

基于深度残差网络的垃圾识别分类应用

梁言周 丽

四川大学锦城学院 计算机与软件学院 四川 成都 611731

【摘要】随着科技的进步,人们越来越多地思考关于科技和生活的问题,其中一大难题就是越来越来的城市垃圾处理能力,所以开始推行垃圾分类(也就是将垃圾按规定依次进行分类储存、投放和搬运)这一政策,以此展示并实现垃圾的最大经济价值。因此,对垃圾的分类和回收具有巨大的研究价值与重要意义。于是解决垃圾识别分类问题也成为了我们的首要任务。基于此,笔者借助ResNet18网络结构以及迁移学习的思想设计了一款垃圾分类系统。实验表明,基于这款垃圾分类系统模型,垃圾分类准确率能达到90%以上,很大程度上地解决了垃圾识别分类问题。

【关键词】ResNet18; 迁移学习; 深度学习; 垃圾分类

1 绪论

随着当前我国的城市化聚集和智能化生活的步伐加快,人们的物质生活水平也在不断地提升,生活需求不断加大,导致城市日常生活中的垃圾总量始终呈现出稳定增长的态势。近年来,我国已发展成为全球范围内城市日常生活垃圾处理工作压力最大的国家之一^[1]。据初步统计,目前中国97%的地区由于无法收集到城市垃圾,所以只能在地下室堆放或者场外填埋,且传统垃圾处理方式存在各种环境污染和资源浪费的弊端。日本的垃圾治理主要着重于发展循环经济,着重于资源再利用型的垃圾分类处理;德国的垃圾分类的法律十分完善也很严明,而且德国特意有一套自己专门的惩罚措施,来特地针对个人垃圾分类处理方面;在美国,乱丢垃圾被认定为一种违法行为,惩罚不仅是罚款,严重时甚至还要入狱或社区服务,最高要收监一年^[2]。而自2000年起,我国便开始试点实行垃圾分类,直至如今,我国的垃圾分类处理进程仍然处在初级阶段。近年来,我国一直在大力宣传进行垃圾分类工作。但对于大多数人来说,很多人只能借助直觉和常识来进行垃圾分类,面对各类不同的垃圾很难做出正确的判断。在2018年这一

年期间,我国国内的各种生物垃圾清运量超过了2亿吨,也就是人均垃圾废物产量超过0.4吨,人们的生命和身体健康被大量垃圾所影响、所危害,同时,这些垃圾也很大程度地侵蚀着人类的正常活动领域和土壤面积^[3]。

而对于垃圾处理的过程中,确定垃圾的位置以及识别垃圾的种类时实现自动化分拣和处理的前提。在此前提下,传统图像处理和识别方法需要我们逐一提取图片特征,且受数据集的影响很大,我们完全可以利用深度学习的方法于垃圾图像识别分类的领域。

2 实验理论知识

2.1 garbage_classify 数据集

在这个模型训练中,我们使用的garbage_classify数据集包含了14802张图片以及14802个由对应图片名字和相对应的标签组合的文本文档组成,文本文档中的图片名与标签用半角逗号分隔开,在其他垃圾、厨余垃圾、可回收物、有害垃圾四个大范围内,分为了40个不同的类别,分别对应数字0~39,各类别的其数量关系如下图所示:

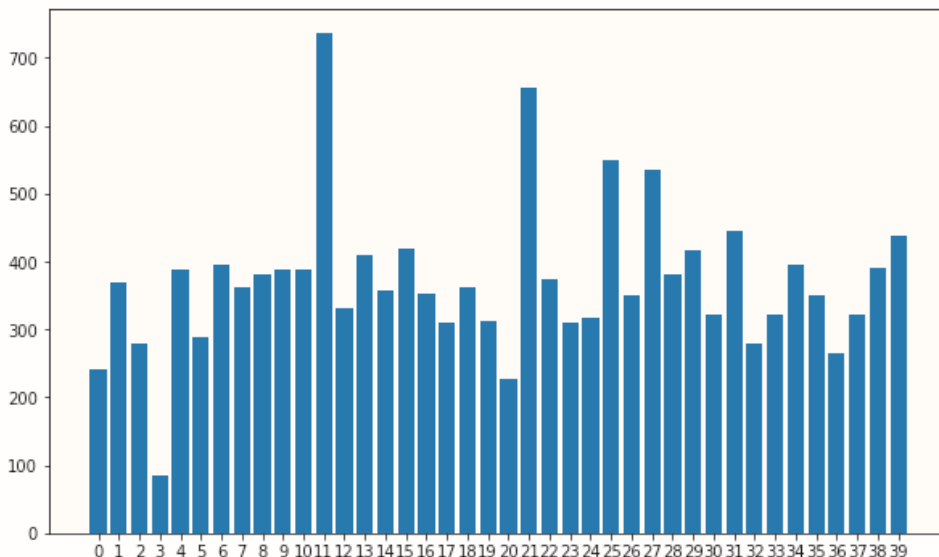


图1 四十类垃圾的数量关系

在数据集划分时, 本文以 0.8 为比例进行划分, 分别将 40 个种类的图片划分为 80% 的训练集和 20% 的测试集。

2.2 平原网络与深度残差网络

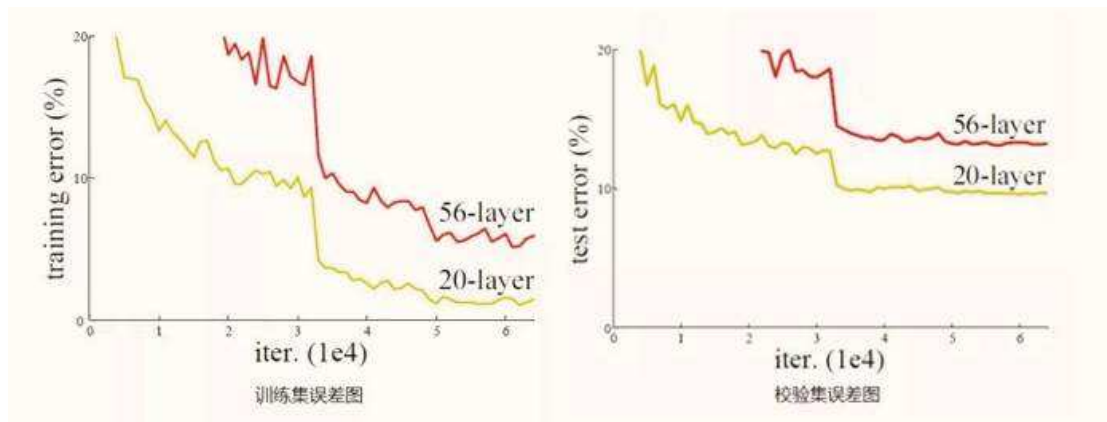


图 2 多层平原网络训练集测试集误差

可以看出, 在网络层达到一定深度时, 模型效果所有降低, 误差率陡然提高。此时我们不难看出这样的结论: 模型的精度随着网络的层数的不断增加而提升, 而当网络叠加到一定的层数以后, 训练集和测试集的准确率都迅速下降。由此可见, 当网络层数达到一定深度以后, 深度网络的各种功能有所减退, 因此也变得更加难以训练了^[4]。

而 ResNet (Residual Neural Network, 残差网络) 网络使用了残差网络结构, 在需要把网络层数设定得很深的时候, 我们就可以使用这种残差网络结构。这种残差网络结构目前最多可以使我们的网络达到一千多层。通过这样的结构进行训练得来的最终的分分类效果也达到了较好的程度。下图展示的即是残差块网络块的基本结构, 不言而喻, 残差块是带有跳跃结构的:

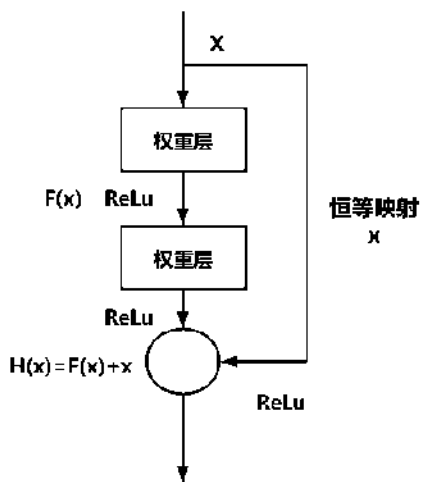


图 3 残差块结构

而深度残差网络的优势就在于残差网络学习恒等函数比较容易, 可将添加的网络层看成一个残差块^[5], 利用短路机制将仿射变换的矩阵乘法变成了加法性质, 这使得计算相对更加稳定, 且短路机制可以无损地传播梯度, 梯度不会在加权层之后直接传递过来, 从而进一

步抑制了梯度衰减。通过被抑制的梯度衰减以及加法性质的综合计算, 大大提升了训练的稳定性, 并极大简化了训练过程。

针对传统神经网络有着第 n 层的输出, 只能给接下来的第 $n+1$ 层作为输入的牵制, 这种深度残差网络可以使某一层的输出越过几层权重层, 直接作为几层后的权重层的输入^[6], 以这样的方式完成叠加多层网络的效果, 这既没有引入新的参数, 也没有增加计算复杂度, 具有很强的推广性, 大大解决了多层模型在深层时准确率下降的难题。本文中用到的 ResNet18 网络是由 17 个卷积层以及 1 个全连接层组成的。

