

基于监控视频流的手持探针探测位置检测算法的设计

张建鹏¹, 徐云¹, 杨承翰¹, 林奇洲²

(1. 浙江理工大学 信息科学与工程学院, 杭州 310018;

2. 浙江理工大学 机械工程学院, 杭州 310018)

摘要: 针对监控视频流开展生产流程的有效性检测, 存在计算量大、耗时长等问题; 根据采集的手持探针探测的视频流图像, 构建数据集, 训练人工手持探针的探测模型; 采用 KNN 算法分析前后帧的监控视频流, 实现视频流图像前景和背景的分离; 利用人工手持探针模型实时提取监控视频中的探针, 获得手持探针的前景图像; 提出基于像素搜索的手持探针的位置探测算法实现视频图像中人工手持探针探测位置的点推算, 并对比理论应检测的真实位置, 从而判断手持探针检测的有效性; 工厂监控视频流实际测试结果表明, 设计的基于监控视频流的手持探针探测位置检测算法的平均准确率约 93.26%, 召回率约 81.11%, F_1 值约 86.76%, 检测速度约 9.66 fps/s, 能够实现工厂监控视频流中人工手持探针的有效性检测。

关键词: 背景分离; 探针检测; KNN 算法; 机器视觉; 有效性检测

Design for Position Detection Algorithm of Hand-Held Probe Based on Video Surveillance

ZHANG Jianpeng¹, XU Yun¹, YANG Chenghan¹, LIN Qizhou²

(1. School of Information Science and Engineering, Zhejiang Sci-Tech University, Hangzhou 310018, China;

2. School of Mechanical Engineering, Zhejiang Sci-Tech University, Hangzhou 310018, China)

Abstract: Aiming at the problem of large calculation and long time consumption for the validity detection for the production process of video surveillance. The dataset is established based on the collected detection video surveillance images for the hand-held probe, and the detection model for the hand-held probe is trained; The K-Nearest-Neighbors (KNN) algorithm is applied to analyze the continuous frames of the video surveillance and separate the foreground and background for the video frame, the artificial handheld probe model is used to extract the probe from surveillance videos in real time, and then the foreground image of the hand-held probe can be obtained; Based on the foreground image of the hand-held probe, the algorithm based on the pixel search is proposed to calculate the detection position of the hand-held probe. Then, compared with the actual position need to be theoretically detected for the probe, the validity for detection of the hand-held probe will be evaluated. The experiment is carried out in the factory surveillance video, and the results show that the average accuracy for the validity detection of the hand-held probe in the video surveillance is about 93.26%, the recall rate is about 81.11%, and the F_1 score is about 86.76% with the detection speed of 9.66 fps/s, which achieves the validity detection of the hand-held probe in the factory surveillance video.

Keywords: background separation; probe detection; KNN algorithm; machine vision; validity detection

0 引言

智能制造水平直接关乎着一个国家制造业发展的水平^[1]。目前, 计算机视觉、工业控制数据、物联网技术等极大化地推动了制造业的发展。制造业数字化转型、网络化协同、智能化变革是制造强国建设的主攻方向。随着智能工厂中监控设备的普及, 针对工厂监控视频开展生产流程的有效性检测, 提高工厂的生产质量、生产效能、减少

资源消耗是推动产业技术变革和优化升级的必经之路^[2-7]。

空调外机结构复杂, 出厂前需人工探测空调外机中铜管焊接的可靠性, 从而保障压缩机壳体与储液器的密封性。该过程中, 人工检测的不确定性将导致空调外机中铜管焊点的检测存在漏检、误检等情况的发生。利用工厂的监控视频设备, 开展基于监控视频流的手持探针探测位置检测算法的研究, 利用计算机实时分析人工手持探针探测

收稿日期: 2023-08-01; 修回日期: 2023-09-07。

基金项目: 国家自然科学基金(62203393); 浙江省自然科学基金(LQ20F030019)。

作者简介: 张建鹏(2000-), 男, 硕士研究生。

通讯作者: 徐云(1987-), 女, 博士研究生, 讲师。

引用格式: 张建鹏, 徐云, 杨承翰, 等. 基于监控视频流的手持探针探测位置检测算法的设计[J]. 计算机测量与控制, 2024, 32(3): 112-117.

