

基于改进的卷积神经网络的道路井盖缺陷检测研究

姚明海, 隆学斌

(浙江工业大学 信息工程学院, 杭州 310023)

摘要: 道路井盖缺陷检测对于道路维护与安全至关重要, 论文提出了一种改进的卷积神经网络算法, 可实现井盖缺陷的快速、准确检测; 算法对卷积神经网络的激活函数模型进行了改进, 针对 Relu 激活函数在输入小于零时输出设为零, 导致部分缺陷信息丢失问题, 设计了 MReLU 和 BReLU 两种改进激活函数; 在此基础上, 为了增强神经网络模型的特征表达能力, 提出了双层激活函数模型; 最后, 在公共数据集 MNIST, CIFAR-10 上进行了比较实验, 网络主要参数有批处理大小 (batch size) 为 32, 最大迭代次数为 1 000 次, 学习率为 0.000 1, 每经过 5 000 次迭代衰减 50%; 实验结果表明, 基于改进后的激活函数和应用双层激活函数所构造的卷积神经网络, 大大减少了训练参数, 不仅收敛速度更快, 而且可以更加有效地提高分类的准确率。

关键词: 井盖缺陷; 卷积神经网络; 激活函数; 神经元

Research on Manhole Cover Detection Using Improved Convolutional Neural Network

Yao Minghai, Long Xuebin

(College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China)

Abstract: The defect detection of road manhole cover is very important for road maintenance and safety. The paper proposes an improved convolutional neural network algorithm to achieve rapid and accurate detection of manhole cover defects. The algorithm improves the activation function model of convolutional neural network. For the Relu activation function, when the input is less than zero, the output is set to zero, which results in losing most of the input information. Therefore, two improved activation functions, MReLU and BReLU, are designed. On this basis, in order to enhance the feature expression ability of neural network model, a two-layer activation function model is proposed. Finally, a large number of comparative experiments were performed on the proposed algorithm in the public data set MNIST, CIFAR-10, and the main parameters of the network are batch size of 32, the maximum number of iterations is 1000, the learning rate is 0.000 1, and the attenuation is 50% after 5 000 iterations. The experimental results show that the convolutional neural network based on the improved activation function and the application of the two-layer activation function greatly reduces the training parameters, not only the convergence speed is faster, but also can improve the classification accuracy more effectively.

Keywords: manhole cover detection; convolutional neural networks; activate function; neuron

0 引言

道路井盖是现代城市建设和交通管理的必需设备, 井盖的缺失和损坏会造成严重的交通事故和经济损失。所以道路井盖的管理变得愈发重要。然而, 道路井盖妥善的管理必须建立在精确的井盖缺陷检测的基础上。人工排查丢失或者有破损的井盖会产生诸如效率低、耗时多、精确度不足等问题, 而且由于城市道路整改多, 人工检查车道以及草丛内等地方的井盖非常不安全, 因此人工排查显然不是一个好的井盖缺陷检测选择, 实现对井盖状态的自动检测具有重要意义。为了克服人工检测带来的这些问题以达到更安全、更快速、更精确地进行井盖缺陷检测的目的, 一些研究人员在传感器和图像处理方法上做出了一些研究。

比如, 文献 [1] 提出了基于无线传输、电机驱动井盖锁结构技术来实现对丢失井盖的动态检测的方法; 文献 [2] 通过提取井盖的颜色、边缘轮廓等特征, 并使用改进的霍夫变换方法, 计算窨井盖的区域范围, 进而判断是否有破损或丢失的情况。通过分析这些文献, 可以发现, 基于传感器和图像处理的方法虽然能实现井盖的监测, 但算法的性能不够好, 计算量较大, 对破损且有安全隐患的井盖无法实现自动检测, 难以应用到实际检测当中。由于深度学习在特征提取上有显著的优势, 基于传感器和图像处理的井盖缺陷识别方法慢慢被卷积神经网络所取代。文献 [3] 提出了基于 Faster R-CNN^[4] 的卷积神经网络模型对井盖进行检测的方法。该方法证明, 卷积神经网络跟之前基于传感器和图像处理的方法相比, 在准确率和检测效率上确实取得了飞跃性的进步, 但如果在改进激活函数和网络结构上做进一步改进的话, 性能会更好, 这也就是本文的主要研究内容。

