

基于并行深度残差网络的堆场烟雾检测方法

王正来, 黄敏, 朱启兵, 蒋胜

江南大学轻工过程先进控制教育部重点实验室, 江苏 无锡 214122

摘要 堆场烟雾检测对于火灾预警、保障人员与财产安全具有重要意义。针对传统烟雾检测方法特征提取不充分, 误报率偏高以及稳健性较差的问题, 提出一种基于并行深度残差网络的堆场烟雾检测方法。该方法利用目标场景烟雾 RGB 图像的 R 、 G 、 B 分量以及图像 HSI 变换的 H 、 S 、 I 分量构建并行深度残差网络, 自适应获得烟雾特征; 同时通过样本扩边、负样本强化学习策略来加强模型对类烟物体的判别能力。实验结果表明, 该算法能有效降低因类烟物体产生的误报率, 且提升了网络的检出率和稳健性。

关键词 图像处理; 图像识别; 堆场; 烟雾检测; 并行深度残差网络

中图分类号 TP391.4

文献标识码 A

doi: 10.3788/LOP55.051008

Smoke Detection in Storage Yard Based on Parallel Deep Residual Network

Wang Zhenglai, Huang Min, Zhu Qibing, Jiang Sheng

Key Laboratory of Advanced Process Control for Light Industry, Ministry of Education,
Jiangnan University, Wuxi, Jiangsu 214122, China

Abstract Smoke detection in storage yard has great signification for fire early warning and protecting the safety of personnel and property. To solve the problem of insufficient features extraction, high false positive rates and poor robustness of traditional smoke detection methods, a new method of smoke detection in storage yard based on the parallel deep residual network is proposed. This method builds the parallel deep residual network with R , G , B components of the smoke RGB image and H , S , I components of the HSI transform image to adaptively extract the features. Meanwhile, the discriminant ability for the target like-smoke of the model is enhanced by the strategy including expanding the sample scale and reinforcement learning of the negative samples. The experimental results show that the proposed algorithm can effectively reduce the false positive rate caused by target like smoke and improve the detection rate and robustness of network.

Key words image processing; image recognition; storage yard; smoke detection; parallel deep residual network

OCIS codes 100.4996; 150.1135; 100.3008

1 引言

堆场的安全问题, 尤其是火灾预警一直是消防安全领域的一个重要课题。据公安部消防局统计, 2015年, 全国共接报火灾 33.8 万起, 造成 1742 人死亡、1112 人受伤, 直接财产损失达 39.5 亿元^[1]。火灾发生前期通常会产生烟雾, 如果能够及时地检测

到烟雾, 则可以提供更早的火灾预警, 减少人员伤亡和财产损失^[2]。现有的烟雾检测方法通常基于颜色、纹理、形状、运动等特征。Yu 等^[3]基于颜色和运动特征的烟雾检测方法, 首先通过背景提取和颜色确定候选烟雾区域, 然后计算 Lucas-Kanade 光流将烟雾与其他候选区域区分开来, 最后采用反向传播神经网络完成分类识别。Yuan^[4]通过构造三层多

收稿日期: 2017-10-17; 收到修改稿日期: 2017-11-23

基金项目: 江苏省研究生培养创新工程(SJLX16_0498)、江苏省政策引导类计划(产学研合作)-前瞻性联合研究项目(BY2016022-32)

作者简介: 王正来(1993—), 男, 硕士研究生, 主要从事检测与传感技术方面的研究。E-mail: wang930508@163.com

导师简介: 黄敏(1974—), 女, 博士, 教授, 主要从事图像处理方面的研究。E-mail: huangmzqb@163.com (通信联系人)

尺度的图像金字塔,然后利用基于变量的局部二值模式检测纹理。Ho^[5]利用烟雾区域的周长和面积之比来表达烟雾形状的复杂性,然后综合颜色规则、运动检测,利用模糊逻辑对烟雾进行识别。

尽管现有的方法在烟雾检测方面取得了一定的成果,但在堆场环境的实际应用中仍然存在一些问题,主要原因在于烟雾的静态特征如颜色、纹理等受天气、光照、时间等影响变化范围太大,而动态特征如背景建模或者帧间差分通常基于阈值,阈值对检测结果影响较大。烟雾检测目前所使用的方法对某些特定场合可能非常有效,而对其他场合效果不佳,没有提取出烟雾的本质特征^[6]。

