

基于改进残差网络的扬尘图像识别方法^①



王 艳¹, 张游杰²

¹(太原科技大学 计算机科学与技术学院, 太原 030024)

²(中国电子科技集团公司第三十三研究所, 太原 030032)

通讯作者: 王 艳, E-mail: 1530942248@qq.com

摘 要: 当前利用深度学习方法进行扬尘图像识别的研究较少, 一些传统的方法使得扬尘图像的识别率较低. 针对这种情况, 提出一种基于改进残差网络的扬尘识别方法. 该方法将 ResNet-50 网络应用到扬尘数据集中, 并对其网络结构进行了改进. 加入空间金字塔池化以解决输入图像尺寸不固定的问题, 并且将金字塔池的策略改为平均池化, 将扩大特征图的方法应用到主干网络中, 有利于提取到更加细粒度的特征, 提升模型的性能, 从而提高识别率. 实验结果表明, 该方法具有很高的精确度, 为扬尘识别提供了一种有效的方案.

关键词: 扬尘识别; 深度学习; 残差网络; 金字塔池化; 卷积神经网络

引用格式: 王艳, 张游杰. 基于改进残差网络的扬尘图像识别方法. 计算机系统应用, 2021, 30(5): 202-207. <http://www.c-s-a.org.cn/1003-3254/7909.html>

Dust Image Recognition Method Based on Improved Residual Network

WANG Yan¹, ZHANG You-Jie²

¹(College of Computer Science and Technology, Taiyuan University of Science and Technology, Taiyuan 030024, China)

²(The 33rd Research Institute of China Electronics Technology Group Corporation, Taiyuan 030032, China)

Abstract: At present, there are few studies on dust image recognition using the deep learning method, and the recognition rate of dust images is low due to the application of some traditional methods. In view of this situation, a dust identification method based on an improved residual network is proposed. The method applies ResNet-50 network to a dust data set, and the network structure is improved. Then, spatial pyramid pooling is added to solve the problem that the size of the input images is not fixed. In addition, the pyramid pooling is changed to average pooling, and the method of expanding a feature graph is applied to the backbone network, which is conducive to extract more fine-grained features, improve the performance of the model, and increase the recognition rate. In conclusion, the proposed method has high accuracy and provides an effective scheme for dust identification.

Key words: dust identification; deep learning; residual network; pyramid pooling; Convolutional Neural Network (CNN)

随着我国经济迅速的发展, 环境的问题也变得越来越明显, 其中扬尘污染也越来越成为环境污染的重要来源. 扬尘污染的存在方式是多种的, 在土质和碎石道路上存在, 在建筑工地和水泥等路面上也会存在. 同时, 因为重型车辆和一些大型设备对路面的影响, 扬尘量也变得越来越, 扬尘污染也将会变得更加严重. 因

此, 对于扬尘识别的研究具有重大的实际意义.

传统的扬尘识别方法中大多采用物理信号进行监测. 李晋惠等^[1]充分利用了 Mie 散射原理, 在这一基础上研究出了一种多光路扬尘监测方法. 雷俊花^[2]提出了一种高层建筑装修施工扬尘污染远程监测方法. 牛志云等^[3]设计了基于 STM32 微控制器的光散射式扬

① 基金项目: 山西省重点研发计划 (201903D111002)

Foundation item: Key Research and Development Program of Shanxi Province (201903D111002)

收稿时间: 2020-09-17; 修改时间: 2020-10-13; 采用时间: 2020-10-21; csa 在线出版时间: 2021-04-28

尘监测仪。Yu 等^[4]根据实验室和教室实际环境中空气质量的监测要求,设计了粉尘浓度监测系统。为了能够实时监测大气扬尘,张所容^[5]利用 Mie 散射原理和激光散射法,导出了计算扬尘物质量浓度的方法。随后搭建监测系统的实物架构,并采用基于 B/S 的结构进行了相应的软件系统设计。袁冬琴^[6]设计了基于遥感技术的工地扬尘污染自动监测方法。传统方法的缺点在于比较依赖环境,而且识别的距离是非常有限的,需要大量人力和物力。

