

## 基于改进图像识别的绝缘子缺陷检测研究

徐铭申<sup>1,\*</sup>, 章涵睿<sup>2</sup>, 莫千雪<sup>3</sup>, 李啸天<sup>1</sup>, 张腾<sup>1</sup>

1. 华北电力大学, 数理系, 保定 河北, 071000

2. 华北电力大学, 计算机系, 保定 河北, 071000

3. 华北电力大学, 电力工程系, 保定 河北, 071000

**摘要:** 本文提出了一种优化后的 YOLOv5 算法, 旨在识别绝缘子缺陷。通过结合 Transformer 模型、BiFPN (双向特征金字塔网络) 和 CBAM (卷积块注意力模块) 注意力机制, 显著提升了识别的准确度、稳定性和泛化能力。此算法能精确识别和定位各种形态和尺寸的缺陷, 为绝缘子缺陷检测提供了高效且可靠的解决方案。实验数据显示, 优化后的 YOLOv5 算法在绝缘子故障识别中表现出色, 检测准确率高达 76.8%, 相较于传统 YOLOv5 算法, 性能提升超过 20%。在处理复杂绝缘子缺陷和小样本目标识别时, 该算法表现出更强的能力, 在未来具有巨大的应用潜力。

**关键词:** 检测; 深度学习; 绝缘子故障; Yolov5; Transformer; BiFPN; CBAM

## Research on Insulator Defect Detection Based on Edge-Cloud Collaboration and Image Recognition

Mingshen Xu<sup>1,\*</sup>, Hanrui Zhang<sup>2</sup>, Qianxue Mo<sup>3</sup>, Xiaotian Li<sup>1</sup>, Teng Zhang<sup>1</sup>

1. Department of Mathematics and Physics, North China Electric Power University, Baoding 071003,  
Hebei, China.

2. Department of Computer, North China Electric Power University, Baoding 071003, Hebei, China.

3. Department of Electrical Engineering, North China Electric Power University, Baoding 071003, Hebei,  
China.

**Abstract:** This paper proposes an improved YOLOv5 algorithm for insulator defect identification. By integrating Transformer model, BiFPN(Bi-directional Feature pyramid networks) and CBAM(Convolutional Block Attention Module) attention mechanism, The recognition accuracy, robustness and generalization ability are improved. The algorithm can accurately locate and identify defect targets of various shapes and sizes, and provides an efficient and reliable solution for insulator defect detection. The experimental results show that the improved YOLOv5 algorithm in this paper achieves good performance in insulator fault identification task. The detection accuracy rate reached 76.8%, and the experimental performance was improved by more than 20%. Compared with the traditional YOLOv5 algorithm, the improved algorithm has a stronger ability in the identification of complex insulator defects and small sample targets, and has a wider application prospect in the future.

**Keywords:** Target detection; Deep learning; Insulator failure; Yolov5; Transformer; BiFPN; CBAM

在我国的高压输电线路持续大规模的建设过程中, 输电线路的绝缘故障是各输配电企业不得不

去面对的严峻问题。而绝缘子作为输电线路绝缘防护的关键元件，在解决和预防此问题中起到了关键作用。然而，绝缘子长时间暴露在外部环境中，容易受到雷电、污秽、高温和鸟类活动等环境条件的影响，导致其憎水、防污及防闪络等性能下降，从而影响输电线路的绝缘防护安全。因此，对绝缘子的缺陷故障进行及时的检测极其重要<sup>[1]</sup>。

目前，绝缘子的故障检测方法主要分为传统的故障检测方法和基于图像的检测方法，传统的检测方法主要是基于绝缘子的各项指标进行检测，如泄露电流检测法<sup>[2]</sup>、绝缘电阻检测法<sup>[3]</sup>和水性能检测法<sup>[4]</sup>等。赵琳<sup>[5]</sup>等利用时差定位法对绝缘子母线内部放电位置进行精准定位，极大提高盆式绝缘子检测的可靠性。黄继盛<sup>[6]</sup>通过电容分压检测配电网线路故障的暂态电压信号，进而实现配电线故障检测的新方法。师涛<sup>[7]</sup>针对绝缘子内部缺陷检测困难的问题，通过超声波相控阵检测方法实现绝缘子平板硅橡胶和芯棒钻孔的检测，但该方法仅适用于停电作业。可知，传统的绝缘子检测方法能有效的检测绝缘子故障，但由于其普遍是对绝缘子的各项指标进行检测，在检测时需消耗大量劳力成本、且效率低下，易受干扰、准确率不足。

而近几年来，随着图像处理技术的快速发展与应用，其在绝缘子检测方面的作为备受关注。输电线路持续大规模的建设也增加了绝缘子缺陷故障检测的需求，这也成为了图像处理技术研究的重点内容。基于图像的绝缘子缺陷故障方法主要包括传统光学图像检测法<sup>[8]</sup>、红外图像检测法<sup>[9]</sup>和深度学习图像识别检测法<sup>[10]</sup>。传统的光学图像检测是基于可见光图像对绝缘子特征进行直接提取，廖圣<sup>[11]</sup>等利用望远镜、无人机、照相机等通过可见光观测技术对整条线路的绝缘子缺陷故障进行总结，分析故障原因。孙磊等<sup>[12]</sup>通过图像检测法提取多个绝缘子特征描述污秽状况，实现了铁路接触网绝缘子的污秽判别。然而，这两种方法均需要一定的拍摄角度，在输电线路进行大规模的部署有一定困难。也难以实现大规模的应用。徐嘉臻<sup>[13]</sup>提出了盆式绝缘子闪络的典型异物红外图像识别算法，提取红外图像分割后图像的的异物类别特征，迭代训练后，可以识别绝缘子周围典型异物的类别，并减小定了位边界框与异物矩形框的间距，提高了异物目标定位精度和识别效率。但由于是红外图像检测，对检测时的环境温度一定要求。由此，目前基于图像的绝缘子检测方法在图像的拍摄角度和图像类别方面都具备一定限制，难以实现大规模有效的绝缘子检测。

