

# 基于 PU-Faster R-CNN 的手机屏幕 缺陷检测算法研究

李伟朝<sup>1</sup>, 陈志豪<sup>1</sup>, 张 颢<sup>2</sup>, 查云威<sup>1</sup>

(1. 广东工业大学 自动化学院广东省物联网信息技术重点实验室, 广州 510006;

2. 华南理工大学 电力学院, 广州 510630)

**摘要:** 手机屏幕缺陷检测是手机生产的重要环节, 实现准确而高效的屏幕缺陷检测对于提高手机工业产能具有重要意义; 在实际生产过程中, 手机屏幕图像缺陷特征隐晦、缺陷尺寸差异大等问题, 加大了手机屏幕缺陷检测的难度; 为解决上述问题, 提出了一种基于 Preprocessing operations are combined with U-Net-Faster R-CNN (PU-Faster R-CNN) 的手机屏幕缺陷检测模型; 针对手机屏幕图像的特征信息隐晦的问题, 提出多层特征增强模块, 有效的对目标缺陷特征信息进行增强; 构建多尺度特征提取网络, 有效提取多尺度的缺陷特征信息; 为了生成拟合性更好的 Anchor box, 提出了自适应区域建议网络, 通过自迭代聚类算法生成尺寸更准确的 Anchor box 模板; 实验结果表明, 基于 PU-Faster R-CNN 的手机屏幕检测框架在手机屏幕数据集上优于主流的手机屏幕缺陷检测框架。

**关键词:** 手机屏幕; 缺陷检测; PU-Faster R-CNN; 多层特征增强模块; 自适应区域建议网络

## PU-Faster R-CNN-Based Defect Detection Model for Mobile Phone Screen

LI Weichao<sup>1</sup>, CHEN Zhihao<sup>1</sup>, ZHANG Xie<sup>2</sup>, ZHA Yunwei<sup>1</sup>

(1. School of Automation, Guangdong Key Laboratory of IoT Information Technology, Guangdong

University of Technology, Guangzhou 510006, China;

2. School of Electric Power Engineering, South China University of Technology, Guangzhou 510630, China)

**Abstract:** Mobile phone screen defect detection is an important part of mobile phone production. A accurate and efficient defect detection is of great significance for improving the productivity of mobile phone industry. In the actual production process, there are the problems of no obvious features in the screen defect image and great difference in the defect size, which increases the difficulty of mobile phone screen defect detection. A mobile phone screen defect detection model based on PU-Faster R-CNN was proposed to solve above problems. For the problem of obscure feature information of mobile phone screen images, a multi-layer feature enhancement module was proposed to effectively enhance the target defect feature information. A multi-scale feature extraction network was constructed to effectively extract multi-scale defect feature information. In order to generate Anchor boxes with better fitting performance, an adaptive region proposal network was proposed to generate Anchor box templates with more accurate size by the self-iterative clustering algorithm. The experimental results showed that the framework was superior to the mainstream mobile phone screen defect detection framework in mobile phone screen datasets.

**Keywords:** mobile phone screen; defect detection; PU-Faster R-CNN; multi-layer feature enhancement module; adaptive region proposal network

## 0 引言

在手机屏幕的制作工艺过程中, 由于环境和技术等因素影响, 一些产品难免出现不可避免的缺陷。如果让存在缺陷的手机屏幕流入市场, 将会严重影响使用者的体验, 甚至危害生命。因此手机屏幕的缺陷检测是手机生产的关键环节。传统的人工检测方法, 存在着筛检成本高、耗时长、准确率低等一系列问题。

随着机器学习的发展, 其应用场景涉及多个领域。基于机器学习的缺陷检测方法也不断被提出。Huang 等人<sup>[1]</sup>提出了一种检测手机屏幕表面缺陷的框架。该框架包含连续图像切割规范化、原始图像规范化、图像块分割、特征提取和机器学习等方法。Jian 等人<sup>[2]</sup>提出了一种改进的手机屏幕玻璃缺陷识别和分割检测算法。该算法基于轮廓对齐方法生成模板图像。根据对准结果, 采用减法和投影相结合的方法识别图像中的缺陷。针对具有模糊灰边界的噪声

收稿日期: 2023-01-17; 修回日期: 2023-03-01。

基金项目: 科技创新 2030-“新一代人工智能”国家级重大项目(2020AAA0108300)。

作者简介: 李伟朝(1995-), 男, 广东肇庆人, 硕士研究生, 主要从事深度学习、目标检测、图像处理方向的研究。

通讯作者: 张 颢(1975-), 女, 四川达州人, 博士, 副教授, 主要从事深度学习, 智能化系统方向的研究。

引用格式: 李伟朝, 陈志豪, 张 颢, 等. 基于 PU-Faster R-CNN 的手机屏幕缺陷检测算法研究[J]. 计算机测量与控制, 2023, 31(7): 99-

图像中的缺陷,提出了一种改进的模糊 c-均值聚类算法。尽管基于机器学习的特征提取部分能通过专家设定有效的特征,但与深度学习相比,所提取的特征信息丰富度欠缺。此外,基于机器学习的缺陷检测方法需要大量的先验知识,并且缺乏鲁棒性。

