



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院空天信息创新研究院
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
12
VOL.25

ISSN1006-8961
CN11-3758/TB



数字浮雕建模 P2647



第25卷第12期 (总第296期)
2020年12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 中国科学院空天信息创新研究院

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 吴一戎

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 Aerospace Information Research Institute, CAS

China Society of Image and Graphics

Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Wu Yirong

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

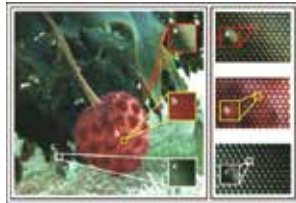
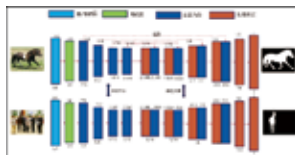
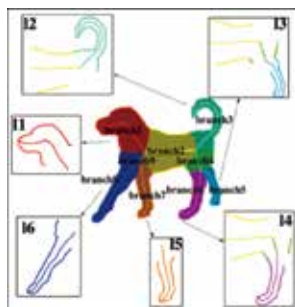
ISSN 1006-8961

CODEN ZTTXFZ

国外发行代号 M1406

国内邮发代号 82-831

国内定价 60.00元

光场显著性检测研究综述
(第2465页)融合注意力机制与知识蒸馏
的孪生网络压缩(第2563页)面向可解释性的物体拓扑结
构骨架表征方法(第2587页)

综述

光场显著性检测研究综述

刘亚美, 张骏, 张旭东, 孙锐, 高隼 2465

图像处理和编码

自适应卷积的残差修正单幅图像去雨

王美华, 何海君, 李超 2484

局部特定空间关系统计特征的RVIN噪声检测器

于海雯, 易昕炜, 徐少平, 林珍玉, 刘蕊蕊 2494

全局与局部属性一致的图像修复模型

孙劲光, 杨忠伟, 黄胜 2505

图像分析和识别

关键语义区域链提取的视频人体行为识别

马淼, 李贻斌, 武宪青, 高金凤, 潘海鹏 2517

针对形变与遮挡问题的行人再识别

史维东, 张云洲, 刘双伟, 朱尚栋, 暴吉宁 2530

多特征融合的行为识别模型

谭等泰, 李世超, 常文文, 李登楼 2541

结构特征下的可撤销人脸识别

孙浩浩, 邵珠宏, 尚媛园, 陈滨, 赵晓旭 2553

融合注意力机制与知识蒸馏的孪生网络压缩

耿增民, 余梦巧, 刘峡壁, 吕超 2563

联合聚焦度和传播机制的光场图像显著性检测

李爽, 邓慧萍, 朱磊, 张龙 2578

图像理解和计算机视觉

面向可解释性的物体拓扑结构骨架表征方法

危辉, 余莉萍 2587

自监督深度残差函数映射网络的3维模型对应关系计算

杨军, 马中昌 2603

融合自编码器和one-class SVM的异常事件检测

胡海洋, 张力, 李忠金 2614

抗高光的光场深度估计方法

王程, 张骏, 高隼 2630

计算机图形学

不同类型高度图控制浅浮雕模型生成方法

尚菁, 余文杰, 王美丽 2647

遥感图像处理

残差密集空间金字塔网络的城市遥感图像分割

韩彬彬, 张月婷, 潘宗序, 台宪青, 李芳芳 2656

结合判别相关分析与特征融合的遥感图像检索

葛芸, 马琳, 储珺 2665

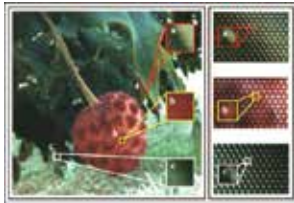
改进卷积网络的高分遥感图像城镇建成区提取

侯博文, 闫冬梅, 郝伟, 黄青青, 苏秀琴, 李青雯 2677

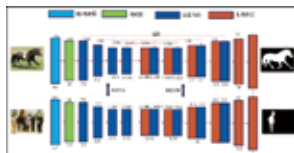
总目次 I

CONTENTS

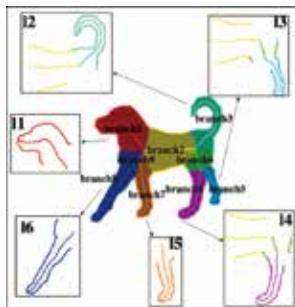
JOURNAL OF IMAGE AND GRAPHICS



Review of saliency detection on light fields(P2465)



Combining attention mechanism and knowledge distillation for Siamese network compression (P2563)



Skeleton characterization of object topology toward explainability (P2587)

Review

Review of saliency detection on light fields

Liu Yamei, Zhang Jun, Zhang Xudong, Sun Rui, Gao Jun 2465

Image Processing and Coding

Single image rain removal based on selective kernel convolution using a residual refine factor

Wang Meihua, He Haijun, Li Chao 2484

Random-valued impulse noise detector using local spatial structure statistics

Yu Haiwen, Yi Xinwei, Xu Shaoping, Lin Zhenyu, Liu Ruirui 2494

Image inpainting model with consistent global and local attributes

Sun Jinguang, Yang Zhongwei, Huang Sheng 2505

Image Analysis and Recognition

Human action recognition in videos utilizing key semantic region extraction and concatenation

Ma Miao, Li Yibin, Wu Xianqing, Gao Jinfeng, Pan Haipeng 2517

Person re-identification based on deformation and occlusion mechanisms

Shi Weidong, Zhang Yunzhou, Liu Shuangwei, Zhu Shangdong, Bao Jining 2530

Multi-feature fusion behavior recognition model

Tan Dengtai, Li Shichao, Chang Wenwen, Li Denglou 2541

Cancelable face recognition with fusion of structural features

Sun Haohao, Shao Zhuhong, Shang Yuanyuan, Chen Bin, Zhao Xiaoxu 2553

Combining attention mechanism and knowledge distillation for Siamese network compression

Geng Zengmin, Yu Mengqiao, Liu Xiabi, Lyu Chao 2563

Saliency detection on a light field via the focusness and propagation mechanism

Li Shuang, Deng Huiping, Zhu Lei, Zhang Long 2578

Image Understanding and Computer Vision

Skeleton characterization of object topology toward explainability

Wei Hui, Yu Liping 2587

Correspondence Calculation of 3D Models by a Self-Supervised Deep Residual Functional Maps Network

Yang Jun, Ma Zhongchang 2603

Anomaly detection with autoencoder and one-class SVM

Hu Haiyang, Zhang Li, Li Zhongjin 2614

Anti-specular light-field depth estimation algorithm

Wang Cheng, Zhang Jun, Gao Jun 2630

Computer Graphics

Bas-relief generation controlled by different height maps

Shang Jing, Yu Wenjie, Wang Meili 2647

Remote Sensing Image Processing

Residual dense spatial pyramid network for urbanremote sensing image segmentation

Han Binbin, Zhang Yueting, Pan Zongxu, Tai Xianqing, Li Fangfang 2656

Remote sensing image retrieval combining discriminant correlation analysis and feature fusion

Ge Yun, Ma Lin, Chu Jun 2665

Urban built-up area extraction using high-resolution remote sensing images with an improved convolutional neural network

Hou Bowen, Yan Dongmei, Hao Wei, Huang Qingqing, Su Xiuqin, Li Qingwen 2677

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2020)12-2665-12

论文引用格式: Ge Y, Ma L and Chu J. 2020. Remote sensing image retrieval combining discriminant correlation analysis and feature fusion. Journal of Image and Graphics, 25(12): 2665-2676 (葛芸, 马琳, 储珺. 2020. 结合判别相关分析与特征融合的遥感图像检索. 中国图象图形学报, 25(12): 2665-2676) [DOI: 10.11834/jig.200009]

