Vol.47, No.3 Jun.2021

文章编号: 1671-0444(2021)03-0090-06

DOI: 10.19886/j.cnki.dhdz.2020.0168

基于改进 Bilinear-CNN 的服装图像风格识别

李 扬^a,黄 荣^{a, b},董爱华^{a, b}

(东华大学 a.信息科学与技术学院; b.数字化纺织服装技术教育部工程研究中心, 上海 201620)

摘要:为解决服装风格的细粒度识别问题,提出一种改进的 Bilinear-CNN (convolutional neural network)模型。在 VGG16 特征提取网络的输出特征图中引入空间注意力机制,提升对图像中服装区域的关注程度。设计一种单特征通路的双线性池化方法,在减小参数量和计算量的同时,采用全局平均池化和全局最大池化两种操作充分挖掘细粒度特征。试验结果表明,改进的 Bilinear-CNN 模型在 FashionStyle14 数据集上达到 76.4%的识别准确率。相比原始 Bilinear-CNN 模型,改进模型的识别准确率提升 2 个百分点,减少 40% 的参数量和 52% 的计算量,并对含扰动服装图像具有风格识别的稳健性。

关键词:服装风格识别;细粒度识别;卷积神经网络;空间注意力;双线性池化

中图分类号: TP 351.4 文献标志码: A

Fashion style recognition based on an improved Bilinear-CNN

LI Yang^a, HUANG Rong^{a, b}, DONG Aihua^{a, b}

(a. College of Information Science and Technology; b. Engineering Research Center of Digitized Textile & Apparel Technology, Ministry of Education, Donghua University, Shanghai 201620, China)

Abstract: To solve the problem of recognizing fine-grained fashion style from apparel images, an improved Bilinear-CNN (convolutional neural network) model was proposed. The spatial attention mechanism was introduced into the last feature map of a VGG16-based feature extractor, so as to encourage the network to pay more attention to the apparel area. The bilinear pooling method over single feature path was designed, which could fully mine fine-grained features and reduce the number of parameters and the computational burden. Experimental results show that the improved model achieves the recognition accuracy of 76.4% for the FashionStyle14 dataset. Compared with the original Bilinear-CNN model, the proposed model improved the recognition accuracy by 2 percentage points, and reduced the number of parameters and computational burden by 40% and 52%, respectively. The proposed model can robustly recognize the fashion style with the same accuracy level even when the apparel images are distorted by some disturbance.

Key words: fashion style recognition; fine-grained recognition; convolutional neural network; spatial attention; bilinear pooling

收稿日期: 2020-06-26

基金项目: 国家重点研发计划资助项目(2019YFC1521300)

通信作者:黄荣,男,讲师,研究方向为 AI 安全、服装搭配与智能推荐、图像修复与质量评价等,E-mail:rong.huang@dhu.edu.cn

引用格式: 李扬,黄荣,董爱华,基于改进 Bilinear-CNN 的服装图像风格识别[J]. 东华大学学报(自然科学版), 2021,47(3):90-95.

LIY, HUANG R, DONG A H. Fashion style recognition based on an improved Bilinear-CNN[J]. Journal of Donghua University (Natural Science),2021,47(3): 90-95.

随着计算机技术的发展和电子商务的成熟,研究者们开始从服装图像数据中挖掘与服装相关的信息[1-2],其中服装风格是一种重要的服装信息,由服装的结构、色彩、纹理等多种要素综合而成,可表征服装整体的视觉感受[3]。在服装行业中,服装设计者将品牌理念转化为产品风格呈现给消费者[4-5]。服装图像风格的自动识别,不仅可以帮助服装设计者把握潮流趋势,也便于消费者根据风格检索目标服装[6]。

传统的服装图像风格识别方法主要聚焦于如何设计有效的特征描述子。Qian 等^[8] 采用 Canny 算子提取服装轮廓,再设计 3 层神经网络作为分类器以识别服装风格。Chao 等^[9] 将图像的局部二值模式(LBP)和方向梯度直方图(HOG)作为描述服装风格的特征,通过度量特征之间的相似性实现服装风格推荐。吴圣美等^[10] 将尺度不变特征变换(SIFT)等 5 种图像局部描述子与自定义的服装语义属性进行匹配,采用多任务学习模式,实现面向少数民族服装的风格识别。但这些传统的服装图像风格识别方法依赖于专家知识,需要手动设计特征描述子,耗费的人力与时间成本较高。

