计算机系统应用 ISSN 1003-3254, CODEN CSAOBN
[doi: 10.15888/j.cnki.csa.009815] [CSTR: 32024.14.csa.009815]

E-mail: csa@iscas.ac.cn
http://www.c-s-a.org.cn
Tel: +86-10-62661041

基于深度哈希的多模态临床数据相似病例检索①

谢明朗[1], 袁贞明[1], 施军平[1,2], 田　昕[1,2]

[1](杭州师范大学, 杭州 311121)
[2](杭州师范大学附属医院, 杭州 310015)
通信作者: 田　昕, E-mail: 59712781@qq.com

摘　要: 随着电子健康档案 (EHR) 的普及, 相似患者检索已成为支持辅助诊断和制定治疗计划等临床决策的重要任务. 然而, EHR 数据具有高维度、异构性且数量大的特点. 为了有效整合多模态临床数据并实现高效检索, 本文提出了一种基于深度哈希的多模态临床数据相似病例检索模型——MCDF. 该模型根据不同模态数据的特性, 将结构化文本数据、非结构化文本数据、和图像数据分别使用多层感知机 (multi-layer perceptron, MLP) 模型、BioBERT、BioMedCLIP 进行特征提取, 并通过自注意力机制模块进行特征融合. 利用三元组损失函数引导模型直接生成能够有效代表样本的哈希码, 通过哈希码快速比对实现样本检索, 不仅能提高检索的准确性, 还能显著提升检索效率. 本文基于公开数据集 MIMIC-III, 采用归一化折扣累计收益均值 (*MNDCG*) 和均值平均精度 (*MAP*) 作为评价指标, 将 MCDF 模型与传统哈希方法 (如 spectral hashing) 和先进哈希方法 (如 deep hashing network) 进行比较. 实验结果显示, MCDF 模型的表现优于所有基线模型, 验证了本文提出模型的优越性.

关键词: 相似患者检索; 多模态数据融合; 深度哈希网络

引用格式: 谢明朗,袁贞明,施军平,田昕.基于深度哈希的多模态临床数据相似病例检索.计算机系统应用. http://www.c-s-a.org.cn/1003-3254/9815.html

Similarity Patient Retrieval of Multimodal Clinical Data Based on Deep Hashing

XIE Ming-Lang[1], YUAN Zhen-Ming[1], SHI Jun-Ping[1,2], TIAN Xin[1,2]

[1](Hangzhou Normal University, Hangzhou 311121, China)
[2](The Affiliated Hospital of Hangzhou Normal University, Hangzhou 310015, China)

Abstract: With the widespread adoption of electronic health record (EHR), retrieving similar cases has become a critical task in supporting clinical decision-making, such as in auxiliary diagnosis and treatment planning. However, EHR data is characterized by high dimensionality, heterogeneity, and large volume. To effectively integrate multimodal clinical data and achieve efficient retrieval, this study proposes a multimodal clinical data retrieval model for similar cases based on deep hashing—MCDF. This model employs different methods for feature extraction tailored to the characteristics of various modalities, utilizing multi-layer perceptron (MLP) for structured text data, BioBERT for unstructured text data, and BioMedCLIP for image data, followed by feature fusion through a self-attention mechanism. A triplet loss function guides the model to directly generate hash codes that effectively represent the samples, enabling rapid comparisons for sample retrieval. This not only enhances retrieval accuracy but also significantly improves efficiency. Using the publicly available MIMIC-III dataset, the MCDF model is evaluated against traditional hashing methods (such as spectral hashing) and advanced hashing methods (such as deep hashing network) using mean normalized discounted cumulative gain (*MNDCG*) and mean average precision (*MAP*) metrics for evaluation. Experimental results demonstrate that the MCDF model outperforms all baseline models, validating the superiority of the proposed approach.

① 基金项目: 杭州市生物医药和健康产业发展扶持科技专项 (2022WJC197)
收稿时间: 2024-09-27; 修改时间: 2024-10-23; 采用时间: 2024-11-07; csa 在线出版时间: 2025-03-31

Key words: similar patient retrieval; multimodal data fusion; deep hashing network

在过去的几十年中, 随着电子健康档案 electronic health record, EHR) 的迅速发展, 基于 EHR 的个性化医疗越来越受到人们关注.个性化医疗要求医生综合考虑个体的遗传、环境和生活方式等信息, 为每位患者量身定制个性化的治疗方案.在这一过程中, 相似病例检索发挥了关键作用.相似病例检索的具体过程如图 1 所示, 从复杂的海量病例中精准找到与样本病例最相似的若干病例, 不仅能为临床医生提供有效的参考, 还能优化医疗资源的分配, 实现降本增效.

实现精准且高效的病例检索的前提是确定合适的患者相似性比较方法. 早期病例检索方法一般采用统计方法来学习具有临床意义的相似性度量. Parimbelli 等[1]在综述中调查了 279 篇以患者相似性作为上游任务的文章, 发现大部分研究均采用统计方法实现患者相似性比较 (如聚类分析、主成分分析等). 随着算力的发展, 机器学习[2–4]和深度学习[5–7]因其强大的特征提取能力及良好的可解释性, 逐渐成为患者相似性比较的有效工具. 如文献[2,3]中使用机器学习方法来评估患者之间的相似性, 并在各自的下游任务中取得了显著效果. 而文献[4]利用低秩稀疏投影方法对数据进行降维, 从而消除冗余特征, 提升相似比较的精准度. 深入学习一般用于对样本进行特征提取, 如文献[5–7]均使用深度学习模型进行病例的特征提取, 以期望在病例比较时获得更高的精度.

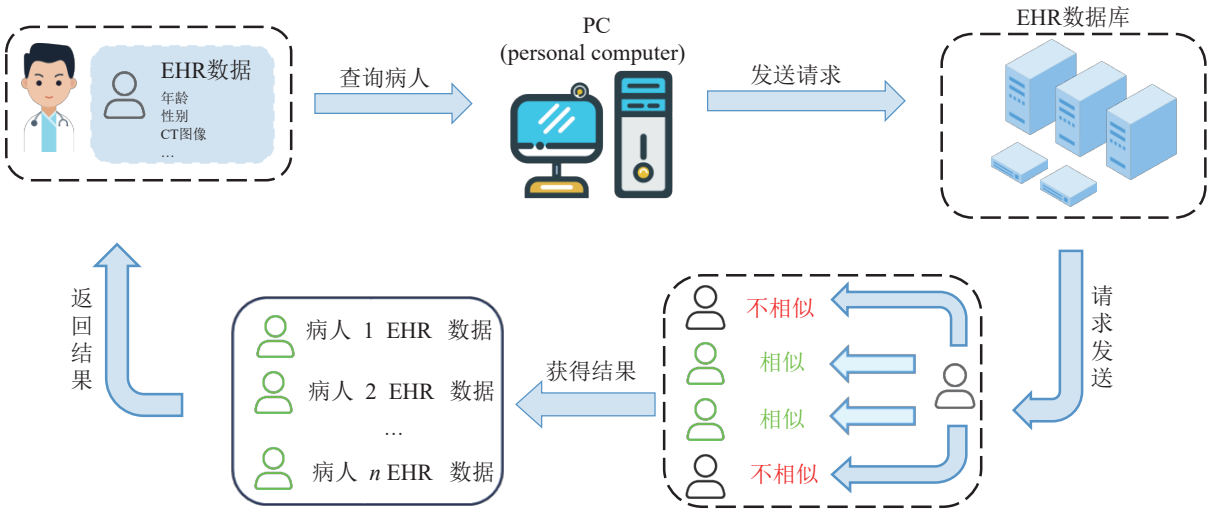


图 1 患者检索图

还有一些研究依赖医生的手工标注或反馈. 例如, 文献[8]利用医生标注的成对标签进行检索, 而文献[9]则基于可视化系统进行检索, 并收集医生在使用过程中的标注. 为了提高模型在特定数据上的预测能力, 文献[10,11]等研究利用了图的方法组织样本进行检索. 这些方法虽然提高了检索精度, 但仅适用于固定数量的患者数据库, 并且未标注的样本也无法参与检索. 随着病例数量的增加和手动标注成本的上升, 这些方法的局限性也逐渐显现.

