

DOI: 10.13382/j.jemi.B2306964

随机森林算法在超声缺陷识别中的应用研究*

魏新园^{1,2} 周京欢¹ 钱牧云¹ 李丹¹ 黄三傲¹

(1.安徽省智能破拆装备工程实验室 马鞍山 243032; 2.安徽工业大学电气与信息工程学院 马鞍山 243032)

摘要: 超声波检测是一种常见的钢材缺陷检测方法,通过机器学习算法建立分类模型能够实现有效的缺陷识别。神经网络是目前最常采用的一种算法,但存在模型结构复杂且需要大量训练数据的问题。对此,提出一种基于随机森林的超声缺陷识别方法,能够实现对缺陷类型的智能、准确识别,以解决模型结构复杂和训练数据需求大的问题。首先对方体试件中的不同形状、尺寸和深度的缺陷进行超声检测实验,基于实验数据利用随机森林算法建立超声缺陷识别模型;进而对模型的缺陷识别效果进行分析,并与支持向量机、K-近邻分类算法、AdaBoosting算法和卷积神经网络比对分析缺陷识别效果;然后利用验证试件进行缺陷识别验证实验,以进一步验证所建立缺陷识别模型的有效性。结果表明,所提缺陷识别方法相比其他算法具有最高的准确率,验证实验中缺陷分类准确率达到94.6%。

关键词: 随机森林; 超声探伤; 缺陷识别; 不同形状和尺寸; 分类准确率

中图分类号: TN05; TG115.28 **文献标识码:** A **国家标准学科分类代码:** 520.60

Research on the application of random forest algorithm in ultrasonic defect recognition

Wei Xinyuan^{1,2} Zhou Jinghuan¹ Qian Muyun¹ Li Dan¹ Huang San'ao¹

(1. Anhui Province Engineering Laboratory of Intelligent Demolition Equipment, Ma'anshan 243032, China;

2. School of Electrical and Information Engineering, Anhui University of Technology, Ma'anshan 243032, China)

Abstract: Ultrasonic detection is a common method of steel defect detection. The classification model established by machine learning algorithm can realize effective defect identification. Neural network is the most commonly used algorithm at present, but it has the problem of complex model structure and large amount of training data. In this paper, an ultrasonic defect recognition method based on random forest is proposed, which can realize intelligent and accurate identification of defect types to solve the problems of complex model structure and large training data requirements. Firstly, ultrasonic detection experiments were carried out for defects of different shapes, sizes and depths in the specimen. Based on the experimental data, an ultrasonic defect recognition model was established using random forest algorithm. Then, the defect recognition effect of the model is analyzed, and compared with support vector machine, K-nearest neighbor classification algorithm, AdaBoosting algorithm and convolutional neural network. Then the defect identification verification experiment is carried out with the verification specimen to further verify the validity of the established defect identification model. The results show that the proposed method has the highest accuracy compared with other algorithms, and the accuracy of defect classification reaches 94.6% in the verification experiment.

Keywords: ultrasonic defect detection; defect identification; random forest; different shapes and sizes; classification accuracy

收稿日期:2023-10-16 Received Date: 2023-10-16

* 基金项目:安徽省重点研究与开发计划项目(2022f04020005)、安徽省高等学校科研研究重点项目(2022AH050313)、安徽省智能破拆装备工程实验室开放基金(2022f04020005)项目资助

0 引言

钢材已成为现代工业生产不可缺少的原材料,但钢材在生产过程中有可能在材料内部形成缺陷。这些缺陷的存在会影响最终产品的性能和质量,因此对钢材内部存在的缺陷进行检测十分重要。超声波探伤是检测钢材内部质量的重要手段之一。随着用户对钢材内部质量的重视,对钢材进行超声波探伤检测的需求不断增加。研究人员通过不同的方法和手段,不断改进和完善超声波探伤技术。2018年张勇等^[1]利用超声水浸聚焦检测方法提出了一种对中厚板结构件激光焊缝内部缺陷的检测与识别方法。2020年赵一凡等^[2]验证了超声波检测技术对车辆金属材料中的铝制材料零部件探伤的准确性与合理性。2021年高翔等^[3]介绍了超厚壁无缝钢管的接触法和水浸法检测方法,分析了超厚壁无缝钢管的声束覆盖区域。2022年杨林等^[4]设计了一种基于 KUKA 机器人的超声波自动探伤系统,实现了对挖掘机大臂焊缝超声波自动探伤的功能。2022年杨秀文等^[5]利用超声波对巴氏合金和轴承衬体结合面处的缺陷实现对脱壳缺陷的判定。这些研究成果对于提高产品的性能和质量具有积极意义。

随着机器学习和人工智能的发展,研究者们逐步利用算法对超声检测数据进行缺陷识别。2016年 Filho 等^[6]则提出了通过不同变换方法(离散傅里叶变换、小波变换和余弦变换)提取超声信号特征,并使用 BP 网络进行训练分类检测。2016年施成龙等^[7]使用小波包变换提取缺陷特征信息,通过深度信念网络进行分类识别,准确率达到 98.83%。2017年 Meng 等^[8]提出了分层组合小波变换系数,利用 CNN 对不同的特征提取方法进行分类,准确率达到 98.15%。2018年 Hou 等^[9]提出了一种基于深度卷积神经网络的模型,从 X 射线图像中直接提取深度特征,并达到了 97.2% 的准确率。2019年侯怀书等^[10]利用短时傅里叶变换方法,对超声检测技术在金属管道对接焊缝缺陷中的缺陷类别识别问题进行了探究,利用短时傅里叶变换实现金属管道焊缝中的 4 种缺陷类型的识别。2019年王永利等^[11]提出一种基于深度学习的 PCB 缺陷识别算法,对 10 类 PCB 缺陷可以得到 96.67% 的识别准确率。2020年鲍光海等^[12]运用深度学习针对汽车高度调节器提出了一种基于改进型卷积网络的智能检测方法,识别准确率在 99% 以上。2022年李丹等^[13]提出了一种轻量级部分深度混合可分离网络用于带钢表面缺陷识别,识别准确率达到 99.78%。2022年高子洋等^[14]使用卷积神经网络对超声相控阵缺陷信号进行识别,并通过选择合适的 CNN 模型提高了缺陷识别的准确率。2023年陈彦杰等^[15]设计了一种具有分支

