

◎图形图像处理◎

基于轻量级卷积神经网络的载波芯片缺陷检测

周天宇¹, 朱启兵¹, 黄敏¹, 徐晓祥²

1. 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

2. 无锡市创凯电气控制设备有限公司, 江苏 无锡 214400

摘要:载波芯片(chip on carrier, COC)是光发射次模块(transmitter optical subassembly, TOSA)的重要组成部分,被广泛应用于光通信领域,实现光电转换。针对载波芯片崩口、定位柱破损以及波导污渍三种不同类别缺陷的实时检测问题,提出了一种基于轻量级卷积神经网络的载波芯片缺陷检测算法YOLO-Efficientnet。为了减少网络参数,缩短检测时间,采用轻量级卷积神经网络Efficientnet作为主干网络对图像进行特征提取,在移动翻转瓶颈卷积(MBConv)的基础上,引入了压缩与激发网络(SENet)的注意力思想,在通道维度上引入注意力机制;为了解决下采样的过程中导致信息丢失的问题,引入空间金字塔池化(SPP)结构来增大图像的感受野,分离出更加显著的上下文特征。针对COC缺陷多尺度以及波导区域污渍小目标难以检测的问题,引入了PANet结构进行多尺度特征融合。实验结果表明,提出的算法对COC缺陷检测的准确率达到98.5%,检测时间达到每张图片0.42 s,满足实时检测的需求。

关键词:计算机视觉;载波芯片缺陷检测;目标检测;YOLO-Efficientnet;光发射次模块

文献标志码:A **中图分类号:**TP391.4 **doi:**10.3778/j.issn.1002-8331.2104-0299

Defect Detection of Chip on Carrier Based on Lightweight Convolutional Neural Network

ZHOU Tianyu¹, ZHU Qibing¹, HUANG Min¹, XU Xiaoxiang²

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

2. Wuxi CK Electric Control Equipment Co., Ltd., Wuxi, Jiangsu 214400, China

Abstract: Chip on carrier (COC) is an important component of transmitter optical subassembly (TOSA) and is widely used in the field of optical communication to realize photoelectric conversion. Aiming at the problem of real-time detection of three different types of defects on COC, such as collapse, positioning column damage and waveguide stain, a defect detection algorithm of COC based on lightweight convolution neural network, YOLO-Efficientnet, is proposed. Firstly, in order to reduce network parameters and shorten detection time, lightweight convolution neural network, efficientnet, is used for image feature extraction as a backbone network. On the basis of MBConv, the attention idea of SENet is introduced, and the attention mechanism is introduced in the channel dimension. Secondly, in order to solve the problem of information loss during the downsampling process, the spatial pyramid pooling (SPP) structure is introduced to increase the receptive field of the image and separate more significant context features. Finally, to solve the problems of multi-scale COC defects and difficult detection of small targets in waveguide region, PANet structure is introduced for multi-scale feature fusion. Results of the experiments show that the algorithm proposed in this paper has an accuracy of 98.5% for COC defect detection, and the detection time reaches 0.42 seconds per picture, meeting the requirements of real-time detection.

Key words: computer vision; defect detection of chip on carrier; object detection; YOLO-Efficientnet; transmitter optical subassembly

基金项目:国家自然科学基金(61775086)。

作者简介:周天宇(1996—),男,硕士研究生,主要研究方向为深度学习目标检测;朱启兵(1973—),通信作者,男,博士后,教授,博士生导师,研究方向为传感与检测技术、信息感知与智能处理、物联网系统集成, E-mail: zhuqib@163.com。

收稿日期:2021-04-25 **修回日期:**2021-07-08 **文章编号:**1002-8331(2022)07-0213-07

光发射次模块(transmitter optical subassembly, TOSA)可以实现光电信号的转换,在光通信领域具有广泛的应用。作为TOSA的重要组成部分,载波芯片(chip on carrier, COC)的质量好坏对TOSA是否能够长期稳定地工作有着重大的影响。由于制造工艺和生产环境的影响,崩口、定位柱破损以及波导区域污渍是COC生产过程中经常出现的三种典型缺陷。研究高效、快速而准确的COC缺陷自动检测方法,对COC质量控制与故障检测具有重要的实际意义。

随着深度学习的发展,越来越多的基于深度卷积神经网络的目标检测方法被提出^[1-4],这些方法主要可以分为两类:一类是双阶段(two stage)目标检测算法;另一类是单阶段(one stage)目标检测算法。其中,双阶段算法是基于区域推荐的分类方法实现的,代表性算法有R-CNN^[5]、Fast R-CNN^[6]、Faster R-CNN^[7]、Mask R-CNN^[8]、Cascade R-CNN^[9]和Sparse R-CNN^[10]等。单阶段算法是基于端到端的回归方法实现的,代表性算法有SSD^[11]、YOLO^[12]、YOLOv2^[13]、YOLOv3^[14]和YOLOv4^[15]等。双阶段分类算法具有较高的检测准确性而单阶段回归算法具有更快的检测速度,随着众多算法的提出,这些方法的检测效果在不断提高,尤其是YOLO系列算法经过不断的更新优化,网络性能进一步提升。