为了在可扩展样本库中实现高效检索, 一种有效方法是使用哈希编码. 哈希编码将每个样本的特征映射为固定长度的比特码, 通过异或运算计算样本间的距离, 从而实现高效检索. 哈希方法可以显著提升检索效率, 但在生成哈希码的过程中也可能导致信息损失, 甚至引发哈希冲突等问题, 进而影响整个系统的准确性. EHR 中多模态的数据蕴含着丰富信息, 如果能充分利用多模态数据, 即可实现在保证精度的同时显著提升检索效率.

综合以上两个想法, 本文提出了一种基于深度哈希的多模态临床数据相似病例检索模型, 即 MCDF (multimodal clinical data fusion hashing model). 该模型使用不同方法对不同模态的数据进行特征提取, 并通

过三元组损失函数[12]引导模型进行自适应权重的特征融合, 并生成能够有效代表样本的哈希码. 通过哈希码快速比对实现样本检索, 不仅提高了检索的准确性, 还显著提升了检索的效率.

1 相关工作

1.1 患者相似性比较与相似病例检索

随着 EHR 的普及, 患者数据大多以结构化形式存储, 为患者相似性研究奠定了数据基础. 患者相似性比较的核心在于找到合适方法来判断两个患者的相似性. 早期研究中, 这一过程通常采用无监督的方法来实现. Kenney 等[13]使用基于欧几里得距离的相似性度量来判断患者的相似性; Panahiazar 等[14]则采用马氏距离, 通过计算样本点的标准差倍数进行评估. Sun 等[15]提出局部监督度量学习 (LSML), 结合医生反馈评估患者相似性. 余弦相似度方法[16–18]因其简洁高效也被广泛应用于相似病例检索. Zhang 等[19]在分析患者与药物相似度时使用 Jaccard 相似性系数. 有监督的方法则更依赖于临床专家和医生的经验判断. 特别的, 为了降低人工标注的成本, 在一些方法中, 使用患者诊断或药品作为标签, 这些方法属于监督学习中的弱监督类型. 交互式相似标记系统[9]提出了一种基于交互式患者标记和自动模型更新的检索方法, 并设计了辅助医生标记的视觉系统, 支持检索并收集数据以促进后续研究. Wang 等[20]提出动态贝叶斯神经网络 (TDBNN) 来捕获医疗指标间的条件依赖关系, 整合到多变量时间序列分析中以学习细粒度患者相似度. Zhang 等[21]通过统一框架学习局部和全局患者状态, 先测量相似度再预测住院死亡率. 优质的相似性比较方法能有效引导模型学习样本间的区别和联系, 本文对多种相似性方法进行了比较, 并综合考虑了模型及数据集的结构特点, 最终选择了基于患者多标签诊断的广义 Jaccard 系数来评估患者的相似性.

相似病例检索是患者相似性比较方法的常见下游任务. 根据实现方式的不同, 检索方法可以分为 3 类: 基于词向量的检索、基于树模型的检索以及深度学习模型的检索. 早期的检索任务通常基于词向量. Jia 等[22]将输入拆分并用 FastTex 生成词向量, 再与病例库中的词向量进行匹配来实现相似患者检索. 基于树模型的检索具有天然的效率优势, Wang 等[8]提出 ART 检索树, 采用半监督方法构建检索树, 降低计算成本并提升检索效率. 深度学习因强大特征提取能力被广泛应用于病例检索. Wang 等[23]提出带注意力的多任务神经网络深度哈希方法, 结合连续值嵌入和二进制哈希码进行相似患者检索. Gu 等[24]利用图卷积网络 (GCN) 提取患者图标签特征, 与多层感知器 (MLP) 提取的特征融合, 生成带标签信息的哈希码并实现高效检索.

以上模型和方法在相似病例检索任务中已取得一定成果. 然而, 病例样本具有复杂性和异构性的特点, 如果可以从多模态角度出发, 并针对不同结构的数据特异性进行处理, 理论上可能会有更好效果.

1.2 基于多模态数据的病例检索

利用多模态数据进行检索任务通常能够显著提高准确性和泛化能力. 许多医学检索任务已尝试采用多模态数据来训练模型, 以期获得更优的性能表. Quelleca 等[25]提出了一种基于决策树的多模态 (图像和语义信息) 医学检索系统, 评估结果显示其性能显著优于单幅图像检索. 在后期的工作中, 他们结合贝叶斯网络和 Dezert-Smarandache 理论又对检索模型进行进一步优化, 使该系统在糖尿病视网膜病数据库上的检索精度达到了 80.5%, 进一步提升了检索精度. Silva-Leite 等[26]提出了一种用于 CT 图像检索的系统——ChestFinder. 该系统首先对放射学报告和 CT 图像进行特征提取, 利用预训练的编码器生成两份排名列表. 随后, 采用阈值算法 (TA) 对这两份列表进行融合, 从而获得最终的检索结果. Zhang 等[27]设计了一种高效的聚合策略, 通过综合多种相似性度量, 构建了更加精确的亲和矩阵, 以全面捕捉数据之间的关系. 该方法集成了模态特定编码器、图卷积网络 (GCN) 和融合模块, 实现了跨模态的统一二进制编码, 在病例检索任务中取得了不错的效果.

利用多模态临床数据实现病例检索任务可以有效提升检索精度. 然而, 更大的计算与存储开销意味着检索效率的降低, 如果可以使用哈希方法进行优化, 不仅可以通过缓存机制降低存储成本, 还能显著加快检索速度, 进一步提升检索效率.

1.3 哈希方法

在大规模的 EHR 数据中实现高效的相似病例检索, 一种可行的方案是利用哈希方法来加速检索.

早期的哈希方法通常是数据无关的, 其基本思想是通过某种方法比较两个样本之间的距离, 从而实现检索. 其中, 最具代表性的方法是局部敏感哈希 (LSH). LSH 通过一系列随机投影的哈希函数将原始数据映射

到哈希桶中. 在高维空间中, 样本越是接近, 它们被映射到同一个桶中的概率就越高. 然而, 由于这些方法缺乏对数据特性的考虑, 它们在实际应用中往往表现不佳.

学习哈希 (learning-to-hash) 方法能够充分利用数据的内在特性, 因此通常能取得更优的效果. 该方法主要分为两类: 无监督方法和有监督方法. 无监督方法依据数据的结构和分布来学习哈希函数, 适用于没有标签的数据. 典型的无监督方法包括谱哈希 (SH)[28], 通过谱松弛技术高效地计算哈希码; 此外, 还有同质哈希 (IsoHash) 和可扩展图哈希 (SGH) 等方法. 有监督方法则利用监督信息 (如点对点相似性、成对样本之间相似性、三元组或列表级别的标签) 来生成哈希码. 代表性的算法有基于核的监督哈希 (KSH) 和监督离散哈希 (SDH)[29]. 其中, KSH 通过保持成对样本之间的相似性来引导模型学习基于核的哈希函数, 而 SDH 则通过优化线性分类的目标函数来生成高质量的哈希码.

深度哈希 (deep hashing) 方法因其强大的特征提取能力, 已成主流的学习哈希技术. Zhu 等[30]提出的深度哈希网络 (DHN) 通过同时优化特征表示和量化误差, 实现对哈希编码的有效学习. 为缓解连续表示二值化过程中检索质量的下降, Cao 等[31]提出 HashNet, 该方法利用具有收敛性保证的延拓方法几乎精确地学习到了代表样本的哈希码. 虽然这些方法 (如 DSH、DPSH 和 ADSH 等) 利用成对相似性作为监督信息, 但通常仅粗略判定相似性, 无法有效处理多标签实例的细粒度相似性. 因此, Zhang 等[32]提出了改进的深度哈希网络 (IDHN), 通过基于归一化语义标签的百分比量化方法, 提升了多标签实例的检索质量.

