

基于 Yolov5s 的作业现场阀门规范操作识别方法

赵崇志¹, 卢俊², 田源¹, 邢金台¹, 董金平¹, 张蕾¹, 田枫²

(1. 中国石油天然气股份有限公司冀东油田分公司, 河北 唐山 063004;

2. 东北石油大学 计算机与信息技术学院, 黑龙江 大庆 163318)

摘要:将基于深度学习算法的图像识别技术应用到油田作业现场监控领域,解决阀门操作的规范性问题。采用 YOLOv5s 网络作为作业现场阀门规范操作识别的模型,通过图像增强方法解决阀门周围的栏杆遮挡问题,制作阀门操作的数据集,保证了数据集的丰富性。然后在自制阀门操作分类的数据集上对网络进行训练,利用训练好的 YOLOv5s 网络提取作业人员不同的阀门操作图像的特征和位置信息,实现作业现场阀门规范操作的识别。经过测试,验证该模型最终检测准确率达到93%,检测速度能达到实时的效果。基于 YOLOv5s 网络的作业现场阀门规范操作识别的模型在不同光照和视角等条件下,检测准确率高,鲁棒性好、模型计算速度快。满足了油田作业现场实际需求,解决了油田作业现场员工在阀门操作上的安全问题。

关键词:深度学习;作业现场;YOLOv5s;目标检测;阀门操作识别

中图分类号:TP399

文献标识码:A

文章编号:1673-629X(2022)05-0216-05

doi:10.3969/j.issn.1673-629X.2022.05.036

Yolov5s-based Method for Identifying Valve Specification Operations at Job Sites

ZHAO Chong-zhi¹, LU Jun², TIAN Yuan¹, XING Jin-tai¹, DONG Jin-ping¹,

ZHANG Lei¹, TIAN Feng²

(1. Jidong Oilfield Company, PetroChina, Tangshan 063004, China;

2. School of Computer & Information Technology of Northeast Petroleum University, Daqing 163318, China)

Abstract: To apply the image recognition technology based on deep learning algorithm to the field of oilfield operation site monitoring to solve the problem of valve operation specification. In this paper, the YOLOv5s network is used as a model for the recognition of valve normative operations at operation sites, and the problem of obscuring the railings around the valves is solved by image enhancement methods to produce a dataset of valve operations, which ensures the richness of the dataset. The network is then trained on the dataset of homemade valve operation classification, and the trained YOLOv5s network is used to extract the features and location information of different valve operation images of operators to achieve the recognition of valve regulation operation at the job site. After testing, it was verified that the final detection accuracy of the model reached 93% and the detection speed could achieve real-time results. The YOLOv5s network-based model for the recognition of regulated valve operations at the job site has high detection accuracy, strong robustness and fast model computation under different lighting and viewing angles, which meets the actual needs of oilfield operation sites and solves the safety problems of oilfield operation site employees in valve operation.

Key words: deep learning; operation sites; YOLOv5s; object detection; valve operation identification

0 引言

深度学习以及新一轮人工智能的发展对计算机视觉领域相关技术的进步起到了极大的推动作用。计算机视觉应用不断地深入到各行各业,作为国之利器的

石油行业也在逐步向智慧油田管理迈进,智慧油田的管理需要更加智能的监控视频技术作为支撑。目前油田生产过程中主要依靠人工对作业现场进行监控,摄像头存在路数众多,监督成本高等问题。为了快速定

收稿日期:2021-05-06

修回日期:2021-09-07

基金项目:黑龙江省省属本科高校基本科研业务费项目(东北石油大学优秀中青年科研创新团队)(KYCXTD201903);东北石油大学研究生教育创新工程项目(JYCX_11_2020);黑龙江省省属本科高校基本科研业务费项目(东北石油大学引导性创新基金)(2020YDL-11)

作者简介:赵崇志(1986-),男,工程师,研究方向为石油行业数字化转型;通信作者:田枫(1980-),男,博士,教授,CCF会员(51268M),研究方向为计算机视觉。

