

多源异构医疗大数据融合与分析技术

崔立真^{1,2}, 孙晓芳^{1,2*}, 刘宁², 徐庸辉², 何伟^{1,2}

(1.山东大学软件学院, 山东 济南 250101; 2.山东大学-南洋理工大学人工智能国际联合研究院, 山东 济南 250101)



崔立真, 山东大学教授, 博士生导师。现任山东大学软件学院院长, 山东大学-南洋理工大学人工智能国际联合研究院(C-FAIR)联合院长。主要研究方向为软件与数据工程。在国际顶级期刊和会议上发表高水平学术论文100余篇, 多篇入选ESI高被引论文, 全球前2%顶尖科学家。以第一完成人获国家教学成果二等奖、山东省科技进步一等奖等奖励。

摘要:医疗健康数据作为现代医学研究与实践的核心要素, 具有多源异构、碎片化和利用率低等特征, 难以有效挖掘蕴含的关联关系与知识价值。如何突破多源异构数据的融合瓶颈, 成为实现人体健康管理从被动治疗向主动干预转型的关键挑战。聚焦医疗健康数据的核心价值与融合难题, 系统梳理研究进展与技术突破方向。综合分析多模态多源异构数据融合、可解释性知识发现、跨模态关联关系挖掘, 创新提出多源异构医疗大数据前沿技术体系, 助力医疗健康数据体系向多模态化转型、知识图谱化升级、可解释化革新三重进化, 充分释放医疗健康数据作为国家战略性资源的倍增效应。

关键词: 医疗大数据; 多源异构; 多模态融合; 数据驱动与知识引导; 可解释模型

中图分类号: TP391 **文献标志码:** A

引用格式: 崔立真, 孙晓芳, 刘宁, 等. 多源异构医疗大数据融合与分析技术[J]. 山东大学学报(工学版), 2026, 56(1): 1-13.

CUI Lizhen, SUN Xiaofang, LIU Ning, et al. Multi-source heterogeneous medical big data fusion and analysis technology[J]. Journal of Shandong University (Engineering Science), 2026, 56(1): 1-13.

Multi-source heterogeneous medical big data fusion and analysis technology

CUI Lizhen^{1,2}, SUN Xiaofang^{1,2*}, LIU Ning², XU Yonghui², HE Wei^{1,2}

(1. School of Software, Shandong University, Jinan 250101, Shandong, China; 2. Joint SDU-NTU Centre for Artificial Intelligence Research, Shandong University, Jinan 250101, Shandong, China)

Abstract: Healthcare data, as a core element of modern medical research and practice, was characterized by multi-source heterogeneity, fragmentation, and low utilization rate, making it difficult to effectively uncover the implicit correlations and knowledge value. Overcoming the integration challenges of multi-source heterogeneous data became a key obstacle in shifting health management from passive treatment to proactive intervention. Focusing on the core value and integration difficulties of healthcare data, the research progress and technological breakthroughs were systematically reviewed. Through comprehensive analysis of multimodal and multi-source heterogeneous data fusion, interpretable knowledge discovery, and cross-modal correlation mining, an advanced technological framework for multi-source heterogeneous medical big data was innovatively proposed. This framework supported the triple evolution of healthcare data systems toward multi-modal transformation, knowledge graph upgrading, and interpretability innovation, thereby fully unleashing the multiplier effect of healthcare data as a national strategic resource.

Keywords: medical big data; multi-source heterogeneity; multimodal fusion; data driven and knowledge guided; interpretable model

收稿日期: 2025-09-09

基金项目: 国家自然科学基金重大研究计划资助项目(91846205)

第一作者简介: 崔立真(1976—), 男, 甘肃兰州人, 教授, 博士生导师, 博士, 主要研究方向为软件与数据工程。E-mail: clz@sdu.edu.cn

* 通信作者简介: 孙晓芳(1996—), 女, 山东潍坊人, 博士研究生, 主要研究方向为医疗大数据挖掘与分析。

E-mail: xiaofangsun@mail.sdu.edu.cn

0 引言

在数字经济与健康中国战略交汇的新发展阶段,医疗大数据作为数据要素正从基础资源向新质生产力核心载体跃迁。中共中央、国务院在《关于构建数据基础制度更好发挥数据要素作用的意见》中将数据纳入生产要素范畴^[1]。国家数据局联合16个部门印发《“数据要素×”三年行动计划(2024—2026年)》^[2],针对医疗健康领域,布局“数据要素×医疗健康”行动,提出“加强医疗数据融合创新,支持公立医疗机构在合法合规前提下向金融、养老等经营主体共享数据,支撑商业保险产品、疗养休养等服务产品精准设计,拓展智慧医疗、智能健康管理等数据应用新模式新业态”等具体行动计划。这一系列国家战略布局不仅确立了医疗大数据的要素属性,更凸显出推进医疗数据融合创新的现实紧迫性。

随着医疗大数据和数据要素资源流通体系的快速发展,医疗领域的数据类型呈现出爆炸性增长^[3]。医疗大数据普遍面临多源异构(涵盖患者基本信息、诊断数据、影像报告等多模态类型)^[4]、碎片化存储(分散于医院信息系统(hospital information system, HIS)、医学影像存档与通信系统(picture archiving and communication systems, PACS)、实验室信息系统(laboratory information system, LIS)等独立系统)^[5]及质量参差(术语标准缺失、录入格式混乱等)^[6]等核心数据挑战,导致数据利用率低^[7],形成信息孤岛与分析断层并存的困境^[8],数据价值难以释放,海量数据因分散存储形成知识碎片孤岛,跨部门协作效率低,且低质量信息进一步降低健康管理的有效性。本研究将医疗大数据的基本特点和挖掘痛点总结为数据种类多、数据碎片化、数据利用率低及数据分散广。这些痛点不仅加剧诊疗流程中的信息断层(如影像特征与病历文本难以时空对齐),更使多模态数据的深度关联分析面临技术瓶颈^[9-10]。因此,打通多源异构医疗大数据的语义壁垒,实现跨系统多模态融合创新,已成为提升诊疗精准度与医疗资源协同效率的必然需求。与此同时,临床场景对高效诊断工具的迫切需求持续升级^[11-12]。一方面,医生面临跨系统调阅多源数据(如影像、检验结果、病历等)联合判读耗时的难题,基层医疗机构更急需自动化工具弥补专业资源缺口;另一方面,远程医疗、精准诊疗等新兴场景要求实现多模态数据的时空对齐与语义

关联,例如对影像特征与电子病历文本中的症状描述联合推理。在此背景下,如何突破多源异构数据的融合瓶颈,是构建智能化医疗生态的核心命题,成为实现全人群健康管理从被动治疗向主动干预转型的关键挑战。

1 多源异构医疗大数据融合分析

医疗大数据作为现代医学研究的核心战略资源,其多源异构性、时空碎片化特征与潜在的医学知识价值正深度重构医工交叉领域的研究范式。随着大数据与人工智能(artificial intelligence, AI)技术的革新,医疗大数据已突破传统单一模态的局限,形成跨尺度、高维度、多模态融合的复杂生态系统。在此背景下,医工交叉研究通过AI技术驱动的方法论革新正在突破传统经验医学的边界,多源异构医疗大数据的重要性愈发凸显,吸引研究人员的广泛关注与深入探索。

1.1 多源异构医疗大数据融合

多模态融合通过整合影像组学^[13]、基因组学^[14]、电子病历^[15]及可穿戴设备^[16]等多源异构医疗大数据,突破单一模态的信息局限,实现跨尺度特征互补与语义协同^[17]。多模态融合的核心作用体现在以下两个方面:通过跨模态特征对齐,弥补单模态数据对复杂疾病表征的片面性;异构数据的动态耦合可缓解噪声干扰与数据缺失问题,增强模型的鲁棒性。本研究将聚焦医疗大数据多模态融合的关键技术,对当前主流研究展开详细分析,按照融合层在模型的不同位置划分为数据级、特征级、决策级3个融合层级,探讨不同层级融合方式在特征提取、医学影像与临床数据协同分析中的优势及局限,揭示多模态融合技术在提升医疗数据理解能力、优化临床决策支持系统和推动精准医疗发展的重要作用。

