

基于姿态的判别属性学习及在细粒度识别中的应用

宋凤义, 张守东, 杨 明

(南京师范大学计算机科学与技术学院, 江苏 南京 210023)

[摘要] 姿态变化造成同一对象或同类对象的视觉信息差异巨大, 成为计算机视觉中对象识别的一大挑战因素. 属性表示重在刻画较高的抽象语义特性, 具有应对包括姿态变化的复杂环境变化的鲁棒性, 但也给属性学习自身带来了较大难度. 如何降低属性学习的难度同时提高属性表示的判别力, 成为基于属性表示的识别模型的关键, 尤其面临对判别属性要求较高的细粒度识别任务. 显式地对姿态建模, 在不同姿态下学习能够最大化类别间隔的视觉判别属性, 最终作为中间表示用于类别识别. 最后, 在细粒度公开数据集 CUB 上验证了所提出的基于姿态的判别属性在细粒度识别任务中的有效性.

[关键词] 属性学习, 判别属性, 分散式表示, 细粒度识别

[中图分类号] TP391.4 [文献标志码] A [文章编号] 1001-4616(2017)01-0065-08

Pose-Based Discriminative-Attributes Learning for Fine-Grained Recognition

Song Fengyi, Zhang Shoudong, Yang Ming

(School of Computer Science and Technology, Nanjing Normal University, Nanjing 210023, China)

Abstract: Commonly existed various posture of object makes great challenges for object recognition in computer vision literature. Attribute representation shows robust describable ability with clear semantic meaning invariant to changes of environment factors including posture. However, the inherent description advantages of attributes also result big challenges for itself to learn well worked attribute predictor. Consequently, the key issues in attribute learning are to alleviate the difficulty of predicting attributes and enhance the discriminant ability at the mean time, which especially important for fine-grained recognition task. By explicitly modeling the posture states and learning discriminative attribute with respect to different postures, describable and discriminative attribute can be built for final category recognition. The proposed pose-based discriminative attribute is verified on publicly available fine-grained dataset CUB with advanced performance.

Key words: attribute learning, discriminative attribute, distributed representation, fine-grained recognition

对象识别领域中, 即使对于同一对象或同类对象, 姿态变化导致呈现出来的外观信息差异较大, 造成类内差异明显, 给特征表示学习、度量学习带来很大挑战. 相应地, 几何对齐在模式识别系统中扮演着重要角色, 比如, 人脸对齐对人脸识别的性能影响巨大^[1], 也一直是视觉领域中的活跃研究课题, 比如 3D 情况下的对象对齐^[2-3].

属性表示, 由于具备较高的抽象语义含义, 使其天然具备较好的应对包括姿态变化在内的复杂环境变化的优势, 但这也恰恰成为属性学习的难点. 属性定义和属性学习中的视觉证据的模糊性, 成为属性学习中的典型问题, 比如属性概念漂移^[4]. 由此, 属性表示的可学习性、判别能力成为属性学习的关键问题, 尤其面临对判别属性要求较高的细粒度识别任务.

细粒度识别是最近备受关注的研究问题, 是图像高质量检索、图像场景理解中的关键问题. 细粒度识别是一类典型的多类分类问题, 其特殊性在于类别之间的差异较小, 比如, 待识别的类别为同科或同宗物种的子类别, 如虎皮鹦鹉、牡丹鹦鹉、金刚鹦鹉等. 类别之间的差异存在于细微之处, 因此, 结构分析和关键部件定位成为解决细粒度识别的重要着手点, 也是目前研究者的主要关注点^[5-6]. 同时, 面对细粒度识

收稿日期: 2016-08-20.

基金项目: 江苏省自然科学基金项目 (BK20161020)、江苏省高校自然科学研究项目 (15KJB520023).

通讯联系人: 宋凤义, 讲师, 研究方向: 计算机视觉与模式识别. E-mail: f.song@njnu.edu.cn

别这样的复杂任务,人类尤其是专家仍然能够清晰地描述子类别之间的差异,这些描述性词汇体现了精细的视觉语义信息,即视觉属性,因此,属性表示仍然是解决细粒度识别的有潜力的途径.

