

Recognition Technology of Clothing Attribute Label Based on Deep Learning

Xudong Xu, Xin Liu

Beijing University of Technology, Beijing
Email: 951527202@qq.com

Received: Mar. 17th, 2020; accepted: Apr. 1st, 2020; published: Apr. 8th, 2020

Abstract

On the basis of deep exploration and understanding of deep learning field and neural network structure, in view of the problem that the existing methods are mostly based on the whole dress and ignore the characteristics of dress details, in this paper, we design and implement a clothing attribute label recognition method based on deep learning, using the online clothing image database from Ali e-commerce data, the neural network based on deep learning can quickly and efficiently extract the key feature information in clothing image, and modeling and analysis based on extracted data features, so as to achieve accurate identification and classification of clothing attributes. The experimental results show that this method can effectively realize the recognition and classification of clothing attribute labels, improve the classification accuracy effectively, effectively solve the difficult problem of clothing attribute label recognition and classification, and realize the efficient recognition and classification of local attributes of clothing image.

Keywords

Deep Learning, CNN, Inception Network, Clothing Attribute Label Identification

基于深度学习的服饰属性标签识别技术

徐旭东, 刘 鑫

北京工业大学, 北京
Email: 951527202@qq.com

收稿日期: 2020年3月17日; 录用日期: 2020年4月1日; 发布日期: 2020年4月8日

摘 要

在对深度学习领域和神经网络结构深入探究及理解基础之上, 针对现有方法大都基于服饰整体而忽略服

饰细节要素特征的问题, 本文设计并实现一种基于深度学习的服装属性标签识别的方法, 采用来自阿里电商数据的网络服饰图像数据库, 基于深度学习的神经网络快速高效地完成服饰图像中关键特征信息的提取, 并基于提取的数据特征进行建模分析, 从而达到对服饰属性的精准识别与分类。实验结果表明: 该方法能有效地实现服饰属性标签的识别与分类, 分类准确率有效提高, 可有效解决服饰属性标签识别分类困难的问题, 实现服饰图像局部属性的高效识别分类。

关键词

深度学习, 卷积神经网络, Inception网络, 服饰属性标签识别

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

服饰款式的区分主要依据轮廓, 以及不同的内部形状、长度, 但不包括衣服的材质、色彩等等。互联网技术突飞猛进, 图像分辨率不断提高, 数据呈爆发式增长, 由于人工区分效率低、误差大, 从而激发了图像识别技术的快速发展。在服饰领域, 现有方法大都基于服饰整体进行处理识别而忽略细节部位特征[1]; 通常服装图像的常规表示方法, 并不能从整体层面展示出服装本身所具备的所有有效特征; 并且人工标注的低效性, 以及主观上的判断不同, 会产生一些不可避免的误差, 这都会导致用户不能获得精准的检索[2], 使得用户体验不佳, 尤其在服饰检索领域。本文的重点就是对服饰图像特征进行细粒度的提取, 并对其进行快速的高效分类。

使用图像中提取的有效信息组成所需的特征。目前, 基于深度学习的方法在图像特征提取[3]方面正大放光彩。Wang X. C等[4]能够根据服装实物图数据中完成服装款式的识别, 使用模糊理论将款式轮廓点自动分类, 主要是针对服饰局部轮廓进行识别, 不是服饰的整体轮廓。Cheng Ci [5]等提取形状特征, 是使用边缘检测技术及颜色特征, 来展示服装款式信息从而完成分类, 实现图像检索功能。近年来, 出现了很多解析人物着装的研究, 使用不同的颜色分割标记出人物穿着的不同区域, 为提取穿着在人物身上的服装信息扫除了障碍[6] [7]; 同时多特征融合的服装分类识别领域的学术研究比较多。Zhang H.等[8]提出对于具有差异性的服装款式, 可以从中提出不同的特征来完成分类。Bossard 等[9]则充分利用了服饰本身的大体轮廓, 以及直观的视觉效应, 将服饰图像划分为等大小的密集网格, 并进一步提取各网格中的材质、色彩等特征, 最后结合双模型方法来组合成最终的特征表示。

