

doi:10.11835/j.issn.1000-582X.2020.298

基于深度学习的交通拥堵检测

丁 洁, 刘晋峰, 杨祖葭, 阎高伟

(太原理工大学 电气与动力工程学院, 太原 030024)

摘要:针对交通拥堵检测,提出一种利用深度学习中卷积神经网络(convolutional neural network,CNN)对交通拥堵图像进行检测的方法。首先,使用基于 TensorFlow 框架设计了含有三层卷积层的神经网络分类模型。其次,利用道路拥堵与非拥堵图片对分类模型进行训练,并进行评估。最后,利用训练完成的模型进行道路拥堵检测,与多种深度学习分类模型分类方法进行对比实验,表明该卷积神经网络模型能够更有效地进行拥堵检测,检测准确率达到 98.1%。

关键词:拥堵检测;卷积神经网络;深度学习;TensorFlow

中图分类号:TP751

文献标志码:A

文章编号:1000-582X(2021)04-107-10

Traffic congestion detection based on deep learning

DING Jie, LIU Jinfeng, YANG Zuliang, YAN Gaowei

(College of Electrical and Power Engineering, Taiyuan University of Technology,
Taiyuan 030024, P. R. China)

Abstract: Aimed at traffic congestion detection, a method of detecting traffic congestion images using convolutional neural network (CNN) was proposed. First, a neural network classification model with three layers of convolutional layers was designed based on the TensorFlow framework. Then, the classification model was trained and evaluated using road congestion and non-congestion pictures. Finally, the well-trained model was used to carry out road congestion detection. Compared with many other deep learning classification models, the proposed convolutional neural network model showed high efficiency in congestion detection, and the detection accuracy reached 98.1%.

Keywords: traffic congestion; CNN; deep learning; TensorFlow

随着国民经济的快速发展,近年来机动车保有量大幅度增加,造成了严重的交通拥堵^[1]。有关研究表明,拥堵发生时低速行驶的车辆会排放更多废气。当车速为 5 km/h 以下时,尾气氮氧化物排放系数为 7;而当车速为 35 km/h 以上时,尾气氮氧化物排放系数为 3^[2]。

对于交通拥堵的检测,目前主要采用的是通过 GPS 定位检测、微波检测、红外检测和线圈检测等物理检测手段来获取交通流参数^[3],通过如车辆数量、车辆平均行驶速度等参数进行交通拥堵判断。然而,这些传统方法均存在造价高、对路面破坏较大等问题。

收稿日期:2019-12-13 网络出版日期:2020-03-17

基金项目:国家自然科学基金资助项目(61604105)。

Supported by National Natural Science Foundation of China(61604105).

作者简介:丁洁(1986—),女,太原理工大学副教授,主要从事图像处理、新型半导体器件建模研究,(E-mail)dingjie2015@foxmail.com。

随着计算机视觉的发展及公路上全路段视频监控的实现,为基于视频图像的拥堵检测提供可能^[4]。目前基于视频图像的交通拥堵检测方法大部分是对车辆数量进行检测。例如对感兴趣区域进行处理^[5],提取车辆目标以统计车辆数,再根据车辆数进行交通拥堵的判断。再如,利用背景差分法和帧差法分析图像的纹理变化情况,从而得出道路空间占有率和时间占有率以判断道路拥堵的情况^[6]。以上拥堵检测方法易受环境变化的影响,如光照强度及路面反光等,无法应用于不同的场景中。

