

DOI: 10.13382/j.jemi.B1902853

基于 XGBoost 特征重要度的储罐 缺陷 ANN 面积量化模型*

唐圳雄¹ 唐东林¹ 丁超² 侯军¹

(1. 西南石油大学 机电工程学院 石油天然气装备教育部重点实验室 成都 610500;

2. 成都工业学院 机械工程系 成都 611730)

摘要:针对超声波检测的储罐缺陷的面积量化问题,提出一种改进的储罐腐蚀缺陷面积量化模型。该模型利用 XGBoost 的特征重要度对人工神经网络(ANN)的参数进行先验初始化实现 ANN 模型的改进。该模型可以更快的收敛,并且提高准确率。按照国家标准设计实验平台,获取实验信号,并提取信号的统计特征得到特征数据集,利用数据集训练和测试改进的模型,与传统模型进行对比。通过实验验证得出,改进的 ANN 模型能够更快的收敛,并且准确量化缺陷面积,相比于 ANN 量化模型,在训练集上准确率提高了 17.9%,达到了 98.3%,在测试集上提高了 16.6%,达到了 92.2%。

关键词:储罐腐蚀缺陷;面积量化;超声波检测;XGBoost 特征重要度;ANN 模型

中图分类号: TN98;TG115.28+5 **文献标识码:** A **国家标准学科分类代码:** 510.40

Quantitative model of ANN area of tank defects based on XGBoost feature importance

Tang Zhenxiong¹ Tang Dongling¹ Ding Chao² Hou Jun¹

(1. Key Laboratory of Petroleum and Natural Gas Equipment, Ministry of Education, Mechanical & Electrical Department, Southwest Petroleum University, Chengdu 610500, China 2. Mechanical Engineering Department, Chengdu Technological University, Chengdu 611730, China)

Abstract: In order to solve the problem that quantifying the area of tank defects detected by ultrasonic wave, an improved quantitative model of tank corrosion defect area is proposed. This model uses the feature importance of XGBoost to initialize the parameters of artificial neural network (ANN) a priori to improve the ANN model. The model can converge faster and improve the accuracy. Design an experimental platform according to national standards, obtain experimental signals, and extract the statistical features of the signals to obtain a data set. Use the data set to train and test improved models, and compare them with traditional models. The experimental results show that the improved ANN model can converge faster and quantify the defect area accurately. Compared with the ANN quantization model, the accuracy in the training set has been improved by 17.9%, reached 98.3%. and increased by 16.6% on the test set, reached 92.2%.

Keywords: tank corrosion defects; area quantization; ultrasonic detection; XGBoost model; ANN

0 引言

金属储罐容易在自然环境中或者和存储介质发生化

学作用而产生金属缺陷^[1-2]。为规避以及降低安全风险,相关部门需要对储罐进行定期检查、评估,然后对储罐进行质量定级,而质量定级与储罐的缺陷面积相关^[3]。目前针对储罐的缺陷检测一般使用超声波检测方式,通过

收稿日期:2020-12-23 Received Date: 2020-12-23

* 基金项目:四川省科技厅科技支撑计划(2017FZ0033)、成都市技术创新研发项目(2018-YF05-00201-GX)、西南石油大学国家重点实验室项目(PLN201828)资助

