

基于知识图谱的变压器匝间短路故障辨识研究

查易艺,王 翀,张明明

(国网江苏省电力有限公司信息通信分公司,江苏 南京 210000)

摘 要:变压器出现的故障数据间的关联性没有被较好地利用,会直接影响变压器匝间短路故障的辨识准确性。为此,提出基于知识图谱的变压器匝间短路故障辨识研究。基于柔性策略采集变压器数据,根据实际运行情况及数据采集目标需求,调整数据采集量及时间间隔。根据采集到的变压器数据,采用本体构建、实体抽取、关系抽取和图谱构建的步骤构建知识图谱。将提取的知识图谱故障样本特征,输入到你只看一次(YOLO)v4检测模型中。通过YOLOv4检测模型与知识图谱结合的检测方法,完成变压器匝间短路故障的自动辨识。试验结果表明:变压器匝间短路故障自动辨识的准确率、召回率和 F 值均较高,因而辨别及时性高、自动辨别效果好。该研究解决了传统方法中存在的问题,具有重要的现实意义。

关键词:知识图谱;变压器;匝间短路故障;实体抽取;你只看一次v4检测模型;柔性策略;关系抽取

中图分类号:TH17

文献标志码:A

DOI:10.16086/j.cnki.issn1000-0380.2023010011

Research on Transformer Inter-Turn Short Circuit Fault Identification Based on Knowledge Map

ZHA Yiyi, WANG Chong, ZHANG Mingming

(Information and Communication Branch, State Grid Jiangsu Electric Power Co., Ltd., Nanjing 210000, China)

Abstract: The correlation between the fault data appearing in the transformer is not well utilized, which will directly affect the accuracy of transformer inter-turn short circuit fault identification. For this reason, research on transformer inter-turn short circuit fault identification based on knowledge map is proposed. Based on the flexible strategy to collect transformer data, the amount of data collection and time interval are adjusted according to the actual operating conditions and data collection target demand. Based on the collected transformer data, the knowledge map is constructed using the steps of ontology construction, entity extraction, relationship extraction and map construction. The extracted knowledge map fault sample features are input into the you only look once (YOLO) v4 detection model. The automatic identification of transformer inter-turn short circuit faults is accomplished by the detection method combining the YOLOv4 detection model and the knowledge map. The test results show that the accuracy, recall and F -value of automatic transformer inter-turn short circuit fault identification are high, so the identification timeliness is high, and the automatic identification effect is good. The research solves the problems existing in the traditional method, and has important practical significance.

Keywords: Knowledge map; Transformer; Inter-turn short circuit faults; Entity extraction; You only look once (YOLO) v4 detection model; Flexible strategy; Relational extraction

0 引言

变压器具有电力运输、电压转换与电能分配的作用,是保证电力系统正常运行的重要组成部分之一。匝间短路故障往往发生在变压器故障初期。其故障程度轻且不易被准确识别。一旦没有及时处理,故障会迅速扩大,从而造成严重的安全运行问题。现阶段的变压器匝间故障自动辨识方法存在辨识不及时、辨识

效果不佳的问题。数据挖掘技术的发展可解决传统方法中故障辨识不及时的问题。为了获取更好的变压器匝间故障自动辨识效果^[1-2],相关学者提出了一些方法。

马鑫等^[3]首先通过回溯算法提高残差收缩网络的完整性;其次通过权重交叉熵函数提高残差收缩网络的精度;最后基于数据特征增强方法结合优化后的残差收缩网络模型完成变压器短路故障的自动辨识。

收稿日期:2023-01-16

作者简介:查易艺(1991—),女,硕士,工程师,主要从事数据技术支持、数据应用支撑等方向的研究,E-mail:lh201023@126.com

该方法存在故障辨识错误率较高的问题。张嗣铂^[4]提出电力变压器故障自动化检修技术,并利用核主成分分析方法实现了对变压器故障的自动辨识。但该方法存在故障辨别不及时的问题。曹伟嘉等^[5]首先通过改进遗传算法-极端梯度提升,提取变压器诊断故障特征向量;其次采用遗传算法实现对电力系统分级故障的建模;最后通过改进遗传算法优化变压器故障诊断模型并完成短路故障自动辨识。但该方法存在变压器短路故障辨识精度低的问题。

