常新, 毛永青, 董玉华. 基于 YOLOv5 的盲道及坑洞检测系统设计[J]. 智能计算机与应用,2026,16(5):88-93. DOI:10.20169/j.issn.2095-2163.24090403

# 基于 YOLOv5 的盲道及坑洞检测系统设计

常 新, 毛永青, 董玉华
(大连民族大学 信息与通信工程学院, 大连 116600)

**摘 要**: 盲人出行问题是社会关注热点。针对目前应用中盲道及坑洞检测系统功能单一和准确率较低的问题,本文采用 YOLOv5 深度学习网络模型,设计了对盲道及坑洞可进行识别的系统。通过实景拍摄和网络爬虫下载等方式构建出本设计的盲道及坑洞数据集,训练得到的盲道及坑洞检测模型可以完成识别盲道和坑洞的任务。将训练出的模型部署到 Jetson Nano 开发板中,结合深度相机对道路状况进行实时检测,并能按照系统识别结果进行相应的语音提示。实验结果表明,该系统可以准确识别出盲道及坑洞的位置,具有较高的准确率,可以在一定程度上为盲人出行提供辅助作用。

**关键词**: 深度学习; YOLOv5; 特征提取; 盲道及坑洞检测
**中图分类号**: TP391.41 **文献标志码**: A **文章编号**: 2095-2163(2026)05-0088-06

## Design of tactile paving and pothole detection system based on YOLOv5

CHANG Xin, MAO Yongqing, DONG Yuhua
(College of Information and Communication Engineering, Dalian Minzu University, Dalian 116600, Liaoning, China)

**Abstract**: The issue of blind people's travel is a hot spot of social concern. Aiming at the problem of single function and low accuracy of the tactile paving and pothole detection system in current applications, a system that can identify tactile paving and potholes is designed based on YOLOv5 deep learning network model. The tactile paving and pothole dataset for this project is constructed through real-life shooting and web crawler downloading. The trained model can accomplish the task of recognizing the tactile paving and pothole. Deploying the trained model into the Jetson Nano development board, it is combined with the depth camera to detect the road conditions in real time and provide corresponding voice prompts according to the system recognition results. The experimental results show that the system can accurately identify the locations of tactile paving and pothole with high accuracy, which can provide assistance for the blind to travel to a certain extent.

**Key words**: deep learning; YOLOv5; feature extraction; tactile paving and pothole detection

## 0 引 言

随着需要视力辅助支持的人群逐渐增多以及交通的快速发展,盲人出行安全问题越来越受到社会的关注,导盲产品的需求不断扩大[1]。作为城市道路中重要的无障碍设施,盲道对于视力障碍人群的出行安全有着重要的保障意义,因此在导盲产品的发展过程中,盲道、坑洞的检测至关重要。在实际生活中,盲道设施存在着不连续或道路不平等情况,盲道被占、破损是普遍现象,该现象严重影响了视障人士的出行[2]。因此,使用具有盲道和坑洞检测功能的导盲产品十分重要。目前应用中的盲道或坑洞检测方法,大多存在检测精确度偏低或功能单一等问题,无法有效解决盲人出行安全问题[3-4]。因此,设计一个精准检测盲道及坑洞位置的系统迫在眉睫。

计算机视觉技术提高了目标检测的准确性和实时性,通过深度学习和卷积神经网络方法,计算机视觉技术能够自动提取图像中的特征,从而更准确地识别和定位目标物体[5-6]。相比传统的手工特征提取方法,具有更好的泛化能力和更高的检测精度,可以准确实时地进行盲道和坑洞的检测。本文利用计算机视觉技术,采用 YOLOv5(You Only Look Once version5)深度学习网络模型[7-9],设计出可以准确识别出盲道和坑洞的系统,并结合语音提示,在出行

