

基于 Xception 改进的卷积神经网络服装分类算法^①



任永亮¹, 宋 田², 毋 涛¹

¹(西安工程大学 计算机科学学院, 西安 710048)

²(山东如意毛纺服装集团股份有限公司 公司调度中心, 济宁 272004)

通信作者: 宋 田, E-mail: songtian@chinaruyi.com

摘 要: 针对服装图像分类模型的参数量过大, 时间复杂度过高和服装分类准确度不高等问题, 提出了一种利用网络剪枝方法和网络稀疏约束, 减少卷积神经网络 Xception 中从卷积层到全连接层的冗余参数, 增加网络的稀疏性和随机性, 减轻过拟合现象, 在保证不影响精度的前提下尽可能降低模型的时间复杂度和计算复杂度。此外在卷积层引入了注意力机制 SE-Net 模块, 提升了服装图像分类的准确率。在 DeepFashion 数据集上的实验结果表明, 使用网络剪枝方法缩减的网络模型在空间复杂度上和时间复杂度上均有所降低, 服装图像分类准确率和运行效率与 VGG-16, ResNet-50 和 Xception 模型相比均有所提升, 使得模型对设备的要求更低, 深度卷积神经网络在移动端、嵌入式设备中使用成为可能, 在实际服装领域的电商平台的应用中有比较高的使用价值。

关键词: 服装分类; Xception 模型; 网络剪枝; 注意力机制; 卷积神经网络; 图像分类

引用格式: 任永亮, 宋田, 毋涛. 基于 Xception 改进的卷积神经网络服装分类算法. 计算机系统应用, 2022, 31(6): 381-387. <http://www.c-s-a.org.cn/1003-3254/8511.html>

Improved Clothing Classification Algorithm Based on Xception in Convolutional Neural Network

REN Yong-Liang¹, SONG Tian², WU Tao¹

¹(School of Computer Science, Xi'an Polytechnic University, Xi'an 710048, China)

²(Company Dispatch Center, Shandong Ruyi Woolen Garment Group Co. Ltd., Jining 272004, China)

Abstract: The clothing image classification model has many parameters, high time complexity, and low accuracy of clothing classification. In response, a network pruning method and network sparsity constraints were proposed to reduce the redundant parameters from the convolutional layer to the full connection layer in the Xception of the convolutional neural network (CNN), increase the sparsity and randomness of the network, reduce the over-fitting phenomenon, and lower the time complexity and computational complexity of the model as much as possible without affecting the accuracy. In addition, the attention mechanism SE-Net module is introduced into the convolution layer to improve the accuracy of clothing image classification. The experimental results on the DeepFashion dataset show that the network model reduced by the network pruning method has brought down both the spatial complexity and the time complexity. Compared with VGG-16, ResNet-50, and Xception models, the proposed model, with higher clothing image classification accuracy and operation efficiency, lowers the requirements on equipment and paves the way for the application of the deep CNN on mobile terminals and embedded devices. It has a high value in the application of e-commerce platforms in real-world clothing.

Key words: clothing classification; Xception model; network pruning; attention mechanism; convolutional neural network (CNN); image classification

① 基金项目: 陕西省科技成果转化与推广计划 (2019CGXNG-018).

收稿时间: 2021-08-30; 修改时间: 2021-09-26; 采用时间: 2021-10-09; csa 在线出版时间: 2022-05-26

随着电子商务的崛起,人们选择网购的方式购买商品变得越来越普遍,据数据统计,服装行业在电子商务的市场销售额占比已经高达近30%,中国消费者服装通过线上购买的方式占比近80%。目前,购物网站比如淘宝,拼多多主要通过人工输入文本关键字的形式搜索商品链接,但是,当用户需要的商品周边信息不明确时,这种检索的方式很难达到消费者的真实需求^[1]。而需要快速地找到同类型的服装的前提是需要先进行服装的识别和精确分类,识别不同的服装类别在不断增长的时尚业中具有很高的潜在价值,可以监测和分类供商业使用的服装。由于商业市场上服装产品繁多,人工识别服装既费时又费力,并且对同一幅图片受主观性的影响不同的人可能产生不一致的理解,进而影响检索结果。因此亟需要一种准确而又高效的服装分类方法。