在对深度学习领域和神经网络结构深入探究及理解基础之上, 为解决现今服饰图像领域识别分类精度较低的难题, 本文设计并实现一种基于深度学习的服装属性标签识别方法。采用深度卷积神经网络从数据集中自动提取并学习服饰的局部属性特征, 实现服饰图像分类的分析建模, 完成服饰属性标签的快速高精度分类。此方法可以在众多领域应用, 如服装款式的识别及查找、服装的搭配组合、搭配推荐, 还可以结合人脸识别技术, 应用于人员查找追踪。

2. 理论与方法

2.1. 卷积神经网络

卷积神经网络的使用在图像识别任务中尤为重要, 本文所用的 Inception 结构来源于 GoogLeNet,

GoogLeNet 为第一个版本, 因此定义为 Inception-v1; 在之后的完善中引入 BatchNormalization, 定义为 Inception-v2; 之后加入因子分解思想进一步优化为 Inception-v3; 如今, 诞生了最新的 Inception-v4 网络, Inception-v4 的设计不是用来提高深度, 进而提高准确度的, 而是用来提高速度。

Inception-v4 的网络结合了前三个版本的优点, 并进行进一步的优化, 使其不使用 residual learning 也可以达到与 Inception_Resnet-v2 相似的准确率。在服饰属性识别方面, 选择 Inception_v4 模型, 因为其结构所具备的对特征的多尺度提取的特性在服饰的特征识别上可以更加有利于提取出主要信息, 提高识别效率。研究发现 resnet 网络能够有效防止网络过深导致的梯度消失现象, 提高识别率, 经过验证使用 Inception + resnet 网络结构得到了更高的识别率, 因此也采用 InceptionResNet_v2 网络结构搭建模型。

2.2. 优化器

在优化器的选择上, 相比 SGD 学习率固定不边, Adam 更加科学, 可以自动调优。其主要公式如下。

$$m_t = \mu * m_{t-1} + (1 - \mu) * g_t \quad (1)$$

$$n_t = v * n_{t-1} + (1 - v) * g_t^2 \quad (2)$$

$$\hat{m}_t = \frac{m_t}{1 - \mu^t} \quad (3)$$

$$\hat{n}_t = \frac{n_t}{1 - v^t} \quad (4)$$

$$\Delta\theta_t = -\frac{\hat{m}_t}{\sqrt{\hat{n}_t + \epsilon}} * \eta \quad (5)$$

使用 m_t 定义一阶动量, 表示过去的各个时刻梯度方向上的指数变化平均值, 即当前时刻的梯度下降方向除轻微偏向当前梯度下降的方向外, 很大比重依据过去时刻的下降方向, 极大的增加了梯度下降过程的稳定性。但仍然无法解决限于局部最优的问题。因此二阶动量 n_t 的出现, 才使得学习率实现自适应变化。学习率从 α 转变为 $\alpha/\sqrt{n_t}$ 使得学习率从最初的 α 逐渐减小到 0, 这也就大大降低了人工操作的繁琐性, 但不可避免的面临一个问题, 学习率降为 0 时提前终止训练。结合本实验, 鉴于 Adam 能够更好的快速收敛以及动态调整学习率, 所以首先用 Adam 优化, 但由于 Adam 无法找到全局最优点, 再结合 SGD 进行微调, 找到最优结果。

2.3. 方案

目前服饰属性标签的识别方面的研究有限, 如何较为全面且准确的识别服饰属性特征, 并进行属性分类是重难点。可以采用多种辅助操作, 改进优化神经网络, 增加网络结构的层数, 从而得到在不同网络结构对特征获取的差异性, 进而选择最优网络结构。

图像分类的研究主要集中在特征识别和分类器优化两个方面, 为解决现今服饰图像领域识别分类精度较低的难题, 本文设计并实现一种基于深度学习的服装属性标签识别方法。相比现有的图像识别方法, 本方法要达到快速且精准的目标, 主要从以下几个方面入手:

(1) 对数据集进行初步分类, 完成数据预处理。首先根据数据标签将数据集按属性维度分为长度和设计两个层面, 完成对数据集的初步分类; 然后进行数据预处理操作, 包括数据标准化处理, 图像分辨率的选取及数据增强的使用。

(2) 完成长度层面平铺数据和非平铺数据的分类。将包含人体结构的图像数据分类为非平铺数据, 反之则为平铺数据。设计并实现一种区分平铺数据与非平铺数据的分类策略。因为图像数据包含有平铺数据

和非平铺数据两种, 实验证明, 对两者进行区分之后, 分类精确度得到较大提高。

(3) 多任务识别模型的建立。在长度层面和设计层面分别建立识别模型, 保存模型权重。

(4) 预测模型, 集成训练模型权重, 从而完成对服饰属性标签的精准分类。

3. 实施过程

3.1. 数据处理

3.1.1. 初步分类

本文采用来自阿里电商数据的网络服饰图像数据库, 选取八个属性维度, 每个维度包含若干属性值, 基于数据集所定义的服装特征标签属性, 对属性值以及维度两方面完成对服饰图像(单人平铺或单人模特)的识别分类。依据服饰图像识别的复杂性, 将八个属性维度分为长度层面和设计层面两种方案, 八种属性维度及属性值如表 1 所示。

Table 1. Data set attribute dimensions and attribute values

表 1. 数据集属性维度和属性值

属性维度	属性值
颈线设计	V 领、圆领、深 V 领、方领、不规则领、抹胸领、一字领、露肩领、半开领、桃型领
领子设计	娃娃领、清道夫领、衬衫领、飞行员领、不可见
脖颈设计	荷叶半高领、常规半高领、堆堆领、高常规领、不可见
翻领设计	西装领、一片领、青果领、直线领、不可见
袖长	无袖、杯袖、短袖、五分袖、七分袖、九分袖、长袖、超长袖、不可见
衣长	高腰、正常、长款、加长款、及膝、超长、及地、不可见
裙长	短裙、中裙、七分裙、九分裙、长裙、不可见
裤长	短裤、五分裤、七分裤、九分裤、长裤、不可见

数据集初步分类如图 1 所示。

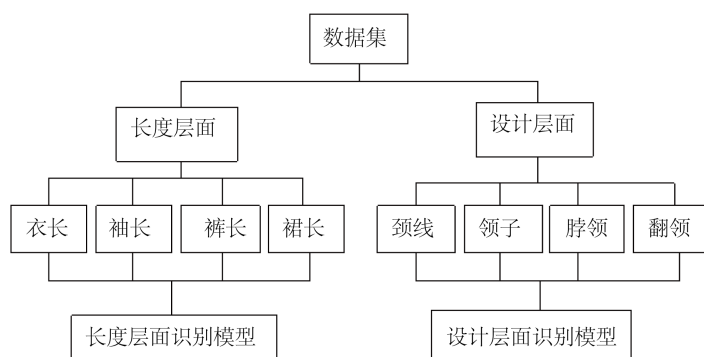


Figure 1. Preliminary classification of data sets

图 1. 数据集初步分类

3.1.2. 数据预处理

(1) 图像标准化

由于服饰图像数据拍摄来源不同, 加上拍摄的角度、光线强弱、以及图像分辨率等不同, 导致数据

样本之间存在较大差异,这就导致识别可能出现较大的误差。所以进行识别之前先做标准化处理,经过图像标准化可得到一致的输入数据,从而加快收敛速度,加速训练过程,提高识别准确率。

使用单独计算的本数据集的 std (标准差)与 mean (平均值),代替直接用 ImageNet 的 std 与 mean 的方式,训练集一共 79573 张图片,计算得到数据集的均值和标准差分别为:

$$\text{mean} = [0.6400365359276026, 0.6001295694316631, 0.5877540247206047],$$

$$\text{std} = [0.08403916220581975, 0.08801511335355927, 0.08918949844575702],$$