围绕属性表示的可学习性、判别力,研究者们做出了许多探索性工作.通常,属性学习依赖于精确的属性标注,不仅仅标注样本上有没有该属性,而且还要给出属性的视觉证据(visual evidence)区域,比如,CUB 数据集^[7]中直接把属性定义在局部部件上,诸如“喙是不是钩状,腹部颜色是否为黑色”,而属性的判别力强弱则依赖于属性集的选取,通常受限于专家领域知识.如果不考虑属性定义的成本,为每个类别定义一组属性,很大程度上会提升属性描述的判别力,因为所构造的属性集合更加专注于该类别对象的视觉特征刻画,但这样却牺牲了属性在类别间的共享性,且在样本收集过程中需要耗费更多的人工标注成本.另一种截然不同的方式是由数据驱动的属性学习. Farhadi^[29]提出在类别之间构建判别分类器,期望捕获能够最大化两组类别之间的视觉属性. Tom 和 Peter^[9]使这种做法限定了每组类别仅包含一个类别,并拓展到局部部件上^[10]. 另外一种思路则是考虑利用额外先验信息,比如姿态信息. Poselets^[11]是一种典型的做法,该方法从局部的角度来建模姿态,重点分析对象的部件的典型几何结构,并对局部部件的几何结构分布进行离散化表示,选取了几种代表性几何结构,如代表性姿态,然后在这些代表性几何结构状态下,分析属性以及对象的分布,由此降低了属性分析的难度,增强了属性学习的针对性,也提升了属性的判别力.

不同于以上方法,本文提出无监督的姿态分析方法,把待分析的对象按照姿态情况进行分组(姿态簇效果图参见图 1),然后在不同的姿态簇内进行判别属性学习. 相对于 poselets^[11]的做法,姿态情况的分析更加灵活,不局限于从姿态空间中采样的个别典型几何结构,具有较好的规模可扩展性. 相对于简单的判别属性,把属性学习限定到不同姿态状态下,可以大大降低待分析对象的分布的复杂性,从而既降低了属性学习的难度,同时保障了属性较好的类别判别性,适用于处理细粒度识别问题.



不同的姿态簇体现了不同的结构先验信息,同一姿态簇内外观特征并不一样,但外观特征所刻画的部件具有相似之处. 在不同姿态簇内分析对象的视觉属性,增强了属性分析的针对性.

Various pose clusters embody kinds of geometric structural prior information, while one pose cluster consists of images with similar explicit seen local parts but with different appearance feature. Learning visual attributes in a specified posture states will enhance its description specification ability.

图 1 姿态簇示意图

Fig. 1 Illustration of pose clusters

1 相关工作

属性学习的相关工作很多,这里选取与本文内容最相关的工作来回顾一下所关注问题的研究现状. 具体而言,将从以下三个方面来回顾与本研究内容密切相关的现有工作的思路和方法.

1.1 几何结构分析

计算机视觉中分析的视觉对象通常是现实世界中的真实物体,由于不同视角,以及对象自身表现出的姿态变化(通常非刚体对象表现出更多的结构变化),造成对象的几何结构差别巨大.因此,特定点定位和几何对齐一直是构建良好视觉模型的重要基础步骤,几何结构信息成为建立对象外观分布的重要先验信息.

描述几何结构的主要方式是通过关键部件的坐标位置来刻画(通过人工标定或者特征点定位模型来实现),建立在这些信息之上的模型称为点分布模型,比如ASM模型^[12]用来建模关键点的分布,并结合外观分布模型实现关键点的定位.图模型(pictorial model)^[23-24]是另外一种分析对象几何结构的方法,关键部件被表示为图中的节点,关键部件之间的结构变形度和外观匹配度则体现为图中的边.基于此,变形模板技术(DPM)^[13]成为对象检测和识别的有效方法,也被用于解决细粒度识别问题和属性学习中^[25].

