

文章编号:1671-6833(2019)02-0018-05

基于卷积神经网络的道路拥堵识别研究

罗荣辉,袁航,钟发海,聂上上

(郑州大学 物理工程学院,河南 郑州 450001)

摘要:为解决城市道路日益严峻的拥堵问题,结合深度学习和图像处理技术,提出了一种基于卷积神经网络的道路拥堵检测方法.此方法相对于传统机器视觉方法,无需前期提取道路背景,不受光照亮度和实际环境的影响,具有识别速度快、占用计算资源少、泛化性好等特点.现已在实际项目中得以应用,并取得了较好的效果.

关键词:卷积神经网络;深度学习;图像识别;拥堵检测;智慧城市

中图分类号:TU528.1 文献标志码:A doi:10.13705/j.issn.1671-6833.2019.02.008

0 引言

随着经济的快速进步和人均汽车保有量的提高,城市道路拥堵已成为阻碍城市现代化发展首要问题.国情所致,我国的一些城市发展过快,城市道路建设远远跟不上经济发展的脚步,使得道路拥堵问题更为严峻.拥堵问题已成为智慧城市建设道路上不可避免的障碍.而解决拥堵问题的首要条件便是实现城市道路拥堵的快速检测.

目前,道路拥堵的检测手段较多,在机器视觉方向主要有基于图像特征的传统检测手段<sup>[1]</sup>和基于深度学习技术的新型手段<sup>[2-3]</sup>.目前尚有欠缺的是使用深度学习技术针对城市复杂路况的研究.笔者采用卷积神经网络对环境千差万别的城市道路进行拥堵检测.此方法相对于传统检测手段,计算量小、鲁棒性好、泛化性强,具有良好的应用前景.

1 卷积神经网络

卷积神经网络(convolutional neural network, CNN)的概念出自科学家在研究动物视觉皮层细胞时提出的感受野(receptive field)<sup>[4]</sup>.后由日本科学家提出的认知机(neocognitron)概念,可算作CNN最初的实现原型<sup>[5]</sup>.此后便诞生了大量著名的卷积神经网络模型,如LeCun的LeNet5<sup>[6]</sup>、Alex

Krizhevsky的AlexNet<sup>[7]</sup>、牛津大学计算机视觉组(visual geometry group)的VGGNet<sup>[8]</sup>、Google的InceptionNet<sup>[9]</sup>以及微软研究院的ResNet<sup>[10]</sup>等.

卷积神经网络是一种基于特征的学习算法,它区别于常规的特征提取算法,不需要人工的提取特征,只需要大范围学习标注好的数据便可自我完成特征提取工作<sup>[11]</sup>.解决了因某些数据特征人工理解不准确、不到位或特征提取不足造成预测准确率较低的问题,如今已不仅仅在图像领域,更在视频<sup>[12]</sup>、音频<sup>[13]</sup>和自然语言处理<sup>[14]</sup>等领域得到了广泛的应用,并取得了优秀的成绩.

图1为一个经典的卷积神经网络结构.它由1个输入层、2个卷积层、2个池化层、2个全连接层和1个输出层组成.输入层将已预处理好的32×32格式的图片数据加载至神经网络结构中,第一个卷积层C1使用含6个卷积通道的5×5卷积核,以步长为1对输入的数据进行卷积操作,最终得到一个28×28×6的张量.之后是一个2×2的平均池化层S2对C1中产生的张量进行降采样得到14×14×6的张量.后面卷积层和池化层与前面原理类似,最终经过S4后得到一个5×5×16的张量,然后使用120个与之前尺寸一样的5×5的卷积核将数据转换为1×120的张量,因此构成全连接,也可以称作全连接层.最后经过下一个全连接层和输出层最终得到分类结果.

