融合 BiLSTM 与改进减法平均优化器的矿山设备异常检测研究*

吕宗岩，苑喜文

（矿冶科技集团有限公司， 北京 100160）

摘要：针对传统矿山设备异常检测效率低、精度不足的问题，提出融合双向长短记忆网络与改进减法平均优化器的矿山设备异常检测模型。引入注意力机制实现关键特征动态加权，结合卷积神经网络提取局部时序特征，采用引导正则化随机森林完成特征选择，降低数据维度并抑制噪声干扰。改进减法平均优化器引入混沌映射初始化与黄金正弦步长自适应调整策略，显著提升了模型的收敛速度与训练稳定性。结果表明，在矿山设备多传感器时序数据集与矿山机械故障模拟数据集中，本文所提模型的准确率分别达到94.17%和95.48%，综合性能优于现有主流检测模型；消融试验中，集成所有关键模块时，模型的准确率、精准率、召回率和 F_1 分数分别达到98.25%、97.62%、98.05%、97.84%。在某大型露天矿场的皮带输送机和破碎机等关键设备实测中，模型的检测准确率提升至97.32%，平均检测时间仅为0.39 s，表明该模型可高效、精准地检测矿山设备异常状态，为保障矿山生产安全提供了可靠的技术支持。

关键词：双向长短记忆网络(BiLSTM)；减法平均优化器；卷积神经网络；深度学习；矿山设备异常检测

中图分类号：TD407　　**文献标识码**：A　　**文章编号**：1005-2763(2026)06-0163-11

Research on Mining Equipment Anomaly Detection Integrating BiLSTM and an Improved Subtraction Average Optimizer

Lyu Zongyan, *Yuan Xiwen*

(BGRIMM Technology Group Co., Ltd., Beijing 100160, China)

Abstract: To address the issues of low efficiency and insufficient accuracy in traditional mining equipment anomaly detection, a mining equipment anomaly detection model integrating bidirectional long short memory network and an improved subtraction average optimizer was proposed. Attention mechanism was introduced to achieve dynamic weighting of key features. Convolutional neural network was combined to extract local temporal features, and guided regularized random forest was employed to perform feature selection, thereby reducing data dimensionality and suppressing noise interference. The improved subtraction average optimizer introduced chaotic mapping initialization and golden-sine step-size adaptive adjustment strategy, which significantly enhanced the convergence speed and training stability of the model. The results show that on the mining equipment multisensor dataset and the mining machinery fault simulation dataset, the accuracy of the proposed model reaches 94.17% and 95.48%, respectively, and its comprehensive performance surpasses that of existing mainstream detection models. In ablation experiments, when all key modules are integrated, the accuracy, precision, recall, and F_1 score of the model reach 98.25%, 97.62%, 98.05% and 97.84%, respectively. In field tests on key equipment such as belt conveyors and crushers at a large open-pit mine, the detection accuracy of the model increases to 97.32%, with an average detection time of only 0.39 s. This indicates that the model can detect abnormal states of mining equipment efficiently and accurately, providing reliable technical support for ensuring mining production safety.

Key words: Bidirectional long short memory network(BiLSTM), Subtractive average optimizer, Convolutional neural network, Deep learning, Mining equipment anomaly detection

* **收稿日期**：2025-07-02

作者简介：吕宗岩(1979—)，男，黑龙江齐齐哈尔人，硕士，高级工程师，主要从事矿山自动化方面的研究。E-mail：13513341232@163.com

通信作者：苑喜文(1979—)，男，黑龙江绥化人，正高级工程师，主要从事矿山电气方面的研究。E-mail：yuanxiwen@bgrimm.com

0 引言

随着矿山行业向智能化、自动化方向快速发展,矿山设备的高效稳定运行已成为保障矿山生产安全与效率的核心要素。矿山设备是矿山资源开采的重要组成部分,设备运行状态直接影响矿山安全和生产效率[1]。由于矿山作业环境复杂恶劣,设备长期处于高负荷、高磨损状态,易产生性能退化与机械故障[2]。因此,及时、准确地检测矿山设备的异常状态,对于预防故障发生、保障矿山生产的安全性和稳定性具有重要的研究意义。然而,传统矿山设备检测方法以人工巡检、阈值判断和信号处理为主,存在检测效率低、精度不足、泛化性能差等缺陷[3]。近年来,随着深度学习技术的迅速发展,双向长短记忆网络(Bidirectional Long Short Memory Network, BiLSTM)可融合双向时序上下文信息,全面挖掘设备运行数据的动态规律,进而提升异常检测的准确性,已逐渐成为故障检测领域的研究热点[4]。

郝磊[5]提出了基于卷积神经网络(Convolutional Neural Network, CNN)和 BiLSTM 的电缆故障自动化定位方法,通过 CNN 提取电缆信号特征并利用 BiLSTM 进行时序数据建模,显著提高了故障定位精度,使定位误差大幅降低。Nacer 等[6]提出了基于 BiLSTM 神经网络的轴承故障诊断新方法,通过 BiLSTM 网络对轴承振动信号的时序特征进行建模,自动提取故障特征,在多种工况下对不同故障类型的诊断准确率超过了 95%。覃仕明等[7]提出了基于 CNN 与 BiLSTM 的无刷直流电机故障诊断方法,采用 CNN 提取故障信号的时空特征,利用 BiLSTM 对时序特征进行建模,该方法对多种故障类型的识别效果优异。Chen 等[8]提出了基于 BiLSTM 的铁路车辆悬挂故障检测方法,通过优化 BiLSTM 网络结构,增强了其对悬挂系统振动信号的特征提取能力,使对早期故障的识别能力显著增强。张寿岩等[9]提出了基于 BiLSTM 与改进鲸鱼优化算法的电力变压器故障诊断方法,通过改进 BiLSTM 网络结构,增强对故障特征的提取能力,并引入混合策略对鲸鱼优化算法进行改进,所提方法的故障诊断准确率分别提高了 10.42%和 7.85%。

优化算法直接决定了深度学习模型的训练性能与泛化能力。传统梯度下降类优化器易陷入局部最优解,存在收敛速度慢、泛化能力弱等缺陷[10]。减法平均优化器作为一种新兴的优化方法,通过动态可自适应调整参数更新步长,兼顾全局探索与局部开发能力,可有效提升模型在复杂数据下的鲁棒性和收敛效率[11]。陈闯等[12]提出了基于柔性残差神经网络的滚动轴承智能故障诊断方法,通过残差模块增强网络对复杂故障特征的提取能力,并引入减法平均优化器提升训练效率,使得故障诊断准确率显著提高。Dai 等[13]提出了基于改进减法平均优化器的配电网故障定位方法,通过优化减法平均算法,增强对故障特征的提取能力,并结合网络拓扑进行故障区段定位,使得故障定位精度显著提高。费红博等[14]提出了基于变分模态分解与多周期相关峭度滤波的微弱故障信号降噪及特征增强方法,通过变分模态分解信号,利用多周期相关峭度滤波进行特征提取与降噪处理,使得故障诊断准确率显著提高。Trojovsky 等[15]提出了基于减法平均优化器的新型算法,用于解决优化问题,通过模拟群体中个体的减法平均行为,实现对解空间的高效搜索和优化,使得收敛速度更快,全局搜索能力更强。

