

## 基于目标检测的VOC泄漏区域识别技术研究

杨承霖, 陈海永, 岳学智, 李华曜, 邓立齐, 郭东歌, 王海超, 刘欢

**Research on VOC leakage area identification based on target detection**

引用本文:

杨承霖, 陈海永, 岳学智, 李华曜, 邓立齐, 郭东歌, 王海超, 刘欢. 基于目标检测的VOC泄漏区域识别技术研究[J]. *激光技术*, 2024, 48(6): 922–930.

YANG Chenglin, CHEN Haiyong, YUE Xuezhi, et al. Research on VOC leakage area identification based on target detection[J]. *Chinese Journal of Liquid Crystals and Displays*, 2024, 48(6): 922–930.

---

## 您可能感兴趣的其他文章

### 1. 基于改进YOLOv4算法的水果识别检测研究

引用本文: 裴瑞景, 王硕, 王华英. 基于改进YOLOv4算法的水果识别检测研究[J]. 激光技术, 2023, 47(3): 400–406.

### 2. 激光成像引信的目标识别方法研究

引用本文: 周瑜, 贺伟. 激光成像引信的目标识别方法研究[J]. 激光技术, 2023, 47(2): 267–272.

### 3. 稀疏训练指纹库融合MMPSO-ELM室内可见光定位

引用本文: 张慧颖, 梁誉, 卢宇希, 等. 稀疏训练指纹库融合MMPSO-ELM室内可见光定位[J]. 激光技术, 2022, 46(6): 788–795.

### 4. 基于双邻域对比度的红外小目标检测算法

引用本文: 朱金辉, 张宝华, 谷宇, 等. 基于双邻域对比度的红外小目标检测算法[J]. 激光技术, 2021, 45(6): 794–798.

### 5. 改进型卷积神经网络焊点缺陷识别算法研究

引用本文: 常颖, 常大俊. 改进型卷积神经网络焊点缺陷识别算法研究[J]. 激光技术, 2020, 44(6): 779–783.

文章编号: 1001-3806(2024)06-0922-09

## 基于目标检测的 VOC 泄漏区域识别技术研究

杨承霖<sup>1</sup>, 陈海永<sup>1,2\*</sup>, 岳学智<sup>1</sup>, 李华曜<sup>2</sup>, 邓立齐<sup>1</sup>, 郭东歌<sup>1</sup>, 王海超<sup>1</sup>, 刘欢<sup>2\*</sup>

(1. 汉威科技集团股份有限公司, 郑州 450001, 中国; 2. 华中科技大学 集成电路学院 湖北光谷实验室, 武汉 430074, 中国)

**摘要:** 为了解决红外气体成像仪在挥发性有机化合物(VOC)泄漏识别中存在的误识别率高、漏检率高、算法执行效率低以及模型泛化能力差等问题,提出了一种基于运动特征增强的VOC泄漏区识别方法。采用视频序列投影变化率统计的方法确定视频稳定性判定阈值,提取稳定状态下运动背景和运动前景;采用优化线性拉伸的方法对运动前景进行特征增强和异常值过滤;将运动前景与原始帧进行图像融合,并利用目标检测算法进行VOC泄漏区域识别;通过模型预训练和迁移学习的方法,以烟雾数据集和少量VOC泄漏数据集进行了识别模型训练,并将模型迁移至RK3588S嵌入式开发板上进行了执行效率测试。结果表明,该算法在交并比为0.5的情况下,平均精度均值为0.88;交并比在0.5~0.95范围内,平均精度均值为0.51,单帧平均识别时间为49 ms,具有较高的识别精度和识别效率,能够满足实时监测需求。本文中的算法能够保持稳定的模型性能且具有一定的抗干扰能力,为VOC泄漏识别提供了一定的参考。

**关键词:** 图像处理; 泄漏区域识别; 目标检测; 挥发性有机化合物泄漏; 红外气体成像仪

**中图分类号:** TN219      **文献标志码:** A      **doi:**10.7510/jgis.issn.1001-3806.2024.06.019

## Research on VOC leakage area identification based on target detection

YANG Chenglin<sup>1</sup>, CHEN Haiyong<sup>1,2\*</sup>, YUE Xuezhi<sup>1</sup>, LI Huayao<sup>2</sup>, DENG Liqi<sup>1</sup>,  
GUO Dongge<sup>1</sup>, WANG Haichao<sup>1</sup>, LIU Huan<sup>2\*</sup>

(1. Hanwei Electronics Group Corporation, Zhengzhou 450001, China; 2. Optics Valley Laboratory, School of Integrated Circuits, Huazhong University of Science and Technology, Wuhan 430074, China)

**Abstract:** To solve the problem of high misidentification rate, high missed detection rate, low algorithm execution efficiency, and poor model generalization ability in the recognition of volatile organic compound (VOC) leakage area of infrared gas imager, a VOC leakage area recognition method based on motion feature enhancement was proposed. The video stability threshold was determined by using the statistical method of projection change rate of video sequence, and the moving background and moving foreground were extracted under stable state. Optimized linear stretching was used to perform feature enhancement and outlier filtering on the moving foreground. The motion foreground was fused with the original frame, and VOC leakage area identification was performed using the target detection algorithm. Through the method of model pre-training and transfer learning, the smoke dataset and a small amount of VOC leakage dataset were used to train the recognition model, and the model was transferred to the RK3588S embedded development board for execution efficiency test. Experimental results show that the mean average precision of the proposed algorithm is 0.88 when the intersection over union ratio is 0.5, and the mean average precision is 0.51 when the intersection over union ratio ranges from 0.5 to 0.95. The average recognition time of a single frame is 49 ms, which has high recognition accuracy and recognition efficiency, and can meet the requirements of real-time monitoring. The algorithm in this article can maintain stable model performance and has certain anti-interference capabilities providing some reference for VOC leak identification.

**Key words:** image processing; leakage area identification; target detection; volatile organic compounds leakage; infrared gas imager

## 0 引言

挥发性有机化合物(volatile organic compound,

基金项目:华中科技大学湖北光谷实验室创新科研项目(OVL2023ZD003);郑州市重大科技专项基金资助项目(2021KJZX0061)

\*通信作者:hai@hust.edu.cn; huan@hust.edu.cn

收稿日期:2024-01-12;修回日期:2024-03-14

VOC)是一类在常温条件下易挥发的有机化合物,在工业生产、交通运输、建筑材料以及室内气体环境中均广泛分布,会对环境和健康造成重大危害,其光化学反应会形成臭氧和其它有毒大气污染物,对环境造成严重污染。长期暴露于高浓度VOC会导致眼鼻喉刺激、呼吸道疾病、头痛、恶心、疲劳等健康问题<sup>[1-2]</sup>,甚至可能引发潜在的癌症风险<sup>[3]</sup>。因此,监测VOC泄漏对于保护人体健康和环境质量至关重要,在早期阶段发现并

