



面向医学图像检索的模糊注意力协同哈希网络

周琳琳, 丁卫平, 张炜, 黄嘉爽, 武倩楠

引用本文:

周琳琳, 丁卫平, 张炜, 等. 面向医学图像检索的模糊注意力协同哈希网络[J]. *智能系统学报*, 2026, 21(4): 932–942.

ZHOU Linlin, DING Weiping, ZHANG Wei, et al. Fuzzy attentional collaborative hashing network for medical image retrieval[J]. *CAAI Transactions on Intelligent Systems*, 2026, 21(4): 932–942.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202508013>

您可能感兴趣的其他文章

自步稀疏最优均值主成分分析

Sparse optimal mean principal component analysis based on self-paced learning
智能系统学报. 2021, 16(3): 416–424 <https://dx.doi.org/10.11992/tis.201911028>

利用残差密集网络的运动模糊复原方法

Image restoration with residual dense network
智能系统学报. 2021, 16(3): 442–448 <https://dx.doi.org/10.11992/tis.201912002>

基于注意力机制的显著性目标检测方法

Salient object detection method based on the attention mechanism
智能系统学报. 2020, 15(5): 956–963 <https://dx.doi.org/10.11992/tis.201903001>

基于注意力融合的图像描述生成方法

An image caption generation method based on attention fusion
智能系统学报. 2020, 15(4): 740–749 <https://dx.doi.org/10.11992/tis.201910039>

基于自编码器的特征迁移算法

Feature transfer algorithm based on an auto-encoder
智能系统学报. 2017, 12(6): 894–898 <https://dx.doi.org/10.11992/tis.201706037>

行人重识别研究综述

Survey on pedestrian re-identification research
智能系统学报. 2017, 12(6): 770–780 <https://dx.doi.org/10.11992/tis.201706084>

DOI: 10.11992/tis.202508013

网络出版地址: <https://link.cnki.net/urlid/23.1538.tp.20260403.0847.004>

面向医学图像检索的模糊注意力协同哈希网络

周琳琳, 丁卫平, 张炜, 黄嘉爽, 武倩楠

(南通大学人工智能与计算机学院, 江苏 南通 226019)

摘要: 近年来, 医学影像技术的革新推动图像数据激增, 高效检索成为提升诊断效率的关键。深度哈希算法以其高效编码广泛应用于图像检索, 但在处理不确定的医学图像及不平衡数据时存在局限。为此, 本文提出一种基于模糊系统的模糊注意力协同哈希网络, 用于高效的医学图像检索。设计一个注意力增强的模糊特征协同器, 通过计算模糊特征的注意力权重, 提高类间差异性; 基于 Transformer 的后件协同编码器利用多头自注意力机制增强图像语义建模, 处理医学图像的非线性关系; 最后, 综合成对损失、平衡损失、量化损失与分类损失对模型进行优化, 生成紧凑哈希码。在两个医学图像数据集上的大量实验证明了该算法的有效性。

关键词: 医学图像; 图像检索; 模糊系统; 注意力机制; 深度学习; 特征提取; 深度哈希; 数据处理

中图分类号: TP391 **文献标志码:** A **文章编号:** 1673-4785(2026)04-0932-11

中文引用格式: 周琳琳, 丁卫平, 张炜, 等. 面向医学图像检索的模糊注意力协同哈希网络 [J]. 智能系统学报, 2026, 21(4): 932-942.

英文引用格式: ZHOU Linlin, DING Weiping, ZHANG Wei, et al. Fuzzy attentional collaborative hashing network for medical image retrieval[J]. CAAI transactions on intelligent systems, 2026, 21(4): 932-942.

Fuzzy attentional collaborative hashing network for medical image retrieval

ZHOU Linlin, DING Weiping, ZHANG Wei, HUANG Jiashuang, WU Qiannan

(School of Artificial Intelligence and Computer Science, Nantong University, Nantong 226019, China)

Abstract: In recent years, innovations in medical imaging technology have driven a surge in image data, with efficient retrieval becoming a critical factor in enhancing diagnostic efficiency. Deep hashing algorithms are widely used in image retrieval due to their efficient encoding. However, these algorithms have limitations when applied to uncertain medical images and imbalanced datasets. To address these challenges, this paper proposes a fuzzy attention collaborative hashing network (FACH) for efficient medical image retrieval. First, fuzzy feature coordinator enhanced by attention is introduced, which improves inter-class discrimination by calculating attention weights for fuzzy features. Then, Transformer-based posterior collaborative encoder employs the multi-head self-attention mechanism to enhance image semantic modeling and capture the nonlinear relationships in medical images. Finally, the model is optimized by integrating pairwise loss, balance loss, quantization loss, and classification loss to generate compact and discriminative hash codes. Extensive experiments conducted on two medical image datasets validate the effectiveness of the proposed algorithm.

Keywords: medical image; image retrieval; fuzzy systems; attention mechanisms; deep learning; feature extraction; deep hashing; data processing

近年来, 医学影像技术迅速发展, 临床和科研

团队已积累大量医学影像资源。例如, 上海人工智能实验室构建的超大规模数据库包含 460 万张影像和 1970 万张标注掩膜^[1]。在此背景下, 构建高效的医学影像检索系统具有重要临床价值, 能为诊断和研究提供支持^[2]。哈希算法作为提升大规模数据检索效率的核心工具, 已在数据库管理^[3]、分布式架构^[4]、文件系统优化^[5]等领域表现出优

收稿日期: 2025-08-10. 网络出版日期: 2026-04-03.

基金项目: 国家重点研发计划项目 (2024YFE0202700); 国家自然科学基金项目 (62576178, U2433216, 62471259, 62506176); 江苏省自然科学基金项目 (BK20231337); 江苏省高等学校自然科学基金项目 (24KJB520032); 江苏省研究生科研与实践创新计划项目 (SJCX25_2004, KYCX25_3741).

通信作者: 丁卫平. E-mail: dwp9988@163.com.

势。然而,传统哈希算法依赖手工特征提取,存在语义差距、检索速度慢和准确性低等问题^[6]。

随着深度学习技术的快速发展,一些联合深度学习技术的深度哈希算法被提出^[7]。这类方法通常先利用卷积神经网络等深度学习模型进行特征提取,然后通过哈希层生成紧凑二进制编码。例如 Ding 等^[8]提出了一种基于视觉 Transformer 哈希和残差模糊生成的图像检索方法,引入生成对抗网络和模糊逻辑策略,显著提升了检索的鲁棒性;Kou 等^[9]提出了注意引导对比哈希网络,引入加权特征增强模块有效缓解了降维过程中的信息损失;He 等^[10]提出了一种结合卷积神经网络和自注意力机制的深度哈希方法,采用分阶段架构并利用块聚合函数实现局部自注意力;Shen 等^[11]通过阶段式注意力机制逐步定位图像的判别区域,同时抑制已激活区域以发现其他互补区域。

虽然当前深度哈希算法在图像检索中取得了一定进展,但仍然面临以下局限性。1)深度模糊哈希算法通过模糊层处理样本间的关系,但忽略了样本间的语义信息和特征间的非线性关系,导致其难以有效处理医学影像中高度不确定的病理特征。因此,增强深度哈希算法对不确定性数据的建模能力,对于提升检索性能至关重要。2)基于注意力的深度哈希算法利用注意力机制捕捉图像中的关键信息,但这些方法未充分考虑数据不平衡性。因此,增强数据的类间区分度和自适应能力有助于提高哈希码的判别性。3)目前,深度哈希算法与模糊理论在医学图像检索中的结合应用仍未得到充分探索,亟需进一步研究。

模糊系统是一个具有较强不确定性数据挖掘能力的智能模型,其主要通过模糊隶属度实现对不确定性信息的自适应捕捉^[12]。当前模糊系统在多个不同领域被广泛应用,在医学图像分类领域,Zhang 等^[13]提出可解释的 Takagi-Sugeno-Kang (TSK) 模糊系统,通过放射组学特征分类和引入软标签矩阵,成功实现医学图像的多分类任务;Song 等^[14]提出的标签引导层次 TSK 模糊分类器,在分类急性放射性皮炎时,采用标签干预技术提高了分类准确性。在脑电图信号分类方面,Ta 等^[15]引入筛选机制,提出的多层次 TSK 模糊系统实现了高效的脑电图信号分类;Li 等^[16]则提出了一种基于模糊系统的多视角信息保留转移表示学习方法,有效解决了癫痫脑电图分析中的数据不足和模型不可解释性问题。目前,模糊系统在医学领域主要应用于特定疾病的诊断分析,针对不同类型的医学信号和图像,利用其处理不确定性和复杂性优

