

DOI:10.19651/j.cnki.emt.1802265

一种改进池化模型对卷积神经网络性能影响的研究

刘梦雅 毛剑琳

(昆明理工大学 信息工程与自动化学院 昆明 650500)

摘要: 池化模型作为卷积神经网络模型中至关重要的一部分,具有降维、提高模型泛化能力等作用。为了进一步提高卷积神经网络模型的准确率,优化模型的学习性能,提出了一种基于最大池化和平均池化的改进池化模型,并在全球手写数字数据集 MNIST 和 CIFAR-10 上分别对改进池化模型的有效性进行了验证。通过与常见池化模型的对比实验发现,采用改进池化模型的卷积神经网络的学习性能较优,一次迭代情况下,在 MNIST 和 CIFAR-10 数据集上,错误率分别下降了 4.28% 和 2.15%。

关键词: 卷积神经网络;池化模型;图像识别;MNIST;CIFAR-10

中图分类号: TP389.1; TN919.81 **文献标识码:** A **国家标准学科分类代码:** 520.1040

Research on the influence of an improved pooling model on the performance of convolutional neural networks

Liu Mengya Mao Jianlin

(School of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China)

Abstract: As a vital part of the convolutional neural network model, the pooling model has the functions of dimension reduction and generalization of the model. In order to further improve the accuracy of the convolutional neural network model and optimize the learning performance of the model, this paper proposes an improved pooling model based on maximum pooling and average pooling, and the global handwritten digital datasets MNIST and CIFAR-10 data. The effectiveness of the improved pooling model was verified on the two dataset. Comparing with the common pooling model, it is found that the learning performance of the convolutional neural network with improved pooling model is better. In one iteration, the error rate decreases by 4.28% on the MNIST and decreases by 2.15% on CIFAR-10 datasets.

Keywords: convolutional neural network; pooling model; image recognition; MNIST; CIFAR-10

0 引言

近年来,由于深度学习在特征提取上具有的显著优势,使其在很多图像识别任务中都发挥了很大的作用。卷积神经网络(convolutional neural network, CNN)对深度学习在图像处理等方面的研究具有十分瞩目的作用。它通过模拟人类视觉系统逐层提取图像深层特征的机制提高网络识别率。而且利用这种方式提取到的特征具有较强的泛化能力,解决了由于图像变换引起的特征提取不准确的问题^[1]。

Alex 等^[2]使用基于最大池化模型的卷积神经网络实现了图像的快速扫描;Lee 等^[3]使用分层生成模型改进卷积神经网络来处理高维图像,大大提高了准确率;Gong 等^[4]利用多尺度无序池化模型解决了全局激活缺少几何不变性的问题。文献^[5]提出了利用主成分分析(principal

component analysis, PCA)算法优化初始权值的方法。除此以外,卷积神经网络的结构、卷积核的大小、池化模型、特征图维数和损失函数等都会影响卷积神经网络的收敛速度和模型性能。文献^[6]比较了不同大小的卷积核对卷积神经网络性能的影响,得出如下结论:较小的卷积核能够提取到更多的原始特征,因而能更好地表示原始图像,但是计算量陡增,模型收敛需要的时间也较多。文献^[7]比较了不同特征图个数下的识别率,发现当特征图的个数较少的时候,网络能够提取到的特征较少,使得识别率较低;但如果特征图的个数太多,识别率会有所提高,但会使网络的结构变复杂,训练时耗也会增大,还有可能出现过度训练现象^[8]。

卷积神经网络在处理数据时,需要经过卷积和池化操作,而池化模型决定了子采样层提取输入数据特征的有效性。本文提出了一种基于最大池化和平均池化的改进池化

算法,对比实验的结果表明,算法可以提高识别率,由此验证了改进算法的有效性。

1 卷积神经网络

卷积神经网络作为一种深度学习的神经网络,它的输入一般为未经过预处理的原始图片。其基本结构包括输入层、卷积层、子采样层、全连接层和输出层,其中,卷积层和子采样层可重复出现。每个特征层由多个二维平面构成,这些二维平面分别由多个神经元构成^[9]。卷积神经网络的网络结构如图1所示。

