

基于改进 YOLOv8n-Pose 的轨道作业人员跨轨安全动作识别

叶彦斐 胡龙葵 张成龙

(河海大学人工智能与自动化学院 南京 211106)

摘要:针对轨道作业人员跨轨安全动作监督方法存在效率低、漏检率高等问题,引入改进的人体姿态估计算法 YOLOv8n-Pose 对跨轨安全动作进行识别和监督。对 YOLOv8n-Pose 算法改进方法为在网络中添加注意力机制并轻量化网络结构,并改进网络的 bbox 损失函数和关键点损失函数,以提高网络的识别精度和速度。使用高斯滤波和 ColorJitter 算法对自制数据集增强。在训练前使用遗传算法对训练超参数进行自适应调整,在训练时使用迁移学习和知识蒸馏方法,提高网络训练速度、识别精度和泛化能力。将训练好的模型对轨道现场作业人员图像进行检测,可成功识别出作业人员姿态并根据关键点位置信息识别安全动作,人体关键点识别精确度为 94.3%,推理速度为 238.1 fps,验证模型改进研究取得了有益效果,提高了模型识别精度、识别速度和鲁棒性。

关键词:人体姿态估计;深度学习;YOLOv8n-Pose;目标检测

中图分类号: TP391.4; TP183 **文献标识码:** A **国家标准学科分类代码:** 510.8

Safety actions recognition of rail workers crossing the track based on improved YOLOv8n-Pose

Ye Yanfei Hu Longgui Zhang Chenglong

(College of Artificial Intelligence and Automation, Hohai University, Nanjing 211106, China)

Abstract: Aiming at the problems of low efficiency and high missed detection rate of the rail workers' cross-track safety actions supervision method, an improved human pose estimation algorithm YOLOv8n-Pose is introduced to detect and supervise the cross-track safety actions. The improvement method of YOLOv8n-Pose algorithm is to add an attention mechanism to the network and lighten the network structure, and improved the bbox loss function and the keypoint loss function of the network in order to improve the network's recognition accuracy and speed. The self-made dataset is enhanced by Gaussian filtering and ColorJitter algorithm. Genetic algorithm is used to adaptively adjust the training hyperparameters before training, and migration learning and knowledge distillation methods are used during training to improve the network training speed, recognition accuracy and generalization ability. The trained model is used to detect the images of the workers, which can successfully recognize the workers' keypoints and identify the safety actions based on the keypoints. The human keypoints recognition accuracy is 94.3% at the speed of 238.1 fps, which verifies that the model improvement research has achieved beneficial effects and improved recognition accuracy, recognition speed and robustness of the model.

Keywords: human pose estimation; deep learning; YOLOv8n-Pose; object detection

0 引言

铁路维护过程中,铁路工作人员需频繁跨越铁轨,为保障其生命安全,工作人员应遵守严格的标准来跨越铁

轨,即只有在特定的、受摄像头监控的道口,完成“一停、二看、三确认、四通过”标准安全动作后才能跨越铁轨。为对该安全动作进行监督,传统方法为人工通过摄像头监督,该方法效率低下,易出现纰漏,存在安全隐患。文献[1]引

入目标检测算法,对轨道作业人员进行识别,跟踪其行进路径并保存这段视频,但无法识别其动作,上述安全动作监督仍需人工查看保存的视频,实时性差且效率较低。

随着人工智能的发展,深度学习在目标检测领域应用广泛,文献[2]基于 YOLOv4 的轻量级目标检测网络 R-YOLO 实现了精确快速的铁路人员入侵识别。文献[3]通过将 YOLOv5 网络结合卷积注意力机制,改进损失评价函数,提高了轨道探伤检测的实时性和准确性。

基于深度学习的人体姿态估计算法应用也非常广泛。文献[4]为解决高精度识别存在推理速度慢的问题,创新性提出了轻量级通道空间注意力网络 LCSA-Net,并在损失函数上使用基于距离自适应加权策略,同时提高了识别精度和速度。文献[5]将 YOLOv8 的主干网络替换为 ShuffNetV2,并引入混合注意力机制(shuffle attention, SA),解决了特征提取不充分、实时性差的问题,但使用 ShuffNetV2 让网络特征提取能力变差。文献[6]在 OpenPose 算法上增加了一个小型的 6 层网络提取人体姿态全局特征,并融合 YOLO v3 算法提取的人体手部特征,对学生动作进行识别分析,但增加的 6 层网络对识别精度提升较低,不如注意力机制。文献[7]使用改进的 OpenPose 算法对人体动作进行识别,其将网络的 VGG19 网络替换为 ResNet18 网络,并使用深度可分离卷积替代了网络中的部分卷积,虽然该方法在不损失过多识别精度的前提下减少了网络参数,但其识别速度仍然较慢。文献[8]基于热力图积分,提出了回归方法 soft-argmax,该方法将热力图转化为概率分布并积分计算关键点坐标,一定程度上提高了关键点回归精度,但使用 soft-argmax 会导致概率分布多峰,不利于关键点回归的准确率和置信度。3D 动作识别旨在从流视频中实时识别正在进行的动作,文献[9]提出了轻量高效的多尺度动作学习 Transformer,并进一步提出了用于捕捉多尺度动作特征、具有多个编码分支的分层编码器。但是基于流视频的 3D 动作识别方法占用磁盘空间大,识别速度慢,不利于部署到边缘计算设备。