除了传统方法外,应用图像处理技术监测扬尘污染也是一个很好的思路。王建鹏^[7]提出一种融合了无线传感器技术、激光前散射粉尘监测技术、图像识别技术和云平台技术的扬尘监测系统。马国鑫等^[8]充分使用了无人机和图像识别技术,将 HSV 特征提取,直方图对比以及非零像素点计算作为核心,设计了施工扬尘污染源自动监测系统,并且测试了该系统的运行效果。任婷婷^[9]运用图像处理技术中的图像滤波去噪,图像分割,图像特征提取和分类识别的方法,对道路扬尘污染监测与识别算法做了研究。这类方法相对于传统方法来讲,可以降低人力和物力的需求,提升识别精度和稳定性,但这一类方法的缺点是比较针对于某一个特定的场景,训练出的模型具有很大的局限性。

基于深度学习的扬尘识别方法能够很好地解决上述存在的问题。因此,提出一种基于改进残差网络的扬尘识别方法,将图像通过改进的 ResNet-50 网络进行特征提取,更好的表达图像信息,提高识别精确度。

1 相关工作

1.1 卷积神经网络

卷积神经网络^[10-12]是一类包含卷积计算并且具有深度结构的前馈神经网络,是深度学习的代表算法之一。它具有表征学习的能力,它能够以原始图像作为输入,通过卷积网络结构来形成更加抽象的高层次特征,从而可以获得更深层次的数据特征表示,避免了人工提取图像特征的复杂^[13]。其中的经典网络有 LeNet-5, AlexNet, ZFNet, VGG-16, ResNet 等。

1.2 残差网络

模型使用的基础网络 ResNet 是 2015 年由 He 等^[14]提出的,在 ImageNet 测试集上的误差达到 3.57%,这个结果在 ILSVRC 2015 分类任务中取得第一名。ResNet 网络主要是为了解决网络中的梯度消失和退化问题,有利于网络更快的收敛。目前,残差网络主要应用于图像分类,物体识别,自然语言处理,目标检测等领域。

