

基于改进卷积神经网络的动态障碍物检测方法

孙 凯 何明祥* 张 红 蒋纪威
(山东科技大学计算机科学与工程学院 山东 青岛 266590)

摘 要 针对传统卷积神经网络在感兴趣目标较小的情况下对动态障碍物检测会出现结果不准确的问题,提出一种基于改进卷积神经网络的运动障碍物检测方法。该方法在层层抽取细节特征的基础上融入全局特征,利用全局特征修正细节特征的提取,利用 Softmax 层进行分类来获取图像的整体信息。实验结果表明,与传统卷积神经网络相比,改进卷积神经网络具有较低的时间复杂度,以及较高的识别率。

关键词 深度学习 卷积神经网络 障碍物检测 感兴趣区域

中图分类号 TP391 **文献标识码** A **DOI**:10.3969/j.issn.1000-386x.2019.02.042

DYNAMIC OBSTACLE DETECTION BASED
ON IMPROVED CONVOLUTION NEURAL NETWORK

Sun Kai He Mingxiang* Zhang Hong Jiang Jiwei
(College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, Shandong, China)

Abstract Traditional convolution neural network can produce inaccurate results in dynamic obstacle detection when the target of interest is small. To solve this problem, we proposed moving obstacle detection method based on improved convolution neural network. The method integrated global features into detail feature extraction layer by layer. We used the global features to modify the extraction of detail features, and used the Softmax layer to classify to obtain the overall information of the image. The experimental results show that the improved convolution neural network has lower time complexity and higher recognition rate than the traditional convolution neural network.

Keywords Deep learning Convolutional neural network Obstacle detection Region of interest

0 引 言

障碍物检测是无人驾驶汽车和机器人研究方面的关键技术之一。目前,国内外研究人员在障碍物检测方面进行了大量的研究,但是由于检测所用的工具和应用环境的不同,其检测的具体方法也不尽相同。Zhou 等^[1]提出了一种基于多普勒频移和高度的障碍物检测方法;邹斌等^[2]提出一种基于三维激光雷达的动态障碍物检测和追踪方法;金立生等^[3]提出了一种基于 Adaboost 算法的日间前方车辆检测方法;Delp 等^[4]提出一种基于贝叶斯集群学习的分类方法;王铮等^[5]提出了一种基于双目视觉的 AGV 障碍物检测方法。

此外,卷积神经网络也是一种很好的障碍物检测方法,它擅长处理大型图像信息,目前已经成功应用于人脸识别和交通信号识别等领域。传统的卷积神经网络大多是采用卷积层和池化层循环抽样的多层级联的结构^[6-7],此类结构主要利用卷积层提取图像的细节特征,利用池化层来降低特征图的分辨率,最后通过全连接层进行分类整合得出结果^[8]。但是,在运动障碍物检测识别过程中过于考虑局部细节特征,容易受限于图像本身。例如需要识别的障碍物较远时,即当感兴趣的目标过于小时,按照多层级联的结构只靠提取细节来识别会影响准确率,反而适量的增加感兴趣目标的全局特征能达到比较好的效果^[9-10]。针对以上问题,本文通过增加全局特征来修

正传统卷积神经网络,在层层抽取细节特征的基础上融入了全局特征,利用全局特征修正细节特征的提取,进而提高对动态障碍物识别的准确率。

1 卷积神经网络

卷积神经网络(CNN)是一种多层非全连接的深度学习网络结构。CNN 通常由输入层、卷积层、池化层、全连接层以及输出层五层结构组成。传统的卷积神经网络 LenNet-5 是 LeCun 等^[11]提出的一种基于卷积神经网络的文字识别模型,其网络结构如图 1 所示。

