

中文引用格式:张淼,王孝军,雷经发,等. 轻量化神经网络结合深度相机的矿工目标检测与定位[J]. 中国安全科学学报, 2025, 35(3):115-124.

英文引用格式:ZHANG Miao, WANG Xiaojun, LEI Jingfa, et al. Lightweight neural network combined with depth camera for miner target detection and localization [J]. China Safety Science Journal, 2025, 35(3):115-124.

轻量化神经网络结合深度相机的矿工目标 检测与定位*

张淼^{1,2}讲师, 王孝军^{1,2}, 雷经发^{**1,2,3}教授, 赵汝海^{1,2}副教授, 李永玲^{1,2}讲师
(1 安徽建筑大学机械与电气工程学院, 安徽合肥 230601; 2 工程机械智能制造安徽省
教育厅重点实验室, 安徽合肥 230601; 3 过程装备与控制工程四川省高校重点实验室,
四川自贡 643000)

中图分类号: X924.3 文献标志码: A DOI: 10.16265/j.cnki.issn1003-3033.2025.03.0863

基金项目: 安徽高校自然科学研究重大项目(KJ2021ZD0068); 安徽高校协同创新项目(GXXT2022-019); 过程装备
与控制工程四川省高校重点实验室开放基金资助(GK202308)。

【摘要】 为防止矿工误入危险区域, 提出一种YOLOv5s-MPD轻量化井下矿工目标检测模型, 并结合深度相机定位矿工目标, 实时检测矿工是否进入危险区域。首先, 使用MobileNetV3轻量化神经网络作为主干特征提取网络, 大幅降低模型体积; 其次, 引入极化自注意力模块(PSA), 增强目标的感知能力; 最后, 采用可变形卷积网络(DCNv2)替代特征融合层中C3模块的标准卷积, 解决常规卷积丢失部分特征信息的问题, 利用改进模型结合深度相机获取的彩色图像检测矿工目标, 并得到目标中心点的空间三维坐标。结果表明: 改进模型相比于YOLOv5s, 参数量和计算量分别减少83.54%和77.03%, 模型体积大小仅为3.4 MB, 检测速度为70.2帧/s, 提升54.97%, 平均精度均值(mAP)为0.825。与主流目标检测模型相比, 改进模型的参数量、计算量、模型体积、检测速度和mAP较为均衡。在实际定位精度试验中, 1~8 m范围内测得相机与矿工目标间距离的平均绝对误差和平均相对误差分别为0.11 m和1.74%; 最大绝对误差和最大相对误差分别为0.25 m和2.96%。在动态检测中, 均能检测到矿工目标并输出其位置信息, 检测成功率达97.5%。

【关键词】 轻量化; 神经网络; 深度相机; 目标检测; 目标定位; 安全预警

Lightweight neural network combined with depth camera for miner target detection and localization

ZHANG Miao^{1,2}, WANG Xiaojun^{1,2}, LEI Jingfa^{1,2,3}, ZHAO Ruhai^{1,2}, LI Yongling^{1,2}

(1 School of Mechanical and Electrical Engineering, Anhui Jianzhu University, Hefei Anhui 230601, China;
2 Key Laboratory of Intelligent Manufacturing of Construction Machinery, Anhui Education Department,
Hefei Anhui 230601, China; 3 Sichuan Provincial Key Laboratory of Process Equipment and Control
Engineering, Zigong Sichuan 643000, China)

Abstract: To prevent miners from mistakenly entering dangerous areas, a lightweight underground miner

* 文章编号: 1003-3033(2025)03-0115-10; 收稿日期: 2024-10-23; 修稿日期: 2024-12-25

** 通信作者: 雷经发(1978—), 男, 安徽巢湖人, 博士, 教授, 主要从事视觉检测、人因工程等方面的研究。E-mail: rain78828@163.com。

object detection model based on YOLOv5s-MPD was proposed, which combined with depth camera to locate miner targets and detect whether miners had entered dangerous areas in real time. Specifically, the MobileNetv3 lightweight neural network was used as the backbone feature extraction network to significantly reduce the model size. Secondly, Polarized Self-Attention (PSA) module was introduced to enhance the perception of targets. Finally, Deformable Convolution Network v2 (DCNv2) was used to replace the standard convolution in the C3 module of the feature fusion layer, solving the problem of partial feature information loss in conventional convolution. The improved model was used in combination with the color images obtained by the depth camera to detect miner targets and obtain the spatial three-dimensional coordinates of the target center points. The results show that compared with YOLOv5s, the improved model reduces the number of parameters and computation by 83.54% and 77.03%, respectively. The model size is only 3.4 MB, and a detection speed of 70.2 f/s, which is increased by 54.97%. The mean average precision is 0.825. Compared with mainstream object detection models, the improved model has a more balanced number of parameters, computation, model size, detection speed, and mean average precision. In the actual positioning accuracy test, within a range of 1–8 meters, the average absolute error and average relative error of the distance between the camera and the miner target were 0.11 meters and 1.74%, respectively. The maximum absolute error and maximum relative error were 0.25 meters and 2.96%, respectively. In the dynamic detection, the miner target could be detected and its location information output, with a detection success rate of 97.5%.

