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
05
VOL.25

ISSN1006-8961
CN11-3758/TB

# 中国图象图形学报

刊名题字：宋健 | 月刊（1996年创刊）

第25卷第5期（总第289期）
2020年5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

**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



## 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.

2020年5月 第25卷 第5期（总第289期）
中国图象图形学报
Zhongguo Tuxiang Tuxing Xuebao

# 目次

多相图像分割Vese-Chan模型连续最大流方法(第0926页)

保持区域平滑的3维动画形状高效编辑(第1019页)

树形结构卷积神经网络优化的城区遥感图像语义分割(第1043页)


## 学者观点

多源空—谱遥感图像融合方法进展与挑战
肖亮，刘鹏飞，李恒 …… 0851

## 综述

中国图像工程：2019
章毓晋 …… 0864

## 图像处理和编码

相同编码参数HEVC视频重压缩检测
潘鹏飞，姚晔，王慧 …… 0879
密集连接卷积网络图像去模糊
吴迪，赵洪田，郑世宝 …… 0890
密集网络图像哈希检索
王亚鸽，康晓东，郭军，李博，张华丽，刘汉卿 …… 0900

## 图像分析和识别

域自适应城市场景语义分割
张桂梅，潘国峰，刘建新 …… 0913
多相图像分割Vese-Chan模型连续最大流方法
王洁，潘振宽，魏伟波，徐子森 …… 0926
结合注意力机制和多属性分类的行人再识别
郑鑫，林兰，叶茂，王丽，贺春林 …… 0936
多分辨率特征注意力融合行人再识别
沈庆，田畅，王家宝，焦珊珊，杜麟 …… 0946
全卷积网络电线识别方法
刘嘉玮，李元祥，龚政，刘心刚，周拥军 …… 0956
双核压缩激活神经网络艺术图像分类
杨秀芹，张华熊 …… 0967

## 图像理解和计算机视觉

高效检测复杂场景的快速金字塔网络SPNet
李鑫泽，张轩雄，陈胜 …… 0977
结合自底向上注意力机制和记忆网络的视觉问答模型
闫茹玉，刘学亮 …… 0993
统筹图像变换与缝合线生成的无参数影像拼接
高炯笠，吴军，刘祺昌，徐刚 …… 1007

## 计算机图形学

保持区域平滑的3维动画形状高效编辑
冼楚华，杨煜，黄俊贤，李桂清 …… 1019

## 医学图像处理

肌骨超声图像特征检测及拼接
颜焕欢，张培镇，王伊侬，李璇，阳维，王青 …… 1032

## 遥感图像处理

树形结构卷积神经网络优化的城区遥感图像语义分割
胡伟，高博川，黄振航，李瑞瑞 …… 1043
融合累积变异比和集成超限学习机的高光谱图像分类
尹玉萍，魏林，刘万军 …… 1053


封面图片由武汉大学张良培教授团队提供

Volume 25, Number 5
Published May 16, 2020

# CONTENTS
## JOURNAL OF IMAGE AND GRAPHICS

Continuous max-flow method for multiphase segmentation Vese-Chan model(P0926)

Efficient 3D animation shape manipulation with region smoothness preservation(P1019)

Semantic segmentation of urban remote sensing image based on optimized tree structure convolutional neural network (P1043)


### Scholar View
Progress and challenges in the fusion of multisource spatial-spectral remote sensing images
Xiao Liang, Liu Pengfei, Li Heng ····· 0851

### Review
Image engineering in China: 2019
Zhang Yujin ····· 0864

### Image Processing and Coding
Detection of double compression for HEVC videos with the same coding parameters
Pan Pengfei, Yao Ye, Wang Hui ····· 0879
Motion deblurring method based on DenseNets
Wu Di, Zhao Hongtian, Zheng Shibao ····· 0890
Image Hash retrieval with DenseNet
Wang Yage, Kang Xiaodong, Guo Jun, Li bo, Zhang Huali, Liu Hanqing ····· 0900

### Image Analysis and Recognition
Domain adaptation for semantic segmentation based on adaption learning rate
Zhang Guimei, Pan Guofeng, Liu Jianxin ····· 0913
Continuous max-flow method for multiphase segmentation Vese-Chan model
Wang Jie, Pan Zhenkuan, Wei Weibo, Xu Zisen ····· 0926
Improving person re-identification by attention and multi-attributes
Zheng Xin, Lin Lan, Ye Mao, Wang Li, He Chunlin ····· 0936
Multi-resolution feature attention fusion method for person re-identification
Shen Qing, Tian Chang, Wang Jiabao, Jiao Shanshan, Du Lin ····· 0946
Power line recognition method via fully convolutional network
Liu Jiawei, Li Yuanxiang, Gong Zheng, Liu Xingang, Zhou Yongjun ····· 0956
Art image classification with double kernel squeeze-and-excitation neural network
Yang Xiuqin, Zhang Huaxiong ····· 0967

### Image Understanding and Computer Vision
Snap Pyramid Network: Real-time complex scene detecting system
Li Xinze, Zhang Xuanxiong, Chen Sheng ····· 0977
Visual question answering model based on bottom-up attention and memory network
Yan Ruyu, Liu Xueliang ····· 0993
Stitching of parametric-free images based on coordinated image transformation and seam-line generation
Gao Jiongli, Wu Jun, Liu Qichang, Xu Gang ····· 1007

### Computer Graphics
Efficient 3D animation shape manipulation with region smoothness preservation
Xian Chuhua, Yang Yu, Huang Junxian, Li Guiqing ····· 1019

### Medical Image Processing
Feature detection algorithm of musculoskeletal ultrasound image and its application of image stitching
Yan Huanhuan, Zhang Peizhen, Wang Yinong, Li Xuan, Yang Wei, Wang Qing ····· 1032

### Remote Sensing Image Processing
Semantic segmentation of urban remote sensing image based on optimized tree structure convolutional neural network
Hu Wei, Gao Bochuan, Huang Zhenhang, Li Ruirui ····· 1043
Ensemble extreme learning machine with cumulative variation quotient for hyperspectral image classification
Yin Yuping, Wei Lin, Liu Wanjun ····· 1053

E-mail: jig@radi.ac.cn
Website: www.cjig.cn
Tel: 010-58887035

中国图象图形学报
JOURNAL OF IMAGE AND GRAPHICS
©中国图象图形学报版权所有

**中图法分类号**:TP391　　**文献标识码**:A　　**文章编号**:1006-8961(2020)05-0900-13

**论文引用格式**:Wang Y G, Kang X D, Guo J, Li b, Zhang H L and Liu H Q. 2020. Image Hash retrieval with DenseNet. Journal of Image and Graphics,25(05):0900-0912(王亚鸽,康晓东,郭军,李博,张华丽,刘汉卿. 2020. 密集网络图像哈希检索. 中国图象图形学报,25(05):0900-0912)[DOI:10.11834/jig.190416]

# 密集网络图像哈希检索

王亚鸽,康晓东,郭军,李博,张华丽,刘汉卿

天津医科大学医学影像学院,天津　300203

**摘　要**:**目的**　为提取可充分表达图像语义信息的图像特征,减少哈希检索中的投影误差,并生成更紧致的二值哈希码,提出一种基于密集网络和改进的监督核哈希方法。**方法**　用训练优化好的密集网络提取图像的高层语义特征;先对提取到的图像特征进行核主成分分析投影,充分挖掘图像特征中隐含的非线性信息,以减少投影误差,再利用监督核哈希方法对图像特征进行监督学习,将特征映射到汉明空间,生成更紧致的二值哈希码。**结果**　为验证提出方法的有效性、可拓展性以及高效性,在 Paris6K 和 LUNA16(lung nodule analysis 16)数据集上与其他 6 种常用哈希方法相比,所提方法在不同哈希码长下的平均检索精度均较高,且在哈希码长为 64 bit 时,平均检索精度达到最高,分别为 89.2% 和 92.9%;与基于卷积神经网络的哈希算法(convolution neural network Hashing,CNNH)方法相比,所提方法的时间复杂度有所降低。**结论**　提出一种基于密集网络和改进的监督核哈希方法,提高了图像特征的表达能力和投影精度,具有较好的检索性能和较低的时间复杂度;且所提方法的可拓展性也较好,不仅能够有效应用到彩色图像检索领域,也可以应用在医学灰度图像检索领域。