综上所述,使用人体姿态估计的动作识别算法在识别速度和精度上无法很好地权衡,而 3D 动作识别算法模型大,不利于部署到边缘计算设备。有关跨轨安全动作识别的研究较少,为解决铁路工作人员跨越铁轨安全动作监督存在的问题,本文将改进人体姿态估计算法,用于实时识别人员关键点,并根据关键点位置信息识别安全动作,减少对人工的依赖。这对于保护铁路现场工作人员生命安全,保障铁路安全畅通运行有着重大意义。

1 跨轨安全动作监督算法

标准安全动作需要人员完成左右两侧安全动作和一个前向抬手的动作,其核心在于使用姿态估计算法识别出人体关键点,利用关键点位置关系判断人员在跨越铁轨前是否完成安全动作。为此本文采用改进的 YOLOv8n-Pose 算法进行姿态估计,跨轨安全动作监督算法流程如

图 1 所示。

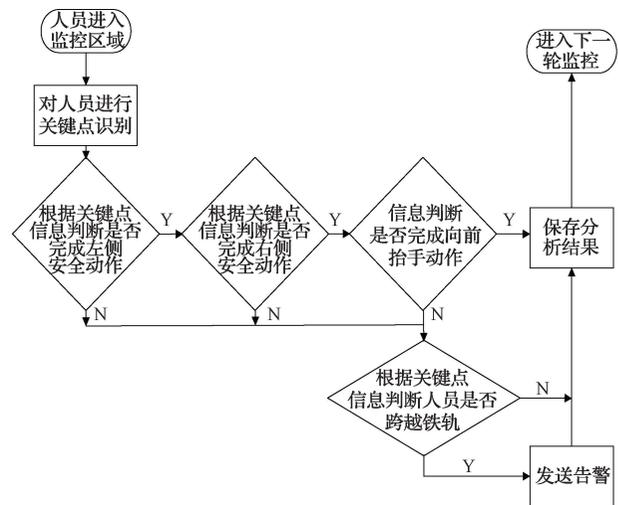


图 1 监督算法流程

Fig. 1 Flowchart of the supervision algorithm

用姿态估计算法识别人员关键点是根据关键点位置信息分析人员脸部朝向、左右手臂与身体关系,以识别安全动作。安全动作的判定基于同侧手腕、手肘、肩膀连线与肩膀、胯部连线的夹角, $\geq 90^\circ$ 为安全动作,正常走路时夹角较小。头部状态的判断依据于双耳、眼睛、鼻子关键点连线状态,正脸时只能识别双耳,左转头时能识别左耳、左眼,右转头时能识别右耳、右眼。若做安全动作时为左转头,则为左向安全动作;若为右转头,则为右向安全动作;若为正脸且一侧手肘、肩膀、胯部连线夹角超过 150° ,则为前向抬手动作。跨轨动作通过判断两脚踝关键点连线与预设轨道线的相对位置关系来确定,若踝部连线始终在轨道线同侧,则认为未越线;若出现在轨道线两侧,则认为越线。

YOLOv8n-Pose 的结构如图 2 所示,由主干网络(Backbone)、颈部网络(Neck)和头部网络(Head)组成。主干网络包含 Conv、C2F 和 SPPF 模块。颈部网络采用路径聚合网络(path aggregation network, PAN)以融合不同尺度的特征信息^[10]。头部网络采用目前流行的解耦头结构,将来自颈部网络的特征分别进行目标框坐标、类别信息和关键点坐标预测。

2 数据集制作与处理

2.1 数据采集

采用自制数据集,从某铁路局轨道作业现场监控视频中截取含有现场作业人员共 1 911 张分辨率为 $1\ 920 \times 1\ 080$ 的图像制作数据集,原始图像如图 3(a)所示。并对原始数据集进行数据增强,使其数量扩充到 3 822 张,同时将 COCO val 2017 关键点数据集中随机 1 346 张图像加入数据集以提高模型泛化能力,得到数据集共 5 168 张图像。最后按 8:1:1 比例将数据集划分为训练集、验证

