



DOI:10.12404/j.issn.1671-1815.2403663

引用格式:洪杨,吴钦木.基于自适应本地融合联邦学习的电机轴承故障诊断[J].科学技术与工程,2025,25(8):3217-3225.

Hong Yang, Wu Qinmu. Motor bearing fault diagnosis based on adaptive local collaboration federated learning[J]. Science Technology and Engineering, 2025, 25(8): 3217-3225.

机械、仪表工业

基于自适应本地融合联邦学习的电机轴承故障诊断

洪杨, 吴钦木*

(贵州大学电气工程学院, 贵阳 550025)

摘要 工业电机轴承的故障诊断对设备性能和寿命至关重要。传统的诊断方法是将多个工厂的数据汇集在一起,这存在数据隐私和标注成本高的问题。为了解决这些问题,提出一种基于自适应本地融合(adaptive local collaboration, ALC)联邦学习的故障诊断策略。在该方法中,不同工况轴承数据将存储于多个客户端,中心服务端与各个客户端协同工作,以建立联邦学习诊断模型。采用改进的 ResNet-18 网络作为分类器,在个性化联邦学习框架下进行训练,ALC 联邦学习方法使每个客户端能有效融合全局和局部模型,提取全局信息优化本地训练结果。实验证明,该方法在保护数据隐私的同时与其他方法相比,提高了故障诊断准确性,特别在多工厂环境中表现出更高的故障分类精度。

关键词 电机轴承; 联邦学习(FL); 自适应本地融合(ALC); 故障诊断

中图分类号 TH17; **文献标志码** A

Motor Bearing Fault Diagnosis Based on Adaptive Local Collaboration Federated Learning

HONG Yang, WU Qin-mu*

(School of Electrical Engineering, Guizhou University, Guiyang 550025, China)

[Abstract] Fault diagnosis of industrial motor bearings is crucial for equipment performance and lifespan. Traditional diagnostic methods aggregate data from multiple factories, leading to issues with data privacy and high annotation costs. To address these problems, a fault diagnosis strategy based on adaptive local collaboration (ALC) federated learning was proposed. In this approach, bearing data under different working conditions was stored across multiple clients, with a central server collaborating with each client to build a federated learning diagnostic model. An improved ResNet-18 network was used as the classifier, which was trained within the personalized federated learning framework. The ALC federated learning method enables each client to effectively integrate global and local models, extracting global information to optimize local training results. Experiments demonstrate that this method enhances fault diagnosis accuracy while protecting data privacy, showing higher fault classification precision compared to other methods, especially in multi-factory environments.

[Keywords] motor bearings; federated learning(FL); adaptive local collaboration(ALC); fault diagnosis

电机轴承作为工业电机的核心旋转部件,广泛应用于电动汽车、风力发电和航天科技中,其健康状况对整个机械系统的性能和寿命具有决定性影响。然而,轴承在运行过程中由于负载、疲劳、润滑不足以及其他操作条件不当等因素,可能会引发一系列故障。因此,研究电机轴承故障诊断技术,对保证电机正常运行具有重要意义^[1]。

常见的故障诊断方法有:基于模型分析、基于信号和数据驱动分析^[2]。基于数据驱动的电机故障诊断近年来被广泛应用^[3]。特别是基于深度学习和电机电流分析的图像识别方法来诊断永磁同步电机故障^[4]。虽然深度学习故障诊断方法可准确识别出故障类别,但是深度学习故障诊断方法需要大量高质量标记数据,在收集和处理这些数据需

收稿日期:2024-05-17; 修订日期:2025-01-03

基金项目:国家自然科学基金(52267003)

第一作者:洪杨(1998—),男,侗族,贵州铜仁人,硕士研究生。研究方向:联邦学习、故障诊断。E-mail:183259651@qq.com。

*通信作者:吴钦木(1976—),男,侗族,贵州铜仁人,博士,教授,博士研究生导师。研究方向:电机控制、深度学习、联邦学习、故障诊断。

E-mail:wqm_watlei@163.com。

要花费大量的时间和精力^[5],这其中还涉及数据共享,以及各个参与方数据隐私等问题,在一定程度上限制了该方法的应用^[6]。

联邦学习(federated learning, FL)是近年来提出的一种分布式机器学习方法,用于解决数据隐私和数据共享的问题,在故障诊断的应用中,联邦学习技术可以帮助实现不同用户之间的协同训练,防止各用户之间隐私泄露^[7]。联邦学习是,客户端不直接共享其持有的敏感个人数据,只有客户端的本地模型更新的权值被安全地传送到中心服务器以进行汇总处理。这种机制使得各客户端能够在确保个人数据隐私的前提下,共同建立高效的协同训练模型。Mehta等^[8]提出开发了一种用于多工厂混合故障诊断的联邦学习框架,对所有参与的工厂产生良好的混合故障分类精度。Zhang等^[9]制定了一种具有动态验证方案的FL方法,该方法在凯斯西储大学(Case Western Reserve University, CWRU)轴承数据集和转向架数据集上显示出有良好的故障分类。Ma等^[10]利用异步更新范式对传统的FL方法进行了扩充,以提高FL处理故障数据的效率。Lu等^[11]提出一种用于分散风力发电机组故障诊断的类不平衡隐私保护联邦学习框架。

上述方法中每进行一次全局训练都要进行一次全局初始化,未考虑各个客户端进行个性化训练,在耗费大量通信时间的同时,也难以实现高精度的故障诊断训练。鉴于此,提出一种基于自适应本地融合(adaptive local collaboration, ALC)的电机轴承分布式故障诊断方法。利用改进ResNet-18网络作为分类器,并且在个性化联邦学习框架下进行训练。自适应本地融合是客户端自适应地将全局模型和局部模型融合到局部目标上,从全局模型中捕获所需的信息,提高故障诊断精度。为分布式电机轴承故障诊断提供了一种高效、精准且具有实际应用价值的解决方案。

