

本文引用格式: 黄艺航,蔡凯武,黄晓智,等.基于机器学习的通信电缆故障检测与定位方法[J].自动化与信息工程, 2024,45(1): 35-41;60.

HUANG Yihang, CAI Kaiwu, HUANG Xiaozhi, et. al. Fault detection and localization method of communication cables based on machine learning[J]. Automation & Information Engineering, 2024,45(1):35-41;60.

基于机器学习的通信电缆故障检测与定位方法

黄艺航¹ 蔡凯武² 黄晓智¹ 袁澄¹ 梁恩源¹ 林智海¹

(1.广东工业大学机电工程学院, 广东 广州 510006

2.广东工业大学先进制造学院, 广东 揭阳 522000)

摘要: 为解决传统的通信电缆故障检测与定位方法存在的灵敏性不足和智能化程度低等问题, 提出基于机器学习的通信电缆故障检测与定位方法。首先, 基于行波法检测原理搭建通信电缆故障仿真模型来采集实验数据样本; 然后, 提出基于粒子群优化-支持向量机(PSO-SVM)的通信电缆故障检测模型, 其故障识别准确率达 99.4%; 接着, 提出基于卷积神经网络-长短时记忆(CNN-LSTM)的通信电缆故障定位模型, 该模型对故障点定位的平均绝对误差为 0.334 9, 均方根误差为 0.320 8; 最后, 通过对比实验验证 CNN-LSTM 的网络准确率较单独使用 CNN 和 LSTM 模型分别提高了 9.47%和 6.2%。

关键词: PSO-SVM 模型; CNN-LSTM 模型; 行波法; 通信电缆; 故障检测; 故障定位

中图分类号: TN391.4

文献标志码: A

文章编号: 1674-2605(2024)01-0006-08

DOI: 10.3969/j.issn.1674-2605.2024.01.006

Fault Detection and Localization Method of Communication Cables Based on Machine Learning

HUANG Yihang¹ CAI Kaiwu² HUANG Xiaozhi¹ YUAN Cheng¹
LIANG Enyuan¹ LIN Zhihai¹

(1.School of Electromechanical Engineering, Guangdong University of Technology, Guangzhou 510006, China

2.School of Advanced Manufacturing, Guangdong University of Technology, Jieyang, 522000, China)

Abstract: To address the issues of insufficient sensitivity and low intelligence in traditional communication cable fault detection and localization methods, a fault detection and localization method of communication cables based on machine learning is proposed. Firstly, based on the principle of traveling wave detection, a communication cable fault simulation model is constructed to collect experimental data samples; Then, a communication cable fault detection model based on Particle Swarm Optimization Support Vector Machine (PSO-SVM) is proposed, with a fault recognition accuracy of 99.4%; Next, a communication cable fault location model based on Convolutional Neural Network Long Short Term Memory (CNN-LSTM) is proposed. The average absolute error of the model for fault location is 0.334 9, and the root mean square error is 0.320 8; Finally, through comparative experiments, it was verified that the network accuracy of CNN-LSTM was 9.47% and 6.2% higher than that of using CNN and LSTM models alone, respectively.

Keywords: PSO-SVM model; CNN-LSTM model; traveling wave method; communication cable; fault detection; fault localization

0 引言

随着信息社会的进步, 通信电缆作为信息传输的

主要媒介, 其稳定性和可靠性非常重要。然而, 许多因素可能导致通信电缆出现故障, 如机械损伤、断线、

接头故障、绝缘老化、水淹和电磁干扰等，从而产生通信中断、数据传输错误或网络拥堵等问题。因此，通信电缆故障检测与定位具有重要意义。

传统的通信电缆故障检测与定位方法可分为阻抗法和行波法。其中，阻抗法通过通信电缆的电流、电压、故障距离来确定故障的类型和位置；行波法基于行波放射理论，向通信电缆发射脉冲，通过反射脉冲波的极性和时间来进行故障的定性及定位^[1-2]。但這些方法依靠人工经验，其准确性与效率难以保证。

