

基于 MobileNetV2 的干电池底盖缺陷检测研究

高学秋¹ 张 祺¹ 吴泽强¹ 钟善机¹ 刘 磊¹

GAO Xueqiu ZHANG Qi WU Zeqiang ZHONG Shanji LIU Lei

摘 要

干电池是用于小型电子产品的电源,在移动电源市场上有一定的市场份额,具有一定的经济价值。然而,在干电池的生产中,底盖缺陷问题很常见,而且用肉眼进行人工分拣效率较低。为了快速准确地分拣底盖缺陷产品,采集了 1568 张干电池底盖图像,共分为 4 种类别,并将其随机划分为训练集、验证集和测试集,比例分别为 70%、10% 和 20%。通过将 CBAM 注意力机制嵌入 MobileNetV2 网络模型的倒残差结构中,成功构建了 CBAM_MobileNetV2 模型。实验结果表明,CBAM_MobileNetV2 模型的准确率达到了 95.78%,相较于 MobileNetV2 提高了 1.19%,优化效果较为明显。

关键词

MobileNetV2; CBAM; 缺陷检测; 干电池底盖; 深度学习

doi: 10.3969/j.issn.1672-9528.2024.03.040

0 引言

在干电池的生产过程中,由于生产工艺和原材料等因素,会出现不合格产品的底盖问题。这些问题可能导致电池接触不良、漏液,甚至引发电池爆炸等严重后果。因此,在生产工艺中,需要增加人工肉眼检测环节来筛选不合格的干电池底盖产品。然而,人工肉眼检测方式不仅容易受到主观影响,而且需要支付高昂的人工费用。长时间进行人工肉眼检测还会导致工作疲劳,进而降低检测效率。此外,由于生产工艺已实现自动化,人工肉眼检测将会影响生产进度。因此,提高干电池底盖缺陷检测效率变得十分紧迫,需要找到一种方法来解决这个问题。

近年来,在电池领域,卷积神经网络(convolutional neural networks, CNN)的应用越来越广泛。学者们利用 CNN 技术进行电池目标检测和图像分类等研究。阎旭^[1]提出了一种基于机械视觉的电池尾端缺陷检测算法,该算法可以有效替代人工肉眼检测。周颖等人^[2]改进了 YOLOv8 模型,用于光伏电池缺陷检测。他们在主干网络和特征融合网络中引入上下文聚合模块,并使用 GauGAN 进行数据增强,以提高模型的泛化能力。实验证明,该模型的平均精度达到了 89.9%。田杰等人^[3]提出了一种基于改进的 YOLOv4 的柱形锂电池缺陷检测方法,他们将 CSPDarkNet53 替换为轻量化网络 MobileNetV1。改进后的模型不仅提高了 0.2% 的检测精度,而且检测速度提高了 32.97 帧/s。

MobileNetV2^[4]是 Google 在 2018 年推出的一种轻量级卷

积神经网络模型。张鹏程等人^[5]使用改进的 MobileNetV2 模型对柑橘害虫图像进行分类,并取得了 93.63% 的准确率。另外,黄乾峰等人^[6]将改进后的 MobileNetV2 模型应用于番茄叶片病害种类识别,实验结果显示,相比原始模型,准确率提高了 2.29%,达到了 98.57%,有效地对番茄病叶片进行了分类。陈宗阳等人^[7]在 MobileNetv2 网络模型基础上设计出递进式分类器和跨局部连接结构,实现了高精度的快速识别操作。实验结果显示,该方法在涂层表面缺陷检测中取得了很好的准确率表现,其准确率达到 99.43%。巩文静等人^[8]提出对 MobileNetv2 网络进行去除部分网络层并调整通道数,通过数据降维和增加全连接层得到分类结果,并设计了一个融合网络结构,将声光图像作为输入特征进行有效提取,实验数据表明,改进的 MobilenetV2 网络在水下目标分类方面性能更好,融合网络相比之前的方法有更高的分类准确率,更适用于水下目标分类任务。

本文将干电池底盖缺陷检测研究成果部署到工业相机中实现缺陷检测,考虑到模型大小等因素,本文选择了轻量级神经网络 MobileNetV2 作为研究对象,并在其基础上添加了 CBAM 注意力机制。利用通道注意力模块提高特征图的表征能力,同时利用空间注意力模块强化模型对干电池底盖图像局部区域的关注程度,从而提高对干电池底盖的检测精度。

