

基于多标签集成学习的螺旋 CT 机故障诊断研究

闫小如^{1,2}

(1. 连云港市第一人民医院 临床医学工程部, 江苏 连云港 222000; 2. 中南大学 计算机学院, 长沙 410083)

摘要: 医学应用领域计算机 X 线断层摄影螺旋机由于复杂的结构和较高的集成度在实际故障定位和检测中具有极高的难度; 为了解决这个问题, 研究对螺旋 CT 机故障定位与检测问题进行了分析, 提出一种多标签集成学习方法; 该方法采用了折半查找算法获取螺旋 CT 机的故障数据, 同时有效结合现有的卷积神经网络和循环神经网络的文本表征网络, 通过自适应标签关系增强方法找出标签间的依赖关系, 并利用加权约简标签集的不平衡学习能有效杜绝模型可扩展性低和模型泛化性弱等问题; 经损失值、准确度、运行时间、精准率、灵敏度 5 个指标的实例测试结果表明, 研究所给出的方法均相对于其他 3 种较为创新的多标签集成学习方法更具优势, 且提升数值均超过 2%, 训练集的各个指标数据均比测试集相应数值更高; 训练集和测试集中空时网络聚类约简的多标签集成学习方法的精准率分别为 93.12% 和 87.26%, 召回率分别为 86.35% 和 84.25%; 该方法能精准快速查找螺旋 CT 机的故障类型和故障部位, 极大程度降低维修成本和延长设备的使用年限。

关键词: 螺旋 CT 机; 多标签集成学习; 故障检测; 折半查找算法; 空时网络聚类约简

Research on Fault Diagnosis of Spiral CT Machine Based on Multi Label Ensemble Learning

YAN Xiaoru^{1,2}

(1. Department of Clinical Medical Engineering, Lianyungang First People's Hospital, Lianyungang 222000, China;

2. College of Computer, Central South University, Changsha 410083, China)

Abstract: Computed tomography camera (CT) spiral machines in the field of medical applications face extremely high difficulties in actual fault localization and detection due to their complex structure and high integration. To address this issue, an analysis was conducted on the fault localization and detection of CT spiral machines, and a multi label ensemble learning method was proposed. This method uses a half search algorithm to obtain fault data of CT spiral machines, while effectively combining existing convolutional neural networks and recurrent neural networks for text representation. Through an adaptive label relationship enhancement method, the dependency relationships between the labels are identified, and the imbalanced learning of weighted reduction label sets can effectively eliminate problems such as low model scalability and weak model generalization. The test results of five indicators, including loss value, accuracy, running time, accuracy, and sensitivity, show that the proposed methods have more advantages over the other three innovative multi label ensemble learning methods, and the improvement values all exceed 2%. The various indicators of the training set are higher than those of the test set. The accuracy of the multi label ensemble learning method for spatiotemporal network clustering reduction in the training set and test set is 93.12% and 87.26%, respectively, with the recall rates of 86.35% and 84.25%. This method can accurately and quickly identify the types and locations of faults in CT spiral machines, greatly reducing maintenance costs and extending the service life of equipment.

Keywords: CT spiral machine; multi label ensemble learning; fault detection; half search algorithm; space time network clustering reduction

0 引言

中国医疗体制改革下医疗健康事业得到了蓬勃的发展, 社会大众也对医疗仪器设备提出了更高的要求。计算机 X 线断层摄影 (CT, computed tomography camera) 螺旋机作为医院临床诊断疾病的重要设备, 其故障定位和检测与医疗系统的服务质量紧密关联。研究和分析螺旋 CT 机的故障定位和检测能力, 不仅有助于降低设备出现故障的概率, 同时也能明显改善医院服务病人的能力, 还能改善螺旋 CT 机的工作

性能和延长仪器的使用周期, 进而节省医院在设备维修运行方面的开销。螺旋 CT 机具有较高的集成程度和复杂的结构, 这增加了故障定位和检测工作的开展难度^[1-3]。

对于螺旋 CT 机故障诊断研究领域而言, 王剑等研究学者为了解决螺旋 CT 机故障精准定位方面存在的效率低、耗时长等问题, 提出通过折半查找算法确定故障类型和相关部件。该方法将从一系列关联元素中找到故障数据。仿真分析结果表明, 该方法能精准查找到螺旋 CT 机的故障, 同时大幅度降低维修成本^[4]。国内外其他学者也对螺旋 CT 机

收稿日期: 2024-04-11; 修回日期: 2024-05-24。

作者简介: 闫小如(1974-), 女, 大学本科, 高级工程师。

引用格式: 闫小如. 基于多标签集成学习的螺旋 CT 机故障诊断研究[J]. 计算机测量与控制, 2024, 32(11): 48-55.

故障定位与检测进行了多角度和全方位的分析，但得到的螺旋 CT 机故障定位与检测结果仍然无法取得满意的结果，这也是困扰医学领域医疗设备专家的一大难题^[5-6]。Zhang J 等人针对螺旋 CT 机的旋转故障诊断，利用神经网络的非线性映射能力，提出了一种融合不平衡学习方法。同时利用无量纲参数化结合时频变换方法提取数据特征，构建不同的证据源。结果显示，该方法的故障诊断率有待提升^[7]。Liu Z H 等研究人员为了实现螺旋 CT 机的有效故障诊断，提出了一种基于堆栈自动编码器的部分对抗域自适应模型来解决螺旋 CT 机的故障诊断问题。大量实验结果表明，在处理场景一部分域自适应问题时，研究所提方法的诊断性能较优，但仍然存在一定的进步空间^[8]。多标签集成学习作为一种经典的学习方法，其可被用于实际案例中包含多种标签数据的学习问题，且数据具有不均衡性、高低维度不同等特点，在不同领域数据分类和查找中具有突出的优势。传统的多标签集成学习通常只关注到较少的标签量，但在大量标签数据中该方法具有较大的局限性。极端多标签集成学习的研究得到标签分类学者们的重点关注。但当前的极端多标签集成学习方法的泛化能力低，且算法的可扩展性弱。空时网络聚类约简的多标签集成学习方法凭借着的数据处理没有约束、泛化能力强、可扩展性强等优势在多种类型数据标签分类中得到了应用。

