

控制与决策

Control and Decision

融合稀疏编码与深度学习的草图特征表示

赵鹏, 高杰超, 冯晨成, 韩莉

引用本文:

赵鹏, 高杰超, 冯晨成, 等. 融合稀疏编码与深度学习的草图特征表示[J]. *控制与决策*, 2021, 36(3): 699–704.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2019.0941>

您可能感兴趣的其他文章

Articles you may be interested in

一种基于深度学习的时间序列预测方法

A time series prediction method based on deep learning

控制与决策. 2021, 36(3): 645–652 <https://doi.org/10.13195/j.kzyjc.2019.0809>

基于卷积神经网络的云雾遮挡舰船目标识别

Obscured ship target recognition based on convolutional neural network

控制与决策. 2021, 36(3): 661–668 <https://doi.org/10.13195/j.kzyjc.2019.0781>

结合注意力机制的循环神经网络复述识别模型

Recurrent neural networks based paraphrase identification model combined with attention mechanism

控制与决策. 2021, 36(1): 152–158 <https://doi.org/10.13195/j.kzyjc.2019.0638>

基于联合知识表示学习的多模态实体对齐

Multi-modal entity alignment based on joint knowledge representation learning

控制与决策. 2020, 35(12): 2855–2864 <https://doi.org/10.13195/j.kzyjc.2019.0331>

基于改进堆叠自动编码器的循环冷却水系统工艺介质温度预测控制方法

Predictive control method of process medium temperature in circulating cooling water system based on improved stacked auto encoders

控制与决策. 2020, 35(12): 2835–2844 <https://doi.org/10.13195/j.kzyjc.2019.0694>

融合稀疏编码与深度学习的草图特征表示

赵鹏[†], 高杰超, 冯晨成, 韩莉

(安徽大学 计算机科学与技术学院, 合肥 230601)

摘要: 针对小数据集下单纯使用深度学习方法的草图特征提取可分辨性低下的问题, 提出一种融合稀疏编码和深度学习的草图特征表示方法. 该算法首先对草图进行语义分割; 然后迁移深度学习方法, 分别提取草图特征和草图部件特征, 之后将部件特征降维聚类, 获取聚类中心; 最后利用部件聚类中心向量初始化稀疏编码中的字典, 交替迭代求得最终的草图特征. 不同于以往的草图特征表示方法, 将迁移深度学习获得的草图部件特征引入到稀疏编码中, 作为字典的初始基向量, 将语义信息融入到稀疏编码, 在提升草图特征表示性能的同时, 使得稀疏表示具有更好的可解释性. 实验结果显示, 所提方法下的草图识别率高于采用传统非深度学习和深度学习方法提取的草图特征的草图识别率.

关键词: 草图特征表示; 稀疏编码; 深度学习; 特征提取; 字典学习; 部件分割

中图分类号: TP273

文献标志码: A

DOI: 10.13195/j.kzyjc.2019.0941

开放科学(资源服务)标识码(OSID):



引用格式: 赵鹏, 高杰超, 冯晨成, 等. 融合稀疏编码与深度学习的草图特征表示[J]. 控制与决策, 2021, 36(3): 699-704.

A feature representation of sketch based on fusion of sparse coding and deep learning

ZHAO Peng[†], GAO Jie-chao, FENG Chen-cheng, HAN Li

(College of Computer Science and Technology, Anhui University, Heifei 230601, China)

Abstract: In order to overcome that the performance of sketch feature representation and recognition based on purely deep learning is not very well especially in limited sketch dataset, this paper proposes a feature representation of sketch based on the fusion of sparse coding and deep learning (FSCDL). Firstly, this method divides the sketch into components and extracts the features of sketch and sketch components with transfer deep learning. Then, it reduces the feature dimensions of the sketch and sketch components and clusters the sketch components. The cluster centers of sketch components are utilized to initialize the dictionary of sparse coding. Finally, the sketch feature representation is obtained by solving the objective function of sparse coding. Different from the previous work, this paper transfers deep learning to extract the features of sketches and sketch components, which are introduced to the sparse coding. The dictionary is initialized with the features obtained from the above transfer deep learning, which combines the semantic information obtained from deep learning and sparse coding. The proposed method not only improves the performance of the representation of the sketch, but also makes the sparse coding more interpretable. The experimental results show that the sketch recognition accuracy of the proposed method is higher than that of the traditional sketch feature representation methods and the sketch representation methods based on deep learning.