为了解决上述方法中存在的问题,本文提出基于知识图谱的变压器匝间短路故障辨识方法。该方法首先根据柔性策略,对变压器数据进行采集,并根据实际运行情况和数据采集目标需求调整采集量和时间间隔;然后利用采集到的数据构建知识图谱,提取知识图谱中故障样本特征,输入到你只看一次(you only look once, YOLO)v4 检测模型中,以完成变压器匝间短路故障的自动辨识。试验结果表明,该方法的准确率、召回率和 F 值均较高,具有较好的自动辨别效果。

1 变压器故障关联数据采集

通过传感器配合柔性策略采集变压器数据,可以提高数据采集的准确性^[6-7]。为解决传统采集策略造成大量冗余数据且故障数据难以区分的问题,本文使用柔性策略采集变压器实时数据。柔性策略即根据实际运行情况及数据采集目标需求,调整数据采集量及时间间隔等,以实现最终的数据采集。

本文以区间为单位判断数据波动情况,并以此作为应激式采集间隔调整的依据。本文假设 \bar{g} 为区间内采样平均值,则第一重判断指标为区间内数据标准差 Δg_o 。

$$\Delta g_o = \frac{\sum_{o=1}^l (g_o - \bar{g})^2}{l} \quad (1)$$

式中: g_o 为区间内第 o 个采样点的采样数据; l 为采集区间。

第二重判断指标采用移动平均法判断。变压器设备有瞬时干扰、不可回复性两种特殊状态。根据设备状态量 Δg_1 的计算,排除瞬时干扰。

$$\Delta g_1 = \sqrt{\frac{\sum_{o=1}^l (g_o - g_{o-1})^2}{l-1}} \quad (2)$$

式中: g_{o-1} 为区间内第 $(o-1)$ 个采样点的采样数据。此时采集间隔时间不变,将区间内的一个数据作为目标准则,以判断之后的数据是否存在数据波动。

设备状态不可回复性可通过设备另一状态量 Δg_2 的计算完成。

$$\Delta g_2 = \sqrt{\frac{\sum_{o=1}^l (g_o - g_{o-2})^2}{l-2}} \quad (3)$$

式中: g_{o-2} 为区间内第 $(o-2)$ 个采样点的采样数据。此时,变压器数据采集间隔不需发生改变。根据变压器数据变化量完成数据波动程度的判定,减少了数据采集时间。同时,两种特殊设备状态减少了对数据采集间隔的干扰,提高了数据采集精度。

根据历史数据及数据精度要求,本文设置变压器设备运行数据的最大变化量 ζ_{\max} 和最小变化量 ζ_{\min} 。原始数据平均变化量 $\Delta \bar{g}$ 为:

$$\Delta \bar{g} = \frac{\sum_{o=1}^m (g_o - g_{o-1})}{m-1} \quad (4)$$

式中: m 为原始数据变化的上限值。

本文依据最大变化量、最小变化量设计应激调整方案。具体步骤如下。

(1) 当 $\zeta_{\min} \leq \Delta g_o \leq \zeta_{\max}$ 时,变压器数据采集间隔不变。

(2) 判断设备是否处于如下三种特殊状态。

① 当 $\zeta_{\min} \leq \Delta g_1 \leq \zeta_{\max}$ 时,变压器数据采集间隔不变。

② 当 $\Delta g_1 > \zeta_{\max}$ 且 $\Delta g_2 > \zeta_{\max}$ 时,降低变压器采集间隔。

③ 当 $\Delta g_o \leq \zeta_{\max}$ 时,实行步骤(3)。

(3) 判断变压器是否处于如下两种特殊状态。

① 当 $\zeta_{\min} \leq \Delta g_1 \leq \zeta_{\max}$ 或 $\zeta_{\min} \leq \Delta g_2 \leq \zeta_{\max}$ 时,变压器数据采集间隔不变。