ResNet-50 结构图如图 1(a) 所示。

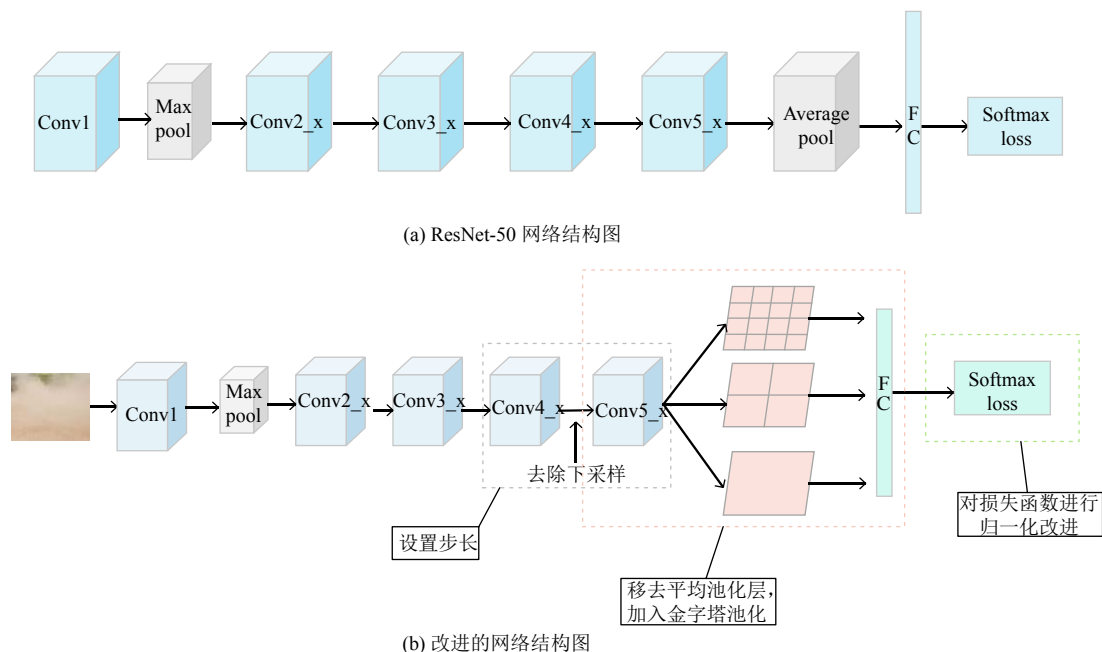


图 1 网络结构图

图 1 中输入图像是大小为 $224 \times 224 \times 3$ 的三维数组,第一个卷积层是独立卷积层 Conv1,使用 64 个大

小为 7×7 ,步长为 2 的卷积核,输出的大小为 112×112 。之后是一个最大池化层,池化窗口和步长分别为 3×3

和2, 输出大小为 56×56 . 接着是4个不同的卷积残差块 Conv2_x, Conv3_x, Conv4_x 和 Conv5_x, 每个卷积残差块由2~3个卷积层和跨越它们的连接组成, 并且每个块的输出大小分别为 56×56 , 28×28 , 14×14 , 7×7 . 最后是平均池化层, 全连接层和损失函数 Softmax. ResNet-50 网络中的残差块主要采用的是 $1 \times 1 + 3 \times 3 + 1 \times 1$ 的卷积层形式, 相比常规网络的2个 3×3 卷积层形式, 使用三层残差块结构不仅可以保持网络精度, 同时还能减少网络的计算量. 两种残差块的结构如图2(a)和图2(b)所示.

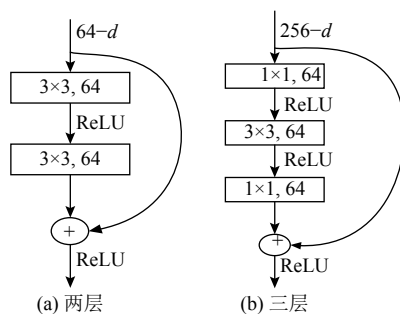


图2 残差块

1.3 残差网络的学习算法

通过一个假设将残差的概念引入进来: 假设可以通过恒等映射的方式来构建所增加的层, 这样就可以把一个深层网络变成了一个浅层的网络. 至少可以保证深层网络的训练误差不会大于所基于的较浅模型. 将 $H(x)$ 假设为所需的底层映射, 第一层的输入为 x , 让堆叠的非线性层来拟合另一个映射: $F(x)=H(x)-x$. 因此原来的映射转化为: $F(x)+x$.

在残差网络中, 对于网络中每一个堆叠层都将采用残差学习的方式. 残差块如图2所示. 并将残差块定义为:

$$y = f(x, \{W_i\}) + x \quad (1)$$

其中, x 和 y 是考虑层的输入和输出向量, $f(x, \{W_i\})$ 是要学习的残差映射, W_i 是 i 层的权重. 在式(1)中 x 和 f 的维度是必须要相等的. 如果维度不同, 可以通过快捷连接方式执行线性投影 W_s 以匹配维度:

$$y = f(x, \{W_i\}) + W_s x \quad (2)$$

2 模型设计

ResNet-50 网络在许多领域都得到了广泛的应用, 比如在人脸识别等比较成熟的领域应用较为理想, 但应用于扬尘识别时, 识别准确率较低. 因此, 对

ResNet-50 网络进行改进, 形成新的网络模型, 以适应扬尘图像识别, 改进如下: (1) 将网络中最后一次空间下采样操作的卷积层步长从原来的2设置为1; (2) 移去了网络中最后的平均池化层. 同时, 在该网络的基础上引入了多尺度的金字塔池化方法; (3) 在全连接层后加入改进的 Softmax 损失函数, 进一步提高模型的识别性能. 整体的网络结构如图1(b)所示.

