计算机系统应用 ISSN 1003-3254, CODEN CSAOBN
Computer Systems & Applications,2023,32(2):339−346 [doi: 10.15888/j.cnki.csa.008881]

E-mail: csa@iscas.ac.cn
http://www.c-s-a.org.cn
Tel: +86-10-62661041

基于改进 YOLOv5s 的大尺寸导光板缺陷检测①

刘　霞, 张环宇

(东北石油大学 电气信息工程学院, 大庆 163318)
通信作者: 张环宇, E-mail: 932939991@qq.com

摘　要: 导光板 (LGP) 是液晶显示器 (LCD) 背光模组的主要部件. 导光板的缺陷将直接影响液晶显示器的显示效果. 针对导光板图像纹理背景复杂、低对比度、缺陷尺寸小等问题, 本文提出了一种用于大尺寸导光板缺陷检测的 AYOLOv5s 网络. 首先, 将导光板图像进行分图处理, 然后在主干部分和特征融合部分集成 Transformer 和注意力机制 coordinate attention, 并选择 Meta-ACON 激活函数. 最后, 基于自建数据集 LGPDD 进行了大量实验. 实验结果表明, LGP 缺陷检测算法的平均精度 (*mAP*) 可以达到 99.20%, 并且 FPS 可达 77, 可以实现在 12 s/pcs 内对尺寸为 17 英寸的导光板中的亮点、划伤、异物、磕碰伤、脏污等缺陷具有较好的实际检测效果.

关键词: 导光板缺陷检测; 分图处理; AYOLOv5s; 注意力机制; 深度学习

引用格式: 刘霞,张环宇.基于改进 YOLOv5s 的大尺寸导光板缺陷检测.计算机系统应用,2023,32(2):339–346. http://www.c-s-a.org.cn/1003-3254/8881.html

Defect Detection of Large-size Light Guide Plate Based on Improved YOLOv5s

LIU Xia, ZHANG Huan-Yu

(School of Electrical Information Engineering, Northeast Petroleum University, Daqing 163318, China)

Abstract: A light guide plate (LGP) is the main component of the backlight module of a liquid crystal display (LCD), whose defects can directly affect the display effect of LCD. To address the problems of complex texture background, low contrast, and small defect size of LGP images, this study proposes an AYOLOv5s network for defect detection of large-size LGP images. First, the LGP image is divided into different images. Then, Transformer and the attention mechanism coordinate attention are integrated in the main part and feature fusion part, and the Meta-ACON activation function is selected. Finally, massive experiments are carried out on the basis of the self-built data set LGPDD. The experimental results indicate that the defect detection algorithm for LGP enjoys the mean average accuracy (*mAP*) of up to 99.20% and FPS of 77, which can realize good effects in the practical detection of bright spots, scratches, foreign bodies, bumps, dirt, and other defects in the 17-inch LGP in 12 s/pcs.

Key words: defect detection of light guide plate; figure processing; AYOLOv5s; attention mechanism; deep learning

1　前言

随着工业现代化的飞速发展, 国内的导光板厂商开始陆续尝试自动光学检测方法 (AOI) 进行导光板质检. AOI 是以机器视觉技术为基础的缺陷检测技术, AOI 是解决目前传统检测效率低、工人成本高问题的最有效方法. 广泛应用于 PCB 等相关行业的检测中. 稳定的 AOI 检测是目前导光板检测的发展方向.

2019 年 Ming 等人提出了一种基于机器视觉的导光板 (LPG) 缺陷检测新方法. 在图像分割和增强之后, 对所提出的具有动态权重的组合分类器 (CCDW) 进行训练[1]. 导光板图像有背景复杂、对比度低、亮度不均、各种缺陷特征差异大等特点. 基于机器视觉和机

① 收稿时间: 2022-04-27; 修改时间: 2022-06-01; 采用时间: 2022-06-10; csa 在线出版时间: 2022-11-14
CNKI 网络首发时间: 2022-11-15

器学习的传统检测方法需要提取手动特征, 鲁棒性差. 难以满足导光板在线缺陷检测的抗干扰性、检测精度、实时性等方面的需求.

近年来, 许多深度学习缺陷检测方法被广泛应用于 LCD 等工业领域. 并且在导光板缺陷检测领域也开始了一些初步的尝试. 2021 年 Li 等人提出了一种用于手机导光板缺陷检测的端到端多任务学习网络架构. 同过 U 型编码结构获取多尺度特征, 并采用特征融合与多尺度特征交互[2]. Hong 等人提出了一个密集双线性卷积神经网络 (BCNN), 一个端到端缺陷检测网络[3]. 何炎森针对车载导航导光板缺陷, 结合轻量化与级联网络提出了一种缺陷快速检测方法[4].