Keywords: lightweight; neural network; deep camera; target detection; target localization; security warning

0 引言

目前,我国煤矿采掘行业正处于综合机械化采矿向智能化采矿的过渡阶段,逐渐形成了煤炭智能化理论体系^[1]。ZHANG Kexue 等^[2]提出煤矿智能化系统的架构,认为矿井目标检测是煤矿智能化高速通信及信息获取的基础,并指出应以人员检测为主。由于煤矿井下存在多种危险区域,如易发生垮落区、采空区、有害气体超标区、高压电区以及大型机械设备工作区等^[3],为防止矿工误入危险区域,煤矿井下人员定位是安全生产的重要组成部分^[4-5]。因此,在煤矿井环境中实现煤矿工人实时高精度的定位对保护矿工人身安全、提高安全生产效率具有重要意义。

在国内外关于井下矿工目标检测与定位的研究中,视觉感知方法因其信息量丰富和检测效率高,具有广泛的应用前景。基于深度相机的煤矿井下目标定位技术,通过融合彩色图像与深度信息,在彩色图像中检测矿工目标,并结合检测结果与深度数据,实现井下矿工的精确定位^[6]。近年来,许多学者致力于将深度学习目标检测算法应用于矿井安全检测领域^[7-9]。张春堂等^[10]通过更换单发多框检测器(Single Shot MultiBox Detector, SSD)算法的特征提取网络,提高了矿工安保穿戴情况检测速度。郭永

存等^[11]改进掩码区域卷积神经网络(Mask Region-Convolution Neural Network, Mask R-CNN)算法,提高了黑暗环境中无人电车小目标障碍物检测精度。李伟山等^[12]通过改进快速区域卷积神经网络(Faster Region-Convolution Neural Network, Faster R-CNN)算法中的候选区域网络,提高了矿工检测精度。王科平等^[13]提出一种基于改进 YOLOv4 的目标检测算法用于综采工作面对关键设备及人员的准确检测,其平均精度均值(mean Average Precision, mAP)为 92.59%。郝明月等^[14]提出一种基于改进 YOLOv5 的矿工排队检测方法,在复杂矿井环境中能够准确识别矿工排队情况。杨豚等^[15]通过采用部分卷积替换 YOLOv5 模型 C3 模块中的传统卷积和解耦模型预测头,为井下无人驾驶的轨道电机车提供了更精准的安全预警。上述方法虽然能够实现对矿工目标的检测,但模型复杂度高、参数量和计算量较大,无法在检测精度和检测速度之间实现有效平衡。在检测出目标后,为进一步实现精确的定位,多采用单目或双目视觉空间定位技术。郭曦等^[16]使用双目相机跟踪目标并进行测距,防止矿工进入危险区域。ZHOU Cheng 等^[17]基于双目相机,将其应用于机器设备的安全保护,实现人员与设备的分离和预警。韩江洪等^[18]基于单目相机定位井下巷道行人,实现了矿工与机车之间的距离检测,从而减少机车运行

事故发生。但此类方法的定位精确度和速度仍需进一步提升。

因此,笔者拟针对矿井作业人员安全预警需要,提出一种改进的 YOLOv5s 轻量化井下矿工目标检测模型,着重于优化检测速度与精度之间的平衡,并结合深度相机以实时检测和定位矿工目标。

1 改进的 YOLOv5s 轻量化矿工目标检测模型

选用 YOLOv5 算法中的 YOLOv5s 模型作为基础模型,该模型框架由输入层、主干网络层、特征融合层和输出层组成。

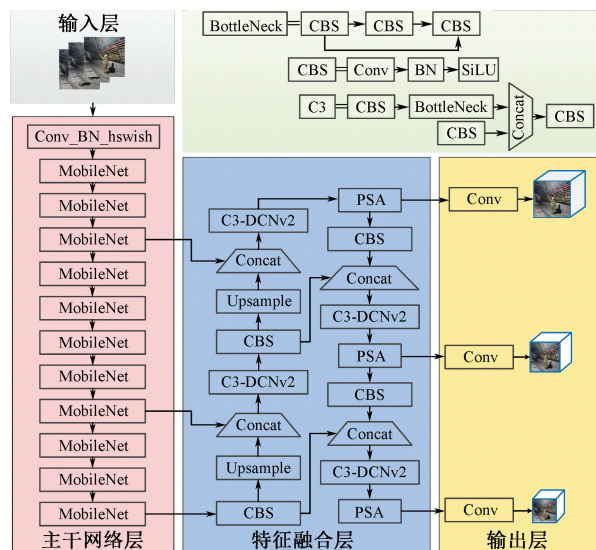
输入层应用 Mosaic 增强技术提升数据多样性;主干网络层通过 CBS (Convolution - BatchNormalization - SiLU)、C3 和快速空间金字塔池化等模块进行特征提取;特征融合层由特征金字塔网络 (Feature Pyramid Networks, FPN) 和路径聚合网络 (Path Aggregation Network, PAN) 构成,FPN 自底向上传递语义信息,PAN 实现跨层级特征交互以增强多尺度检测能力;输出层最终生成目标定位与分类信息,为检测跟踪任务提供支持。