2.1 扩大特征图

在 ResNet-50 网络中, 网络中的每个块都会有一个空间下采样的操作, 为了增大扬尘特征图的空间大小, 去除网络最后一次空间下采样. 当输入一张 224×224 的扬尘图像后, 网络会输出一个 7×7 的特征图. 将网络中最后一次空间下采样卷积层的步长从2设置为1, 将特征图的输出尺寸扩大为 14×14 . 这样不仅不会增加模型的参数量, 而且还能够使得模型的参数结构得以保持, 增大特征图的空间大小有利于提取到更加细粒度的扬尘图像特征, 进而提升扬尘识别模型的性能.

2.2 金字塔池化层加入

在针对扬尘识别的 ResNet-50 网络中, 卷积层接受任意大小的输入, 但它们产生的输出大小不一, 而全连接层需要固定长度的向量作为输入. 那么对于结构已经确定的 ResNet-50 网络, 就需要对扬尘图像进行裁剪或缩放等操作, 这么做会降低扬尘识别的精度. 在2015年 He 等^[15]提出的金字塔池化(SPP)方法解决了这一问题. 金字塔池化方法能够让网络适应不同大小尺寸的扬尘图片, 不仅增加了扬尘识别模型的尺度不变性, 而且还可以抑制训练过程中过拟合现象的发生. 它可以通过将空间信息集中在本地的空间容器中来维护空间信息, 这些空间存储单元的大小与扬尘图像大小成比例, 无论图像大小, 存储单元的数量都是固定的. 因此, 改进的 ResNet-50 网络模型去掉全连接层之前的平均池化层, 并将金字塔池化的方法引入到扬尘识别中. 金字塔池化层如图3所示, 具体操作是将经过 ResNet-50 网络输出的特征映射采用不同的池化尺度($1 \times 1, 2 \times 2, 4 \times 4$), 得到21个固定的特征向量, 然后把这些向量平铺组成一个新的特征向量并送入全连接层.

金字塔池化层就是在最后一个残差块 conv5_x 卷积层的扬尘特征图的每个图片上进行了3个卷积操作. 最下边为原图像特征图, 中间的是分成大小为4的特征图, 最上边的是分成大小为16的特征图. 因此, 每一个特征图就变为 $1+4+16=21$ 个特征图. 其中每一个卷

积操作的本质都是平均池化,不同的只是窗口大小和步长.参数设置如图4所示.

