

**毒性病理学图像数据标注的
质量控制要求（征求意见稿）**

**Quality Control Requirements for
Image Data Annotation in
Toxicologic Pathology**

目 次

前言	II
引言	III
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 人员要求和主要职责	2
5 数据采集与预处理	3
6 图像数据标注的任务实施及质量控制	4
7 数据的安全管理及存储	5
附录 A（规范性） 病理图像扫描质量判定条件	7
附录 B（资料性） 毒性病理学典型病变标注及工具操作示例	9
附录 C（规范性） 业务架构示例	13
参考文献	14

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本标准不涉及专利。

本文件由中国药学会提出。

本文件由中国药学会归口。

本文件起草单位：中国药学会毒性病理专业委员会。

本文件主要起草人（排名不分先后）：

林 志（中国食品药品检定研究院）	屈 哲（中国食品药品检定研究院）
王 浩（中国食品药品检定研究院）	靳洪涛（中国医学科学院药物研究所）
王和枚（康龙化成（北京）生物技术有限公 司）	胡春燕（国家成都新药安全性评价中心/成都华 西海圻医药科技有限公司）
吕建军（湖北天勤鑫圣生物科技有限公司）	孔庆喜（青岛德朔生物技术有限公司）
陆姮磊（中国科学院上海药物研究所）	邱 爽（国家成都新药安全性评价中心/成都华 西海圻医药科技有限公司）
尹纪业（军事医学研究院）	黄明姝（上海中医药大学）
杜 牧（北京昭衍新药研究中心股份有限公 司）	吴晓静（北京赛赋医药研究院）
杨 沁（北京协和建昊医药技术开发有限责 任公司）	戴锦龙（广东莱恩医药研究院有限公司）
霍桂桃（中国食品药品检定研究院）	李双星（中国食品药品检定研究院）
杨艳伟（中国食品药品检定研究院）	张 頔（中国食品药品检定研究院）
李言川（湖北天勤鑫圣生物科技有限公司）	李一昊（湖北天勤鑫圣生物科技有限公司）
范 维（北京协和建昊医药技术开发有限责 任公司）	何 杨（国家成都新药安全性评价中心/成都华 西海圻医药科技有限公司）
韩 刚（北京赛赋医药研究院）	

引 言

毒性病理学评价是药物安全性评价的关键环节,直接影响创新药物的研发成败。当前,人工智能(AI)在病理学中的应用日益广泛,尤其在早期识别、亚型分类及预后预测领域发挥重要作用。然而,公开的动物数据资源稀缺,且存在物种差异、标注体系缺失等问题。为规范图像数据标注流程,建立高质量、可审计的标准化病理数据集,特制定本文件,旨在提高病理诊断的准确性和一致性。

毒性病理学图像数据标注的质量控制要求

1 范围

本文件规定了毒性病理学人工智能（AI）诊断系统中图像数据标注的人员要求、标注规则、工具及环境、具体操作步骤、任务实施与质量控制，以及数据的安全管理及存储要求。

本文件适用于通过数字病理扫描仪获得病理图像数据，并借助病理图像分析软件所带的标注工具进行数据标注工作。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 42755-2023 人工智能 面向机器学习的数据标注规程

YY/T 1833.1—2022 人工智能医疗器械 质量要求和评价 第1部分：术语

YY/T 1833.3—2022 人工智能医疗器械 质量要求和评价 第3部分：数据标注通用要求

DB21/T 4419-2026 组织病理学全切片图像标注技术规程

3 术语和定义

YY/T 1833.1—2022、YY/T 1833.3—2022、DB21/T 4419-2026界定的以及下列术语和定义适用于本文件。

3.1

病理专业人员 Pathologist

具有临床医学或兽医学等相关专业教育背景，接受过系统的病理学专业培训，具备相应的资质或专业认证，能够独立进行大体及组织病理学检查、诊断、数据分析与评价的专业技术人员。在图像数据标注流程中，病理专业人员主要负责依据标准执行病理图像的精准标注、特征勾画、数据核查及质量把控。

