# 面向超大尺寸病理图的高效标注平台

## 系统背景:

随着医疗信息化的快速发展，数字化病理图像（Whole Slide Image, WSI）逐渐成为医疗诊断和科研领域的重要数据来源。病理图像的分辨率极高、信息量极为丰富，通常由上亿像素组成，单张图像体积可达数百GB。这类图像被广泛用于疾病诊断、癌症筛查、病变分析、以及药物研发等多个场景，是推动精准医疗和个性化治疗的关键数据基础。然而，面对如此超大尺寸和复杂结构的病理图，传统的标注方法面临诸多挑战，迫切需要开发高效且灵活的病理图标注平台，以提高数据标注的准确性和效率。

病理图标注是医疗数据管理和算法开发的重要环节。由于病理图像的专业性和超高分辨率，标注工作不仅对工具的性能提出了极高要求，同时还需要兼顾标注精度与操作便捷性。具体挑战如下：

1）数据量巨大，无法直接加载。传统图像标注工具多面向小型图片设计，而单张 WSI 图像的数据体量极大，直接加载和操作会占用大量内存，并导致系统卡顿甚至崩溃。

2）细粒度特征复杂，操作精度高。病理图像中不同的细胞组织类型（如肿瘤区域、坏死组织、正常组织）常呈现相似的颜色或形态，需要进行高度精细化标注。标注时不仅要区分微小的病变区域，还要避免漏标、误标。

3）动态范围和层次信息处理困难。病理图像包含大量的层次化信息，如细胞级和组织级的特征，需要不同的缩放比例来观察和标注。然而，在不同尺度间频繁切换容易导致信息断层和标注误差。

4）医生标注成本高。病理图像标注需要专业医生参与，并且由于需要长时间高精度操作，工作强度高且效率较低。特别是标注肿瘤或细胞类型时，容易因疲劳造成失误。

现有的标注工具（如 VGG Image Annotator、LabelMe 等）虽然在自然图像和小型医学图像领域得到广泛使用，但在超大尺寸病理图标注中缺乏对超高分辨率病理图像的支持，无法实现流畅的缩放和浏览，且操作延迟明显，加载和浏览大尺寸图像时，工具的响应速度慢，医生在切换视图和标注时容易中断思路。

为了应对上述挑战，病理图高效标注平台需要在系统设计和功能上进行针对性优化。关键需求包括：

1. 病理图像过大无法一次性加载，因此系统应采用分块加载策略，按需加载局部区域，并支持多层次缩放，使医生可以无缝地切换不同视角。同时，需实现快速渲染以保证流畅浏览。
2. 平台应支持多种标注工具（如多边形、自由手绘、矩形框等），帮助医生根据需要对病理特征区域进行标注。此外，应提供标签管理功能，方便对不同类型区域进行分类与统计。
3. 病理数据常来自不同的医院、科室或不同的病种研究项目。平台需支持按数据集、分组对病理图进行分类管理。例如，将肺癌某个病人的多张切片数据归入到肺癌数据集下改用户分组中。

## 主要功能

1. 用户登陆注册功能。
2. 数据集创建（同时为数据集添加标签）、编辑、删除、条件查询。
3. 数据集内分组创建、编辑、删除、条件查询。
4. 分组内病理图上传、移动分组、编辑图像元信息、删除、查询。
5. 标注文件创建、更新、下载、删除、条件查询。
6. 病理图数据转化为deepzoom格式用于展示，转化任务创建、查询、任务进度更新。
7. 病理图金字塔层级展示用户可以从宏观视角（整体图像结构）逐渐缩放到微观层级（单个细胞或病变区域），实现细致观察（见图一）；通过优化的分块加载技术，用户缩放或拖动视图时无延迟，实现流畅操作；支持在不同层级间快速跳转，并通过缩略图导航条辅助用户找到感兴趣的区域；在金字塔展示下进行的标注会自动记录所在层级，并同步到其他层级视图，确保不同分辨率下的数据一致性。
8. 支持多种形状（点、矩形、椭圆、多边形和自由路径）标注，标签颜色映射以及单独图像自定义颜色标签，对标注内容添加备注，单张图像多个标注文件标注内容任意切换更新。