而随着深度学习技术的快速进步，深度学习图像识别检测法在绝缘子缺陷故障检测方面得到了广泛应用。典型的深度学习方法可以分为单阶段和双阶段目标检测<sup>[14]</sup>。Faster Region-based Convolutional Network (Faster R-CNN) 是最具代表性的双阶段检测算法之一，它使用区域建议网络 (Region Proposal Network, RPN) 预测边界框，极大节省了计算资源，达到了接近实时检测的速度<sup>[15]</sup>。贺元帅<sup>[16]</sup>便基于改进的 Faster R-CNN 算法融合 ResNet50 网络解决了现有算法对绝缘子检测精度普遍不高的问题。张庆<sup>[17]</sup>针对无人机拍摄的绝缘子图像存在目标相对背景较小、特征不明显、绝缘子破损区域较小识别难度大等问题，便设计计了一种基于 CNN 的输电线路破损绝缘子识别算法降低上述影响。翟永杰<sup>[18]</sup>针对无人机检测遮挡而存在的漏检问题，在 Faster R-CNN 模型上选择 ResNet101 作为骨干网络，引入 FPN 结构提高对被遮挡的小目标绝缘子的检测精度，极大降低了受遮挡影响的目标的漏检率。虽然双阶段目标检测的精度高但训练资源需求大，难以实际部署。单阶段检测算法无需区域建议，直接预测图像中物体的位置和类别。两个常见例子是 You only look once(YOLO)<sup>[19]</sup> 和 Single Shot MultiBox Detector (SSD)<sup>[20]</sup>快速算法。这两类算法虽然精度相对较低，但是训练速度更快，更适合移动设备部署进行实时检测。王道累等<sup>[21]</sup>针对玻璃绝缘子自爆缺陷问题，在 YOLO v5

模型中采用深度分离卷积和新的骨干网络 L-CSPDarknet53，满足实际巡检需求。同时，不少学者不断改进 YOLO 算法提高其精度：池小波<sup>[22]</sup> 针对现有绝缘子缺失检测模型计算复杂度高和小目标难以检测等问题，在 YOLO v5 模型中移除 C3 模块减少模型参数量并引入卷积块注意力机制来提高复杂背景下的特征提取能力，提出一种基于改进的 YOLOv5s 轻量级检测模型，在提高精度的同时极大地减少了计算量。杨龙欢<sup>[23]</sup> 在 YOLOv5s 的基础上，将浅层的 Fused-MBconv 和深层的 MBconv 模块替代了 YOLOv5s 中的特征提取网络，极大地减小了模型的参数量和计算量。BiFPN 替代原有的 PANet，提升了特征融合能力，保障了检测精度。而本文旨在提升绝缘子检测的准确性，为此我们本文改进了 YOLO v5 的方法。该方法不仅保持了原有的检测速度，还进一步提升了精度。以下是本文对 YOLO v5 进行改进的主要贡献：

- (1) 运用图像增强技术，多角度采集等技术，增强数据集的多样性。
- (2) 在 YOLO v5 网络模型的 Backbone 部分，本文整合了轻量级的注意力机制 CBAM，这种机制能够从通道和空间两个维度更有效地捕获特征，从而减少漏检的可能性。
- (3) 针对绝缘子所处背景复杂以及特征不足的问题，通过在 YOLO v5 网络模型中添加 BiFPN 双向特征金字塔网络；同时，同时引入 Transformer 模型以捕捉特征之间的长距离依赖关系，提高模型对绝缘子故障缺陷的检测精度。

## 1 相关技术概述

### 1.1 YOLO v5 算法简介

YOLO (You Only Look Once) 系列算法<sup>[24]</sup> 是一种基于深度模型的目标检测算法。该算法将特征图划分为网格，并一次性检测所有网格中的目标，通过预测目标边界框、定位信度以及各类别的概率分布，从而实现高效的目标检测。

YOLO v5 是 YOLO 系列的第五个版本<sup>[25]</sup>。分为 YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x 四种，具备根据不同的应用领域动态调整网络结构和深度的能力。本文以 Yolov5s 为 baseline 进行改进。

YOLO v5 算法模型的网络结构由输入端、骨干网络(Backbone)、颈部(Neck)以及预测部分(Prediction)组成。其中，骨干网络主要由跨阶段的局部网络 (Cross Stage Partial network, CSP) 和金字塔网络 (Spacial Pyramid Pooling, SPP) 组成，负责输入图像中不同尺度上特征的提取<sup>[26]</sup>。颈部采用路径聚合网络 (Path Aggregation Network, PAN) 进行提取特征的进一步处理，形成不同尺度的特征图像<sup>[27]</sup>。预测部分负责输出，利用卷积神经网络在不同尺度的特征图上进行对应目标的预测。

完成目标检测后，YOLO v5 采用非极大抑制 (NMS) 机制，通过计算不同边界框的重叠程度来筛选多个目标框，并保留置信度最高的边界框。为了避免对同一目标生成多个重复的检测结果，将与该边界框高度重叠的其他边界框剔除。具体结构如图 1 所示：

### 1.2 BiFPN 双向特征金字塔网络

在输电线路中，由于绝缘子体积小巧且特征不明显，同时其检测图像的背景又相当复杂，导致现有的绝缘子检测模型普遍面临检测精度不高的问题。因此，本文加入关于多尺度特征的有效表示和处理模块 BiFPN。图 2 展示了 BiFPN 双向特征金字塔网络的多尺度特征的处理架构，其中 (a) 部分描绘了 FPN<sup>[28]</sup> 的特征处理结构。该方法融合了多尺度的特征，利用复合特征层和更多的图像信息进行预测。并以该想法为基础，(b) 的 PANet<sup>[29]</sup> 在 FPN 的基础之上添加了一个自下而上的路径聚合