**基金项目**: 辽宁省大学生创新训练项目(202412026492)。
**作者简介**: 常 新(2002—),男,学士。主要研究方向:图像信号处理;毛永青(2003—),男,本科生。主要研究方向:图像信号处理。
**通信作者**: 董玉华(1978—),女,硕士,副教授。主要研究方向:智能信息处理与应用。Email:dongyuhua@dlnu.edu.cn。
**收稿日期**: 2024-09-04

问题上为盲人带来一定便利。

## 1　YOLOv5 模型架构

深度学习领域中的目标检测算法主要可以归类为两大类：两阶段（Two Stage）方法和单阶段（One Stage）方法。本文中应用的 YOLOv5 算法是常见的 One Stage 算法，相较于两阶段算法结构简单，速度更快。相较于 YOLOv4 网络模型而言，YOLOv5 具有轻量级、速度快、精度高的特点，支持根据自己的需求自定义数据集和模型训练，是一种准确、高效、轻便的检测算法[10]，适合应用于本文。YOLOv5 的网络结构主要由四部分组成，分别是输入部分（Input）、主干网络（Backbone）、颈部网络（Neck）和头部网络（Head）[11]。YOLOv5 网络结构如图 1 所示。

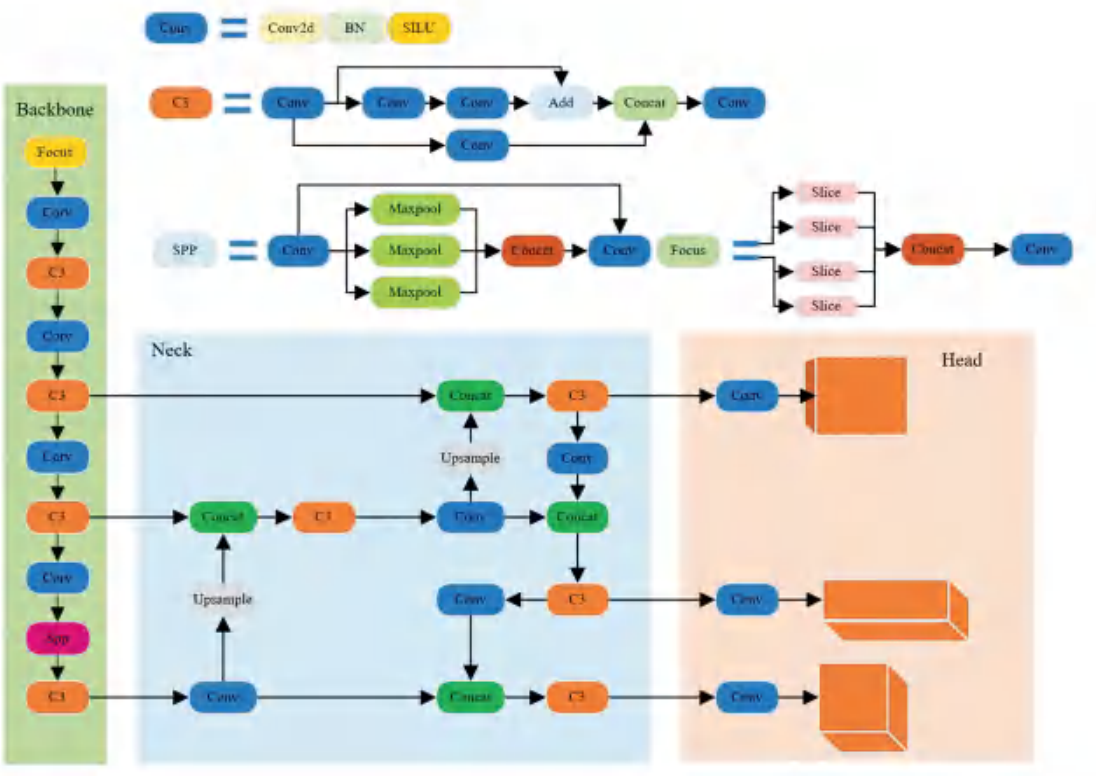



**图 1　YOLOv5 网络结构图**

**Fig. 1　YOLOv5 network structure diagram**

### 1.1　输入部分

在 YOLOv5 网络结构中，输入（Input）部分主要负责图像数据的读取、解压、预处理和增强。其将原始的输入图像转化为多层特征图，以便后续的目标检测任务使用。在这一部分，YOLOv5 采用 Mosaic 数据增强技术[12]、自适应锚框计算技术、自适应图片缩放技术来提高模型的性能。

