

嵌入式设备高效卷积神经网络的电力设备检测^①



林唯贤

(中国石油大学(华东) 计算机与通信工程学院, 青岛 266580)

通讯作者: 林唯贤, E-mail: linvhian@163.com

摘要: 随着大型图像集的出现以及计算机硬件尤其是 GPU 的快速发展, 卷积神经网络 (CNN) 已经成为人工智能领域的一种成功算法, 在各种机器学习任务中表现出色. 但 CNN 的计算复杂度远高于传统算法, 嵌入式设备上有限资源的限制成为制造高效嵌入式计算的挑战性问题. 在本文中, 我们提出了一种基于嵌入式设备的高效卷积神经网络用于电力设备检测, 根据处理速度评估这种高效的神经网络. 结果表明, 该算法能够满足嵌入式设备实时视频处理的要求.

关键词: 嵌入式设备; 深度学习; 卷积神经网络 (CNN); 电力设备检测

引用格式: 林唯贤. 嵌入式设备高效卷积神经网络的电力设备检测. 计算机系统应用, 2019, 28(5): 238–243. <http://www.c-s-a.org.cn/1003-3254/6917.html>

Efficient Convolutional Neural Networks for Electrical Equipment Inspection on Embedded Devices

LIN Wei-Xian

(College of Computer & Communication Engineering, China University of Petroleum, Qingdao 266580, China)

Abstract: With the emergence of large image sets and the rapid development of computer hardware especially GPU, Convolutional Neural Network (CNN) has become a successful algorithm in the region of artificial intelligence and exhibit remarkable performance in various machine learning tasks. But the computation complexity of CNN is much higher than traditional algorithms, however, the restrict of limited resources on embedded devices become a challenging issue for making efficient embedded computing. In this study, we propose a efficient convolutional neural networks based on embedded devices for electrical equipment inspection, this efficient neural network is evaluated in term of processing speed. The results show that the proposed algorithm can meet the requirement of real-time video processing on embedded devices.

Key words: embedded devices; deep learning; Convolutional Neural Networks (CNN); electrical equipment detection

红外热像仪很早就应用于电力设备的检修上, 但是, 目前电力系统中最主要的状态检测方式是由技术人员手持热像仪现场诊断或者采集图片分析. 对于规模庞大的变电站, 仅靠人工诊断任务繁重、效率低, 并且具有一定的安全风险. 据统计, 电力系统故障事故中高达 90% 是由于电力设备故障引起, 而发生事故的电力设备中将在故障早期阶段呈现出异常的过热症状^[1],

因此为保证电力系统安全运行, 对电力设备的状态检修、故障诊断提出了更高的要求. 红外成像技术是通过探测器检测目标物体表面发射出来红外辐射, 将它转化成与之相应的电信号, 再通过专门的电信号处理系统, 获得物体表面的温度分布, 因此红外成像技术十分适用于电力设备因异常过热症状所导致故障状态检测.

① 收稿时间: 2018-12-07; 修改时间: 2018-12-25; 采用时间: 2019-01-03; csa 在线出版时间: 2019-05-01

随着深度学习和新型传感器于信息技术的融合得到广泛研究和认可,并应用于实际场景中^[2],因此,可以通过巡检机器人搭载低功耗嵌入式设备来实现实时视频处理的无人电力设备智能巡检,降低作业人员的安全风险,代替繁重的人工巡检,以提高巡检的自动化智能化。低功耗的嵌入式设备可以搭载实时视频框架来实现监控、巡检等任务,但是嵌入式设备有限的硬件资源使得无人离线巡检、监控等任务难以推广。在本文中,我们基于深度可分离卷积设计一个高效的卷积模块,并改进主流的目标检测算法,可在嵌入式设备上实现实时电力设备目标检测。

本文分为四个章节,第1节描述现有的红外图像目标检测算法和深度学习模型压缩优化;第2节描述适用于单个嵌入式设备的高效卷积神经网络设计;第3章节为实验评估;第4章节给出了结论和展望。