位作业人员,减少违规作业并保证油田的生产安全,智能视频监控因其24小时监控、准确率高、响应速度快的优势应运而生。油田场景下,作业人员对阀门的操作频繁,一旦操作失误将会造成不可估计的损失,危害作业人员生命安全。基于此,需要研究阀门的错误操作识别方法,从而实现作业人员对阀门操作的智能监控,保障工作人员的安全和减少事故的发生。

目前,油田阀门的规范操作识别的研究尚未相关报道,阀门规范操作的识别虽然属于行为识别,但是一方面,行为识别拥有其独特的特点,包括处理难度大、数据集样本少、检测速度慢、过于依赖于场景等问题,也使得行为识别与目标检测相比发展缓慢。当前行为识别领域也涌现出了各种不同的技术方法,比如Cao等人^[1]提出了用于行为识别的最新OpenPose模型,与RMPE(regional multi-person pose estimation)模型^[2]和Mask R-CNN(region-based convolutional neural network)模型^[3]相比,该网络模型的优势在于大大缩短了关键点的检测周期,并保持检测精度基本不变。但是,它缺点是参数共享率低、冗余度高、时间成本长、模型太大,仍不能解决实际应用需求。

另一方面,基于深度学习的目标检测技术在识别行为方面展现了其优越性,邢予权使用基于手势识别目标检测的监控场景下抽烟识别,取得了不错的识别效果^[4]。王春辉、王全民^[5]使用改进的YOLOv3识别学生的“抬头”“低头”“说话”3种姿态,除满足识别的精度外,速度也能满足实时检测。邓永川^[6]使用DSC-Net+YOLOv3的漏油检测方法,检测精度达到了96.22%,满足油田企业检测要求。基于以上两方面,该文选用深度学习目标检测方法解决阀门的规范操作识别,不仅保证了识别的精确度,并且速度也达到实时的视频智能监控。

目前,目标检测算法在深度学习领域可以分为两类,一是双阶段目标检测算法,其中包括最具有代表性的Faster R-CNN(faster region-CNN)^[7];二是单阶段目标检测算法,代表的有YOLO(you only look once)^[8]、单次多框检测(single shot multibox detector, SSD)^[9]算法。2016年,Redmon J^[8]提出了Yolo算法,可以一次性地识别图片内多个物品的类别和位置,实现了端到端的图像识别,运行速度快。此后,YOLO系列算法^[10-12]逐步更新换代,最新提出的YOLOv5s相比于Faster R-CNN不仅速度快而且参数小,仅27 MB。

基于YOLOv5s的网络结构,该文提出了一种采用YOLOv5s深度学习算法的作业现场阀门规范操作识别的模型。自主拍摄阀门操作的数据集,并利用数据增强方式扩充数据集,然后将处理好的数据输入

YOLOv5s网络训练,保存最优的训练权重数据,并用作测试集测试。测试结果表明,所用算法模型的检测速度和准确性都达到了油田作业现场的要求,智能监控系统部署后,阀门安全操作监管防护效果很好。

1 作业现场阀门规范操作识别方法

1.1 阀门识别原理

开关阀门是采油工生产中经常进行的最基本的操作内容,正确开关阀门不仅仅是确保人身安全的保障措施,还可以使阀门始终处于良好的使用状态,确保设备的安全运行。正确操作阀门时,人体应位于阀门手轮的侧面,随阀门位置高低不同,身体适当弯曲,头部在阀门侧上方30 cm左右。错误操作最危险的一种就是人体占位不同,应禁止人体直面阀门,这样会危及操作人员的人身安全。正确操作和错误操作阀门如图1所示。二者差别很大,可以使用目标检测进行识别。



图1 阀门操作示意图

油田作业现场工作人员阀门的操作识别主要由数据集获取、模型训练和错误操作报警三部分完成。数据集获取是使用海康威视摄像头采集工作人员阀门操作的视频,接下来将获取到的视频利用python代码逐帧截取成图片,然后对图片进行数据增强和归一化处理;模型训练是指采用YOLOv5s网络训练处理好的数据集,且将训练好的模型用于智能监控视频系统,进行实时的检测任务;错误操作报警是指出现异常操作就进行报警提示,并将异常出现后的第一帧开始进行图片保存和视频保存,以备后续调查备案。识别流程如图2所示。

