

基于混合属性的零样本图像分类

程玉虎, 乔 雪, 王雪松

(中国矿业大学信息与控制工程学院, 江苏徐州, 221116)

摘 要: 对于具有相似属性的类别而言, 在有限维度的语义属性下, 基于属性的零样本图像分类器难以对它们进行正确区分. 考虑到语义属性描述类别的有限性, 在直接属性预测(Direct Attribute Prediction, DAP)模型的基础上, 提出一种基于混合属性的零样本图像分类模型(Hybrid Attribute-Based DAP, HA-DAP). 首先, 对样本的底层特征进行稀疏编码并利用编码后的非语义属性来辅助现有的语义属性; 将非语义属性与语义属性构成混合属性并将其作为DAP模型的属性中间层, 利用属性预测模型的思想进行混合属性分类器的训练; 最后, 根据预测的混合属性以及属性与类别之间的关系进行测试样本类别标签的预测. 在OSR、Pub Fig以及Shoes数据集上的实验结果表明, HA-DAP的分类性能优于DAP, 不仅能够取得较高的零样本图像分类精度, 而且还获得了较高的AUC值.

关键词: 零样本图像分类; 混合属性; 语义属性; 非语义属性; 稀疏编码

中图分类号: TP181 **文献标识码:** A **文章编号:** 0372-2112 (2017)06-1462-07

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2017.06.026

Hybrid Attribute-Based Zero-Shot Image Classification

CHENG Yu-hu, QIAO Xue, WANG Xue-song

(School of Information and Control Engineering, China University of Mining & Technology, Xuzhou, Jiangsu 221116, China)

Abstract: When the dimensionality of the semantic attributes is limited, it is difficult for attribute-based zero-shot image classifiers to distinguish the objects with similar attributes. Aiming at the limitation of describing objects with semantic attributes, an improved direct attribute prediction(DAP) model for zero-shot image classifying based on hybrid attribute(HA) is proposed, which is called HA-DAP. At first, we carry out the sparse coding on the low-level features to obtain the non-semantic attributes that are used to assist the existing semantic attributes. Then, we take the hybrid attributes including the learned non-semantic attributes and the manually specified semantic attributes as the mid-layer of DAP model and use the idea of attribute prediction to train the hybrid attribute-based classifier. At last, according to the predicted hybrid attributes and the relationship between the attributes and classes, we can recognize the class label for the testing sample. Experimental results on the OSR, Pub Fig and Shoes datasets show that, the HA-DAP outperforms the DAP in the classification performance, i. e., when compared with the DAP, the proposed HA-DAP yields much higher zero-shot image classification accuracy and AUC value.

Key words: zero-shot image classification; hybrid attribute; semantic attribute; non-semantic attribute; sparse auto-encoding

1 引言

属性是指物体固有的特性, 视觉属性^[1]是指可以通过人工标注并且能在图像中观察到的特性(例如, 有翅膀、黑头发). 视觉属性和特征都可以对图像的内容进行描述, 不同之处在于特征是图像内容的底层描述, 它能够被机器识别但是不能被人理解, 而视觉属性是图像内容的高层描述, 能够同时被机器和人理解. 大量的研究已经显示了视觉属性在目标识别^[2,3]、图像描

述^[4,5], 以及图像分类^[6]方面的作用. 一个对象通常具有许多特性, 例如, 老鹰是有翅膀的, 亚洲人有黑头发, 马路是开阔的, 高跟鞋有很高的鞋跟等等, 通过这些视觉属性可以认识对象的外观并且把该对象描述给其他人. 此外, 不同的对象类别往往具有共同的属性, 将它们作为分类器的属性中间层允许不同类别之间共享到这些有关联的属性, 将已知的属性知识从一个类别迁移到另一个新的类别上面, 这使得分类器能够对那些没有训练样本的类别进行识别.

收稿日期: 2015-05-13; 修回日期: 2015-12-12; 责任编辑: 梅志强

基金项目: 国家自然科学基金(No. 61472424, No. 61273143); 中央高校基本科研业务费(No. 2013RC10, No. 2013RC12, No. 2014YC07)

利用图像的语义属性实现零样本图像分类^[7,8]是当前属性应用领域的研究热点,与传统的图像分类问题不同,零样本图像分类在测试阶段所分类和识别的样本未参与分类器模型的训练.在零样本图像分类问题中,为了实现从可见类别到未见类别的知识迁移,分类模型就需要通过属性来搭建一座从底层特征到类别标签的桥梁^[9].最近的研究工作中提出了很多基于属性学习的图像分类方法,具有代表性的是文献[10]中提出的直接属性预测模型 DAP 和间接属性预测模型 (Indirect Attribute Prediction, IAP).在基于语义属性的零样本图像分类器模型中,语义属性考虑了样本是否具有某一种属性,根据属性的“有”、“无”可以确定样本在属性空间的位置,进而确定样本的类别标签.但是,对那些属性很相似的类别而言(属性空间位置临近),语义属性就很难对它们进行区分.为此,拟利用稀疏编码^[11]对图像的底层特征进行重构,可以得到图像的另一种低维、紧凑的表示方式.由于这种重构后的特征没有语义信息,因此可命名为非语义属性.通过这种额外维度的非语义属性对原有的语义属性进行补充和辅助,将属性空间加以扩展并构成混合属性.非语义属性可以增加语义属性的差异性,从而能够使语义属性相似的类别更加容易区分^[12].进一步,将构造的混合属性应用于 DAP,提出一种基于混合属性的零样本图像分类方法.

