

基于卷积神经网络的面料检索系统

王彪, 毋涛

(西安工程大学 计算机科学学院, 陕西 西安 710600)

摘要:针对目前市场上纺织面料图像种类多、视觉特征复杂、数据量大的特点,为解决传统图像检索方法存在的检索速度慢、检索精度低的问题,提出一种基于深度学习的面料检索方法。该方法采用迁移学习,微调 VGG16 网络的结构,加入 BN 层提高模型的泛化能力,调整 FC 层神经元的数量减少计算量。同时优化损失函数以约束模型学习相似或相同的面料的特征也相似。以纺织企业提供的面料为数据集训练模型,得到提取面料特征的网络模型。进行在线面料检索时,使用 faiss 向量检索工具,快速计算待检索的面料的特征与特征库的相似度,得到相似度 top-k 的检索结果。经过实验证明,在企业面料数据集上,该系统检索 mAP 可达到 0.892,检索时间仅为 0.012 s,均优于以往的算法,从而验证了其可行性。

关键词:迁移学习; VGG16; 损失函数; faiss; 面料检索

中图分类号: TP391.41

文献标识码: A

文章编号: 1673-629X(2023)09-0052-05

doi: 10.3969/j.issn.1673-629X.2023.09.008

Fabric Retrieval System Based on Convolutional Neural Network

WANG Biao, WU Tao

(School of Computer Science, Xi'an Polytechnic University, Xi'an 710600, China)

Abstract: Aiming at the characteristics of many types of textile fabric images on the market, complex visual features and large amount of data, in order to solve the problems of slow retrieval speed and low retrieval accuracy of traditional image retrieval methods, a fabric retrieval method based on deep learning is proposed. It adopts transfer learning, fine-tunes the structure of the VGG16 network, adds the BN layer to improve the generalization ability of the model, and adjusts the number of neurons in the FC layer to reduce the amount of computation. At the same time, the loss function is optimized to constrain the model to learn the characteristics of similar or the same fabrics are also similar. Using the fabric provided by the textile enterprise as the data set to train the model, the network model for extracting fabric features is obtained. When conducting online fabric retrieval, the faiss vector retrieval tool is used to quickly calculate the similarity between the characteristics of the fabric to be retrieved and the feature library, and obtain the retrieval result of similarity top-k. Experiments show that on the enterprise fabric data set, the retrieval mAP of this system can reach 0.892, and the retrieval time is only 0.012 seconds, which is better than that of the previous algorithm, thus verifying its feasibility.

Key words: transfer learning; VGG16; loss function; faiss; fabric retrieval

0 引言

近些年来,随着互联网经济的快速发展,电子商业也进入了蓬勃发展的阶段,消费者对于以纺织面料为原材料的商品,如服装、鞋等的需求量大大增加。面料生产企业为了能满足市场需求,不得不研究生产新的面料,这样也就导致市场上的面料种类越来越多,从而给面料生产企业带来了新的问题,即如何从种类繁杂的面料中快速、精准地检索到目标面料^[1]。

传统的纺织企业在进行面料检索时,一般都是采取人工的方式进行。该方式不仅费时费力,检索结果

还附带主观性的影响,不能达到用户对检索速度和准确度的要求^[2]。市场上虽然存在成熟的图像检索方法和系统,但是由于面料图像的特殊性,并不适用于面料图像的检索需求,需要进行较大的改进。

在图像检索领域,主要存在两种技术:基于文本的图像检索(Text-Based Image Retrieval, TBIR)和基于内容的图像检索(Content-Based Image Retrieval, CBIR)^[3-6]。TBIR,需要人为地对图像进行文本标注,具有效率低、查找不准的缺点。CBIR,提取图像的浅层视觉特征或深层语义特征,结合相似度度量方法,从

收稿日期: 2022-10-13

修回日期: 2023-02-16

基金项目: 国家自然科学基金(青年基金)(61806160)

