

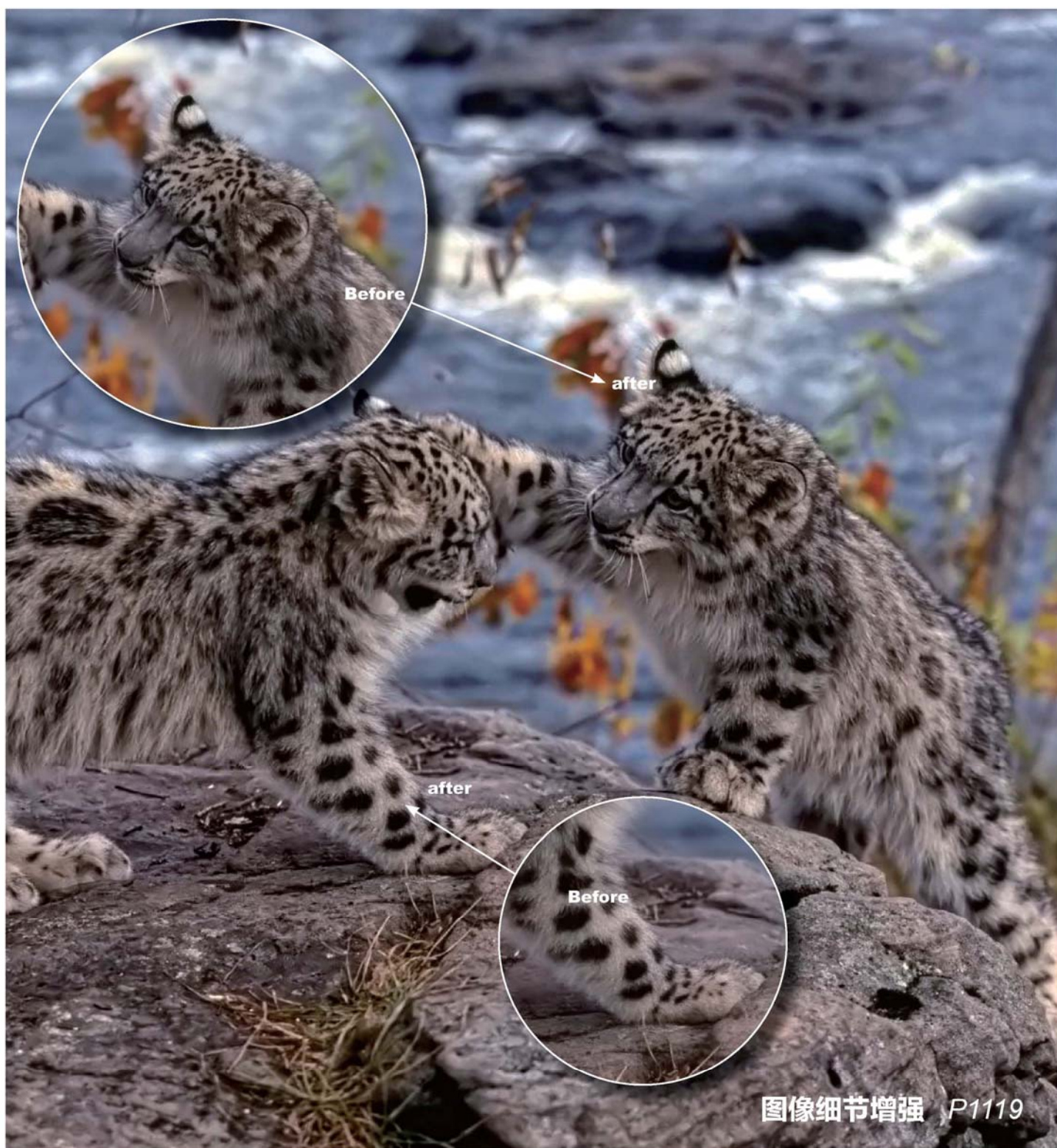


主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2016
09
VOL.21

ISSN1006-8961
CN11-3758/TB



图像细节增强 P1119



第21卷第9期（总第245期）
2016年9月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院
主办单位 中国科学院遥感与数字地球研究所
中国图象图形学会
北京应用物理与计算数学研究所

主 编 顾行发
编辑出版《中国图象图形学报》编辑出版委员会
邮政信箱 北京9718信箱
邮 编 100101
电子信箱 jig@radi.ac.cn
电 话 010-64807995
网 址 www.cjig.cn

广告经营许可证 京朝工商广字第0361号
总 发 行 北京报刊发行局
订 购 全国各地邮局
海外发行 中国国际图书贸易集团有限公司
(邮政信箱: 北京399信箱 邮编: 100048)
印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

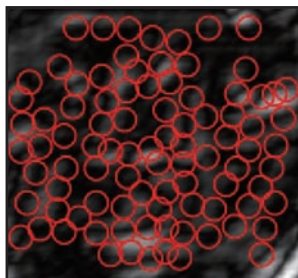
Title inscription: Song Jian Monthly, Started in 1996

Superintended by Chinese Academy of Sciences
Sponsored by Institute of Remote Sensing and Digital Earth, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

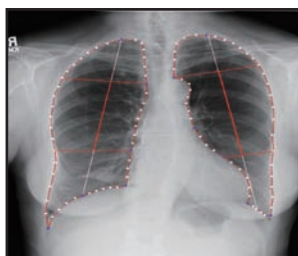
Editor-in-Chief Gu Xingfa
Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics
P.O.Box 9718, Beijing, P.R.China
Zip code 100101
E-mail jig@radi.ac.cn
Telephone 010-64807995
Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB
ISSN 1006-8961
CODEN ZTTXFZ
国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



基于频率簇模型的人脸识别
(第1166页)



胸部解剖结构回归模型的虚拟双能量X线减影方法(第1247页)



参考1维光谱差异的区域生长种子点选取方法(第1256页)

图像处理和编码

融合梯度信息的改进引导滤波

谢伟, 周玉钦, 游敏 1119

用于屏幕图像编码的索引图快速预测算法

陈规胜, 宋传鸣, 王相海, 刘丹 1127

图像分析和识别

哈希编码结合空间金字塔的图像分类

彭天强, 栗芳 1138

自上而下注意力分割的细粒度图像分类

冯语姝, 王子磊 1147

结合类内和类间距离的可能聚类分割算法

刘璐, 吴成茂 1155

基于频率簇模型的人脸识别

袁姮, 王志宏, 姜文涛 1166

图像理解和计算机视觉

不同池化模型的卷积神经网络学习性能研究

刘万军, 梁雪剑, 曲海成 1178

惰性随机游走视觉显著性检测算法

李波, 卢春园, 金连宝, 冷成财 1191

自适应邻域相关性的背景建模

万剑, 洪明坚, 赵晨丘 1202

实时鲁棒的特征点匹配算法

陈天华, 王福龙 1213

暗通道先验图像去雾的大气光校验和光晕消除

赵锦威, 沈逸云, 刘春晓, 欧阳毅 1221

计算机图形学

定点容量限制质心Power图生成

郑利平, 郜文灿, 李尚林, 曹力 1229

医学图像处理

多相期增强CT中肝内血管的自动分割

袁戎, 石姝玥, 谢庆国 1238

胸部解剖结构回归模型的虚拟双能量X线减影方法

陈胜, 张茗屋 1247

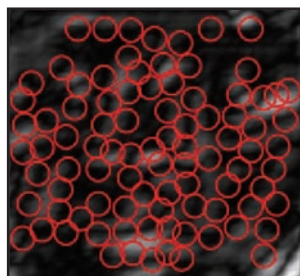
遥感图像处理

参考1维光谱差异的区域生长种子点选取方法

李修霞, 荆林海, 李慧, 唐韵玮, 戈文艳 1256

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



Face recognition method based on frequency cluster (P1166)



Virtual dual-energy subtraction method for X-ray radiographs by using regression model based on chest anatomical structure(P1247)



Seed extraction method for seeded region growing based on one-dimensional spectral differences(P1256)

Image Processing and Coding

- Improved guided image filtering integrated with gradient information
Xie Wei, Zhou Yuqin, You Min 1119
- Fast prediction algorithm of index maps for screen image coding
Chen Guisheng, Song Chuanming, Wang Xianghai, Liu Dan 1127

Image Analysis and Recognition

- Image classification algorithm based on hash codes and space pyramid
Peng Tianqiang, Li Fang 1138
- Fine-grained image categorization with segmentation based on top-down attention map
Feng Yushan, Wang Zilei 1147
- Possibilistic clustering segmentation algorithm based on intra-class and inter-class distance
Liu Lu, Wu Chengmao 1155
- Face recognition method based on frequency cluster
Yuan Heng, Wang Zhihong, Jiang Wentao 1166

