

面向大规模图像检索的哈希学习综述*

张雪凝¹, 刘兴波², 宋井宽³, 聂秀山², 王少华², 尹义龙¹

¹(山东大学 软件学院, 山东 济南 250101)

²(山东建筑大学 计算机科学与技术学院, 山东 济南 250101)

³(电子科技大学 计算机科学与工程学院, 四川 成都 611731)

通信作者: 刘兴波, E-mail: sclxb@mail.sdu.edu.cn; 尹义龙, E-mail: ylyin@sdu.edu.cn



摘要: 随着互联网空间中图像数据的爆发式增长和图像应用领域的拓宽, 大规模图像检索的需求与日俱增. 哈希学习为大规模图像检索提供显著的存储与检索效率, 并成为近年来一个研究热点. 现有哈希学习综述存在着时效性弱与技术路线不清晰的问题, 即多总结 5–10 年前的研究成果, 且较少总结哈希学习算法各组成部分间的关联关系. 鉴于此, 通过总结近 20 年公开发表的哈希学习文献, 对面向大规模图像检索的哈希学习进行系统的综述性研究. 首先, 介绍哈希学习的技术路线和哈希学习算法的主要组成部分, 包括损失函数、优化策略及样本外扩展映射. 其次, 将面向图像检索的哈希学习算法分为无监督哈希方法和监督哈希方法两类, 并分别梳理每类方法的研究现状和演化过程. 然后, 介绍哈希学习算法评估通用的图像数据集与评估指标, 并通过对比实验分析部分经典算法的性能. 最后, 结合哈希学习的局限性与新挑战对其发展前景进行阶段性总结与展望.

关键词: 图像检索; 大规模数据; 近似最近邻检索; 哈希学习; 相似性保持

中图法分类号: TP18

中文引用格式: 张雪凝, 刘兴波, 宋井宽, 聂秀山, 王少华, 尹义龙. 面向大规模图像检索的哈希学习综述. 软件学报. <http://www.jos.org.cn/1000-9825/7141.htm>

英文引用格式: Zhang XN, Liu XB, Song JK, Nie XS, Wang SH, Yin YL. Survey on Hash Learning for Large-scale Image Retrieval. Ruan Jian Xue Bao/Journal of Software (in Chinese). <http://www.jos.org.cn/1000-9825/7141.htm>

Survey on Hash Learning for Large-scale Image Retrieval

ZHANG Xue-Ning¹, LIU Xing-Bo², SONG Jing-Kuan³, NIE Xiu-Shan², WANG Shao-Hua², YIN Yi-Long¹

¹(School of Software, Shandong University, Jinan 250101, China)

²(School of Computer Science and Technology, Shandong Jianzhu University, Jinan 250101, China)

³(School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China)

Abstract: As image data grows explosively on the Internet and image application fields widen, the demand for large-scale image retrieval is increasing greatly. Hash learning provides significant storage and retrieval efficiency for large-scale image retrieval and has attracted intensive research interest in recent years. Existing surveys on hash learning are confronted with the problems of weak timeliness and unclear technical routes. Specifically, they mainly conclude the hashing methods proposed five to ten years ago, and few of them conclude the relationship between the components of hashing methods. In view of this, this study makes a comprehensive survey on hash learning for large-scale image retrieval by reviewing the hash learning literature published in the past twenty years. First, the technical route of hash learning and the key components of hashing methods are summarized, including loss function, optimization strategy, and out-of-sample extension. Second, hashing methods for image retrieval are classified into two categories: unsupervised hashing methods and supervised ones. For each category of hashing methods, the research status and evolution process are analyzed. Third, several image

* 基金项目: 国家自然科学基金 (62176141, 62176139, 61876098, 62206160); 山东省自然科学基金青年项目 (ZR2022QF082); 山东省自然科学基金重大基础研究项目 (ZR2021ZD15); 山东省青年泰山项目 (tsqn202103088); 山东省杰出青年基金 (ZR2021JQ26); 山东建筑大学特聘教授专项

收稿时间: 2022-09-26; 修改时间: 2023-02-11, 2023-08-22; 采用时间: 2023-12-06; jos 在线出版时间: 2024-05-08

benchmarks and evaluation metrics are introduced, and the performance of some representative hashing methods is analyzed through comparative experiments. Finally, the future research directions of hash learning are summarized considering its limitations and new challenges.

Key words: image retrieval; large-scale data; approximate nearest neighbor search; hash learning; similarity preserving

随着数据采集、分布式存储等技术的飞速发展与互联网的日益普及,数据呈现爆发式增长态势.据统计,2020年全球产生、复制的数据总量达64.2 ZB^[1].而根据《数据时代2025》报告预测,2025年的数据产生规模将增至175 ZB,相当于每天产生491 EB的数据^[2].《中国互联网络发展状况统计报告》显示,截至2021年12月,我国已有网站数量达418万,网页数量达3350亿,网民规模达10.32亿,互联网普及率达73.0%^[3].因此,可以预见互联网空间中将会累积大量的媒体数据,其中一类重要的数据是图像数据.以社交媒体领域为例,Facebook每天产生的数据达4 PB,其中照片上传数目超过3.4亿.海量图像数据为人们日常工作和生活带来便利的同时,也对传统的信息检索方法和数据库技术提出了挑战.面对这样一个大规模的、高维的、实时扩展的数据库,如何从中快速检索出用户感兴趣的图像,已成为多媒体检索和计算机视觉领域研究的热点问题之一^[4-6].此外,对大规模图像检索技术的研究也是国家战略规划的重要组成部分.在《新一代人工智能发展规划》^[7]提及的城市大数据平台建设和《政府工作报告》^[8]提及的全国一体化大数据中心体系建设中都发挥着关键作用.经过近10年的发展,大规模图像检索技术不仅为搜索引擎、电子商务等商业应用提供了重要的技术支撑^[9],而且在医学、纺织业、版权保护、地图定位等领域得到了广泛的应用.例如,用户可以向拍卖淘等购物应用上传商品图片,并即时获取相似商品的购买链接;医生可以在海量医学影像库中快速检索到多个相似部位的病灶图像,从而辅助病情诊断;制衣厂可以利用纹理图片检索便捷地管理布料样本;高德地图等应用还提供了基于地标检索实现定位的服务等.

2022年发布的《人工智能白皮书》指出,当前,在算法层面,预训练模型的训练数据规模按照每年300倍的趋势增长,在端、边等资源受限情形下,具备低内存和低计算量优势的技术成为业界需求^[10].因此,在设计大规模图像检索算法时,存储成本和检索效率是两个重要的考虑因素.为了实现准确而快速的检索,近似最近邻(ANN)检索方法通过牺牲可接受范围内的精度来加快检索速度^[11-13].现有的近似最近邻检索方法主要包括类:空间分割方法^[14,15]、矢量量化方法^[16,17]和哈希方法^[18-20].空间分割方法通常利用各种树形结构对特征空间进行分割^[14,15],以减少查询次数,从而提升检索效率.虽然此类方法在应用于低维数据时表现出良好的性能,但其在面对高维数据时往往退化为遍历搜索^[21].矢量量化方法通常利用子空间分解对数据进行压缩表示^[16,17],从而加速查询过程中的距离计算.相较于空间分割方法,矢量量化方法在对高维数据的处理方面表现出优势.然而,许多矢量量化方法在进行数据压缩时仅考虑了单个数据的特征重构,而忽略了原始数据间的相似性.与之不同,哈希方法对保相似性(即保持原始数据间的近邻关系)的低维表示进行了探索^[18-20].与前两类方法相比,哈希方法通过将高维特征映射为简洁的二值哈希码,大大提升了近似最近邻检索的效率.一方面,哈希方法有效地降低了海量高维数据的存储成本,例如,用32位哈希码可以在16 GB的内存空间中存储5亿张图像;另一方面,数据在海量空间中的距离计算可以借助硬件上的按位异或操作来完成,从而大大加快了检索速度.上述优点使得哈希方法在大规模高维数据的存储和检索任务中具备显著的优势.目前,哈希方法已成为大规模图像检索领域中的一个研究热点^[4-6].

哈希方法包括数据独立(data-independent)方法^[11,22,23]和数据依赖(data-dependent)方法^[24,25]两类,其中数据依赖方法又被称为哈希学习.数据独立方法通常采用随机投影作为哈希函数,而不考虑数据分布及内部结构.以局部敏感哈希(locality sensitive hashing, LSH)^[11]为例,它用一组随机线性映射作为哈希函数,确保原始空间中余弦距离小的数据点以足够高的概率落入相同的哈希桶中. LSH能够为特定度量空间中的相似性保持提供很好的渐近理论保证,在一定程度上保证了检索性能.因此,许多基于LSH的变体^[22,23]被提出,以进一步提高哈希编码的准确性.然而,在实际检索应用中,上述数据独立方法存在局限性.一方面,为了获得良好的检索精度和召回率,这些方法必须使用足够长的哈希码构造多个哈希表,导致其在处理大规模检索任务时面临巨大的存储和计算成本;另一方面,由于LSH的理论保证仅适用于某些简单的距离或相似性度量,无法编码更复杂的语义相似性表示,所以数据独立方法无法有效地利用实际检索应用中提供的标记信息.区别于数据独立方法,哈希学习在生成哈希函数时考虑了数据分布和语义信息,因而能够用更短的码长编码更多依赖于数据和特定于任务的信息,从而加强了检索精度与

存储开销两方面的优势^[19,20].

哈希学习通常借助机器学习技术,学习能够保持原始空间相似性的哈希函数和精简的哈希码.这里的相似性可以由原始特征间的距离度量(如欧氏距离)定义,也可以由监督信息(如类别标签、成对相似度、三元组信息)定义.根据训练过程是否利用了监督信息,哈希学习算法大致可以分为无监督哈希方法^[6,24-65]和监督哈希方法^[4,66-106]两类.其中,无监督哈希方法通常采用特征空间变换,使得原始特征空间中相似样本间的海明距离较近,而原始特征空间中不相似样本间的海明距离较远;监督哈希方法通常在监督信息的指导下学习哈希函数,使得同类样本间的海明距离较近,而异类样本间的海明距离较远.一般来说,监督哈希方法通过借助于更精确的语义信息,能够取得比无监督哈希方法更高的检索准确率^[107].

针对哈希学习的研究现状,已有学者做了综述性研究^[19,20,108-110].其中,文献[20]从学习范式和哈希函数特性等多个角度对哈希学习算法进行分类,并择取经典的哈希学习算法进行介绍;文献[19]将哈希学习算法按相似性保持方式进行了分类整理,且较多地关注了量化方法;文献[108]根据近似最近邻检索的过程将哈希方法分为哈希编码方法和哈希排序方法两类,并分别对两类方法进行了更细化的分类;文献[109]首先介绍了哈希方法的组成部分,然后介绍了经典的无监督哈希方法;文献[110]介绍了深度哈希学习的最新研究进展和发展趋势.然而,上述研究较少探索哈希学习算法各组成部分之间的关联关系,也较少关注每一类目下不同算法之间的演化关系与演变过程.此外,上述研究多以英文的形式梳理了数年前的研究进展.近年来,随着哈希学习算法的不断迭代与推广,哈希学习在技术手段、损失函数、应用场景等方面有了很多新进展,并且面临高效训练、精细化检索、理论分析等方面的新挑战.因此,有必要对哈希学习领域的进展与前景进行更为细致的整理,以促进相关领域的研究工作.具体来说,本文首先从损失函数、优化策略及样本外扩展映射3个角度系统地梳理了哈希学习的技术路线和算法要素;其次回顾了无监督图像哈希方法和监督图像哈希方法的研究现状,并揭示了每一类内代表性哈希学习算法的演化关系;然后展示了部分代表性哈希学习算法的实验结果与分析;最后结合目前哈希学习的不足与目前哈希领域关注的热点问题,展望了哈希学习未来的研究方向.该研究旨在为初学者提供详尽易懂的引导,并为哈希学习算法的研究提供可能的切入点和突破点,进而促进面向大规模图像检索的哈希学习研究,并推动哈希学习在相关研究领域与现实场景中的应用.

1 面向图像检索的哈希学习

哈希学习的主要目标是学习形如 $\mathbf{b} = h(\mathbf{x})$ 的哈希函数,以将原始数据的高维特征 $\mathbf{x} \in \mathbb{R}^D$ 映射至海明空间,得到保相似性的简短二值表示 $\mathbf{b} \in \{0, 1\}^K$,从而实现数据降维并加速检索.值得注意的是,在算法的实施过程中,为了方便计算和阈值化操作,通常用 $\mathbf{b} \in \{-1, 1\}^K$ 替代 $\mathbf{b} \in \{0, 1\}^K$,作为哈希码的优化目标.根据数据模态的不同,哈希学习算法也可以划分为文本哈希、图像哈希、视频哈希等.本文主要以图像作为输入的形式展开.

如图1所示,哈希学习的技术路线包括哈希学习模型的训练和基于哈希的近似最近邻检索两个阶段.其中,哈希学习模型的训练一般需要构建浅层或深度模型,并设计某种哈希学习算法来训练模型.区别于其他机器学习方法,哈希学习需要在训练阶段学习样本外扩展映射(即哈希函数),来为检索阶段的新样本生成哈希码.为了实现保相似性的目标,并保证哈希编码的质量,现有的研究主要着眼于哈希学习算法的设计,包括损失函数和优化策略的设计^[19,110,111],以及样本外扩展映射的学习^[19,36,112].这三者在训练阶段的作用如图1中虚线箭头所示.其中,样本外扩展映射是训练的结果.损失函数反映哈希学习模型的优化目标,尤其是针对特定场景、特定类型数据的相似性保持目标.优化策略是对模型参数进行求解的依据,其中最重要、同样也备受研究者关注的是对哈希码的求解方法^[66,113,114].该方法对于模型训练过程中的误差有着显著的影响,因而也大大影响了样本外扩展映射的质量.在检索阶段,样本外扩展映射能够高效地生成给定图像的哈希码,包括查询样本的哈希码和检索数据库的哈希码,如图1中实线箭头所示.通过计算查询样本的哈希码与检索数据库的哈希码之间的海明距离,即可快速找到其近似最近邻.

接下来,本节从损失函数、优化策略和样本外扩展映射3个方面详细介绍哈希学习的算法要素,以及其中针对大规模图像检索的设计要点.

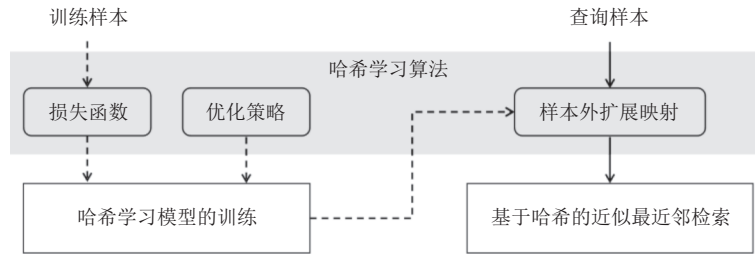


图1 哈希学习的技术路线

1.1 损失函数

哈希学习算法的损失函数通常由惩罚项和约束项两部分组成. 惩罚项是针对输入及采用的具体模型构造的. 根据其物理意义, 惩罚项可分为保相似性惩罚项、量化损失和正则化惩罚项. 与惩罚项不同, 约束项通常针对模型中的各优化变量构造, 以保证哈希函数在数据划分方面的平衡性、简洁性和哈希码的离散性等^[25]. 这些性质的满足有助于提升哈希学习的精度. 最常见的一类约束项是对哈希比特进行约束^[4, 25]. 哈希学习算法的损失函数的一般形式为:

$$\begin{cases} \min_{\mathbf{B}, f(\cdot)} \sum_{i,j} L_s(s_{ij}^o, s_{ij}^h) + L_q(\mathbf{B}, f(\mathbf{X})) + L_r \\ \text{s.t. } \mathbf{B} \in \{-1, 1\}^{K \times N}, \mathbf{B}\mathbf{1}_N = \mathbf{0}_K, \mathbf{B}\mathbf{B}^T = \mathbf{I}_K \end{cases} \quad (1)$$

其中, \mathbf{B} 是哈希码, $f(\cdot)$ 是哈希函数. 保相似性惩罚项 L_s 用于衡量原始空间中相似性 s_{ij}^o 与哈希码相似性 s_{ij}^h 间的差距^[19, 110]. 通过最小化该惩罚项, 可使哈希码尽可能多地保留原始空间中的相似性. 量化损失 L_q 用于衡量将实值哈希表示转化为二值哈希码时的信息损失, 一般用实值哈希表示 $f(\mathbf{X})$ 与二值哈希码 \mathbf{B} 间的量化误差来表示^[66, 115]. 通过最小化该惩罚项, 可使哈希码尽可能多地保留原始特征信息. 正则化惩罚项 L_r 通常表示为模型参数的范数^[36, 66]. 通过最小化该惩罚项, 可达到增强模型泛化能力与鲁棒性的目的. $\mathbf{B} \in \{-1, 1\}^{K \times N}$ 、 $\mathbf{B}\mathbf{1}_N = \mathbf{0}_K$ 、 $\mathbf{B}\mathbf{B}^T = \mathbf{I}_K$ 分别是哈希码的离散约束、比特平衡约束和比特独立约束.

在上述惩罚项和约束项中, 保相似性惩罚项的设计是充分利用原始相似性的关键, 而哈希比特约束的满足有助于得到更精简的哈希码. 因此, 为了更好地适应检索需要, 现有的图像哈希学习算法多聚焦于保相似性惩罚项的设计和哈希比特性质的约束, 并在相似性度量、相似性保持方式以及哈希比特约束 3 方面开展了大量的研究工作^[19, 109, 110]. 本文重点对这 3 方面进行介绍.

(1) 相似性度量是指相似度的定义, 用于判定样本之间的相似程度. 由于哈希学习是将数据从原始空间映射到编码空间, 所以需要选择原始空间中的相似性度量和编码空间中的相似性度量^[19, 110]. 其中, 针对图像检索, 原始空间中的相似性度量可以定义为原始图像特征间的距离度量, 如欧氏距离^[4, 102]、核距离^[6]、不同图像对间的余弦距离^[55]等. 具体的距离度量根据所处理数据集的特点确定. 原始图像特征间的距离越小, 其相似度就越高^[109]. 此外, 原始空间中的相似性度量也可以由训练图像样本包含的额外监督信息定义^[6], 具体形式包括类别标签、成对相似度、三元组信息等^[20]. 由于在图像检索任务中, 样本间的相关性往往由人来判断和定义, 因此, 由监督信息定义的相似性度量通常能保证更高的图像检索精度^[4, 108]. 编码空间中的相似性度量一般基于海明距离定义^[108], 其中一种常见的形式是哈希码内积^[67, 68]. 另外, 编码空间中的相似性度量还可定义为欧氏距离, 这种定义方式多见于一些基于量化的哈希学习^[116, 117]. 除了上述基于已有距离度量的相似性度量外, 还有一些研究工作提出了基于学习的相似性度量, 如对编码空间中的相似性度量进行学习, 以进一步提高检索精度^[5, 118-120].

