

基于改进 SSD 的输电场景绝缘子识别

秦金磊,王兆楠,李 整

(华北电力大学计算机系,河北 保定 071003)

摘要:为了满足输电场景下智能巡检系统构建的准确和快速要求,文中提出了一种基于轻量级单步多框检测器(single shot multibox detector,SSD)网络的输电场景绝缘子目标检测方法。首先,以 SSD 为基础,采用轻量级网络 MobileNetV3-large 改进模型结构,引入通道注意力结构和新的激活函数,增强特征提取能力的同时降低模型复杂度;其次,结合高效通道注意力(efficient channel attention,ECA),有效捕获跨通道交互的信息;最后,构建包含不同背景、不同种类输电线路绝缘子图像数据集,对模型开展训练、测试和验证工作。实验结果表明,与原 SSD 算法相比,对于检测绝缘子图像,模型参数量减少了 82.93%,目标检测速度提升 96.38%,且能达到 92.70%的准确率,为方便部署且保证快速有效地识别提供了一种可行方案。

关键词:SSD;绝缘子;深度学习;注意力机制;目标检测

中图分类号:TP391.41;TM755 **文献标志码:**A **文章编号:**1001-9898(2026)01-0041-09

Recognition of Insulators in Power Transmission Scenarios Based on Improved Single Shot Multibox Detection

QIN Jinlei,WANG Zhaonan,LI Zheng

(North China Electric Power University,Department of Computer,Baoding 071003,China)

Abstract:In order to meet the requirements of accuracy and speed for the construction of intelligent inspection systems in power transmission scenarios,this paper proposes an insulator target detection method for power transmission scenarios based on an improved lightweight Single Shot Multibox Detector (SSD) network.Firstly,based on SSD,the model structure is improved using the lightweight network MobileNetV3-large,introducing a channel attention mechanism and a new activation function,which enhances feature extraction capability while reducing model complexity.Secondly,by integrating Efficient Channel Attention (ECA),information from cross-channel interactions is effectively captured.Finally,a dataset containing images of transmission line insulators with different backgrounds and types is constructed for model training,testing,and validation.Experimental results showed that compared with the original SSD algorithm,for detecting insulator images,the number of model parameters is reduced by 82.93%,the target detection speed was increased by 96.38%,and an accuracy of 92.70% was achieved,providing a feasible solution for convenient deployment while ensuring fast and effective recognition.

Key words:SSD;insulators;deep learning;attention mechanism;object detection

0 引言

绝缘子在电力系统中承担电气绝缘和机械固定的功能,但其长期暴露在户外,容易受到外界环境的影响而损坏,进而给电网的运行带来极大的隐患。因此,定期对绝缘子进行状态检测和故障诊断尤为重要,而状态检测的前提是绝缘子的识

别。近年来,我国电网利用无人机对高压线路进行检查,对拍摄到的图像进行状态的识别和检测,减轻了电力工人工作的强度和危险性。因此,利用图像处理算法识别绝缘子,是当前的优质解法和良好工具。

近年来,深度学习被广泛地应用在绝缘子等电力设备的检测中。龚钢军等^[1]调整卷积神经网络

收稿日期:2025-02-17 修回日期:2025-05-01

作者简介:秦金磊(1979-),男,副教授,研究方向为深度学习、能源互联网等。

基金项目:河北省自然科学基金项目(F2014502081);中央高校基本科研业务费专项资金资助(2020MS120)。

络参数,获得图像识别模型,巡检能力明显优于传统方法。杨达伟等^[2]提出改进 YOLOv7 的变电站电力设备红外图像的多目标检测方法,利用较为流行的 YOLOv7 网络和轻量型注意力机制改善误检、漏检等问题。

基于深度学习的目标检测模型有两个分支:两阶段和单阶段检测模型。两阶段模型将整个过程分为两部分,检测精度高,但耗时太长,难以满足检测的实时性要求,代表算法有区域卷积神经网络(regions with convolutional neural networks, R-CNN)^[3]、Fast R-CNN^[4]、Faster R-CNN^[5]等。胡晓东等^[6]利用 Oriented-RCNN 对单串瓷绝缘子进行方向性检测定位,能够检测出倾斜目标。Faster R-CNN 模型具有强大的特征提取能力以及比传统方法更好的识别精度和效率,特别适合电力设备的图像识别^[7]。程海燕等^[8]应用 Faster R-CNN 框架,识别航拍图像中的绝缘子,对于遮挡条件下的绝缘子识别效果不理想。两阶段方法的准确率相对较高,但在检测速度以及计算量等方面不占优势。与两阶段模型相比,单阶段模型可以实现端到端的检测,检测速度更快,但精度有所降低,代表算法有“你只看一次”(you only look once, YOLO)^[9]、单步多框检测器(single shot multibox detector, SSD)^[10]、CenterNet^[11]等。朱有产等^[12]将改进的 YOLOv3 算法应用于绝缘子检测领域,提升聚类算法效果并增加了尺度预测,仍存在部分漏检、错检问题。这些单阶段模型在检测速度上有明显优势,但相对于两阶段模型精度仍显不足。