近年来, 部分研究者们尝试将深度哈希技术应用于相似患者检索. Wang 等[23]提出了一种深度哈希方法, 该方法采用了多任务神经网络和注意力机制, 通过将网络中获得的连续值嵌入向量与二进制哈希码结合, 实现了从粗到细的相似患者检索. Xu 等[33]则提出了联邦患者哈希框架, 该框架利用联邦深度学习的方式生成患者的二进制哈希码, 从而缓解隐私问题. 然而, 这些方法未考虑多模态数据中所蕴含的丰富信息, 相似度判断方法也比较粗糙.

2 基于深度哈希的多模态临床数据相似病例检索模型

本节将详细介绍基于深度哈希的多模态临床数据相似病例检索模型 MCDF. MCDF 模型由 3 个输入模块、1 个融合模块和 1 个输出模块组成, 具体结构如图 2 所示.

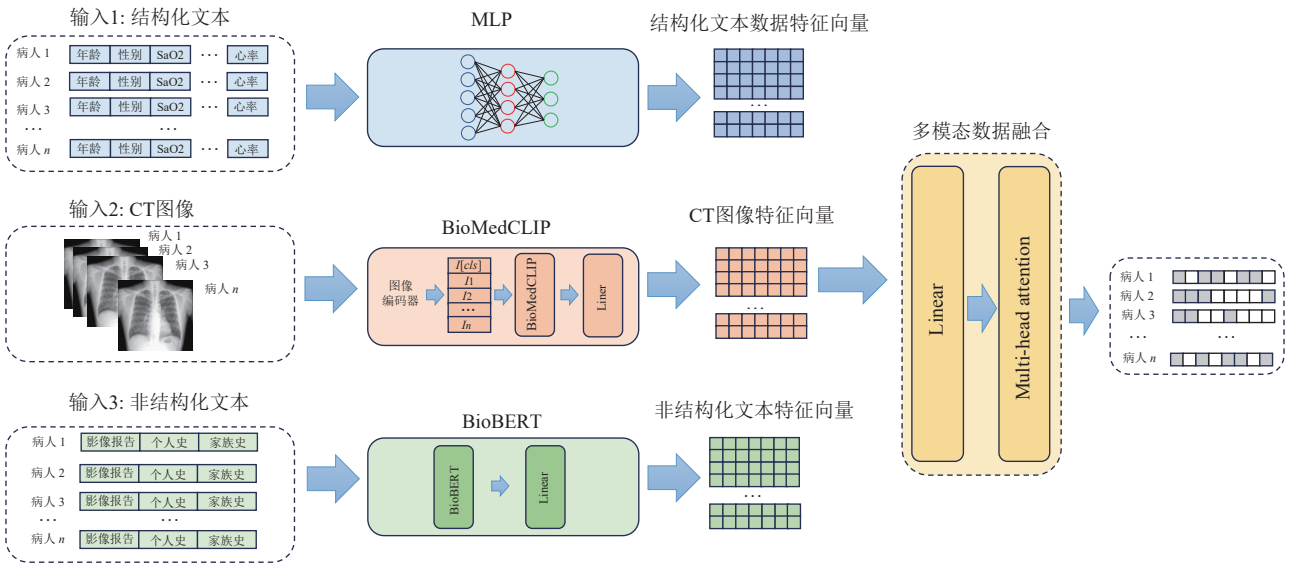


图 2 MCDF 模型结构图

2.1 相关符号

本研究将患者数据记作$X=\{x_i\}_{i=1}^{N}\in\mathbb{R}^N$, 其中 N 为患者数量, 而x_i表示第 i 位患者的数据特征. 考虑到 EHR 数据的结构差异, 每位患者的特征x被表示为$x=\{x^k\}_{k=1}^{M}$, 其中, M 代表数据的不同部分, 而x^k代表某位患者的第 k 部分数据. 本文基于患者的诊断来评估患者之间

的相似性, 而诊断是多标签的, 将其表示为: $L=\{l_i\}_{i=1}^{N}\in\{0,1\}^{N\times C}$, 其中$l_i=[l_{i1},l_{i2},\cdots,l_{ic}]$表示第 i 位患者的标签, C 是不同诊断的总数. 如果第 i 位患者患有第 j 种疾病, 则$l_{ij}=1$, 否则$l_{ij}=0$. 为充分利用多标签数据中所蕴含的丰富信息, 采用广义 Jaccard 系数 (generalized Jaccard coefficient) 来衡量患者之间的相似性, 公式为:

$$J(l_i,l_j)=\frac{\langle l_i,l_j\rangle}{\langle l_i,l_i\rangle+\langle l_j,l_j\rangle-\langle l_i,l_j\rangle} \tag{1}$$

其中, $\langle\cdot,\cdot\rangle$表示内积. 根据式 (1) 可以得到一个用来衡量患者x_i与患者x_j之间的相似性的值$J(l_i,l_j)$. $J(l_i,l_j)$越接近 1 表示两个样本之间的相似度越高, 越接近 0 则相似度越低. 如果是判断患者是否相似, 可以把它简化为二值问题. 本文用S_{ij}来表示患者x_i和x_j之间是否相似, 则S_{ij}可以描述为式 (2), 其中β为阈值.

$$S_{ij}=\begin{cases}1, & J(l_i,l_j)>\beta\\ 0, & J(l_i,l_j)\leqslant\beta\end{cases} \tag{2}$$

2.2 特征提取及特征融合方法

本文希望通过全面整合各模态的信息来更加精确地提取患者特征, 并实现高效检索. 然而, 每位患者的 EHR 数据不仅是多模态的, 而且同一模态的数据也可能存在结构性差异. 因此, 本文对不同模态的数据采取针对性的处理方法, 并利用自适应权重模型对各模态特征进行融合.

2.2.1 文本数据

文本数据包括结构化数据和非结构化数据, 其中结构化数据通常指数值型数据 (如布尔型、二元型、整数型、浮点型等), 而非结构化数据则主要指长段医学文本. 为了更有效地从数据中提取特征, 本文分开处理结构化和非结构化文本数据.

(1) 结构化文本数据

对于数值型、类别型、二元型等结构化医学文本数据, 本文将每个医学指标视为一个独立的维度. 这些数据缺乏明确的上下文依赖和空间含义, 因此, 本文放弃了擅长捕捉数据空间关系的网络 (如卷积神经网络), 而选择了传统的多层感知机 (multi-layer perceptron, MLP) 模型进行特征提取. 根据文献[5]中的建议, 将网络层数设定为 3 层, 并且输入层的神经元数量由数值特征的数量 O 决定. 为了在保证特征提取效果的同时避免过多的参数, 将第 2 层神经元的数量设置为 $2O$ 个. 输出层包含 N 个神经元, 以确保与其他特征提取模块的输出向量维度一致. 采用修正线性单元 (rectified linear unit, ReLU) 作为模型的激活函数来引入了非线性因素, 从而增强模型的学习能力和表达能力. 结构化文本的特征提取过程可以表示为:

$$H_1=E_{P_1}(X^1;\Theta_{P_1}) \tag{3}$$

其中, E_{P_1}是对结构化文本数据进行特征提取, X^1表示使用第 1 部分样本数据, Θ_{P_1}表示E_{P_1}中的可训练参数, $H_1=\{h_i\}_{i=1}^{N}$表示对样本结构化文本数据的特征提取结果.

(2) 非结构化文本数据

结构化文本数据通常为长段的医学文本数据, 数据内部存在着一定的序列性, 蕴含着丰富的上下文信息, 同时有着结构复杂, 噪声较多的特点. 这些数据通常为长段落, 具有明显的序列性和丰富的上下文信息, 内部结构复杂, 信息量大, 同时可能包含较多噪音.