(2) 相似性保持方式是指如何将原始空间中的相似性嵌入哈希学习过程. 相似性保持方式的选择和定义具体体现在保相似性惩罚项的设计上. 文献^[19]依据相似性保持方式将哈希学习算法分为基于成对相似性保持的方法、基于多重相似性保持的方法、基于隐式相似性保持的方法和基于量化的方法四类. 在该分类框架下, 文献^[110]进一步分化出面向分类的方法这一类别. 不同于上述基于具体技术的分类, 本文从相似性保持目标的角度出发, 将

相似性保持方式归纳为特征空间相似性保持和语义相似性保持两类. 其中, 特征空间相似性保持主要指图结构保持, 语义相似性保持主要包括类别标签嵌入、成对相似度嵌入和三元组关系保持.

图结构保持即让哈希表示保持原始特征之间的距离关系. 通常情况下, 为了实现图结构保持, 首先需要对图像分布信息作一定的假设, 其中最常见的是聚类假设^[33-35]和流形假设^[24,25,54-65,121]. 然后, 基于该假设和原始特征间的距离构造邻接图, 具体的构造形式有全局距离矩阵^[25,39]、 k 近邻图 (k -nearest neighbor graph, k -NNG)^[40,55]、锚点图^[24,41]、局部重构权重矩阵^[43]等. 该邻接图能够反映样本在原始特征空间中的相似性. 最后, 将邻接图以拉普拉斯矩阵或成对相似度矩阵的形式嵌入损失函数.

类别标签嵌入通过对样本哈希表示与类别标签之间的关系进行建模, 达到保相似性的目的^[20,122]. 类别标签嵌入在浅层哈希学习模型中一般以回归损失的形式出现^[66,72-75,115,123], 而在深度哈希学习模型中往往通过对哈希层的分类层构造分类损失 (如交叉熵损失等) 实现^[76-79,81,84].

成对相似度嵌入是哈希学习使用最广泛的相似性保持方式, 其原理在于让原始空间和编码空间中样本对的距离 (或相似性) 关系保持一致, 即为相似的样本对生成相近的哈希码, 为不相似的样本对生成差异较大的哈希码^[19,110]. 成对相似度嵌入最直观的构造形式是谱关系保持 (如 $\text{Tr}(\mathbf{B}\mathbf{L}\mathbf{B}^T)$ 或 $\|\mathbf{K}\mathbf{S} - \mathbf{B}^T\mathbf{B}\|$, \mathbf{K} 为哈希码 \mathbf{B} 的码长, \mathbf{S} 和 \mathbf{L} 分别为成对相似度矩阵和其对应的拉普拉斯矩阵)^[4,67-70]. 另一种常见的构造形式是差异最小化损失, 即最小化原始空间和编码空间中相似性关系之差 $\|\mathbf{S}^o - \mathbf{S}^e\|$ (\mathbf{S}^o 和 \mathbf{S}^e 分别代表原始空间和编码空间中的相似性)^[30-32,124]. 此外, 还有一类基于概率分布一致性的构造方法, 旨在让哈希码拟合语义空间中的联合概率分布, 具体构造形式包括最大化似然损失、最小化交叉熵损失和最小化 KL 散度损失等^[85-94].

三元组关系保持一般可视为排序损失的一种特化^[20], 旨在保持样本三元组在原始空间中基于相似性水平的排序关系. 其中的样本三元组通常由锚点、正样本和负样本组成. 对于单标签数据集, 负样本是与锚点不相似的样本; 而对于多标签数据集, 负样本也可选择与锚点次相似的样本, 以探索样本之间的多级语义结构^[99-103].

为了更充分地挖掘样本在原始空间中的相似性信息, 获取样本对更为准确有效的语义关联关系, 许多哈希学习算法综合采用了上述多种相似性保持方式, 进而有效提升了哈希码的检索精度^[111,125,126].

(3) 哈希比特约束是使得哈希码自身满足一定的特性, 以提高其表达能力. 例如, 比特平衡约束旨在保证各哈希比特的平衡性, 即令每一位哈希码取值为-1 或 1 的概率相等, 以最大化每一比特提供的信息^[4,25]; 比特独立约束旨在减少哈希表示的冗余性^[91]; 稀疏性约束旨在获得更精简的哈希表示^[121,127]; 离散约束保证了哈希码取值的二值属性^[66,115].

目前已有许多研究工作对上述不同要素进行侧重研究, 或探索各要素之间的有机组合, 并产生了许多经典的哈希学习算法, 本文将在第 2、3 节择取其中部分经典算法进行介绍.

1.2 优化策略

通常来讲, 损失函数中涉及每个优化变量对应一个子优化问题. 整体损失函数的优化通过迭代地求解各子优化问题来完成. 其中, 哈希码 \mathbf{B} 对应的子优化问题的定义如公式 (2) 所示^[128]:

$$\begin{cases} \min_{\mathbf{B}} L(\mathbf{B}) \\ \text{s.t. } \mathbf{B} \in \{-1, 1\}^{K \times N}, \mathbf{B}\mathbf{1}_N = \mathbf{0}_K, \mathbf{B}\mathbf{B}^T = \mathbf{I}_K \end{cases} \quad (2)$$

该优化问题通常既包含离散变量 $\mathbf{B} \in \{-1, 1\}^{K \times N}$, 又包含其他连续变量, 因而是一个 NP 难的混合整数规划问题. 为了有效地求解该问题, 需针对离散约束设计合适的优化策略.

根据处理离散约束的方式, 现有哈希学习算法的优化策略可以大致分为实值优化^[4,24,25,40,67]、离散优化^[66,68,113-115,129]和反向传播^[69-71,130-132]这 3 种. 其中, 前两种主要是针对浅层哈希学习模型的优化, 而第 3 种是针对深度哈希学习模型的优化. 各优化策略总结如下.

(1) 实值优化是指首先将离散约束放松为连续约束, 然后求解连续优化问题得到实值解, 最后对求得的实值解进行阈值化处理, 从而得到离散的哈希码. 由于该策略简单易行, 所以广泛地应用于早期的哈希学习算法^[4,24,25]. 然而, 这种策略中涉及的松弛操作通常会带来累积量化误差与局部最优解的问题, 进而降低了哈希码的检索精度.

(2) 离散优化是指在优化过程中保持离散约束. 一个经典的离散优化策略是 Shen 等人^[66]提出的离散循环坐标下降 (discrete cyclic coordinate descent, DCC) 算法. 该方法通过对哈希码进行逐位求解, 有效地减小了量化误差, 但存在耗时和解不稳定的问题. 因此, 快速监督离散哈希 (fast supervised discrete hashing, FSDH)^[115]在其基础上进行了改进, 通过改变哈希码与标签之间映射的回归方向 (即由 $\|\mathbf{Y} - \mathbf{B}\mathbf{W}\|$ 变为 $\|\mathbf{B} - \mathbf{Y}\mathbf{W}\|$, \mathbf{W} 代表哈希码 \mathbf{B} 与标签 \mathbf{Y} 之间的映射), 提高了哈希码求解的速度与稳定性. 另一种常用的离散优化策略是基于离散矩阵分解技术学习哈希码. 例如, 基于列采样的离散监督哈希 (column sampling-based discrete supervised hashing, COSDISH)^[68]利用采样技术降低复杂度, 并将二元二次规划 (binary quadratic program, BQP) 问题转化为与之等价的聚类问题, 从而直接学习离散的哈希码; 哈达玛编码监督离散哈希 (hadamard coded supervised discrete hashing, HC-SDH)^[129]从哈达玛矩阵中采样列, 作为候选哈希码. 除上述离散优化策略外, 还有一些方法对哈希码的离散约束作等价变换, 并设计出更为通用的离散优化框架. 例如, 离散近端线性最小化 (discrete proximal linearized minimization, DPLM)^[113]将离散约束变换为损失函数中的 L1 范数, 并利用近端梯度下降法来求解哈希码; 交替二元矩阵优化 (alternating binary matrix optimization, ABMO)^[114]将离散约束和其他约束 (包括线性约束如 $\mathbf{B}\mathbf{1}_N = \mathbf{0}_K$ 、正交约束如 $\mathbf{B}\mathbf{B}^T = \mathbf{I}_K$) 变换为两个闭集的交集, 进而得到哈希码的解析解. DPLM 和 ABMO 提出的离散优化框架不仅适用于任意形式的损失函数, 而且兼容哈希码的比特平衡约束、比特独立约束.

(3) 反向传播通常采用小批量梯度下降策略来优化模型中的参数^[63,70,104], 其在求解哈希码时所面临的主要挑战是符号函数带来的梯度为零问题和不可导问题. 具体来讲, 为了获得二值哈希码, 通常需使用符号函数对网络的实值输出 \mathbf{c} 进行量化处理. 然而, 由于符号函数对所有非零输入的梯度都为零, 所以神经网络无法使用梯度下降法进行优化. 为了应对这一挑战, 最广泛采用的策略是连续松弛. 其中一种做法是丢弃符号函数, 直接对实值 \mathbf{c} 进行优化; 另一种做法是用光滑激活函数如双曲正切函数 $\tanh(\cdot)$ 或 Sigmoid 函数近似符号函数 $\text{sgn}(\cdot)$, 将输出映射至区间 $[-1, 1]$ ^[130]. 为了避免上述连续松弛策略带来的量化误差问题, 一些深度哈希方法将离散约束替换为损失函数中的一个量化惩罚项 (如 $\|\|\mathbf{c}_i - \mathbf{1}\|_1$), 以使网络输出值 \mathbf{c} 接近 1 或 -1^[69,70]. 这些方法在求解时通常采用平滑近似, 如 $|x| \approx \log(\cosh x)$. 为了进一步减小量化误差, 一些方法将离散优化策略引入深度哈希学习模型. 常见的做法是基于交替方向乘子法 (alternating direction of method of multipliers, ADMM) 的坐标下降, 其用现有的离散优化算法求解哈希码的子优化问题^[71]. 该做法虽然保持了哈希码的离散约束, 但要求各子问题均可解析表达. 此外, 一些基于梯度估计的反向传播策略也被提出^[131,132].

虽然现有的离散优化策略均在一定程度上缓解了由离散约束导致的优化困难问题, 然而, 离散优化算法的收敛性、稳定性与快速性等理论研究仍是值得探索的方向.

1.3 样本外扩展映射

样本外扩展映射旨在为新数据高效地生成哈希码. 为了兼顾数据特性和查询效率, 需要选择合适的哈希函数形式. 现有的哈希学习大都采用单一形式的哈希函数, 通常定义为形如 $\mathbf{b} = h(\mathbf{x}) = \text{sgn}(f(\mathbf{x}))$ 的形式. 其中 $\text{sgn}(z)$ 代表符号函数, 当 $z \geq 0$ 时, $\text{sgn}(z) = 1$; 否则, $\text{sgn}(z) = -1$. $f(\cdot)$ 是预先指定的哈希函数.

具体形式上, $f(\cdot)$ 可以采用线性投影^[4,133]、核函数^[124]、谱函数^[25]、无参数函数^[116,117]和神经网络^[49-51]等. 其中, 线性哈希函数 (如 $f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + t$, \mathbf{w} 代表投影向量, t 代表偏移变量) 由于具有计算高效性, 所以得到广泛的应用. 无参数函数一般基于最近向量分配等方法构造. 几乎所有的线性哈希函数都可以扩展至非线性哈希函数 (如核函数、神经网络等). 一般来说, 非线性哈希函数能够更好地捕获高维图像数据中隐含的复杂结构信息, 从而提升哈希码的检索精度, 但也往往面临着复杂度高、优化困难的问题.

学习策略上, 样本外扩展映射可采用一步哈希策略或两步哈希策略来生成^[111]. 其中, 一步哈希策略是指在训练过程中同时学习哈希函数和哈希码; 而两步哈希策略是在得到训练集哈希码的基础上, 进一步设计损失函数来学习哈希函数. 相较于一步哈希策略, 两步哈希策略在结合不同的损失函数和哈希函数方面更加灵活, 但往往需要更多的训练时间^[86]. 此外, 对于深度哈希学习模型, 采用一步哈希策略通常能获得更优的性能^[134], 其原因可能在于两步哈希策略所采取的分离的学习方式打破了深度模型端到端学习的优势.

尽管目前已有哈希学习综述针对哈希学习算法的上述组成部分展开讨论^[19,109-110],但是,现有研究对各组成之间关系的关注较少.鉴于此,本文将哈希学习算法各个组成之间的关联关系总结如下.

损失函数与样本外扩展映射相互影响.一方面,对于采用一步哈希策略的方法,样本外扩展映射是损失函数的重要组成部分;另一方面,损失函数的降低程度影响着样本外扩展映射的性能.损失函数收敛时,习得的样本外扩展映射通常具有较高且稳定的性能.优化策略影响训练误差,进而影响损失函数的学习效果和样本外扩展映射的性能.一般而言,离散优化策略有助于提升样本外扩展映射的性能,而实值优化策略会导致性能的下降.总体而言,损失函数是哈希学习模型的学习准则,指导模型的优化方向;优化策略影响哈希学习算法的效率和上限;样本外扩展映射的性能是衡量哈希学习算法有效性的重要标准.合适的损失函数、高效的离散优化策略和简捷的样本外扩展映射都有助于哈希学习算法性能的提升.

接下来,本文在第2、3节分别回顾无监督图像哈希方法和监督图像哈希方法,并依照哈希学习算法的组成部分,对不同的哈希学习算法进行分类和介绍.

2 无监督图像哈希方法

本节详细介绍无监督图像哈希方法的研究现状.无监督哈希方法通常仅使用训练样本的特征来进行哈希学习,旨在为原始特征相似的样本生成相近的哈希码,为原始特征差距大的样本生成较远的哈希码.根据样本外扩展映射的不同,无监督哈希方法可以进一步分为浅层无监督哈希方法^[6,16,24-48,121,124]和深度无监督哈希方法^[49-65,135].其中,浅层无监督哈希方法通常采用线性或非线性哈希函数作为样本外扩展映射,而深度无监督哈希方法通常采用神经网络作为样本外扩展映射.下面从哈希学习算法要素的角度出发,回顾每一类别下的部分经典无监督哈希算法,同时阐述其演化关系.

2.1 浅层无监督哈希方法

浅层无监督哈希方法通过对特征空间进行变换实现降维.目前主流的思路有两类:一类未显式地保持样本之间的相似性,而是直接对特征空间进行划分,如一些基于量化的哈希方法^[6,26-29,33-35]和基于矩阵分解的哈希方法^[36-38];另一类则显式地构建了样本间的关系,旨在通过保持样本间距离关系来保持图像流形,如一些基于图结构保持的哈希方法^[24,25].

与大多数学习映射函数的哈希方法不同,量化方法旨在学习一个重构字典,并用该字典对原始高维数据进行编码,从而实现降维^[19].为了构建该字典,现有的基于量化的浅层无监督哈希方法主要从维度选择与最近向量分配两个角度开展探索^[16,33].其中,维度选择的目的在于直接从原始维度中提取出主要的、互不相关的分量(如基于PCA的方法^[6]).最近向量分配则通常先根据距离的远近对原始维度进行分组,再用每个组的中心向量来构建字典并完成编码,其中最常见的是基于聚类的方法^[33].

针对维度选择,典型的策略是基于各投影维度相互独立的假设,将特征空间沿投影维度的方向划分为若干不相交的子集^[16].例如,各向同性哈希(isotropic hashing, IsoH)^[26]通过让划分后各个维度的方差相等,来平衡各个维度包含的信息量.为了同时平衡各维度方差和减小量化误差,迭代量化(iterative quantization, ITQ)^[6]首先利用主成分分析对零均值化的数据进行降维,然后对降维数据施加随机正交旋转来生成二值哈希码. IsoH和ITQ都只能对每个投影维度进行单比特编码.为了进一步减少量化过程中的信息丢失,多比特量化(multi-bit quantization, MBQ)方法^[29,30]将每个投影维度划分为若干子区域,并用一个或多个比特对各子区域进行编码.例如,基于维度分析的量化(dimension analysis based quantization, DAQ)^[29]首先对投影得到的维度进行重要性分析,并选出信息量最大的维度,然后将选定的投影维度划分为若干区域,并分别对每个区域进行量化编码,最后利用量化编码之间的曼哈顿距离来估计相似性.最小重构偏差哈希(minimal reconstruction bias hashing, MRH)^[32]基于投影和量化两个阶段,进一步探索了维度可以灵活调整的哈希码.

针对最近向量分配, k 均值量化方法(k -means hashing, KMH)^[33]基于 k -means聚类学习 k 个中心向量,并用其二值索引来构建字典.在KMH的基础上,乘积量化(product quantization, PQ)^[16]采用半结构化量化,通过增加中心

向量的个数,减小了量化误差,得到了更精确的欧氏距离近似.进一步地,笛卡尔 k 均值 (Cartesian k -means, CKM)^[34] 对 PQ 进行扩展,利用了欧氏距离的旋转不变性,通过引入旋转矩阵,能够在不改变编码间距离的前提下寻找更优的子空间划分.此外,文献 [35] 也采用了最近向量分配的思想,在 LSH^[11] 的基础上学习稀疏表示.

基于量化的哈希方法具有较强的可解释性,但当量化编码较长时,其数据压缩效果和哈希码的检索性能都面临退化.为了解决这一问题,文献 [36-38] 采用矩阵分解技术进行哈希学习,将其原始特征矩阵分解为两个低维矩阵的乘积,并用分解得到的隐向量来近似描述原始矩阵,从而得到了更加精简的哈希码.其中,文献 [36] 将原始特征矩阵分解为基向量矩阵和潜在表示矩阵,并给出了矩阵分解技术在保相似性方面的理论证明;文献 [37] 将哈希码的学习目标构造为非负矩阵分解的形式,进而转化为求解一个可并行计算的乘法更新算法,提升了学习效率;在上述单向矩阵分解方法的基础上,双向离散矩阵分解哈希 (bidirectional discrete matrix factorization hashing, BDMFH)^[38] 进一步设计了逆因子分解模型,交替执行从数据中学习二值哈希码和从二值哈希码中恢复数据两个过程,以使哈希码保持原始图像的内在结构. BDMFH 利用最小化最大值 (minorization-maximum, MM) 算法进行离散优化.

与基于量化或矩阵分解的方法不同,基于图结构保持的浅层无监督哈希方法显式地利用样本间的相似性关系 (或距离关系) 来学习哈希码,以使哈希码保持原始特征空间的局部邻域结构或全局拓扑结构.常用的损失函数有局部邻域保持、正负对保持和细粒度关系保持.其中,正负对保持和细粒度关系保持均由局部邻域保持演化而来,可看作局部邻域保持的特例.