3.2

标注工具 Annotation Tool

供病理专业人员对病理全切片图像进行全景浏览、感兴趣区域的勾画、分类标识以及结构化语义信息录入等操作的专用软件系统或计算机程序。

注：标注工具通常具备图像多层次缩放、多边形或自由曲线绘制、标签体系管理及数据按特定格式导出的功能，是实施图像数据标注全流程的必备载体。

3.3

数据标注 Data Annotation

借助标注工具，由标注人员或辅助算法对病理图像数据中的特定区域、组织结构、细胞或病灶等目标进行识别、分类、位置勾画，并添加结构化语义信息的过程。[来源：GB/T 42755-2023，有修改]

3.4

初级标注人员 Primary Annotator

具备毒性病理学诊断基础知识，主要负责图像类别确定以及病变区域勾画标注的人员。

3.5

审核人员 Reviewer

具备丰富的毒性病理学诊断知识与实践经验，主要负责对初级标注人员完成的任务进行审核的人员。

3.6

仲裁人员 Arbitrator

具备资深毒性病理学诊断经验的毒性病理学家，主要负责对标注与审核过程中存在分歧的疑难切片进行最终审核与仲裁的人员。

4 人员要求和主要职责

4.1 总体要求

标注人员应具备毒性病理学诊断的初级知识或高级知识，能够独立执行或审核标注任务。标注团队应分为三类人员：初级标注人员、审核人员和仲裁人员。建议改为标注团队应包含三类人员：初级标注人员、审核人员与仲裁人员。其中，初级标注人员需具备毒性病理学诊断的基础知识，能够独立执行标注任务；审核人员需具备丰富的经验，能够独立完成标注结果的审核。

4.2 初级标注人员

应具备初级职称或在毒性病理学领域领域工作1年以上。其主要职责为：

- a) 确定图像病变类别；
- b) 执行病变区域的初步勾画与标注。

4.3 审核人员

应具备中级及以上职称的病理专业人员，或在毒性病理学领域全职工作5年以上。其主要职责为：对初级标注人员完成的标注任务进行审核与精细修订。

4.4 仲裁人员

应具备副高级及以上职称的病理专业人员，或在毒性病理学领域工作10年以上的毒性病理学家。其主要职责为：

- a) 对存在分歧的疑难切片进行审核与仲裁；
- b) 最终确定病变类别以及病变区域的标注界线。

4.5 标注责任人

单一项目中可设置项目组长作为标注责任人。项目组长可由审核人员或仲裁人员担任，负责单一项目的全部数据标注管理工作。

5 数据采集与预处理

5.1 伦理批准

病理全切片图像来源于动物相关药效及毒性研究项目，试验开始前需得到相关动物使用与管理委员会的认可。

5.2 数据质量控制、查重要求

5.2.1 图像质量控制的具体要求

标注项目责任方应制定详细的图像质量控制标准，在数据正式进入标注流程前开展图像质量控制操作，重点核查并剔除以下不合规内容：

- a) 格式与分辨率异常：应确认图像为病理全切片图像（WSI）的数据格式（如SVS、NDPI、TIFF 或 OME-TIFF 等）；对于数据格式不符或空间分辨率未满足项目预设技术指标（如未达 0.25 $\mu\text{m}/\text{pixel}$ ）的图像，应予以剔除；
- b) 文件毁损与解析障碍：对于在数据传输或存储过程中发生数据损坏、文件结构破坏，或无法被常规数字病理软件正常读取与解析的图像文件，应予以剔除；
- c) 图像内容不合理：剔除切片数据缺失或标识信息错配的图像；单张 WSI 的扫描视场应完整覆盖物理玻片上的有效组织区域，对于存在扫描截断或边缘遗漏的图像，应予以剔除；
- d) 内容无关与严重伪影：应剔除与目标评估器官无关的非靶区组织图像；保留的 WSI 应完整呈现目标组织切面。同时，应重点核查并剔除存在严重制片伪影或扫描缺陷（如大面积组织折叠、组织撕裂、多发气泡、广泛失焦或对比度异常等）且足以影响病理学独立评价的图像。

5.2.2 数据查重与唯一性核验

为保证图像的唯一性，防止同源数据在后续算法测试中造成数据污染，应执行严格的图像查重流程：

- a) 查重范围：确保本批次待标注的图像内部不得出现重合文件，且不应与其他已公开的图库或历史研究项目中的图像发生重叠。所有保留的图像应当真实、唯一；
- b) 技术比对手段：在原始数据入库处理的过程中，应针对图像文件（如 NDPI 格式文件）计算其 MD5 散列值（或校验值），并基于该值进行全库交叉比对。

