

# 基于改进 YOLOv4 的电力设备红外图像识别

段中兴<sup>1\*</sup>, 张雨明<sup>1</sup>, 马佳豪<sup>2</sup>

<sup>1</sup>西安建筑科技大学信息与控制工程学院, 陕西 西安 710055;

<sup>2</sup>国网西安供电公司, 陕西 西安 710032

**摘要** 针对电力设备红外图像存在的图像视觉效果差、尺度差异性大、数据类别不平衡等问题,提出了一种基于改进 YOLOv4 的目标检测模型。利用该模型并结合 MSRCR 图像增强算法对电力设备红外图像进行识别。首先,构建电力设备红外图像目标检测数据集,使用 MSRCR 算法对原始红外图像进行增强处理,改善雨雾天气下红外图像存在的对比度低、像素模糊等问题,进而提高模型在雨雾天气下对电力设备的检测能力;其次,在 YOLOv4 骨干网络中引入多尺度卷积模块,通过不同大小的卷积核获取输入图像的多尺度特征,增强对初始特征的代表能力;为进一步提升检测精度,采用 Focal loss 函数,解决了红外图像数据不均衡导致的分类困难问题。经测试,所提方法对 8 类电力设备的平均识别精度达 96.31%,检测速度达 71 frame/s。实验结果验证了所提方法的有效性。

**关键词** 目标检测; 图像增强; 红外图像; 电力设备识别

中图分类号 TP391.4

文献标志码 A

DOI: 10.3788/LOP202259.2410002

## Infrared Image Recognition of Power Equipment Using Improved YOLOv4

Duan Zhongxing<sup>1\*</sup>, Zhang Yuming<sup>1</sup>, Ma Jiahao<sup>2</sup>

<sup>1</sup>College of Information and Control Engineering, Xi'an University of Architecture and Technology,  
Xi'an 710055, Shaanxi, China;

<sup>2</sup>State Grid Xi'an Power Supply Company, Xi'an 710032, Shaanxi, China

**Abstract** In this paper, we propose an improved YOLOv4 target detection model combined with an MSRCR image enhancement algorithm to mitigate the poor visual effect, large-scale difference, and unbalanced data category in infrared images of power equipment. First, we constructed a target detection dataset of infrared images of power equipment and applied the MSRCR algorithm to enhance the original infrared images and improve the low contrast and blur pixel of infrared images in rain and fog weather. This improves the detection ability of the model to power equipment in rain and fog weather. Second, a multi-scale convolution module was introduced in the YOLOv4 backbone network to obtain multi-scale features of input images using convolution kernels with different sizes to enhance the initial feature representation. To further improve the detection accuracy, the Focal loss function was used to solve the difficult problem of classification caused by unbalanced infrared image data. The test results show that the average recognition accuracy of the proposed method for eight power equipments is 96.31%, and the detection speed is 71 frame/s. The experimental results verify the effectiveness of the proposed method.

**Key words** target detection; image enhancement; infrared image; identification of power equipment

## 1 引言

红外成像技术在变电站运营维护中有着广泛的应用<sup>[1]</sup>,通过及时获取电力设备的状态信息,能够有效预防设备过热引起的电器故障。传统的红外图像识别主要依靠人工进行分析、判断,存在耗时长、工作量大等

问题,难以做到实时监测。因此,快速、准确地对红外图像中的目标设备进行定位与识别对提高变电站智能化水平有着重要研究意义。

针对电力设备图像检测与识别问题,研究人员先后提出了各种各样的方法,目前这些方法主要可以分为基于传统机器学习的方法和基于深度学习的方法两

收稿日期: 2021-09-06; 修回日期: 2021-10-17; 录用日期: 2021-10-27

基金项目: 国家自然科学基金(51678470)

通信作者: \*zhx\_duan@163.com

大类。基于传统机器学习的电力设备图像识别方法主要通过类似滑动窗口的探测器检测图像局部区域是否包含目标设备<sup>[2]</sup>或通过提取图像特征信息进行聚类分析<sup>[3-5]</sup>,但此类方法特征辨识度弱、检测效率较低。随着我国智能电网的快速发展,对信息技术的全面感知有了更高的要求,这些方法已无法满足识别精度和速度的要求。

近年来,基于深度学习的方法在图像分类、目标检测等领域有着广泛的应用,同样被研究人员用于电力设备图像识别领域<sup>[6-7]</sup>。文献[8]将 GoogLeNet 网络应用于电力设备图像识别中,通过 Softmax 分类层输出最终结果,但该网络容易受背景信息干扰,鲁棒性较差。文献[9]在 Faster R-CNN 模型<sup>[10]</sup>中引入双特征映射,通过参数的动态调整实现对不同尺寸变压器的检测,但此方法在识别过程中会先产生若干个目标候选域,再对候选域进行细分类及位置回归,计算成本过大,导致识别速度慢。文献[11]以经典目标检测算法 SSD<sup>[12]</sup>为基准,通过将模型原特征提取网络替换为轻量化的 MobileNets 结构<sup>[13]</sup>降低网络的参数量,进而提高识别速度,但该网络用于检测小目标的特征层所含语义信息少,容易导致尺度较小的电力设备漏检或误检。文献[14]提出一种端到端的目标识别和定位算法(YOLO),该算法将检测任务与分类任务集于一体,最大的特点是运行速度快,满足实时性要求。随后,文献[15-17]提出多种改进策略,使得 YOLO 的目标检测精度和速度有了明显提高。但是这些改进策略应用于电力设备红外图像检测<sup>[18]</sup>时存在以下缺点:对特征尺度差异性较大的目标检测不够理想;难以应对天气不同导致的目标差异;数据样本不均导致模型泛化能力差等。