结合判别相关分析与特征融合的遥感图像检索

葛芸, 马琳, 储珺

南昌航空大学软件学院, 南昌 330063

摘要: **目的** 高分辨率遥感图像检索中, 单一特征难以准确描述遥感图像的复杂信息。为了充分利用不同卷积神经网络(convolutional neural networks, CNN)的学习参数来提高遥感图像的特征表达, 提出一种基于判别相关分析的方法融合不同 CNN 的高层特征。**方法** 将高层特征作为特殊的卷积层特征处理, 为了更好地保留图像的原始空间信息, 在图像的原始输入尺寸下提取不同高层特征, 再对高层特征进行最大池化来获得显著特征; 计算高层特征的特征间散布矩阵, 结合判别相关分析来增强同类特征的联系, 并突出不同类特征之间的差异, 从而提高特征的判别力; 选择串联与相加两种方法来对不同特征进行融合, 用所得融合特征来检索高分辨率遥感图像。**结果** 在 UC-Merced、RSSCN7 和 WHU-RS19 数据集上的实验表明, 与单一高层特征相比, 绝大多数融合特征的检索准确率和检索时间都得到有效改进。其中, 在 3 个数据集上的平均精确率均值(mean average precision, mAP)分别提高了 10.4%~14.1%、5.7%~9.9% 和 5.9%~17.6%。以检索能力接近的特征进行融合时, 性能提升更明显。在 UC-Merced 数据集上, 融合特征的平均归一化修改检索等级(average normalized modified retrieval rank, ANMRR)和 mAP 达到 13.21% 和 84.06%, 与几种较新的遥感图像检索方法相比有一定优势。**结论** 本文提出的基于判别相关分析的特征融合方法有效结合了不同 CNN 高层特征的显著信息, 在降低特征冗余性的同时, 提升了特征的表达能力, 从而提高了遥感图像的检索性能。

关键词: 遥感图像检索; 卷积神经网络; 高层特征融合; 判别相关分析; 最大池化

Remote sensing image retrieval combining discriminant correlation analysis and feature fusion

Ge Yun, Ma Lin, Chu Jun

School of Software, Nanchang Hangkong University, Nanchang 330063, China

Abstract: **Objective** With the rapid development of remote sensing technology, numerous high-resolution remote sensing images have become available. As a result, the effective retrieval of remote sensing images has become a challenging research topic. Feature extraction is key to determining the retrieval performance of high-resolution remote sensing image retrieval tasks. Traditional feature extraction methods are mainly based on handcrafted features, whereas such shallow features are easily affected by artificial intervention. Convolutional neural networks (CNNs) can learn feature representations automatically, and thus are suitable to deal with high-resolution remote sensing images with complex content. However, the parameters of CNNs are difficult to train fully due to the small scale of currently available public remote sensing datasets. In

收稿日期: 2020-01-08; 修回日期: 2020-02-19; 预印本日期: 2020-02-26

基金项目: 国家自然科学基金项目(41801288, 61663031, 41261091, 61762067, 61866028); 江西省自然科学基金项目(2020BAB212011); 南昌航空大学博士科研启动金项目(EA201920276)

Supported by: National Natural Science Foundation of China(41801288, 61663031, 41261091, 61762067, 61866028)

this case, the transfer learning of CNNs has attracted much attention. CNNs pretrained on large-scale datasets have good generalization ability, and parameters can be transferred to small-scale data effectively. Therefore, extracting CNN features on the basis of transfer learning has become an effective method in the field of remote sensing image retrieval. Given the abundant and complex visual content of high-resolution remote sensing images, it is difficult to accurately express the content of remote sensing images using a single feature. Thus, feature fusion is a useful method to improve the feature representation of remote sensing images. To maximize the learning parameters of different CNNs to represent the content of remote sensing images, a method based on discriminant correlation analysis (DCA) is proposed to fuse the high-level features of different CNNs. **Method** First, CNN parameters from VGGM (visual geometry group medium), VGG (visual geometry group) 16, GoogLeNet, and ResNet50 are transferred for high-resolution remote sensing images, and the high-level features are adopted as special convolutional features. To preserve the original spatial information of the image, the high-level features are extracted under the original input image size, and the output form of three-dimensional tensor is retained. Then, max pooling is adopted on the high-level features to extract salient features. Second, DCA is adopted to enhance the feature representation. The DCA is the first to incorporate the class structure into the feature level fusion and has low computational complexity. To maximize the correlation of corresponding features across the two feature sets and in the same time decorrelates features that belong to different classes within each feature set, the between-class scatter matrices of the two sets of high-level features are calculated, and matrix diagonalization and singular value decomposition are adopted to transform the features. The transformed matrix contains the important eigenvectors of the between-class scatter matrix, and the dimension of the transformed matrix is reduced accordingly. Thus, the transformed feature vectors have strong discriminative power and low dimension. Lastly, two methods of concatenation and summation are selected to perform the fusion of transformed feature vectors, and the fused features are normalized via Gaussian normalization. The similarities between the query and dataset features are calculated using the Euclidean distance method, and the retrieval results are returned in accordance with the sort of similarities. **Result** Experiment results on the UC-Merced, RSSCN7, and WHU-RS19 datasets show that the retrieval accuracy and retrieval time of most fusion features are effectively improved in comparison with a single high-level feature; the mean average precision (mAP) of the fusion feature is improved by 10.4% ~ 14.1%, 5.7% ~ 9.9%, and 5.9% ~ 17.6%, respectively. The retrieval results of the fused features using the concatenation method are better than that using the summation method. Multifeature fusion experiments show that the best result on the UC-Merced dataset is obtained from the fusion of four features, whereas the best results on the RSSCN7 and WHU-RS19 datasets are obtained from the fusion of three features. This finding indicates that a larger number of fused features does not translate into better performance; selecting the appropriate features is crucial for feature fusion. Especially, when the different features have good representation and similar retrieval capabilities, the fusion of these features can achieve good retrieval performance. Compared with other state-of-the-art approaches, the average normalized modified retrieval rank (ANMRR) and mAP of the proposed fused feature on the UC-Merced dataset reach 0.132 1 and 84.06%, respectively. Experimental results demonstrate that our method outperforms state-of-the-art approaches. **Conclusion** The feature fusion method based on discriminant correlation analysis combines the salient information of different high-level features. This method reduces feature redundancy while improving feature discrimination. Features with equivalent retrieval capabilities can be fused by the proposed method well, thus effectively improving the retrieval performance of high-resolution remote sensing images.

Key words: remote sensing image retrieval; convolutional neural network (CNN); high-level feature fusion; discriminant correlation analysis (DCA); max pooling

0 引 言

随着遥感技术的不断发展,可以获取的遥感图像数量和种类逐渐增多,为了更有效地管理和利用数量庞大、类别丰富的高分辨率遥感图像,基于内容

的遥感图像检索(content-based remote sensing image retrieval, CBRSIR)成为研究热点。

CBRSIR 主要包括特征提取和相似性度量。提取合理有效的特征来准确表达遥感图像的内容是提高图像检索性能的关键。传统的特征提取主要采用手工方式提取纹理、形状和颜色等特征。手工特征

的提取方式简单,但是主观性较强,且难以表达图像的高层语义信息。这种特征的信息表达与高层语义理解之间的不一致导致了比较严重的语义鸿沟问题。

目前,深度学习中的卷积神经网络(convolutional neural network, CNN)发展迅速,在语音识别、动作识别和图像识别等领域应用广泛。CNN采用了多层神经网络的训练机制来模拟人脑神经的认知和表达,自动学习不同层次的特征。图像检索中引入CNN的相关技术,能够有效缩小语义鸿沟,提高图像的检索性能(Husain 和 Bober, 2019; Radenović 等, 2019)。遥感图像检索中,由于现有公开的高分辨率遥感数据集规模较小,难以充分训练CNN,而基于大规模数据集训练的CNN泛化能力强,能够有效迁移到其他小规模数据集中,因此,在高分辨率遥感图像检索中引入CNN的迁移学习是一种有效方法。

基于CNN迁移学习的遥感图像检索主要是将大规模数据集上预训练CNN的参数通过直接或者微调的方式迁移到高分辨率遥感图像,然后提取高层特征或者中层特征用于检索。高层特征表达了图像的全局信息,迁移到遥感图像中的效果好,但是为了提取图像的高层特征,输入图像的尺寸必须调整为CNN需要的默认输入尺寸(Zhou 等, 2017; Wang 等, 2017)。中层特征通常是卷积层特征,表达了图像的局部信息,该特征的提取不受输入图像尺寸限制,但一般要经过聚合之后才能更好地表达图像的内容(Hu 等, 2017; Tong 等, 2017)。为了改进CNN迁移学习在遥感图像检索中的性能, Ye 等人(2018)和彭晏飞等人(2019)引入了反馈机制以优化检索结果。Zhou 等人(2018)基于分割图提取图像的区域卷积特征,采用多标签的方法优化检索性能。Imbriaco 等人(2019)利用乘法和加法注意力机制选择具有代表性的CNN特征,提高检索精度。

