

基于红外热像的YOLOV8n-TOD列车障碍物检测算法

赵守俊^{1*}, 陈 嘉², 谢兰欣³, 张轩雄¹

¹上海理工大学光电信息与计算机工程学院, 上海

²上海申通地铁集团有限公司技术中心, 上海

³苏州同睿兴科技有限公司, 江苏 苏州

收稿日期: 2024年12月25日; 录用日期: 2025年1月18日; 发布日期: 2025年1月24日

摘 要

热成像可在低光环境下检测障碍物。针对列车颠簸影响图像质量的问题, 基于ORB特征提取算法与Farneback、Lucas-Kanade光流法加权平均设计一种EIS算法, 对采集的数据进行EIS及CLAHE预处理。同时, 针对红外图像低分辨率、高噪声敏感性的问题, 提出一种列车障碍物检测算法YOLOV8n-TOD, 该算法从3个方面进行改进: 在YOLOV8n算法中使用MobileNetV3网络替换原主干, 通过轻量级结构和深度可分离卷积操作提高算法的计算效率和特征提取能力; 在颈部网络中使用FasterBlock网络重构C2f模块, 优化特征融合及增强信息传递, 提高算法的稳定性与检测精度; 优化CIoU损失函数, 提高算法的泛化能力。测试结果显示: 经预处理后YOLOV8n算法的mAP提高了2.4%; 采用YOLOV8n-TOD算法后mAP又提升了7.2%, 显著增强了障碍物检测能力。

关键词

EIS, Farneback, Lucas-Kanade, MobileNetV3, FasterBlock

YOLOV8n-TOD Algorithm for Train Obstacle Detection Based on Infrared Thermal Imaging

Shoujun Zhao^{1*}, Jia Chen², Lanxin Xie³, Xuanxiong Zhang¹

¹School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

²Technology Center, Shanghai Shentong Metro Group Co., Ltd., Shanghai

³Suzhou Tongruixing Technology Co., Ltd., Suzhou Jiangsu

*第一作者。

文章引用: 赵守俊, 陈嘉, 谢兰欣, 张轩雄. 基于红外热像的YOLOV8n-TOD列车障碍物检测算法[J]. 建模与仿真, 2025, 14(1): 1086-1099. DOI: 10.12677/mos.2025.141099

Received: Dec. 25th, 2024; accepted: Jan. 18th, 2025; published: Jan. 24th, 2025

Abstract

Thermal imaging can detect obstacles in low-light environments. To address the issue of image quality degradation caused by train vibrations, an EIS algorithm has been designed based on ORB feature extraction and weighted averaging using Farneback and Lucas-Kanade optical flow methods. The collected data undergoes EIS and preprocessing with CLAHE. To address the low resolution and high noise sensitivity of infrared images, a train obstacle detection algorithm, YOLOv8n-TOD, is proposed. The algorithm enhances YOLOv8n in three ways: replacing the original backbone with MobileNetV3 for efficient feature extraction using its lightweight structure and depthwise separable convolutions; by using FasterBlock networks to reconstruct the C2f module in the neck network, optimizing feature fusion and enhancing information transfer to improve model stability and detection Accuracy; and by refining the CIOU loss function to boost model generalization capability. Experimental results show that after preprocessing, the mAP of the YOLOv8n algorithm increased by 2.4%; with the YOLOv8n-TOD model, the mAP further improved by 7.2%, significantly enhancing obstacle detection performance.

Keywords

EIS, Farneback, Lucas-Kanade, MobileNetV3, FasterBlock

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

列车障碍物检测任务主要由可见光摄像机和激光雷达承担,可见光摄像机捕捉前方轨道图像,通过图像处理和机器学习算法进行障碍物识别、分类,激光雷达提供三维点云数据进一步验证和精确障碍物位置。文献[1]提出自适应反射强度阈值过滤无效点云数据,再结合体素滤波器、聚类半径差异化等方法进行聚类,提升全自动列车的障碍物检测速度与精度。文献[2]通过检测与轨道平行的参照物间接提取轨道区域,利用曲线拟合推断出中心线并生成二维轨道区域,再结合地面检测获取三维轨道区域,获取更长的探测区域。然而,可见光摄像机受限于光线,在低光环境下得到的图像容易模糊,检测效果往往不佳。相比之下,热成像[3]对红外辐射具有敏感性,不受天气和光照限制,能够在低光环境下清晰成像,可以作为可见光摄像机的辅助手段实时获取目标物的红外辐射信息,监测包括人、动物在内可能对列车行驶安全造成威胁的障碍物,保障列车运行安全。

障碍物主动检测系统根据检测设备的安装位置分为轨旁障碍物检测和车载障碍物检测[4]。轨旁主动障碍物检测将设备安装在重点防护区段进行障碍物检测。检测对象相对简单,算法复杂度相对较低。文献[5]基于红外热成像设计了一套全天候轨道行人闯入检测系统来辅助驾驶员发现轨道上的行人并及时发出警报。文献[6][7]基于目标增强融合红外和可见光图像配准算法解决异源图像配准问题,提高障碍物检测的准确率。车载主动障碍物检测通常将传感器安装在列车两端头车上,以便随列车行进实时监测。YOLO [8]-[10]系列、SSD [11]、超声波[12]等目标检测技术都在铁路障碍物检测上有广泛应用。其中,YOLO 系列凭借其高速、高效、良好的可移植性和高灵活性,广泛应用于小目标和密集目标等相关检测

任务[13] [14], 而 YOLOV8 在平衡速度和精度的基础上, 进一步增强了实时性。对于检测速度和准确性有较高要求的列车障碍物检测任务, 可有效提升检测系统的响应速度和精确度, 为列车运行安全提供关键支持。

然而, 红外图像的分辨率低、细节少, 与背景温差小的目标难以检测, 且存在噪音干扰。此外, 车载主动障碍物检测过程中, 视频随列车行进产生的抖动容易造成图像模糊, 影响检测效果。针对上述两个问题, 采集不同天气状态下列车前方数据, 基于 ORB 算法设计一种电子防抖(Electronic Image Stabilization, EIS)算法, 并结合对比限制自适应直方图均衡化(Contrast Limited Adaptive Histogram Equalization, CLAHE)对数据进行预处理, 为后续障碍物检测提供更清晰、稳定的数据集。之后, 使用 MobileNetV3 网络替换 YOLOV8n 原有的主干网络 CSPDarkNet53, 提高算法计算效率和特征提取能力; 使用 FasterBlock 网络重构颈部网络中的 C2f 模块, 提高算法检测精度和稳定性; 优化 CIOU 损失函数, 提高算法泛化能力。实现提高障碍物检测精度的要求。