② 当 $\Delta g_1 < \zeta_{\max}$ 且 $\Delta g_2 < \zeta_{\max}$ 时,增大变压器采集间隔。

基于以上步骤,本文完成变压器数据的采集。

2 故障关联知识图谱的构建

故障关联知识图谱的构建过程如下。

① 通过采集到的变压器数据,构建面向变压器故障辨识的知识图谱^[8-9]。本文选择自上而下的知识图谱构建方法。该方法分为本体构建、实体抽取、关系抽取和图谱构建四个步骤^[10-11]。

面向变压器匝间短路故障自动辨别的知识图谱本体可根据专家先验知识获取,并将本体作为实体抽取和关系抽取的规范准则,完成变压器知识图谱的实体抽取。实体抽取的对象为非结构数据。实体抽取选取

原始语料,并将其划分为训练集和测试集两个部分,以分析原始语料中字符的实体信息;以原始语料作为训练集的输入,使用训练集标注标签结果。

语料抽取标签和人工标注标签的准确率一致,均为 A 。

$$A = \frac{Y_A}{Y_A + G_A} \quad (5)$$

式中: Y_A 为实际正确并且预测正确的数据量,个, G_A 为实际正确但预测错误的数据量,个。

语料抽取标签和人工标注标签的召回率一致,均为 R 。

$$R = \frac{Y_A}{Y_A + G_M} \quad (6)$$

式中: G_M 为实际错误但被预测为正确的数据量,个。

知识图谱构建步骤如图 1 所示。

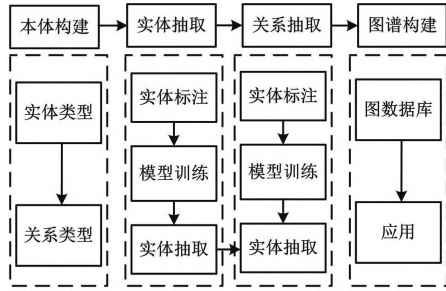


图 1 知识图谱构建步骤

Fig. 1 Construction process of knowledge map

②根据训练集训练效果评估完成实体抽取。针对变压器匝间短路故障的自动识别,本文采用了一种基于注意力机制的双向长短期记忆网络算法,以实现故障的自动辨识。注意力机制模型包含注意力概率分布和注意力分布最终特征计算两个部分。对于最终状态,本文设输入序列元素个数为 M ,则 n 时刻的输出数据注意力概率 s_m 为:

$$s_m = \frac{j_n^T \times I \times G}{\sum_{n=1}^M j_n^T \times A} \times R \quad (7)$$

式中: j_n 为双向隐藏层状态值; I 为权重矩阵; T 为类别标签数量,个; G 为特征向量。

注意力分布的最终特征 G_{att} 为:

$$G_{att} = \sum_{m=1}^M s_m \times j_n \quad (8)$$

经 softmax 函数获取输入注意力机制的数据特征

向量分类标签的概率分布 u 为:

$$\begin{cases} u = \text{softmax}(G'_{att}) = \frac{\exp(G'_{att(o)})}{\sum_{k=1}^T \exp(G'_{att(k)})} \\ G'_{att} = B \times G_{att} \end{cases} \quad (9)$$

式中: $G'_{att(o)}$ 为第 o 个注意力分布的最终特征分量; $G'_{att(k)}$ 为第 k 个注意力分布的最终特征分量; B 为模型输出层权重矩阵。

交叉熵损失函数 $R(U, u)$ 的表达式为:

$$R(U, u) = U \times \lg(u) \quad (10)$$

式中: U 为真实类别分布。

③根据 $R(U, u)$ 构建非结构化知识的实体关系,以抽取构建知识图谱。Neo4j 是一种高度可扩展、基于图的数据库管理系统,以节点和边的形式存储数据,并允许使用索引快速查询。存储原生图数据的 Neo4j 可通过遍历图数据以避免节点查询算法。Neo4j 具有高性能及敏捷性。在 Neo4j 中,每个节点都可以包含任意数量的属性。这些属性可以通过键值对的方式进行存储。节点之间的关系通过边来表示。每条边都可以包含任意数量的属性,并且可以指定方向和权重。通过这种方式,Neo4j 能够清晰、直观地表示出数据的结构和关系。Neo4j 提供了清晰、直观的数据存储和查询方式,可以快速、高效地进行实体关系抽取和知识图谱构建。

