



文章编号:1001-9081(2018)07-1853-04

DOI:10.11772/j.issn.1001-9081.2017122920

## 基于多特征组合的细粒度图像分类方法

邹承明<sup>1,2</sup>, 罗莹<sup>1,2\*</sup>, 徐晓龙<sup>1,2</sup>

(1. 交通物联网技术湖北省重点实验室(武汉理工大学), 武汉 430070; 2. 武汉理工大学 计算机科学与技术学院, 武汉 430070)

(\*通信作者电子邮箱 luoyingwork@foxmail.com)

**摘要:**针对单一特征表示的局限性会导致细粒度图像分类准确度不高的问题,提出了一种基于卷积神经网络(CNN)和尺度不变特征转换(SIFT)的多特征组合表示方法,综合考虑对目标整体、关键部位和关键点的特征提取。首先,分别以细粒度图像库中的目标整体和头部区域训练CNN得到两个网络模型,用来提取目标的整体和头部CNN特征;然后,对图像库中所有目标区域提取SIFT关键点并通过K均值(K-means)聚类生成码本,再将每个目标区域的SIFT描述子通过局部特征聚合描述符(VLAD)参照码本编码为特征向量;最后,组合多种特征作为最终的特征表示,采用支持向量机(SVM)对细粒度图像进行分类。使用该方法在CUB-200-2011数据库上进行实验,并与单一的特征表示方法进行了比较。实验结果表明,该方法与基于单一CNN特征的细粒度图像分类相比提升了13.31%的准确度,证明了多特征组合对细粒度图像分类的积极作用。

**关键词:**卷积神经网络;尺度不变特征转换;K均值聚类;局部特征聚合描述符;细粒度图像分类

**中图分类号:** TP391.413; TP18    **文献标志码:**A

### Fine-grained image classification method based on multi-feature combination

ZOU Chengming<sup>1,2</sup>, LUO Ying<sup>1,2\*</sup>, XU Xiaolong<sup>1,2</sup>

(1. Hubei Key Laboratory of Transportation Internet of Things (Wuhan University of Technology), Wuhan Hubei 430070, China;

2. College of Computer Science and Technology, Wuhan University of Technology, Wuhan Hubei 430070, China)

**Abstract:** As the limitation of single feature representation may cause low accuracy of fine-grained image classification, a multi-feature combination representation method based on Convolutional Neural Network (CNN) and Scale Invariant Feature Transform (SIFT) was proposed. The features were extracted from the entire target, the key parts and the key points comprehensively. Firstly, two CNN models were trained with the target-entirety regions and the head-only regions in the fine-grained image library respectively, which were used to extract the target-entirety and the head-only CNN features. Secondly, the SIFT key points were extracted from all the target-entirety regions in the image library, and the codebook was generated through the K-means clustering. Then, the SIFT descriptors of each target-entirety region were encoded into a feature vector by using the Vector of Locally Aggregated Descriptors (VLAD) along with the codebook. Finally, Support Vector Machine (SVM) was used to classify the fine-grained images by using the combination of multiple features. The method was evaluated in CUB-200-2011 database and compared with the single feature representation method. The experimental results show that the proposed method can improve the classification accuracy by 13.31% compared with the single CNN feature representation, which proves the positive effect of multi-feature combination on fine-grained image classification.

**Key words:** Convolutional Neural Network (CNN); Scale Invariant Feature Transform (SIFT); K-means clustering; Vector of Locally Aggregated Descriptors (VLAD); fine-grained image classification

### 0 引言

一直以来,细粒度图像分类都是计算机视觉和模式识别等领域的许多研究人员感兴趣的课题,但由于细粒度图像往往拥有较大的类内差异和细微的类间差异,其分类难度远远高于普通的图像分类<sup>[1]</sup>。

细粒度图像的样本通常属于同一大类,不同子类间拥有较高的相似性,相互的区别往往体现在难以察觉的局部细节之上,而同一子类下的样本又会由于拍摄距离、目标姿势、复杂背景和遮挡物等因素的干扰产生极大的差异,彼此的共性可能存在旋转、缩放、变形和缺失的情况,因此细粒度图像分

