

DOI: 10.13382/j.jemi.B2306235

基于 MBDC 和双重注意力的变电站人员穿戴检测*

纪超^{1,2} 侯威^{1,2} 高鸣江^{1,2} 张凡^{1,2} 杨鹏³ 李小兵⁴

(1. 西安工程大学电子信息学院 西安 710048; 2. 西安市电气设备互联感知与智能诊断重点实验室 西安 710048;
3. 金属挤压与锻造装备技术国家重点实验室 西安 710018 4. 西安金源电气股份有限公司 西安 710000)

摘 要: 安全帽与工作服是变电站工作人员安全的重要保障,为解决现有检测模型对其检测精度低的问题,本文提出了 MBDC 和双重注意力的变电站人员穿戴检测算法。该算法提出了多分支深度卷积(multi branch deep convolution, MBDC)网络增加深度可分离卷积层以增强特征提取的完备性;然后提出多通道交互注意力(multimodal interaction attention, MIA)增加模型对小目标的检测能力,并将 MIA 机制结合高效通道注意力(efficient channel attention, ECA)机制构成双重注意力机制,增强模型对于小目标和遮挡目标的识别精度;最后引入焦点损失函数和 SIOU(scylla intersection over union)作为损失函数以解决正负样本不平衡问题并加快收敛速度。实验表明,本文算法全类平均精度达到 84.88%,比原算法高 9.92%,总体性能优于对比算法。
关键词: 变电站人员穿戴;多分支深度卷积;双重注意力机制;多通道交互注意力;高效通道注意力
中图分类号: TM93;TN0 **文献标识码:** A **国家标准学科分类代码:** 470.40

Wear detection of substation staff based on MBDC and dual attention

Ji Chao^{1,2} Hou Wei^{1,2} Gao Mingjiang^{1,2} Zhang Fan^{1,2} Yang Peng³ Li Xiaobing⁴

(1. School of Electronic Information, Xi'an Polytechnic University, Xi'an 710048, China; 2. Xi'an Key Laboratory of Interconnection Perception and Intelligent Diagnosis of Electrical Equipment, Xi'an 710048, China; 3. State Key Laboratory of Metal Extrusion and Forging Equipment Technology, Xi'an 710018, China;
4. Xi'an Jinyuan Electric Co., Ltd., Xi'an 710000, China)

Abstract: Helmets and work clothes are important guarantees for the safety of substation staff. In order to solve the problem of low detection accuracy of existing detection models, this paper proposes a substation staff wear detection algorithm based on multi-branch deep convolution and dual attention. The algorithm proposes a multi-branch deep convolution (MBDC) network to add a deep separable convolution layer to enhance the completeness of feature extraction. Then, it is proposed that multimodal interaction attention (MIA) can increase the detection ability of the model to small targets, and combine the MIA mechanism with efficient channel attention (ECA) mechanism to form a dual attention mechanism to enhance the recognition accuracy of the model for small targets and occluded targets. Finally, the focus loss function and SIOU (scylla intersection over union) are introduced as the loss function to solve the problem of positive and negative sample imbalance and accelerate the convergence speed. The experiment shows that the average accuracy of the algorithm in this paper is 84.88%, 9.92% higher than the original algorithm, and the overall performance is better than the comparison algorithm.
Keywords: wearing of substation personnel; multi branch depth convolution; dual attention mechanism; multimodal interaction attention; efficient channel attention

0 引言

安全帽和工作服是变电站工作人员安全的有效保障,传统的人员穿戴检测是监察人员在监控前通过人工观察的方式检测,这种方式误判率高,效率低。随着图像监控系统的普及,基于图像处理的规范穿戴识别技术得到快速发展^[1-5]。该技术一方面能规范站内检修和建造人员着装,并在事故发生时快速定位被困人员位置;另一方面,通过检测是否规范穿戴,辅助监测未经授权人员非法进入。因此,变电站人员穿戴的智能检测方法对于工作人员的安全和变电站安全运行具有重要的意义。

常见的安全着装检测算法可以分为基于运动检测的背景建模算法、基于人工特征和传统分类器的机器学习算法以及基于卷积神经网络的深度学习算法^[6-8]。文献[9]采用蒙特卡洛法随机采样训练样本邻域中的点以扩充样本的数量并降低分类器输出的波动性,并采用高斯牛顿法迭代求解隐含层到输出层的权值,建立以高斯函数为基函数的径向基神经网络分类器^[9]。文献[10]提出了一种基于 Faster R-CNN 的安全帽检测方法并根据视觉条件对图像帧进行分类,将图像帧输入到 Faster R-CNN 模型中进行训练,得到了比传统安全帽检测算法更高效、能有效检测不同场景的安全帽检测算法。文献[11]提出单人关键点检测模型和区域定位策略,降低着装特征提取难度并引入注意力机制,进一步提升准确率。文献[12]提出了一种分层正样本选择 (HPSS) 机制,提高了 YOLOv5 的拟合能力,同时提出了一种基于盒

子密度的后处理算法,有效地抑制了虚假检测的出现。文献[13]通过人脸特征识别和头部识别的交叉验证定位构造者的头部区域,解决了复杂姿态下构造者头部位置的确定问题,并采用 I-YOLOv3 目标检测网络对安全帽佩戴情况进行检测。