1 MobileNetV2 网络

1.1 MobileNetV2 模型介绍

MobileNetV2 是在 MobileNetV1 基础上提出的 CNN 架构,主要由深度可分离卷积和倒残差结构组成。如图 1 所示,深度可分离卷积将卷积操作分解为深度卷积(Depthwise

1. 广东工业大学 广东广州 510006

Convolution) 和逐点卷积。深度卷积是指在输入数据的每个通道上分别进行卷积操作, 即对每个通道应用一个独立的卷积核。这样可以减少模型中参数的数量, 这种分离的卷积操作可以更好地捕捉输入数据在通道维度上的特征。逐点卷积 (Pointwise Convolution) 是指在深度卷积之后, 使用 1×1 的卷积核对通道维度上的特征进行整合和组合。再对深度卷积的结果进行线性组合, 通过引入非线性激活函数, 进一步提取和组合特征。通过深度可分离卷积, 大大减少了计算量和参数数量。

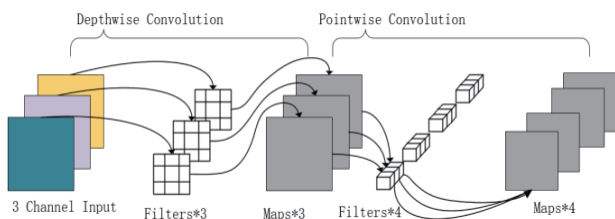


图1 深度可分离卷积

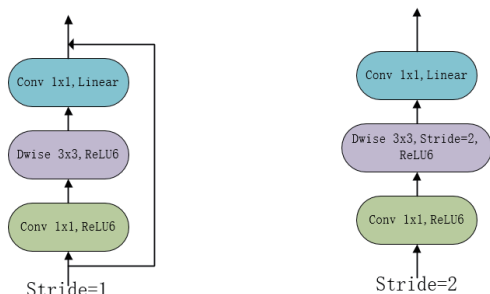


图2 倒残差结构

如图2所示, 倒残差结构采用先升维后降维的设计, 它能够有效地捕捉更丰富的特征信息。其中, 该结构由两个深度可分离卷积组成, 第一个深度可分离卷积用于增加通道数, 第二个深度可分离卷积用于恢复通道数。这种设计策略能够将输入特征图的通道数扩张到更多, 通过逐点卷积恢复通道数, 确保输入和输出通道数相同, 避免信息丢失, 从而提升了模型的表达能力。此外, 在步长为1的倒残差块中, 输入和输出的维度相同, 通过将最后一次逐点卷积的结果与输入相加, 实现了残差连接, 避免了梯度消失的问题。MobileNetV2 网络模型中的倒残差结构呈现两头小中间大的形式, 因此, 模型最后输出的是一个低维的特征信息。为了避免大量丢失低维特征信息, 模型在最后一个卷积层采用线性激活函数。

1.2 CBAM 注意力机制

注意力机制是一种模拟人类视觉的机制, 通过关注重要特征并抑制不必要的特征来提高模型性能, 如图3所示。CBAM (convolutional block attention module) 作为一种轻量级的卷积注意力模块, 在引入卷积神经网络中可以显著提升

模型性能。同时, CBAM 的引入对网络的参数量和计算量的增加非常小, 因此可以嵌入到大多数卷积神经网络中。

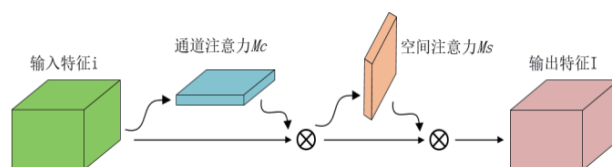


图3 CBAM 结构图

特征图的不同维度承载着不同的信息。通道维度主要用于表达特征的抽象性, 而空间维度则更关注物体的位置信息。CBAM 是一种融合了通道注意力模块 (channel attention module, CAM) 和空间注意力模块 (spatial attention module, SAM) 的机制。这两个子模块以串行的方式相互结合, 依次在通道维度和空间维度上生成注意力特征图信息。CAM 利用特征图中不同通道之间的关系生成通道注意力图, 然后将其与 SAM 连接起来, 得到最具信息量的特征图。

