

基于循环卷积神经网络的目标检测与分类

艾玲梅, 叶雪娜

(陕西师范大学 计算机科学学院 陕西 西安 710119)

摘 要: 卷积神经网络模仿人类的视觉识别能力, 提取图像目标的显著抽象特征, 在图像目标检测与分类的应用上效果良好。在当前比较流行的批量随机梯度训练算法训练卷积神经网络的过程中, 当神经元处于饱和状态时, 会出现梯度下降缓慢和过度拟合问题, 易使神经网络模型训练陷入困难。结合卷积神经网络和循环神经网络的特点, 提出了构造浅层循环卷积神经网络, 且在训练循环卷积神经网络模型时, 分别采用进退法、黄金分割法自适应地改变批量随机梯度下降算法的规范化参数和学习率。实验结果表明, 改进算法能够较好地避免梯度下降缓慢和过度拟合问题, 在训练循环卷积神经网络模型时具有较好的目标检测分类效果和更快的收敛性。

关键词: 物体检测; 进退法; 黄金分割算法; 随机梯度算法; 神经网络

中图分类号: TP311

文献标识码: A

文章编号: 1673-629X(2018)02-0031-05

doi: 10.3969/j.issn.1673-629X.2018.02.008

Object Detection and Classification Based on Circular Convolutional Neural Network

AI Ling-mei, YE Xue-na

(School of Computer Science, Shaanxi Normal University, Xi'an 710119, China)

Abstract: Convolutional neural network simulates human vision recognition and extracts the abstract characteristics significantly of the image target, with better effects on the application of image target detection and classification. In the currently popular training of convolution neural network by batch stochastic gradient algorithm, when neurons in a saturated state, there will be a slow gradient descent and excessive fitting which lead to the difficulties in training of the neural network model. In this paper, we propose a simple circular convolutional neural networks combined with the characteristics of the convolutional and circular neural networks. In the training of circular convolutional neural network model, advance and retreat method and golden section are used to adaptively change the normalized parameter and the learning rate of batch stochastic gradient descent algorithm. Experiment shows that the proposed algorithm, with a better effect on detection and classification with faster convergence, can avoid the problem of slow gradient decent and excessive fitting to some extent.

Key words: object detection; advance and retreat method; golden section algorithm; stochastic gradient method; neural network

0 引 言

卷积神经网络识别二维图像目标具有很好的检测分类效果和稳定性。近年来, 卷积神经网络在一些图像目标检测分类中取得了一些突破性的进展^[1], ILS-VRC2014 比赛第一名 GoogLeNet 使用 22 层深度卷积神经网络结构, 错误率达到 6.67%^[2]。对于手写数字字符和 CIFAR-10 图像目标检测与分类的方法, arXiv201 的部分最大池化方法错误率分别下降到 0.32%、3.47%^[3]; 2014ImageNet 挑战赛中取得不错的分类检测效果^[4]; 深度监督网络训练的结果错误率为 0.53%、8.22%^[5]; Zhang 等结合贝叶斯分类器提出深度卷积神

经网络, 对图像目标进行分类^[6]; 大部分卷积神经网络模型采用随机梯度下降算法^[7]。传统的卷积神经网络模型由卷积层、抽样层和一个全连接层来实现图像目标的检测与分类, 而深度卷积神经网络模型参数规模较大。对于提高神经网络的训练速度和分类检测准确率, 目前主要有两个方向: 第一个是改进训练算法, Xavier Glorot 使用 Tanh 激活函数^[8], Hinton 使用弃权技术^[9], Bergstra 利用网格参数优化^[10]逐层使用标签信息预训练; 第二个是简化神经网络模型的结构, 使得简单的神经网络模型能够达到目前复杂的神经网络模型在图像目标检测与分类的准确率。Thomaxs 采纳异

收稿日期: 2017-01-21

修回日期: 2017-05-24

网络出版时间: 2017-10-19

基金项目: 国家自然科学基金(61672021); 陕西省自然科学基金基础研究计划资助项目(2015JM6296)

作者简介: 艾玲梅(1965-), 女, 副教授, 研究方向为信号处理和图像处理; 叶雪娜(1989-), 女, 硕士研究生, 研究方向为图像处理。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.TP.20171019.1624.032.html>