上述算法中的传统机器学习算法在变电站面对小目标和复杂场景下的着装检测识别率较低且易过拟合,而 Faster R-CNN 等双阶段算法虽然检测效果相对于传统算法识别率高,但实时性较差且对硬件要求高^[14-15]。上述的单阶段算法的出现解决了部分的问题,但在检测穿戴安全帽这类小目标以及遮挡目标时检测精度依然不佳,因此本文提出了基于多分支深度卷积和双重注意力的变电站人员穿戴检测算法。本文首先提出了 MBDC 网络增加卷积层来增加整体检测效果,然后提出双重融合注意力机制增加网络对小目标的检测能力,最后通过改进损失函数平衡正负样本,最终实现了变电站场景下安全帽和工作服高效的智能识别。

1 算法设计

1.1 本文算法模型

变电站的安全帽和工作服的检测具有以下几个难题:目标小,在远距离拍摄时尤其明显;人员之间或电气设备与人员之间遮挡严重;现场设备和建筑过多导致检测背景复杂;多目标检测效果差^[16]。针对上述难点,在 YOLOv7 的基础上提出了本文模型^[17],本文算法总体结构图如图 1 所示。

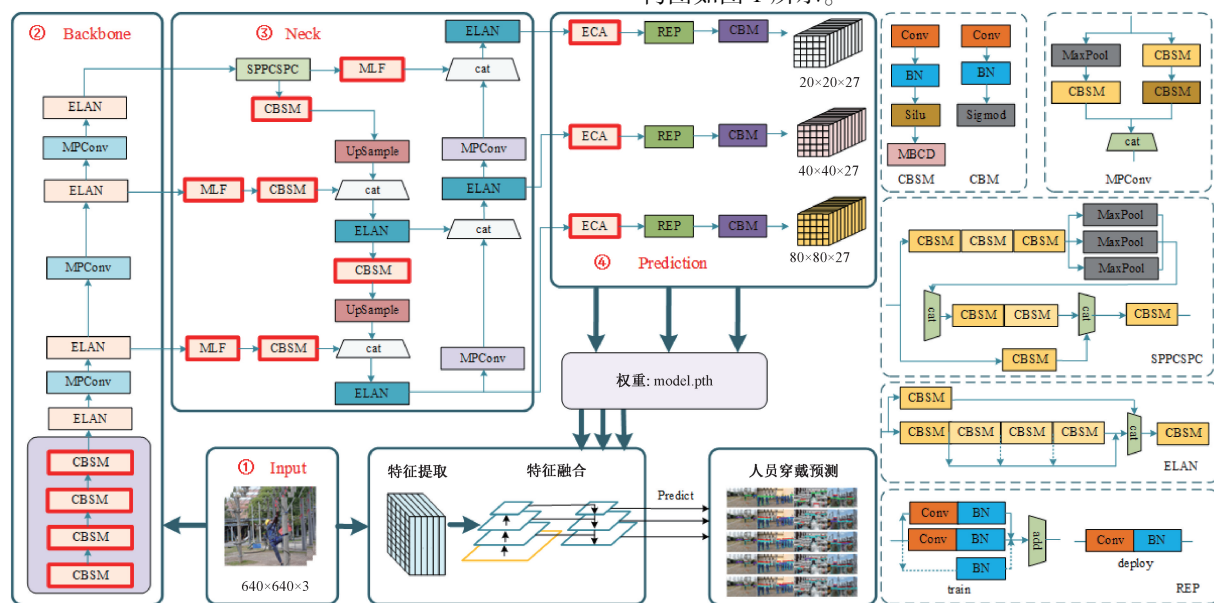


图 1 本文算法总体结构

Fig. 1 Overall structure of the algorithm in this paper

由图 1 可知,本文算法主要有输入层(Input)、骨干层(Backbone)、颈部(Neck)和预测层(Prediction)4部分组成,其中粗线框为本文改进模块。本文算法提出 MBDC 网络并插入到每次卷积之后,通过堆叠深度可分离卷积层来提取更加复杂、抽象的信息,并通过使用多分支结构提取图像不同尺度的信息,得到更好的图像表征能力;然后,提出双重注意力机制将更多的注意力集中到人的头部和身体区域;最后,使用焦点损失(focal loss)函数和 SIOU 优化损失函数减小正负样本不平衡带来的负

面影响,同时加快模型训练收敛速度。

1.2 多分支深度卷积网络的设计

为提高网络的图像表征和特征提取能力并减小因加入注意力机制对特征提取的影响,本文提出了多分支深度卷积网络。该网络首先使用深度可卷积网络作为基本卷积模块增加卷积层数,再利用多分支构建网络结构,最终实现在增加有限参数的情况下大幅增加网络的特征提取和图像表征能力,其总体结构图如图 2 所示。

