

文章编号: 2095-2163(2020)05-0014-05

中图分类号: TP311.52

文献标志码: A

基于深度学习的服装风格识别问题的研究

张艺凡¹, 刘国华¹, 盛守祥², 王国栋², 王力民²

(1 东华大学 计算机科学与技术学院, 上海 201620; 2 华纺股份有限公司, 山东 滨州 256602)

摘要: 随着电子商务的发展,网络购物已成为选购服装商品的主要渠道,使得服装图像的准确分类变得愈发重要。近年来,对服装图像分类的研究大多专注于服装种类,而对服装风格识别的研究相对较少。随着线上服装交易量的不断增加,各类平台累积了大量无法得到充分利用的未标注风格的服装图像。针对这一问题,本文运用基于 Inception v3 的迁移学习技术,构建服装风格识别模型,并通过 Deepfashion 服装数据集对该模型进行重训练,生成服装风格分类器,而后将此服装风格分类器应用于对服装图像数据集进行批量标注。

关键词: 服装图像数据集; 服装风格识别; Inception v3

Research on clothing style recognition based on deep learning

ZHANG Yifan¹, LIU Guohua¹, SHENG Shouxiang², WANG Guodong², WANG Limin²

(1 School of Computer Science and Technology, Donghua University, Shanghai 201620, China;

2 Huafang Co., Ltd., Binzhou Shandong 256602, China)

[Abstract] With the development of e-commerce, online shopping has become the main channel for people to buy clothing products, and accurate classification of clothing images has become very important. In recent years, researches on clothing image classification have mostly focused on clothing types, while clothing style recognition research has been relatively rare. With the continuous progress of online clothing transactions, various platforms have accumulated a large number of unlabeled clothing images that cannot be fully utilized. In response to this problem, this paper uses transfer learning technology to build a clothing style recognition model based on Inception v3, and uses Deepfashion The clothing dataset is retrained to generate a clothing style classifier. Finally, this clothing style classifier is applied to batch label the clothing image dataset.

[Key words] Clothing image dataset; clothing style recognition; Inception v3

0 引言

随着经济的发展,人们在穿着打扮方面愈发追求个性化风格。服装的流行趋势随着时代的发展不断更迭,风格也在不断扩展。随着电子商务的蓬勃发展,在网络选购自己心仪的服装商品已然成为一种潮流和趋势。2019 年度中国消费者网购服装情况如图 1 所示。

商品。服装类电商平台包含大量的服装图像,对这些图像进行服装风格的分类,有助于帮助用户快速定位到满足自己需求的商品区域^[1]。除此之外,对服装图像进行服装风格属性标注,可有效利用网络服装图像数据。已标注风格属性的服装图像在各类服装检索及推荐系统中有良好的应用前景。

由于服装风格的评判标准个性化强且风格特征发展迅速^[2],其识别和分类问题非常复杂。服装风格通常被认为是一个相对模糊的概念,一般采用主观评价的方法,一直以来缺少指标和主观感受之间的对应关系,因此通过总结和提取各类服装图像特征,对服装风格进行量化具有重要的研究意义^[3]。最原始的服装风格标注方法是文本手动标注,即通过人为判断来对服装图像的风格进行标注,费时费力,且人为参与标注加入了主观因素。

随着计算机技术的发展,在图像识别领域出现

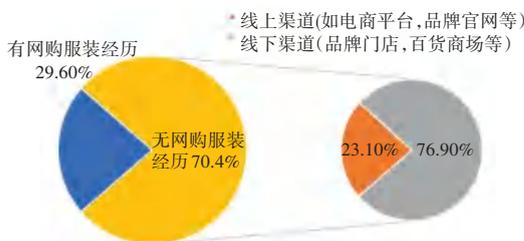


图 1 2019 年中国消费者网购服装比例和渠道偏好

Fig. 1 Chinese consumers' clothing purchase ratio and channel preferences in 2019