早期的局部邻域保持方法通常以某种方式构造样本的邻接图,并保持其中的邻域结构.例如,谱哈希 (spectral hashing, SH)^[25] 基于加权高斯核构造全局距离矩阵,并将哈希学习过程转化为图割过程.具体而言,SH 对哈希比特约束采用松弛操作,并通过对拉普拉斯矩阵进行特征值分解来求解哈希码. SH 存在若干局限性:首先,基于特征值分解得到的各个哈希比特之间往往存在高阶相关性,不满足比特独立约束;其次,松弛操作带来了量化误差;此外,SH 所做的数据服从均匀分布的假设在许多现实场景中是不成立的.针对 SH 的不足,文献 [39] 提出了改进方案:一方面,在量化损失 $\|\mathbf{F}-\mathbf{B}\|$ 中,通过对实值哈希表示 \mathbf{F} 施加正交旋转 (将 \mathbf{F} 变为 \mathbf{FQ} , 其中 \mathbf{Q} 为旋转矩阵),减少了哈希码 \mathbf{B} 的冗余性;另一方面,通过对哈希码进行离散求解,减小了量化误差.近年来,很多更为灵活的邻接图构造方式被提出.例如,诱导流形哈希 (inductive manifold hashing, IMH)^[40] 在哈希学习中引入了流形学习,基于样本的 k 近邻构造邻接图,并用几种无参数方法对数据进行降维,进而学习能够保持数据潜在流形的哈希函数.考虑到完整 k 近邻图的计算复杂度较高,锚点图哈希 (anchor graph hashing, AGH)^[24] 和离散图哈希 (discrete graph hashing, DGH)^[41] 采用计算更加简洁的锚点图来近似原始数据的邻接图,并结合一个层次化阈值策略学习哈希码.与 IMH 相比,AGH 和 DGH 更适用于大规模图像检索.为了进一步降低图构造过程所带来的时间开销,可扩展图哈希 (scalable graph hashing, SGH)^[42] 用特征表示的变换来近似邻接图,并采用顺序学习策略生成哈希函数.此外,局部线性哈希 (locally linear hashing, LLH)^[43] 首先基于局部线性嵌入 (locally linear embedding, LLE) 学习局部重构权重矩阵,从而捕获图像中隐含的非线性结构,然后通过最小化重构误差和量化损失,学习局部敏感的稀疏哈希编码;文献 [121] 基于数据之间存在稀疏重构关系这一先验信息,通过学习一个稀疏权重矩阵来刻画训练样本的稀疏成对相似性关系.除了通过构造邻接图来保持样本间相似性外,二值重构嵌入 (binary reconstructive embedding, BRE)^[124] 和角度重构嵌入 (angular reconstructive embeddings, ARE)^[31] 提出通过最小化原始特征距离 d_{ij}^o 与相应哈希码的海明距离 d_{ij}^h 之间的误差,达到保相似性的目的.其中, BRE 基于欧氏距离计算 d_{ij}^o ; ARE 基于余弦相似度计算 d_{ij}^o ,并设计了实值优化和离散优化两种优化方案.

上述局部邻域保持方法仅考虑了正相关性,而忽略了负相关性,这种语义偏差可能会削弱哈希码的判别性.为了解决该问题,正负对保持方法在局部邻域保持的基础上,进一步保持了样本间的负相关性.例如,判别无监督图哈希 (discriminative unsupervised graph hashing, DUGH)^[44] 基于概率方法构造相似对和不相似对,即在 k 近邻图上同时构造权值为正的边和权值为负的边;无监督离散哈希 (unsupervised discrete hashing, UDH)^[45] 制定相似对损失和不相似对损失,并将比特平衡约束和比特独立约束转化为惩罚项,采用 DCC 算法进行求解.

进一步地,细粒度关系保持方法旨在保持多层次的局部邻域结构^[46]和样本间的多级排序关系^[47,48].其中,邻域

金字塔保持哈希 (neighbourhood pyramid preserving hashing, NPH)^[46]利用不同长度的邻域列表来表示每个数据点的最近邻, 从而构成不同层次的邻域金字塔, 其中每一层的邻域大小由该层所涉及最近邻的平均距离定义. NPH学习能够同时保持数据点最近邻顺序和邻域平均距离的哈希函数, 并采用实值优化策略和随机梯度下降法来求解. 文献 [47,48] 利用数据点之间的排序信息构造有序图, 并让哈希码保持图中的排序关系.

总体而言, 各类浅层无监督哈希方法各有所长, 互为补充. 其中, 基于量化的哈希方法从学习重构字典的角度出发, 分化出维度选择和最近向量分配方法. 为了让量化编码能够更精确地近似原始空间中特征间的距离, 通常需要学习足够大的码本^[16]. 因此, 维度选择方法经历了从单比特编码到多比特编码的演化^[26,29], 最近向量分配方法也在平衡码本大小与内存成本方面做了不少有意义的探索^[16,34]. 量化编码在极端情况下退化为原始特征, 此时能够得到精确的距离近似, 但失去了降维的意义, 导致图像检索效率大大降低. 相比之下, 基于矩阵分解的哈希方法在划分特征空间时, 能够更灵活地控制编码的维度 (即码长), 但矩阵分解在训练样本较少时往往效果不佳, 并且, 从保相似性的角度看, 矩阵分解得到的低维表示没有很好的可解释性. 基于图结构保持的哈希方法从保持样本间相似性关系 (或距离关系) 的角度出发, 主要针对相似度的建模展开讨论, 并经历了由局部邻域保持到正负对保持、细粒度关系保持的演化. 自适应的码长设置、提高模型及矩阵的语义可解释性、平衡算法精度与计算复杂度都是浅层无监督哈希方法值得探索的演化方向.

2.2 深度无监督哈希方法

由于神经网络具有优异的特征学习能力和训练方式上的优势, 所以深度无监督哈希方法通常能获得比浅层无监督方法更高的精度. 深度无监督哈希方法主要采用特征重构和图结构保持两种思路^[55,135]. 两种思路在方法设计的侧重点上有所不同. 特征重构主要是基于特殊的网络结构, 如生成对抗网络 (generative adversarial network, GAN)^[136]和自编码器; 图结构保持则更多地关注损失函数的设计, 而不依赖于特定的网络结构.

基于特征重构的哈希方法通常利用所学的哈希码重构原始特征, 并采用重构误差最小化作为损失函数, 来提高模型的特征提取能力. 重构误差最小化损失可以表示为 $\|\mathbf{x} - \hat{\mathbf{x}}\|$, 其中 \mathbf{x} 代表原始图像特征, $\hat{\mathbf{x}} = f_d(\mathbf{b}; \theta_d)$ 代表由哈希码 \mathbf{b} 重构的特征. Semantic hashing^[135]是一种典型的特征重构哈希算法, 其利用多个预训练的限制玻尔兹曼机 (restricted boltzmann machines, RBMs)^[137]构造自编码网络来提取样本特征, 并让哈希码重构输入的原始特征. 在 Semantic hashing 自编码思想的基础上, 一些方法进一步考虑了哈希比特约束. 例如, 文献 [49] 关注了哈希码的二值属性, 通过采用离散优化策略求解, 提升了自编码哈希方法的性能. 文献 [50] 对网络结构进行了改进, 借助一个新的无参数哈希编码层来嵌入比特平衡约束. 除了自编码器外, 重构误差最小化与哈希比特约束也广泛地应用于其他深度神经网络中. 例如, HashGAN^[51]利用 GAN 生成比特平衡、比特独立和保一致性的哈希码; DeepBit^[52]通过最小化重构误差保持特征的旋转不变性, 并考虑了比特平衡约束和比特独立约束; 基于二值深度神经网络的无监督哈希 (unsupervised hashing with binary deep neural network, UH-BDNN)^[53]通过最小化重构误差保持语义相似性, 并在网络中的一层基于 DCC^[66]算法直接生成离散的哈希码.

与浅层无监督哈希方法类似, 基于图结构保持的深度无监督哈希方法常从局部邻域保持的角度出发设计损失函数. 例如, 多层潜在结构保持哈希 (multi-layer latent structure preserving hashing, ML-LSPH)^[54]首先利用多层非负矩阵分解对原始数据进行降维, 并基于概率分布一致性构造图结构, 然后通过最小化 KL 散度损失, 学习哈希码的松弛表示. 无监督对抗哈希 (unsupervised adversarial hashing, UADH)^[55]改进了图构造方式, 并用 $\tanh(\cdot)$ 函数近似 $\text{sgn}(\cdot)$ 函数以便优化. 具体来说, UADH 借助 GAN 网络生成特征, 并基于余弦距离提供的局部信息和公共邻居个数提供的全局信息构造 k 近邻图. 然而, 这些方法大多停留在特征层面的局部邻域保持, 而对潜在的高层语义信息的挖掘能力有限.

为了探索潜在语义信息, 近来, 一些伪标签嵌入方法^[56,57]提出将伪标签作为监督信息来训练网络. 其中, 文献 [56] 用 k -means 生成伪标签; 基于伪标签的无监督深度哈希 (unsupervised deep hashing with pseudo labels, UDPH)^[57]用 CCA-ITQ^[6]生成伪标签, 并基于 Sigmoid 函数进行实值优化.

除伪标签嵌入外, 一些伪图结构保持方法也被提出, 其通过构造伪成对相似度来反映语义相似性信息. 相较于

构造伪标签,这种做法的应用更为广泛.例如,基于语义结构的无监督深度哈希 (semantic structure-based unsupervised deep hashing, SSDH)^[58]用两个高斯分布拟合距离直方图,并基于距离计算和阈值划定来构造伪成对相似度.无监督深度 k 均值哈希 (unsupervised deep k -means hashing, UDKH)^[59]在图结构的基础上引入聚类思想,首先用预训练的 VGG-F 模型提取特征,并基于二元 k -means 分解获取簇标签,然后基于簇标签构造伪成对相似度.UDKH 同时提升了聚类的准确性和哈希码的判别性.无监督深度多相似度哈希 (unsupervised deep multi-similarity hashing with semantic structure, UDMSH)^[60]基于两个阈值构造实值细粒度成对相似度.文献 [61] 设计了一个通用的基于分布的度量,首先用余弦相似度和球散度来构造伪成对相似度并对其进行加权,然后利用对比损失增强哈希码的稳定性、平衡性及判别性.文献 [62] 同样采用对比学习捕获语义相似性信息,进而提升哈希码的判别性.

在伪图结构保持的基础上,一些工作进一步提出对伪图结构进行优化和更新.该思想最早由相似度自适应深度哈希 (similarity-adaptive deep hashing, SADH)^[63]提出. SADH 基于 CNN 模型提取深度特征,进而构造拉普拉斯矩阵,并采用 ABMO^[114]的离散优化框架求解哈希码.随后,双瓶颈哈希 (twin-bottleneck hashing, TBH)^[64]引入双瓶颈策略,从重构角度学习哈希码,并基于学到的哈希码对伪成对相似度矩阵进行动态更新.为了进一步提高图结构更新的有效性,深度无监督自演进哈希 (deep unsupervised self-evolutionary hashing, DUSH)^[65]引入课程学习策略,通过在低维空间中由易到难地选择伪相似对来构造伪成对相似度,同时对阈值进行迭代更新.

上述深度无监督哈希方法中,基于特征重构的哈希方法借助网络结构的优势,在保持哈希比特约束、设计离散优化策略等方面进行了诸多探索;基于图结构保持的哈希方法则利用深度神经网络在语义挖掘方面的优势,经历了由局部邻域保持到伪标签嵌入、伪图结构保持,再到图结构更新的演变过程.

通过比较各类无监督哈希方法的特点可以发现,在性能方面,基于量化或图结构保持的方法较低,基于矩阵分解或深度学习的方法相对较高;在复杂度方面,基于量化或矩阵分解的方法相对较低,基于图结构保持或深度学习的方法较高;在可解释性方面,基于量化或图结构保持的方法较为直接,而基于矩阵分解或深度学习的方法可解释性相对较弱.这是由于矩阵分解中隐向量的维度往往基于经验或实验获取,而深度学习模型自身具有较高的复杂度,很难赋予每一层、每个参数具体的意义.本文将无监督图像哈希方法的特点和性能汇总在表 1 中.

表 1 无监督图像哈希方法的特点和性能

无监督哈希方法分类	样本外扩展映射	技术手段	损失函数	性能	复杂度	可解释性	参考文献
浅层无监督哈希方法	线性 非线性	量化	维度选择 最近向量分配	最低	低	强	[6,16,26–30,32] [33–35]
		矩阵分解	矩阵分解	次高	低	弱	[36–38]
		图结构保持	局部邻域保持 正负对保持 细粒度关系保持	次低	高	强	[24,25,31,39–43,121,124] [44,45] [46–48]
		特征重构	重构误差最小化 哈希比特约束	次高			[49–53,135]
深度无监督哈希方法	深度神经网络	图结构保持	局部邻域保持 伪标签嵌入 伪图结构保持 图结构更新	最高	高	弱	[54,55] [56,57] [58–62] [63–65]

总体而言,矩阵分解框架与深度学习框架由于精度上的优势成为近些年无监督图像哈希方法研究的焦点问题.此外,无监督图像哈希方法越来越注重挖掘细粒度的特征空间相似性信息.将投影、量化、图结构保持、矩阵分解、深度神经网络、离散优化等多种技术手段相结合也是无监督图像哈希方法的重要演化方向之一.

尽管无监督哈希方法对于图像数据集的要求较低,适用于标注成本昂贵的现实应用场景,但是,由于其无法利用图像样本的标签信息,所以在模式分类等任务上可能会有一定的精度损失.相比之下,监督哈希方法能够处理更为复杂的语义相似性,并且,其哈希码在识别能力、判别能力和图像检索精度等方面表现更为突出.因此,监督图

像哈希方法在国内外学界和业界获得了更多的关注。

3 监督图像哈希方法

本节详细介绍监督图像哈希方法的研究现状。相较于无监督哈希方法, 监督哈希方法在进行哈希学习时, 更多关注语义相似性保持, 旨在为同类样本生成相近的哈希码, 为异类样本生成较远的哈希码。由于其借助了语义信息的指导, 缩小了底层图像特征与高层语义之间的鸿沟^[138], 因此监督哈希方法能够在诸多现实应用中返回更精确的图像检索结果。根据损失函数中相似性保持方式的不同, 监督哈希方法可进一步分为逐点监督哈希方法^[66,72-84,115,123]、成对监督哈希方法^[4,67-70,85-98]、三元监督哈希方法^[99-103]和多重监督哈希方法^[71,104-106,111]。其中, 逐点监督哈希方法、成对监督哈希方法、三元监督哈希方法分别关注不同层次的语义相似性保持, 而多重监督哈希方法通常采用多种相似性保持方式来克服单一相似性保持方式的局限性, 从而进一步提升检索精度^[111,125-126]。下面以部分经典算法为例对每一类监督哈希方法进行回顾, 并分析其中的演化关系。

3.1 逐点监督哈希方法

逐点监督哈希方法通常将类别标签嵌入作为相似性保持方式。根据样本外扩展映射的不同, 逐点监督哈希方法可进一步分为浅层逐点监督哈希方法和深度逐点监督哈希方法。

浅层哈希方法一般借助哈希码与标签之间的回归实现逐点监督。例如, 文献^[139]基于实例级语义标签设计哈希函数, 并用于姿态估计; 文献^[140]用样本属性实现逐点监督。监督离散哈希 (supervised discrete hashing, SDH)^[66]将哈希码与类别标签之间的关系建模为线性分类器, 并采用 DCC 算法进行离散优化。基于 SDH, Gui 等人^[115]通过将回归方向改为从标签到哈希码, 大大提高了优化效率。自然监督哈希 (natural supervised hashing, NSH)^[72]把类别标签视为二值哈希码, 并基于内积空间保持变换获取近似的标签和哈希码。在 SDH 框架的基础上, 许多方法针对标签回归做了改进, 以提升检索性能。例如, 松弛监督离散哈希 (supervised discrete hashing with relaxation, SDHR)^[73]和松弛局部保持监督离散哈希 (relaxed locality preserving supervised discrete hashing, RLPSDH)^[141]对类别标签进行重新编码, 通过加大类间距, 提高了哈希学习的灵活性和判别性; 鲁棒旋转监督离散哈希 (robust rotated supervised discrete hashing, R^2 SDH)^[74]通过对类别标签施加随机正交旋转来实现鲁棒回归, 从而得到稳定的哈希码; 双向线性回归监督离散哈希 (supervised discrete hashing with mutual linear regression, SDHMLR)^[123]基于两个回归方向学习标签与哈希码间的映射。为了适应半监督场景, Zhang 等人^[75]在标签回归的基础上引入协同过滤, 通过保留原始标注信息隐含的语义关联, 在一定程度上削弱了噪声的影响。

深度哈希方法一般用分类损失实现逐点监督。例如, 监督语义保持深度哈希 (supervised semantics-preserving deep hashing, SSDH)^[76]假设类别标签由若干标记为选中或未选中的潜在属性控制, 并依赖这些属性构造分类误差; 角度深度监督矢量量化 (angular deep supervised vector quantization, ADSVQ)^[77]将 Softmax 分类视作具有角度决策边界的矢量量化, 并学习判别性特征表示和码本; Cui 等人^[78]提出对训练样本进行逐点语义提纯。许多深度哈希方法聚焦于网络结构的设计, 旨在寻找更优的图像特征或图像语义表示, 进而提升逐点监督的质量。例如, 归纳结构一致性哈希 (inductive structure consistent hashing, ISCH)^[79]用一个多层堆叠类别编码器构造更为灵活的归纳语义空间, 并基于语义字典学习模型弥合视觉特征和语义表示之间的鸿沟; 卷积神经网络哈希 (convolutional neural network-based hashing, CNNH)^[142]利用卷积神经网络提取特征, 通过最小化成对相似度矩阵与哈希码内积之间的误差学习哈希函数。

为了得到更加适用于图像检索的哈希码, 逐点监督哈希方法常与特征空间相似性保持结合^[80], 以同时保持语义信息和图像流形。例如, 图神经网络哈希 (graph convolutional network-based hashing, GCNH)^[81]基于图神经网络保持特征空间中的锚点图结构, 并对网络输出进行逐点监督。语义指导哈希 (semantic guided hashing, SGH)^[82]在一个统一的框架中集成了图像流形保持、语义挖掘和哈希学习, 其中, 图像流形保持通过自适应的图结构学习实现, 语义挖掘通过二元矩阵分解实现。

在许多大规模图像检索场景中, 模型往往需要识别新类别的图像。针对这类应用场景, 近年来, 一些研究者尝

试将样本属性作为逐点监督信息,并设计了一系列零样本图像哈希方法^[83,84].这些方法旨在让哈希学习模型利用属性信息进行推理和知识迁移,从而既能识别已知类别的图像,也能识别训练过程中未见类别的图像.

简言之,逐点监督哈希方法从构建哈希码 \mathbf{B} 与标签 \mathbf{Y} 之间的映射 \mathbf{W} 的角度出发,进行了一系列探索,并得到以下经验:通过改变回归方向,将 $\|\mathbf{Y}-\mathbf{B}\mathbf{W}\|$ 变为 $\|\mathbf{B}-\mathbf{Y}\mathbf{W}\|$,可以提升优化效率,并得到更加稳定的哈希码^[115];双向回归的利用可以进一步提升检索精度^[123];在半监督图像检索场景中,通过对标签 \mathbf{Y} 进行过滤,可以有效地降噪^[75].逐点监督哈希方法的一个重要演化方向是对更优的图像语义表示的探索.一般而言,通过引入优化变量 $\hat{\mathbf{Y}}$ 来替换给定的标签 \mathbf{Y} ^[73,141]、对标签施加正交旋转(将 \mathbf{Y} 变为 $\mathbf{Y}\mathbf{R}$,其中 \mathbf{R} 为旋转矩阵)^[74]等方式,可以拉大类间距离,并得到更具判别性的图像哈希码.除此之外,对于深度模型,还可以借助多种网络结构来生成新的语义表示^[79].近年来,自适应的标签表示也得到了探索^[143],使得自适应的逐点监督哈希学习成为可能.

