



服装图像分类技术综述

陈巧红, 陈 翊, 孙 麒, 贾宇波

(浙江理工大学信息学院, 杭州 310018)

摘 要: 服装图像分类技术的发展对电子商务行业影响巨大, 准确的服装图像分类不仅可以完成大量自动标注任务, 减少人力成本, 而且有助于提高服装检索的效果。文章通过对现有国内外文献的领域研究进行梳理, 概述了服装图像分类方法的技术框架, 分析了基于传统图像内容和基于深度学习的服装分类方法之间的差异、优缺点以及适用情况, 探讨了当前研究存在的主要问题和未来可能的发展方向, 为该领域的进一步研究提供参考。综述结果显示: 不同类型的分类任务应选择恰当的方法; 对于分类类别数量少且区分度高的分类任务, 利用传统图像内容方法可以取得较高准确率; 对于分类类别数量多且区分度低的分类任务, 采用深度学习更为合适。基于深度学习的分类方法更具广泛性、有效性, 是未来发展的主要方向。

关键词: 服装图像分类; 服装属性分类; 服装识别; 深度学习; 机器学习

中图分类号: TP181

文献标志码: A

文章编号: 1673-3851(2019)09-0631-13

Overview of clothing image classification technology

CHEN Qiaohong, CHEN Yi, SUN Qi, JIA Yubo

(School of Information Science and Technology, Zhejiang Sci-Tech University,
Hangzhou 310018, China)

Abstract: The development of clothing image classification technology has great influence on e-commerce. The precise clothing image classification technology not only can complete quantities of automatic labeling tasks and reduce manpower cost, but also contributes to improving the efficiency of clothing retrieval. Through collating domestic and foreign researches, this paper gives an overview of technical frameworks clothing classification methods and analyses the differences, strengths, weaknesses, application scenarios of clothing classification methods based on traditional image content and deep learning. We also discuss the main problems and the possible development direction in future, which provides some references for the researches of clothing classification. Meanwhile, the results show that appropriate methods should be chosen for different types of classification tasks. For the classification tasks with a small number of classification types and high discrimination, traditional image content method can achieve higher accuracy, while deep learning is more suitable for the classification tasks with a large number of classification types and low discrimination. In summary, the classification method based on deep learning is of more validity and universality, and it is the main development direction in the future.

Key words: clothing image classification; clothing attribute classification; clothing identification; deep learning; machine learning

收稿日期: 2019-01-26 网络出版日期: 2019-06-04

基金项目: 国家自然科学基金项目(51775513)

作者简介: 陈巧红(1978—), 女, 浙江临海人, 副教授, 博士, 主要从事计算机辅助设计及机器学习技术方面的研究。

0 引言

近年来,电子商务发展迅速,网络购物已经成为人们日常生活中的一种重要购物方式,对社会经济发展和个人生活均产生了巨大的影响。服装是电子商务销售品类中的第一大类目,在电商网购市场中的占比约 20.6%^[1],2016 年 12 月中国电子商务研究中心发布的《2015—2016 年度中国服装电商行业报告》显示,服装网购交易额逐年增长,2013—2015 年交易额分别为 4349、6153、7457 亿元^[2]。由此可以看出,服装是一个非常有潜力的电商行业,且仍处于上升阶段。检索是网购商品的第一步,快速高效地检索对服装电商行业具有重要作用,当前电商检索服装主要是通过搜索框中输入与服装相关的关键字来得到商品的链接,该方法需要事先对服装图像进行人工分类,并记录与之相关的文字信息,当用户检索的时候,通过关键字确定服装类别,然后返回同类服装作为检索结果。人工服装图像分类是服装检索的基础性工作,有助于缩小检索范围,提高检索速度和效果,但其存在比较明显的缺点:a)大量的服装需要人工进行分类,费时费力;b)由于人工分类主要凭个人经验进行分类,因此分类结果存在主观性,不可避免地存在误差^[3];c)服装时尚行业发展较快,图像分类人员对新出现的服装种类认知有限,不能快速准确地分辨。

为了解决人工服装图像分类存在的问题,国内外学者研究图像处理技术来代替人工分类,进行了大量研究,并且取得了一定的研究成果。当前服装图像分类的方法可以分为两类:a)基于传统图像内

容的服装图像分类方法^[4];b)基于深度学习的服装图像分类方法^[5]。前者的发展时间较早,服装的图像特征主要由人工设计,存在一定的局限性;后者的发展时间较晚,服装的图像特征由神经网络自动学习获得,是目前主要的发展方向。服装图像分类除了应用在服装检索^[6]中,还被广泛应用于其他领域,如服装搭配推荐系统^[7]、服装自动分拣^[8]、人员识别^[9-11]以及性别分类^[12-14]等。

服装图像分类是当前的研究热点,但相关综述文献较少。本文首先梳理了近几年来服装图像分类方法的相关文献,概述了服装图像分类的基本流程和技术框架;然后分析了基于传统图像内容和基于深度学习的服装图像分类方法的技术组成、优缺点以及适用场景;最后探讨未来可能的发展方向及其面临的挑战。

1 服装图像分类概述

服装图像分类实质是利用图像特征和分类模型来确定服装的类别,按目标任务可以分为服装类别分类(例如款式分类、风格分类、情感分类等)和服装属性分类(例如领型分类、颜色分类等)两种。分类基本流程如图 1 所示,首先对待分类的服装图像进行预处理,然后利用传统图像处理或深度学习提取图像特征,获得特征向量,最后将特征向量输入到分类器进行分类,输出服装的类别。待分类的服装图像包含两种:一种是线上服装图像,背景一般为纯色且服装占据图像的绝大部分区域;另一种是人体服装图像,例如买家秀照片,背景复杂、噪声多且服装占图像区域相对较小。

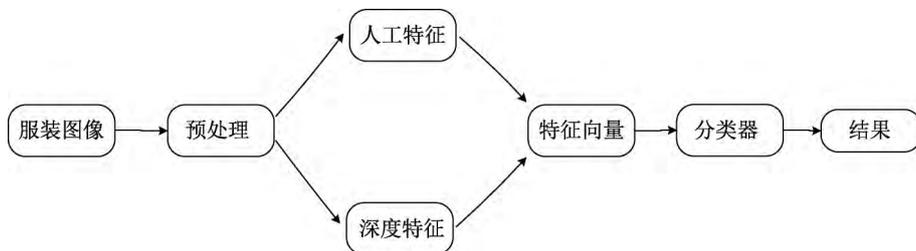


图 1 服装图像分类基本流程

本文对服装图像分类的技术框架进行了总结,技术框架如图 2 所示。传统图像内容方法的图像特征包括全局特征和局部特征,常用分类模型分有支持向量机模型(Support vector machine, SVM)、极限学习机模型(Extreme learning machine, ELM)、随机森林模型(Random forest, RF)等。深度学习方法的图像特征根据提取深度不同分为浅层特征、中层特征以及深层特

征,学习方式分为有监督、半监督和无监督三种,常用的深度网络模型主要有卷积神经网络(Convolutional neural networks, CNN)^[15]、深度信念网络(Deep belief network, DBN)^[16]以及循环神经网络(Recurrent neural network, RNN)^[17]。与传统图像内容方法不同,深度学习方法是一种端对端网络,一般图像特征的提取与分类模型为一个整体。

