

基于挤压激励残差神经网络的垃圾分类

池硕 李丹

四川大学锦城学院计算机与软件学院 四川 成都 610065

【摘要】为了简单高效地对生活垃圾进行识别分类，本文提出了一种基于挤压激励残差神经网络的垃圾分类识别方法。该方法是使用挤压激励块对残差网络进行的变形，这种网络继承了二者的优点。本文设计一组对照实验，并通过实验观察对比表明，挤压激励残差网络有效的提高了网络的效率，加快网络收敛速度，准确率比同等参数的残差网络高 1%。

【关键词】残差网络；垃圾分类；图像分类；深度学习；挤压激励网络

引言

随着我国经济不断的飞速增长，城市化进程不断的快速向前推进，城市居民生活垃圾产量快速的增加。有媒体报道，目前我国有 30% 以上的城市已经被垃圾所包围，另外还有数据表示，我国除了县城之外的 668 个城市当中，有大约 60% 的城市已经被垃圾包围，25% 左右的城市已经没有场地来填埋堆放垃圾^[1]。因为长时间的将垃圾露天进行堆放，对大气环境、地下水和土壤等已经造成了非常严重的威胁和危害。当今传统的垃圾处理方式已经无法适应和满足我国经济建设的快速发展，随着垃圾排放量的不断增长，垃圾环境问题已经危害到了人们的身心健康，同时也束缚了经济的发展。

2019 年 7 月，上海市率先开始实施垃圾分类政策，跟据上海绿化市容局报告，上海 1.3 万多个居住区中的垃圾分类达标率已经由 2018 年底的 15% 提升到了 2019 年底的 90%，基本的形成了垃圾全程分类收运的体系^[2]。今年 5 月，上海全市湿垃圾分出量达到了自从垃圾分类实施以来的最高峰。截至 2020 年 6 月，全国共有 46 座城市计划在年底前先行实施生活垃圾强制分类政策。

目前垃圾识别分类使用的主流技术，是利用传统的计算机视觉算法，或者采用传感器进行筛选识别^[3]。本文采用了基于残差网络的挤压激励神经网络来进行图像分类，并与传统的残差网络进行对比分析。

1 残差神经网络

深度学习网络的深度对最后的分类和识别的效果有着非常大的影响，所以很多时候我们就想的是把网络设计的越深越好，但是事实上并不是这样，常规的网络的堆加在网络很深的时候，效果变得越来越差。其原因之一是由于网络层次越深，梯度消失或爆炸的现象就

变的更加严重，网络的训练效果也会变差。为了解决在随着网络层数增加的情况下出现的梯度消失或爆炸问题，2015 年何凯明等人题出了残差网络^[4]（Residual Network）。

残差网络使用了新的思想^[4]，即假设我们涉及的一个网络层，存在最优化的网络层次结构，我们所设计的这个深层次的神经网络，一般是有很多的网络层是冗余层的。这种情况下我们希望这些冗余的层是能够进行恒等映射的，确保经过该恒等层的输入和输出完全相同的。残差网络是由残差块构建的，在残差块当中， x 为残差块的输入， $F(x)$ 是当前块最后一层的线性变化后的输出，先让 $F(x)$ 和这一层的输入值 x 相加，然后再进行激活并输出。在输出值激活前与 x 相加，这条路径称作短连接（shortcut）。 $F(x)$ 为残差。残差网络的结构和原理都非常的简单，但非常有效的解决了随着深度的增加而效果变差的问题。

2 挤压激励神经网络

卷积神经网络（CNN）的核心操作是卷积操作，它让网络能够通过每一层的局部感受野当中，融合空间和通道信息来构造信息特征^[5]。挤压和激励（SE）块，它通过显式地建模各通道间的相互依赖性，并自动调整通道特征响应。这些块可以堆叠在一起形成挤压激励神经网络结构（SENet），也可以插入到目前已经存在的网络当中，并且在不同的数据集上都有着很不错的泛化能力。SE 块以微小的额外计算成本为现有的最新 CNN 带来显著的性能改进。