为了降低螺旋 CT 机设备维修人员的工作强度，以及准确分析到设备的故障问题，研究创新性地提出通过空时网络聚类约简的多标签集成学习方法，该方法有效结合现有的卷积神经网络和循环神经网络的文本表征网络，通过自适应标签关系增强方法找出标签间的依赖关系，并利用加权约简标签集的不平衡学习能有效杜绝模型可扩展性低和模型泛化性弱等问题，以此来实现螺旋 CT 机的定位和检测。与以往的研究相比，该方法采用的多标签集成学习方法有助于提升故障诊断精度及效率，有望为后续医疗设备的故障定位和检测提供新的研究思路。

1 螺旋 CT 机故障诊断技术

1.1 折半查找算法的螺旋 CT 机故障查找

在医院医疗疾病诊断过程中，螺旋 CT 机作为一种重要检测设备通过 X 射线扫描人体相应部位而确诊疾病类型，其具备扫描成像精度高、扫描技术成熟、扫描时间短等优点。

螺旋 CT 机作为一种高精密仪器，可在短时间内对病人的身体进行准确检查，进而降低患者自身由于呼吸运动造成的图像伪影。当出现故障时，则螺旋 CT 机的使用寿命会降低，同时也会间接对患者的生命产生威胁^[9-11]。计算机故障主要表现为鼠标硬盘等硬件设备的损坏，以及数据存储信息的完整性和隐匿性。除此之外，滑环和碳刷故障也是常见的类型。螺旋 CT 机故障定位和检测的依据为查找数据元素间存在次序性关联，在此基础上才可以借助多标签集成学习实现精准故障诊断。螺旋 CT 机故障查找通过折半查找算法来实现，同时螺旋 CT 机故障定位和检测利用空时网络聚类约简的多标签集成学习方法来达成。该方法利用折半查找算法获取螺旋 CT 机的故障数据，同时有效结合现有的卷积神经网络和循环神经网络的文本表征网络，通过自适应标签关系增强方法找出标签间的依赖关系，并利用加权约简标签集的不平衡学习能有效杜绝模型可扩展性低和模型泛化性弱等问题。折半查找算法在应用于螺旋 CT 机故障检测之前，需通过前期处理技术对非关联数据进行剔除，这样做的目的是降低标签数据在查找过程中所花费的大量时间，进而提升数据标签在查找过程中的运行效率。非必要元素剔除方式为遍历所有元素，将所有元素按照顺序排列，找到每个元素之间的差异和不同，同时将其与螺旋 CT 机不同部件的故障数据进行分析，找出与故障数据不匹配的数据并进行移除。

依据螺旋 CT 机的故障性质对故障类别进行划分，构建相应的故障类别序列。研究所采用的螺旋 CT 机为佳能医疗系统（中国）有限公司的新一代 640 层动态容积 CT，型号为 Aquilion ONE 640。表 1 是指机架、扫描器、高压电源、量子探测器、球管 5 个部件的指标参数。机架包括采集图像层数、采集单元数、倾斜角 3 个指标，扫描器包括重建矩阵、平均层厚、扫描均时、扫描模式、采样技术 5 个指标，高压电源包括电流、电压、功率 3 个指标，量子探测器包括通信采集能力、密度分辨率、采集角度、覆盖角度。球管包括散热率、阳极热容量、等效热容量 3 个指标。故障类型可分为 3 种类型，分别为软件故障、硬件故障、扫描参数设置问题。软件故障的主要原因为软件程序损坏和软件间的通讯故障，将其划分为类别 A 和类别 B。硬件故障包括焊脚不稳、元件接触不灵和元件损坏，将其划分为类

表 1 新一代 640 层动态容积 CT 五个部件的指标参数

| 部件部分 | 指标 | 数值大小 | 部件部分 | 指标 | 数值大小 |
|------|----------|-------------|-------|---------|-----------------|
| 机架 | 采集图像层数 | 320 层 | 高压电源 | 电压 | 100 kV 和 220 kV |
| | 采集单元数 | 640 排×896 排 | | 功率 | 2 200 kW |
| | 倾斜角 | ±40° | 量子探测器 | 通信采集能力 | 640 排×0.5 mm |
| 重建矩阵 | 628×628 | 密度分辨率 | | 2 mm±3% | |
| 平均层厚 | 0.973 mm | 采集角度 | | 360° | |
| 扫描均时 | 0.68 s | 覆盖角度 | | 16 cm | |
| 扫描器 | 扫描模式 | 320 层超螺旋扫描 | 球管 | 散热率 | 1 273 kHU/min |
| | 采样技术 | 双源 CT | | 阳极热容量 | 5 MHU |
| 高压电源 | 电流 | 50~600 mA | | 等效热容量 | 30 MHU |

别 C 和类别 D。扫描参数设置问题包括遭遇系统攻击或没有根据使用范围调整对应的参数数值，通过类别 E 表示。设置螺旋 CT 机的故障数量为 m 个，故障点 l 和故障点 o 的坐标分别为 (y_l, z_l) 和 (y_o, z_o) ，两个故障点间的距离通过欧几里得计算公式得到，用式 (1) 指代：

$$d_{lo} = \sqrt{\sum_{i=1}^2 (y_i - z_i)^2} \quad (1)$$

螺旋 CT 机故障定位问题可通过数学模型式 (2) 表示：

$$f = \min_{(y_i, z_i), i=1}^m (\sqrt{(y_l - z_l)^2 + (y_o - z_o)^2} - d_{lo}) \quad (2)$$

式 (2) 中， $l, o = 1, 2, \dots, m$ 。数学模型的表达式包括大量计算过程，这导致计算所耗费的时间和人力成本极大提升。当计算量较大时，整体运行过程中将会产生大量的随机误差，进而影响结果的精准性。折半查找算法是一种在有序数组中查找某一特定元素的搜索算法。搜索过程从数组的中间元素开始，如果中间元素正好是要查找的元素，则搜索过程结束。图 1 是指折半查找算法的流程图。该图中仅展示了 4 次查找过程，每次查找的数据均为 11 个。折半查找算法在第一次查找过程中，最小值和最大值分别为序号为 1 和 11 的数据，中间值的数据序号为 6。在第二次查找过程中，最小值和最大值分别为序号为 7 和 11 的相应数据，中间值的数据序号为 9。在第一次和第二次查找过程中，最大值均为 11，而第二次查找的最小值序号从 1 增加到 7。

