



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2020
07
VOL.25

ISSN1006-8961
CN11-3758/TB





第25卷第7期（总第291期）
2020年7月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院
主办单位 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

主 编 顾行发
编辑出版 《中国图象图形学报》编辑出版委员会
通信地址 北京市海淀区北四环西路19号
邮 编 100190
电子信箱 jig@radi.ac.cn
电 话 010-58887035
网 址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218号
总 发 行 北京报刊发行局
订 购 全国各地邮局
海外发行 中国国际图书贸易集团有限公司
(邮政信箱: 北京399信箱 邮编: 100048)
印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian | Monthly, Started in 1996

Superintended by Chinese Academy of Sciences
Sponsored by Institute of Remote Sensing and Digital Earth, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Gu Xingfa
Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics
Address No. 19, North 4th Ring Road West, Haidian District, Beijing, P. R. China
Zip code 100190
E-mail jig@radi.ac.cn
Telephone 010-58887035
Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB
ISSN 1006-8961
CODEN ZTTXFZ

国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



非真实感绘制技术研究现状与展望(第1283页)



图卷积网络下牙齿种子点自动选取(第1481页)



迁移学习下高分快视数据道路快速提取(第1501页)

学者观点

非真实感绘制技术研究现状与展望

钱文华, 曹进德, 徐丹, 吴昊 1283

综述

深度哈希图像检索方法综述

刘颖, 程美, 王富平, 李大湘, 刘伟, 范九伦 1296

严肃游戏中虚拟角色行为建模综述

刘翠娟, 刘箴, 柴艳杰, 刘婷婷, 陈效奕 1318

智能制造中的计算机视觉应用瓶颈问题

雷林建, 孙胜利, 向玉开, 张悦, 刘会凯 1330

图像处理和编码

浅层CNN网络构建的噪声比例估计模型

徐少平, 林珍玉, 李崇禧, 崔燕, 刘蕊蕊 1344

烧结断面火焰图像退化模型及断面图像复原

王福斌, 刘贺飞, 武晨 1356

gyrator变换域的高鲁棒多图像加密算法

王丰, 邵珠宏, 王云飞, 姚启钧, 刘西林 1366

基于透射率修正的湍流模型与动态调整retinex的水下图像增强

汤新雨, 李敏, 徐灵丽, 郝真, 张学武 1380

图像分析和识别

俯视深度头肩序列行人再识别

王新年, 刘春华, 齐国清, 张世强 1393

结合混合池化的双流人脸活体检测网络

汪亚航, 宋晓宁, 吴小俊 1408

增强型灰度图像空间实现虹膜活体检测

刘明康, 王宏民, 李琦, 孙哲南 1421

基于水动力学、分数阶微分及Steger算法的线状目标提取

陈卫卫, 王卫星, 李润青, 蔡晓琰, 郑思凡 1436

时域候选优化的时序动作检测

熊成鑫, 郭丹, 刘学亮 1447

图像理解和计算机视觉

视频中多特征融合人体姿态跟踪

马淼, 李贻斌, 武宪青, 高金凤, 潘海鹏 1459

计算机图形学

插值给定数据点的四次PH曲线构造

方林聪, 阳诚砖, 邸文钰, 刘芳 1473

医学图像处理

图卷积网络下牙齿种子点自动选取

李占利, 孙志浩, 李洪安, 刘童鑫 1481

乳腺超声图像中易混淆困难样本的分类方法

杜章锦, 龚勋, 罗俊, 章哲敏, 杨菲 1490

遥感图像处理

迁移学习下高分快视数据道路快速提取

张军军, 万广通, 张洪群, 李山山, 冯旭祥 1501

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



Research status and progress of non-photorealistic rendering technology(P1283)



Automatic selection of tooth seed point by graph convolutional network(P1481)



Rapid road extraction from quick view imagery of high-resolution satellites combined with transfer learning(P1501)

Scholar View

Research status and progress of non-photorealistic rendering technology

Qian Wenhua, Cao Jinde, Xu Dan, Wu Hao 1283

Review

Deep Hashing image retrieval methods

Liu Ying, Cheng Mei, Wang Fuping, Li Daxiang, Liu Wei, Fan Jiulun 1296

Virtual character behavior modeling in serious games: a review

Liu Cuijuan, Liu Zhen, Chai Yanjie, Liu Tingting, Chen Xiaoyi 1318

Bottleneck issues of computer vision in intelligent manufacturing

Lei Linjian, Sun Shengli, Xiang Yukai, Zhang Yue, Liu Huikai 1330

Image Processing and Coding

A shallow CNN-based noise ratio estimation model

Xu Shaoping, Lin Zhenyu, Li Chongxi, Cui Yan, Liu Ruirui 1344

Degradation model and restoration of flame image of sintering section

Wang Fubin, Liu Hefei, Wu Chen 1356

Multiple image encryption of high robustness in gyrator transform domain

Wang Feng, Shao Zhuhong, Wang Yunfei, Yao Qijun, Liu Xilin 1366

Underwater image enhancement based on turbulence model corrected by transmittance and dynamically adjusted retinex

Tang Xinyu, Li Min, Xu Lingli, Hao Zhen, Zhang Xuewu 1380

Image Analysis and Recognition

Person re-identification based on top-view depth head and shoulder sequence

Wang Xinnian, Liu Chunhua, Qi Guoqing, Zhang Shiqiang 1393

Two-stream face spoofing detection network combined with hybrid pooling

Wang Yahang, Song Xiaoning, Wu Xiaojun 1408

Enhanced gray-level image space for iris liveness detection

Liu Mingkan, Wang Hongmin, Li Qi, Sun Zhenan 1421

Linear object extraction based on hydrodynamics, fractional differential and Steger algorithm

Chen Weiwei, Wang Weixing, Li Runqing, Cai Xiaoyan, Zheng Sifan 1436

Temporal proposal optimization for temporal action detection

Xiong Chengxin, Guo Dan, Liu Xueliang 1447

Image Understanding and Computer Vision

Human pose tracking based on multi-feature fusion in videos

Ma Miao, Li Yibin, Wu Xianqing, Gao Jinfeng, Pan Haipeng 1459

Computer Graphics

Construction of quartic PH curves via interpolating given points

Fang Lincong, Yang Chengzhuan, Di Wenyu, Liu Fang 1473

Medical Image Processing

Automatic selection of tooth seed point by graph convolutional network

Li Zhanli, Sun Zhihao, Li Hongan, Liu Tongxin 1481

Classification method for samples that are easy to be confused in breast ultrasound images

Du Zhangjin, Gong Xun, Luo Jun, Zhang Zheming, Yang Fei 1490

Remote Sensing Image Processing

Rapid road extraction from quick view imagery of high-resolution satellites with transfer learning

Zhang Junjun, Wan Guangtong, Zhang Hongqun, Li Shanshan, Feng Xuxiang 1501

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2020)07-1296-22

论文引用格式: Liu Y, Cheng M, Wang F P, Li D X, Liu W and Fan J L. 2020. Deep Hashing image retrieval methods. Journal of Image and Graphics, 25(07): 1296-1317 (刘颖, 程美, 王富平, 李大湘, 刘伟, 范九伦. 2020. 深度哈希图像检索方法综述. 中国图象图形学报, 25(07): 1296-1317)
[DOI:10.11834/jlg.190518]

深度哈希图像检索方法综述

刘颖^{1,2,3}, 程美³, 王富平^{1,3}, 李大湘^{1,3}, 刘伟^{1,3}, 范九伦^{1,3}

1. 电子信息现场勘验应用技术公安部重点实验室, 西安 710121; 2. 无线通信与信息处理技术
国际联合研究中心, 西安 710121; 3. 西安邮电大学图像与信息处理研究所, 西安 710121

摘要: 随着网络上图像和视频数据的快速增长, 传统图像检索方法已难以高效处理海量数据。在面向大规模图像检索时, 特征哈希与深度学习结合的深度哈希技术已成为发展趋势, 为全面认识和理解深度哈希图像检索方法, 本文对其进行梳理和综述。根据是否使用标签信息将深度哈希方法分为无监督、半监督和监督深度哈希方法, 根据无监督和半监督深度哈希方法的主要研究点进一步分为基于卷积神经网络(convolutional neural networks, CNN)和基于生成对抗网络(generative adversarial networks, GAN)的无监督/半监督深度哈希方法, 根据数据标签信息差异将监督深度哈希方法进一步分为基于三元组和基于成对监督信息的深度哈希方法, 根据各种方法使用损失函数的不同对每类方法中一些经典方法的原理及特性进行介绍, 对各种方法的优缺点进行分析。通过分析和比较各种深度哈希方法在 CIFAR-10 和 NUS-WIDE 数据集上的检索性能, 以及深度哈希算法在西安邮电大学图像与信息处理研究所(Center for Image and Information Processing, CIIP)自建的两个特色数据库上的测试结果, 对基于深度哈希的检索技术进行总结, 分析了深度哈希的检索技术未来的发展前景。监督深度哈希的图像检索方法虽然取得了较高的检索精度。但由于监督深度哈希方法高度依赖数据标签, 无监督深度哈希技术更加受到关注。基于深度哈希技术进行图像检索是实现大规模图像数据高效检索的有效方法, 但存在亟待攻克的技术难点。针对实际应用需求, 关于无监督深度哈希算法的研究仍需要更多关注。

关键词: 图像检索; 无监督; 监督; 深度学习; 哈希; 深度哈希

Deep Hashing image retrieval methods

Liu Ying^{1,2,3}, Cheng Mei³, Wang Fuping^{1,3}, Li Daxiang^{1,3}, Liu Wei^{1,3}, Fan Jiulun^{1,3}

1. Key Laboratory of Electronic Information Application Technology for Scene Investigation, Ministry of Public Security, Xi'an 710121, China;
2. International Joint Research Center for Wireless Communication and Information Processing, Xi'an 710121, China;
3. Center for Image and Information Processing, Xi'an University of Posts and Telecommunications, Xi'an 710121, China

Abstract: The efficient processing of massive amounts of data obtained as a result of the rapid growth of image and video data transmission is becoming increasingly difficult for traditional image retrieval methods. The feature-Hashing technology, which can achieve efficient feature compression and fast feature matching and image retrieval, is introduced to address this issue. The deep learning technology also has unique advantages in feature extraction and compact description. The deep Hashing technology, which combines feature Hashing with deep learning, has become an interesting research topic in the

收稿日期: 2019-10-23; 修回日期: 2020-01-11; 预印本日期: 2020-01-18

基金项目: 国家重点研发计划项目(2017YFC080380); 国家自然科学基金项目(61671377, 61802305); 公安部科技强警项目(2016GABJC51)

Supported by: National Key Research and Development Program of China (2017YFC080380); National Natural Science Foundation of China (61671377, 61802305); Science and Technology Project Fund under Ministry of Public Security of China (2016GABJC51)

area of large-scale image retrieval in solving the problem of large-scale image retrieval. Image retrieval methods based on deep Hashing have attracted increasing attention. Extensive research on image retrieval technologies using deep Hashing has been conducted in recent years and is reported in this paper. First, the deep Hashing method is divided into unsupervised, semisupervised, and supervised deep Hashing methods according to whether label information is used. Second, unsupervised and semisupervised deep Hashing methods are further divided into two types, namely, unsupervised/semisupervised deep Hashing based on deep network models and GANs (generative adversarial networks). In the unsupervised deep Hashing based on the deep network models, the DeepBit algorithm and the SADH (similarity-adaptive deep Hashing) algorithm are mainly introduced. In the GAN-based unsupervised deep Hashing method, we illustrate the principles of HashGAN, BGAN (binary generative adversarial networks) and PGH (progressive generative Hashing) algorithms. In the semi-supervised deep Hashing method, the SSDH (semi-supervised discriminant Hashing) algorithm based on the depth models and the SSGAH (semi-supervised generative adversarial Hashing) algorithm based on the generated adversarial network are mainly interpreted. Third, the supervised deep Hashing algorithms are divided into deep Hashing methods based on triple labels and data pairs depending on the different types of label information used. Designing loss functions and controlling quantization errors occupies important parts of deep Hashing image retrieval, hence the algorithms are classified in more detail according to different loss functions in several supervised deep Hashing methods. In the deep Hashing methods based on paired supervision information, the algorithm are further classified as deep Hashing methods using square loss function, using cross-entropy loss function, or designing a new loss function. 1) In the Hash method using the square loss function, CNNH (convolutional neural network Hashing) is introduced in detail. 2) In the Hash method using the cross entropy loss function, we mainly describe DPSH (deep supervised Hashing with pairwise labels), DSDH (deep supervised discrete Hashing), HashNet and HashGAN four algorithm models. 3) DSH (deep supervised Hashing) and DVSQ (deep visual-semantic quantization) algorithms design new loss functions in their research. Among the deep Hashing methods based on triple labels, 1) deep Hashing methods using triple loss function are mainly illustrated: NINH (network in network Hashing), DRSCH (deep regularized similarity comparison Hashing), DTQ (deep triplet quantization). And the triple loss function is actually improved from the hinge loss function; 2) the deep Hashing methods using the triple entropy loss function: DTSH (deep supervised Hashing with triplet). Because triple labels require a lot of image preprocessing, there is little research about it. After introducing principles and characteristics of selected classical algorithms, and the advantages and disadvantages of each deep supervised algorithm are analyzed. Fourth, we compare the retrieval performances of each algorithm on two commonly used large-scale datasets, namely, CIFAR-10 and NUS-WIDE. We also investigate the performance of the DPSH algorithm on two specialized datasets, namely, CIIP (Center for Image and Information Processing)-CSID (crime scene investigation image database) and CIIP-TPID (tread pattern image dataset), and summarize existing deep Hashing-based retrieval technologies. Finally, we discuss the future development of deep Hashing-based retrieval algorithms. Hashing has improved the image retrieval speed on very-large-scale datasets, but the overall retrieval performance remains low. Hashing with deep learning has been extensively used in recent years to extract features of high-level semantic information. The CNNH algorithm is the first of such attempts. The excellent performance of CNNH has opened a new chapter for Hashing-based image retrieval methods. Deep Hashing methods based on paired supervisory information or triplet supervised information have caused improvements in algorithm structure, Hashing function, loss function, and control quantization error. However, the improvement of triple-based deep Hashing methods is limited by their requirement of numerous image-preprocessing works. Deep Hashing methods based on pairwise label information provide some insights into the way to enhance triple-based deep Hashing methods. For example, NINH improves the network structure from CNNH, and the DTSH algorithm is based on the algorithm structure of DPSH. Deep Hashing-based image retrieval methods have their own advantages in retrieval performance. The existing methods have achieved superior retrieval precision, but a space for improvement in controlling quantization error and learning image representation remains. Labeling images one by one will require high labor and time costs because the supervised deep Hashing method highly depends on data labels but the data scale in reality is expanding. Scholars have paid increasing attention to unsupervised deep Hashing technologies and achieved significant performance improvements by combining such technologies with GANs or deep mod-

