

基于 S 变换和深度学习的多特征融合的电电压暂降源识别方法

张峰 陈雷

(东北石油大学电气信息工程学院 大庆 163000)

摘要:随着工业和科技的发展,用户对电压暂降的关注度不断提高,识别电压暂降产生的原因愈显得越来越重要。针对引起电压暂降的单一暂降源和复合暂降源,提出了将 S 变换提取特征和深度学习自动提取特征相结合的识别方法。首先利用数值模型框架产生单一暂降源和复合暂降源数学模型,进而得到 9 种故障类别的暂降数据集并作为原始数据,其次对原始数据进行处理,即在标准的 S 变换基础上引入两个调节因子得到改进的 S 变换,得到 S 变换数据,引入 16 个指标对 S 变换数据进行特征提取并作为指标特征,将上述原始数据和 S 变换数据作为模型输入,利用卷积神经网络(convolutional neural network, CNN)对暂降数据进行空间特征提取,同时将数据分为多个一维向量输入到双向长短期记忆网络(bi-directional long-short-term memory networks, Bi-LSTM)提取时序特征,最后建立指标特征、空间特征以及时序特征的多特征融合的 S-CNN-LSTM 识别模型。仿真结果表明,未经过特征融合与经过多特征融合的认可准确率分别为 98.36%、99.08%,说明多特征融合能够提高识别准确率。

关键词:电压暂降源;改进 S 变换;CNN-LSTM 模型;特征融合

中图分类号: TP181; TM743 **文献标识码:** A **国家标准学科分类代码:** 470.4054

Multi-feature fusion based on S-transform and deep learning for voltage transient source identification

Zhang Feng Chen Lei

(College of Electrical Information Engineering, Northeast Petroleum University, Daqing 163000, China)

Abstract: With the development of industry and science and technology, the user's concern about voltage dips is increasing, and it is more and more important to identify the causes of voltage dips. Aiming at the single and composite sources of voltage dips, this paper proposes a recognition method that combines the S-transform feature extraction and deep learning automatic feature extraction. Firstly, the numerical modeling framework is used to generate mathematical models of single and composite transient sources, and then the transient datasets of nine fault categories are obtained and used as the original data, and secondly, the original data are processed, i. e., two adjustment factors are introduced to obtain the improved S-transform based on the standard S-transform to obtain the S-transform data, and 16 metrics are introduced to extract features from the S-transform data and are used as the metrics' features. The original data and the S-transformed data are used as model inputs, and the spatial features are extracted from the transient data by using convolutional neural network (CNN), while the data are divided into several one-dimensional vectors and input into bi-directional long short-term memory networks (Bi-LSTM), which is the most suitable method to extract the spatial features of the S-transformed data. Memory networks (Bi-LSTM) to extract temporal features, and finally establish a multi-feature fusion S-CNN-LSTM recognition model with indicator features, spatial features and temporal features. The simulation results show that the recognition accuracies without feature fusion and after multi-feature fusion are 98.36% and 99.08%, respectively, indicating that multi-feature fusion can improve the recognition accuracy.

Keywords: voltage transient source; improved S-transform; CNN-LSTM model; feature fusion

0 引言

电压暂降是指供电电压有效值快速下降至额定值的10%~90%的事件,其主要原因包括短路故障、变压器投切、电动机启动,典型持续时间为0.5~30个周波,尽管持续时间很短,但对用户造成的损失却极为严重,据统计80%服务器出现瘫痪以及用户端45%左右数据丢失均与电压暂降有关。因此准确识别电压暂降源对解决供电部门和用户之间的纠纷具有重要意义^[1]。

目前,电压暂降数据集获取主要是利用 Simulink/PSCAD 仿真得到^[2-3],但是此方法需要不断调整参数,并且得到的数据量较少,不足以进行深度学习。特征提取主要是通过人工特征提取和深度学习自动提取两种提取方式,经验模态分解(empirical mode decomposition, EMD)、S变换是人工特征提取的典型方式,但是,EMD^[4]会因信号存在噪声干扰出现模态混叠,导致错假的时频分布,通过小波阈值^[5]可以去噪,但是会导致信号失真。S变换^[6]虽然可以检测突变点,但是其高斯窗的大小不能根据波形的变化而自动调节,突变点时刻存在误差。S变换模矩阵方法^[7]绘制基频幅值以及频率幅值包络线虽然能直观显示不同暂降源的差异,但是不能量化这些差异。深度学习自动提取特征主要是利用卷积神经网络(convolutional neural network, CNN)、深度置信网络(deep belief network, DBN)以及多模型集成学习网络, CNN 中引入多个卷积核^[8-13]对数据信息进行处理,虽然在一定程度上能够提高准确率,但是增加了模型的复杂性,延长了训练时间,并且池化层的存在会导致信息丢失。深度置信网络模型^[14-15]可以将人工提取转变为自动提取,解决了过度依赖专家经验问题,但是该模型容易过拟合,并有因含有多层RBM,计算量较大,收敛速度较慢。多模型集成学习^[16-17]虽然可以提高识别率,但是数据集很大需要更多的计算资源和时间,数据集过小会过度依赖训练数据,导致在测试

数据上性能下降而且超参数选择时需要大量实验和调参,增加模型的复杂性。

针对以上数据量不足,本文通过数值模型框架构建暂降源数据集^[18],获取多种暂降源时域数据并作为原始数据,对于模型复杂以及信息丢失,本文将人工经验及深度学习相结合即基于改进的S变换得到时频域数据^[19]并作为S变换数据,引入16个暂降指标^[20]提取S变换数据特征获得指标数据作为深度学习补充信息,将卷积神经网络和长短期记忆(long-short term memory, LSTM)网络相结合并进行优化作为学习模型得到简化模型,最后,原始数据、S变换数据以及指标数据作为模型输入,对数据进行特征提取,然后经过训练验证测试,获得暂降源类别。通过人工提取和深度学习自动提取相结合方法,在一定程度上减少了信息丢失,拓展了特征的多样性,提升了模型的识别准确度,实现多特征融合的暂降源识别,完成不同暂降源的分类。

1 电压暂降数值模型

线路短路故障、感应电动机启动及变压器投运是引起电压暂降的3种典型因素。通过搭建数值模型方法来模拟电压暂降成因,并在此基础上将各种故障模型组合形成复合模型,由此形成一种由单一暂降源和复合暂降源共同构成符合实际暂降事件的数据集。由于单相短路接地故障在电力系统中相对于其他两种短路故障发生频次最高,对于复合暂降本文仅考虑单相短路故障与其他故障的复合,包括单相短路接地与变压器投切的复合,单相短路接地与感应电动机启动的复合以及单相短路接地与变压器投切、感应电动机启动复合。设置每种信号的总时间为0.4 s,生成1 000条波形,每条波形样本数量为5 000,基准频率为50 Hz,暂降持续时间随机生成。具体的模型如表1所示。