过程的有效性, 对提高工厂的生产效能具有重要的研究意义^[8-10]。

目前, 国内外研究学者针对工厂监控视频流开展了丰富和广泛的研究工作^[11-14]。2018年, Eduardo Cermeño 等人基于背景建模系统, 添加机器学习全局特征模块, 设计了一种能学习和识别事件的入侵者探测器, 实现了监控视频中恶意入侵的探测与警报^[15]。2019年, 金肖莹针对工厂员工安全帽有效佩戴的问题提出一种基于 Faster-RCNN 的检测方法^[16], 实现了工厂监控视频中工作人员安全帽佩戴的实时检测与预警。2020年, 王海州采用 KNN 算法实现了监控视频流中飞机目标的检测与跟踪^[17], 并利用树莓派操作系统进行了实验测试, 并证明了该方法的有效性; Artur Wilkowski 等人使用背景分离算法丰富数据集, 进行了训练数据特征提取, 构建了两级检测器确保检测系统的速度和检测的可靠性, 实现了基于监控视频流的数据自动处理与目标探测^[18]。2021年, 李林等人提出一种基于 K-SSD-F 的融合背景分离和目标检测算法^[19], 对视频前帧的时空特征进行提取, 利用标注好的数据训练 SSD 模型, 并对视频进行监测, 提取视频的静态特征, 提高东亚蝗虫计数的准确率。利用补帧算法识别漏计数的帧, 最终实现了监控视频中东亚蝗虫的实时、连续自动监测和计数。2022年, 郎松等人提出一种改进的 YOLOv5 算法^[20], 应用融合卷积注意力机制模块的 YOLOv5 算法, 实现了反射片靶标的广角镜头识别与监测, 应用目标自动照准算法实现反射片靶标中心的长焦镜头精确照准, 进而实现目标点位置坐标的全自动测量。实验测试证明了该方法能够实现了反射片靶标的广角镜头识别与检测, 完成目标点的全自动测量。目前, 基于监控视频的探测技术已逐渐广泛应用于工业领域^[21], 但工厂监控视频因生产环境的光照强度不均、探测算法的实时性和有效性等因素均极大化地影响了生产过程中检测的精度。因此, 针对工厂监控视频流, 开展有效信息检测算法的设计研究, 对提升产品生产质量、提高工厂生产效率具有极其重要的研究意义。

本文以空调外机气密性探测为研究背景, 利用人工智能技术和机器视觉技术对空调外机生产企业的工厂监控视频流进行研究, 采集监控视频流中人工手持探针的图像并对其进行前景和背景分离研究, 构建基于像素搜索的探针位置探测模型, 对比理论应检测的真实位置, 从而判断人工手持探针检测的有效性, 保障出厂前空调外机铜管焊点气密性检测的有效性。

1 基于机器视觉的探针识别算法

空调外机出厂前, 机身上铜管的焊缝(如图1(a)所示)需进行气密性检测, 从而保证空调外机压缩机壳体与储液器的密封性。一般而言, 需人工对其焊缝点进行手工探测。该过程中监控视频捕获的图像如图1(b)所示。

图1(a)可以看出, 外机待检测的焊缝数量多、分布零散、背景复杂, 且部分待测焊点相对较密集, 极大化地



图1 空调外机焊缝检测

增加了手持探针探测的难度。图1(b)给出的工厂监控视频流捕获的图像中, 工人手持探针进行探测时, 探针细长, 图像背景复杂, 探针特征难以直接提取, 不利于开展手持探针探测位置的检测。

