

DOI: 10.13382/j.jemi.B2508117

基于机器视觉的煤矿多场景关键目标检测方法研究*

苏国用^{1,2,3} 帅洪铎³ 邓海顺^{1,2,3} 王鹏或^{2,3} 赵东洋³ 庞子金³

(1. 安徽理工大学煤炭无人化开采数智技术全国重点实验室 淮南 232001; 2. 安徽理工大学矿山智能技术与装备省部共建协同创新中心 淮南 232001; 3. 安徽理工大学机电工程学院 淮南 232001)

摘要:针对煤矿井下复杂作业场景中高粉尘、低照度、人机多目标混杂与跨尺度变化等因素导致作业人员与装备的目标检测效果不佳问题,提出一种基于机器视觉的煤矿多场景关键目标检测方法。首先,采用 CGNet (context guided network) 特征提取模块、SlimNeck 特征融合模块与 Dyhead 动态检测头对 YOLOv5s 算法进行优化,以构建 YOLOv5s-CSD 网络模型。其次,基于自建煤矿数据集,围绕 YOLOv5s-CSD 模型开展消融实验、对比实验与嵌入式检测实验。实验结果表明,在煤矿井下掘进、支锚、采煤与辅助运输 4 种复杂作业场景中, YOLOv5s-CSD 的检测精度达 91.0%,相较于 YOLOv5s 算法提升了 3.5%,并且其与 YOLOv9s、YOLOv11s、YOLOv12s 等 6 种主流目标检测算法相比,模型复杂度适中且检测精度最高。在实验测试平台上, YOLOv5s-CSD 模型对工作人员、支护装置、电机车等 7 类关键目标的实时检测精度均在 90.0% 以上,并且其实时检测速度达 38.6 fps,检测精度高且实时性强,可为煤矿井下复杂环境的视觉动态感知提供技术支撑。

关键词: 复杂作业场景; 机器视觉; 目标检测; YOLO; 视觉动态感知

中图分类号: TP183; TN919.5 **文献标识码:** A **国家标准学科分类代码:** 440.6099

Research on multi-scene key target detection method for coal mine based on machine vision

Su Guoyong^{1,2,3} Shuai Hongxin³ Deng Haishun^{1,2,3} Wang Pengyu^{2,3} Zhao Dongyang³ Pang Zijin³

(1. State Key Laboratory of Digital Intelligent Technology for Unmanned Coal Mining, Anhui University of Science and Technology, Huainan 232001, China; 2. Collaborative Innovation Center for Mining Intelligent Technology and Equipment, Anhui University of Science and Technology, Huainan 232001, China; 3. School of Mechanical Engineering, Anhui University of Science and Technology, Huainan 232001, China)

Abstract: Aiming at the problem of poor target detection of operating personnel and equipment due to high dust, low illumination, human-machine multi-target mixing and cross-scale changes in the complex operation scene of coal mine underground, we propose a multi-scene key target detection method based on machine vision for coal mine. Firstly, the YOLOv5s algorithm is optimised using CGNet (context guided network) feature extraction module, SlimNeck feature fusion module with Dyhead dynamic detection head in order to construct the YOLOv5s-CSD network model. Secondly, based on the self-constructed coal mine dataset, ablation experiments, comparison experiments and embedded detection experiments were carried out around the YOLOv5s-CSD model. The experimental results show that YOLOv5s-CSD achieves a detection accuracy of 91.0% in four complex operation scenarios of underground coal mine tunneling, anchor support, coal mining, and auxiliary transport, which is 3.5% higher than YOLOv5s algorithm, and compared with six mainstream target detection algorithms, such as YOLOv9s, YOLOv11s, and YOLOv12s, it has the moderate model complexity and the highest detection accuracy. On the experimental test platform, the real-time detection accuracy of YOLOv5s-CSD model for seven types of key targets, such as person, support, and electric locomotive, is above 90.0%, and its real-time detection speed is up to

收稿日期: 2025-01-10 Received Date: 2025-01-10

* 基金项目: 安徽省高等学校科学研究项目 (2022AH050834)、国家自然科学基金项目 (52304166)、深部煤矿采动响应与灾害防控国家重点实验室开放基金 (SKLMRDPC22KF24)、安徽理工大学矿山智能技术与装备省部共建协同创新中心开放基金 (CICJMITE202206)、安徽理工大学引进人才科研启动基金 (2022yjrc61)、安徽省高校优秀科研创新团队项目 (2022AH010052)、安徽理工大学研究生创新基金 (2025cx2056) 项目资助

38.6 frames/s, which is high in detection accuracy and real-time, and it can provide technical support for the visual dynamic perception of the complex environment of underground coal mines.

Keywords: complex operational scenarios; machine vision; target detection; YOLO; visual dynamic perception

0 引言

煤炭是全球重要的基础能源之一,当前煤炭工业正转向高质量发展阶段,煤矿智能化感知技术是煤炭工业高质量发展前提与保障^[1-3]。煤矿智能感知系统中由于视觉感知具有直观性、实时性、非接触性等特点,广泛应用于皮带煤流与异物检测^[4]、作业人员定位^[5-6]、轨道障碍物识别^[7]等众多领域,其对加快煤矿智能化发展,提升煤矿安全生产水平具有重要意义。

目前,煤矿井下作业场景中对关键目标视觉感知的方法主要有传统机器学习与深度学习两种。传统机器学习为基于小样本数据在多尺度图像上提取特征向量,再采用分类器进行学习判别。华钢等^[8]提出了一种基于DM642的多目标运动状态感知系统,并利用多阈值分割法、帧间差分法与边缘检测法实现了对多种监控目标的运动状态感知。王勇^[9]设计了一种基于背景差分法的井下运动人员检测算法,从而在序列视频中实现对井下工作人员的高效检测与移动跟踪。巩固等^[10]提出基于运动视频的光流场目标视觉识别方法,实现煤矿灾后复杂环境中救援机器人对障碍物的快速识别定位与实时环境监测。然而上述方法会生成大量冗余的建议框,导致学习效率低下,难以捕捉目标对象的高级语义特征和有效信息。相较于传统机器学习算法,深度学习算法能够将输入图像中的像素数据转化为更高阶、更抽象化的层级特征,从而更有效的对关键目标对象进行识别与定位,并逐步成为煤矿视觉感知的重要手段^[11-12]。

基于深度学习的目标检测算法可分为单阶段检测算法和双阶段检测算法,其中,双阶段检测算法包括 Faster R-CNN^[13]、Mask RCNN^[14]、R-FCN^[15]等,但双阶段检测算法存在模型参数冗余、算法复杂度高、检测速度慢等缺陷,导致其难以在煤矿井下应用。而以 SSD^[16]、RetinaNet^[17]、YOLO 系列^[18-19]等为代表的单阶段目标检测算法以其模型简单、泛化能力强、检测速度快等优势,更适用于煤矿井下采煤、掘进、辅助运输、支锚等作业场景,以满足轻量化、实时性、高精度等需求。Xue 等^[20]提出一种 ResNet18-YOLOv3 煤矸石检测算法,以解决人工煤矸分选劳动强度大、分选效率低等问题,实现煤矸石分选机器人对煤矸石的实时精准分选。李飞等^[21]以 YOLOv4 为基线模型,通过引入 MobileNetv3 网络、空间域注意力机制构建了轻量级 ECSMv3_YOLOv4 检测模型,提升了煤矿井下输送带纵向撕裂状态的检测精度。Zhao

等^[22]提出了一种融合 C3-Dense 模块、解耦头部、ECIoU 损失函数和加权双向特征金字塔网络的多目标检测与跟踪算法,解决了煤矿掘进工作面运动模糊、多目标混合等恶劣环境因素造成的关键目标对象实时感知问题,实现了对关键目标对象的实时精准感知。Pan 等^[23]提出一种融合 GhostNet 卷积模块、参数整流线性单元激活函数(rectified linear unit, RELU)的 YOLOv5s 目标检测算法,以解决煤矿综采工作面液压支架的相对位置检测难题。Yang 等^[24]提出了一种融合 InceptionNeXt 块、新空间金字塔池(new spatial pyramid pooling, NSPP)模块、基于特殊任务的上下文解耦检测头(TSCODE)的 YOLO-Region 区域障碍物检测方法,解决了煤矿井下无人驾驶电机车因障碍物检测不准确所引起的频繁启停问题。王宏伟等^[25]以 YOLOv7-tiny 为基线模型,通过引入自适应图像增强(image adaptive, IA)模块和注意力机制 SimAM 构建 IA-SimAM-YOLOv7-tiny 模型,以解决煤矿复杂环境中的锚护孔位定位难题,实现了机载式锚杆钻机的精准锚固作业。薛小勇等^[26]在 YOLOv8n 检测算法的基础上,通过引入 DSConv 模块、极化自注意力机制与 4 层检测输出模块,以解决采掘工作面的小目标检测难题;Xu 等^[27]提出一种融合 t 特征提取模块 C2f-FasterNe 与注意力模块 SimAM 的 YOLO-CFS 煤流识别算法,实现了煤矿输送带的节能控制并提升了煤流识别精度。张婧等^[28]针对煤矿工作面喷雾除尘场景下的模糊目标检测难题,提出一种基于 SAC-YOLOv9 的安全帽识别算法,利用监督空洞卷积获取不同尺度的感受野,并最终提高煤矿复杂作业背景中的安全帽识别精度。

近年来,国内外众多研究机构和学者提出了多种研究方法,提高了煤矿井下视觉感知能力,但受煤矿井下复杂环境影响,多场景视觉感知仍存在以下问题。1)由于煤矿井下复杂工作面存在高粉尘浓度、低光照强度的恶劣环境因素,造成关键目标成像出现颜色纹理信息缺失、轮廓边缘模糊的问题,不利于检测模型对关键目标复杂特征信息的学习与提取;2)煤矿井下移动终端的存储资源有限,传统检测模型通过自顶向下与自低向上传递信息时重复计算中间特征图,导致模型参数量和计算量大幅增加,不利于轻量化检测模型的部署;3)煤矿井下复杂工作环境中采煤机、掘进机与电机车在工作过程中位置多变且作业姿态各异,传统检测模型的检测头存在多尺度特征感知与表达能力不足问题,导致各关键目标的分类与定位损失急剧上升。

