

# 基于 CBAM 改进 YOLOv7 的电力设备 红外图像分类检测

陈 佳<sup>1,2</sup>, 余成波<sup>1,2</sup>, 王士彬<sup>2,3</sup>, 蒋启超<sup>1,2</sup>, 何 鑫<sup>1,2</sup>, 张 未<sup>1,2</sup>  
(1. 重庆理工大学 电气与工程学院, 重庆 400054; 2. 重庆市能源互联网工程技术研究中心, 重庆 400054;  
3. 国网重庆市电力公司市南供电分公司, 重庆 401336)

**摘要:** 针对复杂环境下电力设备红外图像的深度学习目标检测数据繁杂、检测精度较低等问题, 本文提出一种基于卷积块注意模块 (Convolutional Block Attention Module, CBAM) 改进 YOLOv7 的电力设备红外图像分类算法。首先将已有数据集进行标注, 并按一定比例划分成训练集、验证集以及测试集, 然后在 YOLOv7 的主干网络中引入 CBAM, 使模型能对感兴趣的区域进行强调并抑制无用信息, 其次将划分好的数据集放入改进后的 YOLOv7 进行模型训练, 同时对比了 6 种改进的 YOLOv5s 模型。实验结果表明, 在相同实验条件下改进 YOLOv7 模型优于 YOLOv7 模型、YOLOv5s 模型和基于 YOLOv5s 的 6 种注意力模型。改进 YOLOv7 性能有明显提升, 可实现快速、精准的红外图像分类。  
**关键词:** 电力设备; YOLOv7; 红外图像; CBAM  
中图分类号: TN219      文献标识码: A      文章编号: 1001-8891(2025)01-0072-09

## IR Image Classification and Detection of Power Equipment Based on CBAM Improvement

CHEN Jia<sup>1,2</sup>, YU Chengbo<sup>1,2</sup>, WANG Shibing<sup>2,3</sup>, JIANG Qichao<sup>1,2</sup>, HE Xin<sup>1,2</sup>, ZHANG Wei<sup>1,2</sup>  
(1. School of Electrical and Electronic Engineering, Chongqing University of Technology, Chongqing 400054, China;  
2. Chongqing Energy Internet Engineering Technology Research Center, Chongqing 400054, China;  
3. State Grid Chongqing Electric Power Company Shinan Power Supply Branch, Chongqing 401336, China)

**Abstract:** To address the problems of complicated data and low detection accuracy for deep-learning target detection of IR images of power equipment in complex environments, this study proposes a convolutional block attention module (CBAM) based on YOLOv7 to improve the classification algorithm for IR images of power equipment. First, the existing dataset is labeled and divided into training, validation, and test sets in a certain proportion and then introduced into the backbone network of YOLOv7 to enable the model to emphasize the region of interest and suppress useless information. Second, the divided dataset is put into the improved YOLOv7 for model training, and six improved YOLOv5s models are compared. The experimental results show that the improved YOLOv7 model outperforms YOLOv7, YOLOv5s, and six attention models based on YOLOv5s under the same experimental conditions. The improved YOLOv7 exhibits significantly improved performance and achieves fast and accurate IR image classification.  
**Key words:** power equipment, YOLOv7, infrared image, CBAM

### 0 引言

电力行业逐步朝着自动化与智能化方向发展, 对

智能设备的需求也随之增加。红外热成像仪作为一种测温设备能有效地呈现设备的温度情况, 并反映出其健康状况, 被广泛用于电力设备巡检中。但目前的巡

收稿日期: 2023-09-06; 修订日期: 2023-11-22。  
作者简介: 陈佳 (1996-), 女, 硕士研究生, 研究方向为图像处理。E-mail: 823331255@stu.cqut.edu.cn。  
通信作者: 余成波 (1964-), 男, 教授, 硕士研究生导师, 研究领域为远程测试技术、网络化仪器、信息处理等领域。E-mail: yuchengbo@cqut.edu.cn。  
基金项目: 重庆市自然科学基金面上项目 (cstc2021jcyj-msxmX0251); 重庆市教育委员会科学技术研究计划 (KJQN202101115, KJQN202201157); 重庆理工大学研究生教育高质量发展行动计划资助成果 (gzlxc20233121)。

检方式主要依靠人工对已采集的红外图像进行分析<sup>[1]</sup>, 面对海量数据时, 仅依靠人工难以完成。当前研究者对红外图像分类研究并不多见, 因此有必要提出一种智能化分类方法, 实现对电力设备红外图像快速、精准的分类。

