

文章编号: 1005-0523(2026)02-0104-11



基于 FedGSC 和 RT-DETR 的电力设备缺陷检测技术

韦宝泉, 刘龙平, 邓芳明, 刘彦志, 曾建军, 薛宪法

(华东交通大学电气与自动化工程学院, 江西 南昌 330013)

摘要: 现有电力设备缺陷检测算法难以同时保证检测精度与速度, 而且模型参数冗余、规模大, 对部署在边缘端嵌入式设备中构成了挑战。提出一种基于联邦梯度评分修正(FedGSC)算法和实时端到端目标检测器(RT-DETR)的电力设备缺陷检测技术。首先, 采用轻量化主干网络 GhostNet 替换 RT-DETR 的原有主干网络, 利用通道剪枝对模型体积进一步压缩, 大幅减少冗余参数, 提高推理速度; 构建基于云端服务器的联邦学习架构对边缘端轻量级 RT-DETR 模型进行分布式训练, 为解决联邦学习训练过程中存在的非独立同分布(Non-IID)数据, 引入 FedGSC 对每轮模型更新的梯度进行修正。实验结果显示, 轻量级 RT-DETR 与传统 RT-DETR 以及 YOLOv8 相比较, 算法模型大小仅 47 MB, 均值平均精度(mAP)为 90.46%, 能快速精准识别电力设备缺陷; 提出的 FedGSC 算法在训练精度和收敛性上都明显优于联邦平均算法(FedAvg)和联邦公平平均算法(Fed-FV), 并且 FedGSC 算法相较于 FedAvg 和 FedFV 算法分别节省 40% 和 20% 左右的通信成本。

关键词: 电力设备; 缺陷检测; 轻量化; 分布式训练; 联邦学习; 梯度评分修正算法

中图分类号: TM50; TP391

文献标志码: A

本文引用格式: 韦宝泉, 刘龙平, 邓芳明, 等. 基于 FedGSC 和 RT-DETR 的电力设备缺陷检测技术[J]. 华东交通大学学报, 2026, 43(2): 104-114.

Research on Defect Detection Technology for Power Equipment Based on Efficient Federated Learning and RT-DETR

Wei Baoquan, Liu Longping, Deng Fangming, Liu Yanzhi, Zeng Jianjun, Xue Xianfa

(School of Electrical and Automation Engineering, East China Jiaotong University, Nanchang 330013, China)

Abstract: Existing defect detection algorithms for power equipment are difficult to ensure both detection accuracy and speed, and the large scale of model parameter redundancy is a challenge for deployment in edge-side embedded devices. In this paper, we propose a power equipment defect detection technique based on federated gradient score correction(FedGSC) algorithm and real-time-detection Transformer(RT-DETR). First, the lightweight backbone network GhostNet is used to replace the original backbone network of the model in RT-DETR, and the model volume is further compressed using channel pruning, which significantly reduces the redundant parameters and improves the inference speed; on this basis, the federated learning architecture is introduced into the cloud server for the distributed training of the lightweight RT-DETR model at the edge end in order to solve the problem of non-independent homogeneous distribution(NID) in the federated learning training process. Non-independent and identically distributed(Non-IID) data exists, FedGSC is introduced to correct the gradient of each round of model update. Comparing the lightweight RT-DETR with traditional RT-DETR and YOLOv8, the algo-

收稿日期: 2024-12-24

基金项目: 国家自然科学基金项目(52377103, 52567009); 江西省自然科学基金项目(2023BAB204064)

rithm model size is only 47 MB, and the mean average precision(mAP) is 90.46%, which can quickly and accurately identify the defects of power equipment; and the FedGSC algorithm reduces communication costs by approximately 40% and 20% compared to the FedAvg and FedFV algorithms, respectively.

Key words: power equipment; defect detection; lightweighting; distributed training; federated learning; gradient score correction algorithm

Citation format: WEI B Q, LIU L P, DENG F M, et al. Research on defect detection technology for power equipment based on efficient federated learning and RT-DETR[J]. Journal of East China Jiaotong University, 2026, 43(2): 104-114.

近年来,随着我国能源互联网的发展,电力设备装机容量不断增大,中国已经建成世界上电压等级最高,体量最大的电网系统^[1],而持续可靠供电是发展的前提。电力系统的安全稳定运行高度依赖于电力设备的可靠性水平^[2]。因此,必须定期对电力设备进行缺陷检测,以确保设备的稳定运行。

传统电力设备缺陷检测主要依赖人工巡检^[3],这种方式不仅费时费力,并且检测效率很低。随着机器视觉的发展,越来越多的研究人员将深度学习目标检测技术应用在电力设备缺陷检测中^[4]。国内外研究人员基于Transformer提出了不同的目标检测模型^[5]。张乃雪等^[6]提出一种Smooth-DETR算法,基于Smooth-L1损失函数的优势改进GIOU损失函数,对设备表面小尺寸缺陷检测精度有很大的提高。李刚等^[7]提出将先验知识融合DETR目标检测算法,增强模型的识别能力。欧阳名三等^[8]采用Transformer stage和残差网络(residual networks, ResNets)中间输出特征来改进DETR算法,能更好地捕捉绝缘子不同尺度特征信息,增强泛用性。但上述方法在收敛速度上存在明显不足,并且模型参数还是过于庞大,在训练过程中难以满足实际要求。

受限于数据隐私保护要求,各训练参与者之间会形成数据孤岛,难以收集大量电力设备数据,导致检测模型训练十分困难^[9]。联邦学习在保护数据隐私的同时极大地推动了电力设备检测领域的发展,然而在实际训练中,广泛存在的非独立同分布(non-independent identically distribution, Non-IID)数据^[10-11],会降低模型表现并且严重影响网络传输中的通信效率。Manasa等^[12]引入层次聚类技术改

进联邦平均算法,根据本地模型和全局模型的相似度对其聚类分离,减少通信频次。Karimireddy等^[13]通过预估本地模型更新差异来缓解Non-IID数据的问题。韩沛秀等^[14]提出了带个性化层的联邦学习(federated learning with personalization layers, FedPer)算法,划分全局模型部分,然后分别训练网络层和个性化层解决非独立同分布问题。张瑞麟等^[15]引入生成对抗网络去Non-IID化,使得参与训练的数据都趋于独立同分布,提高了通信效率。Wang等^[16]提出的FedFV通过梯度投影在聚合之前修正客户端梯度更新方向,缩减不同客户端之间因Non-IID数据带来的梯度差异。但目前,联邦学习在电力设备缺陷检测模型训练中的Non-IID数据问题还需不断探索和研究。