采用机器学习技术进行通信电缆的故障检测与定位，可更好地应对故障信号的多样性和复杂性，提高故障检测与定位的准确性和效率^[3]。然而这类方法需要大量的样本数据，且采用不同的模型和参数配置，故障检测与定位的效果也存在差异^[4]。

为此，本文通过搭建通信电缆故障仿真模型来采集足够多的样本数据；提出基于粒子群优化（particle swarm optimization, PSO）-支持向量机（support vector machine, SVM）的通信电缆故障检测模型，实现通信电缆的故障检测及故障类型分类；提出基于卷积神经网络（convolutional neural networks, CNN）-长短时记忆（long short-term memory, LSTM）网络的通信电缆故障定位模型，实现通信电缆故障点的定位。

1 通信电缆故障分类

通信电缆故障的主要原因是电缆芯线的绝缘阻抗发生变化，即故障点处的等效阻抗与电缆的特性阻抗不同^[5]。通信电缆的故障主要分为短路、断路、高阻故障（高阻泄露、高阻闪络）等^[6]，如表 1 所示。其中， f 为绝缘阻抗， n 为额定绝缘阻抗， c 为特性阻抗。

表 1 通信电缆故障类型及阻抗特征

故障类型	绝缘阻抗值 X_f	连续性
短路	$X_f < 10X_c$	连续
断路	$X_f = X_c$	不连续
高阻泄露	$10X_c < X_f < X_n$	连续
高阻闪络	与电压呈线性关系	不连续

短路和断路是通信电缆的基础故障，故障点处阻抗不匹配现象较为明显；高阻故障（高阻泄露、高阻

闪络）是通信电缆的复杂故障，需通过特定的方法进行检测与定位。

2 通信电缆故障仿真模型

本文采用 MATLAB R2022b 仿真软件建立的通信电缆故障仿真模型来采集数据样本。通信电缆故障仿真模型由脉冲信号源发生器、测量元件、传输线元件和故障信号产生模块组成，如图 1 所示。

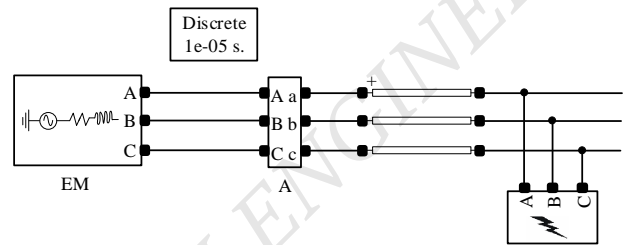


图 1 通信电缆故障仿真模型

脉冲信号源发生器输出正弦波信号，频率为 50 Hz。测量元件采集发射脉冲信号、反射脉冲信号的数据，并调用 Scope 工具实现行波数据可视化。传输线元件用于模拟通信电缆。故障信号产生模块用于产生短路、断路及高阻故障信号，并通过设置故障信号发生点、传输线的位置及发生时间，模拟通信电缆故障发生的位置及故障信号在传输过程中的时序特性。

3 基于 PSO-SVM 的通信电缆故障检测模型

3.1 行波法检测原理

通信电缆发生故障时，故障点处会出现阻抗不匹配现象。向故障通信电缆线一端发射低压脉冲信号，当脉冲信号到达故障点（断路点、短路点、电缆老化和中间接头）时会发生反射，导致故障点处电流、电压的传播方向和幅值发生变化。行波法通过检测反射脉冲信号的极性可以进行故障分类；通过发射脉冲信号、反射脉冲信号回到入射点位置的时间可以进行故障点定位^[7]。断路、短路故障时通信电缆的发射脉冲信号与反射脉冲信号的波形如图 2 所示。

