

基于 YOLOv5 和生成对抗网络的 塑料标签缺陷检测

庄昌乾¹, 李璟文^{1,2}

(1. 江南大学 理学院, 江苏 无锡 214122; 2. 江西盛泰光学有限公司, 江西 新余 336600)

摘要: 塑料标签物的缺陷检测与识别是工业过程控制和质量控制的关键; 为了克服现有塑料标签缺陷检测方法的局限性, 使用了单阶段目标检测模型 YOLOv5 对其瑕疵进行实时检测与分类; 此外, 为解决由于样本缺陷数量不足造成的模型识别准确率低等问题, 采用了一种基于 Defect-GAN 的生成对抗网络对小样本进行数据增强和扩增; 该方法通过模拟缺陷生成和缺陷图像重建的过程, 可以高效合成大量具有高保真度和多样性的缺陷样本, 尤其适用于形状不规则、分布随机且尺寸不同的瑕疵生成; 实验结果表明, 通过使用扩增数据集训练目标检测器, 并对网络的超参数进行优化, 可以显著提高目标检测器的准确率和精度, 其平均精度 mAP 可达 99.5%; 此外, 为了模拟该方法在实际生产中的应用场景, 设计并定制了一台半自动的图像采集机械平台用于采集圆柱样品表面的印刷标签, 以及一个自主开发的图像处理和统计分析软件用于样本采集、图像处理及统计分析; 该方法和平台可以很容易地推广并应用到其他工业质量控制和缺陷检测系统中。

关键词: 缺陷检测; YOLOv5; Defect-GAN; 数据增强; 图像处理

Industrial Defect Detection of Plastic Labels Based on YOLOv5 and Generative Adversarial Networks

ZHUANG Changqian¹, LI Jingwen^{1,2}

(1. School of Science, Jiangnan University, Wuxi 214122, China;

2. Jiangxi Shengtai Optics Co., Ltd., Xinyu 336600, China)

Abstract: The defect detection and identification of plastic labels is the key of industrial process control and quality control. In order to overcome the limitations of exiting plastic label defect detection methods, the single-stage target detection model YOLOv5 is used to detect and classify the defects in real time. In addition, in order to solve the problem of low model recognition accuracy due to insufficient sample defects, a generative adversarial networks based on Defect-GAN is used to enhance the data and amplify the small samples. By simulating the processes of defect generation and defect image reconstruction, the method can efficiently synthesize a large number of defect samples with high fidelity and diversity. It is especially suitable for the defects of irregular shape, random distribution and different size. The experimental results show that the accuracy of the detector can be significantly improved by using the amplified data set to train the detector and optimizing the network hyperparameters, and the mean average precision (mAP) of the detector can reach 99.5%. In addition, in order to simulate the application scenario of this method in actual scenario, a semi-automatic image acquisition machine platform is designed and customized for collecting the printed labels on the surface of cylindrical samples, and a self-developed image processing and statistical analysis software is used for the sample collection, image processing and statistical analysis. The method and platform can be easily extended and applied to other industrial quality control and defect detection systems.

Keywords: defect detection; YOLOv5; Defect-GAN; data augmentation; image processing

0 引言

自动化缺陷检测对于先进制造中产品的质量控制在至关重要, 现已被广泛应用于检测各种产品生产过程中的异常和缺陷^[1-2]。印刷塑料标签作为一种独特的标识和质量指标, 展现了商品的信息、制造商的设计理念, 也代表了产品的质量。但由于生产设备参差不齐、生产工艺标准不同、操作人员误

操作等外部因素, 在实际制造过程中, 塑料标签表面可能会出现各种类型的缺陷。一些最常见的缺陷包括异物、重影、划痕、墨点、指纹、线条缺失。这些缺陷不可避免地会影响产品的质量和外观, 因此在制造过程中检测和识别这些缺陷, 从而确保产品的整体质量是非常必要的。

目前, 最常见的检测方法仍依赖人工筛选^[3]。然而, 这种方法不仅劳动强度大, 而且容易受到主观因素、疲劳

收稿日期: 2023-03-01; 修回日期: 2023-03-09。

基金项目: 国家自然科学基金青年基金(11904135); 中国博士后科学基金第 70 批面上资助一等(2021M700039)。

作者简介: 庄昌乾(1996-), 男, 山东枣庄人, 硕士, 主要从事机器视觉方向的研究。

通讯作者: 李璟文(1985-), 男, 甘肃人, 博士, 副教授, 主要从事自动化光学检测与人工智能方向的研究。

引用格式: 庄昌乾, 李璟文. 基于 YOLOv5 和生成对抗网络的塑料标签缺陷检测[J]. 计算机测量与控制, 2023, 31(7): 91-98.

程度和检验员经验的影响,导致误检漏检现象频繁发生,整体准确率和效率较低。此外,另一个明显的缺点是,这种技术从根本上难以与自动化制造过程无缝衔接与兼容。

机器视觉是一种更好的缺陷检测方法,该方法依靠高速相机采集样本图像,通过算法进行图像处理,从而识别和分类各种缺陷^[4-6]。目前,大多数研究都局限于基于经典图像处理方法(如振幅阈值、图像去噪与增强、图像分割和特征提取)的决策算法^[6]。另一类缺陷检测依赖于基于频谱的算法,它将图像信息转换到频域(例如 Gabor 变换,傅里叶变换等)^[7]。该方法在发现纹理特征背景下的瑕疵具有一定的优势,已应用于织物瑕疵检测。然而,由于塑料标签缺陷的复杂性,单幅图像中可能存在多个、多种缺陷,使用传统的图像处理技术对多尺度缺陷进行定位和识别具有很大的挑战性。