导光板的缺陷检测要求相对高, 大部分缺陷, 如亮点、暗点、线划痕和压痕小于 0.1 mm. 检测的精度需要大于 99%, 每个导光板需要在 12 s 内检测到. 一般来说, 基于深度学习分类网络的缺陷检测只能获得粗略的位置. 相对而言, 目标检测网络可以获取目标的准确位置和类别信息. 目标检测模型可以分为两大类: 两阶段和一阶段网络. 两阶段网络需要在目标检测之前生成可能包含缺陷的候选区域, 主要是包括 R-CNN[5]、Faster R-CNN[6] 等网络. 一阶段网络通过从网络中提取的特征直接进行缺陷检测和识别, 主要包括 YOLO[7,8] 系列和 SSD[9] 等网络. 检测速度更快, 目前在导光板缺陷检测问题上已经取得了很好的效果, 2022 年 Yao 等人提出一种用于 LGP 缺陷检测的 AYOLOv3-Tiny 网络[10]. 对传统的 YOLOv3-Tiny 主干部分的卷积操作进行改进, 满足了工业检测需求, 验证了 YOLO 模型在导光板缺陷检测任务上的优秀性能.

YOLOv5 融合了 YOLO 系列版本的优点, 在检测速度和精度上都更优秀. 但是为了实现本文大尺寸导光板图像背景下对小目标缺陷的检测, 还需要对 YOLOv5 的网络结构进行相应的调整和改进. 因此本文结合导光板的缺陷特征和检测要求, 提出了一种基于改进 YOLOv5s 的大尺寸导光板缺陷检测算法.

2 改进的 YOLOv5s 算法

YOLOv5 根据不同 CSP 结构的深度和宽度主要将模型分为 4 种: YOLOv5s、YOLOv5l、YOLOv5m、YOLOv5x. 模型依次增大, 推理时间依次延长, 考虑到算法的实时性, 本文选择 YOLOv5s 为基础架构进行改进. 图 1 为改进后的 YOLOv5s 网络结构图, 由 Input、Backbone、Neck、Head 组成, 本文将 Backbone 部分倒数第 2 层的 C3 模块替换为 C3TR 模块并将 coordinate attention 注意力机制[11] 集成到 Neck 部分, 在激活函数方面选择了 Meta-ACON[12], 在损失函数方面选择*CIoU* 作为目标框回归的损失函数. 使得网络训练速度的速度提高的同时, 特征提取能力也得到了进一步的增强, 从而提高了对细微缺陷的检测能力. 下面从输入端 (Input)、主干特征提取网络 (Backbone)、特征融合 (Neck)、激活函数、损失函数 (*Loss*) 这 5 个方面进行详细介绍.