1 红外图像目标检测

要实现视频实时电力设备检测的首要步骤就是将视频流解码成帧,运用目标检测算法从图像中检测出电力设备。

传统的方法是从红外图像中找出感兴趣的区域,中区域中提取信息进行分析匹配。使用传统方法关键在于,获得感兴趣的区域的质量决定了能否较好的获得电力设备状态的信息。Huda等^[3]使用简单的阈值分割来获得感兴趣区域。Jaffery等^[4]对红外图像采用数学形态学处理和阈值分割的方法来获得感兴趣区域。门洪等^[5]先对红外图像采用变换结合的方法来增强图片,增强后的图片采用CA算法进行边缘提取并使用二值化处理,最后和OTSU算法分割后的红外图像进行融合。Korman等^[6]提出FAsT-Match(Fast Affine Template Matching),在二维放射变换下减小SAD误差,达到近似全局最优的快速仿射变换匹配算法。可以看出,目前的电力设备红外图像检测的算法较少,传统的方法需要人工提取和设计每个设备的特征,将这些特征应用到分类器来获得匹配结果,但是人工提取和设计特征消耗大量的时间和经历,且传统算法不很好的将目标和背景区分出来。

深度学习能够克服传统方法的缺陷,通过逐层的特征变换,将蕴含于数据中的信息映射到新的特征空间上,让计算机自动的学习到高层的特征^[7],随着数据量的增加而增强学习的效果,获得更高的匹配精度,这

些特征并且拥有较好的鲁棒性。

R-CNN(Region CNN)^[8]最先使用深度学习进行目标检测。R-CNN提取某些区域,对这些区域提取特征,相比较传统目标检测算法使用人工设计的特征,R-CNN使用海量数据来训练卷积神经网络,并使用卷积神经网络来提取特征。SPP-Net^[9]提出金字塔池化,接在最后一个卷积层,实现网络模型输入任意图片并且输出固定大小。Fast-R-CNN^[10]训练方式简单,基于多任务的损失函数设计,不需要再设计SVMs。Faster-R-CNN^[11]使用RPN候选框提取网络,之后是用Fast-R-CNN进行目标检测,训练时RPN和Fast-R-CNN共享卷积层权重降低时间消耗并提高候选框的质量。YOLO(You Only Look Once)^[12]对目标检测使用回归方式,采用一个单对单的卷积神经网络模型来实现端对端(End-to-End)的目标检测,YOLO9000^[13]可以识别超过9000中目标类别。

深度神经网络复杂、训练大、计算量、参数量大,模型的部署需要非常强大的硬件设备,嵌入式设备由于硬件限制只能部署浅层神经网络,目标检测精度受到限制。要实现工程化,需要研究深度网络模型的压缩、降低内存占用、降低功耗、减少计算量和参数量。剪枝(Pruning)是对网络连接的裁剪,只保留重要性高的连接,对模型的参数进行量化(Quantization)以减少模型的体积,还可以使用哈夫曼编码(Huffman Coding)压缩模型^[14]。卷积神经网络计算大部分集中在卷积操作,因此压缩模型、提高效率需要设计新型的网络结构,谷歌公司提出的MobileNet^[15]根据Sifre^[16]提出深度可分离卷积(Depthwise Separable Convolutions)将常规的卷积分解,对每个通道单独计算后再进行线性组合,显著降低了模型的计算量和参数量。

2 目标检测模型设计

在这一部分内容介绍设计高效的深度神经网络,基于深度可分离卷积与组逐点群卷积和目标检测算法部署在Jeston TX1上,以实现离线红外热像仪实时电力设备检测。

2.1 深度可分解卷积

常规的卷积神经网络大部分计算都集中在卷积操作上,因此减少网络模型的复杂度、减少模型的计算量和减少模型大小的关键是设计高效的卷积层。深度可分离卷积^[16](Depthwise Separable Convolutions),它是

