



DOI:10.12404/j.issn.1671-1815.2402551

引用格式:汪发现,石中华,张魁.基于YOLOv8的道岔路口环境下轨道巡检识别技术[J].科学技术与工程,2025,25(13):5476-5483.

Wang Faxian, Shi Zhonghua, Zhang Kui. Track inspection and recognition technology for turnout intersections based on YOLOv8[J]. Science Technology and Engineering, 2025, 25(13): 5476-5483.

自动化技术、计算机技术

基于YOLOv8的道岔路口环境下轨道巡检识别技术

汪发现¹,石中华²,张魁^{3,4*}

(1.北京北九方轨道交通科技有限公司,北京100089;2.长沙润伟机电科技有限责任公司,长沙410006;

3.湘潭大学机械能与力学学院,湘潭411105;4.湘潭大学复杂轨迹加工工艺及装备教育部工程研究中心,湘潭411105)

摘要 社会车辆运行状态、交通灯管控状态及轨道部件的精准识别成为制约轨道巡检机器人推广应用的技术瓶颈。针对轨道健康状态巡检需求,首先给出了轨道巡检机器人的视觉检测系统及基于“北斗+5G”的导航系统的技术方案;其次,基于YOLOv8算法搭建了视觉检测系统模型,并创新性地采用网络爬虫技术从网络开源视频资源中抓取获得有关交通信号灯、汽车尾灯等样本数据来训练视觉检测模型;随后,对训练后的模型,进一步采用迁移学习法和早停法优化模型检测精度。研究表明:采用YOLOv8算法并使迁移学习法和早停法针对模型进行优化后,巡检机器人能有效巡检轨道部件及道岔路口的车辆以及交通信号灯,同时有效提高巡检效率和准确性。

关键词 YOLOV8; 信号灯; 道岔路口; 轨道; 巡检机器人

中图分类号 TP242.6;

文献标志码 A

Track Inspection and Recognition Technology for Turnout Intersections Based on YOLOv8

WANG Fa-xian¹, SHI Zhong-hua², ZHANG Kui^{3,4*}

(1. Beijing North Nine Fang Rail Transit Technology Co., Ltd., Beijing 100089, China;

2. Changsha Runwei Electromechanical Technology Co., Ltd., Changsha 410006, China;

3. School of Mechanical Engineering and Mechanics, Xiangtan University, Xiangtan 411105, China;

4. Engineering Research Center of Complex Tracks Processing Technology and Equipment of Ministry of Education, Xiangtan University, Xiangtan 411105, China)

[Abstract] In the intersection scenario, the running state of social vehicles, the control state of traffic lights and the accurate identification of track components have become the technical bottlenecks restricting the promotion and application of track inspection robots. Aiming at the requirements of track health inspection, firstly, the vision inspection system of the track inspection robot and the technical scheme of the navigation system was presented based on “Beidou + 5G”. Secondly, the vision detection system model was built based on YOLOv8 algorithm, and the web crawler technology was innovatively used to capture sample data about traffic lights and car taillights from open source video resources to train the vision detection model. Then, transfer learning method and early stop method were used to optimize the detection accuracy of the trained model. The research results show that after adopting YOLOv8 algorithm and optimizing the model with transfer learning method and early stop method, the inspection robot can effectively detect the track components, vehicles and traffic lights at the switch junction, and effectively improve the inspection efficiency and accuracy.

[Keywords] YOLOV8; traffic light; crossing; railway; inspection robot

目前,以巡检机器人为代表的巡检系统广泛应用于电力、能源和工业领域,包括变电站、火电站、工厂及航空发动机等重要设备的巡检等工作,巡检机器人可全天候自主巡检,完成上述场合的急、难、

险和重复性工作^[1]。

轨道交通作为现代城市交通的重要组成部分,其安全性和运营效率日益受到重视,轨道交通的安全性和轨道的健康状况密切相关。以往都是由人

收稿日期:2024-04-09 修订日期:2025-01-10

基金项目:湖南省教育厅科学研究项目(24A0126);湖南省自然科学基金面上项目(2023JJ30589)

第一作者:汪发现(1980—),男,汉族,山东济南人,硕士,高级经济师。研究方向:电力机车智能运维技术。E-mail:wfx34@126.com。

*通信作者:张魁(1985—),男,汉族,湖南岳阳人,博士,教授。研究方向:机械设计制造自动化。E-mail:zhangk@xtu.edu.cn。

工巡检轨道,近年来,为解决当前轨道巡检方式过度依赖工人技术和效率低等问题,业界开始把巡检机器人运用到轨道巡检领域。

国外对巡检机器人的研究起步较早,早在 20 世纪 80 年代,东京电力有限公司的泽田教授就成功研发了首款巡线机器人,这标志着巡检机器人技术的诞生;加拿大水电研究院生产的 LineScout 新型巡检机器人,能够在带电的导线上工作,并具备可见光和红外视频的检测功能,为电力巡检提供了强有力的支持。相对于国外,国内巡检机器人的研究起步较晚,但近年来发展迅速。随着国家对智能制造和人工智能技术的重视,国内科研机构和企业纷纷投入巡检机器人的研发和应用中^[2-3]。