类问题成为了机器视觉领域一个极大的挑战<sup>[2]</sup>。

早期的细粒度分类方法使用简单的特征提取和编码方式生成特征表示,这些表示方法相当粗糙,采集到的有效信息十分有限。比如Wah等<sup>[3]</sup>在细粒度数据集CUB-200-2011上作基准测试时,通过对图像提取RGB颜色直方图特征和尺度不变特征转换(Scale Invariant Feature Transform, SIFT)<sup>[4]</sup>特征,运用词袋模型(Bag of Words, BoW)生成特征编码,最终分类精度最高仅为17.3%。这项工作表明,简单的特征表示方法描述性较弱,不足以支撑更高的细粒度图像分类准确度。但这并不意味着早期提取的特征是完全无用的,随着更为强大的编码方式的出现,如Fisher Vector编码<sup>[5]</sup>、局部特征聚合描

收稿日期:2017-12-13;修回日期:2018-02-08;录用日期:2018-03-16。    基金项目:中央高校基本科研业务费专项(2017-zy-084)。

**作者简介:**邹承明(1975—),男,广东徐闻人,教授,博士,CCF会员,主要研究方向:计算机视觉、嵌入式系统、软件理论与方法;罗莹(1993—),女,湖南益阳人,硕士研究生,主要研究方向:图形图像处理;徐晓龙(1995—),男,安徽宿州人,硕士研究生,主要研究方向:图形图像处理。



述符(Vector of Locally Aggregated Descriptors, VLAD)<sup>[6]</sup>等,与之结合的SIFT特征的分类准确度也有所提高。

近年来卷积神经网络( Convolutional Neural Network, CNN)<sup>[7-10]</sup>在一般的图像分类任务中取得了令人瞩目的成绩,证明了卷积特征对图像的描述能力,这一结论也为细粒度图像分类带来了新的发展方向,研究人员开始选择CNN特征作为图像表示用于细粒度图像分类<sup>[11]</sup>。Zhang等<sup>[12]</sup>依次对目标整体、头部和身体区域训练CNN模型分别用来提取三种卷积特征,最后将三种特征级联成一个向量用于分类;Branson等<sup>[13]</sup>对目标整体和不同部位分别提取不同层的CNN特征,综合多项卷积特征进行分类。上述算法在CUB-200-2011数据集上都达到了75%以上的分类准确度。这些工作证明了卷积特征在细粒度图像分类中可以发挥较好的作用,但是分类准确度仍然存在上升空间。

从现有研究成果来看,目前细粒度图像分类的主流框架大多是基于单一特征,这往往导致一定的片面性。以SIFT和CNN特征为例,前者注重图像底层信息,针对关键点提取的特征太过细化且极易被噪声干扰,自身无法支撑较高的全局分类准确度;后者拥有强大的抽象能力,能够捕捉图像的高层语义,可针对目标整体区域提取的全局特征却过于泛化,忽视了底层和局部的细节,不利于区分细粒度图像库中十分相似的异类样本。针对这些问题,本文提出的多特征组合方法综合考虑高层和底层、全局和局部的信息,采用针对目标整体、关键部位和关键点层层递进的特征提取方法,将各个层次的特征组合起来,得到更丰富的特征表示。本文算法在CUB-200-2011鸟类数据集上进行实验,结果表明,本文提出的特征表示方法对于细粒度图像分类准确度的提升有积极作用。

## 1 特征提取模型

图像特征提取是解决图像分类问题的重要步骤,特征提取方法的优劣决定了所提取特征的描述能力,从而影响图像分类性能。细粒度分类由于其分类目标更加细致,对图像特征的要求也更加苛刻,往往需要在描述全局特征的同时不忽视局部特征,因此,一个综合考虑全局特征和局部特征的特征提取模型对细粒度分类尤为重要。

CNN是LeCun等<sup>[7,14]</sup>提出的一种经典的神经网络特征提取模型,它可以直接处理原始图像,还能够根据不同的分类问题从图像中提取出相应有效的特征,目前在计算机视觉领域中应用广泛,其基本的网络结构如图1所示。