在激活函数角度上, 目前使用最多的是 ReLU 激活函

收稿日期: 2019-06-21; 修回日期: 2019-07-19。

基金项目: 国家自然科学基金项目(61871350)。

作者简介: 姚明海(1963-), 男, 浙江嘉善人, 教授, 博士生导师, 主要从事模式识别、图像识别方向的研究。

数。一方面,在一定程度上克服了 sigmoid 和 tanh 激活函数的梯度弥散的缺点,为训练后的网络带来了适度的稀疏特性,而且由于 ReLu 表达式中没有除法和指数运算,因此计算速度更快。但是,另一方面,ReLu 激活函数也没有真正解决输入小于零时输出也为零的缺点,同样也在一定程度上导致了神经元信息的丢失,特别是当输入很多为负值时。因此,本文提出一种基于 ReLu 改进的激活函数 MReLu 和 BReLu,其克服了 ReLu 的信息丢失的缺点,保留更多的输入为负值的信息并且能够传输更多信息到下一个网络层。在网络结构角度上,之前的神经网络多数是只用一个激活函数,不能最大限度地表达神经元的非线性特性。因此,本文运用两个激活函数对卷积神经网络模型进行优化。实验表明,本文提出的方法跟现有方法相比,不仅具有更高的识别准确率,而且同时提高了收敛速度和鲁棒性,并且可以应用于各种深度神经网络体系结构中。

1 涉及相关算法

1.1 MCCNN 神经网络模型

改进的卷积神经网络是专门用于道路井盖缺陷检测的卷积神经网络模型 (manhole cover detection convolutional neural network, MCCNN)。如图 1 所示,本网络结构由 3 个卷积层,两个最大池化层以及一个全连接层加上 softmax 层组成。从图可知,第一个网络层接收输入井盖图片,然后网络层 1 和 2 执行卷积操作,并经过激活函数将神经元特征激活,然后再通过最大池化层降低参数数量;最后将经过只有卷积操作的网络层 3 输出进行全连接,综合前几层采集到的井盖特征,经过 softmax 层将全连接层的输出映射成概率分布,用于井盖识别的分类。本文以非线性激活函数为唯一变量,来研究不同激活函数对井盖识别快速性和准确率的影响。

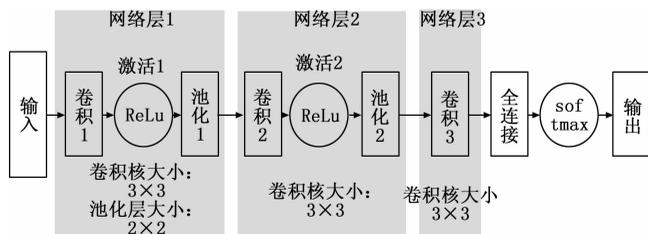


图 1 神经网络模型 MCCN 结构图

1.2 卷积神经网络和激活函数

卷积神经网络 (convolutional neural network, CNN)^[5]提供了一种端到端的学习模型,模型中的参数可以通过反向传播中的梯度下降算法进行反复训练以使输出结果接近真实值。经过训练的卷积神经网络能够学习到图像中的特征,比如纹理特征、颜色特征、形状特征和空间特征等,并完成对图像特征的提取和分类。作为深度学习的一个重要研究分支,卷积神经网络的特点在于每一层的特征都由上一层部分区域通过权值共享的卷积核以及激活函数激励得到。这一特点使卷积神经网络在图像缺陷识别^[6]的应用

