

引用格式: 郜春艳, 何秀娟, 黄文美, 等. 基于2-D范围扫描的室内场景识别方法[J]. 电光与控制, 2018, 25(12): 30-34. GAO C Y, HE X J, HUANG W M, et al. An indoor scene recognition method based on 2-D range scanning[J]. Electronics Optics & Control, 2018, 25(12): 30-34.

## 基于2-D范围扫描的室内场景识别方法

郜春艳, 何秀娟, 黄文美, 刘卓锟

(河北工业大学省部共建电工装备可靠性与智能化国家重点实验室, 天津 300132)

**摘要:** 对于在室内工作的机器人而言, 要完成不同环境的自主导航就必须能够对其所处场景进行有效识别。传统的方法是通过视觉或雷达等传感器对所处环境进行匹配以实现场景识别。提出了一种基于2-D范围扫描的室内场景识别的方法。该方法对激光雷达的范围扫描信息进行特征提取, 利用所提取的样本训练基于局部感受野的极限学习机, 对多种室内场景进行分类识别。在Gazebo搭建的仿真环境中采集虚拟范围扫描数据, 对室内场景识别方法进行了研究。利用DR Dataset提供的测距数据对所提出的方法进行了实验验证。实验结果表明: 该方法的室内场景识别准确率高于传统方法。基于2-D范围扫描场景识别的研究也为机器人实现自主导航提供理论依据和实验数据。

**关键词:** 移动机器人; 自主导航; 2-D范围扫描; 室内场景识别; 基于局部感受野的极限学习机

**中图分类号:** TP181 **文献标志码:** A **doi:**10.3969/j.issn.1671-637X.2018.12.007

## An Indoor Scene Recognition Method Based on 2-D Range Scanning

GAO Chun-yan, HE Xiu-juan, HUANG Wen-mei, LIU Zhuo-kun

(State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology, Tianjin 300132, China)

**Abstract:** The robots working indoors must be able to effectively identify their surroundings to complete the autonomous navigation in different scenes. Traditional approaches realize scene recognition by using visual or radar sensors to match the scene. A method of indoor scene recognition based on 2-D range scanning is proposed. This method extracts the features of range scanning information of the lidar, and Extreme Learning Machine Based on Local Receptive Fields (ELM-LRF) is trained by using extracted samples to classify and identify various indoor scenes. In the simulation environment built by Gazebo, the virtual range scanning data is collected, and then the indoor scene recognition methods are studied. The proposed method is verified by experiments based on the range data provided by DR Dataset. The results show that the recognition accuracy of the proposed method is higher than that of traditional methods. The study of scene recognition based on 2-D range scanning also provides theoretical support and experimental data for autonomous robot navigation.

**Key words:** mobile robot; autonomous navigation; 2-D range scanning; indoor scene recognition; extreme learning machine based on local receptive field

### 0 引言

场景识别是室内移动机器人研究的一个重要分支<sup>[1]</sup>, 移动机器人的诸多功能如定位、导航、路径规划的实现都基于其良好的场景识别能力<sup>[2]</sup>。利用2-D测

距信息进行机器人场景识别技术的研究是当前一大热点<sup>[3]</sup>。不同的室内场景由于结构和布置不同而有着不同的形状与大小特征, 比如走廊呈现规则的矩形, 又窄又长; 卧室的床与墙之间存在窄缝; 客厅面积宽阔且较为杂乱, 卫生间面积最小。这些特点都是基于2-D扫描进行室内场景识别的重要依据<sup>[4-5]</sup>。

安装有激光雷达传感器的移动机器人在室内行驶的过程中, 范围扫描数据也会被实时记录下来, 这些数据描述了移动机器人所处环境的轮廓特征。基于激光雷达扫描的室内场景识别问题可以描述为: 给定一系

收稿日期: 2017-12-19

修回日期: 2018-02-12

基金项目: 国家自然科学基金(51777053); 河北省自然科学基金(E2017202035); 河北省高层次人才项目(C2015003037)

作者简介: 郜春艳(1991—), 女, 江苏盐城人, 硕士, 研究方向为机器学习和模式识别、新型磁性材料与器件。

列来自不同类型室内场景的激光雷达测距样本,利用机器学习算法对已知样本进行学习,归纳总结出各类场景特有的数据特征,当再给予陌生的测距样本时,可判断出其所属室内场景的类型<sup>[6-8]</sup>。