**关键词**:密集卷积网络(DenseNet);监督核哈希;图像特征;投影误差;核主成分分析

# Image Hash retrieval with DenseNet

Wang Yage, Kang Xiaodong, Guo Jun, Li bo, Zhang Huali, Liu Hanqing

*School of Medical Imaging, Tianjin Medical University, Tianjin 300203, China*

**Abstract**: **Objective** To extract image features that can fully express image semantic information, reduce projection errors in Hash retrieval, and generate more compact binary Hash codes, a method based on dense network and improved supervised Hashing with kernels is proposed. **Method** The pre-processed image data set is used to train the dense network. To reduce the over-fitting phenomenon, L2 regularization term is added into the cross entropy as a new loss function. When the dense network model is training, batch normalization (BN) algorithm and root mean square prop (RMSProp) optimization algorithm are used to improve the accuracy and robustness of the model. High-level semantic features of images with trained and optimized dense network model are removed to enhance the ability of image features to express image information and build an image feature library of the image dataset. The kernel principal component analysis projection is then performed on the extracted image features. The nonlinear information implicit in the image features is fully exploited to reduce the projection error. The supervised kernel Hash method is also used to supervise the image features, enhance the resolution of the linear inseparable image feature data, and map the features to the Hamming space. According to the correspondence between the inner product of Hash code and Hamming distance and the semantic similarity monitoring matrix composed of

**收稿日期**:2019-06-14;**修回日期**:2019-09-28;**预印本日期**:2019-10-05

**基金项目**:京津冀协同创新项目(17YEXTZC00020)

**Supported by**:Beijing-Tianjin-Hebei Collaborative Innovation Project(17YEXTZC00020)

© 中国图象图形学报版权所有

image label information, the Hamming distance is optimized to generate a more compact binary Hash code. Next, the image feature Hash code library of the image dataset is constructed. Finally, the same operation is performed on the input query image to obtain the Hash code of the query image. The Hamming distance between the Hash code of the query image and the Hash code of the image feature in the image dataset is compared to measure the similarity. The retrieved similar images are returned in ascending order. **Result** To verify the effectiveness, expansion, and efficiency of the proposed method, our method is used respectively in Paris6K and lung nodule analysis 16 (LUNA16) datasets. It is also compared with other six commonly used Hashing methods. The average retrieval accuracy is compared in 12, 24, 32, 48, 64, and 128 bits of code length. Experimental results show that the average retrieval accuracy increases with the increase of Hash code length. When the Hash code length increases to a certain value, the average retrieval accuracy decreases. The average retrieval accuracy of the proposed method is always higher than that of the other six Hash methods. Except for the semantic Hashing method, the average retrieval accuracy value reaches the maximum when the Hash code length is 48 bits. Other Hash methods, including the proposed method, have the maximum average retrieval accuracy value when the Hash code length is 64 bits, and the retrieval accuracy is better. When the Hash code length is 64 bits, the average retrieval accuracy value of the proposed method is as high as 89.2% and 92.9% in the Paris6K and LUNA16 datasets, respectively. The time complexity of the proposed method and the convolutional neural network (CNN) Hashing method is compared in the Paris6K and LUNA16 data sets when the Hash code length is 12, 24, 32, 48, 64, and 128 bits. Results show that the time complexity of the proposed method is reduced under different Hash code lengths and is efficient to a certain degree. **Conclusion** A method based on dense network and improved supervised Hashing with kernels is proposed. This method improves the expression ability of image features and projection accuracy and is superior to other similar methods in average retrieval accuracy, recall rate, and precision rate. It improves the retrieval performance to some extent. It has a lower time complexity of algorithm than the method of CNN Hashing method. In addition, the proposed method has better extensibility, which can be used not only in the field of color image retrieval but also in the field of medical gray scale image retrieval.

**Key words**: dense convolutional network (DenseNet); supervised Hashing with Kernels; image features; projection error; kernel principal component analysis (KPCA)

## 0 引　言

图像检索技术主要是通过分析图像特征，达到查找相似图像的目的(康晓东，2009)。随着技术的进步，图像数据增长迅猛，如何在海量图像数据中实现快速有效的检索已成为亟待解决的问题。目前主流的图像检索方法是基于内容的图像检索(content-based image retrieval，CBIR)，其核心是从数据集中找到与给定图像最相似的图像，考虑到其在存储空间与检索时间的局限，近似最近邻(approximate nearest neighbor，ANN)检索方法广泛应用于基于内容的图像检索中(Shen 等，2016)。基于哈希的方法是 ANN 中重要且有效的方法，它通过将数据嵌入到汉明空间来减少存储消耗并抑制查询复杂度的线性增长(李武军和周志华，2015)，克服了传统检索算法对大规模数据在存储空间和检索时间的不合理要求。图像哈希算法是一种有效快速的图像检索方法(Vadlamudi 等，2017)，它不需要存储原始图像的特征数据，只需存储相应特征图像的哈希码即可，可以大大减少存储空间。哈希算法分为数据独立哈希和数据依赖哈希。数据独立哈希的典型算法是局部敏感哈希(locality-sensitive Hashing，LSH)(Har-Peled 等，2012)和平移不变核哈希(shift invariant Kernel Hashing，SIKH)(Raginsky 和 Lazebnik，2009)，数据独立哈希算法的缺陷是没有考虑到数据分布的空间关系，由随机确定的哈希函数得到的哈希码不能很好地保持原始特征空间的相似关系，也不能很好地处理线性不可分问题，往往需要较长的哈希码才能达到较好的检索精度，对存储空间的需求较高。与数据独立哈希算法相比，在相同的训练集上数据依赖的哈希算法可以得到与数据更加适配的哈希函数，在使用相同的比特数下，可以取得更好的检索精度。依据给定的数据训练集是否有标签，数据依赖哈希可以分为非监督数据依赖哈希和监督数据依赖哈希。非监督数据依赖哈希的代表性算法是迭代量化哈希(iterative quantization，ITQ)(Gong 等，2013)和谱哈希(spectral Hashing，SH)(Weiss 等，2008)。

© 中国图象图形学报版权所有

ITQ 算法通过学习旋转矩阵,有效解决了信息量分布不均衡的问题,获得了较好的检索精度;SH 算法在哈希码学习过程中应用图划分技术,并采用谱松弛技术得到最后的二值哈希码,在一定程度上避免了二值哈希码的冗余问题;它们的优点是检索速度快,缺点是不能很好地利用图像包含的丰富语义信息。监督数据依赖哈希可以在一定程度上避免图像语义信息的丢失,代表性算法有半监督哈希(semi-supervised Hashing,SSH)(Wang 等,2010)和监督学习核哈希(supervised Hashing with Kernels, KSH)(Liu 等,2012)。SSH 使用小部分图像数据标签来提供监督信息,并采用正则化方法来减少映射矩阵的参数数量,在一定程度上避免了过拟合问题,同时缩短了检索时间;KSH 利用内积乘积和汉明距离之间的关联,构造在核空间中学习非线性的哈希函数,很好地处理了数据线性不可分的情况。

图像哈希的检索精度很大程度上取决于充分高效的图像特征提取,随着深度学习技术的飞速发展,图像检索领域也广泛应用深度学习的方法提取图像特征(Liu 等,2012)。Razavian 等人(2014)首先使用 ImageNet 训练的卷积神经网络模型,提取其第 1 个全连接层输出的深度特征用于图像检索。虽然卷积神经网络模型可以提取到图像的深层语义特征,但也存在网络设计复杂、训练时间长等问题,因此通过设计先进合理的神经网络结构可以有效提高图像特征提取。AlexNet(Alex network)(Krizhevsky 等,2012)和 VGG(visiual geometry group)(Simonyan and Zisserman,2014)是两个重要的深度卷积神经网络,它们证明了用较小的卷积核构建更深层次的网络可以有效提高卷积神经网络的学习能力。之后 He 等人(2015)提出了神经网络层数达数百层的残差网络(residual network, ResNet),降低了优化难度。Huang 等人(2016)提出的密集卷积网络(dense convolutional network, DenseNet)将输入特征图和输出特征图通过密集块连接起来,而不是简单的相加,这也是其与 ResNet 的主要区别。

随着人工智能的快速兴起,基于深度学习的图像哈希方法得到了广泛应用。Salakhutdinov 和 Hinton(2009)提出的语义哈希(semantic Hashing,SH)最早使用深度学习技术进行无监督哈希,该方法输入特征的本质仍属人工提取,与现有的深度哈希有所不同。Xia 等人(2014)提出了一种基于卷积神经网络的哈希算法(convolutional neural network Hashing, CNNH),该算法是 CNN 与哈希结合的一次全新尝试,与 SH 算法相比,该算法采用 CNN 模型提取的特征作为输入,检索性能得到了显著提升。随后出现了很多 CNN 与哈希结合的深度哈希算法,如 Lin 等人(2015)提出了深度学习二进制哈希码(deep learning of binary Hash codes, DLBHC),Lai 等人(2015)提出了深度神经网络哈希(deep neural network Hashing, DNNH),Yang 等人(2018)提出了有监督语义深度哈希(supervised semantics deep Hashing, SSDH)。

考虑到深度哈希在图像检索领域的广泛应用,本文针对基于内容的图像检索存在的特征提取不足和图像检索精度低的问题,提出一种结合 DenseNet 和改进的监督核哈希的图像检索方法。首先,用 DenseNet 提取图像特征,以增强图像特征对图像信息的表征能力;其次,引入改进的监督核哈希方法,先对图像特征进行核主成分分析(Kernel principal component analysis,KPCA),以减少映射误差,再使用监督核哈希增强对线性不可分的特征数据的分辨力,依据哈希码内积与汉明距离的对应关系以及由图像标签信息构成的语义相似性监督矩阵,来优化汉明距离,生成高质量的哈希编码,并在汉明空间中通过相似性度量来完成有效检索。

