

文章编号 1004-924X(2024)09-1420-12

## 面向头颈部肿瘤转移性 淋巴结分割网络

周涛<sup>1,2</sup>, 石道宗<sup>1,2\*</sup>, 薛佳文<sup>3</sup>, 彭彩月<sup>1,2</sup>, 党培<sup>1,2</sup>, 周忠伟<sup>3</sup>

(1. 北方民族大学 计算机科学与工程学院, 宁夏 银川 750021;

2. 北方民族大学 图像图形智能处理国家民委重点实验室, 宁夏 银川 750021;

3. 宁夏医科大学 口腔学院, 宁夏 银川 750004)

**摘要:** 头颈部肿瘤是我国常见的恶性肿瘤, 其预后主要受颈部淋巴结转移的影响, 医学上通过核磁共振成像技术对转移性淋巴结成像后再进行诊断, 然而, 核磁共振技术成像存在病灶形态信息丢失, 病灶区域对比度低和病灶边界模糊的问题。针对这些问题, 提出面向头颈部肿瘤转移性淋巴结分割网络协助医生进行诊断。首先, 设计跨层跨视野注意力模块, 其接收深浅层的特征信息后利用自注意力机制分别突显深浅层的转移淋巴形状, 通过不同感受野的深层特征图学习到更好的语义上下文特征, 将浅层特征图与深层特征图逐像素融合, 增强转移淋巴病灶区域的形态信息。其次, 设计多尺度特征融合模块, 在特征金字塔的初始位置融合不同尺度的特征图, 丰富转移淋巴病灶区域的形态信息。然后, 设计增强注意力预测头模块, 通过对预测前的特征图使用并行的自注意力与门控通道转换模块, 凸出病灶区域, 细化病灶边界。最后, 使用临床淋巴结转移医学图像数据集验证网络的有效性。实验结果表明, 所提网络对于淋巴结转移病灶分割的 APdet, APseg, ARdet, ARseg, mAPdet 和 mAPseg 分别为 74.88%, 74.12%, 63.11%, 62.28%, 74.64% 和 74.04%。该网络实现对淋巴结转移病灶区域的精确检测分割, 对辅助淋巴结诊断具有积极意义。

**关键词:** 医学图像处理; 头颈部肿瘤; 淋巴结转移; 实例分割; 注意力机制

**中图分类号:** TP391.41 **文献标识码:** A **doi:** 10.37188/OPE.20243209.1420

## Segmentation network for metastatic lymph nodes of head and neck tumors

ZHOU Tao<sup>1,2</sup>, SHI Daozong<sup>1,2\*</sup>, XUE Jiawen<sup>3</sup>, PENG Caiyue<sup>1,2</sup>, DANG Pei<sup>1,2</sup>, ZHOU Zhongwei<sup>3</sup>

(1. College of Computer Science and Engineering, North Minzu University, Yinchuan 750021, China;

2. Key Laboratory of Image and Graphics Intelligent Processing of State Ethnic Affairs Commission,  
North Minzu University, Yinchuan 750021, China;

3. College of Oral Cavity, Ningxia Medical University, Yinchuan 750004, China)

\* Corresponding author, E-mail: shidaozong167@163.com

**Abstract:** Head and neck tumors are prevalent malignant tumors in China, with prognosis significantly influenced by cervical lymph node metastasis. In medical practice, magnetic resonance imaging (MRI) is employed to identify metastatic lymph nodes. However, MRI images often suffer from blurred edges and low

收稿日期: 2023-11-10; 修订日期: 2023-12-12.

基金项目: 国家自然科学基金资助项目 (No. 62062003); 宁夏自然科学基金资助项目 (No. 2022AAC03149, No. 2023AAC03293)

contrast between the lesion and surrounding tissue. This paper introduces a segmentation network tailored for metastatic lymph nodes in head and neck tumors. Initially, a cross-layer and cross-field attention module is developed, integrating features from both deep and shallow layers to enhance the shape representation of metastatic lymph nodes through a self-attention mechanism. This module improves contextual semantic understanding across different receptive fields, allowing for pixel-level fusion of shallow and deep feature maps, thereby enhancing the morphological details of metastatic lymphatic nodes. Subsequently, a multi-scale feature fusion module is designed to amalgamate features across various scales in the feature pyramid, enriching the morphological details of the lymph nodes. Furthermore, an enhanced attention prediction head module is implemented, combining parallel self-attention and gate channel transformation to accentuate the lesion area and refine its boundaries on the feature map. The network's effectiveness is confirmed using a clinical dataset of lymph node metastasis medical images. The performance metrics, APdet, APseg, ARdet, ARseg, mAPdet, and mAPseg for lymph node metastasis lesion segmentation are 74.88%, 74.12%, 63.11%, 62.28%, 74.64%, and 74.04%, respectively. This network provides precise detection and segmentation of lymph node metastasis lesions, offering significant benefits for lymph node diagnosis.

**Key words:** medical image processing; head and neck tumors; lymph node metastasis; instance segmentation; attention mechanism

## 1 引言

根据美国国家生物技术信息中心的数据<sup>[1]</sup>,癌细胞转移是导致死亡的主要原因,肿瘤倾向于以淋巴系统为目标,因为淋巴系统是重要的抵抗病毒入侵的防线。头颈部肿瘤是我国常见的肿瘤疾病<sup>[2]</sup>,恶性肿瘤扩散时,癌细胞会进入淋巴系统,通过淋巴液流向周围的淋巴结,癌细胞在淋巴结内停留并继续增殖,淋巴结可能会肿大,并形成转移性淋巴结<sup>[3]</sup>。颈部淋巴结是头颈部肿瘤常见的转移部位。因此,在手术中准确识别和切除颈部转移淋巴结对于减少头颈肿瘤复发至关重要。随着医学图像设备与成像技术的发展,转移性淋巴结采用影像学检查辅助诊断的比例越来越高。核磁共振成像(Magnetic Resonance Imaging, MRI)<sup>[4]</sup>可以提供淋巴结影像信息,对软组织和淋巴结的分辨率较高,有助于判断淋巴结的状况,已成为颈部淋巴结常用的诊断措施。

传统的转移淋巴结诊断是由医生进行人工阅片,存在漏诊和误诊风险<sup>[5]</sup>。利用深度学习技术辅助转移性淋巴结诊断不仅可以优化工作流程,而且可以显著提高转移性淋巴结的诊断精确度<sup>[6]</sup>。Gurney-Champion等<sup>[7]</sup>通过训练深度卷积神经网络,在接受放疗的头颈部癌症(异氢氨酸)

患者的弥散加权(Diffusion Weighting, DW)MRI上自动勾画所涉及的淋巴结。Tong<sup>[8]</sup>等提出一种淋巴浸润检测框架,利用基于生成性和对抗性网络的图像到图像转换来实现淋巴浸润的智能定量检测。Luo等<sup>[9]</sup>提出一种半监督的多视图网络来诊断淋巴结转移,它结合粗视图和细视图来获得全面的特征信息,有效地提取全局和局部特征,从而得到准确的诊断结果。Zhao等<sup>[10]</sup>提出一种基于深度学习的全自动淋巴结检测和分割模型,该模型通过迁移学习为Mask R-CNN<sup>[11]</sup>提供输入参数,从而实现全自动淋巴结检测和分割的auto-LNDS模型。上述研究表明,通过对转移性淋巴结图像进行病灶信息检测与分割可以准确识别转移淋巴结,从而实现辅助诊断的作用。