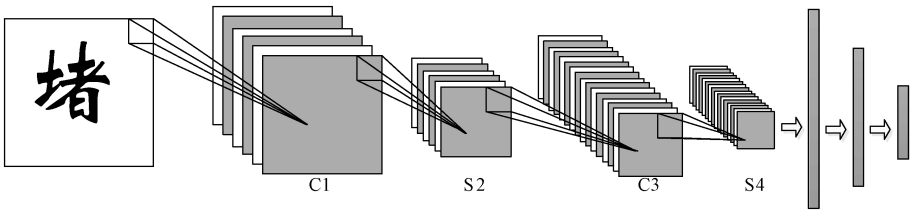


图 1 经典卷积神经网络结构

Fig. 1 Structure of classic convolutional neural network

与传统的深度神经网络结构相比,卷积神经网络具有局部连接、权值共享和池化层降采样的特点. 这些特点使得卷积神经网络结构参数量少、训练复杂度低、过拟合风险小并赋予了模型对平移和形变的容忍性,提高了模型的泛化能力.

2 模型介绍

2.1 模型设计与选择

卷积神经网络的构造对最终的实验结果有着重大的影响. 考虑训练数据、模型参数,硬件环境和应用场景. 最终在保证高准确率的情况下,笔者选择基于 AlexNet 的模型来构造本文网络.

AlexNet 模型与其他模型的区别主要是:①使用 Relu 作为卷积神经网络的激活函数,解决了在网络较深时的梯度弥散问题;②训练时使用 dropout 随机忽略掉部分神经元,防止过拟合;③使用重叠的最大池化,避免平均池化的模糊化效果;④提出了 LRN 层,使响应较大的神经元变得更大并忽略反馈较小的神经元<sup>[15]</sup>.

由于显存不足问题,AlexNet 原形在训练时使用了两块 GPU,将模型拆作两个部分. 笔者将 AlexNet 模型整合至一个 GPU 上训练,并根据数据实际情况,调整了输入图片的大小,并修改了部分卷积核尺寸,最终产生的模型如图 2 所示.

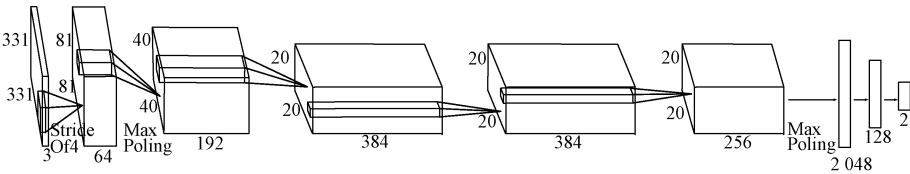


图 2 笔者卷积神经网络结构图

Fig. 2 The convolutional neural network structure of this paper

如图 2 所示,模型的输入图片尺寸为  $331 \times 331 \times 3$ . 第一个卷积层使用  $11 \times 11$  尺寸的较大卷积核,步长为 4,共有 64 个卷积通道,使用 Relu 激励函数,后接 LRN 层. 第二个卷积层使用  $5 \times 5$  尺寸的卷积核,步长为 1,共用 192 个卷积通道,使用 Relu 激励函数,后接 LRN 层和  $3 \times 3$  步长为 2 的最大池化层. 之后的 3 个卷积层结构与前面类似,卷积核尺寸均为  $3 \times 3$ ,使用的卷积通道数量分别为 384 个、384 个和 256 个,与前面不同的是这 3 个卷积层不再使用 LRN 层和 Max\_pool 层. 在所有卷积完成后连接一个  $2 \times 2$  的最大池化层且步长为 2,最终得到一个尺寸为  $10 \times 10 \times 256$  的数据. 后面使用 3 个包含 dropout 的全连接层进行全连接,最终输出预测值.

2.2 数据增强

本文中的数据集来自互联网数据,从数据分布来看,样本数据极不均衡,使用此数据训练出的模型泛化性较差,识别准确率较低. 为解决这一问题,笔者对样本数较少的非堵车数据使用了过采

样技术<sup>[16]</sup>. 为防止因过采样而造成测试集数据与训练集重复使得测试准确率虚高,笔者仅对训练集中的非堵车数据使用了随机过采样技术,使其达到与堵车图片相近的数量.

所用数据集除分布不均衡以外,还存在样本总数偏少、相似度较高的问题. 针对此问题,笔者使用如下数据增强手段.

(1) 根据应用场景,对数据进行左右镜面对称,使数据集数量翻倍.

