

基于深度哈希算法的极光图像分类与检索方法

陈昌红* 彭腾飞 千宗良

(南京邮电大学通信与信息工程学院 南京 210003)

摘要: 面对形态万千、变化复杂的海量极光数据, 对其进行分类与检索为进一步研究地球磁场物理机制和空间信息具有重要意义。该文基于卷积神经网络(CNN)对图像特征提取方面的良好表现, 以及哈希编码可以满足大规模图像检索对检索时间的要求, 提出一种端到端的深度哈希算法用于极光图像分类与检索。首先在CNN中嵌入空间金字塔池化(SPP)和幂均值变换(PMT)来提取图像中多种尺度的区域信息; 其次在全连接层之间加入哈希层, 将全连接层最能表现图像的高维语义信息映射为紧凑的二值哈希码, 并在低维空间使用汉明距离对图像对之间的相似性进行度量; 最后引入多任务学习机制, 充分利用图像标签信息和图像对之间的相似度信息来设计损失函数, 联合分类层和哈希层的损失作为优化目标, 使哈希码之间可以保持更好的语义相似性, 有效提升了检索性能。在极光数据集和 CIFAR-10 数据集上的实验结果表明, 所提出方法检索性能优于其他现有检索方法, 同时能够有效用于极光图像分类。

关键词: 极光图像; 分类与检索; 卷积神经网络; 哈希编码; 多尺度特征融合

中图分类号: TN911.73

文献标识码: A

文章编号: 1009-5896(2020)12-3029-08

DOI: 10.11999/JEIT190984

Aurora Image Classification and Retrieval Method Based on Deep Hashing Algorithm

CHEN Changhong PENG Tengfei GAN Zongliang

(College of Communication and Information Technology, Nanjing University of Posts and
Telecommunications, Nanjing 210003, China)

Abstract: It is of great significance to classify and retrieve the vast amount of aurora data with various forms and complex changes for the further study of the physical mechanism of the geomagnetic field and spatial information. In this paper, an end-to-end deep hashing algorithm for aurora image classification and retrieval is proposed based on the good performance of CNN in image feature extraction and the fact that hash coding can meet the retrieval time requirement of large-scale image retrieval. Firstly, Spatial Pyramid Pooling(SPP) and Power Mean Transformation(PMT) are embedded in Convolutional Neural Network (CNN) to extract multi-scale region information in the image. Secondly, a Hash layer is added between the fully connected layer to Mean Average Precision(MAP) the high-dimensional semantic information that can best represent the image into a compact binary Hash code, and the hamming distance is used to measure the similarity between the image pairs in the low-dimensional space. Finally, a multi-task learning mechanism is introduced to design the loss function by making full use of similarity information between the image label information and the image pairs. The loss of classification layer and Hash layer are combined as the optimization objective, so that a better semantic similarity between Hash code can be maintained, and the retrieval performance can be effectively improved. The results show that the proposed method outperforms the state-of-art retrieval algorithms on aurora dataset and CIFAR-10 datasets, and it can also be used in aurora image classification effectively.

Key words: Aurora image; Classification and retrieval; Convolutional Neural Network(CNN); Hash coding; Multi-scale feature fusion

收稿日期: 2019-12-09; 改回日期: 2020-08-09; 网络出版: 2020-08-13

*通信作者: 陈昌红 chenqh@njupt.edu.cn

基金项目: 国家自然科学基金(61501260), 江苏省研究生科研与实践创新计划(KYCX17_0776)

Foundation Items: The National Natural Science Foundation of China (61501260), The Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX17_0776)

1 引言

极光(aurora)是由太阳带电粒子流在极区电离层发生物理电离作用时呈现出的自然发光现象。极光形态特征变化与日地间电磁活动有着内在联系,这对太阳地球间的影响规律和太阳风暴等研究奠定了基础,因此极光图像的分类和检索成为日地物理学研究的重要途径。而面对这些海量极光数据,若通过手工选择或传统的基于SIFT的方法来进行极光图像的分类与检索效果不佳,所以如何有效合理利用这些重要数据已成为极光研究人员急需解决的问题。

一直以来研究人员都在进行着极光方面的研究,图像分类研究主要有文献[1]提出基于局部二值模式和分块化融合思想的极光图像分类方法;文献[2]提出基于融合显著性信息的文档主题生成模型对极光图像进行分类。检索研究主要有文献[3]提出利用极光的形状特征的极光图像检索方法;文献[4]将全局特征和局部特征融合实现极光图像检索;文献[5]利用CNN提取多尺度特征,并针对极光结构设计极区划分方案实现极光图像检索。以上方法大都采用传统的特征提取方法,对于极光图像特征的代表不够全面,存在性能不是很好、检索耗时长、存储开销大等问题。

近年来得益于深度学习在图像处理方面的应用^[6],使得人们开始不断地对新的应用场景进行思考和尝试,深度哈希算法就是研究者们提出的一种新的探索。深度哈希算法利用CNN提取的高层语义特征代替传统哈希算法使用手工特征来进行哈希编码,不仅有效避免了“语义鸿沟”问题,还提高了检索性能。文献[7]提出通过构造相似矩阵的深度

