

引用格式:王梓舟,赵海英,任文超.基于多层字典学习的传统服饰图像标注算法[J].中国传媒大学学报(自然科学版),2022,29(04):02-07+18.

文章编号:1673-4793(2022)04-0002-07

基于多层字典学习的传统服饰图像标注算法

王梓舟¹,赵海英^{2*},任文超¹

(1.北京邮电大学计算机学院,北京 100876;2.北京邮电大学人工智能学院,北京 100876)

摘要:中国传统服饰图像是中华优秀传统文化重要的组成部分,图像内涵丰富,时空跨度大,但理解存在歧义,急需一套语义解读方法,而大量图像标注算法主要关注在各自的垂直领域,传统服饰图像仍然面临着标注精度亟需提高的挑战。本文以中国传统服饰图像作为研究对象,以字典学习多标签标注方法作为研究方法,提出了融合深度多层结构框架的多标签字典学习算法,通过结合字典学习与多层结构框架来提高标注性能。最后通过对比实验验证了该思路的正确。

关键词:传统服饰图案;多标签标注;多层字典学习;字典相关性

中图分类号:TP391

文献标识码:A

Traditional clothing image annotation algorithm based on multi-layer dictionary learning

WANG Zizhou¹, ZHAO Haiying^{2*}, REN Wenchao¹

(1. School of Computer, Beijing University of Post and Telecommunication, Beijing 100876, China;

2. Artificial Intelligence Institute, Beijing University of Post and Telecommunication, Beijing 100876, China)

Abstract: Chinese traditional clothing image is an important part of Chinese excellent traditional culture. The image connotation is rich, the space-time span is large, and the understanding is ambiguous. There is an urgent need for a set of semantic interpretation methods for it. A large number of image annotation algorithms mainly focus on their respective vertical fields, and traditional clothing images are still facing the challenge of improving the annotation accuracy. This paper takes Chinese traditional clothing images as the research object, takes dictionary learning multi label tagging method as the research method, and proposes a multi label dictionary learning algorithm integrating deep multi-layer structure framework, which improves the tagging performance by combining dictionary learning and multi-layer structure framework. Finally, the correctness of this idea is verified by comparative experiments.

Keywords: traditional dress patterns; multi label labeling; multi level dictionary learning; dictionary relevance

1 引言

中华民族作为一个拥有五千年历史的民族,在漫长的历史长河形成了辉煌的中华文化,不仅是中华民族存在的证明,更是有别于其他民族的重要标识。在数字智

能时代,亟需我们向世界弘扬中华传统文化。如何精炼地提取并标识文化内涵,继而去发展和弘扬,受到越来越多的关注。中国传统服饰图像是中华优秀传统文化重要的组成部分,是在中国历史过程中逐步发展出来的,

基金项目:揭榜挂帅重点研发课题(课题编号:2021YFF0901701)

作者简介(*为通讯作者):赵海英(1972-),女,副教授,博士,主要从事文化计算与媒体数据挖掘研究。Email:zhaohaiying@bupt.edu.cn

具有继承性、稳定性的一种文化符号。因此,研究中国传统服饰图像,有助于中华优秀传统文化的研究和发展。

然而,传统的图像标注法大多是通过人工观察数据集的规律来确定模型,没有利用数据集的标签信息,所以往往只是在单一类别的数据上能取得比较好的分类效果。随着应用场景的复杂化,分类算法中二元分类算法逐步发展为多元分类算法和多标签分类算法。其中,二元分类问题是指判断某样本是否属于某个具体的类别,其结果只有两种:是或者不是。多元分类问题,则将类别扩大为多个,即判断某样本属于多个类别中的哪一个。而多标签分类算法,指某个事物同时属于多个类别,比如一张风景图,可能包含小鸟、天空、树等多种元素,因此标注的时候需要给该图片打上多个标签。

