**技术交底书**

**发明名称： 大模型辅助的前列腺癌症T分期系统**

**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**

**大模型辅助的前列腺癌症T分期系统**

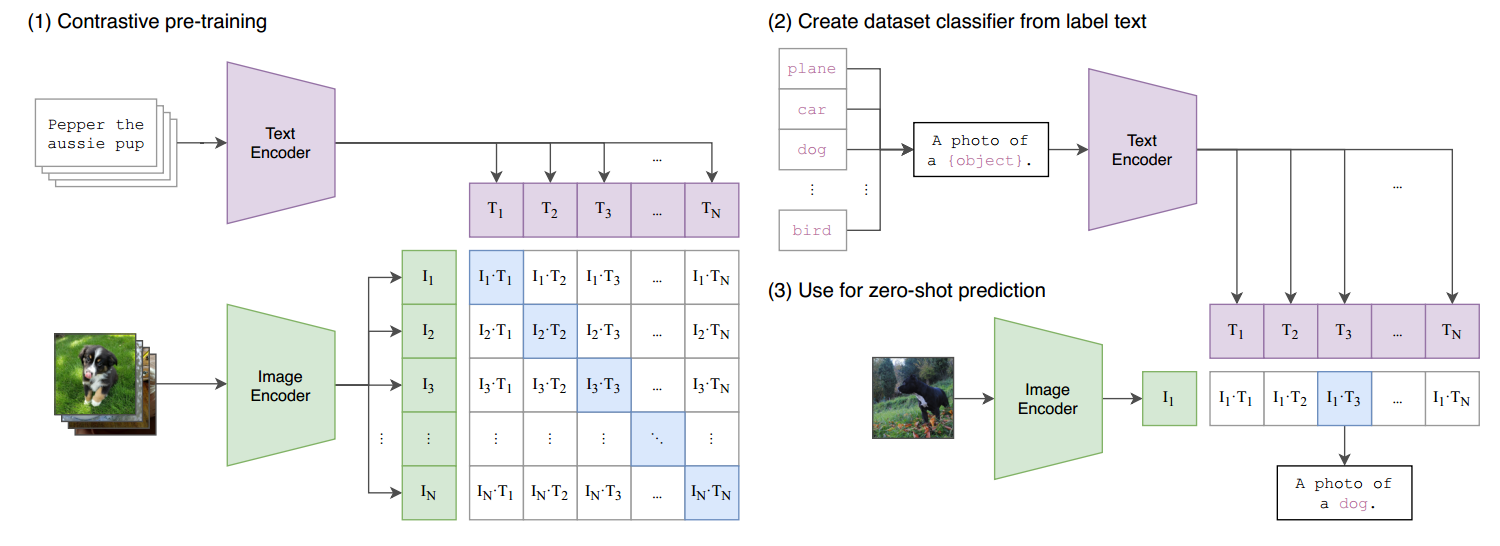
**技术领域**

本发明涉及前列腺癌症T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。

基础的clip图片-文本对齐的方法为医学信息和医学图像对齐提供了一种对齐思想，应用到视频方

****

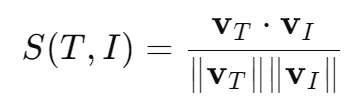
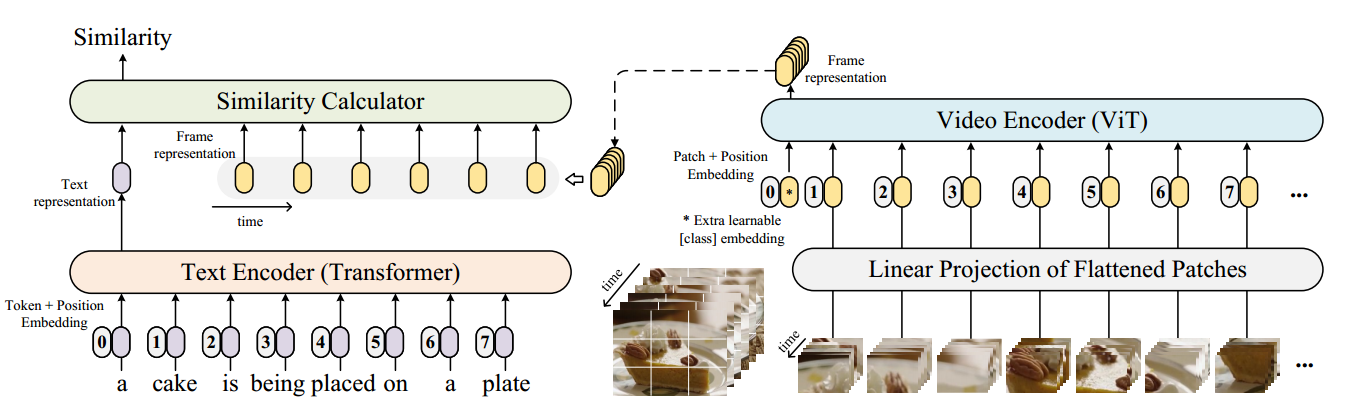
模态中需要提取每一帧和文本做对齐，对于视频特征使用ResNet或者Vision Transformer，I代表输入的图像，VI代表提取出的特征，维度为。n为batch size，h为图片高度，w为图片宽度，c为图片通道数。

（1）

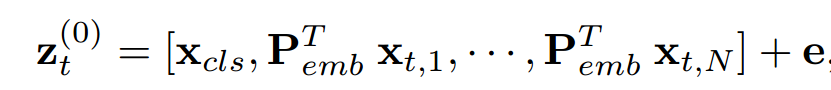
在文本特征的提取中，使用预训练的文本编码器，T代表图像对应label的文本，VT表示获取的文本特征，特征维度为(n,l)。n为batch size，l为文本长度。

（2）

获取两个模态的特征后通过对比学习拉进对应图像和文本的相似度，其中S(T,I)表示文本T和图像I的相似度，VI⋅VT代表向量点积，||VI||和||VT||为各自向量的欧几里得范数。最后通过InfoNCE损失更新参数。

（3）

那么在视频中我们考虑对t帧采样，对一个视频video 其中T代表采样的帧数，H和W代表视频帧的尺寸，按照VIT的要求要把向量安装PxPx3的大小扁平化成N个向量,加入了额外的Xcls之后整个每帧即将输入VIT的token如下