为了解决基于机器学习的检测方法的局限性,不少研究人员提出优秀的基于深度学习的检测框架。Guo 等人<sup>[3]</sup>提出了一种高效的缺陷检测网络(EDD-Net)。该模型的骨干网络是高效网络。该模型提出了一种基于手机表面缺陷的特征金字塔模块 GCSA-BiFPN。此外,该模型利用一个盒/类预测网络进行有效的缺陷检测。文献<sup>[4]</sup>提出了一种基于 Faster R-CNN 小样本学习的手机屏幕表面缺陷检测模型框架。设计了一个深度卷积生成对抗网络(DCGAN),用于自动提取和融合缺陷特征,以增强和生成缺陷样本。Wang 等人<sup>[5]</sup>提出了一种用于手机屏幕缺陷检测的孪生网络。该网络结合对比度损失和交叉熵损失来提高模型的识别能力。Ren 等人<sup>[6]</sup>提出了一种基于“分类网络+带注意力机制的 U-Net”的小目标分割和非明显缺陷检测方法。该方法提出了一种改进的解决方案,在分类网络中添加了一个分割网络,并在经典的 U-Net 中添加了注意力机制,提高了小目标分割和无意义缺陷检测的性能。Zhu 等人<sup>[7]</sup>提出了分层多频注意力网络(HMFCA-Net)。值得注意的是,其提出了一种使用多频率信息和局部跨通道交互的注意力机制来表示加权缺陷特征。另外介绍了一种基于变形卷积的 ResNeSt 网络,旨在处理各种缺陷形状。以上基于深度学习的缺陷检测框架能达到令人满意的效果,但是对于手机屏幕缺陷检测识别准确率与实际应用的要求还存在一定的差距,仍然存在缺陷目标检测精度不高的问题,特别是对尺寸小的缺陷目标检测准确率偏低。

### 1 手机屏幕特点分析及缺陷检测问题难点

原始手机屏幕检测样品为透明玻璃面板,需放在指定工控机上,通过光源照射,在背景板用摄像头收集对应的手机屏幕原始图像到指定服务器。手机屏幕图像为灰度图像,背景为黑色像素,白色像素根据不同特征为不同的缺陷。如图 1 所示,分别是缺陷为气泡、划痕、锡灰和无缺陷的手机屏幕图像。方框仅为缺陷所在位置方便读者阅,其非预测结果。

手机屏幕缺陷检测是一个自动化的目标检测过程,是一个集手机屏幕缺陷特征提取、缺陷分类和缺陷定位的多任务过程。目的是准确找到手机屏幕上的缺陷,确定其类别和位置。具体而言,检测算法要实现 3 个方面的功能:1)判断输入的手机屏幕图像有无缺陷;2)对有缺陷的手机屏幕图像中的缺陷进行定位;3)对定位的手机屏幕缺陷进行分类。

通过对手机屏幕图像检查过程进行技术分析,手机屏幕缺陷检测主要存在以下难点:1)缺陷特征提取困难。如图 1 (e) (f) 所示手机屏幕图像缺陷特征极其隐晦,缺陷大多数为几十像素甚至十几像素,且缺陷分布不规律。该

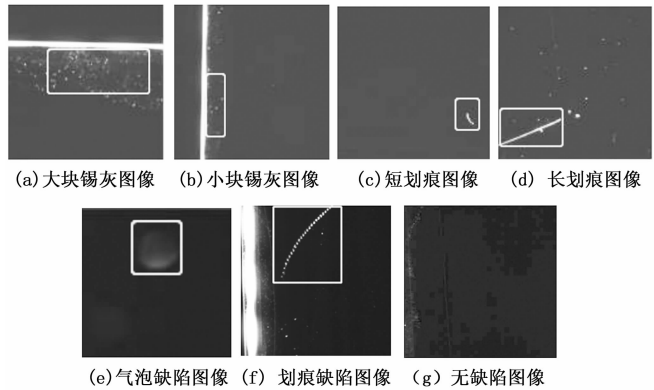


图 1 手机屏幕图像

特性造成缺陷信息复杂,提取特征极度困难。2)如图 1 (a) (b) 大块锡灰和小块锡灰图像和 (c) (d) 短划痕和长划痕图像所示缺陷尺寸差异大。同种缺陷形状各不相同,尺寸跨度较大,影响了候选框的边框回归效果。3)缺陷特征相似。手机屏幕中一些缺陷特征极其相似,肉眼难以分辨,如划痕纹理与锡灰相似度高。

针对手机屏幕缺陷数据集带来的检测难点,使得基于机器学习和深度学习的检测框架难以达到令人满意的检测效果。因此,本文在 Faster R-CNN 目标检测模型,提出 PU-Faster R-CNN 检测框架,针对手机屏幕缺陷图像的检测难点,实现手机屏幕缺陷的高效检测。

### 2 PU-Faster R-CNN 算法模型

本文在 Faster R-CNN 目标检测的基础上加以改进,提出 PU-Faster R-CNN。PU-Faster R-CNN 框架结构如图 2 所示。该模型在原 Faster R-CNN 模型中提出多层特征增强模块,特征提取网络层采用多尺度特征提取网络,提出了自适应区域建议网络作为区域建议网络。

