Pontifícia Universidade Católica
do Rio de Janeiro

Guilherme Rezende Bessa Ferreira

Data-driven ultrasonic non-destructive evaluation of pipes and welds in the context of the oil and gas industry

Dissertação de Mestrado

Dissertation presented to the Programa de Pós–graduação em Engenharia Mecânica of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Engenharia Mecânica.

Advisor : Prof. Helon Vicente Hultmann Ayala
Co-advisor: Prof. Alan Conci Kubrusly

Rio de Janeiro
December 2021

Pontifícia Universidade Católica
do Rio de Janeiro

Guilherme Rezende Bessa Ferreira

Data-driven ultrasonic non-destructive evaluation of pipes and welds in the context of the oil and gas industry

Dissertation presented to the Programa de Pós–graduação em Engenharia Mecânica of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Engenharia Mecânica. Approved by the Examination Committee:

Prof. Helon Vicente Hultmann Ayala
Advisor
Departamento de Engenharia Mecânica – PUC-Rio

Prof. Alan Conci Kubrusly
Co-advisor
Centro de Estudos em Telecomunicações – PUC-Rio

Prof. Arthur Martins Barbosa Braga
Departamento de Engenharia Mecânica – PUC-Rio

Prof. Moisés Luiz Lagares Júnior
Universidade Federal de Juiz de Fora - UFJF

Rio de Janeiro, December the 6th, 2021

Guilherme Rezende Bessa Ferreira

Majored in mechanical engineering by the Federal University of Juiz de Fora (Juiz de Fora-MG,Brazil) in 2019.

Bibliographic data

Ferreira, Guilherme Rezende Bessa

Data-driven ultrasonic non-destructive evaluation of pipes and welds in the context of the oil and gas industry / Guilherme Rezende Bessa Ferreira; advisor: Helon Vicente Hultmann Ayala; co-advisor: Alan Conci Kubrusly. – 2021.

90 f: il. color. ; 30 cm

Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Mecânica, 2021.

Inclui bibliografia

1. Engenharia Mecânica – Teses. 2. Avaliação não destrutiva. 3. Ondas guiadas. 4. Ultrassom. 5. Aprendizado de máquina. 6. Oleoduto. 7. Soldagem. I. Vicente Hultmann Ayala, Helon. II. Conci Kubrusly, Alan. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Mecânica. IV. Título.

CDD: 621

To my parents, for their support
and encouragement.

Acknowledgments

It is a genuine pleasure to express my deep sense of thanks and gratitude to those that contributed to this dissertation.

No words of thanks can sum up the gratitude that I owe to my supervisors. Without their aid, the accomplishment of this document would not be possible. Firstly, I would like to thank my advisor, Professor D.Sc. Helon Ayala, for his careful supervision and continuous guidance, for the opportunities provided, and for all the patience and availability during the last couple of years. I would also like to express my profound gratitude to Professor D.Sc. Alan Kubrusly, my co-Advisor, whose contribution, knowledge, and support were of extreme importance.

I owe a deep sense of gratitude to Professor D.Sc. Arthur Braga, for providing the opportunity to work on a challenging and promising research project at a reputed laboratory in PUC-Rio.

I would like to thank my parents, Monária e José Carlos, my sister Nayara and my girlfriend Ivy, without their emotional support, encouragement, and love I would not be able to accomplish my goals, especially this work. Thank you for being such incredible people and always pushing me up.

I cannot forget to thank my friends Mateus Ribeiro e Pedro Calderano, for encouraging me to pursue my master's degree in PUC-Rio. Special thanks go to Mateus, who helped in the first contribution of this work.

I would also like to thank Paula Aida Sesini and Luis Paulo Brasil, for providing the data used in the third contribution of this work.

In addition, I would like to thank the professors who made up the panel, Professor D.Sc. Arthur Braga and Professor D.Sc. Moisés Lagares, for kindly evaluating my dissertation providing comments that significantly improved the final version of this document.

Finally, I thank the Brazilian National Agency for Petroleum, Natural Gas and Biofuels (ABP) for providing financial support through the Human Resources Program for the Petroleum, Natural Gas and Biofuels sector - PRH-ANP. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001

Abstract

Ferreira, Guilherme Rezende Bessa; Vicente Hultmann Ayala, Helon (Advisor); Conci Kubrusly, Alan (Co-Advisor). **Data-driven ultrasonic non-destructive evaluation of pipes and welds in the context of the oil and gas industry**. Rio de Janeiro, 2021. 90p. Dissertação de Mestrado – Departamento de Engenharia Mecânica, Pontifícia Universidade Católica do Rio de Janeiro.

Ultrasonic non-destructive evaluation is of extreme importance in the oil and gas industry, especially for assets and structures subjected to conditions that accelerate failure mechanisms. Despite being widely spread, ultrasonic non-destructive methods depend on a specialized workforce, thus being error-prone and time-consuming. In this context, pattern recognition methods, like machine learning, fit conveniently to solve the challenges of the task. Hence, this work aims at applying artificial intelligence techniques to address the interpretation of data acquired through ultrasonic non-destructive evaluation in the context of the oil and gas industry. For that purpose, this dissertation involves three case studies. Firstly, ultrasonic guided wave signals are used to classify defects present in welded thermoplastic composite joints. Results have shown that, when using features extracted with autoregressive models, the accuracy of the machine learning model improves by at least 72.5%. Secondly, ultrasonic image data is used to construct an automatic weld diagnostic system. The proposed framework resulted in a lightweight model capable of performing classification with over 99% accuracy. Finally, simulation data was used to create a deep learning model for estimating the severity of corrosion-like defects in pipelines. R^2 results superior to 0.99 were achieved.

Keywords

Non-destructive evaluation; Guided waves; Ultrasound; Machine learning; Pipeline; Welding.

Resumo

Ferreira, Guilherme Rezende Bessa; Vicente Hultmann Ayala, Helon; Conci Kubrusly, Alan. **Avaliação não-destrutiva de dutos e soldas baseada em dados ultrassônicos no contexto da indústria de óleo e gás**. Rio de Janeiro, 2021. 90p. Dissertação de Mestrado – Departamento de Engenharia Mecânica, Pontifícia Universidade Católica do Rio de Janeiro.

A avaliação não destrutiva ultrassônica é de extrema importância na indústria de óleo e gás, principalmente para ativos e estruturas sujeitos a condições que aceleram os mecanismos de falha. Apesar de amplamente difundidos, os métodos ultrassônicos não destrutivos dependem de uma força de trabalho especializada, sendo, portanto, suscetíveis a erros e demorados. Nesse contexto, métodos de reconhecimento de padrões, como o aprendizado de máquina, se encaixam convenientemente para solucionar os desafios da tarefa. Assim, este trabalho tem como objetivo a aplicação de técnicas de inteligência artificial para abordar a interpretação de dados adquiridos por meio de avaliação não destrutiva ultrassônica no contexto da indústria de óleo e gás. Para tanto, esta dissertação envolve três estudos de caso. Primeiramente, sinais de ondas guiadas ultrassônicas são usados para classificar os defeitos presentes em juntas soldadas de compósito termoplástico. Os resultados mostraram que, ao usar atributos extraídos com modelos autoregressivos, a acurácia do modelo de aprendizado de máquina melhora em pelo menos 72,5%. Em segundo lugar, dados ultrassônicos em formato de imagens são usados para construir um sistema de diagnóstico de solda automático. A estrutura proposta resultou em um modelo computacionalmente eficiente, capaz de realizar classificações com acurácia superior à 99%. Por fim, dados obtidos por simulação numérica foram usados para criar um modelo de aprendizado profundo visando estimar a severidade de defeitos semelhantes à corrosão em dutos. Resultados de R^2 superiores a 0,99 foram alcançados.

Palavras-chave

Avaliação não destrutiva; Ondas guiadas; Ultrassom; Aprendizado de máquina; Oleoduto; Soldagem.

Table of contents

List of Figures

List of Tables

List of Abreviations

AI Artificial Intelligence

AR Autoregressive Models

CNN Convolutional Neural Network

CV k-Fold Cross-validation

DT Decision Tree

FEM Finite Element Method

GNB Gaussian Naive Bayes

KNN k-nearest Neighbors

LGR Logistic Regression

MCCV Monte Carlo Cross-validation

ML Machine Learning

MLP Multilayer Perceptron

NAR nonlinear autoregressive models

NDE Non-Destructive Evaluation

OGI Oil and Gas Industry

PCA Principal Component Analysis

PC Principal Component

RF Random Forest

ReLu Rectifier Linear Unit

SHM Structural Health Monitoring

SVM Support Vector Machine

TFD Time and Frequency Domain Features

UGW Ultrasonic Guided Wave

UT-images ultrasonic Images

eMARS Major Accident Reporting System

Our intelligence is what makes us human, and
AI is an extension of that quality

Yann LeCun, .

1
Introduction

The concept of Artificial Intelligence (AI) dates back to 1956 and it refers to "the science and engineering of making intelligent machines, especially intelligent computer programs", as defined by John McCarthy [1, 2]. Recently, AI has attracted attention from several sectors of the world economy, especially from the Oil and Gas Industry (OGI) [3, 4]. The use AI methods in the OGI is mostly related applications of Machine Learning (ML) [5] involving several knowledge fields such as hydraulic fracturing [6], seismic exploration [7], petrophysics [8], geology [9], reservoir engineering [10], drilling [11], pipeline integrity [12], and many others.

The OGI plays a vital role in the global economy as it is responsible for the primary fuel sources used around the globe [13]. The production and distribution of oil and gas involve complex and costly operations, requiring the employment of high-end technology. Therefore, it is necessary to assure the availability and reliability of assets and structures, especially those whose operational conditions can accelerate failure mechanisms [14]. Some of these components are storage tanks, thin-walled vessels, heat exchangers, offshore risers, sealing cement in wells, and pipelines [15, 14, 16]. In this context, the use of Non-Destructive Evaluation (NDE) and monitoring techniques is critical in the OGI.

NDE refers to characterization, discrimination, and prediction of defects without sacrificing the inspected specimen. Several are the types of non-destructive tests such as x-ray [17], acoustic emission [18], eddy current [19], and ultrasound [20, 21]. The latter is known to be less expensive and more reliable [22].

1.1
Motivation

Recently, the use of composite materials has gained strength in the OGI due to the advantages offered over steel and aluminum [23]. Composite materials have a lower density than conventional metallic materials, providing weight reduction of structures. Composite materials have outstanding mechanical properties, such as high fatigue resistance, high stiffness, high-

temperature wear, oxidation resistance capabilities, and cost-effective manufacturing [24, 25]. Amongst these several advantages, corrosion resistance stands out since, according to Major Accident Reporting System (eMARS), corrosion is the principal cause of accidents in the OGI [14]. However, composites are subjected to cracks and delaminations caused by fluctuating and impact stresses [26]. Ultrasonic Guided Waves (UGWs) have been fruitfully employed to investigate the presence of defects in structures made of composite materials [27, 28, 29].

Pipelines are considered the main mean of conveying petroleum products, thus being of utmost importance in the OGI [30]. These assets require great attention to their structural integrity as the presence and propagation of damages can lead to a risk factor for either the environment or human life [31, 32]. One of the most hazardous damage mechanisms that occur in pipelines is corrosion [33], and ultrasonic NDE has been employed for corrosion inspection [34, 35].

Additionally, manufacturing limitations cause pipelines to be constructed by welding several pipes with relatively short lengths. The welding process is controlled by multiple parameters, being influenced by many factors. Given this, a weld is susceptible to various types of flaws like slag inclusion, porosity, lack of fusion, lack of penetration, hydrogen cracks, solidification crack, and reheat crack [36, 37, 38, 39, 40]. Such vast possibilities of defects make the weld a critical point of the pipeline, and research regarding the inspection of these structures is paramount. Ultrasonic testing is one of the most popular non-destructive evaluation methods for damage assessment in welded joints [41, 42, 43].

Traditionally, the interpretation of the data acquired during ultrasonic testing is accomplished by a certified inspector through visual examination, turning the efficiency of the process dependant on the experience and knowledge of the inspector [44]. Therefore, the task is error-prone and time-consuming, especially when there is plenty of data to evaluate [45]. Automatic ultrasonic flaw classification systems are strongly needed, and ML, as a pattern recognition method, has been successfully employed to solve the challenges of interpreting ultrasonic data [46, 47, 48, 49, 50].

1.2 Critical Literature Review

UGWs are elastic waves that propagate along the length of a bounded medium when its cross-section dimension is in the order of the wavelength [51]. UGW based inspection provides several advantages such as the ability to

scan large areas, use of lightweight transducers, enable detection of small-sized damages, and immunity to the interference of low-frequency ambient vibration [52].

Notwithstanding the advantages, it is generally recognized that interpreting elastic waves signals towards modeling damage indexes is not straightforward, due to the complex nature of the propagating phenomenon that includes attenuation, dispersion, refraction [53], mode conversion [54, 55, 56], and mode mixing [57]. ML has been widely employed to interpret UGWs in structures made of composite materials [58, 59, 60]. Seno et al.[61] proposed a feature engineering approach, based on the Time of Arrival extraction for the development of an artificial neural network model for impact localization in composite plates. System identification techniques such as Autoregressive Models (AR) and nonlinear autoregressive models (NAR) models were used as a feature extraction procedure by the authors in [62]. The models' coefficients were then used as input to artificial neural networks for detecting dot defects in carbon fiber reinforced polymer plates. In [63], deep learning techniques were used for Structural Health Monitoring (SHM) in aircraft structures, relying upon wavelet transform and fast Fourier transform, respectively, to extract features from UGW signals.