## 1 基于 DenseNet 的图像特征提取

图像特征提取是基于内容图像检索的关键环节,对图像检索的精度依赖于能否充分有效地提取图像特征以及能否最大限度地表征图像中包含的信息。随着深度学习的发展,通过卷积神经网络提取图像特征得到了广泛应用,与传统的尺度不变特征变换(scale invariant feature transform,SIFT)(Lowe, 2004)和方向梯度直方图(histogram of oriented gradient, HOG)(Dalal 和 Triggs,2005)特征提取方法相比,深度学习的方法具有强大的特征自学和特征表达能力,提取的特征对图像内容的表现力强,能够很好地反映图像间的语义相似性,并且深度学习模型运用局部连接和权值共享策略,有良好的容错能力、泛化性和鲁棒性。

鉴于深度学习在图像特征提取方面的优势,本文采用 DenseNet 提取图像特征。DenseNet 脱离了加深网络结构和加宽网络结构来提升网络性能的定

© 中国图象图形学报版权所有

式思维，通过特征重用和旁路（bypass）设置，大幅度减少了网络的参数量，加强了特征传播，鼓励了特征重用，也在一定程度上缓解了梯度消失的产生（Huang 等，2016）。

### 1.1　DenseNet 结构

DenseNet 主要由密集块（dense block）和过渡层（transition layer）交替连接组成。DenseNet 最大化了网络中所有层之间的信息流，每一层都把前面所有层的输出特征图连接起来作为自己的输入，然后再把自己的输出输送给之后的所有层。本文采用的 DenseNet 模型结构如图 1 所示，由 4 个密集块和 3 层过渡层组成。密集块的结构如图 2 所示，密集块的主要作用是提取更多的图像隐含特征并增强图像特征表达，每个密集块包含多层 BN_ReLU_Conv，4 个密集块包含的 BN_ReLU_Conv 层数分别为 6、12、48 和 32，其中的卷积操作分别为 1 × 1 和 3 × 3，1 × 1 的卷积操作又称瓶颈层（bottleneck layer），把它放在 3 × 3 的卷积操作前的作用是降维，减少计算量并融合各个通道的特征。密集块中，每个 BN_ReLU_Conv 的输出和输入连接在一起传输给下一个 BN_ReLU_Conv，使得每个 BN_ReLU_Conv 都彼此相连，并把特征图都连接起来，而不是简单的加起来，DenseNet 的特征连接方式如公式（1）所示，具有正则化的效果，可以在一定程度上抑制过拟合现象。

$$x_\varphi = H_\varphi([x_0, x_1, \cdots, x_{\varphi-1}]) \tag{1}$$

式中，$H_\varphi$ 代表复合函数集合，包含 BN_ReLU_Conv 层所有函数的组合操作。$[x_0, x_1, \cdots, x_{\varphi-1}]$ 表示将 0 到 $\varphi - 1$ 层的输出特征图进行串联连接。

过渡层由批量归一化层（batch normalization，BN）、瓶颈层和池化层组成。瓶颈层在过渡层的作用是进一步压缩参数数量；池化层为池化核大小为 2 × 2 、步长为 2 的平均池化。在密集块后加入过渡层来压缩其输出的特征图，从而保证训练的高效性。

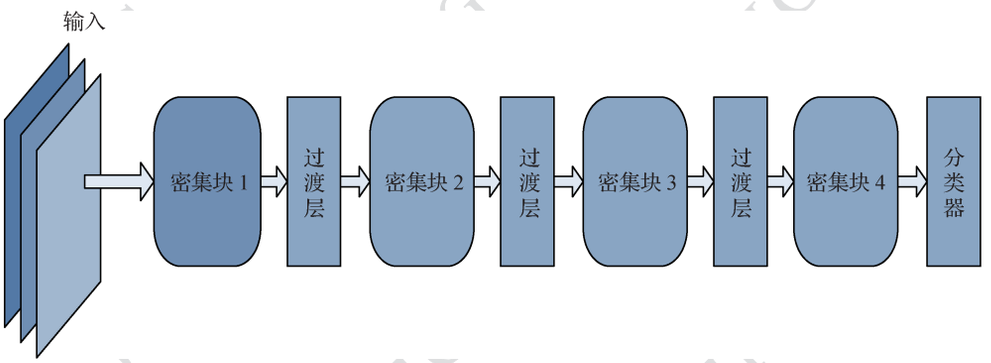



图 1　DenseNet 模型结构图

Fig. 1　DenseNet model structure diagram

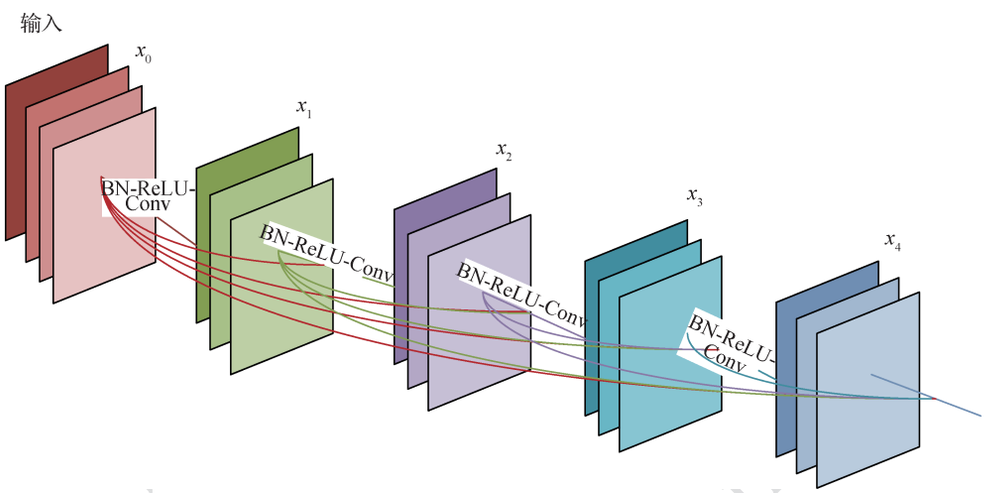



图 2　密集块结构图

Fig. 2　Structural diagram of dense block

© 中国图象图形学报版权所有

### 1.2 BN 算法

图 1 中用到的 BN 算法(Ioffe 和 Szegedy,2015)是对每一个神经元进行归一化处理,即对输入激活函数 ReLU 的数据进行归一化处理,可以很好地解决输入数据发生偏移和增大的影响,即

$$\hat{x^{(k)}} = \frac{x^{(k)} - E[x^{(k)}]}{\sqrt{var[x^{(k)}]}} \tag{2}$$

式中,$\hat{x^{(k)}}$ 表示经 BN 算法处理后的数据,$E[x^{(k)}]$ 表示每个 batch 训练数据神经元 $x^{(k)}$ 的平均值,$\sqrt{var[x^{(k)}]}$ 表示每个 batch 训练数据的标准差。直接使用式(2)可能会导致数据分布发生破坏,因此在式(2)的基础上加上两个可学习参数 $\gamma$、$\beta$,优化公式的鲁棒性,从而保持模型的表达能力,即

$$y^{(k)} = \gamma^{(k)} \hat{x^{(k)}} + \beta^{(k)} \tag{3}$$

在训练过程中,每个神经元都有可学习重构参数 $\gamma$、$\beta$,当 $\beta^{(k)} = E[x^{(k)}]$,$\gamma^{(k)} = \sqrt{var[x^{(k)}]}$ 时,网络可以学习恢复出原始的某一层网络所要学习的特征分布。将该算法用在本文网络中,通过参数共享策略以每个特征图为单元求取可学习参数 $\gamma$、$\beta$,然后对特征图进行神经元的归一化。

### 1.3 损失函数

损失函数是用来估量模型的预测值与真实值差异程度的非负实值函数,损失函数越小,模型的鲁棒性就越好。交叉熵(binary cross-entroy)是机器学习中常用的针对概率之间的损失函数,对网络训练过程中得到的概率分布和真实分布的差异情况进行评估,交叉熵损失越小,网络模型的预测准确率越高,特征提取与表征能力也越好。本文 DenseNet 模型采用交叉熵作为损失函数,为减少过拟合现象和提升模型的鲁棒性,在 DenseNet 模型的交叉熵损失函数中加入 L2 正则化项,即

$$L_1 = L_0 + \frac{\lambda}{2n}\sum_w w^2 \tag{4}$$

式中,$L_1$ 表示加入 L2 正则化项后的损失函数,$L_0$ 是交叉熵函数,即

$$L_0 = -\sum_{i=1}^{n} \hat{y}_i \ln y_i + (1 - \hat{y}_i)\ln(1 - \hat{y}_i) \tag{5}$$