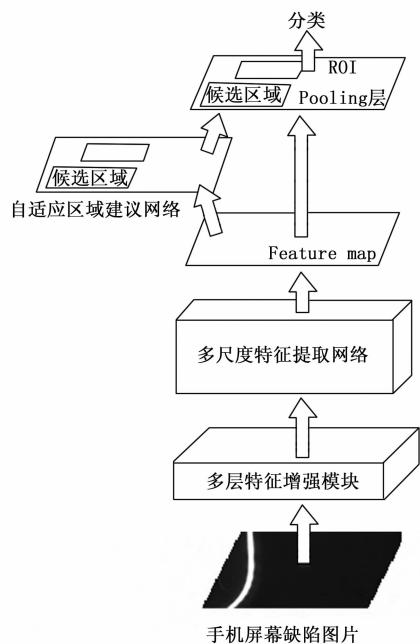


图 2 PU-Faster R-CNN 框架图

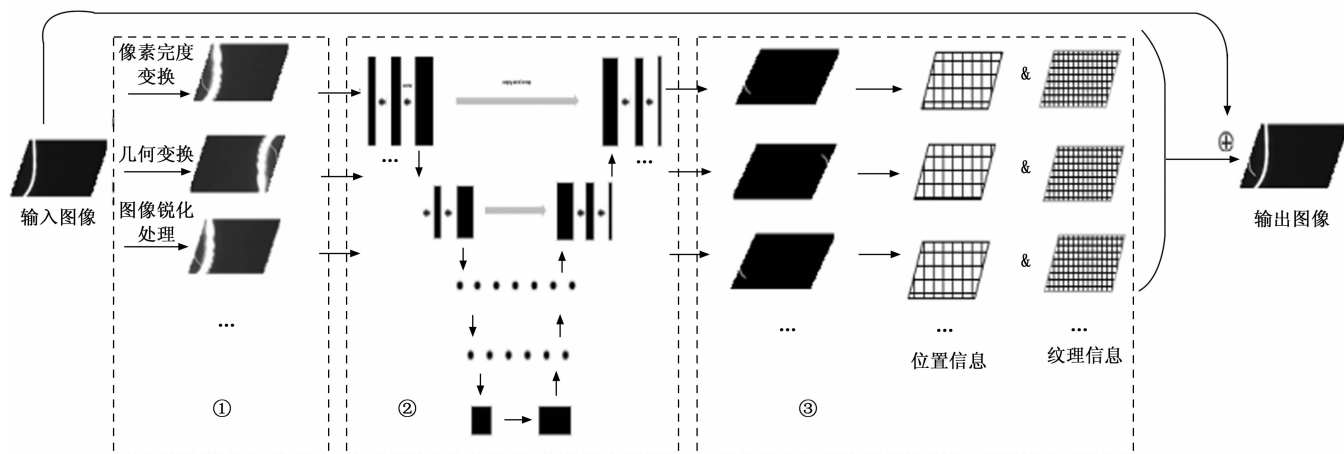


图 3 多层特征增强模块

### 2.1 多层特征增强模块

手机屏幕图像为灰度图像, 其中的缺陷通常范围为几十像素甚至十几像素。缺陷特征隐晦, 缺乏丰富的特征信息, 特征提取层提取到的特征信息缺少该缺陷应具有的关键特征信息, 严重影响后续网络层回归和分类, 导致检测效果很差。因此, 改进网络让特征提取层提取到丰富特征信息, 是手机屏幕缺陷检测的关键。

多层特征增强模块主要采用多组图像预处理和 U-Net 的串联, 将手机屏幕缺陷图像中隐晦的缺陷特征信息进行针对性像素加强, 进而达到增强图像特征的效果。多层特征增强模块结构如图 3 所示, 该模块分为 3 个阶段。

第一个阶段对输入图像分别进行像素亮度变换、几何变换、图像锐化等图像处理, 得到多组不同图像处理下的原始图像。不同缺陷的特征会因为对应的预处理操作而有效提高了可视程度, 各种缺陷特征更清晰, 不同类型的缺陷的特征信息得到了不同程度的特征增强, 有效为第二阶段的 U-Net 语义分割提供特征明显的原始图像。

第二阶段, 各组图像输入已训练好的 U-Net 进行语义分割操作, 得到多组对应的预测的掩模图。记录各组图像的掩模所在位置信息和对应原图上的纹理特征。

第三阶段, 根据对应的位置信息和纹理特征, 在原始输入图像对应的位置叠加对应的纹理信息, 对同一张原始图像进行相应的特征增强, 得到缺陷特征明显增强的输出图像。

手机屏幕图像经过多层特征增强模块处理的先后效果对比如图 4 所示。容易看到, 相比原手机屏幕图 (左), 手机屏幕缺陷特征变得更加明显, 更加突出。其原理如下:

- 1) 图像预处理的使用, 扩大图像中不同目标特征之间的差别, 抑制不感兴趣的特征, 使之改善图像质量, 丰富信息量, 加强图像判读和识别效果。
- 2) U-Net 串联在图像预处理之后, 可以根据 U-Net 判断每张图像每个像素点的类别得到精确的分割图。记录下分割图的掩膜位置信息以及位置信息对应相应原始图像的

位置纹理信息。最终, 利用这些位置信息和对应的纹理信息在原始图像上进行叠加, 针对性的对原始图像进行特征增强。

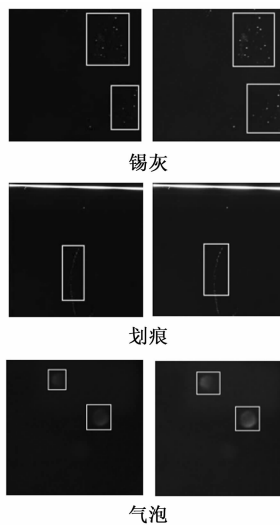


图 4 效果对比

左列为未处理的手机屏幕缺陷图, 右列为经过多层特征增强模块后的手机屏幕缺陷图。方框仅为缺陷所在位置方便读者阅读, 非预测结果。

### 2.2 多尺度特征提取网络

手机屏幕缺陷图像的缺陷目标尺寸很小, 多为小目标, 一般只有几十甚至几个像素。对于小目标, 当卷积池化到最后一层, 对于一个 RoI 区域映射到特征图的特征信息已经很少了。另外, 卷积网络中, 深层网络能响应高语义特征, 但特征图尺寸小, 拥有太少定位信息; 浅层网络虽然包含定位信息多, 但高语义特征比较少, 不利于分类。这个问题在手机屏幕缺陷检测中尤为突出, 严重影响检测效果。

改进 Faster R-CNN 的特征提取层, 加入特征金字塔网络 (FPN, feature pyramid network), 改进为多尺度特征提

取网络。通过融合上采样后的高语义特征和浅层的定位细节特征，使模型能够融合多个卷积层的多尺度信息，增强特征的表达能力，从而获得更多关于手机屏幕小目标缺陷的有效特征信息，进而提高网络性能。

多尺度特征提取网络结构如图 5 所示。语义信息增强部分，用 ResNet50 作为特征提取网络 {C2、C3、C4、C5}，采用 1×1 卷积进行横向连接，自上而下地传递语义信息构造 {T2、T3、T4、T5}， $T_i$  的计算过程为：

$$T_i = \begin{cases} F_{1 \times 1}(C_i); i = 5 \\ F_{1 \times 1}(C_i) \oplus Up(T_{i+1}, 2); i = 2, 3, 4 \end{cases} \quad (1)$$

式中， $F_{1 \times 1}(C_i)$  是对  $C_i$  进行 1×1 卷积操作， $Up(T_{i+1}, 2)$  表示对  $T_{i+1}$  进行 2 倍上采样。表示 merge 操作。

