

DOI: 10.13382/j.jemi.B2307008

# 深度学习的接触网小目标缺陷识别研究\*

顾桂梅 王小亮

(兰州交通大学自动化与电气工程学院 兰州 730070)

**摘要:**吊弦线夹螺栓是铁路接触网供电线路的重要器件,其状态会影响电力机车受流质量,于是对SSD算法进行改进;首先引入一种轻量级神经网络 MobileNetV3 用于前端特征提取,降低模型复杂度,以提高检测速度;其次采用CA注意力机制替换反向残差结构线性瓶颈层的SE模块,使位置信息沿空间两个方向聚合,调整后的特征层能够捕获全局远程特征信息;最后设计了特征融合模块以重构特征层,优化小目标检测层以提高对小目标的识别效果。还用 CycleGAN 等方法扩充训练样本,解决数据集不足的问题。实验结果表明,改进算法的模型复杂度下降,mAP@0.5和FPS分别达到95.5%和81 fps,该研究有助于接触网检测仪器向小型移动嵌入式设备转变。

**关键词:**深度学习;小目标;注意力机制;接触网检测

**中图分类号:** U225.4;TP391.4

**文献标识码:** A

**国家标准学科分类代码:** 520.2060

## Identification of the catenary small target defects in deep learning

Gu Guimei Wang Xiaoliang

(School of Automation and Electrical Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China)

**Abstract:** The dropper clamp bolt is an important component of railway power supply line, which can affect the flow quality of electric locomotive. Therefore, this paper improves the SSD algorithm: Firstly, a lightweight neural network MobileNetV3 is introduced for front-end feature extraction to reduce the model complexity and improve the detection speed; secondly, CA attention mechanism to replace the SE module of the linear bottleneck layer with inverted residuals structure, aggregate the position information in the two directions of space, and the adjusted feature layer can capture the global remote feature information. Finally, the feature fusion module for reconstructing the feature layer is designed to adjust the small target detection layer to improve the detection effect of small targets. This paper also expands the training sample with CycleGAN to solve the problem of insufficient data set. The experimental results show that the model complexity of the improved algorithm decreased, and mAP @ 0.5 and FPS reached 95.5% and 81 fps, respectively. This study helps the transformation of catenary detection instruments to small mobile embedded devices.

**Keywords:** deep learning; small target; attention mechanism; catenary detection

## 0 引言

中国经济高质量发展离不开铁路安全、可靠的运输保障,接触网供电线路可靠运行是电气化铁路的重要一环。吊弦线夹螺栓是接触悬挂的重要组成器件,其状态会影响电力机车的受流质量。电气化列车在行进过程中会造成弓网间高频次、大强度的相互作用,损耗接触网供

电系统机械强度,引发吊弦线夹螺栓松动、螺母缺失等故障,进而影响列车的正常运行。现如今,一方面,随着接触网悬挂状态检测监测装置(4C)引入接触网检测车,检修效率大大提高,人工检修压力得到释放;另一方面,接触网检测车一般在晚上作业,异常曝光等因素不利于图像采集,螺栓缺陷目标区域较小,现有目标检测算法通常会造误检、漏检等问题,需要经验丰富的核验人员对采集的图像进行二次筛查,耗时耗力的重复作业不利于铁

收稿日期:2023-10-29 Received Date: 2023-10-29

\* 基金项目:甘肃省科技计划资助项目(20JR10RA216)资助

路行业的发展,须另辟蹊径。

2006 年 NVIDIA 推出了通用计算平台 CUDA,使开发者更加注重计算逻辑的实现。现如今,GPU 和并行技术的进一步革新助力深度学习迅速发展,特别是在计算机视觉领域,目标检测算法取得了突破性发展,使得自动、精准、快速检测在图像识别中成为可能,这为铁路接触网检测提供了新思路。基于深度学习的目标检测算法主要可以分为两类:一类是基于回归的 one-stage 目标检测算法,如 YOLO 系列<sup>[1-3]</sup>、SSD 等;另一类是基于区域建议的 two-stage 阶段算法,如 R-CNN、Faster R-CNN<sup>[4]</sup>等。目标检测算法在缺陷识别类的工作中有很多应用:贾耀华等<sup>[5]</sup>通过改进 YOLOv5s 算法识别铁路接触网吊弦缺陷目标,该方法引入卷积块注意力模块(convolutional block attention module, CBAM)<sup>[6]</sup>,但该模块只能提取局部关系,缺乏长距离关系提取的能力;李雪峰等<sup>[7]</sup>通过更换特征提取网络、融合特征层来改进 Faster R-CNN (FPN)算法检测输电线路销钉缺陷目标,由于该方法在重构特征层时增加模型复杂度,检测速度变慢;张姝等<sup>[8]</sup>采用分级检测原则,采用 SSD 算法定位、切割出缺陷螺栓连接部位,利用 YOLOv3 算法检测缺陷螺栓;康高强等<sup>[9]</sup>通过 Faster R-CNN 网络对旋转双耳整体进行定位,再使用多个深度卷积神经网络提取开口销钉的多种特征,最后利用由多个线性 VSM 构成的集成分类器实现开口销钉缺失故障检测;顾超越等<sup>[10]</sup>基于 Faster R-CNN,以 ResNet101 为特征提取网络,以 K-means 算法优化初始锚