为进一步满足移动端及嵌入式设备对检测模型的轻量化和高实时性要求,相关学者开始探索利用轻量级网络结合高效注意力机制来改进目标检测模型。比如, Li 等^[13]指出, MobileNetV2、MobileViT^[14]等现有的轻量化模型,虽然能够降低参数量,但在复杂场景下的精度损失显著,达到 5%~8%。Maaz 等^[15]提出结合 CNN 与 Transformer 的轻量化模型,参数量仅 6.2 M,与 CNN 相比,Transformer 模型通常伴随着更大的计算开销,在轻量化应用方面, CNN 通常更占优势。

SSD 算法的优点在于,作为一个单阶段模型,在保障运行速度的前提下,检测精度可以和 Faster RCNN 媲美。马建宝等^[16]提出基于 SSD 和

ResNet 的光伏板检测算法,加入注意力机制,增强了网络的识别精度,故类似方法也可用于绝缘子串检测中。李伟性等^[17]采用 SSD 卷积神经网络方法检测图像中的绝缘子缺陷,通过调整网络,增强了特征提取能力。王磊^[18]对绝缘子图像进行多种方法的增强扩充,并采用 SSD 检测模型进行图像训练与测试,可以解决绝缘子图像数据不足的问题。以上方法利用 SSD 算法进行绝缘子缺陷的检测,但是缺乏针对检测速度提升的研究。

目前大多数研究提出的目标检测模型对输电场景下的绝缘子目标检测精度较低,且参数量和计算量过大,导致检测时间过长。针对以上问题,结合绝缘子识别的特点,对 SSD 网络进行改进:

1) 采用 MobileNetV3-large 网络作为模型的特征提取网络,使用通道可分离卷积代替常规卷积,大大减少了参数数量;

2) 在网络模型的各个阶段加入高效通道注意力机制(ECA)^[19],通过有效学习通道注意力机制,提升模型对绝缘子目标的检测性能;

3) 在已有数据集的基础上,构建了输电线路绝缘子数据集,包括多种材质、多种背景下的绝缘子图像。基于该数据集的实验结果表明,该方法优于其他目标检测模型,减少了网络参数数量,提高了检测速度,同时保持了较高的识别精度。检测结果可用于进一步的绝缘子状态分析。

本研究的核心思路聚焦于轻量化主干网络与注意力机制协同优化提升 SSD 算法的检测效率。首先基于理论分析构建 MobileNet 与 4 种注意力机制的组合模型,设计多尺度特征强化模块以平衡计算复杂度与特征表达能力。随后通过控制变量实验系统验证不同主干网络与注意力机制组合的性能差异,确定最优基础架构。进一步在 PAS-CAL VOC 数据集上开展对比实验,从检测精度、推理速度、参数量 3 个维度评估改进方案的优越性。所提出的改进架构有效平衡检测精度与计算效率,在不显著增加计算负载的前提下维持了目标检测任务的竞争优势。

1 SSD 算法

SSD 算法使用一个单独的深度神经网络来检测图片中的目标,其核心是在特征图上使用小卷积滤波器来预测一组固定的默认边界框的类别分

数和边界框偏移量。SSD 网络基于前馈神经网络,生成固定大小的边界框集合和边界框中目标的类别得分,最后通过非极大值抑制(non-maximum suppression, NMS)得到最终的预测结果。基础网络采用图像分类的架构,然后向其中添加辅助结构来产生具有关键特征的预测。SSD 算法包括基础网络和辅助结构两部分,具体模型结构见图 1。

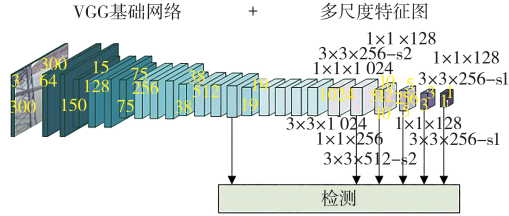


图 1 SSD 模型结构

1.1 SSD 网络结构介绍

SSD 算法包含基础网络和辅助结构两部分,其中辅助结构又包含多尺度特征图、卷积预测器和默认框生成过程 3 个重要特点,SSD 算法的整体流程如图 2 所示,该图共分为输入阶段、特征提取阶段、预测阶段、训练阶段和测试阶段五部分。