针对上述模型参数过于庞大并且收敛速度不足的缺陷,文中旨在降低占用边缘服务器的空间并提高推理速度。将原始特征网络中的ResNet替换成GhostNet,并结合通道剪枝策略对模型进一步压缩,提出一种轻量级实时端到端目标检测器(real-time-detection transformer, RT-DETR)算法,并在边缘客户端搭载该算法。然后,在云端引入联邦学习框架打破传统集中式训练下电力数据孤岛壁垒,引入联邦梯度评分修正(federated gradient score correction, FedGSC)算法进行梯度修正,并将修正后的梯度发送至云端,解决联邦学习训练过程中Non-IID数据的问题。最后,通过模型仿真验证所提缺陷检测技术的有效性。

1 轻量化RT-DETR缺陷检测模型

1.1 RT-DETR网络基本架构

RT-DETR是一个基于Transformer架构改进的

端到端实时目标检测模型^[17],消除了置信度阈值和IoU阈值对检测速度的延迟影响,在速度和精度上获得了非常好的性能。RT-DETR整体模型网络结构如图1所示,它一共包含三个主要部分:第一部分为传统的CNN网络当作主干网络用于提取原始图像的特征向量,采用了经典的ResNet和可缩放HGNetv2两种网络;第二部分是颈部网络Efficient hybrid encoder,包含Attention-based intra-scale feature interaction(AIFI)和CNN-based cross-scale feature-fusion module(CCFM)两个模块;第三部分为带辅助检测头的解码器,用于优化查询对象以生成预测方框和所属类别的置信度分数。

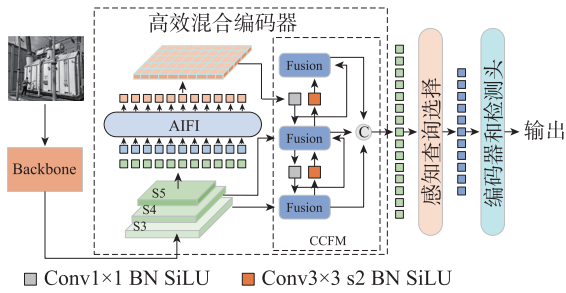


图1 RT-DETR网络结构

Fig. 1 RT-DETR network structure

1.2 轻量化主干网络GhostNet

为了减少RT-DETR-R34模型的体积,提高检测速度,本文将轻量化GhostNet代替原来的基础主干网络,以降低模型初始复杂度。相较于传统卷积神经网络的逐点卷积降维,GhostNet融合了线性运算和深度卷积运算^[18]。通过常规卷积、Ghost生成和特征图拼接进行特征信息提取^[19],如图2所示。

在训练中,给定输入数据 $X \in \mathbb{R}^{c \times h \times \omega}$,其中 c 为输入通道数, h 、 ω 分别为高度和宽度,在任意卷积层产生 n 个特征映射的操作过程可以表示为

$$Y = X * f + b \quad (1)$$

式中: b 为偏置项; $f \in \mathbb{R}^{c \times k \times k \times n}$ 为当前层的卷积滤波器。其中 h' 和 ω' 分别为输出数据的高度和宽度, $k \times k$ 分别为卷积滤波器 f 的核大小。在卷积过程中,所需要的浮点数(FLOPs)可计算为 $n \times h' \times \omega' \times c \times k \times k$,在这个过程中会包含大量的冗余参数。原输出的特征为某些内在特征映射的Ghost向量,这些通过卷积滤波器产生的映射的Ghost向量一般较小,数量为 m 的内在特征映射使用一次卷积生成

$$Y' = X * f' \quad (2)$$

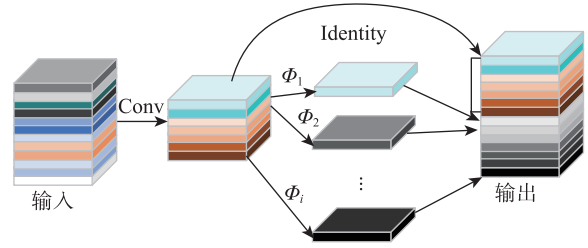


图2 Ghost卷积过程

Fig. 2 Ghost convolution process

式中: $f' \in \mathbb{R}^{c \times k \times k \times m}$ 为使用的滤波器, $m \leq n$,为简化式子,省略偏置项。通过对 Y' 中的每个特征进行简洁的线性操作,获得所需数量为 n 的特征映射,根据以下公式生成 s 个Ghost特征图

$$y_{ij} = \Phi(y'_i), \forall i = 1, \dots, m, j = 1, \dots, s \quad (3)$$

式中: y'_i 为在特定特征图中的第 i 个特征图, Φ_{ij} 为线性变换函数,表示序列 i 的特征图经过的第 j 个线性变换, y'_i 可以有一个或者多个Ghost映射,根据式(3)可得 $n = m(s-1)$ 个特征映射 $Y = [y_{11}, y_{12}, \dots, y_{ms}]$ 作为Ghost模块的输出数据。为保留固有特征图,通过添加恒等映射 Φ_{ij} ,叠加特征图到经过线性变换的特征图上。

因为Ghost模块生成的特征映射与普通卷积层输出的特征映射数量相同,为进一步分析其理论加速比,假设一个固有单位映射, $n = m(s-1)$ 个线性操作,每个通过线性操作的平均核大小尺寸为 $d \times d$,其具体实现过程如下式

$$r_s = \frac{n \times h' \times \omega' \times c \times k \times k}{\frac{n}{s} \times h' \times \omega' \times c \times k \times k + (s-1) \times \frac{n}{s} \times h' \times \omega' \times d \times d} \quad (4)$$

其中 $k \times k$ 与 $d \times d$ 程度几乎相同,又 $s \ll c$,所以理论上的参数压缩比为

$$r_s = \frac{s \times c}{s + c - 1} \approx s \quad (5)$$

1.3 基于GhostNet与通道剪枝的轻量化RT-DETR模型

本文将轻量级主干网络GhostNet与原始RT-DETR-R34模型进行结合,替换后的网络结构如图3所示。

在上述轻量化RT-DETR模型中,GhostNet先将输入的原始图片通过卷积操作得到部分真实的特征层,再将真实特征层通过线性操作得到剩余的Ghost层,最后将这两个特征层拼接到一起形成一个完整的特征层,GhostNet所提取的3个特征图用