使用以下公式进行图像标准化处理:

$$\text{image_standardization} = \frac{X - \text{mean}}{\text{image_std}} \quad (6)$$

$$\text{image_std} = \max\left(\text{std}, \frac{1.0}{\sqrt{N}}\right) \quad (7)$$

其中 X 表示要处理图像的像素矩阵, $\text{image_standardization}$ 表示标准化后的图像像素矩阵, mean 是整体数据集的均值, std 表示整体数据集的标准方差, image_std 每张图片的标准偏差, N 表示图像 X 的像素数量。

(2) 图像分辨率

数据集中图片的分辨率是不一样的,对图像识别来说,图像的大小对识别结果至关重要,更高的分辨率显然可以获取到更多的有效信息,可以提取到更多的有效特征。提高图像分辨,增大图像输入尺寸可提高分类准确率,但是分辨率的增加无疑会使训练过程更加复杂,增加卷积核和显存的压力,因此需选择一个合适的尺寸。如图 2 所示,选用 ImageNet,进行分辨率选取实验,根据实验结果最后选定 size 为 480,采用 RGB 三通道进行实验。

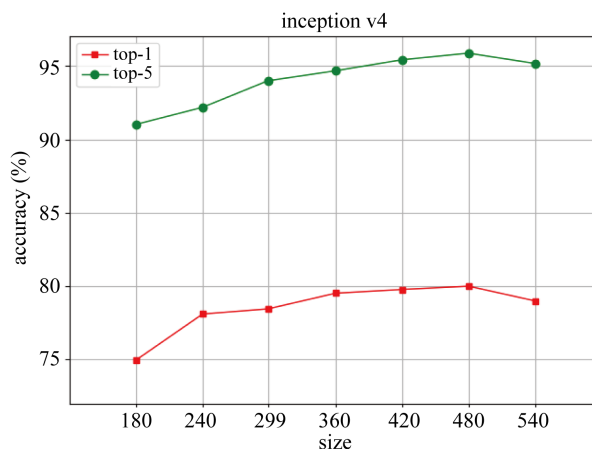


Figure 2. Select the image resolution size

图 2. 选取图像分辨率尺寸

(3) 数据增强

从大量图像数据中了解得到,图像本身并不是规则的,为了从中提取到有效信息,采用 Imgaug 技术,它提供了多种数据增强方式,并且可以方便快捷的使用其提供的接口,并根据自己需要对增强方式进行自定义组合,从而增强不同大小的图片,使得在背景处理中通过此种增强方式来大大提高模型性能。

显然仅执行一种增强作用并不明显, 因此, 通过多种增强组合的方式来得到更加鲁棒的效果。通过实例化单个 `augment_images()` 来分次实现图像增强, 无疑是繁琐的。因此本文使用 `Imgaug` 生成变换序列 (Sequential), 将图像 batch 传入, 并在 Sequential 中加入所需的图像增强方式。本实验选用了 `Affine`, 将放射转应用于图像, 又增加了 `AdditiveGaussianNoise`, 在初期加入了裁剪图像 (Crop), 但经过实验发现, 若是裁剪不当可能会导致图片特征被裁减掉, 影响最终的识别效果, 因此这里不加入裁剪图像的方式。此外由于颜色特征对于本文的识别分类是无意义的, 为了避免颜色对图像的识别分类产生影响, 加入了颜色通道的随意变化。另外, 数据集中存在不同旋转角度的图像, 因此添加水平翻转及随机角度翻转来增强模型的泛化能力, 实验证明模型经过数据增强后有了较大提升。

3.2. 长度层面平铺和非平铺数据分类

Joint Body Parsing & Pose Estimation Network, 简称 JPPNet, 是由中山大学研究[10], 可以很好对人体部位进行识别分类, 人体解析的目的是将人体图像用细粒度语义分割成多个部分, 如身体部位和衣服。