Keywords: sketch feature representation; sparse coding; deep learning; feature extracting; dictionary learning; component segmentation

0 引言

随着手写设备的普及, 手绘草图获取越来越便捷, 基于草图的各类应用不断涌现, 例如草图建模^[1]、交互式草图识别^[2]、基于草图语义的图像检索^[3]等.

目前已有的草图数据集规模十分有限, 尤其在特定领域的研究中, 有标注信息的草图数据集十分稀缺, 这使得基于草图部件与整体关系的草图特征表示存在很大的挑战.

收稿日期: 2019-07-03; 修回日期: 2019-09-09.

基金项目: 国家自然科学基金项目(61602004); 安徽省高校自然科学研究重点项目(KJ2018A0013, KJ2017A011); 安徽省自然科学基金项目(1908085MF188, 1908085MF182); 安徽省重点研究与开发计划项目(1804d08020309).

[†]通讯作者. E-mail: zhaopeng_ad@163.com.

稀疏编码是数学中的线性组合方法在计算机科学中的应用,在诸多研究领域都有广泛的应用.在稀疏编码中,构建字典的方法主要分成3类^[4-7]:1)基于概率的方法;2)基于向量量化或聚类的方法;3)基于特定结构的方法.在图像特征表示中,文献[8]提出的基于稀疏表示的分类(sparse representation based classification, SRC)算法,从理论上和实验上证明了稀疏表示对于提高高维度数据识别精度的必要性.

在草图识别领域,草图的特征提取方法主要分成两大类:1)手工设计草图特征^[9-11];2)利用深度学习提取草图特征^[12-15].基于手工特征提取方法虽然在某些数据集中取得的效果较好,但是手工特征提取费时费力,而且多数是基于局部的特征,缺少全局的特征链接.基于深度学习的特征提取方式在诸多模式识别领域取得了较大的成就.文献[12]对 Alex-Net^[16]的结构进行了微调,使之适合 sketch 分类;文献[13]结合数据集的特点,提取时序的子图,分别训练网络,最后采用特征融合;文献[14]提出了一种多尺度多通道的深度卷积神经网络,并结合贝叶斯分类器处理多尺度的融合,从而提高识别度.

草图作为人类对现实世界物体的一种抽象描述,是由一个个线条笔画构建而成,其本身就具有一定的稀疏性.人类在画一幅草图的时候,首先会概括物体具有的主要部件,然后对这些主要部件进行抽象表示,再进行某种形式的组合.因此,草图由各个部件组成,可以看作是由各个部件的线性组合.而稀疏编码的过程就是将字典中的基向量进行线性组合,从而近似表示数据样本,这种稀疏表示的过程与人类抽象构建草图的过程具有一定的相似性.因此本文提出用迁移深度学习获得的部件特征来初始化字典,融合稀疏编码和深度学习来获取草图特征表示.由于已标注草图数据集较小,尤其在某些特定领域,无法直接利用深度学习方法,将自然图像中训练学习到的深度学习模型迁移到草图和草图部件集,然后用深度学习获得的部件特征来初始化稀疏编码字典,从而使得稀疏编码中的字典具有较好的可解释性,提高算法的稳定性.

本文提出的融合稀疏编码与深度学习的草图特征表示方法(a feature representation of sketch based on the fusion of sparse coding and deep learning, FSCDL),首先将草图分割为语义部件,利用迁移深度学习,获取草图和部件的特征表示,再利用迁移深度学习获得的草图部件特征初始化稀疏编码模型中的字典,最后交替迭代求解稀疏编码模型,获取最终的草图特征表

示.