考虑到局部结构分析中通常涉及到局部部件位置的标注以及详细特征点标注,需要耗费大量的人力代价.因此,如何以较低的成本实现自动化的局部信息以及几何结构信息的获取成为关键问题.由此,不进行局部部件标注,以自动化的方式实现局部部件检测,成为具有重要价值的研究目标.利用局部部件在多个样例中的并发出现特点,Krause等人^[26-27]提出了基于无监督手段的自动部件发现和定位技术,大大提升了几何结构信息利用的效率,并用于提升细粒度识别的精度.

1.2 属性学习

属性学习是融入领域先验知识到特征表示和分类模型中的有效方式.同时,属性的语义可理解性、类别共享性,以及编码的高效性被广泛应用到多个领域,如人脸识别、图像检索、场景理解、零样本学习等,以至于目前多数视觉数据集发布时都提供属性标注信息,比如Labeled Faces in the Wild (LFW)^[14],Animals with Attributes (AWA)^[8],Caltech-UCSD Birds (CUB)^[7],以及继ImageNet^[28]之后的又一大规模视觉数据集Visual Genome^[15].

但是属性学习并不是一件容易的事.首先,属性学习过程中的属性集选取,属性标注常依赖于人工方式来实现,既耗费了大量人力成本,同时也难以避免引入主观偏见,同时也存在属性的视觉证据不明确的问题,进而导致属性学习中的复杂问题,如概念漂移^[4].鉴于这些问题,研究者们尝试利用数据驱动的方式来规避上述问题,比如,利用第三方附带文本信息^[16],以及词典库来自动建立属性集;把属性看作隐变量^[17],联合学习属性表示和类别分类器,以避免属性定义和标注,同时,融入最终分类任务的目标要求到属性学习过程中,以此提升属性判别力.

另一方面,提升属性的判别力离不开类别信息的引入,结合利用类别信息可以实现一般意义上的判别属性学习.比如,Farhadi^[29]考虑到动物之间的区分度小于动物与交通工具的可分性,而尝试在不同组的类别之间学习判别属性.Tom和Peter^[9]则进一步把区分对象限定到两两类别之间,以此期望学习出能够最大化类别分类间隔的判别视觉属性,并进一步考虑在不同的局部几何结构情况下进行属性学习^[10].

1.3 细粒度识别

随着图像理解技术的进步,智能图像理解不应满足于简单的对象检测和识别.给定一张图片,智能视觉系统不仅仅要回答图中是否出现一只鸟,更要回答图中是一种什么鸟;同样,看到公园里的各种鲜花,人们禁不住想知道花的品种.细粒度识别正是要解决这类问题,可以看出,细粒度识别任务要辨识的类别间存在较大的混淆性,且用于区分类别的关键信息通常存在于细微之处.因此,细粒度识别任务的独特性要求视觉识别算法和模型应具有较好判别力的特征表示和模型学习能力.

考虑到对象的复杂非刚体变形,关键部件定位和几何结构分析是获取良好判别特征的基础.由此,许多工作关注实现较好的关键部件的定位,以及部件之上的特征提取和属性分析^[7],这一点充分体现在CUB数据集^[7]上的属性集设计上,在第4节数据集介绍中有详细阐述.围绕该基本思路,典型的方法包括,Wang Yaming^[18]挖掘具有判别力的局部图像块,在特征利用层面上提升细粒度模型的判别性.Zhang Han^[19]提出了由语义部件的位置信息来提升细粒度识别的性能.Akata^[20]则提出利用多种信息以强监督的方式来进行细粒度识别.其他方式还包括深度特征学习^[21]、度量学习^[22]来应对细粒度识别中对高判别性的要求.

2 基于姿态的判别属性学习

结合以上分析,本文提出基于姿态的判别属性学习方法,以无监督的方式充分挖掘利用姿态先验信息来提升属性学习的判别性,并最终提升基于属性表示的细粒度识别任务的性能. 本节从以下两个方面来介绍所提出的基于姿态的判别属性学习的思路及详细过程. 接下来,将首先解释姿态先验下的判别属性的动机和含义,随后介绍进行判别属性学习和利用.

