

# 基于图像增强与样本平衡优化的行人属性识别<sup>①</sup>



韦学艳, 吴春雷, 王雷全, 吴杰, 李阳

(中国石油大学(华东) 计算机科学与技术学院, 青岛 266580)

通信作者: 吴春雷, E-mail: wuchunlei@upc.edu.cn

**摘要:** 由于采用的监控图像尺寸较小、分辨率低等原因, 行人属性识别一直是一个极具挑战性的任务, 而低分辨率的图像往往又导致数据集存在识别行人主体区域不突出、背景噪音干扰严重等问题. 之前的方法大多将未作处理的原始图像作为输入, 使得属性识别效果一直不够理想, 并且, 属性识别的主流数据集通常存在正负样本不平衡的问题, 例如, 许多行人的服装属性分布有着季节性或习俗性的偏差. 因此, 本文提出一个新的深度学习网络——图像增强与样本平衡优化模型 IEBO (image enhancement and sample balance optimization). 该模型通过色彩增强与提取行人主体区域的噪音抑制方法, 在突出行人核心特征的同时消除无用背景信息, 防止其对属性识别造成干扰. 另外模型通过权重调节针对样本不平衡的属性进行优化, 提高不平衡属性的识别能力. 实验最终表明, 新的行人属性识别模型在 Market-1501-attribute 数据集中取得了较好的性能.

**关键词:** 行人属性识别; 图像增强; 色彩增强; 背景噪音抑制; 样本平衡优化; 深度学习

引用格式: 韦学艳, 吴春雷, 王雷全, 吴杰, 李阳. 基于图像增强与样本平衡优化的行人属性识别. 计算机系统应用, 2022, 31(6): 202-209. <http://www.c-s-a.org.cn/1003-3254/8512.html>

## Pedestrian Attribute Recognition Based on Image Enhancement and Sample Balance Optimization

WEI Xue-Yan, WU Chun-Lei, WANG Lei-Quan, WU Jie, LI Yang

(College of Computer Science and Technology, China University of Petroleum (East China), Qingdao 266580, China)

**Abstract:** Pedestrian attribute recognition has been a challenging task due to the small size and low resolution of the surveillance images used, and the low-resolution of the images often leads to a dataset with problems such as indistinct main part of recognized pedestrians and serious background noise interference. Most of the previous methods use unprocessed raw images as inputs, so the attribute recognition results are often unsatisfactory. Moreover, the mainstream datasets for attribute recognition usually suffer from imbalanced positive and negative samples. For example, many pedestrians manifest seasonal or customary deviations in the distribution of clothing attributes. Therefore, this study proposes a new deep learning network, namely, the image enhancement and sample balance optimization (IEBO) model. In this model, methods of color enhancement and noise suppression in the main part of the extracted pedestrians are applied to highlight the core features of the pedestrians and eliminate useless background information to prevent it from interfering with attribute recognition. In addition, the model is optimized through weight adjustment to deal with sample imbalanced attributes and thereby improve the recognition of imbalanced attributes. The experiments show that the new pedestrian attribute recognition model achieves better performance on the Market-1501-attribute dataset.

**Key words:** pedestrian attribute recognition; image enhancement; color enhancement; background noise suppression; sample balance optimization; deep learning

<sup>①</sup> 基金项目: 山东省自然科学基金 (ZR2020MF136); 国家自然科学基金 (62072469)

收稿时间: 2021-09-03; 修改时间: 2021-09-26; 采用时间: 2021-10-09; csa 在线出版时间: 2022-05-26

## 1 引言

行人属性识别,其任务的主要目的是挖掘给定图像中行人的属性特征,进而对属性进行分类,最终能够预测测试集中行人的属性标签,比如性别、年龄、衣服颜色、衣服样式等。属性识别一直是计算机视觉中比较热门的领域,在许多方向都有着广泛的应用,比如基于属性的人脸相关技术<sup>[1-3]</sup>已经相当成熟,应用于日常生活中的各个领域。随着近几十年监控设备在全世界普及,监控场景下的行人属性识别<sup>[4-6]</sup>任务也越来越受到重视,但是由于监控设备和普通设备所拍摄的图像的差异,往往存在以下几个问题:(1)图像尺寸较小,分辨率低;(2)行人图像为抓拍图像,可能存在姿态变换、遮挡、模糊等情况;(3)摄像头为固定摄像头,受天气、光线、角度等外界因素影响;(4)数据集往往从某一区域某一时间段内获取,通常存在严重的特殊性。

