

改进残差网络以增强垃圾分类

唐子煊 周 丽

四川大学锦城学院 计算机与软件学院 四川 成都 611731

【摘要】随着计算机技术的提升,垃圾分类问题成为当今热门的话题。垃圾分类的目的在于将垃圾种类分类处理,各种垃圾的处理方式不同,将提高可回收垃圾的再利用率。本文通过 ResNet18 原有的网络对垃圾分类进行识别,由于该模型是基于 ImageNet 实现分类,因此需要达到更好的效果,则需要对垃圾分类的数据集作迁移学习,以增强识别垃圾分类的能力。在 ResNet18 中,增加当前神经网络层数,最终达到的减少训练次数的同时,达到提高准确率,并且需要防止过拟合。

【关键词】ResNet18; 垃圾分类残差网络

1 引言

在当今社会的发展现况,对垃圾混杂处理进行焚烧,填埋等措施,极大地浪费了资源,不仅在能源消耗方面,而且在垃圾本身的可利用价值上。因此需要对垃圾进行有效的分类,不同种类的垃圾进行不同处理。同时政府已经颁发政策来进行垃圾分类的管理。对个体,垃圾分类对于个人生活有重要的作用。对于政府,垃圾分类将有效促进经济发展。

为了实现在日常生活中辅助人们对于垃圾类别的判别,本文采用基于监督学习的卷积神经网络来实现垃圾图像的分类。

整个分类和建立的思想在将采取标签标准方式进行收集数据,将收集数据进行数据筛选,进行处理,最终通过修改后的残差网络进行训练。残差网络选择 ResNet18,因为 ResNet18 是基于 ImageNet 进行图像 1000 类别的分类,因此垃圾分类可以基于 ResNet18 对垃圾图像进行识别并分类。

2 残差网络

2.1 网络结构

CNN 神经网络整个的层数包含:卷积层、池化层、全连接层、输出层。

卷积层将提取到每一层上的多个特征值,第一层上所提取到的边缘、轮廓的特征点,传输到下一层进行深入的提取,这样的层在 ResNet18 中存在有 17 层。

ResNet18 使用最大池化层,将在卷积层后面进行,有效地降低了计算量,扩大感知野,最后进行二元自适应平均池化操作。

全连接层 ResNet18 中,原有是分为 1000 分类,因此这里需要将 1000 分类改为需要分类的垃圾数目。

2.2 改进网络

ResNet 残差网络全称 Residual Network,不同于一般的 LeNet 网络等网络,由于它不断通过增加卷积层,可以使得本身精度得到了更多的特征值,从而进一步的增加,但是与此同时它并没有因为层数的增加出现了过拟合的现象,这是由于其残差映射进行拟合。核心思想为:对一个准确率达到饱和的网络进行了完全拟合恒等映射函数 $H(x)=x$,产生了残差块结构^[1]。

残差块的结构原理如下图 1 所示:

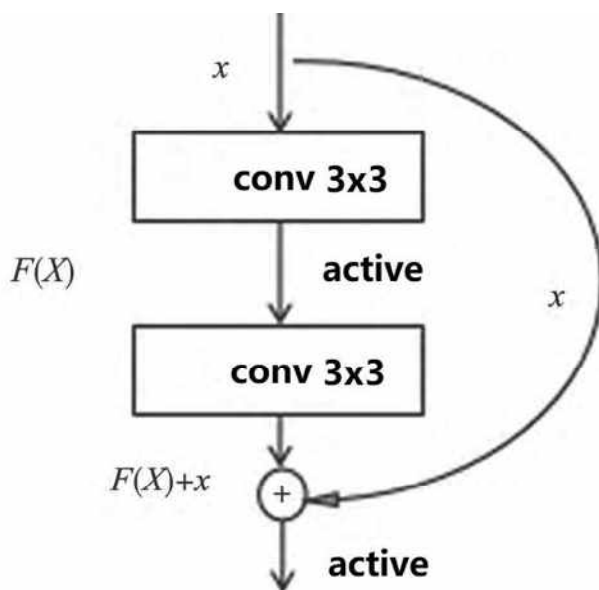


图 13*3 的激活层残差块

$H(x)=F(x)+x$, 其中 x 为前一层的输入, $F(x)$ 为残差学习后的输出, $H(x)$ 表示为残差模型的整体输出^[2]。

这里 3*3 的激活层残差块,源自 ResNet18 中的 layer1 第 1 层的简化图,中间仍有两次的 BN 和 1 层 ReLU 激活函数。BN 为标准化和归一化的方式,增加了批归一化,提高数据进度同时也可以加快训练速度,标准化输入可以提高训练的速度,增强泛化能力^[4]。为了保持在整个训练过程中不会因为卷积,池化等操作使得训练的图片大小发生变化。