哈希算法CNNH;文献[8]提出基于图像标签对来设计损失函数的DPSH算法;文献[9]提出基于图像对间的相似度的DSH算法。上述算法对图像检索的性能都有一定程度上的改善,但CNNH算法的特征提取和哈希学习是分割开的,在训练过程中不能相互反馈;而DPSH和DSH算法都是基于图像对的相似信息来训练网络,标签信息没有全部发挥作用。

基于上述分析,本文结合CNN和哈希编码的优势,提出一种基于深度哈希算法的极光图像分类与检索方法,该算法的主要特点有:

(1) 在CNN之间加入哈希层实现了数据降维,还减少直接利用高维特征向量检索耗时、存储空间大的缺点,同时还将哈希学习和特征提取融合到同一个框架中,实现一种端到端的极光图像分类与检索算法;

(2) 多尺度特征融合,将空间金字塔池化(Spatial Pyramid Pooling, SPP)和幂均值变换(Power Mean Transformation, PMT)嵌入到CNN中,SPP不仅可以解决CNN结构对输入尺寸的限制而导致信息损耗的问题,还能对多个不同区域的特征信息融合,同时加入PMT,利用数学函数引入非线性变换,使得模型可以从数据中学习得更复杂的关系;

(3) 引入多任务学习机制,充分利用样本图像的标签信息和图像对之间的相似度信息来进行模型的训练,将哈希层的损失作为优化目标之一来保持二值哈希码之间良好的语义相似性。

2 基于深度哈希算法的极光图像分类与检索

本文方法的框架如图1所示,上下两部分分别为训练网络和测试网络。

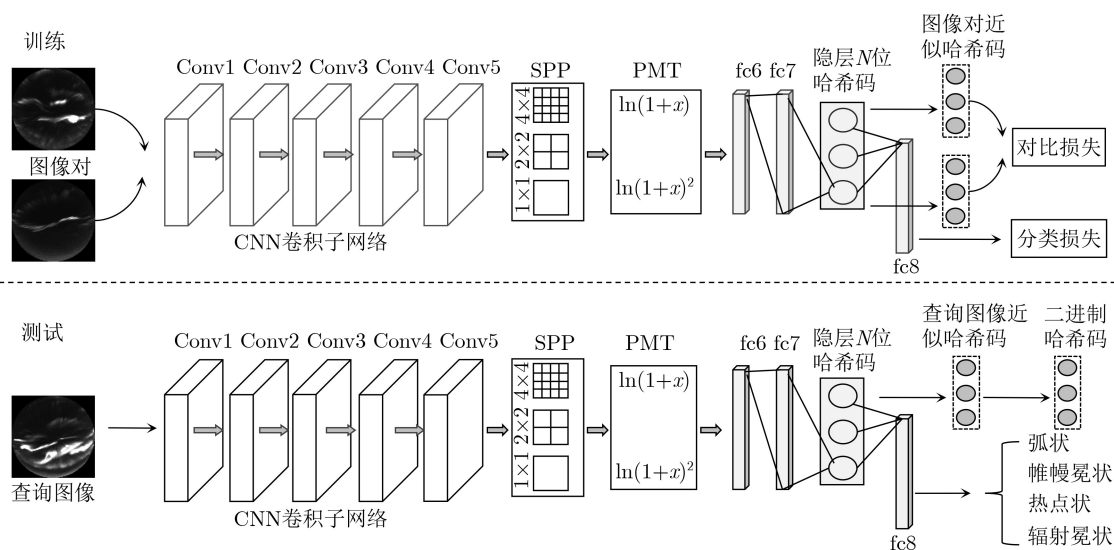


图1 本文算法的训练和测试框图

其中训练网络主要由3部分组成：(1)CNN卷积子网络；(2)空间金字塔池化(SPP)和幂均值变换(PMT)；(3)损失函数。由于要用到图像之间的相似度信息，因此网络的输入层要求以图像对的形式输入。首先图像对经CNN卷积子网络提取局部特征；然后，最后一个卷积层输出若干特征图经过SPP和PMT融合图像多尺度的区域特征；接着将多尺度特征送入全连接层，哈希层中使用激活函数Sigmoid将输出的编码值映射到[0,1]区间，分别计算分类损失、对比损失，并反向传播进行模型训练。测试网络基本上与训练网络相同，只需去掉损失函数部分。

2.1 CNN卷积子网络

CNN卷积子网络用来提取图像的特征，本文采用AlexNet^[10]作为基本网络结构，为了获得更有效的特征，将池化层pool3替换为SPP并加入PMT。其中Conv1-Conv5为AlexNet中结构，SPP的参数是金字塔的高度为3和池化方式为MaxPooling。

为了提高性能以及减少模型参数、加快训练速度，本文对AlexNet卷积子网络进行改进(用改进的AlexNet表示，记为Im-AlexNet)。经过大量网络结构的调整实验，主要从以下几点改进：

(1) Conv1 的卷积核由 11×11 改为 5×5 。采用较小的卷积核网络可以提取更加细致、深层次的图像特征，进而提高性能，同时可以减少网络参数；

(2) Conv2-Conv5 每层前加入1层 1×1 卷积，并且其后加 ReLU 激活函数。 1×1 卷积为2次提取图像特征，可以更好地表征图像信息；加深网络的同时实现通道的信息整合，同时ReLU激活函数可以加快训练速度，增加网络的非线性特性，增强数据拟合准确性；