深度学习出现以后,大量的行人属性识别模型对上述问题做出了贡献。Fabbri等人<sup>[7]</sup>基于对抗生成网络的模型对遮挡图片进行生成补充,扩充行人的特征信息;Li等人<sup>[8]</sup>基于图来建模区域和属性间的空间关系和语义关系。这些方法有效地改进了遮挡以及分辨率较低情况下的属性识别,但是处理方式仅仅针对图像特征提取方式进行改进处理,忽视了数据集本身存在的问题,比如行人主体区域不突出、背景噪音干扰严重以及现实生活中季节性以及习俗性导致的属性正负样本分布不平衡的问题。

从数据集角度来说,虽然图像都为单行人图像,但是由于监控拍摄的随机性,图像的分辨率通常比较低并且仍然存在着许多无用的背景噪音信息,如图1,背景中的汽车、自行车等信息往往会对属性识别造成不利影响。另外,随着跨任务模型<sup>[9,10]</sup>的出现,许多其他任务的数据集尝试被标注上属性信息,应用于行人属性识别,比如基于传统行人重识别数据集改进的Market-1501-attribute数据集,把属性信息与行人身份编号信息相结合,将其应用于行人属性识别和行人重识别任务。然而,由于数据集的提出并不是以属性识别作为基础解决任务,因此数据集通常包含着严重的正负样本不平衡的问题,如图2,由于社会习俗原因,下衣为黑色的样本占比较多。

因此,基于上述观察,本文提出了一种图像增强与样本平衡优化模型 IEBO (image enhancement and sample balance optimization),在图像特征提取之前,先

进行色彩增强并通过提取行人主体区域的方式减少背景噪音对行人信息的干扰,再经过深层神经网络提取行人核心特征,然后基于身份信息针对不平衡属性进行样本平衡优化,最终达到提高属性识别性能的目的。



图1 背景中的汽车、自行车对行人特征造成影响



图2 下衣为黑色的正样本数量远大于负样本数量

综上所述,本文的贡献是:

(1) 提出基于色彩增强与噪音抑制的图像增强方法,突出行人核心特征并挖掘行人整体图像,抑制背景噪音信息防止其对属性识别造成干扰。

(2) 提出基于身份信息融合的行人属性样本平衡优化算法,调节跨任务数据集正负样本不平衡的缺陷,提升模型的识别能力。

(3) 在跨任务 Market-1501-attribute 数据集的样本不平衡属性中取得了较好的结果。

## 2 相关工作

行人属性识别主要是将给定图像中行人的属性特征挖掘出来,进而对各个属性进行分类,最终生成的识别模型能够预测图像中行人的属性标签,比如性别、年龄、衣服颜色、衣服样式等。其中头发、衣服、鞋子等属性,可以具体定位到图像中的某一区域,被称为是低级属性,而像性别、年龄等属性是较为抽象的概念,并不针对某一区域,则被称为是高级属性。与传统手工提取的特征相比,神经网络提取出的属性特征可

以看作高级语义信息,从而能够判断出更深层次的语义含义.因此,计算机视觉领域中许多的任务通过结合属性识别技术来辅助其执行从而获得更好的性能,如行人重识别、人脸检测等.但是由于视角变换、遮挡、低分辨率、数据集样本不平衡等原因,行人属性识别的进一步应用仍是一个亟待解决的难题.

## 2.1 属性识别

近些年深度学习的稳步发展,使基于神经网络的深度学习算法逐渐广泛应用于行人属性识别任务当中<sup>[11,12]</sup>.与传统的手工特征方法相比,基于卷积神经网络的深度学习方法在行人属性识别精度上获得了更好的结果.Li 等人<sup>[9]</sup>考虑到属性任务的统一性,基于深度学习方法提出了针对单属性和多属性分类的模型 DeepSAR、DeepMAR,单属性分类一次只预测一个属性,而多属性分类一次可以预测多个属性,并同时提出了数据集样本不平衡的问题,但是该模型只考虑了正样本过少的情况,忽略了正样本过多以及不同类别的属性分布不一致的情况.Fabrizi 等人<sup>[7]</sup>针对行人遮挡以及分辨率低等问题,提出了基于对抗生成网络的模型 GAM,利用生成网络对遮挡图像进行生成补充,扩充图像特征信息;Li 等人<sup>[8]</sup>考虑到属性之间的相关性,提出了基于推理框架的模型 VSGR,针对区域到区域,属性到属性,区域到属性的空间关系和语义关系进行建模并预测属性信息,在一定程度上减小了行人遮挡以及分辨率低带来的影响.但是这些方法大多针对图像特征提取的方式进行改进处理,并没有考虑到原始图像本身存在的问题,比如行人核心特征不突出、背景噪音信息干扰严重以及正负样本不平衡等.

