

文章编号: 2095-2163(2019)05-0075-05

中图分类号: TP391

文献标志码: A

基于改进 LeNet-5 的人脸识别研究

周官皓, 孙海洋

(南京大学金陵学院, 南京 210089)

摘要: 近些年来, 深度学习逐渐走入大众视野, 并在譬如语音识别、人脸检测等方面发挥着举足轻重的作用。本文利用深度学习的优势, 根据学术界的发展, 使用 Tensorflow 在 LeNet-5 的基础上, 创新地将 2 个卷积层的卷积核数目增长到 15 与 20 个, 并且增加了批归一化与 Dropout 两种模型处理手段, 减少过拟合的出现, 提升模型泛化程度。实验结果从测试准确率来分析, 结果表明增加了批归一化与 Dropout 的效果最好, 单一地增加二者之一对精度损失微乎其微; 验证了改进后的 LeNet-5 模型对校园内单一年龄段的人脸识别有着良好的效果。

关键词: 卷积神经网络; 深度学习; 人脸识别

Face recognition research based on improved LeNet-5

ZHOU Guan hao, SUN Hai yang

(Nanjing University Jinling College, Nanjing 210089, China)

[Abstract] In recent years, deep learning has entered the public's vision and played a pivotal role in fields like speech recognition and face detection, etc. This paper takes advantage of deep learning, using Tensorflow, on the basis of the LeNet-5 model, according to the development of academia and innovatively increases the number of convolution kernels of two convolutional layers to 15 and 20. Moreover, the two methods of batch normalization and Dropout are added to reduce the occurrence of over-fitting and improve the generalization of the model. The experimental results are analyzed from the test accuracy. The results show that the addition of batch normalization and Dropout is the best. Simply increasing one of them has little precision loss. The results verify that the improved LeNet-5 model has a good effect on face recognition in a single age group on campus.

[Key words] CNN; deep learning; face recognition

0 引言

自 2006 年卷积神经网络模型架构^[1]的提出与 2012 年卷积神经网络在 ImageNet 大赛上凭其出色表现而获学界赞许后, 深度学习^[2]便受到了全世界机器学习方向学者的密切关注。2016 年 AlphaGo 击败人类围棋大师李世石之后, 深度学习再次引发热潮, 直到今天, 深度学习已逐渐发展演变为一个相对独立的方向。深度学习隶属于人工神经网络体系, 相对于传统的人工神经网络, 深度学习更类似人脑, 其网络结构与分析判别具体事物的流程也更接近于人类分析的过程。目前, 深度学习已然衍生出运用于不同方向的网络架构, 本文主要分析并使用 CNN^[3]。典型的卷积神经网络架构有 AlexNet^[4]、ZFNet^[5]、VGGNet^[6]、GoogLeNet^[7]和 ResNet^[8]等。卷积神经网络通过结合局部感受野、池化和下采样

等方法来优化网络结构, 一定程度上保证位移不变性, 对于无约束条件下的特征提取, 其算法也有较强的鲁棒性。卷积神经网络拥有的权值共享特点这一方面, 也类似于深度学习所有的类似于生物神经网络的重要特点。

1 神经网络结构

本文模型搭建的基础 LeNet-5^[9]网络是一个用于手写数字识别的 7 层 CNN 网络, 其模型设计如图 1 所示。虽然结构与此后的其它网络相比规模较小, 但是包含了深度学习网络架构的基本模块。

LeNet-5 每层都包含可训练参数, 且每层都有多个特征图, 每个特征图通过一种卷积核提取输入的一种特征, 每个特征图有多个神经元。下面即以 LeNet-5 为示例来阐释说明卷积神经网络的层次结构。

基金项目: 南京大学金陵学院重点课题(0010521806); 教育部产学研合作协同育人项目(201802003026, 201802026004)。

作者简介: 周官皓(1999-), 男, 本科生, 主要研究方向: 大数据与人工智能; 孙海洋(1982-), 男, 硕士, 副教授, 主要研究方向: 大数据与人工智能。

