

基于 SHAP 可解释性的焊缝缺陷类型超声识别 XGBoost 模型

陈明良¹, 马志远¹, 张东辉², 付冬欣², 廖静瑜², 林莉¹

(1. 大连理工大学 无损检测研究所, 大连 116024; 2. 中国核工业二三建设有限公司, 北京 101300)

摘要: 针对焊缝缺陷机器学习超声识别过程中存在特征冗余、可解释性差等问题, 提出了一种基于 SHAP 可解释性的焊缝缺陷超声识别 XGBoost(极限梯度提升)模型。在碳钢焊缝试样上加工 4 类典型缺陷, 采用横波斜入射法采集超声反射回波信号, 分别提取 16 个时域特征、16 个频域特征以及 3 个信息熵特征。计算 SHAP 值并选择其前 8 个高贡献特征构建特征子集, 利用交叉验证和网格搜索优化 XGBoost 模型进行缺陷识别。试验结果表明, 4 种缺陷识别的平均准确率为 96.7%; 其中, 横通孔的识别效果最佳, 精确率、召回率和 $F_{1-score}$ 均达到 100%, 三角槽次之, 方形槽略差, 矩形槽的识别结果最差, 其精确率、召回率和 $F_{1-score}$ 均为 93.3%。最后, 讨论了高贡献特征与缺陷类别之间的相关性, 并对特征贡献差异及其原因进行了分析。

关键词: 超声检测; 缺陷分类; XGBoost 模型; 特征选择; SHAP

中图分类号: TG115.28; TB553

文献标志码: A

文章编号: 1000-6656(2024)06-0036-07

XGBoost model for ultrasonic recognition of weld defect types based on SHAP interpretability

CHEN Mingliang¹, MA Zhiyuan¹, ZHANG Donghui², FU Dongxin², LIAO Jingyu², LIN Li¹

(1. NDT & E Laboratory, Dalian University of Technology, Dalian 116024, China;

2. China Nuclear Industry 23 Construction Co., Ltd., Beijing 101300, China)

Abstract: Aiming at the problems of feature redundancy and poor interpretability in the process of ultrasonic recognition of weld defects based on machine learning, an XGBoost model for ultrasonic recognition of weld defects was proposed based on SHAP interpretability. Four kinds of typical defects were machined on carbon steel weld samples. The ultrasonic reflection echo signal was collected by shear wave oblique incidence method, and 16 time-domain features, 16 frequency-domain features and 3 information entropy features were extracted respectively. SHAP values were calculated and 8 high-contribution features were selected to build a feature subset. Cross-validation and grid search were utilized to optimize the XGBoost model. The feature subset was used as input to identify the defect types. The results showed that average recognition accuracy of the four defects was 96.7%. Among them, the recognition effect of the transverse hole was the best, and the precision, recall and $F_{1-score}$ all reached 100%, followed by the triangular groove, and the square groove was slightly worse. The recognition result of the rectangular groove was relatively poor, and its precision, recall and $F_{1-score}$ were all 93.3%. Finally, the correlation between high contribution features and defect categories was discussed, and the difference of feature contribution and its causes were analyzed.

Key words: ultrasonic testing; classification of defect; XGBoost model; feature selection; SHAP

收稿日期: 2024-01-29

作者简介: 陈明良(1999—), 男, 硕士研究生, 主要研究方向为超声检测缺陷识别

通信作者: 林莉(1970—), 女, 教授, 博士生导师, 主要研究方向为超声检测与评价, linli@dlut.edu.cn

超声检测具有灵敏度高、操作简单、适用范围广等优点^[1], 是金属材料焊接质量控制的重要手段, 根据缺陷回波特征识别焊缝缺陷类型是质量评定的重要环节^[2]。随着人工智能和信号处理技术的发展, 缺

陷超声检测的智能化识别已成为未来的发展趋势^[3]。大量研究利用神经网络^[4]、支持向量机(SVM)^[5]、极限梯度提升(XGBoost)^[6]等机器学习模型提高超声检测的识别准确率和效率^[7]。其中,XGBoost是结合多个学习器的集成学习模型,具有正则化、并行处理运算、内置交叉验证等优势,能够获得优于单一模型的泛化性能和准确率,已成为近年来倍受关注的机器学习算法之一^[8]。

