

Flower image Classification Algorithm Based on Improved Residual Network^{*}

PEI Xiaofang^{1,2}, ZHANG Yang³

(1. Binjiang College, Nanjing University of Information Science & Technology, Nanjing 210044, China; 2. Jiangsu Collaborative Innovation Center of Atmospheric Environment and Equipment Technology, Nanjing University of Information Science & Technology, Nanjing 210044, China; 3. College of Electronic and Information Engineering, Nanjing University of Information Science & Technology, Nanjing 210044, China)

Abstract: Traditional flower classification usually obtains a significant area by image segmentation and manually extracting or choosing suitable features. This method generally lacks in high precision, low cost and strong generalization ability. Aiming at the application of deep learning in fine-grained image classification, an end-to-end method of flower classification based on residual network is proposed. Firstly, the method uses ResNet18 as the basis model. Secondly, this paper applies the idea of full convolution structure to the network model. The convolution layer replaces the fully connected layer of ResNet18 to optimize the network model (Res18_conv). Finally, the residual attention network is incorporated into the optimized ResNet18 (Res18_conv_a). The residual attention network studies a mixed domain, focusing on the spatial domain as well as the channel domain. After the entire network extracts all flower features, the flower is classified by the softmax function. This paper selected two data sets of Oxford17 flowers and Oxford102 flowers to verify the proposed image classification method based on improved residual network. The method of this paper has achieved the desired effect compared with the previous flower image classification method. For Oxford17 data set, the accuracy of Res18_conv_a is 99.26%. For Oxford102 data set, the accuracy of Res18_conv_a is 99.02%. This paper proposes an improved residual structure based on attention mechanism. Compared with the predecessor's floral image classification method, this method can extract the characteristics of key information more effectively and suppress the information of the interference area. This method has a significant effect on the classification of flower images, which is suitable for fine-grained image classification.

Key words: image classification; flower recognition; residual network; full convolution structure; attention mechanism
EEACC: 0200; 7260D doi: 10.3969/j.issn.1005-9490.2020.03.041

基于改进残差网络的花卉图像分类算法^{*}

裴晓芳^{1,2}, 张 杨³

(1. 南京信息工程大学滨江学院, 江苏 南京 210044; 2. 南京信息工程大学江苏省大气环境与装备技术协同创新中心, 江苏 南京 210044; 3. 南京信息工程大学电子与信息工程学院, 江苏 南京 210044)

摘 要: 传统的花卉图像分类都是基于人工手动选择单一特征或者多特征融合再分类, 这种方法普遍存在精度低、成本高、泛化能力弱等缺陷。针对目前深度学习在细粒度图像分类中的应用, 提出一种基于残差网络、实现端到端的花卉图像分类方法。首先以 ResNet18 为基础模型; 其次将全卷积结构的思想应用于网络模型中, 将 ResNet18 的全连接层替换成卷积层以优化网络模型; 最后在优化后的 ResNet18 中融入混合域注意力机制, 由 Softmax 层进行分类。本文选取了 Oxford17flowers 和 Oxford102flowers 两个花卉图像数据集做对比试验, 与前人的花卉图像分类方法对比, 本文的方法取得了理想的效果, 在 Oxford17 和 Oxford102 上分别取得了 99.26% 以及 99.02% 的正确率。提出了一种基于注意力的残差结构改进方法, 相较于前人的花卉图像分类方法, 该方法能够更有效地提取关键信息的特征, 抑制干扰区域的信息, 对花卉图像分类具有显著性效果, 适用于细粒度图像分类。

关键词: 图像分类; 花卉识别; 残差网络; 全卷积结构; 注意力机制

中图分类号: TP391.4

文献标识码: A

文章编号: 1005-9490(2020)03-0698-07

随着经济的发展和生活质量的提高, 花卉已经成为了人们生活不可或缺的一部分。花卉的需求量增大, 自然对花卉种植业提出了新的要求。然而对

于传统的花卉种植业而言, 种植人员的专业性是阻碍花卉种植业发展的关键因素。若大量引入专业的种植人员, 势必会增大种植成本, 而且专业的种植人

项目来源: 南京信息工程大学滨江学院课题项目 (2019bjyng006)