网络，使其能够同时考虑特征层上层的语义信息和下层的位置信息。FPN 和 PAN 等特征结构能够被 YOLO 框架有效整合，然而由于二者在不同尺度的贡献不均衡，本文在 YOLO v5 框架中引入更简洁的双向特征金字塔网络结构 (BiFPN)，以提升检测精度。

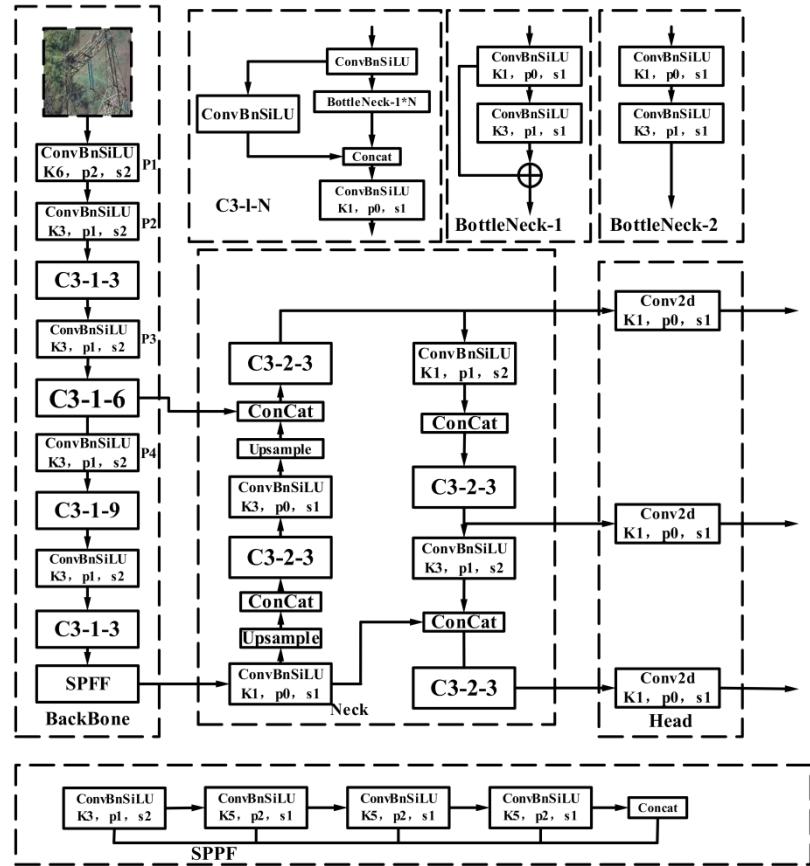


图 1 Yolov5 结构示意图

Fig. 1 Schematic diagram of Yolov5 structure

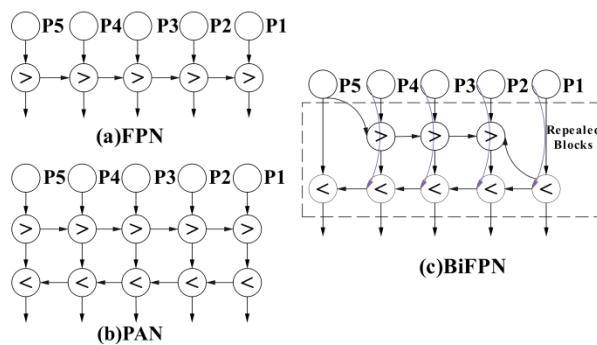


图 2 BiFPN 结构示意图

Fig. 2 Schematic diagram of the BiFPN structure

如图 2 所示，双向特征金字塔网络 (BiFPN) 相当于 FPN 和 PAN 的结构的融合强化，将不同的路径视作独立的特征结构层来处理，并反复进行这一处理过程。通过对不同层次特征进行处理，BiFPN 能够对不同尺度特征的重要性进行区别，并将其进行融合分类。

双向特征金字塔网络 (BiFPN) 对不同尺度特征进行融合分类的方法是归一法, 具体如式(1)所展示。此外, BiFPN 还采用了深度可分离卷积来对不同尺度的特征进行区分融合, 每次操作完成后都会对其进行归一化, 极大提高了深度网络学习模型的检测效率。以第三层为例, 定义两个融合后的特征式如式(1)-(3):

$$O = \sum_i \frac{W_i}{\varepsilon + \sum_j W_j} \bullet I_i \quad (1)$$

$$P_3^{\text{td}} = \text{Conv}\left(\frac{W_1 \bullet P_3^{\text{in}} + W_2 \bullet \text{Resize}(P_4^{\text{in}})}{W_1 + W_2 + \varepsilon}\right) \quad (2)$$

$$P_3^{\text{out}} = \text{Conv}\left(\frac{W_1' \bullet P_3^{\text{in}} + W_2' \bullet P_3^{\text{td}} + W_3' \bullet \text{Resize}(P_2^{\text{out}})}{W_1' + W_2' + W_3' + \varepsilon}\right) \quad (3)$$

式中,  $P_3^{\text{td}}$  表示在第三层中, 从上至下所提取的中间特征,  $P_3^{\text{out}}$  则表示在第六层中, 由下至上提取的输出特征。而  $W_i$  则是权重, 通过 ReLU 激活函数, 为确保数值的稳定性, 还设定了一个极小的正值  $\varepsilon$ 。借助不同层级间的相互关联与结合, BiFPN 成功实现了不同尺度的迅速交叉融合。鉴于上述优势, 本文将双向特征金字塔网络(BiFPN)融入 YOLOv5 架构后, 对特征金字塔结构进行了优化, 加强了不同尺度特征之间的融合, 从而显著提升模型的检测精度。