通讯作者: 孙海洋 Email: sunhy_jlxy@163.com

收稿日期: 2019-07-25

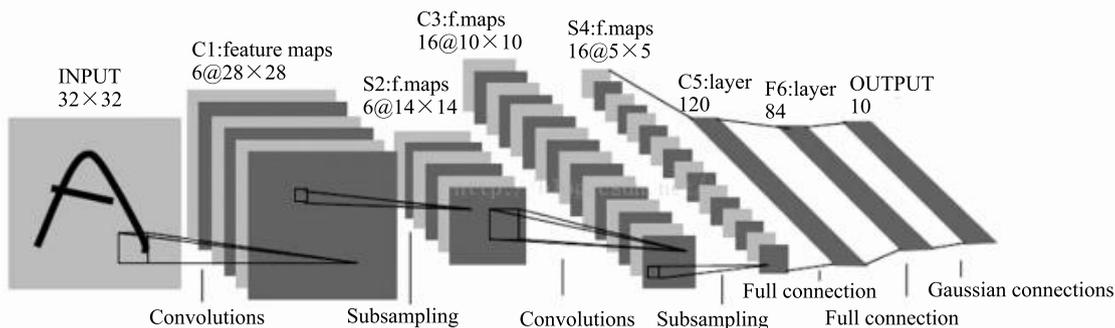


图1 LeNet-5 网络模型

Fig. 1 LeNet-5 network mode

1.1 卷积层

卷积神经网络中,每一层卷积层都有一个或多个卷积核,这些卷积核识别图像中某些特定特征,每个卷积核将会去滑动卷积上一层的特征图。在此过程中,卷积核的参数不变且共享,这样训练过程中只需要很少的参数就可以完成训练,且效果相对传统神经网络来说更加优异。卷积神经网络使用 BP (Back Propagation) 算法^[10] 逆向传递残差,对算法设计流程可得重点论述如下。

(1) 将输入示例提供给输入层神经元,逐层将信号前传,直到产生输出层的结果。

(2) 计算输出层的误差,再将误差逆向传播至隐层神经元,计算隐层神经元的误差,数学公式为:

$$\delta_k \leftarrow o_k(1 - o_k)(t_k - o_k), \quad (1)$$

$$\delta_h \leftarrow o_h(1 - o_h) \sum_{k \in \text{outputs}} w_{kh} \delta_k, \quad (2)$$

其中, δ_k 是输出层的误差; o_k 是第 k 个神经元的输出; δ_h 是隐层的误差。

(3) 根据隐层神经元的误差来调整连接权和阈值,数学公式为:

$$w_{ji} \leftarrow w_{ji} + \Delta w_{ji} = w_{ji} + \eta \delta_j x_{ji}. \quad (3)$$

其中, w_{ji} 是第 j 层到第 i 层之间更新后的权值。

1.2 池化层

池化层对上一层卷积层进行拟合,使上一层卷积层输入特征图大小在经过该池化层变换后减小,从而减少特征数量及参数数量。对局部区域取平均值或者最大值,得到可表示此区域的显著特征,保证参数减少的同时特征图不变。常用的有最大池化 (Max Pooling), 还有全局平均池化^[11] (Global Average Pooling), 作为层变换特征的表示。

1.3 全连接层

全连接层的目的是为了加强网络的非线性映射能力,同时将网络大小固定在某个范围内。该层神经元与前一层的每个神经元相互连接,同层神经元

之间互不相连,该数学公式可表现为:

$$o_j^l = f\left(\sum_{i=1}^n x_i^{(l-1)} \cdot w_{ji}^{(l)} + b_j^{(l)}\right). \quad (4)$$

其中, n 是上层神经元个数; l 是当前层; $w_{ji}^{(l)}$ 是层与层间的权值; $b_j^{(l)}$ 是当前层神经元 j 的偏置; $f(\cdot)$ 是激励函数。