为更准确地提取非结构化文本数据中的信息, 本文选择了经过生物医学数据集预训练的 BioBERT[34] 模型来提取特征. BioBERT 基于 Transformer 架构, 在自然语言处理任务中表现优异, 经过医学领域的专门预训练后, 能在生物医学任务中展现出卓越的性能. 根据文献[13]的建议, 本文将最大输入长度设置为 512 个 token, 以确保模型能够有效处理上下文之间的关系. 为了减少过拟合并提高模型训练效率, 在训练时冻结了模型参数, 并在输出层后附加了一个包含 N 个神经元的全连接层, 与结构化数据一致, 采用 ReLU 激活函数. 类似地, 这一特征提取过程可以表达为:

$$H_2=E_{P_2}(X^2;\Theta_{P_2}) \tag{4}$$

其中, E_{P_2}是对非结构化文本数据进行特征提取, X^2表示使用第 2 部分的样本数据, Θ_{P_2}表示E_{P_2}中的可训练参数, $H_2=\{h_i\}_{i=1}^{N}$表示对样本非结构化文本数据的特征提取结果.

2.2.2 图像数据

医学图像能够以较高的分辨率呈现身体内部结构, 包括骨骼、软组织和血管系统, 帮助医生有效的识别细微的病变和异常, 在疾病检测和临床诊断中发挥着重要作用. BioMedCLIP[35]在多个标准生物医学图像任务中表现出色, 尤其在医学图像信息提取方面性能卓越. 因此, 选择 BioMedCLIP 模型提取医学图像特征, 使用模型提供的图像预处理工具将图像统一处理成 224×224 的尺寸. 为了更好地满足图像特征提取的需

求, 模型去除了文本编码器, 仅保留图像编码器作为模型输入. 与 BioBERT 类似, 冻结了模型的参数, 并在模型的输出层后增加了一个包含 N 个神经元的全连接层, 为了进一步增强模型的表达能力, 使用 ReLU 作为激活函数以增加其非线性因素. 与之前类似, 这一特征提取过程可以表示为:

$$H_3 = E_{P_3}(X^3;\Theta_{P_3}) \tag{5}$$

其中, E_{P_3} 是对图像数据进行特征提取, X^3 表示使用第 3 部分的样本数据, Θ_{P_3} 表示 E_{P_3} 中的可训练参数, $H_3 = \{h_i\}_{i=1}^N$ 表示对样本图像数据的特征提取结果.

2.2.3 特征融合

在对多模态数据进行特征提取后, 采用适当的融合方法通常能够显著提升性能. 传统的数据融合方法 (如简单拼接、加权平均等) 可能在特征表达能力、灵活性和可解释性方面有所不足, 难以充分挖掘多模态数据的潜在信息. 此外, 这些传统方法通常无法有效处理不同模态间的复杂交互和长距离依赖, 限制了其在复杂任务中的表现. 为了更好捕捉不同模态数据之间的关系, 本文在选取多头注意力机制模块进行特征融合. 多头注意力机制使模型能够在不同的注意力空间中学习和理解数据, 从而显著增强特征表示能力. 在第 2.2.1 及 2.2.2 节中, 选择合适的方法对不同模态的数据进行了特征提取, 并通过修改其输出层的尺寸使模型输出的向量维度均是 $1\times N$. 为了消除数据特征在尺度和分布上的差异, 本文对 3 个特征提取模块的输出进行了归一化处理. 为了实现更有效的特征融合, 首先使用具有 N 个神经元的全连接层进行初步融合, 然后再通过一个 8 头注意力机制模块[36]进行进一步融合. 由于希望引导模型直接输出能够精确代表样本的哈希码, 本文将注意力机制模块的输出层修改为 M 个神经元的全连接层, 并使用双曲正切函数 (hyperbolic tangent function, tanh) 作为激活函数将模型的输出映射到 (−1, 1), 这个过程可以被描述为:

$$H = E_P(H_1,H_2,H_3;\Theta_{P_4}) \tag{6}$$

其中, E_P 是对多模态数据提取的特征进行融合, Θ_{P_4} 表示 E_P 中的可训练参数, $H = \{h_i\}_{i=1}^N$ 表示多模态数据融合后的特征.

根据文献[37]中的方法, 我们利用二值化函数将融合向量 H 转化为哈希码 B:

$$b_{ij} = \mathrm{sgn}(h_{ij}) = \begin{cases} 1, & h_{ij} > 0 \\ -1, & h_{ij} \leqslant 0 \end{cases} \tag{7}$$

其中, h_{ij} 表示第 i 个病人的融合特征向量中第 j 位的值. sgn(·)是符号函数, 利用式 (7), 我们将融合向量 H 中的每一位 h_{ij} 转换为该位所对应的哈希码 b_{ij}, 完成转换后, 我们得到了用于相似病例检索的哈希码 B.

2.3 目标函数及训练策略

为了让模型尽可能精准地学习不同患者之间的区别与联系, 从而更有效地生成能够充分代表患者样本的哈希码, 以实现高效检索. 三元组损失函数是监督学习中用于学习嵌入空间的损失函数. 它要求模型建立样本之间的正确关系, 使得相似样本在嵌入空间中的距离更近, 而不相似样本的距离更远. 其输入是一个由锚样本 x_i, 正样本 x_p, 负样本 x_n 组成的三元组, 其中 x_i 与 x_p 相似, x_i 与 x_n 不相似. 利用式 (2) 作为判断标准并遵循文献[12]的意见在训练时为每个 x_i 动态的构建三元组, 并希望:

$$d(x_i,x_p) + \alpha < d(x_i,x_n) \tag{8}$$

其中, $d(x_i,x_p)$ 表示样本 x_i 和样本 x_p 之间的距离, α 为边距超参数, 用于在正样本对和负样本对之间施加最小距离约束. 所以, 最终要优化的 Loss 可以表示为式 (9):

$$L = \sum_{i=1}^{N}\left[\max\left(d\left(x_i,x_p\right) - d\left(x_i,x_n\right) + \alpha, 0\right)\right] \tag{9}$$

如果锚样本与正负样本之间距离差大于 α, 就使用最大值函数将其 Loss 值设为 0. 式 (9) 是可以微分的, 所有的模型参数都可以用标准的反向传播算法和梯度下降算法进行优化, 且 3 个编码器可以实现端到端联合训练. MCDF 训练过程的伪代码如算法 1 所示.

算法 1. MCDF 模型训练算法

输入: 病例数据 X (结构化文本数据 X^1, 非结构化文本数据 X^2, CT 图像数据 X^3).
输出: 融合特征向量 H.

1. 初始化: 编码参数 Θ_P; 超参数 α; 批大小 N_b; 学习率 μ; 相似性参数 β; 目前迭代次数 $iter = 1$;
2. 最大迭代次数 $T_{iter} = \left\lceil \frac{N}{N_b} \right\rceil$; 目前轮次编号 $epoch = 1$.
3. **repeat**
4. **for** $iter = 1, 2, \cdots, T_{iter}$ **do**
5. 从批数据里中选择一个病例 x_i;
6. 通过式 (3)–式 (6) 获得融合特征向量 H;
7. 通过式 (9) 计算 Loss;
8. 使用反向传播算法和梯度下降算法更新 Θ_P;

9. **end for**
10. **until** 收敛;
11. 通过式 (7) 获得哈希码 B;
12. 返回 B 和Θ_P;

病例数据 X, 包括结构化文本数据 X^1, 非结构化文本数据 X^2, CT 图像数据 X^3, 利用式 (3)–式 (5) 进行特征提取并通过式 (6) 进行特征融合, 接着通过式 (9) 计算 Loss, 并使用反向传播算法与梯度下降算法来更新模型参数, 直至收敛.最终, 利用式 (7) 获得哈希码 B 并保存模型Θ_P.