折半查找算法的实施过程如下，数列元素为螺旋 CT 机的所有等待查找的项目，设置等待查找且可能出现故障的项目的数量为 u ，序列依据递增顺序可表示为 $U = (a_1, a_2, \dots, a_u)$ 。数列元素需满足式 (3)：

$$a_1 < a_2 < \dots < a_u \quad (3)$$

序数列的中间位置元素 a_p 通过式 (4) 得到：

$$a_p = \frac{a_1 + a_u}{2} \quad (4)$$

将中间位置元素 a_p 与待查找元素 a_q 的取值进行对比，根据对比结果和式 (5) 创建查找序列 U' ：

$$\begin{cases} a_p < a_q, U' = (a_1, a_2, \dots, a_p) \\ a_p > a_q, U' = (a_p, a_{p+1}, \dots, a_u) \end{cases} \quad (5)$$

式 (5) 中，当两个值的取值相同时，则说明螺旋 CT 机故障目标数据已被找到。在查找螺旋 CT 机故障类型的过程中，该算法可通过 1/2 倍的幅度不断降低查找范围，不仅

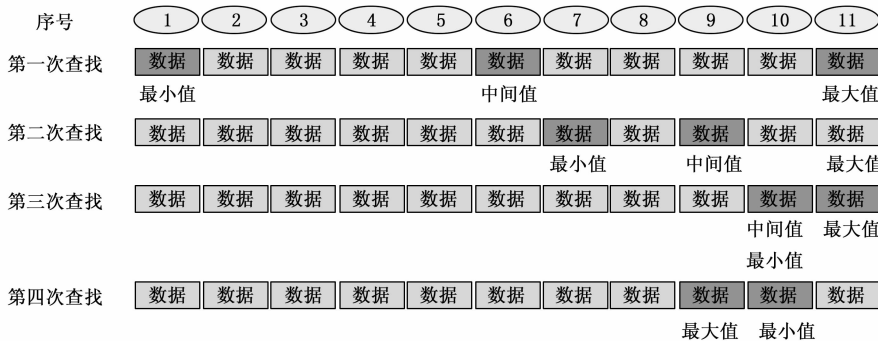


图 1 折半查找算法的流程图

能提升查找的精度，同时还能在较短时间内发现故障的位置，进而加快检修工作的进程。当序列中存在相同数据时，找到中间索引值，不能立即返回。同时向中间索引值的左边扫描，将所有与查找值相同的元素的下标加入集合。向中间索引值的右边扫描，将所有与查找值相同的元素的下标加入集合。

1.2 空时网络聚类约简的多标签集成学习方法构建

1.2.1 空时网络注意力集成表征

在利用折半查找算法获取螺旋 CT 机的故障数据后，包括软件故障、硬件故障、扫描参数设置问题，研究构建空时网络聚类约简的多标签集成学习方法，进而完成螺旋 CT 机的定位和检测。空时网络聚类约简的多标签集成学习方法的流程图主要包括 3 个阶段。图 2 为空时网络聚类约简的多标签集成学习方法。第一个阶段为空时网络注意力集成表征，该阶段集成了多标签文本中标签、短语、中词三者间的交互注意力，分析三者之间的依赖关系，极大程度提升极端多标签文本表征能力。在空时网络注意力集成表征阶段，研究首先获取原始极端多标签数据，同时针对不同螺旋 CT 机疾病诊断中获取的极端多标签数据进行学习。然后对短语级、词级进行表征。设置给定文档为 $x_i \in \mathbf{R}^{d \times n}$ ， $\mathbf{R}^{d \times n}$ 是指横和列分别为 d 和 n 的 \mathbf{R} 矩阵，短语级表征过程中可使用卷积神经网络卷积核 $W_i \in \mathbf{R}^{w \times d}$ 和偏差项 b_i 完成短语级不同权重的学习，从而进行表征，其中， $w \in [0, n]$ 。通过向量 c_i 表示词联合对特征进行表征，则特征可用式 (6) 指代：

$$p_i = \delta(\text{Conv1D}(W_i, c_i) + b_i) \quad (6)$$

式中， p_i 是指多标签文本中标签、短语、中词的特征表示， δ 是指激活函数。词级表征通过循环神经网络进行表征，其可以对双向的词级信息进行学习，用 $H = [H^f; H^b]$ 指代。其次是表征空间语义信息。利用混合的注意力机制得到空间语义信息表征，模块包括卷积交互注意力模块和卷积多头自注意力模块，可用 $C = \text{Concat}(S, I_1)$ 指代。该类型表征需分析不同短语之间的关系，同时还需分析标签和短语的关系。多头注意力模块的具体实现过程如下。首先是算单头注意力。根据点积的注意力，query、key、value 三个嵌入矩阵分别为 $Q \in \mathbf{R}^{2r \times 1}$ 、 $K \in \mathbf{R}^{2r \times 1}$ 、 $V \in \mathbf{R}^{2r \times 1}$ ，注意力输出矩阵 $\text{Attention}(Q, K, V)$ = 可用式 (7) 指代：

$$\text{Attention}(Q, K, V) = \text{soft}\left(\frac{QK}{\sqrt{2r}}\right)V \quad (7)$$

其次是计算多头注意力。根据单头注意力的计算结果，多头注意力的计算结果可获得。最后是计算卷积的多头自注意力 Multi-head Attention (Q, K, V)，可用式 (8) 指代：

$$\text{Muti-head Attention}(Q, K, V) = \text{Concat}(\text{head}_1, \text{head}_2, \dots, \text{head}_h) \quad (8)$$