基于深度卷积神经网络的目标检测算法被广泛应用于工业产品的表面缺陷检测。黄凤荣等人^[16]提出了一种基于Faster R-CNN的发动机零件表面缺陷检测算法;Chen等人^[17]将生成对抗网络与YOLOv3算法进行结

合,研究了一种芯片晶粒的表面缺陷检测方法。Perez等人^[18]提出了一种基于VGG-16的预训练卷积神经网络分类器,并结合CAM网络实现建筑物表面缺陷的检测和定位;金颖等人^[19]研究了一种基于Mask R-CNN的柱塞式制动主缸内槽表面缺陷检测系统。

现有基于深度卷积网络的表面缺陷检测算法为了取得更好的检测准确性,通常依赖于大型深度网络,存在参数量过多、检测时间慢的问题,难以满足COC缺陷检测实时性的需求。针对此问题,本文提出了一种基于轻量级卷积神经网络的COC缺陷检测算法YOLO-Efficientnet。

首先,借鉴了YOLOv3网络端到端单阶段检测的思想,引入了轻量级特征提取网络Efficientnet^[20]作为主干网络对图像进行特征提取。其次,为了解决下采样的过程中导致信息丢失的问题,引入SPP-Net^[21]结构来增大图像的感受野,分离出更加显著的上下文特征。最后,针对COC缺陷多尺度以及波导区域污渍小目标难以检测的问题,引入了PANet^[22]结构进行多尺度特征融合。本文提出的基于轻量级卷积神经网络的COC缺陷检测算法YOLO-Efficientnet在保证检测准确性的同时,减少了网络的参数量,缩短了检测时间,能够对COC缺陷进行高效、快速而准确地检测。

