

文章编号 1004-924X(2024)08-1241-11

面向工业目标检测的密集控制阀零件数据集

王琳毅¹, 白 静^{1,3*}, 李彦梅², 李文静¹

(1. 北方民族大学 计算机科学与工程学院, 宁夏 银川 750021;

2. 六盘山实验室, 宁夏 银川 750021;

3. 北方民族大学 图像图形智能处理国家民委重点实验室, 宁夏 银川 750021)

摘要:工业生产中的自动化智能化离不开自动目标检测,而高准确性的自动目标检测则依赖于与实际场景相适应的数据集。本文针对工业实际场景,发布了一个密集控制阀零件数据集,命名为PD4CV(Part Detection for Control Valve) 2023。该数据集的图像全部来源于控制阀生产车间,图像采集完成后,首先对数据集图片进行预处理操作,接着对数据集图片中的零件目标进行标注,然后再对数据集图片进行训练集、验证集以及测试集的划分。PD4CV2023数据集共涵盖9类零件,包括510张工盘图像和15 015个零件样本,平均每张图像含有约29个零件样本。与现有的目标检测数据集相比,该数据集具有零件摆放密集、遮挡,零件尺寸差异大,部分零件外形相似,零件样本数量不均衡等特点。最后,在不同类型数据集上的预训练对比实验表明,一般场景数据集、特定工业场景数据集只适用于一般和特定任务,而代表实际控制阀生产工况下的PD4CV2023数据集,可用于控制阀零件目标检测,其具有其特殊性和不可替代性;一系列目标检测算法在该数据集上的综合对比则验证了PD4CV2023数据集在一般性目标检测、多尺度目标检测、小规模、不均衡数据下目标检测中的有效性。PD4CV2023数据集可用于面向工业的目标检测的相关研究。

关键词:深度学习;工业目标检测;数据集;控制阀零件

中图分类号:TP391.41;TG95 **文献标识码:**A **doi:**10.37188/OPE.20243208.1241

Dense control valve parts dataset for industrial object detection

WANG Linyi¹, BAI Jing^{1,3*}, LI Yanmei², LI Wenjing¹

(1. Department of Computer Science and Engineering, North Minzu University,
Yinchuan 750021, China;

2. Mount Liupan Laboratory, Yinchuan 750021, China;

3. Key Laboratory of Image and Graphics Intelligent Processing of State Ethnic Affairs Commission,
North Minzu University, Yinchuan 750021, China)

* Corresponding author, E-mail: baijing_nun@163.com

Abstract: Automated intelligence in industrial production is inseparable from automatic object detection, and high-accuracy automatic object detection relies on datasets adapted to the actual scene. This article published a dense control valve parts dataset for industrial practical scenarios, named PD4CV (Part Detection for Control Valve) 2023. The image of this dataset came from the control valve production workshop, and after the image collection was completed, it underwent steps such as dataset preprocessing, da-

收稿日期:2023-09-22;修订日期:2023-11-14.

基金项目:国家自然科学基金项目(No. 62162001);宁夏自然科学基金项目(No. 2022AAC02041);宁夏优秀人才支持计划项目;北方民族大学研究生创新项目(No. YCX23150)

taset annotation, and dataset partitioning. The images of this dataset were all from the control valve production workshop. After the image collection was completed, the dataset images were first preprocessed, followed by labeling the part targets in the dataset images. Then, the dataset images were divided into training, validation, and testing sets. The PD4CV2023 dataset covered a total of 9 types of parts, including 510 workstation images and 15 015 part samples, with an average of approximately 29 part samples per image. Compared with the existing object detection datasets, this dataset had the characteristics of dense placement and occlusion of parts, large size differences of parts, similar shapes of some parts, and unbalanced number of parts samples. Finally, pre training comparative experiments on different types of datasets show that general scenario datasets and specific industrial scenario datasets are only suitable for general and specific tasks, while the PD4CV2023 dataset, which represents the actual production conditions of control valves, can be used for target detection of control valve parts, and has its particularity and irreplaceability; a comprehensive comparison of a series of algorithms on this dataset verifies the effectiveness of PD4CV2023 dataset in general object detection, multi-scale object detection, and object detection under small-scale and imbalanced data. The PD4CV2023 dataset can be used for research on industrial oriented object detection algorithms.

Key words: deep learning; industrial object detection; dataset; control valve parts

1 引 言

随着计算机视觉的快速发展,目标检测作为其中的核心问题之一,已经被广泛应用于医疗、交通、农业以及工业等领域^[1-3]。其中,零件检测为目标检测在工业领域的一个基本问题,旨在识别密集空间范围内(如工盘)的零件类型及其位置,以有效支持零件自动分拣、智能装配等后续任务,促进了工业互联网^[4]的发展。然而,由于零件种类多样,特定空间内零件摆放密集、无序、彼此遮挡,且部分零件相似度高,人工分拣费时费力且易产生错分^[5],传统的目标检测算法则存在手工标注要求高、适应性差、定位及识别能力弱的问题,严重影响了装配环节的生产效率。如今,深度学习在目标检测领域不断取得突破性进展,学者们开始考虑将其引入至零件检测任务中,以进一步提高检测的准确性。