分解卷积的一种形式, 它将常规卷积分解为 Depthwise 卷积和称为 Pointwise 卷积的 1×1 卷积. Depthwise 卷积将单个滤波器应用于每个输入通道中. 随后 Pointwise 卷积使用 1×1 卷积来组合 Depthwise 卷积. 常规卷积在一个步骤将过滤并将输入合并成一组新的输出. 而深度可分离卷积将常规卷积分成两层, 一层用于过滤, 另一层用于合并. 这种分解的卷积层在减少模型的计算量和模型大小效果显著.

常规的卷积层输入一个特征图 (Feature Map) F 的大小为 $f \times f \times m$ 生成一个 $f \times f \times n$ 的特征图 G . 其中 f 是输入的特征图的宽度和高度, m 是输入通道数, n 是输出通道数. 常规的卷积层通过卷积核 K 参数化, K 的大小为 $k \times k \times m \times n$, k 为空间维度.

假设 Padding 和 Stride 为 1, 常规卷积的计算公式为:

$$G_{k,l,n} = \sum_{i,j,m} K_{i,j,m} \times F_{k+i-1,l+j-1,m} \quad (1)$$

所以常规的卷积计算消耗为:

$$k \times k \times m \times n \times f \times f \quad (2)$$

常规卷积操作具有通过卷积核过滤特征并结合特征来生成新的效果. 深度可分离卷积将常规卷积分成两个步骤来减少计算量和压缩模型体积. 深度可分离卷积由 Depthwise 卷积和 Pointwise 卷积这两层组成, depthwise 卷积对每个输入通道使用单个 Filter, 随后使用 1×1 卷积也就是 Pointwise 卷积来创建一个深度层输出的线性组合. 对这两个层都使用 batchnorm 和 ReLU.

Depthwise 卷积对每个输入通道使用单个 Filter 计算公式为:

$$\hat{G}_{k,l,m} = \sum_{i,j} \hat{K}_{i,j} \times F_{k+i-1,l+j-1,m} \quad (3)$$

所以 Depthwise 卷积的计算量为:

$$k \times k \times m \times f \times f \quad (4)$$

Depthwise 卷积相对于常规卷积只过滤输入通道不会将它们组合起来成一个新的特征. 因此, 为了得到这些新特征, 需要使用 Pointwise 卷积也就是 1×1 卷积将 Depthwise 卷积的输出再进行结合.

所以深度可分离卷积的计算量为 Depthwise 卷积计算量与 Pointwise 卷积计算量之和:

$$k \times k \times m \times f \times f + m \times n \times f \times f \quad (5)$$

所以深度可分离卷积与常规卷积的计算量之比为:

$$\frac{k \times k \times m \times f \times f + m \times n \times f \times f}{k \times k \times m \times n \times f \times f} = \frac{1}{n} + \frac{1}{k \times k} \quad (6)$$

2.2 卷积模块设计

本文基于 Depthwise 卷积与 Pointwise Group 卷积和 Channel Shuffle^[17]的基础上设计一个高效的卷积模块. 图 1(a) 表示常规的卷积操作, 常规的卷积层后面接 BatchNorm^[18]和 ReLU, 图 1(b) 表示 Depthwise 卷积与 Pointwise 卷积的使用方式, 一般在 Depthwise 卷积与 Pointwise 卷积后面连接 BatchNorm 和 ReLU.

图 1(c) 是本文设计的卷积模块. Depthwise 卷积没有改变通道数的能力, 上一层通道数量少, 则 Depthwise 卷积只能在低维空间中提取特征, 效果不显著. 为了改善这个问题, 在每个 Depthwise 层前加一个 Pointwise 卷积用来提升维度, 获得更多的特征.

