

马露, 丁晓阳, 陈俊豪, 等. 基于密度估计的非小细胞肺癌全身代谢肿瘤体积估计[J]. 智能计算机与应用, 2025, 15(4): 86-93. DOI: 10.20169/j.issn.2095-2163.24101104

基于密度估计的非小细胞肺癌全身代谢肿瘤体积估计

马露¹, 丁晓阳², 陈俊豪¹, 祁婧¹, 刘秀婷¹

(1 兰州财经大学 信息工程与人工智能学院, 兰州 730030; 2 甘肃农业大学 信息科学技术学院, 兰州 730070)

摘要: 全身代谢肿瘤体积 (whole-body Metabolic Tumor Volume, MTV_{WB}) 被证明是非小细胞肺癌患者独立的预后指标, 与患者的总生存期相关。计算 MTV_{WB} 需医生手动分割肿瘤, 这一过程既耗时又费力, 且现有 MTV_{WB} 的自动估计方法由于全身范围过大而性能欠佳。因此, 提出了基于密度估计的两阶段方法, 用于估计非小细胞肺癌的 MTV_{WB} 。第一阶段采用 nnSAM 模型对全身肿瘤进行分割, 找到含有肿瘤的影像切片; 第二阶段将人群计数中的密度估计方法应用在医学图像中, 采用 TransU-Net 网络通过密度估计的方法对第一阶段筛选得到的切片进行估计得到 MTV_{WB} 。该方法在芝加哥大学 480 例非小细胞肺癌患者的 PET/CT 数据上进行了测试, R^2_{score} 达到了 0.71, 性能超过了现有研究方法。实验表明提出的方法能够有效预测 MTV_{WB} , 辅助医生制定治疗方案、预测病人预后。

关键词: 全身代谢肿瘤体积; 密度估计; 医学图像分割; 非小细胞肺癌; PET/CT

中图分类号: TP391

文献标志码: A

文章编号: 2095-2163(2025)04-0086-08

Whole-body Metabolic Tumor Volume estimation of non-small cell lung cancer based on density estimation

MA Lu¹, DING Xiaoyang², CHEN Junhao¹, QI Jing¹, LIU Xiuting¹

(1 College of Information Engineering and Artificial Intelligence, Lanzhou University of Finance and Economics, Lanzhou 730030, China; 2 College of Information Science and Technology, Gansu Agricultural University, Lanzhou 730070, China)

Abstract: Whole-body Metabolic Tumor Volume (MTV_{WB}) has been proven to be an independent prognostic indicator for patients with non-small cell lung cancer, which is associated with the overall survival of patients. The calculation of MTV_{WB} requires doctors to manually segment the tumor, which is both time-consuming and laborious. In addition, the existing automatic estimation methods for MTV_{WB} perform poorly due to the large scope of the whole body. Therefore, a two-stage method based on density estimation is proposed to estimate MTV_{WB} in non-small cell lung cancer. In the first stage, the nnSAM model is used to segment the whole-body tumor and find the image slices containing the tumor. In the second stage, the tumor area is regarded as a dense crowd area, and the density estimation method in crowd counting is applied to medical images. The TransUNet network is used to estimate the MTV_{WB} of the slices containing the tumor obtained from the first stage through density estimation. This method is tested on PET/CT data of 480 patients with non-small cell lung cancer from the University of Chicago, with an R^2_{score} of 0.71, outperforming existing research methods. The experiments show that the proposed method can effectively predict MTV_{WB} , assist doctors in formulating treatment plans, and predict patient prognosis.

Key words: whole-body Metabolic Tumor Volume; density estimation; medical image segmentation; non-small cell lung cancer; PET/CT

0 引言

肺癌是一种恶性肿瘤, 其源自肺组织内的异常细胞增殖, 是全球癌症导致相关死亡最常见的原因之一^[1]。肺癌可以分为两大类: 小细胞肺癌

(Small Cell Lung Cancer, SCLC) 和非小细胞肺癌 (Non-Small Cell Lung Cancer, NSCLC), 其中 NSCLC 是肺癌的主要类型, 占有肺癌病例的 80~85%^[2]。

全身代谢肿瘤体积作为 18F-FDG-PET/CT 代谢性肿瘤负荷指标是非小细胞肺癌患者独立的预后

作者简介: 马露 (2001—), 女, 硕士研究生, 主要研究方向: 医学图像处理。

通信作者: 丁晓阳 (1969—), 男, 教授, 主要研究方向: 软件工程, 计算机视觉。Email: 1033592996@qq.com。

收稿日期: 2024-10-11

哈尔滨工业大学主办 ◆ 学术研究与应用

指标^[3],与非小细胞肺癌患者的总生存时间(Overall Survival, OS)相关^[4]。准确估计全身代谢肿瘤体积对于个体化治疗和生存预测至关重要。然而,在实际临床应用中,计算全身代谢肿瘤体积需要医生手动勾勒肿瘤区域,这一过程既耗时又繁重。因此,研究一种能够自动计算全身代谢肿瘤体积的方法在实际临床应用中具有重要意义。