如图4所示, 在通道注意力 CAM 模块中, 通过平均化层和最大化层对输入特征的每个通道特征进行平均和最大值提取。然后, 这些特征值通过多层感知机 (由卷积和激活函数组成) 进行处理, 从而得到每个通道的新特征表示。最后, 将这些新特征进行相加融合, 并经过 sigmoid 激活函数操作, 生成最终的通道注意力模块特征。

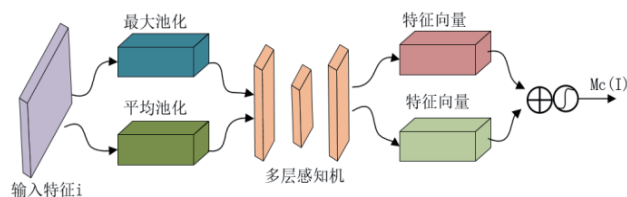


图4 通道注意力模块

如图5所示, 在空间注意力模块中, 第一层结构包括最大化层和平均化层, 将通道注意力模块的特征转换为两种不同的特征, 并将其扩展为双倍的通道数。第二层结构是一个二维卷积层, 将上一层得到的特征进行降维, 使其恢复为原始的通道数。随后, 经过第三层的 sigmoid 激活函数, 得到空间注意力模块的特征, 这个特征表示了特征图中每个空间位置的权重。最后, 将空间注意力权重与原始输入特征相乘, 得到最终的特征表示。

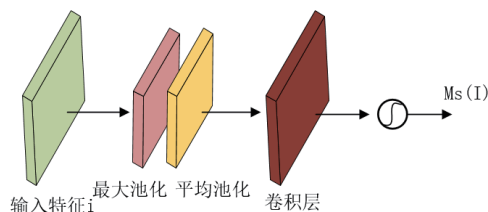


图5 空间注意力模块

本文在 MobileNetV2 的倒残差结构中引入了 CBAM

注意力机制，构建了 CBAM-MobileNetV2 模型，如图 6 所示。该模型能够增强每个倒残差块内特征的学习和表示能力，进一步提高对复杂特征的捕获能力。同时，它还能够更有效地利用每个残差块内的信息，提高特征的区分性和表达能力。

2 实验过程与分析

2.1 实验数据集来源

本文所使用的干电池底盖实验数据集由 QuickThinkVC® Type-B 工业相机采集，采集地点位于东莞市某干电池工厂。总共收集 4 种干电池底盖缺陷类别，如底盖生锈、收缩破损、外壳生锈、无底盖，如图 7 所示。由于工厂生产环境出现不良品数量占比较小，总共采集 1568 张，其中各种类数量如表 1 所示。

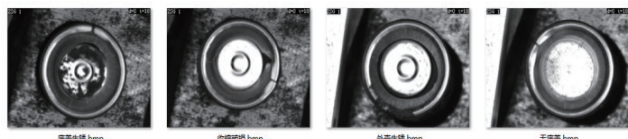


图 7 干电池底盖缺陷分类

本文将干电池底盖数据集随机划分为训练集、验证集和测试集，占比分别为 70%、10%、20%，其中训练集包含 1097 张图片，验证集包含 157 张图片，测试集包含 314 张图片。

表 1 干电池各种类数量

序号	类别	数量
1	底盖生锈	432
2	收缩破损	376
3	外壳生锈	408
4	无底盖	352

2.2 数据增强

由于本文收集的干电池底盖数据存在数量有限和数据多样性较低等问题，为了提高训练模型的泛化能力和分类识别性能，并避免训练样本过拟合问题，本文对 1097 张训练集图片进行了数据增强。共采用了 5 种数据增强方法，其中前两种是常见的数据增强方法，可扩充数据集而不影响图像本身，后三种方法基于图像颜色通道变化，旨在填补工厂生产环境中可能存在的光照问题，减少环境光照和图片亮度之间的差异对模型的干扰。鉴于随机裁剪或平移数据增强方法可能导致干电池底盖图片缺失或不完整，本文未采用这些方法。通过数据增强，最终获得 6856 张干电池底盖样本，以应对数据量不足和缺乏多样性的问题。

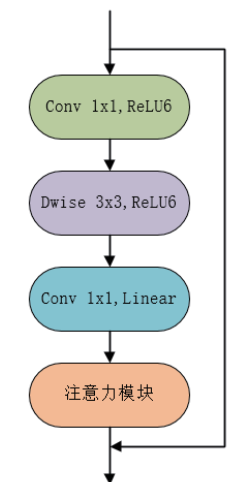


图 6 改进后 MobileNetV2 网络的倒残差结构

表 2 数据增强操作方法

操作名	操作说明	随机范围
Flip	对称翻转	1
Rotate	旋转角度	$[-35, 35]$
Sharpness	锐化值	$[0.2, 0.8]$
Brightness	亮度	$[0.2, 0.8]$
Color Transformation	色彩变换	$[3, 9]$

