

面向智能医疗的人工智能辅助诊断系统优化研究

郝志丹

中软炫达（北京）科技有限公司，北京市东城区，100011；

摘要：围绕人工智能辅助诊断系统的可用性与可托付性构建优化框架，从数据治理、模型能力、计算架构、临床集成到合规治理形成贯通链条，提出以统一表征与分层学习为主线，用数据谱系与质量量化托底，输入可靠，以自适应结构与多任务目标提升泛化与稳健，引入不确定度管理与解释机制，增强临床可审计性，在边缘云协同与时延约束下进行推理编排，通过标准化接口与流程对接临床工作流，同时以风险分级与伦理规则约束模型与人员行为，给出从离线评估到在线监测的闭环路径，以期在不同机构与设备条件下获得稳定可迁移的诊断表现。

关键词：人工智能辅助诊断；智能医疗；数据治理；不确定度；边缘云协同

DOI：10.69979/3060-8767.25.06.058

引言

人工智能在医学影像、生理信号与临床文本等场景的应用不断扩展，诊断精度与时效随之提升，同时也暴露数据异构、标注稀缺与部署成本高等问题。系统若仅关注单一模型指标，往往难以满足临床对稳定、透明与可追溯的要求。为了在真实环境中长期运行，需要把数据流程、模型结构、计算资源、安全规范与临床工作方式纳入同一优化目标，并以可验证指标与审计记录支撑改进。本文提出一套可迁移的理论路径，从数据底座、模型优化、计算与部署到集成与治理逐层展开，目标是在有限算力与严格合规的条件下获得可靠且高效的辅助诊断能力。

1 数据治理与标准化底座

1.1 多模态医疗数据统一表征方法

诊断系统对输入一致性的依赖极强，不同医院与设备产生的影像、信号与文本往往存在分布差异。为降低域间漂移，可在采集环节记录设备指纹与采样协议，在数据层引入统一数据字典与编码规则，将影像组织为多分辨率体素与解剖标签的组合，将心电图与生命体征表达为多通道时序与频带要素的拼合，将病历记录转写为分层语义单元与标准术语映射，并建立跨模态的对齐坐标，以便在下游共享同一嵌入空间^[1]。表征学习阶段引入域不变成分与域特定残差，前者支撑跨场景迁移，后者记录本地语境差异，通过一致性约束让影像、信号与文本在共享空间中保持几何邻近，同时保留模态特性以支持细粒度判断。数据清洗流程覆盖去伪影、配准与时间对齐，缺失片段在规则边界内插补，并附带来源健康度，使模型能够依据可信信息动态分配注意。结构化后的样

本绑定数据谱系，任何加工步骤都有版本与校验记录，便于回溯差错并在机构之间复用规则。统一表征不仅是格式的统一，更是语义层级的统一，需要在词项命名、单位换算与坐标系映射上给出严格规则，通过公共本体把检查项目、药物处置与解剖结构联系起来，让不同科室与不同机构在同一语言下协同训练。样本在进入模型前完成规范化切片，切片边界由临床事件与器官解剖点共同决定，保证片段内部的状态相对一致。对于体素级标签稀缺的场景，可采用弱监督与伪标签引导区域学习，再用少量精标样本进行校正。统一坐标还需要时间轴上的锚点，把就诊事件、检查时间与用药时序整合到同一时间线，让模型能把动态变化与诊断结论联系起来。在接口层面提供可执行的示例与校验器，任何接入方都能快速验证字段与单位的一致，降低集成成本并减少差错。

1.2 标注分歧管理与质量评估

医学标注存在经验差与任务边界模糊等难题，单一标签容易掩盖分歧，更稳健的做法是采用分布化表达，记录多名专家在强度与部位上的判断差异，训练时以分布回归替代单点分类，推理阶段输出区间与可信度，让临床能够结合情境做出最终判定。样本不均衡可通过重加权与难例采样缓解，少见病种在物理约束下进行增广，例如基于解剖形变与成像物理的强一致变换，在不破坏关键结构的前提下扩展样本^[2]。质量评估不仅看噪声与遮挡，还应度量配准误差与时间漂移，以来源健康度作为训练权重的调节因子，对高风险样本给予更严格的校验与较低的学习强度。标注工作需要形成带有对照样例与复核清单的流程，高分歧片段进入二次讨论并保留差异信息，模型在此类片段上接受一致性约束，减少对偶然特征的依赖。长周期运维阶段以数据质检报告反向驱

动采集改造与术语规范的更新,使输入端逐步向稳定与可解释方向收敛。分歧管理可以借助带实例的指南,在复杂边界给出正反样例与评分细则,减少因认知差异造成的漂移。质量评估加入跨天与跨周的稳定性统计,识别季节性与排班对数据结构的影响,在采集端设置自动质检规则,如视野遮挡比例、成像参数偏离与同步丢失,触发告警并给出可操作建议。对于文本记录引入术语映射与拼写纠错,减少同义词与简写导致的语义分裂,在训练样本页面提供分层可视,让标注者查看来自影像与信号的证据,促进跨模态的一致理解,在质控环节引入双盲抽检,以比例化规则持续评估一致水平。