近年来, 基于深度学习智能检测方法具有检测精度高、检测速度快等优点, 应用广泛。赵庆生等人采用SURF算法寻找特征点, 并利用K-means算法处理特征描述算子, 最后将训练集图像的高维特征向量进行机器学习, 实现了对电力设备不同图像的快速分类<sup>[2]</sup>。为了提高电力设备识别的准确率, Jiang等提出一种基于可见光图像深度学习的红外图像识别方法, 提升了电力设备识别准确率<sup>[3]</sup>。吴慧海等人针对电力设备红外图像遮挡、分类不准确等问题, 提出一种基于TA-YOLO的电力设备红外图像检测方法, 通过迁移学习把可见光电力设备的数据集应用于电力设备红外图像中, 对遮挡问题修改了损失函数DIOU\_NMS, 并把Triplet注意力机制引入特征提取网络中, 提高了电力设备红外图像的检测精度和速度, 但该训练模型样本较小, 对于深度学习问题样本需进一步增强<sup>[1]</sup>。王建强等人提出一种基于小样本学习方法Meta-Network的变压器红外图像分类方法研究, 通过微调原理改进了VGG19模型, 增强了小样本情况下的学习效果, 但模型的准确率不理想<sup>[4]</sup>。周可慧等人提出一种改进CNN的电力设备红外图像分类模型构建研究, 对CNN-Alex进行改进, 验证了改进CNN红外图像分类模型的良好性能, 但该模型存在分类不均的问题<sup>[5]</sup>。针对电力设备红外图像数据类别不均衡等问题, 段中兴等人将YOLOv4模型进行改进, 在YOLOv4骨干网络中加入多尺度卷积模块, 并采用Focal loss函数, 解决了电力设备红外图像数据不均衡导致的分类困难问题<sup>[6]</sup>。随后, Cong等人提出一系列改进方法, 实验发现, 改进算法的识别准确率显著提高<sup>[7-8]</sup>。余一聪等人提出基于改进YOLOv7的绝缘子缺失检测方法, 将CBAM引入YOLOv7的头部架构中, 实现了有效区分绝缘子缺失与否的问题。但对于采集图像时外界环境对图像的影响问题, 仍存在研究空间<sup>[9]</sup>。上述国内外研究中, 这些改进方法应用于电力设备红外图像检测均取得了不错的效果, 但如何将小样本与复杂环境下电力设备红外图像的目标检测数据繁杂、检测精度较低的问题结合起来仍是一个亟需解决的问题。

因此, 本文综合以上特点对电力设备红外图像分类情况展开进一步研究。由于YOLOv7模型具有当下目标检测领域较好的检测速度和精度<sup>[10]</sup>, 且YOLOv7算法在进行预处理操作时, 可对图片进行随机的裁

剪、翻转以及图片间的任意拼接与叠加等操作, 所以本文基于优秀的YOLOv7算法展开研究。首先将已收集的数据集在LabelImg中进行标注, 然后按一定比例划分成训练集、验证集以及测试集, 为使模型能对感兴趣的区域进行强调并抑制无用信息, 将CBAM引入YOLOv7模型中, 为实现精准、快速的红外图像分类, 使模型从输入端进入骨干架构网络时进一步筛选出模型感兴趣的区域, 本文将CBAM引入添加在YOLOv7的主干中, 然后将划分好的数据集放入改进后的YOLOv7进行模型训练, 同时对比了6种改进的YOLOv5s模型。实现了对电力设备红外图像的分类研究。

## 1 数据集构建

本文的红外图像数据采集来源于重庆某电网公司红外数据图库。数据集中设备类型共有5种, 分别是绝缘子串 (Insulator String)、隔离开关 (Disconnect Switches)、避雷器 (Lightning arrester)、断路器 (Circuit Breaker)、电流互感器 (Current Transformer), 共计1110张图像。图1为数据集部分典型样本。

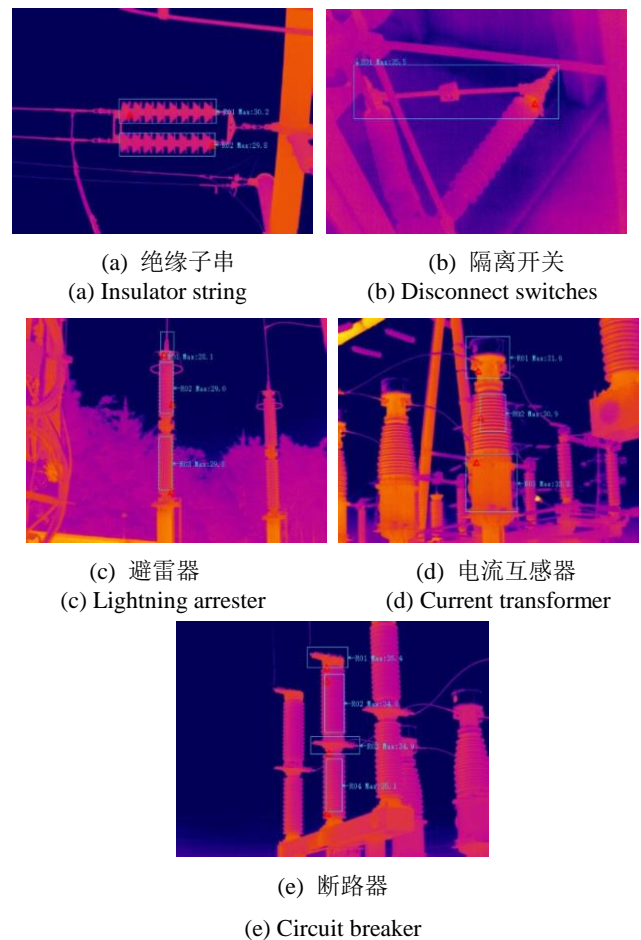


图1 电力设备红外图像数据集示例

Fig.1 Example of an infrared image dataset for power equipments

本文使用图片标注工具（LabelImg）对数据集进行标注，将红外图像中每种类型用矩形框标记（图 2），标记后的红外图像以 XML 文件存储，并将 XML 文件存储作为对应红外图像标签。将该数据集以 8:1:1 的比例划分为训练集、验证集、测试集。其中，训练集 888 张，验证集 111 张，测试集 111 张。

2 模型介绍

2.1 YOLOv7 模型

YOLOv7 是 2022 年 7 月初推出的目标检测算法，相较于以往 YOLO 系列<sup>[11-16]</sup>，YOLOv7 模型在检测速度和精度上存在较好平衡<sup>[17]</sup>。YOLOv7 主要由输入端（Input）、主干架构（Backbone）、头部架构（Head）以及预测端（Detection）组成，其模型结构如图 3 所示。

