

基于迁移学习的垃圾图像分类应用

雷豪洋 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】随着社会的不断发展,人们生活水平的不断提高,生活方式的极大改善,使得垃圾处理问题越来越受到重视。如果没有对垃圾进行科学分类处理,反而是随意丢弃,那么对于土壤,水资源等人类赖以生存的环境会有很大程度的污染。垃圾分类则是有效地缓解环境污染问题的关键环节。迁移学习可以有效地解决在没有足够数据的情况下模型训练会产生过拟合的问题,本文使用基于预训练网络的迁移学习方法,将垃圾图像数据集中的80%作为训练集,剩下的作为测试集,来训练模型以实现垃圾图像分类。通过对不同模型在相同参数下的实验数据的对比,选择ResNet152作为预训练网络模型,测试准确率最高可达到82%。

【关键词】迁移学习;ResNet152;垃圾分类

1 引言

随着社会经济的发展,人民收入的提高,消费的能力也逐渐提高,购买力的增强随之而来的是产生的垃圾会越来越多,但是如果将产生的生活垃圾随处丢弃或者时处理方式不当,将会对人们赖以生存的土壤,水源,空气等环境因素造成极大的污染,所以对产生的垃圾进行合理地处理,是减少垃圾污染环境的有效手段。目前根据2021年正式施行的《成都市生活垃圾管理条例》将垃圾的分类划分为可回收垃圾,厨余垃圾,有害垃圾和其他垃圾四类。但目前国内关于这些垃圾类别的图像识别和分类的相关技术还不太成熟,关于垃圾分类算法的应用都不广泛。早期有不少研究提出基于SEM模型,树莓派等技术来实现机器人对垃圾图像的识别,但存在着精度低,运行慢,难训练等问题^[1]。随着卷积神经网络的日益成熟,在图像识别的应用中也加入了深度学习,虽然有了深度学习算法的帮助,Region-CNN, YOLO等模型或多或少的解决了上述的一些早期问题,但是依旧存在着难训练的困扰。因此本文采用对比GoogLeNet

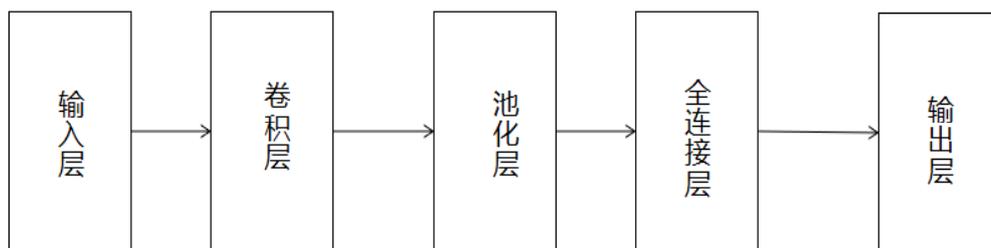
和ResNet152两种容易训练的神经网络模型,选择其中拥有较高精度的一种作为预训练网络模型。

2. 理论知识

2.1 卷积神经网络

卷积神经网络是在传统的神经网络上进行改进的多层神经网络,同时它也是一类由包含卷积计算和具有深度结构的前馈神经网络^[2]。经过不断的改进和完善,成为了深度学习的代表算法之一。卷积神经网络的每一层都是有单个且独立的神经元组成,它的隐含层由卷积层,池化层和全连接层组成。卷积神经网络提供的特征具有一定程度上的空间不变性是它和其他神经网络最大的不同。造成这种不同是由于卷积神经网络的卷积操作和采样操作是基于人工神经网络^[3]引入的。正因为提取到的特征的空间不变性所以卷积神经网络在图像识别和分类处理的领域应用非常广泛。并且卷积神经网络与前馈神经网络等其他网络相比,所需要考虑的参数就显得较少^[4]。以下为卷积神经网络的结构,如图1所示。

图1 卷积神经网络结构示意图



2.2 GoogLeNet

GoogLeNet是巨头互联网公司Google设计的划时代意义的基于Inception模块的深度学习神经网络模型,在2014年的ImageNet比赛中脱颖而出。GoogLeNet的创新性在于它的框架不同于以往的线性模型的框架,而是采用的并行的分支结构^[5]。有效解决了传统网络模型提取的特征不足的情况。