针对服装图像的分类问题,研究人员提出了很多的服装分类算法和网络模型,在基于图像内容的服装分类方法领域方面,Willimon等^[2]提出了一种称为L-C-S-H (low level, characteristics, selection mask, high level)的中层服装分类方法,该方法将问题分解为高(H)、低(L)和多个中层(特征(C)、选择掩模(S))层,并产生“局部”解决方案来解决全局分类问题。Huo等^[3]还提出了一种基于服装局部和整体特征进行融合的分类方法对自然环境场景生活中的各种服装造型进行融合分类,实验证明了此方法的有效性。Wang等^[4]提出了一种知识引导的时尚网络来解决视觉时尚分析的问题,例如时尚地标定位和服装品类分类。建议的时装模型利用了该领域的高级人类知识,在大规模时尚数据集上的实验结果证明时尚语法网络的优越性能。另外基于深度学习的卷积神经网络,研究者对卷积神经网络通过不同方式做了优化改进,目的是能够在服装分类的性能,准确率或者速度上有所提升。张振焕等^[1]提出通过调整批量归一化层、激活函数层、卷积层的排列顺序和池化层+卷积层的排列顺序来改变参差的优化卷积神经网络提高服装的精度。王雅静等^[5]提出了基于BP神经网络算法的女性服装方面的款式分类方法,但是没有具体说明服装的分类种类和对比性不足等问题,而且仅限于女装。刘童童^[6]提出了通过计算每一个卷积层的卷积核容量和覆盖率来评估网络分类效率的方式,使得深层卷积神经网络可以缩短训练时间,同时提高了资源利用率。高妍等^[7]将HyperNet网

络、空间变换网络(STN),卷积神经网络R-FCN算法相融合的方式进行对服装图像的识别和分类,提高了分类准确率,减少了训练时间,但是存在分类属性不够精确,复杂背景下对服装图像的检测影响比较大。杨天祺等^[8]通过批量标准归一化,改进了卷积层组织结构、增加了冗余分类器的方式大大提高了服装的图像归一化分类的准确率和速度。严安等^[9]为了解决复杂背景下的服装图片的分类,在AlexNet模型的结构上增加一个全连接层,提出了一种九层的卷积神经网络,但是在数据集上的选择的每类服装的数量不均匀,缺乏说服力。陈巧红等^[10]由于使用多尺度深度可分离卷积来提升模型特征数据信息的丰富度,多尺度SE-Xception模型在复杂的场景下可以稳定较高的准确率,提高了鲁棒性,但是所提模型的结构更加复杂,增加了一定的参数量和计算量。上述针对服装图像分类,从多个方向对卷积神经网络进行了改进,如增加卷积层,批量归一化,调整网络结构顺序等,但是由于网络深度的不断增加,参数量过大,运行时间长等问题依然存在。

本文针对卷积神经网络在处理图像分类时参数过多,训练时间过长等问题提出了利用一种通道剪枝方法^[11]来加速深度可分离卷积的轻量级卷积神经网络Xception,并嵌入SE-Net模块^[12]提升服装图像分类准确率,使模型更加稳定。

1 Xception 剪枝服装图像分类

卷积神经网络(convolutional neural network, CNN)是一种比较常用的人工神经网络,常被应用在图像识别、图像分类和图像分割等领域。它是由卷积层、池化层和全连接层是相互堆叠在一起形成的CNN架构^[13],如图1所示。