YOLOv7 的特征提取层由若干 Conv 层、E-ELAN 层以及 MPConv 层组成，其中 Conv 层由卷积层和 BN 层以及激活函数组成。YOLOv7 的头部层通过 SPPCPC 层、若干 Conv 层、MPConv 层、Concat 层

以及后续输出 3 个头部层的 REP 层组成。首先，输入层对输入的图片进行数据增强等一系列预处理操作后，其结果被送入主干架构中，主干架构对处理后的图片进行特征提取，随后，提取到的特征经过特征融合处理得到大、中、小三种尺寸的特征，最后，融合后的特征被送入检测头部，经过检测后输出得到最终结果。

2.2 注意力机制

注意力机制的实质是检测感兴趣的信息并抑制无用信息。3 种主要的注意力机制是基于对特征空间和通道进行加权，如空间注意力机制，通道注意力机制，空间和通道混合注意力机制。

卷积块注意模块（Convolutional Block Attention Module, CBAM）是一种空间和通道混合的注意力机制模块<sup>[18]</sup>，包含通道注意力机制以及空间注意力机制，其中，空间注意的重点是图像中物体的位置，通道注意的重点是图像中的物体。卷积块注意模块结构如图 4 所示，通道注意模块和空间注意模块分别如图 5 和图 6 所示。