对通信电缆一端发射脉冲信号时，若在入射点处检测到反射脉冲信号，且与发射脉冲信号极性相同，则表示大阻抗，可能是断路或高阻故障；若与发射脉

冲信号极性不同,则表示小阻抗,可能是短路故障^[8]。

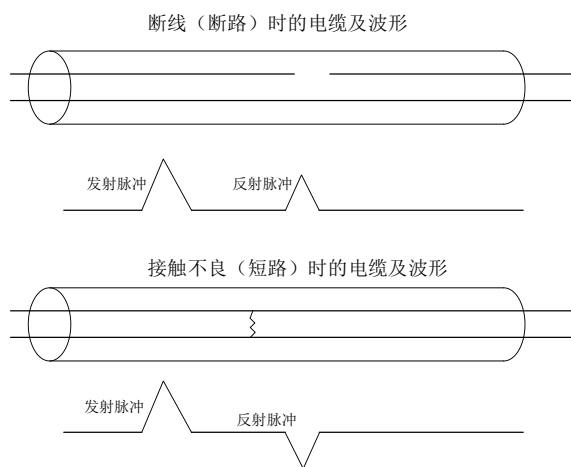


图2 断路、短路脉冲信号的波形图

根据脉冲信号在通信电缆中的传播速度 V 和发射脉冲信号与反射脉冲信号到达入射点的时间差 Δt , 计算故障点与入射点的距离 L 为

$$L = V \frac{\Delta t}{2} \quad (1)$$

当脉冲信号存在时间宽度 τ 时, 如果反射脉冲信号与发射脉冲信号在 τ 内重叠, 会导致发射点的信号接收器无法区分信号来自故障点还是发射源, 形成盲区。盲区范围对应的距离 L_τ 为

$$L_\tau = \frac{V\tau}{2} \quad (2)$$

根据行波法检测原理, 利用 MATLAB 搭建通信电缆故障仿真模型来采集样本数据。在 Simulink 工具箱中调用传输线元件构建通信电缆电路, 在电路一端调用脉冲信号发生器模块产生脉冲信号, 调用 Signal Processing Toolbox 中的故障信号生成工具和处理器设置故障点。测量元件可以测量传输线参数和传输线上的行波信号, 并通过 Scope 工具观察脉冲信号在通信电缆中的传播与反射情况。行波信号包括电压、电流信号的幅度、频率、时域特性(峰值、峰峰值和波形)、频域特性(功率谱密度、频谱分布)、时频域特征(小波变换系数)等数据, 将以上数据按与仿真

时间点对应的格式绘制图表并保存。

3.2 PSO-SVM

SVM 是由 VAPNIK 提出的监督学习算法^[9], 其基本思想是基于最大间隔分类原理, 通过构建一个最优化的超平面来分割两个类别的样本, 如图 3 所示。

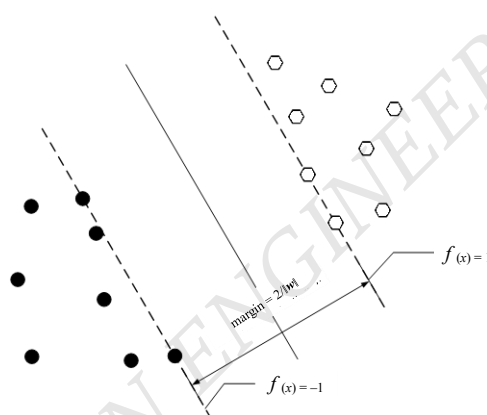


图3 SVM 最优分类超平面

在 SVM 模型中, 超平面可表示为 $(\mathbf{w} \cdot \mathbf{x}) + b = 0$,

其中 \mathbf{w} 为法向量, 决定了超平面的方向; b 为位移量, 决定了超平面与原点之间的距离。设训练集样本 $P = \{(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_N, y_N)\}$, 其中, $i = 1, 2, \dots, N$, $x_i \in R^n, y_i \in \{-1, 1\}$ 。找到一个超平面能够分开样本, 可以转化为凸二次规划问题, 求解能够满足约束条件的 \mathbf{w} 和 b , 使样本与超平面的间距最大化。约束条件为

$$\begin{cases} \min_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}\|^2 \\ \text{s. t. } y_i * (\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1 \end{cases} \quad (3)$$