3.2 成对监督哈希方法

成对监督哈希方法通常将成对相似度嵌入作为相似性保持方式.不同于关注单个样本分类性能的逐点监督哈希方法,成对监督哈希方法主要关注样本二元组的语义相似关系保持.根据检索阶段对查询样本和检索数据库的编码方式是否相同,现有的成对监督哈希方法可大致分为对称成对监督哈希方法和非对称成对监督哈希方法.其中,对于对称成对监督哈希方法,查询样本和检索数据库的哈希码通常都采用样本外扩展映射生成;对于非对称成对监督哈希方法,只有查询样本的哈希码采用样本外扩展映射生成,而检索数据库的哈希码通常在训练阶段直接学习得到^[144].

谱关系保持损失 $\|\mathbf{K}\mathbf{S}-\mathbf{B}^T\mathbf{B}\|^2$ 是对称成对监督哈希方法的基础损失项之一^[4,67,68,85-93],其物理意义是用 K 维哈希码的内积 $\mathbf{B}^T\mathbf{B}$ 近似成对相似度 \mathbf{S} .在基于核的监督哈希(kernel-based supervised hashing, KSH)^[67]中,核函数被用作哈希函数,相应地, $\mathbf{B}^T\mathbf{B}$ 被构造为核化特征的内积. KSH 借助按位优化策略逐位生成哈希码.深度成对监督哈希(deep pairwise supervised hashing, DPSH)^[89]在一个参数共享的孪生网络框架下,用最大似然估计求解谱关系保持损失.监督多层次深度哈希(supervised hierarchical deep hashing, SHDH)^[92]采用分层数据标记方式,通过对网络的每一层进行加权,计算成对相似度.深度语义重构哈希(deep semantic reconstruction hashing, DSRH)^[93]结合了量化损失与成对相似度嵌入,其中成对相似度由语义标签定义.谱关系保持损失广泛地用于半监督图像哈希方法中.例如,半监督哈希(semi-supervised hashing, SSH)^[4]设计了一个通用的哈希学习框架,其中,谱关系保持损失被用于最小化标记数据上的经验误差.为了同时处理有监督数据和无监督数据,SSH 利用全体数据构造正则化惩罚项,并结合了特征空间相似性保持和语义相似性保持.在 SSH 的基础上,基于 Bootstrap 顺序投影学习的半监督非线性哈希(semi-supervised nonlinear hashing using bootstrap sequential projection learning, Bootstrap-NSPLH)^[85]采用非线性哈希函数和 Bootstrap 顺序学习策略生成哈希函数.两步哈希(two-step hashing, TSH)^[86]基于谱关系保持损失学习哈希码,并提出了灵活的两步哈希策略.随后出现的哈希学习算法广泛地采用了这种两步哈希策略.例如,快速监督哈希(fast supervised hashing, FastHash)^[87]应用非线性哈希函数,首先基于块搜索学习哈希码,然后训练一个决策树模型作为样本外扩展映射.与基于核函数的方法相比, FastHash 可以更高效地应用于复杂高维的图像特征和大规模图像数据.为了平衡相似对和不相似对,差异最小化深度哈希(discrepancy minimizing deep hashing, DMDH)^[88]引入一个放缩参数对成对相似度进行加权,同时最小化实值映射矩阵与二值映射矩阵间的误差.基于谱关系保持损失,国内外研究者还提出了一系列离散监督哈希方法.例如, COSDISH^[68]和深度离散监督哈希(deep discrete supervised hashing, DDSH)^[90]利用列采样技术对谱关系保持损失进行离散优化.进一步地,强约束离散哈希(strongly constrained discrete hashing, SCDH)^[91]将比特平衡约束与比特独立约束集成到一个辅助变量中,学习能够保相似性和哈希比特约束的哈希码.

对称成对监督哈希方法常用的另一类损失函数是概率分布一致性损失^[69,70,94],其通常令海明空间中的联合概率分布与语义空间中的联合概率分布(即成对相似度)相一致.例如,基于潜在因子模型的监督哈希(supervised hashing with latent factor models, LFH)^[94]基于潜在变量模型和对数最大后验概率学习保相似性的哈希码.为了缩短参数学习时间,并适应大规模数据,LFH 采用随机采样技术进行优化.深度哈希网络(deep hashing network,

DHN)^[69]通过最小化量化损失和成对交叉熵损失, 学习卷积神经网络的参数, 并用训练得到的网络为查询样本生成哈希码. 深度柯西哈希 (deep cauchy hashing, DCH)^[70]基于柯西分布构造成对交叉熵损失和量化损失, 并制定了一个新的激活函数来提升检索效率和精度.

在大规模图像检索应用中, 成对监督哈希方法由于涉及成对相似度矩阵的构建, 所以往往受限于 $O(n^2)$ 级别的训练时间复杂度. 此外, 在检索阶段, 上述方法采用的对称策略无法充分利用训练数据所提供的监督信息. 因此, 非对称离散图哈希 (asymmetric discrete graph hashing, ADGH)^[95]和非对称深度监督哈希 (asymmetric deep supervised hashing, ADSH)^[96]设计了一种非对称的成对相似度保持策略, 一方面, 利用检索集与数据规模相对较小的查询集之间的成对相似度来降低训练时间复杂度; 另一方面, 在训练阶段同时学习样本外扩展映射和检索数据库的哈希码, 仅将习得的样本外扩展映射用于生成查询样本的哈希码, 从而更充分地保留了训练数据的相似性信息. 深度解耦离散哈希 (deep uncoupled discrete hashing, DUDH)^[144]进一步从检索数据库中采样一个较小的图像集, 并用该图像集构造相似度迁移矩阵, 从而将查询样本与检索数据库解耦, 大大降低了优化成本.

上述对称成对监督哈希方法和非对称成对监督哈希方法仅关注了样本间的相似性关系. 近年来, 随着网络与社交媒体的发展, 多标签图像已不再罕见. 相应地, 一些成对监督哈希方法在保持样本间相关性的基础上, 进一步探索了类别标签之间的语义关系. 例如, 基于标签的弱监督模态协作哈希网络 (tag-based weakly-supervised modally cooperative hashing network, TelecomNet)^[97]提出了一种从标签中挖掘语义词的稀疏编码方案. 该方法用获取的语义词来定义语义向量空间, 从而缓解了标签噪声和标签模糊问题. Shen 等人^[98]基于概率方法实现相似性保持, 并制定了样本特征间的成对损失、标签间的成对损失以及样本特征与标签间的成对损失.

简言之, 成对监督哈希方法从保持样本对的语义相似性的角度出发, 在针对特定场景的成对相似度矩阵构造、损失函数设计、离散优化、降低成对监督自身复杂度等方面进行了探索. 其中, 一个重要的演化方向是降低成对相似度矩阵带来的计算复杂度. 一般而言, 利用采样等技术可以有效地提升计算效率^[68,144], 但同时也可能带来一定的精度损失. 因此, 有必要在算法的精度与效率之间进行折衷, 并挖掘相对重要的样本来提高采样的质量. 此外, 尽管成对相似度矩阵包含丰富的语义相似性信息, 但是在检索阶段构造该矩阵并不现实, 导致训练阶段和检索阶段可能存在语义信息上的差距. 如何弥补这一差距也是近来成对监督哈希方法研究的一个焦点.

3.3 三元监督哈希方法

三元监督哈希方法通常将三元组关系保持作为相似性保持方式, 构建基于三元组排序关系的哈希学习模型. 由于三元组的构造包含了度量学习的思想, 且每一个三元组都可隐式地分解得到两个成对标签, 所以相较于逐点监督哈希方法和成对监督哈希方法, 三元监督哈希方法往往能保持更丰富且更具判别性的相似性信息.

典型的三元监督哈希方法如网中网哈希 (network in network hashing, NINH)^[99]采用排序损失来指导卷积神经网络的参数学习, 并利用一个分离编码结构来减少哈希码的冗余性. 文献 [100] 引入了相对相似度的概念, 显式地在正负样本对之间设置相对距离阈值. 深度三元量化 (deep triplet quantization, DTQ)^[101]设计了一个三元组选择模块, 并采用弱正交约束控制深度特征的量化. 概率顺序保持语义哈希 (probability ordinal-preserving semantic hashing, POSH)^[102]基于无参数贝叶斯理论提出了顺序保持哈希损失, 并将概率顺序相似性保持损失、概率量化损失以及概率语义保持损失融合在一个统一的哈希学习框架中. 深度列表三元哈希 (deep listwise triplet hashing, DLTH)^[103]首先利用知识提炼模块得到能够反映类别间相对相似度的类标向量, 然后基于这些类标向量构造三元组, 最后在深度哈希模块中用列表三元组损失来实现三元监督. 离散语义排序哈希 (discrete semantic ranking hashing, DSeRH)^[145]利用原始特征间的距离关系及语义标签构造三元组, 并将三元组包含的语义排序信息嵌入哈希码. DSeRH 同时支持浅层哈希学习模型和深度哈希学习模型.

总体而言, 三元监督哈希方法从保持排序关系的角度出发, 在三元组构造、三元组损失设计等方面进行了探索. 深度三元监督哈希方法还重点关注了网络结构的设计. 然而, 三元监督哈希方法的精度严重依赖于构造的三元组. 如果三元组中正样本和负样本的差别过大, 那么模型几乎不更新; 而如果三元组中正样本和负样本的差别过小, 那么模型极易拟合噪声. 因此, 更有效的三元组构造方式仍然需要进一步的探索.

3.4 多重监督哈希方法

基于上述 3 种基本的监督哈希方法, 多重监督哈希方法旨在嵌入更丰富的多级语义相似性关系. 例如, 深度监督哈希 (deep supervised hashing, DSH)^[104]用对比损失实现成对相似度嵌入, 用三元组排序损失实现三元组关系保持, 并结合量化误差最小化学习网络参数和哈希表示. 深度监督离散哈希 (deep supervised discrete hashing, DSDH)^[71]在深度哈希学习模型中结合了哈希码的二分类与成对相似度保持, 并采用 DCC 算法生成离散的哈希码. 快速可扩展监督哈希 (fast scalable supervised hashing, FSSH)^[111]和深度锚点图哈希 (deep anchor graph hashing, DAGH)^[105]在学习哈希码时同时使用了类别标签嵌入和成对相似度嵌入. Li 等人^[106]设计了一个通用的深度监督离散哈希框架, 可同时实现逐点监督、成对监督或者三元监督.

总体而言, 在性能方面, 成对监督哈希方法和三元监督哈希方法较低, 逐点监督哈希方法和多重监督哈希方法较高, 这是由于逐点监督哈希方法保留了更具判别性的信息, 多重监督哈希方法保持了更多的相似性信息; 在复杂度方面, 逐点监督哈希方法较低, 成对监督哈希方法和三元监督哈希方法由于在优化过程中涉及成对相似度矩阵和三元组, 所以复杂度较高; 在可解释性方面, 各类监督哈希方法都具备较好的可解释性, 从图像样本间相似性保持的角度来看, 成对监督哈希方法最为直观, 三元监督哈希方法次之, 而逐点监督哈希方法和多重监督哈希方法略有下降. 表 2 概要地总结了监督图像哈希方法的特点和性能.

表 2 监督图像哈希方法的特点和性能

监督哈希方法分类	技术手段	损失函数	相似性保持方式	性能	复杂度	可解释性	参考文献
逐点监督哈希方法	浅层	回归损失	类别标签嵌入	次高	低	强	[66,72–75,80,82,83,115,123,139–141]
	深度	分类损失					[76–79,81,84,142]
成对监督哈希方法	对称	谱关系保持损失 概率一致性损失	成对相似度嵌入	低	高	最强	[4,67,68,85–93]
	非对称	非对称成对损失					[69,70,94]
	多标签	标签语义挖掘					[95,96,144]
三元监督哈希方法	三元组关系保持	排序损失	三元组关系保持	次低	高	次强	[99–103,145]
多重监督哈希方法	多种相似性保持	图结构保持 标签嵌入 成对损失 排序损失	多种相似性保持	最高	视具体情况而定	强	[71,104–106,111]

无监督图像哈希方法所采用的量化模型、矩阵分解模型及深度学习模型等, 均可推广到监督图像哈希方法, 并且, 各类监督信息的嵌入均能够显著地改善图像检索性能. 弱监督场景下的语义挖掘^[97]、针对具体应用场景的监督图像哈希方法^[146]、多种相似性保持方式的融合^[104]、模型的效率及鲁棒性提升^[74,144]等都在近些年来得到了有意义的探索, 且是监督图像哈希方法的重要演化方向.

4 实验

本节归纳总结了哈希学习算法评估通用的部分图像数据集与评估指标, 并通过对比实验结果分析和比较了各方法的性能.

4.1 图像数据集

在学习给定图像的哈希码之前, 一般需要提取该图像的判别性特征, 例如 GIST 特征^[147], SIFT 特征^[148], LBP 特征^[149], Fisher 特征^[150], VLAD 特征^[151], CNN^[152]和视觉 Transformer^[153]等. 在图像检索领域有许多广泛使用的基准数据集, 包括单标签数据集和多标签数据集. 这些图像数据集有不同的规模, 对应不同的检索任务. 本文简单介

绍如下几个通用的图像数据集.

(1) MNIST^[154]. MNIST 是一个包含 70 000 张图片的单标签手写数字数据集, 每张图片像素大小是 28×28, 并用从 0–9 的数字标签进行标记, 共 10 类.

(2) CALTECH^[155,156]. CALTECH-101^[155] 是一个包含 101 个类别、9 146 张图片的单标签数据集, 每张图片的像素大小大约是 28×28, 每个类别由 40 至 800 张图片组成. CALTECH-256^[156] 是一个包含 256 个类别、30 607 张图片的单标签数据集, 每个类别超过 80 张图片.

(3) CIFAR^[157]. CIFAR-10^[157] 是一个包含 10 个类别、60 000 张小型彩色图片的单标签数据集, 每张图片的像素大小约为 32×32, 每类图片数为 6 000. 这 10 个类别分别代表飞机、汽车、鸟、猫、鹿、狗、青蛙、马、船和卡车. CIFAR-100^[157] 是一个包含 20 个超类、100 个小类、60 000 张图片的单标签数据集, 每类图片的数目是 600.

(4) MS-COCO^[158]. MS-COCO 是一个包含 91 个类别、82 081 张图片的多标签数据集, 其内容主要取自复杂的日常场景. 平均每张图片包含 3.5 个类别和 7.7 个实例目标, 仅包含一个类别的图片占比不到 20%, 仅包含一个实例目标的图片占比 10%.

(5) NUS-WIDE^[159]. NUS-WIDE 是一个包含 1 000 个类别、269 648 张图片的大规模多标签数据集, 这些图片来自 Flickr 网站, 内容侧重于真实世界中人们的日常生活和事件. 图像检索任务中通常选择 21 个包含样本数最多的类别(如动物、人类、建筑物等)对应的包含 195 834 张图片的子集.

(6) ImageNet^[160]. ImageNet 是一个用于大规模视觉识别挑战赛 (ILSVRC 2015) 的单标签图像数据集. 该数据集包含 1 000 个手工标注类别、超过 120 万张图片, 每张图片的像素大小大约为 500×400. 图像检索任务中通常会使用由文献 [130] 整理的包含 100 个类别、超过 14 万张图片的子集.

(7) Tiny Images^[161]. Tiny Images 数据集包含 75 062 个类别、7 900 万张图片, 每张图片的像素大小为 32×32.

4.2 评估指标

哈希学习算法的性能一般要从准确率和效率两个角度来衡量. 其中, 用于检验准确率的评估指标包含海明排名和哈希查找两类. 海明排名类指标侧重于整个检索结果列表的准确率, 而哈希查找类指标仅关注检索结果中排名靠前的项. 令 N 表示检索到的样本总数, N_r 表示检索到的查询样本的真实邻居个数, N_a 表示数据集中查询样本的邻居总数. 本文介绍如下几个评估指标.

(1) 精度 (*Precision*) 是指利用哈希学习算法检索得到的样本中, 查询样本的真实邻居所占的比例.

$$Precision = \frac{N_r}{N} \quad (3)$$

(2) 召回率 (*Recall*) 是指利用哈希学习算法检索出的真实邻居数占数据集中所有邻居数的比例, 用于评估哈希学习算法检索的全面性.

$$Recall = \frac{N_r}{N_a} \quad (4)$$

在评估哈希模型优劣时, 通常希望 *Precision* 和 *Recall* 的值都是越高越好, 但两者间往往存在此消彼长的关系, 即当 *Precision* 高时 *Recall* 可能较低, 当 *Recall* 高时 *Precision* 可能较低.

(3) F 值 (F_β) 可以综合 *Precision* 和 *Recall* 的结果, 对两者的关系进行加权调和.

$$F_\beta = \frac{(1+\beta^2) \times Precision \times Recall}{\beta^2 \times (Precision + Recall)} \quad (5)$$

当 $\beta > 1$ 时, 说明当前的检索任务更关注 *Recall*; 当 $\beta < 1$ 时, 说明当前的检索任务更重视 *Precision*; 当 $\beta = 1$ 时, 得到 $F1$ 值, *Precision* 与 *Recall* 同等重要, 需同时考虑. $F1$ 值越高, 表明哈希模型越理想.

(4) 平均精度均值 (mean average precision, *mAP*) 属于海明排名评价, 用于评估哈希学习算法的平均检索性能, *mAP* 越大, 表明方法的性能越好. *mAP* 是所有查询样本 $Q = \{q_1, \dots, q_m\}$ 的 *AP* 的均值.

$$mAP = \frac{\sum_{q \in Q} AP_q}{m} \quad (6)$$

平均精度 (average precision, AP) 表示哈希学习算法对查询样本 q 返回的检索结果的 $Precision$ 均值:

$$AP_q = \frac{\sum_{i=1}^k Precision_i \times \sigma_i}{N} \quad (7)$$

其中, $Precision_i$ 表示检索结果中前 i 个样本的 $Precision$, σ_i 指示第 i 个返回结果是否与 q 相关, 相关则 $\sigma_i=1$, 不相关则 $\sigma_i=0$.

(5) 海明半径 2 以内的精度 ($Precision@H \leq 2$) 属于哈希查找评价, 指以查询样本为球心、半径为 2 的海明球内的检索精度.

(6) 海明半径 2 以内的平均精度均值 ($mAP@H \leq 2$) 也属于哈希查找评价, 指与查询样本的海明距离不超过 2 的检索结果样本的 mAP .

(7) 归一化累计收益折扣 (normalized discounted cumulative gain, $NDCG$) 一般用于衡量方法的排序效果, 适用于语义排序的情况. 给定查询样本 q 及其包含 p 个样本的排序结果列表, $NDCG$ 可以计算为:

$$NDCG@p = \frac{1}{Z} \sum_{i=1}^p \frac{2^{s_i} - 1}{\log(i + 1)} \quad (8)$$

其中, s_i 为排序结果列表中第 i 个样本与查询样本 q 的相似程度 (如公共标签个数), Z 为归一化因子.

上述评估指标侧重于评估哈希学习算法的 ANN 检索精度, 而未考虑哈希学习算法的效率. 这里效率包括空间成本和时间效率. 其中, 空间成本仅取决于哈希码的码长, 时间效率包括训练效率和检索效率.

(8) 训练时间用于评估训练效率, 通常取决于损失函数及优化策略的设计; 查询的平均检索时间用于衡量检索效率, 通常取决于学到的样本外扩展映射.

(9) 常用的基本统计指标

均值是指所有样本之和再除以样本总个数, 用于描述样本集的中心点, 可以在一定程度上反映算法的平均性能.