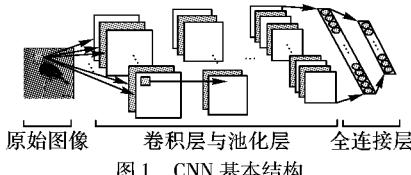


Fig. 1 Basic structure of CNN

网络整体可分为三个部分:输入层、中间层和输出层<sup>[15]</sup>。输入层为原始图像,输出层为特征向量,中间层由若干个卷积层和池化层交替组成,其作用是对输入层的数据进行逐层抽象,最终通过全连接得到含有高层语义的输出向量。对于CNN而言,不同的卷积参数和池化策略将导致网络最终所提取特征的不同意义,而多样化的特征又伴随着强大的描述能力和迁移能力,因此,卷积层和池化层是CNN提取更具区分度的特征的关键所在。

其中,卷积层由大小固定的卷积核对上层的输出进行卷积操作,其核心原理如图2所示:图2(a)展示了多种卷积核的卷积操作,每个卷积核都拥有相同尺寸的局部感受野,感受野内部是各自的权值参数,参数值的差异决定了卷积核的不同种类,卷积核依次沿着x轴和y轴方向在图层上进行滑动和卷积,得到的值组成该卷积核的特征响应图;图2(b)是卷积的计算过程,一个卷积核的内部权值参数对于它要卷积的对象是共享的,也就是它在滑动和卷积过程中不会发生权值的变化,而卷积就是将感受野所覆盖的图层数据与自身相应位置的权值进行加权求和,所得结果作为输出数值。

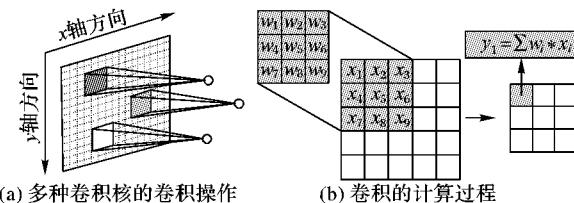


图2 卷积层核心原理

Fig. 2 Core principle of convolutional layer

池化层操作原理与卷积层类似,区别在于池化层的感受野会根据不同池化策略来取对应位置的最大值或者平均值。

从CNN的基本结构来看,它所提取的特征属于全局属性,而通过CNN中卷积和池化的操作过程也不难看出,随着中间层数的增加,CNN将原始图像的底层信息逐步合并和归纳,进而向高层过渡,但一些局部细节也随之逐层丢失,因此CNN虽然在一般的分类任务中出类拔萃,却在细粒度分类中表现欠佳。受这一现象的启发,本文提出了利用SIFT关键点来增强局部点特征的方法,获得更强大的特征描述力。

SIFT是一种局部特征描述子,它可以定位和筛选图像尺度空间中的极值点,并用128维的向量描述每个关键点,这种特征具有旋转不变性和尺度不变性,是极为可靠的局部描述符。然而,不同图像的SIFT关键点数目往往差异较大,因此SIFT特征不便于直接用于图像分类,此时就需要特定的编码方式在考虑到每张图像SIFT特征的同时,生成统一维度的特征向量。VLAD就是极为优秀的一种特征编码方式,它相对于BoW编码而言损失较少的信息,相对于Fisher Vector编码而言计算更为简单,而且在一些分类任务中它的效果还优于前两者。它保存了每个SIFT特征点到离它最近的聚类中心的距离,且考虑了特征点每一维的值的影响,对图像局部内容有更细致的刻画,因此本文选择用VLAD模型编码SIFT特征,获取的信息用来强化其他全局特征的局部描述性。

综上,本文所使用的特征提取模型兼顾了CNN和SIFT两种特征的提取,以CNN提取的高层语义特征把控全局方向,以VLAD算法编码的SIFT关键点特征补充局部细节,得到对细粒度图像目标更具判别力的特征表示。

## 2 多特征组合

由第1章CNN的原理介绍可知,CNN特征提取模型针对的是图像块,它所输出的特征是将图像块中所有像素以某种方式进行融合后的产物。这样的特征充分反映了该图像块的全局属性,然而图像块的尺度是可变的,也就是说同样大小的图像块,它的内容可能是整个目标,也可能只是目标的一部分,例如鸟类的整体和头部区域本质都是图像块,都可以通过CNN模型进行特征提取,因此CNN模型本身也可以提取出多种特征,这些特征与VLAD模型提取的SIFT点特征组合在一