针对上述问题, 本文设计了基于前景背景分类模型和目标检测相融合的手持探针探测位置检测方案, 探测流程如图2所示。

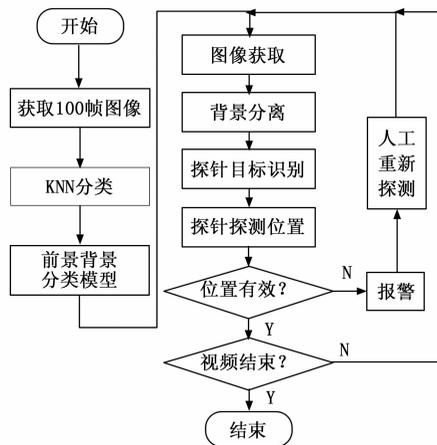


图2 探针有效性探测算法流程

图2所示的探测算法流程主要可分为4大部分: 1) 前景背景分类模型; 2) 探针目标识别部分; 3) 探针位置推算部分; 4) 探测位置有效性辨别。

在前景背景分类模型中, 需要提取视频流中前100帧的图像, 然后利用KNN分类模型对该100帧图像进行训练与分析, KNN分类模型具备精度高的优势, 利用该前景背景分类模型对实时采集的监控视频中前后帧视频图像进行前景和背景的分离, 捕获工人手持探针的前景图像, 并将其应用于后期探针的目标识别。

探针目标识别部分, 旨在实现视频图像中人工手持探针的识别。基于YOLOv5算法对工厂监控视频流采集的图像中人工手持探针进行目标跟踪与识别。该算法主要包括输入端、主干网络层、颈部层以及预测端层。根据工厂监控视频流采集的图像, 利用主干网络层、颈部层实现人工手持探测的图像特征加强, 从而最终实现监控视频流采集的图像中人工手持探测的识别与跟踪。

探针位置推算部分是实现探针有效性辨别的核心。本文提出基于像素搜索的探针位置推算的算法。该算法基于像素搜索的方式对视频图像中人工手持探针探测位置进行搜索与推算,在此基础上利用最小二乘算法对搜索与推算获得的探针位置进行拟合,在保证拟合相关系数的前提下,对监控视频流采集的图像中手持探针的长度进行补偿,从而实现监控视频流采集的图像中人工手持探针探测位置的判断。

探测位置有效性辨别根据当前监控视频流采集的图像中人工手持探针探测的位置,结合实际待检测的位置进行判别,求解手持探针检测的实际位置与待检测位置的交并比,从而实现人工手持探针探测位置有效性的辨别。

下面针对上述各部分的模型和算法进行详细阐述与分析。

1.1 前景和背景分类模型

前景和背景的分旨在实现对监控视频流采集图像进行前景和背景的分,从而获得人工手持探针探测的前景图像。根据监控视频流中前后帧视频图像的时空特征,去除复杂的工作环境背景,即可获得工人手持探针探测的前景图像。本文监控视频流中的每帧视频图像像素为 $1\ 920 \times 1\ 080$ 。

1) 前景、背景分类模型训练。提取当前视频图像的前 100 帧视频图像,利用 KNN 分类算法对每个像素的特征做前景、背景训练,从而获得视频图像中前景和背景的分模型。

2) 前景提取。利用该前景、背景的分模型对当前视频图像进行前景提取,设当前第 i 帧视频图像中第 n 个像素点像素为:

$$X_i^n = \{x_i^n(1), x_i^n(2), x_i^n(3)\} \quad (1)$$

第 $i-100$ 到第 $i-1$ 帧视频图像中第 j 帧视频图像中第 n 个像素点像素为:

$$Y_i^n = \{y_i^n(1), y_i^n(2), y_i^n(3)\} \quad (2)$$

分别计算第 i 帧图像与前 100 帧图像中第 n 个像素点

X_i^n 和 Y_j^n 的欧式距离:

$$d_j = \sqrt{\sum_{k=1}^3 [x_i^n(k) - y_j^n(k)]^2}, (j = i - 100, i - 1) \quad (3)$$

根据式 (3) 计算的 100 帧视频图像的欧式距离,寻找与第 i 帧像素点最近邻的像素点。应用训练得到的前景背景分类模型对第 i 帧图像中所有的点进行前景和背景的分类,获得当前第 i 帧图像中的前景图像,并将其应用于后续视频图像中的人工手持探针的识别。