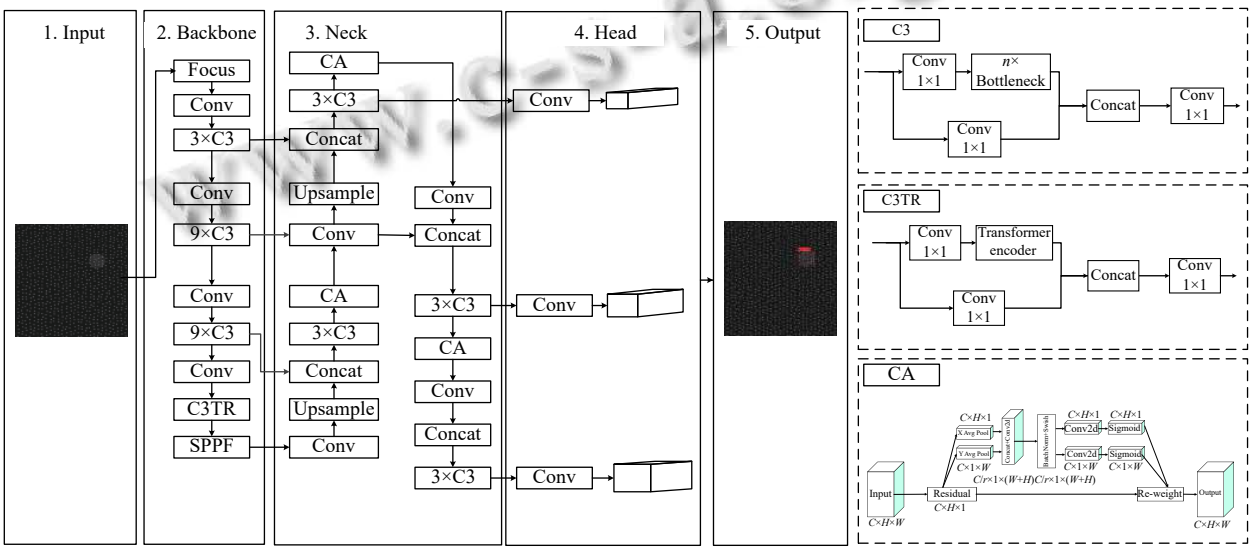


图 1 AYOLOv5s 网络结构

2.1 分图输入

本文待检导光板的尺寸为 12–17 英寸. 导光板的缺陷检测要求相对较高, 且导光板尺寸太大. 因此, 本文采用 16k 线阵相机获取高质量的图像. 图 2 为采集后的导光板图像, 为了适应不同尺寸导光板, 采集后的导光板图像分辨率统一为 25000×16384≈4.1 亿个像素, 而小的缺陷仅为 7 个像素左右, 缺陷尺寸和图像背景差距悬殊, 无法直接进行检测, 因此考虑到算法的可实现性, 需要采用逐行扫描的方式进行检测. 本文首先将获得的完整导光板图像进行分图, 裁剪成数张 640×640 的小图后再通过算法进行滑窗检测. 从图 2 采集的导光板图像可以看出, 我们只需裁剪真正的导光板区域即可, 即 18064×13548 个像素的感兴趣区域. 值得注意的是由于导光板进料时非常稳定, 所以提取感兴趣区域时不需复杂操作, 只需根据相应尺寸提前设置好区域坐标即可. 为了消除边界缺陷的漏检, 设置的大小为 10 个像素的重叠区域 overlap, 那么步长就为 630 个像素点. 从左上角开始切图, 切出来图像的左上角记为 x, y, 那么可以容易想到 y 依次为: 0, 630, 1 260, ⋯, 17 010 但接下来却并非是 17 640, 因为 17640+640>18064, 所以这里要对切图的 overlap 做一个调整, 最后一步的 y=18064–640, 基于此就可完成将所有感兴趣区域切成小图, 将裁剪后的小图通过 Mosaic 数据增强后输入主干特征提取网络.

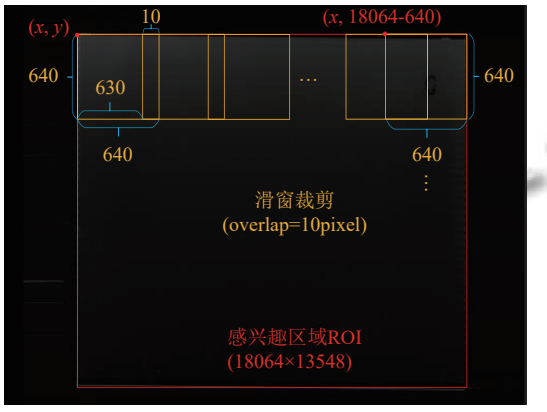


图 2 采集的导光板图像

2.2 基于 Transformer 的主干特征提取网络

YOLOv5 的主干部分采用了 Focus 下采样、改进 CSP 结构、SPPF 池化金字塔结构提取图片的特征信息. YOLOv5 所使用的主干特征提取网络为 CSPDarknet, 都由残差卷积构成, 残差网络能够通过增加相当的深度来提高准确率. 其内部的残差块使用了跳跃连接, 缓解了在深度神经网络中增加深度带来的梯度消失问题. 图 3 为 C3 模块结构图, YOLOv5 使用的 C3 模块结构是将原来的残差块的堆叠进行了一个拆分, 如图 3 所示由两个分支组成, 一个分支使用了多个 Bottleneck 堆叠和 1 个标准卷积层, 另一支仅经过一个基本卷积模块, 最后将两支进行 Concat 操作.

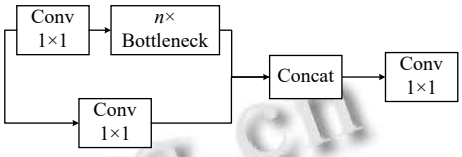


图 3 C3 模块结构

Zhu 等人[13] 在 TPH-YOLOv5 中用 Transformer 编码块替代了 YOLOv5 中的一些卷积块和 CSP bottlenck blocks, 其结构如下图 4 所示. Transformer 中的基本成分, 包括多头自注意力机制 (multi-head self-attention, MSA)、多层感知机 (multi-layer perceptron, MLP) 和层标准化 (layer normalization, LN) 与 CSPDarknet53 中的 bottleneck blocks 相比, Transformer 可以捕获全局信息和丰富的上下文信息. TPH-YOLOv5 通过这种集成 Transformer 的结构使 YOLOv5 在小目标检测的任务上取得了更好的性能, 但是提升模型精度的同时也不可避免的增加了图像的推理时间.

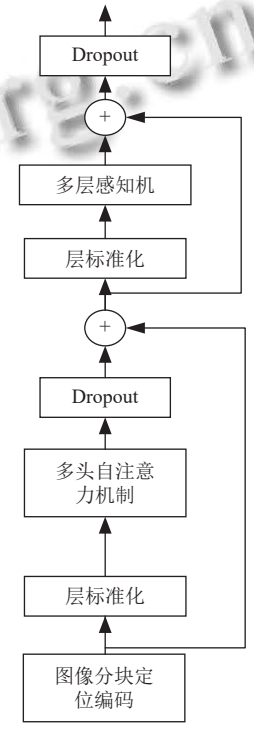


图 4 Transformer 编码器结构

对于小目标来说, 一般在卷积的深层才能体现特征, 所以本文的思路是将 Transformer 加到 YOLOv5s 的主干网络的最后一个 C3 模块中, 这样就可以利用 Transformer 的特性及将注意力的方式更好捕获对小目标有利的信息上, 从而更好地方便后期特征提取和检测. 使用 Transformer 编码器结构替换原 YOLOv5 中 C3 模块中的 Bottleneck , 组成新的 C3TR 模块, C3TR 的结构如图 5 所示.