1.4 Softmax 层

LeNet-5 最后一层是 RBF 分类器,每个 RBF 输出单元计算输入特征和参数矢量之间的欧几里德距离。但是人脸特征较数字而言更复杂,所以采用非线性分类能力较强的 Softmax 函数作为分类器。Softmax 层的最终结果拟设置为乘性,将多个神经元的输出,映射到 $(0, 1)$ 范围内,这就可将其视作概率,在此基础上再选取概率最大的神经元输出的结果作为最终的结果。过程中将用到的数学公式为:

$$h_\theta(x(i)) = \begin{bmatrix} p(y(i) = 1 | x(i); \theta) \\ \vdots \\ p(y(i) = k | x(i); \theta) \end{bmatrix} = \frac{\begin{bmatrix} \exp(\theta_1^T x(i)) \\ \vdots \\ \exp(\theta_k^T x(i)) \end{bmatrix}}{\sum_{j=1}^k \exp(\theta_j^T x(i))}, \quad (5)$$

其中, $p(y^{(i)} = j | x(i); \theta)$ 代表样本 $x^{(i)}$ 属于第 j 类的概率; θ_j^T 代表模型参量。其代价函数可定义为:

$$J(\theta) = -\frac{1}{m} \left[\sum_{i=1}^m \sum_{j=1}^k 1\{y^{(i)} = j\} \log \frac{e^{\theta_j^T x^{(i)}}}{\sum_{i=1}^k \exp(\theta_i^T x^{(i)})} \right]. \quad (6)$$

2 数据采样及预处理

研究可知,人脸识别中,至关重要的一环就是数据采样和预处理。为此,研究收集了来自地区各个

高校、共计 100 位同学的各 10 张照片,总计 1 000 张。考虑到拍摄背景、人占比与摄像机器等因素之间的差异,文中对于数据的采样与预处理,将做研究分述如下。

- (1) 使这 100 位同学拍摄时的实时背景为白色系。
- (2) 使用相同品牌型号的手机进行摄像,拍摄分辨率为 4 K,无多余设置。
- (3) 统一截取面部。多余的拍摄到的身体部位或背景丢弃,上下只保留自脖颈至头顶,左右从左耳至右耳轮廓最外侧。
- (4) 压缩图片像素大小为 100×100。
- (5) 对每张初步处理过的图片做像素归一化,使像素范围在[0,255]之间。