1 YOLO-Efficientnet 网络结构

本文提出的YOLO-Efficientnet算法的网络结构如图1所示,为了减少网络参数,缩短检测时间,主干网络

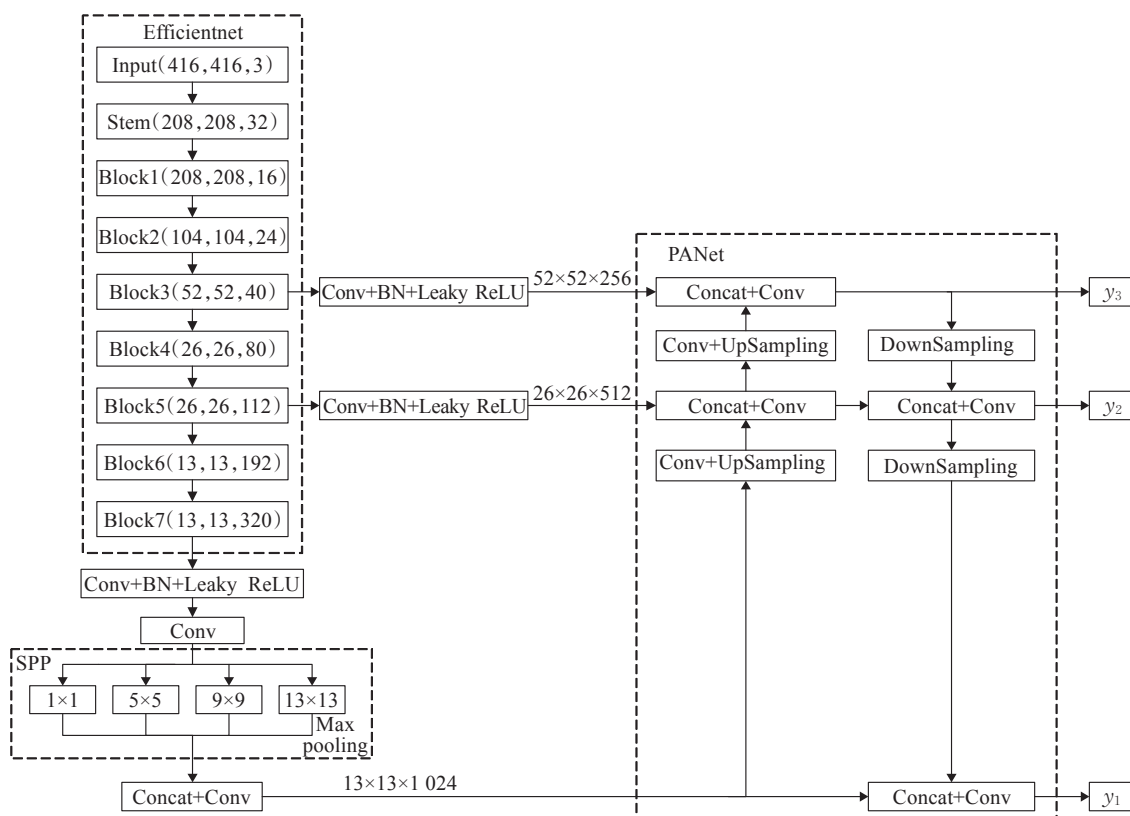


图1 YOLO-Efficientnet 网络结构

Fig.1 Network structure of YOLO-Efficientnet

采用轻量级卷积神经网络 Efficientnet, 对尺寸为 416×416 的输入图片进行 8 倍、16 倍以及 32 倍的下采样特征提取, 得到 52×52 、 26×26 和 13×13 这 3 种不同尺度的特征图。为了解决 Efficientnet 对输入图片进行特征提取下采样过程导致的信息丢失问题, 增大图像的感受野, 在 Block7 之后引入 SPP-Net 结构, 分别使用大小为 13×13 、 9×9 、 5×5 和 1×1 的池化核进行最大池化处理, 分离出更加显著的上下文特征。针对 COC 缺陷多尺度以及波导区域污渍小目标难以检测的问题, 引入 PANet 对 52×52 、 26×26 和 13×13 这 3 个不同尺度的特征图进行增强特征融合, 在原有的特征金字塔自上而下 (top-down) 特征融合的基础上, 增加自下而上 (bottom-up) 的增强路径, 通过对浅层特征图和深层特征图的反复提取, 更加充分地融合不同深度的特征信息。最终得到 3 个不同尺度的特征图输出 y_1 、 y_2 和 y_3 , 实现多尺度预测。

在预测阶段, 对于每张输入图像, 预测包括目标边界框位置、置信度和目标类别的三维张量。将 y_1 、 y_2 和 y_3 这 3 种尺寸分别为 13×13 、 26×26 和 52×52 特征图划分为 $S \times S$ 个单元格 (S 分别取 13、26 和 52), 给每个网格设置 3 个先验框来预测 3 个边界框, 输出维度为 $S \times S \times (3 \times (4 + 1 + M))$, 即 4 个边界框的偏移量、1 个缺陷目标置信度和 M 个缺陷类型。最后通过非极大值抑制 (non-maximum suppression, NMS) 的方法剔除重复的边界框, 从而实现目标检测。

2 Efficientnet 轻量级网络

近年来, 随着卷积神经网络的发展, 为了提升网络性能, 提高模型的泛化能力, 通常通过网络的深度、网络的宽度或者网络的分辨率这 3 个维度实现网络缩放。其中, 网络的深度是指增加网络的层数, 通过更深的神经网络来增强特征提取能力; 网络的宽度是指加深网络的通道数, 增加每一层提取的特征层数, 得到更多的特征; 网络的分辨率是指通过增大输入图片的分辨率来减少图像信息损失, 使得网络能够学习更加丰富的特征。然而, 在前人的研究中, 大多只通过放大其中的一个维度以达到更高的准确率, 例如目前常用的残差网络 ResNet^[23], 从 ResNet-18 到 ResNet-200, 通过将网络深度从 18 层增加到 200 层来提高网络的准确率。3 个维度中任一维度的放大都可以带来精度的提升, 但随着倍率越来越大, 提升却越来越小, 同时会引入更多的参数, 增大模型的计算量, 模型准确率也会很快达到饱和。

Google2019 提出的轻量级网络 Efficientnet^[20], 利用复合模型缩放 (compound model scaling) 方法, 对网络的深度、宽度以及分辨率进行同时缩放, 综合优化 3 个维度提升网络指标, 在提高网络准确率的情况下, 大大

减少了模型的数量和计算量。Efficientnet 网络在不同的缩放比例条件下, 可得到不同的模型结构 (Efficient-B0~Efficient-B7)。

2.1 Efficientnet-B3

综合考虑检测时间和检测准确性, 本文选用 Efficientnet-B3 作为主干网络对输入图像进行特征提取。Efficientnet-B3 的网络结构如表 1 所示, Efficientnet-B3 由 1 个 Stem 层和 7 个 Block 层组成。Stem 层是 1 个 3×3 的卷积, Block 层由不断堆叠的倒残差模块 MBConv-Block 组成, 其中, MBConvBlock1 和 MBConvBlock6 分别表示其扩展比例为 1 和 6。在 Block3、Block5 和 Block7 后输出 52×52 、 26×26 和 13×13 这 3 种不同尺度的特征图。