### 1.2　主干网络

主干网络（Backbone）是用于特征提取的主要网络结构，负责从输入图像中提取特征，为后续的目标检测任务提供基础[13]。与 YOLOv4 相比，YOLOv5 的 Backbone 新增了 Focus 结构[14]。YOLOv4 主干网络仅使用了单一 CSP 结构。而 YOLOv5 设计了两种 CSP 结构，其中 CSP1_x 结构应用于 Backbone 主干网络，通过共享权重减少计算量，CSP2_x 结构应用于 Neck 模块。

### 1.3　颈部网络

颈部网络（Neck）是连接 Backbone 和 Head 的中间模块，核心作用是对 Backbone 提取的特征进行进一步融合、提炼，为 Head 模块提供优化后的特征进行目标检测[15]。与 YOLOv4 相同，YOLOv5 的 Neck 部分采用了特征金字塔网络（FPN）与路径聚合网络（PAN）结构，实现多层特征图的整合和优化，提高检测准确性与效率。

### 1.4　头部网络

头部网络（Head）负责完成最终的回归预测，主要处理 Neck 模块传递的特征图，执行分类和回归任务，生成目标物体的类别概率和位置坐标。其基于锚框进行检测，通过计算预测框与真实框的交并比来筛选和输出最终的检测结果。检测结果包括目标框的位置、大小和类别等信息。

## 2 模型训练及性能评估

### 2.1 数据集

由于没有公开的盲道数据集,本文通过网络爬虫下载,实景拍摄等方式制作盲道及坑洞数据集,其中包含 2 392 张图像。对标注完的 2 392 张图像进行训练,得到的模型识别准确率较低。经分析是数据集规模不足引发了过拟合。为此,本文通过缩放、水平翻转等数据增强方式对数据集进行扩大,最终得到了包含 5 158 张图像的数据集。

为提升模型的泛化能力和鲁棒性,需对数据集进行划分:训练集用于模型训练,验证集用于在训练过程中评估模型性能以选择最优模型配置。本课题数据集总量为 5 158 张,将数据集的 80%设置为训练集,共 4 126 张。20%设置为验证集,共 1 032 张。保证了模型训练及测试的准确性与合理性[16]。

### 2.2 评估指标

TP(True Positive):实际为正例,预测为正例的样本数。

FP(False Positive):实际为负例,预测为正例的样本数。

TN(True Negative):实际为负例,预测为负例的样本数。

FN(False Negative):实际为正例,预测为负例的样本数。

根据这 4 个统计量计算出对应的精确率和召回率[17]。

精确率(Precision):预测为正例的样本中,真实为正例的概率,反映预测结果的正确性。公式如下:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{1}$$

召回率(Recall):真实值为正例的样本中,被正例预测为正例的概率,反映模型对正样本的识别能力。公式如下:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{2}$$