虽然国内外的巡检机器人已经实现了从简单巡检到复杂任务处理的转变,但是,如前所述,国内外的巡检机器人大多设计为运行在封闭或半封闭区域内,对于巡检过程包含社会活动区域的巡检机器人涉及相对较少。

铁路轨道虽然大部分为封闭区域,但也存在与社会道路交汇的区域,巡检机器人运行到这类区域必须具有准确检测轨道健康状况的同时检测周边社会车辆的行驶意图及交通信号灯状态并据此评估周边交通状况来提供越障导航信息的能力,这就对巡检机器人的核心能力之一视觉识别检测能力要求非常高^[4]。因此,在道岔路口场景下,精准识别轨道部件健康状况的同时能够自主越障导航成为轨道巡检机器人推广应用的难题。

针对现有运行于封闭或半封闭区域内的巡检机器人在精准检测轨道部件的健康状况同时自主越障导航方面的不足,现提出一种基于“北斗 + 5G + YOLOV8 + 云电脑”的巡检机器人方案设计,以期实现稳定的巡检数据远程高速传输和实时智能分析。

1 巡检机器人总体方案设计

对于巡检机器人而言,视觉检测系统是其在复杂巡检路况环境下实现自主行走、交通流检测、越障导航、故障异常识别与故障诊断等关键功能的核心,通过配合摄像头、云端电脑等组件对实时路况和作业对象进行分析检测,包括交通信号灯红灯与绿灯检测、前车尾灯亮灭检测,轨道健康状况等。为此,搭建了如图 1 所示电力机车巡检机器人用视觉检测系统。视觉检测系统包括如下部分:用于图像采集的摄像头、用于图像处理与检测识别的云电脑、用于信息反馈的机器人控制器;摄像头完成影像采集后,利用 5G 网络技术将视频数据传输给云

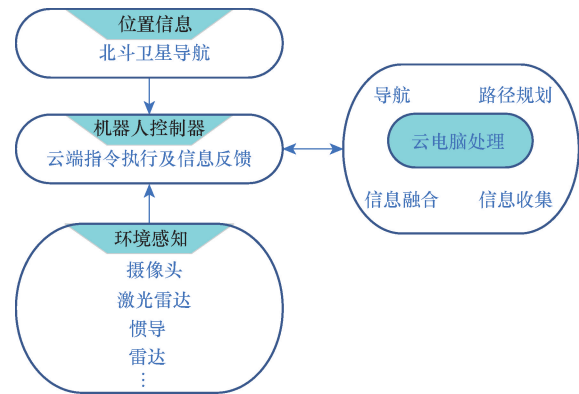


图 1 道岔路口场景下巡检机器人视觉检测方案设计示意图

Fig. 1 A schematic diagram of the overall visual design plan for the inspection robot in the turnout intersection scenario

电脑;由云电脑进行交通影像的视频分帧处理,使动态视频转变为静态图片,将静态图片输入 YOLOv8 网络进行识别检测,识别检测的结果再次利用 5G 网络技术传回巡检机器人的主控系统^[5-7]。

同时,机器人的自主巡航,自主避障,采用“北斗 RTK (Real-Time Kinematic) + 5G”定位的技术方案,能实时获取机器人的位置和方向信息,且定位的精度水平方向达到 3 cm,垂直方向为 6 cm,能满足巡检机器人巡航过程位置精度的要求^[8]。

2 视觉检测模型算法选择

2.1 网络模型选择

目前,目标检测主要采用深度学习模型,包括一级和二级两种网络模型。前者将输入图像分割成 $N \times N$ 图像块,每个图像块有 M 个固定大小的锚定盒,用于识别输出锚定盒的位置和分类标签。因此,一级网络不需要经过区域建议阶段便可直接生成对象的类别概率和位置坐标值,典型算法有 YOLO、SSD 等^[9]。二级网络是与之相对应的区域建议算法,分为传统算法和神经网络;在输入图像中生成目标候选区域,并将所有候选区域输入分类器进行分类^[10]。对比之下,一级网络速度远快于二级网络,故采用一级网络模型检测目标。