2.3 实验准备及设备

在对模型进行训练之前，需要进行一系列预处理步骤。首先，对图像尺寸进行统一设置，以确保输入图像具有相同的尺寸。其次，选择优化器为 Adam，通过历史梯度数据对模型的变量进行优化调整。为了防止梯度爆炸的问题，初始学习率 (learning rate) 被设置为 0.001。

为了训练出具有更好分类能力的模型，本文选择了 100 个训练轮次 (epochs)，每个轮次中的训练样本批次大小 (batch size) 为 32。

在实验过程中，使用了特定的软硬件设备，其详细信息如表 3 所示。

表 3 软硬件实验设备

硬件	软件
CPU: 12th Gen Intel(R) Core(TM) i5-12400 2.50 GHz	TensorFlow_GPU 2.2.0
GPU: NVIDIA GeForce GTX 970	CUDA 10.1
内存: 32 GB	CUDNN 7.6
显存: 4 GB	Python 3.8
OS: Windows10 专业版	Anaconda3

2.4 分类评价指标

在本文的干电池底盖分类实验中，需要对训练好的模型进行性能评估。为了准确地度量模型预测结果与真实结果之间的误差，采用了以下四个常见的二分类评价指标：准确率 (Accuracy, Acc)、精确率 (Precision, Pre)、召回率 (Recall, Rec) 和 F_1 分数。这些评价指标的计算公式为：

$$\text{Accuracy} = \frac{T_p + T_n}{T_p + F_p + F_n + T_n} \quad (1)$$

$$\text{Precision} = \frac{T_p}{T_p + F_p} \quad (2)$$

$$\text{Recall} = \frac{T_p}{T_p + F_n} \quad (3)$$

$$F_1 = \frac{2 \otimes \text{Precision} \otimes \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

式中： T_p 表示模型正确预测为正样本的数量，即将干电池底盖图片分类到真正的类别中； T_n 表示模型正确预测为负样本的数量，即将干电池底盖图片分类到错误的类别中； F_p 表示

模型错误地将负样本分类为正样本的数量，即将错误的干电池底盖图片分类到真正的类别中； F_N 表示模型错误地将正样本分类为负样本的数量，即将真正的干电池底盖图片分类到错误的类别中。

2.5 结果分析

将 MobileNetV2 网络添加 CBAM 注意力机制后，对网络性能的影响进行了探究。对 MobileNetV2 和 CBAM_MobileNetV2 两种网络模型在本文数据集上进行了测试，得到它们的准确率和损失值如图 8 和图 9 所示。从图 8 可以看出，CBAM_MobileNetV2 在 epoch 为 50 时逐渐收敛，而 MobileNetV2 在 epoch 为 58 时才逐渐收敛，这表明改进后的模型收敛速度更快，具有更强的泛化能力。相较于 MobileNetV2 网络模型准确率曲线，CBAM_MobileNetV2 网络模型准确率曲线居上，表明具有更高的准确率。从图 9 可以得出结论，改进后的 CBAM_MobileNetV2 网络模型的损失曲线收敛更快，表现出更好的稳定性。