随着计算机视觉的快速发展,深度学习模型开始被应用于工业产品检测。到目前为止,已经有许多基于不同网络模型和算法的目标检测器用于缺陷检测^[8-10]。这些检测器主要分为单阶段和双阶段检测器。其中,基于双阶段检测算法的典型模型有:Region-CNN (R-CNN)^[8]、Fast Region-based CNN (Fast R-CNN)^[9]、Faster Region-based CNN (Faster R-CNN)^[10]。在这些方法中,涉及两个阶段(即 generating candidate regions 和 classifying proposal regions),这通常需要较长的处理时间。另一方面,单阶段检测器只需要通过一次神经网络,并在短时间内预测所有边界框及其类别概率,因此更适合有实时需求的场景。最常见的单阶段检测器有 YOLO^[11-15]、SSD^[16]、和 Corner-Net^[17]。其中,YOLO 是一种非常先进的目标检测器,因为它能够实时检测物体,并具有良好的精度和鲁棒性,已在许多工业场景中用于缺陷检测。例如,在文献[18]中,作者展示了一个多注意力深度学习网络,以解决纺织图像中多个和多尺度缺陷共存的问题。在文献[19]中,作者开发了一个基于 YOLO-v3 的端到端缺陷检测模型。在文献[20]中,作者展示了一种改进的 MS-YOLOv5 网络,该网络采用多流网络,旨在提高对铝表面缺陷的识别,对 7 种类型缺陷的总体准确率达到 87.4%,且检测性能没有明显的不平衡。尽管研究人员在表面缺陷检测方面进行了大量的研究,但很少有人尝试解决工业应用中对塑料标签瑕疵缺陷的高精度、高效率检测需求的日益增长。

此外,基于深度学习的缺陷检测技术的另一个难点在于,它们往往需要大量带注释的训练样本,而这些样本本身数量较少,或收集起来很耗时。在某些情况下,特别是在训练深度神经网络模型时,缺陷样本较少,可用的数量远远不能满足需求。缺陷样本的不足已成为各种自动化缺陷检测系统设计和部署的瓶颈。最近的一些研究尝试采用 few-short learning,即从极其有限的训练样本中学习来解决这一问题^[21-22]。然而,它们的性能相当有限,离实际应用还很遥远。解决缺陷样本不足的一种最直接的方法是使用数据增强方法来丰富数据集的数量和多样性,从而训练出

更好的缺陷检测网络模型。

因此,在本文的工作中,为了克服现有塑料标签缺陷检测方法的局限性,并解决可用缺陷样本不足的问题,我们使用单阶段检测器(即 YOLOv5)进行缺陷检测,并使用生成对抗网络(GAN, generative adversarial network)进行数据增强和缺陷样本合成。由于传统的 GAN 在合成缺陷图像时存在局限性,因此,我们采用了一种专门用于生成形状不规则、分布随机且尺寸不同缺陷样本的 Defect-Gan,通过模拟缺陷生成(defacement)和缺陷图像重建(restoration)的过程,我们可以巧妙地利用大量的正常样本来生成具有高保真度和多样性的缺陷样本。通过使用扩增数据集训练目标检测器,并对网络的超参数进行优化,可以显著提高目标检测器的准确率和精度。此外,为了模拟该方法在实际生产中的应用场景,我们设计并定制了一台半自动的图像采集机械平台用于采集圆柱样品表面的印刷标签,以及一个自主开发的图像处理和统计分析软件。本文所开发的方法和平台可以很容易地推广并应用到其他工业质量控制和缺陷检测系统中。

1 方法

1.1 基于 YOLOv5 的目标检测

自 2016 年首次亮相以来,YOLO 已经有了很大的发展和升级。即使最新的版本是 YOLOv7(仍在更新中),YOLOv5 仍然被认为是迄今为止最受欢迎的实时目标检测算法之一,由于其在卷积神经网络领域的最佳优化策略,它在精度、效率和识别模型等方面都优于其他目标检测算法。YOLOv5 的网络结构图可分为 Input、Backbone、Neck、Head 四部分,如图 1 所示。

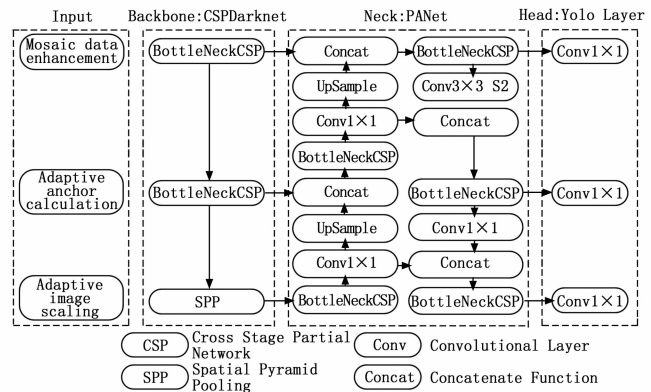


图 1 YOLOv5 架构

具体来说,YOLOv5 的 Input 主要包括三种策略:马赛克数据增强(mosaic data enhancement)、自适应锚点计算(adaptive anchor calculation)和自适应图像缩放(adaptive image scaling)。首先,采用马赛克数据增强算法^[23],通过翻转、裁剪、缩放等操作,对 4 张缺陷图片进行随机拼接,提高模型训练效率,增强网络的鲁棒性。为了适应不同的数据集,YOLOv5 在输入上集成了自适应锚框计算。因此,当使用不同的数据集进行训练时,初始锚盒的计算可以自

动执行和更新。

YOLOv5 的 Backbone 主要使用 Focus 结构和 CSP (cross-stage particle network) 结构, 通过多次卷积和池化操作从输入图像中提取不同维度的特征图。采用 Bottleneck CSP, 将基础层的特征图分成两部分, 再通过 CSP 进行合并, 减少了计算量。最后, SPP (spatial pyramid pooling) 结构从不同尺度提取特征, 生成三尺度特征图, 以提高检测精度^[24-26]。

在 Neck 中, YOLOv5 使用 FPN (feature pyramid network) 和 PAN (pixel aggregation network) 来融合特征层, 并在三个不同尺度的特征层上实现预测。其中 FPN^[27] 是一个自顶向下的结构, 它将高层特征通过上采样和底层特征做融合得到进行预测的特征图。PAN^[28] 是一个自底向上的特征融合层, 对 FPN 进行补充, 并将底层的定位特征传递上去。这两种方法都有助于增强特征融合能力, 增强定位的可移植性, 提高检测精度。

