**技术交底书**

**发明名称： 多专家机制的前列腺癌症T0/T1期分类**

**本专利发明人： 倪张凯 朱俊泽 肖润宇 王瀚漓 ­**

**技术交底书撰写人及技术联系人：\_倪张凯**

**电话： 13646019080 FAX: \_ \_ E-MAIL：\_zkni@tongji.edu.cn \_**

注意：

1.**第2部分和第4部分**，**一定要写的全面、清楚**。

2.英文缩写要有中文译文和英文原词，最好在术语解释部分给出。

3.**全文对同一事物的叫法应统一，避免出现一种东西多种叫法。**

4.应该阐述发明目的是通过什么技术方案来实现的，不能只有原理，也不能只做功能介绍。

**术语解释：**

解释一些跟技术方案有关的专业术语

1. **本发明要解决的技术问题是什么？**（务必明确一个最主要的技术问题）

**2、详细介绍技术背景,并描述已有的与本发明最相近似的实现方案**（与本发明最接近的技术方案的说明（对于方法，应说明现有方法的步骤，对于装置，应当说明结构组成及其关系））

**3、以因果关系推理的方式推导出现有技术的缺点是什么？**

**针对这些缺点，说明本发明的目的。**（缺点可以有多个，但一定要分清主次。最主要的缺点原则上只能有一个。）

**4、本发明技术方案的详细阐述，应该结合流程图、原理框图、电路图、时序图进行说明**（所有英文缩写都应有中文注释；**所有附图都应该有详细的文字描述，以别人不看附图即可明白技术方案为准**；同时附图中的关键词或方框图中的注释都尽量用中文；方法专利都应该提供流程图，并提供相关的系统装置）。