势,提出创新解决方案,显著提升分类精度和系统可解释性^[17]。因此,如何基于模糊系统开发出高效的医学图像检索系统,成为一个亟待关注的研究方向。

针对医学图像检索中存在的特征不确定性挑战,本文提出了一种全新的模糊注意力协同哈希网络(fuzzy attentional collaborative hashing network, FACH),该模型融合 TSK 模糊系统的建模优势与深度学习的特征表达能力,实现两个关键阶段的协同优化。首先,在模糊特征提取阶段,摒弃传统基于固定触发强度的模糊化方法,创新地设计注意力引导的模糊特征协同器,通过注意力权重矩阵对医学图像特征进行空间重组:该模块利用可学习的注意力分数增强关键病灶区域特征响应,同时抑制非显著背景噪声,使同类样本在模糊空间内的相似度显著提升,异类样本相似度有效降低。其次,在模糊推理优化阶段,本文引入了 Transformer 网络^[18],构建基于 Transformer 的后件协同编码器替代传统线性函数,其多头自注意力机制通过建立跨区域长程依赖关系,进而提升了非线性特征映射的精度。最后,设计由成对损失、平衡损失、量化损失和分类损失构成的损失函数对模型进行训练。本文的主要贡献:

1)基于 TSK 模糊系统提出了模糊注意力协同哈希网络,通过融合 TSK 模糊系统的不确定性建模与深度学习的非线性表征优势,显著增强模型对医学图像不确定信息的处理能力。

2)设计注意力增强的模糊特征协同器与基于 Transformer 的后件协同编码器的联合优化框架,前者通过注意力机制聚焦特征空间关键区域,后者捕捉图像深层非线性语义关联,有效解决复杂医学图像建模能力不足的问题。

3)在 ChestX-ray14 和 ISIC2018 医学图像数据集上验证了本文提出的算法的有效性,对比先进的深度哈希算法,本文提出的算法在图像检索应用中取得了较好的效果。

1 相关工作

1.1 TSK 模糊系统

TSK 模糊系统的工作机制包含输入模糊化、模糊规则执行及输出解模糊化 3 个阶段,适用于复杂非线性系统的建模与控制^[19]。该系统在保持较强数据驱动的建模能力的同时,具有较强的可解释性,因而在特征识别^[20]、迁移学习^[21]和大数据处理^[22]等智能计算场景中具有重要应用价值。

TSK 模糊系统的模糊推理规则定义为

$$\begin{aligned} &\text{IF } x_1 \text{ is } A_1^r \text{ and } x_2 \text{ is } A_2^r \text{ and } \cdots \text{ and } x_d \text{ is } A_d^r, \\ &\text{THEN } f^r(\mathbf{x}) = p_0^r + p_1^r x_1 + p_2^r x_2 + \cdots + p_d^r x_d \quad (1) \\ &r = 1, 2, \dots, R \end{aligned}$$

式中: R 代表规则库中模糊规则数, $\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_d]^T$ 代表规则的前件输入向量, $A_i^r (i = 1, 2, \dots, d)$ 代表第 i 个特征 x_i 的第 r 条规则所对应的模糊集, $f^r(\mathbf{x})$ 代表第 r 条规则的输出, $p_j^r (j = 1, 2, \dots, d)$ 代表规则的后件参数。

隶属度函数是 TSK 模糊系统的核心组件, 定义了输入变量的模糊分布, 直接影响系统的推理机制与输出特性。常用模糊隶属度函数是高斯函数, 定义为

$$\mu_{A_i^r}(x_i) = \exp\left(\frac{-(x_i - c_{r,i})^2}{2\delta_{r,i}}\right) \quad (2)$$

式中: $c_{r,i}$ 、 $\delta_{r,i}$ 为前件参数, 分别代表高斯隶属函数的中心和宽度。通常采用聚类算法或者空间划分技术估计得到, 例如模糊 C 均值聚类^[23]。对于输入 $\mathbf{x}$, 第 r 条规则的触发强度 $\mu^r(\mathbf{x})$ 以及归一化为

$$\mu^r(\mathbf{x}) = \prod_{i=1}^d \mu_{A_i^r}(x_i) \quad (3)$$

$$\tilde{\mu}^r(\mathbf{x}) = \frac{\mu^r(\mathbf{x})}{\sum_{r=1}^R \mu^r(\mathbf{x})} \quad (4)$$

最后, TSK 模糊系统的输出为

$$y = \sum_{r=1}^R \tilde{\mu}^r(\mathbf{x}) f^r(\mathbf{x}) \quad (5)$$

在获得前件参数后, TSK 模糊系统的输出由式(5)重新表示为

$$y = \mathbf{x}_g \mathbf{p}_g \quad (6)$$

式中: $\mathbf{p}_g$ 代表 TSK 模糊系统的后件参数, $\mathbf{x}_g$ 代表通过 TSK 模糊系统的前件进行模糊映射得到的特征向量, $\mathbf{p}_g$ 和 $\mathbf{x}_g$ 具体定义为

$$\begin{aligned} \mathbf{x}_e &= (1, \mathbf{x}) \in \mathbb{R}^{1 \times (d+1)} \\ \tilde{\mathbf{x}}^r &= \tilde{\mu}^r(\mathbf{x}) \mathbf{x}_e \in \mathbb{R}^{1 \times (d+1)} \\ \mathbf{x}_g &= [\tilde{\mathbf{x}}^1 \ \tilde{\mathbf{x}}^2 \ \cdots \ \tilde{\mathbf{x}}^R] \in \mathbb{R}^{1 \times R(d+1)} \\ \mathbf{p}^r &= [p_0^r \ p_1^r \ \cdots \ p_d^r] \in \mathbb{R}^{1 \times (d+1)} \\ \mathbf{p}_g &= [\mathbf{p}^1 \ \mathbf{p}^2 \ \cdots \ \mathbf{p}^R]^T \in \mathbb{R}^{R(d+1) \times 1} \end{aligned} \quad (7)$$

式中: $\tilde{\mathbf{x}}^r$ 代表输入 $\mathbf{x}$ 基于第 r 条规则前件的特征映射, $\mathbf{p}^r$ 代表第 r 条规则的后件参数。

1.2 深度哈希算法

深度哈希算法通过融合深度学习模型提取高维数据的特征, 生成紧凑的二进制哈希码。该算法不仅显著提升了检索速度和准确性, 还降低了存储和计算成本, 广泛应用于图像、视频和文本

等多媒体检索领域^[24]。深度哈希算法一般分为无监督方法、半监督方法、监督方法^[25]。无监督方法利用数据的内在结构生成哈希码, 无需标签信息^[26-27]。半监督方法结合少量标签信息和大量无标签数据, 学习优化哈希函数^[28-29]。

监督方法通过充分利用丰富的标签信息, 采用监督学习直接优化哈希码, 因此检索精度通常高于无监督和半监督方法^[30]。经典的深度监督哈希算法是由 Liu 等^[31]提出的, 通过卷积神经网络和损失函数实现了高效的图像检索; Jiang 等^[32]通过采用对称学习策略, 提出了非对称深度监督哈希方法, 实现大规模最近邻搜索; 为提高哈希码的灵活性。随着 Transformer 模型在计算机视觉领域的广泛应用, 其在哈希算法中的应用也取得了显著进展; Dubey 等^[33]率先将视觉 Transformer 引入哈希算法领域, 采用预训练模型作为特征提取模块, 显著提升了图像检索的性能; 为进一步增强哈希码的判别能力, Chen 等^[34]提出了基于双流特征的视觉 Transformer 架构, 该架构通过并行学习全局特征和局部特征, 有效提高了哈希码的表征能力; 刘传升等^[35]将空间金字塔池化和幂均值变换引入视觉 Transformer 哈希模型, 提出了面向医学图像检索的视觉 Transformer 哈希改进算法, 显著提升了特征表征能力与检索精度。