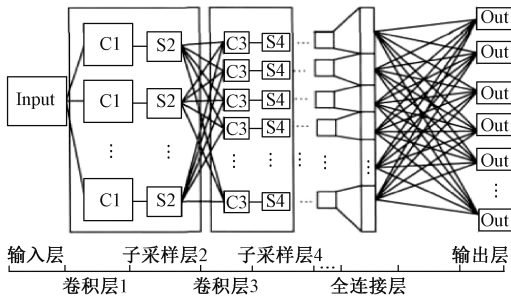


图1 卷积神经网络结构

卷积过程当中,多个卷积核分别对图像的不同感受野进行卷积,从而提取出多个特征图,此时,卷积核参数是与所有的感受野共享的,所以大大减少了参数量,提升了网络的学习性能^[10]。

2 池化模型理论

对卷积神经网络算法而言,加入池化操作,可以在保留主要特征的同时大大减少需要训练的参数。通过减少下一层的参数和计算量,也可以防止过拟合,增强旋转、平移、尺度不变性^[11]。由于池化操作减少了许多参数,所以对特征的提取必然产生很大的影响。本文介绍几种保留主要特征的池化方法,并提出一种基于最大池化和平均池化的改进池化算法^[12]。

2.1 平均池化

平均池化即对池化域内的元素求平均值,然后用每块池化域的平均值构成池化层。平均池化模型的计算公式如式(1)所示。

$$s_{ij} = \frac{1}{c^2} \left(\sum_{i=1}^c \sum_{j=1}^c F_{ij} \right) + b \quad (1)$$

式中: F 表示输入特征图中的一块池化域; F_{ij} 表示 F 中第 i 行第 j 列对应的像素值; c 为池化域的边长; b 为偏置。

平均池化通过计算池化域内的平均值提取特征图的特征,并对输入特征图进行降维。在特征提取的过程中,池化域大小受限会造成估计值方差增大,从而引起误差,平均池化可以减小这方面的误差,更多地保留图片的背景信息。

2.2 最大值池化

最大值池化的主要功能和平均值池化一致,即降维。在删除冗余信息的同时,尽量不影响识别结果。最大值池化即对池化域内的元素求最大值,然后用每块池化域的最大值构成池化层。最大值池化模型的计算公式如式(2)所示,式中的符号表示和平均值池化模型一致,以下不再赘述。

$$s_{ij} = \max_{i=1,j=1}^c (F_{ij}) + b \quad (2)$$

最大池化保留了每一小块池化域中的最大值,以最大值表示这一块区域最佳匹配的结果。所以卷积神经网络经过池化操作之后,不再关注池化域内匹配特征的具体位置,而是只关注是否存在匹配特征。所以,池化操作可以增强旋转、平移、尺度不变性。此外,对于最大值池化,可以减少由于卷积层参数误差造成估计均值偏移引起的特征提取误差,保留特征图更多的纹理信息^[13]。

2.3 中间值池化

中间值池化是基于平均值池化和最大值池化的一种折中算法。它选取池化域内最大值和平均值的中间值作为该池化域的最佳匹配结果。这种池化模型兼顾了平均值池化和最大值池化获取的最佳匹配值,使得识别精度有了一定的提高。中间值池化模型的计算公式如式(3)和(4)所示。

$$s_{ij} = T/2 + b \quad (3)$$

$$T = \frac{1}{c^2} \left(\sum_{i=1}^c \sum_{j=1}^c F_{ij} \right) + \max_{i=1,j=1}^c (F_{ij}) \quad (4)$$

2.4 改进的池化模型

卷积神经网络在训练的过程中,会产生很多特征图,为了更好地提取特征,池化模型的作用至关重要。实验发现,传统的池化模型虽然效果还可以,但仍然有必要进行改进且有很大的改进空间。为了进一步提高池化模型,本文基于平均值池化模型和最大值池化模型,提出了一种改进的池化模型。该池化模型利用平均值池化和最大值池化获取结果的乘积,再引入池化因子 ρ ,则改进的池化模型表达式如式(5)所示。

$$s_{ij} = \rho \cdot \frac{1}{c^2} \left(\sum_{i=1}^c \sum_{j=1}^c F_{ij} \right) \cdot \max_{i=1,j=1}^c (F_{ij}) + b \quad (5)$$

式中: ρ 为池化因子,是本文引进的一个超参数,它的取值受数据集的影响。针对不同的数据集,根据二分法选择最适合该数据集的超参数 ρ 。此式为改进池化模型的基本表达式,其本质是引入了池化因子 ρ ,使得池化模型能针对不同的数据集提取更合适的特征,从而提高卷积神经网络的识别精度。