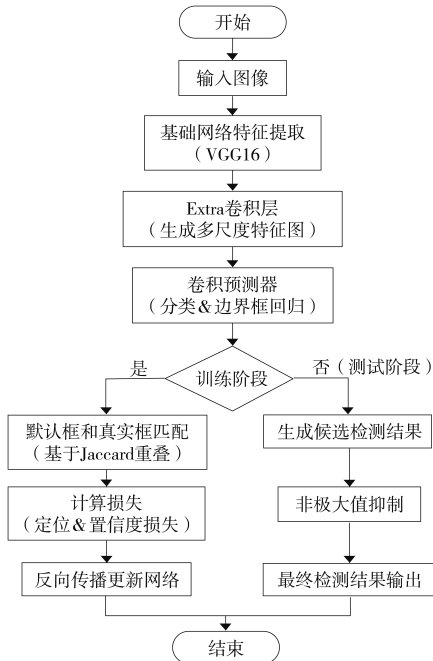


图 2 SSD 算法流程

1.1.1 基础网络

SSD 基于预训练的视觉几何组 (visual geometry group, VGG) 网络^[20],借鉴 DEEPLab-LargeFOV^[21],将 VGG16 中的前两个全连接层替

换为卷积层,并对其参数进行二次采样,更改了最后一个最大池化层的步长,移除了所有的随机失活层和最后一个全连接层,并根据数据集进行了微调。

1.1.2 辅助结构

SSD 网络通过多级特征图实现多尺度检测,其中包含具有 6 个不同尺度特征层的特征金字塔。各层独立应用 3×3 卷积核生成固定的预测结果。假设给定位置有 k 个框,类别数为 c ,对于每个框,计算 c 类的得分和相对于默认框的 4 个偏移量,因此对于特征图的每个位置,应用 $(c+4) \times k$ 个过滤器。在默认框的生成过程中,基于宽高比 $a_r = \{1, 2, 3, 1/2, 1/3\}$ 生成先验框,中心坐标通过特征图网络位置进行计算。

1.2 训练

在 SSD 算法的训练过程中,真实框与默认框的匹配有其独特的策略,为默认框选择合适的尺度和纵横比,最后调整正负比例以完成默认框的生成。

1.2.1 匹配策略

Jaccard index^[22],又称为 Jaccard 相似系数,用于比较有限样本集之间的相似性与差异性。对于每个真实框,选择默认框与其匹配。首先,将每个真实框与具有最佳 Jaccard 重叠的默认框匹配;然后,将默认框与任何 Jaccard 重叠高于阈值的真实框进行匹配。允许网络预测多个重叠默认框的高分,而不是只要求它选择重叠值最大的那个。

1.2.2 样本比例调整

经过上述的匹配策略,大多数的默认框都会被划分为负例,导致了严重的正负样本不平衡问题。为了解决这个问题,SSD 算法按照最高置信度损失为每个默认框进行排序,从高到低选择,使正负样本比例达到 1:3。

2 改进的 SSD

为了提升绝缘子检测速度,在 SSD 基础上进行改进。其包含 MobileNets 系列网络为特征提取的主干网络模块、用于聚焦绝缘子特征信息的注意力机制、用于上下文信息融合的中间层以及预测模块。核心在于选取轻量级的主干网络来提高速度并且融入注意力机制以保证特征的提取效果。

2.1 MobileNet

MobileNets 系列网络使用深度可分离卷积来构建轻量级的深度神经网络,在识别时延和准确性之间做出了有效的权衡。MobileNet^[23]模型基于深度可分离卷积,将标准卷积分解为深度卷积和 1×1 的逐点卷积,分别用于特征的过滤和组合,大大减少了计算量和模型规模。MobileNetV2^[24]引入了线性瓶颈结构和反向残差结构,加深了网络层数,但是模型的体积更小,速度也有所提升,该版本又应用于目标检测框架中,将 SSD 预测层中的标准卷积替换为可分离卷积,即 SSDLite。MobileNetV3^[25]使用了神经架构搜索(neural architecture search, NAS)技术和 NetAdapt 算法生成网络结构,引入了“挤压和激励”模块和 h-swish 激活函数,提出了 Large 和 Small 两个版本,分别对应高资源和低资源的使用场景。

深度可分离卷积相对于标准卷积参数量和计算量都有大幅度减少。假设输入特征图的通道数为 C_{in} ,输出通道数为 C_{out} ,卷积核的尺寸为 $K \times K$,特征图的尺寸为 $H \times W$,那么进行一次标准卷积操作的参数量是 $C_{in} \times C_{out} \times K^2$,计算量为 $C_{in} \times C_{out} \times K^2 \times H \times W$ 。对于深度可分离卷积来说,深度卷积是对每个输入通道单独进行空间卷积,所以卷积核个数和输出通道数均等于输入通道数,参数量为 $C_{in} \times K^2$,计算量为 $C_{in} \times H \times W \times K^2$,逐点卷积是通过 1×1 卷积融合通道信息,将输入通道数映射到输出通道数,参数量为 $C_{in} \times C_{out}$,计算量为 $C_{in} \times C_{out} \times H \times W$,故深度可分离卷积的参数总量为 $C_{in} \times (K^2 + C_{out})$,计算总量为 $C_{in} \times H \times W \times (K^2 + C_{out})$,相比于标准卷积明显减少。