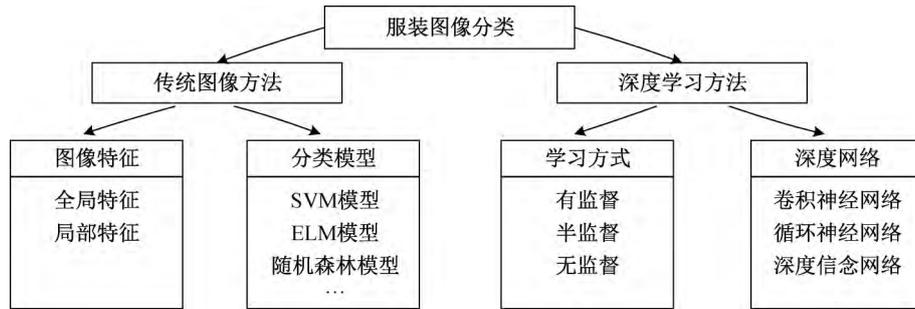


图 2 服装图像分类技术框架

服装图像分类方法的常用评估指标有准确率 (Accuracy)^[4]、Top3 准确率 (Top3 accuracy, Top3)^[18]、精准度 (Precision)^[19]、召回率 (Recall)^[20] 和均值平均精确度 (mean average precision, mAP)^[21] 等。

2 基于传统图像内容的服装分类方法

在早期服装图像分类研究中,如何提取图像特征、提取什么样的特征、提取一种或多种特征均由人工设定。提取方法主要使用数字图像处理和模式识别技术。服装图像特征包括全局和局部两类特征。

2.1 基于全局特征的服装分类

2.1.1 基于形状特征的服装分类

形状特征反映服装的整体结构信息,主要用于识别服装的款式,服装分类领域常用的形状特征提取方法有傅里叶描述子、几何不变距等。傅里叶描述子是运用极其广泛的形状特征之一,其原理是物体外部形状一般为一条闭合曲线,曲线上坐标变化为一个周期函数可用傅里叶级数展开,有计算较方便、速度快、轮廓描述能力强的特点。李东等^[22]提出了基于傅里叶描述子的服装款式分类方法,这种方法优势在于速度快,但是要求更准确的边缘检测算子;An 等^[23]提出了小波傅里叶描述子的服装款式分类方法,小波傅里叶描述子是傅里叶描述子的扩展,优点是支持多分辨率检测,但图像特征之间的匹配时间较长,为了提高匹配速度,该文献进行了两个方面的改进:使用了线性判断分析 (Linear discriminant analysis, LDA)对特征向量进行降维,选择更快的分类器—ELM 模型。这类仅基于形状特征的分类方法显然会受到形状特征的诸多限制,可适用分类任务范围狭小,对服装的功能性分类(即帽子、上衣、裤子以及鞋子等)效果好,而对服装外形区分不明显的分类任务效果差;同时,当衣服出现折叠扭曲等变形情况时,分类准确度难以保证。

2.1.2 基于颜色特征的服装分类

颜色特征是一种基于像素点描述整个图像色彩信息的特征,通过统计图像中的颜色值提取,是应用

极为广泛的视觉特征,服装分类领域常用的颜色特征提取方法有颜色直方图、颜色距^[24]等。颜色直方图是一种常见的颜色特征,它统计全局图像像素点的色彩值来表达颜色的分布情况,即表示不同颜色在图像中比例,优点是对图像平移、旋转具有鲁棒性。Gao 等^[25]提出了基于颜色特征的服装颜色分类方法,目标任务是识别服装颜色,显然只需提取颜色特征即可,而提取一种颜色空间不足以抵抗光线变换,多种颜色空间直方图拥有更好地识别能力,该方法融合了 4 种颜色空间 (RGB 颜色空间、HSV 颜色空间等)的直方图进行分类,研究结果显示,该方法对服装颜色分类有一定程度的提升。由于颜色特征是统计的方式,无法描述服装的颜色分布情况,且对尺度变化敏感,所以该方法一般应用于固定尺度以及色彩跨度大的服装分类任务。

2.1.3 基于纹理特征的服装分类

纹理特征是一种反映服装图像表面特性的特征,通过像素和空间邻域的灰度分布情况来描述图像纹理,具有旋转不变性和良好的抗噪性能。服装图像在某些情况下会出现遮挡、变形等,给仅基于形状特征识别的方法带来一定挑战,而纹理特征对于这种干扰具有较强健壮性。服装分类领域常用的纹理特征提取方法有灰度共生矩阵、局部方向模式 (Local derivative pattern, LDP) 以及局部二值模式 (Local binary patterns, LBP)^[26]等。Thewsuan 等^[27]提出基于两种纹理特征 (LBP 和 Gabor 变换)的服装款式分类,与颜色直方图的统计方式不同,LBP 计算的是每个灰度像素点与周围灰度像素点的关系,通过固定大小邻域内产生的二进制数来表示纹理信息,使其对光线变化具有鲁棒性。纹理特征通过统计服装图像局部重复出现的方式来描述图像,具有良好的旋转不变性和抗噪能力。采用这种方法,可能难以区分局部纹理信息不规则的服装,且当图像尺度不一带来的分辨率改变,有可能影响准确度,纹理特征对侧重于面料方向的分类任务表现较好。

2.1.4 基于多种全局特征的服装分类

以上方法均只提取单一类型全局图像特征进行分类,这类方法存在太多限制难以保证准确率,表现在两个方面,一方面要求服装类别间的区分点突出,如形状、颜色亦或纹理差异性大;另一方面,对服装图像实验数据有较高要求—分辨率统一,光照充足等,为了适应复杂环境下的分类任务,同时扩大分类任务适用范围,需要融合多种全局特征的方法。Yamaguchi等^[20]提出一种基于颜色直方图和 Gabor 小波的服装款式分类;Kalantidis等^[28]提出一种基于 LBP 和颜色特征的服装款式分类;Zheng等^[29]同样提出基于颜色直方图和 Gabor 小波的视频行人服装分类,该文献以行人服装款式分类为目标任务,行人场景极为复杂,服装所在的环境不定,服装款式多种多样,款式的区分不能仅凭一种特征详尽描述,服装款式区分的要素从三种图像特征上考虑,利用图像特征进行一一对应表达:形状特征—结构、颜色特

征—元素、纹理特征—质地;Zhang等^[30]提出基于灰度共生矩阵和颜色特征的服装情感分类理论,服装情感信息在质地与色彩上均有所体现,例如黑色皮质服装更多表达人一种沉稳冷静的感情色彩,纹理特征和颜色特征的组合方式有助于提高表达服装情感能力;此外,Hidayati等^[19]利用多种全局图像特征对含有细粒度的服装款式分类取得了超越深度学习方法的效果,其框架如图3所示,该方法首先利用人体姿势估计得到人脸、躯干、手臂和腿部等人体躯干部分,然后提取颜色、纹理等特征识别局部风格元素,包括领型、图案、袖长等,最后融合各个局部特征分类 Formal shirt、Henley shirt、Informal shirt、Long-sleeved T-shirt、Polo shirt、Spaghetti shirt、Tank top、T-shirt 八个上装细粒度类别,同样的下装也包括八个类别,最后分类精准度为 88.76% 超越了深度学习特征的 62.58%,但该方法的各种局部特征需要人工设计提取,费时费力,融合后的特征维度可能较高。