作者简介: 王彪(1996-),男,硕士研究生,通信作者,研究方向为计算机视觉;毋涛(1976-),男,教授,硕导,博士,研究方向为智能信息处理。

数据库中检索出最为相似的 top-k 张图片,这是目前国内研究的热点^[7]。徐佳等人^[8]提出一种基于全局和局部相位特征相融合的图像检索算法。王妙^[9]提出基于深度学习的印花织物图像检索系统设计,采用哈希算法的粗检索和卷积神经网络的细致检索相结合的分级检索,在提高精度的基础上进一步提高了速度。Xia 等人^[10]提出的 CNNH 通过对相似度矩阵进行分解,得到样本的哈希码,再利用卷积神经网络对哈希码进行拟合。Lai 等人^[11]提出 DNNH。该方法对网络结构做了针对性的设计:用部分连接取代全连接,引入分段量化函数。Zhang 等人^[12]提出的 DRSCN 用加权的汉明距离代替普通的汉明距离,这种方法可以提高计算距离的效率和精度,但是时间复杂度也会增加。Lin 等人^[13]在 AlexNet 的 F7 和 F8 之间加入一个全连接层,基于由粗到细的策略,利用学习到的类哈希二进制码和 F7 层特征,实现图片检索。该方法降低了特征匹配的计算量,加速了检索速度,但是不能保证哈希码相似的图像在语义上也相似,因此检索精度不高。何彬等人^[1]提出了一种基于 InceptionV3 迁移学习的面料图像特征提取算法。该算法在 Inception 中加入哈希编码层,并优化了损失函数,将模型输出的哈希编码作为面料图像的特征向量,采用分级检索的策略进行检索。该方法虽然在精度和时间上优于一般的检索算法,但是由于采用的暴力检索策略,在速度上仍有较大的提升空间。刘瑞昊等^[7]通过改进 LresNet50E-IR 网络结构,将其迁移学习到面料检索上,结合 faiss 进行面料图像检索可以取得很好的效果。但是算法采用的损失函数只能限制模型具有较好的分类效果,不能保证相似的面料图像的特征也相似。

该文主要针对目前面料图像检索领域存在的“找料慢”“找料难”等问题^[14],通过结合深度学习和 faiss 向量检索解决以上问题,为纺织企业提供高效的面料管理策略。

1 面料检索系统框架

1.1 检索系统模块

面料图像的检索是以纺织企业的实际需求为依据,利用卷积神经网络对面料图像进行特征学习、表示和匹配,从面料数据库中检索出相似面料图像。系统分为模型训练、构建面料特征数据库和面料检索3个模块。模型训练是对 VGG16^[15] 网络结构进行微调,同时优化损失函数,以企业提供的面料作为训练数据训练模型参数,以表达面料图像的特征;构建面料特征数据库是用训练好的模型提取面料特征,以向量的形式保存下来;面料检索是将待检索的面料和数据库中图像进行相似度匹配,排序输出检索结果。

1.2 检索整体流程

基于 CNN 的面料检索系统框架结构如图 1 所示。

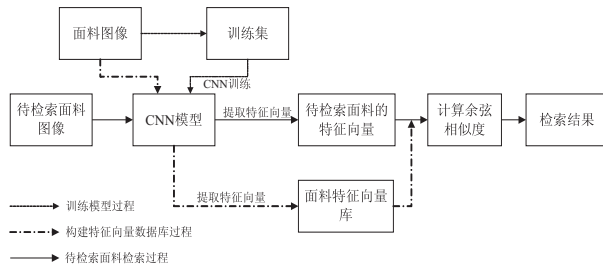


图1 面料检索整体框架流程

该系统首先以企业提供的面料为训练集对模型进行训练,得到提取面料特征的模型,对所有的面料图形进行特征提取,将特征向量保存在本地作为特征向量库;然后将待检索的面料送到检索系统中,提取其特征,以余弦相似度作为特征向量之间相似度的度量,将相似度最高的 top-k 检索结果返回,完成一次检索过程。

2 相关工作

2.1 准备数据集

企业提供真实的面料样本,通过图像采集设备获得每一张面料的图像,一共有 37 002 张,部分面料图像见图 2。根据不同纺织机机型,对面料进行批量的归类,一共分为 10 类,分别是 12E、14E、16E、18E、20E、22E、24E、26E、28E、36E。按照 8:2 的比例,将数据集划分为训练集和测试集,其大小分别是 29 602 和 7 400。