结构的可融合的二值卷积神经网络,通过降低参数和计算量,实现了 98.11% 的准确率。2024年陈赛赛等^[16]提出基于双阶段支持向量机(support vector machine, SVM)与小波核扩散的轴承复合故障分类方法,比 SVM 分类准确率平均提升 7.5%。

这些研究为超声缺陷检测方法提供了宝贵的经验,但当前研究大多基于神经网络算法实现缺陷类型识别,存在模型复杂、需要数据量大等问题,且需要在预处理阶段都使用统计或信号处理方法来提取特征以提高识别准确率。

集成学习是一种旨在将弱学习器转化为强学习器的机器学习方法,常见的集成学习方法包括随机森林算法、Boosting 算法等。随机森林算法具有训练速度快、抗噪声能力强、高准确率和不易出现过拟合优点。因此,随机森林算法已成为一种重要的缺陷分类和故障诊断工具,并广泛应用于各个行业。如 2015年胡峻峰等^[17]使用随机森林分类器对板材表面缺陷图像进行识别分类,准确率达到 95.67%。2016年 Patel 等^[18]使用随机森林方法对低碳钢缺陷进行诊断,识别准确率高达 95%。2017年 Wang 等^[19]针对钢表面分布式缺陷提出了一种改进的随机森林算法,并通过实验证明了其有效性。2018年刘传泽等^[20]提出了一种基于随机森林算法的分类模型,实现了纤维板表面大刨花、油污和杂物的快速准确识别,正确率达到了 99%。2022年李冬等^[21]提出了基于多重分形理论和改进随机森林的模式识别方法,对故障识别的正确率达到 100%。2023年唐宇等^[22]提出了一种基于分数阶变异麻雀优化随机森林的软件缺陷预测算法,具有更高的预测精度和稳定性。

这些研究大多需要对待分类原始数据进行特征提取,进而建立分类模型。而本文实验研究发现,基于超声波原始探伤信号建立随机森林缺陷识别模型,同样能够取得高准确率。因此,本文提出一种基于随机森林的超声缺陷识别方法,直接使用原始超声信号构建缺陷识别模型,省去特征提取从而简化建模过程。首先,对试件中不同大小、形状和深度的缺陷进行超声检测,利用随机森林算法构建缺陷类型识别模型;进而,对模型的缺陷识别效果进行分析,并与支持向量机、K-近邻分类算法、AdaBoosting 算法和卷积神经网络比对其缺陷类型识别准确率;最后,利用验证试件验证所建立模型的有效性。

1 超声缺陷识别建模方法

1.1 基于随机森林的缺陷识别方法

随机森林利用多棵决策树对样本进行训练并预测。设对钢材进行超声检测的原始数据为:

$$\mathbf{H} = \{(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2), \dots, (\mathbf{x}_i, \mathbf{y}_i), \dots, (\mathbf{x}_m, \mathbf{y}_m)\} \quad (1)$$

式中: \mathbf{x}_i 表示超声原始数据, 其维度为 M ; \mathbf{y}_i 表示缺陷类型, m 是样本数据个数。

本文在训练随机森林缺陷识别模型时, 首先用 bootstrap 方法有放回从超声信号数据集 \mathbf{H} 中随机抽取 d 个样本, 并构建 d 个决策树; 然后每棵树根据“Gini”指数进行分裂, 直到无法分裂为止, 基尼系数定义如式(2)所示。

$$\text{Gini}(p) = 1 - \sum_{i=1}^r p_i^2 \quad (2)$$

式中: p 为数据集的概率分布, r 为数据集中的类别数量。

重复上述过程, 生成多棵分类树, 建立随机森林模型。在使用该随机森林模型对超声检测数据进行分类预测时, 随机森林中的每个决策树作为一个预测器对超声信号数据进行分类, 进而通过聚合随机森林中所有决策树的分类结果, 得到最终的缺陷识别结果。

1.2 基于 SVM 的缺陷识别方法

SVM 通过寻找具有最大间隔的超平面来将不同类别的样本进行划分, 从而使得不同类别样本之间的距离尽可能大。对于一个二分类问题, 设数据集为:

$$\mathbf{D} = \{(\mathbf{a}_1, \mathbf{b}_1), (\mathbf{a}_2, \mathbf{b}_2), \dots, (\mathbf{a}_i, \mathbf{b}_i), \dots, (\mathbf{a}_n, \mathbf{b}_n)\} \quad (3)$$

式中: \mathbf{a}_i 是一个 N 维向量, 类别标签 \mathbf{b}_i 的取值为 -1 或 1, n 是样本数据个数。则任意超平面可以被表示为:

$$\mathbf{w}^T \mathbf{a} + c = 0 \quad (4)$$

式中: \mathbf{w} 代表法向量, 决定了超平面的方向, c 代表位移项, 决定了超平面与原点之间的距离。假设超平面 (\mathbf{w}, c) 能将训练样本正确分类, 即对于 $(\mathbf{a}_i, \mathbf{b}_i) \in \mathbf{D}$, 都有:

$$\begin{cases} \mathbf{w}^T \mathbf{a}_i + c \geq +1, \mathbf{b}_i = 1 \\ \mathbf{w}^T \mathbf{a}_i + c \leq -1, \mathbf{b}_i = -1 \end{cases} \quad (5)$$