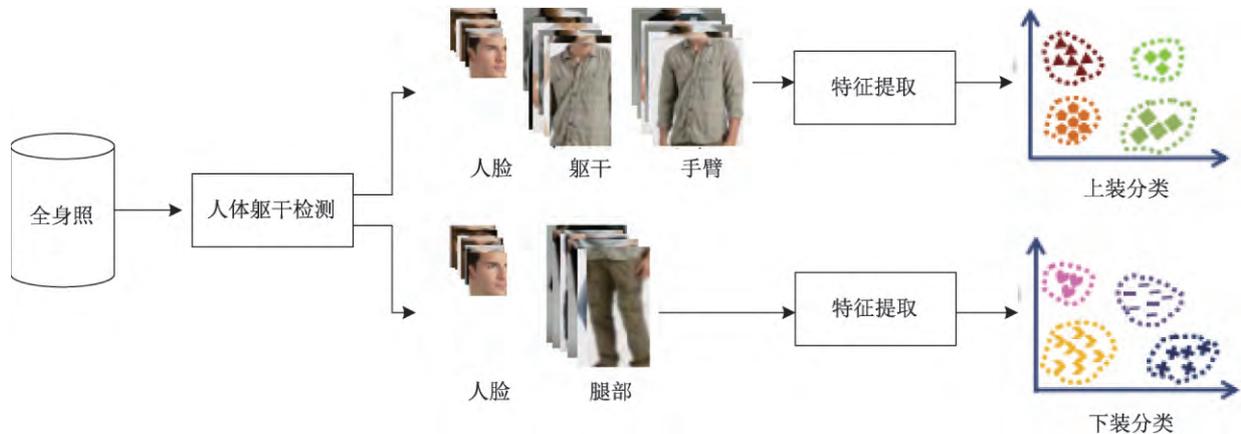


图3 服装款式分类方法^[19]

2.2 基于局部特征服装分类

全局特征主要描述的是服装整体,所以对服装整体信息较敏感而对局部信息不敏感,然而在某些情景下从图像中获取的信息有限,虽然不一定完整,但易提取的服装局部信息较多,不易受到形变、光照等因素影响,仅提取一定量的局部特征就可以完成分类。稳定的局部特征对服装图像的光照、形变、拍摄角度、镜头缩放等干扰有较强的鲁棒性,因此得到了广泛应用。常用的局部特征尺度不变特征变换 (Scale-invariant feature transform, SIFT)^[31]、加速鲁棒特征 (Speeded up robust features, SURF) 和方向梯度直方图 (Histogram of oriented gradient, HOG) 等。

SIFT 特征也称为尺度不变特征,具有尺度不变性,且对旋转、拍摄角度均具有良好鲁棒性,因而

成为使用最广泛的局部特征之一。但这种特征计算复杂,不依靠硬件加速难以达到实时性,在实时性要求高的服装分类任务上不适用。Di等^[32]提出一种基于 SIFT 特征的服装分类方法,该方法利用 SIFT 特征构建词包特征,对多种服装属性进行分类,实验结果显示,与使用全局特征相比,准确度取得了提升;Huo等^[33]在针对民族服装分类任务上融入 SIFT 特征,准确性也得到了提高;Chen等^[34]采用的特征中融入了 SIFT 特征,取得较好的分类效果。

当 SIFT 检测的特征点较多时,整个过程中的计算量变大、速度变慢,针对这个问题,借鉴 SIFT 特征的设计思路, SURF 特征继而提出。SURF 特征运用积分图和海森矩阵行列式的特征点检测方法提高计算速度。Surakarin等^[35]提出一种基于

SURF 特征和纹理特征的服装款式分类方法,采用的纹理特征对光线具有鲁棒性,但其他方面表现一般,该方法加入 SURF 特征,以增强特征的抗干扰能力。另外其他文献^[36-37]采用了类似的方法。

HOG 特征是常见的局部特征,与前两种局部特征不同,它对光线变化具有较强鲁棒性,但不具备旋转、尺度不变性。HOG 特征通过计算图像中成块像素梯度方向直方图提取特征,可以较好地描述物体外形,找出服装在图像中的位置,基于该特征的分类方法更偏向形状方向的服装分类任务,对于那种有层次感、存在遮挡、噪声多的场景表现可能不理想,如识别杂物间中的服装。Zhang 等^[38]提出了一种基于 HOG 的服装款式分类方法,该研究中服装图像均来自天猫买家秀类的正面照,场景复杂度较低,分类准确率较高;Li 等^[39]提取服装图像的 HOG 特征,构建字典学习进行服装属性分类,因为在稀疏矩阵和字典重建过程中丢失的信息主要是噪声,所

以基于稀疏字典学习的方法有一定的去噪能力;Lorenzo-Navarro 等^[40]提出了一种服装属性分类方法,结果表明 HOG 特征准确率高于其他全局特征;类似的研究见文献^[41-42]。

2.3 小结

上文介绍了基于传统图像内容的服装分类方法,各种人工图像特征各有所长,在同种分类任务中表现不尽相同。传统图像特征的优缺点分析结果见表 1。但随着服装行业的发展,基于传统图像内容的服装分类方法的缺点渐渐凸显。首先,随着服装类别不断增多,分类任务会变得多种多样,面对复杂的服装分类任务传统特征难以详尽描述;其次,图像数量也不断增多,当分类大量服装图像时,预处理、描述、匹配等过程耗费太多时间。总而言之,基于传统图像特征分类准确度严重依赖特征的选择,特定的分类任务应视情况而定的选择特征、组合特征,才能获得较高的分类准确率。

表 1 传统图像特征分类方法的对比与分析

特征类别	图像特征	文献	分类任务	数据集	场景复杂度	评价指标	优点	缺点	适用场景
全局特征	单一特征	文献[22]	款式分类	线上图像	低	Accuracy95%	特征提取方便、直观	仅从一个角度对服装描述,适用范围小	图像完整,特征区分度高
		文献[23]	设计图款式分类	线上图像	低	Accuracy100%			
		文献[25]	颜色分类	VIPeR ^[43]	高	Accuracy96%			
		文献[27]	款式分类	摄像机拍摄	中	Accuracy78%			
	多种全局特征	文献[20]	款式分类	Fashionista ^[20]	中	Recall69%	可从多个角度对服装进行描述	多种特征提取难度大,特征维度高	图像完整,类别较多
局部特征	尺度不变特征	文献[28]	款式分类	Fashionista	中	—	对尺度、光线和视角变化不敏感	特征点过多时,匹配时间长	存在遮挡、视角不定等复杂环境下的识别
		文献[29]	款式分类	Fashionista	高	—			
		文献[19]	款式分类	线上图像	中	Precision88%			
	加速鲁棒特征	文献[32]	属性分类	线上图像	低	Accuracy40%~86%	具有 SIFT 的优点且计算速度快	图像尺度和旋转变鲁棒性比 SIFT 强	复杂环境下识别,实时分类
		文献[35]	款式分类	线上图像	中	Precision64%			
		文献[36]	款式分类	线上图像	中	Precision73%			
		文献[37]	款式分类	ACS ^[37]	中	Accuracy41%			
方向梯度直方图	文献[33]	民族服装分类	线上图像	高	Accuracy78%	对光线变化具有鲁棒性,很好捕捉外形特征	难处理遮挡、背景复杂情况;需要统一尺度	侧重于服装外形的分类任务	
	文献[34]	属性分类	线上图像	中	Accuracy55%~92%				
	文献[38]	款式分类	线上图像	中	—				
		文献[39]	属性分类	线上图像	高	Precision54%			
		文献[40]	属性分类	线上图像	中	Accuracy78%~86%			