然而,工业零件检测领域有效数据集的构建成为制约发展的主要瓶颈。目前,研究人员提出了多个自然场景数据集,以便帮助视觉研究者更好地开发、改进和评价目标检测算法。其主要包括 MS COCO^[6], PASCAL VOC^[7], Open Images^[8], ImageNet^[9]等数据集;但针对工业场景的目标检测数据集较少,工业目标检测数据集通常是

为了解决特定的工业场景中的任务而构建的,相对于自然场景数据集,其适用范围较小、针对性强,主要有支持三维检测的 MVTec ITODD^[10]和 T-LESS^[11],以及支持缺陷检测的 KolektorS-DD^[12]、DeepPCB^[13]和 NEU-CLS^[14]。

总体来说,自然场景的目标检测数据集是在自然环境下的采集的图像数据,样本内容形式简单,类型单一,基本上与工业场景无关,并且目标类型与零件检测任务要求不符,无法为零件检测提供训练数据支持。现有的工业场景的目标检测数据集主要用于解决工业上的缺陷检测^[15-17],只有少数与零件相关的数据集,通常适用于三维检测^[18-19],难以适用于特定需求的零件检测。上述的数据集对于特定场景的目标检测任务无法提供较好数据支持。

针对以上需求,为了弥补控制阀零件数据集的空白,提高工业生产效率和促进工业智能化生产,本文构建了一个关于控制阀的工业场景数据集;并通过 PD4CV2023 数据集必要性和有效性测试实验进行分析,验证了该数据集的特殊性和不可替代性以及有效性,以供相关领域学者使用。(PD4CV2023 数据集链接为: <https://github.com/wangliny01/PD4CV2023>)

2 PD4CV2023数据集

2.1 数据集收集和预处理

PD4CV2023数据集的所有图像均源自某控制阀生产企业的实际生产车间。图1为PD4CV2023数据集收集和预处理过程。

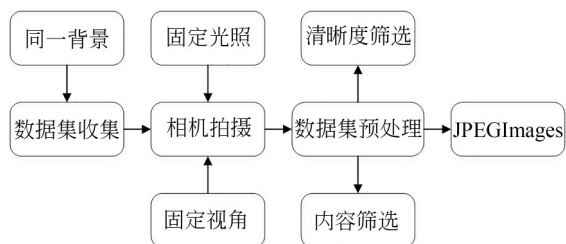


图1 PD4CV2023数据集收集和预处理

Fig. 1 PD4CV2023 dataset collection and preprocessing

在装配过程中,工作人员使用相机从固定视角对零件工盘进行垂直拍摄。拍摄对象是摆放在工盘中的控制阀零件。初始图像共采集了253张,但由于工盘中的零件类型、尺寸、位置和数量存在差异,每张图像都呈现出目标密集、背景复杂的特征。

为确保数据质量,首先对初始图像的清晰度进行筛选,保留清晰完整的图像。接着,对图像内容进行筛选,确保图像内容为工盘中的控制阀零件。经过筛选后,共有170张图像被编号并保存在名为“JPEGImages”文件夹中。

2.2 数据集标注

根据PASCAL VOC数据集的构建标准,对经过预处理的图像进行标注。利用LabelImg工具,在PD4CV2023数据集中的每张图像上为每个零件绘制一个最小的矩形框,同时标注出它们

所属的零件类别。随后,按照顺序将完成标注的文件保存至名为“Annotations”的文件夹中。图2为零件标记完成后的展示结果。

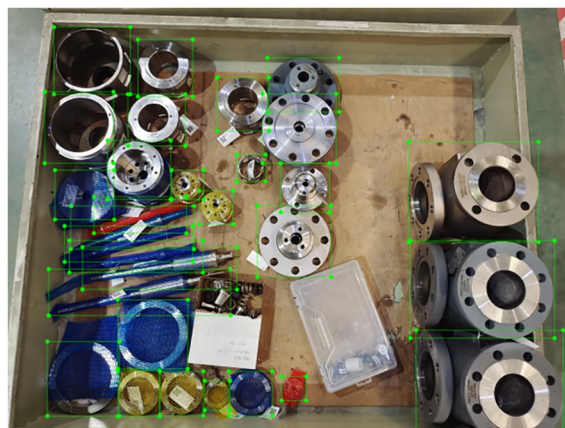


图2 PD4CV2023数据集标注

Fig. 2 PD4CV2023 dataset annotation

2.3 数据集扩增

在深度学习的训练过程中,数据集的多样性和丰富性对于模型的性能和泛化能力至关重要。因此,在初始图像筛选完成后,如图3所示,对原有数据随机地进行了镜像、旋转和裁剪操作,以扩充数据集。具体来说,镜像操作是指将图像水平翻转,产生与原图像对称的新图像。旋转操作则是将图像旋转一定的角度,产生新的视角和构图。裁剪操作则是对图像进行切割,提取出其中的一部分,产生具有部分内容缺失的图像。通过这些操作使数据集样本扩增为原来的3倍。