2.2 算法选择

进一步选用一级网络中运行速度相对较快的 YOLO 算法开展目标程序设计,其特点介绍如下。

YOLO 算法属于单阶段目标检测,仅需如图 2 所示神经网络结构处理一次图片,即同时预测位置和类别信息,这就使目标检测任务定义为端到端的回归问题,提升了计算速度,其检测过程主要分为预处理、模型推断和后处理 3 个阶段。

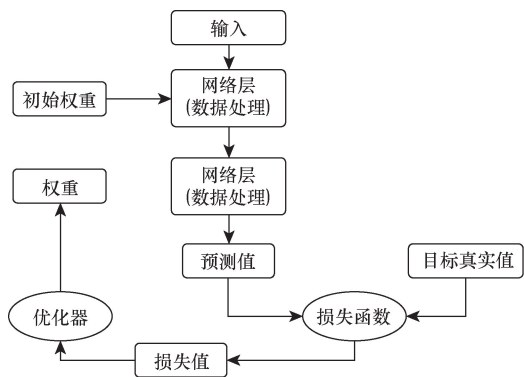


图2 神经网络结构图

Fig. 2 Neural network structure diagram

具体的,选择 YOLOv8 算法,相比于其前身 YOLOv5,网络结构上两者都是 backbone (负责特征提取) + PANet(负责特征融合) + Head(预测输出结果)结构,但 YOLOv8 具有新的骨干网络(如 C3(增强感受野用)模块被替换成了 C2f 模块,更为轻量化)、一个新的 anchor-free 检测头以及新的损失函数,因此可以构建更深层次的网络模型,更好地提取图像特征^[11]。

3 程序实现与验证

为了实现所述巡检机器人在途径道岔路口等真实道路场景下的自主巡检,其前提条件为可靠识别交通信号灯和社会车辆尾灯并判断社会车辆的通行意图,以及准确检测轨道部件健康状况。现以上述基本功能为例,目标程序设计过程如下。

3.1 数据集获取

为了重现巡检机器人在道岔路口等真实道路场景下自主行驶过程中的道路交通状态,以便客观检验巡检机器人的目标识别能力,从视频网站上通过网络爬虫方式抓取由行车记录仪拍摄的行车影像,涉及大连、广东、洛杉矶等国内外城市,时间跨度为 8 个月,视频分辨率为 1 920 × 1 080,视频帧率为 25。随后,筛选获得涵括道岔路口等典型场景下共计 5 911 张的交通图像,作为制作数据集的样本。如表 1 所示,为交通目标样本数量的统计表。表 1 中,有关交通信号灯的图片有 3 724 张,汽车尾灯的图片有 1 596 张;按 9:1 的比例确定训练集与验证集的数量。

表 1 交通目标样本数量

Table 1 Number of traffic target samples			
类别	训练集	验证集	合计
图片	5 320	591	5 911
交通信号灯	3 724	414	4 138
汽车尾灯	1 596	177	1 773

轨道图片的数据集来自网上较为系统的科罗拉多州普韦布洛交通技术中心的约 500 张轨道图片,并按 7:3 的比例分别用于训练集和验证集。

3.2 基于视频分帧技术的图像获取

云电脑运用基于 python 和 Opencv 库的视频分帧程序,将摄像头获取到的行车记录视按每 30 帧一张图片的频次获取图像数据。经分帧处理后的图像大小为 1 920 × 1 800,文件格式为 JPG。以某路口视频为例,经分帧处理后的图片如图 3 所示,仅以对交通状况的视频分帧为例,检测轨道部件状况的视频分帧过程相同。

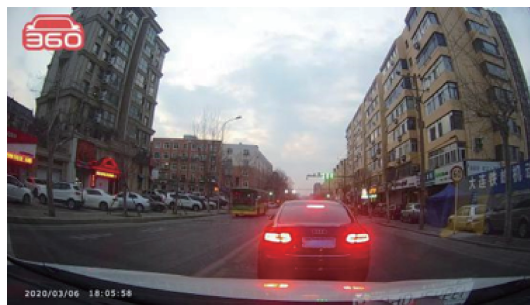


图3 道路路口拍摄视频经分帧程序处理后所获图像效果

Fig. 3 The image effect obtained after processing the video captured at a certain intersection through a frame separation program

3.3 数据集结构设计及图片标注

在建立用于训练网络模型的数据集时,给出了与 COCO 数据集(一个大规模的、丰富的物体检测、分割和字幕数据集)相似的数据文件夹结构,其中: Annotations 文件夹里存在着 xml 文件,其用于记录每张图片已经标注过的物体的类别与位置坐标; JPEGImages 文件夹中包含着所有的未经过标注的原始图片,所有的原始图片都有单独对应的名称; Train.txt 中记录着训练集中图片的列表文件名与路径。选用图片标注软件 labellmg,人工标注图片中目标的类别,并确定物体边框的位置信息。每一次标注一张图片,便会自动生成一个对应的 xml 文件。

