DOI:10.20033/j.1003-7241.(2025)06-0095-05

基于LSTM的电气化铁路牵引供电故障识别研究

翟政中[1]，黄永涛[2]，刘华云[3]

(1.中国铁路上海局集团有限公司建设部,上海 200000;
2.北京市轨道交通建设管理有限公司,北京 100000;
3.成都唐源电气股份有限公司,四川 成都 610000)

摘 要: 为了减少牵引供电故障对电气化铁路运行安全的影响,需要提供更准确、快速的牵引供电故障识别方法。为此研究在牵引供电故障分析的基础上,基于长短时记忆神经网络,构建故障识别模型,并引入卷积神经网络进一步加快识别速度。测试结果表明,研究所提模型的准确率达到了0.98,而单一的卷积神经网络为0.81,长短时记忆神经网络为0.82。相较于其他两个模型,研究所提模型的准确率至少提高了20%;此外,在数据量为100时,相较于长短时记忆神经网络,研究所提模型的运行时间减少了2.36 s,验证了研究的有效性。研究结果为保障牵引供电安全提供了参考。

关键词: 牵引供电;电气化铁路;故障识别;长短时记忆网络;卷积神经网络
中图分类号:TP183 文献标识码:A 文章编号:1003-7241(2025)06-0095-05

Research on Fault Identification of Electrified Railway Traction Power Supply Based on LSTM

ZHAI Zheng-zhong[1], HUANG Yong-tao[2], LIU Hua-yun[3]

(1. Construction Department of China Railway Shanghai Bureau Group Co., Ltd., Shanghai 200000, China;
2. Beijing Rail Transit Construction Management Co., Ltd., Beijing 100000, China;
3. Chengdu Tangyuan Electric Co., Ltd., Chengdu 610000, China)

Abstract: In order to reduce the impact of traction power supply failure on the operation safety of railway electrification, it is necessary to provide a more accurate and rapid identification method of traction power supply failure. For this reason, based on the analysis of traction power supply fault, a fault identification model is built based on the Long and Short Term Memory neural network, and the Convolutional Neural Network is introduced to further accelerate the identification speed. In the test results, the accuracy of the proposed model reaches 0.98, while the single Convolutional neural network is 0.81, and the Long and Short Term Memory neural network is 0.82. Compared to the other two models, the accuracy of the proposed model is improved by at least 20%. In addition, when the data volume is 100, compared to the Long and Short Term Memory neural network, the running time of the model proposed in the study decreases by 2.36 seconds, verifying the effectiveness of the study. The research aims to provide reference for ensuring the safety of traction power supply.

Keywords: traction power supply; railway electrification; fault identification; long short term memory network; convolutional neural network

0 引言

随着全球经济的快速发展和城市化进程的加速,铁路运输作为国民经济的大动脉,其重要性日益凸显[1]。电气化铁路的高效运行离不开稳定可靠的牵引供电系统。牵引供电系统的故障不仅会导致列车运行中断,影响旅客出行,还可能引发严重的安全事故,造成巨大的经济损失和社会影响[2]。因此,牵引供电故障的识别对于电气化铁路系统的正常运行极其重要。随着人工智能技术的快速发展,特别是深度学习技术的出现,为电气化铁路牵引供电故障识别提供了新的解决方案。其中,循环神经网络(recurrent neural network,RNN)因其强大的非线性特征提取能力和长时间记忆能力,成为当前故障诊断领域中的热门研究方向之一[3]。翁楦乔和文成林两位学者在电网故障诊断领域中,提出了基于深度特征聚类和RNN的智能诊断方法[4]。该方法通过卷积神经网络来提取时序数据的高层特征,然后进行半监督聚类,从而对无标签样本的故障类别进行识别。在与常用分类方法的对比实验中,验证了该方法优秀的故障诊断效果。然而,由于牵引供电故障类型复杂、特征参数不确定等因素,传统RNN

