

一种基于深度学习的半监督分层模型

王江晴^{1,2}, 张 蕾^{1,2}, 孙 翀^{1,2}, 帖 军^{1,2}, 周玮瑜^{1,2}, 孟 凯^{1,2}

(1. 中南民族大学计算机科学学院, 武汉, 430074; 2. 湖北省制造企业智能管理工程技术研究中心, 武汉, 430074)

摘 要: 依照图像识别出的对象标签, 通过层次结构来分类图像集是图像自动化分类的重要研究问题之一。现有的方法实现了对象标签已知情况下的层次结构构建, 仅存在少量方法考虑部分对象标签未知的情况。本文对经典方法进行了扩展和优化, 实现了存在部分对象标签未知情况下的层次结构构建和更新。利用卷积神经网络(Convolutional neural network, CNN)对图像编码, 提出半监督学习方法, 根据传统算法构建类标签已知图像集的层次结构, 通过周期性相似性比较, 对层次结构中标签未知图像进行聚类, 实现对半监督分层模型(Semi-supervised layer-wise model, SLM)的构建。本文采用了真实公开的数据集, 实验结果表明, 该方法能够有效地实现层次结构的构建和更新, 并且能够在较小规模的数据集上取得好的预测分类效果。

关键词: 卷积神经网络; 图像分类; 层次结构; 聚类; 半监督

中图分类号: TP391 **文献标志码:** A

A Semi-supervised Layer-Wise Model Based on Deep Learning

WANG Jiangqing^{1,2}, ZHANG Lei^{1,2}, SUN Chong^{1,2}, TIE Jun^{1,2}, ZHOU Weiyu^{1,2}, MENG Kai^{1,2}

(1. College of Computer Science, South-Central University for Nationalities, Wuhan, 430074, China; 2. Hubei Provincial Engineering Research Center for Intelligent Management of Manufacturing Enterprises, Wuhan, 430074, China)

Abstract: Using hierarchical structure to classify image set which has the object labels identified by images is one of the important research issues in image automation classification. The previous researches have already implemented the hierarchical structure construction for the labeled images, and now there are only a few methods to consider the influence of the part of the unlabeled images. In this paper, the classical method is extended and optimized, and the hierarchical structure construction and update are realized when some object labels are unknown. The convolutional neural network (CNN) is used to encode these images, and the semi-supervised learning method is proposed. The hierarchical structure of the image set which has known the object labels is constructed according to the traditional algorithm. Through the periodic similarity comparison, the unlabeled images in the hierarchy are clustered. The construction of the semi-supervised layer-wise model (SLM) is realized. This paper adopts the real public data sets. The experimental results show that the SLM can effectively realize the construction and update of the hierarchical structure, and can achieve good prediction classification effect on the smaller scale data sets.

Key words: convolutional neural network; image classification; hierarchical structure; clustering; semi-supervised

基金项目: 国家社会科学基金重大(17ZDA166)资助项目; 湖北省技术创新专项重大(2019ABA101)资助项目; 中央高校基本科研业务费专项资金(CZY18015)资助项目。

收稿日期: 2019-10-22; **修订日期:** 2020-01-10

引 言

图像集分类是计算机视觉领域的研究热点之一。实际应用中,图像集中图像数量大、类别多且不同类别的图像间也可能存在相似性^[1]。扁平化的图像集分类方法不足以刻画图像间的相似关系,利用层次结构可以解决该问题。在层次结构中,自顶向下对图像集中的图像进行从粗到细的分类。如图1所示,每个节点代表一个图像类别集合,根节点包含所有的图像类别,每个叶节点只包含一个特定的图像类别。对于一个给定的内节点,只包含父节点图像类别的一个子集,这些图像类别可能在某方面具有一定的相似性^[2]。现有图像集层次结构分类研究局限于对标签已知的图像集,当出现标签未知图像时,这些方法都可能存在重新学习层次结构的问题,针对标签未知图像在层次结构中分类的研究目前很少。