除机器学习模型外,特征选择也对模型性能具有重要影响。胡宏伟等^[9]对夹渣、气孔和未焊透3类焊缝缺陷提取20个时域及时频域特征,采用主成分分析法选取贡献率超过90%的前6个主成分作为特征向量,SVM的分类准确率达96.7%。经特征选择可以提高模型识别的准确率和效率,但由于机器学习模型为“黑箱”模型,普遍存在模型透明度低^[10]、可解释性差^[11]等问题,若模型决策发生错误,其执行代价将难以承受。为了解锁“黑箱”模型,2017年提出了基于博弈论的SHAP (SHapley Additive exPlanation)方法,该方法通过量化贡献程度来辅助优化特征选择^[12],同时也能增强机器学习模型的可解释性^[13],尤其适合XGBoost等集成学习模型的事后解释^[14]。GHOLIZADEH等^[15]使用经SHAP选择后的5个关键特征对CFRP(碳纤维基增强复合材料)试件损伤程度进行多模型分类,其中XGBoost模型的识别分类准确率为95.7%。

文章针对碳钢焊缝人工缺陷类型的超声识别问题,基于SHAP可解释性方法进行特征优选,构建具有可解释性的缺陷类型超声识别XGBoost模型,结合超声与缺陷作用分析了不同特征贡献差异及产生原因,并根据试验结果,进一步借助SHAP方法讨论了高贡献特征与缺陷类别之间的相关性。

1 试验原理

1.1 XGBoost模型

XGBoost算法是基于决策树的集成机器学习算法,其将所有决策树的结果相加作为模型最终输出^[16],即

$$\hat{y}_i = \sum_{l=1}^L f_l(W_i), f_l \in F \quad (1)$$

式中: \hat{y}_i 为第*i*个样本的预测值; f_l 为第*l*棵决策树, L 为决策树的总数; F 为所有决策树的集合; W_i 为第*i*个信号样本的特征数据集。

XGBoost算法的目标函数为

$$O^{(t)} = \sum_{i=1}^n l\left(y_i, \hat{y}_i^{(t-1)} + f_t(W_i)\right) + \Omega(f_t) \quad (2)$$

$$\Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{\alpha=1}^T \omega_{\alpha}^2 \quad (3)$$

式中: $\hat{y}_i^{(t-1)}$ 为第*t*-1次迭代时第*i*个样本的预测值; $\sum_{i=1}^n l(y_i, \hat{y}_i)$ 为损失函数,用来衡量样本的预测值 \hat{y}_i 和真实值 y_i 之间的误差; $\Omega(f_t)$ 为减少过拟合风险的惩罚函数; γ, λ 为正则化系数; T 为叶节点数; ω_{α} 为第*α*个叶节点的权重。

1.2 SHAP方法

SHAP方法通过多个线性模型对复杂机器学习模型进行逼近,以此给出预测过程中每个特征对预测值的贡献度,即SHAP值^[17],其计算过程可写为

$$S_{ij} = \sum_{Q \in N \setminus \{j\}} \frac{q!(n-q-1)!}{n!} \times \left[\hat{y}_i(\omega_i^{Q \cup \{j\}}) - \hat{y}_i(\omega_i^Q) \right] \quad (4)$$

式中: N 为所有特征的集合,共*n*个特征; Q 为集合*N*中部分特征(至少1个特征)组成的子集; q 为子集*Q*中包含的特征数量; $\hat{y}_i(\omega_i^Q)$ 为在不加入特征*j*时,使用子集*Q*中特征构建的模型得到的第*i*个样本的预测值; $\hat{y}_i(\omega_i^{Q \cup \{j\}})$ 为在加入特征*j*后,对应模型得到的预测值; S_{ij} 为第*i*个信号样本的第*j*个特征 w_{ij} 的SHAP值(若 $S_{ij} > 0$,则表示该特征的SHAP值与模型预测值成正相关;反之,成负相关)。

2 试验方法

2.1 试样制备

试样材料为X52低合金高强度钢,焊接方法为熔化极惰性气体保护焊,其尺寸为200 mm×40 mm×18 mm(长×宽×高)。采用线切割的方法在试样焊缝中加工矩形槽、三角槽、方形槽、横通孔等4类典型缺陷。其中,矩形槽代表面积型缺陷,其尺寸为2 mm×0.5 mm(长×宽);横通孔代表体积型缺陷,其直径为2 mm;三角槽和方形槽代表由面积型向体积型缺陷的过渡,三角槽边长为2 mm,方形槽尺寸为2 mm×2 mm(长×宽)。每类缺陷各制作5个不同的试样,即缺陷中心位置距底面距离分别为4.0, 4.5, 5.0, 5.5, 6.0 mm,总共20个缺陷试样。4种碳钢缺陷试样实物如图1所示。