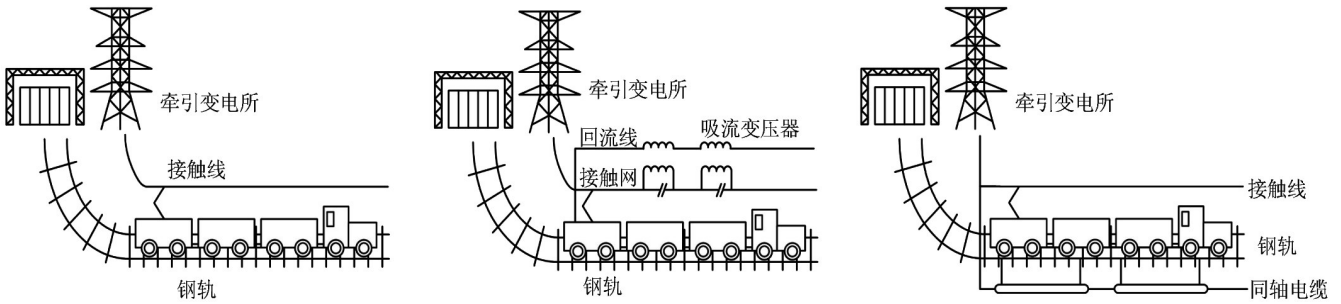
*基金项目:中央发改委深化电力体制改革专项(09-44B)
收稿日期:2023-10-23

模型识别效果有限。而在处理复杂故障识别的研究中，长短期记忆神经网络(long short-term memory，LSTM)展现出了独特的优势。戴森柏等学者提出了一种基于LSTM和稳态时间序列的诊断模型，用于复杂的光伏列阵故障诊断[5]。该方法通过稳态时间序列过滤暂态过程和干扰噪声，然后通过LSTM实现故障识别。在仿真实验中，该方法展现出了良好的精度和泛化性能。在上述研究的基础上，此次研究提出了一种基于LSTM的牵引供电故障诊断模型。并且为了提高LSTM的处理数据的速度，研究还引入了卷积神经网络(convolutional neural net work，CNN)来以降低数据长度，从而加快LSTM模型学习和训练的速度。此外，研究还设置了与CNN-LSTM和CNN以及LSTM的对比实验，以验证研究方法的有效性。此研究结果提高了牵引供电故障的识别效率，为电气化铁路安全运行提供了保障。

1 基于LSTM的电气化铁路牵引供电故障识别

1.1 电气化铁路牵引供电故障分析

牵引供电系统是电气化铁路的重要组成，该系统由一次供电网络、牵引变电所、接触网、分区所和开闭所等组成[6]。其中一次供电网络由高压输电线路和变电所组成，电压等级一般为110 kV或者是220 kV。牵引变电实现了供电网中交流电的变换，以满足电气化铁路的用电需求。牵引网为电气化铁路提供电力。分区所在同一电气化铁路线上的不同牵引变电所中，用于实现相邻供电区段的并联或单独工作。开闭所在变电站的下一级，用于实现分区供电，提高供电的可靠性[7]。直接供电、吸流变压器供电和同轴电缆供电的方式如图1所示。



(a) 直接供电　(b) 吸流变压器供电　(c) 同轴电缆供电

图1 直接供电、吸流变压器供电与同轴电缆供电对比

由图1可见，在进行直接供电时，牵引变电所通过两条馈线连接接触线和钢轨，实现对列车的供电。这种供电方式会形成较高的电位，所产生的电磁感应效应还会对附近的通信造成干扰，因此我国逐渐转用对通信影响更小的吸流变压器供电[8]。吸流变压器供电模式如图1(b)所示，在直接供电的基础上，额外加入回流线，并将吸流变压器设置在回流线和接触线之间。通过以1:1为变比的吸流式变压器串联方法，消除了电力供应时的电磁干扰。然而，吸流变压器供电方式，也存在着可靠性差、易产生电弧、末端电压下降明显等问题。因缺陷明显，同轴电缆供电方式被提出。图1(c)为同轴电缆供电方式示意图，该方式将同轴电缆铺设在铁路沿线，由于馈线和回流线使用同轴电缆，二者之间拥有更大的互感系数，能够抵消二者产生的磁场，起到降低通信影响的作用。但是同轴电力电缆造价昂贵，对铺设环境的要求较高，因此一般只在重要枢纽城市或一些特殊路段应用。而目前性价比最高的供电方式为自耦变压器供电，如图2所示。