作为最后的检测步骤, Head 用于检测物体的位置和类别。在 Precision 中, 使用 GIoU 作为边界框的损失函数, 表示为:

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (1)$$

$$GIoU = IoU - \frac{|C/A \cap B|}{|C|} \quad (2)$$

式中, A 、 B 为 GT 框 (the ground truth box) 和预测框, C 为 A 、 B 的最小闭合框。 IoU 由预测框与 GT 框的交互作用定义, 它代表了预测的边界框和 GT 在目标区域的大小。 $GIoU$ 是减去 C 框中与 A 和 B 不重叠的面积之比得到的值。因此, $GIoU$ 越大, 性能就越好。

1.2 合成数据的生成

足量的注释数据集在基于深度学习的目标检测和目标分类中至关重要。缺乏缺陷样本会导致训练过程过拟合、检测精度差、泛化能力差等问题。然而, 在某些情况下, 很难获得用于训练模型的足够数据集。为了解决这一问题, 可以采用数据增强算法来提高网络的泛化能力。在本节中, 我们展示了两种合成图像生成方法 (即生成缺陷图像的传统方法和基于深度学习的 Defect-GAN 方法)。

生成缺陷图像最直接的方法是使用传统技术, 包括几何变换 (例如, 旋转、翻转、平移)、颜色空间变换 (例如, 随机改变图像的色调、饱和度、亮度和对比度)、Kernel 滤波器、随机擦除、混合图像和复制/粘贴^[29]。

本工作中使用的第二种方法是使用 GAN 自动合成缺陷样品^[30-32]。其基本思想是将两个神经网络结合起来, 同时训练一个生成器来生成逼真的假图像, 以及一个鉴别器来区分生成的图像和真实的图像。

生成器 (G, generator) 将一个固定长度的随机向量作为输入, 并学习模拟生成原始数据集包含的缺陷样本。然后将生成的样本和原始数据集的样本输入鉴别器 (D, discriminator), 鉴别器学习区分合成图像和真实图像。生成器的目标是创建与原始图像尽可能相似的样本图像, 从而欺

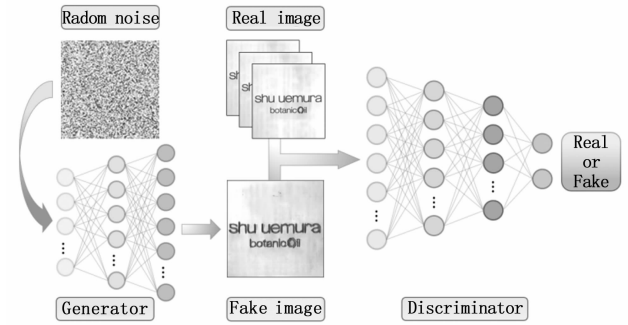


图 2 GAN 架构

骗鉴别器, 使其相信合成样本是真实的。因此, 生成器和鉴别器可以被认为是在进行极大极小值的博弈, 其中生成器的目标是最大化鉴别器将生成的图像标记为真实样本的概率, 鉴别器的目标是最小化将真实数据标记为真实样本的概率。

迄今为止, 研究者已经提出了许多不同的 GAN 架构^[33-40], 用于面部表情编辑、图像修补、风格翻译、超分辨率等。其中包括用于生成人脸图像的 StyleGan2^[33], 用于图像到图像转换任务的 pix2pixGan^[34], 将黑白照片转换为彩色, 以及用于扩大生物医学图像数据集的 CycleGAN^[35]。然而, 这些方法在用于生成高保真缺陷样本时存在局限性。在最近的另一篇文章中, 作者提出了 Defect-GAN^[36], 它是专门为生成缺陷样本而设计的。在这个体系结构中, 通过模拟缺陷生成 (即在正常样本上生成缺陷) 和缺陷图像重建 (即将缺陷样本恢复到正常样本) 的过程, 我们可以巧妙地利用大量的正常样本来生成缺陷样本。Defect-Gan 的结构如图 3 所示。

生成器采用了一个编码器-解码器帧, 它首先对输入图像进行编码, 步幅为 4, 然后将其解码为原始大小。此外, 由于塑料标签上的一些缺陷通常具有复杂且不规则的形状, 且具有随机变化, 因此使用传统的 GAN 建模具有困难。为了缓解这种情况, 我们采用了一种自适应噪声引入机制, 该机制结合了随机性来模拟缺陷内的随机变化。有了这样的配置, 我们可以生成更真实的缺陷样本, 提高保真度和多样性。判别器的主要架构沿用了 StarGan^[37], 它使用 PatchGan^[34] 和一个分类指标来预测生成的缺陷类别来区分假样本和真实样本。

在数学上, D 和 G 使用极小极大博弈来生成与现实图像难以区分的图像, 我们采用的对抗性损失为^[37]:

$$\mathcal{L}_{adv} = E_x [\log D_{src}(x)] + E_{x,c} [\log(1 - D_{src}(d(x, c)))] \quad (3)$$

其中: G 生成一个图像 $d(x, c)$, 条件是输入图像 x 和目标缺陷标签 c , 而 x 试图区分真实和虚假的图像。 $D_{src}(x)$ 为 D 给出的源的概率分布。

对于给定的输入图像 x , 目标是生成属于目标类别的缺陷图像。为此, 在 D 的基础上增加一个辅助分类器, 在 D 和 G 的优化迭代过程中都引入了分类损失, 因此目标可以分解为两个独立的项, 即 \mathcal{L}_{ds}^v 和 \mathcal{L}_{ds}^f , 我们定义为:

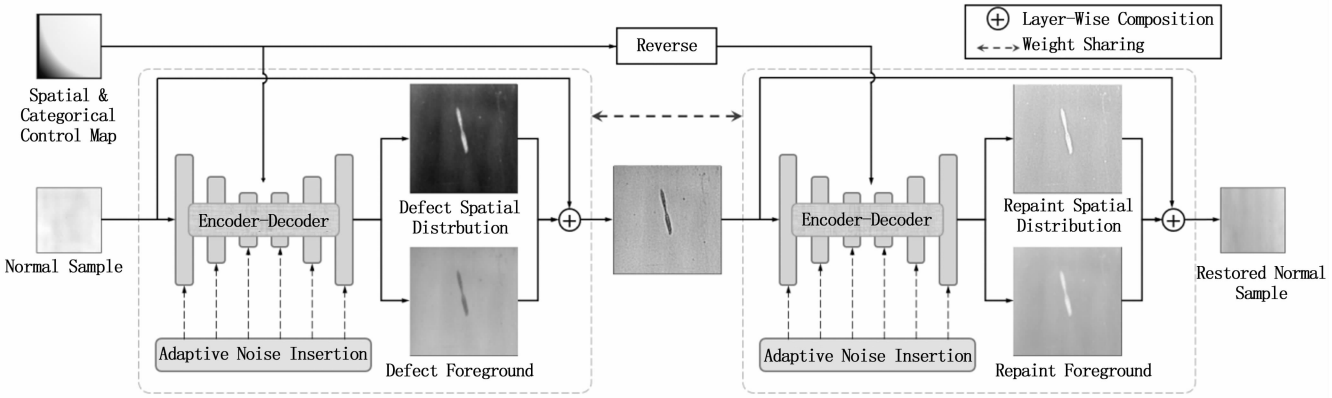


图 3 Defect-Gan 架构。采用编码器-解码器结构，通过模拟破坏和恢复过程来合成缺陷

$$\mathcal{L}_{ds}^e = E_{x,c'}[-\log D_{ds}(c' | x)] \quad (4)$$

$$\mathcal{L}_{ds}^g = E_{x,c}[-\log D_{ds}(c | G(x, c))] \quad (5)$$

前者是对 D 进行优化，将真实的缺陷样本分类到目标类别，后者是对 G 进行优化，生成目标类别 c 的缺陷样本。因此，通过最小化对抗和分类损失，我们可以在目标类别中生成真实的样本图像。

1.3 模型评价指标

在本节中，我们将简要描述用于评估模型性能的指标。我们使用精度 P （定义为正确分类的缺陷产品占分类器所划分的所有缺陷产品的比例）和召回率 R （定义为正确分类的缺陷产品占缺陷产品数量的比例）来计算模型的平均精度（mAP, mean average precision）：

$$mAP = \frac{1}{N} \sum_1^N AP \quad (6)$$

式中， N 表示 N 个分类， AP 表示平均准确率，计算方法如下：

$$AP = \int_0^1 P(R) dR \quad (7)$$

$$P = \frac{TP}{TP + FP} \quad (8)$$

$$R = \frac{TP}{TP + FN} \quad (9)$$

其中： TP 为真正例，表示本文算法提取的缺陷与人工标注的缺陷相符合。类似的， FP 为假正例，表示本文算法能提取到缺陷，而实际无缺陷； FN 为假反例，表示本文算法未提取到缺陷，而实际有缺陷。此外，利用精准率-召回率曲线（PR 曲线）和 F_1 分数来评估我们的模型性能。其中 PR 曲线是精准率与召回率的关系图，精确率和召回率是一对相互矛盾的指标，一般来说高精准率往往意味着低召回率，反之亦然。为了获得优秀的模型，需要综合 P 和 R ，因此曲线越靠近图的右上角，说明模型的性能越好。但一般来说，算法之间的 PR 曲线彼此交错，很难分辨优劣，所以实际中常用 F_1 分数来衡量。 F_1 定义为：

$$F_1 = 2 \times \frac{P \times R}{P + R} \quad (10)$$

2 实验装置

2.1 实验装置

现实场景中涉及塑料标签种类繁多。在我们的初步研

究中，我们将主要针对印刷在圆柱表面上的半透明塑料标签。它主要有六种不同的缺陷类型，即异物、重影、划痕、墨点、线条缺失、指纹。由于基材的透明性，事实上，其中一些缺陷很难通过简单的人工检查来识别。当检查人员经过几个小时的操作后出现视觉疲劳时，情况会变得更糟。在本节中，我们首先介绍自动采集样本图像的实验设置。我们的装置主要由机械平台、光学照明、图像采集、图像处理与缺陷检测与显示五个模块组成，如图 4（a~e）所示。

为了模拟真实的工业检测场景，该物体被放置在定制的机械台上，可以同时在一维平移和 360 度旋转。具体地说，该物体被放置在可旋转的支架上，如图 4（c）所示。当物体被传送到指定位置时，编码器提供一个基于距离的触发信号。为了用线扫描相机（Linea2, Dalsa, Canada）获得圆柱形物体上标签的清晰 2D 图像，物体的旋转速度和扫描速率必须完美同步。当物体到达指定位置时，旋转电机单元开始旋转圆柱形物体。同时将图像通过 GigE 接口发送给计算机，然后重构成展开的塑料标签二维图像，如图 4（d）所示。物体旋转与相机之间的精确同步是在水平轴和垂直轴上获得可重复分辨率和相同分辨率的关键。为了保障采集的标签图像不失真，同轴度必须得到保证。最后，自主开发的图形用户界面（GUI, graphical user interface）上可对样本进行图像的自动采集、处理和统计分析，如图 4（e）所示。当用户从 GUI 界面发出“采集图像”的请求时，平台开始旋转，并同时打开线扫光源进行照明。在界面上展示了原始图像、检测到的缺陷以及不同缺陷的实时更新和直方图。在我们的初步研究中，我们发现，即便使用半自动化的处理，采集足够的样本图像也是一个耗时的过程，当需要大量缺陷图像时，该方法仍有一定的局限性。事实上，正如我们将在下面的工作中展示的那样，通过使用数据增强，可以方便地合成缺陷图像。此外，需要说明的是，根据样品的特性和检验要求，光学照明可以采用不同的配置。特别是为了突出塑料表面的轻微划痕，可采用低角度照明。此外，通过在光源前安装偏振滤光片，我们可以避免来自反射表面的眩光。在图 5 中，我们展示了通常出现