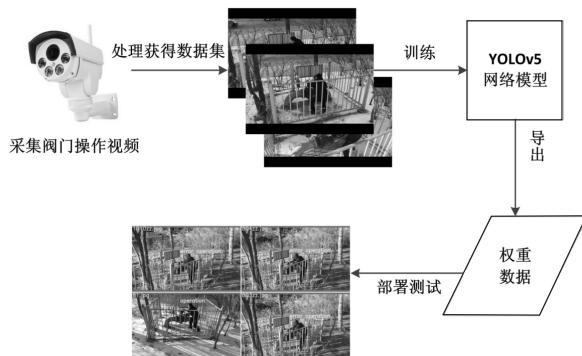


图2 识别流程

1.2 YOLOv5s 网络框架

Ultralytics 团队在 2020 年提出了 YOLOv5 模型,其目标检测速度高达 0.007 s,每秒可以处理 140 帧,同时结构更为小巧,有四个不同的版本^[13],使得在模型的快速部署上具有极强的优势,只需简单调节一些

网络的深度和宽度即可。与传统的网络相比^[14],YOLOv5 模型将网络分为检测区和候选区两个阶段,是一种直接利用回归网络实现目标检测的算法^[3]。其在大部分目标检测任务中同时兼顾了速度和准确度。YOLOv5s 的网络结构如图 3 所示。

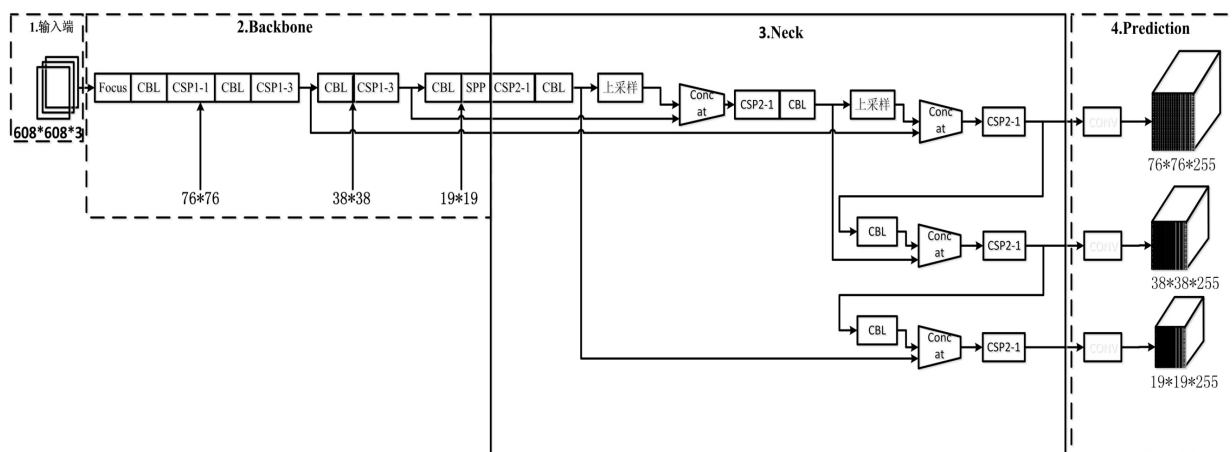


图 3 YOLOv5s 网络结构

YOLOv5s 目标检测网络以 Focus 结构为主干,而 Focus 结构,在 YOLOv3 和 YOLOv4 中并没有这个结构,其中比较关键的是切片操作,并且 YOLOv5s 有两种 CSP 模块,一种是 CSP1_X 结构,其应用于 Backbone 主干网络,另一种是 CSP2_X 结构,其应用于 Neck 网络中,二者结合后进一步增强整个网络的学习能力;YOLOv5s 网络颈部是采用 PAN 和 FPN 的结构,FPA 是自上而下的,将高层特征通过上采样和低层特征作融合得到用于预测的特征图;PAN 是一个自底向上的特征金字塔,其自底向上进行强定位特征的提取。二者相互促进,从不同的主干层对不同的检测层进行参数聚合;YOLOv5s 头部继续沿用 YOLOv4 的检测头,