由式(4)对权重 $w$ 和偏置 $b$ 求偏导,得

$$\frac{\partial L_1}{\partial w} = \frac{\partial L_0}{\partial w} + \frac{\lambda}{n} w \tag{6}$$

$$\frac{\partial L_1}{\partial b} = \frac{\partial L_0}{\partial b} \tag{7}$$

式中,$n$ 表示训练集的样本数,$\lambda$ 是 L2 正则化项的参数。参数更新为

$$w = w - \frac{\eta\lambda}{n} w - \frac{\eta}{m}\sum_m \frac{\partial L_m}{\partial w} \tag{8}$$

$$b = b - \frac{\eta}{m}\frac{\partial L_m}{\partial b} \tag{9}$$

式中,$m$ 表示梯度随机下降的 batch size,$\eta$ 表示参数更新时的步长。

### 1.4 RMSProp 优化算法

RMSProp(root mean square prop)算法(Tieleman 和 Hinton, 2012)采用小批量随机梯度,并以元素平方的指数加权移动平均来调整学习率,不需要设置初始学习率,从而进一步优化损失函数在更新过程中摆动幅度大的问题并加快收敛速度。RMSProp 算法分别计算了权重 $w$ 和偏置 $b$ 梯度的微分平方加权平均数,有利于消除摆动幅度大的方向,修正摆动幅度,从而使各维度的摆动幅度都较小并加快网络收敛速度,计算为

$$s_{dw} = \gamma s_{dw} + (1 - \gamma)(dw)^2 \tag{10}$$

$$s_{db} = \gamma s_{db} + (1 - \gamma)(db)^2 \tag{11}$$

$$w = w - \delta\frac{dw}{\sqrt{s_{dw}} + \varepsilon} \tag{12}$$

$$b = b - \delta\frac{db}{\sqrt{s_{db}} + \varepsilon} \tag{13}$$

式中,若当前的迭代轮数为 $q$,则 $s_{dw}$ 和 $s_{db}$ 分别为损失函数在前 $q-1$ 轮迭代过程中累积的权重和偏置梯度动量,$\gamma$ 是梯度累积指数,$dw$ 和 $db$ 分别是损失函数反向传播时所求得的权重和偏置梯度,$\delta$ 是网络的学习率,$\varepsilon$ 是为防止分母为零的取值很小的平滑数值,一般取值为 $10^{-8}$。

## 2 改进的监督核哈希检索

图像哈希检索可以分为投影和量化两个阶段(栾婷婷 等, 2019),在大规模的图像检索中,对投影阶段进行优化,可以有效地减少存储空间和降低计算代价,因此本文提出一种改进的监督核哈希算法。将 DenseNet 提取的高维图像特征先进行 KPCA 投影,来减少映射矩阵的参数数量,并可在一定程度上减少映射误差,再对投影后的特征数据进行 KSH 监督学习,生成高质量的哈希编码,从而提高检索精度。

© 中国图象图形学报版权所有

### 2.1 KPCA 投影

KPCA 是一种可以保留全局性质的非线性投影方法，能够挖掘出数据中蕴含的非线性信息（Hotelling，1993）。它的基本思想是通过某种隐式方式将输入空间映射到某个特征空间，并在特征空间中实现 PCA。KPCA 首先引入隐式的非线性映射函数 $\boldsymbol{\phi}$，将样本 $x_i(i=1,2,\cdots,N,x_i\in\mathbf{R}^d)$ 转换到更高维的 $K$ 维子空间，$\boldsymbol{\phi}$ 的空间定义为

$$\boldsymbol{\phi}:\mathbf{R}^d\rightarrow\mathbf{R}^k(k\gg d) \tag{14}$$

$K$ 空间中样本的协方差矩阵为

$$\boldsymbol{C}=\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\phi}(x_i)\boldsymbol{\phi}^{\mathrm{T}}(x_i) \tag{15}$$

对 $\boldsymbol{C}$ 进行特征分解，并求得分解后的特征向量 $\boldsymbol{v}$，即

$$\lambda\boldsymbol{v}=\boldsymbol{C}\boldsymbol{v} \tag{16}$$

$$\boldsymbol{v}=\frac{1}{N}\sum_{i=1}^{N}\alpha\boldsymbol{\phi}(x_i) \tag{17}$$

式中，$\boldsymbol{v}$ 为 $\boldsymbol{C}$ 的特征向量，$\lambda$ 是与特征向量对应的特征值，$\alpha=\frac{1}{\lambda}\boldsymbol{\phi}^{\mathrm{T}}(x_i)$，是公式变换中的变量。

将式（16）左右同时乘以 $\boldsymbol{\phi}(x_i)$，则有

$$\lambda(\boldsymbol{\phi}(x_i)\boldsymbol{v})=\boldsymbol{\phi}(x_i)\boldsymbol{C}\boldsymbol{v} \tag{18}$$

定义一个 $N\times N$ 的矩阵 $\boldsymbol{k}$，且有

$$k(x_i,y_i)=\boldsymbol{\phi}(x_i)\boldsymbol{\phi}(x_j) \tag{19}$$

由式（14）、（15）计算在特征向量 $\boldsymbol{v}$ 上的投影为

$$y(i)=\sum_{j=1}^{N}\alpha_i(j)\cdot k(x_i,x_j) \tag{20}$$

式中，$k(x_i,y_i)$ 为核函数。

### 2.2 KSH 哈希编码

鉴于监督核哈希可以很好地增强线性不可分数据的分辨力，采用 KSH 方法将经 KPCA 投影后的图像特征映射为二值哈希码，若哈希码长为 $n$，则需设计 $n$ 个哈希函数，即 $\boldsymbol{H}=\{h_1,h_2,\cdots,h_n\}$，哈希函数的一般形式为

$$h_k(x_i)=\mathrm{sgn}(\boldsymbol{w}_k^{\mathrm{T}}x_i+b_k) \tag{21}$$

式中，$\boldsymbol{w}$ 为需要学习的系数向量，$b$ 表示偏差，为生成均衡的哈希码，$b$ 通常被赋值为样本中全部图像特征的期望值，由于本文先对图像特征进行了 KPCA 投影，其中已包含了零均值化操作，故将 $b$ 赋值为 0，所以可用 KPCA 投影后的特征 $y_i$ 代替式（21）中的 $x_i$，简化为

$$h_k(\boldsymbol{y}_i)=\mathrm{sgn}(\boldsymbol{w}_k^{\mathrm{T}}\boldsymbol{y}_i) \tag{22}$$

利用图像的标签信息求解式（22）中的系数向量 $[\boldsymbol{w}_1,\boldsymbol{w}_2,\cdots,\boldsymbol{w}_k]$，过程如下：

1）用图像标签信息来表征样本间的语义相似性，采用 $l$ 个样本标签构建标签矩阵 $\boldsymbol{s}\in\mathbf{R}^{l\times l}$，矩阵元素为

$$s_{ij}=\begin{cases}1 & label(\boldsymbol{y}_i,\boldsymbol{y}_j)=1\\ 0 & \text{其他}\\ -1 & label(\boldsymbol{y}_i,\boldsymbol{y}_j)=-1\end{cases} \tag{23}$$

式中，$s_{ij}=1$，$label(\boldsymbol{y}_i,\boldsymbol{y}_j)=1$，表示图像 $\boldsymbol{y}_i$，$\boldsymbol{y}_j$ 相似；相反，$s_{ij}=-1$，$label(\boldsymbol{y}_i,\boldsymbol{y}_j)=-1$，表示图像差异大；$s_{ij}=0$ 表示图像 $\boldsymbol{y}_i$，$\boldsymbol{y}_j$ 之间的相似性不确定。为了增强哈希码的分辨力，以便在汉明空间可以有效地判断图像间的相似性，应使图像 $\boldsymbol{y}_i$，$\boldsymbol{y}_j$ 的 $s_{ij}$ 和汉明距离 $d(\boldsymbol{y}_i,\boldsymbol{y}_j)$ 满足

$$d(\boldsymbol{y}_i,\boldsymbol{y}_j)=\begin{cases}0 & s_{ij}=1\\ r(r\neq 0) & s_{ij}=-1\end{cases} \tag{24}$$

2）用内积法计算相似度。考虑到用汉明距离相似度反推系数向量时，异或运算很难求导，因此需要将异或运算转化为汉明码间的内积运算，来计算哈希码间的距离，记图像 $\boldsymbol{y}$ 的哈希码长为 $code_r(\boldsymbol{y})=[h_1(\boldsymbol{y}),\cdots,h_r(\boldsymbol{y})]\in\{1,-1\}^{l\times r}$，转化公式为

$$\begin{aligned}D(\boldsymbol{y}_i,\boldsymbol{y}_j)&=code_r(\boldsymbol{y}_i)\circ code_r(\boldsymbol{y}_j)=\\ &\left|\{k\,|\,h_k(\boldsymbol{y}_i)=h_k(\boldsymbol{y}_j),1\leqslant k\leqslant r\}\right|-\\ &\left|\{k\,|\,h_k(\boldsymbol{y}_i)\neq h_k(\boldsymbol{y}_j),1\leqslant k\leqslant r\}\right|=\\ &r-2\left|\{k\,|\,h_k(\boldsymbol{y}_i)\neq h_k(\boldsymbol{y}_j),1\leqslant k\leqslant r\}\right|=\\ &r-2d(\boldsymbol{y}_i,\boldsymbol{y}_j)\end{aligned} \tag{25}$$

