

基于深度学习的行人属性识别

袁配配, 张良*

中国民航大学天津市智能信号与图像处理重点实验室, 天津 300300

摘要 针对监控场景的背景杂乱及行人被遮挡等问题, 提出一种基于背景抑制的行人属性识别方法, 该方法可以减小背景对行人属性识别的影响。首先, 改进卷积神经网络以生成三个分支, 将分支分别用于行人图像、人体区域、背景区域的特征提取; 然后, 将区域对比损失函数和加权交叉熵损失函数作为网络的联合代价函数。在此联合代价函数的约束下, 神经网络学习到的特征具有背景杂乱不变性, 从而提高了行人属性识别的准确度。将所提方法在 PETA 和 RAP 两个行人属性数据集上进行验证。与其他现有方法相比, 所提方法在平均精度、准确度、精确度等指标上性能均有所提升, 证明了所提方法的有效性。

关键词 图像处理; 卷积神经网络; 行人属性识别; 语义分割; 联合代价函数

中图分类号 TP391

文献标志码 A

doi: 10.3788/LOP57.061001

Pedestrian Attribute Recognition Based on Deep Learning

Yuan Peipei, Zhang Liang*

Tianjin Key Laboratory of Advanced Signal and Image Processing, Civil Aviation University of China,
Tianjin 300300, China

Abstract In this study, we propose a pedestrian attribute recognition method based on background suppression to solve the problems of background clutter and object occlusion associated with the monitor scene. The proposed method can reduce the impact of the background on pedestrian attribute recognition. First, three branches are generated by improving the convolutional neural network. These three branches are used to extract the features of the pedestrian images, human body regions, and background regions. Then, the regional contrast loss function and weighted cross-entropy loss function are considered to constitute the joint cost function of the network. The features learned by the neural network exhibit background clutter invariance under the constraint of the joint cost function. Therefore, the proposed method can improve the pedestrian attribute recognition accuracy. The proposed method was verified using the PETA and RAP pedestrian attribute datasets. The results denote that the proposed method exhibits improved the mean accuracy, accuracy, precision, and other performance indicators when compared with those exhibited by the remaining methods, confirming its effectiveness.

Key words image processing; convolutional neural network; pedestrian attribute recognition; semantic segmentation; joint cost function

OCIS codes 100.4996; 150.1135; 100.5010; 110.2960

1 引言

行人属性识别^[1]旨在挖掘监控场景中行人可被观察到的一些高级语义信息, 如性别、年龄、服饰、携带品等。通常, 这些信息可通过人工观察的方式进行提取, 这会导致工作量巨大且十分耗时。行人属

性识别可以无需人工操作地将这些信息转换成可以用于检索的高级语义信息。目前, 行人属性识别存在 2 个主要的挑战。1) 模型将背景区域特征误判为某个属性的特征, 从而产生一定的特征提取损失; 2) 行人间的互相遮挡会使模型混淆提取。

随着深度学习^[2-3]在计算机视觉任务中的表现

收稿日期: 2019-06-26; 修回日期: 2019-08-02; 录用日期: 2019-08-21

基金项目: 国家自然科学基金(61179045)、民航安全能力建设项目(20600523)