平均精确度(Average Precison,AP):用于计算单个类别在不同召回率下的平均精确率。公式如下:

$$\text{AP} = \sum_{k=0}^{k=n-1} [\text{Recall}(k) - \text{Recall}(k+1)]\text{Precision}(k) \tag{3}$$

mAP:对所有类别的 AP 值求平均值,用于衡量检测模型的整体性能。公式如下:

$$\text{mAP} = \frac{1}{n}\sum_{k=1}^{k=n} \text{AP}_k \tag{4}$$

mAP@0.5:0.5 表示 IoU(Intersection over Union)阈值,当预测框与真实框的交并比大于 0.5 时,判定该检测结果正确。

### 2.3 模型训练及仿真分析

本文采用 YOLOv5s.pt 预训练权重开展模型训练,输入图像尺寸为 640×640,批次数量 batch size 为 8,迭代次数 epoch 为 100。按照本文训练需求,设置好文件路径、训练集和验证集图片路径,根据自己的数据集配置模型参数后开始训练,训练之后的理论效果如图 2 所示。



图 2 理论效果图

Fig. 2 Theoretical rendering

图 3 表示精确率与召回率之间的关系图，随着召回率的增加，精确率逐渐下降[18-19]。分析图 3 可知，本模型对坑洞的检测性能最佳，达到 0.908。对盲道的检测性能表现较好，达到 0.842。总体来看，检测性能较为理想。

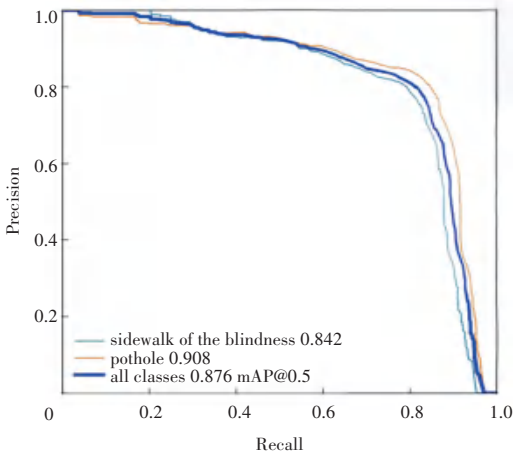



**图 3　精确率-召回率（*P*-*R*）曲线**

**Fig. 3　Precision-Recall（*P*-*R*）curve**

损失函数是衡量模型预测值与实际值之间差异的一个指标，其值越低通常意味着模型的预测效果越好[20]。图 4—6 为训练过程中的损失函数曲线，展示了本课题中盲道和坑洞检测模型的训练过程。

图 4 为训练集中 bounding box 回归损失曲线，该曲线用于评估模型预测的边界框位置与真实边界框位置之间的误差。较低的“train/box_loss”表示模型通过预测精确的边界框坐标来准确定位盲道和坑洞。随着训练轮次的增加，回归损失逐渐降低，整体下降趋势明显，说明通过训练，本模型定位较准确，且未出现明显过拟合现象。

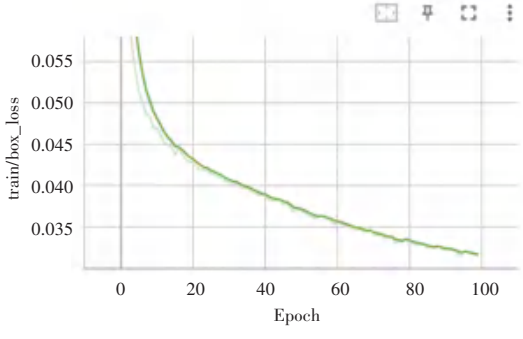



**图 4　训练集边界框回归损失（box_loss）曲线**

**Fig. 4　Training set bounding box regression loss（box_loss）curve**

图 5 为训练集分类损失（train/cls_loss）曲线，该曲线用于评估模型对每个目标类别概率估计的准确性。最小化“train/cls_loss”表明模型对盲道及坑洞实现了正确分类。训练前期损失快速下降，最终趋于稳定，表明模型具备良好的目标分类。