Image Understanding and Computer Vision

- Learning performance of convolutional neural networks with different pooling models
Liu Wanjun, Liang Xuejian, Qu Haicheng 1178
- Saliency detection based on lazy random walk
Li Bo, Lu Chunyuan, Jin Lianbao, Leng Chengcai 1191
- Background modeling based on adaptive neighborhood correlation
Wan Jian, Hong Mingjian, Zhao Chenqiu 1202
- Real-time robust feature-point matching algorithm
Chen Tianhua, Wang Fulong 1213
- Dark channel prior-based image dehazing with atmospheric light validation and halo elimination
Zhao Jinwei, Shen Yiyun, Liu Chunxiao, Ouyang Yi 1221

Computer Graphics

- Generation method for a centroidal capacity constrained Power diagram with fixed sites
Zheng Liping, Gao Wencan, Li Shanglin, Cao Li 1229

Medical Image Processing

- Automatic hepatic vessel segmentation algorithm in multi-phase contrast-enhanced CT
Yuan Rong, Shi Shuyue, Xie Qingguo 1238
- Virtual dual-energy subtraction method for X-ray radiographs by using regression model based on chest anatomical structure
Chen Sheng, Zhang Mingwu 1247

Remote Sensing Image Processing

- Seed extraction method for seeded region growing based on one-dimensional spectral differences
Li Xiuxia, Jing Linhai, Li Hui, Tang Yunwei, Ge Wenyan 1256

中图分类号: TP391 文献标识码: A 文章编号: 1006-8961(2016)09-1178-13

论文引用格式: Liu W J, Liang X J, Qu H C. Learning performance of convolutional neural networks with different pooling models[J]. Journal of Image and Graphics, 2016, 21(9): 1178-1190. [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.] [DOI:10.11834/jig.20160907]

不同池化模型的卷积神经网络学习性能研究

刘万军, 梁雪剑, 曲海成

辽宁工程技术大学软件学院, 葫芦岛 125105

摘要: 目的 基于卷积神经网络的深度学习算法在图像处理领域正引起广泛关注。为了进一步提高卷积神经网络特征提取的准确度, 加快参数收敛速度, 优化网络学习性能, 通过对比不同的池化模型对学习性能的影响提出一种动态自适应的改进池化算法。方法 构建卷积神经网络模型, 使用不同的池化模型对网络进行训练, 并检验在不同迭代次数下的学习结果。在现有算法准确率不高和收敛速度较慢的情况下, 通过使用不同的池化模型对网络进行训练, 从而构建一种新的动态自适应池化模型, 并研究在不同迭代次数下其对识别准确率和收敛速度的影响。结果 通过对比实验发现, 使用动态自适应池化算法的卷积神经网络学习性能最优, 在手写数字集上的收敛速度最高可以提升 18.55%, 而模型对图像的误识率最多可以降低 20%。结论 动态自适应池化算法不但使卷积神经网络对特征的提取更加精确, 而且很大程度地提高了收敛速度和模型准确率, 从而达到优化网络学习性能的目的。这种模型可以进一步拓展到其他与卷积神经网络相关的深度学习算法。

关键词: 深度学习; 卷积神经网络; 图像识别; 特征提取; 算法收敛; 动态自适应池化

Learning performance of convolutional neural networks with different pooling models

Liu Wanjun, Liang Xuejian, Qu Haicheng

College of Software, Liaoning Technical University, Huludao 125105, China

Abstract: **Objective** Deep learning algorithms based on convolutional neural networks are attracting attention in the field of image processing. To improve the accuracy of the feature extraction process and the convergence rate of parameters, as well as optimize the learning performance of the network, an improved dynamic adaptive pooling algorithm is proposed, which compares the effect of different pooling models on learning performance. **Method** A convolutional neural network model, which is trained with different pooling models, is constructed. The results of the trained model are verified in different iterations. To compensate for low accuracy and slow convergence speed, a dynamic adaptive pooling model is proposed, which trains the network with different pooling models. The effect of the model on the accuracy and convergence rate in different iterations are then studied. **Result** Contrast experiment shows that the dynamic pooling model has optimal learning performance. The maximum improvement of the convergence rate on handwritten database is 18.55% and the maximum decrement of the accuracy rate is 20%. **Conclusion** A dynamic adaptive pooling algorithm can improve the accuracy of feature extraction, convergence rate, and accuracy of the convolutional neural network, thereby optimizing network learning performance.

收稿日期: 2015-11-24; 修回日期: 2016-04-20

基金项目: 国家自然科学基金项目(61172144); 辽宁省教育厅科学技术研究一般项目(L2015216)

第一作者简介: 刘万军(1959—), 男, 教授, 博士生导师, 1991年于阜新矿业学院获电力传动及其自动化专业工学硕士学位, 主要研究方向为数字图像处理、运动目标检测与跟踪。E-mail: liuwanjun@lntu.edu.cn

Supported by: National Natural Science Foundation of China(61172144)

The dynamic adaptive pooling model can be further extended to other deep learning algorithms related to convolutional neural networks.

Key words: deep learning; convolutional neural network; image recognition; feature extraction; algorithm convergence; dynamic adaptive pooling

0 引言

近年来,深度学习使机器学习在人工智能领域取得了重要突破。深度学习的本质是为了使机器学习的过程更加接近于人工智能^[1]。深度学习技术由于对图像特征提取的独特优势而被计算机视觉、模型匹配与模式识别等领域广泛研究。深度学习强化了机器学习的学习能力,简化了学习的复杂度,赋予机器学习强大的数据表示与特征表达能力。作为人工智能的重要标志,深度机器学习成为人工智能领域的研究热点,并促进了计算机视觉、语音识别和自然语言处理等领域的发展。机器学习从产生开始经历了由浅层学习到深度学习的两次浪潮^[2]。Hinton 等人^[3]提出多隐层结构的神经网络特征提取能力更优异的观点,使得深度学习被大数据和机器学习领域的学术界广泛讨论与研究。大数据除了数据量非常大和数据种类繁多之外,数据的结构也比较复杂,矩阵数据的维数往往也非常高,这样对于很多算法来说处理起来都会比较困难,可以实现的算法结果通常也并不能令人满意。而深度学习算法试图找到数据的内部结构,从而发现数据变量之间真正的关系形式^[4]。深度学习可以通过降维等手段把数据简化,并且使用梯度下降算法对海量数据进行优化,提高模型的准确率。深度学习对海量数据的处理能力^[5]和对多层次间非线性关系的学习能力是其他的算法无法比拟的,深度学习只使用简洁的参数就可以学习到复杂的函数^[6]。

卷积神经网络(CNN)对深度学习在图像处理等方面的研究与发展具有十分重要的作用。使用卷积神经网络的深度学习算法使得很多图像处理问题取得了成功,而大规模深度学习网络在 GPU 上的运行又进一步促进了其发展。利用卷积神经网络对图像进行处理,不但可以模拟人视觉系统中逐层提取特征的处理机制去提取图像较深层次的特征,而且可以使用端到端的特征提取方式,使得提取到的特征具有比较强的泛化能力,避免了由于图像变换产

生的特征提取不精确的问题。除此之外,卷积神经网络既可以学习到图像的全局特征,又可以通过多层的隐层结构学习到图像的局部特征,而这些局部特征之间是具有一定的关联性的。由于存在的多种优越性,卷积神经网络在图像分类与识别、自然语言识别、语音的识别^[7-8]与合成、信号处理以及声音信号场景分析和人脸识别^[9-10]等方面有着广泛的应用。卷积神经网络对数据的学习需先经过卷积和子采样等步骤来构建模型,然后通过检测模型的准确率以确定其是否符合要求,而池化算法的选择决定了子采样提取到特征的有效性,从而对模型的准确率产生影响。因此,卷积神经网络往往存在训练时间长、收敛速度慢或者模型准确率不够高等问题。为了满足实际应用对卷积神经网络的要求, Lee 等人^[11]使用分层生成模型使卷积神经网络处理高维图像具有很高的准确率; Matsugu 等人^[12]提出了卷积神经网络在人脸检测时具有鲁棒性的观点;而 Huang 等人^[13]通过构建局部卷积 RBM 模型来达到对网络学习性能的提升;余永维等人^[14]通过对特征提取层和子采样层的研究保证了其实验模型的有效性。卷积神经网络对缩放、旋转和平移等各种扭曲不变性图像具有良好的鲁棒性^[15],因此其对 2 维图像进行处理也更加容易。特征提取的有效程度往往决定卷积神经网络的学习性能,所以对特征提取过程(卷积和池化)进行优化与改进来提高算法的收敛速度和准确率等特性^[16]具有可行性。Alessandro 等人^[17]通过使用最大池化模型的卷积神经网络实现了图像的快速扫描,而 Jawad 等人^[18]同样使用最大池化模型实现了手势识别并取得了良好的效果。Gong 等人^[19]使用多尺度无序池化解决了卷积神经网络全局激活缺乏几何不变性的问题。为了加快卷积神经网络的收敛速度和提高图像识别能力。本文通过在卷积神经网络子采样过程中使用不同的池化模型,并对经典池化模型进行改进,对比其在特征提取阶段的优化程度,分析对网络学习性能的影响,进而提出一种基于最大池化算法的动态自适应改进池化模型,以提高模型的准确率及网络的学习性能。