表1 电压暂降源模型

Table 1 Voltage temporary source model

故障名称	数学模型	参数设置
单相短路接地		$t=0.4\text{ s}, f_1=50\text{ Hz}, s=0.01$ $d\text{iptime}=5\text{ ms}\sim 0.1\text{ s}, t_1=0.08\sim 0.18\text{ s}$ $\theta_A=\theta, \theta_B=\theta-120, \theta_C=\theta+120$ $u_n=1: \forall t \text{ if } t-t_n>0 \quad u_n=0: \forall t \text{ if } t-t_n<0$
两相短路	$V=[1-ms(u_1-u_2)]\sin(2pf_1t+q)$	$m_{A1}=0.5\sim 0.6, m_{B1}=-0.1\sim 0, m_{C1}=-0.2\sim -0.1$ $m_{A2}=0.5\sim 0.6, m_{B2}=0.45\sim 0.55, m_{C2}=-0.2\sim -0.1$
三相短路		$m_{A3}=0.5\sim 0.6, m_{B3}=0.45\sim 0.55, m_{C3}=0.4\sim 0.5$

续表

故障名称	数学模型	参数设置
变压器投切	$V = [1 - m_s(u_1 \cdot e^{-\alpha(t-t_1)})] \sin(2\pi f_1 t + \theta)$	$\alpha = 4 \sim 5, m_A = 0.09 \sim 0.14, m_B = 0.06 \sim 0.11, m_C = 0.01 \sim 0.06$
感应电动机启动	$V = [1 - m_s(M_s - M_r)] \sin(2\pi f_1 t + \theta)$ $M_s = u_1(1 - e^{-\alpha t}) + u_1(r \sin(2\pi f_r t_1) e^{-\rho t})$ $M_r = u_2(1 - e^{-\beta t})$	$m_A = 0.11 \sim 0.16, m_B = 0.1 \sim 0.15, m_C = 0.09 \sim 0.14$ $r = 0.1 \sim 0.2, \alpha = 100 \sim 106, \beta = 10 \sim 11$ $f_r = 10 \sim 11, \rho = 100 \sim 101$
单相短路+变压器投切	$V = [1 - V_2 - V_3] \sin(2\pi f_1 t + \theta)$	$V_1 = m_1 s_1(M_s - M_r) V_2 = m_2 s_2(u_3 - u_4)$
单相短路+感应电动机启动	$V = [1 - V_1 - V_2] \sin(2\pi f_1 t + \theta)$	$V_3 = m_3 s_3(u_1 \cdot e^{-\alpha(t-t_1)}) s_1 = s_2 = s_3 = 0.01,$ $\alpha = 4 \sim 5$
变压器投切+感应电动机启动	$V = [1 - V_1 - V_3] \sin(2\pi f_1 t + \theta)$	$m_{1A} = 0.11 \sim 0.16, m_{1B} = 0.1 \sim 0.15,$ $m_{1C} = 0.09 \sim 0.14$ $m_{2A} = 0.5 \sim 0.6, m_{2B} = -0.1 \sim 0, m_{2C} = -0.2 \sim -0.1$
所有故障	$V = [1 - V_1 - V_2 - V_3] \sin(2\pi f_1 t + \theta)$	$m_{3A} = 0.09 \sim 0.14, m_{3B} = 0.06 \sim 0.11,$ $m_{3C} = 0.01 \sim 0.06$

表1中, t 为总时间, t_1 和 t_2 分别为电压暂降开始和结束时间, u 为阶跃函数用来模拟电压暂降, m 为暂降幅度, s 为开关时间, f_1 为基准频率, θ 为初相角, α 为电压正常水平下降为骤降水平的凹陷衰减率, β 为电压骤降水平回升到标准水平的恢复上升率, M_s 为电压跌落电动机启动函数, M_r 为电压跌落恢复函数, r 为电动机启动扰动系数, f_r 电动机启动扰动频率, ρ 三相正弦电压的衰减系数。

2 基于改进 S 变换获取暂降特征

2.1 引入调节因子改进 S 变换

为了更好地控制窗口的形状, 以及达到更好的时-频分辨率, 因此, 在 S 变换高斯窗函数的基础上引入调节因子 a 和 b , 这使得时窗宽度随信号频率变化的速度更加灵活。该改进的 S 变换表达式如下:

$$S(t, f) = \int_{-\infty}^{+\infty} h(\tau) \frac{a |f|^b}{\sqrt{2\pi}} e^{-\frac{a^2(t-\tau)^2 |f|^{2b}}{2}} e^{-j2\pi f\tau} d\tau \quad (1)$$

式中: $h(\tau)$ 为信号函数; t 为时间; f 为频率; τ 为高斯窗函数的中心位置; a 为高斯窗幅度调节系数; b 为高斯窗指数调节系数。不同的 a 和 b 对应不同的时频分辨率, 通过调整 a 和 b 可以使电压暂降特征提取效果达到最佳。

设置 T 为采样间隔, N 为总采样点数, 对式(1)进行离散化处理, $f = n/(NT)$, $\tau = jT$ 。信号 $x(t)$ 的改进 S 变换的离散形式表示如下:

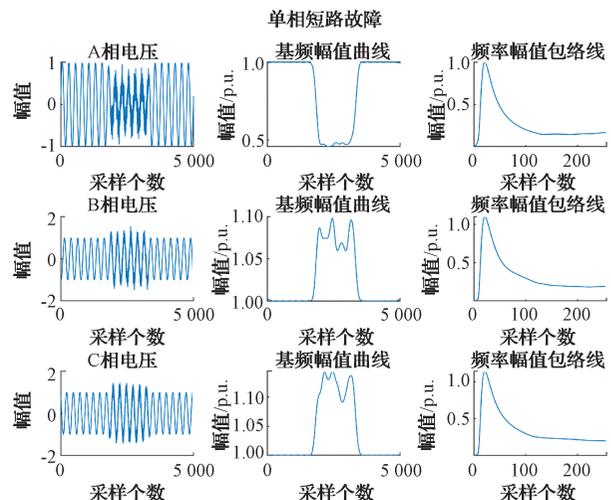
$$\begin{cases} S^{a,b}[iT, \frac{n}{NT}] = \sum_{m=0}^{N-1} X[\frac{n+m}{NT}] \cdot \exp\left(-\frac{2\pi^2 m^2}{an^{2b}} + j\frac{2\pi mi}{N}\right), & n \neq 0 \\ S^{a,b}[iT, 0] = \frac{1}{N} \int \sum_{m=0}^{N-1} X[\frac{m}{NT}], & n = 0 \end{cases} \quad (2)$$