Joining composite materials is an indispensable step due to manufacturing limitations imposed by the material or the geometry required. In these circumstances, ultrasonic welding is a promising joining method thanks to advantages such as fast joining process and material strength, among others [64]. However, welds with reduced fatigue resistance can occur, and efficient weld quality inspection of composite materials is required. Some works have addressed the task employing different methods [65, 66, 67]. Li et al. [68] proposed a quality inspection method combining the prediction of the failure load and weld quality level output simultaneously by an artificial neural network and a random forest model, respectively. Features consisting of the duration and energy at each welding stage were extracted from the process signatures and used as input for the machine learning models. In [69], a CNN was used to perform weld quality inspection in ultrasonically welded composite joints. The input of the proposed CNN consisted of signals extracted from the welding process. Ochôa et al. [70] investigated the propagation of UGWs in two composite plates that were ultrasonically welded. Although they did address the defect detection task, a time-consuming graphical analysis was adopted and the diagnostic parameters used were found to be dependent on the excitation frequency, providing a reasonable detection accuracy only for a specific frequency.

Regarding the welding of metals, several works have proposed machine learning frameworks to characterize weldment defects from ultrasonic data [71, 72, 73, 74, 75, 76, 77, 78, 79, 80]. In this context, more recently, the authors in [81] compared the performance of a deep learning model with human inspectors in the task of interpreting ultrasonic data towards crack detection in a welded austenitic stainless steel pipe. A data augmentation procedure was adopted to overcome the problem of data scarcity. The results showed that the adopted CNN was able to reach human-level performance. There is also an increasing interest in developing machine learning-based embedded systems for many applications [82, 83, 84, 85, 86], including welding inspection [87, 88, 89]. In [87], a Support Vector Machine (SVM) model was embedded in a Raspberry PI device to enhance visual welding inspection. In order to address the problem of flaw detection in steel-welded joints, the authors in [88, 89] developed a novel segmented analysis of ultrasound signals in an embedded system using extreme learning machines.

As previously discussed, the reliability of pipelines is not only related to the integrity of the weld joint but also the integrity of its external surface. Monitoring corrosion and leaks in oil and gas pipelines is essential to maintaining operational conditions and mitigating risks [90]. Recent studies present several methods to inspect corrosion and leaks in such structures [91, 92, 93, 94, 95]. Liu et al. [96] propose a technique based upon the propagation of acoustic waves to address leakage point location in natural gas pipelines. Different signal processing methods were used, but only the localization of the defect was tackled. The data was obtained with a laboratory-scale gas pipeline. In [97], the authors also used acoustic signals to locate leakage in gas pipelines. Wavelet transforms jointly with SVM were used to detect and estimate the severity of the leak. An experimental setup was used to acquire the data. Yaacoubi et al. [98] use the UGW technique for the safe monitoring of tubular structures. A full-scale pipe with machined corrosion-like defects was used as the test setup. In [99], the authors introduce a new application of the optical frequency domain reflectometry technique to monitor both corrosion and leakage. To verify this method corrosion, and leakage experimental tests were carried out. Liu et al. [100] propose a leak detection method based on Markov feature extraction and a two-stage decision scheme. Pressure data collected from the industrial and experimental fields were used.

As seen in the extensive literature review presented, the use of ultrasound inspection techniques is relevant for the OGI. On this basis, the application of ML to interpret ultrasonic data is helpful as it can provide automatic diagnosis methods, reducing human errors and the time taken to complete the task.

The literature review on inspection of composite materials has shown that the state of art lacks to address the use of UGW coupled with ML to defect detection in ultrasonically welded thermoplastic composite joints. Moreover, the use of AR, often employed in the predictive modeling context, to extract features from UGW signals is scarce. Also, no work enables the use of small data evaluation, which is important as the acquisition process is expensive, time-consuming and usually required for data-driven diagnosis in real-world applications.

Additionally, the discussion has shown that monitoring the health of pipelines is mainly related to the inspection of the welds present in the structure and monitoring the presence of corrosion and leakage points. With respect to welding inspection, there is an increasing interest in research regarding embedded machine learning systems. However, literature addressing the problem is still restricted. Many works employ deep learning, specifically addressing the problem of model workflow building, putting aside optimization towards smaller models for embedded systems. It is also worth noticing that most of the works focus on interpreting A-scan signals (an ultrasonic wave displayed in a graph whose y-axis represents the echo amplitude and the x-axis the time), while ultrasonic phased array systems can readily provide richer data in a B-scan format (a 2D image combining several A-scans collected along with the displacement of the transducer).

Finally, research on pipeline corrosion inspection has shown that several works address the problem using different methods and data. Some papers present a framework to perform defect detection and localization, while others also estimate the severity of the corrosion, which is crucial to schedule maintenance interventions. Concerning the methods employed, few works make use of ultrasonic NDE coupled with ML to interpret the resulting ultrasonic data. Moreover, it is also worth mentioning that the vast majority of the works perform experimental tests, which are costly and time-consuming. Given this, using data obtained through numerical simulation is advantageous, and the literature review showed that the state of art lacks works that use data obtained through numerical simulations of ultrasonic inspection.

1.3 Objectives

The general objective of this work is the application of ML techniques for interpreting data obtained through ultrasonic NDE in the context of the OGI . The work is composed of three case studies with the following specific objectives.

1. To develop an approach, based on supervised machine learning and ultrasonic guided wave inspection, capable of successfully detecting defects in ultrasonically welded thermoplastic composite joints.

2. To propose an efficient machine learning model construction workflow for obtaining a model suitable to be embedded in an automatic diagnostic system to perform pipeline welding inspection using ultrasonic imaging.

3. To develop a deep learning-based framework for corrosion-like defect estimation in oil pipelines, indicating not only the presence of defect but its severity.

The case study (1) is accomplished with experimental data provided by the authors of [70]. Case study (2) is assessed through the use of experimental ultrasonic data provided by [81]. For the third case study (3), simulated data were used.

1.4 Contributions

The contributions of this dissertation are developed based on the case studies previously mentioned. The specific contributions for each work are:

1. Improved Feature Extraction of Ultrasonic Guided Wave Signals for Defect Detection in Welded Thermoplastic Composite Joints;

   A ML-based approach is devised to address the problem of defect diagnosis in ultrasonically welded thermoplastic composite joints. Different feature extraction techniques, traditionally employed to ultrasonic guided wave or vibration signals, and supervised learning paradigms were evaluated. UGW signals with multiple excitation frequencies and resampling-based model validation were used, enabling evaluation with small data. It is worth mentioning that using multi-frequency signals and train/test resampling is relevant as data acquisition is a costly task.

2. Efficient Pipeline Crack Detection with Ultrasonic Data and Machine Learning;

   Several modeling paradigms for crack detection in butt-weld of stainless steel pipes are evaluated, indicating the ones more successful for the task at hand. To do so, we propose metrics for the optimization of the model construction procedure. Such metrics take into account the model size and allow obtaining smaller and less computationally intense models, which are more suitable for being embedded in dedicated hardware.

Additionally, the adoption of an appropriate pre-processing procedure contributed to the use of a simpler model to address the problem. It is also worth mentioning that a more in-depth analysis of the deep learning model adopted in [81] was provided by evaluating its performance with different test sets and fairly comparing it to other model architectures.

3. Corrosion-like Defect Severity Estimation in Pipelines Using Convolutional Neural Networks;

   A deep learning-based framework was developed to estimate the severity of corrosion-like defects in oil pipelines. The proposed framework was constructed based upon a CNN. The data used was obtained through numerical simulation, which is cost-effective and faster when compa to data acquisition through experiments. A resampling-based model validation scheme was used, providing a reliable estimate of model performance.

Each of the aforementioned contributions was submitted either as a journal or a conference paper, as detailed in the following.

1. **FERREIRA, G. R. B.**; RIBEIRO, M. G. C.; KUBRUSLY, A. C.; AYALA, H. V. H.; **Improved Feature Extraction of Guided Wave Signals for Defect Detection in Welded Thermoplastic Composite Joints** (manuscript submitted to Ultrasonics).

2. **FERREIRA, G. R. B.**; G. R. V. TONUSSI; C. H. LLANOS; KUBRUSLY, A. C.; AYALA, H. V. H.; **Efficient Embedded Pipeline Crack Detection with Ultrasonic Data and Machine Learning** (manuscript submitted to IEEE Transactions on Industrial Electronics).

3. **FERREIRA, G. R. B.**; SESINI, P. A.; SOUZA, L. P. B.; KUBRUSLY, A. C.; AYALA, H. V. H.; **Corrosion-like Defect Severity Estimation in Pipelines Using Convolutional Neural Networks** In: 2021 IEEE Symposium Series on Computational Intelligence (SSCI). Orlando, USA: IEEE, 2021.

   The following was produced during the period of the dissertation and is not correlated to this thesis.

4. **FERREIRA, G. R. B.**; AYALA, H. V. H.; **Improved image-based TIG weld defect classification with dimensionality reduction and shallow learning** (manuscript submitted to Experimental Mechanics).

1.5 Organization

This document is divided in four main parts.

Part is dedicated to introducing the research theme explored in this dissertation. Chapter 1 provides contextualization, motivation, state-of-art review, objectives, and contributions of this dissertation.

Part II discuss the ML methods employed in this work. Chapter 2 presents the concepts of ML and the fundamentals of the supervised learning algorithms adopted .The feature extraction techniques are presented in Chapter 3. In Chapter 4 the model construction and validation strategies are stated.

The Part III is devoted to the original contributions of this dissertation. Chapter 5 focus on the problem of defect detection in ultrasonically welded thermoplastic composite joints , and in Chapter 6 efficient pipeline crack detection with machine learning and ultrasonic data is tackled. In Chapter 7, corrosion-like defect severity estimation in pipelines is addressed.

Afterwards, in Part IV, Chapter 8 presents the conclusions of the devised case studies and respective future work suggestions.

Part II

Theoretical Background

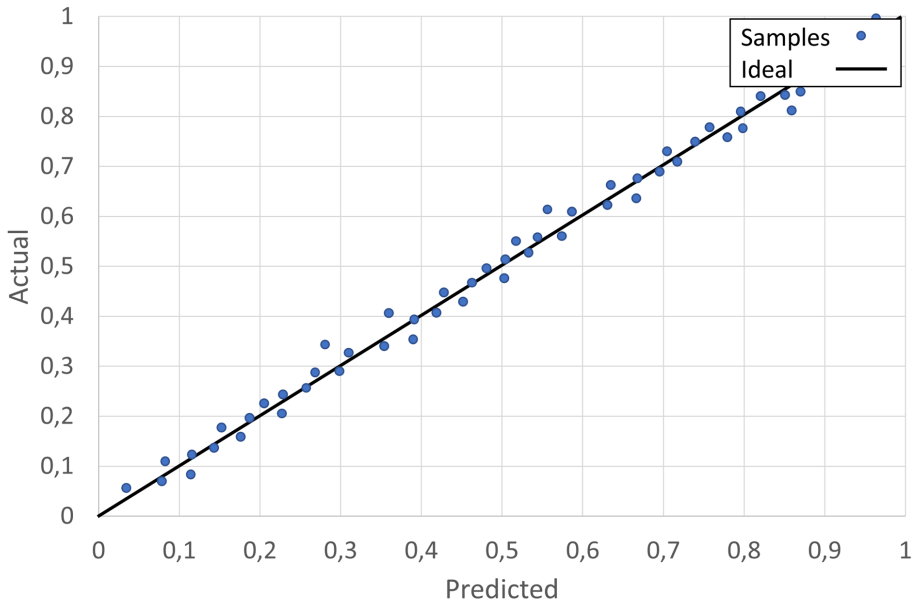
2
Machine Learning

ML is a branch of AI that uses algorithms that iteratively learn from data without being explicitly programmed, improving its capacity for making inferences on new observations. ML algorithms are classified according to the type of train they receive. If an algorithm is trained with labeled data, input/output pairs, it is classified as supervised learning. On the other hand, an algorithm trained with unlabeled data, only inputs, is defined as unsupervised learning. There still exists a third group that comprehends algorithms that learn through trial-and-error, the so-called reinforcement learning. Despite the existing types, the majority of machine learning applications in the engineering field use supervised learning.

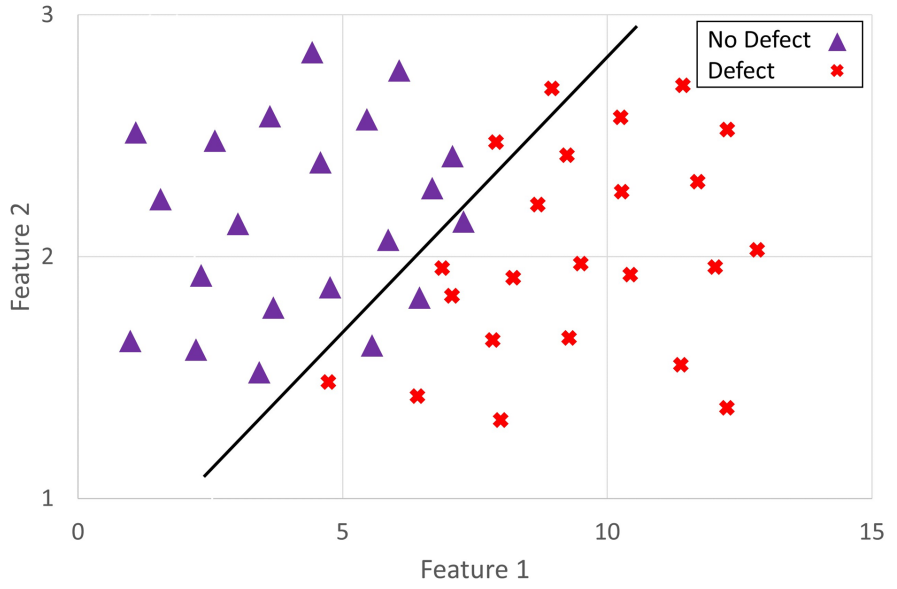
2.1
Classification and Regression Problems

Supervised learning problems can be of two types, namely, classification and regression. A problem is said to be of classification or regression based on the output given by the algorithm. When the output is continuous, like the depth or area of a defect, we have a regression task. Fig 2.1(a) depicts a simplified example of a regression problem. The blue dots correspond to the outcomes of the regression model, and the black line ($y = x$) represents an ideal model. The closest the blue dots lay to the black line, the better is the regression model.