1 卷积神经网络

卷积神经网络作为深度学习算法在人工神经网络中的应用,得到了越来越多的重视并被不断地改进以适用于各种实际情况。卷积神经网络的输入是没有经过预处理的原始图像,并且在输入层与输出层之间没有精确的数学计算表达式,无论是特征提取还是目标模型构建的过程都是在内部根据训练进行的,这样就可以避免人工提取特征的主观性和局限性,从而使生成的模型也更加接近于理想值。卷积神经网络中引入感受野和权值共享的概念,可以降低网络结构的复杂性,同时也使其更加适合并行计算。

1.1 卷积神经网络的结构

卷积神经网络由输入层、隐层和输出层构成。在隐层中包含多个由卷积层(convolutional layer)和子采样层(sub-sampling layer)构成的重复结构。卷积神经网络结构如图1所示。

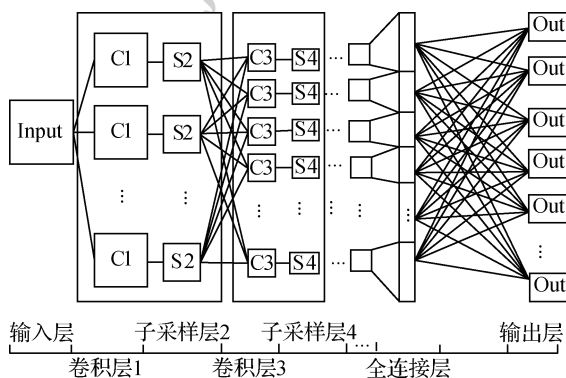


图1 卷积神经网络结构图

Fig. 1 Convolutional neural network structure

图1中卷积神经网络的输入层是原始图像 Input,并不需要人工提取特征。卷积层 C1 中包含的多个卷积核分别对 Input 进行卷积,生成各自对应的卷积特征图(feature map)。子采样层 S2 对 C1 中的特征图进行局部特征提取,生成对应的子采样特征图。C3 和 S4 重复 C1 和 S2 的结构,这种重复结构大量存在于卷积神经网络的隐层之中。使用“卷积+子采样”的结构进行特征提取,可以使网络对一些符合扭曲不变性的形变图像有较好的容忍能力,同时减小了特征图像的分辨率,增加生成特征图的数量,获取更多的特征信息。把最后得到的子采

样特征图全部展开构成1维矩阵(全连接层),输出层与1维矩阵采用全连接的方式得到最后的输出结果。

1.2 卷积过程

卷积层(C层)卷积的过程如图2所示。设输入层为 $m \times m$ 的矩阵 M ,卷积核为 $n \times n$ 的矩阵 C ,其值为可训练参数(权重),偏置为 b_1 ,得到的卷积特征图为矩阵 F ,卷积核的移动步长为1,使用 sigmoid 函数作为激活函数 $S(\cdot)$,因此得到的卷积特征图 F 的大小为 $(m-n+1) \times (m-n+1)$ 。卷积计算为

$$F_{ij} = S\left(\sum_{i=1}^n \sum_{j=1}^n (M_{ij}C_{ij}) + b_1\right) \quad (1)$$

式中, F_{ij} 为矩阵 F 中元素, M_{ij} 表示卷积时输入层中与卷积核相对应的元素,并非 M 中实际的第 i 行 j 列的值。输入层可以是初始图像,也可以是子采样层生成的特征图。卷积核内部的参数是可训练的,卷积核分别与输入层同等大小的矩阵进行点乘,求和之后再加偏置,激活后得到的结果就是对应卷积特征图的值。图2为偏置 $b_1 = 0$ 时激活前的卷积过程,得到的特征图为 3×3 的矩阵。

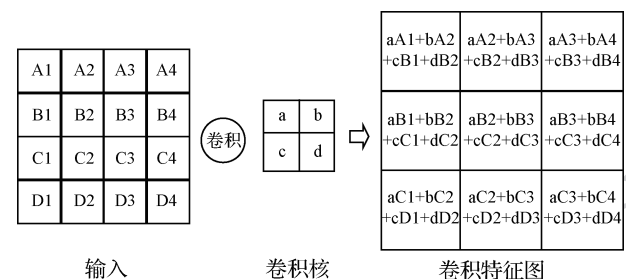


图2 卷积过程示意图

Fig. 2 Convolutional process schematic

复杂图形往往由基本的结构组成,局部特征往往可以表征全局特征,所以全局特征提取并不是十分必要,而且根据人的视觉结构,更高层的神经元可以把低层获取的局部特征综合起来,这样就可以得到全局的特征信息。因此,在卷积神经网络中提出以感受野作为卷积的基本对象进行局部特征提取,相比较全局特征提取就达到了参数减少的目的。在进行卷积运算的过程中,使用同一个卷积核对图像的不同感受野进行计算,这个过程中所有的感受野是共享卷积核参数的,这种权值共享的方法进一步减少了训练参数,提升了网络的学习性能。

1.3 子采样过程

子采样的目的是进行二次特征提取,其中最重

要的过程就是池化(pooling)。在计算图像某个区域特征值的时候,往往需要对这个区域的特征进行分析与统计,用新的特征来代表这个区域的总体特征。这个区域叫池化域,这个过程叫做池化。池化后得到的高层特征图不但可以降低原特征图的维度和分辨率,还可以避免过拟合等问题。池化的方法有平均池化法和最大池化法等。池化过程如图3所示。

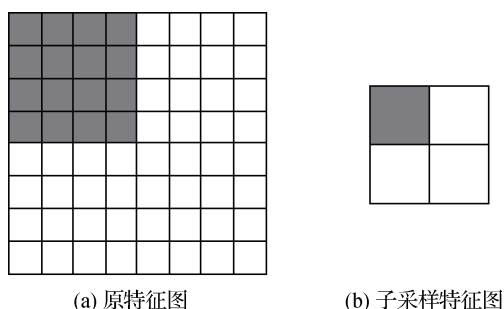


图3 池化过程

Fig. 3 Pooling process ((a) original feature map; (b) sub-sampling feature map)

图3(a)原特征图为 8×8 的矩阵,其中 4×4 的灰色区域代表池化域。对原特征图池化域进行池化时的移动步长为4,得到图3(b)子采样特征图为 2×2 的矩阵。子采样特征图的灰色区域为原特征图灰色区域对应的池化结果。

池化的实质就是池化函数 $P_{\circ}()$ 作用于池化域的过程。设原特征图为矩阵 F ,特征图的池化域为矩阵 P ,将池化后的结果加偏置 b_2 进行计算,遍历原特征图的池化域后,得到子采样特征图 S ,即

$$S = P_{\circ}(F) + b_2 \quad (2)$$

1.4 卷积神经网络模型的构建

为了研究不同池化模型对卷积神经网络的优化性能,首先需要建立卷积神经网络。卷积神经网络的结构基本包含建立模型、训练模型和测试模型3个部分,在训练模型过程中包含训练的前向过程、计算误差反向传播、权值修改与模型更新3个过程。训练过程实现了特征提取和有监督调整误差权值,完成了对输入数据的处理与目标模型的建立。卷积神经网络模型的结构如图4所示。