式中: $i = 0, 1, 2, \dots, N-1, m, n$ 同。

2.2 故障源数据进行改进 S 变换

对上述单一故障和复合故障模型进行改进 S 变换, 得到各个故障三相复数样本并且拟合出各个故障基准频率下的样本曲线, 将变换后该频率下的数据可视化, 并且在大量的实验基础上可以得出不同故障的调整因子的大小^[21]。短路故障和感应电动机启动对应的 $a = 5, b = 0.45$; 变压器投切对应的 $a = 1, b = 0.63$ 。各个故障的时域波形曲线、基频幅值以及频率幅值曲线如图 1~3 所示。

从图 1 可以看出, 三相短路故障引发电压暂降电压幅值相等, 某一相发生暂降的同时另两相发生电压暂升, 电压暂降发生和恢复的波形陡, 基波电压幅值变化过程呈矩形, 暂降开始和结束瞬间, 幅值均发生突变, 暂降过程中电压幅值变化不大, 整个过程并没有谐波产生。



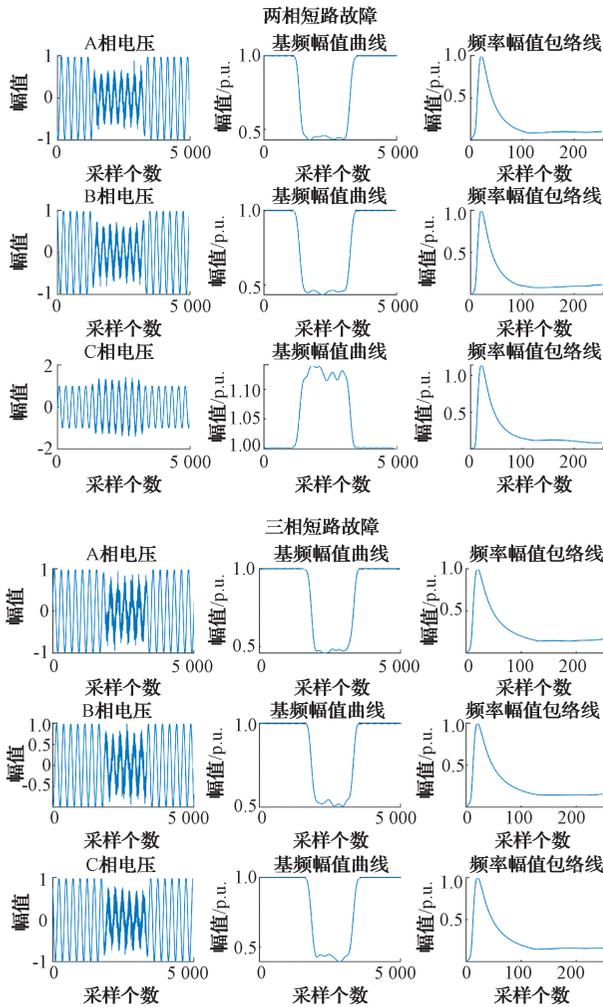


图1 短路故障的三相电压、基频幅值以及频率幅值曲线
Fig.1 The three-phase voltage, base frequency value, and frequency amplitude curve of short-circuit failure

从图2可以看出,三相电压暂降不平衡且幅值不相等,电压暂降幅值不会低于0.9 p.u.;恢复过程相对于暂降过程会有突变;电压信号中有高次谐波。

从图3可以看出,三相电压暂降不平衡且暂降坡度较陡,三相暂降恢复过程是不相同的,A相恢复之后的电压幅值约为1 p.u.,B、C两相恢复后的电压幅值约为0.98 p.u.和0.99 p.u.。

考虑到复合暂降模型是上述单一暂降模型互相叠加,因此不同单一暂降复合后的变化可以视为上述变化的叠加,复合暂降包含不同单一暂降的所有特征。

2.3 基于改进S变换特征提取

为了能够量化表示出上述三相短路,两线短路、单相短路、变压器投切以及电动机启动的时域波形、基频幅值曲线以及频率幅值包络线的信息,引入均值、标准差、均方根值、暂降深度、能量、峰度、突变点个数、香农熵、对数能量熵、基频幅值上升和下降斜率、波峰系数和暂降时间

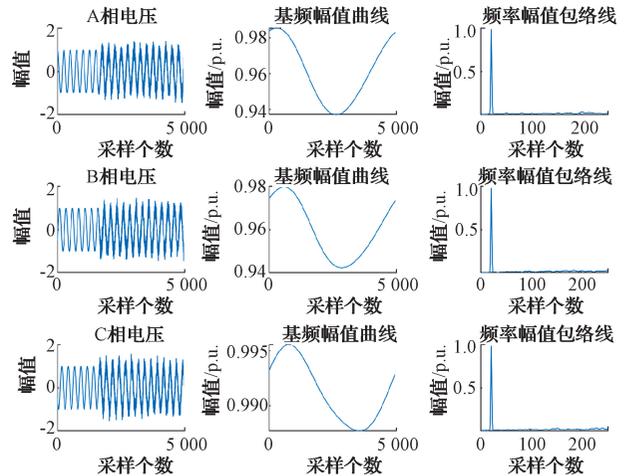


图2 变压器投切的三相电压、基频幅值以及频率幅值曲线
Fig.2 The three-phase voltage, base frequency value, and frequency amplitude curve of the transformer are cut

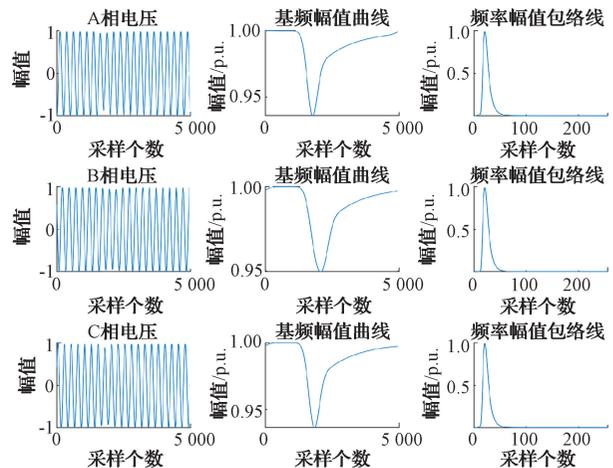


图3 电动机启动的三相电压、基频幅值以及频率幅值曲线
Fig.3 The three-phase voltage, base frequency value, and frequency amplitude curve of the motor start

