

DOI: 10.13382/j.jemi.B2205785

融合注意力机制与逆残差结构的 轻量级钻机目标检测方法*

张 栋¹ 姜媛媛^{1,2}

(1. 安徽理工大学电气与信息工程学院 淮南 232001;
2. 安徽理工大学环境友好材料与职业健康研究院(芜湖) 芜湖 241003)

摘 要:为实现煤矿下定向钻进钻机钻孔深度的精准测量,提出一种融合注意力机制与逆残差结构的轻量级钻机目标检测网络(GCI-YOLOv4),通过自动、快速及准确检测记录钻机的运动轨迹,获取打入钻杆数量,计算出钻孔深度。针对煤矿下色域区分度低问题,采用 GhostNet 作为特征提取网络去除复杂背景的冗余特征,同时轻量化模型,加快推理速度。针对煤矿井下光照不均导致钻机目标显著度低的问题,引入注意力模块增强钻机在复杂背景中的显著度。针对钻机高速运动时难以被准确检测的问题,引入逆残差结构,提取更丰富语义特征的同时保持速度与精度的均衡。为保证模型的准确性和可靠性,将提出的检测算法与5种经典目标检测算法进行对比。实验结果表明,GCI-YOLOv4 可以较好的解决煤矿下背景色域区分度低、钻机高速运动以及受光照不均等问题,平均检测精度达到99.49%,检测速度达到58.10 FPS,性能优于经典目标检测算法。将 GCI-YOLOv4 部署在工作面现场进行测试,能够准确获取钻机的运动轨迹,通过滤波处理统计上升沿计算钻杆数量,钻杆计数精度达到99.4%,精确计算出钻孔深度,验证了该方法的可行性和实用性。

关键词: 目标检测;YOLOv4;GhostNet;注意力模块;逆残差结构;钻杆计数

中图分类号: TP183;P634.3[†]1 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Lightweight target detection method of drilling rig based on attention mechanism and inverse residual structure

Zhang Dong¹ Jiang Yuanyuan^{1,2}

(1. School of Electrical and Information Engineering, Anhui University of Science and Technology, Huainan 232001, China;
2. Institute of Environment-Friendly Materials and Occupational Health, Anhui University of
Science and Technology, Wuhu 241003, China)

Abstract: In order to realize the accurate measurement of drilling depth of directional drilling rig under coal mine, a lightweight drilling rig target detection network integrating attention mechanism and inverse residual structure (GCI-YOLOv4) is proposed. Through automatic, rapid and accurate detection, the movement track of drilling rig, the number of driven drill rods and the drilling depth are obtained. Aiming at the problem of low color gamut discrimination in coal mine, GhostNet is used as the feature extraction network to remove the redundant features of complex background, lighten the model and accelerate the speed of model reasoning. Aiming at the problem of low target saliency of drilling rig caused by uneven illumination in coal mine, the attention module is introduced to enhance the saliency of drilling rig in complex background. Aiming at the problem that it is difficult to detect accurately when the drilling rig moves at high speed, the inverse residual structure is introduced to extract richer semantic features while maintaining the balance between speed and accuracy. In order to ensure the accuracy and reliability of the model, the proposed detection algorithm is compared with five classical target detection algorithms. The experimental results show that the proposed detection algorithm can better solve the problems of low background gamut discrimination, high-speed movement of drilling rig and uneven illumination under coal mine. The average detection accuracy is 99.49% and the detection speed is 58.10 FPS. The performance is better than the classical target detection

收稿日期: 2022-08-29 Received Date: 2022-08-29

* 基金项目:安徽省重点研究与开发计划(202104g01020012)、安徽理工大学环境友好材料与职业健康研究院研发专项基金(ALW2020YF18)项目资助

algorithm. The proposed detection algorithm is deployed in the field of the working face for testing, which can accurately obtain the motion trajectory of the drilling rig. The number of drill pipes is calculated by filtering and counting the rising edge. The counting accuracy of drill pipes is 99.4%. The drilling depth is accurately calculated, which verifies the feasibility and practicability of this method.

Keywords: object detection; YOLOv4; GhostNet; attention module; inverted residuals; drill pipe count

0 引 言

突出煤层掘进煤巷之前需要对预掘煤巷区域采取煤与瓦斯突出防止措施,最主要的措施是瓦斯抽采^[1]。瓦斯抽采对提高煤矿井下采掘效率、保障矿井生产安全有着至关重要的作用^[2],能够降低煤层瓦斯的含量和压力,消除或降低采掘工作面煤与瓦斯突出危险性。煤层瓦斯预抽主要通过并在井下施工抽采钻孔来实施,在钻机抽采过程中,煤层内部原有的应力平衡被打破,围岩应力沿巷道径向往煤层深部依次呈现出应力降低区、应力升高区、原始应力区^[3]。当钻孔深度到达应力降低区时,孔内瓦斯抽取时会在钻孔周围产生泄露通道,导致瓦斯抽采效率降低。当钻孔深度到达应力升高区时,钻孔周围煤层在应力作用下发生变形和破碎,导致钻孔塌孔,形成应力屏障,阻断钻孔末端到应力升高区的瓦斯流动,造成抽采盲区。很多矿区由于突出煤层,存在钻进深度不达标的情况,钻孔深度的实时测量一直是瓦斯抽采环节的重点和难点,极大制约煤矿的高效生产和瓦斯的高效抽采。

钻孔深度的测量方法主要是通过钻杆计数计算出钻孔深度,包括人工计数和随钻测量技术^[4]。人工计数通过打入钻杆数量计算出钻孔深度,耗时耗力,自动化水平低。路拓等^[5]通过弹性波在钻杆中的传播速度分别从时间域和频率域计算钻杆的长度,从而确定钻孔深度。王岚等^[6]基于光电编码器设计,将钻杆位移量转换成计数脉冲送入单片机计算获得钻孔实时深度。王小龙等^[7]采用基于钻孔静水压力数据结合钻孔倾角数据的方法确定钻孔深度。随钻测量技术需额外安装测量设备,安装布线困难,成本高昂。