在塑料标签上的六种主要缺陷类型。



图 4 5 个模块示意图

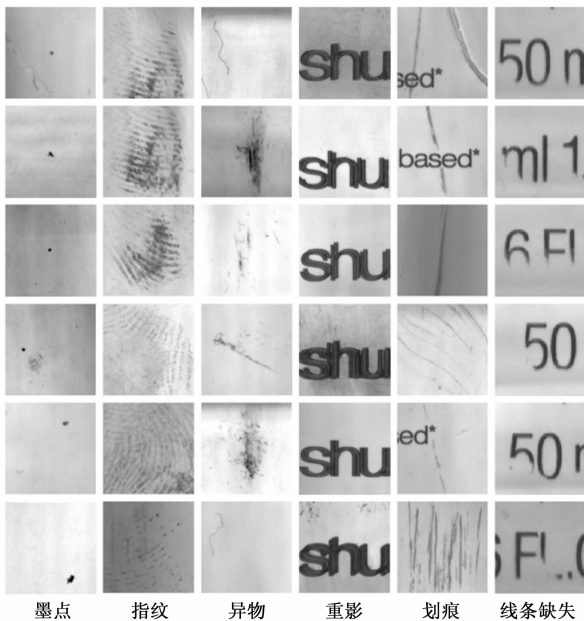


图 5 塑料标签的 6 种主要缺陷类型

2.2 塑料标签数据集

原始数据集是使用图 4 所示的装置获得的 343 张图像 (512×512 像素), 其中共包括 6 种类型的缺陷 (即墨点、指纹、异物、重影、划痕和线条缺失)。考虑到样本不足, 我们采用传统数据增强方法和生成对抗网络 Defect-GAN 进行数据增强, 相应的图像和注释数量如表 1 所示。

表 1 每个数据集中图像和注释的数量 个

| 数据集 | 总数 | 训练集 | 验证集 |
|--------|-------|-------|-----|
| 原数据集 | 343 | 309 | 34 |
| 传统方法扩增 | 988 | 889 | 99 |
| GAN 扩增 | 1 826 | 1 643 | 183 |

2.3 计算机环境

图像处理和计算的计算机环境定义如下: CPU, Intel

(R) Core (TM) i7-11700KF@3.60GHz 8-core CPU; GPU, NVIDIA GeForce RTX3060Ti; SSD, 1TB; programming languages, Python3.9; Framework, Pytorch 1.7.1; deep learning accelerator, CUDA 11.0 and CUDNN 11.1; IDE: Spider 3.9。

3 结果与讨论

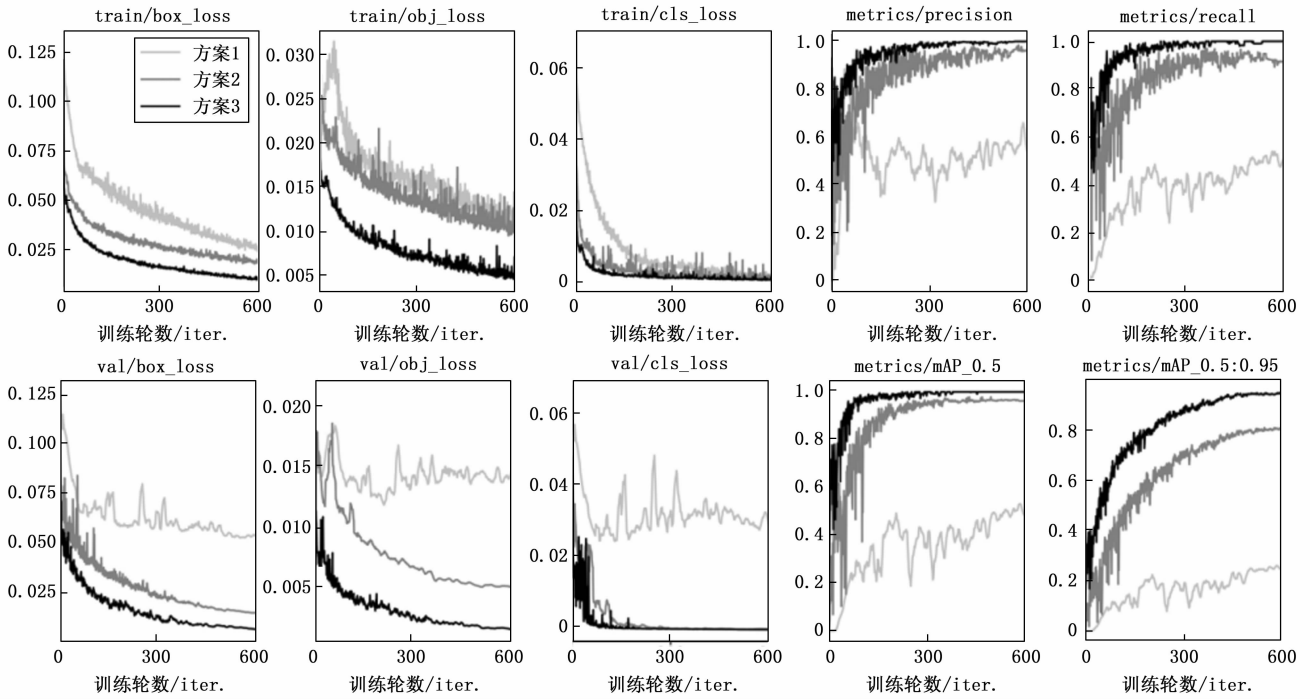
在这项工作中, 训练前的参数设置如下: image size: 512×512; learning rate: 0.001; Optimizer: stochastic gradient descent (SGD); momentum, 0.937; weight decay: 0.005; batch size: 64; epoch: 600。

为了可视化 YOLOv5 在数据增强后的提升, 我们分别给出了原始数据集训练的 YOLOv5 模型, 以及传统增强和 GAN 增强后的 YOLOv5 模型的度量曲线。目标检测算法利用训练损失和验证损失来判断模型是否稳定。YOLOv5 的损失可以进一步分为 box_loss、obj_loss 和 cls_loss, 它们是算法进行目标预测效果的指标。如图 6 所示, 用原始数据集训练的模型, 曲线不仅不稳定, 而且不收敛。相比之下, 当使用增强数据集训练时, 模型在 100 次迭代后变得稳定并且变化最小, 并且在两种情况下, mAP 都有很大的改进。