2.3 迁移学习

在机器学习中, 迁移学习 (Transfer Learning) 是其中常见的模型训练方法之一, 它是利用已存有的基础知识来分析并解决不同但相似技术领域的问题的机器学习方法^[7]。简单来说就是举一反三, 根据已有知识的实际情况和原理, 推导出新的知识; 那么在机器学习中, 就是通过已有的神经网络模型数据, 找到共通点, 研究出如何学习新模型的模式, 其中最关键的一点就在于如何发现神经网络模型中已有知识和新信息两者之间的相似性。迁移学习的主要优点在于对数据和标签的依赖性小, 放宽了独立同分布约束条件^[8]。

所以, 使用迁移学习的方法不但能够节省大量时间, 降低我们使用 GPU 所需要的算力, 也是在各种方面上获得更好性能的捷径。尤其针对刚接触深度学习的学者来说, 由于自身知识和能力有限, 很难训练出一个较好的模型, 那么我们就可以利用迁移学习, 对前人训练好的模型加以迁移使用, 就可以更快更轻松地完成我们的需求。总的来说, 迁移学习就是利用源域和目标域两者之间的相似性和相关性, 将源模型部分参数或特征表示迁移到目标模型, 并自适应地自行调整目标模型。因此, 本文利用迁移学习的思想, 将 ResNet18 模型结构进行调整, 传入本文的 40 类垃圾图片进行训练。

3 系统实现

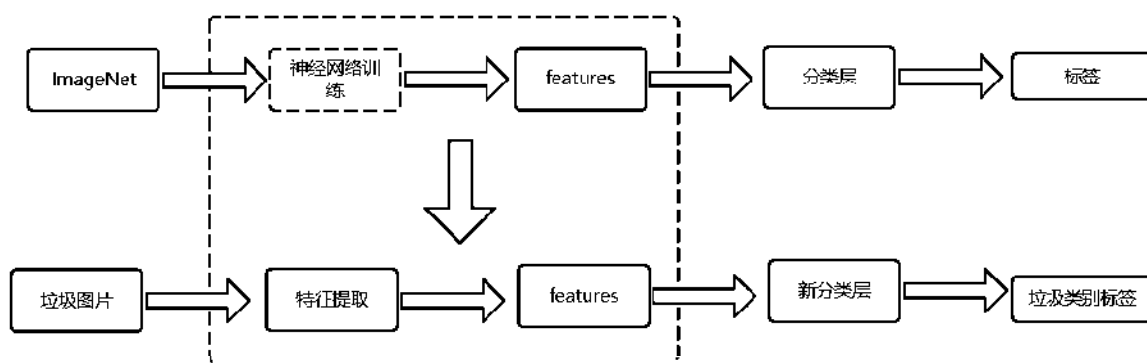


图 4 系统流程

本文使用 Pytorch 系统框架，是基于 torch 的 python 机器学习库，由 Facebook 开源，专门针对 GPU 加速含有张量计算的深度神经网络（DNN）编程^[9]。其特点在于张量计算引擎，自动求导机制和神经网络高级接口。Pytorch 中的计算基本都是基于张量 Tensor 的，它可以说是 Pytorch 中的基本计算单元。所以本文选择利用张量形式来对图片进行处理，以及以此计算实验最

后的准确率。基于 ResNet18 网络模型，并在其基础上进行部分改进。

关于图像处理与增强，本文利用 transforms 函数进行一系列操作处理，主要进行了中心剪裁、随机旋转、变换为张量并标准化等方法，使图片适用于已有的 ResNet18 网络，并且经过图片数据处理后，模型的泛化能力有所提高，收敛速度也有所加快。



```
tensor([[[[[-0.4226, -0.4226, 0.1597, ..., 1.2557, 1.6667, 1.5982],
[-0.5767, -0.2856, -0.1657, ..., 1.5982, 2.0605, 2.0434],
[-0.7479, -0.2513, -0.2684, ..., 1.9578, 2.0948, 2.0948],
...,
[-0.5938, -0.6109, -0.5424, ..., -0.6965, -0.9020, -0.8507],
[ 0.3823, 0.4337, 0.4679, ..., -0.6965, -0.8507, -1.0048],
[ 0.9988, 0.9817, 1.0502, ..., -0.7479, -0.3164, -1.1075]],
...,
[[[-0.4076, -0.3901, 0.2402, ..., 1.6232, 2.0084, 1.8859],
[-0.5126, -0.2150, -0.0574, ..., 1.8859, 2.3235, 2.2710],
[-0.6877, -0.1975, -0.1975, ..., 2.2010, 2.3235, 2.3060],
...,
[-0.6001, -0.6176, -0.5476, ..., -0.8452, -1.0378, -0.9503],
[ 0.5378, 0.5903, 0.6429, ..., -0.8277, -0.9678, -1.0903],
[ 1.2731, 1.2556, 1.3256, ..., -0.8627, -0.8978, -1.1604]],
...,
[[[-0.2532, -0.2358, 0.4091, ..., 1.9603, 2.3088, 2.1868],
[-0.4101, -0.1138, 0.0779, ..., 2.1868, 2.5877, 2.5529],
[-0.6018, -0.1312, -0.0790, ..., 2.4831, 2.5703, 2.5703],
...,
[-0.5147, -0.5321, -0.4624, ..., -0.7413, -0.9504, -0.8807],
[ 0.6705, 0.7054, 0.7576, ..., -0.7064, -0.8633, -1.0027],
[ 1.4722, 1.4374, 1.5071, ..., -0.7413, -0.7761, -1.0550]]]]]])
```