3 实验与结果分析

本文选用MNIST数据集和CIFAR-10数据集分别对提出的改进池化模型进行训练和性能分析。众所周知,卷积神经网络在数字手写体的识别上效果显著,所以本文首先利用MNIST数据集对几种常见的池化模型和本文提出

的改进池化模型进行测试和对比研究。其次,为了验证改进池化模型在其他数据集上对卷积神经网络性能的影响,本文选用 CIFAR-10 数据集进行了进一步的测试,以验证本文提出的改进池化模型具有一定的普适性。

3.1 MNIST 数据集实验与分析

MNIST 数据集共有 70 000 条数据,选取 60 000 条数据进行网络的训练,其余 10 000 条数据进行测试,输入的图片都是 28×28 的灰度图像,卷积神经网络的识别结果为输入图片中的数字^[14]。训练时采用小批次梯度下降法作为网络的优化算法,设置批度为 50,所以共有 1 200 个批度,即卷积神经网络迭代一次,网络就训练了 1 200 次。实验平台为 Windows 7,运行内存 6 G,算法使用的软件为 MATLAB(2012a)。

针对传统的卷积神经网络(网络结构与 LeNet5 一致)分别采用平均值池化模型、最大值池化模型、中间值池化模型、本文提出的改进池化模型进行实验。其中,对于改进的池化模型,经过实验,池化因子取值 2.45。对每一种池化模型根据迭代次数进行实验,训练的迭代次数分别为 1、2、3 次。记录 4 种池化模型训练完后的误识率如表 1 所示,训练过程中均方误差随训练批次的变化曲线如图 2 所示。

表 1 4 种池化模型对 CNN 性能影响的比较(MNIST)
(%)

迭代次数	平均值池化	中间值池化	最大值池化	改进池化
1	11.19	11.58	10.03	6.91
2	7.88	7.38	5.80	4.24
3	6.28	5.31	5.00	3.46

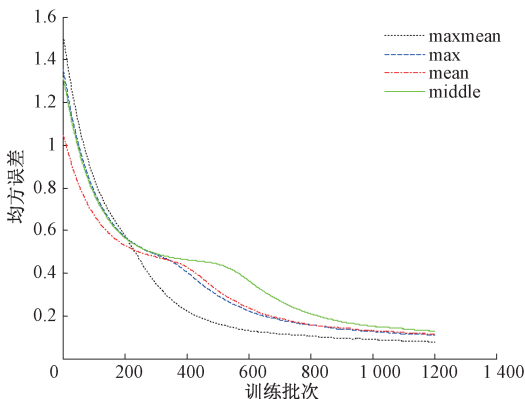


图 2 4 种池化模型在 MNIST 数据集上的均方误差曲线(MNIST)

由表 1 可见,随着迭代次数的增加,每一种池化模型都会不断地降低误识率,但改进池化模型的效果很显然比其他模型好很多。卷积神经网络迭代一次时,平均值池化、中间值池化、最大值池化的误识率分别是 11.19%、11.58%、10.03%,而改进池化模型的误识率仅为 6.91%。迭代两

次时,平均值池化、中间值池化、最大值池化的误识率分别是 7.88%、7.38%、5.8%,而改进池化模型的误识率仅为 4.24%。迭代 3 次时,平均值池化、中间值池化、最大值池化的误识率分别是 6.28%、5.31%、5%,而改进池化模型的误识率仅为 3.46%。与当下各种网络、框架等最常用的最大值池化模型相比,改进后的池化模型在迭代 1、2、3 次情况下,误识率分别下降了 31.106 7%、26.896 6%、30.800 0%。与平均值池化、中间值池化相比,效果更是好很多。