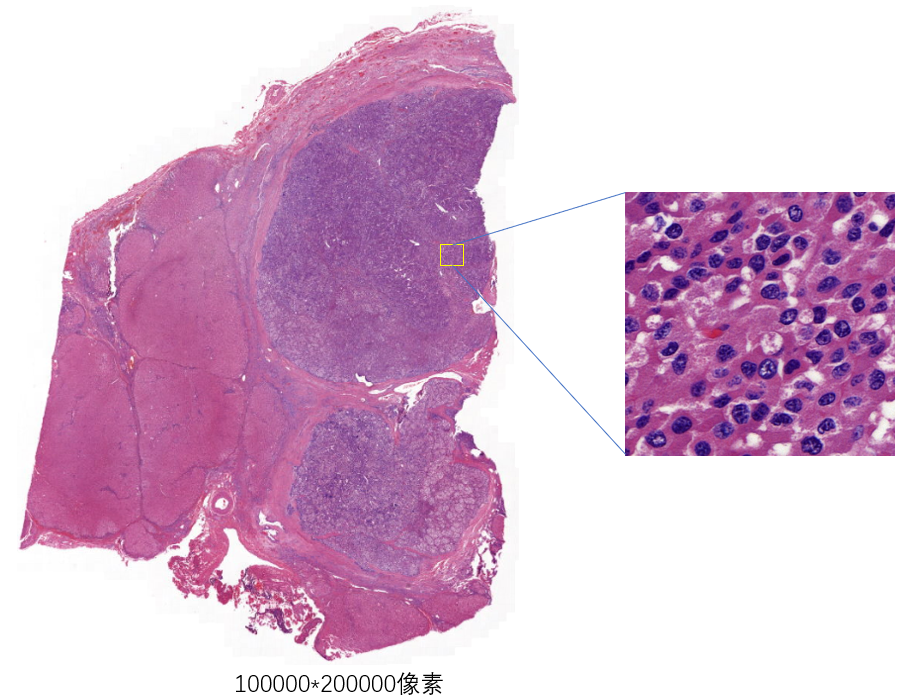


图 1 肝癌病理图像样例

## 开发环境

操作系统及版本：Ubuntu 18.04.5

支撑软件及版本列表：JDK17、Maven3.9.8、SprintBoot8.0.37、RabbitMQ3.13.6、Python3.8、Flask3.0.2、Nginx1.20.1、Nodejs14.17.3、React17.0.2、openseadragon2.4.2、fabric4.5.1

数据库环境：mysqlRedis6.2.14、MongoDB6.0.16、MySQL8.0.37

## 基本的架构体系

面向超大尺寸病理图高效标注平台是一个专门为病理图像提供高效标注的系统，整合多种标注工具和数据管理模块以提升标注效率。主要包括：用户管理、数据集管理、分组管理、图片数据管理、标注数据管理和任务管理六大模块。