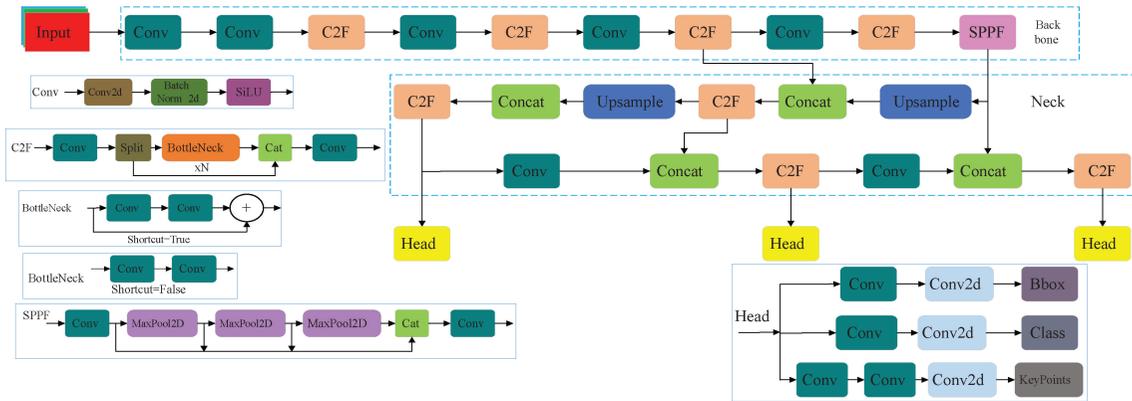


图2 YOLOv8n-Pose 网络结构
Fig. 2 YOLOv8n-Pose network structure

集和测试集,用于模型的训练、验证和测试。

2.2 数据集标注

使用 Labelme 对数据集图像标注,并将标注好的标签转为 YOLO 的 txt 格式。标签中内容为人体所在位置的锚框坐标及其 17 个关键点的坐标,数据集标注好后则可用于模型的训练、验证和测试。

2.3 数据增强

由于原始数据集图片数量较少,容易导致模型出现过拟合、泛化能力差等问题,故对原始数据集进行图像增强以扩充数据集,增强后的图片与原始图片的对比如图 3 所示,其中图 3(a)为原始图像,图 3(b)为增强后的图像。



图3 原始图像与增强后图像对比
Fig. 3 Original image vs enhanced image

随后将 COCO val 2017 人体关键点数据集中随机 1 346 张图像加入自制数据集,以提高模型的鲁棒性和泛化性。

数据集中不少图片存在高斯白噪声,故先对原始图像进行高斯滤波去噪,使图像更加清晰平滑。然后使用 ColorJitter 算法对原始图像进行增强,具体为对图像亮度在 0.5~1.5 范围内随机变换,对比度在 0.6~1.4 范围内随机变换、色彩饱和度及色域在 0.7~1.3 范围内随机变换,从而得到新的图像扩充数据集,提高网络的鲁棒性和泛化性。最后将上述增强后图像、原始图像和 COCO val 2017 中随机 1 346 张图像共 5 168 张图像作为本文自制数据集,以提升关键点检测精度,防止网络过拟合。

3 YOLOv8n-Pose 算法改进

3.1 添加 EMA 注意力机制

为使网络更精炼地提取特征信息,提高网络对人体及其关键点区域注意力,在网络的颈部位置加入高效多尺度注意力模块 (efficient multi-scale attention, EMA)^[11],如图 4 所示。EMA 模块是一种轻量级并行结构的注意力模块,其结构如图 5 所示。

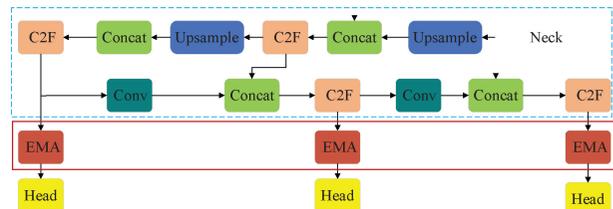


图4 添加注意力机制的方式
Fig. 4 The way of adding attention mechanism

EMA 注意力机制的核心是重构部分通道至批次维度,并将通道维度细分为多个子特征,旨在保存每个通道信息的同时降低计算负担。EMA 模块利用编码全局信息的方式重新调整每个并行分支中的通道权重,并通过跨维度交互捕捉像素级关系。

它的并行子结构和跨空间特征融合,提高了人体及其关键点区域在网络预测过程中的计算权重,使网络更好地利用人体及其关键点相关的特征信息,减少样本分布不均及背景对识别带来的影响,使网络对人体及其关键点更敏感。