基于以上分析,本文选择目前工业领域应用较为广泛的 YOLOv4 目标检测算法<sup>[19-20]</sup>作为研究对象,提出一种基于改进 YOLOv4 的电力设备红外图像识别方法。针对雨雾天气下红外图像对比度低、像素模糊等问题,在网络输入端引入 multi-scale Retinex with color restoration(MSRCR)图像增强算法,提高模型在雨雾天气下对电力设备的检测能力;针对视角变化导致电力设备尺度差异性大的问题,在特征提取阶段引入多尺度卷积模块,利用不同大小的卷积核获取红外图像的多尺度特征信息,弱化视角变化带来的干扰;为进一步提升检测精度,采用 Focal loss 函数作为模型的分误差损失函数,以解决红外图像数据类别不均问题,增强模型的泛化能力。

## 2 MSRRCR 图像增强算法

电力设备所处环境是影响电力设备红外图像识别精度的因素之一。由于大气介质对户外成像有一定的干扰,雨雾天气下所拍摄的电力设备红外图像存在对比度低、像素模糊、边缘特征信息弱、纹理丢失等问题,

将对红外图像中电力设备的定位与识别带来干扰。基于此,本文采用 MSRRCR 算法对红外图像进行数据增强处理,通过改善红外图像的视觉效果,提高模型在雨雾环境下对电力设备的检测能力。

MSRRCR 算法是基于 Retinex 理论和 multi-scale Retinex(MSR)算法发展而来的一种图像增强方法。Retinex 理论<sup>[21]</sup>利用 HVS 系统在不同光照条件下的颜色恒常性,通过高斯函数对图像色彩通道进行滤波,进而降低光照变化对物体成像的干扰,该理论在图像去雾、色彩恢复等领域有着广泛的应用。由于单尺度 Retinex 算法在图像动态压缩范围过大时难以平衡图像的颜色保真度和细节保持度,MSR 算法<sup>[22]</sup>引入了尺度权重,通过对一幅图像在不同尺度上进行滤波,并对获取的滤波结果进行平均加权,能够较好地平衡图像的对比度和色阶层次,但在图像动态压缩过程中,容易引起三基色的比例失调,导致图像偏色。为解决这一问题,MSRRCR 算法引入了色彩恢复参数  $C$ ,按照图像原有色彩的比例,调整增强后的图像色彩比例,从而在一定程度上解决图像的偏色问题,具体过程涉及的表达式为

$$R_{\text{MSRRCR}}(x, y) = C_i(x, y) R_{\text{MSR}}(x, y), \quad (1)$$

$$R_{\text{MSR}}(x, y) = \sum_{k=1}^N \theta_k \left\{ \log I_i(x, y) - \log [I_i(x, y) * G_k(x, y)] \right\}, \quad (2)$$

$$C_i(x, y) = \kappa \cdot \log \left[ \sigma \frac{I_i(x, y)}{\sum_{i=1}^3 I_i(x, y)} \right], \quad (3)$$

式中: $I_i(x, y)$ 表示第  $i$  个通道的输入图像;\*表示卷积运算; $G(x, y)$ 表示高斯核函数; $N$ 为尺度个数, $N=3$ ; $\theta_k$ 表示第  $k$  个尺度的权重,且满足  $\sum_{k=1}^N \theta_k = 1$ ; $C_i(x, y)$ 为色彩恢复参数; $\kappa$ 和  $\sigma$ 为矫正参数; $R_{\text{MSRRCR}}$ 和  $R_{\text{MSR}}$ 分别表示经 MSRRCR 算法和 MSR 算法处理后的输出图像。

经大量实验验证,MSRRCR 图像增强算法中的参数设置如表 1 所示。

表 1 MSRRCR 算法参数取值表  
Table 1 Parameter value of MSRRCR algorithm

Parameter	$N$	$C_1$	$C_2$	$C_3$	$\theta_1$	$\theta_2$	$\theta_3$	$\kappa$	$\sigma$
Value	3	15	75	225	0.33	0.33	0.34	43	131

表 1 中的参数是经过大量实验验证得出的最优值,由参数  $C$  可知,本文将红外图像划分为大、中、小三个尺度,通过综合三个不同尺度大小的色彩通道对原始图像进行增强。图 1 为原始红外图像。图 2 为经 MSRRCR 增强处理后的图像。对比图 1 和图 2 可以看出,雨雾天气下采集到的原始图像视觉效果较差,利用 MSRRCR 算法对原始红外图像进行数据增强后,图像的雾化程度、边缘特征有了明显的改善。