通过这些扩充措施,可以提高模型的泛化能力,使其在预测新样本时更加准确和稳定。同时,也能够提高模型的鲁棒性,使其更好地应对各种复杂情况和噪声干扰。这对于实际应用中

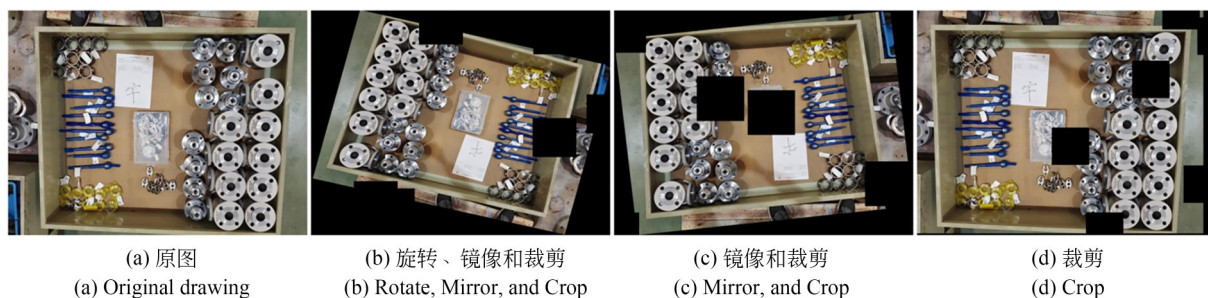


图3 PD4CV2023数据集扩增

Fig. 3 PD4CV2023 dataset amplification

的模型性能和准确度至关重要。

2.4 数据集划分

PD4CV2023数据集共包含 510 张图像,分为 9 个类别。该数据集的零件示例如图 4 所示,从左至右依次为阀体(Valve Body)、上阀盖(Upper Valve Cover)、套筒(Sleeve)、阀杆(Valve Stem)、阀座(Valve Seat)、柱状阀芯(SpoolAP)、杆状阀芯(SpoolEAP)、

上套筒(Upper Sleeve)和导向套(Guide Sleeve)。

本文按照图像数量随机地将该数据集划分为 80% 的训练集、10% 的测试集和 10% 的验证集来训练不同的目标检测方法。从表 1 可以看出,训练集、验证集和测试集的图像数量为 408 张、51 张和 51 张,零件样本数量分别为 12 089 个、1 523 个和 1 403 个。



图 4 PD4CV2023 数据集零件类型

Fig. 4 PD4CV2023 dataset part types

表 1 PD4CV2023 数据集数据统计

Tab. 1 PD4CV2023 dataset data statistics (pieces)

	图像(张)	样本	阀体	上阀盖	套筒	柱状阀杆	杆状阀杆	阀座	阀芯	上套筒	导向套
训练集	408	12 089	1 820	2 381	2 295	1 921	590	2 210	449	195	228
验证集	51	1 523	276	283	305	278	47	260	35	25	14
测试集	51	1 403	232	285	280	240	59	221	47	17	22
总数	510	15 015	2 328	2 949	2 880	2 439	696	2 691	531	237	264

2.5 数据集特点分析

PD4CV2023 中的所有图像均来自工业场景。鉴于工业场景的实际复杂性,该数据集呈现出一些独特的特点。

(1)零件摆放密集且存在遮挡。观察图 5 所示的零件密集遮挡摆放实例可以发现,本数据集内样本对象之间存在多样化的遮挡。图 5 第一行,由于零件摆放的密集程度不同,工盘内零件之间存在一定程度的遮挡,零件摆放稀疏的工盘中,零件遮挡的影响较小,而零件摆放密集的工盘中,零件遮挡造成的影响较大。

图 5 第二行,有些零件表面覆盖着一层塑料网状保护膜,形成另一种形式的遮挡;零件自带的 QR 标签也会对其本身及周围零件造成不同程

度的遮挡;零件垂直或者水平摆放也会影响零件的识别效果,导致误检或者错检;工盘中的其他杂物也会对零件造成一定程度的遮挡。

从表 2 的统计数据来看,MS COCO,PASCAL VOC,Open Images,ImageNet,DOTA^[20]和 LEVIR^[21]数据集中单张图像的平均样本数为 7.19 个、2.89 个、5 个、1.37 个、3.5 个和 0.5 个,而 PD4CV2023 数据集中单张图像平均样本数约为 29 个,远高于其他目标检测数据集,存在明显的样本密集性。

(2)零件尺寸差异大。控制阀产品中不同零件尺寸差异较大,没有相对固定的尺寸。如图 6 所示,左上方阀体的尺寸要明显大于右下方上阀盖的尺寸。



图 5 零件密集、遮挡摆放实例

Fig. 5 Example densely obstructed placement of parts

表 2 各个数据集的平均样本数	
Tab. 2 Average number of samples per dataset(pieces)	
数据集	平均样本数
MS COCO	7.19
PASCAL VOC(07++12)	2.89
Open Images	5.00
ImageNet	1.37
DOTA	3.50
LEVIR	0.50
PD4CV2023	29.00

(3)部分零件相似度较高。如图 7 所示,塑料包装膜下的阀杆和柱状阀芯形态相似,二者均为杆状物体,因而具有相似的特征;正面观察下,上套筒和导向套的形状也非常相近。