1.2 探针目标识别

针对工厂监控视频中的人工手持探针,开展探针探测位置的有效性检测。基于 Yolov5 算法对工厂视频中的探针探测的位置进行跟踪与识别,从而判断工人进行空调外机铜管焊缝气密性检测的有效性。基于 Yolov5 算法的人工手持探针跟踪与识别算法构架如图 3 所示。

基于 Yolov5 的算法具备快速高效、准确率高、模型轻量化、易于部署和使用的优点。图 3 中的算法构架主要分为 4 个部分。

1) 输入端:采集监控视频中的图像,并将其输入至该算法的输入端并进行输入端的前期处理。输入端进行前期处理的算法主要包括 Mosaic 数据增强算法、自适应锚框计算、自适应图像缩放算法等预处理。其中, Mosaic 数据增强算法通过随机缩放、随机裁剪、随机排布的方式对输入的图像进行拼接以增强图像数据,增加训练集的多样性和难度,同时提高检测模型的鲁棒性和泛化能力,降低过拟合的风险;在每次图像训练时,自适应锚框计算用于自适应地计算不同训练集中的最佳锚框值,提高检测的精度和鲁棒性;自适应图像缩放算法是将原始采集的图像自适应地缩放到一个标准尺寸,再输入到检测网络的主干网络中进行处理。经过上述图像的预处理,还能够有效减少工厂监控视频因探测环境光照亮度不足对图像采集的影响。

2) 主干网络:由下采样层、卷积层、瓶颈层和池化层组成。对于输入为 $608 \times 608 \times 3$ 的图像而言,经过下采样层的操作,可将其变为 $304 \times 304 \times 12$ 的特征图。再经过卷

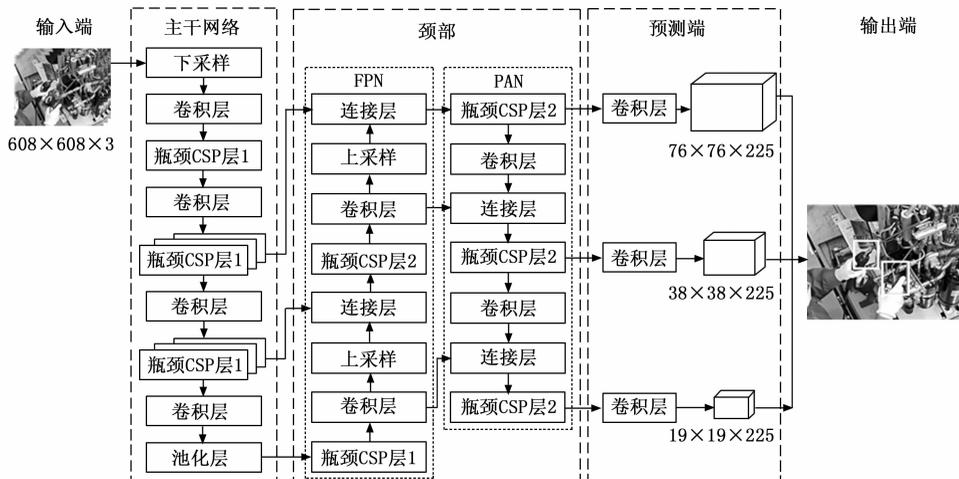


图 3 基于 Yolov5 算法的人工手持探针跟踪与识别算法构架图

积层, 利用 32 个卷积核的操作, 将其变为 $304 \times 304 \times 32$ 的特征图。瓶颈层为 YOLOv5 算法的重要部分, 将输入特征图分为两部分, 一部分进行子网络处理, 先使用卷积层将输入的特征图像进行压缩, 然后再进行卷积处理, 从而提取出相对较少的高层次的特征, 另外一部分直接输入颈部层进行处理。YOLOv5 算法中随着网络的不断加深, 可增加网络特征提取和特征融合能力, 因此图 3 中主干网络层采用了多层网络。瓶颈 CSP 层 1 的作用可有效减少网络参数和计算量, 提高特征提取的效率。主干网络旨在实现输入图像的特征提取。