起可以构建出丰富的特征表示,不同成分的组合特征在细粒度图像分类中也有着不同的结果,因此即使已经确定了多特征提取的模型,各种特征的选取也至关重要。

本文结合人眼分辨细粒度图像的经验,对 CNN 模型和 VLAD 模型提取的多种特征的选择和组合策略进行了分析。在面对需要辨别的细粒度图像时,人们通常会先根据目标整体区域的信息判断大致类别,这一过程说明了目标整体区域的特征的必要性,这类特征可通过 CNN 模型提取;接下来就是根据目标关键部位和关键点的细节差异来区分细粒度类别,这一过程说明了目标关键部位区域和关键点的特征对细粒度图像最终辨识结果的重要性,这两类特征可分别通过 CNN 和 VLAD 模型提取,因此,本文提出的多特征组合方法同时选取目标整体、关键部位和关键点三重特征层层递进组合成对细粒度图像具有更高描述性和区分性的最终特征。

图 3 为相应的多特征组合结构示意图:分别用各自的模型提取目标整体、关键部位和关键点特征,三者级联之后用 SVM 分类器进行分类。其中关键部位取用的是目标的头部区域,原因有二:其一是头部区域往往蕴含大量细粒度图像分类所需信息;其二是图像样本中头部区域的完整性最能得到保障。

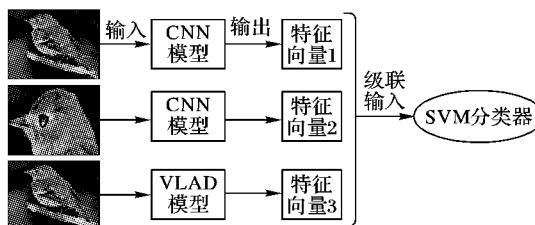


图 3 多特征组合结构

Fig. 3 Multi-feature combination structure

在实现本文提出的多特征组合方法时,需要提前对三种特征提取模型进行训练,图 4 为本文细粒度图像分类模型的训练过程示意图。首先,分别以细粒度图像库训练集中的目标整体和头部区域训练 CNN 得到两个网络模型  $M_1$  和  $M_2$ ,提取训练集中所有目标的整体和头部 CNN 特征  $F_1$  和  $F_2$ ;然后,对训练集中所有目标区域提取 SIFT 特征,通过 K 均值聚类生成聚类中心集合作为码本  $C$ ,再将每个目标区域的 SIFT 特征点通过 VLAD 参照码本  $C$  编码为特征  $F_3$ ;最后,组合  $F_1$ 、 $F_2$  和  $F_3$  作为最终的特征表示,训练 SVM 分类器  $S$ 。

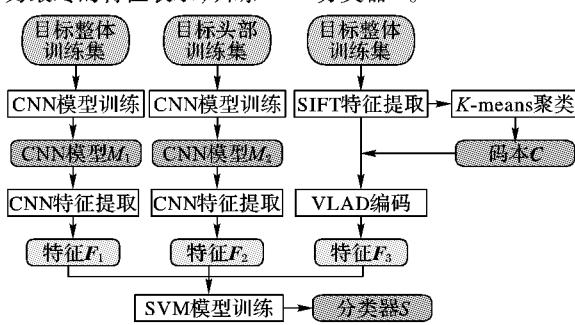


图 4 模型训练过程

Fig. 4 Training process of models

模型的测试过程如图 5 所示:对测试目标整体和头部分别用  $M_1$  和  $M_2$  提取特征  $F_1'$  和  $F_2'$ ,对测试目标整体提取 SIFT 特征点,再参照码本  $C$  进行 VLAD 编码,得到特征  $F_3'$ ,将  $F_1'$ 、 $F_2'$  和  $F_3'$  级联,输入到分类器  $S$  中得到结果。