卷积神经网络从建立网络结构到最后确定均方

误差曲线和模型准确率的流程图如图5所示。

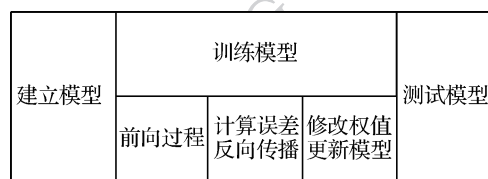


图4 卷积神经网络模型结构

Fig. 4 Convolutional neural network model structure

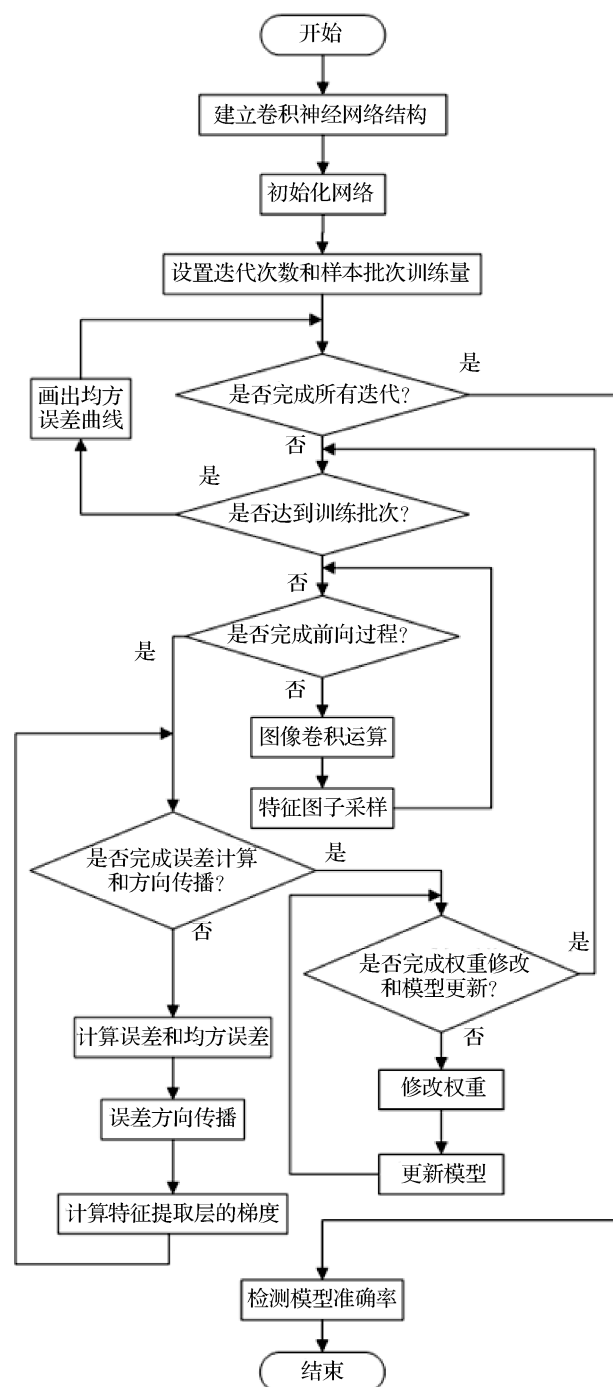


图5 卷积神经网络流程图

Fig. 5 Convolutional neural network flow chart

2 池化模型理论

2.1 经典池化模型

经典池化模型中最常见的两种模型为平均池化模型和最大池化模型。平均池化模型指在池化过程中,对池化域内所有值求和并取其平均数作为子采样特征图中的值;而最大池化模型则是取池化域中的最大值作为子采样的特征值。

设输入特征图矩阵 F , 子采样池化域为 $c \times c$ 的矩阵 P , 偏置为 b_2 , 得到的子采样特征图为 S , 设池化过程移动步长为 c 。平均池化和最大池化的算法表达式分别为

$$S_{ij} = \frac{1}{c^2} \left(\sum_{i=1}^c \sum_{j=1}^c F_{ij} \right) + b_2 \quad (3)$$

$$S_{ij} = \max_{i=1, j=1}^c (F_{ij}) + b_2 \quad (4)$$

式中, $\max_{i=1, j=1}^c (F_{ij})$ 表示从输入特征图 F 的大小为 $c \times c$ 的池化域中取出的最大元素。

2.2 改进的中间模型

经典的池化模型均不能很好的对池化域的特征进行提取。若取两种不同的池化域,如图6所示,图6中阴影部分表示值不相同的像素,空白部分表示像素值为0。整个池化域图6(a)的特征集中在最大值 v 处,如果对特征图图6(a)使用最平均池化方法,则会减弱图6(a)的特征;同理,图6(b)中在其他两个值 v_1 、 v_2 与最大值 v_3 的关系未知的前提下,用最大值表示这个池化域的特征值也会对特征有一定的弱化。所以使用平均池化或者最大池化提取特征会对全局特征的表示及模型的精确度有一定的损伤。

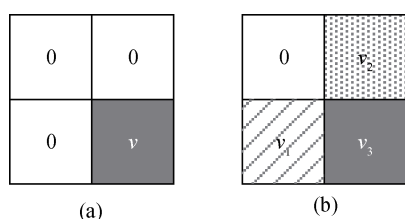


图6 不同池化域

Fig. 6 Different pooling area

对于以上问题,可以在经典池化模型的基础上进行改进,即通过选取两种中间模型来减少对卷积神经网络学习性能的影响。

1)最大二均值池化法。最大二均值池化是对最大池化的一种改良算法。因为最大池化法可能会忽略其他一些影响因素较大的值,所以在池化时选取池化域中值依次最大的两个元素并求其值的和,取和的平均数作为子采样的特征值。最大二均值池化算法表达式为

$$S_{ij} = \frac{1}{2} \sum_{i=1, j=1}^c M2(F_{ij}) + b_2 \quad (5)$$

式中, $\sum_{i=1, j=1}^c M2(F_{ij})$ 表示从输入特征图 F 的大小为 $c \times c$ 的池化域中取出值依次最大的两个元素并且求和。

2)中间值池化法。中间值池化法是对平均池化法和最大池化法的一种折中算法。这种算法可以兼顾平均池化法与最大池化法获取到的特征值,从而使模型产生的误差较小,稳定性较高,因而普遍适用于一般性图像。中间值池化算法表达式为

$$S_{ij} = T/2 + b_2 \quad (6)$$

$$T = \frac{1}{c^2} \left(\sum_{i=1}^c \sum_{j=1}^c F_{ij} \right) + \max_{i=1, j=1}^c (F_{ij}) \quad (7)$$

2.3 动态自适应池化模型

改进池化模型的目的是为了对特征提取过程进行优化。在整个卷积神经网络学习过程中会产生很多不同的特征图和各式各样的池化域,无论是经典的池化模型还是改进的中间模型,在面对这些特征图和池化域时都很难取得一个较为满意的效果。为了进一步改进池化模型,提出一种基于最大池化算法的动态自适应池化模型。这种模型可以根据不同的特征图,动态调整其池化过程,并且根据各个池化域的内容,自适应地调整池化权值。若池化域只有一个值,则这个值既是最大值,也是其特征的表示。如果这个池化域的特征值都相同,则其最大值也可以表示为池化域的特征值。因此,在最大池化算法的基础上,根据插值原理,构建数学模型进行函数模拟。设 μ 为池化因子,则动态自适应池化算法表达式为

$$S_{ij} = \mu \max_{i=1, j=1}^c (F_{ij}) + b_2 \quad (8)$$

此式为动态自适应算法的基本表达式。其本质是使用池化因子 μ 对最大池化算法进行优化。优化后的特征能更准确地表达特征。其余各项参数遵循最大池化模型的参数设置。

$$\mu = \rho \frac{a(v_{\max} - a)}{v_{\max}^2} + \theta \quad (9)$$

式中, a 为池化域元素除最大值外的平均值, $v_{\max}$ 为池化域元素中的最大值, θ 为校正误差项, ρ 为特征系数, 计算表达式为

$$\rho = \frac{c}{1 + (n_{\text{epo}} - 1)c^{n_{\text{epo}} + 1}} \quad (10)$$

式中, n_{epo} 为训练时的迭代次数。

式(8)~(10)中, 特征系数 ρ 取决于池化域的边长 c 和迭代次数 n_{epo} ; 而特征系数与池化域中各项的值决定了池化因子 μ 的取值。在池化域大小确定的情况下, 保持迭代周期不变, 池化因子会根据池化域的不同自适应取值; 而面对同一池化域时, 池化因子则会根据迭代次数的不同, 动态地调整以达到最优。因池化因子 $\mu \in (0, 1)$, 这样既可以兼顾最大池化和平均池化, 在处理最大值特征明显的池化域时不会丢失精度, 又可以在处理其余池化域时弱化最大池化的影响, 从而使卷积神经网络在不同的迭代次数下处理不同的池化域时都可以提取到更为精确的特征。

2.4 动态自适应池化的实现

因池化过程与卷积有相似之处, 所以可以利用卷积的原理实现池化, 其实现过程如图7所示。图7是对 8×8 的特征矩阵进行池化的过程, 池化域为 2×2 的矩阵, 子采样得到的特征图为 4×4 的矩阵。因为卷积的移动步长为1, 池化的移动步长为2, 卷积会对图7(a)中的输入特征图从第2行和第2列开始的偶数行与列重复卷积计算, 而池化域中没有重叠部分, 所以图7(b)对卷积结果中重复计算的行与列进行删除操作, 最后得到的结果就是子采样特征图。