3 实验分析

3.1 数据集

在本节中, 首先介绍实验评估所用数据集, 然后介绍实验设置 (包括消融实验及对比试验) 和评价指标, 最后对实验结果进行分析和展示. MIMIC-III[38]是一个公开数据库, 存储着来自贝斯以色列女执事医疗中心 (Beth Israel deaconess medical center) 重症监护病房 (ICU) 患者的 EHR 数据. 通过多个关键指标对 MIMIC-III 数据库进行联表查询, 并剔除了部分缺失数据, 最终保留了 592 例 EHR 数据作为实验样本. 由于 MIMIC-III 数据集具有时序性, 为了消除时间对样本数据的影响, 选择单次急诊的数据作为一个样本, 以确保各模态数据的时间差不超过 24 h. EHR 样本中包含文本数据和图像数据, 在数据预处理阶段, 依据临床医生和专家的建议进行处理. 对于结构化文本数据进行以下处理: 布尔值被映射为 1 和 0, 其余数值型数据 (包括整数和浮点数) 保留至小数点后两位, 并对其进行归一化处理, 以消除不同量纲对模型的影响. 结构化文本数据处理时, 首先将长段文本拼接, 去除特殊字符和停用词, 将所有字母转换为小写, 最后使用 BioBERT 的预训练分词器将其转化为 token. CT 图像数据则利用 BioMedCLIP 提供的工具将其尺寸调整为 224×224, 以适配模型的输入要求. 数据集共有 19 种诊断标签, 利用独热编码方式对这些诊断进行编码, 得到 19 维标签向量. 在完成以上预处理之后, 单个样本的多模态特征及标签维度大小如表 1 所示.

本文将数据集按照 8:1:1 的比例随机分为训练集、验证集和测试集, 并将训练集和验证集中的样本合并作为检索的病例.

表 1 数据集详细信息

数据集	实例	训练	验证	测试 (查询)	检索	结构化文本	非结构化文本	CT 图像	多标签
MIMIC-III	592	474	59	59	533	11	256	224×224	19

3.2 实验准备工作

3.2.1 评价指标

在评估相似患者检索的质量时, 本文采用两个在检索任务中广泛使用的评价指标: 归一化折扣累计收益均值 (mean normalized discounted cumulative gain, *MNDCG*)[39]及均值平均精度 (mean average precision, *MAP*)[40].

MNDCG 是一种有效评估多标签数据检索的指标. 对于给定查询 q, 前 n 个检索的实例的 *DCG* 分数定义为:

$$DCG@n=\sum_{i=1}^{n}\frac{2^{Rs(q,i)}-1}{\log(1+i)} \tag{10}$$

其中, $Rs(q,i)$表示查询 q 与实例 i 之间的相关性评分, 使用式 (1) 来实现. 为了更准确地评估检索效果, 使用理想 *DCG* (ideal discounted cumulative gain, *IDCG*) 作为归一化因子对 *DCG* 进行归一化处理, 从而得到归一化 *DCG* (*NDCG*). 通过计算测试集上所有样本的 *NDCG* 的平均值, 最终得到 *MNDCG*. 这个过程可以描述为:

$$MNDCG@n=\frac{1}{Q}\sum_{q=1}^{Q}\frac{DCG@n}{IDCG@n} \tag{11}$$

其中, Q是测试集样本的数量. $MNDCG@n$的取值范围为[−1, 1], 值越高则模型表现越佳, 反之则表现较差.

均值平均精度 (mean average precision, *MAP*) 是一种广泛用于评估检索方法性能的指标, 它表示平均精度 (average precision, *AP*) 的均值. 对于给定查询 q, 计算前 n 个检索实例的 *AP* 分数的公式如下:

$$AP(q)@n=\frac{1}{N_{Rd}(q)@n}\sum_{i=1}^{n}\left(Rd(q,i)\frac{N_{Rd}(q)@i}{i}\right) \tag{12}$$

其中, $Rd(q,i)\in\{0,1\}$是一个相关性判断函数. 本文遵循先前研究惯例, 采用宽松标准来确定$Rd(q,i)$. 如果实例i和查询q共享至少一个诊断结果, 或两者均未受该疾病影响, 就认为这两个病例是相关的, 并赋值$Rd(q,i)=1$; 否则$Rd(q,i)=0$. $N_{Rd}(q)@i$表示前i个检索结果中, 与查询q相关的实例数量. 类似于 *NDCG*, 计算平均分数如式 (13) 所示, 较高的 *MAP* 分数表示模型性能优越.

$$MAP = \frac{1}{Q}\sum_{q=1}^{Q} AP(q) \tag{13}$$

3.2.2 评价指标

本文将模型 MCDF 与几种基线方法进行了比较, 其中包括一些传统的哈希方法和当前领先的深度哈希方法, 下面简要介绍这些方法及其相应的参数.

SH (spectral hashing)[28]是一种基于谱图理论和图论技术的无监督哈希方法. 它将原始数据构建为特定形式的图, 并运用谱松弛技术将高维数据映射为低维的二进制哈希码, 从而实现高效的信息检索.

KSH (kernel supervised hashing)[41]是一种基于核函数的监督哈希方法, 它利用哈希码之间的内积与汉明距离之间的等价性来引导函数生成准确的哈希码. 在构建 KSH 函数时, 使用了高斯径向基函数 (Gaussian radial basis function, RBF) 作为核函数.

SDH (supervised discrete hashing)[29]是一种离散哈希方法, 其目标是生成适用于线性分类的有效哈希码. SDH 将输入数据映射到紧凑且高效的二进制哈希码空间中, 依靠近似最近邻搜索实现高效检索. 在实现该方法时, 设定正则化参数λ=1, 惩罚参数$\beta = 1\times10^{-5}$以及最大迭代次数t=5.

HN (HashNet)[31]是一个著名的深度哈希框架, 它通过连续放松技术有效从不平衡的相似性数据中学习精确的哈希码, 且保证了模型的收敛性. 在实现过程中, 将调节自适应 Sigmoid 函数带宽的超参数α设置为 0.1/K, 其中 K 表示哈希码长度.

DHN (deep hashing network)[30]是一个深度哈希框架, 它能够精确学习适合哈希编码的特征表示, 并有效管理量化误差. 实验中将量化惩罚参数α设置为 0.001.

DPSH (deep pairwise supervised hashing)[42]是一种用于图像检索和大规模数据查询的深度学习哈希方法. 它通过最大化成对标签的似然来学习样本的特征表示和哈希码生成. 实验中, 将正则化参数α设置为 0.1.

IDHN (instance-dependent hashing network)[32]是一种专为多标签实例检索设计的深度哈希框架. 该框架区分了“硬相似度”和“软相似度”. 其中“硬相似度”用于交叉熵损失函数“软相似度”用于均方误差损失函数. 在模型实现时, 约束带宽α设置为 5/K, 均方误差损失系数β设置为 0.1/K, 量化损失系数 n 设置为 0.1, 其中 K 表示哈希码长度.

考虑到实验中这些基线方法在训练时所用数据集均是基于文本的, 为公平起见, 将多模态数据用 MCDF 主干模型进行特征提取并拼接, 得到的特征向量作为其他竞争对手的输入. 在 MCDF 模型中, 将批量的大小为 32, 学习率采用 1E–4, 权重衰减为 5E–4 的均方根传播优化器. 根据验证集结果, 将式 (9) 中的超参数设置为 0.6. 为了避免过拟合, 所有方法均使用了“早停策略”, 其超参数 *patience* 设置为 25, 并且训练次数不超过 400 次. 所提模型的实现基于 PyTorch 框架. 系统在 Ubuntu 18.04 LTS 操作系统上实现, 使用 Intel(R) Xeon(R) CPU E5-2650 v4@2.20 GHz 处理器和 128 GB 内存, 配备 NVIDIA Corporation GV100 GPU. 所有实验均在该系统上进行.