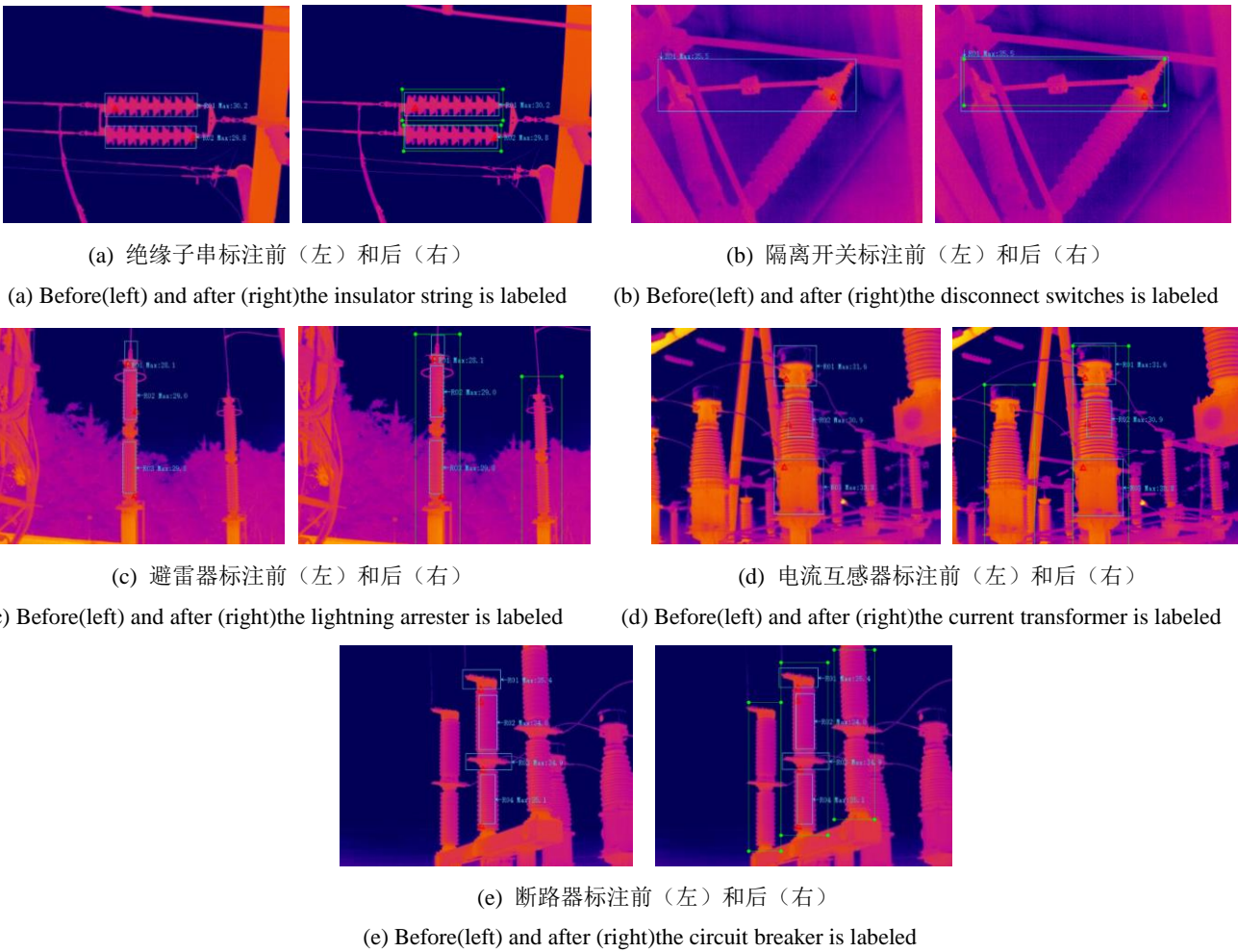


图 2 电力设备红外标注示意图

Fig.2 Infrared labeling of power equipment schematic

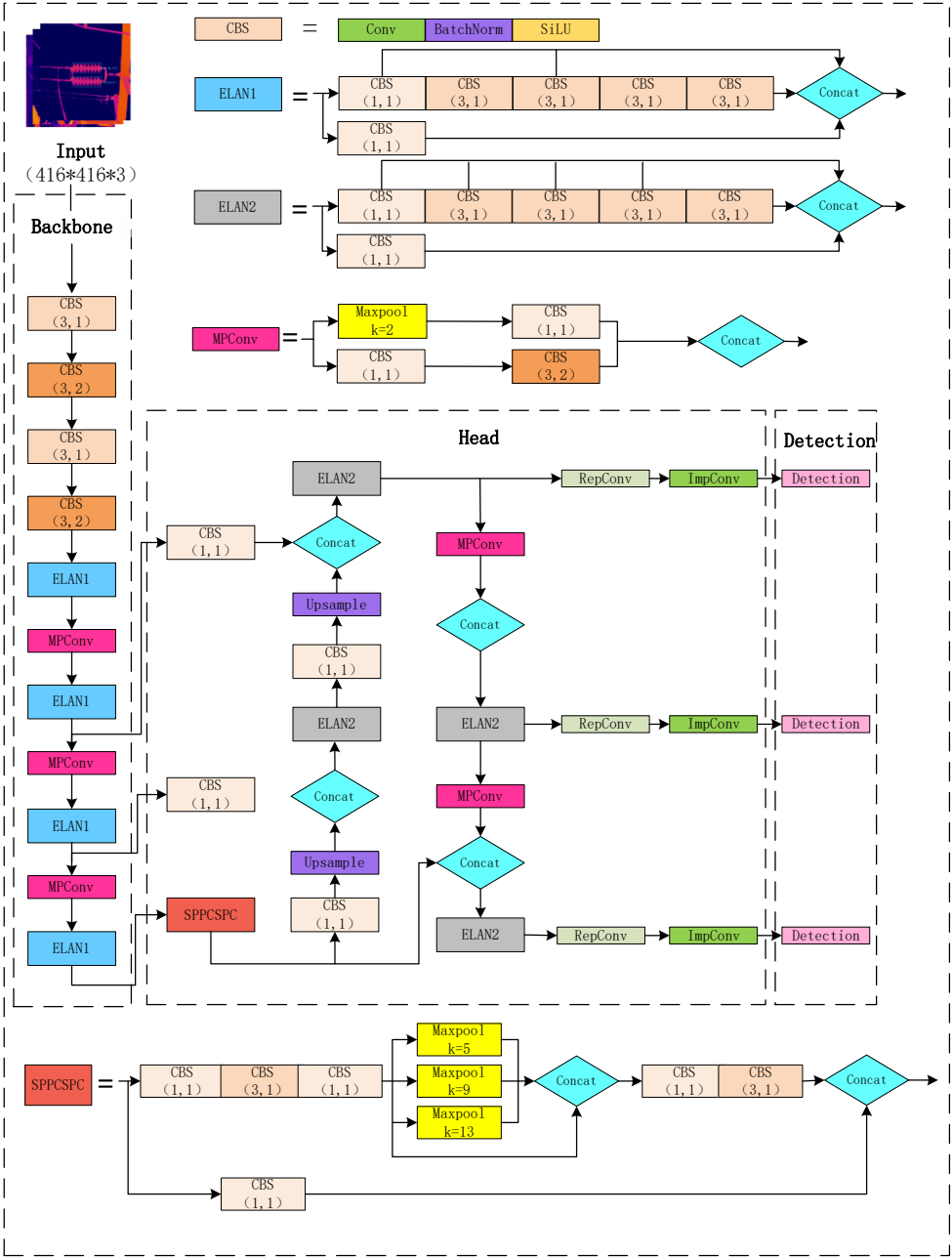


图 3 YOLOv7 模型结构

Fig.3 YOLOv7 model structure

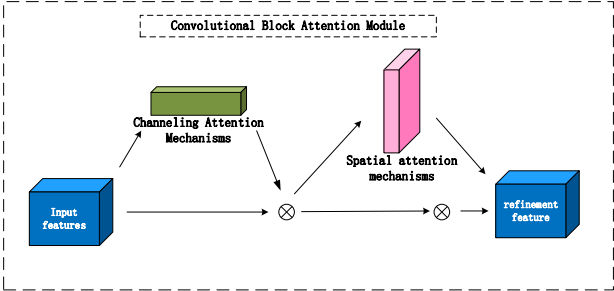


图 4 卷积块注意模块结构

Fig.4 Convolutional block attention module structure

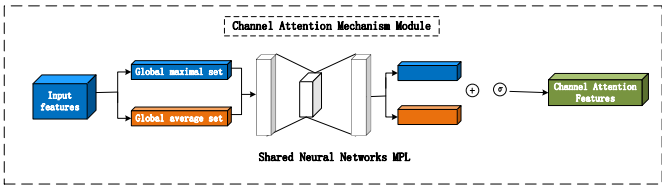


图 5 通道注意力模块结构

Fig.5 Channel attention module structure

具体流程如下：将输入的特征图分别根据宽度和高度进行全局最大集合和全局平均集合，得到两个  $1 \times 1 \times C$  的特征图，然后将其送入两层的神经网络MLP，第一层的神经元数量为  $C/r$ ，其中  $r$  为还原率，激活函数为 ReLU，第二层的神经元数量为  $C$ 。两层



的神经网络共享,通道注意特征和输入特征图按元素相乘,生成最终的通道注意特征。最后,输入特征图被逐级相乘,生成空间注意模块所需的输入特征。随后,在空间注意模块中,该模块首先进行基于通道的全局最大集合和全局平均集合,得到两个特征图。在这两个基于通道的特征图上进行连接操作,然后进行卷积,将维度降低到一个通道,然后通过一个 Sigmoid 来生成一个空间注意力特征。