(4)零件样本数量不均衡。对比表 2 中各类的样本数量,可以发现上套筒和导向套的数量



图 6 零件尺寸差异实例

Fig. 6 Example part size differences

最少(为 237 和 264);杆状阀杆数量居中,约为前两类样本的 3 倍;剩余类别样本数量超过了 2 100 个,约为前两类样本的 10 倍;即,PD4CV2023 数据集存在着明显的类内样本不均衡的问题。

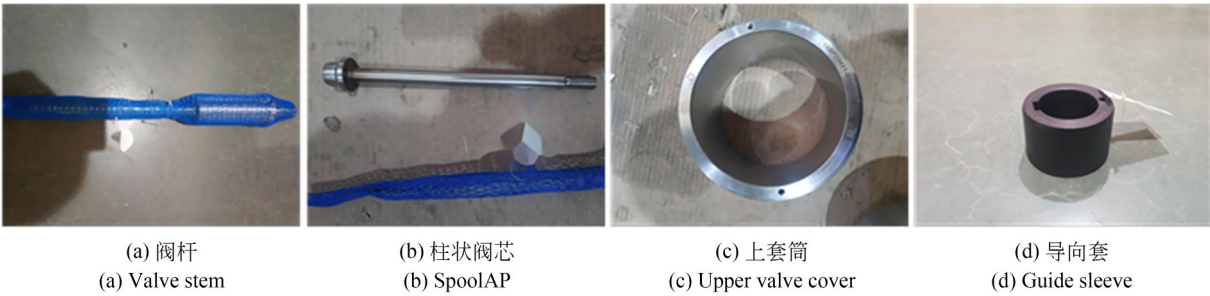


图 7 相似零件对比

Fig. 7 Comparison similar parts

综上所述,PD4CV2023 数据集具有零件摆放密集且存在遮挡,零件尺寸差异大,部分零件相似度较高,零件样本数量不均衡的特点。现有的数据集很难同时兼具工业场景数据集的特点,该数据集更符合工业场景的实际情况。

2.6 数据集对比

为了验证目标检测算法的效能,研究者们精心构建了多个与目标检测密切相关的数据集,这些数据集旨在为视觉研究者提供帮助,以便他们能够更好地开发、改进和评估目标检测算法。以下是对八个常用数据集和 PD4CV2023 数据集进行的对比分析。表 3 为各个数据集在

类别、图像数量、用途以及样本标记数方面的差异。

在类别方面,这些数据集涵盖了各种不同的类别,包括车辆、行人、飞机、舰船、电路板缺陷以及钢板缺陷等。然而,PD4CV2023 数据集更加专注于控制阀零件的检测,提供了更具体的应用场景。

在图像数量上,不同的数据集有着显著的区别。虽然 PD4CV2023 数据集的图像数量较少,但 PD4CV2023 数据集包含了超过 15 000 个样本,部分零件的检测难度较大,为研究者提供了充足的数据资源。

表 3 部分相关数据集对比

Tab. 3 Comparison of relevant datasets

名称	类别	图像数	用途	样本标记数
MS COCO	80	123 287	一般目标检测	886 266
PASCAL VOC(07++12)	20	21 503	一般目标检测	62 199
ImageNet	200	349 319	一般目标检测	478 806
DOTA	15	2 806	航空目标检测	188 282
LEVIR	3	22 000	遥感目标检测	11 000
KolektorSDD	1	399	工业缺陷检测	52
DeepPCB	6	3 000	电路板缺陷检测	8 853
NEU-CLS	6	1 800	钢板缺陷检测	1 800
PD4CV2023	9	510	控制阀目标检测	15 015

此外,各个数据集的用途也各不相同。一些数据集主要用于一般目标检测,评估目标检测算法的性能,还有一些数据集用于航空目标检测和缺陷目标检测等。而 PD4CV2023 数据集则更注重于实际应用中的控制阀零件检测,具有更强的实用性。

3 PD4CV2023 数据集的验证及分析

3.1 实验环境

本文实验环境为 Window10 系统,模型采用 Pytorch 框架实现。具体的实验硬件环境为 1 块 Intel9(R) Xeon(R) Gold 6154 处理器,1 张显存大小为 12 G 的 NVIDIA TITAN V 显卡。模型统一训练 300 个 Epoch, Batch Size 为 8,输入图像的尺寸为 640×640 pixels。

3.2 评价指标

目标检测常用的评价指标有 FPS(Frame Per Second)、精确率(Precision)、召回率(Recall)、平均精确率均值(Average Precision, AP)和 mAP(mean Average Precision)等。本文采用 FPS 和 mAP 作为评价指标。

3.3 验证方法

为了对 PD4CV2023 数据集的有效性和可用性进行充分的验证评估,本文使用 Faster-RCNN^[22], SSD^[23], DETR^[24], YOLOv3^[25], YOLOv5^[26]以及 YOLOv8^[27]等目标检测算法进行了对比实验。其中:

(1)Faster R-CNN 通过引入候选区域生成网络(Region Proposal Network, RPN)实现了端到端的目标检测,相较于传统的 R-CNN 算法,具有更快的速度和更高的准确率。

