

遮挡与缺失场景下屏幕缺陷视觉检测

尹东富¹, 杜明臣², 胡天昊³, 李又明², 张笑虹², 于非¹

1) 人工智能与数字经济广东省实验室(深圳), 广东深圳 518107; 2) 深圳大学微众银行金融科技学院, 广东深圳 518060;
3) 深圳市国电科技通信有限公司, 广东深圳 518109

摘要: 为提高缺陷检测的智能化水平, 解决复杂场景下缺陷检测难题, 结合目标检测与目标匹配、图像差分技术, 搭建一种遮挡与缺失场景下屏幕缺陷视觉检测系统. 该系统包括手机屏幕检测、遮挡检测、缺失检测及屏幕内容变化检测4部分. 使用YOLOv8n模型检测出图像中的手机屏幕位置, 使用多台相机及检测框过滤方式检测出遮挡的手机屏幕, 使用目标匹配算法判断手机屏幕是否存在缺失, 再通过改进的图像差分算法判断手机屏幕显示内容是否发生变化. 通过生产现场采集数据, 并对数据进行标注、增强及过滤等处理, 然后对目标检测模型进行训练与测试, 所得手机屏幕平均检测准确率为96.8%. 在生产现场部署并应用算法模型, 通过对应目标匹配及差分计算, 实现了多路视频同时处理下手机屏幕缺陷的实时准确检测. 相关算法的应用既可有效减少人工成本又能提高检测效率. 该检测技术可用于针对手机、电脑和电视等电子屏幕的质量检测.

关键词: 人工智能; 图像处理; 目标检测; 缺陷检测; 目标匹配; 图像差分

中图分类号: TP311

文献标志码: A

DOI: 10.3724/SP.J.1249.2023.06631

Visual detection of screen defects in occlusion and missing scenes

YIN Dongfu¹, DU Mingchen², HU Tianhao³, LI Youming²,
ZHANG Xiaohong², and YU Fei Richard¹

1) Guangdong Laboratory of Artificial Intelligence and Digital Economy (SZ), Shenzhen 518107, Guangdong Province, P. R. China

2) WeBank Institute of FinTech, Shenzhen University, Shenzhen 518060, Guangdong Province, P. R. China

3) China Gridcom Co. Ltd., Shenzhen 518109, Guangdong Province, P. R. China

Abstract: To improve the intelligent level and solve the difficult problem of defect detection in complex scenarios, a visual detection system for screen defects in occluded and missing scenes is constructed based on object detection and matching, as well as image difference technology. The modules for mobile phone screen detection, occlusion detection, missing detection, and screen content change detection are established in this system. The YOLOv8n model is used to detect the position of mobile phone screens in images. Multiple cameras and detection boxes are used to filter out obstructed mobile phone screens. Target matching algorithms are used to determine whether there are any missing mobile phone screens, and an improved image difference algorithm is designed to determine whether the displayed content of mobile phone screen has changed. The data is collected at the practical production site, and then annotated, enhanced, and filtered. The target detection model is trained and tested, and the average detection accuracy of mobile screen is 96.8%. The deployment and application of developed model on the production site

Received: 2023-05-04; **Accepted:** 2023-06-21; **Online (CNKI):** 2023-10-22

Foundation: National Natural Science Foundation of China (62271324); Fund of Shenzhen Association for Science and Technology (2023KCSZ01); Fund of China Gridcom Co. Ltd. (23220008)

Corresponding author: Professor YU Fei Richard (yufei@gml.ac.cn)

Citation: YIN Dongfu, DU Mingchen, HU Tianhao, et al. Visual detection of screen defects in occlusion and missing scenes [J]. Journal of Shenzhen University Science and Engineering, 2023, 40(6): 631-639. (in Chinese)



have been completed. Through corresponding target matching and differential calculation, real-time and accurate detection of mobile phone screen defects under simultaneous processing of multiple videos has been achieved. The application of proposed algorithm can effectively reduce labor costs and improve detection efficiency. The proposed method can be used for electronic screen quality detection in fields such as mobile phones, computers and televisions.

Key words: artificial intelligence; image processing; target detection; defect detection; target matching; image differentiation

显示屏幕是智能手机、平板电脑、笔记本电脑和电视等电子产品中最重要的部件之一,它直接影响到产品的视觉效果和用户的使用体验.在生产过程中对电子屏幕进行质量检测是一道重要的工序.传统的视觉检测方法大多依靠人工先验知识进行特征提取,对光照、相机位置等变化较敏感,且当被检测物体发生变化后,特征提取规则需重新设计和开发,导致算法的泛化性和复用性不佳.与之相比,深度学习神经网络方法的适应性更好,通用性更广.目标检测指依靠深度学习神经网络找出图像中感兴趣的目标,并确定这些目标的类别和位置.代表性的目标检测模型包括经典的双阶段目标检测模型 Fast R-CNN (fast region-based convolutional neural network)^[1]、兼顾速度和精度的 YOLO (you only look once) 系列模型^[2]、基于无锚点的 CenterNet 模型^[3]、基于 Transformer 架构的 Swin Transformer 模型^[4]、基于神经网络结构搜索的 NAS-FPN 模型^[5]、适合移动端应用的轻量化模型 MobileDenseNet^[6]等.这些模型在检测精度、速度和鲁棒性方面各具优势,因此被广泛应用于各类缺陷检测中.