为构造误差最小的超平面, 通过方程(4)可得到唯一的最优解。

$$\min J(\omega, \xi) = \frac{1}{2} \omega^T \omega + C \sum_{i=1}^n \xi_i \quad (4)$$

式中: C 为惩罚系数(或正则化参数), 表示对样本分类错误的容忍度。 C 值越高, 越不能容忍出现分类错误的情况, 模型容易过拟合; 相反 C 值越小, 模型容易欠拟合。

针对非线性分类问题, SVM 利用核函数将低维非线性数据映射到高维空间中, 对转换后的数据进行

线性分类。SVM 核函数种类包括径向基核函数(radial basis function, RBF)、 d 次多项式核和 Sigmoid 核^[10], 其中 RBF 可以直接反映数据之间的距离, 其分类效果优于其他核函数^[11]。本文选取 RBF 作为 SVM 核函数, 其计算公式为

$$k(X_i, X_j) = \exp\left(-\frac{\|X_i - X_j\|^2}{2\sigma^2}\right) \quad (5)$$

式中: X_i 为空间中的任意一点; X_j 为样本的中心点; $\|X_i - X_j\|$ 是从空间中任意一点到样本中心的距离; σ 为内核参数, 决定 RBF 的作用范围, 其值过大模型会更复杂, 过度拟合训练数据导致泛化性能较差。

正则化参数 C 和核函数内核参数 σ 的取值, 直接影响 SVM 的复杂度和泛化能力。在传统分类任务中, 设置超参数需要一定的经验和调优过程, 如交叉验证方法, 但该方法较为繁琐, 且存在模型过拟合的风险。

PSO 算法通过模拟鸟群觅食行为不断地迭代粒子速度和位置, 以寻找最优解^[12]。PSO 算法将一群粒子随机分布在搜索空间, 每个粒子代表一种 SVM 模型的超参数组合并作为一个候选解, 计算每个粒子位置对应的适应度值, 即训练好的模型在测试集的分类准确率。适应度值用于评价粒子解的好坏程度。PSO 算法中每个粒子都有自己的速度和位置。根据 PSO 原理, 粒子的速度和位置受个体经验和群体协作的影响, 粒子更新速度和位置的公式为

$$\begin{cases} V_{id}^{k+1} = \omega v_{id}^k + c_1 r_1 (pbest_{id}^k - x_{id}^k) + \\ \quad c_2 r_2 (gbest_{id}^k - x_{id}^k) \\ x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1} \end{cases} \quad (6)$$

式中: k 为当前迭代次数; ω 为惯性因子, 其值越小, 局部寻优能力越强, 全局寻优能力越弱; c_1 、 c_2 为学习因子; r_1 、 r_2 为随机数, 介于 (0,1) 之间; $i = 1, 2, \dots, n$ 为粒子数; $d = 1, 2, \dots, D$ 为当前空间维度; v 为迭代速度; x 为迭代位置; $pbest_{id}^k$ 和 $gbest_{id}^k$ 分别为第 i 个粒子的个体最优解和全局最优解。