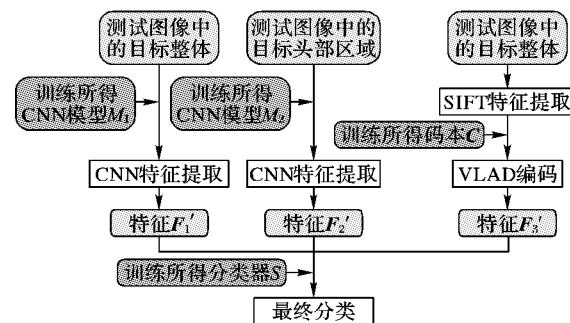


图 5 模型测试过程

Fig. 5 Testing process of models

### 3 实验结果与分析

#### 3.1 实验设置

本文在细粒度图像分类领域中最为经典和常用的数据集 CUB-200-2011 上进行实验,该数据集包含 200 种不同鸟类的类别,共计 11788 张图像,实验选取数据集中的 5994 张作为训练集,其他 5794 张作为测试集。实验主要硬件环境为 Nvidia Tesla P100 计算显卡,软件环境为 Linux 系统下 Caffe 和 Matlab R2014a。

为了便于目标整体和关键部位区域 CNN 特征的统一提取,将所有图像数据的尺寸预处理为  $227 \times 227$ 。实验中 CNN 结构设置如下:1~5 层为卷积层,每层卷积后都有 ReLU 激活函数处理和最大值池化操作;6~7 层为全连接层,本文算法取用的 CNN 特征就是第 7 层的 4096 维特征。

为了得出本文算法中 VLAD 特征编码时最适合的  $k$  均值聚类中心数,分别在 8、16、32、64 四种  $k$  值下进行实验,表 1 为在细粒度数据集 CUB-200-2011 上使用本文算法时选用不同  $k$  值所对应的分类准确度。

表 1 不同聚类数下的分类准确度

Tab. 1 Classification accuracy of different clustering numbers

聚类数	分类准确度/%	聚类数	分类准确度/%
8	81.41	32	81.53
16	81.60	64	81.46

从表 1 可以看出,当聚类数等于 16 时,细粒度图像识别准确度达到了最高值。分析其原因,当聚类数过小时,聚类中心之间相距较远,每个中心辐射的聚类范围较广,各个聚类的关键点分布也较为分散,也就是聚类很可能并没有达到较好的收敛状态,因此凭借这个聚类中心集合作为码本提取的信息没有足够的可靠性;而聚类数过大时,聚类中心较为拥挤,每个中心辐射的聚类范围重叠,关键点可能同时归入多个聚类,造成某些特征的重复提取,从而导致过拟合,因此,本文算法提取 VLAD 特征时的关键点聚类数设置为 16,相应的特征维度为  $16 \times 128$ 。

上述两种特征确定后,将它们进行级联组合后,选择一对多方式下的多类 SVM 分类器对数据进行训练和测试。

#### 3.2 实验结果

本文算法结合 CNN 特征和 VLAD 编码的 SIFT 特征,使用“目标整体 + 关键区域 + 关键点”的特征提取和组合方法,最后通过一对多 SVM 进行分类。

为了方便对比分析,在进行本文提出的三重特征组合的细粒度分类实验的同时,也分别在同样的数据集上进行了多种单一 CNN 特征的分类和多种组合 CNN 特征的分类。



实验结果如表 2 所示, 表中字母的含义如下:  $F_{\text{entirety-cnn}}$ 、 $F_{\text{head-cnn}}$  和  $F_{\text{body-cnn}}$  分别表示在目标整体、头部和身体区域提取的 CNN 特征,  $F_{\text{entirety-cnn}} + F_{\text{head-cnn}}$ 、 $F_{\text{entirety-cnn}} + F_{\text{body-cnn}}$  和  $F_{\text{head-cnn}} + F_{\text{body-cnn}}$  分别表示三种 CNN 特征的两两组合,  $F_{\text{entirety-cnn}} + F_{\text{head-cnn}} + F_{(16 \times 128)-\text{vlad}}$  则表示目标整体 CNN 特征、头部 CNN 特征和以 16 个聚类中心为码本编码的 VLAD 特征的组合表示。