此外, 我们还比较了不同网络模型的 F1_curve 和 PR_curve。如图 7 (a~c) 所示, 我们绘制了 F1_curve, 这是一个同时考虑了模型的准确性和召回率的度量。曲线越高, 模型的性能越好。我们可以看到, 在所有类别的数据增强后, YOLOv5 的 F1_curve 变得更高。为了评估目标检测器在各种阈值下的性能, 我们还绘制了 P_R 曲线, 如图 7 (d~f) 所示。值得注意的是, 我们可以看到, 使用 GAN 增强数据集训练的 YOLOv5 所包围的区域比使用原始数据集和传统增强数据集训练的 YOLOv5 所包围的区域要大, 这表明模型获得了明显的增强。

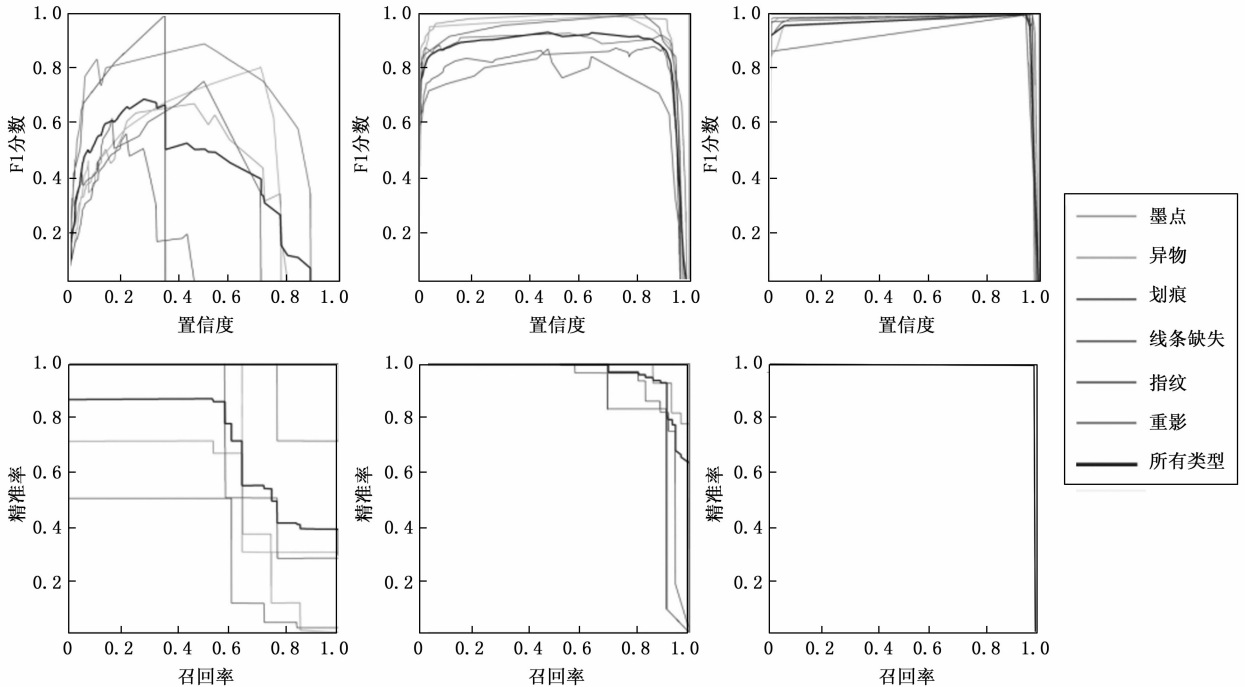
最后, 利用不同的网络模型进行实际检测分类的结果如图 8 所示。图中的标签显示了分类结果及其置信度分数。第一行是用原始数据集 (方案 1) 训练的 YOLOv5 对每个缺陷的检测结果, 第二行和第三行分别是用传统数据增强方法 (方案 2) 和 Defect-GAN 生成对抗网络 (方案 3) 训练的 YOLOv5 对每个标签缺陷的检测结果。从图 8 中可以清楚地看到, 使用方案 1 时未发现的一些缺陷在使用方案 2 和方案 3 时可被发现。此外, 对于已识别的缺陷, 方案 3 的置信度评分普遍高于方案 1 和方案 2, 这表明分类精度有所提高。结果表明, 方案 3 检测性能最好, 几乎所有缺陷类型都能被准确识别和分类。

为了使实验结果更加直观, 我们在表 2 中列出了三个主要性能指标。可以很好地观察到, 与使用原始数据集相比, 使用合成图像训练模型的精度、召回率和 mAP 都得到了提高。当采用生成对抗网络合成的数据进行训练后, 检测器的 mAP 已经超达到 99.5%, 可以满足绝大多数工业应用场景的需求。我们的实验证明, 为了达到预期的性能,



前三列是 box loss, objectness loss, 和 classification loss。右边的两列是训练过程中的 precision 和 recall, 以及验证过程中的 mAP。

图 6 训练和验证过程的输出



第一、第二和第三列分别对应原始数据集、传统增强后和 GAN 增强后的结果。

图 7 YOLOv5 下不同数据集的 F1_curve 和 PR_curve 比较

有足够的的数据是很重要的。此外, Defect-GAN 能够生成具有高保真度和多样性的缺陷图像, 从而提高目标检测器的精度。本工作的实验结果也证实了 Defect-GAN 在缺陷合成的保真度、多样性和可转移性方面的优越性, 尤其适用于工业质量控制和缺陷检测。

表 2 方案 1、2、3 性能比较

| 方法 | Precision/% | Recall/% | mAP/% |
|--------|-------------|----------|-------|
| 原始数据集 | 61.2 | 71.8 | 71.7 |
| 传统方法增强 | 91.2 | 95.8 | 96.0 |
| GAN 增强 | 99.7 | 99.9 | 99.5 |