3.2 改进交并比 (IoU) 损失函数

YOLOv8n-Pose 使用 Ciou 损失函数,虽效果不错,但它未考虑真实框与预测框之间方向不匹配的问题,在训练中会出现预测框与真实框在某一方向上震荡回归导致无法进一步收敛或收敛效果变差,这会导致网络训练速度慢,识别精度低。传统 IoU 也有其局限性,它无法根据训

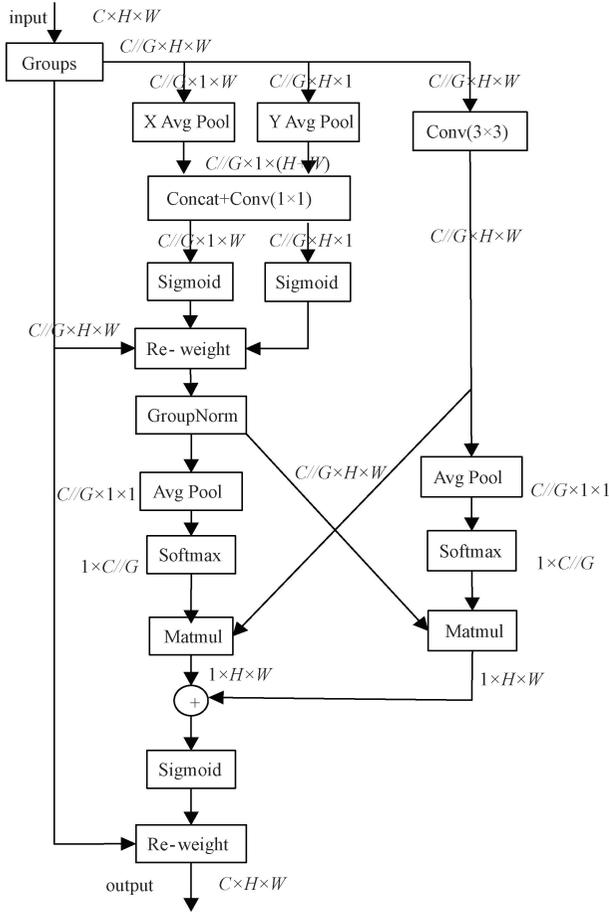


图5 EMA 模块结构

Fig. 5 EMA module structure

练状态自动调整计算值大小,泛化能力弱。为加快网络收敛速度,提升网络识别精度和泛化性,本文将 Inner-IoU^[12]与 SIoU^[13]结合,提出了 α -InnerSIoU 损失函数,其表达式如下:

$$L_{\alpha\text{-InnerSIoU}} = 1 - (1 - \alpha)IoU - \alpha IoU^{inner} + \frac{\Delta + \Omega}{2} \quad (1)$$

其中 α 为可学习参数,限定在 0~1 之间,用于调节预测框(图 6 绿色实线框)与真实框(图 6 黑色实线框)交并比和辅助预测框(图 6 绿色虚线框)与辅助真实框(图 6 黑色虚线框)交并比权重;预测框与真实框的 IoU 公式如式(2)所示;为辅助预测框与辅助真实框的交并比 IoU^{inner} 如式(3)所示;预测框中心点与真实框中心点距离惩罚项如式(4)所示;预测框与真实框形状相似度惩罚项 Ω 如式(8)所示。

$$IoU = \frac{intersection}{union} \quad (2)$$

式中: $intersection$ 表示真实框与预测框相交部分面积; $union$ 表示真实框与预测框相并部分面积。

$$IoU^{inner} = \frac{intersection^{inner}}{union^{inner}} \quad (3)$$

式中: $intersection^{inner}$ 表示辅助真实框与辅助预测框相交

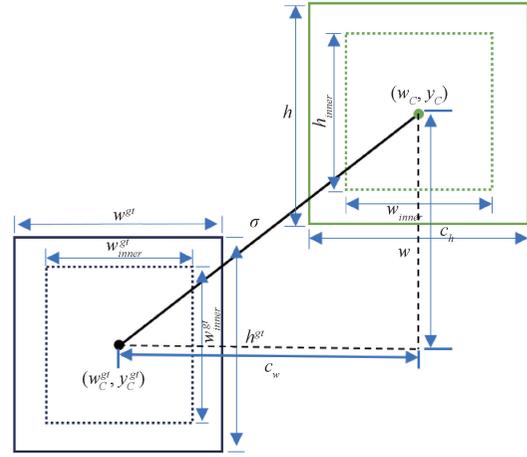


图6 改进 IoU 计算图

Fig. 6 Improved IoU computational descriptive graph

部分面积; $union^{inner}$ 表示辅助真实框与辅助预测框相并部分面积。

图 6 中,真实框和辅助真实框中心点坐标为 (x_c, y_c) , (x_c^{gt}, y_c^{gt}) 分别为真实框宽高, w, h 分别为预测框宽高, w^{gt}, h^{gt} 分别为真实框宽高, w^{inner}, h^{inner} 分别为辅助真实框宽高, w^{gt}, h^{gt} 分别为辅助预测框宽高, $ratio$ 为辅助框的比例超参数,根据文献[12]将其设置为 0.75。