Inception就是把多个卷积或池化操作放在一起组

成一个网络模块,然后通过组成的网络模块为单位形成整个网络结构来设计神经网络。在实际情况下,不同尺度的图片需要不同大小的卷积核使得网络模型的性能更好,而Inception模块中并列地提供了多种卷积核大小的操作,能满足网络自己选择不同尺寸的卷积核的要求。而且网络中池化操作必不可少,所以池化层也加入到网络中与不同的卷积核并列。以下为Inception模块结构,如图2所示。

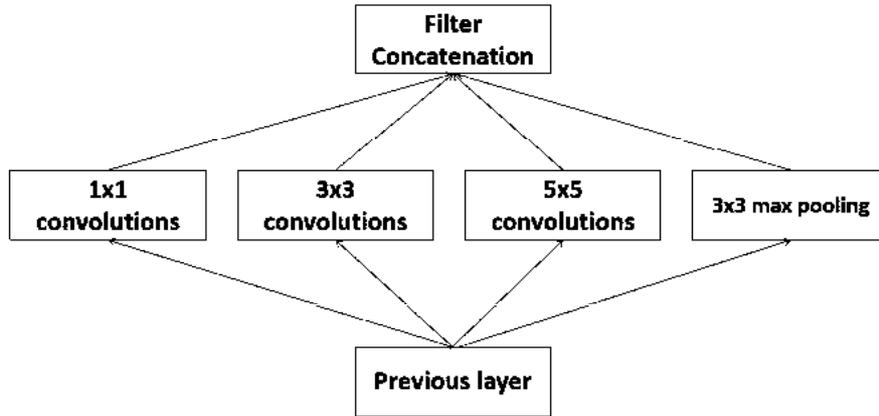


图 2 Inception 模块结构示意图

2.3 ResNet

ResNet 引入了残差学习解决难以优化深度网络的问题，是 Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun^[6] 提出的。它是一种获得比以往更深的网络模型结构^[7]。一般情况下，我们对网络结构采用将卷积层和池化层两两叠加的方法加深网络结构或者通过加深特征图来增加网络宽度，是为了能够获得鲁棒性

更强 [8] 的特征，但是增加宽度或者是深度都需要一个上限，否则就会造成网络结构过深而造成的梯度消失和梯度爆炸，亦或者是网络过宽时产生的降低运算效率等问题。而 ResNet 的提出就是为了解决增加深度所带来的网络退化问题，使得能够单纯的通过加深网络深度从而提供网络性能。如图 3 所示为 ResNet152 网络结构示意图

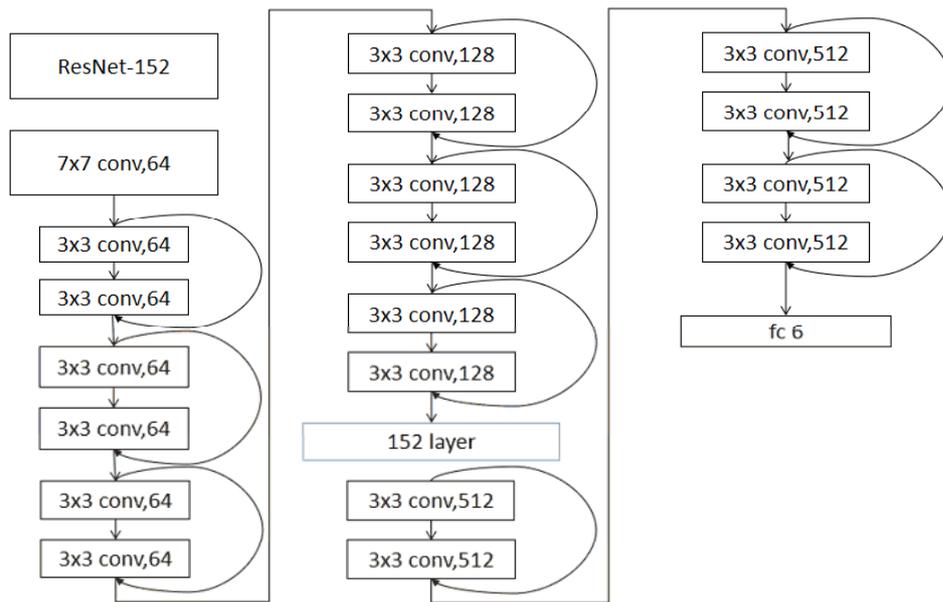


图 3 ResNet152 网络结构示意图

3. 应用实现

3.1 迁移模型训练流程

首先将数据集中的包含所有图像信息按照 80% 作为训练集样本，20% 作为测试集样本，实现对整个数据集的划分。将训练集和测试集中的图像信息分别写入 train.csv 和 test.csv 文件中。根据文件名分别读取训练集图像信息与测试集图像信息，并对对应的图像进