(3) 全连接层fc6和fc7节点数由4096改为1024，可以很大程度上减少网络参数，加快模型训练速度。经过对AlexNet网络结构进行改进，在泛化性、模型训练速度以及整体的鲁棒性等方面有所提高。

2.2 空间金字塔池化(SPP)和幂均值变换(PMT)

2.2.1 空间金字塔池化

空间金字塔池化(SPP)^[11]采用了多尺度分块的方法，通过使用多个不同大小的池化操作保证固定的特征向量输出，进而提升网络输入自适应能力，实现不同尺寸的输入，其中选择大小分别为 4×4 , 2×2 , 1×1 的池化框。SPP可以对最后一个卷积层输出的每一个特征图进行不同尺度的特征提取，最后生成融合多个区域信息的多尺度特征^[12]，同时避免了CNN对输入图像尺寸的限制而造成极光图像经

过裁剪、缩放等操作丢失部分重要信息。图2为空间金字塔池化示意图。

SPP的原理如图2所示：

(1) 首先是输入层，其大小可以是任意尺寸，输入图像经过CNN卷积子网络进行卷积运算，到最后一个卷积层输出若干个特征图；

(2) 然后进入SPP层，从左往右3个格子表示池化框大小分别为 1×1 , 2×2 , 4×4 ，先看第3个格子，假设最后一个卷积层的输出为512个 7×7 的特征图，则将其分成 16×256 份，其他两个池化框同理，即将特征图分别分成 4×256 份和 1×256 份；

(3) 将特征图分成若干等份后分别进行池化操作，一般选择Max Pooling，这样最后一个卷积层得到的特征图通过SPP层后被转化成了 $16 \times 256 + 4 \times 256 + 1 \times 256 = 21 \times 256$ 的矩阵，在送入全连接层时扩展成1维矩阵，即 1×5376 ，所以第1个全连接层的节点数就设置成5376。

2.2.2 幂均值变换

在传统的CNN模型中非线性变换主要来源于激活函数和池化层。为了增加和多样化网络中的非线性变换，在SPP层和之后的全连接层之间引入幂均值变换(PMT)^[13]，文献[13]验证了引入PMT增加模型非线性的多样性能够有效地提高模型的性能。PMT的原理是给定一个输入 \mathbf{x} ， \mathbf{x} 为图像特征向量。PMT将 \mathbf{x} 转换为 $\ln(\mathbf{x} + \mathbf{b})$, $\ln^2(\mathbf{x} + \mathbf{b})$ ，其中 \mathbf{b} 是参数矩阵。在前向传递时，使 $\partial l / \partial \mathbf{y}$ 等于PMT输出的梯度值，其中 l 为 \mathbf{x} 经过PMT后的输出值。因为 $\mathbf{y} = [\ln(\mathbf{x} + \mathbf{b}), \ln^2(\mathbf{x} + \mathbf{b})]$ ，由链式法则得到

$$\frac{\partial l}{\partial \mathbf{y}} = \frac{\partial l}{\partial [\ln(\mathbf{x} + \mathbf{b}), \ln^2(\mathbf{x} + \mathbf{b})]} \quad (1)$$

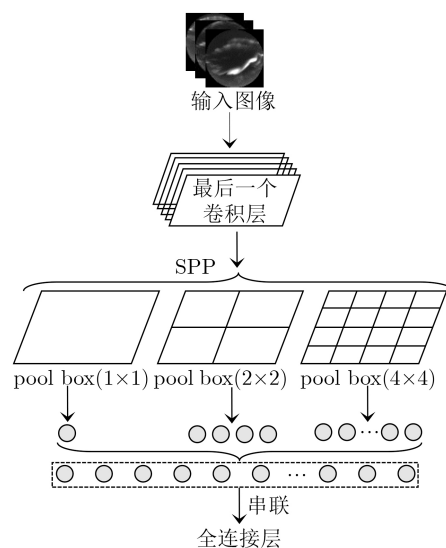


图2 空间金字塔池化示意图

$$\frac{\partial l}{\partial \mathbf{x}^T} = \frac{\partial l}{\partial \mathbf{y}^T} \frac{\partial \mathbf{y}}{\partial \mathbf{x}^T} = \frac{\partial l}{\partial (\ln(\mathbf{x} + \mathbf{b}))^T} \frac{\ln(\mathbf{x} + \mathbf{b})}{\partial \mathbf{x}^T} + \frac{\partial l}{\partial (\ln^2(\mathbf{x} + \mathbf{b}))^T} \frac{\ln^2(\mathbf{x} + \mathbf{b})}{\partial \mathbf{x}^T} \quad (2)$$

有了前向和反向传播公式, PMT就比较容易集成到CNN模型中, 并使用现有的优化器进行端到端的训练, 从而可以从数据中学习到更复杂的信息, 有利于提取全面准确的特征。

2.3 二进制哈希码的生成

CNN的卷积层提取的是局部特征, 而全连接层提取的是能够很好表征输入图像的高层语义信息。AlexNet的fc7层最能表现图像的高层语义信息, 但fc7层提取的特征是一个4096维向量, 直接用来做相似性计算时面临着计算耗时、占用大量存储空间的问题。