迅速控制泄漏,能够有效预防长期暴露可能带来的危害,同时也符合相关法规和标准要求。

随着气体探测技术的不断发展,VOC 泄漏监测已由非成像模式逐步转向成像模式。比如,Bertin 公司的 SecondSight MS<sup>[4]</sup>产品通过精心设计多个滤光片,依据气体在不同波段的吸收特性,实现对特定气体的精确识别。FLIR 公司的光学气体图像摄像机(optical gas imaging, OGI<sup>[5]</sup>)则采用窄带滤光片和斯特林制冷系统,使其能够可视化气体在红外吸收波段的分布情况,通过差分图像技术,该设备能够有效识别可疑的 VOC 泄漏区域,进一步提高了监测的准确性和可靠性。

相较于传统的非成像气体检测设备,光学气体成像仪不仅在作业效率和准确定位泄漏点方面具有显著优势,还能远距离评估泄漏区域的大小,帮助巡检人员精确掌握泄漏趋势,降低人员暴露于 VOC 泄漏区域的风险。随着红外图像分析技术的不断发展,越来越多的学者开始将其与红外成像技术相结合,以提高 VOC 泄漏的自动化识别效果。较为典型的分析方法包括帧差法<sup>[6-7]</sup>、光流法<sup>[8-9]</sup>和背景建模法<sup>[10]</sup>,这些方法的核心思想是利用运动目标的运动特征来识别目标。然而,由于 VOC 的泄漏通常以不稳定的模式移动,使得上述方法难以区分气体运动区域和非气体运动区域。为减少非气体运动区域干扰、提升识别效果,研究者们从背景建模算法,运动区域的轮廓、面积、纹理和历史特征等方面进行了深入研究,例如,WANG 等人<sup>[11]</sup>将多幅帧间差分和背景差分后图像进行了融合处理,经形态学滤波后得到了疑似 VOC 泄漏区域;LIU 等人<sup>[12]</sup>则基于支持向量机分类算法,设计了一种基于轮廓特征分类的泄漏 VOC 识别方法。此外,多特征融合识别的方法也备受关注,例如,HONG 等人<sup>[13]</sup>通过优化高斯背景建模,并引入气体不规则性、气体连通区域面积和移动方向等特征,显著提高了高浓度 VOC 泄漏区域的识别效果;BADAWI 等人<sup>[14]</sup>则利用 VOC 动态纹理区域的空间和时间结构特征,结合视频图像深度神经网络和时间序列神经网络的方法,实现了从视频中快速识别 VOC 泄漏区域。

上述研究中,VOC 泄漏区域识别算法多集中在优化背景建模、改善 VOC 疑似区域分类模型和多特征融合等方面,在 VOC 泄漏识别中存在误识别率高、漏检率高、算法执行效率低以及模型泛化能力差等问题,并且在 VOC 泄漏区域的特征提取方面缺少自主学习能力。因此,本文作者提出了一种基于增强运动区域特征并利用目标识别算法识别 VOC 泄漏区域的算法,包括视频稳定性判定、运动背景提取、运动前景特征增强、图像融合与 VOC 泄漏区域识别等 5 个步骤。通过

分析相邻视频帧的灰度投影矩阵变化率来判断视频是否稳定,减少干扰因素的影响;利用高斯混合模型背景建模方法提取出静止区域作为运动背景;通过减去运动背景,提取出视频中的运动区域,并进行图像增强,突出细节特征,将增强后的运动图像与原始帧进行融合,并利用运动目标检测算法识别 VOC 泄漏区域,实现对泄漏区域的识别。该算法的优势在于结合了视频稳定性判定和混合高斯背景建模,提高了算法的执行效率和背景建模质量;运动图像增强处理在保留原始帧数据特征的同时,融合了运动区域的细节特征,提高了 VOC 泄漏区域的识别精度和可靠性。

## 1 算法介绍

### 1.1 视频稳定性判定

本文中将灰度投影矩阵变化率作为判定视频稳定性的指标。较小的投影变化率表示视频处于稳定状态,较高的投影变化率表示视频处于非稳定状态,或者视频中有快速移动物体的干扰。该算法计算效率高,能够有效过滤非稳定状态的视频数据,减少镜头晃动、人员快速移动等因素对结果的干扰。下面介绍灰度投影矩阵变化率计算方法。

1.1.1 灰度投影矩阵计算 分别计算视频第  $k$  帧行方向 1 维灰度投影矩阵  $\mathbf{X}_k$ 、列方向 1 维灰度投影矩阵  $\mathbf{Y}_k$  和上一帧行方向 1 维灰度投影矩阵  $\mathbf{X}_{k-1}$ 、列方向 1 维灰度投影矩阵  $\mathbf{Y}_{k-1}$ ,计算公式见下:

$$X_k(x) = \frac{1}{w} \sum_{y=1}^w f_k(x, y) \quad (1)$$

$$Y_k(y) = \frac{1}{h} \sum_{x=1}^h f_k(x, y) \quad (2)$$

$$X_{k-1}(x) = \frac{1}{w} \sum_{y=1}^w f_{k-1}(x, y) \quad (3)$$

$$Y_{k-1}(y) = \frac{1}{h} \sum_{x=1}^h f_{k-1}(x, y) \quad (4)$$

式中: $h$  为视频帧的行数; $w$  为视频帧的列数; $f_k(x, y)$ 、 $f_{k-1}(x, y)$  分别为视频第  $k$ 、 $k-1$  帧图像的 2 维灰度信息; $X_k(x)$ 、 $Y_k(y)$  为第  $k$  帧图像第  $x$  行和第  $y$  列的灰度投影值; $X_{k-1}(x)$ 、 $Y_{k-1}(y)$  为第  $k-1$  帧图像第  $x$  行和第  $y$  列的灰度投影值。

1.1.2 灰度投影矩阵标准差计算 第  $k$  帧行方向的灰度投影标准差矩阵  $\mathbf{S}_{x,k}$ 、列方向灰度投影标准差矩阵  $\mathbf{S}_{y,k}$  和上一帧行方向灰度投影标准差矩阵  $\mathbf{S}_{x,k-1}$ 、列方向灰度投影标准差矩阵  $\mathbf{S}_{y,k-1}$ ,计算公式为:

$$S_{x,k}(x) = \sqrt{\frac{1}{w} \sum_{y=1}^w (f_k(x, y) - X_k(x))^2} \quad (5)$$

$$S_{x,k}(y) = \sqrt{\frac{1}{h} \sum_{x=1}^h (f_k(x,y) - Y_k(y))^2} \quad (6)$$

$$S_{x,k-1}(x) = \sqrt{\frac{1}{w} \sum_{y=1}^w (f_{k-1}(x,y) - X_{k-1}(x))^2} \quad (7)$$

$$S_{y,k-1}(y) = \sqrt{\frac{1}{h} \sum_{x=1}^h (f_{k-1}(x,y) - Y_{k-1}(y))^2} \quad (8)$$