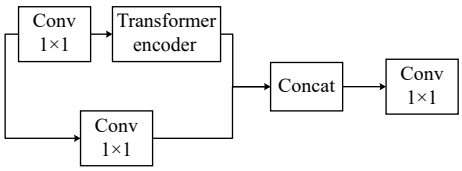


图 5 C3TR 模块结构

2.3 基于 coordinate attention 的特征融合网络

YOLOv5 的特征融合主要采用 FPN+PAN 的特征金字塔结构, 实现了不同尺寸目标特征信息的传递, 解决了多尺度问题.

本文在使用卷积神经网络去处理导光板图像的时候, 更期望卷积神经网络去注意存在缺陷的地方, 而不是关注整个导光板图像背景, 这种期望无法通过手动调节来实现, 因此, 该思考的就是如何让卷积神经网络去自适应的注意导光板图像中的缺陷位置. 注意力机制就是实现思考的一种方式. Coordinate attention (CA) 是 Hou 等人[1] 提出的新的注意力机制. 图 6 为 CA 注意力模块结构图, 从图中可以看出 CA 分别在垂直和水平两个方向通过平均池化的方式得到两个一维向量, 在空间维度上 Concat 连接和 1×1 卷积来压缩通道, 然后通过 BN 和 Non-linear 来编码水平和垂直两个方向的空间信息, 接着进行分裂, 再各自通过 1×1 卷积得到和输入特征图一样的通道数, 最后归一化加权.

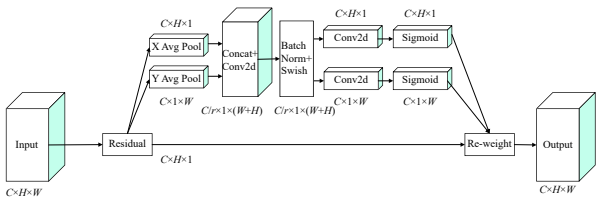


图 6 CA 注意力模块结构

注意力机制是个即插即用的模块, 可以集成到任何 CNN 架构中, 并且不会对网络增加很大负担. 本文中在特征融合网络中插入 CA 注意力模块, 使用 CA 可以更加注意线缺陷的特征信息, 进而提高线缺陷的准确率.

2.4 激活函数选择

Ma 等人[12] 提出了一个简单、有效、通用的激活函数, 称之为 ACON, 它可以学习并决定是否要激活神经元, 然后又进一步提出了 Meta-ACON, 它明确地学习优化非线性 (激活) 和线性 (非激活) 之间的参数切换. 作者提出一种新颖的 Swish 函数解释: Swish 函数是 ReLU 函数的平滑近似 (smooth maximum), 并基于这个发现, 进一步分析 ReLU 的一般形式 Maxout 系列激活函数, 利用 smooth maximum 将 Maxout 系列扩展得到简单且有效的 ACON 系列激活函数: ACON-A、ACON-B、ACON-C. 表 1 为 Maxout 和 ACON 系列激活函数的总结. 把$\eta_a(x)$, $\eta_b(x)$代入不同的值, 就可以得到表 1 的不同形式.

表 1 Maxout 和 ACON 系列激活函数

输入1	输入2	Maxout系列	ACON系列
$\eta_a(x)$	$\eta_b(x)$	$\max(\eta_a(x),\eta_b(x))$	$(\eta_a(x)-\eta_b(x))\cdot\sigma(\beta(\eta_a(x)-\eta_b(x)))+\eta_b(x)$
x	0	$\max(x,0)$: ReLU	ACON-A(Swish) : $x\cdot\sigma(\beta x)$
x	px	$\max(x,px)$: PReLU	ACON-B: $(1-p)x\cdot\sigma(\beta(1-p)x)+px$
p_1x	p_1x	$\max(p_1x,p_2x)$	ACON-C: $(p_1-p_2)x\cdot\sigma(\beta(p_1-p_2)x)+p_2x$

ACON 激活函数通过 β 来控制是否激活神经元 (β 为 0 则不激活). Meta-ACON 设计了自适应函数来计算平滑因子. 提出一个 $G(x)$ 模块由输入特征图 $x(C\times H\times W)$ 来动态的学习 β, 以达到自适应控制函数线性/非线性能力, 这种定制的激活行为有助于提高泛化和传递性能. 自适应函数包含了 layer-wise, channel-wise, pixel-wise 这 3 种空间, 分别对应的是层, 通道, 像素. Meta-ACON 选择了通道空间, 首先分别对 H, W 维度求均值, 然后通过两个卷积层, 使得每一个通道所有像素共享一个权重, 公式如下:

$$\beta_c=\sigma W_1W_2\sum_{h=1}^{H}\sum_{w=1}^{W}x_{c,h,w} \tag{1}$$

其中, 每个通道共享参数 β, 得到的 β 大小为 $C\times1\times1$, 其中 W 是 1×1 卷积的参数, 有 $W1(C, C/r)$, $W2(C, C/r)$, r 为缩放参数, 设置为 16. x 首先在 H 和 W 维度上求均值, 然后经过两层 1×1 卷积, 最后由 Sigmoid 激活函数得到一个 (0, 1) 的值, 用于控制是否激活.