对于全身代谢肿瘤体积的现有估计方法通常是先对全身肿瘤进行分割,然后计算代谢肿瘤体积。然而肿瘤分割需要从全身寻找肿瘤,搜索范围大,且PET图像中肿瘤边界模糊、假阳性区域也较多,因此采用分割的方法存在一定的困难。故实现对全身代谢肿瘤体积的精准估计是一项具有挑战性的任务。

针对上述问题,本文提出了一种基于密度估计的两阶段全身代谢肿瘤体积的自动估计方法。本文的贡献如下:

(1)提出了通过密度估计得到 MTV_{WB} 的方法,把肿瘤区域视为密集的人群区域,将自然图像中人群计数的方法借鉴到了医学图像领域之中。

(2)提出了两阶段方法估计 MTV_{WB} 。在第一阶段通过分割任务筛选出含有肿瘤的影像切片,解决了由于全身范围较大、信息复杂带来的难题,有助于在第二阶段对 MTV_{WB} 进行估计。

(3)在芝加哥大学480例非小细胞肺癌患者的PET/CT数据集上进行了实验,验证了该两阶段方法的有效性,性能超过了现有的研究方法。

1 相关工作

1.1 肿瘤分割

高效、准确的医学图像分割是疾病诊断、预后、治疗计划与监测、治疗结果随访等现代临床工作流程的重要组成部分^[5]。卷积神经网络已经成为医学图像分割的主流,虽然具有很强的特征提取能力,但却只关注局部信息,无法提取到全局信息。所以对全局信息有关关注能力的Transformer在医学图像分割方面得到了广泛的应用。

现如今,基于CNN和Transformer研究开发了各种分割网络架构。如:TransUNet^[6]融合了U-Net^[7]和Transformer^[8]的优点,定义了医学图像分割的新方向。SwinUNet^[9]引入了另一种transformer驱动的医学图像分割方法,利用U形编码器-解码器架构和跳过连接来增强局部-全局语义特征学习。Manzari等学者^[10]提出了一种创新的U型网络,称为BEFU-Net,将Transformer和CNN结合在一起,

BEFU-Net增强了身体和边缘信息的融合,以实现精确的医学图像分割。Zhang等学者^[11]提出了一个结合传播和转换器架构的分割网络,该网络只需要少量的标记数据,同时利用基本未标记的图像即可实现医学图像分割。

然而,许多分割工作需要人为地进行架构修改和超参数调优,以适应不同的数据集。基于这一挑战,nnUNet^[12]框架被提出,nnUNet通过适当的预处理和后处理组合,即使是基础的网络架构也可以在各种医疗分割任务中实现最先进的性能。虽然nnUNet提升了医学图像分割项目方法的灵活性,但这类分割网络的性能仍然依赖于充足的训练数据。Segment Anything Model (SAM)^[13]是一个在1 100万张图像和超过10亿个分割掩码(SA-1B训练数据集)上训练过的模型,模型出现促进了少标签数据和零标签数据的发展。但将SAM直接应用于医学图像时,其实实验精度会受到影响且SAM需要提示作为输入,这就使其在临床应用中受到限制。因此,在受nnUNet和SAM的优缺点启发后,一种新的即插即用的解决方案nnSAM^[14]被提出,旨在提高医学图像的分割精度。nnSAM将SAM强大的特征提取和泛化能力与nnUNet以数据为中心的自动配置能力相结合。如今,在借助大模型实现医学图像分割研究中已取得不错的成果,如:Zhang等学者^[15]利用SAM生成的掩码、特征和稳定性评分来增强常用医学图像分割模型的图像输入。本文在第一阶段利用nnSAM进行肿瘤分割,再利用输出结果挑选出含有肿瘤的图像切片,降低了全身信息的复杂度。

1.2 人群计数

随着城市化的加速和人口密度的增加,人群管理成为了一个倍受关注的问题,人群计数已经成为了当下的热点研究^[16]。关于人群计数的方法在现阶段已经非常成熟,目前已成功应用于安防、大型活动、大型比赛的场景。传统的人群计数方法主要是基于回归和检测,如Chan等学者^[17]提出了一种基于回归的方法,通过先分离背景,再从前景中提取特征来预测人数。Lin等学者^[18]提出了一种基于检测的方法,该方法使用2个连续的视频帧序列来预测人数。