3) 颈部层: 由特征金字塔网络 (FPN, feature pyramid network) 和聚合网络 (PAN, path aggregation network) 组成。FPN 为自顶向下的设计, 通过上采样操作, 将高层的特征信息和底层的特征信息进行融合, 计算出预测的特征图。FPN 中顶部信息流须通过主干网络逐层往下传递, 由于层数较多将导致计算量较大。PAN 层在 FPN 层的基础上提取网络内特征层次的结果, 引入自底向上的路径, 经过自顶向下的特征融合后, 再进行自底向上的特征融合, 实现底层位置信息的深层传递, 从而增加多个尺度的定位能力, 将底层的特征信息和高层特征进行融合, 最后输出预测的特征图给预测端进行预测。颈部层利用瓶颈 CSP 层 2 加强网络特征的融合能力。对步骤 2) 中提取的有效特征进行上采样和下采样的特征融合, 实现视频图像中特征的加强提取。

4) 预测端: 通过卷积层输出特征后, 利用 CIOU Loss 损失函数计算预测框和真实框的损失值, 采用加权非极大值抑制方法针对多目标框进行筛选, 实现手持探针的识别与跟踪。

1.3 探针位置探测推算

基于分离的前景图像以及目标识别捕获的手持探针的位置, 开展人工手持探针位置探测的推算分析, 本文提出基于像素搜索的探针探测位置推算算法, 主要由以下 5 部分组成:

1) 设手持探针的初始位置为 $O(x, y)$, 记背景分离后的前景图像中探针宽度所占像素为 a 。根据探针的方向, 分别沿 x 轴和 y 轴搜索 5 个检测框, 每个检测框的像素为 $a \times a$, 共 9 个检测框, 如图 4 所示。

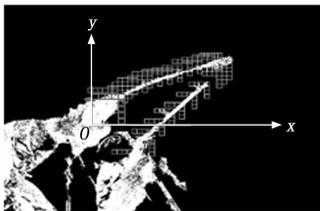


图 4 探针位置探测搜索示意图

2) 根据搜索的 9 个检测框, 分别计算各个检测框中的像素之和 P_n :

$$P_n = \sum_{i=1}^a \sum_{i=1}^a P_x(i), (n = 1 \sim 9) \quad (4)$$

其中: $P_x(i)$ 为背景分离后各检测框中第 i 个点的像素值。基于像素搜索的方式开展探针探测位置的推算。

3) 判断 P_n 是否超过设定阈值, 如果未超过, 则在起点位置的基础上增加 a 个像素, 沿探针所在的方向建立 x 轴和 y 轴继续搜索。重复 5 次, 仍然未超过设定的阈值, 则跳出搜索转向下一帧图像处理; 如果超过设定的阈值, 判断该检测框对应的原视频图像中的颜色与探针的颜色是否相符, 若相符, 则认定该点为探针上的点, 否则在起点位置的基础上增加 a 个像素, 并分别在探针所在的方向沿 x 轴和 y 轴继续搜索。

4) 当该点判定为探针上的点时, 则以当前探针上的点为起点 $O(x, y)$, 返回 1) 继续搜索并进行探针探测位置的推算。

5) 采用最小二乘法对获得的探针探测位置上所有点进行拟合, 保证拟合的相关系数 r 值大于 0.98。根据探针标准长度, 对拟合算法获得的探针长度进行补偿, 从而最终判定人工手持探针探测的位置。

1.4 探测位置辨别

在探针位置探测推算的基础上, 结合实际待检测的位置进行判断, 求解探针位置和待检测位置的交并比 (IOU, Intersection over Union), 从而判断人工手持探针探测的位置是否有效。

$$IOU = \frac{TP}{TP + FP + FN} \quad (5)$$

其中: TP 为被判定为正样本, 实际也是正样本。 FP 为被判定为正样本, 但实际是负样本。 FN 为被判定为负样本, 但实际为正样本。

2 实验测试

2.1 实验测试环境与说明

利用 python3.7 编写探针探测有效性检测算法的代码, 选用的主板为 Inter (R) Core (TM) i7-11700, 显卡为 RTX3070Ti, 内存为 16 GB, 操作系统为 Windows10 的计算机开展实验测试的研究。根据工厂提供的实时监控视频, 选取其中一段监控视频数据, 每秒提取一帧视频图像, 共提取包含探针的 1 645 张图像作为数据集。随机划分该数据集, 其中训练集、验证集和测试集的比例为 8: 1: 1。采用精确度、召回率、F1 值对本文设计的人工手持探针的有效性检测算法进行综合评判。