SE 块的结构中，对于将任意给定的输入 x ，映射到特征映射 U 的变换 F_{tr} ，如卷积操作等，其中 $U \in \mathbb{R}^H \times \mathbb{W} \times \mathbb{C}$ 。特征 U 首先通过挤压操作，挤压操作

是通过全局平均池化提取其空间维度 ($H \times W$) 上的特征映射, 来完成对通道的描述^[6]。该描述的功能是, 生成通道特征响应的全局分布的信息嵌入, 让这个网络的全局感受野的信息可以被其所有的网络层使用。提取之后是一个激励操作, 它采用一个简单的门控机制的形式, 该机制以信息嵌入作为输入, 并生成每个通道权重的集合。这些权重应用于特征映射 U , 以生成 SE 块的输出, SE 块的输出可直接传入网络的下一层。

本文将 SE 块与残差网络相结合, 在两者的基础上实现挤压激励残差网络 (SE-ResNet)。挤压激励残差网络的构成块的结构是, 在残差块的卷积操作之后, 进入 SE 块, 再将结果与残差块的卷积操作后的结果相乘, 构成残差项, 再于原始输入项相加。我们也可以通过改变 SENet 挤压部分的网络结果来再次的对网络进行升级^[7], 以再次提高网络效率。

3 实验设计与结果分析

所使用数据集来自百度 AI Studio 的开源数据集。原始数据集中的数据按照垃圾分类的四大标准分为四个文件夹, 每个文件夹里面又根据具体的垃圾分为多个小的文件夹, 比如说可回收物的文件夹中有袋子和充电宝等垃圾, 在袋子的文件夹里面又分为打包袋、档案袋、购物纸袋、尼龙袋和纸袋。整个数据集一共包含 246 种图片名称 (有的名称下还会细分为更细致的名称如上文中提到的包), 一共 80961 张图片, 文件压缩包的大小是 9.64GB。为了使用方便, 我们将多层目录转变为单层目录, 同时删除掉出现损坏的文件, 最终数据集共有 247 个分类, 80943 张图片, 对其随机分割成训练集和数据集, 训练集占比 80%, 共 64754 张, 测试集占比 20%, 共 16189 张。

我们共设计了一组对比实验, 具体实验的过程中, 采用相同的参数 (如表 1) 分别用残差网络和挤压激励残差网络对数据集进行训练, 对比两者的实验产生的结果 (如表 2, 图 1, 图 2)。实验在 RTX 1080Ti、PyTorch 1.12、CUDA 10.0 和 Python 3.6 的环境下进行。

表 1 实验参数

参数名称	参数值
学习率	0.1
批处理大小	64
网络层数	34
迭代次数	50
降低率 (仅 SE-ResNet)	16

表 2 实验结果对比

实验结果	ResNet	SE-ResNet
准确率 %	59.160	60.250
训练平均单次迭代用时 s	614	303
验证平均单次迭代用时 s	63	63

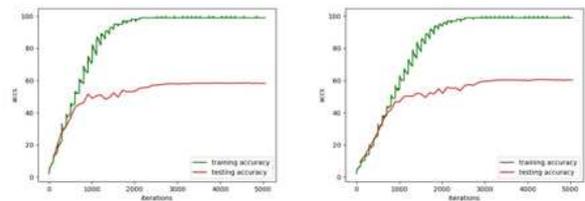


图 1 34 层 SE-ResNet (左) 和 34 层 ResNet (右) 准确率走势图

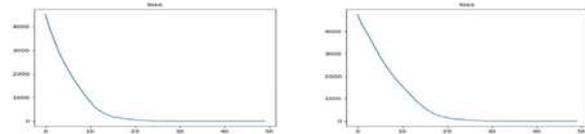


图 2 34 层 SE-ResNet (左) 和 34 层 ResNet (右) 损失值走势图

由于批处理大小为 64, 所以 64754 张训练集共分成 1011 份进行训练, 训练过程中每 10 次记录一次训练准确率, 迭代 50 次共记录训练准确率 5050 次, 而测试集每验证一次记录一次准确率, 迭代 50 次共记录测试准确率 50 次。图 1、图 2 为实验所记录数据图表, 在两张图中, 其中左边图表为 34 层 SE-ResNet 实验数据, 右边图表为 34 层 ResNet 实验数据。图 4 为训练过程中的准确率走势图, 绿色为训练准确率, 红色为测试准确率。图 2 为训练过程中的损失值走势图。