动态自适应池化是在最大池化的基础上进行改进与优化的。最大池化模型的输入部分为图7(a)中输入特征图这样的2维矩阵, 卷积核也与图7(a)中相同, 使用4个不同的卷积核, 如权值分别为 $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $C = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ 以及 $D = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ 的矩阵对输入的特征图进行卷积, 分别得到池化域中4个不同的值对应的卷积结果。对卷积结果进行连接, 连接之后会得到一个 $4 \times 7 \times 7$ 的3维矩阵 T_1 。用求最大值函数对 T_1 进行运算, 得到一个 $1 \times 7 \times 7$ 的3维矩阵 T_2 , 对这个3维矩阵进行转置, 得到了 $7 \times 7 \times 1$ 的3维矩阵, 也就是 7×7 的3维矩阵, 这个矩阵就是最大值的卷积结果。再进行图7

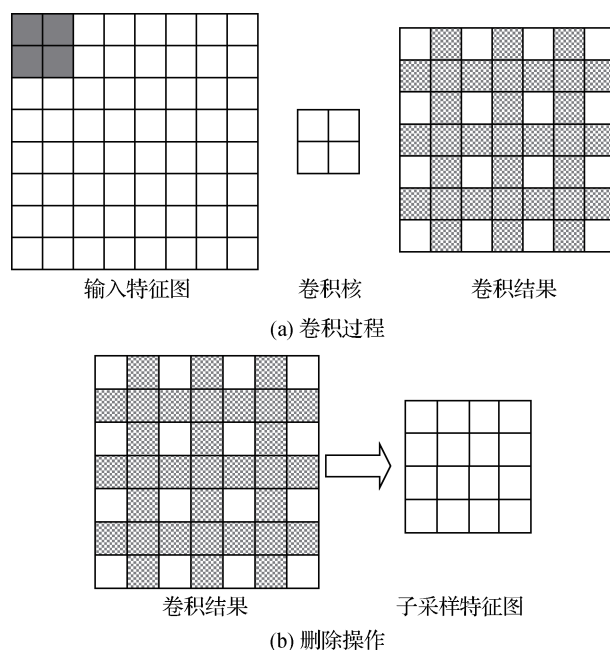


图7 池化的实现过程

Fig. 7 The process of realizing pooling

((a) convolution process; (b) delete operation)

(b)中的删除操作, 最后得到的结果就是最大池化模型的子采样特征图。

动态自适应池化算法在最大池化算法求得3维矩阵 T_1 时, 进一步对其进行求和运算, 即对池化域中的所有元素求和, 得到了 $1 \times 7 \times 7$ 的3维矩阵 T_3 。 T_3 和 T_2 进行相减, 再取平均, 得到的矩阵中的元素就是式(9)中平均值 a 。进一步计算即可求得动态自适应池化模型的子采样特征图。

3 实验结果与分析

实验共分为两部分, 第1部分是在手写数字数据集上进行研究性实验, 第2部分是在其他数据集上进行验证性实验。由于卷积神经网络对手写数据有着良好学习性能, 所以使用手写数字集对不同池化模型进行训练, 研究动态自适应池化算法对卷积神经网络的收敛速度、识别率和学习性能的提升效果; 为了验证动态自适应池化模型在其他数据集上对卷积神经网络的提升效果, 分别在车牌数据集和人脸数据集上进行实验, 以验证动态自适应池化模型对卷积神经网络相关算法的普适性。

3.1 手写数据集实验与分析

实验在 Matlab2012a 环境下实现, 计算机的

CPU 型号为 Intel(R) Core(TM) i7-3610QM, 内存为 8 GB。实验数据使用 MNIST 手写数字库 (<http://yann.lecun.com/exdb/mnist/>), 其中有 7 万条手写数字图片数据, 6 万条作为训练数据, 每次训练选择 50 条数据, 所以一个迭代周期训练 1 200 次; 其余 1 万条作为检测数据, 对模型进行检测以计算其识别准确率。输入的图像大小为 28×28 , 实验过程中校正误差项 θ 取 0 值, 分别对使用平均池化模型、最大池化模型、中间值池化模型和动态自适应池化模型的卷积神经网络进行实验, 选择迭代次数并记录各个周期耗时, 并绘制均方误差曲线。

3.1.1 平均池化模型实验

卷积神经网络使用平均池化模型进行学习, 根据迭代次数进行 3 次实验, 每次训练的迭代次数分别为 1、2 和 3 次。记录每次实验的耗时和学习后模型的误识率, 如表 1 所示。均方误差曲线如图 8 所示。

表 1 平均池化周期耗时与模型误识率
Table 1 Time and model error rate of mean pooling

迭代次数	周期 1/s	周期 2/s	周期 3/s	误识率/%
1	89.002 3			11.13
2	87.624 1	86.196 6		7.75
3	86.438 0	89.531 2	88.572 4	6.10

由表 1 可以看出, 随着训练迭代次数的增加, 完成实验所使用的时间也在增加, 并且与迭代次数基本成正比。迭代次数越多, 模型准确率越高。

图 8 是 1~3 个周期内均方误差随着训练批次的变化曲线。从 (1, 1.463) 到 (1 200, 0.121 3) 为迭

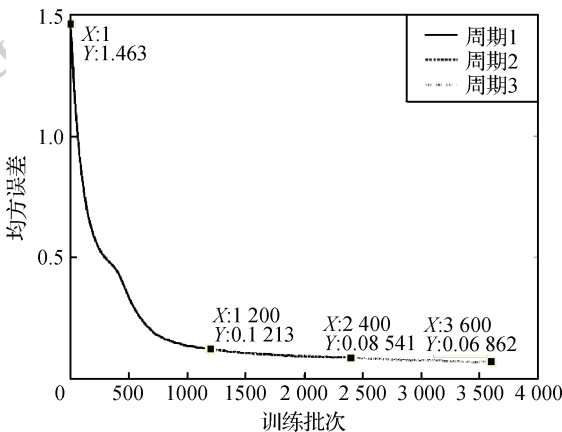


图 8 平均池化均方误差曲线

Fig. 8 Mean square error curve of mean pooling

代 1 次时 1 个周期的均方误差曲线; 从 (1, 1.463) 到 (2 400, 0.085 41) 为迭代 2 次时 2 个周期的曲线; 从 (1, 1.463) 到 (3 600, 0.068 62) 为迭代 3 次时 3 个周期的曲线。由图 8 可知, 迭代次数不同时, 平均池化模型的相同周期内的误差曲线基本吻合, 随着训练批次的增加, 误差值不断下降。

3.1.2 最大池化模型实验

卷积神经网络使用最大池化模型进行学习, 分别迭代 1、2、3 次并获得周期时间与误识率, 结果如表 2 所示, 实验过程均方误差曲线变化如图 9 所示。

表 2 最大池化周期耗时与模型误识率
Table 2 Time and model error rate of max pooling

迭代次数	周期 1/s	周期 2/s	周期 3/s	误识率/%
1	97.331 9			10.25
2	95.049 6	95.662 1		6.27
3	95.170 1	95.222 6	95.971 2	4.99

由表 2 可以看出, 随着迭代次数的增加, 最大池化与平均池化算法都有周期耗时正比增长和模型准确率增加的趋势, 而最大池化法识别的准确率要明显比平均池化法高。

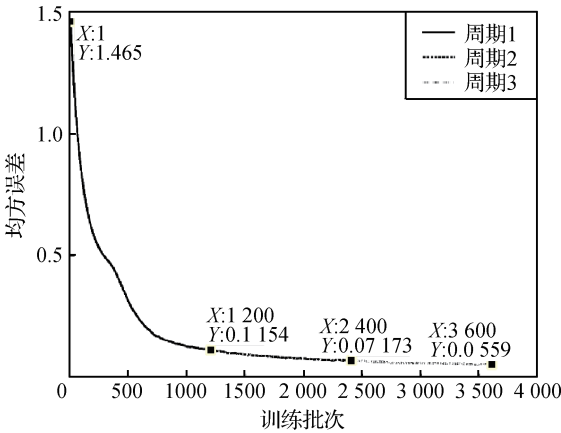


图 9 最大池化均方误差曲线

Fig. 9 Mean square error curve of max pooling

图 9 中从 (1, 1.465) 到 (1 200, 0.115 4) 为迭代 1 次时 1 个周期的均方误差曲线; 从 (1, 1.465) 到 (2 400, 0.071 73) 为迭代 2 次时 2 个周期的曲线; 从 (1, 1.465) 到 (3 600, 0.055 9) 为迭代 3 次时 3 个周期的曲线。相比较平均池化模型, 最大池化法使得误差下降更快。