随着深度学习的发展和应用,目前越来越多的研究工作使用神经网络来解决人群计数问题。如:Huang等学者^[19]提出了一种多尺度融合网络SRNet,通过密度图的方法对人群进行计数,在密集人群场景下取得了具有竞争力的计数性能。Zhang

等学者^[20]提出了一个同时解决拥挤检测和计数问题的两阶段卷积神经网络的新框架,通过峰值置信度图对人群进行计数。与密度图相比,峰值置信度图不仅可以解决人群计数问题,而且可以准确预测人的位置。与精心设计的特征提取方案相比,基于 CNN 的方法使用简单,性能出众。基于上述研究现状,本文将肿瘤区域视为密集的人群区域从而借助密度估计的方法对全身代谢肿瘤体积进行估计。

1.3 全身代谢肿瘤体积自动化计算

在医学领域中,主要是由医生手动勾画全身肿瘤区域进行全身代谢肿瘤体积的计算,但是整个过程繁琐费时,时间成本大。也有学者借助医学知识对全身代谢肿瘤体积进行估计,如 Ilyas 等学者^[21]利用 SUV 阈值划分对弥漫性大 B 细胞淋巴瘤代谢肿瘤体积进行计算。具体来说,首先通过 SUV 的阈值勾画出肿瘤区域,再计算得到全身肿瘤代谢总体积。该方法简单高效,在淋巴瘤的数据中有不错表现,但在非小细胞肺癌的数据中却难以达到较好的效果。

在深度学习领域中,对全身代谢肿瘤体积的估计依赖于全身肿瘤的分割。如:Sadik 等学者^[22]直接使用基于 AI 的工具 PARS (来自西门子) 和 RECOMIA (recomia.org) 自动识别肿瘤,无需任何手动调整,最后得出使用人工智能工具辅助医生测量

MTV_{WB} 以评估临床实践中的预后是可行的。Blanc-Durand 等学者^[23]利用具有 PET 和 CT 两个输入通道的 3D U-Net 在弥漫性大 B 细胞淋巴瘤 (DLBCL) 患者的数据集中自动分割总代谢肿瘤体积。Kawaguchi 等学者^[24]在针对腺癌的工作中通过考虑到每个感兴趣区域的形态学特征和位置信息,以此来预测全身肿瘤代谢总体积。Capobianco 等学者^[25]通过自动全身高摄取分割算法识别出所有三维感兴趣区域 (RoI),再使用在独立队列上训练的卷积神经网络将生成的 RoI 分类为非可疑或可疑摄取区域。全身代谢肿瘤体积的计算被估计为可疑区域 RoI 值的总和。上述研究对于全身代谢肿瘤体积的估计均依赖于肿瘤的勾画,且对 MTV_{WB} 估计的精度不高,故本文提出了一种基于密度估计的方法对 MTV_{WB} 进行估计,相比于现有的方法而言性能得到了提升。

2 研究方法

本文采用了基于密度估计的两阶段方法对全身代谢肿瘤体积进行估计。第一阶段通过 nnSAM 对肿瘤进行分割,根据分割结果筛选出带肿瘤的切片,减少了全身信息的复杂度;第二阶段使用 TransUNet 网络通过密度估计的方法对全身代谢肿瘤体积进行估计,总体流程如图 1 所示。

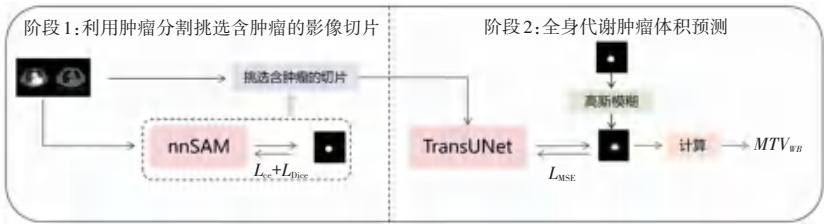


图 1 总体流程

Fig. 1 Overall process

2.1 筛选含肿瘤的图像切片

在第一阶段,本文通过 nnSAM 对非小细胞肺癌 PET/CT 图像进行肿瘤分割,根据最终的分割结果挑选出含有肿瘤的图像切片。

nnSAM 由 2 个并行编码器,即 nnUNet 编码器和 SAM 编码器组成。SAM 编码器是一个预训练的视觉变压器 (Vision Transformer, ViT)^[26]。将来自 2 个编码器的特征拼接在一起,然后馈送到 nnUNet 的解码器以输出最终的分割结果,如图 2 所示。