els. Experimental results on two special databases show that the DPSH algorithm performs efficiently on CIIP-CSID and competitively on CIIP-TPID. The deep Hashing technology is an effective method for large-scale image retrieval, but major problems remain unsolved. On the basis of the needs of practical applications, the research on unsupervised deep Hashing algorithms requires further attention. Network models and feature learning should also be improved in different ways depending on dataset characteristics and case used. The potential applications of the deep Hashing technology are wide, including biometrics and multimodal retrieval. The experimental results of the DPSH algorithm on two special databases reveal the need to customize network models and feature-learning algorithms in accordance with their cases used. Such a need renders the deep Hashing technology for critical image retrieval research areas and presents a great potential for various specialized industries.

Key words: image retrieval; unsupervised; supervised; deep learning; Hashing; deep Hashing

0 引言

图像检索任务是在图像数据集中搜索与待查询图像在视觉或语义上相似的图像。假设数据库中的图像和待查询图像都用实值特征表示,那么图像检索过程就是计算待查询图像特征值与数据库中所有图像特征值之间的距离,距离最近的图像就是与待查询图像最相似的图像(Liu等,2016)。然而,随着因特网技术和多媒体的飞速发展,因特网上的图像视频数据量迅速增长,图像检索需要在海量图像数据中进行,面对百万级规模的图像数据库,这种遍历搜索方法效率极低,且浪费大量时间和巨大内存(Zhang等,2014)。

面对大规模图像数据库的检索需求,传统图像检索方法面临重大挑战(Hadsell等,2006),杨安邦和寿震宇(2018)总结了两类解决方法。第1类是基于树的检索方法,具有代表性的是基于k-d(k-dimensional)树(Meyer-Bäse等,2004)的算法及其改进算法,但在训练集特征数量大且维数高时,易发生过拟合,极易产生“维数灾难”(Friedman,1997),直接影响分类器训练和最终检索性能;第2类是基于哈希的图像检索方法,将图像的高维特征映射到二值空间,生成低维的二进制编码特征,极大降低了特征维数,避免了“维数灾难”问题,通过二进制模式匹配或汉明距离测量实现图像编码特征的快速搜索,大幅提高了检索效率,与基于树的检索方法相比,更适应大规模图像检索。

基于哈希的图像检索方法包括传统哈希算法和深度哈希算法。孙瑶(2018)按网络结构的不同将深度哈希算法分为3类,即基于RBM(restricted

Boltzmann machine)(Hinton和Salakhutdinov,2006)的深度哈希、基于多层前馈神经网络的深度哈希和基于卷积神经网络(convolutional neural network, CNN)的深度哈希,并对每一类中的代表性算法进行了简单介绍。李武军和周志华(2015)对大数据哈希学习的研究进展和发展趋势进行了介绍。本文首先对传统哈希技术进行简单回顾,然后针对是否使用监督信息,将深度哈希方法分为无监督、半监督和监督深度哈希方法,并按网络结构的不同对无监督和半监督深度哈希方法进一步分类,按各种方法使用损失函数的不同对监督深度哈希方法进一步分类,如图1所示。最后对各种哈希方法的检索性能进行对比与总结,并对深度哈希的未来提出展望。

1 传统基于哈希的图像检索方法

哈希方法大致分为两类,即数据无关方法和数据依赖方法(Li等,2016)。在数据无关的哈希方法中,模型中的哈希函数通常随机生成,且独立于任何训练数据,但检索性能的提高需要用哈希码的长度换取。数据依赖的哈希方法是Li等人(2016)提出的,试图从一些训练数据中学习哈希函数,称为学习哈希算法(Kong和Li,2012)。与数据无关的方法相比,学习哈希算法可以用更短的哈希码实现更高的准确性。因此,在实际应用中学习哈希算法比数据无关方法更流行。

传统基于哈希的图像检索方法主要使用手工设计的图像描述子特征(杨妹和陈宇,2018),如SIFT(scale-invariant feature transform)特征(Lowe,1999)、LBP(local binary pattern)特征(Ojala等,2002)、HOG(histogram of oriented gradient)特征

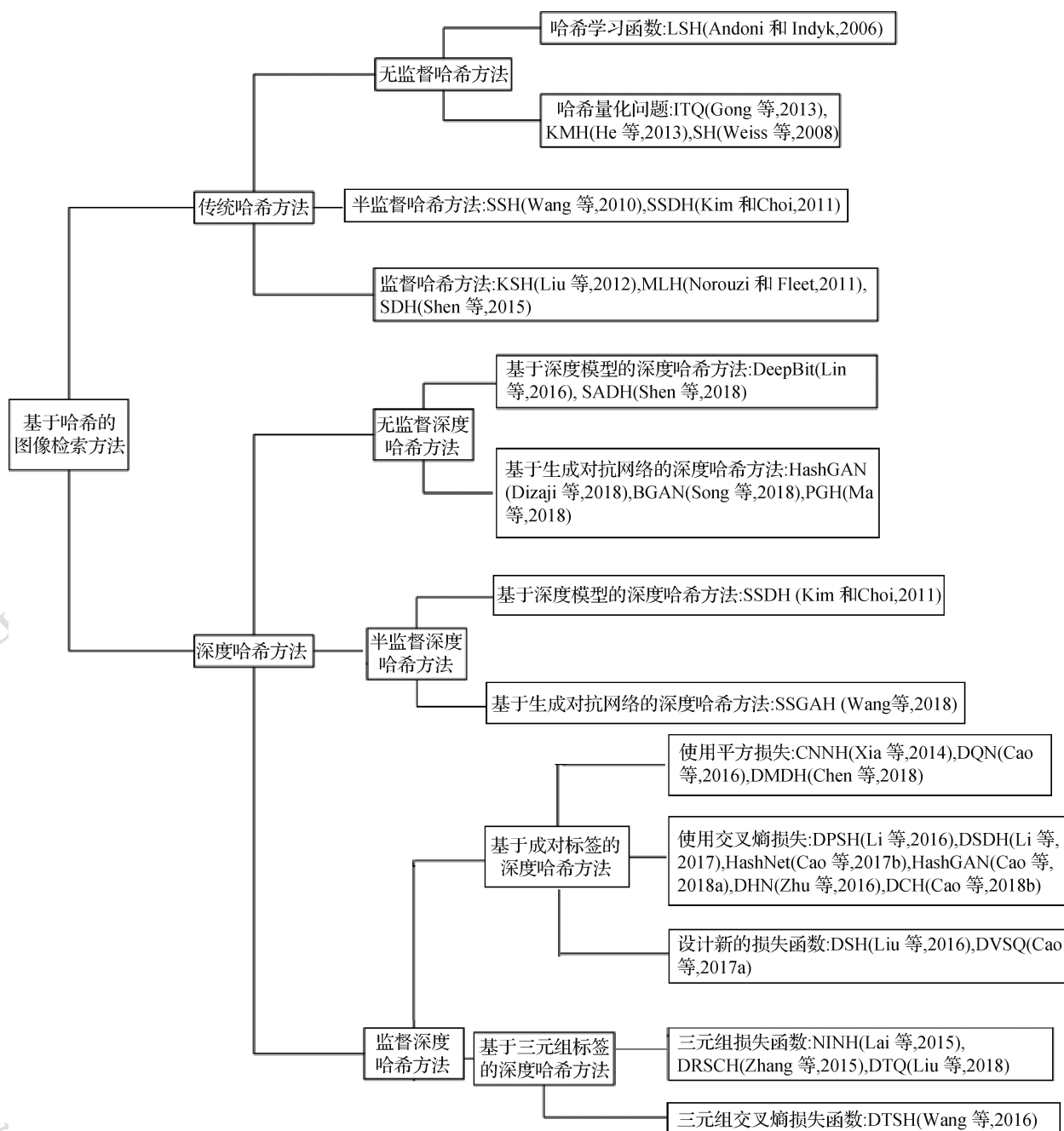


图1 基于哈希的图像检索方法

Fig. 1 Methods of image retrieval based on Hashing

(Dalal 和 Triggs, 2005)、GIST (generalized search trees) (Oliva 和 Torralba, 2001) 描述子等。根据是否使用标签信息, 传统哈希方法可分为有监督、无监督和半监督的哈希方法。

1.1 无监督哈希方法

无监督哈希方法不使用数据集标签信息, 仅依据图像数据学习哈希函数。传统无监督哈希方法的研究方向主要集中在哈希函数学习和哈希码量化

问题。

1.1.1 哈希函数学习

哈希函数学习本质上是一个分类问题, 通过创建线性哈希函数和非线性哈希函数进行哈希码学习。线性哈希函数一般采用线性判别分析、线性回归或线性支持向量机 (support vector machine, SVM) (Burges, 1998) 学习一系列超平面作为哈希映射函数。最经典的无监督线性哈希方法是局部敏感哈希

(locality sensitivity Hashing, LSH) (Andoni 和 Indyk, 2006), 使用随机映射产生哈希码。而 LSH 需要以码长为代价换取检索性能的提升, 因此 LSH 不能生成紧凑的哈希码。为此学者们提出了一系列改进方法, 产生了基于逻辑回归 (Liu 等, 2017) 和基于核函数 (Yang 等, 2015) 的非线性哈希学习模型。

1.1.2 哈希码量化问题

二进制哈希码量化问题主要通过控制量化误差提升检索性能。常用的二进制码量化方法使用“符号量化”, 将符号函数的返回值作为哈希码, 但通常需要对符号函数进行连续化松弛, 容易产生误差, 影响检索性能 (Wu 等, 2013)。

Gong 等人 (2013) 受多类谱聚类松弛解 (Yu 和 Shi, 2003) 离散优化算法的启发, 提出 ITQ (iterative quantization) 迭代量化方法, 减小了量化误差。He 等人 (2013) 提出了基于 K-means 的量化算法 KMH (K-means Hashing), 可以在进行 K-means 聚类的时候学习量化单元的索引值, 同时考虑量化及距离近似。与 ITQ 相比, KMH 可对超立方体进行伸缩, 实现更小的量化误差。Weiss 等人 (2008) 提出 SH (spectral Hashing), 将编码过程看做图分割过程, 对高维度数据集进行谱分析, 同时通过放松约束条件将问题转换为 graph Laplacians 的降维问题, 从而获得图像数据的哈希码, 在一定程度上提高了检索精度。与 ITQ 算法相比, SH 要求数据集必须服从均匀分布, 在实际中显然很难满足。Gray 和 Neuhoff, (1998) 提出基于向量量化技术的二值量化方法 (Jégou 等, 2011), 可以更好地控制量化误差, 实现较好的图像检索结果。

1.2 半监督哈希方法

半监督哈希方法同时运用少量的数据集标签和图像数据本身学习哈希函数。Wang 等人 (2010) 提出的半监督哈希 (semi-supervised Hashing, SSH) 算法属于经典的半监督哈希算法, 对有标签数据最小化经验误差, 正则化处理所有数据。Kim 和 Choi (2011) 基于线性判别分析的思想, 提出半监督判别哈希 (semi-supervised discriminant Hashing, SSDH), 最大化不同类之间的二进制编码的关联性, 将没有类标的数据作为正则项处理。

1.3 监督哈希方法

为了更充分地利用标签信息, 学者们提出了监督哈希方法, 得到了比无监督、半监督方法更好的检

索效果。

Liu 等人 (2012) 提出的基于核的哈希方法 (supervised Hashing with kernels, KSH)、Norouzi 和 Fleet (2011) 提出的最小损失哈希 (minimal loss Hashing, MLH) 均根据标签信息建立数据之间的相似关系, 并据此训练模型, 得到了较好的实验结果。其中, KSH 使用二进制代码的内积与汉明距离之间的等价关系优化模型, 构建紧凑的哈希码。MLH 则通过将铰链损失函数 (hinge loss function) (Norouzi 等, 2012) 最小化生成二进制哈希码。Shen 等人 (2015) 提出的 SDH (supervised discrete Hashing) 算法通过直接采用逐位优化哈希码的方法代替松弛方法, 获得了高质量哈希码。