为实现轻量化和速度提升目标,将 MobileNets 系列网络作为 SSD 的主干网络进行替换,分别使用了 MobileNetv2、MobileNetv3-large 和 MobileNetv3-small。

2.2 注意力机制

注意力机制的基本思想是为特征的各个维度加权。本文使用了 4 种注意力模块进行实验,可以灵活地插入到经典网络中使用。

2.2.1 SE 模块

挤压和激励(squeeze and excitation, SE)模

块^[26]通过明确通道之间的相互依赖性,来自适应地重新校准通道相关的特征响应,包括挤压操作和激励操作。首先通过全局平均池化处理特征图($H \times W \times C$),得到 $1 \times 1 \times C$ 的特征图。然后,应用两个全连接层增加非线性处理过程,拟合通道之间的相关性,接着进行 sigmoid 激活,获得另一个 $1 \times 1 \times C$ 的特征图,最后将输入的特征图与该特征图进行全乘。

2.2.2 CBAM

卷积块注意力模块(convolutional block attention module, CBAM)^[27]是一个轻量级的通用模块。该模块结合通道注意力和空间注意力,将两个注意力特征图串联起来,生成通道加权和空间加权特征图,进行自适应特征细化。

2.2.3 ECA 模块

ECA 模块通过分析 SE 模块的通道注意力,注意到避免降维对于学习通道注意力十分重要,适当的跨通道交互可以在显著降低模型复杂性的同时提升性能。ECA 模块在不降维的前提下使用全局平均池化(global average pooling, GAP)聚合卷积特征,自适应确定卷积核的大小。ECA 的通道权重 w 计算公式如式(1)所示,先通过 GAP 得到每个通道的统计信息,再使用一维卷积快速完成计算,其中一维卷积的核大小 k 计算公式如式(2)所示:

$$w = \sigma(\text{Conv1d}_k(y)) \quad (1)$$

$$k = \left\lceil \frac{\log_2(C)}{\gamma} + \frac{b}{\gamma} \right\rceil_{\text{odd}} \quad (2)$$

式中: y 为输入特征; Conv1d 为一维卷积; σ 为 Sigmoid 激活函数; b 和 γ 为超参数; odd 表示保证卷积核大小是奇数。

在计算过程中,不需要升维降维的操作,保留了原始通道信息,提升了特征依赖表达能力。

2.2.4 CA 模块

坐标注意力(coordinate attention, CA)模块^[28]采用坐标信息嵌入和坐标注意力生成捕获特征的通道和位置之间的关系,将通道注意力分解为两个一维特征的编码过程,捕获远程依赖关系并且保持精确的位置信息,这些过程可以互补地增强对感兴趣对象的表示能力。

2.3 网络结构

在改进的 SSD 网络绝缘子检测模型中,将 ECA 模块与轻量级网络 MobileNetV3-large 相结合,其网络框架如图 3 所示,主要由主干网络 Backbone、收集特征图网络层 Neck 和预测模块 Prediction 三部分组成,能够实现端到端实时的绝缘子检测。

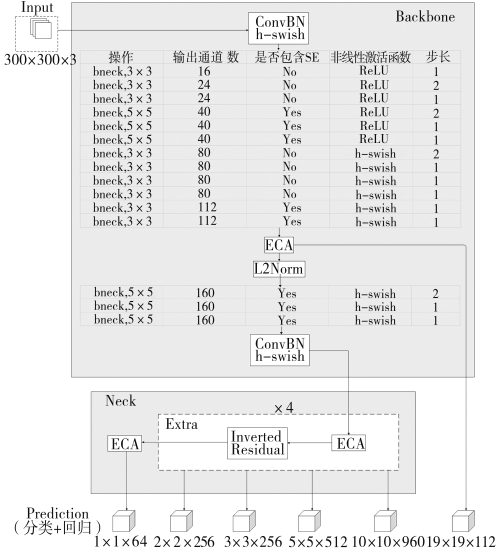


图 3 网络结构

首先,将原始图像按照 300×300 的尺寸输入到网络中。然后,利用主干网络 MobileNetV3-large 对输入的图像进行特征提取,将 ECA 模块加入到 12 个 bneck 之后,用于获取 $19 \times 19 \times 112$ 大小的特征图。该特征图输入到 L2Norm 层进行归一化,在经过 3 个 bnecks 和 Conv BN h-swish 模块后输入到 Neck 部分,经过添加的连续 4 个 Extra 特征提取层和 ECA 模块,获得不同大小的特征图输出。最后,在 Prediction 模块中,为获得的 6 个有效特征层添加分类预测和回归预测,获取绝缘子位置和分类的输出。