1 数据处理

1.1 连续小波变换

连续小波变换(continuous wavelet transform, CWT)^[12],是一种非常有效的时频域分析工具,常用于处理、分析和特征提取信号中的非稳定性和非线性特征,尤其适用于处理电机故障信号。连续小波变换是通过一系列的小波母函数来对信号进行分析,小波母函数可以通过平移和伸缩来改变其位置和尺度,从而匹配信号中的不同特征。

CWT的计算公式为

$$\begin{aligned} T_{\text{CWT},x}(g,h) &= \int_{-\infty}^{+\infty} x(t)\omega_{g,h}(t)dt \\ &= \frac{1}{\sqrt{h}} \int_{-\infty}^{+\infty} x(t)\varphi\left(\frac{t-g}{h}\right)dt \end{aligned} \quad (1)$$

式(1)中: $T_{\text{CWT},x}(g,h)$ 为小波变换系数; g 为平移因子; h 为尺度因子; $x(t)$ 为被分析的信号; t 为时间变量; $\omega_{g,h}(t)$ 为小波基函数,其表达式为

$$\omega_{g,h}(t) = \frac{1}{\sqrt{h}} \varphi\left(\frac{t-g}{h}\right), \quad h > 0; g \in \mathbf{R} \quad (2)$$

1.2 实验数据集处理

采用的数据来自凯斯西储大学(CRWU)轴承试验台采集的数据^[13]。试验台由电机、传感器、轴承和负载组成如图1所示。

采用驱动端为SKF6205规格轴承,风扇端为SKF6203规格轴承,设置传感器采样频率均为12 kHz,驱动端电机和轴承分别在1 772 r/min、1 750 r/min和1 730 r/min下运行。驱动端和风扇端轴承在上述负载条件下运行,使用传感器采集轴承的振动信号。包含正常(N)、内圈故障(IR)、外圈故障(OR)和滚动体(B)4种工况,3种不同程度故障。总共得到1类正常故障数据和9类不同故障数据通过可视化得到振动信号波形,部分波形如图2所示。

基于一定的取样样本点间隔,将信号数据分成多个窗口,每个窗口包含一定数量的数据点,将每个窗口中的数据实施CWT,在不同的平移因子和尺度因子重复计算得到小波系数矩阵和对应的频率,小波系数矩阵是一个复数矩阵,其实部和虚部分别表示信号在不同尺度和时间位置上的局部特性。通常取其绝对值,得到振幅矩阵,代表信号在不同尺度和时间位置上的强度或能量分布,来生成时间-频率图像。设置以下参数,图片大小为128×128,取样间隔为30个样本点,由于每类有12 000个样本点,通过连续小波变换图片数据集样本数如表1所示,得到的部分二维图像如图3所示。

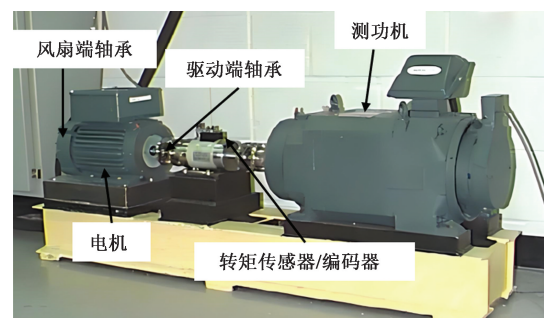


图1 轴承故障数据采集试验台

Fig. 1 Bearing fault data acquisition experimental bench

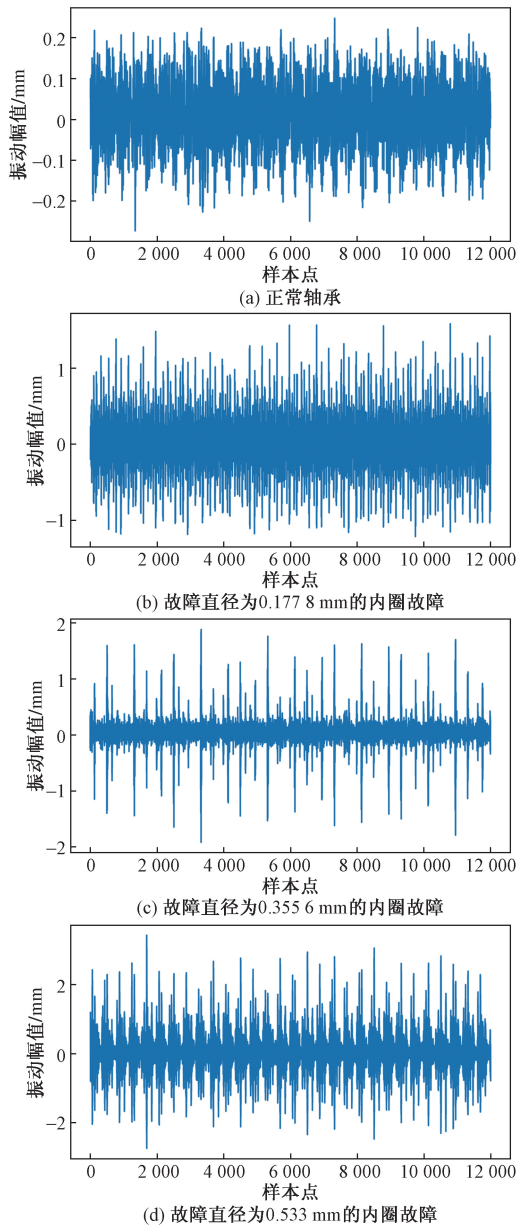


图2 不同轴承故障的信号图

Fig. 2 Signal diagram of different bearing faults

表1 生成的图片数据集样本数

Table 1 Number of samples in the generated image dataset

轴承故障名称	标签	样本数量
正常轴承	de_normal	936
故障直径为 0.177 8 mm 滚动体故障	de_7_inner	936
故障直径为 0.177 8 mm 内圈故障	de_7_ball	936
故障直径为 0.177 8 mm 外圈故障	de_7_outer	936
故障直径为 0.355 6 mm 滚动体故障	de_14_inner	936
故障直径为 0.355 6 mm 内圈故障	de_14_ball	936
故障直径为 0.355 6 mm 外圈故障	de_14_outer	936
故障直径为 0.533 mm 滚动体故障	de_21_inner	936
故障直径为 0.533 mm 内圈故障	de_21_ball	936
故障直径为 0.533 mm 外圈故障	de_21_outer	936