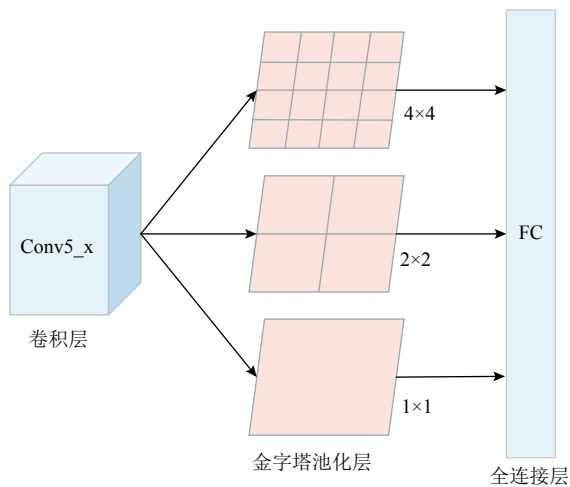


图3 金字塔池化层

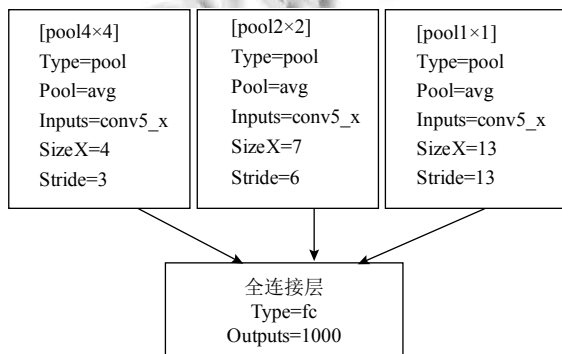


图4 金字塔池化参数设置

2.3 损失函数的改进

为了充分发挥 ResNet-50 网络体系结构学习表示的识别能力,受人脸识别中损失函数的启发,在扬尘识别模型训练阶段加入了改进的 Softmax 损失函数. Softmax 对于类间距离的优化效果很明显,对于所学习的特性, Softmax 损失函数公式为:

$$L_{\text{Softmax}} = -\frac{1}{m} \sum_{i=1}^m \log \frac{e^{W_{y_i}^T x_i + b_{y_i}}}{\sum_{j=1}^p e^{W_j^T x_i + b_j}} \quad (3)$$

其中, m 是训练过程中的样本的数目, p 是训练数据集中的类数. x_i 是第 i 个样本的特征, y_i 是对应于 x_i 的标签, W 和 b 是 Softmax 损失之前的最后内积层的权重矩阵和偏置向量, x 是网络最后一层输出的特征,经过一个全连接的分类层,即乘以权重 W 并加上偏置 b , 输出 score. W_j 是 W 的第 j 行, 对应于 j 类. Softmax 可以有

效区分类间差异,但是对于类内的分布没有很好的约束,因此监督效果不够.

$$z_j = W_j^T x = \|W_j\| \|x\| \cos \theta_j \quad (4)$$

原始的 Softmax 损失在空间分布比较随意,输出的 score 为式 (4), 即神经元的权重向量的模 $\|W_j\|$ 乘以特征向量的模 $\|x\|$, 再乘以两个向量夹角的余弦值 $\cos \theta_j$, 因此, 在分类过程中 score 会受到模和夹角的影响.

$$W_j^* = \frac{W_j}{\|W_j\|}, X^* = \frac{x}{\|x\|} \quad (5)$$

为了消除在扬尘识别过程中模和夹角的影响,将权重向量 W_j 和特征向量 x 归一化, 从而消除模的影响. 如式 (5) 所示, 归一化之后的 Softmax 只需要比较两个角度的大小, 这样可以减少损失函数优化参数的数量, 只需要考虑角度, 大大提升了 Softmax 分类器的分类性能, 提高了扬尘识别的准确率.

3 实验分析

3.1 实验数据集

由于扬尘图像识别目前没有公开的数据集. 因此, 通过在互联网上收集各类有关扬尘场景的图片, 整理形成训练集和测试集, 然后使用数据增强在训练集和测试集中生成更多的数据, 解决了该研究中没有数据集以及数据量不足的问题. 为了增强模型的泛化能力, 采用数据增强的操作扩充所需的数据样本, 让网络模型更好的适应各种应用场景.

首先收集含有扬尘以及非扬尘的图像制作训练和测试数据集, 并使用数据扩增技术增加数据量, 部分实验图像如图 5、图 6 所示. 实验数据集分为训练集和测试集, 每个集合均含有扬尘图像与非扬尘图像. 将 10000 张扬尘图像与 5000 张非扬尘图像作为实验数据, 将其中的 13000 张作为训练集, 2000 张作为测试集.

3.2 实验设置

实验使用的操作系统是 64 位的 Ubuntu18.04 操作系统, 在训练和测试过程中使用了 NVIDIA GTX 1660Ti GPU 进行加速, 算法使用深度学习框架 PyTorch 实现, Python 版本为 3.6.