式（25）表明哈希码内积运算与汉明距离运算的一致性，且内积 $D(\boldsymbol{y}_i,\boldsymbol{y}_j)\in[-r,r]$，为将其归一化到 $[-1,1]$，需让内积除以 $r$。

3）用样本标签的内积矩阵和标签矩阵来定义目标函数，即

$$\min_{\boldsymbol{H}_l\subset\{1,-1\}^{l\times r}}\left\|\frac{1}{r}\boldsymbol{H}_l\boldsymbol{H}_l^{\mathrm{T}}-r\boldsymbol{S}\right\|_{\mathrm{F}}^2 \tag{26}$$

式中，$\|\cdot\|_{\mathrm{F}}^2$ 表示求矩阵 Frobenius 范数，$\boldsymbol{H}_l\boldsymbol{H}_l^{\mathrm{T}}$ 表示样本标签的内积矩阵，$\boldsymbol{S}$ 为标签矩阵，$l$ 为样本数目，$r$ 为哈希码长。

4）用贪心算法求解目标函数，并在求解过程中依次对目标函数进行频谱化宽松和 Sigmoid 平滑处理，从而减小求解难度。

获得系数向量 $\boldsymbol{w}$ 后，生成哈希函数和哈希表 $\boldsymbol{H}$，通过对查询图像的图像特征进行哈希函数映射，可生成特征长度的哈希码 $code_r(\boldsymbol{y})$，计算 $code_r(\boldsymbol{y})$

© 中国图象图形学报版权所有

和哈希表 $\boldsymbol{H}$ 中哈希码的汉明距离，并以升序返回的相似图像集作为检索结果。

## 3 本文方法

### 3.1 数据集及其处理

选用 Paris6K（Philbin 等，2008）图像检索数据集，如图 3 所示。Paris6K 数据集共有 6 412 幅图像，实际可用 6 392 幅。数据集中有 11 种巴黎地标建筑物图像，每种建筑物有 5 幅查询图像，共计 55 幅查询图像，每幅图像都有一个标签。

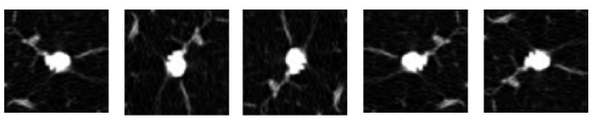
图 3 图像检索数据集示例图像

Fig. 3 Image retrieval dataset sample image

为进一步验证本文方法的有效性和可拓展性，还选用 Kaggle 算法的标准数据集 LUNA16（lung nodule analysis 16），它在美国癌症中心公开的 LIDC/IDRI 数据（lung image database consortium/image database resource initiative）（Armato 等，2011）的基础上，将切片厚度大于 3 mm 的、切片 space 不一致以及部分切片缺失的肺部 CT 扫描去除，包括 888 个病例，1 186 个结节。LUNA16 共有 10 个子文件夹，每个文件夹里都是病例，每个病例对应两个文件名相同但后缀不同文件，后缀为 .mhd 和 .raw 的文件分别存放 CT 基本信息和实际 CT 数据。

考虑到不同设备采集的 CT 图像的动态范围有很大不同，致使 CT 图像的最大值和最小值有较大差距以及 DenseNet 对输入图像数据的要求，需要对数据进行预处理。首先对数据进行体素归一化处理，考虑到体素值的物理意义，将肺结节可能出现的体素范围从原来的（－1 000，400）归一化到（0，1），以便 DenseNet 从中提取到有效的图像特征，并将 CT 图像的 HU（hounsfield unit）值转为灰度值。其次将依据 LUNA16 数据集的标注信息提取出的包含结节信息和候选结节信息的图像，保存为 .jpg 格式。

考虑到 DenseNet 网络对输入图像的要求是三通道图像，通过计算机视觉库 OpenCV 将灰度图像转换为三通道图像。最后为提升 DenseNet 模型的鲁棒性和泛化性以及均衡正负样本比例，对正样本进行数据扩增操作，即对正样本数据分别进行向右和向左旋转 90°、水平翻转、垂直翻转；按照随机种子数为 42、训练集占比为 0.7 来划分数据集。图 4 为随机选取的肺结节样本，及其经数据扩增后的图像，图 4（a）为肺结节样本图像，图 4（b）—（e）为数据增强后的图像。



(a) 肺结节图像 (b) 右旋90° (c) 左旋90° (d) 水平翻转 (e) 垂直翻转

图 4 数据扩增后的肺结节图像

Fig. 4 Augmented images of pulmonary nodules

((a) pullmonary nodule image; (b) rotate the image 90 degrees tothe right; (c) rotate the image 90 degrees to the left; (d) flip the image horizontally; (e) flip the image vertically)

### 3.2 算法流程

本文方法流程如图 5 所示。为提高图像特征的表征能力和提升哈希码质量，提出了基于 DenseNet 和改进的监督核哈希图像检索方法。首先，用预处理过的图像数据集对 DenseNet 进行训练，为减少过拟合现象，在交叉熵中加入 L2 正则化项作为新的损失函数；为增加模型的准确性和鲁棒性，还在训练网络模型时使用了 BN 算法和 RMSProp 优化算法；用得到的最优 DenseNet 模型提取数据集的图像特征，并构建数据集的图像特征库。其次，用改进的监督核哈希图像检索方法对图像特征进行哈希编码，即先对图像特征进行 KPCA 投影，以减少投影误差；再对投影后的图像特征进行 KSH 编码，以生成高质量的二值哈希码，并构建数据集的图像特征哈希码库；最后对输入的查询图像进行相同操作，得到查询图像的哈希码，通过比较查询图像的哈希码与数据集的图像特征哈希码库中的哈希码间的汉明距离进行相似性度量，将检索的数据集图像按照汉明距离大

© 中国图象图形学报版权所有

小以升序返回，从而获得与查询图像相似度高的前 $k$ 幅图像。汉明距离计算公式为

$$d(x,y) = \sum x \oplus y \tag{27}$$

式中，$x$，$y$ 分别表示查询图像和数据集图像的哈希编码，$\oplus$ 表示异或运算。汉明距离越小，两幅图像相似度越高。

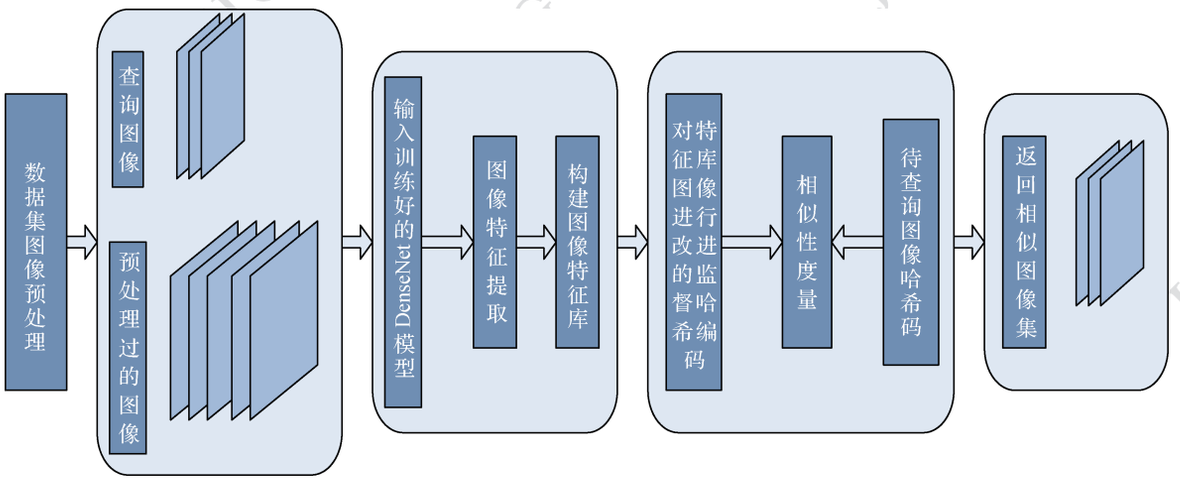



图 5　本文方法流程图

Fig. 5　Flow chart of proposed algorithm

## 4　实验结果与分析

### 4.1　实验环境与评价指标

实验环境为戴尔 Precision Tower5810 图形工作站、Intel(R) Xeon(R) CPU E5-1603 v4@ 2.8 GHz 处理器、Windows10 64 位操作系统。在 Python3.6 平台上采用以深度学习框架 Tensorflow 为后端的开源库 Keras 实现。