1 本文方法

1.1 基于深度学习的草图及部件特征提取

现有已标注草图集通常规模较小,尤其是在一些特定领域,无法用深度学习直接训练模型.由于草图数据集可以看作是自然图像的一个抽象表示,本文通过使用迁移大规模自然图像训练出来的 AlexNet 的网络结构,提取草图的整体和草图部件特征.本文将迁移而来的 AlexNet 进行微调,采用的微调策略为:对低层特征图的卷积核权值不作调整,对中、高层卷积核进行微调,同时将最后3层全连接层替换为新的全连接层.全连接层逐层剔除冗余特征,将特征进一步的精纯化,以方便分类.

草图是由多个具有语义概念的部件组成的,草图语义分割是为了更好地细分部件,提高计算机对图像的理解.草图部件特征就是草图分解后的各个部件图的特征.在草图图像分割领域,由于图片所含信息量非常少,无法使用自然图像中的分割方法,对其进行语义分割难度非常大.文献[17]提出了一种数据驱动的草图分割标注算法,将草图拆分到笔画层,再将笔画的序列进行组合,构成笔画组,并与预构建的3D数据库中的3D图形进行各个部件的迭代比对,最终可以得到基于笔画层的 sketch 草图的分割标注.

在文献[17]提供的草图部件集(sketch_dataset)上,本文进行如下处理:首先对 sketch_dataset 中提供的基于笔画层的标注信息进行预处理,然后按标签语义及最大连通性原则,将 sketch_dataset 中的草图分割成一张张部件图.具体处理方法为:假设有两个笔画 stroke1 和 stroke2,它们对应的标签为 label₁ 和 label₂,如果 label₁ = label₂,并且 stroke1 的终点坐标正好是 stroke2 的起点坐标,则认为它们是同属于一个部件中的笔画序列,否则,stroke1 和 stroke2 分属不同的部件.对于部件特征的提取,本文同样迁移 AlexNet 网络结构,并使用与草图整体特征提取的操作一致的微调方式微调网络参数.

1.2 基于部件字典的稀疏编码模型

首先使用 PCA 方法,将迁移深度学习获取的草图和草图部件特征维度降至 256 维,然后再采用 Kmeans 聚类方法对所有的部件特征进行聚类,每个聚类中心作为稀疏编码中部件字典的一个向量,构造并初始化部件字典.

本文基于部件字典的稀疏编码模型的目标函数如下所示:

$$\begin{aligned} \min_{B,S} & \|X - BS\|_F^2 + \alpha \text{tr}(SLS^T) + \lambda \|S\|_1; \\ \text{s.t.} & \|B\|_F^2 \leq v. \end{aligned} \quad (1)$$

其中: X 为草图特征矩阵; B 为草图部件特征组成的部件字典矩阵, 其初始值为部件特征的聚类中心; S 为草图特征在该部件字典下的稀疏编码. 本文引入拉普拉斯矩阵 L 来保留草图数据集的几何特性, 使得草图稀疏编码符合同一结构分布. 根据给定的 m 个样本, 构建含有 m 个顶点的最邻近图 G , W 为 G 的权重矩阵. 如果在 G 中, x_i 与 x_j 是最邻近关系, 则 $W_{ij} = 1$, 否则 $W_{ij} = 0$. x_i 的度为 $d_i = \sum_{j=1}^m W_{ij}$, 记 $D = \text{diag}(d_1, d_2, \dots, d_m)$, 则图拉普拉斯矩阵 $L = D - W$.

本文使用交替迭代求解法求解该目标函数(1), 交替迭代求解过程如下:

首先, 固定字典 B , 目标函数为

$$f = \min_S \|X - BS\|_F^2 + \alpha \text{tr}(SLS^T) + \lambda \|S\|_1. \quad (2)$$

然后, 固定稀疏表示为 S , 微调字典为 B , 目标函数为

$$\begin{aligned} f &= \min_B \|X - BS\|_F^2; \\ \text{s.t.} & \|B\|_F^2 \leq v. \end{aligned} \quad (3)$$

1.3 融合稀疏编码和深度学习的草图特征表示方法

本文提出的融合稀疏编码与深度学习的草图特征表示方法(FSCDL)框架如图1所示.