近年来,卷积神经网络(convolutional neural network,CNN)在图像处理领域的应用逐渐增多,其有着自动提取图像特征、鲁棒性很强等优点^[7]。Lecun 等^[8]针对手写字符识别,首次提出了 LeNet-5 图像识别模型,获得了极佳的效果。Alex 通过加深网络层数、使用 CUDA 进行训练等方式获得了效果更好的 AlexNet^[9]。而后,谷歌提出的 GoogLeNet 与牛津大学提出的 VGG(visual geometry group)模型以较小的卷积核尺寸增强了网络的实用性及简洁性,分别获得 2014 年 ILSVRC 竞赛的第一名、第二名,但是 VGG 在多个迁移学习任务中的表现优于 GoogLeNet^[10]。之后基于 CNN 的分类模型有了长足发展,He 等^[11]在 2016 年提出了用于图像识别的深度残差模型 ResNet,通过使用残差结构优化了深层 CNN 模型。Howard 等^[12]于 2017 年提出了轻量级模型 MobileNet,其通过深度可分离卷积优化了模型结构,使检测速度大大提高。目前,已有研究人员将经典卷积神经网络结构应用于道路拥堵识别中,如罗荣辉等^[13]通过改进 AlexNet 模型的网络结构来检测道路拥堵情况,检测速度达到了 0.012 s/帧,但是其存在模型臃肿、检测速度较慢的缺点。因此,卷积神经网络在道路拥堵识别中的应用仍有待发展。

基于以上研究,笔者提出一种利用深度学习卷积神经网络的方法进行道路拥堵判断,实现对道路拥堵的准确检测,并且检测速度得到改进。在文章第一部分设计了所使用的卷积神经网络模型结构,并对模型进行训练。在第二部分进行了对比实验,利用多种卷积神经网络模型的结果进行比较,并对网络结构进行了消融实验。第三部分对实验结果进行了讨论。

1 设计 CNN 分类模型

本文中所设计的基于 CNN 的分类模型主要由 3 个模块组成,如图 1 所示分别为:图片预处理模块、卷积神经网络模块和结果优化模块。下面将对模块内容进行详细介绍。

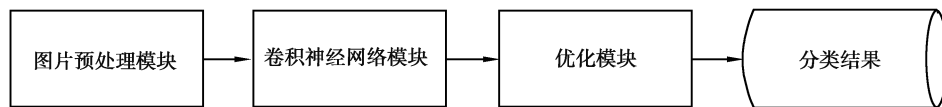


图 1 模块组成

Fig. 1 Module composition

1.1 图片预处理模块

图片预处理模块主要进行的步骤为:

- 1) 对图片数据进行增强,包括了随机翻转、变换亮度等,以丰富数据集;
- 2) 对图像统一缩放为 64×64 大小,以便于输入至卷积层中。

其中,训练集中某图片操作结果如图 2 所示。



图 2 图片预处理效果图

Fig. 2 The preprocessing effect of a picture

1.2 卷积神经网络模块

本文中所设计的三层卷积神经网络结构如图 3 所示。

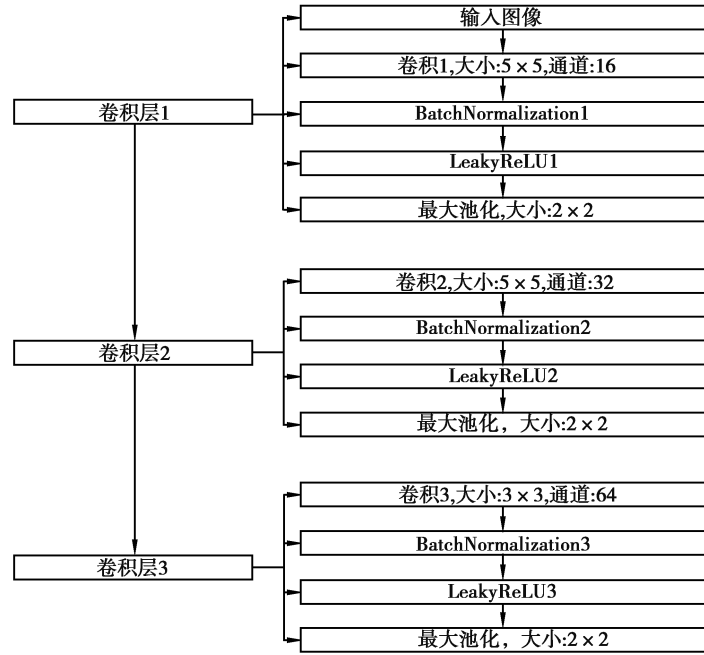


图 3 CNN 模型结构

Fig. 3 CNN model structure

在图 3 的结构中,选用了 5×5 与 3×3 的小卷积核进行特征提取,减少了参数数量,加快了训练与识别的速度。在每一层中使用 Leaky-ReLU 激活函数与 Batch-Normalization。图 4 为 ReLU 与 Leaky-ReLU 函数图像,相较于普通使用的 ReLU 激活函数,Leaky-ReLU 更全面地保留了特征信息。