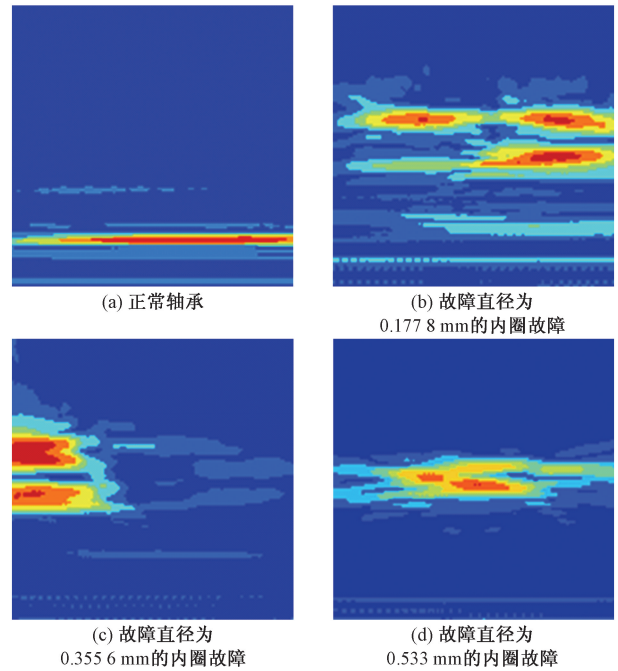


图3 不同轴承故障的时频信号图

Fig. 3 Time domain signal diagrams of different bearing faults

2 联邦学习模型

2.1 传统联邦学习

联邦学习是一种新的机器学习方法,它允许多个参与者协同训练一个共享模型,同时保持各自数据的隐私和安全。在这种训练过程中,原始数据保留在本地,无需直接共享数据,这主要通过将模型训练过程分布到多个参与者上来实现。参与者各自在自己的数据上训练模型,然后仅共享模型更新,例如权重或梯度,而非原始数据。这些更新随后用于改进中心模型或共享模型。联邦学习模型所涉及的步骤如图4所示,具体如下。

步骤1 中心服务器对全局模型进行初始化,并且将模型下发给客户端共享该模型。

步骤2 每个客户端在自己的数据集上独立训练模型,更新模型参数。

步骤3 每个客户端将其模型更新(神经网络的权值)发送回中心服务器。

步骤4 中心服务器聚合这些更新以产生新的全局模型参数。

重复上述步骤,直到达到全局模型收敛准则,因此联邦学习都是在本地客户端进行,有效的防止了数据的隐私泄露。

联邦平均算法是从本地客户端数据学习全局机器学习模型的核心。假设数据集划分给 m 个客户端, D_i 为客户端 i 上的联邦学习数据集,其中每个数据由故障的图片信息和对应的故障标签相组成,

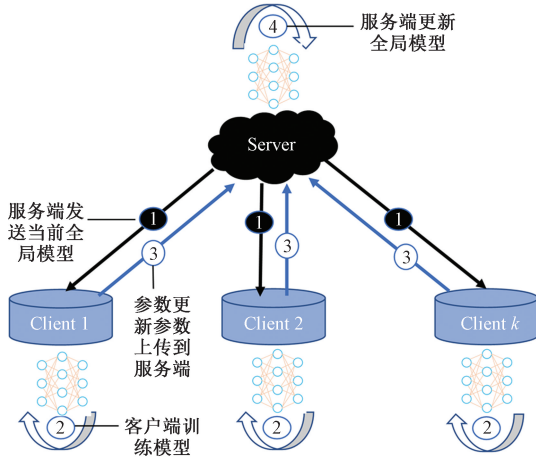


图4 联邦学习训练程序示意图

Fig. 4 Schematic diagram of Federated Learning training program

设 n_i 为第 i 个客户的数据信息, $n = \sum_{i=1}^m n_i$ 为联邦学习中训练的总的的数据信息。目标是使全局目标最小化,得到合理的局部模型可表示为

$$\{\theta_1, \theta_2, \dots, \theta_m\} = \operatorname{argmin} \left(\sum_m \frac{n_m}{n} L_m \right) \quad (3)$$

式(3)中: $L_m = \frac{1}{n_m} \sum_{i \in D_m} l(x_i, y_i; \theta_m)$, 其中, x_i, y_i 为输入数据和标签, $l(\cdot)$ 为每个客户端所使用的损失函数。

设 θ_m 为全局模型集合, D_m 为客户端数据集合, c 为客户端比例, E 为本地迭代次数, η 为本地学习率, B 为本地批量大小, S 为全局模型训练次数。联邦平均算法 (FedAvg)^[14] 的伪代码如下。其中 ∇_{θ} 为对函数 l 中的 θ 求偏导。

算法 1 (FedAvg)

服务端:

1. 初始化全局模型参数 θ^0 ;
2. for 全局模型执行 $t=1, 2, \dots, S$ 次更新;
3. 从客户端 c_m 中随机选择 m_i , 将当前全局模型 θ^{t-1} 发送到所选客户端;
4. for 对客户端 $i \in m_i$, 并行执行;
5. $\theta_i^t \leftarrow$ 客户端更新 (i, θ^{t-1});
6. 接收从客户端 i 发送的 θ_i^t ;
7. 计算加权平均并更新全局模型

$$\theta^t \leftarrow \sum_{i=1}^m \frac{n_i}{n} \theta_i^t$$

客户端: 在第 i 个客户端上进行更新 (i, θ);