2 基于混合属性的零样本图像分类

在如图 1 所示的基于混合属性的零样本图像分类模型中,语义属性和特征编码后的非语义属性均采用二值属性,即语义属性空间 $A = \{0, 1\}^M$,非语义属性空间 $B = \{0, 1\}^N$.算法的基本思想为:对所有类别样本的特征进行稀疏编码,得到非语义属性与类别标签的对应关系,将语义属性和非语义属性构成的混合属性作为 DAP 模型的属性中间层,利用属性预测模型的思想进行混合属性分类器的训练,然后根据预测的混合属性以及属性与类别之间的关系进行样本类别标签的预测.在训练阶段,利用训练样本学习语义属性分类器和非语义属性分类器.在测试阶段,利用混合属性分类器对测试样本的语义属性和非语义属性进行预测,根据混合属性中间层与各类别标签之间的关系,获得样本的类别标签.

2.1 混合属性

以零样本图像分类数据集 Shoes 为例阐述混合属性构造的基本思想,如图 2 所示.假设给定 5 个语义属性 $(a_1, a_2, a_3, a_4, a_5) \in A$ 用以描述三类类别:高跟鞋、婚鞋和运动鞋.由图可知,高跟鞋和婚鞋仅在最后一个属性“Shiny”上有所区别,其余属性则均相同.也就是说,我们

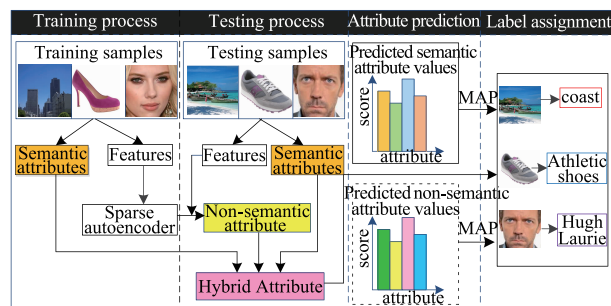


图1 基于混合属性的零样本图像分类

利用 DAP 或者 IAP 预测得到的高跟鞋和婚鞋的属性是非常相似的.因此,当在进行零样本图像分类的时候,基于这些预测的属性,基于属性的零样本分类器就很容易混淆这两类鞋子.然而,用以描述运动鞋的属性与高跟鞋和婚鞋的属性之间存在着较大的差别,因此,在进行属性预测后,零样本分类器不容易将其与其他类别的鞋子混淆.考虑到属性的有限性,提出对样本的特征进行稀疏编码,将编码后的特征作为非语义属性 $(b_1, b_2, b_3) \in B$ 对有限的语义属性进行补充并组成混合属性,进而能够更好地对相似类别进行区分.至于运动鞋的样本,原有的语义属性足以将它与其他的类别进行区分,因此辅助的非语义属性并不会对其分类产生过多的影响.

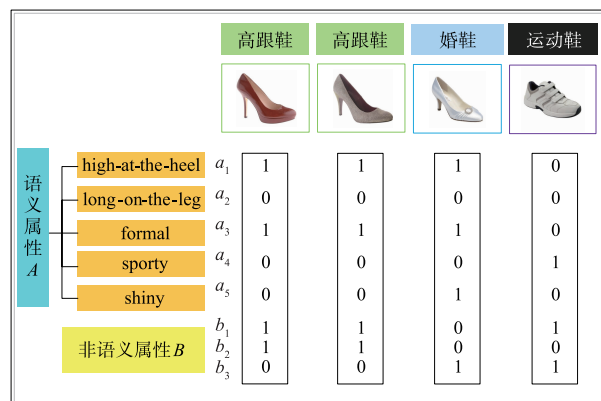


图2 混合属性示意图

2.2 基于稀疏编码的非语义属性学习

假设共有 H 类图像,将其中 K 类作为可见的训练类别参与混合属性分类器的训练,剩下的 $L = H - K$ 类作为测试类别, L 不参与混合属性分类器的训练,即不可见类别.利用稀疏编码学习图像非语义属性的主要思路是:用一组无标签的训练样本的特征学习一组基向量,所有训练样本的特征均可以用这组基向量线性表示,同时测试样本的特征也可以用这些基向量进行稀疏编码.假设 $X = (x_1, x_2, \dots, x_K) \in R^{d \times K}$ 表示 K 类训练图像集的特征,使用稀疏编码算法来学习得到一组基向量集合 $\Phi = \{\varphi_1, \varphi_2, \dots, \varphi_N\}$,然后用这些基向量来

对原始的输入特征进行重构 $\tilde{x}_i \approx \sum_{j=1}^N b_j \varphi_j$, 其中 b_j 称为激活量, 也就是本文提出的非语义属性. 稀疏编码由训练和编码两个过程共同完成.