* E-mail: l-zhang@cauc.edu.cn

越来越出色,研究人员开始尝试利用卷积神经网络提取监控场景中行人图像的属性特征。卷积神经网络使用多层非线性变换进行自动特征提取,能更有效地表征行人图像的特征。

总体来说,国内外基于深度学习的行人属性识别算法有4种:1)对损失函数进行改进,以解决数据不平衡的问题;2)基于局部特征的行人属性识别方法结合了全局和细粒度部件的特征;3)基于注意力机制的属性识别方法;4)利用联合循环学习提取上下文信息,以辅助属性识别。李亚鹏等^[4]通过加入正样本比例因子改进损失函数的方法使网络的性能得到提高。Li等^[5]提出可用于学习不同属性间相关性的卷积神经网络模型DeepMAR。这两种算法均属于第一类算法,模型较为简单,难以进一步提升识别效果。Li等^[6]使用poselet部件检测器得到分割后的行人部件,再将其与行人整体结合,提取深度特征进行行人属性识别,这种算法属于第二类算法,但该部件分割算法在监控场景中效果不佳。Liu等^[7]提出一个基于注意力机制的深度网络HydraPlus-Net,这属于第三类算法,其能够捕获细节和局部特征,但模型参数较多且训练过程较为繁琐。Wang等^[8]提出联合循环学习模型,该模型可挖掘属性上下文信息及属性间的相互关系,可提升识别准确率,属于第四类算法,但该模型复杂度较高,当训练集较小时,易出现过拟合现象。

除上述算法外,凌弘毅^[9]使用知识蒸馏方法挖掘被丢弃的属性信息,以辅助属性识别,但被丢弃的属性样本非常少,准确率提升效果并不明显。知识蒸馏法常用于模型压缩,教师网络为复杂模型,学生网络为精简模型。教师网络可利用学习到的隐藏知识指导学生网络训练,实现“知识迁移”。针对行人属性识别现存的问题,受Song等^[10]在研究行人再识别问题时所使用方法的启发,本研究提出基于背景抑制的行人属性识别方法。以残差网络^[11-13]为基础网络,增加两个分支,辅助网络提取不同区域的特征,同时利用联合代价函数减小行人图像特征与背景区域特征的相似度,增大行人图像特征与人体区域特征的相似度,以此来减小背景对网络特征提取的影响。通过在公开数据集上的实验验证,所提方法能取得良好的识别效果。

2 行人属性识别网络构建

2.1 训练网络的整体结构

背景抑制的行人属性识别网络结构如图1所

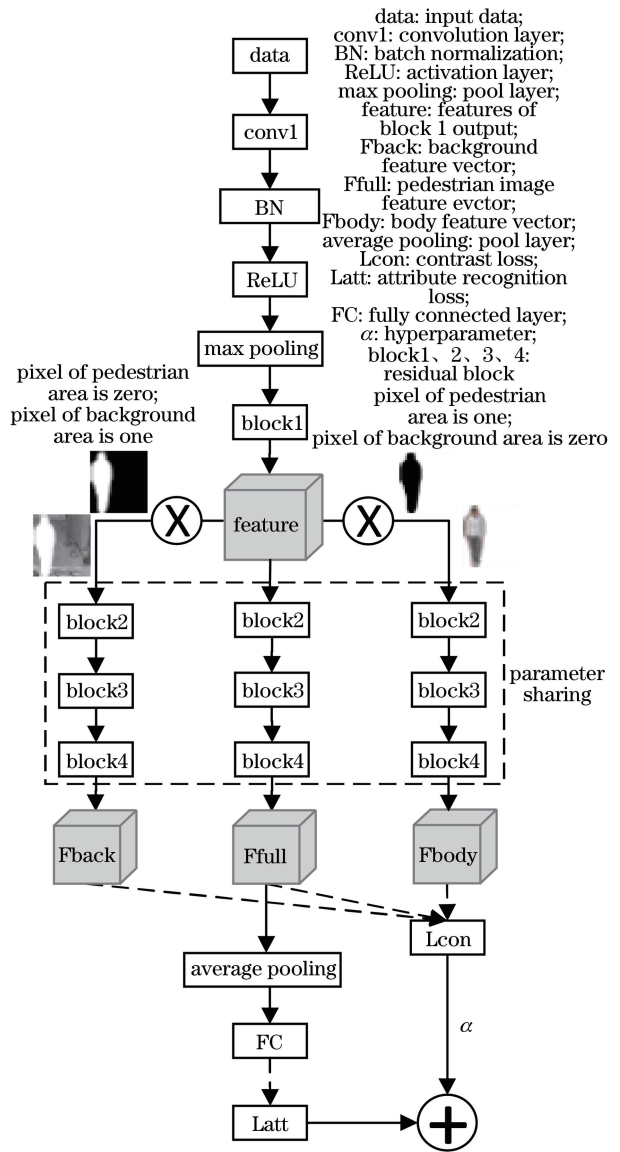


图1 行人属性识别网络

Fig. 1 Pedestrian attribute recognition network

示。中间分支对行人图像进行特征提取,左侧分支对背景区域进行特征提取,右侧分支对人体区域进行特征提取,三个分支参数共享。重新定义网络全连接层的神经元数量,使其与行人属性数据集的属性数量一致。网络分类层的联合代价函数由加权交叉熵损失函数和区域对比损失函数组成。

