

基于残差网络的互联网终端设备状态智能诊断与AI辅助故障修复研究

张子樵¹

ZHANG Ziqiao

摘要

随着互联网终端设备数量的大幅增长,传统的故障诊断方法和修复方法有着响应速度较慢、低诊断准确率等缺点。文章提出了一种基于改进残差网络的互联网终端设备状态智能诊断与AI辅助的故障修复方法。通过引入长短期记忆网络(LSTM)和注意力机制来增强残差网络对时序数据的特征提取能力,同时利用Dropout层缓解了过拟合的问题,提高了故障诊断的精度与收敛速度。在故障修复方面,构建了基于强化学习决策的AI辅助故障修复系统,实现了从故障诊断到修复的闭环管理。实验结果表明,所提方法故障诊断准确率达到98.2%,修复效率提升86%,为互联网终端设备的智能运维提供了有效的解决方案。

关键词

残差网络;故障诊断;人工智能;故障修复;互联网终端;注意力机制;LSTM

doi: 10.3969/j.issn.1672-9528.2026.06.022

0 引言

近年来,互联网的各种终端设备已经完全融入日常生活与工作中。2025年,我国的互联网宽带接入的用户数达到了6.86亿户^[1]。由于各类型终端设备的大量增长,传统故障诊断方法是依赖人工的,存在诸多挑战^[2]。传统的故障诊断方法主要是依据规则或较为简单的机器学习模型,但提取不到复杂时序的故障特征。另外,目前的故障修复方法很大程度上依赖工程师的经验,导致修复效率低,尤其是复杂故障^[3-4]。近年来,深度学习逐渐应用到故障诊断方面^[5]。残差网络是一种典型的深度神经网络模型,为解决深层网络存在的梯度消失的问题,引入了残差连接^[6-7]。

因此,本研究在残差网络基础上,加入LSTM与注意力机制,目的在于提高对设备的时序特征的提取能力。另外,本文构建了一个AI辅助的故障修复系统,实验证明了该系统在实际场景的有效性。

1 基于改进残差网络的诊断模型

1.1 总体架构

本研究提出的改进残差网络模型包括:输入预处理、LSTM-ResNet混合特征提取、多重注意力机制、正则化与特征融合、分类输出等模块。

(1) 输入预处理:本模型输入的是多源时序数据,比如设备的性能指标(CPU、内存占用率)、事件日志、网络

流量的统计等。采用Z-score标准化对数值类型的指标进行处理,对日志事件嵌入表示,映射为128维稠密向量。

(2) LSTM-ResNet混合层:此层是本文模型的核心组成部分,把LSTM的时序处理能力与ResNet的深度特征提取能力结合起来,在每个残差块内部嵌入LSTM单元。设残差块输入为 x_t ,LSTM单元在 t 时刻的隐藏状态与细胞状态用公式表示为:

$$\begin{aligned}f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\\tilde{C}_t &= \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \\C_t &= f_t \times C_{t-1} + i_t \times \tilde{C}_t \\o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\h_t &= o_t \times \tanh(C_t)\end{aligned}\quad (1)$$

式中: x_t 表示 t 时刻的输入, h_t 表示隐藏状态, C_t 表示细胞状态。

残差块采用3层卷积,卷积核大小3×3,步长1,输出通道数依次为64、128、256。LSTM单元隐藏层维度设为256,细胞状态维度与隐藏层一致。残差块的卷积输出与LSTM单元的隐藏状态 h_t 通过元素加权求和实现融合,权重系数 $\omega \in [0.3, 0.7]$ 由验证集自适应学习得到,最终优化值 $\omega=0.52$,用公式表示为:

$$F_{\text{fusion}} = \omega \cdot F_{\text{ResNet}} + (1 - \omega) \cdot h_t \quad (2)$$