1.1.1 数据级融合

数据级融合是数据预处理阶段、预处理之后直接融合原始数据,属于早期策略。文献[18]针对目前大多数医学图像分割技术仅限于空间领域(例如仅利用磁共振成像(magnetic resonance imaging, MRI)、计算机断层扫描(computed tomography, CT)、正电子发射断层显像(positron emission tomography, PET)等扫描图像提取空间域特征)的问题,提出一种创新联合空间-光谱信息融合方法,将现有的单一模态数据转化为一个新的域,从替代空间中提取特征,在跨模态特征空间同步信息前引入一个对

比损失函数,实现多模态特征对齐。多模态的生理信号为多模态医疗大数据的分支,不同模态信号的采样率和幅值尺度等存在差异,为数据级融合的实现带来困难。为此,文献[19]提出一种基于小波变换的脑电(electroencephalography, EEG)和功能性近红外光谱(functional near-infrared spectroscopy, fNIRS)的多模态数据融合方法,对两种信号进行小波分解,将两组小波系数通过基于 Fisher 值的融合规则生成新的小波系数,通过小波重构得到融合信号,提取融合信号的共空间模式特征,实现多模态生理信号的特征级融合,获取不同生理信号的多模态互补特征。文献[20]将文本和医学成像数据融合成电子健康记录(electronic health record, EHR),提高诊断精度及治疗效果,加速医学研究,利用最先进的算法和技术,以互补的方式整合文本和成像数据,生成患者健康状况的全面情况。

1.1.2 特征级融合

特征级融合是多源数据融合的中层策略。在特征提取阶段,整合不同模态或来源的数据,结合各模态的高层次抽象特征增强模型的表达能力。特征级融合的核心在于挖掘跨模态特征间的语义关联,降低原始数据冗余。文献[21]设计一种混合偏移轴向自注意力机制的脑胶质瘤分割多层级轴向注意力网络(multi-level axial-attention net, MLANet)算法,提高脑胶质瘤 MRI 图像分割精度及质量,采用混合偏移轴向自注意力机制和混和损失函数,提取更精确的全局相对位置关系,提升网络对细节结构特征的敏感程度,实现精确分割胶质瘤模糊边界。针对图像和文本数据,现有多模态预训练模型大多在大规模数据集上使用外部目标检测器提取图像区域特征,与相应文本进行联合预训练。

表1 多源异构医疗大数据融合方法对比

Table 1 Comparison of multi-source heterogeneous medical big data fusion methods

融合层级	典型方法	异构数据处理策略	策略优点	策略缺点	适用场景
数据级融合	空间-光谱融合 ^[18] 、小波变换融合 ^[19] 、文本-影像整合 ^[20]	直接融合原始异构数据,通过格式统一、特征对齐实现融合	最大程度保留原始数据细节	对异构数据格式差异敏感,计算开销大	模态差异较小的场景,需保留原始数据细节的任务
特征级融合	多层级轴向注意力网络 ^[21] 、弱监督多模态预训练 ^[22] 、多模态自适应分配融合权重 ^[23]	提取各模态高层抽象特征,通过注意力机制、蒸馏学习等挖掘跨模态语义关联	降低原始数据冗余,适配多模态异构性	依赖高质量特征提取模型,小样本场景易出现特征偏移	多模态互补性强的场景
决策级融合	EHR 多模态综合融合框架 ^[24]	无需处理原始异构数据,仅整合各单模态模型的决策结果	计算效率高,隐私保护性强	依赖单模态精度,易丢失跨模态细节关联	隐私敏感、多源决策互补场景

综上,如何选择多模态融合方法的本质是信息完整性、计算效率、隐私安全性的三角权衡。数据

然而,传统多模态学习方法在训练数据不足、无边界框标注、存在噪声等场景下并不适用。针对上述问题,文献[22]提出一种融合弱监督和动量蒸馏的多模态预训练模型,使用弱监督目标定位获得目标对象的边界框,提取区域图像特征,引入动量蒸馏模块,创建一个与学生模型结构相同的教师模型,将区域图像和文本对输入教师模型生成伪标注,作为学生模型的额外监督,用于执行预训练任务。文献[23]提出一种结构先验引导的多模态信息融合分割算法,设计的多模态图像编码模块(multi-modality encoding module, MMEM)同时对不同模态的加权图像进行语义特征提取,跨模态体素融合模块(cross-modality voxel fusion module, CMVF)在融合过程中为各模态图像特征自适应分配融合权重,实现对椎体及椎间盘的精准图像分割。

1.1.3 决策级融合

决策级融合是在模型预测结果的最终阶段整合多个独立模型或模态的输出,属于后融合策略。EHR 数据拥有丰富的异构数据,包含多种来源的患者记录、医学测试、医学成像、诊断、药物、程序、临床记录等,共同提供患者健康状况的整体视图并相互补充。结合本质上不同的多模态数据在深度学习中评估 EHR 具有挑战性。为评估多模态数据的期望,文献[24]引入一个多模态融合框架,旨在整合 EHR 中的时间变量、医学影像和临床记录,增强临床风险预测性能,分别训练不同模态的分类器,将单模态预测结合,形成全局多模态预测,将早期、联合和晚期融合策略用于多模态数据融合。

综合分析上述 3 个层级的定义和应用,总结不同融合层级的策略优缺点和适用场景,如表 1 所示。

级融合以高计算成本换取原始信息保留,仅适配同机构内模态差异小的场景;特征级融合作为当前主

流,通过抽象特征平衡适配性与效率,但需解决小样本下的特征偏移问题;决策级融合以牺牲细节关联为代价实现隐私保护,更适用于多中心协作场景。现有多模态融合研究仍存在跨层级融合协同不足的共性问题,多模态数据融合的跨模态标准化、异质数据因果推理等挑战仍需突破。多模态融合的难点本质在于如何从多源异构医学数据中提炼跨尺度、跨模态的医学语义一致性,需结合领域知识驱动与算法创新。

1.2 医疗大数据的可解释性知识发现

数据驱动的深度学习技术在疾病预测、分类及检测等医疗任务中展现出显著优势,但固有的黑盒特性导致预测结果可解释性不足,制约临床决策中的应用^[25]。当前医学知识表示学习方法主要呈现两大研究范式^[26]:基于医学属性的知识表示学习,聚焦疾病特征与病理机制的表征建模^[27];基于医学知识图谱的知识表示学习,着力于构建多源医学知识的语义关联网^[28]。这两种方法均致力于建立可计算的知识表征体系,为智慧医疗提供兼具专业性与可解释性的决策支持。

1.2.1 基于医学属性的知识表示学习方法

医学属性知识是由症状、体征、实验室指标、病理特征等结构化医学信息构成的专业知识体系^[29]。结合深度学习技术,可将医学属性知识编码为可计算的特征向量,增强模型对疾病关键特征的辨识能力与临床可解释性。尽管处理时间序列 EHR 数据的深度学习技术已取得重大进步,但目前的方法忽视了领域知识与患者数据的整合。文献[30]提出一种基于时间序列 EHR 的心力衰竭预测方法,采用领域知识引导的卷积神经网络(convolutional neural network, CNN)和长短期记忆网络(long short-term memory, LSTM),针对 EHR 数据中的不规则时间间隔问题,LSTM 处理心力衰竭患者特征监测结果,引入位置编码,从医学知识图中提取心力衰竭特征与疾病的关联关系并结合实验室测试结果,CNN 模型提取疾病与特征之间关系的高维特征,整合构建级联预测模型。一些研究试图从文本中提取实体作为简要特征,为临床文本分类提供特定领域的知识,但是只注入单一模态模型是不够的。文献[31]提出实体增强的基于 Transformer 双向编码器表示模型(entity-enhanced bidirectional encoder representations from Transformer, E-BERT),利用 BERT 的结构属性进行医学专业预测,包含实体嵌入层和实体感知关注,以注入特定领域的知识,专注于序列中医学相