### 1.3 Transformer 模型

Transformer 是一种基于注意力机制的神经网络架构, 由 Vaswani 在 2017 首次提出<sup>[30]</sup>, 最初在自然语言处理 (natural language processing, NLP) 中取得了巨大的成功, 由此受到启发, 图像领域的学者尝试将其应用到计算机视觉任务中, 由 Alexey 等<sup>[21]</sup>提出的用于图像分类的 ViT(Vision Transformer)模型, 在各大视觉任务上都发挥了出色的性能。

Transformer 模型由编码器和解码器组成, 其编码器包括 L 个 Transformer 模块, 每个模块含有 多头注意力机制 (MSA) 和多层次感知机层 (MLP), 此外还包括层归一化 (LN), 位置编码, 残差链接等技术, 结构如图 3 所示。

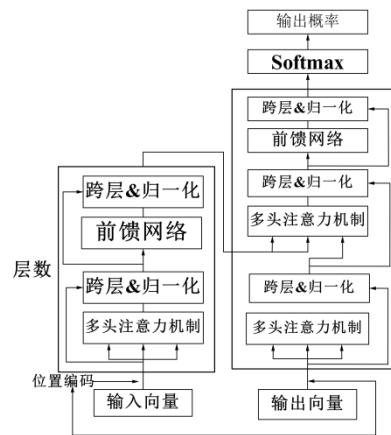


图 3 Transformer 结构示意图  
Fig. 3 Schematic diagram of Transformer structure

在接收到输入图像时, Transformer 先将其转化为一维序列, 随后通过位置编码给每个元素加上一个有顺序特征的位置向量, 使模型可以处理序列中的顺序关系, 通常位置向量由公式(4)-(5)表示:

$$PE(p, 2i) = \sin(p / 10000^{2i/d}) \quad (4)$$

$$PE(p, 2i+1) = \cos(p / 10000^{2i/d}) \quad (5)$$

其中 P 表示元素在序列中的位置，d 表示位置编码的维度，2i 表示偶数的维度，2i+1 表示奇数维度。

Transformer 中注意力机制可通过学习序列之间关系，有效捕捉全局图像之间联系，编码的表示通过多头注意力块传递，其输出结果如下表示：

$$Multi-Head(Q, K, V) = \text{Concat}(\text{head}_1, \dots, \text{head}_n)W^O \quad (6)$$

$$\text{head}_n = \text{Attention}(Q_n, K_n, V_n) \quad (7)$$

其中 Q, K, V 是输入序列分别与 WQ, WK, WV 相乘得到结果，WQ, WK, WV 是可训练参数矩阵， $\text{Concat}(\text{head}_1, \dots, \text{head}_n)W^O$  是拼接操作，在输入下一层时，需将 n 个输出拼接在一起，然后乘以 WO，将维度降低至所需， $\text{Attention}(Q_n, K_n, V_n)$  为注意力机制。

最终，获得的结果与经过残差链接和层归一化，在所有向量之间共享。除了输入层添加了掩码多头注意力机制之外，解码器的结构和编码器完全相同。

#### 1.4 CBAM 注意力机制

为解决 YOLO v5 在绝缘子目标检测过程中检测精度较低的问题，引入了结合通道注意力机制（CAM）和空间注意力机制（SAM）的一种用于前馈卷积神经网络的简单而有效的注意力模块--CBAM，区别于 CAM 仅对不同通道和 SAM 对不同空间提升关注<sup>[3]</sup>，CBAM 从通道和空间上两个方面更关注目标重要的特征，使二者有效结合，其结构如图 4 所示。

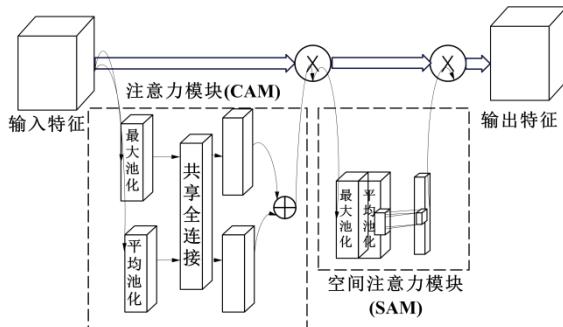


图 4 CBAM 结构示意图  
Fig. 4 Schematic diagram of CBAM structure

由图 4 可知，输入特征会先得到一个通道注意映射，随后被输入到下一模块并进行加权运算得到最终特征。具体如式 (8)-(9)所示：

$$F' = Mc(F) \oplus F \quad (8)$$

$$F'' = Ms(F') \oplus F' \quad (9)$$

在通道注意力模块进行全局最大池化（MaxPool）和全局平均池化（AvgPool）提取特征，之后对结果利用多层感知器(MLP)进行处理，得到最终注意力向量，最后，通过 Sigmoid 函数 ( $\sigma$ )，产生通道注意力权重，将得到的权重与原始特征图每个通道相乘，作为空间注意力模块的输入特征，具体表达式如式 (10)：

$$Mc(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \quad (10)$$

空间注意力模块在每个特征点的通道上分别进行全局最大池化和全局平均池化，再将两结果叠加，再通过卷积层处理，生成空间注意力权重，最后应后对每个空间位置的特征进行加权，得到最终输出特征，具体表达式如式 (11)：