使用颈部的特征来预测边界框和类标签;YOLOv5s 使用一个在线的数据增强策略增加输入图像可变性,丰富图像特征信息,包括 Mosaic 数据增强和自适应锚框计算以及自适应图片缩放。设计的目标检测模型可以实现更高的鲁棒性,同时,采用标签平滑和学习率余弦退火衰减等技术来优化网络训练过程^[15]。

1.3 损失函数

YOLOv5s 网络训练中的损失函数由三部分组成:边界框回归损失 L_{ciou} 、分类的损失 L_{class} 、置信度损失 L_{conf} 。当某个边界框内没有目标时,只计算置信度损失 L_{conf} ,若存在一个目标时,三个损失同时计算。三种损失函数数学表达式为^[16]:

$$L_{ciou} = \sum_{i=0}^{S^2} \sum_{j=0}^B I_{i,j}^{obj} \left[1 - IOU + \frac{\rho^2(b, b^{gt})}{c^2} \right] + \frac{\frac{16}{\pi^4} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^4}{1 - IOU + \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2}$$

$$L_{conf} = - \sum_{i=0}^{S^2} \sum_{j=0}^B I_{i,j}^{obj} \left[C_i^j \log(C_i^j) + (1 - C_i^j) \log(1 - C_i^j) \right] - \lambda \sum_{i=0}^{S^2} \sum_{j=0}^B I_{i,j}^{nobj} \left[C_i^j \log(C_i^j) + (1 - C_i^j) \log(1 - C_i^j) \right]$$

$$L_{class} = - \sum_{i=0}^{S^2} \sum_{c \in \text{classes}} I_{i,j}^{obj} \left[P_i^c \log(P_i^c) + (1 - P_i^c) \log(1 - P_i^c) \right]$$

$$\text{Loss} = L_{ciou} + L_{conf} + L_{class}$$

上述公式中: S^2 、 B 为特征图尺度与先验框; λ 为权重系数;如果第 i 个网格的第 j 先验框处有目标, $I_{i,j}^{obj}$ 、 $I_{i,j}^{nobj}$ 分别取 1 和 0,若没有目标,则相反,分别取 0 和 1; $\rho(\cdot)$ 为欧几里得距离; c 为实际边框和预测边框之间的闭合区域的对角线距离; b 、 w 、 h 分别为预测边框的中心坐标及其宽度和高度; b^{gt} 、 c^{gt} 、 h^{gt} 分别为实际边框的中心坐标及其宽度和高度; C_i^j 、 C_i^j 为预测

边框和被标注边框的置信度; P_i^j 、 P_i^j 为预测边框和标注边框的类别概率。

通过交叉熵的方法计算置信度损失与分类损失,通过 L_{ciou} 损失函数计算边界框回归损失,相比传统的损失函数, L_{ciou} 损失函数更多地关注预测边框和实际边框二者之间的位置和大小以及距离,进而提高了预测边框的准确度和速度。

2 数据预处理

2.1 数据采集

该文使用的是一个自制数据集,其被用在在作业现场阀门规范操作识别模型上,使用油田现场专业的海康威视摄像头拍摄工人对阀门操作,拍摄视频剪辑后,通过 python 代码处理视频获取每帧的照片,按 1:1 的正样本和负样本选取 2 000 张阀门操作照片。

由于收集到的数据量很小,为了避免模型的过度拟合,提高模型的泛化能力,使用了随机角度旋转和随机亮度增强的方法以及大小不等的黑色遮挡策略扩充数据样本,因此训练后的网络模型能够满足栏杆遮挡,不同姿势的阀门操作识别需求,其数据增强示例见图 4。

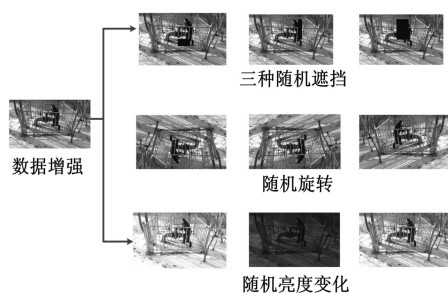


图 4 数据增强

通过数据增强获得阀门操作图片 5 000 张,包含随机黑块遮挡、随机方向选择、随机亮度变化下阀门操作图片。图片数据中有两类阀门操作图片,总计 2 500 张正确操作图片,2 500 错误操作图片,可以保证数据的丰富性,以及正负样本的平衡。