训练过程: 稀疏编码采用训练数据生成基向量集, 再利用基向量集的线性组合来表示输入向量. 因此, 稀疏编码的优化问题可表示为^[13]:

$$\min_{\varphi, b} \sum_{i=1}^K \|x_i - \sum_{j=1}^N b_{i,j} \varphi_j\|^2 + \lambda \sum_{i=1}^K \sum_{j=1}^N \|b_{i,j}\|_{L_1} \quad (1)$$

$$\text{s. t. } \|\varphi_j\|^2 \leq c, \forall j=1, 2, \dots, N$$

其中, $b_{i,j}$ 为第 i 类训练样本特征 x_i 的第 j 个稀疏编码表示; λ 表示权重衰减系数; $\|b_{i,j}\|_{L_1}$ 是稀疏正则项, 该稀疏项潜在地迫使约束函数具有唯一解, 并且保证了输入 x_i 只由比较显著的特征模式来表示. 通过求解式(1)的优化问题, 就可以得到输入特征向量 x_i 的稀疏编码激活量 $b_{i,j}$ 和基向量 φ_j . 当 $b_{i,j}$ 和 φ_j 同时变化时, 目标函数不一定是凸优化问题. 然而, 固定基向量 φ_j , 优化问题是关于激活量 $b_{i,j}$ 的凸优化问题; 固定激活量 $b_{i,j}$, 它是关于基向量 φ_j 的凸优化问题. 因此, 可以通过固定一个变量求解另一个变量的交替迭代优化求解方法^[14,15] 解决稀疏编码问题, 每个迭代过程分为两步: (1) 首先固定基向量 φ , 调整激活量 b 使得式(1)最小; (2) 然后固定激活量 b , 调整基向量 φ 使得目标约束函数最小. 不断迭代, 直至收敛, 这样就可以找到一组能够良好表示样本特征的基向量了.

编码过程: 若给定新的样本特征 x_i , 由于训练阶段已求得基向量 φ , 则此时的稀疏编码约束函数可表示为:

$$\min_b \sum_{i=1}^K \|x_i - \sum_{j=1}^N b_{i,j} \varphi_j\|^2 + \lambda \sum_{i=1}^K \sum_{j=1}^N \|b_{i,j}\|_{L_1} \quad (2)$$

$$\text{s. t. } \|\varphi_j\|^2 \leq c, \forall j=1, 2, \dots, N$$

只需调整激活量 $b_{i,j}$ 使得式(2)最小, 则此时的激活向量 $b = (b_1, b_2, \dots, b_N) \in \{0, 1\}$ 就是这个输入特征的稀疏编码表示, 亦即图像样本的非语义属性. 通过稀疏编码获得非语义属性的具体算法流程如算法 1 所示.

算法 1 基于稀疏编码的非语义属性学习

输入: 训练样本和测试样本的底层特征 x ; 权重衰减系数 λ .

Step 1: 将训练样本的底层特征 x 带入约束方程式(1)中, 并随机初始化基向量 φ ;

Repeat:

固定上一步的基向量 φ , 求解能够使约束方程式(1)最小化的激活量 b ;

固定上一步的激活量 b , 求解能够使约束方程式(1)最小化的基向量 φ ;

不断重复以上步骤直至收敛, 求得一组能够良好表示样本特征的基向量 φ ;

Step 2: 将上一步求得的基向量 φ 以及测试样本的底层特征 x 带

入约束方程式(2)中;

Step 3: 调整激活量 b 使得式(2)最小, 此时的激活量即为测试样本的非语义属性.

输出: 训练样本及测试样本的非语义属性.

2.3 基于混合属性的零样本图像分类

将构造的混合属性应用于 DAP 模型, 提出一种基于混合属性的 DAP (Hybrid Attribute-Based DAP, HA-DAP) 并将其用于解决零样本图像分类问题. 图 3 给出了 HA-DAP 的示意图, 其中 $x = (x^1, x^2, \dots, x^d)$ 代表样本的 d 维底层特征, a 和 b 分别表示样本的语义属性和非语义属性, y 和 z 表示标签, $y = (y_1, y_2, \dots, y_k)$ 代表训练样本类(可见类)的标签集合, $z = (z_1, z_2, \dots, z_L)$ 代表测试样本类(不可见类)的标签集合. 用于训练的类别 y_1, y_2, \dots, y_k 和不可见的测试类别 z_1, z_2, \dots, z_L 与混合属性 $(a_1, a_2, \dots, a_M; b_1, b_2, \dots, b_N)$ 之间的关系通过一个二值矩阵给出, 矩阵中的值表示对于一个给定的类别 y 或者 z , 混合属性对于分类是有效还是无效(1 为有效, 0 为无效). 语义属性 $a_m \in \{0, 1\}$ 是通过人工有监督给出的, 非语义属性 $b_n \in \{0, 1\}$ 是在人工监督下通过稀疏编码学习得到的. 混合属性预测模型主要分为两个部分: 语义属性预测部分和非语义属性预测部分.