- c) 复项处理机制：一旦在比对核验中发现重复图像，应对此部分异常数据进行详细的标识与登记，予以物理或逻辑封存，并坚决从最终交付标注的合格数据池中排除。

5.2.3 质量控制记录与规程管理

图像的质量控制与查重活动应当建立对应的标准操作规程（SOP）（图像质量控制流程见附录A）。所有质量控制、查重与核验的执行过程必须形成详实的操作记录，确保文件剔除、封存的每一步均具备清晰的可追溯性。

5.2.4 设备和数据采集技术要求

所有组织样本宜采用数字病理扫描仪对所有组织样本进行扫描，常规推荐放大倍数为20×或40×。实际操作中暂不作强制定限，可根据不同病变类型的诊断需求以及数据存储条件进行灵活调整。并将数据导入计算机中。设备方面主要应考虑制造厂家、型号、同一台设备的不同成像参数配置，这些因素影响图像的对比度、分辨率、信噪比、细节丰富程度等基本参数，同时也会影响病理专业人员的标注或读片结论。

基于市场有国内外多种数字病理扫描仪种类和型号，为确保获得高质量的组织数字病理图像，建议考虑以下因素：

- a) 设备厂家：已获医疗器械注册证的国内外厂家；
- b) 扫描参数：20×或者40×物镜，采用20×或40×扫描模式；
- c) 扫描速度：无限定要求；
- d) 通量要求：无限定要求；
- e) 聚焦方式：手动/自动；
- f) 采集方式：明场扫描；
- g) 数据输出格式：SVS、NDPI、TIFF 或 OME-TIFF 等格式。

6 图像数据标注的任务实施及质量控制

6.1 标注流程

标注任务分解为两个，包括分类和分割（描边界，即描画病变区域边界）。按照初级标注、审核和仲裁两个模块配置标注人员和工作流程。标注任务的输入为组织病理学图像。在初级标注阶段，任务启动后，由两名初级标注人员独立执行图像识别与病变类别的定性判定，并对双方的分类结果开展交互复核。若在此环节中出现判定分歧，相关疑难切片将提交至仲裁专家进行专向审核与最终定类。完成类别确证后，标注人员继而执行病变区域的分割与边界勾画，并对勾画结果再次进行交互修订。

随后，数据流转入审核和仲裁阶段。审核人员首先针对初级复核中无分歧的样本，按30%的比例对其类别判定进行抽样审查。在抽查过程中，若发现类别判定存在差异，该样本须作为争议项转交仲裁专家重新审核分类；若判定一致，则直接进入最终核验环节。最后，由仲裁人员对所有样本的病灶边界勾画结果进行30%的抽样终审。待所有质量控制节点均核验通过后，该项数据标注任务正式闭环完成。（流程图见附录C）

6.2 任务实施的过程组织

6.2.1 任务生成及分配

标注责任方为项目负责人，由标注项目组长（责任人）根据项目内容分配任务给初级标注人员，设置操作权限，确定标注流程、决策机制与工作量，下发待标注的数据。初级标注人员对收集的待标注数据进行核查，确定项目所需的样本符合进行AI病理数据库构建的基本要求，如切片质量良好，染色均匀，组织细胞形态清晰可读等。审核人员和仲裁人员根据分工参与标注任务的审核和仲裁工作。

6.2.2 任务实施及质量控制

标注人员应根据标注规则执行标注任务。当标注责任人分配好任务后，还需对标注人员进行培训和考核，并形成相应记录。考核通过后，确定为项目实施的标注人员。

在标注过程中，标注责任人应对标注人员的标注质量进行监督，评估标注人员的能力，设置重复性指标和准确性指标（如在标注样本中重复 10% 样本）。当标注人员表现出质量显著下降时，标注责任人应对标注人员进行休整、培训和再评估。对重复性指标的评价采用埋题验证的方式，统计同一个标注人员在每次连续标注过程中对同一个数据的标注结果，计算重复标注一致或误差在允许范围内的样本在重复标注样本中的比例。具体操作：每完成 30 个病理切片的肿瘤区域标注后，随机抽选一张切片重新进行标注。计算重复标注的相似度。

对准确性指标的评价可对比标注人员与仲裁结论，计算仲裁人员认为正确的初级标注样本比例。标注分类可使用灵敏度、特异度、准确率进行评价；标注分割的准确性可采用 Dice 系数进行评价。