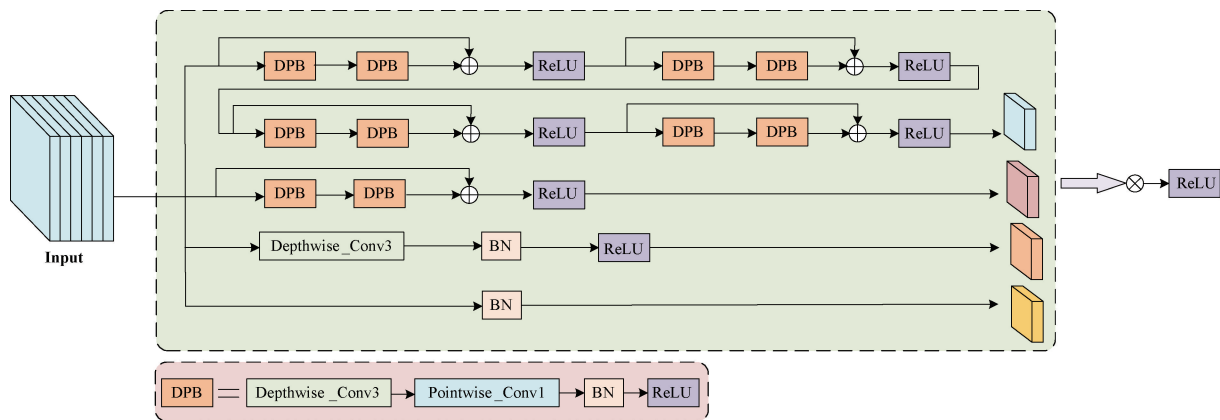


图 2 多分支深度卷积网络结构

Fig. 2 MBDC network structure

由图 2 可知,MBDC 网络结构具有 4 个分支,不同分支增加了梯度流动,减少了因网络深度增加导致梯度消失的问题。同时,残差结构没有涉及跨模块之间的连接关系,相对于跨模块的跳联结构在训练时更加节省资源。DPB 模块中的深度卷积(depthwise convolution)、逐点卷积(pointwise convolution)两种卷积共同组成了深度可分离卷积,深度可分离卷积结构如图 3 所示^[18]。

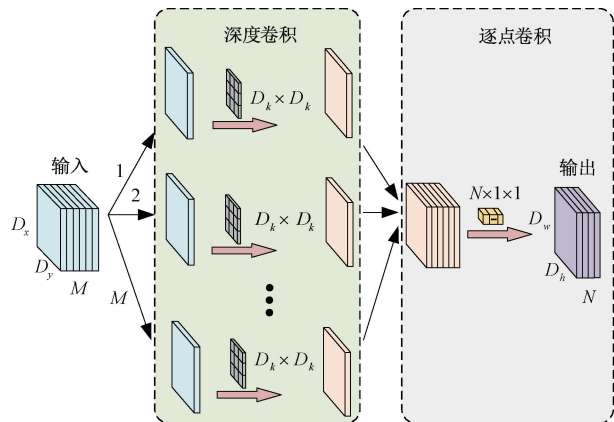


图 3 深度可分离卷积网络结构

Fig. 3 Deep separable convolution network structure diagram

图 3 中: M 为输入图像通道数, N 为输出图像通道

数; D_x 为输入图像长度, D_y 为数据输入图像宽度; D_k 为卷积核大小; D_w 为输出图像长度, D_h 为输出图像宽度。网络中原有卷积主要是对通道特征图进行卷积操作,其计算量如式(1)所示。

$$Q_1 = D_k^2 \cdot M \cdot N \cdot D_w \cdot D_h \quad (1)$$

深度可分离卷积计算量 Q_2 如式(2)所示。

$$Q_2 = D_k^2 \cdot M \cdot D_w \cdot D_h + M \cdot N \cdot D_w \cdot D_h \quad (2)$$

通过引入深度卷积比直接引用标准卷积计算量和参数的数量大幅减,深度可分离卷积与标准卷积之间的计算量比值如式(3)所示。

$$\frac{Q_2}{Q_1} = \frac{D_k^2 \cdot M \cdot D_w \cdot D_h + M \cdot N \cdot D_w \cdot D_h}{D_k^2 \cdot M \cdot N \cdot D_w \cdot D_h} = \frac{1}{N} + \frac{1}{D_k^2} \quad (3)$$

1.3 双重融合注意力机制的设计

为了解决难以检测未佩戴安全帽目标的问题,本文算法提出了 MIA 和 ECA 组成的双重注意力机制。MIA 注意力机制部署于骨干层与头部间,ECA 注意力机制部署于头部特征融合与预测层之间,融合每个特征图的通道和空间权重值,使得模型训练时更加专注于安全帽和工作服的特征并提高最终识别的精度以及效率。

1) 多通道交互注意力机制

MIA 模块融合每个特征图的通道和空间权重值,使

得模型训练时更加专注于安全帽和工作服的特征并提高最终识别的精度以及效率。该模块通过建立不同卷积结构的密集连接,最大化了前后层的信息交互,实现了特征在通道维度上的复用,不但减缓了梯度消失的现象,也使其可以在参数与计算量更少的情况下实现更优的性能。其总体框架图如图 4 所示。

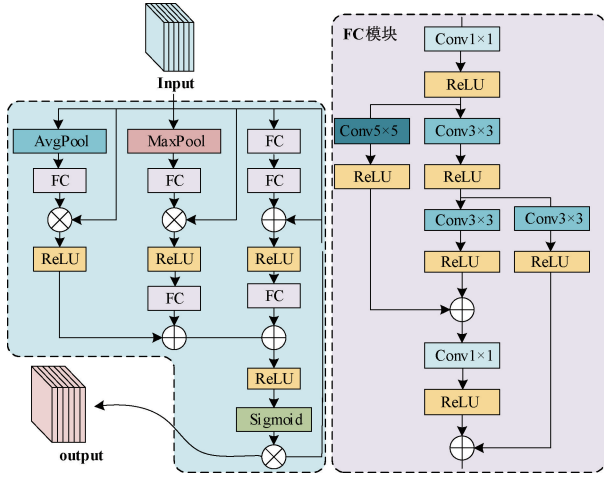


图 4 多通道交互注意力网络架构

Fig. 4 MIA network structure

由图 4 可知,特征图输入后分为 3 个分支,网络通过融合卷积(fusion convolution, FC)模块强化特征,在特征图的空间上累计注意力权重。同时将输出的特征与未经 FC 模块强化特征的权重相加融合,减少在深度卷积过程中特征信息的损失程度。输入的特征通过全局平均池化得到 1×1 的特征向量,整合输入特征的全局通道信息,增强网络的鲁棒性。为了避免单独采用一种池化结构使提取到的特征丧失部分信息故将输入的特征通过全局最大池化得到一维的特征向量并在整合通道信息后与输入特征逐元素相乘。