重复更新粒子的速度和位置, 直到迭代次数达到

上限或者满足停止条件为止。在每次迭代过程中, 根据适应度值来适更新历史最优解和全局最优解。

基于 PSO-SVM 的通信电缆故障检测流程如图 4 所示。

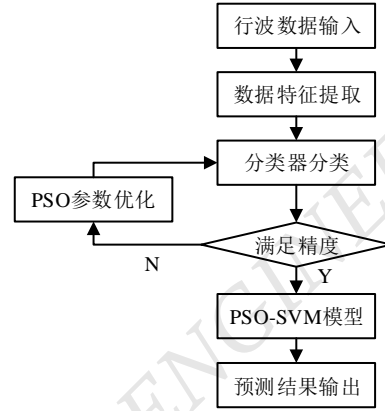


图 4 基于 PSO-SVM 的通信电缆故障检测流程

将通信电缆故障仿真模型采集的样本数据转化为特征向量并进行标记, 构成通信电缆行波数据集。将该数据集按 8 : 2 的比例划分为训练集和验证集。SVM 模型超参数配置时, 正则化参数 C 设置为 0.01, 核函数选择 RBF, 同时将核函数 σ 设置为 0.1。将训练集数据输入到 SVM 模型进行训练, 迭代过程中根据 SVM 模型在训练集上的性能评估其适应度值, 即精度是否满足要求。当精度不满足要求时, 通过 PSO 算法优化 SVM 模型超参数, 重复此过程直至达到最大迭代次数或者精度达到设定的阈值。最后通过验证集验证 SVM 模型性能是否满足要求。

4 基于 CNN-LSTM 的通信电缆故障定位模型

4.1 CNN

CNN 是一种拥有深度结构的神经网络模型, 主要包括输入层、卷积层、池化层、全连接层和输出层等, 基本结构如图 5 所示。

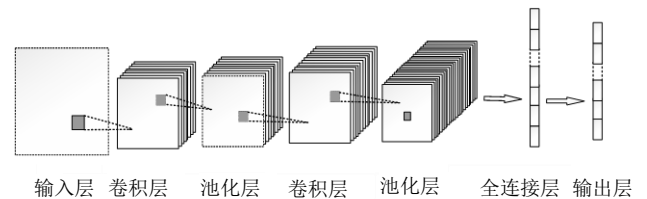


图 5 CNN 基本结构

输入层、卷积层和池化层组成特征提取器，用于遍历输入数据，提取高维特征；全连接层和输出层构成分类识别器，用于判断目标类别并将其映射为预测结果^[13]。输入数据经预处理后前向传播，依次进行卷积操作和池化操作，提取样本数据的特征信息。其中，卷积操作利用卷积核对输入数据进行扫描并提取特征；池化操作对卷积后的特征图进行降采样，减少特征图的维度并保留重要信息。全连接层融合提取的特征并降低维度，最终由输出层将全连接层的输出映射到类别标签上，完成分类任务。

4.2 LSTM

LSTM 网络是循环神经网络的一种特殊变体，它通过记忆单元和门控机制，提升了从长输入序列中学习特征的能力，不仅克服了训练循环神经网络的诸多困难，还增强了网络的稳定性。当多个 LSTM 网络单元相互连接，形成更复杂的网络结构时，排除了梯度消失问题，进一步提升了模型的性能和准确性，使 LSTM 网络在处理复杂时间序列特征的波形数据时，可以有效地捕捉到信号中的时序特征和周期性变化。LSTM 网络单元结构如图 6 所示。

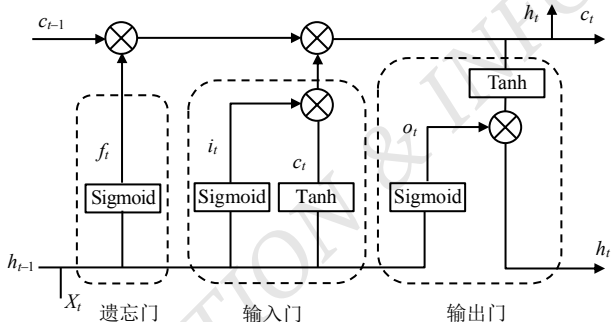


图 6 LSTM 单元结构图

图 6 中： f_t 、 i_t 、 o_t 分别为遗忘门、输入门和输出门的计算结果，Sigmoid 激活函数的输出值为 0 或 1，Tanh 激活函数输出值的范围为 $[-1, 1]$ 。 c_{t-1} 和 h_{t-1} 为上一层的输出， c_t 和 h_t 为本层的输出，以 c_{t-1} 和 h_{t-1} 为例， h_{t-1} 为上一层输出到这一层的全部信息， c_{t-1} 表示上一层与本层的交互信息量，其决定本层将会保留哪些信息。