标准差即方差的算术平方根, 它描述的是样本集中各样本与均值之间的平均距离. 标准差可以反映样本集的离散程度, 标准差越大, 代表离散程度越大.

极差即样本集中最大值与最小值之差距, 体现了样本集中样本的波动范围. 与标准差相似, 极差也可以反映样本集的离散程度. 标准差和极差都可以在一定程度上反映算法的稳定性.

4.3 对比实验结果与分析

本文选取了 4 个代表性的无监督哈希学习算法 SH^[25]、PCA-ITQ^[6]、PCA-RR^[6]、MFH^[36]和 13 个代表性的监督哈希学习算法 NSH^[72]、FSSH^[111]、SSLH^[162]、SDHMLR^[123]、SCDH^[91]、RSLH^[163]、DSH^[104]、DPSH^[89]、DSDH^[71]、DCH^[70]、DDSH^[90]、ADSH^[96]和 DMRH^[164], 在基准图像数据集 CALTECH-101^[155]、CIFAR-10^[157]、ImageNet^[160]、NUS-WIDE^[159]和 MS-COCO^[158]上进行了对比实验. 其中, DSH、DPSH、DSDH、DCH、DDSH、ADSH 和 DMRH 为深度算法, 其他算法为浅层算法. 通常情况下, 深度哈希学习算法由于采用了端到端的学习方式, 所以其性能普遍远优于浅层哈希学习算法. 鉴于此, 我们分别对浅层算法和深度算法开展实验. 所有浅层哈希学习算法在一个配置为 Intel Core i9-12900K 3.19 GHz 4 处理器, 64 GB RAM 的工作站上进行实验, 软件环境为 Windows 10 和 Matlab R2020a. 所有深度哈希学习算法在 4 TITAN X GPU 上进行实验, 软件环境为 Linux 和 PyTorch. 各算法的超参数设置参考了方法原文, 并在此基础上进行了微调. 对比实验的结果如表 3–表 5 和图 2 所示. 其中涉及指标 mAP 、 $mAP@H \leq 2$ 、 $Precision$ 在本文第 4.2 节有详细介绍说明.

通过分析实验数据和综合各算法的整体性能, 我们总结出以下经验, 其中大多已被现有的工作讨论过.

(1) 监督信息越多, 检索性能越好. 一般来说, 监督方法的性能优于无监督方法. 同为监督方法, 通过引入更多的监督信息可以进一步提升检索精度. 各算法在海明排名和哈希查找指标下的结果均可验证这一点. 例如, 观察表 3 和表 4 可知, FSSH 的结果在多数情况下好于 NSH, 可见成对相似度信息的嵌入能够提升分类模型的精度; 又如, NSH 的结果始终好于无监督哈希方法, 说明类别标签信息的嵌入同样有助于进一步改善性能.

表 3 3 个基准数据集上各算法的 mAP 结果

Method	CALTECH-101				CIFAR-10				ImageNet			
	8 bits	10 bits	12 bits	14 bits	4 bits	6 bits	8 bits	10 bits	8 bits	10 bits	12 bits	14 bits
SH ^[25]	0.2043	0.2053	0.2450	0.2744	0.2564	0.2801	0.2772	0.2838	0.0316	0.0338	0.0357	0.0424
PCA-ITQ ^[6]	0.0535	0.0715	0.0920	0.1001	0.1589	0.2117	0.2286	0.2658	0.0266	0.0269	0.0271	0.0276
PCA-RR ^[6]	0.0999	0.1184	0.1138	0.1295	0.1696	0.2302	0.2796	0.2674	0.0328	0.0331	0.0331	0.0334
MFH ^[36]	0.2030	0.2208	0.2395	0.2538	0.2434	0.2549	0.2561	0.2702	0.0314	0.0342	0.0368	0.0388
NSH ^[72]	0.3516	0.3887	0.4129	0.4323	0.3449	0.4842	0.5395	0.5798	0.1065	0.1290	0.1514	0.1752
FSSH ^[111]	0.3502	0.3915	0.4287	0.4475	0.4015	0.5176	0.5708	0.6088	0.0926	0.1307	0.1545	0.1792
SSLH ^[162]	0.3695	0.4078	0.4330	0.4458	0.4356	0.5215	0.5722	0.6058	0.1015	0.1345	0.1592	0.1813
SDHMLR ^[123]	0.3584	0.3859	0.4129	0.4116	0.3417	0.4848	0.5875	0.6027	0.1032	0.1222	0.1627	0.1736
SCDH ^[91]	0.3833	0.4080	0.4361	0.4505	0.4324	0.5693	0.6127	0.6295	0.0985	0.1347	0.1627	0.1891
RSLH ^[163]	0.4263	0.4469	0.4712	0.4837	0.4688	0.5934	0.6163	0.6473	0.1358	0.1749	0.2089	0.2330

表 4 3 个基准数据集上各算法的 $mAP@H \leq 2$ 结果

Method	CALTECH-101				CIFAR-10				ImageNet			
	8 bits	10 bits	12 bits	14 bits	4 bits	6 bits	8 bits	10 bits	8 bits	10 bits	12 bits	14 bits
SH ^[25]	0.0915	0.1130	0.1204	0.1254	0.1958	0.2177	0.2449	0.2670	0.0239	0.0289	0.0345	0.0501
PCA-ITQ ^[6]	0.0405	0.0408	0.0418	0.0453	0.2268	0.2331	0.2359	0.2462	0.0129	0.0129	0.0130	0.0131
PCA-RR ^[6]	0.0572	0.0700	0.0612	0.0745	0.2326	0.2143	0.2149	0.2191	0.0134	0.0162	0.0156	0.0177
MFH ^[36]	0.1169	0.1257	0.1390	0.1568	0.1905	0.2136	0.2375	0.2573	0.0206	0.0253	0.0322	0.0397
NSH ^[72]	0.2924	0.3140	0.3234	0.3480	0.3385	0.4045	0.4956	0.5540	0.0776	0.1131	0.1534	0.1964
FSSH ^[111]	0.3137	0.3704	0.4024	0.4350	0.3801	0.4626	0.5367	0.5892	0.0755	0.1173	0.1621	0.2013
SSLH ^[162]	0.3051	0.3339	0.3530	0.3706	0.3877	0.4772	0.5515	0.5869	0.0853	0.1190	0.1629	0.2086
SDHMLR ^[123]	0.3298	0.3628	0.3891	0.4211	0.3821	0.4463	0.5414	0.6028	0.0686	0.0971	0.1424	0.1885
SCDH ^[91]	0.3521	0.3833	0.4171	0.4446	0.4014	0.4945	0.5749	0.6280	0.0801	0.1234	0.1709	0.2203
RSLH ^[163]	0.3807	0.4242	0.4560	0.4897	0.4132	0.5196	0.6002	0.6370	0.1121	0.1524	0.2063	0.2207

表 5 3 个基准数据集上各深度哈希算法的 mAP 结果

Method	CIFAR-10				NUS-WIDE				MS-COCO			
	24 bits	48 bits	64 bits	128 bits	24 bits	48 bits	64 bits	128 bits	24 bits	48 bits	64 bits	128 bits
DSH ^[104]	0.7864	0.7830	0.7834	0.7835	0.6598	0.6653	0.6587	0.6598	0.5135	0.5069	0.5147	0.5072
DPSH ^[89]	0.8821	0.8853	0.8858	0.8876	0.8390	0.8429	0.8423	0.8468	0.6623	0.6871	0.6965	0.7073
DSDH ^[71]	0.8985	0.9004	0.9002	0.8970	0.8225	0.8328	0.8347	0.8415	0.6988	0.7191	0.7220	0.7227
DCH ^[70]	0.8753	0.8752	0.8749	0.8273	0.7552	0.7632	0.7647	0.7602	0.5858	0.5954	0.5948	0.5953
DDSH ^[90]	0.8681	0.8875	0.8922	0.8995	0.7672	0.8171	0.8161	0.8008	0.5807	0.6004	0.6127	0.6292
ADSH ^[96]	0.9043	0.9073	0.9073	0.9072	0.8962	0.9030	0.9035	0.8927	0.6605	0.6596	0.6648	0.6762
DMRH ^[164]	0.9256	0.9243	0.9245	0.9255	0.8577	0.8645	0.8659	0.8736	0.7510	0.7831	0.7890	0.8061

(2) 离散约束越严格, 检索性能越好. 从图 2 中可以看出, 同为监督方法, 采用离散优化策略的 FSSH、SSLH、SDHMLR、SCDH 和 RSLH 的性能均在整体上优于采用实值优化策略的 NSH. 这是由于通常情况下, 离散优化能够避免累积量化误差的问题, 从而提升哈希码的检索精度. 这一点已在文献 [66] 的消融研究中得以验证. 所以, 严格执行哈希码的离散约束可以提升哈希学习的精度. 此外, 比特平衡和比特独立等约束的保持也有助于获得稳健的性能^[165,166].

(3) 多数情况下, 长哈希码的检索性能比短哈希码的检索性能好, 这是因为长哈希码编码了更多的样本信息. 然而, 少数情况下, 长哈希码的检索精度反而低于短哈希码的检索精度. 如在 NUS-WIDE 数据集上, 对 ADSH 使

用 128 位哈希编码取得的检索性能略低于 64 位, 这可能是由于长哈希码会在嵌入中引入更多的噪声。

(4) 对于深度哈希学习算法, 损失函数的设计、离散优化以及网络结构的设计都会对方法的性能产生重要的影响. 如表 5 中所示, DMRH 在整体上优于 ADSH, 这是因为 DMRH 针对网络结构做了特殊设计; 绝大多数情况下, DSDH 的性能好于 DSH 和 DPSH, 这是因为 DSDH 采用了离散优化。

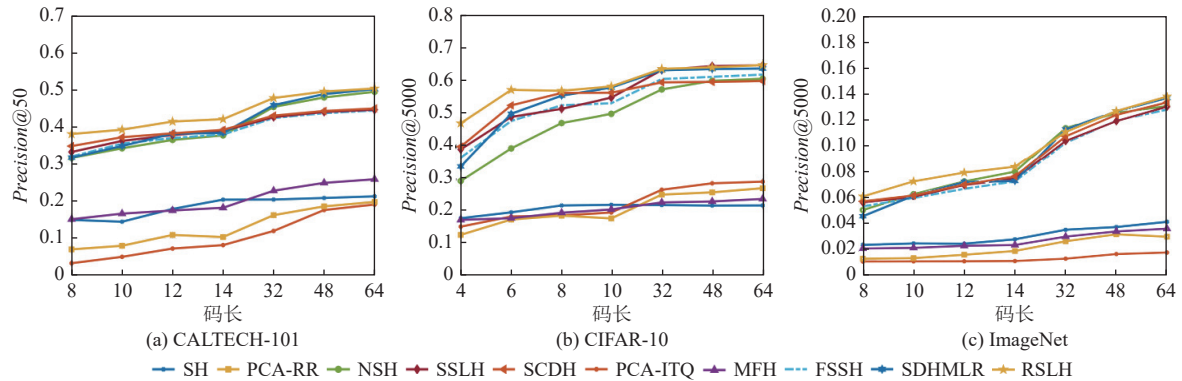


图 2 3 个基准数据集上各算法的 Precision 结果

(5) 没有观测结果表明, 某一种哈希学习算法在所有数据集上的各项指标的表现都最好. 这可能是由于不同的算法适用于不同的数据分布. 此外, 输入的特征表示与采用的优化策略也会影响性能. 这启发我们, 针对特定的数据集, 可以有针对性地设计哈希学习算法, 以更好地拟合数据分布信息。

5 未来研究展望

哈希学习强大的数据压缩功能和突出的检索性能使其在近几年得到了越来越多的关注, 哈希学习领域的研究深度与广度也不断得到扩展. 目前看来, 哈希学习依然存在诸多亟待解决的问题, 本节针对其中的关键问题展开讨论, 并从理论与实践两方面提出 8 个哈希学习的研究趋势。

(1) 参数自适应的哈希学习. 现有的哈希学习在设计损失函数时, 为了提升性能, 往往会引入各种不同的惩罚项, 并用不同的超参数进行加权. 在寻找各惩罚项的最优平衡系数参数与最优超参数时, 现有的方法大都借助交叉验证或者网格搜索的方式来手动设置. 目前针对自适应参数设置的研究较少^[167,168], 所以, 如何自适应地设置哈希学习中的超参数, 为模型快速训练与调整提供便利是一个亟待解决的问题。

(2) 哈希学习的理论分析. 现有哈希学习研究仅对收敛性、复杂度与稳定性做了较为简单的分析^[115], 而对于可学性、泛化界、一致性、收敛率、遗憾界等重要机器学习理论尚未有详尽的分析. 为哈希学习算法提供完善的理论保障, 进而指导哈希学习模型的构造、优化和推理过程, 并对哈希学习算法的质量和模型的有效性进行全面而清晰的验证, 仍是未来一个重要发展方向。

(3) 哈希学习的可解释性. 早期的哈希学习模型如 SH^[25]、KSH^[67]等, 具有较强的可解释性, 但其性能表现一般; 近些年取得显著研究成果的矩阵分解模型^[112,169,170]、量化模型^[171-173]、深度学习模型^[174-176]等, 较多地关注如何挖掘样本的语义信息以提升模型的检索精度, 而未给出内在的可解释性模型分析. 如何从提高哈希学习的可解释性角度降低模型的黑盒特性是一个值得深入研究的问题。

(4) 开放动态环境下的哈希学习. 已有的哈希学习基本上都是在给定的训练集上学习哈希函数, 进而为新数据生成哈希码并完成后续的检索. 然而, 在许多现实应用场景中, 随着数据的实时到达和不断增长, 需对哈希学习模型进行重新训练或增量学习. 虽然一些学者对在线场景中的哈希学习做了研究改进^[177-180], 然而, 处理动态数据集对哈希学习来说仍是一个不小的挑战. 对于流数据中的新类别、异常样本与特征分布变化等开放动态因素, 现有的哈希学习尚未展开充分的研究. 因此, 将哈希学习推广到与实际应用更为贴近的开放动态场景是一个热点研究

方向.

(5) 哈希学习在特定领域图像数据集上的检索. 随着现实应用中图像采集技术的进步与检索需求的增加, 图像检索面临的场景也越来越多、越来越复杂. 在现实场景中, 往往需要对某个精细领域的图像数据进行大规模检索^[181], 比如各种类型的医学影像数据集, 细粒度的动植物数据集, 种类繁多的商品数据集等. 目前已有研究者将哈希学习算法应用于人脸图像^[182]、医学图像^[183]和遥感图像^[184]. 如何将哈希学习与不同类型的图像检索场景有机结合是一个重要的研究方向.

(6) 跨数据集的哈希模型. 经典的哈希学习模型大都局限于特定领域数据的检索, 如商品检索、地标检索与车辆检索等. 然而, 对于不同类型的检索任务, 往往需要重新训练哈希学习模型, 这带来了巨大的时间和空间成本. 如今, 大规模预训练模型为下游任务的快速部署提供了极大的便利, 因而能够为哈希学习模型的跨数据集部署提供技术支撑. 所以, 面向大规模跨模态检索的哈希学习也是当下研究的一个热点^[185-188]. 其中, 如何弥补不同模态特征间的异构鸿沟是一个亟需解决的关键问题.

(7) 精细化的哈希检索模型. 现有的哈希学习技术主要是针对近似最近邻检索过程设计的, 对于同一哈希桶内的样本缺少重排序的功能. 如何结合量化与加权等技术, 对检索到的样本进行进一步精细化的排序是一个可探索的研究方向.

(8) 哈希学习在大规模机器学习中的应用. 哈希学习的优势在于它可以将高维数据转化为精简的二值表示, 而二值表示在存储与计算方面的开销极低. 因此, 将哈希学习与经典的分类模型、聚类模型相结合, 为大规模机器学习模型的轻量化提供支撑是一个可能的思路.

6 总 结

本文对面向大规模图像检索的哈希学习进行了系统的综述性研究, 总结了无监督哈希方法与监督哈希方法的研究现状, 并给出部分经典哈希学习算法的理论分析与实验结果. 与已有的哈希学习综述相比, 本文探索了哈希学习算法各组成部分间的关系, 并从损失函数、优化策略和样本外扩展映射 3 方面系统地梳理了哈希学习的算法要素与技术路线; 整理了哈希学习的经典研究成果与近期研究成果, 并分析了每个类目下代表性哈希学习算法的演化关系; 分析了哈希学习的局限性与新挑战, 并指出目前哈希学习的热点研究问题和重点应用方向.

目前看来, 面向大规模图像检索的哈希学习仍然是一个广泛、开放且具有挑战性的研究领域. 尽管各类哈希学习算法已经取得丰富的研究成果, 但日益高效的软硬件协同编程环境和不断涌现的具体应用需求也对哈希学习算法的检索性能、能效比与针对性设计提出了更高的要求. 因此, 无论是从学术研究的角度还是产业发展的角度, 哈希学习依然需要更加完善的理论保障体系与更多细致的方向性探索. 随着研究的深入与硬件技术的发展, 可以预见, 哈希学习在基础理论、效能提升、应用扩展等方面均可实现进一步的突破.