4 图像检测与处理过程

4.1 图像检测

模型识别预处理图像识别的第一阶段,准确处理好图像数据对后续非常重要。包括几个步骤,确保用于识别的图片符合模型要求,首先调整好图片的尺寸以符合神经网络预期输入的图像尺寸,归一化像素值,以及判断是否需要转变图像灰度值,或者是否需要图像增强(如对比度和细节是否足够、颜

色信息的重要性、后续处理步骤等)。

模型推理阶段, 即对传入神经网络的图像按照训练得到的权重, 对图像中的特征基于卷积神经网络(convolutional neural networks, CNN)进行提取及预测, 输出一组边界框, 每个边界框都附有关联的对象类标签和置信度分数。图像处理是基于 NVIDIA 的硬件, 而 TensorRT 推理框架具有优秀的低精度推理优化以及很好的对 NVIDIA 的硬件适配, 即使是大量数据也能处理迅速, 延时低适合实时处理数据, 由此选用 TensorRT 推理框架进行模型推理。

后处理阶段主要是对模型推理阶段生成的输出进行处理, 得到一组高置信度的边框, 同时每个框标出检测对象的类, 便于后续分析。

4.2 图像识别结果

利用 3.1 节所述数据训练模型, 训练模型选择从 YOLOv8s 到 YOLOv8l (YOLOv8s 用于速度要求较高的场景, 但检测精度相对较低, YOLOv8l 相反), 同时按照从 0.01、0.001 到 0.000 1 的阶梯式学习率计划^[12]。且从 YOLOv8s 到 YOLOv8l 都进行了 150 批次大小为 8 的训练, 随后的推理阶段, 分数阈值设为 0.25, 交集/联合阈值设置为 0.2。

如图 4 所示为正确识别信号灯及尾灯时输出页面截图。然而, 网络模型也因如下原因, 发生了小概率误检。

(1) 原视频数据中, 交通信号灯的像素群占整个页面画幅比例偏小。



(a) 交通信号灯状况



(b) 尾灯检测刹车灯

图 4 正确识别效果

Fig. 4 Correct recognition effect

(2) 复杂交通道路情况下, 摄像头在采集交通信号灯颜色时出现失真。

(3) 视频采集过程中, 因车辆运动导致的画面模糊, 也会增加识别检测的难度。

(4) 交通信号灯数据集数量不够大, 且对数据集图片进行手工标注时可能产生了错误。

如图 5 所示为巡检机器人正确识别轨道时输出页面截图。类似地, 在部分情况下, 网络模型也出现误判、漏判的情况, 尤其是当视频拍摄同时处于运动状态及不平坦地形时, 摄像头在采集轨道图像时会出现更为严重的失真以及抖动。

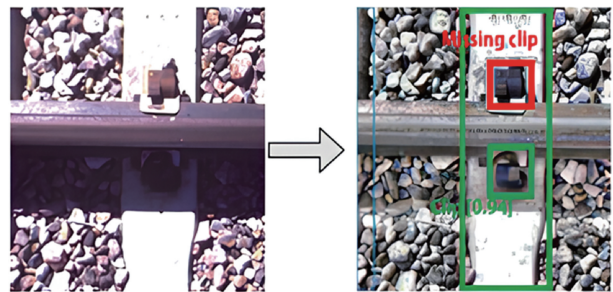


图 5 轨道零件识别效果

Fig. 5 Recognition effect of track components

4.3 识别精度分析

训练模型 (YOLOv8s 到 YOLOv8l) 的选择从准确率、召回率及精确率 3 个方面来判断, 这些需要通过如表 2 所示的混淆矩阵进行计算^[13]。

表 2 混淆矩阵

Table 2 Confusion matrix

矩阵元素	实际为阳性	实际为阴性
	($P = TP + FN$)	($N = FP + TN$)
预测为阳性	事实阳性 (TP)	假阳性 (FP)
预测为阴性	假阴性 (FN)	事实阴性 (TN)

注: P 为同类样本的实际总数; N 为负类样本实际总数; TP 为正确预测的正样本数量; FP 为预测为正样本, 但实际却是负样本数量; FN 为未被预测为正样本, 但实际却是正样本; TN 为未被检测成正样本, 实际也不是正样本。

$$A = \frac{TP + TN}{ALL} \tag{1}$$

$$P = \frac{TP}{TP + FP} \tag{2}$$

$$R = \frac{TP}{TP + FN} \tag{3}$$

式中: A 为准确率; ALL 为所有样本数; P 为精确率; R 为召回率。

以召回率为横坐标, 精确度为纵坐标做出如图 6 所示的 $P-R$ 曲线, 平均精度 (average precision, AP) 就是 $P-R$ 曲线下的面积的值。