将 YOLOv5s 的主干网络替换为 MobileNetv3 轻量化特征提取网络,大幅降低模型的参数量和计算量,提升检测速度,实现模型的轻量化设计。但对于井下强光干扰、灰尘干扰、遮挡干扰和低光照等复杂环境导致的图像质量较差时,YOLOv5s-MobileNetv3 模型检测效果难以保证。因此,在特征融合层末端嵌入极化自注意力 (Polarized Self-Attention, PSA) 模块,关注复杂环境中的目标,提高模型对井下矿工目标的检测精度,将可变形卷积 (Deformable Convolution Network v2, DCNv2) 替代特征融合层中 C3 模块的标准卷积增强模型对井下矿工目标形变的适应能力,改进后的 YOLOv5s-MPD (MobileNetv3 PSA DCNv2) 模型结构如图 1 所示。

1.1 模型轻量化改进

MobileNetv3 结合了深度可分离卷积、线性瓶颈和倒置残差结构,删除计算成本较高的网络层,融入压缩激励注意力模块 (Squeeze-and-Excitation, SE)。此外,MobileNetv3 网络采用结构重参数化方法融合了卷积层和批量归一化层,引入易于量化模型的 h-swish 激活函数。使得模型保持较高精度的前提下,减少参数和计算量。

在主干网络中由于使用 MobileNetv3 深度可分



注:Conv 为卷积;BN 为批量归一化;SiLU 为激活函数;CBS 由 Conv、BN 和 SiLU 组成;BottleNeck 为瓶颈模块;C3 由 3 个 CBS 和 BottleNeck 组成;MaxPool 为最大池化操作;Upsample 为上采样;Concat 为张量拼接。

图 1 改进后的 YOLOv5s-MPD 模型结构

Fig. 1 Improved YOLOv5s-MPD model structure

离卷积替代 YOLOv5s 传统卷积。与传统卷积不同,该方法分开处理通道和空间信息,首先对每个通道分别进行卷积操作,然后再通过逐点卷积进一步减小参数量和计算复杂度。设输入特征图高为 H ,宽为 W ,输入通道数为 D_1 ,卷积核高宽的尺寸为 $K \times K$,输出通道数为 D_2 ,标准卷积参数量 P_o 和计算量 F_o 的计算如下:

$$P_o = D_1 D_2 K K \quad (1)$$

$$F_o = D_1 D_2 H W K K \quad (2)$$

替代为深度可分离卷积后,参数量 P_l 和计算量 F_l 的计算公式:

$$P_l = D_1 K K + D_1 D_2 \quad (3)$$

$$F_l = D_1 H W K K + D_1 D_2 H W \quad (4)$$

深度可分离卷积和标准卷积参数量比值和计算量比值:

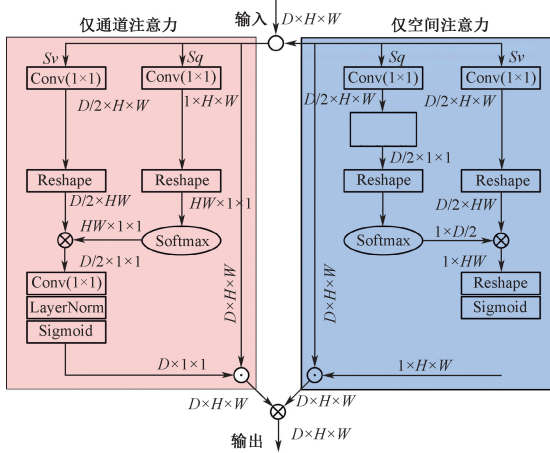
$$\frac{P_l}{P_o} = \frac{F_l}{F_o} = \frac{1}{D_2} + \frac{1}{K^2} \quad (5)$$

由式 (5) 可知:可分离卷积替代传统卷积后,参数量和计算量均减少到标准卷积的 $\frac{1}{D_2} + \frac{1}{K^2}$,即当 MobileNet 使用 3×3 的深度可分离卷积时,相比标准卷积的参数量和计算量减少 8~9 倍。

1.2 注意力模块

针对实际应用中的复杂矿井环境,引入 PSA 模块,

该模块是结合通道和空间的注意力模块,可以增强目标的特征,减弱复杂环境与无关信息干扰,使得检测更准确,同时减少计算开销,模块结构如图2所示。



注:Reshape 为数组重组;Global Pooling 为全局池化;LayerNorm 为归一化;Softmax 和 Sigmoid 为激活函数。

图2 PSA 模块结构

Fig.2 PSA module structure

引入仅通道注意力,可以帮助网络获取不同通道之间的关联性,进而能够更充分地捕获输入特征的深层语义信息,计算式如下:

$$A^{ch}(X) = F_{SG} [S_{z1}(\sigma_1(S_v(X))) \times F_{SM}(\sigma_2(S_q(X)))] \quad (6)$$

式中: S_q 、 S_v 和 S_z 为不同的 1×1 卷积操作; σ_1 和 σ_2 为张量重塑算子; $F_{SM}(\cdot)$ 为 Softmax 操作; θ_1 为通道卷积的中间参数; $F_{SG}(\cdot)$ 为 Sigmoid 操作。