尽管这些深度哈希算法考虑了局部与全局特征的融合以及非线性表征能力的提升, 但忽视了图像特征固有的不确定性和数据不平衡, 导致在面对不确定和复杂图像时检索性能受限。

2 模糊注意力协同哈希网络

如图 1 所示, 模糊注意力协同哈希网络主要包含以下模块: 首先, 引入注意力增强的模糊特征协同器, 计算模糊化特征的注意力权重, 在保留医学图像区域关键信息的同时, 显著提升模糊空间的特征判别性; 其次, 设计基于 Transformer 的后件协同编码器, 捕捉图像深层语义关联, 强化模型对复杂医学图像非线性关系的表达能力; 最后, 通过包含成对损失、平衡损失、量化损失和分类损失的损失函数进行优化。

2.1 符号定义

给定医学图像数据库 $\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^n$, 其中 n 为数据库图像的数量。 $\mathbf{L} = \{\mathbf{l}_i\}_{i=1}^n \in \{0, 1\}^{n \times c}$ 代表数据库 $\mathbf{X}$ 的标签, 其中 c 为数据库的类别数。从数据库 $\mathbf{X}$ 中随机抽取 m 张图像作为查询集 (训练集) $\mathbf{Q} = \{\mathbf{q}_i\}_{i=1}^m$, $\mathbf{L}' \subseteq \mathbf{L}$ 为对应的标签信息。数据库和查询集的相似性表示为 $S_{ii} = \frac{2||\mathbf{l}_i \cap \mathbf{l}_i||}{||\mathbf{l}_i|| + ||\mathbf{l}_i|| - ||\mathbf{l}_i \cap \mathbf{l}_i||}$,

且 $S_{ii} \in [0, 2]$ 。如果 $S_{ii} = 2$, 则 q_i 与 x_i 相似; 如果 $S_{ii} = 0$, 则 q_i 与 x_i 不相似; 如果 $S_{ii} \in (0, 2)$, 则 q_i 与 x_i 具有一定的相似度。

用于图像检索的深度哈希算法目标是针对查询集学习一个哈希函数 $h(q_i)$, 将输入图像映射为二进制哈希码 $B = \{b_i\}_{i=1}^m \in \{-1, 1\}^{m \times k}$, k 为哈希码长。

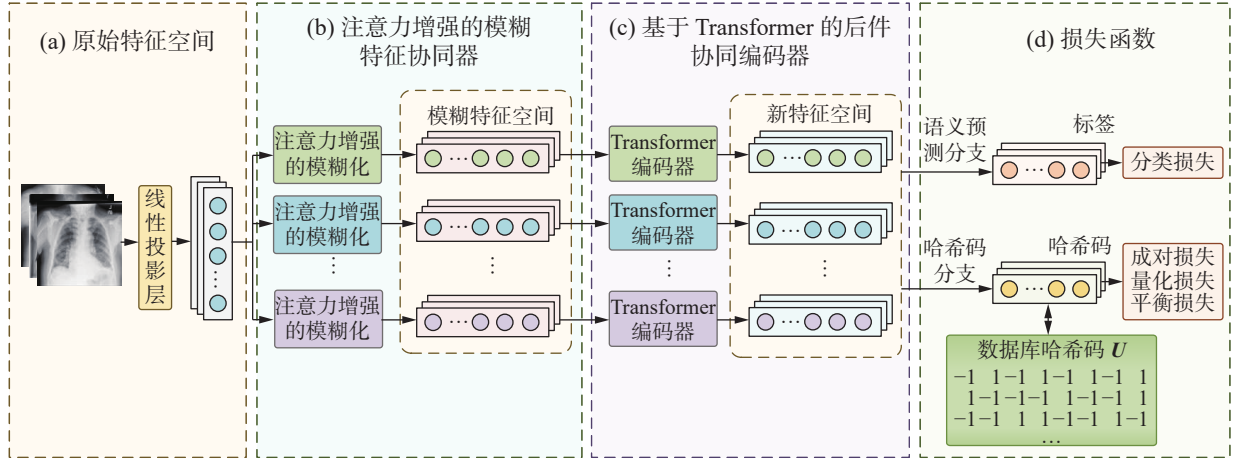


图 1 模糊注意力协同哈希网络框架

Fig. 1 Framework of fuzzy attention collaborative hashing network

2.2 模糊注意力协同哈希网络的规则

传统的模糊系统具有强大的不确定性数据处理能力, 其通过构建模糊规则库并利用模糊推理操作有效挖掘不确定性数据的内在关联。在此基础上, 本文基于 TSK 模糊系统设计了模糊注意力协同哈希网络。首先, 参考经典 TSK 模糊系统及文献 [36], 定义 FACH 的模糊推理规则库:

$$\text{IF } q_i \text{ is AEFFC}^r(q_i), \text{ THEN } f^r(q_i) = \text{TCCE}^r(q_i), \quad (8)$$

$$r = 1, 2, \dots, R$$

式中: q_i 为查询集中第 i 个样本, AEFFC $^r(\cdot)$ 为第 r 个规则前件的注意力增强的模糊特征协同器, TCCE $^r(\cdot)$ 为第 r 个规则后件的基于 Transformer 的后件协同编码器, $f^r(q_i)$ 为第 r 个规则的输出, R 为规则数。

基于式 (8) 构造了模糊注意力协同哈希网络。如图 1 所示, FACH 由两个模块组成: 注意力增强的模糊特征协同器、基于 Transformer 的后件协同编码器。在模糊化阶段, 传统方法直接采用式 (7) 所示的隶属函数进行特征映射, 难以捕捉医学图像内部的结构相似性。为此, 本文提出注意力增强的模糊特征协同器, 通过在模糊空间中构建注意力权重矩阵, 实现特征区域的差异化增强。在后件推理阶段, 经典的 TSK 模糊系统采用式 (6) 线性函数进行去模糊化, 导致对复杂医学图像的非线性关系建模能力不足。对此, 本研究设计基于 Transformer 的后件协同编码器, 更好地适应复杂的非线性关系, 提高了处理图像不确定性的能力。

2.3 注意力增强的模糊特征协同器

本文提出的模糊注意力协同哈希网络延续了传统模糊系统的设计, 采用高斯隶属度函数作为隶属度函数。为获得前件参数, 基于模糊 C 均值聚类算法估计各高斯隶属度函数的中心位置与方差参数。当前件参数确定后, 依据式 (2)~(4) 的映射关系, 输入样本 q_i 在模糊特征空间中可表示为

$$\tilde{q}_i^r = \tilde{\mu}^r(q_i) q_{i,e} \in \mathbf{R}^{1 \times (d+1)} \quad (9)$$

其中 $q_{i,e} = (1, q_i) \in \mathbf{R}^{1 \times (d+1)}$, 则查询集 Q 在模糊特征空间的表示为 $\tilde{Q}^r = [\tilde{q}_1^r \ \tilde{q}_2^r \ \dots \ \tilde{q}_m^r]^T \in \mathbf{R}^{m \times (d+1)}$ 。

传统模糊系统的模糊化过程依赖于输入数据与触发强度的乘积运算, 该方法破坏了原始数据空间的拓扑保持性, 导致相似样本在模糊空间中的距离关系失真。为此, 本文提出基于注意力机制的模糊化框架, 通过评估不同样本之间的相关性得分, 为特征向量分配相应的相似度权重。针对查询集 Q , 改进后的模糊映射表示为

$$\tilde{Q}_{\text{atten}}^r = \text{softmax} \left(\frac{\tilde{Q}^r \cdot (\tilde{Q}^r)^T}{\sqrt{d+1}} \right) \tilde{Q}^r \quad (10)$$

式中: $d+1$ 代表了 $\tilde{Q}^r$ 向量的维度, $\frac{1}{\sqrt{d+1}}$ 是为了防止 softmax 函数计算时上溢出; $\tilde{Q}^r \cdot (\tilde{Q}^r)^T$ 反映了查询集样本之间的相似度, 值较高表明两个样本在特征空间中具有较强的相似性, 而值较低则表明样本之间的差异性。基于式 (10), 查询集改进后模糊映射表示为 $\tilde{Q}_{\text{atten}}^r = [\tilde{q}_{\text{atten},1}^r \ \tilde{q}_{\text{atten},2}^r \ \dots \ \tilde{q}_{\text{atten},m}^r]^T$ 。