3 变压器短路故障自动辨识方法设计

本文将上述构建的面向变压器故障辨识的知识图谱与 YOLOv4 检测模型相结合,完成变压器短路故障自动辨识^[12]。本文提取知识图谱故障样本的特征,将其输入到 YOLOv4 检测模型中,以实现变压器故障识别。YOLOv4 检测模型是一种运行速度快的回归功能实现模型,是通过对 YOLO 模型的跨阶段局部网络改造和金字塔池化优化而来的,具有更好的特征提取能力和特征融合能力。

本文设置信度误差为 Z_{conf} ,则阈值函数 Z 为:

$$Z = (Z_{loc} + Z_{cls} + Z_{conf}) \times R(U, u) \quad (11)$$

式中: Z_{loc} 为回归框预测阈值; Z_{cls} 为分类阈值。

本文假设两框中心点的知识图谱关联注意力距离为 $\sigma^2(S_{ctr}, N_{ctr})$ 、两框最小封闭区域的知识图谱类间距注意力距离为 q ,则 Z_{loc} 为:

$$Z_{loc} = 1 - O_{PI}(S, N) + \frac{\sigma^2(S_{ctr}, N_{ctr})}{q^2} + \frac{\frac{4}{\pi^2} \left(\arctan \frac{e^{hy}}{j^{hy}} - \arctan \frac{e}{j} \right)^4}{\frac{4}{\pi^2} \left(\arctan \frac{e^{hy}}{j^{hy}} - \arctan \frac{e}{j} \right)^2 + [1 - O_{PI}(S, N)]} \quad (12)$$

式中: $O_{pi}(S, N)$ 为预测框 S 和真实框 N 的交并比; e^{hy} 、 e 分别为真实框和预测框的宽度, m ; j^{hy} 、 j 分别为真实框和预测框的高度, m 。

$$Z_{conf} = (O_{ok}^{obj} + O_{ok}^{nobj}) \times \bar{V}_o^k \times \log(V_o^o) \times V_o^k + \mu_{nobj} + Z_{loe} \quad (13)$$

式中: O_{ok}^{obj} 为位于预测边界框内的目标; O_{ok}^{nobj} 为处于预测边界框外的目标; μ_{nobj} 为设置参数; \bar{V}_o^k 、 V_o^k 分别为预测置信度和实际置信度。

4 试验与分析

为了验证基于知识图谱的变压器匝间短路故障自动辨识方法的有效性,本文完成以下测试。本文在计算机系统上使用 Simulink 软件搭建变压器匝间短路模型。

本文采用基于知识图谱的变压器匝间短路故障辨识方法(所提方法)、基于数据特征增强和残差收缩网络的变压器匝间短路故障辨识方法(文献[3]方法)和基于核主成分分析的电力变压器故障自动化检修技术