综上,深度学习方法已在设备的故障检测领域广泛应用,但现有模型在矿山设备异常检测中仍存在检测效率低、精度不足等问题,难以适配矿山非平稳、高噪声的时序监测数据,模型的适应性和鲁棒性不足。为此,本研究构建了一种基于 BiLSTM 和减法平均优化器的矿山设备异常检测模型,以实现矿山设备的高效、精准检测。本研究的创新性在于:融合注意力机制动态加权关键特征,结合卷积神经网络提取局部时序特征,采用引导正则化随机森林进行特征选择,降低数据维度与噪声干扰,提升模型对复杂高维数据的处理能力和检测精度;对减法平均优化器进行改进,提出混沌映射初始化和黄金正弦步长自适应调整策略,提升模型的收敛速度与稳定性,增强矿山在非平稳、高噪声数据下的鲁棒性和适应性。

1 矿山设备异常检测模型构建

1.1 BiLSTM 模型构建

矿山设备的运行状态检测是保障矿山安全生产、提升开采效率的核心环节。传统检测方法存在效率低、精度差等问题,且难以提取高维非线性时序数据的深层特征。随着深度学习技术的快速发展,BiLSTM 可融合前向和后向时序信息,能够更全面地挖掘时序数据的动态特征[16]。为此,本研究构建了一种基于 BiLSTM 的矿山设备异常检测模型,充分利用 BiLSTM 在处理时序数据方面的优势,通过

融合前向与后向时序信息,对设备运行状态的变化规律进行精准捕捉。

BiLSTM 模型由前向、后向两层长短记忆网络(LSTM)组成,可双向解析输入时序序列,其表达式如式(1)所示[17]。

$$\begin{cases}\vec{c}_t = L_{\mathrm{STM}}(\boldsymbol{x}_t, \vec{c}_{t-1}) \\ \overleftarrow{c}_t = L_{\mathrm{STM}}(\boldsymbol{x}_t, \overleftarrow{c}_{t-1}) \\ C_t = \boldsymbol{W}_{x\vec{c}}\vec{c} + \boldsymbol{W}_{\overleftarrow{c}y}\overleftarrow{c} + \boldsymbol{b}_y \end{cases} \tag{1}$$

式中:$\vec{c}_t$ 和 $\overleftarrow{c}_t$ 分别表示时间步为 t 的前向与后向 LSTM 单元的隐藏状态;$\boldsymbol{x}_t$ 表示时间步为 t 的输入特征向量;$\vec{c}_{t-1}$ 和 $\overleftarrow{c}_{t-1}$ 分别表示时间步为 $t-1$ 的隐藏状态;C_t 为前向与后向 LSTM 单元的输出组合;$\boldsymbol{W}_{x\vec{c}}$ 和 $\boldsymbol{W}_{\overleftarrow{c}y}$ 分别为输入到隐藏层的权重矩阵和隐藏到输出层的权重矩阵;$\boldsymbol{b}_y$ 为输出层的偏置向量。BiLSTM 模型的结构如图 1 所示。

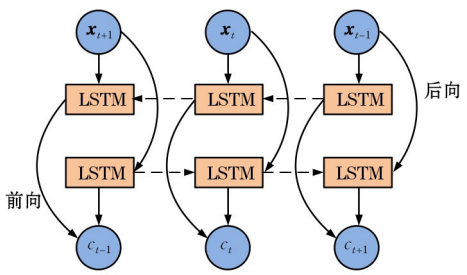


图 1 BiLSTM 模型的结构

Fig.1 Structure of BiLSTM model

图 1 包含两个 LSTM 层,分别处理输入序列的正向和反向信息。$\boldsymbol{x}_{t-1}$、$\boldsymbol{x}_t$ 和 $\boldsymbol{x}_{t+1}$ 为时间序列数据在不同时间步的输入特征向量,c_{t-1}、c_t 和 c_{t+1} 为在不同时间步的细胞状态。细胞状态作为 LSTM 网络中的一种信息存储机制,允许信息在网络中跨越多个时间步长保持不变。各 LSTM 单元融合上一个时间步的输出和当前时刻输入,生成当前时间步的输出和细胞状态。前向 LSTM 从左到右处理序列,后向 LSTM 从右到左处理序列,以提供更全面的序列特征表示。

1.2 注意力机制的引入

在矿山设备异常检测中,BiLSTM 模型虽然能够处理时序数据并提取特征,但在处理复杂的高维数据时,其难以自适应区分特征重要度,易造成关键异常信息淹没。注意力机制可实现特征自适应加权,强化关键异常特征,抑制冗余特征的干扰,提高模型识别精度[18]。为此,本研究在 BiLSTM 模型中引入了注意力机制,以增强模型对异常状态的识别能力。注意力机制通过聚焦任务相关关键信息,实现高维特征的自适应筛选,弱化无关信息干扰。其基本结构包含三个关键组件:查询、键和值。通过计算查询和键的相似度,生成注意力分数,并将其归一化为权重,最后按权重加权聚合值,得到上下文相关的输出。注意力机制的结构如图 2 所示。

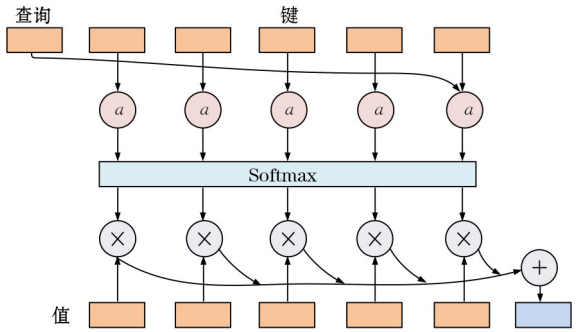


图 2 注意力机制的结构

Fig.2 Structure of attention mechanism

图 2 中,查询为当前时间步的隐藏状态,用于计算关注点;键为所有时间步的隐藏状态;值为待加权的特征表示,与注意力权重相乘后生成上下文向量。Softmax 函数用于注意力得分归一化。a 为注意力权重,+为所有加权后的值向量求和,生成最终的上下文向量。×用于将注意力权重与对应的值向量相乘。注意力分数由查询与键的相似度计算得到,并归一化为权重后加权聚合值,输出上下文相关的特征。注意力机制为 BiLSTM 输出的每个时间特征分配一个权重,权重的大小反映了该特征在当前任务中的重要性。本研究通过加权求和的方式,使模型能够关注重要的特征,提高模型检测性能。设输入的序列为 $C=\{c_1, c_2, \cdots, c_t\}$,其中 $c_t \in \boldsymbol{R}^d$,d 为隐藏层维度。采用 Softmax 函数对权重进行归一化处理,并计算其注意力得分,计算公式如式(2)所示[19]。

$$\begin{cases} a_t = \dfrac{\exp(e_t)}{\sum_{k=1}^{T}\exp(e_k)} \\ e_t = \boldsymbol{v}^{\mathrm{T}}\tan\boldsymbol{c}(\boldsymbol{W}c_t + \boldsymbol{b}) \end{cases} \tag{2}$$