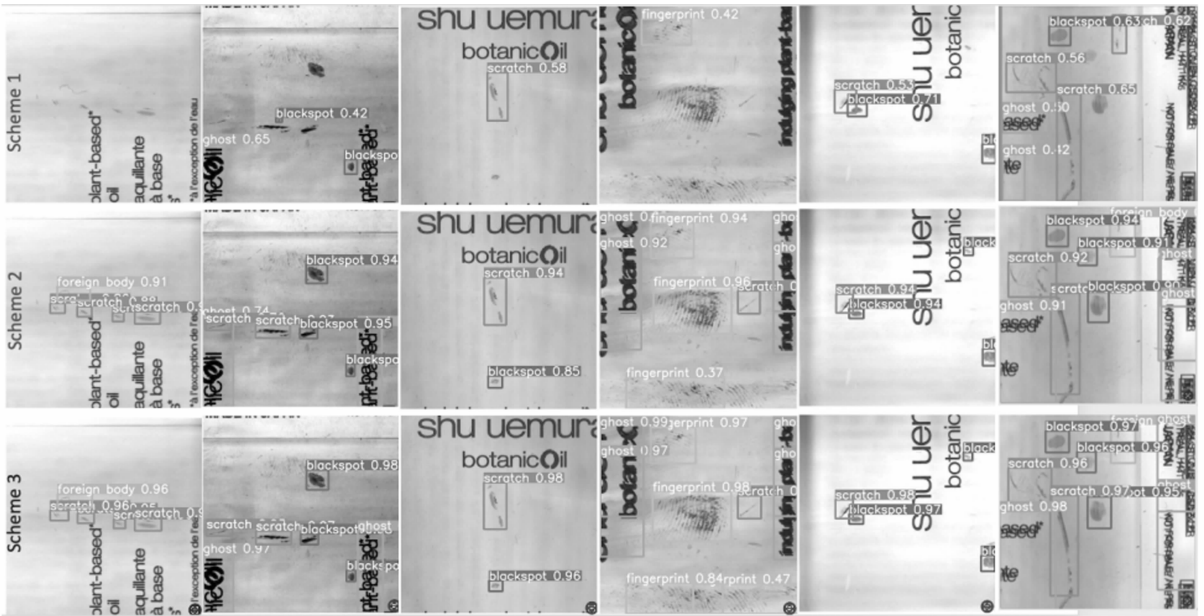


图 8 使用原始数据集训练模型 (方案 1), 经过传统增强 (方案 2) 和 GAN (方案 3) 增强后, 同一数据集的实际检验和分类结果

此外, 为了进一步验证 YOLOv5 在工业缺陷检测方面的优越性和适用性, 我们将其性能与 Faster-RCNN 算法进行了比较。实验结果如图 9 所示。可以看出, 在每个数据集的检测结果中, YOLOv5 在整体性能上优于 Faster RCNN。特别是在原始数据集上训练模型时, mAP 从 48.6% 提高到 71.7%, 达到了 23.1% 的显著提高。同样, 当模型使用合成数据集训练时, 分别有 20.6% 和 16.1% 的改善。

程中样本不足的问题, 从而提高预测和分类精度。我们相信合成数据的生成可以推广到数据集有限的其他场景, 因此特别适用于需要实时缺陷检测和分类的工业应用。

在后续研究中, 我们将继续提高超小目标检测算法的精度和泛化能力, 并进一步开发复杂的集成软件, 用于各种工业场景下的实时缺陷检测和质量控制。

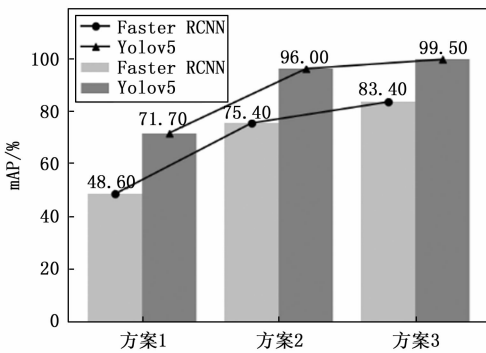


图 9 使用两种不同算法 (YOLOv5 和 Faster RCNN) 时 mAP 对比图

4 结束语

总之, 在本文中, 我们研究并开发了一个用于圆柱表面塑料标签的自动检测和分类平台。我们的主要结论是, 通过数据增强合成缺陷图像可以显著提高目标检测器的准确性和泛化能力。我们选择 YOLOv5 作为目标检测器, 因为它具有良好的效率、精度和鲁棒性。针对实际中缺陷样本不足的问题, 我们探究了两种数据增强方法。通过使用传统的数据增强方法和 Defect-GAN 网络训练模型, 整体精度得到了很大的提高。所展示的数据增强方法可以合成具有良好多样性和保真度的缺陷图像, 能够解决网络训练过

参考文献:

- [1] PEREZ L, RODRIGUEZ I, RODRIGUEZ N, et al. Robot Guidance Using Machine Vision Techniques in Industrial Environments: A Comparative Review [J]. Sensors (Basel), 2016, 16 (3): 1-26.
- [2] REN Z, FANG F, YAN N, et al. State of the Art in Defect Detection Based on Machine Vision [J]. International Journal of Precision Engineering and Manufacturing-Green Technology, 2021, 9 (2): 661-691.
- [3] WU Q, ZHU M, VENKATESAN C. Application of Image Processing Sensor and Pattern Recognition in Detection of Bearing Surface Defects [J]. Journal of Sensors, 2022, 2022: 1-7.
- [4] LIN Q, ZHOU J, MA Q. Detection of the fluff fabric surface quality based on machine vision [J]. The Journal of The Textile Institute, 2021, 113 (8): 1666-1676.
- [5] LU C, LI Z, SHI S, et al. Research on Target Defect Detection Based on Machine Vision [J]. Journal of Physics: Conference Series, 2022, 2219 (1): 1-7.
- [6] SAKHARE K, KULKARNI A, KUMBHAKARN M, et al. Spectral and spatial domain approach for fabric defect detection and classification [C] //Proceedings of the International Conference on Industrial Instrumentation & Control, 2015: 640-644.
- [7] SHU Y, TAN Z. Fabric defects automatic detection using Ga-