(2)SSD 通过使用多个卷积层和特征图金字

塔,以及利用先验框和多尺度预测的方法,实现了高效准确的目标检测。

(3)DETR 网络结构主要由 Transformer 模型和附加的编码器-解码器结构组成。通过将编码器和解码器结合在一起,DETR 能够端到端地学习图像中目标的位置和类别。相较于传统的目标检测算法,DETR 能够通过 Transformer 的注意力机制同时处理所有目标,从而更好地捕捉目标之间的关系和上下文信息。

(4)YOLOv3 在不同尺度上进行了目标检测,通过分别预测不同尺度下的边界框,有效地解决了小目标检测的问题。

(5)YOLOv5 通过特征金字塔网络和检测头实现了在不同尺度上进行目标检测的能力,结合了高效的基础网络和有效的损失函数,使它在目标检测任务中取得了较好的性能。

(6)YOLOv8 在 YOLOv5 的基础上将 C3 模块更换为 C2F 模块,减少了不必要的卷积层,将预测阶段的目标识别与位置预测进行了解耦,是 YOLO 系列的最新版本。与前几个版本相比,YOLOv8 在准确性、速度和效率方面都有所提高。

综上,本文既选择了双阶段的 Faster R-CNN、单阶段 SSD 和基于 Transformer 的 DETR,又选择了经典的 YOLOv3 和 YOLOv5 以及

先进的 YOLOv8。通过上述方法在 PD4CV2023 数据集的实验结果来验证数据集的必要性和有效性。

3.4 实验结果分析

3.4.1 PD4CV2023 数据集的必要性测试

为了验证构建 PD4CV2023 数据集的必要性,本节设计了该实验。选择了两种典型的目标检测算法,YOLOv5s 和 YOLOv8s 作为待训练模型,并使用现有的目标检测数据集 MS COCO, KolektorSDD, NEU-CLS 以及本文构建的 PD4CV2023 数据集作为训练数据对模型进行预训练。

从表 4 的结果可以清晰地看出,不同训练数据所得到的对比结果存在明显差异。当模型分别使用 MS COCO、KolektorSDD 和 NEU-CLS 作为训练集,使用 PD4CV2023 作为测试集,其 mAP 值都为 0;而当使用 PD4CV2023 数据集进行训练,用 PD4CV2023 数据集来测试时,YOLOv5s 和 YOLOv8s 的 mAP 值分别为 88.20% 和 85.40%。

这充分表明现有的目标检测数据集无法对控制阀零件进行有效地检测和识别。在这种情况下,构建一个与控制阀相关的专门数据集是必要的。

表 4 不同训练集训练模型在 PD4CV2023 数据集上的零件检测结果对比
Tab. 4 Comparison of part detection results of different training sets of models on the PD4CV2023 dataset (%)

	YOLOv5s				YOLOv8s			
训练集	MS COCO	KolektorSDD	NEU-CLS	PD4CV2023	MS COCO	KolektorSDD	NEU-CLS	PD4CV2023
测试集	PD4CV2023	PD4CV2023	PD4CV2023	PD4CV2023	PD4CV2023	PD4CV2023	PD4CV2023	PD4CV2023
mAP	0	0	0	88.20	0	0	0	85.40

3.4.2 PD4CV2023 数据集的有效性测试

本节中,设计了三个实验对 PD4CV2023 数据集的有效性进行综合测试。

实验一:不同目标检测方法的性能对比
首先,以整个数据集上的目标检测指标作为评价对象,测试数据集的有效性。本实验中,选用了 3.3 节提供的六种目标检测模型,在 PD4CV2023 数据集上进行目标检测任务的训练和测试,得到了代表准确率的 mAP 值和检测速度的 FPS 值。

训练结果如表 5 所示,从速度指标来看,YOLOv5s 的检测速度最快,FPS 为 47,Faster-RCNN 的速度最慢,FPS 为 10;从检测准确率来看,YOLOv5s 的检测结果最高,mAP 值为 88.20%,DETR 的检测结果最低,mAP 值为 55.11%,而其余模型的 mAP 值都达到了 75% 以上。

以上结果初步说明 PD4CV2023 数据集能够有效区分不同模型的目标检测质量,同时,也能区分不同模型的检测速度,即,PD4CV2023 数据

表 5 不同目标检测方法在 PD4CV2023 数据集的检测性能对比

Tab. 5 Comparison of detection performance of different object detection methods in the PD4CV2023 dataset (%)

年份	模型	mAP	FPS
2016	Faster-RCNN	75.38	10
2016	SSD	75.83	23
2017	DETR	55.11	20
2018	YOLOv3	80.67	17
2020	YOLOv5s	88.20	47
2023	YOLOv8s	85.40	42

集能够甄别不同目标检测算法的优劣,说明了该数据集的有效性。

实验二:不同零件尺寸的性能对比

表 6 不同目标检测方法在 PD4CV2023 数据集不同尺寸零件上的对比

Tab. 6 Comparison of different object detection methods on parts of different sizes in the PD4CV2023 dataset (%)

