

# 基于机器学习的压力容器缺陷检测与预测技术研究

胡立洪

杭州皓华压力容器有限公司，浙江杭州，311401；

**摘要：**本文探讨了基于机器学习的压力容器缺陷检测与预测技术。介绍了机器学习在缺陷检测中的应用背景，分析了传统检测方法的局限性，详细阐述了机器学习在压力容器缺陷检测中的具体实现，包括数据预处理、特征提取、模型选择与训练等关键环节。同时，对机器学习在缺陷预测中的应用进行了展望，强调了其在提高检测效率、准确性和智能化水平方面的潜力。

**关键字：**机器学习；压力容器；缺陷检测；缺陷预测

**DOI：**10.69979/3041-0673.25.06.098

## 1 机器学习在缺陷检测中的应用背景

### 1.1 传统缺陷检测方法的局限性

在工业生产领域，压力容器作为承载高压、高温等极端条件的关键设备，其安全性至关重要。然而，由于长期运行、介质腐蚀、疲劳损伤等多种因素，压力容器内部不可避免地会出现各种缺陷，如裂纹、腐蚀坑、变形等。这些缺陷若未能及时发现和处理，极有可能引发严重的安全事故，造成人员伤亡和财产损失。因此，对压力容器进行定期、准确的缺陷检测是保障其安全运行的重要环节。

传统的压力容器缺陷检测方法主要包括超声检测、射线检测、磁粉检测、渗透检测等。超声检测利用超声波在材料中的传播特性来检测缺陷，具有检测灵敏度高、可检测厚度大等优点，但对缺陷的取向和位置较为敏感，且需要专业的操作人员进行精确的操作和判断。射线检测则通过X射线或 $\gamma$ 射线穿透材料，根据射线在缺陷处的衰减情况来成像，能够直观地显示缺陷的形状和大小，但该方法对人体有一定的辐射危害，且设备昂贵，检测成本较高。磁粉检测和渗透检测主要用于检测表面缺陷，对于内部缺陷的检测能力有限<sup>[1]</sup>。

传统检测方法还存在检测效率低、难以实现自动化检测等问题。在大型压力容器的检测中，需要耗费大量的人力和时间，且检测结果容易受到操作人员经验和技能水平的影响，导致检测结果的准确性和可靠性难以保证。随着工业生产规模的不断扩大和对压力容器安全性能要求的不断提高，传统缺陷检测方法已经难以满足实际需求，迫切需要寻找一种更加高效、准确、智能化的检测方法。

### 1.2 机器学习技术的发展与优势

近年来，机器学习技术作为人工智能领域的核心技术之一，取得了飞速的发展。机器学习通过从大量数据中学习并自动改进算法，能够识别和理解数据中的复杂模式和规律。它不需要明确的编程指令，而是通过对数据的学习和分析来自主做出决策和预测<sup>[2]</sup>。

在缺陷检测领域，机器学习技术具有显著的优势。首先，机器学习算法具有强大的数据处理能力，能够处理海量的检测数据。对于压力容器来说，其检测数据可能包括图像数据、传感器数据等多种形式，机器学习算法可以对这些数据进行高效的分析和处理，从中提取出有价值的信息。其次，机器学习算法具有自适应学习能力，能够根据不同的检测对象和检测环境自动调整模型参数，提高检测的准确性和可靠性。例如，在不同的压力容器材质、不同的运行工况下，机器学习模型可以通过学习新的数据来不断优化自身的性能。

机器学习技术还可以实现缺陷检测的自动化和智能化。通过构建合适的机器学习模型，可以实现对压力容器缺陷的自动识别和分类，大大减少了人工干预，提高了检测效率。同时，机器学习模型还可以对检测结果进行实时反馈和预警，及时发现潜在的安全隐患，为压力容器的安全运行提供有力保障。

### 1.3 机器学习在缺陷检测中的应用案例与成效

目前，机器学习技术已经在多个领域的缺陷检测中得到了广泛应用，并取得了显著的成效。在航空航天领域，机器学习算法被用于飞机结构件的缺陷检测，通过对大量的无损检测数据进行学习，能够准确地识别出结构件中的微小裂纹和腐蚀缺陷，大大提高了飞机的安全

性。在电子制造领域，机器学习技术被应用于印刷电路板（PCB）的缺陷检测，能够快速、准确地检测出PCB上的短路、断路等缺陷，提高了电子产品的质量和生产效率。

在压力容器缺陷检测方面，也有相关的研究和应用案例。例如，研究人员利用卷积神经网络（CNN）对压力容器的X射线检测图像进行分析，通过对大量标注好的图像数据进行训练，CNN模型能够准确地识别出图像中的裂纹缺陷，其检测准确率明显高于传统的人工检测方法。此外，还有一些研究利用支持向量机（SVM）等机器学习算法对压力容器的超声检测信号进行处理，实现了对缺陷的自动分类和识别。