FC 模块是 MIA 的核心组成模块。将 F_{IN} 作为输入 MFC 模块的初始特征图像,利用 1×1 的卷积对初始特征图进行变换,提高获取图像特征的泛化能力,同时实现各个特征的跨通道的信息交互。在卷积后增加激活函数 ReLU,增加激活函数能将卷积后的线性特征转变为非线性特征,激活函数表达式如式(4)所示。

$$\phi(x) = \max(0, x) \quad (4)$$

2) ECA 注意力机制

本文算法使用双重注意力机制,为避免增加过多参数导致模型计算量增加,故本文选用 ECA 网络作为另一个注意力机制^[19]。通过添加 ECA 注意力模块,不同的权重分配给不同的卷积通道突出人员穿戴特征,通过适当的跨信道交互在显著降低模型复杂度的同时保持性能,避免降维对学习通道的影响。

ECA-Net 通过空间维度使用全局平均池化对输入特征图进行空间特征压缩得到 $1 \times 1 \times C$ 的矩阵,然后对压缩后的特征图通过大小为 k 的快速一维卷积进行通道特征学习,最后将通道注意力结合。特征通道 C 与卷积核大小 k 成正相关,计算如式(5)所示:

$$C = \varphi(k) = \gamma \times k - b \quad (5)$$

式中: C 为特征通道数; γ 和 b 表示一次函数 $\varphi(k)$ 的比例系数和截距; k 表示卷积核大小。为增加 C 与 k 关系范围,故将线性函数 $\varphi(k)$ 拓展到非线性函数且通道数 C 通常设置为 2 的倍数,故 C 最终表示为式(6)所示。

$$C = \phi(k) = 2^{(\gamma \times k - b)} \quad (6)$$

式中: γ 和 b 中设置为 2 和 1,用于改变通道数 C 和卷积核大小和之间的比例。在给定特征通道数 C 后,大小为 k 的快速一维卷积的自适应函数如式(7)所示。

$$k = \psi(C) = \left\lfloor \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}} \quad (7)$$

式中: $\lfloor \cdot \rfloor_{\text{odd}}$ 表示 k 只能取奇数。

1.4 损失函数的改进

为解决训练时正负样本不平衡的问题并加快模型收敛速度,在目标置信度损失和分类损失引入焦点损失(focal loss)函数代替原来的二元交叉熵损失(binary cross entropy loss, BCE loss)函数,坐标损失采用 SIOU 损失函数^[20-21]。

1) Focal loss

原算法 YOLOv7 使用的 BCE Loss 迭代缓慢且有一定概率偏离正确的优化方向,无法调整至最优,同时也无法解决训练过程中正负样本不平衡问题。而 Focal Loss 函数引入 δ 权重系数以及权重因子 $(1-x)^\beta$,通过改变正负样本的相对频率和降低简单样本的贡献权重来解决二元交叉熵损失函数中的类不平衡问题,其表达式如式(8)所示:

$$\text{FocalLoss} = \begin{cases} -\delta(1-x)^\beta \log(x), & (a=1) \\ -(1-\delta)x^\beta \log(1-x), & \text{其他} \end{cases} \quad (8)$$

式中:权重系数用于平衡正负样本; $(1-x)^\beta$ 权重因子用于平衡样本的难易程度; β 为调制系数,实验中 β 取 2 效果最佳。

2) SIOU

CIUO 等传统的目标检测损失函数依赖于边界框回归指标的聚合,没有考虑期望的真实框和预测框之间不匹配的方向,该缺陷可能会因为预测框在训练过程中无法快速定位导致收敛速度较慢且效率较低。SIOU 重新定义了惩罚度量并考虑到期望回归之间的向量夹角, SIOU 示意图如图 5 所示。

图 5 中: B 和 B^{GT} 分别代表预测框和真实框; C_w 和 C_h 分别为二者中心点的横向距离和纵向距离; C_x 和 C_y 分别为二者的外侧边框的横向和纵向距离; α 和 β 二者中心点

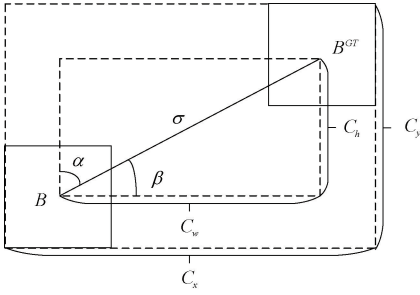


图5 SIOU 损失函数示意图

Fig.5 Schematic diagram of SIOU Loss function

连线与水平线和垂直线的夹角; σ 为二者中心点的距离。

该损失函数由角度成本、距离成本、形状成本和 IOU 成本 4 种成本函数组成, SIOU Loss 添加了真实框 B^{GT} 和预测框 B 中心间的角度感知, 减少距离相关的额外变量, 角度成本计算公式如下所示。

$$x = \frac{c_h}{\sigma} = \sin(\alpha) \quad (9)$$

$$d = \sqrt{(b_{c_x}^{gt} - b_{c_x})^2 + (b_{c_y}^{gt} - b_{c_y})^2} \quad (10)$$

$$c_h = \max(b_{c_y}^{gt}, b_{c_y}) - \min(b_{c_y}^{gt}, b_{c_y}) \quad (11)$$

$$\Lambda = 1 - 2 \times \sin^2(\arcsin(x) - \frac{\pi}{4}) \quad (12)$$

式中: Λ 是角度计算的最终结果; x 是图中真实框和预测框中心点的角度 α 的正弦值; d 是它们之间的距离; c_h 是两点间的相对高度差。

距离成本计算公式如下所示:

$$\rho_x = \left(\frac{b_{c_x}^{gt} - b_{c_x}}{c_w} \right)^2, \rho_y = \left(\frac{b_{c_y}^{gt} - b_{c_y}}{c_h} \right)^2, \gamma = 2 - \Lambda \quad (13)$$

$$\Delta = \sum_{t=x,y} (1 - e^{-\gamma \rho_t}) \quad (14)$$

式中: ρ_x 和 ρ_y 是真实框和预测框的中心点在 x 轴和 y 轴上的相对距离与它们的最小外接矩形的宽高的比值的平方, e 为欧拉常数。

形状成本的计算公式如下所示:

$$\omega_w = \frac{|w - w^{gt}|}{\max(w, w^{gt})}, \omega_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})} \quad (15)$$

$$\Omega = \sum_{t=w,h} (1 - e^{-\omega_t})^\theta \quad (16)$$

式中: (w, w^{gt}) 和 (h, h^{gt}) 分别是预测框和真实框的宽高; θ 为形状计算公式中的注意力系数, 不同的数据集它的值也不相同, 本文取 2。最后是 IOU 成本的计算, 也是 SIOU 的最终表达式, 其公式如下式。

$$IOU = \frac{|B \cap B^{GT}|}{|B \cup B^{GT}|} \quad (17)$$

$$SIOU = 1 - IOU + \frac{\Delta + \Omega}{2} \quad (18)$$

2 实验结果与分析

2.1 实验设置

1) 实验环境与性能评价

本实验在 Pytorch 框架下进行, 实验中 batch size 设置为 16, 动量大小 0.9, 权重衰减设置 0.0005, 学习率设置为 0.001。本实验使用显卡型号为 NVIDIA RTX3090, 处理器为 Intel i9-11900k, 操作系统为 ubuntu 18.04。

为客观评价本文算法的优势, 选取全类平均精度 (mean average precision, mAP)、召回率 (recall, R)、准确率 (precision, P) 和每秒处理帧数 (frames per second, FPS) 作为评价指标。

2) 数据集制作

本文所使用的变电站数据集有两部分组成, 一部分自行拍摄于西安某大型变电站, 拍摄场景包括常见的复杂背景、多目标、小目标等共计 7 000 张图像; 另一部分来自于网络的变电站人员穿戴图像, 用于增强数据集的广泛性, 共计 4 500 张图像, 所有数据集加起来共 11 500 张。数据集使用 Labellmg 软件对数据集中的 4 种目标进行标注, 将数据集按 8 : 1 : 1 随机划分为训练集、验证集和测试集, 最终训练集包含 9 200 张图像, 验证集以及测试集中分别含有 1 150 张图像, 部分数据如图 6 所示。



图6 部分数据集展示

Fig.6 Partial dataset display

2.2 消融实验

为了验证本文所采取的两种不同的机制的改进效果, 将所改进的机制依次嵌入至原始算法中, 使用相同测试集对各算法进行测试的实验结果如表 1 所示。

由表 1 可知, 4 个不同模块分别加入到原算法后效果均有提升, MBDC 是在原有的卷积模块上通过在增加轻量级的多分支深度卷积大幅提高网络的特征提取和图

表 1 消融实验结果对比

Table 1 Comparison of ablation test results

MBDC	MIA	ECA	LOSS	mAP@ 0. 5/%	FPS/(t/s)
✓	✓	✓	✓	74. 96	20. 51
				77. 99	16. 17
				79. 86	14. 63
				78. 75	20. 75
✓	✓	✓	✓	76. 18	20. 48
				81. 56	14. 58
				84. 88	11. 43

像表征能力,在原算法的基础上提升明显。同时,将两种注意力机制融合之后的算法在原来的算法基础上增加了 6. 6%,优于单独加入 MIA 或 ECA 注意力机制。本文算法相对于原始算法提升明显,相较于改进前 mAP 值增加了 9. 92%。各部分的 FPS 值中,ECA 的加入与损失函数的改进对 FPS 几乎没有影响,其他各部分则通过牺牲部分检测速度提升了检测精度。

改进算法前后对小目标、多目标、复杂背景和人员遮挡 4 种场景下预测结果如图 7 所示,戴安全帽、未戴安全帽、穿工作服和未穿工作服 4 种目标分别用标签 helmet、no_helmet、workclothes 以及 no_work-clothes 表示。

图 7 中 4 幅图分别为小目标、多目标、复杂背景和人员遮挡图片。由图可知,改进后算法性能在各场景下明显优于改进前算法。其中,改进后的算法对于各目标中未戴安全帽目标的优势最为明显。除此之外,改进后的算法漏检率与错检率也明显降低。

2.3 算法对比

为了更好的比较本文算法相较于其他算法的改进之处,实验将本文的训练集分别用 6 种主流目标检测算法以及本文进行训练,并使用相同的测试集进行验证,其结果如表 2 所示。

表 2 不同算法性能对比

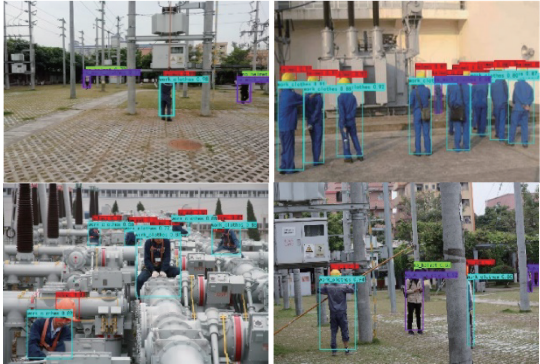
Table 2 Performance comparison of different algorithms