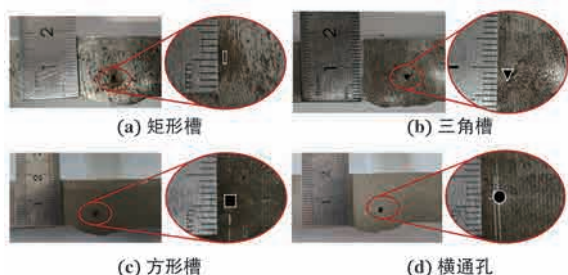


图1 4种碳钢焊缝缺陷试样实物

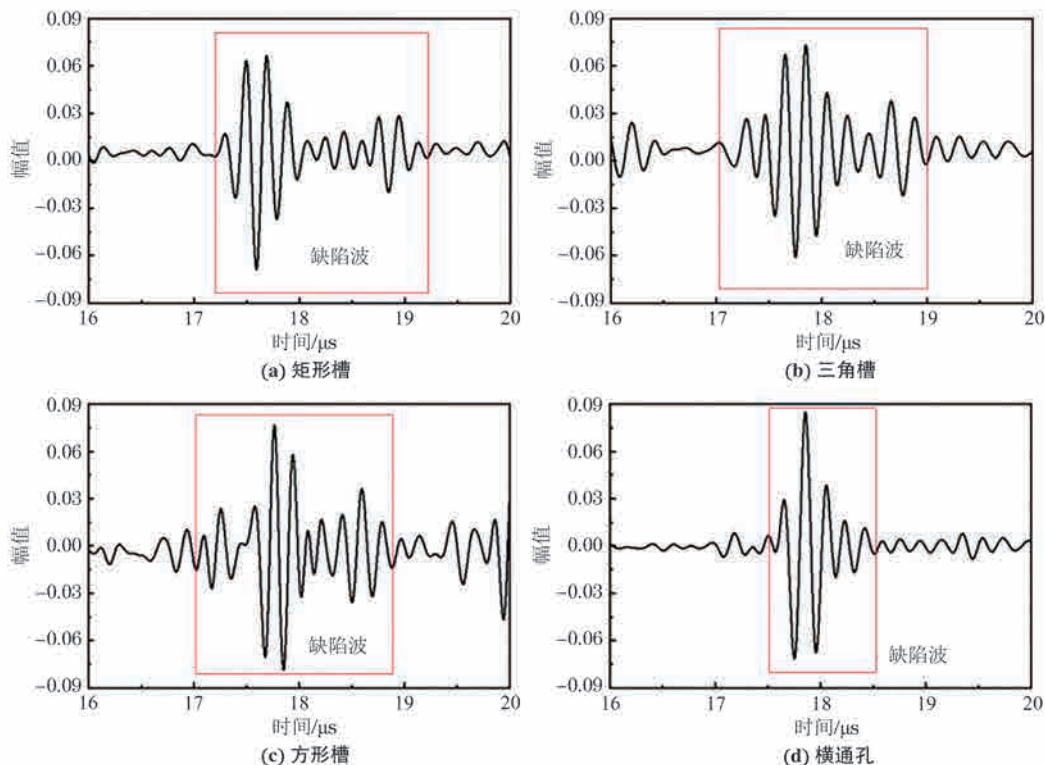


图2 4类缺陷的典型超声回波信号

2.3 特征提取

2.3.1 时域特征

对时域信号 $y(n)$ 进行希尔伯特变换可以获得其包络谱(见图3),图中黑色曲线为时域信号,蓝色曲线为包络线,其两端的时间差为持续时间。包络线的最高点处幅值为峰值,20%峰值与包络线有左、右两个交点,左交点与峰值点的时间差为上升时间,右交点与峰值点的时间差为下降时间,包络谱的偏斜和尖峭程度分别用偏度和峭度来描述。提取峰值、持续时间、上升时间、下降时间、偏度和峭度等6个描述时域波形特征的参量,分别记为 $T_1 \sim T_6$ 。

除上述波形特征外,还提取了描述信号幅值特征的10个时域参量,即均值、绝对均值、标准差、均方根值、方根幅值、峰值因子、脉冲因子、波形因子、裕度因子和能量,分别记为 $T_7 \sim T_{16}$ 。

2.2 信号采集

采用Olympus 5800型脉冲发射接收仪,选择横波斜入射法进行缺陷超声检测。探头中心频率为5 MHz, K 值为1.5,晶片尺寸为13 mm×13 mm(长×宽),采样频率为40 MHz。每隔2 mm采集一个信号,每个试样采集20个信号,共20个试样,故一共采集400个信号。4类缺陷的典型超声回波信号如图2所示(图中幅值数据无量纲)。