则两类间的间隔距离为:

$$\gamma = \frac{2}{\|\mathbf{w}\|} \quad (6)$$

引入拉格朗日乘子, 将求解 γ 最大值问题转换为一个求解最小化目标函数和满足约束条件的问题; 求解拉格朗日乘子, 求得最优的超平面参数 (\mathbf{w}, c) 。

本文进行超声探伤识别时, 每次选择一个类作为正样本, 负样本只用选其余的一个类, 这样对于具有 g 个类别的多分类问题, 就构建 g 个二分类模型。对于每对类别只使用属于这两个类别的训练样本训练模型。通过计算样本与超平面之间的距离或使用核函数来找到最佳的分类超平面。对于新的待分类数据, 将待分类样本分别用多个二分类模型进行预测, 并选择获得最多投票的类别作为最终的预测类别。

1.3 基于 K-近邻的缺陷识别方法

K-近邻算法的原理基于通过测量样本间的距离来进行分类。对于超声信号数据集为 \mathbf{H} , 通过使用 K-近邻算法, 将每个类别作为一个分类标签, 并将待分类信号与训练集中的样本进行距离计算。通过使用合适的距离度量方法衡量待分类信号与各个类别的相似度, 如欧氏距离或曼哈顿距离。本文中选用的是欧几里德距离, 即在二维平面中, 设两点 $A(x_p, y_p)$ 和 $B(x_q, y_q)$, 则欧几里德距离用式(7)表示为:

$$d = \sqrt{(x_p - x_q)^2 + (y_p - y_q)^2} \quad (7)$$

选择 k 个与其最近的样本作为邻居, 通过投票机制或加权计算来确定其所属类别。在确定邻居时, 选择不同的 k 值会产生不同效果。较小的 k 值可能会导致对噪声敏感, 而较大的 k 值可能会导致分类错误的增加。

1.4 基于 AdaBoosting 的缺陷识别方法

AdaBoosting 是一种用来提高弱分类算法准确度的方法, 这种方法通过构造一个预测函数系列, 然后以一定的方式将他们组合成一个预测函数。对于超声信号数据集 \mathbf{H} , 在训练的初始阶段每一个数据样本的权值如式(8)所示。

$$\varphi(t) = (\mu_1, \mu_2, \dots, \mu_t) = \left(\frac{1}{t}, \frac{1}{t}, \dots, \frac{1}{t}\right) \quad (8)$$

式中: $\varphi(t)$ 表示在数据集中的权值分布, μ_t 为每一个数据样本所分配的权值。

将数据集 \mathbf{H} 中的数据样本平均划分为 s 个子集, 并在初始阶段为每个子集赋予相同的权重; 进而通过训练这 s 个子集来构建基分类器, 计算每个基分类器的错误率; 然后根据错误率, 调整基分类器的权重, 并进行新一轮迭代, 这个迭代过程会持续直到处理完所有的数据样本; 最后, 根据基分类器的错误率, 计算出它们在集成中所占的权重。综合所有基分类器的结果, 得到最终分类结果。

1.5 基于卷积神经网络的缺陷识别方法

卷积神经网络是一种参数量较小, 容易训练, 不易过拟合的神经网络模型。它的基本结构主要包括输入层、卷积层、池化层和全连接层。

1) 输入层

本文中, 输入层的输入是超声信号数据集 \mathbf{H} 。

2) 卷积层

采用卷积核对输入数据进行卷积计算, 提取和学习数据集中的特征, 在计算时, 卷积核与输入数据的局部区域进行卷积操作, 再对所有结果进行加权求和并加上偏差量。卷积计算公式如式(9)所示。

$$F(x) = \sum_{i=1}^L x(i) \Psi(i) \quad (9)$$

式中: $x(i)$ 表示数据的第 i 个特征, L 表示特征数量, $\Psi(i)$ 表示对应的权重, $F(x)$ 表示 $x(i)$ 和 $\Psi(i)$ 相乘求和的函数。

3) 池化层

进行特征筛选, 通过下采样的方式减少特征的空间大小, 以缩小模型规模和计算量, 并且提高处理速度。本文选用的是最大池化层, 最大池化层挑选目标空间中最大值为特征值输出。

4) 全连接层

将池化层输出特征通过重新非线性映射, 转成与类别数目相同的向量的形式, 实现不同缺陷类别的分类。同时, 全连接层还可以通过调整权重参数的大小来提取特征的重要性, 使网络更加关注对于正确分类最有贡献的特征。其计算公式如式 (10) 所示。

$$x^i = f(\Psi \cdot x^{i-1} + z^i) \tag{10}$$

式中: Ψ 是权重矩阵, x 是特征向量, z 是偏置向量, $f(\cdot)$ 是激活函数。

本文采用 Softmax 逻辑回归函数对数据特征进行分类预测。

2 缺陷检测实验

2.1 实验装置

超声缺陷检测实验装置包括数字式超声波探伤仪 (堀扬 KY-CT350B 型号)、超声探伤数据采集软件、探头和试件 (304 不锈钢材料), 如图 1 所示。本实验中使用的探头是超声波单晶直探头, 频率为 2.5 MHz。使用连接线将超声探头与探伤仪连接, 通过探伤仪可获得超声缺陷检测的原始信号。