算法	mAP@ 0. 5/%	FPS/(f·s ⁻¹)
Faster R-CNN	68. 42	7. 80
EfficientDet	70. 55	15. 74
SSD	60. 88	14. 54
YOLOv3	73. 51	22. 67
YOLOv5	72. 38	26. 23
YOLOv7	74. 96	20. 51
本文	84. 88	11. 44

由表 2 可知,上述算法中,本文算法通过增加 MBDC 网络、双重融合注意力机制大幅增加了检测检测效果,比其他算法高出 9% 以上,提升显著。而本文算法由于提高检测精度牺牲了部分检测速度,但速度明显高于 Faster



(a) 改进前模型检测结果
(a) Model test results before improvement



(b) 改进后模型检测结果
(b) Improved model test results

图 7 不同模型预测结果

Fig. 7 Forecast results of different models

R-CNN,且与 EfficientDet、SSD 接近。因此,总的来说本文算法虽检测速度有所下降,但检测精度相较于主流算法明显提升。不同模型训练时的 mAP 值在交并比为 0. 5 时的对比曲线如图 8 所示。

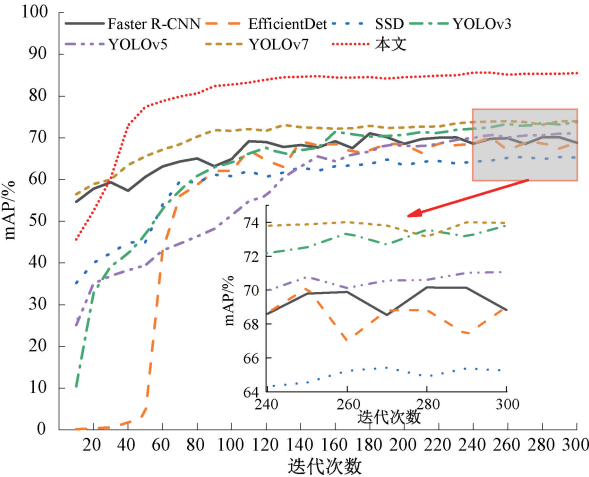


图 8 模型训练 mAP 变换曲线

Fig. 8 Model training mAP transformation curve

由图 8 可知,随着迭代次数的增加,各算法的 mAP 值随之增加,并最终均能在 250 轮之前上升至最大值。其中,本文算法 mAP 最高且在 140 轮已经稳定,由此可见经过改进过的损失函数,其收敛能力得到了明显提升。训练时不同算法的损失值变化曲线如图 9。

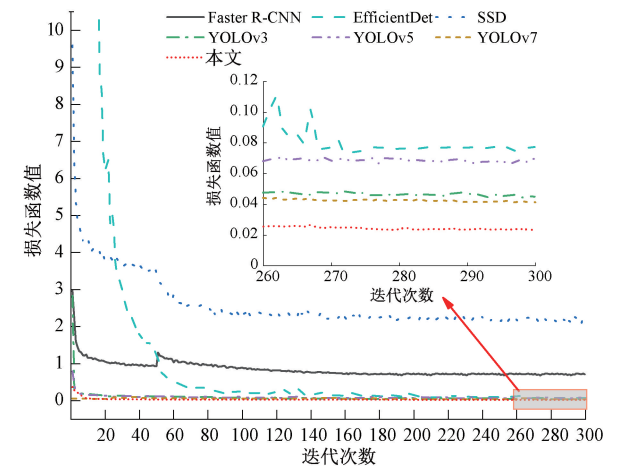
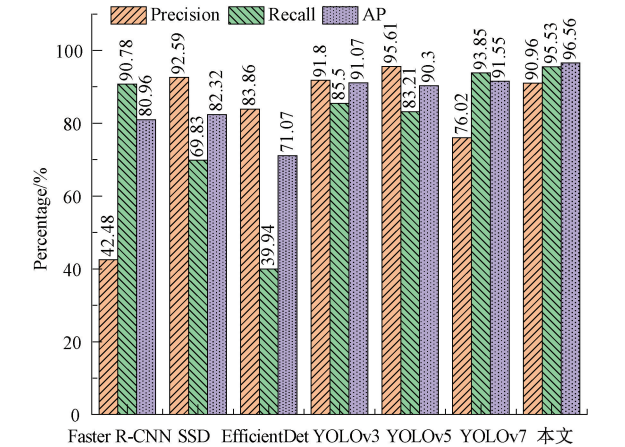


图 9 模型训练损失函数值变换曲线

Fig. 9 Loss function value transformation curve

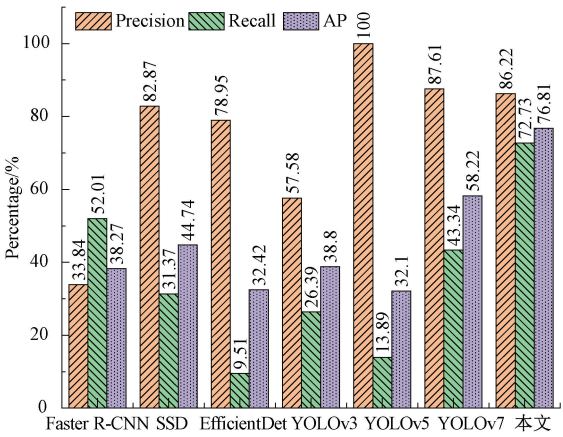
由图 9 可知,各模型均可收敛且损失值变化在 180 轮之前基本已经稳定。本文算法在开始时损失值较高,但在迭代 14 轮之后已经降至 0.052,在对比算法中效果最优。该图验证了 Focal Loss 与 SIOU 的引入相对于原算法可以实现更好的训练和检测效果。