CNN的标准来自LeCun等的作品《Gradient-based learning applied to document recognition》^[14],他们首次在图像识别和分类领域成功地使用卷积神经网络,1998年,LeCun创建了一种最早的卷积神经网络架构,称为LeNet5架构。2012年,Krizhevsky等^[15]提出的AlexNet网络模型在ILSVRC(image net large scale visual recognition challenge)网络图像分类竞争中获得第一名,错误率比第二名低约10%,AlexNet在卷积数据层种采用ReLU作为数据激励函数,使用了随机数据的失活和动态数据属性增强技术,缩短了模型的收敛时间。2013年,Zeiler等^[16]提出的ZF-Net在2013

年 ILSVRC 图像分类算法的比赛中取得第一名, 是在 AlexNet 的基础上做了将第一个卷积核大小调整为 7×7 , 步长减半的处理, 取得较好的分类效果. 2015 年, Simonyan 等^[17] 随后提出了一种使用较小卷积核可以代替大卷积核的一种 VGG-Net 分类模型, 同时大大增加了模型层的深度; Szegedy 等^[18] 提出了采用不同类型大小卷积核可以提升分类准确度的 GoogLeNet 模型. 2016 年 He 等^[19] 关于 CNN 的退化问题提出了 Res-Net 模型. 2017 年, Chollet^[20] 在 CVPR 2017 会议中提出的 Xception 模型在服装图像分类上取得了优于其他模型的效果.

1.1 Xception 模型

本文研究的框架是基于 Xception 模型构建, 该框架由输入层, 中层和输出层 3 个主要组成部分共同组成, 具体结构如图 2 所示. 它是基于 InceptionV3 模型的基础

上做出的改进. Xception 包括 36 个卷积层 14 个模块, 采用深度可分离卷积来进行替换 Inception V3 中卷积模块, 同时加入 residual learning 结构, 除了第一和最后一个卷积模块, 中间每一个模块都有残差模块连接. Xception 网络与 Inception 网络比较相似, 但有不同之处. 前者首先执行通道空间卷积操作, 每次卷积完成操作之后它都加入了分批归一化和激活函数 ReLU. 中间模块采用残差直连结构, 可以减少网络复杂性, 在数据量比较大的服装图像数据集上, 训练效果比 Inception V3 更好.