本文通过 nnSAM 对非小细胞肺癌的 PET/CT 图像进行肿瘤分割,根据得到的输出结果来实现对含有肿瘤切片的挑选,最终只将含有肿瘤的切片输

入到下一阶段,缩小了全身信息的范围,降低了由于全身信息复杂冗余而带来的影响。

2.2 全身代谢肿瘤体积估计

第二阶段,本文采用人群密度估计的方法通过 TransUNet 网络对全身代谢肿瘤体积进行估计。将肿瘤区域视为密集的人群区域,通过高斯模糊将肿瘤 Mask 转化为密度图,用到的公式可写为:

$$G(x,y)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{1}$$

其中, x,y 分别表示原点到 x 轴或者 y 轴的距离, σ 表示正态分布的标准差。

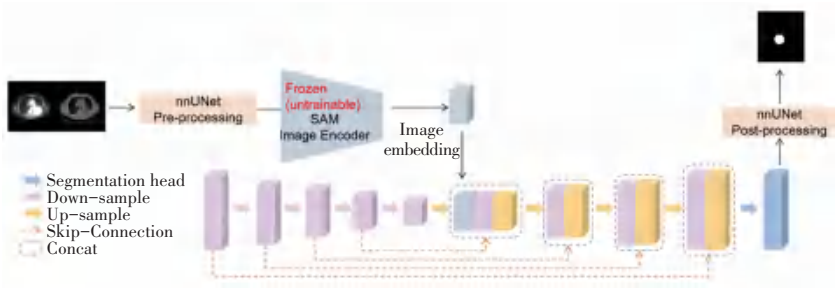


图 2 nnSAM 结构图

Fig. 2 Structure diagram of nnSAM

将密度图视为该阶段的“金标准”,因而在人群计数的应用中对网络最后的输出结果求和被视为人群数量,在本文中就可以视为是全身肿瘤的体素之和。在得到体素之和后再根据 PET/CT 的 *voxel size* 大小计算得到 MTV_{WB} ,定义公式为:

$$MTV_{WB} = \sum_j^n F(x_i^j) \times 2a \times 2b \times c \times \frac{n}{1\,000H} \quad (2)$$

其中, x_i^j 表示输入的图片; i 表示第几个病人; j 表示第 i 个病人的第 j 张切片; F 表示训练的网络模型; n 表示一个病人的总切片数; a 表示图像每个体素的长; b 表示图像每个体素的宽; c 表示图像每个体素的高; H 表示图像的像素大小。

voxel size 是指图像中每个体素的大小。体素是医学图像中最小的数据单元,通常表示为三维空间中的点。*voxel size* 通常以毫米(mm)为单位,表示图像中每个体素的长、宽、高。例如,如果 *voxel size* 是 $1\text{ mm} \times 1\text{ mm} \times 1\text{ mm}$,那么就意味着图像中的每个体素在长、宽、高三个方向上的大小都是 1 mm 。该方法相比于分割思路来看,更适合于计数的任务,降低了对全身代谢肿瘤体积估计的难度。

TransUNet 采用 CNN-Transformer 混合架构,将 CNN 提取的详细高分辨率空间信息和 Transformer 提取的全局信息进行结合,避免了单个使用的缺陷。TransUNet 通过使用 Transformer 将自关注机制引入到了 UNet 结构的编码器设计中。一方面,Transformer 对来自 CNN 特征图的标记化图像补丁进行编码,将其作为提取全局上下文的输入序列。另一方面,使用 UNet 框架的解码器对编码特征进行采样,通过恢复局部空间信息来增强更精密的细节。具体来说,首先将输入的图片经过 ResNet50 进行特征提取,保留其中 3 个阶段的输出,用于此后的跳过连接。然后对 ResNet50 输出的特征图进行序列化,再输入到 Transformers 中进行序列预测,接着将其结果合并、Reshape 成一个新的特征图,再经过 UNet 的编码器部分,通过不断的上采样和与 ResNet50 的输出做跳过连接来还原到原图的大小,整体结构如图 3 所示。本文用 TransUNet 作为主干网络,以 *MSELoss* 为损失函数,高斯模糊后的肿瘤区域为“金标准”,在得到全身肿瘤的体素之和后通过计算得到 MTV_{WB} 。