本文在 Ubuntu 系统下,利用机器人操作系统(ROS)和仿真软件 Gazebo 搭建仿真环境,并采集虚拟激光雷达测距信息,对范围扫描数据提取二值化图像信息后,利用基于局部感受野的极限学习机(ELM-LRF)对特征信息进行学习与分类。最后,利用公共数据集 DR Dataset 所提供的从 123 个真实生活场景中所采集到的激光雷达范围扫描信息对所提出的方法进行了验证。

## 1 仿真环境中的数据采集

### 1.1 搭建仿真环境

ROS 以其点对点设计、模块化程度高、代码可重用性强、采用分布式架构等优点成为目前国内外机器人研究者的主要操作平台。在 ROS 操作平台中利用仿真软件 Gazebo 能够模拟现实生活场景搭建 3-D 仿真环境。相比普通 2-D 仿真环境,如 MobileSim, Gazebo 能够更加真实地模拟事物的各种属性以及物体之间在物理上合理的相互作用,如摩擦、碰撞、打滑等。两种仿真环境的直观展现如图 1 所示。

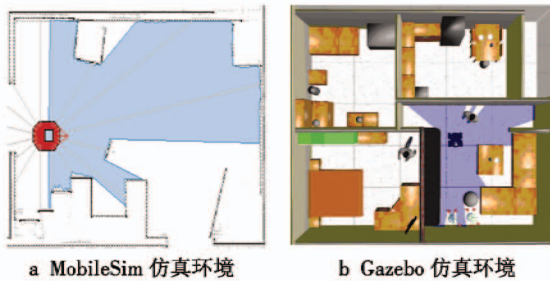


图 1 仿真环境

Fig. 1 Simulation environment

本文共搭建了 4 种仿真环境,包括客厅、卧室、厨房、卫生间,每个房间的大小、形状、位置都参考真实生活场景,每个场景中都有相应的摆设,如客厅摆有茶几、沙发等家具。由于 Gazebo 的编辑界面仅包含一些简单模型,衣柜、花盆等模型需在 solidworks 中进行编辑,再将文件导入 Gazebo。

### 1.2 数据采集

在仿真场景中,模拟移动机器人 Pioneer 3-AT 进行数据采集,在该移动机器人平台上安装激光雷达传感器 RP LIDAR A1,其最大扫描距离为 6 m,扫描范围为 360°,每进行一次范围扫描,将提取到一组包含 400 个采样点的距离数据的测距信息。

通过 ROS 节点发布命令消息控制移动机器人的行驶方向与时间,使移动机器人在仿真环境中无碰撞

前行,且所经路途均无重复。同时对激光雷达主题进行订阅,实时记录并保存所采集到的测距信息。由于各房间大小不同,在每个房间中所采集到的样本数量也有所差别,卧室样本为 751、厨房样本为 678、客厅样本为 1338、卫生间样本为 572。

## 2 数据特征提取

每一组原始的激光雷达测距数据均为单纯的离散数据集,各个数据的关联性很弱,边缘特征和局部特征均很模糊,无法进行有效的特征提取,故需要对数据进行处理。特征提取过程如图 2 所示。

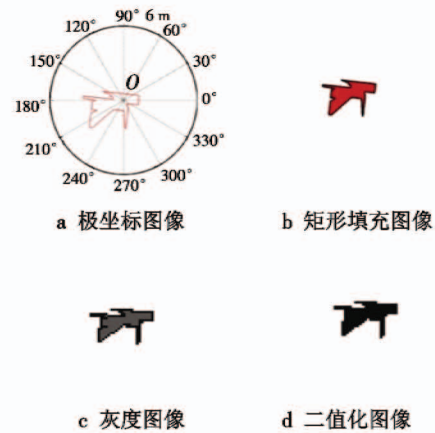


图 2 特征提取过程

Fig. 2 Feature extraction process

以任意一组激光雷达所测得的距离数据为例,利用 Matlab 将其转化为极坐标图,如图 2a 所示,图像的中心点为  $O$ ,进行范围扫描时,将装有激光雷达或声纳传感器的移动机器人所处的位置相邻的点连接起来便得到红色边缘部分,表示该时刻所扫描到的环境的轮廓信息。