最后以平均精度均值 (mean average precision,

mAP)来评估其训练效果,其计算公式如式(4)、式(5)所示,评估结果如表3所示,从(YOLOv8s)到(YOLOv8l)的所有模型的mAP都徘徊在79%,YOLOv8s的mAP在验证集最高,达到79.76%,优于其较大的对应版本YOLOv8m和YOLOv8l,这个结果可能由于大模型虽然增强了学习能力,但更容易出现过拟合。

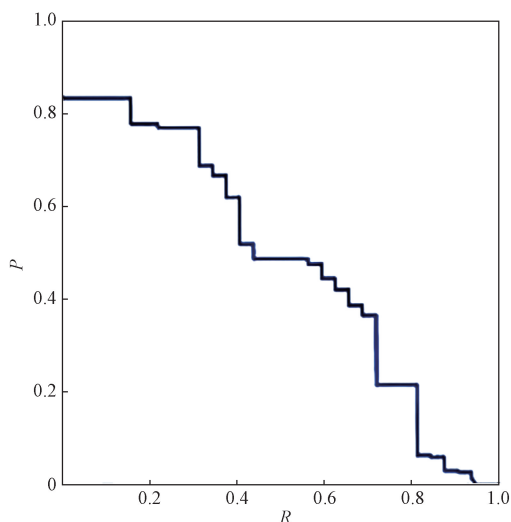


图6 P-R曲线

Fig. 6 P-R curve

表3 YOLOv8模型的验证图和推理速度
Table 3 Verification diagram and inference speed of the YOLOv8 model

项目	YOLOv8s	YOLOv8m	YOLOv8l
mAP/%	79.76	78.75	79.30
延迟/ms	9.60	11.44	13.61

$$AP = \int_0^1 P(R) dR \quad (4)$$

$$mAP = \frac{1}{C} \sum_{i=1}^c AP_i \quad (5)$$

式中: c 为类的数量; AP_i 为类的平均精度; P_i 为某个框的真实标签; $P(R)$ 为某个召回率 R 下模型的精度。

在完成网络模型的建立后,对识别精度进行评估分析是优化机器学习模型的关键^[14]。为此,基于Loss函数,对前述所获预测值和真实值之间的关系进行了定量分析。Loss函数包括类别分类损失和边框回归损失。在YOLOv8中,类别损失采用多类别交叉熵损失,其中每个预测框输出的是所有类别的概率分布,模型通过最小化预测类别和真实类别之间的交叉熵来优化类别预测精度。对于边框回归损失,YOLOv8采用DFL(distribution focal loss)和CIoU(generalized intersection over union) Loss。DFL通过优化边框的概率分布来提高回归精度,CIoU则

结合了目标框的中心距离、长宽比和重叠区域,进一步优化边框的预测精度^[15-16]。DFL loss如式(6)所示,CIoU Loss如式(7)所示。

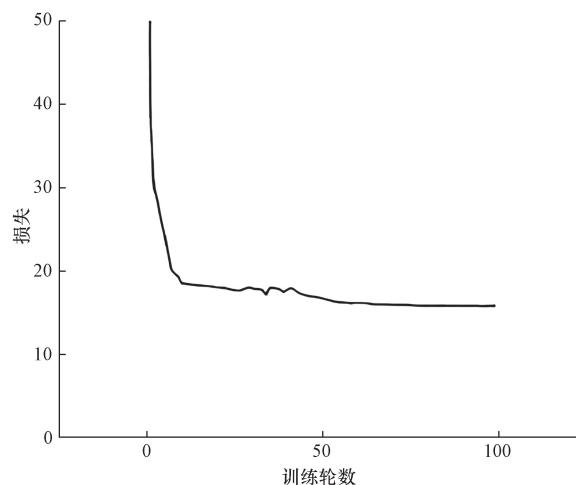
$$L_{DFL} = - \sum_{i=1}^c p_{i,c} \ln(\hat{p}_{i,c}) (1 - \hat{p}_{i,c}) \quad (6)$$