3.3 实验结果

为了更全面地评估模型 MCDF 相对于其他方法或模型在 MIMIC-III 数据集上的性能差异, 本文设计了 4 种不同长度的哈希码进行实验, 分别为 32、64、128 和 256 位. 在实验中, 使用检索结果的前 10 个病例来计算 *NDCG*@10 和 *AP*@10, 然后对测试集中所有实例的 *NDCG*@10 和 *AP*@10 结果进行求和并取平均, 最终得到 *MNDCG*@10 和 *MAP*@10 指标. 为了验证实验结果的统计学显著性, 采用 Wilcoxon 秩和检验, 增强实验结果的可靠性和说服力. 为了确保公平性, 考虑到其他方法或模型中使用的两两相似度量方法各不相同且部分方法较为粗糙, 决定统一采用本文中提出的两两相似度方法进行实验. 特别的, 鉴于 IDHN 方法的独特性, 不会对其相似度评判方法进行调整.

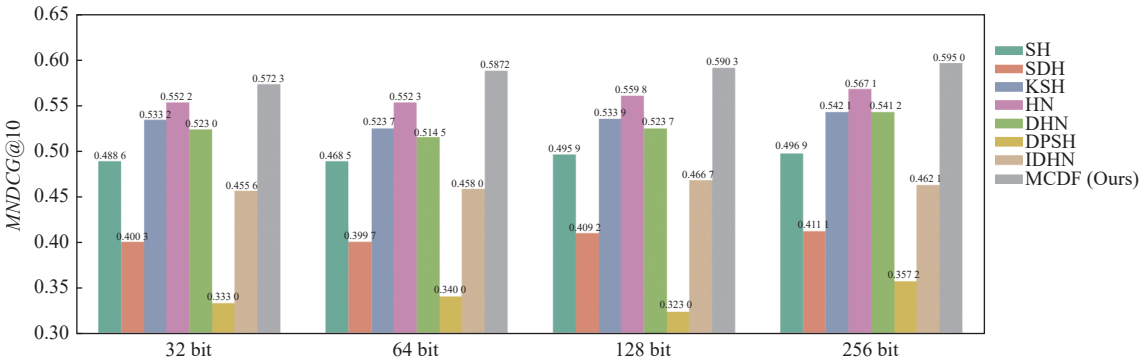
3.3.1 MCDF 在 *MNDCG* 指标中的性能比较

表 2 与图 3 展示的 *MNDCG*@10 指标实验结果表明, 本文提出的模型 MCDF 在相同码长下相较于其他最优秀指标分别提高了 3.64%、6.34%、5.48% 和 4.92%. 与深度哈希方法相比, 传统哈希方法的性能整体较弱, 这主要归因于深度学习模型在特征提取方面的强大能力. 需要注意的是, DPSH 的效果偏差可能是由于相似性评估算法的变化.

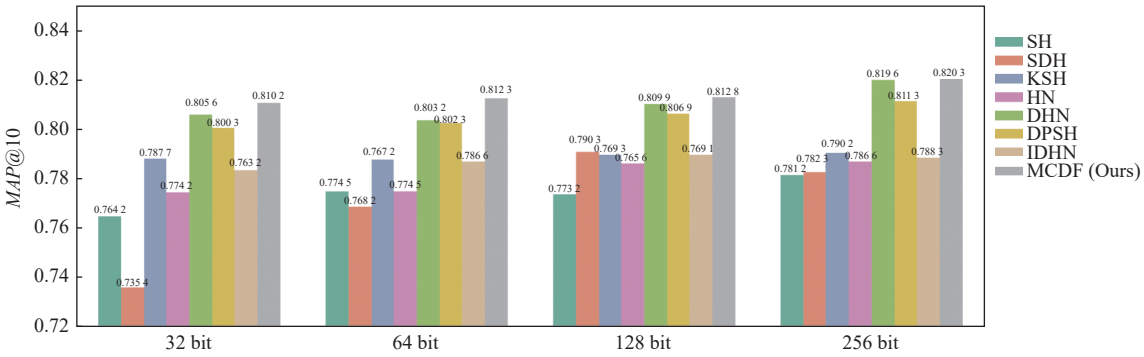
为了评估结果的统计学意义, 本文使用 Wilcoxon 秩和检验对各竞争对手在测试集上的 *NDCG*@10 值进行比较, 并以 MCDF 作为控制方法. 当 p 值小于 0.05 时, 认为两种方法的实验结果存在显著差异. 为了更清晰地展示实验结果, 当 p 值小于 1.08E–5 时, 统一将其计为 1.08E–5.

表 2 在 MIMIC-III 数据集上 *MNDCG*@10 与 *MAP*@10 实验结果展示

评价指标	*MNDCG*@10				*MAP*@10			
	32 bit	64 bit	128 bit	256 bit	32 bit	64 bit	128 bit	256 bit
SH	0.4886	0.4885	0.4959	0.4969	0.7642	0.7745	0.7732	0.7812
SDH	0.4003	0.3997	0.4092	0.4111	0.7354	0.7682	0.7903	0.7823
KSH	0.5332	0.5237	0.5339	0.5421	0.7877	0.7872	0.7893	0.7902
HN	0.5522	0.5523	0.5598	0.5671	0.7742	0.7745	0.7856	0.7866
DHN	0.5230	0.5145	0.5237	0.5412	0.8056	0.8032	0.8099	0.8196
DPSH	0.3330	0.3400	0.3230	0.3572	0.8003	0.8023	0.8059	0.8113
IDHN	0.4556	0.5480	0.4667	0.4621	0.7832	0.7866	0.7891	0.7883
MCDF (Ours)	0.5723	0.5872	0.5903	0.5950	0.8102	0.8123	0.8128	0.8203



(a) 不同方法在不同bit下的*MNDCG*@10得分



(b) 不同方法在不同bit下的*MAP*@10得分

图 3 *MAP*@10 和 *MNDCG*@10 实验结果柱状图

不同码长下的 p 值见表 3, 观察到所有 p 值均小于 0.05. 结合 *NDCG*@10 差异, 可以得出结论: MCDF 在性能上优于其他竞争对手.

3.3.2 MCDF 在 *MAP* 指标中的性能比较

表 2 与图 3 展示了关于 *MAP*@10 的实验结果. 本文的 MCDF 方法在所有码长下均表现出提升, 特别是 256 位时, *MAP* 值达到了 0.8203. 相较于 *MNDCG*@10 指标提升 4%–5%, MCDF 模型在 *MAP*@10 指标上平均约提升 1%, 这可能是因为式 (12) 中的相似度判断仅限于“相似”和“不相似”两种选项, 忽略了许多标签信息, 从而限制了 MCDF 模型的 *MAP* 指标提升. 同样的, 本文也使用 Wilcoxon 秩和检验对各竞争对手在测试集上的 *AP*@10 进行了比较, 并以 MCDF 作为控制方法. 结果如表 3 所示, p 值均小于 0.05, 进一步验证了 MCDF 模型的优越性.

4 讨论

在本节中, 首先通过消融实验验证本文所设计模型的有效性, 随后分析式 (9) 中的超参数对 MCDF 的影响.

4.1 消融实验

为了验证 MCDF 模型中各模块的有效性, 本文设置了消融实验. 模型性能的提升主要归因于两个关键

因素: 一是充分利用了 EHR 中的多模态数据, 并选择了合适的方法进行特征提取. 二是利用多头注意力机制进行特征融合, 显著增强了模型对多模态信息的整合能力和表达能力.