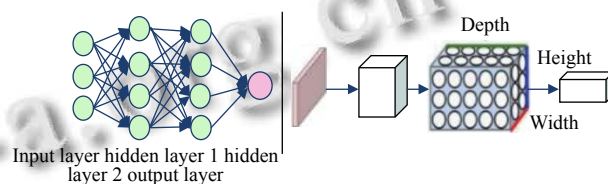


图 1 一个规则的三层神经网络与一个规则的 CNN

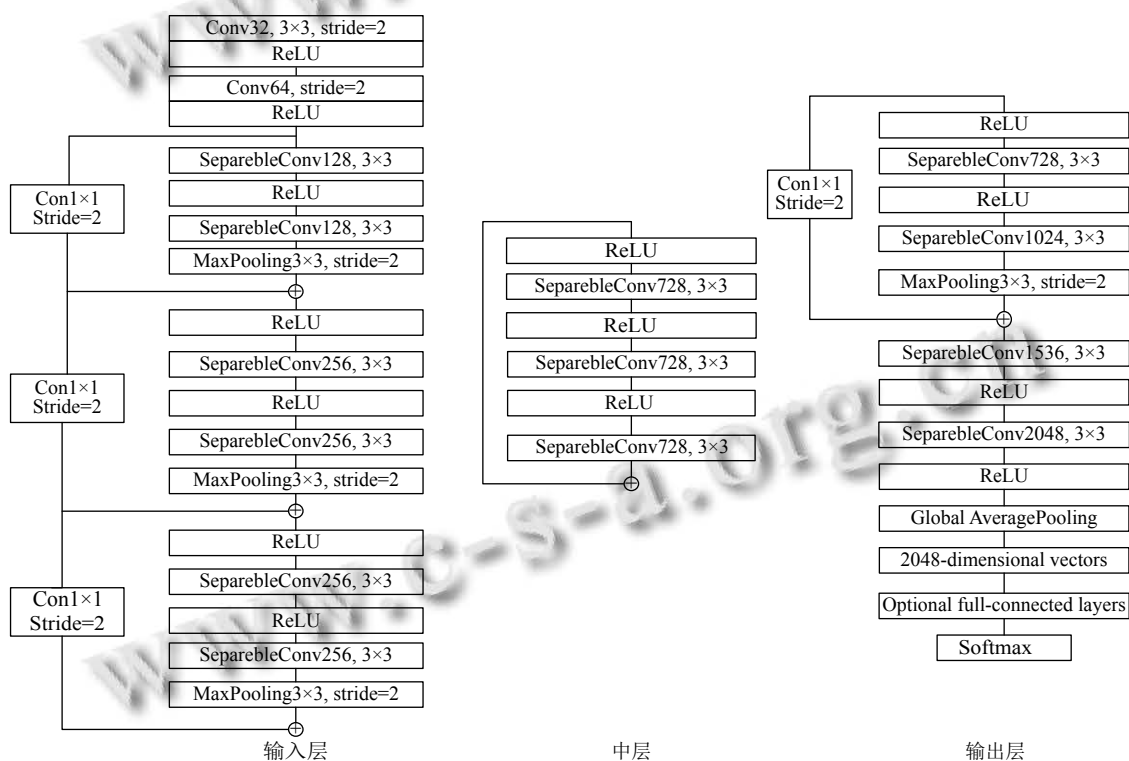


图 2 Xception 模型的主要网络模块示意图

1.2 SE-Net 模型

SE-Net (squeeze-and-excitation networks) 模型是由 Momenta 公司的研发团队的高级工程师 Hu 等^[12] 提出, 该网络模块在 2017 年 ILSVRC 技术挑战赛国际图像处理分类比赛中获得了第一名. SE-Net 模块不是一个全新、完整的 CNN 网络模型, 是一种可以直接

嵌入其他各种网络模型中的子模块. SE 模块的结构图如图 3 所示, 结构操作流程通过使用 Squeeze, Excitation 和 Reweight 三个基本操作流程来重标定前面得到的特征. 给定一个新的输入特征 x , 特征通道数用 C_1 表示, 通过一系列卷积等一般函数变换后得到一个特征通道数为 C_2 的输入特征.

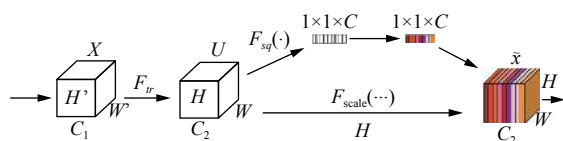


图3 SE模块结构图

(1) Squeeze 信道特征向量压缩统计操作, 为了有效解决信道特征向量依赖性的问题, 首先将一个全局空间信息特征压缩到一维信道描述符中, 通过使用全局平均池来直接生成一维信道特征统计数据公式实现, 得到包含 c 个实数的一维信道特征向量, Squeeze 通过式 (1) 计算:

$$\mathbb{F}_{sq}(u_c) = \frac{1}{W_{sq} \times H_{sq}} \sum_{i=1}^{W_{sq}} \sum_{j=1}^{H_{sq}} u_c(i, j) \quad (1)$$

其中, $\mathbb{F}_{sq}(\cdot)$ 为 Squeeze 函数; u_c 为输入特征中第 c 通道的特征曲线图; W_{sq} 和 H_{sq} 分别表示输入特征图的宽和高; $u_{(i,j)}$ 为输入特征图中的点在坐标位置 (i, j) 处的取值。