在网购的过程中,人们会根据服装图像来取舍

基金项目: 科技部国家重点研发计划(2017YFB0309800); 上海市工业互联网创新发展专项项目(2019-GYHLW-004)。

作者简介: 张艺凡(1996-),女,硕士研究生,主要研究方向:人工智能、图像处理;刘国华(1966-),男,博士,教授,博士生导师,主要研究方向:人工智能、大数据、关系数据库。

通讯作者: 张艺凡 Email: m18522608197@163.com

收稿日期: 2020-03-30

了大量提取图像特征的算法。传统特征提取方法分为针对图像全局和局部进行特征提取两类。常见的全局特征包括颜色特征^[4]、纹理特征^[5]和形状特征^[6]。针对基于颜色特征的服装图像识别问题,构建了一个新的时尚图像数据集,称为 Colorful-Fashion^[7]。除此之外,有文献提出了一种基于颜色特征的服装分割方法,该方法利用形状和颜色信息对服装商店数据库中服装图像进行服装分割^[8],基于此类全局特征尤其形状特征可以辨别出服装的类别。但是由于服装的主观风格量化难度较高,通过形状、纹理等简单特征无法准确识别服装风格。

常见局部特征提取方法有: SURF^[9]、SIFT(尺度不变特征变换)^[10]、HOG(方向梯度直方图)等。但这类传统的特征抽取算法有无法处理遮挡和无法提取光滑边缘等缺点,其提取出的特征难以用于服装图像的分类。针对服装图像的场景多样杂乱以及遮挡等问题,利用基于多模态语义关联规则(MMSAR)的自动融合关键词和视觉特征的 Web 图像检索方法加以解决^[11]。

近年来,随着深度学习的快速发展,深度学习在计算机视觉领域的应用取得了突破性进展。P 使用 BP 神经网络来识别机织物的图案^[12],从分类中提取正常织物的质地和结构特征,输入到神经网络中,以完成学习过程。针对当前越来越多“以图搜图”的需求,期望通过街拍图片来搜索到同一服装的卖家图片^[13]。

传统的特征提取方法难以利用大数据优势,依赖手工调整参数,因此特征中参数量较少;而深度学习可以从海量数据中学习包含多参数的特征,因此深度学习提取出的特征表达效果更好,更适用于识别服装

风格这类需要大量特征对比来判断的属性。

本文采用基于深度学习的特征提取模型对服装图像的风格进行识别,根据提取出的特征训练服装风格分类器,通过将预测结果写入数据库记录分类结果,实现服装图像风格的批量识别与标注。

1 基于卷积神经网络的服装分类方法

1.1 服装风格分类体系

服装风格评定因素分为客观评定因素和主观评定因素^[14]。客观评定因素即服装的直观样式,包括面料、袖长、领形等易识别特征。服装的主观评定因素,如:淑女风格、运动风格等等。

服装的主观风格的定义主要包括 3 个方面:(1)时代特色和民族传统。(2)面料,颜色,图案纹路等。(3)服装的功能性。当代主流的服装风格包括:都市、优雅、复古、淑女、可爱、简约、民族、职业等。

本文对六类服装风格进行区分和识别。(1)简约风格。服装在色彩特征方面以黑、白、灰色为主,较少使用图案,常采用廓形设计,形状宽松。(2)优雅风格。服装在色彩特征方面以深色为主,常用色为红色、黑色、紫色,给人感觉低调且华丽,形状修身且偏长。(3)复古风格。服装以深棕、卡其色等为主体色调,通常采用素色图案,形状特征主要有荷叶边领、方形领、灯笼袖等突出形状。(4)民族风格。服装的颜色通常以艳丽的红色和蓝色为主,多采用复杂且重复的纹路和图案。(5)可爱风格。服装代表颜色为淡粉色,也常采用对比强的颜色,款式层叠较多且造型复杂。(6)职场风格。服装多采用黑、白、深蓝色,款式固定,形状对称。六大服装风格指定特征总结如表 1 所示。