Res Net 网络结构如表 1 所示:

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
conv2_x	56×56	3×3 max pool, stride 2				
		$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix} \times 8$
conv4_x	14×14	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix} \times 36$
conv5_x	7×7	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10 ⁹	3.6×10 ⁹	3.8×10 ⁹	7.6×10 ⁹	11.3×10 ⁹

表 1 ResNet 网络结构

这其中 18 层和 34 层均没有采用 1*1 的卷积层进行操作，这样操作的目的是为了减少相应的参数。而 18 层和 34 层由于数目不大相比不用进行 1*1 的操作。

中的右侧箭头分析出，均为残差块，作用是为了防止过拟合出现。因此基于 ResNet18 进行垃圾分类问题的改进，垃圾分类中需要分类出 40 种的垃圾。

ResNet18 是由 17 层卷积层以及 1 层全连接层构成。随着深度学习与神经网络层的不断改进，而整个过程，从深度学习中的卷积神经网络模型 (CNN) 由最初的 LeNet 模型到 AlexNet、VGGNet、InceptionNet 再到 ResNet，模型的深度逐渐加深，识别准确率逐渐提高^[3]。



(1) 其他垃圾



(2) 厨余垃圾



(3) 可回收垃圾



(4) 有害垃圾

3 实验及分析结果

3.1 数据集

本次实现总共将 14780 个数据进行训练，其中将训练数据和测试数据进行分离，用于训练的数据占总共 80%，并且将数据的顺序进行乱序。训练数据总共 11824，测试数据总共 2956。并将数据依次存入训练和测试的文件目录中。垃圾分类中总共分为 4 个类别，细分为 40 种的小类，类别如图 2 所示

图 2 垃圾分类类别

3.2 数据预处理

将用于训练的数据通过缩放，中心裁剪，旋转，翻转的方式进行处理。缩放到 256*256，中心裁剪为 224，测试集数据进行图片属性随机调整，亮度 0.5，对比度 0.5，饱和度 0.5。随机旋转角度为 5，并进行默认随机翻转。对于测试的数据进行缩放和中心裁剪。

此处没有将测试数据与训练数据进行统一化处理。使数据进行一定的随机化处理,训练数据进行随机化处理可以有效增强训练的效果,测试数据没有进行改变,这种方式相对于训练数据没有处理,而测试数据随机化的方式能让模型更加准确处理。

3.3 模型训练

基于 ResNet18 的基础将全连接层的输出端改为 40, 因为 ResNet18 本身的训练结果为 1000 类别的分类, 而在这个垃圾分类中, 需要 40 种类的分类。在 ResNet18 本身的训练过程中, 发现收敛速度较慢, 而此时的学习率较大, 因此从这个角度出发, 初始的准确率较低, 将模型进行改变, 在 ResNet18 原有的全连接层中, 增加一层全连接层。初始准确率较低, 则会直接影响训练效果, 因此需要通过减少相应参数的目的来实现快速下降的目的。

学习率为 0.01, 迭代次数为 10 次, 总体的框架为 17 层的卷积层以及改进后 2 层的全连接层。并且在 17 层的全连接层当中, 增加有 1*1 的卷积层, 因为 1*1 的卷积层虽然对于 18 层网络的影响不大, 但本身可以减少相应的参数, 则可以达到进行训练的目的和要求。改进网络层图 3 所示:

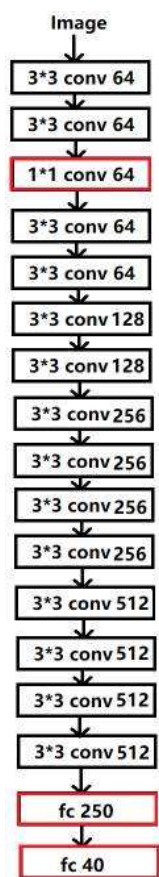


图 3 改进网络层

3.4 结果和分析

在实验中, 首先对于原有的 ResNet18 模型, 将垃圾分类中的训练集进行训练, 然后将训练集进行测试, 来进行对照, 然后将改进后的 ResNet18 模型用相