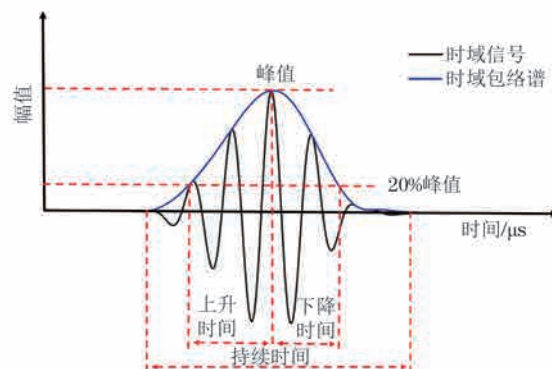


图3 时域信号曲线及其包络谱

2.3.2 频域特征

采用傅里叶变换将时域信号转换为频域,得到其幅度谱,提取频谱峰值、重心频率、频率方差和均方频率等表征频率分布和集中度的4个特征,分别记为 $F_{17} \sim F_{20}$,特征计算公式为

$$F_{17} = |Y(\omega)|_{\max} \quad (5)$$

$$F_{18} = \sum_{n=1}^N (\omega \cdot Y(\omega)) / \sum_{n=1}^N Y(\omega) \quad (6)$$

$$F_{19} = \sum_{n=1}^N ((\omega - F_{18})^2 \cdot Y(\omega)) / \sum_{n=1}^N Y(\omega) \quad (7)$$

$$F_{20} = \sum_{n=1}^N (\omega^2 \cdot Y(\omega)) / \sum_{n=1}^N Y(\omega) \quad (8)$$

式中: $Y(\omega)$ 为时域信号 $y(n)$ 的幅度谱; N 为采样点数; ω 为频率。

除以上特征外, 还提取了以下表示幅度谱能量分布的 12 个频域特征, 即偏度、峭度、均值、绝对均值、标准差、均方根值、方根幅值、峰值因子、脉冲因子、波形因子、裕度因子和能量, 分别记为 $F_{21} \sim F_{32}$ 。

2.3.3 信息熵特征

超声波与缺陷作用复杂, 仅通过时域和频域特征难以准确辨识缺陷类型, 此处结合信息熵理论来进一步提取多域特征。信息熵是评价系统状态不确定性的指标, 对于随机变量 $X = \{x_1, x_2, \dots, x_n\}$, 信息熵 $H(X)$ 可定义为^[18]

$$H(X) = -\sum_{i=1}^n p_i \log p_i \quad (9)$$

式中: p_i 为每一个事件 x_i 发生的概率。

信息熵 $H(X)$ 受概率分布的影响, 当系统的概率分布不确定性越大时, 系统的信息熵也越大。笔者提取了奇异谱熵、功率谱熵和小波包能量熵 3 个信息熵特征, 分别记为 $S_{33} \sim S_{35}$ 。

奇异谱熵能够定量描述信号时域能量的分布情况, 其计算方法为: 对缺陷时域回波信号进行空间重构得到轨迹矩阵, 再对轨迹矩阵进行奇异值分解获得对应的奇异值谱 $\{\lambda_1, \lambda_2, \dots, \lambda_m\}$, 将各个奇异值在奇异值谱中所占的比重 $p_i = l_i / \sum_{i=1}^m l_i$ 代入式(9)即可求出。

功率谱熵能够定量描述信号频域能量分布的复杂程度, 其计算方法为: 对回波信号进行离散傅里叶变换, 根据能量与功率的关系, 得到功率谱 $\{S_1, S_2, \dots, S_k\}$, 将各个频段的能量占总能量的比重 $p_i = S_i / \sum_{i=1}^k S_i$ 代入式(9)即可求出得到。

小波包能量熵能够定量描述信号在时频域的子频带能量分布特征, 其计算方法为: 对回波信号进行小波包分解, 获得各个节点上的分解序列 $\{E_1, E_2, \dots, E_j\}$, 将信号的各个节点能量占总能量的比重 $p_i = E_i / \sum_{i=1}^j E_i$ 代入式(9)即可求出^[19]。

2.4 XGBoost 模型建立

2.4.1 特征选择

将 4 类缺陷共 400 个信号样本按 7:3 的比例划分训练集和测试集, 每个信号样本提取 16 个时域、16 个频域和 3 个信息熵特征, 共计 35 个特征, 构建 400×35 的特征矩阵。对各类特征依次进行 Z-score 归一化处理后, 将所有特征输入到 XGBoost 模型中, 根据模型输出结果计算 SHAP 值并对特征进行量化排序, 35 个特征的平均绝对 SHAP 值(除部分数据已标出单位, 其余数据无量纲)如图 4 所示。图 4 中由上到下按特征贡献程度降序排列, 矩形颜色代表缺陷类型, 不同类型缺陷平均绝对 SHAP 值的和为该特征的总体平均绝对 SHAP 值, 特征对应的 SHAP 值越大, 说明其对缺陷识别的贡献越大。