(2) Excitation 特征激励, 利用在 Squeeze 操作中所聚集的特征信息, 进行第二个操作, 该操作的目的是完全自动捕获通道间的依赖。首先它必须考虑是灵活的和非完全互斥的关系, 为了满足这些标准我们采用两个全连接层和 Sigmoid 激活函数, 公式如下。

$$\mathbb{F}_{ex}(z, W) = \sigma[W_2 \delta(W_1 z)] \quad (2)$$

其中, z 为 Squeeze 操作获得的全局描述, δ 表示 ReLU 函数, 保证输出为正, W_1, W_2 为两个全连接层, 其中 $W_1 \in R^{r \times c}$, $W_2 \in R^{c \times r}$ 其中 r 为缩放参数, 主要用于减轻网络的计算复杂度和参数量, $\sigma(\cdot)$ 为 Sigmoid 函数。

(3) Reweight 特征重标定, 将 Excitation 的输出权重看作是经过特征选择后的每个特征通道的重要性, 然后通过乘法逐通道加权到先前的特征上, 完成在通道维度上的对原始特征的重标定, 公式如下。

$$\mathbb{F}_{scale}(u_c, s_c) = u_c \times s_c \quad (3)$$

其中, $\mathbb{F}_{scale}(\cdot)$ 表示 Reweight 函数, s_c 为第 c 个特征图的权重值。

1.3 网络模型剪枝

随着卷积神经网络深度的不断加深, 导致了大量的浮点数量的增加, 全连接层也产生了大量的训练浮点参数^[11]。近几年, 研究者们针对这样的问题, 在 CNN 加速上做出改进, CNN 加速工作分为 3 类: 优化实现 (如 FFT^[21])、量化 (如 BinaryNet^[22]) 和结构化简化将

CNN 转换为精简的^[23]。结构化简化主要包括: 张量因子分解、稀疏连接、通道剪枝。张量分解法是将一个卷积层分解成了几个比较有效的卷积层。但是 feature map 宽度 (channel number) 不能减少, 这使得分解现代网络 (如 GoogLeNet, ResNet, Xception) 的 1×1 卷积层结构变得困难。这种类型的方法还会增加额外的计算开销。稀疏连接使神经元或通道之间的连接失效。虽然它可以达到很高的理论加速比, 但稀疏卷积层的“不规则”形状, 这种实现是欠佳的。相比之下, 通道剪枝直接减少了特征图的宽度。因为不需要特别的实现, 它在 CPU 和 GPU 上的表现都是比较高效的。

由于传统的剪枝方法需要长时间的再训练程序, 通过多次的随机试验选择通道, 在深度网络上需要较长时间评估每一次实验, 这限制了在深的模型和大的数据集上的实验。因此, 为了正确的选择最具代表性的渠道, 本文采用一种迭代两步算法的方法, 使用基于线性 LASSO 回归的方法公式为:

$$\mathbb{F}(x) = a_1 \varphi_1(X) + a_2 \varphi_2(X) + \dots + a_m \varphi_m(X) \quad (4)$$

其中, $\varphi_1(X)$ 是事先已经选定的一组线性空间无关的函数, φ_k 是待定向量系数 $k=1, 2, \dots, m, (m < n)$, 拟合准则指的是使 y ($i=1, 2, \dots, n$) 与 $\mathbb{F}(x_i)$ 之间的距离 φ_i 的平方和最小, 称为最小二乘拟合准则。

1.4 改进的 Xception 网络模型

前面已经详细介绍本次研究所使用的模型, 本文改进的 Xception 剪枝模型与原始 Xception 模型大体一致, 整体结构如表 1 所示, 卷积层的前两层标准卷积不变, Conv_3–Conv_7 层进行信道剪枝以减少信道数量, 为了保证实验准确率的稳定和运行时间上的平衡性, 在 Conv_3–Conv_7 层根据实验结果灵活选取是否需要剪枝操作, 使用基于线性 LASSO 回归的方法提取最小且最有代表性的通道, 删除冗余变量通道, 另外, 使用线性最小二乘重构提取剩余通道的数据输出, 两步交替进行。与原来相比, 改变了卷积神经网络层的内部结构, 在 Conv_3–Conv_7, Conv_8 和 Conv_9 层中分别加入 SE-Net 模块, 增加了对卷积后的特征图大小包括宽高以及通道数的特征压缩, 特征激励和重标定的操作使得可以得到维度相同的特征图, 保留了原全局均值池化层 GPA_10 与全连接层 FC_11, 通过实验验证和分析了所改进模型的优势和可行性, 如表 1 所示。