比共16个指标来构建电压暂降特征体系,流程如图4所示。

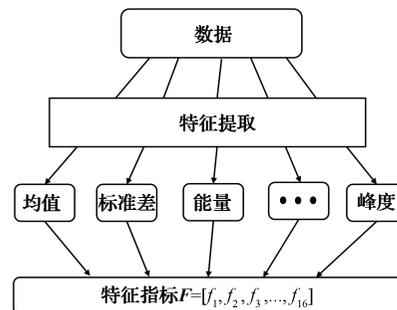


图4 特征指标体系

Fig.4 Feature indicator system

3 多特征融合模型

3.1 模型框架

为了实现电压暂降多故障分类,本文设计基于深度学习和改进 S 变换的特征融合并行优化模型—S-CNN-LSTM。对三相原始数据,S 变换数据以及 16 个特征指标

数据作为模型的输入,通过 CNN 和 LSTM 分别进行特征提取,然后经融合后输入到全连接层,最终实现对故障的分类。该模式下实现了对多维度空间、时序、S 变换指标的特征融合,提升模型的泛化能力、分类效率以及准确率。框架如图 5 所示。

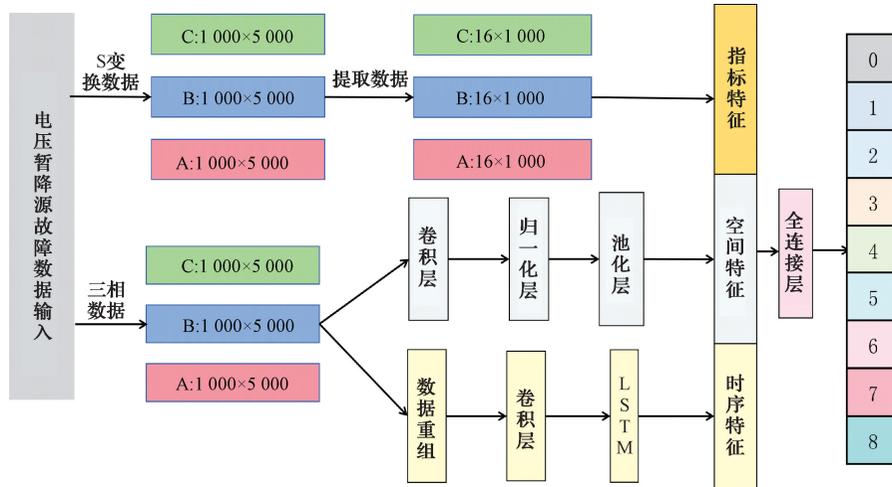


图 5 特征融合模型框架

Fig. 5 Feature fusion model framework

3.2 CNN 模型

CNN 模型如图 6 所示,本文是以一维 CNN 为基础,并结合故障数据的输入长度和数据形式进行 5 次卷积,批量归一化,最大池化,提高了模型收敛速度和分类效果。其中添加批量归一化可以简化调参过程,缓解梯度消失和梯度爆炸,具有一定的正则化效果,提升模型的学习速度;最后的池化层输出结果通过 Flatten 层可以压缩多维的空间数据,形成一维张量,有利于后续的特征融合。

本文对电压暂降故障信号进行特征提取,其本质上是对一维时间序列进行处理,因此采用一维卷积神经网络。

由于信号的输入维度为 n 的离散型时间序列输入层可以表示为:

$$\mathbf{X}_n = [x_1, x_2, \dots, x_n] \quad (3)$$

其次,经过卷积层时,移动的卷积核窗口通过非线性激活函数对数据中的元素进行卷积计算,本文采取 ReLU 函数,相比 Sigmoid 函数可以防止梯度消失和梯度爆炸,并且加快计算速度,长度为 k 的卷积核移动时,每次输出的 x_{conv} 计算公式为:

$$x_{conv} = f\left(\sum_{i \in n-k+1} W \cdot x_i + B\right) \quad (4)$$

式中: W 和 B 分别表示该层的卷积核权重和偏置,每一层

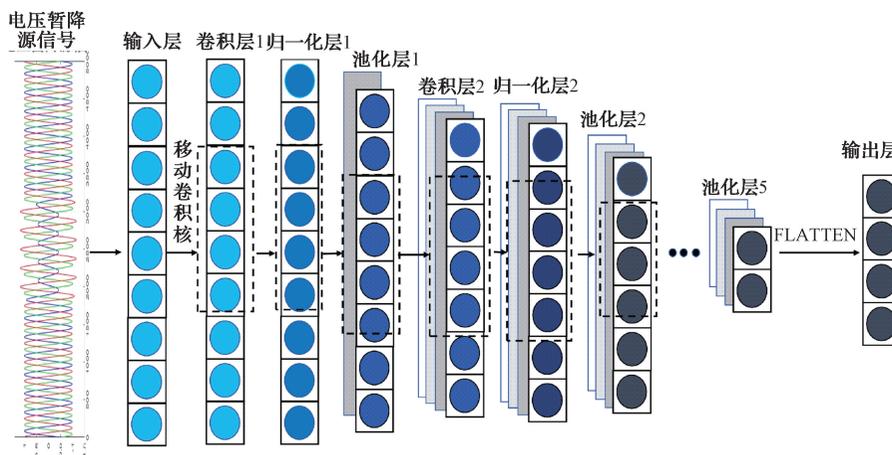


图 6 CNN 模型

Fig. 6 CNN model

的参数会随着卷积核的大小和步长而发生改变; n 为输入序列的长度; $f(\ast)$ 代表激活函数, 该函数的计算公式如下:

$$ReLU(x) = \max(0, x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (5)$$

数据经过卷积层之后要经过批量归一化层(batch normalization, BN), BN层使每一层拥有服从相同分布的输入数据, 从而可以使用大的学习率进行训练, 加快网络的训练速度和收敛速度, 控制梯度爆炸, 防止梯度消失和过拟合, 增强网络的泛化能力。其公式为:

$$\begin{cases} \hat{y}_k^l = \frac{y_k^l - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \\ z_k^l = \gamma_k^l \hat{y}_k^l + \beta_k^l \end{cases} \quad (6)$$

式中: y_k^l, z_k^l 表示 BN 层的输入和输出; $\mu_B = E[y_k^l]$, 其中 E 表示期望; γ_k^l 表示第 l 层第 k 个神经元的缩放系数; β_k^l 表示 BN 层的偏置; σ_B^2 表示方差; ϵ 为常数项。

批量归一化层将数据控制在特定范围内之后进入到池化层, 在该区域内数据可以进行空间合并, 有效减少模型参数, 提升模型运算速度。本文采用最大池化 h_{\max} 。

$$h_{\max} = \max \sum \{x_i | x_i \in M\} \quad (7)$$

式中: M 为池化域内的所有序列数据组成的集合; $\max(\ast)$ 表达式为输出集合的最大值。

输入 X_n 经过若干次的卷积层、批量归一化层, 最大池化层后, 在最后一次池化之后加入 FLATTEN 层, 可以使前面的特征压缩为一维张量, 从而得到一维特征向量。