近年来,深度卷积神经网络在人脸识别、图像分类等方面取得了突破性进展。深度卷积神经网络能以原始图像作为输入,通过组合浅层特征形成更加抽象的高层特征,以发现数据的更深层次分布式特征表示,避免人工提取特征工作的繁复冗杂^[7-8]。但常规的深度卷积神经网络在训练更高层的网络时,

易性能退化导致精度下降,深度残差网络(ResNet)的出现解决了这一问题。本文将在深度卷积神经网络基础上发展而来的深度残差网络引入到烟雾纹理特征提取,提出一种基于并行深度残差网络的烟雾检测方法,将图像的不同表达形式通过深度残差网络进行特征提取并合并的方式,以更全面地表达图像信息。实验结果显示,本文方法在烟雾识别的检出率、准确率,以及误检率上均表现良好。

2 面向堆场烟雾检测的并行深度残差网络

由于烟雾在不同环境、不同时间的变化差异很大,导致现有的检测方法误报率偏高,且通用性不好,对训练集中未出现的烟雾样本检出率偏低。为降低烟雾检测的误报率,提升对未知场景烟雾的检出率,本文从网络输入输出、样本构造,以及分类等方面进行改进,同时引入深度残差网络。本文方法的整体框架如图1所示。

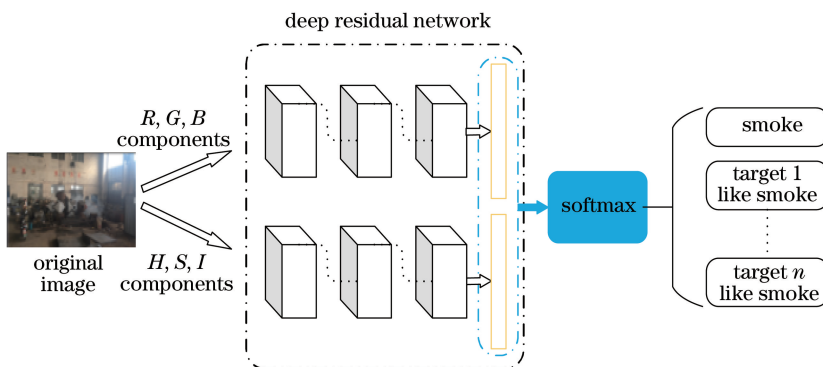


图1 本文方法的整体框架

Fig. 1 Overall framework of the method in this study

2.1 并行深度残差网络的输入及输出

现有的烟雾检测方法通常采用RGB图像,考虑到烟雾具有一定的与颜色相关的视觉特征,比如人眼区分烟雾与非烟雾物体通常会用到色调、亮度等信息,因此本文将图像的不同编码方式作为网络的输入,将网络学习到的特征结合,更全面地表达烟雾信息。Krstinic等^[9]通过实验发现RGB和HSI颜色空间是区别烟与非烟效果较好的颜色空间。因此本文选择RGB和HSI图像作为并行深度残差网络的输入,具体做法是将RGB和HSI图像分别输入到单路深度残差网络中,训练得到深度残差网络模型,并将两个单路深度残差网络模型获得的特征结

合,作为Softmax分类器的输入特征进行训练,得到烟雾识别结果。

传统的烟雾识别仅考虑烟雾和非烟雾两种类别(即网络的输出只有两类)。当检测环境存在天空中的云、静止的白色物体等类烟物体,网络通常会误报,如图2所示。卷积神经网络的检测结果受类别标签影响,误报的原因在于训练集样本中通常只标注正样本,网络分类器中通常只对正样本进行分类,忽略了负样本(比如类烟物体)对网络判别的干扰。为解决这一问题,在网络的分类器中增加负样本(类烟物体)分类,将相似的类烟物体给予相同的负样本标签,对差异较大的类烟物体给予不同的负样本标签。