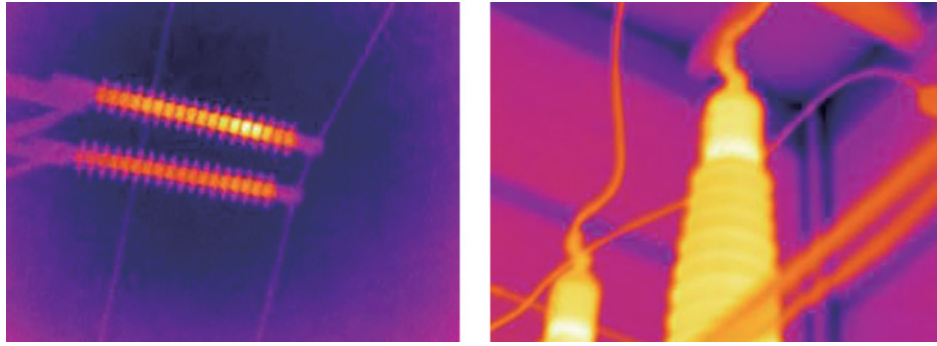


图 1 原始红外图像

Fig. 1 Original infrared images

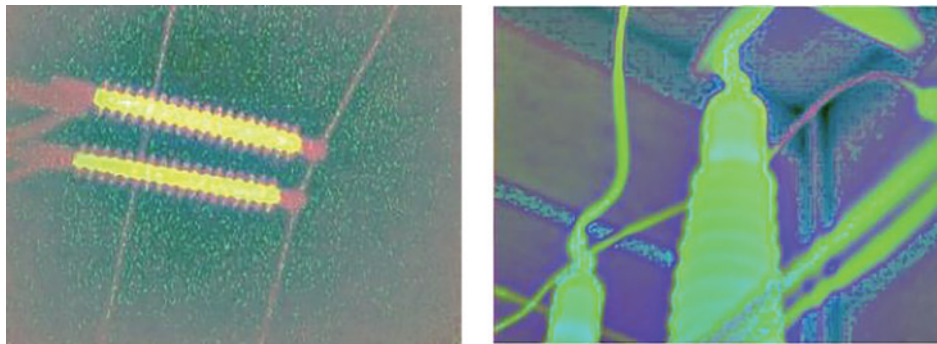


图 2 经 MSRCR 增强处理后的图像

Fig. 2 Images enhanced by MSRCR

### 3 YOLOv4 目标检测算法

YOLOv4 是目前实时性最优的目标检测算法,网

络结构如图 3 所示,该算法由基础骨干网络 Backbone、颈部网络 Neck、头部网络 Head 三部分组成。该算法首先利用基础骨干网络对输入红外图像进行初步特征