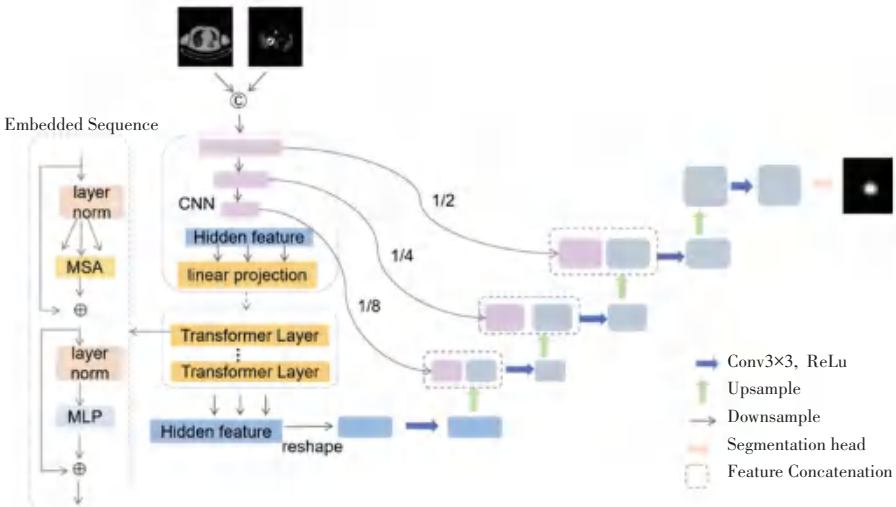


图 3 第二阶段网络结构图

Fig. 3 Network structure diagram of the second stage

2.3 损失函数

本项目研究在第一阶段利用交叉熵损失函数 (Cross-Entropy Loss)^[27] 和 Dice 损失函数 (Dice Loss)^[28] 对 nnSAM 进行训练。交叉熵损失函数主要衡量了模型预测的概率分布与真实概率分布之间的差异,基于信息论中的交叉熵概念,通过计算预测概率分布与真实概率分布之间的“距离”来评估模型的性能,由此推得的公式为:

$$L_{\text{cross}} = -y \log(p) + (1-y) \log(1-p) \quad (3)$$

其中, y 表示真实标签、取值为 0 或 1, p 表示模型预测为正类的概率。

Dice 损失函数是一种常用于图像分割任务的损失函数。具体基于 Dice 系数 (Dice Coefficient) 进行定义,用于评估预测分割图像与真实分割图像之间的相似度。Dice 系数越接近 1,表示预测结果与真实标签越相似。在医学图像分割中,Dice 系数常用于衡量模型预测的分割掩码与真实标签之间的相似度,计算公式为:

$$L_{\text{Dice}} = 1 - \frac{2|X \cap Y|}{|X| + |Y|} \quad (4)$$

其中, X 表示真实目标, Y 表示预测目标。

本研究在第二阶段利用 MSE Loss (Mean Squared Error Loss)^[29]、即均方误差损失对 TransUNet 进行训练。 MSE Loss 是机器学习和深度学习常用的一种损失函数。主要用于衡量模型预测值与真实值之间的差异, MSE Loss 通过计算预测值与真实值之间差异的平方和,然后取平均值,来衡量模型的拟合程度,计算公式为:

$$L_{\text{MSE}} = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2 \quad (5)$$

其中, N 表示预测值和真实值的数量; $\hat{y}_i$ 表示预测值; y_i 表示真实值。

3 实验

3.1 数据介绍及数据预处理

数据集包括 2004~2014 年在芝加哥大学收集的 480 例经机构审查委员会批准的非小细胞肺癌患者的识别全身 PET/CT 图像。其中,352 例 PET/CT 扫描来自 2012 年 3 月 15 日之前的 Reveal HD 扫描仪,128 例 PET/CT 扫描来自 2012 年 3 月 15 日及之后的西门子 ct 扫描仪。

在输入网络之前,需要对 PET/CT 影像数据进行预处理。首先,对 480 名非小细胞肺癌患者的 PET/CT 图像进行配准,其次计算出 PET 的 SUV

值,最后再对 PET 和 CT 图像进行归一化,将归一化之后的图像输入网络进行训练。

3.2 实验设置及参数

本研究将 480 名非小细胞肺癌患者按 3:1:1 划分为训练集、验证集和测试集。实验环境是使用 Python 3.9 和 Torch 1.13.0 构建的。图形处理单元为 NVIDIA 4090,具有 24 GB 内存。在训练阶段,将学习率设置为 0.001, $epochs$ 设置为 100, $batch$ size 大小为 64,优化器是 NAdam。

3.3 实验评价标准

本文依据 $Accuracy$ 、 $Precision$ 和 $Recall$ 三种评价指标来对第一阶段的筛选结果进行评估。

(1) 准确率 ($Accuracy$) 是指分类模型正确预测的样本数占总样本数的比例,计算公式如下:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

准确率反映了模型整体预测的正确性。

(2) 精确率 ($Precision$) 又称查准率,表示模型预测为正例的样本中真正为正例的比例,计算公式如下:

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

精确率越高,说明模型预测为正例的样本中实际为正例的样本越多,模型也能越准确地识别出目标样本。

(3) 召回率 ($Recall$) 又称查全率,表示所有实际为正例的样本中被模型正确预测为正例的比例,计算公式如下:

