

同济大学大学生创新训练项目计划

申请书

项目编号 _____

项目名称 一种基于多专家机制的前列腺癌肿瘤分期分类方法 _____

项目负责人 高晗博 联系电话 15333687527 _____

所在学院 计算机科学与技术学院 _____

学号 2453585 专业班级 计算机拔尖班 _____

指导教师 倪张凯 _____

E-mail 15333687527@163.com _____

申请日期 2025 年 4 月 17 日 _____

项目期限 一年期 _____

同济大学 创新创业学院

填写说明

1. 本申请书所列各项内容均须实事求是，认真填写，表达明确严谨，简明扼要。
2. 申请人可以是个人，也可为创新团队，首页只填负责人。
“项目编号”一栏不填。
3. 本申请书为大 16 开本（A4），左侧装订成册。可网上下载、自行复印或加页，但格式、内容、大小均须与原件一致。
4. 负责人所在学院认真审核，经初评和答辩，签署意见后，将申请书（一式两份）报送同济大学项目管理办公室。

一、 基本情况

项目名称	一种基于多专家机制的前列腺癌肿瘤分期分类方法						
所属学科	学科一级门： 工学 学科二级类： 计算机类						
项目来源	<div><input type="radio"/> A、学生自主选题，来源于自己对课题的长期积累与兴趣</div> <div><input checked="" type="radio"/> B、导师推荐课题</div> <div><input type="radio"/> C、SITP 直升课题 （原课题名称： ）</div>						
项目申请级别	SITP						
申请金额	3000		项目期限		一年制		
负责人	高晗博	性别	男	民族	汉族	出生年月	2006年3月
学号	2453585	联系电话	宅： 手机： 15333687527				
指导教师	倪张凯	联系电话	宅： 手机： 13646019080				
项目简介	<p>为了解决前列腺癌针对超声影像的研究较少的问题，本项目于大模型开发一种方法研究以下内容：1. 设计针对前列腺超声影像病灶区域的遮掩自动编码预训练框架；2. 基于多专家优化特征学习策略，提升预测准确性和鲁棒性；3. 对构建的方法进行系统化验证，并开发临床可用的辅助诊断平台。创新点包括：1. 针对MAE 遮掩自动编码预训练中超声影像无关区域过多的特征设计的预训练流程；2. 开发适配不同长度超声数据的自适应分期预测框架；3. 实现低成本、高效的分期诊断，推动其在临床应用转化。</p>						
负责人曾经参与科研的情况	负责人曾参与 SITP 创新项目，在指导老师指导下参与课题						

指导教师承担科研课题情况		国家重点研发计划子课题，国家自然科学基金（青年）项目，上海市浦江人才计划项目，上海市领军人才（海外）青年人才计划项目			
指导教师对本项目的支持情况		指导教师每周和项目组成员开会，定期单独和项目负责人交流进度			
项目组 主要成员	姓名	学号	专业班级	所在学院	项目中的分工
	高晗博	2453585	计算机拔尖1班	计算机科学与技术学院	项目管理
	朱俊泽	2351114	大数据1班	计算机科学与技术学院	算法设计，模型搭建
	吉稼熠	2352976	大数据1班	计算机科学与技术学院	数据整理

二、 立项依据（可加页）

（1）研究目的

前列腺癌是全球男性中第二常见的癌症类型，也是导致癌症相关死亡的主要原因之一。而检测前列腺癌症使用 CT 则成本较高，使用超声影像则过于依赖专业医生的评估内容，本项目研究目的就是通过深度学习的方法，结合遮掩自动编码器预训练方法和下游微调的 MAE 框架利用前列腺超声影像视频对前列腺癌分期。

（2）研究内容

本项目主要以遮掩自动编码为核心结构，添加遮掩得分网络模块。应用了相关医学分割大模型，来解决随机遮掩的 MAE 中对关键病灶区域的表征学习的忽略，并且提高下游分类任务水平，实现人机协同。