哈希算法作为当前主流的数据降维方法之一, 主要目的是将任意长度的输入投影成一串固定长度的二进制哈希码, 并且利用汉明距离来度量二值编码之间的相似性, 使得不同哈希码能保持图像之间的语义相似性。本文在AlexNet的fc7和fc8层之间加入哈希层, 使得fc7层学习到的高维语义特征映射为哈希码, 这样哈希码中也包含了图像的语义信息, 可以明显提高图像检索的效果。

如图1中测试网络所示, 查询图像输入测试网络, 会依次通过CNN卷积子网络、SPP_MPT层、全连接层以及哈希层, 提取哈希层中Sigmoid输出的值, 记为 $P(i)$, 阈值设定为0.5, 把 $[0, 1]$ 之间的输出值通过 H_i 转换为二进制哈希码用来检索。本文检索使用的哈希码长度为 N , 对于 $i = 1, 2, \dots, N$, 每一位哈希码为

$$H_i = \begin{cases} 1, & P(i) \geq 0.5 \\ 0, & \text{其他} \end{cases} \quad (3)$$

2.4 损失函数设计

多任务学习MTL(Multi-Task Learning)就是利用额外的信息来提升各个任务的学习性能而提出的一种机制。本文在设计损失函数时充分利用极光图像的标签信息和图像对之间的相似性信息, 避免图像标签信息的浪费, 同时将哈希层的损失作为优化目标之一, 能够进一步保持二值哈希码之间的语义相似性。本文方法的损失函数表示为

$$L_t = \mu L_{\text{cel}} + \lambda L_{\text{cl}} \quad (4)$$

式中 L_{cel} 和 L_{cl} 分别表示交叉熵损失函数(CrossEntropyLoss)和对比损失函数(ContrastiveLoss), μ 和 λ 为权重因子。

式(4)符号右侧第1部分为交叉熵损失函数。分类损失层以每一个图像预测的类别概率以及其真实

标签作为输入, 用来衡量样本预测类别与真实标签之间的相似度。交叉熵损失函数的公式为

$$L_{\text{cel}} = -\frac{1}{n} \sum_x [y \ln y^* + (1-y) \ln(1-y^*)] \quad (5)$$

其中 x 表示样本, n 表示样本的总数, y 为期望的输出, y^* 为神经元的实际输出。

式(4)符号右侧第2部分为对比损失函数。假设有 N 对训练样本 $(I_{i,1}, I_{i,2})$, $i = 1, 2, \dots, N$, y_i 衡量样本对之间的相似度(y_i 等于0表示相似, y_i 等于1表示不相似), 则优化目标是尽量减小相似样本间的距离并拉大不相似样本间的距离, 即

$$L_{\text{cl}} = \sum_{i=1}^N \left\{ \frac{1}{2} (1-y_i) \|b_{i,1} - b_{i,2}\|_2^2 + \frac{1}{2} y_i \max(m - \|b_{i,1} - b_{i,2}\|_2^2, 0) \right\} \quad (6)$$

其中, $b_{i,1}$ 和 $b_{i,2}$ 表示图像对由哈希层输出的近似哈希码, 用欧式距离来进行优化; 边界参数 $m > 0$; 前一项是衡量相似图像对应的近似哈希码之间所存在的距离而计算出损失; 后一项表示在图像对不相似的情况下, 如果对应的近似哈希码之间的距离小于 m 的值就会产生损失。

3 实验结果与分析

3.1 数据集与性能评价

为了验证本文方法的有效性, 在极光图像数据集和公共数据集CIFAR-10上分别进行评估。本文使用的极光图像数据集, 来自中国极地科考站北极黄河站的极光观测数据, 总共有8001幅极光图像, 其中包含3934幅弧状极光(arc)、1786幅帷幔冕状极光(drapery)、1497幅辐射冕状极光(radial)和784幅热点状极光(hot-spot)。如图3所示为4类极光类型图像。CIFAR-10数据集包含有60000张大小为 32×32 的图像, 总共有10个类别, 每个类别有6000张图像, 包括50000张训练集和10000张测试集。

本文图像检索的性能评价采用Top-k检索返回的准确率曲线、MAP和查准率-查全率(Precision-Recall, P-R)曲线3项指标。Top-k检索返回的准确率曲线是衡量检索返回的与查询图像二值哈希码汉明距离最小的前 K 张图像中相似图像所占的比例。P-R曲线是指按照汉明距离从小到大的排序, 所有测试图像的平均查全率和平均查准率的曲线图。MAP是指P-R曲线所包围的面积。

图像分类的性能评价指标是平均准确率(Accuracy)和4分类归一化混淆矩阵。

3.2 实验结果与分析

3.2.1 极光数据集上的实验

首先分别验证加入哈希层损失和SPP_PMT层的有效性。实验中使用48位哈希码来进行说明，CNN卷积子网络使用AlexNet的前5个卷积层。表1是考虑哈希层损失(同时使用对比损失和交叉熵损失)与不考虑哈希层损失(只使用交叉熵损失)两种情况下的MAP值和准确率的对比结果。表2表示有无SPP_PMT层两种情况(都加入哈希层损失)下的MAP值和准确率的对比结果。

