

引用格式:周悦,曾上游,杨远飞,等. 基于分组模块的卷积神经网络设计[J]. 微电子学与计算机,2019,36(2):68-72. [ZHOU Y,ZENG S Y,YANG Y F,et al. Design of convolutional network based on slice[J]. Microeletronics & Computer,2019,36(2): 68-72.]

基于分组模块的卷积神经网络设计

周 悦,曾上游,杨远飞,冯燕燕,潘 兵
(广西师范大学 电子工程学院,广西 桂林 541004)

摘 要: 本文提出了一种基于分组模块的卷积结构,先将输出特征图分成数量相等的两组,每一组采用不同的卷积核进行操作以提取更充分的信息,而后将分组得到的多样性特征图像进行级联,最后通过 1×1 的卷积核将所有特征图整合. 与传统的 CNN 比较,使用本文模块的网络在数据集 Caltech256 上将识别率由 50.1% 提升至 52.2%. 在数据集 101_food 上将识别精度由 66.3% 提升至 68.9%. 实验结果表明网络在识别精度性能上有所提升.

关键词: 卷积神经网络;分组;网络性能;多卷积核

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 1000-7180(2019)02-0068-05

Design of Convolutional Network Based on Slice

ZHOU Yue,ZENG Shang-you,YANG Yuan-fei,FENG Yan-yan,PAN Bing

(College of Electronic Engineering, Guangxi Normal University,Guilin 541004, China)

Abstract: This paper proposes a grouping method of convolution structure for feature extraction of images. The model increases the depth of network, and the network parameters changes smaller. Compared with the traditional CNN, the network using the module of this paper increases the recognition rate from 50.1% to 52.2% on the data set Caltech256. On the data set 101_food, the recognition accuracy increased from 66.3% to 68.9%. Experimental results show that the network structure in the paper improves the network performance.

Key words: convolutional neural network; Grouping; Network performance; Multiple convolution kernels

1 引言

目前,深度学习算法已成为人工智能领域重要的算法^[1],无论是在图像识别还是自然语言处理等方面,深度学习算法的表现都远远优于其他算法.深度学习所代表的人工智能技术,已经开始慢慢的改变人们的日常生活,研究者已经在自动驾驶,实时语音处理等多个方面取得了重大突破,对人工智能的研究进入了一个新的阶段.

CNN^[2]是深度学习算法中的重要的一部分.

传统的卷积神经网络是通过卷积核对上层输出进行卷积操作得到卷积层的输出特征图,但是卷积

操作只有一条通道,提取的图像特征不充分^[3].传统的层叠式网络,基本上都是一个卷积层的堆叠,每层只用一个尺寸的卷积核,例如 VGG 结构中使用了大量的 3×3 卷积层.本文为了提高网络的分类精度,设计出一种新的网络模块代替传统的 3×3 卷积核,提高了网络的分类精确度.

2 卷积神经网络

卷积神经网络的基本网络结构可以分为五个部分:卷积层、激励层、下采样层、BN^[4]层、全连接层.

2.1 卷积层

卷积神经网络的卷积层,主要作用是提取输入

特征图的特征.不同的卷积核提取的特征不同,多个小卷积核的使用可以实现和较大卷积核一样的效果,CNN在进行卷积提取特征的过程中采用权值共享的策略,同一种卷积核之间权值相互共享,大量的减少了网络参数,加快了网络训练时间.其计算公式为:

$$x_m^l = f(\sum x_n^{l-1} * k_{mn}^l + b) \quad (1)$$

式中, $f(\cdot)$ 代表激活函数; x_m^l 表示下层输出; x_n^{l-1} 表示上层的输出; k 代表该层卷积核的权值; b 代表该层卷积核的偏置.

2.2 激励层

卷积神经网络的激励层,主要作用是对已经提取的特征进行非线性映射.激励层采用的激励函数一直在变化,从早期的 sigmoid 或 tanh 等饱和和线性函数到现在大量使用的 Relu^[5] 不饱和和非线性函数. Relu 激励函数的使用在一定程度上解决的梯度弥散的现象. Relu 函数较为简单计算,节约了网络训练过程中的计算开销,使得训练时间大量减少.

2.3 下采样层

卷积神经网络的下采样层又叫做 Pooling 层,主要作用是利用图像局部相关性原理,对图像进行子抽样.下采样层的计算公式为公式(2).经过下采样后,主要特征仍会保留,但特征图的维度有所减少,减少了网络的数据处理量.在卷积神经网络中经常使用的下采样操作有三种:平均池化,最大池化,全局池化.平均池化:采样后的值取采样区域内所有像素值的平均值.最大池化:采样后的值取采样区域内所有像素值的最大值.全局池化:采样后的值取整张特征图所有像素值的平均值或者最大值.

$$x_m^l = f(\beta_m^l \text{down}(x_m^{l-1}) + b) \quad (2)$$

式中, $f(\cdot)$ 代表激活函数; x_m^l 表示下层输出; x_m^{l-1} 表示上层的输出; β 代表权重系数; $\text{down}()$ 代表下采样函数; b 代表偏置项.