本文通过多标签分类算法对中国传统服饰的图像进行分类,以解决目前中国传统服饰图像标注中存在的问题。

2 相关工作

将字典学习算法用于图像分类的应用上,通常有两种思路:第一是通过样本在不同字典上的重建误差来进行分类;第二种思路则是利用样本在统一字典上的重建系数来进行分类。在第一种方法中,算法模型将为每个类别分别学习一个对应的字典。该字典将只包含该类图像的特征信息,因此当本类样本图像在其对应的字典上进行重建时,其重建误差将会很小。而当该样本在非对应的字典上进行重建时,其重建误差将会很大。因此可以根据该重建误差的大小,来判断该样本属于哪个字典对应的类别。重建误差的大小,代表的是该样本属于该类别的可能性大小。误差越小,可能性越大。基于重建误差来进行分类的字典学习算法流程如图1所示:

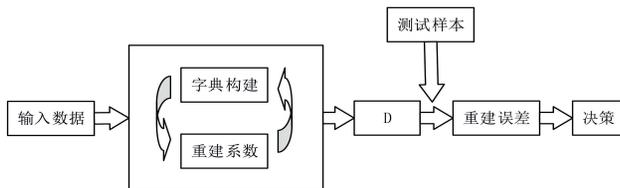


图1 基于重建误差的字典学习

在第二种方法中,所有关于字典学习(Dictionary Learning, DL)的先前研究都是“浅层的”学习模型,使用单层字典的学习框架学习过程中,计算机视觉领域的研究人员观察到,可以通过更深层次的结构得到更好的表示(更抽象和紧凑)。深信网络(Deep Belief Network, DBN)由一台受限玻尔兹曼机堆叠(Restricted Boltzmann Machine, RBM)在另一台RBM之后而成^{[1][2]},类似地,堆

叠自动编码器(Stacked Autoencoder, SAE)由一个自动编码器(Auto Encoder, AE)在其他AE之后而成^{[3][4]}。由于DBN和SAE成功,多层结构优化的字典学习算法也被纷纷提出,在图像分类应用中收获了很大的成功。

Snigdha^[5]提出,将图2左边的浅层字典学习结构拓展为图2右边的多层字典学习结构。

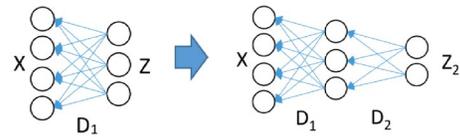


图2 将浅层字典学习拓展到多层字典学习

学习这两个字典以及最深层次的特征是一个难题,原因有两个:

(1)字典学习是一个双线性(因此是非凸的)问题。学习多层字典以及特征使问题更加难以解决,单层字典学习的收敛证明很难复制到多层结构。

(2)当同时学习多层字典时,需要求解的参数数量会增加。由于训练数据有限,这可能会导致过度拟合。

因此以贪婪的方式来学习字典,如图3所示:

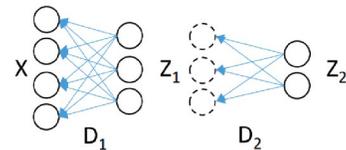


图3 用贪婪的方式学习多层字典

在这些网络中,在前面的字典学习层中并不需要使用稀疏编码,仅在最后一层字典学习层需要使用稀疏编码。例如,对于两个字典学习层的结构,第一层字典学习层的编码矩阵将会是稠密的,而第二层得到的编码矩阵则是稀疏的。两个字典学习层并不能坍塌为一个字典学习层,因为其整个学习过程并不是线性的。假设样本特征维数是 m ,第一个字典大小为 $m * n_1$,第二个字典大小为 $n_1 * n_2$,那么不可能学习到一个等价的 $m * n_2$ 字典。其中, n_1 表示第一层字典原子个数, n_2 表示第二层字典原子个数。

3 深度结构优化的多标签字典学习算法

3.1 浅层多标签字典学习算法及其缺点

D-KSVD^[6]是一种在字典学习过程中优化线性分类器的方法,它使用了带有重建约束项的类别标签,并通过KSVD算法求解其函数。其目标函数如下:

$$\min_{D, W, A} \|X - DA\|_F^2 + \lambda \|A\|_1 + \mu \|H - WA\|_F^2 \quad (1)$$