ACON 可以很自然地应用于对象检测和语义分

割, 这表明 ACON 在各种任务中是一种有效的替代方案. 因此本文将 Meta-ACON 激活函数应用到改进后的缺陷检测网络模型之中.

2.5 AYOLOv5s 损失函数

*Loss*实际上是网络的预测结果和网络的真实结果的对比, 是衡量模型性能的重要指标. YOLOv5 模型损失函数包括定位损失函数L_{box}、置信度损失函数L_{conf}、分类损失函数L_{cls}, 因此损失函数可表示为:

$$Loss = \lambda_{box}L_{box} + \lambda_{conf}L_{conf} + \lambda_{cls}L_{cls} \tag{2}$$

其中, λ_{box}、λ_{conf}、λ_{cls}分别是定位损失权重, 置信度损失权重和分类损失权重.

本文选择*CIoU_Loss*[14] 作为 BoundingBox 回归的损失函数, 其计算式为:

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b,b^{gt})}{c^2} + \alpha\upsilon \tag{3}$$

$$\alpha = \frac{\upsilon}{(1-IoU)+\upsilon} \tag{4}$$

$$\upsilon = \frac{4}{\pi^2}\left(\arctan\frac{\omega^{gt}}{h^{gt}} - \arctan\frac{\omega}{h}\right)^2 \tag{5}$$

其中, α是一个平衡参数, 不参与梯度计算; υ是用来衡量长宽比一致性的参数. *CIoU_Loss*综合考虑了真实框与预测框之间的重叠率、中心点距离、长宽比, 使得网络有更快更好的收敛效果.

$$L_{conf} = \lambda_{obj}\sum_{i=0}^{s^2}\sum_{j=0}^{B} I_{ij}^{obj}(C_i - \hat{C}_i)^2 + \lambda_{noobj}\sum_{i=0}^{s^2}\sum_{j=0}^{B} I_{ij}^{noobj}(C_i - \hat{C}_i)^2 \tag{6}$$

$$L_{cls} = \sum_{i=0}^{s^2}\sum_{j=0}^{B} I_{ij}^{obj}\sum_{c\in classes}[\hat{p}_i(c)\log(P_i(c)) + (1-\hat{p}_i(c))\log(1-p_i(c))] \tag{7}$$

其中, C_i表示预测框的置信度得分, $\hat{C}_i$表示真实框的置信度得分; I_{ij}^{obj}和I_{ij}^{noobj}分别代表第i个单元的j个锚点; c是当前检测到的目标类别, *classes*是所有目标的类别; $\hat{p}_i(c)$和$P_i(c)$分别是预测类别和真实类别.

3 实验结果与分析

3.1 导光板缺陷数据集

本文主要针对导光板的亮点、压伤、划伤、脏污、异物、黑点、曲面斑纹等缺陷进行研究. 由于这些导光板的分类主观性较强, 实际成像上, 缺陷之间没有较明显的界限, 因此本文根据生产厂家的技术要求、一线工人的工作经验和导光板的成像特点, 将这些缺陷分为点缺陷和线缺陷两类, 图 7 为导光板缺陷分类情况.

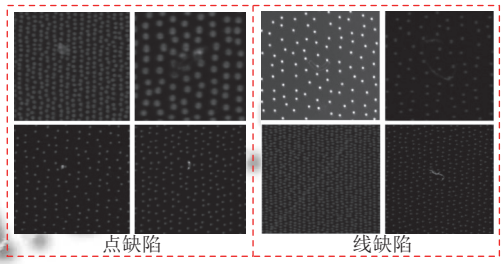


图 7 导光板缺陷分类

导光板缺陷数据集 (LGPDD) 来自于真实的工业领域. 包含 12–17 英寸内的共 500 张导光板, 经过分图预处理后, 获得 640×640 的小图后制作数据集, 本文通过数据扩充来平衡两种缺陷的数量. 数据增强包括以 50% 的概率旋转, 120%–150% 的亮度增强. 扩展后数据集具体参数如表 2 所示.