式中:a_t 为注意力权重;e_t 为注意力得分;e_k 为第 k 个输入元素的注意力得分;$\boldsymbol{v}^{\mathrm{T}}$ 为权重向量的转置;$\boldsymbol{c}$ 为上下文向量;$\boldsymbol{W}$ 为权重矩阵;$\boldsymbol{b}$ 为可学习的偏置向量。本研究将隐藏状态按注意力权重加权求和,得到上下文向量,如式(3) 所示。

$$F = \sum_{t=1}^{T} a_t c_t \tag{3}$$

式中,F 为加权后的特征表示。融合注意力机制的 BiLSTM 模型结构如图 3 所示。

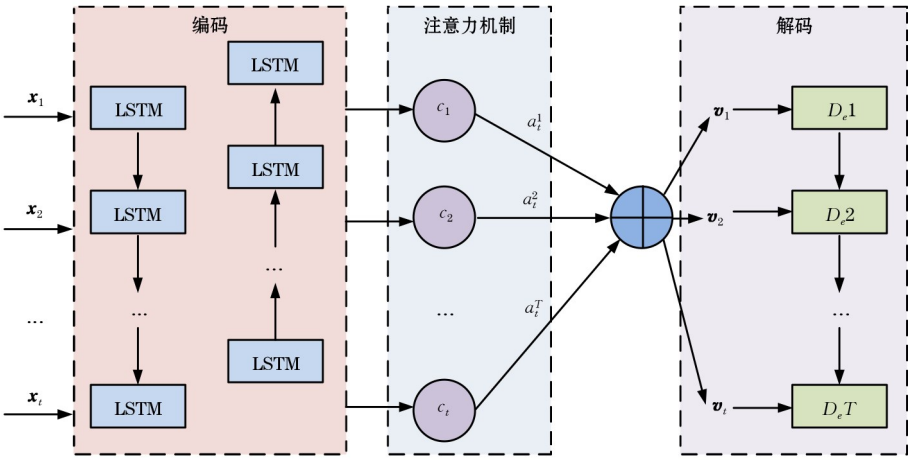


图 3 融合注意力机制的 BiLSTM 模型结构

Fig.3 Structure of BiLSTM model integrating attention mechanism

图 3 中，$\boldsymbol{x}_1$、$\boldsymbol{x}_2$ 和 $\boldsymbol{x}_t$ 分别为输入序列中不同时间步的输入特征向量；c_1、c_2 和 c_t 分别为在编码阶段不同时间步的细胞状态；a_i^T 为时间步 T 对应细胞状态的注意力权重；$\boldsymbol{v}_1$、$\boldsymbol{v}_2$ 和 $\boldsymbol{v}_t$ 分别为在解码阶段不同时间步的输出向量；D_e1、D_e2 和 D_eT 分别为解码阶段生成的最终输出。该模型融合了 BiLSTM 和注意力机制的优势，由编码、注意力机制和解码三部分构成。编码阶段通过 BiLSTM 层提取输入序列，生成隐藏状态。注意力机制随后计算权重，对隐藏状态进行动态加权，强化关键信息表征。解码层基于加权后的特征进行预测或生成。在本研究提出的矿山设备异常检测模型中，注意力机制被引入到 BiLSTM 模型中，用于动态加权关键特征。BiLSTM 模型首先从时序数据中提取双向时序隐藏特征，然后通过注意力机制为每个时间步的特征分配自适应权重。通过加权求和的方式，模型能够更加关注重要的特征，显著提升模型对设备异常状态的识别精度。

1.3 基于引导正则化随机森林的特征选择方法

矿山设备异常检测中的传感器数据具有高维度、强噪声和非线性等特性。直接采用原始特征训练模型易导致计算复杂度升高、过拟合风险增大。特征选择是提升模型性能的关键步骤，可有效降低数据冗余、抑制噪声干扰，进而提高模型的泛化能力与检测精度。随机森林具有良好的特征重要性评估能力，被广泛应用于特征选择领域。但传统随机森林仍存在过拟合风险，且自助采样方式易造成特征重要性评估不稳定等问题[20]。为此，本研究在传统随机森林的基础上进行改进，提出一种引导正则化随机森林方法，通过引入正则化项和时序引导策略，提升特征选择过程的鲁棒性和针对性。

本研究通过计算各特征的重要性，实现特征重要性评估，为降低噪声特征带来的干扰，引入了正则化项对特征重要性进行修正，如式(4)所示[21]。

$$I_j^{\mathrm{GR}} = I_j - \lambda \cdot \sum_{k \in S} s_{\mathrm{im}}(j,k) \cdot I_k \tag{4}$$

式中：I_j^{GR} 为引入正则化项的特征重要性得分；I_j 为特征 j 的原始重要性得分；λ 为正则化系数；S 为已选择的特征子集；$s_{\mathrm{im}}(j,k)$ 为特征 j 与特征 k 的相似度；I_k 为特征 k 的原始重要性得分。在每棵树的训练过程中，本研究按照概率优先选取高重要性特征，如式(5) 所示。

$$P_j = \frac{I_j^{\mathrm{GR}}}{\sum_{i=1}^{M} I_i^{\mathrm{GR}}} \tag{5}$$

式中：P_j 为特征 j 被选择的概率；M 为总特征数。

通过上述方式，正则化项可有效抑制相似特征间的冗余性，突出关键特征的贡献度。针对时序数据特征，采用滑动窗口技术分时段计算特征重要性。首先将时序数据划分为多个滑动窗口，各窗口包含固定数量的时间步；在各滑动窗口内独立计算特征重要性得分。随后融合所有滑动窗口的特征重要性得分结果，得到全局重要性得分。该方法可有效捕捉时序数据中的局部特征变化规律，同时避免全局特征选择易忽略局部有效信息的问题。最终融合得到全局特征重要性，如式(6) 所示。

$$I_j^{\mathrm{fi}} = \frac{1}{N_w} \sum_{w=1}^{N_w} I_{i,w}^{\mathrm{GR}} \tag{6}$$

式中：I_j^{fi} 为特征 j 的最终全局重要性得分；N_w 为滑动窗口数；$I_{i,w}^{\mathrm{GR}}$ 为窗口 w 内的正则化重要性得分。在矿山设备检测系统中，各传感器实时采集设备在不同时刻的运行状态数据。数据中可能存在许多冗余特征或噪声特征，直接用于模型训练会导致模型