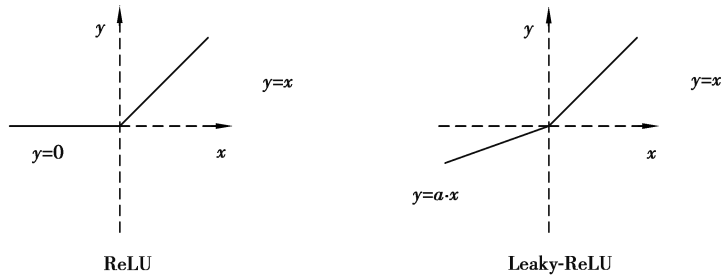


图 4 ReLU 与 Leaky-ReLU 的函数图像

Fig. 4 Function image of ReLU and Leaky-ReLU

式(1)为 Leaky-ReLU 函数公式, a_i 是 $(1, +\infty)$ 区间内的固定参数。Leaky-ReLU 给所有负值赋予一个非零斜率,保留了一定的特征效果。而结构图 3 中的 Batch-Normalization 层通过规范化卷积的输出,使得激活函数的输入值能进入对输入比较敏感的区域,避免了梯度消失问题^[14]。

$$y_i = \begin{cases} x_i & \text{if } x_i \geq 0, \\ \frac{x_i}{a_i} & \text{if } x_i < 0. \end{cases} \quad (1)$$

1.3 优化模块

本文中所设计的优化模块网络结构如图 5 所示。

在图 5 的结构中,Flatten 层可将卷积层 3 的输出“压平”,把多维输入一维化,实现卷积层到全连接层的过渡;Dense 层则为全连接层,对卷积的结果进行特征的进一步提取。其中,第一个 Dense 层含有 128 个神经

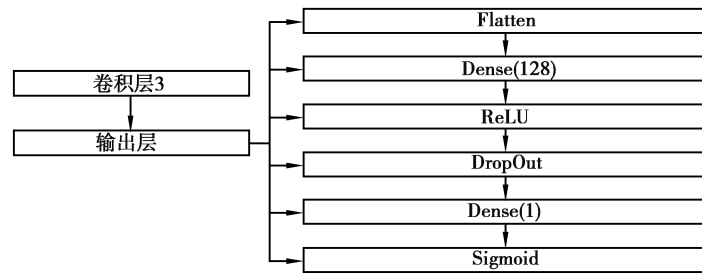


图 5 优化模块结构

Fig. 5 Optimization module structure

元,可将卷积层 3 中卷积层提取的局部特征进行全连接,实现特征的融合;由于本文中针对的交通拥堵分类是二分类问题,因此第二个 Dense 层只含有一个神经元,作用为融合所有特征,便于之后进行类别概率计算。损失函数为二分类交叉熵损失 Sigmoid_cross_entropy,其公式为:

$$S(t) = \frac{1}{1 + e^{-t}} \quad (2)$$

式中: t 为 Dense 层的输出, $S(t)$ 为输出类别概率,这样就将卷积层的输出变换为二分类概率。

优化模块主要采用了 DropOut 算法以减轻过拟合的产生,并提高模型的泛化型。DropOut 算法示意图如图 6 所示。

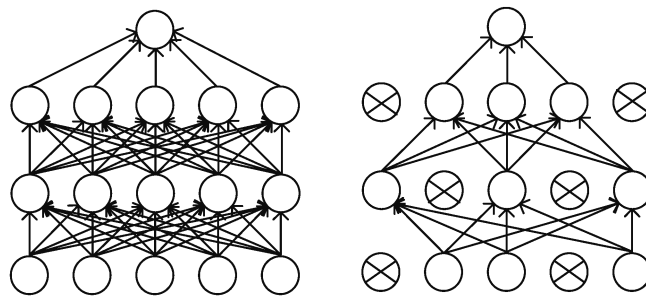


图 6 DropOut 在神经元中的作用

Fig. 6 The role of DropOut in neurons

如图 6 所示,左图为未使用 DropOut 时正常的神经网络传递方式,其传递公式为:

$$y_i^{(l+1)} = f(w_i^{(l+1)} y^{(l)} + b_i^{(l+1)}) \quad (3)$$

式中: $y_i^{(l+1)}$ 为第 $l+1$ 层的输出, $y^{(l)}$ 为第 l 层的输出, $w_i^{(l+1)}$ 为第 $l+1$ 层的系数矩阵, $b_i^{(l+1)}$ 为第 $l+1$ 层的偏置, f 为激活函数,下同。

图 6 右图为随机放弃 40% 神经元后的传递方式,其传递公式为^[15]:

$$r_j^{(l)} = \text{Bernouli}(p), \quad (4)$$

$$\tilde{y}^{(l)} = r^{(l)} * y^{(l)}, \quad (5)$$

$$y_i^{(l+1)} = f(w_i^{(l+1)} \tilde{y}^{(l)} + b_i^{(l+1)}) \quad (6)$$

式中的 Bernouli 函数会生成随机向量 0 或者 1,起到 DropOut 的作用。

DropOut 可以解决过拟合的原因在于^[16]:

1) 训练时可产生多种网络:在训练过程中,如果使用相同的数据去训练不同的网络,会得到不同的结果,那么将这些结果取平均后,可以有效避免过拟合。DropOut 可以在每次训练时随机放弃一些神经元,模拟产生多种不同的网络结构;

2) 减少某些神经元的敏感性:神经网络应该在学习时更具有鲁棒性,DropOut 能让一些神经元不是每次都在网络中出现,这样就有效减轻了网络对特定信息的依赖。

1.4 模型训练

在对模型进行训练时,使用 9 457 张图片作为数据集,其中,5 000 张图片为训练集,2 320 张图片为验证集,2 137 张图片为测试集。所用图片采集于山西省太原市街道内,数据集中包括了不同角度、不同光照条件下的场景图片,如图 7 所示,包含了不同角度、在白天夜晚条件下的场景。



图 7 不同角度、不同光照条件下的场景

Fig. 7 Scenes under different angles and different lighting conditions

在训练神经网络模型时,学习速率为 0.001,每次迭代选择 32 张训练集图片进行训练,训练轮次 epochs=1 000,图 8 为训练过程中的损失与准确率指标的展示。

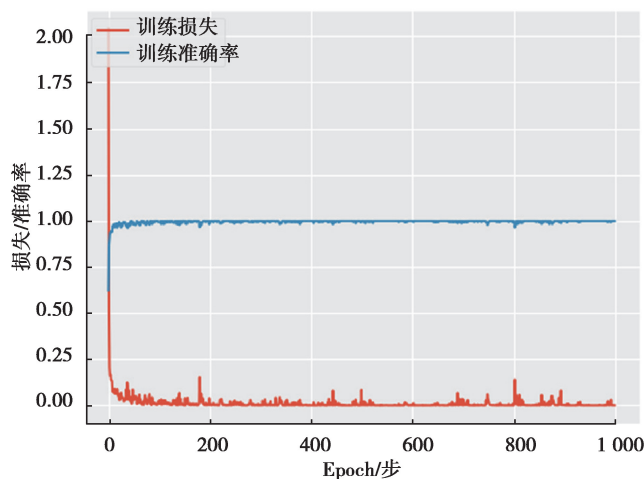


图 8 训练过程

Fig. 8 Training process

图 8 结果表明,经过 1 000 轮训练后,本文模型所达到的验证准确率为 99%。

1.5 消融实验

为了验证优化模块的有效性,对模型结构进行了消融实验,如表 1 所示,验证模型 1 不使用 DropOut 方法,其余结构不变;验证模型 2 不使用 Leaky-ReLU 激活函数而使用普通的 ReLU 激活函数,其余结构不变。

表 1 消融实验

Table 1 Ablation experiment

	本文模型	验证模型 1	验证模型 2
LeakyReLU	✓	✓	
DropOut	✓		✓

对验证模型 1、验证模型 2 在同样的训练集下,以学习速率为 0.001,每次迭代选择 32 张训练集图片进行训练,训练轮次 epochs=1 000,图 9、图 10 为验证模型的训练结果展示。