上更优于其他神经网络方法。随着卷积神经网络的流行,一些改进算法也不断出现,如有基于自适应学习率的研究^[7],有将 DrouOut 层运用于池化阶段的方法^[8],有致力于预训练与随机初始化模型^[9],这些方法都在一定程度上提高了卷积神经网络的表达能力。本文主要研究在道路井盖缺陷识别问题上,激活函数的优化对提高分类准确率和降低损失率的影响。激活函数主要作用是提供网络的非线性建模能力,如果网络中只包含线性卷积和全连接运算,即使网络层数再多也仅能表达线性映射,难以有效解决实际环境中的非线性分布的数据。当激活函数应用于卷积神经网络时,主要在前向传播和反向传播两个过程中对网络参数的训练产生影响和作用。

1.3 ReLu 激活函数

因为曾经一度被广泛使用的经典激活函数 Sigmoid 和 Tanh 都具有致命的梯度弥散的缺点,所以当前在神经网络模型中激活函数使用的趋势就是非饱和修正线性函数 ReLu。修正线性单元 (rectified linear unit, ReLu) 是这样一种激活函数,如果输入的值小于或者等于零,那么输出就为零;如果输入大于零,那么就会保持原值不变,原样输出。其用数学表达式定义为: $y(x) = \max(0, x)$,如图 2 所示。ReLu 这种直接地强制某些数据为 0 的做法,在一定程度上,为训练后的网络带来了适度的稀疏特性。它不但减少了参数之间的相互依存关系,而且缓解了过拟合问题的发生。相比较传统的 S 型激活函数,ReLu 不含除法和指数运算,计算速度更快。但是,经过卷积操作的输出值多数都会产生大量的负值,如果所有的负值在经过修正线性单元后都被取代为零的话,就会丢失大量的输入信息,对特征提取将产生很大的不良影响,最后导致识别准确率下降。

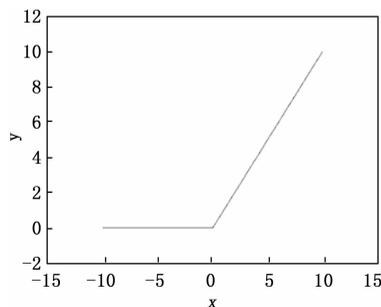


图 2 ReLu 函数

1.4 双边 ReLu 激活函数

在近期的卷积神经网络的研究中,多种基于 ReLu 的激活函数不断涌现。其中,双边修正线性单元 (Bilateral ReLU) 就是一个典型的代表,该激活函数被应用于 DehazeNet 网络模型^[10],用来增强模糊的原始图片。

2 改进算法

在本文的研究中,我们将基于改进的修正线性单元和改进的双边修正线性单元来改善卷积神经网络结构 MCCNN 性能。

2.1 改进的修正线性单元

改进的修正线性单元激活函 (Modified ReLu, MReLu) 的数学表达式 (1) 如下:

$$f(y) = \max(t_{min}, y) \tag{1}$$

其中: y 表示输入值, $f(y)$ 表示输出, t_{min} 在 MReLu 激活函数中是常量。MReLu 激活函数的如图 3 所示。由图可知, 当输入 $f(y) \geq 0$ 时, 输出保留输入的值不变; 当输入 $t_{min} < f(y) < 0$ 时, 输出也保持该负值不变; 当输入 $f(y) < t_{min}$ 时, 输出为 t_{min} 。对比 ReLu 可以看出, 该激活函数克服了 ReLu 因为梯度弥散导致的信息严重丢失的缺陷。