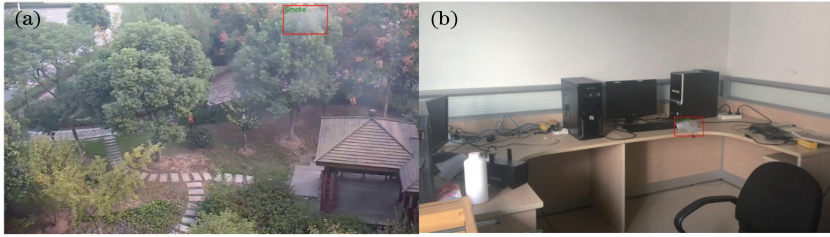


图 2 类烟物体示意图

Fig. 2 Schematic of the like-smoke object

2.2 深度残差网络的结构及学习算法

深度卷积神经网络的深度与训练样本数量存在着互相适应的关系,当训练样本数量增加时,网络的深度也需要增加,而常规的深度卷积神经网络通常采用平原网络,训练层数更多的网络易产生性能退化问题,即网络的准确度在达到饱和后迅速下降^[10-12]。研究结果表明,网络性能的退化并不是由过度拟合造成的^[13-14],其原因在于常规的深度卷积神经网络难以通过多层网络拟合出所需的同等函数。深度残差网络的出现,在保留深度卷积神经网络自适应提取特征优点的同时,较好地解决了网络性能退化问题,该网络通过增加一个近似同等函数层,减小拟合的难度,减少计算量^[15]。

2.2.1 深度残差网络的结构

采用 ResNet 50 作为网络的基本结构,其各层参数配置如表 1 所示。ResNet 50 采用多个残差块串联操作的形式实现图像特征的提取。残差块示例如图 3 所示。残差块以 $1 \times 1 + 3 \times 3 + 1 \times 1$ 卷积层的形式代替常规网络的 2 个 3×3 的卷积层。相比于常规网络,采用残差块结构,可在保持网络精度的同时,减少网络计算量。以 conv3_x 为例,输入为 $256 \times 56 \times 56$ (分别为维度 d 、宽 w 、高 h) 的特征图,首先通过 $128 \times 1 \times 1$ 的卷积层降维为 $128 \times 56 \times 56$,然后通过 $128 \times 3 \times 3$ 的卷积层,最后通过 1×1 的卷积层升维。

表 1 ResNet 50 网络结构

Table 1 Network structure of ResNet 50

Layer name	Output size	50-layer
conv1	112×112	74×74 , 64, stride2
		3×3 max pool, stride2
conv2_x	56×56	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 2$, stride1
		$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 1$, stride2
conv3_x	28×28	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 3$, stride1
		$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 1$, stride2
conv4_x	14×14	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 5$, stride1
		$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 1$, stride2
conv5_x	7×7	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$, stride1
	1×1	Average pool, 1000-d fc, softmax

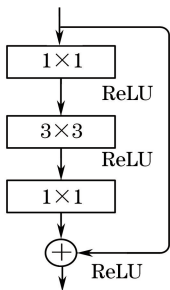


图3 残差块

Fig. 3 Residual block

2.2.2 深度残差网络的学习算法

残差概念的引入基于如下假设:如果能够以恒等映射的方式来构建所增加的层,那么一个深层网络则退化为浅层网络,至少深层网络的训练误差不会大于它所基于的较浅模型。如图4所示,假设期望输出是一个理想映射 $H(x)$, 平原网络中网络学习的则是 $H(x)$, 其期望权重层能与 $H(x)$ 拟合即学习到恒等变换, 网络输出为

$$y = w_{c3} \sigma[w_{c2} \sigma(w_{c1} x)]. \quad (1)$$

而在残差网络中, x 通过恒等映射后与网络输出合并, 网络输出为

$$y = w_{c3} \sigma[w_{c2} \sigma(w_{c1} x)] + w_s x. \quad (2)$$

残差网络学习的是残差函数 $F(x)$, 它是一个残差映射, 比恒等变换更容易拟合。

$$F(x) = H(x) - w_s x, \quad (3)$$

式中 w_{c1} 、 w_{c2} 、 w_{c3} 分别为各层权重, σ 为修正线性单元(ReLU)函数, 并且为了简化符号省略了偏置。线性投影 w_s 用以解决维数匹配的问题, 在 F 与 x 维数相同时, $w_s = 1$, 维数不同时, 通过 1×1 的卷积层来匹配维数。