收稿日期: 2019-09-02 修改日期: 2019-09-27

员短缺也是不得不考虑的问题。因此,智能化的种植系统就变得十分必要,而智能化的种植系统中,对花卉的智能识别分类自然是最基本的要求。

花卉图像分类^[1]属于精细图像分类,由于不同花卉种类间的相似性较高,同一类别存在差异性较大问题,给花卉分类任务带来了很大的困难。传统的花卉分类方法大都是基于手动选择单一特征或者多特征融合再用分类器进行分类。Saitoh 等人^[2-3]提出了一种自动分割的方法,利用该方法找到像素点之间最小的代价函数路径;Hong 等人^[4-5]提出了利用花卉的颜色和形状特征来提取出花卉区域;Nilsback 等人^[6]提出了通过花卉的颜色、纹理等特征构建了一个模型,最后通过 KNN 算法加权得到了分类结果;Angelova 等人^[7]提出了一种提取 4 类尺度的方向梯度直方图 (HOG) 特征,使用 LLC (locality constrained linear coding) 对特征编码;谢晓东等^[8]提出了一种将花卉的显著性特征与图割算法 (Grabcut) 相结合的方法,同时利用多特征融合,最后采用支持向量机 (SVM) 分类;Fernando 等人^[9]提出通过提取花卉的 HUB 颜色信息和尺度不变特征变换 (SIFT) 特征,采用逻辑回归算法计算权重值,最后采用 SVM 分类;武慧琼等人^[10]提出一种基于三支决策的分类方法,因其获取事物更多信息后再进行延迟决策,成为了一种有效的分类模型。

近年来,大数据以及云计算等技术的飞跃发展导致了卷积神经网络的崛起。卷积神经网络^[11]采用端到端的策略,实现了图像特征的自动学习和分类。在近几年的 ImageNet 比赛中也产生了很多优秀的网络模型,从一开始的 AlexNet^[12]、GoogleNet^[13]到后来的 VGG^[14]、ResNet^[15]、Inception^[16]等,都验证了卷积神经网络的可靠性和有效性。在花卉识别领域,Xia 等人^[17]采用 Inception-v3 结构实现花卉识别,取得了良好的效果;Gogul 等人^[18]先对图片进行预处理提取出花卉区域再用卷积神经网络提取花卉特征分类;Siraj 等人^[19]在花卉分类方面对卷积神经网络及逻辑回归算法进行了实验对比,得出了卷积神经网络优于逻辑回归算法的结论;Liu 等人^[20]提出的方法和 Gogul 等人提出的方法相似,都是对图片进行处理提取显著区域再采用卷积神经网络;Azizpour 等人^[21]对影响深度网络性能的几种因素进行分析比较,调优之后分类的正确率有了显著性的提高,在 Oxford102flowers 数据集上达到了 91.3% 的准确率;韩猛^[22]通过实验对比了 3 种网络模型在花卉分类上的分类效果,分别为 AlexNet、VGG16、VGG19,得出了网络越深分类效果越好的结论;尹红等人^[23]提出选择性卷积特征融合的方

法,在花卉分类工作中取得了较高的精确度。

结合前人的工作,本文采用 ResNet18 作为基础模型,在优化网络结构的基础上融入注意力机制,提出了一种新的花卉分类方法。实验表明,相较于前人的分类方法,本文提出的花卉分类方法具有收敛性好、精度高、泛化能力强等特性。

1 残差网络

1.1 基本网络结构

卷积神经网络早在 20 世纪 80 年代就已被提出,但由于当时计算资源的匮乏,所以卷积神经网络并没有快速发展。随着时间的推移,数值计算设备的改进以及深度学习理论的提出,深度卷积神经网络成为了研究热潮并且快速地发展,被大量地应用到计算机视觉、自然语言处理等领域^[24]。科研人员也提出了更多的深度学习网络,但网络的加深也导致了网络退化等问题的出现。为此,He 等人^[15]在 2015 年提出了残差网络的概念,该网络一经问世,便在 2016 年的 ImageNet 竞赛中取得了冠军。实验表明,残差网络更容易优化,收敛更快。由于本文的数据量有限,实验对比了常见的 18 层,34 层,50 层的残差网络模型,结果表明 ResNet18 在此花卉图像分类上效果最好,所以选用 ResNet18。其结构如图 1 所示,其中 Relu 为激活函数。