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

其中, TP 表示真正例,模型预测为正、且实际为正的数; TN 表示真负例,模型预测为负、且实际为负的数量; FP 表示假正例,模型预测为正、但实际为负的数量; FN 表示假负例,模型预测为负、但实际为正的数。

召回率越高,说明模型找出的实际正例越多,即模型在检测相关实例方面的能力越强。

本文依据 R^2_score 、 MSE 、 MAE 来评价第二阶段对 MTV_{WB} 估计的实验性能。

(1) 决定系数 (R^2_score)。反映因变量的全部变异能通过回归关系被自变量解释的比例,公式具体如下:

$$R^2 = 1 - \frac{\sum_i (\hat{y}^{(i)} - y^{(i)})^2}{\sum_i (\bar{y}^{(i)} - y^{(i)})^2} \quad (9)$$

(2) *MSE*: 数理统计中均方误差是指参数估计值与参数真值之差平方的期望值,公式具体如下:

$$MSE = \frac{1}{m} \sum_{i=1}^m (y_{\text{test}}^{(i)} - \hat{y}_{\text{test}}^{(i)})^2 \tag{10}$$

MSE 可以评价数据的变化程度。*MSE* 的值越小,说明预测模型描述实验数据具有更好的精确度。

(3) *MAE*:是绝对误差的平均值,能更好地反映预测值误差的实际情况,公式具体如下:

$$MAE = \frac{1}{m} \sum_{i=1}^m |y_{\text{test}}^{(i)} - \hat{y}_{\text{test}}^{(i)}| \tag{11}$$

3.4 实验结果及对比分析

本文使用 3 种主流的分割网络对第一阶段筛选含有肿瘤的图像切片性能和第二阶段对 MTV_{WB} 估计的性能进行对比分析,这 3 种方法分别是:UNet、SwinUNet、Co-learning^[30],其中 UNet 为经典模型、SwinUNet 是将 CNN 与 Transformer 结合起来的新兴模型、Co-learning 则是针对于 PET/CT 图像分割的模型,表 1 和表 2 展示了各自的性能。可以看出在第一阶段本文采用的 nnSAM 方法性能最好,准确率达到了 0.96,精确率达到了 0.85,召回率达到了 0.74。相比于其他 3 种方法,性能有着明显的提升。

表 1 第一阶段切片挑选性能评价表

Table 1 Performance evaluation table for slice selection in the first stage

模型	<i>Acc</i>	<i>pre</i>	<i>recall</i>	<i>Dice</i>
UNet	0.93	0.68	0.67	0.46
SwinUNet	0.94	0.82	0.56	0.59
Co-learning	0.93	0.69	0.70	0.53
nnSAM	0.96	0.85	0.74	0.61

表 2 与主流分割方法对比结果分析

Table 2 Analysis of comparison results with mainstream segmentation methods

模型	R^2_score	<i>MAE</i>	<i>MSE</i>
UNet	0.51	135.40	276.90
SwinUNet	0.64	99.57	189.52
Co-learning	0.55	101.32	220.89
ours	0.71	97.44	176.60

在第二阶段,本文所采用的基于密度估计两阶段方法性能最好。 R^2_score 达到了 0.71,相比于使用 SwinUNet 网络通过分割来预测 MTV_{WB} 的方法提升了 7%,相比于 UNet 和 Co-learning 则提升了 20% 和 16%,在性能方面本文方法得到了很大的提升。