2. 数据采集及处理

使用某典型热像仪采集不同天气下列车前方数据, 对数据进行 EIS、帧提取和 CLAHE 预处理, EIS 通过检测和补偿视频帧之间的相对运动来减少抖动引起的图像不稳定, 提升图像质量。CLAHE 将图像划分为多个小块, 并对每个块进行均衡化处理, 增强局部对比度, 同时避免过度增强噪声, 提升图像的辨识度。进行障碍物检测之前先对数据预处理有助于之后 YOLOV8n 算法更准确地检测障碍物, 提升列车的运行安全。流程如图 1 所示。

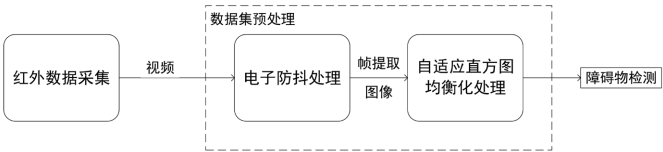


Figure 1. Data collection and preprocessing process
图 1. 数据采集及预处理流程

2.1. 数据采集

模拟列车运动轨迹, 使用推车搭载热像仪录制其在轨道上前进时轨道前方视频, 如图 2 所示。采集的视频数据包含: 人(person)、车(car)、狗(dog)、火车(train)、骑行者(Cyclists)、自行车(bicycle)、小推车(cart)、电动车(Electric bicycle)共 8 个类别。

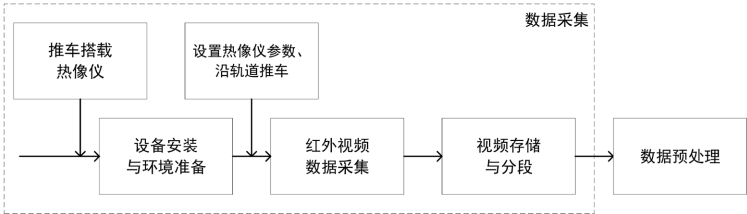


Figure 2. Data collection flowchart
图 2. 数据采集流程图

2.2. 红外视频电子防抖处理

为了减少列车运行中颠簸和震动对红外视频图像造成的影响, 提高视频稳定性和清晰度, 基于 ORB

算法设计了一段 EIS 算法, 流程如图 3 所示。

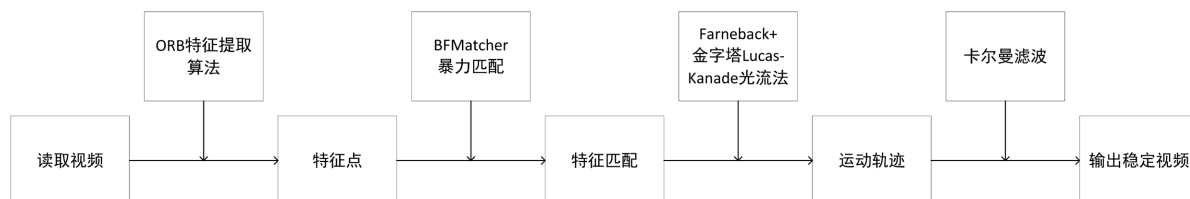


Figure 3. Electronic image stabilization algorithm flow

图 3. 电子防抖算法流程

使用 ORB 检测器从每一帧中提取特征点和描述符, 通过暴力匹配(Brute-force matcher, BFMatcher)进行前后帧特征点匹配, 再加权平均融合 Farneback 经典光流法和金字塔 Lucas-Kanade 稀疏光流法计算匹配特征点的光流来估计帧间仿射变换矩阵以提高运动估计的准确性。使用卡尔曼滤波对计算出的运动轨迹进行平滑处理, 生成一个平滑且稳定的轨迹。再根据轨迹对每一帧进行仿射变换补偿, 去除由抖动、噪声等因素导致的小幅度波动、不规则运动, 使视频画面更加稳定、平滑。

2.2.1. ORB 特征提取算法

ORB [15] 是一种结合了 FAST 特征点检测与 BRIEF 特征描述符, 实时性更强, 具有旋转不变性且对噪声有一定鲁棒性, 与红外图像有较好的匹配性能。

对于像素 p 及其周围 16 个邻域像素 p_i , 通过 FAST 特征点检测来判断 p 是否是特征点。FAST 特征点检测的计算公式为

$$\text{if } \sum_{i=1}^N |I(p_i) - I(p)| > t, \text{ then } p \text{ is a corner point} \quad (1)$$

其中, $I(p)$ 是像素 p 的灰度值, $I(p_i)$ 是邻域像素 p_i 的灰度值, N 是连续像素的数量, t 是阈值。

得到的特征点 p 的质心坐标 (C_x, C_y) 及二者间角度 θ 的计算公式为

$$C_x = \frac{\sum_i x_i \cdot I(p_i)}{\sum_i I(p_i)}, \quad C_y = \frac{\sum_i y_i \cdot I(p_i)}{\sum_i I(p_i)} \quad (2)$$

$$\theta = \arctan\left(\frac{C_y - y_p}{C_x - x_p}\right) \quad (3)$$

其中, (x_i, y_i) 是像素 p_i 的坐标, (x_p, y_p) 是特征点 p 的坐标。

根据 θ 旋转像素 p_i , 对旋转后的像素点 p_1 和 p_2 进行重复比较, 生成一个长度为 n 的二进制描述符, 即 BRIEF 特征描述符, 计算公式为

$$\text{BRIEF}(p_1, p_2) = \begin{cases} 1, & I(p_1) < I(p_2) \\ 0, & I(p_1) \geq I(p_2) \end{cases} \quad (4)$$

2.2.2. 融合光流法

Farneback 光流法[16]基于多项式展开估计每个像素的光流, 计算图像中所有像素的运动矢量, 提供全图密集光流信息。计算公式为

$$v_F(x, y) = \frac{\partial I}{\partial x} u + \frac{\partial I}{\partial y} v + \frac{\partial I}{\partial t} \quad (5)$$