## 2.2 跨任务模型

随着深度学习的崛起,跨任务研究也开辟了新的热点方向,行人属性识别任务与行人重识别任务的结合,使各自性能都有了很大的提升.行人属性识别的通用数据集一般以 PETA<sup>[13]</sup> 为主,后来由于多任务融合的出现,传统的数据集已不能满足需求.Lin 等人<sup>[14]</sup>在行人重识别任务的传统数据集 Market-1501<sup>[15]</sup> 上标注了行人属性,并提出了以行人属性辅助行人重识别的方法;Sun 等人<sup>[10]</sup>也同时将行人属性集成到行人重识别模型当中,辅助行人的辨识任务;腾讯优图实验室的 Liu 等人<sup>[16]</sup>提出了一个新的模型 JCM,将属性和图像映射为一个序列统一进行学习,并使用行人身份信息辅助行人属性识别;Yang 等人<sup>[17]</sup>根据属性层次和身份

层次异同,提出层次特征模型 HFE,通过行人身份信息聚合同一行人的图像,以此提高属性识别任务的性能.但是以行人身份信息辅助行人属性识别,其任务核心是属性识别,如果不合理利用行人身份信息,将会导致模型忽略掉属性识别任务的最初目标,给模型带来不利的影响.虽然跨任务研究进一步提升了行人属性识别的效果,然而在自然场景下,行人属性识别仍然面临着多重挑战.

## 3 图像增强与样本平衡优化网络

本文提出一个新的深度学习网络图像增强与样本平衡优化模型 IEBO,如图 3,该模型包含 IE、SBO 两部分,给定 1 张输入图像,经过 IE 模块对图像进行优化处理,使行人特征更加突出,之后经过传统卷积网络 ResNet50 提取行人特征,最后经过 SBO 模块,对不平衡属性进行优化处理,最终预测行人的整体属性值.本节分别从图像增强模块、样本平衡优化模块分别对模型进行展开讨论.

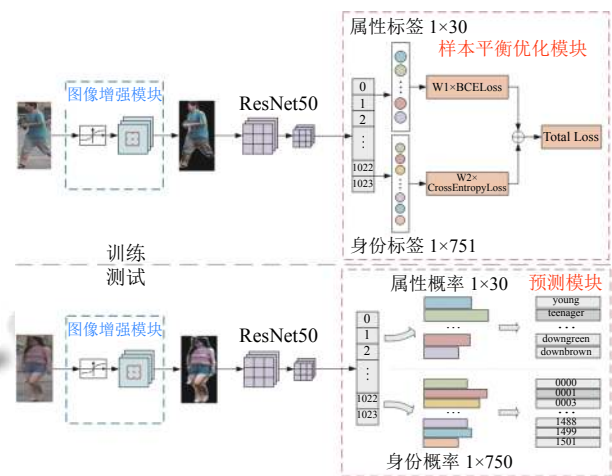


图3 IEBO模型整体结构图

### 3.1 图像增强模块

基于监控系统的行人属性识别,由于数据集图像尺寸较小、分辨率低,在进行特征提取时往往存在主体区域不突出、背景噪音存在等问题,这通常会导致行人核心特征提取不充分,识别结果不理想.因此,特征提取前的图像处理至关重要.输入行人图像之后,首先通过色彩增强<sup>[18]</sup>对图像进行处理,目的是使行人核心特征更加突出,色彩更加饱和,有利于行人特征信息的提取.



第一阶段适应局部图像对比度, 首先将像素值映射到 [0, 1] 范围, 然后分别对彩色图像的 RGB 通道单独进行处理, 计算每个像素点的  $F(X)$  值. 公式如下:

$$F(X) = \sum_{x_0 \in \Omega \setminus x} \frac{G(I(X) - I(X_0))}{\|x - x_0\|} \quad (1)$$

其中,  $\Omega$  表示当前像素及其相邻 8 个像素的邻域范围,  $x_0$  表示邻域范围中除当前像素以外的其余像素点, 等价于  $x \cup x_0 = \Omega$ ,  $\|x - x_0\|$  表示两个像素点之间的欧氏距离,  $I(x) - I(x_0)$  表示两个像素值之差,  $G(t) = \min\{\max\{\alpha t, -1\}, 1\}$ , 为坡度函数, 其中,  $\alpha$  为增强因子, 值越大, 细节增强越明显, 默认  $\alpha \geq 1$ .