由此,即可得到研究数据示例如图 2 所示。

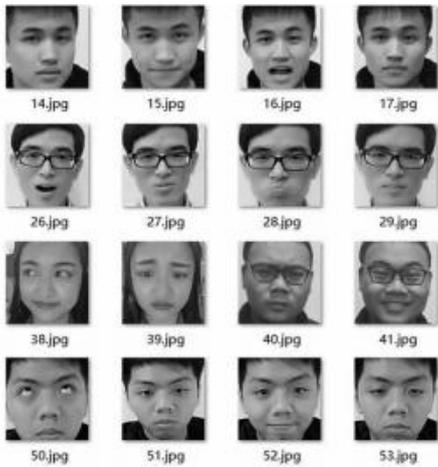


图 2 数据集示例
Fig. 2 Data set example

3 实验

3.1 LeNet-5 预测表现

本文模型的构造基于 LeNet-5 网络,LeNet-5 网络在本数据集上的表现如图 3 所示。

曲线图及最终结果显示,LeNet-5 模型在本数据集上的预测准确率为 0.841。初步估计准确率较低的原因则在于 LeNet-5 的参数设置是建立在

MNIST 数据集样本之上,而手写数字数据无论从图片像素、还是特征复杂度来说,都远远低于人脸数据。

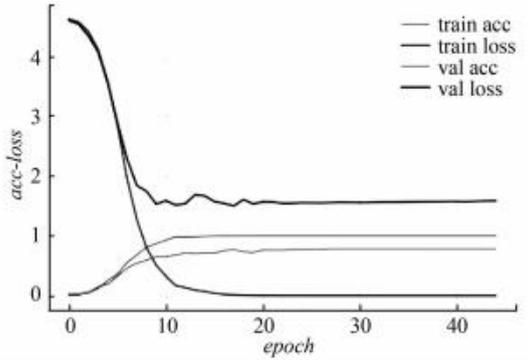


图 3 LeNet-5 预测表现

Fig. 3 LeNet-5 forecast performance

3.2 模型构建

3.2.1 流程设计

基于前述分析成果,研究可得模型设计流程如图 4 所示。

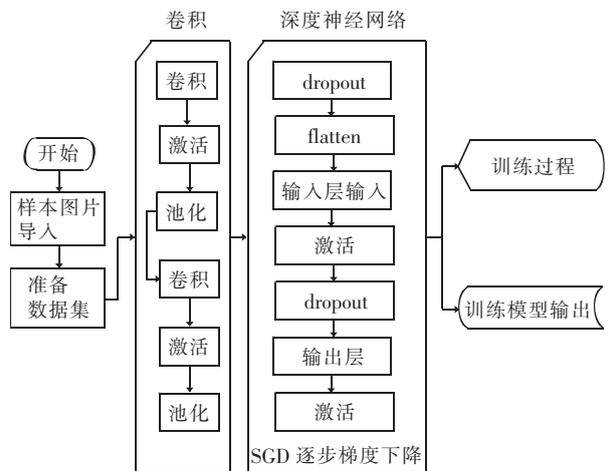


图 4 模型流程图

Fig. 4 Model flow chart

3.2.2 搭建卷积网络

卷积网络构造如图 5 所示。本文采用的卷积神经网络共使用了 2 层批归一化层^[12] (Batch - Normalization)、2 层卷积层、2 层池化层、1 层全连接层、2 层 Dropout 层^[13]以及 1 层 Softmax 输出层。

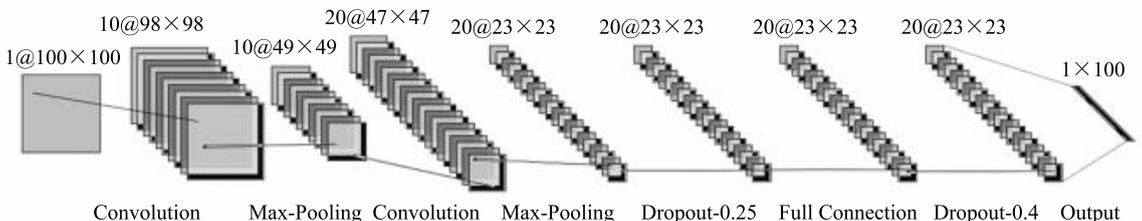


图 5 模型结构图
Fig. 5 Model structure

3.2.3 批归一化

对于人脸数据集来说,使用 1 000 张图片来训练仍显不足,因此本次研究选择在每一次卷积之前,即使用批归一化来做数据处理。批归一化层的公式可顺次表示如下。

Input: Values of x over a minibatch: $B = \{x_i, \dots, x_m\}$;

Parameters to be learned: γ, β

Output: $\{y_i = BN_{\gamma, \beta}(x_i)\}$

$$u_B \leftarrow \frac{1}{m} \sum_{i=1}^m x_i; \quad (7)$$

$$\sigma_B^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - u_B)^2; \quad (8)$$

$$\bar{x}_i \leftarrow \frac{x_i - u_B}{\sqrt{\sigma_B^2 + \varepsilon}}; \quad (9)$$

$$y_i \leftarrow \gamma \bar{x}_i + \beta \equiv BN_{\gamma, \beta}(x_i). \quad (10)$$

综上所述,可以将批归一化视作对数据的尺度缩放和偏移等系列操作流程,最终实现数据的恒等变换,保持数据特征的均衡分布。

3.2.4 Dropout

Dropout 是 Hinton 在 2012 年提出的一种防止过拟合的方法。Dropout 可以被视为正则化的一种,通过随机丢弃指定卷积层的部分神经元达到降低过拟合的目的,减少神经元之间的相互影响,使权重变化趋于独立,减小相互间的干扰,增强区分度。同时,Dropout 也可视作为一种增加模型复杂度的方法,通过训练含有不同神经元的网络,增强了模型的泛化能力。缺陷是训练时间长于未使用 Dropout 的网络。含有 Dropout 的神经网络更新公式具体如下:

$$r_j^{(l)} \sim \text{Bernoulli}(p); \quad (11)$$

$$\tilde{y}^{(l)} = r^{(l)} * y^{(l)}; \quad (12)$$

$$z_i^{(l+1)} = w_i^{(l+1)} \tilde{y}^l + b_i^{(l+1)}; \quad (13)$$

$$y_i^{(l+1)} = f(z_i^{(l+1)}). \quad (14)$$

其中,引入伯努利函数是用来设置每个结点随机失活的概率。

3.3 参数设置

根据网络结构,研究最终选择初始化参数批处理块大小为 100,轮训 45 次,动量 0.9,学习率恒为 0.008,根据前文提到的 BP 算法更新参数。通过对数据集进行训练辅以数据可视化,由此可以看到训练集和测试集的正确率与损失率的变化。