表 1 为改进的 Xception 网络模型的整体框架, 从第 3 卷积层开始加入剪枝操作, 即基于 LASSO 正则化

剔除冗余卷积核与其对应的特征图,然后重构剩余网络,剪枝操作可选表示根据具体实验结果在准确度和效率上选取一个平衡点,在第8,9层上嵌入了 SE-Net 模型用以增强有用通道,减弱无用通道,第10和第11层分别保留了原有的全局均值池化和全连接层。

表1 改进的 Xception 网络模型的整体框架

卷积层级 序号	参差链接	卷积操作	输出尺寸
Conv_3	Conv1×1, stride = 2	深度可分离卷积128	55×55×128
		剪枝操作	
Conv_4	Conv1×1, stride = 2	最大池3×3, stride=2, padding=1	28×28×256
		深度可分离卷积256	
Conv_5	Conv1×1, stride = 2	剪枝操作(可选)	14×14×728
		最大池3×3, stride=2, padding=1	
Conv_6_x	直连	深度可分离卷积728	14×14×728
		剪枝操作(可选)	
Conv_7	Conv1×1, stride = 2	深度可分离卷积728	7×7×1024
		深度可分离卷积1024	
Conv_8	—	剪枝操作(可选)	7×7×1024
		最大池3×3, stride=2, padding=1	
Conv_9	—	深度可分离卷积1536	7×7×1024
		嵌入SE-Net模块	
GAP_10	—	深度可分离卷积2048	1×1×2048
		嵌入SE-Net模块	
FC_11	—	全局均值池化	1×1×7
		全连接层	

2 实验结果与分析

2.1 数据集

本实验选取公开 DeepFashion 数据集^[24], 包含的服装图片背景一般为纯色, 复杂度较低, 每个服装包含从不同角度进行拍摄的图像, 又利用特定模型的训练, 部分服装进行了服装位置的精确裁剪, 数据集一共包含了 80 多万张服装图像, 图片标记了种类, 属性等关键点的位置坐标. 本文选取其中的一部分作为本实验的数据集, 选取 10 种熟数量相差不大的服装类别, 每种选取 2 000 张左右, 一共选取 2 万多张, 其中选取 80% 作为训练集, 20% 作为验证集, 部分图片如图 4 所示.

2.2 实验准备

实验环境配置及工具如下: 在带有 NVIDIA CUDA 8.0 的 GPU 机器上安装 MXNet (0.11.0) 深度学习库,

采用笔记本 Windows 10 系统, 12 GB 运行内存, PyCharm 2020 开发工具, GPU 为 Intel(R) HD Graphics 520, Python 版本为 3.6. 由于 DeepFashion 数据集的尺寸大小不同, 将数据集进行归一化预处理, 统一尺寸大小设置为 256×256, 做图像的归一化处理可以加快模型的收敛速度.