3.1.3 中间值池化模型实验

卷积神经网络使用中间值池化模型进行学习,

学习过程同3.1.1节和3.1.2节相同,3次迭代的周期时间和最后模型的误识率如表3所示。实验过程均方误差随训练批次下降曲线如图10所示。

表3 中间值池化周期耗时与模型误识率
Table 3 Time and model error rate of intermediate value pooling

迭代次数	周期 1/s	周期 2/s	周期 3/s	误识率/%
1	98.395 0			10.73
2	97.475 3	97.577 0		7.02
3	101.590 9	101.271 0	100.538 8	5.34

由表3可以看出,中间值池化法相比较于前两种池化方法,周期耗时与准确率随迭代次数增加的趋势基本一致。平均池化耗时最少,但准确率也最低,其余两种池化提升准确率都以增加时间为代价。中间池化法的准确率要高于平均池化而低于最大池化算法。

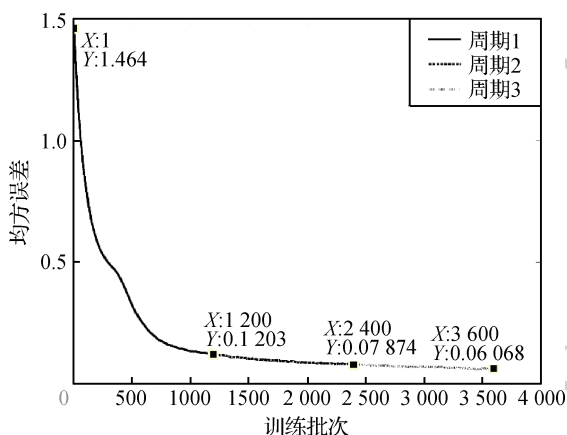


图10 中间值池化均方误差曲线

Fig. 10 Mean square error curve of intermediate value pooling

图10中从(1,1.464)到(1200,0.1203)为迭代1次时1个周期的均方误差曲线;从(1,1.464)到(2400,0.07874)为迭代2次时2个周期的曲线;从(1,1.464)到(3600,0.06068)为迭代3次时3个周期的曲线。从3条曲线的整体变化情况可以看出,中间池化模型的误差下降速度比平均池化要快,而比最大池化要慢。

3.1.4 动态自适应池化模型实验

卷积神经网络使用动态自适应池化模型进行学习,学习过程同前一致。学习结果的周期耗时与模型误识率如表4所示,实验过程中的均方误差随训练批次的下降曲线如图11所示。

表4 动态自适应池化周期耗时与模型误识率

Table 4 Time and model error rate of dynamic adaptive pooling

迭代次数	周期 1/s	周期 2/s	周期 3/s	误识率/%
1	98.292 8			9.59
2	96.536 0	98.865 2		6.20
3	95.896 5	96.011 8	96.079 4	4.99

由表4可以看出,动态自适应池化模型相比较其他3种模型,学习后的模型准确率达到最高。中间池化算法虽然同平均池化算法相比准确率有所提高,但同时耗时也增加不少,而动态自适应池化算法并没有以更多的时间消耗为代价去提高准确率。

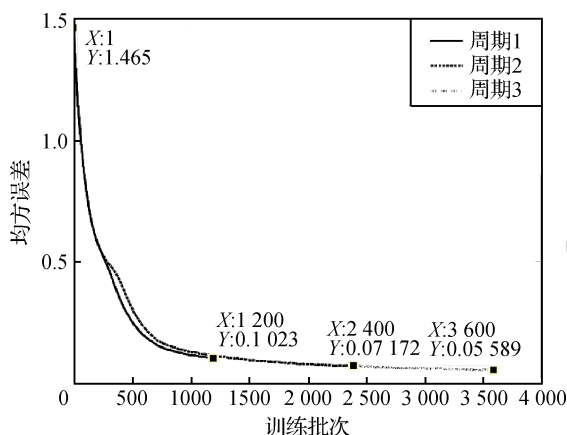


图11 动态自适应池化均方误差曲线

Fig. 11 Mean square error curve of dynamic adaptive pooling

图11中从(1,1.465)到(1200,0.1023)为迭代1次时1个周期的均方误差曲线;从(1,1.465)到(2400,0.07172)为迭代2次时2个周期的曲线;从(1,1.465)到(3600,0.05589)为迭代3次时3个周期的曲线。动态自适应池化模型迭代次数为2和3时误差下降趋势基本一致,但是迭代次数为1时下降速率明显比较快,而且相比较其他模型,其模型的准确率也比较高。

3.1.5 实验分析

1) 迭代次数对实验模型准确率的影响。通过上面的实验可以看出,随着迭代次数的增加,模型的准确率随之增加,而均方误差逐渐减小。各迭代次数下模型的误识率如表5所示,识别率的提升效果如表6所示。

通过表5和表6的数据可以看出,相同池化算法的误识率随着迭代次数的增加而降低,卷积神经网络迭代次数越多,模型识别的准确率越高。

表5 不同迭代次数下各模型的误识率

Table 5 Error rate of each model with different iteration times

迭代次数	平均池化	中间值池化	最大池化	动态自适应池化
1	11.13	10.73	10.25	9.59
2	7.75	7.02	6.27	6.20
3	6.10	5.34	4.99	4.99

表6 迭代 n 次比迭代 m 次误识率减少百分比Table 6 Error rate percentage reduction of iterations n than m

n	m	平均池化	中间值池化	最大池化	动态自适应池化
2	1	30.37	34.58	38.83	35.35
3	2	21.29	23.93	20.41	19.52

2) 动态自适应池化模型对收敛速度的影响。根据各模型的均方误差曲线及其在周期 1~3 内的误差下降程度和动态自适应池化模型对收敛速度的影响,分析动态自适应池化模型较其他模型对收敛速度的提升效果,结果如表 7 所示。

表7 动态自适应池化对收敛速度的提升

Table 7 The Improvement of convergence speed

迭代次数	较平均池化	较中间值池化	较最大池化
1	15.66	14.96	11.35
2	16.03	8.92	0.01
3	18.55	7.89	0.02

由表 7 可以看出,动态自适应池化算法较其余池化算法均方误差率下降,收敛速度提高,最高可提升 18.55%。因此可知,动态自适应池化模型可以较大程度降低均方误差并较大幅度提升卷积神经网络的收敛速度。

3) 动态自适应池化模型对准确率的影响。分析表 5 数据可知,动态自适应池化模型与其余池化模型相比较,对图像的误识率有一定的降低,降低效果如表 8 所示。

由表 8 可知,相比较于其他 3 个池化模型,动态自适应池化模型学习后识别的准确率最高,而平均池化最低。在 2 次迭代的情况下前者比后者模型的误识率降低了 20%。在迭代次数为 3 次时,动态自

表8 动态自适应池化对误识率的降低

Table 8 The error rate reduction of dynamic adaptive pooling

迭代次数	较平均池化	较中间值池化	较最大池化
1	13.83	10.62	6.45
2	20.00	11.68	1.12
3	18.20	6.55	0.00

适应池化算法与最大池化算法的准确率基本一致。

4) 实验结果。随着迭代次数的增加,动态自适应池化算法的准确率趋近于最大池化算法,而平均池化算法的准确率最低。这种实验结果与实验数据有关。实验数据样本如图 12 所示。

图 12 为手写数字 0~9 的 10 个图像样本数据,样本的输入大小为 28×28 ,由图 12 可知,大部分像素点并没有数据,手写数字只占用了很小的一部分。因此使用平均池化法会降低样本特征选取的准确性,所以平均池化法的模型准确率最低。随着迭代次数的增加,不同池化模型对准确率的绝对影响逐渐减弱,并且在实验程中通过计算误差反向传播算法更新权值,使得在特征提取阶段的误差逐渐减小。所以迭代次数越大,动态自适应算法与最大池化算法的模型准确率越接近。

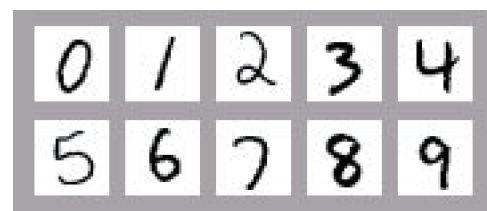


图12 实验数据样本

Fig. 12 Experimental data sample

5) 实验总结。综合比较各个池化模型的均方误差曲线可知,误差值随着训练批次增长的下降趋势基本一致,但是收敛速度却不相同,特别是在 1 次迭代的情况下,图 11 中的曲线收敛速度提高明显。动态自适应的收敛速度比其他模型最高可以提升 18.55%。所以动态自适应池化模型的收敛性更好。