质的神经网络模型使用步长为 2 的循环卷积层取代传统的卷积神经网络模型中的卷积层和池化层,利用批量随机梯度下降算法进行训练,取得了当前复杂卷积神经网络模型分类检测的效果^[11]。

为了提高神经网络模型在图像数据集中检测与分类的训练速度和准确率,文中分别利用进退法、黄金分割算法自适应调整批量随机梯度下降算法的规范化参数和学习率来训练神经网络模型。深度复杂的神经网络模型训练过程比较缓慢甚至出现过拟合问题,结合循环神经网络和卷积神经网络的特点构造简单的循环神经网络模型,在训练少量数据集的情况下达到与较复杂的深度卷积神经网络同等的效果。

1 深度循环卷积神经网络模型

1.1 感知器单元

基于对生物学神经系统中视觉神经元工作机理的研究,提出了单个神经元为基础的“感知器”概念。这种模仿人类视觉机理的方法,在处理人类视觉和语音识别问题上表现出了明显优势。感知器模仿人类视觉神经元工作机理,当神经元受到刺激,相当于一个感知器从外界获取信号,然后通过加权求和的方式构成一个神经信号,再将这个输入信号通过非线性激活函数进行输出。

$$y = \text{sigmoid}(w_1x_1 + w_2x_2 + w_3x_3) \quad (1)$$

1.2 交叉熵代价函数

神经网络由若干神经元组成,在神经网络模型能够达到很好的分类效果时,交叉熵代价函数的输出值就会趋于 0。神经网络模型使用批量随机梯度下降算法,利用交叉熵代价函数可以避免神经元饱和、学习速率缓慢等问题^[12],因此采用交叉熵代价函数作为神经网络在分类时的代价函数。

$$c = -\frac{1}{n} \sum_x [y \ln a + (1 - y) \ln(1 - a)] \quad (2)$$

其中, n 为训练数据的总数,求和是在所有训练输入 x 上进行; y 为对应的目标输出。如果对于所有的训练输入 x ,神经元实际的输出值接近目标值,交叉熵接近于 0。

1.3 循环卷积神经网络

标准卷积神经网络通过卷积层提取输入层图像特征,然后输出到池化层作进一步处理。循环卷积神经网络使用循环卷积层来代替卷积层,从输入层提取图像目标,与标准的卷积神经网络中的卷积层功能一致。不同的是,循环卷积层提取输入层特征后并不输出给池化层,而是用变化的循环神经网络进行处理,采用对特征层数据添加空数据的方法,进行特征的抽象抽取。实验中限制的循环次数为 2,通过 2 次从特征层二次

抽取数据作为池化层的输入。

循环卷积神经网络模型的结构如图 1 所示。

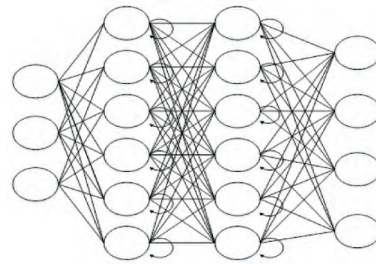


图 1 循环卷积神经网络模型结构

2 规范化参数

2.1 过度拟合

神经网络在学习过程中训练数据集较少时,训练速度较快。虽然神经网络学习到的模型适用于训练数据集,但是没有学习到数据的本质特征,神经网络模型泛化能力并不是很高。另一方面增加训练数据集学习到数据的本质特征,然而在增加训练数据集的过程中使得神经网络模型训练速度变得缓慢,且训练数据集在实际中很难获取。选择适量的数据集作为训练数据集来训练神经网络模型,随着神经网络模型训练过程中迭代次数的不断增加,神经网络模型在验证数据集和测试数据集中都会表现出过度拟合问题。

选取 MNIST 数据集,使用批量的随机梯度下降算法训练神经网络模型。如表 1 所示,分别用不同的隐藏层神经元数目和学习率,用相同的训练数据集来训练神经网络,随着迭代次数的增长,隐藏层神经元数目越多,批量随机梯度下降算法学习率越大,卷积神经网络模型出现过拟合时的迭代次数越小。当隐藏层神经元数目相同,批量随机梯度下降算法的学习率相同时,训练数据集越大,卷积神经网络模型训练过程中不容易出现过拟合问题,虽然增加了训练时间,但是训练的卷积神经网络模型在测试数据集上表现出良好的分类检测效果。