引入仅空间注意力,可以利用输入特征之间的空间位置关系,更充分地捕获局部细节信息,计算式如下:

$$A^{sp}(X) = F_{SG} [\sigma_3(F_{SM}(\sigma_1(F_{GP}S_q(X)))) \times \sigma_2(S_v(X))] \quad (7)$$

式中: σ_3 为张量重塑算子; $F_{GP}(\cdot)$ 为全局池化操作。将2个分支的输出并行连接后,极化自注意力公式如下:

$$PSA(X) = Z^{ch} + Z^{sp} = A^{ch}(X) \odot^{ch} X + A^{sp}(X) \odot^{sp} X \quad (8)$$

式中: \odot^{ch} 和 \odot^{sp} 分别为 channel-wise 和 spatial-wise 乘法运算符; Z^{ch} 和 Z^{sp} 分别为经过仅通道和仅空间注意力的输出特征图。

1.3 DCNv2

在卷积神经网络中,采用卷积核作为特征提取器,对输入样本进行卷积操作来捕获其内在特征。然而,在矿下,由于矿工目标通常处于运动状态,传

统卷积很难适应目标变化,导致模型的检测精度下降。DCNv2 是在卷积层通过卷积操作,捕获输入特征图中目标的变形信息,如图3所示。假设卷积核的大小为 3×3 ,采样点有9个,每采样点都被赋予偏移量,在提取图像特征时,保持相同的感受野大小和卷积计算的有效性。DCNv2 的参数比传统卷积更加紧凑,并且卷积核的形状可以根据输入特征自适应调整,因此,更适合处理不同形状和位置的特征。

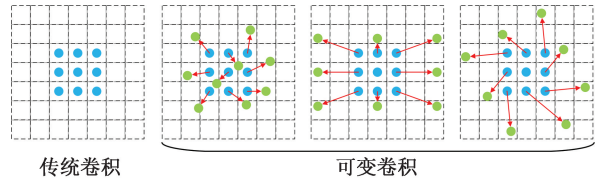


图3 传统卷积与 DCNv2 采样对比

Fig.3 Comparison between traditional convolution and DCNv2 sampling

DCNv2 在某一采样点 l_0 的计算过程如下:

$$y(l_0) = \sum_{l_n \in R} w(l_n) X(l_0 + l_n + \Delta l_n) \Delta m_k \quad (9)$$

式中: X 为输入特征图; w 为采样位置的权重; R 为卷积核的区域; l_n 为 R 中相对于 l_0 的相对位置; Δl_n 为偏移量。偏移量 Δl_n 根据目标特征的分布来移动区域 R 中的点,通常情况下,偏移量 Δl_n 为小数,这导致输入特征图上像素位置 X 的值无法直接获得,需要通过双线性插值估计,将可变形卷积的公式转换如下:

$$x(l) = \sum_q G(q, l) X(q) \quad (10)$$

式中: $l=l_0+l_n+\Delta l_n$ 为偏移区域的任意位置; q 为特征图映射中的所有积分空间位置; $X(q)$ 为特征图中全部整数位置上点的取值; $G(q, l)$ 为单个二维卷积核的双线性插值函数。

文中提出改进方案,使用 DCNv2 替代特征融合层 C3 模块中的常规卷积,提取不同尺度目标更多的特征信息,改进后的 C3-DCNv2 模块如图4所示。

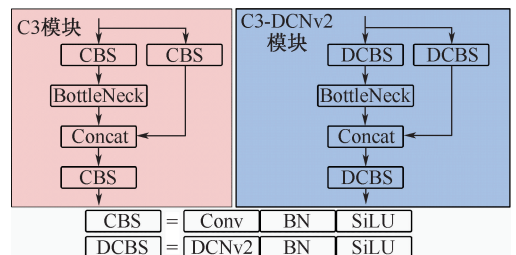


图4 C3 模块与 C3-DCNv2 模块结构

Fig.4 C3 module and C3-DCNv2 module structure

2 目标定位算法设计

2.1 检测与定位系统框架

采用高性能几何关系定位模型。首先,基于成像原理推导坐标系转换关系;然后,结合红外相机与彩色相机的内外参数标定结果,构建矿工目标测量模型;最后,根据彩色相机的检测结果选取定位点,并融合深度信息计算矿工目标的三维坐标,具体定位流程如图5所示。