随着神经网络技术的发展,目标检测模型在电子产品检测^[7-10]、屏幕检测等领域得到了广泛应用. JIANG 等^[11]提出基于 U-Net 模型对屏幕弱划痕和变色缺陷进行检测的方法. LI 等^[12]设计了一种适用于各种类型屏幕检测的预检测和粗精缺陷提取算法,并为了避免误检测或漏检,提出一种结合多层感知器 (multilayer perceptron, MLP) 和深度学习技术的分类算法. YANG 等^[13]提出一种基于 YOLOv5s 和 Ghost 骨干网络的手机屏幕缺陷检测模型,其注意力模块使用协调注意力和深度可分离卷积来减少参数.为快速准确地检测智能仪表液晶显示屏 (liquid crystal display, LCD) 的缺陷, PEI 等^[14]提出一种结合线段检测器 (line segment detector, LSD) 和深度学习的方法,检测液晶屏字符缺陷的准确率约为 96%. ÇELİK 等^[15]提出一种图像采集系统,并使用基于深度学习的对象检测器,实现了对 LCD 像素级别缺陷

的实时检测.但是,上述研究大都属于单帧图像目标检测领域,未有效利用多帧相邻图像内容,不适用于检测多帧图像间屏幕内容变化等场景.

由于检测现场环境复杂,检测时常会出现目标被遮挡及目标缺失等现象.目前常见的对目标遮挡的研究主要集中在解决单幅图像中的遮挡问题,一般是通过修改网络模型、损失函数或添加遮挡目标权重等方式来实现遮挡检测^[16-17].对目标缺失的研究则主要集中在标注框缺失与类别缺失等方面^[18-19].本研究主要针对的场景是多部手机倾斜地连续摆放、且受相机拍摄角度的限制所产生的屏幕遮挡场景,以及因手机的放入或取出等操作而产生的屏幕缺失场景.不同的应用场景使得现有目标检测领域的遮挡及缺失检测方法难以被直接应用.

根据业务场景的需求,在获取屏幕目标后,需判断图像对应位置处屏幕显示内容是否发生变化.相对于传统的图像像素分布统计和直方图比对等检测方法,基于不同帧间图像的差分计算方法可取得更好的效果.目前,图像间差分计算的方法已广泛应用于图像压缩^[20]和目标跟踪^[21]等技术领域,以及生物医学诊断^[22]、工业质检^[23]和光学遥感^[24]等应用领域.本研究通过对图像差分算法进行改进,提升了检测系统的鲁棒性.

本研究结合采用深度学习的目标检测及图像差分计算方法,对手机及屏幕是否存在缺陷进行了研究.首先使用目标检测模型检测出视野中的手机位置;然后使用多台相机及检测框过滤方式对被遮挡的手机屏幕进行检测;使用目标匹配算法对两幅图像中的检测框进行匹配,以判断手机屏幕是否存在缺失,并获得手机屏幕的绝对位置;最后通过屏幕变化检测算法判断手机屏幕是否发生变化.

1 目标检测模型与图像差分方法

1.1 YOLOv8 模型

YOLO 是一种单阶段的目标检测算法,将图像

数据输入到该网络模型可得到检测框的坐标和物体的类别, 其对目标的识别速度比两阶段的目标检测算法要快. 目前, YOLOv8 是 YOLO 系列模型中的最新版本, 支持图像分类、物体检测和实例分割任务. YOLOv8 较之前的版本在目标检测方面的改进主要包括: ① 删除了特征多尺度融合的上下采样路径聚合网络-特征金字塔网络(path aggregation network - feature pyramid network, PAN-FPN)模块, 将 C3 模块替换为了 C2f 模块, 令模型在保证轻量化的同时获得更加丰富的梯度流信息; ② 将耦合头变为解耦头, 分别输出类别与位置, 并重新融入无锚框的思想; ③ 使用分布焦点损失(distribution focal loss, DFL)与完全交并比(complete intersection over union, CIoU)损失作为分类损失函数, 并引入任务对齐分配样本匹配方式. YOLOv8n 是 YOLOv8 系列中的一个轻量化的版本, 其识别速度快、检测精度高, 占用内存小, 且能够在中央处理器(central processing unit, CPU)和图形处理器(graphics processing unit, GPU)等硬件平台上运行. 因此, 本研究主要使用 YOLOv8n 模型进行手机屏幕的检测.

1.2 图像差分算法

差分图像是目标场景在连续时间点图像相减所构成的图像. 广义的差分图像定义为目标场景在不同时间点所成图像的差别. 差分图像由目标场景在相邻时间点的图像相减得到的. 两图像间的差值为

$$\Delta x = x_1 - x_2 \quad (1)$$

其中, x_1 和 x_2 分别为目标场景在 t_1 和 t_2 时刻的图像.

帧间差分法广泛用于运动目标轮廓的检测, 它可削弱图像中的相似部分, 同时突出显示图像的变化部分. 当监控场景中出现异常运动目标时, 相邻两帧图像之间会出现较明显的差异. 两帧相减可求得图像对应位置像素值差的绝对值, 通过判断其是否大于阈值, 进而分析视频或图像序列中的物体运动特性.

图像差分算法具有实现简单、程序设计复杂度低、运行速度快、动态环境适应性强、对场景光线变化不敏感等优势. 本研究通过加入双重阈值验证过程来改进图像差分算法, 以提高缺陷检测结果的准确性.