近年来,深度神经网络被广泛应用于图像识别领 域[11-12],其将表示学习与分类学习级联,通过反向传 播算法,能够自适应地从数据中挖掘出有效的特征信 息,克服了传统方法依赖人工设计特征描述子的缺 点。目前,国内外基于深度神经网络的服装图像风格 识别方法主要有两类。一是以 CD-CNNs(componentdependent convolutional neural networks)[13] 为代表的 基于服装组件的方法。该类方法利用卷积神经网络 提取不同部位服装组件的特征,通过数据驱动学习各 组件特征的融合方式,最后用支持向量机(SVM)进行 服装风格识别;其以服装组件为单位部署特征提取模 型,需要多种与服装组件相关的监督信息,整个模型 的复杂程度高、参数多、计算量大。二是基于图像全 局的端到端方法。该类方法将整幅图像作为深度神 经网络的输入,自动学习能够表示服装风格的特征。 Aoki 等[14] 用 PSPNet(pyramid scene parsing network) 和 SSD(single shot multibox detector)定位图像场景 中的服装区域,通过 ResNet50(residual network50)识 别服装风格。Takagi 等[15] 收集 13 126 幅服装图像, 涵盖礼服、古典、洛丽塔等 14 种风格,并利用 VGG (visual geometry group)等 5 种通用图像识别网络测 试服装风格识别的性能。刘正东等[16]提出一种基 于尺寸分割和负样本的 SSD 检测方法,克服了小目 标识别的困难,提高了西装图像分类的准确率。相

比基于服装组件的方法,该类方法以单个类别标签 为监督信息,将整幅图像作为识别对象,所需的参数 量和计算量较少。

服装风格具有细粒度特性。细粒度特性通常指 目标对象的视觉感知相似,属于同一个粗粒度元类, 具有较为显著的类间相似性[17]。服装风格的细粒 度特性体现在:所有服装图像均属于"服装"元类,而 不同风格的服装是"服装"元类中的不同子类。因 此,对服装图像风格进行识别存在一定的挑战 性[18]。对于细粒度图像识别问题,Lin 等[19]提出一 种 Bilinear-CNN(convolutional neural network)模 型,该模型通过两条并行的网络通路提取图像特征, 并使用双线性池化方法[20] 计算并行特征间的相关 性,筛选能够有效区分细粒度子类的特征。然而, Bilinear-CNN 模型存在两个问题: 一是缺乏关注细 粒度目标的机制,未能有效抑制图像背景区域对识 别性能的影响;二是因使用两条并行的网络通路提 取特征,需成倍的参数量和计算量。针对上述问题, 提出一种改进的 Bilinear-CNN 模型,并将其用于服 装图像风格的识别。

1 改进的 Bilinear-CNN 模型

Bilinear-CNN 模型改进步骤:首先,引入空间注意力机制,通过数据驱动学习注意力的分配权重值,使得网络能够关注体现服装风格细粒度差异的空间区域,达到抑制背景干扰的目的;其次,设计一种单特征通路的双线性池化方法,该方法用两种不同的池化操作从单通道特征中挖掘细粒度特征。改进后的 Bilinear-CNN 模型结构如图 1 所示。

将 VGG16 中 conv1 <u>1</u> 至 conv5 <u>3</u> 的网络部分作为单个网络通路中的特征提取器^[21]。其中,下划线"_"前的数字代表卷积模块序号,下划线"_"后的数字代表某个卷积模块中卷积层的序号。通过ImageNet 数据集预训练的 VGG16 模型能够提取丰富的图像特征。因此,采用这些预训练的参数对特征提取器进行初始化。

使用交叉熵损失和 L_1 范数正则化项的加权和作为代价函数,如式(1)所示。

$$L = \frac{1}{N} \sum_{i=1}^{N} \hat{y}_i \log y_i + \lambda \parallel W \parallel_1$$
 (1)

式中: N 为样本数量; \hat{y}_i 为第 i 个输入图像的类别预测值; y_i 为第 i 个输入图像的真实类别标签; W 为网络参数; λ 为 L_1 范数正则化项的权重参数, 本文将 λ 设为 0.1。