性能下降。采用引导正则化随机森林方法,可高效地筛选出对异常检测贡献度高的重要特征。

1.4 结合 CNN 的特征提取与降维方法

虽然通过引导正则化随机森林方法可从原始数据中筛选出重要性较高的特征,但矿山设备多变量时序数据仍包含复杂的局部关联特征与时序依赖关系,仅依靠特征选择难以充分挖掘数据深层信息。CNN 具有出色的局部特征提取能力和权值共享机制,被广泛应用于图像和时序数据处理领域[22]。因此,本研究将 CNN 与 BiLSTM 网络进行结合,实现时序特征的高效提取与数据降维。

本研究将多变量的时序信号作为输入数据,记为 $X \in R^{T\times D}$,采用一维卷积核完成输入数据的局部特征提取,卷积操作的计算如式(7) 所示。

$$h_i^{(l)} = \sigma\left(\sum_{j=1}^{k} w_j^{(l)} \cdot x_{i+j-1}^{(l-1)} + b(l)\right) \tag{7}$$

式中:$h_i^{(l)}$ 为第 l 层第 i 个位置的卷积输出;$w_j^{(l)}$ 为第 l 层卷积核的第 j 个权重;$x_{i+j-1}^{(l-1)}$ 为第 $l-1$ 层输入数据中,第 $i+j-1$ 个位置的输入特征值;$b(l)$ 为第 l 层的偏置;σ 为激活函数。

经多层卷积运算后,采用最大池化对特征进行下采样处理,降低数据维度,池化操作的计算如式(8) 所示。

$$p_i^{(l)} = \max(h_{i\times s}^{(l)}, h_{i\times s+1}^{(l)}, \cdots, h_{i\times s+s-1}^{(l)}) \tag{8}$$

式中:$p_i^{(l)}$ 为池化输出;$h_{i\times s+t}^{(l)}$ 为第 l 层卷积输出中,第 $i\times s+t$ 个位置的卷积特征值,$t=0,1,\cdots,s-1$;s 为池化步长。

为进一步增强特征表征能力,本研究将CNN提取的局部特征与原始时序数据进行融合,生成低维融合特征表示,如式(9) 所示。

$$F_{\mathrm{fi}} = \sigma(W_{\mathrm{fc}} \cdot F_{\mathrm{fused}} + b_{\mathrm{fc}}) \tag{9}$$

式中:F_{fi} 为生成的低维特征表示;F_{fused} 为融合特征;W_{fc} 和 b_{fc} 分别为全连接层的权重和偏置。本研究将降维后的特征输入 BiLSTM 模型,进一步捕捉时序数据的动态规律。

1.5 减法平均优化器的改进

在矿山设备异常检测任务中,优化算法的性能对模型训练效率和检测精度具有决定性作用。传统的优化器易陷入局部最优,且存在收敛速度慢等问题,进而导致模型泛化能力降低。减法平均优化器可通过动态调整参数更新步长,实现全局探索与局部开发能力的平衡,但其在复杂时序数据下的鲁棒性与收敛效率仍有待提升[23]。因此,结合矿山设备检测数据特性,本研究对减法平均优化器进行改进,以增强模型训练的稳定性与检测异常精度。基于 BiLSTM 与改进减法平均优化器的矿山设备异常检测模型结构如图 4 所示。

图 4 中,输入层接收多传感器检测数据,经特征选择模块筛选有用信息。优化层采用减法平均优化器处理数据,进一步输入卷积层与池化层进行特征提取。特征经过展平后,输入融合注意力机制的 BiLSTM 层进行编码和解码,最终输出设备异常检测结果。传统减法平均优化器采用随机初始化方法生成初始参数,导致搜索空间覆盖不均匀。为此,本研究引入一种 Piecewise 混沌映射方法生成初始参数,增强搜索空间的全局遍历性[24],其表达式如式(10)所示。

$$x_{n+1} = \begin{cases} x_n / p_c, & 0 \leqslant x_n < p_c \\ (1-x_n)/(1-p_c), & p_c \leqslant x_n < 1 \end{cases} \tag{10}$$

式中:x_{n+1} 为下一个混沌状态;x_n 为当前混沌状态;p_c 为控制参数。

传统减法平均优化器的步长更新依赖于平均运算,难以适配矿山设备数据的复杂特性。黄金正弦算法融合黄金比例和正弦函数的全局搜索能力,可在优化过程中实现步长自适应调整,有效平衡全局探索与局部开发能力。为进一步提升优化器的综合性能,本研究引入黄金正弦策略,实现步长的自适应动态调整,强化算法全局探索与局部开发能力[25],该方法的步长更新如式(11)所示。

$$\Delta\theta_t = \frac{\pi}{2} \cdot (\sin(\varphi_1) - \gamma_1 \cdot \sin(\varphi_2)) \cdot (\theta_{t-1} - \theta_{t-2}) \tag{11}$$

式中:$\Delta\theta_t$ 为在第 t 次迭代中步长的更新值;φ_1 和 φ_2 为黄金比例相关的相位角,具体值为 $1-\frac{\sqrt{5}-1}{2}$;γ_1 为增加步长更新的随机数;θ_{t-1} 和 θ_{t-2} 分别为第 $t-1$ 次和第 $t-2$ 次迭代中的步长值。

通过黄金正弦策略,优化器可在全局探索阶段快速调整步长,有效规避陷入局部最优;在局部开发阶段,自适应减小步长,提高局部搜索精度。在改进减法平均优化器中,本研究首先使用混沌映射生成初始参数,保证初始搜索点在搜索空间内均匀分布,避免陷入局部最优。在每次迭代中,结合当前的优化状态,实现步长动态调整。黄金正弦步长调整策略可根据优化过程自适应调整步长大小,实现全局探索与局部开发的动态平衡。随着迭代次数的增加,步长逐渐减小,优化器逐渐聚焦于局部最优解进行精细搜索,显著提升了搜索精度。