2.1 姿态先验下的判别属性定义

判别属性旨在挖掘利用能够最大化区分类别间隔的视觉证据(visual evidence). 从特征选择的角度而言,判别属性代表着能够最大化区分两类的特征,可能属于正类也可能属于负类,或者是组合两类的典型特征. 从模型的角度而言,判别属性提取器用于计算样本中的相应属性出现的程度描述,具体则表示判定该样本与两类别相似的置信度. 从特征分布式表示(distributed representation)的角度而言,相当于把待编码样本投射到类别空间中,参照每个类别而给出测试样本的一种类别归属度的分布式描述. 判别特征(discriminative attribute)最早由 Farhadi^[29]提出用于刻画区分一组类别与另一组类别的视觉信息,不同组的类别可以随机划分产生,也可以只关心部分类别之间的判别信息分析. 进一步 Tom 和 Peter^[9]把类别划分规则定义在任意两两类别之间,如此在一个更具体的类别划分下进行判别信息分析.

从上述分析可以看出,判别属性既有一定的语义可解释性(与待辨识的两类对象的相似程度),同时又融入了类别信息,大大提升了属性表示的判别力,但是不足的是上述做法没有考虑计算机视觉中普遍存在的复杂变化因素对模型性能的影响. 鉴于此,本文重点分析姿态先验信息下的判别属性学习. 姿态先验信息很大程度上决定着对象所呈现出来的外观信息. 不同的姿态代表着对象呈现出的外观信息不同,姿态信息是建立对象外观分布的重要先验信息.

鉴于此,基于姿态的判别属性的出发点在于把姿态信息融入到寻找能够最大化类别分类间隔的视觉属性的过程中. 最终目标是期望能够降低判别属性学习的难度,提高判别属性抽取的精准度. 要利用姿态先验信息并不是一件容易的事. 典型的思路包括对齐,在人脸识别流程中成为重要的预处理步骤. 这种思路可以看作是对姿态变化导致的对象的几何结构空间的归一化处理,但很难应对大姿态下的变形. 一个理想的处理思路是建立对象的几何结构分布,但通常涉及到大量的参数,模型复杂度很高,模型估计中需要的样本数量也呈指数级增长. 一个比较可行的方式是采用离散化的方式来近似刻画对象的几何结构分布. 比如, Bourdev 提出的 Poselets^[11]的思路,把人体的几何结构分布离散化为一组典型的结构分布,并在其中考虑引入局部建模的思想,即,采用局部结构刻画结合离散刻画的方式,然后在相应的结构分布先验下进行后续的识别工作. 这种方式也符合应对复杂任务的“分而治之(divide and conquer)”的模式识别思想. 本文将结合这种思想到判别属性学习过程中,接下来将给出详细实现步骤.

2.2 姿态先验下的判别属性学习

基于 2.1 节对基于姿态的判别属性的介绍,可知判别属性学习的关键步骤是如何进行姿态划分. 这里为了降低人工代价,而采用无监督聚类方式,即,依据对象的关键部件位置信息(可借助于部件检测器实现自动化信息收集),通过聚类技术对不同姿态的样本进行划分,由此建立了不同姿态下的样本集合,用于进一步构建姿态先验下的判别属性提取器.

本文方法不同于之前介绍的 Poselets^[11]方法中对姿态信息的利用方式,在于以聚类的方式自动化地建立离散化的姿态分布,既近似刻画了对象的几何空间分布,也无需耗费大量的人力,同时可以避免姿态空间离散化过程中的人为主观偏见. 从图 1 所示聚类效果可以看出,簇内呈现出较好的姿态统一性,这比较符合鸟类这一自然对象的运动规律和生活习性,比如,飞翔的典型方式,栖息的典型方式等.