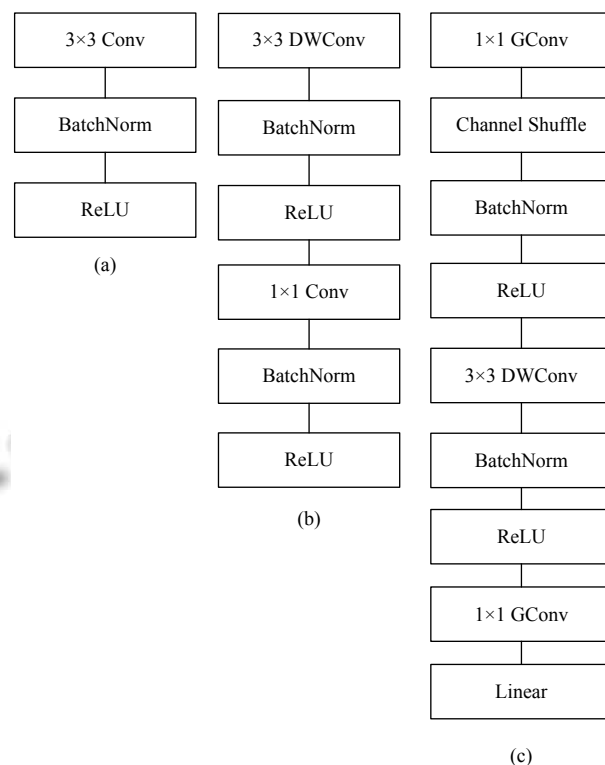


图1 卷积模块对比

如表 1 所示, MobileNets 在 Pointwise 卷积也就是 1×1 卷积中花费 94.86% 的计算时间也包含了整个网络 74.59% 的参数.

Pointwise 卷积耗时且容易造成有限的通道之间充满约束损失精度, 使用 Pointwise Group 卷积, 也就是带

组的卷积核为 1×1 的卷积来建立通道稀疏连接. 在多层 Pointwise 卷积堆叠, 模型的信息流被分割在各个组内, 引入 Channel Shuffle^[17], 使各组间信息能相互交流, 并且降低卷积运算的跨通道计算量和错误率. 第二个 Pointwise Group 卷积用来降维, 激活函数在高位空间能够有效增加非线性, 而在低维空间会破坏特征^[19], 于是如图 1(c) 第二个 Pointwise Group 卷积后接一个 Linear Bottleneck. 卷积神经网络大部分计算都集中在卷积操作上, 设计出来的卷积模块包含一个 Depthwise 卷积与两个 1×1 卷积, 因此该卷积模块与常规的卷积计算量之比为:

$$\frac{k \times k \times m \times f \times f + 2 \times m \times n \times f \times f}{k \times k \times m \times n \times f \times f} = \frac{2}{n} + \frac{1}{k \times k} \quad (7)$$

表 1 MobileNets 各网络层类型的资源

Type	Mult-Adds(%)	Parameters(%)
1×1 Conv	94.86	74.59
3×3 DW Conv	3.06	1.06
3×3 Conv	1.19	0.02
Fully Connected	0.18	24.33

2.3 目标检测

如今先进的目标检测算法都会生成一系列假设的

候选框, 从这些候选框中提取特征, 最后将这些特征输入到分类器中来判断候选框分目标类别. 对于资源有限的嵌入式设备来说, 这些目标检测的方法计算量大, 耗时长, 对于单个嵌入式设备难以实现实时检测.

SSD(Single Shot MultiBox Detector)^[20]取消中间的候选框和像素特征的重采样过程, 保证速度的同时保证了检测的精度. SSD 输出一系列离散化的候选框, 候选框生成在不同层上的特征图且长宽比不同, 经过卷积神经网络的前馈操作, SSD 生成一系列固定大小的候选框, 使用小卷积 Filter 来预测候选框位置中的目标类别和偏移即候选框中包含目标种类的概率, 最后通过极大值抑制方法得到最终的预测结果.