第二阶段对图像进行调整以获得全局白平衡, 首先通过最大最小归一化将  $F(x)$  拉伸到 [0, 1] 范围, 然后恢复到 RGB 像素值. 公式如下:

$$S(x) = \frac{F(x) - \min F}{\max F - \min F} \quad (2)$$

其中,  $\min F$  为  $F(x)$  函数的最小值,  $\max F$  为  $F(x)$  函数的最大值. 通过色彩增强算法, 提高了行人图像的动态范围和色彩饱和度, 使核心特征更加明显, 有利于更充分的提取行人特征信息.

然而, 图像中仍然包含一些无用的背景信息, 这些信息有可能会对特征提取造成干扰. 因此, 需要对图像进行噪音抑制处理. 图像色彩增强之后, 使用谷歌 Chen 等人<sup>[19]</sup> 提出的基于 ResNet101 的 DeepLabv3 模型对行人背景进行分割, 模型尝试通过级联与并行的空洞卷积的方式对图片进行处理, 如图 4, 前边卷积模块采用传统的 ResNet101 网络结构, 之后通过一层  $rate=2$  的空洞卷积, 并且经过金字塔池化 ASPP 模型, 同时对特征进行  $rate=1, 6, 12, 18$  四个空洞卷积和一个全局池化的并行处理, 其中,  $rate=1$  为普通卷积, 最后将并行特征图进行连接, 并输入到传统的卷积神经网络中进行上采样恢复到原图.

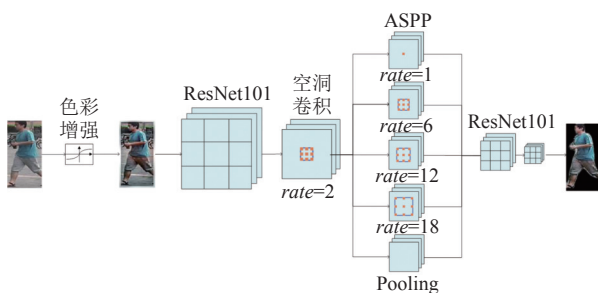


图 4 图像增强模块

在进行行人分割的同时, 将检测出的实例框以外的背景区域置 0. 通过背景置 0 的方式, 将无用的信息去掉, 可以有效防止背景噪音在特征提取过程中对行人核心特征造成干扰.

### 3.2 样本平衡优化模块

近几年, 由于深度学习的发展, 跨任务研究逐渐进入大众视野, 研究人员将行人重识别领域的传统数据集 Market-1501 数据集标注上行人属性信息, 作为可以结合行人身份信息与行人属性信息的数据集进行使用. 因此, 模型利用身份信息进行辅助属性识别, 如图 5.

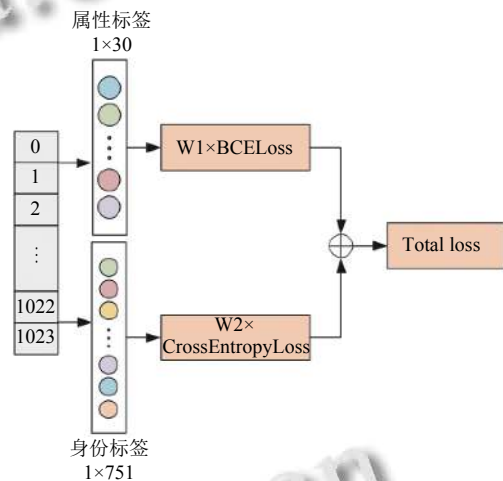


图 5 样本平衡优化模块

然而, 由于数据集不是以属性识别任务进行创建的, 所以在属性方面, 数据集存在正负样本不平衡的问题. 比如在“上衣类型”这一属性中, 正样本“上衣为短袖”远比负样本“上衣为长袖”的样本数量要多; “下衣颜色为红色”这一属性的负样本“颜色不为红色”远比正样本“颜色为红色”的样本数量要多. 所以在进行训练时, 模型往往会出现由于正负样本不平衡而造成的训练不充分、准确率不高等问题. 因此, 我们针对正负样本不平衡属性进行了权重优化处理. 由于行人属性在 Market-1501-attribute 数据集被转换为 30 个二分类问题, 因此属性标签采用二分类交叉熵损失函数进行计算, 公式如下:

$$Loss_F = -\frac{1}{N} \sum_{n=1}^N \sum_{l=1}^L W_l [y_{nl} \times \log x_{nl} + (1 - y_{nl}) \times \log(1 - x_{nl})] \quad (3)$$

$$W_1 = \log \left( \frac{\sum_{n=1}^N y_{nl} \times \sum_{n=1}^N y_{nl}}{L_M} \right) \quad (4)$$

其中,  $n \in 1, 2, \dots, N$ ,  $N$ 表示训练集图像个数,  $l \in 1, 2, \dots, L$ ,  $L$ 表示属性个数,  $x_{nl}$ 表示第 $n$ 张图像第 $l$ 个属性的预测值,  $y_{nl}$ 表示真实值,  $L_M$ 表示各属性样本合理比例, 具体如下:

$$L_M = \frac{t_i}{\sum_{k=1}^K t_{ki}} \quad (5)$$

其中,  $i \in 1, 2, \dots, I$ ,  $I$ 表示属性种类个数, 例如普遍的二分类以及多分类中的四分类、八分类,  $k \in 1, 2, \dots, K$ ,  $K$ 表示各种分类属性中的属性值个数. 本文计算了数据集中各属性正负样本的数量以及每类属性的正样本合理占比. 所谓正样本合理占比, 即该类属性正样本的合理概率值, 例如“性别”属性为二分类属性, 包含“男性”和“女性”, 正样本合理占比为 50%, 而“年龄”属性为多分类属性, 包含“儿童”“青少年”“成年人”“老年人”4 种选项, 为了计算方便, 将多分类属性分为 4 个二分类属性, 即“是否为儿童”“是否为青少年”“是否为成年人”“是否为老年人”, 因此正样本合理占比为 25%. 具体值如下:

$$L_M = \begin{cases} 0.25 & \text{age} \\ 0.125 & \text{C.up} \\ 0.11 & \text{C.low} \\ 0.5 & \text{else} \end{cases} \quad (6)$$

训练过程中, 提高正负样本不平衡且偏离合理占比属性的权重, 力求此类属性可以进行更多的训练, 以此来达到更好的识别效果.

另外, 传统的属性识别可能会由于光线、遮挡、分辨率低等因素造成某些图像的属性识别率较低, 但是对于基于监控的行人属性数据集来说, 行人图像是从不同监控视频的不同序列片段所截取的不同帧, 每一组序列片段, 都有着行人身份信息相同, 行人属性信息相同, 行人图像不相同的特点. 因此, 利用行人身份信息辅助样本平衡优化可以进一步提高行人属性识别能力. 数据集包含多个行人身份信息, 所以身份标签采用多分类交叉熵损失函数. 公式如下:

$$Loss_I = -\frac{1}{N} W_2 \left[ -y_n + \log \sum_{i=1}^I \exp x_{ni} \right] \quad (7)$$

$$W_2 = \lambda \times \exp(-\lambda \times epoch) \quad (8)$$

$$Loss = Loss_F + Loss_I \quad (9)$$

其中,  $i \in 1, 2, \dots, I$ ,  $I$ 表示训练集行人身份个数,  $x_{ni}$ 表示第 $n$ 张图像第 $i$ 个行人身份的预测值,  $y_n$ 表示第 $n$ 张图像行人身份的真实值,  $\lambda$ 取 0.3.

样本平衡损失与身份损失融合作为模块总损失. 基于身份信息融合的行人属性样本平衡优化方法, 可以弥补跨任务数据集正负样本不平衡的缺陷, 提升模型的识别能力.

## 4 实验

本文在大规模跨任务数据集 Market-1501-attribute 数据集上进行了实验.

### 4.1 数据集

Market-1501-attribute 数据集的图像来自清华大学, 是在传统的行人重识别 Market-1501 数据集<sup>[18]</sup>的基础上标注行人属性信息所得, 以此用于行人属性识别任务. 该数据集的训练集包含 19 732 张图像和 751 个行人身份信息, 测试集包含 13 328 张图像和 750 个行人身份信息. 其中每一张图像, 都注释了 27 个属性, 包含 9 个二分类属性、1 个多分类属性和 2 个组合二分类属性, 最终转换为 30 个二分类属性.