表 2 LGPDD 数据集参数

缺陷类型	标签名称	数量
点缺陷	dot	2362
线缺陷	line	2140

3.2 实验环境与模型训练

本文实验的具体配置为: Windows 10 操作系统、单 RTX2060 显卡、Python 3.7 虚拟环境、PyTorch 1.7 深度学习开源框架, CUDA11 进行加速. 本文在自建数据集上将改进后的 YOLOv5s 模型与原始 YOLOv5s 进行对比实验, 代码参考 YOLOv5 官方代码 6.1 版本. 迭代次数为 300 epoch, 优化器选择 SGD, 初始学习率为 0.01, 动量参数为 0.937 并采用 epoch 为 3、动量参数为 0.8 的 warmup 方法预学习, 在 warmup 阶段, 采用一维线性插值更新学习率, 预学习结束后采用余弦退火对学习率进行更新. 训练过程中每完成一次迭代就会进入一次验证阶段. 每一次迭代后自动更新权重文件.

3.3 评价指标与实验结果分析

本文评估指标采用平均精度均值 (*mAP*) 和每秒检测图片的帧数 (FPS) 这两种在目标检测算法中较为常见的评价指标来评估本文算法的性能. 平均精度与精

确率 (*precision*, *P*) 和召回率 (*recall*, *R*) 有关, 精确率是指预测数据集中预测正确的正样本个数除以被模型预测为正样本的个数; 召回率是指预测数据集中预测正确的正样本个数除以实际为正样本的个数. *mAP* 计算中的 *AP* 即 *P*-*R* 曲线下方面积, 具体计算基于式 (8)–式 (11):

$$mAP=\frac{\sum_{i=1}^{N}AP_i}{N} \tag{8}$$

$$AP=\int_0^1 P(R)dR \tag{9}$$

$$P=\frac{TP}{TP+FP} \tag{10}$$

$$R=\frac{TP}{TP+FN} \tag{11}$$

其中, *AP* 值是指 *P*-*R* 曲线面积; *mAP* 的值是通过所有类别的 *AP* 求均值得到; *N* 表示检测的类别总数, 本实验中 *N*=2, *mAP* 的值越大, 表示算法检测效果越好, 识别精度越高; *TP*、*FP* 和 *FN* 分别表示正确检测框、误检框和漏检框的数量.

为了评估 AYOLOv5s 网络的准确性和实时性本文在自建数据集 LGPDD 上进行了大量的实验, 图 8 为 YOLOv5s 和 AYOLOv5s 在训练过程中 *box_loss*、*cls_loss*、*mAP*@0.5、*mAP*@0.5:0.95 的曲线图. *mAP*@0.5 代表在 *IoU* 阈值为 0.5 时的平均 *AP*, *mAP*@0.5:0.95 代表在 *IoU* 阈值为从 0.5 到 0.95, 步长为 0.05 时各个 *mAP* 的平均值. *mAP*@0.5 主要用于体现模型的识别能力, *mAP*@0.5:0.95 由于要求的 *IoU* 阈值更高, 主要用于体现定位效果以及边界回归能力, 通过观察训练过程中不同模型的曲线可以看出本文改进的 YOLOv5s 网络收敛效果优于原版本 YOLOv5s 模型, 说明 AYOLOv5s 面对导光板缺陷这种小目标的特征具有更好的学习能力.