接下来,将在每个姿态簇下实现 3.1 节中介绍的判别属性学习过程. 给定姿态簇划分,在每个簇中,可以利用样本类别标号训练 1-vs-all 分类器作为判别属性提取器. 考虑到在不同的姿态簇中,可能涵盖的类别数量不定,每类下的样本数目也不定,因此不适合采用 1-vs-1 多类别判别分析方式,同时为了防止 1-vs-all 中的正类样本数太少,而采用正类样本补偿的方案,即,使用该类的所有的训练样本作为 1-vs-all 分类器的正例集合,聚类簇中的其他类别的样本作为反例集合. 本文方法不同于之前介绍的判别属性学习方式,在于在同一姿态先验信息下(同一姿态簇下)进行判别属性的学习,采用 1-vs-all 的类别判别方

案,同时引入正类样本补偿方案,提升模型的判别性以及应对类别不平衡和类别样本深度问题(同类样本的数量).

具体过程整理成如下算法:

算法 基于姿态的判别属性学习

输入: N 个训练样本,每个样本的关键点坐标以及类标(K 类)

输出: 属性提取器,及类别分类性能

- 1 依据训练样本结构信息进行聚类为 M 个簇 $\{C_1, C_2, \dots, C_M\}$
- 2 for $i = 1:M$
- 3 for $j = 1:K$
- 4 选择第 j 类的所有样本为正类,
- 5 簇中除第 i 类之外的所有样本为负类,
- 7 训练线性 SVM 分类器,作为属性提取器
- 8 end
- 9 end
- 10 在所有样本上计算 $M \times K$ 个分类器的响应值
- 11 基于响应值的新编码表示,训练 1-vs-all 线性 SVM 类别分类器
- 12 在测试样本上评估分类性能.

3 实验

为了验证提出的姿态先验下的判别属性学习方法在细粒度识别任务中的有效性,在典型的细粒度识别公开数据集 Caltech-UCSD Birds 200-2011(CUB)^[7]上展开了详细的实验验证.

3.1 数据集和实验协议

CUB 数据集^[7]建立了鸟的子类别辨识任务,共收集了 11 788 张鸟的图片,涵盖 200 种不同种类的鸟,每张图片上给出的标注信息包括类别、局部部件的位置、属性,以及鸟出现在图片中的位置. 其中,类别涵盖 200 种鸟,比如黑足信天翁、北美红雀、灰猫鹊等,局部部件共 15 个,比如背部、眼部、喙、翅膀等,属性共 312 个用来刻画 15 个局部部件的特性包括颜色、形状等.

CUB 数据集^[7]上给定了标准的训练和测试数据划分协议,共分作 5 994 张训练样本,5 794 张测试样本. 训练样本和测试样本中的类别上的样本数保持均衡,平均每类 30 个样本分别用于训练和测试. 本文遵从标准的实验协议,所提方法中涉及到的聚类、编码分类器训练,以及最终的类别分类器的训练过程都是在训练样本上进行的,最终在测试集上汇报性能. 多类别识别性能指标采用了平均类别识别正确率,特征采用数据库提供的 312 维属性表示.

3.2 基准算法

在第 2 节相关工作介绍中,回顾了与本文方法最相关的工作 Tom-vs-Peter^[10]编码器的想法,该方法尝试在两两类别之间学习判别分类器,期望能够捕捉关键视觉信息以最大化两类分类间隔,在人脸验证和鸟类识别上均取得了较好的性能. 考虑到不同的实验细节会严重影响最终的性能结果,本文中依照 Tom-vs-Peter 编码器的想法,在遵从与本文相同的实验设定的情况下,重新评估了模型性能,使得算法比较在相对公平的条件下进行,从而对不同算法的效果建立较明确的认识.

考虑到本文提出的判别属性的落脚点在于建立一种特征表示——兼具较好的语义描述能力和类别判别能力,由此,直接使用 CUB 数据集上提供的 312 个属性表示成为本文要对比的另一对比算法.