本研究选取 2 名患者对其预测结果进行了具体展示,如图 4 所示。通过对比分析,本研究所得的预测结果与金标准基本保持一致。

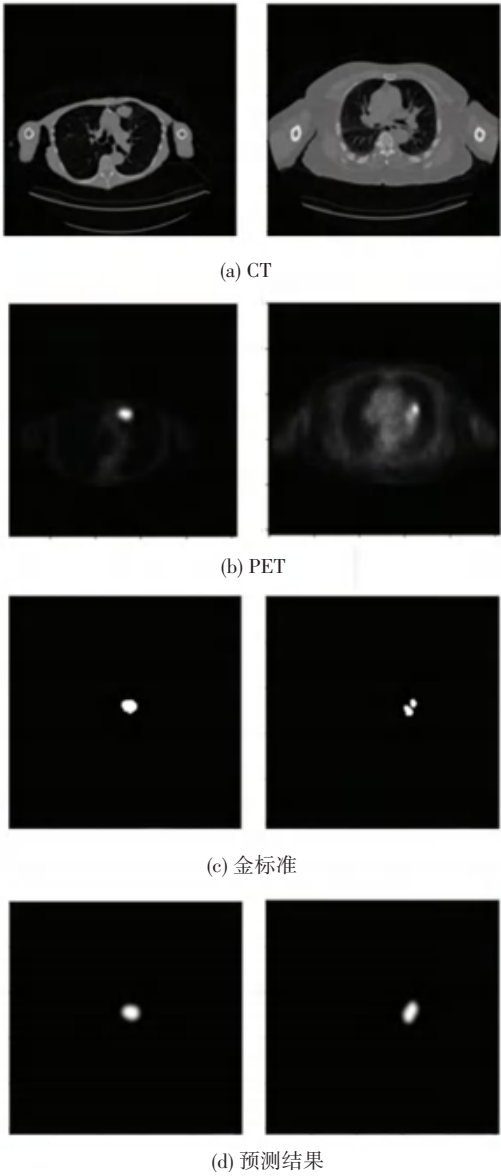


图 4 实验结果演示图

Fig. 4 Demonstration of experimental results

3.5 消融实验

为验证实验各阶段的有效性,本文设计了消融实验。具体而言,分别仅采用 nnSAM 进行肿瘤分割以估算 MTV_{WB} ,以及单独运用密度估计方法来预测 MTV_{WB} 。实验结果详列于表 3 中。表 3 中数据表明,仅依赖第一阶段或第二阶段的方法,其性能均显著低于本文提出的两阶段综合方法。这一发现有力证明了本研究所提出方法的优越性及其各阶段的必要性。

表 3 消融实验

Table 3 Ablation experiments

阶段方法	<i>R2_score</i>	<i>MAE</i>	<i>MSE</i>
☒ 第一阶段 ☒ 第二阶段	0.66	97.53	182.14
第一阶段 ☒ 第二阶段	0.50	140.27	297.56
☒ 第一阶段 ☒ 第二阶段	0.71	97.44	176.60

4 结束语

考虑到对于全身代谢肿瘤体积估计的常用方法受到全身信息复杂度高的干扰且分割方法并不能准确估计 MTV_{WB} , 本文采用了基于密度估计的两阶段方法对全身代谢肿瘤体积进行预测, 经过实验证明该方法相较于现有的估计方法有了显著提升。该两阶段方法的优势在于借助分割减少了全身信息复杂度的干扰且将分割问题转化为了计数问题, 更符合 MTV_{WB} 的计算原理。本文将肿瘤区域视为自然界中密集的人群区域, 并将肿瘤的每个像素点类比为人群中的个体, 将自然界常见的人群计数方法引入到医学领域。这个创新思路说明了在研究过程中, 视觉领域的方法在一定程度上可以相互借鉴和融合, 需要学会跨学科思考。

全身代谢肿瘤体积作为一项独立于其他指标的重要预后指标, 在实现良好的预测效果方面对医生的诊断和患者的后续治疗可提供有效助力。然而, 本研究也存在一些缺陷。例如, 在第一阶段进行切片挑选时, 采用 2D 网络可能导致切片挑选不连贯的问题。为了弥补这一不足, 在后续的研究中, 会尝试采用 3D 网络来解决这个问题, 并在此基础上对全身肿瘤代谢体积进行估计, 将性能进一步提升。

参考文献

[1] WANG Shuai, KANG Bo, MA Jinhu, et al. A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19)[J]. European Radiology, 2021, 31: 6096-6104.

[2] WANG T, CAO C, FAN Y, et al. GABPB1 plays a cancer-promoting role in non-small cell lung cancer[J]. Discover Oncology, 2024, 15(1): 72.

[3] ZHANG Hao, WROBLEWSKI K, APPELBAUM D, et al. Independent prognostic value of whole-body metabolic tumor burden from FDG-PET in non-small cell lung cancer[J]. International Journal of Computer Assisted Radiology and Surgery, 2013, 8: 181-191.

[4] ZHANG Chenpeng, LIAO Chuanhong, PENNEY B C, et al. Relationship between overall survival of patients with non-small cell lung cancer and whole-body metabolic tumor burden seenon postsurgical fluorodeoxyglucose PET images[J]. Radiology, 2015, 275(3): 862-869.

[5] LIU Xiangbin, SONG Liping, LIU Shuai, et al. A review of deep-

learning-based medical image segmentation methods[J]. Sustainability, 2021, 13(3): 1224.

[6] CHEN Jieneng, MEI Jieru, LI Xianhang, et al. TransUNet: Rethinking the U-Net architecture design for medical image segmentation through the lens of transformers[J]. Medical Image Analysis, 2024, 97: 103280.

[7] ISENSEE F, JAEGER P F, KOHL S A A, et al. nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation[J]. Nature Methods, 2021, 18(2): 203-211.

[8] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30: 5998-6008.

[9] CAO Hu, WANG Yueyue, CHEN J, et al. Swin-UNet: Unet-like pure transformer for medical image segmentation[C]//European Conference on Computer Vision. Cham: Springer, 2022: 205-218.