$$Ms(F') = \sigma(f^{7*7}([AvgPool(F'); MaxPool(F')])) \quad (11)$$

其中， $f^{7*7}$  表示卷积核大小为  $7*7$  的卷积操作。

利用 CBAM 对 CAM 和 SAM 的整合，能在不同维度上捕获特征之间相关性，在目标检测中保留更多特征信息，提高模型检测速度和精度。

### 1.5 改进后网络模型结构

由于 CBAM 模块的特性使其在 YOLO 网络架构中的效果与其插入位置有关。本文将 CBAM 模块插入到 YOLOv5 架构的 C3-1 模块中的特定位置（如图 5 中的 C3-1-N 所示），并将其置于 Backbone 结构的最后一个 C3-1 与 SPPF 层之间，从而构建了 YOLOv5s-CBAM 模型。同时，鉴于 YOLOv5 对复杂环境背景下的绝缘子等小目标的检测存在漏检的问题，本文选择引入 Transformer 模型对 YOLOv5 架构的主干网络进行优化，增强对全局上下文信息的处理能力并提高计算效率。Transformer 模型对长距离信息具有较强的处理能力，将其如图 5 所示引入 CBAM 模块前方，极大增强了 YOLOv5 对全局信息的建模能力，能更好地提取复杂背景图像的重要特征，极大降低了绝缘子等小目标的漏检率，并由于其位于 CBAM 的前方，两者得到更加充分的利用，极大降低了运算复杂度，并提高了模型的泛化能力。基于原始 YOLOv5 结构，本文还引入了 BiFPN（双向特征金字塔网络）模块，进一步优化了特征融合过程。BiFPN 模块被插入到 Neck 部分，用于在不同尺度的特征图之间进行有效的信息融合和传递，从而增强了对不同尺度目标的检测能力。具体改进结构如图 5 所示：

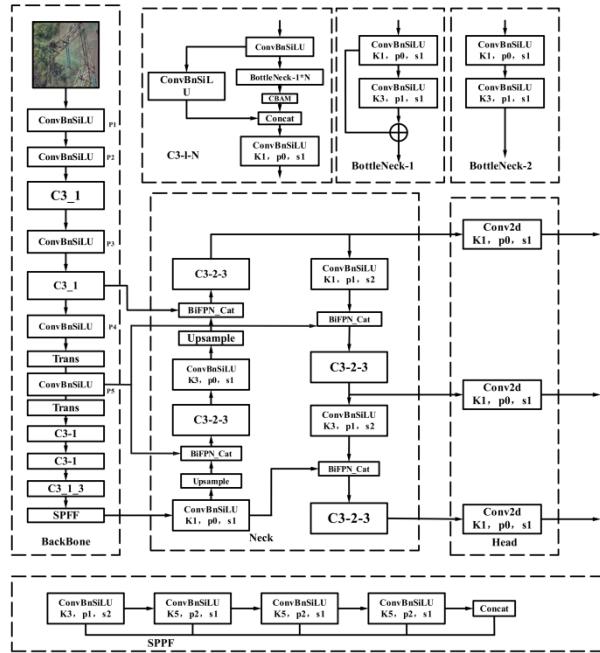


图 5 改进后 Yolov5 结构

Fig. 5 Schematic diagram of Improved YOLOv5 structure

## 2 实验准备与数据集

本文使用了 2000 张原始数据集，其中包含大小各异的绝缘故障图像，每张图像的背景也各不相同，而单帧图像尺寸大小为  $1280 \times 1280$ 。图像种类及数量如下：



图 6 实验所用数据集  
Fig. 6 The dataset used in the experiment

表 1 绝缘子站缺陷识别数据集  
**Table 1 Insulator Station Defect Identification Dataset**

| 类别    | 标签               | 数量  |
|-------|------------------|-----|
| 绝缘子   | insulator        | 609 |
| 防震锤   | hammer           | 713 |
| 防震锤缺陷 | hammer_defect    | 330 |
| 异物    | foreign_body     | 20  |
| 闪络    | discharge        | 14  |
| 鸟巢    | bird_nest        | 99  |
| 绝缘子缺陷 | insulator_defect | 214 |

由上可知，不同类别的被绝缘子在出现频率和取样难度上存在差异，这可能导致样本占比不均衡。为了解决这一问题，本文将原始数据按照 8:1:1 的比例划分为训练集、验证集和测试集，在划分过程中确保各类别的占比保持一致。最终，我们得到了包含 1600 张图片的训练集，以及各包含 200 张图片的验证集和测试集。

### 3 实验结果与对比分析

#### 3.1 模型评估指标

为确保能够对改进后的 YOLO v5 算法进行客观准确的评价<sup>[32]-[34]</sup>，本文采用如式(12)-(14)的指标来评估改进后算法的网络模型：

$$Precision = \frac{TP}{TP + FP} \quad (12)$$

$$Recall = \frac{TP}{TP + FN} \quad (13)$$

$$mAP = \frac{\sum P_A}{N_C} \quad (14)$$

其中， $TP$ ,  $FP$ ,  $FN$  分别代表：预测正确的正确样本数量、预测错误的错检样本数量以及预测错误的正确样本数量。此外，本文用精度来表示网络模型对单一检测目标的准确性，而用平均精度来衡量网络模型中所有检测目标的整体准确性。

同时，本文采用帧率（FPS，即每秒帧数）来衡量网络模型处理图片的速度，并针对部分网络及逆行对比试验。其中，帧率越高，意味着模型在单位时间内能处理的图片数量越多，处理速度及检测速度均更快。此外，本文还通过参数量（Parameter）来评估各网络模型的权重大小，以及使用损失值（Loss）来评估网络模型的训练情况。

### 3.2 实验结果与分析

本文通过选取不同背景下的国网输电线路绝缘子作为检测对象，以验证所提出的改进后 YOLO v5 算法的准确性。实验在 Windows 系统下进行，采用 Intel i9-13900HX CPU、NVIDIA GeForce RTX4060 显卡、8GB 显存、16GB 内存，并使用 Python 3.9 和 PyTorch 1.12.1 版本作为实验环境。