表 1 Efficientnet-B3 网络结构

Table 1 Network structure of Efficientnet-B3

Stage	Operator	$k \times k$	Stride	Outputs	Layers
Stem	Conv	3×3	2	$208 \times 208 \times 32$	$\times 1$
Block1	MBConv Block1	3×3	1	$208 \times 208 \times 16$	$\times 2$
Block2	MBConv Block6	3×3	2	$104 \times 104 \times 24$	$\times 3$
Block3	MBConv Block6	5×5	2	$52 \times 52 \times 40$	$\times 3$
Block4	MBConv Block6	3×3	2	$26 \times 26 \times 80$	$\times 5$
Block5	MBConv Block6	5×5	1	$26 \times 26 \times 112$	$\times 5$
Block6	MBConv Block6	5×5	2	$13 \times 13 \times 192$	$\times 6$
Block7	MBConv Block6	3×3	1	$13 \times 13 \times 320$	$\times 2$

2.2 MBConv Block

作为 Efficientnet 的基本模块, 倒残差模块 MBConv-Block 在移动翻转瓶颈卷积 (mobile inverted bottleneck convolution, MBConv) 模块的基础上, 引入了压缩与激发网络 (squeeze-and-excitation network, SENet) 的注意力思想, 在通道维度上引入注意力机制, 这种注意力机制让模型可以更加关注信息量最大的通道特征, 而抑制那些不重要的通道特征。

MBConv Block 的网络结构如图 2 所示。首先对输入图片进行 1×1 的卷积操作并根据扩展比例改变输出通道维度 (如扩展比例为 6 时, 会将通道维度提升 6 倍; 如果扩展比例为 1, 则直接省略该 1×1 的卷积和其之后批归一化和激活函数)。在此基础上, 利用 $k \times k$ 的卷积核对待扩展后的特征图在二维平面内进行逐通道深度卷积 (depthwise Conv) 操作; 并在通道维度上引入 SE module 注意力机制, 获得不同通道的权重对特征图进行通道相乘操作, 使模型可以更加关注信息量最大的通道特征, 而抑制那些不重要的通道特征。利用多个 1×1 的卷积核对待特征图进行逐点卷积 (pointwise Conv) 操作, 恢复原通道维度。最后进行输入的残差跳跃连接, 得到最终的特征图输出。

相比于常规卷积操作, MBConv Block 网络利用深度卷积和逐点卷积操作来提取特征, 具有更少的参数量

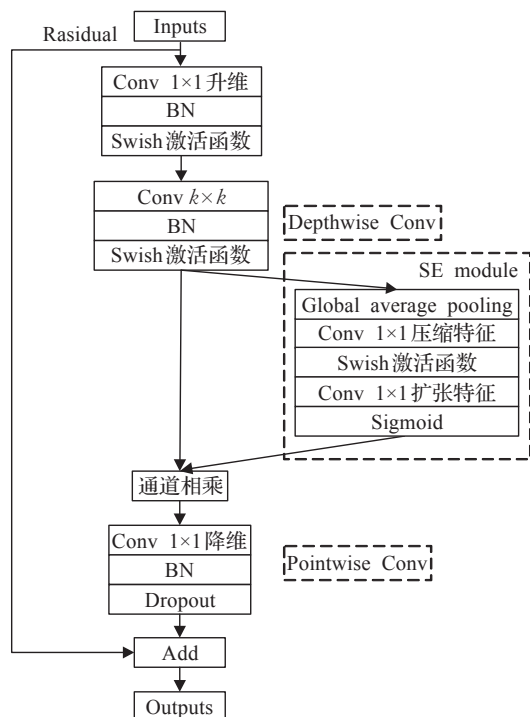


图2 MBConv Block网络结构

Fig.2 Network structure of MBConv Block

和运算量。同时,SE module注意力机制的引入有助于提高模型的精度。

3 SPP-Net

为了解决Efficient-B3对输入图片进行特征提取下采样的过程中导致信息丢失的问题,增大图像的感受野,如图3所示,在最后一层Block7之后引入SPP结构,分别使用大小为 13×13 、 9×9 、 5×5 和 1×1 的池化核进行最大池化处理,分离出更加显著的上下文特征,有利于实现多尺度预测。

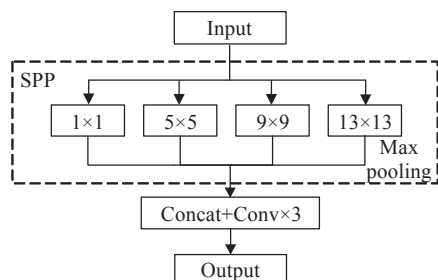


图3 SPP-Net网络结构

Fig.3 Network structure of SPP-Net

4 PANet增强特征融合

深层特征图包含更多的语义信息,注重轮廓特征,适合大目标的检测;浅层特征图位置信息,注重细节特征,适合小目标的检测。针对COC缺陷多尺度以及波导区域污渍小目标难以检测的问题,引入PANet对浅层特征图和深层特征图进行增强特征融合。