2 方案设计

2.1 总体方案

屏幕缺陷视觉检测系统整体流程如图1. 针对采集到的视频提取其中的关键帧, 并使用 YOLOv8n 模型进行手机屏幕目标检测. 结合存储的图像模板与目标框位置信息, 采用缺失检测方法判断是否存在屏幕缺失现象, 并输出屏幕在工作台上的绝对位置. 在间隔固定时间后, 再次从采集到的视频中提取关键帧, 重复执行上述过程, 并输出手机屏幕在该帧图像中的绝对位置. 使用目标匹配方法对不同帧间的手机屏幕进行匹配, 再在对应的手机屏幕上进行屏幕变化检测. 当屏幕发生变化时, 判定屏

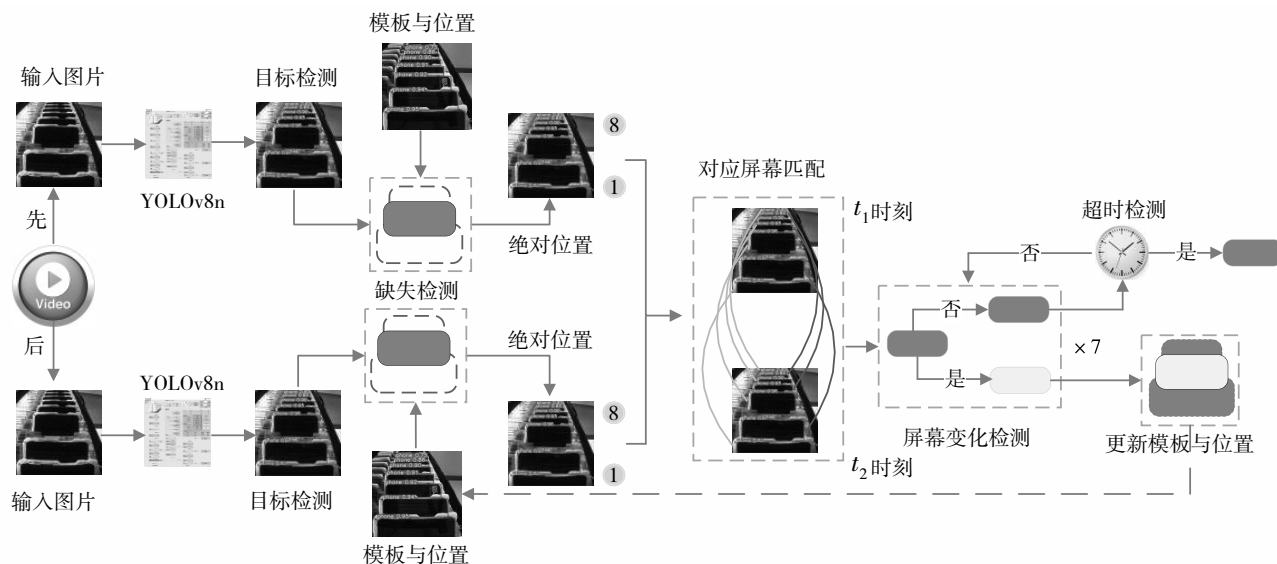


图1 屏幕缺陷视觉检测流程

Fig. 1 Visual inspection flow chart for screen defects.

幕未出现缺陷, 保存此时的图像及屏幕的位置坐标, 并更新预存的图像模板及屏幕的位置坐标. 若超过设定时间屏幕仍未变化, 则判定手机屏幕出现缺陷, 系统将输出此时手机屏幕的位置坐标并报警.

2.2 遮挡目标检测

在实际应用场景中, 将手机有规律地摆放在检测台上, 检测台每层最多可摆放 32 部手机. 为节省检测台长度方向的空间, 一般倾斜放置手机, 这使得相邻手机间的距离较近, 因此前后手机会存在较严重的遮挡现象. 为节省检测台高度方向的空间, 检测台层间高度通常较小, 顶部距离下方的手机较近, 这使安装的检测用相机(摄像头)只能拍摄到少量手机屏幕, 且手机距离相机越远, 其屏幕被遮挡现象越严重, 检测相机采集到的图像中显示的视屏幕内容也越少. 由于使用单个普通相机或广角相机均难以实现好的拍摄效果, 本研究主要通过布置多个相机(图 2)及采用阈值过滤的方法来解决这一问题.

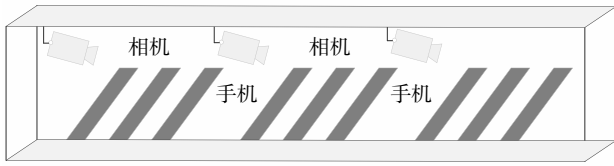


图 2 遮挡场景下相机与手机布置图

Fig. 2 Camera and phone layout in occlusion scenes.

多台相机同时采集图像并进行处理, 必然会在两台相机检测到同一手机屏幕的问题, 即某个手机屏幕同时出现在前一相机采集到的图像的尾部及后一相机采集到的图像的头部. 为此, 本研究分别在数据标注和结果筛选阶段都进行二次处理, 以此来保证每个手机检测结果的唯一性. 在进行数据标注时, 只标注出位于图像前部的数个手机屏幕, 神经网络只提取这几个前部手机屏幕的特征, 这样模型将不会检测出位于图像后方较小区域的屏幕. 为实现此目的, 系统通过设置检测框大小及高宽比对检测结果进行过滤, 去除远处误检的小目标, 保证手机屏幕不被重复检测识别. 本研究设置每个图像只检测前 8 部手机的屏幕.

2.3 目标匹配

由于生产现场的检测相机可能存在位置变动, 依靠检测框在图像中的绝对位置进行定位与匹配会使得结果不够准确. 本研究通过目标框的位置及相

互间的重合度来进行匹配, 使得小目标与小目标匹配、大目标与大目标匹配, 以提升目标匹配的准确性.