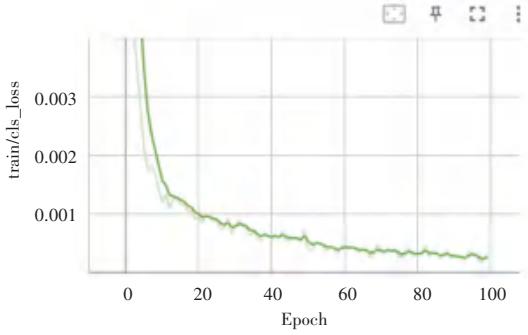



**图 5　训练集中分类损失（cls_loss）曲线**

**Fig. 5　Training set classification loss（cls_loss）curve**

图 6 为训练集目标置信度损失（objectness_loss）曲线，该曲线用于评估模型对目标存在性的识别准确率。

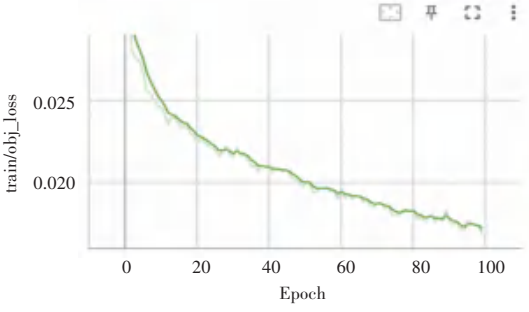



**图 6　训练集目标置信度损失（obj_loss）曲线**

**Fig. 6　Training set obj_loss（obj_loss）curve**

曲线的下降表明模型能有效地将盲道及坑洞与其他物体或背景区域区分开来。训练初期下降速度较快，后期趋于稳定，表明模型已完成对检测目标的有效学习。整体而言，该模型在训练过程中 3 种损失均随着迭代次数的增加而收敛，降低后趋于平稳，表明模型检测性能在逐步提升。

本课题模型在验证集上损失曲线的变化趋势如图 7—9 所示，反映了验证过程与盲道、坑洞检测模型的准确性。图 7 为验证过程中的方框回归损失（val/box_loss）曲线。该曲线用于衡量预测边界方框坐标与地面实际盲道及坑洞坐标之间的差异。方框回归损失越小，说明模型对盲道及坑洞位置的定位越准确。初始损失较大，随着轮数的增加，最终趋于平稳状态。

图 8 为验证过程中的分类损失（val/cls_loss）曲线，该分类损失曲线衡量的是模型为检测到的物体分配正确类别盲道和坑洞标签的准确性，较低的分类损失表明模型能够正确识别盲道及坑洞，并将其与其他物体类别有效区分开来，最终曲线保持在较低的水平，这表明模型的分类能力表现良好。

图 9 为 val/obj_loss 缺失曲线,图描述了验证过程中的对象损失,代表了模型准确区分盲道及坑洞区域的能力。曲线呈下降趋势,表明该模型在验证集上可以很好地检测对象。