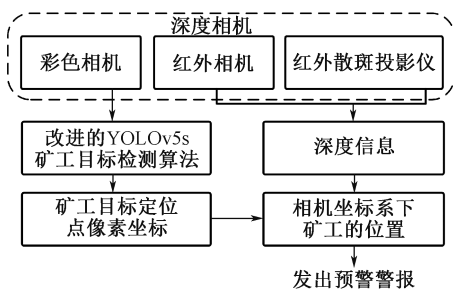


图5 矿工目标检测与定位流程

Fig. 5 Miner target detection and localization process

2.2 深度测量原理

深度相机使用投影仪投射红外激光,从而在物体表面形成随深度变化的散斑图案。随后,红外相机捕获这些图案,以生成包含距离信息的深度图,测量原理如图6所示。

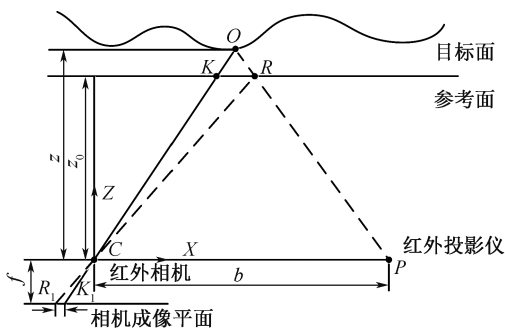


图6 红外相机测距原理

Fig. 6 Infrared camera ranging principle

由图6可知:对于激光发出的同一射线 PO ,经过距离为 z_0 的参考面上 R 点,在红外相机上的成像点为 R_1 ;经过距离为 z 的目标面上 O 点,在红外相机上的成像点为 K_1 。已知世界坐标系中的点 (x, y, z) 转换成像素坐标系的点为 (u, v) ,焦距为 f ,视差 d ,基线距离为 b ,则该点距离红外相机的深度 z 计算公式:

$$z = \frac{z_0}{1 + \frac{z_0}{fb}d} \quad (11)$$

根据式(11)及相机内部参数,得到该点在相机坐标中的 x, y 的值,计算公式:

$$x = \frac{(u - c_x)}{f}z, y = \frac{(v - c_y)}{f}z \quad (12)$$

式中 c_x 和 c_y 为相机光心在像素坐标系下的水平和垂直方向上的位置,根据相机内参,计算出相机坐标系矿工目标的空间坐标。

深度相机由彩色相机和红外相机组成,彩色相机用于捕捉彩色图像,而红外相机则用于捕获由红外投影仪投射的红外散斑图,生成深度图像。然而,彩色图像和深度图之间存在一定的偏移,导致彩色图像上的像素点与正确的深度信息之间存在不一致。为解决此问题,通过彩色相机和红外相机的外参矩阵,将2个相机的坐标系重合,使得捕捉的彩色图像对应正确的深度信息。

2.3 目标测距算法

YOLOv5s-MPD模型可以输出检测矿工目标的检测框和置信度,检测框的左上角顶点坐标为 (u_{\max}, v_{\max}) ,右下角顶点坐标为 (u_{\min}, v_{\min}) 。选取矿工目标检测框中心的坐标,记作 (u_t, v_t) ,将其作为距离定位点,计算公式如下:

$$\begin{cases} u_t = (u_{\max} + u_{\min})/2 \\ v_t = (v_{\max} + v_{\min})/2 \end{cases} \quad (13)$$

在目标测距模型中,相机中心点的世界坐标为 (x_0, y_0, z_0) ,目标特征点的世界坐标为 (x_t, y_t, z_t) 。通过求解空间中2点的欧氏距离完成矿工目标距离的计算,计算公式如下:

$$L_0 = \sqrt{(x_t - x_0)^2 + (y_t - y_0)^2 + (z_t - z_0)^2} \quad (14)$$

3 模型训练与结果分析

3.1 数据集的构建及训练环境

在Windows10操作系统下完成模型训练,学习框架为PyTorch,试验配置见表1。

试验采用自制数据集,数据采集地点为安徽省淮南市某矿井。经过筛选后,共得到3674张图像的原始数据集,每张图像至少包含1个目标。为保证数据集划分的均衡性与合理性,采用分层采样方法将数据集按8:1:1的比例划分为训练集、验证集和测试集,确保三者之间的样本独立性和分布一致性。

3.2 评价指标

模型的评价的指标包括查准率 P 、召回率 R 、

表1 配置信息

Table 1 Configuration information

名称	配置
CPU	Intel Core i5-13600KF
GPU	NVIDIA GeForce RTX4060
PyTorch	1.12.1
CUDA	12.2

注:中央处理器(Central Processing Unit, CPU);图形处理器(Graphic Processing Unit, GPU);统一计算设备架构(Compute Unified Device Architecture, CUDA)。

mAP及帧率(Frames Per Second, FPS),各指标的计算式:

$$P = \frac{TP}{TP + FP} \quad (15)$$

$$R = \frac{TP}{TP + FN} \quad (16)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n \int_0^1 P(R) dR \quad (17)$$

$$FPS = \frac{Q}{T} \quad (18)$$

式中:TP(True Precision)、FP(False Precision)、FN(False Negative)分别为真正例、假正例、假反例对应的样例数; n 为类比总数; Q 为总帧数,幅; T 为检测时间,s。

3.3 训练过程

在模型训练前,将输入图片设置为640pixel×640pixel,批处理大小为16,初始学习率为0.01,动量参数为0.937,权重衰减系数为0.0005,训练迭代次数为300,使用交并比作为目标检测的评价指标,并设置0.6作为阈值来区分正负样本,采用随机

梯度下降作为优化器。

3.4 施加的不同注意力模块对模型效果的影响

为验证文中引入注意力机制的有效性,在网络特征融合层末端引入不同类型的注意力模块,包括卷积注意力(Convolutional Block Attention Module, CBAM)、坐标注意力(Channel Attention, CA)、通道注意力(Effective Channel Attention, ECA)、SE和PSA,试验结果见表2。

