结的智能分割DSC分别为72.4、79.3和60.5,均优于其他3种网络,且分割效果最接近于真实标签,有助于食管肿瘤和转移淋巴结轮廓信息的快速提取,从而有助于后期食管癌的快速T和N分期。

表2 不同模型在食管癌数据集上的结果比较

模型	食管癌	主动脉	正常食管	纵隔淋巴结	心脏	平均DSC
U_net	52.2	78.4	57.3	38.2	93.7	70.81
Transunet	66.7	87.6	74.8	52.4	96.7	78.48
Swin-Unet	71.0	86.2	78.9	52.3	96.8	80.13
DAEUnet	72.4	87.5	79.3	60.5	96.8	82.08

为了更直观地比较DAEUnet与手动分割的分割效果,本研究将DAEUnet智能分割和真实标签的食管癌、纵隔淋巴结及毗邻结构在Arima软件进行三维重建。结果表明,DAEUnet与手动分割在三维重建效果上差距不大,其分割效果已接近真实标签,适用于基于CT图像的食管癌及其毗邻结构的快速三维重建、可视化及量化。这对于评估肿瘤侵袭性及淋巴结转移程度至关重要,有助于食管癌的快速精准诊断与治疗决策(图5)。

2.3 食管癌的智能分期诊断结果

本研究基于DAEUnet食管癌分割结果,搭建了ConvNeXt、Swin Transformer和ResNet 3种智能T分期诊断模型,使用5折交叉验证评估其性能,计算精确率、召回率、特异性、召回率的调和均值和准确率,并对结果进行综合评估。各个模型的性能如表3所示。ConvNeXt网络对EC智能分期诊断AUC值为0.892,T1~T4准确率分别为0.65、0.727、0.889、0.92;Swin Transformer网络AUC值为0.857,T1~T4期准确率分别为0.62、0.667、0.882、0.89;ResNet网络AUC最低,为0.620,T1~T4期的准确性分别为0.32、0.565、0.603、0.626。结果显示,ConvNeXt网络在T分期的准确性上优于ResNet和Swin Transformer。

通过分析3个分类网络的食管癌T分期结果(表3)、混淆矩阵和ROC曲线(图6),可以得到ConvNeXt网络在所有T分期中表现最好,且3个网络分类准确性随着T分期的升高有不同幅度的提升。ConvNeXt网络