检测信号来推断缺陷的面积。众多学者针对缺陷量化进行了大量的研究,如李东升等^[4]对超声检测回波和缺陷的对应关系,建立了对管道壁缺陷量化模型。周正干等^[5]针对浸没式超声波探伤,研究了一种新的基于频域合成孔径聚焦技术(FSAFT-DQM)的 6 dB 跌落缺陷量化方法。这些方法均具有一定的优势,但是相比于缺陷量化的另一些方法,其量化过程和模型结构不够方便和快捷。如程迪等^[6]提出一种基于最小二乘支持向量机的缺陷漏磁检测的缺陷量化方案,建立了缺陷的三维漏磁信号与缺陷的长度、宽度、高度的映射关系,仿真结果表明具有有效性。王太勇等^[7]针对油管缺陷利用人工神经网络解决了信号特征量与缺陷几何外形特征之间的非线性映射问题,建立了基于特征分类的油管缺陷量化识别模型,实验表明能够得到较好量化结果。Liu 等^[8]利用 CNN-LSTM 模型对在线检测的 CO₂ 焊接缺陷进行了量化识别。Wu 等^[9]利用 BP 神经网络量化钢棒中的缺陷。Ding 等^[10]提出了一种基于 ADABOOT-SVM 集成学习模型的缺陷检测与识别框架用来识别量化印刷桶的缺陷问题。唐东林等^[11]针对的管道腐蚀超声内检测研究了基

于主成分分析(PCA)和支持向量机(SVM)的量化模型。梁远远等^[12]BP 神经网络量化针对脉冲涡流检测工业缺陷进行了量化研究。但是这些方法针对超声波检测缺陷面积的量化并不能有理想的效果,所以需要一种针对面积的量化模型去实现超声检测的缺陷面积量化。

本文针对超声波检测方式,提出了一种基于 XGBoost 特征重要度的 ANN 量化模型,将该模型用于超声数据集进行实验,结果显示该模型能够准确量化储罐缺陷,并且和其他无损检测的量化模型做了对比,展示了其优良性能。

1 超声波检测介绍

超声波检测信号利用超声波技术发射超声脉冲进入工件内部,由工件内部的反射信号来判断是否有缺陷。其检测形式如图 1 所示,图 1(a)为超声直探头检测示意图,超声波的反射信号如图 1(b)所示。目前针对超声波检测方式的缺陷量化模型精度普遍不高^[13-15],所以提出一种改进的模型来提高精度是必要的。

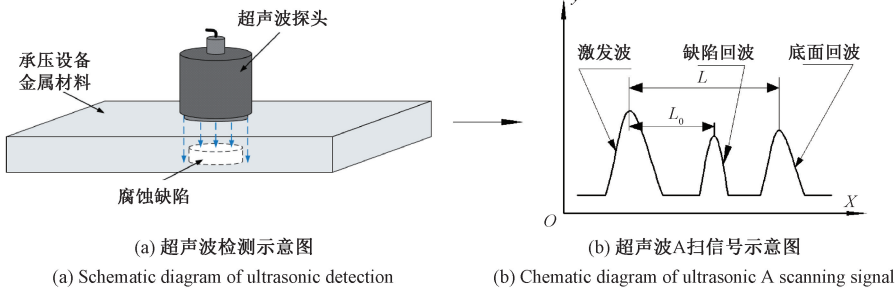


图 1 超声波检测

Fig. 1 Ultrasonic detection

2 基于 XGBoost 特征重要度的单神经元

设有训练集: $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$, 其中 x_i 为一特征组成的向量, y_i 为样本 x_i 对应的输出, 以此类推。假设一棵 CART 回归树的模型为:

$$\hat{y}_i = f(x_i) \tag{1}$$

则 XGBoost 的目标函数可表达为:

$$\text{Objective} = \sum_{i=1}^n \text{loss}(y_i, \hat{y}_i) + \sum_{i=1}^I (f(x_i)) \tag{2}$$

其中:

$$\begin{cases} \text{loss}(y_i, \hat{y}_i) = \frac{1}{2m} \sum_{j=1}^m (\hat{y}_i - y_i)^2 \\ \Omega(f(x_i)) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T f(x_i^j)^2 \end{cases} \tag{3}$$

其中 y_i 、 \hat{y}_i 分别代表真实值和模型的预测值。将目标函数经由贪心算法和去掉常数项,最终目标函数可表达为:

$$\text{Objective} = \sum_{j=1}^T \left[G_j w_j + \frac{1}{2} (H_j + \lambda) w_j^2 \right] + \gamma T \tag{4}$$