Backbone 部分经过实验比较和分析选择了速度与精度权衡效果最佳的 MobileNetV3-large 网络,该网络具有 MobileNets 系列网络的优势,将其组合起来构建最有效的模型。MobileNetV3 分为 large 和 small 两个版本,本文使用 large 版本。图 3 中,Backbone 首尾使用的 Conv BN h-swish 模块,具体结构如图 4(a)所示,MobileNetV3 在倒置残差模块中加入 SE 模块,并根据不同层使用不同的非线性激活函数,使用 ReLU

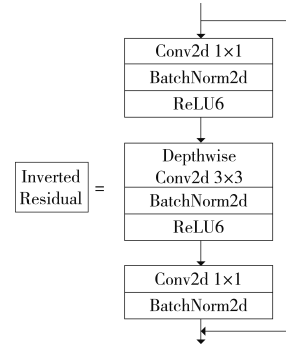
或 swish 函数来保持精度,ECA 模块具体结构如图 4(b)所示。Neck 部分用到的 ECA 模块与 Backbone 部分中添加的 ECA 相同,InvertedResidual 模块的具体组成如图 4(c)所示。

$$\text{ConvBN h-swish} = \text{Conv2d} + \text{BatchNorm2d} + \text{Hardswish}$$

(a) Conv BN h-swish 模块结构

$$\text{ECA} = \text{AdaptiveAvgPool2d} + \text{Conv1d} + \text{Sigmoid}$$

(b) ECA 模块结构



(c) InvertedResidual 模块结构

图 4 网络组件说明

在 Backbone 的首尾卷积层使用 h-swish 函数减少过滤器的数量。ECA 模块在不降维的情况下使用 GAP 对卷积特征进行聚合后,首先以自适应的方式确定核大小 k ,然后进行一维卷积,利用 sigmoid 函数学习通道注意力。InvertedResidual 结构将输入的低维压缩表示扩展到高维,再通过线性瓶颈层将特征投影回低维压缩表示,其与 ECA 模块一同组成了 Neck 部分的 Extra 结构。InvertedResidual 模块来源于 MobileNetV2,Backbone 中 MobileNetV3 的 bneck 也是根据该结构改进的。

2.4 损失函数

本研究使用的总体目标损失函数定义为定位损失 L_{loc} 和置信度损失 L_{conf} 的加权和,具体公式如式(3)所示:

$$L(x_{ij}^p, c, l, g) = \frac{1}{N} (L_{\text{conf}}(x_{ij}^p, c) + \alpha L_{\text{loc}}(x_{ij}^p, l, g)) \quad (3)$$

式中: $x_{ij}^p = \{1, 0\}$ 为第 i 个默认框与第 j 个类别 p 的真实框是否相匹配; c 为类别置信度的预测值; l 为模型输出的预测框; g 为真实框; α 为用于调整定位损失和置信度损失之间的比重,默认为 1; N

为匹配到的默认框的数量。

3 实验与分析

3.1 数据预处理

3.1.1 图片收集

本文实验的数据集包括前人收集的合并绝缘子数据集(merged insulators datasets, MID)和由 3 种绝缘子(玻璃绝缘子、陶瓷绝缘子和复合绝缘子)组成的自制数据集(self-made insulator dataset, SID),数据集样本分别如图 5(a)和 5(b)所示。MID 是由中国电力线路绝缘子数据集(chinese power line insulator dataset, CPLID)^[29]和 Tomaszewski 等提供的绝缘子数据集^[30]合并后,使用 albumentations 图像增强库进行数据增强得到的。



(a) 公开数据集绝缘子图像



(b) 自建数据集绝缘子图像

图 5 数据集样本

公开数据集包括无人机拍摄的正常绝缘子图像、合成的缺陷绝缘子图像、不同光照条件下以及不同背景下的高分辨率绝缘子图像。自建数据集列出的样本中分别包括了玻璃、陶瓷和复合绝缘子的图像,其中玻璃绝缘子图像是由实景拍摄得到的,而陶瓷和复合绝缘子则是通过将建模绝缘子图像和卫星背景相结合得到的。

3.1.2 数据标注

本文用到的两个数据集都采用 PASCAL VOC^[31]格式,数据集文件夹的主要组成如图 6 所示。

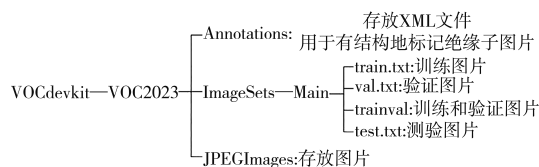


图 6 数据集文件夹结构

图像标签标注工具使用的是 LabelImg,它是用 Python 编写的,使用 Qt 作为图形界面。图 7 包括了两个数据集样本的标记示例。



图 7 数据集标记示例

图 7 展示了在 LabelImg 中使用矩形框标记绝缘子目标的过程和标记完成后标签在图片上的呈现方式。

3.1.3 训练集和测试集

数据集包括训练集、验证集和测试集三部分。测试集约占数据集的 1/10,训练集与验证集的比例约为 9 : 1,如表 1 所示。将两个数据集合并,训练集包括 6 289 张图像,验证集包括 699 张图像,测试集包含 777 张图像。