实验的主干网络采用了在 ImageNet^[16] 上预训练好的 ResNet-50 模型, 完成模型的权重初始化, 这样可以减少训练时间. 扬尘图片的输入尺寸将其调整为 224×224 . 在训练过程中, 设定初始学习率为 0.001, 迭代次数为 120 次, batch 大小设置为 32.



图5 非扬尘图像



图6 扬尘图像

3.3 实验结果

为了说明改进 ResNet-50 网络的有效性, 选择了 3 种方法进行对比, 方法 1 是使用原始的 ResNet-50 网络, 方法 2 是将最后一次空间下采样操作的卷积层步长从原来的 2 设置为 1, 不使用金字塔池化的 ResNet-50 网络, 方法 3 是只使用金字塔池化, 最后一次空间下采样操作的卷积层步长不变的 ResNet-50 网络, 方法 4 是使用改进的 ResNet-50 网络. 此外, 在训练中 4 种方法均使用 Softmax 损失函数进行优化. 4 种方法的实验对比结果如表 1 所示. 各个实验的训练曲线如图 7 所示.

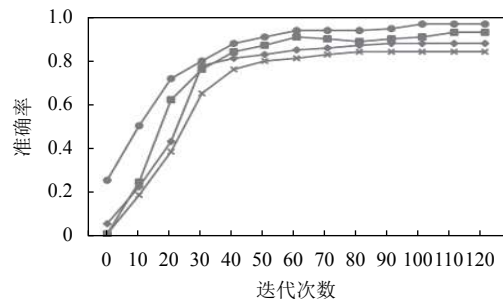
表 1 不同网络实验对比结果

方法	描述	训练次数	准确率(%)
方法1	基础网络	120	88.7
方法2	去除最后空间下采样	120	89.2
方法3	加入金字塔池化	120	89.5
方法4	改进ResNet-50	120	93.4

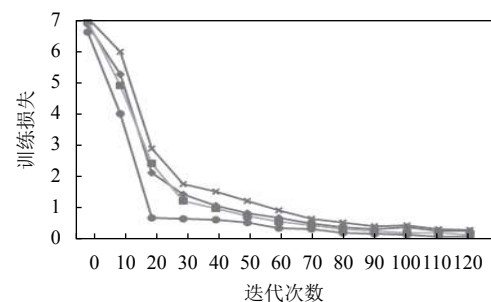
通过对这 4 种网络结构的实验结果进行比较可以看出, 使用初始 ResNet-50 网络的准确率最低. 与方法 1 相比, 通过增大特征图尺寸来提高识别效果的方法可以提高 2.5%, 使用金字塔池化后可以提高 5.1%, 而方法 4 要比方法 1 提高 8%, 说明了将增大特征图尺寸和金字塔池化的方法加入到 ResNet-50 网络中是有效的, 证明了改进 ResNet-50 网络的有效性.

从图 7 可以看出, 准确率提高和损失率下降最快的是方法 4, 相比其他 3 种方法, 方法 4 的收敛速度更快, 在训练次数达到 60 次时趋于稳定在 0.94 左右. 方

法 2 和方法 3 的收敛速度有交叉的部分, 方法 1 的速度是最慢的.



(a) 训练准确率



(b) 训练损失

—×— 方法1 —◆— 方法2 —■— 方法3 —●— 方法4

图 7 训练曲线

为了验证 ResNet-50 网络的性能, 选取了其他两种卷积神经网络 Vgg16^[17] 和 Alexnet^[18] 训练扬尘数据集, 并和 ResNet-50 网络的训练准确率作比较, 比较结果如表 2 所示. 通过表中数据可以发现, 在这 3 种网络结构