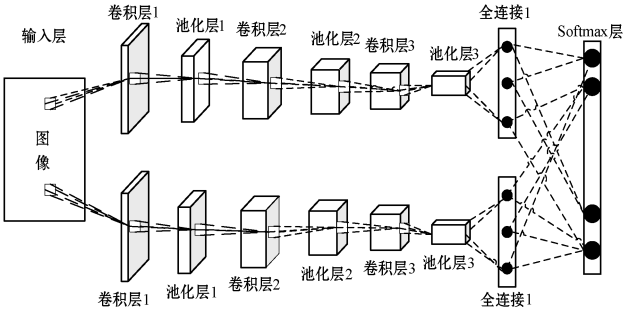


图 1 卷积神经网络结构示意图

1.1 卷积层

卷积层的主要作用是提取图像的特征信息,具备自学习能力的卷积核能够对上一层获得的特征图进行卷积计算,并加上一个合适的偏置数值,然后经过激活函数后输出形成这一层的特征图。

卷积层的计算公式如下:

$$\chi_j^i = f(\sum_{\alpha \in M_j^i} \chi_{j-1}^\alpha \times K_j^{\alpha i} + P_j^i) \tag{1}$$

式中: χ_j^i 表示第 i 层神经元的第 j 个特征图的输入, M_j^i 表示 j 层神经元第 i 个特征图对应的输入图像集, χ_{j-1}^α 表示第 $j-1$ 层神经元中第 α 个特征图的输出, $K_j^{\alpha i}$ 表示第 $j-1$ 层神经元第 α 个特征图与第 j 层神经元的第 i 个特征图间的权值, P_j^i 为 j 层神经元第 i 个特征图的偏置值, $f(\cdot)$ 表示激活函数。

1.2 池化层

池化层是对输入进行不重叠的采样操作,进一步提取图像特征,来降低特征图的分辨率,除去冗余的数据,减少图像特征维数,同时保持位移不变。

池化层的计算公式如下:

$$\chi_j^i = f(down(\chi_{j-1}^i + P_j^i)) \tag{2}$$

式中: $down(\cdot)$ 表示池化函数。

1.3 全连接层

全连接层是用来连接经过卷积和池化操作的全部

特征,并对其进行特征分类,再交由输出层输出最后的结果。

全连接层的计算公式如下:

$$\chi_j^i = f(\sum_{\alpha \in \chi^{i-1}} (\chi_j^{i-1} \times K_{\alpha j}^i + P_j^i)) \tag{3}$$

1.4 Softmax 层

Softmax 层是卷积神经网络中的最后一层,用来进行最后的分类和归一化,得到最终的图像信息。

Softmax 层的计算公式如下:

$$f(z_j) = \frac{e^{z_j}}{\sum_{i=1}^n e^{z_i}}$$

式中: z_i 表示第 i 类预测结果。

2 改进卷积神经网络

针对传统卷积神经网络在特定的情况下,因过度关注细节特征而影响识别的准确率这个问题,本文对传统卷积神经网络进行了改进。该网络结构是由上下两层并行组成,最后由一个全连接层整合两层的识别结果。上层的卷积神经网络主要对图像的细节特征进行采样,因此不改变原来的网络结构,下层卷积神经网络主要对上层采样的结构进行修正,对图像的全局特征进行采样,因此改进点主要在下层,网络结构如图 2 所示。

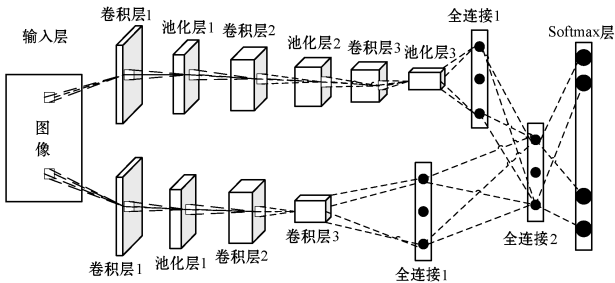


图 2 改进卷积神经网络结构图

如图 2 所示,该网络结构上半部分是一个由 3 层卷积层和 3 层池化层组成的传统卷积神经网络结构。该结构的主要作用是对细节的采样和表达图像的特征。为了保证对细节的足够采样,所有卷积层步长均为 1。第一个卷积层的卷积核为 8×8 ,第一层池化操作提取的特征图数为 16,第二个和第三个卷积层的卷积核为 3×3 ,第二层和第三层池化层提取的特征图数为 32。