#### 3.2.1 模型参数量与计算量分析

为了更全面、细致地评估本文提出的算法，我们不仅进行了性能上的测试，还深入分析了模型参数规模和计算量这两个关键指标。模型的总参数规模直接关系到其复杂度和所需存储空间，较小的参数规模往往意味着模型更为轻量，这对于嵌入式系统、移动设备以及其他资源受限的应用场景来说至关重要，因为它们通常无法承受大型模型带来的计算和存储负担。此外，我们还特别关注了各模型在推理阶段的浮点操作数（GFLOPs），这一指标能够直观地反映出模型在实际运行时所需的计算资源。通过与传统目标检测模型对比分析，我们能够更准确地了解本文算法在计算效率方面的表现，进而为实际应用中的模型选择提供有力依据。具体的参数规模和计算量对比数据详见表 2，从中我们可以看出本文算法在保持性能的同时，在模型轻量化和计算效率方面也取得了不错的表现。

表 2 参数规模和计算量对比

Table 2 The dataset used in the experiment

| 模型                        | 参数规模    | GFLOP  |
|---------------------------|---------|--------|
| YOLOv5s-BIFPN-Transformer | 50.21M  | 36.9G  |
| CenterNet                 | 62.665M | 69.6G  |
| Faster-RCNN               | 137.77M | 370.4G |

从表 2 的数据中可以清晰地观察到，本文提出的算法在模型参数规模及计算量方面均显著优于 Faster R-CNN 和 CenterNet。综上所述，经过详尽的模型参数和计算量对比，本文提出的方法在模型轻量化和计算效率方面展现出显著优势。特别在资源受限的场景，如移动设备或边缘计算环境中，该方法表现得更为出色，更具实用价值。

图 7 展示了随着迭代次数的增加，准确率、召回率以及平均精确度的变化趋势。其中，mAP.0.5 曲线代表了当 IOU 阈值为 0.5 时的平均精确度变化，而 mAP.0.5:0.95 曲线则反映了 IOU 阈值在 0.5 到 0.95 范围内时平均精确度的变化。通过这些曲线，我们可以清晰地观察到模型在训练过程中的性能提升和收敛情况。

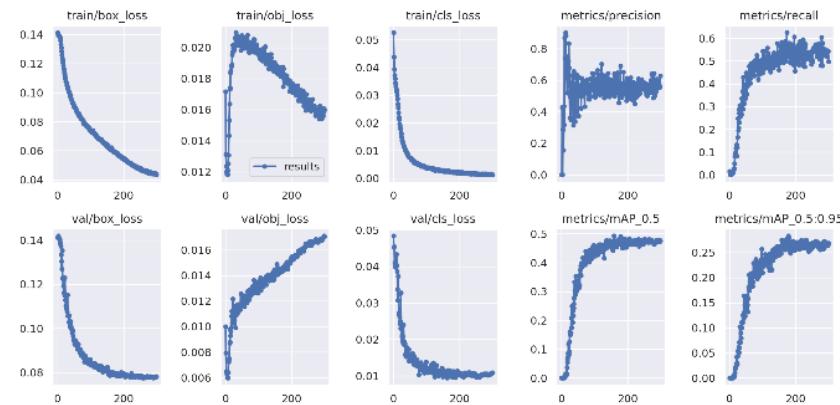


图 7 模型性能可视化

Fig. 7 Visualization of Model Performance

### 3.2.2 对比实验

为了全面评估本文所提出模型的性能，本文进行了对比实验。在实验中，本文以精确率和召回率为评价指标，将 YOLO v5 与其他目标检测模型如 YOLO v4、v7 以及 SSD、EfficientDet 等进行了比较。实验结果如表 3 和表 4 所示。

表 3 不同目标检测模型性能对比

Table 3 Performance comparison of different object detection models

| 类别    | YOLO v5 |       | YOLO v3 |       | YOLO v4 |       | YOLO v7 |       | CENTERNET |      |
|-------|---------|-------|---------|-------|---------|-------|---------|-------|-----------|------|
|       | 精确率     | 召回率   | 精确率     | 召回率   | 精确率     | 召回率   | 精确率     | 召回率   | 精确率       | 召回率  |
| 整体效果  | 0.768   | 0.503 | 0.603   | 0.516 | 0.578   | 0.447 | 0.564   | 0.592 | 0.39      | 0.21 |
| 绝缘子   | 0.357   | 0.624 | 0.356   | 0.56  | 0.331   | 0.554 | 0.371   | 0.655 | 0.62      | 0.16 |
| 防震锤   | 0.391   | 0.603 | 0.374   | 0.562 | 0.343   | 0.539 | 0.347   | 0.612 | 0.44      | 0.02 |
| 防震锤缺陷 | 1       | 0     | 0.184   | 0.107 | 0       | 0     | 0.314   | 0.179 | 0         | 0    |
| 异物    | 0.923   | 0.154 | 0.643   | 0.154 | 0.81    | 0.154 | 0.138   | 0.231 | 0.15      | 0.1  |
| 闪络    | 1       | 0.824 | 0.937   | 0.833 | 0.794   | 0.667 | 0.999   | 1     | 0.1       | 0.01 |
| 鸟巢    | 0.835   | 0.804 | 0.843   | 0.821 | 0.912   | 0.768 | 0.867   | 0.821 | 0.28      | 0.64 |
| 绝缘子缺陷 | 0.861   | 0.514 | 0.884   | 0.575 | 0.857   | 0.449 | 0.915   | 0.646 | 0.06      | 0.42 |