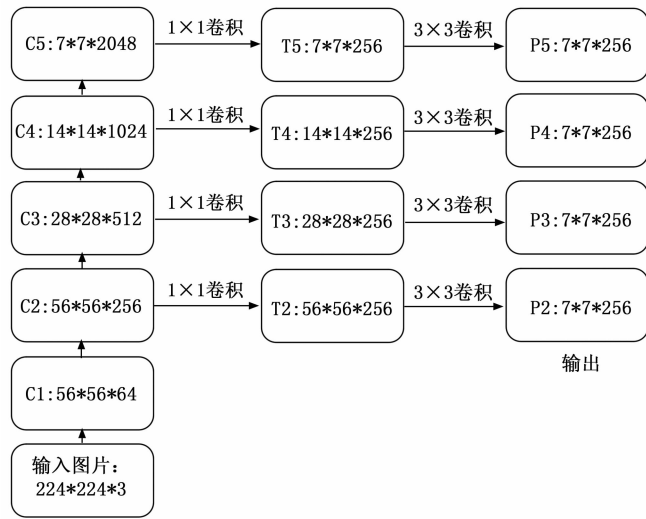


图 5 多尺度特征提取网络

$P_i$  为融合特征部分，其计算过程为：

$$P_i = F_{3 \times 3}(T_i); i = 2, 3, 4, 5 \quad (2)$$

$F_{3 \times 3}(T_i)$  是对  $T_i$  进行 3×3 卷积操作。

### 2.3 自适应区域建议网络

手机屏幕图像中，同种缺陷的形状大小各不相同，尺寸跨度极大。因此以人工设置方式在区域建议网络设置的 Anchor box 模板尺寸和数量与 ground truth box 存在一定偏差，会影响检测效果。

为提升检测效果，RPN (region proposal network) 层引入了 Isodata<sup>[14]</sup> 自迭代聚类算法，改进为自适应区域建议网络。通过对 Anchor box 数据进行自迭代聚类，得到合适的 Anchor box 模板。灵活数量和大小 Anchor box 模板可以有效提高手机屏幕缺陷检测的精度。

Anchor box 基本概念为以特征图上可以映射回原图像上的一个点为中心，预先人为设定  $x$  个 Anchor box 模板，称为在这个点上生成的  $x$  个 Anchor box。Anchor box 的生成方式如图 6 所示，在原始 Faster R-CNN 中，经过特征提取网络处理之后，得到一个尺寸为  $m \times n$  的特征图，对应将原图像划分为  $m \times n$  个区域，即原图的每个区域的中心由特征图上的一个像素点坐标表示。通过上述 Anchor 机制，在

每个区域生成  $x$  个 Anchor box， $x$  一般人为设置。本文改进的自适应区域建议网络中， $x$  为 isodata 自迭代聚类算法根据训练集标注尺寸样本聚类得到的聚类中心的数目，各个聚类中心的结果作为 Anchor box 的模板尺寸。

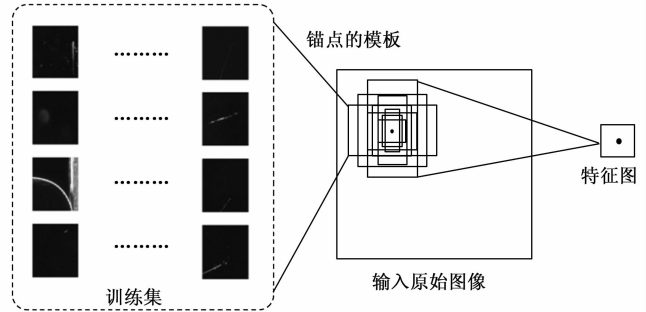


图 6 anchor box 的生成方式

自适应区域建议网络算法流程如图 7 所示。自适应区域建议网络基本步骤如下：

- 1) 初始设定  $K_0$  个 Anchor box 模板类，每个模板类中最少的样本数为  $N_{min}$ ，输入  $n$  个训练集标注尺寸样本，最大迭代次数  $I$ 。
- 2) 针对每个训练集标注尺寸样本，计算它与模板类中心的距离，将其分入距离最小的模板类中。此时模板类的数目为  $K_1$ 。
- 3) 判断每个模板类中样本数目  $N$  是否低于最少样本数  $N_{min}$ ，低于则取消该模板类，类中样本按最小距离原则重新分配到剩余类，令  $K_1$  减 1。
- 4) 根据各模板类中数据，重新计算模板类聚类中心。
- 5) 若  $K_1 < K_0/2$ ，说明类数量不够，执行分裂操作；若  $K_1 > 2K_0$ ，说明类数量太多，执行合并操作。
- 6) 达到最大迭代次数时终止，输出  $K = K_1$ ，否则返回步骤 2 继续迭代。 $K_1$  即为设置 Anchor box 的模板数量，各模板类聚类中心则为 Anchor box 模板。

算法根据计算样本与类中心的欧式距离进行分类。距离越小，相似度越大。计算公式为：

$$d(S_i, u_j) = \|S_i - u_j\|_2 \quad (3)$$

式中， $S_i$  是第  $i$  个样本， $u_j$  为第  $j$  个聚类。

在原 RPN 网络引入 isodata 聚类算法，通过迭代聚类训练手机屏幕缺陷样本数据，能够获得合适的 Anchor box 模板和数量，增强 Anchor box 的适应性，解决屏幕缺陷尺寸跨度大的问题，提高检测精度。