学习多标签文本过程中，鉴于每个文档可能包含多个标签，研究使用多标

签注意力机制分析不同的标签关系。根据上一步矩阵输出多标签注意力。卷积交互注意力的具体实现过程包括嵌入标签图和计算卷积的交互注意力。通过图嵌入向量获得标签共现图向量，以此来探究标签的结构信息。将每个标签看成一个节点，假设一个文档中出现两个随意的标签，则这两个标签间有个边相连。在随机游走策略的支持下，图嵌入技术能获得标签依赖关系的高阶层次，其中，每个标签都可用 $2r$ 维向量表示，第 i 个标签可用 $L_i \in R^{2r}$ 指代，整个标签可用 $L \in R^{k \times 2r}$ 指代。而对于计算卷积的交互注意力，可表示为 $I_1 = V \times \text{softmax}(LK)^T$ 。再者，表征时序语义信息。表征过程采用混合的注意力机制，涉及的模块包括循环的交互注意力模块和循环的自注意力模块，两者可表示为 $R = \text{Concat}(B, I_2)$ 。对于前者而言，为了更好确定上下文档词级的依赖关系，研究通过加权自注意力表示文档的不同属性，这不仅可学习到文档数据中的各类别信息，同时还能获得较长时间内的数据依赖性。循环的自注意力模块可用 $B \in R^{2r \times k}$ 指代。循环的交互自注意力模块和卷积的交互注意力类似，该模块可通过交互注意力获取细粒度的词级信号，据此获得标签和单词间的匹配分数，循环的交互自注意力模块可用 $I_2 = H \times \text{softmax}\left(\begin{bmatrix} L \\ L \end{bmatrix} \begin{bmatrix} H^f \\ H^b \end{bmatrix}\right)^T$ 指代。最后，集成空时网络注意力表征。根据 $R = \text{Concat}(B, I_2)$ 和 $C = \text{Concat}(S, I_1)$ ，研究提出了自适应加权集成策略。其一，对两者进行 I_2 标准化处理。其二，利用全连接层和多层感知层转换获得相应的权重，可用 $\alpha \in R^{k \times 1}$ 和 $\beta \in R^{k \times 1}$ 指代。其三，对权重进行标准化处理得到空时网络注意力的集成表征。对比于现阶段极端多标签文本学习方法，研究所提出的方法可以有效结合现有的卷积神经网络和循环神经网络的文本表征网络。同时也兼顾了多标签文本中标签、短语、中词相互间的交互注意力，并且也考虑了标签、短语、词之间的依赖性^[12-14]。

1.2.2 自适应增强和聚类约简学习

第二个阶段为标签关系的自适应增强和聚类约简学习，提出了标签关系的自适应增强，可以很大概率挖掘数据标签间的依赖关系，同时提升算法的泛化性能。通过聚类约简学习理论，能通过现有的模型训练不同的量级标签。该阶段包括聚类标签树、强化标签关系、聚类约简学习 3 个部分。对于聚类标签树，内积求和标签文本特征和包含标签的稀疏文本特征，同时对每个标签的特征通过标准化的形式表示。接着利用 K-Means 聚类算法对数据标签进行聚类，通过持续的迭代过程使其符合以下条件。原始 K-means 算法最开始随机选取数据集中聚类中心，而研究使用的聚类算法选取聚类中心，假设已经选取了初始聚类中心，则在选取下一个聚类中心时距离当前聚类中心越远的点会有更高的概率被选为下一个聚类中心。当属于某个类别的样本数过少时把这个类别去除，当属于某个类别的样本数过多、分散程度较大时把这个类别分为两个子类别。设置每个簇的最大标签量，依据要求将其划分到簇中，簇

的数量为 S 个，簇内标签的数量需超过最大值的 $1/2$ ，或者低于最大标签量。通过空时网络注意力集成表征结果，即可利用一个全连接层获得 S 维度的向量 P 。对于强化标签关系，在原本的预测向量 P 基础上，通过标签关系的传达添加一个 bottleneck 层，该层的目的可实现标签数据的自适应增强 \hat{P} 。聚类约简学习而言，通过标签数聚类关系获得不同簇的标签索引识别，和标签数据的自适应增强 \hat{P} 构建了一个数据的学习机制 L_s ，可用式 (9) 指代：

$$L_s = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^S [y_{ij}^u \log(\hat{P}_{ij}) + (1 - y_{ij}^u) \log(1 - \hat{P}_{ij})] \quad (9)$$

1.2.3 加权的约简标签集成学习

第三个阶段为加权的约简标签集成学习，该机制解决了标签稀疏和数据不平衡导致的问题。第三个阶段包括嵌入约简标签集和不平衡数据的加权学习。加权约简标签集的不平衡学习能有效杜绝模型可扩展性低和模型泛化性弱等问题，该方法能提升极端多标签文本学习的精度^[15-17]。不平衡学习机制能在一定程度杜绝空时网络注意力集成表征过程中造成的数据缺失，同时通过学习降低不平衡度对结果的影响^[18-19]。

对于前者而言，通过簇内标签得到标签的约简集，根据第一阶段得到的空时网络注意力集成表征，能进一步得到约简标签集的嵌入向量 Q 。根据前者获得每个标签集对应的标签 $y^u \in [0, 1]^u$ 。为了克服多标签中的数据存在的失衡问题，研究通过加权方法对数据进行学习，结果式 (10) 指代。加权学习方法是指在机器学习中，对不同样本或不同任务赋予不同的权重，以便更好地适应模型。这种权重可以用于调整模型在训练过程中对于不同样本的关注度，从而影响模型的学习行为。该方法在解决数据在线学习、领域适应、噪声数据、类别不平衡等问题中具有较强的实用价值。研究对损失函数进行加权，即对不同样本的损失赋予不同的权重。在模型更新的过程中，高权重的样本模型需更加关注，从而更有效地进行学习。

$$L_Q = -\frac{1}{N} \sum_{i=1}^N \sum_{k=1}^U \begin{cases} (1 - Q_k)^{\gamma^+} \log(Q_k), y_k^u = 1 \\ Q_k^{\gamma^-} \log(1 - Q_k), y_k^u = 0 \end{cases} \quad (10)$$

式 (10) 中，多标签不平衡数据的加权处理结果为 L_Q ，约简标签集的真实标签为 y_k^u ，嵌入数据为 Q_k ，正负标签样本的相应权重为 γ^+ 和 γ^- 。

2 螺旋 CT 机故障诊断效果分析

为了评估空时网络聚类约简的多标签集成学习方法在螺旋 CT 机故障诊断效果，研究设置对比方法为张量网络的多标签学习方法、双拉普拉斯正则化与因果推断的多标签学习、多标签分类的性质识别方法^[20-22]。为了便于图中方法的展示，研究采用方法 1—方法 3 指代。实验测试环境如下，操作系统为 Ubuntu20.04 操作系统，深度学习框架为 Pytorch1.8，编程语言为 Python3.7，系统内存为 64 GB，硬件为 Intel (R) Core (TM) i9-10980XE CPU @ 3.00