式中: γ 和 λ 为常数系数; T 为决策树的总叶子节点个数; j 代表叶子节点的序号; W_j 代表树模型关于叶子节点序号的函数; G_i 和 H_i 为贪心算法关于第 i 颗树的一阶导和二阶导常数。如果树的结构确定,则:

$$\begin{cases} w_j^* = -\frac{G_i}{H_j + \lambda} \\ \text{loss}^* = -\frac{1}{2} \sum_{j=1}^T \frac{G_i^2}{H_j + \lambda} + \gamma T \end{cases} \tag{5}$$

因为损失函数越小越好,即是 $\frac{G_i^2}{H_j + \lambda}$ 越大越好;从而,对于一个叶子节点的分裂的分裂,分裂前后的信息增

益定义为:

$$Gain = \frac{1}{2} \left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{(H_L + H_R) + \lambda} \right] - \gamma \quad (6)$$

Gain 值越大,分裂后减少的损失值越大。所以对于一个叶子节点分割时,计算所有候选的特征对应的 Gain,选择 Gain 最大特征进行分割。划分结束后,即回归树模型构建完毕,特征 j 在单棵树中的重要度计算公式为:

$$\hat{J}_j^2(T) = \sum_{t=1}^{L-1} \hat{v}_t^2 I(v_t = j) \quad (7)$$

式中:L 为树的叶子节点数量;L-1 为树的非叶子节点的数量;v_t 是和节点 t 相关联的特征;v_t² 是节点 t 分裂之后平方损失的减少值;I 为 sign 函数,即 v_t = j 时其值取 1,不等时取 0。则特征 j 在 XGBoost 模型下的全局重要度通过特征 j 在所有树中的重要度的平均值来衡量,其计算公式如下:

$$\hat{J}_j^2 = \frac{1}{M} \sum_{m=1}^M \hat{J}_j^2(T_m) \quad (8)$$

式中:M 是树的数量。ANN 的单个神经元的输出 h_{w,b}(x_i) 可表达为:

$$h_{w,b}(x_i) = f(W^T x_i + b) = f\left(\sum_{j=1}^n W_j x_i^j + b\right) \quad (9)$$

其中 f 为激活函数,本文利用 Relu 函数作为激活函数, W^T 为 x_i 的权重参数,b 为输入的偏置。则将 XGBoost 的特征重要度加入 ANN 模型可表达为:

$$h_{w,j,b}(x_i) = f(W^T, \hat{J}^{2T}, b) = \max\left(0, \sum_{j=1}^n \hat{J}_j^2 W_j x_i^j + b\right) \quad (10)$$

式(10)即为基于 XGBoost 的 ANN 模型的一个神经元的计算公式。

3 基于 XGBoost 特征重要度的 ANN 模型的构建

ANN 即人工神经网络,其大致可分为 3 部分组成,即输入层、隐藏层和输出层。其中输入层即一个数据样本,其由多个特征组成,隐藏层为多个神经元组成,其可为多层结构,输出层即为隐藏层神经元的输出^[16]。

基于 XGBoost 特征重要度的 ANN 模型由基于 XGBoost 特征重要度的神经元组成。如第 i 个样本输入到单个神经元的示意图如图 2 所示,输入到神经网络如图 3 所示。