从表1可以看出在哈希层中引入损失作为优化目标之一使得二进制哈希码之间保持良好的语义相

似性，MAP值和准确率分别提高了9.91%和3.68%，说明联合哈希损失和分类损失可以充分利用图像的标签信息和图像对之间的相似性信息，效果有一定的提升，从而验证了哈希损失的有效性。

从表2可以看出在CNN中嵌入了SPP和PMT，可以提取到多尺度的特征，使得网络可以从图像中捕获更复杂的关系。MAP值和准确率分别提高了4.09%和2.94%，验证了SPP_PMT的有效性。

然后分别用AlexNet, AlexNet-SP(使用AlexNet并加入SPP_PMT)和Im-AlexNet-SP(使用Im-AlexNet并加入SPP_PMT)3种方法进行对比。都考虑了哈希层损失，并设置不同的哈希码长度(bit)。表3为不同哈希码长度下的MAP以及模型参数大小和训练时间对比。图4为MAP, P-R以及Top-k检索返回的准确率曲线。表4为在不同哈希码长度下的分类准确率。图5从左到右分别为AlexNet, AlexNet-SP和Im-AlexNet-SP在bit=48哈希码下的4分类混淆矩阵。

从表3和图4可以看到3种方法MAP, P-R曲线和Top-k检索返回的准确率曲线以及模型参数大小与训练时间对比。经过对AlexNet进行改进后使得新提取的图像特征表达效果更好，在加深网络的同时实现通道的信息整合，不仅MAP和精度有一定程度提高，最高MAP值达到91.73%，而且模型参数和训练时间相比之下都有所减小，在提升检索性

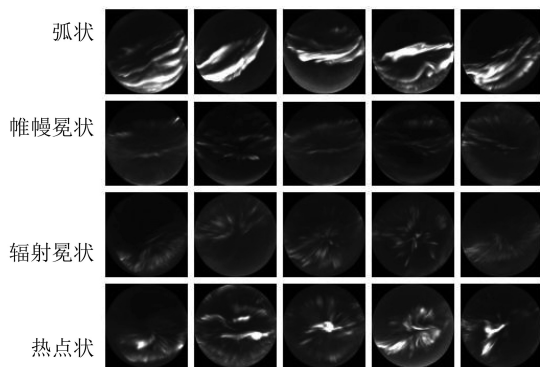


图3 4类极光类型图像

表1 有无哈希层损失两种方法对比

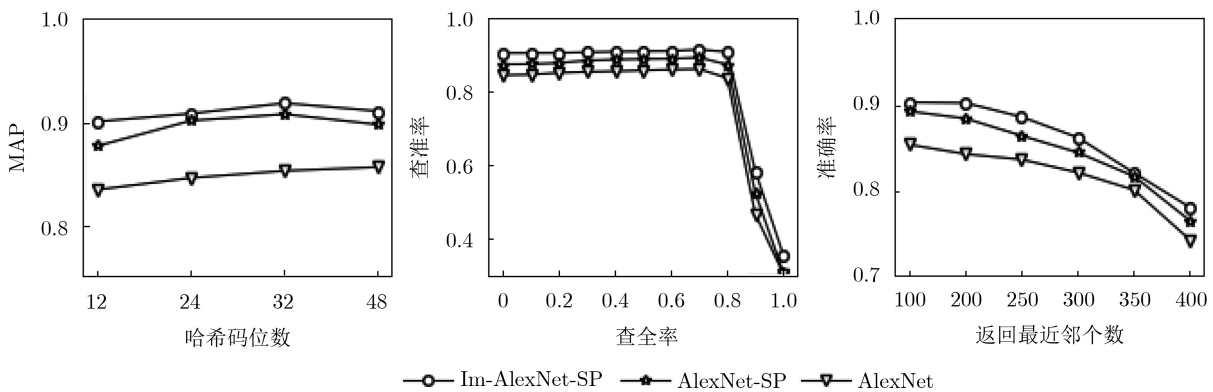
方法	MAP	准确率
不考虑哈希层损失	0.7563	0.8705
考虑哈希层损失	0.8554	0.9073

表2 有无SPP_PMT层两种方法对比

方法	MAP	准确率
不加SPP_PMT	0.8554	0.9073
加入SPP_PMT	0.8963	0.9367

表3 3种方法的MAP以及在bit=48下模型参数大小(MB)和训练时间(min)

方法	不同哈希码长度(bit)下的MAP				参数大小	训练时间
	12	24	32	48		
AlexNet	0.8336	0.8450	0.8518	0.8554	218.20	158
AlexNet-SP	0.8729	0.9004	0.9066	0.8963	179.15	115
Im-AlexNet-SP	0.8995	0.9072	0.9173	0.9095	100.77	80



(a) 不同哈希码长度时的MAP曲线

(b) bit=48时的P-R曲线

(c) bit=48时的Top-k检索返回的准确率曲线

图4 3种方法的MAP, P-R以及Top-k检索返回的准确率曲线

表4 3种方法在不同哈希码长度下的准确率

方法	不同哈希码长度(bit)下的准确率			
	12	24	32	48
AlexNet	0.8964	0.8995	0.8988	0.9073
AlexNet-SP	0.9312	0.9298	0.9325	0.9367
Im-AlexNet-SP	0.9320	0.9305	0.9410	0.9384