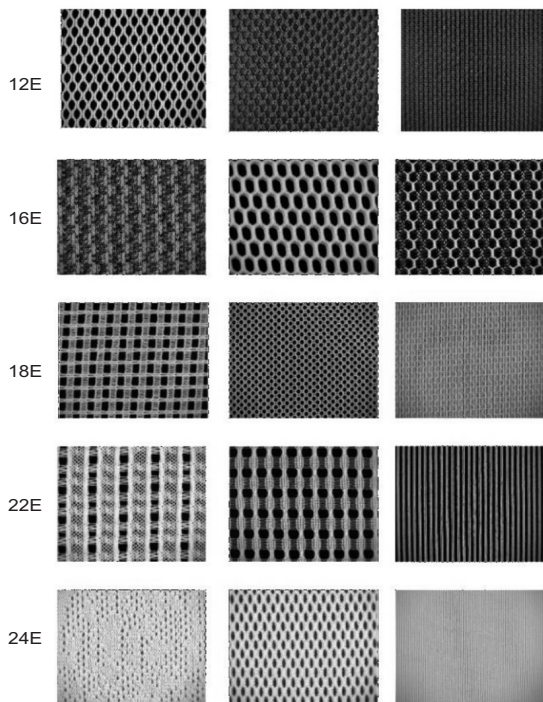


图2 部分面料

在模型训练阶段,考虑到检索情景的多样性,提高模型的泛化能力,使检索系统具有较强的鲁棒性,对训练集进行随机水平翻转、随机颜色抖动等数据增强操作。

2.2 模型选择

面料特征的提取是整个面料检索系统最核心的环节。优秀的模型提取出的特征向量可以很好的表示图像,也是后续进行特征匹配的关键,所以模型选择至关重要^[7]。

2.2.1 VGG16 网络结构

VGG 网络是 2014 年 ImageNet 大规模图像识别大赛的亚军,VGG16 是其中的一类模型,常用于图像分类、目标检测和图像分割等任务,其结构如图 3 所示。VGG16 一共有 16 层,包括卷积层、池化层和全连接层。

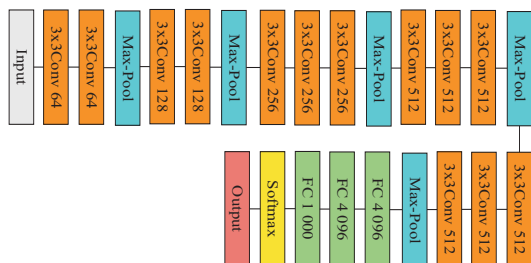


图 3 VGG16 网络结构

VGG16 在卷积层采用 3×3 的小卷积核,减少了参数量,同时加快了模型的训练速度。在每个卷积层之后使用 ReLU 激活函数对特征图进行非线性映射以提高模型的代表能力。

池化层的池化窗口大小为 2×2,步长为 2,在不减少特征图数量的前提下,缩小了特征图的尺寸,压缩了参数量。

网络最后是 3 层全连接层,融合了卷积层提取的局部特征,表达了输入图像的全局特征。

2.2.2 改进 VGG16

虽然使用 3×3 的卷积核减少了参数量,但是模型还是很复杂,有可能会出现过拟合的问题;在检索过程中计算两个面料的特征向量的相似度时,由于特征向量维度过高,带来较大的计算量;交叉熵损失函数是用来分类的损失函数,只能保证模型能够对图像进行正确的分类,不能让模型学习相似的面料图像的特征向量也相似。本面料检索系统以 VGG16 为基础网络,迁移学习其网络结构,对其做出以下改进:为了避免模型发生过拟合,提高模型的泛化能力,在卷积层之后,激活层之前加入 BN 层;为了减少检索过程中向量相似度计算的计算量,修改 classifier 部分的 layer3 的输出神经元数量为 128 个;为了使模型学习到相似面料图像的特征向量也相似,对损失函数进行优化;在原来分

类损失函数的基础上加入相似度损失函数 similarity。要计算相似度损失函数,首先要构建两个相似度矩阵 A 和 B ,然后再计算 A 和 B 对应位置元素差的平方和的均值。 A 和 B ,以及相似度损失函数的计算公式见式(1)、(3)、(4)。