表 1 合并数据集

数据集	训练集	验证集	测试集
公开	5 550	616	694
自建	739	83	83
合计	6 289	699	777

3.1.4 评价指标

本研究使用的模型评价指标包括:平均精确度(average precision, AP)、模型参数量(parameters)、每秒 10 亿次的浮点计算量(giga floating-point operations per second, GFLOPs)以及每秒内可处理的图像数量(frames per second, FPS),同时列出精确率 P (precision)和召回率 R (recall)结果。其中,精确率 P 、召回率 R 的计算式分别如式(4)、式(5)所示:

$$P = \frac{T_P}{T_P + F_P} \times 100\% \quad (4)$$

$$R = \frac{T_P}{T_P + F_N} \times 100\% \quad (5)$$

式中: T_P (true positive) 为被预测为正的正样本数; F_P (false positive) 为被预测为正的负样本数; T_N (true negative) 为被预测为负的负样本数; F_N (false negative) 为被预测为负的正样本数。

A_P 是以召回率 R 、精确率 P 分别为横纵坐标构成的 $P-R$ 曲线在 $0 \sim 1$ 区间内的积分面积, 计算式如式(6)所示:

$$A_P = \int_0^1 P(R) dR \quad (6)$$

3.2 实验环境

本文中的实验基于 Pytorch 深度学习框架构建改进后的 SSD 网络。采用 Pytorch1.12.1 深度学习框架, Pycharm2022.1.3 软件, CUDA11.6 开发环境, Windows 11 操作系统完成模型的训练和测试。实验硬件配置为 Intel Core i5-12500H 12 核 16 线程的 CPU 以及 NVIDIA GeForce RTX-3050 Ti GPU。

3.3 训练过程

算法训练的参数设定为: 优化器种类为随机梯度下降 (stochastic gradient descent, SGD) 优化器, 模型的最大学习率为 0.002, 批处理大小设置为 8, 迭代次数为 200, 权值衰减为 0.000 04, 动量因子为 0.937。输入图像尺寸调整为 300×300 。训练过程中的损失下降曲线如图 8 所示。

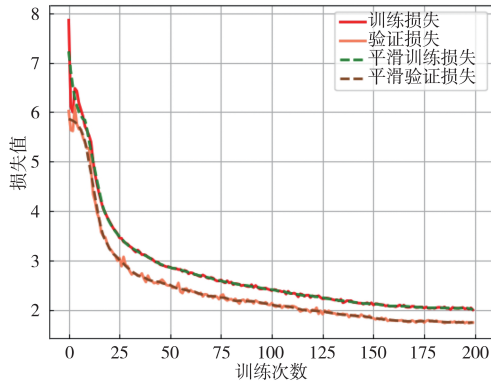


图 8 训练损失函数曲线

图 8 显示了最后确定的网络模型在训练过程中的损失下降情况, 包括训练集上的损失和验证集上的损失, 并对曲线做了平滑操作。可以明显看出, 伴随着训练次数的不断增加, 损失函数值不断下降, 在前 25 轮, 损失下降最快, 在 175 轮之后, 逐渐趋于平缓, 网络训练效果良好。

3.4 实验结果

3.4.1 不同主干的检测结果

保持 SSD 算法的基本结构不变, 改变主干网

络, 以其 A_P 、参数量、GFLOPs 和 FPS 作为评估指标, 分别从精确度、模型体积、浮点运算次数和速度 4 个方面对比分析不同主干网络对模型检测性能的影响, 实验结果如表 2 所示。

表 2 不同主干网络的检测结果

主干网络	$A_P/\%$	参数量/MB	GFLOPs	FPS
VGG16	93.54	23.61	60.8	56.90
MobileNetV2	85.74	3.54	1.4	134.50
MobileNetV3-large	90.09	4.03	1.0	122.10
MobileNetV3-small	87.57	1.71	0.3	135.40

3.4.2 不同注意力模块的检测结果

本实验利用上述评价指标对比分析了几种不同的注意力模块添加到网络中的性能, 选取效果最好的注意力机制, 实验结果如表 3 所示。

表 3 不同注意力模块的检测结果

注意力模块	$A_P/\%$	参数量/MB	GFLOPs	FPS
不添加	90.09	4.03	1.0	122.10
SE	92.58	4.20	1.0	113.18
CBAM	91.66	4.20	1.0	80.21
ECA	92.70	4.03	1.0	111.74
CA	91.75	4.20	1.0	88.59

由表 3 可以看出, 当在网络中加入各种注意力模块时, 精确度均有提升, 而参数量、浮点运算次数以及检测速度都没有受到太大影响。故选择效果最佳的 ECA 作为注意力机制的添加模块。