如图 3, 以 3 个检测出的屏幕目标匹配为例, 详细介绍对应的目标匹配过程. 假设在 t_1 时刻检测出的 3 个目标框分别为 A_1 、 B_1 和 C_1 , 在 t_2 时刻检测出的 3 个目标框分别为 A_2 、 B_2 和 C_2 , 对应目标匹配要解决的问题是判断 A_1 与 A_2 、 B_1 与 B_2 、 C_1 与 C_2 位置处于一一对应的关系.

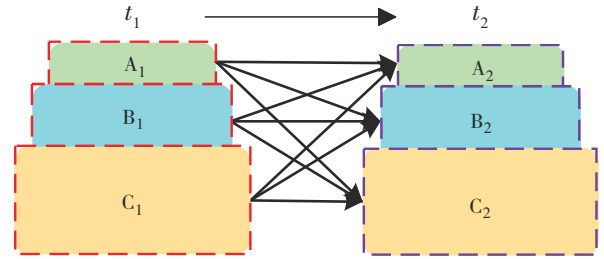


图 3 不同时刻目标重合度匹配示意

Fig. 3 (Color online) Schematic diagram of target coincidence matching at different times.

针对 t_1 与 t_2 时刻检测到的每个目标, 求它们的交并面积(intersection over area, IOA). IOA 值大于设定阈值时的匹配对即为对应的两个目标. 设 t_1 和 t_2 时刻检测到的目标(矩形框)面积分别为 S_1 和 S_2 , 则两目标的 IOA 值(S_{IOA})为

$$S_{IOA} = \max\left(\frac{S_1 \cap S_2}{S_1}, \frac{S_1 \cap S_2}{S_2}\right) \quad (2)$$

目标检测框为未旋转的矩形框, 目标算法输出每个检测框的长、宽及位置坐标信息. 矩形框面积通过其长乘以宽直接计算得出, 即

$$S_1 = (x_1^1 - x_0^1)(y_1^1 - y_0^1) \quad (3)$$

$$S_2 = (x_1^2 - x_0^2)(y_1^2 - y_0^2) \quad (4)$$

其中, (x_0^1, y_0^1) 和 (x_1^1, y_1^1) 分别为某检测框在 t_1 时刻的左上角和右下角坐标; (x_0^2, y_0^2) 和 (x_1^2, y_1^2) 分别为另一检测框在 t_2 时刻的左上角和右下角坐标. 因此, 有

$$S_1 \cap S_2 = (x_1 - x_0)(y_1 - y_0) \quad (5)$$

其中, (x_0, y_0) 和 (x_1, y_1) 分别为两检测框交集图形的左上角和右下角坐标; $x_0 = \max(x_0^1, x_0^2)$; $y_0 = \max(y_0^1, y_0^2)$; $x_1 = \min(x_1^1, x_1^2)$; $y_1 = \min(y_1^1, y_1^2)$.

由式(5)可见, 当 $x_1 < x_0$ 或 $y_1 < y_0$ 时, $S_{IOA} = 0$, 此时两个检测框不相交.

2.4 缺失目标检测

系统首次运行时, 存储手机屏幕未缺失时的图像作为模板, 并保存各屏幕检测框位置. 为保证结果的准确性, 采集多帧图像的数据计算并保存检测框的平均位置. 后续每次重启运行时, 都会重新采集图像及检测目标框的位置数据, 并替换掉原来的图像及目标框数据, 以保证模板处于最新状态.

当手机屏幕有缺失时, 不同时刻检测出的目标框数量会不同. 先根据不同时刻检测出的目标框数量的差值计算出缺失的屏幕数量, 再对不同帧之间的目标框进行匹配. 当其中一帧图像中的目标框无法在另一帧图像中找到匹配的目标框时, 则判定该位置处的手机屏幕缺失.

由相机成像原理可知, 物体距离相机越远, 其成像越小. 由于被测手机的屏幕尺寸相同, 且相邻两手机屏幕的间距相同, 故检测图像中的手机屏幕尺寸由近到远也呈比例缩小, 同时检测框的左上角坐标也将按照线性规律改变. 因此, 可根据检测框的左上角的坐标位置进行目标粗匹配. 依据目标框位置进行目标匹配判断虽然简单有效, 但因远处目标框较小, 易出现误匹配情况. 本研究采用重合度阈值对目标粗匹配结果进行验证, 来提升缺失位置检测的准确度.

同样以3个检测框为例, 本研究提出的缺失目标位置检测示意图如图4. 其中, 实线矩形框为检测出的目标框, 虚线矩形框为缺失的目标; A、B和C为不同的手机屏幕目标; A_U 、 B_U 和 C_U 分别指对应目标的上部分; A_L 、 B_L 和 C_L 分别指对应目标的下部分; 实心圆为目标框左上角; 带圈数字表示目标的绝对位置.

当后方目标A消失时, 目标B的面积(S_B)不变; 当前方目标C缺失时, 检测出的目标B的面积变大. 如图4(b)所示, 当前后目标同时缺失时, 目标B'的面积为 $S_{B'} = S_{B_U} + S_{B_L}$. 此时, 将图4(a)中3个目标分别与图4(b)中的目标进行匹配时, 目标A的面积与目标B'的面积不重合, 目标B的面积与目标B'的上部分面积 S_{B_U} 重合, 目标C的面积与目标B'的下部分面积 S_{B_L} 重合. 然后, 通过比对重合面积的大小以及 B_U 和 B_L 的位置, 可确定目标B与目标B'相对应, 进一步确定目标B'的绝对位置为②. 同理, 当前方连续目标缺失时, 将图4(a)中目标分别与图4(c)中目标A'进行匹配, 基于重合面积及坐标位置可确定目标A'的绝对位置为③. 由于后方缺