表 1 训练过度拟合实验数据表

数据集	神经元数	学习率	迭代数	正确率/%
1 000	30	0.5	60	83.38
1 000	30	0.6	50	82.63
1 000	100	0.5	100	86.64
1 000	100	0.6	120	82.75
2 000	30	0.5	100	87.13
2 000	30	0.6	85	86.57
2 000	100	0.5	70	88.24
2 000	100	0.6	60	85.39
5 000	30	0.5	180	90.23
5 000	30	0.6	145	92.32
5 000	100	0.5	150	92.10
5 000	100	0.6	130	92.49

图2展示了在神经元数目不同和批量随机梯度下降算法学习率相同的情况下,使用相同的训练数据集进行卷积神经网络模型的训练过程。卷积神经网络模型使用200个隐藏层神经元的检测分类效果不如使用100个隐藏神经元的效果,因此当训练数据集较少而神经元数目过多时,卷积神经网络模型的检测分类效果并不是很好。训练数据集较少时,简单的卷积神经网络模型表现出良好的检测分类效果,复杂的卷积神经网络模型在训练过程中出现过拟合问题,检测分类效果较差。

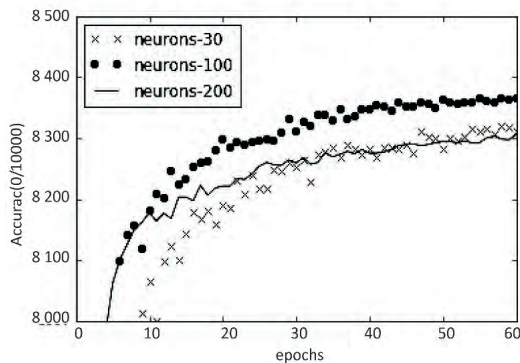


图2 不同神经元数目训练过程

2.2 参数规范化

神经网络模型训练特定的数据集,当迭代次数逐渐增加时会出现过度拟合问题。增加训练样本数量可以消除过拟合问题,然而训练样本在现实生活中很难获取,且会降低神经网络训练速度。文中采取固定的网络和固定的训练数据集,使用规范化权重衰减技术,在代价函数中增加一个额外的代价函数的规范化项。

$$c = -\frac{1}{n} \sum_{ij} [y_j \ln a_j^i + (1 - y_j) \ln(1 - a_j^i)] + \frac{\lambda}{2n} \sum_w w^2 \quad (3)$$

代价函数在交叉熵代价函数的基础上增加了规范化项,所用的规范化项为所有权重的平方和,其中 $\lambda > 0$ 可以称为规范化参数,而 n 就是训练集合的大小。增加了规范化项之后,权重的学习规则发生了微小的变化,这和通常的梯度下降学习规则相同,权重在原来学习的基础上重新调整 w ,这种调整通常称为权重衰减。

使用加入规范化项的交叉熵代价函数,随着迭代次数的增加,神经网络训练过程中在验证数据集上的分类准确率不断提高,降低了神经网络的过度拟合情况。对于固定的数据集和固定的网络结构,如何选择规范化参数可以使得神经网络在学习过程中能够得到更好的神经网络模型。

2.3 学习率

批量随机梯度下降算法在训练神经网络过程中,

学习率和代价函数值比较大时,神经网络模型训练速度较快;学习率和代价函数值较小时,神经元会出现饱和,使神经网络模型在测试和验证数据集中分类准确率提高得并不多,且学习速度缓慢,甚至很难找到最好的分类效果。传统的神经网络学习过程中,批量随机梯度下降算法使用固定的学习率训练神经网络,当学习率设置过大时,神经网络在训练过程中能够快速下降到全局最优解附近。然而由于学习率过大,神经网络学习过程在全局最优解附近徘徊。当学习率设置过小时,神经网络在学习过程中学习速度缓慢,并且容易陷入局部最优解。