图 2 展示了迭代一次时,均方误差随批次变化的曲线。图例中的 mean 代表平均值池化对应的均方误差曲线;max 代表最大值池化对应的均方误差曲线;middle 代表中间值池化对应的均方误差曲线;maxmean 代表本文提出的改进池化模型对应的均方误差曲线。由图 2 可知,随着训练批次的增加,每一种池化模型对于的均方误差都不断下降。但是训练批次在 200 到 600 之间时,相比于其他池化模型,改进池化模型对应的均方误差的下降趋势得到了改善,从图 2 可以很直观地发现,改进后的算法均方误差下降迅速,收敛速度也有所提高。平均值池化、中间值池化、最大值池化在批次 200 到 600 之间出现了一个平台期,这段时间均方误差的下降趋势大大减缓,而改进后的池化模型在批次 200 到 400 之间时均方误差的下降趋势比较平滑,下降得更快,平台现象已经基本消失。此外,改进后的池化模型对应的均方误差在迭代一次后较其他池化模型小。

3.2 CIFAR-10 数据集实验与分析

CIFAR-10 是由 Hinton 的两个学生 Alex 和 Ilya 收集的一个用于普适物体识别的数据集。CIFAR-10 由 60 000 张 32×32 的 RGB 彩色图片构成,共 10 个分类^[15]。本次实验采用数据集中的 50 000 张图片进行训练,10 000 张图片进行测试。这个数据集最大的特点在于将识别迁移到了普适物体,而且应用于多分类(姊妹数据集 CIFAR-100 达到 100 类,ILSVRC 比赛则是 1 000 类)。同已经成熟的手写体识别相比,普适物体识别面临的挑战更艰巨。相比于 MNIST 数据集,CIFAR-10 数据集中的图片含有大量的特征以及噪声,需要识别的物体比例也比较多样。因此,CIFAR-10 相对于传统图像识别数据集是相当有挑战的。

本次实验训练时采用和 MNIST 数据集实验时一致的小批次梯度下降法作为网络的优化算法,设置批度为 50,所以共有 1 000 个批度,即卷积神经网络迭代一次,网络就训练了 1 000 次。实验环境同 MNIST 数据集实验。

CIFAR-10 试验中设置的卷积神经网络结构与参数与 MNIST 实验采用的完全一致。利用设置好的卷积神经网络分别采用平均值池化模型、最大值池化模型、中间值池化模型、本文提出的改进池化模型在 CIFAR-10 数据集上进行实验。其中,对于改进的池化模型,经过实验,池化因子取值 2。对每一种池化模型根据迭代次数进行实验,训练的迭代次数分别为 1、2、3 次。记录 4 种池化模型训练完后

的误识率如表2所示,训练过程中均方误差随训练批次的变化曲线如图3所示。

表2 4种池化模型对CNN性能影响的比(CIFAR-10) (%)

迭代次数	平均值池化	中间值池化	最大值池化	改进池化
1	77.85	79.45	79.20	75.70
2	75.65	73.00	73.25	70.15
3	71.35	70.55	70.15	68.00

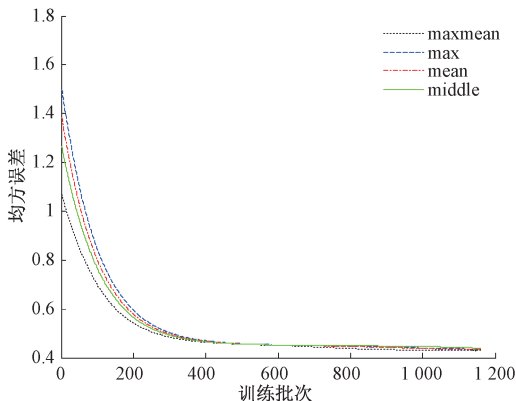


图3 4种池化模型在MNIST数据集上的均方误差曲线(MNIST)

首先可发现相比于MNIST数据集,在CIFAR-10数据集上误识率增大了很多。这是因为CIFAR-10数据集比MNIST数据集庞大了许多,且CIFAR-10数据集多了很多背景信息和噪音,在图片的特征上也复杂了许多。MNIST数据集是 28×28 的无背景下的手写数字黑白图片,而CIFAR-10是复杂背景下 32×32 大小的复杂物体的彩色图片,所以在这样一个卷积核较少,网络较浅的模型下,仅迭代1~3次,误识率会相对较高。