3.4.3 与主流目标检测方法对比分析

为测试融合 ECA 的 MobileNetV3-large 轻量级网络的表现, 使用现有数据集与本文网络同经典目标检测算法进行对比实验, 并选用上述评价指标来衡量不同算法的综合性能。与其它算法对比结果如表 4 所示。

表 4 不同模型的对比结果

模型	$A_P/\%$	参数量/MB	GFLOPs	FPS
Faster R-CNN	95.40	137.10	370.2	22.01
CenterNet	92.19	191.24	585.1	12.87
YOLOv5	94.42	37.20	105.1	41.14
YOLOv11 ^[32]	91.17	9.44	21.5	92.40
SSD	93.54	23.61	60.8	56.90
本文方法	92.70	4.03	1.0	111.74

由表 4 可知, 模型大小方面, 本文模型参数量为 4.03 MB, 检测速度最快, 达到 111.74 FPS, A_P 为 92.70%, 与其它模型相差不大。与最新算法 YOLOv11 对比, 在精确度、参数量以及检测速度方面, 仍然占据一定优势。与原始 SSD 算法相比, 参数数量从 23.61 MB 减少到 4.03 MB, 减少

了 82.93%；检测速度从 56.90 FPS 提升到 111.74 FPS,提升了 96.38%。综合结果表明,本文算法对绝缘子的检测在精度和速度之间有更好的平衡。

3.4.4 绝缘子检测结果展示

为验证模型泛化性,本文在自制数据集 SID 和公开数据集 MID 上进一步测试。为保证结果准确性,选择未参加训练的测试集和验证集,MID 测试验证集的图片数量为 1 310 张,SID 测试验证集数量为 166 张。在相同实验设置下测试模型性能,以最终得到的 A_P 值作为衡量指标,结果如表 5 所示。

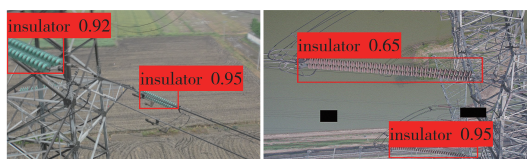
表 5 不同数据集的结果展示

A_P /%	SSD	YOLOv11	本文方法
SID	98.81	94.19	95.40
MID	93.42	90.75	92.84

由表 5 可知,在公开数据集 MID 上,本文提出的方法高达 92.84%的精确率,与 SSD 原始模型仅差 0.58%,且与先进模型 YOLOv11 对比也占有绝对优势,保持了轻量级特点和检测精度。选取测试集中的原始图像进行测试,分别选取了公开数据集和自制数据集,并与原始 SSD 算法进行比较,最终检测结果如图 9 所示。



(a) 自制数据集 SSD 检测结果 (b) 公开数据集 SSD 检测结果



(c) 自制数据集本文方法检测结果 (d) 公开数据集本文方法检测结果

图 9 绝缘子检测结果

其中红色框代表对应区域的子块被判定为属于该类目标,通过对比可以看出,SSD 算法能够检测出的绝缘子位置,本文算法也能够检测准确,可见与 SSD 的检测能力相差不大,在参数量和速度方面占有的优势证明了本算法的价值。

4 结论

为方便复杂输电场景下输电线路绝缘子缺

陷的检测,提升绝缘子检测算法性能,提出了一种基于 SSD 的改进绝缘子检测算法,经实验验证可得到以下结论:

1)在算法轻量化和提高实时性方面,将轻量级卷积神经网络 MobileNetV3-large 代替 VGG16 作为 SSD 的基础网络,该网络的模型体积更小、检测速度更快。

2)在提高算法准确性方面,采用 ECA 注意力机制使模型能够专注于学习绝缘子特征,提高了模型对绝缘子检测的灵敏度及辨别能力。在制作的绝缘子数据集上的 A_P 达到 92.70%。

本文方法在模型大小、检测速度方面的优势,为移动端和嵌入式等计算能力、存储资源有限的设备移植提供了可能。虽然能够快速实时地对绝缘子进行检测,但是准确率仍有较大提升空间。当前模型尚无法直接推断绝缘子类别、电压等级、群伞数等具体信息,未来将通过采集并重制数据集等方式进行多任务预测,进一步提升该方法的实用性。另外,绝缘子检测服务于绝缘子缺陷的检测和判别,下一步的研究工作将考虑对于识别到的绝缘子区域进行缺陷检测。

参考文献:

- [1] 龚钢军,张 帅,吴秋新,等.基于 TensorFlow 的高压输电线路异物识别[J].电力自动化设备,2019,39(4):204-209,216.
- [2] 杨达伟,杨明圣,付 博.改进 YOLOv7 的电力设备红外图像多目标检测[J].红外技术,2025,47(3):326-334.
- [3] GIRSHICK R,DONAHUE J,DARRELL T,et al.Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus,OH,USA:IEEE,2014:580-587.
- [4] GIRSHICK R.Fast R-CNN[C]//Proceedings of the 2015 IEEE International Conference on Computer Vision.Santiago,Chile:IEEE,2015:1440-1448.
- [5] 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.
- [6] 胡晓东,于传维,王世强,等.一种自适应的红外图像劣化瓷绝缘子诊断方法[J].河北电力技术,2024,43(5):83-88.
- [7] 李 浩.基于计算机视觉的电力设备红外检测研究[D].长沙:长沙理工大学,2020.
- [8] 程海燕,翟永杰,陈 瑞.于 Faster R-CNN 的航拍图像中绝缘子识别[J].现代电子技术,2019,42(2):98-102.