为解决上述难题,实现煤矿井下采煤、掘进、支锚、辅

助运输等多种复杂作业场景中关键装备与作业人员的实时监测。根据煤矿井下关键目标的特性与研究现状,提出了一种基于机器视觉的煤矿多场景关键目标检测算法(YOLOv5s-CSD)。

1) 采用特征提取模块(CGNet)^[29]替换原 YOLOv5s 算法主干网络中的普通卷积模块,利用其局部-全局的残差学习机制与上下文引导策略,增强主干网络对作业人员、支护装置、电机车(person、support、E-locomotive)等目标关键特征信息的学习与提取能力。

2) 引入颈部特征融合模块(SlimNeck)^[30]对原算法颈部网络的特征融合方式进行优化,通过 GSConv (grouped sparse convolution) 模块的通道混洗策略与 VoVGSCSP (VOVNet-GSConv) 模块的一次聚合策略,降低检测模型参数量和计算量,并提高模型对关键目标的检测精度。

3) 采用动态检测头(Dyhead)^[31]替换原算法的耦合检测头,利用其多维感知注意力模块以增强检测头的信息感知与表达能力,以提升检测网络在煤矿复杂背景中对关键检测目标的分类与定位精度。

1 煤矿多场景关键目标检测方法

1.1 整体架构

由于煤矿井下采煤、掘进、支锚与辅助运输等作业场景存在高粉尘、低照度、人机多元素混杂分布以及目标跨尺度变化等复杂煤矿环境^[32-33],导致针对关键装备与作

业人员的检测过程中常出现误检、漏检及检测精度低等问题,极大地影响了煤矿井下的安全高效生产。为解决上述问题,本文提出一种基于机器视觉的煤矿多场景关键目标检测方法,该方法的整体架构如图 1 所示。煤矿关键目标检测方法主要分为数据样本采集、检测网络构建与嵌入式平台测试 3 个部分。首先,利用煤矿井下的工业计算机调用多台矿用本安型防爆摄像头对采煤、掘进、支锚与辅助运输 4 种作业场景进行全时段、全方位的视频数据采集,并将其整理转化为煤矿复杂作业场景数据集,用于后续的网络训练与测试;其次,将数据集输入到所构建的 YOLOv5-CSD 检测模型中,并利用特征提取、特征融合与检测输出 3 个子网的迭代训练进行学习并重构视觉数据中多样化的特征信息;最后,将训练完成的检测网络迁移部署到实验测试平台,以实现煤矿井下作业场景中作业人员、支护装置、电机车等关键(支护架、机载锚杆钻机、作业人员等)目标的实时监测。

1.2 YOLOv5s-CSD 模型

YOLOv5 系列算法是 Ultralytics 公司所推出的 YOLO 系列算法的基础版本,其稳定性突出,后续 YOLOv8、YOLOv11 与 YOLOv12 检测模型也均以其为基线模型进行改进。YOLOv5 根据不同的宽度因子和深度因子可划分为 n、s、m、l、x 5 种大小不同的目标检测模型,其中, YOLOv5s 是 5 种模型中的轻量化版本,适用于资源受限的计算环境。该模型主要由 3 部分组成,分别为主干网络(backbone)、颈部网络(neck)和检测头(head),其中,主干网络通过卷积层和轻量化的 CSP 模块逐层提取图

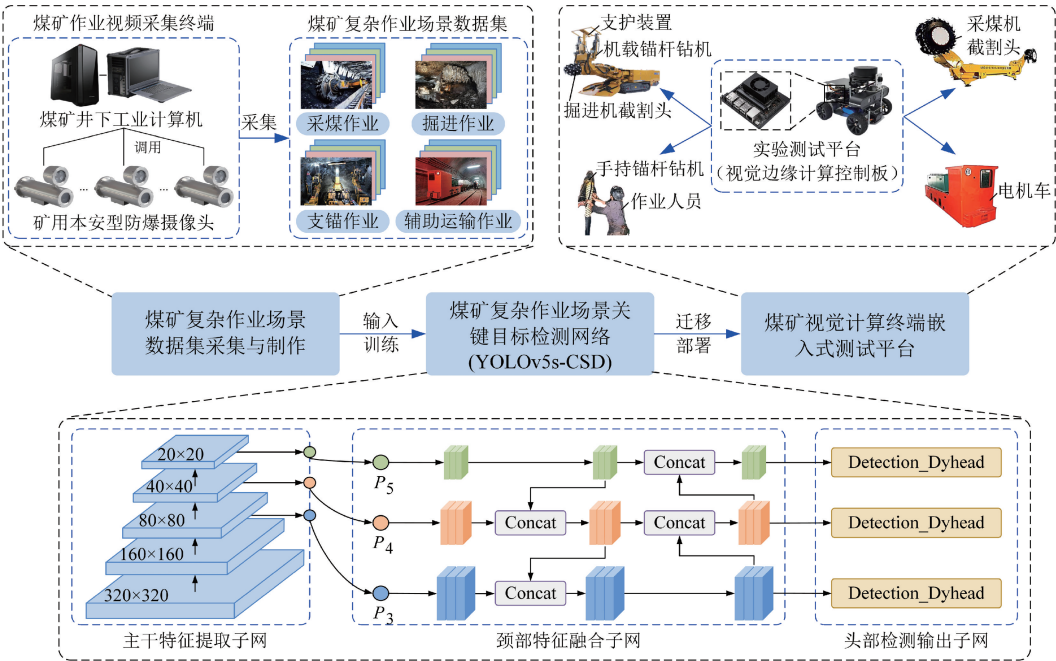


图 1 基于机器视觉的煤矿多场景关键目标检测方法整体架构

Fig. 1 Overall architecture of machine vision-based multi-scene key target detection method for coal mine

像的多尺度特征;颈部网络采用 FPN+PAN 结构双向融合不同层级的特征;Head 通过 3 个检测特征层输出目标检测结果。由于煤矿井下移动终端的存储资源有限,且对检测模型的稳定性具有较高要求,因此,本文选用兼具稳定性与轻量化特点的 YOLOv5s 模型作为基线模型,采用 CGNet 主干特征提取模块、SlimNeck 颈部特征融合模块与 Dyhead 动态检测头 3 种优化策略以提升 YOLOv5s 算法的检测性能,并最终构建出 YOLOv5s-CSD 检测模型。YOLOv5s-CSD 检测模型的整体架构如图 2 所示。

1) 特征提取模块优化,采用 CGNet 模块替换原 YOLOv5s 算法主干网络中的卷积模块 CBS,从而利用其局部残差学习(local residual learning, LRL)与全局残差学习(global residual learning, GRL)方式将目标局部特征与周围环境特征相关联,以得到联合特征信息,并通过上下文引导策略对联合特征进行加权融合,进一步增强主干

网络的特征表达能力。

2) 特征融合网络优化,采用 SlimNeck 中的 GSConv 模块与结合了 GSConv 和 CSP(cross-stage partial network)结构的 VoVGSCSP 模块分别代替颈部中的 Conv 模块和 C3 模块,从而利用 GSConv 将标准卷积(standard convolution, SC)与深度可分离(depthwise separable convolution, DSC)所提取的特征信息进行通道混洗,以降低空间维度压缩和通道维度扩展所造成的部分语义信息损失,并通过 VoVGSCSP 的一次聚合策略实现多层次特征图的跨阶段融合,以提升颈部网络的特征融合能力。

3) 检测输出模块优化,采用 Dyhead 替换原 YOLOv5s 算法的耦合检测头,并利用 Dyhead 中的尺度感知注意力模块、空间感知注意力模块和任务感知注意力模块增强检测头的多维信息感知能力,以提升网络在煤矿复杂背景中对关键检测目标的分类与定位精度。

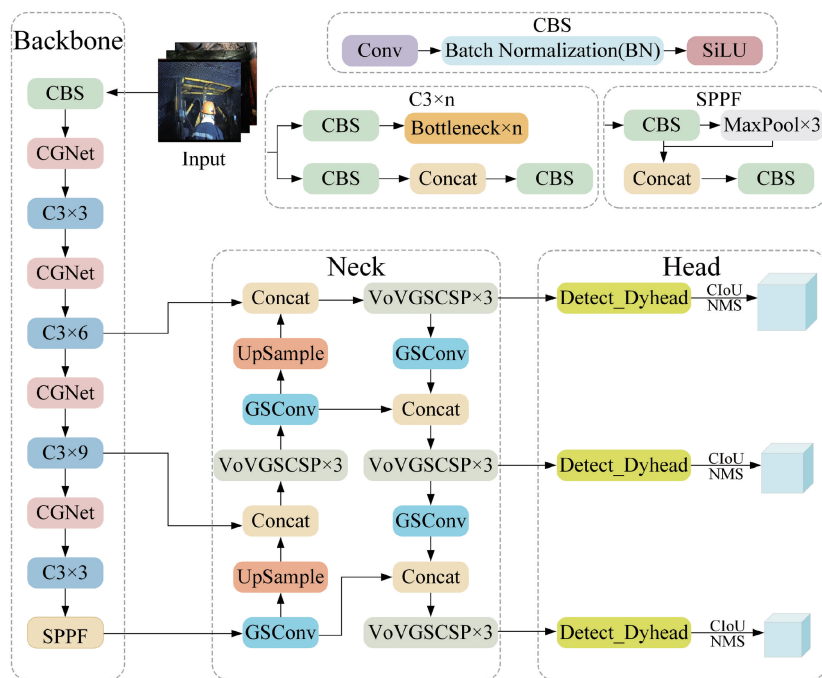


图 2 YOLOv5s-CSD 整体架构

Fig. 2 Overall architecture of YOLOv5s-CSD

2 YOLOv5s-CSD 模型优化策略

2.1 CGNet 特征提取模块

针对煤矿复杂工况条件所导致的目标颜色纹理信息缺失、轮廓边缘模糊、特征复杂度上升等问题,仅凭借原 YOLOv5s 主干网络中堆叠的普通卷积块,难以在煤矿复杂背景中对目标的关键特征信息进行有效提取。为解决上述问题,本文采用 CGNet 特征提取模块替换原 YOLOv5s 主干网络中的 CBS,以凭借其局部-全局的残差