式中:  $S_{x,k}(y)$ 、 $S_{y,k}(y)$  分别为第  $k$  帧图像第  $x$  行和第  $y$  列的灰度投影标准差;  $S_{x,k-1}(x)$ 、 $S_{y,k-1}(y)$  分别为第  $k-1$  帧图像第  $x$  行和第  $y$  列的灰度投影标准差。

1.1.3 灰度投影矩阵变化率计算 利用下式分别计算行、列方向灰度投影矩阵变化率  $X_c$  和  $Y_c$ :

$$X_c = \frac{\sum_{x=1}^h |X_{k-1}(x) - X_k(x)|}{\frac{1}{w} \sum S_{x,k-1}} \quad (9)$$

$$Y_c = \frac{\sum_{y=1}^w |Y_{k-1}(y) - Y_k(y)|}{\frac{1}{h} \sum S_{y,k-1}} \quad (10)$$

式中:  $(\sum S_{x,k-1})/w$  和  $(\sum S_{y,k-1})/h$  分别为视频第  $k-1$  帧行方向和列方向灰度投影标准差的均值。如果  $X_c$  和  $Y_c$  都较低, 则说明第  $k$  帧图像处于稳定视频序列中。

## 1.2 运动背景提取

本文中基于背景减法实现对运动前景的提取。常用背景模型的构建方法有中值建模法<sup>[15]</sup>、均值建模法<sup>[16]</sup>和基于混合高斯模型的建模<sup>[17]</sup>方法。下面介绍采用混合高斯背景建模算法作为背景建模和前景提取方法的步骤。

1.2.1 模型参数初始化 在混合高斯背景建模算法中, 选取连续  $N$  帧图像, 用于初始化高斯分布的权重、均值和方差:

$$\boldsymbol{\omega}_{m,0} = \frac{1}{m} \quad (11)$$

$$\boldsymbol{\mu}_{m,0} = \frac{1}{N} \sum_{i=0}^N \mathbf{f}(i) \quad (12)$$

$$\boldsymbol{\sigma}_{m,0}^2 = \frac{1}{N} \sum_{i=0}^N (\mathbf{f}(i) - \boldsymbol{\mu}_{m,0})^2 \quad (13)$$

式中:  $m$  是高斯分量的数量; 初始权重矩阵  $\boldsymbol{\omega}_{m,0}$  采用平均分配的方式;  $\boldsymbol{\mu}_{m,0}$  表示高斯分布的初始化均值矩阵;  $\boldsymbol{\sigma}_{m,0}^2$  表示各高斯分布的初始化方差矩阵;  $\mathbf{f}(i)$  表示第  $i$  帧图像的灰度矩阵。

1.2.2 背景建模和更新 对于新输入第  $k$  帧图像位置  $(x, y)$  处的灰度值  $f(x, y, k)$ , 利用式(14)计算其与当前  $m$  个混合高斯模型匹配度, 若满足匹配条件则标

记  $p_{m,k} = 1$ ; 否则, 标记  $p_{m,k} = 0$ 。

$$|f(x, y, k) - \boldsymbol{\mu}_{m,k-1}(x, y, k-1)| \leq c \times \sigma_{m,k-1}(x, y, k-1) \quad (14)$$

式中:  $c \in [2, 3]$  是一个经验常数。随后, 根据匹配标记结果利用式(15)更新权重参数, 如果存在  $p_{m,k} = 1$ , 采用式(16)和式(17)分别更新第  $m$  个高斯模型的均值和方差。

$$\boldsymbol{\omega}_{m,k}(x, y, k) = (1 - \alpha) \cdot \boldsymbol{\omega}_{m,k-1}(x, y, k-1) + \alpha \cdot p_{m,k} \quad (15)$$

$$\boldsymbol{\mu}_{m,k}(x, y, k) = (1 - \alpha) \cdot \boldsymbol{\mu}_{m,k-1}(x, y, k-1) + \alpha \cdot f(x, y, k) \quad (16)$$

$$\boldsymbol{\sigma}_{m,k}^2(x, y, k) = (1 - \alpha) \cdot \boldsymbol{\sigma}_{m,k-1}^2(x, y, k-1) + \alpha \cdot p_{m,k} \cdot [f(x, y, k) - \boldsymbol{\mu}_{m,k-1}(x, y, k-1)]^2 \quad (17)$$

式中:  $\alpha \in [0, 1]$  是一个学习率参数, 用于更新各高斯模型的参数。若  $f(x, y, k)$  与所有模型都未得到匹配, 则用  $f(x, y, k)$  取代最后一个高斯模型对应位置的均值, 并重新为该模型的  $(x, y)$  处分配一个较大的方差和较小的权重。

为进一步计算第  $k$  帧视频数据的背景模型, 对前  $k-1$  帧已完成参数更新的模型进行如下调整: 利用式(18)对视频帧各像素位置的模型权重进行归一化处理; 随后, 按由大到小的排序规则对  $\boldsymbol{\omega}_{m,k}(x, y, k)/\boldsymbol{\sigma}_{m,k}^2(x, y, k)$  的计算结果对各高斯模型进行排序, 利用式(19)选择前  $B$  个模型作为背景分布,  $\tau$  为权重阈值; 最后, 利用式(20)对第  $k$  帧视频数据  $f(x, y, k)$  各像素点的值和前  $B$  个背景模型进行匹配判断, 若匹配到一个模型满足匹配要求, 则认为该像素点为背景信息, 否则为前景信息。

$$\boldsymbol{\omega}_{m,k}(x, y, k) = \frac{\boldsymbol{\omega}_{m,k}(x, y, k)}{\sum_{i=1}^m \boldsymbol{\omega}_{i,k}(x, y, k)} \quad (18)$$

$$B = \operatorname{argmin}_m \left( \sum_{m=1}^M \boldsymbol{\omega}_{m,k}(x, y) \geq \tau \right) \quad (19)$$

$$|f(x, y, k) - \boldsymbol{\mu}_{m,k}(x, y, k)| \leq c \times \sigma_{m,k}(x, y, k) \quad (20)$$

## 1.3 运动前景提取

帧差法<sup>[18]</sup>是一种常用的运动目标检测算法, 它通过分析相邻帧像素灰度值的差异, 能够精确地检测出视频中的运动目标。第 1.3.1 节中采用连续两帧图像与背景图像进行灰度值差值分析, 取两者的均值作为帧差结果; 然后将平均值与预设阈值进行比较, 如果平均值超过阈值 0, 则相应区域被判定为运动区域。这种方法能够有效地检测出运动目标, 并且在实际操作中具有较高的准确性和可靠性。