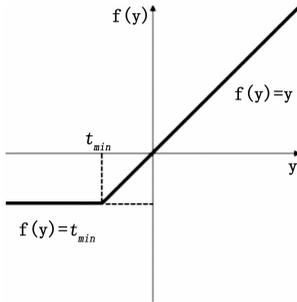


图 3 MReLU 激活函数

神经网络模型 MCCNN 在网络层 2 的卷积操作后使用 MReLU 作为其激活函数示意图如图 4 所示。

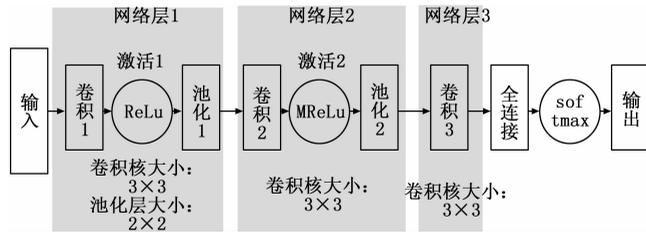


图 4 使用 MReLU 激活函数的网络结构图

2.2 双边修正线性单元

双边修正线性单元 (Bilateral Rectified ReLu, BReLU) 可以用下面的数学等式 (2) 来描述:

$$f(y) = \min(t_{max}, \max(t_{min}, y)) \tag{2}$$

其中: y 表示图片输入, $f(y)$ 表示输出, t_{min} 和 t_{max} 是 BReLU 激活函数的边缘常量值。BReLU 激活函数如图 5 所示。由图可知, 当 $t_{min}=0$ 且 t_{max} 不存在时, 就可以退化成正线性单元 ReLu; 在 BReLU 激活函数中, t_{min} 和 t_{max} 都是常量, 且相应地, $t_{min} < 0, t_{max} > 1$ 。

相应地, 神经网络模型 MCCNN 在网络层 2 的卷积操作后使用 BReLU 作为其激活函数示意图如图 6 所示。

3 实验结果与分析

本文中, 我们两类不同的实验。一个是在不同数据集上对提出的改进的激活函数的对比验证; 另一个是利用第一个实验的验证结果, 将改进的激活函数放到本研究所要解决的道路井盖识别的具体应用上。两个实验都是基于主流框架 TensorFlow^[11], 实验设备是一台基于 Tesla K40C

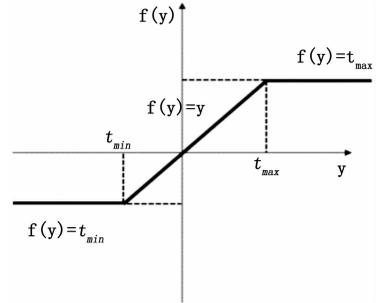


图 5 BReLU 激活函数图

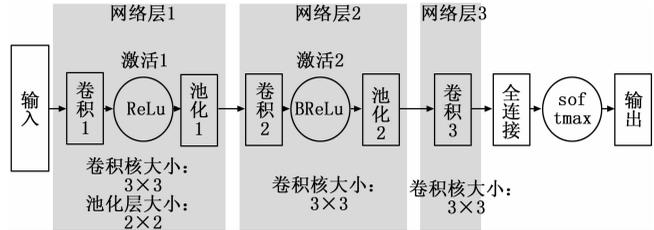


图 6 使用 BReLU 激活函数的网络结构图

的服务器, 具体设备参数如表 1 所示。并且训练时使用 Xavier 初始化网络参数, 将 AdamOptimizer^[12] 作为优化器的优化算法, softmax^[13] 交叉熵损失函数作为损失函数, 网络训练的其他参数还有: 批处理大小 (batch size) 为 32, 输入井盖尺寸为 208×208 , 最大迭代次数为 1 000 次, 学习率为 0.000 1, 每经过 5 000 次迭代衰减 50%。表 1 为实验硬件参数。

表 1 硬件参数

硬件环境	CPU Intel(R) Xeon(R) CPU E5-2609 Gpu NVIDIA Tesla K40c
内存	128G
操作系统环境	Ubuntu 14.04.5 LTS

3.1 不同公共数据集上的验证

本实验中, 我们将基于 1.1 所述的 MCCNN 的神经网络模型, 以激活函数为唯一变量来在两个典型的数 (其中 t_{min} 取 -0.25 , t_{max} 取 2.0) 数据集 MNIST, CIFAR-10 上做对比实验。