2 模型优化与可信推理

2.1 诊断模型的结构改进与自适应学习

模型设计需要在精度、可解释与可部署之间取得平衡。影像通道可采用多尺度卷积与残差连接,捕获器官层级结构,信号通道以时频变换与门控单元捕捉节律变化与瞬态异常,文本通道以分层语义嵌入结合指代解析维持上下文一致。跨模态的统一变量作为锚点,把各通道的瓶颈向量投影到共享空间,通过一致性损失稳定对齐,多任务目标把病灶检出、分型与风险评分联合优化,在参数共享下通过正交约束降低相互干扰^[3]。小样本场景依靠迁移学习与轻量适配层实现快速收敛,以冻结主干并仅更新低维侧枝的方式减少过拟合风险。训练过程中引入难例挖掘与边际正则,提升类间间隔并抑制对噪声标注的过拟合,对抗扰动与形变增广用于提升鲁棒性,在离线评估中绘制退化曲线与临界点,为后续部署护栏提供依据。蒸馏与低秩分解把大模型的知识转移到紧凑结构中,保证边缘设备上的推理成本可控,同时维持主要指标在可接受区间。结构改进还体现在层级组织的组织方式,局部注意捕捉细小征象,全局注意承载器官级与病例级关联,二者以门控方式结合。在长序列中引入记忆压缩与滑动缓存,减少重复计算并降低延迟。自适应学习通过领域对齐与参数共享实现,以小样本校准快速拟合本地分布,再用一致性约束限制漂移幅度,对类别极少的病种设置阈值提升策略,当置信不足时自动升级为建议级输出,并提示补充检查。在评估阶段引入任务分解与错因归类,把定位错误、分型错误与评分偏移拆开计算,提供针对性的改进方向,对不同模态的权重演化进行记录,分析失败样本中的证据冲突,把高风险模式加入训练禁忌列表,在混合精度训练下结合损失缩放提升吞吐,同时保持数值稳定。对于长程依赖可加入稀疏注意与局部窗口,在不牺牲关键信号的前提下降低

计算量,在跨任务迁移时设置冻结名单与可变名单,关键层保持稳定,适配层灵活调整,减少对既有能力的扰动。

2.2 不确定度管理与可解释决策

临床更关切可靠度与可追溯信息,模型需要对自己的把握程度给出清晰表达,可通过温度校准与采样估计构建不确定度量,在融合层按可信度分配权重,当某一通道质量下降时自动让渡决策权,并触发额外复核或追加模态。解释机制不止展示显著性图,还需给出证据路径与时间演化轨迹,说明判定依赖的关键片段与跨模态呼应关系,因果建模用于剥离虚假相关,以解剖结构与生理先验构建约束图,让推理遵循可验证的医学逻辑^[4]。在线阶段设置异常监测,当输入脱离训练分布时给出警示并降级输出,把不确定度与来源健康度同时呈现给用户,便于制定后续检查与会诊安排。在多机构协作中以统一的解释模板与指标字典实现可比与可审计,减少因话语系统差异引起的沟通障碍,解释不仅服务医生,也服务患者,沟通系统应生成便于理解的提示语,配合关键切片与时序曲线,让非专业人员也能把握要点。在团队协作中提供证据共享视图,支持多学科会诊时的同步标注与讨论。不确定度在路径规划中扮演重要角色,当风险高又证据不足时系统建议补充检查或延后决策,把潜在风险转化为明确行动,对于边界病例提供多方案备选,并标注各自的依据与注意事项。解释模块采用模板化与可扩展规则共存的设计,新任务上线时快速接入现有范式,同时保留灵活修改空间,在教学医院的实践环境中解释结果还可用于带教与知识沉淀,把典型病例与失败样本整理入库,作为后续训练与评审的公共材料。模型在不断更新时,解释模板与阈值策略也随之调整,保证不同版本之间的可比较性,对外部质询提供清晰证据,让系统具备经得起复盘的能力。当出现跨模态矛盾时,系统展示候选路径与冲突点,并给出继续采集或人工核验的建议,把不一致变成可操作的改进线索,解释与决策应形成结构化报告,关键依据排序呈现,附带版本与阈值信息,便于后续追踪与横向比较。

3 计算架构与临床集成

3.1 边缘云协同与时延约束下的推理编排

临床现场对响应时间与稳定性的要求严格,系统需要在边缘与云之间动态分工,边缘侧完成预处理与低成本筛查,云端承担重计算与模型更新,二者通过任务图与缓存机制协同,在网络不稳时仍能保持核心能力。推理流程采用并行与流水结合的方式,影像与信号短窗并