行预处理，包括将图像大小调整为相同大小，对图像进行增强处理等，然后加载 ResNet152 预训练模型，接着执行冻结网络，最后替换最后一个全连接层，将模型放在 GPU 上，然后将训练数据和测试数据也放在 GPU 上进行训练。迁移模型训练的流程如下图 4 所示。

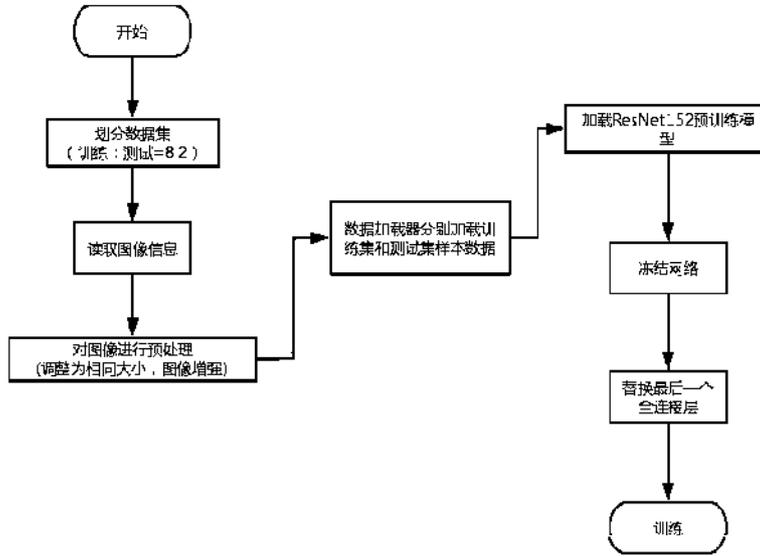


图 4 迁移模型训练流程图

3.2 实验数据与预处理

采用华为垃圾分类挑战赛数据集作为样本数据集，数据集中共有 14802 张样本图片，包含其他垃圾，厨余垃圾，可回收垃圾和有害垃圾四个大类，其中其他垃圾包括一次性快餐盒，污损塑料等 6 个小类别，厨余垃圾包括剩饭剩菜，水果果皮等 8 个小类别，可回收垃圾包括充电宝，易拉罐，纸板箱等 23 个小类别，有害垃圾包括干电池等 3 个小类别。将图像数据集中的 80% 作为训练集，剩下的 20% 作为测试集。为达到实验要求，主要根据居民日常生活中所产生的垃圾进行分析。由于日常垃圾种类复杂多样，并且这些垃圾的颜色，形状等有

很大的差别，所以我们不细分该垃圾的类别，根据选取的数据集分为常见的四个大类。这四个种类的垃圾分别是可回收垃圾，有害垃圾，厨余垃圾和其他垃圾。数据集内垃圾图像种类如图 5 所示。

根据实验要求对选取的训练集和测试集对图像进行预处理，首先将训练集和测试集图片的大小都调整到 256x256，并对训练集图片使用在图像的中心区域进行裁剪、随机水平翻转、修改对比度和饱和度、归一化、标准化等方式进行图像增强，而测试集的图像增强则是按照中心裁剪，归一化、最后将图像数据标准化的方式进行。



图 5 数据集内垃圾图像种类

3.3 网络对比

将 GoogLeNet 与 ResNet152 实验结果进行对比，通过比较两个网络模型在相同的 epoch, batch size, learning rate 的参数下的训练准确率，选择准确率较高的一个作为实验的网络模型。下表 1 为两种网络模型实验数据对比。

表 1 GoogLeNet 与 ResNet152 实验数据简略对比

	Epoch	Batch_ Size	Learning_Rate	Test_ Acc
GoogLeNet	3	32	0.005	0.6499
ResNet152	3	32	0.005	0.7208