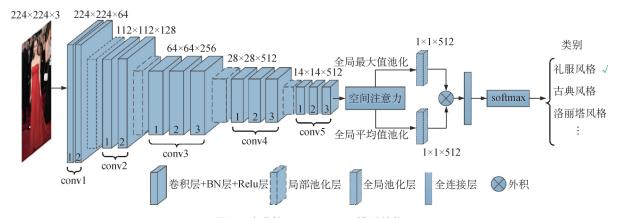


图 1 改进的 Bilinear-CNN 模型结构

Fig.1 Improved Bilinear-CNN model structure

1.1 空间注意力机制

空间注意力机制使得神经网络能够自适应地增强对前景区域的感知,抑制背景部分的干扰,提高识别性能。本文在单特征通路之后嵌入一种可学习的空间注意力网络,从而实现上述机制 $^{[22]}$ 。空间注意力网络的输入为一个 $h \times w \times c$ 的特征图 F,输出为一个 $h \times w \times 1$ 的注意力图 M。该注意力图中每个元素取值范围为[0,1],表示注意力的分配权重值,其大小表征关注程度的强弱。注意力图 M 的计算流程如图 2 所示。

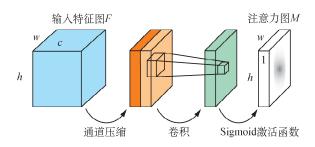


图 2 空间注意力网络 Fig.2 Spatial attention network

注意力图 M 的计算步骤如下:

- (1) 通道压缩。在输入特征图 F 中,以空间位置为单位,分别取每个 c 维特征的最大值和平均值,得到两个具有全局通道信息的 $h \times w$ 矩阵。将这两个矩阵在通道维度上进行拼接,得到 $h \times w \times 2$ 的张量,其中 c > 2,实现通道压缩。
- (2) 卷积。对上一步输出的 $h \times w \times 2$ 的张量做卷积操作,采用一个大小为 $3 \times 3 \times 2$ 、步长为 1、参数可学习的卷积核(共 18 个参数)。为保证卷积前后张量空间尺寸(即 h 和 w)的一致性,对输入张量的边界进行全零填充:上下/左右边界各填充一行/列。
- (3) Sigmoid 激活。使用 Sigmoid 激活函数将 上一步输出张量的取值范围映射到[0,1],生成注意

力图 M。 其中,元素的取值越大,表示网络更加关注该元素所对应的局部感受野。

在训练开始时,空间注意力网络中卷积核的参数被随机初始化。在训练过程中,通过误差反向传播算法调整空间注意力网络中卷积核的参数。经过训练之后,对于一幅给定的服装图像,空间注意力网络能够输出一张注意力图 M,为空间域中的各局部感受野分配合适的注意力权重值,使得模型能够有效聚焦风格差异的服装区域。图 3 示为不同风格服装图像的注意力图 M (已转换为可视化的热力图),其中,图 3(a) (b) (c) 分别为礼裙风格、古典风格和洛丽塔风格的服装图像,图 3(d) (f) 分别为图 3(a) (c) 的注意力图,图中越高亮的像素代表越高的注意力权重值。

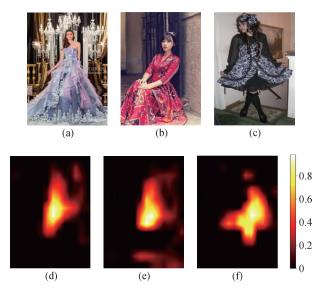


图 3 服装图像及其可视化的注意力图 Fig.3 Apparel images and visualization attention maps

由图 3 可知,注意力图的高亮区域均与原图的服装区域贴合。对于图 3(a),模型关注礼服的胸褶和腰

褶部分;对于图 3(b),模型关注服装的衣身部分;对于图 3(c),模型注重反映洛丽塔风格的蕾丝花纹。这说明所设计的空间注意力网络具有聚焦风格信息、抑制背景干扰的能力。

将注意力图 M 与输入特征图 F 融合,得到经过注意力分配调制的新特征图 F'。注意力图 M 与输入特征图 F 的融合计算如式(2) 所示。首先,将注意力图 M 复制 c 份并拼接成一个 $h \times w \times c$ 的张量,接着,计算该张量与输入特征图 F 的哈达玛积,得到输出特征图 F'。

$$F' = F \circ [\underline{M; M; \cdots; M}]$$
 (2)

式中:";"表示张量拼接操作;"。"表示哈达玛积。

深度神经网络中的深层特征图是浅层特征图的组合,反映图像的高级语义信息。因此,在 VGG16特征提取器的最深层特征图,即 conv5_3 引入上述的空间注意力机制,使得模型能够筛选与服装相关的高级语义信息。在消融试验中,测试了在不同深度的特征图引入空间注意力机制后对细粒度识别性能的影响。