由于核磁共振仪器中的局部磁场不均匀性会导致图像畸变和病灶形态信息丢失,同时,核磁共振图像中的噪声或伪影会降低病灶区域的对比度<sup>[12]</sup>,使得病灶区域与周围正常组织在外观上难以区分,从而模糊病灶边界。因此,本文提出面向头颈部肿瘤转移性淋巴结分割网络(Lymph Node Metastasis Segmentation Network, LNMS Net)。设计跨层跨视野注意力模块(Cross layer Cross field Attention Module, C<sup>2</sup>AM),通过C<sup>2</sup>AM模块获取深浅层的特征信

息,利用自注意力机制分别增强深浅层的转移淋巴的形状和边缘信息,并通过不同感受野的深层特征图学习病灶区域不同语义上下文特征,然后将浅层特征图与深层特征图逐像素融合,从而增强转移淋巴病灶区域的形态信息。设计多尺度特征融合模块(Multi-scale Feature Fusion Module, MFFM),在特征金字塔的初始位置将经过C<sup>2</sup>AM增强病灶区域后的不同尺度特征图进行融合,丰富转移淋巴病灶区域的形态信息,再对融合后的图像进行加权,从而有效地融合前景突出特征和背景突出特征。设计增强注意力预测头模块(Enhanced Attention Prediction Head Module, EAPM),对预测前的特征图使用注意力机制突出病灶区域,同时使用门控通道转换模块,从而达到细化病灶边界的目的。

## 2 整体网络

本文基于单阶段实例分割法<sup>[13]</sup>中的面向头颈部肿瘤转移性淋巴结分割网络LNMS Net<sup>[14]</sup>,

模型整体网络结构如图1所示。颈部淋巴结图像输入网络后,采用ResNet101<sup>[15]</sup>作为主干网络提取不同尺度的淋巴转移特征图,然后分别将邻近的深层特征图与浅层特征图输入C<sup>2</sup>AM中,从而增强转移淋巴病灶区域的形态信息。然后,将增强后的特征图传入特征金字塔。使用MFFM接收来自4个C<sup>2</sup>AM的不同尺度的特征图,从而增强特征金字塔的特征融合能力,同时,在特征金字塔自上而下的融合过程中,C<sup>2</sup>AM3,C<sup>2</sup>AM4传入特征金字塔中,进行淋巴特征融合,从而丰富病灶区域的形态信息。通过构造特征增强预测头,利用特征增强模块(Enhanced Feature Module, EFM)实现淋巴转移病灶的特征增强后,进行掩膜系数预测、锚框定位以及类别判定;然后,将每一个特征图获取的特征预测结果进行融合,并利用非极大值抑制(Non-Max Suppression, NMS)获取目标边界框与掩膜系数的极大值;最后,将掩膜系数和原型分支进行线性组合,并按照目标边界进行切片处理,通过真实标注与预测结果间的阈值判定分割结果。

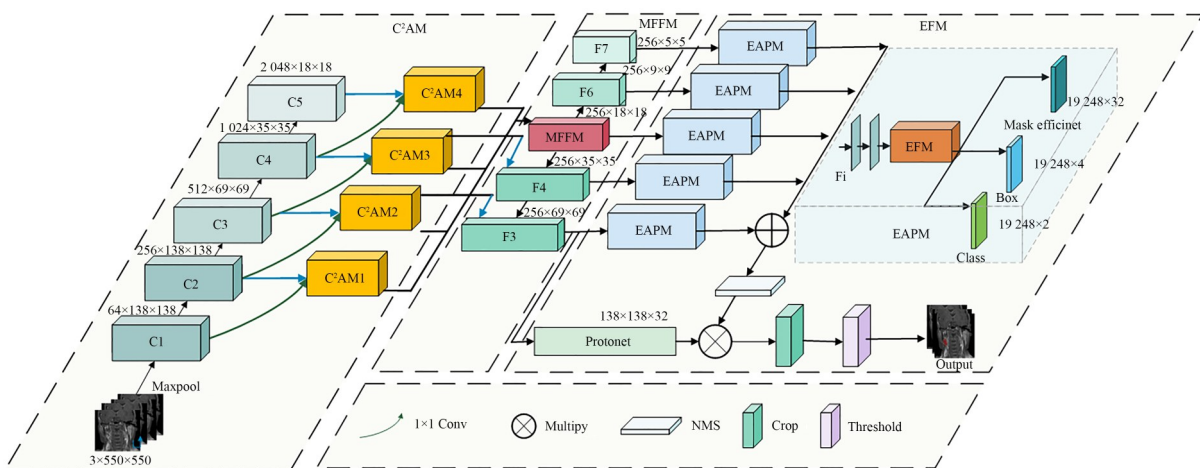


图1 LNMS Net模型整体结构

Fig. 1 Overall structure of LNMS Net model

在LNMS Net模型中,针对转移淋巴中形态特征不突出以及边界模糊的问题,本文在特征提取阶段设计C<sup>2</sup>AM模块,通过深浅层特征交互增强淋巴病灶信息,然后从不同的视野对病灶特征进行融合,达到有效增强病灶边界的目的。同时,在特征融合阶段构建多尺度特征融合模块,将经过C<sup>2</sup>AM增强的特征图进行融合,从而提高

特征金字塔特征融合能力。为了实现对转移淋巴的精准分割与定位,设计增强注意力预测头,通过增强注意力模块有效提高转移淋巴的病灶信息与背景信息的对比度,进一步增强模型的特征表达能力。

### 2.1 C<sup>2</sup>AM 模块

由于转移淋巴病灶区域的形状与尺寸存在

差别且病变区域及其周围组织具有相似的强度,残差神经网络虽然能提取病灶信息,但是捕捉图像全局信息的能力有限,容易造成病灶信息丢失、形变。因此,本文设计跨层跨视野注意力模块,它接收深浅层的特征信息后利用注意力机制分别突显深浅层的转移淋巴形状、边缘信息,并通过不同感受野学习到更好的语义上下文特征,有效地缓解特征提取过程中产生的病灶形态信息丢失问题。

$C^2AM$  模块结构如图2所示,头颈部淋巴结图像经过特征提取后,生成浅层特征图和深层特征图,然后将深浅层特征图分别通过  $1 \times 1$  卷积获取查询( $Q$ )、键( $K$ )以及值( $V$ )对特征进行投影,并与可学习矩阵相乘,得到3个张量  $W_q$ ,  $W_k$  以及  $W_v$ ,将  $W_q$  和  $W_k$  执行点积  $F(Q, K)$ ,增强