通过实验数据可以得出，在上述三个参数相同的情况下，ResNet152 的测试准确率高于 GoogLeNet 的测试准确率，所以选择 ResNet152 作为预训练网络模型。

3.4 框架

模型训练与测试均是在 Python 的深度学习框架下进行的。硬件环境: Intel(R) Core(TM) i7 10870H @2.90GHZ 2.90GHZ, NVIDIA GEFORCE RTX 2060。软件环境: Anaconda3+cuda11.1 +pytorch1.6+Windows10。在本次实验中, 主要对 Epoch(被训练的次数), Batch_Size(每次提交网络数据的数量), Learning_Rate(学习率) 三个参

数进行迁移模型训练。

Batch_Size: 每次提交的网络数据数量, 对 Batch_Size 进行调整可以有效解决数据量过大的问题。Batch_Size 过大时, 会造成模型运行速度变慢。所以最佳的 Batch_Size 可以有效平衡容量和效率, 使得网络模型的性能和运行速度最佳^[9]。

Epoch	Batch_Size	Learning_Rate	Train_Loss	Test_Loss	训练准确率	测试准确率
5	32	0.005	1.4737	1.7679	0.7128	0.7035
5	16	0.005	2.3462	2.7636	0.668	0.7004
5	8	0.005	3.879	3.6871	0.6051	0.696
5	32	0.001	0.7993	0.7283	0.7523	0.7912
5	32	0.01	2.8286	4.0709	0.6804	0.6677
5	32	0.0001	1.1781	1.0701	0.7202	0.7295
10	32	0.001	0.7012	1.041	0.7832	0.7345
20	32	0.001	0.6399	0.6135	0.8006	0.8267

3.5 训练结果与分析

在 ResNet152 预训练网络模型下进行试验, 随机将图片集中的每个类别垃圾图片按照训练集占比 80% 和测试集占比 20% 进行划分, 再将 Epoch, Batch_Size, Learning_Rate 分别选用不同的数值进行试验。得到的实验数据如下表 2 所示。

表 2 实验数据表

从表中数据可以看出 ResNet152 的测试准确率可达到 82%。因为增加 Epoch 必然会使得时间增加, 训练网络时的耗时不足以作为评判标准, 所以综合所有 Epoch=5 的数据的测试准确率相比, 数据组合 (5, 32, 0.001) 的性能最好。且根据实验数据对比可以得出, 当 Epoch, Batch_Size 不变时, Learning_Rate 增大那么测试准确率降低, 反之则提高, 但是如数据组合 (5, 32, 0.0001) 时, Learning_Rate 的值过于小, 那么测试准确率也依旧会降低; 当 Batch_Size, Learning_Rate 不变时, Epoch 值增加, 将会使测试准确率提高, 反之则降低; 当 Epoch, Learning_Rate 不变时, Batch_size 值越小, 由三个数据组合 (5, 32, 0.005), (5, 16, 0.005) 和 (5, 8, 0.005) 对比可以看出。虽然测试准确率没有太大的波动, 但是训练集和测试集的损失函数值将会增加。图 6 为数据组合 (5, 32, 0.001) 的损失曲线。图 7 是数据组合 (5, 32, 0.001) 的训练准确率与测试准确率曲线。