框来对架空线路销钉缺陷进行检测;李瑞生等<sup>[11]</sup>基于 SSD 算法,通过引入残差网络,采用卷积拆分、权值量化压缩网络参数量实现对输电线路销钉缺陷检测;吴闽等<sup>[12]</sup>基于 SSD 算法,采用 ACGAN 进行样本扩充,并在主干网络中引入注意力机制实现对气体绝缘组合电器(GIS)多源局放模式识别;Fu 等<sup>[13]</sup>引入特征提取能力更强的 ResNet-101<sup>[14]</sup>网络,在后端预测网络引入反卷积来增强特征层的表征能力,进一步增强小目标的识别能力;Li 等<sup>[15]</sup>采用“早融合”的思想先在特征提取层进行融合,最后在特征层进行多尺度检测;Jeong 等<sup>[16]</sup>设计了多尺度特征层融合的“彩虹金字塔”来提升小目标检测效果,但改进后模型复杂度提高。上述研究中使用二阶段算法的检测方法虽然提高了识别精度,但检测速度较慢,在要求实时检测的场合并不实用,而使用一阶段算法的改进方法对小目标识别的效果不佳,于是本文针对该问题和接触网吊弦线夹螺栓缺陷识别等问题,提出了改进算法。

## 1 SSD 算法原理

SSD<sup>[17]</sup>属于经典的一阶段目标检测算法,它分别借鉴了 Faster R-CNN 和 YOLO 的锚框机制和回归机制。如图 1 所示,SSD 算法网络框架由两部分组成:第 1 部分为前端特征提取网络 VGG-16,包括卷积层 Conv1~Conv5,并将其后的全连接层 FC6 及 FC7 分别改为卷积层 Conv6、Conv7;第 2 部分是后端多尺度特征检测网络。

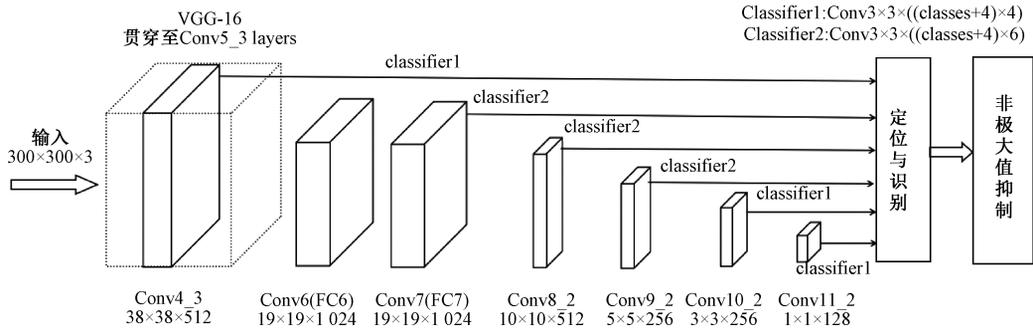


图 1 SSD 算法框架

Fig. 1 SSD algorithm framework

对于后端 6 种不同尺度的卷积特征层(Conv4\_3、Conv7、Conv8\_2、Conv9\_2、Conv10\_2 和 Conv11\_2),分别赋予不同数量的宽高比不定的默认框(default box, dbox)来匹配类别(classes)不同的检测目标;用卷积预测层 Classifier 通过 dbox 与真实框(ground truth box, gtbody)匹配得到类别得分和边界框回归参数,使用非极大值抑制(non-maximum suppression, NMS)得到最终检测结果。

特征层默认框尺度计算公式为:  $s_k = s_{\min} +$

$$\frac{s_{\max} - s_{\min}}{m - 1}(k - 1), k \in \{1, 2, \dots, m\}, \text{其中, } s_{\min} \text{ 取 } 0.2, s_{\max}$$

取 0.9,  $m$  为特征图层数。该算法的损失计算(式(1))包括类别损失(式(2))及定位损失(式(3))两个部分。损失函数如下:

$$L(x, c, l, g) = \frac{1}{N}(L_{\text{conf}}(x, c) + \alpha L_{\text{loc}}(x, l, g)) \quad (1)$$

$$L_{\text{conf}}(x, c) = - \sum_{i \in \text{Pos}} x_{ij}^p \log(\hat{c}_i^p) - \sum_{i \in \text{Neg}} \log(\hat{c}_i^0) \quad (2)$$

$$L_{loc}(x, l, g) = \sum_{i \in Pos_m \in \{cx, cy, w, h\}} \sum_{j \in Neg_m} x_{ij}^k \text{smooth}_{L1}(l_i^m - \hat{g}_j^m) \quad (3)$$

其中:

$$\hat{c}_i^p = \frac{\exp(c_i^p)}{\sum_p \exp(c_i^p)}, \hat{g}_j^{cx} = (g_j^{cx} - d_i^{cx})/d_i^w$$

$$\hat{g}_j^{cy} = (g_j^{cy} - d_i^{cy})/d_i^h, \hat{g}_j^w = \log(\frac{g_j^w}{d_i^w}), \hat{g}_j^h = \log(\frac{g_j^h}{d_i^h}),$$

$$\text{smooth}_{L1}(x) = \begin{cases} 0.5x^2 & |x| < 1 \\ |x| - 0.5 & \text{其他} \end{cases}, N \text{ 为匹配到的正样}$$

本数,  $\alpha$  取 1,  $\hat{c}_p$  是预测的第  $i$  个 dbox 对应 gtbox (类别为  $p$ ) 的类别概率,  $x_{ij}^p = \{0, 1\}$  是第  $i$  个 dbox 匹配到的第  $j$  个 gtbox (类别是  $p$ ),  $l_i^m$  为预测对应第  $i$  个正样本位置回归参数,  $\hat{g}_j^m$  为正样本  $i$  匹配的第  $j$  个 gtbox 的位置回归参数,  $d_i^h$  和  $d_i^w$  分别为第  $i$  个 dbox 的高和宽。Pos 指正样本, Neg 指负样本。{cx, xy, w, h} 为位置坐标。