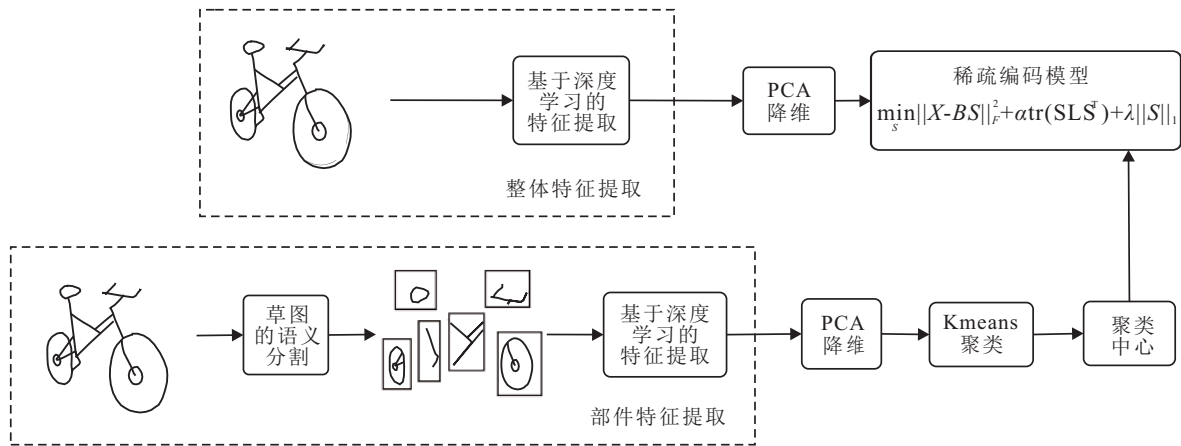


图1 FSCDL框架结构

本文算法流程如下.

输入: 完整的草图样本 I , 草图 I 对应的部件图 I_{comp} , 图拉普拉斯矩阵 L , 正则化参数 α 、 λ , 迭代次数 n .

输出: 草图 I 的稀疏编码特征表示 S .

step 1: 读入图像, 利用式(2)提取强度分量;

step 2: 使用草图图像 I 微调 AlexNet, 得到草图完整图像微调后的网络结构 AlexNet_Whole;

step 3: 使用部件图 I_{comp} 微调 AlexNet, 得到部件图像微调后的网络结构 AlexNet_Comp;

step 4: 使用 AlexNet_Comp 网络结构下的前馈运算, 提取部件图 I_{comp} 的部件特征 $I_{comp_feature}$;

step 5: 将草图部件特征 $I_{comp_feature}$ 进行聚类, 用聚类中心初始化部件字典 B ;

step 6: 交替迭代求解 S

for $k = 1$ to n

1) 根据式(2)优化求解稀疏表示 S ,

2) 根据式(3)微调部件字典 B ,

end for

2 实验结果与分析

2.1 数据集和实验设置

sketch_dataset^[17] 是目前主流的草图分割数据集, 广泛应用于草图分割以及草图检索中. 该数据集包含 10 个类别, 每个类别各 30 张草图, 共 300 张草图以及对应于每张草图的 xml 描述文件, 该文件中记录了每一笔笔划的信息, 包括点坐标以及当前笔画的标签. 本文将 10 个类别共 300 张图像按照 1.1 节所述的数据预处理方法, 将草图笔划组合成一张张部件图, 共得到 2 374 张部件图, 64 个部件类别. 其中: 飞机类别有 6 个部件类, 自行车类别共有 9 个部件类, 烛台类别共有 6 个部件类, 椅子类别共有 11 个部件类, 4 条腿的动物类别共有 5 个部件类, 人类别共有 6 个部件类, 灯类别共有 3 个部件类, 步枪类别共有 7 个部件类, 桌子类别共有 7 个部件类, 花瓶类别共有 4 个部件类. 本文对完整的草图(300 张)以及部件图(2 374 张)分别采用旋转和裁剪的方式进行数据增强.