3 基于深度学习的服装分类方法

随着近年来深度学习的兴起,深度学习在各个领域都取得了突破性进展,完成了许多以往解决不了的任务。深度学习也开始在服装图像分类领域崭露头角,出现了大量基于深度学习的服装分类方法。

3.1 基于卷积神经网络的服装分类

深度学习中的卷积神经网络^[15]适合图像相关的任务,如图像分类、图像分割及图像检索等任务,

与传统方法不同,深度卷积网络的特征提取和模型训练为一体,基本网络结构由输入层、卷积层、池化层和全连接层组成(图 4),整体可以看成是一个非线性函数的组合,不同层的函数有着不同的作用。它的优点在于局部区域感知和权值共享,这样的方式可以减少神经网络训练需要的参数数量,加速训练。训练后的图像特征分为浅层特征、中层特征和深层特征,常见的深度卷积模型有 AlexNet^[44]、VGG-NET^[45]、Google-Net^[46]以及 ResNet^[47]等。

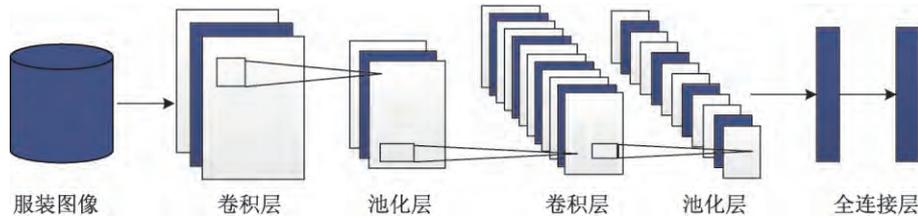


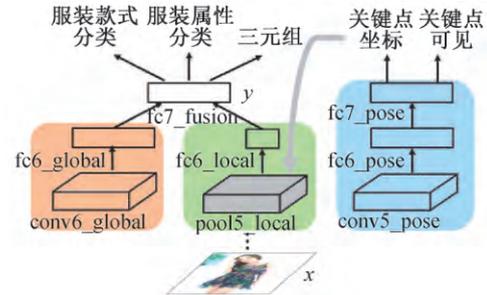
图4 卷积神经网络基本结构

Nawaz 等^[48]提出一种基于 Inception 模型和优化算法的民族服装分类方法,该方法设计了 CNN 网络结构,加入了 Inception 模块,以减少计算量提高特征表达能力,同时还采用梯度优化均方根反向传播(Root mean square prop, RMSProp)方法,实验准确度为 89.22%。Cychnerski 等^[49]对比了 SqueezeNet^[50]和 ResNet 两种网络的服装图像分类结果,结果表明小型 CNN 网络准确度不如一般网络;同时对服装跨域问题进行了改进,提高模型泛化能力,准确度在 70%~81%之间。Dong 等^[51]加入空间金字塔池化(Spatial pyramid pooling, SPP)以解决 CNN 需要固定图像输入尺度的问题。由于生活中的服装图像尺寸不尽相同,而通常 CNN 最后一层的卷积与全连接层之间的输入输出固定,导致需要对输入图像进行缩放或裁减才能作为网络输入,最后准确度为 79.5%。Li 等^[52]提出了一种改进的卷积神经网络方法,以解决数据不平衡问题。Ivanov 等^[53]针对特定服装分类任务,采用经过调整的 CNN 模型参数,获得了优于 SIFT 特征和 SURF 特征的结果,准确度为 84%。Chen 等^[54]采用不同的 CNN 网络结构进行服装分类实验,均取得良好的结果,最高准确度为 92.02%。类似的研究见文献^[55-58]。Liu 等^[18]提出一种基于 VGG-16 和关键点池化的 FashionNet 的服装款式与属性分类方法,该研究收集了 80 万张服装图像建立了数据集 DeepFashion。FashionNet 框架如图 5 所示,fc7_fusion 特征向量有两个输入管道,分别是来自 fc6_global 的全局特征和 fc6_local 的局部池化特征,另外, pool5_local 特征图通过对每个预测的关键点坐标周围信息进行最大池化来获得,关键点坐标回归和可见与否的预测由 fc7_pose 输出。

FashionNet 的损失计算包含有服装款式分类、服装属性分类以及关键点坐标回归等损失,其中关键点坐标损失 L_{land} 可用下式计算:

$$L_{land} = \sum_j^D \|v_j \times (\bar{\varphi}_j - \varphi_j)\|_2^2,$$

其中: D 为训练样本数目; v_j 为关键点可见值; $\bar{\varphi}_j$

图5 服装款式与属性分类方法^[18]

为关键点真实坐标值; φ_j 为关键点预测坐标值。

由上述文献的准确率可看出,CNN 网络在服装分类领域的表现非常好,并且完成了对于传统图像内容方法需要大量工作量的任务,主要体现在:CNN 可以直接将图像信息输入网络获得结果,省去了基于传统图像内容的预处理和特征选择过程。但是一个具有良好分类能力的 CNN 模型,需要大量准确的服装标注信息来训练,数据太少或网络深度不够会导致过拟合和欠拟合问题。另外,由于池化层的特点,CNN 对服装空间位置变化不敏感,例如服装图像中某个图案或者拉链的位置变化对输出结果不会造成太大影响,但这种位置变化可能导致服装风格完全不一样,因此对侧重于空间位置变化的服装分类任务的效果可能不理想。

3.2 基于循环神经网络的服装分类

循环神经网络属于记忆型深度学习模型,在处理与时序、上下关系紧密任务上有突出表现,例如语音识别、文字情感分析、机器翻译等领域应用广泛。CNN 在处理时序模式的任务时表现一般,因为输入输出相对独立,RNN 不仅会输入本次数据,还会加入之前记忆数据一起输出结果。Wang 等^[59]提出一种基于双向卷积循环神经网络的服装关键点检测与款式属性分类方法,该方法首先使用 VGG-16 得到 conv4 layer 的卷积特征图,再借助循环神经网络的特点和服装关键点的对称关系,对关键点位置的进行修正,修正过程如图 6^[59]所示。将修正后的关键点信息做平均池化处理得到注意力特征图与卷积特征图做乘积进行更新:

$$G_c^L = A^L \times F_c,$$

其中： G_c^L 为修正后的特征图； A^L 为 28×28 的注意力特征图； F_c 为原 conv4 layer 卷积特征图； c 为

512 个特征图的数量。采用这样的方式构建了一种注意力机制，从而提高分类准确度。该方法利用关键点之间的连接关系需要较多监督信息和计算量。

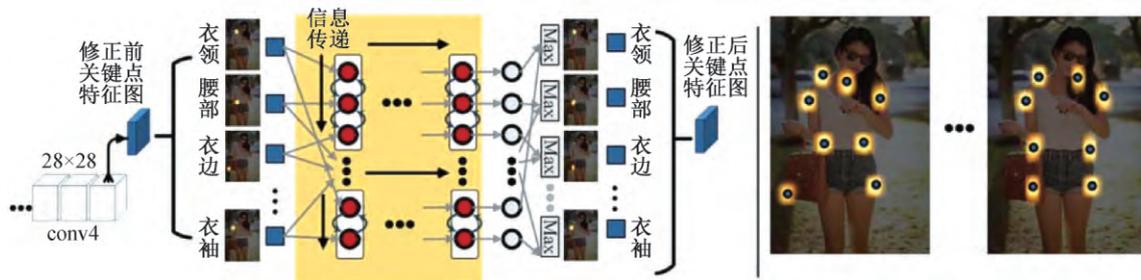


图 6 关键点修正过程^[59]