考虑到数据集中的图片会有平铺数据和非平铺数据两种, 研究对比发现, JPPNet 可以很好对人体部位进行识别, 结合本文的研究目标, 发现 JPPNet 可以很好实现平铺数据和非平铺数据的精确分类。对长度层面的图片进行测试并分类。首先使用 JPPNet 进行初步分类, 若图片中含有头、脸、左臂、右臂、左腿、右腿、左脚、右脚其中之一部位, 则判定为非平铺, 反之则为平铺数据, 得到初步的平铺数据和非平铺数据。

由于 JPPNet 框架的局限性, 分析初筛结果发现非平铺数据中仍存在大量的平铺数据, 且不同属性维度的分类结果有差异, 针对差异进行二次筛选, 采用 MaskRCNN 和 JPPNet 做对衣长和袖长并行训练进行联合复筛; 由于 JPPNet 检测出现大量错误, 因此使用 MaskRCNN 对裙长和裤长进行复筛。

对复筛的结果使用优化器进行优化处理, 先用 Adam 优化器进行优化, 学习率设置为 0.001, 以交叉熵作为损失函数, 以准确率作为评估指标, 进行三轮训练准确率达到 0.9024, 然后又将学习率设置为 0.000025 和 0.000075 进行两轮训练, 准确率达到 0.9411, 最后只有 SGD 进行微调, 学习率设置为 0.0000045 和 0.000001 分别进行训练, 最终准确率达到 0.957。完成平铺数据和非平铺数据的分类, 得到最终的平铺数据和非平铺数据。

3.3. 多任务识别模型训练优化

为了将多属性维度的数据集一起输入到网络中进行训练, 引入了多任务训练的设计。整个属性识别网络中的识别任务的卷积层共享, 但是由于其任务的差异性, 并不能使用完全相同的模型来对全部数据进行统一分类, 如果不进行区分会对识别结果产生较大影响, 因此将数据按属性分成设计层面和长度层面两类, 由于服饰所具备的多个属性维度需要提取各自不同的特征才能进行识别, 在设计与长度两个层面的差异性最为巨大, 所以使用分离的多任务训练方法。

3.3.1. 设计层面识别模型

由于平铺数据和非平铺数据对设计层面相关属性的识别分类没有明显影响, 所以在设计层面对包含的多个属性维度构造了一个多任务识别分类模型。

建模过程: 输入设计层面数据集, 使用整体数据集的均值和标准差将批处理的图像进行标准化, 再处理数据使之符合 TensorFlow 格式, 然后导入基网络, 冻结网络层, 添加全局平均池化层和 Dropout 来防止过拟合, 全连接层选择 SoftMax 分类器进行分类, 之后使用 Adam 结合 SGD 分类器完成优化微调, 生成设计层面识别模型, 保存模型权重。

以 Inception-V4 和 InceptionResNet-V2 为基网络建模, 得到两个设计层面识别模型, 保存权重。

3.3.2. 长度层面识别模型

考虑到平铺数据和非平铺数据对长度层面相关属性识别分类有明显影响, 因此在长度层面分别生成平铺识别模型和非平铺识别模型, 再将两个模型进行融合形成长度层面的识别模型。

首先在长度层面训练了平铺识别模型和非平铺识别模型。训练过程为, 首先读入数据, 使用整体数据集的均值和标准差将批处理的图像进行标准化, 经过处理使数据规格符合 Tensorflow 的数据格式, 之后导入基网络, 冻结网络层, 引入全局平均池化和 Dropout, 这两个操作是防止过拟合, 最后的全连接层选择 SoftMax 分类器进行多任务的分类, 最终形成多任务输出模型, 最后使用 Adam 结合 SGD 分类器完成优化微调, 生成识别模型。分别读入平铺数据和非平铺数据, 生成平铺识别模型和非平铺识别模型。