References:

- [1] i-SCOOP. How the pandemic impacted data creation and storage. 2022. <https://www.i-scoop.eu/big-data-action-value-context/data-storage-creation>
- [2] i-SCOOP. Data Age 2025: The datasphere and data-readiness from edge to core. 2022. <https://www.i-scoop.eu/big-data-action-value-context/data-age-2025-datasphere>
- [3] CNNIC. The 49th statistical report on China's Internet development. 2022. <https://www.cnnic.cn/n4/2022/0401/c88-1131.html>
- [4] Wang J, Kumar S, Chang SF. Semi-supervised hashing for scalable image retrieval. In: Proc. of the 2010 IEEE Conf. on Computer Vision and Pattern Recognition. San Francisco: IEEE, 2010. 3424–3431. [doi: 10.1109/CVPR.2010.5539994]
- [5] Tian X, Ng WWY, Wang H, Kwong S. Complementary incremental hashing with query-adaptive re-ranking for image retrieval. IEEE Trans. on Multimedia, 2021, 23: 1210–1224. [doi: 10.1109/TMM.2020.2994509]
- [6] Gong YC, Lazebnik S, Gordo A, Perronnin F. Iterative quantization: A procrustean approach to learning binary codes for large-scale image retrieval. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2013, 35(12): 2916–2929. [doi: 10.1109/TPAMI.2012.193]
- [7] The State Council, the People's Republic of China. Circular of the State Council on printing and issuing the plan for development of the new generation of artificial intelligence. 2017. https://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm

- [8] The State Council, the People's Republic of China. Government work report. 2022. http://www.gov.cn/premier/2022-03/12/content_5678750.htm
- [9] Hansen C, Hansen C, Simonsen JG, Alstrup S, Lioma C. Content-aware neural hashing for cold-start recommendation. In: Proc. of the 43rd Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. Xi'an: ACM, 2020. 971–980. [doi: [10.1145/3397271.3401060](https://doi.org/10.1145/3397271.3401060)]
- [10] CAICT. White paper on artificial intelligence (2022). Technical Report, Beijing: CAICT, 2022.
- [11] Indyk P, Motwani R. Approximate nearest neighbors: Towards removing the curse of dimensionality. In: Proc. of the 30th Annual ACM Symp. on Theory of Computing. Dallas: ACM, 1998. 604–613. [doi: [10.1145/276698.276876](https://doi.org/10.1145/276698.276876)]
- [12] Li W, Zhang Y, Sun YF, Wang W, Li MJ, Zhang WJ, Lin XM. Approximate nearest neighbor search on high dimensional data—experiments, analyses, and improvement. IEEE Trans. on Knowledge and Data Engineering, 2020, 32(8): 1475–1488. [doi: [10.1109/TKDE.2019.2909204](https://doi.org/10.1109/TKDE.2019.2909204)]
- [13] Jo J, Seo J, Fekete JD. PANENE: A progressive algorithm for indexing and querying approximate k -nearest neighbors. IEEE Trans. on Visualization and Computer Graphics, 2020, 26(2): 1347–1360. [doi: [10.1109/TVCG.2018.2869149](https://doi.org/10.1109/TVCG.2018.2869149)]
- [14] Ram P, Sinha K. Revisiting KD-tree for nearest neighbor search. In: Proc. of the 25th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. Anchorage: ACM, 2019. 1378–1388. [doi: [10.1145/3292500.3330875](https://doi.org/10.1145/3292500.3330875)]
- [15] Li S, Hu YM, Hao XH, Zhang LP, Hao ZX. Approximate k -nearest neighbor query of high dimensional data based on dimension grouping and reducing. Journal of Computer Research and Development, 2021, 58(3): 609–623 (in Chinese with English abstract). [doi: [10.7544/issn1000-1239.2021.20200285](https://doi.org/10.7544/issn1000-1239.2021.20200285)]
- [16] Jégou H, Douze M, Schmid C. Product quantization for nearest neighbor search. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2011, 33(1): 117–128. [doi: [10.1109/TPAMI.2010.57](https://doi.org/10.1109/TPAMI.2010.57)]
- [17] Babenko A, Lempitsky V. Additive quantization for extreme vector compression. In: Proc. of the 2014 IEEE Conf. on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014. 931–938. [doi: [10.1109/CVPR.2014.124](https://doi.org/10.1109/CVPR.2014.124)]
- [18] Li WJ, Zhou ZH. Learning to hash for big data: Current status and future trends. Chinese Science Bulletin, 2015, 60(5): 485–490 (in Chinese with English abstract). [doi: [10.1360/N972014-00841](https://doi.org/10.1360/N972014-00841)]
- [19] Wang JD, Zhang T, Song JK, Sebe N, Shen HT. A survey on learning to hash. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2018, 40(4): 769–790. [doi: [10.1109/TPAMI.2017.2699960](https://doi.org/10.1109/TPAMI.2017.2699960)]
- [20] Wang J, Liu W, Kumar S, Chang SF. Learning to hash for indexing big data—A survey. Proc. of the IEEE, 2016, 104(1): 34–57. [doi: [10.1109/JPROC.2015.2487976](https://doi.org/10.1109/JPROC.2015.2487976)]
- [21] Qin JB, Wang W, Xiao C, Zhang Y, Wang YS. High-dimensional similarity query processing for data science. In: Proc. of the 27th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. Singapore: ACM, 2021. 4062–4063. [doi: [10.1145/3447548.3470811](https://doi.org/10.1145/3447548.3470811)]
- [22] Kulis B, Grauman K. Kernelized locality-sensitive hashing for scalable image search. In: Proc. of the 12th IEEE Int'l Conf. on Computer Vision. Kyoto: IEEE, 2009. 2130–2137. [doi: [10.1109/ICCV.2009.5459466](https://doi.org/10.1109/ICCV.2009.5459466)]
- [23] Raginsky M, Lазebnik S. Locality-sensitive binary codes from shift-invariant kernels. In: Proc. of the 22nd Int'l Conf. on Neural Information Processing Systems. Vancouver: Curran Associates Inc., 2009. 1509–1517.
- [24] Liu W, Wang J, Kumar S, Chang SF. Hashing with graphs. In: Proc. of the 28th Int'l Conf. on Machine Learning. Washington: Omnipress, 2011. 1–8.
- [25] Weiss Y, Torralba A, Fergus R. Spectral hashing. In: Proc. of the 21st Int'l Conf. on Neural Information Processing Systems. Vancouver: Curran Associates Inc., 2008. 1753–1760.
- [26] Kong WH, Li WJ. Isotropic hashing. In: Proc. of the 25th Int'l Conf. on Neural Information Processing Systems. Lake Tahoe: Curran Associates Inc., 2012. 1646–1654.
- [27] Kong WH, Li WJ. Double-bit quantization for hashing. In: Proc. of the 26th AAAI Conf. on Artificial Intelligence. Toronto: AAAI, 2012. 634–640. [doi: [10.1609/aaai.v26i1.8208](https://doi.org/10.1609/aaai.v26i1.8208)]
- [28] Xie HT, Mao ZD, Zhang YD, Deng H, Yan CG, Chen ZN. Double-bit quantization and index hashing for nearest neighbor search. IEEE Trans. on Multimedia, 2019, 21(5): 1248–1260. [doi: [10.1109/TMM.2018.2872898](https://doi.org/10.1109/TMM.2018.2872898)]
- [29] Cao Y, Qi H, Gui J, Li KQ, Tang YY, Kwok JTY. Learning to hash with dimension analysis based quantizer for image retrieval. IEEE Trans. on Multimedia, 2021, 23: 3907–3918. [doi: [10.1109/TMM.2020.3033118](https://doi.org/10.1109/TMM.2020.3033118)]
- [30] Wang Z, Duan LY, Lin J, Wang XF, Huang TJ, Gao W. Hamming compatible quantization for hashing. In: Proc. of the 24th Int'l Conf. on Artificial Intelligence. Buenos Aires: AAAI, 2015. 2298–2304.
- [31] Hu MQ, Yang Y, Shen FM, Xie N, Shen HT. Hashing with angular reconstructive embeddings. IEEE Trans. on Image Processing, 2018,

- 27(2): 545–555. [doi: [10.1109/TIP.2017.2749147](https://doi.org/10.1109/TIP.2017.2749147)]
- [32] Duan LY, Wu YW, Huang YC, Wang Z, Yuan JS, Gao W. Minimizing reconstruction bias hashing via joint projection learning and quantization. *IEEE Trans. on Image Processing*, 2018, 27(6): 3127–3141. [doi: [10.1109/TIP.2018.2818008](https://doi.org/10.1109/TIP.2018.2818008)]
- [33] He KM, Wen F, Sun J. K-means hashing: An affinity-preserving quantization method for learning binary compact codes. In: *Proc. of the 2013 IEEE Conf. on Computer Vision and Pattern Recognition*. Portland: IEEE, 2013. 2938–2945. [doi: [10.1109/CVPR.2013.378](https://doi.org/10.1109/CVPR.2013.378)]
- [34] Norouzi M, Fleet DJ. Cartesian k -means. In: *Proc. of the 2013 IEEE Conf. on Computer Vision and Pattern Recognition*. Portland: IEEE, 2013. 3017–3024. [doi: [10.1109/CVPR.2013.388](https://doi.org/10.1109/CVPR.2013.388)]
- [35] Ryali CK, Hopfield JJ, Grinberg L, Krotov D. Bio-inspired hashing for unsupervised similarity search. In: *Proc. of the 37th Int'l Conf. on Machine Learning*. PMLR, 2020. 8295–8306.
- [36] Ding GG, Guo YC, Zhou JL. Collective matrix factorization hashing for multimodal data. In: *Proc. of the 2014 IEEE Conf. on Computer Vision and Pattern Recognition*. Columbus: IEEE, 2014. 2083–2090. [doi: [10.1109/CVPR.2014.267](https://doi.org/10.1109/CVPR.2014.267)]
- [37] Mukherjee L, Ravi SN, Ithapu VK, Holmes T, Singh V. An NMF perspective on binary hashing. In: *Proc. of the 2015 IEEE Int'l Conf. on Computer Vision*. Santiago: IEEE, 2015. 4184–4192. [doi: [10.1109/ICCV.2015.476](https://doi.org/10.1109/ICCV.2015.476)]
- [38] He SY, Wang BK, Wang Z, Yang Y, Shen FM, Huang Z, Shen HT. Bidirectional discrete matrix factorization hashing for image search. *IEEE Trans. on Cybernetics*, 2020, 50(9): 4157–4168. [doi: [10.1109/TCYB.2019.2941284](https://doi.org/10.1109/TCYB.2019.2941284)]
- [39] Hu D, Nie FP, Li XL. Discrete spectral hashing for efficient similarity retrieval. *IEEE Trans. on Image Processing*, 2019, 28(3): 1080–1091. [doi: [10.1109/TIP.2018.2875312](https://doi.org/10.1109/TIP.2018.2875312)]
- [40] Shen FM, Shen CH, Shi QF, van den Hengel A, Tang ZM. Inductive hashing on manifolds. In: *Proc. of the 2013 IEEE Conf. on Computer Vision and Pattern Recognition*. Portland: IEEE, 2013. 1562–1569. [doi: [10.1109/CVPR.2013.205](https://doi.org/10.1109/CVPR.2013.205)]
- [41] Liu W, Mu C, Kumar S, Chang SF. Discrete graph hashing. In: *Proc. of the 27th Int'l Conf. on Neural Information Processing Systems*. Montreal: MIT Press, 2014. 3419–3427.
- [42] Jiang QY, Li WJ. Scalable graph hashing with feature transformation. In: *Proc. of the 24th Int'l Conf. on Artificial Intelligence*. Buenos Aires: AAAI, 2015. 2248–2254.
- [43] Irie G, Li ZG, Wu XM, Chang SF. Locally linear hashing for extracting non-linear manifolds. In: *Proc. of the 2014 IEEE Conf. on Computer Vision and Pattern Recognition*. Columbus: IEEE, 2014. 2123–2130. [doi: [10.1109/CVPR.2014.272](https://doi.org/10.1109/CVPR.2014.272)]
- [44] Ma C, Gong C, Li X, Huang XL, Liu W, Yang J. Toward making unsupervised graph hashing discriminative. *IEEE Trans. on Multimedia*, 2020, 22(3): 760–774. [doi: [10.1109/TMM.2019.2931808](https://doi.org/10.1109/TMM.2019.2931808)]
- [45] Jin S, Yao HX, Zhou Q, Liu Y, Huang JQ, Hua XS. Unsupervised discrete hashing with affinity similarity. *IEEE Trans. on Image Processing*, 2021, 30: 6130–6141. [doi: [10.1109/TIP.2021.3091895](https://doi.org/10.1109/TIP.2021.3091895)]
- [46] Wang M, Zhou WG, Tian Q, Li HQ. Neighborhood pyramid preserving hashing. *IEEE Trans. on Multimedia*, 2020, 22(6): 1507–1518. [doi: [10.1109/TMM.2019.2943778](https://doi.org/10.1109/TMM.2019.2943778)]
- [47] Liu H, Ji RR, Wu YJ, Huang FY. Ordinal constrained binary code learning for nearest neighbor search. In: *Proc. of the 31st AAAI Conf. on Artificial Intelligence*. San Francisco: AAAI, 2017. 2238–2244. [doi: [10.1609/aaai.v31i1.10906](https://doi.org/10.1609/aaai.v31i1.10906)]
- [48] Liu H, Ji RR, Wang JD, Shen CH. Ordinal constraint binary coding for approximate nearest neighbor search. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2019, 41(4): 941–955. [doi: [10.1109/TPAMI.2018.2819978](https://doi.org/10.1109/TPAMI.2018.2819978)]
- [49] Carreira-Perpiñán MÁ, Raziperchikolaei R. Hashing with binary autoencoders. In: *Proc. of the 2015 IEEE Conf. on Computer Vision and Pattern Recognition*. Boston: IEEE, 2015. 557–566. [doi: [10.1109/CVPR.2015.7298654](https://doi.org/10.1109/CVPR.2015.7298654)]
- [50] Li YQ, van Gemert J. Deep unsupervised image hashing by maximizing bit entropy. In: *Proc. of the 35th AAAI Conf. on Artificial Intelligence*. AAAI, 2021. 2002–2010. [doi: [10.1609/aaai.v35i3.16296](https://doi.org/10.1609/aaai.v35i3.16296)]
- [51] Dizaji KG, Zheng F, Nourabadi NS, Yang YH, Deng C, Huang H. Unsupervised deep generative adversarial hashing network. In: *Proc. of the 2018 IEEE/CVF Conf. on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 3664–3673. [doi: [10.1109/CVPR.2018.00386](https://doi.org/10.1109/CVPR.2018.00386)]
- [52] Lin K, Lu JW, Chen CS, Zhou J, Sun MT. Unsupervised deep learning of compact binary descriptors. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2019, 41(6): 1501–1514. [doi: [10.1109/TPAMI.2018.2833865](https://doi.org/10.1109/TPAMI.2018.2833865)]
- [53] Do TT, Doan AD, Cheung NM. Learning to hash with binary deep neural network. In: *Proc. of the 14th European Conf. on Computer Vision*. Amsterdam: Springer, 2016. 219–234. [doi: [10.1007/978-3-319-46454-1_14](https://doi.org/10.1007/978-3-319-46454-1_14)]
- [54] Liu L, Yu MY, Shao L. Latent structure preserving hashing. *Int'l Journal of Computer Vision*, 2017, 122(3): 439–457. [doi: [10.1007/s11263-016-0931-4](https://doi.org/10.1007/s11263-016-0931-4)]
- [55] Deng C, Yang EK, Liu TL, Li J, Liu W, Tao DC. Unsupervised semantic-preserving adversarial hashing for image search. *IEEE Trans. on Image Processing*, 2019, 28(8): 4032–4044. [doi: [10.1109/TIP.2019.2903661](https://doi.org/10.1109/TIP.2019.2903661)]

- [56] Hu QH, Wu JX, Cheng J, Wu LF, Lu HQ. Pseudo label based unsupervised deep discriminative hashing for image retrieval. In: Proc. of the 25th ACM Int'l Conf. on Multimedia. Mountain View: ACM, 2017. 1584–1590. [doi: 10.1145/3123266.3123403]
- [57] Zhang HF, Liu L, Long Y, Shao L. Unsupervised deep hashing with pseudo labels for scalable image retrieval. IEEE Trans. on Image Processing, 2018, 27(4): 1626–1638. [doi: 10.1109/TIP.2017.2781422]
- [58] Yang EK, Deng C, Liu TL, Liu W, Tao DC. Semantic structure-based unsupervised deep hashing. In: Proc. of the 27th Int'l Joint Conf. on Artificial Intelligence. Stockholm: AAAI, 2018. 1064–1070.
- [59] Dong X, Liu L, Zhu L, Cheng ZY, Zhang HX. Unsupervised deep K-means hashing for efficient image retrieval and clustering. IEEE Trans. on Circuits and Systems for Video Technology, 2021, 31(8): 3266–3277. [doi: 10.1109/TCSVT.2020.3035775]
- [60] Qin QB, Huang L, Wei ZQ, Xie KZ, Zhang WF. Unsupervised deep multi-similarity hashing with semantic structure for image retrieval. IEEE Trans. on Circuits and Systems for Video Technology, 2021, 31(7): 2852–2865. [doi: 10.1109/TCSVT.2020.3032402]
- [61] Luo X, Wu DQ, Ma ZY, Chen C, Deng MH, Huang JQ, Hua XS. A statistical approach to mining semantic similarity for deep unsupervised hashing. In: Proc. of the 29th ACM Int'l Conf. on Multimedia. Chengdu: ACM, 2021. 4306–4314. [doi: 10.1145/3474085.3475570]
- [62] Li CS, Min QX, Cheng YR, Yuan Y, Wang GR. Local semantic structure captured and instance discriminated by unsupervised hashing. Ruan Jian Xue Bao/Journal of Software, 2021, 32(3): 742–752 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/6178.htm> [doi: 10.13328/j.cnki.jos.006178]
- [63] Shen FM, Xu Y, Liu L, Yang Y, Huang Z, Shen HT. Unsupervised deep hashing with similarity-adaptive and discrete optimization. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2018, 40(12): 3034–3044. [doi: 10.1109/TPAMI.2018.2789887]
- [64] Shen YM, Qin J, Chen JX, Yu MY, Liu L, Zhu F, Shen FM, Shao L. Auto-encoding twin-bottleneck hashing. In: Proc. of the 2020 IEEE/CVF Conf. on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020. 2815–2824. [doi: 10.1109/CVPR42600.2020.00289]
- [65] Zhang HF, Gu YF, Yao YZ, Zhang Z, Liu L, Zhang J, Shao L. Deep unsupervised self-evolutionary hashing for image retrieval. IEEE Trans. on Multimedia, 2021, 23: 3400–3413. [doi: 10.1109/TMM.2020.3025000]
- [66] Shen FM, Shen CH, Liu W, Shen HT. Supervised discrete hashing. In: Proc. of the 2015 IEEE Conf. on Computer Vision and Pattern Recognition. Boston: IEEE, 2015. 37–45. [doi: 10.1109/CVPR.2015.7298598]
- [67] Liu W, Wang J, Ji RR, Jiang YG, Chang SF. Supervised hashing with kernels. In: Proc. of the 2012 IEEE Conf. on Computer Vision and Pattern Recognition. Providence: IEEE, 2012. 2074–2081. [doi: 10.1109/CVPR.2012.6247912]
- [68] Kang WC, Li WJ, Zhou ZH. Column sampling based discrete supervised hashing. In: Proc. of the 30th AAAI Conf. on Artificial Intelligence. Phoenix: AAAI, 2016. 1230–1236. [doi: 10.1609/aaai.v30i1.10176]
- [69] Zhu H, Long MS, Wang JM, Cao Y. Deep hashing network for efficient similarity retrieval. In: Proc. of the 30th AAAI Conf. on Artificial Intelligence. Phoenix: AAAI, 2016. 2415–2421. [doi: 10.1609/aaai.v30i1.10235]
- [70] Cao Y, Long MS, Liu B, Wang JM. Deep Cauchy hashing for hamming space retrieval. In: Proc. of the 2018 IEEE/CVF Conf. on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 1229–1237. [doi: 10.1109/CVPR.2018.00134]
- [71] Li Q, Sun ZN, He R, Tan TN. Deep supervised discrete hashing. In: Proc. of the 31st Int'l Conf. on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017. 2479–2488.
- [72] Liu Q, Lu HT. Natural supervised hashing. In: Proc. of the 25th Int'l Joint Conf. on Artificial Intelligence. New York: AAAI, 2016. 1788–1794.
- [73] Gui J, Liu TL, Sun ZN, Tao DC, Tan TN. Supervised discrete hashing with relaxation. IEEE Trans. on Neural Networks and Learning Systems, 2018, 29(3): 608–617. [doi: 10.1109/TNNLS.2016.2636870]
- [74] Gui J, Li P. R²SDH: Robust rotated supervised discrete hashing. In: Proc. of the 24th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. London: ACM, 2018. 1485–1493. [doi: 10.1145/3219819.3219955]
- [75] Zhang HW, Zhao N, Shang XD, Luan HB, Chua TS. Discrete image hashing using large weakly annotated photo collections. In: Proc. of the 30th AAAI Conf. on Artificial Intelligence. Phoenix: AAAI, 2016. 3669–3675. [doi: 10.1609/aaai.v30i1.10453]
- [76] Yang HF, Lin K, Chen CS. Supervised learning of semantics-preserving hash via deep convolutional neural networks. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2018, 40(2): 437–451. [doi: 10.1109/TPAMI.2017.2666812]
- [77] Zhou C, Po LM, Ou WF. Angular deep supervised vector quantization for image retrieval. IEEE Trans. on Neural Networks and Learning Systems, 2022, 33(4): 1638–1649. [doi: 10.1109/TNNLS.2020.3043103]
- [78] Cui H, Zhu L, Li JJ, Yang Y, Nie LQ. Scalable deep hashing for large-scale social image retrieval. IEEE Trans. on Image Processing, 2020, 29: 1271–1284. [doi: 10.1109/TIP.2019.2940693]
- [79] Zhang Z, Liu LY, Luo YD, Huang Z, Shen FM, Shen HT, Lu GM. Inductive structure consistent hashing via flexible semantic