3.1.1 MNIST 数据集

采用的 MNIST 数据集^[14] 是由美国中学生手写阿拉伯数据集, 共有 60 000 个 28×28 维的 0-9 的灰度手写数字图片, 其中 50 000 个样本作为训练集, 10 000 个样本作为测试集。在该数据集上, 应用图 1 所示的 MCCNN 神经网络作为训练模型, 以 ReLu, MReLU, BReLU 分别作为第双层网络的激活函数来验证不同激活函数的性能。为了观察采用不同激活函数的网络表现出的性能, 对每次实验中的测试精度和损失函数进行表统计。表 2~3 分别代表激活函数与迭代次数对应的损失函数与精度统计。之所以要对损失函数进行统计是因为损失函数是衡量数据属于某个类别的概率, 损失函数越小说明网络收敛越快, 而测试精度则是直观反映网络性能的量。

表 2 MNIST 数据集上激活函数与迭代次数对应的损失函数值(Loss)

迭代次数(次)	1 000	2 000	3 000	40 000	50 000	60 000	70 000	80 000	90 000	10 000
ReLU	0.057 3	0.041 4	0.034 8	0.030 5	0.030 2	0.027 8	0.028 5	0.028 9	0.028 6	0.028 3
MReLU	0.067 7	0.045 0	0.040 2	0.033 1	0.034 9	0.027 6	0.033 2	0.025 5	0.028 2	0.027 7
BReLU	0.060 5	0.050 3	0.042 2	0.032 1	0.031 6	0.028 8	0.027 6	0.021 5	0.021 1	0.020 9

表 3 MNIST 数据集上激活函数与迭代次数对应的损失函数值(Accuracy/%)

迭代次数(次)	1 000	2 000	3 000	40 000	50 000	60 000	70 000	80 000	90 000	10 000
ReLU	0.950 4	0.957 7	0.958 6	0.959 5	0.960 3	0.961 0	0.962 2	0.970 3	0.972 7	0.975 4
MReLU	0.950 6	0.954 4	0.955 6	0.957 5	0.960 1	0.961 0	0.971 2	0.976 5	0.982 2	0.985 5
BReLU	0.951 6	0.954 5	0.956 5	0.957 7	0.969 9	0.961 4	0.970 9	0.977 6	0.982 1	0.982 5

从表 3 出, 采用 MReLU 激活函数的网络取得了最高的准确率 0.985 5, BReLU 次之, 为 0.982 5, 而 ReLU 为 0.975 4。说明改进的激活函数不仅解决了梯度消失问题, 还大大提高了网络的识别率。同时, 从表 2 中可以发现, MReLU 不仅是精度最高的, 而且是收敛速度最快的, 这是因为 MReLU 的输出更接近 0 均值, SGD 更接近 Natural gradient。因此可知, 我们改进的激活函数有更高的精度和更少的损失, 呈现出更好的性能。

3.1.2 CIFAR-10 数据集

作为本实验的第二个数据集, 我们考虑使用 CIFAR-10 数据集^[15]。它是由 CIFAR (Canadian Institute For Advanced Research) 收集整理的一个用于机器学习和图像识别的数据集。该数据集包含 60 000 个 32×32 的彩色图像, 共有 10 类。其中 50 000 个训练图像和 10 000 个测试图像。实验采用的参数与网络模型与 MNIST 数据集一样。同理, 为了观察采用不同激活函数的网络表现出的性能, 对每次实验中的测试精度和损失函数进行表统计, 并且作图以直观显示。表 4~5 分代表激活函数与迭代次数对应的损失函数与精度统计。

从表 5 看出, 采用 MReLU 激活函数的网络取得了最高的准确率 0.988 5, BReLU 次之, 为 0.986 5。这两者与准确率为 0.975 4 的 ReLU 激活函数相比, 有了比较明显的提升。说明改进的激活函数不仅解决了梯度消失问题, 还

大大提高了网络的识别准确率。同时, 从表 4 中可以发现, BReLU 与 MReLU 的收敛速度差不多一样快, 比 ReLU 激活函数快一些。因此可知, 我们改进的激活函数有更高的精度和更少的损失, 呈现出更好的性能。