## 2 机器学习在压力容器缺陷检测中的实现

### 2.1 数据收集与预处理

数据收集是机器学习在压力容器缺陷检测中的首要步骤。为了构建准确有效的检测模型，需要收集大量与压力容器缺陷相关的数据。这些数据可以来自多个渠道，包括无损检测设备（如超声检测仪、射线探伤仪等）获取的图像数据、传感器采集的运行参数数据（如压力、温度、振动等），以及历史检测记录和维修报告等文本数据。在收集图像数据时，要确保图像的清晰度和分辨率能够满足后续分析的需求。对于运行参数数据，需要保证数据的完整性和准确性，记录不同工况下压力容器的各项参数变化。同时，历史检测记录和维修报告可以提供缺陷的类型、位置、大小以及修复情况等重要信息，有助于建立全面的缺陷数据库。

收集到的数据需要进行标注，以便机器学习模型能够识别和学习缺陷特征。对于图像数据，专业的检测人员需要在图像中标记出缺陷的位置和类型，如裂纹、腐蚀坑等。对于运行参数数据，可以根据历史检测记录将数据分为有缺陷和无缺陷两类，并标注出对应的缺陷类型。数据标注是一个耗时且需要专业知识的过程，标注的质量直接影响模型的训练效果。因此，需要建立严格的标注规范和审核机制，确保标注的准确性和一致性。

原始数据中往往存在噪声、异常值和缺失值等问题，需要进行数据清洗。噪声数据可能是由于设备误差、环境干扰等因素引起的，可以通过滤波、平滑等方法进行处理。异常值是指与大多数数据明显不同的数据点，可能是由于测量错误或特殊情况导致的，可以通过统计分析等方法进行识别和剔除。缺失值则需要根据数据的分

布和特征进行填充或删除。

不同类型的数据可能具有不同的量纲和取值范围，为了消除量纲的影响，提高模型的训练效率和稳定性，需要对数据进行归一化处理。常见的归一化方法包括线性归一化、Z-score归一化等。通过归一化处理，将数据映射到一个特定的范围内，使模型能够更好地学习和处理数据。

### 2.2 特征提取与选择

特征提取是从预处理后的数据中提取出能够反映缺陷特征的信息。对于图像数据，可以使用卷积神经网络（CNN）等深度学习模型进行自动特征提取。CNN通过卷积层、池化层等结构，能够自动学习图像中的边缘、纹理、形状等特征，并将这些特征表示为高维向量。对于运行参数数据，可以使用时域分析、频域分析等方法提取特征。例如，通过计算压力、温度等参数的均值、方差、峰值等统计特征，以及进行傅里叶变换提取频域特征，能够反映压力容器的运行状态和潜在缺陷信息。

提取的特征可能存在冗余和不相关的信息，需要进行特征选择。特征选择的目的是从众多的特征中选择出对缺陷检测最有价值的特征，减少模型的复杂度，提高模型的泛化能力。常见的特征选择方法包括过滤式、包裹式和嵌入式方法。过滤式方法根据特征的统计特性（如相关性、方差等）进行筛选；包裹式方法通过不断尝试不同的特征组合，根据模型的性能选择最优的特征子集；嵌入式方法则在模型训练过程中同时进行特征选择，如决策树算法在构建树的过程中会自动选择重要的特征。

### 2.3 模型选择与训练

根据特征的类型和数量、数据的规模和分布等因素选择合适的机器学习模型。对于图像数据，卷积神经网络（CNN）是常用的模型，它能够自动学习图像中的复杂特征，在图像分类和识别任务中表现出色。对于运行参数数据，支持向量机（SVM）、随机森林（Random Forest）等模型具有较好的分类性能。此外，还可以考虑使用集成学习方法，如梯度提升树（GBDT）、XGBoost等，通过组合多个弱学习器来提高模型的性能。

使用标注好的数据对选择的机器学习模型进行训练。在训练过程中，需要将数据分为训练集、验证集和测试集。训练集用于模型的参数学习，验证集用于调整

模型的超参数，测试集用于评估模型的最终性能。训练过程中需要选择合适的损失函数和优化算法。损失函数用于衡量模型预测结果与真实标签之间的差异，常见的损失函数包括交叉熵损失函数、均方误差损失函数等。优化算法则用于调整模型的参数以最小化损失函数，常见的优化算法包括随机梯度下降（SGD）、Adam 等。

训练完成后，需要使用测试集对模型的性能进行评估。常用的评估指标包括准确率、召回率、F1 值等。准确率是指模型预测正确的样本数占总样本数的比例；召回率是指模型正确预测出的正样本数占实际正样本数的比例；F1 值是准确率和召回率的调和平均数，综合考虑了模型的精确性和召回能力。如果模型的性能不满足要求，需要调整模型参数或重新选择模型进行训练。同时，还可以使用交叉验证等方法来进一步评估模型的稳定性和泛化能力。