2.4 BN 层

卷积神经网络的激励层,主要作用是将每一层的输入归一化. BN 层将输入数据的分布归一化到均值为 0,方差为 1,可以提高网络的泛化能力,一定程度的避免过拟合.加入 BN 层之后,输入的数据被归一化到特定范围,因此在训练过程中可以减少对初始化的依赖,训练时可以选择比较大的初始学习

率,使训练速度加快. BN 层计算公式如下:

$$\mu_B \leftarrow \frac{1}{m} \sum_{k=1}^m x_k \quad (3)$$

$$\sigma_B^2 \leftarrow \frac{1}{m} \sum_{k=1}^m (x_k - \mu_B)^2 \quad (4)$$

$$\hat{x}_k \leftarrow \frac{x_k - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad (5)$$

$$y_k \leftarrow \gamma \hat{x}_k + \beta \equiv \text{BN}_{\gamma, \beta}(x_k) \quad (6)$$

2.5 全连接层

卷积神经网络的激励层,主要作用是将提取的特征图转化为一维特征向量.转化的一维特征向量即为学习到的图片的空间分布向量.将其作为分类器的输入,通过分类器进行分类.

$$S_j = f(\sum x_i * w_{ji} + b_j) \quad (7)$$

式中, $f(\cdot)$ 代表激活函数; w 代表全连接层神经元的权值; x 代表上层神经元的输出; b 代表神经元的偏置参数.

3 基于分组与级联模块的卷积神经网络模型

3.1 模型原理

传统的 CNN 通过单一大小的卷积核对特征图进行卷积操作,提取图片特征.由于通道的单一性,可能导致提取的图像特征不够充分.图 1 为传统 3×3 卷积核的卷积结构.

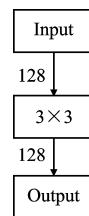


图 1 传统 3×3 卷积核卷积结构

本文提出的分组的卷积模块将输入特征图等分成两条支路,一条支路采用 3×3 的卷积核进行卷积提取特征;另一条支路采用等量 5×5 的卷积核进行卷积提取特征.而后,通过级联操作将两条支路上的特征图结合在一起,形成一组新的特征图,使得采集到的特征信息更加充分.提升卷积神经网络识别精度的一种有效的方法是增加网络深度^[6],相比于单层卷积模型,多层卷积能够对每个局部感知域进行更为繁琐的计算,从而逼近期望得到的特征结果^[7].最后,将新的特征图通过 1×1 的卷积核整合特征图并增加网络的深度.图 2 为本文分组的卷积模块.

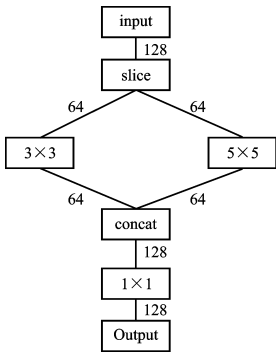


图 2 分组模块

用一个 128 张特征图输入的例子来解释一下本文分组卷积模块:把 128 个输入通道平均分为 2 组,每组拥有 64 个通道,并分别对 2 组单独做卷积运算.一组使用 3×3 的卷积核进行卷积运算,另一组采用 5×5 的卷积核进行卷积操作.分组操作可以减少参数,但同时,由于每组卷积之间不存在交互,不同组的输出通道与输入通道并不相关.为了让每一个输入通道实现关联,在将两个支路的输出级联之后,使用 1×1 的卷积核进行第二次卷积.第二次卷积过程中,对所有的特征图进行卷积操作以达到整合特征图和增加网络深度的目的.虽然增加深度神经网络深度能一定程度提升网络的性能,但是大的网络结构需要学习更多的参数,容易造成网络对训练数据集的过拟合^[8].本文模块与传统的 3×3 卷积结构相比参数改变较小,在输入和输出均为 128 张特征图时,传统的 3×3 卷积结构参数为 $128 \times 3 \times 3 \times 128=147456$,本文使用的卷积模块参数为 $64 \times 3 \times 3 \times 64+64 \times 5 \times 5 \times 64+128 \times 1 \times 1 \times 128=155648$.在参数量变化较小的情况下,本文分组模块增加网络深度,同时使用两种卷积核采样,增加了提取特征的多样性.