图4 DeepFashion 服装示例图

2.3 实验结果分析

为了验证所改进的 Xception 模型的有效性, 将 DeepFashion 数据集先在 Xception 模型上作训练和验证, 记录相关的性能数据, 然后 Conv_3–Conv_9 层保持不变, 在 Conv_8 和 Conv_9 层中加入 SE-Net 模块, 进行再训练, 最后在基础上在 Conv_3–Conv_9 层进行信道剪枝操作, 由于剪枝操作在准确率和运行时间上具有互斥性, 在 Conv_3–Conv_7 层根据实验结果调整剪枝操作的次数和卷积层, 进行反复训练对比, 最终实验证明在分别在 Conv_3、Conv_5 和 Conv_7 层上做剪枝操作进行通道的选取可以在两者之间达到一个比较好的平衡. 分别记录了 3 种模型的 Top3 准确率, 消耗时间和消耗内存大小如表 2 所示. 3 次训练验证均在相同条件下, 从实验结果数据中可以看出, 相对于原始的 Xception 模型, 加入 SE-Net 模块的 Xception 在时间和内存的消耗上边均有小幅度所增加, 但是 Top3 准确率提升, 3.12 个百分点, 而之后加入的 LASSO 回归的方法提取最有代表性的通道, 删除冗余通道和最小二乘重构剩余通道, 使得模型在精度只降低了 0.21 的情

况下,大幅度减少了训练时间和内存空间的占用,缩短了将近3倍.这证明了所提出方法的可行性和有效性.

表2 不同改进的模型实验数据性能对比

网络模型	Top3准确率 (%)	训练时间约 (h)	消耗内存 (MB)
Xception	74.55	10.5	509
SE-Net+Xception	77.67	11.2	591
剪枝+SE-Net+Xception	77.46	3.6	205

为了进一步证明所提出模型的优势和有效性,本实验也和 VGG-6, ResNet-50 和 Xception 模型在最终的平均分类准确率做了对比,在相同的数据集和实验条件下,几种不同方法模型的平均服装分类准确率如表3所示, VGG-16 模型的平均分类准确率最低为 70.23%, 然后 ResNet-50, Xception 依次升高. 每种均选取均匀种类的服装,在不同的由数据可知,提出的通过剪枝通道的 SE-Net+Xception 模型平均分类精确度最高. 此外,实验数据集按大小分成了3种,分别和其他3种模型进行实验对比,如图5所示,由图可直观地看出,随着数据集的增大,每种模型的服装分类准确率均有不同程度的提升,相比于其他3种模型,所提出改进的 Xception 模型在不同大小的数据集下分类准确率均高于其他模型,而且在运行效率上均比其他模型用时少. 实验从不同的维度证明了本文所提出模型的优越性,是一种性能比较好的服装图像分类网络模型.

表3 不同网络模型平均分类准确率对比

网络模型	R (%)
VGG-16	70.23
ResNet-50	73.65
Xception	74.81
剪枝+SE-Net+Xception	79.88

3 结束语

本次研究提出了一种改进的 Xception 模型算法,在 Xception 的基础上,加入了 SE-Net 模块,这是一种新的架构单元.它在数据集上实现了比较好的性能. SE-Net 模块诱发的特征重要性有助于其他相关领域的压缩,如网络剪枝等.为了解决参数过大和内存消耗问题,在此基础上,利用 LASSO 回归的方法提取最有代表性的通道,删除冗余通道和最小二乘重构剩余通道的方法对 Xception 模型卷积层进行了剪枝操作,实验证明,这样做在很小精度损失的条件下提高了训练速

度,缩小了内存的占用空间.最后实验证明了,本文所提出模型和其他模型相比,大幅度缩短了运行时间,服装的平均分类准确率也得到了进一步的提升.但是也有一些不足,比如如何对复杂背景下服装的精确分类问题,以及服装图像分类准确率在参数减少的前提下如何仍可以更好地保持精确度的问题等.