2.3 样本构造

为进一步提高网络的性能, 减少对负样本的误报率, 本文从样本扩边以及类烟负样本强化学习两个方面进行改进。

1) 样本扩边。烟雾具有模糊背景的特性, 即烟



图5 部分实验视频

Fig. 5 Part of the experimental videos

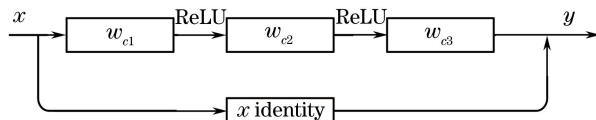


图4 残差网络

Fig. 4 Residual network

雾与背景交界不明显。常规的深度卷积神经网络通常忽视这一点, 本文为了有效利用这一特征, 将训练集中所有正负样本标注的真实框采取 1.2 倍的扩边操作后进行训练, 使网络能够学习烟雾边缘模糊的特征, 增强模型的判别能力。

2) 类烟负样本强化学习。在训练过程中, 用于训练的类烟负样本数量往往偏少, 导致网络对这类样本的识别能力不够。本文通过复制该类负样本以及仿射变换制作负样本的方式, 扩充训练集, 加强网络对这些负样本的学习。

通过上述方法, 网络能有效地区分烟雾与类烟物体, 误报率明显降低。

3 实验及分析

3.1 样本选取及实验环境

本实验所用视频来源于网络及实验拍摄, 选取的视频为各种场景下的烟雾视频及非烟雾视频, 部分实验视频如图5所示。由于样本来源复杂, 其原始尺寸不同, 本文统一调整为 $600 \text{ pixel} \times 1000 \text{ pixel}$, 作为网络的输入。训练图片数为 14180 张含烟雾图片, 计算机配置为 Intel(R) Core(TM) i7-6700HQ CPU 2.60 GHz, 显卡为 NVIDIA GeForce GTX 1060。网络参数更新方法均为随机梯度下降法(SGD), 学习率均为 0.001, 训练模型的最大迭代次数为 100000 次。现有的烟雾检测方法通常采用 RGB 图像, 因此本文中单路深度残差网络除输入 RGB 图像外, 其网络参数设置与并行网络一致。

测试集从 29 个不同场景挑选 1243 张图像进行测试, 包含 1124 张含烟雾图像和 119 张无烟雾

图像,其中 22 个场景用于验证实验,7 个用于测试实验。验证实验为已知场景(场景部分样本参与网络训练),从 22 个验证场景中挑选 1032 张未参与网络训练的图片进行测试,比较并行网络与单路网络的检测效果。测试实验为未知场景(场景所有样本均未参与网络训练),从 7 个测试场景挑选 211 张图片进行测试,比较并行网络与单路网络的稳健性以及漏报、误报情况,更全面地评估模型性能。现有的烟雾检测方法通常采用 RGB 图像,因此本文中单路深度残差网络除输入 RGB 图像外,其余设置与并行网络一致。每帧图像的平均检测时间为 200 ms,基本满足堆场实时烟雾检测的需求。

3.2 置信值及评价指标

对检测到的目标,分类器会给出目标类别以及是这个类别的概率,这个概率称之为置信值。鉴于堆场火灾预防的重要性,即使置信值较低的检测目标也可能是真实烟雾,因此需要考虑在内。本文将

置信值分为以下几级:0.1、0.2、0.3、0.4、0.5、0.6、0.7、0.8、0.9、0.95、0.98。

实验的指标为检出率、误报率,以及受试者工作特征(ROC)曲线,由于实际检测中仅考虑真实烟雾是否检出,不需要 True Negative 值,所以检出率与漏报率之和为 1。模型的漏报率和误报率越低越好,故 ROC 曲线图改为不同置信值下漏报率与误报率的变化情况,曲线越接近坐标原点说明模型的性能越好。烟雾检出率、误报率和漏报率公式分别为

$$R_{\text{detection}} = N_{\text{detection}} / N_{\text{exist}}, \quad (4)$$

$$R_{\text{positive}} = N_{\text{positive}} / N_{\text{total}}, \quad (5)$$

$$R_{\text{negative}} = 1 - R_{\text{detection}}, \quad (6)$$

式中 $N_{\text{detection}}$ 是检出烟雾张数, N_{exist} 是存在烟雾图片张数, N_{positive} 是误报张数, N_{total} 是总测试样本数。