根据人工手持探针有效性探测算法的流程, 针对算法的测试效果进行测试, 主要包括前景背景分类模型的效果测试, 探针探测位置的有效性测试。1) 前景背景分类模型的效果测试用于对监控视频流中采集的图像进行基于 KNN 算法的前景和背景的分离, 并对分离的效果进行测试和分析。2) 探针探测位置的有效性测试用于对监控视频流中手持探针探测位置的有效性进行测试与分析。

2.2 前景背景分类模型的效果测试

分析监控视频中前后帧图像的时序特征，去除复杂工作环境背景，采用 KNN 算法，获得前景图像。截取视频中的一帧图像，以图 5 (a) 为例，设前景图像为白色，背景图像为黑色，对其原视频图像进行前景、背景分离，效果如图 5 (b) 所示。

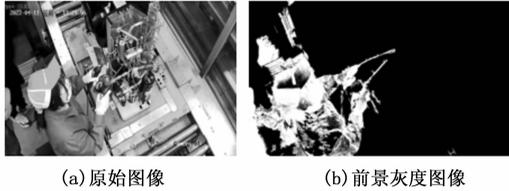


图 5 背景分离效果图

图 5 (b) 可以看出，监控视频中的前景和背景图像已充分分离，能够获得人工手持探针的前景灰度图像。

2.3 探针探测位置的有效性测试

探针探测位置的有效性测试步骤如下：1) 每秒采集监控视频中的图像，基于每帧视频图像建立包含人工手持探针的数据集，按照 Yolov5 算法的要求对数据集进行探针的标注；2) 随机划分数据集，其中训练集、验证集和测试集的占比为 8: 1: 1；3) 利用 Yolov5 算法对训练集进行训练，利用验证集对训练的模型进行效果验证，从而获得最优的人工手持探针探测位置的检测模型；4) 在测试集中，采用训练的人工手持探针探测位置的检测模型进行测试，并输出探针探测位置的辨识结果。

根据训练获得的人工手持探针探测位置的检测模型，对工厂监控视频中工人手持探针探测的位置进行实时检测。记工人手持探针的检测框为 Tool_box，目标概率记为 Tool_score，其阈值设为 0.5。筛选所有目标概率大于 0.5 的检测框，并且筛选目标概率最大的两个检测框，区分左右手检测框，并分别记为 Tool_box_L 和 Tool_box-R，如图 6 所示。



图 6 手持探针检测效果图

图 6 可以看出，工人手持的探针能够有效识别出该目标，为后续探针探测位置的检测奠定了基础。

根据分离的前景图像以及基于 Yolov5 算法的手持探针检测获得的检测框 Tool_box_L 和 Tool_box_R，开展基于像素搜索的探针探测位置的检测。根据 1.3 节给出的

手持探针探测位置的推算流程，图 7 给出了人工手持探针探测位置推算的示意图。

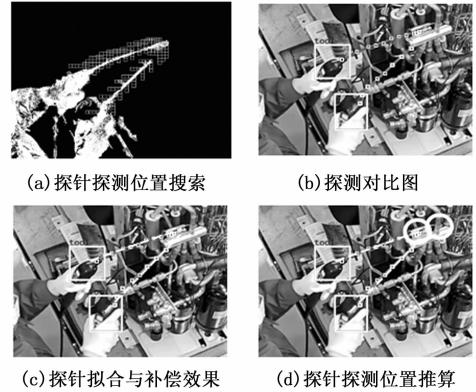


图 7 探针探测位置推算过程图

图 7 (a) 为利用分离的前景图像，开展基于像素搜索的探针探测位置的推算过程图。图 7 (b) 为搜索的检测框与原始图像中的探针进行对比的效果，其中小方块的标注为探测的探针上的点。图 7 (c) 为采用最小二乘法对检测到的探针进行拟合和补偿的效果图。图 7 (d) 为根据拟合与补偿后的探针推算出的探针检测的位置推算结果，图中采用圆圈标注。