表2 注意力模块不同算法训练对比结果

Table 2 Comparison results of different algorithm training for the attention module

模型	查准率	召回率	mAP
YOLOv5s-MobileNetv3	0.859	0.688	0.79
+SE	0.868	0.677	0.792
+CBAM	0.864	0.674	0.785
+CA	0.881	0.661	0.787
+ECA	0.880	0.664	0.787
+PSA	0.872	0.702	0.811

由表2可知:CBAM、CA和ECA模块对mAP造成轻微下降,分别降低0.63%、0.38%和0.38%,当引入SE和PSA模块时,mAP分别提高0.25%和2.66%。因此,引入PSA模块的P、R和mAP总体优于其他改进模型。

3.5 消融试验

为验证YOLOv5s-MPD各改进策略的有效性,在数据集、训练参数、试验环境相同的基础上进行消融对比,结果见表3。

表3 消融试验结果

Table 3 Ablation test results

模型	改进策略			查准率	计算量/ G	参数量/ M	模型体积/ MB	mAP	FPS/ (帧·s ⁻¹)
	MobileNetv3	PSA	DCNv2						
YOLOv5s	—	—	—	0.891	15.8	7.01	13.7	0.848	45.3
	√	—	—	0.859	2.5	1.39	3.0	0.79	77.2
	√	√	—	0.872	2.9	1.54	3.3	0.811	67.4
	√	—	√	0.869	2.2	1.45	3.1	0.806	71.8
YOLOv5s-MPD	√	√	√	0.884	2.6	1.61	3.4	0.825	70.2

由表3可知:使用MobileNetv3代替原YOLOv5s网络主干,模型参数量下降80.17%,计算量减少84.18%,模型体积减小78.10%,推理速度提升70.42%,说明更换主干网络以降低模型复杂度是实现模型轻量化、提升推理速度的有效方法。此外,引入PSA模块,使查准率和检测精度分

别提高1.51%和2.66%,而模型的参数量、计算量和模型体积并没有大幅度增加,因为PSA模块将空间注意力和通道注意力相结合,增强了模型对矿工目标特征学习能力,一定程度上弥补了轻量化带来的精度上的损失。同样,将DCNv2替代特征融合网络层中C3模块的标准卷积,虽然改进模

3.6 不同模型对比

为验证改进模型的优越性,选取 YOLOv5s、YOLOv5n、YOLOv5m、YOLOv7tiny、YOLOv8s、Faster R-CNN 和 SSD 共 7 个主流目标检测模型进行训练,从参数量、计算量、模型体积、mAP 和检测速度对比改进模型,以验证改进模型性能,结果见表 4。

表 4 不同模型对比结果

Table 4 Comparison results of different models

模型	参数量/M	计算量/G	模型体积/MB	mAP	FPS/(帧·s ⁻¹)
YOLOv5s	7.01	15.8	13.7	0.848	45.3
YOLOv5n	1.76	4.1	3.8	0.792	61.8
YOLOv5m	20.85	47.9	42.1	0.857	24.5
YOLOv7tiny	6.01	13.2	12.0	0.833	30.2
YOLOv8s	11.13	28.4	22.0	0.851	39.2
Faster R-CNN	136.69	401.7	110.8	0.795	13.4
SSD	23.61	273.2	92.8	0.764	17.2
YOLOv5s-MPD	1.61	2.6	3.4	0.825	70.2

由表 4 可知:改进模型参数量与计算量最小,模型参数量仅为 1.61 M,与主流模型相比分别减少 77.0%、8.5%、92.3%、73.2%、85.5%、98.8% 和 93.2%;模型计算量为 2.6 G,与主流模型相比分别减少 83.5%、36.6%、94.6%、80.3%、90.8%、99.4% 和 99.0%;从模型大小来看,改进模型的体积是最小的,仅为 3.4 MB。改进模型的 mAP 为 0.825,与 YOLOv5n、Fater R-CNN 和 SSD 相比,分别提高 4.2%、3.8% 和 8.0%,略低于其他模型。改进模型在检测时间上花费最短,同时在参数量、计算量和模型体积方面表现出更明显的优势,但在检测效果方面与主流模型相当,实现了检测速度和精度的有效平衡,能够实现对井下矿工目标的实时准确检测。

4 矿工目标定位

4.1 相机标定

在矿工目标定位中,选用深度相机标定板规格为 12×9(单元边长为 10 mm)。利用深度相机的彩色相机和红外相机,分别对标定板进行拍摄,以采集不同姿态的图像,每种相机各拍摄 30 张图像。为了保证所拍摄的图像较为清晰,需要将标定板置于相机的视场范围之内,并补充光源。采用张正友标定法建立了三维世界坐标点与二维图像坐标点的映射关系,并计算相机的内外参数。

4.2 试验结果与分析

进行矿工目标检测与定位试验时,在 1~8 m 距

离范围内,成功检测出矿工目标后,记录矿工目标中心点的定位坐标,以目标测距算法式(14)对测量数据进行评价分析,采用高精度测距仪(激光测距仪,精度为 10 m 范围内最大测量误差为±2 mm)获取到矿工目标中心点的距离作为真实距离。因为真实矿井环境下电子设备需要做防爆炸处理,所以在模拟矿井环境试验场地进行测试,并对测量矿工目标的距离误差进行分析,结果见表 5。