1.6 矿山设备异常检测方法

对减法平均优化器进行改进后,本研究将其应用于矿山设备异常检测任务,以提升模型整体性能

和检测精度。基于 BiLSTM 与减法平均优化器的 矿山设备异常检测模型的流程如图 5 所示。

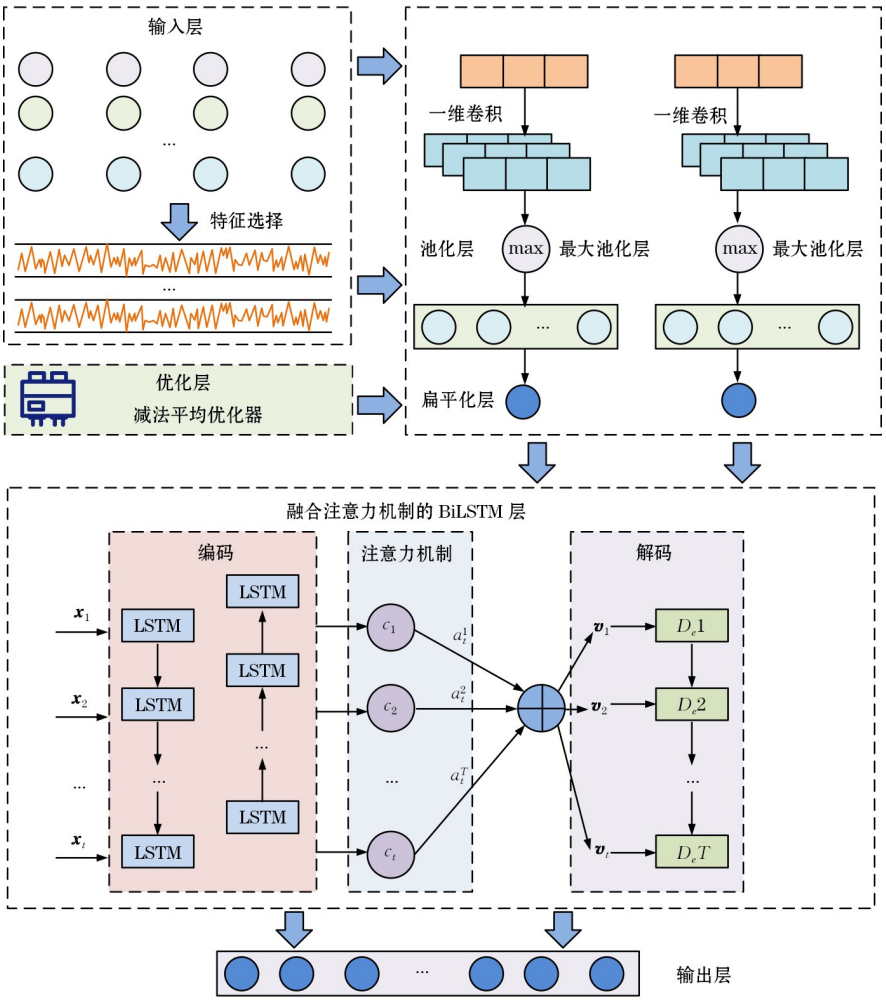


图 4 基于 BiLSTM 与减法平均优化器的矿山设备异常检测模型结构

Fig.4 Structure of mining equipment anomaly detection model based on BiLSTM and subtraction average optimizer

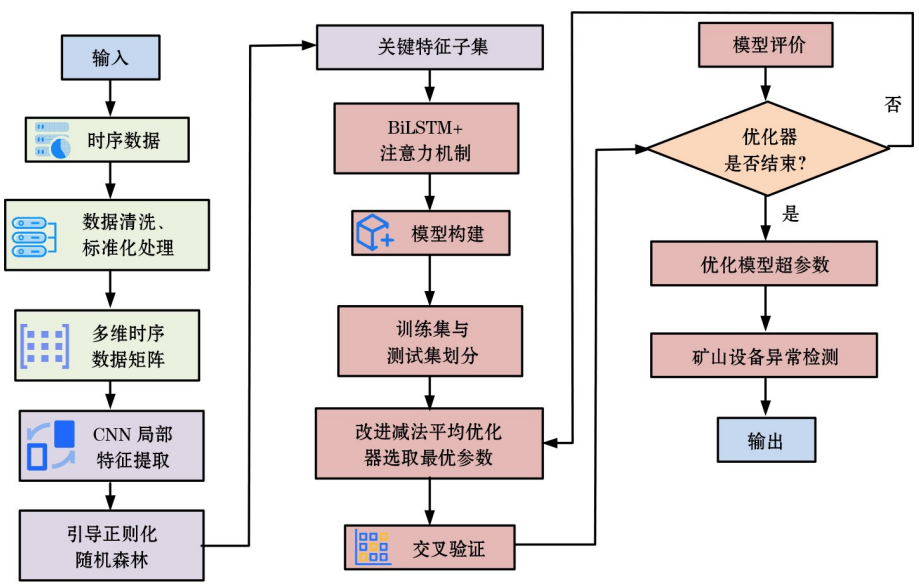


图 5 基于 BiLSTM 与减法平均优化器的矿山设备异常检测模型的流程

Fig.5 Process of mining equipment anomaly detection model based on BiLSTM and subtraction average optimizer

图 5 中,输入时序数据并完成数据清洗与标准化处理,构建多维时序数据矩阵;利用 CNN 进行局部特征提取,并引入正则化随机森林实现特征优选。该方法采用改进减法平均优化器完成最优参数选择,并开展交叉验证。最后,对模型进行评估并执行矿山设备异常检测,输出最终结果。

2 试验结果与分析

2.1 试验环境及参数设置

为验证基于 BiLSTM 与减法平均优化器的矿山设备异常检测模型的有效性和可行性,本研究在以下试验环境中开展试验,见表 1。

表 1 试验硬件环境与软件环境

Table 1 Hardware environment and software environment of experiments

硬件环境		软件环境	
名称	信息	名称	信息
处理器	Intel Core i9-12900K 3.2 GHz	操作系统	Windows 10
内存	64 GB DDR4	Python	3.8.10
显卡	NVIDIA GeForce RTX 3090	深度学习框架	PyTorch 1.10.0
存储	2 TB SSD 4 TB HDD	数据处理	NumPy 1.21.2 Pandas 1.3.3

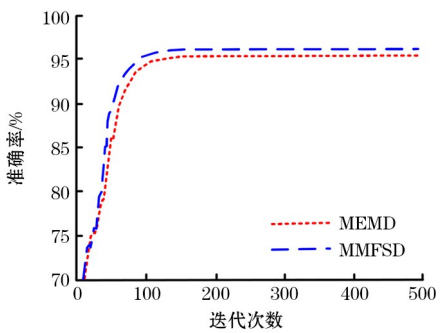
2.2 试验数据集及评估指标

为验证所提模型对矿山设备的异常检测性能,采用两组具有代表性的数据集进行试验:矿山设备多传感器时序数据集(Mining Equipment Multisensor Dataset, MEMD)和矿山机械故障模拟数据集(Mining Machinery Fault Simulation Dataset, MMFSD)。其中 MEMD 取自某矿业集团设备真实监测数据,涵盖多种矿山设备实际运行过程中的多传感器时序数据,采样频率为 10 Hz;矿山设备包括电动挖掘机、破碎机、通风机、矿用卡车和带式输送机等;具体的监测参数包括振动信号、温度数据、电流信号、电压信号、油压数据和油温数据等。MMFSD 为基于数字孪生平台和物理仿真模型生成的模拟数据集。该平台采用 SolidWorks 构建设备三维模型,利用 ANSYS 开展多物理场仿真,模拟皮带输送机、球磨机等设备在 6 种典型故障模式(轴承磨损、齿轮断裂、电机过热等)下的运行状态;仿真参数设置与实际设备保持一致,采样频率为 5 kHz,包含振动、温度、电流等 8 类监测信号。数据集中 70%的数据为训练集,15%的数据为测试集,剩余 15%的数据为验证集。模型训练设置大小为 64,最大迭代次数为 500。