### 2.4 损失函数

PU-Faster R-CNN 框架的损失函数由两部分组成，RPN 部分的损失和 Fast R-CNN 部分的损失。RPN 网络的损失函数是分类损失和回归损失的总和。损失函数计算公式如下：

$$L(p_i, t_i) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (4)$$

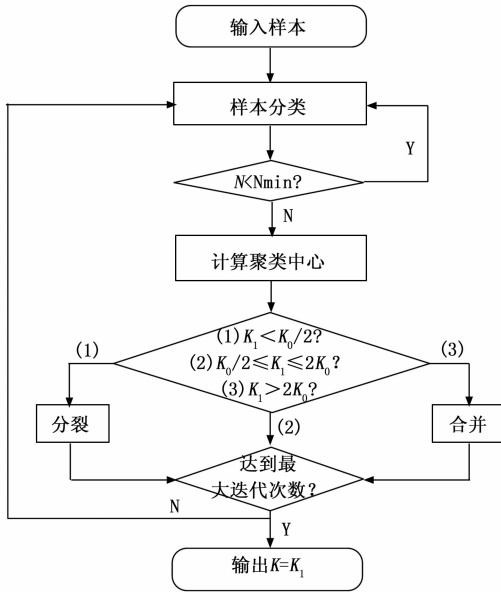


图 7 算法流程图

其中:  $P_i$  表示第  $i$  个 Anchor box 被预测为真实标签的概率。正样本时  $P_i$  为 1, 负样本时为 0。  $t_i$  表示预测第  $i$  个 Anchor 的边界框回归参数。  $t_i$  表示第  $i$  个 Anchor 对应的 ground truth box。  $N_{cls}$  表示一个小批次 256 中所有样本的数量。  $N_{reg}$  表示 Anchor box 位置的数量, 约为 2 400 个。

$$L_{cls}(p_i, p_i^*) = -\log[p_i^* p_i + (1 - p_i^*)(1 - p_i)] \quad (5)$$

其中:  $L_{cls}$  是分类损失。  $P_i$  表示第  $i$  个 Anchor box 被预测为真实标签的概率。正样本时  $P_i$  为 1, 负样本时为 0。

$$L_{reg}(t_i, t_i^*) = \sum_i smooth_{L1}(t_i - t_i^*) \quad (6)$$

$$smooth_{L1}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise} \end{cases} \quad (7)$$

$$t_i = [t_x, t_y, t_w, t_h] \quad (8)$$

$$t_x = (x - x_a)w_a \quad (9)$$

$$t_y = (y - y_a)h_a t_y \quad (10)$$

$$t_h = \log\left(\frac{h}{h_a}\right) \quad (11)$$

$$t_w = \log\left(\frac{w}{w_a}\right) \quad (12)$$

$$t_x^* = \frac{(x^* - x_a)}{w_a} \quad (13)$$

$$t_y^* = \frac{(y^* - y_a)}{h_a} \quad (14)$$

$$t_w^* = \log\left(\frac{w^*}{w_a}\right) \quad (15)$$

$$t_h^* = \log\left(\frac{h^*}{h_a}\right) \quad (16)$$

其中:  $P_i$  为正样本时为 1, 为负样本时为 0。  $t_i$  表示预测第  $i$  个 Anchor box 的边界框回归参数。  $t_i$  表示第  $i$  个 Anchor box 对应的 ground truth box。

Fast R-CNN 的损失函数部分是 PU-Faster R-CNN 第二部分的损失函数。Fast R-CNN 损失函数由分类损失和边界盒回归损失组成。

$$L(p, u, t^u, v) = L_{cls}(p, u) + \lambda[u \geq 1]L_{loc}(t^u, v) \quad (17)$$

其中:  $p$  是分类器预测的 softmax 概率分布。  $u$  对应于目标的实际类别标签。  $t^u$  则对应于由边界框回归器预测的相应类别  $u$  的回归参数。  $v$  对应于实际目标的边界框回归参数。

$$L(p, u, t^u, v) = -\log p_u \quad (18)$$

$$L_{loc}(t^u, v) = \sum_{i \in \{x, y, w, h\}} smooth_{L1}(t_i^u - v_i) \quad (19)$$

$$smooth_{L1}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise} \end{cases} \quad (20)$$

其中:  $t^u$  对应于相应类别  $u \cdot v$  的边界框回归因子预测的回归参数, 而  $v$  对应于实际目标的边界盒回归参数。

### 3 实验结果与分析

#### 3.1 数据集构建与实验设置

实验的数据集来自广东省某手机屏幕生成企业, 原手机实体屏幕为透明玻璃面板, 需要在指定工控机上通过光源照射, 在背景板采集黑白的手机屏幕图像。原始手机屏幕图像如图 8 所示。



图 8 原始手机屏幕图像

原始手机屏幕图像分辨率为  $6\,400 \times 6\,400$  像素, 图像尺寸太大, 而缺陷只有十几像素甚至几像素, 导致目标检测效果不佳。为了获得更加丰富多样的数据, 同时深度学习模型的训练要求大量的样本, 实验中对数据集进行了数据增强。实验中所采用的数据增强方法为: 1) 对原始图片进行  $10 \times 10$  的切割, 最终图片分辨率为  $640 \times 640$  像素; 2) 对于切割后的数据集进行像素变换方法: 调节亮度, 饱和度, 直方图均衡化, 高斯模糊; 3) 对于切割后的数据集进行几何变换方法: 翻转, 旋转, 裁剪。

数据集的标注采用 labelImg 目标检测数据集标注工具, 对每一张数据增强后的图中的缺陷进行标注, 每张图像形成一个 xml 文件, 每个 xml 文件记录了对应图像所包含的缺陷的种类及缺陷位置信息。

数据集的生成按照 PASCAL VOC 2012 数据集的相关

目录结构和格式生成，最终生成 txt 文件分别记录每个 jpg 图像名和 xml 文件名。

实验的数据集包含 3 种类别的缺陷：划痕 (scratch)、气泡 (bubble) 和锡灰 (tin\_ash)。经过处理后的，数据集图片数量共 6 659 张，大致按 5:1 的比例随机选取 5 407 张图片为训练集，1 252 张图片为测试集，如表 1 所示。在有缺陷的图片中，具体缺陷的数量分布见表 2。

表 1 训练集、测试集数量分布

|     | 有缺陷图片   | 无缺陷图片 | 总计      |
|-----|---------|-------|---------|
| 训练集 | 5 142 张 | 265 张 | 5 407 张 |
| 测试集 | 1 191 张 | 61 张  | 1 252 张 |

