

基于迁移学习的手写数字识别优化

张世鸿 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】在一个数据类型和数据量的不断增加的时代,对要求机器学习模型的要求越来越高,需要模型拥有能够快速进行构建的能力和很强的泛化能力,迁移学习就能适用与这些情况。发现 MNIST 数据集采集的是外国人手写数字,而不同国家的书写习惯也不一样,所以如果把在 MNIST 上训练的模型来预测我们手写的数字,在实际应用中的发现准确率不高,于是本文加入迁移学习来提升模型,首先收集笔者自己的手写数字图片数据,然后对已在 MNIST 数据集上训练好的卷积神经网络模型进行自采集数据集的迁移学习,最后实验证明进行迁移学习后在实验者采集数据集上,模型准确率提高了 7%。

【关键词】卷积神经网络 (CNN); 手写数字; 迁移学习

1 引言

人工智能在当今时代被广泛关注,字符识别技术广泛应用于实际生活中。其中手写数字识别技术是相对基础但却重要的部分,在平时手写数字笔记的各种问题中,比如想对已记录数字信息的修改,查询记录数字段中的某个部分,或者复制其中一部分,这些在纸上都无法很好的实现,有了手写数字识别将手写输入的数字存在电脑里就可以很好解决这些问题。由于不同国家书写习惯不同,发现 MNIST 数据集中的手写数字与实验者自己采集的手写数字在形状上有不小的差别,所以将 MNIST 数据集训练的模型用在对写法有差异的数字上实际效果没有预期效果好,采用迁移学习来对模型进行优化可以得到很好处理。

2 关键技术

2.1 卷积神经网络

卷积神经网络^[1]在 1989 年就已经被 LeCun 使用在多层神经网络中,但在后面的十年中因为当时的硬件计算能力完全不够,一直处于低谷状态,在后面随着硬件的发展,而且计算量越来越大使得卷积神经网络再一次回到人们的视野中。卷积神经网络利用卷积核来降低参数目录,这样后的参数还是会过多,再利用权值共享来进一步降低参数目录,减小了权值的数量和降低了过拟合的风险。卷积神经网络由一系列的输入层,卷积层,池化层,全连接层组成^[2]。卷积神经网络相对与其他深度、前馈神经网络拥有更高的性能但却只使用了相对更少的参数,也因此使其在近些年来发展迅速,是备受重视的一种高效识别算法。

2.2 卷积层

增强特征信息,过滤不相干的信息^[3]是卷积层最主要的作用,从而有卷积层的出现。卷积神经网络中最重要的部分就是其中的卷积层,卷积层的核心是卷积。卷积就是两个函数的一种运算,这种运算叫卷积运算。在处理图像识别问题里卷积核与二维图像进行卷积操作,也就是卷积核在二维图像每个位置上与该位置做内积。利用不同的卷积核可以卷积操作图像处理领域中提

取图片中不同的特征。卷积拥有局部感知的机制和权值共享^[4]的特性。

2.3 池化层

由于直接使用这些在卷积层获得到图像特征训练分类器计算量会非常大,而且很容易产生过拟合现象,所以引入池化层来解决这一问题。池化 (Pooling)^[5]又称下采样,池化的主要作用是在保留主要特征的同时通过减少网络的参数从而达到减少计算量的目的,并且能够在一定程度减小上过拟合的情况和提高模型泛化能力。常用的池化方式有最大池化、均值池化、全局最大(或均值)池化^[6]这三种。

2.4 全连接层

全连接层处于卷积神经网络网络的尾部,全连接层对之前卷积层或池化层中有类别区分性的部分信息进行整合。全连接层之前各层的作用是提取局部特征,而全连接层的作用则是将之前层提取的局部特征通过权值矩阵再次装成完整的图,其主要作用就是能减少不同位置的特征对整个分类结果的影响。

2.5 迁移学习

利用迁移学习来解决难获取卷积标记是很好的方法^[7]。计算机在处理视觉任务和自然语言中利用迁移学习是很常用的,这些问题的模型往往需要大量数据和复杂的网络结构。简单来说,利用数据、任务、模型间的相似性简单来将训练好的内容应用到新的任务上就是迁移学习。其中被迁移的对象被称为源域,被赋予“经验”的领域被称为目标域。不是在任意的源域和目标域之间都可以进行迁移学习,如果源域与目标域完全不相似,反而可能造成“负迁移”,就是指在源域上学的东西在目标域上反而起到负面的作用。与传统的机器学习相比,迁移学习的训练和测试数据服从不同的分布,并且也不需要像传统机器学习那么大的数据标注来训练模型,传统机器学习每个任务都要建里不同的模型,但是有了迁移学习,一个模型可以在多个不同的任务之间进行迁移。