表 3 在实验结果 *MAP* 和 *NDCG* 上进行 Wilcoxon 秩和检验的 p 值结果展示 (以 MCDF 为控制方法)

评价指标	*MNDCG*@10				*MAP*@10			
	32 bit	64 bit	128 bit	256 bit	32 bit	64 bit	128 bit	256 bit
SH	1.08E–5	1.08E–5	1.08E–5	1.08E–5	1.08E–5	1.08E–5	1.08E–5	1.08E–5
SDH	1.08E–5	1.08E–5	1.08E–5	1.08E–5	1.08E–5	1.08E–5	1.08E–5	1.08E–5
KSH	1.08E–5	1.08E–5	1.08E–5	1.08E–5	1.08E–5	1.08E–5	1.08E–5	1.08E–5
HN	1.08E–5	2.16E–5	4.38E–5	1.08E–5	1.08E–5	1.08E–5	1.30E–5	1.08E–5
DHN	1.08E–5	1.08E–5	1.08E–5	5.79E–5	1.08E–5	1.08E–5	1.90E–5	1.08E–5
DPSH	1.08E–5	1.08E–5	1.08E–5	1.08E–5	1.08E–5	1.08E–5	1.08E–5	1.08E–5
IDHN	1.08E–5	1.08E–5	1.08E–5	1.08E–5	1.73E–5	1.50E–5	1.33E–5	1.60E–5

本文将两个变体 MCDF-I 与 MCDF-II 与基准模型 MCDF 进行了比较. MCDF-I 删除了 CT 图像处理模块 (BioMedCLIP), 其他模块保持不变, 以验证多模态数据的有效性. MCDF-II 则移除了多头注意力机制模块, 并用一个新的全连接层代替, 以验证通过多头注意力机制进行特征融合的有效性.

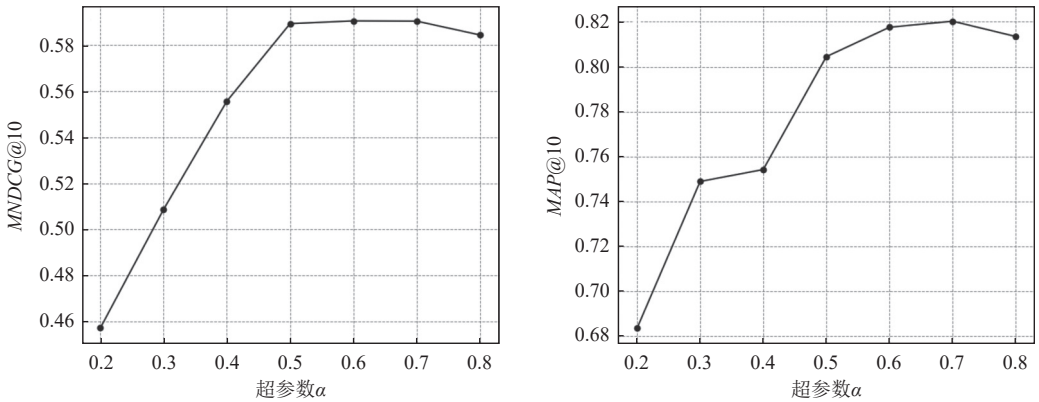
实验结果如表 4 所示. 与完整模型相比, MCDF-I 和 MCDF-II 在不同程度上均表现出性能下降. 值得注意的是, MCDF-I 在 3 个模型中的表现最差, 在 *MAP*@10 指标上, 相比于不同比特长度的基准模型, 平均下降约 16%. 这充分说明了多模态特征对哈希码生成的重要作用. 在对 MCDF-II 的评估中发现, 使用普通全连接层进行特征融合导致模型性能有所下降. 这进一步验证了多头注意力机制模块在特征融合中的有效性.

4.2 超参数

式 (9) 中的α是训练模型损失函数中的超参数, 其表示锚样本与正负样本之间距离之差的最小值. 选择合适的α不仅可以加快模型收敛的速度, 防止过拟合, 还可以进一步提高模型的性能. 本文仅对性能最优的模型 (MCDF-256 bit) 进行了绘图, 因为在其他比特数下观察到的趋势相同. 根据文献[12]中的经验, 将α的初始值设置为 0.2, 并以 0.1 为步长绘制了α与各评价指标之间的变化图 (图 4). 从图 4 中可以清晰地观察到, α在 0.6 之前增长迅速, 0.7 以后开始下降, 因此将α设置为 0.6.

表 4 在 MIMIC-III 上的消融实验结果展示

评价指标	方法	32 bit	64 bit	128 bit	256 bit
MNDCG@10	MCDF-I	0.4905	0.5073	0.5146	0.5178
	MCDF-II	0.5662	0.5688	0.5724	0.5835
	MCDF	**0.5723**	**0.5872**	**0.5903**	**0.5905**
MAP@10	MCDF-I	0.6754	0.6687	0.6764	0.6814
	MCDF-II	0.8110	0.8040	0.7911	0.8011
	MCDF	**0.8125**	**0.8103**	**0.8128**	**0.8203**



(a) 超参数α对*MNDCG*@10的影响 (b) 超参数α对*MAP*@10的影响

图 4 超参数α与评价指标 *MNDCG*@10 和 *MAP*@10 的折线统计图

5 结论与展望

本文提出了 MCDF 模型, 用于相似病例的检索. MCDF 由 3 个不同的输入模块和 1 个融合模块组成. 该模型根据不同模态数据的特性, 将结构化文本数

据、非结构化文本数据和图像数据分别使用多层感知机 (multi-layer perceptron, MLP) 模型、BioBERT、BioMedCLIP 进行特征提取, 并通过自注意力机制模块进行特征融合. 最终, 利用三元组损失函数引导模型直接生成能够有效代表样本的哈希码. 利用样本生成的哈希码进行检索, 不仅提高了检索的准确性, 还显著提升了检索效率. 在公开数据集 MIMIC-III 上进行了广泛的对比实验和消融实验, 以评估 MCDF 模型在相似病例检索任务中的效果, 实验结果验证了 MCDF 的优越性. 本文工作也存在一些局限性. 尽管 EHR 中包含大量多模态数据 (如基因检测报告、病理检查、CT 图像等), 由于数据集大小和模型规模的限制, 仅使用了文本数据和图像数据. 此外, 本文使用的仅是单次急诊的数据, 而未考虑数据中蕴含的时序性因素. 对比实验模型均选取哈希相关的方法, 未考虑与其他领域先进多模态检索模型的比较. 在未来的工作中, 将充分利用电子健康记录 (EHR) 中的多模态数据, 并综合考虑样本数据的时序性因素. 鉴于许多大型语言模型在不同领域中展现了出色的适应能力和卓越的效果, 还计划利用这些模型进行微调实验, 以实现更高效的检索.

参考文献

1 Parimbelli E, Marini S, Sacchi L, *et al*. Patient similarity for precision medicine: A systematic review. Journal of Biomedical Informatics, 2018, 83: 87–96. [doi: 10.1016/j.jbi.2018.06.001]

2 Gottlieb A, Stein GY, Ruppin E, *et al*. A method for inferring medical diagnoses from patient similarities. BMC Medicine, 2013, 11: 194. [doi: 10.1186/1741-7015-11-194]

3 Masud MM, Hayawi K, Mathew SS, *et al*. Effective patient similarity computation for clinical decision support using time series and static data. Proceedings of the 2020 Australasian Computer Science Week Multiconference. Melbourne: ACM, 2020. 33.

4 Zhan MT, Cao SL, Qian BY, *et al*. Low-rank sparse feature selection for patient similarity learning. Proceedings of the 16th International Conference on Data Mining. Barcelona: IEEE, 2016. 1335–1340.

5 Ni JZ, Liu J, Zhang CX, *et al*. Fine-grained patient similarity measuring using deep metric learning. Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. Singapore: ACM, 2017. 1189–1198.