### 4.2 实验细节

本文提出的模型使用 SGD 优化算法, 批度尺寸为 16, 权重衰减系数为  $5E-4$ , 特征提取层的基本学习率为 0.01, 分类层的基本学习率为 0.1, 使用 StepLR 学习率调整策略, 学习率衰减为 0.1, 每 5 个循环衰减一次, 总共训练了 20 个循环.

该模型在 Market-1501-attribute 数据集上进行实验, DeepMAR<sup>[9]</sup> 是基于深度学习的行人属性识别的经典模型, APR<sup>[15]</sup> 是 Lin 等人提出的行人重识别任务中的属性识别框架, UF<sup>[10]</sup>, JCM<sup>[11]</sup> 都是使用行人身份信息辅助行人属性识别的跨任务模型. 表 1 表示了基于类的评估, 表 2 表示了基于实例的评估, 可以发现, 本文提出的模型在最新的 Market-1501-attribute 数据集中取得了较好的性能.

Market-1501-attribute 数据集中, “L.slv”“L.low”“S.clth”“B.pack”“H.bag”“C.up”“C.low”分别表示的是 length of sleeve, length of lower-body clothing, style of clothing, backpack, handbag, color of upper-body

clothing and color of lower-body clothing. 其中, “C.up”和“C.low”为两组单分类属性, “C.up”包含 up black, up white, up red, up purple, up yellow, up gray, up blue, up green 八个单分类属性, 几乎涵盖了日常上衣所有颜色, 合理的属性分布应各为 1/8 约 12.5%, 但是数据集中分布为 up white 占比 30%, up purple 占比 4%, up yellow 占比 6%, up blue 占比 6%, up green 占比 8%, 均属于正负样本不平衡属性, “C.low”包含 down black, down white, down pink, down purple, down yellow, down gray, down blue, down green, down brown 九个单分类属性, 几乎涵盖了日常下衣所有颜色, 合理的属性分布应各为 1/9 约 11.1%, 但是数据集中分布为 down black 占比 38%, down pink 占比 4%, down purple 占比 1%, down yellow 占比 1%, down green 占比 20%, 均属于正负样本不平衡属性. 另外, 普通的单分类属

性中, 合理的属性分布应各为 1/2 约 50%, 但是数据集中分布为 style of clothing 占比 63%, hat 占比 2%, 同样属于正负样本不平衡属性. 通过实验发现, 该论文提出的模型在正负样本不平衡的属性中取得了较好的结果.

本文在 Market-1501-attribute 数据集上进行了基于实例的评估, 其中, *prec* 表示精确率, 意味着预测为正的样本中有多少是真正的正样本, *recall* 表示召回率, 意味着有多少正样本被正确预测. 理想状态下, 本文力求 *prec* 和 *recall* 都尽可能获得较高的值, 但事实上这两者在很多情况下是相互矛盾的, *prec* 升高会导致 *recall* 降低, 反之亦然. 因此本文考虑使用一个中间数  $F1=2 \times prec \times recall / (prec + recall)$  来平衡两个值, 尽可能使 *prec* 和 *recall* 都获得较优解, 如表 2, 可以发现, 本文提出的模型在 *F1* 上仍然保持了较好的结果.

表 1 Market-1501-attribute 数据集基于类评估实验

Model	Gender	Hair	L.slv	L.low	S.clth	Hat	B.pack	Bag	H.bag	Age	C.up	C.low	Avg
DeepMAR <sup>[9]</sup>	85.24	85.48	92.79	91.37	89.37	95.93	84.56	71.57	86.53	<b>94.56</b>	93.78	92.92	88.68
APR <sup>[15]</sup>	85.64	85.62	92.87	92.80	89.91	97.13	78.12	75.41	<b>90.53</b>	93.18	93.18	92.27	88.93
UF <sup>[10]</sup>	88.94	78.26	93.53	92.12	84.79	97.06	85.46	67.28	88.4	84.76	87.5	87.21	86.28
JCM <sup>[11]</sup>	89.7	82.5	<b>93.7</b>	<b>93.3</b>	89.2	97.2	85.2	<b>86.9</b>	86.2	87.4	92.4	93.13	89.73
IEBO	<b>89.87</b>	<b>86.39</b>	93.57	92.72	<b>92.62</b>	<b>97.89</b>	<b>86.15</b>	76.11	87.25	93.56	<b>94.88</b>	<b>94.08</b>	<b>90.43</b>

表 2 Market-1501-attribute 数据集基于实例评估实验

Model	acc	prec	recall	F1
DeepMAR	69.65	82.60	80.24	81.40
APR	70.25	83.52	78.96	81.81
IEBO	<b>73.48</b>	<b>85.50</b>	<b>82.13</b>	<b>83.78</b>

### 4.3 消融实验

IEBO 模型包括图像增强模块 IE、样本平衡优化模块 SBO, 表 3 表示了这些模块分别在 Market-1501-attribute 数据集上进行的消融实验.