8. 用 θ 初始化本地模型;
9. for 本地模型执行从 1 到迭代次数 E 每次迭代;
对数据集 D_i 进行分片, 分为大小为 B 的 batches;
10. for 对批量数 B 中的批量序号 b 执行
 $\theta \leftarrow \theta - \eta \nabla_{\theta} l(b)$

2.2 个性化联邦学习

个性化联邦学习分为以下三类:①学习单个全局模型并对其进行微调的方法;②学习附加个性化

模型的方法;③通过个性化(局部)聚合学习局部模型的方法。Zhang 等^[15]提出了一种自适应局部聚合 (ALA) 模块,该模块在联邦学习中可以自适应地将下载的全局模型和局部模型聚合到每个客户端的局部目标上,以便在每次迭代训练之前初始化局部模型。因此将此算法应用到电机故障诊断,来提高诊断精度。

在传统的联邦学习中,在全局训练过程中,服务器下发的模型将覆盖旧的模型,使用自适应局部聚合模块聚合全局模型和本地模型,可表示为

$$\hat{\theta}_i^t \leftarrow \theta^{t-1} \odot W_{i,l} + \theta^{t-1} \odot W_{i,g} \quad (4)$$

式(4)中: $\hat{\theta}_i^t$ 为覆盖旧的局部模型得到初始化的局部模型; \odot 为哈达玛积运算; $W_{i,l}$ 和 $W_{i,g}$ 为聚合权重。

$w_{i,l}^q + w_{i,g}^q = 1$, 其中 $w_{i,l}^q$ 与 $w_{i,g}^q$ 为各模型中第 q 个元素的权重,且都大于 0。通过基于梯度的学习方法很难学习到带有约束的 $W_{i,l}, W_{i,g}$ 。因此将其合并为 W_i 得

$$\hat{\theta}_i^t \leftarrow \theta^{t-1} + (\theta^{t-1} - \theta_i^{t-1}) \odot W_i \quad (5)$$

式(5)中: $\theta^{t-1} - \theta_i^{t-1}$ 实现模型更新。

将使用元素加权裁剪 $\sigma(w) = \max[0, \min(1, w)]$ 进行正则化,一般情况下,深度神经网络模型,浅层的网络用于学习普通的故障特征,深层的网络负责提取更复杂的故障特征。为了减少计算开销,引入一个超参数 p 来控制 ALA 的范围,将其应用于 p 个较高的层,并覆盖联邦平均算法等较低层的参数进行局部初始化式(6)。

$$\hat{\theta}_i^t \leftarrow \theta^{t-1} + (\theta^{t-1} - \theta_i^{t-1}) \odot [1^{\lfloor \theta_i \rfloor^p}; W_i^p] \quad (6)$$

式(6)中: $\lfloor \theta_i \rfloor$ 为模型总层数; p 为模型中深层网络结构层数; $[1^{\lfloor \theta_i \rfloor^p}; W_i^p]$ 为一个自适应因子函数,初始化 W_i^p 中的每个元素的值为 1,在每次迭代中在旧的 W_i^p 中学习 W_i^p 。

为了减少计算开销,在第 t 次迭代中随机选取 D_i 中的 s (单位: %), 因此基于随机梯度下降的训练 W_i^p 可表示为

$$W_i^p \leftarrow W_i^p - \eta \nabla_{W_i^p} l(\hat{\theta}_i^t, D_i^{s,t}; \theta^{t-1}) \quad (7)$$

式(7)中: η 为学习率权重; $\nabla_{W_i^p}$ 为对 l 中的 W_i^p 求偏导; $D_i^{s,t}$ 为在第 t 次迭代中随机选取 D_i 中的 $s, \%$ 。

根据式(6)、式(7)可得

$$\nabla_{W_i^p} l_i^t = \eta (\theta^{t-1} - \theta_i^{t-1}) \odot \nabla_{\theta} l_i^t \quad (8)$$

式(8)中: $\nabla_{W_i^p}$ 为对函数 l_i^t 中的 W_i^p 求偏导; ∇_{θ} 为对函数 l_i^t 中的 $\hat{\theta}$ 求偏导。

l_i^t 表示 $l(\hat{\theta}_i^t, D_i^{s,t}; \theta^{t-1})$, 可推导得

$$\hat{\theta}_i^t \leftarrow \theta_i^{t-1} - \eta (\theta^{t-1} - \theta_i^{t-1}) \odot (\theta^{t-1} - \theta_i^{t-1}) \odot \nabla_{\theta} l_i^t \quad (9)$$

自适应局部融合联邦学习算法(ALCFL)的伪代码如下。其中 $\nabla_{\hat{\theta}}$ 为对函数 l_i 中的 $\hat{\theta}$ 求偏导。

算法 2 (ALCFL)

1. 初始化全局模型参数 θ^0 ;
 2. 客户端初始化 W_i^0 ; 3. for 全局模型执行 $t=1, 2, \dots, S$ 次更新;
 4. 从客户端 cm 中随机选择 m_t , 将当前全局模型 θ^{t-1} 发送到所选客户端;
 5. for 对客户端 $i \in m_t$, 并行执行;
 6. 客户端 i 采样 $s\%$ 的本地数据;
 7. 客户端 i 由式(7)训练 W_i^t ;
- 并且使用裁剪函数 $\sigma(\cdot)$ 对 W_i^t 进行裁剪;
8. 客户端 i 通过使用式(6)得到 $\hat{\theta}_i^t$;
- 客户端 i 通过本地模型训练获得 θ_i^t

$$\theta_i^t \leftarrow \hat{\theta}_i^t - \eta \nabla_{\hat{\theta}} l(\hat{\theta}_i^t, D_i^{t,t}, \theta^{t-1})$$

9. 客户端 i 发送 θ_i^t 到服务端;
10. 服务端计算加权平均并更新全局模型;

$$\theta^t \leftarrow \sum_{i=1}^m \frac{n_i}{n} \theta_i^t$$

11. return $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_N$;