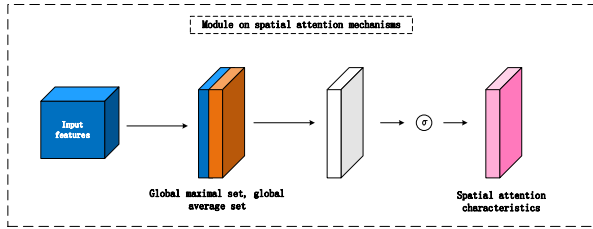


图 6 空间注意力模块

Fig.6 Spatial attention module

### 2.3 改进 YOLOv7 模型

将卷积块注意模块添加到主干之前<sup>[19]</sup>,使模型从输入端进入主干架构网络时首先筛选出模型感兴趣的区域,通过将 CBAM 与 YOLOv7 进行整合,在特征图中分配通道特征和空间特征的权重,以提高检测视觉物体的注意力,抑制对非物体的注意力,从而提高了物体检测模型整体的准确性,改进后 YOLOv7 模型结构如图 7 所示。

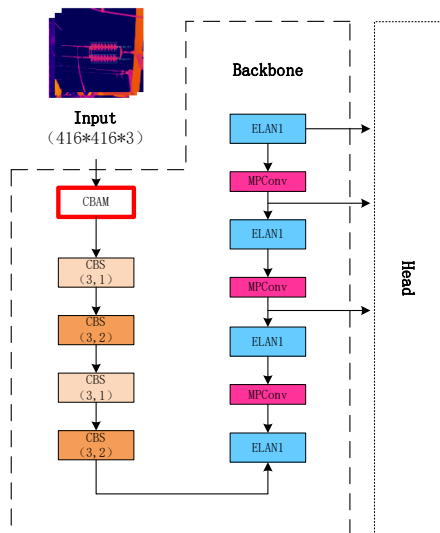


图 7 改进后 YOLOv7 模型中卷积注意模块的位置

Fig.7 Position of the convolutional attention module in the improved YOLOv7 model

将大小为  $416 \times 416 \times 3$  的预处理图像输入到骨干网络中,输出的特征图首先通过 CBAM 中的全局最大集合和全局平均集合进行处理,然后通过共享权重的多层感知器和 Sigmoid 激活函数进行基于两个特征图的加法操作。在通道注意模块完成后,特征图被输

入到空间注意模块。两个特征图通过使用全局最大集合和平均集合进行组合,然后通道的数量减少到  $7 \times 7$  的卷积。应用 Sigmoid 激活函数以获得空间注意力特征图。最后,通过注意模块和空间注意模块的输出相乘,得到 CBAM 的输出特征图。CBAM 的特征图被送入 CBS 模块,最终输出预测结果,以实现模型的物体检测。

## 3 实验与结果分析

### 3.1 实验环境

本次实验硬件方面采用了 i7-11700k 处理器, RTX3090 显卡,软件方面采用了 Win10 操作系统, CUDA11.7, Python3.7 以及 PyTorch1.12;参数设置为输入图像尺寸  $640 \times 640$ , batch size 为 16, epoch 为 300。

### 3.2 评价指标

精度 ( $P$ )、召回率 ( $R$ ) 和平均精度 (mAP) 被作为评价指标。具体公式为:

$$P = \frac{TP}{TP+FP} \quad (1)$$

$$R = \frac{TP}{TP+FN} \quad (2)$$

$$AP_i = \int_0^1 P(R) d(R) \quad (3)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (4)$$

式中:  $P$  为预测正确的结果数量占总预测数量的比值; TP、FP 分别表示正确、错误预测的样本数。FN 为未检测到的正确目标的数目, AP 为 PR 曲线下的面积,表示模型对某个类别识别效果的好坏; mAP 为所有类别的平均准确率的均值,本文采用了两个 mAP 指标,分别为 mAP@0.5 和 mAP@0.95。mAP@0.5 表示将 IOU 分配为 0.5,并对所有类别的图像进行平均。mAP@0.95 则是指不同的 IOU 阈值 (从 0.5 到 0.95,步长为 0.05) 上的平均 mAP。IOU 表示对象预测框和真实框的交叉联合比例。其中:  $N$  表示检测目标类别数量;  $n$  为模型处理图片的张数。

### 3.3 消融实验

为验证改进模型的有效性,引入了 YOLOv5s 六种改进的 YOLOv5s 模型和 YOLOv7 模型与改进 YOLOv7 模型进行对比实验,其中, YOLOv5s 六种改进模型分别是: YOLOv5s\_SE、YOLOv5s\_ECA、YOLOv5s\_CBAM 以及替换后的 YOLOv5s\_C3SE、YOLOv5s\_C3ECA、YOLOv5s\_C3CBAM。

具体操作如下:

1) 在YOLOv5s模型中Backbone中的SPFF模块前插入SE<sup>[20]</sup>、ECA<sup>[21]</sup>、CBAM三种注意力机制, 并将改进后的算法命名为YOLOv5s\_SE、YOLOv5s\_ECA、YOLOv5s\_CBAM。