虽然基于CNN迁移学习提取的特征能够有效检索遥感图像,但是遥感图像内容复杂,有些图像类间差异小,类内差异却很大,仅靠单一特征难以准确表达遥感图像的内容,因此将多种特征融合可以进一步提高特征表达的准确性。传统的特征融合主要是融合不同的手工特征,陆丽珍等人(2004)提出了基于权重分配融合颜色和纹理特征的遥感图像检索方法,检索性能明显优于单一特征。Klaric 等人

(2012)对形状特征和光谱特征进行多特征索引,得到了更理想的检索结果。Wang 等人(2016)提出3层图的方法来融合局部特征与全局特征,有效提高了遥感图像的检索性能。

除了手工特征的融合外,对CNN特征进行融合是目前发展的重要趋势。目前,针对遥感图像检索中CNN特征融合的研究较少,大部分研究集中在遥感图像分类。遥感图像中的CNN特征融合主要分为两类,一类是同一个网络中不同层次特征的融合。杨珂等人(2019)选择同一CNN的中层特征与高层特征进行融合,有效结合了中层特征细粒度识别的优势与高层特征语义表达的优势。Li 等人(2019a)提出自适应方法来融合同一CNN的高层特征与中层特征,强调有用的信息并抑制无用的信息,提取的特征具有更好的判别性。另一类是不同网络中的特征融合。Li 等人(2019b)将ResNet和VGG(visual geometry group)16的高层特征直接串联在一起,该融合方法充分利用从不同网络中学到的参数来更全面地表达遥感图像,从而提高了遥感图像的分类结果。另外,不同CNN特征融合的对比研究中,不管是基于权重分配的融合方法(Ge 等, 2018),还是基于特征空间学习的融合方法(Du 等, 2019),研究结果都表明,不同网络的CNN特征进行融合优于同一个CNN中不同层次特征的融合。因此,CNN特征融合是提高遥感图像特征表达的重要手段,其中以不同网络中的CNN特征融合效果更好。

为了进一步优化特征融合的效果,选择具有代表性的特征进行融合是直接有效的方法。基于典型相关分析(canonical correlation analysis, CCA)的特征融合方法(Sun 等, 2005),通过分析两组特征的相关性得到具有代表性的融合特征。基于CCA的融合方法在图像检索中有很好的应用。刘润杰等人(2015)针对内容复杂的图像,利用CCA将图像纹理特征与颜色特征融合,削弱了复杂图像的不利影响。Li 等人(2017)基于CCA的思想来学习图像视觉特征与标签信息之间的相关性,获得了更好的检索性能。然而,CCA仅关注两组变量之间的相关性,忽视了类别信息。除此以外,对于小样本数据集来说,特征维数过高也为融合带来不便,因此Haghighat 等人(2016)在CCA的基础上提出了判别相关分析(discriminant correlation analysis, DCA)的方法。DCA在增强同类特征联系的同时,突出不同类特征

之间的差异,在多模态生物识别领域取得了较好的效果,并且在遥感图像分类领域取得了成效。Farah 等人(2018)引用 DCA 来融合高光谱图像的光谱特征和空间特征,增加类间方差并减少类内方差,有效提高了遥感图像的分类精度。Chaib 等人(2017)提取同一个 CNN 的两个全连接层特征,然后通过 DCA 来融合这两个特征,减少了融合特征的冗余信息,并显著提高了分类精度。因此,DCA 能够有效融合遥感图像的不同特征。受此启发,本文结合 CNN 迁移学习能力强的特点,提出基于 DCA 特征融合的遥感图像检索方法。为了综合不同 CNN 的学习参数来更全面地表达高分辨率遥感图像的内容,本文方法选择不同 CNN 的高层特征进行融合。为了保留特征原始的空间信息,将高层特征作为卷积层特征处理,在图像的原始输入尺寸下提取高层特征,并对高层特征进行最大池化操作以提取图像的显著特征。最后进行 DCA 融合,根据两组特征之间的相关性对特征进行选择,在降低特征维数的同时,保留判别能力强的特征。实验结果表明,融合后的特征对遥感图像的表达式更准确,在多个高分辨率遥感数据集上的检索能力都有所提升。

1 本文方法

本文提出一种基于 DCA 特征融合的方法,用于遥感图像检索。检索流程图如图 1 所示:1)使用迁移学习的方法提取图像原始尺寸下不同 CNN 的高层特征,并对高层特征进行最大池化以提取显著信息;2)选择不同高层特征进行 DCA 融合,既保留了两组特征集对应特征之间的相关性,又能有效区分分类间差异;3)对融合特征进行归一化,以方便计算图像的相似度;4)对图像相似度排序,返回相似图像结果。

1.1 特征提取

高分辨率遥感图像内容复杂多样,CNN 的中层特征主要表达图像的局部信息,迁移到高分辨率遥感图像中的检索性能不如高层特征(Zhou 等, 2017)。为了更全面地表达高分辨率遥感图像的内容,本文使用不同的网络模型来提取遥感图像的不同高层特征,并采取 DCA 的方法对不同的高层特征进行融合。

传统提取高层特征的方法需要调整图像的输入

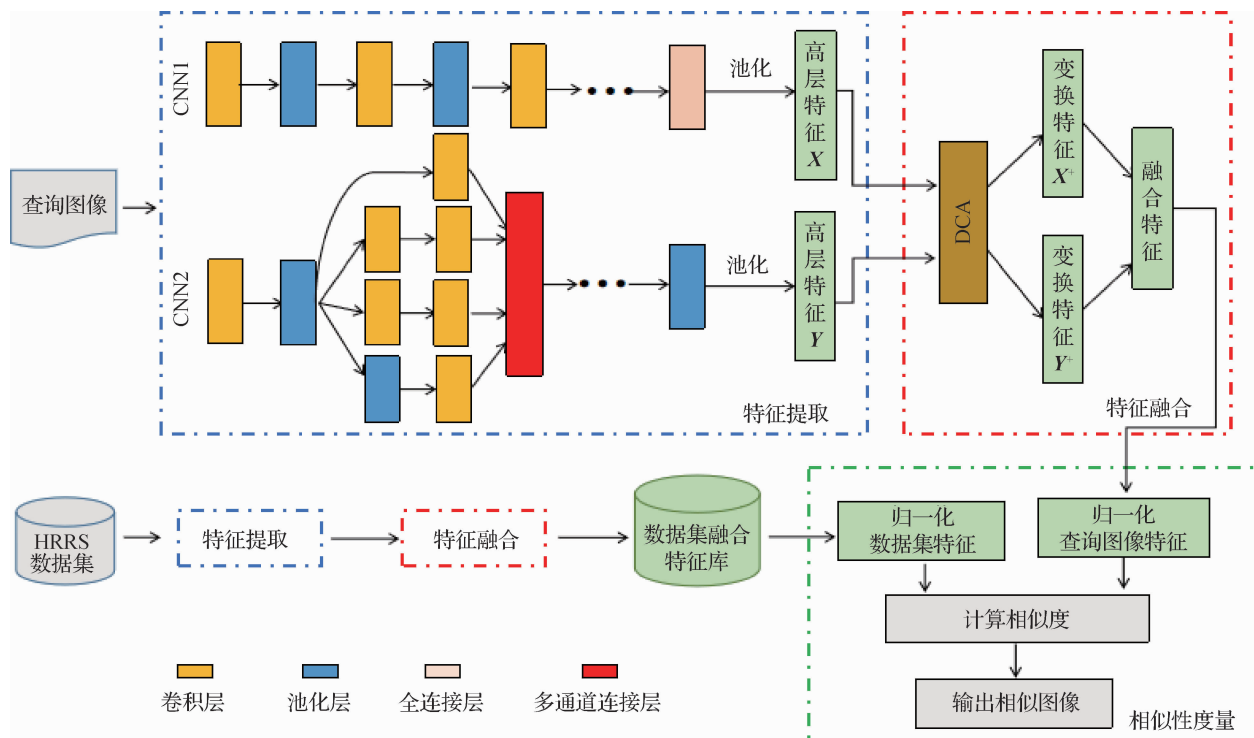


图 1 基于 DCA 特征融合的图片检索流程图

Fig. 1 Image retrieval flowchart based on DCA feature fusion

尺寸为 CNN 的默认尺寸, 再将高层特征转换为特征向量, 容易丢失图像的原始空间信息。Shelhamer 等人(2017)指出, 全连接层实际上是一种特殊的卷积层, 将全连接层作为卷积层处理, 输入图像的尺寸可不受限制, 能够更好地保留原始输入图像的空间信息。受此启发, 本文将高层作为特殊的卷积层处理, 在输入图像的原始尺寸下提取高层特征, 提取特征计算为