通过使用自耦变压器，能够显著提高牵引网的电压，从而降低电压损失和功率损失，提升整体的供电效率。对于高速电气化铁路而言，这意味着列车可以在更高的速度下稳定运行，同时减少因供电问题导致的延误和中断，确保旅客的出行体验和货物运输的时效性。自耦变压器供电分为单线供电和复线供电。其中单线供电如图2所示。复线供电在单线供电的基础上，并联了上下行牵引网中的自耦变压器。作为我国高速铁路供电的主要方式，复线自耦变压器的供电故障将直接影响到电气化铁路的运作。根据调查研究可知，复线自耦变压器供电方式下的线路故障多为接触网直接短路故障，包括了接触线与正馈线、接触线与钢轨、正馈线与钢轨在内的三种短路类型。此次研究以上述短路类型为研究对象，提出了基于LSTM的电气化铁路牵引供电故障识别模型。

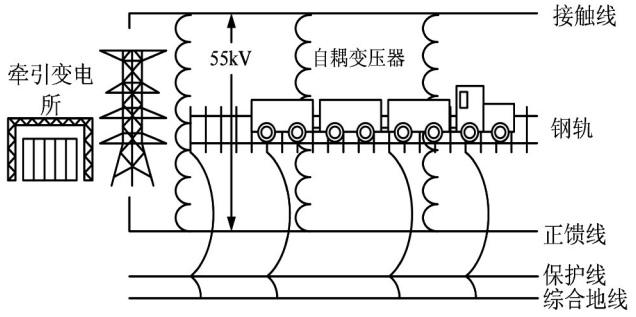


图2 自耦变压器供电方式示意图

1.2 基于LSTM的电气化铁路牵引供电故障识别模型

在故障识别领域，RNN是常见的深度学习方法之一。RNN的特殊性在于其具有独特的记忆功能，能够优化序列信息的处理。RNN通过在标准神经网络架构中引入额外的隐藏层，实现了信息在连续输入步骤间的传递，赋予了网络处理时间序列数据的能力[9]。是在训练过程中，RNN可能会遇到梯度消失或梯度爆炸的问题，这些问题通常与网络中信息的长距离传递有关。为了解决这些问题，LSTM应运而生[10]。LSTM网络是对RNN的一种改进，额外加入了遗忘门，输入门和输出门，使网络能记住更长的输入过程。LSTM输入门公式如式(1)所示。

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{1}$$

式中，i_t表示t时刻的输入门，W_i表示输入门的权重矩阵，$[h_{t-1}, x_t]$表示将两个向量连接成更长的向量，h_{t-1}和x_t分别表示两个矩阵中的输出项目，b_i表示输入门的偏置项，σ表示Sigmoid函数。遗忘门的公式如式(2)所示。

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{2}$$

式中，f_t表示t时刻的遗忘门，W_f表示遗忘门的权重矩阵，b_f表示遗忘门的偏置项，σ表示Sigmoid函数。遗忘门记忆公式如式(3)所示。

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{3}$$

式中，$\tilde{C}_t$表示t时刻的遗忘门记忆，W_C表示遗忘门记忆的权重矩阵，b_C表示遗忘门记忆的偏置项，tanh表示双曲正切函数。输出门公式如式(4)所示。

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{4}$$

式中，o_t表示t时刻的输出门，W_o表示输出门的权重矩阵，b_o表示输出门的偏置项，σ表示Sigmoid函数。长记忆、短记忆和公式如式(5)所示。

$$\begin{cases} C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \\ h_t = o_t \cdot \tanh(C_t) \end{cases} \tag{5}$$

式中，h_t表示长记忆，C_t表示短记忆，f_t表示遗忘门，$\tilde{C}_t$表示遗忘门记忆，o_t表示输出门，tanh表示双曲正切函数。通过上式表明，模型中的输出由长记忆决定，长记忆受短记忆影响，因此输出受到当前输入和长时间输入的影响。长短时网络模型中，每个单元之间都会传递长记忆向量和短记忆向量[11]。研究收集了牵引网络短路故障后的暂态信号，训练基于LSTM的短路故障识别模型。从牵引网络短路故障后的暂态信号中可以看出，不同故障的短路电流流向不同，因此具有较强的时序特征和频率特征，能够使用LSTM处理故障信号。但是在LSTM处理中，随着数据长度增加，LSTM训练消耗的时间也会显著增加，不利于模型的学习训练。为了减少数据长度，此次研究采用CNN与LSTM结合的方式，利用CNN提取数据特征以降低数据长度，从而实现模型学习训练速度的加快[12]。由于CNN是较为基础的网络模型，故在此处不再展开描述。研究构建的CNN−LSTM电气化铁路牵引供电故障识别模型，如图3所示。