中间分支提取行人图像特征后通过全局平均池化层进行降采样,然后通过全连接层将网络提取的特征映射到样本标记空间中,最后通过加权交叉熵损失函数计算网络预测结果与真实标签之间的误差,该误差即为属性识别的损失。通过区域对比损失函数计算网络特征提取的损失,然后将两个损失按比例相加作为网络总的损失,再通过随机梯度下降算法对参数进行优化更新。block1、block2、

block3、block4 为四个残差块,具体实现细节可查阅文献[11],这里不再赘述。

2.2 残差块结构

残差网络^[11]由何凯明团队提出,其在 ILSVRC2015 比赛中取得冠军,残差网络在前向网络中增加了一些快捷连接。这些连接会跳过某些层,将原始数据直接传到之后的层。这些连接即可称为残差块,结构如图 2 所示。

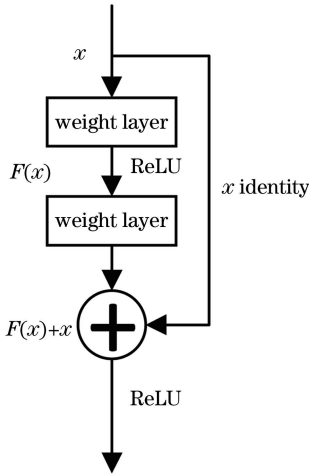


图 2 残差块结构

Fig. 2 Structure of residual block

2.3 人体区域及背景区域的特征提取

由于行人属性数据集中的图像分辨率较低,清晰度较差,以往很少有利用行人掩码来辅助提高识别准确率的属性识别方法。随着大量基于深度学习的图像分割算法^[14-16]被提出,对行人掩码的提取成为一种可能。

图像分割分为语义分割和实例分割。语义分割是将图中每一点像素标注为某个类别,同一物体的不同实例不需要单独分割出来,相对于语义分割,实例分割则需要将每一个物体的不同实例分割开,如图 3 所示。



图 3 图像分割示例。(a)待分割图像;(b)语义分割结果;(c)实例分割结果

Fig. 3 Examples of image segmentation. (a) Image to be segmented; (b) result of semantic segmentation; (c) result of instance segmentation

如果要将行人与背景分离,并且无需对每个行人单独分割,可采用图像语义分割方法中较为经典

的 Deeplab-v2^[14]算法进行行人掩码提取。对于数据集中少量错误提取的图像,可采用手工标注的方法提取掩码。

行人掩码为一个与行人图像等高等宽等通道的二值图像。将行人图像中人体区域的像素值置为 1,背景区域的像素值置为 0,即可获得行人掩码。人体区域特征提取如图 1 右侧分支所示,具体过程可分为 3 步:

1)对行人掩码进行尺度变换,使变换结果的维度与 block1 的输出维度相同。

2)将行人图像作为网络的输入,将尺度变换得到的行人掩码与 block1 输出的特征图相乘,得到人体区域的特征图。

3)将人体区域的特征图作为 block2 的输入,继续提取特征可得到人体区域特征。

背景区域特征提取方式与人体区域特征提取方式相似,利用一个与行人掩码等维度的全 1 矩阵减去行人掩码,得到一个人体区域像素值为 0,背景区域像素值为 1 的二值图像,然后对其尺度变换,再将变换结果与 block1 输出的特征图相乘得到背景区域特征图,将其作为 block2 的输入,继续进行特征提取即可得到背景区域特征。

2.4 加权交叉熵损失函数

如果一张图片中出现某个属性,那么这张图片对于该属性为正样本,否则为负样本。在行人属性数据集中,每个属性的正样本和负样本数目相差很大,PETA^[17]数据集就是其中一例。