本研究采用多种常用的评估指标,从多维度衡量模型的准确性、鲁棒性和检测效率,保证模型在实际应用中的可靠性和有效性。评估指标包括准确率、精准率、召回率、F_1 分数、模型收敛性能与检测时间。

2.3 矿山设备异常检测模型性能验证分析

首先在 MEMD 和 MMFSD 两组数据集中,对本文所提模型的训练准确率与训练损失进行分析,结果如图 6 所示。图 6(a)中,当迭代次数为 50 时,模型在 MEMD 和 MMFSD 数据集下的准确率分别为 87.36%和 89.24%。当迭代次数达到 200 时,模型在两组数据集中的准确率分别为 94.17%和 95.48%。图 6(b)中,迭代次数为 50 时,模型在两组数据集中的损失值分别为 0.115 和 0.096。迭代次数达到 200 时,模型损失值分别为 0.042 和 0.035。结果显示,基于 BiLSTM 与改进减法平均优化器的矿山设备异常检测模型在两组数据集中均表现出良好的训练性能与收敛稳定性。



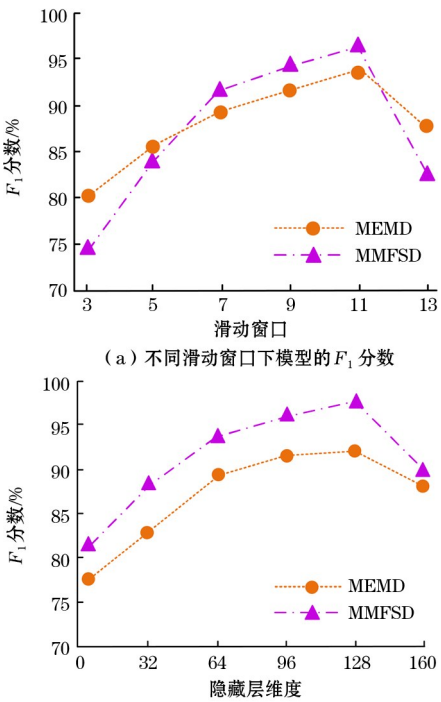
(a) 两组数据集中模型的准确率

(b) 两组数据集中模型的损失值

图 6 MEMD 和 MMFSD 数据集中模型的训练准确率与损失结果

Fig.6 Training accuracy and loss results of the model in MEMD and MMFSD datasets

为验证不同参数配置下模型的检测性能,本研究在两组数据集中,分析不同滑动窗口与隐藏层维度下的 F_1 分数,结果如图 7 所示。图 7(a)中,滑动窗口为 3 时,MEMD 和 MMFSD 数据集中的 F_1 分数分别为 80.25%和 74.91%;当滑动窗口达到 11 时,模型在 MEMD 和 MMFSD 数据集中的 F_1 分数最高,分别为 92.56%和 95.08%。图 7(b)中,隐藏层维度为 32 时,模型在 MEMD 和 MMFSD 数据集中的 F_1 分数分别为 82.84%和 88.59%。当隐藏层维度达到 128 时,MEMD 和 MMFSD 数据集中的 F_1 分数最高,分别为 91.36%和 96.24%。结果表明,当滑动窗口为 11、隐藏层维度为 128 时,模型在两组数据集中的 F_1 分数均最高,综合检测性能最佳。



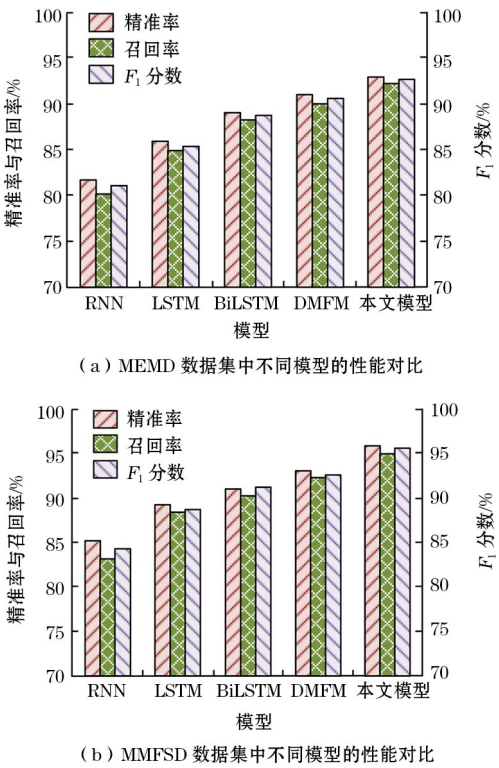
(a) 不同滑动窗口下模型的 F_1 分数

(b) 不同隐藏层维度下模型的 F_1 分数

图 7 不同滑动窗口与隐藏层维度下模型的 F_1 分数

Fig.7 F_1 scores of the model under different sliding windows and hidden layer dimensions

为验证所提模型的综合优越性,在两组数据集中,将其与现有主流检测模型的精准率、召回率和 F_1 分数进行对比分析。包括循环神经网络(Recurrent Neural Network, RNN)、LSTM、BiLSTM 和深度多模态融合模型(Deep Multimodal Fusion Model, DMFM),结果如图 8 所示。由图 8(a)可知,在 MEMD 数据集中,RNN、LSTM、BiLSTM 和 DMFM 的精准率分别为 82.15%、86.73%、88.92%和 90.24%,召回率分别为 80.34%、85.22%、87.61%和 89.15%,F_1 分数分别为 81.23%、85.97%、88.26%和 89.69%。本文所提模型的精准率、召回率和 F_1 分数分别为 93.58%、92.47%和 93.02%。由图 8(b)可知,在 MMFSD 数据集中,5 种模型的精准率分别为 85.62%、89.47%、91.85%、93.17%和 96.12%,召回率分别为 83.91%、88.01%、90.67%、92.08%和 95.03%,F_1 分数分别为 84.75%、88.74%、91.25%、92.65%和 95.57%。结果显示,本文所提模型在两组数据集中均具备优异的检测性能。