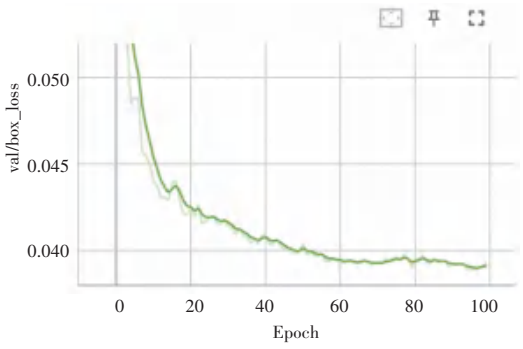



**图 7 验证集边界框回归损失(val/box_loss)曲线**

**Fig. 7 Validation bounding box regression loss(val/box_loss) curve**

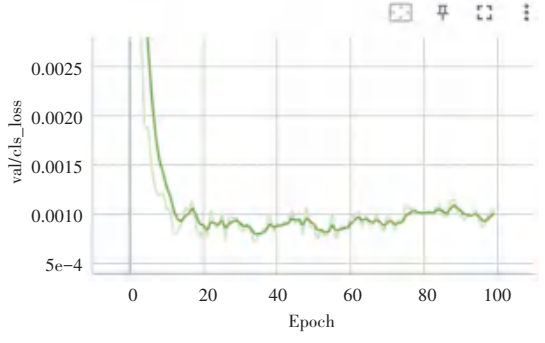



**图 8 验证集分类损失(val/cls_loss)曲线**

**Fig. 8 Validation set classification (val/cls_loss) loss curve**

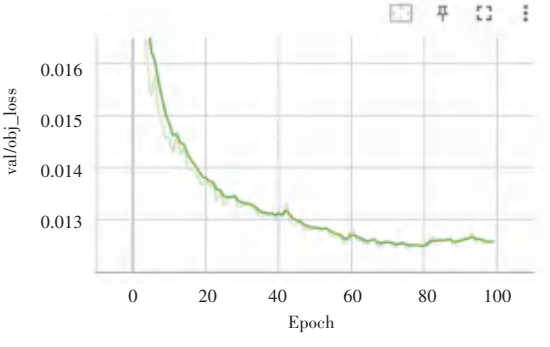



**图 9 验证集目标置信度损失(val/obj_loss)损失曲线**

**Fig. 9 Validation set objectness loss(val/obj_loss) curve**

整体而言,模型在验证集上的各项损失随着训练轮次增加而收敛,都呈先下降后平稳的趋势,可见该模型的泛化能力较好,并未出现明显的过拟合。

## 3 模型边缘端部署

### 3.1 硬件平台选择

边缘节点通常具有有限的计算和存储资源,因此需要根据具体要求选择合适的硬件。本课题需要使用深度相机进行图像捕捉,该设备的运行离不开 GPU 的支持,Jetson Nano 在这类场景下表现尤为出色,可支持深度学习任务,实用性更强,因此本课题选择使用 Jetson Nano B01 作为一款高性能的嵌入式计算平台,Jetson Nano 拥有 4 核 GPU 与 128 个 NVIDIA CUDA 核心,是一款高性能、低功耗、丰富接口、易于扩展的特点,完全适配本设计的需求。

### 3.2 总体系统方案

本系统采用深度相机进行实时图像采集,在 Jeston Nano 中经过视频流处理以及目标检测之后获取周围环境信息,进行模型分析。扬声器根据识别结果进行相对应的语音提示。本系统如图 10 所示。

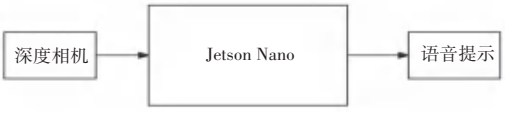



**图 10 系统框图**

**Fig. 10 System block diagram**

### 3.3 系统验证

将训练完的模型在 Jetson Nano 板中部署完毕后,进行系统测试。由图 11 可知,当系统识别出盲道后,发出语音提示“前方道路是盲道”。



**图 11 检测盲道效果图**

**Fig. 11 Effect picture of tactile paving detection**

由图 12 可知,当系统检测出坑洞后,系统会发出语音提示“前方有坑洞,请小心”。




**图 12 检测坑洞效果图**

**Fig. 12 Effect picture of pothole detection**

## 4 结束语

针对盲人出行问题,本文采用 YOLOv5 深度学

习网络模型，设计了一个盲道及坑洞检测系统，构建出盲道及坑洞数据集并完成模型训练。该模型可实现盲道与坑洞的位置检测，将其部署在 Jetson Nano 开发板中可以根据具体识别情况，给予盲人相应的语音提示，为其出行带来便利。实验结果表明，本设计能实现高准确率的盲道及坑洞识别，可以在短时间内处理和分析采集到的图像数据，及时反馈结果、发出语音提示，有助于保障盲人的出行安全。目前，本设计仍有改善空间，可继续深度优化模型性能，提高模型的泛化能力以及识别对象的准确率。

## 参考文献


[1] 牛恒，陈智利，周泉，等. 基于图像处理技术的盲人避障系统研究与实现[J]. 自动化技术与应用，2023，42(1)：1-4. DOI：10.20033/j.1003-7241.(2023)01-0001-05.