4.3 CNN-LSTM

基于行波法得到的通信电缆行波数据不属于传

统意义上的结构数据，而是一维或多维的时间序列数据^[14]。通过融合 CNN 与 LSTM 模型能有效提取通信电缆行波数据的空间特征和时序特性，实现故障的精确定位。基于 CNN-LSTM 的通信电缆故障定位模型结构如图 7 所示。

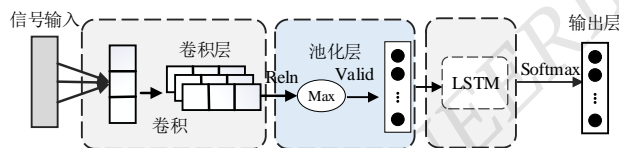


图 7 基于 CNN-LSTM 的通信电缆故障定位模型结构

基于 CNN-LSTM 的通信电缆故障定位模型包含信号输入层、卷积层、池化层、LSTM 层和输出层。为了防止过拟合，在每个卷积层后加入批量归一化 (BN)，并应用 Dropout 操作。Dropout 随机将神经元输出置零，减少神经元间的依赖，降低过拟合风险。Dropout 删除比例设为 0.1，初值设为 0.001^[15]。使用自适应矩估计算法 (adaptive moment estimation, Adam) 将训练误差进行反向传播，逐层逐步地更新模型参数，进行学习率自适应。

基于 CNN-LSTM 的通信电缆故障定位模型训练流程如图 8 所示。

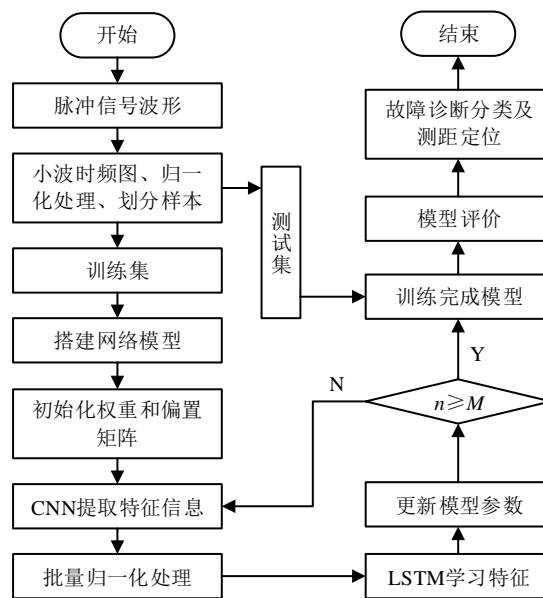


图 8 基于 CNN-LSTM 的通信电缆故障定位模型训练流程

配置 CNN-LSTM 网络初始结构参数、权重值和学习率后, 将输入脉冲反射信号样本进行连续小波变换^[16], 生成每段小波时频数据。当故障数据和正常数据相差过大时, 进行归一化处理:

$$y' = \frac{y - y_{\min}}{y_{\max} - y_{\min}} \quad (7)$$

式中: y 为原始数据; y' 为归一化结果, 值域在 $[0,1]$ 之间。

根据网络输入尺寸对输入样本进行调整后, 按 8 : 2 的比例将其划分为训练集和测试集。首先, 将训练集数据输入 CNN-LSTM 网络进行训练, CNN 通过卷积核滑动提取训练样本的局部空间特征, 生成不同尺度和抽象的空间特征映射; 然后, 经过池化层减少特征映射的尺寸, 并保留重要的特征信息做归一化处理; 接着, 将 CNN 输出的特征序列作为 LSTM 网络的输入序列, LSTM 网络通过记忆单元和门控机制学习数据的时间序列信息, 并捕获数据的长期依赖关系^[17]; 最后, 判断当前网络的训练次数 n 是否达到初始设定值 M , 训练完成后通过在测试集上的表现评价模型。

5 实验验证

为验证本文方法在通信电缆故障检测与定位任务中的有效性, 对基于 PSO-SVM 的通信电缆故障检测模型、基于 CNN-LSTM 的通信电缆故障定位模型进行实验验证。实验软硬件条件如表 2 所示。