其中, (u, v) 是光流矢量, $\frac{\partial I}{\partial x}$ 、 $\frac{\partial I}{\partial y}$ 是图像的空间梯度, $\frac{\partial I}{\partial t}$ 是时间梯度, $I(x, y)$ 是图像亮度, 满足约束方程

$$I(x, y, t) = I(x + u, y + v, t + 1) \quad (6)$$

金字塔 Lucas-Kanade 光流法[17]先在低分辨率图像中估计光流, 再逐层细化到高分辨率图像, 对特征点的跟踪更为精准。在局部窗口 W 内, 计算公式为

$$V_{L-K}(x, y) = \min_v \sum_{(x', y') \in W} \left(\frac{\partial I}{\partial x} u + \frac{\partial I}{\partial y} v + \frac{\partial I}{\partial t} \right)^2 \quad (7)$$

其中, (x', y') 是窗口 W 内像素坐标。

对于具有噪声和低对比度区域特性的红外图像, 结合 Farneback 全局运动信息和金字塔 Lucas-Kanade 局部精细运动估计, 平衡计算效率与精度的同时可以提高鲁棒性。最终运动矢量计算公式为

$$V_{final} = aV_F + (1-a)V_{L-K} \quad (8)$$

其中, a 是权重因子, 用于平衡两种光流法融合后的结果。

根据加权平均后的最终运动矢量可以计算得出更精确的仿射变换矩阵, 进一步得到更准确的运动补偿。

3. 红外热像列车障碍物检测算法

3.1. YOLOV8n-TOD 算法结构

考虑实时性需求, 选择最小、最轻量的 YOLOv8n 算法[18], 将其改进为列车障碍物检测算法 YOLOV8n-TOD, 如图 4 所示。该算法的改进主要包括: 使用 MobileNetV3 网络替换算法原有主干网络, 增强特征提取能力, 提升推理速度。颈部网络 C2f 模块结合 FasterBlock 模块重构, 提升算法检测精度。优化 CIOU 损失函数, 提升算法泛化能力。

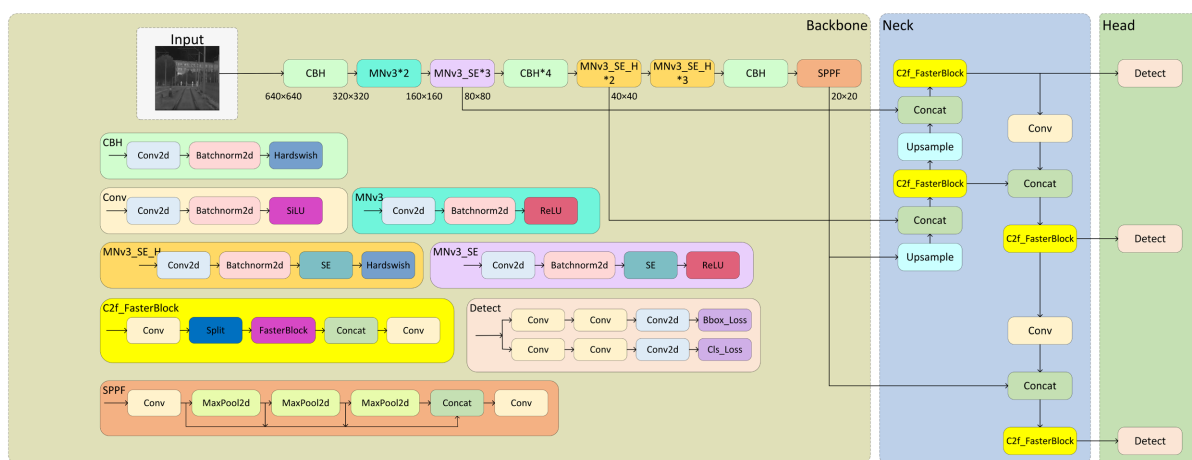


Figure 4. YOLOV8n-TOD model architecture diagram

图 4. YOLOV8n-TOD 算法结构图

3.2. 具有 MobileNetV3 的改进型 Backbone 网络

红外图像具有较低对比度和有限颜色信息, 温度和光照变化都有可能造成误检, 降低检测精度, 因

此对特征提取的有效性要求更高；此外，车载障碍物检测受车载环境影响，通常无法使用算力较高的计算设备[4]。障碍物检测的延迟可能会影响列车即时响应，威胁乘客安全。因此，为了提升检测的速度和精度，需要更高的计算效率和更强的特征提取能力。

MobileNetV3 是一种由 MobileNetV2 改进而来的轻量化网络[19]，如图 5 所示，其核心原理是整合深度可分离卷积、倒残差结构、SE (Squeeze-and-Excitation)模块、Hard-swish 激活函数，以及 NAS (Neural Architecture Search)技术，保证准确率的同时，增强特征提取能力、降低计算复杂度、减少计算参数数量。替换 YOLOV8n 原有主干网络可以降低算法推理时间，提升算法检测精度。