[2] 张展浩，刘昊杰，文浩然，等. 基于互联网的智慧盲道导航系统的设计[J]. 黑龙江科学，2024，15(2)：122-124.

[3] 郑芷涵，许旭兵. 基于无障碍理念的智能导盲背心设计研究[J]. 西部皮革，2024，46(3)：84-86. DOI：10.20143/j.1671-1602.2024.03.084.

[4] 顾素梅，王鑫，李雨桐. 感官代偿在盲人产品设计中的应用研究[J]. 设计，2023，36(11)：76-79. DOI：10.20055/j.cnki.1003-0069.000849.

[5] AHN H，SON S，ROH J，et al. SAFP-YOLO：Enhanced object detection speed using spatial attention-based filter pruning[J]. Applied Sciences，2023，13(20)：268.

[6] 李锐，王鑫，杨威朋，等. 计算机视觉在道路成像和坑洞检测中的应用研究综述[J]. 汽车工程师，2023(9)：1-8. DOI：10.20104/j.cnki.1674-6546.20230308.

[7] 刘昕斐，张荣芬，刘宇红，等. 基于 YOLOv5s 的导盲系统障碍物检测算法[J]. 智能计算机与应用，2023，13(11)：220-226.

[8] CHUNLING Y. Computer interactive gesture recognition model based on improved YOLOv5 algorithm[J]. International Journal of Computational Intelligence Systems，2024，17(1)：735.

[9] WEN H，LI Y，WANG Y，et al. Usage of an improved YOLOv5 for steel surface defect detection[J]. Materials Testing，2024，66(5)：726-735.

[10] 李岳，夏益民，蔡述庭，等. 基于 YOLOv5s 的盲人出行辅助系统研究[J]. 工业控制计算机，2023，36(12)：15-17.

[11] 郑凯东，马莉娜. 基于改进 YOLOv5 模型的安全帽佩戴检测方法研究[J]. 智能计算机与应用，2024，14(7)：140-144. DOI：10.20169/j.issn.2095-2163.240721.

[12] 刘思诚，李嘉琛，邓皓，等. 基于 YOLOv5 改进的小目标检测算法[J]. 兵工自动化，2022，41(12)：78-82.

[13] 陈倩倩，王欢，朱敏. 基于 YOLOv5 的盲道和斑马线检测[J]. 信息技术与信息化，2022(7)：10-14.

[14] 李在瑞，郑永果，东野长磊. 基于 YOLOv5 的高分辨率遥感图像目标检测算法[J]. 智能计算机与应用，2023，13(10)：83-87.

[15] 李雨诗. 基于场景识别与障碍物检测的导盲系统研究[D]. 武汉：华中师范大学，2022.

[16] 陈建瑜，邹春龙，王生怀，等. 改进 YOLOv5 的路面缺陷快速检测方法研究[J]. 电子测量技术，2023，46(10)：129-135. DOI：10.19651/j.cnki.emt.2211195.

[17] 冉险生，李锐，贺帅. 基于改进 YOLOv5s 的道路障碍物检测算法[J]. 电子测量技术，2023，46(22)：177-185. DOI：10.19651/j.cnki.emt.2313079.

[18] 蔡鑫楠，丁学文，张子怡，等. 基于 YOLOv5 的铁轨异物入侵检测[J]. 智能计算机与应用，2022，12(11)：161-167.

[19] 段中兴，王剑，丁青辉，等. 基于深度学习的盲道障碍物检测算法研究[J]. 计算机测量与控制，2021，29(12)：27-32.

[20] 周研逸，周月娥，沈琳芸，等. 基于改进 YOLOv5 的路面坑洞检测设计[J]. 现代计算机，2024，30(6)：61-64.