表 2 实验软硬件条件

操作系统	Windows 10 Pro (64 位)
CPU	Intel Core i5-12490F @4.6 GHz
GPU	NVIDIA GeForce RTX 3070 8 GB
内存	16 GB
仿真软件	MATLAB R2022b

通信电缆故障仿真模型采集了 10~200 m 范围内间距为 10m 的正常、短路、断路和高阻故障, 共 400 组通信电缆行波数据样本。以可视化短路故障行波数据图 (图 9) 和 200 m 高阻故障行波时间序列图 (图

10) 为例, 反映通信电缆行波信号 (电压信号) 随时间的变化趋势及反射情况。

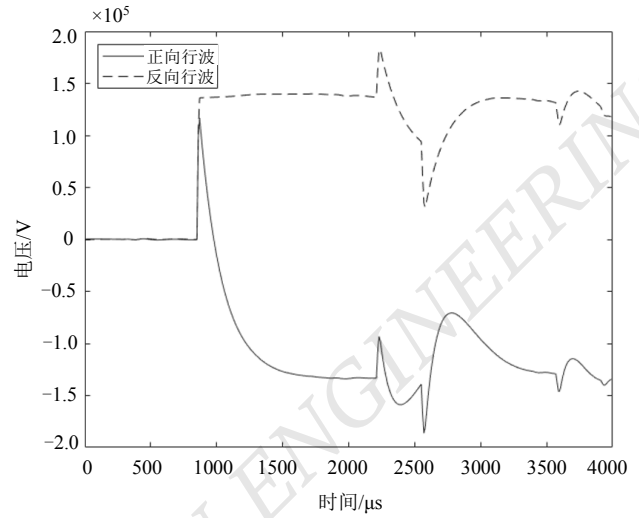


图 9 可视化短路故障行波数据图

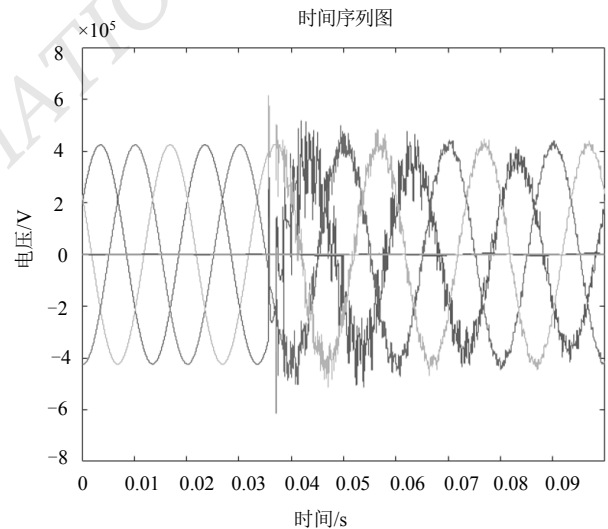


图 10 可视化 200 m 高阻故障行波时间序列图

将通信电缆故障仿真模型采集的通信电缆行波数据转化为特征向量并进行标注, 构建通信电缆行波数据集。对输入 CNN-LSTM 通信电缆故障定位模型的数据进行小波变化及归一化处理。经 1 200 轮训练迭代后, 基于 PSO-SVM 的通信电缆故障检测模型在测试集上的识别准确率达 99.4%。

平均绝对误差 (mean absolute error, MAE) 用于评估模型预测值与实际观测值之间的平均绝对差异。均方根误差 (root mean squared error, RMSE) 用于评

估预测结果与真实值之间的差异程度。MAE、RMSE 值越小,模型的预测精度和准确性越好,经过 1 200 轮训练迭代后,基于 CNN-LSTM 的通信电缆故障定位模型在测试集上的 MAE 为 0.334 9, RMSE 为 0.320 8。

采用单独的 CNN 模型和 LSTM 模型对相同数据集进行对比实验,实验结果如表 3 所示。

表 3 模型实验结果对比

网络模型	Accuracy	RMSE	MAE
CNN	89.76%	0.907 5	0.698 1
LSTM	93.03%	0.702 5	0.541 9
CNN-LSTM	99.23%	0.320 8	0.334 9