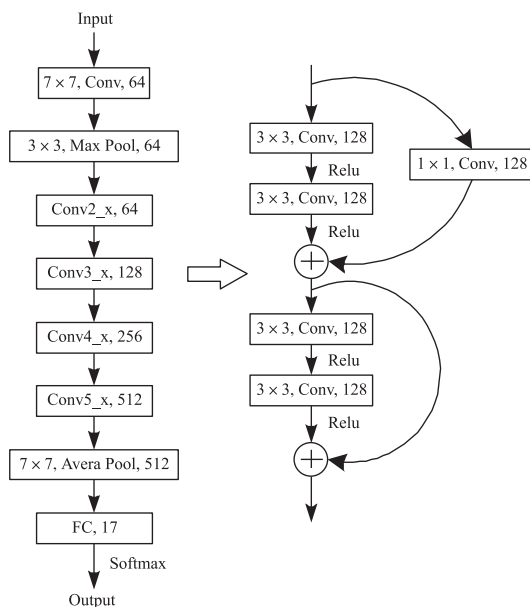


图1 ResNet18 结构图

1.2 残差块原理

深层网络中直接让一些层去拟合恒等映射函数 $H(x)=x$ 是不可取的,这就是深层网络难以训练的原因。残差的设计理念就在于着重学习相同主体之外的微变化部分,残差网络的核心就是“shortcut connection”(捷径连接),结构如图 2 所示。捷径连

接最大的优点在于加强了特征的传递,但是网络的参数和计算的复杂度并没有增加,而且该结构可以移植到其他网络中。假设 $H(\mathbf{x})$ 是神经网络的输出,即 $H(\mathbf{x}) = F(\mathbf{x}) + \mathbf{x}$,其中 $\mathbf{x}$ 为输入, $F(\mathbf{x})$ 为残差映射。残差学习通过残差块转换为学习一个残差函数 $F(\mathbf{x}) = H(\mathbf{x}) - \mathbf{x}$,假设 $F(\mathbf{x}) = 0$,就构成了一个恒等映射 $H(\mathbf{x}) = \mathbf{x}$,相对于学习 $H(\mathbf{x})$, $F(\mathbf{x})$ 更容易学习。Shortcut Connection 有两种方式:

shortcut 同等维度映射

$$\begin{aligned} \mathbf{y} &= F(\mathbf{x}, \mathbf{W}_i) + \mathbf{x} \\ F &= \mathbf{W}_2 \sigma(\mathbf{W}_1 \mathbf{x}) \end{aligned} \quad (1)$$

shortcut 不同维度映射

$$\begin{aligned} \mathbf{y} &= F(\mathbf{x}, \mathbf{W}_i) + \mathbf{W}_s \mathbf{x} \\ F &= \mathbf{W}_2 \sigma(\mathbf{W}_1 \mathbf{x}) \end{aligned} \quad (2)$$

式中: $\mathbf{y}$ 是输出向量; $F(\mathbf{x}, \mathbf{W}_i)$ 是需要学习的残差映射, $\mathbf{W}_i$ 是神经网络层的权重参数; σ 为激活函数; $\mathbf{W}_s$ 是一个线性映射,来满足维度变化的需求; $\mathbf{W}_1$ 、 $\mathbf{W}_2$ 分别是神经网络 1、2 层的权重参数。

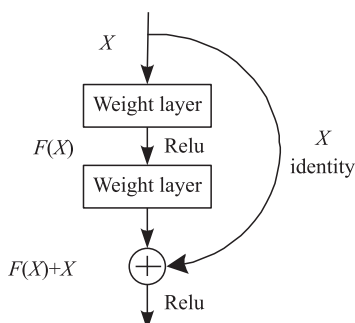


图 2 捷径连接

2 改进型残差网络

首先,在原有的 ResNet18 的基础上,将全卷积结构的思想融入到此网络中,用卷积层代替全连接层优化网络。由于深度学习方法需要通过大量的数据支持,过小的数据集训练出的模型容易过拟合;且模型越简单,小数据集的分类效果越好,所以实验对比了 Resnet18、Resnet34 和 Resnet50,实验结果表明 Oxford17 和 Oxford102 在 Resnet18 上的分类结果最好。用卷积层代替全连接层可以减轻过拟合的现象。其次,在优化网络之后加入了注意力机制,该注意力机制为 6 层,注意力机制的特性使得以注意力机制构建的模型能够找到输入中的关键部分,特别是在细粒度图像分类方面,能够找出具有区分度的部分,适合解决弱监督的学习问题。综上,改进型的残差网络可以定义为 24 层。具体步骤是:

(1) 网络首先是一个大小为 7×7 的卷积层以及一个大小为 3×3 的最大池化层,对大小为 $224 \times 224 \times$

3 的图片进行浅层特征的提取。

(2) 将步骤(1)得到的结果输入到组合块,每个组合块的结构类似,其中包含残差块和注意力模块。先由残差块对特征进行提取,每个残差块都有 4 个大小为 3×3 的卷积层,每两个卷积层都有捷径连接,其中第 1 个捷径上有一个大小为 1×1 的卷积层,卷积核数量同该组 3×3 卷积层的卷积核数量相同,第 2 个捷径则为直接捷径。经过第 1 个残差块后,图像被表达为一个 $56 \times 56 \times 64$ 的三维矩阵。

(3) 将残差块的输出作为注意力模块的输入,在分裂为主干分支和掩膜分支之前,网络先经过一个残差块进行预处理,再进入到不同的分支中。在主干分支中,包含了两个残差块,而在掩膜分支中,最大池化层和残差块结合使用。两个分支处理结束后,将两个分支特征进行结合。将得到的结合特征再送入一个残差块,最终得到注意力模块的输出还是一个 $56 \times 56 \times 64$ 的三维矩阵。

(4) 每个组合块中的残差块结构不同,不同之处在于卷积层的卷积核数量,其分别为 64、128 和 256。在 3 个组合块之后,再次连接了一个残差块,这里的残差块结构和组合块中的残差块结构相同,只是此残差块每个卷积层的卷积核数量为 512,特征图尺寸为 1×1 ,图像最终被表达为 $1 \times 1 \times 17$ 或者 $1 \times 1 \times 102$ 的三维矩阵。在得到输出后,网络将输出送入失活概率为 0.5 的 Dropout 层。接着通过大小为 1×1 的卷积层进行分类。

2.1 全卷积结构

全卷积网络^[25] (Fully Convolutional Network, FCN) 顾名思义,就是将传统的卷积神经网络中的全连接层转化成卷积层,故称全卷积。全卷积网络的提出最初是应用于图像语义分割领域,受到此网络结构的启发,本文使用卷积层代替原有的全连接层,使网络模型具有更好的收敛效果。

传统的卷积神经网络通常是若干层的卷积层加上全连接层,将卷积层产生的特征图强制映射成一个固定长度的向量。卷积层区别于全连接层的地方在于卷积中的神经元只与输入数据的部分区域连接,并且参数共享,有效减少大量参数以及过拟合的现象。卷积层和全连接层可以直接相互转化,因为它们的计算方式以及函数形式一致。将全连接层换成卷积层后,可以接收任意尺度的图像,通过反卷积层对卷积层提取的特征进行处理,使其与输入时的图像尺度一致,同时保留了原始图像的空间信息。所以本文将 ResNet18 最后的全连接层替换成卷积再分类,以此优化网络结构。

2.2 注意力机制

注意力机制^[26] 源于对人类视觉的研究,该机制

是研究人员根据人类视觉信息处理的方法研究出的一类机制模型。注意力机制的基本原理就是对关键信息加以关注,抑制干扰信息。注意力机制可以分为两大类:一是硬注意力(hard attention),二是软注意力(soft attention)。硬注意力在图像领域里可以很好地筛选出重要区域作为输入数据,是一个随机的动态的预测过程。但是硬注意力的一大缺陷是不可微的,通常需要强化学习来训练,也就造成了通用性差的缺点。软注意力更加关注空间或者通道信息,而且软注意力最重要的一点是可微,这样就可以直接采用端到端的学习方式了,在网络前向传播和后向反馈的时候学习注意力的权重。

本文所采用的残差注意力机制^[27-28]是一种混合域的注意力机制,即注意力同时关注空间域以及通道域。残差注意力机制主要由多层残差注意力模块堆叠而成,残差注意模块(Residual Attention Module)是一种可堆叠的注意力结构,与残差网络中的残差块相似,可以很好地嵌入到网络模型中达到易学习易优化的效果。其他利用注意力机制的网络都是需要单独一个分支进行训练,而残差注意模块在网络前向训练时就可以提取注意力模型,使得网络学习变得更简单。本文所用注意力模块结构图如图3所示。