（3）国、内外研究现状和发展动态

随着深度学习技术的普及，国内外不仅仅是前列腺癌症，很多很多的病情诊断也都渐渐有融合深度学习的方法进行自动诊断检测。问题是大多数基于深度学习的方法依赖医生提供的精标签，细到每一帧的分割。这样对医生提供的数据要求极高。视频质量参差不齐，标注过粗都能够影响到模型的精确性。

比如 VideoMAE 工作采用了基于遮掩自动编码 (Masked Autoencoder, MAE) 的自监督预训练方法，基于视频的 MAE 预训练流程如下图所示：

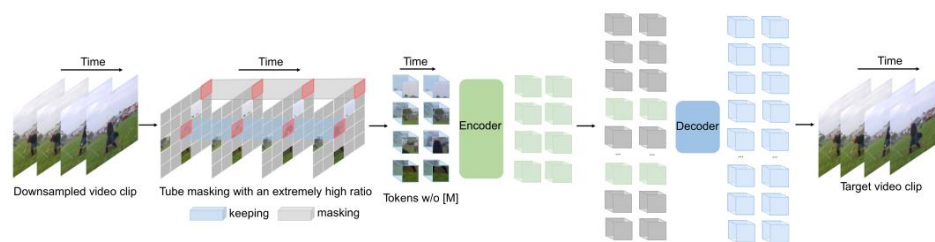


图 1.VideoMAE 预训练流程框架

将下采样帧作为输入，使用跨行时间采样策略来进行更有效的视频预训练。形式上，首先从原始视频 V 中随机采样一个由 t 个连续帧组成的视频剪辑。然后，使用时间采样将剪辑压缩为 T 帧，每个帧的分辨率为 $H \times W \times 3$ ，并使用立方体嵌入来获得视频令牌。其中我们将每个大小为 $2 \times 16 \times 16$ 的立方体视为一个令牌嵌入。因此，立方体嵌入层获得 $\frac{t}{2} \times \frac{H}{16} \times \frac{W}{16}$ 个三维令牌，并将每个令牌映射到通道维度 d 。这种设计可以降低输入的时空维度，有助于缓解视频中的时空冗余。

对于获得的三维令牌，VideoMAE 随机选择 90% 的信息进行遮掩。然后把没有遮掩的令牌输入到了视觉编码器 (Vision Encoder) 中，在 VideoMAE 工作中使用了视觉转换器 (Vision Transformer, ViT) 作为主干。最后利用解码器进行重建训练，优化重建损失 $\mathcal{L}_{\text{recon}}$ 。

随后针对微调过程，VideoMAE 利用预训练好的编码器，将三维令牌输入之前预训练好的视觉转换器 (Vision Transformer, ViT) 的视觉编码器 (Vision Encoder) 中获得特征，取最深维度的特征作为获取的特征输入分类头，通过分类损失训练分类头和视觉编码器

(4) 创新点与项目特色

本研究项目致力于设计遮掩得分网络，来迫使模型在遮掩自动编码的预训练任务中学习病灶区域的表征，提高重建学习内容和效率。

本发明提出了一种基于多专家的大模型丰富基础知识指导的可泛化的前列腺癌症 T0/T1 期诊断系统。该方法针对 MAE 流程中随机遮掩的重建预训练未让模型学习到关键区域的特征问题，设计了一个基于病灶分割内容的遮掩得分网络。迫使模型学习关键区域的特征也就是关键的局部特征。

在下游微调时，利用在重建训练阶段预训练好的遮掩得分网络，将输入的视频中的重要区域的块作为额外提示添加给特征辅助分类头训练。来解决多医生采样之间的个体差异、诊断风格和数据质量波动带来的模型效果问题。

(5) 技术路线、拟解决的问题及预期成果

技术路线：

步骤 1 数据预处理

将视频切分为多个等长的剪辑片段，片段长度固定。将片段中的每一帧通过 MedSAM2, Deeplabv3, nnUNet2++, 三个预训练的分割大模型获得三个模型对帧的分割结果。