随着计算机视觉技术的发展,越来越多的学者在已有的视频监控系统上利用图像处理算法实现智能化生产^[8]。彭业勋等^[9]提出了一种角点检测的方法,结合金字塔光流法对运动钻机进行光流场估计,采用跳帧的思想实现计数功能;董立红等^[10]提出使用改进的 Camshift 算法实时捕捉钻杆目标进行钻杆计数;方杰等^[11]使用 ECO-HC 算法实时跟踪钻杆目标,从而实现钻杆计数。以上方法主要依赖人工特征提取,但限于矿井下环境复杂,存在光照不均、煤尘干扰严重等问题,容易丢失目标,造成统计结果偏差。

近年来,深度学习的发展及其在图像任务上的成功应用,部分学者将其应用在钻杆计数任务。高端等^[12]提出通过改进 ResNet 网络进行二分类,采用积分法对视频分类置信度进行滤波,统计置信度曲线下降沿数量实现钻杆计数。党伟超等^[13]提出利用三维卷积神经网络对钻机的卸杆动作进行识别。以上方法主要是对工人装卸杆动作进行区分,但在不同工作面下的普适性较弱,须在瓦斯钻取工艺稳定的工况下才能实现有效应用。

在钻机工作过程中,钻机卡盘在直线方向上做往复周期运动,每当钻机卡盘完成一个来回,钻机钻进或者退出一根钻杆,故可以通过目标检测算法反馈钻机的运动轨迹,根据钻机往复次数计算出钻杆数量,从而测量出钻孔深度。目前,目标检测算法主要分为两类:第 1 类是 two-stage,代表算法有 R-CNN 系列算法,该类算法先生成候选框,再对候选框进行分类和回归,通常速度较慢,并不能满足钻机实时目标检测的要求;第 2 类是 one-stage,代表算法有 YOLO(you only look once)^[14] 系列算法,该类算法直接对网络进行端到端的训练,实时性高。YOLOv4^[15]作为目前主流网络之一,在检测精度和检测速度上表现优良。张明路等^[16]提出用逆残差结构代替 YOLOv4 主干网络降低模型参数提高电子元器件检测精度;刘艳菊等^[17]提出使用 MobileNetv3 替换主干特征提取网络使 YOLOv4 实现热轧钢条表面实时缺陷检测;彭继慎等^[18]在 YOLOv4 主干网络中加入注意力机制,运用深度可分离卷积替换传统卷积,并且压缩模型使其适用于电力巡检无人机避障。

综上所述,针对钻孔深度难以实时测量问题,提出一种改进的 GCI-YOLOv4 (Ghost-CBAM-Inverted residuals-YOLOv4)用于钻机目标检测,通过获取钻机的实时位置信息获取钻孔深度,提高瓦斯抽采效率,保障矿井生产安全。为了解决煤矿下色域区分度低问题,采用 GhostNet 作为特征提取网络,去除复杂背景的冗余特征,同时轻量化模型,加快模型推理速度。针对钻机运动过程中受光照不均影响使其难以被准确检测的问题,引入注意力模块增强钻机在复杂背景中的显著度,从而提高钻机的特征表达能力,进而提高检测的精度。针对钻机高速运动时钻机难以被准确检测的问题,引入逆残差结构,通过 1×1 卷积升维浓缩特征,深度可分离卷积精简原检测网络模块,达到检测速度和精度平衡。

1 YOLOv4 目标检测算法

YOLOv4 模型主要由输入端、特征提取网络、特征融合颈部、检测头 4 个部分组成。YOLOv4 网络在训练时加入 mosaic 数据增强, 分别对 4 张图片进行翻转、缩放、色域变化等操作后拼接成一张图片, 丰富检测图片的背景, 增加学习样本的多样性。YOLOv4 以 CSPDarkNet 网络作为特征提取网络, 由 CBM(由卷积、批归一化和 Mish 激活函数组成) 和 CBL(由卷积、批归一化和 Leaky-ReLU 激活函数组成) 模块组成 CSP 模块进行特征提取。YOLOv4 在特征提取网络和最后的输出层之间加入空间金字塔池化模块(SPP)与路径聚合网络结构(PAN)构成

特征融合颈部。

SPP 模块主要利用不同大小的池化核增加感受野, 分离出显著的上下文特征。通过 PAN 完成从底到顶的特征提取和从顶到底的特征信息融合。YOLOv4 在特征提取部分提取多特征层进行目标检测, 一共提取 3 个 shape 为 $(52, 52, 256)$ 、 $(26, 26, 512)$ 、 $(13, 13, 1024)$ 的特征层, 通过特征融合颈部获得 3 个 shape 为 $(13, 13, 75)$ 、 $(26, 26, 75)$ 、 $(52, 52, 75)$ 的输出。YOLOv4 的 3 个特征层将图片分为 13×13 、 26×26 、 52×52 的网格, 每个网格负责一个区域的检测, 获得图片中物体的具体位置, 采用 CIOU_Loss (complete-IOU_Loss) 来估算检测目标框的识别损失。YOLOv4 整体网络结构如图 1 所示。