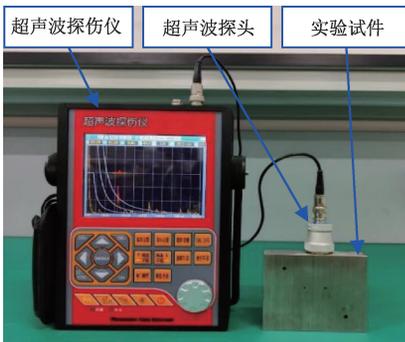


图 1 探伤实验系统

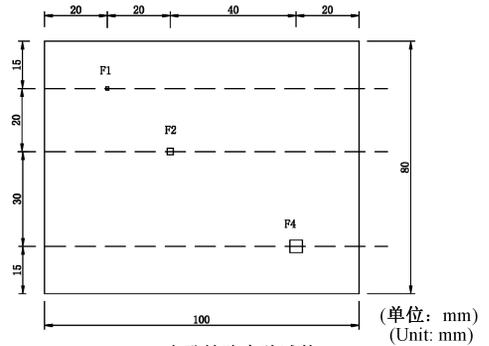
Fig. 1 Test system for flaw detection

在本文中, 设置了 3 个试件用于实验, 包含了 10 类不同的缺陷类型, 具体情况如表 1 所示。试件尺寸及各类缺陷的位置分布如图 2 所示。

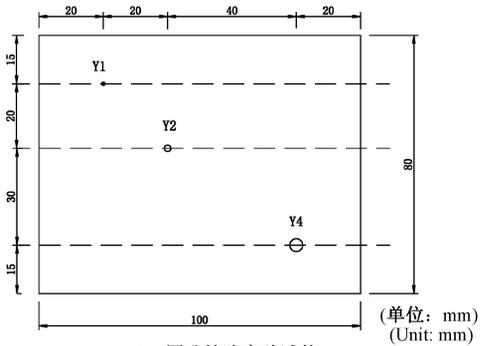
表 1 试件及其所含缺陷情况

Table 1 Details of the specimen and the types of defects contained therein

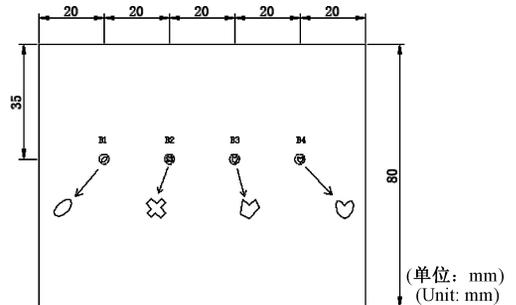
| 编号 | 形状 | 尺寸/mm | 深度/mm | 所处试件 |
|-----|-------|---------|-------------|-----------|
| F-1 | 方孔 | 边长 1 | 15、20、65、80 | 方孔缺陷实验试件 |
| F-2 | 方孔 | 边长 2 | 35、40、45、60 | |
| F-4 | 方孔 | 边长 4 | 15、20、65、80 | |
| Y-1 | 圆孔 | 直径 1 | 15、20、65、80 | 圆孔缺陷实验试件 |
| Y-2 | 圆孔 | 直径 2 | 35、40、45、60 | |
| Y-4 | 圆孔 | 直径 4 | 15、20、65、80 | |
| B-1 | 不规则-1 | 外接圆直径 2 | 35、45 | 不规则缺陷实验试件 |
| B-2 | 不规则-2 | 外接圆直径 2 | 35、45 | |
| B-3 | 不规则-3 | 外接圆直径 2 | 35、45 | |
| B-4 | 不规则-4 | 外接圆直径 2 | 35、45 | |



(a) 方孔缺陷实验试件
(a) Square hole defect test specimen



(b) 圆孔缺陷实验试件
(b) Round hole defect test specimen



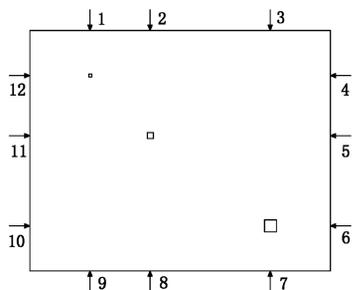
(c) 不规则缺陷实验试件
(c) Irregular defect test specimen

图 2 缺陷实验试件尺寸及位置示意图

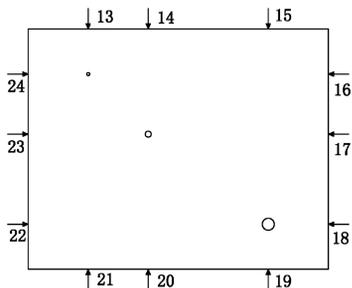
Fig. 2 Schematic diagram of size and position of defect specimen

2.2 实验过程

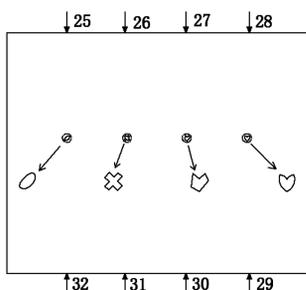
为了获取原始探伤数据,使用如图 1 所示的实验装置对图 2 中试件进行实验。考虑到缺陷所处深度对超声波探伤信号的影响,使用超声波探头按照图 3 中箭头 1~32 标注位置进行超声缺陷检测,得到受缺陷形状、尺寸和深度 3 个因素影响的共 32 类不同超声缺陷检测信号。每个箭头位置分别进行 10 次重复检测,共测得 320 组数据,记为 K1。