灵敏度：用于评价标注人员对正类样本的识别能力，计算方法见公式（1）

$$Sensitivity = \frac{TP}{TP + FN} \dots\dots\dots (1)$$

式中：TP 为真阳性样本数（待评价的标注区域为正、专家确定的真值区域为正）；FN 为假阴性样本数（待评价的标注区域为负、专家确定的真值区域为正）。

特异度：用于评价标注人员对负类样本的正确识别能力，计算方法见公式（2）

$$Specificity = \frac{TN}{TN + FP} \dots\dots\dots (2)$$

式中：TN 为真阴性样本数（待评价的标注区域为负、专家确定的真值区域为负）；FP 为假阳性样本数（待评价的标注区域为正、专家确定的真值区域为负）。

准确率：用于评价分类模型整体预测正确的比例，计算方法见公式（3）

$$Accuracy = \frac{TN + TP}{TP + TN + FN + FP} \dots\dots\dots (3)$$

式中：TP、TN、FP、FN 分别为真阳性、真阴性、假阳性、假阴性样本数。

Dice 系数：用于计算两个分割区域的重合度，计算方法见公式（4）

$$Dice = 2 \times \frac{|A \cap G|}{|A| + |G|} \dots\dots\dots (4)$$

式中：A 为待评价的标注区域，G 为专家确定的真值区域

7 数据的安全管理及存储

7.1 安全管理

在标注前，标注责任方应确保待标注数据信息可对外公开；应建立待标注数据的独立备份，确保该备份数据不被修改、删除。执行数据标注、计算和存储的设备在停用、退役时，将其中所有数据完全移出，然后将其中所有数据彻底删除，并无法恢复。标注责任方应保证标注过程的网络安全，如采用防火墙、边界防护、入侵防护等安全措施。

7.2 数据存储

数据采集设备为数字病理扫描仪，获得毒性病理数字图像，储存格式可为SVS、NDPI、TIFF 或 OME-TIFF等其他图像格式。标注前的原始图像数据以及带有标注信息的更新数据，均应严格执行双重备份机制。具体要求如下：

- a) 存储介质要求（双备份）：上述原始图像与标注数据均须保留两份物理拷贝。一份存储于独立硬盘中，另一份备份至数字病理扫描仪配套的服务器上，以实现备份存储；
- b) 备份执行策略：原始图像数据生成后应即时进行一次性全量备份；针对标注过程中持续更新的标注信息数据，应设定明确的周期，执行定期的增量备份（或配合周期性全量备份），确保数据动态更新的安全性。

AA

附 录 A
(规范性)
病理图像扫描质量判定条件

A.1 判定不合格的情形

A.1.1 图像信息完整性不合格

A.1.1.1 组织区域缺失：扫描区域未能完整覆盖玻片上的所有目标组织区域。

A.1.1.2 重要标记缺失：扫描生成的数字图像须保证玻片关键标记（如病理号、动物编号、实验编号等）的清晰可读；针对标签缺失或模糊不可辨认的异常样本，需进行人工介入核验。此外，凡采用非原始信息（如脱敏编码、AI 专用标识）进行扫描建档的样本，必须具备完善的映射溯源功能，确保非原始信息能够逆向追踪至原始标记信息。

A.1.2 图像清晰度/分辨率不合格

A.1.2.1 整体模糊：在整个扫描区域或大部分区域，图像缺乏锐利度，细胞边界、核膜、染色质结构等关键细节模糊不清，即使在最高放大倍率（40×）下也无法分辨。

A.1.2.2 局部失焦：图像中存在显著的区域性失焦，显著的区域性失焦是指图像特定区域内因聚焦不准确导致的清晰度严重下降，其核心判据在于造成了实质性的诊断信息损失。具体表现为：在诊断所需的高倍率（40×）下，该区域内关键的诊断性细节变得模糊不清或完全无法识别，包括细胞核的精细结构，如核膜的完整性、染色质的粗糙或细腻分布、核仁的数量大小及清晰度、核分裂象的准确计数与识别、单个细胞边界的清晰界定。

