

[文章编号] 1003—4684(2020)05-0062-03

# 基于改进 Inception-Resnet v2 图像分类算法

舒 军, 李振亚

(湖北工业大学电气与电子工程学院, 湖北 武汉 430068)

[摘 要] 根据工业生产要求,需要对不同等级的产品进行分类,Inception-ResNet v2 算法模型对其分类准确率相对较低,且检测速度慢,不能够满足工业现场实时性检测分类的需要。针对这些问题提出基于 Inception-ResNet v2 改进的算法。该方法在网络的 Inception-ResNet v2 结构中采用密集串联的方式,加强原始图像的特征传递,提取更多表面特征;又对网络的卷积方式进行改进,采用深度可分离卷积,同时降低特征图数量,用以降低网络的运算量从而提升整体网络的训练速度。实验结果表明,改进后的 GoogLeNet 算法准确率提高 7%,运算量降低 19%。

[关键词] 工业图片; Inception-ResNet v2; 深度可分离卷积

[中图分类号] TP751.1 [文献标识码] A

随着人工智能的发展,基于深度学习的图像分类算法有着广泛的应用前景和重要的研究价值,近年来更是成为了热点领域,无论是工业应用还是学术研究都已经取得许多有价值的成果。对于表面特征复杂的物体,提高分类准确率,同时加快检测速度仍是一项艰巨挑战<sup>[1]</sup>。

目前基于深度学习的图像分类算法主要为 VGG, ResNet, GoogLeNet 等。Inception-ResNet v2 为 GoogLeNet 后续改进版本,其将 ResNet 中提出的“残差”结构融入到 Inception 结构模块中去,能够使网络更快地收敛<sup>[2]</sup>,但该方法识别准确率不能够达到工业生产的要求,检测分类速度不能适应实时性的需求。为了解决这些问题,需要对 Inception-ResNet v2 网络进行相应的改进<sup>[3]</sup>。

## 1 GoogLeNet 算法

过去经典深度学习网络模型(如 AlexNet, VGG, ResNet 等)通常存在着收敛速度慢,训练参数过多,容易出现梯度弥散等问题难以优化模型<sup>[4]</sup>。为解决这些问题,GoogLeNet 模型提出 Inception 结构,是一种同时具有稀疏性和高性能的网络结构。Inception v1 将  $1 \times 1, 3 \times 3, 5 \times 5$ , 池化操作( $3 \times 3$ )的卷积堆叠在一起,同时分别在  $3 \times 3, 5 \times 5$ , 池化层之前增加  $1 \times 1$  卷积核进行降维<sup>[5]</sup>。这种 Inception 结构一方面增加网络的宽度,也同时增加了网络对不同尺度输入的适应性。Inception v2 突出改进将  $5 \times 5$  卷积核改为两个  $3 \times 3$  卷积核,在保证相同效

果的前提下,使用卷积分解能够减少参数量,加快计算速度<sup>[6]</sup>。Inception v3 提出分解思想,将  $n \times n$  卷积核分为两个  $n \times 1, 1 \times n$  使得网络进一步加深同时增加网络非线性。Inception v4 在网络中加入“残差”结构使其性能能进一步提升。Inception-Resnet v1 网络设计主要用来与 Inception v3 模型性能进行比较,其减少了子网络的网络计算量,获得与 Inception v3 相近的性能<sup>[7]</sup>。

Inception-ResNet v2 在 Inception v4 的基础上进行更加精巧的设计,为了使残差计算可行,卷积后的输入输出必须具有相同的维度<sup>[8]</sup>,因此在卷积之后加入  $1 \times 1$  卷积进行维度的匹配,便于相加。为了避免深层的网络对模型稳定性的影响,对网络中的“残差”模块的输出进行缩放,使整个训练过程更加稳定高效<sup>[9]</sup>。

## 2 Inception-ResNet v2 算法的改进

为了能够更加精细地提取原始输入图像的特征,在 Inception 结构模块中采用更加紧密的连接结构<sup>[10]</sup>,由于密集连接块特性,特征复用率增高,网络能够更全面学习原始数据。但密集块的增加也相应地加大了计算量,影响计算效率。为解决此问题,对特征图数量进行减半,同时采用深度可分离卷积。

整个 Inception-Resnet v2 模型由不同作用的 Inception 模块构成,图 1 为改进后的 Inception-Reduction 模块,作用为改变输入数据宽度与高度。

深度可分离卷积将一个卷积过程分两步进行卷

积;Depthwise 过程和 Pointwise 过程<sup>[7]</sup>。对于一个标准卷积过程,假设输入有  $N \times H \times W \times C$  的输入,同时有  $K$  个  $3 \times 3$  的卷积,若设置  $\text{pad}=1, \text{stride}=1$ ,那么标准卷积输出为  $N \times H \times W \times C$ 。对于可分离卷积,在 Depthwise 阶段将输入的  $N \times H \times W \times C$  分为  $C$  组,然后针对每一组进行  $3 \times 3$  卷积,提取每个通道的空间特征;在 Pointwise 阶段对输入的  $N \times H \times W \times C$  做  $K$  个普通  $1 \times 1$  卷积,提取图片每个点的特征。