表 1 各类服装风格特征总结

Tab.1 Summary of various clothing style characteristics

服装风格	颜色特征	形状特征	其他特征
简约风格	主黑、白、灰	廓形设计,形状宽松	纯色为主较少使用图案
优雅风格	常红、黑、紫	修身且偏长	低调且华丽
复古风格	主深棕、卡其	较多突出形状特征	荷叶边、方形领、灯笼袖
民族风格	主艳丽	繁复且偏长	复杂且重复的纹路图案
可爱风格	淡粉色,对比强	层叠较多且复杂	色彩活泼,图案空间大,
职场风格	黑、白、深蓝	款式固定,形状对称	典型为西装

1.2 Inception 深度学习模型

GoogLeNet 团队提出的 Inception 结构,是用基础神经元搭建的一个稀疏性、高计算性的网络结构^[15]。Inception 网络结构经历了 V1, V2, V3, V4 这 4 个版本, Inception V1 的网络结构如图 2 所示。

该结构将卷积神经网络中的卷积,池化操作叠加在一起,增加了网络对尺寸的适应性,并在 3 * 3 卷积、5 * 5 卷积和 max pooling 前加上了 1 * 1 卷积来降低特征图的厚度并修正 ReLU。与传统 CNN 一样, V1 在每一次卷积操作之后,通过 ReLU 激励函数来为网络添加非线性特征。

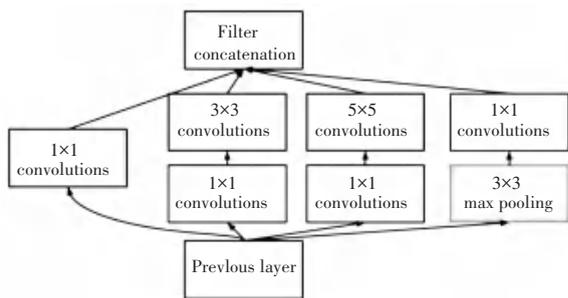


图2 Inception V1的网络结构

Fig. 2 Network structure of Inception V1

Inception V1 虽然准确率较高,但是堆叠网络使 Inception V1 的计算效率很低。因此在 V2 网络中,用更多的小尺寸卷积核来代替大尺寸的卷积核来减少参数量^[16]。但在训练过程中,每一层输入的分布由于网络中参数的更新而发生变化,导致训练变难,此现象被称为 internal covariate shift^[16]。为了解决这一问题,在 V2 网络中添加了 Batch Normalization 层来对每一层神经元的输入做归一化。

与 V2 的改进不同, V3 提出了非对称分解的概念,即将 $n \times n$ 的卷积层分解为两个 $1 \times N$ 和 $N \times 1$ 的卷积层。例如:将 7×7 分解成两个一维的卷积($1 \times 7, 7 \times 1$),这样既可以进一步加速计算,又可以将 1 个卷积分解为 2 个卷积,使得网络深度进一步增加。实验证明 V3 的非对称分解方式训练结果优于 V2。

对 V3 进一步分析和优化, Inception v4 引入了专用的缩减块(reduction block),添加更多一致的模块从而提高性能^[17]。Inception-ResNet 则参考了 ResNet 的设计思想,利用残差连接(Residual Connection)来改进 V3 和 V4 结构^[18],将 Inception 模块的卷积运算输出添加到输入上,残差运算要求卷积计算的输入输出具有相同的维度,因此 Inception-ResNet 在初始卷积后使用 1×1 卷积来匹配深度。

2 服装风格分类模型

2.1 ImageNet 数据集分类任务的迁移

ImageNet 提供了易于访问的图像数据库,共有 14 197 122 幅图像,分为 21 841 个类别(synsets)。

迁移学习是一种机器学习方法,是指在一个特定的数据集上,重新利用已经训练过的卷积神经网络,将其改造或迁移到一个不同的数据集中。

计算机视觉中使用深度学习技术进行图像处理,通常需要大量标记的训练数据和超强的计算能力。在训练过程中,数据量缺少或者运行速度过慢,使训练过程十分艰难。因此,在训练时,常将通过海量数据训练出的模型应用于新的实验中。