3.3.1 激励函数

研究可知,elu 及 tanh 函数图像如图 6、图 7 所示。对于本文使用的数据集,使用三层 elu^[14] 激励函数的训练正确率有 0.966,而使用三层 tanh 激励

函数的训练正确率略低,只达到 0.943,故而使用 elu 函数进行训练的速度要比 tanh 函数更快。进一步分析可知,在更大的数据集上训练结果的差距则会更明显。

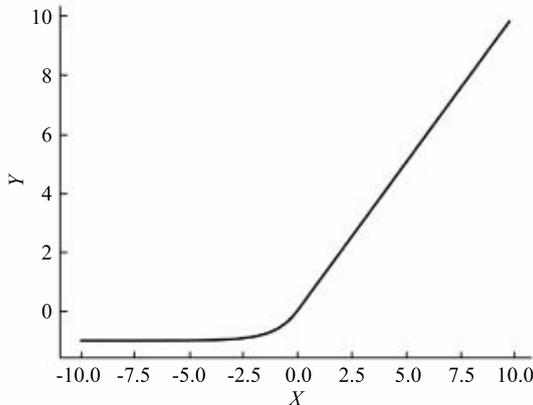


图 6 elu 函数图像

Fig. 6 elu function image

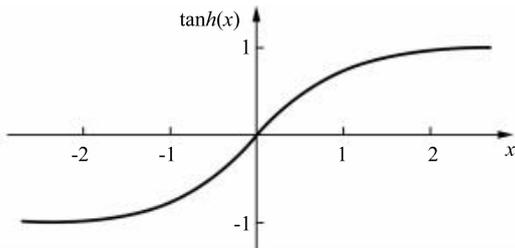


图 7 tanh 函数图像

Fig. 7 tanh function image

3.3.2 卷积核

网络架构共有两层卷积层,保持卷积核大小为 3×3 固定,结果见表 1。

人脸分类比数字识别更复杂,人脸的尺寸更大,分类所需的细节更繁复,因此相比于 LeNet-5 网络来说,就采用了更小、更多的卷积核。由表 1 中可以清晰看出,卷积核数量与精度大小的关系是先正比、后下降的,这即表明过多的卷积核容易造成过拟合。

表 1 卷积核个数对精度影响

Tab. 1 The influence of the number of convolution kernels on accuracy

参数	训练集	测试集	参数	训练集	测试集
10-10	98.22	91.53	15-20	99.78	95.10
10-15	98.71	93.57	20-20	99.99	94.36
10-20	99.29	96.71			

同时研究可知,在 10-20 的卷积核数目设置下,测试集正确率达到最大。

3.3.3 全连接层

模型中除了最后的分类层,只有一层全连接层。保持卷积核大小为 3×3 固定,激励函数为 elu,卷积核数量为 10~20,改变全连接层的神经元数量,实验结果见表 2。

表2 全连接层神经元数量对精度影响

Tab. 2 The influence of the number of neurons in the fully connected layer on accuracy %

参数	训练集	测试集	参数	训练集	测试集
1 100	99.76	94.53	1 250	99.87	96.71
1 150	99.88	94.64	1 300	99.98	91.26
1 200	99.89	94.33			

以上测试都是使用了批归一化与 Dropout 运行处理后得出的结果。而在单一使用 Dropout 或者单一使用批归一化的模型测试中,最终两者间的准确率结果数值已然非常接近。

4 结束语

本文提出了一个基于改进 LeNet-5 的多层神经网络用于对校园范围或处于单一年龄段范围内人群的人脸识别,对测试样本的分辨准确率最终达到 96.71%,对含有光照差异、表情变化、遮挡物等干扰因素具有鲁棒性。本次研究中,以 Windows10 专业版为测试系统,GTX 1060 6 G 的显卡作 GPU 加速,内存为 8 G 的计算机作为测试环境。在此环境下该模型运行效果良好,平均检测人脸用时低于 0.03 s。

参考文献

- [1] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507.
- [2] 陈超,齐峰.卷积神经网络的发展及其在计算机视觉领域中的应用综述[J].计算机科学,2019,46(3):63-73.
- [3] LECUN Y, KAVUKCUOGLU K, FARABET C. Convolutional networks and applications in vision[C]//Proceedings of 2010 IEEE International Symposium on Circuits and Systems. Paris, France:IEEE, 2010: 253-256.