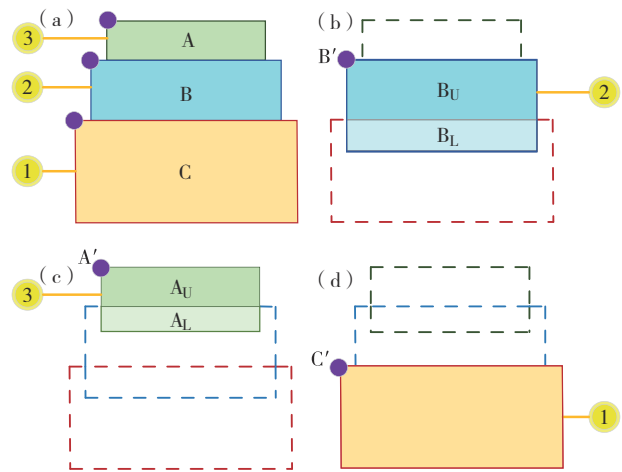


图4 缺失目标位置检测示意 (a)目标未缺失;(b)前后目标同时缺失;(c)前方连续多目标缺失;(d)后方连续多目标缺失(其中,①、②和③分别为前、中、后目标的绝对位置;实线矩形框为检测出的目标框;虚线矩形框为缺失的目标;A、B和C分别为后、中前屏幕目标;A'、B'和C'分别为缺失状况下的后、中、前屏幕目标; A_U 、 B_U 和 C_U 分别指后、中、前目标的上部分; A_L 、 B_L 和 C_L 分别指后、中、前目标的下部分)

Fig. 4 (Color online) Schematic diagram of missing target position detection. (a) Target not missing, (b) missing both front and rear targets, (c) missing multiple consecutive targets ahead, and (d) rear continuous multiple target missing. ①, ②, and ③ are the absolute positions of the front, middle, and rear targets, respectively. The solid rectangular box represents the detected target box. The dashed rectangular box represents the missing target. A, B and C are the rear, middle, and front screen targets, respectively. A', B' and C', respectively, refer to the rear, middle, and front screen targets in the absence situation. A_U , B_U , and C_U , respectively, refer to the upper part of the rear, middle, and front targets. A_L , B_L , and C_L , respectively, refer to the lower part of the rear, middle, and front targets.

失对目标大小没有影响, 故目标C'的面积与目标C的面积大小相等, 其绝对位置为①.

2.5 屏幕变化检测

在进行屏幕变化测试时, 手机屏幕会动态展示纯色背景、图案及文字等内容, 如果间隔预设时间后, 屏幕显示内容仍未发生变化, 则判定屏幕损坏.

本研究通过对比方向梯度直方图方法、像素分布统计法和图像差分法的判别效果, 最终选用图像差分方法来进行手机屏幕变化检测. 梯度方向直方图方法通过计算局部图像梯度方向信息的统计值来生成目标的描述特征, 但它对目标区域的纯色变化不敏感. 像素分布统计法不需要计算梯度, 而是将小区域内的像素分布转为特征表示. 但是, 由于屏

幕遮挡导致离检测相机越远的屏幕能被检测到的显示区域越小, 以及离相机越远的屏幕成像尺寸越小, 这两种因素使得基于像素统计的方法难以准确获取远端屏幕的统计特征. 帧间差分法是一种通过对视频的图像序列中两帧作差分运算后获得两帧差异的方法. 该方法检测的是不同帧之间的相同位置是否存在差异, 对不同大小的目标及纯色区域均适用, 因此对手机屏幕变化的检测效果比另两种方法更好.

本研究设计的帧间差分法是一种基于目标区域帧差分法的双阈值图像差分判定方法. 设函数 P 的值为像素变化值超过设定阈值的概率, 即

$$P(|M_1 - M_2| > \theta_1) > \theta_2 \quad (6)$$

其中, M_1 和 M_2 分别为 t_1 和 t_2 时刻图像目标区域内对应像素点在 R(红)、G(绿)和 B(蓝)颜色通道上的像素值; θ_1 为两幅图像对应像素变化的阈值; θ_2 为两幅图像之间所有对应像素点差值大于 θ_1 所占比例的阈值. 由式(6)可见, 该帧间差分法使用双重阈值进行判断, 当两帧中对应目标区域的像素之间的变化量及发生变化像素的比例均超过设定阈值时, 则判定该目标区域内的手机屏幕内容发生了变化. 使用双阈值判定法可较好地解决生产现场存在的因不同帧间的目标框不完全重合, 以及屏幕包含外包装等问题导致的不同帧间屏幕内容不完全变化现象. 相比传统的单阈值帧间差分法, 双阈值差分法在像素点突变等含噪声场景、行人走动等原因产生阴影的场景, 以及光照变化等场景中均取得了更好的效果, 鲁棒性和适用性更佳.

判定手机屏幕显示内容是否变化的屏幕变化检测具体过程为: ① 检测间隔预设时间前后两帧图像中手机屏幕的位置及缺失情况, 并完成对应手机屏幕的匹配. ② 采用基于目标区域的双阈值帧间差分法判断屏幕内容是否变化, 当目标区域的像素变化量及发生变化像素的比例均超过设定阈值时, 则判定该屏幕内容发生变化. ③ 执行图像内手机屏幕区域遍历操作, 完成两帧中所有对应手机屏幕内容是否发生变化的判断. 当屏幕显示内容发生变化时, 表明屏幕处于正常工作状态. 当手机屏幕显示内容未发生变化时, 表明屏幕处于疑似异常状态, 记录此刻手机屏幕的位置及时间, 继续执行间隔时间图像采样及手机屏幕变化检测, 并重复步骤①到③. 当预设时间内手机屏幕内容发生变化时, 则将该位置处手机屏幕的疑似异常状态转为正常状

态; 反之, 当超过预设时间手机屏幕内容仍未变化时, 则判定该位置处手机屏幕为异常状态.