A classification problem occurs when the desired output is a categorical variable that represents a class or state, e.g., *defective* and *non-defective*. A simplified example of a classification problem is exhibited in Fig 2.1(b). In such a problem, the model tries to map a hyperplane (black line) that separates the samples according to their respective classes. The case depicted refers to a binary classification problem with the data having only two features. More complex problems involving several classes and features can occur, but the idea remains the same.



(2.1(a))



(2.1(b))

Figure 2.1: Simplified regression (a) and classification (b) problems. In (a), the blue dots correspond to the model prediction, and the black line represents an ideal model. In (b), the black line corresponds to the hyperplane that represents the model outcome.

2.2 Supervised Learning Algorithms

A machine learning model is a result obtained after training a machine learning algorithm, which can be of different families such as linear in the parameters, instance-based, tree-based, neural networks, bayesian, and ensemble. This work considers several supervised learning algorithms, which are described

in the following.

2.2.1 Decision Tree

Decision Trees (DTs) are core supervised learning algorithms [101] that classifies the data according to a series of questions, each one contained in a node, about the features associated with the instances of a dataset [102]. The root node, hidden nodes, and leaf nodes are the three types of nodes that compose the hierarchical structure of a decision tree. Basically, the classification is achieved following a path that starts at the root node, passes through the hidden nodes, until reaching a leaf node associated with a class [5].

There are several types of decision tree algorithms, but the most popular is known as CART (classification and regression tree). It relies on searching, at each data split, a feature (k) and an associated threshold (t_k) that produces the purest subsets [103]. The cost function that the algorithm tries to minimize is defined as:

$$J\left(k, t_k\right) = \frac{m_{left}}{m} G_{left} + \frac{m_{right}}{m} G_{right}, \tag{2-1}$$

where $m_{left/right}$ is the number of instances in the left/right subsets, respectively. $G_{left/right}$ is the impurity of each subset and it is given by:

$$G_i = 1 - \sum_{k=1}^{m} p_{i,k}^2, \tag{2-2}$$

being $p_{i,k}$ the ratio of class k instances among all the instances in the i^{th} node.

2.2.2 Random Forests

Random Forests (RFs) can be defined as a random combination of decision trees in a way that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest [104]. In summary, each tree in the forest is trained on a random subset obtained from the original dataset. The sampling can be performed with replacement, in what is called *bagging* [105], or without replacement, known as *pasting* [106]. Additionally, it is possible to sample the features for each predictor [107]. In this case, the trees are trained on a random subset of the input features. Finally, after all the individual trees are built, considering bagging or pasting and feature sampling or not, the final estimation is performed considering the predictions of each tree.

2.2.3 Gaussian Naive Bayes

The Gaussian Naive Bayes (GNB) algorithm relies on estimating the class of a new instance based on the prior probability distributions of the labeled data [101]. It is a probabilistic model based on Bayes' theorem. The main idea is to calculate the conditional probability for a class (y_i) with a given instance (x_i), which is defined, according to the Bayes Theorem, as:

$$P\left(y_i \mid a_1(x_i), a_2(x_i), \ldots, a_r(x_i)\right) = \frac{P\left(y_i\right) \cdot P\left(a_1(x_i), a_2(x_i), \ldots, a_r(x_i) \mid y_i\right)}{P\left(a_1(x_i), a_2(x_i), \ldots, a_r(x_i)\right)}, \tag{2-3}$$

where $a_r(x_i)$ is the value of the r^{th} attribute of x_i. $P\left(y_i\right)$ is the prior probability, $P\left(a_1(x_i), a_2(x_i), \ldots, a_r(x_i)\right)$ is the probability of the predictor values and $P\left(a_1(x_i), a_2(x_i), \ldots, a_r(x_i) \mid y_i\right)$ is the conditional probability of the observation based on the class. The first is easily inferred from the dataset, while the last two require complex calculations.

Therefore, the naive assumption of independence between every pair of attributes (predictors) simplifies Eq. 2-3 to

$$P\left(y_i \mid a_1(x_i), a_2(x_i), \ldots, a_r(x_i)\right) = P(y_i) \prod_{k=1}^{r} P(a_k(x_i) \mid y_i) \tag{2-4}$$

$P\left(y_i\right)$ can be estimated as the number of instances labeled as y_i divided by the total number of instances in the dataset. $P(a_k(x_i) \mid y_i)$, also called likelihood of the attribute and is assumed to be Gaussian distributed:

$$P(a_k(x_i) \mid y_i) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left\{\frac{-\left[a_k(x_i) - \mu_y\right]^2}{2\sigma_y^2}\right\}, \tag{2-5}$$

where the mean (μ_y) and the standard deviation (σ_y) are estimated using maximum likelihood [108].

Finally, the predicted class can be defined as the one with the highest probability amongst the possible classes. This approach is known as the maximum a posteriori (MAP) decision rule [5].

2.2.4 k-nearest Neighbors

The k-nearest Neighbors (KNN) is one of the simplest supervised learning algorithms [101]. It is an instance-based method where a sample is classified according to the class assigned to the majority of the k instances nearest to the query instance [101]. The definition of the nearest neighbors is given by a distance metric, usually the Euclidean distance. This metric is defined as:

$$d\left(x_i, x_j\right) = \sqrt{\sum_{r=1}^{n}(a_r(x_i) - a_r(x_j))^2}, \tag{2-6}$$

where $a_r(x)$ denotes the value of the r^{th} attribute (feature) of instance x [5].

2.2.5 Logistic Regression

Logistic Regression (LGR) is one of the go-to methods for binary classification problems. Basically, the task of classification is to learn a hypothesis h that takes an instance $\boldsymbol{x}$ and returns a value between 0 and 1 [108]. This hypothesis is defined as:

$$h_{\boldsymbol{\theta}}(\boldsymbol{x}) = \sigma(\boldsymbol{x}^T\boldsymbol{\theta}), \tag{2-7}$$

where $\boldsymbol{\theta}$ is the parameter vector and $\sigma(\cdot)$ is the logistic function (sigmoid), given by:

$$\sigma(t) = \frac{1}{1+e^{-t}}. \tag{2-8}$$

The value $h_{\boldsymbol{\theta}}(\boldsymbol{x})$ represents the estimated probability, $\hat{p}$, that an instance $\boldsymbol{x}$ belongs to the positive class (labeled as "1"). After obtaining $\hat{p}$, the prediction $\hat{y}$ is easily made by:

$$\hat{y} = \begin{cases} 0 \;\text{ if }\; \hat{p} < 0.5 \\ 1 \;\text{ if }\; \hat{p} \geq 0.5 \end{cases} \tag{2-9}$$

2.2.6 Multilayer Perceptron

The Multilayer Perceptron (MLP) is one type of an artificial neural network, a mathematical model that mimics the functioning of the human brain [103]. Tipically, the MLP consists of layers of neurons, namely input layer, hidden layers, and output layer [109]. The neuron is the main part of a neural network and it can be mathematically represented by:

$$y_k = \phi\left(b_k + \sum_{j=1}^{m} w_{kj}x_j\right), \tag{2-10}$$

where $\phi(\cdot)$ is the activation function. w_{kj}, b_k, x_j, and y_k are, respectively, the weights, bias, inputs and output of the k-th neuron. The activation function is the source of the nonlinearity of the model and it can be of many types such as threshold function, the logistic function, the hyperbolic tangent function, and the Rectified Linear Unit function [103].

In an MLP, every layer, with exception of the output layer, includes a bias neuron and is fully connected to the next layer. Basically, the computation

of the output is done by the propagation of the values from the input to the output layer. A node collect the outputs of the nodes in the previous layers, compute the value defined in Eq. 2-10, and passes this value to every node in the subsequent layer.

2.2.7 Support Vector Machines

The SVM is a popular supervised learning algorithm frequently used for applications where the best classification scores are required [101]. Recently, it has been successfully employed as a tool for assessing structural defects [110, 58, 34].

The algorithm tries to classify the d-dimensional data by finding a hyperplane that separates the instances into groups of the same class. There are plenty of possible hyperplanes for a given dataset. It can be defined as the one that maximizes the distance to the nearest data points, what is called hard margin classification, or it can be defined as the one that has the smaller distance to the nearest data points, known as soft margin classification. A hard margin classification gives more importance on classifying all the training data correctly, possibly causing overfitting. On the other hand, soft margin classification allows some training data to be misclassified. The definition between soft or hard margin is given by the regularization hyperparameter C [111].

The construction of the model is defined as the optimization problem [112]:

$$\min_{w,b,\xi} \left(\frac{1}{2} \boldsymbol{w}^T \boldsymbol{w} + C \sum_{i=1}^{l} \xi_i \right), \tag{2-11}$$

subject to

$$y_i(\boldsymbol{w}^T \phi(\boldsymbol{x_i}) + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, i = 1, \ldots, n, \tag{2-12}$$

where $y_i \in \{-1, 1\}$, $\boldsymbol{w}$ and b parametrize the hyperplane, $C > 0$ is the regularization hyperparameter, ξ_i is the margin distance, and $\phi(\boldsymbol{x_i})$ is a kernel function that maps $\boldsymbol{x_i}$ into a higher-dimensional space, where the data is separable.

2.2.8 Convolutional Neural Network

A CNN is a robust deep learning algorithm capable of dealing with data of different dimensions like 1D signals and sequences; 2D images or spectrograms; and 3D data such as videos or RGB images [113]. Given the flexibility of dealing with input data of different dimensions and the capacity of processing data in

its raw state, without the need for handcrafted feature extraction, CNNs have been widely employed in diverse fields such as computer vision [114], natural language processing [115], time series classification[116], time series forecasting [117] and structural health monitoring [118].

A typical CNN can be divided into two main blocks. The first part, composed of convolutional and pooling layers, functions as a features extractor, while the second part, composed of fully connected layers, builds a relationship between the extracted features and the target variable. In synthesis, the convolutional layer generates a feature map by applying convolutional filtering operations, with small sliding windows, to an entity. The output of this layer is then reduced by a pooling layer while preserving the relevant information. Traditionally, the pooling operations consist in calculating the maximum (max pooling) or the average (average pooling) value of a subarray. These convolution and pooling operations can be repeated several times before concatenating the last feature map into a 1D vector provided to the fully connected layers. The dense layers, another name for fully connected layers, are composed of artificial neurons and activation functions. They are the same as the layers that compound shallow learning neural networks. The activation function is the source of non-linearity of the models, and it can be of many types being the Rectifier Linear Unit (ReLu) commonly used in deep learning applications [119].

2.3 Discussion

The present chapter was dedicated to discussing fundamental ML concepts, and the supervised learning algorithms used throughout this dissertation. This dissertation addresses three supervised learning problems, two of them are of classification, and one is of regression. Several architectures were presented, going from shallow to deep learning methods. According to the *No Free Lunch Theorem*, no algorithm stands out as the best when evaluated across all possible problems [120]. Hence, one should test architectures from different families before defining the most adequate for the application in question.

3
Feature Extraction

Feature extraction refers to the transformation of the data into a new set of variables, ideally reducing the dimensionality and preserving the information present in the original data. An effective feature extraction remove irrelevant and redundant data, improving the learning process and result comprehensibility [121]. Additionally, feature extraction can enrich data understanding and visualization providing valuable insights for the model construction. There are several feature extraction approaches such as automatic extraction through convolutions [122], signal processing techniques [123], image feature extraction [124], linear dimensionality reduction [125], non-linear dimensionality reduction [126], and many others. In the remainder of this chapter we present the mathematical formulation of the methods adopted in this dissertation.

3.1
Principal Component Analysis

The Principal Component Analysis (PCA) is a dimensionality reduction technique. It relies on transforming high-dimensional interrelated data into a novel dataset composed of uncorrelated variables, known as Principal Components (PCs). The transformation is performed preserving the largest amount of the variation present in the original dataset [127]. The PCA can be computed as described hereafter.

Being $\boldsymbol{X} \in \Re^{j \times i}$ the dataset matrix with j samples and i measurements, the mean-subtracted data matrix $\boldsymbol{B}$ is defined as [101]

$$\boldsymbol{B} = \boldsymbol{X} - \bar{\boldsymbol{X}}, \tag{3-1}$$

where $\bar{\boldsymbol{X}} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \bar{\boldsymbol{x}}$ and $\bar{\boldsymbol{x}}_j = \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{X}_{ij}$.