(a) MEMD 数据集中不同模型的性能对比

(b) MMFSD 数据集中不同模型的性能对比

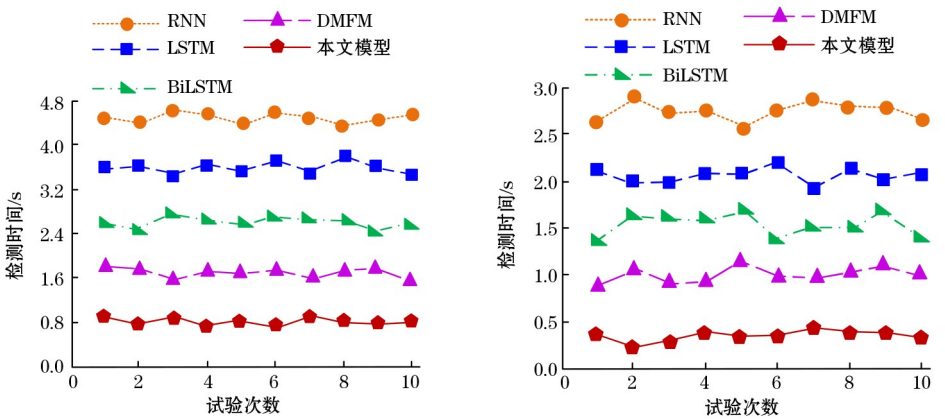
图 8 不同检测模型的性能对比结果

Fig.8 Performance comparison results of different detection models

本研究进一步在两组数据集中,对所提模型的设备异常检测效率进行了分析,以验证模型的实时性,结果如图 9 所示。图 9(a)中,RNN、LSTM、BiLSTM 和 DMFM 模型在 MEMD 数据集中的平均检测时间分别为 4.58 s、3.62 s、2.57 s 和 1.74 s。与上述模型相比,本文所提模型的平均检测时间仅 0.86 s,检测耗时分别缩短了 81.22%、76.24%、66.54%和 50.57%。图 9(b)中,RNN、LSTM、BiLSTM 和 DMFM 模型在 MMFSD 数据集中的平均检测时间分别为 2.72 s、2.08 s、1.55 s、0.93 s。与上述模型相比,本文所提模型的平均检测时间仅 0.38 s,检测耗时分别缩短了 86.03%、81.73%、75.48%和 59.14%。结果表明,本文所提模型在两组数据集中均表现出优越的检测效率,验证了本文

模型的高效性。

为量化各个关键模块对基于 BiLSTM 与改进减法平均优化器的矿山设备异常检测模型性能的实际贡献，本研究进行了消融试验，结果见表 2。由表 2 可知，模型在集成所有关键模块后综合性能最佳，准确率、精准率、召回率和 F_1 分数分别达到 98.25%、97.62%、98.05%和 97.84%，表明各模块均对提高模型性能至关重要。



（a）MEMD 数据集中不同模型的检测时间对比　（b）MMFSD 数据集中不同模型的检测时间对比

图 9　不同模型的设备异常检测效率对比

Fig.9　Comparison of equipment anomaly detection efficiency among different models

表 2　消融试验结果

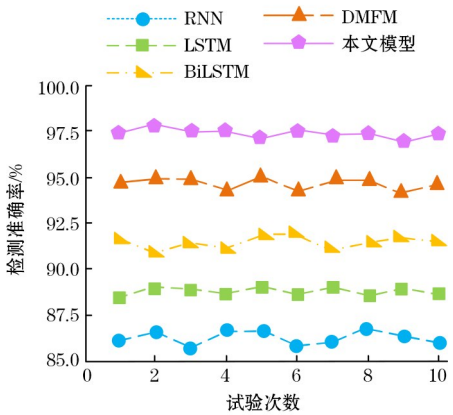
Table 2　Results of ablation experiments

BiLSTM	注意力机制	引导正则化随机森林	CNN	减法平均优化器	准确率/%	精准率/%	召回率/%	F_1 分数/%
√					76.54	77.44	76.37	76.91
	√				79.63	80.12	80.25	80.19
		√			81.57	82.96	82.58	82.77
			√		85.48	85.25	84.69	84.97
				√	87.62	88.24	87.79	88.02
√	√				89.44	90.16	88.57	89.37
√	√	√			92.12	92.95	92.58	92.77
√	√	√	√		96.31	96.19	96.90	96.55
√	√	√	√	√	98.25	97.62	98.05	97.84

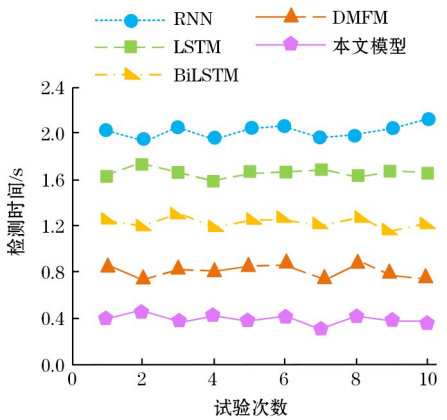
2.4　实际应用效果分析

为验证本文所提模型的工程实际应用效果，本研究在某大型露天矿场的关键设备上开展了实时检测试验。试验选取 WK-12 型矿用挖掘机和 MT5500 型矿用卡车作为检测对象；在设备关键测点安装 PCB 352C33 型加速度传感器、PT100 铂电阻温度传感器和 LEM CKSR 系列电流传感器，搭建边缘计算节点并部署各检测模型；通过 5G 专网实现数据实时传输，其中振动信号为 10 kHz。试验连续采集设备 30 d 的运行数据，每 24 h 开展一次全模型性能测试，记录 5 种模型在相同测试设备上的检测时间和准确率；采用滑动窗口方法处理时序数据，对比不同模型的实时检测性能。

本研究对不同模型的检测时间与准确率进行了分析，结果如图 10 所示。图 10（a）中，RNN、LSTM、BiLSTM、DMFM 和本文所提模型的平均检测准确率分别为 86.12%、88.45%、91.27%、94.48%和 97.32%。与其他 4 种模型相比，本文所提模型的准确率分别提升了 13.01%、10.03%、6.63%和 3.01%。图 10（b）中，RNN、LSTM、BiLSTM、DMFM 和本文所提模型的实际平均检测时间分别为1.92 s、1.68 s、1.22 s、0.78 s 和 0.39 s。与其他 4 种模型相比，本文所提模型的检测时间分别减少了 79.69%、76.79%、68.03%和 50.00%。结果显示，基于 BiLSTM 与改进减法平均优化器的矿山设备异常检测模型在实际应用中具有优异的检测精度和检测效率。



(a) 不同模型的实际检测准确率对比



(b) 不同模型的实际检测时间对比

图 10 不同模型的实际检测时间与检测准确率对比

Fig.10 Comparison of actual detection time and detection accuracy of different models

3 结论

本研究提出一种基于 BiLSTM 与改进减法平均优化器的矿山设备异常检测模型,经理论分析和试验验证,得到如下主要结论。

(1) 在 MEMD 和 MMFSD 数据集中,本文所提模型的检测准确率分别达到 94.17%和 95.48%,损失值分别为 0.042 和 0.035,表明该模型在矿山设备异常检测任务中表现出优异的检测精度和收敛稳定性。

(2) 消融试验结果表明,集成所有关键模块时,模型的准确率、精准率、召回率和 F_1 分数分别达到 98.25%、97.62%、98.05%和 97.84%,表明注意力机制、引导正则化随机森林方法和改进减法平均优化器均对模型性能的提升具有显著正向作用。

(3) 工程实时检测试验中,本文所提模型的准确率为 97.32%,检测时间为 0.39 s,综合性能均优于各对比模型。表明该模型在矿山设备检测场景中具有良好的应用价值,可满足实时检测要求,为设备安全运行提供了可靠的技术支持。

此次研究的局限性在于模型实时检测中对硬件资源的要求较高,不利于在低算力资源场景下部署。未来研究可以开展轻量化模型设计,降低硬件需求,推动模型在边缘设备上的应用。

参考文献(References):

[1] 李富伟,孙玉澄,李广,等.非煤地下矿山机电设备全生命周期安全技术研究与应用[J].矿业研究与开发,2022,42(8):165-172.
Li Fuwei, Sun Yucheng, Li Guang, et al. Research and application of safety technology for the whole life cycle of electromechanical equipment in non-coal underground mines [J]. Mining Research and Development, 2022, 42(8): 165-172.