3 应用实现

3.1 数据集对比

MNIST 数据集中收集了 60000 张训练图像和 10000 张测试图像。每张图片分别显示 0-9 中的某一个数字，样本中还包含了各个训练数据相应的标签，其中标签集包含了 0-9 这 10 个分类数据^[8]。其中的数据已经经过归一化处理了，采用大小相等的灰度图像，MNIST 数据集中的部分手写字符展示如图 1 所示。

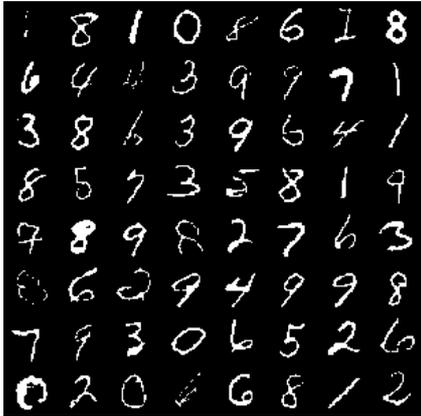


图 1 MNIST 数据集图片展示

实验者通过手写输入采集了训练集图片 0-9 各十张，测试集 0-9 各 5 张。部分手写字符如图 2 所示。创建一个 512*512 的画布，设置其背景色为黑色，通过点

击并拖动鼠标在画布上留下拖动痕迹为白色，进而可以实现对手写数字的输入，松开鼠标即可停止画布，再将整个画布以图片的形式保存在本地文件夹，将作为训练集和测试集的图片分别放在两个文件夹中，两个文件夹中又分别有对应 0-9 十个数字的文件夹，将对应手写数字图片放入对应的文件夹，作为实验者手写数字的训练



图 2 实验者手写数据集图片展示

观察对比 MNIST 数据集和实验这手写数据集中的图片可以明显发现，两个数据集中的数字拥有显著差异，相同的数字其他国家的人和我们的人写法上有很大的区别，从而用 MNIST 训练集训练出来的模型对于我们手写输入的数字的准确率相对会降低，从而引出迁移学习将训练好的模型作为新模型的起点，从而提高准确率。

3.2 模型网络结构

本文采用 LeNet-5 网络模型^[9]结构如图 3 所示，1998 年，卷积神经网络大师 LeCun 提出 LeNet，用 LeNet 来实现手写数字识别^[10]。LeNet-5 在 LeNet 的基础上加入了池化层来对输入进去的特征进行筛选。LeNet-5 模型结构以输入层作为结构的开端，中间是一层卷积层加一层池化层再是一层卷积层加一层池化层组成，然后是若干全连接层，最后是输出层。

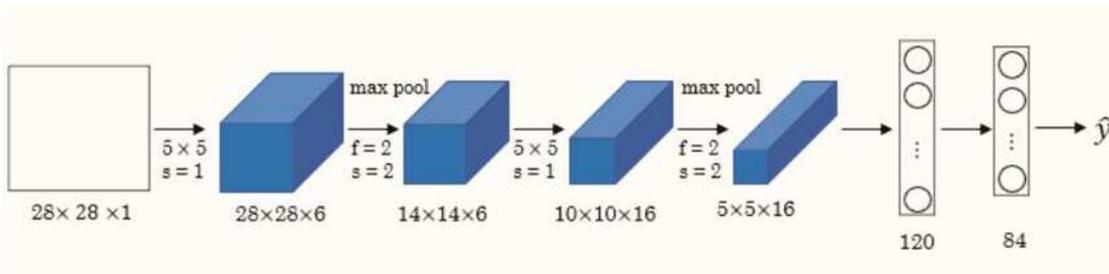


图 1 实验所用 LeNet-5 模型

3.3 训练步骤及测试结果

用 MNIST 数据集进行训练，全连接层输出为 10，选择 SGD 优化器，设置学习率为 0.05，设置 batchsize 为 64，训练 20 轮得到的 loss 值曲线图如图 3 所示。实际的识别效果并没有想象中那么好，如图 4 所示。

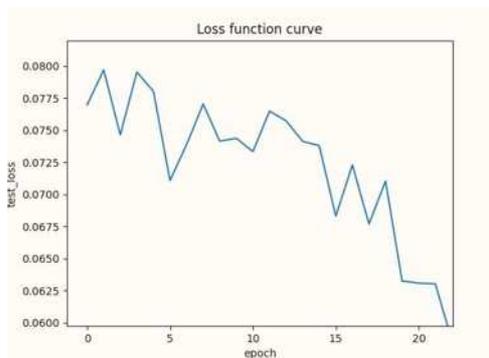


图 4 实际识别结果

实验者手写输入数字对用 MNIST 数据集训练完成的模型进行测试其准确率，得到实际的测试准确率如表 1。

Epoch	1	2	3	4	5	6	7	8	9	10
Acc	0.60	0.64	0.64	0.62	0.68	0.66	0.68	0.68	0.70	0.66

表 1 实验准确率

3.4 迁移学习实验者手写数字数据集

发现在训练多轮后准确率仍不理想,因此在之前建立的模型训练网络的基础上采用迁移学习的方式达到提高识别准确率的目的,将利用 MNIST 数据集训练出的

模型迁移至实验者自己收集的数据集,冻结全连接层之外的所有网络,全连接层输出设置为 10,选择 SGD 优化器,设置学习率为 0.0002,设置 batchsize 为 6,经 1000 轮训练得到的 loss 值曲线图如图 4 所示。