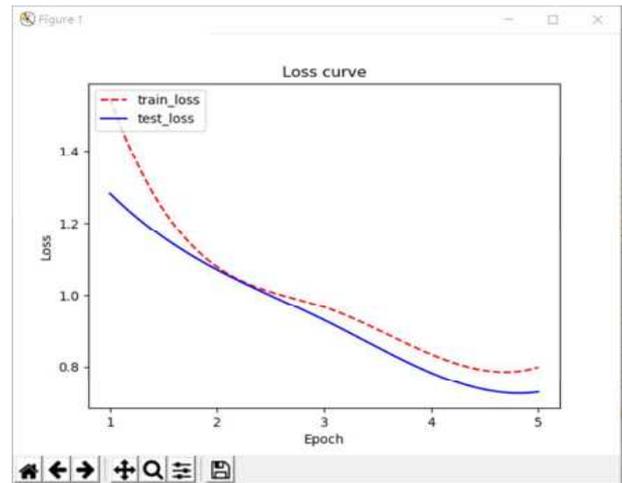


图 6 数据组合 (5, 32, 0.001) 的损失曲线

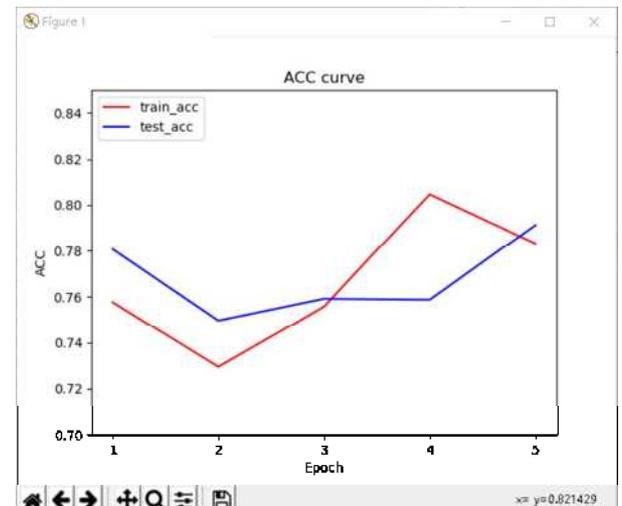


图 7 数据组合 (5, 32, 0.001) 的训练准确率和测试准确率曲线

4. 总结

4.1 优缺点

优点: 根据实验数据可以看出, ResNet152 的训练准确率比较高, 有效地解决了网络退化问题, 也就是当网络深度增加的时候, 性能反而下降的问题, 以及使得前向传播和反向传播算法非常顺利进行, 在极大程度上, 使得优化较深层模型更为简单。

缺点: 根据实验结果可以看出, ResNet152 的学习结构对网络权重的波动很敏感。网络权重的稍微变化就会大幅度降低训练效率。

4.2 应用前景

垃圾分类是一个有着良好前景的环保产业 [], 随着各省市政府对垃圾分类的政策出台, 人工智能, 大数据等技术不断创新。需求就会带来市场, 垃圾分类的市场将会因此得到极大的拓展。而人工智能 + 垃圾分类将是垃圾分类行业一种新的发展趋势, 它可以有效地解决现有垃圾分类行业中存在的诸多问题, 比如可以提高垃圾分类的效率, 降低垃圾分类处理的人力成本等。用人工智能进行垃圾分类是不能好高骛远的, 它的成效也是不立竿见影, 这项技术需要不断地创新, 不断地突破, 才能使垃圾分类行业蓬勃发展。

【参考文献】

[1] 戴冰燕, 郑志硕. 基于迁移学习的可回收垃圾识别分类研究 [J]. 中国科技信息, 2021 (05): 86-88.

[2] Goodfellow, I., Bengio, Y., Courville, A. Deep learning (Vol. 1) [M]. Cambridge: MIT press, 2016: 326-366.

[3] 王志明. 数字图像处理与分析 [M]. 北京: 清

华大学出版社, 2012: 184-204.

[4] Collobert, Ronan; Weston, Jason; Bottou, Leon; Karlen, Michael; Kavukcuoglu, Koray; Kuksa, Pavel. Natural Language Processing (Almost) from Scratch [J]. Journal of Machine Learning Research, 2011 (12): 2493-2537.

[5] 李英, 宋丽娟. 基于 GoogLeNet 模型的遥感图像识别分类研究 [J]. 电脑知识与技术, 2021 (12): 4-6.

[6] He K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.

[7] 杜云梅, 黄帅, 梁会营. 基于深度卷积神经网络的脑电图异常检测 [J]. 华南师范大学学报 (自然科学版), 2020, 52 (2): 122-128.

[8] 易峰, 胡馨莹. 基于深度残差网络的行人人脸识别算法研究 [J]. 电脑知识与技术, 2018, 14 (23): 233-235.

[9] 郑誉煌, 戴冰燕, 熊泽琿, 陈钊. 基于迁移学习的可回收生活垃圾图像分类识别研究 [J]. 广东第二师范学院学报, 2020, 40 (03): 94-100.

[10] 王颖, 江苏省城市垃圾分类回收的现状 & 政策探讨 [J]. 资源节约与环保, 2021 (01): 75-78.