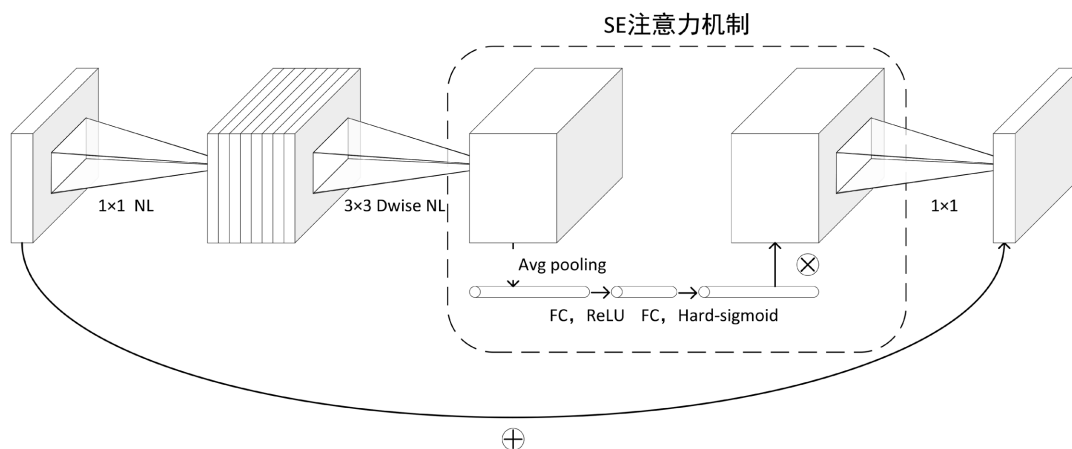


Figure 5. MobileNetV3 network architecture diagram

图 5. MobileNetV3 网络结构图

SE 模块通过自适应重新调整每个通道的权重，增强关键特征的表达并抑制无关特征，提高算法对障碍物的识别能力。Hard-swish 激活函数更加敏锐地捕捉微小特征变化，有效保留重要特征信息，减少对小幅信号信号的忽略，进一步提高检测的准确性。同时，Hard-swish 具有较高的计算效率，在列车障碍物检测中，能够更快做出反应，保障行车安全。此外，NAS 技术优化后的 MobileNetV3 网络结构能够有效识别找出最适合的层结构和超参数设置，进而提升算法的识别精度和鲁棒性。

MobileNetV3 有 large 和 small 两种结构，large 结构有更多参数和更复杂的网络结构，特征提取能力更强、精度更高、更适应复杂场景，在兼顾速度与精度方面更有优势。

3.3. 利用 FasterBlock 重构 C2f 模块改进 Neck 网络

多尺度特征融合与提取，有助于提升算法的检测精度与稳定性。同时，由于列车制动距离较远，障碍物检测的感知范围也需要相应提升[4]。

FasterBlock 是 FasterNet 网络中的核心构建模块[20]，如图 6 所示，是一种用于优化深度学习网络结构的轻量化模块，其核心原理是在网络结构设计过程中通过部分卷积(Partial Conv3)和多层感知机(MLP)等轻量化结构，减少网络的计算复杂度和参数量，提升推理速度。

Partial Conv3 在部分通道上执行卷积操作，保留其余通道原始信息，在保存和强化重要特征信息的同时，避免冗余计算对效率的影响。MLP 通过扩展和恢复特征维度，深度处理输入特征，提高算法表达能力，保持高效计算性能。此外，DropPath 在训练时随机丢弃部分计算路径，防止过拟合，以此增强算法的泛化能力，Layer Scale 则进一步对不同通道的特征进行自适应缩放调整，提升特征表达的准确性。通过 FasterBlock 重构的 C2f_FasterBlock 模块采用 CSP(Cross Stage Partial)结构思想，将特征分为经卷积

操作与未经卷积操作两部分进行处理，并在不同尺度上特征融合，在保持原始信息的同时，提取更丰富的多尺度特征，进一步提高小型和远距离上的障碍物检测精度。

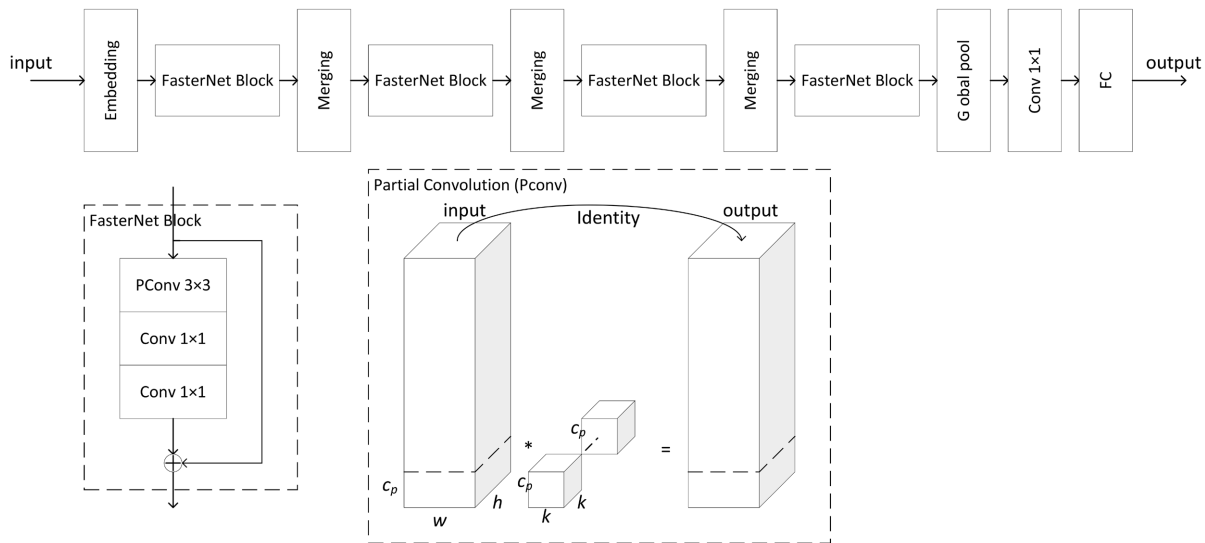


Figure 6. FasterNet network architecture diagram

图 6. FasterNet 网络结构图

3.4. 损失函数优化

YOLOV8n 算法配置的损失函数为 CIOU (Complete-IoU)，在红外图像中，目标边界较为模糊，且障碍物类型不同，宽高比概率也不相同，CIOU 在 DIOU 基础上考虑宽高比[21]，有更准确的相似性度量和更强的鲁棒性，提供更稳健的边界框回归效果，减小障碍物误检、漏检的概率，提高算法的检测精度。