其中 $X \in \mathbf{R}^{M \times N}$ 是由所有训练样本组成的数据矩阵,

$D \in \mathbf{R}^{M \times K}$ 是该算法学习得到的字典, $A \in \mathbf{R}^{K \times N}$ 是编码向量矩阵, $H \in \mathbf{R}^{C \times N}$ 是关于训练样本的标签矩阵, $W \in \mathbf{R}^{C \times K}$ 则是分类器参数, λ 和 μ 均为正则化参数。

该函数可以用 KSVD 算法求解并同时学习字典和分类器。LC-KSVD 可以看作是 D-KSVD 方法的扩展, 在学习过程中进一步引入了判别码错误项。因此, 学习到的编码向量的判别能力明显提高。

浅层结构的字典学习算法有一些固有的缺陷:

(1) 由于数据收集的成本较大, 在实际应用中算法训练集的规模通常是有限的。对于单层架构, 字典的表示能力通常不是最优的, 它的原子经常受到训练集中噪声的影响。

(2) 特征空间中单层编码向量可能不包含真正的

$$\begin{aligned} \min_{D_m, G_m, A_m, W, b} & \|X - D_1 A_1\|_F^2 + \alpha_1 \text{Tr}(A_1^T G_1 A_1) + \dots + \|A_{l-1} - D_l A_l\|_F^2 \\ & + \alpha_l \text{Tr}(A_l^T G_l A_l) + 2\lambda \sum_{c=1}^C \mathcal{L}(A_l, y_c, w_c, b_c) \\ \text{s.t. } & \psi(D_m), (m = 1, 2, \dots, l) \end{aligned} \quad (2)$$

其中 $y_c (c = 1, 2, \dots, C)$ 表示来自 c 类的训练样本的类标签。上述函数中, $\|X - D_1 A_1\|_F^2$ 为第一层重构误差, 第二项表示局部性约束项, 其中图拉普拉斯矩阵 $G_1 \in \mathbf{R}^{K_1 \times K_1}$ 继承了原始样本的流形结构。因此, 它的第一项和第二项对第一层的局部约束下的重建进行编码。根据这两项, 可以重写下一层的等价项, 直到第 l 层。其中图拉普拉斯矩阵 $G_m \in \mathbf{R}^{K_m \times K_m}$ 是基于字典 $D_m \in \mathbf{R}^{K_{m-1} \times K_m}$ 构建的。 $\psi(D_m)$ 表示字典 D_m 的每一个原子满足二范数约束。

还有很多其他深度字典学习算法, Chun^[8] 提出了一种使用 Majorizer 进行卷积字典学习的 Block Proximal Gradient 方法。Hu^[9] 提出了一种非线性字典学习方法并将其应用于图像分类任务。Xiao^[10] 提出了一种用于对象识别任务的两层局部坐标编码框架。Zhang^[11] 引入了一种分析判别式用于图像分类任务的字典学习框架。Nguyen^[12] 提出了一个使用稀疏和分层网络的域适应框架, 该框架与 DDLCN^[13] 有着类似的思路。

多层结构的字典学习算法有几个重要的优点:

(1) 由于数据收集的成本较大, 在实际应用中数据通常是有限的。对于单层架构, 字典的表示能力通常不是最优的, 它的原子经常受到训练集中噪声的影响。这种负面影响在多层词典中会减少。

(2) 特征空间中单层编码向量可能不包含真正的分类信息, 而多层编码向量可以给出一些更好的近似值。

分类信息, 例如两个编码向量类别在单层中可能不是线性可分的。

3.2 多层字典学习算法以及优势

MDDL^[7] 算法是以多层字典学习层的结构来进行学习。其每一个字典学习层设为 $X \approx D_l A_l$ 。理论上, 每层重构误差的概率分布为相同, 所以可以得到整个重建误差为 $\|X - D_1 A_1\|_F^2 + \dots + \|A_{l-1} - D_l A_l\|_F^2$ 。最后一层编码向量使用分类器进行分类, 这里采用多类 SVM 分类器。使用 one-vs-all 策略学习 C 个超平面 $W = [w_1, w_2, \dots, w_C] \in \mathbf{R}^{K_l \times C}$ 和偏差 $b = [b_1, b_2, \dots, b_C] \in \mathbf{R}^{1 \times C}$ 。为了保留数据的局部性信息, 基于图拉普拉斯矩阵学习字典上的约束被用作编码向量 (A_1, A_2, \dots, A_l) 的约束。