淋巴结病灶区域的相似度,如下:

$$F(Q, K) = QK^T. \quad (1)$$

将点积后的特征图进行 Softmax 归一化处理,获取头颈部淋巴特征图的注意力权重,并将权重与  $W_v$  张量进行加权获取需要注意的病灶区域特征,从而突显深浅层转移淋巴的病灶区域。

接着深层特征图通过  $1 \times 1$  卷积、 $3 \times 3$  卷积以及  $5 \times 5$  卷积获取不同感受野特征图,然后将不同尺寸的感受野通过  $1 \times 1$  卷积调整尺寸后进行融合,从而强化深层特征图的病灶形态信息,并将融合后的特征图进行 Softmax 处理,获取注意力权重后对输入的深层特征图进行加权,有效地融合前景显著特征和背景显著特征。最后,将浅层特征图与深层特征图逐像素融合,增强转移淋巴病灶区域的形态信息。

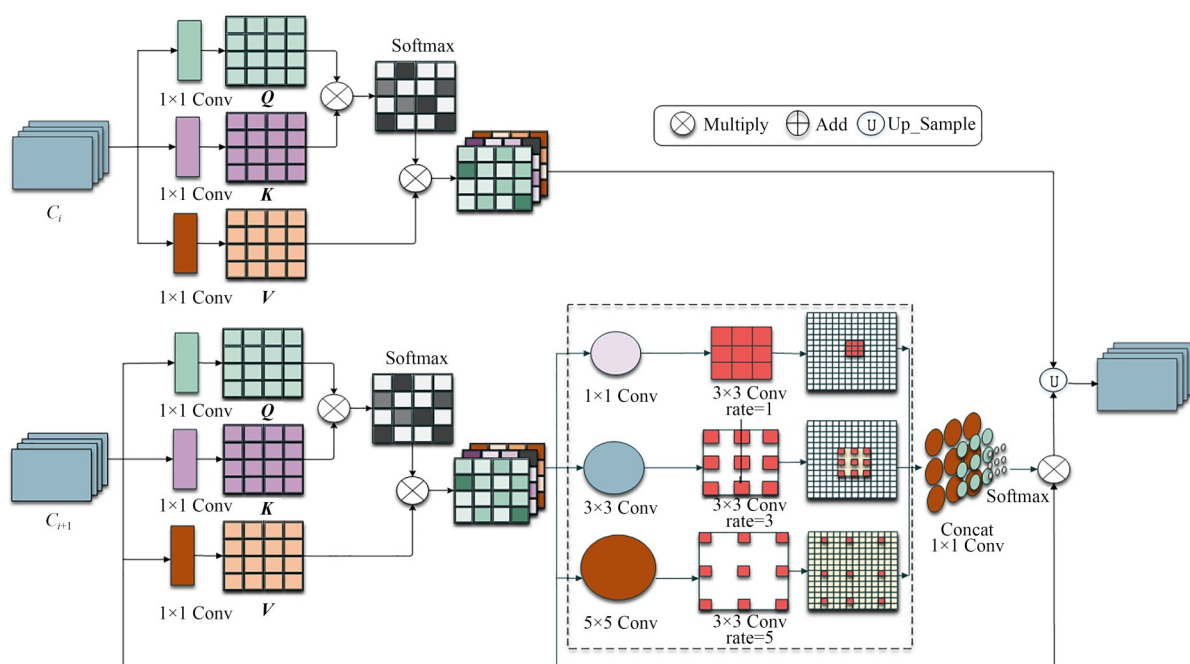


图2 跨层跨视野注意力模块

Fig. 2 Cross layer and cross field attention module

## 2.2 MFPM 模块

转移淋巴经过核磁共振成像后,由于局部磁场不均匀,病灶区域存在形态信息丢失和畸变问题。特征金字塔<sup>[16]</sup>通过低分辨率、语义强的特征与高分辨率、语义弱的特征结合达到特征图融合的目的,因此,提高特征金字塔初始位置的特征图所蕴含信息的准确度以及丰富程度,能有效降低融合过程中的背景噪声因子并增强病灶的纹

理、轮廓等形态信息。本文设计多尺度特征融合模块,在特征金字塔的初始位置融合经过跨层跨视野注意力模块处理过的不同层的特征图,有效融合前景突出特征和背景突出特征,丰富转移淋巴病灶区域的形态信息,如图3所示。

首先,将经过跨层跨视野注意力模块处理的  $C^2AM1$ ,  $C^2AM2$ ,  $C^2AM3$ ,  $C^2AM4$  分别传入深度分别为  $D_1$ ,  $D_2$ ,  $D_3$ ,  $D_4$  的4条扩张卷积链,



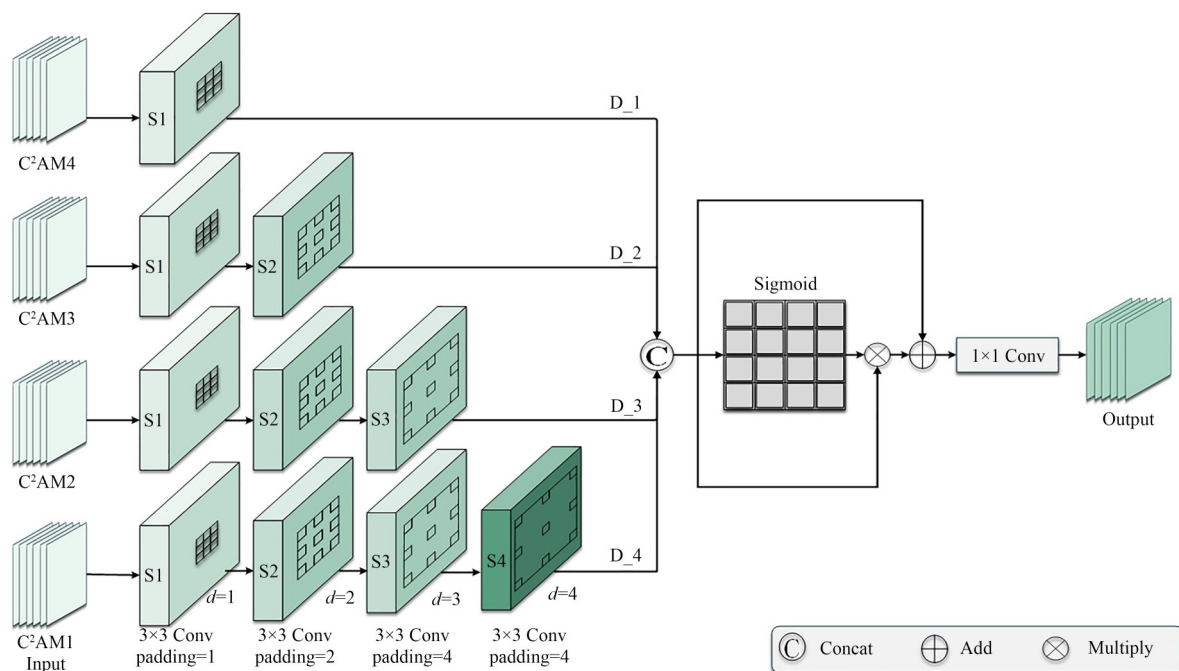


图3 多尺度特征融合模块

Fig. 3 Multi-scale feature fusion module

扩张卷积链由 $3 \times 3$ 扩张卷积,不同扩张率 $d$  ( $d=1, 2, 3, 4$ )以及不同padding(padding=1, 2, 3, 4)堆叠而成,将4条扩张卷积链的输出进行逐元素相加融合,充分融合不同尺度的病灶区域信息,从而丰富特征图的特征信息。然后,通过Sigmoid激活函数将融合后的特征图权重压缩映射在0~1之间,系数越大表示更加关注该像素特征,并将权重注意图与融合特征图相乘,从而凸显病灶区域。最后,将加权后的特征图与融合后的特征图进行逐像素相加得到最终融合特征图,从而达到增强病灶区域的同时丰富病灶区域形态信息的目的。