(a) 方孔缺陷实验试件
(a) Square hole defect test specimen



(b) 圆孔缺陷实验试件
(b) Round hole defect test specimen



(c) 规则缺陷实验试件
(c) Irregular defect test specimen

图 3 超声缺陷检测缺陷实验试件位置示意图

Fig. 3 Position diagram of ultrasonic defect test specimen

为了验证所建缺陷识别模型的效果,对图 3 中箭头 1~32 位置分别进行 2 次额外的重复检测,得到 64 组数据,记为 K2。各批次实验具体情况如表 2 所示。

2.3 实验数据分析

对实验获得的超声检测原始信号进行分析。首先将超声波探头放在图 3 中箭头 2 和 3 之间进行探测,可得到无缺陷时的超声波检测信号,如图 4 所示。

表 2 各批次实验说明

Table 2 Illustrates the experiments

| 实验批次 | 编号 | 数据组数 | 说明 |
|--------|----|------|---------------------------|
| 第 1 批次 | K1 | 320 | 对图 3 中 32 类缺陷各进行 10 次重复检测 |
| 第 2 批次 | K2 | 64 | 对图 3 中 32 类缺陷各进行 2 次重复检测 |

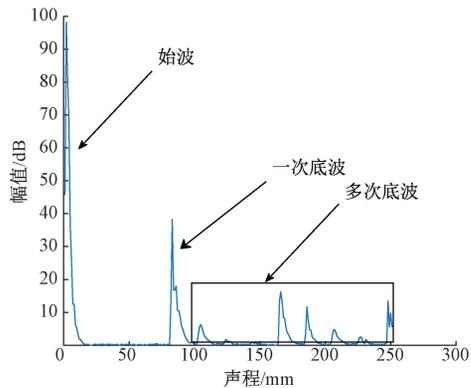


图 4 无缺陷超声波检测信号

Fig. 4 Ultrasonic detection signal without defect

通过分析发现,超声检测信号主要包含 3 部分:始波、一次底波和多次底波。始波是超声波传到试件检测表面发生反射形成,一次底波是超声波第一次传到试件底部形成,多次底波是一次底波回传到探测面时,又会产生反射,又会向底面传播,如此来回反射形成。对图 4 中超声信号进一步分析可知始波波峰的横坐标是(3,0),一次底波波峰的横坐标是(83,0),两个峰值距离为 80 mm,与实际试件尺寸(高 80 mm)一致。

绘制图 3 中箭头 1、2 和 13 处的超声缺陷检测信号图,如图 5~7 所示,以进一步分析不同缺陷形状、尺寸和深度的影响。

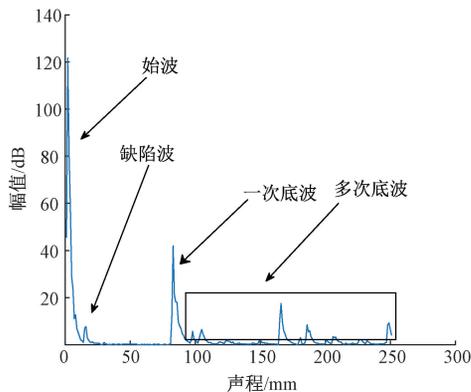


图 5 箭头 1 处超声波信号

Fig. 5 Ultrasonic signal at arrow 1

与图 4 中超声检测信号对比可发现,图 5~7 中出现了缺陷波。且图 5~7 中缺陷波峰值出现的位置与实际

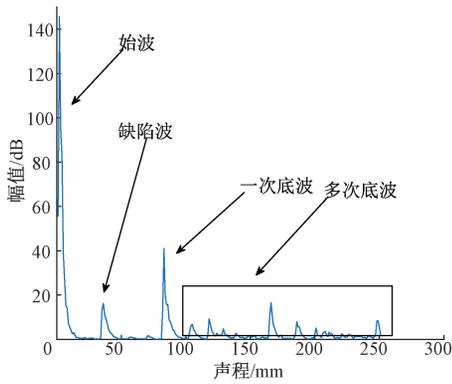


图 6 箭头 2 处超声波信号

Fig. 6 Ultrasonic signal at arrow 2

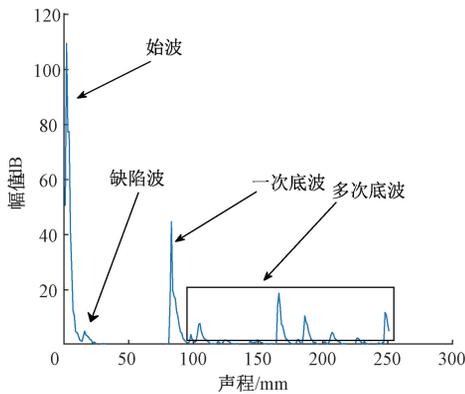


图 7 箭头 13 处超声波信号

Fig. 7 Ultrasonic signal at arrow 13

试件中缺陷位置一致,例如图 5 中缺陷波的峰值与始波信号的峰值距离为 15 mm,与实际缺陷位置至试件检测表面的距离一致。对比图 5 和 6 可发现,图 6 中的缺陷波的波峰峰值为 16.25 dB,高于图 5 中的缺陷波的波峰峰值 7.75 dB,符合相同缺陷形状尺寸越大缺陷波的波峰峰值越大的规律。对比图 5 和 7 可发现,图 5 中的方形缺陷波的波峰峰值为 7.75 dB,高于图 7 中圆形缺陷波的波峰峰值 5 dB。这是由于当缺陷大小相同时,方形缺陷反射的超声波信号比圆形缺陷多^[23]。以上分析结果说明本文的实验结果与理论一致。