对正负样本的损失乘以同样的权重系数,训练得到的网络会把大部分正样本预测成负样本。为了应对这种属性分布不平衡的问题,Li 等^[18]提出加权交叉熵损失函数,该函数可为每个属性的正负样本赋予不同的惩罚权重。当某个属性的正样本数目较少时,该函数可为该属性的正样本赋予一个较大的惩罚权重,同时可为该属性的负样本赋予一个较小的惩罚权重,即

$$L_{\text{att}} = -\frac{1}{N} \sum_{n=1}^N \sum_{l=1}^L [\omega_{\text{pos}} y_{n,l} \ln(p_{n,l}) + \omega_{\text{neg}} (1 - y_{n,l}) \ln(1 - p_{n,l})], \quad (1)$$

$$\omega_{\text{pos}} = \exp[(1 - r_l)/\sigma^2], \quad (2)$$

$$\omega_{\text{neg}} = \exp(r_l/\sigma^2), \quad (3)$$

式中: L 为属性的个数; N 为行人图像的数量; $y_{n,l}$ 为第 n 个样本的第 l 个属性标签, $y_{n,l} \in \{0, 1\}$, $y_{n,l} = 1$ 代表第 n 个样本包含第 l 个属性,反之不包含; $p_{n,l}$ 为属性识别网络对第 n 个样本的第 l 个属性

的预测结果; r_l 为第 l 个属性的正样本比率, 即该属性正样本数目在训练集中所占的比率; ω_{pos} 为第 l 个属性的正样本权重; ω_{neg} 为负样本权重; σ 设置为 1。

2.5 区域对比损失函数

为了减小行人间的互相遮挡以及嘈杂的背景对属性识别带来的影响, 引入区域对比损失函数, 该损失函数以欧氏距离来衡量特征之间的相似度。

$$L_{\text{con}} = \|\mathbf{F}_{\text{full}} - \mathbf{F}_{\text{body}}\|_2^2 + \max\{(m - \|\mathbf{F}_{\text{full}} - \mathbf{F}_{\text{background}}\|_2^2), 0\}, \quad (4)$$

式中: m 为一个超参数, 指行人图像的特征向量与背景特征向量的欧氏距离; \mathbf{F}_{full} 、 \mathbf{F}_{body} 、 $\mathbf{F}_{\text{background}}$ 均为由网络提取的 2048 维的特征向量, 表征行人图像、人体区域、背景区域的特征。

2.6 联合代价函数

为了在完成属性识别的同时, 提高网络在嘈杂背景下特征提取的稳定性, 减小背景影响, 可将加权交叉熵损失函数和区域对比损失函数联合作为网络的损失函数, 即

$$L_{\text{all}} = L_{\text{att}} + \alpha L_{\text{con}}, \quad (5)$$

式中: α 为一个超参数, 用来控制区域对比损失函数在整体损失中所占的比例; L_{att} 为加权交叉熵损失函数计算的行人属性识别的损失; L_{con} 为区域对比损失函数所计算的网路特征提取的损失。

3 实验与结果分析

3.1 行人属性数据集

PETA 数据集包含 19000 个行人图像, 分辨率最小为 17×39 , 最大为 169×365 , 每个行人用 61 个二分类属性标签和 4 个多分类属性标签进行标注。由于某些属性的正样本数量极少, 所以大多数方法主要还是关注 35 个正样本比例较高的属性。该数据集中 9500 个行人图像用于训练, 1900 个行人图像用于验证, 7600 个行人图像用于测试, 在本实验中, 将训练集与验证集合并用于训练。RAP^[19] 数据集包含 41585 张图片, 其中 33268 张图片用于训练, 其余图片用于测试, 分辨率范围从 36×92 到 344×554 , 每个行人用 69 个二分类属性标签和 3 个多分类属性标签进行标注。对于 RAP 数据, 大多数方法只关注其正样本比例较大的 51 个属性。