残差注意力模块包含了两个分支:掩膜分支和

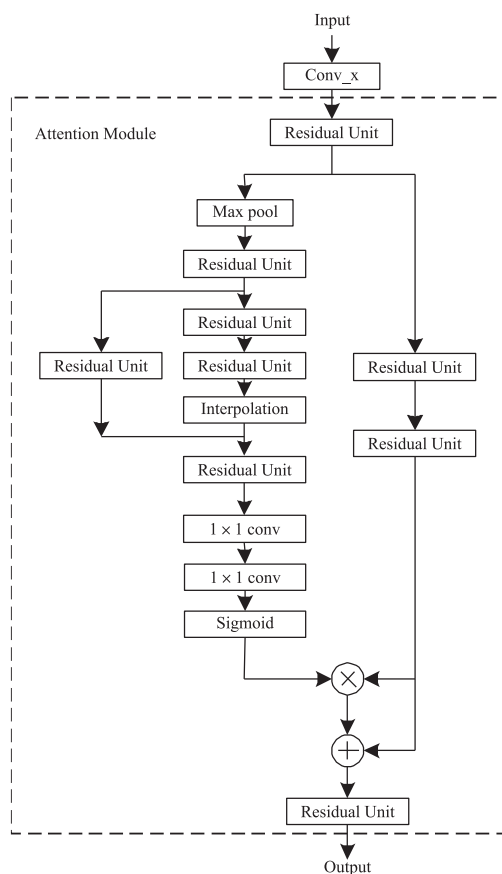


图3 注意力模块结构图

主干分支。其中掩膜分支采用了一种自底向上和自顶向下相结合的采样结构,在此结构中降采样过程获取了特征图的全局高维特征。接着,结构中的下采样过程将提取出来的全局高维特征的尺寸放大到与原始输入一样大,最终通过一个sigmoid激活函数将输出归一化到0到1之间。残差注意力网络的最终输出是掩膜分支和主干分支的点乘,但由于掩膜分支的输出在0到1之间,这必然导致特征图的输出变弱,特别是在经过几次模块的堆叠之后,最终输出的特征图的值会变得很小。为此,在这里借鉴残差网络的思路,采用恒等映射的思想,增强了最终的输出。因此残差注意力机制的最终输出为:

$$H_{i,c}(x) = (1 + M_{i,c}(x)) * F_{i,c}(x) \quad (3)$$

式中: $F_{i,c}(x)$ 表示深卷积网络产生的特征,即主干分支的输出。 $M_{i,c}(x)$ 为掩膜分支的输出,取值[0, 1]区间内。 i 取值于所有的空间位置上, $c \in \{1, 2, \dots, C\}$ 为通道索引。

3 实验及结果分析

3.1 实验环境及参数设置

本文实验采用的计算机环境是Linux系统,CPU E5-2670, GeForce GTX 1080 Ti 显卡, 8 GB 显存, 64G 内存, 软件框架是基于TensorFlow的Keras深度学习框架, 编程语言为Python。

在参数设置方面,选择实现简单且计算高效的优化器Adam,学习率设置为0.000 1, Batch_size为32,损失函数选择Categorical_crossentropy, Dropout设为0.5,迭代轮数设为150轮。

3.2 实验数据

目前网上公开的花卉数据集较少,所以本文采用由英国牛津大学公开的花卉数据,该数据集是由Visual Geometry Group采集制作,选取英国当地常见的花卉种类,该数据集包括Oxford17和Oxford102。Oxford17包括17类花卉,每类80张图片,总共1 360张图片。Oxford102包括102类花卉,每类由40张~258张图片组成,总共8 189张图片。示例如图4所示。

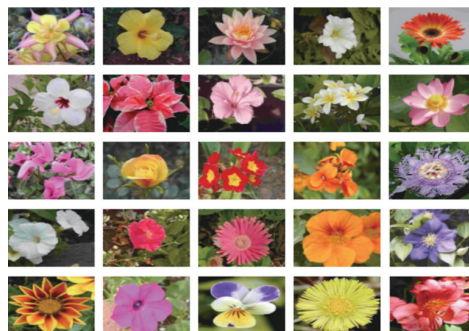


图4 Oxford数据集

数据增强是扩充数据常用的一种方法,可以增加数据量和噪声数据,有效减少过拟合提升网络模型的鲁棒性和泛化能力。数据增强的方式包括旋转、裁剪、翻转、缩放、尺度变换等。数据集 Oxford17 和 Oxford102 的图像数据量较少,所以需要数据增强。为了达到良好的增强效果,考虑到花卉图像的实际应用场景,本文采用了旋转、翻转、缩放、平移 4 种增强方式,有效地扩充了训练数据。