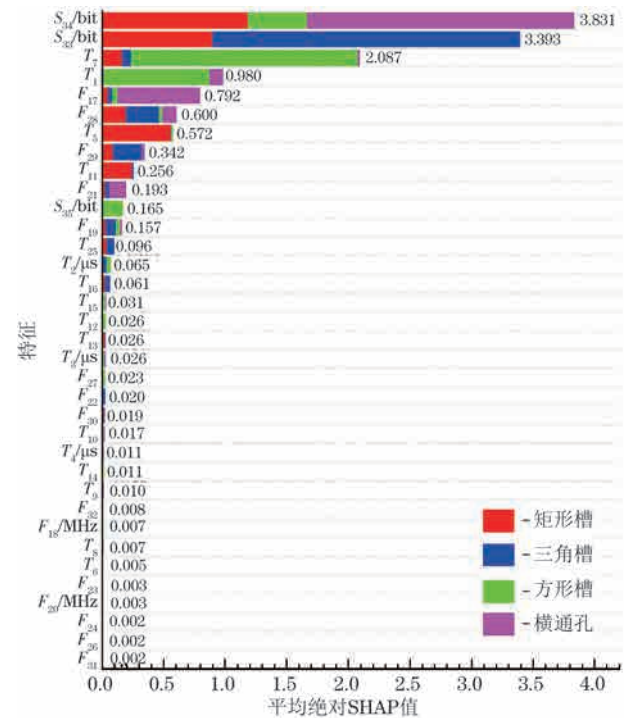


图 4 35 个特征的平均绝对 SHAP 值

由图 4 可知, 尽管 35 个特征对 4 类缺陷的识别均有贡献, 但贡献程度差异很大。笔者选择 SHAP 值累计占比为 91.0% 的前 8 个高贡献特征来构建特征子集, 按照贡献度由大至小分别为功率谱熵 (S_{34})、奇异谱熵 (S_{33})、均值 (T_7)、峰值 (T_1)、频谱峰值 (F_{17})、峰值因子 (F_{28})、偏度 (T_5) 和脉冲因子 (F_{29})。其中, 时域特征包括 T_1, T_5, T_7 , 频域特征包括 F_{17}, F_{28}, F_{29} , 信息熵特征包括 S_{33}, S_{34} 。

2.4.2 超参数优化

XGBoost 模型常见超参数包括 booster、learning_

rate、max_depth等7个参数,各个参数的含义及取值见表1。

表1 XGBoost模型各超参数的含义及取值

超参数	含义	取值
booster	弱学习器的类型	gbtree
learning_rate	学习率	0.05
max_depth	树最大深度	4
n_estimators	弱学习器的数量	150
gamma	最小损失函数值	0
reg_lambda	正则化权重项	1
reg_alpha	正则化权重项	1

3 试验结果与分析

3.1 XGBoost模型建立

将高贡献特征子集作为XGBoost模型的输入,采用准确率、精确率、召回率和 $F_{1-score}$ 等4个指标评价模型的分类性能。

准确率指在所有样本中,被预测正确的样本所占的比例。准确率数值越高,说明模型的分类效果越好。结果表明,采用XGBoost模型对矩形槽等4类缺陷120个测试集样本进行缺陷识别,正确识别了其中116个样本,平均准确率达96.7%。

精确率指在所有被预测为某类缺陷的样本中,实际为该缺陷的样本所占的比例。

召回率指在所有实际为某类缺陷的样本中,预测为该缺陷的样本所占的比例。

$F_{1-score}$ 是基于精确率与召回率的调和加权平均。精确率、召回率和 $F_{1-score}$ 的数值越高,分类效果越好^[20]。XGBoost模型识别不同类型缺陷的精确率、召回率和 $F_{1-score}$ 结果如表2所示。

表2 XGBoost模型识别不同类型缺陷的精确率、召回率和 $F_{1-score}$

缺陷类型	评价指标		
	精确率	召回率	$F_{1-score}$
横通孔	100	100	100
三角槽	100	96.7	98.3
方形槽	93.5	96.7	95.1
矩形槽	93.3	93.3	93.3

比较可知,横通孔的识别效果最佳,精确率、召回率和 $F_{1-score}$ 均达到100%,三角槽次之,方形槽略差,矩形槽最差,但就整体而言,XGBoost模型对4种缺陷类型均有较强的分类识别性能。

3.2 模型的可解释性分析

3.2.1 基于SHAP的可解释性分析

为了分析高贡献特征对各类型识别缺陷的贡献,针对4种缺陷,分别绘制了400个样本特征SHAP值的密度散点图,其结果如图5所示,其中每个点对应一个样本,图中横轴为SHAP值(未标出单位的数据为无量纲数据),其正负分别代表特征值与预测结果之间为正相关或负相关关系,SHAP数值越大,相关性越强。纵轴为高贡献特征,由上到下按特征的贡献程度降序排列。样本点的颜色代表特征值的大小,颜色越红则特征值越大。如果红色样本点对应SHAP正值,蓝色样本点对应SHAP负值,则样本为某类缺陷的可能性与特征值成正相关,即特征值较大的样本为某类缺陷可能性越大,反之成负相关^[21]。每一行表示一个特征,样本点在横向越分散,说明特征对模型识别的贡献越大,纵向较宽的区域表示有大量样本聚集。