- calibration. *IEEE Trans. on Neural Networks and Learning Systems*, 2021, 32(10): 4514–4528. [doi: [10.1109/TNNLS.2020.3018790](https://doi.org/10.1109/TNNLS.2020.3018790)]
- [80] Zhu L, Huang Z, Li ZH, Xie L, Shen HT. Exploring auxiliary context: Discrete semantic transfer hashing for scalable image retrieval. *IEEE Trans. on Neural Networks and Learning Systems*, 2018, 29(11): 5264–5276. [doi: [10.1109/TNNLS.2018.2797248](https://doi.org/10.1109/TNNLS.2018.2797248)]
- [81] Zhou X, Shen FM, Liu L, Liu W, Nie LQ, Yang Y, Shen HT. Graph convolutional network hashing. *IEEE Trans. on Cybernetics*, 2020, 50(4): 1460–1472. [doi: [10.1109/TCYB.2018.2883970](https://doi.org/10.1109/TCYB.2018.2883970)]
- [82] Li ZC, Tang JH, Zhang LY, Yang J. Weakly-supervised semantic guided hashing for social image retrieval. *Int'l Journal of Computer Vision*, 2020, 128(8–9): 2265–2278. [doi: [10.1007/s11263-020-01331-0](https://doi.org/10.1007/s11263-020-01331-0)]
- [83] Yang Y, Luo YD, Chen WL, Shen FM, Shao J, Shen HT. Zero-shot hashing via transferring supervised knowledge. In: *Proc. of the 24th ACM Int'l Conf. on Multimedia*. Amsterdam: ACM, 2016. 1286–1295. [doi: [10.1145/2964284.2964319](https://doi.org/10.1145/2964284.2964319)]
- [84] Ji Z, Sun YX, Yu YL, Pang YW, Han JG. Attribute-guided network for cross-modal zero-shot hashing. *IEEE Trans. on Neural Networks and Learning Systems*, 2020, 31(1): 321–330. [doi: [10.1109/TNNLS.2019.2904991](https://doi.org/10.1109/TNNLS.2019.2904991)]
- [85] Wu CX, Zhu JK, Cai D, Chen C, Bu JJ. Semi-supervised nonlinear hashing using bootstrap sequential projection learning. *IEEE Trans. on Knowledge and Data Engineering*, 2013, 25(6): 1380–1393. [doi: [10.1109/TKDE.2012.76](https://doi.org/10.1109/TKDE.2012.76)]
- [86] Lin GS, Shen CH, Suter D, van den Hengel A. A general two-step approach to learning-based hashing. In: *Proc. of the 2013 IEEE Conf. on Computer Vision*. Sydney: IEEE, 2013. 2552–2559. [doi: [10.1109/ICCV.2013.317](https://doi.org/10.1109/ICCV.2013.317)]
- [87] Lin GS, Shen CH, van den Hengel A. Supervised hashing using graph cuts and boosted decision trees. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2015, 37(11): 2317–2331. [doi: [10.1109/TPAMI.2015.2404776](https://doi.org/10.1109/TPAMI.2015.2404776)]
- [88] Chen ZX, Yuan X, Lu JW, Tian Q, Zhou J. Deep hashing via discrepancy minimization. In: *Proc. of the 2018 IEEE/CVF Conf. on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 6838–6847. [doi: [10.1109/CVPR.2018.00715](https://doi.org/10.1109/CVPR.2018.00715)]
- [89] Li WJ, Wang S, Kang WC. Feature learning based deep supervised hashing with pairwise labels. In: *Proc. of the 25th Int'l Joint Conf. on Artificial Intelligence*. New York: AAAI, 2016. 1711–1717.
- [90] Jiang QY, Cui X, Li WJ. Deep discrete supervised hashing. *IEEE Trans. on Image Processing*, 2018, 27(12): 5996–6009. [doi: [10.1109/TIP.2018.2864894](https://doi.org/10.1109/TIP.2018.2864894)]
- [91] Chen Y, Tian ZB, Zhang H, Wang J, Zhang D. Strongly constrained discrete hashing. *IEEE Trans. on Image Processing*, 2020, 29: 3596–3611. [doi: [10.1109/TIP.2020.2963952](https://doi.org/10.1109/TIP.2020.2963952)]
- [92] Wang D, Huang HY, Lu C, Feng BS, Wen GH, Nie LQ, Mao XL. Supervised deep hashing for hierarchical labeled data. In: *Proc. of the 32nd AAAI Conf. on Artificial Intelligence*. New Orleans: AAAI, 2018. 7388–7395. [doi: [10.1609/aaai.v32i1.12247](https://doi.org/10.1609/aaai.v32i1.12247)]
- [93] Wang YB, Ou XF, Liang J, Sun ZN. Deep semantic reconstruction hashing for similarity retrieval. *IEEE Trans. on Circuits and Systems for Video Technology*, 2021, 31(1): 387–400. [doi: [10.1109/TCSVT.2020.2974768](https://doi.org/10.1109/TCSVT.2020.2974768)]
- [94] Zhang PC, Zhang W, Li WJ, Guo MY. Supervised hashing with latent factor models. In: *Proc. of the 37th Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval*. Gold Coast: ACM, 2014. 173–182. [doi: [10.1145/2600428.2609600](https://doi.org/10.1145/2600428.2609600)]
- [95] Shi XS, Xing FY, Xu KD, Sapkota M, Yang L. Asymmetric discrete graph hashing. In: *Proc. of the 31st AAAI Conf. on Artificial Intelligence*. San Francisco: AAAI, 2017. 2541–2547. [doi: [10.1609/aaai.v31i1.10831](https://doi.org/10.1609/aaai.v31i1.10831)]
- [96] Jiang QY, Li WJ. Asymmetric deep supervised hashing. In: *Proc. of the 32nd AAAI Conf. on Artificial Intelligence*. New Orleans: AAAI, 2018. 3342–3349. [doi: [10.1609/aaai.v32i1.11814](https://doi.org/10.1609/aaai.v32i1.11814)]
- [97] Zhao W, Xu C, Guan ZY, Wu XL, Zhao WQ, Miao QG, He XF, Wang Q. TelecomNet: Tag-based weakly-supervised modally cooperative hashing network for image retrieval. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2022, 44(11): 7940–7954. [doi: [10.1109/TPAMI.2021.3114089](https://doi.org/10.1109/TPAMI.2021.3114089)]
- [98] Shen XB, Dong GH, Zheng YH, Lan L, Tsang IW, Sun QS. Deep co-image-label hashing for multi-label image retrieval. *IEEE Trans. on Multimedia*, 2022, 24: 1116–1126. [doi: [10.1109/TMM.2021.3119868](https://doi.org/10.1109/TMM.2021.3119868)]
- [99] Lai HJ, Pan Y, Liu Y, Yan SC. Simultaneous feature learning and hash coding with deep neural networks. In: *Proc. of the 2015 IEEE Conf. on Computer Vision and Pattern Recognition*. Boston: IEEE, 2015. 3270–3278. [doi: [10.1109/CVPR.2015.7298947](https://doi.org/10.1109/CVPR.2015.7298947)]
- [100] Zhuang BH, Lin GS, Shen CH, Reid I. Fast training of triplet-based deep binary embedding networks. In: *Proc. of the 2016 IEEE Conf. on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016. 5955–5964. [doi: [10.1109/CVPR.2016.641](https://doi.org/10.1109/CVPR.2016.641)]
- [101] Liu B, Cao Y, Long MS, Wang JM, Wang JD. Deep triplet quantization. In: *Proc. of the 26th ACM Conf. on Multimedia*. Seoul: ACM, 2018. 755–763. [doi: [10.1145/3240508.3240516](https://doi.org/10.1145/3240508.3240516)]
- [102] Zhang Z, Zhu XF, Lu GM, Zhang YD. Probability ordinal-preserving semantic hashing for large-scale image retrieval. *ACM Trans. on Knowledge Discovery from Data*, 2021, 15(3): 37. [doi: [10.1145/3442204](https://doi.org/10.1145/3442204)]
- [103] Liang YC, Pan Y, Lai HJ, Liu W, Yin J. Deep listwise triplet hashing for fine-grained image retrieval. *IEEE Trans. on Image Processing*, 2022, 31: 949–961. [doi: [10.1109/TIP.2021.3137653](https://doi.org/10.1109/TIP.2021.3137653)]

- [104] Liu HM, Wang RP, Shan SG, Chen XL. Deep supervised hashing for fast image retrieval. In: Proc. of the 2016 IEEE Conf. on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016. 2064–2072. [doi: [10.1109/CVPR.2016.227](https://doi.org/10.1109/CVPR.2016.227)]
- [105] Chen YD, Lai ZH, Ding YJ, Lin KY, Wong W. Deep supervised hashing with anchor graph. In: Proc. of the 2019 IEEE/CVF Int'l Conf. on Computer Vision. Seoul: IEEE, 2019. 9795–9803. [doi: [10.1109/ICCV.2019.00989](https://doi.org/10.1109/ICCV.2019.00989)]
- [106] Li Q, Sun ZN, He R, Tan TN. A general framework for deep supervised discrete hashing. *Int'l Journal of Computer Vision*, 2020, 128(8-9): 2204–2222. [doi: [10.1007/s11263-020-01327-w](https://doi.org/10.1007/s11263-020-01327-w)]
- [107] Zhang J, Peng YX, Yuan MK. SCH-GAN: Semi-supervised cross-modal hashing by generative adversarial network. *IEEE Trans. on Cybernetics*, 2020, 50(2): 489–502. [doi: [10.1109/TCYB.2018.2868826](https://doi.org/10.1109/TCYB.2018.2868826)]
- [108] Cao Y, Qi H, Zhou WR, Kato J, Li KQ, Liu XL, Gui J. Binary hashing for approximate nearest neighbor search on big data: A survey. *IEEE Access*, 2018, 6: 2039–2054. [doi: [10.1109/ACCESS.2017.2781360](https://doi.org/10.1109/ACCESS.2017.2781360)]
- [109] Fei LK, Qin JY, Teng SH, Zhang W, Liu DN, Hou Y. Hashing for approximate nearest neighbor search on big data: A survey. *Journal of Guangdong University of Technology*, 2020, 37(3): 23–35 (in Chinese with English abstract). [doi: [10.12052/gdutxb.190123](https://doi.org/10.12052/gdutxb.190123)]
- [110] Luo X, Wang HX, Wu DQ, Chen C, Deng MH, Huang JQ, Hua XS. A survey on deep hashing methods. *ACM Trans. on Knowledge Discovery from Data*, 2023, 17(1): 15. [doi: [10.1145/3532624](https://doi.org/10.1145/3532624)]
- [111] Luo X, Nie LQ, He XN, Wu Y, Chen ZD, Xu XS. Fast scalable supervised hashing. In: Proc. of the 41st Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. Ann Arbor: ACM, 2018. 735–744. [doi: [10.1145/3209978.3210035](https://doi.org/10.1145/3209978.3210035)]
- [112] Wang D, Gao XB, Wang XM, He LH. Label consistent matrix factorization hashing for large-scale cross-modal similarity search. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2019, 41(10): 2466–2479. [doi: [10.1109/TPAMI.2018.2861000](https://doi.org/10.1109/TPAMI.2018.2861000)]
- [113] Shen FM, Zhou X, Yang Y, Song JK, Shen HT, Tao DC. A fast optimization method for general binary code learning. *IEEE Trans. on Image Processing*, 2016, 25(12): 5610–5621. [doi: [10.1109/TIP.2016.2612883](https://doi.org/10.1109/TIP.2016.2612883)]
- [114] Xiong H, Yu MY, Liu L, Zhu F, Qin J, Shen FM, Shao L. A generalized method for binary optimization: Convergence analysis and applications. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2022, 44(9): 4524–4543. [doi: [10.1109/TPAMI.2021.3070753](https://doi.org/10.1109/TPAMI.2021.3070753)]
- [115] Gui J, Liu TL, Sun ZN, Tao DC, Tan TN. Fast supervised discrete hashing. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2018, 40(2): 490–496. [doi: [10.1109/TPAMI.2017.2678475](https://doi.org/10.1109/TPAMI.2017.2678475)]
- [116] He JF, Liu W, Chang SF. Scalable similarity search with optimized kernel hashing. In: Proc. of the 16th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. Washington: ACM, 2010. 1129–1138. [doi: [10.1145/1835804.1835946](https://doi.org/10.1145/1835804.1835946)]
- [117] Strecha C, Bronstein AM, Bronstein MM, Fua P. LDAHash: Improved matching with smaller descriptors. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2012, 34(1): 66–78. [doi: [10.1109/TPAMI.2011.103](https://doi.org/10.1109/TPAMI.2011.103)]
- [118] Wang JF, Shen HT, Yan SC, Yu NH, Li SP, Wang JD. Optimized distances for binary code ranking. In: Proc. of the 22nd ACM Int'l Conf. on Multimedia. Orlando: ACM, 2014. 517–526. [doi: [10.1145/2647868.2654898](https://doi.org/10.1145/2647868.2654898)]
- [119] Zhang L, Zhang YD, Tang JH, Lu K, Tian Q. Binary code ranking with weighted hamming distance. In: Proc. of the 2013 IEEE Conf. on Computer Vision and Pattern Recognition. Portland: IEEE, 2013. 1586–1593. [doi: [10.1109/CVPR.2013.208](https://doi.org/10.1109/CVPR.2013.208)]
- [120] Gordo A, Perronnin F, Gong YC, Lazebnik S. Asymmetric distances for binary embeddings. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2014, 36(1): 33–47. [doi: [10.1109/TPAMI.2013.101](https://doi.org/10.1109/TPAMI.2013.101)]
- [121] Ye RZ, Li XL. Compact structure hashing via sparse and similarity preserving embedding. *IEEE Trans. on Cybernetics*, 2016, 46(3): 718–729. [doi: [10.1109/TCYB.2015.2414299](https://doi.org/10.1109/TCYB.2015.2414299)]
- [122] Long J, Sun LZ, Guo L, Hua LJ, Yang Z. Label embedding semantic-guided hashing. *Neurocomputing*, 2022, 477: 1–13. [doi: [10.1016/j.neucom.2021.12.073](https://doi.org/10.1016/j.neucom.2021.12.073)]
- [123] Liu XB, Nie XS, Zhou Q, Yin YL. Supervised discrete hashing with mutual linear regression. In: Proc. of the 27th ACM Int'l Conf. on Multimedia. Nice: ACM, 2019. 1561–1568. [doi: [10.1145/3343031.3351091](https://doi.org/10.1145/3343031.3351091)]
- [124] Kulis B, Darrell T. Learning to hash with binary reconstructive embeddings. In: Proc. of the 22nd Int'l Conf. on Neural Information Processing Systems. Vancouver: Curran Associates Inc., 2009. 1042–1050.
- [125] Li N, Li C, Deng C, Liu XL, Gao XB. Deep joint semantic-embedding hashing. In: Proc. of the 27th Int'l Joint Conf. on Artificial Intelligence. Stockholm: AAAI, 2018. 2397–2403.
- [126] Jang YK, Cho NI. Generalized product quantization network for semi-supervised image retrieval. In: Proc. of the 2020 IEEE/CVF Conf. on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020. 3417–3426. [doi: [10.1109/CVPR42600.2020.00348](https://doi.org/10.1109/CVPR42600.2020.00348)]
- [127] Chen ZX, Lu JW, Feng JJ, Zhou J. Nonlinear sparse hashing. *IEEE Trans. on Multimedia*, 2017, 19(9): 1996–2009. [doi: [10.1109/TMM.2017.2705918](https://doi.org/10.1109/TMM.2017.2705918)]
- [128] Chen Q, Shen ZW, Chen Z. A penalty function semi-continuous thresholding methods for constraints of hashing problems. *Journal of Visual Communication and Image Representation*, 2022, 87: 103552. [doi: [10.1016/j.jvcir.2022.103552](https://doi.org/10.1016/j.jvcir.2022.103552)]