1.4 传统哈希方法比较

在传统无监督、半监督、监督哈希方法中, 无监督哈希方法缺失了标签信息, 无法对图像进行准确描述, 但检索速度快, 且不需要图像标注, 因此更适用于实际应用。半监督哈希算法大多是基于图的算法, 存在的问题是建图代价高, 无法利用数据的天然图结构学习哈希函数。与无监督哈希方法相比, 由于半监督哈希方法运用了一部分标签信息, 检索性能有一定提升。相较于无监督和半监督哈希方法, 监督哈希方法利用图像数据的标签信息获得了更好的检索性能。但由于传统的监督哈希方法使用手工特征, 并将特征提取与哈希函数学习分开进行, 导致检索效果并不理想, 同时由于监督哈希需要对每幅图像进行标注, 耗费了巨大的人工标注资源, 因此传统的监督哈希方法的研究多局限在理论层面。

传统基于哈希的图像检索主要通过改进哈希函数、相似性度量准则和损失函数等方法提高检索精度。面对大规模数据集时, 此类方法在降低存储空间和加快检索速度方面具有较强的优越性。但由于算法对底层特征描述不足, 难以表达图像丰富的语义信息, 致使传统哈希方法的检索性能一直未有较大突破。

2 基于深度哈希的图像检索方法

在 2012 年 ImageNet 图像分类竞赛中, 深度学习网络 AlexNet (Simonyan 和 Zisserman, 2014) 表现优异, 分类精度远超传统方法, 从此基于 CNN 的图像检索逐渐成为一种趋势。然而在面对大规模数据

集时,CNN提取的特征虽然语义丰富,但需要付出高昂的时间成本,而基于传统哈希的检索具有速度快且占用内存少的优点,研究者将哈希与CNN融合,产生了基于深度哈希的图像检索技术,并按照是否使用标签信息,分为无监督、半监督和监督深度哈希方法,且对各种方法进行了更细致的分类。

2.1 无监督深度哈希方法

传统的非线性无监督哈希方法在一定程度上提高了预测哈希码的能力,但仍不能满足实际应用的需求,而具有强大拟合能力的深度网络引起了研究者的关注,提出了无监督深度哈希。在早期的无监督深度哈希模型中,Wang等人(2015)利用由多个分离的DBN(deep belief network)及一个RBM构成的MDBN(multimode deep belief network)学习多模式数据的紧凑哈希码。随着深度学习技术的不断成熟,无监督深度哈希方法主要通过CNN模型或生成对抗网络(generative adversarial networks, GAN)(Goodfellow等,2014)提升检索性能。

2.1.1 基于CNN模型的无监督深度哈希方法

随着深度学习技术的飞速发展,涌现了 Alex-

Net、VGG(visual geometry group)(Simonyan和Zisserman,2014)、GoogLeNet(Szegedy等,2015)、ResNet(He等,2016)等一系列性能优异的CNN模型。Lin等人(2016)和Shen等人(2018)利用VGG-16模型极大提升了无监督哈希学习方法的性能。

Lin等人(2016)开发了一种深度神经网络DeepBit,以无监督的方式和非线性哈希函数学习二进制哈希码,通过最小化量化误差、强制二进制代码均匀分布优化网络参数,并通过位与位不相关的哈希码从输入图像中捕获更多不相关的信息。

Shen等人(2018)提出了一种无监督深度哈希框架SADH(similarity-adaptive deep Hashing),在3个主要组件上进行了交替优化:1)基于ADMM(alternating direction method of multipliers)(Wei和Ozdaglar,2012)的二进制代码改良,为具有二值约束的哈希模型提供了通用的求解器;2)深度哈希模型训练,用于与数据相似度计算有关的其他哈希问题;3)使用学习到的模型进行数据相似度图更新,有效保留数据相似度,优化学习到的二进制哈希码。算法结构如图2所示。

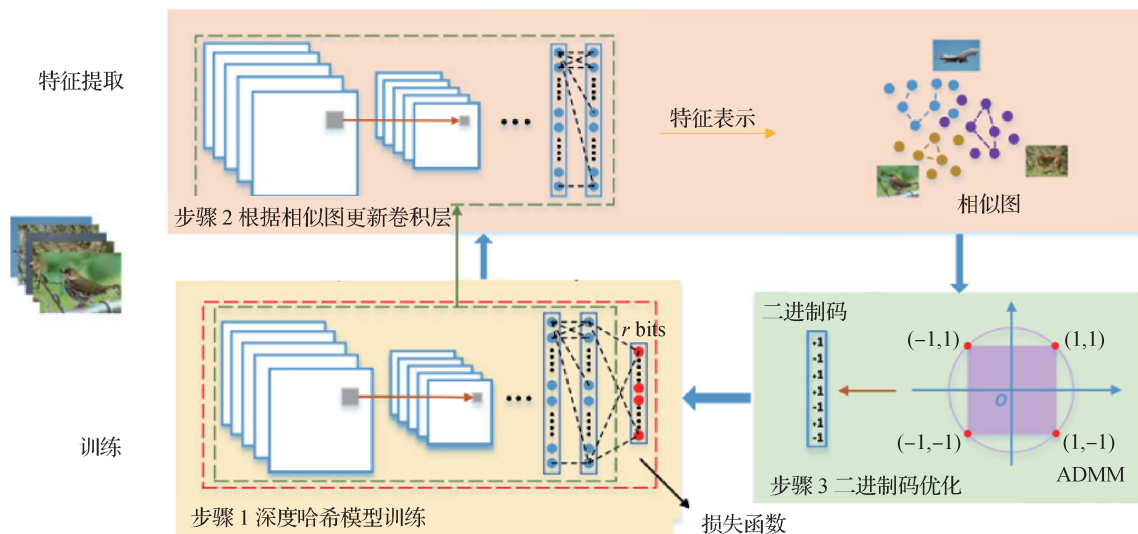


图2 SADH算法结构(Shen等,2018)

Fig. 2 Architecture of SADH algorithm(Shen et al., 2018)

2.1.2 基于GAN的无监督深度哈希方法

为了使训练样本更具多样性,使用GAN生成合成样本,进一步提高模型的泛化能力。

Dizaji等人(2018)提出一种新的深度非监督哈希框架HashGAN,由生成器、鉴别器和编码器组成。首先定义一个新的目标函数,有效训练深度哈希函

数。然后使鉴别器和编码器共享网络参数,降低对抗损失,并将对抗损失作为依赖数据的正则化进行哈希函数的无监督学习。最后设计新的损失函数从真实图像学习到量化、平衡、一致和独立的哈希码。此外,引入了协作损失,在训练哈希函数时使用合成图像。

Song 等人(2018)提出的 BGAN(binary generative adversarial networks)深度框架解决了两个在图像散列中未能解决的问题:1)提出两个等效但平滑的符号激活函数,并设计一种学习策略,解决了不经连续变量二进制化直接生成二进制代码的问题;2)设计一个新的损失函数,包括对抗损失、内容损失和邻域结构损失,解决除了通过 GAN 合成生成图以外,使用其他方法为二进制哈希码配备准确的图像检索功能的问题。因此,与无监督哈希方法相比, BGAN 的检索性能提高了两倍。

Ma 等人(2018)结合 GAN 网络设计一种新的无监督渐进式生成哈希网络(progressive generative Hashing, PGH),将哈希码视为生成相似图像的一种语义条件,将原始图像及其哈希码输入到 GAN 中,利用真实图像与合成图像进一步训练基于三元组损失的深度哈希网络。通过将学习到的二进制哈希码逐步馈入网络, PGH 可以逐渐获得更好的深度哈希函数。PGH 的算法结构如图 3 所示。

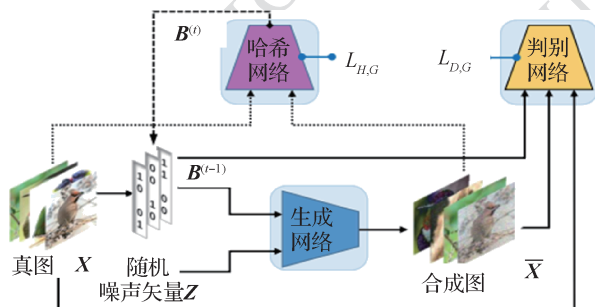


图 3 PGH 算法结构(Ma 等, 2018)

Fig. 3 Architecture of PGH algorithm(Ma et al., 2018)

2.2 半监督深度哈希方法

深度学习的快速发展使得学者开始尝试将半监督哈希方法与深度学习相结合。与无监督深度哈希方法类似,半监督深度哈希方法包括基于 CNN 模型的半监督深度哈希方法和基于 GAN 的半监督深度哈希方法。

2.2.1 基于 CNN 模型的半监督深度哈希方法

Zhang 和 Peng(2019)提出了大规模图像检索下的半监督深度哈希方法(semi-supervised deep Hashing, SSDH),提出半监督损失,通过最小化有标签数据的经验误差以及有标签和无标签数据的嵌入误差,保留图像间的语义相似性。设计半监督式深度哈希网络,通过利用有标签和无标签的数据,结合提出的“在线图构建方法”,在训练过程中获取与数据

相关的深度特征。SSDH 是第 1 个以半监督方式同时执行哈希码学习和特征学习的深度哈希方法。

2.2.2 基于 GAN 的半监督深度哈希方法

在无监督深度哈希与 GAN 结合取得较好检索效果的基础上, Wang 等人(2018)将半监督深度哈希与 GAN 网络结合,提出了 SSGAH(semi-supervised generative adversarial Hashing)算法。首先,在使用三元组标签信息和未标记数据的框架中统一生成模型、判别模型和深度哈希模型。其次,设计了生成模型和判别模型的新颖结构,以半监督的方式学习三重态信息的分布。此外,提出了半监督排序损失和对手排序损失,学习保留标记数据和未标记数据的语义相似性的二进制代码。最后,以对抗训练的方式优化整个模型,学习到的二进制代码可以捕获所有数据更好的语义信息。SSGAH 的算法结构如图 4 所示。

2.3 监督深度哈希方法

与无监督深度哈希方法相比,监督哈希方法应用语义标签先验信息训练网络模型,可以获得更好的性能。深度哈希中监督信息通常以成对或三元组的形式给出,按监督信息不同,基于深度哈希的图像检索方法分为基于成对监督信息的深度哈希方法和基于三元组监督信息的深度哈希方法。

深度哈希方法通过损失函数保留图像间的相似性,在生成二进制哈希码时,大部分首先学习连续的图像表示,再通过分离的二进制化步骤将连续的图像表示量化为离散的二进制码,该过程会产生量化损失。设计损失函数并控制量化误差是深度哈希中的重要一环,按损失函数的不同,可对监督深度哈希方法进行更细致的分类。

2.3.1 基于成对监督信息的深度哈希方法

基于成对监督信息的深度哈希方法通过包含相似信息的标签对生成相似性矩阵,并以此描述图像对之间的相似性信息,利用相似性信息学习哈希函数(Lin 等, 2015),主要使用平方损失函数和交叉熵损失函数。按损失函数的不同,基于成对监督信息的深度哈希方法分为使用平方损失、使用交叉熵损失、设计新损失函数的深度哈希方法。

1)使用平方损失函数的深度哈希方法。Xia 等人(2014)提出了 CNNH(convolutional neural network Hashing)算法,通过两阶段法将 CNN 与哈希编码融合,第 1 阶段学习哈希编码,第 2 阶段通过训练 CNN

以输出连续的哈希码。CNNH 算法通过 CNN 自动学习图像特征和一组哈希函数,对二进制代码进行拟合,检索性能取得了显著提升。但该框架输入的

除了原始图像数据外,还包括第1阶段学习到的哈希码,不是端到端的方法,不能完全发挥 CNN 的学习能力。算法结构如图5所示。

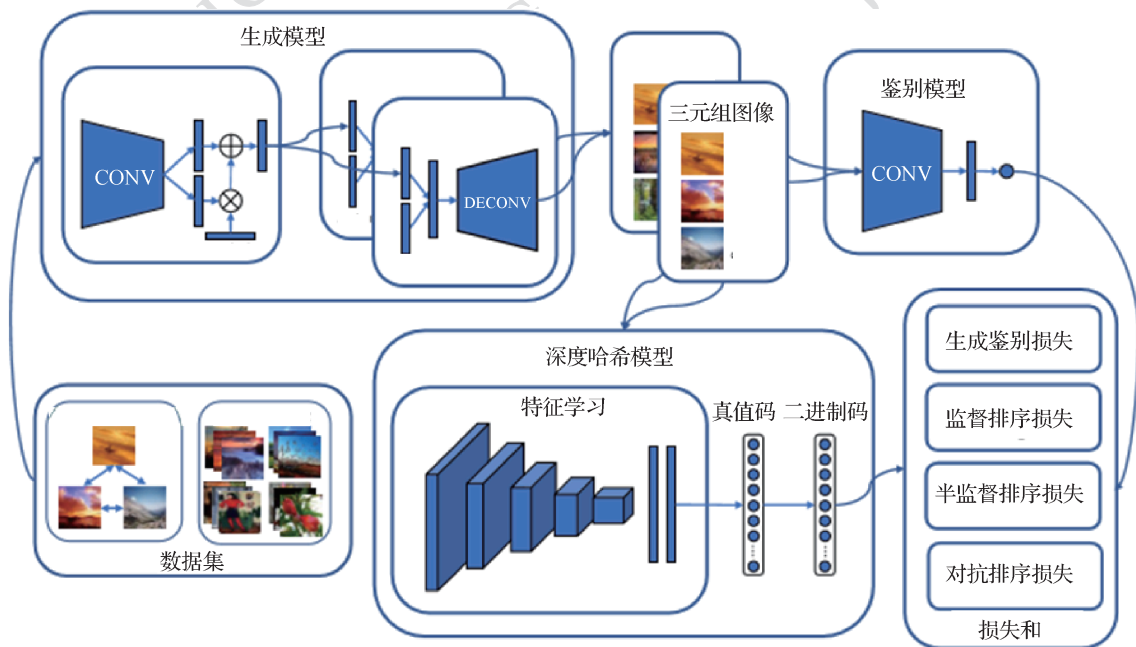


图4 SSGAH 算法结构(Wang 等, 2018)