(2) 根据此数据集的实际情况,图片的边缘地带包含的信息较少,于是采用从原始图像中截取部分主要区域,增加数据量. 下式为可增加的数据量.

$$C = (m - i) \times (n - j), \quad (1)$$

式中: $C$  为数据可增加的倍数; $m, n$  为原始图像的尺寸; $i, j$  为从中截取的图像尺寸.

(3) 随机改变部分图片的亮度和对比度,提高模型处理来自不同光照条件图像的泛化性能<sup>[17]</sup>.

2.3 基于虚警及漏警的损失权重

对于拥堵与非拥堵的二分类问题,可根据真实类别与模型预测类别的组合划分为真正例(true positive)、假正例(false positive)、真反例(true negative)和假反例(false negative),如表 1 所示.

表 1 预测情况表  
Tab.1 Table of forecast result

真实情况	预测结果	
	正例	反例
正例	TP(真正例)	FN(假反例)
反例	FP(假正例)	TN(真反例)

根据实际情况,该模型应用面向智慧交通,其判断结果将直接影响到城市交通资源的调度,所以在考察模型时不仅仅要考虑模型的准确率(accuracy),漏警率(missing alarm)和虚警率(false alarm)也十分重要<sup>[18]</sup>.

$$A = \frac{TP + TN}{TP + FP + FN + TN}, \tag{2}$$

$$MA = \frac{FN}{TP + FN}, \tag{3}$$

$$FA = \frac{FP}{TP + FP}, \tag{4}$$

式中:A 为准确率;MA 为漏警率;FA 为虚警率.

在此模型中,漏警情况对应道路发生拥堵而模型判断为非拥堵,虚警情况对应道路未发生拥堵而模型判断为拥堵的情况.上文已经提到,该模型将直接影响到城市交通资源的调整情况,相对于漏警率,若虚警率较高将会造成大量的交通资源浪费.因此在模型的训练过程中,人为地调高虚警情况在损失计算中所占的权重,使模型在反向传播时更关注虚警情况,从而降低虚警率.

3 实验与对比

该实验硬件环境使用 Intel Xeon E5-2650 v4 CPU,NVIDIA 1080TI GPU,内存 128 GB.操作系统使用 Centos 6.5,框架采用 TensorFlow,并将数据生成为 TFrecord 格式,方便后续调用.实验数据来自互联网,其中拥堵图片 811 张,非拥堵照片 464 张,共 1 275 张.其中 80% 作为训练集,20% 作为测试集.经对训练集进行数据增强后,训练集数据 10 200 条,测试集数据 255 条.图 3 为该数据集中实例图像,上面三张为拥堵图像,下面三张为非拥堵图像.



图 3 数据集内部分图片  
Fig.3 Part of images in data set

3.1 对比常规神经网络

此实验将使用经过数据增强的数据与包含 3 个隐藏层的全连接网络、LeNet5 作为对比,结果如表 2 所示.

表 2 与常规神经网络对比结果

Tab.2 The result of comparison between the method in this paper and conventional methods

算法模型	准确率/%	参数数量/个	1 万次迭代耗时(纯 CPU 计算)/s
全连接网络	65.7	约 25 亿	3 773
LeNet5	83.2	约 1 300 万	624
笔者网络结构	94.6	约 2 亿	1 156

由表 2 可以看出,与全连接网络相比,卷积神经网络参数较少、计算量小、特征提取能力强,相对于全连接网络模型在图像识别领域有明显的优势.另外通过基于笔者网络结构与 LeNet5 对比,可以看出在避免梯度弥散的前提下,使用更加深层的网络结构,对提高准确率有明显作用.

3.2 对比数据增强前后

此实验分别使用原始训练数据集、过采样均衡分布训练数据集和经过了所有数据增强步骤的训练数据集,在图 2 网络结构模型上进行对比实验,实验结果正确率如图 4 所示.

从最终结果可以看出在经过数据平衡和数据增强后,准确率分别上升到 88.6% 和 94.3%.由此说明在小样本、非均衡数据集的深度学习问题中,数据增强手段对提高模型判断准确率和模型泛化性能有重要作用.