以横通孔和矩形槽为例,对缺陷识别进行解释分析。由图5(a)可知,在高贡献特征中,对横通孔的识别有明显贡献的特征有5个,按贡献度由大至小分别为功率谱熵(S_{34})、频谱峰值(F_{17})、峰值因子(F_{28})、峰值(T_1)和脉冲因子(F_{29})。其中,特征 S_{34} , F_{17} , F_{28} 的蓝色样本点对应SHAP正值,表明特征值较小的样本为横通孔的可能性大。特征 T_1 的红色样本点对应SHAP正值,表明特征值较大的样本为横通孔的可能性大。而特征均值 T_7 、偏度 T_5 和奇异谱熵 S_{33} 的大多数样本点的SHAP值分布在0附近,说明其对识别横通孔的贡献小。

由图5(d)可知,所有高贡献特征对矩形槽的识别均有明显贡献,由大至小分别为功率谱熵(S_{34})、偏度(T_5)、奇异谱熵(S_{33})、均值(T_7)、峰值因子(F_{28})、脉冲因子(F_{29})、频谱峰值(F_{17})和峰值 T_1 。其中,特征 S_{34} , S_{33} , F_{17} 的红色样本点对应SHAP正值,表明特征值较大的样本为矩形槽的可能性大。特征 T_5 , F_{28} , T_1 的蓝色样本点对应SHAP正值,表明特征值较小的样本为矩形槽的可能性大。

3.2.2 超声与缺陷作用分析

以典型缺陷横通孔和矩形槽为例,分析高贡献特征中的峰值(T_1)、频谱峰值(F_{17})和功率谱熵(S_{34})的贡献差异及产生原因。横通孔和矩形槽的时域信号、幅度谱和9 μs时刻下的波场快照如图6所示(图中幅值数据无量纲)。