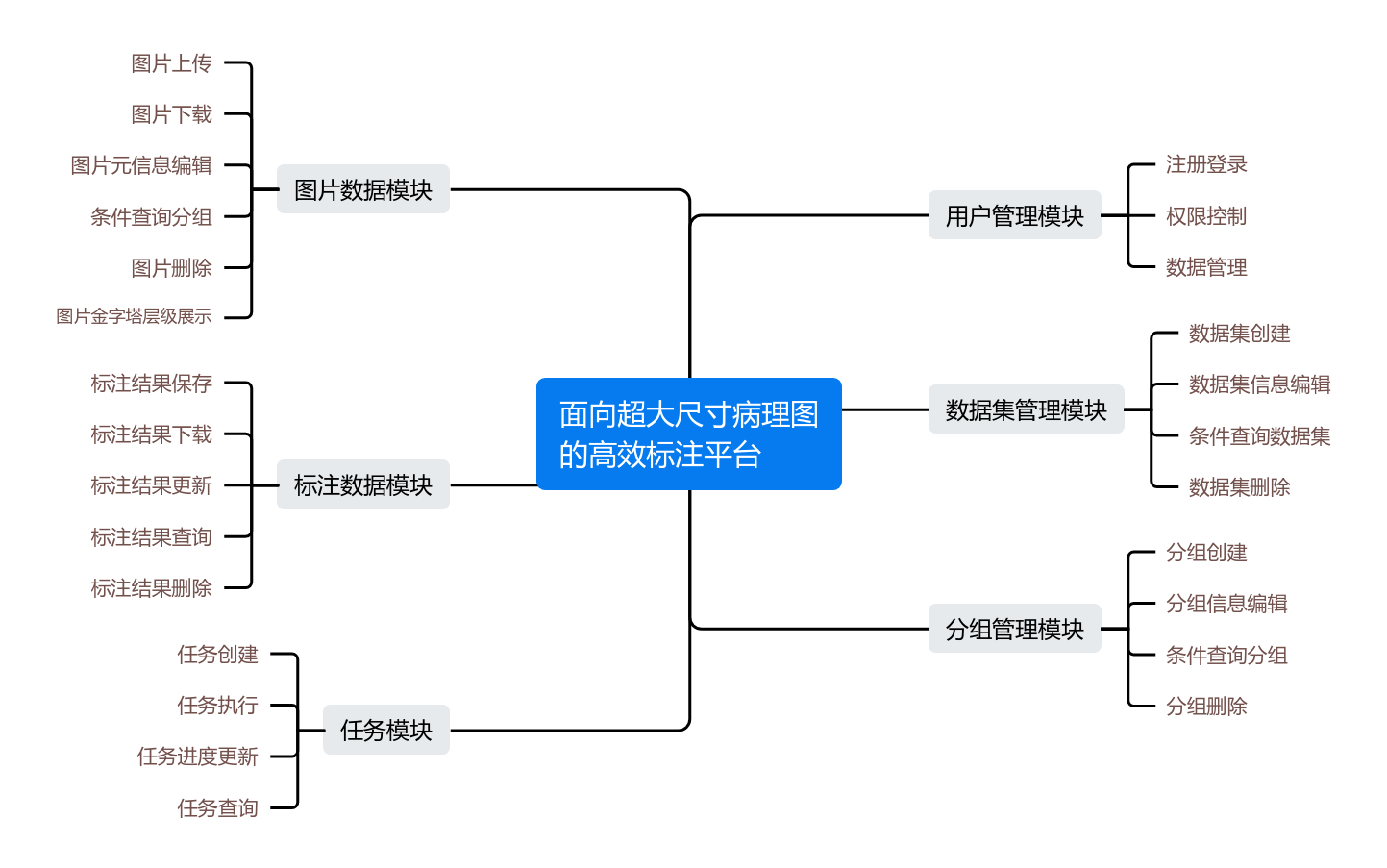


图 2 面向超大尺寸病理图高效标注平台功能结构图

图片数据模块：负责图片的上传、下载、元信息编辑、按条件查询分组以及图片的分块查看和显示，确保对超大尺寸病理图像的高效管理和展示。

标注数据模块：处理标注结果的保存、下载、更新、查询和删除，帮助用户有效管理病理图像的标注信息。

任务模块：提供任务创建、执行、进度更新和查询功能。

用户管理模块：负责用户的注册登录、权限管理，以及相关数据的管理。

数据集管理模块：支持数据集的创建、编辑、条件查询和管理，帮助用户按不同项目或类别管理海量的病理图数据。

分组管理模块：提供分组的创建、编辑、条件查询和管理，方便对数据集进行更细致的分组和管理。