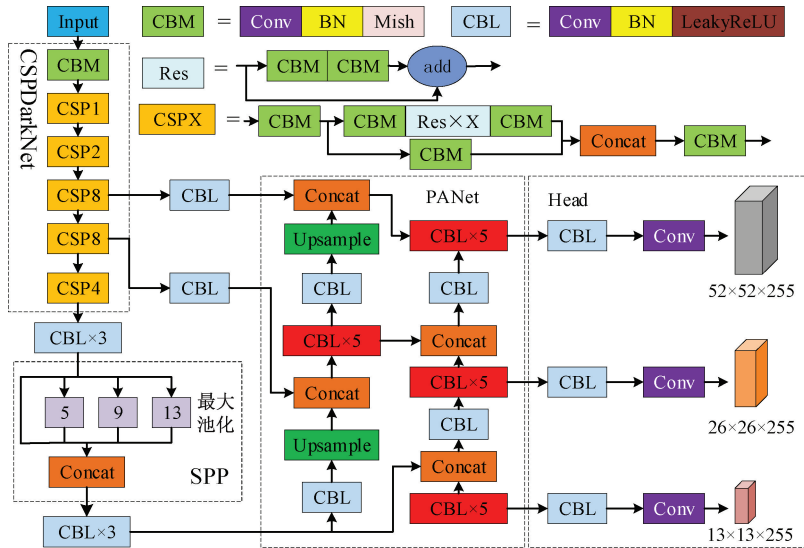


图 1 YOLOv4 网络结构

Fig. 1 YOLOv4 network structure

2 YOLOv4 算法改进

钻机目标检测的核心任务是准确且快速的检测出煤矿下实时钻机的运动轨迹。而 YOLOv4 在煤矿井下这一特殊环境里的检测效果还不够理想。因此针对煤矿下钻机检测任务对 YOLOv4 做出改进, 提出了一种增强型轻量级钻机目标检测网络 GCI-YOLOv4。

2.1 特征提取网络改进

尽管 YOLOv4 的特征提取网络 CSPDarkNet 特征提取能力突出, 但其在网络优化过程中的梯度信息重复将导致计算量较大、推理成本过高。在钻机目标检测任务中, 煤矿下环境昏暗, 颜色单一, 不像地面检测任务有着明显的色域区分度, 导致大量重复的特征。这些冗余的特征对于钻机检测效果提升有限, 而且增加模型的计算

量。本文将 YOLOv4 特征提取网络 CSPDarkNet 替换为 GhostNet^[19], 在模型轻量化的同时能够保持较强的特征提取能力。

Ghost 卷积分为常规卷积、Ghost 生成和特征图拼接 3 步。

1) 首先, 假设输入特征图尺寸是 $H \times W \times c$, 输出特征图的尺寸是 $H' \times W' \times n$, 卷积核大小为 $k \times k$ 。用常规卷积得到本征特征图 $Y_{H' \times W' \times m}$, 这部分的计算量约等于 $H \times W \times c \times m \times W' \times H'$ 。

2) 然后, 将本征特征图 $Y_{H' \times W' \times m}$ 中每一个通道的特征图用 Φ_i 操作来产生 Ghost 特征图。在线性变换 (Φ_i) 中, 假设特征图的通道数为 m , 变换的数量为 s , 最终得到的新的特征图的数量为 $n = m \cdot s$ 。

由于 Ghost 模块的变换过程中最后存在一个恒等变换 (Identity), 所以实际有效的变换数量是 $(s - 1)$, 可以

得到:

$$m(s-1) = \frac{n}{s(s-1)} \quad (1)$$

3) 最后, 将步骤 1) 得到的本征特征图和步骤 2) 得到的 Ghost 特征图拼接 (Identity 连接) 得到最终结果。

当输出特征图的通道数量远大于本征特征图的通道数时 (即 $n \gg m$), 可以计算得到普通卷积与 Ghost 卷积的计算量比值 $r_s \approx s$ 。Ghost 模块结构如图 2 所示。

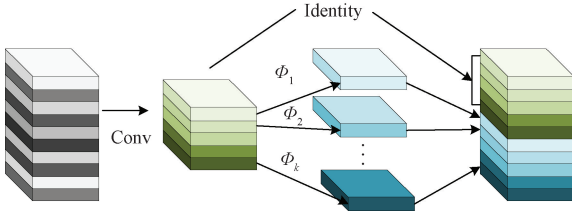


图 2 Ghost 模块

Fig. 2 Ghost module

相比于常规卷积, Ghost 卷积可以去除复杂环境里的冗余特征, 而且计算量大幅度降低。利用 Ghost 卷积这些特征, 采用 GhostNet 模块作为特征提取网络提取 3 个 shape 为 (52, 52, 40)、(26, 26, 112)、(13, 13, 160) 的特征层, 使整体网络结构在保持适度深度的同时轻量化模型获取有效特征, 更适合钻机目标检测。

2.2 加入注意力模块

为了解决煤矿井下光照不均导致的钻机目标显著度低的问题, 引入卷积注意力模块 (convolutional block attention module, CBAM)^[20], 获取特征图的通道信息和空间信息重构特征图, 提取钻机目标的显著特征, 并加强显著特征的权重, 弱化不必要特征的权重, 从而获得全面可靠的注意力信息, 提升模型精度。CBAM 结构如图 3 所示。

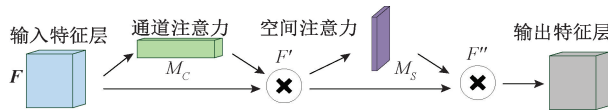


图 3 CBAM 结构

Fig. 3 Structure of convolutional block attention module

假设给定一个输入特征矩阵 $F \in R^{C \times H \times W}$ 。首先通过通道注意力模块, 同时对输入按通道进行全局最大池化和全局平均池化, 分别得到两个不同的 $C \times 1 \times 1$ 一维向量, 通过共享的全连接层将两个一维向量相加, 最后经过 Sigmoid 激活函数, 得到通道注意力向量 $M_C \in R^{C \times 1 \times 1}$, 计算如式 (2) 所示。

$$M_C = \sigma(\text{MLP}(\text{Avgpool}(F))) + \text{MLP}(\text{MaxPool}(F)) \quad (2)$$

式中: σ 表示 Sigmoid 函数, MLP 表示全连接层, AvgPool

和 MaxPool 分别表示对特征图进行全局平均池化和全局最大池化操作。再将通道注意力向量与输入特征矩阵相乘, 获得重构后的特征映射 $F' \in R^{C \times H \times W}$ 。

然后将 F' 作为空间注意力模块的输入, 沿着通道维度同时进行全局平均池化和全局最大池化, 分别得到两种不同的 $1 \times H \times W$ 向量, 将两者进行拼接, 通过一个卷积核为 7×7 的卷积和激活函数, 得到空间注意力向量 $M_S \in R^{1 \times H \times W}$, 计算如式 (3) 所示。

$$M_S = \sigma(f^{7 \times 7}([\text{AvgPool}(F'); \text{MaxPool}(F')])) \quad (3)$$