例如, 两个编码向量类别在单层中可能是非线性可分的, 但它们可以通过更深层的结构转化为近似线性可分。

(3) 利用监督信息, 进一步提高学习编码向量的可辨别性。

(4) 后层的编码向量可以看作是前层的稀疏表示。这种机制可以鼓励来自同一类的样本具有更多相似的编码向量, 并降低分类器对少数几个的敏感度数据点。

3.3 融合多层结构框架的多标签字典学习算法

本文充分利用多层字典学习分类应用上的各种优势, 在 DDLCN 框架上将其调整改进, 使其与 ERCDLSI 结合, 从而更高效的解决多标签分类问题。

DDLCN 网络的结构如图 4 所示。

该框架具有大部分标准深度学习层(例如输入/输出层、池化层、全连接层等), 但基本的卷积层被替换为字典学习层和编码层。字典学习层学习一个过完备字典输入训练数据。在深度编码层, 局部性约束保证激活的字典原子彼此靠近。

经过上述步骤, 算法可以从浅字典层学习边缘、线和角表示, 更加复杂的“分层”特征表示可以从更深的字典层中学习。由于引入了字典学习和编码层, 算法对输入有更好的逼近能力, 它利用了流形几何结构, 将来自底层数据流形的数据点嵌入到低维深层结构空间中。算法可以将一个非常困难的非线性学习问题转化为一个更简单的线性学习问题。

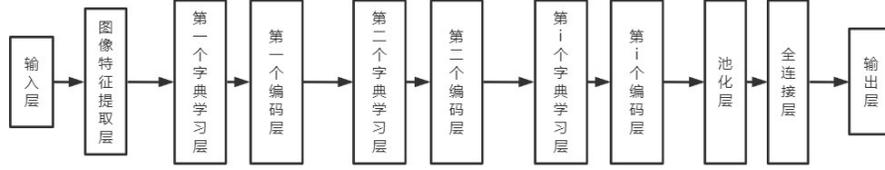


图4 DDLN框架图

在字典学习层,使用ERCDLSI算法中的思路,其整体模型定义为

$$\min_{D,A} \sum_{i=1}^q \left\{ \begin{aligned} & \left\| Y_i - D_i A_i \right\|_F^2 + \lambda_1 \|A_i\|_1 + \lambda_2 \left(\sum_{j \neq i} R_{ij} \|D_j^T D_i\|_F^2 \right) \\ & + \lambda_3 \sum_i \sum_j \left[\min(0, (1 - y_k(w_k a_i + b_k))) \right]^2 \end{aligned} \right\} \quad (3)$$

首先要构建 C 个二元分类器,其中 C 为类别数。每个二元分类器使用数据集中某类数据与非该类数据之间的平面来构成。分类器将对重建系数进行一次判断,对于判断错误的,则做相应程度惩罚。

在编码层,学习 D 后,每个局部特征通过几个最近的原子编码生成编码。通过这样做,特征编码层将每个局部描述符 y_i 转换为向量 $\gamma^i = [\gamma_1^i, \gamma_2^i, \dots, \gamma_l^i]$ 。具体来说,每个编码可以使用以下公式获得:

$$\min_{\gamma_i^1} \left[\sum_{i=1}^l \frac{1}{2} \|y_i - D \gamma_i^1\|_2^2 + \beta \|\gamma_i^1 \odot \zeta_i^1\|_1 \right] \quad (4)$$

s.t. $\mathbf{1}^\top \gamma_i^1 = 1$

其中 γ_i^1 是距离向量,测量距离在 y_i 和 d_i 之间, \odot 表示元素乘法或Hadamard乘积,它使两个向量(γ_i^1 和 ζ_i^1)的对应项能够相乘。通常 ζ_i^1 可以使用 l_2 范数获得,即

$$\zeta_i^1 = \left[\|y_i - d_1\|_2, \|y_i - d_2\|_2, \dots, \|y_i - d_{s_1}\|_2 \right]^\top$$