鉴于平铺域和非平铺域的不一致性和复杂性, 在长度层面, 为平铺识别模型和非平铺识别模型设计了一个双通道的服饰属性识别网络进行融合, 最终完成属性标签识别。对于非平铺数据, 可以根据人体相对位置得到更好的特征识别, 而对于平铺数据的识别更复杂, 使用神经网络的自行优化来进行识别容易产生过拟合, 因此模型融合过程为, 抛弃平铺识别模型顶层的全连接层, 之后加入一个 Dropout 层; 抛弃非平铺识别模型顶层的全连接层, 之后引入 Dense(128); 利用 concat 方法将处理后的平铺识别模型和非平铺识别模型进行融合; 最后使用优化器优化, 使用交叉熵作为损失函数进行训练, 生成长度层面识别模型, 得到模型权重。

以 Inception-V4 和 InceptionResNet-V2 为基网络建模, 得到两个长度层面识别模型, 保存权重。

3.4. 预测模型

预测模型搭建流程, 输入图像数据, 导入基网络, 冻结网络层, 开始进行微调, 由于数据量的限制, 虽经过了数据增强, 数据打乱, 正则规范化等一系列预处理操作, 但是经过基网络的处理之后, 必然面临着数据过拟合的情况发生, 因此引入池化, 由于平均池化更适用于网络的深层, 因此这里使用了 mean-pooling。

通常来说, 平均池化与 Dropout 结合使用, 效果会更佳出众。因此又引入了 Dropout 层。经过实验, 发现隐含节点 Dropout 率为 0.5 的时候效果最好, 此时 Dropout 随机生成的网络结构最多。验证后发现, 确实在准确率上有了轻微的提升效果。

载入对于对应基网络训练模型保存的权重, 在最后一层引入全连接网络, 由于每个层面各自包括多个属性维度, 且每个属性维度又分别包含不同数量的属性值。因此在最后的全连接层引入的 SoftMax 多分类器中的类别是需要动态变化的。例如在进行长度层面识别训练时, 在长度层面的列表字典中, 通过索引不同的属性维度, 取到该属性维度对应的属性值数量, 将其传给 SoftMax 进行输出。

最后通过自定义的 process 函数, 得到输入数据原图与水平翻转两种状态下的图像, 输入到预测模型, 得到输入数据的两种状态在该模型的两个预测结果, 两个结果相加取平均得到输入数据在该模型下的最终预测结果。

分别使用 Inception_v4 和 Inception_ResNet_v2 作为基网络建立两个预测模型, 进行上述预测, 最后将两个预测模型的识别结果进行集成, 即将两个模型的预测结果分别乘以相应的权重再相加, 权重和为 1, 这里采用两个预测模型, 权重各为 0.5, 得到最终的属性标签识别结果。

4. 实验结果及分析

4.1. 长度层面双通道网络的应用

采用本文的平铺和非平铺分类策略, 将长度层面数据集分成平铺数据集和非平铺数据集之后, 在长度层面的识别上, 首先建立了长度平铺识别模型和长度非平铺识别模型, 然后又设计了双通道网络将两

个模型进行融合, 形成最终的长度层面识别模型; 若搭建一个训练模型, 直接输入长度层面数据集进行训练, 得到的是单通道识别模型。如图 3 所示: 实验结果表明使用本文的双通道网络比使用单通道网络, 长度层面各属性维度识别的准确率(acc)有明显的提升。

$$\text{acc} = \frac{\text{预测正确的属性值数量}}{\text{预测属性值总量}} \quad (8)$$

4.2. 各属性维度 AP

多标签图像分类任务中图片的标签不止一个, 因此评价不能用普通单标签图像分类的标准, 本文采用的是平均精确率均值—mAP (mean Average Precision), P 表示查准率, R 表示查全率, AP 表示平均精确率, AP 能够更加有效的降低样本不均衡问题带来的影响, 因此使用 AP 来衡量给定维度的好坏, 而使用 mAP 来衡量整个模型在所有维度的好坏。