在本实验中, 为了保证草图样本以及部件样本在经过 AlexNet 微调之后, 依旧保持一致的样本特性, 对

于草图样本和部件样本均迁移 AlexNet, 并分别微调, 在微调过程中, 采用一致的微调策略和迭代次数. 对于新加入的全连接层, 采用随机初始化其参数的方法, 迭代的次数均为 20 次, 这是因为草图样本和部件样本在 20 次迭代之后模型均处于一个稳定的状态. 实验随机选择 3/4 的草图特征样本作为基于草图的稀疏编码的训练集, 剩余 1/4 的草图特征样本作为测试集. 特征的分类采用的是 SVM 分类器.

2.2 实验结果与分析

本文设计 4 组实验, 用以验证所提出的融合草图部件和整体关系的草图特性提取方法的有效性: 第 1 组实验验证本文提出的基于部件字典的方法与不采用部件特征的随机字典方法, 对模型分类精度的影响; 第 2 组实验验证不同参数值对本文提出的模型分类精度的影响; 第 3 组实验验证部件字典的基数取不同值时对模型分类精度以及时耗的影响; 第 4 组实验是本文提出的方法同其他 6 种主流特征表示方法的对比, 用以验证本文所提出方法的有效性.

第 1 组实验分别使用迁移后的 AlexNet 的网络不同层输出作为模型的输入特征. 该迁移后的 AlexNet 的网络参数是在草图集上训练微调而得. 本文分别选取卷积层后不同的全连接层的特征作为模型的输入, 设置模型参数为 $\alpha = 0.1$, $\lambda = 0.01$, 迭代次数 $n = 50$. 首先使用 PCA (principal component analysis) 对所有迁移深度学习获得的特征向量降维至 256 维, 然后输入到稀疏编码模型中, 实验结果如表 1 所示.

表 1 不同层特征对模型分类准确率的影响

特征维数	特征层	基于随机字典下的分类准确率 / %	基于部件字典下的分类准确率 / %
4096	$f_{1,1}$ 层(conv 后)	66.67	69.33
	$f_{1,2}$ 层(relu 后)	73.33	74.67
1024	$f_{2,1}$ 层(conv 后)	92.00	94.67
	$f_{2,2}$ 层(relu 后)	93.33	94.67
	$f_{2,3}$ 层(dropout 后)	91.47	96.00
512	$f_{3,1}$ 层(conv 后)	93.33	98.67
	$f_{3,2}$ 层(relu 后)	93.60	97.33
	$f_{3,3}$ 层(dropout 后)	91.95	98.67

在表 1 中: $f_{1,1}$ 层为第 1 个全连接层的输出, $f_{1,2}$ 层为 $f_{1,1}$ 层经过 relu 后的输出, $f_{2,1}$ 层为第 2 个全连接层的输出, $f_{2,2}$ 层为 $f_{2,1}$ 层经过 relu 之后的结果, $f_{2,3}$ 层为 dropout 值为 0.5 后的输出结果, $f_{3,1}$ 层为第 3 个全连接层的输出, $f_{3,2}$ 层为 $f_{3,1}$ 层经 relu 的结

果, $f_{3,3}$ 层为 dropout 之后的结果. 表 1 中: 第 1 列为迁移深度学习获得的特征向量的维数, 第 2 列为输入特征迁移深度学习的特征层, 第 3 列为不同输入特征下不使用部件特征的随机字典的稀疏编码的草图分类正确率, 第 4 列为本文提出的基于部件字典的稀疏编码的草图分类正确率.

由表 1 可知, 随着迁移深度学习选择的特征层维数的降低, 模型分类准确率呈上升趋势, 当特征维数为 4096 维, 1024 维, 512 维时, 本文方法的平均分类准确率分别为 72%、95.11% 和 98.22%. 当输入特征为 4096 维时, 在 PCA 降维过程中, 维数压缩过大, 丢失了部分非冗余特征, 导致分类正确率较低. 当特征维度为 4096 时, $f_{1,1}$ 层与 $f_{1,2}$ 层特征下, 相对于不使用部件特征的随机字典方法, 本文提出的基于部件字典的方法的分类准确率分别高出 2.66% 和 1.34%. 当特征维度为 1024 时, $f_{2,1}$ 层、 $f_{2,2}$ 层以及 $f_{2,3}$ 层特征下, 相对于不使用部件特征的随机字典方法, 本文提出的基于部件字典的方法的分类准确率分别高出 2.67%、1.34% 和 4.53%. 当特征维度为 512 时, $f_{3,1}$ 层、 $f_{3,2}$ 层以及 $f_{3,3}$ 层特征下, 相对于随机字典方法, 本文提出的基于部件字典的方法的分类准确率分别高出 5.34%、3.73% 和 6.72%. 说明本文提出的基于部件的字典方法相较于不使用部件特征的随机字典方法, 能够更好地表达草图集的可分辨语义信息, 辅助完整的草图图像获得更好的分类效果.