### 1.3.1 计算帧差

$$d(x, y, k-1) = |f(x, y, k-1) - b(x, y, k-1)| \quad (21)$$

$$d(x, y, k) = |f(x, y, k) - b(x, y, k)| \quad (22)$$

$$g(x, y, k) = 0.5(d(x, y, k - 1) + d(x, y, k)) \quad (23)$$

式中: $f(x, y, k - 1)$ 和 $f(x, y, k)$ 为影像第 $k - 1$ 和第 $k$ 帧图像处的灰度值; $b(x, y, k)$ 为第 $k$ 帧的运动背景; $d(x, y, k - 1)$ 和 $d(x, y, k)$ 为第 $k - 1$ 和第 $k$ 帧的帧差结果; $g(x, y, k)$ 为第 $k$ 帧求均值后的帧差结果,也就是运动前景。

**1.3.2 图像增强** 为提高图像质量、突出图像细节,常采线性拉伸<sup>[19]</sup>、细节滤波<sup>[20]</sup>、多尺度变换<sup>[21]</sup>等算法对红外图像进行特征增强。其中,线性拉伸是一种常见的图像增强技术,用于改善图像的对比度和亮度。它基于简单的线性变换,将图像像素的灰度值从原始范围映射到更广的范围,这样可以使原本较暗的区域变得更明亮,同时保留图像中的细节。

为解决 VOC 泄漏运动前景灰度值低、存在异常噪声、对比度不明显的问题,本文作者提出了一种优化的线性拉伸算法来提升图像质量。首先计算图像的相对累积直方图,查找最小百分比和最大百分比相对应的数据值,将其分别标记为 $a$ 值和 $b$ 值;接着利用最小调整百分比 0.1 和最大调整百分比 0.5 更新 $a$ 值(见式(24))和 $b$ 值(见式(25)),得到新的 $a'$ 和 $b'$ ,将小于 $a'$ 和大于的 $b'$ 值标记为最小噪声和最大噪声;然后使用公式(见式(26))对图像进行拉伸和噪声去除。

$$a' = a - (b - a) \times 0.1 \quad (24)$$

$$b' = b + (b - a) \times 0.5 \quad (25)$$

$$E(x, y, k) = \begin{cases} 0, & (g(x, y, k) < a') \\ 255, & (g(x, y, k) > b') \\ 255 \frac{g(x, y, k) - a'}{(b' - a')}, & (a' \leq g(x, y, k) \leq b') \end{cases} \quad (26)$$

式中: $g(x, y, k)$ 为原始运动前景; $E(x, y, k)$ 为第 $k$ 帧图像经过拉伸和图像增强后的运动前景。

#### 1.4 图像融合

为利用视频的场景特征,利用式(27)将经过特征增强的运动前景与原始图像进行加权融合,使动态运动区域的特征与原始图像结合在一起。该方法在提升运动区域形状、纹理和位置特征的同时,保留了图像的上下文信息,有助于目标检测算法更好地理解和识别运动区域。

$$A(x, y, k) = \alpha \times f(x, y, k) + \beta \times E(x, y, k) \quad (27)$$

式中: $A(x, y, k)$ 为融合后像素的灰度值; $\alpha$ 和 $\beta$ 为权重系数,本文中 $\alpha$ 和 $\beta$ 均为 0.5; $f(x, y, k)$ 为原始图像的灰度值; $E(x, y, k)$ 为第 $k$ 帧图像经过拉伸和图像增强后的运动前景。

#### 1.5 模型选择

在充分考虑后续识别任务对多尺度、时效性和准确性的要求后,本文中选择了目标检测模型 YOLOv7。YOLOv7<sup>[22]</sup>是一种基于深度学习的单阶段目标检测网络,具备高性能和实时检测的特点。该模型采用多尺度特征提取的骨干网结构,能够专注于从深层获取目标的高级描述信息,在多尺度目标检测中表现出色。

模型的 Head 部分通过并行结构和多级特征融合技术,如空间金字塔池化(spatial pyramid pooling, SPP)、空间金字塔卷积(spatial pyramid convolution, SPC)、上采样、最大池化以及重复卷积等,进一步细化特征并有效整合了不同分辨率的特征图。另外, YOLOv7 的 head 结构通过多尺度预测能够灵敏捕获不同尺寸的物体,准确预测它们的位置和类别。

## 2 实验与分析

### 2.1 模拟实验

模拟实验在设有罐状设备的工厂内进行,以原油气作为 VOC 泄漏源。泄漏源右侧为林木,实验当天的天气条件为微风,树枝有明显晃动,符合移动干扰的特征。实验现场的情况如图 1 所示。



图 1 实验地点

Fig. 1 Experiment site

实验数据采集设备为红外气体成像仪,该成像仪采用了浙江超晶晟锐光电有限公司的Ⅱ类超晶格中波制冷型机芯,型号为 406B,实物如图 2 所示,其核心参数如表 1 所示。



图 2 中波制冷型机芯实物图

Fig. 2 Mid-wave integrated detector cooler assembly

表 1 红外气体成像仪核心参数

Table 1 Core parameters of infrared gas imaging device

item	technical indicators
resolution	640 pixel × 512 pixel
pixel size	15 $\mu\text{m}$
spectral response range	3.2 $\mu\text{m}$ ~ 3.5 $\mu\text{m}$
noiseequivalent temperature difference	22 mK
cooling method	stirling refrigeration machine
data type	8 bit/14 bit

为便于数据分析,对成像仪采集的原始数据进行了处理,包括将 YUV 格式的二进制数据转换为通用的数据格式、调整分辨率和帧率设置,最终生成了一段分辨率为 640 pixel × 512 pixel、帧率为 25 的高质量视频。

## 2.2 视频稳定性阈值获取

实验中本文作者对原始视频序列进行了投影矩阵变化率的统计分析,并得到了如图 3 所示的时序序列。根据该序列的分布和统计特性,采取了一定的处理措施。首先计算投影矩阵变化率序列的标准差,并将两倍标准差的数值定义为异常视频帧的判定阈值,剔除了大于该阈值的视频数据,这样可以过滤掉异常值或异常波动,保留相对平稳的数据;接着对剩余数据取均值,以得到一个稳定的投影矩阵变化率阈值。在本实验中,该阈值设置为 3,即当投影矩阵的变化率不大于 3 时,认为视频处于稳定状态,可以进行后续处理和识别操作,而一旦超过阈值,则判定图像处于运动状态或有快速移动物体的干扰,直接返回原始图像。通过以上处理和阈值设定,可以筛选出稳定的视频帧,以确保后续的处理和识别操作的准确性和可靠性。