- [129] Koutaki G, Shirai K, Ambai M. Hadamard coding for supervised discrete hashing. *IEEE Trans. on Image Processing*, 2018, 27(11): 5378–5392. [doi: [10.1109/TIP.2018.2855427](https://doi.org/10.1109/TIP.2018.2855427)]
- [130] Cao ZJ, Long MS, Wang JM, Yu PS. HashNet: Deep learning to hash by continuation. In: *Proc. of the 2017 IEEE Conf. on Computer Vision*. Venice: IEEE, 2017. 5609–5618. [doi: [10.1109/ICCV.2017.598](https://doi.org/10.1109/ICCV.2017.598)]
- [131] Su SP, Zhang C, Han K, Tian YH. Greedy hash: Towards fast optimization for accurate hash coding in CNN. In: *Proc. of the 32nd Int'l Conf. on Neural Information Processing Systems*. Montreal: Curran Associates Inc., 2018. 806–815.
- [132] Huang LK, Chen JD, Pan S. Accelerate learning of deep hashing with gradient attention. In: *Proc. of the 2019 IEEE/CVF Int'l Conf. on Computer Vision*. Seoul: IEEE, 2019. 5270–5279. [doi: [10.1109/ICCV.2019.00537](https://doi.org/10.1109/ICCV.2019.00537)]
- [133] Nie XS, Liu XB, Xi XM, Yin YL. Model parameter re-optimization for linear hashing based on similarity drive. *Journal of Software*, 2020, 31(4): 1039–1050 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/5918.htm> [doi: [10.13328/j.cnki.jos.005918](https://doi.org/10.13328/j.cnki.jos.005918)]
- [134] Mu YD, Liu Z. Deep hashing: A joint approach for image signature learning. In: *Proc. of the 31st AAAI Conf. on Artificial Intelligence*. San Francisco: AAAI, 2017. 2380–2386. [doi: [10.1609/aaai.v31i1.10858](https://doi.org/10.1609/aaai.v31i1.10858)]
- [135] Salakhutdinov R, Hinton G. Semantic hashing. *Int'l Journal of Approximate Reasoning*, 2009, 50(7): 969–978. [doi: [10.1016/j.ijar.2008.11.006](https://doi.org/10.1016/j.ijar.2008.11.006)]
- [136] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial networks. *Communications of the ACM*, 2020, 63(11): 139–144. [doi: [10.1145/3422622](https://doi.org/10.1145/3422622)]
- [137] Nair V, Hinton GE. Rectified linear units improve restricted boltzmann machines. In: *Proc. of the 27th Int'l Conf. on Machine Learning*. Haifa: Omnipress, 2010. 807–814.
- [138] Zhang W, Yang XQ, Teng SH, Wu NQ. Semantic-guided hashing learning for domain adaptive retrieval. *World Wide Web*, 2023, 26(3): 1093–1112. [doi: [10.1007/s11280-022-01072-9](https://doi.org/10.1007/s11280-022-01072-9)]
- [139] Shakhnarovich G, Viola P, Darrell T. Fast pose estimation with parameter-sensitive hashing. In: *Proc. of the 9th IEEE Int'l Conf. on Computer Vision*. Nice: IEEE, 2003. 750–757. [doi: [10.1109/ICCV.2003.1238424](https://doi.org/10.1109/ICCV.2003.1238424)]
- [140] Rastegari M, Farhadi A, Forsyth D. Attribute discovery via predictable discriminative binary codes. In: *Proc. of the 12th European Conf. on Computer Vision*. Florence: Springer, 2012. 876–889. [doi: [10.1007/978-3-642-33783-3_63](https://doi.org/10.1007/978-3-642-33783-3_63)]
- [141] Xu XX, Lai ZH, Chen YD. Relaxed locality preserving supervised discrete hashing. *IEEE Trans. on Big Data*, 2022, 8(4): 1118–1128. [doi: [10.1109/TBDATA.2020.3027379](https://doi.org/10.1109/TBDATA.2020.3027379)]
- [142] Xia RK, Pan Y, Lai HJ, Liu C, Yan SC. Supervised hashing for image retrieval via image representation learning. In: *28th AAAI Conf. on Artificial Intelligence*. Québec City: AAAI, 2014. 2156–2162. [doi: [10.1609/aaai.v28i1.8952](https://doi.org/10.1609/aaai.v28i1.8952)]
- [143] Yang HF, Tu CH, Chen CS. Learning binary hash codes based on adaptable label representations. *IEEE Trans. on Neural Networks and Learning Systems*, 2022, 33(11): 6961–6975. [doi: [10.1109/TNNLS.2021.3095399](https://doi.org/10.1109/TNNLS.2021.3095399)]
- [144] Wu DY, Dai Q, Li B, Wang WP. Deep uncoupled discrete hashing via similarity matrix decomposition. *ACM Trans. on Multimedia Computing, Communications, and Applications*, 2023, 19(1): 22. [doi: [10.1145/3524021](https://doi.org/10.1145/3524021)]
- [145] Liu L, Shao L, Shen FM, Yu MY. Discretely coding semantic rank orders for supervised image hashing. In: *Proc. of the 2017 IEEE Conf. on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017. 5140–5149. [doi: [10.1109/CVPR.2017.546](https://doi.org/10.1109/CVPR.2017.546)]
- [146] Chen YB, Guo KC, Liu FX, Huang YS, Qi ZW. Supervised contrastive vehicle quantization for efficient vehicle retrieval. In: *Proc. of the 2022 Int'l Conf. on Multimedia Retrieval*. Newark: ACM, 2022. 44–48. [doi: [10.1145/3512527.3531432](https://doi.org/10.1145/3512527.3531432)]
- [147] Oliva A, Torralba A. Modeling the shape of the scene: A holistic representation of the spatial envelope. *Int'l Journal of Computer Vision*, 2001, 42(3): 145–175. [doi: [10.1023/A:1011139631724](https://doi.org/10.1023/A:1011139631724)]
- [148] Lowe DG. Distinctive image features from scale-invariant keypoints. *Int'l Journal of Computer Vision*, 2004, 60(2): 91–110. [doi: [10.1023/B:VISI.0000029664.99615.94](https://doi.org/10.1023/B:VISI.0000029664.99615.94)]
- [149] Guo ZH, Zhang L, Zhang D. A completed modeling of local binary pattern operator for texture classification. *IEEE Trans. on Image Processing*, 2010, 19(6): 1657–1663. [doi: [10.1109/TIP.2010.2044957](https://doi.org/10.1109/TIP.2010.2044957)]
- [150] Perronnin F, Liu Y, Sánchez J, Poirier H. Large-scale image retrieval with compressed fisher vectors. In: *Proc. of the 2010 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*. San Francisco: IEEE, 2010. 3384–3391. [doi: [10.1109/CVPR.2010.5540009](https://doi.org/10.1109/CVPR.2010.5540009)]
- [151] Jégou H, Douze M, Schmid C, Pérez P. Aggregating local descriptors into a compact image representation. In: *Proc. of the 2010 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*. San Francisco: IEEE, 2010. 3304–3311. [doi: [10.1109/CVPR.2010.5540039](https://doi.org/10.1109/CVPR.2010.5540039)]
- [152] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: *Proc. of the 25th Int'l*

- Conf. on Neural Information Processing Systems. Lake Tahoe: Curran Associates Inc., 2012. 1097–1105.
- [153] Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai XH, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, Uszkoreit J, Houshby N. An image is worth 16x16 words: Transformers for image recognition at scale. In: Proc. of the 9th Int'l Conf. on Learning Representations. OpenReview.net, 2021.
- [154] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc. of the IEEE, 1998, 86(11): 2278–2324. [doi: [10.1109/5.726791](https://doi.org/10.1109/5.726791)]
- [155] Fei-Fei L, Fergus R, Perona P. Learning generative visual models from few training examples: An incremental Bayesian approach tested on 101 object categories. Computer Vision and Image Understanding, 2007, 106(1): 59–70. [doi: [10.1016/j.cviu.2005.09.012](https://doi.org/10.1016/j.cviu.2005.09.012)]
- [156] Griffin G, Holub A, Perona P. Caltech-256 object category dataset. Technical Report 7694. Pasadena: California Institute of Technology, 2007.
- [157] Krizhevsky A. Learning multiple layers of features from tiny images. Technical Report, Toronto: University of Toronto, 2009.
- [158] Lin TY, Maire M, Belongie S, Hays J, Perona P, Ramanan D, Dollár P, Zitnick CL. Microsoft COCO: Common objects in context. In: Proc. of the 13th European Conf. on Computer Vision. Zurich: Springer, 2014. 740–755. [doi: [10.1007/978-3-319-10602-1_48](https://doi.org/10.1007/978-3-319-10602-1_48)]
- [159] Chua TS, Tang JH, Hong RC, Li HJ, Luo ZP, Zheng YT. NUS-WIDE: A real-world Web image database from national university of Singapore. In: Proc. of the 2009 Int'l Conf. on Image and Video Retrieval. Fira: ACM, 2009. 48. [doi: [10.1145/1646396.1646452](https://doi.org/10.1145/1646396.1646452)]
- [160] Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma SA, Huang ZH, Karpathy A, Khosla A, Bernstein M, Berg AC, Fei-Fei L. ImageNet large scale visual recognition challenge. Int'l Journal of Computer Vision, 2015, 115(3): 211–252. [doi: [10.1007/s11263-015-0816-y](https://doi.org/10.1007/s11263-015-0816-y)]
- [161] Torralba A, Fergus R, Freeman WT. 80 million tiny images: A large data set for nonparametric object and scene recognition. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2008, 30(11): 1958–1970. [doi: [10.1109/TPAMI.2008.128](https://doi.org/10.1109/TPAMI.2008.128)]
- [162] Liu XB, Nie XS, Zhou Q, Xi XM, Zhu L, Yin YL. Supervised short-length hashing. In: Proc. of the 28th Int'l Joint Conf. on Artificial Intelligence. Macao: AAAI, 2019. 3031–3037.
- [163] Liu XB, Nie XS, Dai Q, Huang YP, Lian L, Yin YL. Reinforced short-length hashing. IEEE Trans. on Circuits and Systems for Video Technology, 2021, 31(9): 3655–3668. [doi: [10.1109/TCSVT.2020.3040863](https://doi.org/10.1109/TCSVT.2020.3040863)]
- [164] Zhou Q, Nie XS, Shi Y, Liu XB, Yin YL. Deep multi-region hashing. In: Proc. of the 2020 IEEE Int'l Conf. on Acoustics, Speech and Signal Processing. Barcelona: IEEE, 2020. 1938–1942. [doi: [10.1109/ICASSP40776.2020.9054178](https://doi.org/10.1109/ICASSP40776.2020.9054178)]
- [165] Chen Y, Zhang H, Tian ZB, Wang J, Zhang D, Li XL. Enhanced discrete multi-modal hashing: More constraints yet less time to learn. IEEE Trans. on Knowledge and Data Engineering, 2022, 34(3): 1177–1190. [doi: [10.1109/TKDE.2020.2995195](https://doi.org/10.1109/TKDE.2020.2995195)]
- [166] Lian DF, Xie X, Chen EH. Discrete matrix factorization and extension for fast item recommendation. IEEE Trans. on Knowledge and Data Engineering, 2021, 33(5): 1919–1933. [doi: [10.1109/TKDE.2019.2951386](https://doi.org/10.1109/TKDE.2019.2951386)]
- [167] Liu HW, Li XL, Zhang SC, Tian Q. Adaptive hashing with sparse matrix factorization. IEEE Trans. on Neural Networks and Learning Systems, 2020, 31(10): 4318–4329. [doi: [10.1109/TNNLS.2019.2954856](https://doi.org/10.1109/TNNLS.2019.2954856)]
- [168] Zheng CQ, Zhu L, Cheng ZY, Li JJ, Liu AA. Adaptive partial multi-view hashing for efficient social image retrieval. IEEE Trans. on Multimedia, 2021, 23: 4079–4092. [doi: [10.1109/TMM.2020.3037456](https://doi.org/10.1109/TMM.2020.3037456)]
- [169] Liu X, Hu ZK, Ling HB, Cheung YM. MTFH: A matrix tri-factorization hashing framework for efficient cross-modal retrieval. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2021, 43(3): 964–981. [doi: [10.1109/TPAMI.2019.2940446](https://doi.org/10.1109/TPAMI.2019.2940446)]
- [170] Chen ZD, Li CX, Luo X, Nie LQ, Zhang W, Xu XS. SCRATCH: A scalable discrete matrix factorization hashing framework for cross-modal retrieval. IEEE Trans. on Circuits and Systems for Video Technology, 2020, 30(7): 2262–2275. [doi: [10.1109/TCSVT.2019.2911359](https://doi.org/10.1109/TCSVT.2019.2911359)]
- [171] Hong WX, Meng JJ, Yuan JS. Distributed composite quantization. In: Proc. of the 32nd AAAI Conf. on Artificial Intelligence. New Orleans: AAAI, 2018. 61–68. [doi: [10.1609/aaai.v32i1.11294](https://doi.org/10.1609/aaai.v32i1.11294)]
- [172] Xu D, Tsang IW, Zhang Y. Online product quantization. IEEE Trans. on Knowledge and Data Engineering, 2018, 30(11): 2185–2198. [doi: [10.1109/TKDE.2018.2817526](https://doi.org/10.1109/TKDE.2018.2817526)]
- [173] Liu Q, Zhang J, Lian DF, Ge Y, Ma JH, Chen EH. Online additive quantization. In: Proc. of the 27th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. Singapore: ACM, 2021. 1098–1108. [doi: [10.1145/3447548.3467441](https://doi.org/10.1145/3447548.3467441)]
- [174] He ZY, Feng XP, Liu LJ, Huang QS. Deep highly interrelated hashing for fast image retrieval. Journal of Computer Research and Development, 2020, 57(11): 2375–2388 (in Chinese with English abstract). [doi: [10.7544/issn1000-1239.2020.20190498](https://doi.org/10.7544/issn1000-1239.2020.20190498)]
- [175] Tu RC, Mao XL, Guo JN, Wei W, Huang HY. Partial-Softmax loss based deep hashing. In: Proc. of the 2021 World Wide Web Conf. Ljubljana: ACM, 2021. 2869–2878. [doi: [10.1145/3442381.3449825](https://doi.org/10.1145/3442381.3449825)]
- [176] Li ZM, Zhang P, Liu YJ, Li H. Deep hashing retrieval algorithm combing attention model and bimodal Gaussian distribution. Journal of

- Computer-aided Design & Computer Graphics, 2020, 32(5): 759–768 (in Chinese with English abstract). [doi: 10.3724/SP.J.1089.2020.17889]
- [177] Lin MB, Ji RR, Liu H, Wu YJ. Supervised online hashing via Hadamard codebook learning. In: Proc. of the 26th Int'l ACM Conf. on Multimedia. Seoul: ACM, 2018. 1635–1643. [doi: 10.1145/3240508.3240519]
- [178] Lin MB, Ji RR, Chen S, Sun XS, Lin CW. Similarity-preserving linkage hashing for online image retrieval. IEEE Trans. on Image Processing, 2020, 29: 5289–5300. [doi: 10.1109/TIP.2020.2981879]
- [179] Wang D, Wang Q, An YQ, Gao XB, Tian YM. Online collective matrix factorization hashing for large-scale cross-media retrieval. In: Proc. of the 43rd Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. Xi'an: ACM, 2020. 1409–1418. [doi: 10.1145/3397271.3401132]
- [180] Tian X, Ng WWY, Wang H. Concept preserving hashing for semantic image retrieval with concept drift. IEEE Trans. on Cybernetics, 2021, 51(10): 5184–5197. [doi: 10.1109/TCYB.2019.2955130]
- [181] Shen Y, Sun XH, Wei XS, Jiang QY, Yang J. SEMICON: A learning-to-hash solution for large-scale fine-grained image retrieval. In: Proc. of the 17th European Conf. on Computer Vision. Tel Aviv: Springer, 2022. 531–548. [doi: 10.1007/978-3-031-19781-9_31]
- [182] Dong Z, Pei MT. Cross-modality face retrieval based on heterogeneous hashing network. Chinese Journal of Computers, 2019, 42(1): 73–84 (in Chinese with English abstract). [doi: 10.11897/SP.J.1016.2019.00073]
- [183] Yang EK, Liu MX, Yao DR, Cao B, Lian CF, Yap PT, Shen DG. Deep Bayesian hashing with center prior for multi-modal neuroimage retrieval. IEEE Trans. on Medical Imaging, 2021, 40(2): 503–513. [doi: 10.1109/TMI.2020.3030752]
- [184] Li P, Han LR, Tao XW, Zhang XY, Grecos C, Plaza A, Ren P. Hashing nets for hashing: A quantized deep learning to hash framework for remote sensing image retrieval. IEEE Trans. on Geoscience and Remote Sensing, 2020, 58(10): 7331–7345. [doi: 10.1109/TGRS.2020.2981997]
- [185] Wang YX, Chen ZD, Luo X, Li R, Xu XS. Fast cross-modal hashing with global and local similarity embedding. IEEE Trans. on Cybernetics, 2022, 52(10): 10064–10077. [doi: 10.1109/TCYB.2021.3059886]
- [186] Liong VE, Lu JW, Duan LY, Tan YP. Deep variational and structural hashing. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2020, 42(3): 580–595. [doi: 10.1109/TPAMI.2018.2882816]
- [187] Liu XB, Nie XS, Yin YL. Mutual linear regression based supervised discrete cross-modal hashing. Journal of Computer Research and Development, 2020, 57(8): 1707–1714 (in Chinese with English abstract). [doi: 10.7544/issn1000-1239.2020.20200122]
- [188] Wang YM, Xue B, Cheng Q, Chen YH, Zhang LJ. Deep unified cross-modality hashing by pairwise data alignment. In: Proc. of the 30th Int'l Joint Conf. on Artificial Intelligence. Montreal: IJCAI.org, 2021. 1129–1135.

附中文参考文献:

- [3] 中国互联网络信息中心. 第49次《中国互联网络发展状况统计报告》. <https://www.cnnic.cn/n4/2022/0401/c88-1131.html>, 2022-02-25.
- [7] 国务院. 国务院关于印发新一代人工智能发展规划的通知. 2017. https://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm
- [8] 中华人民共和国中央人民政府. 政府工作报告. 2022. http://www.gov.cn/premier/2022-03/12/content_5678750.htm
- [10] 中国信息通信研究院. 人工智能白皮书(2022年). 北京: 中国信息通信研究院, 2022.
- [15] 李松, 胡晏铭, 郝晓红, 张丽平, 郝忠孝. 基于维度分组降维的高维数据近似 k 近邻查询. 计算机研究与发展, 2021, 58(3): 609–623. [doi: 10.7544/issn1000-1239.2021.20200285]
- [18] 李武军, 周志华. 大数据哈希学习: 现状与趋势. 科学通报, 2015, 60(5): 485–490. [doi: 10.1360/N972014-00841]
- [62] 李长升, 闵齐星, 成雨蓉, 袁野, 王国仁. 捕获局部语义结构和实例辨别的无监督哈希. 软件学报, 2021, 32(3): 742–752. <http://www.jos.org.cn/1000-9825/6178.htm> [doi: 10.13328/j.cnki.jos.006178]
- [109] 费伦科, 秦建阳, 滕少华, 张巍, 刘冬宁, 侯艳. 近似最近邻大数据检索哈希散列方法综述. 广东工业大学学报, 2020, 37(3): 23–35. [doi: 10.12052/gdutxb.190123]
- [133] 聂秀山, 刘兴波, 袁肖明, 尹义龙. 基于相似度驱动的线性哈希模型参数再优化方法. 软件学报, 2020, 31(4): 1039–1050. <http://www.jos.org.cn/1000-9825/5918.htm> [doi: 10.13328/j.cnki.jos.005918]
- [174] 贺周雨, 冯旭鹏, 刘利军, 黄青松. 面向大规模图像检索的深度强相关散列学习方法. 计算机研究与发展, 2020, 57(11): 2375–2388. [doi: 10.7544/issn1000-1239.2020.20190498]
- [176] 李宗民, 张鹏, 刘玉杰, 李华. 结合注意力模型与双峰高斯分布的深度哈希检索算法. 计算机辅助设计与图形学学报, 2020, 32(5): 759–768. [doi: 10.3724/SP.J.1089.2020.17889]

- [182] 董震, 裴明涛. 基于异构哈希网络的跨模态人脸检索方法. 计算机学报, 2019, 42(1): 73–84. [doi: [10.11897/SP.J.1016.2019.00073](https://doi.org/10.11897/SP.J.1016.2019.00073)]
- [187] 刘兴波, 聂秀山, 尹义龙. 基于双向线性回归的监督离散跨模态散列方法. 计算机研究与发展, 2020, 57(8): 1707–1714. [doi: [10.7544/issn1000-1239.2020.20200122](https://doi.org/10.7544/issn1000-1239.2020.20200122)]



张雪凝(1999—), 女, 硕士生, CCF 学生会员, 主要研究领域为哈希学习, 弱监督学习.



聂秀山(1981—), 男, 博士, 教授, 博士生导师, CCF 杰出会员, 主要研究领域为多媒体检索, 计算机视觉, 机器学习, 数据挖掘.



刘兴波(1995—), 男, 博士, 教授, 博士生导师, CCF 专业会员, 主要研究领域为多媒体检索, 计算机视觉.



王少华(1979—), 女, 高级工程师, CCF 专业会员, 主要研究领域为模式识别, 智能系统, 哈希学习.



宋井宽(1986—), 男, 博士, 教授, 博士生导师, CCF 专业会员, 主要研究领域为大规模多媒体检索, 基于哈希的图像(或视频)分割与图像(或视频)理解, 图学习, 深度学习.



尹义龙(1972—), 男, 博士, 教授, 博士生导师, CCF 杰出会员, 主要研究领域为机器学习, 数据挖掘, 计算机视觉.