S3, S4, S5 标记,它们分别对图片进行了8倍、16倍以及32倍的下采样。将3个特征图与RT-DETR多尺度特征进行拼接,进入编码器部分,AIFI只处理S5特征图,经过多头自注意力与FFN,将其调整成二维,生成F5。F5经过 1×1 的卷积与S4融合,第1次融合特征图经过 1×1 的卷积与S3再融合输出第1个结果;生成的第一个结果特征图经过 3×3 的下采样再与第一次融合图进行计算输出第2个结果;生成的第2个结果图经过 3×3 的下采样后与F5的 1×1 卷积特征图进行融合计算生成第3个结果,最后将3个结果特征图相加后输出。解码器接收到编码器输出的特征序列,通过IoU-aware Query Selection,训练期间约束检测器对高IoU的特征产生高分类分数,对低IoU的特征产生低分类分数,从而使得模型根据分类分数选择的Top-k特征对应的预测框同时具有这两种类型的IoU特征。

轻量化RT-DETR网络损失函数 l 由两个部分组成:框回归重建的损失函数 L_1 loss和GIoU loss、分类损失函数Focal loss,其定义如下

$$l_1(x, y) = \frac{1}{n} \sum_{i=1}^n |y_i - f(x_i)| \quad (6)$$

$$l_{\text{GIoU}} = \text{GIoU} - \frac{A^c - U}{A^c}, \quad -1 \leq \text{GIoU} \leq 1 \quad (7)$$

$$l_f(p_i) = -\alpha_i (1 - p_i)^i \log(p_i) \quad (8)$$

式中: l_{GIoU} 为样本 i 的真实值; x_i 为模型对样本 i 的预测标签;GIoU loss中IoU表示预测框与真实框的相似程度; A^c 为真实框和预测框最小外接矩形的面积; U 为预测框与真实框之间重合以外的面积;分类损失函数Focal loss中 p_i 表示模型预测类别的概率(置信度); α_i 用来平衡正负样本数量,样

本数量多即赋予更小的 α_i 值,样本数量少的则与之相反; λ 用于调节难分易分样本不均衡问题,一般情况下 $\lambda \geq 1$ 。

尽管替换了轻量化的GhostNet主干网络,然而在实际训练过程中还是会产生较大的冗余参数。为解决这个问题,在训练具有GhostNet主干网络的RT-DETR模型时,对其采用通道剪枝的方法进一步轻量化。在对模型进行剪枝时,需要判断每个特征图的重要程度,将得分比较低的特征图去掉,在本文方案中将 γ 值的大小作为判断特征图重要性的标准。将正则化 L_1 引入BN层,让重要的特征图 γ 趋于更大,不重要的特征图 γ 值趋于0,从而更有效判断哪个特征图更加重要。在模型训练中 L_1 正则化具有稀疏参数的作用,将其添加到权重中,将权重稀疏化, L_1 正则化具体公式如下

$$L_1 = \mu \sum_{j=1}^n |w_j| \quad (9)$$

联合训练的损失函数如下

$$L = \sum_{(x,y)} l(f(x, W), y) + \mu \sum_{y \in L} g(\gamma) \quad (10)$$

式中: x, y 分别为训练的输入与输出, W 为可训练的权重参数, $l(f(x, W), y)$ 为常规的卷积神经网络训练损失函数, $g(\cdot)$ 为对缩放因子的惩罚函数,在本文中取 $g(s) = |s|$,权重系数 μ 用来平衡损失函数和正则化项在系数训练过程中,由于正则化项的存在,迫使 γ 趋向于0。具体裁剪通道数据由实验设定的阈值决定,由于剪枝深层网络带来的误差要大于浅层网络,将浅层网络的阈值设为0.8,深层网络的阈值设为0.2。本文基于GhostNet结合通道剪枝的轻量化RT-DETR模型检测流程如图4所示。剪枝完

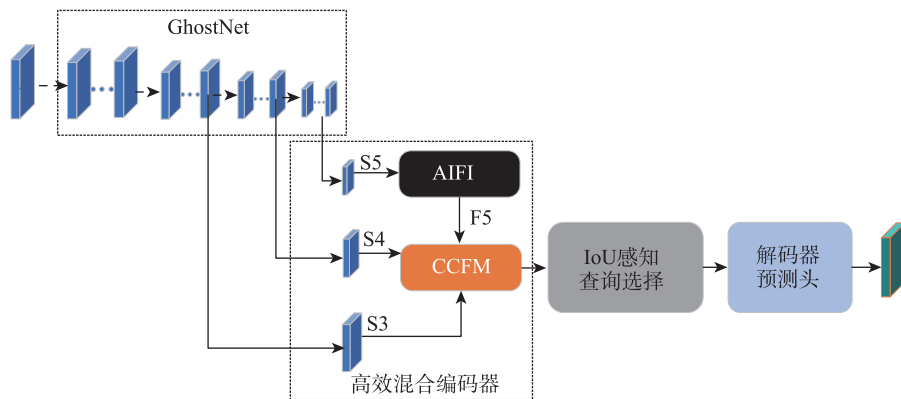


图3 轻量化RT-DETR网络结构图

Fig. 3 Network structure of lightweight RT-DETR

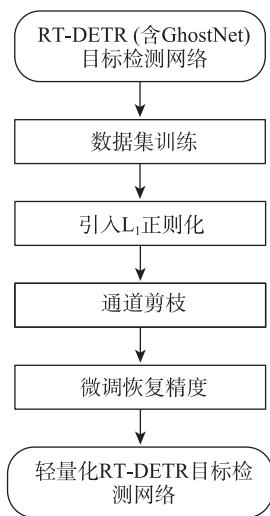


图4 轻量化流程图

Fig. 4 Flowchart of lightweight

成后模型参数和运算量大幅减少,运行速度更快,能更好地在智能移动设备端进行实时检测。

2 基于联邦学习的缺陷检测系统

2.1 基于联邦学习的分布式训练框架

联邦学习作为一种分布式训练方法,可以在不泄露隐私数据的同时训练出高性能的检测模型。本文将目标检测技术和联邦学习结合,提出一种全新的分布式模型训练架构系统,用于电力系统设备缺陷检测。系统架构如图5所示。