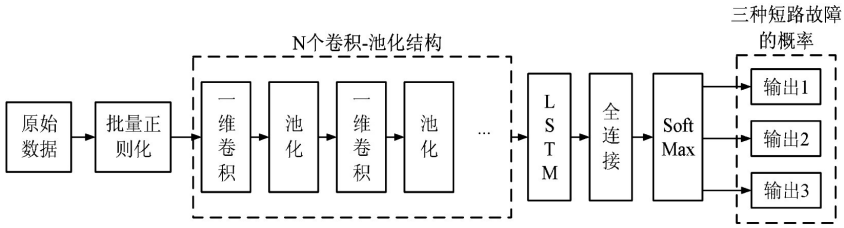


图3 CNN-LSTM的电气化铁路牵引供电故障识别模型

由于电气化铁路牵引供电故障信号为一维信号，因此研究选用一维CNN进行特征提取，以降低运算量。此外，一维CNN还能在训练过程中自我优化，以区分有效数据和冗余数据。在基于LSTM的电气化铁路牵引供电故障识别模型中，研究选择交叉熵作为评估预测误差的指标，并建立出训练时使用的损失函数。此次研究选用CNN网络的卷积−池化结构的数量为4。LSTM的参数量与参数的表达式如式(6)所示。

$$N_{params} = 4 \times [S_h \times (S_h + D) + S_h] \tag{6}$$

式中，N_{params}表示LSTM的参数量；S_h表示隐藏层尺寸；D表示输入数据的特征深度。在故障识别领域中，D的大小受到卷积神经网络的影响，研究根据之前经验设置D的数值为8，因此N_{params}与S_h的关系如式(7)所示。

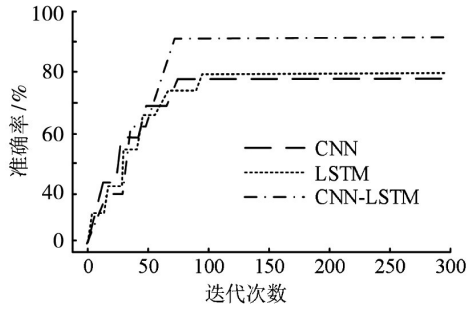
$$N_{params} = 4 \times S_h^{\,2} + 36 \times S_h \tag{7}$$

由式(7)可知，LSTM的参数量与隐藏层尺寸呈现正相关关系，所以可以通过扩大隐藏层尺寸的方式增加LSTM参数量，进而实现模型性能提升的目的。研究通过预实验，发现隐藏层尺寸为60时训练效果更好，因此将60作为CNN−LSTM电气化铁路牵引供电故障识别模型的隐藏层尺寸。

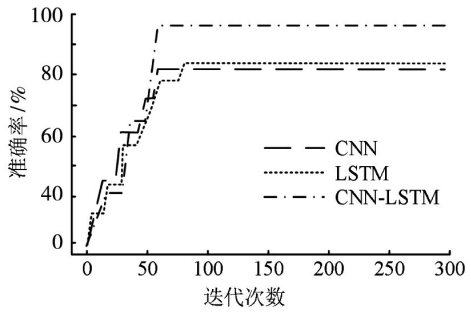
2 CNN-LSTM电气化铁路牵引供电故障识别模型

为了减小系统误差对实验结果的干扰，研究在同样的测试环境下进行测试，用于测试的计算机显卡为GTX

1080ti；CPU为Inter Xeon E5；内存为64 GB；操作系统为Windows 10；GPU为两块Quardro4000 8 GB。研究使用TensorFlow作为深度学习框架。基于LSTM的电气化铁路牵引供电故障识别模型中，卷积−池化结构的数量为4；输入数据的特征深度为8；隐藏层尺寸为60。训练与测试的数据来自此次研究收集的牵引网络短路故障后的暂态信号构成，其中70%的数据作为训练集，30%的数据作为测试集。为了验证此次研究的有效性，选用CNN、LSTM和研究提出的CNN−LSTM模型进行对比，得到模型训练准确率与测试准确率，如图4所示。