3 实验

3.1 数据收集与处理

为保证所设计的屏幕缺陷视觉检测系统的训练数据与真实的生产数据分布一致, 在收集数据时, 采集了 1 112 幅来自不同摄像头、场地、时间及光照条件下的图像. 所有被测手机的型号、手机屏幕外包装、电源线及连接形式、安装摆放方式、安装固定台均与应用现场一致. 在采集图像时, 通过调整相机位置及角度来模拟不同相机拍摄角度, 通过摆放不同位置及数量的手机屏幕来模拟遮挡及缺失等场景.

采用 LabelMe 软件对采集到的图像进行标注. 由于离相机越远的手机被检测到的屏幕尺寸越小, 且被遮挡的面积越多, 因此, 在标注时只标注图像中未被遮挡的屏幕部分. 通过预先设置检出框面积及高宽比阈值, 过滤掉远处较小的目标框, 从而降低误检率.

在模型训练前, 随机采用灰度化、颜色变换、旋转、翻转、缩放、平移、剪切和噪声添加中的 4 种方式对图像数据进行增强, 将单幅图像扩充到 5 幅. 然后, 将模型在数据集上进行迁移学习, 即先在 COCO 数据集^[28]上对模型进行预训练, 然后在此前现场采集及标注的自有数据集上对模型进行微调及二次训练, 以利用已有的数据知识辅助模型尽快地学习新领域知识.

3.2 模型训练与测试

在 COCO 及自有数据集上训练 YOLOv8n 模型. 其中, 训练批大小设为 2, 较小的批可提高模型的泛化性能; 采用自适应学习率进行训练, 初始学习率设为 0.01; 动量大小设为 0.937, 权重衰减设为 5×10^{-4} . 在 RTX 3090Ti 平台上进行模型训练与推理测试, 训练迭代次数为 100 次.

分别使用位置损失、类别损失及分布焦点损失(distribution focal loss, DFL)来衡量每次迭代的前向计算结果与真实值的差距, 从而指导下一步的训练向正确的方向进行. 其中, DFL 函数将检测框的位置建模为概率分布, 使网络能够从概率最大的角度快速定位到标注位置附近, 从而使模型获得更高的精度与更快的训练速度. 模型训练及测试集上各项

性能指标变化结果如图5。由图5(a)可见, 各项损失均处于下降趋势, 训练结果收敛。从图5(b)中可见, 在测试集上, 模型预测准确率为99.9%, 召回率为100%, 平均准确率(mean average precision, mAP)为96.8%。较高的准确率主要是因为该任务为单分类目标检测任务, 且场景相对单一。

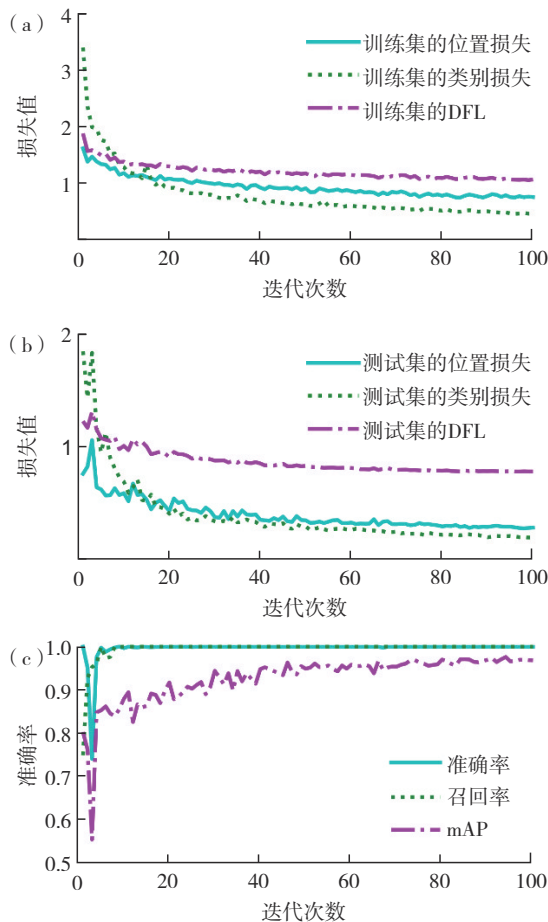


图5 训练及测试集上各主要指标结果 (a)训练集上位置损失(实线)、类别损失(虚线)及分布焦点损失(点划线); (b)测试集上位置损失(实线)、类别损失(实线)及分布焦点损失(点划线); (c)测试集上准确率(实线)、召回率(虚线)及平均准确率(点划线)结果

Fig. 5 (Color online) Main indicator results on the training and testing set. (a) Position loss (solid line), category loss (dotted line), and distribution focal loss (dash dotted line) on the training set. (b) Position loss (solid line), category loss (dotted line), and distribution focal loss (dash dotted line) on the test set; (c) precision (solid line), recall (dotted line), and mean average precision (dash dotted line) results on the test set.