式中: F_{ResNet} 为残差块的特征输出,该融合方式既保留ResNet的深层空间特征,又融入LSTM的时序依赖信息。

(3) 多重注意力机制:本文采用通道注意力+时序注

1. 中国移动通信集团湖北有限公司 湖北武汉 430000

意力的双重注意力模块，具体结构如下：

通道注意力分支：对 LSTM-ResNet 融合后的特征图 $F \in \mathbb{R}^{C \times T}$ ，通过全局平均池化得到通道特征向量 $v_c \in \mathbb{R}^C$ ，再经 2 层全连接层和 Sigmoid 激活，生成通道注意力权重 $\alpha_c \in \mathbb{R}^C$ ，公式为：

$$\alpha_c = \sigma(W_2 \cdot \delta(W_1 \cdot v_c + b_1) + b_2) \quad (3)$$

式中： δ 为 ReLU 激活函数， W_1 、 W_2 为可学习权重矩阵。

时序注意力分支：采用多头自注意力机制，头数 $h=8$ ，将特征图按时间步展开为 $F_t \in \mathbb{R}^{T \times C}$ ，通过查询向量 $Q = F_t \cdot W_q$ 、键向量 $K = F_t \cdot W_k$ 、值向量 $V = F_t \cdot W_v$ 计算时序注意力权重，用公式表示为：

$$\alpha_t = \text{Softmax}\left(\frac{Q \cdot K^T}{\sqrt{C/h}}\right) \cdot V \quad (4)$$

通道注意力加权后的特征与时序注意力输出进行元素相乘，得到双重增强特征 $F_{\text{att}} = \alpha_c \odot F \odot \alpha_t$ ，有效聚焦故障相关的通道和时间步特征。

(4) Dropout 正则化层：为缓解高维特征空间中易出现的过拟合问题，在每个残差块的全连接层后插入 Dropout 层，随机丢弃率为 0.3。该设置基于网格搜索实验确定，能在保留有效特征的同时提升模型泛化能力。

(5) 分类输出层：最后层采用 Softmax 函数将提取的高层特征映射为不同设备状态的概率分布，输出维度对应故障类型的数量。

1.2 模型训练与优化

模型训练采用两阶段策略，首先在大型公开数据集上进行预训练，然后在特定领域数据上进行微调。预训练采用渐进式学习策略，第一阶段：冻结 LSTM 单元，仅训练 ResNet 主干网络，学习率设置为 5×10^{-4} ，训练 50 个 epoch；第二阶段：解冻全部参数，以 1×10^{-3} 学习率联合优化，训练 100 个 epoch；第三阶段：精细调优注意力机制参数，学习率降至 5×10^{-5} ，训练 20 个 epoch。预训练的目标函数为多任务联合损失：

$$L_{\text{pretrain}} = \lambda_1 L_{\text{type}} + \lambda_2 L_{\text{severity}} + \lambda_3 L_{\text{contrastive}} \quad (5)$$

式中： $\lambda_1=0.6$ 、 $\lambda_2=0.3$ 、 $\lambda_3=0.1$ 。

自适应微调阶段即在目标领域数据上采用课程学习策略，按故障难度逐步调整训练样本分布。简单样本：前 20 个 epoch；中等样本：中间 30 个 epoch；困难样本：最后 20 个 epoch。微调过程中采用判别性学习率，不同网络组件设置不同的学习率，ResNet 主干为 1×10^{-4} ，LSTM 时序模块为 3×10^{-4} ，注意力机制为 5×10^{-4} 。

优化过程中使用自适应矩估计算法，该算法的权重衰减策略有利于控制模型复杂度。将初始学习率设置为 0.001，并采用余弦退火策略动态调整，更新规则为：

$$\eta_t = \eta_{\min} + \frac{1}{2}(\eta_{\max} - \eta_{\min})(1 + \cos(\frac{T_{\text{cur}}}{T_{\max}} \pi)) \quad (6)$$