为了使 SSD 适用于嵌入式设备, 我们对 SSD 进行改进. 如图 2 所示 SSD 以 VGG-16 网络为基础, 第 6 层和第 7 层由全连接层改为卷积层, 后接三个卷积层与平均池化层作为特征提取层. 我们改进的 SSD 使用 Mobilenet 取代 VGG-16 作为 SSD 的基础网络结构, 并且保留 Conv7, 并将第 7 至第 10 层标准的卷积层换成 2.2 节所提出的卷积模块, 并且不断调整 SSD 网络的参数, 调整迭代次数和学习率等超参数, 以获得最优的检测精度.

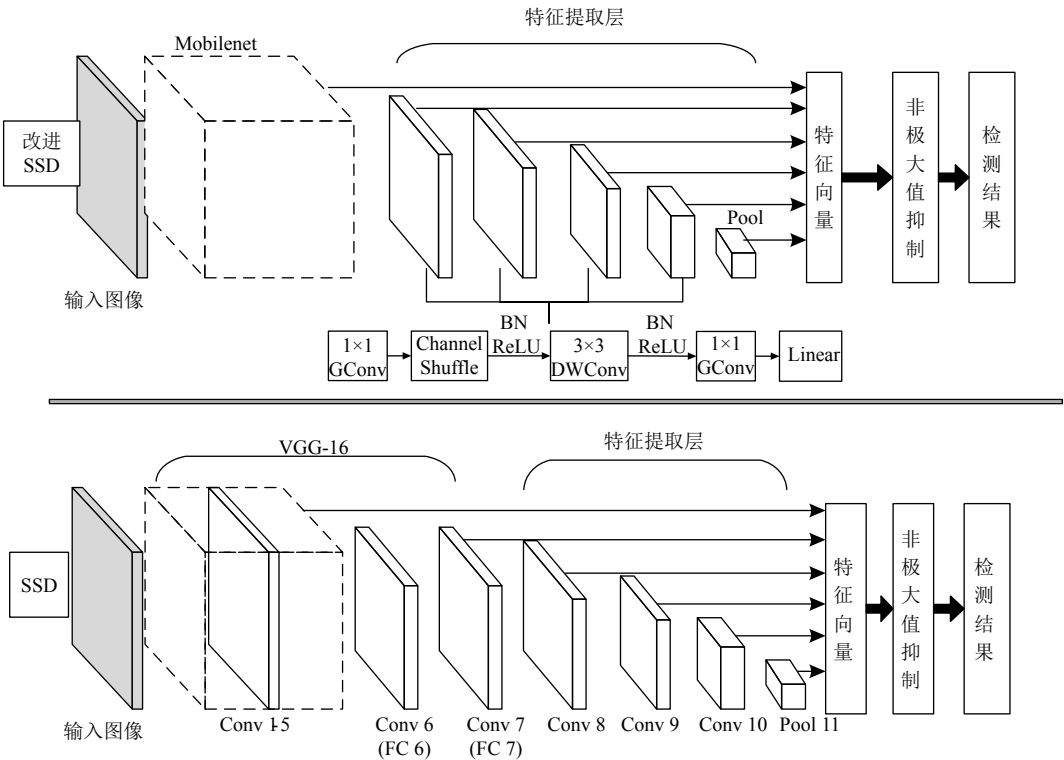


图 2 改进 SSD 与 SSD 结构对比

3 实验分析

3.1 数据集

实验数据集由变电站采集而来, 包含红外图像 11835 张. 为了增加训练样本的数量以及数据样本的多样性, 对采集到红外图像数据集进行增强, 将数据集中的图像水平和垂直方向分别平移 0, +1, +2, +4, +6, +8, -1, -2, -4, -8 个像素, 图像内正时针、逆时针旋转 0°, 15°, 30°, 45°, 用人工标注方法对数据集进行标注, 将数据集分为训练集、验证集、测试集三部分.