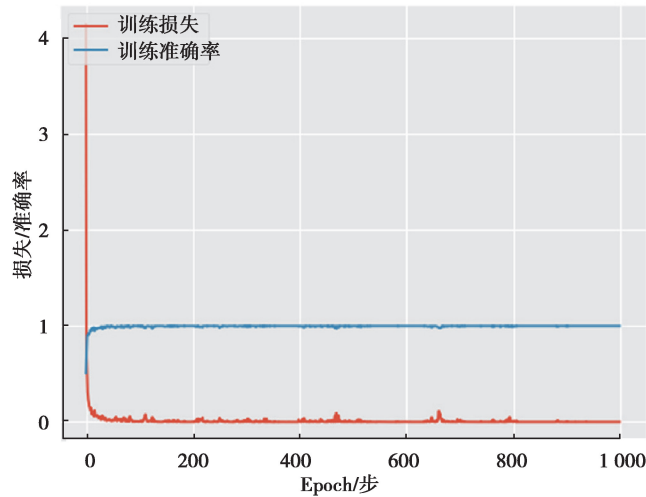


图 9 验证模型 1 训练过程

Fig. 9 Training process of model 1

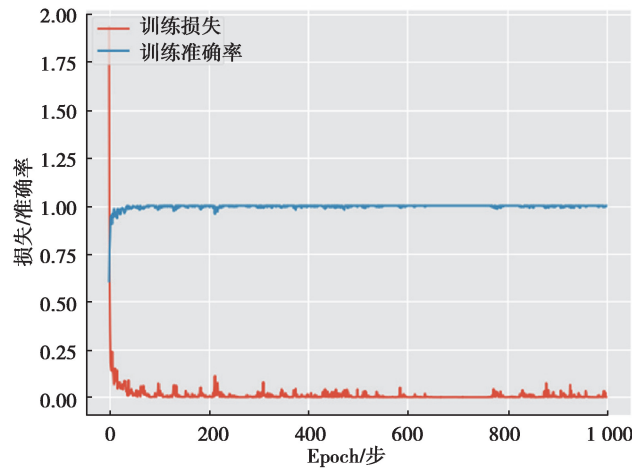


图 10 验证模型 2 训练过程

Fig. 10 Training process of model 2

图 9、图 10 表明,验证模型 1、验证模型 2 在训练后均已收敛,且训练集的验证准确率达 99% 以上。

2 实验与评估

本文的实验环境为 Intel Core i7-8750H CPU(2.20 GHz)、Nvidia GeForce GTX 1050TI GPU(4G 显存)和 16 GB RAM。对测试集中 1 123 张交通拥堵图片、1 014 张交通通畅图片共计 2 137 张图片进行分类后,

所得结果如表 2 所示。

表 2 本文模型分类结果

Table 2 Classification results of our model

	分类正确	分类错误	准确率/%
拥堵	1 103	20	98.2
通畅	994	20	98.0
总计	2 097	40	98.1

由表 2 中测试集图片的分类结果可得本文使用的模型分类准确率为 98.1%。经测试,分类模型处理单张图片的速度为 0.004 s。

使用验证模型 1 对测试集图片进行分类的结果如表 3 所示。

表 3 验证模型 1 分类结果

Table 3 Model 1 classification results

	分类正确	分类错误	准确率/%
拥堵	1 028	95	91.5
通畅	908	106	89.5
总计	1 936	201	90.6

由表 3 中测试集图片的分类结果可得验证模型 1 的分类准确率为 90.6%,效果明显劣于使用了 DropOut 的本文模型。归结其原因在于,对于使用了 DropOut 的网络结构,其网络在训练时,相当于进行了训练数据的扩充。这是由于总可以在训练集中找到一个样本,使得该样本在原始网络上的输出结果也能达到 DropOut 后的效果,所以每一次 DropOut 都相当于增加了样本,使得网络能够多尺度、更丰富地使用特征,分类效果变得更好。