式中： T_{cur} 为当前迭代次数， $T_{\max}$ 为总迭代次数。该策略有利于模型在训练后期更稳定地收敛至局部最优。

2 AI 辅助故障修复系统

2.1 总体架构

本文设计了一套 AI 辅助故障修复系统。如图 1 所示，系统采用分层架构，包含数据采集层、分析诊断层和决策执行层三个层级。

(1) 数据采集层：该层负责多源数据的收集和初步处理，通过 eBPF 技术实现零侵入的内核级指标采集，资源占用比传统方法降低了 60%。且可以从设备日志、性能监控平台和用户端应用获取数据。在设备端完成初步的数据压缩和特征提取，只将关键特征上传至中心分析节点，这样可以减少网络传输开销。

(2) 分析诊断层：该层集成改进残差网络诊断模型，对采集到的数据进行实时分析，识别设备的故障类型并定位故障发生的原因。同时引入多模态大语言模型的认知能力，实现自然语言交互式诊断。

(3) 决策执行层：该层是根据诊断结果生成个性化的修复方案，通过自动化脚本或指导用户操作完成故障修复。对于复杂故障，支持远程介入和修复。

2.2 基于强化学习的修复决策

故障修复决策模块利用深度强化学习 (DRL) 算法，把故障修复过程形式化为一个五元组 (S, A, P, R, g) 马尔可夫决策过程。

状态空间： $S = \{\text{设备状态}\} \times \{\text{故障类型}\} \times \{\text{环境上下文}\} \times \{\text{历史操作}\}$ ；

动作空间： $A = \{a_1, a_2, \dots, a_n\}$ ，其中 $a_i \in \{\text{重启服务, 回滚配置, 资源调整}\dots\}$ ；

状态转移概率： $P(s_{t+1}|s_t, a_t)$ 是基于历史数据学习的动态模型。

智能体依据当前设备的状态选择修复动作，并在执行后获得环境反馈，根据不断试错优化修复策略。奖励函数补充故障严重程度权重和动作有效性惩罚，使得奖励与运维的实际需求深度匹配：

$$R(S_t, A_t, S_{t+1}) = \quad (7)$$

$$\alpha \cdot w_f \cdot P_{\text{修复成功}} - \beta \cdot \frac{T_{\text{修复时间}}}{T_{\max}} - \eta \cdot \frac{C_{\text{修复成本}}}{C_{\max}} - \delta \cdot I(A_t \text{无效})$$

式中： $\alpha=0.62$ 、 $\beta=0.28$ 、 $\eta=0.10$ 、 $\delta=0.05$ ，通过网格搜索和验证集优化确定。

w_f 为故障严重程度权重，分为 1~5 级：完全故障 (5 级， $w_f=2.0$)、严重性能下降 (4 级， $w_f=1.6$)、中度故障 (3 级， $w_f=1.2$)、轻微故障 (2 级， $w_f=0.8$)、潜在故障 (1 级， $w_f=0.4$)。 $T_{\max}=30 \text{ min}$ ，传统人工修复最长耗时； $C_{\max}=1$ ，归一化后的最大资源消耗，如重启设备 $C=0.8$ 、清理缓存

$C = 0.1$ 。 $I(A_i \text{ 无效})$ 为动作无效指示函数，动作执行后状态无改善则为 1，否则为 0，避免智能体重复执行无效动作。 $P_{\text{修复成功}}$ 是指示函数，修复成功时为 1，否则为 0。