为维持原始模糊化特征的完整性并缓解模型训练中的梯度衰减问题, 本文引入残差连接架构

与层归一化技术, 构建稳健的特征传播路径^[37]。因此, 最终的模糊映射表示为

$$\tilde{\mathbf{q}}_{\text{atten},r}^r = \text{LayerNorm}(\tilde{\mathbf{q}}_{\text{atten},r}^r + \tilde{\mathbf{q}}_i^r) \quad (11)$$

其中 $\text{LayerNorm}(\cdot)$ 代表层归一化操作。基于式 (11), 可以得到查询集 $\mathbf{Q}$ 的最终模糊映射表示为 $\tilde{\mathbf{Q}}_{\text{atten}}^r = [\tilde{\mathbf{q}}_{\text{atten},1}^r \ \tilde{\mathbf{q}}_{\text{atten},2}^r \ \cdots \ \tilde{\mathbf{q}}_{\text{atten},m}^r]^T$ 。

2.4 基于 Transformer 的后件协同编码器

在前件模糊化阶段, 我们引入注意力机制, 保留医学图像中的关键区域, 生成保留原始空间结构的模糊化特征。传统模糊系统通常使用线性函数作为后件参数, 这在面对复杂和非线性的数据特征时, 往往无法充分捕捉数据之间的潜在关系^[38]。在后件参数学习阶段, 传统线性回归模型因无法捕捉复杂医学图像的深层语义依赖, 导致哈希码判别性不足。为此, 本文设计基于 Transformer 的后件协同编码器。

本文构建的基于 Transformer 的后件协同编码器由一个 Transformer 编码器与两个多层感知机分支组成。其中, 第 1 个多层感知机生成哈希码, 将高维特征压缩为低维二进制哈希码; 第 2 个多层感知机构成语义预测分支, 其输出的分类损失通过反向传播优化哈希码的语义可区分性。通过该编码器, 哈希码不仅能够保持图像的相似性信息, 同时也能增强其对类别区分的能力, 确保在图像检索中的高效性和准确性。对于融合注意力思想的模糊表示 $\tilde{\mathbf{q}}_{\text{atten},r}^r$, 可以得到输出:

$$\psi_h(\tilde{\mathbf{q}}_{\text{atten},r}^r) = \text{MLP}_h^r(\text{Transformer}^r(\tilde{\mathbf{q}}_{\text{atten},r}^r))$$

$$\psi_c(\tilde{\mathbf{q}}_{\text{atten},r}^r) = \text{MLP}_c^r(\text{Transformer}^r(\tilde{\mathbf{q}}_{\text{atten},r}^r))$$

式中: $\text{Transformer}^r(\cdot)$ 为第 r 个规则的 Transformer 编码器操作, $\text{MLP}_h^r(\cdot)$ 为第 r 个规则的用于生成哈希码的多层感知机, $\text{MLP}_c^r(\cdot)$ 为第 r 个规则的用于预测类别的多层感知机。因此, 查询集样本 $\mathbf{q}_i$ 的哈希码和预测类别分别为

$$\mathbf{b}_i = h(\mathbf{q}_i; \theta_h) = \text{sign} \left(\text{softsign} \left(\sum_{r=1}^R \psi_c(\tilde{\mathbf{q}}_{\text{atten},r}^r) \right) \right) \quad (12)$$

$$\mathbf{y}_i = f(\mathbf{q}_i; \theta_c) = \text{softmax} \left(\sum_{r=1}^R \psi_c(\tilde{\mathbf{q}}_{\text{atten},r}^r) \right)$$

其中 $\text{softsign}(x) = \frac{x}{1+|x|}$ 将值映射到 $[-1, 1]$ 范围内, 符号函数为 $\text{sign}(x) = \begin{cases} -1, & x < 0 \\ 1, & x \geq 0 \end{cases}$, θ_h 和 θ_c 分别是哈希码分支和预测类别分支的网络参数。

2.5 损失函数

为提升哈希码的可判别性和相似性保持能力, 本文采用成对损失、分类损失、平衡损失和量

化损失 4 种损失函数对模型进行优化。

成对损失 L_p : 通过最小化相似样本对的汉明距离, 同时最大化不相似对的距离, 保持原始样本间相似关系, 定义为

$$L_p = \sum_{i=1}^m \sum_{j=1}^n ||\text{softsign}(\psi_h(\tilde{\mathbf{q}}_{\text{atten},r}^r)) \mathbf{u}_i - k \mathbf{S}_{ij}||^2$$

其中 $\mathbf{U} = [\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_n]^T \in \{-1, +1\}^{n \times k}$ 为数据库的哈希码。

分类损失 L_c : 针对数据集的标签信息个数存在差异的问题, 本文设计了分别适用于单标签数据集和多标签数据集的两种分类损失函数。

1) 单标签分类损失 L_{single} :

$$L_{\text{single}} = -\frac{1}{m} \times \sum_{i=1}^m \{l_i \times \log(\mathbf{y}_i) + (1-l_i) \times \log(1-\mathbf{y}_i)\}$$

2) 多标签分类损失 L_{multi} :

$$L_{\text{multi}} = -\frac{1}{c} \times \sum_{i=1}^m \left\{ l_i \times \log[(1 + \exp(-\mathbf{y}_i))^{-1}] + (1-l_i) \times \log \left[\frac{\exp(-\mathbf{y}_i)}{1 + \exp(-\mathbf{y}_i)} \right] \right\}$$

平衡损失 L_b : 确保位哈希码中出现 +1 和 -1 的概率尽可能相同, 定义为

$$L_b = \sum_k \left(\left| \sum_{i=1}^m \text{mean}(\mathbf{b}_i) \right|^2 \right)$$

量化损失 L_q : 通过添加一个正则项来约束实值输出和哈希码之间的差异, 定义为

$$L_q = \sum_{i=1}^m ||\text{softsign}(\psi_h(\tilde{\mathbf{q}}_{\text{atten},r}^r)) - \mathbf{b}_i||^2$$

因此, 模糊哈希网络的损失函数为

$$\mathbf{L} = L_p + \alpha L_c + \beta L_b + \lambda L_q$$

其中 α 、 β 、 λ 为超参数。

2.6 算法实现步骤

本文在算法 1 中详细描述了模糊注意力协同哈希网络。训练模糊注意力协同哈希网络后, 根据数据库的哈希码 $\mathbf{U} = [\mathbf{u}_1 \ \mathbf{u}_2 \ \cdots \ \mathbf{u}_n]^T \in \{-1, +1\}^{n \times k}$ 进行图像检索。具体的, 对于给定医学图像查询集 $\mathbf{Q}$ 之外的图像样本 $\mathbf{x}_j \in \mathbf{R}^{1 \times d}$, 首先通过式 (9) 得到在第 r 个模糊特征空间中的表示 $\tilde{\mathbf{x}}_j^r \in \mathbf{R}^{1 \times (d+1)}$, 然后基于注意力机制得到模糊映射表示为 $\tilde{\mathbf{x}}_{\text{atten},j}^r \in \mathbf{R}^{1 \times (d+1)}$, 最后图像样本 $\mathbf{x}_j$ 的哈希码为

$$\mathbf{b}_j = \text{sign} \left(\text{softsign} \left(\sum_{r=1}^R \text{MLP}_h^r(\text{Transformer}^r(\tilde{\mathbf{x}}_{\text{atten},j}^r)) \right) \right)$$

接着, 将其哈希码 $\mathbf{b}_j$ 与数据库哈希码 $\mathbf{U}$ 计算汉明距离 $\text{dist}_H = \frac{1}{2}(k - \langle \mathbf{b}_j, \mathbf{u}_i \rangle)$, $i = 1, 2, \dots, n$, 其中, $\langle \mathbf{b}_j, \mathbf{u}_i \rangle$ 表示向量 $\mathbf{b}_j$ 与 $\mathbf{u}_i$ 的内积。最后将汉明距离 dist_H 按照升序排序, 并按要求返回检索到的前