为了评估本文方法的检索性能，选取查准率(precision)、查全率(recall)、平均检索精度(mean average precision，MAP)和算法的时间复杂度作为评价指标。表 1 列出了可能出现的检索结果。

**表 1　检索结果情况描述**

**Table 1　Description of retrieval results**

| 检索情况 | 相关 | 不相关 |
|---|---|---|
| 检索出 | 相似图像数 $A$ | 不相似图像数 $C$ |
| 未检索出 | 相似图像数 $B$ | 不相似图像数 $D$ |

precision 和 recall 分别被定义为

$$p_{\text{all}} = \frac{A}{A+C}$$

$$r_{\text{all}} = \frac{A}{A+B}$$

式中，$p_{\text{all}}$为 precision，表示一次查询返回的相似图像数占所有返回图像数目比，$r_{\text{all}}$为 recall，表示返回结果中相似图像数占数据集中所有相似图像数目比，MAP 是一组检索图像中每个检索图像的准确度平均值。MAP 值的计算过程分为两步，首先计算平均准确率 AP，若经检索后返回 $M$ 个相关图像，位置分别为 $x_1$，$x_2$，…，$x_m$，则单个类别的平均准确率 AP 可表示为

$$A_i = \frac{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_m}}{M} \tag{28}$$

然后，求 AP 的算术平均，图像类别数为 $N$，则 MAP 的计算公式为

$$M_{\text{all}} = \frac{1}{N}\sum_{i=1}^{N} A_i \tag{29}$$

### 4.2　结果分析

为验证本文方法的有效性，将实验结果分别与 DNNH，CNNH，KSH，BRE(binary reconstructive embeddings)(Kulis 和 Darrell，2009)，ITQ 和 SH 这 6 种常见哈希方法进行了对比。其中 DNNH 和 CNNH 是基于 CNN 的哈希方法，KSH 和 BRE 是传统的监督学习哈希方法，ITQ 和 SH 是典型的非监督学习哈希方法。在 Paris6K 数据集上，为验证哈希码长对检索精度的影响，实验比较了不同方法在 12 bits、24 bits、32 bits、48 bits、64 bits 和 128 bits 下的 MAP 值，结果如图 6 所示。从图 6 可看出，有监督学习哈希方法的平均检索精度优于非监督学习哈希方法的

© 中国图象图形学报版权所有

平均检索精度。随着哈希码长的增加，MAP 值也逐渐增大，当哈希码长增加到一定值后，MAP 值呈下降趋势，且本文方法的 MAP 值始终高于其他方法；除了 SH 方法在哈希码长为 48 bits 时，MAP 值达到最大，包括本文方法在内的其他方法均在哈希码长为 64 bits 时，MAP 值最大，检索精度较好。与检索精度较好的 DNNH 和 CNNH 方法相比，本文方法的 MAP 值分别提高了 1.4%、1.0%、1.8%、1.9%、2.8%、1.5% 和 2.9%、2.4%、2.9%、2.7%、4.5%、5.2%，在一定程度上表明本文方法的检索精度较好。

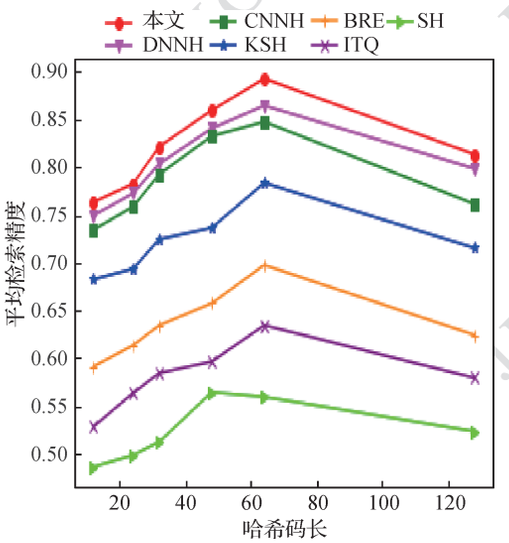



图 6　不同哈希码长下的 MAP 值

Fig. 6　MAP values with different Hash code lengths

由图 6 可看出，本文方法与 DNNH、CNNH、KSH、BRE 和 ITQ 方法，均在哈希码长为 64 bits 时平均检索精度最高，为进一步比较本文方法与其他方法的检索性能，绘制了不同方法在哈希码长为 64 bits 时的 P-R(precision-recacl)曲线，如图 7 所示。从图 7 可以看出，随着 recall 值的增大，precision 值逐渐减小，且在相同的查全率下，本文方法的查准率始终高于其他方法，即本文方法在保证一定的查全率下，可以取得较好的查准率；在相同的查准率下，本文方法的查全率也始终高于其他方法，在一定程度上表明了本文方法的检索性能较好，可以兼顾查全率与查准率。

图 8 和图 9 为在哈希码长为 64 bits 时，本文方法和其他方法的返回图像数目与 recall 值和 precision 值的关系曲线。图 8 和图 9 表明随着返回图像数的增多，查全率逐渐增高，查全率逐渐降低，且本文方法始终高于其他方法，从而表明本文方法的检

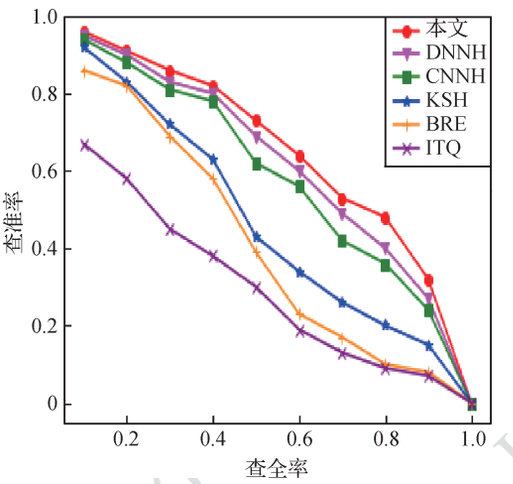



图 7　哈希码为 64 bits 时不同哈希方法的 P-R 曲线

Fig. 7　P-R curves of different Hashing methods at 64 bits

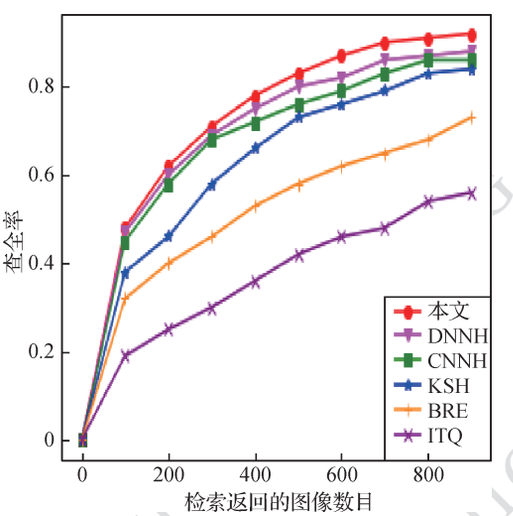



图 8　哈希码为 64 bits 时返回不同图像数量时的查全率

Fig. 8　Recall value of returning different images at 64 bits

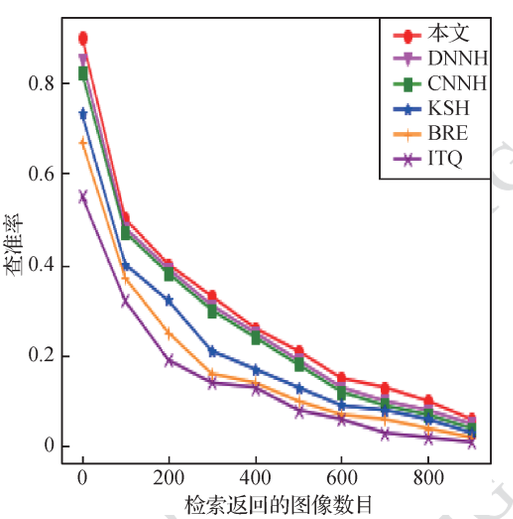



图 9　哈希码为 64 bits 时返回不同图像数量时的查准率

Fig. 9　Precision value of returning different images at 64 bits

© 中国图象图形学报版权所有

索性能有所提升。

由图 6—图 9 及其结果分析可以看出，本文方法在 Paris6K 数据集上取得较好的检索结果，验证了本文方法的有效性和可行性。图 10 显示了本文方法在 Paris6K 数据集上，哈希码长为 64 bits 时的查询图像及其相应的检索返回的前 10 幅图像，其中每行的首幅图像为查询图像，从中可以看出返回图像均与查询图像相关。

为进一步验证本文方法的有效性和可拓展性，在 LUNA16 数据集上验证哈希码长对检索精度的影响，实验对比了不同方法在不同哈希码长下的平均检索精度，结果如表 2 所示。