不同深度的特征层有着各自的感受野和分辨率。深层特征图感受野较大但分辨率较低,用于识别大目标;浅层特征图感受野较低,但分辨率较高,含有丰富的底层特征信息,用于识别小目标。小目标<sup>[18]</sup> 基于相对尺度通常是指目标边界框的宽高与图像的宽高比小于 0.1, 基于绝对尺度是指小于 32 pixels×32 pixels。由图 2 可知,本文所识别的吊弦线夹螺栓缺陷目标宽高与图像(5120×5120)的宽高比主要聚集在中心点(0.030, 0.025)处,符合小目标的定义。

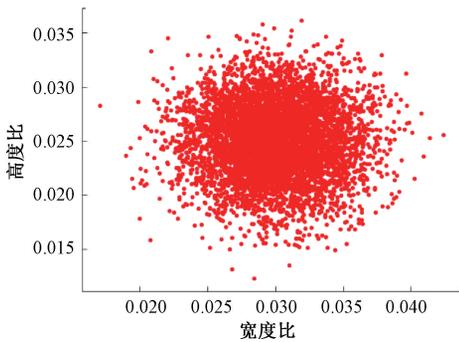


图 2 训练集样本尺寸与图像尺寸之比

Fig. 2 Ratio of sample size to image size in the training set

## 2 改进方法

### 2.1 引入 MobileNetV3 网络

MobileNetV3<sup>[19]</sup> 是基于深度可分离卷积的 MobileNet 系列<sup>[20-21]</sup> 最新变体,拥有更小的体积、计算量、参数量,在轻量级神经网络中拥有极大的优势,被广泛应用于移动终端和嵌入式设备当中。该算法借助网络结构自动搜索算法能够优化网络模块的超参数;引入反向残差结构线

性瓶颈层 (the linear bottleneck layers with inverted residuals structure, bneck) 优化网络结构,使得网络层次加深,模型计算量和参数量减少,高维信息在传输过程中得以保留;引入 SE 模块加强通道特征,提高模型精度。

如图 3 所示, bneck 层先经过 Conv 1×1 扩充通道数,再通过 Dwise 进行卷积变换及 SE 模块传输,最后用 Conv 1×1 压缩通道数, NL 是非线性激活函数 Relu 及 h-swish (式(4))。

$$h-swish[x] = x \frac{\text{ReLU6}(x + 3)}{6} \quad (4)$$

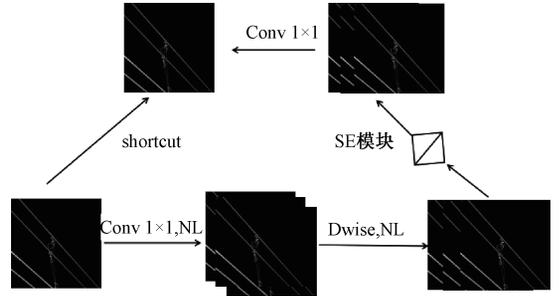


图 3 反向残差结构的线性瓶颈层

Fig. 3 The Linear bottleneck layers with inverted residuals structure

如图 4 所示, SE 模块<sup>[22]</sup> 首先进行压缩操作,通过全局平均池化将特征图输出为 1×1×C 矩阵;再进行激励操作,通过两个全连接层及 Scale 操作完成特征信息传输。ReLU 及 hard-σ (式(5)) 是激活函数。

$$\text{hard-}\sigma[x] = \frac{\text{ReLU6}(x + 3)}{6} \quad (5)$$

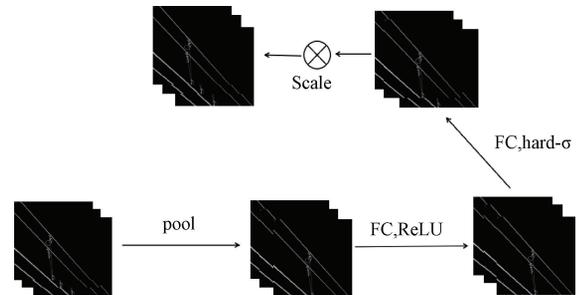


图 4 SE 模块

Fig. 4 SE module

### 2.2 引入坐标注意力机制

注意力机制注重目标特征信息,抑制无关干扰信息。近年来对轻量级网络设计的研究表明<sup>[23]</sup> 通道注意力模块使得模型性能得到显著提升,但该模块只考虑通道间信息,忽略了位置信息。随后改进的卷积块注意力机制 CBAM 试图通过减少输入张量的通道维度,然后使用卷积计算空间注意力来利用位置信息,然而卷积只能捕获

局部关系,而无法对视觉任务必不可少的远程依赖进行建模。

坐标注意力 (coordinate attention, CA) 模块能够将通道注意力分解为两个一维特征编码过程,并分别沿 2 个空间方向聚合特征,通过这种方式,可以沿一个空间方向捕获远程依赖关系,同时沿另一个空间方向保留精确的位置信息。因此本文把 MobiltNetV3 中反向残差结构的线性瓶颈层的 SE 模块替换为 CA 模块。CA 模块如图 5 所示。

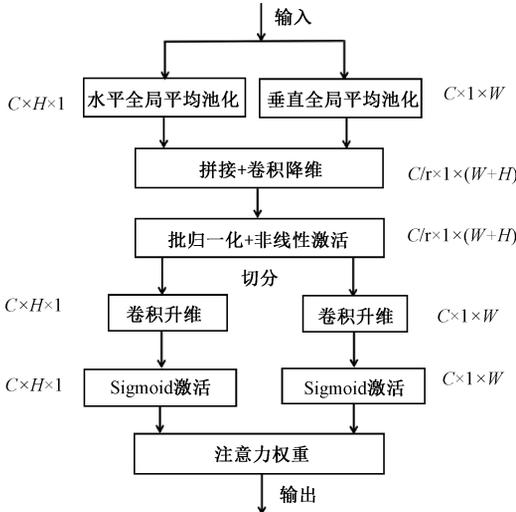


图 5 CA 模块  
Fig. 5 CA module

CA 注意力机制将输入特征信息从两个方向进行分解来获取位置信息,首先对输入的特征层进行垂直 (式 (6)) 和水平 (式 (7)) 的全局平均池化,接着把两个方向的特征进行拼接,采用卷积 1x1 压缩通道数 (式 (8)),经过批归一化和非线性回归并进行向量切分,再用卷积 1x1 扩充通道数,最后结合 sigmoid 激活函数得到的注意力向量权重 (式 (9)、(10)),进而得到将精确位置信息嵌入通道信息的特征图 (式 (11))。