表 5 目标定位测距结果

Table 5 Target localization ranging results

编号	定位坐标	预测距离/m	测量距离/m	绝对误差/m	相对误差/%
1	(-45,11,1112)	1.11	1.11	0.00	0.27
2	(-61,37,2125)	2.13	2.12	0.01	0.24
3	(-430,124,3056)	3.09	3.13	0.06	1.25
4	(354,-162,4174)	4.19	4.12	0.07	1.77
5	(-412,325,5012)	5.04	4.96	0.08	1.70
6	(601,412,6213)	6.26	6.44	0.18	2.88
7	(1054,-141,7019)	7.10	7.30	0.20	2.88
8	(-723,-211,8622)	8.66	8.41	0.25	2.96

由表 5 可知:在 1~8 m 距离范围对矿工目标进行定位,测得距离的平均绝对误差和平均相对误差分别为 0.11 m 和 1.74%,最大绝对误差和最大相对误差分别为 0.25 m 和 2.96%。同时,由表 5 可以看出,随着矿工目标与相机之间的距离增加,其误差也相应的增加。主要原因是红外相机的分辨率有限,影响了测量的定位精度,但整体测量误差仍满足矿工目标定位的要求。

在动态检测定位中,确保单矿工目标和多矿工目标随机出现在相机视场范围内。测试过程中目标出现 80 次,成功检测出 78 次,检测成功率达 97.5%。如图 9 所示,检测对象为矿工目标,检测框中心的圆点为矿工目标的定位点,坐标值表示该定位点在三维空间中的具体位置。

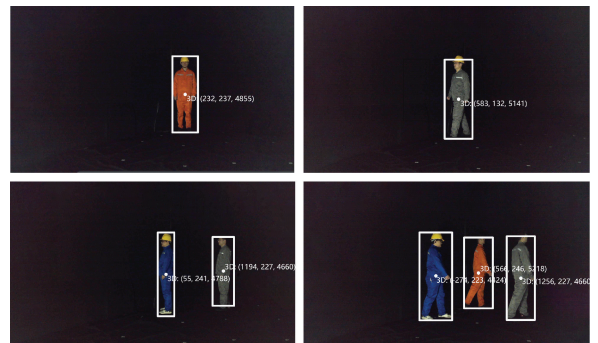


图 9 矿工目标实时检测与定位可视化结果
Fig. 9 Real-time detection and localization visualization of miner targets

5 结 论

1) 提出一种深度学习结合深度相机的目标实时检测和精准定位方法,该方法能够提高检测速度并确保矿工目标定位的高精度,自动检测靠近危险区域的矿工目标,实现安全预警。

2) 在测试样本上,改进模型在复杂矿井环境下对矿工目标具有良好的检测效果,相较于 YOLOv5s 模型,改进模型体积缩减至 3.4 MB,缩小 75.18%,模型的计算量和参数量分别减少 83.54% 和

77.03%,mAP 为 0.825,检测速度为 70.2 帧/s,提升 54.97%,与主流目标检测模型相比较,改进模型实现了在参数量、计算量、模型体积、检测速度和检测精度的有效平衡。在矿工目标定位试验中,将矿工目标与相机间距离保持在 1~8 m 范围内,测得矿工目标与相机间距离的平均绝对误差和平均相对误差分别为 0.11 m 和 1.74%;最大绝对误差和最大相对误差分别为 0.25 m 和 2.96%,动态目标检测率达 97.5%,能满足实际应用场景下对矿工目标的实时检测与定位需求。