为除去冗余信息,进一步提取有效特征,如图 2b 所示,将极坐标图提取到直角坐标系中,生成矩形图,矩形图的中点为极坐标图的原点。为增强其特征信息,对极坐标图中的范围轮廓内部进行颜色填充,此时所得到的图像为以白色为背景的三通道彩色图像。

图像处理对计算机的性能要求较高,所占内存较大,为简化计算,减少空间占用,先对彩色图像进行灰度化处理,将每个像素点的 R, G, B 分量分别求平均值,得到灰度图像的像素信息矩阵,如图 2c 所示。

随后使用最大类间方差法找到灰度图像像素点的一个合适的阈值,将大于阈值的像素点的值转换为 1,小于阈值的像素点的值转换为 0,完成灰度图像到二值图像的转换,如图 2d 所示。最后,二值化图像像素矩阵将作为输入基于局部感受野的极限学习机的训练样本和测试样本。

### 3 基于局部感受野的极限学习机

基于局部感受野的极限学习机(ELM-LRF)是一种快速有效的机器学习算法。该算法在输入层节点和隐藏层节点引入了根据连续概率分布函数产生的局部感受野,提出利用组合节点将多个隐藏层节点的输出进行组合。并且可随机生成卷积层节点与输入层节点之间的输入权重,各网络参数在学习过程中均不需要进行迭代调整,最终可产生唯一的最优解,因此,与SVM相比,ELM-LRF的学习速度要快几十倍<sup>[9]</sup>。

为了充分提取输入的特征,ELM-LRF对同一输入进行多次卷积,生成多个互异的特征图,假设为 $K$ 个,需要产生 $K$ 个不同的输入权重,并且不同特征图的输入权重互异,同一特征图的输入权重共享。如图3所示,设输入的大小为 $d \times d$ ,局部感受野的大小为 $r \times r$ ,隐藏层节点即为卷积节点。随机生成初始权重矩阵 $\hat{A}^{init} \in \mathbf{R}^{r \times K}$ , $\hat{A}^{init} = [\hat{a}_1^{init}, \hat{a}_2^{init}, \dots, \hat{a}_K^{init}]$ ,其中, $\hat{a}_k^{init} \in \mathbf{R}^r$ , $k=1, \dots, K$ ,并利用奇异值分解运算(SVD)对初始权重 $\hat{A}^{init}$ 进行正交化,若 $r^2 < K$ ,先将 $\hat{A}^{init}$ 进行转置,经正交化运算后再转置回来,记所得输入权重矩阵为 $\hat{A}$ 。

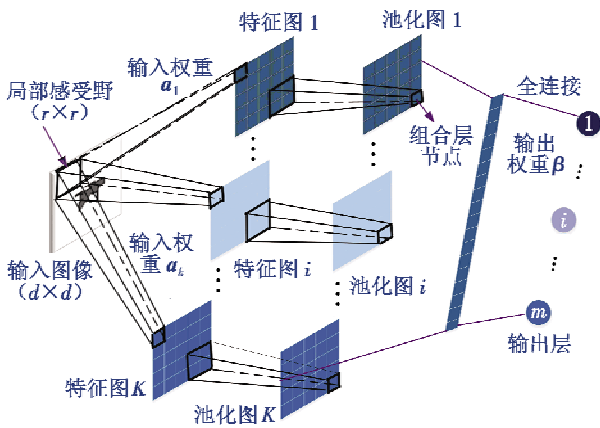


图3 基于局部感受野的极限学习机原理图

Fig.3 Extreme learning machine based on the local receptive field

第 $k$ 个特征图的输入权重为 $a_k \in \mathbf{R}^{r \times r}$ ,由 $\hat{a}_k$ 逐列排成。设该特征图的大小为 $(d-r+1) \times (d-r+1)$ ,卷积节点 $(i,j)$ 的值 $c_{i,j,k}$ 为

$$c_{i,j,k}(x) = \sum_{m=1}^r \sum_{n=1}^r (x_{i+m-1,j+n-1} \cdot a_{m,n,k}) \quad (1)$$