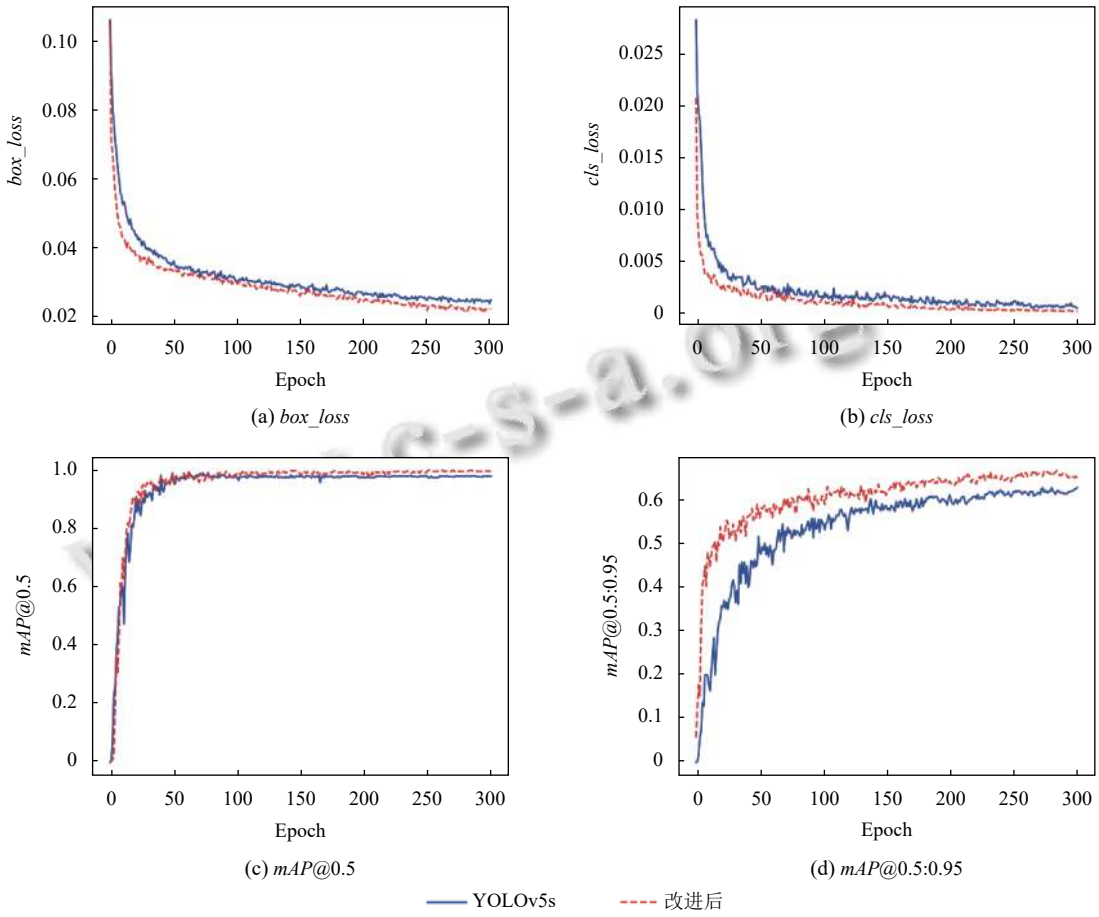
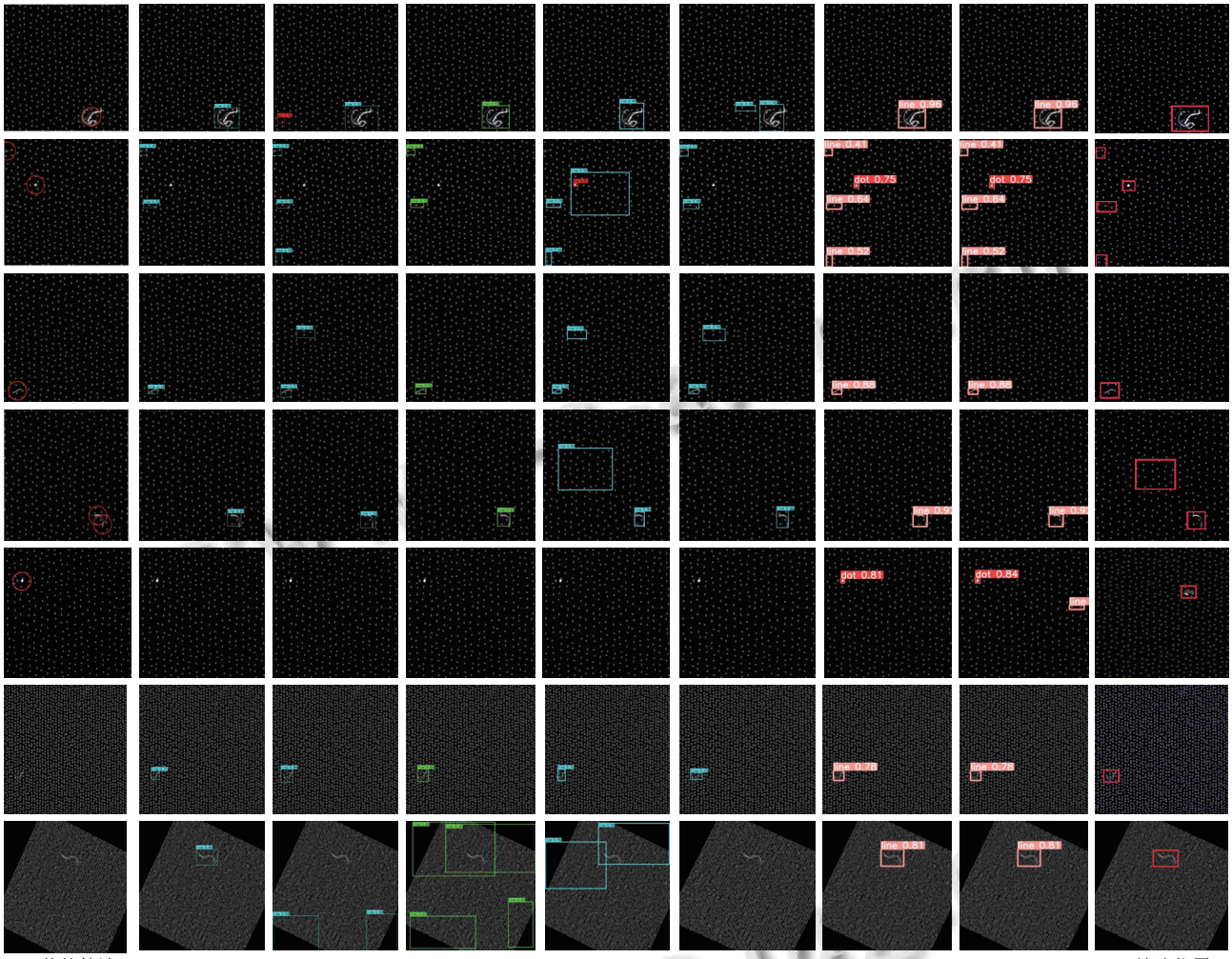


图 8 训练过程中 *box_loss*、*cls_loss*、*mAP*@0.5、*mAP*@0.5:0.95 曲线

此外, 通过 *mAP* 和 FPS, 将 AYOLOv5s 网络与 SSD、YOLOv3、EfficientDet[15]、Faster R-CNN 等目标检测网络进行了比较. 图 9 为各模型在 LGPDD 上部分测试结果, 表 3 为实验结果汇总.