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1j} \\ \vdots & \ddots & \vdots \\ a_{i1} & \cdots & a_{ij} \end{pmatrix} \quad (1)$$

式中, a_{ij} 表示第 i 个特征向量 x^i 和第 j 个特征向量 x^j 的余弦相似度,计算公式见式(2)。

$$a_{ij} = \frac{\sum_{k=1}^n x_k^i \times x_k^j}{\sqrt{\sum_{k=1}^n (x_k^i)^2} \times \sqrt{\sum_{k=1}^n (x_k^j)^2}} \quad (2)$$

式中, x_k^i 、 x_k^j 分别代表向量 x^i 和 x^j 的各个分量。

$$B = \begin{pmatrix} b_{11} & \cdots & b_{1j} \\ \vdots & \ddots & \vdots \\ b_{i1} & \cdots & b_{ij} \end{pmatrix} \quad (3)$$

式中, b_{ij} 仅有 0 和 1 两种取值。0 表示向量 x^i 和向量 x^j 不属于同一种类别,即不相似;1 表示两个特征向量属于同一种类别,即相似。

$$\text{similarity} = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n (a_{ij} - b_{ij})^2 \quad (4)$$

式中, n 表示特征向量的数目。

2.3 模型训练

准备数据集、选择模型和构建损失函数这些前期工作做好后,下一步是训练模型。训练参数设置见表 1。

表 1 模型训练参数

参数	数值
epoch	80
batch_size	64
learning_rate	0.001
optimizer	SGD

在迭代 80 个 epoch 后,训练集和测试集的损失和准确率都已基本收敛,将此时的网络保存到本地作为构建特征数据库和检索时提取面料特征的模型。

2.4 基于 faiss 的图像检索

faiss 是 Meta(原 Facebook)AI 团队为了解决海量稠密向量的检索问题,结合高效相似度检索和聚类方法,提出的一种开源的相似度检索库^[16]。例如,从给定的面料图像库中检索出与待检索的面料图像相似度最高的前 K 张图像,称为 K 近邻检索^[17]。本研究借助 faiss 向量检索工具,为每一张面料图像的特征向量绑定一个唯一的索引,在损失极小的精度的情况下,大大

提高了检索的速度。

3 对比实验

为证明文中方法的可行性,与多种图像检索算法进行对比。多特征融合方法结合了图像的形状、颜色和纹理 3 种特征,检索效果要优于单一的特征^[18]。Inception_HashOP^[1]通过在倒数第 2 层加入哈希编码层并优化损失函数,采取分级检索的策略进行图像检索。VGG19^[15]相比于 VGG16 在结构上没有太大的区别,只是比 VGG16 多了三层卷积层。

3.1 评价指标

要比较这几种算法的性能,首先要确定统一的评价指标。该文采用图像检索领域常用的平均查准率(mAP)和检索一张面料平均消耗的时间 T 作为算法性能的评估指标。

查准率(Precision)是指一张面料图像经过检索,检索结果列表中与该图像相似的数量在检索结果列表中所占的比例^[19]。Precision 的计算公式见式(5)。

$$\text{Precision} = \frac{TP}{TP + FP} \quad (5)$$

式中,TP 是指检索到的相关面料,FP 是指检索到的不相关面料。

平均查准率(Average Precision, AP)表示一张面料图像经过检索之后,检索结果列表中,每个召回率点上的查准率的均值。计算公式见式(6)。

$$AP = \frac{\sum_{i=1}^R P(r_i)}{R} \quad (6)$$

式中, r_i 是指第 i 召回点, R 表示检索结果中与检索面料相似的个数, $P(r_i)$ 表示第 i 个召回点的查准率。

平均查准率均值(mAP)是多张查询面料的平均查准率的均值^[20]。计算公式见式(7)。

$$\text{mAP} = \frac{\sum_{i=1}^n AP_i}{n} \quad (7)$$

式中, AP_i 表示第 i 个查询面料的平均查准率, n 表示查询面料的个数。

3.2 实验环境

模型训练是在服务器上进行。服务器 CPU(E5-2678 v3)运行内存 122 GB, GPU(RTX3090)显存为 24 GB,使用 Pytorch 深度学习框架。