同的训练集进行训练, 并且进行测试, 以验证改进后的 ResNet18 在初始的准确率, 以及后续的梯度下降中相较于原有 ResNet18 有了很大的提升和改变。在测试集和训练集的准确率的图像中, 我们可以清晰地看出相比于原有 ResNet18 模型, 改进后的模型在第一轮进行训练的时候, 准确率均比原有训练的第一轮准确率要高一些, 改进后的 ResNet18 能够快速进行下降, 因此在后续的迭代过程中, 改进后的模型能够更好地进行快速下降, 并且能够得到较高的准确率, 在图 4 的 ResNet18 最终的准确率为 80.43%, 在图 4 中的改进后的 ResNet18 的准确率为 81.18%。

图 4 中红色和蓝色为不同的模型下的准确率。这里采用 SGD, SGDR 即带重新启动的随机梯度下降的学习率, 使用余弦函数作为周期函数, 并在每个周期最大值时重新开始学习速率^[5]。为使神经网络的误差尽量小, 损失函数要取到最小值, 这个过程可以近似看作求取损失函数最优解的过程。对损失函数最小值的寻找方向一定是其下降幅度最大的方向, 即损失函数初始点位处梯度向量的方向^[6]。因此通过这个实验, 我们可以很明显看出, 针对于不同的训练集, ResNet18 需要因训练集的特征而做出相应的改变和调整, 以便于我们能够适应现实环境下的条件得到更好的训练结果。

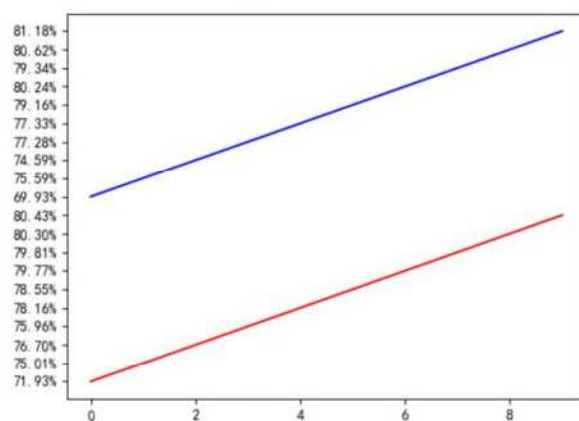


图 4 ResNet18 和改进 ResNet18 准确率

4 结论

对于 ResNet18 作为残差网络, 本身可以避免因为训练特征过多而导致的过拟合现象, 因此选择 ResNet18 是基于 17 层卷积层以及 1 层的全连接层的层数并不多, 而在本身图片大小进行标准化处理的结果后, 所需要训练的轮数不多, 就可以达到快速训练出结构的要求。对于 ResNet18 原有模型的基础上, 由于开始的准确率不高, 因此通过增加一层全连接层的方式, 来进一步提高其准确率。在基于原有 ResNet18 的基础上, 根据数据集的特点, 相应对网络结构进行调整, 符合实际生产中的环境特点。这同时表明, 并非所有的模型都可以原封不动进行迁移使用, 就可以得到很好的效果, 相对于数据集的特点进行有效的调整, 则是更为关键的步骤和方法。

【参考文献】

- [1] 管淑贤, 葛万成. 基于 ResNet18 的减速带识别及其环境影响研究 [J]. 通信技术, 2021, 54(03):597-603.
- [2] Jiasong Bao and Xixia Huang and Ke Ji. Recognition of Floating Objects on Water Surface--Based on Improved Residual Network[J]. International Core Journal of Engineering, 2021, 7(5)
- [3] 王浩. 基于深度残差网络 ResNet 的废料瓶分类系统 [J]. 科技与创新, 2020(14):71-72.
- [4] 杨剑, 刘方涛, 张涛, 张启尧, 任宇杰. 基于改进型残差网络烟雾图像识别 [J]. 科学技术与工程, 2019, 19(32):236-243.
- [5] 郑佑顺, 林珊玲, 林志贤, 周雄图, 郭太良. 基于残差结构和幻象模块的垃圾图片分类算法 [J]. 信息技术与网络安全, 2021, 40(01):50-55.
- [6] 谢裕睿, 董建娥. 基于 ResNet 网络的东巴象形文字识别研究 [J]. 计算机时代, 2021(01):6-10.