3.3 基于深度信念网络的服装分类

深度信念网络是一种概率生成模型，深度网络模型的初始参数选择对模型的训练存在一定的影响，DBN 可以为网络提供一个更好的初始参数，DBN 过程分为两步：第一步，通过显层（输入层）与隐层之间逐层无监督学习初始权重，然后计算显层和隐层神经元被激活的概率分布，使用能量函数作为目标函数，经过逐层训练优化后从而赋予整个网络一个较好的参数值；第二步，经过预训练的模型加入数据标签进行有监督的学习。Lin 等^[60]提出一种基于深度信念网络的服装款式分类方法，该方法首先从线上爬取十类服装款式，经过无监督训练后再进行有监督的学习，得到了较好的表现。DBN 的训练过程包含了无监督学习和有监督学习，由于无监督的学习方式使模型在进行有监督学习之前已具备一定的提取特征能力，因此在缺少标注数据的情况下可提高模型的提取特征能力；缺点是需要固定输入服装图像尺寸，而且如果进行大规模的训练 DBN 算法复杂度且过高时间长。

3.4 基于迁移学习的服装分类

通常一个具有良好分类能力的深度学习模型需要大量服装图像进行训练，这样巨量的训练数据获取成为阻碍服装分类发展的难点。某些特定服装分类任务的服装图像很少，甚至仅有几百或几千张，远远不能满足训练的要求，若直接使用少量数据训练则会导致模型过拟合。迁移学习是解决深度学习数据集稀少的常用方法，其一般过程为：首先将模型在源领域中进行一定的训练，然后使用目标域进行二次微调训练，源领域的数据充足，目标域数据稀少，实际上是把一个领域学习的知识迁移至另一个领域，提高了模型的泛化能力。

Eshwar 等^[61]提出一种基于迁移学习的服装图

像分类方法，该方法使用 GoogLeNet 作为基础特征提取网络，替换并重新训练原全连接层与 softmax 分类层，取得了不错的分类结果，平均准确度为 98%。Lao 等^[62]使用预训练好的 AlexNet 网络进行微调，对服装款式和属性进行分类；Chen 等^[63]在大型数据集上预训练微调后实现服装款式分类；Seo 等^[64]预训练 GoogLeNet 后，对收集 1000 左右的人工标注服装数据进行微调，然后进行服装属性分类。

迁移学习有效解决了一些服装分类任务数据稀少的问题，扩大了深度学习适用范围，但仍存在不足，主要表现为：预训练域与目标域的相似度存在依赖关系，若相似度低可能导致欠适配的情况，甚至可能会比不迁移的时候表现要更差，两域之间相似度高则模型分类效果较好，这需要研究人员考量。

3.5 基于弱监督的服装分类

迁移学习一定程度解决了数据稀少方面的问题，但其实从本质上看是一种跨域的学习方式，跨域学习的时候两域之间的差异会阻碍模型学习能力，弱监督学习作为另一种解决方法在深度学习应用广泛，与迁移学习不同的是，弱监督不会存在跨域问题，“弱”表示数据标注信息可信度不是很高，但也可用于学习。深度学习模型的训练需要大量强标注信息，可以看出，对数据上的要求体现在“量”和“强”两方面。在目前互联网电商行业发达的时代，从网上获取大量服装图像较为容易，“量”的任务从而得以解决。如何获得强标注、大量的数据人工标注是非常困难的，但获得弱标注较为容易，例如线上服装产品页面通常含有与服装对应的文字信息，可以提取文字中关键字作为图像的标注。

Corbiere 等^[65]从线上爬取大量的服装图像及文字信息（如服装名称、类别、描述），并对文字信息用文本处理技术抽取关键字作为服装的标注。另

外,为了加快训练速度,在公开数据集 ImageNet 上预训练 ResNet50 模型初始权重。Inoue 等^[21]提出了一种基于少量人工标注数据的服装分类方法,对网上获取的服装图像和文字信息进行文本技术处理后,依然存在噪声信息,这给模型的训练带来负面影响。为减少噪声信息有助于提高分类效果,该方法将少量数据人工清洗后,输入清洁网络^[66]进行训练获得降噪能力,再输入剩余数据降噪。Xiao 等^[67]提出了一种基于噪声概率模型的服装分类方法,该研究显示,线上的文字信息中存在无明确的类别标签或标签信息存在人为主观误差,对分类结果有影响,为此,该方法采用一种噪声概率模型以验证分类结果的正确性,这可以理解为一种后验的纠错机制,该方法不仅输出真实标签,而且输出错误分类的概率。

表2 深度学习分类方法的对比与分析

学习方式	文献	分类任务	数据集	场景复杂度	评价指标和结果	优点	缺点	适用场景
卷积神经网络	文献[48]	民族服装分类	线上图像	低	Accuracy89%	权值共享参数较少,具有一定的平移、尺度不变性	需要大量已标注数据,对空余位置信息不敏感	丰富的已标注服装数据
	文献[49]	属性分类	线上图像	中	Accuracy70%~81%			
	文献[51]	风格分类	线上图像	低	Accuracy79%			
	文献[53]	款式分类	ACS	中	Accuracy84%			
	文献[54]	款式分类	线上图像	中	Accuracy92%			
	文献[18]	款式和属性分类	DeepFashion ^[18]	中	Top382%			
深度信念网络	文献[60]	款式分类	线上图像	中	Accuracy98.75%	较好的初始参数;无监督的方式	图像尺度固定;数据量大,训练时间长	服装数据量适中,人工标注少
递归神经网络	文献[59]	款式分类	DeepFashion	中	Top390%	权值共享参数较少,具有记忆能力	常见的梯度消失和爆炸问题	需要时序上的分类任务
迁移学习	文献[61]	款式分类	线上图像	低	Accuracy98%	需要的服装数量小;提高模型泛化能力	源域与目标域的差异阻碍模型学习	服装训练数据集少
	文献[62]	属性分类	ACS	中	Accuracy74%			
	文献[63]	款式分类	ACS	中	Accuracy59%			
	文献[64]	风格分类	线上图像	中	Accuracy62%			
弱监督学习	文献[65]	款式分类	DeepFashion	中	Top386%	数据集收集方便;无需标注或少量标注信息,人工标注成本低	需要文本处理技术;文本中的噪声影响模型学习	服装训练数据集少,从线上获得服装数据较为容易
	文献[21]	属性分类	Fashion144k ^[69]	高	mAP 64%			
	文献[67]	款式分类	线上图像	中	Accuracy78%			
	文献[68]	风格分类	Fashion144k	高	Accuracy75%			

4 面临的挑战及发展方向

从目前的研究情况来看,分类任务以款式分类和属性分类为主,还融入了某些特定的分类任务,例如民族服装分类等,总体趋势向着更复杂的实际场景分类任务发展。在诸多款式分类文献中,分类任务的类别数量不同难度也不一样。通常类别数量越多,分类难度越大,因为类别内会存在一些细粒度类别。无论是基于传统图像内容还是基于深度学习的

Simo-Serra 等^[68]提出了一种基于弱标签和度量学习的服装分类方法。综上可以看出,基于弱监督学习服装分类优点是不需要人工标注或仅需少量标注,缺点是需要额外的文本处理过程;同时弱标注中的噪声对模型学习存在负面影响。