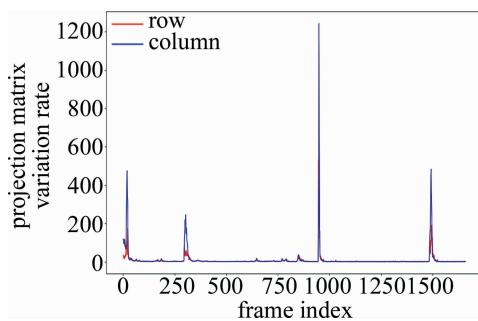


图 3 投影矩阵变化率

Fig. 3 Change rate of video sequence projection matrix

## 2.3 VOC 泄漏数据集获取

基于第 2.2 节中获取的稳定性判定阈值,对稳定状态的视频序列进行后续背景建模、运动前景提取、运动前景图像增强和图像融合处理。在处理后的图像中,人工筛选出具有明显泄漏气体特征的图像,作为 VOC 泄漏样本,本实验中共获取 VOC 泄漏样本 323 个。

在数据集标准方面,本文作者采用数据标注工具对样本进行标注。样本数据标注前,将已融合图像的

上下两端各填充 64 行 0 值,使其符合目标检测数据集对图像尺寸 640 pixel × 640 pixel 的要求。部分已标注数据集图 4 所示。

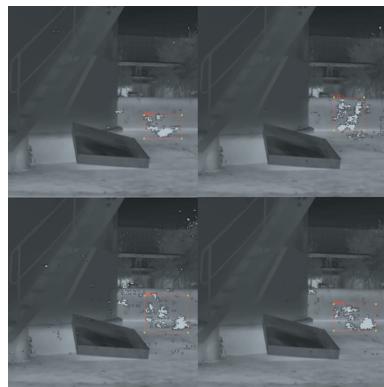


图 4 部分已标注数据集

Fig. 4 Part of the annotated dataset

## 2.4 模型训练与精度评价

本实验中使用 PyTorch 2.1.1 框架来训练基于 YOLOv7 算法的目标检测网络模型。由于真实样本有限,为了提高模型的泛化能力并降低小样本数据带来的过拟合问题,作者采用了迁移学习中模型预训练的策略。在预训练阶段,使用了与 VOC 泄漏数据集具有相似特征的烟雾数据集进行训练,以获得一个具有良好烟雾特征提取能力的模型。这一策略旨在使模型具备一定的先验知识,有助于模型更好地处理和提取 VOC 泄漏数据集中的目标信息。在获得预训练模型后,作者又使用真实的 VOC 泄漏数据集进行模型微调,使模型的参数能够更好地适应实际的检测任务,提高模型的准确性和泛化能力。

预训练数据集为网络公开数据集<sup>[23]</sup>,共 15000 个,使用前将图片尺寸统一调整为 640 pixel × 640 pixel,训练、验证和测试集比例为 8 : 1 : 1,即 80% 的数据用于训练模型,10% 的数据用于验证模型的性能和调优,另外 10% 的数据用于测试模型的泛化能力和评估最终性能。真实 VOC 数据集 323 个,图片尺寸 640 pixel × 640 pixel,数据集划分方法与预训练数据集划分方法一致。模型训练时的部分参数设置如表 2 所示。

表 2 模型训练关键参数

Table 2 Key parameters of model training

item	smoke recognition model	VOC recognition model
number of categories	1	1
anchor point	[10 13 16 30 33 23] [30 61 62 45 59 119] [116 90 156 198 373 326]	[10 13 16 30 33 23] [30 61 62 45 59 119] [116 90 156 198 373 326]
initial learning rate	0.01	0.01
batch size	20	10
epochs	10	100
number of samples	15000	323

本实验中,模型评估遵循训练与验证两个阶段:在

训练阶段,使用检测边框损失、物体置信度损失、准确率和召回率来衡量模型对目标边界框的识别能力;在验证阶段,使用检测边框损失、物体置信度损失、平均精度均值 (mean average precision, mAP) mAP@ 0.5 和 mAP@ 0.5 ~ 0.95 指标综合评价模型性能。其中, mAP@ 0.5 是指固定交并比 (intersection over union, IOU) 阈值为 0.5 情况下的 mAP, mAP@ 0.5 ~ 0.95 是指 IOU 阈值从 0.5 ~ 0.95 时的平均精度均值,更全面

地评估目标检测模型的性能。图 5 和图 6 分别为烟雾和 VOC 识别模型训练过程的可视化结果。从训练结果可以看出,模型迁移对 VOC 泄漏识别模型的训练效率提升明显。

为评价模型精度,以测试集为输入数据,以 mAP@ 0.5 和 mAP@ 0.5 ~ 0.95 为评价指标对模型精度进行评价,结果表明在 IOU 为 0.5 时, mAP 为 0.88;当 IOU 阈值从 0.5 ~ 0.95 时, mAP 为 0.51。