3 建模与效果分析

3.1 缺陷识别模型的建立

将 K1 数据集按照 8 : 2 的比例划分为训练集和测试集,此时训练集数据组数和测试集数据组数分别为 256 和 64。使用训练集数据建立识别模型,并使用测试集分析模型效果。

随机森林算法缺陷识别模型的建立:使用 MATLAB

自带函数 `TreeBagger` 进行随机森林算法建模。用参数 `NumTrees` 设置随机森林中的决策树棵树为 100,并将参数 `OOBPrediction` 设置为 `On` 来计算袋外误差,并基于实验数据对模型结构参数进行了优化。

支持向量机算法缺陷识别模型的建立:使用 MATLAB 自带函数 `fitcecoc` 进行支持向量机算法建模,并用参数 `Learners` 设置支持向量机中的核函数为 `linear`, `C` 和 `Gamma` 值通过交叉验证的方式,在一定范围内进行搜索并选择最佳的参数值。

K-近邻算法缺陷识别模型的建立:建模时使用 MATLAB 自带函数 `fitcknn` 进行 K-近邻算法建模,并通过交叉验证的方式,在一定范围内进行搜索并选择最佳的 `k` 值,并将参数 `Distance` 设置为 `Euclidean` 来指定用于计算邻居之间距离的度量方法为欧氏距离。

AdaBoosting 算法缺陷识别模型的建立:建模时使用 MATLAB 自带函数 `fitcensemble` 进行 AdaBoosting 算法建模,使用决策树作为基分类器,使用参数 `Method` 设置集成方法为 `AdaBoostM1`,使用参数 `NumLearningCycles` 设置学习周期为 100。

卷积神经网络算法缺陷识别模型的建立:建模时定义卷积神经网络结构,设置训练参数。在本文中,卷积神经网络结构包括输入层、卷积层(神经元数量为 16)、激活层、最大池化层、全连接层(神经元数量为 32)、归一化层和分类层;设置最大训练轮数为 500、初始学习率为 0.001、学习率规则的下降因子为 0.1。

3.2 缺陷识别效果分析

基于 K1 数据集的模型识别准确率如表 3 所示。此外,为了进一步验证所建立模型的有效性,使用 K1 数据集所建立的模型对 K2 数据集进行预测,缺陷识别准确率如表 3 所示。(表 3 中各算法的识别准确率为经过 15 次运行结果的平均值)。

表 3 K1 交叉验证和 K2 验证结果

Table 3 Results of K1 cross-validation and K2 verification

| 算法类型 | K1 交叉验证结果 | K2 验证结果 |
|-------------|-----------|---------|
| | 识别准确率 | 识别准确率 |
| 随机森林 | 0.994 | 1.000 |
| 支持向量机 | 0.948 | 0.860 |
| K-近邻算法 | 0.837 | 0.928 |
| AdaBoosting | 0.952 | 0.969 |
| 卷积神经网络 | 0.996 | 1.000 |

通过表 3 的结果可以得出:基于 K1 数据集的识别结果中,随机森林算法和卷积神经网络算法的预测准确率较高;基于 K2 数据集的识别结果中,随机森林算法和卷积神经网络之间的预测准确率都是 100%,而支持向量机和 K-近邻分类算法的准确率较低。整体来看,随机森林算法和卷积神经网络算法缺陷识别准确率较高,而支持

向量机算法、K-近邻算法和 AdaBoosting 算法的准确率略低。

4 缺陷识别模型的应用效果分析

4.1 应用效果验证试件设计

在超声缺陷检测方法的实际应用中,缺陷的形状、尺寸及深度往往是随机出现的,即包含缺陷形状、尺寸和深度等因素的缺陷类型不是有限且确定的。但若能够将较为接近的缺陷识别为一类以满足实际应用需求,则缺陷识别模型是可行的。

对此,本文设计一个新的试件(验证试件),如图8所示。该验证试件中的缺陷并不属于上述缺陷中的任意一种,但较为接近部分缺陷类型。该试件中共有4种缺陷,分别为直径为1.8 mm的圆通孔、直径为2.2 mm的圆通孔、直径为3.8 mm的圆通孔和直径为4.2 mm的圆通孔,分别记为y-1、y-2、y-3和y-4。

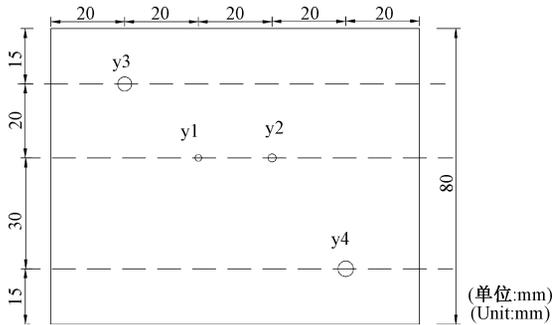


图8 验证试件尺寸及缺陷位置分布

Fig. 8 Verifies the specimen size and defect location distribution

缺陷 y-1 和 y-2 与图2中缺陷 Y-2 形状和位置相同,孔径大小不同但较为接近,用于验证所建模型能否正确将这两个缺陷识别为与缺陷 Y-2 同类别。同理,缺陷 y-3 和 y-4 与缺陷 Y-4 具有相同的形状和位置,和较为接近的尺寸。

4.2 应用效果验证实验和结果分析

使用超声探头对图8中验证试件中的缺陷进行超声检测,按照图9中每个箭头方向进行2次检测,共24组实验数据,记作K3。

使用3.2节中基于K1数据所建里的缺陷识别模型对K3的超声检测数据进行缺陷识别,各算法识别准确率如表4所示(运行15次后的平均值)。

通过表4的结果可知,使用各类算法建立的缺陷识别模型对验证试件的缺陷识别结果中,随机森林算法表现出最高的预测准确率,达到了94.6%。而支持向量机和K-近邻分类算法的预测准确率则低于随机森林算法,为79.2%和83.3%。而AdaBoosting算法和卷积神经网络