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} x_c(h, i) \quad (6)$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w) \quad (7)$$

$$f = \delta(F_1([z_c^h, z_c^w])) \quad (8)$$

$$g^h = \delta(F_h(f^h)) \quad (9)$$

$$g^w = \delta(F_w(f^w)) \quad (10)$$

$$y_c(i, j) = x_c(i, j) \times g^h(i) \times g^w(j) \quad (11)$$

$z_c^h$  指垂直特征,  $z_c^w$  指水平特征,  $x_c$  指输入特征,  $c$  指特征向量的通道。  $f$  指注意力关联图,  $f^h$  指切分后的两个方向的向量,  $g^h, g^w$  指注意力权重,  $y_c$  指经过 CA 模块修正的特征图。

### 2.3 重构特征层

鉴于 SSD 算法特征层在卷积层预测时未进行融合,暴露出小目标识别效果差的问题,为提升模型对接触网吊弦线夹螺栓缺陷识别的效果,在原特征层构建特征融合模块 (feature fusion module, FFM)。对于接触网吊弦线夹螺栓尺寸来说,大多数目标采用  $N_1, N_2$  特征图来做识别。如图 6 所示,本文基于原有特征层  $\{N_1, N_2, N_3, N_4, N_5, N_6\}$ , 将  $N_1, N_2$  分别和  $N_3, N_4$  进行融合,将深层、浅层特征融合后,使其具备更加强劲的语义信息和位置信息。特征层融合具体实现操作为:特征层先通过卷积 1x1 操作、上采样统一通道数及尺寸,再进行 Concat 操作得到重构特征层。重构特征层后,相应的默认框尺寸调整为  $\{21, 45, 99, 153, 207, 261\}$ 。  $N_1, N_2$  用来检测小目标,  $N_3, N_4$  用来检测尺寸中等目标,  $N_5, N_6$  用来检测尺寸偏大的目标。相应的,默认框宽高比 (aspect ratio,  $a_r$ ) 调整为  $\{N_1, N_5, N_6\} \in (0.5, 1, 2), \{N_2, N_3, N_4\} \in (1/3, 1/2, 1, 2, 3)$ , 当宽高比为 1 时,新增尺度为  $(\sqrt{45 \times 99}, \sqrt{99 \times 153}, \sqrt{153 \times 207}, \sqrt{207 \times 261}, \sqrt{261 \times 315})$  的默认框。值得一提的是,本文所采用的特征融合策略不是 BiFPN<sup>[24]</sup>、FSSD 及 RSSD 的密集型融合,而是选择最靠近吊弦线夹螺栓真实目标框尺寸的特征层进行融合。

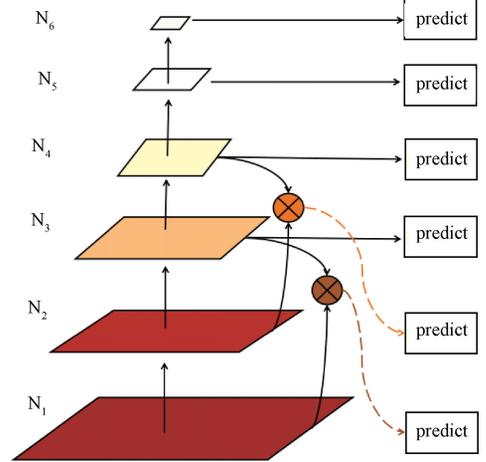


图 6 重构特征层  
Fig. 6 Reconstruct the feature layer

### 2.4 图像增强

在现有铁路接触网供电系统中,吊弦线夹螺栓缺陷是小概率事件,因此存在缺陷样本不足的情况,进而导致模型训练不充分,泛化性弱等问题。针对上述问题,对训练集进行扩充:首先对低灰度值的图像进行亮度增强,再采用中值滤波增强语义信息间的关联,其次对像素值分布过于集中或松散的图像采用自适应直方图均衡化增强对比度以提升图像的视觉效果,最后用 CycleGAN 扩充训练样本,处理过程如图 7 所示。