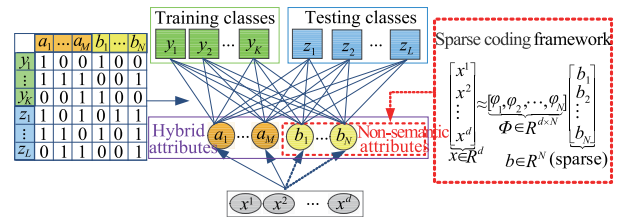


图3 HA-DAP示意图

在语义属性预测部分, 采用直接属性预测模型的思想, 将所有可见的 K 类样本的底层特征 $x = (x^1, x^2, \dots, x^d)$ 和语义属性标签 $a = (a_1, a_2, \dots, a_M) \in \{0, 1\}$ 作为训练样本, 为每一个语义属性训练一个属性分类器. 在测试阶段, 测试样本的底层特征 x 的有效语义属性 a 可以通过已训练好的属性分类器进行预测, 并用后验概率 $p(a|x)$ 表征^[1]:

$$p(a|x) = \prod_{m=1}^M p(a_m|x) \quad (3)$$

根据贝叶斯定理, 可以得到从预测属性 a 到测试类标签 z 的表示:

$$p(z|a) = \frac{p(z)}{p(a)} p(a|z) \quad (4)$$

以判别式的方式确定测试类别 z 的属性分布:

$$p(a|z) = \begin{cases} 1, & \text{if } a = a^z \\ 0, & \text{else} \end{cases} \quad (5)$$

那么, 从测试样本的底层特征 x 到测试样本类标签 z 的

预测可以表示为:

$$p(\mathbf{z}|\mathbf{x}) = \sum_{\mathbf{a} \in \{0,1\}^M} p(\mathbf{z}|\mathbf{a})p(\mathbf{a}|\mathbf{x}) = \frac{p(\mathbf{z})}{p(\mathbf{a}^z)} \prod_{m=1}^M p(a_m^z|\mathbf{x}) \quad (6)$$

在上式中,由于缺乏先验知识,那么假设测试类别被分为任何类的概率值都是相等的,则在进行测试类别标签预测的时候就可以忽略 $p(\mathbf{z})$ 的影响. 至于 $p(\mathbf{a}) = \prod_{m=1}^M p(a_m)$, 其中 $p(a_m) = \frac{1}{K} \sum_{k=1}^K a_m^{y_k}$ 是利用训练样本的属性概率作为先验属性. 事实上,先验属性对于分类的影响不大,并且当 $p(a_m) = \frac{1}{2}$ 时的分类效果较好^[10].

那么, $p(\mathbf{z}|\mathbf{x})$ 可表示为:

$$p(\mathbf{z}|\mathbf{x}) = \sum_{\mathbf{a} \in \{0,1\}^M} p(\mathbf{z}|\mathbf{a})p(\mathbf{a}|\mathbf{x}) = \prod_{m=1}^M \frac{p(a_m^z|\mathbf{x})}{p(a_m^z)} \quad (7)$$

在非语义属性预测部分,利用训练阶段学习的稀疏编码器对测试样本的非语义属性进行预测,获得样本的非语义属性的后验概率 $p_b(\mathbf{b}|\mathbf{x})$:

$$p_b(\mathbf{b}|\mathbf{x}) = \prod_{n=1}^N p(b_n|\mathbf{x}) \quad (8)$$

由自动编码器获得非语义属性的后验概率后,可由非语义属性与类别标签之间的关系获得从测试样本的底层特征到类别的预测,过程与语义属性预测过程相同,则可得到:

$$p_b(\mathbf{z}|\mathbf{x}) = \sum_{\mathbf{b} \in \{0,1\}^N} p(\mathbf{z}|\mathbf{b})p(\mathbf{b}|\mathbf{x}) = \prod_{n=1}^N \frac{p(b_n^z|\mathbf{x})}{p(b_n^z)} \quad (9)$$

在标签分配阶段,通过最大后验(Maximum A Posteriori, MAP)估计实现由训练类别标签的后验分布来推知测试样本类标签的概率分布,即:

$$f(\mathbf{x}) = \arg \max_{l \in \{1, \dots, L\}} \left\{ \prod_{m=1}^M \frac{p(a_m^z|\mathbf{x})}{p(a_m^z)} + \prod_{n=1}^N \frac{p(b_n^z|\mathbf{x})}{p(b_n^z)} \right\} \quad (10)$$

综上所述,给出利用 HA-DAP 模型进行零样本图像分类的算法步骤如算法 2 所示.

算法 2 基于混合属性的零样本图像分类

输入:训练类 K 和测试类 L 的底层特征 \mathbf{x} 和语义属性 \mathbf{a} .

Step 1: 利用训练类的底层特征学习稀疏编码器,并利用稀疏编码器获得训练类和测试类的非语义属性 \mathbf{b} , 由此便可以人工有监督地给出类别-非语义属性的二值矩阵;

Step 2: 利用训练样本为每一个语义属性 \mathbf{a} 学习一个分类器,并对测试样本的语义属性进行预测 $p(\mathbf{a}|\mathbf{x})$, 然后根据式(7)获得从测试样本底层特征到标签的预测 $p(\mathbf{z}|\mathbf{x})$;

Step 3: 利用 Step 1 学习的稀疏编码器对测试样本的非语义属性进行预测 $p_b(\mathbf{b}|\mathbf{x})$, 根据式(9)获得从测试样本底层特征到标签的预测 $p_b(\mathbf{z}|\mathbf{x})$;

Step 4: 对于测试样本,根据式(10)计算该样本属于不同类别的

后验概率,找出使后验估计最大的类别并分配标签.