根据手持探针检测的位置，对标实际待检测的位置，求解探针检测的位置与实际待检测位置的交并比，若交并比大于 50%，则视为检测正确，否则视为检测错误，最终实现人工手持探针检测位置有效性的辨别。本文采用精确度、召回率、 F_1 值对设计的手持探针有效性检测算法进行综合评判。召回率旨在评价探针探测位置为正确的待探测位置的概率， F_1 值是用于评价探针探测位置模型精确度的一种指标。精确度、召回率越高越好， F_1 值同时考虑精确度和召回率，让两者同时达到最高，取得平衡， F_1 值越大越好。

针对工厂提供的监控视频，计算机利用计时器对探针有效性探测算法进行测试，记录一帧视频图像从采集到有效性检测完成所需的时间，计算获得探针有效性探测算法的平均检测速度为 9.66 fps/s。此外，对手持探针有效性的检测进行了 10 组测试，检测结果如表 1 所示。

表 1 可以看出，采用本文所提的手持探针有效性检测算法的检测精确度达到 93.26%，召回率约 81.11%， F_1 值约 86.76%，实现了工厂监控视频流中人工手持探针位置的有效性检测，检测精度和检测速度均能够满足工厂实际的生产需求。

3 实验数据分析

工厂实际监控视频的测试结果可以看出：1) 采用工厂监控视频流训练的基于 KNN 的前景背景分类模型能够有效地实现视频流中采集图像的前景和背景分离，获得人工手持探针的前景灰度图像，为后续开展探针探测位置的检测奠

表1 探针有效性检测结果统计表

| 组序 | 精确度/% | 召回率/% | F ₁ 值/% |
|----|-------|-------|--------------------|
| 1 | 93.29 | 81.11 | 86.78 |
| 2 | 93.33 | 81.67 | 87.11 |
| 3 | 93.27 | 80.83 | 86.61 |
| 4 | 93.59 | 81.11 | 86.90 |
| 5 | 92.99 | 81.11 | 86.65 |
| 6 | 93.31 | 81.39 | 86.94 |
| 7 | 93.31 | 81.39 | 86.94 |
| 8 | 92.97 | 80.83 | 86.48 |
| 9 | 93.89 | 81.11 | 87.03 |
| 10 | 92.65 | 80.56 | 86.18 |
| 平均 | 93.26 | 81.11 | 86.76 |

定了良好的基础; 2) 探针探测位置的有效性测试中, 根据分离的人工手持探针的前景灰度图像, 基于像素搜索的探针探测位置推算算法, 结合最小二乘拟合算法能够有效实现探针的拟合和补偿, 在此基础上获得人工手持探针探测位置的推算; 3) 基于人工手持探针探测位置推算的结果, 结合实际待检测的位置进行分析, 最终实现了工厂监控视频中手持探针探测位置的有效性检测。

4 结束语

以空调出厂前空调外机中铜管焊缝的气密性检测为研究背景, 开展基于工厂监控视频中人工手持探针探测位置的有效性检测算法的设计研究。首先, 采用KNN算法对监控视频中前后帧图像的时空特征进行提取, 实现视频图像中前景和背景图像的分离; 其次, 基于Yolov5算法开展探针探测位置的检测模型研究, 提出基于像素搜索的探针探测位置推算算法, 结合最小二乘拟合算法实现探针检测位置的有效推算; 最后, 对标实际待检位置, 判断工人手持探针探测的有效性。工况实测测试结果表明, 本文设计的手持探针检测算法的检测速度可达9.66 fps/s, 平均检测精确度约为93.26%, 召回率约81.11%, F1值约86.76%, 能够满足工厂实际应用的需求。

参考文献:

- [1] WANG B, TAO F, FANG X, et al. Smart manufacturing and intelligent manufacturing: a comparative review [J]. *Engineering*, 2021, 7 (6): 738-757.
- [2] 李东洁, 李东阁, 杨柳. 卷积神经网络在动态手势跟踪中的应用 [J]. *计算机科学与探索*, 2020, 14 (5): 841-847.
- [3] YANG Y, SARKIS R A, EL ATRACHE R, et al. Video-based detection of generalized tonic-clonic seizures using deep learning [J]. *IEEE Journal of Biomedical and Health Informatics*, 2021, 25 (8): 2997-3008.
- [4] MOHTAVIPOUR S M, SAEIDI M, ARABSORKHI A. A multi-stream CNN for deep violence detection in video sequences

- using handcrafted features [J]. *The Visual Computer*, 2022, 38 (6): 2057-2072.
- [5] 张兴旺, 刘小雄, 林传健, 等. 基于Tiny-YOLOV3的无人机地面目标跟踪算法设计 [J]. *计算机测量与控制*, 2021, 29 (2): 76-81.
- [6] 王殿伟, 杨旭, 翰鹏飞, 等. 复杂背景下全景视频运动小目标检测算法 [J]. *控制与决策*, 2021, 36 (1): 249-256.
- [7] 王迪聪, 白晨帅, 邬开俊. 基于深度学习的视频目标检测综述 [J]. *计算机科学与探索*, 2021, 15 (9): 1563-1577.
- [8] HAO Y, LI J, WANG N, et al. Spatiotemporal consistency-enhanced network for video anomaly detection [J]. *Pattern Recognition*, 2022, 121: 108232.
- [9] 宋艳艳, 谭励, 马子豪, 等. 改进YOLOv3算法的视频目标检测 [J]. *计算机科学与探索*, 2021, 15 (1): 163-172.
- [10] 刘艺, 李蒙蒙, 郑奇斌, 等. 视频目标跟踪算法综述 [J]. *计算机科学与探索*, 2022, 16 (7): 1504-1515.
- [11] FATHY C, SALEH S N. Integrating deep learning-based iot and fog computing with software-defined networking for detecting weapons in video surveillance systems [J]. *Sensors*, 2022, 22 (14): 5075.
- [12] ASAD M, JIANG H, YANG J, et al. Multi-Stream 3D latent feature clustering for abnormality detection in videos [J]. *Applied Intelligence*, 2022, 52 (1): 1126-1143.
- [13] SADIQ M, MASOOD S, PAL O. FD-YOLOv5: A fuzzy image enhancement based robust object detection model for safety helmet detection [J]. *International Journal of Fuzzy Systems*, 2022, 24 (5): 2600-2616.
- [14] CHANG Y, TU Z, XIE W, et al. Video anomaly detection with spatio-temporal dissociation [J]. *Pattern Recognition*, 2022, 122: 108213.
- [15] CERMENO A E, Perez A, SIGUENZA J A. Intelligent video surveillance beyond robust background modeling [J]. *Expert Systems with Applications*, 2017, 2018 (91): 138-149.
- [16] 金肖莹. 工厂监控视频中的安全帽检测 [D]. 武汉: 华中科技大学, 2019.
- [17] 王海州. 基于树莓派的飞机跟踪系统设计与实现 [D]. 沈阳: 沈阳航空航天大学, 2020.
- [18] WILKOWSKI A, STEFANCZYK M, KASPRZAK W. Training data extraction and object detection in surveillance scenario [J]. *Sensors*, 2020, 20 (9): 363-372.
- [19] 李林, 柏召, 刁磊, 等. 基于K-SSD-F的东亚飞蝗视频检测与计数方法 [J]. *农业机械学报*, 2021, 52 (s1): 261-267.
- [20] 郎松, 曹选, 张艳微, 等. 融合改进YOLOv5算法的图像全站仪全自动测量方法 [J]. *仪器仪表学报*, 2022, 43 (5): 120-127.
- [21] GUO Y, LU Y, LIU R W. Lightweight deep network-enabled real-time low-visibility enhancement for promoting vessel detection in maritime video surveillance [J]. *The Journal of Navigation*, 2022, 75 (1): 230-250.