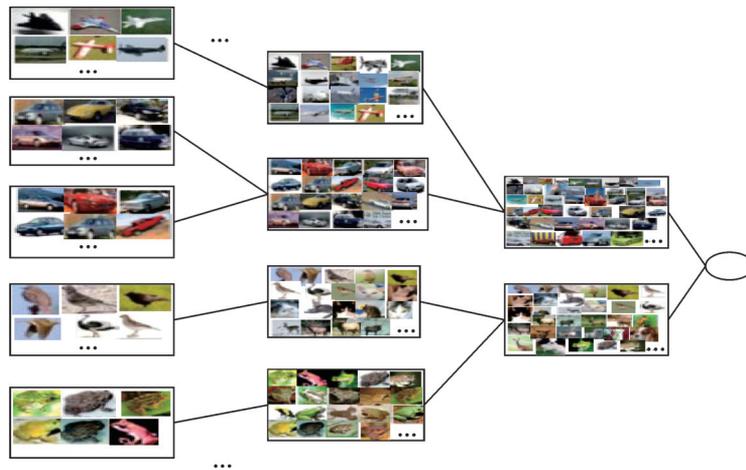


图1 图像集分类层次结构

Fig.1 Image set classification hierarchy

为了解决标签未知图像在树结构中的无法定位及标签已知新图像在树结构添加的问题,本文提出了一种半监督分层模型(Semi-supervised layer-wise model, SLM)。该模型实现对标签已知数据集进行分类,并有效地学习标签未知图像特征,找到其在树节点中的相应位置,更新层次结构。SLM具有3方面的特点。(1)图像集编码。即在层次结构中,图像的分类是在特征空间而不是在语义标签空间中实现的。对比传统图像处理算法,卷积神经网络(Convolutional neural network, CNN)优点是其可以直接对1张输入图像进行处理。另一方面,深度学习强大的建模能力允许通过特征表示来处理复杂的高维数据集^[3];(2)层次结构。根据图像集中各个类别的图像编码,聚类得到层次结构,自底向上地将标签已知图像集划分出不同的组,实现在层次结构图像自动化分类;(3)针对标签未知图像更新层次结构。将标签未知的图像编码与预训练层次结构的节点编码自底向上进行相似性比较,将匹配图像归入层次结构。实验表明该方法能有效地更新层次结构,在较小训练集中取得好的效果。

1 相关工作

图像集层次分类研究工作主要涉及到3个方面:图像特征提取、层次结构学习以及针对标签未知图像更新的问题。利用深度学习提取图像特征能够更抽象、更深层次描述图像的属性特征;根据图像之间的相关性进行层次结构的构建,在图像集应用中前景巨大;随着图像集中动态出现新图像,层次结构更新也出现新的解决方案。

近年来,深度学习^[4-6]在图像特征提取应用上发展迅速。文献[7]通过引入卷积神经网络,研究网络

深度对效果的影响,当增加到16层与19层时,效果最好。VGG-16与VGG-19两种网络提高神经网络在图像分类中的准确性。VGGNet卷积核大小只有 3×3 ,两层的卷积核可以达到 7×7 卷积核的效果,使得模型参数更少。深度学习展示了学习图像特征、提高识别正确率等方面的突出能力。扁平化的图像分类方法忽略了图像类别之间存在的关联性,将对象类分层组织成层次结构可以解决这个问题。文献[8]提出将图像集合组织成树形层次结构,利用层次树对对象类进行分组;文献[9]提出一种用于学习大规模分类的标签树,同时确定了树结构分类器;文献[10]提出了一种支持大规模植物物种识别的层次学习算法,利用多任务结构学习算法联合训练相关分类器,以提高识别能力;文献[11]提出支持大规模视觉识别的分层深度多任务学习(Hierarchical deep multi-task learning, HD-MTL)算法,从神经网络中不同层提取多组多级深度特征,利用类间相似性来学习分组的深度表示法,从而有效地区分视觉上相似的对象类。文献[12]提出针对新类别检测的基于支持向量机(Support vector machine, SVM)多任务分类器,将给定的测试图像示例找到一个最匹配的父节点,如果测试图像属于某个子节点,将其分配到对应子节点,否则在父节点下插入一个新的子节点。该方法为监督学习算法,每个节点都有一个分类器,出现新类别图像时,训练成本高。并且还可能会层间错误传播的影响,即父节点的错误将传递到其子节点,直到叶节点。基于上述存在的问题,本文提出基于半监督学习的算法,利用带标签图像聚类构建层次结构,利用不带标签图像自底向上更新层次结构。