构建基于 XGBoost 特征重要度的 ANN 模型的步骤如下。

1) 由数据集 D 构建第一棵树模型,此模型的构建方法由式(5)、(6)决定。由式(7)计算此棵树模型的所有

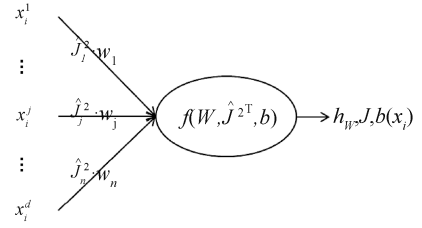


图 2 单神经元示意图

Fig. 2 Single neuron diagram

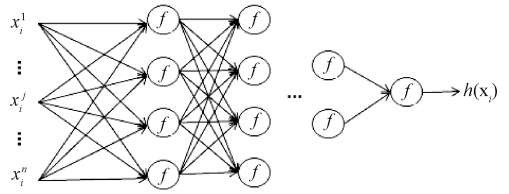


图 3 深度神经网络示意图

Fig. 3 Schematic diagram of deep neural network

特征的重要度。

2) 由第 1 棵树模型预测 (X, Y) 样本, 得到预测值 \hat{Y} 。以 (X, Y - \hat{Y}) 为样本构建第 2 棵树模型。

3) 重复步骤 1) ~ 2), 直至构建的模型使得 Y - \hat{Y} = 0, 或者达到给定阈值。

4) 将构建时的计算的特征重要度按式(8)进行计算统计, 得到最终所有特征重要度。

5) 按照式(10)得到基于 XGBoost 特征重要度的单个神经元。

6) 根据需求由多个基于 XGBoost 特征重要度的神经元构建 ANN 模型。

4 实验

4.1 超声波信号的采集和处理

实验平台由含缺陷的符合储罐要求的高强度 Q235B 金属钢板, 2.5 MHz 超声波探头, RIGOL DS4024 型示波器, 信号激发和前端处理电路板, 电源, 导线等构成, 如图 4 所示。

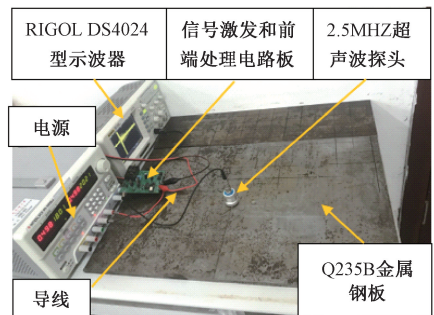
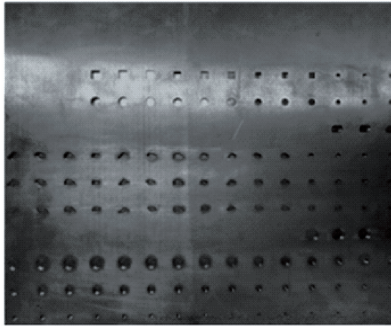


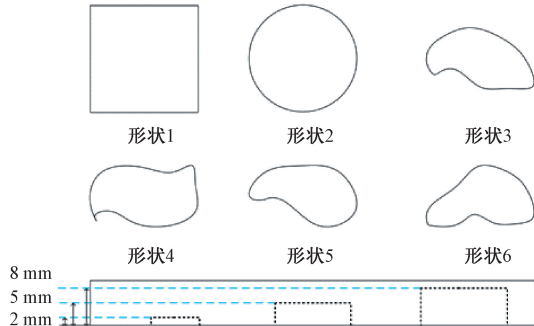
图 4 实验装置图

Fig. 4 Experimental device

本文针对缺陷的面积量化设计了对应的实验样本,实验样本依据中华人民共和国能源行业标准^[3]设计了 1 220 个缺陷的样本库,缺陷面积在 0~314 mm²,具有 6 种不同形状和 3 种不同深度,部分缺陷实物如图 5(a)所示