按照网络算法流程如下:

特征提取层:令 Y 表示一组 m 维局部描述符, $Y = [y_1, \dots, y_N] \in \mathbb{R}^{m \times N}$,其中 N 为总数局部描述符。

第一个字典学习层:通过最小化目标函数,得到字典 D 和编码系数矩阵 A 。

第一个特征编码层:通过公式计算得到 γ_i^1 。

第二个字典学习层:如第一字典学习层,只是将 Y_i 替换成第一个字典 D ,得到字典 U 。

第二个特征编码层:如第一个特征编码层,将 y_i, D 换成 d_i, U ,通过公式得到 γ_i^2 。

池化层:在特征编码层之后,池化然后层接管。对于每张图像,采用 1×1 、 2×2 和 4×4 的空间金字塔进行最大池化处理。

全连接层:第一个中的每个编码向量 γ_i^l 可以进行扩

充,例如,第 j 项可以扩充为:

$$\left[\gamma_i^1(v_j), \gamma_i^1(v_j) \left[\gamma_j^2(u_1), \gamma_j^2(u_2), \dots, \gamma_j^2(u_{s_2}) \right] \right]^\top \quad (5)$$

最后的输出层,使用svm来分类。

算法1:融合深层结构框架的多标签标注算法整体计算流程

输入:训练样本集 $Y \in \mathbb{R}^{m \times l}$

输出:编码向量 γ_i

1. 第一层字典学习: $V \leftarrow Y$

2. 第一层的局部约束项的计算:

$$\zeta_i^1 = \left[\|y_i - d_1\|_2, \|y_i - d_2\|_2, \dots, \|y_i - d_{s_1}\|_2 \right]^\top$$

3. 第一层特征编码:

for i=1 to l :

$$\gamma_i^1 \leftarrow \min_{\gamma_i^1} \left[\frac{1}{2} \|y_i - D \gamma_i^1\|_2^2 + \beta \|\gamma_i^1 \odot \zeta_i^1\|_1 \right]$$

s.t. $\mathbf{1}^\top \gamma_i^1 = 1$

end

4. 第二层字典学习: $U \leftarrow D$

5. 第二层的局部约束项的计算:

$$\zeta_i^2 = \left[\|d_i - u_1\|_2, \|d_i - u_2\|_2, \dots, \|d_i - u_{s_2}\|_2 \right]^\top$$

6. 第二层特征编码:

for i=1 to D_i :

$$\gamma_i^2 \leftarrow \min_{\gamma_i^2} \left[\frac{1}{2} \|d_i - U \gamma_i^2\|_2^2 + \beta \|\gamma_i^2 \odot \zeta_i^2\|_1 \right]$$

s.t. $\mathbf{1}^\top \gamma_i^2 = 1$

end

7. 编码增强

for i=1 to l :

for j=1 to D_i :

$$\gamma_i^1(v_j) \leftarrow \left[\gamma_i^1(v_j), \gamma_i^1(v_j) \left[\gamma_j^2(u_1), \gamma_j^2(u_2), \dots, \gamma_j^2(u_{s_2}) \right] \right]^\top$$

end

end

8. 使用multiclass SVM对编码向量进行分类,并排序。

算法2:字典学习层的具体步骤

输入:训练样本集合 Y 。

输出:字典 $D_i, i = 0, 1, \dots, C$; 超平面 $U_i^0, i = 0, 1, \dots, C$; $b_i^0, i = 0, 1, \dots, C$

1. 初始化字典 $D_i^0, i = 0, 1, \dots, C$, 重建系数 A^0 ;

2. 循环:

计算 U_i^t, b_i^t , 共 c 个超平面;

更新重建系数

$$a = \arg \min_a \left\| x - D_i a \right\|_F^2 + \lambda_1 \|a\|_1 + \lambda_3 \sum_c \left[\min(0, (1 - \gamma(U_c^t a + b_c^t))) \right]^2$$