A.1.3 图像伪影不合格

A.1.3.1 污垢/灰尘伪影：扫描图像中存在非组织来源的、显著的、影响观察的污点、灰尘、纤维、气泡影像，覆盖关键诊断区域。

A.1.3.2 扫描条纹/条带：图像中出现明显的、贯穿扫描区域的明暗交替条纹或条带，严重影响图像均匀性和观察。

A.1.3.3 组织折叠/刀痕伪影：制片过程产生的组织折叠、刀痕等，在扫描图像中因光线折射或阴影形成大面积、深色、遮挡重要结构的伪影区域。

A.1.3.4 照明不均：图像整体或局部存在明显的明暗不均，影响对组织染色深浅的判断。

A.2 判定流程与说明

A.2.1 初筛：扫描人员完成扫描后，应首先依据本标准进行初步质量检查，数字切片组织扫描完整，无缺损，图像无色差，预览图象采集完整清晰，聚焦良好，图像清晰，无污，无杂质，颜色均一，发现明显不合格项（如大面积失焦、组织缺失、严重伪影、无法打开）应立即标记并安排重新扫描或检查制片。

A.2.2 正式质控：由经过培训的初级标注人员，使用专业的数字病理图像浏览器或标注工具，在符合校准要求的显示器上，按照本标准逐项检查。

A.2.3 记录与反馈：所有判定为不合格的图像，初级标注人员应记录不合格的原因，记录表格见表A.1。

A.2.4 分级处理：可分为以下几类

- a) 严重不合格：影响核心诊断信息获取（如关键区域失焦、组织缺失、严重色偏、无法打开）。
- b) 一般不合格：存在影响阅片体验或次要区域的问题（如非关键区域小污点、轻微照明不均、不影响诊断的微小瑕疵）。可根据具体情况决定是否重新扫描或备注说明后使用。
- c) 可接受瑕疵：存在轻微、明确标注的瑕疵（如玻片边缘小气泡、非诊断区小灰尘），不影响整体诊断。

A.3 结论

任何病理数字图像，若存在本判定标准中分级处理中严重不合格项，应判定为质量不合格；若存在本判定标准中分级处理中一般不合格项，且该问题影响图像用于其预定目的的可靠性和准确性时，应判定为质量不合格，若不影响图像用于其预定目的的可靠性和准确性时，应判定为质量合格；若存在本判定标准中分级处理中可接受瑕疵，应判定为质量合格。

表 A.1 病理图像质量控制表

样本编号	质量问题	结果	标注人员签字

附录 B

(资料性)

毒性病理学典型病变标注及工具操作示例

B.1 毒性病理学典型病变标注

B.1.1 标注任务分类

标注任务分类	数据模态: 毒性病理全切片图像 (WSI)
	执行主体: 半自动标注 (AI 辅助与人工修订结合)
	标注结果格式: GeoJSON、xml
	标注结果性质: 参考标准
	标注结果形式: 分类、分割 (区域勾画)

B.1.2 标注对象

标注对象是毒性病理图像中的典型病变组织与细胞，遵循国际通用的毒性病理学诊断标准，具体诊断和标注要点如下。

B.1.2.1 胃鳞状细胞癌

- a) 呈现外生性和 (或) 内生性生长方式;
- b) 单个肿瘤细胞或者小簇肿瘤细胞突破基底膜;
- c) 细胞分化缺失, 有间变, 可侵袭黏膜下层、肌层和浆膜;
- d) 分化良好型: 形态接近正常的鳞状上皮伴有不规则结构的乳头体, 中央常显示过度角化灶 (角化珠); 侵袭区域可见多角形和多形性细胞;
- e) 分化差型 (间变型): 实体片状分布或条索状排列的梭形细胞, 呈促纤维增生性特征;
- f) 细胞核深染、体积增大且含有明显的核仁, 较多核分裂象;
- g) 常见浅表溃疡和炎症、固有层或黏膜下层纤维增生, 可转移至腹腔、区域淋巴结或肺脏。

B.1.2.2 肝脏混合细胞浸润

- h) 单核细胞伴有中性粒细胞浸润。
- i) 如果在 CART 细胞治疗产品研究中, 可能存在增殖的 CAR-T 细胞。
- j) 活跃的肝细胞变性区和/或坏死区存在混合细胞浸润。
- k) 肝细胞再生存在混合细胞浸润。
- l) 小叶内随机分布。
- m) 可能有中性粒细胞和单核细胞浸润的共同诊断要点