学习机制与上下文引导策略,实现对图像局部与全局特征的关联学习并增强对目标关键特征信息的表达能力。

CGNet 特征提取模块的整体架构如图 3 所示,其主要由 3×3 Conv 模块、CGblock(context guided block)模块、 1×1 Conv 模块和 Upsample 模块构成,其整体运算步骤可分为 3 个阶段。

1) 通过堆叠 3 个 3×3 Conv 的标准卷积层充分提取目标特征并将特征图分辨率下采样至原来的 $1/2$,以逐步降低图像的空间维度并提升计算效率;

2) 利用两组相同数量的 CGblock 模块,对输入图像

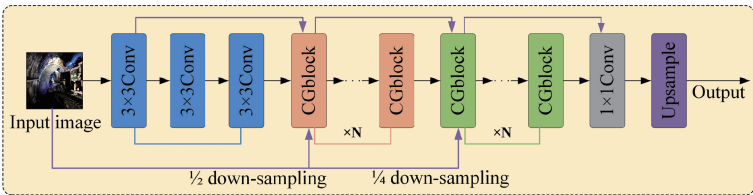


图 3 CGNet 结构框架
Fig. 3 Structural framework of CGNet

进行连续两次下采样并将其与上一层输入特征结合,以利用特征重用方法在保留低层次细节特征的同时进一步提取高层次的抽象特征信息;

3) 采用 1×1 Conv 与 Upsample 模块分别从通道维度与空间维度对所有的特征图进行整合,以实现图像增强特征信息的全方位重建。

此外,如图 4 所示,CGblock 是 CGNet 的核心组成部

分,其利用局部与全局的残差学习方式,将局部目标特征和周围环境特征整合为联合特征,并通过重新加权方式对上下文的联合特征信息进行融合输出,以增强模型的特征表达能力。其中,CGblock 使用了两种类型的剩余连接,即局部残差学习 LRL 与全局残差学习 GRL,二者分别利用联合特征提取器 $f_{jio}(\ast)$ 和全局特征提取器 $f_{glo}(\ast)$ 对输入特征与提炼特征进行拼接。

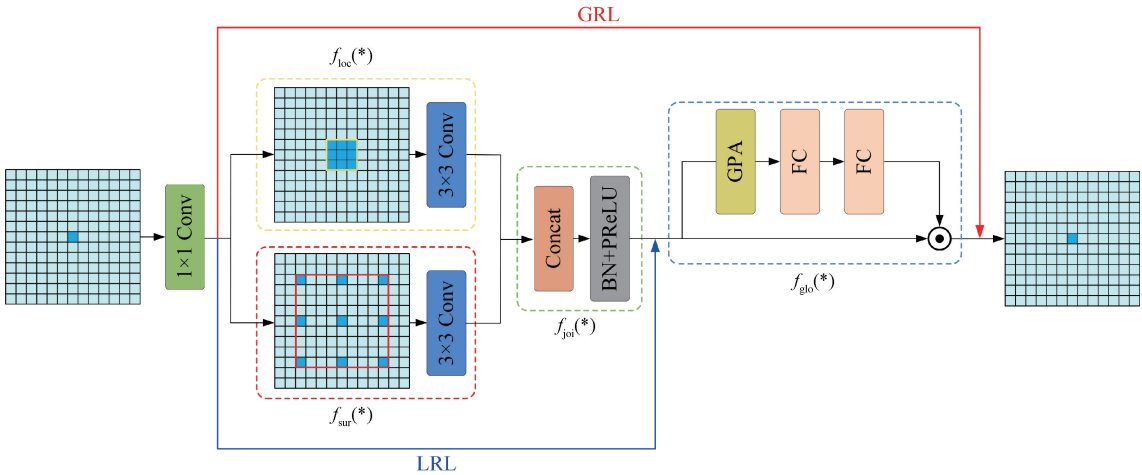


图 4 CGblock 结构框架
Fig. 4 Structural framework of CGblock

CGblock 具体运行步骤如下。

1) 利用 1×1 Conv 调整图像特征维度并进行初步特征提取。

2) 将局部特征提取器 $f_{loc}(\ast)$ 实例化为一个 3×3 的标准卷积层,并从相邻特征向量中学习局部特征,以捕捉到图像局部区域的特征信息;同时将周围上下文提取器 $f_{sur}(\ast)$ 实例化为一个 3×3 的空洞卷积层,从而获得更广泛的周围区域特征信息。

3) 采用联合特征提取器 $f_{jio}(\ast)$ 对 $f_{loc}(\ast)$ 和 $f_{sur}(\ast)$ 输出的局部特征和周围上下文特征信息进行拼接,并通过批量归一化模块 (batch normalization layer, BN)^[34] 与 PReLU 激活函数^[35] 处理后输出。

4) 利用全局特征提取器 $f_{glo}(\ast)$ 中的全局平均池化 (global average pooling, GAP) 操作和两层全连接 (fully

connected, FC) 操作捕捉输入联合特征中的上下文信息与非线性关系,并通过点积运算实现对联合特征信息的加权融合与输出。

2.2 SlimNeck 特征融合网络

由于原 YOLOv5s 颈部网络所采用的路径聚合网络 (path aggregation network, PANet) 结构^[36] 虽然能够通过自底向上与自顶向下路径传递信息,但跨阶段融合信息能力不足,且煤矿井下移动终端的存储资源有限,从而对检测模型的参数量与复杂度提出了更高的要求。为解决上述问题,本文采用 SlimNeck 网络对颈部的特征融合方式进行优化,并通过 GSConv 模块的通道混洗策略与 VoVGSCSP 模块的一次聚合策略降低语义特征信息损失,并实现多层次特征图的跨阶段融合。其中,GSConv 的整体架构如图 5 所示。

1) 通过对输入特征图进行标准卷积操作,得到通道数为 $C_2/2$ 的特征图,普通卷积计算量计算公式如式(1)所示。

$$GFLOPs = W \times H \times K \times K \times C_{in} \times C_{out} \quad (1)$$

式中: W 和 H 分别为特征图宽度和高度; K 为卷积核大小; C_{in} 和 C_{out} 分别为输入和输出特征通道数;

2) 通过对上一操作得到的特征图进行深度可分离卷积操作,得到另一个通道数为 $C_2/2$ 的特征图,深度可分

离卷积计算量的计算公式如式(2)所示。

$$GFLOPs = W \times H \times K \times K \times 1 \times C_{out} \quad (2)$$

3) 将 SC 部分的特征图和 DSC 部分的特征图按通道数进行拼接,然后通过通道混洗操作输出一定通道数的目标特征图,GSConv 卷积的计算公式如式(3)所示。

$$GFLOPs2 = W \times H \times K \times K \times 1 \times \frac{C_{out}}{2} (C_{in} + 1) \quad (3)$$

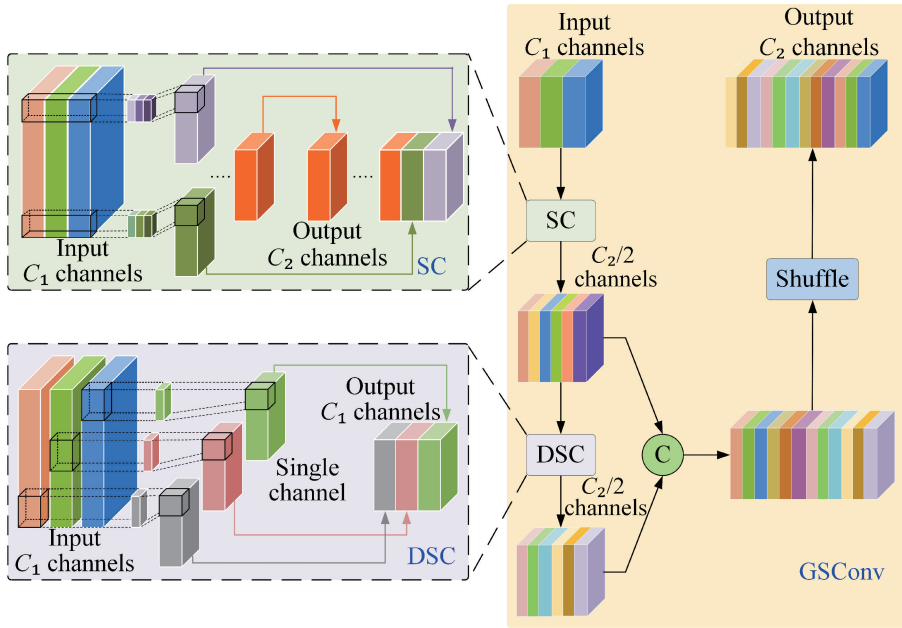


图 5 GSConv 结构框架

Fig. 5 Structural framework of GSConv

VoVGSCSP 模块主要由 SC 模块、GSbottleneck 模块、Concat 模块组成,其能将不同阶段的特征图信息进行有效融合提升颈部网络特征融合能力,V0VGSCSP 的整体结构如图 6 所示。