3.3 实验结果

部分检测结果如图 6 所示。

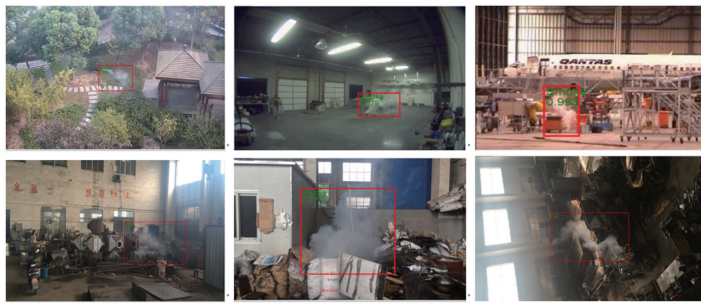


图 6 部分测试结果

Fig. 6 Part of the test results

3.3.1 验证实验

将本文方法与单路深度残差网络在 22 个验证场景下进行实验比对,结果如表 2 及图 7 所示。结果表明,除 0.9 置信值外,并行网络在其余各级的

检出率较单路网络均有小幅提升,同时在误报率上均优于单路网络。从曲线图中可以看出网络的训练效果接近,但并行网络的曲线更接近坐标原点。

表 2 验证场景单路网络与并行网络的检出率和误报率

Table 2 Detection rate and false positive rate of validation scenarios with single and parallel network %

Confidence level	Single network		Parallel network	
	Detection rate	False positive rate	Detection rate	False positive rate
0.1	93.764	1.357	95.768	0.581
0.2	92.650	0.678	94.655	0.194
0.3	91.759	0.581	92.984	0.097
0.4	90.646	0.484	91.314	0.000
0.5	89.978	0.291	90.535	0.000
0.6	88.864	0.097	89.421	0.000
0.7	87.751	0.097	88.085	0.000
0.8	85.412	0.000	85.746	0.000
0.9	81.849	0.000	81.849	0.000
0.95	78.953	0.000	79.065	0.000
0.98	73.385	0.000	73.608	0.000

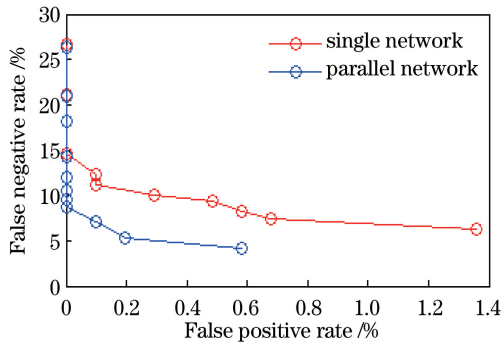


图7 验证场景并行网络与单路网络的ROC曲线图

Fig. 7 ROC curves of validation scenarios with parallel and single network

3.3.2 测试实验

将本文方法与单路深度残差网络在7个测试场

表3 测试场景单路网络与并行网络的检出率和误报率

Table 3 Detection rate and false positive rate of test scenarios with single and parallel network

%

Confidence level	Single network		Parallel network	
	Detection rate	False positive rate	Detection rate	False positive rate
0.1	43.363	26.540	69.912	36.967
0.2	37.168	15.640	65.487	16.588
0.3	34.513	15.640	61.504	13.744
0.4	31.858	15.640	58.850	12.322
0.5	30.973	15.166	56.637	11.848
0.6	29.204	13.744	54.867	10.427
0.7	27.876	12.322	52.212	8.531
0.8	26.549	9.953	49.115	7.109
0.9	24.779	0.474	45.575	0.000
0.95	19.027	0.000	38.496	0.000
0.98	12.389	0.000	26.549	0.000

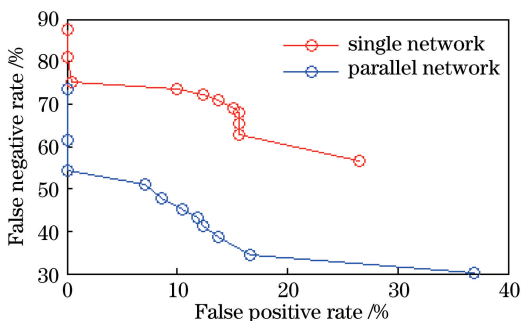


图8 测试场景并行网络与单路网络的ROC曲线图

Fig. 8 ROC curves of test scenarios with parallel and single network

4 结 论

深度学习在人工智能领域取得重要的突破,尤其在语音识别、自然语言处理、计算机视觉、图像与视频分析等领域均取得了巨大成功。为此提出了一

种基于并行深度残差网络的堆场烟雾检测方法,通过多形式输入构建并行深度残差网络模型,自适应得到烟雾特征,同时对网络输入、样本构造及分类等方面进行了改进。实验结果验证了本文方法的可行性和实时性,不仅提升了烟雾检出率,而且降低了误报率,提升了模型的稳健性。