几张图像。

3 实验结果与分析

本节给出了模糊注意力协同哈希网络在 ChestX-ray14 和 ISIC2018 数据集上的实验结果。本文的所有实验均在基于 Windows 服务器上使用 PyTorch 深度学习框架完成(其中, CPU 的配置为 13th Gen Intel(R) Core(TM) i9-13900K(32 CPUs)、3.0 GHz, GPU 的配置为 NVIDIA GeForce RTX 4090。

3.1 实验数据集和评价指标

本文采用的两个医学图像数据集:

ChestX-ray14 数据集: 包含 30 805 名患者的 112 120 张正面 X 线胸片图像, 涵盖 14 种胸部病理。为优化数据分布, 移除疝气类别及正常图像, 并对低频病理样本进行旋转和翻转增强, 最终获得 47 723 张病理图像。实验从中选取 13 类病理各 100 张(共 1 300 张)作为测试集, 剩余 46 423 张作为数据库, 并从中随机抽取 5 000 张作为查询集。

ISIC2018 数据集: 包含 10 208 张皮肤病镜图像, 覆盖 7 种皮肤病类型。剔除皮肤纤维瘤和血管病变后, 对剩余 5 种病理样本进行数据增强, 最终构建包含 16 345 张图像的实验数据集。从中选取 5 类病理各 100 张(共 500 张)作为测试集, 剩余 15 845 张作为数据库, 并随机抽取 5 000 张作为查询集。

为评估检索性能, 本文采用整体平均精度 (mAP) 作为评价指标。

算法 1: 模糊注意力协同哈希网络 (FACH)

输入 医学图像数据库 $X = \{x_i\}_{i=1}^n$, 模糊规则数量 R , 哈希码长 k , 批处理大小 b , 迭代次数 $iters$ 和 $epochs$;

输出 数据库哈希码 U , 模糊哈希网络参数 θ_h 和 θ_c 。

1) 初始化网络参数 θ_h 、 θ_c , 数据库哈希码 U ;

2) for $iter = 1$ to $iters$ do

3) 从数据库中随机生成查询集 $Q = \{q_i\}_{i=1}^m$, 对应标签为 L' ;

4) 使用模糊 C 均值聚类算法确定模糊哈希网络的前件参数后, 通过式 (8~11) 得到在第 r 个模糊特征空间中融合注意力机制的模糊映射表示 Q'_{atten}^r ;

5) for $epoch = 1$ to $epochs$ do

6) for $batch = 1$ to $\frac{m}{b}$ do

7) 根据式 (12), 通过基于 Transformer 的后件协同编码器生成哈希码;

8) 通过反向传播更新网络参数 θ_h 和 θ_c ;

9) end

10) end

11) 采用逐比特更新的策略更新 U ;

12) end

3.2 实验结果与分析

本文将所提出的模糊注意力协同哈希网络算法和当前主流的深度哈希方法进行对比, 包括 DPSH^[30] (deep pairwise-supervised hashing)、IDHN^[39] (improved deep hashing network)、VTS^[33] (vision Transformer hashing)、SADH^[40] (deep hashing with self-supervised asymmetric semantic excavation and margin-scalable constraint)、DPN^[41] (deep polarized Network)、DAMH^[42] (deep adaptive-mapping-based hashing)、VITH^[35] (vision Transformer hashing)、ACHNet^[9] (attention-guided contrastive hashing network)、HybridHash^[10] (hybrid convolutional and self-attention deep hashing)。表 1 和图 2 给出了本文提出的模糊注意力协同哈希网络与对比算法在 ChestX-ray14 和 ISIC2018 数据集上的 mAP 对比结果。从表中可以看出在两个数据集上, 除 DAMH 外, 模糊注意力协同哈希网络在所有码长下的 mAP 值均高于其他算法。虽然 DAMH 在 ISIC2018 数据集上 36 位的 mAP 达到了 0.754, 但从图 3 中可以看出 FACH 整体性能优于 DAMH。

表 1 两个数据集上不同码长的汉明排序的 mAP
Table 1 The mAP of Hamming ranking with different code lengths on two datasets

方法	ChestX-ray14				ISIC2018			
	8 位	12 位	24 位	36 位	8 位	12 位	24 位	36 位
DPSH ^[30]	0.201	0.206	0.217	0.217	0.596	0.663	0.694	0.694
IDHN ^[39]	0.220	0.223	0.231	0.233	0.619	0.659	0.694	0.658
VTS ^[33]	0.215	0.219	0.243	0.256	0.632	0.715	0.736	0.712
SADH ^[40]	0.272	0.286	0.297	0.301	0.654	0.647	0.702	0.688
DPN ^[41]	0.236	0.258	0.283	0.287	0.617	0.640	0.657	0.682

续表 1

方法	ChestX-ray14				ISIC2018			
	8 位	12 位	24 位	36 位	8 位	12 位	24 位	36 位
DAMH ^[42]	0.261	0.265	0.276	0.290	0.652	0.709	0.672	0.754
VITH ^[35]	0.257	0.279	0.301	0.317	0.716	0.740	0.689	0.729
ACHNet ^[9]	0.284	0.300	0.305	0.311	0.601	0.643	0.687	0.691
HybridHash ^[10]	0.267	0.273	0.291	0.292	0.608	0.602	0.680	0.714
FACH	0.321	0.383	0.302	0.410	0.728	0.742	0.748	0.739