第 2 组实验是验证稀疏编码模型不同参数取值对模型性能的影响. 本文分别调整了图正则化参数 α , 以及稀疏正则化参数 λ , 用来验证各参数对于模型的影响, 实验结果如图 2 所示.

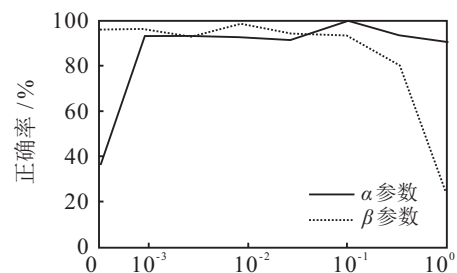


图 2 α 和 λ 对模型分类准确率的影响

首先, 实验中设置稀疏正则化参数 $\lambda = 0.01$, 迭代次数 $n = 50$, α 的取值为 $\{0, 0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1\}$. 当 $\alpha = 0$ 时, 即没有保留草图集的流形结构时, 忽略了原数据集的几何特性, 模型的识别能力低下, 分类准确率为 37.33%. 而当 α 取值逐渐增大时, 新特征空间下数据集原几何特性的保留在最小化目标函数中所占的权重逐渐增大, 模型分类准确率

逐步提高,当 $\alpha = 0.1$ 时,模型的分类准确率达到最大值98.67%,当 α 取值继续增大时,原几何特性的保留在最小化目标函数中所占的权重逐渐增大,稀疏正则化所占比重过小,导致模型的分类准确率逐步下降.然后,实验中设置图正则化参数 $\alpha = 0.1$,迭代次数 $n = 50$, λ 的取值为 $\{0, 0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1\}$.当 $\lambda = 0$ 时,模型分类准确率为94.67%,这是由草图本身的稀疏性决定的;而当 $\lambda = 1$ 时,稀疏度惩罚占的权重过大,忽略了原数据集的几何特性,导致分类正确率下降;当 $\lambda = 0.01$ 时,分类正确率为98.67%,相较于 $\lambda = 0$ 时的分类正确率提升了4%,表明此时的参数设置既能较好地保留原草图集的几何特性,又较好地保留了稀疏性,导致获得较高的特征表示效果.

第3组实验是验证字典基数对模型性能的影响.由于本文稀疏编码方法是基于部件字典的,而部件字典是由Kmeans聚类中心构造的,本实验选取的字典基数就是聚类中心的个数,取值分别设为300,500,700,900,1000,实验结果如表2所示.表2中第2列为在相应字典基数下获得的分类正确率,第3列为不同的字典基数对于稀疏编码每一次的迭代计算需要的时间.

表2 字典基数不同取值下的分类正确率

字典基数/个	分类正确率/%	耗时/s
300	93.33	8.92
500	94.67	22.32
700	98.67	38.69
900	97.33	112.39
1000	98.67	166.39

由表2可知,随着字典基数的增加,分类准确率逐渐增加,同时稀疏编码求解时每次迭代的耗时也逐渐增加.这是由于当字典基数较少时,占内存空间小,计算量也小,但基数过小时,不能很好地将不相似的部件分开,导致基于部件的字典分类正确率较低.字典基数增加可以更好地将相似的部件聚为同一类,不相似的部件分为不同类,但是也给内存计算等带来压力,导致耗时成倍增加.当字典基数增至700时,分类准确率基本稳定,但耗时增幅很大.当字典基数为500时,相较于基数300,分类正确率增长了1.34%,耗时增加了约2.5倍;当字典基数为700时,相较于基数为500时的分类正确率增长了4%,耗时增加了约1.73倍;而当字典基数为900时,它相较于基数为700时,分类正确率下降了1.34%,耗时却增加了

约2.9倍;当字典基数为1000时,它相较于基数为900时,分类正确率虽然增长了1.34%,但是耗时增加了约1.48倍,而字典基数为1000时的识别精度与字典基数为700时的分类正确率持平,耗时却是4.3倍.因此,为了平衡分类准确率与时耗,字典基数可以设置在500~700之间.