3.3 结果分析

实验 1 卷积层替换全连接层。全连接层参数量大,且容易造成过拟合现象,鉴于本文数据集小,所以用卷积层代替全连接层。实验结果如图 5 及图 6 所示,分别为 Oxford17 和 Oxford102 的测试结果。Res18 代表 Resnet18 的测试曲线图,Res18_conv 代表用卷积层替换全连接层的优化网络的测试曲线图。从上述图中可以看出 Res18_conv 曲线比 Res18 曲线平滑,说明用卷积层代替全连接层能使网络收敛的更快,效果更好。表 1 对比了两种网络结构的准确率,从上述表格可以看出 Res18_conv 在 Oxford17 上的准确率和 Res18 的准确率一致,在 Oxford102 上提升了 0.06%。实验结果表明,用卷积层代替全连接层可以加快网络的收敛速度,减少曲线震荡幅度,使得模型效果更优。

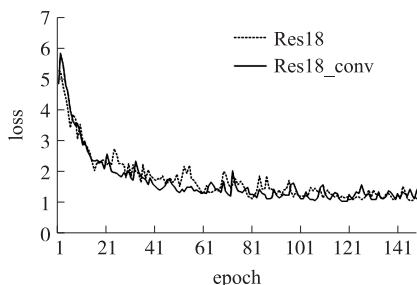


图 5 Oxford17 对比结果

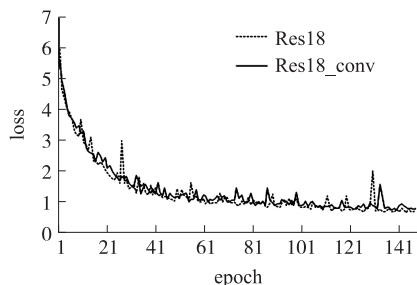


图 6 Oxford102 对比结果

实验 2 融入注意力机制。在实验 1 的基础上加入注意力机制,关注混合域特征,有效提取更多的重要特征。实验结果如图 7 及图 8 所示,分别为 Oxford17 和 Oxford102 的测试结果。Res18_conv 代表优化网络结构后的测试曲线图,Res18_conv_a 代表在优化过的残差网络上加入注意力机制的测试曲

线图。从图中对比结果可以看出,Res18_conv_a 的 loss 值更小,这就表示 Res18_conv_a 这条曲线的效果更好。如表 1 所示,Res18_conv_a 模型的正确率相较于 Res18_conv 在 Oxford17 和 Oxford02 上分别提升了 2.1% 和 0.79%。实验结果表明,加入注意力机制能够更好地提取出重要信息的特征,从而更好地提高分类任务的精度。

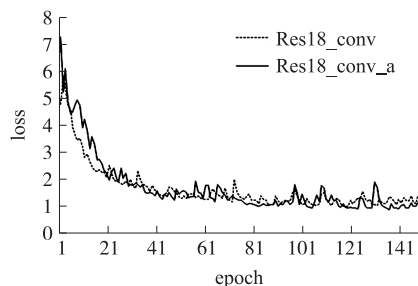


图 7 Oxford17 对比结果

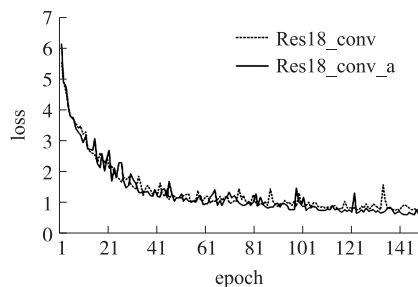


图 8 Oxford102 对比结果

表 1 不同改进方法的花卉图像分类准确率

方法	Oxford17	Oxford102
Res18	97.16	98.17
Res18_conv	97.16	98.23
Res18_conv_a	99.26	99.02

本文以 ResNet18 作为基准,为了对本文提出的方法进行验证和分析,本文选取了文献[6-9]的分类结果作为对比,对比结果如表 2 和表 3 所示。文献[6-9]都是基于传统的学习算法对花卉图像进行特征提取,如颜色、形状等特征,再将提取的各种特征融合达到分类的目的。从上述表格可以看出,采用传统的方法进行分类虽然准确率有所提升,但相比于深度卷积神经网络还是有所欠缺,这也足以证明了深度卷积神经网络的优越性。