Fig. 4 Architecture of SSGAH algorithm(Wang et al. , 2018)

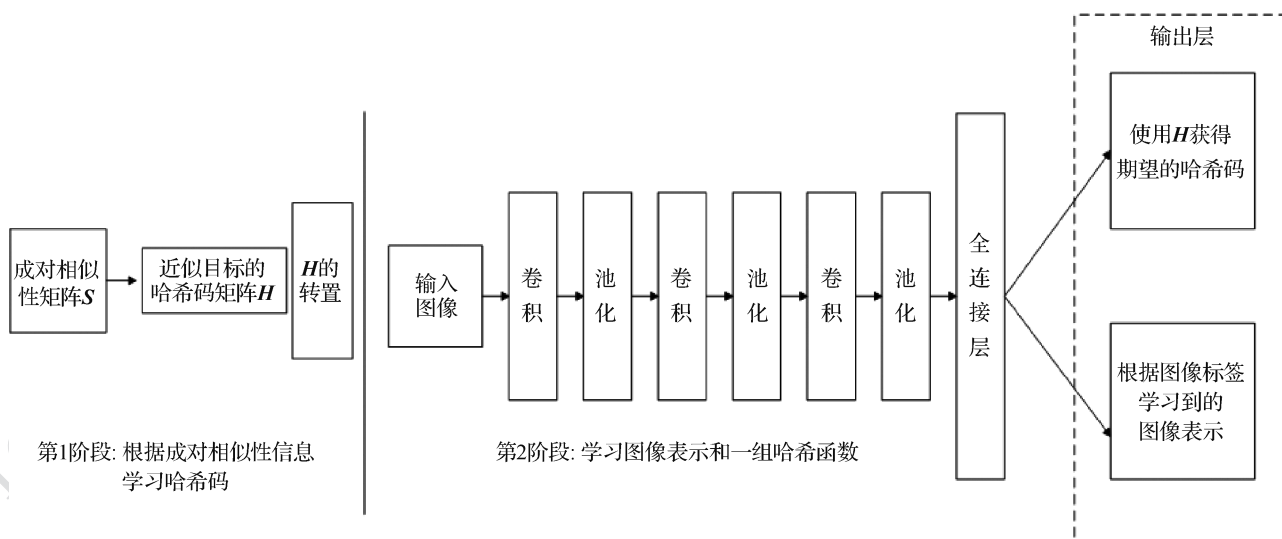


图5 CNNH 算法结构(Xia 等, 2014)

Fig. 5 Architecture of CNNH algorithm(Xia et al. , 2014)

除 CNNH 算法外,Cao 等人(2016)提出的 DQN (deep quantization network)、Chen 等人(2018)提出的差异最小化模型 DMDH(deep Hashing via discrepancy minimization)通过将二元离散哈希码松弛为连续变量,适应基于梯度的哈希函数学习。

2) 使用交叉熵损失函数的深度哈希方法。

DPSH(deep supervised Hashing with pairwise labels)是 Li 等人(2016)提出的一种新的深度哈希算法,由顶部 CNN-F(fast convolutional neural networks)(Chatfield 等, 2014)和底部 CNN-F 组成,两个 CNN-F 具有相同的结构和权重,并以一对图像作为模型的输入。CNN-F 由 5 个卷积层和 2 个全连

接层组成,用于对成对图像数据同时进行特征学习和哈希码学习。DPSH 包含 3 个关键组件:1) 深度神经网络,用于学习图像精确表示;2) 哈希函数,用于将学习到的图像标签映射到哈希码;3) 损失函数,用于衡量基于成对标签引导的哈希码质

量,同时引入正则化项减小量化误差。DPSH 算法以端到端的方式将图像从像素映射到成对标签,且不同的组件可以在 DPSH 中相互提供反馈,使得检索性能得到显著提升。算法结构如图 6 所示。

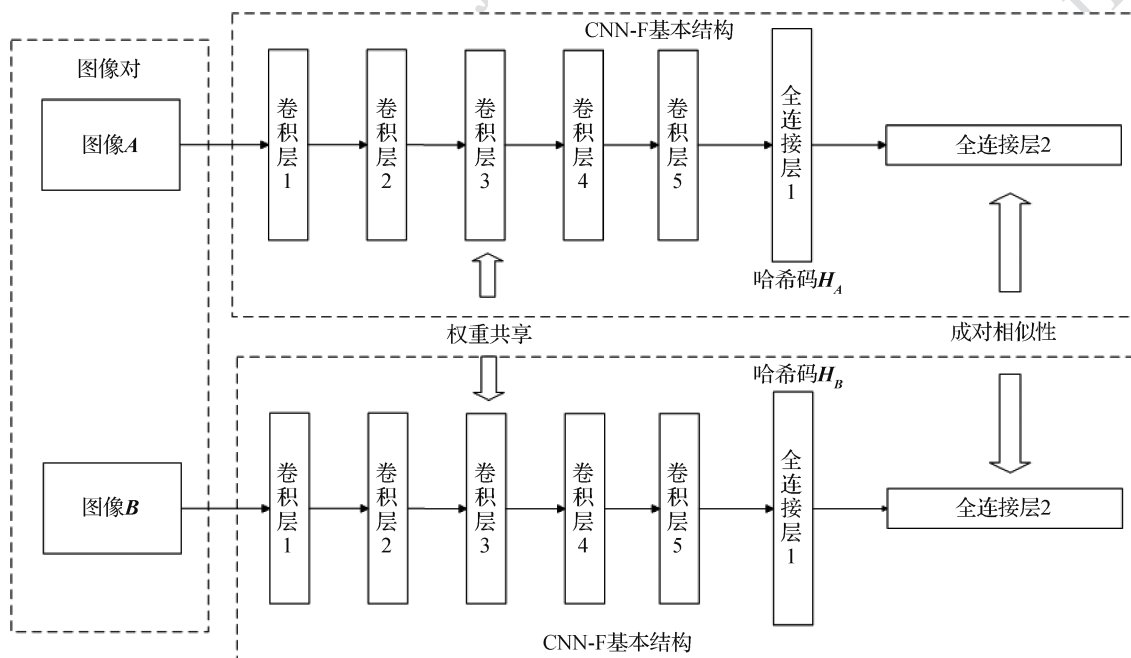


图 6 DPSH 算法结构(Li 等,2016)

Fig. 6 Architecture of DPSH algorithm(Li et al. , 2016)

DSDH(deep supervised discrete Hashing)算法(Li 等,2017)将学习到的二进制码用于图像分类,结构与 DPSH 类似,以 CNN-F 作为基本结构,并在同一框架下同时使用成对标签信息,不同的是 DSDH 生成的哈希码可用于分类。实验证明,该算法学习到的二进制编码既能保持图像之间的相似关系,又能和标签信息保持一致;此外,为了减少量化误差,算法保留了哈希编码的离散化,并使用离散循环坐标下降法优化哈希函数。

大多数深度哈希方法在生成哈希码时,首先学

习连续的图像表示,然后通过一个分离的二进制化步骤产生二进制哈希码,会导致潜在的检索质量损失。Cao 等人(2017b)提出 HashNet 框架,通过收敛的连续方法直接进行哈希码学习,从连续的相似数据中学习准确的二进制码。此外,为了保持图像间的相似性,HashNet 基于交叉熵损失函数设计了加权成对交叉熵损失函数。实验证明,HashNet 可以准确生成二进制哈希码,提升了检索性能。算法结构如图 7 所示。

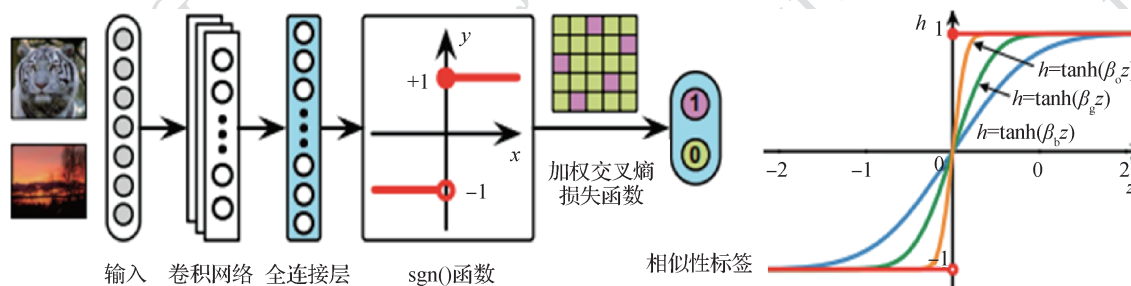


图 7 HashNet 算法结构(Cao 等,2017b)

Fig. 7 Architecture of HashNet algorithm(Cao et al. , 2017b)

深度学习哈希通过端对端的方式学习图像特征,并通过对特征进行哈希编码提升检索性能。针对图像数据相似性信息不足,Cao 等人(2018a)提出 HashGAN 架构,利用实际图像与生成模型合成的图像进行训练,学习出紧凑的二进制哈希码。主要思

想是以成对图像相似性信息为条件,利用 Wasserstein GAN(PC-WGAN)合成生成图增强训练数据的相似性,从真实图像和生成图像中学习紧凑的二进制哈希码,获得优越的检索性能,算法结构如图 8 所示。

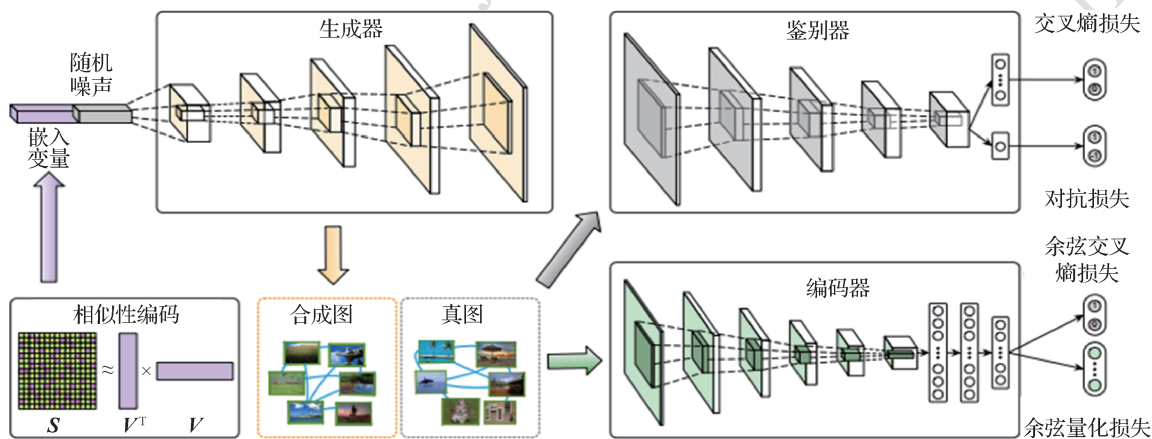


图8 HashGAN 算法结构(Cao 等,2018a)

Fig. 8 Architecture of HashGAN algorithm(Cao et al., 2018a)

除上述 4 种算法外,Zhu 等人(2016)提出 DHN (deep Hashing network),实现了同时控制量化误差和优化损失函数;Cao 等人(2018b)提出 DCH (deep cauchy Hashing)架构,通过联合优化柯西交叉熵损失和柯西量化损失生成紧凑且集中的二进制哈希码,实现了高效的汉明空间检索。由此可见,基于成对标签的深度哈希方法大部分通过交叉熵损失函数或对交叉熵损失函数进行改进保留图像的相似性。

3)设计新损失函数的深度哈希方法。除了使用平方损失函数和交叉熵损失函数,还有一些算法通过设计新的损失函数获得了高质量的哈希码。Liu 等人(2016)在 DSH(deep supervised Hashing)框架中根据汉明距离设计了一种新的损失函数,对来自输入图像对的监督信息进行编码,对实值输出进行正则化,使连续的图像表示更逼近离散值($+1/-1$),最大化哈希码的可区分性,使得不相似的图像间哈希码差异更大。Cao 等人(2017a)首次将深度学习从图像监督信息和常规文本域提取的语义信息中量化出来,提出了 DVSQ (deep visual-semantic quantization)深度哈希框架,通过设计网络结构和新的自适应边距铰链损失函数,同时学习图像特征并提取文本域中的语义信息,最后将图像和文本域中提取的语义信息进行匹配。DVSQ 模型通过在混合