PANet的结构如图4所示,在原有的特征金字塔FPN自上而下(top-down)特征融合的基础上,增加自下而上(bottom-up)的增强路径,通过对浅层特征图和深层特征图的反复提取,更加充分地融合不同深度的特征信息。引入PANet结构,对 13×13 、 26×26 和 52×52 这3个不同尺度的特征图进行增强特征融合,最终得到3个不同尺度的输出 y_1 、 y_2 和 y_3 ,实现多尺度预测。

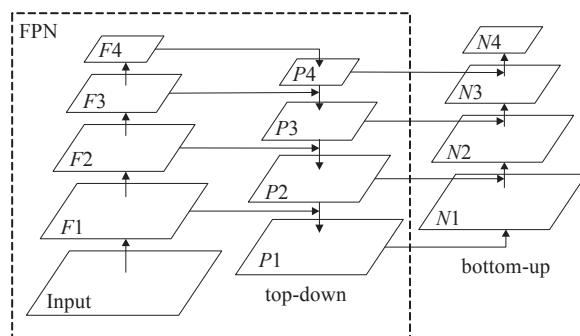


图4 PANet网络结构

Fig.4 Network structure of PANet

5 边界框预测

将3种尺度的特征图 y_1 、 y_2 和 y_3 分别划分为 13×13 、 26×26 和 52×52 个单元格,给每个单元格设置3个先验框来预测缺陷目标的边界框。边界框的预测过程如图5所示,虚线矩形框是预设先验框,宽为 p_w ,高为 p_h ,中心点在特征图上的坐标为 (c_x, c_y) 。网络通过预测缺陷的位置坐标的偏移量 t_x, t_y, t_w, t_h 从而得到缺陷的实际位置。实线矩形框是通过网络预测偏移量得到的预测边界框,中心点坐标为 (b_x, b_y) ,宽为 b_w ,高为 b_h 。计算公式如下:

$$b_x = \sigma(t_x) + c_x \quad (1)$$

$$b_y = \sigma(t_y) + c_y \quad (2)$$

$$b_w = p_w e^{t_w} \quad (3)$$

$$b_h = p_h e^{t_h} \quad (4)$$

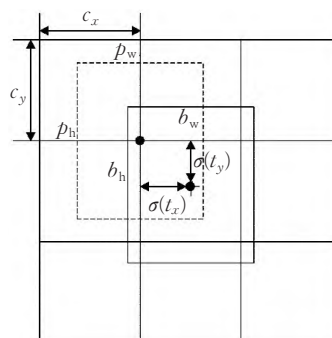


图5 边界框预测

Fig.5 Bounding box prediction

6 损失函数

本文提出的YOLO-Efficientnet算法的损失函数由

边界框误差 $loss_{coord}$ 、置信度误差 $loss_{IOU}$ 和类别误差 $loss_{class}$ 三部分组成,其中边界框误差采用的是均方误差,置信度误差和类别误差采用交叉熵误差。将尺寸分别为 13×13 、 26×26 和 52×52 的特征图 y_1 、 y_2 和 y_3 划分成 $S \times S$ 个单元格 (S 分别取 13、26 和 52),每个单元格预测 3 个边界框,损失函数为:

$$\begin{aligned} loss &= loss_{coord} + loss_{IOU} + loss_{class} = \\ \lambda_{coord} &\sum_{i=0}^{S \times S} \sum_{j=0}^K l_{ij}^{obj} \left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right] + \\ \lambda_{coord} &\sum_{i=0}^{S \times S} \sum_{j=0}^K l_{ij}^{obj} (2 - w_i h_i) \left[(w_i - \hat{w}_i)^2 + (h_i - \hat{h}_i)^2 \right] - \\ \lambda_{obj} &\sum_{i=0}^{S \times S} \sum_{j=0}^K l_{ij}^{obj} \left[\hat{c}_i \ln(c_i) + (1 - \hat{c}_i) \ln(1 - c_i) \right] - \\ \lambda_{noobj} &\sum_{i=0}^{S \times S} \sum_{j=0}^K l_{ij}^{noobj} \left[\hat{c}_i \ln(c_i) + (1 - \hat{c}_i) \ln(1 - c_i) \right] - \\ &\sum_{i=0}^{S \times S} \sum_{j=0}^K l_{ij}^{obj} \sum_{c \in M} \left[\hat{p}_i(c) \ln(p_i(c)) + (1 - \hat{p}_i(c)) \ln(1 - p_i(c)) \right] \quad (5) \end{aligned}$$