$$y_j^l = r(\sum_{i \in M_j} y_i^{l-1} \times w_{ij}^l + b_j^l) \quad (1)$$

式中, r 为激活函数, y_j^l 表示 l 层第 j 个特征图, M_j 表示 $l-1$ 层特征图的集合, w_{ij}^l 表示上一层对第 l 层的权重, b_j^l 表示特征图 y_j^l 的偏置。为了更好地反映原始图像的空间信息, 输出的高层特征保留原来的 3 维张量形式, 具体为

$$f = [y_1, y_2, \dots, y_j, \dots, y_p] \quad (2)$$

式中, y_j 为第 j 个特征图, p 为特征图的数目, f 为对应的高层特征。

上述方法提取的高层特征虽然能够更有效地反映原始图像的空间信息, 但也增加了特征的维数。为了进一步提取高层特征中的显著特征, 采用最大池化的方法来对高层特征进行压缩。假定高层特征 f 的尺寸为 $w \times h \times p$, 其中 w 和 h 分别代表高层特征图的宽度和高度。最大池化操作的计算式为

$$P_i^{\max} = \max f(\cdot, \cdot, i), i = 1, 2, \dots, p \quad (3)$$

经过最大池化后, 高层特征转换为一个 $1 \times 1 \times p$ 维特征向量。最大池化后该特征的维数与传统高层特征的维数相同, 且特征的鲁棒性更强。

1.2 特征融合

基于 DCA 的特征融合是通过分析两组变量的相关关系来增强同类特征之间的相关性, 并降低不同类别特征之间的冗余相关性。该方法在降低特征维数的同时提高了特征的判别能力, 适用于小规模数据集中特征维数较高的特征融合。因此本文选用 DCA 的方法来融合维数较高的高层特征。

假定数据集 D 的类别数为 c , X, Y 表示图像数据集 D 的两组高层特征矩阵。则

$$\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^n x_{ij} \quad (4)$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^c n_i \bar{x}_i \quad (5)$$

式中, $x_{ij} \in X$, 表示从第 i 类的第 j 个图像提取的特

征向量, n_i 表示第 i 类图像的数目, $n = \sum_{i=1}^c n_i$ 。 $\bar{x}_i$ 和 $\bar{x}$ 分别表示第 i 类下的 x_{ij} 向量的平均值和整个特征矩阵 X 的平均值, 即 $\bar{x}_i$ 是类内特征向量平均值, $\bar{x}$ 是类间特征向量平均值。

根据 $\bar{x}_i$ 和 $\bar{x}$ 可以计算出代表不同特征类间关系的散布矩阵 S_{bx} , 计算式为

$$\Phi_{bx} = [\sqrt{n_1}(\bar{x}_1 - \bar{x}), \sqrt{n_2}(\bar{x}_2 - \bar{x}), \dots, \sqrt{n_c}(\bar{x}_c - \bar{x})] \quad (6)$$

$$S_{bx} = \sum_{i=1}^c n_i (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T = \Phi_{bx} \Phi_{bx}^T \quad (7)$$

式中, 矩阵 Φ_{bx} 表示 X 中不同类别特征之间的差异度。当类别数 c 小于特征向量维数 p 时, $\Phi_{bx} \Phi_{bx}^T$ 的主要特征向量可以转化为计算 $\Phi_{bx}^T \Phi_{bx}$ 来获取, 因此问题就变成求 $c \times c$ 维的 $\Phi_{bx}^T \Phi_{bx}$ 。当特征矩阵中不同类别特征之间差异性较好时, 此时协方差矩阵 $\Phi_{bx}^T \Phi_{bx}$ 将会是一个对称的对角矩阵。因此, 为了不同类别之间可以很好地进行分离, 将其对角化, 具体为

$$Q^T (\Phi_{bx}^T \Phi_{bx}) Q = \Lambda \quad (8)$$

式中, Q 是 $\Phi_{bx}^T \Phi_{bx}$ 的特征向量矩阵。取 Q 中前 k 个特征向量表示为 $\varphi_{(c \times k)}$, 为了得到 k 阶最大非零特征, 有

$$(\Phi_{bx} \varphi)^T S_{bx} (\Phi_{bx} \varphi) = \Lambda_{(k \times k)} \quad (9)$$

对特征按照降序排序, 并取前 k 个最显著特征, 由于最多有 $c-1$ 个非零广义特征值, 因此 k 的上限为 $c-1$ 。定义 $W_{bx} = \Phi_{bx} \varphi \Lambda^{-1/2}$ 为转换矩阵来归一化 S_{bx} , $W_{bx}^T S_{bx} W_{bx} = I$ 。其中, I 是经过 W_{bx} 转换降维后的散布矩阵。利用转换矩阵 W_{bx} 将特征矩阵 X 的维数从 p 降到 k , 经过以上变换, 极大削弱了高层特征中不同类别之间的联系, 具体为

$$X'_{(k \times n)} = W_{bx}^T (k \times p) X_{(p \times n)} \quad (10)$$

同理, 可求得另一特征集变换矩阵 Y' 。为了增加 X 中的特征与 Y 中对应同类特征之间的相关性, 使用奇异值分解将变换后两个高层特征矩阵 X' 和 Y' 的类间协方差矩阵 $S'_{xy} = X' Y'^T$ 对角化, 具体为

$$S'_{xy} = U \Sigma V^T \quad (11)$$

$$U^T S'_{xy} V = \Sigma \quad (12)$$

式中, Σ 为奇异值矩阵, 仅在主对角线上有非零奇异值, 也是类间协方差特征矩阵 S'_{xy} 的特征值, 除对角线元素外, 其他元素均为零。 U 、 V 分别为左、右

奇异矩阵,则有 $(U\Sigma^{-1/2})^T S'_{xy} (V\Sigma^{-1/2}) = I$ 。同理,定义 $W_{cx} = U\Sigma^{-1/2}$, $W_{cy} = V\Sigma^{-1/2}$ 为转换矩阵,通过计算得到最终的变换特征组 X^* 和 Y^* ,具体为

$$X^* = W_{cx}^T X' = W_{cx}^T W_{bx}^T X = W_x X \quad (13)$$

$$Y^* = W_{cy}^T Y' = W_{cy}^T W_{by}^T Y = W_y Y \quad (14)$$

式中, $W_{cx}^T W_{bx}^T$ 与 $W_{cy}^T W_{by}^T$ 分别是两组高层特征 X 和 Y 的最终转换矩阵,即 W_x 和 W_y 。 X 和 Y 两组高层特征经过 DCA 降维,使得相同类别对应特征之间的相关性最大,不同类别特征之间的区分度较好,因此基于 DCA 的融合特征在表达高分辨遥感图像内容时具有较好的判别性。

串联(concatenation)和相加(summation)是特征融合中常用的两种方式。串联融合后所得特征的维数增加,相加融合后所得特征的维数不变。

串联融合的具体计算为

$$Z = \begin{bmatrix} X^* \\ Y^* \end{bmatrix} = \begin{bmatrix} W_x & 0 \\ 0 & W_y \end{bmatrix}^T \begin{bmatrix} X \\ Y \end{bmatrix} \quad (15)$$