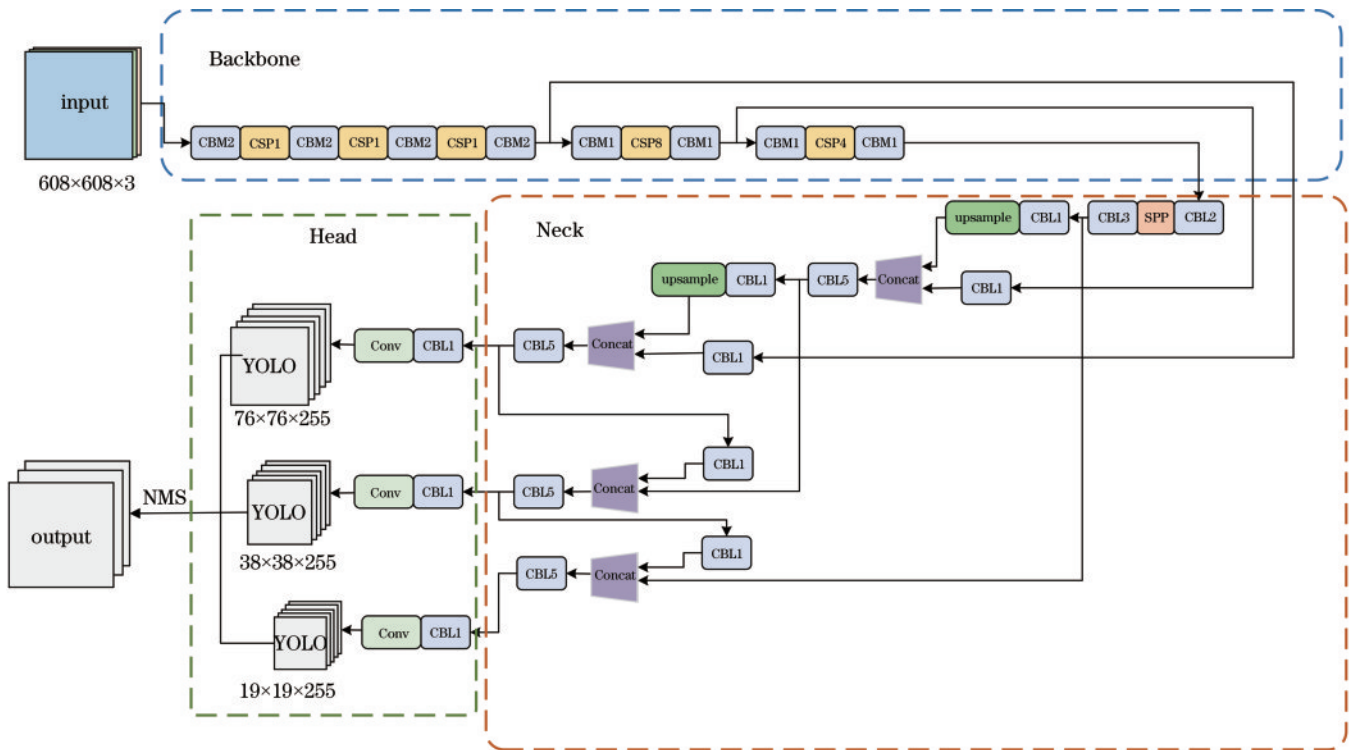


图 3 YOLOv4 网络结构

Fig. 3 YOLOv4 network structure



提取;其次,通过颈部网络聚合图像的多层特征信息;之后,使用头部网络预测目标的位置和类别,生成多个预测框;最后,利用非极大值抑制(NMS)算法生成最佳预测框。

### 3.1 基础骨干网络

YOLOv4 采用 CSPDarkNet53 作为基础骨干网络,该网络是在 DarkNet53 基础上,借鉴 CSPNet 的优化策略改进而来的特征提取网络。其主要优点在于:通过引入残差结构消除卷积层数增加时引起的梯度消失、梯度爆炸问题,进而增强卷积神经网络的学习能力;在保证准确率的前提下实现了模型的轻量化改进,降低模型参数量;采用了 Mish 激活函数,有助于稳定网络梯度流,增强信息传播,提高模型的准确率和泛化能力。

### 3.2 颈部网络

颈部网络由空间金字塔池化(SPP)模块和路径汇聚网络(PANet)组成。其中,SPP 模块用于对 CSPDarkNet53 最后一个特征层输出的特征图进行不同尺度的最大池化操作,之后对生成的特征图与原特征图进行聚合输出,以提取不同感受野下的特征信息。PANet 在特征金字塔的基础上增加了自下而上的特征融合结构,通过将高层语义信息引入空间细节信息,低层特征引入高级语义信息,避免了特征传播过程中

低层细节信息丢失问题,提高目标检测的识别精度。

### 3.3 头部网络

头部网络通过对特征层进行不同尺度的下采样操作,得到三种不同尺度的特征图,分别对应大尺度、中尺度、小尺度三种电力设备目标特征,在特征图的每个网格单元上预测三个边界框,每个边界框均包含置信度、类别分数和位置信息。

## 4 改进 YOLOv4 的电力设备红外图像识别方法

### 4.1 多尺度特征提取

在目标检测任务中,初始特征的提取对网络的最终检测效果有着直接影响。由于红外成像仪拍摄图像时的视角变化,红外图像中同一类电力设备的尺度大小不一。单列卷积每层使用固定大小的卷积核,难以应对电力设备尺度变化的问题。针对此问题,设计了一种用于全局特征提取的基于多列结构的多尺度卷积模块。采用 3 个不同大小的卷积核从输入的红外图像中学习尺度相关的特征,实现多尺度信息的有效获取,采用的多尺度卷积模块结构如图 4 所示。经实验验证,采用 3 个并列的  $3 \times 3$ 、 $5 \times 5$ 、 $7 \times 7$  卷积核能够有效聚合红外图像的多尺度信息。

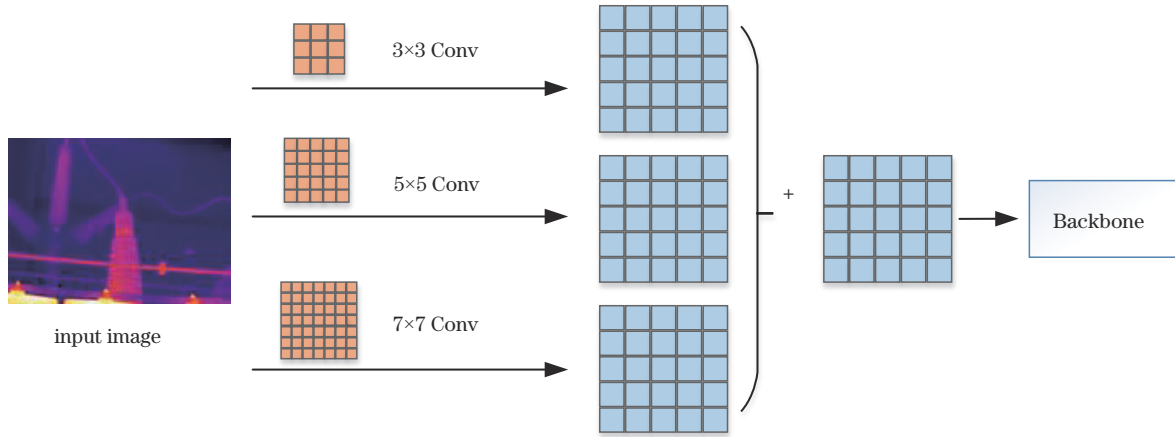


图 4 多尺度卷积模块

Fig. 4 Multi-scale convolution module

### 4.2 损失函数改进

YOLOv4 算法的损失函数由预测框回归误差  $L_{obj}$ 、置信度误差  $L_{conf}$ 、分类误差  $L_{class}$  三部分构成:

$$L = L_{obj} + L_{conf} + L_{class}, \quad (4)$$

式中: $L_{obj}$  用于确定目标所在的具体位置; $L_{conf}$  用于确定预测框存在目标的概率; $L_{class}$  用于确定预测框的类别概率。其中,预测框回归误差采用 CIoU 损失<sup>[23]</sup> 计算方式,置信度误差和分类误差采用交叉熵损失形式,具体表示为