能的同时, 加快模型训练速度。

从表4可以看到3种方法的极光图像分类的情况, 前两种方法对比可以看出SPP_PMT层对极光分类效果有较大提高; 后两种方法对比可以看出对AlexNet改进后对于准确率略有提高, 最高达到94.10%。但从图5的混淆矩阵图看出弧状和帷幔冕状极光图像的分类准确率都较高, 热点状极光主要被错分成了弧状类和辐射冕状类, 经分析主要原因是热点状极光具有亮斑、不规则块、涡旋、射线束等多种复杂结构, 与弧状和辐射冕状极光具有局部

相似性, 且热点状极光样本数量较其他几种偏少, 所以容易导致错分情况。

最后将本文方法与其他几种极光图像检索算法(HE^[14], VLAD^[15], MAC^[16]和MS-RMAC^[17])进行对比。表5为本文方法(bit=48)与其他几种算法的MAP以及平均查询时间对比。从性能来看, 与基于传统SIFT特征的方法(HE和VLAD)相比, 基于CNN特征检索方法(MAC, MS-RMAC以及本文方法)的MAP有明显提高, 尤其是本文方法取得最高90.95%的MAP值, 这得益于多尺度CNN特征融合和联合损失来训练。从效率来看, MAC和MS-RMAC是最耗时的, 因为它们使用高维的CNN特征; 相反, 由于PCA对特征进行降维操作, VLAD获得了惊人的效率, HE模型紧随其后。而本文方法使用哈希算法和CNN结合将高维特征向量映射为低维紧凑的二值哈希码, 达到了数据降维的作用, 减少了检索时间。

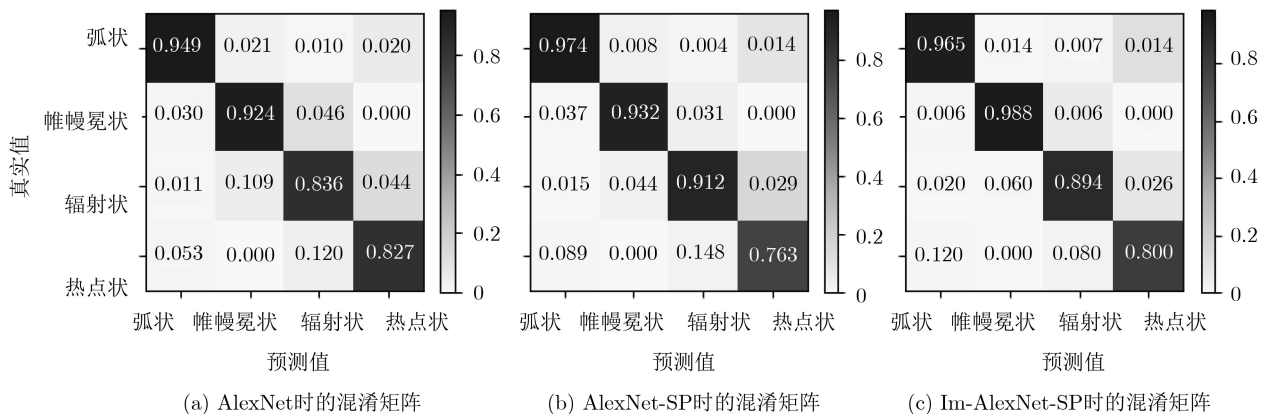


图5 3种方法在哈希码长度为48 bit时的四分类混淆矩阵

3.2.2 CIFAR-10数据集上的实验

为了验证本文方法的检索效果对其他数据集的可扩展性, 同时与其他主流的哈希算法(如LSH^[18], ITQ^[19], KSH^[20], CNNH^[8], DSH^[10], DPSH^[9]等)进行对比。表6是不同哈希算法在CIFAR-10不同哈希码长度下的MAP对比。图6是不同哈希算法在CIFAR-10上MAP, PR以及Top- k 检索返回的准确率曲线。

表5 本文方法与其他极光检索算法的MAP以及平均查询时间对比(s)

方法	MAP	平均查询时间
HE	0.5253	0.65
VLAD	0.5868	0.52
MAC	0.6558	1.22
MS-RMAC	0.6901	2.89
本文Im-AlexNet-SP	0.9095	0.43

从表6和图6中可以看出基于CNN的深度哈希算法(CNNH, DSH, DPSH和本文方法)总体上优于传统的哈希算法(LSH, ITQ和KSH), 其中CNNH算法性能不算好, 由于该算法不属于端到端的学习, 其哈希编码和CNN的训练过程是相互独立

表6 不同哈希算法在CIFAR-10不同哈希码长度下的MAP

方法	不同哈希码长度(bit)下的MAP			
	12	24	32	48
本文Im-AlexNet-SP	0.902	0.904	0.912	0.907
DPSH	0.713	0.727	0.744	0.757
DSH	0.673	0.685	0.690	0.694
CNNH	0.439	0.511	0.509	0.522
KSH	0.303	0.337	0.346	0.356
ITQ	0.162	0.169	0.172	0.175
LSH	0.127	0.137	0.141	0.149