$$\Delta = \sum_{t=x,y} (1 - e^{-\gamma \rho_t}) \quad (4)$$

$$\rho_x = \left(\frac{x_c^{gt} - x_c}{c_w} \right), \rho_y = \left(\frac{y_c^{gt} - y_c}{c_h} \right), \gamma = 2 - \Delta \quad (5)$$

$$c_w = |x_c^{gt} - x_c|, c_h = |y_c^{gt} - y_c| \quad (6)$$

$$\Lambda = 1 - 2\sin^2\left(\arcsin\left(\frac{c_h}{\sigma}\right) - \frac{\pi}{4}\right) \quad (7)$$

$$\sigma = \sqrt{c_w^2 + c_h^2}$$

式中: Λ 表示真实框与预测框之间的角度惩罚项。

$$\Omega = \sum_{t=x,y} (1 - e^{-\omega_t})^\theta \quad (8)$$

$$\omega_w = \frac{|\omega - \omega^{gt}|}{\max(\omega, \omega^{gt})}, \omega_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})} \quad (9)$$

式中: θ 为超参数,用于控制网络对 Ω 的计算权重。根据文献[13], θ 取值范围宜为 2~6,将其设置为 4。

3.3 改进关键点损失函数

在 YOLOv8n-Pose 算法中,使用关键点相似度(object keypoint similarity, OKS)作为关键点预测评价标准,其取值范围为 0~1,预测结果越准确,OKS 越接近于 1,则关键点损失函数公式为:

$$L_{kp} = 1 - OKS = 1 - \frac{\sum_i \exp\left(\frac{-d_i^2}{2s^2k_i^2}\right) \delta(v_i > 0)}{\sum_j \delta(v_j > 0)} \quad (10)$$

式中: $d_i^2 = (x_i - x_i^{gt})^2 + (y_i - y_i^{gt})^2$ 表示第 i 个人体关键

点预测位置与真实位置的欧氏距离。由于该损失函数只考虑了预测点和真实点的欧氏距离,未考虑方向匹配问题,训练时会出现预测点在真实点附近徘徊回归现象,导致网络收敛速度慢,预测精度低等问题。本文改进了关键点损失函数,在原有基础上加入轴最大距离惩罚项,其公式为:

$$L_{kp} = 1 - \frac{\sum_i \exp\left(\frac{-(d_i^2 + \mu * D_{\max})}{2s^2 k_i^2}\right) \delta(v_i > 0)}{\sum_j \delta(v_j > 0)} \quad (11)$$

式中: $D_{\max} = \max(|x_i - x_i^{gt}|, |y_i - y_i^{gt}|)$ 为轴最大距离惩罚项, μ 为可训练参数,用于调节惩罚项的权重,该惩罚项使预测点优先向真实点垂直、水平距离最大的方向回归,减少徘徊现象,提高网络预测精度。为进一步提升网络的非线性特征表达能力,将网络激活函数改为 Swish 激活函数,其公式为:

$$Swish(x) = \frac{x}{1 + e^{-\beta x}} \quad (12)$$

式中: β 为可学习参数。

3.4 模型轻量化改进

为对网络轻量化,加快推理速度,本文对网络中的 C2F 模块进行了改进,将其 BottleNeck 替换成了 FasterNet^[14] 中的 FasterBlock 模块,以轻量化 C2F 模块。FasterBlock 模块创新性提出了部分卷积(partial convolution, Pconv),该卷积操作只对输入进行部分卷积操作,以减少冗余特征实现高效空间特征提取,使网络参数量、计算量和内存访问量有效降低,加快了网络推理速度。

3.5 迁移学习与知识蒸馏

为提高训练效率,将迁移学习和知识蒸馏方法用于训练。主要步骤为,将基于 COCO 关键点数据集训练好的 YOLOv8m-Pose 和 YOLOv8n-Pose 模型分别作为教师和学生网络的训练初始化参数;然后使用本文数据集训练 YOLOv8m-Pose 教师网络。将本文网络作为学生网络,使用本文数据集训练,训练中将教师网络输出与学生网络输出的交叉熵损失作为软损失加入总损失中,使学生网络学习教师网络的知识以提高识别精度。使用迁移学习和知识蒸馏的方法,提高训练速度的同时弥补了自制数据集图像少的缺点,提高了模型精度和泛化性。