参 考 文 献

- [1] LI Ailing, ZHANG Jixiong, ZHOU Nan, et al. A model for evaluating the production system of an intelligent mine based on unascertained measurement theory [J]. *Journal of Intelligent & Fuzzy Systems*, 2020, 38(2): 1 865-1 875.
- [2] ZHANG Kexue, KANG Lei, CHEN Xuexi, et al. A review of intelligent unmanned mining current situation and development trend [J]. *Energies*, 2022, 15(2): DOI: 10.3390/en15020513.
- [3] 付玉平, 陈兆波, 赵振保, 等. 煤矿安全事故的综合致因模型[J]. *安全与环境学报*, 2024, 24(7): 2 731-2 740. FU Yuping, CHEN Zhaobo, ZHAO Zhenbao, et al. Research on the comprehensive causative model of coal mine safety accidents [J]. *Journal of Safety and Environment*, 2024, 24(7): 2 731-2 740.
- [4] 王国庆, 赵鑫, 杨春雨, 等. 煤矿地下空间定位技术研究进展[J]. *工程科学学报*, 2024, 46(10): 1 713-1 727. WANG Guoqing, ZHAO Xin, YANG Chunyu, et al. Recent advances in research on underground space positioning technology for coal mining [J]. *Chinese Journal of Engineering*, 2024, 46(10): 1 713-1 727.
- [5] 孙哲星. 煤矿井下人员精确定位方法[J]. *煤炭科学技术*, 2018, 46(3): 130-134. SUN Zhexing. Personnel position method in underground coal mine [J]. *Coal Science and Technology*, 2018, 46(3): 130-134.
- [6] 李梅, 姜展, 姜龙飞, 等. 三维可视化技术在智慧矿山领域的研究进展[J]. *煤炭科学技术*, 2021, 49(2): 153-162. LI Mei, JIANG Zhan, JIANG Longfei, et al. Research progress on 3D visualization technology for intelligent mine [J]. *Coal Science and Technology*, 2021, 49(2): 153-162.
- [7] 左明成, 焦文华. 面向煤矿井下作业场景的安全帽佩戴识别算法[J]. *中国安全科学学报*, 2024, 34(3): 237-246. ZUO Mingcheng, JIAO Wenhua. Helmet-wearing recognition algorithm for coal mine underground operation scenarios [J]. *China Safety Science Journal*, 2024, 34(3): 237-246.
- [8] 张磊, 李熙尉, 燕倩如, 等. 基于改进 YOLOv5s 的综采工作面人员检测算法[J]. *中国安全科学学报*, 2023, 33(7): 82-89. ZHANG Lei, LI Xiwei, YAN Qianru, et al. Personnel detection algorithm in fully mechanized coal face based on improved YOLOv5s [J]. *China Safety Science Journal*, 2023, 33(7): 82-89.
- [9] XIN Fangfang, HE Xinyu, YAO Chaoxiu, et al. A real-time detection for miner behavior via DYS-YOLOv8n model [J]. *Journal of Real-Time Image Processing*, 2024, 21(3): DOI: 10.1007/s11554-024-01466-0.
- [10] 张春堂, 管利聪. 基于 SSD-MobileNet 的矿工安保穿戴设备检测系统[J]. *工矿自动化*, 2019, 45(6): 96-100. ZHANG Chuntang, GUAN Licong. Detection system of miners' wearable security equipments based on SSD-MobileNet [J]. *Industry and Mine Automation*, 2019, 45(6): 96-100.
- [11] 郭永存, 董佳乐, 王爽. 井下无人驾驶电机车行驶场景中多目标检测研究[J]. *工矿自动化*, 2022, 48(6): 56-63. GUO Yongcun, TONG Jiale, WANG Shuang. Research on multi-object detection in driving scene of underground

- unmanned electric locomotive [J]. Journal of Mine Automation, 2022, 48(6): 56-63.
- [12] 李伟山, 卫晨, 王琳. 改进的 Faster RCNN 煤矿井下行人检测算法[J]. 计算机工程与应用, 2019, 55(4): 200-207.
LI Weishan, WEI Chen, WANG Lin. Improved Faster RCNN approach for pedestrian detection in underground coal mine [J]. Computer Engineering and Applications, 2019, 55(4): 200-207.
- [13] 王科平, 连凯海, 杨艺, 等. 基于改进 YOLOv4 的综采工作面目标检测[J]. 工矿自动化, 2023, 49(2): 70-76.
WANG Keping, LIAN Kaihai, YANG Yi, et al. Target detection of the fully mechanized working face based on improved YOLOv4 [J]. Journal of Mine Automation, 2023, 49(2): 70-76.
- [14] 郝明月, 闵冰冰, 张新建, 等. 基于改进 YOLOv5s 的矿工排队检测方法[J]. 工矿自动化, 2023, 49(11): 160-166.
HAO Mingyue, MIN Bingbing, ZHANG Xinjian, et al. A miner queue detection method based on improved YOLOv5s [J]. Journal of Mine Automation, 2023, 49(11): 160-166.
- [15] 杨豚, 郭永存, 王爽, 等. 煤矿井下无人驾驶轨道电机车障碍物识别[J]. 浙江大学学报:工学版, 2024, 58(1): 29-39.
YANG Tun, GUO Yongcun, WANG Shuang, et al. Obstacle recognition of unmanned rail electric locomotive in underground coal mine [J]. Journal of Zhejiang University: Engineering Science, 2024, 58(1): 29-39.
- [16] 郭曦, 谢炜, 朱红秀, 等. 井下目标跟踪与测距方法研究[J]. 煤炭工程, 2019, 51(3): 117-121.
GUO Xi, XIE Wei, ZHU Hongxiu, et al. Target tracking and ranging method in underground mine [J]. Coal Engineering, 2019, 51(3): 117-121.
- [17] ZHOU Cheng, REN Dacong, ZHANG Xiangyan, et al. Human position detection based on depth camera image information in mechanical safety [J]. Advances in Mathematical Physics, 2022, 2022(1): DOI: 10.1155/2022/9170642.
- [18] 韩江洪, 袁稼轩, 卫星, 等. 基于深度学习的井下巷道行人视觉定位算法[J]. 计算机应用, 2019, 39(3): 688-694.
HAN Jianghong, YUAN Jiakuan, WEI Xing, et al. Pedestrian visual positioning algorithm for underground roadway based on deep learning [J]. Journal of Computer Applications, 2019, 39(3): 688-694.



作者简介: 张森 (1986—),男,河南项城人,博士,讲师,主要从事机器视觉、人机工程等方面的研究。E-mail: miao@ahjzu.edu.cn。