输出:测试样本的属性和类别标签.

3 实验结果分析

3.1 实验数据集

实验选取户外场景识别数据集(OSR)、公开人脸数据集(Pub Fig)和属性发现数据库--鞋类数据集(Shoes)进行测试. 上述数据集涵盖了风景、人脸和物品三个截然不同的领域,可以充分验证算法对于不同应用领域的适用性. OSR 数据集共包含 2688 幅图片,8 个场景类以及 6 种属性,使用 512 维的 gist 特征来表示图像;Pub Fig 数据集共包含 8 位名人的 772 幅头像图片,对每幅图片样本提取 512 维 gist 特征以及 30 维全局颜色特征,该数据集拥有 11 种语义属性;Shoes 数据集共包含 14658 幅图片,10 类鞋子以及 10 种属性. 与 Pub Fig 相同,Shoes 数据集也使用 gist 特征和全局颜色特征来表示图像. 表 1 给出了三个数据集的具体描述.

表 1 数据集描述

数据集	样本量	类别数	属性数	特征维数
OSR	2688	8	6	gist(512)
Pub Fig	772	8	11	gist(512); color(30)
Shoes	14658	10	10	gist(960); color(30)

3.2 零样本图像分类

本实验重点讨论 HA-DAP 在零样本学习上的图像分类效果,对比方法为 DAP 模型. 在每一个数据集上均进行多次零样本图像分类实验,每次实验选取不同的训练类和测试类进行分类实验. 另外,为了消除随机因素对实验结果的影响,实验采取了 C_h^h 折交叉验证的方法(H 为数据集类别总数, h 为参与训练的类别数),交叉验证使得每次实验中几乎所有样本均参与模型训练,所得评估结果更加可靠. 也就是说,在实验过程中,在 OSR 数据集和 Pub Fig 数据集上分别进行了 5 次交叉验证实验,即: $C_8^6 = 28$ 折、 $C_8^5 = 56$ 折、 $C_8^4 = 70$ 折、 $C_8^3 = 56$ 折、 $C_8^2 = 28$ 折;在 Shoes 数据集上进行 6 次交叉验证实验,即: $C_{10}^8 = 45$ 折、 $C_{10}^7 = 120$ 折、 $C_{10}^6 = 210$ 折、 $C_{10}^5 = 252$ 折、 $C_{10}^4 = 210$ 折、 $C_{10}^3 = 120$ 折.

图 4 给出了在不同训练类别数的情况下,零样本图像分类精度与非语义属性维数 N 之间的关系曲线,其中 N 在 0~10 范围内取值. 值得指出的是,当 $N=0$ 的时候,HA-DAP 就退化为 DAP. 由图 4 可以看出:随着训练类别数 h 的减少,DAP 和 HA-DAP 的零样本分类精度均有所降低. 这是由于当训练类别数减少时,参与训练的属性会减少,导致对于测试样本中出现而训练样本中没有出现的一些属性(训练样本的属性空间无法涵

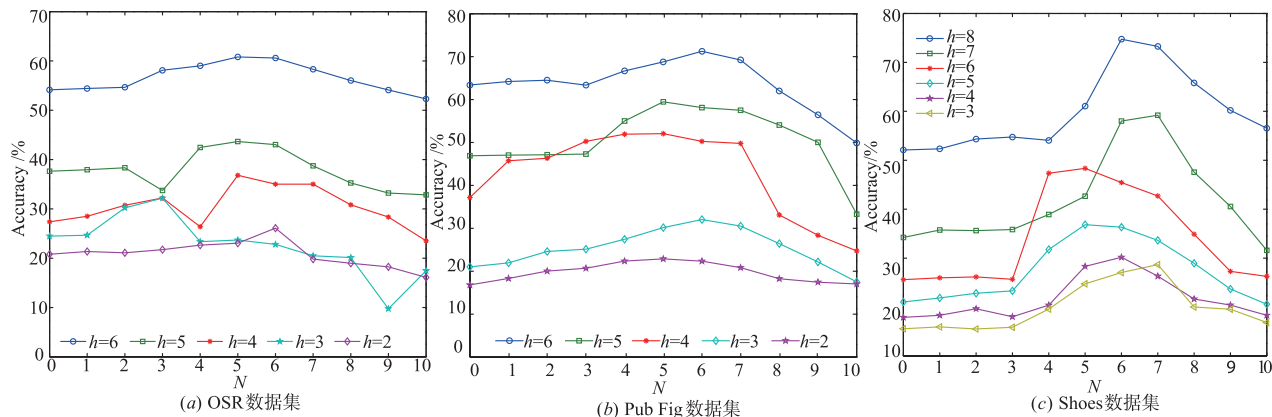


图4 分类精度与非语义属性空间维数间关系曲线

盖测试样本的属性)的预测精度会偏低,进而导致在未见测试样本上的分类精度下降。

表2给出了DAP和HA-DAP在三个数据集上的零样本图像平均分类精度比较。由图4和表2可以看出:随着非语义属性维数的增加,零样本图像的分类精度有先升高后下降的趋势。这是由于少量的非语义属性的参与可以有效地增加属性空间的差异性,进而提高分类精度。然而,随着非语义属性维数的不断增加,过多的非语义属性反而会造成属性的冗余,进而造成分类精度的下降。