其中, λ_{coord} 是边界框预测的惩罚系数, λ_{obj} 和 λ_{noobj} 分别表示包含和不包含目标时置信度的惩罚系数。 l_{ij}^{obj} 和 l_{ij}^{noobj} 分别表示第 j 个候选边界框所在的第 i 个网格负责和不负责检测该目标; (x_i, y_i, w_i, h_i) 和 $(\hat{x}_i, \hat{y}_i, \hat{w}_i, \hat{h}_i)$ 分别表示预测边界框和真实边界框的中心点横坐标、纵坐标,边界框的宽度和高度; c_i 和 \hat{c}_i 分别表示预测置信度和真实置信度; M 表示目标类别的数量; $p_i(c)$ 和 $\hat{p}_i(c)$ 分别表示网格中的目标属于某个类别 c 的预测概率和真实概率。

通过构建边界框误差、置信度误差和类别误差的损失函数对网络进行训练,预测目标的位置信息和类别信息,最后,通过置信度阈值的设置来过滤掉得分较低的边界框,对剩下的边界框进行非极大值抑制操作来剔除重复的边界框,从而实现COC缺陷检测。

7 实验及结果分析

本文实验是基于TensorFlow的Keras深度学习框架实现。在硬件配置为Inter® Core™ i7-5930K CPU @ 3.50 GHz, 32 GB 内存, NVIDIA GeForce GTX TITAN X操作系统为Windows 10(64位)的服务器下进行。总训练轮次(epoch)为300,初始学习率设置为0.001,批量大小设置为16,权重衰减系数为0.000 5,动量设置为0.9,在第100次epoch之后减小为0.000 1。

7.1 数据集

本文的数据是COC经过200倍显微相机采集得到的图像,COC缺陷图像如图6所示,其中蓝色实线框表示崩口缺陷,黄色虚线框表示波导污渍缺陷,红色点划线框表示定位柱破损缺陷。

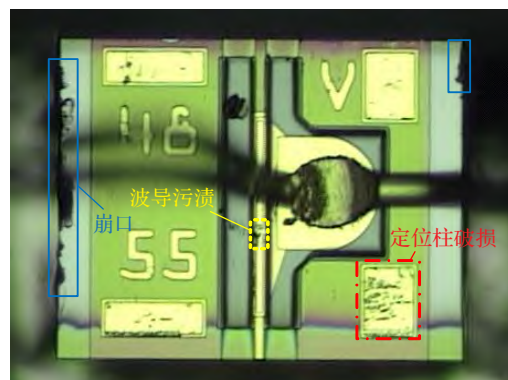


图6 COC缺陷

Fig.6 COC defect

COC缺陷图像共200张,选取150张进行训练,50张用作测试。对150张训练图像进行水平翻转、垂直翻转、提高亮度、引入高斯噪声以及锐化操作实现数据集扩充,得到4 800张训练样本。使用Labellmg图像标注软件进行人工标注,数据集格式为VOC2007。

7.2 实验结果与分析

将50张待检测COC缺陷图像输入训练好的模型进行测试,采用检测准确性、误检率和漏检率作为性能指标,其中误检率表示对缺陷的类别进行误判的概率,漏检率表示对缺陷未能检测到的概率。对波导区域污渍、崩口以及定位柱破损这3类缺陷进行检测统计的结果如表2所示。本文提出的YOLO-Efficientnet算法对COC波导区域污渍、崩口以及定位柱破损3类缺陷的检测准确率分别达到了100%、98%和96.3%,未出现误检情况,对COC缺陷的平均检测准确率达到98.5%。

表2 缺陷检测统计

Table 2 Defect detection statistics

缺陷类别	总数	正确检测	误检	漏检	准确率	误检率	漏检率
波导污渍	69	69	0	0	1.000	0	0
崩口	98	96	0	2	0.980	0	0.020
定位柱破损	27	26	0	1	0.963	0	0.037
合计	194	191	0	3	0.985	0	0.015

为了进一步验证YOLO-Efficientnet算法对COC缺陷检测的有效性,将本文算法与常用的目标检测算法YOLOv3、YOLOv4以及现有的轻量级算法YOLOv4-tiny在COC缺陷数据集上进行对比实验,实验参数采用默认参数。考虑到实际生产需求,采用COC缺陷平均检测准确率、参数量以及每张COC缺陷图像的平均检测时间作为衡量指标,实验结果如表3所示。

相比较传统的Faster R-CNN以及SSD算法,本文YOLO-Efficientnet算法的检测准确性大幅提高,同时缩短了检测时间。

与YOLOv3和YOLOv4算法相比,由于轻量级主干

表3 不同算法COC缺陷检测结果

Table 3 COC defect detection results of different algorithms

模型	主干网络	平均准确率/%	参数量/(params)/10 ⁶	检测时间/ms
Faster R-CNN	VGG-16	82.5	140.1	189
SSD	VGG-16	85.6	33.2	54
YOLOv3	Darknet-53	92.3	62.0	64
YOLOv4	CSPDarknet53	94.8	64.4	67
YOLOv4-Tiny	CSPDarknet53-Tiny	83.5	59.2	25
YOLO-Efficientnet	Efficientnet-B3	98.5	47.1	39