(文献[4]方法)完成试验。

变压器匝间短路模型如图2所示。

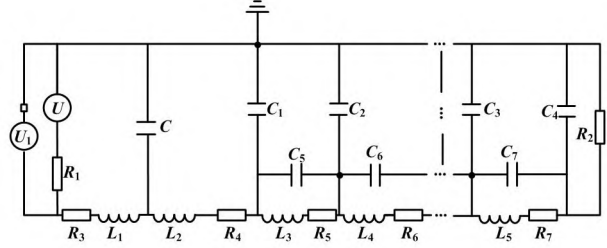


图2 变压器匝间短路模型

Fig. 2 Transformer inter-turn short circuit model

4.1 故障响应特征曲线

本文设置匝1、匝2、匝3短路故障,采用所提方法、文献[3]方法和文献[4]方法完成变压器匝间短路故障检测,并记录变压器匝间短路响应特征曲线。

匝1、匝2和匝3短路故障响应特征曲线分别如图3、图4和图5所示。

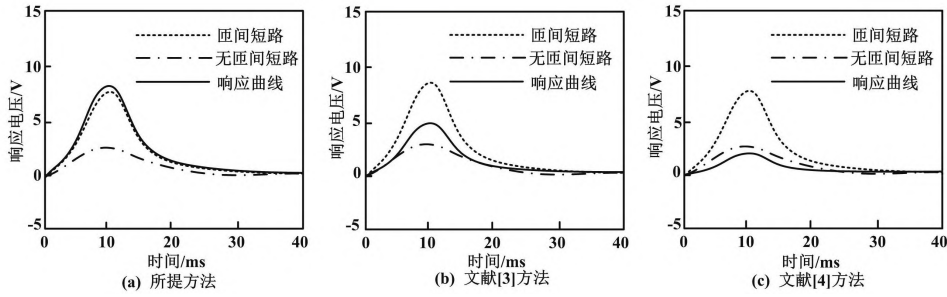


图3 匝1短路故障响应特征曲线

Fig. 3 Turn 1 short circuit fault response characterization curves

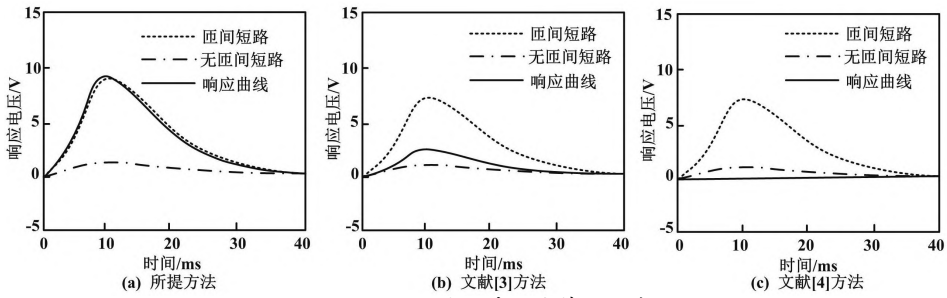


图4 匝2短路故障响应特征曲线

Fig. 4 Turn 2 short circuit fault response characterization curves

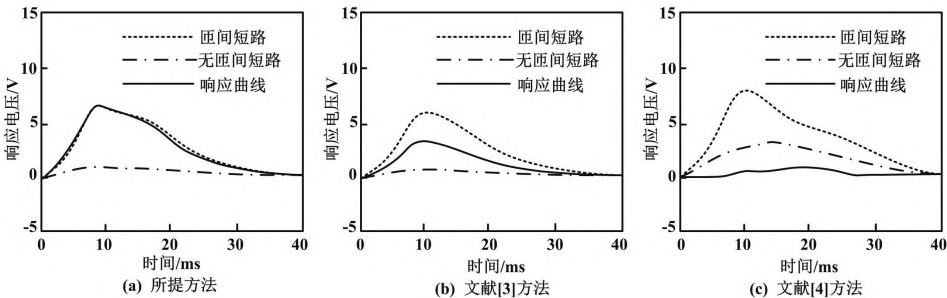


图5 匝3短路故障响应特征曲线

Fig. 5 Turn 3 short circuit fault response characterization curves

由图 3~图 5 可知,文献[3]方法和文献[4]方法在匝 1、匝 2 和匝 3 的短路故障时,响应特征曲线较匝间短路和无匝间短路故障曲线偏离较大。所提方法在匝 1、匝 2 和匝 3 的短路故障时,响应特征曲线与匝间短路故障曲线较为接近。该结果表明,所提方法可以较好地检测到变压器匝间短路故障。其原因是所提方法通过采集到的变压器数据构建面向变压器故障辨识的知识图谱,并选择自上而下的知识图谱构建方法。该方法分为本体构建、实体抽取、关系抽取和图谱构建四个步骤,有利于较好地检查变压器匝间短路故障。

4.2 准确率、召回率、 F 值

本文引入准确率、召回率、 F 值三个指标评估所提方法、文献[3]方法和文献[4]方法的故障自动辨别性能。

F 值是一种衡量变压器匝间短路故障自动辨识精度的指标,在计算过程中综合了辨识方法的准确率和召回率两个指标,具有综合性和平衡性的特点。

F 值为:

$$F = \frac{2 \times A \times R}{A + R} \times 100\% \quad (14)$$

通过 F 值的计算,可以判断故障自动辨别性能是否具有全面性。采用三种方法完成变压器匝间短路故障检测,得到三种方法的检测准确率、召回率、 F 值的结果如表 1 所示。