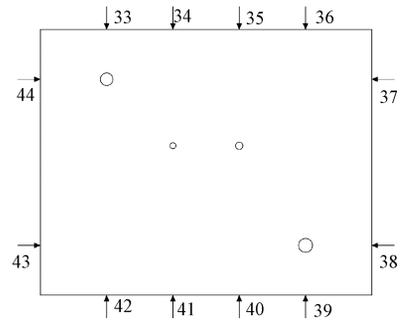


图9 应用效果验证试件位置示意图

Fig. 9 Application effect verification specimen position diagram

的预测准确率最低,且预测准确率不理想。由此说明本文所提基于随机森林的超声缺陷识别方法具有较高的准确率和实用性。

表4 K1、K3数据集运行结果

Table 4 Operation results of K1 and K3 data sets

| 算法类型 | 识别准确率 | 识别准确个数 |
|-------------|-------|--------|
| 随机森林 | 0.946 | 22.7 |
| 支持向量机 | 0.792 | 19.0 |
| K-近邻算法 | 0.833 | 20.0 |
| AdaBoosting | 0.500 | 12.0 |
| 卷积神经网络 | 0.500 | 12.0 |

5 结论

本文针对超声缺陷类型识别问题构建了基于随机森林算法缺陷识别模型,并基于实验数据对比分析其与支持向量机、K-近邻算法、AdaBoosting算法和卷积神经网络的识别准确率。随机森林算法对K1数据集的预测准确率为99.4%,对K2数据集的预测准确率为100%,对验证试件的实验数据K3的预测准确率为94.6%。实验结果表明,本文所提基于随机森林的超声缺陷识别方法能够实现较高的准确率,实用性强。整理来看随机森林算法的预测结果优于其他算法,说明了本文所提基于随机森林的超声缺陷检测方法的优越性。

本文主要考虑形状、尺寸和深度的不同缺陷类型的识别,所建立的缺陷识别模型能够将相似的缺陷识别为一类。考虑到实际应用中钢材中缺陷的具体位置通常是未知的,因此后续将进一步研究针对随机出现的缺陷进行智能和准确检测识别的方法,以弥补当前方法的不足。此外,本研究相比传统B超成像算法具有准确率高和智能化水平高的优点,但无法直观显示缺陷图像信息。实现智能缺陷识别的可视化也是本团队的后续重要的研究方向之一。

参考文献

- [1] 张勇,汪小凯,关山月,等. 中厚板结构件激光焊缝水浸聚焦超声检测及缺陷识别研究[J]. 武汉理工大学学报,2018,40(10):44-50.
ZHANG Y, WANG X K, GUAN SH Y, et al. Research on detection and defect identification of laser welding seam with water immersion focused ultrasound for medium thick plate structure [J]. Journal of Wuhan University of Technology,2018,40(10):44-50.
- [2] 赵一凡,张晓静. 基于超声波检测技术车辆铝制材料无损探伤分析研究[J]. 科技创新与应用,2020(16):1-5.
ZHAO Y F, ZHANG X J. Research on non-destructive testing of aluminum materials for vehicle based on ultrasonic detection technology [J]. Science and Technology Innovation and Application,2020(16):1-5.
- [3] 高翔,彭俊,李冬华,等. 超厚壁无缝钢管纵向缺陷超声波检测探讨[J]. 钢管,2021,50(6):76-79.
GAO X, PENG J, LI D H, et al. Study on ultrasonic detection of longitudinal defects of ultra-thick seamless steel pipe [J]. Steel Pipe,2021,50(6):76-79.
- [4] 杨林,陈春,金福凯. 大型金属结构件自动探伤系统设计[J]. 汽车工艺师,2022(7):46-50.
YANG L, CHEN CH, JIN F K. Design of automatic flaw detection system for large metal structural parts [J]. Automotive Technologist,2022(7):46-50.
- [5] 杨秀文,张明胜,翟玉奇,等. 超声波探伤在磨机轴瓦脱壳检测中的应用[J]. 现代制造技术与装备,2022,58(8):170-173.
YANG X W, ZHANG M SH, ZHAI Y Q, et al. Application of ultrasonic flaw detection in the shell removal detection of mill bearing [J]. Modern Manufacturing Technology and Equipment,2022,58(8):170-173.
- [6] FILHO E F S, SILVA M M, FARIAS P C M A, et al. Flexible decision support system for ultrasound evaluation of fiber-metallaminates implemented in a DSP[J]. NDT&E International,2016,79:38-45.
- [7] 施成龙,师芳芳,张碧星. 利用深度神经网络和小波包变换进行缺陷类型分析[J]. 声学学报,2016,41(4):499-506.
SHI CH L, SHI F F, ZHANG B X. Defect type analysis using deep neural network and wavelet packet transform [J]. Acta Acoustica Sinica,2016,41(4):499-506.
- [8] MENG M, CHUA Y J, WOUTERSON E, et al. Ultrasonic signal classification and imaging system for composite materials via deep convolutional neural networks[J]. Neurocomputing,2017,257:128-135.
- [9] HOU W, WEI Y, JIN Y, et al. Deep features based on a DCNN model for classifying imbalanced weld flaw types[J]. Measurement,2018,131:482-489.
- [10] 侯怀书,任慧霞. 金属管道对接焊缝超声检测缺陷类型识别研究[J]. 热加工工艺,2019,48(21):143-146,152.
HOU H SH, REN H X. Research on defect type identification in ultrasonic detection of butt weld of metal pipeline [J]. Hot Working Technology,2019,48(21):143-146,152.
- [11] 王永利,曹江涛,姬晓飞. 基于卷积神经网络的 PCB 缺陷检测与识别算法[J]. 电子测量与仪器学报,2019,33(8):78-84.
WANG Y L, CAO J T, JI X F. PCB defect detection and identification algorithm based on convolutional neural network [J]. Journal of Electronic Measurement and Instrumentation,2019,33(8):78-84.
- [12] 鲍光海,林善银,徐林森. 基于改进型卷积网络的汽车高度调节器缺陷检测方法[J]. 仪器仪表学报,2020,41(2):157-165.
BAO G H, LIN SH Y, XU L S. Defect detection method of automotive height regulator based on improved convolutional network [J]. Chinese Journal of Scientific Instrument,2020,41(2):157-165.
- [13] 李丹,王慢慢,刘俊德,等. 基于轻量级卷积神经网络的带钢表面缺陷识别[J]. 仪器仪表学报,2022,43(3):240-248.
LI D, WANG M M, LIU J D, et al. Strip surface defect recognition based on lightweight convolutional neural network [J]. Chinese Journal of Scientific Instrument,2022,43(3):240-248.
- [14] 高子洋,师芳芳,张碧星,等. 卷积神经网络的缺陷类型识别分析[J]. 应用声学,2022,41(2):301-309.
GAO Z Y, SHI F F, ZHANG B X, et al. Analysis of defect type recognition using convolutional neural network [J]. Journal of Applied Acoustics,2022,41(2):301-309.
- [15] 陈彦杰,吴丽君,陈志聪. 基于二值卷积神经网络的钢铁缺陷检测[J]. 工业控制计算机,2023,36(7):45-47.
CHEN Y J, WU L J, CHEN ZH C. Steel defect detection based on binary convolutional neural network [J]. Industrial Control Computers,2023,36(7):45-47.
- [16] 陈赛赛,杨晨曦,陈超,等. 基于小波核扩散与双阶段 SVM 的轴承复合故障分类方法[J]. 仪器仪表学报,2023,44(10):179-188.
CHEN S S, YANG CH X, CHEN CH, et al. Bearing compound fault classification Method based on wavelet proliferation and two-stage SVM [J]. Chinese Journal of