受硬件条件限制, 模型训练结果并不是很理想, 但我们观察实验结果仍然可以发现, 挤压激励残差网络的训练结果比普通残差网络的训练速度有很明显的提高, 34 层的网络准确率提高 1.09%。我们通过上面 2 张图可以发现, 34 层的挤压激励残差网络在迭代 25 次左右开始收敛, 而 34 层的残差网络在迭代 30 次左右才开始收敛, 挤压激励残差网络有更快的收敛速度。我们通过模型的训练用时发现, 普通的残差网络要比插入 SE 块的残差网络慢将近一倍, 而训练中的验证过程却大致相同, 笔者推测残差网络的时间消耗在网络的反向传播过程中。

挤压激励网络通过强调特征的重要的通道并且抑制不重要的通道来提高网络的性能, 残差网络通过短链接有效的解决了网络退化和梯度消失或爆炸的问题, 两者结合而成的挤压激励残差网络很好的继承了他们的优点, 在分类问题上表现出了优异的成绩。

4 应用设计

为了更加方便的进行垃圾分类, 对训练好的模型进行封装, 设计成垃圾分类查询软件。应用包含上传图片查询和拍照查询两个功能, 应用 UI 采用 PyQt5 来设计^[8], 共包括首页 (如图 3 左)、结果展示页 (如图 3 右)、拍照页 (如图 4 左) 和拍照结果页 (如图 4 右) 4 个页面。

打开应用后进入首页, 在首页可以选择要查询的

方式，上传图片查询和拍照查询，选择上传图片查询后弹出图片选择框，选择要查询的图片后，会转到结果展示页，该页面包含查询的图片和对应的结果，结果为所属的垃圾分类和具体的物品名称两部分，我们可以点击页面左上角的返回首页继续查询。选择拍照查询后会转到拍照页面，我们点击拍照后，会转跳到拍照结果页，我们在此页可以选择使用此照片查询或重新拍照两个选项，若选择重新拍照，则会跳转会拍照页，若选择使用此照片查询则会跳转到结果页。



图3 应用展示 (1)



图4 应用展示 (2)

5 结束语

本文通过对比残差网络和挤压激励残差网络对垃圾分类进行了初步的研究，并做成一个简单的应用。由于挤压激励网络可以自适应地重新校准通道特征响应，强调重要的特征，抑制不重要的特征，所以它非常有效的提高了网络的表示能力和性能。并且因为残差网络能有效的解决网络退化和梯度消失或爆炸的问题，将二者结合所构建的挤压激励残差网络能更加有效的提高网络的效率。

垃圾分类是件非常有意义的事情，若将软件与硬件相结合，形成垃圾自动分类产品，能更加有效的推广和进行垃圾分类。

【参考文献】

- [1] 何林华. 浅谈我国城市生活垃圾分类收集的意义及对策 [J]. 江西化工, 2013(04):326-328.
- [2] 钱培坚. 上海垃圾分类，一周年答卷如何? [N]. 工人日报, 2020-06-18(003).
- [3] 秦斌斌, 何级. 基于卷积神经网络的垃圾分类研究 [J]. 无线通信技术, 2019,28(03):51-56.
- [4] He, Kaiming et al. "Deep Residual Learning for Image Recognition." 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016): 770-778.
- [5] Hu, Jie et al. "Squeeze-and-Excitation Networks." 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (2018): 7132-7141.
- [6] 张佳怡. 基于多尺度特征融合与挤压激励模型的目标跟踪 [D]. 西安: 西安电子科技大学, 2019.
- [7] 常昊, 陈晓雷, 张爱华, 李策, 林冬梅. 嵌入改进 SENet 的卷积神经网络连续血压预测 [J/OL]. 计算机工程与应用 :1-6[2020-06-20].
- [8] 肖文鹏. 用 PyQt 进行 Python 下的 GUI 开发 [J]. 中文信息, 2002(07):73-75.