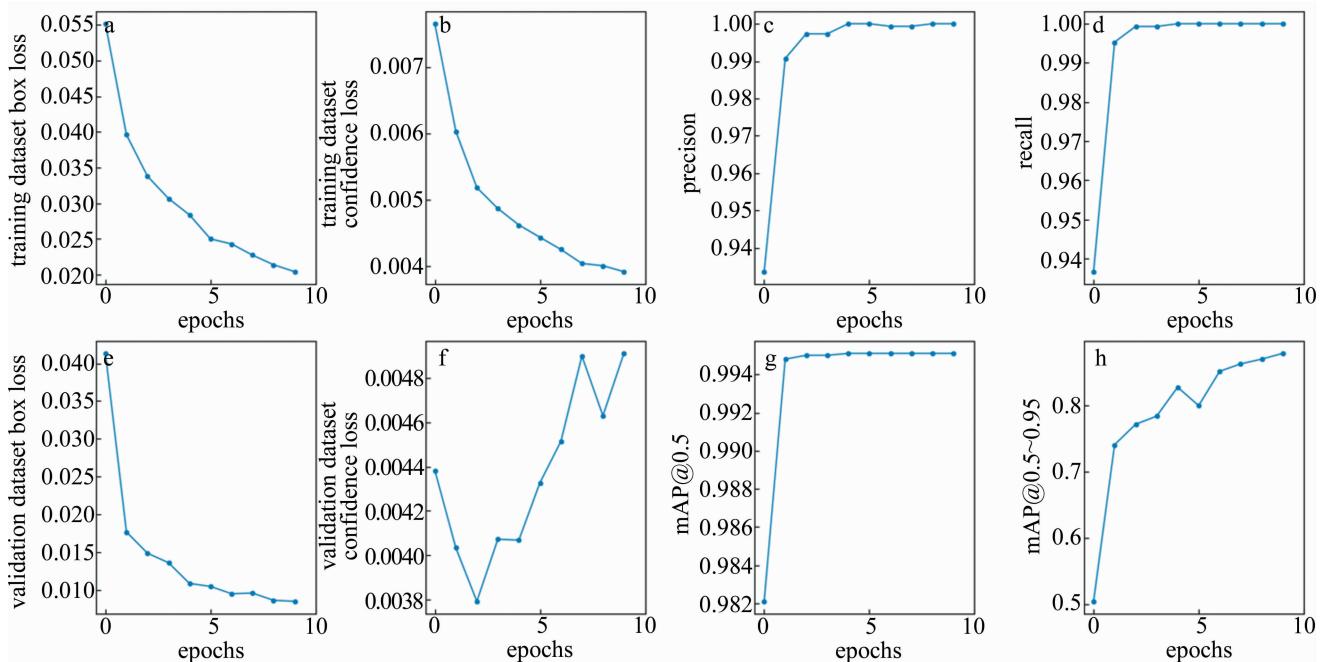


图 5 烟雾识别模型的训练结果

Fig. 5 Training results of smoke recognition model

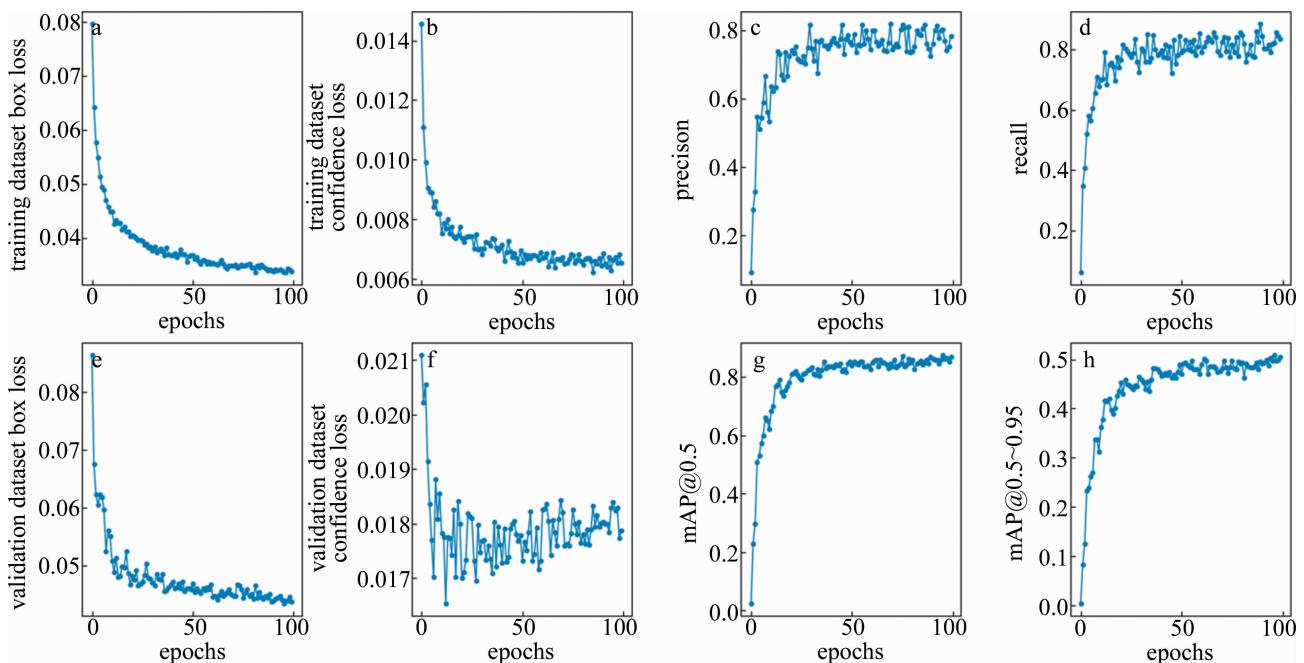


图 6 VOC 识别模型的训练结果

Fig. 6 Training results of VOC recognition model

## 2.5 算法执行效率

为了验证移动端算法的执行效率,本文中采用瑞芯微电子有限公司的 RK3588S 开发板作为识别效率的验证平台。该开发板搭载 4 核 A76 和 4 核 A55 的处理器,主频高达 2.4 GHz,并且集成了 ARM-MaliG610MP4 GPU 以及高性能的 3-D 和 2-D 图像加速模块。此外,该开发板还内置了具备高达 6 T/s 算力的 AI 加速 NPU,支持 INT4、INT8、INT16 和 FP16 混合计算,能有效加速网络模型的推理过程。开发板如图 7 所示。

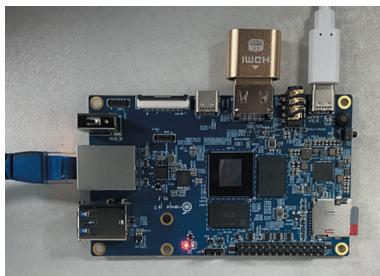


图 7 RK3588S 嵌入式开发板

Fig. 7 RK3588S embedded development board

测试所用视频为第 2.1 节中预处理后的视频,文件大小 9.76 Mbytes,分辨率 640 pixel  $\times$  512 pixel,帧率 25,视频时长 55 s。该视频包含人员移动、镜头移动、林木晃动等典型干扰,最大限度模拟了真实 VOC 泄漏场景中的各种干扰。算法执行过程中,记录各帧投影矩阵变化率,算法运行耗时等关键参数。当视频稳定算法判定视频处于稳定状态,满足背景建模要求后,识别算法开始计时,将运动背景建模、运动前景提取、运动前景图像拉伸、图像增强和 VOC 泄漏区域识别过程,均记录为算法耗时。实验中对 1284 帧数据进行了算法耗时统计,最大、最小和平均耗时分别为 83 ms、38 ms 和 49 ms,各帧详细统计参数如图 8 所示。

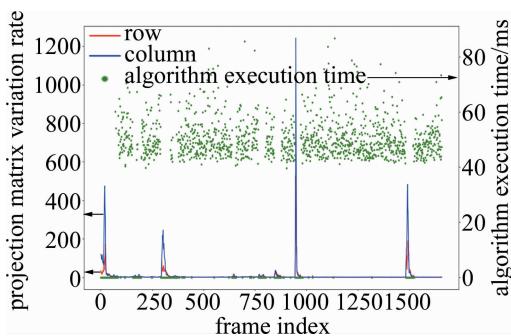


图 8 算法运行参数统计图

Fig. 8 Statistical chart of algorithm runtime parameter

为更直观地展示算法在移动端的检测效果,随机选择 4 帧具有代表性的图片进行可视化表达,识别结果如图 9 所示。可以看出,聚拢和非聚拢形态的 VOC

泄漏区,检测框都予以了准确标注。

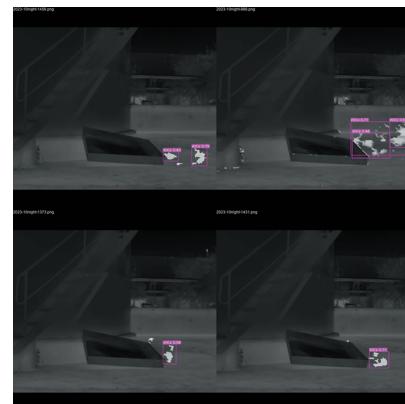


图 9 VOC 识别结果

Fig. 9 VOC recognition results

## 2.6 与其它算法的对比

为比较本文中的算法与帧差法<sup>[6-7]</sup>、轮廓识别法<sup>[18]</sup>的识别效果,在相同实验条件下,包括数据预处理、视频稳定性判定、模型训练策略和算法执行环境,利用第 2.1 节中预处理后的视频数据进行对比实验。实验中选取准确率、算法执行效率和单帧算法耗时标准差为关键对比指标,实验结果如表 3 和图 10 所示。