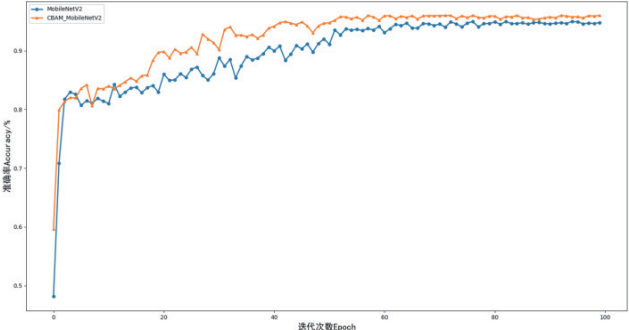


图 8 不同模型的准确率

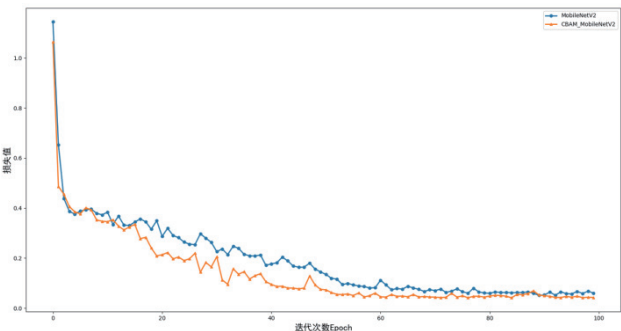


图 9 不同模型的损失值

通过对比两种网络模型的性能评价指标，如表 4 所示，可以观察到 CBAM_MobileNetV2 模型在干电池底盖缺陷检测分类任务中相较于 MobileNetV2 模型有了一定的提升。具体而言，CBAM_MobileNetV2 模型的准确率、召回率和精确率分别提高了 1.19%、2.54% 和 1.82%。这意味着通过添加 CBAM 注意力机制，改进后的模型在识别干电池底盖缺陷方面的性能得到了改善。不过，改进后的 CBAM_MobileNetV2 模型的参数量增加了 0.39 MB，这导致了模型的运行速度略微下降。在实际应用中，需要权衡模型性能和运行速度之

间的平衡。如果时间要求紧迫，考虑使用 MobileNetV2 模型；而如果追求更高的分类准确率和综合性能，可以选择 CBAM_MobileNetV2 模型。

表 4 改进后的模型性能评价指标对比

模型	准确率	召回率	精确率	参数量
MobileNetV2	94.59%	93.27%	94.03%	2.17 MB
CBAM_MobileNetV2	95.78%	95.81%	95.85%	2.56 MB

本文对比了 EfficientNet-B0^[9]、ShuffleNetV2^[10]、ResNet50^[11]、SqueezeNet^[12]、MobileNetV2 和添加 CBAM 注意力机制后的 CBAM_MobileNetV2 这 6 种卷积神经网络模型，并进行了性能评价。根据实验结果，如表 5 所示，可以得出 CBAM_MobileNetV2 表现最好，具有最高的准确率，达到了 95.78%。相比于 ShuffleNetV2、EfficientNet-B0、ResNet50、SqueezeNet 和 MobileNetV2，CBAM_MobileNetV2 的准确率分别提高了 11.22%、2.06%、17.11%、3.92% 和 1.19%。

表 5 不同模型性能评价指标对比

模型	准确率	召回率	精确率	F_1 分数
ShuffleNetV2	84.56%	83.53%	85.28%	84.40%
EfficientNet-B0	93.72%	93.74%	93.73%	93.73%
ResNet50	78.67%	77.95%	79.34%	78.64%
SqueezeNet	91.86%	91.97%	91.93%	91.95%
MobileNetV2	94.59%	93.27%	94.03%	93.65%
CBAM_MobileNetV2	95.78%	95.81%	95.85%	95.83%

3 结论

本文通过在 MobileNetV2 网络基础上添加 CBAM 注意力机制，改进了一种用于干电池底盖缺陷检测的模型。根据实验结果，CBAM_MobileNetV2 网络模型达到了 95.78% 的准确率和 95.85% 的精确度，能够高效地对干电池底盖缺陷进行分类识别。该网络模型具有轻量级特性，适合部署在嵌入式设备上，如工业流水线上的工业相机，用于检测干电池底盖的缺陷。因此，本文所改进的 CBAM_MobileNetV2 网络模型具有重要的研究意义和实际应用价值。

参考文献：

[1] 阎旭. 基于机器视觉的电池尾端缺陷检测算法的研究 [D]. 成都: 电子科技大学, 2013.

[2] 周颖, 颜毓泽, 陈海永, 等. 基于改进 YOLOv8 的光伏电池缺陷检测 [J/OL]. 激光与光电子学进展, 1-17 [2023-11-19]. <http://kns.cnki.net/kcms/detail/31.1690.tn. 20230821. 1446.128.html>.

[3] 田杰, 胡昊, 周华健, 等. 基于改进 YOLOv4 的柱形锂电池缺陷检测研究 [J]. 机械工程师, 2023(3):16-18.