3.3 实验结果

首先对比验证了所提方法与两个基准算法之间的性能差别,详细结果展示在图 2 中. 可以看出此类方法表现出比两种对比算法都要好的性能. 在设置聚类簇个数为 15 的情况下,该方法取得了 44.72% 的多类别识别平均精度,相对于 312 维属性表示的 41.34% 的性能提高了 3 个百分点以上,相对于基于 Tom-vs-Peter^[10]类别判别编码方案的性能(36.19%)提高了 8 个百分点以上.

本文所提方法中,如何设置姿态簇的个数成为一个关键参数. 由此,本文还验证了不同姿态簇个数对

性能的影响,结果如图 3 所示,在较小的聚类个数基数下,性能逐渐提升,当到达一定程度之后提升则较慢. 对这一现象的解释有三点:第一,聚类数目较少时,姿态的划分比较粗糙,难以形成准度较高的划分;第二,编码长度随着聚类个数的增强表现出倍数扩增趋势,应用到更大聚类个数下会造成最终类别分类器的维数太大,增加模型训练的复杂度;第三,聚类个数增加导致每个聚类中的样本数减少(CUB 上训练样本共 5994 个),难以充分刻画对象的分布特点.

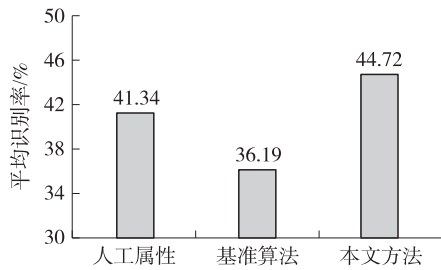


图 2 不同算法的比较

Fig. 2 Performance comparison of several approaches

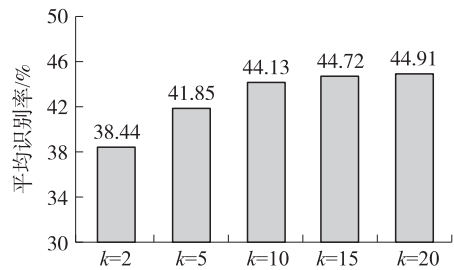


图 3 展示所提出的方法性能受不同姿态簇个数的影响

Fig. 3 Illustrating how different number of clusters affects the performance of the proposed approach

对于第一点解释,还做了另外一个实验来充分验证其合理性,即,测试随机划分簇下的判别属性的性能,结果展示在图 4 中. 对比图 4 和图 3 可以看出,当聚类个数较少时($k=2$),依照几何结构进行划分簇和随机划分簇下的判别属性的性能相差无几,这说明聚类个数很少时,难以保证簇内的几何结构的统一性,也充分说明了几何结构统一性对后续判别属性的表达能力的影响. 另外,在不同的簇个数下,对比姿态先验下的判别属性与随机划分簇下的判别属性在后续细粒度辨识中的性能差异,从图 3 和图 4 可以看出: (1) 在任一簇个数设置下,基于几何结构分析的姿态簇下的复合属性用于后续任务的性能都高于随机簇下的判别属性用于后续任务的性能; (2) 两种簇划分方式下,簇个数的增加都带来性能的提升. 对于(1)充分证明了姿态先验信息对于提升判别属性的表达力的贡献,对于(2)可以归结为随机划分也能把原来复杂的姿态分布给简单化(分而治之),更重要的是在本文提出的判别属性学习方法中,姿态先验是重要信息,而类别判别信息也起着重要的作用,这一点保证了随机划分簇下的判别属性在后续任务的性能.

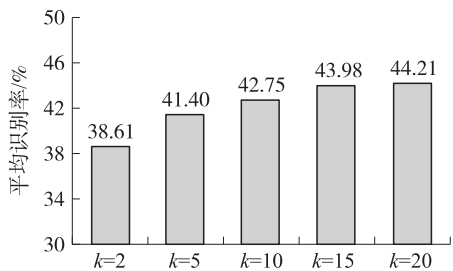


图 4 展示随机簇划分下的判别属性用于细粒度辨识的性能及其受不同簇个数的影响

Fig. 4 Illustrating the performance of fine-grained recognition accuracy with discriminative attribute learned under random clusters, and also the performance tendency with increasing number of clusters