网络 Efficientnet-B3 引入, 本文算法具有更少的参数量, 缩短了检测时间; 同时, 由于引入 SPP-Net 结构和 PANet 增强特征融合, 本文算法取得了更高的准确率。

此外, 由于现有轻量级网络 YOLOv4-tiny 算法采用了主干网络 CSPDarknet53-tiny 以及简化版的图像金字

塔, 只对两种尺度的特征图进行特征融合, 因此, 虽然检测时间更短但是检测准确率大幅降低, 不能满足 COC 缺陷检测的需求。

以上实验表明, 本文提出的 YOLO-Efficientnet 算法, 对 COC 缺陷检测的准确率达到 98.5%, 每张 COC 缺陷图像的平均检测时间达到 39 ms, 满足 COC 缺陷实时检测的需求, 检测结果如图 7 所示。

8 结束语

本文针对 COC 的 3 类缺陷的实时检测问题, 提出了一种基于轻量级卷积神经网络的 COC 缺陷检测算法 YOLO-Efficientnet。借鉴 YOLOv3 目标检测的思想, 引入了轻量级卷积网络 Efficientnet 进行特征提取, 减少了网络参数, 缩短了检测时间。同时, 引入 SPP 结构增大

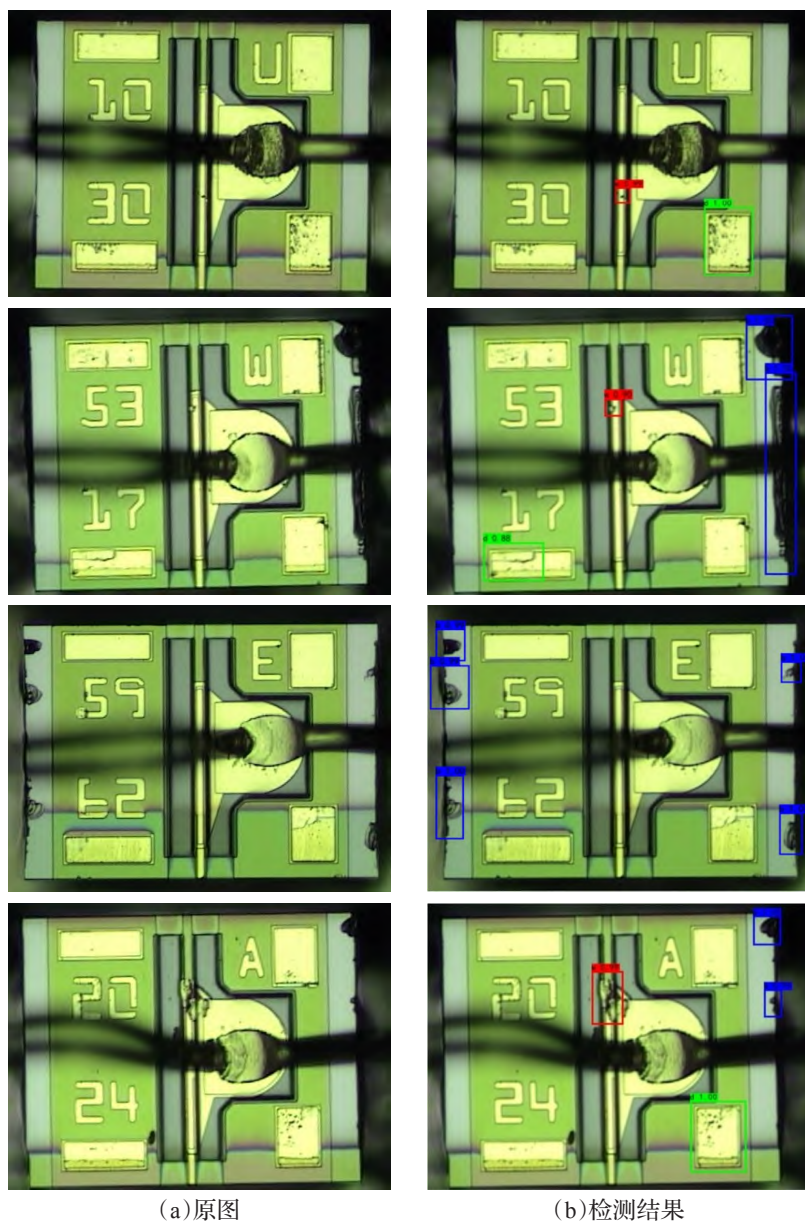


图7 基于 YOLO-Efficientnet 的 COC 缺陷检测结果图

Fig.7 COC defect detection results based on YOLO-Efficientnet

图像感受野,采用PANet进行增强特征融合,从而实现多尺度预测。实验结果表明,本文提出的算法对COC缺陷检测的准确率达到98.5%,检测时间达到每张图片0.42 s,满足COC缺陷高效、快速而准确检测的需求。