注: 加粗表示结果在本列最好。

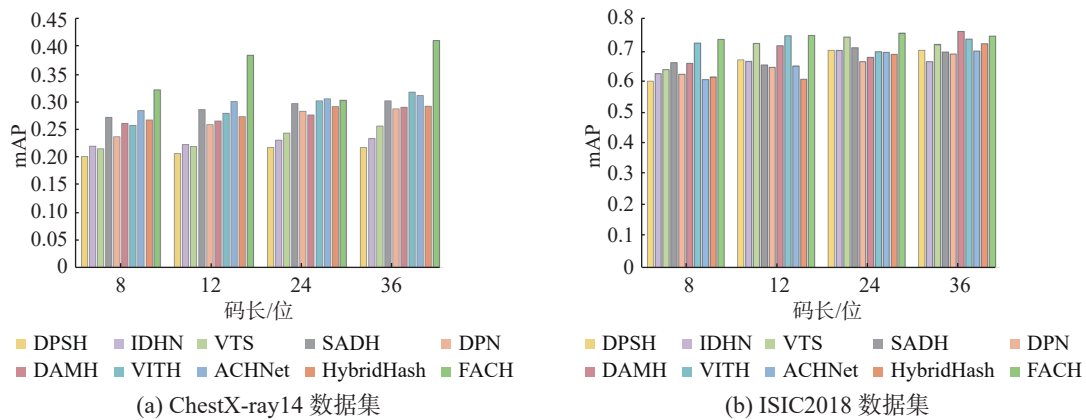


图 2 不同算法在两个数据集上的对比实验

Fig. 2 Comparative experiments of different algorithms on two datasets

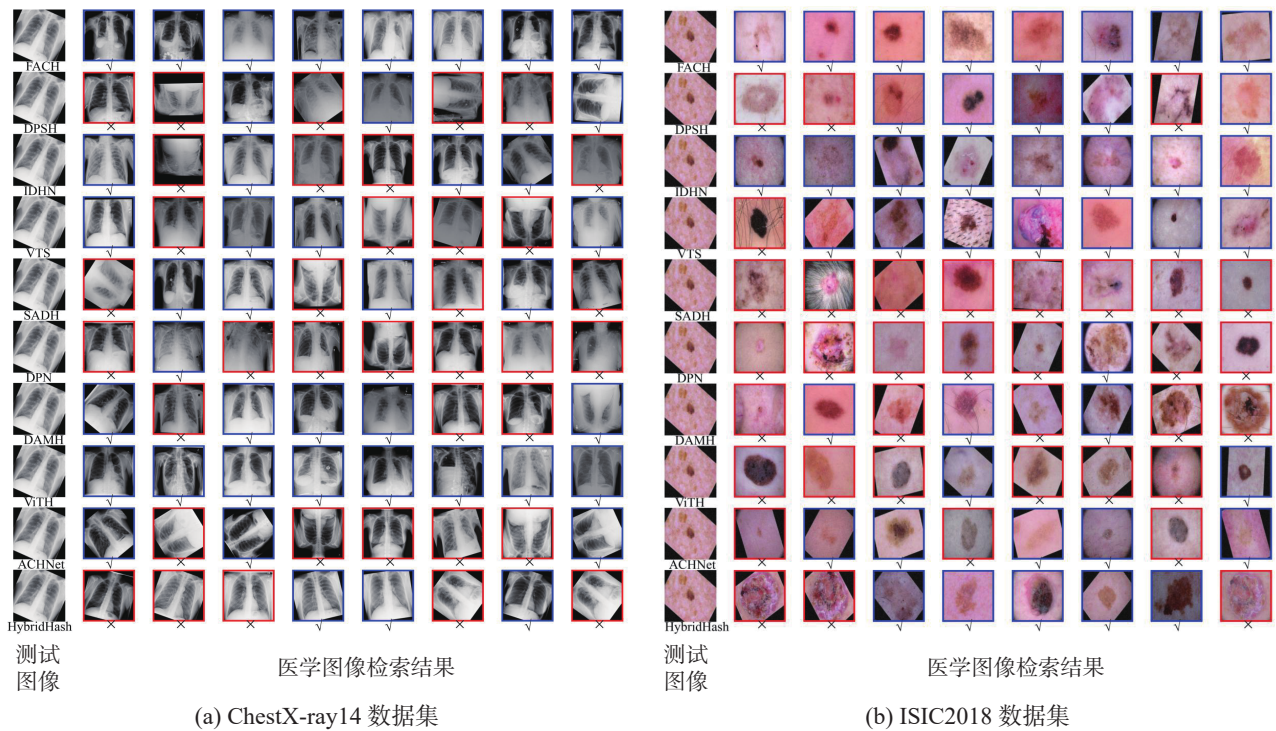


图 3 两个数据集上的医学图像检索结果

Fig. 3 Medical image retrieval results on two datasets

为进一步验证检索性能, 本文将 FACH 与对比方法应用于医学图像检索, 基于 36 位哈希码返回前 8 张检索结果。图 3 给出了在 ChestX-ray14

与 ISIC2018 数据集上的检索效果, 其中红色边框标记错误结果, 蓝色边框标记正确结果。实验结果表明, FACH 在两个数据集上均成功检索到

8张正确图像,展现出更优的检索效果。此外,表2给出了不同算法在两个数据集上8位的检索时间对比结果。从中可以看出,FACH算法在检索速度方面表现优异,尽管在ChestX-ray14数据集上略慢于SADH算法,但其具有更高的检索精度。

表2 两个数据集上的检索时间对比分析

Table 2 Comparative analysis for retrieval time on the two datasets

数据集	ChestX-ray14		ISIC2018	
方法	时间/s	排序	时间/s	排序
DPSH	28.286	8	3.939	4
DBDH	22.608	4	3.826	2
IDHN	26.294	7	3.891	3
VTS	25.173	6	3.972	6
SADH	18.726	1	14.847	8
DPN	21.122	3	4.264	7
DAMH	23.998	5	3.948	5
ACHNet	1094.316	10	127.676	10
HybridHash	71.586	9	18.454	9
FACH	20.121	2	3.812	1

注:加粗表示结果在本列最好。

3.3 消融实验

为深入分析FACH算法的整体架构设计,本文通过消融实验验证各模块的有效性。具体实验设置如下:通过逐步移除算法中的关键模块,分别评估其对模型性能的影响。设计两种FACH变体:FACH 1移除注意力机制,取消对模糊化特征的相似度加权;FACH 2去除深度TSK模糊系统,以验证其对特征建模的贡献。

表3和表4给出了ChestX-ray14和ISIC2018数据集上的消融实验结果。通过分析这些结果,可以得出以下结论:首先,在两个数据集上,FACH方法的mAP精度整体优于其他两个变体,表现出更高的准确性。其次,随着各模块的逐步去除,FACH 1和FACH 2的mAP精度下降。这说明本文提出的双模块对算法性能起到积极的促进作用,每个模块都对提升检索精度有重要贡献。最后,FACH 2的mAP精度显著低于FACH 1和FACH,这凸显了本文所设计模糊特征协同器与Transformer后件编码器的联合框架,不仅能够有效提升哈希图像检索的精度,还能增强算法在面对医学图像数据中固有的不确定性和模糊性时的表现。综上所述,消融实验证明本文提出的模糊注意力协同哈希网络显著提升算法在医学图像检索任务中的表现。

表3 在ChestX-ray14数据集上消融实验的比较分析

Table 3 Comparative analysis for ablation experiments on the ChestX-ray14 dataset

方法	8位	12位	24位	36位
FACH	0.321	0.383	0.302	0.410
FACH 1	0.310	0.334	0.310	0.305
FACH 2	0.257	0.277	0.301	0.301

注:加粗表示结果在本列最好。

表4 在ISIC2018数据集上消融实验的比较分析

Table 4 Comparative analysis for ablation experiments on the ISIC2018 dataset

方法	8位	12位	24位	36位
FACH	0.728	0.742	0.748	0.739
FACH 1	0.724	0.729	0.721	0.741
FACH 2	0.711	0.694	0.693	0.704

注:加粗表示结果在本列最好。

3.4 鲁棒性分析

为进一步分析FACH在图像干扰条件下的鲁棒性表现,本文在ISIC2018数据集上引入随机噪声干扰,即每幅图像注入10000个噪声点。如表5所示,在噪声干扰条件下,所有对比算法的mAP值均出现不同程度下降。值得注意的是,在不同哈希码长设置下,FACH始终保持着最优的mAP性能,这充分证实了FACH模型对图像噪声干扰具有更强的鲁棒性,展现了其在噪声环境下的稳定特征保持能力。此外,除FACH外,VITH同样展示了卓越的医学图像检索性能。这归因于VITH通过引入空间金字塔池化模块,能够有效地提取多尺度的特征表示,从而增强其在捕捉医学图像中微小病理特征方面的能力。图4给出了不同算法在噪声干扰下的检索时间对比分析结果。从表5中可以看出,尽管FACH的检索速率略低于IDHN、VTS,但在实际检索中FACH表现出了更高的检索精度。

表5 在ISIC2018数据集上的噪声实验

Table 5 Noise experiments on the ISIC2018 dataset

方法	8位	12位	24位	36位
DPSH	0.491	0.575	0.628	0.632
IDHN	0.572	0.590	0.598	0.616
VTS	0.565	0.584	0.612	0.631
SADH	0.553	0.570	0.592	0.594
DPN	0.498	0.565	0.618	0.623
DAMH	0.530	0.540	0.614	0.641
VITH	0.574	0.593	0.621	0.649
ACHNet	0.524	0.565	0.593	0.610
HybridHash	0.538	0.563	0.610	0.619
FACH	0.642	0.674	0.656	0.684