3.2 模型设置

本文设计了两种网络进行对比,分别对比传统的 CNN 网络 traNet 和使用本文分组模块的 mysliceNet 网络在两个数据集上的性能优劣.为了使本文对比实验结果更具有说服力,两种网络的整体结构的层次深度保持一致.传统的 traNet 结构结构如表 1,由 6 个卷积层、4 个池化层构成.第一个卷积核为 7×7 ,卷积之后接一个 pooling 层进行下采样,后为两个卷积层和 pooling 层的交替使用.在第三个 pooling 层之后,采用一个 3×3 的卷积层进行采样,之后为全局池化层和全连接层.

表 1 traNet 结构

层级结构	卷积核大小/步长/pad	输出维度
Input	—	$227 \times 227 \times 3$
Conv1/BN	$7 \times 7/2/0$	$111 \times 111 \times 48$
MaxPool1	$3 \times 3/2/0$	$55 \times 55 \times 48$
Conv2/BN	$3 \times 3/1/1$	$55 \times 55 \times 128$
Conv3/BN	$3 \times 3/1/1$	$55 \times 55 \times 128$
MaxPool2	$3 \times 3/2/0$	$27 \times 27 \times 128$
Conv4/BN	$3 \times 3/1/1$	$27 \times 27 \times 128$
Conv5/BN	$3 \times 3/1/1$	$27 \times 27 \times 128$
Max/Pool3	$3 \times 3/2/0$	$13 \times 13 \times 128$
Conv6/BN	$3 \times 3/1/1$	$13 \times 13 \times 384$
global_pooling/AVE	—	$1 \times 1 \times 384$
Fc1	—	256/101

使用本文分组模块的 mysliceNet 网络如表 2.将 traNet 除 Conv1 和 Conv6 之外的卷积层均使用本文分组模块代替.

表 2 mysliceNet 结构

层级结构	卷积核大小/步长/pad	输出维度
Input	— — —	$227 \times 227 \times 3$
Conv1/BN	$7 \times 7/2/0$	$111 \times 111 \times 48$
Max/Pool1	$3 \times 3/2/0$	$55 \times 55 \times 48$
分组模块	$3 \times 3/1/1$ $5 \times 5/2/1$	$55 \times 55 \times 128$
分组模块	$3 \times 3/1/1$ $5 \times 5/2/1$	$55 \times 55 \times 128$
Max/Pool2	$3 \times 3/2/0$	$27 \times 27 \times 128$
分组模块	$3 \times 3/1/1$ $5 \times 5/2/1$	$27 \times 27 \times 128$
分组模块	$3 \times 3/1/1$ $5 \times 5/2/1$	$27 \times 27 \times 128$
Max/Pool3	$3 \times 3/2/0$	$13 \times 13 \times 128$
Conv6/BN	$3 \times 3/1/1$	$13 \times 13 \times 384$
global pooling/AVE	— —	$1 \times 1 \times 384$
Fc1	—	256/101

4 实验

4.1 实验设置

本实验采用 caffe^[9]深度学习框架进行,计算机配置为 i7-6700K 四核 CPU、Ubuntu14.04 操作系统、32 GB 内存以及 NVIDIA-GTX1070 的 GPU.

数据集:本实验采用 101 类的 101_food 数据集

以及将背景类除去后的 256 类的 Caltech256 数据集作为数据集. 随机将该数据集分为 23 919 张训练图片与 5 862 张测试图片. 101_food 数据集每类图片 1 000 张图片,共 75 750 张训练图片和25 250张测试图片.

参数设置:实验在 caltech256 数据集上训练时, solver 文件配置为:初始学习率为 0. 01,学习率变化方式为 multistep,gamma 为 0. 1,stepvalue 为 24 000 和 48 000,最大迭代次数设为 60 000,每训练 500 次测试一次网络性能. 实验在 101_food 数据集上训练时, solver 文件配置为:初始学习率为 0. 005,学习率变化方式为 multistep,gamma 为 0. 1,stepvalue 为 40 000、80 000 和 120 000,最大迭代次数设置为 150 000,每训练 500 次测试一次网络性能.

图像预处理:训练之前对所有图片进行数据增强^[10],进行归一化和去均值操作,将所有的图片转化为 256 × 256 的规格,并对图片的每个像素都减去训练图片的平均值. 所有图片分别 5 个方位(左上角,右上角,左下角和右下角以及中间)随机裁剪为 227 × 227 大小,并进行水平翻转,以增加大量的数据.

4.2 实验结果及分析

本实验对比了传统的 CNN 和采用本文分组模块模块的网络在 Caltech256 以及 101_food 数据集上识别性能的好坏. 图 3 给出了两种网络在 Caltech256 训练时测试精度的变化曲线,当到达设置的最大迭代次数时,采用本文分组模块的 mysliceNet 识别精度明显高于传统的 CNN. 由表 3 可以看出 traCNN 在 caltech256 可达到 50. 1%的识别精度,而改进的 mysliceCNN 可到到 52. 2%的识别精度,在 traCNN 上有了 2. 1% 的提升.