1.2 单特征通路的双线性池化方法

原始 Bilinear-CNN 模型通过两条并行的网络通路提取特征,并利用双线性池化操作计算格拉姆矩阵,提取能够有效区分细粒度子类的特征。然而,并行特征通路的双线性池化方法导致模型参数量和计算量的成倍增加。设计一种单特征通路的双线性池化方法,如图 4 所示,在减少参数量和计算量的同时利用两种不同的池化操作充分挖掘细粒度特征。由图 4 可知,单个特征通路输出一个 h $w \times c$ 的特征图,对该特征图分别进行全局平均池化和全局最大池化,得到两个描述特征图全局信息的向量,再计算这两个向量的外积,得到格拉姆矩阵 G。

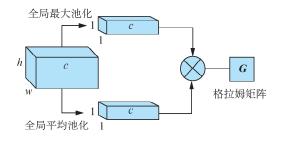


图 4 单特征通路的双线性池化方法
Fig.4 Bilinear pooling method over single feature path

综上所述,对单个网络通路的输出特征图采用 全局最大池化和全局平均池化提取了两种不同的统 计量,并通过格拉姆矩阵来刻画特征通道之间的相 关性。与并行特征通路相比较,单特征通路的双线 性池化方法在降低模型参数量与计算量的同时仍能 较好地挖掘细粒度特征。

2 试验与结果分析

2.1 试验准备

所有试验均采用公开的 FashionStyle14 数据集^[15]。该数据集包含 13 126 幅由服装设计领域专家收集并标注的服装图像,涵盖礼服、古典、洛丽塔等 14 种服装风格标签。该数据集每一类的图像数量均在区间[816,1106]内,其数据具有均衡性。因此,将所有类别识别准确率的均值作为最终总体准确率。将 80%的数据集作为训练集,剩余的 20%作为测试集,分别计算训练集和测试集中样本图像的均值和方差,并对数据集中每一幅样本图像进行去均值和方差标准化处理。

试验环境:CPU 为 Inteli7 8700,内存 16G;GPU 为 Nvidia GTX 2070;操作系统为 Ubuntu18.04;使 用 PyTorch 深度学习框架。

2.2 训练步骤

训练 过程分为两个阶段。第一阶段冻结 VGG16 特征提取器的初始参数,只训练空间注意力 网络和最后的全连接层。初始学习率设为 0.05,批 大小设为 32,训练 50 轮。在此基础上,第二阶段使用较小的学习率微调包括 VGG16 特征提取器在内的全部参数。初始学习率设为 0.001,批大小设为 16,共训练 50 轮。训练过程采用余弦函数衰减学习率 l,如式(3)所示。

$$l = l_0 \times \cos\left(\frac{\pi}{2n}\right) \tag{3}$$

式中: l_0 为初始学习率;n 为训练轮次。使用自适应矩估计优化器(adaptive moment estimation,ADAM)^[23]作为网络参数更新的优化算法,其一阶和二阶动量项分别设为 0.90 和 0.99。

2.3 试验结果分析

设计模型性能对比试验、消融试验和稳健性试验来验证改进 Bilinear-CNN 模型的细粒度识别性能,并采用 Top-1 准确率作为衡量识别性能的指标。参数量统计网络中参与训练的参数个数之和,单位以百万(M)计,作为衡量参数量的指标。采用浮点运算次数 (floating point of operations, FLOPs)作为衡量计算量的指标,单位以十亿(B)计。使用 PyTorch 框架中的 torchstat 软件库计算

模型的参数量和 FLOPs。

2.3.1 模型对比试验

在识别效果方面,将改进的 Bilinear-CNN 模型和 CD-CNNs 模型^[13]、ResNet50 模型^[15]、原始Bilinear-CNN 模型^[19]进行比较,其中,CD-CNNs 是基于检测服装组件的图像识别模型,ResNet50 是通用的图像识别模型。4 种模型经训练后对于服装风格的识别效果如表 1 所示。

表 1 服装图像风格识别效果对比

Table 1 Performance comparison for fashion style recognition

模型	参数量/ M	FLOPs/	Top-1 准确率 /%
$\text{CD-CNNs}^{\llbracket 13 \rrbracket}$	24	4.2	72.0
$ResNet50^{[15]}$	146	25.2	77.7
原始 Bilinear-CNN ^[19]	30	32.0	74.4
改进 Bilinear-CNN	18	15.3	76.4