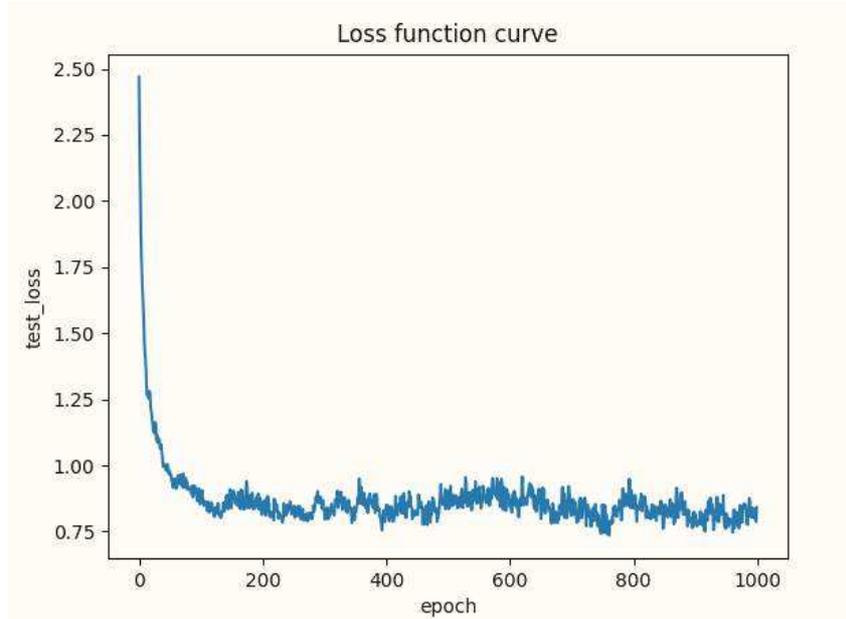


图 5 迁移学习后 loss 值曲线图

迁移学习后新模型的准确率如表 2 所示:

Epoch	100	200	300	400	500	600	700	800	900	1000
Acc	0.63	0.57	0.70	0.75	0.72	0.72	0.73	0.75	0.74	0.77

表 2 实验准确率

3.5 实验结论

对比实验准确率表 1 和表 2,可发现在加入迁移学习后的实验者手写数字数据集上准确率更高。所以利用迁移学习对手写数字识别进行优化,相比与用 MNIST 数据集训练模型的效果有显著提高,手写输入的数字识别准确率提高了 7%,迁移学习减小了训练过程所需的时间,使其不再需要庞大的数据集同时提高深度模型的性能。

论文总结

本文简要介绍了卷积神经网络相关结构以及迁移学习的基本概念,并对卷积池化等操作做了相应的介绍并解释其作用。引用了 LeNet-5 网络模型来实现了手写数字识别的应用,并采集实验者的手写数字图片数据,送入训练好的模型进行预测,检验模型的准确率并不理想,基于此再利用迁移学习达到提高准确率的目的。对于实验的改进可选择优化网络结构,调整学习率,增加实验者收集的数据集大小,增加训练的次数,改变模型的优化器等操作来提高模型准确率。

【参考文献】

[1] 李彦冬,郝宗波,雷航.卷积神经网络研究综述[J].计算机应用,2016,36(09):2508-2515+2565.
[2] 周凯利.应用于手写数字识别的卷积神经网络算法与 FPGA 实现[D].重庆邮电大学,2019.
[3] 韩卫雪,周亚同,池越.基于深度学习卷

积神经网络的地震数据随机噪声去除[J].石油物探,2018,57(06):862-869+877.

[4] 丁小刚. BP 神经网络与卷积神经网络在文字识别中的应用研究[D].华中科技大学,2014.

[5] 朱书眉.基于卷积神经网络的手写体数字识别[J].数字技术与应用,2018,36(11):44+46.

[6] 李彦冬.基于卷积神经网络的计算机视觉关键技术研究[D].电子科技大学,2017.

[7] 来学伟.TensorFlow 读取数据在简单图像识别中的应用[J].现代信息科技,2019,3(12):98-99.

[8] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11):2278-2324.

[9] 吕红.基于卷积神经网络的手写数字识别系统的设计[J].智能计算机与应用,2019,9(02):54-56+62.

[10] 彭先伦.动量梯度下降法的收敛性[D].华东理工大学,2020. [