经过对数据集的观察发现, PETA 数据集存在一个问题, 即该数据集中一个行人存在多张图像, 仅通过随机选择一个示例图像进行一次注释, 然后对该人的多张图像共享相同的注释属性, 会造成一些

图像的标签存在错误, 如图 4 所示。



图 4 行人图像样例。(a)~(f) 样例 1~6

Fig. 4 Samples of pedestrian images. (a)-(f) Sample 1-6

表 1 为图 4 中 6 幅图像的标签及每个属性正样本的比例, 图 4(a)~(c) 标签一致, 图 4(d)~(f) 标签一致。标签 1 代表含有该属性, 0 代表不含有该

表 1 行人图像标签及属性比例

Table 1 Label of pedestrian image and the ratio of attribute

Attribute	Fig. 4(a)-(c)	Fig. 4(d)-(f)	Ratio/%
Age 16-30	0	0	49.7
Age 31-45	0	1	32.9
Age 46-60	0	0	10.2
Age above 61	1	0	6.2
Backpack	0	0	19.7
Carrying other	0	0	19.9
Casual lower	1	1	86.1
Casual upper	1	1	85.3
Formal lower	0	0	13.8
Formal upper	0	0	13.4
Hat	1	1	10.2
Jacket	0	0	6.9
Jeans	0	0	30.6
Leather shoes	1	0	29.6
Logo	0	0	4.0
Long hair	0	0	23.8
Male	1	0	54.9
Messenger bag	0	1	29.6
Muffler	0	1	8.4
No accessory	0	0	74.9
No carrying	1	0	27.6
Plaid	0	0	2.7
Plastic bag	0	0	7.7
Sandals	0	0	0.2
Shoes	0	1	36.3
Shorts	0	0	3.5
Short sleeve	0	0	14.2
Skirt	0	0	4.6
Sneaker	0	0	21.6
Stripes	0	0	1.7
Sunglasses	0	0	2.9
Trousers	1	1	51.5
T-shirt	0	0	8.4
Upper other	1	1	45.6
V-neck	0	0	1.2

属性。虽然图 4(a)、(b)携带塑料袋、鞋子、裤子等属性的标签为 1,但无法通过图像观察到这些特征,而图 4(d)、(e)没有行人出现,这些标签的错误会极大地影响模型的准确率。因此,本文增加 RAP 数据集进行实验验证。

3.2 参数初始化及数据预处理

本实验在构建属性识别网络时使用的深度学习框架是 PyTorch。不同于大多数深度学习框架,PyTorch 在运行期间是使用反向模式自动微分来构建计算图表的。

由于行人属性数据集的数据量较小,而图 4 所示的行人属性识别网络参数量巨大,若单纯使用行人属性数据集对属性识别网络进行训练,网络会出现过拟合现象,所以将在 ImageNet 数据集上预训练的 ResNet-50 模型参数迁移作为网络的初始权重,然后利用行人属性数据集对网络参数进行微调。模型参数迁移过程分为 3 步:

1) 去除预训练模型参数中全连接层的参数。

2) 利用该参数对属性识别网络的 conv1 层及 block1、block2、block3、block4 块进行参数初始化。

3) 全连接层使用均值为 0、标准差为 0.01 的高斯分布进行随机初始化。

在训练前需要对数据进行预处理(如尺度变换、随机水平翻转、数据归一化操作等)。随机水平翻转会随机产生一个 0 到 1 的数,然后判断该数是否小于 0.5,若小于 0.5,则将行人图像与行人掩码翻转,若大于等于 0.5,则不进行翻转。在测试时,只需要作尺度变换及数据归一化操作。

3.3 评估标准

Deng 等^[17]采用平均精度(E_{mA})来评估属性识别算法。对于每个属性,分别计算正样本和负样本的分类准确度,然后将它们的平均值作为属性的识别结果,最后,通过对所有属性取平均值来获得识别率。平均精度的表达式为

$$E_{mA} = \frac{1}{2N} \sum_{i=1}^L \left(\frac{N_{TP_i}}{P_i} + \frac{N_{TN_i}}{N_i} \right), \quad (6)$$