3.6 小结

以上介绍了一些基于深度学习模型的服装分类的方法,从总体上可以看出,这些方法以卷积神经网络为主,相比于传统图像内容方法,分类准确度有了明显提升,且分类任务更偏向于实际应用场景;从单项研究看,学者们针对输入图像尺度不一、服装数据集稀少、模型泛化能力弱等问题进行了改进研究,取得了一定成果。各类基于深度学习的服装分类优缺点分析见表2。

服装分类方法,都已取得了一定的成果,但该领域的研究还处在一个正在发展的阶段,技术尚未完全成熟,还存在诸多领域难点和挑战需要解决,需要探索新的发展方向。

4.1 面临的挑战与难点

a) 服装定位不准确。服装定位主要用于复杂场景下的分类任务定位工作,如若将整个含有人和复杂背景的服装图像输入到网络,会严重影响分类性能,所以需要高准确度的服装定位技术。在基于传

统图像内容的服装分类研究中,主要依靠人脸识别^[35]和人体姿态分析技术^[29,34],这类方法会出现目标框过大或过小的情况,过大使得背景噪声多,过小使得一些服装区域被拆剪掉,而且采用人脸识别来确定服装位置存在一定的盲目性;在基于深度学习的服装分类研究中,主要依靠区域推荐网络^[70-74]。Cychnerski等^[49]对区域推荐网络进行了相关实验,但准确度和速度有待提高。

b)添加新类别困难。服装时尚行业变化日新月异,服装类别的迭代速度非常快,为了迎合行业发展趋势,研究可新增分类类别的方法具有实际意义。目前无论是基于传统图像技术,还是基于深度卷积模型的方法,均无法在原有模型基础上对新类别进行识别。当出现新种类时,需要输入所有服装图像数据重新训练模型,随着种类的不断增多,图像数量急剧上升,大数据时代海量的图像数据使得模型的训练越来越困难。虽Kumar等^[55]提出的多层次方法可以初步解决添加新的服装类别,但该方法只能将新类别加在顶层,还不是一个完善的解决方法。

c)如何选择特征。图像特征是解决分类任务的基石,使用不同的图像特征直接影响服装分类任务准确度,对于不同的分类任务,图像特征选择也应不一样,如何有效组合这些特征是难点。基于传统图像技术的分类效果依赖于特征设计人员的相关知识和经验,非常考验人员对图像特征的理解;基于深度卷积特征的方法通常采用深层次语义特征进行分类,一般使用网络中的最后一个全连接层,该层中神经元是经过一系列卷积池化操作的高度聚合值,卷积池化会丢失图像的细节信息,而对于某些分类任务来说,这些信息非常重要,如在细粒度服装分类中浅层或中层的卷积特征效果更好。

d)数据集稀缺。数据集的获取是训练模型的第一步,获得大量含有完整标注信息的服装数据非常困难,当前服装数据的来源有两种方式:一种是使用已公开的数据集;另一种是利用爬虫爬取电商服装网页中的信息。但这两种方式并不能完全解决数据集稀缺的问题,因为公开数据集类别数量有限且一般不会扩充新的类别,而且通用的服装类别无法满足对特定的服装图像分类任务的要求。而从线上爬取大量服装图像需要进行人工标注和清理,这又是一项艰巨的任务。

e)类别数量不平衡。已有研究的服装数据集中不同类别图像数量通常是不均衡的,在这种环境下学习得到的模型在测试时会使结果偏向数量多的种

类,影响准确度。在目前公开服装数据集或自行清理标注的数据集中,一些日常生活服装数量明显多于一些特定的服装,如在Deepfashion数据集中女衬衫的图像数量远比滑雪衫的数量要多得多,前者有24557张图像,后者仅有160张。

f)细粒度识别准确度低。现有的服装图像分类研究的准确度层次不齐,主要体现在粗粒度分类和细粒度分类上。由于粗粒度分类的服装类别之间差异性较大,所以整体上取得了比较好的效果,如对T恤、连衣裙、鞋子、帽子等这类穿在人体不同部位的类别准确性接近100%。但服装的细粒度分类还有较大的提升空间,以往文献对含有细粒度类别的款式分类任务中分类准确度偏低,细粒度分类注重图像中细微的差别,所以分类难度更高,如对绗缝夹克、飞行员夹克、骑士夹克等进行分类。基于深度卷积特征的细粒度分类比普通识别能力强,但与时尚专业的人相比还具有一定差距^[75],普通人依靠纹理和显著差异来做决定,而时尚专业人会考虑到一些细微的差别。

g)领域依赖性。在定义服装图像的场景复杂度时,将其分为低、中和高三种,“低”表示为白色背景、平铺、无形变,“中”表示为电商服装图像(经过人为PS处理),“高”表示为买家秀,在不同场景训练得到的模型通常仅对该场景分类具有适用性,迁移至其他场景会严重影响分类准确度。当前已有文献对领域依赖性的研究较少,均只适用于与训练数据集相同场景下的服装进行识别。

4.2 未来研究方向

a)基于无监督的服装分类。随着时尚产业不断发展,服装种类必然越来越多,已有研究都是采用监督的学习方式,使得增加类别需要重新训练模型,过于繁琐,如能采用无监督自主学习方式似乎可以一定程度上解决拓展类别的问题。此外,无监督学习不需要标注,还解决了大量服装图像标注的问题,无监督自主学习的可行性和有效性已被相关研究证实^[76]。因此,如何拓展模型对新类别识别能力,从图像数据中自动挖掘知识是服装分类领域需要研究的方向。

b)基于细粒度的服装分类。细粒度分类通常是对某大类别下的子类别进行分类,依靠图像中的细微差异,所以细粒度分类要比粗粒度分类难度大得多。当前服装的细粒度分类平均准确度不高,依然存在有一定的提升空间。另外,细粒度分类对检索具有重要作用,因为细粒度分类的结果可以进一

步缩小检索范围,更准确地找到与之相似的服装。

c)可自增数据的服装分类。可自增数据的需求对深度学习有着重要的作用,主要体现在数据过少容易过拟合、服装类别数量不平衡导致分类偏向数量多的一方、新增的类别缺少图像数据等问题上,缺少数据的问题限制了深度学习服装分类的发展,随着近两年深度学习的不断发展,出现了一些可增数据的深度学习网络,例如生成式对抗网络^[77],该网络可以从数据中生成新数据,是否能解决服装领域数据稀少的问题有待研究。

d)基于弱监督的服装分类。大数据时代背景下,这样海量的服装图像数据难以通过人工进行标注,而一般方法都是以强监督的学习方式训练模型,海量标注信息则成为了强监督学习的瓶颈。弱监督学习可以作为突破此瓶颈的力量,从上述有关基于弱监督的服装分类文献来看,弱监督学习得到的模型分类性能不弱于强监督学习,而且平均准确度较高。目前基于弱监督的服装分类方法还处于发展的初级阶段,仍存在上升空间,因此,可以预测未来一段时间内依然会是研究的重点。

e)基于胶囊网络的服装分类。基于卷积神经网络的服装分类经历了一段时期的发展,已取得了不错的表现,但卷积神经网络存在一个明显的缺点,即对服装细节空间位置变化不敏感,例如调整图案、口袋、拉链等风格元素的位置和旋转角度对输出结果影响不大,然而对某些细粒度风格分类任务来说,这些信息的变化极为重要,影响服装的风格决策。近来胶囊网络^[78]渐渐开始受到学者的关注,该网络对空间变化敏感,可以输出位移、旋转等信息,有助于更准确地风格分类。另外,卷积神经网络训练识别一个物体需要多个角度的图像,造成数据集太大,而胶囊网络对姿态信息进行了学习,因此或许可以通过小样本学习模型,解决数据稀少的问题。因此,可以探索基于胶囊网络的服装分类。