每个客户端都拥有自己的数据集,通过无人机实时采集上传,并在本地采用轻量化RT-DETR算法进行电气设备缺陷检测模型的训练。训练完成后,客户端将模型关键指标和局部参数模型上传至云端服务器,云端服务器接收到所有参与训练的客户端梯度后经过加权聚合得到全局模型参数,以模型关键指标为基准通过云端搭载的联邦学习算法得到全局参照梯度,之后再将全局模型参数和参照梯度发送到各个客户端,各客户端接收后,利用全局参数更新本地模型,并结合参照梯度修正下一轮训练,经过多次迭代直到损失函数收敛。

结合图5系统架构,建立一个虚拟的电力系统联网场景,若该场景中有 N 个变电站,无人机利用蜂窝网络将拍摄的电力设备图像上传到边缘客户端服务器中。具体检测流程如下。

1) 移动智能检测设备(无人机)对数据进行采集后将所需有效图片上传至客户端服务器。

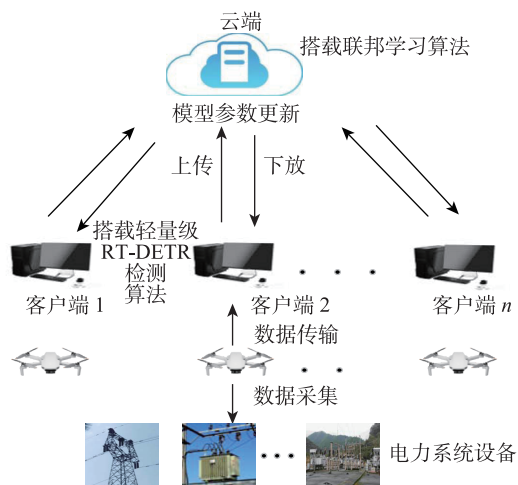


图5 系统框架图

Fig. 5 System frame diagram

2) 客户端下载初始化的云端全局模型,通过本地轻量化RT-DETR算法训练局部模型,并将局部模型参数上传至云端服务器。

3) 云端服务器接收各个客户端的局部模型参数,利用联邦学习算法聚合局部模型,再更新下发新的全局模型。

4) 重复上述3个步骤,检验全局模型参数,多次迭代训练直至参数达标收敛,结束训练。

2.2 面向Non-IID数据的FedGSC算法

目前通常采用FedFV算法聚合各个客户端参数来计算全局模型,该方法在独立同分布条件下十分有效^[20]。然而实际模型训练时,通常会存在Non-IID数据,并且在每一轮通信过程中,服务器只会随机选择部分客户端去训练。不同终端设备所拥有的数据分布不同,各个客户端梯度差异性较大^[21]。这种情况下,由于数据的分布差异,直接将各个终端设备的数据进行聚合可能会导致模型性能下降。为了尽可能缩小各个客户端梯度之间的差异,需要更多关键指标来提高全局梯度的质量,使其成为参照梯度。本文选取3个最重要的指标,即模型准确率(model accuracy, MA),数据质量(data quality, DQ)和模型差异(model differences, MD)来描述局部模型梯度对全局梯度的影响,为了更好地量化训练指标,给出关键训练指标的定义,使用式(11)中的标准Min-Max归一化操作将其量化到相同的取值范围

$$\text{Normalize}(z) = \frac{z - \min_{\text{val}}}{\max_{\text{val}} - \min_{\text{val}}} \quad (11)$$

式中: \min_{val} 和 \max_{val} 分别为训练中指标的最大值和最小值, z 为当前轮次中的指标值, 各关键指标的定义如下所示。

MA:

$$C_{\text{MC}}^k = \text{Normarize}(acc^k) \quad (12)$$

式中: 客户端 k 上的模型准确率 C_{MC}^k 为 acc^k 的归一化值。

DQ:

$$C_{\text{DQ}}^k = \text{Normarize} \left(\frac{\sum_{i=1}^M \sum_{j=1}^N d_{ij} f_{ij}}{\sum_{i=1}^M \sum_{j=1}^N f_{ij}} \right) \quad (13)$$

式中: M 和 N 分别为局部数据标准和全局数据标准分布的标签号; d_{ij} 为 i 和 j 之间的标签距离; f_{ij} 代表局部数据分布中第 i 个标签与全局数据分布中第 j 个标签的权重差。

MD:

$$C_{\text{MD}}^k = \text{Normarize} \left(\sqrt{(\mathbf{e} - \mathbf{f})^T (\mathbf{e} - \mathbf{f})} \right) \quad (14)$$

式中: C_{MD}^k 为客户端 k 上局部模型和全局模型参数差, 其为归一化值; \mathbf{e} 和 \mathbf{f} 分别为全局模型和局部模型的输出层矩阵。数据分布的差异会导致模型的权重参数不同, 因此局部梯度与全局梯度的差异也在一定程度上反映了局部数据与全局数据分布的差异。

由于客户端模型的训练需要经过多次迭代与同步才能收敛, 且在每一轮迭代中3个关键指标的数值可能动态变化, 难以直接评估模型质量。为此, 本文综合考虑关键指标及其相互依赖关系, 定义了一个模型质量评分, 该评分能够有效反映局部模型的质量。具体如式(15)

$$M_k = \omega_1 \cdot C_{\text{MC}}^k + \omega_2 \cdot C_{\text{DQ}}^k + \omega_3 \cdot C_{\text{MD}}^k \quad (15)$$

s.t.

$$\begin{cases} \omega_1 = 1 \\ \omega_2 = \omega_1 \cdot C_{\text{MC}}^k \\ \omega_3 = \omega_2 \cdot C_{\text{DQ}}^k = \omega_1 \cdot C_{\text{MC}}^k \cdot C_{\text{DQ}}^k \end{cases}$$

式中: M_k 为模型质量分数, C_{MC}^k 、 C_{DQ}^k 和 C_{MD}^k 分别为客户端 k 局部模型指标 MC、DQ、MD 的值, ω_1 、 ω_2 和 ω_3 分别为3个关键指标的权重。