[10] MANZARI O N, KALEYBAR J M, SAADAT H, et al. BEFUNet: A hybrid CNN-Transformer architecture for precise medical image segmentation[J]. arXiv preprint arXiv, 2402.08793, 2024.

[11] ZHANG Jie, LI Fan, ZHANG Xin, et al. Automatic medical image segmentation with vision Transformer[J]. Applied Sciences, 2024, 14(7): 2741.

[12] ISENSEE F, JAEGER P F, KOHL S A A, et al. NNU-Net: a self-configuring method for deep learning-based biomedical image segmentation[J]. Nature Methods, 2021, 18(2): 203-211.

[13] KIRILLOV A, MINTUN E, RAVI N, et al. Segment anything[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway, NJ: IEEE, 2023: 4015-4026.

[14] LI Y, JING B, FENG X, et al. NNSAM: Plug-and-play segment anything model improves nnunet performance[J]. arXiv preprint arXiv, 2309.16967, 2023.

[15] ZHANG Yizhe, ZHOU Tao, WANG Shuo, et al. Input augmentation with SAM: Boosting medical image segmentation with segmentation foundation model[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2023: 129-139.

[16] LI Yongchao, JIA Ruisheng, HU Yingxiang, et al. A weakly-supervised crowd density estimation method based on two-stage linear feature calibration[J]. IEEE/CAA Journal of Automatica Sinica, 2024, 11(4): 965-981.

[17] CHAN A B, VASCONCELOS N. Bayesian poisson regression for crowd counting[C]//2009 IEEE 12th International Conference on Computer Vision. Piscataway, NJ: IEEE, 2009: 545-551.

[18] LIN Zhe, DAVIS L S. Shape-based human detection and segmentation via hierarchical part-template matching[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(4): 604-618.

[19] HUANG L, ZHU L, SHEN S, et al. SRNet: Scale-aware representation learning network for dense crowd counting[J]. IEEE Access, 2021, 9: 136032-136044.

[20] ZHANG Jin, CHEN Sheng, TIAN Sen, et al. A crowd counting framework combining with crowd location[J]. Journal of Advanced Transportation, 2021, 2021: 1-14.

[21] ILYAS H, MIKHAEI N G, DUNN J T, et al. Is there an optimal method for measuring baseline metabolic tumor volume in diffuse large B cell lymphoma?[J]. European Journal of Nuclear Medicine and Molecular Imaging, 2019, 46: 520-521.

[22]SADIK M, BARRINGTON S F, TRAGARDH E, et al. Metabolic tumour volume in Hodgkin lymphoma–A comparison between manual and AI–based analysis[J]. Clinical Physiology and Functional Imaging, 2024, 44(3): 220–227.

[23]BLANC–DURAND P, JEGOU S, KANOUN S, et al. Fully automatic segmentation of diffuse large B cell lymphoma lesions on 3D FDG–PET/CT for total metabolic tumour volume prediction using a convolutional neural network[J]. European Journal of Nuclear Medicine and Molecular Imaging, 2021, 48: 1362–1370.

[24]KAWAGUCHI Y, SHIMADA Y, MURAKAMI K, et al. Prognostic impact of artificial intelligence – based volumetric quantification of the solid part of the tumor in clinical stage 0–I adenocarcinoma[J]. Lung Cancer, 2022, 170: 85–90.

[25]CAPOBIANCO N, MEIGNAN M, COTTEREAU A S, et al. Deep – learning 18F – FDG uptake classification enables total metabolic tumor volume estimation in diffuse large B – cell lymphoma[J]. Journal of Nuclear Medicine, 2021, 62 (1): 30–36.

[26]DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16×16 words; Transformers for image recognition at scale[J]. arXiv preprint arXiv,2010.11929, 2020.

[27]LE P B, NGUYEN Z T. ROC curves, loss functions, and distorted probabilities in binary classification[J]. Mathematics, 2022, 10(9): 1410.

[28]MILLETARI F, NAVAB N, AHMADI S A. V – net: Fully convolutional neural networks for volumetric medical image segmentation[C]//2016 Fourth International Conference on 3D Vision (3DV). Piscataway,NJ; IEEE, 2016: 565–571.

[29]HUANG Huaibo, HE Ran, SUN Zhenan, et al. Wavelet–SRNet: A wavelet–based CNN for multi–scale face super resolution[C]// Proceedings of the IEEE International Conference on Computer Vision. Piscataway,NJ; IEEE, 2017: 1689–1697.

[30]KUMAR A, FULHAM M, FENG D, et al. Co–learning feature fusion maps from PET – CT images of lung cancer[J]. IEEE Transactions on Medical Imaging, 2019, 39(1): 204–217.