$$L_{obj} = \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{obj} \left[ 1 - R_{IoU} + \frac{\rho^2(b, b^{gt})}{m^2} + \alpha v \right], \quad (5)$$

$$L_{conf} = \sum_{i=0}^{S^2} \sum_{j=0}^B (I_{ij}^{obj} + \lambda_{noobj} I_{ij}^{noobj}) \times \left[ \hat{F} \log F + (1 - \hat{F}) \log(1 - F) \right], \quad (6)$$

$$L_{class} = \sum_{i=0}^{S^2} I_{ij}^{obj} \sum_{c \in C_{class}} \left[ \hat{P} \log P + (1 - \hat{P}) \log(1 - P) \right], \quad (7)$$

$$\alpha = \frac{v}{(1 - R_{IoU}) + v}, \quad (8)$$

$$v = \frac{4}{\pi^2} \left( \arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2, \quad (9)$$

式中: $S^2$  表示输入图像被划分的网格数; $B$  表示预测框个数; $I_{ij}^{obj}$  和  $I_{ij}^{noobj}$  表示预测框是否包含目标,取值为 0 或

1;  $R_{IoU}$  为预测框和真实框之间的交并比;  $\lambda_{noobj}$  表示交并比的误差权重;  $\rho(b, b^{gt})$  为预测框中心点和真实框中心点之间的欧氏距离;  $m$  为预测框和真实框的最小包围框的对角线距离;  $\alpha$  为权重参数;  $v$  表示用于度量长宽比相似性的参数;  $F$  和  $\hat{F}$  为预测框和真实框的置信度;  $P$  和  $\hat{P}$

为预测框和真实框的类别概率;  $b, w, h$  和  $b^{gt}, w^{gt}, h^{gt}$  分别表示预测框和真实框的中心坐标、宽度、高度值。

由于电力设备所处环境的背景差异性, 不同电力设备之间的外形相似性, 红外图像中电力设备分类的难易程度不同, 如图 5 所示。

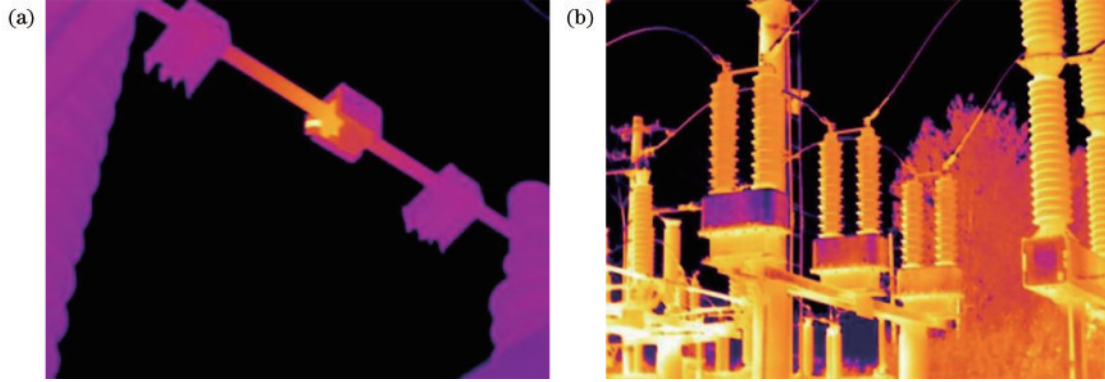


图 5 电力设备红外图像。(a)易分类样本;(b)难分类样本

Fig. 5 Infrared images of power equipment. (a) Easy to classify sample; (b) difficult to classify sample

针对电力设备红外图像类别不平衡问题, 将分类交叉熵损失函数替换为 Focal loss 函数<sup>[24]</sup>, 通过减小易分类样本的权重, 模型在训练时更专注于难分类的样本, 从而达到增强模型泛化能力的目的。Focal loss 函数的具体表达式为

$$L_F = -\beta(1 - y')^\gamma \log y', \quad (10)$$

式中:  $y'$  表示预测值;  $\beta$  和  $\gamma$  为调制参数, 其中  $\beta$  用于平衡样本本身的比例不均,  $\gamma$  用于调节不同样本对损失函数的权重, 取值范围为 (0, 1)。改进后的 YOLOv4

算法的总损失函数为

$$L' = L_{obj} + L_{conf} + L_F. \quad (11)$$

## 5 实验与分析

### 5.1 实验数据集

使用的数据集为自建数据集, 来源于某电网公司在多次巡检过程中通过红外成像仪采集的红外图像数据, 包括绝缘子、套管、隔离开关、断路器、避雷器、电压互感器、电流互感器、变压器共 8 类变电站常见设备, 合计 3200 张红外图像, 设备数量如表 2 所示。

表 2 红外图像数据集中的设备数量

Table 2 Number of devices on infrared image dataset

Parameter	Insulator	Bushing	Disconnecter	Breaker	Arrester	Voltage transformer	Current transformer	Transformer
Number	496	451	418	401	405	382	334	313