3.3 基于数据分割的双向 LSTM

基于数据分割的双向 LSTM 模型如图 7 所示, 由于时间序列长度过大, 不宜直接输入网络进行训练。首先引入折叠率, 将一维时间序列数据进行等长分割, 然后对采样点重组成 X_1, X_2, \dots, X_n , 分批量输入 LSTM 中可以有

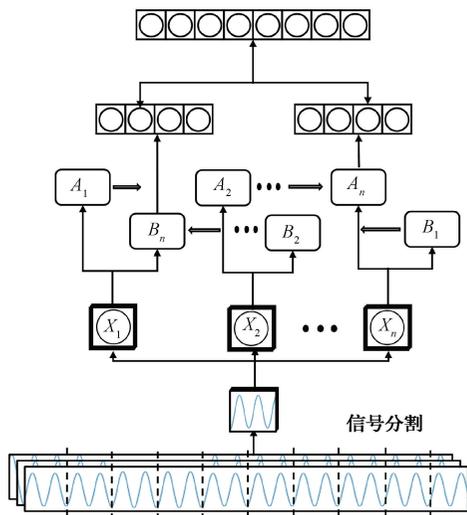


图 7 基于数据分割的双向 LSTM 模型
Fig. 7 Data-based two-way LSTM model

效缓解 LSTM 梯度消失也可以提升模型的训练速度, 本文采用双向 LSTM 网络划分为前馈 LSTM($A_1 \sim A_n$) 和反馈 LSTM($B_1 \sim B_n$)。把上述 X_n 维向量的每个向量按照正向顺序输入到前馈 LSTM($A_1 \sim A_n$), 再按照反向顺序输入到反馈 LSTM($B_1 \sim B_n$) 同时提取特征。最后将双向 LSTM 输出等长的特征序列进行拼接, 整个模型训练结束。

LSTM 网络的内部结构如图 8 所示, 是在 RNN 基础上为了避免梯度消失和梯度爆炸引入了门控装置, 可以对每一个时刻的信息进行选择保留, 保留重要信息, 丢弃不重要信息, 从而为序列数据赋予依赖性^[19]。

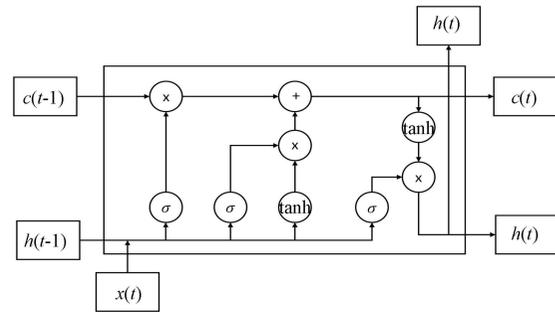


图 8 LSTM 内部网络结构

Fig. 8 LSTM internal network structure diagram

3.4 特征融合

三相初始数据一方面经过卷积神经网络提取空间特征, 经过一系列卷积池化, 最后得到 64 个特征图, 然后将它展开为一维张量作为全连接层的输入, 全连接层输入特征数量为 4 096(神经元个数), 其中有 $4\ 096 \times 512$ 权重矩阵, 输出向量长度为 512。另一方面由于输入信号的长度很长, 将数据进行切分处理, 即引入折叠率将 5 000 数据分成 10 份, 每一份是单独的一维向量, 然后输入到 LSTM 网络中提取时间序列特征经过训练输出向量长度为 512, 最后将上述提取的指标特征通过拼接的方式与上述两种特征进行融合, 共同输入到权重矩阵 $512 \times 1\ 024$ 的全连接层, 输出向量长度为 512, 图 9 所示为 S-CNN-LSTM 模型流程。

考虑到神经网络层数很多, 为了简化计算, 防止梯度弥散和加快收敛速度, 该层使用 ReLU 作为激活函数, 主要作用是将输入值限制在一个非负的范围, 当输入的值 < 0 时, 输出值为 0; 当输入值 ≥ 0 时, 输出值等于输入值本身。其计算公式如下:

$$f(x) = \max(0, x) \quad (8)$$

4 仿真实验

4.1 实验环境及模型参数设置

实验采用基于 Pytorch 的深度学习框架以及 Python 编程语言, 构建电压暂降识别模型。学习率为 0.000 1, 信号长度 5 000, 信号通道 3, 输入维度为 30, 隐藏神经元为

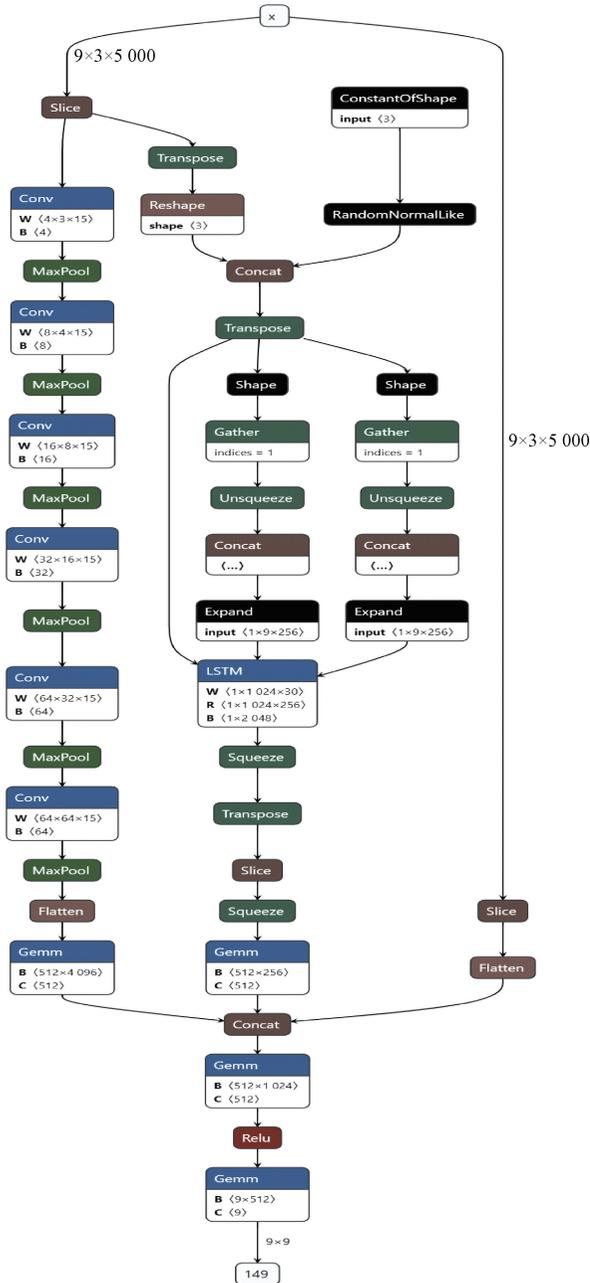


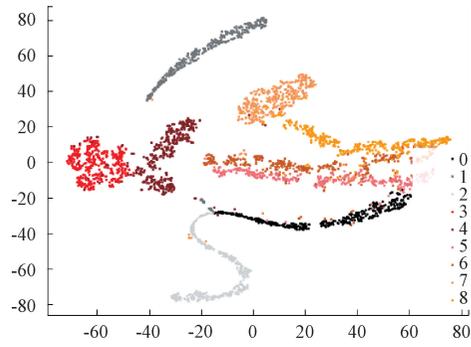
图9 S-CNN-LSTM 模型流程
Fig.9 S-CNN-LSTM model flow chart