The covariance matrix of mean-subtracted data matrix is defined as $C = B^T B$ and its eigenvectors ($\boldsymbol{V}$) and eigenvalues ($\boldsymbol{D}$) matrices are computed with $\boldsymbol{CV} = \boldsymbol{VD}$. Once the eigenvectors of the covariance matrix corresponds to the directions of the principal components, the projection of the data in the principal components space is given by

$$\hat{\boldsymbol{B}} = \boldsymbol{BV} \tag{3-2}$$

On the other hand, the Singular Value Decomposition (SVD) of the mean-subtracted data matrix is defined as

$$\boldsymbol{B} = \boldsymbol{U\Sigma V}^T, \tag{3-3}$$

where $\boldsymbol{U}$ is the matrix of left singular vectors of $\boldsymbol{B}$, $\boldsymbol{V}$ is the matrix of right singular vectors of $\boldsymbol{B}$, and the diagonal elements of $\boldsymbol{\Sigma}$, which are in descending order, are called singular values.

Comparing Eq. (3-2) to Eq. (3-3), we can see that right-hand singular vectors correspond to eigenvectors of the covariance matrix. Hence, the projection of the data in the principal component space is also given by:

$$\hat{\boldsymbol{B}} = \boldsymbol{U\Sigma} \tag{3-4}$$

Finally, Eq. (3-4) shows that the PCA can be performed by applying the SVD to the mean-subtracted data matrix.

3.2 Autoregressive Model

The AR is a well-known system identification technique capable of extracting information about the dynamics of a system. Given this capacity, it has been used for feature extraction in many works [128, 129, 62, 130]. The extracted features are the coefficients of the model.

The AR can be defined by the following equation [131]

$$y\left(k\right) + a_1 y\left(k-1\right) + a_2 y\left(k-2\right) + \cdots + a_{n_a} y\left(k-n_a\right) = \xi(k), \tag{3-5}$$

where n_a indicates the order of the model, $y\left(k\right)$ is the discrete signal to be modeled, $\xi(k)$ corresponds to the measurement error or noise and a_i are the AR coefficients.

Defining the regression vector $\phi(k)$, with dimension $n_a \times 1$, as

$$\phi(k) = \left[-y\left(k-1\right) \quad -y\left(k-2\right) \quad \ldots \quad -y\left(k-n_a\right)\right]^T, \tag{3-6}$$

and the parameter vector, also with dimension $n_a \times 1$, as

$$\boldsymbol{\theta} = \left[a_1 \quad a_2 \quad \ldots \quad a_{na}\right]^T, \tag{3-7}$$

Eq. (3-5) can be rewritten as

$$y\left(k\right) = \phi^T(k)\boldsymbol{\theta} + \xi(k), \tag{3-8}$$

which denotes a linear regression model [132].

Once we have N measurements, the regression matrix, with dimension $(N-p)\times n_a$, where p is equal to n_a+1, is denoted as

$$\mathbf{\Phi} = \begin{bmatrix} \phi^T(p) \\ \phi^T(p+1) \\ \vdots \\ \phi^T(N) \end{bmatrix}, \tag{3-9}$$

and the target measurement vector and the noise vector are given by:

$$\boldsymbol{y} = \begin{bmatrix} y(p) \\ y(p+1) \\ \vdots \\ y(N) \end{bmatrix} \quad ; \quad \boldsymbol{\xi} = \begin{bmatrix} \xi(p) \\ \xi(p+1) \\ \vdots \\ \xi(N) \end{bmatrix} \tag{3-10}$$

Finally, Eq. (3-5) in the matrix form is

$$\boldsymbol{y} = \mathbf{\Phi}\boldsymbol{\theta} + \boldsymbol{\xi} \tag{3-11}$$

and the model can be estimated by a least squares approach, like the batch least squares algorithm.

3.3 Time and Frequency domain features

Time and Frequency Domain Features (TFD), namely Time-domain signal energy, characteristic frequency shift, Time-of-Flight (ToF), peak frequency shift, power spectral moment of order 1, and frequency bandwidth, are herein defined according to [70].

3.3.1 Time-domain signal energy

The energy of the time-domain signal, $x(t)$, is defined as

$$Es = \int_{t_i}^{t_f} |x(t)|^2 dt \tag{3-12}$$

3.3.2 Characteristic frequency shift

The characteristic frequency shift, Δf_{ch}, corresponds to the difference between the characteristic frequency of the excitation signal and the received signal. In turn, the characteristic frequency, the average of all frequencies in the FFT signal weighted by the corresponding coefficients, is given by:

$$f_{ch} = \sum_{i=1}^{i=n} FFT(x)_i \cdot f_i / \sum_{i=1}^{i=n} FFT(x)_i \tag{3-13}$$

3.3.3
Peak frequency shift

The peak frequency shift is defined as

$$\Delta f_{peak} = \max\left[FFT(x_s)\right] - \max\left[FFT(x_e)\right], \tag{3-14}$$

where x_s and x_e are the sensed and the excitation signals, respectively.

3.3.4
Power spectral moment of order 1

The power spectral moment of order 1, calculated for the sensed signal, is given by:

$$M_1 = \int_0^{f_n} (f - f_{ch}) W(f) df, \tag{3-15}$$

where f_n is the Nyquist frequency and $W(f)$ is the power spectral density function.

3.3.5
Time-of-Flight and frequency bandwidth

The ToF is inversely proportional to the group speed [133] and it is calculated as the time interval between the maximum amplitude point of the excitation and the sensed signal. Additionally, the frequency bandwidth was calculated considering a 25 dB threshold.

3.4
Discussion

In this chapter, the feature extraction methods used in this work were presented. The approaches herein stated consist in linear dimensionality reduction, system identification strategies and signal processing techniques. The construction of features is one of the key steps in the predictive modeling pipeline as it generally reduces the amount of redundant information from the data, speeding up the model construction. Therefore, adopting a suitable approach is of great importance as it largely conditions the success of a machine learning model [134].

4
Model Construction and Validation

This chapter is dedicated to discussing the methods used in the model construction phase, which is the core of the machine learning modeling pipeline. In such a stage, it is crucial to adopt proper means of assessing the model effectiveness faithfully. Therefore, to achieve an unbiased estimate of the generalization performance, the techniques presented in the following may be considered. They concern the validation scheme, the hyperparameter tuning strategy, and the evaluation metrics.

4.1
Resampling-based Model Construction and Validation

Machine learning models are structures that are built accordingly to the data at hand. Given this, it is necessary to establish means of evaluating how well a model will perform when presented with new data. One traditional way of doing this is to split the data into two subsets, namely, training and test. A common strategy used to split the dataset is the hold-out approach, which consists of randomly splitting the data according to a defined proportion, e.g., 70% and 30%. As the name suggests, the training set is responsible for teaching the model, and the test set is used to evaluate the generalization capabilities of the model. However, using a single test set may produce biased estimates of the model performance. One way to overcome such problem is to adopt resampling techniques, for which the dataset splitting is repeated multiple times enabling the use of all available data to construct the model and establishing statistically significant results.

One of the several existing resampling techniques is the k-Fold Cross-validation (CV) procedure. In this case, the dataset is divided into k equal-sized parts, known as folds, and the model is fit k times. In each fitting round, one fold is kept apart for validation and the other $k - 1$ folds are used for training [135], as depicted in Fig 4.1. With this strategy, each sample in the dataset is tested exactly once and the number of possible partitions evaluated is limited by the k chosen.

When the CV is performed with an unbalanced dataset, a stratified k-fold division, where each fold preserves the proportions of classes in the original

dataset, is commonly used. Additionally, a single run of the CV procedure may not provide a good estimation of model performance as different splits of data lead to different results. The repeated k-fold cross-validation overcome this issue by repeating the procedure multiple times and reporting the mean result. It is possible to combine these two strategies in what is called the Repeated Stratified k-fold Cross-validation.

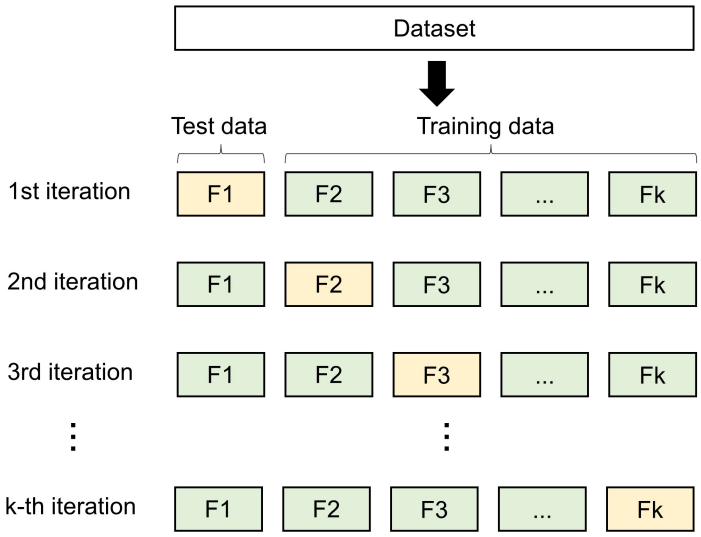


Figure 4.1: Schematic representation of the k-fold cross-validation procedure. The dataset is divided into k parts, and at each iteration k-1 folds are used for training while 1 is kept apart for checking the model performance. The training and test are repeated k times so that each sample in the dataset is tested exactly once.

Another popular resampling method is the Monte Carlo Cross-validation (MCCV), which consists in repeating the holdout approach multiple times. Given the random nature of the sampling process, there is a large number of possible partitions and thus each sample in the dataset is tested arbitrary times. The number of repetitions is defined by the user, and a choice that results in stable estimates of model performance lies in the range 50-200 [136]. Fig 4.2 depicts a schematic representation of the procedure.

4.2 Hyperparameter tuning

During the training of a machine learning algorithm, many parameters related to the optimization problem are tuned, however, some non-trainable parameters cannot be learned by common optimization methods. These parameters are known as hyperparameters, and they need to be set previously to the training process. In many cases, the hyperparameter control the model complexity and a poor definition can lead to overfitting [136]. Hence, defin-

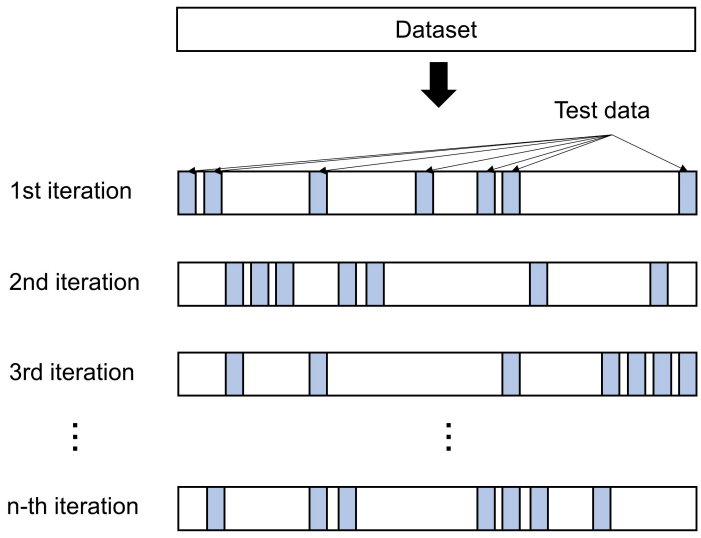


Figure 4.2: Schematic representation of Monte Carlo cross validation procedure. The data is randomly split n times, according to a holdout proportion. A good choice for number of iterations (n) is between 50 and 200.

ing the hyperparameters is a challenging task and a reasonable approach is to search for combinations of values that result in better performance. Such a procedure is called hyperparameter optimization or hyperparameter tuning, and is schematically depicted in Fig. 4.3.

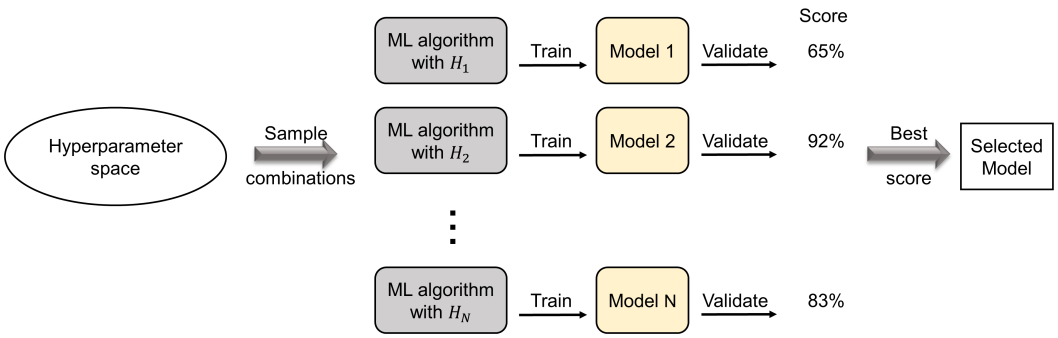
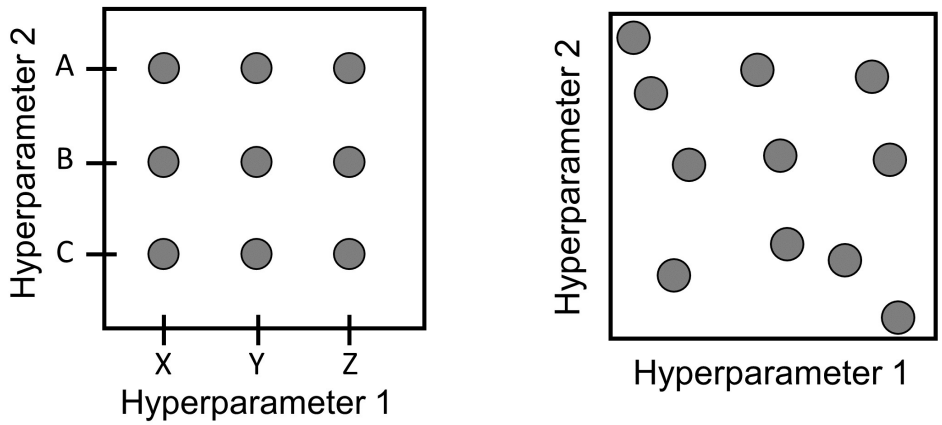


Figure 4.3: Schematic representation of the hyperparameter tuning process.