式中, $i,j=1,2,\dots,(d-r+1)$ 。

组合层节点即为池化节点,对所得到的特征图进行平方根池化处理,这样便将多个卷积节点所得到的特征值组合为一个节点值。设 $e$ 为池化的大小,表示池

化区域的中心到边的距离,若 $h_{p,q,k}$ 表示第 $k$ 个池化图中的组合节点 $(p,q)$ 的值,则 $h_{p,q,k}$ 的算式为

$$h_{p,q,k} = \sqrt{\sum_{i=p-e}^{p+e} \sum_{j=q-e}^{q+e} c_{i,j,k}^2} \quad p,q=1,2,\dots,(d-r+1) \quad (2)$$

对池化层进行全连接处理,即将某个输入所得到的所有组合节点的值连接成一个行向量,假设共有 $N$ 个输入,则可得到 $N$ 个组合的行向量,将它们组合在一起,形成组合层矩阵 $H \in \mathbf{R}^{N \times K \cdot (d-r+1)^2}$ ,采用正则化最小二乘法计算输出权重矩阵 $\beta$ 。

当 $N > K \cdot (d-r+1)^2$ 时,

$$\beta = \left( \frac{I}{C} + H^T H \right)^{-1} H^T T \quad (3)$$

当 $N \leq K \cdot (d-r+1)^2$ 时,

$$\beta = H^T \left( \frac{I}{C} + H H^T \right)^{-1} T \quad (4)$$

式中: $C$ 为可调参数; $T$ 为输入样本所对应的标签矩阵。本文中,输入为二值化图像像素矩阵,其大小为 $43 \times 43$ ,局部感受野的大小为 $5 \times 5$ 。

### 4 实验与结果分析

每个房间随机选取80%的样本进行训练学习,其余样本用于测试。则训练集共包含2670个样本,其中600个取自卧室,542个取自厨房,458个取自卫生间,1070个取自客厅,其余样本全部组成测试集。给不同类型的房间设置不同的标签,卧室为1,厨房为2,客厅为3,卫生间为4。生成训练标签矩阵的大小为 $2670 \times 4$ ,若第 $k$ 个样本取自客厅,则训练标签矩阵的第 $k$ 行,除第3列的值为1,其余列的值都为0。同样,生成测试标签矩阵。

将特征图个数设置为30,调节参数 $C$ 从 $10^{-4}$ 到 $10^4$ ,将训练集以及训练标签矩阵输入ELM-LRF对网络进行训练,获得输出权重 $\beta$ ,随后将测试集和测试标签输入已完成训练的网络结构,得到测试正确率。所得实验结果如图4所示。