式中: $f^{7 \times 7}$ 表示 7×7 的卷积运算。将空间注意力向量与 CBAM 模块的输出映射相乘, 得到最终的特征映射 $F'' \in R^{C \times H \times W}$ 。

CBAM 模块带来的参数变化很小, 却能明显提升 YOLOv4 的模型精确率, 本文通过在 GhostNet 提取出的 (52, 52, 40)、(26, 26, 112)、(13, 13, 160) 3 个特征层后加入 CBAM 模块提升钻机特征的显著度。

2.3 PANet 中引入逆残差结构

针对钻机高速运动时难以被准确检测问题, 将 PANet 中的 $CBL \times 5$ 结构修改为逆残差 (inverted residuals) 结构^[21]。逆残差结构在 3×3 深度卷积 (depthwise convolution, Dwise) 前采用 1×1 卷积进行升维操作, 将低维空间特征映射到高维空间, 使输出的特征图包含更丰富的语义特征, 在 3×3 深度卷积后再次采用 1×1 卷积进行降维, 并将最后一个 ReLU 替换为 Linear 线性激活函数, 解决低维度 ReLU 运算产生的信息丢失问题, 保持了精度与速度的均衡。

逆残差结构中的 3×3 深度卷积和 1×1 卷积降维操作构成深度可分离卷积^[22], 大幅度降低了参数量, 显著提升检测速度。假设输入特征图的大小是 $D_F \times D_F \times M$, 输出特征图大小是 $D_F \times D_F \times N$, 其中 D_F 是特征图的宽度与高度, 假设两者是相同的。M 是输入通道数量, N 是输出通道数量, 卷积核的大小是 $D_K \times D_K$ 。标准卷积的计算量是 $D_K \times D_K \times M \times N \times D_F \times D_F$, 深度可分离卷积的计算量是 $D_K \times D_K \times M \times D_F \times D_F + M \times N \times D_F \times D_F$, 深度可分离卷积与标准卷积的计算量比值为:

$$\frac{1}{N} + \frac{1}{D_K^2} \quad (4)$$

YOLOv4 中 PANet 中含有 4 个 $CBL \times 5$ 结构, 5 次卷积的卷积核分别为 1×1 、 3×3 、 1×1 、 3×3 、 1×1 , 使用逆残差结构代替其中 3×3 的部分。逆残差结构如图 4 所示。

3 实验过程及结果

3.1 数据集及实验环境

本文实验所用数据采集自安徽省淮南市谢家集区朱

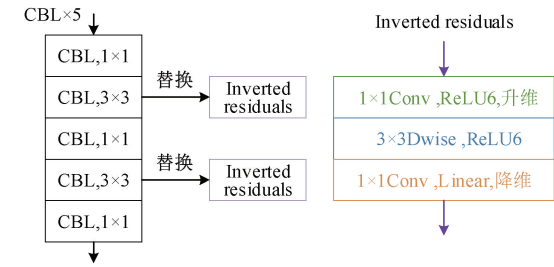


图 4 逆残差结构
Fig. 4 Inverted residuals structure

集东矿和河南省平顶山市大庄矿,采集了煤矿下 ZDY3200 液压钻机 和 ZYWL-4000SY 自动钻机的实时工作视频。视频包括装钻杆、打钻杆、卸钻杆、停止工作 4 种不同工作状态的钻机工作视频,从视频中抽取 2 722 张样本用于实验,图像分辨率为 1 920×1 080。使用 LabelImg 标注工具对数据集图像进行标注,并按照 4 : 1 的比例将数据集划分为训练集、验证集。针对煤矿下复杂场景,抽取常规场景、光照不均、存在部分遮挡以及运动模糊图像各 200 张制作测试集。数据标注示例如图 5 所示。

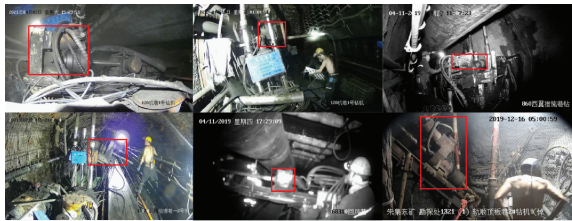


图 5 数据标注示例
Fig. 5 Data annotation example

本次实验的运行环境为 Windows10 (64 位) 系统的服务器,搭载 Intel core i7-10870H 和 NVIDIA GeForce RTX2060 显卡,项目代码基于 Tensorflow 深度学习框架,使用 Python3.7 开发。

3.2 网络训练参数

模型训练过程中,使用的 9 个锚框大小分别为 12×16、19×36、40×28、36×75、76×55、72×146、142×110、192×243 和 459×401 pixels。SGD 优化器动量因子设为 0.937,避免模型陷入局部最优,衰减系数为 0.000 5,防止网络训练过程中出现过拟合。设置 Batchsize 为 2,训练 100 个 epoch,目标类别数为 1,CIoU 阈值为 0.3,置信度阈值为 0.5。初始学习率为 0.01,最小学习率为 0.000 1,采用余弦退火算法更新学习率。

模型训练过程中采用 Mosaic 数据增强,训练时随机抽取 4 张图像,拼接而成,丰富图片背景,增加样本多样性。由于 Mosaic 数据增强后的样本脱离实际应用背景,

本次实验在后 30 个 epoch 关闭 Mosaic 数据增强,让训练样本符合实际应用场景,并保留最优模型。

3.3 特征提取网络对比实验

为验证本文采用 GhostNet 作为 YOLOv4 特征提取网络的优势,将轻量级网络 ResNet50^[23]、MobileNet2、MobileNetv3^[24]、ShuffleNetv2^[25]作为 YOLOv4 特征提取网络进行对比实验。为了验证模型的性能,采用检测精度 (mAP)、模型体积、模型检测速度 (FPS) 进行评价,实验结果如表 1 所示。