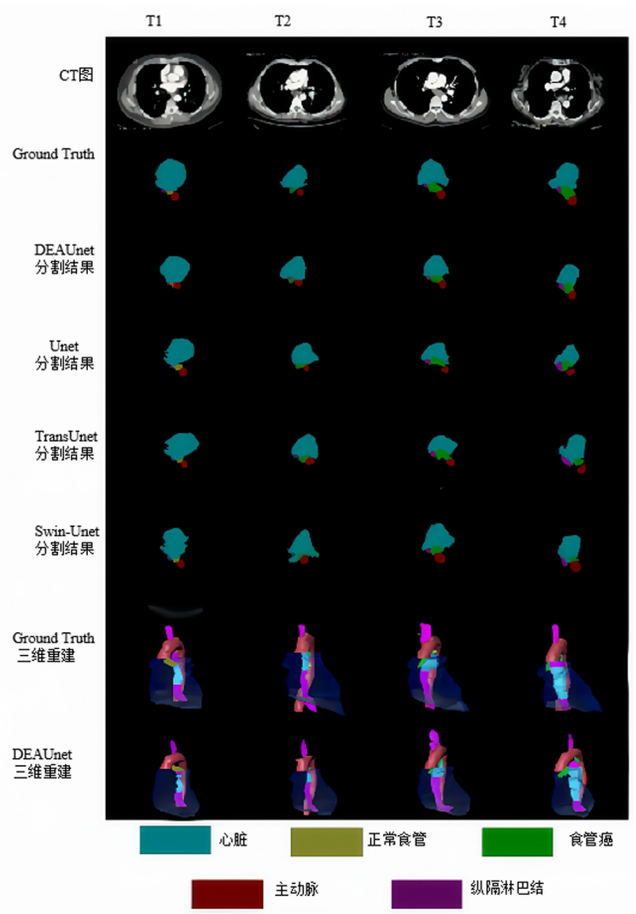


图5 食管癌不同T分期在4个网络的智能分割对比

表3 3个分类网络食管癌T分期结果

不同网络	T分期	精确度	召回率	特异性	NPV	F1-score	AUC
Swin Transformer	T1	0.62	0.84	0.84	0.82	0.78	0.857
	T2	0.667	0.667	0.867	0.76	0.74	
	T3	0.882	0.889	0.629	0.752	0.885	
	T4	0.89	0.921	0.89	0.982	0.84	
ResNet	T1	0.32	0.12	0.693	0.623	0.283	0.620
	T2	0.565	0.138	0.757	0.750	0.249	
	T3	0.603	0.736	0.232	0.545	0.650	
	T4	0.626	0.657	0.778	0.737	0.747	
ConvNeXt	T1	0.65	0.79	0.89	0.854	0.682	0.892
	T2	0.727	0.72	0.882	0.802	0.722	
	T3	0.889	0.90	0.742	0.54	0.915	
	T4	0.92	0.89	0.962	0.989	0.925	

在T1和T4期的表现最佳,其T4期的精确度和召回率分别为0.65和0.92,F1-score达到0.925。尽管在T2和T3期的表现稍有下降,但整体准确率仍然最高。Swin Transformer网络的精确度和召回率从T1期的

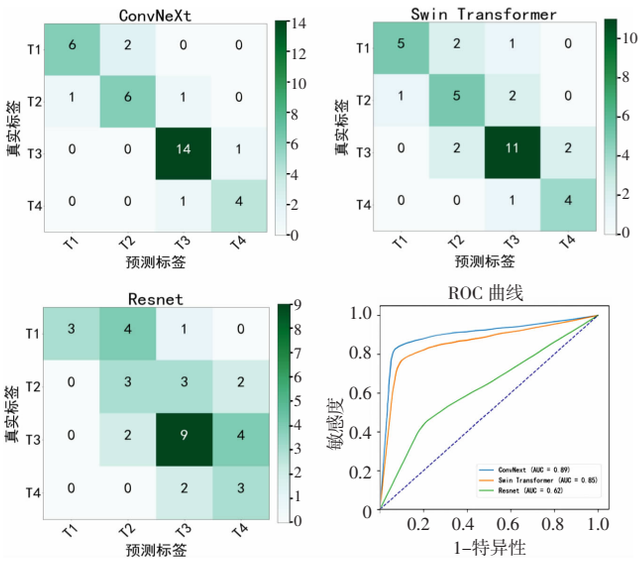


图6 3个网络模型的混淆矩阵和ROC曲线

0.62和0.84提升到T4期的0.89和0.921,F1-score也从0.78上升至0.84,呈现出逐步提升的趋势。ResNet网络的精确度和召回率在T1期起步较低,但在T4期表现相对较好,整体呈现稳定上升的趋势。

从这些结果来看,ConvNeXt网络在所有T分期中的分类准确度最高,表明其对食管癌T分期的诊断效能最佳。进一步分析3个分类模型的混淆矩阵,可以发现ConvNeXt模型的对角线元素明显大于其他2个模型,表明其正确预测的样本更多,因此在T分类上的准确率更高。此外,ConvNeXt模型的非对角线元素数值也明显少于其他模型,表示其误分类情况较少。误分类主要发生在T1和T2期,而T3和T4期几乎没有误判。

3 讨论

3.1 基于CT图像食管癌及其毗邻结构的智能分割

精准分割食管癌及其毗邻结构,能够为病灶量化和三维重建提供基础。分割网络模型实验结果显示,DAEUnet网络在多个关键结构的分割精度上表现优异,分割结果接近标签真值图像。在食管癌病灶的分割上,DAEUnet的DSC系数为72.4%,显著高于其他模型,尤其比UNet高出20.2%,比TransUnet高出5.7%,比Swin-Unet高出1.4%。此外,在淋巴结区域,DAEUnet的DSC为60.5%,大幅超越其他模型,相比UNet提高了22.3%,比TransUnet高8.1%,比Swin-Unet高8.2%,为肿瘤的侵袭性评估和周围组织关系的分析提供了更加精准的数据支持。在主动脉和心脏结构的智能