折扣因子 g ：设置为 0.95，平衡即时奖励（当前修复成功）与长期奖励（避免后续故障复发），鼓励智能体选择修复彻底、复发率低的动作。

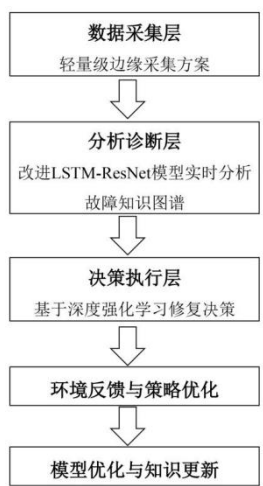


图 1 系统总体架构图

3 实验与案例分析

3.1 实验设计

为了证明本文方法的有效性，采集来自某运营商真实终端设备的运维数据，覆盖路由器、交换机、智能家居设备等常见的终端设备。数据集包含 10 000 个终端设备样本，包括正常、性能下降、完全故障等 7 种状态，故障类型有硬件故障、配置错误、网络连接异常等。数据集按 7:2:1 的比例划分为训练集、验证集和测试集。

设置对照实验，对比方法有：传统机器学习方法：支持向量机（SVM）、随机森林（RF）；深度学习方法：一维卷积神经网络（1D-CNN）、传统残差网络（ResNet）、多重注意力卷积网络（MA-CNN）。对比这些方法在诊断准确率、精确率、召回率、 F_1 值和修复成功率等指标方面的差异。

3.2 实验结果分析

3.2.1 模型诊断性能对比

从表 1 可以看出，本文提出的改进残差网络模型在准确率、精确率、召回率和 F_1 值四项核心指标上均显著优于所有对比模型，实现了最优的综合诊断性能。本文方法的准确率达到 98.2%，相较于性能次优的 MA-CNN 模型（94.1%）提升了 4.1%，与传统 ResNet 相比提升了 5.6%，这证明了引入 LSTM 时序分析单元和多重注意力机制的有效性，充分证明了模型整体架构设计的先进性和有效性，为后续的 AI 辅助故障修复提供了坚实可靠的决策基础。

表 1 不同方法在终端设备故障诊断任务上的性能对比

单位：%				
诊断方法	准确率	精确率	召回率	F_1 值
SVM	84.3	83.2	81.9	82.5
RF	86.7	85.1	84.6	84.8
1D-CNN	90.2	89.1	88.5	88.8
ResNet	92.6	91.5	90.3	90.9
MA-CNN	94.1	93.2	92.8	93.0
本文方法	98.2	96.7	95.3	96.0

3.2.2 消融实验

为评估本模型中各组件的贡献，设计了消融实验，结果如表 2 所示。

表 2 消融实验结果

单位：%			
模型配置	准确率	F_1 值	准确率下降幅度
完整模型	98.2	96.0	—
无注意力机制	95.0	92.8	3.2
无 LSTM 单元	96.1	94.2	2.1
无 Dropout	96.7	94.9	1.5
基础 ResNet	92.6	90.9	5.6

由表 2 可知，删除任一单元都会导致模型性能下降，其中删除注意力机制使准确率下降了 3.2%，表明注意力机制的贡献最多。LSTM 单元对时序特征的建模使准确率提升了 2.1%，Dropout 机制通过防止过拟合进一步提升了 1.5%，这说明本文提出的各个单元都对最终性能有一定实质性的贡献。

3.2.3 修复效果评估

表 3 三种修复方法的实验结果

模式	修复成功率 /%	平均修复耗时 /min
传统人工运维	72.5	15.3
基于规则的修复	83.7	8.6
本文 AI 辅助修复	91.3	2.1

从表 3 可知，本文的 AI 辅助修复方法修复成功率达到了 91.3%，相比基于规则的修复方法提高 7.6%，较传统人工修复方法提高 18.8%，这表明本文智能修复系统可以有效处理较为复杂的故障难题。此外，本文方法的平均修复耗时仅为 2.1 min，比传统人工修复时间缩短了 86%。

3.3 应用案例

为验证本系统的实用性，将该系统运用在一个工厂的物联网终端上，一共 300 台终端。运行 3 个月后，发现故障诊断准确率达到 97.8%，比此前提高了 23.4%；故障修复成功率达到 90.5%，比此前提高了 18.2%。人工运维的工作量减少了 60%，由故障引发的生产中断时间也缩短了 80%，生产效率提升了 12%。