从图6(a), (d)中可以看出,横通孔的缺陷回波信

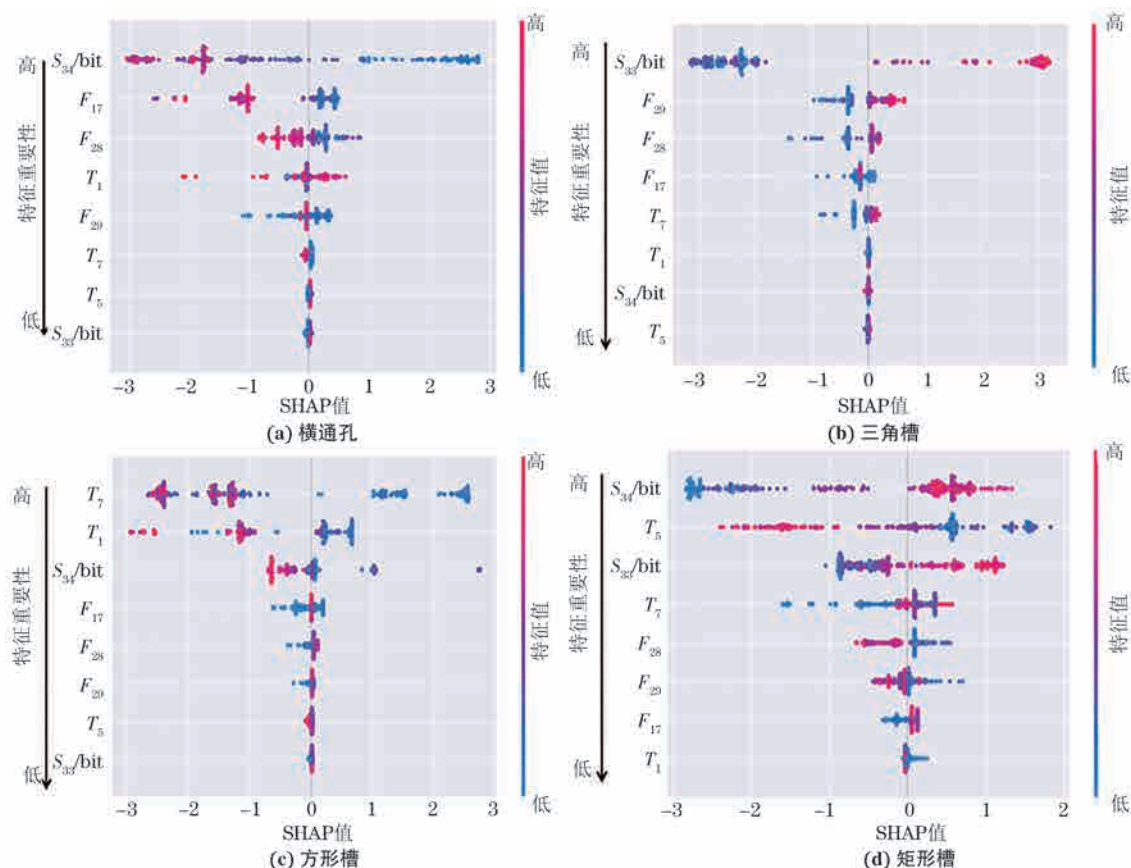


图5 400个样本特征的 SHAP 值密度散点图

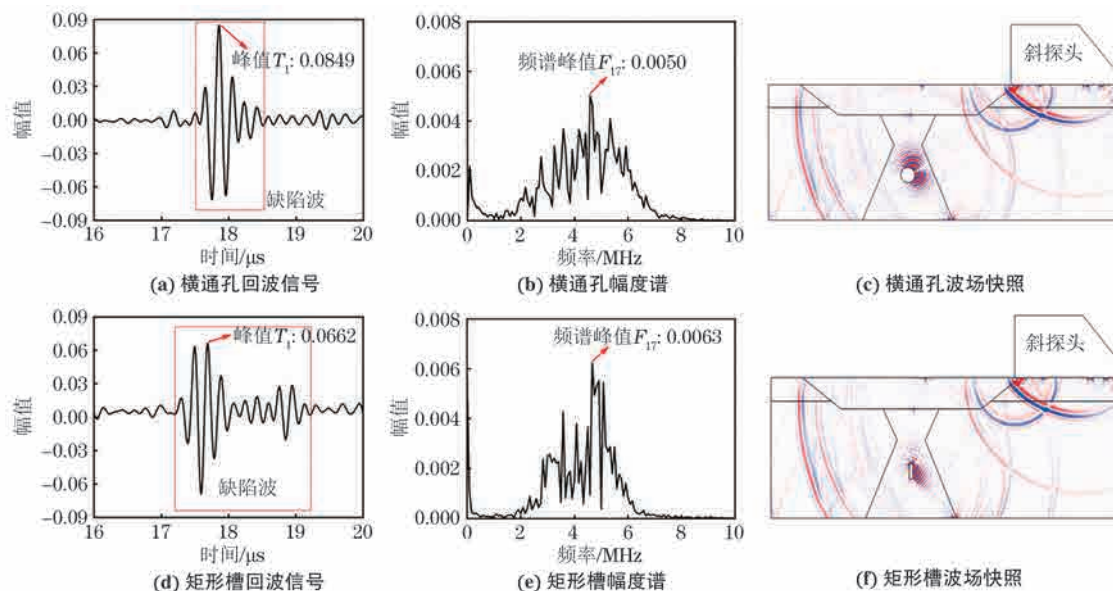


图6 横通空孔和矩形槽的时域信号、幅度谱和9 μs时刻下的超声声场

号比较集中、持续时间短,峰值 T_1 为0.084 9。矩形槽的缺陷回波包括主回波信号及较明显的跟随信号,持续时间长,峰值 T_1 为0.066 2。这是因为横通孔与超声波的作用面是一个连续光滑的反射曲面,探头能够接收到较强的回波信号,分析图6(c)的波场快照也可以得出类似结论。矩形槽有2个平直反射面和1个尖锐

的棱角,其反射波方向性强,探头只能接收到一部分反射信号,能量略低。从图6(b),(e)的幅度谱可以看出,横通孔幅度谱的尖锐峰少,谱峰间隔较均匀,能量分布相对分散,主要分布范围为2.16 MHz~6.34 MHz,频谱峰值 F_{17} 为0.005 0,功率谱熵 S_{34} 为1.37 bit(位)。矩形槽幅度谱的尖锐峰多,跳跃性强且随机,能量主

要分布范围在 2.91 MHz~5.75 MHz, 频谱峰值 F_{17} 为 0.006 3, 其对应的功率谱熵 S_{34} 为 2.27 bit, 明显高于横通孔。

从本质上说, 超声波与不同形状反射体之间的作用会引起缺陷回波信号在波形、幅值等方面的差异, 由此带来频谱、信息熵等多域特征的变化。分析超声波与缺陷的作用机制有助于理解机器学习模型的预测效果、特征贡献差异及产生原因。

4 结论

采用集成学习模型 XGBoost 作为分类器, 利用 SHAP 方法进行特征优选和可解释性分析, 构建了具有可解释性的焊缝缺陷类型超声识别 XGBoost 模型, 主要结论如下。