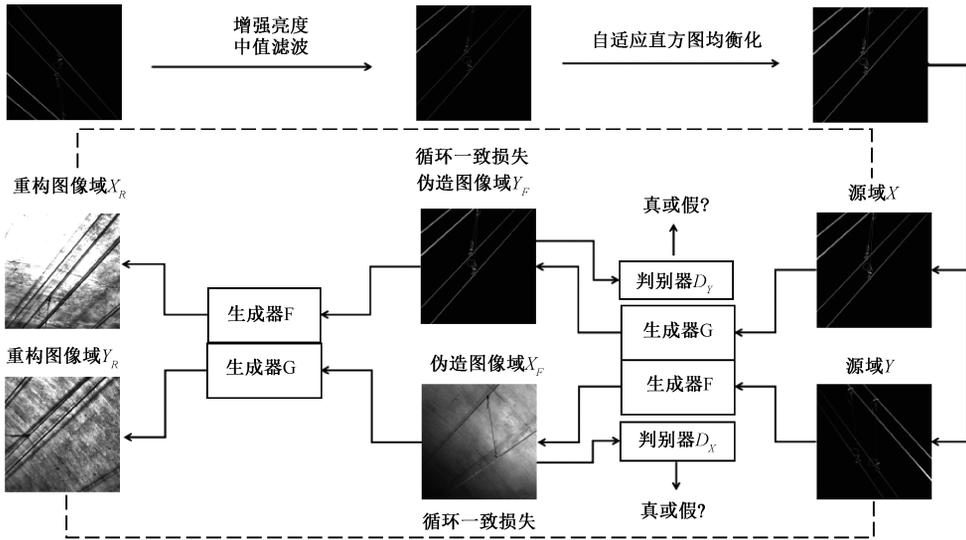


图 7 数据集预处理

Fig. 7 Data set preprocessing

### 3 实验与分析

#### 3.1 实验环境配置

本文实验所使用的环境配置为:CPU 型号:12th Gen Intel(R) Core(TM) i7-12700H(16 GB),GPU 型号:NVIDIA GeForce RTX 3060(8 GB),操作系统为 Windows 11,配置 Cuda11.3,Cudnn8.2.1,基于 Pytorch1.7.1 深度学习框架,Python3.9 实现网络。

#### 3.2 数据集

本文实验所用数据集来源于 4C 检测车,共采集到 882 张原始图像,尺寸为 5 120 pixels×5 120 pixels。经过数据集扩充后得到 3 990 张图像。数据标注格式为 Pascal VOC,标注工具为 labelImg。螺母松动的螺栓标记为“loosing”,螺母缺失的螺栓标记为“missing”,状态正常的螺栓则标记为“normal”。扩充数据集后随机选取 3 232 张图片作为训练集,361 张作为验证集,400 张作为测试集。缺陷样本如图 8 所示。



图 8 缺陷样本

Fig. 8 Defective sample

#### 3.3 模型训练设置

模型训练使用自适应矩估计(adaptive moment

estimation, Adam)作为梯度下降中的优化器,该算法结合了随机梯度下降算法(stochastic gradient descent,SGD)和自适应学习率算法,能够快速收敛并且减少训练时间,动量设置为 0.937,权重衰减系数设置为 0,训练图片输入尺寸调整为 600×600,批量设置为 32,总计训练轮次为 120,模型最大学习率设置为  $6 \times 10^{-4}$ ,最小学习率设置为  $(6 \times 10^{-4}) \times 0.01$ 。学习率下降方式采用余弦退火(cosine annealing),以避免算法陷入局部最优,使得模型损失计算通向全局最小值。

#### 3.4 实验结果分析

本文设计了两组实验进行检测结果评估与验证(每组统计了 3 次实验结果的平均值和标准差):第 1 组实验进行不同改进方式的消融实验来评估各个改进方式;第 2 组实验进行不同主流算法的对比试验,以评估改进模型的整体效果。改进算法与比较算法及 SSD300 在相同的数据集、实验环境配置和训练参数下分别训练了 120 轮次。由于实验涉及多个检测类别,选取平均准确率均值 mAP@0.5(式(15))、选取模型计算量(giga floating-point of operations,GFLOPs)、参数量(Params)和平均检测速度(frames per second,FPS)作为评价指标。

$$P = \frac{T_p}{T_p + F_p} \tag{12}$$

$$R = \frac{T_p}{T_p + F_N} \tag{13}$$

$$P_{AP} = \int_0^1 P(R) d(R) \approx \frac{1}{n} \sum_{i=1}^n P(R_i) \tag{14}$$

$$P_{mAP@0.5} = \frac{1}{C} \sum_{i=1}^C P_{AP}(i) \tag{15}$$

其中,  $P$  为精准率 (precision,  $P$ ),  $R$  为召回率 (recall,  $R$ ),  $T_p$  为预测为正且实际为正的样本数,  $F_p$  为预测为正但实际为负的样本数,  $F_n$  为预测为负但实际为正的样本数。  $P_{AP}$  为平均准确率,  $P_{mAP@0.5}$  为平均准确率均值。  $n$  为取样点的数量,  $C$  为检测类别数。 本文中  $n = 200$ ,  $C = 3$ 。

1) 不同改进方法的消融实验

本文改进算法模型以 SSD300 为基准, 通过分析消融实验结果判断不同改进方法的有效性。 基准模型和改进模型通过测试集对模型进行评估。 图 9 中改进算法模型用虚线表示, 基准模型用实线表示, 改进算法模型迭代至约 40 次时已经趋于收敛, 而基准模型迭代至约 90 次时损失曲线才趋于稳定, 改进模型比原模型损失值更低, 进

而改进模型的训练效果更好。 由图 10 可知: 改进模型在训练至约 40 轮次时已取得 95.5% 左右的优异检测结果, 而 SSD300 模型训练至约 90 轮次才达到 90.2% 左右的检测结果。 不同改进方法的消融实验结果如表 1 所示。 4 种方法的  $mAP@0.5$  分别提升了 0.9%、1.8%、1.3% 和 1.3%, 相比于基准模型, 改进模型的  $mAP@0.5$  提升了 5.3%, 在保证不同缺陷类别识别精度增加的同时, 检测速度略有提升, 进一步得知引入特征提取网络 MobileNetV3-Large、CA 模块、FFM 模块方法及图像增强方法能够提高模型的泛化性, 比较各改进方法的  $mAP@0.5$  曲线 (图 10)、原算法模型识别精度 (图 11) 及改进算法模型识别精度 (图 12) 得知改进模型优于原模型, 进一步验证了改进方法的有效性。