注:加粗表示结果在本列最好。

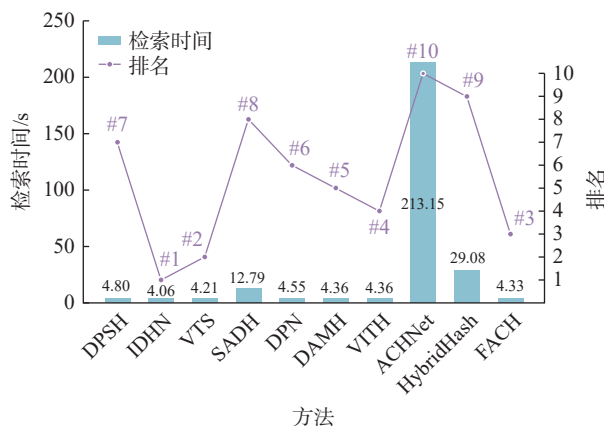


图4 在噪声下的检索时间对比分析

Fig. 4 Comparison of retrieval time under noise

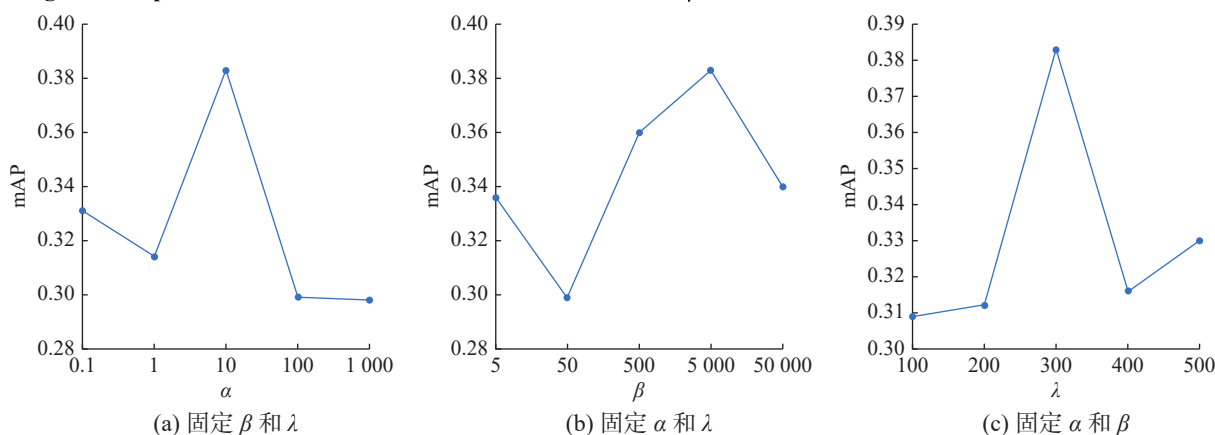


图5 在 ChestX-ray14 数据集上对 12 位哈希码的超参数分析

Fig. 5 Hyperparameter analysis for 12-bit hashing codes on the ChestX-ray14 dataset

4 结束语

本文针对医学图像数据中存在的 uncertainty 以及数据不平衡问题,提出了一种创新的深度哈希算法——模糊注意力协同哈希网络。通过结合 TSK 模糊系统的动态建模与深度学习的强大特征表达能力,本算法采用双模块协同机制,有效提升了医学图像检索的性能。具体而言,在模糊系统的前件,通过引入注意力机制,对模糊化特征进行加权,保留了医学图像关键区域的关键信息,同时增强了类内一致性和类间差异性。基于 Transformer 的后件协同编码器则进一步建模图像的复杂非线性关系,强化了对医学图像深层语义的捕捉能力。通过这种双模块的协同作用,本文所提出的算法显著提升了哈希码在医学图像检索中的表现,在 ChestX-ray14 和 ISIC2018 数据集上均取得了优异的实验结果,证明了该方法在处理复杂医学影像特征时的有效性。

该算法在当前的医学图像检索任务中表现出了优势,但仍有进一步优化的空间。首先,现有算法主要针对单模态数据,未来研究可扩展至跨

3.5 超参数分析

为进一步分析 FACH 中损失函数的 3 个超参数对于算法性能的影响,我们针对超参数 α 、 β 、 λ ,在 ChestX-ray14 数据集上对 12 位哈希码进行超参数分析。其中, α 代表分类损失的权重, β 代表平衡损失的权重, λ 代表量化损失的权重。在实验设计中,仅调整一个参数,而将其他两个参数固定为其最优值。从图 5(a)中可以看出,固定 β 和 λ ,当 $\alpha = 10$ 时 FACH 的 mAP 精度最高;从图 5(b)中可以看出,固定 α 和 λ ,当 $\beta = 5000$ 时 FACH 的 mAP 精度最高;从图 5(c)中可以看出,固定 α 和 β ,当 $\lambda = 300$ 时 FACH 的 mAP 精度最高。

模态医学图像检索任务,提升其在跨模态数据上的适应性,随着医学影像数据的日益多样化,跨模态检索将成为一个亟待解决的研究方向。其次,当前的深度哈希算法在面对数据噪声与恶意攻击时,鲁棒性较弱,因此,未来可探讨对抗性训练的引入,以增强模型的鲁棒性,进一步提升其在实际应用中的稳定性与可靠性。

参考文献:

- [1] YE Jin, CHENG Junlong, CHEN Jianpin, et al. SA-Med2D-20M dataset: segment anything in 2D medical imaging with 20 million masks[EB/OL]. (2023-11-20) [2025-08-10]. <https://arxiv.org/abs/2311.11969>.
- [2] JUSH F K, TRUONG T, VOGLER S, et al. Medical image retrieval using pretrained embeddings[C]//2024 IEEE International Symposium on Biomedical Imaging. Athens: IEEE, 2024: 1-5.
- [3] KARA K, ALONSO G. Fast and robust hashing for database operators[C]//2016 26th International Conference on Field Programmable Logic and Applications. Lausanne: IEEE, 2016: 1-4.
- [4] DEVINE R. Design and implementation of DDH: a distributed dynamic hashing algorithm[C]//Foundations of

- Data Organization and Algorithms. Berlin: Springer, 1993: 101–114.
- [5] LENSING P, MEISTER D, BRINKMANN A. HashFS: applying hashing to optimize file systems for small file reads[C]//2010 International Workshop on Storage Network Architecture and Parallel I/Os. Incline Village: IEEE, 2010: 33–42.
- [6] 李来, 刘光灿, 孙玉宝, 等. 各向同性的迭代量化哈希算法[J]. 电子学报, 2017, 45(7): 1707–1714.
LI Lai, LIU Guangcan, SUN Yubao, et al. Isotropic iterative quantization hashing[J]. *Acta electronica sinica*, 2017, 45(7): 1707–1714.
- [7] LIONG V E, LU Jiwen, WANG Gang, et al. Deep hashing for compact binary codes learning[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015: 2475–2483.
- [8] DING Weiping, LIU Chuansheng, HUANG Jiashuang, et al. ViTH-RFG: vision transformer hashing with residual fuzzy generation for targeted attack in medical image retrieval[J]. *IEEE transactions on fuzzy systems*, 2024, 32(10): 5571–5584.
- [9] KOU Xuan, XU Chenghao, YANG Xu, et al. Attention-guided contrastive hashing for long-tailed image retrieval [C]//Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence. Vienna: International Joint Conferences on Artificial Intelligence Organization, 2022: 1017–1023.
- [10] HE Chao, WEI Hongxi. HybridHash: hybrid convolutional and self-attention deep hashing for image retrieval[C]//Proceedings of the 2024 International Conference on Multimedia Retrieval. Phuket: ACM, 2024: 824–832.
- [11] SHEN Yang, SUN Xuhao, WEI Xiushen, et al. SEMI-CON: a learning-to-hash solution for Large-scale fine-grained image retrieval[C]//Computer Vision–ECCV 2022. Cham: Springer, 2022: 531–548.
- [12] 周天奕, 丁卫平, 黄嘉爽, 等. 模糊逻辑引导的多粒度深度神经网络[J]. 模式识别与人工智能, 2023, 36(9): 778–792.
ZHOU Tianyi, DING Weiping, HUANG Jiashuang, et al. Fuzzy logic guided deep neural network with multi-granularity[J]. *Pattern recognition and artificial intelligence*, 2023, 36(9): 778–792.
- [13] ZHANG Yuanpeng, YANG Dongrong, LAM S, et al. Radiomics-based detection of COVID-19 from chest X-ray using interpretable soft label-driven TSK fuzzy classifier[J]. *Diagnostics*, 2022, 12(11): 2613.
- [14] SONG Guanghao, ZHAI Jia, ZHOU Ta. Classification of acute radiation dermatitis by active-guided-based T-S fuzzy classifier[C]//2023 2nd International Conference on Artificial Intelligence, Human-Computer Interaction and Robotics. Tianjin: IEEE, 2024: 154–160.
- [15] TA Zhou, WEI Xue, NAN Chuxi, et al. A screening mechanism fast-aggregation-based Takagi-Sugeno-Kang fuzzy classification for epileptic electroencephalograms signal[J]. *Journal of medical imaging and health informatics*, 2019, 9(7): 1458–1463.
- [16] LI Andong, DENG Zhaohong, ZHANG Wei, et al. Multiview transfer representation learning with TSK fuzzy system for EEG epilepsy detection[J]. *IEEE transactions on fuzzy systems*, 2024, 32(1): 38–52.
- [17] ONTIVEROS E, MELIN P, CASTILLO O. Comparative study of interval Type-2 and general Type-2 fuzzy systems in medical diagnosis[J]. *Information sciences*, 2020, 525: 37–53.
- [18] 王海鹏, 高自强, 董佳俊, 等. 基于粗糙注意力融合机制与 Group Transformer 的视网膜血管分割网络[J]. 南通大学学报 (自然科学版), 2024, 23(1): 28–37, 48.
WANG Haipeng, GAO Ziqiang, DONG Jiajun, et al. Retinal vessel segmentation network based on rough attention fusion mechanism and Group Transformer[J]. *Journal of Nantong University (natural science edition)*, 2024, 23(1): 28–37, 48.
- [19] JUANG C F. A TSK-type recurrent fuzzy network for dynamic systems processing by neural network and genetic algorithms[J]. *IEEE transactions on fuzzy systems*, 2002, 10(2): 155–170.
- [20] DENG Zhaohong, XU Peng, XIE Lixiao, et al. Transductive joint-knowledge-transfer TSK FS for recognition of epileptic EEG signals[J]. *IEEE transactions on neural systems and rehabilitation engineering*, 2018, 26(8): 1481–1494.
- [21] 蒋亦樟, 邓赵红, 王士同. 0 阶 L2 型 TSK 迁移学习模糊系统[J]. 电子学报, 2013, 41(5): 897–904.
JIANG Yizhang, DENG Zhaohong, WANG Shitong. 0-order-L2-norm-Takagi-Sugeno-Kang type transfer learning fuzzy system[J]. *Acta electronica sinica*, 2013, 41(5): 897–904.
- [22] HERRERA L J, PÉREZ M M, SANTANA J, et al. A data mining approach based on a local-global fuzzy modelling for prediction of color change after tooth bleaching using vita classical shades[C]//2009 Ninth International Conference on Intelligent Systems Design and Applications. Pisa: IEEE, 2009: 1268–1273.
- [23] HASHEMI S E, GHOLIAN-JOUYBARI F, HAJIAG-HAEI-KESHTALI M. A fuzzy C-means algorithm for optimizing data clustering[J]. *Expert systems with applications*, 2023, 227: 120377.
- [24] 陈昌红, 彭腾飞, 干宗良. 基于深度哈希算法的极光图像分类与检索方法[J]. 电子与信息学报, 2020, 42(12): 3029–3036.
CHEN Changhong, PENG Tengfei, GAN Zongliang. Aurora image classification and retrieval method based on deep hashing algorithm[J]. *Journal of electronics & information technology*, 2020, 42(12): 3029–3036.
- [25] LUO Xiao, WANG Haixin, WU Daqing, et al. A survey on deep hashing methods[J]. *ACM transactions on knowledge discovery from data*, 2023, 17(1): 1–50.
- [26] SHEN Fumin, XU Yan, LIU Li, et al. Unsupervised deep