(a) 部分金属缺陷样本实物
(a) Some metal defect samples



(b) 样本缺陷形状及深度
(b) Sample defect shape and depth

图 5 金属缺陷样本

Fig. 5 Metal defect sample

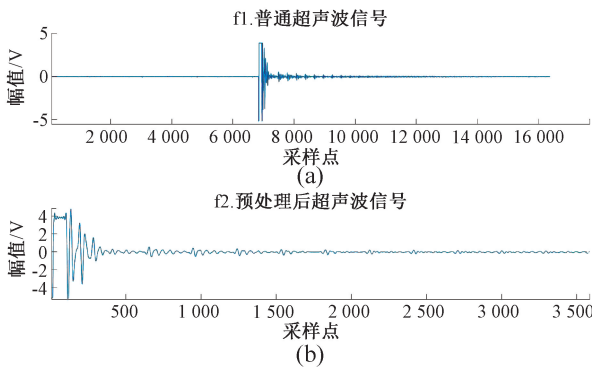


图 6 超声波检测信号

Fig. 6 Ultrasonic detection signal

4.2 特征提取

信号的特征提取是模式识别方法取得良好结果的基础,本文着重研究模型的改进对结果的影响,所以本文大多仅提取超声波的统计学的特征参数来表征超声波信号。特征参数的计算公式如表 1 所示。其中补充了信号能量,信号过零率、信号谱熵 3 个特征来表征信号在能量域,频率域以及功率域的状态信息。

表 2 为部分缺陷信号提取的特征值(所有值保留 4 位有效数字)。一个缺陷对应着 11 个特征数据,其标签即为该缺陷对应的面积。

4.3 模型的训练和测试

在缺陷量化领域,量化的模型常有线性回归模型(LR)、ANN、支持向量机回归(SVR),本文引入了在目前在回归问题上表现比较好的 XGBoost^[17],并且利用 XGBoost 特征重要度来改进 ANN 模型用来量化储罐缺陷面积,利用特征集分别训练和测试这 5 种机器学习模型。

示,其中形状以及深度的设计如图 5(b)所示。实际超声波检测信号如图 6(a)所示,经过去噪和有效信息段截断得到的信号如图 6(b)所示。

表 1 特征类型

Table 1 Feature types

编号	特征名称	描述	公式
1	mean	信号均值	$\sum X/N$
2	std	信号方差	$\sqrt{\sum (X-mean)^2/N-1}$
3	peak	信号峰值	$peak = \frac{\max X }{\sqrt{\frac{\sum Sig^2}{N}}}$
4	gradient	信号梯度	$\sum (X-mean)^3/((N-1) \times std^3)$
5	kurtosis	峰态	$\sum (X-mean)^4/((N-1) \times std^4)$
6	Clear	清除指标	$\frac{\max X }{\sum \sqrt{ X /N}}$
7	shape	信号的形状指标	$\frac{\sqrt{\sum X^2/N}}{\sum X /N}$
8	impulse	脉冲指标	$\frac{\max X }{\sum X /N}$
9	energy	信号能量	$\sum X^2$
10	zerocro	信号过零率	$N(X > 0)/N$
11	spectral	信号谱熵	$\ln \sqrt{2\pi e} + \frac{1}{2} \int_{-\frac{1}{2}}^{\frac{1}{2}} \ln G_x(X) df$

由 4.1 节的实验平台采集 1 220 个缺陷的超声波信号,为增加实验数据以及对抗实验采集偏差带来的数据噪音,每个缺陷采集 10 次超声波信号,最终得到 12 200 个实验信号,按照 4.2 节方法提取每个实验信号的特征,并给每个数据贴上对应的面积标签,最终构成 12 200 个特征数据集,将 80% 用作训练数据其余作为测试数据用

表 2 部分缺陷特征提取结果

Table 2 Partial defect feature extraction results

面积/mm ²	mean	std	peak	gradient	kurtosis	clear	shape	impulse	energy	zerocro	spectral
78.510 0	1.174 0	1.266 0	1.011 0	28.240 0	7.332 0	6.407 0	4.021 0	1.456 0	2.734 0	0.324 6	0.161 3
94.990 0	2.408 0	3.519 0	1.892 0	33.870 0	10.790 0	8.688 0	5.825 0	2.247 0	0.611 9	0.389 5	1.936 0
113.000 0	2.988 0	5.126 0	2.058 0	35.860 0	15.640 0	12.350 0	6.863 0	3.485 0	0.455 6	0.356 4	3.654 0
132.7	3.784 0	6.694 0	3.254 0	39.650 0	19.540 0	14.580 0	9.654 0	4.658 0	0.406 8	0.338 7	5.497 0