- bor filters [C] //Proceedings of the Intelligent Control and Automation, 2004. WCICA 2004. Fifth World Congress on, 2004; 3378 - 3380.
- [8] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation [C] //Proceedings of the CVPR, 2014.
- [9] GIRSHICK R. Fast R-CNN [C] //Proceedings of the CVPR, 2017; 1440 - 1448.
- [10] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [J]. IEEE Trans Pattern Anal Mach Intell, 2017, 39 (6): 1137 - 1149.
- [11] ZHU X, LYU S, WANG X, et al. TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-captured Scenarios [C] //Proceedings of the CVPR, 2021.
- [12] REDMON J, FARHADI A. YOLO9000: Better, Faster, Stronger [C] //Proceedings of the CVPR, 2017; 6517 - 6525.
- [13] REDMON J, FARHADI A. YOLOv3: An Incremental Improvement [C] //Proceedings of the CVPR, 2018.
- [14] BOCHKOVSKIY A, WANG C Y, LIAO H. YOLOv4: Optimal Speed and Accuracy of Object Detection [C] //Proceedings of the CVPR, 2020.
- [15] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection [C] //Proceedings of the CVPR, 2016.
- [16] BERG A C, FU C Y, SZEGEDY C, et al. SSD: Single Shot MultiBox Detector [C] //Proceedings of the ECCV, 2015.
- [17] LAW H, DENG J. CornerNet: Detecting Objects as Paired Keypoints [J]. International Journal of Computer Vision, 2019, 128 (3): 642 - 656.
- [18] WEI B, XU B, HAO K, et al. Textile defect detection using multilevel and attentional deep learning network (MLMA-Net) [J]. Textile Research Journal, 2022, 92 (19 - 20): 3462 - 3477.
- [19] KOU X, LIU S, CHENG K, et al. Development of a YOLO-V3-based model for detecting defects on steel strip surface [J]. Measurement, 2021, 182: 1 - 9.
- [20] WANG T, SU J, XU C, et al. An Intelligent Method for Detecting Surface Defects in Aluminium Profiles Based on the Improved YOLOv5 Algorithm [J]. Electronics, 2022, 11 (15): 1 - 18.
- [21] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative Adversarial Nets [C] //Proceedings of the Neural Information Processing Systems, 2014.
- [22] ZHANG B, LING H, LI P, et al. Multi-Head Attention Graph Network for Few Shot Learning [J]. Computers, Materials & Continua, 2021, 68 (2): 1505 - 1517.
- [23] YUN S, HAN D, CHUN S, et al. CutMix: Regularization Strategy to Train Strong Classifiers With Localizable Features [C] //Proceedings of the ICCV, 2019.
- [24] GUO S, LI L, GUO T, et al. Research on Mask-Wearing Detection Algorithm Based on Improved YOLOv5 [J]. Sensors (Basel), 2022, 22 (13): 1 - 16.
- [25] LIU H, SUN F, GU J, et al. SF-YOLOv5: A Lightweight Small Object Detection Algorithm Based on Improved Feature Fusion Mode [J]. Sensors (Basel), 2022, 22 (15): 1 - 14.
- [26] LI S, WANG X. YOLOv5-based Defect Detection Model for Hot Rolled Strip Steel [J]. Journal of Physics: Conference Series, 2022, 2171 (1): 1 - 6.
- [27] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature Pyramid Networks for Object Detection [C] //Proceedings of the CVPR, 2017.
- [28] LIU S, QI L, QIN H, et al. Path Aggregation Network for Instance Segmentation [C] //Proceedings of the CVPR, 2018.
- [29] ZHONG Z, ZHENG L, KANG G, et al. Random Erasing Data Augmentation [C] //Proceedings of the AAAI Conference on Artificial Intelligence, 2017; 13001 - 13008.
- [30] YANGHUAN X, DONGCHENG W, HONGMIN L, et al. Flatness Defect Recognition Method of Cold Rolling Strip with a New Stacked Generative Adversarial Network [J]. steel research international, 2022, 93 (11): 1 - 17.
- [31] XIANGANG C, HENGYANG W, PENG W, et al. High Quality Coal Foreign Object Image Generation Method Based on StyleGAN-DSAD [J]. Sensors, 2022, 23 (1): 1 - 21.
- [32] JIE K, SHUJIE F. Pavement Cracks Segmentation Algorithm Based on Conditional Generative Adversarial Network [J]. Sensors, 2022, 22 (21): 298 - 302.
- [33] KARRAS T, LAINE S, AITTALA M, et al. Analyzing and Improving the Image Quality of StyleGAN [C] //Proceedings of the CVPR, 2020; 8107 - 8116.
- [34] ISOLA P, ZHU J Y, ZHOU T, et al. Image-to-Image Translation with Conditional Adversarial Networks [C] //Proceedings of the CVPR, 2017.
- [35] ZHU J Y, PARK T, ISOLA P, et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks [C] //Proceedings of the ICCV, 2017.
- [36] ZHANG G, CUI K, HUNG T Y, et al. Defect-GAN: High-Fidelity Defect Synthesis for Automated Defect Inspection [C] //Proceedings of the CVPR, 2021; 2523 - 2533.
- [37] CHOI Y, CHOI M, KIM M, et al. StarGAN: Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation [C] //Proceedings of the CVPR, 2018.
- [38] POSILOVI L, MEDAK D, SUBASIC M, et al. Generative adversarial network with object detector discriminator for enhanced defect detection on ultrasonic B-scans [J]. Neurocomputing, 2021, 459: 361 - 369.
- [39] MIRZA M, OSINDERO S. Conditional Generative Adversarial Nets [EB/OL]. Computer Science, 2014; 11 - 6. <https://arxiv.org/abs/1411.1784>.
- [40] RADFORD A, METZ L, CHINTALA S. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks [EB/OL]. ICLR, 2016; 1 - 7. <https://arxiv.org/abs/1511.06434>.