[2] Lei Zhiyong, Ma Xiaolong, Zhao Shujun, et al. Intelligent collaborative management and control platform for continuous mining equipment in open-pit mines[J]. Coal Science and Technology,2025,53(4):362-372.

[3] Pincheira M, Antonini M, Vecchio M. Integrating the IoT and blockchain technology for the next generation of mining inspection systems[J]. Sensors,2022,22(3):899-913.

[4] Meng Hongyu, Zhang Jianliang, Cai Zhaolong, et al. Research on DC microgrid fault diagnosis based on CNN-BiLSTM-Attention[J]. Proceedings of the CSEE,2023,45(4):1369-1380.

[5] 郝磊.基于 CNN 和 BiLSTM 的电缆故障自动化定位技术[J].自动化与仪表,2025,40(3):118-121+161.
Hao Lei. Cable fault-automation positioning technology based on CNN and BiLSTM [J]. Automation and Instrumentation,2025,40(3):118-121+161.

[6] Nacer S M, Nadia B, Abdelghani R, et al. A novel method for bearing fault diagnosis based on BiLSTM neural networks [J]. The International Journal of Advanced Manufacturing Technology,2023,125(3):1477-1492.

[7] 覃仕明,马鹏.基于 CNN-A-BiLSTM 的无刷直流电机故障诊断方法研究[J].计算机测量与控制,2024,32(9):118-124+148.
Qin Shiming, Ma Peng. Research on diagnosis method of brushless DC motor based on CNN-A-BiLSTM[J]. Computer Measurement & Control,2024,32(9):118-124+148.

[8] Chen Yuejian, Niu Gang, Li Yifan, et al. A modified bidirectional long short-term memory neural network for rail vehicle suspension fault detection [J]. Vehicle System Dynamics,2023,61(12):3136-3160.

[9] 张寿岩,史卫刚,杨利国,等.改进 BiLSTM 在电力变压器故障诊断中的应用研究[J].电测与仪表,2024,61(5):160-165.
Zhang Shouyan, Shi Weigang, Yang Liguo, et al. Research on the application of improved BiLSTM in power transformer fault diagnosis[J]. Electrical Measurement & Instrumentation,2024,61(5):160-165.

[10] Patel V, Berahas A S. Gradient descent in the absence of global lipschitz continuity of the gradients[J]. SIAM Journal on Mathematics of Data Science,2024,6(3):602-626.

[11] 张柄棋,刘云平,王爽,等.改进减法平均优化算法的移动机器人路径规划[J].电子测量技术,2024,47(24):57-64.
Zhang Bingqi, Liu Yunping, Wang Shuang, et al. Improved subtraction-average-based optimizer algorithm for mobile robot path planning [J]. Electronic Measurement Technology,2024,47(24):57-64.

[12] 陈闯,李先锋,史建涛,等.基于柔性残差神经网络的滚动轴承智能故障诊断方法[J].工程科学学报,2025,47(3):480-488.
Chen Chuang, Li Xianfeng, Shi Jiantao, et al. Intelligent fault diagnosis method for rolling bearings based on flexible residual neural network[J]. Journal of Engineering Science, 2025,47(3):480-488.

[13] Dai Jinjin, Zhang Ziyu, Li Shaoyong, et al. Research on fault section location in an active distribution network based on improved subtraction-average-based optimizer [J]. Symmetry,2025,17(1):107-127.

[14] 费红博,张超,吴乐,等.基于 VMD-MCKD 的微弱故障信号降噪及冲击特征增强方法[J].机电工程,2025,42(2):237-246.
Fei Hongbo, Zhang Chao, Wu Le, et al. Method for denoising weak fault signals and enhancing impact characteristics based on VMD-MCKD[J]. Mechanical and Electrical Engineering,2025,42(2):237-246.

[15] Trojovsky P, Dehghani M. Subtraction-average-based optimizer:a new swarm-inspired metaheuristic algorithm for solving optimization problems[J]. Biomimetics,2023,8(2):149-190.

[16] Galdelli A, Narang G, Tomassini S, et al. Data imputation in large and small-scale spatiotemporal time series gaps using BackForward Bi-LSTM[J]. Journal of Big Data,2025, 12(1):1-18.

[17] Xu Yuanyuan, Yang Jiapeng, Cai Xin. Intelligent analysis algorithm for power engineering data based on improved BiLSTM[J]. Scientific Reports,2025,15(1):15320-15335.

[18] Kavianpour P, Kavianpour M, Jahani E, et al. A CNN-BiLSTM model with attention mechanism for earthquake prediction[J]. The Journal of Supercomputing, 2023, 79 (17):19194-19226.

[19] Shobana J, Murali M. An improved self attention mechanism based on optimized BERT-BiLSTM model for accurate polarity prediction[J]. The Computer Journal, 2023,66(5):1279-1294.

[20] Sunarya P A, Rahardja U, Chen S C, et al. Deciphering digital social dynamics: a comparative study of logistic regression and random forest in predicting e-commerce customer behavior[J]. Journal of Applied Data Sciences, 2024,5(1):100-113.

[21] Thakur D, Biswas S. Guided regularized random forest feature selection for smartphone based human activity recognition [J]. Journal of Ambient Intelligence and Humanized Computing,2023,14(7):9767-9779.

[22] Wu Jimmy Mingtai, Li Zhongcui, Herencsar N, et al. A graph-based CNN-LSTM stock price prediction algorithm with leading indicators[J]. Multimedia Systems, 2023, 29 (3):1751-1770.

[23] Jiang Yaxuan, Ding Yipeng. A target localization algorithm for a single-frequency doppler radar based on an improved subtractive average optimizer[J]. Remote Sensing,2024,16 (12):2123-2141.

[24] Xu Wenyuan, Cui Chuang, Ji Yongcheng, et al. A UAV path planning algorithm for bridge construction safety inspection in complex terrain[J]. Scientific Reports,2025,15 (1):13564-13580.

[25] Majumdar P, Mitra S. Enhanced honey badger algorithm based on nonlinear adaptive weight and golden sine operator[J]. Neural Computing and Applications,2025,37(1):367-386.