式(6)中: L_{DFL} 为损失; $p_{i,c}$ 为真实类别 c 的标签值(1或0); $\hat{p}_{i,c}$ 为模型预测的类别概率值。

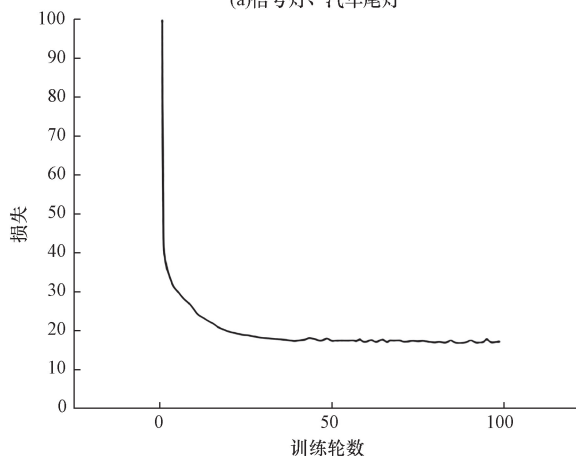
$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \quad (7)$$

式(7)中: IoU 为预测框与真实框的交并比; α 和 v 为调节因子,用于调整宽高比损失项的影响; $\rho^2(b, b^{gt})$ 为预测框和真实框的中心点距离的平方,表示目标框位置的偏差; c^2 为预测框和真实框外接框的对角线长度的平方,用于衡量框的尺度差异。

同时,在训练日志中,可以分别得到如图7所示信号灯、汽车尾灯模型及轨道模型在算法模型下对应的Loss曲线。



(a)信号灯、汽车尾灯



(b)轨道图

图7 交通信号灯和汽车尾灯识别 Loss 曲线变化图
Fig. 7 Loss curve change diagram for identification of traffic lights and car tail lights

由图 7(a)可知,在前 20 轮训练中,所建网络模型的损失值随着训练轮数的增加而呈现下降趋势,这说明训练轮数的增加使得信号灯、汽车尾灯的预测值和真实值的差距变小,网络模型正在进行学习进化;随后,损失值开始产生明显收敛,并几乎停留维持在 15% 之上。由图 7(b)可知,在前 40 轮训练中,损失值随着训练轮数的增加而逐渐下降;之后随着训练轮数增加,损失值产生明显收敛。

对比图 7(a)与图 7(b)可知,有关信号灯、汽车尾灯的机械学习效果相对较好;这主要是由于信号灯、汽车尾灯的数据集更大;此外,如前所述,由于摄像头在运动且地形不平坦过程采集轨道图像时会出现更为严重的失真与抖动,导致标注难度大,最终导致识别难度相对增加。

5 模型优化

5.1 组合优化策略

因为巡检过程光照条件和阴影都会影响到视觉识别效果,因此有必要对图像进行增强和优化,采用对比度有限自适应直方图均衡(contrast limited adaptive histogram equalization, CLAHE)作为图像数据增强技术,以减少上述问题的影响^[17]。

机械学习需要用到大量精准标注的目标物体图片,而制造一个全面高质量的数据集会花费大量的时间和精力。所建网络模型采用了由几千张标注图片构成的数据集进行训练,数据集本身偏小也是导致 Loss 增加的原因。考虑到现有被练过的且具有不同分布的类似数据集,考虑到训练成本以及提高识别精度,采用迁移学习法将已被训练好的模型数据转移至所建模型中。目前,主流的迁移学习法包括迁移学习(transfer learning)、特征提取(extract feature)和模型微调(tune tune),其中:transfer learning 为将预训练模型全部冻结,提取的特征是放弃预训练模型,根据计算得到预训练模型卷起层的特征向量,然后分别定制全连接网络;Fine tune 是对前两种方法的优化,其只冻结预训练模型的前几个卷积层,训练全连接层和相邻输出层的卷积层^[18]。

随着神经网络的不断训练,模型泛化性能得到增强,数据开始拟合。几乎所有的深度学习神经网络都有可能发生过拟合,且数据集越大,过拟合现象发生概率反而越大。为此,选用早停法优化,步骤如下:仍按如前述比例选取训练集和验证集;利用训练集训练时,每 10 个阶段计算训练模型在验证集上的性能;当验证集的性能低于上次测试时,停止训练;最后一次迭代的结果参数作为模型的最终参数。

5.2 优化结果分析

如图 8 所示为信号灯、汽车尾灯模型及轨道模型经优化后所获 Loss 变化曲线图。

对比图 8、图 7(a)与图 7(b)可知,在使用 Fine-tune 和早停法后,信号灯、汽车尾灯识别损失在 100 之后由 15% 下降到了 5%,也即识别准确率提高了 10%。类似地,轨道模型经优化后的损失在 100 轮训练后大幅度下降,于 110 轮训练时进入收敛状态,识别准确率提高了 5%。

如图 9~图 11 所示,为模型优化前后,交通信号灯、尾实际识别效果对比图以及优化后轨道图像效果对比图。可知,对于同一行车记录仪拍摄的同一帧视频,优化后对交通信号灯检测的置信度增加,同时对远距离目标检测的效果提升;优化后对汽车尾灯的检测准确率提升明显,经 CLAHE 优化后的轨道检测图片也更为清晰。综上,利用迁移学习和早停法相结合,对所建立的具有较小规模数据集的网络模型的优化效果十分明显。