表 3 各模型在 caltech256 上的性能

网络模型	准确率 (%)	训练时间 (min)	Caffemode 大小 (MB)
traNet	50. 1	111	4. 2
mysliceNet	52. 2	178	3. 5

图 4 给出了两种网络在 101_food 训练时测试精度的变化曲线,当到达设置的最大迭代次数 15 000 时,mysliceNet 识别精度依然明显高于传统的 CNN. 由表 4 可以看出 traCNN 在 caltech256 可达到 66. 3%的识别精度,而改进的 mysliceCNN 可到

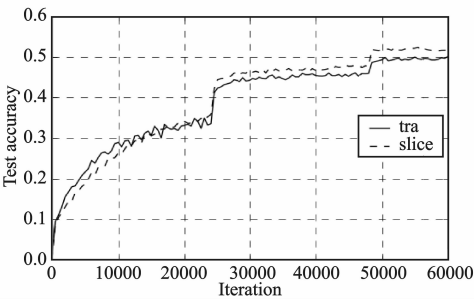


图 3 各模型在 caltech256 上的训练曲线

到 68. 9%的识别精度,在 traCNN 上有了 2. 6%的提升.

表 4 各模型在 101_food 上的性能

网络模型	准确率 (%)	训练时间 (min)	Caffemode 大小 (MB)
traNet	66. 3	381	4. 0
mysliceNet	68. 9	612	3. 2

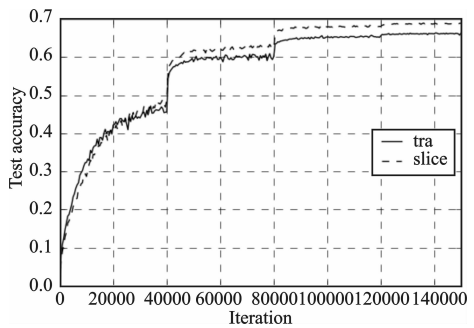


图 4 各模型在 101_food 上的训练曲线

从实验可以看出,改进网络在两个数据集中都有较好程度的分类精度提高. mysliceNet 由于结合了两条支路的特征所得出的分类精度比一个单层的卷积要高,但是因为结构深度加深并且网络中使用了 5 × 5 的较大卷积核而计算量加大,整个实验耗时也有所加长,但随着计算机硬件的提升,网络改进带来的参数增加而导致训练时间加长的的问题,很轻易得以解决,所以在准确率提高一定的前提下,训练时间的改变对整个网络增益没有大的影响.

5 结束语

本文通过设计一种网络模块进行特征提取,该模块使用了两种卷积核对特征图进行特征提取,增加了特征提取的多样性. 并将两种卷积提取的特征图级联在一起,使用 1 × 1 的卷积核进行整合,增加了网络的深度. 但在卷积提取特征之前,先使用

Slice 层将特征图等分为两份,分别使用两种卷积核提取特征,有利于参数的减少.使用该模块搭建卷积神经网络与传统的 CNN 比较在 Caltech256 和 101_food 数据集上的分类精度来对比网络性能.结果表明使用本文分组的卷积网络相比传统的 CNN 识别精度有了明显的提高,证明本文模块能有效的提升网络性能并控制模型大小.

参考文献:

- [1] Hinton G E, Salakhutdinov R R. Reducing the Dimensionality of Data with Neural Networks[J]. Science, 2006, 313(5786):504.
- [2] 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.
- [3] 杨远飞,曾上游. 基于并行和切片的深度卷积网络设计研究[J]. 微电子学与计算机, 2018, 35(3):120-124.
- [4] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]// International Conference on International Conference on Machine Learning. JMLR.org, 2015.
- [5] 周飞燕,金林鹏. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6):1229-1251.
- [6] 杨天祺,黄双喜. 改进卷积神经网络在分类与推荐中

的实例应用[J]. 计算机应用研究, 2018, 35(4): 974-977.

- [7] 张顺,龚怡宏. 深度卷积神经网络的发展及其在计算机视觉领域的应用[J]. 计算机学报, 2017(9).
- [8] Jia Y, Shelhamer E, Donahue J, et al. Caffe: Convolutional Architecture for Fast Feature Embedding[C]// ACM International Conference on Multimedia. ACM, 2014:675-678.
- [9] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks [C]//Advances in Neural Information Processing Systems. 2012: 1097-1105.
- [10] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [C]// International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012: 1097-1105.

作者简介:

周悦男, (1993-), 硕士研究生, 研究方向为深度学习. E-mail: gxnuzhouyue@163.com.

曾上游男, (1974-), 博士, 教授, 研究方向为非线性动力学、计算神经科学.