本文的检测目标为安全帽、未戴安全帽、穿工作服和未穿工作服 4 种模型对于每种目标检测的效果相差较大,其中对于未戴安全帽和未穿工作服的检测难度最大,不同算法对于不同检测目标关于 P、R 与平均精确度 (average precision, AP) 对比如图 10 所示。



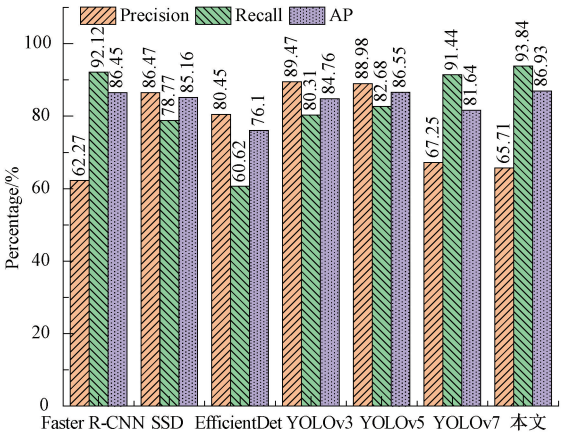
(a) 戴安全帽目标性能对比

(a) Comparison of target performance with helmet



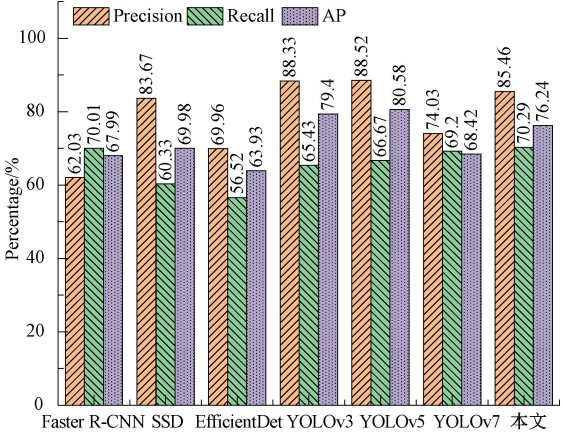
(b) 未戴安全帽目标性能对比

(b) Performance comparison of targets without helmet



(c) 穿工作服目标性能对比

(c) Comparison of target performance of wearing work clothes



(d) 未穿工作服目标性能对比

(d) Target performance comparison without working clothes

图 10 模型 P、R 和 AP 值对比

Fig. 10 Comparison of model P, R and AP values

由图 10 所示,本文算法相对于其他对比算法综合性能最优,且对戴安全帽、未戴安全帽、穿安全服 3 种目标的检测效果优势最为明显。其中对于检测难度最大的未戴安全帽目标,本文算法得益于双重注意力机制的引入,

更加注重小目标的检测,因此精确率、召回率和 AP 值高达 86.22%、77.73% 和 76.81%,AP 值相较于改进前的 YOLOv7 提升了 18.59%,其他算法中表现较好的是 SSD,但相对于本文算法也低 32.07%。对于戴安全帽目标的检测,本文算法 AP 值为 96.56%,比改进前的 YOLOv7 提升了 5.01%,同时高于其他算法。对于穿工作服目标的检测,未改进前的 YOLOv7 算法 AP 值低于 Faster R-CNN 等算法,但在改进之后高于其他对比算法。对于未穿工作服目标的检测,本文算法 AP 值略低于 YOLOv3 和 YOLOv5,但相对于改进前具有 7.82% 的提升,且高于其他 4 种对比算法。本文算法主体模型由 YOLOv7 改进而来,YOLOv7 相对于 YOLOv3 和 YOLOv5 引入高效聚合等网络,以弥补 YOLOv5 对于小目标检测效果差的缺点,但由于过于关注小目标的检测,故在本数据集上对于常规目标检测效果略低于 YOLOv3 和 YOLOv5。本文算法不仅扩大了 YOLOv7 对于小目标检测效果好的优势,还缩小了 YOLOv7 与 YOLOv5 对于未穿工作服目标检测的差距。

3 结 论

本文针对目前变电站人员穿戴安全检测模型检测性能较低的问题,提出了基于多分支深度卷积和双重注意力机制的变电站人员穿戴检测算法。该算法分别使用多分支深度卷积增加卷积层数提高整体检测精度;使用双重注意力机制增加对小目标的检测效率;使用 Focal Loss 和 SIOU 代替原来的 BCE Loss 和 CIOW 损失函数,加快模型收敛速度并解决正负样本不平衡的问题。最终,本文模型相较于其他主流算法具有明显优势,而本文模型对检测难度较高的未穿戴安全帽的小目标类别优势则更为明显,对于保障变电站工作人员安全具有重要作用。但本文模型也存在 FPS 偏低的问题,在之后的研究中会考虑平衡检测精度和检测速度,对模型进行轻量化的处理。