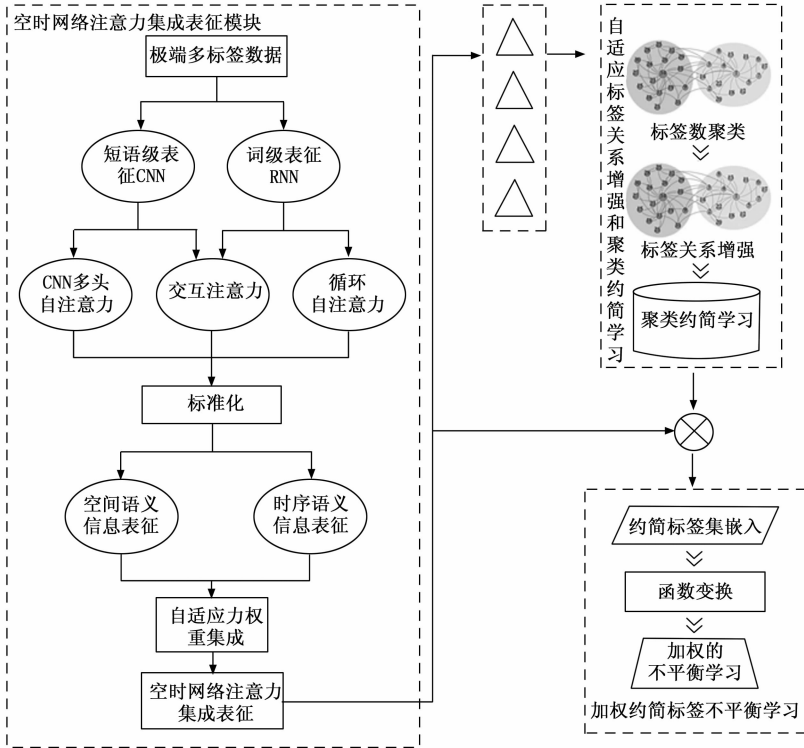


图 2 空时网络聚类约简的多标签集成学习方法

中的消耗问题。相比于不同类型的现有的标签量约简方法，研究给出的方法可降低时间和空间的损耗，也能提升簇内标签的精度。

图 4 (a) 和图 4 (b) 分别为训练集和测试集中不同多标签集成学习方法的运行时间和准确率。对于训练集和测试集，研究所给出的准确率相比于其他多标签集成学习方法的相应数值更高，而运行时间的数值却相比于其他多标签集成学习方法的相应数值更低。随着迭代次数的增加，多标签集成学习方法算法的准确率逐渐增加，而运行时间逐渐降低。训练集和测试集中空时网络聚类约简的多标签集成学习方法的准确率分别为 93.31% 和 90.25%；运行时间分别为 0.39s 和 0.61s。训练集中张量网络的多标签学习方法、双拉普拉斯正则化与因果推断的多标签学习、多标签分类的性质识别方法的准确率分别为 91.02%、90.26%、88.26%，相对于空时网络聚类约简的多标签集成学习方法降低数值为 2.29%、3.05%、5.05%；而测试集中相应多标签集成学习方法的准确率依次为 87.98%、85.26%、83.26%，相对于研究所提出的方法降低数值为 2.27%、4.99%、6.99%。得到这一结果的原因，可能是约简学习方法可通过自适应标签关系增强方法找出标签间的依赖关系，利用不同的量级标签完成模型训练可减少时间和空间中的消耗问题。相比于现有的标签量约简方法，研究给出的方法可降低时间和空间的损耗，也能提升簇内标签的精度。训练集中张量网络的多标签学习方法、双拉普拉斯正则化与因果推断的多标签学习、多标签分类的性质识别方法的运行时间分别为 0.54 s、0.68 s、0.78 s；而测试集中相应多标签集成学习方法的准确率依次为 0.86 s、1.03 s、1.21 s。因此，多标签集成学习方法的准确率和运行时间均较为理想，该方法在螺旋 CT 机故障定位和检测中具有明显的优势。

GHz。评价指标为损失值、准确度、运行时间、精准率、灵敏度。螺旋 CT 机故障数据组成的数据集训练集和测试集的比例为 7 : 3。学习率设置为 0.001，最大训练批次为 10 次，每批次数据规模为 40 个。图 3 是指不同多标签集成学习方法的损失值。随着训练次数的增加，损失值逐渐降低，并逐渐收敛到稳定值。相比于其他多标签集成学习方法，研究所提出的空时网络聚类约简的多标签集成学习方法的收敛损失值更低，数值为 0.023。而张量网络的多标签学习方法、双拉普拉斯正则化与因果推断的多标签学习、多标签分类的性质识别方法的损失值分别为 0.123、0.081、0.046。因此，研究所提出的空时网络聚类约简的多标签集成学习方法在运行过程中产生的误差更低。这些现象产生的原因可能与研究给出的方法采用了约简学习方法有关。该方法通过自适应标签关系增强方法找出标签间的依赖关系，利用不同的量级标签完成模型训练可减少时间和空间

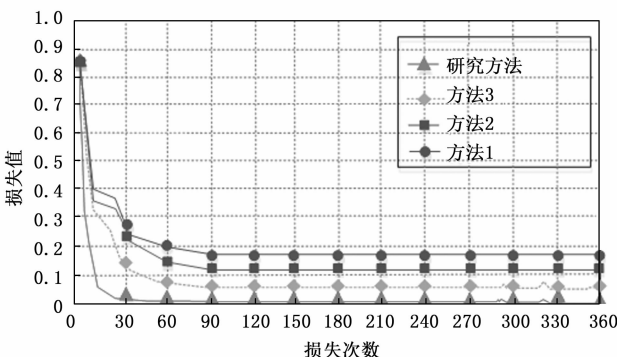


图 3 不同多标签集成学习方法的损失值

对于精准率是衡量方法有效性的关键指标，出现这一结果

图 5 (a) 和图 5 (b) 分别是指训练集和测试集中不同多标签集成学习方法的精准率。对于训练集和测试集，研究所给出的精准率相比较于其他多标签集成学习方法的相应数值更高。随着迭代次数的增加，多标签集成学习方法算法的精准率逐渐增加。训练集和测试集中空时网络聚类约简的多标签集成学习方法的精准率分别为 93.12% 和 87.26%。训练集中张量网络的多标签学习方法、双拉普拉斯正则化与因果推断的多标签学习、多标签分类的性质识别方法的精准率分别为 91.01%、89.16%、87.26%，相对于空时网络聚类约简的多标签集成学习方法降低数值为 2.11%、3.96%、5.86%；而测试集中相应多标签集成学习方法的精准率依次为 85.36%、83.26%、81.20%，相对于研究所提出的方法降低数值为 1.90%、4.00%、6.06%。