1) 通过对输入的特征图进行标准卷积操作,得到通道数为 $C_1/2$ 的特征图。

2) 利用 GSbottleneck 模块对 $C_1/2$ 特征图进行优化梯度处理,其利用标准卷积分支与双重 GSConv 分支分别对特征图进行处理,并通过加法操作对两条分支输出的特征图进行聚合,以得到通道数为 C_2 的特征图。

3) 将输入部分经过标准卷积操作后的特征图与 GSbottleneck 部分的特征图按通道数进行拼接,然后通过标准卷积操作输出通道数为 C_2 的目标特征图。

2.3 Dyhead 检测头

由于煤矿井下复杂工作环境中采煤机、掘进机与电机车(辅助运输电机车)位置多变且作业姿态各异,而 YOLOv5s 中原检测头表达能力和多尺度特征检测能力不足,无法有效降低关键目标的定位损失与分类损失。为

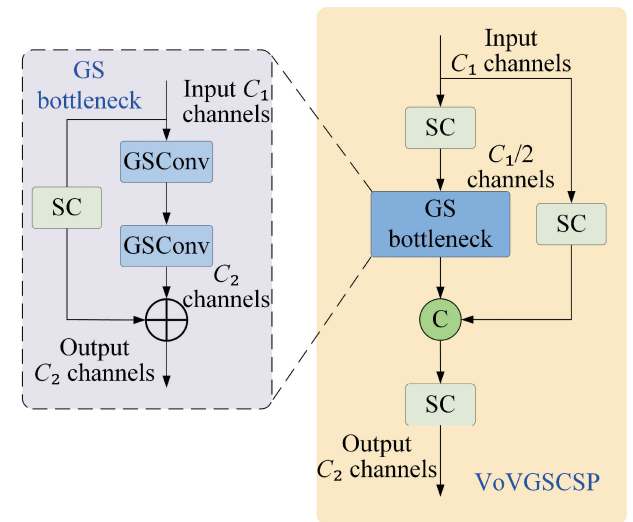


图 6 VoVGSCSP 结构框架

Fig. 6 Structural framework of VoVGSCSP

解决上述问题,本文采用头部 Dyhead 改进:采用一种具

有多维感知能力的动态检测头 Dyhead 替换原检测头,有效解决了原检测头表达与检测能力不足的问题。Dyhead 是一种具有多维感知能力的全新检测头,特征图自上而

下依次通过尺度感知注意力模块(π_L)、空间感知注意力模块(π_S)和任务感知注意力模块(π_C)完成对特征图的感知。其具体计算步骤如图 7 所示。

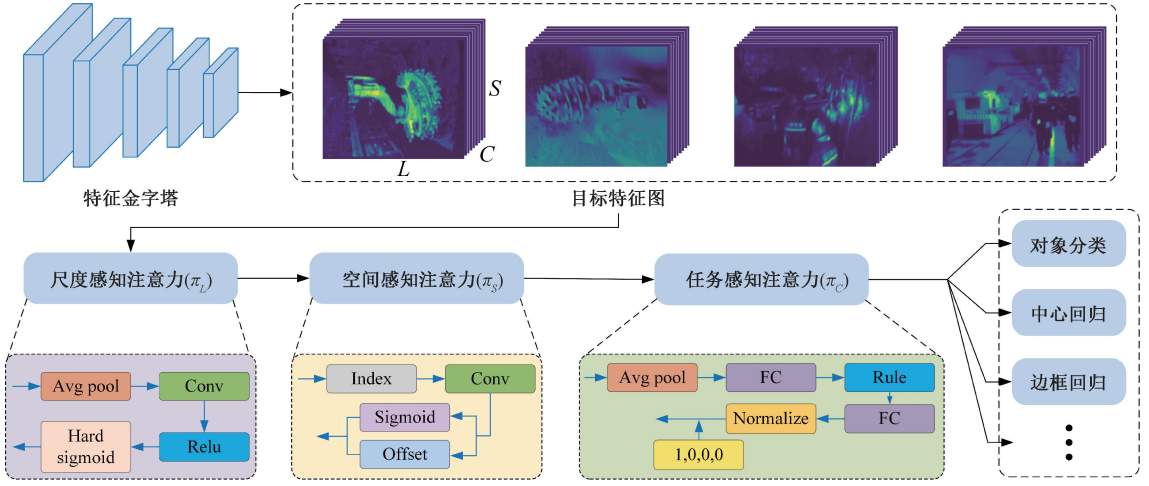


图 7 Dyhead 结构框架

Fig. 7 Structural framework of Dyhead

1) 张量重塑

首先,特征金字塔经缩放后的特征图可以表示为一个四维张量 $F \in R^{L \times H \times W \times C}$, 定义 $S = H \times W$, 将特征图重塑为三维张量 $F \in R^{L \times S \times C}$; 其中 L 表示金字塔中的层数; H 、 W 和 C 分别表示特征图的高度、宽度和通道数;

2) 尺度感知

在尺度感知注意力模块(π_L)中: 首先,特征图通过平均池化操作减少特征图尺寸降低计算量; 其次,通过卷积操作提取特征; 之后,通过 Rule 激活函数加快模型训练速度并减少梯度消失的可能性; 最后,通过 Hard sigmoid 激活函数拟合数据进行特征学习后与原特征图融合。尺度感知注意力计算公式如式(4)所示。

$$\pi_L(F) \cdot F = \sigma(f(\frac{1}{SC} \sum_{s,c} F)) \cdot F \quad (4)$$

式中: $\pi_L(\cdot)$ 表示应用于 L 维度的注意函数; $\sigma(\cdot)$ 为 Hard sigmoid 函数; $f(\cdot)$ 是类似 1×1 卷积的线性函数;

3) 空间感知

在空间感知注意力模块(π_S)中, 首先,通过 Index 快速访问与操作输入张量; 之后,通过可变形卷积使注意力学习变得稀疏; 最后,在相同的空间位置聚合跨级别的特征得到特征图的偏移量值以及特征图偏移量的权值项。空间感知注意力计算公式如式(5)所示。

$$\pi_S(F) \cdot F = \frac{1}{L} \sum_{l=1}^L \sum_{k=1}^K w_{l,k} \cdot F(l; p_k + \Delta p_k; c) \cdot \Delta m_k \quad (5)$$

式中: $\pi_S(\cdot)$ 表示应用于 S 维度的注意函数; K 代表稀疏采样位置的数量; $p_k + \Delta p_k$ 是通过自学习的空间偏移量

Δp_k 聚焦于判别区域的移位位置; Δm_k 是位置 p_k 处的自学习重要性标量;

4) 任务感知

在任务感知注意力模块(π_C)中, 首先,特征图通过平均池化操作降低维数与计算量; 其次,通过 FC 全连接层对提取的特征整合与分类; 然后,通过 Rule 激活函数加快模型训练速度; 再者,再通过一次 FC 全连接层对提取的特征进一步整合分类; 之后,通过 Normalize 标准化层提高模型性能与泛化能力; 最后,根据不同任务输出不同的信道值。任务感知注意力计算公式如式(6)所示。

$$\pi_C(F) \cdot F = \max(\alpha^1(F) \cdot F_c + \beta^1(F), \alpha^2(F) \cdot F_c + \beta^2(F)) \quad (6)$$

式中: $\pi_C(\cdot)$ 表示应用于 C 维度的注意函数; F_c 为第 C 个通道的特征切片; σ, β 为可学习的参数;

5) 检测输出

检测头进行对象分类、中心回归和边框回归等操作。其中一般的注意力公式如式(7)所示。

$$W(F) = \pi(F) \cdot F \quad (7)$$

式中: $\pi(\cdot)$ 表示注意函数; $\pi(F) = \pi_C(\pi_S(\pi_L(F) \cdot F) \cdot F)$ 。

3 实验与分析

3.1 煤矿井下多场景数据集构建

由于煤矿井下作业环境与地面环境不同,传统公共数据集无法体现煤矿井下的复杂作业环境,故本实验所

采用的数据集源自淮南新集煤矿、顾桥煤矿与张集煤矿不同时段掘进、支锚、采煤与辅助运输 4 种作业场景的现场监控视频。通过 OpenCV 中的抽帧操作对监控视频进行处理,以制作包含 3 000 张图像的煤矿复杂作业场景数据集,并按照 7 : 2 : 1 的比例将数据集划分为训练集、验证集与测试集。如图 8 所示,采用 Labelimg 软件对 4 种作业

场景中包括工作人员 (person)、掘进机截割头 (roadhead cutting head, R-cutting)、采煤机截割头 (shearer cutting head, S-cutting)、机载式锚杆钻机 (airborne jumbolter, A-jumbolter)、手持式锚杆钻机 (handling jumbolter, H-jumbolter)、支护装置 (support) 与电机车 (electric locomotive, E-locomotive) 的 7 种关键目标进行标定。

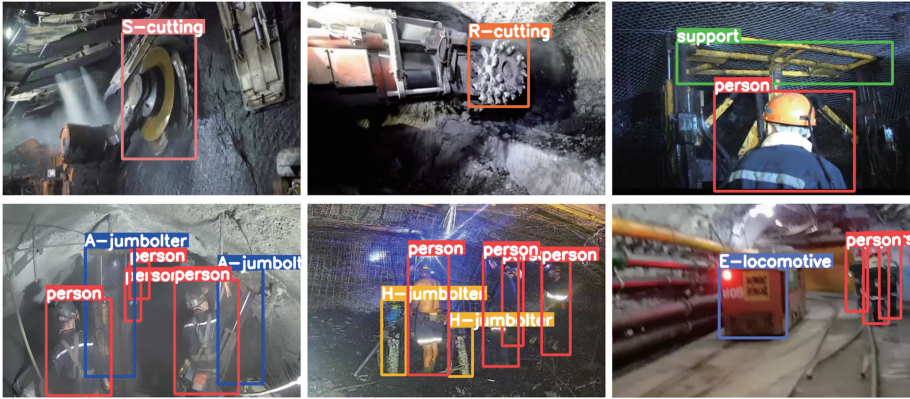


图 8 煤矿多场景数据集标定

Fig. 8 Coal mine multi-scenario dataset calibration

3.2 实验环境

本文研究基于 AutoDL 云服务器对 YOLOv5s-CSD 检测模型进行消融实验与对比实验,并基于 NVIDIA Jetson Xavier NX 视觉控制板开展针对 YOLOv5s-CSD 检测模型的视觉感知终端嵌入式实验。实验过程中遵循控制变量