然后根据模型质量分数 M_k 计算客户端 k 上本地模型的梯度聚合权值 A_k , 如式(16)

$$A_k = \frac{M_k}{\sum_i M_i} \quad (16)$$

将式(15)和式(16)结合后, 得到第 n 轮训练后的全局聚合参照梯度 \mathbf{g}_n^w , 如式(17)

$$\mathbf{g}_n^w = \sum_{k=1}^N \frac{M_k}{\sum_i M_i} w_n^k w_k^{t+1} \quad (17)$$

式中: w_n^k 为本地客户端训练更新参数。

客户端 k 首先从服务器端获取参照梯度方向 \mathbf{g}_n^w 和全局模型 w^t , 之后根据本地数据训练得到新的客户端 w_k^{t+1} , 然后计算客户端更新梯度 \mathbf{g}_k^{t+1} 。随后客户端判断参照梯度方向和客户端更新梯度方向之间的冲突, 进行梯度修正, 并将修正后的梯度发送至服务器。本地客户端在训练时根据式(15)和式(16)计算得到局部模型质量分数和本地模型梯度聚合权值, 服务器接收到所有参与训练的客户端梯度后, 根据式(17)更新全局聚合参照梯度, 再聚合所有局部模型更新的梯度, 最后将新的全局模型和参照梯度下放到客户端。

3 实验过程及结果分析

3.1 轻量化 RT-DETR 算法有效性验证

3.1.1 数据集与实验环境设置

本文的实验数据来自某省电网公司变电站和输电线路无人机巡检采集到的数据集。通过对图像进行翻转、剪裁和平移等操作增加样本多样性, 扩充数据集到4000张。其中包含了挂点螺栓缺失、变压器火灾、鸟巢和避雷器损坏4类设备缺陷图, 图像分辨率为480像素×300像素。为了训练和测试神经网络, 将数据集按8:2的比例分为训练集和测试集, 其中训练集包括正常图像1600张, 缺陷图像1600张, 测试集包括正常图像400张, 缺陷图像400张。

实验基于深度学习开源框架 Paddle 进行, 采用 Python 语言编程。本实验是对电气设备进行实时缺陷检测, 故本文采用性能评价指标、速度评价指标以及模型大小来评估检测模型, 采用全面的平均精度 (mean average precision, mAP) 来评价性能。

3.1.2 实验结果分析

将原始 RT-DETR-R34 中 Backbone 替换为

GhostNet, GhostNet 预训练权重是在 ILSVRC2012 数据集上完成的,利用迁移学习的方式将其用在本实验数据集上对电气设备缺陷进行实时检测。训练分为两个部分:首先,将 AIFI 模块权重冻结,针对主干网络做通道剪枝,批量大小设为 8,学习率设定为 0.000 5。初始模型已经具备较高的准确率和召回率,故剪枝训练只需设置为 50 次迭代,浅层网络裁剪阈值设为 0.8,深层网络的裁剪阈值设为 0.2,训练到 45 次迭代时,将学习率衰减为原始值的 1/10。然后,将 AIFI 模块解冻,所有参数参与训练,这一阶段迭代 300 次,批量大小设为 4,学习率设为 0.000 1,权重衰减设置为 0.000 01,每 50 次迭代进行一次衰减来保证模型更好地收敛。

通道剪枝训练和剪枝训练完成后轻量化 RT-DETR 训练 Loss 曲线图如图 6、图 7 所示。通道剪枝中 Loss 值前期上升,由于正则化的惩罚因子,BN 层的 γ 因子逐渐变小,Loss 值在后续训练中又开始回

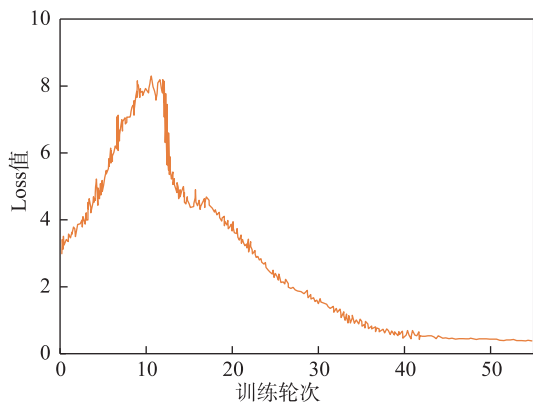


图 6 通道剪枝训练损失

Fig. 6 Loss of channel pruning training

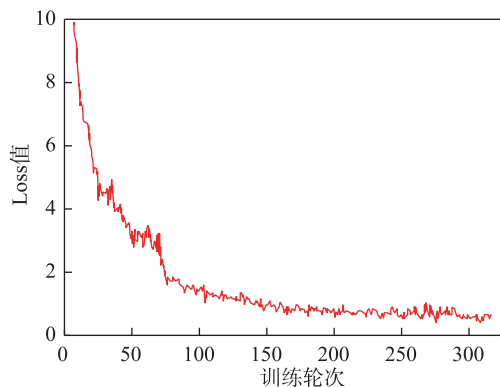


图 7 轻量化 RT-DETR 训练损失

Fig. 7 Training loss of lightweight RT-DETR

落直到稳定收敛。在完成剪枝训练后,解冻 AIFI 模块,随着训练的迭代次数增加,Loss 值整体呈现下降趋势,一直到迭代 300 次后稳定收敛,此时已达到最优检测效果。

图 8 为部分缺陷检测的可视化实验结果图,此图直观地体现了轻量化 RT-DETR 模型的检测效果。



图 8 部分缺陷检测可视化结果

Fig. 8 Visual defect detection results

以上缺陷检测结果图中,黄色框为电气设备预测框,红色框为设备缺陷预测框,右边数值为设备和缺陷类型的概率输出。

实验结果表明,本文提出的轻量化 RT-DETR 检测算法对挂点螺栓缺失、变压器火灾、鸟巢缺陷、避雷器破损的检测准确率分别为 81.13%、84.31%、88.92%、95.44%,可以看出本文提出的算法在对挂点螺栓这种小尺度目标的检测中精度并不算特别高,但也能达到 80% 以上,总之,该算法对于大部分的电力设备缺陷图像都具有良好的检测效果。

为了验证轻量化 RT-DETR 的高精度实时检测性能,将本文算法和 YOLOv8n、RT-DETR-R34 这 2 种算法的性能进行对比,以 mAP、每秒处理图像的张数(FPS)、模型大小和计算量作为评价指标,对比结果如表 1 所示。