表 1 不同改进方法模型的实验结果

Table 1 Experimental results of different improved for method models

MobileNetV3-large	CA	FFM	图像增强	AP/%			mAP@0.5/%	平均检测速度/fps
				normal	loosing	missing		
				88.6±0.18	90.1±0.12	92.1±0.09	90.2	80
✓				89.4±0.38	91.0±0.58	93.1±0.26	91.1	88
	✓			90.4±0.32	91.8±0.36	94.0±0.18	92.1	85
		✓		89.8±0.35	91.4±0.38	93.5±0.22	91.6	83
			✓	89.7±0.5	91.4±0.11	93.6±0.09	91.5	81
✓	✓			91.1±0.21	92.8±0.58	94.9±0.15	92.9	85
✓		✓		90.6±0.15	92.3±0.32	94.5±0.12	92.5	87
✓			✓	90.8±0.14	92.2±0.32	94.4±0.28	92.4	87
✓	✓	✓		94.5±0.28	93.1±0.32	96.2±0.15	94.6	83
✓		✓	✓	91.9±0.26	93.5±0.38	95.8±0.19	93.7	85
✓	✓		✓	92.4±0.15	94.1±0.28	96.2±0.12	94.2	85
	✓	✓	✓	92.6±0.19	94.8±0.38	96.1±0.12	94.5	79
✓	✓	✓	✓	93.1±0.22	97.3±0.21	96.1±0.08	95.5	81

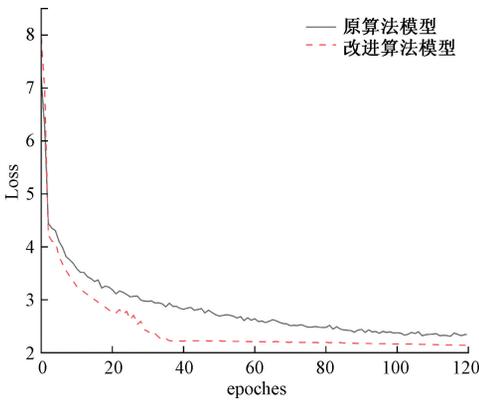


图 9 改进前后损失对比曲线

Fig. 9 The loss comparison curve

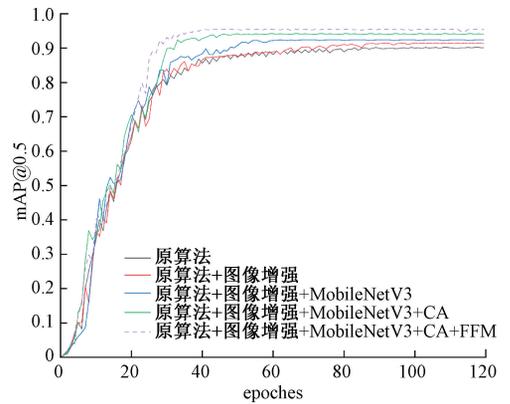


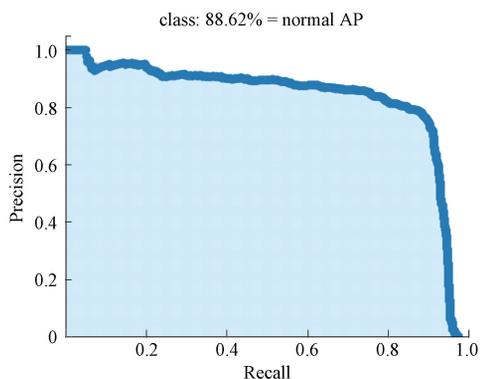
图 10 mAP@0.5 对比曲线

Fig. 10 Comparison curve For mAP @ 0.5

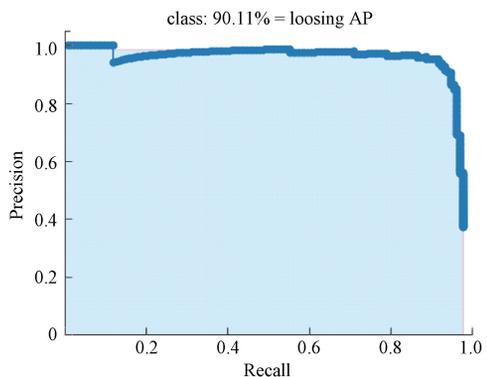
2) 不同算法性能的比较实验

为了更好的测试改进算法的优越性, 选用部分主流的二阶段算法 Faster R-CNN (VGG16)、一阶段算法 YOLOv7、DSSD、FSSD 及 SSD300 与其做对比实验, 通过

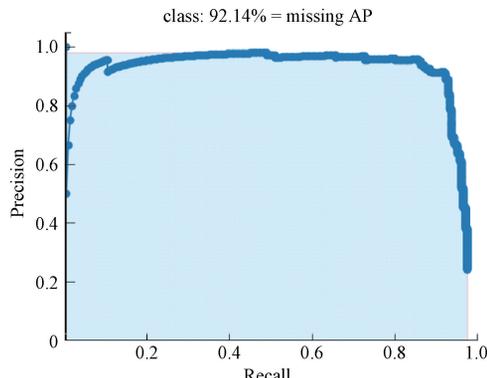
分析各个算法在不同评价指标的实验结果来评估本文所提算法的有效性。 分析图 10、11、12 及表 2 实验数据可知: 改进算法模型相比于其他算法模型, “normal” 处的 AP 值达到 93.1%、 “loosing” 处的 AP 值达到 97.3%、