B.1.2.3 肝脏淋巴瘤

- a) 由不具成熟的核染色质特征的淋巴细胞群组成;

- b) 中至大的淋巴母细胞，细胞质稀少到中度、嗜碱性并可能空泡化；
- c) 核呈圆形、卵圆形、不规则形或卷曲，染色质有细微斑点；
- d) 位于中央的核仁可从不明显到存在 1-3 个清晰的小核仁，核质比高；
- e) 核分裂象数量可变，通常较多；细胞不粘连，但形成均质的片状；
- f) 其他部位出现肿瘤性淋巴细胞聚集，且缺乏炎症刺激。

参考文献；彼得·曼，杨利峰，周向梅. 大鼠和小鼠病理变化术语及诊断标准的国际规范[M]. 中国农业出版社, 2019.

B.2 典型标注工具具体操作步骤

B.2.1 Qupath标注具体步骤

B.2.1.1 创建项目集（Create Project），添加图像，打开所需标注图像文件。

B.2.1.2 应用 QuPath 图像软件中的注释工具（Annotation），建立分类列表（class list），对扫描切片中感兴趣区域（ROI，Region of Interest）的进行初步注释为黄色（Region），针对胃鳞状细胞癌，红色（肿瘤）、蓝色（非肿瘤区域，包括腺胃黏膜以及幽门，non-tumor）；针对肝脏淋巴瘤，红色（肿瘤）、蓝色（空白、坏死或间质）；针对肝脏混合细胞浸润，绿色（混合细胞）、蓝色（空白、坏死或间质）。

B.2.1.3 对相应组织进行标注，选中后所有标注区域后，选择 Analyze 中的 Cell Detection，分析识别组织所有细胞。

B.2.1.4 初级标注人员使用注释工具对病变区域进行标注。

B.2.1.5 选择 Classify 中的 Object Classification，训练分类器。选择 Load Training，添加标注过的图像，点击 live update，查看训练效果，如果准确性较差，重复上一步，重新定义分类器。如果准确性较好，可对分类器进行命名，保存分类器。

B.2.1.6 根据分类结果标记，审核人员进行人工精细修订肿瘤注释。选择注释工具的选择模式（Selection Mode），选中想修改的类别，点击 Auto Set，对于标注错误的细胞进行手动修改。

B.2.1.7 保存所有注释。选择 File 中 Export object as GeoJSON，设定保存路径，保存文件。

B.2.2 HALO AI标注具体步骤

B.2.2.1 打开所需标注图像文件。

B.2.2.2 使用 HALO 图像软件中的注释工具，对扫描切片中的感兴趣区域（ROI，Region of Interest）进行初步注释为（Annotation），建立并保存初级注释区域（layer 1）。

B.2.2.3 在选定的 ROI 内，由病理专业人员使用注释工具，对不同病灶进行标注。针对胃鳞状细胞癌，红色（肿瘤）、黄色（空白）、蓝色（坏死或间质）。针对肝脏淋巴瘤，红色（肿瘤）、蓝色（others：空白、坏死或间质）。针对肝脏混合细胞浸润，绿色（混合细胞）、蓝色（空白、坏死或间质）。分别手动勾画若干具有代表性的典型微小区域（即提供参考示例）。算法将实时提取这些示例区域的图像特征，用于构建基础的图像识别与分割逻辑。

并在分类器操作（Classifier Actions）下的高级定义分类器（Advanced classifier options）中对注释进行标识，即勾选“Mask to Annotation”。

B.2.2.4 对病变区域和其他区域进行标注。

B.2.2.5 对初级注释区域（Annotation, layer 1）进行分类。分类完成后自动将病变组织进行分割并将组织边界生成相应的注释区域。运算结束后，采用 classifier 按钮下的 Real-time Tuning，对完成的分类进行初级评估，如果准确性较差，重复上一步，重新定义分类器。

B.2.2.6 保存分类器。选择分类器操作（Classifier Actions）下拉按钮，然后选择保存（Save）以显示保存分类器（Save Classifier）对话框，对分类器进行命名后保存。

B.2.2.7 根据分类结果标记，病理学专业人员进行人工精细修订。选择 Annotation，打开病变层面。对病变区域或其他区域进行标注，应用注释工具笔（Pen）进行微调，如：错误的注释直接右键删除；需要补充的区域直接用注释笔勾画；需要调整的区域用笔选定病变区域红线进行拖拉。保存所有的病变注释。