利用图像标注软件 LabelImg 对原始红外图像中的电力设备进行标注, 过程如图 6 所示。

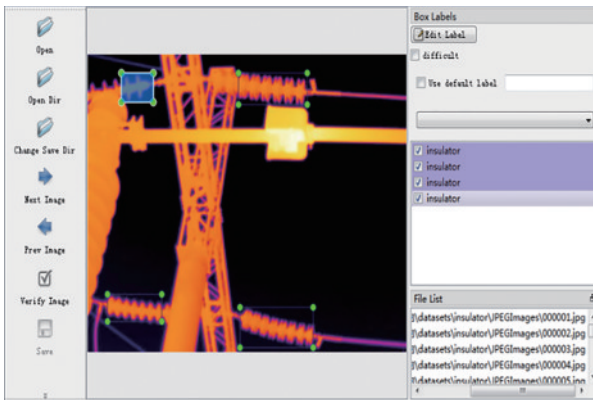


图 6 数据标注过程

Fig. 6 Data annotation process

为了较好地评估模型性能, 将标注好的红外数据按照 3:1 随机划分训练集和测试集。此外, 针对红外图像数据集中多类电力设备样本数量较少导致模型容易出现过拟合的问题, 通过数据增强方式, 对已标注好的红外图像随机进行翻转、放缩等操作, 以扩充数据集样本数, 增强模型的鲁棒性。

### 5.2 实验环境

实验均在 Ubuntu16.04 系统下进行, GPU 型号为 RTX 2080Ti, 实验环境配置为 CUDA9.0+anaconda3+python3.6+pytorch-1.12.0。模型训练过程采用小批量随机梯度下降法, batch\_size 设置为 20, 动量设置为 0.9, 网络初始训练学习率为  $1 \times 10^{-3}$ , 迭代次数为 10000。网络训练过程中, 损失下降曲线如图 7 所示。

由图 7 可以看出, 训练初期网络的损失值下降较

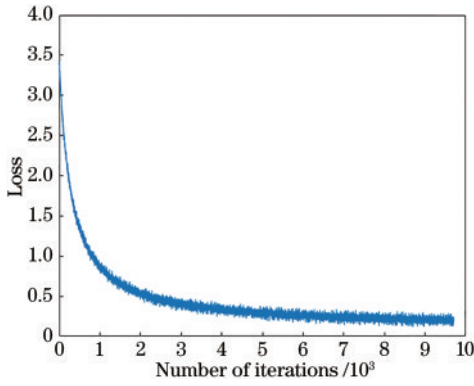


图 7 损失下降曲线

Fig. 7 Loss decline curve

快,迭代到 5000 次左右时,损失曲线逐渐趋于平稳,最终收敛在 0.2 左右,说明所提网络的训练结果较为理想。

### 5.3 评价指标

为了综合评估模型性能,采用每个类别的平均精度(AP)、平均精度均值(mAP)以及检测速度作为衡量模型的评估指标。它们的计算方式分别为

$$P_{\text{precision}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FP}}}, \quad (12)$$

$$R_{\text{recall}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}}, \quad (13)$$

$$P_{\text{AP}} = \int_0^1 P_{\text{precision}}(R_{\text{recall}}) dR_{\text{recall}}, \quad (14)$$

表 4 不同方法的性能对比

Table 4 Performance comparison of different methods

Method	Raw infrared image		Image enhanced by MSRCR algorithm	
	mAP@0.5 / %	Speed / (frame · s <sup>-1</sup> )	mAP@0.5 / %	Speed / (frame · s <sup>-1</sup> )
Faster R-CNN	93.13	16	94.03	17
SSD	90.13	33	91.37	35
YOLOv3	91.54	65	92.71	68
YOLOv4	92.73	74	93.69	75
Proposed method	95.12	69	96.31	71

从表 4 可以看出:相较于对原始红外图像直接进行识别,采用经 MSRCR 算法增强后的图像作为模型输入能得到更好的检测效果;不同方法对 8 种电力设备均具备较好的检测效果;相较于原 YOLOv4 网络,所提方法的平均检测精度均值提高了 2.62 个百分点,表明所提方法在电力设备红外图像识别中具有一定应用价值;在检测速度方面,由于加入多尺度特征的提取模块,所提方法的检测速度略有下降,但相比其他方法,明显具备优势。

网络在训练阶段通过数据增强方式对数量较少样本进行了扩充处理,使得数据集中不同类别的样本数量相对均衡,导致实验精度对 Focal loss 函数中的调制参数  $\beta$  不敏感,故进行  $\gamma$  由小到大的实验。表 5 显示了

$$P_{\text{mAP}} = \frac{\sum_{k=1}^K P_{\text{AP}}(k)}{K}, \quad (15)$$

式中:  $P_{\text{precision}}$  表示精确度;  $R_{\text{recall}}$  表示召回率;  $N_{\text{TP}}$  表示正样本被判断为正确样本的数目;  $N_{\text{FN}}$  表示正样本被判断为错误样本的数目;  $N_{\text{FP}}$  表示负样本被判断为正确样本的数目;  $K$  表示样本总类别数。模型的 IoU 阈值设置为 0.5。

### 5.4 结果与分析

为验证所提方法的有效性,采用 Faster R-CNN、YOLOv3、SSD、YOLOv4 作为对比方法,并在自建数据集上分别对原始红外图像和经 MSRCR 算法增强后的图像进行测试,具体结果如表 3 和表 4 所示。