2 半监督分层模型

每个对象类是包含相同对象的1组图像的集合。标签不同的同一细粒度对象类之间由于具有相似性,会归属于同一个粗粒度标签中。SLM主要可以分为3个部分(图2),首先是对标签已知图像集进行编码,利用CNN的卷积层实现对图像的特征提取;其次是利用图像编码进行聚类构建图像集分类的层次结构;最后根据动态出现的标签未知图像,更新层次结构。

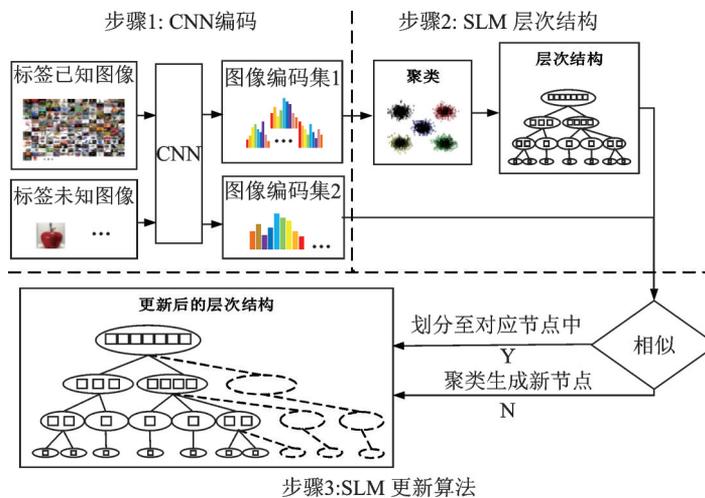


图2 SLM框架

Fig.2 SLM framework

2.1 CNN 编码

给定图像集,选择预训练的VGG网络中全连接层前的卷积层作为特征提取层,提取图像特征得到图像编码集。其中,标签已知图像编码集包含 N 个对象类;标签未知图像编码集为 $U=\{u_1, u_2, \dots\}$,其中第 i 张标签未知图像编码为 u_i 。

2.2 SLM层次结构

深度为 H 的层次结构包含3种节点:根节点、内节点和叶节点。其中,根节点包括所有的对象类;每个内节点仅包含其子节点中所有的对象类;每个叶节点都包含一个特定的对象类。层次结构构建过程如下。

(1)图像集中共有 N 个对象类,利用 K -means对图像编码集聚类,取中心点编码作为层次结构的叶节点编码,第 H 层的叶节点编码集 $V_H=\{v_H^1, v_H^2, \dots, v_H^N\}$ 包括 N 个叶节点。每个叶节点对应的对象类是聚类后每组数量最多的一类对应的对象类。

(2)根据图像集构建先验知识树,引入收缩因子 α, β ,计算每层中内节点的个数,即每次聚类中 k 值,具体计算公式为

$$\text{Count}(l) = (\text{Count}(H) - \alpha)^{\frac{1}{2}} + \beta \tag{1}$$

式中: $\text{Count}(l)$ 代表第 l 层内节点个数, $\text{Count}(H)$ 代表第 H 层叶节点个数。收缩因子 α 和 β 主要作用是来确定每次聚类 k 值的选取,即第 l 层内节点个数 $\text{Count}(l)$ 。根据选取的 k 值对图像编码集进行 K -means聚类,得到第 l 层的内节点编码集 $V_l=\{v_l^1, v_l^2, \dots, v_l^k\}$ 。

(3)层次结构中节点与节点间的边由对应图像集的隶属关系决定,计算每个后辈节点与上一层所有先辈节点对应的图像集交集,以交集最大的集合为主。对于每一个后辈节点,节点间隶属关系概率值 p 计算公式为

$$p = \frac{\text{Count}(P_A \cap P_D)}{\text{Count}(P_D)} \tag{2}$$

式中: $\text{Count}(\cdot)$ 代表图像数量; P_A 代表先辈节点图像集合, P_D 代表某个后辈节点图像集合。当 p 值最大时,对应 p 值最大的先辈节点作为该节点的父节点,连接形成一条新边。同理,对图像集自底向上分层连接形成层次结构。