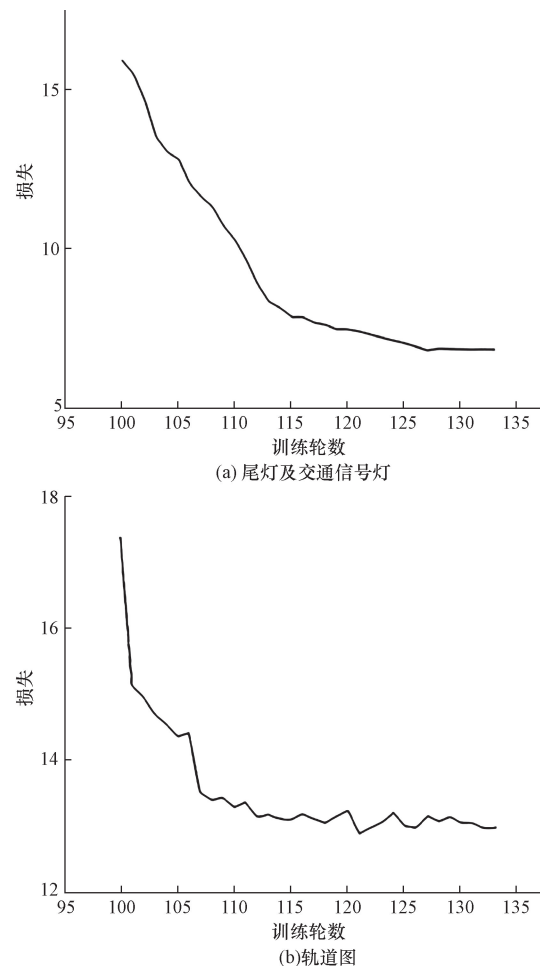


图 8 优化后轨道部件、尾灯及交通信号灯识别 Loss 曲线变化图

Fig. 8 Change in Loss curve for optimized identification of track components, tail lights, and traffic lights



(a)优化前

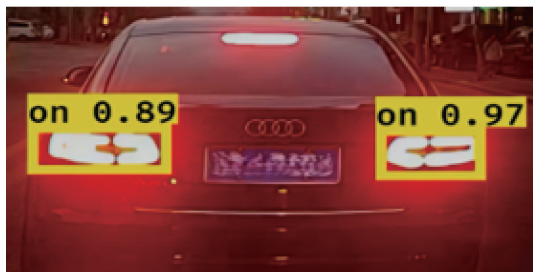


(b)优化后

图9 优化前后信号灯识别效果对比图
Fig. 9 Comparison of traffic light recognition effect before and after optimization



(a)优化前

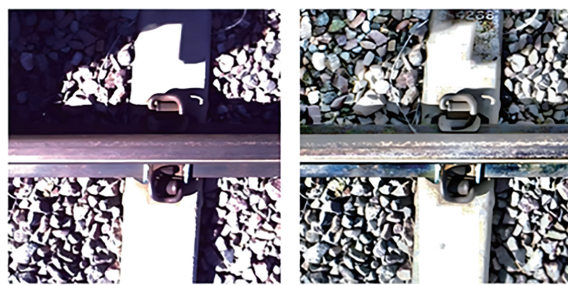


(b)优化后

图10 优化前后尾灯识别效果对比图
Fig. 10 Comparison of tail light recognition effect before and after optimization

6 结论

面向未来全天候危险恶劣环境下长期实时巡检需求,具有视觉检测能力和自行走能力的轨道巡检机器人能够变“原有被动靠问题驱动巡检”为“主



(a)优化前

(b)优化后

图11 优化前后轨道图像效果对比图
Fig. 11 Comparison of track image effect before and after optimization

动智能发现问题”,并能够提升设备管理效能及缺陷管理的及时性,还能降低巡检人员人身伤害风险。然而,在电力机车运行环境下,尤其是在道岔路口或交通信号灯交叉口等复杂交通控制区域内,车辆状态识别技术以及对轨道部状的检测能力是电力机车巡检机器人必须具备的关键能力之一。基于 YOLOv8,开展轨道巡检机器人视觉识别技术研究,主要结论如下。

(1)在设计和实施针对道岔路口车辆状态识别系统时,需考虑多因素,包括视觉设备性能、检测算法准确性以及与其他交通控制和监测系统的集成程度。

(2)采用 YOLOv8 算法检测目标并使迁移学习和早停法针对模型进行优化后可以提高信号灯、汽车尾灯识别准确率 10%,以及提高轨道识别准确率 5%,有效提升道岔路口状态下识别准确率。