由各个池化模型的实验结果可以看出,动态自适应池化模型的准确率要明显高于其余 3 种池化模型,而且模型的误识率最多可以降低 20%。所以使用动态自适应池化算法的卷积神经网络的准确率有明显优势。

综合分析上述实验可知,与其余3个模型比较,动态自适应池化模型是最优模型,且对卷积神经网络的收敛性、模型的准确率和网络的学习性能有较大提升。

3.2 其他数据集实验与分析

通过改变数据集和训练与测试的任务量,验证动态自适应算法对使用卷积神经网络进行分类识别的2维图像具有普适性。现分别使卷积神经网络对车牌数字和人脸图像进行学习与分类,所有实验环境均与手写数字集上实验环境相同,记录均方误差和识别的错误率。

3.2.1 车牌数据集实验结果与分析

为验证动态自适应池化模型为最优模型,分别使用不同池化模型在车牌数据集上进行实验。本次实验所使用的车牌数据均为真实车牌数字。通过收集真实车牌并对其图像分割,选取数字建立车牌数据库以供卷积神经网络训练与测试。车牌数据库数字如图13所示。数据库中共有3 000条数据,其中2 000条作为训练数据,每次训练20条数据,每个迭代周期训练100次;其余1 000条作为测试数据,对模型进行检测以计算其识别准确率。其余各项参数同手写数字集实验相同。分别使用平均池化模型、最大池化模型、中间池化模型和动态自适应池化模型的卷积神经网络进行实验,并选择1次、2次

和3次迭代周期进行训练,记录各模型周期耗时和均方误差。

图13中第1行为车牌真实照片的截取片段,其中黄色框线内的数字依次对应第2行的各个数字,第2行为车牌数据库内的数据,即将车牌数字截取后统一大小,共同构成车牌数据库。因拍摄角度与光线等不同因素的影响,截取的各个数字的角度和比例均不相同。第3行为卷积神经网络的输入图像。

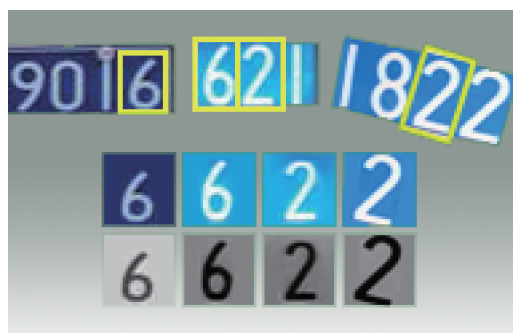


图13 车牌数字

Fig. 13 License plate number

卷积神经网络使用4种池化模型在不同迭代次数下的对车牌数字识别的错误率和各个迭代周期的耗时情况如表9所示。使用动态自适应池化模型的卷积神经网络相比较其余3种模型对车牌数字误差率的降低程度如表10所示。

表9 不同模型在各迭代次数下的误识率和周期耗时

Table 9 Error rate and time of each pooling model in iterations

池化模型	误识率/%			周期耗时/s		
	迭代1次	迭代2次	迭代3次	1个周期	2个周期	3个周期
平均池化	12.7	11.4	9.9	53.401 3	104.292 4	158.731 0
中间值池化	12.3	9.1	7.5	59.036 9	117.031 35	182.045 1
最大值池化	11.9	8.3	7.2	58.399 1	114.427 1	171.818 5
动态自适应池化	11.3	8.2	7.1	58.975 6	117.240 7	172.792 5

表10 动态自适应池化对误识率的降低

Table 10 The error rate reduction of dynamic adaptive pooling

迭代次数	较平均池化	较中间值池化	较最大池化
1	11.02	8.13	5.04
2	28.07	9.89	1.20
3	28.28	5.33	1.38

由表9可知,4种池化模型的周期耗时都随着

迭代次数以倍数增长,增长趋势基本一致。在任意一次迭代下,动态自适应池化模型识别的错误率都最低。而且由表10可知,动态自适应池化模型较其他模型的误识率最多可以降低28.28%。由此可知,动态池化模型在没有增加时间代价的前提下提高了卷积神经网络的识别的准确率。

卷积神经网络使用4种不同池化模型在对车牌数字学习的过程中均方误差的示意图如图14所示,其中收敛指数为均方根误差开方的值,收敛指数越

小,其收敛性越强,各算法起始收敛性都相同。

由图 14 可知,在不同迭代次数下,动态自适应池化算法的均方误差始终最低,达到一定收敛程度时使用的迭代次数最少,收敛速度最快。

根据在车牌数据集上的实验可知,与其余 3 种池化模型比较,使用动态自适应池化模型的卷积神经网络在对车牌数字识别过程中,耗时基本一致,收敛速度最快,识别性能最优。

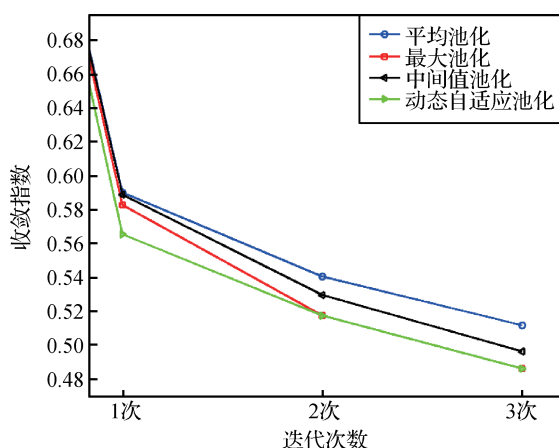


图 14 不同迭代次数下各模型收敛情况

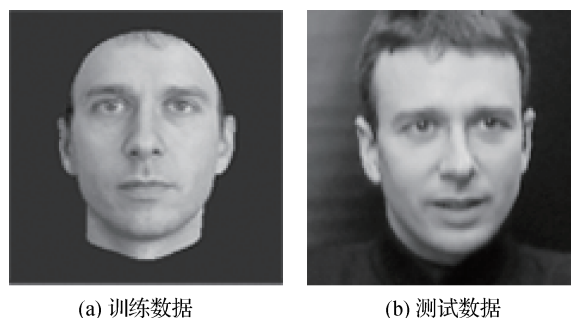
Fig. 14 Convergence of each model in iterations

3.2.2 人脸数据集实验结果与分析

使用人脸数据库来训练使用不同池化模型的卷积神经网络,通过识别的准确率和算法过程的收敛速度来验证动态自适应池化模型对卷积神经网络的学习性能提升效果。

使用的数据库为麻省理工学院生物与计算学习中心提供的人脸数据库 (http://cbcl.mit.edu/software-datasets/heisele_download/)。其中共有 10 个人的有脸图像,每个人的训练图像为 324 幅,共有 3 240 条训练数据;每个人的测试图像为 200 幅,共有 2 000 条测试数据。训练数据和测试数据分别如图 15(a)(b)所示。其余各项参数与 3.1 节和 3.2.1 节实验参数相同。使用平均池化模型、最大池化模型、中间池化模型和动态自适应池化模型的卷积神经网络对人脸库中的训练数据进行学习,然后使用训练后的模型对测试数据进行识别。分别进行 3 次迭代实验,记录各个迭代次数下的周期耗时、识别的错误率和均方误差。

各个模型的卷积神经网络在不同的迭代次数下的误识率如表 11 所示。其中最大降低率指动态自



(a) 训练数据 (b) 测试数据

图 15 人脸训练数据与测试数据

Fig. 15 Training and test data of face ((a) training data; (b) test data)

适应池化模型对其余 3 种模型在各个迭代次数下误识率降低的最大值。

表 11 不同池化模型在各迭代次数下的误识率

Table 11 Error rate of each pooling model in iterations

池化模型	误识率 /%		
	迭代 1 次	迭代 2 次	迭代 3 次
平均池化	19.5	16.7	13.3
中间值池化	18.3	12.9	10.1
最大池化	17.8	11.8	9.6
动态自适应池化	16.6	11.7	9.5
最大降低率	14.87	29.94	28.57

由表 11 可以看出,虽然对人脸的误识率要比对手写数字的误识率高,但是其规律并没有发生变化:随着迭代次数的增加误识率降低,对于不同的迭代次数,使用动态自适应池化模型的卷积神经网络的识别率依然最高。