B.2.2.8 导出标注的数据。选择标注好的图像，右键选择 Manage Annotations。全选标注，选择 Action 中的 Export，选择 geoJSON 格式。设定保存路径，然后运行保存。

B.2.3 Visiopharm标注具体步骤

B.2.3.1 创建项目（Study）并在 Study 文件夹下创建 APP 子文件夹，包括 APP 01-组织识别，APP 02-病变区域识别，APP 03-结果输出（可根据需求增减 APP）。

B.2.3.2 将图像导入至 APP 指定的子文件夹中以构建训练集。训练集具体要求如下：

- a) 样本数量要求：应根据不同的病变类型进行均衡采样，每类病变建议配置 20~50 例有效图像样本；
- b) 模型性能目标：基于该训练集训练后的模型，在后续的性能评估与验证阶段，其分类准确率应达到 85%~90% 的指标要求。

B.2.3.3 APP 01-组织识别：

B.2.3.3.1 APP 构建：在 Image classes for training 添加分类类型，class 1 background（默认为绿色），class 2 tissue（默认为蓝色）；在 Input 栏将 Magnification 调整为 0.5 倍；在 Classification 栏选择 U-Net 算法。

B.2.3.3.2 标记：将整个图像标记为 class 1 绿色背景，将 class 1 绿色背景调整为不可见，再将组织区域标记为 class 2 蓝色，将 class 1 绿色背景调整为可见，即完成将图像标记为蓝色的组织区域和绿色的背景区域，然后点击 training 即可开始训练。

B.2.3.3.3 后处理程序：在 Post Processing 中设置参数，增加识别结果的准确性，如将面积小于某特定值的蓝色组织转换为绿色背景，将最终的蓝色组织输出为 ROI，使后续 APP 在此 ROI 上进行，以缩小运行范围和时间。

B.2.3.4 APP 02-病变区域识别：

B.2.3.4.1 APP 构建：在 Image classes for training 添加分类类型，class 1 background（默认

为绿色), class 2 lesion (默认为蓝色)(分类数量和颜色可根据需求进行选择,如针对胃鳞状细胞癌,红色(肿瘤)、黄色(空白)、蓝色(others:坏死或间质),针对肝脏淋巴瘤,红色(肿瘤)、蓝色(others:空白、坏死或间质),针对肝脏混合细胞浸润,绿色(混合细胞)、蓝色(others:空白、坏死或间质));在 Input 栏将调整 Magnification 的倍数(根据病变的大小进行调整,肝脏淋巴瘤建议调整为 20 倍);在 Classification 栏选择 Deep Labv3+ 算法。

B.2.3.4.2 标记(以肝脏淋巴瘤为例):将整个图像标记为 class 1 蓝色背景,将 class 1 蓝色背景调整为不可见,再将肿瘤区域标记为 class 2 红色,将 class 1 蓝色背景调整为可见,即完成将图像标记为红色的肿瘤区域和蓝色的背景区域(空白、坏死或间质),然后点击 training 即可开始训练。

B.2.3.4.3 后处理程序:在 Post Processing 中设置参数,增加识别结果的准确性,如将面积小于某特定值的红色肿瘤转换为蓝色背景。

B.2.3.4.4 扩大训练集:对前期标注后的图像进行训练,在新的图像上运行 APP 02,病理专业人员对运行结果进行人工精细修订,提高肿瘤标注的准确性,将此修订后的图像作为训练集扩大训练。

B.2.3.5 APP 03-结果输出:

B.2.3.5.1 APP 构建:在 APP 02 的基础上,在 Input 栏将 Magnification 调整为 1 倍(调整为 1 倍,可提高运行速度);在 Classification 栏选择 No Classification;在 Output Variables 里选择 class 2 红色肿瘤的面积或者其他参数作为输出对象。