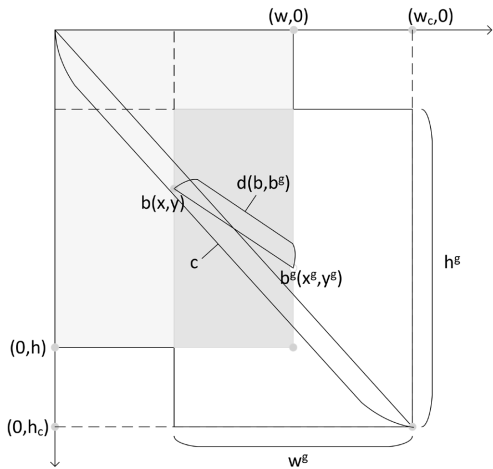


Figure 7. CIOU calculation diagram

图 7. CIOU 计算示意图

参照图 7，CIOU 计算公式为

$$L_{\text{CIOU}} = L_{\text{IOU}} - \frac{d(b, b^g)^2}{c^2} - av \quad (9)$$

其中, v 是预测框和真实框之间的长宽比相似度, 计算公式为

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^g}{h^g} - \arctan \frac{w}{h} \right)^2 \quad (10)$$

w 和 h 是预测框的宽和高, w^g 和 h^g 是真实框的宽和高。

a 是权重系数, 用来平衡长宽比相似度对 CIOU 的影响, 计算公式为

$$a = \frac{v}{(1 - L_{\text{IOU}}) + v} \quad (11)$$

$d(b, b^g)$ 是预测框和真实框之间的中心点距离, 计算公式为

$$d(b, b^g) = \sqrt{(x - x^g)^2 + (y - y^g)^2} \quad (12)$$

(x, y) 和 (x^g, y^g) 分别表示预测框中心点 b 和真实框中心点 b^g 的坐标。

c 是预测框和真实框的最小包围框对角线距离, 计算公式为

$$c = \sqrt{w_c^2 + h_c^2} \quad (13)$$

w_c 和 h_c 是包围矩形的宽和高。

4. 测试验证

4.1. 测试环境

测试在恒源云算力平台进行, 环境配置如表 1 所示。

Table 1. System configuration and versions

表 1. 系统配置及版本

系统配置	版本
操作系统	Linux Ubuntu 20.04.6 LTS
内核	GNU/Linux 5.4.0-155-generic
GPU	NVIDIA GeForce RTX 4060 Ti 16GB
CPU	Intel(R) Xeon(R) CPU E5-2683 v4
CUDA 版本	12.1.1
Python 版本	3.11

模型参数设置为: 批大小(batch size)为 64, 训练轮次(epoch)为 250 轮, 输入图片尺寸(imgsz)为 640×640 pixel, 优化器(optimizer)为自动选择, 初始学习率为根据优化器自动选择。

4.2. 评价指标

测试采用平均精度(mean Average Precision, mAP)、检出率(Recall, R)和漏检率(False Negative Ratio, FNR)作为评价指标。

漏检率 FNR 指模型检测漏掉的目标占有实际目标的比例, 与检出率互补。

$$\text{FNR} = 1 - R \quad (14)$$

结合精度(Precision, P)与 R 计算各个类别的平均精度(Average Precision, AP), 对所有类别 AP 加和再平均得到 mAP, 能够全面评估算法在不同检测阈值下的表现。mAP50 即 IOU 阈值选取 0.5。

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^N \text{AP}_i \quad (15)$$

$$\text{AP} = \sum_{j=1}^n (R_j - R_{j-1}) \cdot P_j \quad (16)$$

式中, N 是类别数, n 是不同的阈值数量, R_j 是每个阈值下的召回率, P_j 是对应的精度。

在 EIS 中常采用结构相似性指数(Structural Similarity Index, SSIM)和峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)作为评价指标。

SSIM 能够衡量 EIS 处理前后视频结构、亮度和对比度等信息的保持情况。结果在[0,1]间, 越接近 1 图像越相似, 结构信息保留得越好。对两幅图像块 x 、 y

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (17)$$

其中, μ 是均值, σ^2 是方差。通常, $C_1 = (K1L)^2$, $C_2 = (K2L)^2$, $K1 = 0.01$, $K2 = 0.03$, $L = 255$ 。

PSNR 能够评估视频防抖后帧的质量相对原始视频帧的偏差, 超过 30 dB 一般被认为是高质量图像。

$$\text{MSE} = \frac{1}{w \cdot h} \sum_{i=0}^{w-1} \sum_{j=0}^{h-1} [I(i, j) - K(i, j)]^2 \quad (18)$$

$$\text{PSNR} = 10 \cdot \lg \left(\frac{\text{MAX}^2}{\text{MSE}} \right) \quad (19)$$

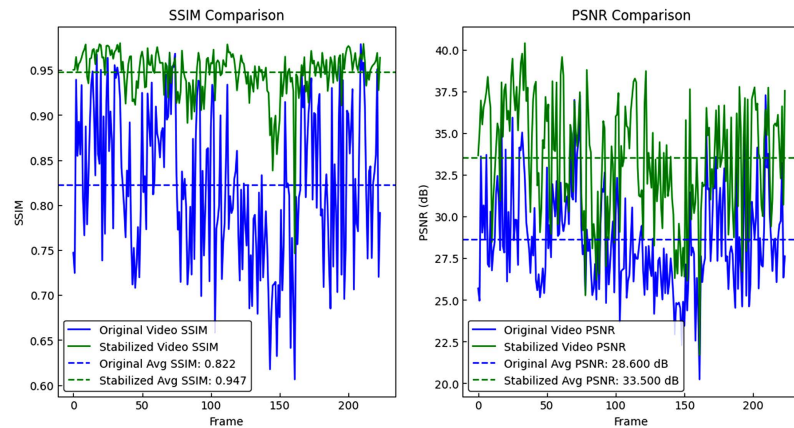
其中, MAX 是图像像素的最大可能值, 通常为 255, $I(i, j)$ 、 $K(i, j)$ 分别是原图像和处理后图像在位置 (i, j) 的像素值, w 和 h 是图像的宽和高。

4.3. 对比测试

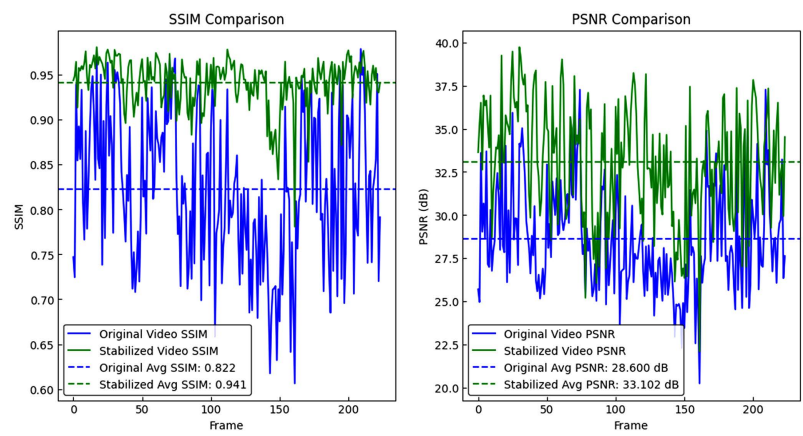
经 EIS 和 CLAHE 预处理后, 得到 2857 张图像, 按照 7.8:1:1.2 比例随机划分为训练集、验证集和测试集。