(1) 对矩形槽、三角槽、方形槽、横通孔等 4 类缺陷提取时域、频域及时频域的 35 个特征, 通过计算 SHAP 值量化特征贡献, 优选出 8 个高贡献特征。

(2) 将高贡献特征作为 XGBoost 模型的输入, 实现了仅用较少特征即可获得较高的机器学习分类识别性能。

(3) 结合 SHAP、超声波与缺陷之间的关系分析了特征贡献差异及产生原因, 提出的 XGBoost 机器学习模型缺陷类型识别方法具有较强的可解释性。

参考文献:

- [1] 张书宁, 刘小川, 马志远, 等. 多相非均质 GFRP 材料的超声检测及其声学特性[J]. 无损检测, 2022, 44(12): 6-11.
- [2] SUN H B, RAMUHALI P, JACOB R E. Machine learning for ultrasonic nondestructive examination of welding defects: a systematic review[J]. Ultrasonics, 2023, 127: 106854.
- [3] VIRKKUNEN I, KOSKINEN T, JESSEN-JUHLER O, et al. Augmented ultrasonic data for machine learning[J]. Journal of Nondestructive Evaluation, 2021, 40(1): 4.
- [4] LIU J, XU G C, REN L, et al. Defect intelligent identification in resistance spot welding ultrasonic detection based on wavelet packet and neural network[J]. The International Journal of Advanced Manufacturing Technology, 2017, 90(9): 2581-2588.
- [5] 原赛男, 于闯, 陈少敏. 飞机典型结构冲击类损伤的智能检测和评估[J]. 无损检测, 2023, 45(5): 51-55.
- [6] LV G L, GUO S F, CHEN D, et al. Laser ultrasonics and machine learning for automatic defect detection in metallic components[J]. NDT & E International, 2023, 133: 102752.
- [7] CANTERO-CHINCHILLA S, WILCOX P D, CROXFORD A J. Deep learning in automated ultrasonic NDE-Developments, axioms and opportunities[J]. NDT & E International, 2022, 131: 102703.
- [8] GUMUS M, KIRAN M S. Crude oil price forecasting using XGBoost[C]//2017 International Conference on Computer Science and Engineering (UBMK). Antalya, Turkey: IEEE, 2017.
- [9] 胡宏伟, 张婕, 彭刚, 等. 基于 LBP-KPCA 特征提取的焊缝超声检测缺陷分类方法[J]. 焊接学报, 2019, 40(6): 34-39, 162.
- [10] 蒲天骄, 乔骥, 赵紫璇, 等. 面向电力系统智能分析的机器学习可解释性方法研究(一): 基本概念与框架[J]. 中国电机工程学报, 2023, 43(18): 7010-7030.
- [11] 孙洪宇, 彭丽莎, 屈凯峰, 等. 机器学习在复合绝缘子缺陷超声检测中的应用与展望[J]. 无损检测, 2021, 43(5): 58-63.
- [12] 金东镇, 郭城楠, 彭芳, 等. 从 SHAP 到概率——可解释性机器学习在糖尿病视网膜病变靶向脂质组学研究中的应用[J]. 中国卫生统计, 2023, 40(4): 511-515.
- [13] 赵婉辰, 郑晨, 肖斌, 等. 基于 Bayesian 采样主动机器学习模型的 6061 铝合金成分精细优化[J]. 金属学报, 2021, 57(6): 797-810.
- [14] 刘巧红, 马雨生, 蔡雨晨. 基于 XGBoost 算法的糖尿病分类预测模型及应用[J]. 现代仪器与医疗, 2023, 29(4): 1-6, 11.
- [15] GHOLIZADEH S, LEMAN Z, BAHARUDIN B T H T. State-of-the-art ensemble learning and unsupervised learning in fatigue crack recognition of glass fiber reinforced polyester composite (GFRP) using acoustic emission[J]. Ultrasonics, 2023, 132: 106998.
- [16] 孟昭亮, 张泽涛, 杨媛, 等. 改进的 XGBoost 杂散电流预测及可解释模型[J]. 激光与光电子学进展, 2022, 59(12): 1215011.
- [17] 陈蒙, 王华. 地震动强度参数估计的可解释性与不确定度机器学习模型[J]. 地球物理学报, 2022, 65(9): 3386-3404.
- [18] SHANNON C E. A mathematical theory of communication[J]. ACM SIGMOBILE Mobile Computing and Communications Review, 2001, 5(1): 3-55.
- [19] 付朕. 矿用带式输送机托辊远程故障诊断系统[D]. 徐州: 中国矿业大学, 2020.
- [20] PASADAS D J, BARZEGAR M, RIBEIRO A L, et al. Guided waves based debonding classification in lap-joints using modified Fisher discriminant criterion[J]. NDT & E International, 2023, 137: 102831.
- [21] 李扬. 基于随机森林模型解释的越野路面识别算法研究[D]. 长春: 吉林大学, 2022.