深度网络上优化自适应余量损失和视觉语义量化损失学习紧凑的二进制代码,能够更有效地进行图像检索。

2.3.2 基于三元组监督信息的深度哈希方法

基于三元组监督信息的深度哈希方法使用图像三元组对 CNN 进行训练,图像三元组由 3 幅图像构成,前两幅相似度较高,第 1 幅与第 3 幅相似度较低(Lai 等,2015),通过对 3 幅图像进行特征提取和哈希编码,使相似的图像具有相似的哈希码。由于基于三元组监督信息的深度哈希方法需要进行大量的图像预处理工作,所以研究成果较少。本文按使用损失函数的不同,将其分为两类:1)使用三元组损失函数(Norouzi 等,2012)的深度哈希方法,三元组损失函数实际上是由铰链损失函数改进而来;2)使用三元组交叉熵损失函数的深度哈希方法。

1)使用三元组损失函数的深度哈希方法。Lai 等人(2015)对 CNNH 网络进行改进,提出了 NINH (network in network Hashing),将三元组图像作为网络输入,通过两个子网络同时训练,通过三元组损失函数使三元组中相似的图像具有相似的哈希码,不相似的图像具有差异较大的哈希码,实现特征提取和哈希编码的同时优化。算法结构如图 9 所示。

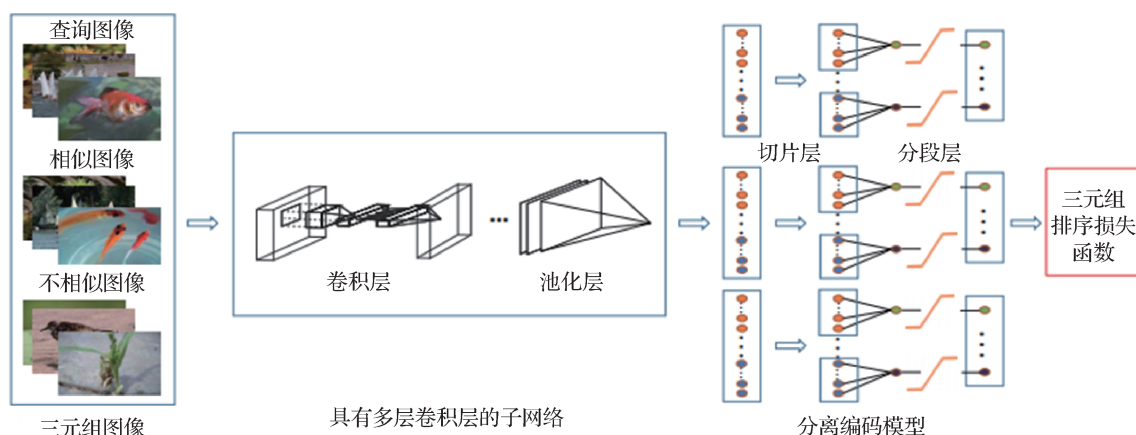


图9 NINH 算法结构(Lai 等,2015)

Fig.9 Architecture of NINH algorithm(Lai et al. , 2015)

Zhang 等人(2015)提出的 DRSCH(deep regularized similarity comparison Hashing)算法将训练图像数据按三元组样本分组,训练模型时可以最大化不相似对之间的汉明距离,在测试阶段引入正则化项确保相似图像具有相似的二进制代码,保证图像对与三元组监督信息之间的确定关系,用加权汉明距离代替标准汉明距离,通过截断无意义的位控制代码长度,生成紧凑的哈希码,并通过去除拉普拉斯正则化项得到 DRSCH 框架的一个简化变体,即 DSCH(deep similarity comparison Hashing)。

为了实现更有效的三元组训练,Liu 等人(2018)设计了 DTQ(deep triplet quantization)框架,

提出一种新的三元组选择方法 group hard,在每个图像组中随机选择三元组。DTQ 与 DQN 和 DVSQ 相似,也是对深度学习量化进行研究。DQN 通过成对的余弦损失和乘积量化损失学习图像表示,生成紧凑的二进制代码。DVSQ 设计了一种探索类标签,逐点自适应边距铰链损失函数和视觉语义量化损失。DTQ 为了生成紧凑的二进制代码,在三重态训练期间进一步应用具有弱正交性的三重态量化,量化损失减少了哈希码冗余,并通过反向传播增强了三元组中相似图像对的哈希码的相似性,增大了不相似图像对的哈希码的差异。DTQ 算法结构如图 10 所示。

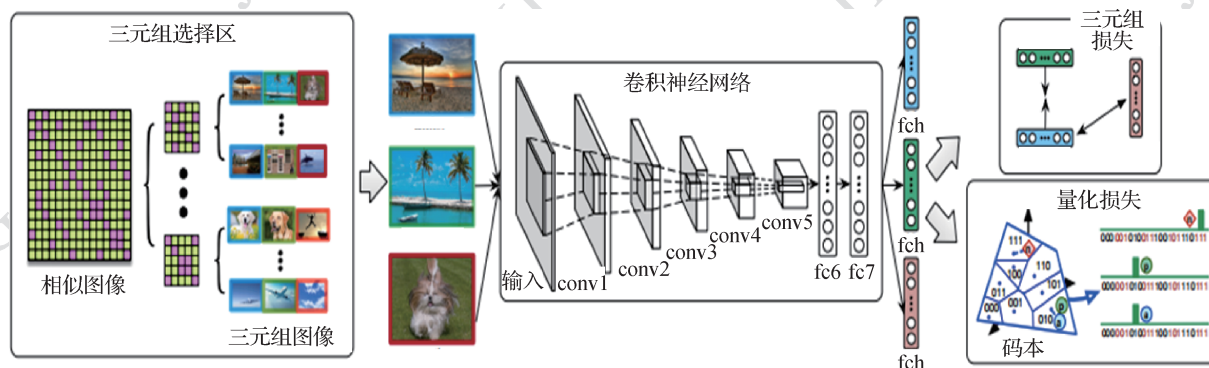


图10 DTQ 算法结构(Liu 等,2018)

Fig.10 Architecture of DTQ algorithm(Liu et al. , 2018)

2)使用三元组交叉熵损失函数的深度哈希方法。Wang 等人(2016)以 DPSH 网络架构为基础,提出了 DTSH(deep supervised Hashing with triplet)算法。与 DPSH 不同,DTSH 以三元组标签作为监督信息同时进行特征和哈希码学习,并基于 DPSH 使

用的成对交叉熵损失函数,提出了基于三元组的交叉熵损失函数。与 DPSH 相似,DTSH 也通过正则化项优化哈希码。

2.3.3 监督深度哈希的图像检索方法比较

表 1 展示了监督深度哈希的各种图像检索方法

表1 监督深度哈希的图像检索方法比较

Table 1 Comparison for methods of image retrieval based on supervised deep-Hashing

分类依据	损失函数	算法	创新点	优点	缺点
基于成对标签	平方损失函数	CNNH (Xia 等, 2014)	两阶段法学习哈希码; 第1阶段学习近似哈希码; 第2阶段输出连续哈希值	将CNN与哈希结合, 相较于传统哈希方法提高了检索精度	不是端到端的方法, 不能完全发挥CNN的学习能力
		DQN (Cao 等, 2016)	控制量化误差的同时优化了语义相似度对上的成对余弦损失和紧凑哈希码上的乘积量化损失	生成紧凑集中的高质量哈希码	采用平方余弦损失保持图像之间的相似性, 无法对语义空间中的真实标签分布建模
		DMDH (Chen 等, 2018)	提出差异最小化模型, 将原始二元优化问题转化为哈希函数的可微优化问题	改善了哈希码学习中的离散优化问题	-
	交叉熵损失函数	DPSH (Li 等, 2016)	通过设计损失函数并改进LFH的目标函数, 基于标签对同时学习图像表示和哈希码	减小了量化损失, 提高了检索精度, 是第1种同时学习图像特征和哈希码的模型	没有充分利用标签信息, 只针对相似性设计损失函数
		DHN (Zhu 等, 2016)	提出成对量化损失, 利用成对量化损失优化成对交叉熵损失	生成紧凑集中的哈希码	无法精确定义二进制化前后的量化损失
		DSDH (Li 等, 2017)	认为哈希码也可用于分类, 将最后一层网络输出直接限制为二值编码以保留哈希码的离散性; 交替最小化方法来进行优化损失函数	生成的哈希码保持了图像间的相似性关系且与标签信息保持一致, 同时降低了量化损失	-
		HashNet (Cao 等, 2017b)	通过设计一种新的加权成对交叉熵损失保留相似性; 不同于之前的哈希方法通过将连续码量化生成二进制码	HashNet可最小化二进制化前后的代码量化误差来精确学习二进制代码	-
		HashGAN (Cao 等, 2018a)	针对现有图像数据的相似性信息不足问题将哈希与GAN网络结合	用GAN网络生成图增强训练数据的相似性, 获得了优越的检索性能	-
		DCH (Cao 等, 2018b)	针对错误指定损失函数问题联合优化损失函数和量化损失函数	生成紧凑集中的哈希码, 实现了高效的汉明空间检索	-
		DSH (Liu 等, 2016)	对网络输出的哈希码进行正则化约束, 使之更接近二值编码(+1或-1)	提高网络收敛速度, 保证了训练过程的稳定性	没有充分利用标签信息, 仅针对相似性设计损失函数
		DVSQ (Cao 等, 2017a)	提出新的损失函数“自适应边距铰链损失函数”保留相似性和“视觉语义量化损失”学习紧凑的哈希码	加快了收敛速度并提高了搜索准确性, 同时减少编码冗余	-
基于三元组标签	三元组铰链损失函数	NINH(DNNH) (Lai 等, 2015)	通过改进CNNH的网络结构, 使算法框架能对3幅图像构成的三元组进行端到端训练	在一个阶段中学习图像表示和其对应的哈希码, 减少了哈希编码信息冗余	事先进行大量图像预处理工作
		DSCH/DRSCH (Zhang 等, 2015)	使用加权汉明距离代替标准汉明距离	生成了具有权重矩阵的哈希码, 对图像间的相似性信息进行了充分利用	事先进行大量图像预处理工作
	三元组交叉熵损失函数	DTQ (Liu 等, 2018)	提出一种新颖的三元组挖掘策略group hard; 提出了一种新的具有弱正交性约束的三重态量化损失, 以减少编码冗余	加快了收敛速度并提高了搜索准确性	事先进行大量图像预处理工作
		DTSH (Wang 等, 2016)	在DPSH的基础上, 以三元组标签作为监督信息同时学习图像表示和哈希; 提出基于三元组的成对交叉熵损失; 引入正则化项减小量化损失	一定程度上提高了检索精度	事先进行大量图像预处理工作

使用的损失函数、创新点和优缺点。由于基于三元组的深度哈希方法需要事先进行大量的图像预处理工作,导致研究较少。而基于成对监督信息的深度哈希方法则为基于三元组的深度哈希方法的改进策略奠定了基础,如 DNNH(deep neural networks Hashing)是由 CNNH 网络结构改进而来,DHN 和 HashNet 缘于 DNNH,DTSH 算法则以 DPSH 的算法结构为基础。由表 1 可以看出,监督深度哈希算法主要针对算法结构、哈希函数、损失函数和控制量化误差等问题进行改进。

3 算法对比及实验结果分析

3.1 实验数据集

使用 CIFAR-10(Krizhevsky,2009)和 NUS-WIDE(Chua 等,2009)数据集对图像检索进行实验分析。CIFAR-10 数据集包括 10 类,每类有 32×32 像素的 6 000 幅图像,如图 11 所示。NUS-WIDE 数

据集是公共多标签图像数据集,共有 269 648 幅彩色图像,5 018 个标签,数据集中每幅图像都带有 1 个或多个类标签,如图 12 所示。



图 11 CIFAR-10 数据集例图

Fig. 11 Example images of CIFAR-10 dataset



图 12 NUS-WIDE 数据集例图

Fig. 12 Example images of NUS-WIDE dataset

3.2 评价指标

使用平均精度(mean average precision,mAP)评估各种算法在 CIFAR-10 和 NUS-WIDE 数据集上的性能表现。通常情况下,使用 AP(average precision)对图像检索算法的性能进行衡量(Zheng 等,2018),AP 相当于精准率—召回率曲线线下的面积,较大的 AP 表示算法具有更好的检索性能。由于图像检索数据集通常具有多个查询图像,平均精度 mAP 就是对多个查询图像各自的 AP 值进行平均。

3.3 算法性能比较与分析

附表 1 和附表 2 列出了不同哈希码长度下各算法在 CIFAR-10 和 NUS-WIDE 数据库上的 mAP 值。

为了更直观地体现各种算法在不同数据库的检索性能,根据附表1和附表2的统计结果绘制了图13—

图 15。

图 13 比较了无监督哈希方法的检索性能,前 3 个和后 3 个方法分别为传统和深度无监督哈希方法。可以看出,深度无监督哈希方法的检索性能优于传统无监督哈希方法。DeepBit 算法首次用 CNN 模型提取图像本身的特征,在没有标签信息的情况下,CNN 模型提取特征能力与传统 ITQ 方法接近。同样基于深度模型 VGG-16 的无监督深度哈希算法 SADH 则通过 ADMM 算法对生成的二进制哈希码进行优化,检索性能比 DeepBit 提高了 1.4 倍。PGH 算法将无监督哈希算法与 GAN 结合,将生成的哈希码作为语义条件作为 GAN 的输入,得到了无监督深度哈希算法的最好性能。图 14 展示了各种监督哈希算法在哈希码长为 32 bits 时的检索性能,其中,SDH、KSH、