相加融合的具体计算为

$$Z = X^* + Y^* = \begin{bmatrix} W_x \\ W_y \end{bmatrix}^T \begin{bmatrix} X \\ Y \end{bmatrix} \quad (16)$$

1.3 检索流程

本文利用基于 ImageNet 数据集预训练的不同 CNN 来提取高层特征,并采用 DCA 方法进行特征融合,检索流程如图 1 所示,具体的检索步骤为:

1) 参数迁移。将预训练的两个 CNN(CNN1 和 CNN2)的参数迁移到高分辨率遥感数据集和查询图像。

2) 提取高层特征。在原始输入尺寸下,对高分辨率遥感数据集和查询图像,都提取两个不同 CNN 的高层特征,并对输出的高层特征进行最大池化。

3) 对不同的高层特征进行特征融合。通过对两组特征进行判别相关分析,找到具有代表性的 k 维特征。接着,采用串联和相加两种方式对不同高层特征融合来提高特征的表达能力。

4) 对融合特征进行高斯归一化。不同 CNN 提取的高层特征在数值上存在差异,为了使不同的特征向量在相似性度量时具有相同的地位,需要对特征进行归一化处理。采取常用的高斯归一化,分别对数据集中的融合特征和查询图像的融合特征进行归一化,使得特征中的绝大部分的值在 $[0, 1]$ 范围内。归一化后让原本差距过大的数值具有可比性,

改进了特征的融合效果。

5) 计算相似度并返回相似图像。采取常用的欧氏距离计算查询图像与数据集中图像的相似度,并按照相似度的大小进行排序,按需返回最相似的若干幅图像作为检索结果。

1.4 检索效率分析

检索流程主要包括特征提取、特征融合以及相似性度量。特征提取首先从预训练的 CNN 中直接提取高层特征,然后对高层特征进行最大池化。最大池化算法的时间复杂度与高层特征图的宽度 w 和高度 h 有关,为 $O(w \times h)$ 。一般情况下,高层特征图的长度和宽度都不大,因此池化速度快。

特征融合阶段采用了 DCA。DCA 算法效率高,运行时间短(Haghighat 等,2016)。在 DCA 算法中,最费时的操作是计算高层特征的类间散布矩阵 S_{bx} 以及奇异值分解。算法的时间复杂度与数据集的训练样本数 n 、类别数 c 、特征的维数 p 相关。计算散布矩阵 S_{bx} 的时间复杂度为 $O(np^2 + c^2 p)$,在进行奇异值分解时,通过将 S_{bx} 转化,矩阵维度由 $p \times p$ 降低为 $c \times c$,所以奇异值分解的时间复杂度由 $O(p^3)$ 降低至 $O(c^3)$,DCA 后续操作在降维后的矩阵上进行,因此算法的效率较高。

相似性度量阶段需要计算查询图像与数据集中候选图像的相似度,并对相似度进行排序,以便返回最相似的图像。相似度计算的时间复杂度与特征维数 c 和候选图像数目 m 有关,为 $O(m \times c)$ 。排序的时间复杂度与返回的 m 个相似度有关,为 $O(m \log m)$ 。

从以上分析可以看出,由于高层特征图尺寸较小并且融合特征维数较低,因此 DCA 算法的总体效率比较高。

2 实验结果及分析

2.1 实验设置与评估标准

为了验证本文方法的有效性,采用 UC-Merced (Yang 和 Newsam, 2013) 数据集、WHU-RS19 (Hu 等,2015) 数据集和 RSSCN7 数据集(Zou 等,2015) 进行实验。UC-Merced 数据集中的图像来源于美国地质勘探局收集的国家城市地区地图,每幅图像大小为 256×256 像素,包含飞机、海滩、建筑物等 21 个类,每类 100 幅图像,示例图像如图 2(a) 所示。

RSSCN7 数据集包含 2 800 幅遥感图像, 每幅图像大小为 400×400 像素, 包含草地、森林、农田、停车场、住宅区、工业区和河湖共 7 个类别, 每个类别包含 400 幅图像, 示例图像如图 2(b) 所示。WHU-RS19 数据集中的图像来自谷歌地球, 每幅图像大小为 600×600 像素, 包含港口、火车站、高架等 19 个类, 每类 50 幅图像, 示例图像如图 2(c) 所示。



图2 示例图像

Fig. 2 Sample images ((a) UC-Merced;
(b) RSSCN7; (c) WHU-RS19)

实验中, 采用 MatConvNet (Vedaldi 和 Lenc, 2015) 框架提取 VGG16 (Simonyan 和 Zisserman, 2015)、VGGM (visual geometry group medium) (Chatfield 等, 2014)、ResNet50 (He 等, 2016) 和 GoogLeNet (Szegedy 等, 2015) 的高层特征, VGG16 和 VGGM 的高层特征为全连接层 ReLU 6 层的输出值, 分别简记为 V 和 M; ResNet50 和 GoogLeNet 的高层特征为最后的池化层对应的特征, 分别简记为 R 和 G。V 与 R 的融合特征简记为 V-R, 其他融合特征采用类似的标记。实验时, 将数据集中 80% 的图像作为训练集, 剩余 20% 作为测试集。

评价指标为平均准确率均值 (mean average precision, mAP) 和平均归一化修改检索等级 (average normalized modified retrieval rank, ANMRR)。mAP 数值越高, 检索性能越好。ANMRR 数值越低, 检索

的正确结果排序越靠前, 即检索能力越强。此外, 还比较了不同融合特征的查准率—查全率。

2.2 不同融合方法比较

特征融合的方法主要有串联和相加两种方式, 为了比较这两种融合方法, 表 1 是在原始输入尺寸下, 在 UC-Merced、RSSCN7 和 WHU-RS19 数据集上的实验结果, 串联简记为 con, 相加简记为 sum。为了与单一特征进行公平比较, 表 1 显示了在原始输入尺寸下提取的单一高层特征的检索结果。

实验还比较了不同特征的检索时间。由于 DCA 融合是在训练数据集中进行, 属于线下操作, 因此检索时间是指检索一幅图像的平均检索时间, 包括计算查询图像与候选图像的相似度以及对相似度进行排序的时间。从表 1 可以看出, 融合特征的检索时间低于单一特征的检索时间, 表明融合特征的检索效率较高。

从实验结果可以看出, 相加的融合方式不会增加维度, 但检索结果不如串联的融合特征。这是因为在 DCA 降维融合的过程中, 对两组高层特征已经进行去冗余操作, 保留的特征具有较好的判别力, 将这两组特征相加融合时, 两组特征之间相互影响, 导致融合效果不佳。

与单一特征相比, 经过 DCA 融合的特征不仅维数大大降低, 并且特征之间的相关性增强, 增加了有效信息。在 UC-Merced 数据集上的实验结果显示, 融合特征的 mAP 比单一特征提高了 12% 左右, AN-MRR 降低了 11% 左右, 其中 R-G-con 的 mAP 值最高, ANMRR 值最小, 检索结果最好。在对 RSSCN7 数据集检索时, 由于数据集的类别数少, 导致特征维数过低, 低维特征不足以全面描述复杂的遥感图像, 因此本文通过适当提高融合特征的维数来改进特征表达。RSSCN7 数据集中融合特征 M-G-con 的 mAP 值比单一特征 M 和 G 分别提高了 8.25% 和 9.95%。在 WHU-RS19 数据集上的实验结果显示, 融合特征 mAP 比单一特征提高了 11% 左右, AN-MRR 降低了 10% 左右。