2.3 SLM更新算法及分析

在更新树的过程中需要对新图像进行预分类,预分类的图像包含标签已知(已存在于树中的标签)图像以及标签未知图像,将预分类的图像编码与各个节点自底向上进行相似性比较。如果找到相似节点,将其分配给对应节点;如果图像不属于其中任意节点,则在层次结构中增加新节点,更新层次结构。具体过程如下。

(1)设置相似度阈值 θ 。

(2)随机选取第 i 个标签未知图像编码 u_i 与层次结构中所有的叶节点编码集 V_H 进行相似度计算,得到一组欧氏距离 $E_H, E_H=\{e_H^1, e_H^2, \dots, e_H^N\}$ 表示在第 H 层中计算 N 次欧氏距离。如果 E_H 中存在小于 θ 的值,则将该图像划分 E_H 最小的叶节点中;如果 E_H 的值全部大于 θ ,则表明在该层中没有匹配节点,继续向上层进行比较。

(3)重新计算图像编码 u_i 与上一层内节点编码 V_l 之间的欧氏距离 $E_l, E_l=\{e_l^1, e_l^2, \dots, e_l^k\}$ 。如果 E_l 中存在小于 θ 的值,则将该图像划分 E_l 最小的内节点中并且划分至新生成子节点池;如果 E_l 的值全部大于 θ ,则在该层中没有匹配节点,重复步骤(3),直到上层节点为根节点。无论该图像是否属于根节点,都将其归入根节点并划分至新生成子节点池。每个先辈节点只生成一个子节点池。

(4)当输入图像达到一定规模时,对层次结构中新增的子节点池进行聚类,更新层次结构。

算法1 SLM更新算法

输入:标签未知图像编码集 $U=\{u_1, u_2, \dots, \}$,层次结构叶节点编码集 $V_H=\{v_H^1, v_H^2, \dots, v_H^N\}$,内节点编码集 $V_l=\{v_l^1, v_l^2, \dots, v_l^k\}$;

输出:更新后的图像分类层次结构

- ① for each u_i from $U=\{u_1, u_2, \dots\}$
- ② 计算标签未知图像编码 u_i 与叶节点编码 V_H 之间的欧氏距离 E_H
- ③ if E_H 中存在小于 θ 的值
- ④ 将该图像划分对应 E_H 最小的叶节点中
- ⑤ 跳出判断,进行步骤⑯
- ⑥ else
- ⑦ 计算 u_i 与上一层内节点编码 V_l 之间的欧氏距离 E_l
- ⑧ if E_l 中存在小于 θ 的值
- ⑨ 将该图像划分 E_l 最小的内节点中,并且划分至新生成子节点池
- ⑩ 跳出判断,进行步骤⑯
- ⑪ else
- ⑫ if l 不等于 1
- ⑬ 跳转至步骤⑦
- ⑭ else
- ⑮ 将其归入根节点并划分至新生成子节点池
- ⑯ if 输入数量小于 Num(Num=200)
- ⑰ 输入下一个标签未知图像编码,重复步骤①至步骤⑮
- ⑱ else
- ⑲ 从上到下对每层生成的子节点池进行聚类操作生成新节点,更新对应祖先节点
- ⑳ end for

该算法描述了 SLM 更新算法流程。在比较阶段,将标签未知图像编码从叶节点开始,自底向上的进行比较,直到根节点为止。层次结构中父节点描述的是所有兄弟子节点之间共享的公共特征,并且相似性总是自底向上进行比较,标签未知图像总是可以在层次结构中找到一个最佳匹配节点。

3 实验环境与结果

3.1 数据集

本文选取两个真实公开数据集对提出的 SLM 算法进行评价,数据集简介如下。

CIFAR-10:该数据集由 10 个类的 60 000 个 32×32 彩色图像组成,每个类有 6 000 个图像。其中有 50 000 个训练图像和 10 000 个测试图像。数据集分为 5 个训练批次和 1 个测试批次,每个批次有 10 000 个图像。测试批次包含来自每个类别的恰好 1 000 个随机选择的图像。