表2 零样本图像平均分类精度比较(OSR数据集)

模型		平均分类精度(%)		
		OSR	Pub Fig	Shoes
DAP		33.08	37.06	27.74
HA-DAP	$N=1$	33.38	39.45	28.37
	$N=2$	35.01	40.52	29.03
	$N=3$	35.60	41.35	28.91
	$N=4$	34.81	44.68	35.35
	$N=5$	37.60	46.67	40.33
	$N=6$	37.51	46.80	45.29
	$N=7$	33.63	45.57	43.97
	$N=8$	31.75	38.77	36.49
	$N=9$	30.20	34.89	31.95
	$N=10$	28.45	28.50	28.36

分类精度能够真实地反映出分类正确的样本数与测试样本总数的关系,但是不能反映误判率(把实际为假值的样本判定为真值的概率)与灵敏度(把实际为真值的样本判定为真值的概率)之间的关系。因此,实验中引入ROC曲线下面积(Area Under Curve, AUC)以便更好地对分类效果进行评价。AUC提供了评价模型平均性能的另一方法^[16],如果模型是完美的,那么它的

AUC值应为1,如果模型是个简单的随机猜测模型,那么它的AUC值为0.5。由表2可知:(1)当 $N=6$ 时,在所有11个零样本图像分类算法中,HA-DAP在三个数据集上基本上能够得到最高的平均分类精度;(2)当 $N=10$ 时,HA-DAP在三个数据集上的零样本图像平均分类精度低于其他9种类型的HA-DAP,甚至低于DAP的平均分类精度。考虑到代表性,此处仅给出DAP、HA-DAP($N=3$)、HA-DAP($N=6$)以及HA-DAP($N=10$)在三个数据集上进行零样本图像分类的AUC值柱状图,如图5所示。可以看出,(1)3种类型的HA-DAP在三个数据集上的AUC值均高于随机试验的AUC值(0.5);(2)当 $N=3$ 和 $N=6$ 时,AUC值远高于DAP的AUC值,这是由于少量的非语义属性增加了属性空间的差异性和提高了分类精度;(3)当 $N=10$ 时,AUC值有所减小,这是由于冗余的非语义属性降低了分类器的性能。

图6~8分别给出了当训练类别为4类(OSR数据集和Pubfig数据集)和5类(Shoes数据集)的时候,DAP和HA-DAP在三个数据集上的某次分类结果混淆矩阵对比图。在混淆矩阵中,主对角线上的方块表示被正确分类的样本数量,颜色越深,表示该类别样本分类正确数量越多。由图8可以看出:(1)由于高跟鞋和婚鞋的语义属性较为相似,DAP难以正确区分这两类样本,如将所有1138幅“婚鞋”图像均错分为“高跟鞋”;(2)由于额外引入了非语义属性,HA-DAP能准确识别部分高跟鞋和婚鞋,对相似度较高的高跟鞋和婚鞋的分类正确率较DAP均有所提高。由图6~8可以得出与前述一样的结论:(1)当 $N=3$ 和 $N=6$ 时,在三个数据集混淆矩阵的主对角线上,HA-DAP正确分类的样本数多于DAP,这是由于少量的非语义属性提高了分类器的性能;(2)当 $N=10$ 时,HA-DAP正确分类的样本数有所减少,甚至低于DAP,这是由于冗余的非语义属性降低了分类器的性能。