更新字典

$$\min_b \sum_{i=1}^c \left\{ \left\| X_i - D_i A_i \right\|_F^2 + \lambda_1 \|A_i\|_1 + \lambda_2 \left(\sum_{j \neq i} R_{ij} \|D_j^T D_i\|_F^2 \right) \right\}$$

;

4 实验结果与分析

实验使用的传统服饰图像数据集所包含的纹样以及样本数量如下表1所示:

表1 传统服饰图像数据集

纹样	样本数量	纹样	样本数量
云纹	251	花纹+蝴蝶纹	317
龙纹	286	鸟纹+花纹+蝴蝶纹	152
花纹	479	鸟纹+云纹	189
蝴蝶纹	448	鸟纹+花纹	331
鸟纹	309	龙纹+云纹	323
鸟纹+花纹+云纹	4	云纹+龙纹+鸟纹	3
合计:3092			

数据集的具体参数如下表2:

表2 传统服饰图像数据集参数

数据集	图像数量	标签数量	LCard	LDen	DL
传统服饰图像	3092	5	1.473	0.218	17

其中LCard表示每个样本标签的平均个数,LDen为标签密度,DL表示标签组合数量。本文的每次实验将随机选择2500张图片作为训练集,其余的图片作为测试集,重复进行十次实验,最终实验结果选择十次实验的平均值。

由于字典原子数 p 对于字典的表达有着很大的影响,若 p 取值过大,即该类图像所提取出来的字典原子数过多,虽然其所包含的更多信息可以更好

的表达本类图像,但是也会带来两个方面的问题,一是会增加计算上的复杂程度,二是增大了本类图像信息的冗余度。当使用这样的字典对其他类的测试样本进行重建时,也会减小重建误差,从而不能很好的区分样本在不同字典下的重建效果,最终导致标注效果变差。若 p 取值过小,由于原子数过少,其所包含的信息就较少,因此可能丢失了一些重建本类样本图像的关键特征信息,字典将不能很好的表达本类图像,标注效果也很变差。

本文将分别设置字典原子数 p 为 [15, 20, 25, 30, 35] 进行多次实验,观察 p 值对于实验的影响,实验的指标使用平均准确度 AP(Average precision)。

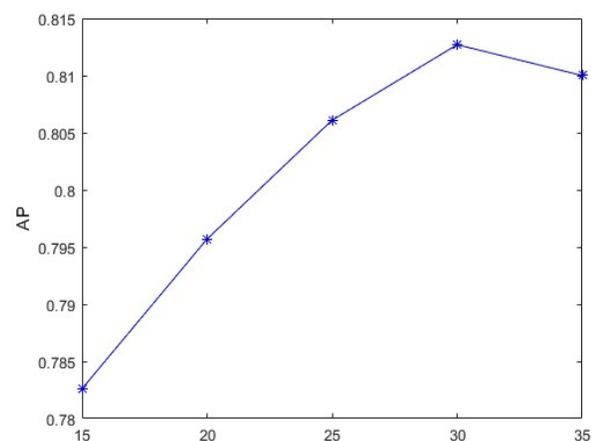


图5 字典原子数对算法AP的影响

实验结果如图5所示,由图5可以看出,随着字典原子数 p 的增加,算法 AP 先逐步增大,在达到了顶点后,又有所下降。因此,选择 $p = 30$ 作为字典原子数。

本实验将本文提出的算法与ML-KNN、Rank-SVM、LLSF、MLNB、SCMIDL、ERCIDL进行对比实验。

ML-KNN算法是将k-Nearest Neighbor算法进行了调整,通过判断特征空间中附近样本的类别,然后利用后验概率最大化来判断该样本最可能属于的类型。

Rank-SVM算法按照标签对是否相关将每组数据集划分开,并计算其间隔其间隔距离,通过使得距离最大化,来提高分类算法性能。

LLSF算法通过引入标签回归系数,来强化算法对于图像信息的表达能力,同时也会考虑到数据类别之间的关联性信息来减少其对图像分类的干扰。另外,该算法还有两个拓展方法,分别为LLSF-BR与