6 Suo QL, Ma FL, YuanY, *et al*. Deep patient similarity learning for personalized healthcare. IEEE Transactions on NanoBioScience, 2018, 17(3): 219–227. [doi: 10.1109/TNB.2018.2837622]

7 Zhu ZH, Yin CC, Qian BY, *et al*. Measuring patient similarities via a deep architecture with medical concept embedding. Proceedings of the 16th International Conference on Data Mining. Barcelona: IEEE, 2016. 749–758.

8 Wang F. Adaptive semi-supervised recursive tree partitioning: The art towards large scale patient indexing in personalized healthcare. Journal of Biomedical Informatics, 2015, 55: 41–54. [doi: 10.1016/j.jbi.2015.01.009]

9 Liu H, Dai HR, Chen JT, *et al*. Interactive similar patient retrieval for visual summary of patient outcomes. Journal of Visualization, 2023, 26(3): 577–592. [doi: 10.1007/s12650-022-00898-9]

10 Gu YF, Yang XB, Tian L, *et al*. Structure-aware siamese graph neural networks for encounter-level patient similarity learning. Journal of Biomedical Informatics, 2022, 127: 104027. [doi: 10.1016/j.jbi.2022.104027]

11 Tashkandi A, Wiese I, Wiese L. Efficient in-database patient similarity analysis for personalized medical decision support systems. Big Data Research, 2018, 13: 52–64. [doi: 10.1016/j.bdr.2018.05.001]

12 Schroff F, Kalenichenko D, Philbin J. FaceNet: A unified embedding for face recognition and clustering. Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston: IEEE, 2015. 815–823.

13 Ng K, Sun JM, Hu JY, *et al*. Personalized predictive modeling and risk factor identification using patient similarity. Proceedings of the 2015 AMIA Joint Summits on Translational Science. San Francisco: AMIA Joint Summits on Translational Science, 2015. 132–136.

14 Panahiazar M, Taslimitehrani V, Pereira NL, *et al*. Using ehrs for heart failure therapy recommendation using multidimensional patient similarity analytics. Studies in Health Technology and Informatics, 2015, 210: 369–373.

15 Sun JM, Wang F, Hu JY, *et al*. Supervised patient similarity measure of heterogeneous patient records. ACM SIGKDD Explorations Newsletter, 2012, 14(1): 16–24. [doi: 10.1145/2408736.2408740]

16 Lee J. Personalized mortality prediction for the critically ill using a patient similarity metric and bagging. Proceedings of the 2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI). Las Vegas: IEEE, 2016. 332–335.

17 Lee J, Maslove DM, Dubin JA. Personalized mortality prediction driven by electronic medical data and a patient similarity metric. PLoS One, 2015, 10(6): e0127428.

18 Li L, Cheng WY, Glicksberg BS, *et al*. Identification of type 2 diabetes subgroups through topological analysis of patient similarity. Science Translational Medicine, 2015, 7(311): 311ra174.

19 Zhang P, Wang F, Hu J, *et al*. Towards personalized medicine: Leveraging patient similarity and drug similarity analytics. Proceedings of the 2014 AMIA Joint Summits on Translational Science. San Francisco: AMIA, 2014. 132–136.
20 Wang Y, Chen W, Li B, *et al*. Learning fine-grained patient similarity with dynamic Bayesian network embedded RNNs. Proceedings of the 24th International Conference on Database Systems for Advanced Applications. Chiang Mai: Springer, 2019. 587–603.
21 Zhang XL, Qian BY, Li Y, *et al*. Learning representations from local to global for fine-grained patient similarity measuring in intensive care unit. Proceedings of the 2022 IEEE International Conference on Data Mining (ICDM). Orlando: IEEE, 2022. 713–722.
22 Jia C, Jia C, Kong L, *et al*. Privacy-aware retrieval of electronic medical records by fuzzy keyword search. Human-Centric Computing and Information Sciences, 2022, 12: 1–15.
23 Wang K, Xia EY, Zhao SW, *et al*. Fast similar patient retrieval from large scale healthcare data: A deep learning-based binary hashing approach. In: Shaban-Nejad A, Michalowski M, Buckeridge DL, eds. Explainable AI in Healthcare and Medicine: Building a Culture of Transparency and Accountability, Cham: Springer, 2021. 11–21.
24 Gu YF, Yang XB, Sun MX, *et al*. Graph-guided deep hashing networks for similar patient retrieval. Computers in Biology and Medicine, 2024, 169: 107865. [doi: 10.1016/j.compbiomed.2023.107865]
25 Quellec G, Lamard M, Bekri L, *et al*. Multimodal medical case retrieval using Bayesian networks and the Dezert-Smarandache theory. Proceedings of the 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro. Paris: IEEE, 2008. 245–248.
26 Silva-Leite J, Fontes C A P, Santos A S, *et al*. Aggregating embeddings from image and radiology reports for multimodal Chest-CT retrieval. Proceedings of the 37th International Symposium on Computer-based Medical Systems (CBMS). Guadalajara: IEEE, 2024. 309–314
27 Zhang PF, Li Y, Huang Z, *et al*. Aggregation-based graph convolutional hashing for unsupervised cross-modal retrieval. IEEE Transactions on Multimedia, 2021, 24: 466–479
28 Weiss Y, Torralba A, Fergus R. Spectral hashing. Proceedings of the 22nd International Conference on Neural Information Processing Systems. Vancouver: Curran Associates Inc., 2009. 1753–1760.
29 Shen FM, Shen CH, Liu W, *et al*. Supervised discrete hashing. Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston: IEEE, 2015. 37–45.
30 Zhu H, Long MS, Wang JM, *et al*. Deep hashing network for efficient similarity retrieval. Proceedings of the 30th AAAI Conference on Artificial Intelligence. Phoenix: AAAI Press, 2016. 2415 –2421.
31 Cao ZJ, Long MS, Wang JM, *et al*. HashNet: Deep learning to hash by continuation. Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017. 5609–5618.
32 Zhang Z, Zou Q, Lin YW, *et al*. Improved deep hashing with soft pairwise similarity for multi-label image retrieval. IEEE Transactions on Multimedia, 2020, 22(2): 540–553. [doi: 10.1109/TMM.2019.2929957]
33 Xu J, Xu ZX, Walker P, *et al*. Federated patient hashing. In Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York: AAAI Press, 2020. 6486–6493.
34 Lee J, Yoon W, Kim S, *et al*. BioBERT: A pre-trained biomedical language representation model for biomedical text mining. Bioinformatics, 2020, 36(4): 1234–1240. [doi: 10.1093/bioinformatics/btz682]
35 Zhang S, Xu YB, Usuyama N, *et al*. BioMedCLIP: A multimodal biomedical foundation model pretrained from fifteen million scientific image-text pairs. arXiv:2303.00915, 2024.
36 Vaswani A, Shazeer N, Parmar N, *et al*. Attention is all you need. Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017. 6000–6010.
37 Singh A, Gupta S. Learning to hash: A comprehensive survey of deep learning-based hashing methods. Knowledge and Information Systems, 2022, 64(10): 2565–2597. [doi: 10.1007/s10115-022-01734-0]
38 Johnson AEW, Pollard TJ, Shen L, *et al*. MIMIC-III, a freely accessible critical care database. Scientific Data, 2016, 3: 160035. [doi: 10.1038/sdata.2016.35]
39 Järvelin K, Kekäläinen J. Cumulated gain-based evaluation of ir techniques. ACM Transactions on Information Systems (TOIS), 2002, 20(4): 422–446. [doi: 10.1145/582415.582418]
40 Beitzel SM, Jensen EC, Frieder O. MAP. Liu L, Tamer Özsu M. Encyclopedia of Database Systems. New York: Springer, 2009. 1691–1692.
41 Liu W, Wang J, Ji RR, *et al*. Supervised hashing with kernels. Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence: IEEE, 2012. 2074–2081.
42 Li WJ, Wang S, Kang WC. Feature learning based deep supervised hashing with pairwise labels. arXiv:1511.03855, 2015.

(校对责编: 王欣欣)