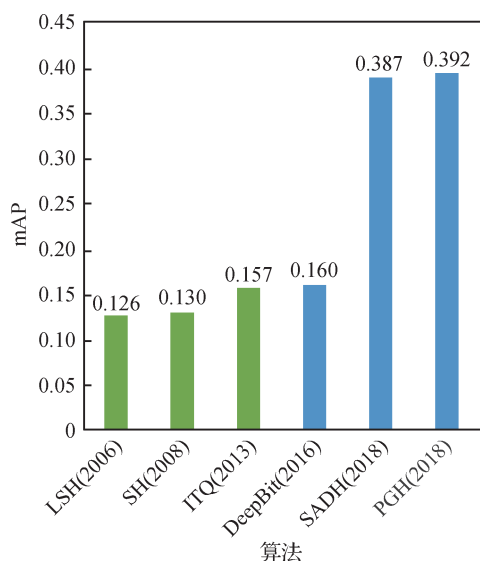


图13 哈希码长为16 bits时,各种无监督哈希算法在CIFAR-10数据集上的性能比较

Fig. 13 Comparison of performance for various unsupervised Hashing algorithms on CIFAR-10 dataset when Hashing code length is 16 bits

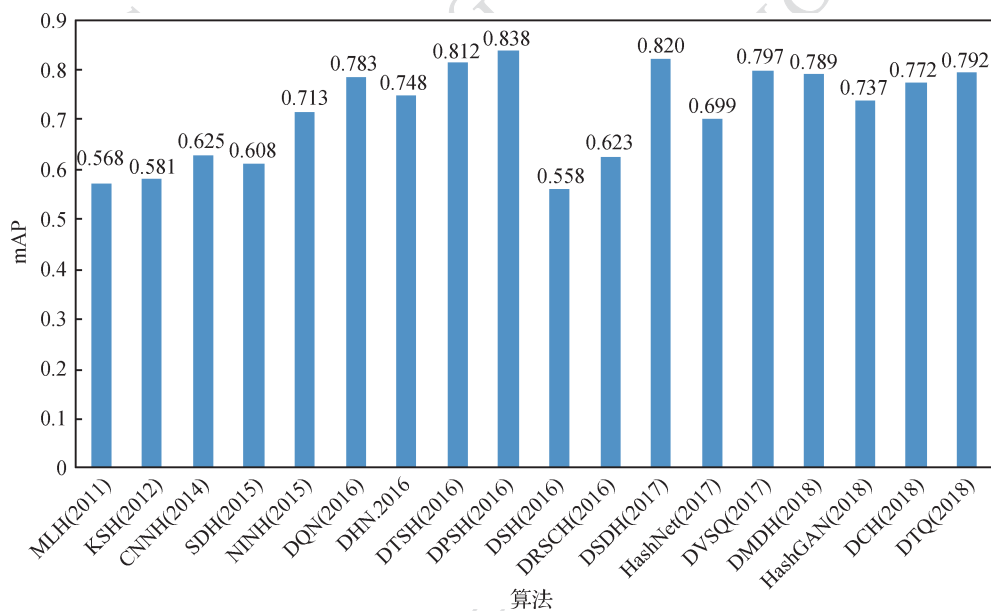


图14 哈希码长为32 bits时,各种监督哈希算法在NUS-WIDE数据集上的性能比较

Fig. 14 Comparison of performance for various supervised Hashing algorithms on NUS-WIDE dataset when Hashing code length is 32 bits

图15是各种从传统到深度哈希算法在哈希码长为32 bits时在CIFAR-10数据集上的检索性能比较,其中,LSH、SH、ITQ、DeepBit、SADH、PGH是传统无监督和深度无监督哈希方法,其余为传统监督和深度监督哈希方法。可以看出,深度监督哈希方法的检索性能显然高于深度无监督哈希方法。在2014—2017年,深度监督哈希方法蓬勃发展,基于成对标签或三元组标签,通过设计网络框架,修改

MLH是传统监督哈希算法,其余为深度监督哈希方法。图中检索性能最好的DPSH算法和第2好的DSDH算法都是利用成对标签作为监督信息,第3好的DTSH算法利用三元组标签作为监督信息,其网络框架和损失函数都是由DPSH算法改进而来,且都引入正则化项对哈希码进行优化。由于DSDH算法从网络中得到的离散哈希码不仅保留了图间的相似性,且与标签信息保持一致,因此取得了显著的检索效果。图中DVSQ、DTQ和DQN算法通过设计不同的量化损失函数对深度学习量化进行研究,DTQ算法是基于三元组监督信息的深度哈希方法,DQN和DVSQ算法则使用成对标签作为监督信息。DQN算法通过成对的余弦损失和乘积量化损失学习图像表示,DVSQ设计了一种新的探索类标签逐点自适应边距较链损失函数以及视觉语义量化损失,DTQ应用具有弱正交性的三重态量化生成了高质量的哈希码,使相似图像的哈希码更相似,不相似图像的哈希码差异更大。

损失函数,提出了一系列深度监督哈希方法,取得了不错的检索性能,但也发现了新的问题,深度哈希算法不错的检索性能是在高度依赖有标签数据的基础上取得的,对于深度神经网络,需要大量的标记数据确保网络提取特征的良好结果,从而获得更高的检索效率。然而在现实世界中,随着网络和移动设备的不断发展,图像数据量每天都在爆炸式增长,如果对这些图像数据一一进行标注,

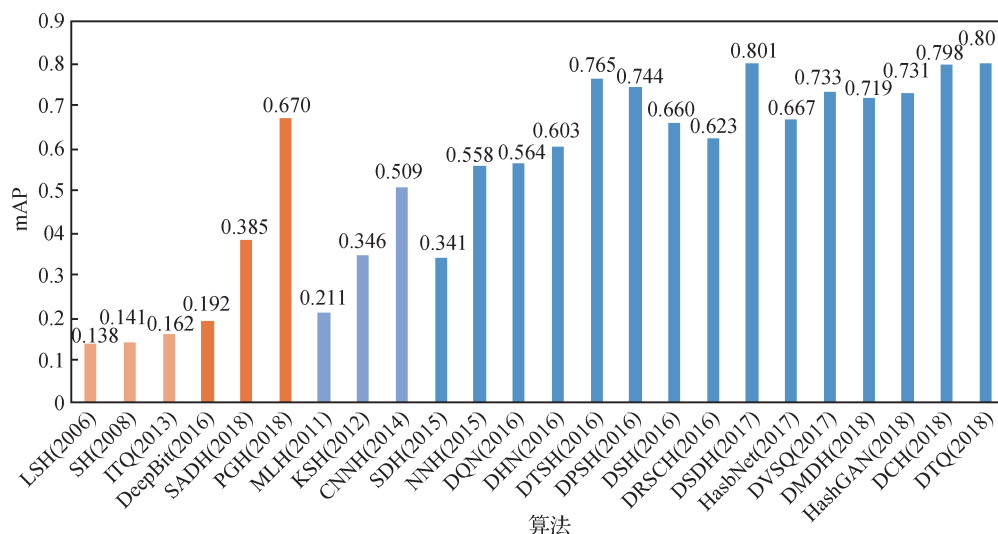


图 15 哈希码长为 32 bits 时,各哈希算法在 CIFAR-10 数据集上的性能比较

Fig. 15 Comparison of performance for various Hashing algorithms on CIFAR-10 dataset when Hashing code length is 32 bits

需要耗费高昂的人工和时间成本。因此不需要图像标注的无监督哈希方法重新受到关注,主要研究方向是将无监督哈希方法与深度学习中的 CNN 模型或 GAN 结合,以期在实际应用中取得更高的检索性能。

3.4 深度哈希算法在特色数据库上的尝试

CIIP-CSID (crime scene investigation image database) (2018) 和 CIIP-TPID (tread pattern image dataset) (2018) 是西安邮电大学图像与信息处理研究所

(Center for Image and Information Processing, CIIP) 自建的两个特色数据库,分别是现勘图像数据库 CIIP-CSID 和刑侦轮胎花纹图像数据库 (CIIP-TPID)。CIIP-CSID 数据库包含生物物证、血迹、车辆、指纹、鞋印、作案工具、轮胎压痕等 46 类案例,共 11 456 幅图像(刘颖等,2018),例图如图 16 所示。CIIP-TPID 数据库是在不同时间、环境、角度、亮度、距离等条件下采集的轮胎花纹图像,每类 50 幅,共 5 000 幅(刘颖等,2019),例图如图 17 所示。



图 16 CIIP-CSID 数据库例图

Fig. 16 Example images for CIIP-CSID database

((a) license plate; (b) palmprint; (c) shoeprint; (d) tattoo; (e) skin; (f) tire; (g) crime tool; (h) hair)

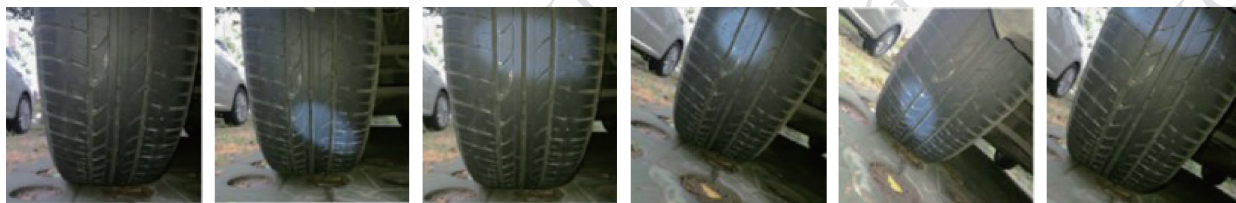


图 17 CIIP-TPID 数据库例图

Fig. 17 Example images for CIIP-TPID database

为验证深度哈希算法在特色数据库上的检索性能,选取在 CIFAR-10 和 NUS-WIDE 数据集上表现

较好的 DPSH 算法在 CIIP-CSID 和 CIIP-TPID 数据库上进行实验,结果如表 2 所示。可以看出,DPSH

算法在 CIIP-CSID 数据库中的 mAP 值介于在 CIFAR-10 和 NUS-WIDE 两个数据集取值之间,检索性能较好,但在 CIIP-TPID 数据库上的 mAP 值较低,主要原因是 CIIP-TPID 数据库与其他数据库数据内容明显不同,该数据库中的轮胎花纹图像不同类之间的视觉差异较小,导致训练中提取的分类信息较少,最终造成 DPSH 算法在 CIIP-TPID 数据库上的检索效果较差,对此可以从两个方面进行改进:1)改进 DPSH 算法的算法结构,更好地学习图像表示,提取到更多的特征;2)扩大数据库的数据量,增加更多的训练样本,使得算法能提取更多的特征进行检索。

表2 DPSH 算法在不同数据库上的 mAP 值比较
Table 2 Comparison of mAP for DPSH algorithm on different databases

数据库	哈希码长度/bits			
	12	24	32	48
CIIP-CSID	0.753	0.771	0.778	0.780
CIIP-TPID	0.200	0.192	0.223	0.173
CIFAR-10	0.713	0.727	0.744	0.757
NUS-WIDE	0.794	0.822	0.838	0.851

实验表明,深度哈希算法针对特色数据库检索时,应根据特色数据库的特点进行适当调整。鉴于 DPSH 算法在 CIIP-CSID 数据库上的优异表现,相信深度哈希算法未来将在更多的特色数据库中发挥作用。

4 结 语

4.1 未来研究方向

深度哈希作为前沿研究方向,尽管在目标函数、网络结构以及模型等方面进行了改进(孙瑶,2018),但在以下方面依然亟待创新:

1)无监督深度哈希方法的性能提升。虽然通过将无监督哈希方法与 CNN 模型或 GAN 结合取得了较传统方法更好的检索效果,但相较于监督哈希方法,无监督哈希缺乏标签信息,缺少了一部分特征来源。因此在无监督哈希方法中如何利用有限的特征获取更高的检索性能是一个值得研究的问题。由于多模态融合(Atrey 等,2010)方法可以利用多模

态之间的互补性剔除模态间的冗余性,从而学习更精确的特征表示。因此在改进无监督深度哈希的图像检索方法时,可以考虑应用多模态融合特征进行特征提取,从而提高检索精度。

2)基于图网络的半监督深度哈希方法。现有的数据都可分为两类,即欧几里德结构数据(Euclidean structure data)和非欧几里德结构数据(non-Euclidean structure data)(Bronstein 等,2017)。在图像检索领域,欧几里德结构数据指图像像素点可以排列成整齐矩阵的数据。鉴于 CNN 离散卷积的特点,可以有效提取这类数据的特征,但无法处理非欧几里德结构数据。针对非欧几里德结构数据中的图数据,Kipf 和 Welling(2017)提出了 GCN(graph convolutional network),用于基于图数据进行半监督图像分类,取得了较好的分类结果。Zhou 等人(2018)将 GCN 与哈希结合,提出了 GCNH(graph convolutional network Hashing)模型进行半监督图像检索,相较于传统的哈希方法,在有限的标签数据下,GCNH 的检索性能得到了很大提升,为半监督深度哈希方法提供了新思路。