表 1 特征提取网络对比实验
Table 1 Feature extraction network comparison experiment

特征网络	mAP/%	模型体积/MB	FPS
ResNet50	95.48	235	32.00
CSPDarkNet	98.61	244	27.06
MobileNet2	97.88	148	51.96
MobileNetv3	98.08	151	54.16
ShuffleNetv2	63.71	39.3	29.07
GhostNet	98.72	150	55.00

由表 1 可知,采用轻量级网络作为 YOLOv4 的特征提取网络时,算法的精度都会有所下降,但是模型体积都更加轻量化,检测速度有不同程度的提高。其中,采用 GhostNet 作为 YOLOv4 的特征提取网络时,模型体积仅有 150 MB,检测精度与检测速度均优于其他轻量级网络。特征提取网络的特征响应输出图如图 6 所示。

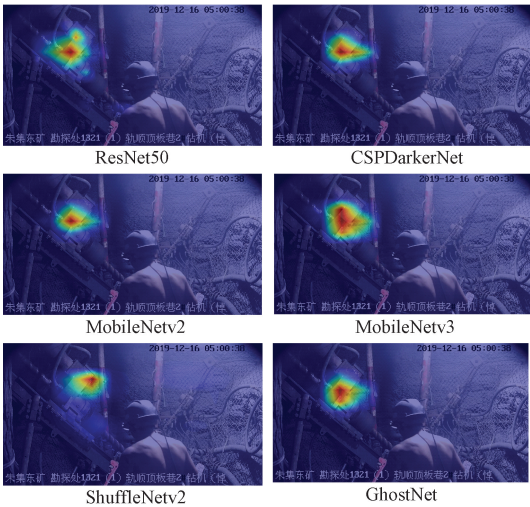


图 6 特征输出响应图
Fig. 6 Characteristic output response diagram

从图 6 中可以看出,ResNet50、Shufflenetv2、MobileNetv3 获得的热力图较为分散,CSPDarkNet、MobileNetv2 获得的热力图的感受野较小,GhostNet 热力图集中于钻机主体,感受野大于其他轻量级网络。综合

考虑,GhostNet 能够有效去除煤矿下背景的冗余特征,在模型轻量化的同时保持较强的特征提取能力。

3.4 消融实验

为了充分验证采用 GhostNet 替换特征提取网络、添加 CBAM 模块、在 PANet 中引入逆残差结构各项改进点对模型性能的提升效果,在测试集上进行消融实验。训练结果如图 7 所示。

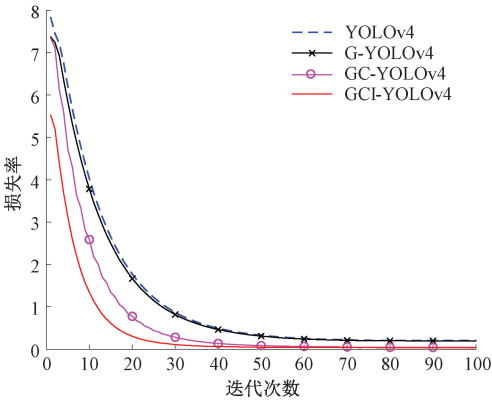


图 7 训练 Loss 曲线
Fig. 7 Training Loss curve

从图 7 中可以看出,随着训练轮数的增加,模型的损失函数不断下降,最终逐渐稳定,整个网络趋于收敛。YOLOv4 损失率在迭代 50 轮后逐渐稳定在 0.2 左右,采用 GhosNet 作为特征提取网络的 G-YOLOv4 损失曲线变化幅度较小,也在 50 轮后逐渐趋于稳定。在引入 CBAM 模块后,GC-YOLOv4 损失率降低,在 35 轮逐渐稳定在 0.041 左右。在加入逆残差结构后,GCI-YOLOv4 模型收敛速度加快,在 25 轮后损失率逐渐稳定在 0.038 左右。消融实验具体数值如表 2 所示。

表 2 消融实验结果
Table 2 Ablation results

模型名称	GhostNet	CBAM	逆残差 结构	mAP/%	模型体积/ MB	FPS
YOLOv4	—	—	—	98.61	244	27.06
G-YOLOv4	✓	—	—	98.72	150	55.00
GC-YOLOv4	✓	✓	—	99.43	150	52.50
GCI-YOLOv4	✓	✓	✓	99.49	114	58.10

从表 2 中可以清晰的看出,G-YOLOv4 的模型体积大大减少,从 244 MB 降到 150 MB,检测速度提升 27.94 FPS,没有损失模型精度,证明 GhostNet 能够减少深度网络模型的体积。加入 CBAM 后的 GC-YOLOv4 模型体积无明显变化,检测速度下降 2.5 FPS, mAP 提升 0.71%,证明 CBAM 模块能增强目标在复杂背景的显著度,从而提升模型精度,同时不会造成参数量和检测速度

的明显变化。加入逆残差结构后的 GCI-YOLOv4 体积进一步下降到 114 MB,检测速度也进一步加快,再次提升 5.6 FPS,同时检测精度可达 99.49%,表明逆残差结构能提取更丰富的语义特征,同时保持速度与精度的均衡。实验证明各改进策略均有一定效果,GCI-YOLOv4 综合了各模块的优点,能够有效提升特征图的利用效率,检测精度达到 99.49%,检测速度达到 58.1 FPS,具有较好的实时性。

3.5 对比实验及结果分析

为验证 GCI-YOLOv4 的可行性和有效性,在相同实验条件下,采用 RetinaNet^[26]、CenterNet^[27]、YOLOv4、YOLOv5、YOLOX^[28] 5 种当前主流的目标检测模型对钻机数据集进行训练与测试,实验结果如表 3 所示。

表 3 对比实验结果
Table 3 Comparison of experimental results

模型名称	mAP/%	模型体积/MB	FPS
RetinaNet	91.02	138	20.66
CenterNet	97.28	125	34.72
YOLOv4	98.61	244	27.06
YOLOv5	98.44	178	18.44
YOLOX	98.56	207	19.09
GCI-YOLOv4	99.49	114	58.10