4 实验

4.1 实验环境

本文的实验硬件环境如下:CPU 为 intel i7-12700H, GPU 为 NVIDIA GeForce RTX3060(6G),内存为海力士 DDR5 16 G 4 800 MHz;软件环境为:操作系统为 Windows 11, Python 版本为 3.8.0, PyTorch 版本为 2.1.1, Cuda 版本为 11.8。

4.2 训练策略及参数设置

训练前,基于本文数据集,使用遗传算法自适应调整

训练相关的超参数用于后续训练。使用迁移学习和知识蒸馏的方法训练本文网络,其中教师网络的精度、mAP@0.5、mAP@0.5:0.95 分别为 98.7%、95.8%、85.5%。训练中优化器选用 AdamW,总训练轮数为 500 epochs,忍耐度(patience)为 50 epochs,输入图片尺寸为 640×640, batch size 为 16。网络训练的损失下降曲线如图 7 所示,训练在第 415 轮后提前结束,各损失均已达到收敛,训练后的模型整体趋于稳定,性能良好。

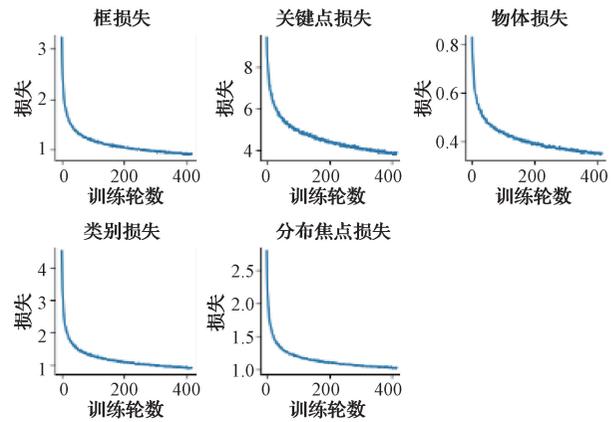


图 7 训练损失曲线

Fig. 7 Training loss curve

4.3 消融实验

为验证本文改进策略的有效性,对改进策略进行消融实验,其结果如表 1 所示,选用原始的 YOLOv8n-Pose 算法作为基准,选用精确度(precision)、平均精确度(mean average precision, mAP) mAP@0.5、mAP@0.5:0.95 和浮点数(floating point operations)作为模型性能评价指标。其中人体关键点预测精确度计算公式为:

$$P_{kp} = \frac{\sum_p \delta(OKS_p > conf_{OKS})}{\sum_p 1} \quad (13)$$

mAP@0.5 表示阈值为 0.5 时平均精确度, mAP@0.5:0.95 表示阈值从 0.5~0.95、步长为 0.05 的平均精确度;浮点运算次数用于衡量模型大小。

表 1 中, BL 表示原始 YOLOv8n-Pose 网络, M1 表示 EMA 模块, M2 表示 α -InnerSIoU 损失函数, M3 表示改进的关键点损失函数和 Swish 激活函数, M4 表示 FasterBlock 模块。由表 1 可知, 4 种改进模块对 YOLOv8n-Pose 网络性能都有所改善。在加入 EMA 模块之后, 得益于其并行化的卷积核并通过跨空间信息聚合方法, 使得输入图像的特征图利用了更丰富的上下文信息, 推理效果提升显著, 在精度、mAP@0.5、mAP@0.5:0.95 分别提高了 1.5%、2.3%、4.9%。 α -InnerSIoU 损失函数考虑了预测框和真实框之间的方向匹配问题, 并在训练时通过辅助框判断预测框的不同阶段回归情况, 自动调整 IoU 的计算值, 使得网络更加准确的检测出人体位置, 有助于网络后

表 1 改进策略对性能的影响

Table 1 Impact of improvement strategies on performance

算法	Precision/%	mAP@0.5/%	mAP@0.5:0.95/%	浮点数/GFLOPs
BL	91.9	90.3	71.4	9.2
BL+M1	93.4	92.6	76.3	9.4
BL+M2	92.73	91.02	73.8	9.2
BL+M3	92.84	91.4	74.0	9.2
BL+M1+M2	94.0	92.7	76.7	9.4
BL+M1+M2+M3	95.2	92.8	76.2	9.4
BL+M1+M2+M3+M4	94.3	93.3	74.8	7.6

续对人体关键点的检测,使用该损失函数后,网络性能在 mAP@0.5:0.95 提升最大,提高了 2.4%。在使用改进的关键点损失函数和 Swish 激活函数后,得益于改进的关键点损失函数提高了预测点与真实点在水平距离和垂直距离中较大部分的计算权重,Swish 激活函数更有利于网络梯度传播,网络在 mAP@0.5 和 mAP@0.5:0.95 指标上较原模型提升明显,分别提高了 1.1%、2.6%。3 种改进方法叠加使用后,网络性能在各个指标上提升幅度均为最大。最后在 3 种改进方法的基础上使用 FasterBlock 模块,由于其