### 2.3 EAPM模块

单阶段实例分割对转移淋巴的特征预测是将经过特征融合后的特征图进行锚框生成与掩膜生成,从而实现对目标定位以及分割。然而,转移性淋巴结周围组织和病变区域之间的模式复杂性和强度相似性,使得模型无法很好地区分医学图像中的病灶区域与背景信息。为了进一步提高特征图背景区域和病灶区域之间的对比度,本文设计增强注意力预测头模块,通过对预测前的特征图使用并行的自注意力机制<sup>[17]</sup>与门控通道转换模块<sup>[18]</sup>,突出病灶区域,细化病灶边

界,能有效提高模型对病灶区域的检测以及分割性能。

如图4所示,增强注意力预测头模块将接收来自增强特征金字塔不同层 $F_i$ ,在经过增强注意力预测头模块时,分别将其传入的特征图进行自注意力机制以及门控通道转换模块,通过使用不同的特征增强方式有效地提高转移淋巴的背景区域与病灶区域的对比度。其中,自注意力机制将传入的特征图通过 $1 \times 1$ 卷积获取查询、键、值,然后与可学习矩阵相乘获得3个线性变换矩阵 $W_q, W_k$ 以及 $W_v$ ,再通过点积计算对 $W_q$ 矩阵和 $W_k$ 矩阵中的每个元素进行相似性匹配,并通过Softmax函数转换成0~1之间的概率分布,最后将转换结果和 $W_v$ 矩阵相乘进行加权,从而增强转移淋巴的病灶区域。与此同时,门控通道转换模块首先通过L2-norm规范化组件来建模通道关系,具体如下:

$$S_c = \alpha_c \|x_c\|_2 = \alpha_c \left\{ \left[ \sum_{i=1}^H \sum_{j=1}^W (x_c^{i,j}) \right] + \epsilon \right\}^{\frac{1}{2}}, \quad (2)$$

其中: $x_c$ 表示输入特征图, $\epsilon$ 为用于避免倒数为0的极小值,参数 $\alpha$ 设定为 $\alpha = [\alpha_1 \cdots \alpha_c]$ ,然后再利用L2范式进行通道建模,并引入通道参数对特

征通道的门控自适应行为进行控制,具体如下:

$$\tilde{S}_c = \frac{\sqrt{C} S_c}{\left[ \left( \sum_{c=1}^C S_c^2 \right) + \epsilon \right]^{\frac{1}{2}}}, \quad (3)$$

其中 $\epsilon$ 作用同前, $s_c$ 为经过规范化处理的特征图。尺度因子 $\sqrt{C}$ 用于规范化 $\tilde{S}_c$ 的尺度以避免尺度过小,然后使用 $1+\tanh(x)$ 门激活,将残差连接

也引入门控适配中,使门控通道转换模块能够进行建模实体映射,并在浅层特征学习病灶形态信息,在深层特征中突出病灶区域,具体如下:

$$\tilde{x}_c = x_c [1 + \tanh(\gamma_c \tilde{s} + \beta_c)], \quad (4)$$

其中 $\gamma_c, \beta_c$ 设定同 $\alpha_c$ 。最后,将增强过的特征图分别进行锚框生成、掩膜系数生成以及类别判定。

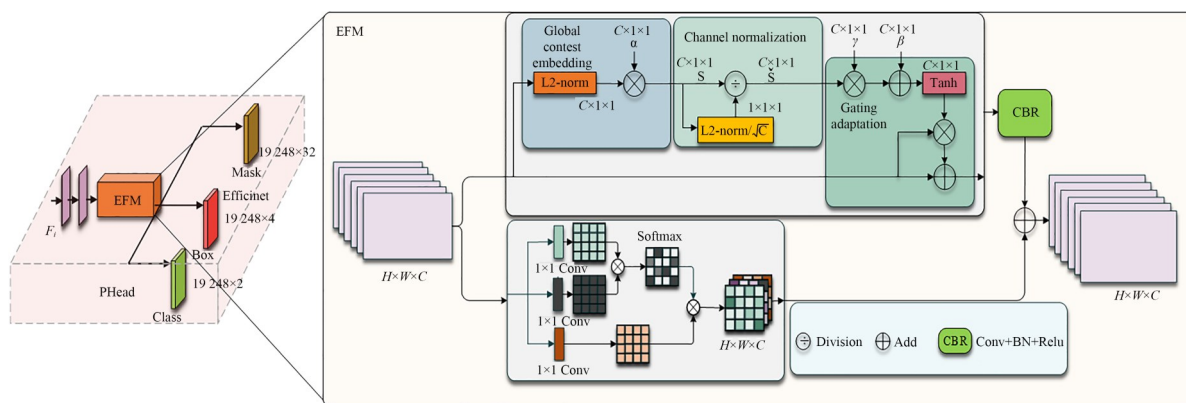


图4 增强注意力预测头

Fig. 4 Enhanced attention prediction head module

### 3 实验

#### 3.1 实验环境与参数设定

实验环境为 Windows Server 2019 系统,计算机内存为 64GB,搭载两块 Intel Xeon Gold 6154, 3 GHz x36 CPU 处理器,并采用两块并行的 TITAN V 显卡加速图像处理。基于 Pytorch (GPU) 框架进行网络搭建,Python 3.7.10,Pytorch 1.7.0 深度学习框架,CUDA 11.3.58。

在模型训练中,采用SGD优化器优化模型参数,初始学习率设置为0.00001,训练周期设置为100,批处理大小设置为2,损失函数包括分类损失 $class\_loss$ 、边框回归损失 $bbox\_loss$ 、掩膜分割损失 $mask\_loss$ 。其中分类损失为交叉熵损失,边框回归损失如式(6)所示,掩膜分割损失为交叉熵损失,具体如下:

$$L_{class} = BCE(C, C_{gt}), \quad (5)$$

$$L_{box} = 1 - IOU(B, B_{gt}), \quad (6)$$

$$L_{mask} = BCE(M, M_{gt}), \quad (7)$$

其中: $C$ 为预测类别, $C_{gt}$ 为真实类别, $B$ 为预测边框, $B_{gt}$ 为真实边框, $M$ 为预测掩膜, $M_{gt}$ 为真实掩膜。