式中: N_{TP_i} 和 N_{TN_i} 分别为正确预测的正样本和负样本的数量; P_i 和 N_i 分别为正样本和负样本的数量。

Yu 等^[20]将上述标准称为基于标签的评估标准,并提出基于实例的评估标准。基于实例的评估标准包括四个指标,即准确度、精确度、召回率和 $F1$ 值,计算方法分别为

$$E_{acc} = \frac{1}{N} \sum_{i=1}^N \frac{|Y_i \cap f(x_i)|}{|Y_i \cup f(x_i)|}, \quad (7)$$

$$E_{prec} = \frac{1}{2N} \sum_{i=1}^N \frac{|Y_i \cap f(x_i)|}{|f(x_i)|}, \quad (8)$$

$$E_{rec} = \frac{1}{2N} \sum_{i=1}^N \frac{|Y_i \cap f(x_i)|}{|Y_i|}, \quad (9)$$

$$F_1 = \frac{2 \times E_{prec} \times E_{rec}}{E_{prec} + E_{rec}}, \quad (10)$$

式中: N 为样本的个数; Y_i 为样本 x_i 的标签; $f(x_i)$ 为样本 x_i 经过网络预测的结果。

3.4 实验结果与分析

在训练网络时,采用随机梯度下降法(SGD)优化网络,将初始学习率设置为 0.001,每经过 50 次迭代,将学习率乘以 0.1,迭代 120 次后网络能够达到良好的效果。动量和权重衰减因子根据经验设置为 0.9 和 0.0005。在测试阶段,只使用图 1 的中间分支,并将参数 α 设置为 0。

采用 3 种实验。实验 1,使用图 1 行人属性识别网络中的右侧分支,将联合代价函数的超参数 α 设为 0,则损失函数仅为加权交叉熵损失函数。实验 2,使用图 1 行人属性识别网络中的中间分支,将联合代价函数的超参数 α 设为 0,则损失函数仅为加权交叉熵损失函数。实验 3,使用图 1 中三个分支,即所提算法,将区域对比损失函数的超参数 m 设为 100,联合代价函数的超参数 α 设为 0.0001,迭代 100 次后,将 α 调整为原来的 1/10。

在 PETA 和 RAP 两个数据集上实验,结果分别如表 2 和表 3 所示。

表 2 PETA 数据集的实验结果

Algorithm	E_{mA}	E_{acc}	E_{prec}	E_{rec}	$F1$
ACN ^[21]	81.15	73.66	84.06	81.26	82.64
DeepMAR ^[5]	82.89	75.07	83.68	83.14	83.41
FSPP ^[20]	81.67	75.72	84.84	83.10	83.96
HP-Net ^[7]	81.77	76.13	84.92	83.24	84.07
PGDM ^[18]	82.97	78.08	86.86	84.68	85.76
Experiment1	81.55	76.03	85.20	82.85	84.10
Experiment2	83.37	78.56	87.63	85.03	86.24
Experiment3	84.49	79.44	87.82	85.94	86.87

由表 2 可知,所提算法与 PGDM^[18]相比,平均精度提高 1.52%,准确度提高 1.36%,精确度提高 0.96%,召回率提高 1.26%, $F1$ 值提高 1.11%。可以看出,区域对比损失函数消除背景特征算法优于现有的几种行人属性识别算法。

表 3 RAP 数据集实验结果

Algorithm	E_{mA}	E_{acc}	E_{prec}	E_{rec}	F1
ACN ^[21]	69.66	62.61	80.12	72.26	75.98
DeepMAR ^[5]	73.79	62.02	74.92	76.21	75.56
FSPP ^[20]	79.64	60.25	69.10	80.16	74.21
HP-Net ^[7]	76.12	65.39	77.33	78.79	78.05
PGDM ^[18]	74.31	64.57	78.86	75.90	77.35
Experiment1	79.01	66.85	79.47	80.04	77.94
Experiment2	79.70	67.02	80.54	80.47	78.38
Experiment3	80.11	67.68	81.60	81.55	79.62