使用了部分卷积操作,有效减少了计算量,性能较使用 3 种改进方法的网络略微下降的情况下,把网络计算量由 9.4 GFLOPs 下降到 7.6 GFLOPs,大幅提高推理速度。

4.4 对比试验

为验证本文算法优越性,选取 YOLO-Pose^[15]、YOLOv7-Pose^[16]、YOLOv8+ShuffleNetV2+SA^[5] 和原始的 YOLOv8s-Pose 进行对比。为保证实验结果可靠性,训练策略均保持一致,使用精度、mAP@0.5、mAP@0.5:0.95 和检测帧率作为评价标准,实验结果如表 2 所示。

表 2 对比实验结果

Table 2 Results of comparative experiments

算法	Precision/%	mAP@0.5/%	mAP@0.5:0.95/%	帧率/fps
YOLOv8s-Pose	95.9	93.4	83.2	192.30
YOLOv7-Pose	91.6	92.8	79.9	42.92
YOLO-Pose	93.8	95.3	84.2	100.00
YOLOv8+ShuffleNetV2+SA	92.21	92.2	73.5	229
本文	94.3	93.3	74.8	238.10

与原始 YOLOv8s-Pose、YOLOv7-Pose 和 YOLO-Pose 相比,本文算法推理速度最快,分别为其 1.24 倍、5.55 倍和 2.38 倍,在识别精度上本文算法与它们相差不多。与文献[5]的改进 YOLOv8 对比,其采用 ShuffleNetV2 作为主干网络以轻量化模型,并在网络颈部,引入 SA 模块提高识别精度,虽较原始 YOLOv8n-Pose 识别精度和推理速度有所提升,但本文算法在主干网络不变的情况下,使用 EMA 模块利用更丰富的上下文信息,采用 α -InnerSIoU 损失函数和改进的关键点损失函数提升关键点检测精度,并使用 FasterBlock 模块减少冗余特征计算以轻量化网络,在识别精度和识别速度上均优于改进 YOLOv8。

综上所述,本文改进的 YOLOv8n-Pose 算法综合性能更好,能准确并更快地检测出人体关键点,本文算法和 YOLO-Pose 算法识别效果对比如图 8 所示,其中图 8(a)为本文算法识别结果,图 8(b)为 YOLO-Pose 算法识别结果。可看出本文算法对轨道作业人员姿态估计

效果更好。

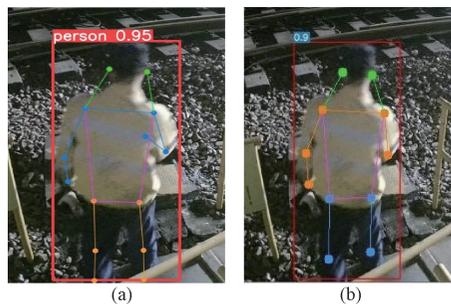


图 8 识别效果对比

Fig. 8 Comparison of recognition effects

最后将本文算法用于识别各安全动作,其识别结果如图 9 所示,其中图 9(a)为左向安全动作识别结果,图 9(b)为右向安全动作识别结果,图 9(c)为前向抬手动作识别结果,图 9(d)为跨轨动作识别结果。

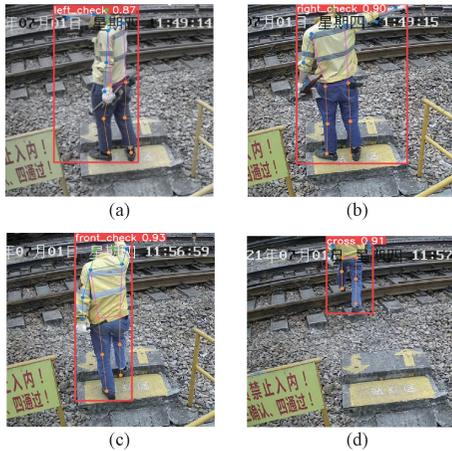


图9 安全动作识别结果

Fig.9 Safety actions recognition results

5 结论

本文针对轨道作业人员跨轨安全动作监督的问题,引入改进的YOLOv8n-Pose算法,其改进方法为在网络的颈部加入EMA注意力机制以提高识别精度,使用Faster-Block模块轻量化模型,使用 α -InnerSIoU作为IoU损失函数,在关键点损失函数引入轴最大距离惩罚项,经实验验证网络性能得到提升。使用高斯滤波和ColorJitter算法增强自制数据集,并将COCO val 2017中部分图像加入数据集以提高模型泛化能力。训练前,使用遗传算法对超参数进行调优,并在训练中使用迁移学习和知识蒸馏技术,提高网络识别精度和鲁棒性,加快网络训练速度。实验结果表明,本文算法识别精度为94.3%,识别速度可达238.1 fps,能准确快速地对轨道作业人员进行姿态识别并根据关键点位置信息识别安全动作,7.6 GFLOPs计算量使网络便于部署到边缘计算设备,使用本文算法可有效对跨轨标准安全动作进行识别和监督。