从表中分析出 GCI-YOLOv4 的检测精度达到 99.49%,比 RetinaNet、CenterNet、YOLOv4、YOLOv5、YOLOX 分别高出 8.47%、2.21%、0.88%、1.09%、0.93%。GCI-YOLOv4 的模型大小只有 114 MB,低于其他 5 种主流目标检测模型。GCI-YOLOv4 检测速度分别高于其他 5 种模型 (RetinaNet、CenterNet、YOLOv4、YOLOv5、YOLOX) 37.44、23.38、31.04、39.66、39.01 FPS。GCI-YOLO 在精度与其他模型接近的情况下,展现出速度与体积方面的优势。

为了验证本文算法的优点,对测试集中常规场景、光照不均、存在部分遮挡以及运动模糊 4 种典型图像进行测试验证。检测结果如图 8 所示。

从第 1 组实验可以看出,常规场景下煤矿下环境复杂,待检测目标占据视频画面较少,存在大量冗余特征,各算法都能检测到待检测目标。但是 CenterNet、YOLOv5 检测目标框尺寸过大,目标框框选多余背景,RetinaNet 检测目标框尺寸过小,目标框丢失钻机主体,YOLOv4、YOLOX 检测精度。而 GCI-YOLOv4 采用 GhostNet 网络作为特征提取网络,去除冗余特征,使得目标检测框更精确。

从第 2 组实验可以看出,受光照影响,待检测的钻机目标区域亮度较高。从实验结果可以看出部分算法存在漏检现象。GCI-YOLOv4 由于引入 CBAM 模块增强钻机

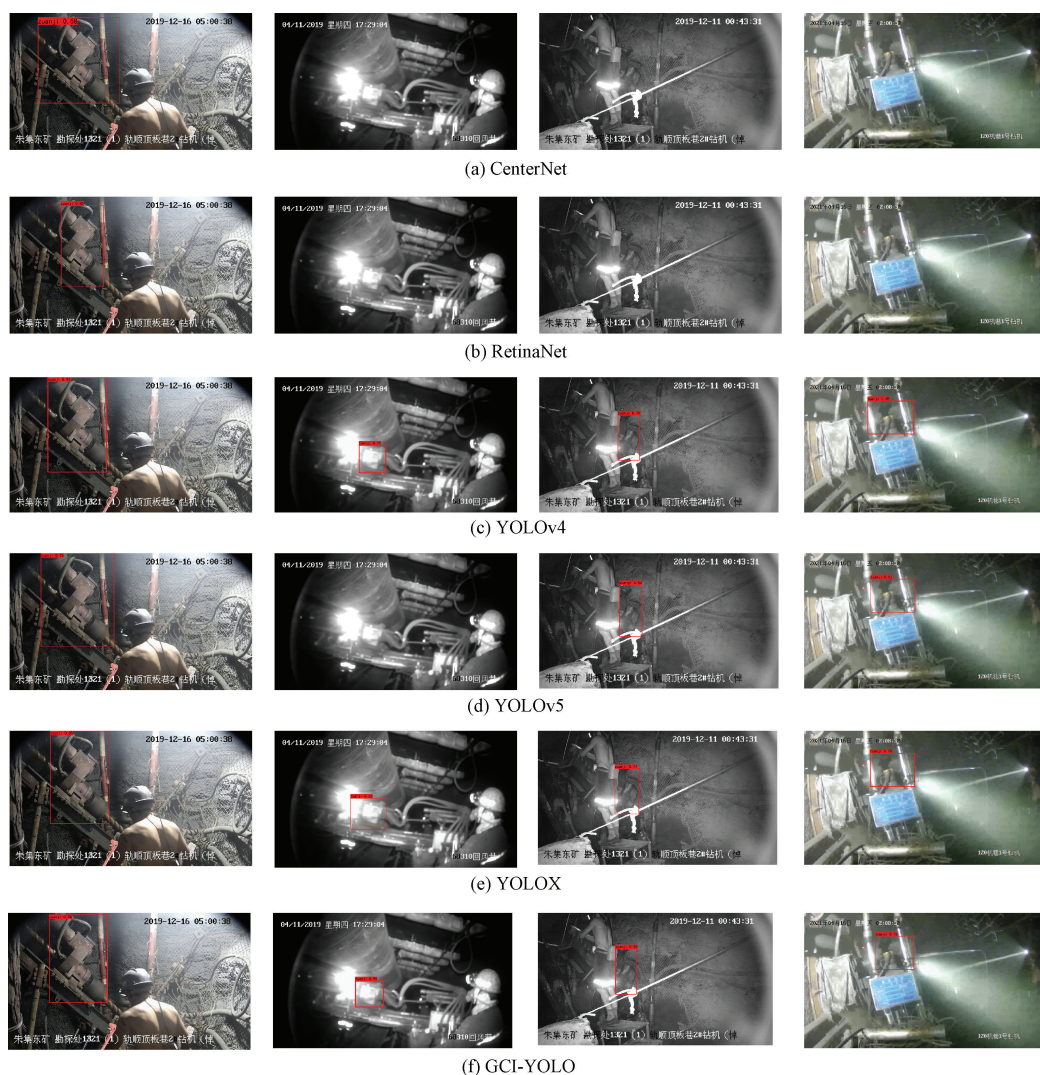


图 8 检测结果对比

Fig. 8 Comparison diagram of test results

在复杂背景中的显著度,提高钻机特征表达能力,使检测网络能够准确检测目标,提高检测精度。

从第 3 组实验可以看出,当钻机运动时存在部分遮挡严重影响钻进目标检测效果,CenterNet、RetinaNet 存在漏检现象,而 YOLOv4、YOLOv5、YOLOX 检测精度低于 GCI-YOLOv4,证明改进后的 GCI-YOLOv4 在部分遮挡时仍然能够很好进行检测。