(下转第 174 页)

体系,使得不同领域的技术能够更好地互联互通^[7]。

GB/T 28181—2022 标准内容繁杂而庞大,其涵盖的范围广泛且复杂,是一份内容密度大、难以轻松消化的文件。相较于 GB/T 28181—2016 标准,新版标准进行了较多的新增与修改,这使得读者需要投入更多的耐心和时间来深入阅读和理解。考虑到标准的全面性和深度,本文的介绍及问题分析只是冰山一角,在有限的篇幅简要介绍了一些在实际应用中可能遇到的相关问题。然而,想要全面掌握和运用新标准的内容,读者仍需深入研读标准文本,逐一理解其中的技术规范和要求,并且结合实际场景进行深入探讨,以更好地应对实际工作中的问题和挑战。在理解和运用 GB/T 28181—2022 标准的过程中,不断学习和实践将有助于读者更好地适应并掌握这一技术规范,从而更有效地推动相关行业的发展。

参考文献:

- [1] 陈朝武. GB/T 28181 国标的重大意义、推广应用及技术发展[J]. 中国公共安全, 2013(19): 208-212.
- [2] IETF. Session Initiation Protocol: RFC 3261[S/OL]. [2023-10-08]. <https://www.rfc-editor.org/rfc/rfc3261>.
- [3] 孔维生, 周鉴刚, 梁雄高, 等. GB/T 28181—2022 解读[J]. 中国安防, 2023(5): 35-42.
- [4] IETF. Use of the Content-Disposition Header Field in the Hypertext Transfer Protocol (HTTP): RFC 6266[S/OL]. [2023-10-12]. <https://www.rfc-editor.org/rfc/rfc6266>.
- [5] IETF. SIP-Specific Event Notification: RFC 6665[S/OL]. [2023-11-26]. <https://www.rfc-editor.org/rfc/rfc6665>.
- [6] 国家市场监督管理总局, 国家标准化管理委员会. 公共安全视频监控联网系统信息传输、交换、控制技术要求: GB/T 28181—2022[S/OL]. [2023-09-26]. <https://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=8BBC2475624A6C31D-C34A28052B3923D>.
- [7] 吴松洋, 齐力, 唐世杰. 视频监控安全发展及展望[J]. 中国安防, 2023(6): 48-51.
-
- (上接第 169 页)
- [4] MARK S, ANDREW H, MENGLONG Z, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2018: 4510-4520.
- [5] 张鹏程, 余勇华, 陈传武, 等. 基于改进 MobileNetV2 的柑橘害虫分类识别方法[J]. 华中农业大学学报, 2023, 42(3): 161-168.
- [6] 黄乾峰, 董琴, 韦静. 改进 MobileNetV2 算法的番茄叶片病害种类识别[J]. 计算机系统应用, 2023, 32(1): 385-391.
- [7] 陈宗阳, 赵辉, 吕永胜, 等. 基于改进 MobileNetV2 网络的涂层表面缺陷识别方法[J]. 哈尔滨工程大学学报, 2022, 43(4): 572-579.
- [8] 巩文静, 田杰, 李宝奇, 等. 基于改进 MobilenetV2 网络的声光图像融合水下目标分类方法[J]. 应用声学, 2022, 41(3): 462-470.
- [9] TAN M X, LE Q V. EfficientNet: rethinking model scaling for convolutional neural networks[C]// Proceedings of the 36th International Conference on Machine Learning (ICML). New York: Curran Associates, Inc., 2019: 10691-10700.
- [10] MA N N, ZHANG X Y, ZHENG H T, et al. ShuffleNet V2: practical guidelines for efficient CNN architecture design[C]// Proceedings of the European Conference on Computer Vision (ECCV). Berlin: Springer, 2018: 122-138.
- [11] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE conference on computer vision and pattern recognition. Piscataway: IEEE, 2016: 770-778.
- [12] FORREST N I, SONG H, MATTHEW W M, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size[EB/OL]. (2016-11-04) [2023-11-23]. <https://arXiv preprint arXiv:1602.07360>.

【作者简介】

高学秋(1997—), 男, 广东湛江人, 硕士研究生, 研究方向: 目标检测与深度学习。

张祺(1975—), 男, 广东韶关人, 博士, 副教授, 研究方向: 工业 AI 机器视觉软件。

吴泽强(1999—), 男, 广东揭阳人, 硕士研究生, 研究方向: 目标检测与深度学习。

钟善机(1998—), 男, 广东湛江人, 硕士研究生, 研究方向: 语音情感识别。

刘磊(1996—), 男, 河南鹤壁人, 硕士研究生, 研究方向: 深度学习。

(收稿日期: 2024-01-05)