如表 2 所示，有监督学习哈希方法的检索精度优于非监督哈希方法的检索精度。随着哈希码长的增加，MAP 值也逐步增大，除了 SH 方法在哈希码长为 48 bits 时，MAP 值最大，检索精度最高，包括本文方法在内的 6 种哈希方法，均在哈希码长为 64 bits 时，MAP 值最大，检索精度最高，且本文方法的检索精度高于其他方法。与 MAP 值较高的 DNNH 和 CNNH 相比，本文方法的 MAP 值分别提高了 9.1%、5.5%、5.3%、5.9%、4.6%、5.9% 和 10.9%、18.3%、15.6%、13.8%、10.3%、13.4%，在一定程度上说明了本文方法的检索效果较好。

图 10　Paris6K 数据集上的查询及返回图像

Fig. 10　Query and return images on Pris6K dataset

**表 2　不同哈希码长下的 MAP 值**

**Table 2　MAP values under different Hash lengths**

| 方法 | 哈希码长/bits | | | | | |
|---|---|---|---|---|---|---|
| | 12 | 24 | 32 | 48 | 64 | 128 |
| 本文 | 0.754 | 0.880 | 0.894 | 0.921 | **0.929** | 0.908 |
| DNNH | 0.663 | 0.825 | 0.841 | 0.862 | **0.883** | 0.849 |
| CNNH | 0.645 | 0.697 | 0.738 | 0.783 | **0.826** | 0.774 |
| KSH | 0.592 | 0.653 | 0.665 | 0.673 | **0.698** | 0.669 |
| BRE | 0.536 | 0.584 | 0.603 | 0.624 | **0.679** | 0.617 |
| ITQ | 0.435 | 0.567 | 0.657 | 0.733 | **0.775** | 0.638 |
| SH | 0.384 | 0.407 | 0.432 | **0.463** | 0.452 | 0.428 |

注：加粗字体为每行最优值。

考虑到不同方法均在哈希码长为 64 bits 时，MAP 值达到最大，检索精度最好，因此，绘制了哈希码长为 64 bits 时，不同方法的 P-R 曲线，如图 11 所示。从图 11 可以看出，随着 recall 值的增大，precision 值呈逐渐减小趋势，且在相同的 recall 值下，本文方法的 precision 值始终高于其他方法，在相同的 precision 值下，本文方法的 recall 值也始终高于其他方法，即本文方法可以兼顾检索的查全率和查准率，说明了本文方法的检索性能较好，本文方法的有效性和可拓展性也得到了一定的验证。

图 12 和图 13 为在哈希码长为 64 bits，返回不同图像数目时，本文方法和其他方法的 recall 值和 precision 值的关系曲线，图 12 表明随着返回图像数目的增加，recall 值逐渐增大，且本文方法的 recall 值始终高于其他方法，图 13 表明随着返回图像数目的增加，precision 值逐渐减小，且本文方法的 precision 值始终高于其他方法。图 11、图 12 和图 13 表明，本文方法的检索性能在一定程度上有所提升。

© 中国图象图形学报版权所有

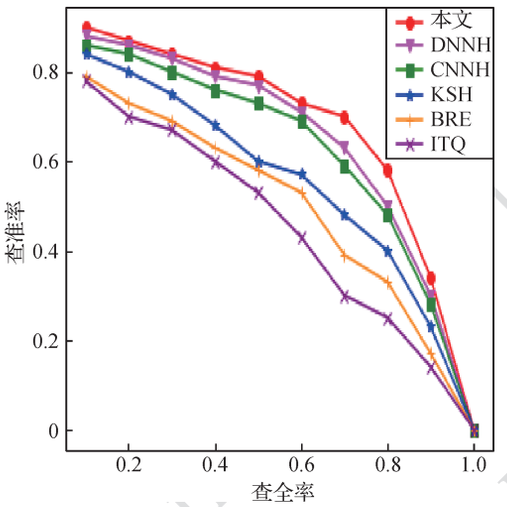



图 11　哈希码为 64 bits 时不同哈希方法的 P-R 曲线

Fig. 11　P-R curves of different Hashing methods at 64 bits

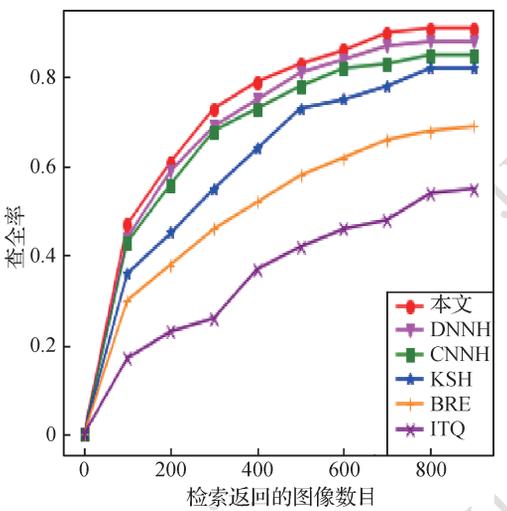



图 12　哈希码为 64 bits 时返回不同图像数的查全率

Fig. 12　Recall value of returning different images at 64 bits

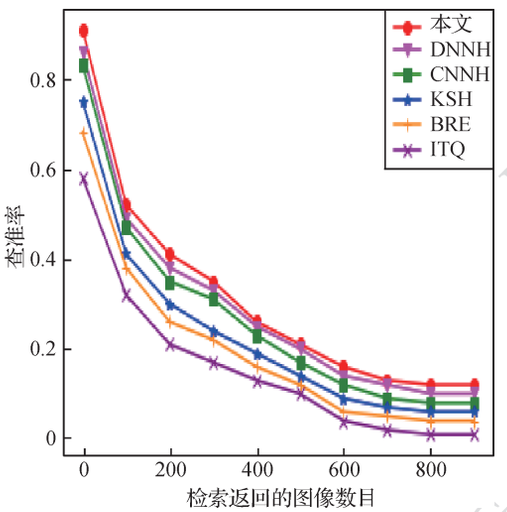



图 13　哈希码长为 64 bits 时返回不同图像数的查准率

Fig. 13　Precision value of returning different images at 64 bits

图 14 显示了在 LUNA16 数据集上，哈希码长为 64 bits 时本文方法的查询图像及其相应的检索返回的前 10 幅图像，其中首幅为查询图像，可以看出返回图像与查询图像均相关。由表 2 和图 11—图 14 及其结果分析可知，本文方法在 LUNA16 数据集上同样取得了较好的检索结果，验证了本文方法的有效性和可拓展性。

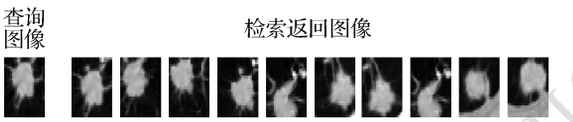



图 14　LUNA16 数据集上的查询及返回图像

Fig. 14　Query and return images on LUNA16 dataset

算法的时间复杂度是评价图像检索性能的重要指标，本文使用哈希函数构造的训练时间（train time）和测试集图像特征生成哈希码的测试时间（test time），来进一步验证本文方法的有效性。考虑到本文方法与 CNNH 方法均是非端到端的哈希方法，哈希函数构造的训练方式相同，所以将本文方法和检索效果较好的 CNNH 方法进行了时间复杂度的比较。表 3 和表 4 分别是在 Paris6K 和 LUNA16 数据集上，在不同哈希码长时本文方法与 CNNH 方法的训练时间和测试时间的比较结果。

**表 3　在 Paris6K 数据集上不同方法时间复杂度的对比**

**Table 3　Comparison of time complexity of different methods on Paris6K dataset**

| 哈希码长/bits | 训练时间/s | | 测试时间/s | |
|---|---|---|---|---|
| | 本文 | CNNH | 本文 | CNNH |
| 12 | 53.46 | 56.97 | 59.84 | 69.97 |
| 24 | 67.03 | 72.64 | 65.67 | 77.21 |
| 32 | 78.45 | 84.75 | 73.96 | 87.26 |
| 48 | 90.47 | 97.01 | 87.73 | 101.36 |
| 64 | 108.36 | 118.94 | 96.55 | 110.82 |
| 128 | 136.54 | 163.47 | 108.02 | 121.73 |

表 3 和表 4 的结果显示，在 Paris6K 和 LUNA16 数据集上，本文方法构造哈希函数的训练时间少于 CNNH 方法，这是因为本文方法在有监督学习哈希函数前，先对图像特征进行了 KPCA 投影，充分挖掘了图像特征中蕴含的非线性信息，使得哈希函数的学习寻优过程更简化，生成哈希码的时间减少，从而