3.3 对比提高虚警损失权重前后

为降低最终模型判断的虚警率,此实验通过 TensorFlow 内部的自定义损失函数,人为地将出现虚警情况的损失值提高为原来的 5 倍,但不改变漏警情况的损失权重,实验结果如表 3 所示.

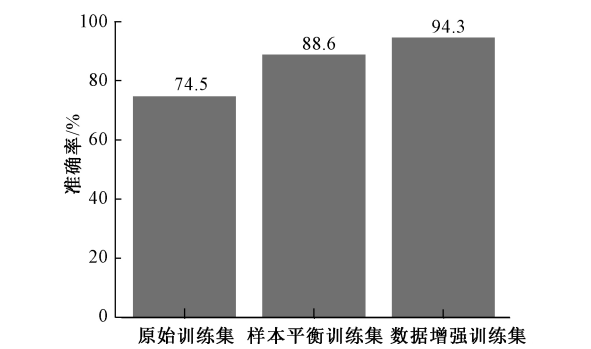


图 4 数据增强结果

Fig. 4 The result of data enhancement

表 3 提高权重后结果  
Tab. 3 The result of weight enhancement

项目	模型准确 率/%	模型漏警 率/%	模型虚警 率/%
正常训练	94. 30	3. 30	2. 80
提高虚警损失权重训练	92. 70	5. 90	1. 94

从表 3 可以看出,提高虚警损失权重后,在模型整体判断错误率略微上升的情况下,虚警率下降,整体虚警数量下降.说明提高虚警损失权重对降低虚警率有明显作用,可明显降低模型将通畅判别为拥堵的情况,从而降低系统后续响应动作的成本.

3. 4 对比传统拥堵检测方式

为凸显笔者方法相对于传统机器视觉检测方法的时间优势,采用基于均值法的拥堵检测模型<sup>[19]</sup>与笔者模型进行对比.其中,基于均值法的拥堵检测模型相对于其他传统方式在检测耗时相似的情况下,背景建模耗时更少.其工作流程为,先对同一地点拍摄到的多张图片直接求平均来得到此地点的近似背景图,再使用此地点的后续图片与背景图对比,根据差异性判断是否发生拥堵.

此实验采用 100 张同一地点的照片进行检测.这些图片首先用传统的方法建模,然后分别使用两种方法进行检测,检测时间对比如表 4 所示.

表 4 检测时间对比表 Tab. 4 The result of test time		
方法	背景计算 耗时/s	单次拥堵检测 平均耗时/s
传统检测方法	0. 856	0. 011
笔者检测方法	0	0. 012

由实验结果可以看出,在单次拥堵检测耗时几乎相同的情况下,传统的基于机器视觉的道路拥堵检测手段,在检测拥堵之前要先计算路面背景,耗时较长,且若发生光照、雨雪等环境因素改变,需重新计算背景,无法实时分析,计算量较大.而基于深度学习的拥堵检测方法,不受外界环境影响,检测速度快,实用性好,符合一些需要快速检测且计算资源较少的场景.

4 模型应用

目前笔者的道路拥堵检测方法已应用于城市道路可移动隔离桩系统原型机的开发,该系统由可移动隔离桩和管控系统组成.可移动隔离桩摆放在四车道以上道路的中央黄线上,每个隔离桩间隔 3 m 使用电缆连接,每 10 个隔离桩组成一个单元并配备摄像头模块.若系统连续检测到某个隔离桩单元一侧发生拥堵,且拥堵时间超过设定阈值时,便会认为道路发生拥堵.此时,若该隔离桩单元附近的另一侧车道未发生拥堵,管控系统便会控制拥堵段附近的隔离桩单元向非拥堵一侧移动,增加拥堵侧的车道数量来缓解拥堵,当拥堵结束后隔离桩会自动复原至道路中央.具体如图 5 所示.