表 3 算法关键指标统计结果

Table 3 Algorithm key performance indicators statistics results

model	accuracy	algorithm execution efficiency/ (frame $\cdot$ ms $^{-1}$ )	algorithm execution efficiency standard deviation/ (frame $\cdot$ ms $^{-1}$ )
frame difference method	—	18	9
contour detection method	0.79	110	126
the algorithm in this article	0.88 (mAP@0.5)	49	22

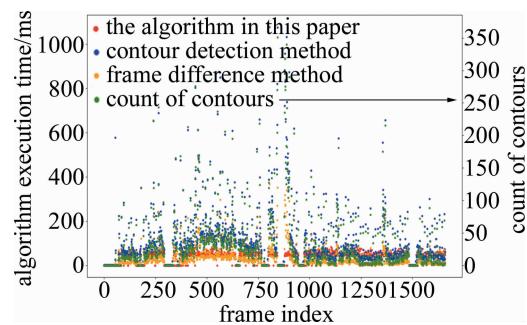


图 10 算法耗时逐帧统计结果

Fig. 10 Algorithm execution time frame-by-frame statistical results

从表 3 中的实验结果可知,本文中算法相较于轮廓识别法,精度提升了 0.09、平均算法耗时降低了 61 ms、算法耗时标准差降低了 104 ms。从图 10 的统计结果来看,轮廓识别法算法稳定性受轮廓数量影响

较大,算法稳定性明显不如本文中算法。

为更直观说明本文中算法的检测效果,选择了视频中原油倾倒前、倾倒过程中和倾倒后3个阶段的视频数据进行对比实验。

原油倾倒前,实验场景内无 VOC 泄漏,但存在人体移动、物体表面辐射变化和林木晃动等干扰,可检验算法的抗干扰能力,实验结果如图 11 所示。帧差法将人体移动、物体表面辐射变化误识别为 VOC 泄漏区;轮廓识别法虽然排除了人体移动干扰,但将物体辐射变化误识别为 VOC 泄漏区;本文中算法则未检测出 VOC 泄漏区。

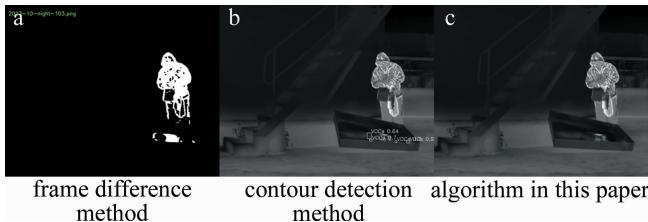


图 11 视频数据第 103 帧检测结果

Fig. 11 Detection result of video data frame 103

原油倾倒过程中,实验场景内包含人体移动、物体表面辐射变化、VOC 泄漏和林木晃动,属复杂场景下的 VOC 泄漏检测场景,实验结果如图 12 所示。帧差法将人体移动、物体表面辐射变化误识别为 VOC 泄漏区;轮廓识别法将部分人体移动干扰和物体辐射变化误识别为 VOC 泄漏区;本文中算法则准确地检测出 VOC 泄漏区。

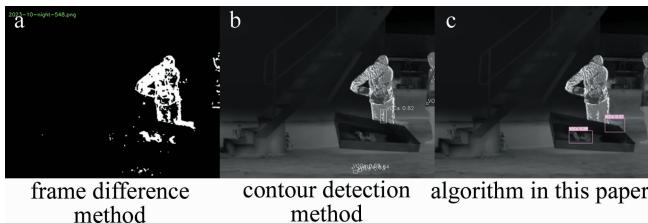


图 12 视频数据第 548 帧检测结果

Fig. 12 Detection result of video data frame 548

原油倾倒后,实验场景内包含物体表面辐射变化、林木晃动和明显的 VOC 泄漏,属于一般场景下的 VOC 泄漏检测场景,实验结果如图 13 所示。帧差法准确识别出 VOC 泄漏区,但同时将物体表面辐射变化误识别为 VOC 泄漏区;轮廓识别法识别出部分 VOC 泄漏区,并将部分物体辐射变化误识别为 VOC 泄漏区;本文中算法则准确地检测出全部 VOC 泄漏区。

综上所述,帧差法执行效率和稳定性最优,但不能识别 VOC 泄漏区和非泄漏区,难以满足自动化检测需求;轮廓识别法虽然可识别具有典型轮廓特征的 VOC 泄漏区,但易受辐射变化、影像噪声、疑似泄漏区数量

等影响,算法稳定性较差;本文中算法则在不同场景下准确识别出 VOC 泄漏区,在算法稳定性、结果可视化方面表现更优,更适用于自动化检测场景。

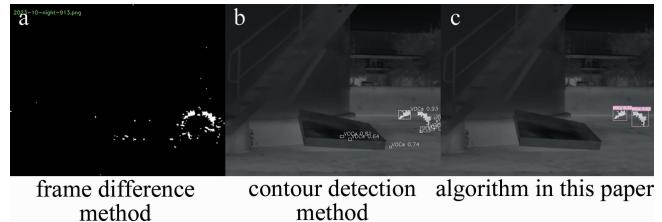


图 13 视频数据第 913 帧检测结果

Fig. 13 Detection result of video data frame 913

### 3 结 论

为了提高 VOC 泄漏区识别效率和识别精度,本文作者设计了一种基于增强运动区域特征并利用目标识别算法识别 VOC 泄漏区的算法。在背景建模方面,该算法将视频稳定性判定技术与高斯背景建模相结合,减少了镜头移动、抖动和检测区内物体快速移动等因素对背景建模和移动区域提取的干扰,提高了背景建模质量和算法执行效率。在泄漏区域特征增强方面,本文中算法利用优化线性拉伸的方法突出了泄漏区域细节,对运动前景进行异常噪声过滤和细节增强,并将其与原始图像进行加权融合,提高了 VOC 泄漏区域的可识别性。在 VOC 泄漏区识别方面,利用目标检测算法,实现了 VOC 泄漏区域的精准识别,泄漏区域识别精度 mAP@0.5 为 0.88, mAP@0.5~0.95 为 0.51, 较轮廓识别法提升了 0.09。在算法执行性能方面,在嵌入式开发板中以红外视频为测试数据,对算法执行效率进行了对比测试,并以原油倾倒前、倾倒过程中和倾倒后 3 个典型场景的检测结果进行了可视化分析。结果表明,本文中算法单帧平均识别时间为 49 ms, 较轮廓识别法降低了 61 ms; 算法耗时标准差为 22 ms, 较轮廓识别法降低了 104 ms。