表 2 不同特征的实验结果对比

Tab. 2 Experimental results of different features

特征	特征组成	测试数据		分类准确度/%
		正确数	总数	
单一 CNN 特征	$F_{\text{entirety-cnn}}$	3 957	5 794	68.29
	$F_{\text{head-cnn}}$	4 073	5 794	70.30
	$F_{\text{body-cnn}}$	3 476	5 794	60.00
组合 CNN 特征	$F_{\text{entirety-cnn}} + F_{\text{head-cnn}}$	4 707	5 794	81.24
	$F_{\text{entirety-cnn}} + F_{\text{body-cnn}}$	3 666	5 794	63.27
	$F_{\text{head-cnn}} + F_{\text{body-cnn}}$	4 667	5 794	80.55
组合 CNN 和 VLAD 特征	$F_{\text{entirety-cnn}} + F_{\text{head-cnn}} + F_{(16 \times 128)-\text{vlad}}$	4 728	5 794	81.60

表 2 中分别记录了上述特征表示方法所对应的细粒度图像分类准确度, 分析其中数据可以得出以下几个结论:

1) 单一的 CNN 特征本身有较强的描述性, 但在细粒度分类问题中还缺乏足够的区分度, 因此其分类准确度只能停留在 60% ~ 70%。

2) 三种单一 CNN 特征中, 头部区域的特征区分度明显高于其他两种, 凡是有这一特征参与的特征组合都可以获得 80% 以上的分类准确度, 这种结果可以证明头部是该数据集的关键部位, 也可以证明关键部位的特征对细粒度分类至关重要。

3) 多特征组合表示的分类准确度大多高于单特征表示, 且优势明显, 除了  $F_{\text{entirety-cnn}} + F_{\text{body-cnn}}$  组合这一特例, 其他组合方法相比单特征表示都有 10% ~ 20% 的准确度提升。而  $F_{\text{entirety-cnn}} + F_{\text{body-cnn}}$  组合的准确度较低的原因除了目标整体区域和身体区域的单一 CNN 特征都不够强大之外, 还存在一些外部因素。结合实验所用的图像数据进行分析可以发现, CUB 数据集中存在一定量的样本完全缺失身体区域, 这种现象更加降低了这类特征组合的有效性, 从而导致分类的失败。

4) 结合目标整体、头部 CNN 特征和关键点 VLAD 特征的表示方法取得了最高的细粒度分类准确度 81.6%, 相较于直接将 CNN 结构套用到细粒度图像分类问题上, 凭借单一的目标整体区域 CNN 特征进行分类, 提高了 13.31% 的准确度, 证明了本文所提出多特征组合的有效性。

综上所述, 本文提出的综合考虑目标整体、关键部位和关键点的多特征组合方法对细粒度图像分类准确度有着明显的提升。

此外, 本文将所提出的方法和其他已有的细粒度图像分类算法在同一个数据集 CUB-200-2011 下进行了实验对比, 结果如表 3 所示, 从分类准确度对比结果可以看出本文提出的方法对细粒度图像分类的确有较好的效果。

## 4 结语

为了解决单一特征表示的局限性会导致细粒度图像分类准确度不高的问题, 本文提出了一种基于 CNN 和 SIFT 的多

特征组合表示方法。基于单一 CNN 特征和多特征组合的细粒度图像分类对比实验表明, 综合考虑对目标整体、关键部位和关键点的特征提取的多特征组合方法拥有最高的分类准确度。由于关键点定位和筛选算法的误差, 导致关键点特征对组合特征的提升不是很明显, 在下一步的研究过程中, 可以针对这一问题进行改进。

表 3 CUB-200-2011 下的各方法分类准确度对比 %

Tab. 3 Classification accuracy contrast of different methods in CUB-200-2011 %

方法	分类准确度	方法	分类准确度
Zhang 等方法 <sup>[12]</sup>	76.4	本文方法	81.6
Branson 等方法 <sup>[13]</sup>	75.7		