5 结束语

服装图像分类技术已经过一段时期的发展,从最初的人工选择性提取图像特征到神经网络自动选取图像特征,准确度不断提升。就当前研究情况看,基于传统图像内容的服装分类方法发展缓慢,而基于深度学习的服装分类方法发展迅速,文献数量较多,原因在于深度学习对分类任务更具广泛性,且相同分类任务的准确度更高。传统图像内容方法也有自身的优势,对数据数量要求较低,当可供使用的数

据稀少时,表现可能优于深度学习技术。

总的来说,服装图像分类仍然是一个极具挑战性的课题,存在诸多可改进的方向。互联网快速发展和硬件性能不断提升,为深度学习方法提供更多发展可能,可以预见在一定时期内,基于深度学习的服装分类方法是未来的主要发展方向。

参考文献:

- [1] 中国国际电子商务中心研究院.中国电子商务报告 2016 [R/OL]. (2017-06-14) [2018-12-17]. <http://images.mofcom.gov.cn/dzsws/201706/20170621110205702.pdf>.
- [2] 中国电子商务研究中心. 2015—2016 年度中国服装电商行业报告 [R/OL]. (2016-12-26) [2018-12-17]. <http://www.100ec.cn/zt/fzyxbg/>.
- [3] 卢兴敬. 基于内容的服装图像检索技术研究及实现[D]. 哈尔滨:哈尔滨工业大学, 2008:1-2.
- [4] 刘聪,丁贵广.基于视觉的服装属性分类算法[J].微电子学与计算机,2016,33(1):28-33.
- [5] Khurana T, Mahajan K, Arora C, et al. Exploiting texture cues for clothing parsing in fashion images[C]// 2018 25th IEEE International Conference on Image Processing. IEEE, 2018: 2102-2106.
- [6] Wang X, Zhang T. Clothes search in consumer photos via color matching and attribute learning[C]//Proceedings of the 19th ACM International Conference on Multimedia.ACM, 2011:1353-1356.
- [7] Sun G L, Cheng Z Q, Wu X, et al. Personalized clothing recommendation combining user social circle and fashion style consistency[J]. Multimedia Tools and Applications, 2018, 77(14): 17731-17754.
- [8] Sun L, Aragon-Camarasa G, Rogers S, et al. Single-shot clothing category recognition in free-configurations with application to autonomous clothes sorting[C]// 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2017: 6699-6706.
- [9] Yang M, Yu K. Real-time clothing recognition in surveillance videos[C]//2011 18th IEEE International Conference on Image Processing. IEEE, 2011: 2937-2940.
- [10] Li A, Liu L, Wang K, et al. Clothing attributes assisted person reidentification[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2015, 25(5): 869-878.
- [11] Kim K J, Park S M, Choi Y J. Clothing identification based on edge information[C]//2008 IEEE Asia-Pacific Services Computing Conference. IEEE, 2008: 876-880.
- [12] Li B, Lian X C, Lu B L. Gender classification by

- combining clothing, hair and facial component classifiers[J]. *Neurocomputing*, 2012, 76(1): 18-27.
- [13] Nazir M, Majid-Mirza A, Ali-Khan S. PSO — GA based optimized feature selection using facial and clothing information for gender classification [J]. *Journal of Applied Research and Technology*, 2014, 12(1): 145-152.
- [14] Ueki K, Komatsu H, Imaizumi S, et al. A method of gender classification by integrating facial, hairstyle, and clothing images [C]//*Proceedings of the 17th International Conference on Pattern Recognition*. IEEE, 2004, 4: 446-449.
- [15] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition [J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [16] Salakhutdinov R, Hinton G. An efficient learning procedure for deep boltzmann machines [J]. *Neural Computation*, 2012, 24(8): 1967-2006.
- [17] Hochreiter S, Schmidhuber J. Long short-term memory [J]. *Neural Computation*, 1997, 9(8): 1735-1780.
- [18] Liu Z, Luo P, Qiu S, et al. DeepFashion: Powering robust clothes recognition and retrieval with rich annotations [C]// 2016 the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016: 1096-1104.
- [19] Hidayati S C, You C W, Cheng W H, et al. Learning and recognition of clothing genres from full-body images[J]. *IEEE Transactions on Cybernetics*, 2018, 48(5): 1647-1659.
- [20] Yamaguchi K, Kiapour M H, Ortiz L E, et al. Parsing clothing in fashion photographs [C]//2012 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2012: 3570-3577.
- [21] Inoue N, Simo-Serra E, Yamasaki T, et al. Multi-label fashion image classification with minimal human supervision[C]// 2017 IEEE International Conference on Computer Vision Workshops. IEEE, 2017: 2261-2267.
- [22] 李东, 万贤福, 汪军. 采用傅里叶描述子和支持向量机的服装款式识别方法[J]. *纺织学报*, 2017, 38(5): 122-127.
- [23] An L, Li W. An integrated approach to fashion flat sketches classification [J]. *International Journal of Clothing Science and Technology*, 2014, 26(5): 346-366.
- [24] Stricker M A, Orengo M. Similarity of color images [C]//*Storage and Retrieval for Image and Video Databases III*. International Society for Optics and Photonics, 1995: 381-392.
- [25] Gao M, Du Y, Ai H, et al. A hybrid approach to pedestrian clothing color attribute extraction[C]//2015 14th IAPR International Conference on Machine Vision Applications. IEEE, 2015: 81-84.
- [26] Ojala T, Pietikainen M, Harwood D. Performance evaluation of texture measures with classification based on Kullback discrimination of distributions [C]//*Proceedings of 12th International Conference on Pattern Recognition*. IEEE, 1994: 582-585.
- [27] Thewsuan S, Horio K. Preprocessing techniques based on LBP and Gabor filters for clothing classification [C]//2016 International Symposium on Intelligent Signal Processing and Communication Systems. IEEE, 2016: 1-6.
- [28] Kalantidis Y, Kennedy L, Li L J. Getting the look: clothing recognition and segmentation for automatic product suggestions in everyday photos [C]//*Proceedings of the 3rd ACM Conference on International Conference on Multimedia Retrieval*. ACM, 2013: 105-112.
- [29] Zheng Q, Chen J, Liang C, et al. Transferring clothing parsing from fashion dataset to surveillance[C]//2017 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2017: 1667-1671.
- [30] 张海波, 黄铁军, 刘莉, 等. 基于支持向量机的面料图像情感语义识别[J]. *天津工业大学学报*, 2013, 32(6): 23-27.
- [31] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110.
- [32] Di W, Wah C, Bhardwaj A, et al. Style finder: fine-grained clothing style detection and retrieval [C]//2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops. IEEE, 2013: 8-13.
- [33] Huo P, Wang Y, Liu Q. A part-based and feature fusion method for clothing classification [C]//*Pacific Rim Conference on Multimedia*. Springer, 2016: 231-241.
- [34] Chen H, Gallagher A, Girod B. Describing clothing by semantic attributes [C]//*European Conference on Computer Vision*. Springer, 2012: 609-623.
- [35] Surakarin W, Chongstitvatana P. Predicting types of clothing using SURF and LDP based on bag of features [C]//*International Conference on Electrical Engineering/Electronics*. IEEE, 2015: 1-5.
- [36] Surakarin W, Chongstitvatana P. Classification of