CIFAR-100:该数据集包含 100 个类,每个类包含 600 个图像。每类各有 500 个训练图像和 100 个测试图像。CIFAR-100 中的 100 个类被分成 20 个超类。每个图像都带有一个“精细”标签(它所属的类)和一个“粗糙”标签(它所属的超类)。

CIFAR-10 数据集与 CIFAR-100 数据集中图像的种类不重复。本文随机选取 CIFAR-10 中每一类的 3 000 张图像构建层次结构,选取 CIFAR-100 中 6 类图像(苹果、火车、皮卡车、老虎、牛和坦克)每一类的 600 张图像作为类标签未知的图像,与 CIFAR-10 剩余图像结合,70% 更新层次结构,30% 测试更新后层次结构对图像分类的准确性。

3.2 实验环境设置

本文使用 VGG-16 网络提取图像的全局特征,其特征维度为 512 维,为了提高计算效率,使用主成分分析(Principal component analysis, PCA)对其进行降维,降维后维度为 20 维。实验在 Python 3.4.2 环境下进行,采用 TensorFlow 框架,实验计算机配置为主频 3.40 GHz 4 核 CPU,8 GB 内存。

3.3 实验结果分析

层次结构的深度 H 对图像分类结果影响较大,如图 3 所示。分层过小,图像对象类之间相似性不能清楚表达,层次结构的分类效果不明显;分层过多会导致分类过细,增加模型的复杂度,找到最优的深度 H 对分类结果至关重要^[13]。为了控制模型的复杂度,假设层次结构有 4 层(包含根节点)。在 SLM 模型构建中,参数的设定对层次结构的更新影响非常大。在进行相似性比较时,需要确定阈值 θ ,阈值的选取对准确率的影响如图 4 所示。从实验结果中可知,当阈值取值为 2.5 时,准确率最高。

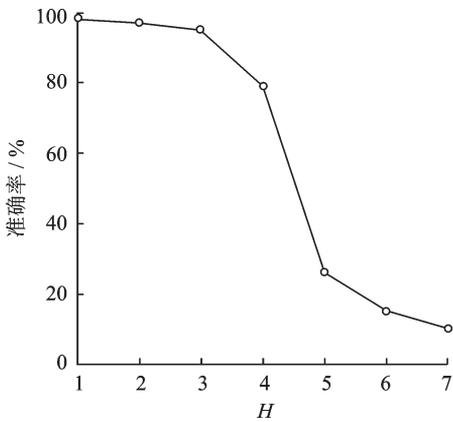


图3 层次结构深度 H 对分类准确率的影响

Fig.3 Effect of hierarchical structure depth H on classification accuracy

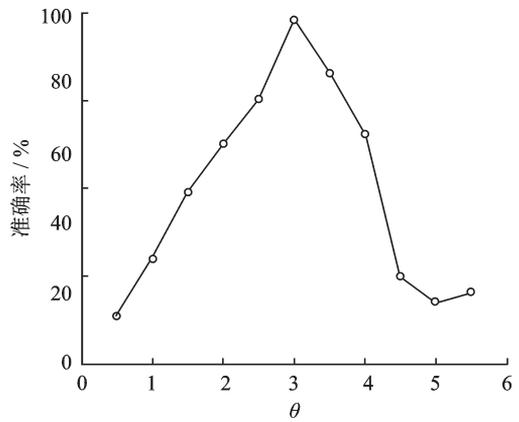


图4 阈值 θ 选取对分类准确率的影响

Fig.4 Effect of threshold θ selection on classification accuracy

将 CIFAR-10 数据集根据图像形态构建先验知识树,如图 5 所示,再将 CIFAR-10 数据集中选取的图像进行 K-means 聚类构建层次结构,如图 6 所示。层次结构中叶节点包含一个对象类,每个对象类是图像集中该类别所有图像的集合,内节点表示聚类后对象类的分组,根节点包含所有对象类。