2.3 网络结构

由于连续小波变换得到的时频图能够揭示细微的故障特征变化,为了有效地从图像中提取不同类型的电机故障特征。在 ResNet-18 网络中,引入注意力机制模块,该模块能够聚焦于输入故障特征中的重要信息,从而提高故障诊断的准确率。Woo 等^[16]提出 CBAM 注意力机制模块,CBAM 依次应用两种类型的注意力机制,通道注意力和空间注意力处理特征图。在最后一个卷积层后面添加一个

CBAM 注意力机制模块。网络结构如图 5 所示。

3 实验结果及分析

3.1 数据分布

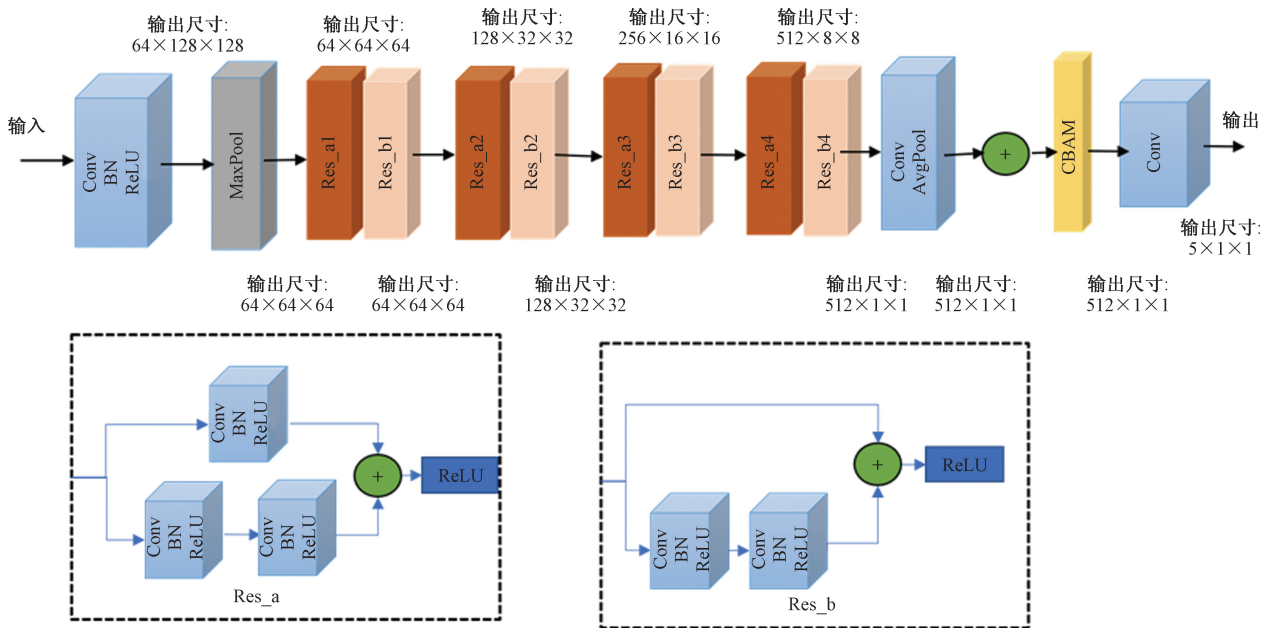
对预处理后的数据集,使用迪利克雷函数进行划分得到非平衡非独立同分布数据集,记为 $\text{Dir}(a)$, 其中 a 为控制数据分布均匀程度的关键变量,越小分布越不均匀,通过两种不同的方式将整个数据集划分给参与训练的 10 个客户端,其中训练集与测试集的比例为 3:1。通过不同分区方法评估联邦学习方法在各个客户端的本地故障可视化。

3.1.1 数据平衡独立同分布(balanced IID)

方案一是模拟理想情况下的数据分布方式,这意味着不同类型的故障在数据集中的样本数量相等,同时数据集中的每个样本都是独立抽取的,并且遵循相同的分布。如图 6(a)所示,每个客户端都具有 10 种不同工况,同时数据分布情况基本相同。此场景为联邦学习训练提供了一个在理想实验场景下训练任务。

3.1.2 数据非平衡非独立同分布(Unbalanced non-IID)

方案二模拟真实的工业场景(不同故障类型与正常状态的样本数量是不均等的),同时,因为来源于不同的操作条件、环境或时间段,同一类别内的样本可能呈现出不同的分布特性。此分布的特征参数为 $\text{Dir}(0.5)$,如图 6(b)所示,每个客户端的数据分布大小为 500 ~ 1 400,同时数据是不平衡的,如