如图 2 所示,下半部分是一个由 3 层卷积层和 1 层池化层组成的新卷积神经网络结构,目的是对图像的全局特征进行采样。第一层卷积层的卷积核为 $5 \times$

5、步长为1,第一层池化操作提取的特征图数为16,第二层和第三层卷积层的卷积核为 3×3 ,但是只保留初次卷积后的池化操作,因为第一次卷积操作会注重图像的边界提取,在此基础上保留池化层可以最大化地保留图像的全局特征。

全连接层在该网络结构中总共有3个,在上下两部分的网络结构中分别有2个具有400个单元的全连接层,分别对上下层网络结构提取的图像特征进行融合,然后由1个100个单元的全连接层将两部分融合在一起,最后经过Softmax层进行分类操作,输出一组图像与原图像相对应。

以下主要是本文对卷积神经网络的改进点的详细描述:

(1) 下半部分的卷积神经网络只保留第一次卷积之后的池化操作。下层卷积神经网络的第一层卷积层的卷积核为 5×5 、步长为1;第二层和第三层卷积层采用的均是 3×3 核。保留第一层的池化层是因为第一层的卷积层是处理图像边缘的操作,而对于目标边缘的识别尤为重要。采用 5×5 的卷积核也是为了最大限度地保留图像的全局特征。当然下半部分的卷积网络最大的缺点是有可能带来过拟合。

(2) 最后一层全连接层融合上下两部分卷积网络提取的特征。上层卷积神经网络通过全连接层1提取的是局部的图像特征,下层卷积神经网络通过全连接层2提取的是全局的图像特征,该改进的算法即是局部特征与全局特征融合在一起得到更为准确的图像识别,最后一步的融合至关重要,因此采用一层100个单元的全连接,最后用Softmax分类层进行分类识别。

3 基于改进卷积神经网络的动态障碍物识别方法

改进卷积神经网络在避障系统中用于对障碍物的识别和分类,其输入是经过处理后的点云图像,识别并分割出的障碍物,如汽车、自行车、行人等,输出为与障碍物相对应的一组对象,用于后续的操作处理。

3.1 映射栅格图坐标

栅格地图^[12]能准确地表达某一时刻的环境情况,比如当前车的位置、障碍物的位置等。在避障系统的研究中,可将栅格图作为最基础的环境建模,在栅格图的基础上进行智能决策和路径规划操作。本文研究的主要是运动的障碍物,因此选用可以更好地描述障碍物动态位置信息的三维障碍物栅格图作为表达环境的

建模方法。在三维坐标系内对环境进行建模,按时间轴表示出一段时间内栅格的占用情况,从而连续地展示环境中所有的运动关系,为接下来的行为决策和路径规划提供依据。

本文采用 512×512 的栅格划分标准,即将能够检测到的环境大小按比例映射到此 512×512 的环境地图中,将得到的某一时刻点的雷达原始数据中所有环境信息同样按比例映射到相应的位置,这样便得到了单点时刻的栅格图。依次按照时间顺序将周围的环境信息映射到同一个三维坐标系中,便得到一个三维的栅格占用图,在这个三维坐标系中可以明显地看出该环境中所有的运动情况。

在传统的二维栅格图中按照单个时间点映射得到的栅格图可得到单帧的动态障碍物栅格图,在本文中三维坐标系中按照时间轴多帧叠加便会得到一个三维的动态障碍物栅格图。图3(a)为单帧的动态障碍物栅格图示意图,图3(b)为多帧叠加的动态障碍物栅格图示意图,由图可知,多帧叠合的栅格图能够看出障碍物的运动状态变化。