（4）

其中e代表时间和空间的位置编码。Pemb代表Linear Projection层的参数。之后取出class token作为图像表征并按照时序排序和文本特征做特征相似度对比。

其中针对前列腺癌症超声影像的数据方面有如下问题：

1. ：需要数据过于精细，详细到每一帧的医学信息描述，这需要消耗大量医生资源。反之如果标签过粗，仅对整个视频执行描述，模型学习到的特征和文本描述对齐的效果就不是非常的理想。在前列腺的超声影像数据采集的过程中，有相对长的寻找病灶部位的过程，这个过程的无关信息被模型捕捉，这造成的一定的干扰。
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语言模型丰富基础知识的可泛化的前列腺癌症T期诊断系统。该方法不需要医生额外的对视频中的每一帧图像进行标注，通过大语言模型分割能力，设计了一个关键帧提取和相似度计算的方法，针对视频中病灶集中的区域加强时序理解。对无病灶内容也就是无关信息进行丢弃，压缩无关信息，初步筛选粗粒度标签下的超声视频，在时序上重点关注病灶区域。避免视频中不存在病灶信息的帧对前列腺癌肿瘤分期分类器训练的干扰，同时将分割大模型的知识迁移到分类器中。

技术方案：

一种基于大模型丰富知识的前处理数据的前列腺癌症T分期系统设计方法。包含以下步骤：

步骤1 数据预处理

将视频进过预训练的大模型MedSAM2，得到分割结果，对有分割结果的视频进行滑动窗口采样，获得一定组数的数据。

步骤2 计算关键帧和相似度

通过对整个视频的分割结果的解析，用聚类的方法选择出n类，每一类的中心帧选择为本类的关键帧，针对几帧的病灶排序选择唯一关键帧。再通过每一帧和关键帧之间pixel和IOU和帧之间距离计算出相对相似度。