MaxPooling为最大值池化; Conv为卷积; AvgPool为平均池化; ReLU为激活函数; CAMB为卷积块注意力模块; BN为BatchNorm

图 5 网络结构

Fig. 5 Network architecture

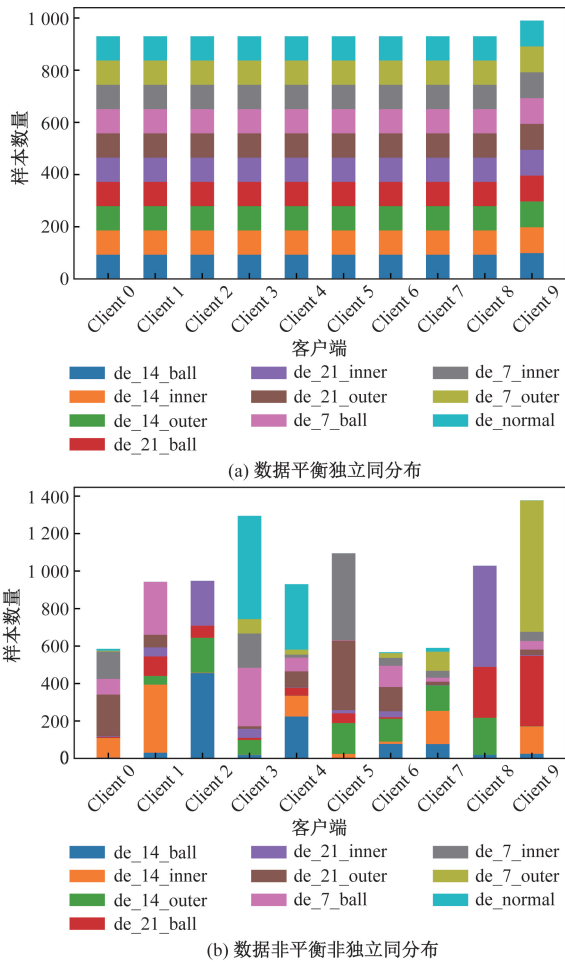


图6 数据分布图

Fig. 6 Data distribution chart

客户端2只有5种工况数据,但客户端9,存在10种工况数据。此场景模拟了在实际工业联盟中可能遇到的一种极端但更现实的情况。

在联邦学习框架下,对上述两种数据分布情况使用不同的算法和网络模型进行全局训练。客户端数量为10,全局训练200次,每轮全局模型本地训练1次,本地学习率设置为0.001,批次大小设置为10,在每个服务器轮结束时,在测试集上评估全局模型的性能。

3.2 结果与分析

3.2.1 超参数影响分析

为了研究超参数 p 和 s 对 ALC 性能的影响,将不同的参数训练得到的平均值进行了对比分析,数据集采用非平衡非独立同分布。由图7和表2可知,超参数 $p = 2, s = 80$ 时 ALC 故障诊断性能达到最优效果:准确率达到97.78%。因此,超参数 p 和在第 t 次迭代中随机选取 D_i 中的 $s\%$,可以控制 ALC 的范围,通过图7、表2可知,设置大的 s 和合适的 p ,可以提高 ALCFL 的故障诊断的精度。

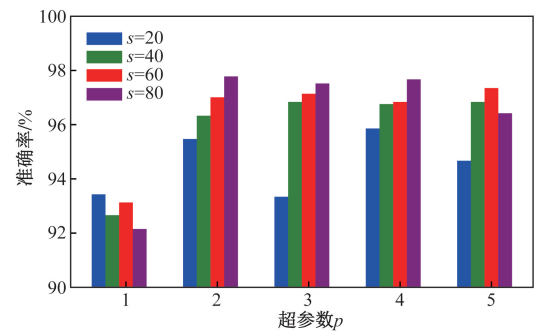


图7 不同超参数故障诊断精度对比

Fig. 7 Comparison of fault diagnosis accuracy with different hyperparameters

表2 不同超参数故障诊断精度
Table 2 Fault diagnosis accuracy with different hyperparameters

p	准确率/%			
	$s = 20$	$s = 40$	$s = 60$	$s = 80$
1	93.43	92.66	93.13	92.15
2	95.47	96.66	97.01	97.78
3	93.34	96.93	97.14	97.52
4	95.86	96.76	96.84	97.67
5	94.67	96.84	97.35	96.42

3.2.2 网络模型对比分析

在 ALCFL 方法中,进行改进的 ResNet-18 网络模型与其他网络模型,如 CNN, Alexnet, ResNet-18, ResNet-10, Mobilenet 在上述最优超参数下,进行对比实验。由图8可知,其他几类除 CNN 网络模型外均有较好的诊断效果,但改进的 ResNet-18 较其他网络模型,有更高的诊断精度,因此选用改进的 ResNet-18 作为本实验的网络模型。

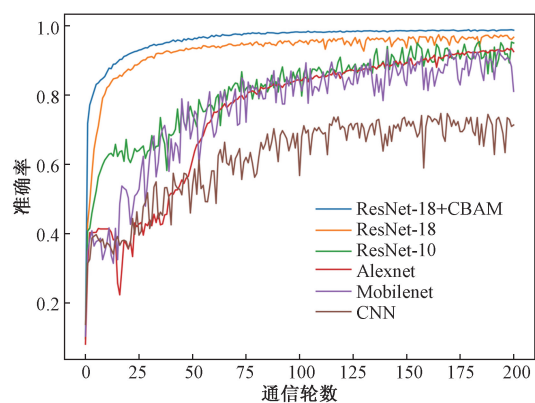


图8 不同网络模型训练结果精度对比

Fig. 8 Different network models training results accuracy comparison

3.2.3 ALC 性能评估

分别使用两种不同分布的数据集进行 ALCFL 方法训练,图9为在上述最优超参数下网络的诊断

精度。其中精度为 10 个客户端的平均值。由图 9 可知,两种不同数据划分情况在 ALCFL 方法下训练的过程中随着全局训练次数的增加而增加。通过 200 次训练结果可知,对于以上两种数据分布,使用 ALCFL 方法实现了 90% 以上的故障分类准确率。两种数据分布的训练结果曲线图和训练结果混淆矩阵如图 9 和图 10 所示。

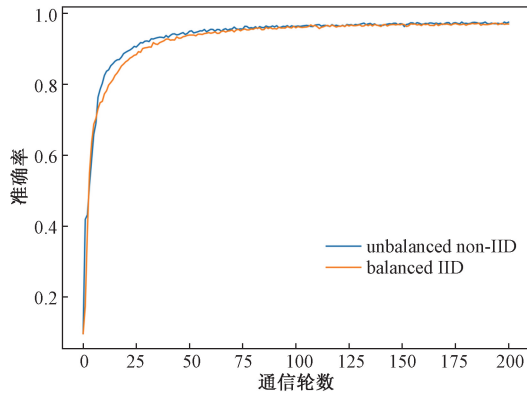


图 9 两种数据分布训练结果的精度

Fig. 9 Accuracy of training results with two different data distributions

3.2.4 算法对比分析

为了验证 ALCFL 方法的优越性,将非平衡非独立同分布数据集的参数分别设置为 0.1、0.3、0.5、0.7。使用 ALCFL 方法与联邦平均算法 (FedAvg)^[15];自适应个性化联邦学习算法 (APFL)^[17];基于正则化的联邦学习算法 (FedDyn)^[18];一阶模型优化个性化联邦学习算法 (FedFomo)^[19];关于非 IID 数据的个性化跨筒仓联邦学习 (FedAMP)^[20] 等算法进行训练对比。全局通信轮数均为 200 次,其他参数设置均相同。表 3 为不同方法的故障诊断精度。图 11 为 Dir(0.1) 时各模型训练精度。