第4组实验是验证本文所提FSCDL方法的有效性.分别将本文方法与相关的6种主流草图特征表示方法进行对比,这6种方法分别为:HOG-SVM^[18],SIFT-Fisher Vector^[10],Deep-sketch^[19],sketch-a-net^[14],SC^[20]和GraphSC^[21],实验结果如表3所示.

表3 本文提出的方法同主流草图识别方法的对比

实验方法	准确率/%
HOG-SVM	56.00
SIFT-Fisher Vector	61.50
Deep-sketch	69.15
sketch-a-net	74.67
SC	88.10
GraphSC	93.33
FSCDL	98.67

由表3可知,本文提出的方法FSCDL获得的分类正确率比传统的草图识别方法HOG-SVM、SIFT-Fisher Vector分别增长了42.67%,37.17%.HOG-SVM方法使用方向梯度直方图提取草图的特征描述子,再使用SVM进行分类识别.SIFT-Fisher Vector方法使用尺度不变特征和费舍尔向量的局部不变性提取草图特征.由于深度学习通常采用的是将完整的草图作为网络的输入,忽略了草图部件的重要性,而且在小数据中,单纯使用深度学习的草图识别效果并不理想.Deep-sketch利用7个卷积层和3个全连接层的深度网络对草图图像进行识别分类,sketch-a-net利用一种多尺度多通道的深度卷积神经网络,并结合贝叶斯分类器处理多尺度的融合.实验结果显示,本文提出的FSCDL方法获得的分类正确率比基于深度学习的方法Deep-sketch、sketch-a-net分别增长了29.52%,24%.实验表明:本文使用部件特征辅助完整草图的识别具有较好的表达效果,尤其在小数据集领域中,FSCDL融合了深度学习与稀疏编码,获得了更高的分类正确率;SC仅考虑了稀疏性,字典是随机初始化的;GraphSC不仅考虑了稀疏性,还保留了原始特征空间下的几何特性,但字典仍然是随机初始化的.本文提出的FSCDL中稀疏编码字典为部件字典,更好地融合了草图的语义信息.FSCDL的分类准确

率比SC和GraphSC分别高出10.57%和5.34%。实验验证了本文提出的融合稀疏编码和深度学习的草图特征表示方法的有效性。

3 结论

深度学习在大数据集样本学习中获得了非常好的效果,然而,对于一些小数据集的应用领域,无法直接利用深度学习进行训练学习。传统的稀疏编码字典主要是采用随机初始化,可解释性较差,且稳定性不高。本文针对小数据集下的草图识别,结合深度学习与稀疏编码的方法,利用深度学习对图片的强感知获取图片的整体特征;同时将草图分割为语义部件,利用迁移深度学习获取其部件特征。将部件语义信息引入到稀疏编码中,构建基于部件字典的稀疏编码模型,实验验证了所提出的特征表示方法具有更好的可分辨性。

参考文献(References)