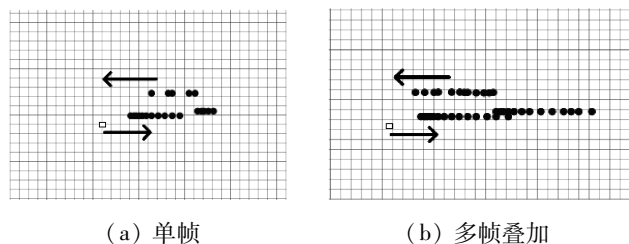


图3 动态障碍物栅格图示意图

3.2 感兴趣区域过滤

感兴趣区域^[13]指为硬件设施检测到的动态障碍物所处的环境区域,感兴趣过滤器处理感兴趣之外的点云数据,去除背景的物体,比如道路上的建筑物和树木等,剩下的就是感兴趣区域的点云,方便用于后续的处理。

此处采用检测到的视线内可操作的最小多边形进行过滤,经过感兴趣区域过滤后,可得到每个雷达点代表的物体是在感兴趣区域内还是区域外。

3.3 障碍物识别模块

假定该卷积神经网络已线下完成训练,获得使最小损失函数最小的权值并保留该权值。在经过感兴趣过滤器识别周围环境之后,识别模块得到了仅在感兴趣区域内的点云,这时大部分的背景障碍物已被去掉。同时通过线下训练完的本文改进卷积神经网络模型的上下两层网络结构,学习点云的特征预测障碍物的相关属性(前景物体的概率、物体的高度、物体分类等),

再经过卷积神经网络的特征识别、提取、聚类等操作,得到一组识别分类结果。

4 实验结果与分析

4.1 时间复杂度对比实验

本实验是在 CPU 为 Intel i5-6300,主频 2.30 GHz、四核,8 GB 内存,64 位 Microsoft Windows 10 操作系统,使用 Matlab 进行实验验证。采用的数据来自 Geiger 开放的 KITTI 数据集^[14],选取 1 000 张道路障碍物图片,其中 900 张用于训练,100 张用于测试。

表 1 中,时间复杂度公式中 $D_e k$ 、 $D_r k$ 表示单层输入层需要访问的参数, $\log_2(|V|)$ 表示输出层 Softmax 层需要访问的参数,其中 $|V|$ 表示词典的大小, D_e 表示原卷积神经网络单层词向量的维度, D_r 表示改进卷积神经网络的单层词向量维数。由于改进的卷积神经网络只保留了一层池化层,并且卷积层的卷积核数也有所降低,即 $D_e > D_r$,所以改进的卷积神经网络的时间复杂度比传统的卷积神经网络低。

表 1 实验检测速率比较

算法名称	时间复杂度	检测时间/s
CNN	$O(2D_e k + \log_2(V))$	0.251
改进 CNN	$O(D_e k + D_r k + \log_2(V))$	0.173

从表 1 给出两种算法的检测速率的比较结果可以看出,改进的卷积神经网络检测速率比传统的卷积神经网络有所降低,表明了本文算法具有较低的时间复杂度,因此本文方法的实时性也能够得到保证。

4.2 障碍物检测实验

本文的仿真实验基于百度开放 Apollo 的平台,采用的数据来自 Geiger 开放的 KITTI 数据集、城市交通数据集 Cityscapes^[15] 和 Apollo 数据集中的相机图片。

此处实验目的为验证改进卷积神经网络算法在图像识别中具有全局优势,因此对改进卷积神经网络上下部分的网络分别输入相同的样本图片,通过上下层网络结构中卷积层输出的特征图结果来验证。

首先手动筛选出一部分较易识别、边缘较明显的图片组成样本图片库,分别通过本文中改进卷积神经网络上下层结构,输出样本图片经过卷积层的特征图。图 4 和图 5 分别展示出了经过上下两层卷积网络得到的特征图,可以看出,上层卷积网络输出的特征图较为抽象,而下层卷积网络输出的特征图较为全局,能够更好地反映出图像的特征。