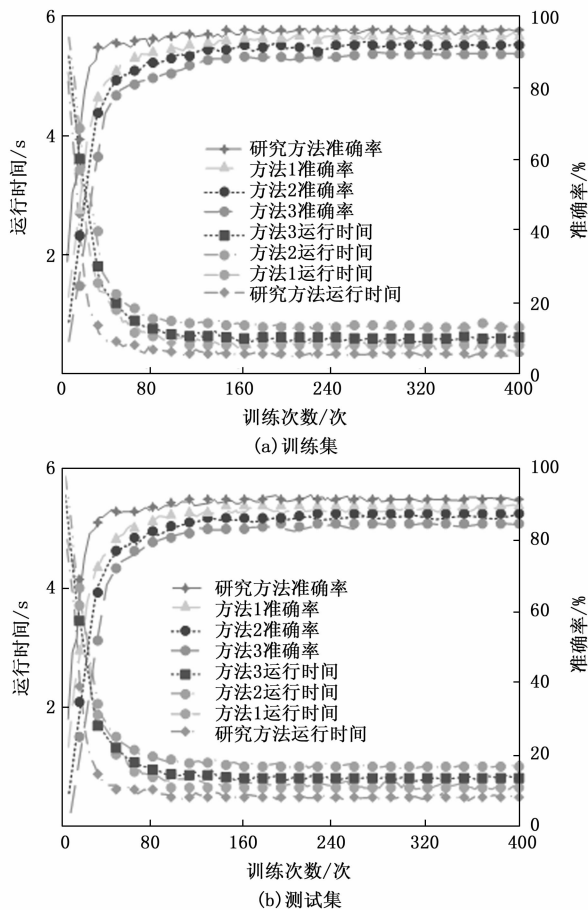


图 4 训练集和测试集中不同多标签集成学习方法的运行时间和准确率

的原因在于加权约简标签集的不平衡学习能有效杜绝模型可扩展性低和模型泛化性弱等问题，该方法能提升极端多标签文本学习的精度。

图 6 (a) 和图 6 (b) 分别是指训练集和测试集中不同多标签集成学习方法的召回率。对于训练集和测试集，研究所给出的召回率相比较于其他多标签集成学习方法的相应数值更高。随着迭代次数的增加，多标签集成学习方法的召回率逐渐增加，这点和准确率和精准率的变化趋势相同。训练集和测试集中空时网络聚类约简的多标签集成学习方法的召回率分别为 86.35% 和 84.25%。训练集中张量网络的多标签学习方法、双拉普拉斯正则化与因果推断的多标签学习、多标签分类的性质识别方法的召回率分别为 83.56%、77.25%、74.06%，相对于空时网络聚类约简的多标签集成学习方法降低数值为 2.79%、9.10%、12.29%；而测试集中相应多标签集成学习方法的召回率依次为 82.36%、76.23%、74.23%，相对于研究所提出的方法降低数值为 1.91%、8.02%、10.02%。出现这一结果的原因在于对比于现阶段极端多标签文本学习方法，研究所提出的方法可以有效结合现有的卷积神经网络和循环神经网络的文本表征网络。同时也兼顾了多标签文本中标签、短

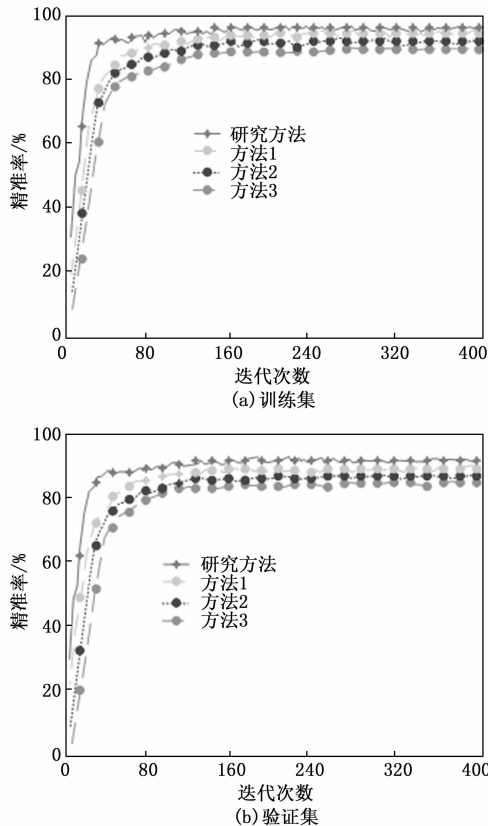


图 5 训练集和测试集中不同多标签集成学习方法的精准率

语、中词相互间的交互注意力，并且也考虑了标签、短语、词之间的依赖性。结合损失值、准确度、运行时间、精准率、灵敏度 5 个指标的测试结果，研究所给出的方法均相对于其他 3 种较为创新的多标签集成学习方法更具优势，且提升数值均超过 2%，且训练集的各个指标数据均比测试集相应数值更高。因此，空时网络聚类约简的多标签集成学习方法对比于较为前沿的集成学习方法更具优势，可应用于螺旋 CT 机故障诊断中，以提升设备仪器的运行稳定性和延长使用寿命。

螺旋 CT 机的常见故障包括 X 线故障、高压逆变器故障、计算机故障以及碳刷与滑环故障。其中，X 线故障参数包括 X 线管寿命、温度、输出功率、管电压、管电流等。高压逆变器故障参数包括逆变器温度、输出电压、输出电流、故障代码等。计算机故障参数包括计算机运行状态、连接状态、响应时间、故障代码等。碳刷与滑环故障参数包括碳刷磨损程度、滑环转动状态以及两者的接触情况等。为验证空时网络聚类约简的多标签集成学习方法对各个故障的诊断效果，研究使用 MATLAB 中的 Simulink 工具建立了螺旋 CT 机的仿真模型。在模拟环境中，实验人员实时采集模拟故障参数数据，包括 X 线管寿命、温度、输出功率、管电压、管电流、逆变器温度、输出电压、输出电流、计算机运行状态、连接状态、响应时间、碳刷磨损程度、滑环转动状态等。将采集到的故障参数数据按照不同故障类型进行标记和处理，形成多标签数据集用于训练和测试。