3.3 屏幕变化检测与模型部署

模型训练结束后被部署运行在 APPLE Silicon M1 MAX 硬件平台上。使用海康威视 DS-2CD2D25

DWD(C)型号相机进行连续图像采集, 采集频率为 60 帧/s, 图像分辨率为 $1\,920 \times 1\,080$ 像素。使用训练好的模型对图像中的目标进行检测, 得到针对遮挡和缺失场景下的检测效果如图 6(a)、6(b)及 6(c)所示。

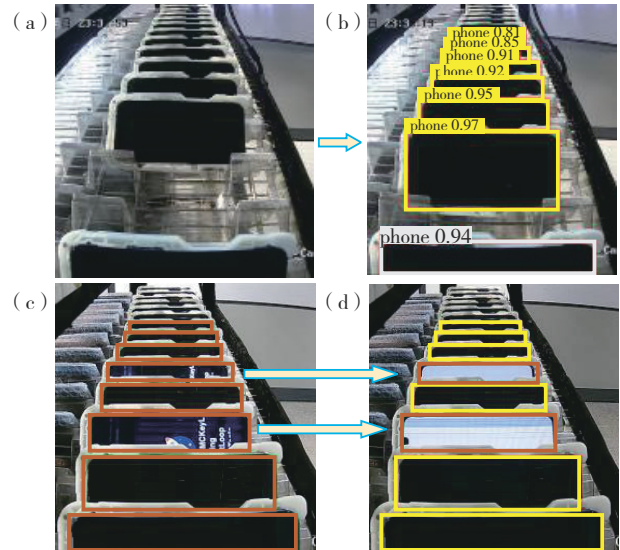


图6 屏幕缺陷视觉检测系统的检测效果图 (a)相机采集的原始图像; (b)检测到的屏幕位置及置信度; (c)屏幕内容初始检测结果; (d)屏幕内容是否变化检测结果

Fig. 6 (Color online) Application rendering of object detection and screen change detection. (a) Original images captured by the camera, (b) the screen position and confidence detected by the model, (c) initial detection results of screen content, and (d) detection results of screen content changes.

在实际应用时, 通过合理设置置信度阈值及对连续帧图像进行检测, 可进一步降低漏检及误检概率。在对远处目标进行检测时, 虽然系统易将背景或不属于该相机检测范围的屏幕检测为正常目标, 但这些检测出的目标通常置信度不高, 因此常被过滤掉, 从而降低了误检概率。本研究设置置信度阈值为 0.25。漏检则主要是因为检测系统未检测出图像远端的手机屏幕。为增加远处屏幕检出的概率, 本研究将对连续 3 帧的图像中的目标进行检测。当 3 帧图像的检测结果不同时, 再取 3 帧图像重新进行检测, 直到连续 3 帧图像检测结果相同或差异小于设定阈值时, 将此 3 帧图像的平均结果作为最终的检测结果输出。

完成目标检测后, 使用目标匹配算法对不同帧间的目标进行匹配。本研究设 IOA 阈值为 0.5。然后每间隔 60 帧抽取 1 帧图像, 并对抽取的图像依次采用屏幕变化检测算法以判断屏幕是否发生变化。

在式(6)中, 设置 $\theta_1 = 25$, $\theta_2 = 0.2$. 对检测图像中的每个目标分别设置计时器, 当目标框内屏幕变化时, 计时器重新计时. 当计时时间超过 10 min, 则判定手机出现异常, 并启动报警程序, 改由人工检测. 对应屏幕检测及变化判断结果如图 6(c)及 6(d)所示.

在实际应用的 CPU 硬件平台上, 该检测系统能够以 20 帧/s 的速度处理图片, 满足工业生产的实时性要求. 为过滤背景等无效信息对检测结果的影响, 在图像输入模型前, 基于图像特征的检测方式, 系统会先识别出手机屏幕的聚集区域, 然后对四周的背景区域进行裁剪, 只保留视野中包含手机的核心区域. 该方法一方面可去除背景干扰且提高系统对目标识别的精度, 另一方面, 由于图像尺寸减小, 可将系统的推理速度提高到 33 帧/s. 尽管系统对实时性要求不高, 但推理速度的提高, 可使得同一硬件服务器能够接入更多的视频输入设备, 实现同步检测更多的手机屏幕. 采用 3.2 节所述硬件平台最多可同时连接及处理 8 路相机输入设备, 即可同步实现对 64 部手机的屏幕进行检测与处理, 检测效率高, 硬件成本低.

本研究与传统目标检测方法相比, 对多帧间图像目标进行匹配, 可有效解决屏幕堆叠摆放引起的遮挡及缺失检测问题. 与传统的图像差分方法相比, 由于采用了大量的数据进行训练, 可有效提高模型检测的精度及对不同场景的泛化能力.

结 语

本研究结合深度学习目标检测、目标匹配, 以及图像差分技术, 搭建了一种能在被检测的手机屏幕存在被遮挡与缺失的情景下, 检测手机屏幕是否存在缺陷的视觉监测系统. 该系统使用 YOLOv8n 模型检测到视野中的手机位置, 使用多台相机及检测框过滤等方式对遮挡的手机屏幕进行检测, 使用目标匹配算法判断手机屏幕是否存在缺失, 通过屏幕变化检测算法判断手机屏幕是否发生变化. 基于生产现场数据, 对目标检测模型进行了训练与测试, 所得手机屏幕平均检测准确率为 96.8%. 通过对应目标匹配及差分计算, 实现了手机屏幕缺陷的准确检测. 与传统的检测方式相比, 本研究提出的目标检测、匹配与图像差分融合方法, 在检测准确率、检测效率及场景适应性等方面均具有优势. 后

续将继续在数据类别平衡、困难样本挖掘、少样本检测等方面对算法模型进行优化.

基金项目: 国家自然科学基金资助项目(62271324); 深圳市科协“科创中国”资助项目(2023KCSZ01); 深圳市国电科技通信有限公司资助项目(23220008)

作者简介: 尹东富(ydf678@hotmail.com), 人工智能与数字经济广东省实验室(深圳)副研究员, 博士. 研究方向: 计算机视觉, 2D 与 3D 目标检测, 图像文本内容生成.