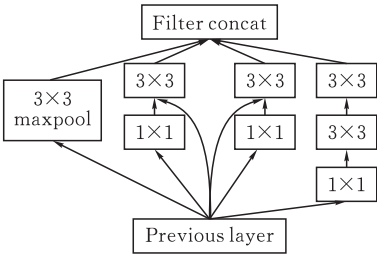


图 1 Inception-Reduction 模块

相比于普通卷积过程,相同的输入,同样得到相同特征图深度可分离卷积参数数量可以大幅减少,提高运算速度<sup>[11]</sup>。

改进后的总体网络结构见图 2。

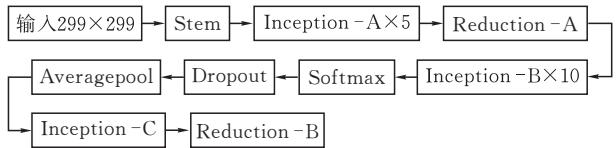


图 2 总体网络结构图

3 实验与分析

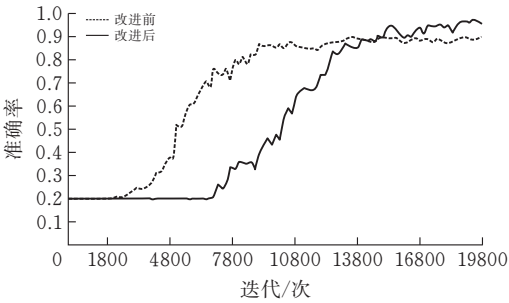
3.1 数据集的介绍

本文使用包含 5 种品类(animals, flowers, guitar, houses, plane)的开源数据集。每个种类包含 500 张图像,分辨率大小不一。同时将网络应用在自制槟榔数据集,该数据集图片采集自工业现场,包含 4 种不同等级共 4973 张图片数据,图片分辨率为  $301 \times 881$ 。在使用 Inception-ResNet v2 进行图片分类时,网络可接受的输入图片大小为  $299 \times 299$ ,所以在进行训练前需要对数据进行随机裁剪,通过这些操作能够获得更多的图像样本,对提高网络模型的准确率和泛化能力非常有帮助<sup>[12]</sup>。所有训练图片 4/5 作为训练集参与网络训练,1/5 作为验证集来分析分类检测的效果。

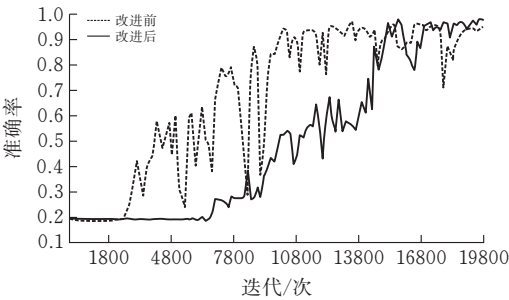
3.2 模型的训练

本研究的实验硬软件环境采用系统平台为 Windows 7,所用 CPU 为 Inter Core i5-4590。同时采用 GPU GTX1080Ti 显卡,CUDA8.0 ,cuDNN6.0 进行计算加速,所使用的编译环境为 TensorFlow

1.14。对于训练参数,学习率 0.01,设置最优优化动量参数为 0.9997,权重衰减正则项为 0.000 04,迭代 20 000 次,训练时每 200 次迭代保存一次模型。训练完毕后,针对开源数据集,网络改进前后训练信息见图 3。



(a) 开源数据集



(b) 自制槟榔数据集

图 3 训练数据准确率

从图 3a 可以看出开源数据集在第 2000 次迭代后准确率开始上升,训练 13 000 次后基本到达稳态,准确率维持在 89%附近。改进后的网络在 7000 次迭代后,准确率开始上升,原因为网络结构对图片原始特征的重复利用增加了训练的强度,所以开始阶段准确率基本不变。在进行 7000 次迭代后,准确率不断升高,最终准确率维持在 97%。相比于原网络,改进后的网络模型 Inception-Resnet v2 准确率提升大约 8%。同时,在计算速度上,改进后的网络虽然特征重复利用增加较多参数,但是在减少特征图一半数量的基础上,再进行深度可分离卷积,网络的运算量降低 19%。没有采用深度可分离卷积计算,仅仅特征图减少一半,相比于改进前网络运算量增加 0.7%。深度可分离卷积在降低网络运算量,提高运算速度上具有非常重要作用。