- hashing with similarity-adaptive and discrete optimization[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2018, 40(12): 3034–3044.
- [27] VENKATESWARA H, EUSEBIO J, CHAKRABORTY S, et al. Deep hashing network for unsupervised domain adaptation[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 5385–5394.
- [28] ZHANG Jian, PENG Yuxin. SSDH: semi-supervised deep hashing for large scale image retrieval[J]. *IEEE transactions on circuits and systems for video technology*, 2019, 29(1): 212–225.
- [29] WANG Guan'an, HU Qinghao, YANG Yang, et al. Adversarial binary mutual learning for semi-supervised deep hashing[J]. *IEEE transactions on neural networks and learning systems*, 2022, 33(8): 4110–4124.
- [30] LI Wujun, WANG Sheng, KANG Wangcheng. Feature learning based deep supervised hashing with pairwise labels[EB/OL]. (2015–11–12)[2025–08–10]. <https://arxiv.org/abs/1511.03855>.
- [31] LIU Haomiao, WANG Ruiping, SHAN Shiguang, et al. Deep supervised hashing for fast image retrieval[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 2064–2072.
- [32] JIANG Qingyuan, LI Wujun. Asymmetric deep supervised hashing[C]//Proceedings of the AAAI conference on artificial intelligence, New Orleans: AAAI Press, 2018.
- [33] DUBEY S R, SINGH S K, CHU Weita. Vision transformer hashing for image retrieval[C]//2022 IEEE International Conference on Multimedia and Expo. Taipei: IEEE, 2022: 1–6.
- [34] CHEN Yongbiao, ZHANG Sheng, LIU Fangxin, et al. TransHash: transformer-based hamming hashing for efficient image retrieval[C]//Proceedings of the 2022 International Conference on Multimedia Retrieval. Newark: ACM, 2022: 127–136.
- [35] 刘传升, 丁卫平, 程纯, 等. ViTH: 面向医学图像检索的视觉 Transformer 哈希改进算法[J]. 西南大学学报(自然科学版), 2024, 46(5): 11–26.
- LIU Chuansheng, DING Weiping, CHENG Chun, et al. ViTH: improved vision transformer hashing algorithm for medical image retrieval[J]. *Journal of Southwest University (natural science edition)*, 2024, 46(5): 11–26.
- [36] HU Fuping, DENG Zhaohong, WANG Guanjin, et al. Graph fuzzy system for the whole graph prediction: concepts, models, and algorithms[J]. *IEEE transactions on fuzzy systems*, 2024, 32(3): 1383–1398.
- [37] HAN Dezhi, SHI Jingya, ZHAO Jiahao, et al. LRCN: layer-residual co-attention networks for visual question answering[J]. *Expert systems with applications*, 2025, 263: 125658.
- [38] LU Huimin, ZHANG Ming, XU Xing, et al. Deep fuzzy hashing network for efficient image retrieval[J]. *IEEE transactions on fuzzy systems*, 2021, 29(1): 166–176.
- [39] ZHANG Zheng, ZOU Qin, LIN Yuewei, et al. Improved deep hashing with soft pairwise similarity for multi-label image retrieval[J]. *IEEE transactions on multimedia*, 2020, 22(2): 540–553.
- [40] YU Zhengyang, WU Song, DOU Zhihao, et al. Deep hashing with self-supervised asymmetric semantic excavation and margin-scalable constraint[J]. *Neurocomputing*, 2022, 483: 87–104.
- [41] FAN Lixin, NG K W, JU Ce, et al. Deep polarized network for supervised learning of accurate binary hashing codes[C]//Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence Main Track. Yokohama: IJCAI, 2021: 825–831.
- [42] SU Hai, FANG Jianwei, LIU Weixing, et al. A deep hashing method of likelihood function adaptive mapping[J]. *Neural computing and applications*, 2023, 35(8): 5903–5921.

作者简介:



周琳琳, 硕士研究生, 主要研究方向为图像检索和多粒度计算。E-mail: 15861759190@163.com。



丁卫平, 教授, 博士生导师, IEEE 高级会员, 中国计算机学会杰出会员。主要研究方向为数据挖掘、多模态机器学习、多粒度计算、演化计算和医学大数据分析。连续 6 年 (2020—2025) 入选美国斯坦福大学全球前 2% 顶尖科学家榜单, 其中 2024 年和 2025 年入选“终身科学影响力排行榜”, 入选国际学术机构 ScholarGPS 2024 年全球前 0.05% 顶尖科学家榜单。江苏省高层次人才培训计划“333 工程”第二层次培养对象, 江苏省“六大人才高峰”高层次人才, 江苏省“青蓝工程”中青年学术带头人, 江苏省“青蓝工程”优秀青年骨干教师, 南通市“226 高层次人才培养工程”第二层次、第三层次培养对象。主持和主要参加国家重点研发计划重点专项、国家自然科学基金民航联合重点项目、国家自然科学基金面上和青年项目、江苏省自然科学基金、江苏省高校自然科学研究重大项目、江苏省现代农机装备与技术推广项目等各类科研项目 20 余项。获发明专利授权 43 项, 其中美国发明专利 3 项, 澳大利亚发明专利 1 项。发表相关学术论文 400 余篇。E-mail: dwp9988@163.com。



张炜, 讲师, 硕士生导师, 博士, 主要研究方向为人工智能、模式识别和多视图学习。E-mail: weizhang@ntu.edu.cn。