尺寸	类别	Faster-RCNN	SSD	DETR	YOLOv3	YOLOv5s	YOLOv8s	均值
较小	导向套	54.19	69.83	55.11	59.49	76.80	66.40	61.86
	套筒	64.07	91.57	37.93	95.47	98.80	98.00	80.97
	柱状阀杆	69.17	65.26	48.65	79.92	91.20	91.60	74.30
	均值	62.48	75.55	47.23	78.29	88.93	85.33	
中等	杆状阀杆	68.51	83.32	57.85	82.69	91.70	87.00	78.51
	阀座	65.21	82.44	34.00	91.47	95.50	96.30	77.49
	阀芯	75.27	37.90	46.47	66.38	55.20	50.70	55.32
	均值	69.66	67.89	46.11	80.18	80.80	78.00	
较大	阀体	94.32	96.52	90.84	99.02	98.60	98.50	96.30
	上阀盖	92.35	96.57	80.46	98.45	96.30	97.10	93.54
	上套筒	95.37	59.09	55.36	53.11	89.50	82.90	72.56
	均值	94.01	84.06	75.55	83.53	94.80	92.83	

(2)对于尺寸中等的杆状阀杆、阀座和阀芯, YOLOv3 的整体检测效果最好,检测结果各项都大于均值, YOLOv5s 在杆状阀杆和阀座的准确率远高于均值, YOLOv5s 在中等尺寸零件上的准确率的均值最高, YOLOv3 次优; DETR 对于中等尺寸零件的检测结果最差,检测结果低于均值,并且在中等尺寸零件上准确率的均值最低。

(3)对于尺寸较大的阀体、上阀盖和上套筒, YOLOv5s 和 YOLOv8s 的检测结果都高于均值, YOLOv5s 在较大尺寸零件上的准确率均值最高, Faster-RCNN 次优;虽然 DETR 在阀体、上阀

根据不同种类零件的尺寸,对 PD4CV2023 数据集进行了细分,并测试和分析了该数据集对多尺度目标检测任务支持程度。如表 6 所示,实验中,将 PD4CV2023 数据集按照该类零件的平均尺寸将其划分成尺寸较小的导向套、套筒和柱状阀杆,尺寸中等的杆状阀杆、阀座和阀芯,尺寸较大的阀体、上阀盖和上套筒。在 6 种经典的目标检测算法中:

(1)对于尺寸较小的导向套、套筒以及柱状阀杆, YOLOv5s 的检测效果最好,其导向套、套筒以及柱状阀杆的检测结果远高于均值,并且 YOLOv5s 在较小尺寸零件上的准确率的均值最高, YOLOv8s 次优; DETR 的检测效果最差,其导向套、套筒以及柱状阀杆的检测结果明显低于均值,在较小尺寸零件上的准确率的均值最低。

盖和上套筒上的准确率低于均值,但是 DETR 在较大尺寸零件上的准确率均值要高于其在较小尺寸和中等尺寸上准确率的均值,说明 DETR 对于较大尺寸的零件有较好的检测能力。

从以上数据可以看到, PD4CV2023 数据集同时包含大中小不同尺度的零件,能够充分地测试不同算法对大、中、小目标的检测能力,即, PD4CV2023 数据集可有效测试多尺度目标检测算法的优劣。

实验三:不同零件数量规模的性能对比

根据不同种类零件的比例对 PD4CV2023 数

数据集进行了细分,并测试和分析了该数据集对小规模、不均衡数据下目标检测任务的支持程度。如表 7 所示,上套筒、导向套、阀芯以及杆状阀杆数量较少,只占整体样本数量的 1.58%,1.76%,3.54% 和 4.64%,比例小于 5% 划分为较小规模;阀体、柱状阀杆、阀座、上阀盖和套筒的数量较多,占整体样本数量的 15.50%,16.24%,17.92%,19.18% 和 19.64%,可以看出该数据集具有明显的样本不平衡性。对比来看,6 种经典的目标检测算法中:

表 7 不同目标检测方法在 PD4CV2023 数据集中不同规模零件上的对比

Tab. 7 Comparison of different object detection methods on parts of different sizes in the PD4CV2023 dataset (%)

规模	类别	数量(个)	比例	Faster-RCNN	SSD	DETR	YOLOv3	YOLOv5s	YOLOv8s	均值
较小	上套筒	237	1.58	95.37	59.09	55.36	53.11	89.50	82.90	72.56
	导向套	264	1.76	54.19	69.83	55.11	59.49	76.80	66.40	61.86
	阀芯	531	3.54	75.27	37.90	46.47	66.38	55.20	50.70	55.32
	杆状阀杆	696	4.64	68.51	83.32	57.85	82.69	91.70	87.00	78.51
			均值	73.34	62.54	53.70	65.43	78.30	71.75	
较大	阀体	2 328	15.50	94.32	96.52	90.84	99.02	98.60	98.50	96.30
	柱状阀杆	2 439	16.24	69.17	65.26	48.65	79.92	91.20	91.60	74.30
	阀座	2 691	17.92	65.21	82.44	34.00	91.47	95.50	96.30	77.49
	套筒	2 880	19.18	64.07	91.57	37.93	95.47	98.80	98.00	80.97
	上阀盖	2 949	19.64	92.35	96.57	80.46	98.45	96.30	97.10	93.54
			均值	77.02	86.47	58.38	92.87	96.08	96.30	