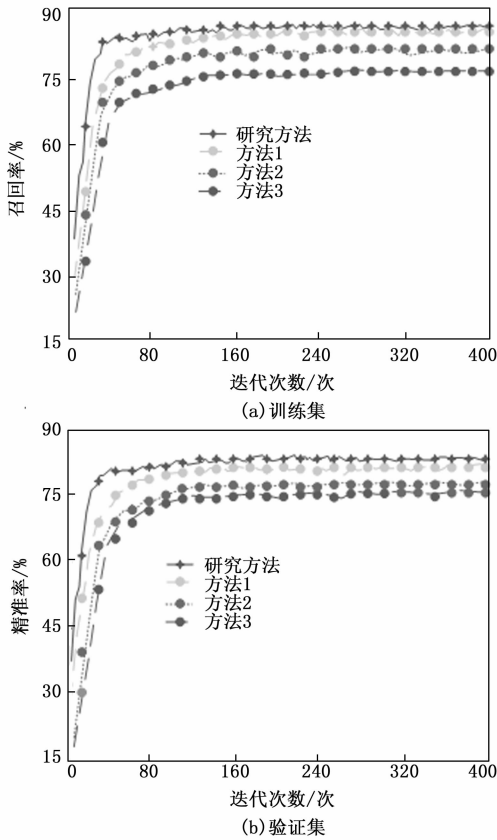


图 6 训练集和测试集中不同多标签集成学习方法的召回率

其中, 训练集占 80%, 测试集为 20%。不同故障类型下的诊断结果如表 2 所示。由表 2 可知, 在模拟结果中。研究所设计的空时网络聚类约简的多标签集成学习方法具有显著的故障性能诊断效果。该方案对 X 线故障、高压逆变器故障、计算机故障以及碳刷与滑环故障的精确率分别高达 92.36%、87.63%、88.96%、86.24%。

表 2 不同故障类型下的诊断结果

| 故障类型 | 精确率/% | 召回率/% | F ₁ 值/% |
|---------|-------|-------|--------------------|
| X 线故障 | 92.36 | 88.96 | 90.13 |
| 高压逆变器故障 | 87.63 | 83.96 | 85.49 |
| 计算机故障 | 88.96 | 90.15 | 89.75 |
| 碳刷与滑环故障 | 86.24 | 87.55 | 86.96 |

3 讨论

为解决螺旋 CT 机故障定位与检测问题, 研究提出一种多标签集成学习方法。其中, 采用折半查找算法能够有效地获取螺旋 CT 机的故障数据, 提供可靠的数据基础。结合现有的卷积神经网络和循环神经网络的文本表征网络能够充分挖掘文本数据中的特征信息, 提高对故障数据的理解和分析能力。在模型的训练中, 自适应标签关系增强方法能够有效地处理标签不平衡导致的模型训练困难和性能下降问题。此外, 加权约简标签集的不平衡学习能够起到提升模型泛化性能的作用。结果显示, 在与张量网络的多标

签学习方法、双拉普拉斯正则化与因果推断的多标签学习、多标签分类的性质识别方法的对比实验中, 空时网络聚类约简的多标签集成学习方法的收敛损失值更低, 数值为 0.023。与其他多标签集成学习方法相比, 空时网络聚类约简的方法的损失值更低, 主要原因在于该方法能够对标签依赖关系的精准建模。相关研究表明, 标签依赖关系的建模有助于提升模型的训练误差^[23]。同时, 空时网络聚类约简的多标签集成学习方法的准确率分别高达 93.31% 和 90.25%; 运行时间仅分别为 0.39s 和 0.61s, 明显优于其余对比算法。原因在于约简学习方法能够找出标签间的依赖关系, 有助于提升效率与空间使用率, 进而提升螺旋 CT 机故障定位和检测性能。这与陈曦等人的研究成果类似^[24]。此外, 空时网络聚类约简的多标签集成学习方法的精准率分别为 93.12% 和 87.26%, 与其他多标签集成学习方法相比, 该方法在精准率方面表现较为优异。表明空时网络聚类约简的多标签集成学习方法具有良好的泛化能力。同时, 通过加权约简标签集的不平衡学习, 能够有效避免模型可扩展性低和模型泛化性弱等问题, 进一步提升精准率。李帆等人的结果显示, 不平衡学习机制能在一定程度杜绝空时网络注意力集成表征过程中造成的数据缺失, 同时通过学习降低不平衡度对结果的影响^[25]。除此之外, 在常见故障的仿真分析中, 研究所设计的空时网络聚类约简的多标签集成学习方法对于 X 线故障、高压逆变器故障、计算机故障以及碳刷与滑环故障的诊断精确率分别高达 92.36%、87.63%、88.96%、86.24%。再次印证了研究所提方法的可行性。

4 结束语

为解决目前螺旋 CT 机在故障定位和查找方面效率过低的问题, 研究提出了一种空时网络聚类约简的多标签集成学习方法, 并将新一代 640 层动态容积 CT 五个部件的故障数据作为基础进行分析。为了评估空时网络聚类约简的多标签集成学习方法在螺旋 CT 机故障诊断效果, 研究将其与张量网络的多标签学习方法、双拉普拉斯正则化与因果推断的多标签学习、多标签分类的性质识别方法进行对比。结果显示研究所提出的空时网络聚类约简的多标签集成学习方法在运行过程中产生的误差更低, 且能够降低时间和空间的损耗, 也能提升簇内标签的精度。同时, 针对螺旋 CT 机的 X 线故障、高压逆变器故障、计算机故障以及碳刷与滑环故障, 研究所设计的空时网络聚类约简的多标签集成学习方法表现出显著的故障性能诊断效果。在查找螺旋 CT 机故障类型的过程中, 研究所给出的方法不仅能提升查找的精度, 同时还能在较短时间内发现故障的位置, 进而加快检修工作的进程。但由于研究条件的限制, 实验仅将新一代 640 层动态容积 CT 五个部件的故障数据为基础进行分析, 下一步研究可将其扩展于其他类型螺旋 CT 机的故障诊断, 以进一步验证空时网络聚类约简的多标签集成学习方法的适用范围, 进而为该技术的推广提供更为科学的

保障。

参考文献:

- [1] 范红意. GE Optima 660 螺旋 CT 扫描床故障分析与检修 [J]. 中国医疗设备, 2019, 34 (10): 177-179.
- [2] 仲建生, 夏云成. 西门子 SOMATOM Spirit CT 故障分析与维修 [J]. 生物医学工程与临床, 2023, 27 (1): 109-113.
- [3] 郎晓华, 王 洲. 西门子 Definition AS 64 排螺旋 CT 的故障分析与维修 [J]. 医疗卫生装备, 2022, 43 (11): 94-97.
- [4] 王 剑, 李彦杰. 基于折半查找算法的多层螺旋 CT 设备故障查找 [J]. 影像科学与光化学, 2023, 41 (1): 130-135.
- [5] 范 卡, 刘 杰. 多层螺旋 CT 滑环装置及准直器的结构原理与故障维修 [J]. 中国医学装备, 2022, 19 (3): 192-194.
- [6] 王裕生. Emotion16 CT 设备探测器数据采集系统工作原理与故障维修 [J]. 中国医学装备, 2021, 18 (5): 202-204.
- [7] ZHANG J, ZHANG Q, HE X, et al. Compound-fault diagnosis of rotating machinery: A fused imbalance learning method [J]. IEEE Transactions on Control Systems Technology, 2020, 29 (4): 1462-1474.
- [8] LIU Z H, LU B L, WEI H L, et al. A stacked auto-encoder based partial adversarial domain adaptation model for intelligent fault diagnosis of rotating machines [J]. IEEE Transactions on Industrial Informatics, 2020, 17 (10): 6798-6809.
- [9] 仲建生, 刘 蕾, 闫 瑾, 等. Brilliance 型 16 排螺旋 CT 的故障分析与维修 [J]. 中国医学装备, 2022, 19 (2): 202-205.
- [10] 李建国, 刘 海, 荀鲁川. 多排螺旋 CT 常见伪影的形成机制及维修 [J]. 中国医学装备, 2020, 17 (10): 219-221.
- [11] 李海洋, 程美英, 高凌浩. 西门子 64 排螺旋 CT 工作原理及其维修案例分析 [J]. 中国 CT 和 MRI 杂志, 2021, 19 (7): 186-188.
- [12] 王 进, 梁 晨, 孙开伟, 等. 基于标签相关性的标签特定特征多标签学习 [J]. 江苏大学学报 (自然科学版), 2023, 44 (5): 554-563.
- [13] 李 强, 陈 潜, 武霁阳, 等. 基于集成学习的高压直流输电系统故障诊断 [J]. 电力系统保护与控制, 2023, 51 (16): 168-178.
- [14] 熊思伟, 刘玉琳. 基于 Borderline-SMOTE 算法与 Stacking 集成学习的前列腺肿瘤风险预测研究 [J]. 现代肿瘤医学, 2023, 31 (16): 3075-3081.
- [15] 马晓宁, 赵东阁. 基于多模型融合的民航领域实体抽取方法 [J]. 计算机工程与设计, 2023, 44 (8): 2516-2522.
- [16] 胡开喜, 李 琳, 吴小华, 等. 一种同伴知识互增强下的序列推荐方法 [J]. 自动化学报, 2023, 49 (7): 1456-1470.
- [17] 董少群, 曾联波, 车小花, 等. 人工智能在致密储层裂缝测井识别中的应用 [J]. 地球科学, 2023, 48 (7): 2443-2461.
- [18] 陆国青, 董少群, 黄立良, 等. 准噶尔盆地玛湖凹陷风城组陆相页岩油储层测井裂缝智能识别 [J]. 地球科学, 2023, 48 (7): 2690-2702.
- [19] 沙明洋, 张思佳, 傅庆财, 等. 基于动态权重的多模型集成水产动物疾病防治事件抽取方法 [J]. 华中农业大学学报, 2023, 42 (3): 80-87.
- [20] 欧阳宵, 陶 红, 范瑞东, 等. 利用标签相关性先验的弱监督多标签学习方法 [J]. 软件学报, 2023, 34 (4): 1732-1748.
- [21] 罗 俊, 高清维, 檀 怡, 等. 基于双拉普拉斯正则化与因果推断的多标签学习 [J]. 计算机工程, 2023, 49 (11): 49-60.
- [22] 王卫军, 宁致远, 杜 一, 等. 基于多标签分类的科技文献学科交叉研究性质识别 [J]. 数据分析与知识发现, 2023, 7 (1): 102-112.
- [23] 邓维斌, 王智莹, 高荣壕, 等. 融合注意力与 CorNet 的多标签文本分类 [J]. 西北大学学报 (自然科学版), 2022, 52 (5): 824-833.
- [24] 陈 曦, 马建敏, 刘权芳. 基于模糊依赖决策熵的多标签特征选择 [J]. 昆明理工大学学报 (自然科学版), 2024, 49 (2): 62-72.
- [25] 李 帆, 张小恒, 李勇明, 等. 基于包络学习和分级结构一致性机制的不平衡集成算法 [J]. 电子学报, 2024, 52 (3): 751-761.
- [16] 谢小川, 周绍军, 黎 力, 等. 基于时序关联智慧城市边缘数据异常检测算法 [J]. 计算机工程与设计, 2022, 43 (7): 1879-1889.
- [17] 杨 娜, 付颖煜, 李天昊. 基于局部离群因子的数据异常识别方法及其在古建筑结构监测中的应用 [J]. 建筑结构学报, 2022, 43 (10): 68-75.
- [18] 李明超, 李明昊, 任秋兵, 等. 基于密度分簇的长周期监测数据异常识别方法 [J]. 水力发电学报, 2021, 40 (3): 124-133.
- [19] 王丽蓉, 郑东健. 基于卷积神经网络的大坝安全监测数据异常识别 [J]. 长江科学院院报, 2021, 38 (1): 72-77.
- [20] 嵇晓燕, 姚志鹏, 杨 凯, 等. 基于 MSLSTM-DA 模型的水质自动监测异常数据报警 [J]. 中国环境科学, 2022, 42 (4): 1877-1883.

(上接第 47 页)

- [11] 何 黎, 陈 磊, 纪莎莎, 等. 基于 K-shape 聚类的连续液位监测数据异常检测方法 [J]. 中国给水排水, 2023, 39 (11): 56-61.
- [12] 何家辉, 程志君, 郭 波. 联合字典学习与 OCSVM 的遥测数据异常检测方法 [J]. 航空学报, 2023, 44 (13): 207-219.
- [13] 夏志祥, 李 准, 徐 伟. 大气电场测量数据的异常检测及校正方法研究 [J]. 电子测量技术, 2023, 46 (1): 90-96.
- [14] 李松轩, 丁 勇, 李登华. 基于影响因子分解法的大坝监测数据异常检测算法 [J]. 人民长江, 2023, 54 (4): 234-240.
- [15] 建剑波, 霍吉祥, 何欣威, 等. 基于小波变换的改进云模型在河口村水库监测数据异常识别中的应用 [J]. 人民黄河, 2022, 44 (11): 149-151.