这表明，本系统在实际工厂中可以良好运行，可提高运维效率，具有实际应用价值。

边缘计算环境下网络切片的动态调度策略研究

宋 莉¹

SONG Li

摘 要

针对边缘计算场景下,资源异构与业务动态变化引发的网络切片调度不稳定问题,文章构建了一套面向多类 SLA 业务的动态调度策略框架,提出一种结合优先级驱动与深度强化学习的多策略调度算法。通过构建资源约束矩阵、设计服务优先级评分函数与多目标优化模型,并引入轻量级分布式控制逻辑,仿真实验验证:所提方法能够有效削减平均响应时延与节点迁移次数,提升资源利用率和服务可用性,同时具备优异的部署适配能力与调度鲁棒性。

关键词

边缘计算;网络切片;动态调度;服务质量(QoS);深度强化学习

doi: 10.3969/j.issn.1672-9528.2026.06.023

0 引言

随着边缘计算在智能制造、智慧园区等场景的广泛部署,网络切片面临跨节点资源动态调度与多业务差异化保障的双重挑战。传统静态映射或贪婪调度策略在复杂边缘环境中难以兼顾低延迟与资源高效利用。本文聚焦异构边缘架构下的

多租户业务,探索融合规则驱动与学习增强策略的调度机制,构建从建模、优化到算法实现的完整路径。

1 网络切片与边缘计算协同需求分析

1.1 网络切片资源模型与边缘系统结构

网络切片在 RAN-Core-Transport 架构中构建虚拟化资源隔离环境,控制面部署 AMF、SMF 等功能实体,数据面基

1. 兰州文理学院 甘肃兰州 730000

4 结论与展望

本文提出了一种基于改进残差网络的终端设备状态智能诊断与 AI 辅助的故障修复方法,在残差网络的基础上引入了 LSTM 和注意力机制,增强对时序数据的特征提取能力。

在改进残差网络基础上设计的 AI 辅助故障修复系统,可以做到发现故障、分析原因到自动修复,能大幅提升运维效率。

实验结果显示,该方法诊断准确率达到 98.2%,修复成功率达到了 91.3%,均优于传统的人工运维。应用案例进一步证明了本文方法的实用性。未来的研究可以进一步提升模型的通用性,让智能运维技术的应用场景更广泛。

参考文献:

- [1] 中华人民共和国工业和信息化部. 2025 年前 7 个月通信业经济运行情况[R]. 北京:工业和信息化部,2025.
- [2] 刘文婷,徐昶,朱茂中,等. 基于 AI Agent 技术实现宽带故障的智能诊断应用研究[J]. 电信工程技术与标准化, 2025, 38(5): 30-36.
- [3] 谢道祥. 智能诊断与故障预判技术在广播电视无线发射台

站的应用研究[J]. 广播与电视技术,2025,52(7):80-83.

- [4] Wu Keyu, Yu Qianjin, Mei Manlin, et al. TN-AutoRCA: benchmark construction and agentic framework for self-improving alarm-based root cause analysis in telecommunication networks[PP/OL].(2025-07-28)[2026-05-23].<https://doi.org/10.48550/arXiv.2507.18190>.
- [5] 谭瑛,黄彬,关俊波,等. 基于 5G 消息和深度学习的宽带故障排查方案研究[J]. 广西通信技术,2022(3):26-29.
- [6] 唐嘉辉,成小乐,孙骞,等. 基于注意力残差神经网络的滚动轴承复合故障诊断方法[J]. 机械设计,2025,42(11):99-105.
- [7] He Kaiming, Zhang Xiangyu, Ren Shaoqing, et al. Deep Residual Learning for Image Recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2016:770-778.

【作者简介】

张子樵(1979—),男,湖北武汉人,本科,高级工程师、高级专家,研究方向:互联网技术、终端与业务、人工智能。

(收稿日期:2025-11-21 修回日期:2026-06-02)