在自制数据集上网络进行 25 000 次训练(图 3b)。在自制数据集上,原网络在进行 3200 次迭代后准确率开始上升,在整个训练过程中准确率并不能稳定上升,出现较大波动,改进后的网络在 8200 次训练准确率开始上升,在后续训练中,准确率也会有所波动,但是总体呈现上升趋势。在进行 22 000 次迭代后,原网络准确率开始下降,改进后网络准确率趋于稳定,说明改进后的 Inception-ResNet v2 网

络较原网络更能够防止过拟合现象的发生。训练完成改进后的网络较原网络准确率提高 4%，达到 98.1%。

由图 4 可以看到,改进前网络模型的 loss 值在前 7000 次迭代中不断上升,然后下降并出现波动。最终达到稳定状态,最后 loss 值稳定在 3 左右。改进后的网络模型 loss 值在前 8000 次迭代中也是不断波动,8000 次后迅速降低到 1.8 左右,同时 loss 更快达到稳定状态,说明改进后网络拟合性能更优,后期有小范围的波动。分析其原因是 batch size 设置的大小对 loss 产生影响。最终网络 loss 值趋于稳定,表明网络模型的性能良好。

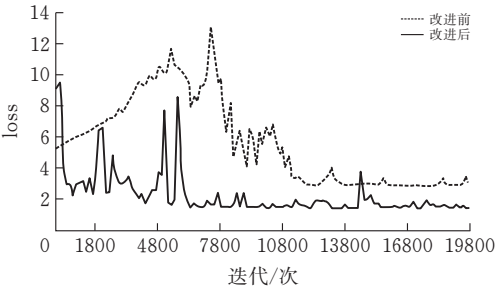


图 4 模型 loss 走势

4 结 论

本文针对物体分类提出的改进了的 Inception-Resnet v2 算法,通过加强输入图片原始特征利用,来获取更多的表面特征信息,从而提升物体识别准确率。常规加强特征重复利用时,通常网络模型的参数量、计算量均会大幅增加。为解决此问题,在卷积方式上进行调整,采用深度可分离卷积替代常规卷积方式。同时对卷积时特征图数量进行缩减。实验表明:改进后的网络模型在获取更优效果的同时,计算量也有所下降。对于物体的识别分类具有一定的参考意义。

由于网络对原始输入分辨率要求过高,因此对输入分辨率较小的数据集训练准确率不高。若训练时对输入图片进行拉伸,难免会破坏原有分辨率特征,导致对输入图片特征学习不足,从而训练效果不尽人意。未来还应在保证准确率的情况下,对小分辨率数据集图像的训练进行优化。

[ 参 考 文 献 ]

[1] 洪伟,李朝锋.基于区域全卷积网络结合残差网络的火焰检测方法[J].激光与光电子学进展,2018,55(4):168-173.

[2] 何新宇,张晓龙.基于深度神经网络的肺炎图像识别模型[J].计算机应用,2009,39(6):1680-1684

[3] 黄双萍,孙超,齐龙,等.基于深度卷积神经网络的水稻穗瘟病检测方法[J].农业工程学报,2017,33(20):169-176.

[4] 葛梦颖,于重重,周兰,等.基于协同半监督的深度学习图像分类算法[J].计算机仿真,2019,36(2):196-200.

[5] 刘翠玲,林珑,于重重,等.基于深度学习的花生高光谱图像分类方法研究[J].计算机仿真,2020,37(3):189-192,283.

[6] 韩要昌,王洁,史通,等.基于改进 GoogLeNet 的遥感图像分类方法[J].弹箭与制导学报,2019,39(5):139-142,153.

[7] Balagourouchetty, Lakshmipriya, K Jayanthi, Pratheeswaran, Pottakkat, Biju, et al. (2019). GoogLeNet based ensemble FCNet classifier for focal liver lesion diagnosis[EB/OL]. [2019-01-04]//www.researchgate.net/publication/335956445\_GoogLeNet\_based\_Ensemble\_FCNet\_Classifier\_for\_Focal\_Liver\_Lesion\_Diagnosis

[8] Kim J H , Seo S Y , Song C G , et al. Assessment of electrocardiogram rhythms by GoogLeNet deep neural network architecture[J]. Journal of Healthcare Engineering, 2019, 2019:1-10.

[9] Smirnov E A , Timoshenko D M , Andrianov S N . Comparison of regularization methods for ImageNet classification with deep convolutional neural networks [C]// 2013 2nd AASRI Conference on Computational Intelligence and Bioinformatics(CIB 2013). 0.

[10] Szegedy C , Liu W , Jia Y , et al. Going deeper with convolutions[C]// 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2015.

[11] He K , Zhang X , Ren S , et al. Deep residual learning for image recognition[C]// IEEE Conference on Computer Vision & Pattern Recognition. IEEE Computer Society, 2016.

[12] Huang G , Liu Z , Maaten L V D , et al. Densely connected convolutional networks [C]// CVPR. IEEE Computer Society, 2017.

(下转第 74 页)