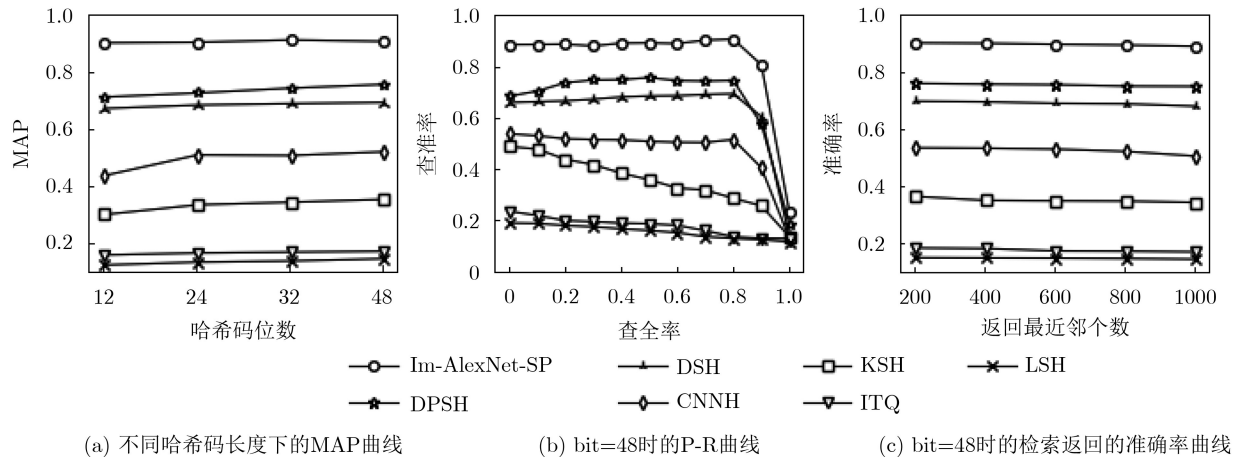


图6 不同哈希算法在CIFAR-10数据集上MAP, P-R以及Top-k 检索返回的准确率曲线

的，不能互相反馈，也就没有发挥出深度网络的强大学习能力。而 DSH算法只使用了图像对的相似度信息，没有加入图像的标签信息，因此检索精度低于本文方法。由以上分析可知，一方面本文方法在 CIFAR-10 上也达到了比较理想的检索效果，其中当哈希长度为32 bit时MAP 值达到91.2%，相较于经典的 DSH 算法相比提升了 22.2%，同DPSH 算法相比MAP值提升了16.8%，从而验证了本文方法的可拓展性和有效性。

4 结束语

本文基于CNN强大的特征学习能力以及哈希编码的数据降维、检索速度快等优势提出了一种新的深度哈希算法，充分利用标签信息和图像对相似度信息来训练网络模型，并将SPP和PMT嵌入到网络中来提取数据中更复杂的特征信息。此外，针对AlexNet做出改进不仅提升了检索性能、减少了模型参数、节省了存储空间、加快了模型训练，还能有效用于极光图像分类。在CIFAR-10上与主流的哈希算法对比中，本文方法实现了最佳的检索性能，进一步验证所提方法的扩展性。同时也有两点需要改进，一方面极光数据集偏小，且类别之间存在不均衡，因此需要针对小数据集以及不均衡的问题使用更合适的模型和一些策略来改进。另一方面由于极光形态结构复杂且是动态变化的，尤其是有些热点状极光图像在视觉上可能是几种类型的混合形态，所以在进行特征提取时应结合极光特有的形态以解决热点状极光容易被错分的问题。

参考文献

[1] WANG Qian, LIANG Jimin, HU Zejun, *et al.* Spatial texture based automatic classification of dayside aurora in all-sky images[J]. *Journal of Atmospheric and Solar-Terrestrial Physics*, 2010, 72(5/6): 498–508. doi:

10.1016/j.jastp.2010.01.011.

- [2] 韩冰, 杨辰, 高新波. 融合显著信息的LDA极光图像分类[J]. *软件学报*, 2013, 24(11): 2758–2766. doi: 10.3724/SP.J.1001.2013.04481.
- HAN Bing, YANG Chen, and GAO Xinbo. Aurora image classification based on LDA combining with saliency information[J]. *Journal of Software*, 2013, 24(11): 2758–2766. doi: 10.3724/SP.J.1001.2013.04481.
- [3] SYRJÄSUO M T, DONOVAN E F, and COGGER L L. Content-based retrieval of auroral images - thousands of irregular shapes[C]. *The 4th IASTED International Conference Visualization, Imaging, and Image Processing*, Marbella, Spain, 2004.
- [4] FU Rong, GAO Xinbo, LI Xuelong, *et al.* An integrated aurora image retrieval system: Aurora Eye[J]. *Journal of Visual Communication and Image Representation*, 2010, 21(8): 787–797. doi: 10.1016/j.jvcir.2010.06.002.
- [5] YANG Xi, GAO Xinbo, SONG Bin, *et al.* Aurora image search with contextual CNN feature[J]. *Neurocomputing*, 2018, 281: 67–77. doi: 10.1016/j.neucom.2017.11.059.
- [6] 葛芸, 马琳, 江顺亮, 等. 基于高层特征图组合及池化的高分辨率遥感图像检索[J]. *电子与信息学报*, 2019, 41(10): 2487–2494. doi: 10.11999/JEIT190017.
- GE Yun, MA Lin, JIANG Shunliang, *et al.* The combination and pooling based on high-level feature map for high-resolution remote sensing image retrieval[J]. *Journal of Electronics & Information Technology*, 2019, 41(10): 2487–2494. doi: 10.11999/JEIT190017.
- [7] 刘治, 潘炎, 夏榕楷, 等. FP-CNNH: 一种基于深度卷积神经网络的快速图像哈希算法[J]. *计算机科学*, 2016, 43(9): 39–46, 51. doi: 10.11896/j.issn.1002-137X.2016.09.007.
- LIU Ye, PAN Yan, XIA Rongkai, *et al.* FP-CNNH: A fast image hashing algorithm based on deep convolutional neural network[J]. *Computer Science*, 2016, 43(9): 39–46, 51. doi: 10.11896/j.issn.1002-137X.2016.09.007.