来训练和测试模型。

本文选用了 Accuracy@ x% 作为评价模型在数据集上的表现,如 Accuracy@ 80% 其表示预测方法可以保证至少 80% 测试数据集的误差区间的半宽度^[18]。Accuracy@ 80%±10% 意味着至少 80% 的数据集的预测误差在真实缺陷面积的±10% 范围内。根据应用任务的敏感性,也可以考虑更高或更低的确定性级别(例如@ 95% 或@ 70%)。本文还采用了均方根误差(RMSE)和拟合优度 R² 来评价模型效果^[19-20]。RMSE 可以度量模型真实值与预测值之间的总损失。拟合优度 R² 可表征模型预测值和真实之间的拟合情况,可以表达模型预测准确率,其范围为(-∞, 1]。RMSE 和 R² 表达式如式(11)所示。本文最终的结果如表 3 所示,图 7 所示为模型的累计误差随数据集大小变化(即 x 的变化)的分布。

式中: N 表示样本的数量, $\bar{y} = \frac{1}{m} \sum_{i=1}^m y_i$ 。

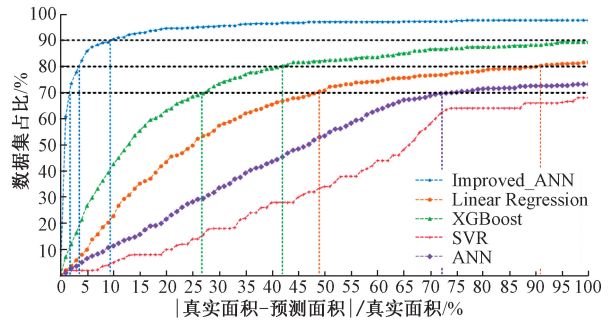


图 7 5 种模型的累计误差分布

5 结果与讨论

图 7 为 5 种模型的数据集比例随累计误差分布的变化曲线,即 Accuracy 曲线,当曲线和横轴之间的面积越大时,表示模型的累计误差越小,模型结果越优。表 3 为 5 种模型在测试集上 Accuracy@ 70%、Accuracy@ 80%、Accuracy@ 90%、拟合优度 R² 和均方根误差 RMSE 的结果。其中拟合优度 R² 越接近数值 1 表示模型拟合效果

$$\left\{ \begin{aligned}
 Accuracy@ x\% &= \frac{x}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \\
 RMSE &= \sqrt{MSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \\
 R^2 &= 1 - \frac{RSS}{TSS} = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}
 \end{aligned} \right. \quad (11)$$

表 3 模型效果

Table 3 Model effects

模型名称	不同数据集缺陷范围/%			拟合优度 R ²		均方根误差 RMSE		收敛时间/s
	Accuracy@ 70%	Accuracy@ 80%	Accuracy@ 90%	训练	测试	训练	测试	
LR	49	91	100	0.706	-4.471	35.637	115.354	135.0
SVR	100	100	100	0.405	0.365	60.126	68.541	153.0
ANN	72	100	100	0.804	0.756	28.455	30.957	176.0
XGBoost	26	42	100	0.954	0.841	1.879 7	21.685	168.0
改进的 ANN	2	4	9	0.983	0.922	1.450 2	11.289 4	94.0

越好,其取值为负数时表明量化模型对数据集的拟合非常差。均方根误差 RMSE 代表了量化模型预测数据集的总损失,其值越接近 0 表示模型效果越好。

从图 7 可以看到,改进的 ANN 模型在 Accuracy 指标上,其曲线和横轴的面积最大,代表改进的模型优于其他 4 种模型。

从表 3 可知,本文提出的改进的 ANN 模型 Accuracy@ 90% 为 9%,意味着至少 90% 的数据集的预测误差在真实缺陷面积的±9% 范围内,Accuracy@ 80% 仅为 4%,而 Accuracy@ 70% 为 2%。反观其他 4 种模型,Accuracy@ 90% 全为 100%,即 90% 的数据集预测误差在真实缺陷面积的 100% 范围内,其中仅有 XGBoost 的 70% 的数据预测误差在真实数据的 26% 范围内,其余表现均