参考文献:

- [1] 黄健,张钢.深度卷积神经网络的目标检测算法综述[J].计算机工程与应用,2020,56(17):12-23.
HUANG J,ZHANG G.Survey of object detection algorithms for deep convolutional neural networks[J].Computer Engineering and Applications,2020,56(17):12-23.
- [2] 罗会兰,彭珊,陈鸿坤.目标检测难点问题最新研究进展综述[J].计算机工程与应用,2021,57(5):36-46.
LUO H,PENG S,CHEN H K,et al.Review on latest research progress of challenging problems in object detection[J].Computer Engineering and Applications,2021,57(5):36-46.
- [3] 方路平,何杭江,周国民.目标检测算法研究综述[J].计算机工程与应用,2018,54(13):11-18.
FANG L P,HE H J,ZHOU G M.Research overview of object detection methods[J].Computer Engineering and Applications,2018,54(13):11-18.
- [4] 李章维,胡安顺,王晓飞.基于视觉的目标检测方法综述[J].计算机工程与应用,2020,56(8):1-9.
LI Z W,HU A S,WANG X F.Survey of vision based object detection methods[J].Computer Engineering and Applications,2020,56(8):1-9.
- [5] GIRSHICK R,DONAHUE J,DARRELL T,et al.Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28,2014:580-587.
- [6] GIRSHICK R.Fast R-CNN[C]//Proceedings of 2015 IEEE International Conference on Computer Vision (ICCV), December 7-13,2015:1440-1448.
- [7] REN S Q,HE K M,GIRSHICK R,et al.Faster R-CNN: Towards real-time object detection with region proposal networks[J].IEEE Transactions on Pattern Analysis and Machine Intelligence,2017,39(6):1137-1149.
- [8] HE K M,GKIOXARI G,DOLLAR P,et al.Mask R-CNN[J].IEEE Transactions on Pattern Analysis and Machine Intelligence,2020,42(2):386-397.
- [9] CAI Z W,VASCONCELOS N.Cascade R-CNN:Delving into high quality object detection[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 18-22,2018:6154-6162.
- [10] SUN P,ZHANG R F,JIANG Y,et al.Y Sparse R-CNN: End-to-end object detection with learnable proposals[J].arXiv:2011.12450,2020.
- [11] LIU W,ANGUELOV D,ERHAN D,et al.SSD: Single shot multibox detector[C]//Proceedings of European Conference on Computer Vision,2016:21-37.
- [12] REDMON J,DIVVALA S,GIRSHICK R,et al.You only look once: Unified, real-time object detection[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30,2016: 6517-6525.
- [13] REDMON J,FARHADI A.YOLO9000: Better, faster, stronger[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26,2017:6517-6525.
- [14] REDMON J,FARHADI A.YOLOv3: An incremental Improvement[J].arXiv:1804.02767,2018.
- [15] BOCHKOVSKIY A,WANG C Y,LIAO H Y M.YOLOv4: Optimal speed and accuracy of object detection[J].arXiv: 2004.10934,2020.
- [16] 黄凤荣,李杨,郭兰申,等.基于Faster R-CNN的零件表面缺陷检测算法[J].计算机辅助设计与图形学学报,2020,32(6):883-893.
HUANG F R,LI Y,GUO S L,et al.Method for detecting surface defects of engine parts based on faster R-CNN[J].Journal of Computer-Aided Design & Computer Graphics, 2020,32(6):883-893.
- [17] CHEN S H,KANG C H,PERNG C B.Detecting and measuring defects in wafer die using GAN and YOLOv3[J].Applied Sciences,2020,10:8725.
- [18] PEREZ H,TAH J H M,MOSAVI A.Deep learning for detecting building defects using convolutional neural networks[J].Sensors,2019,19(16):3556.
- [19] 金颖,王学影,段林茂.基于LabVIEW和Mask R-CNN的柱塞式制动主缸内槽表面缺陷检测[J].现代制造工程, 2020(5):125-132.
JIN Y,WANG X Y,DUAN L M.Surface defect detection of inner groove in plunger brake master cylinder based on LabVIEW and Mask R-CNN[J].Modern Manufacturing Engineering,2020(5):125-132.
- [20] TAN M,LE Q V.EfficientNet: Rethinking model scaling for convolutional neural networks[J].arXiv:1905.11946,2019.
- [21] HE K M,ZHANG X Y,REN S Q,et al.Spatial pyramid pooling in deep convolutional networks for visual recognition[J].IEEE Transactions on Pattern Analysis and Machine Intelligence,2015,37(9):1904-1916.
- [22] LIU S,QI L,QIN H F,et al.Path aggregation network for instance segmentation[J].arXiv:1803.01534,2018.
- [23] HE K M,ZHANG X Y,REN S Q,et al.Deep residual learning for image recognition[J].arXiv:1512.03385,2015.