**5、本发明的关键点和欲保护点是什么？**（对于上一部分给出的详细的完整技术方案，在本部分是提炼出技术方案的关键创新点，列出1、2、3...，以提醒代理人注意，便于专利代理人撰写权利要求书）

**6、用推理方式推导出本发明的优点**（务必与第3部分的现有技术的缺点相对应），可以对应3部分所要解决的技术问题或发明目的来描述。

**7、针对4中的技术方案，是否还有别的替代方案同样能完成发明目的？**

**8、其他有助于专利代理人理解本技术的资料**（给代理人提供更多的信息，可以有助于代理人更好更快的完成申请文件）

1. **案例（实施例）**

**说明书**

**多专家机制的前列腺癌症T0/T1期分类**

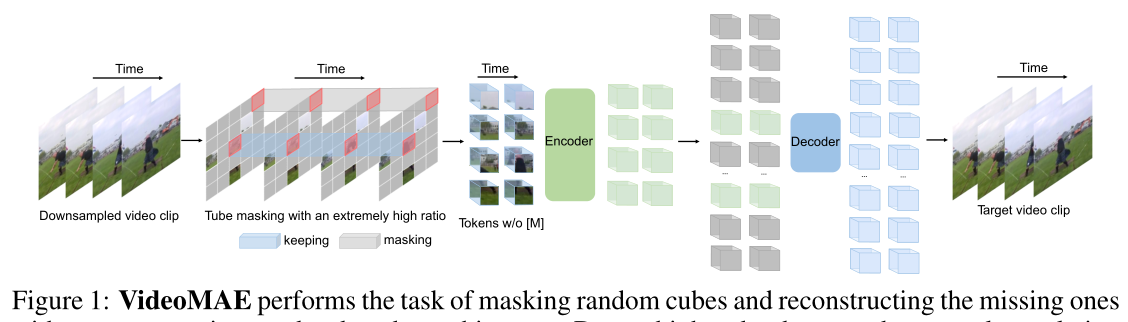
**技术领域**

本发明涉及前列腺癌症T0/T1分期的诊断领域

**背景技术**

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。在《中国前列腺癌筛查与早诊早治指南（2022年）》中，强调了提高筛查效果和规范性。尤其是在低资源地区，低成本且高效的诊断方法尤为关键。作为现代医学诊断的重要工具，医疗影像技术广泛应用于疾病检测中，并发挥着不可替代的作用。医疗影像技术包含磁共振成像（Magnetic Resonance Imaging, MRI）、计算机断层扫描（Computed Tomography, CT）、X光成像以及超声成像技术等。这些技术提供了宝贵的视觉数据，帮助医生诊断和评估疾病。在前列腺癌的诊断中，MRI是最常用的影像学工具。MRI具有较高的特异性，可用于前列腺癌的T分期预测，但其敏感性较低，且受制于价格、禁忌症及设备普及等因素，难以普及到广泛的临床应用中。作为另一种常见的影像诊断工具，超声成像在前列腺癌检测中也有所应用。与MRI相比，超声成像具有成本低、适用人群广和对人体危害小的优势，更适用于低成本且高效的前列腺癌检测。但前列腺癌多发、散灶的特征使得单张影像的诊断价值降低，包绕整个前列腺腺体及周围组织、器官的超声视频弥补了单张影像诊断的不足。然而超声视频成像存在分辨率低的问题，导致难以准确进行诊断。此外，超声的敏感性和特异性存在较大波动，尤其在小病灶和边界模糊的情况下，诊断的准确性依赖于医生的个人经验。因此，如何有效提升超声视频的分辨率，以提供更加清晰且高保真的影像，并提高医生对小病灶及边界模糊等复杂区域的诊断准确性，已成为当前基于超声成像开展前列腺癌诊断的重要研究方向和技术发展重点之一。

自监督学习（Self-Supervised Learning, SSL）是一种机器学习方法，它不依赖于人工标注的数据，而是通过数据本身生成监督信号，在医学中就有了相当的优势，它不消耗医生资源进行额外多标注。这种方法利用数据的内在结构和特性，从无标签的数据中学习表示。目前基于遮掩自动编码(Mask Automation Encoder, MAE)的自监督预训练方法在视频中也逐渐开始运用，比如VideoMAE工作，基于视频的MAE工作如下流程所示：



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，每个帧包含H × W × 3像素，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 × 16 × 16的立方体视为一个令牌嵌入。因此，立方体嵌入层获得个三维令牌，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。

然后对于获得的三维令牌，VideoMAE随机选择90%的信息进行遮掩。然后输入到了视觉编码器中，在VideoMAE工作中使用了视觉转换器(Vision Transformer, ViT)作为主干。然后利用解码器进行重建训练。

其中针对前列腺癌症超声影像的数据方面有如下问题：

1. ：前列腺超声影像的数据中存在大量的无关信息，包括无病灶的区域。随机遮掩会导致模型学习过程中对局部结构的忽略，这在医学影像中尤其重要，因为局部特征往往决定了疾病的早期发现和准确诊断。MAE在遮掩了过多非关键的局部特征时，模型对真正关键的局部特征重建特征的学习就会相对的差。
2. 对医工落地方面，模型对多中心的数据泛化能力不强。不同医院的数据采集设备不一，视频数据中病灶集中区域不一，模型在一个中心学习到的时序特征和局部特征在另一个中心验证时模型性能扰动严重。

**发明内容**

为了解决上述问题，本发明提出了一种基于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。并且在下游微调时，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。

技术方案：

一种基于大模型丰富基础知识指导的重要区域遮掩重建的前列腺癌症T0/T1分期系统设计方法。包含以下步骤：

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

和VideoMAE设计的核心一样，我们开始预训练编码器，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。

**具体实施方式**

下面结合附图和具体实施例对本发明进行详细说明。本实施例以本发明技术方案为前提进行实施，给出了详细的实施方式和具体的操作过程，但本发明的保护范围不限于下述的实施例。

一种基于大模型丰富基础知识指导的重要区域遮掩重建的前列腺癌症T0/T1分期系统设计方法。如图1所示，包括如下具体步骤：

步骤1：数据预处理

针对有T帧的视频V，我们分割为多个等长剪辑，长度l设定为16，每个剪辑采用均匀步长的采样，具体来说我们设定采样间隔为step = ，从第1帧到第step帧。其中每一帧开始往下均匀的采样，每隔step帧就采样进入本剪辑内，最终获得多个等长剪辑片段：

此时经过预训练的分割模型MedSAM2，Deeplabv3，nnUNet2++我们获得对应的分割内容：

步骤2：遮掩得分生成

我们采用联合时空的立方嵌入(Embedding)，具体来说对于一个剪辑Clip大小为，我们在时间上采样2单位，空间上采样单位，就获得了N=()个大小为的块，将每个令牌映射成d=768维度。这样做来缓解时空冗余。