不是很理想,所以从 Accuracy 指标上来看,改进的 ANN 具有最佳性能。

从表 3 的拟合优度 R^2 、均方根误差 $RMSE$ 来看,改进的 ANN 模型在训练集和测试集的拟合优度值分别为 0.983 和 0.922,均方根误差分别为 1.450 2 和 11.289 4,均优于其他 4 种模型,说明改进的 ANN 模型能够很好的拟合数据集,具有较强的稳定性。从表 3 的收敛时间来看,改进的 ANN 模型完成收敛的时间为 94.0 s,相比于其他模型大大缩短了收敛时间。这与理论是相符的,因为基于 XGBoost 特征重要度的先验参数使得模型进行参数寻优时更加容易。

综合分析以上各个评价指标,改进的 ANN 模型具有更好的准确度和收敛速度,模型对储罐缺陷面积量化更具有稳定性和适用性。

6 结 论

本文旨在为储罐缺陷构建一个能准确量化缺陷面积的模型。本文针对传统无损检测模型难以量化超声干耦合检测的缺陷的面积问题,提出了一种改进的 ANN 模型进行面积量化。本文对比了改进的 ANN 和无损检测常用的 3 种模型以及 XGBoost 模型的优劣,结果表明基于 XGBoost 特征重要度的 ANN 模型能更好的适应超声波检测的腐蚀缺陷面积量化。

参考文献

- [1] 王宝轩,沈功田,闫河,等. 大型石油储罐健康管理方法应用研究[J]. 机械工程学报,2017,53(16):125-133.
WANG B X, SHEN G T, YAN H, et al. Application research on health management methods for large oil storage tanks [J]. Journal of Mechanical Engineering, 2017, 53(16): 125-133.
- [2] CHANG J I, LIN C C. A study of storage tank accidents[J]. Journal of Loss Prevention in the Process Industries, 2006,19(1): 51-59.
- [3] 国家能源局. GB/T 47013. 3—2015 承压设备检测第三部分:超声检测[S]. 2015.
National Energy Administration. GB/T 47013. 3—2015 Pressure equipment testing part 3: Ultrasonic testing [S]. 2015.
- [4] 李东升,王昌明,施祖康,等. 管道壁缺陷超声波在役检测的量化分析研究[J]. 仪器仪表学报,2002,23(2):131-134.
LI D SH, WANG CH M, SHI Z K, et al. Quantitative analysis and study on ultrasonic in-service detection of pipeline wall defects [J]. Chinese Journal of Scientific Instrument, 2002,23(2):131-134.