使用验证模型 2 对测试集图片进行分类的结果如表 4 所示。

表 4 验证模型 2 分类结果

Table 4 Model 2 classification results

	分类正确	分类错误	准确率/%
拥堵	1 101	22	98.0
通畅	974	40	96.1
总计	2 075	62	97.1

由表 4 中测试集图片的分类结果可得验证模型 2 的分类准确率为 97.1%,效果略差于使用了 Leaky-ReLU 的本文模型。主要原因在于 Leaky-ReLU 相较于普通的 ReLU 激活函数,通过给所有负值赋予一个非零斜率保留了更多的特征,因此分类效果更佳。

为了进一步表明所使用的神经网络模型结构的优势,使用深度学习中经典的 VGG16、ResNetv1、InceptionV3 和 MobileNet 模型进行对比验证。VGG16、ResNetv1 和 InceptionV3 均是广泛应用的卷积神经网络分类模型,其以网络结构规整、分类性能良好应用在各个领域。而 MobileNet 则是轻量化模型,更注重处理速度。使用训练集对 VGG16、ResNetv1、InceptionV3 和 MobileNet 模型进行训练,学习速率为

0.001, 每次迭代选择 32 张训练集图片进行训练。由于这些模型结构比较庞大, 为了保证网络的收敛, 增加了训练轮次, 训练轮次 $\text{epochs} = 100\ 000$, 并且在同一个测试集上进行测试, 其对交通拥堵的分类结果如表 5~8 所示。

表 5 VGG16 分类结果

Table 5 VGG16 classification results

	分类正确	分类错误	准确率/%
拥堵	1 039	84	92.5
通畅	999	15	98.5
总计	2 038	99	95.4

由表 5 中测试集图片的分类结果可得利用 VGG16 模型的分类准确率为 95.4%。经测试, 使用 VGG16 模型处理单张图片的速度为 0.027 s。

表 6 ResNetv1 分类结果

Table 6 ResNetv1 classification results

	分类正确	分类错误	准确率/%
拥堵	1 031	92	91.8
通畅	988	26	97.4
总计	2 019	118	94.5

由表 6 中测试集图片的分类结果可得利用 ResNetv1 模型的分类准确率为 94.5%。经测试, 使用 ResNetv1 模型处理单张图片的速度为 0.017 s。

表 7 InceptionV3 分类结果

Table 7 InceptionV3 classification results

	分类正确	分类错误	准确率/%
拥堵	1 085	38	96.6
通畅	984	30	97.0
总计	2 069	68	96.8

由表 7 中测试集图片的分类结果可得利用 InceptionV3 模型的分类准确率为 96.8%。经测试, 使用 InceptionV3 模型处理单张图片的速度为 0.025 s。

表 8 MobileNet 分类结果

Table 8 MobileNet classification results

	分类正确	分类错误	准确率/%
拥堵	1 015	108	90.4
通畅	951	63	93.8
总计	1 966	171	92.0

由表8中测试集图片的分类结果可得利用 MobileNet 模型的分类准确率为 92%。经测试,使用 MobileNet 模型处理单张图片的速度为 0.005 s。

3 结果对比

将本文中所使用的模型称为 Our-net。对比表 2~8 的结果,并计算 FPS(frames per second)后如表 9 所示。

表 9 分类结果对比

Table 9 Comparison of classification results

方法名称	准确率/%	FPS
Our-net	98.1	250
Our-net(without Dropout)	90.6	250
Our-net(without Leaky-ReLU)	97.1	250
VGG16	95.4	37
ResNetv1	94.5	59
InceptionV3	96.8	40
MobileNet	92.0	200

由于本文所设计的模型卷积网络层数较少,因此参数数量也较少。能在保持精确分类的条件下,分类速度大大优于 VGG16、ResNetv1 和 InceptionV3 模型,FPS 达到了 250,比其他模型的最佳结果仍高 50;而在维持分类速度的条件下,分类精度大大优于 MobileNet 等模型,模型准确率达到 98.1%,比其他模型的最佳结果仍高 1.3%。