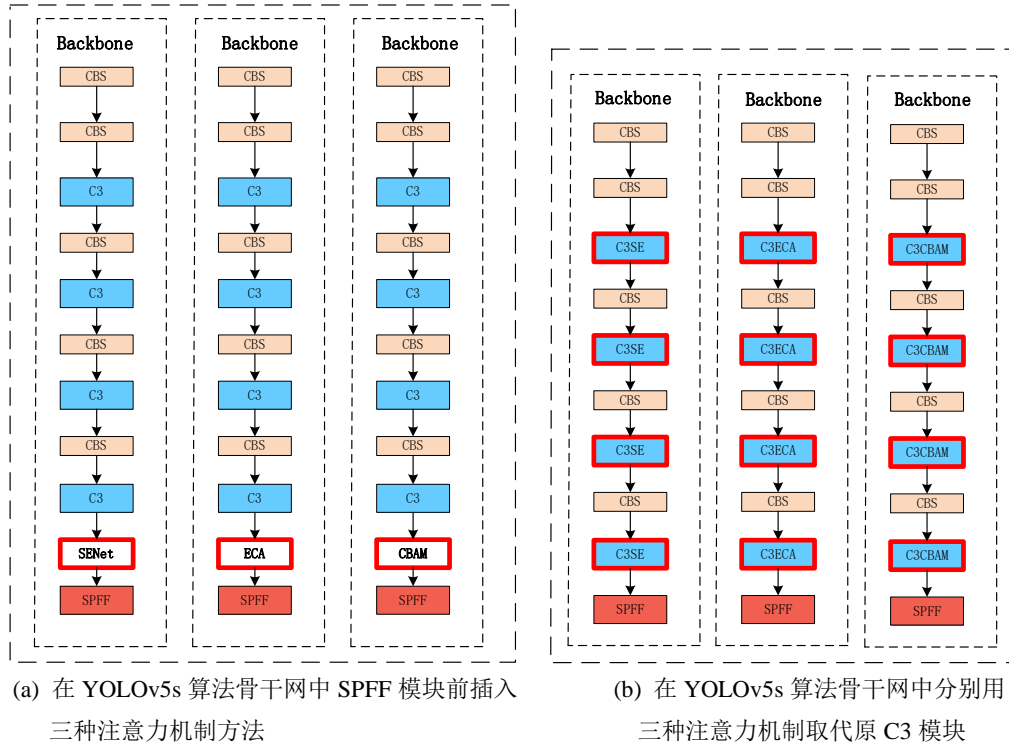
2) 将YOLOv5s原模型中Backbone中的C3模块替换成SE、ECA、CBAM三种注意力机制, 分别将其命名为YOLOv5s\_C3SE、YOLOv5s\_C3ECA、YOLOv5s\_C3CBAM。

3) 将改进的YOLOv7与YOLOv7作对比。

改进的YOLOv5s模型结构图如图8所示, 改进的YOLOv7模型与其他模型实验结果如图9所示, 改进的YOLOv7模型与其他模型测试效果如图10所示,

改进YOLOv7模型与其他模型实验比较结果如表1所示。实验表明, YOLOv7\_CBAM在电力红外图像检测中性能最佳, 比YOLOv5s及其变体更优。加入CBAM的YOLOv7模型精确度提升3.7%, 召回率提升5.9%, mAP@0.5提升1.7%。

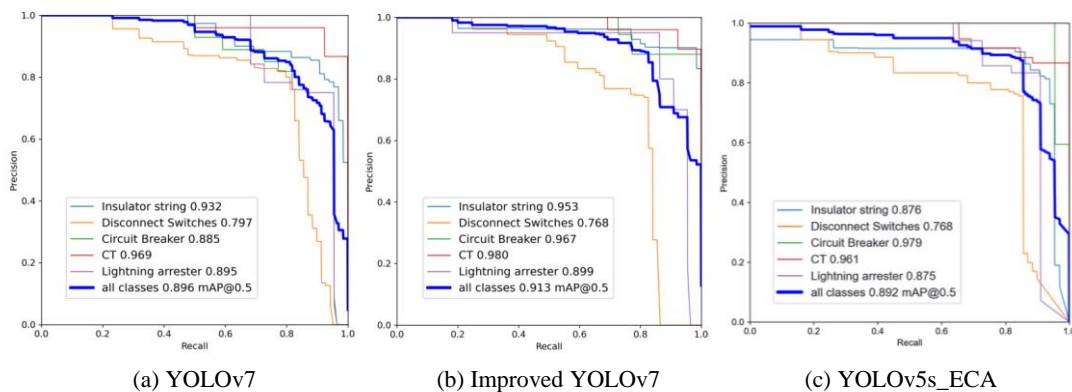
实验发现, YOLOv7\_CBAM模型在相同的实验参数下达到了最佳性能。在本次电力红外图像检测中, YOLOv7和YOLOv7\_CBAM模型的表现优于YOLOv5s和基于YOLOv5s的6种注意力机制的补充模型。与原YOLOv7模型相比, 改进后的YOLOv7模型精确度提升了3.7%, 召回率提升了5.9%, mAP@0.5值提升了1.7%。在YOLOv7模型中加入了CBAM, 可以提高物体的检测性能。



(a) Three attention mechanism methods are inserted in front of the SPFF module in the YOLOv5s algorithm backbone (b) In the YOLOv5s algorithm backbone, three attention mechanisms are used to replace the original C3 module