2.2 数据标注

在训练 YOLOv5s 模型之前,应该对处理后的数据集进行标签标注。在实验中,使用了标签标注工具 LabelImage,设置正确操作阀门标签为 operation,错误操作阀门标签为 error operation,标签标注后被保存为 xml 数格式的文件,两个类别阀门操作的图片标签标记信息如图 5 所示。

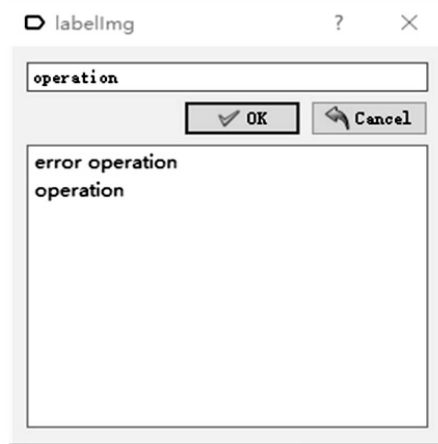


图 5 数据标注

标注后保存的 xml 标签内容见表 1。表格中每一行的数据代表一个标注边框的位置信息,其中 x 、 y 、 w 、 h 分别代表标注框中心点的 x 轴及 y 轴方向的坐标位置以及标注框的宽和高。实验进行模型训练,随机划分 75% 的数据集作为训练集,25% 作为测试集。

表 1 数据标注

标签类别	X	Y	W	H
operation	0.54	0.36	0.14	0.51
error operation	0.52	0.53	0.28	0.50

3 试验与分析

3.1 试验配置

训练平台的参数配置越高,图形计算能力更强,模型训练的效果会更好。为了使模型达到良好的训练效果,在 GPU 上对 YOLOv5s 网络进行训练,YOLOv5s 网络模型的训练平台的配置信息如表 2 所示。

表 2 训练平台配置

名称	配置信息
操作系统	Ubuntu 18.04
系统显存	10 G
系统内存	16 G
使用的 GPU	NVIDIA GeForce RTX 3080
GPU 的加速库	CUDA 10.0, CUDNN 7.6.0

处理好的自制数据集传入神经网络进行训练,训练设置的超参数如表 3 所示。

表 3 超参数设置

学习率	图像大小	迭代次数	权重衰减	动量值
0.01	640 * 640	300	0.000 5	0.9

3.2 模型评估

损失值是指由损失函数计算出的样本预测值和真实值之间的误差。损失值越小,预测效果越高。在网络训练结束后,会生成完整的训练日志,取日志中的部分数据进行折线图绘制,得到损失值折线图如图 6 所示。

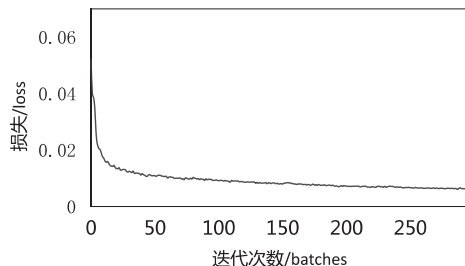


图 6 损失值曲线

从图 6 损失值曲线可以看出,随着迭代次数的增加,损失值开始逐渐减少。经过 300 次的迭代,损失值稳定在 0.01,这证明网络模型的训练效果良好。

3.3 检测效果

为了验证模型的检测效果,在 200 张测试集上测试,得到如表 4 所示结果,结果表明 YOLOv5s 无论是在准确率还是 mAP 上,YOLOv5s 表现能力更好。

表 4 不同算法对比

模型名称	准确率	mAP/%	训练时间/h
YOLOv5s	93.18	96.18	25

为验证基于 YOLOv5s 的作业人员阀门规范操作模型的可行性和实时检测能力,所以将模型嵌入到油田智能视频监控系统,使用海康威视相机在线获取作业人员阀门操作视频流,网络模型对获得的视频流进行实时检测,检测的效果见图 7。