步骤3 压缩无关帧

针对无关信息，在模型训练的时候需要执行丢弃和重新标注。针对全部帧无分割内容的帧或者全部帧相似度低于阈值的窗口，执行丢弃。

说明书附图

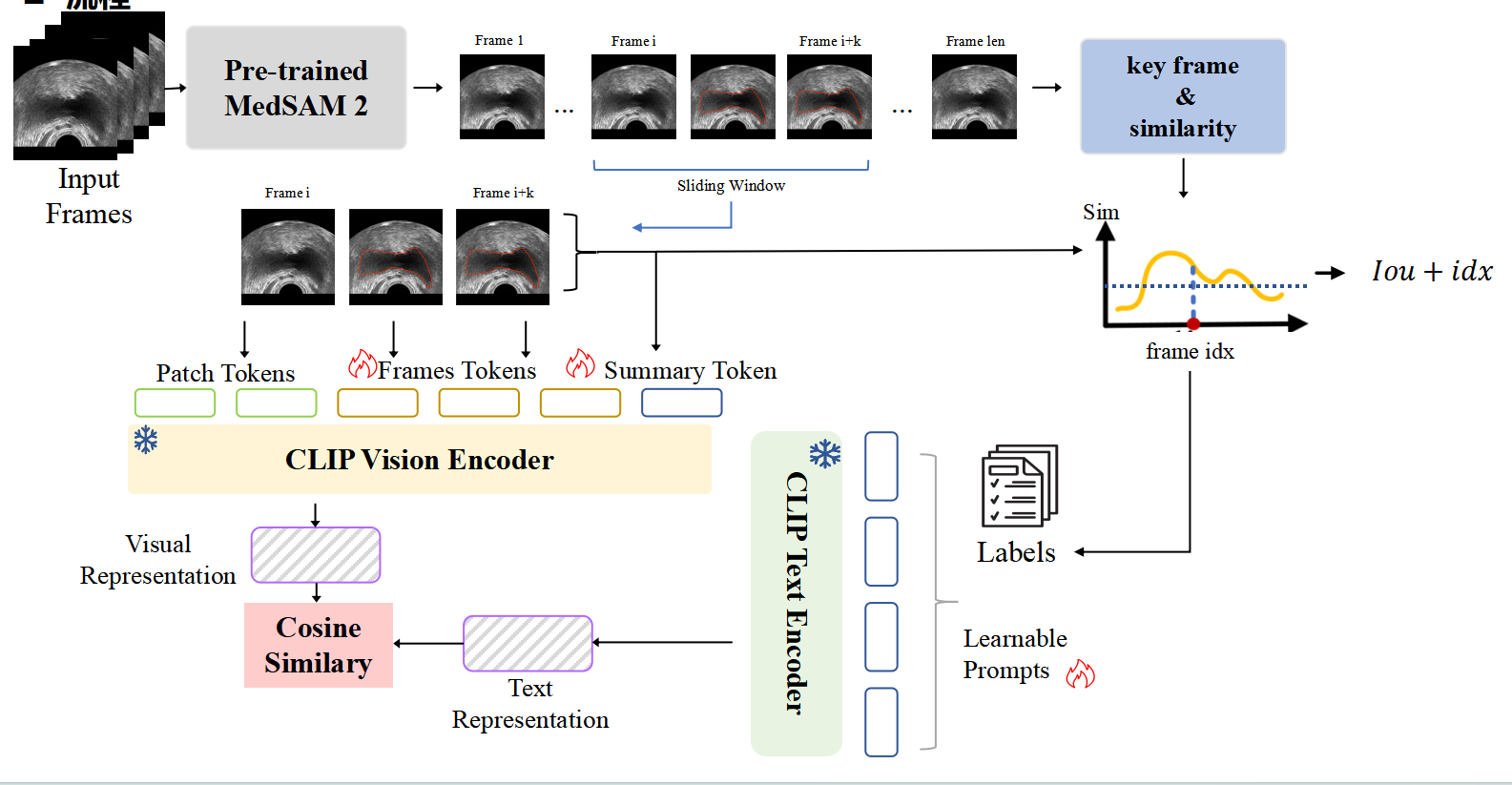


图1