模型中 ResNet-50 的准确率最高。

为了说明模型的泛化能力,采用改进 ResNet-50 网络和 ResNet-50 网络分别在公共数据集 CIFAR-10 上进行了实验。CIFAR-10 是一个和普适物体相近,由 Hinton 的学生整理的一个用于识别普适物体的小型彩色图像数据集,共包含 10 个类别的 RGB 彩色图像,每个类别有 6000 张图像,每张图像的尺寸大小为 32×32 。该数据集有 50 000 张图像用于训练集,10 000 张图像用于测试集。实验准确率对比结果如表 3 所示。

表 2 3 种网络模型准确率对比

网络	迭代次数	准确率(%)
Alexnet	120	82.8
Vgg16	120	84.3
ResNet-50	120	88.7

表 3 CIFAR-10 数据集准确率对比

网络	迭代次数	准确率(%)
ResNet-50	120	80.5
改进ResNet-50	120	91.3

基于改进 ResNet-50 网络的扬尘识别方法在该数据集下的准确率为 91.3%,相比基础的 ResNet-50 网络提高了 10.8%,说明了改进的网络模型有较好的泛化能力。

4 结论

深度学习在图像识别中具有重要意义,相比传统的图像识别方法,基于深度学习的识别方法具有自学习和适应能力,可以更好地对目标进行特征提取。改进的网络模型将 ResNet-50 网络应用到扬尘识别中,并对其进行了改进。经实验验证,改进 ResNet-50 网络对扬尘图像有较高的识别率,同时也有很好的泛化能力,可以有效地识别各种类型扬尘图像以及其他非扬尘图像,在实际中有很大的应用价值。

参考文献

- 李晋惠,于亚琳,田军委.多光路光散射扬尘检测方法研究.电子测量技术,2015,38(10):101-105.[doi:10.3969/j.issn.1002-7300.2015.10.020]
- 雷俊花.高层建筑装修施工扬尘污染远程监测方法研究.环境科学与管理,2019,44(10):145-149.[doi:10.3969/j.issn.1673-1212.2019.10.028]
- 牛志云,杨义新,桂华侨,等.基于 STM32 的光散射式扬尘在线监测仪.仪表技术与传感器,2020,(1):36-41.
- Yu T, Xu MD, Zhang TQ. Design and implementation of a PM_{2.5} dust detection system based on ZigBee. Proceedings of 2019 International Conference on Wireless Communication, Network and Multimedia Engineering. Guilin, China. 2019. 186-189.
- 张所容.基于激光检测法的大气扬尘实时在线监测研究.工业安全与环保,2018,44(7):68-71.[doi:10.3969/j.issn.1001-425X.2018.07.018]
- 袁冬琴.基于遥感技术的施工工地扬尘污染自动监测方法研究.环境科学与管理,2020,45(1):136-141.[doi:10.3969/j.issn.1673-1212.2020.01.027]
- 王建鹏.基于激光前散射的智能扬尘监测系统.机械管理开发,2019,34(12):135-136,138.
- 马国鑫,韩豫,陆建飞,等.基于无人机的施工扬尘污染源自动监测系统设计与实现.中国环境监测,2018,34(1):151-156.
- 任婷婷.基于视频图像处理的城市道路污染检测与识别[硕士学位论文].太原:太原科技大学,2015.
- Szegedy C, Liu W, Jia YQ, et al. Going deeper with convolutions. Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA. 2015. 1-12.
- Howard AG, Zhu ML, Chen B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv: 1704.04861v1, 2017.
- Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA. 2017. 2261-2269.
- 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.激光与光电子学进展,2017,54(9):091501.
- He KM, Zhang XY, Ren SQ, et al. Deep residual learning for image recognition. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. 2016. 770-778.
- He KM, Zhang XY, Ren SQ, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015,37(9):1904-1916.[doi:10.1109/TPAMI.2015.2389824]
- Deng J, Dong W, Socher R, et al. ImageNet: A large-scale hierarchical image database. Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, FL, USA. 2009. 248-255.
- Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv: 1409.1556v6, 2015.
- Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, NV, USA. 2012. 1097-1105.