4 结语

细粒度识别问题作为典型的多类别分类问题,需要更具判别力的特征表示方法和模型学习策略. 属性表示虽然在计算机视觉任务中具有独特的语义可理解性优势,符合人类视觉认知习惯,但面对细粒度识别任务时难以保证较好的性能. 本文提出的姿态先验下的判别属性尝试在利用类别信息的同时考虑引入姿态先验信息,进一步提升基于类别信息的判别属性的表达能力,所提方法在细粒度识别任务中表现出较好的性能. 进一步,考虑到细粒度识别中子类别间辨识的关键在于局部特征判定上,因此,细致的外观结构分析也是提升细粒度识别问题中的关键,融合几何结构信息和外观结构信息到判别属性学习中,将是下一步要开展的重要工作.

[参考文献]

- [1] SHAN S, CHEN X, GAO W. Face misalignment problem[M]. New York, USA: Springer, 2009.
- [2] ZHU X, LEI Z, LIU X, et al. Face alignment across large poses: a 3D solution[C]//Proceedings of the 29th IEEE Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016.
- [3] JOURABLOO A, LIU X. Large-pose face alignment via CNN-based dense 3D model fitting[C]//Proceedings of the 29th IEEE Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016.
- [4] JAYARAMAN D, SHA F, GRAUMAN K, et al. Decorrelating semantic visual attributes by resisting the urge to share[C]//Proceedings of the 27th IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA: IEEE, 2014.
- [5] GAVVES E, FERNANDO B, SNOEK C G, et al. Local alignments for fine-grained categorization[J]. International journal of computer vision, 2015, 111(2): 191–212.
- [6] ZHANG L, YANG Y, ZIMMERMANN R, et al. Fine-grained image categorization by localizing tiny object parts from unannotated images[C]//Proceedings of the 5th ACM on International Conference on Multimedia Retrieval. Shanghai: ACM, 2015.
- [7] WAH C, BRANSON S, WELINDER P, et al. The caltech-UCSD birds-200-2011 dataset[R]. Los Angeles: CIT, 2011.
- [8] LAMPERT C H, NICKISCH H, HARMELING S, et al. Learning to detect unseen object classes by between-class attribute transfer[C]//Proceedings of the 22th IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA: IEEE, 2009.
- [9] BERG T, BELHUMEUR P N. Tom-vs-pete classifiers and identity-preserving alignment for face verification[C]//Proceedings of the 23rd British Machine Vision Conference. Guildford, UK: IEEE, 2012.
- [10] BERG T, BELHUMEUR P N. POOF: part-based one-vs.-one features for fine-grained categorization, face verification, and attribute estimation[C]//Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA: IEEE, 2013.
- [11] BOURDEV L, MAJI S, MALIK J, et al. Poselets: A distributed representation for visual recognition[J]. Journal of vision, 2011, 11(11): 891–891.
- [12] REED J L, PALSSON B O. Thirteen years of building constraint-based in silico models of *Escherichia coli*[J]. Journal of bacteriology, 2003, 185(9): 2 692–2 699.
- [13] FELZENSZWALB P F, GIRSHICK R, MCALLESTER D, et al. Object detection with discriminatively trained part-based models[J]. IEEE transactions on pattern analysis and machine intelligence, 2010, 32(9): 1 627–1 645.
- [14] LFARNED-MILLER E, HUANG G B, ROYCHOWDHURY A, et al. Labeled faces in the wild: a survey[M]//Advances in face detection and facial image analysis. Berlin: Springer, 2016: 189–248.
- [15] KRISHNA R, ZHU Y, GROTH O, et al. Visual genome: connecting language and vision using crowdsourced dense image annotations[J]. International journal of computer vision, 2017, 23(1): 1–42.
- [16] WANG S, JOO J, WANG Y, et al. Weakly supervised learning for attribute localization in outdoor scenes[C]//Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA: IEEE, 2013.
- [17] WANG Y, MORI G. A discriminative latent model of object classes and attributes[C]//Proceedings of the 11th IEEE Conference on European Conference on Computer Vision. Greece: IEEE, 2010.
- [18] WANG Y, CHOI J, MORARIU V I, et al. Mining discriminative triplets of patches for fine-grained classification[C]//Proceedings of the 29th IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016.
- [19] HAN Z, TAO X, MOHAMED E, et al. SPDA-CNN: unifying semantic part detection and abstraction for fine-grained recognition[C]//Proceedings of the 29th IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016.
- [20] AKATA Z, MALINOWSKI M, FRITZ M, et al. Multi-cue zero-shot learning with strong supervision[C]//Proceedings of the 29th IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016.
- [21] REED S, AKATA Z, SCHIELE B, et al. Learning deep representations of fine-grained visual descriptions[C]//Proceedings of the 29th IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016.
- [22] CUI Y, ZHOU F, LIN Y, et al. Fine-grained categorization and dataset bootstrapping using deep metric learning with humans in the loop[C]//Proceedings of the 29th IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016.
- [23] FELZENSZWALB P F, HUTTENLOCHER D P. Pictorial structures for object recognition[J]. International journal of