3)跨模态检索。本文提到的算法都属于以图搜图的哈希方法。在实际生活中,以文本搜图的应用十分广泛。跨模态检索(cross-modal retrieval)方法(Wang 等,2017)是通过利用某一种模态样本搜索近似语义的其他模态样本。以文本搜图属于跨模态检索的应用之一。跨模态哈希(cross-modal Hashing,CMH)由于存储成本低和查询速度快,已广泛用于多媒体检索中的相似性搜索,但几乎所有的 CMH 算法都是基于手工提取的特征进行查询。Jiang 和 Li(2017)将跨模态哈希与深度学习集成到同一框架中,提出了一种新的跨模态哈希算法 DCMH(deep cross-modal Hashing)。Deng 等人(2018)基于哈希的图像检索技术,设计了 TDH(triplet based deep Hashing)网络,基于三元组标签进行跨模态检索,同时设计损失函数优化哈希码,最终取得了比 CMH 算法更好的检索效果。期待跨模态检索与深度哈希的结合能带来更高效的检索性能。

4)生物识别技术与深度哈希的结合。生物识别在实际生活中应用广泛,如支付宝的刷脸支付、智能手机的指纹/人脸解锁等用到了属于生物识别技术的眼纹识别、指纹识别、虹膜识别以及人脸识别

等。人们为了追求方便快捷的生活方式,极易接受采用了生物识别的刷脸支付、人脸解锁等新技术,导致人脸和指纹等数据量急剧增长。而深度哈希技术十分适用于处理海量数据,因此在未来可以考虑将生物识别技术与深度哈希结合,获取更好的检索效果。

4.2 总结

本文介绍了具有代表性的基于深度哈希的图像检索方法,对检索性能进行了比较和分析。相较于传统的基于哈希的图像检索方法,基于深度哈希的图像检索方法的检索精度明显更高。针对是否使用标签信息,将深度哈希方法分为无监督、半监督和监督深度哈希分别介绍。无监督深度哈希方法不使用标签信息,不需要对数据集中的图像进行标注,更适合实际应用。将无监督哈希方法与 CNN 模型或 GAN 网络结合是无监督哈希领域的主要研究方向。半监督深度哈希方法的研究成果较少,没有过多介绍。深度监督哈希方法按使用的监督信息形式的不同,将基于深度哈希的图像检索方法分为基于成对监督信息和基于三元组监督信息的图像检索方法,并进一步根据各算法损失函数的不同分别对各类中的代表性算法进行介绍。鉴于深度哈希在以图搜图架构中的良好表现,可以期待深度哈希技术未来在跨模态检索、生物识别及特色数据库检索中发挥更大的作用。

参考文献 (References)

- Andoni A and Indyk P. 2006. Near-optimal Hashing algorithms for approximate nearest neighbor in high dimensions//Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science. Berkeley, CA, USA; IEEE: 459-468 [DOI: 10.1109/FOCS.2006.49]
- Atrey P K, Hossain M A, El Saddik A and Kankanhalli M S. 2010. Multimodal fusion for multimedia analysis: a survey. *Multimedia Systems*, 16(6): 345-379 [DOI: 10.1007/s00530-010-0182-0]
- Bronstein M M, Bruna J, LeCun Y, Szlam A and Vandergheynst P. 2017. Geometric deep learning: going beyond Euclidean data. *IEEE Signal Processing Magazine*, 34(4): 18-42 [DOI: 10.1109/MSP.2017.2693418]
- Burges C J. 1998. A tutorial on support vector machine for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2): 121-167 [DOI: 10.1023/a:1009715923555]
- Cao Y, Liu B, Long M S and Wang J M. 2018a. HashGAN: deep learning to Hash with pair conditional wasserstein GAN//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT; IEEE: 1287-1296 [DOI: 10.1109/CVPR.2018.00140]
- Cao Y, Long M S, Liu B and Wang J M. 2018b. Deep cauchy Hashing for Hamming space retrieval//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT; IEEE: 1229-1237 [DOI: 10.1109/CVPR.2018.00134]
- Cao Y, Long M S, Wang J M and Liu S C. 2017a. Deep visual-semantic quantization for efficient image retrieval//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA; IEEE: 1328-1337 [DOI: 10.1109/CVPR.2017.104]
- Cao Z J, Long M S, Wang J M and Yu P S. 2017b. HashNet: deep learning to Hash by continuation//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy; IEEE: 5608-5617 [DOI: 10.1109/ICCV.2017.598]
- Cao Y, Long M S, Wang J M, Zhu H and Wen Q F. 2016. Deep quantization network for efficient image retrieval// Proceedings of the 13th AAAI Conference on Artificial Intelligence. Phoenix, Menlo Park, CA; AAAI: 3457-3463 [DOI: 10.1109/cvpr.2017.104]
- Chatfield K, Simonyan K, Vedaldi A and Zisserman A. 2014. Return of the devil in the details: delving deep into convolutional nets//Proceedings of 2014 British Machine Vision Conference. Nottingham, UK; BMVA Press: 1-11 [DOI: 10.5244/C.28.6]
- Chen Z X, Yuan X, Lu J W, Tian Q and Zhou J. 2018. Deep Hashing via discrepancy minimization//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT; IEEE: 6838-6847 [DOI: 10.1109/CVPR.2018.00715]
- Chua T S, Tang J H, Hong R C, Li H J, Luo Z P and Zheng Y T. 2009. NUS-WIDE: a real-world web image database from national university of Singapore//Proceedings of the ACM International Conference on Image and Video Retrieval. Santorini, Fira, Greece; ACM: 48-49 [DOI: 10.1145/1646396.1646452]
- CIIP-CSID. 2018. CIIP-CSID[DB/OL]. [2019-08-27]. <http://www.xuptciip.com.cn/show.html?database-xksjk04> (西安邮电大学图像处理团队现勘图像数据库. 2018. 西安邮电大学图像处理团队现勘图像数据库[DB/OL]. [2019-08-27]. <http://www.xuptciip.com.cn/show.html?database-xksjk04>)
- CIIP-TPID. 2018. CIIP-TPID[DB/OL]. [2019-08-27]. <http://www.xuptciip.com.cn/show.html?database-xksjk03> (西安邮电大学图像处理团队轮胎表面花纹图像数据库. 2018. 西安邮电大学图像处理团队轮胎表面花纹图像数据库[DB/OL]. [2019-08-27]. <http://www.xuptciip.com.cn/show.html?database-xksjk03>)
- Dalal N and Triggs B. 2005. Histograms of oriented gradients for human detection//Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, CA, USA; IEEE: 886-893 [DOI: 10.1109/CVPR.2005.177]
- Deng C, Chen Z J, Liu X L, Gao X B and Tao D C. 2018. Triplet-based deep Hashing network for cross-modal retrieval. *IEEE Trans-*

- actions on Image Processing, 27(8): 3893-3903 [DOI: 10.1109/tip.2018.2821921]
- Dizaji K G, Zheng F, Nourabadi S N, Yang Y H, Deng C and Huang H. 2018. Unsupervised deep generative adversarial Hashing network//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE: 3664-3673 [DOI: 10.1109/CVPR.2018.00386]
- Friedman J H. 1997. On bias, variance, 0/1—loss, and the curse-of-dimensionality. Data Mining and Knowledge Discovery, 1(1): 55-77 [DOI: 10.1023/a:1009778005914]
- Gong Y C, Lazebnik S, Gordo A and Perronnin F. 2013. Iterative quantization: a procrustean approach to learning binary codes for large-scale image retrieval. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(12): 2916-2929 [DOI: 10.1109/TPAMI.2012.193]
- Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A and Bengio Y. 2014. Generative adversarial nets//Proceedings of the 27th International Conference on Neural Information Processing Systems. Cambridge, MA: MIT: 2672-2680
- Gray R M and Neuhoff D L. 1998. Quantization. IEEE Transactions on Information Theory, 44(6): 2325-2383 [DOI: 10.1109/18.720541]
- Hadsell R, Chopra S and Lecun Y. 2006. Dimensionality reduction by learning an invariant mapping//Proceedings of 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York, USA: IEEE: 1735-1742 [DOI: 10.1109/CVPR.2006.100]
- He K M, Wen F and Sun J. 2013. K-means Hashing: an affinity-preserving quantization method for learning binary compact codes//Proceedings of 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland: IEEE: 2938-2945 [DOI: 10.1109/CVPR.2013.378]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Hinton G E and Salakhutdinov R R. 2006. Reducing the dimensionality of data with neural networks. Science, 313(5786): 504-507 [DOI: 10.1126/science.1127647]
- Jégou H, Douze M and Schmid C. 2011. Product quantization for nearest neighbor search. IEEE Transactions on Pattern Analysis and Machine Intelligence, 33(1): 117-128 [DOI: 10.1109/TPAMI.2010.57]
- Jiang Q Y and Li W J. 2017. Deep cross-modal Hashing//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI: IEEE: 3270-3278 [DOI: 10.1109/CVPR.2017.348]
- Kim S and Choi S. 2011. Semi-supervised discriminant Hashing//Proceedings of the 11th IEEE International Conference on Data Mining. Vancouver: IEEE: 1122-1127 [DOI: 10.1109/ICDM.2011.128]
- Kipf T and Welling M. 2017. Semi-supervised classification with graph convolutional networks//Proceedings of the 5th International Conference on Learning Representations, Toulon, France: [s. n.]: 1-14
- Kong W H and Li W J. 2012. Isotropic Hashing//Proceedings of the 25th International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc.: 1646-1654
- Krizhevsky A. 2009. Learning Multiple Layers of Features from Tiny Images. Technical Report TR-2009. University of Toronto
- Lai H J, Pan Y, Liu Y and Yan S C. 2015. Simultaneous feature learning and Hash coding with deep neural networks//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA: IEEE: 3270-3278 [DOI: 10.1109/CVPR.2015.7298947]
- Li Q, Sun Z N, He R and Tan T N. 2017. Deep supervised discrete Hashing//Proceedings of the 22nd Advances in Neural Information Processing Systems. Cambridge, MA: NIPS: 2482-2491
- Li W J and Zhou Z H. 2015. Learning to Hash for big data: current status and future trends. Chinese Science Bulletin, 60(5/6): 485-490 (李武军, 周志华. 2015. 大数据哈希学习: 现状与趋势. 科学通报, 60(5/6): 485-490) [DOI: 10.1360/N972014-00841]
- Li W J, Wang S and Kang W C. 2016. Feature learning based deep supervised Hashing with pairwise labels//Proceedings of the 25th International Joint Conference on Artificial Intelligence. Menlo Park, CA: AAAI: 1711-1717
- Lin K, Lu J W, Chen C S and Zhou J. 2016. Learning compact binary descriptors with unsupervised deep neural networks//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE: 1183-1192 [DOI: 10.1109/CVPR.2016.133]
- Lin K, Yang H F, Hsiao J H and Chen C S. 2015. Deep learning of binary Hash codes for fast image retrieval//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Boston, MA: IEEE: 27-35 [DOI: 10.1109/CVPRW.2015.7301269]
- Liu B, Cao Y, Long M S, Wang J M and Wang J D. 2018. Deep triplet quantization//Proceedings of the 26th ACM International Conference on Multimedia. New York: ACM: 755-763
- Liu H M, Wang R P, Shan S G and Chen X L. 2016. Deep supervised Hashing for fast image retrieval//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 2064-2072 [DOI: 10.1109/CVPR.2016.227]
- Liu L, Yu M Y and Shao L. 2017. Latent structure preserving Hashing. International Journal of Computer Vision, 122(3): 439-457 [DOI: 10.1007/s11263-016-0931-4]
- Liu W, Wang J, Ji R R, Jiang Y G and Chang S F. 2012. Supervised Hashing with kernels//Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, Rhode Island, USA: IEEE: 2074-2081 [DOI: 10.1109/CVPR.2012.6247912]