从第 4 组实验可以看出,由于钻机高速运动造成获取图像较为模糊,CenterNet、RetinaNet 存在漏检现象,YOLOv4、YOLOv5、YOLOX 检测精度较低,GCI-YOLOv4 由于引入逆残差结构,通过其中的 1×1 卷积升维浓缩特征,深度可分离卷积精简原检测网络模块,较好地解决漏检现象。

3.6 钻孔深度测量

为实现钻孔深度测量功能,通过 GCI-YOLOv4 实现

钻机目标检测,获取钻机运动轨迹,计算钻机打入钻杆数量,得到最终的钻孔深度。

通过 GCI-YOLOv4 对钻机工作视频进行实时检测,获取钻机运动轨迹,将钻机的位置坐标信息按时间顺序存储至 CSV 文件。对钻机位置坐标信息进行均值滤波和归一化,再过滤其中因钻机运动导致检测框震荡的坐标信息,统计波形中的上升沿,统计钻杆数量。煤矿下使用的钻杆长度为 2 M,则钻机往复运动两次才能打入一根钻杆,每两个上升沿记为打入一根钻杆,根据钻杆数量精确计算出钻孔深度。钻机位置坐标信息滤波过程如图 9 所示,纵坐标为 GCI-YOLO 获取的钻机坐标信息,横坐标为钻机实时工作视频帧数。

为验证本文提出的钻杆计数方法在矿井下钻机工作面中的实际应用效果,选择安徽省淮南市谢家集区朱集东矿钻机工作面、河南省平顶山市大庄矿钻机工作面进

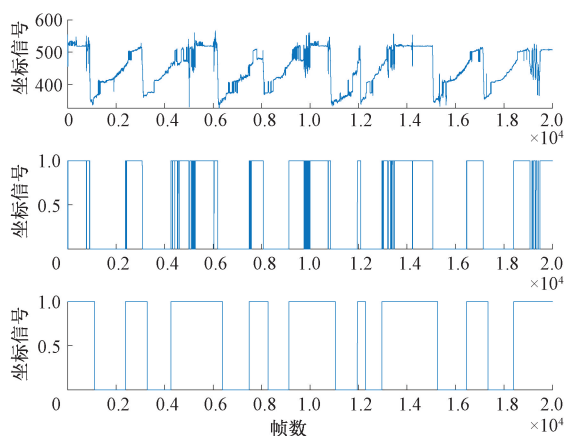


图9 坐标信息滤波过程

Fig. 9 Coordinate information filtering process

行现场测试,测试过程中累计装钻杆 300 根,卸钻杆 200 根,最终计得装钻杆 298 根,卸钻杆 199 根,钻杆计数精度达到 99.4%。

4 结 论

针对钻孔深度测量精度不高,传统图像处理算法易丢失钻机目标等问题,提出了 GCI-YOLOv4 检测钻机运动轨迹计算钻孔深度的方法,主要结论如下:

1) 通过引入 GhostNet 可以去除煤矿下的冗余特征,同时轻量化模型,加快模型推理速度。在 YOLOv4 中引入 CBAM 模块可以增强钻机在复杂背景中的显著度,从而提高钻机的特征表达能力,进而提高检测的精度。通过引入逆残差结构提取更丰富的语义特征,保持了检测模型速度和精度的平衡。

2) 通过实验证明 GCI-YOLOv4 检测精度达到 99.49%,检测速度为 58.10FPS,优于目前主流的目标检测算法。将 GCI-YOLOv4 模型应用于实际钻机工作面的动态视频,对运动轨迹进行滤波获取钻杆数量精度达到 99.4%,能准确测量出钻孔深度,验证了该方法的实用性和可行性。

3) 受煤矿下环境限制,在 GCI-YOLOv4 对钻机视频进行目标检测时,出现额外光源和部分遮挡时,仍然可以准确的进行钻机目标检测任务。当矿井下工作人员探照灯直视摄像头或工作人员长时间遮挡摄像头时,GCI-YOLOv4 无法很好的实现钻机目标检测功能。

参考文献

- [1] 郑仰峰,翟成,辛海会,等. 煤巷掘进工作面强弱耦合能量控制防治煤与瓦斯突出理论与方法[J]. 采矿与安全工程学报,2021,38(6):1269-1280.
ZHENG Y F, ZHAI CH, XIN H H, et al. Theories and

methods of coal and gas outburst prevention by strong-weak coupling energy control in coal roadway driving face [J]. Journal of Mining & Safety Engineering, 2021, 38 (6): 1269-1280.

- [2] 刘见中,孙海涛,雷毅,等. 煤矿区煤层气开发利用新技术现状及发展趋势[J]. 煤炭学报,2020,45(1): 258-267.
LIU J ZH, SUN H T, LEI Y, et al. Current situation and development trend of new technologies for the development and utilization of coalbed methane in coal mining areas [J]. Journal of China Coal Society, 2020, 45 (1): 258-267.
- [3] 李树刚,包若羽,张天军,等. 本煤层瓦斯抽采钻孔合理密封深度确定[J]. 西安科技大学学报,2019,39(2):183-188.
LI SH G, BAO R Y, ZHANG T J, et al. Determining the rational sealing depth for horizontal gas drainage borehole [J]. Journal of Xi'an University of Science and Technology, 2019, 39 (2): 183-188.
- [4] 董小明,褚志伟,张垒,等. 顺煤层随钻测量定向钻进技术在金泰煤矿的应用研究[J]. 煤炭技术,2022,41(1):92-96.
DONG X M, CHU ZH W, ZHANG L, et al. Application of directional drilling technology with measurement while drilling along coal seam in jintai coal mine [J]. Coal Technology, 2022, 41 (1): 92-96.
- [5] 路拓,刘盛东,王勃,等. 深孔钻孔深度的弹性波测量方法[J]. 地球物理学进展,2015,30(5):2176-2180.
LU T, LIU SH D, WANG B, et al. Measurement of deep drilling depth using elastic wave [J]. Progress in Geophysics, 2015, 30 (5): 2176-2180.
- [6] 王岚,陈龙,张鹏,等. 随钻测井钻孔深度测量装置的设计[J]. 煤田地质与勘探,2017,45(3):144-146.
WANG L, CHEN L, ZHANG P, et al. Design of LWD borehole depth-measuring device [J]. Coal Geology & Exploration, 2017, 45 (3): 144-146.
- [7] 王小龙,张军. 钻孔数据处理与三维显示软件设计及应用[J]. 煤矿安全,2021,52(10):130-135.
WANG X L, ZHANG J. Design and application of software for borehole data processing and three-dimensional display [J]. Safety in Coal Mines, 2021, 52 (10): 130-135.
- [8] 徐志强,吕子奇,王卫东,等. 煤矸智能分选的机器视觉识别方法与优化[J]. 煤炭学报,2020,45(6): 2207-2216.
XU ZH Q, LYU Z Q, WANG W D, et al. Machine vision recognition method and optimization for intelligent