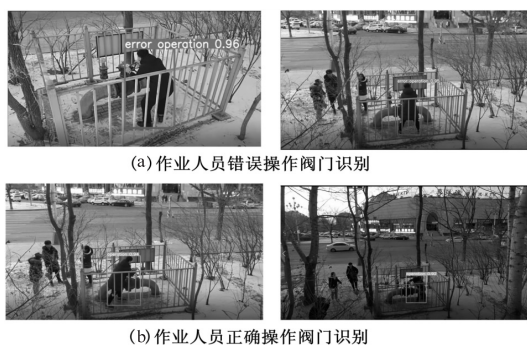


图 7 检测效果图

由图 7 可知,该模型完成了对作业人员错误操作阀门的检测任务,检测准确率达到 93%,能达到实时检测,能够满足智能监控系统下的作业人员阀门操作的识别。

4 结束语

为了解决人工的监控手段不能对阀门操作错误行为及时进行检测和发出预警,使得危险事故不能在较短时间之内得到有效妥善的处理的问题,文中提出了一种基于 YOLOv5s 网络的作业现场阀门规范操作识别的模型,自己制作阀门操作数据集,通过图像增强方法解决阀门周围的栏杆遮挡问题,保证了数据集的丰富性。训练后的模型实际检测效果很好,检测准确率达到了 93%,检测速度能达到实时的效果。

参考文献:

- [1] CAO ZHE, HIDALGO G, SIMON T, et al. OpenPose: real-time multi-person 2D pose estimation using part affinity fields[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 43(1): 172-186.
- [2] FANG H, XIE S, TAI Y W, et al. RMPE: regional multi-person pose estimation[C]//Proceedings of the 2017 IEEE international conference on computer vision. Venice, Italy: IEEE, 2017: 2353-2362.

- [3] HE K, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[C]//Proceedings of the 2017 IEEE international conference on computer vision. Venice, Italy: IEEE, 2017: 2980-2988.
- [4] 邢予权. 基于手势识别的监控场景下抽烟检测[D]. 杭州: 浙江工业大学, 2020.
- [5] 王春辉, 王全民. 基于改进 YOLOv3 网络的学生特定行为识别[J]. 计算机与现代化, 2020(7): 90-96.
- [6] 邓永川. 基于图像处理的油田漏油检测方法设计与研究[D]. 成都: 电子科技大学, 2020.
- [7] REN S, HE K, 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.
- [8] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the 2016 IEEE conference on computer vision and pattern recognition. Las Vegas, NV, USA: IEEE, 2016: 779-788.
- [9] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//Proceedings of the 2016 European conference on computer vision. Amsterdam, The Netherlands: Springer, 2016: 21-37.
- [10] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Honolulu, HI, USA: IEEE, 2017: 6517-6525.
- [11] ZHAO Lique, LI Shuaiyang. Object detection algorithm based on improved YOLOv3[J]. Electronics, 2020, 9: 537-546.
- [12] DENG H, CHENG J, LIU T, et al. Research on iron surface crack detection algorithm based on improved YOLOv4 network[C]//Proceedings of the 2nd international conference on artificial intelligence and computer science (AICS 2020). Hangzhou, China: Hubei Zhongke Institute of Geology and Environment Technology, 2020: 577-585.
- [13] 许德刚, 王露, 李凡. 深度学习的典型目标检测算法研究综述[J]. 计算机工程与应用, 2021, 57(8): 10-25.
- [14] GIRSHICK R. Fast R-CNN[C]//Proceedings of the IEEE international conference on computer vision. Santiago: IEEE, 2015: 1440-1448.
- [15] BOCHKOVSKIY A, WANG Chien-yao, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23) [2020-09-16]. <https://arxiv.org/abs/2004.10934>.
- [16] 许筹艺, 秦贵和, 孙铭会, 等. 基于改进的 ResNeSt 驾驶员头部状态分类算法[J]. 吉林大学学报: 工学版, 2021, 51(2): 704-711.