由表 3 可知,所提算法与现有算法相比,平均精度比 FSPP^[20] 提高 0.47%,准确度比 HP-Net^[7] 提高 2.29%,精确度比 ACN^[21] 提高 1.48%,召回率比 FSPP^[20] 提高 1.39%,F1 值比 HP-Net^[7] 提高 1.57%。同样,所提算法的各指标优于其他现有的算法。

表 2 和表 3 中,实验 2 的结果均优于实验 1,表明直接去除背景并不能提高行人属性识别的准确率,原因主要有 2 点:1)无论使用深度学习还是手工标注,都无法精准地将行人与背景完整分离开,可能会出现行人、人体区域被误认为背景等情况,这会大大影响属性识别的效果;2)直接去除背景,会影响图像的平滑性。因此,可使用区域对比损失函数来提高网络提取特征的稳定性。图 5 反映了两个数据集训练过程中区域对比损失函数损失值的变化趋势。由图 5 可知,随着迭代次数的增加,区域对比损失函数的损失值在逐渐降低并趋于平稳,说明行人图像特征与人体区域特征的相似度越来越高,与背景区域特征的相似度越来越低。

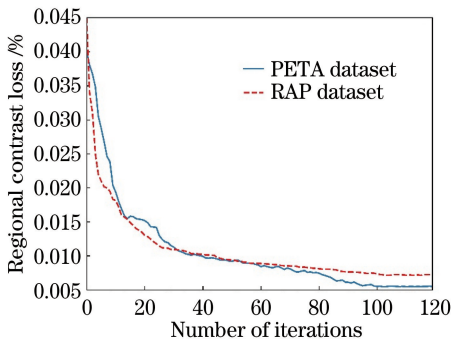
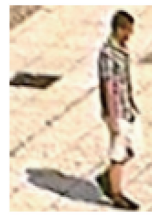


图 5 训练过程中区域对比损失函数的变化趋势

Fig. 5 Variation trend of regional contrast loss function during training

对单张图像进行测试,结果如图 6 所示。由标签和属性识别结果可知,所提方法能比较准确地识别出图像中的行人属性。



Label:	The result of attribute recognition :
Personal less 30	Personal less 30
Lower body casual	Lower body casual
Upper body casual	Upper body casual
Personal male	Personal male
Accessory nothing	Upper body plaid
Upper body plaid	Foot wear shoes
Foot wear shoes	Lower body shorts
Lower body shorts	Upper body shorts sleeve
Upper body shorts sleeve	

图 6 属性识别结果

Fig. 6 Result of attribute recognition

4 结 论

提出基于背景抑制的行人属性识别方法。通过改进残差网络,并引入区域对比损失函数和加权交叉熵损失函数,网络提取的特征会集中于人体区域,能减小背景对于行人属性识别的影响,强制网络提取的特征具有一定的背景杂乱不变性,从而提高属性识别效果。实验结果表明,与现有方法相比,所提方法具有较好的性能。但所提模型仍然存在一些不足,如图像语义分割方法是独立地对行人图像进行掩码提取的,后续可以将掩码的提取过程整合到行人属性识别的网络中,实现一个端到端的网络,以进一步提高网络的性能。

参 考 文 献

- [1] Wang X, Zheng S F, Yang R, et al. Pedestrian attribute recognition: a survey[J/OL]. (2019-01-22) [2019-04-18]. <https://arxiv.org/abs/1901.07474>.
- [2] Zhao J T. Single-image defogging algorithm based on deep learning[J]. Laser & Optoelectronics Progress, 2019, 56(11): 111005.
赵建堂. 基于深度学习的单幅图像去雾算法[J]. 激光与光电子学进展, 2019, 56(11): 111005.
- [3] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [4] Li Y P, Wan S R. Multi-label recognition of pedestrian attributes based on deep learning [J]. Chinese Journal of Biomedical Engineering, 2018, 37(4): 423-428.
李亚鹏, 万遂人. 基于深度学习的行人属性多标签识别[J]. 中国生物医学工程学报, 2018, 37(4): 423-428.
- [5] Li D W, Chen X T, Huang K Q. Multi-attribute learning for pedestrian attribute recognition in surveillance scenarios [C] // 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), November 3-6, 2015, Kuala Lumpur, Malaysia.