- computer vision, 2005, 61(1):55–79.
- [24] TAN X, SONG F, ZHOU Z, et al. Enhanced pictorial structures for precise eye localization under incontrolled conditions[C]//Proceedings of the 22th IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA:IEEE, 2009.
- [25] ZHANG N, FARRELL R, IANDOLA F, et al. Deformable part descriptors for fine-grained recognition and attribute prediction[C]//Proceedings of IEEE Conference on International Conference on Computer Vision. Sydney, Australia:IEEE, 2013:729–736.
- [26] KRAUSE J, GEBRU T, DENG J, et al. Learning features and parts for fine-grained recognition[C]//Proceedings of the 27th IEEE Conference on International Conference on Pattern Recognition. Columbus, USA:IEEE, 2014.
- [27] KRAUSE J, JIN H, YANG J, et al. Fine-grained recognition without part annotations[C]//Proceedings of the 28th IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA:IEEE, 2015.
- [28] DENG J, DONG W, SOCHER R, et al. ImageNet: a large-scale hierarchical image database[C]//Proceedings of the 22th IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA:IEEE, 2009.
- [29] FARHADI A, ENDRES I, HOIEM D, et al. Describing objects by their attributes[C]//Proceedings of the 22th IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA:IEEE, 2009.

[责任编辑:黄 敏]

(上接第64页)

- [8] MARTÍ L, SANCHEZPI N, MOLINA J M, et al. Anomaly detection based on sensor data in petroleum industry applications[J]. Sensors, 2015, 15(2):2 774–2 797.
- [9] YUAN Y, FANG J, WANG Q. Online anomaly detection in crowd scenes via structure analysis[J]. IEEE transactions on cybernetics, 2015, 45(3):562–575.
- [10] KICH J I D F, PEREIRA M F. Appendix: empirical mode decomposition (EMD) method[J]. 海洋学报(英文版), 2015, 19(11):921–940.
- [11] ZHU B, WANG P, CHEVALLIER J, et al. Carbon price analysis using empirical mode decomposition[J]. Computational economics, 2015, 45(2):195–206.
- [12] AKIRA S, EMIKO S. Influence of root conditioning prior to EMD application on periodontal ligament cells of extracted teeth[J]. Applied mechanics and materials, 2014, 568–570(2):1 951–1 954.
- [13] HU J, XIE Q, WANG X, et al. A novel Bi-dimensional EMD algorithm and its application in image enhancement[J]. Information technology journal, 2014, 13(3):469–476.
- [14] CUMRUNVAFA. Extending mirror conjecture to calabi-yau with bundles[J]. Communications in contemporary mathematics, 2012, 1(1):65–70.
- [15] 许宝杰, 张建民, 徐小力, 等. 抑制 EMD 端点效应方法的研究[J]. 北京理工大学学报, 2006, 26(3):196–200.

[责任编辑:黄 敏]