表 2 训练集和测试集的缺陷的数量分布

|     | 气泡      | 划痕      | 锡灰      |
|-----|---------|---------|---------|
| 训练集 | 2 284 个 | 5 430 个 | 4 006 个 |
| 测试集 | 1 297 个 | 952 个   | 1 079 个 |

实验采用的计算机硬件配置为 Intel (R) Xeon (R) CPU E5-2 650 v4 @ 2.20 GHz, NVIDIA GeForce RTX 3 080 GPU。用 Python 编程实现改进的 PU-Faster R-CNN，用于训练和测试。整个实验是使用开源深度学习框架 pytorch 实现的。输入图像尺寸为 640×640 像素。Backbone 选择了 ResNet50。选择优化算法为 Adam 算法。学习率为  $1e-3$ ，使用到的学习率下降方式为 cos。训练步数设置为 3 000。在获得足够的数后，使用自适应矩估计 (Adam) 优化算法进行训练，动量为 0.9，权重衰减为 0.000 1，beta1 为 0.9 和 beta2 为 0.999。训练批量大小设置为 16，以避免局部极小。模型置信度为 0.5，非极大值抑制 nms-iou 参数为 0.3。实验分为两部分，第一部分为 PU-Faster R-CNN 模型和主流的缺陷检测算法的对比实验，参与比较的模型有 Yolo 系列模型，原始 Faster R-CNN 模型和 SSD 模型。第二部分是消融实验，用以验证分析 PU-Faster R-CNN 各模块对提高手机屏幕缺陷检测性能的有效性。

### 3.2 实验结果与讨论

论文实验评价指标为目标检测领域常用的评价指标：准确率 Accuracy，精确率 Precision，召回率 Recall，准确度 Accuracy of Precision (AP)，漏检率 Miss rate，错误率 Error rate。

#### 3.2.1 对比实验

实验中，为验证所提出的 PU-Faster R-CNN 对手机屏幕缺陷检测的有效性，将结合不同的主流目标检测框架在手机屏幕数据集场景下进行对比实验。其中参与实验的框架有 Faster R-CNN、YoloV3、YoloV4 和 SSD 算法。实验采用数据集见前文，PU-Faster R-CNN 针对手机屏幕数据集下的每种缺陷检测效果如图 9 所示。PU-Faster R-CNN 针对手机屏幕数据集下预测每种缺陷的 AP 和总 mAP 如图 10 所示，可以看出，改进后的 PU-Faster R-CNN 模型在各

个类别的手机屏幕缺陷目标检测性能效果都很好。

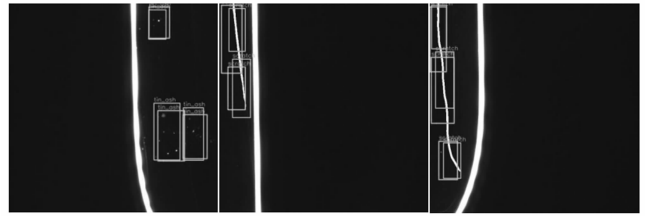


图 9 手机屏幕缺陷检测效果图

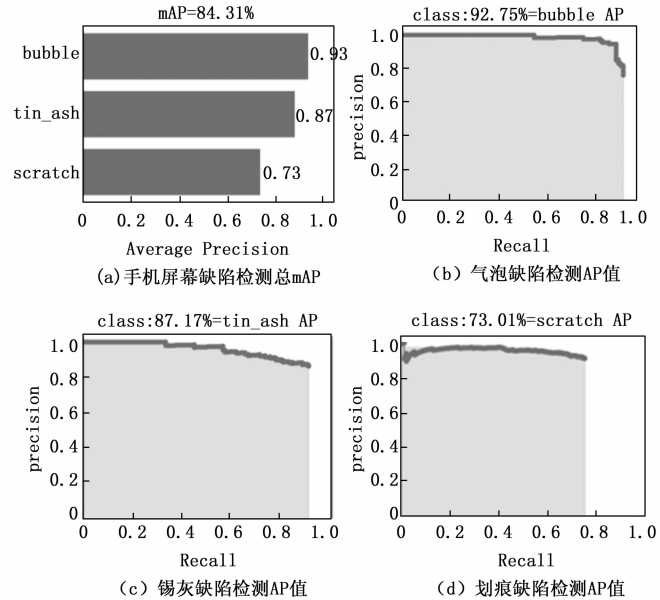


图 10 总 mAP 和每类缺陷预测的 AP 值

表 3 分别显示了 PU-Faster R-CNN 和手机表面数据集上的其他目标检测框架之间的性能比较。与 Faster R-CNN 相比，PU-Faster R-CNN 在 mAP 指标上提高 36.7%、准确率提高了 55.5%、精确度提高了 52.8%、召回率提高了 60.4%、漏检率降低了 60.1%、错误率降低了 55.5%。框架检测性能提升的原因在于：

- 1) 多层特征增强模块可以显著增强缺陷的语义信息，降低后续检测的难度和提高了检测精度；
- 2) 引入多尺度特征提取网络，提高了模型对多尺度目标的特征提取能力；
- 3) 自适应区域建议网络的提出，通过生成尺寸、数量更合适的 Anchor 模板，提高预测的精度和框回归的效率。

与 YoLo 系列模型相比，以 YoLoV4 为例，PU-Faster R-CNN 在 mAP 性能上提高 12.0%、准确率上提高了 4.4%、精确度提高了 21.4%、召回率提高了 7.8%、漏检率降低了 7.5%、错误率降低了 7.5%。原因在于：YoLoV4 骨干部分的 BoF 特征包括特征 CutMix 和 Mosaic 增强、DropBlock 正则化和类别标签平滑化，对于特征明显的缺陷检测图像有优秀的效果，而手机屏幕缺陷图片中的缺陷特征隐晦，YoLoV4 提取到的特征信息掺杂着一些非关

键信息,严重影响回归与分类,导致检测效果略差。PU-Faster R-CNN 提出多层特征特权模块,通过语义分割,让用于区分缺陷类别的轮廓特征和纹理特征变得明显,提升了模块对缺陷特征的表达能力,因而检测性能比 YOLO 系列算法更胜一筹。