注:CD-CNNs 模型的参数量与 FLOPs 仅计入了特征提取部分,未计入检测部分。

由表 1 可知:相比原始 Bilinear-CNN 模型,改进 Bilinear-CNN 模型的识别准确率提升 2 个百分点,减少了 40%的参数量和 52%的计算量;改进Bilinear-CNN 模型的识别准确率(76.4%)接近最高值77.7%,但其参数量仅为 CD-CNNs 模型的 75%。该试验结果证实了本文所提方法的有效性。

2.3.2 消融试验

设计两组消融试验。消融试验一用于对比空间 注意力机制和单特征通道的双线性池化方法对识别 性能的影响,如表 2 所示。消融试验二用于对比在 不同深度的特征图引入空间注意力网络对识别性能 的影响,如表 3 所示。

表 2 消融试验一的结果

Table 2 The results for the first ablation experiment

空间注意力机制	单通道的 双线性方法	参数量/ M	FLOPs/	Top-1 准确率 /%
×	×	15	15.0	72.0
\checkmark	×	15	15.0	73.2
×	\checkmark	18	15.3	75.2
\checkmark	\checkmark	18	15.3	76.4

注:试验选择对应项则用" \checkmark "表示:试验未选用对应项则用" \times "表示。

表 3 消融试验二的结果

Table 3 The results for the second ablation experiment

引入位置	conv1 <u>2</u>	conv2 <u>2</u>	conv3 <u>3</u>	conv4 <u>3</u>	conv5 <u>3</u>
Top-1 准确率 /%	74.3	74.4	74.2	75.5	76.4

对比表 2 中前两行试验数据可知,引入空间注意力机制之后,在未显著增加网络参数量和计算量的前提下,风格识别的准确率上升了 1.2 个百分点。这是因为所提出的空间注意力网络仅使用了一个大小为 $3\times3\times2$ 的卷积核,参数量为 18,远小于百万(M)量纲,可近似忽略。比较表 2 中第 1 行和第 3 行的试验数据可知,单特征通路的双线性池化方法在仅增加 3 M 参数量、0.3 B 计算量的前提下,将识别准确率提升到了 75.2%。由表 2 中的第 4 行数据可知,同时使用空间注意力机制和单特征通道的双线性池化方法使得识别准确率达到最高值 76.4%,均优于表 2 中 $1\sim3$ 行的识别准确率。由此验证了空间注意力机制抑制背景干扰的作用以及双线性方法挖掘细粒度特征的能力。

由表 3 可见:随着深度的增加,识别准确率逐步提升;在 $conv5_3$ 引入空间注意力网络时,识别准确率达到最高值 76.4%。这验证了高级语义信息有助于挖掘细粒度特征。

2.3.3 稳健性试验

在实际应用中,由于拍摄条件的限制,在成像的服装图像中存在旋转、低分辨率和光照等扰动,因此,服装风格识别系统的稳健性是面向实际应用的需求之一。对改进前后的 Bilinear-CNN 模型进行稳健性测试。对测试集中图像分别进行旋转变换、缩放变换和亮度变换,依次增大各变换参数的随机取值范围,分别对测试集中图像进行 5 次试验,识别准确率取 5 次试验均值。

稳健性试验结果如表 4 所示。其中,旋转角度是指旋转变换后的图像与原图像夹角,图像尺寸是指尺度变换后的图像空间分辨率,亮度因子是指亮度变换后的图像相比原始图像的整体亮度增益。由表 4 可知,改进 Bilinear-CNN 模型的识别准确率下降幅度较小,均优于原始 Bilinear-CNN 模型,说明所提方法针对 3 种扰动具有较好的识别稳健性。当进一步扩大参数随机取值范围时,识别准确率将急剧下降,此时需要通过数据扩充或扰动参数估计来提高识别模型的稳健性。

表 4 稳健性试验结果

Table 4 The results for robustness experiment

	分 粉 吹 和	Top-1 准确率 /%		
变换 参数	参数随机 取值范围	原始 Bilinear-CNN	改进 Bilinear-CNN	
旋转角度	(-5°, 5°) (-30°, 30°) (-45°, 45°)	74.1 72.7 71.1	75.2 73.4 72.6	
图像 尺寸	(202,224) (179,224) (157,224)	72.6 71.3 70.6	73.4 72.5 71.0	
亮度 因子	(0.9, 1.1) (0.8,1.2) (0.7,1.3)	74.0 73.3 73.1	75.4 74.6 74.1	