B.2.3.5.2 运行 APP 03 后即可输出结果;点击 Export 即可导出 mld 格式的图像、tsv 格式的数据文档以及 APP。



附录 C
(规范性)
业务架构示例

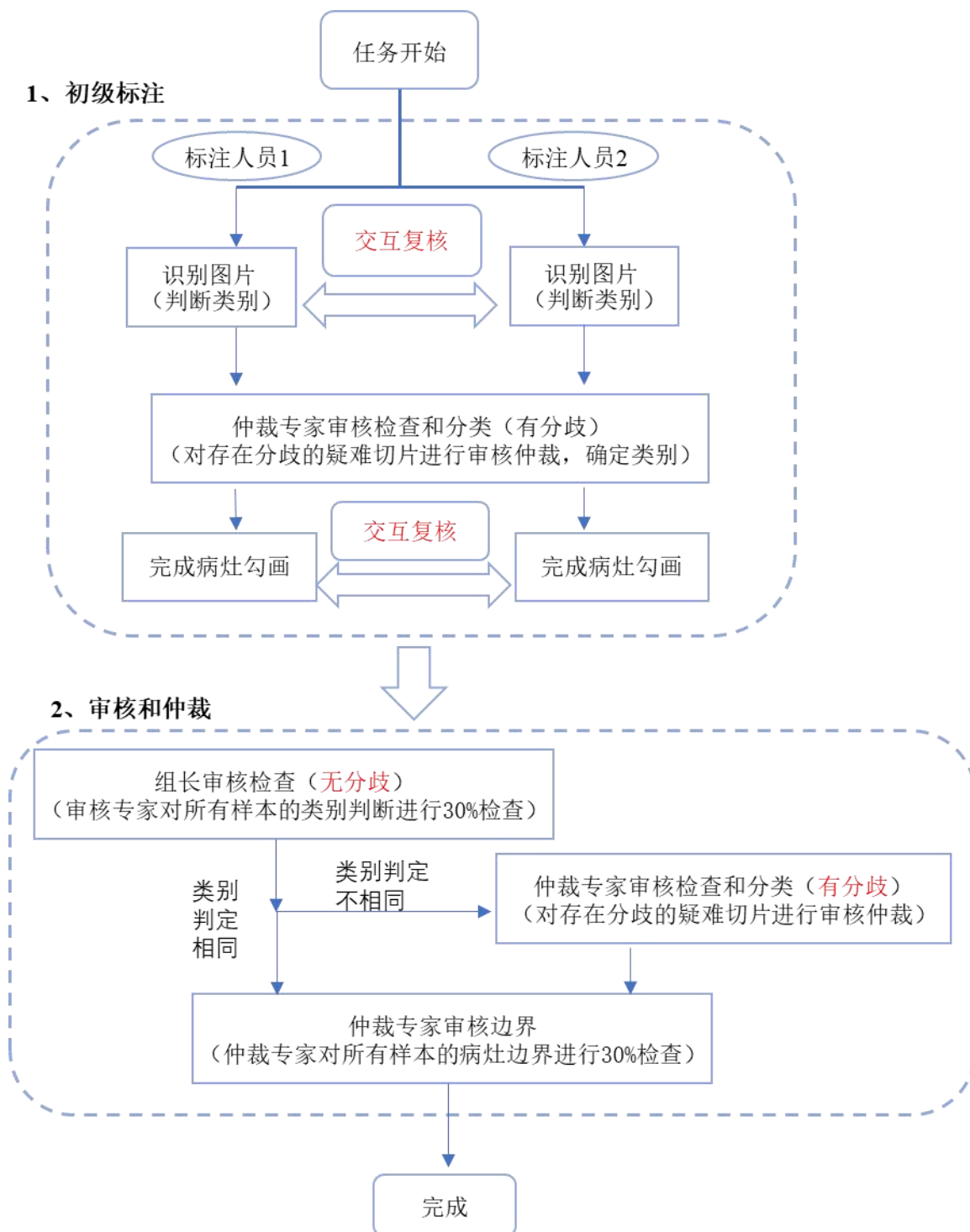


图 C.1 毒性病理学图像数据标注的业务构架图

D

参 考 文 献

- [1]GB/T 42755-2023 人工智能 面向机器学习的数据标注规程
- [2]YY/T 1833.1—2022 人工智能医疗器械质量要求和评价第1 部分：术语
- [3]YY/T 1833.3—2022 人工智能医疗器械 质量要求和评价 第3部分：数据标注通用要求
- [4]DB21/T 4419-2026 组织病理学全切片图像标注技术规程
- [5]《人工智能辅助胃部组织病理学诊断的数据采集和标注专家共识》编写组. 人工智能辅助胃部组织病理学诊断的数据采集和标注专家共识 [J]. 中华病理学杂志, 2024, 53 (09): 893-897.
-