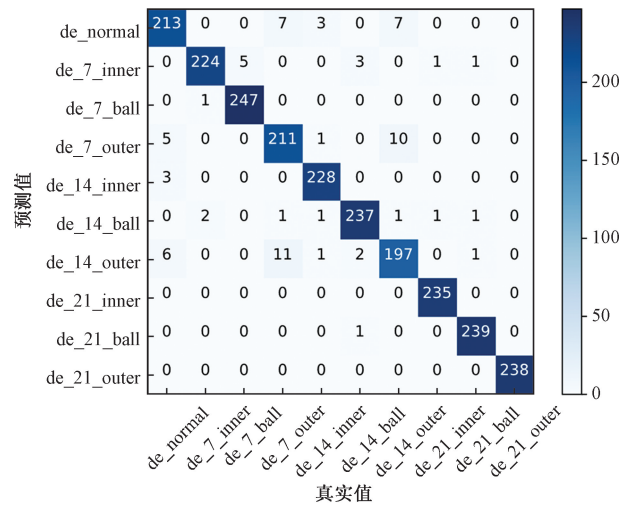
FedAvg 方法在数据异构较大的情况下,诊断效果不是很好,平均诊断精度只有 67.15%,但随着异构差异较小时,诊断效果得到显著提升,平均诊断精度可达到 91.31%。当数据异构性较大时,聚合后的全局模型可能会偏向于那些数据量大或者特

表 3 不同方法的故障诊断精度

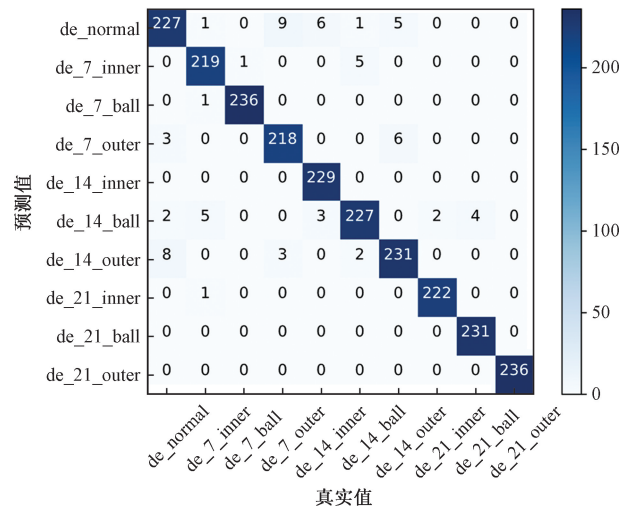
Table 3 Fault diagnosis accuracy of different methods

模型	故障诊断精度/%			
	Dir(0.1)	Dir(0.3)	Dir(0.5)	Dir(0.7)
ALCFL	98.89	98.29	97.78	97.10
FedAvg	67.15	86.72	89.55	91.31
APFL	97.95	96.71	95.57	94.67
FedDyn	98.46	98.38	98.31	98.21
FedFomo	95.46	94.58	92.36	91.09
FedAMP	98.12	96.79	95.64	94.33

征更加突出的客户端,而忽视了少数客户端的特性。这种偏差会使得模型在一些客户端上的表现较好,而在其他客户端上表现不佳。



(a) 数据平衡独立同分布训练结果混淆矩阵



(b) 数据非平衡非独立同分布训练结果混淆矩阵

图 10 不同分布训练结果混淆矩阵

Fig. 10 Confusion matrix of training results with different distributions

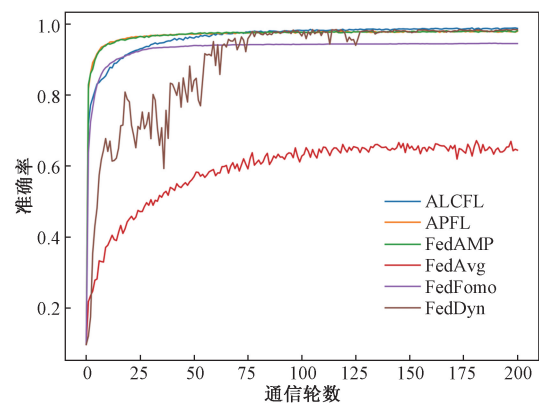


图 11 Dir(0.1) 时各模型训练精度

Fig. 11 Training accuracy of each model at Dir(0.1)

APFL方法、FedDyn方法、FedFomo方法、FedAMP方法,在解决数据异构差异较大的数据时,平均诊断效果都能够达到90%以上,比FedAvg方法更加适用于数据分布不一致、数据异构较大的实验场景。

使用ALCFL方法进行故障诊断对于Dir(0.1)时平均诊断精度可以达到98.89%,对于Dir(0.7)时平均诊断精度可以达到97.10%。与上述其他个性化联邦学习方法相比,解决数据分布不一致、数据异构较大的故障诊断场景有着显著提升。ALCFL方法可以根据每个客户端的数据特性调整本地模型更新的学习率;可以帮助模型更好地适应各个客户端的数据分布,尤其是在数据分布差异大的情况下,通过优化每个客户的学习进程,提升整体学习效率 and 模型性能。