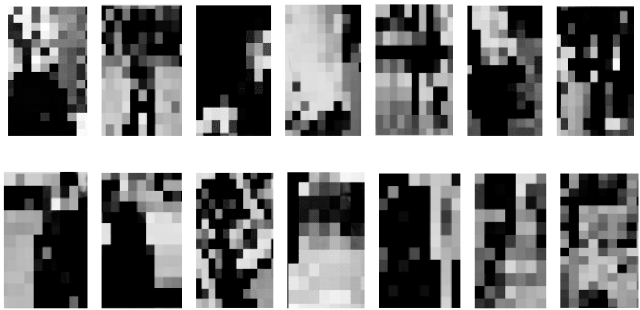


图 4 上层卷积神经网络输出的特征图



图 5 下层卷积神经网络输出的特征图

浅层神经网络具有全局性的特点,本文改进卷积神经网络是受此启发,利用浅层神经网络的全局特性对传统卷积神经网络对图像进行层层抽象的特点进行补充修正,从而改善卷积神经网络过度抽取细节特征从而缺乏全局特征的问题;同时,通过下层卷积神经网络对全局特征的把握,可以改善在训练学习过程中出现过拟合的概率。

为了验证本文提出的改进方法在动态障碍物识别过程中的适用性,本文采用以上数据集的真实数据构建了一个 91 271 个样本的测试库,并将其分为测试集和训练集两部分,样本集包括三个数据集中的数据,包括多种场景下的运动障碍物的样本。表 2 中列出了本文在训练及测试时各个数据库中样本数量的情况。

表 2 实验样本集

数据集名称	训练集数量	测试集数量
KITTI	26 373	1 132
Cityscapes	23 511	5 560
Apollo	32 445	2 250
样本集	82 329	8 942

分别采用表 2 中相同的训练集(82 329 条数据)对传统的卷积神经网络和本文改进卷积神经网络分别进行权值训练,分别保存上线两层卷积神经网络训练后的权值数值。再采用表 2 中的测试集(8 942 条数据)进行实验验证,表 3 列出了本文改进卷积神经网络在各个数据集的测试结果。通过改进 CNN 与 CNN 的算法对比,可以直观地看出本文算法在运动障碍物的

识别上的优势。

表 3 实验测试结果

算法名称	数据集识别准确率/%			
	KITTI	Cityscapes	Apollo	样本集
改进 CNN	98.62	97.24	96.57	97.74
CNN	97.84	96.37	95.62	96.81

上述实验采取的数据集均是采用道路真实的数据,从侧面展示出了改进的算法在不同的环境中具有一定的适应能力,在进行验证实验的同时用改进的算法对三个数据集进行了测试。图 6 为算法改进前后的召回率曲线,从图中可以看出,本文的改进卷积神经网络算法在三个数据集中具有较强的适应能力。

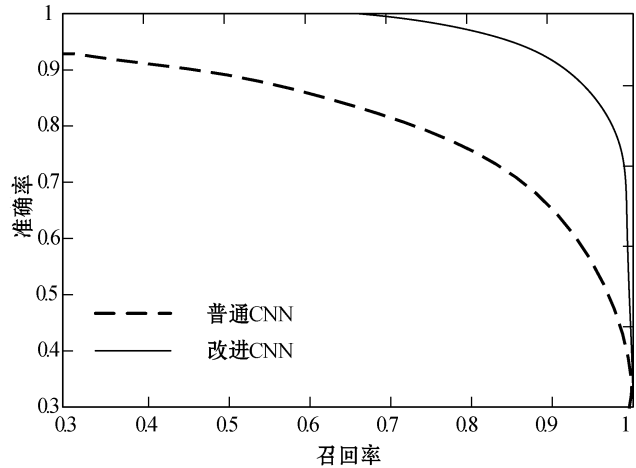


图 6 召回率曲线

为了验证本文改进卷积神经网络在增加全局特性后的有效性,在 Apollo 数据库中分别测试了改进 CNN、下层全局粗特征 CNN 和上层局部细化 CNN,测试结果的 ROC 曲线(受试者工作曲线 Receiver Operating characteristic Curve)如图 7 所示。可以明显地看出改进卷积神经网络算法在识别障碍物的性能上明显比上下层网络结构的性能好。