(a) “正常”精度  
(a) Accuracy of normal



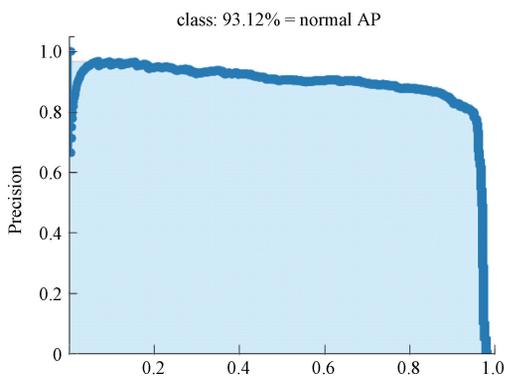
(b) “松动”精度  
(b) Accuracy of losing



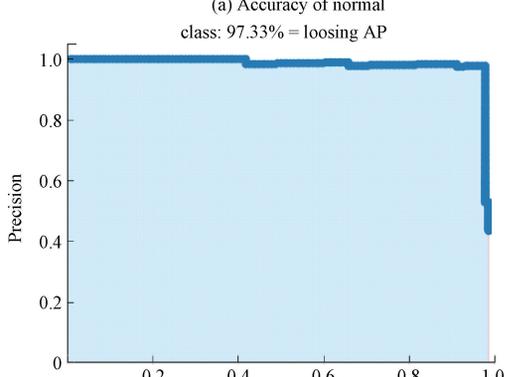
(c) “缺失”精度  
(c) Accuracy of missing

图 11 原算法模型各类别识别精度

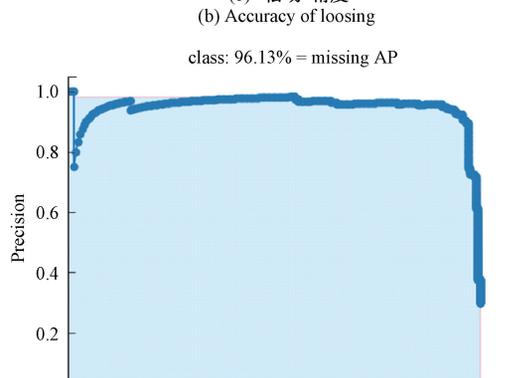
Fig. 11 The classification identification accuracy of the original algorithm model



(a) “正常”精度  
(a) Accuracy of normal



(b) “松动”精度  
(b) Accuracy of losing



(c) “缺失”精度  
(c) Accuracy of missing

图 12 改进算法模型各类别识别精度

Fig. 12 The identification accuracy of the improved algorithm model

“missing”处的 AP 值达到 96.1%，模型计算量低至 55.232 G，参数量低至 17.602 M，检测速度达到 81 fps，召回率均在 70% 以上，在上述主流算法模型中综合性能表现最优。值得注意的是，虽然本文所提算法大量使用深度可分离卷积，但是由于在特征提取网络和特征层分别加入了 CA 模块和 FFM 模块，使得模型检测速度提升幅度不大。图 13 是部分改进算法模型的缺陷识别效果图，达到了无漏检，无误检的要求。

## 4 结论

本文以 SSD 算法为基准做出改进：1) 以轻量型的 MobileNetV3 作为前端特征提取网络，在确保识别精度提高的同时降低计算量和参数量；2) 在反向残差结构的瓶颈层引入能够捕获双向特征信息的坐标注意力模块，使

表 2 各算法模型性能对比

Table 2 Performance Comparison of each algorithm model

模型	AP			mAP@0.5	计算量/G	参数量/M	平均检测速度/fps
	normal	loosing	missing				
Faster-R-CNN(VGG-16)	0.912±0.38	0.921±0.32	0.926±0.21	0.920	369.793	136.750	14
YOLOv7	0.912±0.15	0.883±0.18	0.902±0.12	0.899	103.62	35.232	72
DSSD	0.916±0.58	0.936±0.32	0.926±0.28	0.926	83.654	61.255	16
FSSD	0.901±0.62	0.923±0.26	0.916±0.22	0.913	66.354	28.363	58
SSD300	0.886±0.18	0.901±0.12	0.921±0.09	0.902	61.055	24.013	79
改进算法模型	0.931±0.22	0.973±0.21	0.961±0.08	<b>0.955</b>	55.232	17.602	<b>81</b>

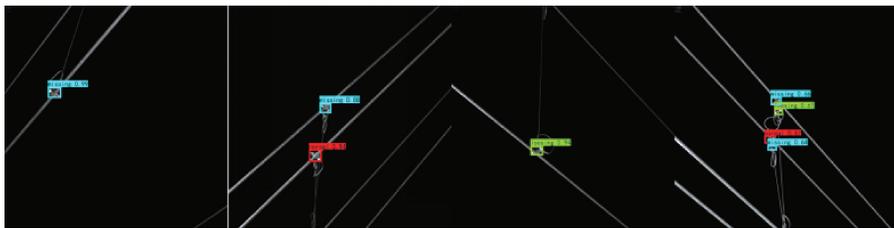
(a) 原始图  
(a) The original figure(b) 识别图  
(b) The identification figure(c) 局部放大图  
(c) Partial enlarged figure

图 13 改进算法缺陷识别结果

Fig. 13 Defect recognition results of the improved algorithm

调整后的注意力图具备远程依赖关系,并对位置信息敏感,增强主干网络的特征提取能力;3)设计特征融合模块重构特征层,使新增特征层具备深、浅层的特征信息,优化小目标识别效果;同时针对数据集缺陷类样本不足的问题进行图像扩充。实验结果表明,改进算法模型复杂度下降,平均准确率均值提升至 0.955,检测速度达到 81 fps,充分证明了本文改进方法的有效性。

### 参考文献

[ 1 ] 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 ( CVPR ). Las Vegas: IEEE, 2016: 779-788.

[ 2 ] FARHADI A, REDMON J. Yolov3: An incremental improvement [ C ]. Computer Vision and Pattern Recognition. Berlin/Heidelberg, Germany: Springer, 2018, 1804: 1-6.