对于固定的网络结构和固定的训练数据集,使用批量随机梯度下降算法在训练神经网络时,选取不同的学习率,神经网络在学习过程中表现出的分类效果和代价函数值区别比较大。

3 改进的批量梯度下降算法

3.1 进退法调整规范化参数

神经网络模型使用批量随机梯度下降算法的学习过程中,使用包含交叉熵规范化项的代价函数,实验证明加入规范化项可以降低神经网络学习过程中的过度拟合问题。对于固定的网络和固定的训练数据集,规范化项中的规范化参数的选择对解决神经网络学习过程中的过度拟合问题有很大帮助。选择一个好的规范化参数,神经网络可以快速学习,因此使用进退法^[13]来确定规范化参数,算法流程如下:

- (1) 给定规范化参数初始值 λ , a , b , 初始化步长 h , $C_1 = C(\lambda)$;
- (2) $\text{epoch} = \text{epochs} + 1$, 计算代价函数值;
- (3) $a = \lambda + h$, $b = \lambda - h$ 计算;
- (4) $c_2 = C(a)$, $c_3 = C(b)$;
- (5) 如果 $C_1 > C_2$, 则 $\lambda = a$, $h = 2h$;
- (6) 如果 $C_1 > C_3$, 则 $\lambda = b$, $h = -h$;
- (7) 当 $\text{epoch} < \text{epochs}$ 返回(2)。

3.2 黄金分割算法调整学习率

神经网络在训练学习的过程中,初始化权重之后,在训练数据集和验证数据集上分类错误率较高,代价函数值较大。为了使神经网络模型能够更快更好地学习参数模型,要不断改进学习率来适应随机梯度下降算法在训练数据集中找到全局最优解。黄金分割算法在给定初始值之后不断减小搜索区间,符合神经网络学习过程中学习率动态变化的需求,流程图见图3。

4 实验结果

4.1 MNIST 数据集

MNIST 数据集包含 70 000 张图像, 60 000 张训练

图像和 10 000 张测试图像。每个 MNIST 图像是一个单一的手写数字字符化图片,每个图像像素大小 28×28 ,像素值为 $0 \sim 255$,0 和 255 分别表示白色和黑色,中间像素值表示灰度级。文中选取 MNIST 数据集中的 1 000 张图像作为训练集,10 000 张图像作为验证集,10 000 张图像作为测试集。

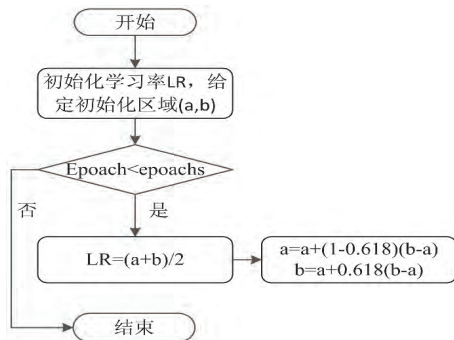


图 3 黄金分割算法流程

4.2 CIFAR-10 数据集

CIFAR-10 是为图像的目标检测而设计的计算机视觉数据库。CIFAR-10 数据库包含 60 000 张图片,由 10 类图像组成,每类图像 6 000 张。文中选取 CIFAR-10 中的 1 000 张图片作为训练集,10 000 张图片作为验证集,10 000 张图片作为测试集。

4.3 深度循环卷积神经网络分类结果

结合卷积神经网络和循环神经网络的各自优点,构造出三层深度神经网络。第一层使用卷积层,从原始图片中提取图片目标的特征,第二层采用最大池化的方法从卷积层进一步获取图像的显著特征,后面几层循环卷积层和池化层交替出现,而池化层分别采用最大池化、平均池化、最大池化和平均池化的方法。

如表 2 所示,深度循环神经网络在 MNIST 数据集上目标检测与分类的错误率为 0.43%,相比于传统的卷积神经网络错误率减少了 0.1%。改进的批量随机梯度下降算法训练的循环卷积神经网络深度只有 3 层,然而在 MNIST 手写数字字符数据集的图像目标检测与分类上比传统深度卷积神经网络性能更好。

表 2 MNIST 分类结果

方法	错误率/%
network in network ^[14]	0.47
stochastic pooling ^[5]	0.47
maxout networks ^[3]	0.45
DSN ^[15]	0.39
CNN ^[4]	0.53
RCNN-32 ^[16]	0.42
recurrent convolution networks(LSGD)	0.43