## 参考文献 (References)

- [1] 罗建豪, 吴建鑫. 基于深度卷积特征的细粒度图像分类研究综述 [J]. 自动化学报, 2017, 43(8): 1306–1318. (LUO J H, WU J X. A survey on fine-grained image categorization using deep convolutional features [J]. Acta Automatica Sinica, 2017, 43(8): 1306–1318.)
- [2] 冯语姗, 王子磊. 自上而下注意图分割的细粒度图像分类 [J]. 中国图象图形学报, 2016, 21(9): 1147–1154. (FENG Y S, WANG Z L. Fine-grained image categorization with segmentation based on top-down attention map [J]. Journal of Image and Graphics, 2016, 21(9): 1147–1154.)
- [3] WAH C, BRANSON S, WELINDER P, et al. The Caltech-UCSD Birds-200-2011 dataset, CNS-TR-2011-001 [R]. Pasadena, CA: California Institute of Technology, 2011.
- [4] LOWE D G. Object recognition from local scale-invariant features [C]// ICCV 1999: Proceedings of the 7th IEEE International Conference on Computer Vision. Piscataway, NJ: IEEE, 1999: 1150.
- [5] SANCHEZ J, PERRONNIN F, MENSINK T. Image classification with the fisher vector: theory and practice [J]. International Journal of Computer Vision, 2013, 105(3): 222–245.
- [6] JEGOU H, DOUZE M, SCHMID C, et al. Aggregating local descriptors into a compact image representation [C]// CVPR 2010: Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2010: 3304–3311.
- [7] LECUN Y, BOTTOU L, BENGIO Y. Gradient-based learning applied to document recognition [J]. Proceeding of the IEEE, 1998, 86(11): 2278–2324.
- [8] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. Advances in Neural Information Processing Systems, 2012, 25: 1106–1114.
- [9] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [C/OL]. ICLR 2015: Proceedings of the 2015 International Conference on Learning Representations. San Diego, CA. [2017-09-12]. <https://arxiv.org/abs/1409.1556>.
- [10] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions [C]// CVPR 2015: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2015: 1–9.
- [11] ZHAO B, FENG J, WU X, et al. A survey on deep learning-based fine-grained object classification and semantic segmentation [J]. International Journal of Automation and Computing, 2017, 14(2): 119–135.

(下转第 1861 页)