表 3 所提方法对不同电力设备的测试结果

Table 3 Test results of the proposed method on different power equipments

Electric power equipment	AP / %	Speed / (frame · s <sup>-1</sup> )
Insulator	97.15	71
Bushing	96.57	
Disconnecter	97.83	
Breaker	98.23	
Arrester	96.51	
Voltage transformer	94.62	
Current transformer	96.35	
Transformer	93.24	
mAP / %	96.31	

不同的调节参数下的实验精度对比。

Focal loss 函数中的两个参数  $\beta$  和  $\gamma$  相互协调,经

表 5 在 Focal loss 的不同参数下的实验精度对比

Table 5 Comparison of experimental accuracy under different parameters of Focal loss

Modulation parameter	mAP / %
$\beta=0.5, \gamma=1$	92.63
$\beta=0.7, \gamma=1$	92.66
$\beta=0.9, \gamma=1$	92.82
$\beta=0.9, \gamma=2$	94.73
$\beta=0.9, \gamma=3$	96.31
$\beta=0.9, \gamma=4$	95.26
$\beta=0.9, \gamma=5$	94.67
$\beta=0.7, \gamma=3$	96.18



实验验证,参数值确定为 $\beta = 0.9, \gamma = 3$ 。

图 8 展示了所提方法在部分测试集上的检测效果

图,可以看出,电力设备均能被很好地检测出来,且检测置信度较高。

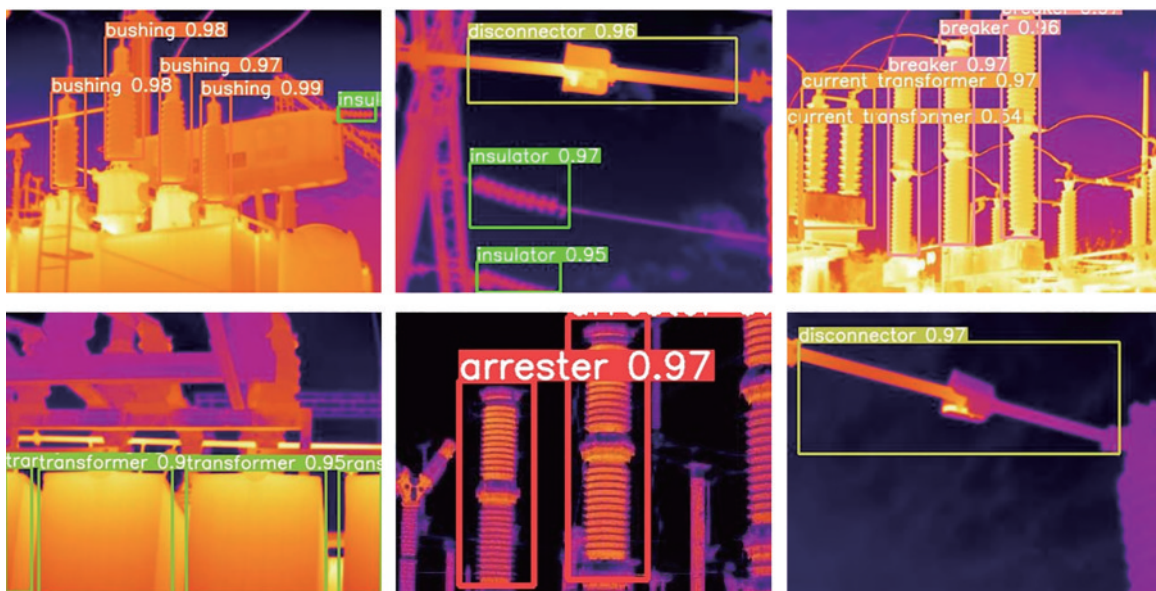


图 8 所提方法对电力设备红外图像的识别结果

Fig. 8 Infrared image recognition result of the proposed method for power equipments

## 6 结 论

提出了一种基于改进 YOLOv4 的电力设备红外图像识别方法。首先对原始红外图像进行 MSRCR 图像增强处理,再将处理后的图像作为模型输入,以提高模型在雨雾天气下对电力设备的检测能力;针对红外图像中目标设备尺度差异性较大造成尺度特征提取困难的问题,提出一种改进的多尺度卷积模块,通过多列结构扩大特征感受野,增强模型的多尺度信息聚合能力;为进一步提升检测精度,采用 Focal loss 函数,解决了红外图像数据不均衡导致的分类困难问题。为使模型得到充分训练,实验部分采用数据增强来扩展训练集。实验结果表明:所提方法对 8 类电力设备的识别精度均高于 93%,具备较好的目标检测效果;模型的平均识别精度为 96.31%,优于 SSD、YOLOv3、YOLOv4 等对比方法;同时检测速度为 71 frame/s,能够快速检测目标所在位置。