4.3.1. EIS 算法处理前后结果比较分析

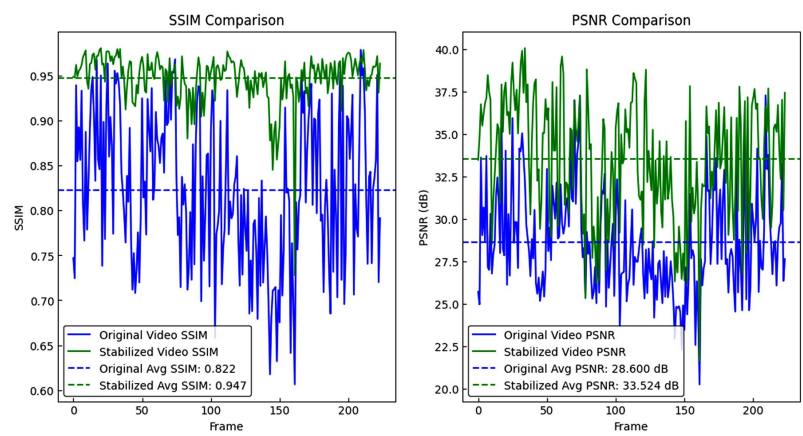
使用 YOLOV8n 算法对 EIS 处理前后的同一段视频进行对比测试, 测试结果如图 8 所示。结果表明, 经过 EIS 算法处理的视频能够很好的实现去抖。



(a) Farneback 光流法



(b) Lucas-Kanade 光流法



(c) Farneback、Lucas-Kanade 融合

Figure 8. Comparison of SSIM and PSNR curves before and after model improvement

图 8. 算法改进前后 SSIM 和 PSNR 曲线对比

4.3.2. EIS 和 CLAHE 算法处理前后性能对比

基于 YOLOV8n 算法对预处理前后数据进行对比测试，实验结果如表 2 所示。结果表明，经过预处理后算法的 mAP50 提升了 2.4%，R 提升了 0.5%，验证了预处理的有效性。

Table 2. Performance comparison before and after data preprocessing

表 2. 数据预处理前后性能对比

算法	mAP50/%	R/%	FNR/%
YOLOV8n	51.8	47.2	52.8
+EIS	53.4	48.1	51.9
YOLOV8n-EA (+EIS + CLAHE)	54.2	47.7	52.3

4.3.3. 主干网络、颈部网络改进前后性能对比

将替换 MobileNetV3 网络、C2f_FasterBlock 模块的 YOLOV8n 算法与原始的 YOLOV8n 算法进行对比测试，测试结果如表 3 所示。MobileNetV3 网络改进后算法的 mAP 提升了 3.7%，R 提升了 2.3%，C2f_FasterBlock 模块改进后算法的 mAP 提升了 1%，R 提升了 3.1%，算法效果得到显著优化。

Table 3. Performance comparison before and after backbone and neck network improvement
表 3. 主干、颈部网络改进前后性能对比

算法	mAP50/%	R/%	FNR/%
YOLOV8n-EA	54.2	47.7	52.3
+MobileNetV3	57.9	50.0	50.0
+C2f_FasterBlock	55.2	50.8	49.2

4.3.4. 优化损失函数前后性能对比

修改 CIOU 的权重系数 a ，将调整 a 后的 YOLOV8n 算法与原 YOLOV8n 算法进行对比测试，测试结果如表 4 所示。结果表明，当 $a = 0.3$ 时，算法的 mAP 提升了 2%。

Table 4. Performance comparison before and after loss function optimization
表 4. 损失函数优化前后性能对比

a	mAP50/%	R/%	FNR/%
0.1	56.1	47.7	52.3
0.2	56.2	47.7	52.3
0.3	56.2	47.7	52.3
0.4	55.8	47.7	52.3
0.5	55.6	47.7	52.3

4.4. 消融实验

将所有改进方式集成到同一算法上，得到列车障碍物检测算法 YOLOV8n-TOD。为了验证 YOLOV8n-TOD 算法效果，进行 5 组消融实验，结果如表 5 所示。OSDaR23 [22] 包含 20 个不同的与铁路环境相关的各种对象类，其测试结果同样见表 5。

Table 5. Ablation study results
表 5. 消融实验结果

MobileNetV3	C2f_FasterBlock	CIOU	mAP50/%		R/%		FNR/%	
			自制	OSDaR23	自制	OSDaR23	自制	OSDaR23
			54.2	34.3	47.7	27.3	52.3	72.2
√			57.9	31.9	50.0	27.6	50.0	72.4
	√		55.2	35.1	50.8	31.4	49.2	68.6
		√	56.2	34.6	47.7	27.3	52.3	72.7
√	√		59.7	35.7	49.8	27.2	50.2	72.8
√	√	√	61.4	38.8	52.2	33.1	47.8	66.9

从表 5 中可以看出，相比原 YOLOV8n 算法，改进后的 YOLOV8n-TOD 算法在处理数据时 mAP 提升了 7.2%、R 提升了 4.5%，在 OSDaR23 上 mAP 提升了 4.5%、R 提升了 5.8%。

图 9 为在预处理后数据上 YOLOV8n 和 YOLOV8n-TOD 算法训练过程 mAP 曲线对比，可以看出将 MobileNetV3、C2f_FasterBlock 集成到同一算法并优化损失函数后，算法的 mAP 有明显提升。