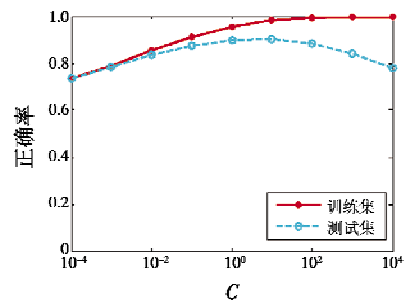


图4 训练与测试正确率规律图

Fig.4 The correct rate of training and testing



从图中可看出,随着  $C$  的增大,训练样本的分类正确率越来越高,最终达到 100%,测试样本的分类正确率逐渐增大达到 90% 左右后逐渐减小,稳定性较好。

图 5 表示各个房间的具体分类情况,可以看到,每个房间的分类正确率都达到了 80% 以上,分类效果很好。

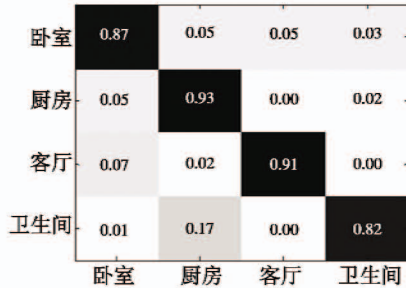


图 5 基于 ELM-LRF 分类的混淆结果

Fig. 5 Confusion results based on ELM-LRF classification

利用 SVM 对激光雷达测距数据进行分类实验,测试结果如图 6 所示。

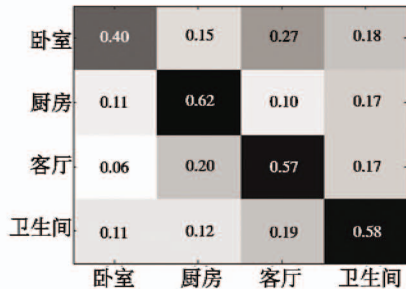


图 6 基于 SVM 分类的混淆结果

Fig. 6 Confusion results based on SVM classification

与图 5 进行对比可见,应用 ELM-LRF 的房间分类测试结果比 SVM 要好,前者每一类房间的认识率都比后者高很多,且前者的训练时间比后者快几十倍。

### 5 基于 DR Dataset 的验证实验

斯洛文尼亚卢布尔雅那大学 Vicos 实验室的 URŠIČ P 等辗转 4 个城市,选择了 24 个不同的家庭进行了激光雷达范围扫描数据采集,其中包括 6 个走廊,12 个卫生间,35 个浴室,21 个客厅,21 个厨房以及 28 个卧室,共采集了 46450 组测量数据,生成了公共数据集 DR Dataset,数据真实全面可靠,受到众多研究者的青睐<sup>[10-12]</sup>。

图 7 展示了数据采集平台,在移动机器人 Pioneer 3-DX 上安装激光测距雷达 Hokuyo URG,其离地高度约为 30 cm,测量半径为 5.6 m,扫描范围为 240°,同时在移动机器人平台上安装里程计,激光雷达测距信息与相应的里程计读数以 10 帧/s 的频率输出到计算机,这样便记录下了该场景下的测量信息。



图 7 DR Dataset 数据采集平台

Fig. 7 DR Dataset data acquisition platform

原始的测量数据包括时间戳、激光雷达测距数据、里程计读数等信息,分别将每组测量数据的激光雷达范围扫描数据提取出来,转化为二值化图像并将像素矩阵作为 ELM-LRF 的输入样本。每类场景中,随机选取 80% 的样本组成训练集,其余样本组成测试集,同时生成相应的训练标签矩阵和测试标签矩阵。

在  $K$  分别为 10, 20, 30, 40 的情况下,调节  $C$  从  $10^{-5}$  增长到  $10^4$ ,实验结果如图 8 所示。

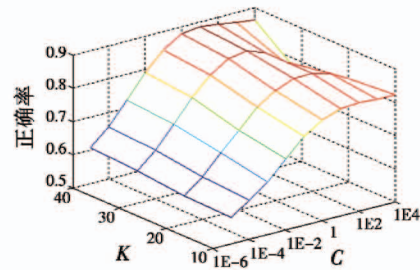


图 8 测试正确率变化规律图

Fig. 8 Correct rate of test

从图中可看出,随着  $C$  的增大,测试正确率逐渐升高,当  $C = 10$  时正确率达到最大值,随后逐渐降低。 $K$  越大,测试正确率越高,当  $C = 10, K = 40$  时,总体的测试正确率最高,达到 89% 左右。表 1 描述了当  $C = 10, K = 40$  时,训练集与测试集的样本组成情况以及各类房间测试集分类结果的具体情况,每类场景的识别率均在 80% 以上。

表 1 样本集组成与分类结果统计表

Table 1 Statistics of sample set composition and classification result

房间	训练集样本数	测试集样本数	分类错误的样本个数	识别率
走廊	1584	396	71	0.82
浴室	8428	2107	126	0.94
卧室	10 218	2554	255	0.90
客厅	8793	2198	374	0.83
厨房	4654	1164	210	0.82
卫生间	3483	871	140	0.84

实验结果说明,本文提出的方法可有效地应用于真实的室内场景。

## 6 结束语

本文将激光雷达测距数据转化为图像信息,用于处理移动机器人的室内场景识别问题,将基于局部感受野的极限学习机作为分类器,大大提高了分类效果。实验结果证明:本文提出的方法可满足移动机器人对室内场景识别的要求,适用于各种室内环境。

