

GoogleNet 迁移学习用于垃圾分类

谭舜天 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】随着人们生活水平的提高,生活中的垃圾比过去也变得更加繁多,怎样通过更加智能的方法进行垃圾的分类工作,来节约更多的人力和物力资源使经济得到更好地发展。本文基于迁移学习来构建图像分类模型,用于识别部分垃圾的类别,按照垃圾分类的标准分成可回收物、厨余垃圾、有害垃圾和其他垃圾四种别,最终模型在测试集上有 70% 的准确率。

【关键词】迁移学习;深度学习;GoogleNet;垃圾分类

1 研究背景

随着人们生活水平的提高,对生活有了更高的追求,人们生活中的垃圾也与日俱增,如果我们产生的生活垃圾被随处丢弃或者回收处理不当,会对我们生活的环境造成很大的破坏。垃圾分类在这两年成为了我们国内的一个热点的问题,将成堆的垃圾分成不同的种类,其中可回收物的循环利用对我国的经济发展有很重要的意义,通过从人力到智能的演变可以提高垃圾回收的效率。通过图像处理和深度学习等方法自动识别生活中常见的生活废品。

迁移学习的核心是卷积神经网络,比机器学习具有更好的特征学习和表达能力,例如: ResNet、VGG19 等,不同的网络结构训练的效果也会有所不同。为此本文主要利用迁移学习的方法将 GoogleNet 的预训练模型来迁移学习垃圾分类,通过对不同参数的调节,希望可以在测试中得出更高的准确率。

2 迁移学习及其网络结构介绍

2.1 迁移学习的概念

迁移学习是一种机器学习方法,就是在一个完成的模型基础上,重新使用在另一个任务的开发模型过程中^[1]。在本研究中我们主要用的就是预训练模型的方法,第一步是选择原本的模型,第二步是重新使用模型,最后一步就是调整模型使其更加适应我们的任务。

2.2 GoogleNet 模型结构

网络的深度是觉得网络好坏的因素之一,随着网络深度的增加输入特征的层数也会变高变高,再增加的

同时,又会出现梯度消失的困难。在加入了正则化处理之后,可以有效的解决无法收敛的问题,但也同时造成了训练集的准确率下降的问题,想要获得质量好的模型最稳妥的做法就是增加模型的深度或者宽度,但是也会出现一定的缺陷。

在 2014 年 ImageNet 比赛中胜出的是由 Google 公司设计的划时代意义的 GoogleNet 模型,其使用了不同于之前的线性模型框架,解决了传统网络特征提取不足的情况。

GoogleNet 是一个 22 层的深度网络是在 AlexNet 的基础上来进行修改的,GoogleNet 通过计算稀疏结构来减少由于层数的增加造成的过拟合问题。GoogleNet 使用了多种数据增强的方式,引入了 BN 正则化技术,使用这个技术可以减少网络冗余的结构,提出了 Inception 模块,Inception 的主要作用就是由网络去自行决定选择使用哪个过滤器,是否需要池化操作等,网络也可以选择自己学习需要的参数。其内的卷积核采用 1×1 , 3×3 和 5×5 的大小是为了方便对齐,该网络结构越到后面特征就更加的抽象,所以说随着层数的增加, 3×3 和 5×5 的卷积比例也要增加。为了减少计算成本,采用 1×1 的卷积核来进行降维。从降维角度来说,一方面解决了计算困难的问题,同时也限制了网络的最大参数,网络才能做的更深。对于输入层,使用 1×1 卷积把通道从 192 个减少到 16 个。然后对这个较小层进行 5×5 卷积,得到最终输出。Inception 其基本的结构如图 1 所示:

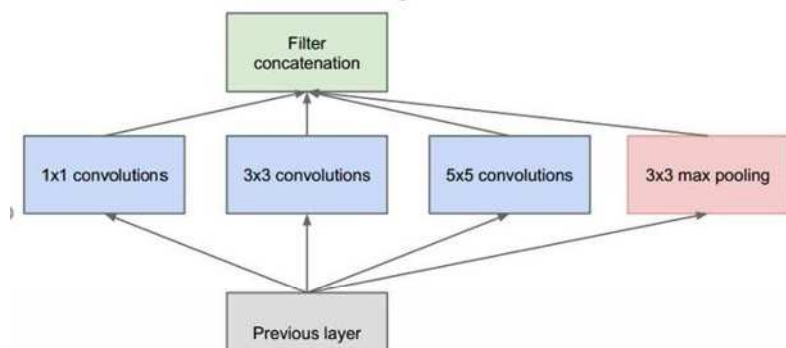


图 1 inception 的基本结构

Googlenet 采用的是模块化的结构方便在进行过程中添补和修改内容网络最后采用了平均池化来代替全连

接层,原理来自 NIN。但是,实际在最后还是加了一个全连接层,主要是为了方便以后大家微调。虽然移除了

全连接层，但是在网络中进行了随机丢弃。为了避免梯度消失，网络额外增加了 2 个辅助的回归模型用于向前传导梯度^[1]。

2.3 ResNet 概念

随着网络结构的增加，出现准确率下降的问题，然而不是由于过拟合造成的，在 2015 年提出了 ResNet 来解决问题。ResNet 是一种全新的网络，又叫深度残

参网络^[2]，主要是为了解决网络退化的问题，残参网络可以看作是一组路径组装而成的集体模型，不同的路径包含了不同的网络层子集，在删去部分网络层或者交换网络模块的顺序之后，可以得出网络的表现与正确网络路径数相关，而在训练中贡献了梯度的是那些相对短的路径。下面图 2 为两种的残参网络的设计结构图。

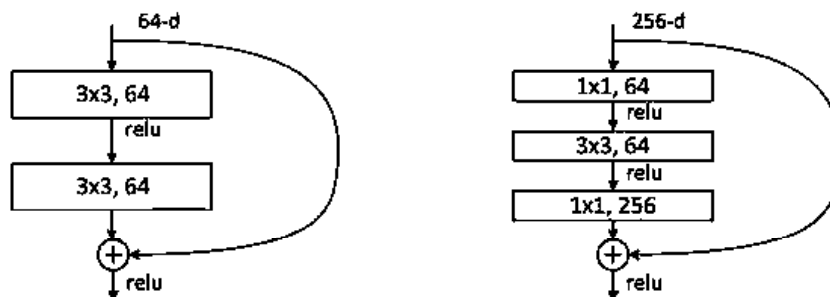


图 2 两种残参结构

3 实验过程及结果

3.1 实验数据集

本研究的数据集以 garbage_classify 数据集为基础，该数据集由共计 14802 张图片组成了 40 种不同物品，

其中又将它们划分成可回收物、厨余垃圾、有害垃圾和其他垃圾四个大类。每种垃圾的图像都有不同种的外观，又将这些图像进行了增强处理，随机水平翻转、随机垂直翻转，随机裁剪以此来最大化数据集的规模。垃圾样本的截图如下：

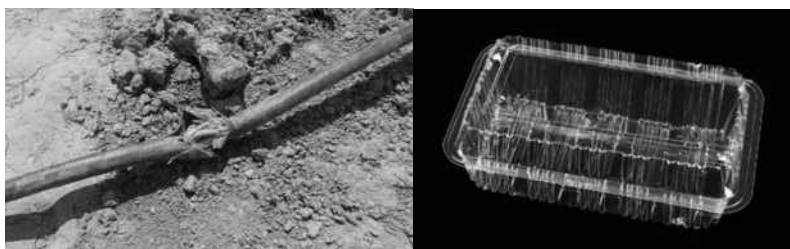


图 3 可回收物（废弃电线）图 4 其他垃圾（打包盒）



图 5 有害垃圾（过期药物）图 6 厨余垃圾（鸡蛋壳）

3.2 实验过程

下面图 7 为实验过程的流程图：

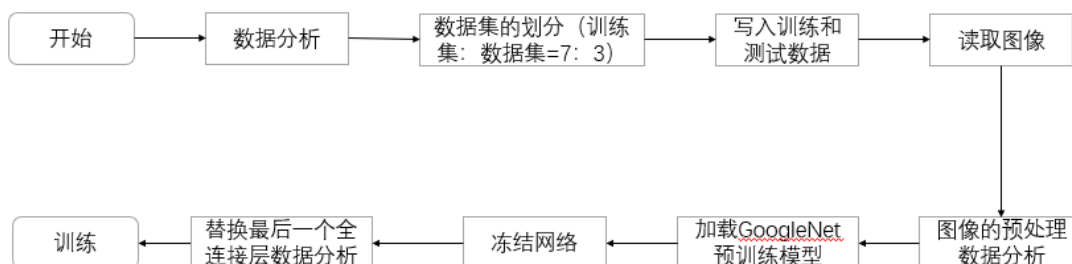


图 7 实验流程图

首先我们要导入我们所需要的库, 首先将数据集中的所有图像按照 70% 作为训练集样本, 30% 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到两个文本文件中。根据文本名分别读取训练集图像标签与测试集图像标签, 首先将训练集和测试集中的图片的大小都调整到 256x256 的大小, 并对训练集图片在图像的中心位置进行裁剪, 概率水平翻转, 修改图片对比度和饱和度, 将图像数据进行归一化、标准化等方式进行图像增强预处理, 而测试集的图像增强预处理则是按照图像预处理中心裁剪, 将图像数据进行归一化处理, 最后将图像数据标准化进行。