256, 折叠率为 10。选取 ReLU 作为激活函数, 采用优化器为 Adam, 迭代次数为 100, 训练集、验证集及测试集的比例为 5 : 1 : 4。

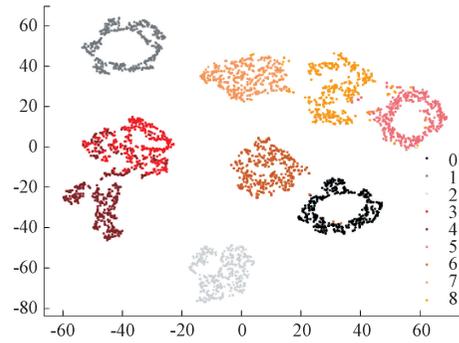
4.2 网络训练及过程分析

为了更加直观体现不同样本数据提取特征, 本文通过 t 分布随机领域嵌入 (t-distributed stochastic neighbor embedding, t-SNE) 降维可视化方法, 将 9 种暂降源的原始数据、S 变换数据以及 S 变换 + 指标数据提取的特征降维到二维空间并可视化, 分别如图 10(a)~(c) 所示。从图 10(a) 可以看出, 9 种故障源区分性比较差, 各种故障类比较

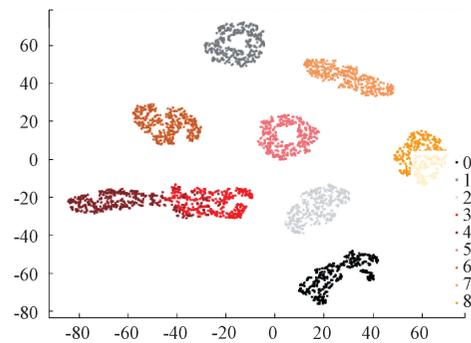
紧凑, 各个类会融合形成一簇, 难以区分; 从图 10(b) 可以看出, 各个类别之间开始渐渐独立, 相对于图 10(a) 会有一些的距离, 个别类之间互相缠绕有浅显的交叉现象; 从图 10(c) 可以看出, 大多类互相独立, 类与类之间有明显的边界, 只有两类之间会有较小的交叉。由此得出, 经过特征融合后, 不同暂降类别区分度更大, 更容易识别暂降源。



(a) 原始数据特征可视化
(a) Visualization of raw data features



(b) 原始数据+S变换数据特征可视化
(b) Raw data + S-transform data feature visualization



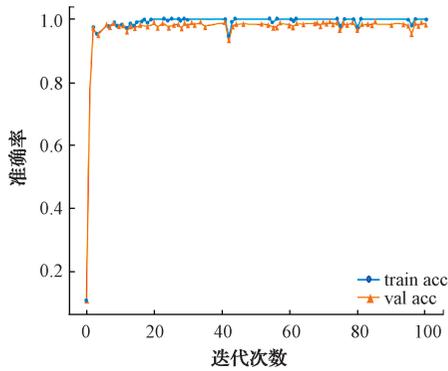
(c) 原始数据+S变换数据+指标数据特征可视化
(c) Raw data + S-transformed data + indicator data feature visualization

图 10 不同数据样本特征可视化

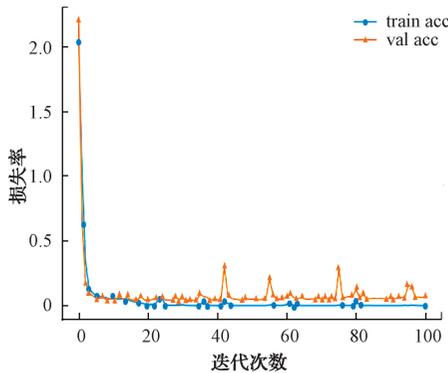
Fig. 10 Different data sample features visualization

首先是原始三相数据作为输入, 电压暂降识别模型训练集和验证集的正确率曲线和损失曲线如图 11 所示。在迭代 20 轮次左右, 训练和验证的正确率和损失率达到稳定, 但随着迭代次数的增加, 训练过程不稳定, 总体正确率可达到 98%。

其次是将 S 变换数据和初始数据同时作为输入, 电压



(a) 原始数据训练和验证集的正确率曲线
(a) Correctness curves for the training and validation sets of the original data



(b) 原始数据训练和验证集的损失率曲线
(b) Loss rate curves for the training and validation sets of the original data

图 11 原始数据训练集和验证集的损失率和正确率曲线

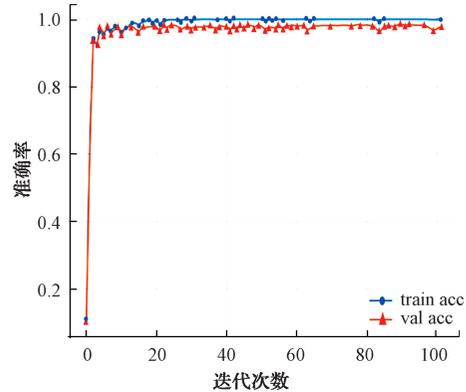
Fig. 11 The loss rate and correct rate curve of the original data training set and verification set

暂降识别模型训练集和验证集的正确率曲线和损失曲线如图 12 所示。在迭代 20 轮次左右,训练和验证的正确率和损失率达到稳定,随着迭代次数增加,训练和验证曲线上毛刺会减少,正确率可达到 99%,证明引入 S 变换数据后能够提高识别的准确率。