原则,即采用相同的超参数与环境配置。模型的统一超参数如下:初始化动量因子为 0.937;输入图像尺寸为 640×640 pixels;批量大小为 32;初始学习率为 0.01;权重衰减系数为 0.000 5;训练总迭代次数为 300;具体实验环境配置如表 1 所示。

表 1 实验环境配置参数

Table 1 Configuration parameters of the experimental environment

Configuration content	Configuration parameters	
	AutoDL cloud server platform	NVIDIA Jetson Xavier NX
Operating system	Ubuntu 20.04 OS	Ubuntu 20.04 OS
GPU	RTX 3090 (24 GB)	384-core NVIDIA Volta™
CPU	Intel(R) Xeon(R) Platinum 8255C	6-core NVIDIA Carmel ARM® v8.2
Algorithm language	Python 3.8	Python 3.8
Accelerator	Cuda 11.3	CUDA 11.4
Algorithm framework	Pytorch 1.11.0	PyTorch 1.11.0

3.3 实验评价指标

在目标检测算法中常用检测准确率 (precision, P)、召回率 (recall, R)、平均检测精度 (average precision, AP)、平均检测精度均值 (mean average precision, mAP)、参数量与每秒浮点运算次数 (floating point operations per second, FLOPs) 作为评价指标检测算法模型的性能,其中 P、R、P_{AP} 与 P_{mAP} 计算公式如式 (8) ~ (11) 所示。

$$P = \frac{TP}{(TP + FP)}$$

(8)

$$R = \frac{TP}{(TP + FN)}$$

(9)

$$P_{AP} = \int_0^1 P(R) dR$$

(10)

$$P_{mAP} = \frac{1}{n} \sum_{i=1}^n P_{APi}$$

(11)

式中: TP 为正类判定为正类; FP 为负类判定为正类; FN 为正类判定为负类; P_{AP} 为 P (R) 曲线的积分即平均精度值; n 为目标类别数; P_{APi} 为第 i 种类别的平均检测精度值。

3.4 消融实验

为验证 3 项优化策略对检测模型性能的增益效果,本文基于相同的数据集与实验环境开展消融实验研究,

实验结果如图 9 和表 2 所示。Modal A 是 YOLOv5s 的基线模型,Modal E 是本文所提出的 YOLOv5s-CSD 检测模型,Modal B 至 Modal D 分别为 3 项优化策略独立作用的检测模型。

由图 9 与表 2 可见,模型 B 的 mAP@0.5 曲线在迭代 270 轮后逐步收敛并超过模型 A,且 mAP@0.5 值也由 87.5% 上升至 88.1%。由图 10 可见,在添加应用 CGNet 模块后,与模型 A 相比,模型 B 在多层级特征提取的各阶段对目标的检测轮廓更加清晰且特征激活点显著增加。

结果表明,CGNet 模块凭借其联合局部和周围环境特征,并将全局上下文应用于通道重新加权联合特征后,有效增强了检测模型的多层级特征提取能力,使得模型 B 提取到了更为丰富的低层形状与高层语义特征信息,有利于促进后续的特征融合操作。

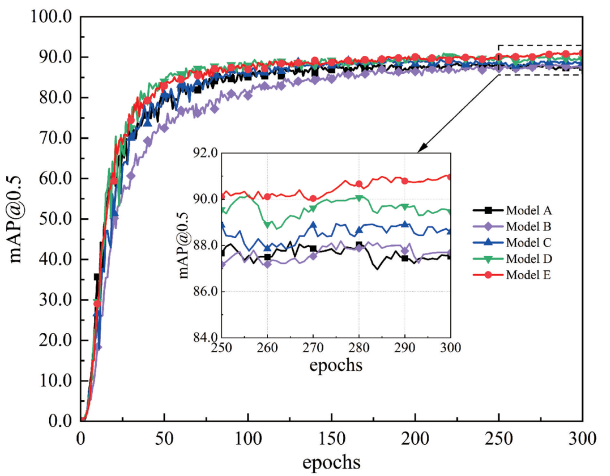


图 9 消融实验 mAP@0.5 曲线

Fig. 9 The mAP@0.5 curves of ablation experiment

表 2 消融实验具体结果分析
Table 2 Results of ablation experiments

模型	优化方法	AP@0.5/%							mAP@0.5/%	参数量/($\times 10^6$)
		R1	R2	R3	R4	R5	R6	R7		
A	YOLOv5s	88.6	89.1	87.9	80.9	82.9	90.4	92.7	87.5	6.70
B	YOLOv5s-CGNet	86.6	90.7	93.7	82.4	81.1	91.4	90.8	88.1	1.50
C	YOLOv5s-SlimNeck	88.3	90.6	91.8	82.7	82.8	92.1	91.4	88.5	5.58
D	YOLOv5s-Dyhead	90.1	90.2	91.6	84.4	83.9	92.7	95.7	89.8	11.74
E	YOLOv5s-CSD	90.5	93.1	96.6	84.1	84.6	92.2	95.5	91.0	10.04

注释: R1-(person); R2-(R-cutting); R3-(S-cutting); R4-(A-jumbolter); R5-(H-jumbolter); R6-(support); R7-(E-locomotive)

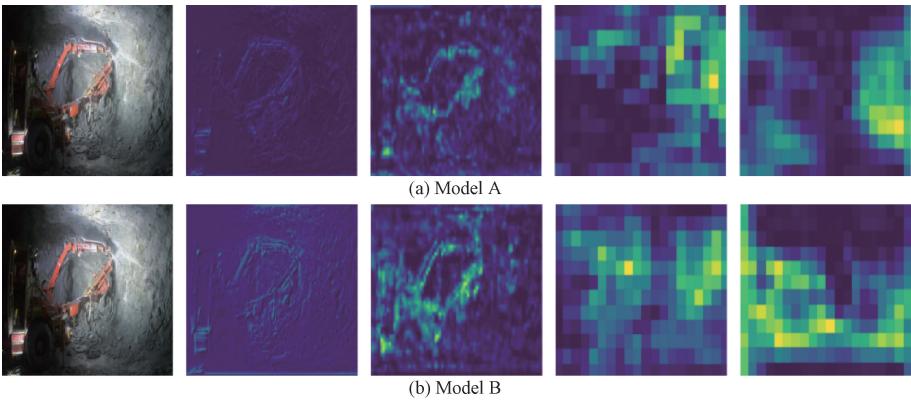


图 10 主干特征提取模块的输出特征

Fig. 10 Output feature map of the backbone feature extraction module

由图 9 与表 2 可见,模型 C 在迭代 100 轮后 mAP@0.5 值逐步收敛并超过模型 A,在采用 GSConv 卷积模块与 VoVGSCSP 模块分别代替颈部中的 Conv 和 C3 层后,模型的参数量由 6.7×10^6 降低至 5.58×10^6 ,同比下降了 16.72%,同时,mAP@0.5 值也由 87.5% 上升至 88.5%。由图 11 颈部模块输出的特征图可知,模型 C 在多层级特征融合后,各阶段对目标的特征激活点显著增加且检测轮廓更加清晰。该结果表明,SlimNeck 模块凭借 GSConv

的通道混洗与 VoVGSCSP 的跨阶段网络模块,能够提升不同阶段特征图的信息融合能力并且降低网络模型的参数量与复杂度。

由图 9 和表 2 可知,采用动态检测头 Dyhead 替换原检测头后,模型 D 在迭代 50 轮后 mAP@0.5 值超过模型 A 并逐步拉开与模型 A 的差距;且 mAP@0.5 值也由 87.5% 上升至 89.8%,提升了 2.3%。由图 12 与 13 可知,Dyhead 检测头能够在煤矿采煤机、掘进机与电机车

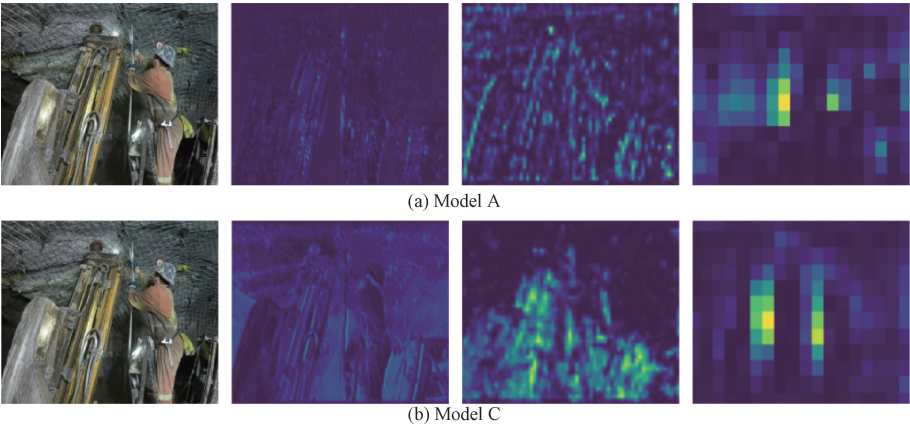


图 11 颈部特征融合模块的输出特征

Fig. 11 Output feature map of the neck feature fusion module

位置多变且作业姿态各异的复杂环境中使模型 D 的定位损失与分类损失显著低于模型 A。该结果表明, YOLOv5s-CSD 检测模型采用头部 Dyhead 改进模块能够增强模型在煤矿井下的多维感知能力,有效降低检测模型的定位损失与分类损失,提高模型检测精度。