分别提高了 5.5% 和 6%。这些实验表明, 特征之间的交互作用能提升特征选择的性能。

未来的研究包括使用并行算法降低 Max-Interaction 的计算复杂度, 以及使用大数据集为信息测度提供更好的估计。

#### 参考文献 (References)

- [1] CAI D, HE X. Manifold adaptive experimental design for text categorization [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2012, 24(4): 707–719.
- [2] 张海龙, 王莲芝. 自动文本分类特征选择方法研究 [J]. *计算机工程与设计*, 2006, 27(20): 3840–3841. (ZHANG H L, WANG L Z. Automatic text categorization feature selection methods research [J]. *Computer Engineering and Design*, 2006, 27(20): 3840–3841.)
- [3] LIU J, SHANG J, HAN J. Phrase Mining from Massive Text and Its Applications [M]. San Rafael, CA: Morgan & Claypool Publishers, 2017: 1–89.
- [4] BROWN G, POCOCK A, ZHAO M J, et al. Conditional likelihood maximisation: a unifying framework for information theoretic feature selection [J]. *Journal of Machine Learning Research*, 2012, 13(1): 27–66.
- [5] YANG Y, PEDERSEN J O. A comparative study on feature selection in text categorization [C]// ICML 1997: Proceedings of the 1997 International Conference on Machine Learning. San Francisco: Morgan Kaufmann, 1997: 412–420.
- [6] TANG B, KAY S, HE H. Toward optimal feature selection in naive Bayes for text categorization [J]. *IEEE Transactions on Knowledge and Data Engineering*, 2016, 28(9): 2508–2521.
- [7] WANG D, ZHANG H, LIU R, et al. Feature selection based on term frequency and t-test for text categorization [C]// CIKM 2012: Proceedings of the 21st International Conference on Information and Knowledge Management. New York: ACM, 2012: 1472–1486.
- [8] 辛竹, 周亚建. 文本分类中互信息特征选择方法的研究与算法改进 [J]. *计算机应用*, 2013, 33(S2): 116–118. (XIN Z, ZHOU Y J. Study and improvement of mutual information for feature selection in text categorization [J]. *Journal of Computer Applications*, 2013, 33(S2): 116–118.)
- [9] VINH N X, ZHOU S, CHAN J, et al. Can high-order dependencies improve mutual information based feature selection? [J]. *Pattern Recognition*, 2016, 53: 46–58.
- [10] BENNASAR M, HICKS Y, SETCHI R. Feature selection using joint mutual information maximization [J]. *Expert Systems with Applications*, 2015, 42(22): 8520–8532.
- [11] ZENG Z, ZHANG H, ZHANG R, et al. A novel feature selection method considering feature interaction [J]. *Pattern Recognition*, 2015, 48(8): 2656–2666.
- [12] JAKULIN A. Machine learning based on attribute interactions [D]. Ljubljana: University of Ljubljana, 2005: 37–38.
- [13] BALACANI K S, PHOHA V V. On the feature selection criterion based on an approximation of multidimensional mutual information [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(7): 1342–1343.
- [14] HAGAR J D, WISSINK T L, KUHN D R, et al. Introducing combinatorial testing in a large organization [J]. *Computer*, 2015, 48(4): 64–72.
- [15] MONTGOMERY D C. Design and Analysis of Experiments [M]. 9th ed. Hoboken, NJ: John Wiley & Sons, 2017: 179–220.
- [16] SHISHKIN A, BEZZUBTSEVA A, DRUTSA A, et al. Efficient high-order interaction-aware feature selection based on conditional mutual information [C]// NIPS 2016: Proceedings of the 30th Annual Conference on Neural Information Processing Systems. Red Hook, NY: Curran Associates, 2016: 4637–4645.
- [17] KLEEREKOPER A, PAPPAS M, POCOCK A, et al. A scalable implementation of information theoretic feature selection for high dimensional data [C]// IEEE BigData 2015: Proceedings of the 2015 IEEE International Conference on Big Data. Piscataway, NJ: IEEE, 2015: 339–346.
- [18] RAMIREZ-GALLEGOS S, MOURIO-TALN H, MARTINEZ-REGO D, et al. An information theory-based feature selection framework for big data under apache spark [J/OL]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2017: 1–13 [2018-01-15]. <http://ieeexplore.ieee.org/document/7970198/>.
- [19] LI J, CHENG K, WANG S, et al. Feature selection: a data perspective [J]. *ACM Computing Surveys*, 2018, 50(6): Article No. 94.

This work is partially supported by the National Natural Science Foundation of China (61602094).

**TANG Xiaochuan**, born in 1986, Ph. D. candidate. His research interests include feature selection, machine learning, big data analysis.

**QIU Xiwei**, born in 1980, Ph. D. His research interests include cloud computing, big data, green computing.

**LUO Liang**, born in 1980, Ph. D., lecturer. His research interests include cloud computing, big data, energy consumption modeling.

(上接第 1856 页)

- [12] ZHANG N, DONAHUE J, GIRSHICK R, et al. Part-based R-CNNs for fine-grained category detection [C]// ECCV 2015: Proceedings of the 2015 European Conference on Computer Vision. Piscataway, NJ: IEEE, 2015: 1143–1151.
- [13] BRANSON S, VAN HORN G, BELONGIE S, et al. Bird species categorization using pose normalized deep convolutional nets [C/OL]. BMVC 2014: Proceedings of the 2014 British Machine Vision Conference. Nottingham, UK. [2017-09-15]. <https://arxiv.org/abs/1406.2952>.
- [14] LECUN Y, BOSER B, DENKER J S. Back propagation applied to handwritten zip code recognition [J]. *Neural Computation*, 1989, 1(4): 541–551.
- [15] 李彦冬, 郝宗波, 雷航. 卷积神经网络研究综述 [J]. *计算机应*

用, 2016, 36(9): 2508–2515. (LI Y D, HAO Z B, LEI H. Survey of convolutional neural network [J]. *Journal of Computer Applications*, 2016, 36(9): 2508–2515.)

This work is partially supported by the Fundamental Research Funds for the Central Universities (2017-zy-084).

**ZOU Chengming**, born in 1975, Ph. D., professor. His research interests include computer vision, embedded system, software theory and methods.

**LUO Ying**, born in 1993, M. S. candidate. Her research interests include graphic and image processing.

**XU Xiaolong**, born in 1995, M. S. candidate. His research interests include graphic and image processing.