从表中数据可知,RT-DETR-R34 算法检测精度和模型大小都是最高的,然而本文提出的轻量化 RT-DETR 算法在 RT-DETR-R34 算法的基础上大大缩减了模型冗余参数,模型大小仅有 47 MB,浮点计算量为 18.5,并且在检测速度上明显提升,达到

表1 各算法性能对比

Tab.1 Performance comparison of algorithms

检测算法	mAP/%	处理速度/ (帧/s)	模型 大小/MB	GFLOP/s
YOLOv8n	83.58	71	37.6	8.9
RT-DETR-R34	92.11	83	68.0	23.2
轻量级RT-DETR	90.46	96	47.0	18.5

96 FPS,成为检测速度最快的算法,mAP值为90.46%,以牺牲少量的精度大幅压缩参数量,减少推理过程中的计算开销,实时检测性能得到显著提高。

3.2 FedGSC 模型性能对比实验

3.2.1 数据集制作

本实验数据来自3.1中所采用的电力数据集,通过对图像数据扩充得到数据集7 800张,包含了正常图像和缺陷图像。通过划分训练集和测试集,随机选取800张(包含缺陷图像)作为全局模型的测试集,剩余的7 000张图片随机作为10个边缘客户端的本地模型训练样本。

3.2.2 数据集制作

1) FedGSC与集中式云计算训练对比实验。为了直观体现出联邦学习算法在模型训练方面的优势,进行FedGSC模型训练模式与传统云计算训练模式的对比。随机选取5个客户端分别进行集中式云计算训练和联邦分布式训练,全局通信频次设置为45次。每个客户端的局部模型采用轻量化RT-DETR检测模型,集中式训练等价于每轮只有一个客户端参与训练,为了得到高性能的配置以及实验的准确性,每个客户端需要迭代训练3次,评价指标为检测精度与通信轮次。

从表2可知,相同的通信轮次下,采用FedGSC模型训练模式的检测精度要比传统集中式训练模式高,另外在达到相同检测精度的前提下,FedGSC训练模式所用的通信轮次明显要比集中式训练少,且集中式训练次数越多,模型的检测精度越接近于联邦学习模型。由此可知,基于FedGSC的联邦分布式训练无论是在通信效率还是检测精度都优于传统集中式云计算。

2) 精度和收敛性能对比实验。将提出的FedGSC与善于处理Non-IID数据的联邦公平平均算法(federated fair averaging, FedFV)算法以及联邦平均算法(federated averaging algorithm, FedAvg)

表2 设备缺陷检测中不同训练方式的mAP对比

Tab.2 Comparison of mAP for different training

通信轮次	FedGSC	集中式云计算训练		
		局部训练 1次	局部训练 2次	局部训练 3次
5	45.2	37.4	38.2	40.1
15	57.6	42.3	46.8	53.2
25	64.3	54.7	58.9	61.5
35	78.9	64.7	69.5	73.2
45	87.2	74.1	77.8	79.3

一同在具有Non-IID数据分布下的电力数据集上展开对比。

为测试不同算法对Non-IID数据的抗性,每个客户端将原始电力数据集通过预处理、清洗和转换得到不同程度的Non-IID数据,并实验对比3种算法在不同程度Non-IID数据下的精度。

如图9所示,FedGSC与FedAvg的测试精度随着Non-IID程度的增加而迅速下降,意味着当迭代次数越来越多时,各个客户端之间差异的增加导致聚合全局模型的性能降低,特别是当Non-IID程度为1时,测试精度低至66%和58%。本文提出的FedGSC全程保持稳定并无太大的浮动,并且其在Non-IID程度达到1时依然能具有92%的精度。因此使用FedGSC训练模型时,梯度修正机制能有效解决Non-IID数据带来模型性能下降的问题。

在原始电力数据集验证3种算法性能的有效性,图10和图11分别是训练的准确率和损失值。由图可知,在性能表现上本文提出的FedGSC算法

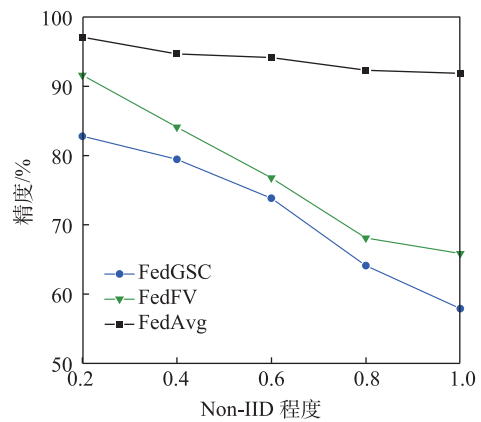


图9 不同非独立同分布程度下各算法的精度

Fig.9 Accuracy of each algorithm under different non-independent co-distribution degree

相较于 FedFV 和 FedAvg 算法都具有明显的优势, FedGSC 算法在 70 轮左右达到收敛, 并且准确率一直在 90% 以上, 对比 FedFV 略有提升, 相较 FedAvg 大幅提高。模型训练过程中, 损失值越小, 模型的鲁棒性越好, 损失值的变化是一直下降的, 证明了 FedGSC 算法具有优秀的性能。

3) 通信开销对比实验。本组实验对比 FedGSC、FedFV 和 FedAvg 的通信开销, 设置 10 个节点数, 选用相同的实验环境和电力数据集来保证实验结果的公平。图 11 展示了不同算法在训练过程中每个节点的通信传输数据尺度, 表 3 是达到收敛后 3 个算法的具体通信开销。

如图 12 所示, FedGSC 算法训练模型传输的数据尺寸小于其他 2 个算法, 全程保持 55 kB 左右。在联邦训练过程中, 初始用户节点 1 的通信效率最高, 随着梯度参数的传递, 数据量的传输尺寸逐渐趋于稳定。结果表明, 从整体角度来说每个参与实