关实体之间的关系。

1.2.2 基于医学知识图谱的知识表示学习方法

医学知识图谱是通过三元组(实体-关系-实体)形式系统整合疾病、症状、药物、治疗方案等医学概念及关联关系的结构化知识库^[32]。文献[33]认为图表示学习将继续推动机器学习在生物医学和医疗保健应用中的发展,包括识别复杂特征背后的基因变异、解缠单细胞行为及对健康的影响、协助患者诊断和治疗及开发安全有效的药物。为了在稀疏数据上自动提出医学建议,文献[34]提出医学建议的知识增强归因多任务学习,引入连接疾病、药物、症状和检查的知识图表及通过共享属性和分子结构连接药物的属性图,通过学习疾病、药物、症状和检查之间的相互关系及医学中的内在关系,获得可用于医学建议的症状和药物的统一嵌入表示,但从不足的信息来源中提取医学知识通常导致数据缺乏完整性,降低所提医学知识图谱在现实世界场景中的有效性。文献[35]建立的系统可以从多源异构信息中提取和管理医学知识,采用命名实体识别和关系提取方法,从医学文本中提取知识三元组,提出一个分层实体对齐框架,进一步完善知识,构建一个大规模、高质量、多源和多语言的医疗知识图谱,包括 13 种实体类型和 17 种关系类型,涵盖 403 784 个实体和 1 225 097 个关系实例,通过广泛试验评估该医疗知识图谱的质量。

在模型可解释性的评估层面,尤其在验证复杂黑盒模型的决策逻辑时,除了模型内在的透明性,事后解释方法同样至关重要。以局部可解释模型无关解释(local interpretable model-agnostic explanations, LIME)和沙普利加性解释(Shapley additive explanations, SHAP)为代表的基于归因的方法在医疗影像和数据分析中得到广泛应用^[36]。这些方法通过构建局部代理模型或计算特征贡献度,生成易于理解的解释。然而,在医疗场景下应用这些方法需审慎评估适用性:LIME 生成的解释因随机采样存在不稳定性;SHAP 虽具有博弈论基础,但计算开销较大^[37]。因此,医疗领域的可解释性评估不应仅依赖单一指标或方法,应结合临床可信度验证与忠诚度等维度进行综合考量,确保解释结果真正服务于临床决策。

综上,两类知识表示学习方法虽从特征具象化与知识结构化两个维度切入可解释性问题,但均未突破从技术表征到医学逻辑的转化瓶颈:基于医学属性的方法聚焦哪些特征是重要的,却未回答为什么它们是重要的;基于医学知识图谱的方法聚焦知

识有哪些关联,却未解决如何追溯关联逻辑的推理透明性问题。这种解释与临床认知脱节的共性矛盾,根源在于现有医学知识表示方法仅实现可计算的知识编码,未达成可理解的医学映射。尽管上述知识表示学习方法通过融合医学专家知识实现了对复杂医学实体的结构化建模,但方法可解释性不足的缺陷仍严重制约临床可信决策辅助系统的构建。当前基于深度学习的知识表示方法虽能挖掘数据中的潜在关联,却难以与医学认知体系中的因果逻辑、病理机制形成语义映射,导致知识推理过程存在黑盒化风险。

1.3 医疗大数据的跨模态关联关系挖掘

医学知识表示学习与知识图谱技术通过图嵌入和图神经网络显式建模“疾病-症状-药物”等显性关联关系,模型可解释性显著提升,但在复杂临床场景中仍需要跨模态关联关系挖掘^[38]。医学关系挖掘存在以下瓶颈:跨模态实体关联(如影像组学特征与病理基因表达的动态映射)缺乏自适应语义对齐机制^[39];病程演进中的隐性时序关联(如并发症风险链式传导)受限于静态知识图谱的离散化表示,难以实现细粒度解耦与联合推理^[40]。本研究主要聚焦当前最新研究进展的两大维度,即实体关系挖掘与时序关系挖掘,系统梳理两者的核心方法、研究现状及突破路径,为破解上述医学关系挖掘瓶颈提供思路。

1.3.1 实体关系挖掘

实体关系挖掘旨在解构医疗实体(如疾病、症状、药物等)间的结构化关联,涵盖从共现依存到病理机制的复杂关系网络。实体关系挖掘的核心挑战在于如何突破跨模态数据(如文本、影像、基因数据等)异质性导致的语义鸿沟,解决深度学习模型因特征解耦不足引发的可解释性缺陷^[41]。为解决传统的实体关系抽取模型在中文医学文本上效果不佳的问题,文献[42]提出中文医学关系实体联合抽取网络,采用多层次语义融合策略,网络核心为关系融合模块,将关系信息融合到句子表示中,实现实体和关系的充分交互,引入双向长短期记忆网络全面捕捉句子特征。医学图像包含大量的医学实体(如不同的解剖区域和病变),在 3D 体积空间中表现出高度不一致的空间模式,导致现有方法中对医学实体偏倚学习,从而在生成报告中产生重复性和不准确性。针对上述问题,文献[43]提出一种医学实体平衡提示网络,利用大语言模型(large language model, LLM)公正地解释各种实体,准确生成脑 CT 报告,为了提取实体的视觉嵌入,基于知

识的联合注意力,利用显式和隐式医学知识探索和提炼实体模式,设计一个学习状态评分器,用于评估实体视觉嵌入的学习情况,为每个实体生成独特的学习状态,将实体视觉嵌入和状态整合到多模态提示中,指导 LLM 的文本生成,使 LLM 能够自适应有偏差实体的学习过程,从而在生成报告中涵盖详细的发现。

1.3.2 时序关系挖掘

时序关系挖掘关注医疗事件在时间轴上的演化规律(如病程进展、用药反应),捕捉数据中的动态依赖^[44]。时序关系挖掘的瓶颈集中于长程时间依赖建模的失效及医疗时序数据中噪声对关键模式提取的干扰^[45]。文献[46]提出一种基于张量权重矩阵的双向门控循环单元网络(tensor-based bidirectional gated recurrent unit, Tensor-BiGRU)和分段注意力机制的关系抽取模型,基于张量权重矩阵改进 BiGRU 网络的编码方式,提升神经网络捕获底层特征的能力,提出两种分段注意力机制,提高模型捕获长句特征的性能。基于图的方法能有效捕捉 EHR 中固有的图结构关系,显著提升下游任务性能,但由于 EHR 数据存在高稀缺性等问题,提取医疗访问的平均时间表示仍面临挑战。文献[47]提出图形变换器-双向编码器表示模型(graph Transformer BERT, GT-BERT),结合图形变换器提取的时间访问嵌入与基于 BERT 的模型获得更具鲁棒性的患者表示,基于图形的方法使 GT-BERT 能够隐式捕获医学观察指标间的内在图形关系,BERT 模型能够提取访问之间的时间关系,近似模拟临床医生的决策过程,引入时间图转换器将时间边特征、全局位置编码和局部结构编码整合到异构图卷积中,捕捉时间信息和结构信息。近年来,随着视觉-语言预训练(vision-language pre-training, VLP)技术的突破,医疗领域的跨模态关联挖掘进入新阶段。这类方法在大规模对齐的图像和文本对上进行预训练,能够学习强大的跨模态通用表示。受对比语言-图像预训练(contrastive language-image pre-training, CLIP)模型的启发,文献[48]开发了医学适配模型 MedCLIP,通过在大规模医学影像报告数据集上进行对比学习,实现医学影像与报告文本的语义对齐,显著提升影像报告生成、零样本医学图像分类等任务的性能。在文本模态方面,基于医学文献和电子健康记录训练的领域预训练模型已成为处理生物医学文本的新基石。跨模态注意力机制等深度学习方法融合深度语言模型与视觉模型已构成前沿的医疗多模态研究范式,能够更精细地挖掘影像中病

灶特征与文本描述临床表现之间的深层语义关联,为精准医疗提供强有力的工具^[49]。

综上,实体关系挖掘的核心矛盾是跨模态覆盖广度与医学关系深度的失衡,虽能实现部分模态的实体关联捕捉,但未触及跨模态病理机制的本质;时序关系挖掘则陷入长序列建模能力与临床时序特异性的错位,技术上可处理超长序列,但无法精准聚焦临床关键节点与语义逻辑。二者共同指向技术建模与临床需求的深层割裂。现有实体关系挖掘方法仅实现数据层面的关系提取,未达成医学层面的关系解读,与临床中多模态证据融合、时序因果推理的决策需求相去甚远。

2 技术应用

随着 AI 技术的突破性发展,数据挖掘在医疗健康领域的核心价值正通过多维应用场景加速释放。作为医疗信息化进程中的关键技术,数据挖掘从海量异构数据中揭示疾病演化规律与诊疗模式,将临床数据的潜在价值转化为可执行的医疗决策^[50-51]。聚焦疾病预测、医学诊断报告生成、诊疗方案推荐三大医疗任务,系统梳理数据挖掘技术在临床知识发现中的创新应用,剖析当前研究在跨模态数据融合、动态决策支持等方面的技术瓶颈,为构建闭环式智慧医疗生态系统提供理论支撑与实践指引。