4 结论

提出结合改进的ResNet-18网络和适应本地融合联邦学习框架,得出以下结论。

(1)设置合适的超参数使ALCFL方法故障诊断性能达到最优效果,提高故障诊断的精度。

(2)在ALCFL方法中,改进的ResNet-18网络模型与其他网络模型在最优超参数下进行对比。实验结果显示,使用改进的ResNet-18网络诊断的精度明显高于其他网络。

(3)在两种不同分布的数据集进行ALCFL方法训练。实验结果显示,两种不同数据分布情况下,随着全局训练轮数的增加,故障分类准确率均超过90%。

(4)通过对比分析,可知ALCFL方法在处理大规模分布式数据时,相较于其他联邦学习算法,能显示出更高的故障诊断精度,在Dir(0.1)条件下的平均诊断精度可达到98.89%,这表明ALCFL方法在电动汽车、风力发电和航天科技等实际工业应用中有很大潜力,特别是在需要处理敏感数据和保证数据隐私的情况下。所提出的ALCFL方法为电机轴承故障诊断提供了一个新的视角和工具,能有效应对现代工业系统中的数据隐私和故障诊断挑战。

参 考 文 献

[1] 郭娜,王健,杨建伟,等. 变工况及小样本情况下滚动轴承故障迁移学习方法综述[J]. 科学技术与工程, 2024, 24(10): 3939-3951.
Wu Na, Wang Jian, Yang Jianwei, et al. A review of transfer learning methods for rolling bearing faults under variable working conditions and small samples[J]. Science Technology and Engi-

neering, 2024, 24(10): 3939-3951.
[2] 何群,余志红,陈志刚,等. 基于改进信息熵和LSTM网络的轴承故障诊断[J]. 科学技术与工程, 2024, 24(12): 4969-4975.
He Qun, Yu Zhihong, Chen Zhigang, et al. Bearing fault diagnosis based on improved information entropy and LSTM network[J]. Science Technology and Engineering, 2024, 24(12): 4969-4975.
[3] Guo Y, Wu Q. Visualization for explanation of deep learning-based fault diagnosis model using class activation map[J]. CMC-Computers Materials & Continua, 2023, 77(2): 1489-1514.
[4] Li Z, Wu Q, Yang S, et al. Diagnosis of rotor demagnetization and eccentricity faults for IPMSM based on deep CNN and image recognition[J]. Complex & Intelligent Systems, 2022, 8(6): 5469-5488.
[5] Mehta M, Shao C. Federated learning-based semantic segmentation for pixel-wise defect detection in additive manufacturing[J]. Journal of Manufacturing Systems, 2022, 64: 197-210.
[6] Jiang L, Tang J, Sun N, et al. A hybrid deep learning model for fault diagnosis of rolling bearings using raw vibration signals[J]. Measurement Science and Technology, 2024, 35(9): DOI:10.1088/1361-6501/ad3fd2.
[7] Sanchita S, Ashlesha H, Kumar A C, et al. A multifaceted survey on privacy preservation of federated learning: progress, challenges, and opportunities[J]. Artificial Intelligence Review, 2024, 57(7): DOI:10.1007/s10462-024-10766-7.
[8] Mehta M, Chen S, Tang H, et al. A federated learning approach to mixed fault diagnosis in rotating machinery[J]. Journal of Manufacturing Systems, 2023, 68: 687-694.
[9] Zhang W, Li X, Ma H, et al. Federated learning for machinery fault diagnosis with dynamic validation and self-supervision[J]. Knowledge-Based Systems, 2021, 213: 106679.
[10] Ma X, Wen C, Wen T. An asynchronous and real-time update paradigm of federated learning for fault diagnosis[J]. IEEE Transactions on Industrial Informatics, 2021, 17(12): 8531-8540.
[11] Lu S, Gao Z, Xu Q, et al. Class-imbalance privacy-preserving federated learning for decentralized fault diagnosis with biometric authentication[J]. IEEE Transactions on Industrial Informatics, 2022, 18(12): 9101-9111.
[12] 陈代俊,陈里里,董绍江. 基于VMD-CWT-CNN的滚动轴承故障诊断[J]. 机械强度, 2023, 45(6): 1280-1285.
Chen Daijun, Chen Lili, Dong Shaojiang. Fault diagnosis of rolling bearings based on VMD-CWT-CNN[J]. Mechanical Strength, 2023, 45(6): 1280-1285.
[13] Smith A W, Randall B R. Rolling element bearing diagnostics using the case Western Reserve University data: a benchmark study[J]. Mechanical Systems and Signal Processing, 2015, 64/65: 100-131.
[14] McMahan B, Moore E, Ramage D, et al. Communication-efficient learning of deep networks from decentralized data[C]//The Proceedings of 20th International Conference on Artificial Intelligence and Statistics. Fort Lauderdale: PMLR, 2017: 1273-1282.
[15] Zhang J, Hua Y, Wang H, et al. Fedala: adaptive local aggregation for personalized federated learning[C]//The Proceedings of 37th the AAAI Conference on Artificial Intelligence. Washington,

- D. C. : AAAI, 2023: 11237-11244.
- [16] Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module [C]//The Proceedings of 15th the European Conference on Computer Vision (ECCV). Munich: AAAI, 2018: 3-19.
- [17] Ren P, Qi K, Li J, et al. CosPer: an adaptive personalized approach for enhancing fairness and robustness of federated learning [J]. Information Sciences, 2024, 675: 120760.
- [18] Acar D A E, Zhao Y, Navarro R M, et al. Federated learning based on dynamic regularization [J]. arXiv Preprint, 2021: arXiv:2111.04263.
- [19] Zhang M, Sapra K, Fidler S, et al. Personalized federated learning with first order model optimization [J]. arXiv Preprint, 2020: arXiv: 2012.08565.
- [20] Huang Y, Chu L, Zhou Z, et al. Personalized cross-silo federated learning on non-iid data [C]//The Proceedings of 35th the AAAI Conference on Artificial Intelligence. Washington, D. C. : AAAI, 2021: 7865-7873.