### 参考文献

- [1] ORABONA F, CASTELLINI C, CAPUTO B, et al. Indoor place recognition using online independent support vector machines [C]//Proceedings of British Machine Vision Conference, 2007:1090-1099.
- [2] 许坤,邹杰,陈谋. 基于联邦滤波的室内移动机器人定位[J]. 电光与控制,2016,23(9):77-83, 110.
- [3] URSIC P, LEONARDIS A, SKOCAJ D, et al. Hierarchical spatial model for 2D range data based room categorization [C]//IEEE International Conference on Robotics and Automation, 2016:4514-4521.
- [4] LIU H P, SUN F C, FANG B, et al. Robotic room-level localization using multiple sets of sonar measurements[J]. IEEE Transactions on Instrumentation and Measurement, 2016, 66(1):2-13.
- [5] PARK S, ROH K S. Coarse-to-fine localization for a mobile robot based on place learning with a 2-D range scan [J]. IEEE Transactions on Robotics, 2016, 32(3):528-544.
- [6] DEMARCO K J, WEST M E, HOWARD A M. A computationally-efficient 2D imaging sonar model for underwater robotics simulations in Gazebo[C]//OCEANS, 2015:1-7.
- [7] TAKAYA K, ASAI T, KROUMOV V, et al. Simulation environment for mobile robots testing using ROS and Gazebo [C]//The 20th International Conference on System Theory, Control and Computing (ICSTCC), 2016:96-101.
- [8] QIAN W, XIA Z Y, XIONG J, et al. Manipulation task simulation using ROS and Gazebo[C]//IEEE International Conference on Robotics and Biomimetics (ROBIO), 2014:2594-2598.
- [9] HUANG G B, BAI Z, KASUN L L C. Local receptive fields based extreme learning machine[J]. IEEE Computational Intelligence Magazine, 2015, 10(2):18-29.
- [10] LYU Q, NIU X, DOU Y, et al. Hyperspectral image classification via kernel extreme learning machine using local receptive fields [C]//IEEE International Conference on Image Processing (ICIP), 2016:256-260.
- [11] 李桂芝,安成万,杨国胜,等. 基于场景识别的移动机器人定位方法研究[J]. 机器人,2005,27(2):123-127.
- [12] 鄢文浩,贺赛先,沈婷婷. 基于随机运动障碍避碰规则的机器人路径规划[J]. 电光与控制,2017,24(5):73-76.

(上接第 29 页)

取目标的特征信息,提高红外图像分割的准确性;与传统基于稀疏矩阵的谱聚类图像分割方法相比,该方法的聚类速度提高了约 16 倍,在一定探测距离下,本文方法能更好地实现目标与背景的温差计算,为完成目标探测打下基础。

### 参考文献

- [1] 韩玉阁,宣益民. 战车红外隐身效果评估方法研究[J]. 红外技术,2003,25(6):22-24.
- [2] 秦娜. 装甲车辆在红外隐身措施下的仿真评估[D]. 南京:南京理工大学,2015.
- [3] 高原,刘剑,张俊举,等. 武器装备红外隐身效果评估方法研究[J]. 红外技术,2017,39(11):1060-1065.
- [4] 程程. 红外成像系统图像建模仿真性能评价研究[D]. 北京:北京理工大学,2015.
- [5] 董一峻,陈铮,刘玉峰,等. 视距估算在舰船红外隐身性能评估中的应用研究[C]//第九届全国光电技术学术交流会论文集,2010:326-329.
- [6] 游瑞蓉,王新伟,任鹏道,等. 约翰逊准则的视频监控目标检测性能评估方法[J]. 红外与激光工程,2016,45(12):1217003-1-1217003-6.
- [7] 阮秋琦. 数字图像处理学[M]. 北京:电子工业出版社,2007.
- [8] 王坤,张杨,宋胜博,等. 改进二维 OTSU 和自适应遗传算法的红外图像分割[J]. 系统仿真学报,2017,29(6):1229-1236.
- [9] 刘仲民,李战明,李博皓,等. 基于稀疏矩阵的谱聚类图像分割算法[J]. 吉林大学学报:工学版,2017,47(4):1308-1313.
- [10] 宋长新,马克,秦川,等. 结合稀疏编码和空间约束的红外图像聚类分割研究[J]. 物理学报,2013,62(4):040702-1-040702-10.
- [11] DENG J, DONG W, SOCHER R, et al. ImageNet: a large-scale hierarchical image database [C]//IEEE Conference on Computer Vision and Pattern Recognition, 2009:248-255.
- [12] LAWRENCE S, GILES C L, TSOI A C, et al. Face recognition: a convolutional neural network approach [J]. IEEE Transactions on Neural Networks, 1997, 8(1):98-113.
- [13] LEA C, FLYNN M D, VIDAL R, et al. Temporal convolutional networks for action segmentation and detection [C]//IEEE Conference on Computer Vision and Pattern Recognition, 2017:1003-1012.
- [14] HE K M, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9):1904-1916.