分割上,DAEUnet网络DSC分别为87.5%和96.8%,相比TransUnet和Swin-Unet表现突出。尤其在心脏区域,尽管与其他模型的差距较小,但双层变压器架构和多尺度特征学习使DAEUnet能够更精准地捕捉复杂区域的细节,从而提升分割精度。相比之下,UNet在处理复杂边界和细小结构时精度较低,仅为52.2%,主要受限于其简单的网络结构,缺乏细节捕捉能力,导致在食管癌及其周围复杂结构的分割上表现不佳。总体而言,DAEUnet通过融合全局上下文和细节特征,大幅提升了分割精度,尤其在复杂病灶的分割中展现了明显优势,具有较高的应用价值。

与先前的研究相比,GUO等[14]设计并实现的SegNet模型分割准确率为82.955%,与本研究的分割准确度相近,但该研究并未对其毗邻结构进行分割,仅展示了食管癌病灶,未体现其空间毗邻结构。WAN等[15]采用2DU-Net、3DV-Net和VUMix-Net对215名食管癌患者的肿瘤进行自动识别,发现VUMix-Net的DSC为0.68,稍高于单一网络模型。相比之下,本研究构建的模型在食管癌的DSC为0.82,显著高于上述模型。因此,本研究所提出的分割模型为食管癌T分期诊断提供了高效的技术支持,有助于食管癌及其毗邻结构的三维可视化展示,精准地辅助临床医师制定个性化的治疗方案。

3.2 基于CT图像食管癌的智能分期诊断

传统影像或胸外科医师难以通过CT图像准确判断食管癌的T分期,本研究的智能T分期诊断模型通过自动分割食管癌及其毗邻结构,实现了精准的T分期诊断,帮助医师更好地评估病情并支持治疗决策。

在T分期诊断结果方面,基于DAEUnet智能分割网络和ConvNeXt分类网络食管癌T分期诊断模型AUC值为0.892,T1~T4期分期准确性分别为0.65、0.727、0.889、0.92,均优于ResNet和Swin Transformer网络的准确性。本研究的深度学习模型证明了CT图像用于食管癌T分期诊断的可行性。尽管当前食管癌T分期诊断主要依赖病理结果,但病理样本获取需要通过胃镜,且存在一定的创伤风险、继发感染及穿刺道种植的风险,因此无创的CT图像T分期智能诊断方法具有重要的临床价值[28]。

与之前的研究相比,本研究模型在食管癌T分期诊断上更具优势。例如,LIU等[11]提出的Inception-ResNet模型仅进行了食管肿瘤的良恶性二分类,未涉及细致的T分期诊断,而本研究提供了更为精准的T分期诊断,有助于做出快速个性化的治疗决策。EBIGBO等[12]提出DeepLabV3+模型于早期食管癌智能检测,其模型的recall为77.58%,精确率为78.06%,而本研究模型的精确率为89.2%,远高于其模型。此外,RAMDHAN ABDALLA等[13]使用经典和量子成分对食管癌进行分类,精确度达到78.06%,但其研究未进行智能分割,也未实现三维可视化展示食管癌及其毗邻结构。本研究采用先分割再分类的方法,既提高了T分期诊断的准确性,也为食管癌的三维可视化和后期手术规划提供了支持。

3.3 不足之处

本研究模型仍存在一定局限性:(1)数据样本量不足,导致分割和分类精度有待提升。计划在后续工作中扩展数据集,构建食管癌的大数据和多中心数据库,以进一步提高模型的分割精度和泛化能力。(2)T1和T4期病例数据较少,数据分布不平衡。本研究将在下一阶段增加T1和T4期病例的数量,力求数据分布的均衡。(3)T1期的智能诊断效果较差,可能由于T1期肿瘤体积较小,且与周围组织的差异不明显,导致分割和分类效果不佳。为此,本研究计划改进分割和分类网络,提升小样本及边界模糊病例的分割效率,以提高T1期分期诊断的准确性。

综上所述,本研究基于DAEUnet构建了食管癌智能分割模型,克服了长距离依赖特征提取的局限,实现食管癌、阳性淋巴结、正常食管、主动脉和心脏的快速智能分割及体积量化。模型通过精确地多区域分割,显著提高了食管癌及相关结构的识别能力,帮助医师更高效地获取肿瘤及相关组织的准确信息,从而优化治疗方案。此外,基于ConvNeXt架构,研究进一步开发了食管癌智能T分期诊断模型,实现了T1~T4期的快速智能分期诊断,减少了诊断中的人为误差和主观判断,提高了分期的准确性。本研究不仅减轻了医师的工作负担,提高了工作效率,还通过自动化和智能化手段,节省了大量医疗资源,特别是在医疗资源匮乏地区,推动智能医疗建设。

利益冲突 本文所有作者声明不存在利益冲突

参考文献:

[1] MA D, ZHOU T L, CHEN J, *et al*. Radiomics diagnostic performance for predicting lymph node metastasis in esophageal cancer: a systematic review and meta-analysis [J]. BMC Med Imaging, 2024, 24(1): 144. DOI: 10.1186/s12880-024-01278-5.