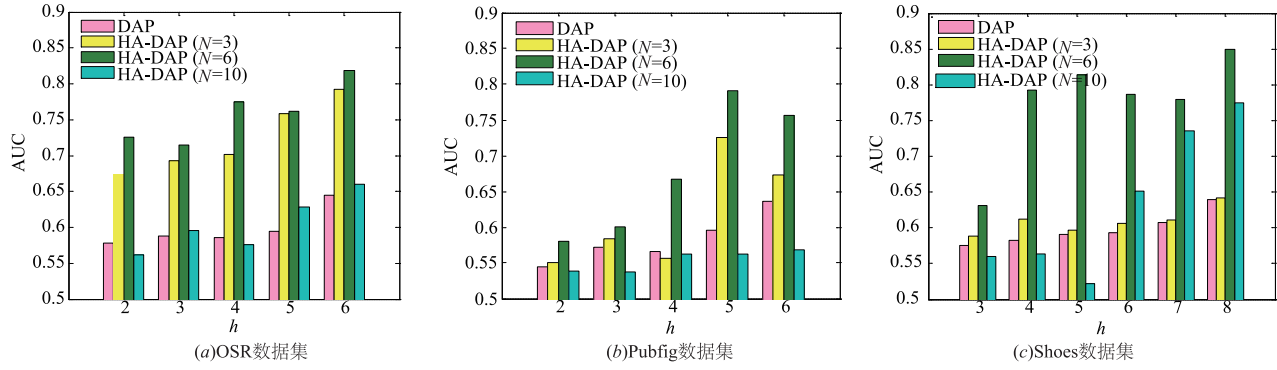


图5 零样本图像分类AUC值

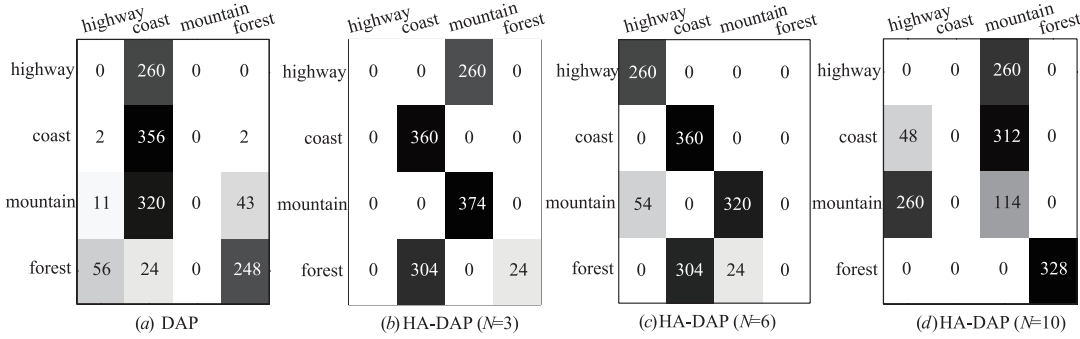


图6 分类结果混淆矩阵 (OSR数据集)

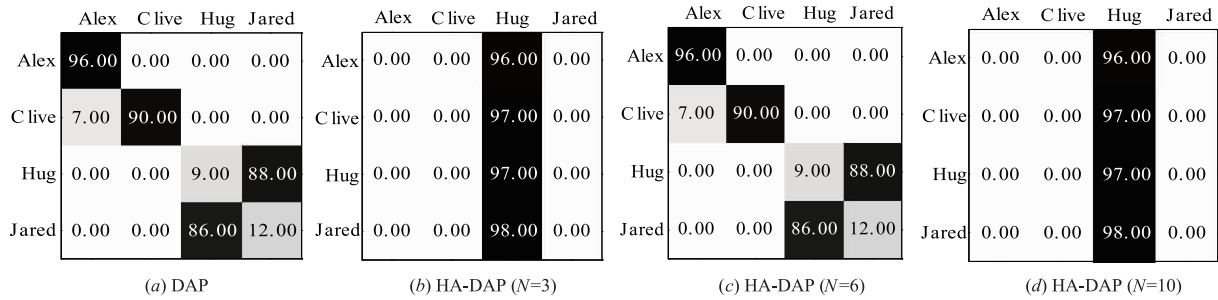


图7 分类结果混淆矩阵(Pub Fig数据集)

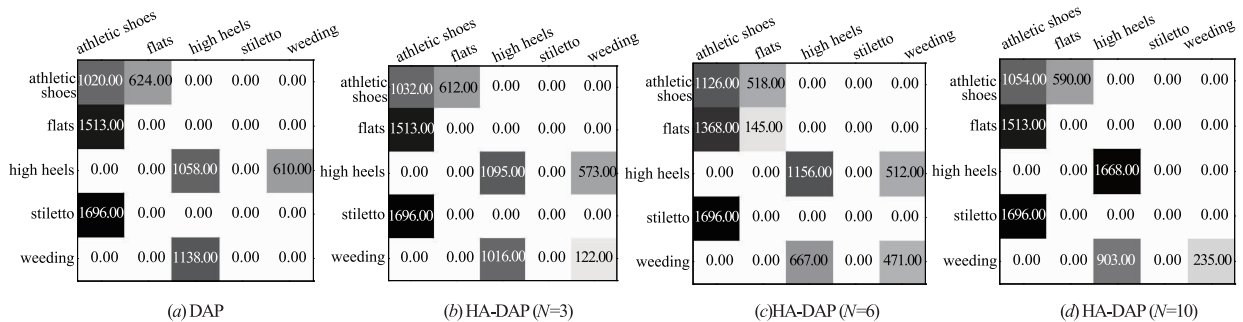


图8 分类结果混淆矩阵 (Shoes数据集)

4 结论

零样本分类主要用于解决标记训练样本不足以涵盖所有对象类的情况下,如何对训练阶段不可见的新对象进行正确分类.利用图像的语义属性是有效解决零样本

图像分类问题的关键.现有的基于属性的零样本图像分类器模型均忽略了样本的底层特征,也就是说,仅根据样本的语义属性来确定测试类样本的类别标签.但是,对那些属性很相似的类别而言,语义属性就很难对它们进行区分.为此,考虑到属性的有限性,提出对样本的底层特

征进行稀疏编码,将编码后的特征作为非语义属性并对有限的语义属性进行补充以组成混合属性.进一步,将构造的混合属性应用于 DAP,提出一种基于混合属性的零样本图像分类方法.将 HA-DAP 与 DAP 在三个数据集上进行了零样本图像分类实验,分类结果表明:由于非语义属性增加了语义属性的差异性,HA-DAP 在零样本图像分类中的分类性能优于 DAP.