3.3 实验过程

从测试集中随机抽取 100 张面料图像,其余作为构建特征数据库的面料图像。当输入一张要查询面料图像,经过改进的 vgg16 模型提取查询面料图像 128 维的特征向量,计算其与特征数据库中的每一个面料的特征向量的余弦相似度,按相似度从高到低排序输

出,然后根据输出结果计算平均查准率均值 mAP。具体流程见图 4。

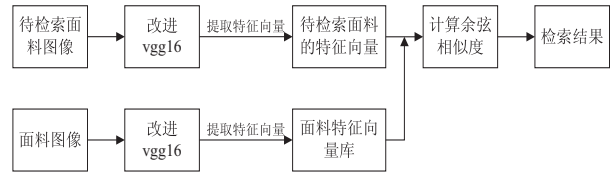


图4 实验流程

3.4 实验结果

表2展示了各种算法在 mAP 和检索时间上的对比。可以看出该文提出的方法无论是在平均查找率均值方面,还是在检索速度上都具有巨大的优势。其中,虽然 Inception_HashOP 和文中方法在 mAP 上相差不多,但是由于其在特征向量相似度计算时采取的是暴力遍历的方法,所以其检索速度远远低于文中方法;VGG19 由于提取的特征向量维度较高,所以检索速度最低。

表2 不同算法对比

方法	mAP	检索时间/(s/张)
特征融合	0.501	2.31
DLBHS	0.721	5.86
Inception_HashOP	0.879	5.93
VGG19	0.704	20.56
ours	0.892	0.012

综上,提出的研究方法在企业面料数据集上同时具备检索速度快和检索精度高的特点,具有良好的检索性能。

4 检索系统

4.1 系统框架设计

面料图像检索系统基于 web 技术,主要由四个模块组成,分别是前端模块、后端模块、算法模块和数据库模块。如图5所示。其中前端模块主要负责面料图像的上传、接受/处理请求和展示面料图像检索的结果,后端模块主要负责整个面料图像检索系统的图像数据的输入与输出,算法模块主要负责对查询的面料图像的检索,面料图像特征数据库保存在 MySQL 数据库中,数据库模块主要由 Redis 构成,主要负责存储面料图像及其相对应的 id,便于前端模块的检索结果的展示。检索流程为:第1步,用户通过前端页面将待检索的面料图像上传至前端;第2步,前端将面料图像传输至后端;第3步,由后端对面料图像进行相关处理,将其输入至算法模块中;第4步,算法模块将检索结果的面料图像 id 返回至后端模块;第5步,后端将算法模块返回的检索结果 id 输入数据库模块;第6步,数据库将图片路由返回给后端;第7步,后端将图片路由

和 id 返回给前端;第 8 步,前端模块接收后端模块返回的面料图像检索结果 id 和相对应的图片路由后,显示给用户。

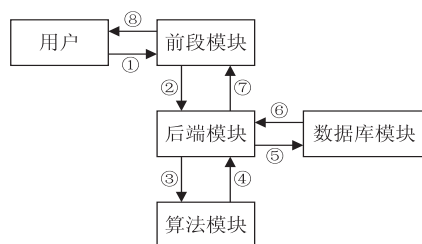


图 5 系统框架

4.2 系统实现

在面料检索的检索页面中,点击上传按钮上传要查询的面料图像,点击检索按钮,经过系统检索后,根据输入的检索数量 N ,系统的前端页面展示出 N 张最相似的面料的图像、id 以及和查询面料的相似度。检索结果如图 6 所示。