- clothing with weighted SURF and local binary patterns [C]//2015 International Computer Science and Engineering Conference. IEEE, 2015: 1-4.
- [37] Bossard L, Dantone M, Leistner C, et al. Apparel classification with style [C]// Asian Conference on Computer Vision. Springer, 2012: 321-335.
- [38] Zhang J, Liu L, Huang D, et al. Clothing co-segmentation based on HOG features and E-SVM classifier [C]//2016 6th International Conference on Digital Home. IEEE, 2016: 16-19.
- [39] Li Z, Li Y, Tian W, et al. Cross-scenario clothing retrieval and fine-grained style recognition [C]//2016 23rd International Conference on Pattern Recognition. IEEE, 2016: 2912-2917.
- [40] Lorenzo-Navarro J, Castrillón M, Ramón E, et al. Evaluation of LBP and HOG descriptors for clothing attribute description [C]//International Workshop on Video Analytics for Audience Measurement in Retail and Digital Signage. Springer, 2014: 53-65.
- [41] 纪娟,秦珂,杨若瑜.基于 HOG 和几何特征的服装细节要素识别与分类[J].图学学报,2016,37(1):84-90.
- [42] Sun G L, Wu X, Peng Q. Part-based clothing image annotation by visual neighbor retrieval [J]. Neurocomputing, 2016, 213: 115-124.
- [43] Satta R, Pala F, Fumera G, et al. People search with textual queries about clothing appearance Attributes [M]. Springer, 2014: 371-389.
- [44] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [C]//Advances in Neural Information Processing Systems 25 (NIPS 2012). NIPS, 2012: 1097-1105.
- [45] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2015-04-10) [2018-12-17]. <https://arxiv.org/abs/1409.1556>.
- [46] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions [C]// 2015 the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2015:1-9.
- [47] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [C]// 2016 the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016: 770-778.
- [48] Nawaz M M T, Hasan R, Hasan M A, et al. Automatic categorization of traditional clothing using convolutional neural network [C]//2018 IEEE/ACIS 17th International Conference on Computer and Information Science. IEEE, 2018: 98-103.
- [49] Cychnerski J, Brzeski A, Boguszewski A, et al. Clothes detection and classification using convolutional neural networks [C]//2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation. IEEE, 2017: 1-8.
- [50] Iandola F N, Han S, Moskewicz M W, et al. Squeezenet: alexnet-level accuracy with 50x fewer parameters and <0.5 mb model size [EB/OL]. (2016-11-04) [2018-12-17]. <https://arxiv.org/abs/1602.07360>.
- [51] Dong C Y, Shi Y Q, Tao R. Convolutional neural networks for clothing image style recognition [C]// 2018 International Conference on Computational, Modeling, Simulation and Mathematical Statistics. Pennsylvania: DEStech Transactions on Computer Science and Engineering, 2018:592-597.
- [52] Li R, Mao Y, Ahmad I, et al. Improving deep convolutional neural networks for real-world clothing image [C]//2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery. IEEE, 2017: 837-843.
- [53] Ivanov A Y, Borzunov G I, Kogos K. Recognition and identification of the clothes in the photo or video using neural networks [C]//2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering. IEEE, 2018: 1513-1516.
- [54] Chen L, Han R, Xing S, et al. Research on clothing image classification by convolutional neural networks [C]//2018 11th International Congress on Image and Signal Processing. IEEE, 2018: 1-5.
- [55] Kumar S, Zheng R. Hierarchical category detector for clothing recognition from visual data [C]// 2017 IEEE International Conference on Computer Vision. IEEE, 2017: 2306-2312.
- [56] 张振焕,周彩兰,梁媛.基于残差的优化卷积神经网络服装分类算法[J].计算机工程与科学,2018,40(2):354-360.
- [57] 包青平,孙志锋.基于度量学习的服装图像分类和检索[J].计算机应用与软件,2017,34(4):255-259.
- [58] Dong Q, Gong S, Zhu X. Multi-task curriculum transfer deep learning of clothing attributes [C]//2017 IEEE Winter Conference on Applications of Computer Vision. IEEE, 2017: 520-529.
- [59] Wang W, Xu Y, Shen J, et al. Attentive fashion grammar network for fashion landmark detection and clothing category classification [C]// 2018 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2018: 4271-4280.

- [60] Lin X, Peng L, Wei G, et al. Clothes classification based on deep belief network [C]//2016 3rd International Conference on Informative and Cybernetics for Computational Social Systems. IEEE, 2016: 87-92.
- [61] Eshwar S G, Gautham G P J, Rishikesh A V, et al. Apparel classification using convolutional neural networks[C]//2016 International Conference on ICT in Business Industry & Government. IEEE, 2016: 1-5.
- [62] Lao B, Jagadeesh K. Convolutional neural networks for fashion classification and object detection [EB/OL]. [2015-09-07]. http://cs231n.stanford.edu.cn/reports/blao_kjag_cs231n_finalpaperfashionclassification.pdf.
- [63] Chen J C, Liu C F. Deep net architectures for visual-based clothing image recognition on large database[J]. *Soft Computing*, 2017, 21(11): 2923-2939.
- [64] Seo Y, Shin K S. Image classification of fine-grained fashion image based on style using pre-trained convolutional neural network [C]//2018 IEEE 3rd International Conference on Big Data Analysis. IEEE, 2018: 387-390.
- [65] Corbière C, Ben-Younes H, Ramé A, et al. Leveraging weakly annotated data for fashion image retrieval and label prediction [C]// 2017 the IEEE International Conference on Computer Vision Workshops. IEEE, 2017: 2268-2274.
- [66] Veit A, Alldrin N, Chechik G, et al. Learning from noisy large-scale datasets with minimal supervision [C]// 2017 the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2017: 6575-6583.
- [67] Xiao T, Xia T, Yang Y, et al. Learning from massive noisy labeled data for image classification[C]// 2015 the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2015: 2691-2699.
- [68] Simo-Serra E, Ishikawa H. Fashion style in 128 floats: Joint ranking and classification using weak data for feature extraction [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016: 298-307.
- [69] Simo-Serra E, Fidler S, Moreno-Noguer F, et al. Neuroaesthetics in fashion: Modeling the perception of fashionability [C]// 2015 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2015: 869-877.
- [70] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]// 2014 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2014: 580-587.
- [71] Dai J, Li Y, He K, et al. R-Fcn: Object detection via region-based fully convolutional networks [C]//30th Conference on Neural Information Processing Systems. NIPS, 2016: 379-387.
- [72] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector [C]//European Conference on Computer Vision. Springer, 2016: 21-37.
- [73] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [C]//Advances in Neural Information Processing Systems 28 (NIPS 2015). NIPS, 2015: 91-99.
- [74] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]// 2016 the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016: 779-788.
- [75] Takagi M, Simo-Serra E, Iizuka S, et al. What makes a style: Experimental analysis of fashion prediction [C]// 2017 the IEEE International Conference on Computer Vision Workshops. IEEE, 2017: 2247-2253.
- [76] Le Q V, Ranzato M A, Monga R, et al. Building high-level features using large scale unsupervised learning [EB/OL]. (2012-07-12) [2018-12-17]. <https://arxiv.org/abs/1112.6209>.
- [77] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets [C]//International Conference on Neural Information Processing Systems. NIPS, 2014: 2672-2680.
- [78] Sabour S, Frosst N, Hinton G E. Dynamic routing between capsules [C]// Advances in Neural Information Processing Systems. NIPS, 2017: 3856-3866.

(责任编辑:康 锋)