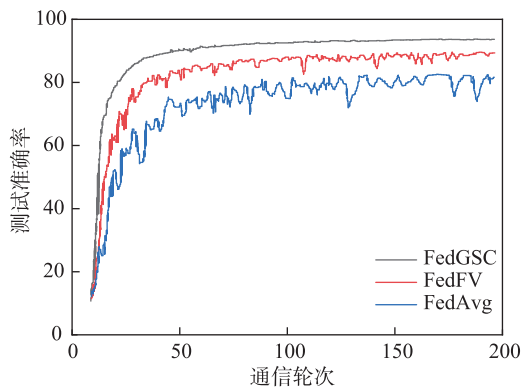


图 10 数据集测试准确率对比

Fig. 10 Comparison of test accuracy of data set

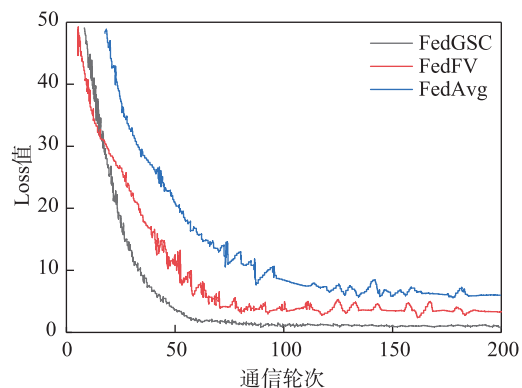


图 11 数据集测试损失值对比

Fig. 11 Comparison of test losses in data sets

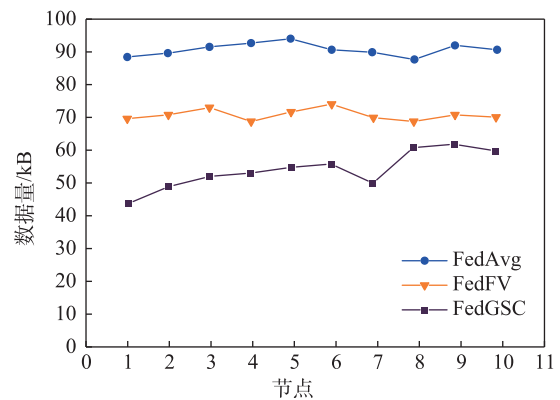


图 12 不同算法每个节点梯度尺寸大小

Fig. 12 Gradient size of each node in different algorithms

表 3 数据集上不同算法的通信开销计算

Tab. 3 Communication overhead for different algorithms on the dataset

训练方法	平均信息量/(kB/节点)			通信轮次
	梯度信息	上下文信息	总信息	
FedAvg	72.2	92.4	198.4	197
FedFV	8.6	69.4	104.5	143
FedGSC	4.1	43.6	92.4	118

验的用户节点通过 FedGSC 算法训练模型在通信效率方面远高于 FedAvg 算法, 略高于 FedFV 算法。

表 3 展示了在数据集下不同算法的平均每个节点的具体通信开销, 从实验结果可以看出, 传统的 FedAvg 算法在通信开销中是最高的, FedFV 作为善于处理 Non-IID 数据的联邦算法, 其基于投影映射方式缓解梯度偏差, 使得该算法在梯度信息上的通信开销比 FedAvg 算法低了 8 倍, 仅有 8.6 kB, 通信轮次 143 轮, 比 FedAvg 少了 54 轮。本文提出的 FedGSC 基于 FedFV 进行改进, 得益于梯度修正机制减少冗余梯度数据的传输, 梯度信息开销只有 4.1 kB, 只有 FedFV 的一半左右, 并且比传统 FedAvg 算法低了 17 倍左右, 通信轮次分别比前 2 个算法少了 79 轮次和 25 轮次, 相比之下 FedGSC 通信效率有着显著提高, 通信开销成本大幅降低。

4 结论

面向电力系统大背景, 为保障电网的稳定运行, 本论文融合目标检测技术和联邦学习框架, 设计了一个基于高效联邦学习的电力设备缺陷检测系统。系统包含云-边端, 边缘端服务器搭载轻量

级 RT-DETR 算法,在云端服务器引入联邦学习的架构对边缘端轻量级 RT-DETR 模型进行分布式训练。通过实验仿真得出以下结论。

1) 轻量化的 RT-DETR,大幅减少冗余参数,推理速度明显提高,减轻了边缘端服务器的运算压力。实验结果表明所提出的轻量级 RT-DETR 检测算法,能够快速且精准地识别常见的电力设备缺陷且鲁棒性好,为后续高效分布式训练提供良好的深度学习基础。

2) 引入联邦学习打破了隐私保护数据的壁垒,实现模型的迭代更新。考虑不同客户端之间设备的差异性导致 Non-IID 数据的存在,提出 FedGSC 算法,根据局部模型的 3 个关键指标生成全局参照梯度,每轮模型更新的梯度按照参照梯度进行修正,缓解 Non-IID 的影响,实验证明此方法能有效地缩减通信轮次,显著提高通信效率。

3) 在泛化性能方面,轻量级 RT-DETR 模型在多种电力设备缺陷检测场景中进行了广泛的实验验证。实验结果表明,模型在不同场景下均能保持高 mAP,显示出良好的泛化能力。表明模型能够适应不同的检测任务,无需大量的场景特定适配工作。

参考文献:

- [1] 林楚. 国家电网: 打造互利共赢新生态融通共建新型电力系统现代产业链[N]. 机电商报, 2024-01-29(A06).
LIN C. State Grid: Creating a new ecology of mutual benefit and win-win cooperation to build a new power system modern industrial chain[N]. Machinery & Electronics Business, 2024-01-29(A06).
- [2] 刘传洋, 吴一全. 基于红外图像的电力设备识别及发热故障诊断方法研究进展[J]. 中国电机工程学报, 2025, 45(6): 2171-2196.
LIU C Y, WU Y Q. Research progress of power equipment identification and thermal fault diagnosis based on infrared images[J]. Proceedings of the CSEE, 2025, 45(6): 2171-2196.
- [3] 李辉, 罗勇, 黄世超, 等. 面向环形管道的电力巡检机器人设计与环境建图[J]. 科学技术与工程, 2024, 24(4): 1546-1552.
LI H, LUO Y, HUANG S C, et al. Design and mapping of electric inspection robot system for annular pipeline[J]. Science Technology and Engineering, 2024, 24(4): 1546-1552.
- [4] 伍颖欣, 刘磊, 肖轶婷, 等. 基于改进注意力机制网络的电力设备图像识别[J]. 中国电机工程学报, 2025, 45(3): 870-883.
WU Y X, LIU L, XIAO Y T, et al. Power device image recognition based on improved attention mechanism[J]. Proceedings of the CSEE, 2025, 45(3): 870-883.
- [5] 刘敏, 周国亮, 王红旭, 等. 基于稀疏重构注意力机制的绝缘子缺陷检测方法[J]. 广东电力, 2024, 37(5): 104-111.
LIU M, ZHOU G L, WANG H X, et al. Insulator defect detection based on sparse reconstruction dual attention mechanism[J]. Guangdong Electric Power, 2024, 37(5): 104-111.
- [6] 张乃雪, 钟羽中, 赵涛, 等. 基于 Smooth-DETR 的产品表面小尺寸缺陷检测算法[J]. 计算机应用研究, 2022, 39(8): 2520-2525.
ZHANG N X, ZHONG Y Z, ZHAO T, et al. Detection method for small-size surface defects based on Smooth-DETR[J]. Application Research of Computers, 2022, 39(8): 2520-2525.
- [7] 李刚, 张运涛, 汪文凯, 等. 采用 DETR 与先验知识融合的输电线路螺栓缺陷检测方法[J]. 图学学报, 2023, 44(3): 438-447.
LI G, ZHANG Y T, WANG W K, et al. Defect detection method of transmission line bolts based on DETR and prior knowledge fusion[J]. Journal of Graphics, 2023, 44(3): 438-447.
- [8] 欧阳名三, 李杰. 基于改进 DETR 的绝缘子缺陷检测算法[J]. 重庆工商大学学报(自然科学版), 2025, 42(1): 20-27.
OUYANG M S, LI J. Insulator defect detection algorithm based on improved DETR[J]. Journal of Chongqing Technology and Business University (Natural Sciences Edition), 2025, 42(1): 20-27.
- [9] 牟元凯. 面向 Non-IID 数据的联邦学习分布式训练优化方法研究[D]. 杭州: 杭州电子科技大学, 2022.
MU Y K. Research on distributed training optimization method of federated learning under non-IID data[D]. Hangzhou: Hangzhou Dianzi University, 2022.
- [10] 唐晓岚, 梁煜婷, 陈文龙. 面向非独立同分布数据的车联网多阶段联邦学习机制[J]. 计算机研究与发展, 2024, 61(9): 2170-2184.
TANG X L, LIANG Y T, CHEN W L. Multi-stage federated learning mechanism with non-IID data in Internet of vehicles[J]. Journal of Computer Research and Develop-

- ment, 2024, 61(9): 2170-2184.
- [11] 曹行健, 孙罡, 虞红芳. 基于Transformer的多模态个性化联邦学习[J]. 电子科技大学学报, 2025, 54(2): 242-249.
CAO X J, SUN G, YU H F. Multimodal personalized federated learning based on Transformer[J]. Journal of University of Electronic Science and Technology of China, 2025, 54(2): 242-249.
- [12] MANASA P, ANANTH P, NATARAJAN P, et al. An analysis of causative factors for road accidents using partition around medoids and hierarchical clustering techniques[J]. Engineering Reports, 2024, 6(6): e12793.
- [13] KARIMIREDDY S P, KALE S, MOHRI M, et al. Scaffold: Stochastic controlled averaging for federated learning[C]//International conference on machine learning. PMLR, 2020: 5132-5143.
- [14] 韩沛秀, 孙卓, 刘忠波, 等. 基于个性化联邦学习的异构船舶航行油耗预测[J]. 计算机集成制造系统, 2025, 31(1): 182-196.
HAN P X, SUN Z, LIU Z B, et al. Heterogeneous ship fuel oil consumption prediction at sea based on personalized federated learning[J]. Computer Integrated Manufacturing Systems, 2025, 31(1): 182-196.
- [15] 张瑞麟, 杜晋华, 尹浩. 跨设备联邦学习中的客户端选择算法[J]. 软件学报, 2024, 35(12): 5725-5740.
ZHANG R L, DU J H, YIN H. Client selection algorithm in cross-device federated learning[J]. Journal of Software, 2024, 35(12): 5725-5740.
- [16] WANG Z, FAN X L, QI J Z, et al. Federated learning with fair averaging[J]. arXiv preprint arXiv, 2021: 2104.14937.
- [17] 姜香菊, 王瑞彤, 马彦鸿. 基于轻量级改进RT-DETR边缘部署算法的绝缘子缺陷检测[J]. 电工技术学报, 2025, 40(3): 842-854.
JIANG X J, WANG R T, MA Y H. Insulator defect detection based on lightweight improved RT-DETR edge deployment algorithm[J]. Transactions of China Electrotechnical Society, 2025, 40(3): 842-854.
- [18] 林德钰, 周卓彤, 过斌, 等. 高斯混合模型与GhostNet结合的YOLO-G遗留物检测方法[J]. 计算机辅助设计与图形学学报, 2023, 35(1): 99-107.
LIN D Y, ZHOU Z T, GUO B, et al. YOLO-G abandoned object detection method combined with Gaussian mixture model and GhostNet[J]. Journal of Computer-Aided Design & Computer Graphics, 2023, 35(1): 99-107.
- [19] 潘沐伽, 陈思光. 面向资源异构的通信高效去中心化联邦学习[J]. 系统工程与电子技术, 2025, 47(4): 1374-1383.
PAN S J, CHEN S G. Communication-efficient decentralized federated learning with resource heterogeneity[J]. Systems Engineering and Electronics, 2025, 47(4): 1374-1383.
- [20] 曹行健, 孙罡, 虞红芳. 基于Transformer的多模态个性化联邦学习[J]. 电子科技大学学报, 2025, 54(2): 242-249.
CAO X J, SUN G, YU H F. Multimodal personalized federated learning based on Transformer[J]. Journal of University of Electronic Science and Technology of China, 2025, 54(2): 242-249.
- [21] 杜江涛, 于家明, 齐辉. 无人机集群不完全信息路径规划方法[J]. 哈尔滨工程大学学报, 2024, 45(11): 2210-2217.
DU J T, YU J M, QI H. Incomplete information path planning method for an UAV cluster[J]. Journal of Harbin Engineering University, 2024, 45(11): 2210-2217.



第一作者: 韦宝泉(1979—),男,教授,博士,硕士生导师,研究方向为电气设备状态监测技术和智能网格技术。E-mail: 48130131@qq.com。



通信作者: 邓芳明(1980—),男,教授,博士,硕士生导师,研究方向为基于人工智能的电力设备状态监测与故障诊断技术、新能源与储能相关技术。E-mail: 2464@ecjtu.edu.cn。

(责任编辑:李根)