下一步就是进行模型选择了, 经过测试已经实际的运行情况选择的是 GoogleNet。首先是加载预训练模型, 接着执行冻结网络, 最后替换最后一个全连接层并将该模型加载到 cuda 上, 在实验的时候使用 GPU 进行加速, 可以提高速度。接着就是来设置损失函数了, 损失函数是将随机事件或随机变量的取值映射为非负实数以表示该随机事件的“风险”或“损失”的函数^[1]。接着对训练集数据和测试集数据进行训练, 分别给对应的数据集进行前向传播^[2]和反向传播^[3]。

最后经过反复测试和修改参数我们得到了下面的表格 1 中, 我们可以将 epoch=10, batchsize=16, lr=0.001 的情况下得到的准确率作为一个模板, 在接下来的实验中我们通过控制变量法, 每次改变一个参数进行训练, 得出了在 epoch=20, batchsize=16, lr=0.001 的参数之下为本次测试最好的准确率。

表 1: 训练参数和测试结果

epoch	batchsize	Learning rate	Train_acc	Test_acc
10	16	0.001	0.6704	0.7029
20	16	0.001	0.6970	0.7107
10	32	0.001	0.7162	0.7029
10	16	0.0001	0.6385	0.6633

4 总结

4.1 优缺点

本文主要采用的是 GoogleNet 预训练模型, 采用迁移学习的方法是在我们自定义的数据集上使用, 替换原模型的全连接层, 构建了基于迁移学习的垃圾分类问题, 训练的准确率虽然不是特别的高, 但是在不停的调参之下, 还是逐渐的趋于一个稳定的值。利用原来模型的学习情况, 之后会考虑扩大数据集或者增加相应的垃圾图片的种类, 进一步来提高模型的准确率。GoogleNet 相比于其他的网络结构其不同之处都是在围绕着减少层数和提高网络结构的性能来设计的。其次, GoogleNet 没有全连接层, 运算的时间减少, 也减少了一些参数。

但是在增加了网络的规模的同时, 即增加了网络的深度和宽度也有了一些新的问题。比如说, 网络的更容易出现过拟合的问题, 其次是计算资源变得更多。

4.2 应用前景

这两年政府越来越重视垃圾分类的相关问题, 模型在产生后可以开发成在线的可回收垃圾的分类系统, 将此模型应用在服务器上只需要加载一次模型, 就可以无限次的使用这个系统, 定期的增加更多的垃圾图片来训练该模型, 从而使其识别的准确率提高到更高的水平。

5 结束语

近年来迁移学习已经得到了广泛的关注和认可, 运用已有的知识对未知的领域进行探索是迁移学习的特点。方便进行信息的分类也是其主要的特点之一, 现在我们所处的社会是网络发展十分迅速的社会, 我们每天都要接收并处理大量的信息, 我们对于信息的处理能力远不及计算机所能处理的速度, 所以说我们就可以借助迁移学习将已经存在的模型加入到我们需要处理的数据中去, 通过分类信息, 可以快速有效的进行处理, 为我们的工作带来更多的便利。

【参考文献】

- [1] Pratt, L. Y.; Thrun, Sebastian (July 1997). "Machine Learning - Special Issue on Inductive Transfer". link.springer.com. Springer. Retrieved 2017-08-10.
- [2] 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, 17(12): 4-6.
- [3] 刘建伟, 赵会丹, 罗雄麟, 许望. 深度学习批归一化及其相关算法研究进展 [J]. 自动化学报, 2020, 46(06): 1090-1120.
- [4] 高珊. 基于高层语义特征的图像检索算法研究 [D]. 长春工业大学, 2018.
- [5] Deep Residual Learning for Image Recognition[C]// IEEE Conference on Computer Vision & Pattern Recognition. IEEE Computer Society, 2016.
- [6] 卢丹丹. 基于深度学习的行人重识别技术 [D]. 南京邮电大学, 2019.
- [7] 谢四江, 许世聪, 章乐. 基于同态加密的卷积神经网络前向传播方法 [J]. 计算机应用与软件, 2020, 37(02): 295-300+312.
- [8] 翟高粤. 基于优化反向传播算法的物体识别技术 [J]. 甘肃科技纵横, 2021, 50(03): 5-7.