参考文献

- [1] TIAN X G, ZHANG G M. Research on surface defect detection technology of wind turbine blade based on UAV image[J]. Instrumentation, 2022, 9(1): 41-48.
- [2] 陈汐, 韩译锋, 闫云凤, 等. 目标物智能跟踪与分割融合算法及其在变电站视频监控中的应用[J]. 中国电机工程学报, 2020, 40(23): 7578-7587.
CHEN X, HAN Y F, YAN Y F, et al. A unified algorithm for object tracking and segmentation and its application on intelligent video surveillance for transformer substation[J]. Proceedings of the CSEE, 2020, 40(23): 7578-7587.
- [3] 程亮, 吴兴辉, 江云华, 等. 基于无人船视觉的水域人员类别识别算法[J]. 电子测量与仪器学报, 2022, 36(8): 43-51.
CHENG L, WU X H, JIANG Y F, et al. Person category identification algorithm in water environment based on unmanned ship vision[J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(8): 43-51.
- [4] ZHENG H, SUN Y, LIU X, et al. Infrared image detection of substation insulators using an improved fusion single shot multibox detector[J]. IEEE Transactions on Power Delivery, 2020, 36(6): 3351-3359.
- [5] 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.
- [6] ZHAO Z Q, ZHENG P, XU S, et al. Object detection with deep learning: A review[J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(11): 3212-3232.
- [7] 刘谋海, 胡军华, 丁国栋, 等. 基于深度学习的电力监控视频超分辨率重建[J]. 电力系统保护与控制, 2021, 49(11): 181-187.
LIU M H, HU J H, DING G D, et al. Super-resolution reconstruction of a power monitoring video based on deep learning[J]. Power System Protection and Control, 2021, 49(11): 181-187.
- [8] LI D, WANG R, CHEN P, et al. Visual feature learning on video object and human action detection: A systematic review[J]. Micromachines, 2021, 13(1): 72.
- [9] 李聪, 胡金磊, 周俊煌, 等. 基于 S-RBFNN 的变电作业人员着装图像识别算法[J]. 广东电力, 2019, 32(9): 124-130.
LI C, HU J L, ZHOU J H, et al. Dressing image recognition algorithm for substation operators based on sensitivity radial basis function neural network[J]. Guangdong Electric Power, 2019, 32(9): 124-130.
- [10] FANG Q, LI H, LUO X, et al. Detecting non-hardhat-use by a deep learning method from far-field surveillance videos[J]. Automation in Construction, 2018, 85: 1-9.
- [11] 何国立, 齐冬莲, 闫云凤. 一种基于关键点检测和注意力机制的违规着装识别算法及其应用[J]. 中国电机工程学报, 2022, 42(5): 1826-1837.
HE G L, QI D L, YAN Y F. An algorithm of irregular

- dress recognition based on key point detection and attention mechanism and its application[J]. Proceedings of the CSEE, 2022, 42(5):1826-1837.
- [12] LI Z, XIE W, ZHANG L, et al. Toward efficient safety helmet detection based on YOLOv5 with hierarchical positive sample selection and box density filtering[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-14.
- [13] PENG H, ZHANG Z. Helmet wearing recognition of construction workers using convolutional neural network[J]. Wireless Communications and Mobile Computing, 2022.
- [14] 陈健. 基于表现特征的人体着装分析与识别[D]. 北京:北京邮电大学, 2009.
- CHEN J. Analysis and recognition of human body dress based on performance characteristics [D]. Beijing: Beijing University of Posts and Telecommunications, 2009.
- [15] 周念成, 廖建权, 王强钢, 等. 深度学习在智能电网中的应用现状分析与展望[J]. 电力系统自动化, 2019, 43(4): 180-191.
- ZHOU N CH, LIAO J Q, WANG Q G, et al. Analysis and prospect of deep learning application in smart grid[J]. Automation of Electric Power Systems, 2019, 43(4): 180-191.
- [16] 袁一丹. 基于图像识别的工作人员穿戴规范性检测技术研究[D]. 成都: 电子科技大学, 2019.
- YUAN Y D. Research on standardized detection technology of staff wear based on image recognition[D]. Chengdu: University of Electronic Science and Technology of China, 2019.
- [17] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [J]. arXiv preprint arXiv:2207.02696, 2022.
- [18] CHOLLET F. Xception: Deep learning with depthwise separable convolutions [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1251-1258.
- [19] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks [C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020.
- [20] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]. Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980-2988.
- [21] GEVORGYAN Z. SIOU Loss: More powerful learning for bounding box regression [J]. arXiv preprint arXiv: 2205.12740, 2022.

作者简介



E-mail: dachao9898@163.com

Ji Chao received his B. Sc. degree from Xi'an Polytechnic University in 2009, M. Sc. degree from Northwestern Polytechnical University in 2012 and Ph. D. degree from Northwestern Polytechnical University in 2015, respectively. Now he is an associate professor in Xi'an Polytechnic University. His main research interests include intelligent detection of transmission line status and image processing and artificial intelligence.



侯威 (通信作者), 2020 年于淮阴工学院获得学士学位, 目前为西安工程大学硕士研究生, 主要研究方向为输电线路状态智能监测、深度学习与计算机视觉。

E-mail: weihouxp@163.com

Hou Wei (Corresponding author) received his B. Sc. degree from Huaiyin Institute of Technology in 2020. Now he is a M. Sc. candidate in Xi'an Polytechnic University. His main research interests include intelligent detection of transmission line status and image processing and artificial intelligence.



高鸣江, 2022 年于西安工程大学获得学士学位, 目前为西安工程大学硕士研究生, 主要研究方向为输电线路状态智能检测、深度学习与机器视觉。

E-mail: 1257483350@qq.com

Gao Mingjiang received his B. Sc. degree from Xi'an Polytechnic University in 2022. Now he is a M. Sc. candidate at Xi'an Polytechnic University. His main research interests include intelligent detection of transmission line status, deep learning and machine vision.