2.1 疾病预测

计算机辅助诊断作为智慧医疗的核心应用之一,其核心价值在于运用 AI 技术提升临床决策精准度。该系统通过整合多模态医疗数据,实现对病灶特征的智能识别与定位,为临床医生提供可视化决策支持。文献[52]提出一个综合机器学习框架,利用一系列生物标志物预测英国生物库中的 3 213 种疾病,通过英国生物库的纵向健康记录数据预测招募的未确诊突发疾病病例,表现优于现有的多基因风险分数方法。对于慢性病诊疗,早期精准发现心力衰竭对临床试验和治疗至关重要,然而现有心脏疾病识别方法仅根据血压、胆固醇水平、心率等特征划分患者的心脏病类型,缺乏互补表示信息。文献[53]提出一种基于自注意力机制的 Transformer 模型,预测心血管疾病(cardiovascular disease, CVD)风险,通过综合方法对组学数据进行有效综合分析,表征遗传变异及与新兴表型相关的表达模式,揭示新的生物标志物,根据个性化风险因素对患者群体进行细分。文献[54]提出一种基于传统生物信息学、经典统计学和多模态机器学习技术的集成方法,利用强大的特征选择方法,确定 27 个转录组特征和单核苷酸多态性特

征,能够有效预测 CVD,通过差异表达分析并结合最小冗余最大相关性特征选择,发现能够解释疾病表型的生物标志物。

2.2 医学诊断报告生成

医疗报告自动生成在提升临床决策效率与保障诊疗规范性方面具有重要价值,核心挑战在于将医学影像特征转化为符合临床规范的结构化文本描述^[55]。当前研究集中于多模态医学数据与专业文本的语义映射建模,尤其在医学影像诊断领域,通过计算机视觉技术解析图像病理特征,结合自然语言处理生成包含病灶定位、形态特征及诊断建议的专业报告。LLM 为医疗专业人员精准预测提供新的方法,但在训练过程中可能存在无意识算法偏见,影响应用效果。针对上述问题,文献[56]基于真实的患者案例,使用 LLMs 生成患者住院情况、费用和死亡率的预测结果,通过人工复核生成预测结果中的大模型算法偏见,在患者背景生成、特定疾病与特定种族和族群的关联及治疗建议的差异等方面能够明显反映现实世界中的医疗差异。在许多临床研究中,高质量医学成像数据集的稀缺性阻碍了 AI 临床应用的潜力。为了应对这一挑战,文献[57]提出一种统一医学图像-文本生成模型,能够根据文字指令生成各种成像模态下不同器官的医学图像,眼科方面的分类精度提高 12%,胸部方面的分类精度提高 15%,大脑方面的分类精度提高 13%,乳房相关任务中的分类精度提高 17%,通过大规模的回归模拟分析和肺癌 CT 图像准确识别靶向治疗敏感的表皮生长因子受体突变,证明了模型的临床潜力。

2.3 诊疗方案推荐

诊疗方案推荐作为智慧医疗领域的核心技术突破点,正推动临床决策从粗放式经验判断向精细化数据驱动范式转型。该系统通过整合多维度医疗数据(包括患者基因组信息、影像学特征、实时生理指标等),结合知识图谱的多模态推理框架与动态优化算法,构建个性化诊疗路径的量化评估体系,其核心价值不仅在于辅助医生优化临床决策路径,更通过揭示诊疗手段与疗效间的隐性关联规律,为循证医学提供可量化的决策依据^[58]。然而,诊疗方案推荐方法目前面临临床动态决策的时序依赖性建模不足、罕见病数据稀疏性导致模型偏差等问题,仍需解决医学先验知识嵌入与数据驱动学习的平衡难题。该领域的突破将加速精准医疗从单病种静态推荐向多并发症动态决策的跨越式发展。为了准确推荐个性化临床路径,文献[59]构建一个以国家标准临床路径为基础的预训练自注意力模型,从实际临床数据集中获取患者的诊疗史序

列,使用自注意力模型进行训练,学习临床项目之间的关系,满足患者的个人需求。

当前研究在疾病预测、医学诊断报告生成与诊疗方案推荐等核心场景中已初步验证多模态医疗数据挖掘的临床价值,但技术瓶颈集中体现在异构数据的动态对齐能力不足、医学知识嵌入的深度受限及决策支持的时空连续性断裂等层面。疾病预测模型虽通过混合神经网络架构提升了时序特征捕获能力,却难以量化多源特征对个体化预测的贡献权重;医学诊断报告生成系统虽借助多模态大模型实现影像-文本的跨模态映射,但在临床语义规范性与因果逻辑一致性上仍存在显著偏差;诊疗方案推荐算法虽引入知识图谱增强可解释性,却尚未突破罕见病数据稀疏性与治疗路径动态演化的建模困境。因此,迫切需要从底层技术架构层面重新审视多源异构医疗大数据的融合机制与价值释放路径。

3 多源异构医疗大数据前沿技术

当前,多源异构医疗大数据的前沿技术体系面临三大核心问题:单一模态处理能力有限,难以有效挖掘多源数据间的深层关联;过度依赖数据统计规律,缺乏对医学领域先验知识的有效建模与利

用,导致资源有限场景下知识表征欠缺,模型泛化受限;黑盒模型决策过程不可知,缺乏可验证性和可追溯性,影响医疗应用的可靠性与信任度。针对上述问题,本研究对现有研究进开展系统调研、分析、总结、创新,形成特征交互驱动的跨模态融合可解释框架技术理论体系,如图 1 所示,提出一种多模态可解释融合框架,通过多模态输入层提取单一模态特征,在融合层借助跨模态记忆矩阵与动态加权单元实现语义对齐,并引入医学知识图谱实施关系约束。该框架支持双向反馈与知识引导,最终由解码层输出结构报告,在保证融合精度的同时增强临床合理性与可解释性。特征交互驱动的跨模态融合可解释框架包含以下 3 个范式:其一,多模态融合范式^[60],整合多模态数据,以充分挖掘数据间关联;其二,数据驱动与知识引导融合范式^[61],突破对纯数据驱动的依赖,利用先验知识缓解小样本约束;其三,可解释模型透明化推理范式^[62],发展可解释机制,确保决策过程可追溯、结果可验证。上述 3 个范式是针对核心问题开展的深度研究,构建数据-知识双轮驱动的智能医疗新范式,不仅显著降低框架对人工标注数据的过度依赖^[63],更能有效应对医疗领域普遍存在的小样本挑战,为精准诊疗和个性化治疗等核心需求提供坚实、可追溯验证的技术支撑。