图8 改进YOLOv5s模型结构图

Fig.8 Improved YOLOv5s model structure diagram



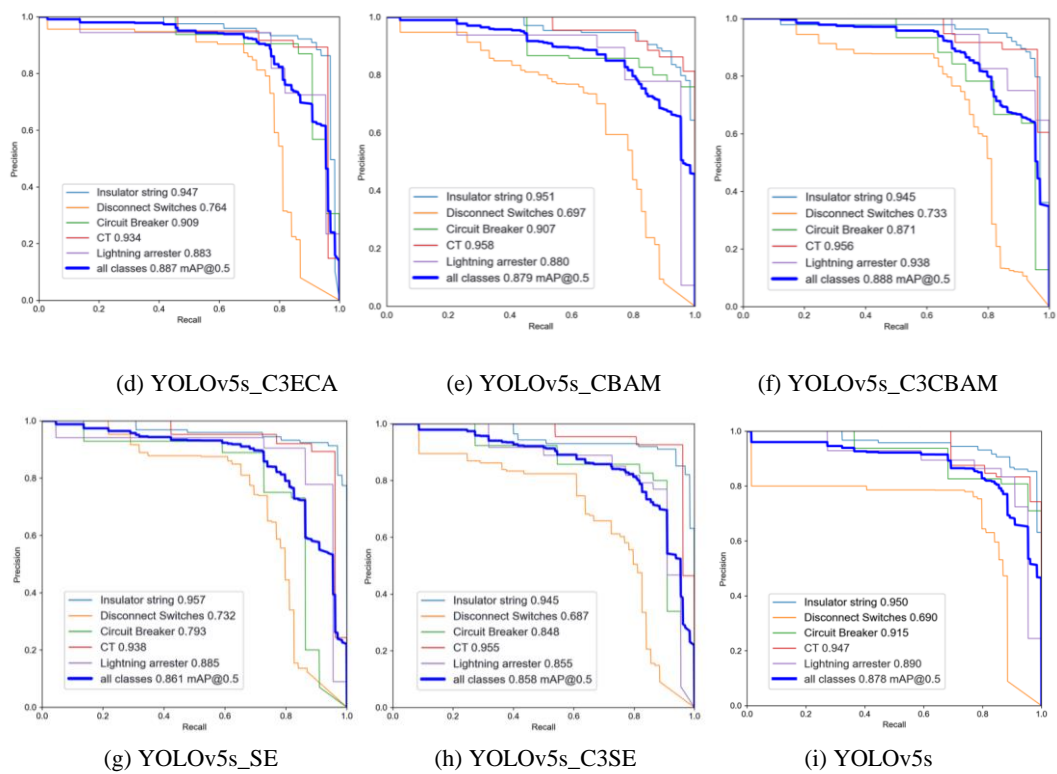


图 9 改进 YOLOv7 模型与其他模型实验结果

Fig.9 Experimental results of improved YOLOv7 model and other models

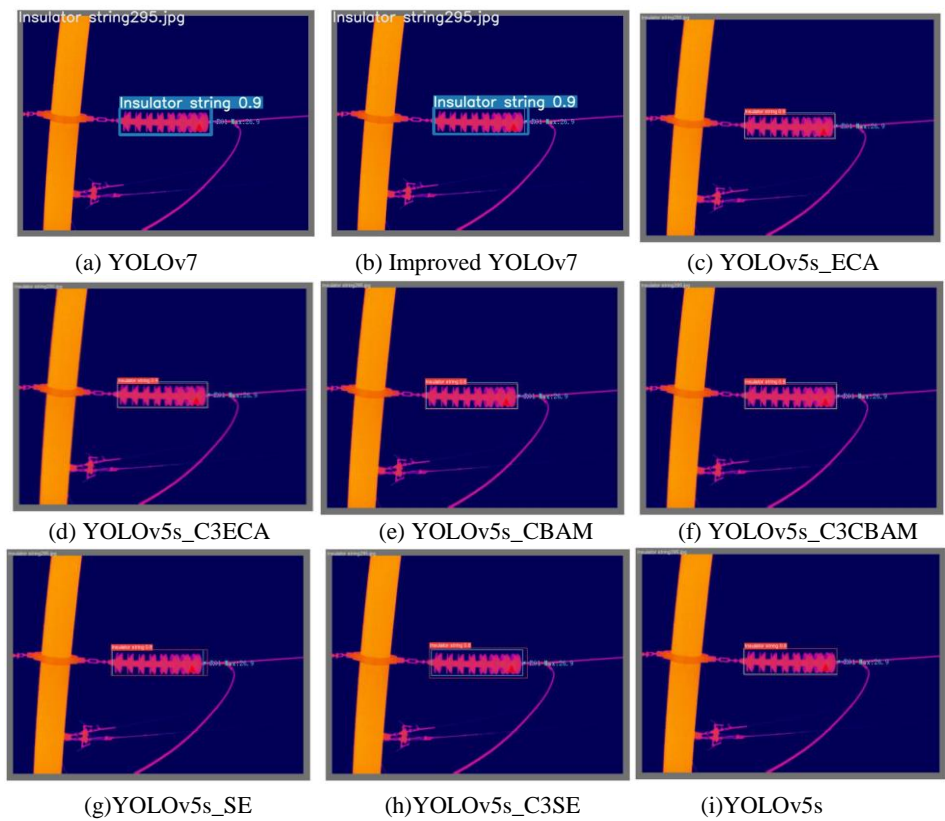


图 10 改进 YOLOv7 模型与其他模型测试效果

Fig.10 Improvement of YOLOv7 model and other model test effect

表1 改进的YOLOv7模型与其他模型实验比较结果

Table 1 Experimental comparison results of improved YOLOv7 model with other models

Models	<i>P</i>	<i>R</i>	mAP@0.5 %
YOLOv5s	79.4	89.7	87.8
YOLOv5s_SE	86.8	80.4	86.1
YOLOv5s_C3SE	83.3	82.9	85.8
YOLOv5s_ECA	85.6	89.1	89.2
YOLOv5s_C3ECA	82.4	87.3	88.7
YOLOv5s_CBAM	77.3	91.7	87.9
YOLOv5s_C3CBAM	82.5	81.9	88.8
YOLOv7	80.5	86.5	89.6
YOLOv7_CBAM	84.2	92.4	91.3