2.2 图像分类器重训练

(1) 图像数据集。计算的服装数据集全部来自于 Deepfashion^[19] 数据集,起初的分类采用人工分类的方式。六类风格中的服装图像个数如表 2 所示。

表 2 服装图像数据集个数分布

Tab. 2 Distribution of clothing image data sets

服装风格	训练集	测试集	验证集
简约风格	4 519	564	565
优雅风格	5 613	701	702
复古风格	4 865	608	608
民族风格	3 217	402	402
可爱风格	4 697	587	587
职场风格	5 139	642	642

(2) 训练。重用了在 ImageNet 数据集上训练的图像分类器的特征提取模块,并在顶部训练新的分类层,特征提取模块采用了 Inception V3 架构。

第一阶段是将所有图片输入到预训练好的模型,计算并存储所有图像的特征向量,分类器根据图像特征向量来区分要识别的所有类。第二阶段是对网络顶层进行训练。根据数据集数量对训练过程的超参数进行调整。在训练时采用的超参数数值如表 3 所示。

表 3 训练过程中的超参数数值

Tab. 3 Hyper parameter values during training

超参数	数值
Learning rate	0.01
Step	8 000
Train_batch_size	256
Validation_batch_size	128
Validation set 占比	10%
Test set 占比	10%

其中, Step 表示训练时会运行 8 000 个步骤,每一步都将训练集图像的特征向量输入到最后一层进行预测,将预测结果与实际进行对比,通过反向传播更新权重。Train_batch_size 和 Validation_batch_size 分别表示每一次训练时随机抽取图像的数目和验证时随机抽取图像的数目。

2.3 训练参数评估

在处理分类问题时,分类结果通常分为以下 4 种情况:(1) TP: 将正类预测为正类数。(2) TN: 将负类预测为负类数。(3) FP: 将负类预测为正类数。(4) FN: 将正类预测为负类数。

(1) 准确率。准确率(Accuracy)是指分类正确的样本占总样本个数的比例,可表达为:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

训练集准确率和验证集准确率如图 3 所示。其

中,橙色和蓝色分别表示训练集的正确率和验证集的正确率,训练的正确率达到了 0.74,验证的正确率大约为 0.72。

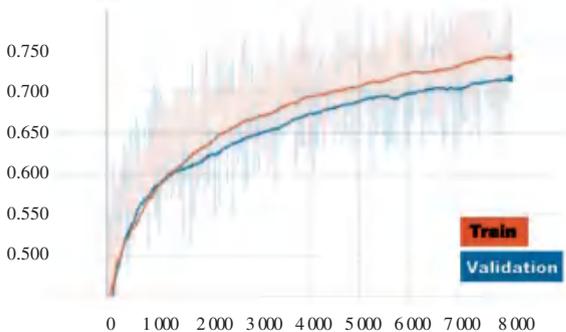


图 3 训练准确率和验证准确率

Fig. 3 Training accuracy and verification accuracy

(2)交叉熵。交叉熵用来描述模型预测值与真实值的差距大小。机器学习算法中,交叉熵代价函数作为模型的优化目标函数,交叉熵代价函数表达式为:

$$L = - [y \log \hat{y} + (1 - y) \log (1 - \hat{y})].$$

训练过程中,训练集交叉熵值和验证集交叉熵值的变化如图 4 所示,经过训练后,训练集的损失达到 0.75,验证集达到 0.82。

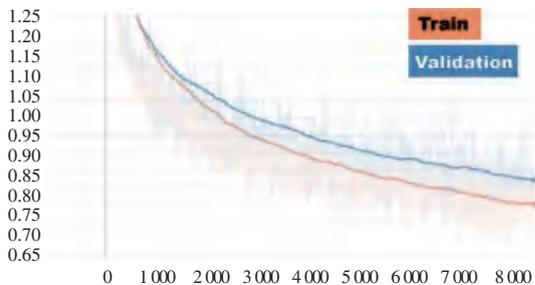


图 4 训练交叉熵和验证交叉熵

Fig. 4 Training cross-entropy and verifying cross-entropy

3 服装风格识别模型应用

在网购过程中,服装风格的分类有助于帮助用户根据自身喜好快速定位到自己所想要找的服装区域。目前,各类网站对服装图像的标注大多只关注服装的种类而忽略了服装的风格,累积了大量未标注风格属性的服装图像,采用服装风格预测模型对服装图像进行批量标记,对未标注服装风格的图像数据集进行属性标注具有重要研究意义。