3 结 语

针对服装风格的细粒度识别问题,提出一种改进的 Bilinear-CNN 模型。在 VGG16 特征提取网络的输出特征图中引入空间注意力机制,增强对图像中服装目标区域的感知能力,抑制复杂背景的干扰。设计了一种单特征通路的双线性池化方法,通过两种不同的池化操作挖掘细粒度特征。试验结果表明,相比原始Bilinear-CNN 模型,所提方法使得识别准确率提升了2个百分点,同时减少了40%的参数量和52%的计算量。下一步研究工作将聚焦于强稳健性的细粒度识别问题及其在服装图像风格识别模型中的应用。

参 考 文 献

- [1] AL-HALAH Z, STIEFELHAGEN R, GRAUMAN K. Fashion forward: forecasting visual style in fashion[C]//IEEE International Conference on Computer Vision. 2017: 388—397.
- [2] KASHILANI D, DAMAHEL B, THAKUR N V. An overview of image recognition and retrieval of clothing items [C]// International Conference on Research in Intelligent and Computing in Engineering. 2018: 1—6.
- [3] 夏帆,刘翔.基于服装设计要素和法则的智能配搭设计[J].纺织学报,2015,36(7):94-99.
- [4] 张霞萍.基于服装品牌风格的吊牌设计研究[D].杭州:浙江理工大学,2016.
- [5] 卢义.服装品牌理念设定与风格定位的依据与流程研究[D].上海:东华大学,2005.
- [6] 孙菲菲,徐平华,丁雪梅.服装风格客观评价的研究现状与不足[J].上海纺织科技,2017,45(10):1-4.
- [7] YAN CR, ZHOU L, WAN Y. A multi-task learning model for better representation of clothing images [J]. IEEE Access, 2019,7: 34499-34507.
- [8] QIAN S Q, JIANG L F, DONG A H. Silhouette shape and

- detail texture based garment style recognition [C]//IEEE International Conference on Computer Science and Automation Engineering. 2011: 441–445.
- [9] CHAO X F, HUISKES M J, GRITTI T, et al. A framework for robust feature selection for real-time fashion style recommendation [C]//ACM International Workshop on Interactive Multimedia for Consumer Electronics. 2009: 35—42.
- [10] 吴圣美,刘骊,付晓东,等.结合人体检测和多任务学习的少数民族服装识别[J].中国图象图形学报,2019,24(4):562—572.
- [11] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet classification with deep convolutional neural networks [C]// Proceedings of the 25th International Conference on Neural Information Processing System. 2012; 1097—1105.
- [12] HEK M, ZHANGX Y, RENS Q, et al. Deep residual learning for image recognition [C]//IEEE Conference on Computer Vision & Pattern Recognition.2016; 770—778.
- [13] YAMAMOTO T, NAKAZAWA A. Fashion style recognition using component-dependent convolutional neural networks [C]//IEEE International Conference on Image Processing. 2019: 3397—3401.
- [14] AOKI R, NAKAJIMA T, OKI T, et al. Accuracy improvement of fashion style estimation with attention control of a classifier [C]//IEEE International Conference on Consumer Electronics. 2019: 289-294.
- [15] TAKAGI M, SIMOSERRA E, IIZUKA S, et al. What makes a dtyle: Experimental snalysis of fashion prediction[C]//IEEE International Conference on Computer Vision Workshop. 2017: 2247—2253.
- [16] 刘正东,刘以涵,王首人,西装识别的深度学习方法[J].纺织学报,2019,40(4):58-64.
- [17] 邓旭冉, 闵少波, 徐静远, 等. 深度细粒度图像识别研究综述[J]. 南京信息工程大学学报(自然科学版), 2019, 11(6): 625-637.
- [18] SEO Y, SHIN K S. Image classification of fine-grained fashion image based on style using pre-trained convolutional neural network[C]//IEEE 3rd International Conference on Big Data Analysis, 2018; 387-390.
- [19] LIN T, ROYCHOWDHURY A, MAJI S. Bilinear CNN models for fine-grained visual recognition [C]//IEEE International Conference on Computer Vision. 2015: 1449—1457.
- [20] CARREIRA J, CASEIRO R, BATISTA J, et al. Semantic segmentation with second-order pooling[C]//IEEE European Conference on Computer Vision. 2012: 430—443.
- [21] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J/OL]. Computer Science, arXiv Prepr. arXiv1409.1556, 2014.
- [22] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]//European Conference on Computer Vision. 2018: 3—19.
- [23] KINGMA D P, BA J. ADAM: a method for stochastic optimization [C]//International Conference on Learning Representations. 2015: 1—15.

(责任编辑:冀宏丽)