景下进行实验比对,结果如表3及图8所示,结果表明,除在0.1置信值上并行网络的误报率高于单路网络,其余情况下并行网络在误报率与漏报率上的表现均优于单路网络,同时从ROC曲线图中可以看出,本文方法训练的网络曲线更接近于坐标原点。综上所述,在训练样本完全一致的情况下,改进后的并行网络模型的训练效果比单路网络好,同时模型性能远优于单路网络,尤其是在未知场景下,大幅度提升检出率的同时降低了误报率。从数据可以看出,无论是并行网络还是单路网络,在未知场景下检出率并不是十分理想,说明基于深度卷积神经网络的方法同样具有缺陷,模型的深度与训练样本数量为匹配关系,存在着欠拟合、过拟合等问题,本文网络模型的训练还有提升的空间。

参 考 文 献

- [1] Huang Z H. Brief discussion unit fire safety training [J]. Marine Fire, 2016, 37(5): 15-17.
黄祖华. 浅议单位消防安全培训 [J]. 水上消防, 2016, 37(5): 15-17.
- [2] Jiang M X, Wang H Y, Cai X Y. Early smoke detection based on codebook model and multiple features [J]. Journal of Image and Graphics, 2012, 17(9): 1102-1108.
姜明新, 王洪玉, 蔡兴洋. 基于码本模型和多特征的

- 早期烟雾检测[J]. 中国图象图形学报, 2012, 17(9): 1102-1108.
- [3] Yu C Y, Fang J, Wang J J, *et al.* Video fire smoke detection using motion and color features[J]. Fire Technology, 2010, 46(3): 651-663.
- [4] Yuan F N. Video-based smoke detection with histogram sequence of LBP and LBPV pyramids[J]. Fire Safety Journal, 2011, 46(3): 132-139.
- [5] Ho C C. Machine vision-based real-time early flame and smoke detection[J]. Measurement Science and Technology, 2009, 20(4): 045502.
- [6] Luo S, Jiang Y Z. State-of-art of video based smoke detection algorithms [J]. Journal of Image and Graphics, 2013, 18(10): 1225-1236.
- 罗胜, Jiang Yuzheng. 视频检测烟雾的研究现状[J]. 中国图象图形学报, 2013, 18(10): 1225-1236.
- [7] Liu F, Liu P Y, Li B, *et al.* Deep learning model design of video target tracking based on TensorFlow platform [J]. Laser & Optoelectronics Progress, 2017, 54(9): 091501.
- 刘帆, 刘鹏远, 李兵, 等. TensorFlow 平台下的视频目标跟踪深度学习模型设计[J]. 激光与光电子学进展, 2017, 54(9): 091501.
- [8] Ye G L, Sun S Y, Gao K J, *et al.* Nighttime pedestrian detection based on faster region convolution neural network [J]. Laser & Optoelectronics Progress, 2017, 54(8): 081003.
- 叶国林, 孙韶媛, 高凯珺, 等. 基于加速区域卷积神经网络的夜间行人检测研究[J]. 激光与光电子学进展, 2017, 54(8): 081003.
- [9] Krstinić D, Stipanićev D, Jakovčević T. Histogram-based smoke segmentation in forest fire detection system[J]. Information Technology and Control, 2009, 38(3): 237-244.
- [10] Bengio Y, Simard P, Frasconi P. Learning long-term dependencies with gradient descent is difficult [J]. IEEE Transactions on Neural Networks, 1994, 5(2): 157-166.
- [11] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks [J]. Journal of Machine Learning Research, 2010, 11(9): 249-256.
- [12] He K M, Zhang X Y, Ren S Q, *et al.* Delving deep into rectifiers: surpassing human-level performance on ImageNet classification [C]. 2015 IEEE International Conference on Computer Vision (ICCV), 2015: 1026-1034.
- [13] LeCun Y A, Bottou L, Orr G B, *et al.* Efficient backprop[J]. Neural Networks: Tricks of the Trade, 2012: 9-48.
- [14] He K M, Sun J. Convolutional neural networks at constrained time cost[C]. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015: 5353-5360.
- [15] 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), 2016: 770-778.