本实验八个属性维度的 AP 如图 4 所示, 大部分维度的 AP 在 95% 以上, 其中翻领和脖领的 AP 低于

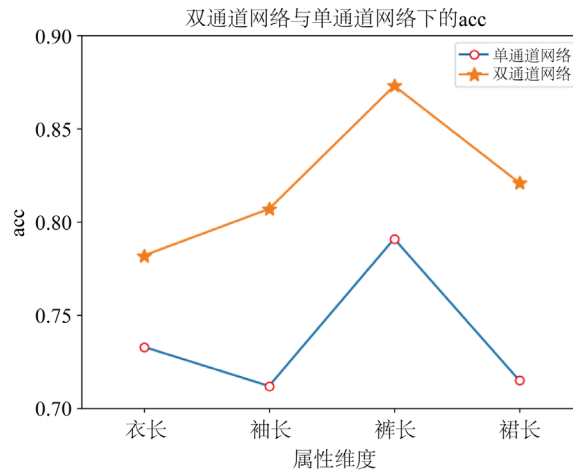


Figure 3. The accuracy of each attribute dimension recognition at the Length level

图 3. 长度层面各属性维度识别的准确率

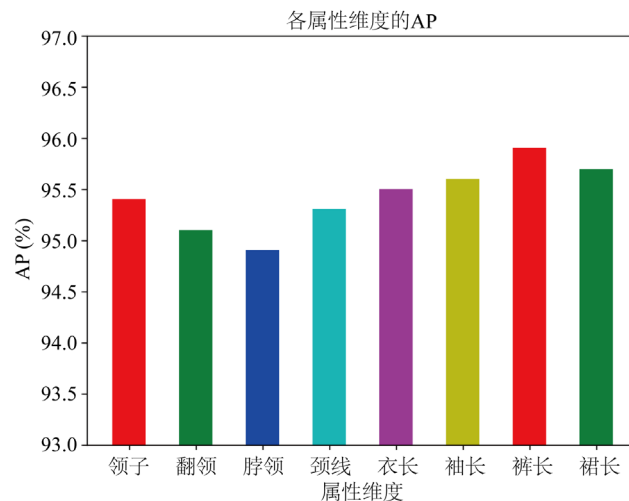


Figure 4. The AP of each attribute dimension

图 4. 各属性维度的 AP

其他维度, 是因为其识别部位往往和其他维度的识别部位产生重叠, 识别难度较大, 后续可以继续研究加入其他方法来进行提高。总体实验结果表明, 使用本文方法进行服饰属性标签识别能够取得了不错的识别效果。

4.3. 模型评估

本实验在模型建立过程中, 对于基网络的选取尤为看中, 使用多个基网络进行了模型的搭建, 大多数情况下, 相比于仅仅使用一个模型, 多个模型结合在一起往往能够提高对数据集拟合的表现, 获得更加鲁棒的效果。图 5 展示了使用不同基网络搭建模型的 acc 和 mAP。在训练阶段, 本实验选取了 Inception_v4 和 Inception_ResNet_v2 分别搭建了训练模型; 在预测阶段, 使用这两个基网络分别建立预测模型来集成训练模型权重, 再将两个预测模型进行融合得到 doubleNet, doubleNet 得到的结果为最终预测结果。实验结果表明, 本文通过多模型融合使得最终识别效果得到了有效提升。