#### 3.2 实验设计思路

本文通过设计消融实验和对比实验来验证 LNMS Net 模型的有效性和合理性。其中,消融实验设计如表1所示。实验一,主干网采用 ResNet101 进行特征采取,然后通过 FPN 进行特征融合,使用预测头进行特征预测;实验二,采用 ResNet101 作为主干网络,同时使用 C<sup>2</sup>AM 模块进行深浅层特征融合,然后将深浅层融合后的特征图传入用 FPN 进行特征融合,再使用预测头进行特征预测;实验三,主干网采用 ResNet101 提取特征,同时通过 C<sup>2</sup>AM 模块将深浅层特征图的特征进行融合,然后将不同 C<sup>2</sup>AM 模块的特征图输入 MFFM,同时 C<sup>2</sup>AM2 和 C<sup>2</sup>AM3 传入特征金字塔的 F3 和 F4 层进行特征交互,再使用预测头进行特征预测;实验四,通过 ResNet101 提取特征后,利用 C<sup>2</sup>AM 将主干网上深浅层特征图进行特征交互,并将交互后的特征图传入 MFFM 中,同时将 C<sup>2</sup>AM2 和 C<sup>2</sup>AM3 传入特征金字塔的 F3, F4 层进行特征融合,最后通过 EAPM 进行特征预测。

表 1 消融实验设计

Tab. 1 Design of ablation experiments

消融实验	ResNet101	FPN	Head	C <sup>2</sup> AM	MFFM	EAPM
一	✓	✓	✓	×	×	×
二	✓	✓	✓	✓	×	×
三	✓	✓	✓	✓	✓	×
四	✓	✓	✓	✓	✓	✓

对比实验设计如表 4 所示,第一类实验是两阶段实例分割网络对比,第二类实验是单阶段实例分割网络对比,共两类七次实验。第一个实验,使用本文模型的基础网络 Yolact 模型作为对比,其中特征提取的主干网为 Resnet50,研究特征提取网络层数较低的对比情况;第二个实验,使用基础模型 Yolact 进行对比,其中主干网为 Resnet101,从而对比本文模型相同条件下基础模型的实验情况;第三个实验,采用 Yolact++<sup>[19]</sup>进行对比,主干网络为 Resnet50,研究同样使用 Yolact 作为基础网络的 Yolact++在 Resnet50 上的实验情况;第四个实验,采用主干网为 Resnet101 的 Yolact++进行对比,从而研究在主干网相同情况下本文实验与 Yolact++的对比情况;第五个实验,采用主干网为 Resnet50 的 Mask RCNN<sup>[20]</sup>进行对比,探究本文模型与两阶段实例分割模型的实验情况;第六个实验,采用主干网为 Resnet101 的 Mask RCNN 进行对比;第七个实验,LNMS Net 采用 Resnet101 作为主干。其中,实验一至实验四是单阶段实例分割的对比,实验五与实验六是两阶段实例分割

的对比。验证实验设计是用训练、测试好的模型分割 150 张头颈部肿瘤的原始图像,并进行可视化结果展示。

4 数据集与评价指标

4.1 数据集

数据集选用 2015 年 1 月至 2023 年 6 月,宁夏某三甲医院进行头颈部恶性肿瘤高分辨率 MRI 扫描以及颈淋巴结活检的 53 例患者数据,包括男性 35 例,女性 18 例,年龄在 6~80 岁,平均年龄为 57.717 岁。提取多序列水平位、冠状位和矢状位图像。本次实验仅标记水平位图像(即 AX T2WI 和 AX T1WI),在标记过程中同时参考冠状面和矢状面来共同判断转移淋巴结。磁共振扫描仪采用飞利浦 Ingenia3.0T 全数字磁共振,患者呈仰卧体位,并尽可能平静呼吸避免咳嗽及吞咽动作减少移动伪影,将患者颈部垫高、双肩下垂以充分暴露甲状腺,采用头颈联合专用表面线圈,扫描范围自外耳孔水平至主动脉弓水平,视具体病变情况酌情扩大扫描范围。在对比增强研究中,患者以 0.2 mL/kg 体重的剂量静脉内注射对比剂。为了保护患者隐私,医生对数据进行剪切处理。本次数据经过两位口腔颌面外科医生的标记,当标记结果不一致时,由第三位影像科专业医生评估并作为最终诊断结果,特别情况下需结合临床综合诊断。数据训练集 1 438 张,测试集 363 张,验证集 150 张,共 1 951 张。图 5 为本次实验的部分原始图像及其对应的标注图像。

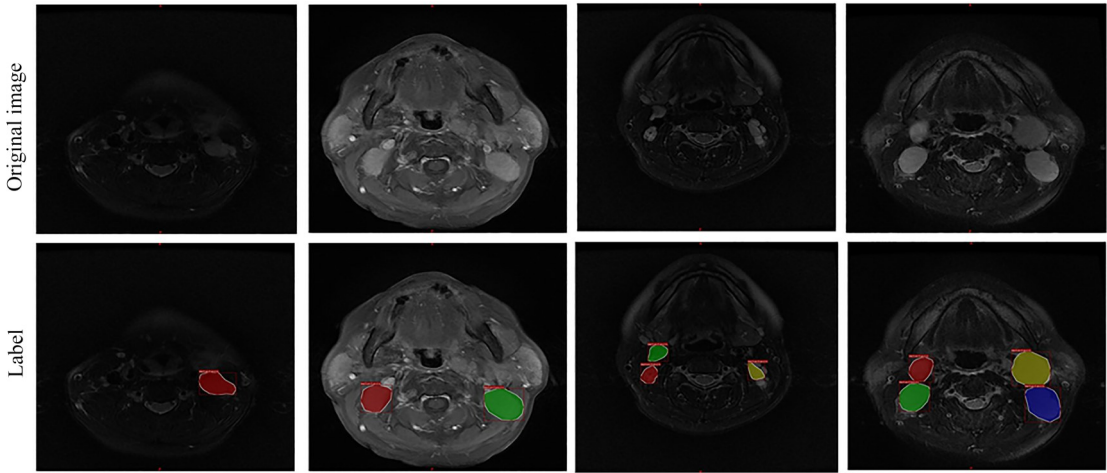


图 5 数据集展示

Fig. 5 Dataset presentation

4.2 评价指标

本文使用交并比 (Intersection over Union, IoU), 平均精度 (Average Precision, AP), 平均召回率 (Average Recall, AR) 以及均值平均精度 (Mean Average Precision, mAP) 作为检测和分割两个部分的评价标准。其中, 检测部分使用的平均精度为 (Average Precision detection, APdet), 平均召回率为 (Average Precision detection, ARdet), 均值平均精度为 (Mean Average Precision detection, mAPdet); 分割部分使用的平均精度为 (Average Precision segmentation, APseg), 平均召回率为 (Average Precision segmentation, ARseg), 均值平均精度为 (Mean Average Precision segmentation, mAPseg), 从而更全面地分析和评