参考文献

- [1] Lampert C H, Nickisch H, Harmeling S. Attribute-based classification for zero-shot visual object categorization[J]. *Pattern Analysis and Machine Intelligence*, 2014, 36(3): 453 - 465.
- [2] Douze M, Ramisa A, Schmid C. Combining attributes and Fisher vectors for efficient image retrieval[A]. *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition [C]*. Colorado Springs: IEEE Press, 2011. 745 - 752.
- [3] B Siddiquie, R Feris, L Davis. Image ranking and retrieval based on multi-attribute queries[A]. *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition [C]*. Colorado Springs: IEEE Press, 2011. 801 - 808.
- [4] Farhadi A, Endres I, Hoiem D, et al. Describing objects by their attributes[A]. *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition [C]*. Miami: IEEE Press, 2009. 1778 - 1785.
- [5] Kumar N, Berg A, Belhumeur P, et al. Describable visual attributes for face verification and image search[J]. *Pattern Analysis and Machine Intelligence*, 2011, 33(10): 1962 - 1977.
- [6] Mahajan D, Sellamanickam S, Nair V. A joint learning framework for attribute models and object descriptions[A]. *Proceedings of the IEEE International Conference on Computer Vision [C]*. Barcelona: IEEE Press, 2011. 1227 - 1234.
- [7] Liu M X, Zhang D Q, Chen S C. Attribute relation learning for zero-shot classification[J]. *Neurocomputing*, 2014, 139(2): 34 - 46.
- [8] Larochelle H, Erhan E, Bengio Y. Zero-data learning of new tasks[A]. *Proceedings of the International Conference on Artificial Intelligence [C]*. Chicago: AAAI Press, 2008. 646 - 651.
- [9] Yu X, Aloimonos Y. Attribute-based transfer learning for object categorization with zero/one training example[A]. *Proceedings of the International Conference on Computer Vision [C]*. Heraklion: Springer Verlag Press, 2010. 127 - 140.
- [10] Lampert C H, Nickisch H, Harmeling S. Learning to detect unseen object classes by between-class attribute transfer[A]. *Proceedings of the IEEE International Computer Vision and Pattern Recognition [C]*. Miami: IEEE Press, 2009. 951 - 958.
- [11] 石光明, 刘丹华, 高大化, 等. 压缩感知理论及其研究进展[J]. *电子学报*, 2009, 37(5): 1071 - 1078.
Shi Guang-ming, Liu Dan-hua, Gao Da-hua, et al. Advances in theory and application of compressed sensing[J]. *Acta Electronica Sinica*, 2009, 37(5): 1071 - 1078. (in Chinese)
- [12] Sharmanska V, Quadrianto N, Lampert C H. Augmented attribute representations[A]. *Proceedings of the International Conference on Computer Vision [C]*. Florence: Springer Verlag Press, 2012. 242 - 255.
- [13] 亓晓振, 王庆. 一种基于稀疏编码的多核学习图像分类算法[J]. *电子学报*, 2012, 40(4): 773 - 779.
Qi Xiao-zhen, Wang Qing. An image classification approach based on sparse coding and multiple kernel learning[J]. *Acta Electronica Sinica*, 2012, 40(4): 773 - 779. (in Chinese)
- [14] 黄丽丽, 肖亮, 韦至辉. 彩色图像去马赛克的非局部稀疏表示方法[J]. *电子学报*, 2014, 42(2): 272 - 279.
Huang Li-li, Xiao Liang, Wei Zhi-hui. A nonlocal sparse representation method for color demosaicking[J]. *Acta Electronica Sinica*, 2014, 42(2): 272 - 279. (in Chinese)
- [15] 孙玉宝, 韦志辉, 肖亮, 等. 多形态稀疏性正则化的图像超分辨率算法[J]. *电子学报*, 2010, 38(12): 2898 - 2903.
Sun Yu-bao, Wei Zhi-hui, Xiao Liang, et al. Multimorphology sparsity regularized image super-resolution[J]. *Acta Electronica Sinica*, 2010, 38(12): 2898 - 2903. (in Chinese)
- [16] Fawcett Tom. An introduction to ROC analysis[J]. *Pattern Recognition Letters*, 2006, 27(8): 861 - 874.

作者简介



程玉虎 男, 1973 年出生于安徽淮南, 中国矿业大学信息与控制工程学院教授, 博士生导师. 主要研究方向为机器学习与模式识别、智能系统等.
E-mail: chengyuhu@163.com



乔雪 女, 1990 年出生于四川成都, 中国矿业大学信息与控制工程学院硕士研究生. 主要研究方向为迁移学习、属性学习等.
E-mail: qiaoxuecumt@163.com

王雪松(通信作者) 女, 1974 年出生于安徽泗县, 中国矿业大学信息与控制工程学院教授, 博士生导师. 主要研究方向为机器学习与模式识别.
E-mail: wangxuesongcumt@163.com