- [5] 周正干,周江华,章宽爽,等. 合成孔径聚焦在水浸超声缺陷定量中的应用[J]. 北京航空航天大学学报,2016,42(10):2017-2023.
ZHOU ZH G, ZHOU J H, ZHANG K SH, et al. Application of synthetic aperture focusing in quantitative analysis of ultrasonic defects in water immersion [J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(10):2017-2023.
- [6] 程迪,黄松岭,赵伟,等. 基于 PSO-LS-SVM 的储罐底板缺陷量化方法研究[J]. 电测与仪表,2018,55(4):87-92.
CHENG D, HUANG S L, ZHAO W, et al. Research on quantification method of tank floor defects based on PSO-LS-SVM [J]. Electric Measurement and Instrument, 2018, 55(4):87-92.
- [7] 王太勇,胡世广,杨涛,等. 一种油管缺陷量化识别技术[J]. 中国机械工程,2005,16(20):1802-1804.
WANG T Y, HU SH G, YANG T, et al. A quantitative identification technique for tubing defects [J]. China Mechanical Engineering, 2005,16(20): 1802-1804.
- [8] LIU T Y, BAO J S, WANG J L, et al. A Hybrid CNN-LSTM algorithm for online defect recognition of CO₂ welding[J]. Sensors, 2018, 18(12):4369.
- [9] WU B, QI W B, HE C F. Recognition of defects on steel rod using ultrasonic guided waves based on neural network [J]. Engineering Mechanics, 2013(2): 470-476.
- [10] DING S, LIU Z, LI C. AdaBoost learning for fabric defect detection based on HOG and SVM [C]. International Conference on Multimedia Technology, IEEE, 2011:2903-2906.
- [11] 唐东林,魏子兵,潘峰,等. 基于 PCA 和 SVM 的管道腐蚀超声内检测[J]. 传感技术学报,2018,31(7):1040-1045.
TANG D L, WEI Z B, PAN F, et al. Ultrasonic internal inspection of pipeline corrosion based on PCA and SVM [J]. Journal of Sensing Technology, 2018, 31(7): 1040-1045.
- [12] 梁远远,杨生胜,文轩,等. 脉冲涡流无损检测中缺陷量化技术研究[J]. 仪器仪表学报,2018,39(11):70-78.
LIANG Y Y, YANG SH SH, WEN X, et al. Research on quantification of defects in pulsed eddy current nondestructive testing [J]. Journal of Instrumentation, 2018, 39(11): 70-78.
- [13] 沙胜义,项小强,伍晓勇,等. 输油管道环焊缝超声波内检测信号识别[J]. 油气储运,2018,37(7):757-761.

- SHA SH Y, XIANG X Q, WU X Y, et al. Recognition of ultrasonic detection signals in girth welds of oil pipelines [J]. Oil & Gas Storage and Transportation, 2018, 37 (7): 757-761.
- [14] 刘松, 顾继俊, 汪颖, 等. 基于超声波检测的 BP 神经网络缺陷识别方法设计 [J]. 压力容器, 2019, 36(8): 62-66, 49.
- LIU S, GU J J, WANG Y, et al. Design of BP neural network defect recognition method based on ultrasonic detection [J]. Pressure Vessel, 2019, 36 (8): 62-66, 49.
- [15] 张鹏, 蒲正元. 管道缺陷漏磁和超声波检测数据的对比分析 [J]. 中国安全科学学报, 2014, 24 (10): 113-119.
- ZHANG P, PENG ZH Y. Comparative analysis of magnetic flux leakage and ultrasonic testing data of pipeline defects [J]. Chinese Journal of Safety Science, 2014, 24 (10): 113-119.
- [16] HINTON G E. Reducing the dimensionality of data with neural networks [J]. Science, 2006, 313 (5786): 504-507.
- [17] CHEN T, GUESTRIN C. XGBoost: A scalable tree boosting system [C]. AcmSigkdd International Conference on Knowledge Discovery & Data Mining, 2016.
- [18] LAYOUNI M, HAMDI M S, TAHAR S. Detection and sizing of metal-loss defects in oil and gas pipelines using pattern-adapted wavelets and machine learning [J]. Applied Soft Computing, 2016, DOI: 10.1016/j.asoc.2016.10.040.
- [19] KRUSKAL J. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis [J]. Psychometrika, 1964, 29(1):1-27.
- [20] CHAI T, DRAXLER R R. Root mean square error (RMSE) or mean absolute error (MAE)? - Arguments against avoiding RMSE in the literature [J]. Geoscientific Model Development, 7(3):1247-1250.

作者简介



唐圳雄(通信作者), 2017 年于西南石油大学获得学士学位, 现为西南石油大学硕士研究生, 主要研究方向为模式识别在无损检测上的应用。

E-mail: 1055953972@qq.com

Tang Zhenxiong (Corresponding author)

received his B. Sc. degree from Southwest Petroleum University in 2017. Now he is a M. Sc. candidate at Southwest Petroleum University. His main research interest includes the application of pattern recognition in non-destructive testing.