步骤 2 遮掩得分生成

将整个视频通过块嵌入 (Patch Embedding) 获得多个三维令牌 (Token)，将三维令牌输入我们设计的遮掩得分网络，该网络由三个独立的多头注意力块 (Multi-Head Attention, MHA) 和全连接层 (Fully Connected, FC) 组成，三个全连接层在输出各自对每一个令牌的得分之后，通过之前获得的三个分割图像，若某一个令牌是在分割图内的话，那么这个遮掩得分就会被加上一个值。以此来达到核心区域遮掩的目的。迫使模型学习核心病灶区域的特征。

步骤 3 预训练编码器

受 VideoMAE 设计的启发，我们开始先预训练编码器，在这里我们选择的是视觉转换器基础版本 (Vision Transformer-base, ViT-B) 作为视觉转换器编码器解码器。通过步骤 2，我们获得了每一个令牌的得分，我们将前 90% 排名的令牌遮掩，之后输入编码器，获得编码过的令牌，再将之前遮掩的令牌添加之后输入解码器，解码器重建出原图像计算重建损失训练 $\mathcal{L}_{\text{recon}}$ 来更新视觉转换器，通过重建内容和原内容差距和遮掩得分网络计算采样损失 $\mathcal{L}_{\text{sample}}$ 来更新遮掩得分网络。

步骤 4 微调编码器、分类头

在经过步骤 3 预训练好视觉编码器和遮掩得分网络后，我们冻结遮掩得分网络，开始进行下游分类任务中编码器和分类头的微调。对于带有标签的少量数据，我们将视频分出的剪辑经过遮掩得分网络后获得每个令牌的得分，将前 10% 的令牌作为额外的提示 (Prompt) 添加给令牌序列，输入预训练好的视觉编码器之后将获得的特征输入给分类头进行分类，通过交叉熵来更新分类头和编码器。

拟解决的问题：

与现有技术相比，本发明具有以下有益效果：

(1) 预训练阶段重建针对性：本发明提出的基于医学分割模型先验知识引导的遮掩重建的预训练策略对不同医生操作的不同数据具有良好的适应性。

(2) 下游微调阶段额外提示针对性：本发明提出的基于预训练过程中训练的遮掩得分网络中重要信息添加作为额外提示，对不同医生操作的数据做下游分类时也能表现出较好的适应性。

预期成果：

1. 通过改进的遮掩得分网络预训练出学习超声影像特征更好的模型
2. 能够设计出上线的应用平台，通过方便的操作来使用模型，同时通过医护人员使用的病例继续优化模型。
3. 相关专利和论文的产出
(6) 项目研究进度安排
第一阶段：（2025.4-2025.7） 阅读论文，研究现有方法
第二阶段：（2025.8-2025.9） 前处理训练，数据集收集，模型复现
第三阶段：（2025.10-2025.12） 模型搭建，算法设计，微调模型
第四阶段：（2025.12-2026.4） 应用平台开发，落地
(7) 已有基础
专利基本撰写已经完成，模型设计和基本实验也已经完成，一些简单的对比实验也已经完成，在指导教师的指导下已经完成了对现有方法的初步调研。
(8) 尚缺少的条件及解决方法：
还有相当的对比实验和消融实验未进行，受到一些算力限制实验进度推进不快。可以使用相关经费租赁算力推进实验。

三、 经费预算

开支科目	预算经费（元）	主要用途
预算经费总额	3000	
专用设备费(购置、维修、租赁)	2000	服务器设备租赁
材料费	500	相关书籍、网盘、腾讯会议等平台会员
测试化验加工费	0	
差旅费	500	校区交通
会议费	0	
国际合作与交流费	0	
出版/文献/信息传递/知	0	

识产权事务费		
劳务费(含专家咨询费)	0	

四、 项目组成员签名

朱俊泽

高晗博

吉缘熠

五、 指导教师意见

同意。

导师（签章）：

倪张甄

2025 年 4 月 17 日

六、 院系大学生创新创业训练计划专家组意见

教学负责人（签章）：

年 月 日

七、 学校大学生创新创业训练计划专家组意见

负责人（签章）：
日

年

月