表 4 不同目标检测模型性能对比

Table 4 Performance comparison of different object detection models

| 类别   | YOLO v8 |       | YOLO v7-tiny |       | FASTER-RCNN |      | SSD  |       | efficientdet |       |
|------|---------|-------|--------------|-------|-------------|------|------|-------|--------------|-------|
|      | 精确率     | 召回率   | 精确率          | 召回率   | 精确率         | 召回率  | 精确率  | 召回率   | 精确率          | 召回率   |
| 整体效果 | 0.475   | 0.397 | 0.564        | 0.592 | 0.26        | 0.34 | 0.33 | 0.084 | 0.23         | 0.19  |
| 绝缘子  | 0.335   | 0.507 | 0.371        | 0.655 | 0.37        | 0.64 | 0.21 | 0.073 | 0.32         | 0.18  |
| 防震锤  | 0.35    | 0.346 | 0.347        | 0.612 | 0.25        | 0.12 | 0.14 | 0.092 | 0.24         | 0.096 |

|       |       |       |       |       |      |       |       |       |      |      |
|-------|-------|-------|-------|-------|------|-------|-------|-------|------|------|
| 防震锤缺陷 | 0     | 0     | 0.314 | 0.179 | 0    | 0     | 0     | 0     | 0.06 | 0.15 |
| 异物    | 0.164 | 0.154 | 0.138 | 0.231 | 0.2  | 0.07  | 0.843 | 0.076 | 0.15 | 0.21 |
| 闪络    | 0.703 | 0.667 | 0.999 | 1     | 0    | 0     | 0     | 0     | 0.1  | 0.1  |
| 鸟巢    | 0.828 | 0.714 | 0.867 | 0.821 | 0.28 | 0.645 | 0.667 | 0.107 | 0.38 | 0.19 |
| 绝缘子缺陷 | 0.945 | 0.389 | 0.915 | 0.646 | 0.16 | 0.05  | 0.451 | 0.075 | 0.04 | 0.18 |

上述实验结果表明，YOLO v5 在精确率和召回率上均表现出色。相较于 YOLO v3，YOLO v5 的精确率提升了约 27%（从 0.603 提升至 0.768），召回率也有所提升，这反映了模型在整体目标检测能力上的显著提升。与 YOLO v4 和 YOLO v7 相比，YOLO v5 同样在精确率和召回率上保持了较高的水平。

其次，针对不同类别的目标检测，YOLO v5 也展现出了良好的性能。在绝缘子、防震锤、鸟巢和绝缘子缺陷等类别的检测中，YOLO v5 的精确率和召回率均较高。特别是在与 YOLO v3 的比较中，YOLO v5 在绝缘子类别的精确率上提升了约 0.3%，在防震锤类别的精确率上提升了约 3%。

此外，本文注意到其他模型在不同类别上的表现各有差异。例如，尽管 YOLO v7 在整体效果上稍逊于 YOLO v5，但其在鸟巢和绝缘子缺陷等类别的检测上表现较好。相反，SSD 和 EfficientDet 在整体效果上的表现不佳，这可能是由于它们在特征提取和目标定位方面的能力相对较弱。

通过对实验和数据分析，本段验证了 YOLO v5 模型在目标检测任务上的优越性能。

### 3.2.3 消融实验

为了验证本文改进 YOLO v5 算法后的准确性和优越性，和不同部分对模型的影响，本文在国网数据集上进行消融实验，应本文改进后的 YOLO v5-BiFPN-Transformer-CBAM 为基础模型，逐步减少模块测试网络模型的性能指标，消融实验结果如表 5 所示。

表 5 不同目标检测模型性能对比

Table 5 Performance comparison of different object detection models

| 模型                             | mAP@0.5 | mAP@0.5:0.95 | 准确率   | 回归率   |
|--------------------------------|---------|--------------|-------|-------|
| YOLOv5s-CBAM-BiFPN-Transformer | 0.527   | 0.331        | 0.768 | 0.593 |
| YOLOv5s-CBAM-BiFPN             | 0.482   | 0.275        | 0.554 | 0.542 |
| YOLOv5s-CBAM                   | 0.446   | 0.255        | 0.546 | 0.505 |
| YOLOv5s-Transformer            | 0.481   | 0.273        | 0.575 | 0.561 |
| YOLOv5s-Transformer-Bifpn      | 0.471   | 0.278        | 0.594 | 0.545 |
| YOLOv5s-BiFPN                  | 0.458   | 0.254        | 0.55  | 0.515 |
| YOLOv5s-Transformer-CBAM       | 0.491   | 0.283        | 0.634 | 0.503 |

实验结果表明，通过逐步引入 BiFPN、Transformer 和 CBAM 等模块，模型的 mAP@0.5 从最初的 0.446 逐步提升至 0.527，提升幅度达到了显著的 18.2%。同时，在更严格的 mAP@0.5:0.95 指标上，模型也呈现出一致的性能提升趋势。准确率和回归率的提升幅度虽然相对较小，但仍然分别提

高了约 6.6% 和 8.8%。

深入分析这些提升幅度，我们可以发现，每个模块的引入都对模型性能产生了积极的影响。BiFPN 通过多尺度特征融合提高了模型对不同大小目标的检测能力；Transformer 的引入则增强了模型对目标上下文信息的捕捉能力；而 CBAM 注意力模块的加入使得模型能够更加聚焦于目标区域，从而提高了检测的准确性。这些模块的有机结合共同推动了模型在各项性能指标上的全面提升，充分验证了本文改进策略的有效性。

## 4 结论

本文提出了一种改进版的 YOLOv5 算法，专用于识别绝缘子缺陷。该算法融合了 Transformer 模型、BiFPN 双向特征金字塔网络及 CBAM 注意力机制，从而在识别精度、稳定性和广泛应用性上均表现优异，能精确定位和识别不同形态、尺寸的缺陷。这为变电站的缺陷检测提供了一个既高效又可靠的解决方案。未来，研究者可利用本文的方法，结合数据挖掘，探索更精细的特征提取和目标定位技术。同时，融入专业知识和专家经验，可进一步提高算法的性能和稳定性。深入分析缺陷数据，还能为变电站的维保管理提供科学决策依据，推动预防性维护和精细化管理，确保电力系统的高效和安全运行。