LLSF-CC,其中BR指的是Binary Relevance二元分类器,而CC表示的是Classifier Chains分类器链表。

MLNB是由朴素贝叶斯分类算法向多标签问题改进适应而来,同样是基于朴素贝叶斯理论来进行分类的,其中使用PCA算法来对特征空间进行了降维,

从而降低了计算的复杂程度,充分利用了图像的特征信息来提高标注效果。

SCMIDL、ERCIMIDL算法在前面已经有详细介绍,这里同样将它们一起进行对比实验。

实验结果如表3所示:

表3 实验结果

	One-error	Coverage	Ranking-loss	Average-precision
ML-KNN	0.4035±0.0174	1.2641±0.1215	0.2351±0.0043	0.7631±0.0219
Rank-SVM	0.3671±0.0236	1.1821±0.01559	0.1977±0.0045	0.7672±0.0085
MLNB	0.4724±0.0259	1.3637±0.041	0.2578±0.0101	0.71±0.0102
LLSF-BR	0.3671±0.0076	1.163±0.051	0.2036±0.0107	0.7661±0.0049
LLSF-CC	0.3896±0.0251	1.3051±0.0961	0.2388±0.021	0.7473±0.0148
LLSF	0.3863±0.0251	1.2609±0.0301	0.2251±0.0108	0.7466±0.0129
SCMIDL	0.3152±0.0376	1.0284±0.0284	0.1693±0.0103	0.7993±0.0138
ERCIMIDL	0.3102±0.0175	1.0096±0.0433	0.1634±0.0115	0.8093±0.0072
本文优化的算法	0.3084±0.0241	1.0095±0.0332	0.1632±0.0132	0.8204±0.0075

通过实验结果可以看出,本文优化的算法在One-error, Average-precision指标上有明显提升,能够更精确的标注出传统服饰图像中的纹样,标注出最优标签的误差率更低,这使得当传统服饰图像中只包含一种纹样的时候,能够更有效地进行标注。在传统服饰图像数据集的所有图片中,有超过一半的图片只包含一种纹样,因此One-error性能的提升,对于传统服饰图像的标注具有重要作用。One-error, Average-precision性能的提升,验证了深度结构网络框架对于字典表达能力的提升,使得字典学习模型可以更有效的识别图片中的重要特征信息。在性能的稳定性方面,该优化算法虽然比不上特定的一些多标签标注算法(LLSF-BR),但是比绝大多数的算法的稳定性都更强,可以胜任大多数的多标注任务。另外,在Coverage和Ranking-loss指标上,本文优化的算法同样具有较好的表现。

综上所述,本文使用深度结构框架优化的算法在四项性能指标上均具有优秀的表现,其中One-error和Average-precision的提升较为明显,这两项指标的提升对于算法在实际场景中的应用也具有更明显的作用,整体0.8204的平均精确度,相比上一章节浅层结构字典优化算法的性能也有了进一步的提升,可以更好地胜任传统服饰图像标注系统的自动标注任务。

算法整体的性能提升也验证了深度结构框架对字典学习算法优化的有效性。

5 总结

本文讨论了基于结构来对字典学习进行优化。文章首先介绍了几种传统的单层结构字典学习算法,它们通过对重建误差与重建系数的优化,在多标签标注问题上有着较好的表现,然而还是存在浅层结构的一些弊端。然后介绍了基于多层深度学习框架的字典学习算法的原理及其优势,并将字典学习多标签标注算法和多层深度学习框架融合,使算法的标注效果得到了进一步提升。最后,本文对提出的算法进行了大量的对比实验,证明了算法的有效性。中国传统服饰文化源远流长,在不同时代,不同种族均保留有大量的图案需要被数字化、传承和发扬。本文工作的字典学习算法中,只考虑到图像信息的稀疏性,而没有从更深的角度去考虑图像特征的局部性,有进一步优化的空间。

参考文献(References):

- [1] Bengio Y, Lamblin P, Popovici D, et al. Greedy layerwise training of deep networks [C]//Advances in Neural Infor-