由以上分析可知, DCA 特征融合能够将不同的高层特征有效地融合在一起, 并以串联方式进行特征融合的效果更好。因此, 接下来的实验选用串联的方式融合不同特征。

2.3 不同输入尺寸下的实验结果

将本文提出的在原始输入尺寸下提取的高层特

表 1 不同数据集上不同融合方式的检索结果

Table 1 Retrieval results of different fusion methods on different datasets

特征	UC-Merced				RSSCN7				WHU-RS19			
	mAP	ANMRR	维数	时间/s	mAP	ANMRR	维数	时间/s	mAP	ANMRR	维数	时间/s
V-M-con	0.823 1	0.151 1	40	0.005 2	0.738 8	0.248 5	24	0.000 3	0.905 0	0.076 3	36	0.001 4
V-G-con	0.822 6	0.146 9	40	0.005 3	0.733 4	0.247 6	24	0.000 3	0.899 0	0.082 0	36	0.001 4
V-R-con	0.827 0	0.143 2	40	0.005 3	0.612 1	0.565 7	24	0.000 2	0.938 6	0.051 5	36	0.001 3
M-R-con	0.838 8	0.134 1	40	0.005 2	0.642 3	0.556 8	24	0.000 2	0.940 0	0.049 5	36	0.001 4
M-G-con	0.815 0	0.155 0	40	0.005 2	0.764 7	0.213 2	24	0.000 2	0.892 5	0.086 6	36	0.001 3
R-G-con	0.840 6	0.132 1	40	0.005 4	0.625 0	0.542 2	24	0.000 2	0.937 3	0.051 4	36	0.001 3
V-M-sum	0.813 4	0.157 9	20	0.005 3	0.734 4	0.251 7	12	0.000 1	0.903 9	0.077 3	18	0.001 3
V-G-sum	0.824 4	0.145 9	20	0.005 2	0.736 0	0.245 1	12	0.000 1	0.905 2	0.077 2	18	0.001 3
V-R-sum	0.817 1	0.151 5	20	0.005 2	0.600 2	0.553 3	12	0.000 2	0.938 0	0.051 6	18	0.001 2
M-R-sum	0.825 0	0.147 6	20	0.005 3	0.627 9	0.558 9	12	0.000 1	0.934 3	0.054 4	18	0.001 2
M-G-sum	0.810 8	0.158 6	20	0.005 2	0.747 0	0.222 1	12	0.000 2	0.899 5	0.081 0	18	0.001 1
R-G-sum	0.838 4	0.134 7	20	0.005 2	0.623 6	0.537 9	12	0.000 1	0.934 8	0.052 3	18	0.001 2
V	0.711 0	0.250 7	4 096	0.078 3	0.664 3	0.250 3	4 096	0.028 6	0.840 1	0.136 0	4 096	0.035 9
M	0.700 0	0.258 2	4 096	0.078 7	0.682 2	0.237 1	4 096	0.028 6	0.799 0	0.174 4	4 096	0.037 1
G	0.699 2	0.256 6	1 024	0.067 7	0.665 2	0.243 6	1 024	0.007 7	0.761 6	0.203 4	1 024	0.007 3
R	0.723 0	0.238 6	2 048	0.071 9	0.314 7	0.571 2	2 048	0.014 9	0.871 9	0.108 6	2 048	0.017 3

注:加粗字体为每列最优结果。

征融合,与传统的在默认输入尺寸下提取高层特征融合相比较。不同输入尺寸下,融合特征的检索结果如图 3 所示。

图 3(a)是在 UC-Merced 数据集上的检索结果,V-M 在默认尺寸和原始尺寸下的结果非常接近,剩下的融合特征中,除了 M-G 外,其他融合特征在原始尺寸(256 × 256 像素)下的检索结果明显优于默认尺寸(224 × 224 像素)。原始输入尺寸下,检索

效果最好的是 R-G,其 mAP 值达到 0.840 6。

图 3(b)和图 3(c)分别是在 RSSCN7 和 WHU-RS19 数据集上的检索结果。可以看出,输入图像在原始尺寸下的检索性能优于传统默认尺寸。RSSCN7 数据集与 WHU-RS19 数据集中图像的原始尺寸分别为 400 × 400 像素和 600 × 600 像素,与默认尺寸 224 × 224 像素相差较大,同 UC-Merced 数据集相比,这两个数据集在原始尺寸下的检索结果

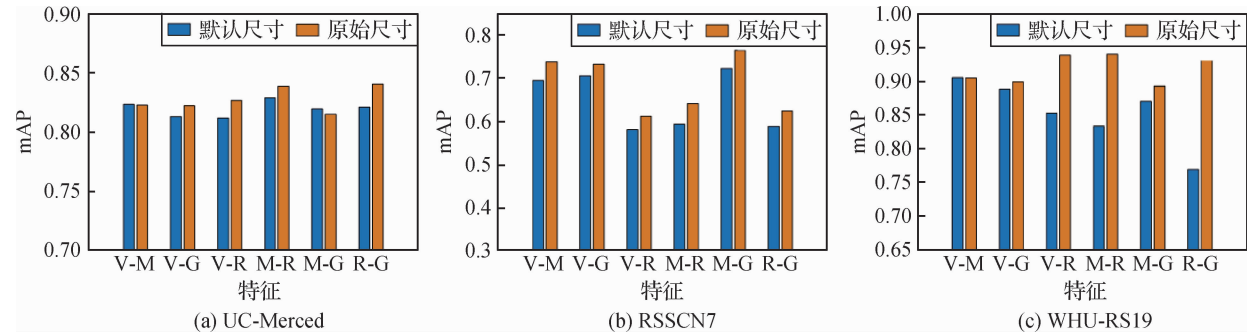


图 3 不同数据集中不同尺寸图像的 mAP 对比

Fig. 3 Comparison of mAP of images for different sizes on the different datasets
((a) UC-Merced ; (b) RSSCN7 ; (c) WHU-RS19)

优势更明显。RSSCN7 数据集在原始输入尺寸下的 mAP 比默认尺寸下的 mAP 提高了 4% 左右, AN-MRR 降低了 3% 左右。在 WHU-RS19 数据集中, 原始输入尺寸下的 mAP 比默认尺寸的 mAP 提高了 7% 左右, ANMRR 降低了 6% 左右。其中检索效果最好的是 M-R, mAP 达到 0.940 0。由此可知, 随着输入图像的原始尺寸增加, 在原始尺寸下提取特征

的检索结果越好。

2.4 查准率—查全率比较

图 4 展示了不同融合特征的查准率—查全率曲线, UC-Merced 数据集返回图像数目最少为 1, 最多为 420。RSSCN7 数据集返回图像数目最少为 1, 最多为 560。WHU-RS19 数据集返回图像数目最少为 1, 最多为 190。

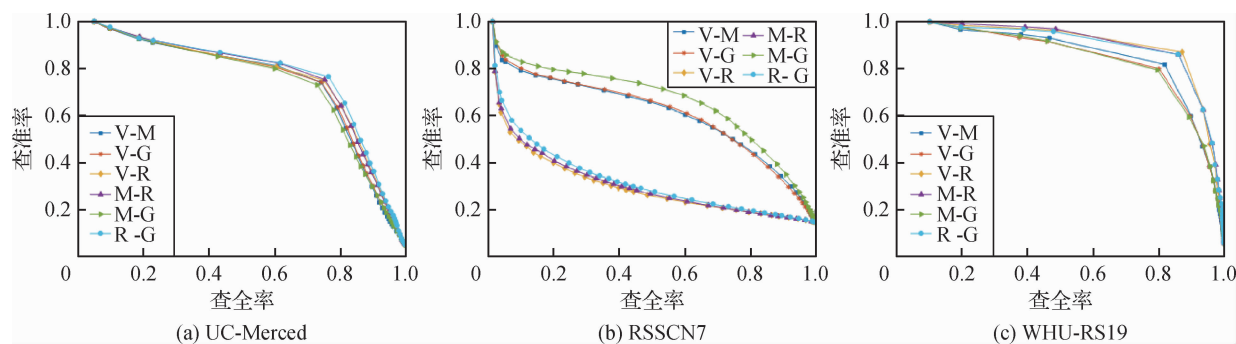


图 4 不同数据集中不同特征的查准率—查全率曲线

Fig. 4 Precision-recall curves for different features on the different datasets

((a) UC-Merced; (b) RSSCN7; (c) WHU-RS19)

在 UC-Merced 数据集中, R-G 曲线位于最顶端, 检索性能最好, 其次是 M-R, 返回图像数目较少时, V-R 检索性能低于 M-R, 随着返回图像数目增加, V-R 的检索性能提升, 与 M-R 持平。在 RSSCN7 数据集中, M-G 曲线的检索性能最好, V-G 与 V-M 检索性能差距不大, 仅次于 M-G, 其他 3 种融合特征性能较差。在 WHU-RS19 数据集中, M-G 曲线位于最低端, 即检索性能最差, R-G、M-R 与 V-R 检索性能差距不大, 三者的检索效果都较好。