引 文: 尹东富, 杜明臣, 胡天昊, 等. 遮挡与缺失场景下屏幕缺陷视觉检测[J]. 深圳大学学报理工版, 2023, 40(6): 631-639.

参考文献 / References:

- [1] LIU Yu. An improved faster R-CNN for object detection [C]// The 11th International Symposium on Computational Intelligence and Design (ISCID). Piscataway, USA: IEEE, 2018: 119-123.
- [2] TERVEN J, CORDOVA-ESPARZA D. A comprehensive review of YOLO: from YOLOv1 and beyond [EB/OL]. (2023-08-07). <https://arxiv.org/abs/2304.00501>.
- [3] DUAN Kaiwen, BAI Song, XIE Lingxi, et al. CenterNet: keypoint triplets for object detection [C]// IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, USA: IEEE, 2019: 6568-6577.
- [4] LIU Ze, HU Han, LIN Yutong, et al. Swin transformer V2: scaling up capacity and resolution [C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2022: 11999-12009.
- [5] GHIASI G, LIN T Y, LE Q V. NAS-FPN: learning scalable feature pyramid architecture for object detection [C]// IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2019: 7029-7038.
- [6] HAJIZADEH M, SABOKROU M, RAHMANI A. Mobile-DenseNet: a new approach to object detection on mobile devices [J]. Expert Systems with Applications, 2023, 215: 119348.
- [7] CHEN I C, HWANG R C, HUANG H C. PCB defect detection based on deep learning algorithm [J]. Processes, 2023, 11(3): 775.
- [8] YANG Jing, XU Yi, RONG Haijun, et al. A method for wafer defect detection using spatial feature points guided affine iterative closest point algorithm [J]. IEEE Access, 2020, 8: 79056-79068.
- [9] HAYASHI N, KOYANAKA S, OKI T. Verification of algorithm for automatic detection of electronic devices mounted on waste printed circuit boards [J]. Journal of the

- Air & Waste Management Association, 2022, 72(5): 420-433.
- [10] KORODI A, ANITEI D, BOITOR A, et al. Image-processing-based low-cost fault detection solution for end-of-line ECUs in automotive manufacturing [J]. *Sensors*, 2020, 20(12): 3520.
- [11] JIANG Jiabin, CAO Pin, LU Zichen, et al. Surface defect detection for mobile phone back glass based on symmetric convolutional neural network deep learning [J]. *Applied Sciences*, 2020, 10(10): 3621.
- [12] LI Changsheng, ZHANG Xianmin, HUANG Yanjiang, et al. A novel algorithm for defect extraction and classification of mobile phone screen based on machine vision [J]. *Computers & Industrial Engineering*, 2020, 146: 106530.
- [13] YANG Ge, LAI Haijian, ZHOU Qifeng. Visual defects detection model of mobile phone screen [J]. *Journal of Intelligent & Fuzzy Systems*, 2022, 43(4): 4335-4349.
- [14] PEI Baoying, PENG Yu, LUO Yong. A method of detecting defects of smart meter LCD screen based on LSD and deep learning [J]. *Multimedia Tools and Applications*, 2021, 80(28): 35955-35972.
- [15] ÇELİK A, KÜÇÜKMANİSA A, SÜMER A, et al. A real-time defective pixel detection system for LCDs using deep learning based object detectors [J]. *Journal of Intelligent Manufacturing*, 2022, 33(4): 985-994.
- [16] ZHOU Chunlun, YUAN Junsong. Occlusion pattern discovery for object detection and occlusion reasoning [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2020, 30(7): 2067-2080.
- [17] KIM J U, KWON J, KIM H G, et al. BBC net: bounding-box critic network for occlusion-robust object detection [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2020, 30(4): 1037-1050.
- [18] QIU Benliu, QIU Heqian, WEN Haitao, et al. Cross-domain object detection with missing classes in target domain [C]//2022 IEEE 24th International Workshop on Multimedia Signal Processing (MMSP). Piscataway, USA: IEEE, 2022: 1-6.
- [19] ZHANG Yongqiang, DING Mingli, BAI Yancheng, et al. Beyond weakly supervised: pseudo ground truths mining for missing bounding-boxes object detection [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2020, 30(4): 983-997.
- [20] AKMAN A, CEKLI S. Efficient 2D DCT architecture based on approximate compressors for image compression with HEVC intra-prediction [J]. *Journal of Real-Time Image Processing*, 2023, 20(2): 22.
- [21] ZOU Tao, SITU Weilun, YANG Wenlin, et al. A method for long-term target anti-interference tracking combining deep learning and CKF for LARS tracking and capturing [J]. *Remote Sensing*, 2023, 15(3): 748.
- [22] GORYAWALA M, ROY B, GUPTA R K, et al. T1-weighted and T2-weighted subtraction MR images for glioma visualization and grading [J]. *Journal of Neuroimaging*, 2021, 31(1): 124-131.
- [23] ULGER F, YUKSEL S E. A standalone open-source system for optical inspection of printed circuit boards [C]//2019 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA). Piscataway, USA: IEEE, 2019: 105-110.
- [24] YU Wentao, BAI Jing, JIAO Licheng. Background subtraction based on GAN and domain adaptation for VHR optical remote sensing videos [J]. *IEEE Access*, 2020, 8: 119144-119157.
- [25] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: common objects in context [C]//Computer Vision - ECCV 2014. Cham: Springer International Publishing, 2014: 740-755.

【中文责编: 英子; 英文责编: 木柯】