表 1 三种方法的检测准确率、召回率、 F 值的结果

Tab. 1 Results of detection accuracy, recall, and F -value for three methods

方法	试验次数/次	准确率/%	召回率/%	F 值/%
所提方法	1	97.8	89.9	93.4
	2	98.7	90.1	94.1
	3	98.6	88.9	93.5
文献[3]方法	1	94.9	84.6	91.0
	2	94.5	86.1	92.3
	3	96.6	86.6	91.3
文献[4]方法	1	97.1	86.4	92.3
	2	95.0	84.9	84.6
	3	94.5	85.5	89.6

分析表 1 可知,所提方法的变压器匝间短路故障检测准确率、召回率和 F 值均高于文献[3]方法、文献[4]方法,表明所提方法的变压器短路故障自动辨识效果更佳。其原因是所提方法对非结构数据实行实体抽取。所提方法选取原始语料,并将其划分为训练集和测试集两个部分,以分析语料中字符的实体信息。

以原始语料作为训练集的输入、标签使用训练集标注结果,在一定程度上有利于提高变压器短路故障自动辨识效果。

通过上述试验可知,所提方法通过知识图谱的构建,提高了变压器故障检测准确率;通过 YOLOv4 检测模型与知识图谱结合的检测方法,提高了变压器匝间短路故障自动辨识性能。

5 结论

变压器匝间短路故障影响着电力系统的运行情况,对变压器实行故障自动辨识检测是发展趋势。因此,本文提出基于知识图谱的变压器匝间短路故障辨识方法。本文首先采集变压器数据;其次构建面向变压器匝间短路故障自动辨识的知识图谱;最后将知识图谱与检测模型结合,以完成变压器匝间短路故障的自动辨识。通过研究可知,所提方法的准确率、召回率和 F 值均较高。该方法解决了传统方法中存在的问题,具有重要的现实意义。

参考文献:

- [1] 雷江开,王秀全,崔强.雷击配电变压器的电磁融合仿真分析[J].自动化仪表,2022,43(5):56-60.
- [2] 张鑫,王衡,卫永鹏,等.基于 SSA-PNN 的电力变压器故障诊断[J].工业仪表与自动化装置,2022(1):86-90.
- [3] 马鑫,尚毅梓,胡昊,等.基于数据特征增强和残差收缩网络的变压器故障识别方法[J].电力系统自动化,2022,46(3):175-183.
- [4] 张嗣铂.基于核主成分分析的电力变压器故障自动化检修技术[J].自动化与仪表,2022,37(10):14-18,23.
- [5] 曹伟嘉,杨留方,徐天奇,等.基于 IGA-XGBoost 的变压器故障诊断研究[J].自动化与仪器仪表,2022(4):1-5.
- [6] 曹也,程亮亮,杨昊,等.基于 FPGA 的激光雷达数据采集卡设计[J].量子电子学报,2022,39(4):620-631.
- [7] 张又文,冯斌,陈页,等.基于遗传算法优化 XGBoost 的油浸式变压器故障诊断方法[J].电力自动化设备,2021,41(2):6136-6144.
- [8] 徐先峰,黄坤,邹浩泉,等.基于 SSAE-SVM 的滚动轴承故障诊断方法研究[J].自动化仪表,2022,43(1):9-14.
- [9] 曹宗胜,许倩倩,李朝鹏,等.基于对偶四元数的协同知识图谱推荐模型[J].计算机学报,2022,45(10):221-224.
- [10] 雷江开,王秀全,崔强.雷击配电变压器的电磁融合仿真分析[J].自动化仪表,2022,43(5):56-60.
- [11] 张凯,刘京菊.一种基于知识图谱的威胁路径生成方法[J].计算机仿真,2022,39(4):350-356.
- [12] 郑玉平,龚心怡,潘书燕,等.变压器匝间短路故障工况下的漏磁特性分析[J].电力系统自动化,2022,46(15):121-127.