- separation of coal gangue [J]. Journal of China Coal Society, 2020, 45 (6): 2207-2216.
- [9] 彭业勋, 符立梅. 基于 Sobel 边缘检测的圆周 Harris 角点检测算法[J]. 西安科技大学学报, 2019, 39(2): 374-380.
- PENG Y X, FU L M. Circle Harris corner detection algorithm based on Sobel edge detection [J]. Journal of Xi'an University of Science and Technology, 2019, 39 (2): 374-380.
- [10] 董立红, 王杰, 余向阳. 基于改进 Camshift 算法的钻杆计数方法[J]. 工矿自动化, 2015, 41(1): 71-76.
- DONG L H, WANG J, SHE X Y, et al. Drill pipe counting method based on improved Camshift algorithm [J]. Journal of Mine Automation, 2015, 41 (1): 71-76.
- [11] 方杰, 李振璧, 夏亮. 基于 ECO-HC 的钻杆计数方法[J]. 煤炭技术, 2021, 40(11): 186-189.
- FANG J, LI ZH B, XIA L. Drill pipe counting method based on ECO-HC [J]. Coal Technology, 2021, 40 (11): 186-189.
- [12] 高瑞, 郝乐, 刘宝, 等. 基于改进 ResNet 网络的井下钻杆计数方法[J]. 工矿自动化, 2020, 46(10): 32-37.
- GAO R, HAO L, LIU B, et al. Research on underground drill pipe counting method based on improved ResNet network [J]. Journal of Mine Automation, 2020, 46 (10): 32-37.
- [13] 党伟超, 姚远, 白尚旺, 等. 煤矿探水卸杆动作识别研究[J]. 工矿自动化, 2020, 46(7): 107-112.
- DANG W CH, YAO Y, BAI SH W, et al. Research on unloading drill-rod action identification in coal mine water exploration [J]. Journal of Mine Automation, 2020, 46(7): 107-112.
- [14] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [15] BOCHKOVSKIY A, WANG C Y, LIAO H. YOLOv4: Optimal speed and accuracy of object detection [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2020.
- [16] 张明路, 郭策, 吕晓玲, 等. 改进的轻量化 YOLOv4 用于电子元器件检测[J]. 电子测量与仪器学报, 2021, 35(10): 17-23.
- ZHANG M L, GUO C, LYU X L, et al. Improved lightweight YOLOv4 for electronic components detection [J]. Journal of Electronic Measurement and Instrumentation, 2021, 35 (10): 17-23.
- [17] 刘艳菊, 王秋霖, 赵开峰, 等. 基于卷积神经网络的热轧钢条表面实时缺陷检测[J]. 仪器仪表学报, 2021, 42(12): 211-219.
- LIU Y J, WANG Q J, ZHAN K F, et al. Real-time defect detection of hot rolling steel bar based on convolution neural network [J]. Chinese Journal of Scientific Instrument, 2021, 42 (12): 211-219.
- [18] 彭继慎, 孙礼鑫, 王凯, 等. 基于模型压缩的 ED-YOLO 电力巡检无人机避障目标检测算法[J]. 仪器仪表学报, 2021, 42(10): 161-170.
- PENG J SH, SUN L X, WANG K, et al. ED-YOLO power inspection UAV obstacle avoidance target detection algorithm based on model compression [J]. Chinese Journal of Scientific Instrument, 2021, 42 (10): 161-170.
- [19] HAN K, WANG Y, TIAN Q, et al. GhostNet: More features from cheap operations[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 1580-1589.
- [20] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [C]. European Conference on Computer Vision, 2018: 3-19.
- [21] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: Inverted residuals and linear bottlenecks[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018: 4510-4520.
- [22] HOWARD A, ZHU M, CHEN B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications [J]. arXiv preprint arXiv: 1704.04861, 2017.
- [23] HE K, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [24] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3 [C]. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), IEEE Press, 2019: 1314-1324.
- [25] MA N, ZHANG X, ZHEN H, et al. ShuffleNetV2: Practical guidelines for efficient CNN architecture design [J]. arXiv preprint arXiv: 1807.11164, 2018.
- [26] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017(99): 2999-3007.
- [27] ZHOU X, WANG D, KRAHENBUHL P. Objects as points[J]. arXiv preprint arXiv: 1904.07850, 2019.
- [28] GE Z, LIU S, WANG F, et al. YOLOx: Exceeding YOLO series in 2021 [J]. arXiv preprint arXiv:

2107.08430, 2021.

作者简介



张栋, 现为安徽理工大学硕士研究生, 主要研究方向为计算机视觉图像处理。

E-mail: 1023857721@qq.com

Zhang Dong is now a M. Sc. candidate at Anhui University of Science and Technology. His main research interests

include computer vision image processing and deep learning.



姜媛媛(通信作者), 2018 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, 主要研究方向为人工智能机器学习。

E-mail: jyyl672@163.com

Jiang Yuanyuan (Corresponding author)

received her Ph. D. from Nanjing University of Aeronautics and Astronautics in 2018. Now she is a professor at Anhui University of Science and Technology. Her main research interest includes artificial intelligence machine learning.