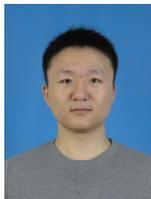
- Scientific Instrument, 2023, 44(10):179-188.
- [17] 胡峻峰,曹军,赵亚凤. 随机森林在板材表面缺陷分类中的应用[J]. 东北林业大学学报, 2015(8):86-90.
HU J F, CAO J, ZHAO Y F. Application of random forest in surface defect classification of plate [J]. Journal of Northeast Forestry University, 2015(8):86-90.
- [18] PATEL S V, JOKHAKAR V N, PATEL S V, et al. A random forest based machine learning approach for mild steel defect diagnosis[C] IEEE International Conference on Computational Intelligence and Computing Research. IEEE, 2016;1-8.
- [19] WANG Y, XIA H, YUAN X, et al. Distributed defect recognition on steel surfaces using an improved random forest algorithm with optimal multi-feature-set fusion[J]. Multimedia Tools and Applications, 2018, 77: 16741-16770.
- [20] 刘传泽,王霄,陈龙现. 基于随机森林算法的纤维板表面缺陷识别[J]. 林业科学, 2018, 54(11):121-126.
LIU CH Z, WANG X, CHEN L X. Surface defect recognition of fiberboard based on random forest algorithm [J]. Scientia Silvae Sinicae, 2018, 54(11):121-126.
- [21] 李冬. 轴承故障高敏感特征提取与随机森林智能识别[J]. 机械设计与制造, 2022(9):157-161.
LI D. Bearing fault highly sensitive feature extraction and random forest intelligent recognition [J]. Machinery Design and Manufacture, 2022(9):157-161.
- [22] 唐宇,代琪,杨志伟,等. 基于优化随机森林的软件缺陷预测算法研究[J]. 计算机工程与科学, 2023, 45(5):830-839.

TANG Y, DAI Q, YANG ZH W, et al. Research on software defect prediction algorithm based on optimal random forest [J]. Computer Engineering and Science, 2023, 45(5): 830-839.

- [23] 范江心. 超声波对钢材缺陷几何特征检测与识别方法研究[D]. 沈阳:东北大学, 2018.

FAN J X. Research on ultrasonic detection and identification of geometric features of steel defects [D]. Shenyang: Northeastern University, 2018.

作者简介



魏新园 (通信作者), 2015年于合肥工业大学获得学士学位, 2020年于合肥工业大学获得博士学位, 现为安徽工业大学讲师, 主要研究方向为精密数控机床误差理论、智能传感技术与机器学习。

E-mail: weixy@ahut.edu.cn

Wei Xinyuan (Corresponding author) received his B. Sc. degree and Ph. D. degree both from the Hefei University of Technology in 2015 and 2020. He is currently a lecturer at the Anhui University of Technology. His main research interests include error modeling theory of precision CNC machine tools, intelligent sensing technology, and machine learning.



周京欢, 现为安徽工业大学硕士研究生, 主要研究方向为超声波无损探伤。

E-mail: 17718276939@163.com

Zhou Jinghuan is currently a M. Sc. candidate at Anhui University of Technology. His research interest includes ultrasonic nondestructive testing.