[ 3 ] 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 [ C ]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern

- Recognition, 2023; 7464-7475.
- [ 4 ] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [ 5 ] 顾桂梅, 贾耀华, 温柏康. 基于YOLOv5s的接触网吊弦线和载流环缺陷识别算法[J]. 铁道科学与工程学报, 2023, 20(3): 1066-1076.
- GU G M, JIA Y H, WEN B K. Defect identification algorithm of dropper line and current-carrying ring of catenary based on YOLOv5s [J]. Journal of Railway Science and Engineering, 2023, 20(3): 1066-1076.
- [ 6 ] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional Block Attention Module [M]. Computer Vision-ECCV 2018. Cham: Springer International Publishing, 2018: 3-19.
- [ 7 ] 李雪峰, 刘海莹, 刘高华, 等. 基于深度学习的输电线路销钉缺陷检测[J]. 电网技术, 2021, 45(8): 2988-2995.
- LI X F, LIU H Y, LIU G H, et al. Transmission line pin defect detection based on deep learning [J]. Power System Technology, 2021, 45(8): 2988-2995.
- [ 8 ] 张姝, 王昊天, 董骁翀, 等. 基于深度学习的输电线路螺栓检测技术[J]. 电网技术, 2021, 45(7): 2821-2829.
- ZHANG S, WANG H T, DONG X C, et al. Bolt detection technology of transmission lines based on deep learning[J]. Power System Technology, 2021, 45(7): 2821-2829.
- [ 9 ] 康高强, 高仕斌, 于龙, 等. 基于深度学习的高铁接触网旋转双耳开口销钉缺失故障检测[J]. 铁道学报, 2020, 42(10): 45-51.
- KANG G Q, GAO S B, YU L, et al. Fault detection of missing split pins in swivel with clevis in high-speed railway catenary based on deep learning[J]. Journal of The Chain Railways Society, 2020, 42(10): 45-51.
- [ 10 ] 顾超越, 李喆, 史晋涛, 等. 基于改进Faster-RCNN的无人机巡检架空线路销钉缺陷检测[J]. 高电压技术, 2020, 46(9): 3089-3096.
- GU C Y, LI Z, SHI J T, et al. Detection for pin defects of overhead lines by UAV patrol image based on improved Faster-RCNN [J]. High Voltage Engineering, 2020, 46(9): 3089-3096.
- [ 11 ] 李瑞生, 张彦龙, 翟登辉, 等. 基于改进SSD的输电线路销钉缺陷检测[J]. 高电压技术, 2021, 47(11): 3795-3802.
- LI R S, ZHANG Y L, ZHAI D H, et al. Pin defect detection of transmission line based on improved SSD[J]. High Voltage Engineering, 2021, 47(11): 3795-3802.
- [ 12 ] 吴闽, 蒋伟, 罗颖婷, 等. 基于改进SSD的GIS多源局放模式识别[J]. 高电压技术, 2023, 49(2): 812-821.
- WU M, JIANG W, LUO Y T, et al. Multi-source partial discharge pattern recognition in GIS based on improved SSD [J]. High Voltage Engineering, 2023, 49(2): 812-821.
- [ 13 ] FU C Y, LIU W, RANGA A, et al. DSSD: Deconvolutional single shot detector [J]. arXiv preprint arXiv:1701.06659, 2017.
- [ 14 ] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C]. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 770-778.
- [ 15 ] LI Z X, ZHOU F Q. FSSD: Feature fusion single shot multi box detector[J/OL]. (2018-05-17)[2020-06-07]. <https://arxiv.org/abs/1712.00960>.
- [ 16 ] JEONG J, PARK H, KWAK N. Enhancement of SSD by concatenating feature maps for object detection [EB/OL]. (2017-05-26) [2019-08-28]. <https://arxiv.org/abs/1705.09587>.
- [ 17 ] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. Proceedings of the 14th European Conference. Amsterdam, Netherlands: Springer, 2016: 21-37.
- [ 18 ] CHEN C, LIU M Y, TUZEL O, et al. RCNN for small object detection [C]. Proceeding of Asian Conference on Computer Vision. Cham: Springer, 2016: 214-230.
- [ 19 ] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3 [C]. Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019: 1314-1324.
- [ 20 ] HOWARD A G, ZHU M, CHEN B, et al. Mobile Ntse: Efficient convolutional neural networks for mobile vision applications [J]. arXiv preprint arXiv: 1704.04861, 2017.
- [ 21 ] SANDLER M, HOWARD A, ZHU M L, et al. MobileNetV2: Inverted residuals and linear bottlenecks [C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, IEEE, 2018: 4510-4520.
- [ 22 ] JIE H, LI S, SAMUEL A, et al. Squeeze-and-excitation networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019: 7132-7141.
- [ 23 ] HOU Q B, ZHOU D Q, FENG J S, et al. Coordinate attention for efficient mobile network design [C]. Computer Vision and Pattern Recognition, 2021.
- [ 24 ] TAN M X, PANG R M, LE Q V. EfficientDet: Scalable and efficient object detection [C]. 2020 IEEE/CVF

Conference on Computer Vision and Pattern Recognition.  
June 13-19, 2020, Seattle, WA, USA. IEEE, 2020:  
10778-10787.

### 作者简介



**顾桂梅**(通信作者), 2005 年于兰州交通大学获得硕士学位, 现为兰州交通大学自动化与电气工程学院教授, 硕士生导师, 主要研究方向为智能信息处理、图像处理、故障诊断。

E-mail: 386509464@qq.com

**Gu Guimei** (Corresponding author) received her M. Sc. degree from Lanzhou Jiaotong University in 2005. Now she is a master tutor in School of Automation and Electrical Engineering,

Lanzhou Jiaotong University. Her main research interests include fault diagnosis, intelligent information processing and image processing.



**王小亮**, 2019 年于兰州交通大学获得学士学位, 现为兰州交通大学自动化与电气工程学院 2022 级在读硕士研究生, 主要研究方向为缺陷识别、图像分类。

E-mail: 2039598780@qq.com

**Wang Xiaoliang** received his B. Sc. degree from Lanzhou Jiaotong University in 2019. Now he is a M. Sc. candidate in the School of Automation and Electrical Engineering, Lanzhou Jiaotong University. His main research interests include defect recognition and image classification.