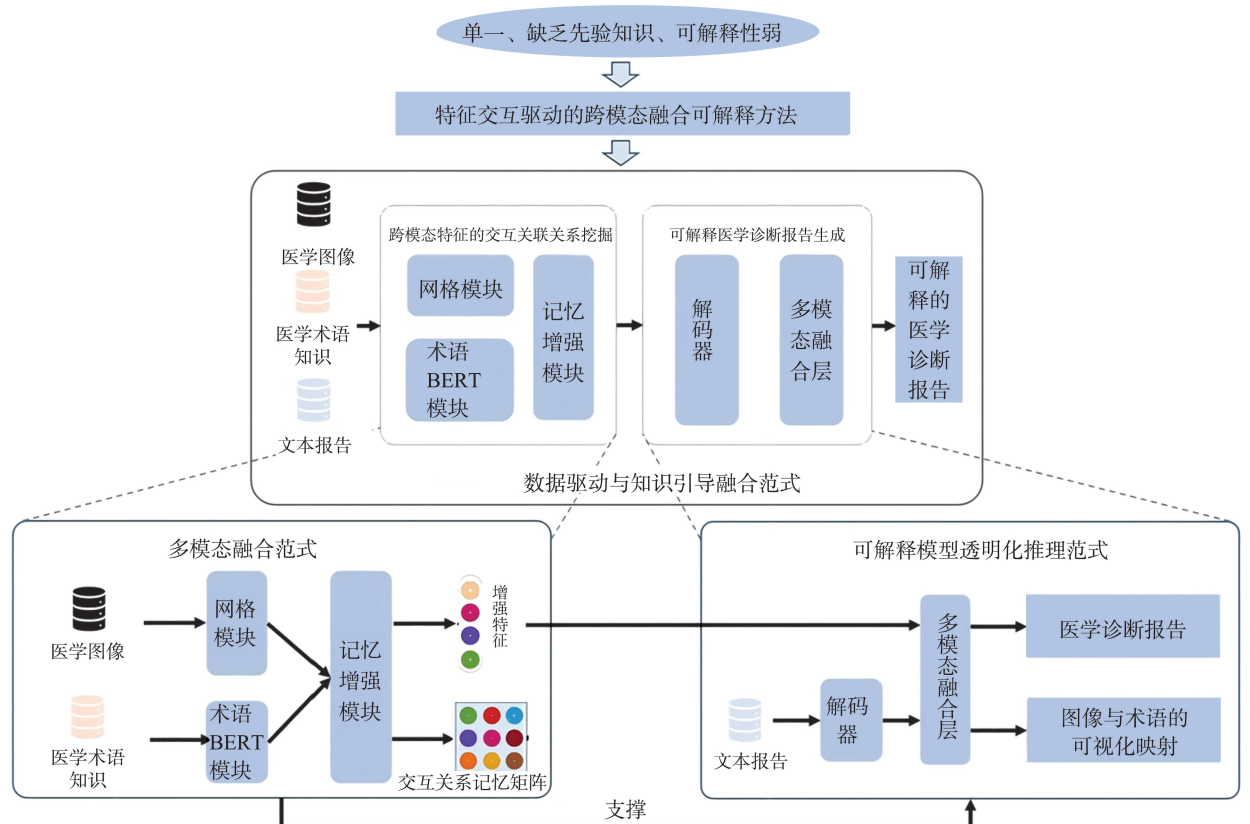


图 1 特征交互驱动的跨模态融合可解释框架

Fig.1 Feature interaction-driven interpretable framework for cross-modal fusion data

3.1 多模态融合范式

作为前沿技术体系的数据基础层,多模态融合范式的核心定位是解决单一模态处理能力有限、多源数据深层关联难挖掘的问题,通过整合影像、文本、基因等多源异构数据,构建跨模态特征交互机制,为后续数据与知识协同提供高质量的整合数据支撑。多模态记忆 Transformer 网络(multi-modal memory Transformer network, MMTN)^[64]、跨模态记忆 Transformer(cross-modal memory Transformer, CMT)^[65]等框架是实现多源数据从割裂到协同的关键技术,输出的统一化多模态特征将直接作为数据驱动与知识引导融合范式的核心输入。

针对医学影像报告生成任务中图像与文本语义割裂、关键术语缺失及描述不一致的问题,文献[64]提出 MMTN 模型,创新设计多模态记忆编码器与跨模态融合机制。首先,通过网格模块和术语 BERT 模块分别建模视觉与知识模态;其次,引入可学习的记忆矩阵,显式关联图像特征与医学术语表征(如将“结肠息肉”语义绑定至对应病变区域),构建视觉-术语对齐桥梁;最后,在解码阶段设计动态加权融合层,自适应平衡视觉特征与语言状态的贡献,确保生成的报告能够同时满足术语精准性和上下文连贯性。该框架通过跨模态记忆对齐与动态特征融合攻克医学报告生成的核心痛点,为临床诊断自动化提供高可靠性工具,大幅降低医生撰写负担,并可扩展至内镜、X 光等多模态医疗场景,推动 AI 辅助诊断向精准化、标准化演进。

针对医学影像报告生成中跨模态交互不足与多模态数据利用不充分的问题,现有基于 Transformer、强化检索生成模型(reinforced retrieval generator, R2Gen)的方法存在两大问题:一是忽略图像与文本特征的深层交互,导致报告关键信息遗漏,如病灶定位偏差或逻辑矛盾;二是未能自适应融合影像、术语、文本三重模态特征,造成术语误用,如混淆 Yamada 息肉与普通息肉。为解决上述问题,文献[65]设计 CMT 编码图像和报告特征之间的关系。利用 CNN 和 Transformer 提取图像特征;医学术语增强模块采用 BERT 编码专业术语库;跨模态记忆解码器通过动态记忆矩阵记录图像到文本的映射关系,在生成词汇时追溯视觉特征,以解决图文不一致现象;多模态融合模块通过自适应加权三模态特征,实现多源信息有效融合。该框架实现影像-术语-文本联合学习,为多模态医疗数据挖掘提供新范式,有效缓解医生撰写负担并提升基层医疗报告标准化水平。CMT 通过跨模态记忆机

制打破特征壁垒,以动态加权实现精准融合,不仅解决特征割裂与术语失准问题,更为医疗 AI 落地提供可扩展路径。

针对医疗报告生成中医生反馈难以融入、关键病灶遗漏和人工依赖过高的挑战,文献[66]提出反馈驱动的自适应优化多模态 LLM 协作框架,创新性构建 3 层架构:第 1 层为生成模块,基于 Transformer 融合视觉图像编码与文本解码,通过跨模态注意力生成初稿;第 2 层为强化学习(reinforcement learning, RL)模块,将医生修正转化为三阶奖励信号(状态信号为修改动作,动作信号为内容调整,奖励信号为满意度评分),以策略梯度反馈优化策略;第 3 层为 LLM 蒸馏模块,用 LLM 代替人工评估,自动生成奖励信号驱动模型迭代,消除传统 RL 对奖励模型训练的依赖。该框架实现修正样本实时回流机制,医生仅需标注错误区域,系统自动构建(原图,修正报告)键值训练对,使模型自迭代提速 3 倍。LLM 奖励引擎将 ChatGPT 等模型转化为虚拟医生评审,降低 80% 人工标注成本。该框架成功解决医疗 AI 落地的核心矛盾,在保障专业性的同时实现高自动化,为重症监护室、急诊等场景提供分钟级报告生成支持,更可扩展至病理诊断、工业质检领域,重塑专业报告生成范式;不仅建立医生反馈驱动的动态医疗报告系统,更开创 RL-LLM 技术路线,为高可靠性人机协作提供新范式,推动 AI 从辅助工具向自主临床工具演进。

综上所述,多模态融合范式的应用提高了模型的透明度和可靠性,为临床医生提供有价值的决策支持。通过整合多模态数据,这一理论体系在提高医疗数据的质量和准确性、减少因错误信息导致的医疗风险方面具有显著作用。这种多模态协同框架正在重塑智能诊疗范式,为临床决策支持系统提供更接近人类专家思维模式的综合分析能力。未来,多个研究方向值得深入探索:开发更智能的跨模态对齐算法,实现更高精度的多模态融合;构建动态可解释模型,使模型根据实时数据动态调整解释逻辑;结合因果推断方法,深入挖掘不同模态数据之间的因果关系,进一步提升模型的可解释性和预测能力。这些研究方向将共同推动医学 AI 从多模态协同学习向更深层次的人工智能迈进。

3.2 数据驱动与知识引导融合范式

作为前沿技术体系的协同核心层,数据驱动与知识引导融合范式的核心定位是突破过度依赖数据统计、小样本场景泛化受限的瓶颈,在多模态融合范式的基础上引入医学领域先验知识(如国际疾

病分类(international classification of diseases, ICD)、疾病-症状关联图谱),构建数据统计规律和知识逻辑约束的双轮驱动机制。异构医疗实体交互感知的图注意力融合框架^[67]、融合知识嵌入与自注意力的ICD编码修正框架^[68]等既承接多模态数据的整合成果,又通过知识注入提升数据表征的医学合理性,为后续可解释模型的可信推理奠定语义基础。

面对EHR数据的复杂异构医疗关系难建模的挑战,大多数研究试图从EHR中提取与疾病风险相关的特征,忽略与疾病风险相关的异构实体(如疾病、症状、药物和其他治疗项目之间的复杂交互),导致疾病风险预测不准确。为了解决上述问题,文献[67]提出异构医疗实体交互感知的图注意力融合框架,构建融合患者、疾病、症状的异构医疗图,设计双重注意力机制,通过节点级注意力动态学习邻居重要性,结合语义级注意力解析元路径,捕捉潜在临床关联。在嵌入计算阶段,采用图卷积分聚合邻居信息,通过投影层统一异构节点表征,最终解码为疾病概率分布。该框架突破时序依赖限制,实现单次就诊的精准预测,注意力权重揭示临床决策依据,高权重症状驱动疾病诊断,增强模型可解释性,元路径分析(如症状到疾病链)有效挖掘潜在风险关联,为异构医疗数据的知识挖掘提供新范式。