- [1] Ding C, Liu L G. A survey of sketch based modeling systems[J]. *Frontiers of Computer Science*, 2016, 10(6): 985-999.
- [2] Renáková M, Tencer L, Cheriet M. Online handwritten gesture recognition based on Takagi-Sugeno fuzzy models[C]. *International Conference on Information Science, Signal Processing and Their Applications*. Montreal: IEEE, 2012: 1247-1252.
- [3] Xu P, Yin Q, Huang Y Y, et al. Cross-modal subspace learning for fine-grained sketch-based image retrieval[J]. *Neurocomputing*, 2018, 278(2): 75-86.
- [4] Toi I, Frossard P. Dictionary learning[J]. *IEEE Signal Processing Magazine*, 2011, 28(2): 27-38.
- [5] Olshausen B A, Field D J. Sparse coding with an overcomplete basis set: A strategy employed by V1?[J]. *Vision Research*, 1997, 37(23): 3311-3325.
- [6] Schmid-Saugeon P, Zakhor A. Dictionary design for matching pursuit and application to motion-compensated video coding[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2004, 14(6): 880-886.
- [7] Yaghoobi M, Daudet L, Davies M E. Parametric dictionary design for sparse coding[J]. *IEEE Transactions on Signal Processing*, 2009, 57(12): 4800-4810.
- [8] John wright J, Yang A Y, Ganesh A, et al. Robust face recognition via sparse representation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(2): 210-227.
- [9] Eitz M, Hildebrand K, Boubekeur T, et al. Sketch-based image retrieval: Benchmark and bag-of-features descriptors[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2011, 17(11): 1624-1636.
- [10] Schneider R G, Tuytelaars T. Sketch classification and classification-driven analysis using fisher vectors[J]. *ACM Transactions on Graphics*, 2014, 33(6): 1-9.
- [11] Li Y, Hospedales T M, Song Y Z, et al. Free-hand sketch recognition by multi-kernel feature learning[J]. *Computer Vision and Image Understanding*, 2015, 137: 1-11.
- [12] Seddati O, Dupont S, Mahmoudi S. Deepsketch: Deep convolutional neural networks for sketch recognition and similarity search[C]. *International Workshop on Content-Based Multimedia Indexing*. Prague: Czech Republic, 2015: 1-6.
- [13] Dupont S, Seddati O, Mahmoudi S. Deepsketch 2: Deep convolutional neural networks for partial sketch recognition[C]. *International Workshop on Content-Based Multimedia Indexing*. Bucharest: IEEE, 2016: 1-6.
- [14] Yu Q, Yang Y X, Song Y Z, et al. sketch-a-net that beats humans[EB/OL]. (2015-01-30)[2015-07-21]. <http://arxiv.org/abs/1501.07873v3>, 2015.
- [15] Yu Q, Yang Y, Liu F, et al. Sketch-a-net: A deep neural network that beats humans[J]. *International Journal of Computer Vision*, 2017, 122(3): 411-425.
- [16] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]. *International Conference on Neural Information Processing Systems*. Lake Tahoe: Curran Associates Incorporation, 2012: 1097-1105.
- [17] Huang Z, Fu H B, Lau R W H. Data-driven segmentation and labeling of freehand sketches[J]. *ACM Transactions on Graphics*, 2014, 33(6): 1-10.
- [18] Eitz M, Hays J, Alexa M. How do humans sketch objects?[J]. *ACM Transactions on Graphics*, 2012, 31(4): 1-10.
- [19] 赵鹏, 王斐, 刘慧婷, 等. 基于深度学习的手绘草图识别[J]. *四川大学学报: 工程科学版*, 2016, 48(3): 94-99. (Zhao P, Wang F, Liu H T, et al. Sketch recognition using deep learning[J]. *Journal of Sichuan University: Engineering Science Edition*, 2016, 48(3): 94-99.)
- [20] Lee H, Battle A, Raina R, et al. Efficient sparse coding algorithms[C]. *International Conference on Neural Information Processing Systems*. Vancouver: Institute of Electrical and Electronics Engineers Incorporation, 2007: 801-808.
- [21] Zheng M, Bu J, Chen C, et al. Graph regularized sparse coding for image representation[J]. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, 2011, 20(5): 1327-1336.

作者简介

赵鹏(1976—),女,副教授,博士,从事机器学习、图像理解等研究, E-mail: zhaopeng_ad@163.com;
高杰超(1995—),男,硕士生,从事图像标注、机器学习的研究, E-mail: 1019069117@qq.com;
冯晨成(1993—),女,硕士生,从事图像标注、机器学习的研究, E-mail: 767254498@qq.com;
韩莉(1975—),女,讲师,博士,从事机器学习的研究, E-mail: hanli98@ahu.edu.cn.

(责任编辑: 闫妍)