由表 3 可知,基于 CNN-LSTM 的通信电缆故障定位模型结合了 CNN、LSTM 两种模型的特点,能够更全面地利用通信电缆行波数据中的空间和时序特征,在故障点位置预测任务中,其 MAE 对比单独使用 CNN 和 LSTM 模型分别降低了 0.363 2 和 0.207, RMSE 分别降低了 0.586 7 和 0.381 7,准确率分别提高了 9.47% 和 6.2%。

6 结论

本文为解决通信电缆故障检测与定位的问题,首先,搭建通信电缆仿真模型来采集数据样本;然后,在基于 PSO-SVM 的通信电缆故障检测模型中,利用 PSO 算法对 SVM 模型的超参数进行优化,实现不同故障的有效分类;接着,在基于 CNN-LSTM 的通信电缆故障定位模型中,先通过 CNN 提取行波数据的空间特征,再通过 LSTM 学习数据中的时间序列信息,实现对通信电缆故障点的精确定位;最后,通过对比实验证明,基于 CNN-LSTM 的通信电缆故障定位模型的网络精度较单独使用 CNN 和 LSTM 模型均有显著提升。未来在进一步优化模型性能的基础上,结合实际场景数据进行更全面的验证和应用。

参考文献

[1] GILANY M, KHALIL IBRAHIM D. Traveling-wave-based fault-location scheme for multi-aged underground cable

system[J]. IEEE Transactions on power delivery, 2006,22(1): 82-89.

[2] 鹿洪刚,覃剑,陈祥训,等.电力电缆故障测距综述[J].电网技术,2004,28(20):58-63.

[3] 袁燕岭,李世松,董杰,等.电力电缆诊断检测技术综述[J].电测与仪表,2016(11):1-7.

[4] JENA S, MISHRA D P, SALKUTI S R. Fault detection, classification, and location in underground cables[M]// Power Quality in Microgrids: Issues, Challenges and Mitigation Techniques. Singapore: Springer Nature Singapore, 2023:195-215.

[5] 高恩嵩.电桥法在电缆低阻故障点探测中的应用[J].电工技术,2017(8):89-90.

[6] 张金亮.基于人工智能的电缆故障精确定位与实现技术[D].山东大学,2023.

[7] YANG X, CHOI M S, LEE S J, et al. Fault location for underground power cable using distributed parameter approach[J]. IEEE Transactions on Power Systems, 2008,23(4): 1809-1816.

[8] JANANI H, KORDI B. Towards automated statistical partial discharge source classification using pattern recognition techniques[J]. High Voltage, 2018,3(3):162-169.

[9] VAPNIK V. The nature of statistical learning theory[M]. Springer science & business media, 1999.

[10] 曹嘉嘉,严圆,陈益等.基于 PSO 优化的 SVM 在心脏病分类上的应用[J].东莞理工学院学报,2022,29(3):50-56.

[11] FENG G H. SVM classification kernel function and parameter selection comparison[J]. Computer Engineering and Applications, 2011,47:123-124.

[12] JIAO L, LI Y, GONG M, et al. Quantum-inspired immune clonal algorithm for global optimization[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 2008, 38(5):1234-1253.

[13] WANG Y, HOU L, PAUL K C, et al. ArcNet: Series AC arc fault detection based on raw current and convolutional neural network[J]. IEEE Transactions on Industrial Informatics, 2021, 18(1):77-86.

[14] WU Y, ZHANG P, LU G. Detection and location of aged cable segment in underground power distribution system using deep learning approach[J]. IEEE Transactions on Industrial Informatics, 2021,17(11):7379-7389.

[15] 冯寅,贾宏涛,杨震强,等.基于 CNN-LSTM 的输电线路故障诊断方法研究[J].电网与清洁能源, 2023,39(11):59-65.

(下转第 60 页)