ICD代码一直是EHR的重要组成部分。然而,ICD中的编码错误严重影响基于EHR分析的准确性。传统方法依赖临床文本分析,易受医生主观干预影响。为了检测EHR中的错误疾病信息,文献[68]提出融合知识嵌入与自注意力的ICD编码修正框架,使用药物结构化信息纠正代码分配方法。基于知识嵌入,利用自注意力机制捕捉模型中药物之间的关系;利用药物的解剖学、治疗学及化学分类系统(anatomical therapeutic chemical classification system, ATC)的结构信息,通过双向门控循环单元融合药物节点与其父类节点的语义,生成鲁棒性嵌入向量,突破低频药物表征瓶颈;提出自注意力药物组合建模模块,将患者用药序列编码为矩阵,通过自注意力机制学习不同药物间的相互作用权重,生成表征药物组合关系的二维矩阵,捕捉非线性关联;结合全连接层与激活函数输出诊断错误概率,通过注意力权重定位关键药物组合,提供决策依据。该框架将知识增强嵌入与自注意力机制结合用于EHR纠错,为医疗数据质量控制提供新范式。通过融合药物知识图谱与注意力机制,解决EHR中疾病编码隐蔽错误检测的难题,在精度与可解释性

上均超越现有方法,对提升医疗数据分析可靠性具有重要实践意义。

综上,数据驱动与知识引导融合范式的提出,是对医学AI领域中传统数据驱动方法难以有效捕捉医学复杂逻辑与因果关系问题的革新性突破,核心意义在于通过将结构化医学知识(如疾病共现、ICD层级、合理用药规则等)系统性地注入模型构建过程,显著增强模型的临床合理性与可解释性,突破黑盒模型限制,利用知识约束提升小样本泛化能力,为精准医疗决策提供更可靠、可追溯的智能支撑。未来,聚焦跨模态深度推理表征对齐、因果驱动融合机制、动态自适应知识图谱构建与融合、融合范式的可解释性理论框架与泛化边界探索,并强化伦理化框架及临床工作流无缝集成,向更深层次的动态认知融合与可信赖智能诊疗迈进。

3.3 可解释模型透明化推理范式

作为前沿技术体系的可信保障层,可解释模型透明化推理范式的核心定位是破解黑盒模型决策不可追溯、医疗应用可靠性不足的难题。基于3.1、3.2节,通过结构化知识嵌入、异构信息解耦等技术,将模型决策过程转化为符合临床思维的可追溯证据链。ICD-10知识驱动层级可解释网络(knowledge-driven interpretable network, KdINet)^[69]、数据与关系双驱动交叉注意力可解释模型^[70]、特征引导逻辑感知的可解释规则学习架构^[71]是对前述技术成果的可信化升级,确保多模态融合与知识协同的输出转化为临床可理解、可验证的决策支持。

现有基于深度学习的医疗诊断方法通常是黑盒模型,隐性决策过程使模型无法解释。为了缓解这一问题,文献[69]提出KdINet,用于医学图像的可解释疾病分类。KdINet将医学图像特征与ICD-10层级语义向量相融合,通过特征聚合层进行注意力加权,依据分类反馈动态调整层级权重。其核心机制在于施加层级语义约束并输出“祖先-父-子”节点的完整推理路径,使诊断决策可追溯至标准医学逻辑。KdINet利用预训练的CNN模块和分层表示模块学习视觉疾病特征和分层疾病特征,疾病分类器学习两种疾病特征的联合训练分层分类标准,以推断诊断,并生成相应的可解释理由,即祖先疾病的诊断路径。所有疾病特征均生成分层解释理由,例如为直肠息肉生成的分层解释理由是消化性疾病、肠道疾病和直肠疾病。这些是ICD-10疾病分类标准中直肠息肉的祖先疾病。上述祖先疾病路径解释了诊断是如何一步步产生的。

EHR疾病预测模型将诊断与治疗信息混同处

理,导致病程轨迹模糊不清,关键特征被湮没,模型解释性严重不足,极大限制模型在实际精准医疗中的应用效果与可信度。为有效破解这一难题,文献[70]提出一种数据与关系双驱动的交叉注意力可解释模型。该模型采用独特的双路分离架构,运用独立的双向循环网络,一路针对诊断信息进行建模,精准反映疾病的实际状态,另一路聚焦治疗信息,表征医疗干预这一关键序列。同时,该模型提出历史就诊交叉注意力机制,在计算历史就诊权重时,将诊断重要性评估与同期治疗信息进行巧妙融合,治疗重要性的判定紧密关联对应诊断。这种双向耦合机制能够精准捕捉诊疗之间的内在关联性,借助权重可视化手段清晰地揭示驱动预测结果产生的核心诊疗组合,为模型的解释性提供强有力支撑。

EHR 数据虽为疾病风险预测提供重要支持,但现有深度学习模型的预测过程缺乏透明性,诊断、用药、手术等异构特征被混杂处理,导致无法量化各类特征的重要性。针对上述问题,文献[71]提出特征引导逻辑感知的可解释规则学习架构,用于健康风险预测。该架构构建两阶段推理框架:第一阶段采用双层逻辑感知(two-layer logical perception, TLLP)单元,从二值化特征中提取可解释规则,将诊断、用药、手术等特征通过二值化函数转化为二进制向量,应用 TLLP 单元生成各类特征对应的逻辑规则向量;第二阶段通过全连接层融合规则向量,形成患者综合逻辑表示,针对不同预测任务动态设计输出层,即疾病预测任务结合疾病-特征矩阵计算疾病概率分布,死亡率预测任务直接输出风险评估。该架构提出逻辑感知网络,实现 EHR 数据规则化表示,提出特征分治架构,量化异构特征重要性,在疾病或死亡率预测中兼顾预测精度与推理透明度,为发现新的生物标志物和开发人类疾病的预测算法提供新机会。

综上所述,可解释模型透明化推理范式通过结构化知识嵌入、异构信息解耦及逻辑规则驱动,实现医疗 AI 决策从黑盒预测向可追溯、可验证的临床思维再现跃迁,核心突破体现为诊断依据的逐步溯源能力、预测规则与医学指南的强对齐性及医生可干预的决策闭环。未来研究将着力构建跨模态认知对齐框架,建立动态知识融合引擎,通过临床验证特征归因、逻辑完备、诊疗行为改变率三阶评价体系,量化解释置信边界,形成符合医疗伦理的闭环决策框架,使 AI 成为具备逻辑透明、错误可溯、认知协同的可信赖智能伙伴,通过国际临床认

证推动可信诊疗革命。

4 结束语

在数字经济与健康中国战略双重驱动下,多源异构医疗大数据正从基础资源向新质生产力跃迁,但多源异构性、碎片化存储及质量参差等特征导致信息孤岛与分析断层问题并存,制约数据价值释放。当前研究围绕三大技术路径突破:多模态融合通过分层整合(数据级-特征级-决策级)增强疾病表征全面性,但跨模态语义对齐与关联断裂仍为核心瓶颈;知识发现结合医学知识图谱与属性嵌入提升模型可解释性,需解决医学逻辑与数据因果推理的脱节;关联关系挖掘从静态实体与动态时序揭示规律,但长程依赖建模与噪声干扰限制隐性模式提取。应用层面,疾病预测、诊断报告生成与诊疗推荐虽借助混合模型、多模态大模型及知识图谱推动精准化,仍面临罕见病数据稀疏、语义规范不足及动态关系复杂等挑战。回顾本研究团队提出的多源异构医疗大数据前沿技术的多模态融合范式、数据驱动与知识引导融合范式及可解释模型透明化推理范式,突破单一模态局限,减轻小样本约束,增强模型可信度,为精准诊疗提供坚实技术支撑。从实验室走向临床科室仍面临多重现实挑战。在数据层面,医疗机构间的数据标准不一,接口封闭,形成“数据烟囱”,日益严格的隐私合规要求使跨机构数据汇聚步履维艰;在系统层面,新研发的 AI 模块与医院部署多年的核心系统难以深度融合,接口复杂,改造周期长,运维成本高,严重制约技术的规模化落地;在应用层面,医生对黑盒模型的信任度普遍不高,算法的决策逻辑若不能转化为符合临床思维、可回溯、可验证的辅助信息,将难以融入实际诊疗流程,使人机协同面临接受度挑战。