表 3 Market-1501-attribute 数据集消融实验

Model	Avg	acc	prec	recall	F1
Baseline	89.93	70.24	86.88	76.84	81.55
Base+IE	90.16	71.22	<b>88.28</b>	77.29	82.42
Base+SBO	90.32	72.72	86.21	80.68	83.35
IEBO	<b>90.43</b>	<b>73.48</b>	85.50	<b>82.13</b>	<b>83.78</b>

baseline 使用最基本的 ResNet50 网络结构作为特征提取方式, SBO 基于身份信息与样本平衡添加了 W1 与 W2 权重参数, 可以发现, 对比 baseline, 每一个模块都使模型的准确率有所提升, 综合所有模块, 可以使模型准确率达到最大值.

### 4.4 参数调优实验

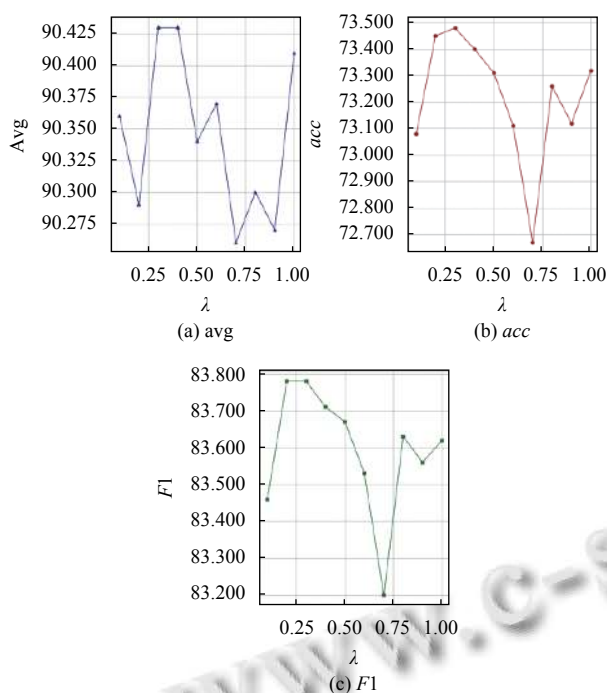
在样本平衡优化中, W2 公式使用  $\lambda$  作为参数进行计算, 我们测试了  $\lambda$  从 0.1 到 1 时模型的运算结果, 如图 6, 可以发现, 当  $\lambda$  取 0.3 时, 模型拥有最优解.

## 5 结论

本文主要介绍了行人属性识别任务的背景知识, 以及新型跨任务数据集的信息, 并且针对新提出的模型讨论了图像增强, 样本平衡等方面的改进. 通过实验比较, 可以发现, 新提出的模型在 Market-1501-attribute 数据集上取得了较好的性能.

行人属性识别任务一直是计算机视觉中热门的研究方向, 从实践应用角度来讲, 行人属性识别在行人检测<sup>[20]</sup>、视觉跟踪<sup>[21]</sup>、行人重识别<sup>[14,22-24]</sup>等领域等都有着重要的作用, 同时这些任务共同在智能监控系统、智能机器人、广告营销、商业零售等领域也拥有着广泛的应用前景. 但是由于基于监控的图像数据集本身获取方式受到限制, 行人属性识别任务的进一步发展仍然是一个巨大的挑战, 模型性能的提高也有待进一步深入研究.