(1)YOLOv5s 和 YOLOv8s 在导向套、上套筒和杆状阀杆这三个零件的检测精度要高于均值,YOLOv5s 在较小规模的零件上检测精度的均值最高,即,YOLOv5s 在少样本类别上拥有较高的目标检测精度;而 DETR 对于这四个零件的检测精度远低于均值,在较小规模的零件上检测精度的均值最低,不适合检测样本数量少的零件。

(2)YOLOv3, YOLOv5s 和 YOLOv8s 在阀体、柱状阀杆、阀座、套筒和上阀盖这五个零件的检测精度要高于均值,YOLOv8s 在较大规模的零件上检测精度的均值最高,YOLOv5s 次高,即,YOLOv8s 和 YOLOv5s 在多样本类别上拥有较高的目标检测精度;同样 DETR 对于这五个零件的检测精度远低于均值,在较大规模的零件上检测精度的均值最低。

(3)通过分析对比,YOLOv5s 的整体检测效果最好,即,YOLOv5s 能够良好适用于样本数量少且不均衡的目标检测任务;而 DETR 的整体检测效果较差,对数据集数量级质量有较高要求。适合数据量大且样本丰富的数据集。

综上所述,PD4CV2023 数据集可有效测试各种目标检测算法在小规模、非均衡数据集上的检测性能。

3.5 实践对比分析

如表 8 所示,通过对比深度学习方法(以 YOLOv5s 为例)和人工分拣控制阀的速度和错误率,可以得出深度学习方法明显提高了控制阀的分拣速度,错误率比人工略有降低。另外,由于 PD4CV2023 数据集中零件摆放密集、尺寸差异大以及部分相似的特点,以及现有深度学习方法没有针对其特点进行相应的优化改进,无法使错误率大幅降低。未来针对工业场景数据集中目标密集、尺寸差异大以及部分相似的特点,设计专门的目标检测算法是极其有必要的。

表 8 不同方式的速度和错误率对比

Tab. 8 Comparison of efficiency and error rates of different methods

方式	分拣速度	错误率
人工	50 个/h	12.00%
YOLOv5s	1 450 个/h	11.80%

4 结 论

目前,工业场景数据集大多是为缺陷检测提供数据支持,没有专门用于控制阀零件检测的数据集,并且现有目标检测数据集也无法满足控制阀生产企业对于零件目标检测的需求。为此,本文构建了一个面向工业目标检测的密集控制阀零件数据集,该数据集收集了 9 种控制阀零件,包含了 510 张数据集图像和 15 015 个样本标签。此外,

通过对不同数据集上的预训练对比实验和不同类型的目标检测算法在 PD4CV2023 数据集上的综合对比实验进行分析,验证了该数据集具有其特殊性和不可替代性以及在一一般性目标检测、多尺度目标检测、小规模、不均衡数据下目标检测中的有效性。在接下来的工作中,将收集更多的待检测图像,进一步更好地来支持工业场景的检测需求;另外,也将针对工业场景的实际情况,研究出更好的目标检测算法,来提高目标检测的效果。

参考文献:

- [1] SUN X Y, YANG S S, ZHAO C. Lightweight industrial image classifier based on federated few-shot learning[J]. *IEEE Transactions on Industrial Informatics*, 2023, 19(6): 7367-7376.
- [2] AGGARWAL A, KUMAR V, GUPTA R. Object detection based approaches in image classification: a brief overview [C]. 2023 *IEEE Guwahati Subsection Conference (GCON)*. Guwahati, India. IEEE, 2023: 1-6.
- [3] YASMINE G, MAHA G, HICHAM M. Overview of single-stage object detection models: from yolov1 to Yolov7 [C]. 2023 *International Wireless Communications and Mobile Computing (IWCMC)*. Marrakesh, Morocco. IEEE, 2023: 1579-1584.
- [4] 庄存波, 刘检华, 隋秀峰, 等. 工业互联网推动离散制造业转型升级的发展现状、技术体系及应用挑战[J]. *计算机集成制造系统*, 2019, 25(12): 3061-3069.
ZHUANG C B, LIU J H, SUI X F, *et al.* Status, technical architecture and application challenges for transformation and updating of discrete manufacturing industry driven by industrial Internet[J]. *Computer Integrated Manufacturing Systems*, 2019, 25(12): 3061-3069. (in Chinese)
- [5] WANG Y T, SHEN W W. Mechanical parts detection algorithm based on enhanced faster R-CNN [C]. 2021 *China Automation Congress (CAC)*. Beijing, China. IEEE, 2021: 4348-4353.
- [6] EVERINGHAM M, ESLAMI S MALI, VAN GOOL L, *et al.* The pascal visual object classes challenge: a retrospective[J]. *International Journal of Computer Vision*, 2015, 111(1): 98-136.
- [7] LIN T Y, MAIRE M, BELONGIE S, *et al.* Microsoft COCO: Common Objects in Context [M]. Computer Vision-ECCV 2014. Cham: Springer International Publishing, 2014: 740-755.
- [8] KUZNETSOVA A, ROM H, ALLDRIN N, *et al.* The open images dataset V4 [J]. *International Journal of Computer Vision*, 2020, 128(7): 1956-1981.
- [9] DENG J, DONG W, SOCHER R, *et al.* ImageNet: a large-scale hierarchical image database [C]. 2009 *IEEE Conference on Computer Vision and Pattern Recognition*. Miami, FL, USA. IEEE, 2009: 248-255.
- [10] DROST B, ULRICH M, BERGMANN P, *et al.* Introducing MVTec ITODD-a dataset for 3D object recognition in industry [C]. 2017 *IEEE International Conference on Computer Vision Workshops (ICCVW)*. Venice, Italy. IEEE, 2017: 2200-2208.
- [11] HODAN T, HALUZA P, OBDRŽÁLEK Š, *et al.* T-LESS: an RGB-D dataset for 6D pose estimation of texture-less objects [C]. 2017 *IEEE Winter Conference on Applications of Computer Vision (WACV)*. Santa Rosa, CA, USA. IEEE, 2017: 880-888.
- [12] TABERNIK D, ŠELA S, SKVARČ J, *et al.* Segmentation-based deep-learning approach for surface-defect detection [J]. *Journal of Intelligent Manufacturing*, 2020, 31(3): 759-776.
- [13] KHALILIAN S, HALLAJ Y, BALOUCHES-TANI A, *et al.* PCB defect detection using denoising convolutional autoencoders [C]. 2020 *International Conference on Machine Vision and Image Processing (MVIP)*. Iran. IEEE, 2020: 1-5.
- [14] HE Y, SONG K C, MENG Q G, *et al.* An end-to-end steel surface defect detection approach via

- fusing multiple hierarchical features [J]. *IEEE Transactions on Instrumentation and Measurement*, 2020, 69(4): 1493-1504.
- [15] CHEN H X, DU Y Z, FU Y Q, *et al.* DCAM-net: a rapid detection network for strip steel surface defects based on deformable convolution and attention mechanism [J]. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72: 5005312.
- [16] WANG W Y, MI C F, WU Z H, *et al.* A real-time steel surface defect detection approach with high accuracy [J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 71: 5005610.
- [17] 郭峰, 朱启兵, 黄敏, 等. 基于改进YOLOV4的陶瓷基板瑕疵检测[J]. *光学精密工程*, 2022, 30(13): 1631-1641.
- GUO F, ZHU Q B, HUANG M, *et al.* Defect detection in ceramic substrate based on improved YOLOV4 [J]. *Opt. Precision Eng.*, 2022, 30(13): 1631-1641. (in Chinese)
- [18] 李宝平, 戚恒熠, 王满利, 等. 联合3D建模与改进CycleGAN的故障数据集扩增方法[J]. *光学精密工程*, 2023, 31(16): 2406-2417.
- LI B P, QI H Y, WANG M L, *et al.* Equipment fault dataset amplification method combine 3D model with improved CycleGAN [J]. *Opt. Precision Eng.*, 2023, 31(16): 2406-2417. (in Chinese)
- [19] CHEN W, ZHAO J, ZHAO W L, *et al.* Shape-aware monocular 3D object detection [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2023, 24(6): 6416-6424.
- [20] XIA G S, BAI X, DING J, *et al.* DOTA: a large-scale dataset for object detection in aerial images [C]. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT, USA. IEEE, 2018: 3974-3983.
- [21] CHEN H, SHI Z W. A spatial-temporal attention-based method and a new dataset for remote sensing image change detection [J]. *Remote Sensing*, 2020, 12(10): 1662.
- [22] 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.
- [23] LIU W, ANGUELOV D, ERHAN D, *et al.* SSD: Single Shot MultiBox Detector [M]. Computer Vision-ECCV 2016. Cham: Springer International Publishing, 2016: 21-37.
- [24] CARION N, MASSA F, SYNNAEVE G, *et al.* End-to-End object detection with transformers [M]. Computer Vision-ECCV 2020. Cham: Springer International Publishing, 2020: 213-229.
- [25] REDMON J, FARHADI A. YOLOv3: an incremental improvement [C]. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2018.
- [26] ULTRALYTICS. YOLOv5 [EB/OL]. (2021-06-10) [2023-09-15] <https://github.com/ultralytics/YOLOv5>, 2021.
- [27] ULTRALYTICS. YOLOv8 [EB/OL]. (2023-01-10) [2023-09-15] <https://github.com/ultralytics/YOLOv8>, 2023.

作者简介:



王琳毅(1997—),男,山西长治人,硕士研究生,2021年于太原理工大学获得学士学位,主要从事机器学习、深度表征学习、计算机视觉应用等方面的研究。E-mail: 627376504@qq.com

通讯作者:



白静(1982—),女,宁夏青铜峡人,博士,教授,2010年于浙江大学获得博士学位,主要从事机器学习、深度表征学习、计算机视觉等方面的研究。E-mail: baijing_nun@163.com