- [9] 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. Las Vegas, NV, USA; IEEE, 2016: 779 – 788.
- [10] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[C]//Proceedings of the Computer Vision-ECCV 2016. Amsterdam, The Netherlands; Springer, 2016: 21 – 37.
- [11] ZHOU X, WANG D, Krähenbühl P. Objects as points[EB/OL]. 2024 – 02 – 08. <https://arxiv.org/abs/1904.07850>.
- [12] 朱有产, 郑 怡, 秦金磊. 基于改进 YOLOv3 的绝缘子目标检测研究[J]. 电瓷避雷器, 2022(3): 166 – 171.
- [13] LI Y, YUAN G, WEN Y, et al. EfficientFormer: Vision Transformers at MobileNet Speed [C]//Proceedings of the Advances in Neural Information Processing Systems, New Orleans, LA, NIPS; 2022.
- [14] MEHTA S, RASTEGARI M. MobileViT: Light-weight, General-purpose, and Mobile-friendly Vision Transformer [J]. 2021. DOI: 10.48550/arXiv.2110.02178.
- [15] MAAZ M, SHAKER A, CHOLAKKAL H, et al. EdgeNeXt: Efficiently Amalgamated CNN-Transformer Architecture for Mobile Vision Applications [C]//Proceedings of the Computer Vision-ECCV 2022 Workshops, Tel Aviv, the State of Israel; 2023.
- [16] 马建宝, 李小飞, 毕 江, 等. 一种改进 SSD 和 ResNet 的光伏组件缺陷识别算法[J]. 粘接, 2025, 52(2): 141 – 145.
- [17] 李伟性, 郑武略, 王 宁, 等. 基于 SSD 算法的输电线路绝缘子缺陷检测方法研究[J]. 仪器仪表用户, 2019, 26(8): 1 – 4.
- [18] 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.
- [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. Seattle, WA, USA; IEEE, 2020: 11531 – 11539.
- [20] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [C]//Proceedings of the International Conference on Learning Representations, San Diego, CA, USA; OpenReview.net, 2015: 1 – 14.
- [21] CHEN L, PAPANDREOU G, KOKKINOS I, et al. Semantic image segmentation with deep convolutional nets and fully connected CRFs [J]. Computer Science, 2014, (4): 357 – 361.
- [22] JACCARD P. The distribution of the Flora in the alpine zone [J]. The New Phytologist, 1912, 11(2): 37 – 50.
- [23] HOWARD A, ZHU M, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications [EB/OL]. 2024 – 02 – 08. <https://arxiv.org/abs/1704.04861>.
- [24] SANDLER M, HOWARD A, ZHU M, et al. MobileNetV2: inverted residuals and linear bottlenecks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA; IEEE, 2018: 4510 – 4520.
- [25] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3 [C]//2019 IEEE/CVF International Conference on Computer Vision, Seoul, Korea (South); IEEE, 2019: 1314 – 1324.
- [26] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA; IEEE, 2018: 7132 – 7141.
- [27] WOO S, PARK J, LEE J, et al. CBAM: Convolutional block attention module [C]//Computer Vision-ECCV 2018. Munich, Germany; Springer, 2018: 3 – 19.
- [28] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, TN, USA; IEEE, 2021: 13708 – 13717.
- [29] TAO X, ZHANG D, WANG Z, et al. Detection of power line insulator defects using aerial images analyzed with convolutional neural networks [J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2020, 50(4): 1486 – 1498.
- [30] TOMASZEWSKI M, RUSZCZAK B, MICHALSKI P. The collection of images of an insulator taken outdoors in varying lighting conditions with additional laser spots [J]. Data in Brief, 2018(18): 765 – 768.
- [31] EVERINGHAM M, M S. ESLAMI A, VAN GOOL L, et al. The Pascal Visual Object Classes Challenge: A Retrospective [J]. International Journal of Computer Vision, 2015, 111(1): 98 – 136.
- [32] K KHANAM R, HUSSAIN M. YOLOv11: An Overview of the key architectural enhancements [EB/OL]. 2025 – 11 – 23. <https://www.arxiv.org/abs/2410.17725>.

本文责任编辑: 王丽斌