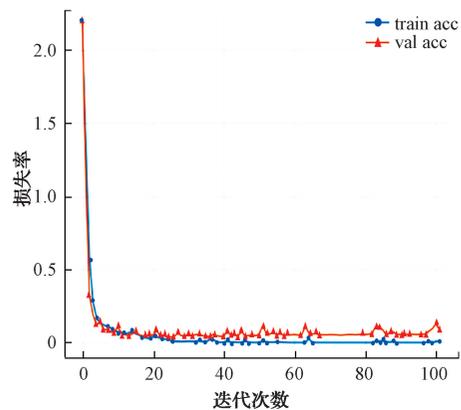
最后是 S 变换数据,原始数据以及指标数据所有数据作为输入,电压暂降识别模型训练集和验证集的正确率曲线和损失曲线如图 13 所示。在迭代 20 轮次左右,训练和验证的正确率和损失率达到稳定,随着迭代次数增加,训练和验证曲线上的毛刺消失,正确率能达到 99.6%,证明 3 种数据形式同时输入识别准确率最高,并且识别训练过程也更稳定。

表 2 是对于测试集不同数据形式下不同类别的识别准确率以及平均准确率。

从表 2 可以得出,原始数据+S 变换数据的识别率达到 98.48%,比原始数据识别率提高了 0.12%,在多数数据特征融合情况下,总体识别率达到 99.08%,与原始数据和原始数据+S 变换数据相比,分别高出 0.72%和 0.6%,证明多数数据特征融合可有效提高暂降源识别率。



(a) 原始+S数据训练和验证集的正确率曲线
(a) Correctness curves for the training and validation sets of the original+S data



(b) 原始+S数据训练和验证集的损失率曲线
(b) Loss rate curves for the training and validation sets of the original+S data

图 12 原始+S 数据训练集和验证集的损失率和正确率曲线

Fig. 12 The loss rate and correct rate curve of the original+S data training set and verification set

表 2 不同数据形式下的 S-CNN-LSTM 识别率
Table 2 S-CNN-LSTM recognition rate in different data forms

故障类 标签	原始 数据/%	原始数据+S 变换/%	原始数据+S 变 换+指标数据/%
0	99.75	100	99.75
1	100	100	100
2	100	100	100
3	93.47	91.77	97.7
4	93.03	94.57	95.6
5	99.26	100	99.01
6	100	100	100
7	99.75	100	99.75
8	100	100	100
平均值	98.36	98.48	99.08

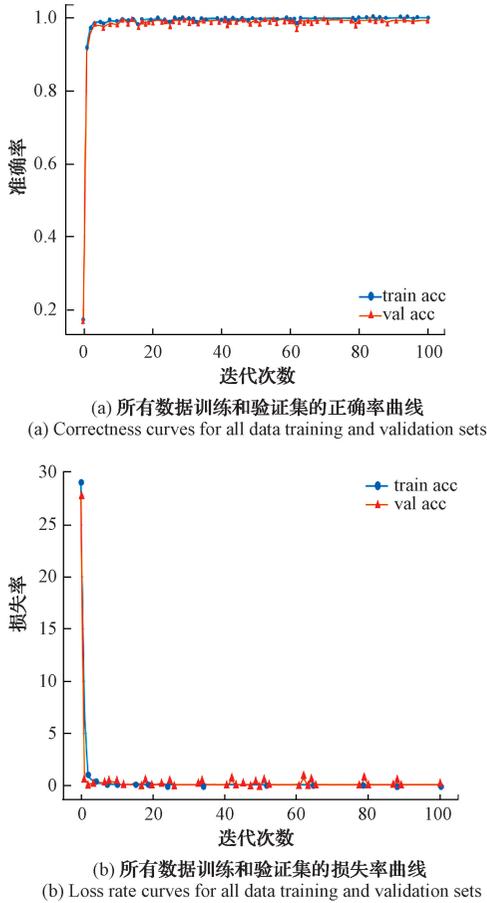


图 13 所有数据训练集和验证集的损失率和正确率曲线
Fig. 13 The loss rate and correct rate curve of all data training sets and verification sets

4.3 对比实验

为了对比不同模型,本文引用了以下评价指标,准确率(accuracy)、精度(precision)、召回率(recall)和 F1 分数^[21]进行模型评估。其表达式为:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (9)$$

$$Precision = \frac{TP}{TP + FP} \quad (10)$$

$$Recall = \frac{TP}{TP + FN} \quad (11)$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (12)$$

式中:Accuracy 为正确样本占总样本比值;Precision 为正样本结果中预测准确度;Recall 为正样本被正确预测的比例;F1 为精度和召回率的调和平均;FP 和 FN 分别表示假阳性和假阴性的数量;TP 和 TN 分别表示真阳性和真阴性的数量。

为了量化所提模型识别准确率,进行了消融实验,即对不同网络组成模型进行对比实验。比较模型包括

CNN、LSTM、CNN-LSTM、CNN-BiLSTM 和本文模型,所有模型隐藏层节点、batch size 和学习率等超参数设置完全一致。经过迭代训练后,各模型验证集识别效果如表 3 所示。

表 3 不同模型实验结果对比

Table 3 Different model experiment results comparison

(%)				
模型类别	精度	召回率	F1 分数	准确率
LSTM	76.5	76.3	76.2	76.3
CNN	87.3	87.5	85.9	87.5
CNN-LSTM	93.81	92.41	90.41	92.41
CNNBiLSM	94.29	94.36	94.36	94.36
本文	98.36	98.48	99.08	98.48

从表 3 可以发现,多尺度特征融合网络模型的识别效果明显好于单尺度特征提取网络模型。本文采用的模型平均检测准确率达到 98.48%,相比于 CNN-LSTM、CNN-BiLSTM 多尺度特征融合网络模型分别提高 6.07%、4.12%。实验结果表明,本文采用的 S-CNN-BiLSTM 模型相对于其他网络模型在特征提取能力和特征融合效果上有一定优势。

5 结论

为了实现对电压暂降源精准识别,本文采用了搭建数值模型框架方法获取电压暂降源数据集,确保有足够多的样本进行深度学习,在原有 S 变换的基础上引入了调整因子,克服了原来获取信息不充分的缺点,并引入 16 个指标提取信号特征得到指标特征,搭建 CNN-LSTM 模型,对数据同时进行空间特征和时序特征的提取,减少了训练参数,降低过拟合风险,最后将指标特征、空间特征以及时序特征融合后输入到全连接层识别,克服了传统神经网络的算法收敛速度慢和人工提取特征数据丢失的问题,提升了电能质量扰动识别辨识度,并且从多种输入数据样本进行比较验证了此模型识别的精准性。

参考文献

- [1] 肖先勇,肖宇,汪颖. 基于用户生产场景辨识的电压暂降经济损失评估[J/OL]. 电网技术, 1-15[2024-07-22].
XIAO X Y, XIAO Y, WANG Y. Voltage based on user production scenarios is temporarily reduced economic loss assessment [J/OL]. Power Grid Technology, 1-15[2024-07-22].
- [2] 郭微,杨家豪. 基于 BP 神经网络的低压变频器电压暂降耐受能力评估[J]. 电工电气, 2023(12): 49-53.
GUO W, YANG J H. The low-voltage inverter voltage of the BP neural network is temporarily reduced tolerance [J]. Electrician Electrical,