Different method can be used to perform hyperparameter search such as exhaustive searching [137, 138], evolutionary optimization [139], bayesian optimization [140], and surrogate models [141]. Two of the most commonly adopted methods are Grid Search and Random Search, which are classified as exhaustive searching. Both methods consists of evaluating combinations of hyperparameters within a predefined universe, as schematically depicted in Fig. 4.4. Regarding the Grid Search, the search space is defined as a grid of hyperparameters and every possible combination in the grid is evaluated. On the other hand, for the Random Search a bounded search space is defined and the algorithm evaluates a predefined number of random combinations of



(4.4(a)) (4.4(b))

Figure 4.4: Comparison between Grid Search and Random Search strategies.

hyperparameters values. Depending on the nature of the problem, available computational resources, and the machine learning algorithm in question, the first approach can take considerable time. Moreover, the Random Search has shown to be more efficient [138]. Therefore, the strategy adopted in this work for hyperparameter tunning is the Random Search scheme.

It is worth mentioning that, simply applying a random search strategy is not sufficient to define the hyperparameters as it would result in a model adapted to a particular dataset, being unlikely to perform well on new data. In view of this, the random search can be combined with a cross-validation procedure. This way, it is guaranteed that the selected hyperparameter combination provides the best model within the possible ones.

4.3 Metrics

This subsection is dedicated to presenting the evaluation metrics adopted in this study. The computation of metrics is essential in every machine learning pipeline, and it can refer both to the accuracy as well as the computational performance of the devised model.

4.3.1 Classification Metrics

Classification models can generate two types of predictions, namely a continuous value representing the probabilities for the class labels or a discrete category representing the predicted class label [136]. Therefore, computing

classification metrics consists in comparing the ground truth class label to the predicted class label.

Classification problems can be divided into binary and multiclass classification, according to the target variable. One example of binary classification is the task of detecting the presence of a defect in a structure, for which the output of the model can be 'defective' and 'non-defective'. Distinguishing which is the nature of the defect, e.g. crack, corrosion, and delamination, is an example of multiclass classification problem. Similar metrics can be used for both but since the nature of the problem changes, the metrics are defined accordingly.

Regarding binary classification, the metrics adopted in this work are accuracy (Acc), error rate (Err), precision (Pr), recall (Re), and F1-score (F_1), defined as:

$$Acc = \frac{TP + TN}{P + N} \tag{4-1}$$

$$Err = 1 - \frac{TP + TN}{P + F} \tag{4-2}$$

$$Pr = \frac{TP}{TP + FP} \tag{4-3}$$

$$Re = \frac{TP}{TP + FN} \tag{4-4}$$

$$F_1 = 2 \cdot \frac{Pr \cdot Re}{Pr + Re}, \tag{4-5}$$

where True Positives (TP), True Negatives (TN), False Positive (FP), and False Negatives (FN) represent, respectively, defective samples correctly detected, non-defective samples correctly classified as non-defective, non-defective samples incorrectly classified as defective, and defective samples incorrectly classified as non-defective.

In the case of multiclass classification, the metrics adopted in this work are balanced accuracy ($BAcc$), weighted precision (Pr), weighted recall (Re), and weighted F1-score (F_1), defined as:

$$BAcc = \frac{1}{\mid L \mid} \sum_{l \in L} R\left(y_l, \hat{y}_l\right) \tag{4-6}$$

$$Pr = \frac{1}{\sum_{l \in L} \mid \hat{y}_l \mid} \sum_{l \in L} \mid \hat{y}_l \mid P\left(y_l, \hat{y}_l\right) \tag{4-7}$$

$$Re = \frac{1}{\sum_{l \in L} \mid \hat{y}_l \mid} \sum_{l \in L} \mid \hat{y}_l \mid R\left(y_l, \hat{y}_l\right) \tag{4-8}$$

$$F_1 = \frac{1}{\sum_{l \in L} \mid \hat{y}_l \mid} \sum_{l \in L} \mid \hat{y}_l \mid F_1\left(y_l, \hat{y}_l\right) \tag{4-9}$$

where L, y_l, $\hat{y}_l$ are the set of labels, the subset of predicted labels of class l, and the subset of true labels from class l, respectively. $P\left(y_l, \hat{y}_l\right)$, $R\left(y_l, \hat{y}_l\right)$, $F_1\left(y_l, \hat{y}_l\right)$

are the precision, recall, and F1-score considering a subset of two classes from L.

4.3.2 Regression Metrics

As discussed in Chapter 2, different from classification problems, regression models output a continuous value. Therefore, the performance of the predictor must be evaluated with error metrics. The regression metrics adopted in this work are root mean squared error ($RMSE$), mean absolute error (MAE), maximum error ($MAXE$), and coefficient of determination (R^2), defined as:

$$RMSE(y, \hat{y}) = \sqrt{\frac{1}{n_{samples}} \sum_{i=1}^{n_{samples}} (y_i - \hat{y}_i)^2} \tag{4-10}$$

$$MAE(y, \hat{y}) = \frac{1}{n_{samples}} \sum_{i=1}^{n_{samples}} \mid y_i - \hat{y}_i \mid \tag{4-11}$$

$$MAXE(y, \hat{y}) = max(\mid y_i - \hat{y}_i \mid) \tag{4-12}$$

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^{n_{samples}} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n_{samples}} (y_i - \bar{y})^2} \tag{4-13}$$

where $\hat{y}_i$ is the predicted value of the i-th sample, y_i is the corresponding ground truth value and $\bar{y} = \sum_{i=1}^{n_{samples}} y_i$.

$RMSE$ and MAE measure the average magnitude of the error between predictions and actual values, with the first giving high weight to higher errors. $MAXE$ corresponds to the maximum error within the test samples and R^2 represents how good a regression model fits a set of observations.

4.3.3 Computational performance metrics

In addition to traditional accuracy-related metrics, it is generally handy to assess the computational performance of the models. Such analysis is relevant, especially when considering the embedding in dedicated hardware, where the optimization towards computationally efficient models is required given architecture limitations.

In this context, three different metrics are used to evaluate the computational complexity of the models in this work. The metrics are training time, prediction time, and model size. The first is defined as the time taken to a model to learn (determine) the values of its parameters given a set of labeled examples and the prediction time corresponds to the time the trained model takes to assign a class label to a set of unlabeled examples. Finally, the model size corresponds to the hard disk space occupied by the model.

4.4 Discussion

In this chapter, the methods involved in the model construction and validation were presented. Firstly, Section 4.1 was dedicated to discussing the resampling-based validation strategies. The hyperparameter tuning procedures were tackled in Section 4.2. Finally, Section 4.3 stated the evaluation metrics. All of the abovementioned methods are key to obtaining a non-biased model. A predictor specialized in a certain train/test split shall not be capable of generalizing when presented to new data, thus being unable to be deployed to production.

Part III

Contributions

5
Improved Feature Extraction of Guided Wave Signals for Defect Detection in Welded Thermoplastic Composite Joints

This chapter investigates different feature extraction methods and ML modeling paradigms for defect detection in ultrasonically welded thermoplastic composite joints. The main contribution is the development of an approach, based on supervised ML and AR feature extraction, capable of detecting weld defects successfully. We show that feature engineering enhances the effectiveness of the models by at least 72.50%.

5.1
Problem Description

In [70] the authors analyze the propagation of ultrasonic guided waves in ultrasonically welded thermoplastic composite joints, using extracted information of the frequency-dependent received signals, namely energy of the time-domain signal, frequency spectral information, time-of-flight, and cross-correlation features considering the nominal condition. The data was made available at `https://data.4tu.nl/articles` and used in the present work.

The specimens were produced by ultrasonically welding pairs of carbon fiber reinforced polymer plates, whose nominal dimensions were 101.6mm x 25.4mm and 1.62mm thickness, with an overlap of 12.7mm nominal length. The welding process was carried out so that specimens with particular weld conditions were produced. Fourteen specimens were generated with 3 different conditions, as depicted in Fig. 5.1. Batch 1 (B1), is composed of 5 specimens with incomplete welds, Batch 2 (B2) comprehends 5 specimens fully welded, i.e., nominal condition, and Batch 3 (B3) with 4 specimens representing the adherend fiber bundle distortion defect. The experimental setup, adopted for the ultrasonic guided wave testing, is schematically depicted in Fig. 5.2. Piezo-ceramic transducers were used to generate and receive the waves. The excitation signal was a sinusoidal tone-burst with a 10-cycle with a Hanning window amplitude modulation. The transmitter was positioned at the left-hand side of the welded overlap, whereas the receiver was positioned at the right-hand side of the weld, as shown in Fig. 5.2. With this configuration, the received ultrasonic waves propagate through the weld. The excitation signal

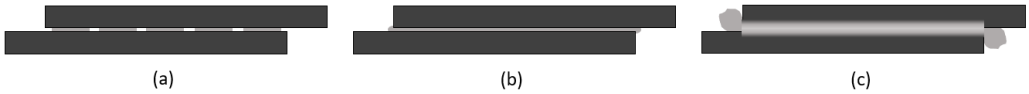


Figure 5.1: Different types of specimens produced with weld defined qualities. (a) specimens with interspersed welded and unwelded areas (Batch 1), (b) specimens with fully welded joints (Batch 2), and (c) specimens with fiber bundle distortion (Batch 3). Batch 2 specimens designate the nominal condition, while batch 1 and 3 represent a defective condition. Figure drawn based on [70].

was centered at four different frequencies, namely 204, 349, 486, and 619 kHz. Further details about the specimen preparation and the experimental setup can be found in [70].

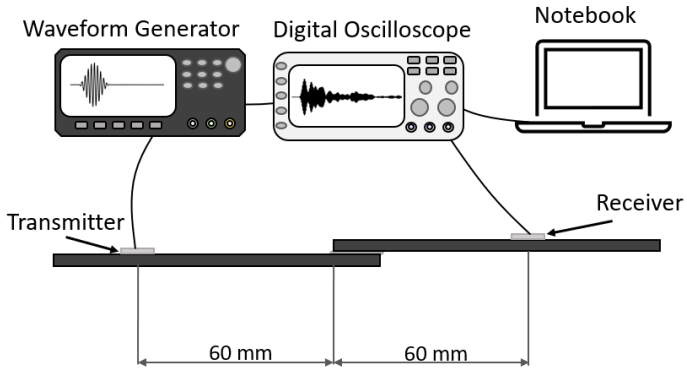


Figure 5.2: Schematic representation of the experimental setup used for ultrasonic guided wave testing. The PZT transducers consisted of discs, 10 mm diameter and 0.4 mm thickness, of APC 850 material. Figure drawn based on [70].

Fig. 5.3 depicts the dispersion curves in the single laminate plate, away from the overlap region, modeled as a 1.62mm thick transversely isotropic plate with elastic constants according to Daggumati et al. [142]. The aforementioned excitation frequencies are represented by vertical dashed lines. As can be seen, virtually only the fundamental modes, namely A0 and S0 modes, can propagate in all the operating frequencies. At the highest operating frequency, namely 619 kHz, the A1 mode is propagating but yet very dispersive. In the overlap region, due to the higher overall thickness, modes of higher order are allowed to propagate, as shown in [70, 143]. The higher-order modes are created in the overlap region by mean of the mode-conversion from the incident fundamental modes when impinging upon the leading edge of the overlap. Within the overlap the wave propagation is complex. Each mode, being from the same type of incident modes or created from mode conversion, propagates

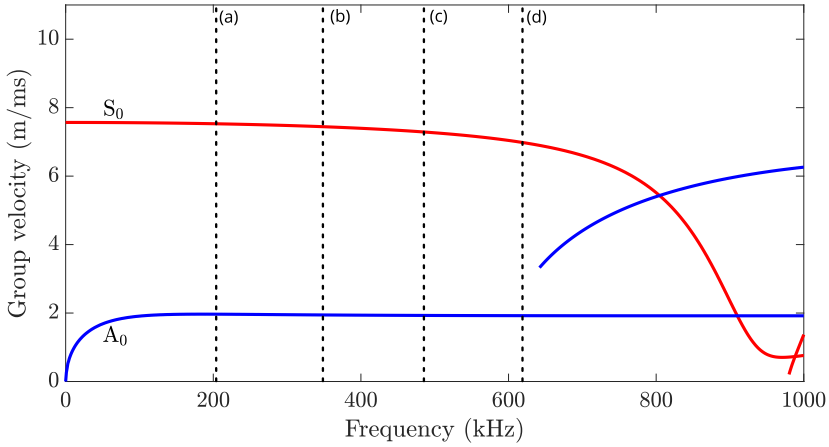


Figure 5.3: Group velocity curves computed with Dispersion Calculator ™. The vertical dashed lines indicates the operation frequencies: (a) 204 kHz, (b) 349 kHz, (c) 486 kHz and (d) 619 kHz.

until the trailing edge of the overlap, where it is partially transmitted to the other single composite plate and partially reflected back to the overlap. Either the transmitted or reflected wavefields can be composed of converted modes. A similar phenomenon occurs when the back-propagating waves impinge on the leading edge of overlap, and so on. This gives rise to a complicated reverberation process. Thus, the signal received is composed of several modes and influenced by the reverberation pattern within the overlap, as shown in Fig. 5.4. This phenomenon is discussed in details in [143].