3.2 井盖数据集上的验证

经过上面两种公共数据集的验证, 我们发现改进的激活函数确实有更好的性能。因此, 为了进一步证明本结论, 我们将其应用到本文要解决的井盖识别问题的应用上。

3.2.1 实验数据

本文训练的样本来源于定期从城市视频监控抓取的视频流集, 共有 1 500 路视频, 将视频分解成帧图片并经过筛选后, 采集其中的 5 000 张图片, 包含有完整井盖、破损井盖和井盖丢失 3 种图片。现将其中 4 000 张图片作为训练集, 剩下 1 000 张作为测试集, 分类结果分为 A、B、C 三类, 分别代表完整井盖、破损井盖和丢失井盖。图片像素作为输入进入 MCCNN 神经网络后统一尺寸为 208×208 。如图 7 所示。

3.2.2 训练和识别过程

将训练集中的实验数据取出, 然后进行人工标注, 分为完整井盖, 井盖缺失, 井盖破损三类。然后将标注好的尺寸为 208×208 的训练集数据输入到本文中 MReLU、MReLU 和 BReLU 作为唯一变量的 MCNN 神经网络, 然后经过 5 000 次正向传播和反向传播的反复迭代来进行网络参

表 4 CIFAR-10 数据集上激活函数与迭代次数对应的损失函数值(Loss)

迭代次数(次)	1 000	2 000	3 000	40 000	50 000	60 000	70 000	80 000	90 000	10 000
ReLU	0.057 3	0.041 4	0.034 8	0.030 5	0.030 2	0.027 8	0.028 5	0.028 9	0.028 6	0.028 3
MReLU	0.067 7	0.045 0	0.040 2	0.033 1	0.034 9	0.027 6	0.033 2	0.025 5	0.028 2	0.027 7
BReLU	0.060 5	0.050 3	0.042 2	0.032 1	0.031 6	0.028 8	0.027 6	0.021 5	0.021 1	0.020 9

表 5 CIFAR-10 数据集上激活函数与迭代次数对应的损失函数值(Accuracy/%)

迭代次数(次)	1 000	2 000	3 000	40 000	50 000	60 000	70 000	80 000	90 000	10 000
ReLU	0.951 1	0.952 8	0.957 6	0.960 1	0.960 3	0.962 0	0.963 1	0.970 3	0.971 7	0.975 4
MReLU	0.951 0	0.953 4	0.954 5	0.957 5	0.961 1	0.962 0	0.971 3	0.976 6	0.981 2	0.988 5
BReLU	0.941 1	0.953 5	0.954 5	0.958 7	0.961 9	0.962 1	0.970 9	0.972 6	0.982 1	0.986 5



图 7 数据集内样本分类

数训练, 最后得出分类结果。其中, 正向传播的作用是提取井盖特征, 反向传播的作用是通过梯度下降的优化算法来减少损失函数, 进而让分类的结果接近真实值。

3.2.3 实验结果分析

将井盖图片进行训练后, 3 种不同井盖类别的识别结果如表 6 所示, 整体井盖识别结果如表 7 所示, 包括识别准确率和损失率。

表 6 不同井盖类型识别结果

	ReLU		MReLU		BReLU	
	精度	损失	精度	损失	精度	损失
完整	97.07	2.80	98.16	2.70	98.55	2.68
破损	95.22	4.15	96.15	3.11	97.01	3.08
丢失	99.22	2.23	99.55	2.10	99.80	2.21

整体识别精度与损失由表 7 统计如下:

表 7 井盖识别结果

	ReLU	MReLU	BReLU
Accuracy(%)	96.56	98.28	98.56
Loss	0.56	0.37	0.29

根据表 6 第二、四、六列的第四、五行数据, 可以看出该城市中井盖出现破损比井盖丢失的情况更多; 由表 7 激活函数 ReLu、MReLU、BReLU 对识别结果的准确率和损失率可知, 利用改进的激活函数检测到的井盖丢失或者破损的精度会比 ReLu 激活函数更高, 且收敛速度更快。由于井盖丢失时特征比较明显, 很容易提取, 比如出现大面积空洞, 而井盖破损程度各异, 大多数没有明显的共同特征, 所以检测效果在 ReLu、MReLU 以及 BReLU 作为激活函数都会相比井盖丢失情况效果稍差; 完整井盖的检测精度是这 3 个识别类别中精度最高的, 特别是应用改进的激活函数, 其中 BReLU 达到 98.56%, MReLU 达到 98.28%, 比 ReLu 的 96.56% 高出许多。由此可见, 改进的激活函数改善了井盖缺陷检测的精度与性能。

4 结束语

本文中, 提出了在道路井盖缺陷识别的应用中, 通过用改进的激活函数 MReLU 和 BReLU 来替代修正单元 ReLu 来作为非线性激活函数。实验证明, 改进的激活函数不仅能够提高井盖缺陷识别的精度, 而且还能加快收敛速度。但算法仍有不足之处, 例如检测过程中出现的漏检、误检, 仍不可避免; 如何改进样本质量, 改变网络参数 t_{min} 和 t_{max} 以再提高检测精度; 此外, 此算法还有一个限制, 就是在网络层 1 和网络层 2 的激活函数不能同时使用 MReLU 或者 BReLU, 因为第二、双层的卷积操作输出值包含大量不同的特征信息。在未来的研究中, 将进一步寻找解决这些问题的方法, 以到达改善道路井盖识别的性能。

参考文献:

- [1] 解兆延, 徐文青, 张延波, 等. 基于 ZigBee 技术的防盗井盖控制系统的设计与实现 [J]. 电子世界, 2014 (17): 145.
- [2] 张丰焰, 陈荣保, 李扬, 等. 基于改进 Hough 与图像比对法的窨井盖疑似破裂检测 (英文) [J]. Journal of Southeast University (English Edition), 2015, 31 (4): 553-558.
- [3] 陶祝, 刘正熙, 熊运余. 基于卷积神经网络的窨井盖检测 [J]. 现代计算机 (专业版), 2018 (2): 79-84.
- [4] Zhao Ruoyan. Butterfly Recognition Based on Faster R-CNN [A]. 博思学术交流中心. Proceedings of 2018 International Seminar on Computer Science and Engineering Technology (SCSET2018) [C]. 博思学术交流中心; 上海筱虞文化传播有限公司, 2018: 6.
- [5] Zhao Xingwen. Classification of Universal Applied Datasets Based on Convolutional Neural Networks [A]. 中国自动化学会控制理论专业委员会. 第 37 届中国控制会议论文集 [C]. 中国自动化学会控制理论专业委员会; 中国自动化学会控制理论专业委员会, 2018.
- [6] Evgeny A. Smirnov. Comparison of Regularization Methods for ImageNet Classification with Deep Convolutional Neural Networks [A]. Proceedings of 2013 2nd AASRI Conference on Computational Intelligence and Bioinformatics (CIB 2013 V6) [C]. Information Engineering Research Institute, USA, 2013: 6.
- [7] Zeiler M D. Adadelta: an adaptive learning rate method [J]. Computer Science, arXiv Preprint arXiv: 1212.5701, 2012.
- [8] 姜枫, 张丽红. 基于随机 Dropout 卷积神经网络的人体识别方法研究 [J]. 测试技术学报, 2016, 30 (1): 17-22
- [9] Duan Jian, Zhai Huimin. Research on deep convolutional neural network for caltech-101 image classification [J]. Journal of Computer Applications and Software, 2016, 33 (12): 165-168. (in Chinese)
- [10] Cai B, Xu X, Jia K, et al. DehazeNet: an end-to-end system for single image haze removal [J]. IEEE Trans. Image Processing, 2016, 25 (11), 5187-5198.
- [11] Yu Liang, Li Binbin, Jiao Bin. Research and Implementation of CNN Based on TensorFlow [J]. IOP Conference Series: Materials Science and Engineering, 2019, 490 (4).