图6  $\lambda$  的参数调优实验图

## 参考文献

- Chen SK, Wang JF, Chen YD, *et al.* Label distribution learning on auxiliary label space graphs for facial expression recognition. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020. 13981–13990.
- Wang K, Peng XJ, Yang JF, *et al.* Suppressing uncertainties for large-scale facial expression recognition. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020. 6896–6905.
- Li RL, Bladin K, Zhao YJ, *et al.* Learning formation of physically-based face attributes. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020. 3407–3416.
- Chen ZY, Li AN, Wang YH. A temporal attentive approach for video-based pedestrian attribute recognition. Proceedings of the 2nd Chinese Conference on Pattern Recognition and Computer Vision (PRCV). Xi'an: Springer, 2019. 209–220.
- Tang CF, Sheng L, Zhang ZX, *et al.* Improving pedestrian attribute recognition with weakly-supervised multi-scale attribute-specific localization. Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul: IEEE, 2019. 4996–5005.
- Tan ZC, Yang Y, Wan J, *et al.* Relation-aware pedestrian attribute recognition with graph convolutional networks. Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York: AAAI, 2020. 12055–12062.
- Fabbi M, Calderara S, Cucchiara R. Generative adversarial models for people attribute recognition in surveillance. 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). Lecce: IEEE, 2017. 1–6.
- Li QZ, Zhao X, He R, *et al.* Visual-semantic graph reasoning for pedestrian attribute recognition. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 8634–8641.
- Li DW, Chen XT, Huang KQ. Multi-attribute learning for pedestrian attribute recognition in surveillance scenarios. 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR). Kuala Lumpur: IEEE, 2015. 111–115.
- Sun CX, Jiang N, Zhang L, *et al.* Unified framework for joint attribute classification and person re-identification. Proceedings of the 27th International Conference on Artificial Neural Networks. Rhodes: Springer, 2018. 637–647.
- Zhao X, Sang LF, Ding GG, *et al.* Recurrent attention model for pedestrian attribute recognition. Proceedings of the 33rd AAAI Conference on Artificial Intelligence. Honolulu: AAAI, 2019. 9275–9282.
- Zeng HT, Ai HZ, Zhuang ZJ, *et al.* Multi-task learning via co-attentive sharing for pedestrian attribute recognition. 2020 IEEE International Conference on Multimedia and Expo (ICME). London: IEEE, 2020. 1–6.
- Deng YB, Luo P, Loy CC, *et al.* Pedestrian attribute recognition at far distance. Proceedings of the 22nd ACM International Conference on Multimedia. Lisboa: ACM, 2014. 789–792.
- Lin YT, Zheng L, Zheng ZD, *et al.* Improving person re-identification by attribute and identity learning. Pattern Recognition, 2019, 95: 151–161. [doi: 10.1016/j.patcog.2019.06.006]
- Zheng L, Shen LY, Tian L, *et al.* Scalable person re-identification: A benchmark. 2015 IEEE International Conference on Computer Vision (ICCV). Santiago: IEEE, 2015. 1116–1124.
- Liu H, Wu J, Jiang J, *et al.* Sequence-based person attribute recognition with joint ctc-attention model. arXiv: 1811.08115, 2018.
- Yang J, Fan JR, Wang YR, *et al.* Hierarchical feature embedding for attribute recognition. Proceedings of the 2020

- IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020. 13052–13061.
- 18 Getreuer P. Automatic color enhancement (ACE) and its fast implementation. *Image Processing on Line*, 2012, 2: 266–277. [doi: [10.5201/ipol.2012.g-ace](https://doi.org/10.5201/ipol.2012.g-ace)]
- 19 Chen LC, Papandreou G, Schroff F, *et al.* Rethinking atrous convolution for semantic image segmentation. arXiv: 1706.05587, 2017.
- 20 Tian YL, Luo P, Wang XG, *et al.* Pedestrian detection aided by deep learning semantic tasks. *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Boston: IEEE, 2015. 5079–5087.
- 21 Danelljan M, Khan FS, Felsberg M, *et al.* Adaptive color attributes for real-time visual tracking. *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus: IEEE, 2014. 1090–1097.
- 22 Su C, Zhang SL, Xing JL, *et al.* Deep attributes driven multi-camera person re-identification. *Proceedings of the 14th European Conference on Computer Vision*. Amsterdam: Springer, 2016. 475–491.
- 23 Layne R, Hospedales TM, Gong SG. Towards person identification and re-identification with attributes. *European Conference on Computer Vision*. Florence: Springer, 2012. 402–412.
- 24 Li AN, Liu LQ, Wang K, *et al.* Clothing attributes assisted person reidentification. *IEEE Transactions on Circuits and Systems for Video Technology*, 2015, 25(5): 869–878. [doi: [10.1109/TCSVT.2014.2352552](https://doi.org/10.1109/TCSVT.2014.2352552)]