The basic principle of using guided waves to detect defects in the welding process lies in the fact that the welding condition changes the reverberation of waves within the overlap region. This, in turn, affects the received signal. Although the welding process was conducted in a controlled way, signals corresponding to the same experimental batch considerably differ as shown in Fig. 5.4(a) to (c). As explained by Ochôa et. al. [70], prior to welding, the adherends had to be manually sanded leading to dimensional and mass variability. Therefore, specimens that belong to the same batch, which ideally were supposed to be identical, have different overlap regions and, thus, the received signal differs considerably from specimen to specimen within each batch. Additionally, Fig. 5.4(d) depicts a comparison between the received signal for one specimen of each batch. One can see that, although the signals correspond to specimens of different batches, they can be quite similar. Therefore, extract quantitative information about the presence of a flaw at the weld joint is not straightforward. Hence, the objective of this work is to establish a supervised ML pipeline capable of dealing with intrinsic differences of the raw signals that are not related to the presence of flaws through the

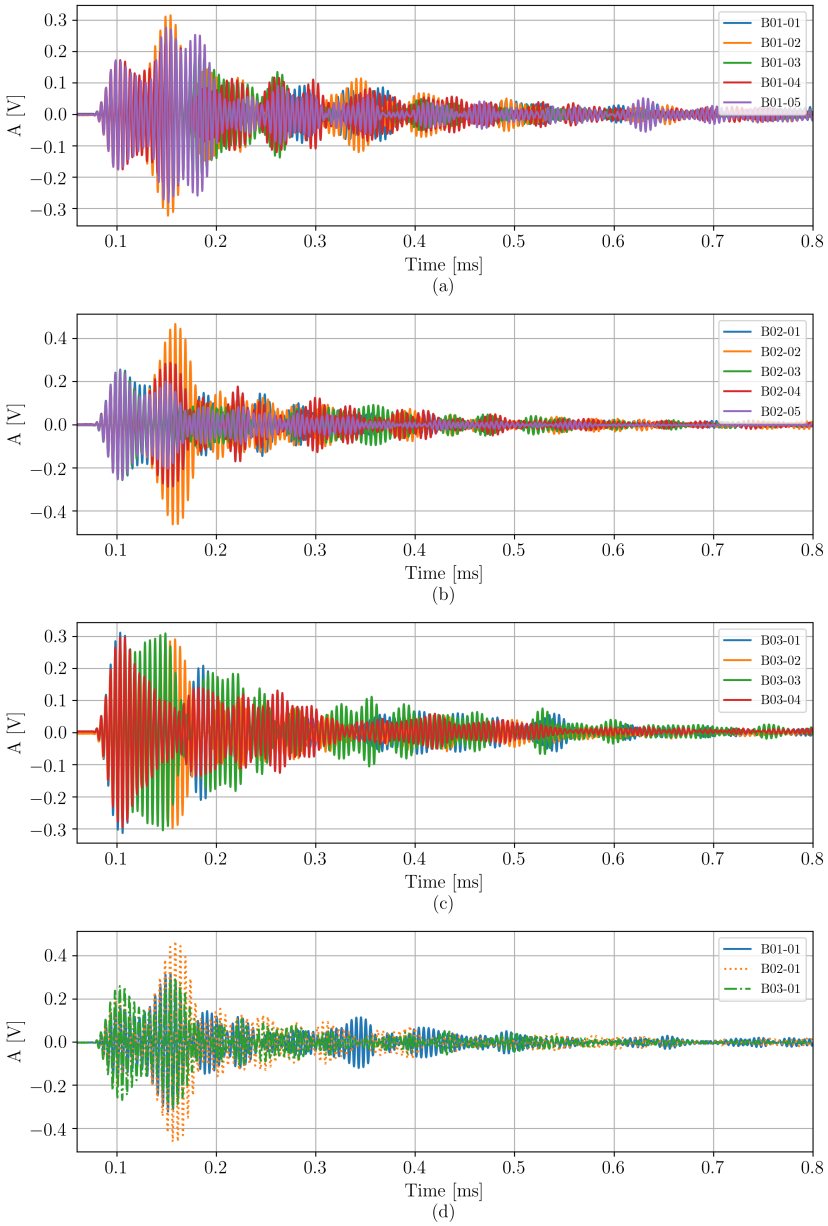


Figure 5.4: UGW signals from all the specimens in (a) Batch 1, (b) Batch 2, (c) Batch 3. (d) depicts a comparison between one specimen of each batch. The UGW signals displayed from (a) to (d) were generated at the excitation frequency of 204 kHz. The signals were labeled according to the convention "B0X-0Y", where X and Y stand for the batch and specimen number, respectively.

employment of a purely data-driven approach to address the defect detection task.

5.2 Modeling Workflow

The modeling workflow proposed in this work to address the defect detection task is depicted in Fig. 5.5. A resampling-based model testing scheme,

namely Monte Carlo cross-validation [144, 145], was adopted to evaluate different supervised learning paradigms and feature extraction strategies. As a consequence of the *No Free Lunch Theorem* [120] instance-based (KNN), tree-based (DT), ensemble (RF) and support vector (SVM) algorithms were considered as candidates. Moreover, since deep learning models require a large amount of processing power [146] and demand a great amount of data to be properly trained [147, 148, 118, 75, 63], only shallow learning models were tested.

The procedure can be summarized as follows. A dataset was created for each one of the feature extraction methods. After feature extraction is processed, the dataset is divided into training and test sets. The models are constructed with a repeated k-fold cross-validation procedure using training data. In the training stage, the hyperparameters are defined by randomly searching [138] for the combination that results in the best model, according to validation metrics. After the architecture is defined with the k-fold cross-validation, the test set assesses the generalization capabilities of the model. Finished the test phase, the dataset is resampled, and the procedure is repeated in a MCCV scheme. It is worth mentioning that, for an unbalanced dataset, it is necessary to adopt a stratified split [149]. The stratification ensures that the test set is representative preserving the proportion of classes in the original dataset. All the resamplings were performed in a stratified fashion.

The proposed framework was conceptualized having in mind that a model can benefit from the data splitting, and thus evaluating it with different train/test splits reduces the uncertainty of the performance estimates [136]. In this context, combining MCCV with CV allows a fair comparison of the modeling paradigms investigated. In addition, the data used in this work is scarce (56 samples), and data acquisition experiments are expensive and time-consuming. Therefore, such a procedure enables the use of small data evaluation resulting in a final model constructed with, ideally, all available data.

5.3 Results

In the present section, the results for the task of damage detection in ultrasonically welded thermoplastic composite joints are presented. Three feature extraction methods are adopted, (a) AR, (b) PCA, and (c) TFD. Furthermore, different supervised learning models were used, being them Support Vector Machines, Decision Trees, k-Nearest-Neighbor and Random Forests. A comparison amongst the diverse feature extraction methods and

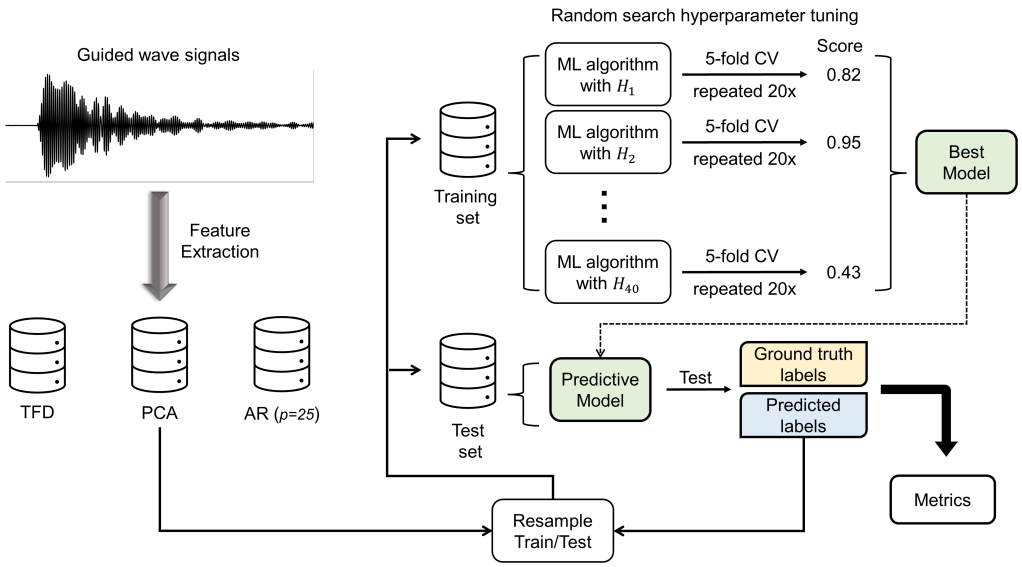
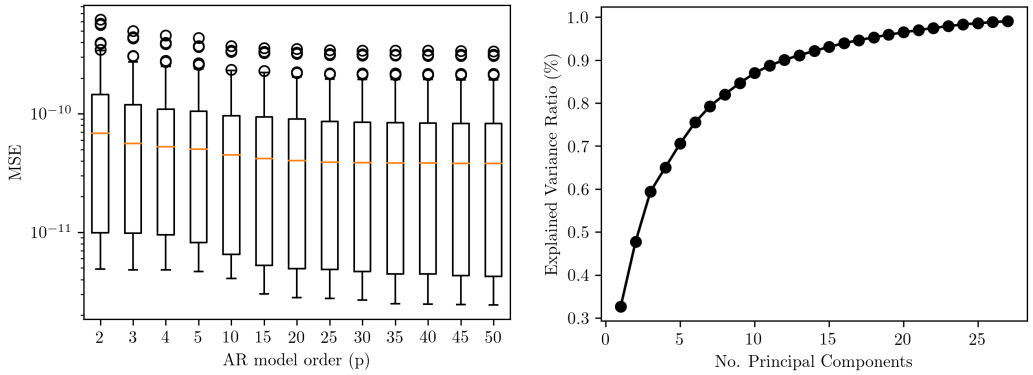


Figure 5.5: Overview of the proposed modeling workflow. Firstly, the features are extracted from the guided wave signals using different methods. After feature extraction, the data is randomly divided into training and test set. The hyperparameters of the ML algorithm are tuned using a procedure that combines randomized search with repeated k-fold cross-validation. The resulting best model is then evaluated with the test set. After storing the predictions, the data is resampled, and the procedure is repeated until reaching a pre-defined number of iterations. Finalized the process, the evaluation metrics are computed.

classifiers is performed. The goal is to highlight the framework most suitable for the problem at hand.

Differently from the TFD approach, both the AR (a) and PCA (b) have hyperparameters that need to be previously defined. For the first, the parameter is the order of the model p. In this work, different p values were tested to determine the best model order. Fig. 5.6 depicts the Mean Squared Error (MSE) obtained by fitting AR models with varied order, ranging between 2 and 50, to the measured guided wave signals. It suggests that the optimal order corresponds to $p = 25$, from which MSE decreases at a significantly lower rate or, in other words, there is no improvement for $p > 25$. Moreover, adopting higher-order AR models would imply increasing the computational cost of the modeling pipeline. Regarding the PCA approach (b), the hyperparameter is the number of selected PCs, and to define it can opt for the explained variance criterion [ref]. In this paper, the PCs were chosen to preserve 99% of the data variance. As shown in Fig. 5.6, 27 components were necessary to achieve the established criterion.

The resampling-based model construction procedure described in section 2 was performed considering 100 iterations with a 10% holdout. For each



(5.6(a)) (5.6(b))

Figure 5.6: AR and PCA hyperparameters definition. MSE statistical properties when varying the AR model order (a), and cumulative explained variance ratio of the PCs. Note that the AR models created do not improve significantly when $p > 25$, and 27 components explain 99% of the variance present in the UGW signals.

resampling iteration, the models were created with a 5-fold cross-validation procedure repeated 20 times. In such a stage, the hyperparameters were tuned by adopting a randomized search with 40 iterations. Table 3 presents the hyperparameters, jointly with their respective distributions, adopted for each strategy tested. Once we evaluate three feature extraction strategies and four supervised learning algorithms, a total of 48,000 models were created during the procedure. The modeling workflow described enables statistical validation of the results as each modeling paradigm was tested on 600 input-output pairs.

Hereafter, the accuracy of supervised learning paradigms tested is investigated. The results regarding the *BAcc* for every holdout realization are presented in Table 5.2. One may notice that, with the TFD approach, none of the models were capable of achieving mean accuracy higher than 0.45. In fact, for 75% of the train/test splits accuracy score was inferior to 0.50, as evidenced by 3rd quartile values. Moreover, the minimum score equal to 0.00 shows us that there were cases where the models misclassified all the test samples. The results for the PCA approach are quite similar to TFD. In this case, the DT,KNN, and RF presented even worse results using PCA features when compared to when using TFD features. The exception was the SVM, for which the PCA provided slightly better results than TFD, as confirmed by the higher mean, median, and 3rd quartile values. Regarding the AR approach, one can see that all models improved their results by at least 72.50% when compared to the TFD and PCA approaches. In addition, the median values show that, for 50% of the train/test splits, the accuracy achieved by the models was higher than 0.67. Lastly, comparing the results of the models, SVM stands out as it

Table 5.1: Hyperparameter settings for the models tested in the present work using randomized search.