总体上看, 6 种融合特征在 UC-Merced 数据集上的检索结果最接近, 这与 4 种单一特征的检索性能接近有关。在 WHU-RS19 数据集上, 不同融合特征的检索结果比较接近, 而在 RSSCN7 数据集中, V-R、M-R 和 R-G 这 3 条曲线位于底端, 检索性能不佳, 这是由于单一特征 R 在 RSSCN7 数据集的检索能力较差, 与其他特征融合后, 反而会降低其他特征的检索能力。因此, 在进行 DCA 融合时, 当不同单一特征的检索能力相近时, 通过 DCA 得到的融合特征检索性能提升较大, 而当不同单一特征的检索能力相差较大时, 融合特征受检索性能较差的单一特征影响较多。

2.5 多种特征融合

为了比较多种特征融合后的检索性能, 将 3 种

和 4 种特征进行 DCA 融合, 实验结果如表 2 所示。

由表 1 与表 2 的对比可知, 当融合特征的数目增加时, 相应的检索能力也随之增强, 这是因为多种不同特征进行融合可以更好地表达内容复杂的遥感图像。在 3 个数据集中, 以 UC-Merced 数据集的检索性能改进最大, 在 RSSCN7 与 WHU-RS19 数据集中, 多特征融合的检索能力也略有提高, 但提升效果不如 UC-Merced 数据集, 由于这两个数据集图像原始尺寸较大, 对特征表达力要求更高, 因此改进效果并不明显。

从表 2 可以看出, 融合特征的数目并不是越多越好。在 RSSCN7 与 WHU-RS19 数据集中, 4 种特征融合的检索性能不如 3 种特征融合的检索性能, 这是因为在 RSSCN7 数据集中, 单一特征 R 的检索结果与其他 3 种单一特征相差较大, 因此 4 种特征融合性能不如 V-M-G。在 WHU-RS19 数据集中, 也是由于单一特征 G 的检索结果不如其他 3 种特征, 导致 4 种特征融合的检索结果不如 V-M-R。而在 UC-Merced 数据集中, 由于 4 种特征的检索能力接近, 因此 4 种特征融合的检索结果优于 3 种特征。

由以上分析可知, 多种特征进行 DCA 融合时的检索性能不仅与特征数目有关, 更与单一特征的检索能力有关。因此, 选择检索能力都比较好且比较

表 2 不同数据集多种特征融合结果
Table 2 Multiple feature fusion results on different datasets

特征	UC-Merced			RSSCN7			WHU-RS19		
	mAP	ANMRR	维数	mAP	ANMRR	维数	mAP	ANMRR	维数
V-M-G	0.841 6	0.133 0	60	0.764 7	0.188 5	36	0.909 5	0.072 4	54
V-M-R	0.846 1	0.129 0	60	0.692 8	0.217 9	36	0.941 5	0.046 5	54
V-G-R	0.848 1	0.126 0	60	0.719 0	0.240 0	36	0.938 1	0.049 1	54
M-G-R	0.852 4	0.122 8	60	0.725 6	0.208 7	36	0.940 6	0.046 7	54
V-G-R-M	0.854 0	0.123 1	80	0.741 4	0.192 0	48	0.940 1	0.047 2	72

注:加粗字体为每列最优结果。

接近的特征进行融合,可以更好地提高遥感图像的检索性能。

2.6 与其他方法比较

现有方法大多在 UC-Merced 数据集中进行实验,表 3 是在 UC-Merced 数据集上,本文方法与其他方法的比较结果。这些方法都将数据集中 80% 的图像用于训练,数据集中剩余的 20% 图像用于评估检索的性能。

在表 3 的方法中,LBP(local binary pattern)和 BoVW(bag of visual words)是两种典型的浅层特征,ReCNN+(Zhou 等,2018)则采用多标签的方法来优化检索性能,以上 3 种是单一特征检索方法,其他方法都采用多特征融合来检索遥感图像。DCA-VGG19(Shelhamer 等,2017)提取了 VGG19 的第 1 个和第 2 个全连接层特征,将这两个特征进行 DCA 融合后再用于遥感图像检索。M_V_G 和 M_V_G_BoVW(Ge 等,2018)则是将多种特征通过距离加权的方式融

合在一起,M_V_G 融合了 VGGM、VGG16 和 GoogLeNet 对应的 CNN 特征,M_V_G_BoVW 是将 VGGM、VGG16 和 GoogLeNet 对应的 CNN 特征和 BoVW 融合在一起。

从表 3 可以看出,本文方法明显优于 LBP 和 BoVW。与 ReCNN+相比,本文方法将 mAP 值提高了 15.26%。除了优于单一特征的检索结果外,本文方法也优于多特征融合的检索结果。与 DCA-VGG19 相比,本文方法将 ANMRR 值降低了 17.20%;与 M_V_G_BoVW 比较时,V-M-con 虽然只融合了其中的两种特征,但是其检索性能更好。实验结果表明,基于 DCA 的特征融合方法能够有效用于高分辨率遥感图像检索。

3 结 论

本文提出一种基于 DCA 的特征融合方法用于高分辨率遥感图像检索。在特征提取阶段,为了综合不同 CNN 的学习参数来更全面地表达高分辨率遥感图像的内容,选择不同 CNN 的高层特征进行融合;为了保留图像更多的原始空间信息,在不同的 CNN 中,提取图像在原始输入尺寸下的高层特征,并保存为特征图的形式;接着,对高层特征采用最大池化的方法来提取图像的显著特征。在特征融合阶段,根据 DCA 得到代表性更强的特征,并通过串联和相加的方式组合在一起。DCA 方法采用对角化散布矩阵对不同类别特征进行分离,并利用奇异值分解对两组特征进行转换来增加同类特征之间的相关性,得到的融合特征既增强了同类特征之间的联系,又突出了不同类特征之间的差异。因此本文

表 3 不同方法的比较
Table 3 Comparison of different methods

方法	ANMRR	mAP
LBP	0.740 0	0.217 0
BOVW	0.538 0	0.398 0
ReCNN +	0.264 0	0.688 0
DCA-VGG19	0.304 1	0.646 7
M_V_G	0.203 8	0.762 7
M_V_G_BoVW	0.203 1	0.763 1
V-M-con	0.151 1	0.823 1
本文	0.132 1	0.840 6

注:加粗字体为每列最优结果。

提出的融合特征不仅维数低,并且判别能力强。实验结果表明,基于DCA的融合特征用于遥感图像检索的效率和准确率都比较高,尤其当检索能力比较接近的特征进行融合时,检索结果提升明显。

本文提出的融合特征在遥感数据集上的检索性能较好,但提取的高层特征来自现有预训练的CNN,适用小规模遥感数据集。下一步工作中,针对规模较大的遥感数据集,研究设计适用遥感图像的深度学习并提取特征进行图像检索。

参考文献 (References)