如表 3 所示,当训练数据集为 10 000 时,深度循环神经网络在训练 CIFAR-10 数据集时,分类结果显示错误率为 2.82%。循环深度神经网络模型错误率比

CNN 模型减少 2% 左右,表明深度循环卷积神经网络在 CIFAR-10 图像的目标检测与分类上表现出良好的分类效果。

表 3 CIFAR-10 分类结果

方法	错误率/%
network in network ^[14]	10.41
stochastic pooling ^[5]	15.13
maxout networks ^[3]	11.68
DSN ^[15]	8.82
RCNN-96(3iters) ^[16]	9.31
recurrent convolution networks(LSGD)	2.82

循环卷积神经网络通过循环卷积层提取图像目标的特征,空间采样层实现图像的位移、旋转不变形。如图 4 所示,选取 CIFAR-10 中的图片 10 张,每类各取 1 张分别经过循环卷积层和空间采样层处理之后得到输出效果。

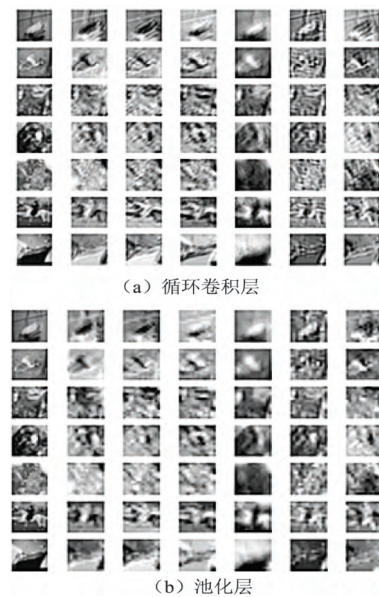


图 4 循环卷积池化得到的可视化特征层(CIFAR-10)

将循环卷积和空间亚采样学习到的特征输出,选取卷积核数目为 9 的 5×5 循环卷积,通过循环卷积得到不同朝向不同频率的边缘特征,循环卷积池化层输出对于位移、形变和旋转不变性的特征。

作为比较,用随机梯度下降算法和改进的批量随机梯度下降算法来训练深度循环卷积神经网络。图 5 显示随机梯度下降算法和改进的批量随机梯度下降算法在训练 CIFAR-10 数据集时,测试集的准确度与训练步数之间的关系。由图 5 可知,改进的批量随机梯度下降算法在训练深度卷积神经网络模型时,比随机梯度下降算法的性能更好,且在训练深度卷积神经网络过程中收敛速度快于传统的随机梯度下降算法。

图 6 显示了传统随机梯度下降算法和改进的批量随机梯度下降算法在训练 MNIST 数据集构建的循环神经网络模型上的不同。由图可知,随机梯度下降算

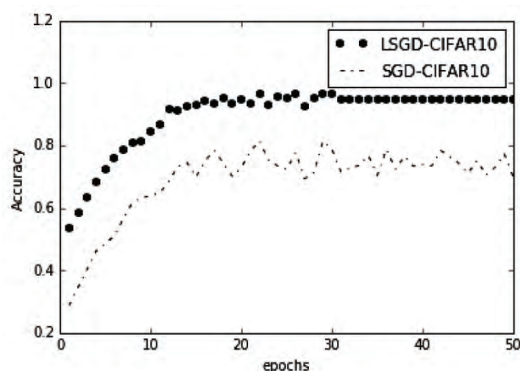


图5 CIFAR-10 分类结果(CIFAR-10)

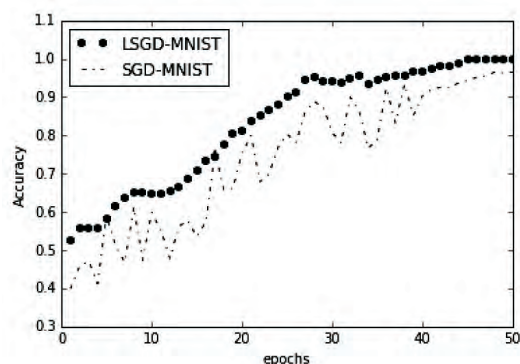


图6 MNIST 训练结果

法和改进的随机梯度下降算法在训练网络时,所表现出的性能基本相同,而改进的批量随机梯度下降算法表现出良好的目标检测分类效果和更快的收敛性。

实验结果表明,通过改进的批量随机梯度下降算法训练的深度循环卷积神经网络,在应用于 CIFAR-10 数据集和 MNIST 数据集的错误率比目前较好的 CNN 方法分别减少了 2% 和 0.1%,表现出了较好的目标检测分类效果,且在训练循环卷积神经网络时收敛速度更快。