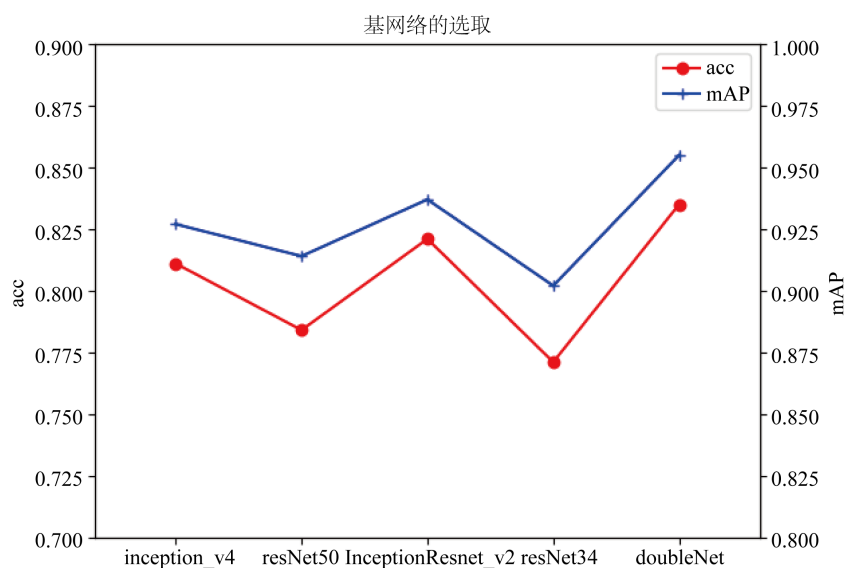


Figure 5. Selection of model base network

图 5. 模型基网络的选取

5. 结论

本实验的识别任务是细粒度的分类任务, 大规模或相关区域的输入对于设计层面的识别任务尤为重要。能够预先对平铺和非平铺数据进行分类, 对于长度层面的识别任务尤为重要。本文设计了两个多任务识别网络来完成设计层面和长度层面的识别任务, 提出了一种区分平铺和非平铺图像数据的分类策略, 使用了联合训练机制, 结合多尺度、多频域、多域训练方法, 对模型进行集成与优化, 最终高效完成服饰属性标签的识别。该方法可以在准确性和效率之间取得很大的平衡, 并且可以应用到大部分的 CV 分类任务中, 实用性很强, 可扩展性较强。

参考文献

- [1] 辛晨. 基于深度学习的图像分类及应用研究[D]: [硕士学位论文]. 北京: 中国科学院大学, 2017.
- [2] 卢兴敬. 基于内容的服装图像检索研究及实现[D]: [硕士学位论文]. 哈尔滨: 哈尔滨工业大学, 2008.
- [3] 青平. 基于深度学习的服装图像分类与检索[D]: [硕士学位论文]. 杭州: 浙江大学, 2017.

- [4] Wang, X.C. and Li, K.J. (2008) Pattern Recognition Based on Fuzzy Cluster for Recognizing Garment Style in the Photo. *Proceedings of 2008 IEEE 9th International Conference on Computer-Aided Industrial Design & Conceptual Design*, Vol. 1, 2008.
- [5] Cheng, C.I. and Liu, D.S.M. (2008) An Intelligent Clothes Search System Based on Fashion Styles. *Proceedings of the 7th International Conference on Machine Learning and Cybernetics*, Kunming, 12-15 July 2008, 1592-1597. <https://doi.org/10.1109/ICMLC.2008.4620660>
- [6] Simo-Serra, E., Fidler, S., Moreno-Noguer, F., *et al.* (2015) A High Performance CRF Model for Clothing Parsing. *Computer Vision ACCV 2014*, Springer International Publishing, 64-81. https://doi.org/10.1007/978-3-319-16811-1_5
- [7] Yang, W., Luo, P. and Lin, L. (2014) Clothing Co-Parsing by Joint Image Segmentation and Labeling. 2014 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Columbus, OH, 23-28 June 2014, 3182-3189. <https://doi.org/10.1109/CVPR.2014.407>
- [8] Zhang, H. and Jiang, Y. (2011) Research of Style-Based Clothing Image Retrieval. 2011 *International Conference on E-Business and E Government (ICEE)*, Shanghai, 6-8 May 2011, 14. <https://doi.org/10.1109/ICEBEG.2011.5882416>
- [9] Bossard, L., Dantone, M., Leistner, C., *et al.* (2012) Apparel Classification with Style. *Asian Conference on Computer Vision*, Springer-Verlag, 321-335. https://doi.org/10.1007/978-3-642-37447-0_25
- [10] Liang, X.D., Gong, K., Shen, X.H. and Lin, L. (2018) Look into Person: Joint Body Parsing Pose Estimation Network and a New Benchmark. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **41**, 871-885. <https://doi.org/10.1109/TPAMI.2018.2820063>