批量标记服装风格将遍历所有的图片,将图片依次输入服装风格识别模型,得到各类别分类的概率,取最大概率的风格类别作为预测结果。预测概率数值及预测如图 5 所示。其中 A-F 代表 A:简约风格;B:优雅风格;C:复古风格;D:民族风格;E:可爱风格;F 职场风格。

ID	imgname	style	style_prob	A	B	C	D	E	F
30	img_033.jpg	A	A	0.10005640	0.37528103	0.12020980	0.13417020	0.25271810	0.02860430
37	img_036.jpg	C	C	0.10005640	0.14018170	0.07852454	0.02370300	0.50920040	0.03034814
41	img_034.jpg	B	B	0.02876770	0.25712810	0.22309020	0.17683340	0.24881170	0.06040170
31	img_041.jpg	A	A	0.02632794	0.08821510	0.11013020	0.26165180	0.03974270	0.17244410
42	img_074.jpg	C	C	0.29142330	0.12430020	0.08181880	0.08543030	0.59440290	0.07032560
40	img_040.jpg	C	C	0.21003240	0.08760420	0.07009910	0.00770300	0.10721510	0.08782010
43	img_039.jpg	C	C	0.22344920	0.17077770	0.14978360	0.06262210	0.17271890	0.04459950

图 5 批量预测服装类别

Fig. 5 Batch prediction clothing category

4 结束语

随着人们个性化的追求和对自身穿衣风格的明确,服装风格已经成为各类服装检索系统和电商平台中服装图像不可缺少的标注属性。因此对服装图像的风格进行精确标注具有重要的研究意义和良好的应用前景。

参考文献

- [1] 高珊. 基于距离度量学习和多视图学习的服装主观风格识别方法[D]. 浙江大学,2016.
- [2] 钟晓东. 基于认知特征的服装风格自主分类的研究与实现[D]. 东华大学,2012.
- [3] GOLDSMITH R E, FLYNN L R, CLARK R A. Materialistic, brand engaged and status consuming consumers and clothing behaviors [J]. Journal of Fashion Marketing and Management, 2012, 16(1): 102-119.
- [4] 王静. 基于颜色特征的图像分类算法研究[D]. 大连海事大学, 2008.
- [5] 苑丽红, 付丽, 杨勇, 等. 灰度共生矩阵提取纹理特征的实验结果分析[J]. 计算机应用, 2009, 29(4): 1018-1021.
- [6] 曹茂永, 孙农亮, 郁道银. 几种边界特征描述方法的比较研究[J]. 光学技术, 2003, 29(3): 284-287.
- [7] LIU S, FENG J, DOMOKOSC, et al. Fashion Parsing with Weak Color-category Labels [J]. IEEE Transactions on Multimedia, 2014, 16(1): 253-265.
- [9] LOWE D G. Distinctive Image Features from Scaleinvariant Keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [8] MANFREDI M, GRANA C, Calderara S, et al. A Complete System for Garment Segmentation and Color Classification [J]. Machine Vision and Applications, 2014, 25(4): 955-969.
- [9] Herbert Bay, Tinne Tuytelaars, Luc J. Van Gool. SURF: Speeded Up Robust Features [C]// Computer Vision - ECCV 2006, 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006, Proceedings, Part I. Springer-Verlag, 2006.
- [10] Eric N. Mortensen, Hongli Deng, Linda G. Shapiro. A SIFT descriptor with global context [C]// 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005), 20-26 June 2005, San Diego, CA, USA. IEEE, 2005.
- [11] HE R, XIONG N, YANG L T, et al. Using Multi-Modal Semantic Association Rules to fuse keywords and visual features automatically for Web image retrieval [J]. Information Fusion, 2011, 12(3): 223-230.
- [12] PAN R, GAO W, LIU J, et al. Automatic Recognition of Woven Fabric Pattern Based on Image Processing and BP Neural Network [J]. The Journal of the Textile Institute, 2011, 102(1): 19-30.