[2] 王润媛.基于深度学习的食管癌智能分割与分期诊断[D].太原:山西医科大学,2023. DOI:10.27288/d.cnki.gsxyu.2023.001615.
WANG R Y. Intelligent segmentation and staging diagnosis of esophageal cancer based on deep learning[D]. Taiyuan: Shanxi Medical University, 2023. DOI: 10.27288/d. cnki. gsxyu.2023.001615.

[3] TALUKDAR F R, DI PIETRO M, SECRIER M, *et al*. Molecular landscape of esophageal cancer: implications for early detection and personalized therapy[J]. Ann N Y Acad Sci, 2018, 1434(1): 342-359. DOI: 10.1111/nyas.13876.

[4] CAI Z J, LIU Q. Understanding the global cancer statistics 2018: implications for cancer control[J]. Sci China Life Sci, 2021, 64(6): 1017-1020. DOI: 10.1007/s11427-019-9816-1.

[5] SAEKI H, SOHDA M, SAKAI M, *et al*. Role of surgery in multidisciplinary treatment strategies for locally advanced esophageal squamous cell carcinoma[J]. Annals Gastroent Surgery, 2020, 4(5): 490-497. DOI: 10.1002/ags3.12364.

[6] 薛龙梅,赵学武,贾霞,等.MSCT对食管癌术前TNM分期的诊断及临床指导价值分析[J].中国CT和MRI杂志,2022,20(9):102-103,123. DOI:10.3969/j.issn.1672-5131.2022.09.038.
XUE L M, ZHAO X W, JIA X, *et al*. Analysis on value of MSCT in the diagnosis and clinical guidance of preoperative TNM staging of esophageal cancer[J]. Chin J CT MRI, 2022, 20(9): 102-103, 123. DOI: 10.3969/j.issn.1672-5131.2022.09.038.

[7] YANG M W, HU P P, LI M L, *et al*. Computed tomography-based radiomics in predicting T stage and length of esophageal squamous cell carcinoma[J]. Front Oncol, 2021, 11: 722961. DOI: 10.3389/fonc.2021.722961.

[8] WU A J, BOSCH W R, CHANG D T, *et al*. Expert consensus contouring guidelines for intensity modulated radiation therapy in esophageal and gastroesophageal junction cancer[J]. Int J Radiat Oncol Biol Phys, 2015, 92(4): 911-920. DOI: 10.1016/j.ijrobp.2015.03.030.

[9] OLD O J, ISABELLE M, BARR H. Staging early esophageal cancer[J]. Adv Exp Med Biol, 2016, 908: 161-181. DOI: 10.1007/978-3-319-41388-4_9.

[10] 王润媛,陈星材,吴蔚,等.基于增强CT图像和Swin Transformer网络的食管癌T分期智能诊断模型的构建与评估[J].陆军军医大学学报,2023,45(16):1770-1778. DOI:10.16016/j.2097-0927.202302021.
WANG R Y, CHEN X C, WU W, *et al*. Construction and evaluation of an intelligent diagnostic model based on enhanced CT images and Swin Transformer network for T staging of esophageal cancer[J]. J Army Med Univ, 2023, 45(16): 1770-1778. DOI: 10.16016/j.2097-0927.202302021.

[11] LIU G S, HUA J, WU Z, *et al*. Automatic classification of esophageal lesions in endoscopic images using a convolutional neural network[J]. Ann Transl Med, 2020, 8(7): 486. DOI: 10.21037/atm.2020.03.24.

[12] EBIGBO A, MENDEL R, PROBST A, *et al*. Computer-aided diagnosis using deep learning in the evaluation of early oesophageal adenocarcinoma[J]. Gut, 2019, 68(7): 1143-1145. DOI: 10.1136/gutjnl-2018-317573.

[13] RAMDHAN ABDALLA W A. Classifying esophageal cancer using a hybrid quantum-classical neural network[C]//2023 IEEE 3rd International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering (MI-STA). IEEE, 2023: 229-234. DOI: 10.1109/MI-STA57575.2023.10169293.