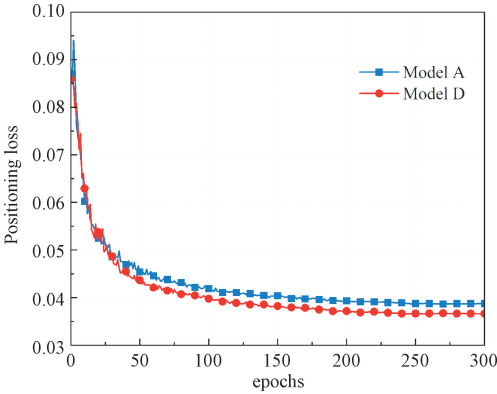


图 12 定位损失的曲线

Fig. 12 Positioning loss curve

3.5 对比实验

为了进一步验证 YOLOv5s-CSD 检测模型在煤矿井下多作业场景复杂环境中检测性能的优越性,本文选取了模型参数量与计算量相近的 YOLOv5s、YOLOv7-tiny、YOLOv8s、YOLOv9s、YOLOv11s、YOLOv12s 6 种目标检测主流算法与 YOLOv5s-CSD 开展对比试验,实验结果如图 14 和表 3 所示。

由图 14 可知,在迭代收敛速度方面,YOLOv5s-CSD、YOLOv12s、YOLOv11s、YOLOv8s、YOLOv5s 算法对应曲线的收敛速度接近且高于 YOLOv7-tiny 和 YOLOv9s 算法曲线,7 种算法曲线均在迭代 160 轮左右后逐渐趋于稳态。在检测精度方面,YOLOv5s-CSD 所对应的 mAP@0.5 曲线在迭代 25 轮后,与 YOLOv12s 所对应的 mAP@0.5 曲

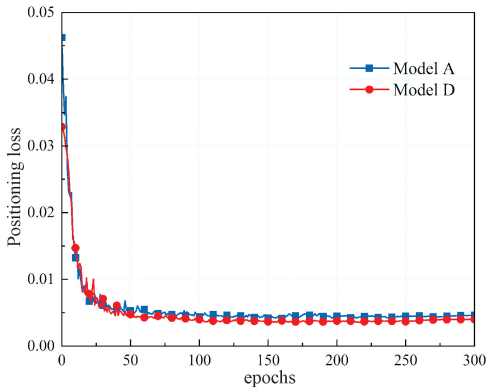


图 13 分类损失的曲线

Fig. 13 Classification loss curve

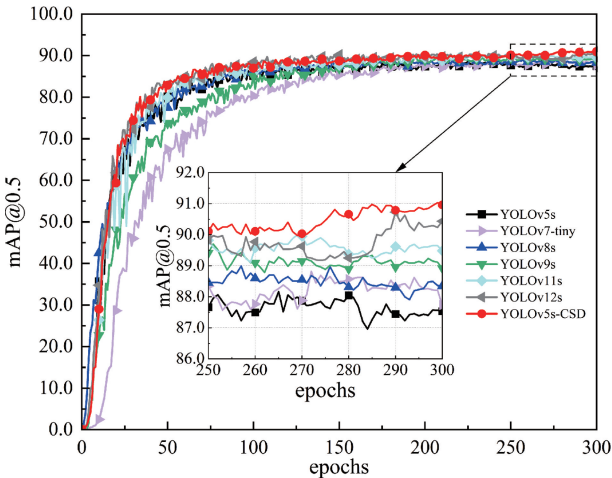


图 14 对比实验 mAP@0.5 曲线

Fig. 14 The mAP@0.5 curves of comparison experiment

线接近,并逐步攀升高于其他 5 种算法,并且在第 250 轮迭代后,5 条曲线均逐渐收敛趋于稳态;其中,YOLOv5s-

CSD 的 mAP@ 0.5 稳态峰值最高且在 90%~91% 之间浮动,而 YOLOv5s 的 mAP@ 0.5 稳态峰值最低且在 86%~

表 3 对比实验结果分析

Table 3 Results of comparative experiments on target detection

模型	AP@ 0.5/%							mAP@ 0.5/%	参数量/ ($\times 10^6$)	浮点数/ GFLOPs
	R1	R2	R3	R4	R5	R6	R7			
YOLOv5s	88.6	89.1	87.9	80.9	82.9	90.4	92.7	87.5	6.70	15.8
YOLOv7-tiny	89.8	90.3	90.9	83.4	82.8	90.1	86.9	87.7	5.75	13.2
YOLOv8s	88.8	94.0	89.9	79.9	84.6	93.3	90.5	88.7	10.62	28.5
YOLOv9s	92.3	82.5	90.8	84.6	84.7	94.7	94.5	89.1	9.33	39.8
YOLOv11s	91.3	87.8	94.5	84.4	84.3	95.1	89.4	89.6	9.43	21.6
YOLOv12s	90.3	92.8	94.9	83.2	85.0	90.3	95.1	90.2	9.26	21.5
YOLOv5s-CSD	90.5	93.1	96.6	84.1	84.6	92.2	95.5	91.0	10.04	15.7

注释: R1-(person); R2-(R-cutting); R3-(S-cutting); R4-(A-jumbolter); R5-(H-jumbolter); R6-(support); R7-(E-locomotive)

由表 3 可知,对于 person、support 和 E-locomotive 等 7 类关键目标的检测精度方面,YOLOv5s 对 support 与 E-locomotive 的检测精度均高于 90.0%,而对 A-jumbolter 与 H-jumbolter 的检测精度较低,相应的 AP@ 0.5 值仅为 80.9%与 82.9%;YOLOv5s-CSD 对于 person、R-cutting、S-cutting、support、E-locomotive 的 5 类关键目标检测精度均高于 90.0%,尤其是对 S-cutting 与 E-locomotive 的检测精度在 5 种算法最高且分别达 96.6%与 95.5%;YOLOv5s-CSD 对于 A-jumbolter 与 H-jumbolter 的检测精度与 YOLOv9s、YOLOv11s、YOLOv12s 相接近,并且高于其他 3 种算法,尤其是相较于 YOLOv5s 算法,其对二者的检测精度分别提升了 3.2%与 1.7%;YOLOv7-tiny 对 H-jumbolter 与 support 的检测精度与 YOLOv5s 相接近,而对 person、R-cutting 与 S-cutting 的检测精度略高于 YOLOv5s,分别为 89.8%、90.3%与 90.9%,但对 E-locomotive 的检测精度低于其余 6 类算法,相应的 AP@ 0.5 值仅为 86.9%;YOLOv8s 对 R-cutting 与 support 的检测精度最高,分别为 94%与 93.3%,而对 A-jumbolter 的检测精度在 7 种算法中最低,相应的 AP@ 0.5 值仅为 79.9%;YOLOv9s 对于 person、S-cutting、support 与 E-locomotive 检测精度均高于 90%,但对 R-cutting 的检测精度在 7 种算法中最低,相应的 AP@ 0.5 值仅为 82.5%。YOLOv11s 对于 S-cutting 与 support 的检测精度最高,分别为 94.5%与 95.1%,但对 E-locomotive 的检测精度与 YOLOv7-tiny 相接近,相应的 AP@ 0.5 值仅为 89.4%,低于其他 5 种算法。YOLOv12s 对于 H-jumbolter 的检测精度最高,略高于 YOLOv5s-CSD,但对其他 6 种关键目标的检测精度均低于 YOLOv5s-CSD。

对于 mAP@ 0.5 数值方面,YOLOv5s-CSD 明显高于其他 6 种算法,相应的 mAP@ 0.5 值为 91.0%,分别比 YOLOv5s、YOLOv7-tiny、YOLOv8s、YOLOv9s、YOLOv11s、YOLOv12s 高 3.5%、3.3%、2.3%、1.9%、1.4%、0.8%。

在模型复杂度方面,对于参数量,YOLOv8s、YOLOv9s、YOLOv11s、YOLOv12s、YOLOv5s-CSD 的参数量想接近,分别为 10.62×10^6 、 9.33×10^6 、 9.43×10^6 、 9.26×10^6 、 10.04×10^6 ;YOLOv5s 与 YOLOv7-tiny 均处于较低水平,分别为 6.70×10^6 和 5.75×10^6 。对于计算量,YOLOv8s、YOLOv9s、YOLOv11s、YOLOv12s 仍处于较高水平,分别为 28.5 GFLOPs、39.8 GFLOPs、21.6 GFLOPs、21.5 GFLOPs;YOLOv5s-CSD、YOLOv5s 与 YOLOv7-tiny 均处于较低水平,分别为 15.7 GFLOPs、15.8 GFLOPs、13.2 GFLOPs。由上述分析可见,YOLOv5s-CSD 由于融合了 CGNet、SlimNeck 与 Dyhead 3 种优化策略,导致其参数量较高,但计算量较低,整体模型复杂度适中,并且得益于 3 种优化策略对检测性能的提升,使其对巷道复杂环境的适应性更强且鲁棒性更高。YOLOv8s、YOLOv9s、YOLOv11s、YOLOv12s 的参数量与计算量均较高,模型架构复杂且冗余信息较多,不利于后续在边缘计算设备上的嵌入式应用。YOLOv5s 与 YOLOv7-tiny 两种算法的模型复杂度较低,但由于检测性能较差,使二者也难以适应煤矿复杂工况条件。综上分析可知,YOLOv5s-CSD 算法模型的参数量和计算量适中且 mAP@ 0.5 值为 91.0%明显高于其他算法,综合检测性能最佳。因此 YOLOv5s-CSD 更适合作为实验测试平台的检测网络,以实现煤矿井下掘进、支锚、采煤与辅助运输多场景复杂作业中关键目标的精准识别与定位。