- New York: IEEE, 2015: 111-115.
- [6] Li Y N, Huang C, Loy C C, et al. Human attribute recognition by deep hierarchical contexts[M] // Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9910: 684-700.
- [7] Liu X H, Zhao H Y, Tian M Q, et al. HydraPlus-Net: attentive deep features for pedestrian analysis[C] // 2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 350-359.
- [8] Wang J Y, Zhu X T, Gong S G, et al. Attribute recognition by joint recurrent learning of context and correlation[C] // 2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 531-540.
- [9] Ling H Y. Pedestrian attribute recognition based on knowledge distillation [J]. Computer Applications and Software, 2018, 35(10): 181-184, 193.
凌弘毅. 基于知识蒸馏方法的行人属性识别研究 [J]. 计算机应用与软件, 2018, 35(10): 181-184, 193.
- [10] Song C F, Huang Y, Ouyang W L, et al. Mask-guided contrastive attention model for person re-identification [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 1179-1188.
- [11] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 770-778.
- [12] Li J, Guan Y P. Pedestrian re-identification based on adaptive weight assignment using deep learning for pedestrian attributes [J]. Laser & Optoelectronics Progress, 2019, 56(14): 141003.
李净, 管业鹏. 基于深度学习行人属性自适应权重分配行人再识别方法[J]. 激光与光电子学进展, 2019, 56(14): 141003.
- [13] Zhang C, Chen Y. Object detection based on hard examples mining using residual network[J]. Laser & Optoelectronics Progress, 2018, 55(10): 101003.
张超, 陈莹. 残差网络下基于困难样本挖掘的目标检测[J]. 激光与光电子学进展, 2018, 55(10): 101003.
- [14] Chen L C, Papandreou G, Kokkinos I, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834-848.
- [15] Pei L, Liu Y, Gao L. Cloud detection of ZY-3 remote sensing images based on fully convolutional neural network and conditional random field [J]. Laser & Optoelectronics Progress, 2019, 56(10): 102802.
裴亮, 刘阳, 高琳. 结合全卷积神经网络与条件随机场的资源3号遥感影像云检测[J]. 激光与光电子学进展, 2019, 56(10): 102802.
- [16] Guo C C, Yu F Q, Chen Y. Image semantic segmentation based on convolutional neural network feature and improved superpixel matching [J]. Laser & Optoelectronics Progress, 2018, 55(8): 081005.
郭呈呈, 于凤芹, 陈莹. 基于卷积神经网络特征和改进超像素匹配的图像语义分割[J]. 激光与光电子学进展, 2018, 55(8): 081005.
- [17] Deng Y B, Luo P, Loy C C, et al. Pedestrian attribute recognition at far distance [C] // Proceedings of the ACM International Conference on Multimedia-MM '14, November 3-7, 2014, Orlando, Florida, USA. New York: ACM, 2014: 789-792.
- [18] Li D W, Chen X T, Zhang Z, et al. Pose guided deep model for pedestrian attribute recognition in surveillance scenarios [C] // 2018 IEEE International Conference on Multimedia and Expo (ICME), July 23-27, 2018, San Diego, CA, USA. New York: IEEE, 2018: 18163829.
- [19] Li D W, Zhang Z, Chen X T, et al. A richly annotated dataset for pedestrian attribute recognition [J/OL]. (2016-04-27)[2019-04-18]. <https://arxiv.org/abs/1603.07054>.
- [20] Yu K, Leng B, Zhang Z, et al. Weakly-supervised learning of mid-level features for pedestrian attribute recognition and localization [J/OL]. (2016-11-17)[2019-04-18]. <https://arxiv.org/abs/1611.05603>.
- [21] Sudowe P, Spitzer H, Leibe B. Person attribute recognition with a jointly-trained holistic CNN model [C] // 2015 IEEE International Conference on Computer Vision Workshop (ICCVW), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 329-337.