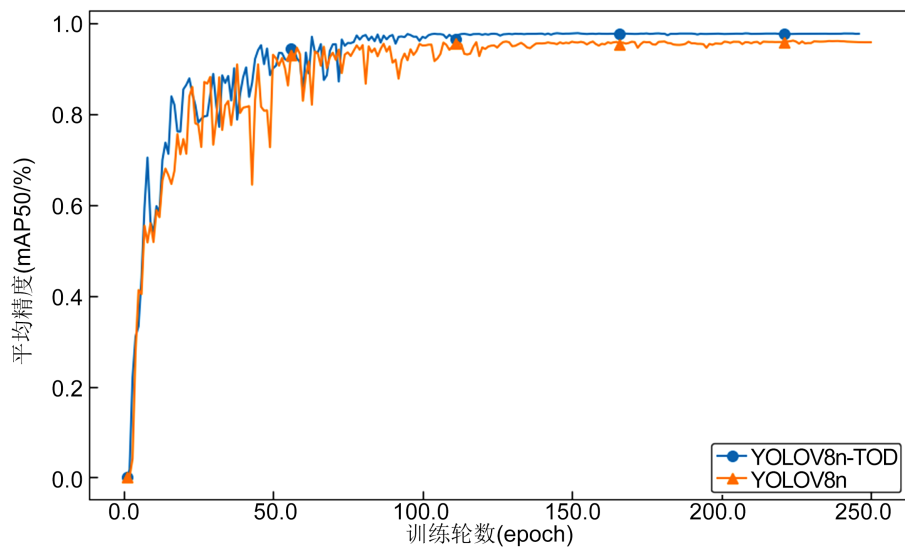


Figure 9. Comparison of mAP curves before and after model improvement
图 9. 算法改进前后 mAP 曲线对比

图 10 为 YOLOV8n、YOLOV8n-EA 和 YOLOV8n-TOD 算法性能对比结果。

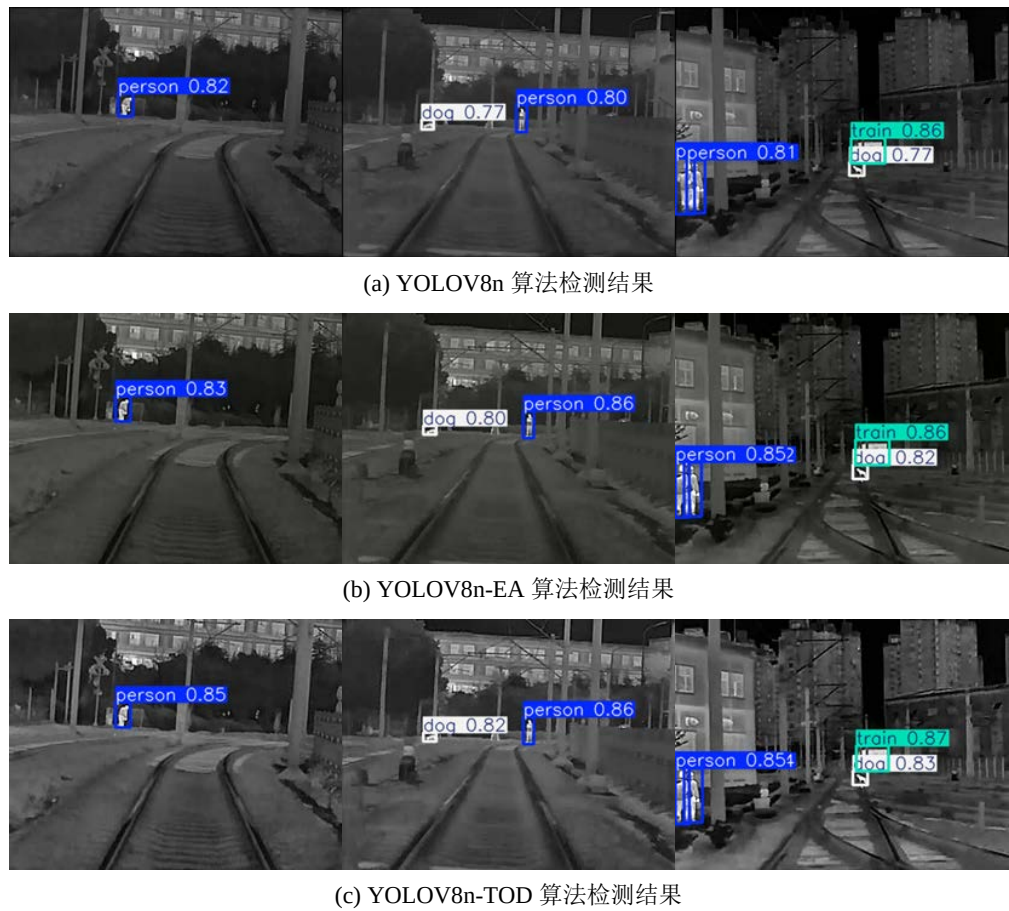


Figure 10. Performance comparison of YOLOV8n, YOLOV8n-EA, and YOLOV8n-TOD models
图 10. YOLOV8n、YOLOV8n-EA 和 YOLOV8n-TOD 算法性能对比

为评估算法性能，将改进后的算法与一些主流算法进行对比测试，表 6 为不同算法对比结果。

Table 6. Performance comparison of different algorithms
表 6. 不同算法的性能对比

算法	mAP50/%	R/%	FNR/%	GFLOPs	FPS
Faster R CNN	51.1	-	-	-	0.08
YOLOV3	53.2	47.5	52.5	282.2	35.46
YOLOV5n	49.7	47.1	52.9	7.2	83.33
YOLOV5m	53.0	48.1	51.9	64.0	72.99
YOLOV8n	51.8	47.7	52.3	8.2	83.33
YOLOV8n-TOD	61.4	52.2	47.8	10.0	88.50

从结果看出，改进后的 YOLOV8n-TOD 算法的 mAP、R 和 FPS 高于其他方法，在 GFLOPs 上高于 YOLOV5n 和 YOLOV8n，但足够满足实时性要求。

5. 结论

针对列车障碍物检测场景，列车行进过程中通常会产生颠簸造成红外视频画面抖动进而影响障碍物检测的检出率及准确性。因此，提出了一种基于 ORB 特征提取算法与 Farneback、Lucas-Kanade 光流法加权平均的 EIS 算法，ORB 特征提取算法与红外图像有较好的匹配性能；Farneback 光流法与金字塔 Lucas-Kanade 光流法加权平均可以在平衡计算效率与精度的同时提高鲁棒性。