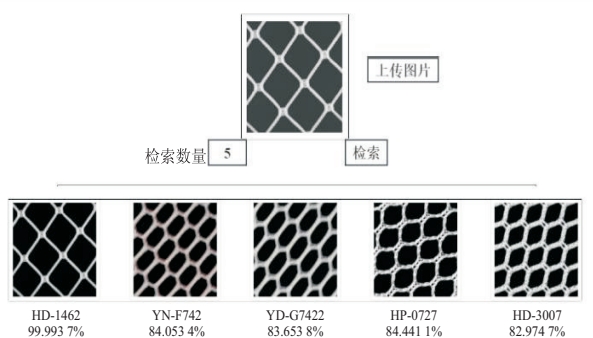


图 6 检索结果

5 结束语

该文提出一种基于卷积神经网络的面料检索系统,解决了企业面料检索速度慢、精度低的难题。通过微调 VGG16 网络结构,迁移学习到面料数据集上,利用 CNN 强大的表征能力,同时借助向量检索工具 faiss 进行特征向量检索,使得检索系统具有很好的性能。系统在面料图像上的 mAP 可达到 0.892,检索时间仅为 0.012 秒,均优于以往的算法。该方法存在的缺点是,模型训练时间耗时长,当面料数据集更新时需要重新训练模型。因此,接下来可在现研究基础上进一步优化网络,降低模型训练的时间,提高检索系统的实用性。

参考文献:

[1] 何 彬. 基于卷积神经网络的面料图像检索研究[D]. 上海: 东华大学, 2021.
[2] 刘嘉唯, 肖勇锋, 白小明, 等. 融合颜色与 LBP 纹理特征的布料色卡图像检索[J]. 软件导刊, 2016, 15(8): 173-176.

[3] 刘翔宇. 基于分块 sim-min-Hash 的近似图像检索[J]. 计算机应用与软件, 2019, 36(7): 259-263.
[4] 钟奇林, 左 劫, 孙频捷. 基于区域的 CBIR 图像检索方法研究进展[J]. 现代计算机, 2019(36): 64-68.
[5] 沈金科. 基于内容的图像检索方法研究与实践[D]. 武汉: 武汉工程大学, 2011.
[6] 董卫军, 周明全, 耿国华, 等. 基于内容的图像检索技术研究[J]. 计算机工程, 2005, 31(10): 162-163.
[7] 刘瑞昊, 于振中, 孙 强. 基于深度学习的织物图像检索系统研究[J]. 棉纺织技术, 2022, 50(5): 42-47.
[8] 徐 佳, 曲怀敬, 张志升, 等. 融合全局和局部相位特征的彩色纹理图像检索[J]. 计算机技术与发展, 2022, 32(9): 56-64.
[9] 王 妙. 基于深度学习的印花织物图像检索系统设计[D]. 西安: 西安工程大学, 2019.
[10] XIA R, PAN Y, LAI H, et al. Supervised hashing for image retrieval via image representation learning[C]//Twenty-eighth AAAI conference on artificial intelligence. Québec: AAAI, 2014: 2156-2162.
[11] LAI H, PAN Y, LIU Y, et al. Simultaneous feature learning and hash coding with deep neural networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Boston: IEEE, 2015: 3270-3278.
[12] ZHANG R, LIN L, ZHANG R, et al. Bit-scalable deep hashing with regularized similarity learning for image retrieval and person re-identification[J]. IEEE Transactions on Image Processing, 2015, 24(12): 4766-4779.
[13] LIN K, YANG H F, HSIAO J H, et al. Deep learning of binary hash codes for fast image retrieval[C]//Proceedings of the IEEE conference on computer vision and pattern recognition workshops. Boston: IEEE, 2015: 27-35.
[14] 于振中, 秦 岭. 基于深度学习的布料图像检索方法研究[J]. 科技与创新, 2021(21): 50-52.
[15] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv:1409:1556, 2014.
[16] SHI Y, JAIN A K. DocFace+: ID document to selfie matching[J]. IEEE Transactions on Biometrics, Behavior, and Identity Science, 2019, 1(1): 56-67.
[17] 戴琳琳, 阎志远, 景 辉. Insightface 结合 Faiss 的高并发人脸识别技术研究[J]. 铁路计算机应用, 2020, 29(10): 16-20.
[18] 计蕴容, 周韦润. 基于多特征融合的面料图像检索系统[J]. 计算机与数字工程, 2021, 49(7): 1460-1464.
[19] 李玉惠. 基于视觉描述子的图像检索与分类方法研究[D]. 新乡: 河南师范大学, 2016.
[20] 崔国影. 基于局部特征的显著区域检测及其在图像检索中的应用[D]. 合肥: 合肥工业大学, 2014.