3.2 硬件环境

Jetson TX1 拥有先进嵌入式计算机视觉系统, 表 2 列出了 Jetson TX1 的硬件资源. 本文的实验将使用单个 TX1 进行开发, 部署本文设计的改进的轻量级 SSD 算法.

表 2 Jetson TX1 硬件资源

GPU	CPU	Memory	Storage
NVIDIA Maxwell 256 CUDA cores	64-bit A57 CPUs	4 GB 64-bit LPDDR4	16 GB eMMC

3.3 结果分析

本次实验通过读取红外热像仪的视频流解码成帧, 在 Jetson TX1 平台运行视频目标检测算法, 对各类目

标检测算法的速度方面进行比较, 电力设备的检测例子如图 3 所示.

表 3 目标检测算法速度 ms/帧

模型名称	实验次数			
	1	2	3	平均
Tiny YOLO	25.85	23.52	23.82	24.40
YOLOv2	90.07	92.31	90.57	90.98
SSD	102.93	106.96	105.07	104.99
MobileNet-SSD	17.63	17.25	18.13	17.67
改进 SSD	18.20	17.74	16.80	17.58

表 3 显示用我们改进的 SSD 算法与 Tiny YOLO YOLOv2、SSD、SSD 与 MoblieNets 结合的目标检测算法在 Jetson TX1 平台上读取红外热像仪进行电力设备检测单帧图像耗时的比较. YOLOv2 和 SSD 都是目前主流的目标检测算法, 由于计算量大耗时多, 不适用于单个 Jetson TX1 上部署模型. Tiny YOLO 在 YOLO 的基础上, 删除了中间的网络层来达到加速的目的, 代价是网络层数少, 损失精度大. MobileNets-SSD 改进了卷积模块, 降低了计算花费和参数量. 而我们改进的 SSD 算法, 设计了高效的卷积模块, 保留了网络的深度的同时显著的提升了速度并且提高各组间信息的交流降低卷积运算的跨通道计算量.

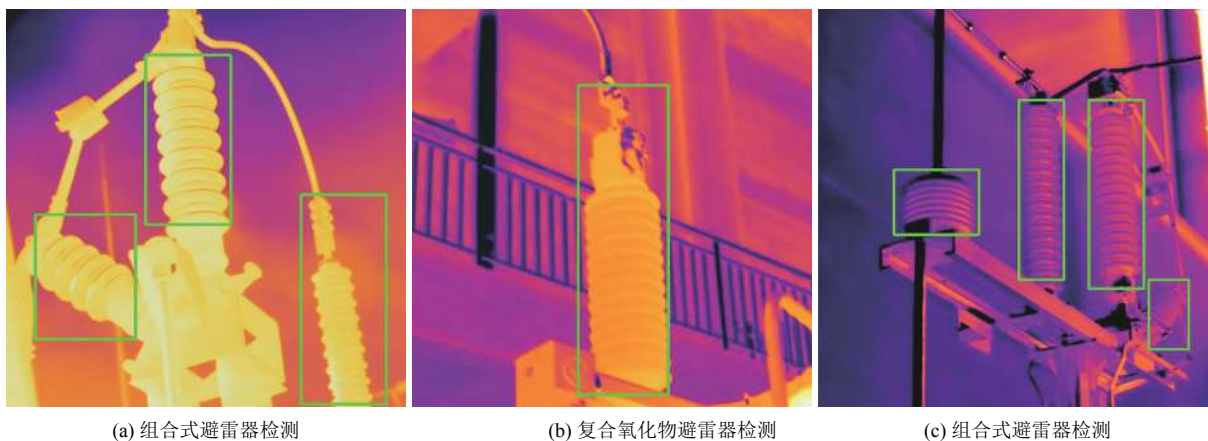


图 3 检测结果

4 结论与展望

小型高效的深度神经网络模型对于检测、识别等计算机视觉任务具有重要意义, 消费类产品很多都是基于嵌入式的终端产品, 计算资源有限, 对于实时视频检测任务来说是一个巨大的挑战. 我们基于深度可分离卷积设计了高效的卷积模块, 并改良了主流的目标检测算法. 实验证明, 在 TX1 上我们的模型能满足电

力设备实时检测的要求, 保持网络深度的同时显著地提高了运行速度. 接下来的研究工作是如何提高模型的精度.