图 5 图片处理

在加载预训练模型后，便冻结网络，冻结网络是为了将指定的图层中的参数（也就是权重）保留在 Pytorch 中，并替换最后一个全连接层，将类别修改为垃圾类别数 40。由于本文中的垃圾识别分类任务为多分类任务，所以我们采用交叉熵损失函数。对比 ResNet18 网络和 GoogLeNet 网络，并且在不同情况下分别对比 Adam 和 SGD 两种加速器、不同的学习率以及 batch_size 后，综合结果发现，基于 ResNet18 网络

的迁移学习的运算速度更快，当各参数设定为 batch_size=32，使用 SGD 加速器，学习率设置为 0.003 时，模型效果最好。

googlenet	batchsize=64 optim=SGD lr=0.003	Train epoch:0 loss:1.6383 acc:0.5489 Test epoch:0 loss:1.1696 test_acc:0.6571 Train epoch:1 loss:1.1440 acc:0.6633 Test epoch:1 loss:1.0161 test_acc:0.7031
resnet18	batchsize=32 optim=SGD lr=0.003	Train epoch:0 loss:1.0020 acc:0.6993 Test epoch:0 loss:0.9544 test_acc:0.7231 Train epoch:1 loss:0.9674 acc:0.7109 Test epoch:1 loss:1.0203 test_acc:0.7045

图表 6 两种网络模型对比结果

此后便将处理好的数据放入模型中进行训练，10 轮训练后准确率就可以达到 90% 以上。

train: acc: 0.9403973509933775 recall: 0.9594594594594594
test: acc: 0.9578947368421052 recall: 0.978494623655914

图表 7 训练 10 轮后的准确率

5 实验总结与心得

本文针对垃圾识别分类问题，使用深度残差网络的 ResNet18 结构，借助迁移学习的思想，实现垃圾的识别与分类。在实验过程中，研究了垃圾分类的现状与存在的问题，以及在垃圾分类领域中的研究现状；创建了垃圾分类的训练集和测试集，并进行图像进行图像处理与图像增强；对比 ResNet18 和 GoogLeNet 两种网络，并找到目前最佳的模型参数。模型训练后期出现了轻微过拟合的现象，所以还需要对参数进行进一步的调节；由于没有包含垃圾的位置标签的数据集，本文完成了垃圾的识别与分类，对垃圾的位置确定这一任务的完成仍为空白，但打标签一事需要研究者逐一标注，工作起来繁琐耗时，目前暂时无法解决。关于垃圾分类一事，笔者认为着重点除了在于居民们能否对垃圾做出正确分类判断的问题，还在于居民参与的积极性不强，政府成体系的监管机制欠缺，以及垃圾分类的设备不够便捷和创新。综上所述，人文素养和科技水平应当共同进步，并驾齐驱，以此解决垃圾分类问题。

【参考文献】

- [1] 余红，梁惠能，方森锋，彭霞锋．“老黄牛”智能回收机的创新及运营模式 [J]．产业创新研究，2020(18):23-24.
- [2] 张雨康，蔡威，陈文浩，刘梦琦，唐梦瑶．智能垃圾分类系统研究文献综述 [J]．中国设备工程，2020(09):31-33.
- [3] 董子源．基于深度学习的垃圾分类系统设计与

实现 [D]．中国科学院大学（中国科学院沈阳计算技术研究所）．2020.

[4] 马进，王健宗．牲畜身份识别方法、电子装置及可读存储介质 [P]．广东省：CN109919005A, 2019-06-21.

[5] 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.

[6] 许泽珊．基于深度神经网络的视频动作识别研究 [D]．华南理工大学，2019.

[7] Pan S J, Yang Q A. A Survey on Transfer Learning [J]. Ieee Transactions on Knowledge and Data Engineering, 2010, 22(10):1345-1359.

[8] 李猛，李艳玲，林民．命名实体识别的迁移学习研究综述 [J]．计算机科学与探索，2021, 15(02):206-218.

[9] 潘振忠．基于神经网络的图像超分辨率重建模型的研究 [D]．中国科学技术大学，2020.