(a) 传统算法 (b) EfficientDet (c) Faster R-CNN (d) SSD (e) YOLOv3 (f) YOLOv4 (g) YOLOv5s (h) AYOLOv5s (i) 缺陷位置

图 9 各类算法结果对比

表 3 实验结果汇总

网络模型	*mAP* (%)	FPS
SSD	96.97	23.1
YOLOv3	89.19	62
YOLOv4	92.54	43
YOLOv5s	97.5	102
EffecientDet	81.55	22
Faster R-CNN	81.65	17.5
AYOLOv5s	99.2	77

从表 3 可以看出, 随着 YOLO 系列网络的发展, YOLOv5s 相较于其他经典缺陷检测算法有着更好的性能. 但是, YOLOv5s 的检测精度并不满足导光板生产的要求, AYOLOv5s 在点和线缺陷的准确性上有显著提高, 但 FPS 略有下降. 与 YOLOv5s 相比, AYOLOv5s 对缺陷的平均检测精度提高了 1.7%, 但 FPS 降低了 24.5%. 由于网络的复杂性较高, 体积增加导致速度降低. 虽然 AYOLOv5s 的检测速度有损失, 但仍然可以满足基于导光板生产过程的实时检测要求, AYOLOv5s 可以实现在 12 s/pcs 内对尺寸为 17 英寸的导光板中的亮点、划伤、异物、磕碰伤、脏污等缺陷具有较好的实际检测效果.

4 总结

本文根据导光板的光学特性、缺陷形成原理和成像特性, 提出了一种用于大尺寸导光板缺陷检测的

AYOLOv5s 网络. 分别从输入端、主干特征提取网络、特征融合网络、激活函数、损失函数 5 个方面进行改进, 有效地提高了 YOLOv5s 网络模型对图像中细微缺陷的检测精度, 改进后的算法检测速率相较于原始 YOLOv5s 算法有所降低, 但仍能满足实时性要求, 可以满足导光板缺陷检测高速和高精度的实际生产需求.

参考文献

1 Ming WY, Shen F, Zhang HM, *et al*. Defect detection of LGP based on combined classifier with dynamic weights. Measurement, 2019, 143: 211–225. [doi: 10.1016/j.measurement.2019.04.087]

2 Li Y, Li JF. An end-to-end defect detection method for mobile phone light guide plate via multitask learning. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1–13.

3 Hong LB, Wu XL, Zhou DB, *et al*. Effective defect detection method based on bilinear texture features for LGPs. IEEE Access, 2021, 9: 147958–147966. [doi: 10.1109/ACCESS.2021.3111410]

4 何炎森. 基于机器学习的车载导航导光板质量视觉检测系统研究 [硕士学位论文]. 杭州: 浙江理工大学, 2021.

5 Girshick R, Donahue J, Darrell T, *et al*. Rich feature hierarchies for accurate object detection and semantic segmentation. Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014. 580–587.

6 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: MIT Press, 2015. 91–99.

7 Redmon J, Farhadi A. YOLOv3: An incremental improvement. arXiv:1804.02767, 2018.

8 Bochkovskiy A, Wang CY, Liao HYM. YOLOv4: Optimal speed and accuracy of object detection. arXiv:2004.10934v1, 2020.

9 Liu W, Anguelov D, Erhan D, *et al*. SSD: Single shot multibox detector. Proceedings of the 14th European Conference on Computer Vision. Cham: Springer, 2016. 21–37.

10 Yao JH, Li JF. AYOLOv3-Tiny: An improved convolutional neural network architecture for real-time defect detection of PAD light guide plates. Computers in Industry, 2022, 136: 103588. [doi: 10.1016/j.compind.2021.103588]

11 Hou QB, Zhou DQ, Feng JS. Coordinate attention for efficient mobile network design. Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021. 13713–13722.

12 Ma NN, Zhang XY, Liu M, *et al*. Activate or not: Learning customized activation. Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021. 8032–8042.

13 Zhu XK, Lyu SC, Wang X, *et al*. TPH-YOLOv5: Improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios. Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision. Montreal: IEEE, 2021. 2778–2788.

14 Zheng ZH, Wang P, Ren DW, *et al*. Enhancing geometric factors in model learning and inference for object detection and instance segmentation. IEEE Transactions on Cybernetics, 2021, 52(8): 8574–8586.

15 Tan MX, Pang RM, Le QV. EfficientDet: Scalable and efficient object detection. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020. 10778–10787.

(校对责编: 孙君艳)