虽然误识率相对较高,但是还是可以从表2得出以下结论:随着迭代次数的增加,每一种池化模型都会不断地降低误识率,但改进池化模型的效果比其他模型更优越。卷积神经网络迭代一次时,平均值池化、中间值池化、最大值池化的误识率分别是77.85%、79.45%、79.2%,而改进池化模型的误识率为75.7%。迭代两次时,平均值池化、中间值池化、最大值池化的误识率分别是75.65%、73%、73.25%,而改进池化模型的误识率为70.15%。迭代3次时,平均值池化、中间值池化、最大值池化的误识率分别是71.35%、70.55%、70.15%,而改进池化模型的误识率为68%。与最大值池化模型相比,改进后的池化模型在迭代1、2、3次情况下,误识率分别下降了4.4192%、4.2321%、3.0649%。与平均值池化、中间值池化相比,效果会好一些。

图3展示了CIFAR-10数据集下迭代一次时,均方误

差随批次变化的曲线。图例表示与MNIST一致。由图3可知,随着训练批次的增加,每一种池化模型对于的均方误差都不断下降。此外,改进后的池化模型对应的均方误差在迭代一次后较其他池化模型小。

4 结 论

本文提出的改进的池化模型,首先计算池化域内最大值和平均值,再引入超参数池化因子,用这三者的乘积对池化域内的特征进行下采样,使得池化模型能针对不同的数据集提取更合适的特征,从而提高卷积神经网络的识别精度。为了研究改进池化模型对卷积神经网络性能的影响,并验证改进池化模型的普适性,本文选用MNIST数据集和CIFAR-10数据集分别对提出的改进池化模型进行训练和性能分析。实验结果表明,无论是在MNIST数据集上还是CIFAR-10数据集上,改进后的池化模型都可以降低卷积神经网络的误识率。因此可以得出结论,改进后的池化模型可以提升卷积神经网络的性能,并具有一定的普适性,可以运用到不同的数据集,也可以进一步地运用到其他基于卷积神经网络的深度学习网络。

参考文献

- [1] 刘万军,梁雪剑,曲海成.不同池化模型的卷积神经网络学习性能研究[J].中国图象图形学报,2016,21(9):1178-1190.
- [2] ALEX K, ILYA S, HINTON G E. ImageNet classification with deep convolutional neural networks[EB/OL].[2015-02-10]. <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>.
- [3] LEE H, GROSSE R, RANGANATH R, et al. Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations[C]. Proceedings of the 26th Annual International Conference on Machine Learning, ACM, 2009: 609-616.
- [4] GONG Y C, WANG L W, GUO R Q, et al. Multi-scale orderless pooling of deep convolutional activation features [C]. Proceedings of the 13th European Conference on Computer Vision, Springer, 2014: 392-407.
- [5] 史鹤欢,许悦雷,马时平,等.PCA预训练的卷积神经网络目标识别算法[J].西安电子科技大学学报,2016,43(3):161-166.
- [6] 张贵勇.改进的卷积神经网络在金融预测中的应用研究[D].郑州:郑州大学,2016.
- [7] 李媛媛.卷积神经网络优化及其在图像识别中的应用[D].沈阳:沈阳工业大学,2016.
- [8] 刘梦雅,毛剑琳.基于灰度关联分析法的卷积神经网络算法优化[J].电子科技,2018,31(6):84-88,95.

- [9] 张文达,许悦雷,倪嘉成,等.基于多尺度分块卷积神经网络的图像目标识别算法[J].计算机应用,2016,36(4):1033-1038.
- [10] 蒋帅.基于卷积神经网络的图像识别[D].长春:吉林大学,2017.
- [11] 余萍,赵继生.基于矩阵 2-范数池化的卷积神经网络图像识别算法[J].图学学报,2016,37(5):694-701.
- [12] 时增林.深度卷积神经网络中基于序的池化方法研究[D].郑州:郑州大学,2017.
- [13] 曹容川.图像分类中基于显著关系的改进特征编码与池化算法研究[D].长春:吉林大学,2017.
- [14] 孙峰,龚晓玲,张炳杰,等.一种基于共轭梯度法的广义

单隐层神经网络[J].郑州大学学报(工学版),2018,39(2):28-32.

- [15] 崔华,刘云飞,宋鑫鑫,等.基于 CNN 深度学习模型的交通图像拥堵识别[J].科技创新与应用,2018(4):19-20,22.

作者简介

刘梦雅,硕士研究生,主要研究方向为计算机视觉、模式识别、深度学习。

E-mail:389146631@qq.com

毛剑琳,博士、教授,主要研究方向为智能优化算法、无线传感器网络、MAC 层资源分配和优化。

E-mail:1318524654@qq.com