参考文献

[1] 叶彦斐,侯翔一,胡龙葵,等.一种基于人员停留时间及其路径跟踪的列检上道规范作业监测方法[P].江苏省:CN202310958634.7,2023-10-31.
YE Y F, HOU X Y, HU L G, et al. A method for monitoring the train inspection on-railway standardized operation based on the dwell time of personnel and their path tracking [P]. Jiangsu Province, CN202310958634. 7, 2023-10-31.

[2] 张永强,李胜男,张子强,等. R-YOLO轨道人员目标检测模型[J]. 河北科技大学学报, 2023, 44(6): 580-588.
ZHANG Y Q, LI SH N, ZHANG Z Q, et al. R-YOLO orbital personnel target detection model[J].

Journal of Hebei University of Science and Technology, 2023, 44(6): 580-588.

[3] 叶彦斐,程立,侯翔一.基于改进YOLO v5的钢轨内部伤损B显图像识别与分类[J]. 国外电子测量技术, 2023, 42(12): 70-76.
YE Y F, CHENG L, HOU X Y. Recognition and classification of internal defects in railway tracks based on improved YOLOv5[J]. Foreign Electronic Measurement Technology, 2023, 42(12): 70-76.

[4] 王名赫,徐望明,蒋昊坤.一种改进的轻量级人体姿态估计算法[J]. 液晶与显示, 2023, 38(7): 955-963.
WANG M H, XU W M, JIANG H K. Improved lightweight human pose estimation algorithm [J]. Chinese Journal of Liquid Crystals and Displays, 2023, 38(7): 955-963.

[5] 王小鹏,石欢.改进型YOLOv8融合关键点的跌倒检测算法[J/OL]. 西安电子科技大学学报, 1-16 [2024-07-14]. <https://doi.org/10.19665/j.issn1001-2400.20240403>.
WANG X P, SHI H. Fall detection algorithm based on the improved YOLOv8 combined with key points [J/OL]. Journal of Xidian University, 1-16 [2024-07-14]. <https://doi.org/10.19665/j.issn1001-2400.20240403>.

[6] 王泽杰,沈超敏,赵春,等.融合人体姿态估计和目标检测的学生课堂行为识别[J]. 华东师范大学学报(自然科学版), 2022, 2022(2): 55-66.
WANG Z J, SHEN CH M, ZHAO CH, et al. Recognition of classroom learning behaviors based on the fusion of human pose estimation and object detection [J]. Journal of East China Normal University(Natural Science), 2022, 2022(2): 55-66.

[7] 李一凡,袁龙健,王瑞.基于OpenPose改进的轻量化人体动作识别模型[J]. 电子测量技术, 2022, 45(1): 89-95.
LI Y F, YUAN L J, WANG R. Improved lightweight human action recognition model based on OpenPose[J]. Electronic Measurement Technology, 2022, 45(1): 89-95.

[8] SUN X, XIAO B, WEI F, et al. Integral human pose regression [C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018: 529-545.

[9] YANG Z, WANG R, TAN Y, et al. MALT: Multi-scale action learning transformer for online action detection [J]. Arxiv preprint arXiv: 2405.20892, 2024.

[10] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]. CoRR, 2018,

- abs/1803.01534.
- [11] OUYANG D, HE S, ZHANG G, et al. Efficient multi-scale attention module with cross-spatial learning[C]. ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2023: 1-5.
- [12] ZHANG H, XU C, ZHANG S. Inner-iou: More effective intersection over union loss with auxiliary bounding box[J]. Arxiv preprint arXiv:2311.02877, 2023.
- [13] GEVORGYAN Z. SIoU loss: More powerful learning for bounding box regression [J]. Arxiv preprint arXiv: 2205.12740, 2022.
- [14] CHEN J, KAO S, HE H, et al. Run, don't walk: Chasing higher FLOPS for faster neural networks[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021-12031.
- [15] MAJID, NAGORI S, MATHEW M, et al. YOLO-pose: Enhancing YOLO for multi person pose estimation using object keypoint similarity loss[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 2637-2646.
- [16] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.

作者简介

叶彦斐, 博士, 副教授, 硕士生导师, 主要研究方向为先进控制、智能制造、信息融合、物联网技术、综合自动化。

胡龙葵(通信作者), 硕士, 主要研究方向为深度学习、目标检测、人体姿态估计。

E-mail: 221306060004@hhu.edu.cn