行处理,文本在到达后异步对齐融合,在证据充足时提前输出,在证据不足时拉取高成本通道,以此在时延与准确之间实现可控取舍。轻量化依赖剪枝、量化与稀疏注意,部署采用灰度发布与影子比对,当指标异常时自动回滚到稳定版本,运行监测记录时延分布、内存曲线与失败率,超阈后触发重建与告警,长期运行阶段定期整理日志与配置档案,形成可重现的运维证据。任务编排把推理流程划分为若干可复用单元,包括预处理、特征抽取、融合、决策与记录存档,单元之间通过明确定义的输入输出连接,便于测量与替换。事件驱动的执行框架可根据输入质量动态决定路径,当影像噪声高时优先拉取信号与文本,降低单一渠道失常带来的波动。边缘节点维持本地缓存与模型快照,在网络恢复后自动对齐版本,保证跨时段结果的一致,负载高峰采取早停与分批返回策略,先给出粗判,再追加细粒度结果,在急救场景中以毫秒级筛查为目标,把高风险病例优先送入人工通道,实现整体效率与生命安全的兼顾。资源编排需要与医院的排班与设备台账对齐,在设备维护与停机窗口内完成模型更新与压力测试,避免对临床产生扰动,对跨园区部署引入就近计算与路由策略,减少跨城网络的波动影响,在安全区域边界设置隔离环,阻断异常流量,推理服务与存储服务分离部署,防止单点故障放大。数据与模型的版本绑定到检查申请与报告,实现端到端的可追踪,便于在复盘时还原当时的运行条件,资源估算基于典型日曲线与峰值场景联合校核,以容量护栏限制突发扩张,并为应急提供弹性池与旁路通道。

3.2 workflow融入与安全合规治理

人工智能要真正落地必须与临床工作方式顺畅衔接,系统应通过标准化接口对接信息系统、影像归档与病历平台,结果以结构化要素与可读文本的联合形态呈现,避免额外负担。在流程上明确告知临床数据来源、模型版本与阈值设置,并提供人工复核与反馈通道,反馈进入模型治理环节,以小批增量与门控策略更新参数。合规治理聚焦最小权限与可追溯原则,数据在采集、传输与存储环节全程加密,访问按岗位授权分级记录,跨机构协作采用联邦训练与隐私计算,保护原始信息。公平性评估关注不同群体与场景的误差差异,通过重加权与去偏方法进行校正,风险管理以分级目录与处置脚本为抓手,对高风险输出强制人工确认,对低风险结果采用批量通过与抽检机制,通过定期评审把制度与技术进

步同步到位。系统融入需要尊重科室节奏与角色分工,门诊、住院与检查科的需求并不相同,输出形式与交互方式应当因地制宜,对结果呈现设置可切换的精简视图与详细视图,精简视图服务快速决策,详细视图服务教学与复盘。合规治理需要形成清晰的证据链,任何版本更新与阈值调整都应留痕,并在固定周期接受第三方评测,数据出域与模型出域均需审批,日志系统记录访问与导出行为,异常行为触发审计流程。公平性与伦理约束纳入日常指标,按人群、地域与设备类型分层统计,持续识别性能差异并推动纠正,患者知情与选择同样重要,系统应提供易懂的说明与退选路径,在临床研究与常规诊疗之间明确界线,避免目的混用,对涉及未成年人与高敏人群的应用额外设置保护线,并通过伦理委员会的持续监督与质量管理体系协同运行,把安全指标与运维指标纳入同一看板,以统一视角联动技术流程与管理流程,通过持续培训与演练提升团队的协同效率,把制度转化为日常习惯,让技术与流程在高压环境下仍能稳定执行。

4 结语

辅助诊断系统的优化不是单点技术的堆叠,而是跨层协同的系统工程,以数据治理与标准化为底座,以结构改进与不确定度管理为核心,再以边缘云协同与流程集成固化运行机制,在稳定与透明的边界内持续迭代。当输入可信、决策可审计与部署可维护三项要求同时满足,人工智能才能成为可托付的基础能力。面向未来应把评价指标、治理规则与资源预算绑定到同一套版本体系,用可复现的流程驱动跨机构协作与长期改进。

参考文献

- [1] 李佳明,张曦,杨丽,等. 人工智能辅助诊断应用中的伦理探讨与哲学思辨[J]. 中国医学伦理学,2024,37(09):1037-1045.
- [2] 彭杰. 基于信息化平台的人工智能辅助诊断技术的研究[J]. 电脑知识与技术,2025,21(13):35-37. DOI:10.14004/j.cnki.ckt.2025.0687.
- [3] 刘超. 基于人工智能的疾病辅助诊断方法的可专利性[J]. 专利代理,2023,(04):62-70+82.
- [4] 朱宇昂. 基于深度学习的根尖片口腔X光影像人工智能辅助诊断研究[D]. 北京交通大学,2023. DOI:10.26944/d.cnki.gbfju.2023.001762.