Model	Hyperparameter	Distribution/option
SVM	C	Log-uniform $[10^{-1}, 10^{3}]$
	γ	Log-uniform $[10^{-4}, 10^{0}]$
	Kernel	Linear, polynomial, radial basis functions, sigmoid
	Degree (polynomial kernel)	Uniform integer $\{2, 3, ..., 6\}$
KNN	No. of neighbors	Uniform integer $\{3, 5, ..., 21\}$
	Weights	Uniform, distance
	Metric	Euclidean, Manhattan, Minkowski
DT	Maximum depth	$\{4, 8, 12\}$
	Maximum leaf nodes	Uniform integer $\{2, 3, ..., 9\}$
	Minimum samples split	Uniform integer $\{1, 2, ...4\}$
	Maximum features	sqrt, log2
RF	No. of estimators	$\{4, 16, 64\}$
	Maximum depth	$\{3, 5, 7\}$
	Maximum features	sqrt, log2

was the best model with the PCA and AR approaches and the second-best with TFD. SVM jointly with AR features achieved perfect performance in 50 out of the 100 train/test splits and a mean $BAcc$ higher than 0.90.

The analysis of the results can be enriched by the confusion matrices depicted in Fig. 5.7. They provide information, not only about the classifier performance but also the type of errors committed. One can notice that for both TFD and PCA features (see Fig. 5.7a to 5.7h), the models presented a high misclassification for all classes, with the leading source of error being B3 samples. The same was not observed for the AR approach (see Fig. 5.7i to 5.7l). Indeed, the B3 class was a minor source of error for the models using the autoregressive features. Such result is following the fact that adherend fiber bundle distortion (B3) affects the measured signals more than unwelded areas (B1) [70] that is, it is easier to differentiate 'B3' cases from the others. Regarding the model that presented the best performance, SVM jointly with AR, the confusion matrix (see Fig. 5.7l) shows B1 samples were classified as B2 and vice versa. Additionally, all the misclassified B3 samples were assigned to the B1 class, which is positive as it indicates the model did not predict the most severe defective samples as non-defective.

The results have shown that, when compared to the other approaches herein presented and the ones previously found in the literature [70], AR features jointly with the SVM stands out as the most suitable strategy for the problem of damage detection in ultrasonically welded thermoplastic composite joints using ultrasonic guided waves. More specifically, the framework combining AR and SVM achieved perfect classification for 50% of the holdout tests performed.

Table 5.2: *BAcc* for all resampling iterations and modeling paradigms tested. The results are separeted according to the feature extraction approach, and the models are organized in ascending order of *BAcc*. It is possible to notice that the AR approach outperformed both TFD and PCA, and the SVM stands out as the model with best overall performance.

Feature	Model	Mean	Std Dev	Min.	1st quartile	Median	3rd quartile	Max	% *BAcc* increase*	
TFD	RF	0.3317	0.1843	0.0000	0.1667	0.3333	0.5000	0.8333	-	-
	DT	0.3533	0.1485	0.0000	0.3333	0.3333	0.5000	0.6667	-	6.51
	SVM	0.4050	0.1957	0.0000	0.3333	0.3333	0.5000	1.0000	-	14.63
	KNN	0.4483	0.1736	0.0000	0.3333	0.5000	0.5000	0.8333	-	28.89
PCA	RF	0.2833	0.1749	0.0000	0.1667	0.3333	0.3333	0.6667	-	-
	DT	0.2850	0.1541	0.0000	0.1667	0.3333	0.3333	0.6667	-	0.60
	KNN	0.4033	0.1776	0.0000	0.3333	0.3333	0.5000	0.8333	-	42.36
	SVM	0.4717	0.1866	0.0000	0.3333	0.5000	0.6667	0.8333	-	66.50
AR	DT	0.6550	0.2028	0.1667	0.5000	0.6667	0.8333	1.0000	85.40	-
	KNN	0.7733	0.1545	0.3333	0.6667	0.8333	0.8333	1.0000	72.50	18.06
	RF	0.8267	0.1534	0.3333	0.7917	0.8333	1.0000	1.0000	133.99	26.21
	SVM	**0.9033**	**0.1189**	**0.5000**	**0.8333**	**1.0000**	**1.0000**	**1.0000**	91.50	37.91

Note: * The values on the left represent the increase with respect to the worst case of the respective model, independently of the feature extraction approach. The right-hand values represent the increase compared to the worst model with the same feature extraction approach.

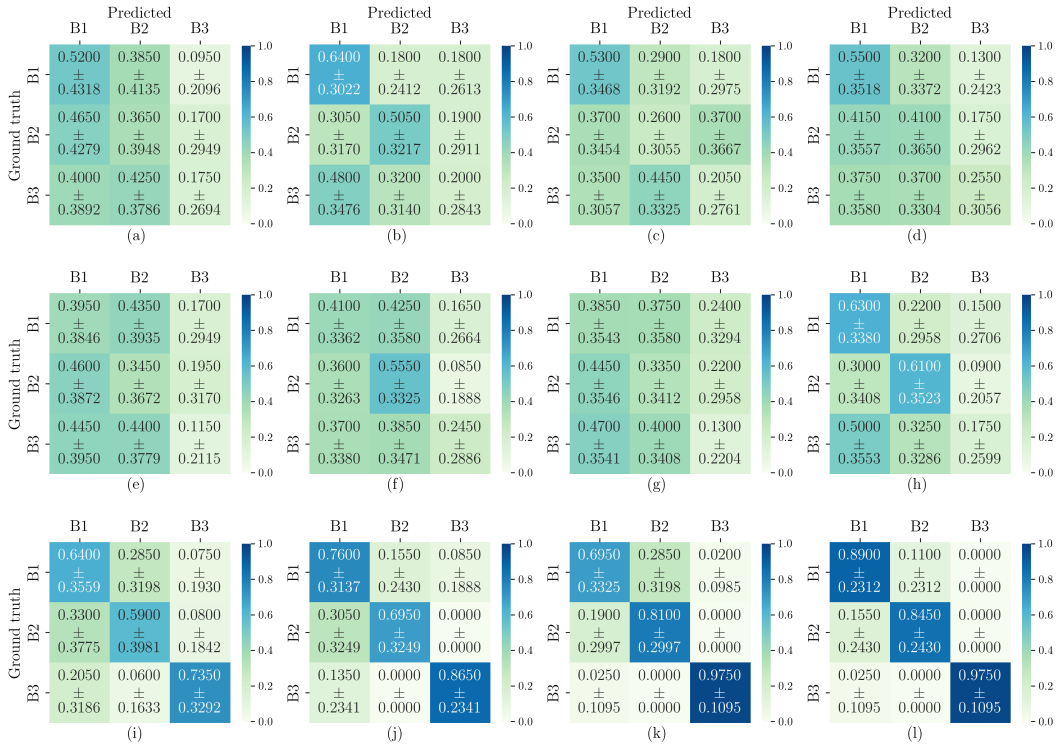


Figure 5.7: Confusion matrices obtained throughout the test phase. From top to bottom we have the matrices for the TFD, PCA, and AR approaches. It can be seen that the accuracy has been greatly improved when AR features were adopted.

5.4 Discussion

In this chapter, several supervised learning paradigms and feature extraction approaches were compared. A resampling-based model construction workflow, combining CV, MCCV, and Random Search was adopted. The proposed framework allowed the use of small data evaluation, without sacrificing the model construction. Moreover multifrequency data was used, maximizing the data generation process. Results have shown that the adoption of proper feature extraction, based on system identification techniques, improved the classifier accuracy by at least 72.50%. Showing that it is beneficial to adopt general feature engineering from machine learning, shall the model be less interpretable in detriment to increasing the classifier accuracy.

6
Efficient Pipeline Crack Detection with Ultrasonic Data and Machine Learning

In this case study, we focuses on developing an efficient ML model construction workflow for an embedded diagnostic system to perform pipeline welding inspection using ultrasonic imaging. First, an extensive comparison between shallow and deep learning architectures is performed, addressing efficiency and computational performance to assess the best compromise in terms of accuracy and hardware consumption for the task at hand. Results show that we were able not only to provide more accurate results when compared to previous works with holdout accuracy greater than 99%, but also to provide a cost-effective model for embedded system implementations.

6.1
Problem Description

In [81], the authors developed a deep learning framework to interpret ultrasonic data generated by inspecting a butt-weld in a stainless steel pipe. A data augmentation step was adopted to suppress the scarcity of representative data, and the performance of the ML model was compared with human performance, previously explored in [150]. Data was made available at `https://github.com/iikka-v/ML-NDT` and used in the present work.

The specimen under investigation consisted of a 30 mm wide U-groove weld in an austenitic 316L stainless steel pipe, with 40 mm thickness and an outer diameter of around 400 mm. The test piece contained three circumferential planar cracks, with depths 1.6, 4.0, and 8.6 mm, introduced near the weld root with thermal fatigue. The ultrasonic data was acquired with a Transmission Receive Shear (TRS) phased array scan, using matrix probes with central frequency at 1.8 MHz. A curved wedge was used, and the coupling between the transducer and the pipe was ensured through a feedwater system. Only one scan angle was used, namely 45°, and it was chosen since it is one of the most common angles used for weld inspection, and it was the one that enabled better visualization of the cracks. The ultrasonic data was acquired in the B-scan format. Fig. 6.1 depicts a schematic representation of the experimental setup. In [151, 81] it is possible to obtain further details about the specimen

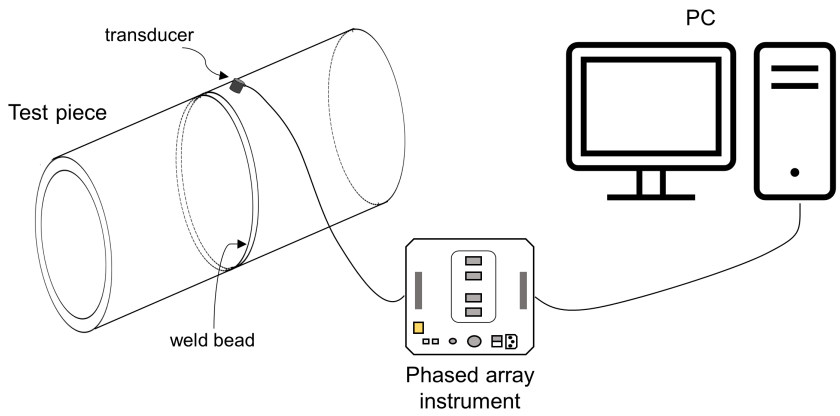


Figure 6.1: Experimental setup used for phased array ultrasonic tests. The phased array instrument used was a Zetec Dynaray 64/64PR-Lite and the probes were the Imasonic 1.5 MHz 1.5M5x3E17.5-9 with a ADUX577A wedge. Drawn based on [151, 81].

and the test setup.

The data were analyzed, and crack indications were identified and extracted from it, providing a defect-free B-scan containing 454 A-scans with 5058 samples. In order to include only the area around the weld, each A-scan was cut, resulting in a defect-free B-scan with 454 x 454 pixels. After that, a resampling procedure was adopted, downsampling the 454 x 454 pixels data to 256 x 256 pixels. In a data augmentation procedure, the extracted flaw signals were introduced to the unflawed scan, with a modified crack size, at random locations. Additionally, the background was changed by random flip to impose data variation and prevent the risk of background memorization.

The dataset is comprised of 20,010 ultrasonic B-scans with 256 x 256 pixels, representing different conditions and different crack sizes. Fig. 6.2 provides a graphical description of the dataset. It is possible to notice that roughly 60% of the data represent a damaged state. Besides that, one can see that the augmented data covers a wide range of crack sizes, which is far more representative than the experimental data that contained only three different depths.

This chapter aims to develop an efficient ML framework to address the problem of crack detection in a butt-weld joint in an austenitic stainless steel pipe. This may be framed as a classification problem, which consists in creating a model that is able to map an observation to an estimated discrete-variable class. In other words, we have $F : \boldsymbol{x} \mapsto y$, where F represents the mathematical model learned through data, $\boldsymbol{x}$ is an UT-image used as input in F, and $y \in$ { "defective", "non-defective"} represents the predicted output which is used to compare with measured data and improve the model through supervised

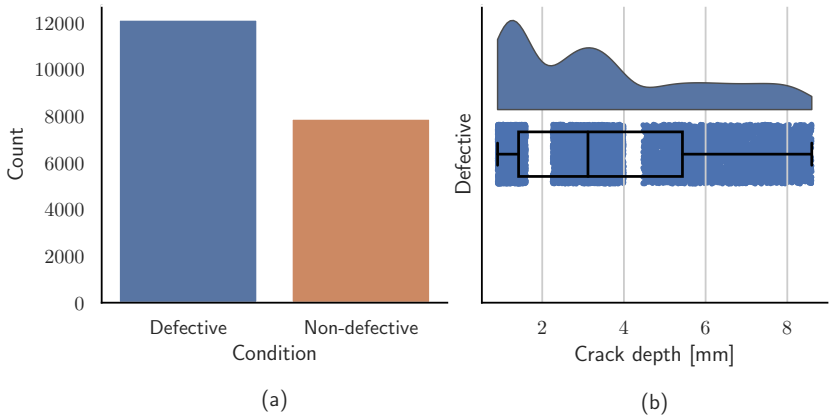


Figure 6.2: Dataset description. (a) depicts the distribution of the samples according to the condition of the weld joint. (b) depicts the distribution of the defective samples according to the depth of the crack.

learning.