估 LNMS Net 模型的识别性能。本文根据图像真实类别和模型预测类别构建混淆矩阵<sup>[21]</sup>, 混淆矩阵反映了模型对各个类别的真实值和预测值之间的差异, 并通过混淆矩阵中的真阳性 (True Positive, TP) 值、假阳性 (False Positive, FP) 值、真阴性 (True Negative, TN) 值和假阴性 (False Negative, FN) 值计算以上评价指标。其中, TP 表示模型正确识别出病灶区域; FP 表示模型将正常组织区域误诊为病灶区域; TN 表示模型正确识别出正常组织区域; FN 表示模型将病灶区域漏诊为正常组织区域。各项评价指标的定义和公式如表 2 所示, 其中  $c$  为检测类别,  $th$  表示每一类别的阈值,  $t$  为检测样本数。

表 2 评价指标及相关公式  
Tab. 2 Evaluation indicators and related formulas

指标	作 用	公 式
IoU	IoU 用于测量标注框和预测框之间的相关度, 数值越大, 表示相关度越高, 模型性能越好。	$IOU = TP / (TP + FP + FN)$
AP	平均精度 (Average Precision, AP) 是正确识别的目标个数占总识别的目标个数的百分数, 用于衡量在每个类别上模型检测器的性能优劣。其中, AP50 表示 IoU 阈值为 0.5 时的 AP 值。数值越大, 模型性能越好。	$AP = \frac{1}{ t } \frac{1}{ th } \sum_t \frac{TP(t)}{TP(t) + FP(t)}$
AR	平均召回率 (Average Recall, AR) 是正确识别的目标个数占测试集中识别的目标个数的百分数。其中, AR50 表示 IoU 阈值为 0.5 时的 AR 值。数值越大, 模型性能越好。	$AR = \frac{1}{ t } \frac{1}{ th } \sum_t \frac{TP(t)}{TP(t) + FN(t)}$
mAP	均值平均精度 (Mean Average Precision, mAP) 是衡量检测精度的指标。其中, mAP50 表示 IoU 阈值为 0.5 时的 mAP 值。数值越大, 该模型性能越好。	$mAP = \frac{1}{ C } \sum_c \left( \frac{1}{ t } \frac{1}{ th } \frac{TP(t)}{TP(t) + FP(t)} \right)$

5 实验结果

5.1 消融实验

本文在头颈部肿瘤淋巴结转移的图像数据集上进行消融实验, 从而对 LNMS Net 各个模块的设计合理性以及有效性进行充分验证, 实验结果如表 3 和图 6 所示。其中, 实验一是基础模型 Yolact 的结果, APdet, APseg, ARdet, ARseg, mAPdet 和 mAPseg 分别为 70.37%, 70.58%, 60.12%, 59.78%, 70.22% 以及 70.43%, 实现对

转移性淋巴结的精准识别和分割, Yolact 能够实现对头颈部肿瘤中同一图像的单个病灶的准确分割, 然而对于同一图像中的多个病灶却无法识别和准确分割。实验二是基于实验一的模型在主干网增加 C<sup>2</sup>AM 模块对深浅层的特征信息进行交互从而丰富转移性淋巴形状以及边缘信息, 其各项实验指标均高于实验一, 证明通过加入 C<sup>2</sup>AM 模块后, 模型对于同一图像中多病灶的分割性能得到提升。实验三是在实验二的基础上增加 MFFM 模块同时将 C<sup>2</sup>AM2, C<sup>2</sup>AM3 特征图



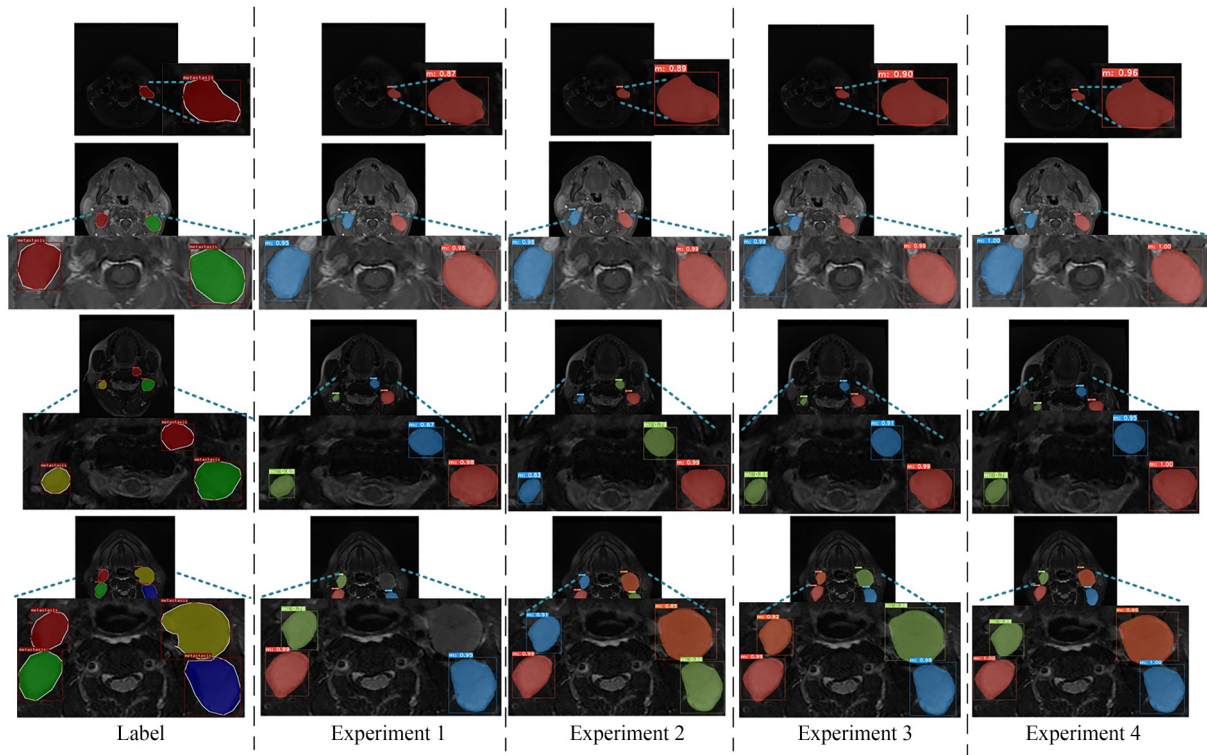


图6 消融实验可视化结果

Fig. 6 Visualization results of ablation experiments

表 3 消融实验结果(IoU=0.50)

Tab. 3 Results of ablation experiment (IoU=0.50)(%)

实验	APdet	APseg	ARdet	ARseg	mAPdet	mAPseg
一	70.37	70.58	60.12	59.78	70.22	70.43
二	71.96	71.77	61.23	61.01	71.67	71.58
三	73.12	72.86	62.40	61.57	73.02	72.66
四	74.88	74.12	63.11	62.28	74.64	74.04