综上所述，本文研究在变电站缺陷识别领域具有重要意义。所提出的改进算法将为实际应用提供有力的技术支持，有助于电力行业的安全和稳定发展。

## 参考文献

- [1] 王德文, 李业东. 基于 WGAN 图片去模糊的绝缘子目标检测[J]. 电力系统保护与控制, 2020, 40(5): 188-198.
- [2] 宁凡, 蒋玉芳. 基于傅里叶变换算法的变电站母线避雷器泄露电流异常检测方法[J]. 电气自动化, 2024, 46(02): 88-90.
- [3] 翁珠奋, 邓永亮, 廖亿峰, 等. 陶瓷绝缘子绝缘电阻检测辅助器在输电线路上的应用[J]. 电子元器件与信息技术, 2020, 4(10): 62-63.
- [4] 孙国华, 孔德瀚, 杨景超, 等. 复合绝缘子憎水性检测技术研究现状[J]. 广西电业, 2023(04): 46-52.
- [5] 赵琳, 邵先军, 王绍安, 等. GIS 盆式绝缘子内部气隙放电检测与诊断[J]. 电力安全技术, 2023, 25(11): 75-78.
- [6] 黄继盛, 刘红文, 唐钊, 等. 基于绝缘子内置电容分压的配电线路故障检测技术研究[J]. 电瓷避雷器, 2023(03): 205-212.
- [7] 师涛. 悬式绝缘子内部人工缺陷的相控阵超声波检测策略研究[J]. 电子制作, 2019(16): 90-92.
- [8] 王小娜. 基于图像特征提取的复合绝缘子粉化状态检测[D]. 华北电力大学(北京), 2024.
- [9] 吴星奇, 付子峰, 郭景武, 等. 基于红外紫外成像检测技术的绝缘子运行状态评估分析[J]. 集成电路应用, 2023, 40(09): 28-30.
- [10] 王磊. 基于深度学习的输电线路绝缘子目标增强识别研究[J]. 铁道建筑技术, 2022, (12): 189-193.
- [11] 廖圣, 蔡伟, 陈俊威, 等. 带电检测及可见光观测技术在配电线路上的应用实例[J]. 电工技术, 2022(15): 109-111.
- [12] 孙磊, 吴文海, 柯坚, 等. 基于可见光图像的铁路接触网绝缘子污秽状况检测[J]. 电瓷避雷器, 2019, 291(5): 240-244.
- [13] 徐嘉臻, 孙国立, 李要锋, 等. 盆式绝缘子闪络的典型异物红外图像识别算法[J]. 信息技术, 2024(02): 160-165.
- [14] L, ZHANG F, LIU F, et al. A survey of Deep Learning-Based Object Detection[J]. IEEE Access, 2019, 7: 128837-128868.
- [15] REN S, HE K, 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.
- [16] 贺元帅, 纪超, 王博雅, 等. 基于改进 Faster R-CNN 的绝缘子缺陷检测识别与定位[J]. 电工技术, 2024(01): 176-179.

- [17] 张庆, 霍婷婷, 苗海东, 等. 基于图像增强与 CNN 的输电线路破损绝缘子识别[J]. 现代信息科技, 2023, 7(19): 151-154+159.
- [18] 翟永杰, 王璐瑶, 郭聪彬. 基于 Faster R-CNN 的复杂背景下绝缘子目标检测[J]. 电子测量技术, 2023, 46(16): 187-194.
- [19] 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 (CVPR). Las Vegas, NV, USA: IEEE, 2016: 779-7882016.
- [20] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot MultiBox Detector[C]//2016 European Conference on Computer Vision (ECCV). Amsterdam, The Netherlands: Springer, 2016: 21-37.
- [21] 池小波, 张伟杰, 贾新春, 等. 基于改进 YOLOv5s 的轻量级绝缘子缺失检测[J]. 测试技术学报, 2024, 38(01): 19-26.
- [22] 杨龙欢, 董效杰, 黄施懿, 等. 基于 YOLOv5 的轻量型绝缘子缺陷检测算法研究[J]. 通信与信息技术, 2023(05): 13-18.
- [23] 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.
- [24] ULTRALYTICS. Ultralytics. yolov5[EB/OL]. 2020 [2022-06-01]. <https://github.com/ultralytics/YOLOv5>.
- [25] 谢政峰. 基于深度学习的新材料地板缺陷检测技术研究[D]. 四川大学, 2021.
- [26] 谷玉海, 曹梦婷, 修嘉芸, 等. 基于 YOLOv4 网络的违章行为检测算法[J]. 重庆理工大学学报(自然科学), 2021, 35(8): 114-121.
- [27] Liu S, Qi L, Qin H, et al. Path aggregation network for instance segmentation [C] // Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 8759-8768.
- [28] 何宇豪, 易明发, 周先存, 等. 基于改进的 Yolov5 的无人机图像小目标检测 [J/OL]. 智能系统学报:1-12[2024-06-13].<http://kns.cnki.net/kcms/detail/23.1538.TP.20231204.1028.002.html>.
- [29] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [EB/OL] . [2017-12-06] . <https://arxiv.org/pdf/1706.03762.pdf>
- [30] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. Transformers for image recognition at scale[A]. 2020.
- [31] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]// ECCV: Proceedings of the European conference on computer vision, LNCS 11211. Cham: Springer, 2018: 3-19.
- [32] 杨乐, 郭一鸣, 霍勇博, 等. 改进 YOLOv5 在电力生产违规穿戴检测中的应用[J]. 电力系统保护与控制, 2023, 51(14): 160-168.
- [33] 张焕龙, 齐企业, 张杰, 等. 基于改进 YOLOv5 的输电线路鸟巢检测方法研究[J]. 电力系统保护与控制, 2023, 51(2): 151-159.
- [34] 施保华, 井任月, 杨超, 等. 用于变电站保护压板状态识别的增强 YOLO 网络[J]. 电力系统保护与控制, 2021, 49(23): 163-170.

<sup>1,\*</sup>第一作者和通讯作者: 徐铭申(2004-), 男, 华北电力大学数理系, 研究方向为深度学习; E-mail: 220221100326@ncepu.edu.cn