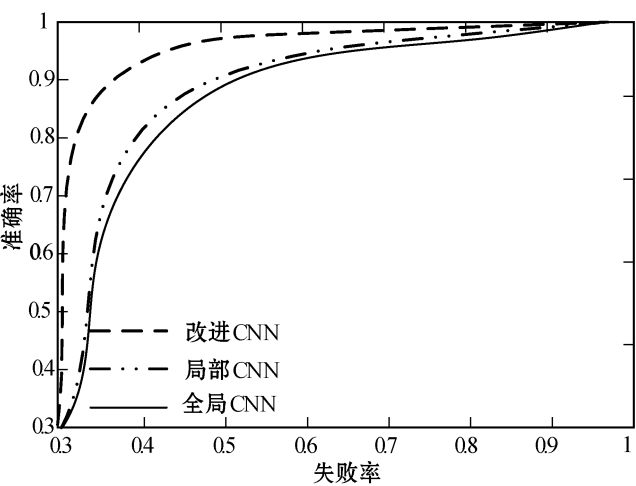


图 7 改进前后 ROC 曲线对比

5 结 语

本文提出了一种基于改进的卷积神经网络对动态障碍物的识别方法。与传统的 CNN 相比,该网络在保留局部特征提取力度的同时增加全局特征的修正,通过局部特征和全局特征的同时提取,再通过一个全连接层将局部特征与全局特征进行整合,最终获取图像的整体信息。实验结果也表明,改进的卷积神经网络在对运动障碍物的识别具有较强的适应性和有效性。同时,由于实验中所使用训练样本数量和种类的限制,对实验的结果会有一定的影响,因此接下来的工作中,要扩大训练样本集的数量和种类来进一步验证改进的 CNN 比传统的 CNN 在动态障碍物识别方面有更高的准确性。

参 考 文 献

[1] Zhou H P, Liu Q B, Chen D P, et al. Doppler Shift and Height Detection of Obstacle Based on FMCW Radar Sensor [C]//2015 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC). IEEE, 2015 :399 - 402.

[2] 邹斌,刘康,王科未. 基于三维激光雷达的动态障碍物检测和追踪方法[J]. 汽车技术, 2017(8) :19 - 25.

[3] 金立生,王岩,刘景华,等. 基于 Adaboost 算法的日间前方车辆检测[J]. 吉林大学学报:工学版,2014,44(6) :1604 - 1608.

[4] Michael D, Nagasaka N, Kamata N, et al. Classifying and passing 3D obstacles for autonomous driving[C]//IEEE International Conference on Intelligent Transportation Systems. IEEE, 2015 :1240 - 1247.

[5] 王铮,赵晓,余宏杰,等. 基于双目视觉的 AGV 障碍物检测与避障[J]. 计算机集成制造系统, 2018, 24(2) :400 - 409.

[6] Tian Y, Luo P, Wang X, et al. Pedestrian detection aided by deep learning semantic tasks[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2015 :5079 - 5087.

[7] Ahmed E, Jones M, Marks T K. An improved deep learning architecture for person re-identification[C]//Computer Vision and Pattern Recognition. IEEE, 2015 :3908 - 3916.

[8] 张永良,张智勤,吴鸿韬,等. 基于改进卷积神经网络的周界入侵检测方法[J]. 计算机科学,2017,44(3) :182 - 186.

[9] Brezak M, Petrovic I. Real-time approximation of clothoids with bounded error for path planning applications[J]. IEEE Transactions on Robotics, 2014, 30(2) :507 - 515.