4 结论

为检测复杂环境中的电力设备红外图像，解决电力设备红外图像的准确定位问题，提出一种基于CBAM改进YOLOv7的方法，用于检测复杂环境中的电力设备红外图像，该方法将CBAM机制与原YOLOv7模型相结合，CBAM模块的加入在减少对无用特征关注的同时，也提高了模型对物体的关注。且将CBAM机制加入YOLOv7的主干架构中，在该应用场景下，能首先筛选出模型感兴趣区域，为实现快速、精准红外图像分类奠定基础。实验结果表明，改进后YOLOv7模型在红外图像检测中比原始YOLOv7性能更佳。改进后YOLOv7模型精确度提升了3.7%，召回率提升了5.9%，mAP@0.5值提升了1.7%，mAP@0.95提升了1.4%。本文的未来工作将进一步优化网络模型结构，并提出一个高准确率和轻量级的改进红外图像检测模型，为进一步应用于实际场景提供理论和参考。

参考文献：

[1] 吴慧海, 沈文忠. 基于TA-YOLO的电力设备红外图像检测方法[J]. 信息技术与信息化, 2022(3): 17-20.

WU Huihai, SHEN Wenzhong. TA-YOLO based infrared image detection method for power equipment [J]. *Information Technology and Informatization*, 2022(3): 17-20.

[2] 赵庆生, 王雨滢, 梁定康, 等. 基于BOF图像检索算法的变电站设备图像分类[J]. 激光与光电子学进展, 2020, 57(18): 138-146.

ZHAO Qingsheng, WANG Yuying, LIANG Dingkang, et al. Image classification of substation equipment based on BOF image retrieval algorithm[J]. *Advances in Laser and Optoelectronics*, 2020, 57(18): 138-146.

[3] JIANG A, YAN N, SHEN B, et al. Research on infrared image recognition method of power equipment based on deep learning[C]//2020 IEEE International Conference on High Voltage Engineering and Application (ICHVE), 2020: 1-4.

[4] 裴泽林, 赵曙光, 王建强. 基于Meta-Network的变压器红外图像分类方法研究[J]. 科技风, 2021(5): 87-88.

PEI Zelin, ZHAO Shuguang, WANG Jianqiang. Research on meta-network based transformer infrared image classification method[J]. *Science and Technology Wind*, 2021(5): 87-88.

[5] 周可慧, 廖志伟, 肖异瑶, 等. 基于改进CNN的电力设备红外图像分类模型构建研究[J]. 红外技术, 2019, 41(11): 1033-1038.

ZHOU Kehui, LIAO Zhiwei, XIAO Yiyao, et al. Research on constructing infrared image classification model for power equipment based on improved CNN[J]. *Infrared Technology*, 2019, 41(11): 1033-1038.

[6] 段中兴, 张雨明, 马佳豪. 基于改进YOLOv4的电力设备红外图像识别[J]. 激光与光电子学进展, 2022, 59(24): 2410002-2410002-8.

DUAN Z X, ZHANG Y M, MA J H. Improved YOLOv4-based infrared image recognition for power equipment[J]. *Laser & Optoelectronics Progress*, 2022, 59(24): 2410002-2410002-8.

[7] CONG S, PU H, WANG X, et al. Application of improved YOLOv5 in infrared image recognition of electrical equipment[C]//2023 8th Asia Conference on Power and Electrical Engineering (ACPEE), 2023: 1836-1843.

[8] LI J, XU Y, NIE K, et al. PEDNet: a lightweight detection network of power equipment in infrared image based on YOLOv4-Tiny[J]. *IEEE Transactions on Instrumentation and Measurement*, 2023, 72: 1-12.

[9] 余一聪, 何领朝, 蔡荣贵, 等. 基于改进YOLOv7的绝缘子缺失检测方法[J]. 黑龙江工业学院学报: 综合版, 2023, 23(7): 83-90.

YU Yicong, HE Lingzhao, CAI Ronggui, et al. Insulator missing detection method based on improved YOLOv7[J]. *Journal of Heilongjiang Institute of Technology: General Edition*, 2023, 23(7): 83-90.

[10] 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.

[11] JIANG P, Ergu D, LIU F, et al. A review of Yolo algorithm developments[J]. *Procedia Computer Science*, 2022, 199: 1066-1073.

[12] Roy A M, Bhaduri J, Kumar T, et al. WilDect-YOLO: An efficient and robust computer vision-based accurate object localization model for automated endangered wildlife detection[J]. *Ecological Informatics*, 2023, 75: 101919.

[13] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]//Proceedings of the IEEE Conference on



- Computer Vision and Pattern Recognition*, 2016: 779-788.
- [14] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 7263-7271.
- [15] Redmon J, Farhadi A. YOLOv3: An incremental improvement[J]. *arXiv preprint arXiv:1804.02767*, 2018.
- [16] Bochkovskiy A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection[J]. *arXiv preprint arXiv:2004.10934*, 2020.
- [17] Olorunshola O E, Irhebhude M E, Evwiekpaefe A E. A comparative study of YOLOv5 and YOLOv7 object detection algorithms[J]. *Journal of Computing and Social Informatics*, 2023, **2**(1): 1-12.
- [18] Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module[C]//*Proceedings of the European Conference on Computer Vision (ECCV)*, 2018: 3-19.
- [19] WANG Xiaorong, XU Yan, ZHOU Jianping, et al. Safflower picking recognition in complex environments based on an improved YOLOv7[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2023, **39**(6): 169-176.
- [20] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018: 7132-7141.
- [21] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020: 11534-11542.