3.6 嵌入式检测实验

为进一步检验 YOLOv5s-CSD 的实时检测性能,采用 8 段包含煤矿井下掘进、支锚、采煤与辅助运输 4 种复杂作业场景的实际生产视频片段作为测试视频,并将训练后的 YOLOv5s-CSD 迁移部署到实验测试平台的视觉边缘计算控制板内,以测试其对 8 段视频片段种中 person、support、E-locomotive 等(支护架、机载锚杆钻机、掘进机截割头、采煤机截割头、手持锚杆钻机、作业人员与电机

车) 7 种关键目标的实时检测效果。

YOLOv5s-SCD 的性能测试结果如表 4 所示, 通过 TensorRT-FP16 引擎加速后, 当数据样本批量 (Batch size) 设定为 32 时, NVIDIA Jetson Xavier NX 控制板的 GPU 可获得最高利用率。YOLOv5s-CSD 对测试集图像的前处理时间最短, 为 0.4 ms, 单图推理时间适中, 且其实时检测速度为 38.6 fps, 接近于 YOLOv11s, 略低于 YOLOv12s、YOLOv5s、YOLOv7-tiny, 但高于 YOLOv8s 与 YOLOv9s 的实时检测速度。由煤矿监控视频实时监测帧率标准可知, 实时检测要求检测速度不低于 24 fps。因此, YOLOv5s-CSD 在视觉感知终端上可以辅助实验测试平台实现对煤矿井下多作业场景复杂环境中关键目标的实时检测。实验测试平台如图 15 所示。本文通过实验测试平台对 YOLOv5s ~ YOLOv12s 6 种 (YOLOv5s ~ YOLOv9s 4 种) YOLO 系列经典检测算法与 YOLOv5s-

CSD 算法进行了对比测试, 进一步验证检测算法的优越性与可靠性。



图 15 实验测试平台

Fig. 15 Experimental platform for test robot

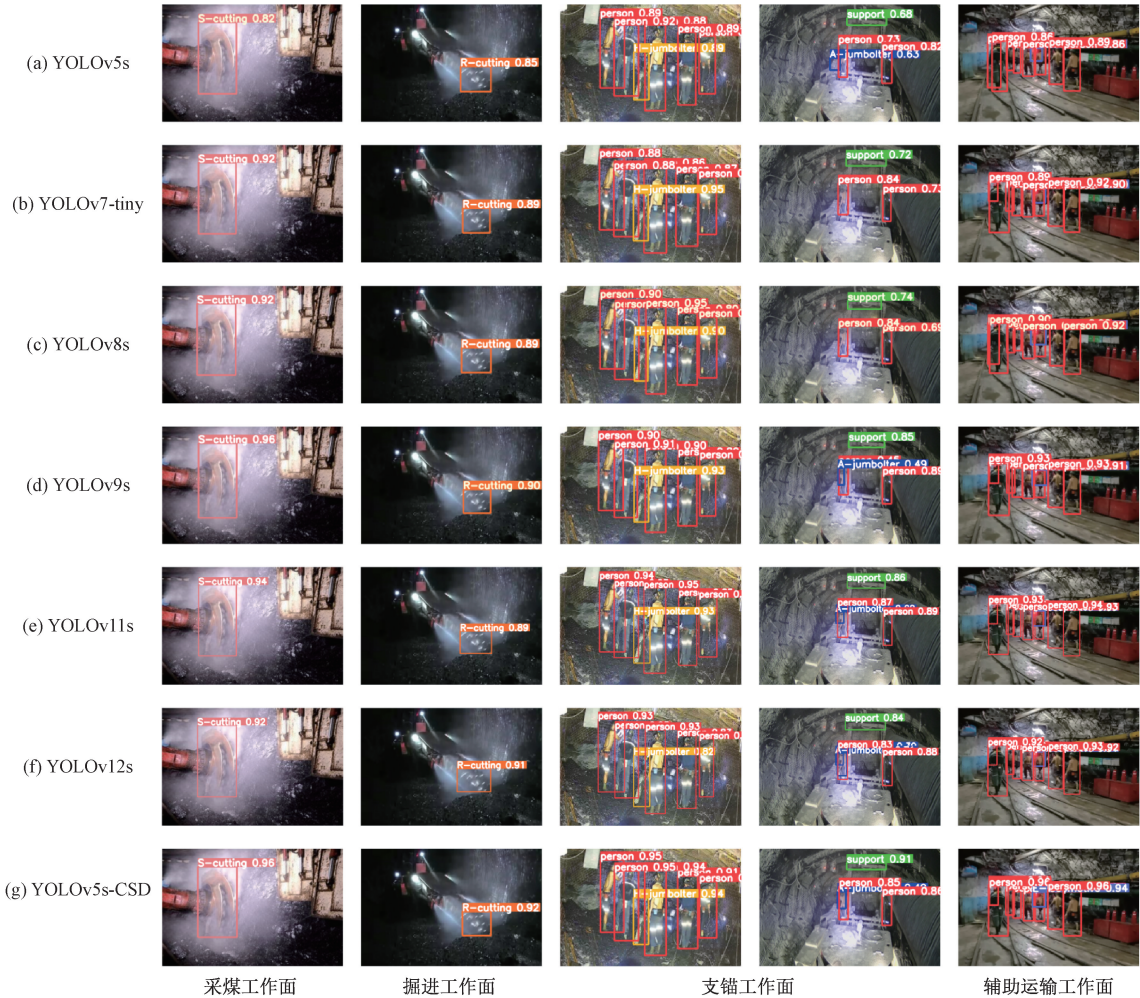


图 16 对比测试结果

Fig. 16 Visual presentation of test results

表 4 YOLOv5s-SCD 的性能测试结果

Table 4 Performance test results of YOLOv5s-SCD

模型	参数量/($\times 10^6$)	浮点数/GFLOPs	实时检测速度/fps	前处理/ms	单图推理时间/ms
YOLOv5s	6.70	15.8	47.6	0.8	19.3
YOLOv7-tiny	5.75	13.2	48.3	1.0	18.7
YOLOv8s	10.62	28.5	35.5	1.7	25.3
YOLOv9s	9.33	39.8	28.4	0.5	33.6
YOLOv11s	9.43	21.6	41.1	1.6	21.3
YOLOv12s	9.26	21.5	44.1	1.5	20.0
YOLOv5s-CSD	10.04	15.7	38.6	0.4	24.8

由图 16 可知,对于高尘雾的采煤作业场景, YOLOv7-tiny~YOLOv5s-CSD 6 种检测算法对 S-cutting 检测精度均高于 90%,并且 YOLOv5s-CSD 对 S-cutting 检测精度最高,达 96%,而 YOLOv5s 对 S-cutting 检测精度最低,仅为 82%。对于低照度的掘进作业场景, YOLOv5s-CSD 对 R-cutting 的检测精度最高,达 92%,而 YOLOv5s、YOLOv7-tiny、YOLOv8s、YOLOv11s 4 种算法对 R-cutting 检测效果较差,均低于 90%。对于人机混杂分布的支锚作业场景, YOLOv5s-CSD 对 support、person、A-jumbolter、H-jumbolter 4 类关键目标的检测精度高于其他 6 种算法, YOLOv5s~YOLOv12s 4 种算法对 support 的检测效果差,均低于 90%,并且 YOLOv7-tiny 与 YOLOv8s 对 A-jumbolter 存在漏检情况。对于存在小目标尺度的辅助运输作业场景, YOLOv5s-CSD 对 E-locomotive 检测精度最高,达 94%。因此,综合上述分析可见,在实验测试平台上, YOLOv5s-CSD 对 person~E-locomotive 7 类关键目标的检测精度均在 90.0%以上,且不存在漏检与误检情况,综合检测性能最佳,更适用于煤矿井下多种复杂作业场景的关键目标识别与定位任务。

4 结 论

针对煤矿井下掘进、支锚、采煤与辅助运输等多场景复杂工矿与恶劣环境因素所导致的关键装备与作业人员等关键目标检测效果不佳的问题,本文提出了一种基于机器视觉的煤矿多场景关键目标检测算法(YOLOv5s-CSD)。首先,利用 CGNet 模块的局部与全局残差学习机制以及上下文引导策略增强主干网络的特征提取能力;利用 SlimNeck 模块的通道混洗与一次聚合策略提升颈部网络的特征融合能力;凭借 Dyhead 模块增强检测头的多维信息感知能力以降低了检测模型的定位损失与分类损失。其次,基于自建的煤矿多目标场景数据集与 AutoDL 云服务器开展消融实验和对比实验,验证了 YOLOv5s-CSD 的优越性能。最后,将 YOLOv5s-CSD 迁移部署到实验测试平台中,检测其在实际多场景复杂工矿下的检测性能,实验表明,其能够实现对煤矿井下多场景

复杂作业环境中关键目标的精准检测。然而通过普通可见光相机采集煤矿数据集时容易受到尘雾、黑暗环境干扰,导致数据集质量降低,今后考虑将普通 RGB 相机与红外热成像相机相融合采集煤矿数据集,避免井下恶劣环境对数据采集的影响,同时将继续探索煤矿井下的目标检测算法。

参考文献

[1] 葛世荣,樊静丽,刘淑琴,等. 低碳化现代煤基能源技术体系及开发战略[J]. 煤炭学报, 2024, 49(1): 203-223.

GE SH R, FAN L J, LIU SH Q, et al. Low carbon modern coal-based energy technology system and development strategy[J]. Journal of China Coal Society, 2024, 49(1): 203-223.

[2] 王国法,张良,李首滨,等. 煤矿无人化智能开采系统理论与技术研发进展[J]. 煤炭学报, 2023, 48(1): 34-53.

WANG G F, ZHANG L, LI SH B, et al. Progresses in theory and technological development of unmanned smart mining system[J]. Journal of China Coal Society, 2023, 48(1): 34-53.

[3] XU J P, GAO W, XIE H P, et al. Integrated tech-paradigm based innovative approach towards ecological coal mining[J]. Energy, 2018, 151: 297-308.