© 中国图象图形学报版权所有

**表 4　在 LUNA16 数据集上不同方法时间复杂度的对比**
**Table 4　Comparison of time complexity of different methods on LUNA16 dataset**

| 哈希码长/ bits | 训练时间/s | | 测试时间/s | |
|---|---|---|---|---|
| | 本文 | CNNH | 本文 | CNNH |
| 12 | 52.38 | 54.26 | 56.02 | 67.53 |
| 24 | 65.23 | 70.21 | 62.18 | 75.62 |
| 32 | 73.51 | 80.57 | 69.93 | 84.73 |
| 48 | 86.93 | 92.06 | 82.21 | 98.73 |
| 64 | 101.23 | 113.46 | 92.47 | 107.16 |
| 128 | 134.72 | 155.21 | 103.19 | 120.47 |

在一定程度上验证了本文方法的高效性。

## 5　结　论

提出一种基于 DenseNet 和改进的监督核哈希的图像检索方法。通过充分提取图像的语义特征，以提升特征图像的表征能力，并通过改进的监督核哈希方法减低哈希编码时的投影误差，提高哈希码质量，从而提升检索性能。首先，对数据进行预处理，训练并优化 DenseNet 模型，用训练好的模型提取高区分力和高表达力的图像特征，然后通过改进的监督核哈希方法对图像特征进行哈希编码，进而比较哈希码间的汉明距离进行相似性度量，并返回相似图像集。在 Paris6K 图像检索数据集上验证本文方法的有效性，实验表明与 DNNH、CNNH、KSH、BRE、ITQ、SH 方法相比，本文方法取得了较好的检索精度；在 LUNA16 数据集上的实验同样取得了较好的检索结果；另外，将本文方法与 CNNH 方法进行算法时间复杂度的对比，结果表明本文方法具有一定的高效性。综上所述，本文方法具有较高的检索精度和较低的时间复杂度，检索性能得到提升，在彩色图像和医学图像检索中都有一定的价值与意义。在今后的研究工作中，将考虑在 DenseNet 模型的最后一层池化层后加入哈希层，同时训练模型和哈希函数，以期待获得学习能力更优的哈希函数，进一步提高检索精度。

## 参考文献(References)

Armato III S G, McLennan G, Bidaut L, McNitt-Gray M F, Meyer C R, Reeves A P, Zhao B S, Aberle D R, Henschke C I, Hoffman E A, Kazerooni E A, MacMahon H, van Beek E J. R, Yankelevitz D, Biancardi A M, Bland P H, Brown M S, Engelmann R M, Laderach G E, Max D, Pais R C, Qing D P Y, Roberts R Y, Smith A R, Starkey A, Batra P, Caligiuri P, Farooqi A, Gladish G W, Jude C M, Munden R F, Petkovska I, Quint L E, Schwartz L H, Sundaram B, Dodd L E, Fenimore C, Gur D, Petrick N, Freymann J, Kirby J, Hughes B, Vande C A, Gupte S, Sallam M, Heath M D, Kuhn M H, Dharaiya E, Burns R, Fryd D S, Salganicoff M, Anand V, Shreter U, Vastagh S, Croft B Y and Clarke L P. 2011. The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. Medical Physics, 38(2): 915-931 [DOI: 10.1118/1.3528204]

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 Computer Society: 886-893 [DOI: 10.1109/CVPR.2005.177]

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]

Hotelling H. 1993. Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology, 24(6): 417-441 [DOI: 10.1037/h0071325]

Har-Peled S, Indyk P and Motwani R. 2012. Approximate nearest neighbor: towards removing the curse of dimensionality. Theory of Computing, 8(1): 321-350 [DOI: 10.4086/toc.2012.v008a014]

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, NV, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]

Huang G, Liu Z, van der Maaten L and Weinberger K Q. 2016. Densely connected convolutional networks [EB/OL]. [2019-05-31]. https://arxiv.org/pdf/1608.06993v5.pdf

Ioffe S and Szegedy C. 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift//Proceedings of the 32nd International Conference on Machine Learning. Lille, France: ACM: 448-456

Kang X D. 2009. Image Informatics. Beijing: People's Medical Publishing House: 191 (康晓东. 2009. 影像信息学. 北京：人民卫生出版社：191)

Krizhevsky A, Sutskever I and Hinton G E. 2012. ImageNet classification with deep convolutional neural networks//Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, Nevada: Curran Associates Inc: 1097-1105 [DOI: 10.1145/3065386]

Kulis B and Darrell T. 2009. Learning to Hash with binary reconstructive embeddings//Proceedings of the 22nd International Conference on Neural Information Processing Systems. Vancouver, British Colum-

© 中国图象图形学报版权所有

bia, Canada: Curran Associates Inc: 1042-1050[DOI: http://espace.library.uq.edu.aulview/UQ: 192948]

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, USA: IEEE: 3270-3278 [DOI: 10.1109/CVPR.2015.7298947]

LeCun Y, Bengio Y and Hinton G. 2015. Deep learning. Nature, 521(7553): 436-444 [DOI: 10.1038/nature14539]

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]

Lin K, Yang H F, Hsiao J 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. Boston, MA, USA: IEEE: 27-35 [DOI: 10.1109/CVPRW.2015.7301269]

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, RI, USA: IEEE Press: 2074-2081 [DOI: 10.1109/CVPR.2012.6247912]

Lowe D G. 2004. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2): 91-110 [DOI: 10.1023/B:VISI.0000029664.99615.94]

Luan T T, Zhu J H, Xu S Y, Wang J X, Shi X and Li Y C. 2019. Hashing method for image retrieval based on product quantization with Huffman coding. Journal of Image and Graphics, 24(3): 389-399 (栾婷婷, 祝继华, 徐思雨, 王佳星, 时璇, 李垚辰. 2019. 哈夫曼编码乘积量化的图像哈希检索方法. 中国图象图形学报, 24(3): 389-399) [DOI: 10.11834/jig.180264]

Philbin J, Chum O, Isard M, Sivic J and Zisserman A. 2008. Lost in quantization: improving particular object retrieval in large scale image databases//Proceedings of 2008 IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, AK, USA: IEEE: 1-8 [DOI: 10.1109/CVPR.2008.4587635]

Raginsky M and Lazebnik S. 2009. Locality-sensitive binary codes from shift-invariant kernels//Proceedings of the 22nd International Conference on Neural Information Processing Systems. Vancouver, British Columbia, Canada: ACM: 1509-1517

Razavian A, Azizpour H, Sullivan J and Carlsson S. 2014. CNN features off-the-shelf: an astounding baseline for recognition//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Columbus, OH, USA: IEEE: 806-813 [DOI: 10.1109/CVPRW.2014.131]

Salakhutdinov R and Hinton G. 2009. Semantic Hashing. International Journal of Approximate Reasoning, 50(7): 969-978 [DOI: 10.1016/j.ijar.2008.11.006]

Shen F M, Zhou X, Yang Y, Song J K, Shen H T and Tao D C. 2016. A fast optimization method for general binary code learning. IEEE Transactions on Image Processing, 25(12): 5610-5621 [DOI: 10.1109/TIP.2016.2612883]

Simonyan K and Zisserman A. 2014. Very deep convolutional networks for large-scale image recognition[EB/OL]. [2019-05-31]. https://arxiv.org/pdf/1409-1556.pdf

Tielemant and Hinton G. 2012. Rmsprop: Divide the Gradient by A Running Average of Its Recent Magnitude, 4(2), 26-31

Vadlamudi L N, Vaddella R P V and Devara V. 2017. Robust image Hashing technique for content authentication based on DWT//Raman B, Kumar S, Roy P P and Sen D, eds. Proceedings of International Conference on Computer Vision and Image Processing. Singapore: Springer: 181-191 [DOI: 10.1007/978-981-10-2104-6_17]

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, CA, USA: IEEE Press: 3424-3431 [DOI: 10.1109/CVPR.2010.5539994]

Weiss Y, Torralba A and Fergus R. 2008. Spectral Hashing//Proceedings of the 22nd Annual Conference on Neural Information Processing Systems. Vancouver: Curran Associates Inc: 1753-1760

Xia R K, Pan Y, Lai H J, Liu C and Yan S C. 2014. Supervised Hashing for image retrieval via image representation learning//Proceedings of the 28th AAAI Conference on Artificial Intelligence. Quebec City, Canada: AAAI: 2156-2162

Yang H F, Lin K and Chen C S. 2018. Supervised learning of semantics-preserving Hash via deep convolutional neural networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(2): 437-451

## 作者简介

王亚鸽,1992 年生,女,硕士研究生,主要研究方向为医学图像处理。
E-mail: wangyage09@163.com

康晓东,通信作者,男,教授,主要研究方向为医学图像处理、医疗信息系统集成。
E-mail: kxd2004@126.com

郭军,男,实验师,主要研究方向为实验技术。
E-mail: 1319466165@qq.com
李博,男,硕士研究生,主要研究方向为医学图像处理。
E-mail: szxfsk123@163.com
张华丽,女,硕士研究生,主要研究方向为医学图像处理。
E-mail: aszhanghuali@163.com
刘汉卿,男,硕士研究生,主要研究方向为医学图像处理。
E-mail: 423065302@qq.com

© 中国图象图形学报版权所有