4 结 语

传统的浅层学习网络过度依赖人工选择火焰特征,无法适应复杂多变的监控场景。本文在DenseNet深度神经网络架构的基础上,提出了一种改进的火灾检测模型。为了增强特征传播的精度,降低存储数据量,本文采取结构化稀疏操作,通过将网络卷积核分为多个组,在训练过程中逐渐减小每个组内不重要的参数连接。同时,本文引入Focal损失函数对分类层进行火灾识别,大大增强了模型的泛化能力。大量的定性定量实验表明本文提出的火灾检测方法能够有效定位到火灾图像位置,对于火焰、烟雾等火灾现象均能有效识别,且不易受干扰环境的影响,适合工程应用。下一步,我们将着手在现有硬件平台下对网络模型进行移植,为油气田安全、高效生产运行提供技防保障。

参 考 文 献

- [1] 洪伟,李朝锋. 基于区域全卷积网络结合残差网络的火焰检测方法[J]. 激光与光电子学进展,2018,55(4): 168-173.
- [2] 朱晓好,严云洋,刘以安,等. 基于深度森林模型的火焰检测[J]. 计算机工程,2018,44(7):264-270.
- [3] 曾思通,吴海彬,沈培辉. 基于多特征融合的视频火焰检测方法研究[J]. 图学学报,2017,38(4):549-557.
- [4] 严云洋,陈垂雄,刘以安,等. 维度加权模式动态纹理特征的火焰检测[J]. 智能系统学报,2017,12(4): 548-555.
- [5] Chen T H, Wu P H, Chiou Y C. An early fire-detection method based on image processing[C]//2004 International Conference on Image Processing. IEEE, 2004:1707-1710.
- [6] Dimitropoulos K, Bampoutis P, Grammalidis N. Spatio-temporal flame modeling and dynamic texture detection analysis for automatic video-based fire[J]. IEEE Transactions on Circuits&Systems for Video Technology,2015, 25(2): 339-351.
- [7] Toreyin B U, Dedeoglu Y, et al. Computer vision based method for real-time fire and flame detection[J]. Pattern Recognition Letters, 2006, 27(1):49-58.
- [8] Gubbi J, Marusic S, Palaniswami M. Smoke detection in video using wavelets and support vector machines[J]. Fire Safety Journal, 2009, 44(8): 1110-1115.
- [9] Listyorini T, Rahim R. A prototype fire detection implemented using the Internet of things and fuzzy logic[J]. World Transactions on Engineering & Technology Education, 2018, 16(1):42-46.
- [10] 钟玲,张兴坤. 基于SVM的视频图像火焰检测[J]. 软件工程,2017,20(6):1-4.
- [11] Emfjälletal. Fire detection system[J]. Optoelectronics Instrumentation & Data Processing, 2018, 45(2): 140-145.
- [12] Sermanet P, Chintala S, Lecun Y. Convolutional neural networks applied to house numbers digit classification[C]//2012 21st International Conference on Pattern Recognition (ICPR 2012). IEEE Computer Society, 2012.
- [13] Springenberg T, Dosovitskiy A, Brox T, et al. Striving for simplicity: The all convolutional net[C]//Proceedings of International Conference on Learning Represenation Workshop Track, 2015.
- [14] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE,2016.
- [15] Huang U, Liu Z, Weinberger KQ, et al. Densely connected convolutional networks[C]//The IEEE Conference on Computer Vision and Pattern Recognition (CPVR). Las Vegas: IEEE, 2017:4700-4708.
- [16] Liu W, Anguelov D, Erhan D, et al. SSD: Single Shot MultiBox Detector[C]//European Conference on Computer Vision, 2016: 21-37.
- [17] Xie S, Girshick R B, Dollar P, et al. Aggregated Residual Transformations for Deep Neural Networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 5987-5995.

(上接第239页)

- [10] 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.
- [11] 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.
- [12] 朱爱斌,刘洋洋,何大勇,等. 解决路径规划局部极小问题的势场栅格法[J]. 机械设计与研究, 2017(5):46-50.
- [13] 钱基德,陈斌,钱基业,等. 基于感兴趣区域模型的车道线快速检测算法[J]. 电子科技大学学报, 2018,47(3):356-361.
- [14] Demetz O, Hafner D, Weickert J. Morphologically invariant matching of structures with the complete rank transform[M]. Kluwer Academic Publishers, 2015.
- [15] Cordts M, Omran M, Ramos S, et al. The cityscapes dataset for semantic urban scene understanding[EB]. eprint arXiv: 1604.01685, 2016.