- 2023(12): 49-53.
- [3] 薛世银,李鹏,周杨.基于特征选择和贝叶斯判别的电压暂降源识别[J].电工技术,2023(17):58-62.
XUE SH Y, LI P, ZHOU Y. Based on characteristic choices and Bayesian, other voltage sources are temporarily lowered [J]. Electrician Technology, 2023(17): 58-62.
- [4] 陈晓华,吴杰康,陈盛语,等.基于EMD和IABC-SVM算法的复合电压暂降源辨识方法[J].广东电力,2022,35(2):11-18.
CHEN X H, WU J K, CHEN SH Y, et al. Composite voltage identification methods based on EMD and IABC-SVM algorithms [J]. Guangdong Power, 2022, 35(2): 11-18.
- [5] 张震,刘明萍,张镇涛,等.基于CEEMDAN和小波软阈值的电能质量扰动信号去噪新方法[J].现代电子技术,2021,44(18):63-68.
ZHANG ZH, LIU M P, ZHANG ZH T, et al. Power quality disturbance signal with CEEMDAN and wavelet soft thresholds [J]. Modern Electronic Technology, 2021, 44(18): 63-68.
- [6] 贾荣,张逸,林鸿伟,等.基于感知哈希序列的电压暂降事件同源识别[J].电力系统保护与控制,2023(3):133-144.
JIA R, ZHANG Y, LIN H W, et al. The voltage of the perception of the sequence of the sequence of perception of the incident was homologous to the same source identification [J]. Power System Protection and Control, 2023(3): 133-144.
- [7] 徐健,王磊.基于S变换模矩阵的电网扰动信号检测[J].电子测量技术,2018,41(21):20-24.
XU J, WANG L. Power grid disturbance signal detection based on the S-transforming mold matrix [J]. Electronic Measurement Technology, 2018,41(21): 20-24.
- [8] WANG S, CHEN H. A novel deep learning method for the classification of power quality disturbances using deep convolutional neural network[J]. Applied Energy, 2019, 235: 1126-1140.
- [9] 王凯,孙贤明,任成昊,等.基于多尺度融合选择卷积网络的电能质量扰动识别[J].现代电子技术,2022,45(4):107-112.
WANG K, SUN X M, REN CH H, et al. Electric energy quality disturbance identification based on multiple-scale integration networks [J]. Modern Electronic Technology, 2022,45(4): 107-112.
- [10] WU L, ZHANG Y, HAO X, et al. Research on a location method for complex voltage sag sources based on random matrix theory[J]. Mathematical Problems in Engineering, 2020(5):1-12.
- [11] 伍济钢,文港,杨康.改进一维卷积神经网络的航空发动机故障诊断方法[J].电子测量与仪器学报,2023,37(3):179-186.
WU J G, WEN G, YANG K. Aviation engine failure diagnosis method to improve the one-dimensional convolutional neural network [J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(3): 179-186.
- [12] 郝洪涛,王凯,张炳建,等.多尺度特征自适应融合的气动控制阀故障诊断[J].仪器仪表学报,2023,44(10):167-178.
HAO H T, WANG K, ZHANG B J, et al. Multi-scale characteristic self-adaptive fusion of pneumatic control valve fault diagnosis [J]. Chinese Journal of Scientific Instrument, 2023,44(10): 167-178.
- [13] 许圳兴,余耀,赵东,等.基于多尺度级联注意网络的肺实质分割[J].国外电子测量技术,2024,43(5):60-69.
XU ZH X, YU Y, ZHAO D, et al. Based on the substantial division of lungs based on the network of multiple-level pay attention [J]. Foreign Electronic Measurement Technology, 2024,43(5): 60-69.
- [14] 李丹奇,梅飞,张宸宇,等.基于深度置信网络的电压暂降特征提取及源辨识方法[J].电力系统自动化,2020,44(4):150-158.
LI D Q, MEI F, ZHANG CH Y, et al. The voltage temporary reduction and source recognition method based on the deep confidence network [J]. Automation of the Power System, 2020, 44(4): 150-158.
- [15] 郑智聪,王红,齐林海.基于深度学习模型融合的电压暂降源识别方法[J].中国电机工程学报,2019,39(1):97-104,324.
ZHENG ZH C, WANG H, QI L H. The voltage of the fusion of deep learning models is temporarily reduced sources [J]. China Electrical Engineering Journal, 2019,39(1): 97-104,324.
- [16] 龚正,邹阳,金涛,等.基于特征融合并行优化模型的电能质量扰动分类方法[J].中国电机工程学报,2023,43(3):1017-1027.
GONG ZH, ZOU Y, JIN T, et al. Electric energy quality disturbance classification method based on feature fusion and line optimization model [J]. China Electrical Engineering Journal, 2023, 43(3): 1017-1027.
- [17] 吴怀诚,刘家强,岳蕾,等.基于多特征融合的卷积神经网络的电能质量扰动识别方法[J].电网与清洁能源,2023,39(9):19-23,31.

- WU H CH, LIU J Q, YUE L, et al. Electric energy quality disturbance recognition method based on multi-characteristic fusion of convolutional neural networks [J]. *Power Grid and Clean Energy*, 2023, 39(9): 19-23,31.
- [18] TAN R H G, RAMACHANDARAMURTHY V K. Numerical model framework of power quality events[J]. *European Journal of Scientific Research*, 2010, 43(1): 30-47.
- [19] 卢彩霞,王新环,刘志超,等. 基于改进 S 变换- TT 变换的电压暂降特性研究[J]. *电子测量技术*, 2019, 42(4): 58-64.
- LU C X, WANG X H, LIU ZH CH, et al. Based on the temporary reduction characteristics of the voltage reduction of the S-transformation- TT transformation [J]. *Electronic Measurement Technology*, 2019, 42(4): 58-64.
- [20] 刘海涛,叶筱怡,吕干云,等. 引入调节因子改进 S 变换电压暂降源识别[J]. *电气传动*, 2021, 51(18): 59-64.
- LIU H T, YE X Y, LYU G Y, et al. Introducing regulation factors to improve the identification of voltage-dropout sources in S-conversion[J]. *Electrical Drives*, 2021, 51(18): 59-64.
- [21] LIU Y, JIN T, MOHAMED M A, et al. A novel three-step classification approach based on time-dependent spectral features for complex power quality disturbances[J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70: 1-14.

作者简介

张峰,硕士研究生,主要研究方向电力系统电压暂降源识别等。

E-mail:2939690619@qq.com

陈雷(通信作者),博士,副教授,主要研究方向为电能质量分析、电气设备状态监测与故障诊断等。

E-mail:10009941@qq.com