参考文献

- 1 Lin J, Lin LH, Liu GQ, *et al.* A substation monitoring and warning system based on infrared technology and image

- separating. Proceedings of 2008 3rd International Conference on Intelligent System and Knowledge Engineering. Xiamen, China. 2008. 66–70.
- 2 Amantea R, Goodman LA, Pantuso FP, *et al.* Progress toward an uncooled IR imager with 5-mK NETD. SPIE's International Symposium on Optical Science, Engineering, and Instrumentation. San Diego, CA, USA. 1998. 647–660.
- 3 Huda ASN, Taib S. Suitable features selection for monitoring thermal condition of electrical equipment using infrared thermography. *Infrared Physics & Technology*, 2013, 61: 184–191.
- 4 Jaffery ZA, Dubey AK. Design of early fault detection technique for electrical assets using infrared thermograms. *International Journal of Electrical Power & Energy Systems*, 2014, 63: 753–759.
- 5 门洪, 于加学, 秦蕾. 基于 CA 和 OTSU 的电气设备红外图像分割方法. *电力自动化设备*, 2011, 31(9): 92–95. [doi: [10.3969/j.issn.1006-6047.2011.09.020](https://doi.org/10.3969/j.issn.1006-6047.2011.09.020)]
- 6 Korman S, Reichman D, Tsur G, *et al.* Fast-match: Fast affine template matching. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Portland, OR, USA. 2013. 2331–2338.
- 7 Hinton GE, Osindero S, Teh YW. A fast learning algorithm for deep belief nets. *Neural Computation*, 2006, 18(7): 1527–1554. [doi: [10.1162/neco.2006.18.7.1527](https://doi.org/10.1162/neco.2006.18.7.1527)]
- 8 Girshick R, Donahue J, Darrell T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation. Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA. 2014. 580–587.
- 9 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](https://doi.org/10.1109/TPAMI.2015.2389824)]
- 10 Girshick R. Fast R-CNN. Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile. 2015. 1440–1448.
- 11 Ren SQ, He KM, Girshick R, *et al.* Faster R-CNN: Towards real-time object detection with region proposal networks. Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal, Canada. 2015. 91–99.
- 12 Redmon J, Divvala S, Girshick R, *et al.* You only look once: Unified, real-time object detection. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. 2016. 779–788.
- 13 Redmon J, Farhadi A. YOLO9000: Better, faster, stronger. arXiv preprint arXiv:1612.08242, 2016.
- 14 Han S, Mao HZ, Dally WJ. A deep neural network compression pipeline: Pruning, quantization, huffman encoding. arXiv preprint arXiv:1510.00149, 2015.
- 15 Howard AG, Zhu ML, Chen B, *et al.* MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861, 2017.
- 16 Sifre L. Rigid-motion scattering for image classification[Ph. D. thesis]. Paris: Ecole Polytechnique, 2014.
- 17 Zhang XY, Zhou XY, Lin MX, *et al.* Shufflenet: An extremely efficient convolutional neural network for mobile devices. arXiv preprint arXiv:1707.01083, 2017.
- 18 Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167, 2015.
- 19 Sandler M, Howard A, Zhu ML, *et al.* MobileNetV2: Inverted residuals and linear bottlenecks. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA. 2018. 4510–4520.
- 20 Liu W, Anguelov D, Erhan D, *et al.* SSD: Single shot MultiBox detector. arXiv preprint arXiv:1512.02325, 2016.