- [4] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [5] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks[M]//FLEET D, et al. ECCV 2014, Part I, LNCS 8689. Switzerland:Springer, 2014: 818-833.
- [6] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]//International Conference on Learning Representations. San Diego, CA:dblp, 2015:1-14.
- [7] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Boston, MA, USA:IEEE, 2015: 1-9.
- [8] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2016: 770-778.
- [9] 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.
- [10] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors[J]. Nature, 1986, 323(6088): 533-536.
- [11] LIN M, CHEN Q, YAN S. Network in network[J]. arXiv preprint arXiv:1312.4400, 2013.
- [12] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[C]//Proceedings of the 32nd International Conference on Machine Learning. Lille:IEEE, 2015: 448-456.
- [13] HINTON G E, SRIVASTAVA N, KRIZHEVSKY A, et al. Improving neural networks by preventing co-adaptation of feature detectors[J]. arXiv preprint arXiv:1207.0580, 2012.
- [14] CLEVERT D A, UNTERTHINER T, HOCHREITER S. Fast and accurate deep network learning by exponential linear units (ELUs)[J]. arXiv preprint arXiv:1511.07289, 2015.

(上接第 74 页)

参考文献

- [1] 安娜.车载自组织网络环境中 AODV 协议的性能研究及改进[D].西安:长安大学,2017.
- [2] SHEN Xiaowei, WU Yi, XU Zhixin, et al. AODV-PNT: An improved version of AODV routing protocol with predicting node trend in VANET[C]//2014 IEEE 7th International Conference on Advanced Infocomm Technology (ICAIT), Fuzhou, China: IEEE, 2014: 91-97.
- [3] 陈哲愚,张建,陈燕.一种基于节点移动性的 AODV 改进协议[J].微电子学与计算机, 2010, 27(9): 155-158.
- [4] 谭小晴.无线自组织网络路由协议研究[D].南京:南京邮电大学,2016.
- [5] 夏志谋.基于 AODV 的高动态稳定路由协议研究[D].哈尔滨:哈尔滨工业大学,2017.
- [6] 刘荣,王东,李晓鸿.移动 Ad Hoc 网络中基于剩余生存时间的链路稳定性路由协议[J].计算机工程与科学, 2012, 34(12): 9-15.
- [7] 唐应辉,唐小我.排队论:基础与分析技术[M].北京:科学出版社, 2006.

- [8] QIU Tie, WANG Lei, GUO He, et al. A new modeling method for vector processor pipeline using queueing network[C]//5th International ICST Conference on Communications and Networking. Beijing, China:IEEE, 2010: 1-6.
- [9] BEUTLER F. Mean Sojourn times in Markov queueing networks: Little's formula revisited[J]. IEEE Transactions on Information Theory, 2003, 29(2): 233-241.
- [10] 章永龙. Dijkstra 最短路径算法优化[J].南昌工程学院学报, 2006, 25(3): 30-33.
- [11] 魏赞,杨晓光,何晓帆.城市道路交通场景下车辆自组织网络建模与仿真[J].兰州交通大学学报, 2018, 37(1): 14-20.
- [12] 马佳荣,赵祥模,马峻岩,等.基于 VANET 的高速公路事故消息快速广播机制[J].计算机工程, 2015, 41(11): 8-12.
- [13] 盘莉莉.基于 OPNET 的移动自组网路由协议的性能仿真[J].桂林航天工业高等专科学校学报, 2010, 15(2): 171-172, 180.
- [14] 任春江.单向三车道高速公路车载自组织网络路由协议研究[D].太原:太原理工大学,2017.