尽管本文作者提出的 VOC 泄漏区域识别算法在实验中取得了较高的识别和定位精度,但受限于数据集场景单一、数量少等问题,需要对算法的视频稳定性判定阈值进行深入优化和验证。为了确保视频稳定性判定的准确性,应引入更多不同场景的视频数据,从而更精确地确定阈值。在识别模型训练方面,需要收集更多具有不同距离、场景和多源气体泄漏源样本,这将有助于增强算法对各类 VOC 以及各类场景的识别能力。同时,考虑到技术的不断进步,应进一步结合循环神经网络等先进的图像处理和机器学习技术,进一步提升算法的性能和智能化程度。

## 参 考 文 献

- [1] YAN M Q, ZHU H K, LUO H N, *et al.* Daily exposure to environmental volatile organic compounds triggers oxidative damage: Evidence from a large-scale survey in China[J]. *Environmental Science & Technology*, 2023, 57(49):20501-20509.
- [2] CHOI Y H, KIM H J, SOHN J R, *et al.* Occupational exposure to VOC and carbonyl compounds in beauty salons and health risks associated with it in South Korea[J]. *Ecotoxicology and Environmental Safety*, 2023, 256:114837.
- [3] HUANGE Y X, D K, XIONG Y, *et al.* One-third of global population at cancer risk due to elevated volatile organic compounds levels [J/OL]. Research Square: (2023-09) [2024-01-11]. <https://doi.org/10.21203/rs.3.rs-3320416/v1>.
- [4] NARANJO E, BALIGA S, BERNASCOLLE P F, *et al.* IR gas imaging in an industrial setting[J]. *Proceedings of the SPIE*, 2010, 7661:7661K.
- [5] RANGEL J, SCHMOLL R, KROLL A. Catadioptric stereo optical gas imaging system for scene flow computation of gas structures [J]. *IEEE Sensors Journal*, 2021, 21(5):6811-6820.
- [6] HUSEIN A M, CALVIN, HALIM D, *et al.* Motion detect application with frame difference method on a surveillance camera[J]. *Journal of Physics: Conference Series*, 2019, 12309(1):012017.
- [7] WANG L M, TONG Zh, JI B, *et al.* Temporal difference networks for efficient action recognition [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, USA: IEEE Press, 2021:1895-1904..
- [8] XIN Y H, HOU J, DONG L M, *et al.* A self-adaptive optical flow method for the moving object detection in the video sequences[J]. *Optik—International Journal for Light and Electron Optics*, 2014, 125(19):5690-5694.
- [9] WANG X, FENG L I, XIN L, *et al.* Moving targets detection for satellite-based surveillance video[C]//IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium. Yokohama, Japan: IEEE Press, 2019:5492-5495.
- [10] HE W, LI J X, QI Q, *et al.* SIM-MFR: Spatial interactions mechanisms based multi-feature representation for background modeling [J]. *Journal of Visual Communication and Image Representation*, 2022, 88:103622.
- [11] WANG J P, LI J Sh, YANG Y W, *et al.* Ethylene gas leakage detection based on infrared imaging[J]. *Journal of Liquid Crystals and Displays*, 2014, 29(4):623-628(in Chinese).  
王建平,李俊山,杨亚威,等.基于红外成像的乙烯气体泄漏检测[J].液晶与显示,2014,29(4):623-628.
- [12] LIU L M G, ZHANG Y Z, LUAN L, *et al.* A shape-based infrared image gas leakage detection method[J]. *Journal of Applied Optics*, 2019, 40(3):468-472(in Chinese).  
刘路民根,张耀宗,栾琳, *et al.* 一种基于形状的红外图像泄漏气体检测方法[J].应用光学,2019,40(3):468-472.
- [13] HONG Sh Zh, YING H, YU H W, *et al.* A VOC gas detection algorithm based on infrared thermal imaging[C]//2019 Chinese Control and Decision Conference (CCDC). Nanchang, China: IEEE Press, 2019:329-334.
- [14] BADAWI D, PAN H Y, CETIN S C C, *et al.* Computationally efficient spatio-temporal dynamic texture recognition for volatile organic compound (VOC) leakage detection in industrial plants[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2020, 14(4):676-687.
- [15] XU Y, DONG J X, ZHANG B, *et al.* Background modeling methods in video analysis: A review and comparative evaluation [J]. *CCAI Transactions on Intelligence Technology*, 2016, 1(1):43-60.
- [16] ZHENG Y, FAN L Zh. Moving object detection based on running average background and temporal difference[C]//2010 IEEE International Conference on Intelligent Systems and Knowledge Engineering. Hangzhou, China: IEEE Press, 2010:270-272.
- [17] MEGHANA R K, CHITKARA Y, MOHANA A. Background-modelling techniques for foreground detection and tracking using Gaussian mixture model[C]//2019 3rd International Conference on Computing Methodologies and Communication (ICCMC). Erode, India: IEEE Press, 2019:1129-1134.
- [18] ZHU M Zh, WANGE H B. Fast detection of moving object based on improved frame-difference method[C]//2017 6th International Conference on Computer Science and Network Technology (ICCSNT). Dalian, China: IEEE Press, 2017:299-303.
- [19] FEI F, KANG J L, LIU J L, *et al.* Implementation of short-wave infrared image gray-level stretching algorithm based on FPGA[J]. *Journal of Terahertz Science and Electronic Information*, 2022, 20(7):713-717(in Chinese).  
费斌,康佳龙,刘俊良,等.基于FPGA的短波红外图像灰度级拉伸算法实现[J].太赫兹科学与电子信息学报,2022,20(7):713-717.
- [20] ZHOU Y K, ZHU Y P, ZENG B Z, *et al.* A review of wide dynamic range infrared image enhancement algorithms[J]. *Laser Technology*, 2018, 42(5):718-726(in Chinese).  
周永康,朱尤攀,曾邦泽,等.宽动态红外图像增强算法综述[J].激光技术,2018,42(5):718-726.
- [21] WEI Y P. Infrared image enhancement using linear transformation and local equalization fusion[J]. *Laser Technology*, 2024, 48(5):705-710(in Chinese).  
魏艳平.线性变换与局部均衡融合的红外图像增强[J].激光技术,2024,48(5):705-710.
- [22] WANG Y C, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-free bies sets new state-of-the-art for real-time object detectors[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, Canada: IEEE Press, 2023:7464-7475.
- [23] CHENG Y H, YIN J L, CHEN B H, *et al.* Smoke 100k: A database for smoke detection[C]//2019 IEEE 8th Global Conference on Consumer Electronics (GCCE). Osaka, Japan: IEEE Press, 2019:596-597.