未来,共同致力于推动医学与工程领域的实质融合,组建涵盖临床医生、数据科学家与医院信息部门的跨学科团队。从真实临床场景出发,共同设计兼顾合规与易用的数据治理方案,开发基于中间件或标准化接口的轻量级系统集成策略,构建以医生工作流为中心、决策过程透明可解释的可信交互机制,系统化打通数据-系统-人 3 层闭环,将医疗大数据潜力转化为提升诊疗效率和人民健康水平的现实力量。

参考文献:

[1] 中共中央, 国务院. 关于构建数据基础制度更好发挥数

- 据要素作用的意见[R/OL]. (2022-12-19)[2025-09-05]. http://www.gov.cn/zhengce/2022-12/19/content_5732698.htm
- [2] 国家数据局, 中央网信办, 科技部, 等. “数据要素×”三年行动计划(2024—2026年)[R/OL]. (2024-01-04)[2025-09-05]. http://www.gov.cn/zhengce/zhengceku/202401/04/content_6925556.htm
- [3] 张振, 周毅, 杜守洪, 等. 医疗大数据及其面临的机遇与挑战[J]. 医学信息学杂志, 2014, 35(6): 2-8. ZHANG Zhen, ZHOU Yi, DU Shouhong, et al. Medical big data and the facing opportunities and challenges[J]. Journal of Medical Informatics, 2014, 35(6): 2-8.
- [4] 曾旺旺, 谢颖夫, 胡光阔. 多源异构数据整合系统在医疗大数据中的应用[J]. 价值工程, 2017, 36(8): 80-82. ZENG Wangwang, XIE Yingfu, HU Guangkuo. Application of multi-source heterogeneous data integration system in large medical data[J]. Value Engineering, 2017, 36(8): 80-82.
- [5] 刘震, 王文桥. 基于区块链的医疗信息共享平台设计与实现[J]. 医疗卫生装备, 2020, 41(8): 36-39. LIU Zhen, WANG Wenqiao. Design and implementation of medical information sharing platform based on blockchain[J]. Chinese Medical Equipment Journal, 2020, 41(8): 36-39.
- [6] AMINIZADEH S, HEIDARI A, DEGHAN M, et al. Opportunities and challenges of artificial intelligence and distributed systems to improve the quality of healthcare service[J]. Artificial Intelligence in Medicine, 2024, 149: 102779.
- [7] JENSEN P B, JENSEN L J, BRUNAK S. Mining electronic health records: towards better research applications and clinical care[J]. Nature Reviews Genetics, 2012, 13(6): 395-405.
- [8] KADER F, NOYES P, IRUKA I U, et al. Addressing maternal and infant health inequities requires improved birth record data collection[J]. Nature Medicine, 2025, 31(2): 358-359.
- [9] RAO V M, HLA M, MOOR M, et al. Multimodal generative AI for medical image interpretation[J]. Nature, 2025, 639(8056): 888-896.
- [10] BEN-MILED Z, SHEBESH J A, SU J, et al. Multi-modal fusion of routine care electronic health records (EHR): a scoping review[J]. Information, 2025, 16(1): 54.
- [11] CAVUTO M L, MALPARTIDA-CARDENAS K, PENNISI I, et al. Portable molecular diagnostic platform for rapid point-of-care detection of mpox and other diseases[J]. Nature Communications, 2025, 16: 2875.
- [12] HU T, KE X X, YU Y Y, et al. NAPTUNE: nucleic acids and protein biomarkers testing via ultra-sensitive nucleases escalation[J]. Nature Communications, 2025, 16(1): 1331.
- [13] 陈驰华, 周婷, 廖凯兵. CT影像特征及影像组学在肺部淋巴瘤与肺浸润性黏液腺癌诊断中的应用[J]. 中国CT和MRI杂志, 2023, 21(9): 82-85. CHEN Chihua, ZHOU Ting, LIAO Kaibing. Application of CT imaging features and radiomics features in the differential diagnosis of pulmonary lymphoma and pulmonary invasive mucinous adenocarcinoma[J]. Chinese Journal of CT and MRI, 2023, 21(9): 82-85.
- [14] JIAO C N, GAO Y L, GE D H, et al. Multi-modal imaging genetics data fusion by deep auto-encoder and self-representation network for Alzheimer's disease diagnosis and biomarkers extraction[J]. Engineering Applications of Artificial Intelligence, 2024, 130: 107782.
- [15] WANG M Y, FAN S Y, LI Y C, et al. Missing-modality enabled multi-modal fusion architecture for medical data[J]. Journal of Biomedical Informatics, 2025, 164: 104796.
- [16] MAHATO K, SAHA T, DING S C, et al. Hybrid multimodal wearable sensors for comprehensive health monitoring[J]. Nature Electronics, 2024, 7(9): 735-750.
- [17] 曹益铭. 多模态数据与知识双驱动的医学诊断报告生成关键技术研究[D]. 济南: 山东大学, 2023: 71-89. CAO Yiming. Research on key technologies for medical diagnosis report generation driven by multi-modal data and knowledge[D]. Jinan: Shandong University, 2023: 71-89.
- [18] DAI S Y, YE K, ZHAN C, et al. SIN-Seg: a joint spatial-spectral information fusion model for medical image segmentation[J]. Computational and Structural Biotechnology Journal, 2025, 27: 744-752.
- [19] 李立柱, 孟明, 高云园, 等. 基于小波变换的 EEG-fNIRS多模态数据融合方法[J]. 传感技术学报, 2023, 36(7): 1064-1072. LI Lizhu, MENG Ming, GAO Yunyuan, et al. A data fusion method for hybrid EEG-fNIRS BCI base on wavelet transform[J]. Chinese Journal of Sensors and Actuators, 2023, 36(7): 1064-1072.
- [20] RELE M, JULIAN A, PATIL D, et al. Multimodal data fusion integrating text and medical imaging data in electronic health records[C]//Innovations and Advances in Cognitive Systems. Cham, Switzerland: Springer, 2024: 348-360.
- [21] 高泽文, 王建, 魏本征. 基于混合偏移轴向自注意力机制的脑胶质瘤分割算法[J]. 山东大学学报(工学版), 2024, 54(2): 80-89.