算法采用两个标准进行评价:(1)标签未知图像分类的准确性;(2)标签已知图像分类的准确性。

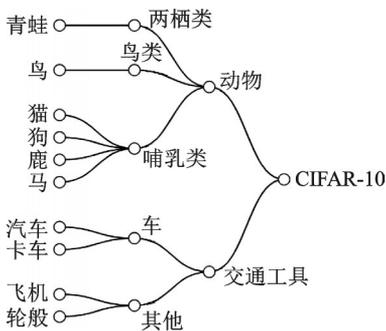


图5 先验知识层次结构

Fig.5 Prior knowledge hierarchy

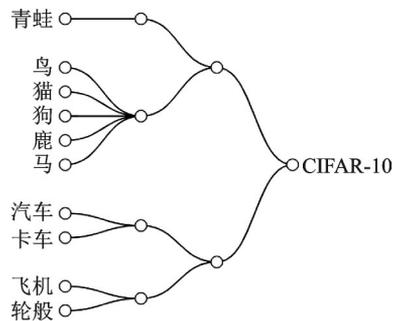


图6 K-means 聚类构建的层次结构

Fig.6 Hierarchical structure of K-means cluster construction

针对测试集的匹配结果设置匹配值,如果实际分类结果与预测对应的节点位置相同,则匹配值为 τ , τ 的初值为1;否则,分类结果可能在正确节点的先辈节点上,继续向上层匹配,每上升一层匹配值 τ 下降 $\frac{1}{L}$,直到0。测试集中正确分类的图像数量为

$$N_T = \sum \tau \quad (3)$$

式中: N_T 代表测试集中正确分类的图像数量。更新后的层次结构对新类别图像分类的准确率定义为

$$\text{prec}_1 = \frac{N_{T1}}{N_{\text{Total1}}} \quad (4)$$

式中: N_{T1} 代表新类别图像正确分类的个数, N_{Total1} 代表新类别图像的总个数。更新后的层次结构对标签已知新图像分类的准确率定义为

$$\text{prec}_2 = \frac{N_{T2}}{N_{\text{Total2}}} \quad (5)$$

式中: N_{T2} 代表标签已知新图像正确分类的个数, N_{Total2} 代表标签已知新图像的总个数。SLM算法的准确性同时参考以上两个标准,它的联合准确率为

$$\text{prec} = \lambda \text{prec}_1 + (1 - \lambda) \text{prec}_2 \quad (6)$$

式中: $\lambda \in (0,1)$,考虑到选择图像的平衡性, λ 的取值为0.5。针对新图像的检测,存在几种传统方法:一种是基于层次分类树的方法^[14];另一种是支持向量机的方法^[15]。SLM模型的性能与传统的几种方法相比准确率略高,准确率比较结果如表1所示。传统方法选择自顶向下进行分类,一旦高层节点分类错误,则下层节点一定会分类错误。SLM算法通过从叶节点开始比较,可以准确地对细粒度对象类进行分类,从而进一步提高了图像集分类性能。

表1 不同方法下的分类准确率对比结果

Table 1 Classification accuracy under different methods %

方法	准确率
SLM	79.36
Hierarchical classification	78.72
One-class SVM	63.68

4 结束语

本文主要提出了半监督分层模型,以支持动态扩展图像层次结构。SLM模型通过学习VGGNet,利用标签已知图像集构建的图像层次结构,通过相似性比较与聚类,动态更新图像层次结构,提升层次结构的鲁棒性。本文选取的数据集均采用真实公开数据集,结果显示在小规模图像集中,SLM模型更具有优势。在标签已知图像层次结构中,主要参考了前期研究者的相关经验,实验采用的数据集范围还不广泛,在面对大型数据集时,各个参数的选取还有待改进。实验在相似性比较过程中,方式单一,后期将会继续研究。

参考文献:

- [1] ENGUEHARD J, O'HALLORAN P, GHOLIPOUR A. Semi-supervised learning with deep embedded clustering for image classification and segmentation[J]. *IEEE Access*, 2019, 7: 11093-11104.
- [2] FAN Jianping, PENG Jinye, GAO Ling, et al. Hierarchical learning of tree classifiers for large-scale plant species identification [J]. *IEEE Transactions on Image Processing*, 2015, 24(11): 4172-4184.
- [3] 赵其鲁, 李宗民. 基于深度多任务学习的层次分类[J]. *计算机辅助设计与图形学学报*, 2018, 30(5): 886-892.
ZHAO Qilu, LI Zongmin. Hierarchical classification based on deep multi task learning[J]. *Journal of Computer-Aided Design and Computer Graphics*, 2018, 30(5): 886-892.