- [8] LI Wujun, WANG Sheng, and KANG Wangcheng. Feature learning based deep supervised hashing with pairwise labels[C]. The 25th International Joint Conference on Artificial Intelligence, New York, USA, 2016: 1711–1717.
- [9] LIU Haomiao, WANG Ruiping, SHAN Shiguang, *et al.* Deep supervised hashing for fast image retrieval[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 2064–2072. doi: [10.1109/CVPR.2016.227](https://doi.org/10.1109/CVPR.2016.227).
- [10] KRIZHEVSKY A, SUTSKEVER I, and HINTON G E. Imagenet classification with deep convolutional neural networks[C]. The 25th International Conference on Neural Information Processing Systems, Lake Tahoe, USA, 2012: 1097–1105.
- [11] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, *et al.* Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1904–1916. doi: [10.1109/TPAMI.2015.2389824](https://doi.org/10.1109/TPAMI.2015.2389824).
- [12] 赵斐, 张文凯, 闫志远, 等. 基于多特征图金字塔融合深度网络的遥感图像语义分割[J]. *电子与信息学报*, 2019, 41(10): 2525–2531. doi: [10.11999/JEIT190047](https://doi.org/10.11999/JEIT190047).
ZHAO Fei, ZHANG Wenkai, YAN Zhiyuan, *et al.* Multi-feature map pyramid fusion deep network for semantic segmentation on remote sensing data[J]. *Journal of Electronics & Information Technology*, 2019, 41(10): 2525–2531. doi: [10.11999/JEIT190047](https://doi.org/10.11999/JEIT190047).
- [13] ZHANG Chenlin and WU Jianxin. Improving CNN linear layers with power mean non-linearity[J]. *Pattern Recognition*, 2019, 89: 12–21. doi: [10.1016/j.patcog.2018.12.029](https://doi.org/10.1016/j.patcog.2018.12.029).
- [14] JEGOU H, DOUZE M, and SCHMID C. Hamming embedding and weak geometric consistency for large scale image search[C]. The 10th European Conference on Computer Vision, Marseille, France, 2008: 304–317. doi: [10.1007/978-3-540-88682-2_24](https://doi.org/10.1007/978-3-540-88682-2_24).
- [15] XIA Yan, HE Kaiming, WEN Fang, *et al.* Joint inverted indexing[C]. 2013 IEEE International Conference on Computer Vision, Sydney, Australia, 2013: 3416–3423. doi: [10.1109/ICCV.2013.424](https://doi.org/10.1109/ICCV.2013.424).
- [16] TOLIAS G, SICRE R, and JÉGOU H. Particular object retrieval with integral max-pooling of CNN activations[C]. The 4th International Conference on Learning Representations, San Juan, Puerto Rico, 2016: 1–12.
- [17] LI Yang, XU Yulong, WANG Jiabao, *et al.* MS-RMAC: Multiscale regional maximum activation of convolutions for image retrieval[J]. *IEEE Signal Processing Letters*, 2017, 24(5): 609–613. doi: [10.1109/LSP.2017.2665522](https://doi.org/10.1109/LSP.2017.2665522).
- [18] DATAR M, IMMORLICA N, INDYK P, *et al.* Locality-sensitive hashing scheme based on p-stable distributions[C]. The 20th Annual Symposium on Computational Geometry, Brooklyn, USA, 2004: 253–262. doi: [10.1145/997817.997857](https://doi.org/10.1145/997817.997857).
- [19] GONG Yunchao and LAZEBNIK S. Iterative quantization: A procrustean approach to learning binary codes[C]. The 24th IEEE Conference on Computer Vision and Pattern Recognition, Providence, USA, 2011: 817–824. doi: [10.1109/CVPR.2011.5995432](https://doi.org/10.1109/CVPR.2011.5995432).
- [20] LIU Wei, WANG Jun, JI Rongrong, *et al.* Supervised hashing with kernels[C]. 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, USA, 2012: 2074–2081. doi: [10.1109/CVPR.2012.6247912](https://doi.org/10.1109/CVPR.2012.6247912).
- 陈昌红: 女, 1982年生, 副教授, 研究方向为智能视频分析、模式识别。
- 彭腾飞: 男, 1994年生, 硕士生, 研究方向为图像处理与图像通信。
- 干宗良: 男, 1978年生, 副教授, 研究方向为分布式视频编码、图像信号视频处理。

责任编辑: 余蓉