- Liu Y, Hu D and Fan J L. 2018. A survey of crime scene investigation image retrieval. *Acta Electronica Sinica*, 46(3): 761-768 (刘颖, 胡丹, 范九伦. 2018. 现勘图像检索综述. *电子学报*, 46(3): 761-768) [DOI: 10.3969/j.issn.0372-2112.2018.03.035]
- Liu Y, Zhang S and Fan J L. 2019. Tread pattern image classification with feature fusion based on transfer learning. *Computer Engineering and Design*, 40(5): 1401-1406 (刘颖, 张帅, 范九伦. 2019. 基于迁移学习及特征融合的轮胎花纹图像分类. *计算机工程与设计*, 40(5): 1401-1406) [DOI: 10.16208/j.issn1000-7024.2019.05.037]
- Lowe D G. 1999. Object recognition from local scale-invariant features// *Proceedings of the 7th IEEE International Conference on Computer Vision*. Kerkyra, Greece: ICCV; 1150-1151 [DOI: 10.1109/ICCV.1999.790410]
- Ma Y Q, He Y, Ding F, Hu S, Li J and Liu X L. 2018. Progressive generative Hashing for image retrieval//*Proceedings of the 27th International Joint Conference on Artificial Intelligence*. San Mateo, CA: IJCAI; 871-877 [DOI: 10.24963/ijcai.2018/121]
- Meyer-Bäse A, Jancke K, Wismüller A and Georgiopoulos M. 2004. Fast k-dimensional tree-structured vector quantization encoding method for image compression. *Optical Engineering*, 43(5): 1012-1013 [DOI: 10.1117/1.1683885]
- Norouzi M and Fleet D J. 2011. Minimal loss Hashing for compact binary codes//*Proceedings of the 28th International Conference on Machine Learning*. Bellevue, Washington, USA: ICML; 353-360
- Norouzi M, Fleet D J and Salakhutdinov R. 2012. Hamming distance metric learning//*Proceedings of the 25th International Conference on Neural Information Processing Systems*. Cambridge, MA: NIPS; 1061-1069
- Ojala T, Pietikäinen M and Mäenpää T. 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7): 971-987 [DOI: 10.1109/TPAMI.2002.1017623]
- Oliva A and Torralba A. 2001. Modeling the shape of the scene: a holistic representation of the spatial envelope. *International Journal of Computer Vision*, 42(3): 145-175 [DOI: 10.1023/a:1011139631724]
- Shen F M, Shen C H, Liu W and Shen H T. 2015. Supervised discrete Hashing//*Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston, MA: IEEE; 37-45 [DOI: 10.1109/CVPR.2015.7298598]
- Shen F M, Xu Y, Liu L, Yang Y, Huang Z and Shen H T. 2018. Un-supervised deep Hashing with similarity-adaptive and discrete optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(12): 3034-3044 [DOI: 10.1109/TPAMI.2018.2789887]
- Simonyan K and Zisserman. 2014. A very deep convolutional networks for large-scale image recognition//*Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition*. Piscataway, NJ, IEEE; 1-14
- Song J K, He T, Gao L L, Xu X, Hanjalic A and Shen T H. 2018. Binary generative adversarial networks for image retrieval//*Proceedings of the 32nd AAAI Conference on Artificial Intelligence*. Menlo Park, CA: AAAI; 394-401
- Sun Y. 2018. The research and development of deep learning Hashing. *Journal of Data Communication*, (2): 49-54 (孙瑶. 2018. 深度学习哈希研究与发展. *数据通信*, (2): 49-54)
- Szegedy C, Liu W, Jia Y Q, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V and Rabinovich A. 2015. Going deeper with convolutions//*Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston: IEEE; 1-9 [DOI: 10.1109/CVPR.2015.7298594]
- Wang B K, Yang Y, Xu X, Hanjalic A and Shen H T. 2017. Adversarial cross-modal retrieval//*Proceedings of the 25th ACM International Conference on Multimedia*. New York: ACM; 154-162 [DOI: 10.1145/3123266.3123326]
- Wang D X, Cui P, Ou M D and Zhu W W. 2015. Learning compact Hash codes for multimodal representations using orthogonal deep structure. *IEEE Transactions on Multimedia*, 17(9): 1404-1416 [DOI: 10.1109/tmm.2015.2455415]
- Wang G A, Hu Q H, Cheng J and Hou Z G. 2018. Semi-supervised generative adversarial Hashing for image retrieval//*Proceedings of the 15th European Conference on Computer Vision (ECCV)*. Munich: Springer; 469-485 [DOI: 10.1007/978-3-030-01267-0_29]
- Wang J, Kumar S and Chang S F. 2010. Semi-supervised Hashing for scalable image retrieval//*Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Francisco: IEEE; 3424-3431 [DOI: 10.1109/CVPR.2010.5539994]
- Wang X F, Shi Y and Kitani K M. 2016. Deep supervised Hashing with triplet labels//*Proceedings of the 13th Asian Conference on Computer Vision*. Taipei, China: Springer; 70-84 [DOI: 10.1007/978-3-319-54181-5_5]
- Wei E M and Ozdaglar A. 2012. Distributed alternating direction method of multipliers//*Proceedings of the 51st IEEE Conference on Decision and Control (CDC)*. Maui: IEEE; 5445-5450 [DOI: 10.1109/CDC.2012.6425904]
- Weiss Y, Torralba A and Fergus R. 2008. Spectral Hashing//*Proceedings of the 21st International Conference on Neural Information Processing Systems*. Vancouver, British Columbia, Canada: NIPS; 1753-1760
- Wu C X, Zhu J K, Cai D, Chen C and Bu J J. 2013. Semi-supervised nonlinear Hashing using bootstrap sequential projection learning. *IEEE Transactions on Knowledge and Data Engineering*, 25(6): 1380-1393 [DOI: 10.1109/tkde.2012.76]
- Xia R K, Pan Y, Lai H J, Liu C and Yan S C. 2014. Supervised Hashing for image retrieval via image representation learning//*Proceed-*

- ings of the 28th AAAI Conference on Artificial Intelligence. Québec, Canada; Menlo Park, CA; AAAI; 2156-2162
- Yang A B and Shou Z Y. 2018. A survey of sorting Hash algorithms in image retrieval. *Journal of Data Communication*, (6): 34-40 (杨安邦, 寿震宇. 2018. 图像检索中的排序哈希算法研究综述. *数据通信*, (6): 34-40)
- Yang M and Chen N. 2018. Cover song identification based on fusion of deep learning and manual design features. *Journal of East China University of Science and Technology (Natural Science Edition)*, 44(5): 752-759 (杨妹, 陈宁. 2018. 基于深度学习和手工设计特征融合的翻唱歌曲识别模型. *华东理工大学学报(自然科学版)*, 44(5): 752-759) [DOI: 10.14135/j.cnki.1006-3080.20170704003]
- Yang Y, Shen F M, Shen H T, Li H X and Li X L. 2015. Robust discrete spectral Hashing for large-scale image semantic indexing. *IEEE Transactions on Big Data*, 1(4): 162-171 [DOI: 10.1109/TBDATA.2016.2516024]
- Yu S X and Shi J. 2003. Multiclass spectral clustering//*Proceedings of the 9th IEEE International Conference on Computer Vision*. Nice, France; IEEE; #313
- Zhang J and Peng Y X. 2019. SSDH: semi-supervised deep Hashing for large scale image retrieval. *IEEE Transactions on Circuits and Systems for Video Technology*, 29(1): 212-225 [DOI: 10.1109/TCSVT.2017.2771332]
- Zhang P C, Zhang W, Li W J and Guo M Y. 2014. Supervised Hashing with latent factor models//*Proceedings of the 37th International ACM SIGIR Conference on Research and Development in Information Retrieval*. Gold Coast, Queensland, Australia; SIGIR: 173-182 [DOI: 10.1145/2600428.2609600]
- Zhang R M, Lin L, Zhang R, Zuo W M and Zhang L. 2015. Bit-scalable deep Hashing with regularized similarity learning for image retrieval and person re-identification. *IEEE Transactions on Image Processing*, 24(12): 4766-4779 [DOI: 10.1109/TIP.2015.2467315]
- Zheng L, Yang Y and Tian Q. 2018. SIFT meets CNN: a decade survey of instance retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(5): 1224-1244 [DOI: 10.1109/TPAMI.2017.2709749]
- Zhou X, Shen F M, Liu L, Liu W, Nie L Q, Yang Y and Shen H T. 2018. Graph convolutional network Hashing. *IEEE Transactions on Cybernetics*, 50(4): 1460-1472 [DOI: 10.1109/TCYB.2018.2883970]
- Zhu H, Long M S, Wang J M and Cao Y. 2016. Deep Hashing network for efficient similarity retrieval//*Proceedings of the 13th AAAI Conference on Artificial Intelligence*. Menlo Park, CA; AAAI; 2415-2421

作者简介



刘颖, 1972年生, 女, 教授, 硕士生导师, 主要研究方向为图像视频检索。

E-mail: liuying_ciiip@163.com

程美, 女, 硕士研究生, 主要研究方向为图像检索。

E-mail: 827127224@qq.com

王富平, 男, 讲师, 硕士生导师, 主要研究方向为图像特征提取与检索。E-mail: wfp1608@163.com

李大湘, 男, 副教授, 硕士生导师, 主要研究方向为图像检索与机器学习。E-mail: 35108809@qq.com

刘伟, 男, 副教授, 硕士生导师, 主要研究方向为图像检索及机器学习。E-mail: bme_liuwei@163.com

范九伦, 男, 教授, 博士生导师, 主要研究方向为模糊信息处理、图像处理、信息安全。E-mail: jiulunf@xupt.edu.cn

附表 (Appendix)

表 1 各种哈希算法在 CIFAR-10 数据集上的 mAP 值
Table 1 Values of mAP for various Hashing algorithms on CIFAR-10 dataset

类别	算法	实验结果参考文献	哈希码长度/bits					
			12	16	24	32	48	64
传统无监督 哈希方法	LSH	(Liu 等, 2016)	—	0.126	—	0.138	—	0.151
	SH		—	0.13	—	0.141	—	0.139
	ITQ		—	0.157	—	0.162	—	0.167
深度无监督 哈希方法	DeepBit	(Shen 等, 2018)	—	0.16	—	0.192	—	0.201
	SADH		—	0.387	—	0.385	—	0.377
	HashGAN	(Dizaji 等, 2018)	—	0.299	—	0.315	—	0.325
	PGH	(Ma 等, 2018)	—	0.392	—	0.67	—	0.68
	BGAN		—	0.403	—	0.408	—	0.516
传统监督 哈希方法	SDH	(Li 等, 2016)	0.285	—	0.329	0.341	0.356	—
	MLH	(Xia 等, 2014)	0.182	—	0.195	0.207	0.211	—
	CNNH		0.303	—	0.337	0.346	0.356	—
	KSH		0.303	—	0.337	0.346	0.356	—
深度监督 哈希方法	CNNH	(Xia 等, 2014)	0.439	—	0.511	0.509	0.522	—
	DQN	(Cao 等, 2016)	0.554	—	0.558	0.564	0.58	—
	DMDH	(Cheng 等, 2018)	—	0.704	—	0.719	0.732	0.737
	DPSH	(Li 等, 2016)	0.713	—	0.727	0.744	0.757	—
	DSDH	(Li 等, 2017)	0.74	—	0.786	0.801	0.82	—
	HashGAN	(Cao 和 Liu, 2018a)	—	0.668	—	0.731	0.735	0.749
	HashNet		—	0.643	—	0.667	0.675	0.687
	DHN		—	0.568	—	0.603	0.621	0.635
	DCH		—	0.79	—	0.798	0.807	0.794
	DSH		0.616	—	0.651	0.66	0.676	—
	DVSQ		—	0.727	0.73	0.733	—	—
	DNNH		0.552	—	0.566	0.558	0.581	—
	DRSCH		—	0.615	0.622	0.623	0.63	0.633
	DTQ		—	0.789	0.79	0.792	—	—
	DTSH		0.71	—	0.75	0.765	0.774	—

注:加粗字体为每列最优值,“—”表示没有实验数据。

表2 各种哈希算法在 NUS-WIDE 数据集上的 mAP 值
Table 2 Values of mAP for various Hashing algorithms on NUS-WIDE dataset

类别	算法	实验结果参考文献	哈希码长度/bits					
			12	16	24	32	48	64
传统无监督 哈希方法	LSH	(Liu 等,2016)	-	-	-	-	-	-
	SH		-	-	-	-	-	-
	ITQ		-	-	-	-	-	-
深度无监督 哈希方法	DeepBit	(Shen 等,2018)	-	-	-	-	-	-
	SADH		-	-	-	-	-	-
	HashGAN	(Dizaji 等,2018)	-	-	-	-	-	-
	PGH		-	-	-	-	-	-
	BGAN	(Ma 等,2018)	-	-	-	-	-	-
传统监督 哈希方法	SDH	(Li 等,2016)	0.568	-	0.6	0.608	0.637	-
	MLH	(Xia 等,2014)	0.5	-	0.514	0.52	0.522	-
	CNNH		0.556	-	0.572	0.581	0.588	-
	KSH		-	-	-	-	-	-
深度监督 哈希方法	CNNH	(Xia 等,2014)	0.611	-	0.618	0.625	0.608	-
	DQN	(Cao 等,2016)	0.768	-	0.776	0.783	0.792	-
	DMDH	(Chen 等,2018)	-	0.751	-	0.789	0.789	0.789
	DPSH	(Li 等,2016)	0.794	-	0.822	0.838	0.851	-
	DSDH	(Li 等,2017)	0.776	-	0.808	0.82	0.829	-
	HashGAN	(Cao 和 Liu,2018a)	-	0.715	-	0.737	0.744	0.748
	HashNet		-	0.662	-	0.699	0.711	0.716
	DHN		-	0.637	-	0.664	0.669	0.671
	DCH	(Cao 等,2018b)	-	0.74	-	0.772	0.769	0.712
	DSH	(Liu 等,2016)	0.548	-	0.551	0.558	0.562	-
	DVSQ	(Cao 等,2017a)	-	0.79	0.792	0.797	-	-
	DNNH	(Liu, 2016)	0.674	-	0.697	0.713	0.715	-
	DRSCH		-	0.618	0.622	0.623	0.628	0.641
	DTQ	(Liu 等,2018a)	-	0.798	0.799	0.801	-	-
	DTSH	(Wang 等,2016)	0.773	-	0.808	0.812	0.824	-

注:加粗字体为每列最优值,“-”表示没有实验数据。