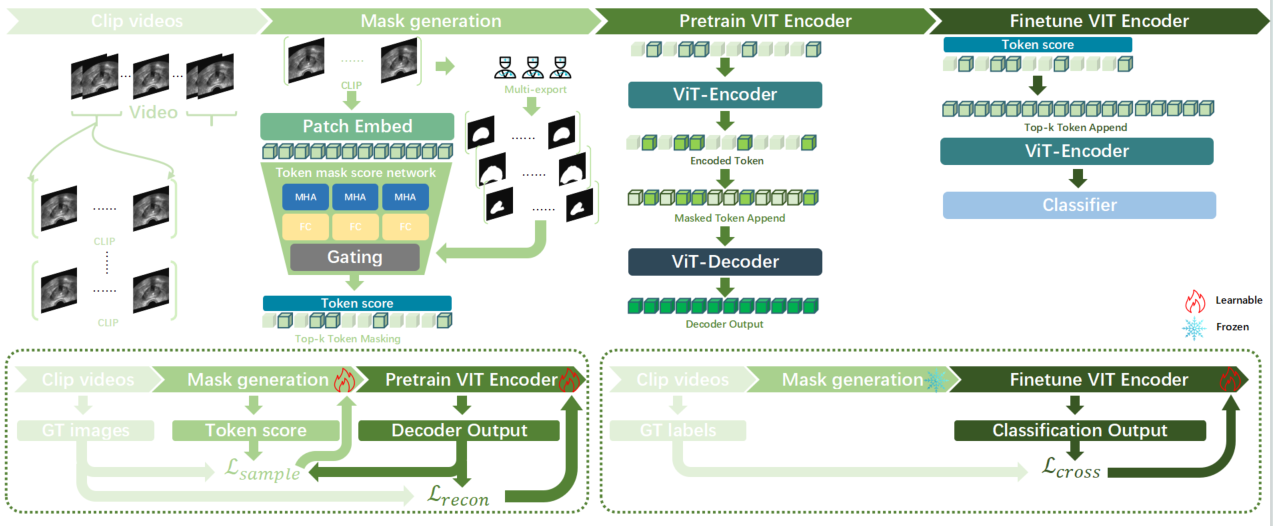
对于这些令牌，我们输入遮掩得分网络，遮掩得分网络分为3个独立的多头注意力块和全连接层，具体计算如下：对于已经嵌入的令牌

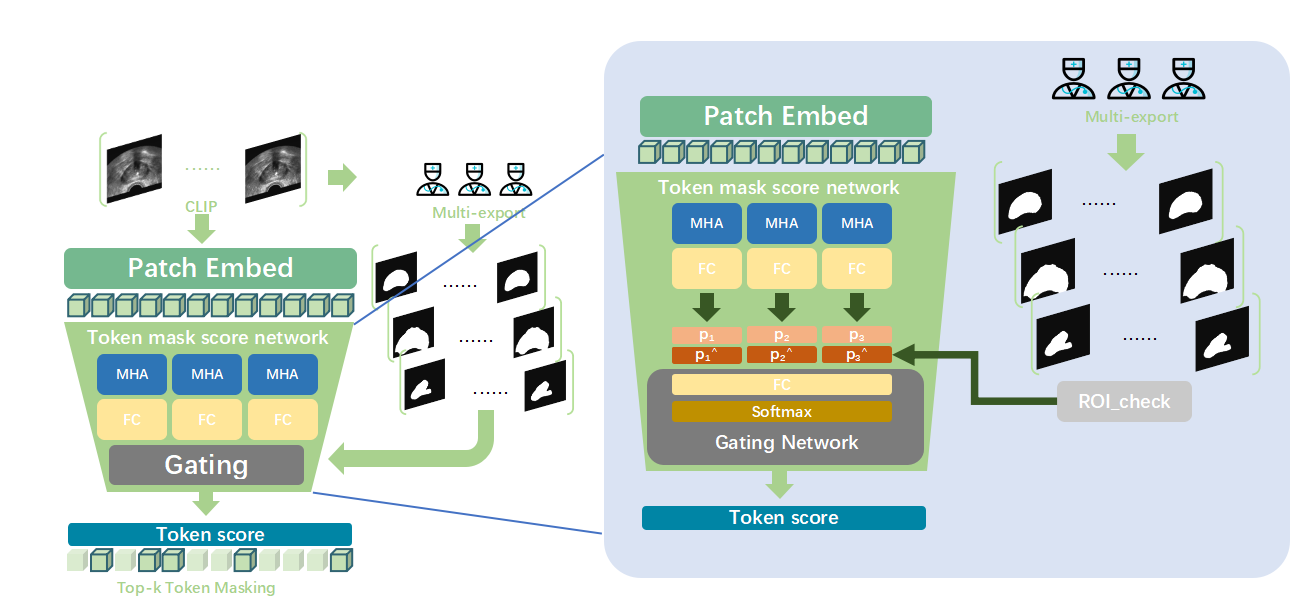
**附图说明**

图1 本发明整体框架流程图

图2 遮掩得分网络具体流程图

**应用实施例**





说明书附图