- Chaib S, Liu H, Gu Y F and Yao H X. 2017. Deep feature fusion for VHR remote sensing scene classification. *IEEE Transactions on Geoscience and Remote Sensing*, 55(8): 4775-4784 [DOI: 10.1109/TGRS.2017.2700322]
- Chatfield K, Simonyan K, Vedaldi A and Zisserman A. 2014. Return of the devil in the details: delving deep into convolutional nets [EB/OL]. [2019-12-23]. <https://arxiv.org/pdf/1405.3531.pdf>
- Du P J, Li E Z, Xia J S, Samat A and Bai X Y. 2019. Feature and model level fusion of pretrained CNN for remote sensing scene classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(8): 2600-2611 [DOI: 10.1109/JSTARS.2018.2878037]
- Farah J, Zhou J, Awrangjeb M and Gao Y S. 2018. Fusion of hyperspectral and LiDAR data using discriminant correlation analysis for land cover classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(10): 3905-3917 [DOI: 10.1109/JSTARS.2018.2868142]
- Ge Y, Jiang S L, Xu Q Y, Jiang C L and Ye F M. 2018. Exploiting representations from pre-trained convolutional neural networks for high-resolution remote sensing image retrieval. *Multimedia Tools and Applications*, 77(13): 17489-17515 [DOI: 10.1007/s11042-017-5314-5]
- Haghighat M, Abdel-Mottaleb M and Alhalabi W. 2016. Discriminant correlation analysis: real-time feature level fusion for multimodal biometric recognition. *IEEE Transactions on Information Forensics and Security*, 11(9): 1984-1996 [DOI: 10.1109/TIFS.2016.2569061]
- 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, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- Hu F, Tong X Y, Xia G S and Zhang L P. 2017. Delving into deep representations for remote sensing image retrieval//*Proceedings of the 13th International Conference on Signal Processing*. Chengdu, China: IEEE: 198-203 [DOI: 10.1109/ICSP.2016.7877823]
- Hu F, Xia G S, Hu J W and Zhang L P. 2015. Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. *Remote Sensing*, 7(11): 14680-14707 [DOI: 10.3390/rs71114680]
- Husain S S and Bober M. 2019. REMAP: multi-layer entropy-guided pooling of dense CNN features for image retrieval. *IEEE Transactions on Image Processing*, 28(10): 5201-5213 [DOI: 10.1109/TIP.2019.2917234]
- Imbriaco R, Sebastian C, Bondarev E and de With P H N. 2019. Aggregated deep local features for remote sensing image retrieval. *Remote Sensing*, 11(5): #493 [DOI: 10.3390/rs11050493]
- Klaric M N, Scott G J and Shyu C R. 2012. Multi-index multi-object content-based retrieval. *IEEE Transactions on Geoscience and Remote Sensing*, 50(10): 4036-4049 [DOI: 10.1109/TGRS.2012.2187353]
- Li G, Li L L, Zhu H, Liu X and Jiao L C. 2019a. Adaptive multiscale deep fusion residual network for remote sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 57(11): 8506-8521 [DOI: 10.1109/TGRS.2019.2921342]
- Li S W, Purushotham S, Chen C, Ren Y Z and Kuo C C J. 2017. Measuring and predicting tag importance for image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(12): 2423-2436 [DOI: 10.1109/TPAMI.2017.2651818]
- Li Y Y, Wang Q, Liang X X and Jiao L C. 2019b. A novel deep feature fusion network for remote sensing scene classification//*Proceedings of 2019 IEEE International Geoscience and Remote Sensing Symposium*. Yokohama, Japan: IEEE: 5484-5487 [DOI: 10.1109/IGARSS.2019.8898900]
- Liu R J, Zhang H X and Kong W J. 2015. Image retrieval method based on feature-associated fusion. *Journal of University of Jinan (Science and Technology)*, 29(5): 327-332 (刘润杰, 张化祥, 孔文杰. 2015. 基于特征关联融合的图像检索方法. *济南大学学报(自然科学版)*, 29(5): 327-332) [DOI: 10.13349/j.cnki.jdxbn.2015.05.002]
- Lu L Z, Liu R Y and Liu N. 2004. Remote sensing image retrieval using color and texture fused features. *Journal of Image and Graphics*, 9(3): 328-333 (陆丽珍, 刘仁义, 刘南. 2004. 一种融合颜色和纹理特征的遥感图像检索方法. *中国图象图形学报*, 9(3): 328-333) [DOI: 10.11834/jig.20040361]
- Peng Y F, Song X N, Wu H and Zi L L. 2019. Remote sensing image retrieval combined with deep learning and relevance feedback. *Journal of Image and Graphics*, 24(3): 420-434 (彭晏飞, 宋晓男, 武宏, 訾玲玲. 2019. 结合深度学习与相关反馈的遥感图像检索. *中国图象图形学报*, 24(3): 420-434) [DOI: 10.11834/jig.180384]
- Radenović F, Tolias G and Chum O. 2019. Fine-tuning CNN image retrieval with no human annotation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(7): 1655-1668 [DOI: 10.1109/TPAMI.2018.2846566]

- Shelhamer E, Long J and Darrell T. 2017. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(4): 640-651 [DOI: 10.1109/TPAMI.2016.2572683]
- Simonyan K and Zisserman A. 2015. Very deep convolutional networks for large-scale image recognition [EB/OL]. [2019-12-23]. <https://arxiv.org/pdf/1409.1556v6.pdf>
- Sun Q S, Zeng S G, Liu Y, Heng P A and Xia D S. 2005. A new method of feature fusion and its application in image recognition. *Pattern Recognition*, 38(12): 2437-2448 [DOI: 10.1016/j.patcog.2004.12.013]
- Szegedy C, Liu W, Jia Y Q, Sermanet R, Reed S, Anguelov D, Erhan D, Vanhoucke V and Rabinovich A. 2015. Going deeper with convolutions [EB/OL]. [2019-12-23]. <https://arxiv.org/pdf/1409.4842.pdf>
- Tong X Y, Xia G S, Hu F, Zhong Y F, Datcu M and Zhang L P. 2017. Exploiting deep features for remote sensing image retrieval: a systematic investigation [EB/OL]. [2019-12-23]. <https://arxiv.org/pdf/1707.07321.pdf>
- Vedaldi A and Lenc K. 2015. MatConvNet: convolutional neural networks for MATLAB//Proceedings of the 23rd ACM International Conference on Multimedia. Brisbane, Australia; ACM: 689-692 [DOI: 10.1145/2733373.2807412]
- Wang G L, Fan B, Xiang S M and Pan C H. 2017. Aggregating rich hierarchical features for scene classification in remote sensing imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(9): 4104-4115 [DOI: 10.1109/JSTARS.2017.2705419]
- Wang Y B, Zhang L Q, Tong X H, Zhang L, Zhang Z X, Liu H, Xing X Y and Mathiopoulos P T. 2016. A three-layered graph-based learning approach for remote sensing image retrieval. *IEEE Transactions on Geoscience and Remote Sensing*, 54(10): 6020-6034 [DOI: 10.1109/TGRS.2016.2579648]
- Yang K, Li C M, Zhou W X, Cheng Q M and Ren Y C. 2019. Remote sensing image retrieval based on multi-layer feature integration of convolution neural networks. *Science of Surveying and Mapping*, 44(7): 9-15, 34 (杨珂, 李从敏, 周维勋, 程起敏, 任应超. 2019. 卷积神经网络多层特征联合的遥感图像检索. *测绘科学*, 44(7): 9-15, 34) [DOI: 10.16251/j.cnki.1009-2307.2019.07.002]
- Yang Y and Newsam S. 2013. Geographic image retrieval using local invariant features. *IEEE Transactions on Geoscience and Remote Sensing*, 51(2): 818-832 [DOI: 10.1109/tgrs.2012.2205158]
- Ye F M, Xiao H, Zhao X Q, Dong M, Luo W and Min W D. 2018. Remote sensing image retrieval using convolutional neural network features and weighted distance. *IEEE Geoscience and Remote Sensing Letters*, 15(10): 1535-1539 [DOI: 10.1109/LGRS.2018.2847303]
- Zhou W X, Deng X Q and Shao Z F. 2018. Region convolutional features for multi-label remote sensing image retrieval [EB/OL]. [2019-12-23]. <https://arxiv.org/pdf/1807.08634.pdf>
- Zhou W X, Newsam S, Li C M and Shao Z F. 2017. Learning low dimensional convolutional neural networks for high-resolution remote sensing image retrieval. *Remote Sensing*, 9(5): #489 [DOI: 10.3390/rs9050489]
- Zou Q, Ni L H, Zhang T and Wang Q. 2015. Deep learning based feature selection for remote sensing scene classification. *IEEE Geoscience and Remote Sensing Letters*, 12(11): 2321-2325 [DOI: 10.1109/LGRS.2015.2475299]

作者简介



葛芸, 1983年生, 女, 讲师, 主要研究方向为遥感图像处理与机器学习。

E-mail: geyun@nchu.edu.cn



储珺, 通信作者, 女, 教授, 主要研究方向为图像处理、计算机视觉、计算机图形学和数据融合。

E-mail: chujun99602@163.com

马琳, 女, 硕士研究生, 主要研究方向为遥感图像处理与机器学习。E-mail: ml960221@qq.com