与 SSD 模型相比,PU Faster R-CNN 在 mAP 性能指标上提高 24.2%、准确率上提高 50.9%、精确度提高了 65.0%、召回率提高了 0.9%、漏检率降低了 0.7%、错误率降低了 50.9%。原因在于:手机屏幕缺陷多为小目标缺陷。在众多目标检测算法中,SSD 对于小目标的检测表现较差,因为经过多层卷积,SSD 对小目标提取的缺陷特征信息丰富度缺乏。PU-Faster R-CNN 多尺度特征提取网络融合上采样后的高语义特征和浅层的定位细节特征,充分提取小目标缺陷的特征信息,故能提高性能。

表3 PU-Faster R-CNN 与主流技术各项指标对比

|                 | Accuracy | Precision | Recall | mAP   | Error rate | Miss rate |
|-----------------|----------|-----------|--------|-------|------------|-----------|
| Faster R-CNN    | 16.5%    | 36.2%     | 19.6%  | 47.6% | 83.5%      | 80.4%     |
| YoloV3          | 33.0%    | 88.4%     | 34.1%  | 47.6% | 67.0%      | 65.9%     |
| YoloV4          | 67.6%    | 67.6%     | 72.2%  | 72.3% | 32.4%      | 27.8%     |
| SSD             | 21.1%    | 24.0%     | 79.1%  | 60.1% | 78.9%      | 21.0%     |
| PU-Faster R-CNN | 72.0%    | 89.0%     | 80.0%  | 84.3% | 28.0%      | 20.3%     |

### 3.2.2 消融实验

本实验将验证分析 PU-Faster R-CNN 各模块对提高手机屏幕缺陷检测性能的有效性。如表 4,实验中使用了 4 种不同的方案。方案一选择原始 Faster R-CNN 作为整个框架;方案二选择基于 Faster R-CNN 引入多尺度特征提取网络;方案三选择基于 Faster R-CNN 引入多尺度特征提取网络和多层特征增强模块;方案四是完整的 PU Faster R-CNN。4 个不同实验方案分别在手机屏幕数据集得到四组实验性能结果。

方案二中基于原始 Faster R-CNN 加入多尺度特征提取网络,较方案一在 mAP 指标上提高了 5.8%,准确率提高了 15.9%,精确率提高了 8.4%。原因在于多尺度特征提取网络融合了浅层网络的定位信息和深层网络的高语义特征,实现了特征的跨层连接,进而得到不同尺度的特征图。相比原始 Faster R-CNN 使用一组基础的卷积层+ReLU 激励层+池化层提取图像特征的方法,能提取到更加丰富的特征信息,用于后续分类与回归,故能提高性能。

方案三中基于原始 Faster R-CNN 加入多尺度特征提取网络和多层特征增强模块,较方案二在 mAP 性能指标上提高了 19.1%,准确率提高了 30.0%,精确率提高了 13.2%。方案三相比于方案一,在 mAP 性能指标上提高了 24.7%,准确率提高了 45.9%,精确率提高了 21.6%。数据集中手机屏幕图片特征非常隐晦,缺陷分布不规律,纹理、背景复杂,提取特征极度困难,且缺陷形状大小等属

性上的特征呈现多样性,即使同一种缺陷,尺寸和形状上也存在多种分布。多层特征增强模块通过图像预处理和 U-Net 网络的结合,对不同类型的缺陷特征信息进行特征增强,再经过语义分割得到特征的位置信息以及纹理特征,叠加到原图上,消除了图像无关信息,增强了缺陷信息的轮廓特征和纹理特征,最终从隐晦的手机屏幕图片中提取到丰富的特征信息。

方案四中为完整的 PU-Faster R-CNN,基于原始 Faster R-CNN 加入了多尺度特征提取网络、多层特征增强模块和自适应区域建议网络,较方案三在 mAP 性能指标上提高了 12.0%、准确率提高了 9.6%、精确率提高了 31.2%。方案四相比于方案二,在 mAP 性能指标上提高了 30.9%,准确率提高了 39.6%,精确率提高了 44.4%。方案四相比于方案一,在 mAP 性能指标上提高了 36.7%,准确率提高了 55.5%,精确率提高了 52.8%。原始 Faster R-CNN 中 RPN 网络的 Anchor box 模板尺寸和数量要求人为设置,而在手机屏幕缺陷检测场景下,缺陷形状大小不一,跨度极大,人为设置的 Anchor box 模板与实际 ground truth box 存在较大差异。PU-Faster R-CNN 的自适应区域建议网络对训练集样本进行自迭代聚类,将聚类得到的聚类中心和聚类中心的数量分别作为 Anchor box 的模板尺寸和模板数量,这样得到的 Anchor box 模板与缺陷尺寸的拟合性更高,有效提高屏幕缺陷检测的性能。

表4 PU-Faster R-CNN 消融实验性能指标

|                         | 多尺度特征提取网络 | 多层特征增强模块 | 自适应区域建议网络 | mAP   | Accyrcacy | Precision |
|-------------------------|-----------|----------|-----------|-------|-----------|-----------|
| Faster R-CNN (Resnet50) | ×         | ×        | ×         | 47.6% | 16.5%     | 36.2%     |
| Faster R-CNN (Resnet50) | √         | ×        | ×         | 53.4% | 32.4%     | 44.6%     |
| Faster R-CNN (Resnet50) | √         | √        | ×         | 72.3% | 62.4%     | 57.8%     |
| Faster R-CNN (Resnet50) | √         | √        | √         | 84.3% | 72.0%     | 89.0%     |

## 4 结束语

长期以来,缺陷检测问题一直困扰着工业制造业。手机表面缺陷检测作为工业生产的重要组成部分,越来越受到重视。然而,图像特征隐晦、缺陷尺寸差异大等问题限制了基于深度学习的检测框架。因此,基于 PU-Faster R-CNN 的手机屏幕缺陷检测框架被提出解决上述问题。针对图像特征不明显,设计了多层特征增强模块,采用一种新颖的方式增强缺陷的特征信息。构建多尺度特征提取网络,有效提取多尺度的缺陷语义信息。论文中设计了自适应区域建议网络,自适应的 Anchor box 模板生成方式,通过聚类的方式生成拟合性更好的 Anchor box 模板。实验结果表明,该模型在手机屏幕检测中优于流行的目标检测模型。