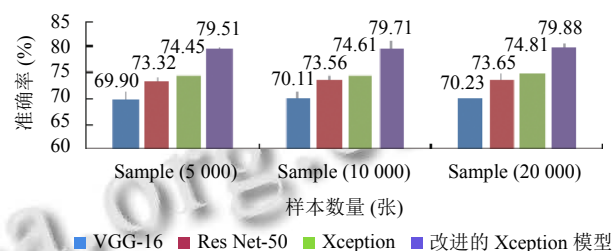


图5 不同数量的数据集在不同模型上的准确率对比

参考文献

- 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. 计算机工程与科学, 2018, 40(2): 354-360. [doi: 10.3969/j.issn.1007-130X.2018.02.023]
- Willimon B, Walker I, Birchfield S. A new approach to clothing classification using mid-level layers. 2013 IEEE International Conference on Robotics and Automation. Karlsruhe: IEEE, 2013. 4271-4278.
- Huo P, Wang YH, Liu QJ. A part-based and feature fusion method for clothing classification. Proceedings of the 17th Pacific Rim Conference on Multimedia. Xi'an: Springer, 2016. 231-241.
- Wang WG, Wang WG, Xu YL, et al. Attentive fashion grammar network for fashion landmark detection and clothing category classification. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 4271-4280.
- 王雅静, 宋丹. 基于 BP 神经网络的女性服装款式分类技术研究. 轻工科技, 2020, 36(4): 107-109.
- 刘童童. 基于卷积神经网络的服装图像精细分类 [硕士学位论文]. 济南: 山东师范大学, 2019.
- 高妍, 王宝珠, 郭志涛, 等. 改进 HSR-FCN 的服装图像识别分类算法研究. 计算机工程与应用, 2019, 55(16): 144-149. [doi: 10.3778/j.issn.1002-8331.1804-0274]
- 杨天祺, 黄双喜. 改进卷积神经网络在分类与推荐中的实例应用. 计算机应用研究, 2018, 35(4): 974-977, 1045.
- 严安, 王姚芬. 基于 CNN 改进模型的服装分类方法. 自动化技术与应用, 2019, 38(11): 114-119. [doi: 10.3969/j.issn.1003-7241.2019.11.026]

- 10 陈巧红, 陈翊, 李文书, 等. 多尺度 SE-Xception 服装图像分类. 浙江大学学报 (工学版), 2020, 54(9): 1727–1735.
- 11 He YH, Zhang XY, Sun J. Channel pruning for accelerating very deep neural networks. Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV). Venice: IEEE, 2017. 1398–1406.
- 12 Hu J, Shen L, Sun G. Squeeze-and-excitation networks. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 7132–7141.
- 13 Madulid JPA, Mayol PE. Clothing classification using the convolutional neural network inception model. Proceedings of the 2019 2nd International Conference on Information Science and Systems. Tokyo: ACM, 2019. 3–7.
- 14 LeCun Y, Bottou L, Bengio Y. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 1998, 86(11): 2278–2324.
- 15 Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe: Curran Associates Inc., 2012. 1097–1105.
- 16 Zeiler MD, Fergus R. Stochastic pooling for regularization of deep convolutional neural networks. Proceedings of the 1st International Conference on Learning Representations. Scottsdale: ICLR, 2013.
- 17 Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv: 1409.1556, 2015.
- 18 Szegedy C, Liu W, Jia YQ, *et al.* Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston: IEEE, 2015. 1–9.
- 19 He KM, Zhang XY, Ren SQ, *et al.* Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016. 770–778.
- 20 Chollet F. Xception: Deep learning with depthwise separable convolutions. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017. 1800–1807.
- 21 Vasilache N, Johnson J, Mathieu M, *et al.* Fast convolutional nets with fbfft: A GPU performance evaluation. arXiv: 1412.7580, 2014.
- 22 Courbariaux M, Hubara I, Soudry D, *et al.* Binarized neural networks: Training deep neural networks with weights and activations constrained to +1 or -1. arXiv: 1602.02830, 2016.
- 23 Jaderberg M, Vedaldi A, Zisserman A. Speeding up convolutional neural networks with low rank expansions. arXiv:1405.3866, 2014.
- 24 Liu ZW, Luo P, Qiu S, *et al.* DeepFashion: Powering robust clothes recognition and retrieval with rich annotations. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016. 1096–1104.