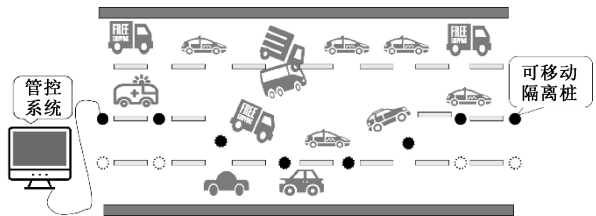


图 5 可移动隔离桩系统结构示意图  
Fig. 5 Model of mobile partition piles system on urban roads

5 结论

基于卷积神经网络设计实现了一种城市道路拥堵检测模型,并结合实际情况使用数据增强和修改虚警率权重等方式对模型实施了进一步优化.与传统基于机器视觉的城市道路拥堵检测方法相比,该方法不受环境因素的影响,可对城市道路实现实时检测,具有计算量小、泛化性好等特点.目前,测试准确率达到 92. 7%,虚警率仅为 1. 94%.但由于数据集来自互联网,样本较少且质量不高,使得模型总体正确性率并未达到理想水平.优化数据集,继续提高模型判断正确率,是下一步需要进行的工作.

## 参考文献:

- [1] 付熊. 基于机器视觉的交通拥堵及运动目标检测[D]. 广州: 华南理工大学自动化科学与工程学院, 2016.
- [2] 崔华, 纪宇, 李盼依. 基于卷积神经网络(CNN)的高速路交通图片拥堵识别[J]. 科技创新与应用, 2018, 224(5): 18-19.
- [3] 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, 224(4): 19-20+22.
- [4] SHERRINGTON C S. Observations on the scratch-reflex in the spinal dog[J]. Journal of physiology, 1906, 34(1/2): 1.
- [5] FUKUSHIMA K. neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position[J]. Biological cybernetics, 1980, 36(4): 193-202.
- [6] 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.
- [7] 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.
- [8] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C] // International Conference on Learning Representations (ICLR). San Diego: arXiv: 109. 1556v6 [cs. V], 2015.
- [9] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C] // IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2015: 1-9.
- [10] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C] // IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2016: 770-778.
- [11] 杨文柱, 刘晴, 王思乐, 等. 基于深度卷积神经网络的羽绒图像识别[J]. 郑州大学学报(工学版), 2018, 39(2): 11-17.
- [12] WANG D L. Deep learning reinvents the hearing aid[J]. IEEE spectrum, 2017, 54(3): 32-37.
- [13] LANE N D, GEORGIEY P, QENDRO L. Deep ear: robust smartphone audio sensing in unconstrained acoustic environments using deep learning[C] // UBI-COMP '15 ACM International Joint Conference on Pervasive and Ubiquitous Computing - September. ACM, 2015: 283-294.
- [14] GE M, LIU Y, JIANG L, et al. Genome-wide analysis of maize NLP transcription factor family revealed the roles in nitrogen response[J]. Plant growth regulation, 2018, 84(1): 95-105.
- [15] 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.
- [16] CHAWLA N V, BOWYER K W, HALL L O, et al. SMOTE: synthetic minority over-sampling technique[J]. Journal of artificial intelligence research, 2002, 16(1): 321-357.
- [17] 黄双萍, 孙超, 齐龙, 等. 基于深度卷积神经网络的水稻穗瘟病检测方法[J]. 农业工程学报, 2017, 33(20): 169-176.
- [18] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016.
- [19] 韩超, 邓甲昊, 邹金慧, 等. 基于差分均值背景提取和矩阵分区目标检测算法的研究[J]. 北京理工大学学报, 2012, 32(12): 1247-1251+1257.

## Traffic Jam Detection Based on Convolutional Neural Network

LUO Ronghui, YUAN Hang, ZHONG Fahai, NIE Shangshang

(School of Physics Engineering, Zhengzhou University, Zhengzhou 450001, China)

**Abstract:** In order to solve the increasingly serious congestion problem of urban roads, a road congestion detection method based on convolutional neural network was proposed in this paper. This method was based on the application value of road congestion rapid detection technology to alleviate the traffic jam problem. Besides, it combined deep learning with image processing technology. Compared with traditional methods, this method did not need to extract the road background in the early stage, and was not affected by the illumination brightness and the actual environment. It had the characteristics of fast recognition speed, less occupied computing resources and good generalization. It has been applied in practical projects and achieved good results.

**Key words:** convolutional neural network; deep learning; image recognition; congestion detection; smart city