最后,该框架适用于低复杂性数据集。然而,当数据集的复杂性增加时,检测结果可能不太令人满意。复杂缺陷的检测也是未来工作的一个新的研究方向。该框架有望成为一种新的目标检测解决方案。它可以应用于手机表面生产,包括应用场景。

#### 参考文献:

- [1] HUANG H, CHAO H, TIAN W, et al. Surface defects detection for mobile phone panel work pieces based on machine vision and machine learning [C] // 2017 IEEE International Conference on Information and Automation (ICIA). IEEE, 2017.
- [2] JIAN C, GAO J, AO Y. Automatic surface defect detection for mobile phone screen glass based on machine vision [J]. Applied Soft Computing, 2017, 52: 348 - 358.
- [3] GUO T Y, ZHANG L L, DING R W, et al. Edd-net: An efficient defect detection network [C] // 2020 25th International Conference on Pattern Recognition (ICPR), IEEE, 2021: 8899 - 8905.
- [4] LV Y, MA L, JIANG H. A Mobile Phone Screen Cover Glass Defect Detection MODEL Based on Small Samples Learning [C] // 2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP). IEEE, 2019.
- [5] WANG Z, LI C, ZHANG Y, et al. Improved Loss Function for Defect Detection of Mobile Phone Screen [C] // Proceedings of the 13th EAI International Conference on Mobile Multimedia Communications, Mobimedia 2020, 2020, 27 - 28.
- [6] REN B Y, LI Z Y, DAI Y. Method for detection of slight scratch of mobile phone screen [J]. Journal of Harbin Institute of Technology, 2021, 53 (1): 29 - 36.
- [7] ZHU Y, DING R, HUANG W, et al. HMFCA-Net: Hierarchical Multi-Frequency based Channel Attention Net for Mobile Phone Surface Defect Detection [J]. Pattern Recognition Letters, 2022, 153: 118 - 125.
- [8] LEI J, GAO X, FENG Z, et al. Scale Insensitive and Focus Driven Mobile Screen Defect Detection in Industry [J]. Neurocomputing, 2018, 294: 72 - 81.
- [9] WANG G, HAN S, DING E, et al. Student-teacher feature pyramid matching for unsupervised anomaly detection [J]. Computer Vision and Pattern Recognition, 2021: 3 - 7 [2023 - 2 - 23]. <https://doi.org/10.48550/arXiv.2103.04257>.
- [10] TAN M X, PANG R M, QUOC V. LE. Efficient Det: Scalable and Efficient Object Detection [J]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, 10781 - 10790.
- [11] CHEN X P, WEI W, YANG X, et al. Successive linear Newton interpolation methods for solving the large-scale nonlinear eigenvalue problems [J]. Applied Mathematics and Computation, 2020: 12 - 15 [2023 - 2 - 23]. <https://doi.org/10.1016/j.amc.2019.124663>.
- [12] QIAO S, CHEN L C, YUILLE A. Detectors: Detecting objects with recursive feature pyramid and switchable atrous convolution [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2021: 10213 - 10224.
- [13] LONG J, SHELHAMER E, DARRELL T. Fully Convolutional Networks for Semantic Segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39 (4): 640 - 651.
- [14] WANG Z W, WANG C B. Application and research on foundation monitoring using ISODATA based fuzzy cluster analysis algorithm [C] // 2013 Fourth International Conference on Digital Manufacturing and Automation, Qingdao, 2013: 264 - 266.
- [15] BERGMANN P, LOWE S, FAUSER M, et al. Improving unsupervised defect segmentation by applying structural similarity to autoencoders [C] // International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP), February 25 - 27, 2019, Prague, Czech Republic. Valtetta: SciTePress, 2019: 372 - 380.
- [16] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117 - 2125.
- [17] LI X D, YANG W D, DEZERT J. An airplane image target's Multi-feature fusion recognition method [J]. Acta Automatica Sinica, 2012, 38 (8): 1298 - 1307.
- [18] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] // Proceedings of the 2016 International Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 770 - 778.
- [19] 李润青, 谢明鸿, 黄冰晶. 一种基于初始点密度最大的改进型 ISODATA 聚类算法 [J]. 软件导刊, 2017, 16 (12): 94 - 98.
- [20] 边肇祺, 张学工. 模式识别 (第 2 版) [M]. 北京: 清华大学出版社, 2000.
- [21] 熊红林, 樊重俊, 赵 珊, 等. 基于多尺度卷积神经网络的玻璃表面缺陷检测方法 [J]. 计算机集成制造系统, 2020, 26 (4): 900 - 909.
- [22] 张红云, 刘向东, 段晓东, 等. 数据挖掘中的聚类算法比较研究 [J]. 计算机应用与软件, 2003, 20 (2): 5 - 6
- [23] 徐振飞, 刘尧峰, 徐 卫. 基于机器视觉的手机白玻璃缺陷检测方法研究 [J]. 中国新通信, 2019 (4): 120 - 121.
- [24] 简川霞, 高 健. 手机玻璃屏表面缺陷视觉检测方法研究 [J]. 包装工程, 2018, 39 (5), 16 - 21.
- [25] 王昌书. 基于机器视觉的手机曲面玻璃典型缺陷检测研究 [D]. 广东: 华南理工大学, 2019.
- [26] 赵 俊, 鹿晓威, 赵 骥, 等. 基于改进 Faster RCNN 的目标检测算法 [J] 辽宁科技大学学报, 2021, 44 (4), 263 - 270.
- [27] 范彩霞, 蒋玉明. 模糊 ISODATA 聚类的 BP 神经网络及其应用 [J]. 计算机工程与应用, 2005, 41 (35): 226 - 228.