传入特征金字塔中,实验结果相比于实验二进一步提升,APdet,APseg,ARdet,ARseg,mAPdet和mAPseg相比实验二分别提高了1.16%,1.09%,1.17%,0.56%,1.35%以及1.08%,表明MFFM模块能够有效增强病灶区域特征,模型对头颈部肿瘤病灶区域的分割更加细致,更贴合病灶的轮廓。实验四是模型经过特征融合后,利用EAPM模块细化转移性淋巴结边界,实验结果相较于实验三得到有效提升,充分证明EAPM模块的有效性,模型对于同一头颈部肿瘤图像的多病灶分割能力得到极大提升,所分割的区域更加接近标注的真实病灶区域。

5.2 对比实验

该文模型与其他实例分割网络在同一数据集上的对比结果如表4所示。其中,Yolact(Resnet50)的APdet,APseg,APdet,APseg,mAPdet,mAPdseg分别为68.89%,69.94%,58.77%,57.68%,69.03%以及70.13%;Yolact(Resnet101)的各项指标数值较Yolact(Resnet50)有一定程度的提升,表明Yolact模型在主干网网络深度提升时,能够提高模型对病灶区域的分割性能;Yolact++(Resnet50)的APdet,APseg,ARdet,ARseg,mAPdet,mAPdseg分别为72.86%,72.45%,61.78%,61.33%,72.97%以及72.56%,相比于Yolact(Resnet101),各项指标得到一定幅度的提升,表明Yolact++(Resnet50)能够对病灶进行精准分割;Yolact++(Resnet101)是在Yolact++(Resnet50)的基础上加深主干网的深度,其APdet,APseg,ARdet,ARseg,mAPdet,mAPdseg相比Yolact++(Resnet50)分别提高1.05%,0.93%,1.1%%,1.23%,1.05%以及1.32%。Mask RCNN(Resnet50)是两阶段实例分割模型Mask RCNN在主干网为ResNet50对转移性淋

表 4  对比实验结果(IoU=0.50)

Tab. 4  Comparative experimental results (IoU=0.50)

(%)

模  型	APdet	APseg	ARdet	ARseg	mAPdet	mAPseg
Yolact(Resnet50)	68.89	69.94	58.77	57.68	69.03	70.13
Yolact(Resnet101)	70.37	70.58	60.12	59.78	70.22	70.43
Yolact++(Resnet50)	72.86	72.45	61.78	61.33	72.97	72.56
Yolact++(Resnet101)	73.91	73.38	62.88	62.56	74.02	73.88
MaskRcnn(Resnet50)	72.77	72.56	61.86	61.35	72.68	72.47
MaskRcnn(Resnet101)	74.05	73.78	62.68	62.02	73.89	73.65
LNMS Net	74.88	74.12	63.11	62.28	74.64	74.04

巴结的分割结果,其各项指标表明 Mask RCNN 能够精准分割转移性淋巴结的病灶区域;Mask RCNN (Resnet101) 相 对 于 Mask RCNN (Resnet50), 各 项 指 标 分 别 提 高 了 1.28%, 1.22%, 0.82%, 0.67%, 1.21% 以 及 1.18%; 本 文 模 型 的 各 项 指 标 分 别 为 74.88%, 74.12%, 63.11%, 62.28%, 74.64% 以 及 74.04%, 相 对 于 其 他 实 例 分 割 方 法, LNMS Net 对 转 移 性 淋 巴 结

病灶区域的分割更准确。

5.3  验证结果可视化

该文模型通过使用 150 张验证集进行模型性能验证,部分结果如图 7 所示。最上层为 5 张原始图像,中间为 5 张标注图像作为可视化结果参照图像,最下层为 5 张分割图像。分割图像中能够精准划分头颈部肿瘤图像中的单病灶区域和多病灶区域,从而证明了本文模型的有效性。

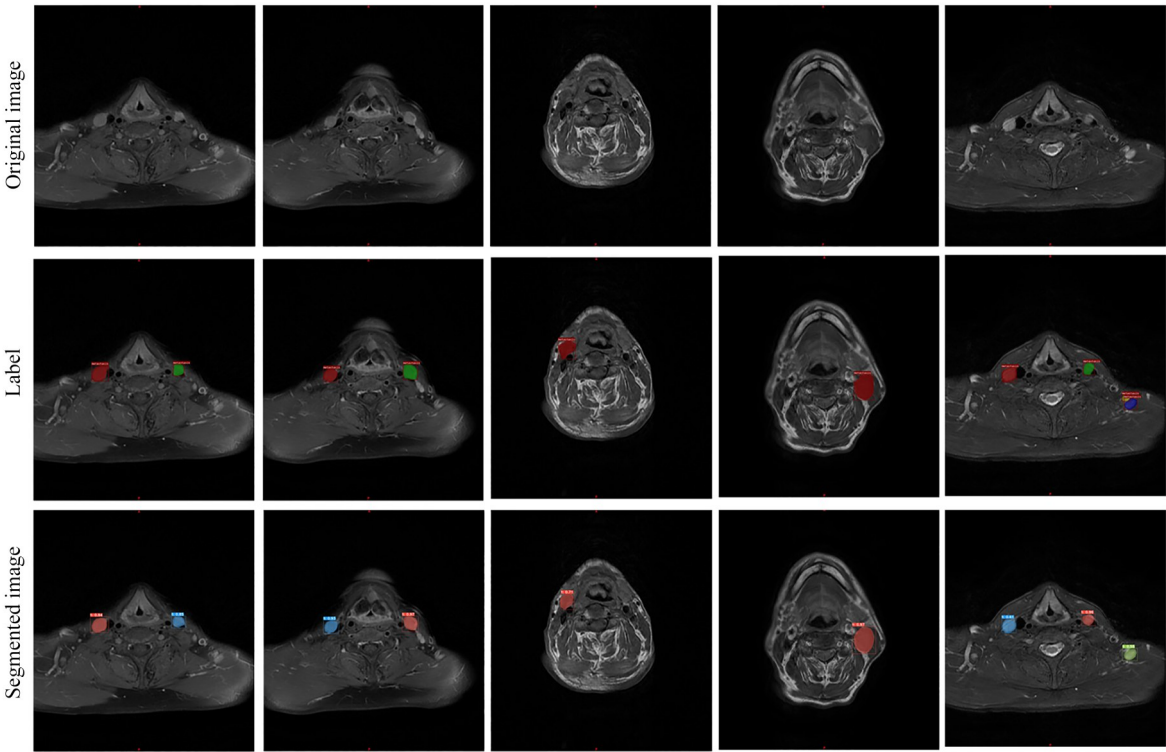


图 7  验证实验可视化结果

Fig. 7  Visualization results of validation experiments

## 6 结 论

本文提出面向头颈部肿瘤转移性淋巴结分割网络。该模型设计了跨层跨视野注意力模块,该模块接收深浅层的特征信息后利用注意力机制分别突显深浅层的转移淋巴形状、边缘信息,并通过不同感受野的深层特征图学习到更好的语义上下文特征,将浅层特征图与深层特征图逐像素融合,增强转移淋巴病灶区域的形态信息。其次,设计多尺度特征融合模块,在特征金字塔的初始位置融合不同尺度的特征图,有效地融合前景突出特征和背景突出特征,丰富转移淋巴病灶区域的形态信息。然后,设