### 3 机器学习在压力容器缺陷预测中的应用展望

#### 3.1 缺陷预测的重要性与潜力

压力容器在化工、能源、制药等众多工业领域扮演着至关重要的角色，其安全稳定运行直接关系到生产过程的连续性和人员的生命安全。传统的压力容器缺陷检测主要侧重于对已存在缺陷的发现，然而，能够在缺陷尚未明显显现时就进行预测，提前采取预防措施，对于避免重大安全事故的发生具有更为关键的意义。机器学习凭借其强大的数据处理和模式识别能力，在压力容器缺陷预测方面展现出了巨大的潜力，有望成为保障工业安全的关键技术手段。

通过机器学习对压力容器缺陷进行预测，可以实现从被动维修到主动预防的转变。在缺陷尚未造成严重后果之前进行维修和保养，不仅能够降低维修成本，避免设备因突发故障而导致的长时间停机，还能有效提高设备的利用率，延长设备的使用寿命。这对于企业提高生产效率、降低运营成本具有重要意义。

#### 3.2 机器学习在缺陷预测中的方法与技术

压力容器的运行参数（如压力、温度、振动等）会随着时间的推移而发生变化，这些变化往往蕴含着设备状态的重要信息。时间序列分析方法可以对这些运行参数的历史数据进行分析，挖掘数据中的趋势、周期性和异常模式。例如，ARIMA（自回归积分滑动平均模型）可以对具有线性特征的时间序列数据进行建模和预测，

通过分析运行参数的变化趋势，预测未来可能出现缺陷的时间点。LSTM（长短期记忆网络）等深度学习模型则能够处理非线性、非平稳的时间序列数据，更好地捕捉运行参数中的复杂模式，提高缺陷预测的准确性。

除了时间序列分析，还可以将缺陷预测问题转化为分类或回归问题。收集大量的压力容器历史数据，包括运行参数、检测数据、维修记录等，并标注出是否发生缺陷以及缺陷的严重程度。然后，使用支持向量机（SVM）、随机森林（Random Forest）、梯度提升树（GBDT）等机器学习算法构建分类或回归模型。这些模型可以学习数据中的特征和模式，对新的压力容器数据进行预测，判断其未来发生缺陷的可能性以及缺陷的大致情况。

集成学习通过组合多个弱学习器来构建一个强学习器，能够提高模型的泛化能力和预测准确性。在压力容器缺陷预测中，可以采用 Bagging、Boosting 等集成学习方法。例如，随机森林就是一种基于 Bagging 的集成学习方法，它通过构建多个决策树并结合它们的预测结果，能够有效降低模型的方差，提高预测的稳定性。XGBoost 则是一种基于 Boosting 的集成学习算法，它能够自动处理缺失值、优化模型参数，在处理大规模数据集和高维特征时表现出色。

#### 3.3 面临的挑战与解决方案

机器学习模型的性能高度依赖于数据的质量。在压力容器缺陷预测中，收集到的数据可能存在噪声、缺失值、异常值等问题，同时，数据的标注也需要专业的知识和经验。为了解决这些问题，一方面需要建立严格的数据采集和存储规范，确保数据的准确性和完整性；另一方面，可以采用数据清洗、插补、异常检测等方法对数据进行预处理。此外，还可以组织专业的团队进行数据标注，并建立标注审核机制，提高标注的质量。

许多机器学习模型，尤其是深度学习模型，具有较高的复杂性，其决策过程往往难以解释。在压力容器缺陷预测中，模型的可解释性至关重要，因为企业和监管部门需要了解模型预测的依据，以便做出合理的决策。为了解决模型可解释性问题，可以采用特征重要性分析、局部可解释模型无关解释（LIME）、SHAP（Shapley Additive exPlanations）等方法，对模型的预测结果进行解释和分析，提高模型的透明度和可信度。

在实际应用中，压力容器缺陷预测需要实时处理大量的运行数据，并及时给出预测结果。这就要求模型具

有较高的计算效率和实时性。然而，一些复杂的机器学习模型，如深度学习模型，需要大量的计算资源和时间来进行训练和推理。为了解决这个问题，可以采用模型压缩、量化、剪枝等技术，减少模型的规模和计算量，提高模型的运行速度。同时，还可以利用云计算、边缘计算等技术，将计算任务分配到多个计算节点上，提高计算资源的利用率。

#### 4 结束语

本文深入探讨了基于机器学习的压力容器缺陷检测与预测技术。通过详细阐述机器学习在缺陷检测中的应用背景、实现过程以及缺陷预测的应用展望，展示了机器学习技术在压力容器安全管理中的巨大潜力。随着技术的不断进步和数据的不断积累，相信机器学习将在

压力容器缺陷检测与预测领域发挥更加重要的作用，为保障工业生产的安全和稳定贡献力量。

#### 参考文献

- [1] 董豪, 李少波, 杨静, 等. 基于特征融合与语义引导的药用胶囊表面缺陷检测[J]. 计算机集成制造系统, 2025, 31(1): 158-170.
- [2] 丁洪涛. 基于超声检测技术的压力容器泄漏检测方法与试验研究[J]. 机械管理开发, 2025, 40(1): 12-15.

作者简介：姓名：胡立洪，1966.09，性别：男，民族：汉族，籍贯：浙江省杭州，职称：工程师，研究方向：智能设备系统及计算机。