5 结束语

结合卷积神经网络和循环神经网络在图像目标检测与分类上的优点,构造深度为 3 层的深度循环卷积神经网络。针对随机梯度下降算法在训练过程中出现的过度拟合和容易陷入局部最优解等缺点,采用改进的批量随机梯度下降算法。对需要靠经验设定梯度下降的学习率(learning rate)和规范化参数问题,分别用进退法和黄金分割算法自适应地改变学习率和规范化参数。实验结果表明,改进算法能较好地避免梯度下降缓慢和过拟合问题,在训练循环卷积神经网络模型时具有较好的目标检测分类效果和更快的收敛性。

参考文献:

[1] LIANG M ,HU X.Recurrent convolutional neural network for object recognition[C]//IEEE conference on computer vision

and pattern recognition. [s.l.]: IEEE, 2015: 3367-3375.

[2] SZEGEDY C ,LIU W ,JIA Y ,et al.Going deeper with convolutions[C]//IEEE conference on computer vision and pattern recognition. [s.l.]: IEEE, 2014: 1-9.

[3] RAYMOND F E ,RAYMOND B W.Fractional max-pooling [J].Theoretical Economics Letters, 2015, 5(2): 225-237.

[4] KRIZHEVSKY A ,SUTSKEVER I ,HINTON G E.ImageNet classification with deep convolutional neural networks[C]//International conference on neural information processing systems. [s.l.]: [s.n.], 2012: 1097-1105.

[5] LEE C Y ,XIE S ,GALLAGHER P ,et al.Deeply-supervised nets[C]//Artificial intelligence and statistics conference. [s.l.]: [s.n.], 2014: 562-570.

[6] ZHANG Y ,SOHN K ,VILLEGAS R ,et al.Improving object detection with deep convolutional networks via Bayesian optimization and structured prediction[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. [s.l.]: IEEE, 2015: 249-258.

[7] LECUN Y ,JACKEL L D ,BOTTOU L ,et al.Comparison of learning algorithms for handwritten digit recognition [C]//International conference on artificial neural networks. [s.l.]: [s.n.], 1995: 53-60.

[8] GLOROT X ,BENGIO Y. Understanding the difficulty of training deep feedforward neural networks [J]. Journal of Machine Learning Research, 2010, 9(2): 249-256.

[9] HINTON G E ,SRIVASTAVA N ,KRIZHEVSKY A ,et al. Improving neural networks by preventing co-adaptation of feature detectors [J]. Computer Science, 2012, 3(4): 212-223.

[10] BERGSTRA J ,BENGIO Y.Random search for hyper-parameter optimization [J]. Journal of Machine Learning Research, 2012, 13(2): 281-305.

[11] SPRINGENBERG J T ,DOSOVITSKIY A ,BROX T ,et al. Striving for simplicity: the all convolutional net [J]. Pattern Recognition Letters, 2014, 3(2): 251-263.

[12] 黄凯奇,任伟强,谭铁牛.图像物体分类与检测算法综述[J].计算机学报, 2014, 37(6): 1225-1240.

[13] 史清录,孔祥莹,康健.进退法在多维非线性有约束优化问题中的应用[J].太原重型机械学院学报, 2001, 22(3): 200-203.

[14] SPRINGENBERG J T ,RIEDMILLER M.Improving deep neural networks with probabilistic maxout units [J]. Computer Science, 2013, 12(3): 495-505.

[15] CHENG D.Controllability of switched bilinear systems [J]. IEEE Transactions on Automatic Control, 2005, 50(4): 511-515.

[16] JARRETT K ,KAVUKCUOGLU K ,RANZATO M ,et al. What is the best multi-stage architecture for object recognition[C]//Proceedings of international conference on computer vision. [s.l.]: [s.n.], 2009: 2146-2153.