- GAO Zewen, WANG Jian, WEI Benzhen. Glioma segmentation algorithm based on hybrid offset axial self-attention mechanism[J]. Journal of Shandong University (Engineering Science), 2024, 54(2): 80-89.
- [22] 张贺贺. 融合弱监督和动量蒸馏的多模态预训练模型研究[D]. 徐州: 中国矿业大学, 2023: 34-47.
- ZHANG Hehe. Multimodal pretraining model with weak supervision and momentum distillation[D]. Xuzhou: China University of Mining and Technology, 2023: 34-47.
- [23] 李伟豪, 王莘莘, 许万博, 等. 结构先验引导的多模态腰MRI图像分割算法[J]. 山东大学学报(工学版), 2025, 55(1): 66-76.
- LI Weihao, WANG Pingping, XU Wanbo, et al. Multimodal lumbar MRI image segmentation algorithm guided by structure priori[J]. Journal of Shandong University(Engineering Science), 2025, 55(1): 66-76.
- [24] WANG Y L, YIN C C, ZHANG P. Multimodal risk prediction with physiological signals, medical images and clinical notes[J]. Heliyon, 2024, 10(5): e26772.
- [25] SHAKIR M A, ABASS H K, JELWY O F, et al. Developing interpretable models for complex decision-making[C]//2024 36th Conference of Open Innovations Association (FRUCT). Lappeenranta, Finland: IEEE, 2024: 66-75.
- [26] ZHANG Y Y, FANG Q, QIAN S S, et al. Multi-modal multi-relational feature aggregation network for medical knowledge representation learning[C]//Proceedings of the 28th ACM International Conference on Multimedia. Seattle, USA: ACM, 2020: 3956-3965.
- [27] JI S X, PAN S R, CAMBRIA E, et al. A survey on knowledge graphs: representation, acquisition, and applications[J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(2): 494-514.
- [28] 何鹏, 姚瑶, 刘秋菊. 时态知识图谱表示学习研究综述[J]. 计算机工程与应用, 2025, 61(14): 37-53.
- HE Peng, YAO Yao, LIU Qiuju. Survey on temporal knowledge graph representation learning[J]. Computer Engineering and Applications, 2025, 61(14): 37-53.
- [29] 叶青, 刘丹红, 杨喆, 等. 临床指南模型构建中医学知识的规范化表达方法[J]. 中国数字医学, 2011, 6(8): 14-17.
- YE Qing, LIU Danhong, YANG Zhe, et al. Representation of medical concepts in computer-interpretable guideline modeling[J]. China Digital Medicine, 2011, 6(8): 14-17.
- [30] WANG Z J, LONG H Y, YU H, et al. DKGC-LSTM: fusion of domain knowledge to guide CNN and LSTM for heart failure risk prediction[C]//2024 IEEE International Conference on Medical Artificial Intelligence (MedAI). Chongqing, China: IEEE, 2024: 552-557.
- [31] LEE S, HAN Y J, PARK H J, et al. Entity-enhanced BERT for medical specialty prediction based on clinical questionnaire data[J]. Plos One, 2025, 20(1): e0317795.
- [32] 王钰涵, 马涪元, 王英. 基于深度学习的细粒度医学知识图谱构建[J]. 计算机科学, 2024, 51(增刊2): 230900157.
- WANG Yuhan, MA Fuyuan, WANG Ying. Construction of fine-grained medical knowledge graph based on deep learning[J]. Computer Science, 2024, 51(Suppl.2): 230900157.
- [33] LI M M, HUANG K X, ZITNIK M. Graph representation learning in biomedicine and healthcare[J]. Nature Biomedical Engineering, 2022, 6(12): 1353-1369.
- [34] ZHANG Y Y, WU X, FANG Q, et al. Knowledge-enhanced attributed multi-task learning for medicine recommendation[J]. ACM Transactions on Information Systems, 2023, 41(1): 1-24.
- [35] YANG P R, WANG H J, HUANG Y Z, et al. LMKG: a large-scale and multi-source medical knowledge graph for intelligent medicine applications[J]. Knowledge-Based Systems, 2024, 284: 111323.
- [36] RIBEIRO M T, SINGH S, GUESTRIN C. "Why should I trust you?": explaining the predictions of any classifier[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco, USA: ACM, 2016: 1135-1144.
- [37] LUNDBERG S M, LEE S I. A unified approach to interpreting model predictions[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA: ACM, 2017: 4768-4777.
- [38] WANG F Y, ZHOU Y Y, WANG S J, et al. Multi-granularity cross-modal alignment for generalized medical visual representation learning[EB/OL]. (2022-10-12) [2025-09-05]. <https://doi.org/10.48550/arXiv.2210.06044>
- [39] MAO L C, WANG H R, HU L S, et al. Knowledge-informed machine learning for cancer diagnosis and prognosis: a review[J]. IEEE Transactions on Automation Science and Engineering, 2025, 22: 10008-10028.
- [40] LIANG K, MENG L Y, LIU M, et al. A survey of knowledge graph reasoning on graph types: static, dynamic, and multi-modal[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(12): 9456-9478.
- [41] BYEON H, RAINA V, SANDHU M, et al. Artificial

- intelligence-enabled deep learning model for multimodal biometric fusion[J]. *Multimedia Tools and Applications*, 2024, 83(33): 80105-80128.
- [42] 盛西方, 赵俊东, 陶青川, 等. MedRelNet: 基于关系融合的中文医学文本实体关系联合抽取模型[J]. *现代计算机*, 2024, 30(22): 49-54.
SHENG Xifang, ZHAO Jundong, TAO Qingchuan, et al. MedRelNet: a joint model for entity-relation extraction in Chinese medical text based on relation fusion[J]. *Modern Computer*, 2024, 30(22): 49-54.
- [43] ZHANG X D, SHI Y Z, JI J Z, et al. MEPNet: medical entity-balanced prompting network for brain CT report generation[EB/OL]. (2025-03-22) [2025-09-05]. <https://arxiv.org/abs/2503.17784>
- [44] WINTER M, LANGGUTH B, SCHLEE W, et al. Process mining in mHealth data analysis[J]. *NPJ Digital Medicine*, 2024, 7: 299.
- [45] ZHANG F, GUO T T, WANG H. DFNet: decomposition fusion model for long sequence time-series forecasting[J]. *Knowledge-Based Systems*, 2023, 277: 110794.
- [46] 隗昊, 唐焕玲, 周爱, 等. 基于双路分段注意力神经张量网络的临床文本关系抽取[J]. *电子学报*, 2023, 51(3): 658-665.
WEI Hao, TANG Huanling, ZHOU Ai, et al. Clinical relation extraction via dual piecewise attention neural tensor network[J]. *Acta Electronica Sinica*, 2023, 51(3): 658-665.
- [47] POULAIN R, BEHESHTI R. Graph Transformers on EHRs: better representation improves downstream performance[C]//The Twelfth International Conference on Learning Representations. Vienna, Austria: [s.n.], 2024: 8705.
- [48] WANG Z F, WU Z B, AGARWAL D, et al. MedCLIP: contrastive learning from unpaired medical images and text[C]// Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP). Abu Dhabi, United Arab Emirates: ACL, 2022: 3876-3887.
- [49] ALSENTZER E, MURPHY J R, BOAG W, et al. Publicly available clinical BERT embeddings[C]// Proceedings of the 2nd Clinical Natural Language Processing Workshop (ClinicalNLP). Minneapolis, USA: ACL, 2019: 72-78.
- [50] CHEN J Y, YIN C C, WANG Y L, et al. Predictive modeling with temporal graphical representation on electronic health records[EB/OL]. (2024-05-07) [2025-09-05]. <https://arxiv.org/abs/2405.03943>
- [51] MENG H F, LIN Z Q, YANG F, et al. Knowledge distillation in medical data mining: a survey[C]//5th International Conference on Crowd Science and Engineering. Jinan, China: ACM, 2022: 175-182.
- [52] GARG M, KARPINSKI M, MATELSKA D, et al. Disease prediction with multi-omics and biomarkers empowers case: control genetic discoveries in the UK Biobank[J]. *Nature Genetics*, 2024, 56(9): 1821-1831.
- [53] RAHMAN A U, ALSENANI Y, ZAFAR A, et al. Enhancing heart disease prediction using a self-attention-based Transformer model[J]. *Scientific Reports*, 2024, 14: 514.
- [54] DEGROAT W, ABDELHALIM H, PEKER E, et al. Multimodal AI/ML for discovering novel biomarkers and predicting disease using multi-omics profiles of patients with cardiovascular diseases[J]. *Scientific Reports*, 2024, 14: 26503.
- [55] REALE-NOSEI G, AMADOR-DOMÍNGUEZ E, SERRANO E. From vision to text: a comprehensive review of natural image captioning in medical diagnosis and radiology report generation[J]. *Medical Image Analysis*, 2024, 97: 103264.
- [56] YANG Y F, LIU X Y, JIN Q, et al. Unmasking and quantifying racial bias of large language models in medical report generation[J]. *Communications Medicine*, 2024, 4: 176.
- [57] WANG J Z, WANG K, YU Y F, et al. Self-improving generative foundation model for synthetic medical image generation and clinical applications[J]. *Nature Medicine*, 2025, 31(2): 609-617.
- [58] ZHU F L, CUI L Z, XU Y H, et al. A survey of personalized medicine recommendation[J]. *International Journal of Crowd Science*, 2024, 8(2): 77-82.
- [59] LIN X J, LI Y, XU Y H, et al. Personalized clinical pathway recommendation via attention based pre-training [C]//2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). Houston, USA: IEEE, 2021: 980-987.
- [60] MOGHTADERI S, EINLOU M, WAHID K A, et al. Advancing multimodal medical image fusion: an adaptive image decomposition approach based on multilevel guided filtering[J]. *Royal Society Open Science*, 2024, 11(4): rsos.231762.
- [61] WAN S, GUAN S H, TANG Y C. Advancing bridge structural health monitoring: insights into knowledge-driven and data-driven approaches[J]. *Journal of Data Science and Intelligent Systems*, 2024, 2(3): 129-140.
- [62] HASSIJA V, CHAMOLA V, MAHAPATRA A, et al. Interpreting black-box models: a review on explainable artificial intelligence[J]. *Cognitive Computation*, 2024, 16(1): 45-74.
- [63] CHEN H R, ZHANG S X, ZHANG L Z, et al. Multi role ChatGPT framework for transforming medical data analysis[J]. *Scientific Reports*, 2024, 14: 13930.