## 参 考 文 献

- [1] 刘喜梅, 马俊杰. 泛在电力物联网在电力设备状态监测中的应用[J]. 电力系统保护与控制, 2020, 48(14): 69-75. Liu X M, Ma J J. Application of the ubiquitous power Internet of Things in state monitoring of power equipment [J]. Power System Protection and Control, 2020, 48(14): 69-75.
- [2] Zhao Z B, Xu G Z, Qi Y C. Representation of binary feature pooling for detection of insulator strings in infrared images[J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2016, 23(5): 2858-2866.
- [3] Hu F K, Chen H B, Wang X F. An intuitionistic kernel-based fuzzy C-means clustering algorithm with local information for power equipment image segmentation[J]. IEEE Access, 2020, 8: 4500-4514.
- [4] 姚建刚, 付鹏, 李唐兵, 等. 基于红外图像的绝缘子串自动提取和状态识别[J]. 湖南大学学报(自然科学版), 2015, 42(2): 74-80. Yao J G, Fu P, Li T B, et al. Algorithm research of automatically extracting the area of insulator from infrared image and state identification[J]. Journal of Hunan University (Natural Sciences), 2015, 42(2): 74-80.
- [5] 赵洪山, 张则言, 孟航, 等. 基于高压绝缘套管纹理特征的红外目标检测[J]. 红外技术, 2021, 43(3): 258-265. Zhao H S, Zhang Z Y, Meng H, et al. Infrared target detection of high voltage insulation bushing based on textural features[J]. Infrared Technology, 2021, 43(3): 258-265.
- [6] Zhao Z B, Fan X Q, Xu G Z, et al. Aggregating deep convolutional feature maps for insulator detection in infrared images[J]. IEEE Access, 2017, 5: 21831-21839.
- [7] 汪权, 易本顺. 基于 Gaussian YOLOv3 的航拍图像绝缘子缺陷识别[J]. 激光与光电子学进展, 2021, 58(12): 1210022. Wang Q, Yi B S. Insulator defect recognition in aerial images based on Gaussian YOLOv3[J]. Laser & Optoelectronics Progress, 2021, 58(12): 1210022.
- [8] 徐凯, 梁志坚, 张懿议, 等. 基于 GoogLeNet Inception-V3 模型的电力设备图像识别[J]. 高压电器, 2020, 56(9): 129-135, 143. Xu K, Liang Z J, Zhang Y Y, et al. Image recognition of electric equipment based on GoogLeNet Inception-V3 model[J]. High Voltage Apparatus, 2020, 56(9): 129-135, 143.
- [9] Liu Z Q, Wang H F. Automatic detection of transformer

- components in inspection images based on improved Faster R-CNN[J]. *Energies*, 2018, 11(12): 3496.
- [10] 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.
- [11] 王旭红, 李浩, 樊绍胜, 等. 基于改进 SSD 的电力设备红外图像异常自动检测方法[J]. *电工技术学报*, 2020, 35(S1): 302-310.  
Wang X H, Li H, Fan S S, et al. Infrared image anomaly automatic detection method for power equipment based on improved single shot multi box detection[J]. *Transactions of China Electrotechnical Society*, 2020, 35(S1): 302-310.
- [12] Liu W, Anguelov D, Erhan D, et al. SSD: single shot MultiBox detector[M]//Leibe B, Matas J, Sebe N, et al. *Computer vision-ECCV 2016. Lecture notes in computer science*. Cham: Springer, 2016, 9905: 21-37.
- [13] Howard A G, Zhu M L, Chen B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[EB/OL]. (2017-04-17)[2021-04-05]. <https://arxiv.org/abs/1704.04861>.
- [14] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 779-788.
- [15] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6517-6525.
- [16] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08)[2021-04-07]. <https://arxiv.org/abs/1804.02767>.
- [17] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23)[2021-04-05]. <https://arxiv.org/abs/2004.10934>.
- [18] 郑含博, 李金恒, 刘洋, 等. 基于改进 YOLOv3 的电力设备红外目标检测模型[J]. *电工技术学报*, 2021, 36(7): 1389-1398.  
Zheng H B, Li J H, Liu Y, et al. Infrared object detection model for power equipment based on improved YOLOv3[J]. *Transactions of China Electrotechnical Society*, 2021, 36(7): 1389-1398.
- [19] 李彬, 汪诚, 吴静, 等. 改进 YOLOv4 算法的航空发动机部件表面缺陷检测[J]. *激光与光电子学进展*, 2021, 58(14): 1415004.  
Li B, Wang C, Wu J, et al. Surface defect detection of aeroengine components based on improved YOLOv4 algorithm[J]. *Laser & Optoelectronics Progress*, 2021, 58(14): 1415004.
- [20] 来文豪, 周孟然, 胡锋, 等. 基于多光谱成像和改进 YOLO v4 的煤矸石检测[J]. *光学学报*, 2020, 40(24): 2411001.  
Lai W H, Zhou M R, Hu F, et al. Coal gangue detection based on multi-spectral imaging and improved YOLO v4 [J]. *Acta Optica Sinica*, 2020, 40(24): 2411001.
- [21] Land E H, McCann J J. Lightness and Retinex theory[J]. *Journal of the Optical Society of America*, 1971, 61(1): 1-11.
- [22] Rahman Z, Jobson D J, Woodell G A. Multi-scale retinex for color image enhancement[C]//Proceedings of 3rd IEEE International Conference on Image Processing, September 19, 1996, Lausanne, Switzerland. New York: IEEE Press, 1996: 1003-1006.
- [23] Zheng Z H, Wang P, Liu W, et al. Distance-IoU loss: faster and better learning for bounding box regression[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, 34(7): 12993-13000.
- [24] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(2): 318-327.