- [4] PATINO-SAUCEDO A, ROSTRO-GONZALEZ H, CONRADT J. Tropical fruits classification using an AlexNet-type convolutional neural network and image augmentation[C]//Proceedings of International Conference on Neural Information Processing. Siem Reap, Cambodia: Springer, 2018: 371-379.
- [5] LEE S G, SUNG Y, KIM Y G, et al. Variations of AlexNet and GoogLeNet to improve Korean character recognition performance[J]. Journal of Information Processing Systems, 2018, 14(1): 205-217.
- [6] REN H Y, EL-KHAMY M, LEE J. DN-ResNet: Efficient deep residual network for image denoising[C]//Proceedings of Asian Conference on Computer Vision. Australia, Perth: Springer, 2018: 215-230.
- [7] LIU Xuan, CHI Mingmin, ZHANG Yunfeng, et al. Classifying high resolution remote sensing images by fine-tuned VGG deep networks[C]//Proceedings of IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium. Spain: IEEE, 2018: 7137-7140.
- [8] MARSZALEK M, SCHMID C. Constructing category hierarchies for visual recognition[C]//Proceedings of European Conference on Computer Vision. Berlin, Heidelberg: Springer, 2008: 479-491.
- [9] LIU B Y, SADEGHI F, TAPPEN M, et al. Probabilistic label trees for efficient large scale image classification[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. USA: IEEE, 2013: 843-850.
- [10] KUANG Zhenzhong, LI Zongmin, ZHAO Tianyi, et al. Deep multi-task learning for large-scale image classification[C]//Proceedings of 2017 IEEE Third International Conference on Multimedia Big Data (BigMM). [S.l.]: IEEE, 2017: 310-317.
- [11] FAN Jianping, ZHAO Tianyi, KUANG Zhenzhong, et al. HD-MTL: Hierarchical deep multi-task learning for large-scale visual recognition[J]. IEEE Transactions on Image Processing, 2017, 26(4): 1923-1938.
- [12] ZHAO Tianyi, ZHANG Baopeng, HE Ming, et al. Embedding visual hierarchy with deep networks for large-scale visual recognition[J]. IEEE Transactions on Image Processing, 2018, 27(10): 4740-4755.
- [13] ZHENG Yu, FAN Jianping, ZHANG Ji, et al. Hierarchical learning of multi-task sparse metrics for large-scale image classification[J]. Pattern Recognition, 2017, 67: 97-109.
- [14] FAN Jianping, ZHANG Ji, MEI Kuizhi, et al. Cost-sensitive learning of hierarchical tree classifiers for large-scale image classification and novel category detection[J]. Pattern Recognition, 2015, 48(5): 1673-1687.
- [15] CAMPBELL C, BENNETT K P. A linear programming approach to novelty detection[C]//Proceedings of Advances in Neural Information Processing Systems. USA: Kernel-Machines, 2001: 395-401.

作者简介:



王江晴(1964-),女,教授,研究方向:人工智能,E-mail: wjqing2000@mail.scuec.edu.cn。



张蕾(1995-),女,硕士研究生,研究方向:机器学习、数据挖掘,E-mail: 2017110240@mail.scuec.edu.cn。



孙翀(1981-),男,博士,研究方向:数据分析、数据挖掘和软件工程,E-mail: nicksun217@mail.scuec.edu.cn。



帖军(1976-),男,教授,研究方向:物联网系统应用、智能制造管理、智慧农业,E-mail: tiejun@mail.scuec.edu.cn。



周玮瑜(1995-),女,硕士研究生,研究方向:数据挖掘、机器学习,E-mail: 2017110236@mail.scuec.edu.cn。



孟凯(1995-),男,硕士研究生,研究方向:自然语言处理、机器学习,E-mail: 2018110260@mail.scuec.edu.cn。

(编辑:刘彦东)