计增强注意力预测头模块,通过对预测前的特征图使用并行的自注意力与门控通道转换模块,突出病灶区域,细化病灶边界,有效提高模型对病灶区域的检测以及分割性能。最后,使用临床淋巴结转移医学图像数据集验证模型的有效性。结果表明,所提模型对于淋巴结转移病灶分割的 APdet, APseg, ARdet, ARseg, mAPdet 和 mAPseg 分别为 74.88%, 74.12%, 63.11%, 62.28%, 74.64%, 74.04%, 能够实现淋巴结转移病灶区域的精确检测分割,对辅助淋巴结诊断具有积极意义。然而,对于淋巴结成像过程中周围组织挤压病灶造成的病灶信息缺失问题,仍需要进一步研究。

## 参考文献:

- [1] PYNNONEN M A. Clinical practice guideline: Evaluation of the neck mass in adults[J]. *Otolaryngol. Head Neck Surg.*, 2017, 157(2):355.
- [2] DILLEKÅS H, ROGERS M S, STRAUME O. Are 90% of deaths from cancer caused by metastases? [J]. *Cancer Medicine*, 2019, 8(12): 5574-5576.
- [3] LEE S, CHO M, LEE W, *et al.* Predicting lymph node metastasis and distant metastasis using differential correlations of miRNAs and their target RNAs in cancer[C]. 2022 *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. Las Vegas, NV, USA. IEEE, 2022: 457-463.
- [4] OYAMA D, ADACHI Y, TSUYUGUCHI N. Magnetic resonance imaging using a magnetoresistive sensor with a flux transformer[J]. *IEEE Transactions on Magnetics*, 2023, 59(11): 5100505.
- [5] WANG Q, QIAO K, QIN R X, *et al.* Predictive model of breast cancer lymph node metastasis based on deep learning (E-Transformer)[C]. 2021 *6th International Symposium on Computer and Information Processing Technology (ISCIPIT)*. Changsha, China. IEEE, 2021: 168-173.
- [6] 周涛, 党培, 陆惠玲, 等. 跨模态跨尺度跨维度的 PET/CT 图像的 Transformer 分割模型[J/OL]. 电子与信息学报, 2022: 1-9. (2022-12-13). <https://kns.cnki.net/kcms/detail/11.4494.TN.20221213.1157.001.html>.
- [7] GURNEY-CHAMPION O J, KIESELMANN J P, WONG K H, *et al.* A convolutional neural network for contouring metastatic lymph nodes on diffusion-weighted magnetic resonance images for assessment of radiotherapy response[J]. *Physics and Imaging in Radiation Oncology*, 2020, 15: 1-7.
- [8] TONG Z X, KANG Y X, YING H, *et al.* A novel framework for lymphatic invasion detection based on efficient medical images translation network [C]. 2021 *6th International Conference on Intelligent Computing and Signal Processing (ICSP)*. Xi'an, China. IEEE, 2021: 669-673.
- [9] LUO Y W, XIN J M, LIU S J, *et al.* Lymph node metastasis classification based on semi-supervised multi-view network [C]. 2020 *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. Seoul, Korea (South). IEEE, 2020: 675-680.
- [10] ZHAO X Y, XIE P Y, WANG M M, *et al.* Deep learning-based fully automated detection and segmentation of lymph nodes on multiparametric-MRI for rectal cancer: a multicentre study [J]. *EBio-Medicine*, 2020, 56: 102780.
- [11] HE K M, ZHANG X Y, REN S Q, *et al.* Deep residual learning for image recognition [C]. 2016 *IEEE Conference on Computer Vision and Pattern*

- Recognition (CVPR). Las Vegas, NV, USA. IEEE, 2016: 770-778.
- [12] ATKINSON D, HILL D L G, STOYLE P N R, *et al.* Automatic correction of motion artifacts in magnetic resonance images using an entropy focus criterion[J]. *IEEE Transactions on Medical Imaging*, 1997, 16(6): 903-910.
- [13] 周涛, 赵雅楠, 陆惠玲, 等. 医学图像实例分割: 从有候选区域向无候选区域[J]. *生物医学工程学杂志*, 2022, 39(6): 1218-1232.
- ZHOU T, ZHAO Y N, LU H L, *et al.* Medical image instance segmentation: from candidate region to no candidate region[J]. *Journal of Biomedical Engineering*, 2022, 39(6): 1218-1232. (in Chinese)
- [14] BOLYA D, ZHOU C, XIAO F Y, *et al.* YOLACT: real-time instance segmentation [C]. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*. Seoul, Korea (South). IEEE, 2019: 9156-9165.
- [15] 周涛, 刘赞璨, 侯森宝, 等. REC-ResNet: 面向 COVID-19 辅助诊断的特征增强模型[J]. *光学精密工程*, 2023, 31(14): 2093-2110.
- ZHOU T, LIU Y C, HOU S B, *et al.* REC-ResNet: feature enhancement model for COVID-19 aided diagnosis [J]. *Opt. Precision Eng.*, 2023, 31(14): 2093-2110. (in Chinese)
- [16] JIAO J, QIN H W. Adaptively weighted balanced feature pyramid for object detection [C]. 2023 *IEEE 6th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*. Chongqing, China. IEEE, 2023: 1214-1217.
- [17] DIAO Y J, LIU S W, GAO X Z, *et al.* CNN based on multiscale window self-attention mechanism for radar HRRP target recognition [C]. 2022 *7th International Conference on Signal and Image Processing (ICSIP)*. Suzhou, China. IEEE, 2022: 281-285.
- [18] YANG Z X, ZHU L C, WU Y, *et al.* Gated channel transformation for visual recognition [C]. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, WA, USA. IEEE, 2020: 11791-11800.
- [19] ZHOU C. *Yolact++ Better Real-Time Instance Segmentation* [M]. Davis: University of California, 2020.
- [20] GAO Y J, WANG X Y, LI Q H. Improved cervical cell segmentation algorithm based on cascade mask RCNN [C]. 2023 *IEEE 7th Information Technology and Mechatronics Engineering Conference (ITOEC)*. Chongqing, China. IEEE, 2023: 1775-1780.
- [21] 梁新宇, 林洗坤, 权冀川, 等. 基于深度学习的图像实例分割技术研究进展[J]. *电子学报*, 2020, 48(12): 2476-2486.
- LIANG X Y, LIN X K, QUAN J C H, *et al.* Research on the progress of image instance segmentation based on deep learning [J]. *Acta Electronica Sinica*, 2020, 48(12): 2476-2486. (in Chinese)

## 作者简介:



周涛(1977—),男,宁夏同心人,博士,教授,2010年于西北工业大学获得博士学位,主要从事医学图像分析处理、深度学习、模式识别等方面的研究。E-mail: zhoutaonxmu@126.com

## 通讯作者:



石道宗(1999—),男,湖南湘西人,硕士研究生,主要从事智能医学影像图像处理,计算机辅助诊断方面等方面的研究。E-mail: shidaozong167@163.com