[14] GUO L J, XIAO X, WU C C, *et al*. Real-time automated diagnosis of precancerous lesions and early esophageal squamous cell carcinoma using a deep learning model (with videos)[J]. Gastrointest Endosc, 2020, 91(1): 41-51. DOI: 10.1016/j.gie.2019.08.018.

[15] WAN J, ZENG Y K. Prediction of hepatic metastasis in esophageal cancer based on machine learning[J]. Sci Rep, 2024, 14(1): 14507. DOI: 10.1038/s41598-024-63213-6.

[16] LIU Z, MAO H Z, WU C Y, *et al*. A ConvNet for the 2020s[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2022: 11966-11976. DOI: 10.1109/CVPR52688.2022.01167.

[17] LI M, WANG J P, CHEN Y, *et al*. Low-dose CT image synthesis for domain adaptation imaging using a generative adversarial network with noise encoding transfer learning[J]. IEEE Trans Med Imag, 2023, 42(9): 2616-2630. DOI: 10.1109/TMI.2023.3261822.

[18] CHEN Y, LIN H P, ZHANG W, *et al*. ICycle-GAN: Improved cycle generative adversarial networks for liver medical image generation[J]. Biomed Signal Process Contr, 2024, 92: 106100. DOI: 10.1016/j.bspc.2024.106100.

[19] AKTER S. Generative AI: a Pix2pix-GAN-based machine learning approach for robust and efficient lung segmentation[EB/OL]. 2024: 2412.10826. https://arxiv.org/abs/2412.10826v1.

[20] 马岩昇,唐娉,赵理君,等.深度学习图像数据增广方法研究综述[J].中国图象图形学报,2021,26(3):487-502.
MA D A, TANG P, ZHAO L J, *et al*. Review of data augmentation for image in deep learning[J]. J Image Graph, 2021, 26(3): 487-502.

[21] LIU W F, WANG S F, YE Z H, *et al*. Prediction of lung metastases in thyroid cancer using machine learning based on SEER database[J]. Cancer Med, 2022, 11(12): 2503-2515. DOI: 10.1002/cam4.4617.

[22] GRANADOS SÁNCHEZ A M, OREJUELA ZAPATA J F. Diagnosis of mesial temporal sclerosis: sensitivity, specificity and predictive values of the quantitative analysis of magnetic resonance imaging[J]. Neuroradiol J, 2018, 31(1): 50-59. DOI: 10.1177/1971400917731301.

[23] KIM S C, CHO Y S. Predictive system implementation to improve the accuracy of urine self-diagnosis with smartphones: application of a confusion matrix-based learning model through RGB semiquantitative analysis[J]. Sensors, 2022, 22(14): 5445. DOI: 10.3390/s22145445.

[24] CHENG Y, LI X C, DAI Y B, *et al*. Identification of an immune-related risk signature and nomogram predicting the overall survival in patients with endometrial cancer[J]. J Gynecol Oncol, 2021, 32(3): e30. DOI: 10.3802/jgo.2021.32.e30.

[25] ZHANG T M, CHEN Z, WANG Z Z, *et al*. Esophagogastroscopy for predicting endoscopic ultrasonography T-stage by utilizing deep learning methods in esophageal cancer[J]. Appl Intell, 2024, 54(19): 9286-9294. DOI: 10.1007/s10489-024-05640-6.

[26] GHATWARY N, YE X J, ZOLGHARNI M. Esophageal abnormality detection using DenseNet based faster R-CNN with Gabor features[J]. IEEE Access, 2019, 7: 84374-84385. DOI: 10.1109/ACCESS.2019.2925585.

[27] DU K P, LI L X, WANG Q, *et al*. Development and application of a dynamic prediction model for esophageal cancer[J]. Ann Transl Med, 2021, 9(20): 1546. DOI: 10.21037/atm-21-4964.

[28] ZHAO X J, CHEN K X, WU G, *et al*. Deep learning shows good reliability for automatic segmentation and volume measurement of brain hemorrhage, intraventricular extension, and peripheral edema[J]. Eur Radiol, 2021, 31(7): 5012-5020. DOI: 10.1007/s00330-020-07558-2.

(收稿:2024-12-13;修回:2025-03-07)

(编辑 张 维)