4 结 语

针对交通拥堵问题,应用卷积神经网络进行道路图片分类。得到的结果表明,本文中设计的 CNN 模型的检测效果大大优于利用 VGG16、ResNetv1、InceptionV3 和 MobileNet 模型分类方法,且检测速度有较大的提高,能够有效地对交通拥堵进行判断分类,最终交通拥堵检测精度达到了 98.1%,检测速度达到 250 FPS,为交通管理者进行准确道路管理提供了一个新的方案。

参考文献:

- [1] 郭继孚,刘莹,余柳.对中国大城市交通拥堵问题的认识[J].城市交通,2011,9(2):8-14.
Guo J F, Liu Y, Yu L. Traffic congestion in large metropolitan area in China[J]. Urban Transport of China, 2011, 9(2): 8-14. (in Chinese)
- [2] 邓顺熙,陈爱侠,曹申存.高速公路汽车污染物排放因子的测试与研究[J].中国公路学报,1999(S1):96-102.
Deng S X, Chen A X, Cao S C. An experimental study of emission factors of vehicular pollutants in a freeway tunnel[J]. China Journal of Highway and Transport, 1999(S1):96-102. (in Chinese)
- [3] Sikirić I, Brkić K, Bevandić P, et al. Traffic scene classification on a representation budget[J]. IEEE transactions on intelligent transportation systems, 2019, 21(1):336-345.
- [4] 黄磊磊,汤一平,孟焱.基于机器视觉的道路拥堵状态检测的研究[J].小型微型计算机系统,2014,35(1):148-153.
Huang L L, Tang Y P, Meng Y. Research on road congestion state detection based on machine vision[J]. Journal of Chinese Computer Systems, 2014,35(1):148-153. (in Chinese)
- [5] Desa S M, Salih Q A. Image subtraction for real time moving object extraction[C]//International Conference on Computer Graphics, Imaging & Visualization. New York, NY, USA:IEEE, 2004:41-45.

- [6] 安纪存, 吕鑫, 季琳雅. 不完全数据下基于时空相关性拥堵预测方法[J]. 计算机工程与应用, 2019, 55(4): 96-100,124.
An J C, Lv X, Ji L Y. Traffic congestion prediction based on spatial-temporal correlation with incomplete data[J].
Computer Engineering and Applications, 2019,55(4):96-100,124. (in Chinese)
- [7] Guillaume R, Soullard Y, Chatelain C, et al. Fully convolutional network with dilated convolutions for handwritten text line segmentation[J]. International Journal on Document Analysis and Recognition, 2018, 3(21):177-186.
- [8] 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.
- [9] Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2012, 25(2): 1097-1105.
- [10] Schlichtkrull M, Kipf T N, Bloem P, et al. Modeling relational data with graph convolutional networks[J/OL]. arXiv, 2017(2017-10-26)[2020-01-25]. <https://arxiv.org/pdf/1703.06103.pdf>.
- [11] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), Las Vegas, NV, USA, June 27-30,2016. IEEE, 2016: 770-778.
- [12] Howard A G, Zhu M, Chen B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[J/OL]. Computer Vision and Pattern Recognition,2017 [2020-01-25]. <https://arxiv.org/abs/1704.04861>.
- [13] 罗荣辉,袁航,钟发海,等.基于卷积神经网络的道路拥堵识别研究[J].郑州大学学报(工学版), 2019,40(2):18-22.
Luo R H, Yuang H, Zhong F H, et al. Traffic jam detection based on convolutional neural network[J]. Journal of Zhengzhou University (Engineering Science), 2019, 40(2):18-22. (in Chinese)
- [14] Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks[C]// NIPS'12: Proceedings of the 25th International Conference on Neural Information Processing Systems. NY, United States: Curran Associates Inc, 2012,1:1097-1105.
- [15] Li Z, Jiang L Y, Zhu Z Q, et al. Vehicle classification for large-scale traffic surveillance videos using Convolutional Neural Networks[J]. Machine Vision and Applications,2017,28(7): 793-802.
- [16] Liu X, Liu W, Ma H, et al. Large-scale vehicle re-identification in urban surveillance videos[C]// IEEE International Conference on Multimedia and Expo (ICME) 2016. New York, USA: IEEE, 2016:1-6.

(编辑 郑洁)