表 2 不同分类方法的 Oxford17 分类准确率

方法	Oxford17
Xie(2014) ^[8]	93.14
Fernando(2012) ^[9]	93
Han(2017) ^[22]	95.1
Res18_conv_a	99.26

表3 不同分类方法的 Oxford102 分类准确率

方法	Oxford102	方法	Oxford102
Nilsback(2006) ^[6]	76.3	Azizpour(2015) ^[21]	91.3
Angelova(2013) ^[7]	80.66	Han(2017) ^[22]	96.5
Xie(2014) ^[8]	79.1	Yin(2019) ^[23]	98.96
Liu(2016) ^[20]	84.02	Res18_conv_a	99.02

同时,也与文献[20-23]基于深度学习方法的分
类结果进行了对比,对比结果如表2和表3所示。
Liu等人先对花卉图像进行显著性检测,再采用神经
网络提取特征,但学习的特征是全连接特征,包含背
景噪声从而影响了分类结果。Azizpour等人针对影
响模型性能的因素在多个数据集上做分类实验,对
这些影响因素优化过后,在花卉图像分类上取得了良
好的效果。Han提出3种网络模型进行花卉图像分
类,其中VGG19取得了最好的分类结果,结果可见表
2和表3,表明了网络深度对特征提取的影响,但网
络深度的增加也造成了浅层特征的不完全利用,所
以影响分类效果。Yin等人选择浅层特征和深层特
征相融合的方法有效的提高了花卉图像分类的精
确度。上述文献方法都在花卉图像分类上取得了理
想的效果,但通过表2和表3的结果对比可以看出,
本文的方法优于其他方法,具有更好的分类效果。

4 结论

本文提出用ResNet18为基础模型进行花卉图
像分类。再用卷积层替换全连接层优化模型,优化
以后的Res18_conv在Oxford17和Oxford102数据
集上都取得了良好的精确度,验证了该方法的分类
结果优于传统的图像分类方法。最后在Res18_conv
的基础上融入残差注意力网络,从实验结果的表1
来看,Res18_conv_a相比于Res18_conv的分类结
果,在Oxford17和Oxford102分别提高了2.1%和
0.79%。所以融入残差注意力网络的Res18_conv_a
能有效提取显著区域特征,抑制干扰信息。实验结
果如表2和表3表明,Res18_conv_a相较于传统的
分类方法以及深度学习方法,Res18_conv_a在花卉
图像分类上具有一定的有效性和可行性,证明了该
方法适用于细粒度图像分类。

参考文献:

- [1] Nilsback Maria-Elena, Zisserman Andrew. Automated Flower Classification over a Large Number of Classes[C]//Proceedings of the sixth Indian Conference on Computer Vision, Graphics & Image Processing. Bhubaneswar, Indian; 2008; 722-729.
- [2] Saitoh T, Aoki K, Kaneko T. Automatic Recognition of Blooming Flowers[C]//Proceedings of the 17th International Conference on Pattern Recognition. Cambridge, UK; 2004; 1051-1055.
- [3] Saitoh T, Kaneko T. Automatic Recognition of Wild Flowers[C]//Proceedings 15th International Conference on Pattern Recognition. Barcelona, Spain; 2000; 1051-1055.
- [4] Hong Anxiang, Chen Gang, Li Junli, et al. A Flower Image Retrieval Method Based on ROI Feature[J]. Journal of Zhejiang University Science, 2004(7): 16-24.
- [5] Hong Anxiang, Chi Zheru, Chen Gang, et al. Region-of-Interest Based Flower Images Retrieval[C]//Proceedings of the International Conference on Acoustics, Speech, and Signal Processing. HongKong, China; 2003; 1520-1524.
- [6] Nilsback M E, Zisserman A. A visual Vocabulary for Flower Classification[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2006; 1447-1454.
- [7] Angelova Anelia, Zhu Shenghuo. Efficient Object Detection and Segmentation for Fine-Grained Recognition[C]//Proceedings of 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, OR, USA; 2013; 811-818.
- [8] 谢晓东. 面向花卉图像的精细图像分类研究[D]. 厦门: 厦门大学, 2014.
- [9] Fernando Basura, Fromont Elisa, Muselet Damien, et al. Discriminative Feature Fusion for Image Classification[C]//Proceedings of Conference on Computer Vision and Pattern Recognition. Providence, RI, USA; 2012; 1063-1069.
- [10] 武慧琼, 张素兰, 张继福, 等. 一种基于三支决策的花卉图像分类[J]. 小型微型计算机系统. 2019, 40(7): 1558-1563.
- [11] 姜芑旭, 傅洪亮, 陶华伟, 等. 一种基于卷积神经网络特征表征的语音情感识别方法[J]. 电子器件, 2019, 42(4): 998-1001.
- [12] Krizhevsky Alex, Sutskever I, Hinton G. ImageNet Classification with Deep Convolutional Neural Networks[C]//Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, Nevada, United States; 2012; 1097-1105.
- [13] Szegedy Christian, Liu Wei, Jia Yangqing, et al. Going Deeper with Convolutions[C]//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA; 2015; 1063-1069.
- [14] Simonyan Karen, Zisserman Andrew. Very Deep Convolutional Networks for Large-Scale Image Recognition[EB/OL]. <https://arxiv.org/pdf/1409.1556.pdf>, 2015-04.
- [15] He Kaiming, Zhang Xiangyu, Ren Shaoqing, et al. Deep Residual Learning for Image Recognition[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA; 2016; 1063-1069.
- [16] Szegedy Christian, Vanhoucke Vincent, Ioffe Sergey, et al. Rethinking the Inception Architecture for Computer Vision[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA; 2016; 1063-1069.
- [17] Xia Xiaoling, Xu Cui, Nan Bing. Inception-v3 for Flower Classification[C]//Proceedings of the 2nd International Conference on Image, Vision and Computing. Chengdu, China; 2017; 783-787.
- [18] Gogul I, Kumar V Sathiesh. Flower Species Recognition System Using Convolution Neural Networks and Transfer Learning[C]//

- Proceedings of the 4th International Conference on Signal Processing, Communication and Networking. Chennai, India; 2017: 1-6.
- [19] Siraj Fadzilah, Salahuddin Muhammad Ashrap, Yusof Shahrul Azmi Mohd. Digital Image Classification for Malaysian Blooming Flower [C]//Proceedings of 2010 Second International Conference on Computational Intelligence, Modelling and Simulation. Tuban, Indonesia; 2010: 2166-8531.
- [20] Liu Yuanyuan, Tang Fan, Zhou Dengwen, et al. Flower Classification via Convolutional Neural Network [C]//Proceedings of 2016 IEEE International Conference on Functional-Structural Plant Growth Modeling, Simulation, Visualization and Applications. Qingdao, China; 2016: 110-116.
- [21] Azizpour Hossein, Razavian Ali Sharif, Sullivan Josephine, et al. From Generic to Specific Deep Representations for Visual Recognition [C]//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Boston, MA, USA; 2015: 2160-7516.
- [22] 韩猛. 基于深度学习的花卉图像种类识别 [D]. 山西: 山西农业大学, 2017.
- [23] 尹红, 符祥, 曾接贤, 等. 选择性卷积特征融合的花卉图像分类 [J]. 中国图象图形学报, 2019, 24(5): 762-772.
- [24] 李彦冬, 郝宗波, 雷航. 卷积神经网络研究综述 [J]. 计算机应用, 2016, 36(9): 2508-2515, 2565.
- [25] Long Jonathan, Shelhamer Evan, Darrell Trevor. Fully Convolutional Networks for Semantic Segmentation [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 39(4): 640-651.
- [26] Cho Kyunghyun, Courville Aaron, Bengio Yoshua. Describing Multimedia Content Using Attention-Based Encoder-Decoder Networks [J]. IEEE Transactions on Multimedia, 2015, 17(11): 1875-1886.
- [27] Raffel Colin, Ellis Daniel P W. Feed-Forward Networks with Attention Can Solve Some Long-Term Memory Problems [EB/OL]. <https://arxiv.org/abs/1512.08756v4>, 2016-09.
- [28] Wang Fei, Jiang Mengqing, Qian Chen, et al. Residual Attention Network for Image Classification [EB/OL]. <https://arxiv.org/abs/1704.06904>, 2017-04.



裴晓芳 (1978-), 女, 硕士, 副教授, 主要研究方向为信号处理与应用, xiaofangpei@163.com;



张 杨 (1995-), 女, 硕士研究生在读, 主要研究方向为图像处理、深度学习, 1604906231@qq.com。