(a) 模型训练结果



(b) 模型测试结果

图4 不同模型的准确率对比

图4(a)反映了3种模型训练时的准确率曲线变化情况。从图中可以看出，CNN在第73次迭代时达到收敛，收敛时的准确率为0.78；LSTM在第84次迭代时达到收敛，收敛时的准确率为0.79；CNN−LSTM在第73次迭代时达到收敛，收敛时的准确率为0.91。从收敛速度来看，CNN−LSTM与CNN同样优秀，显著优于LSTM。从收敛准确度来看，相较于CNN和LSTM，CNN−LSTM收敛时的准确率分别提升了0.13和0.12。可以看出在模型训练时，CNN−LSTM的表现更好。图4(b)反映了3种模型测试时的准确率曲线变化情况。从图中可以看出，LSTM测试时的准确率达到了0.98，而CNN为0.81，LSTM为0.82。相较于其他两个模型，CNN−LSTM的准确率至少提高了20%，验证了此次研究的有效性。

研究进一步对不同故障类型的识别准确率进行测试，实验设置了30个数据组，每个数据中包含了同等数量的接触线与正馈线、接触线与钢轨、正馈线与钢轨3种短路类型。获得实验结果如图5所示。

如图5所示，CNN−LSTM故障识别的准确率平均为0.95，CNN故障识别的准确率平均为0.77，LSTM故障识别的准确率平均为0.81。相较于CNN和LSTM模型，CNN−LSTM识别的准确率有显著的提升。此外，从图中还可以看出，CNN−LSTM仅在第5、15、18、23组数据中出现了较大的准确率偏差，占总组数的13.33%。而CNN在第4、8、9、10、15、23、24组数据中都出现了较大的准确率偏差，占总组数的23.33%。LSTM在第7、8、12、16、18、21、22、23、27组数据中出现了较大的准确率偏差，占总组数的30%。因此可以看出，CNN−LSTM不仅在故障识别中具有较高的准确率，而且有着更为稳定的识别性能。

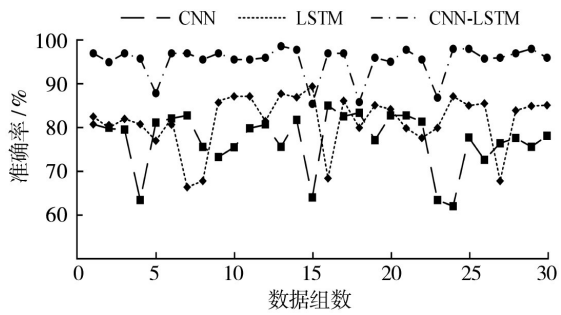


图5 不同故障类型识别准确率结果

此次研究采用CNN结合LSTM的方法，主要是为了去除冗余数据对计算的影响，从而加快LSTM的计算速度。为了验证该方法的有效性，研究对不同模型在不同数据量下的运行时间进行了对比实验。实验结果如图6所示。

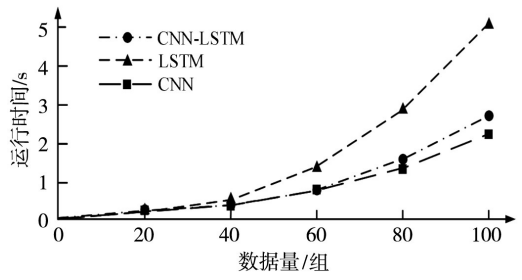


图6 模型运行时间对比

从图6中可以看出，3种模型的运行时间随着数据量的增加而增加。其中LSTM的运行时间随数据量变化而增加的幅度最大，当数据量为60时与CNN和CNN−LSTM产生明显的差异。此时LSTM运行时间为1.19 s；CNN与CNN−LSTM的运行时间均为0.84 s。当数据量为100时，3种模型的运行时间之间的差异比较明显，此时LSTM运行时间为5.04 s；CNN−LSTM的运行时间为2.68 s；CNN的运行时间为2.23 s。计算可知，在数据量为100时，相较于LSTM，CNN−LSTM的运行时间减少了2.36 s；相较于CNN，CNN−LSTM的运行时间增加了0.45 s。可

(下转第179页)