[4] HOU C C, QIAO T Z, DONG H J, et al. Coal flow volume detection method for conveyor belt based on TOF vision[J]. Measurement, 2024, 229: 114468.

[5] ZHEN L, WANG B C, LI M, et al. Inspection routing problem for coal mine safety personnel in underground mines[J]. Computers & Industrial Engineering, 2019, 130: 526-536.

[6] YANG W, WANG S, WU J Q, et al. A low-light image enhancement method for personnel safety monitoring in underground coal mines [J]. Complex & Intelligent Systems, 2024: 1-14.

[7] YANG T, WANG S, TONG J L, et al. Accurate real-time obstacle detection of coal mine driverless electric

- locomotive based on ODEL-YOLOv5s [J]. Scientific Reports, 2023, 13(1): 17441.
- [8] 华钢, 任荣, 徐永刚, 等. 基于 DM642 的多目标运动状态感知系统设计[J]. 工矿自动化, 2015, 41(6): 14-17.
HUA G, REN R, XU Y G, et al. Design of sensing system of multi-target motion state based on DM642[J]. Journal of Mine Automation, 2015 41(6): 14-17.
- [9] 王勇. 煤矿井下人员视频图像识别跟踪的研究与应用[J]. 电子测量技术, 2020, 43(1): 28-31.
WANG Y. Research and application of video image recognition and tracking for underground personnel in coal mine [J]. Electronic Measurement Technology, 2020, 43(1): 28-31.
- [10] 巩固, 朱华. 基于目标识别与避障的煤矿救援机器人自主行走[J]. 南京理工大学学报, 2022, 46(1): 32-39.
GONG G, ZHU H. Autonomous walking of coal mine rescue robot based on target recognition and obstacle avoidance[J]. Journal of Nanjing University of Science and Technology, 2022, 46(1): 32-39.
- [11] 杨辉羽, 李海明. 单聚合 YOLO 航拍小目标检测算法[J]. 国外电子测量技术, 2023, 42(4): 131-140.
YANG H Y, LI H M. Single aggregation YOLO algorithm for airborne small target detection[J]. Foreign Electronic Measurement Technology, 2023, 42(4): 131-140.
- [12] 贾晓芬, 吴雪茹, 赵佰亭. 绝缘子自爆缺陷的轻量化检测网络 DE-YOLO [J]. 电子测量与仪器学报, 2023, 37(5): 28-35.
JIA X F, WU X R, ZHAO B T. Lightweight detection network for insulator self-detonation defect DE-YOLO[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(5): 28-35.
- [13] 包妮沙, 韩子松, 于嘉欣, 等. 基于改进 Faster-RCNN 的露天煤矿开采区遥感识别方法[J]. 东北大学学报(自然科学版), 2023, 44(12): 1759-1768.
BAO N SH, HAN Z S, YU J X, et al. Remote sensing identification method for open-pit coal mining area based on improved Faster-RCNN[J]. Journal of Northeastern University (Natural Science), 2023, 44 (12): 1759-1768.
- [14] WANG X, WANG S, GUO Y C, et al. Multi-scale coal and gangue detection in dense state based on improved Mask R-CNN[J]. Measurement, 2023, 221: 113467.
- [15] TANG C, CHEN S, ZHOU X, et al. Small-scale face detection based on improved R-FCN [J]. Applied Sciences, 2020, 10(12): 4177.
- [16] WANG Y B, WANG Y J, DANG L F. Video detection of foreign objects on the surface of belt conveyor underground coal mine based on improved SSD [J]. Journal of Ambient Intelligence and Humanized Computing, 2023, 14(5): 5507-5516.
- [17] CHENG X, YU J. RetinaNet with difference channel attention and adaptively spatial feature fusion for steel surface defect detection [J]. IEEE Transactions on Instrumentation and Measurement, 2020, 70: 1-11.
- [18] 高嘉, 刘涛, 王显峰, 等. TWRD-Net: 一种用于曳引钢丝绳缺陷的实时检测网络算法[J]. 仪器仪表学报, 2023, 44(6): 223-235.
GAO J, LIU T, WANG X F, et al. TWRD-Net: A real-time detection network algorithm for traction wire rope defects[J]. Chinese Journal of Scientific Instrument, 2023, 44(6): 223-235.
- [19] 张乾闯, 郭晨霞, 杨瑞峰, 等. 基于改进 YOLO 算法的光纤环绕制缺陷检测[J]. 电子测量技术, 2023, 46(10): 32-39.
ZHANG Q CH, GUO CH X, YANG R F, et al. Research on defect detection system for FOC winding based on YOLO algorithm [J]. Electronic Measurement Technology, 2023, 46(10): 32-39.
- [20] XUE G H, LI S X, HOU P, et al. Research on lightweight Yolo coal gangue detection algorithm based on resnet18 backbone feature network [J]. Internet of Things, 2023, 22: 100762.
- [21] 李飞, 胡坤, 张勇, 等. 基于混合域注意力 YOLOv4 的输送带纵向撕裂多维度检测[J]. 浙江大学学报(工学版), 2022, 56(11): 2156-2167.
LI F, HU K, ZHANG Y, et al. Multi-dimensional detection of longitudinal tearing of conveyor belt based on YOLOv4 of hybrid domain attention [J]. Journal of Zhejiang University (Engineering Science), 2022, 56(11): 2156-2167.
- [22] ZHAO D Y, SU G Y, CHENG G, et al. Research on real-time perception method of key targets in the comprehensive excavation working face of coal mine[J]. Measurement Science and Technology, 2023, 35(1): 015410.
- [23] PAN L H, DUAN Y X, ZHANG Y J, et al. A lightweight algorithm based on YOLOv5 for relative position detection of hydraulic support at coal mining faces[J]. Journal of Real-Time Image Processing, 2023, 20(2): 40.
- [24] YANG T, GUO Y C, LI D Y, et al. Vision-Based obstacle detection in dangerous region of coal mine driverless rail electric locomotives [J]. Measurement, 2025, 239: 115514.

- [25] 王宏伟, 李进, 闫志蕊, 等. 基于图像与点云融合的巷道锚护孔位识别定位方法[J]. 煤炭科学技术, 2024, 52(5): 249-261.
WANG H W, LI J, YAN ZH R, et al. Roadway anchor hole recognition and positioning method based on image and point cloud fusion [J]. Coal Science and Technology, 2024, 52(5): 249-261.
- [26] 薛小勇, 何新宇, 姚超修, 等. 基于改进 YOLOv8n 的采掘工作面小目标检测方法[J]. 工矿自动化, 2024, 50(8): 105-111.
XUE X Y, HE X Y, YAO CH X, et al. Small object detection method for mining face based on improved YOLOv8n [J]. Journal of Mine Automation, 2024, 50(8): 105-111.
- [27] XU Z F, SUN Z, LI J Y. Research on coal flow visual detection and the energy-saving control method based on deep learning[J]. Sustainability, 2024, 16(13): 5783.
- [28] 张婧, 冯莹莹, 李洪安, 等. 煤矿工作面喷雾除尘场景下的安全帽识别算法[J]. 矿业安全与环保, 2024, 51(4): 9-16.
ZHANG J, FENG Y Y, LI H AN, et al. Safety helmet recognition algorithm in spray dust removal scenario of coal mine working face [J]. Mining Safety & Environmental Protection, 2024, 51(4): 9-16.
- [29] WU T Y, TANG S, RUI Z, et al. CGNet: A light-weight context guided network for semantic segmentation [J]. IEEE Transactions on Image Processing, 2020, 30: 1169-1179.
- [30] LI H L, LI J, WEI H B, et al. Slim-neck by GSConv: A lightweight-design for real-time detector architectures [J]. Journal of Real-Time Image Processing, 2024, 21(3): 62.
- [31] DAI X Y, CHEN Y P, XIAO B, et al. Dynamic head: Unifying object detection heads with attentions [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Virtual, 2021: 7373-7382.
- [32] ZHOU G, FENG B, YIN W J, et al. Numerical simulations on airflow-dust diffusion rules with the use of coal cutter dust removal fans and related engineering applications in a fully-mechanized coal mining face[J]. Powder Technology, 2018, 339: 354-367.
- [33] SUI Y P, ZHANG L, SUN Z P, et al. Research on coal and gangue recognition based on the improved YOLOv7-

Tiny target detection algorithm [J]. Sensors, 2024, 24(2): 456.

- [34] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift[J]. ArXiv preprint arXiv: 1502.03167, 2015:448-456.
- [35] HE K M, ZHANG X Y, REN S Q, et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification [C]. Proceedings of the IEEE International Conference on Computer Vision, 2015: 1026-1034.
- [36] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.

作者简介



苏国用, 2022 年于安徽理工大学获得博士学位, 现为安徽理工大学讲师, 主要研究方向为多模态智能感知技术。

E-mail: guoyongs005@sina.cn

Su Guoyong received his Ph. D. degree from Anhui University of Science and Technology in 2022. He is now a lecturer in Anhui University of Science and Technology. His main research interest includes multimodal intelligent perception technology.



洪洪铨, 2022 年于河南科技大学获得学士学位, 现为安徽理工大学硕士研究生, 主要研究方向为智能感知与控制。

E-mail: 18403797336@163.com

Shuai Hongxin received his B. Sc. degree from Henan University of Science and Technology in 2022. He is now a M. Sc. candidate at Anhui University of Science and Technology. His main research interests include intelligent perception and control.



邓海顺 (通信作者), 2013 年于南京航空航天大学获得博士学位, 现为安徽理工大学教授, 主要研究方向为矿山智能技术与装备。

E-mail: dhs1998@163.com

Deng Haishun (Corresponding author) received his Ph. D. degree from Nanjing University of Aeronautics and Astronautics in 2013. He is now a professor in Anhui University of Science and Technology. His main research interests include mining intelligent technology and equipment.