同时，针对红外图像分辨率低、易受噪声影响、特征提取能力较弱的问题，提出了列车障碍物检测算法 YOLOV8n-TOD：将 YOLOV8n 算法的 Backbone 替换成 MobileNetV3 主干网络，增强算法特征提取能力，提升算法检测精度和推理速度；将 Neck 中的 C2f 模块结合 FasterBlock 模块重构为 C2f_FasterBlock 模块，保留原始信息的同时进一步增强算法的多尺度特征提取能力，促使小目标特征更易捕捉，进一步扩展算法的检测范围；优化 CIOU 损失函数，提升算法泛化能力。将改进后的算法 YOLOV8n-TOD 在自制数据集和 OSDaR23 数据集上与原 YOLOV8n 算法进行对比测试，YOLOV8n-TOD 的 mAP 分别提升了 7.2%和 4.5%、R 分别提升了 4.5%和 5.8%、FNR 降低了 4.5%和 5.8%，测试结果表明，该算法在保证实时性的基础上，能够明显提高障碍物检测的平均精度和检出率，降低漏检率，具有较好的检测性能，对保障列车行驶安全具有重要意义。未来将进一步优化算法结构以提升检测性能。

基金项目

上海市科委项目(23DZ2204900)。

参考文献

- [1] 沈拓, 钱沿佐, 谢兰欣, 等. 考虑反射强度的全自动运行列车障碍物检测算法研究[J]. 同济大学学报(自然科学版), 2022, 50(1): 6-12.
- [2] Shen, T., Zhou, J., Yuan, T., Xie, Y. and Zhang, X. (2024) Lidar-Based Urban Three-Dimensional Rail Area Extraction for Improved Train Collision Warnings. *Sensors*, **24**, Article 4963. <https://doi.org/10.3390/s24154963>
- [3] 陈钱, 隋修宝. 红外图像处理理论与技术[M]. 北京: 电子工业出版社, 2018: 7-11.
- [4] 师帅. 轨道交通系统主动障碍物检测研究综述[J]. 机电工程技术, 2021, 50(6): 212-216.
- [5] 崔晗. 基于热成像视频的铁路行人闯入检测系统[D]: [硕士学位论文]. 北京: 北京邮电大学, 2022.
- [6] 许鑫龙. 基于目标增强融合的铁路异物入侵检测方法研究[D]: [硕士学位论文]. 北京: 北京交通大学, 2020.

-
- [7] 孙永丽. 基于图像的铁路障碍物自动检测算法研究[D]: [硕士学位论文]. 西安: 西安电子科技大学, 2018.
- [8] 杜开华, 许贵阳, 白堂博. 基于特征降冗余的 Vanilla-YOLOv8 铁路异物侵限检测方法[J/OL]. 北京交通大学学报, 1-12. <http://kns.cnki.net/kcms/detail/11.5258.U.20240903.1206.002.html>, 2024-09-09.
- [9] 吴浩楠, 史宏, 王瑞, 等. 基于改进 YOLO v8 的铁路人员入侵检测方法研究[J/OL]. 铁道科学与工程学报, 1-12. <https://doi.org/10.19713/j.cnki.43-1423/u.T20241118>, 2024-09-09.
- [10] 王辉, 姜朱丰, 吴雨杰, 等. 基于深度学习的铁路异物侵限快速检测方法[J]. 铁道科学与工程学报, 2024, 21(5): 2086-2098.
- [11] 李建国, 陈敬涛, 张伟, 等. 基于改进型 SSD 算法的铁路货场异物侵限小目标检测研究[J]. 铁道通信信号, 2024, 60(7): 57-62.
- [12] 司全龙, 施婷. 基于超声波视频检测技术的列车运行异物检测系统研究[J]. 办公自动化, 2023, 28(4): 59-61+41.
- [13] 何幸, 黄永明, 朱勇. 基于改进 YOLOv5 的路面坑洼检测方法[J]. 电子科技, 2024, 37(7): 53-59.
- [14] 王小铸, 于莲芝. 基于卷积与自注意力聚合的小目标检测[J]. 电子科技, 2024, 37(2): 14-22.
- [15] Rublee, E., Rabaud, V., Konolige, K. and Bradski, G. (2011) ORB: An Efficient Alternative to SIFT or Surf. 2011 *International Conference on Computer Vision*, Barcelona, 6-13 November 2011, 2564-2571. <https://doi.org/10.1109/iccv.2011.6126544>
- [16] Farneback, G. (2003) Two-Frame Motion Estimation Based on Polynomial Expansion. In: *Lecture Notes in Computer Science*, Springer, 363-370. https://doi.org/10.1007/3-540-45103-x_50
- [17] Baker, S. and Matthews, I. (2004) Lucas-Kanade 20 Years On: A Unifying Framework. *International Journal of Computer Vision*, **56**, 221-255. <https://doi.org/10.1023/b:visi.0000011205.11775.fd>
- [18] Varghese, R. and M., S. (2024) YOLOv8: A Novel Object Detection Algorithm with Enhanced Performance and Robustness. 2024 *International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)*, Chennai, 18-19 April 2024, 1-6. <https://doi.org/10.1109/adics58448.2024.10533619>
- [19] Howard, A., Sandler, M., Chen, B., Wang, W., Chen, L., Tan, M., *et al.* (2019) Searching for MobileNetV3. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 1314-1324. <https://doi.org/10.1109/iccv.2019.00140>
- [20] Chen, J., Kao, S., He, H., Zhuo, W., Wen, S., Lee, C., *et al.* (2023) Run, Don't Walk: Chasing Higher FLOPS for Faster Neural Networks. 2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, 17-24 June 2023, 12021-12031. <https://doi.org/10.1109/cvpr52729.2023.01157>
- [21] Zheng, Z., Wang, P., Ren, D., Liu, W., Ye, R., Hu, Q., *et al.* (2022) Enhancing Geometric Factors in Model Learning and Inference for Object Detection and Instance Segmentation. *IEEE Transactions on Cybernetics*, **52**, 8574-8586. <https://doi.org/10.1109/tcyb.2021.3095305>
- [22] Tagiew, R., Klasek, P., Tilly, R., Köppel, M., Denzler, P., Neumaier, P., *et al.* (2023) OSDaR23: Open Sensor Data for Rail 2023. 2023 *8th International Conference on Robotics and Automation Engineering (ICRAE)*, Singapore, 17-19 November 2023, 270-276. <https://doi.org/10.1109/icrae59816.2023.10458449>