3 次迭代实验下,各个池化模型算法的收敛性与耗时情况如图 16 所示。其中左侧纵坐标以均方根误差开方的值作为收敛指数,收敛指数越小,其收

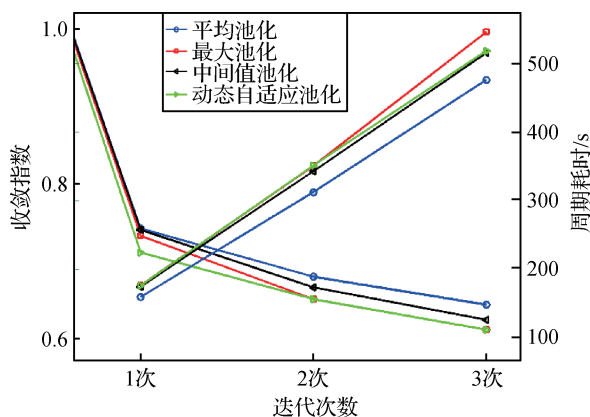


图 16 不同迭代次数下的收敛性与耗时情况

Fig. 16 Convergence and time in iterations

敛性越强。各算法起始收敛性都相同。右侧为周期耗时,显示不同迭代次数下各个池化模型的耗时递增的情况。

图16中每种池化模型随着迭代次数的增加变化趋势基本一致。随着迭代次数的增加,所有池化模型的收敛速度加快,收敛指数变小,收敛性增强;而周期耗时基本保持持续线性增长。

根据在人脸数据集上的实验可知,在各个迭代次数下动态自适应池化模型的收敛速度都为最快,并且相比较于收敛速度与识别率的变化,每次迭代实验的周期耗时变化基本可以忽略,因此动态自适应池化算法提高算法的收敛速度和识别率没有增大时间代价。由此可知,使用动态自适应池化模型的卷积神经网络与使用其余3种池化模型的网络相比较,在其对人脸识别的过程中耗时基本相同,收敛速度最快,识别性能最优。

综合上述可知,无论是在车牌数据集上还是在人脸数据集上,使用动态自适应池化模型的卷积神经网络相比较于其他池化模型都为最优模型,对卷积神经网络的收敛性和识别率以及学习性能都有较大的提升,因此动态自适应池化模型适用于使用卷积神经网络处理2维图像的深度学习算法。

4 结 论

通过在卷积神经网络的特征提取部分对不同池化模型进行选择,对比不同模型对特征提取的优化程度,分析其收敛速度、准确率以及对网络学习性能的影响,提出了一种基于最大池化算法改进的动态自适应池化模型,并对多种池化模型进行对比实验,研究改进池化模型对特征提取的优化程度和对收敛速度、模型准确率的提升效果以及对网络学习性能的影响。实验结果表明,使用动态自适应池化算法的卷积神经网络识别的准确率最高,均方误差下降速度最快,收敛性最好。因此可以得出结论,动态自适应池化模型对卷积神经网络的学习性能提升最大,效果最佳。这种模型也可以进一步应用于其他与卷积神经网络相关的深度学习算法。

参考文献 (References)

[1] Sun Z J, Xue L, Xu Y M, et al. Overview of deep learning[J].

Application Research of Computers, 2012, 29(8): 2806-2810.

[孙志军, 薛磊, 许阳明, 等. 深度学习研究综述[J]. 计算机应用研究, 2012, 29(8): 2806-2810.] [DOI: 10.3969/j.issn.1001-3695.2012.08.002]

[2] Yu K, Jia L, Chen Y Q, et al. Deep learning: yesterday, today and tomorrow[J]. Journal of Computer Research and Development, 2013, 50(9): 1799-1804. [余凯, 贾磊, 陈雨强, 等. 深度学习的昨天、今天和明天[J]. 计算机研究与发展, 2013, 50(9): 1799-1804.] [DOI: 10.7544/j.issn1000-1239.2013.20131180]

[3] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507. [DOI: 10.1126/science.1127647]

[4] Liu J W, Liu Y, Luo X L. Research and development on deep learning[J]. Application Research of Computers, 2014, 31(7): 1921-1930. [刘建伟, 刘媛, 罗雄麟. 深度学习研究进展[J]. 计算机应用研究, 2014, 31(7): 1921-1930.] [DOI: 10.3969/j.issn.1001-3695.2014.07.001]

[5] Wu F, Zhu W W, Yu J Q. Researches on multimedia technology 2014-deep learning and multimedia computing[J]. Journal of Image and Graphics, 2015, 20(11): 1423-1433. [吴飞, 朱文武, 于俊清. 多媒体技术研究: 2014-深度学习与媒体计算[J]. 中国图象图形学报, 2015, 20(11): 1423-1433.] [DOI: 10.11834/jig.20151101]

[6] Zheng Y, Chen Q Q, Zhang Y J. Deep learning and its new progress in object and behavior recognition[J]. Journal of Image and Graphics, 2014, 19(2): 175-184. [郑胤, 陈权崎, 章毓晋. 深度学习及其在目标和行为识别中的新进展[J]. 中国图象图形学报, 2014, 19(2): 175-184.] [DOI: 10.11834/jig.20140202]

[7] Graves A, Mohamed A R, Hinton G. Speech recognition with deep recurrent neural networks[C]//Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Vancouver BC: IEEE, 2013: 6646-6649. [DOI: 10.1109/ICASSP.2013.6638947]

[8] Deng L, Hinton G, Kingsbury B. New types of deep neural network learning for speech recognition and related applications: an overview[C]//Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Vancouver BC: IEEE, 2013: 8599-8603. [DOI: 10.1109/ICASSP.2013.6639344]

[9] Xu P, Bo H. Facial expression recognition based on CNN[J]. Microcomputer & Its Applications, 2015, 34(12): 45-47. [徐鹏, 薄华. 基于卷积神经网络的人脸表情识别[J]. 微型机与应用, 2015, 34(12): 45-47.] [DOI: 10.3969/j.issn.1674-7720.2015.12.015]

[10] Zhang C, Zhang Z Y. Improving multiview face detection with multi-task deep convolutional neural networks[C]//Proceedings of the IEEE Winter Conference on Applications of Computer Vi-

- sion. Steamboat Springs, CO: IEEE, 2014: 1036-1041. [DOI: 10.1109/WACV.2014.6835990]
- [11] 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. Montreal: ACM, 2009: 609-616. [DOI: 10.1145/1553374.1553453]
- [12] Matsugu M, Mori K, Mitari Y, et al. Subject independent facial expression recognition with robust face detection using a convolutional neural network [J]. Neural Networks, 2003, 16(5-6): 555-559. [DOI: 10.1016/S0893-6080(03)00115-1]
- [13] Huang G B, Lee H, Learned-Miller E. Learning hierarchical representations for face verification with convolutional deep belief networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Providence, RI: IEEE, 2012: 2518-2525. [DOI: 10.1109/CVPR.2012.6247968]
- [14] Yu Y W, Yin G F, Yin Y, et al. Defect recognition for radiographic image based on deep learning network [J]. Chinese Journal of Scientific Instrument, 2014, 35(9): 2012-2019. [余永维, 殷国富, 殷鹰, 等. 基于深度学习网络的射线图像缺陷识别方法 [J]. 仪器仪表学报, 2014, 35(9): 2012-2019.]
- [15] Cai J, Cai J Y, Liao X D, et al. Preliminary study on hand gesture recognition based on convolutional neural network [J]. Computer Systems & Applications, 2015, 24(4): 113-117. [蔡娟, 蔡坚勇, 廖晓东, 等. 基于卷积神经网络的手势识别初探 [J]. 计算机系统应用, 2015, 24(4): 113-117.] [DOI: 10.3969/j.issn.1003-3254.2015.04.019]
- [16] Yuan A Q, Bai G, Jiao L J, et al. Offline handwritten English character recognition based on convolutional neural network [C]//Proceedings of the 10th IAPR International Workshop on Document Analysis Systems. Gold Coast, QLD: IEEE, 2012: 125-129. [DOI: 10.1109/DAS.2012.61]
- [17] Giusti A, Cireşan D C, Masci J, et al. Fast image scanning with deep max-pooling convolutional neural networks [C]//Proceedings of IEEE International Conference on Image Processing. Melbourne, VIC: IEEE, 2013: 4034-4038. [DOI: 10.1109/ICIP.2013.6738831]
- [18] Nagi J, Ducatelle F, Caro G A D, et al. Max-pooling convolutional neural networks for vision-based hand gesture recognition [C]//Proceedings of the IEEE International Conference on Signal and Image Processing Applications. Kuala Lumpur: IEEE, 2011: 342-347. [DOI: 10.1109/ICSIPA.2011.6144164]
- [19] 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. Switzerland: Springer, 2014: 392-407. [DOI: 10.1007/978-3-319-10584-0_26]