(3)研究成果可提供电力机车更好检测轨道部件状态,周围社会车辆运动状态和行为意图的信息,有助于提高电力机车响应速度和运行安全性,推动电力机车自动驾驶研究。

参考文献

- [1] 黄山,吴振升,任志刚,等. 电力智能巡检机器人研究综述[J]. 电测与仪表, 2020, 57(2): 26-38.
Huang Shan, Wu Zhensheng, Ren Zhigang, et al. A review of research on intelligent power inspection robots [J]. Electrical Measurement & Instrumentation, 2020, 57(2): 26-38.
- [2] Alhassan A B, Zhang X D, Shen H M, et al. Power transmission line inspection robots: a review, trends and challenges for future research[J]. International Journal of Electrical Power & Energy Systems, 2020, 118(5): 26-32.
- [3] Nagarajan B, Li Y, Sun Z Y, et al. A routing algorithm for inspecting grid transmission system using suspended robot: enhancing cost-effective and energy efficient infrastructure maintenance [J]. Journal of Cleaner Production, 2019, 219(8): 622-638.
- [4] Li R H, Shen Y. YOLOSR-IST: a deep learning method for small target detection in infrared remote sensing images based on super-resolution and YOLO[J]. Signal Processing, 2023, 208(2): 108-

- 124.
- [5] Guo F, Qian Y, Shi Y F. Real-time railroad track components inspection based on the improved YOLOv4 framework[J]. *Automation in Construction*, 2021, 125(4): 22-29.
- [6] 张威, 张鑫中, 王丛佼, 等. 改进灰狼算法的变电站巡检机器人路径规划[J]. *重庆理工大学学报(自然科学)*, 2023, 37(6): 129-135.
- Zhang Wei, Zhang Xinzong, Wang Congjiao, et al. Path planning for substation inspection robot based on improved grey wolf optimization algorithm[J]. *Journal of Chongqing University of Technology (Natural Science Edition)*, 2023, 37(6): 129-135.
- [7] Aydin I, Sevi M, Akin E, et al. A deep learning-based hybrid approach to detect fastener defects in real-time[J]. *Tehnički Vjesnik*, 2023, 30(5): 1461-1468.
- [8] Wang K, Liu M Z, Ye Z J. An advanced YOLOv3 method for small-scale road object detection[J]. *Applied Soft Computing*, 2021, 112(1): 108-119.
- [9] Hong F, Lu C H, Liu C, et al. A traffic surveillance multi-scale vehicle detection object method base on encoder-decoder[J]. *IEEE Access*, 2020, 8(12): 47664-47674.
- [10] Xu W Y, Cui C, Ji Y C, et al. Faster R-CNN: YOLOv8-MPEB small target detection algorithm based on UAV images[J]. *Heliyon*, 2024, 10(8): 28-35.
- [11] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017, PP(99): 2999-3007.
- [12] Yuan Z D, Zhu S Y, Yuan X C, et al. Vibration-based damage detection of rail fastener clip using convolutional neural network: experiment and simulation, engineering failure analysis [J]. *Vehicle System Dynamics*, 2021, 119(1): 191-210.
- [13] Qiu S, Cai B X, Wang W D, et al. Automated detection of railway defective fasteners based on YOLOv8-FAM and synthetic data using style transfer, automation in construction[J]. *Automation in Construction*, 2024, 162: 105-112.
- [14] Redmon J, Farhadi A. YOLOv3: an incremental improvement [J]. *ArXiv Preprint ArXiv*: 2018, 18(4): 57-67.
- [15] Bakirci M. Utilizing YOLOv8 for enhanced traffic monitoring in intelligent transportation systems (ITS) applications [J]. *Digital Signal Processing*, 2024, 152: 32-44.
- [16] 徐慧智, 宋爱秋, 武笑宇. 基于均匀设计的船舶目标检测深度学习模型训练方法[J]. *科学技术与工程*, 2022, 22(25): 11241-11249.
- Xu Huizhi, Song Aiqiu, Wu Xiaoyu. Training method for deep learning model of ship target detection based on uniform design [J]. *Science Technology and Engineering*, 2022, 22(25): 11241-11249.
- [17] Wu Y, Qin Y, Qian Y, et al. Hybrid deep learning architecture for rail surface segmentation and surface defect detection [J]. *Computer-aided Civil and Infrastructure Engineering*, 2022, 37(2): 227-244.
- [18] Dong X, Yan S, Duan C. A lightweight vehicles detection network model based on YOLOv5[J]. *Engineering Applications of Artificial Intelligence*, 2022, 113(6): 74-86.