6.2 Modeling Workflow

Figure 6.3 summarizes the modeling workflow adopted in the present paper. It is divided in three phases, namely (i) *data preprocessing*, (ii) *feature extraction*, and (iii) *model construction* which are detailed below.

Since the original UT-images represents high-dimensional data, a resampling procedure is proposed as the *data preprocessing* step. The aim is to discard potentially irrelevant information, allowing faster data processing in the ML pipeline. The ultrasonic B-scans, with an original resolution of 256 x 256 pixels, were resized to 128 x 128 pixels resolution. The new resolution was chosen so that information was visually preserved after image resizing. In Fig. 6.3 we also show on the left some examples of UT-images. Distinguishing nominal from defective samples is nontrivial, especially for cracks with small depth. Austenitic stainless steel welds are coarse-grained, heterogeneous, and anisotropic material [152]. Additionally, variations in acoustic impedance result in the presence of noisy signals that hinder flaw identification, which justifies the need for an automatic diagnostic tool. Therefore, extracting quantitative information on the presence of flaws at the weld joint by simply visually inspecting the images is not straightforward.

In this context, shallow learning models require an efficient *feature extraction* procedure to reduce the number of variables and thus enable model construction. Therefore, after preparing the flattened and stacked images, the feature extraction method can be applied. In image processing, PCA is a commonly used method, and the analysis devised has shown that this technique

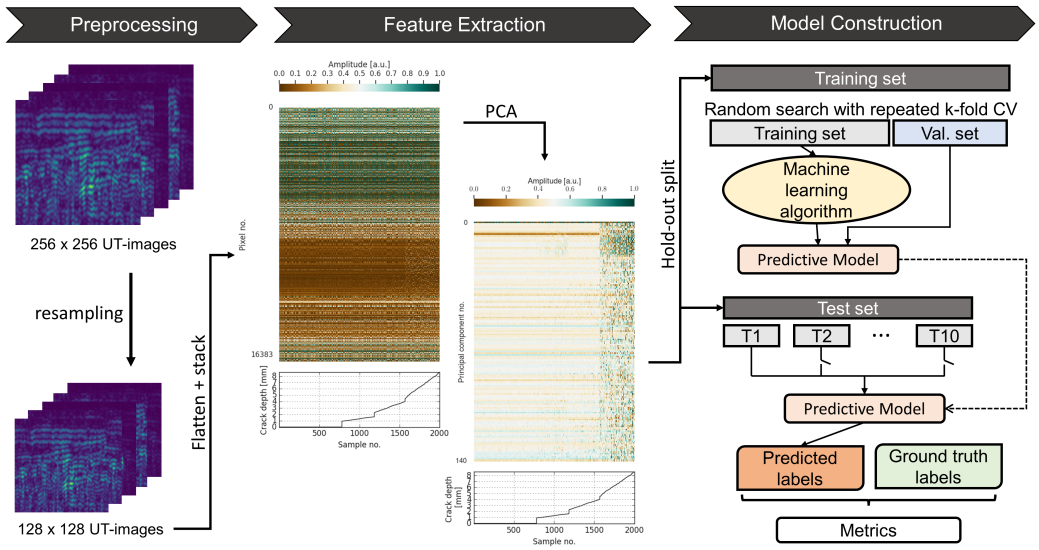


Figure 6.3: Overview of the model construction workflow. In the preprocessing phase, 256 x 256 ultrasonic Images (UT-images) are resized to 128 x 128 pixels, flattened into vectors and stacked into a data matrix. After that, features are extracted from the data by applying the PCA. In the last phase, the data are divided into training and test sets according to a hold-out approach. The training set is used to tune the hyperparameters of the algorithms using a random search with repeated k-fold Cross-validation. After training, ten different subsets of the test data are used to evaluate the performance of the predictive model.

is suitable for the problem at hand. The feature extraction step consists thus in applying the PCA to the data matrix generated after the preprocessing step.

Finally, in the *model construction* phase, a ML model is built according to the data at hand. We use a hold-out dataset for testing the model. During the training process, many parameters related to the optimization problem are tuned. However, hyperparameters cannot be learned by common optimization methods. Defining them is a challenging task since they directly affect the performance of the model. The random search consists of evaluating a specific number of arbitrary combinations and has been shown an efficient option [138] in the case of hyperparameter optimization [153, 154]. It can be combined with a cross-validation procedure, where the training set is divided into k equally-sized parts (k-folds), and the model is fit k times. For every fitting round, $k - 1$ folds are used for training, and one fold is kept apart for validation [135]. The procedure aims at finding the hyperparameters that result in better performance in terms of generalization. In summary, after the feature extraction, the dataset is divided according to a hold-out approach. The resulted training set is used to fine-tune the hyperparameters and define the ideal architecture for each ML model investigated. After the predictive

model is defined, the test set is used to assess the generalization capabilities. In order to enable statistical validation of the results, the model is tested with ten different subsets derived from the test set.

6.3 Results

In the present section, we provide an extensive comparison in terms of model accuracy and hardware use between a deep neural network [81] and the method proposed herein to detect cracks in a butt-weld in an austenitic stainless steel pipe. The compromise between classification and computational performance is evaluated to define which model architecture is the most suitable for deployment in the embedded system. Finally, the model is embedded in hardware in order to evaluate its performance as an online monitoring tool.

Dimensionality reduction is performed using PCA to extract features from the downsampled images. In this case, the number of PCs was chosen to represent 95% of the variance, resulting in a feature vector composed of 141 PCs. This fact represents less than 1% of the original dataset, which indicates that most information in the images is composed of linearly redundant features. The extracted PCs also highlights the differences among samples, especially when the crack depth is higher than 5 mm. Examining the PCA quantitatively, one may conclude that it is suitable for the problem at hand.

The data were then divided into training and test sets with a hold-out approach. Five different splits were considered varying the test set size from 10% to 50% of the 20,010 UT-images. A randomized search with a repeated stratified k-fold procedure was used to fine-tune the hyperparameters of the shallow models. In this procedure, 40 parameter settings were sampled, and the repeated 5-fold division was made with 20 repeats. Thus, in total, 1,200 models were created for selecting the hyperparameter and evaluated in 20 disjoint validation sets sampled randomly from the training set. The results in terms of predictive model performance for the 30% hold-out split are presented in Table 6.1. We have also tested with 10, 20, 40, and 50 % hold-out datasets, whose results are omitted for sabe of brevity as similar conclusions can be drawn for all cases. The CNN [81], DT, GNB, KNN, LGR, MLP, and SVM models were tested using scikit-learn [155]. Their hyperparameter distributions are picked according to each model type. GNB had the worst overall performance. Even though the model achieved 100% precision for all splits, the recall score was considerably lower, bringing accuracy and F1 down. That is, all samples that GNB classified as defective were defective, but a

large number of defective samples were classified as non-defective, which is an undesirable feature. On the other hand, KNN was the model with the best overall performance, achieving perfect, or nearly perfect, scores for every metric. Additionally, both KNN, LGR, and SVM either outperformed or achieved results close to the deep CNN proposed in [81]. Additionally, the results are similar for every hold-out approach. Although the GNB achieved the worst results, the MLP was the model that presented the widest distribution. That is, the performance of the MLP strongly depends on the data it is presented to, which indicates that it fails to generalize well. Moreover, KNN, LGR, and SVM were the ones with the most concise and narrowest error distributions, which indicates their robustness.

We also compared the computational efficiency of each model by calculating the training time, the prediction time, and the model size as shown in Table 6.1. For the sake of brevity, only the results for the 30% hold-out are displayed, as these metrics are proportional to the size of the hold-out dataset. Considering the models that achieved the highest classification performance, namely KNN, LGR, SVM, and CNN, one can see that the deep learning model is by far the slowest to train. The CNN is 198,000 times slower than the KNN, 5,700 times slower than the LGR, and 852 times slower than the SVM. Although the KNN was the fastest model to train, it has the worst results when it comes to the prediction time and size, as the prediction is based on instances. Amongst the four models with the best prediction performance, LGR is the one with the second-lowest training time, the lowest prediction time, and the smaller size. Particularly, both the prediction time and the size are the most important features for real-time embedded use.

Fig 6.4 depicts the performance and computational complexity of the models, providing additional analysis. Considering that classification performance and computational efficiency are important aspects in embedding ML models for online evaluation, LGR stands out as the most suitable model for the problem at hand. We argue that the LGR was able to present better results when compared to the CNN [81] due to the feature extraction procedure adopted. The PCA provided a considerable reduction of the dimensionality of the data and made the samples more easily distinguishable, which finally led to the construction of simpler models.

6.4 Discussion

In this chapter, crack detection in a butt-weld of a stainless steel pipe with ultrasonic phased array data was tackled. Shallow and deep learning mod-

Table 6.1: Metrics obtained as a result of the testing procedure for all the models. These results refer to the 30% hold-out split.

Model	Classification Performance					Computational Performance		
	Acc	*Pr*	*Re*	F_1	*Err*	Training time (s)	Prediction time (s)	Model size (kb)
CNN [81]	0.9935 ± 0.0119	0.9900 ± 0.0180	**1.0000** ± 0.0000	0.9949 ± 0.0092	0.0065 ± 0.0119	5160.9268 ± 55.6928	0.6897 ± 0.0932	885.7240 ± 0.0000
DT	0.9148 ± 0.0154	0.9176 ± 0.0261	0.9458 ± 0.0207	0.9310 ± 0.0106	0.0852 ± 0.0154	0.1305 ± 0.0055	**0.0027** ± 0.0001	6.1080 ± 0.0000
GNB	0.8438 ± 0.0152	**1.0000** ± 0.0000	0.7424 ± 0.0190	0.8520 ± 0.0124	0.1562 ± 0.0152	0.041 ± 0.0014	0.0035 ± 0.0002	8.7230 ± 0.0000
KNN	**1.0000** ± 0.0000	**1.0000** ± 0.0000	**1.0000** ± 0.0000	**1.0000** ± 0.0000	**0.0000** ± 0.0000	**0.0260** ± 0.0018	0.8516 ± 0.0137	8016.2690 ± 0.0000
LGR	0.9990 ± 0.0014	**1.0000** ± 0.0000	0.9984 ± 0.0023	0.9992 ± 0.0012	0.0010 ± 0.0014	0.8922 ± 0.0225	0.0038 ± 0.0018	**5.4930** ± 0.0000
MLP	0.9285 ± 0.0596	0.9756 ± 0.0732	0.9158 ± 0.0704	0.9410 ± 0.0430	0.0715 ± 0.0596	37.9631 ± 0.4418	0.0629 ± 0.0090	32.3408 ± 0.0041
SVM	0.9998 ± 0.0005	**1.0000** ± 0.0000	0.9997 ± 0.0009	0.9999 ± 0.0005	0.0002 ± 0.0005	6.0555 ± 0.1120	0.0140 ± 0.0003	297.6560 ± 0.0000

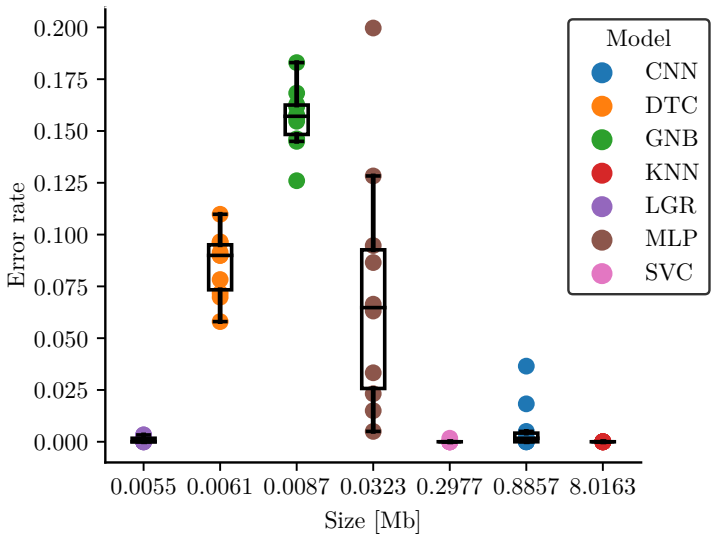


Figure 6.4: Summary comparison between the models combining results from both the performance and computational complexity of the models (in terms of memory size). The chart reports the *Err* distribution against the mean Size of every model for the 30% hold-out. LGR stands out as the model with the best balance between classification performance and memory usage.

els were extensively compared regarding both classification and computational performance. For the latter, several metrics were proposed enabling the evaluation in terms of memory use and processing power. The proposed framework resulted in a lightweight model that outperformed the state-of-art presented in [81]. The results have shown that, the performance of the model can be improved by adopting an approprite dimensionality reduction. The model that provided the best balance between classification and computational performance was the LGR, a linear model.

7
Corrosion-like Defect Severity Estimation in Pipelines Using Convolutional Neural Networks

This chapter aims at developing a deep learning-based framework for estimating the severity of corrosion-like defects in pipelines. The dataset used was generated by finite element simulation of a simple cylinder with different defect depths. A CNN architecture based on LeNet-5 was proposed to address the problem. This architecture was trained and tested 100 times with MCCV. Results show that the proposed architecture achieved good performance, with mean $RMSE$ and R^2 of 0.4448 and 0.9637, respectively.

7.1
Problem Description

Data used to evaluate the ML technique introduced here, were obtained using numerical simulation of a case of study. This consists of a cylindrical structure, mimicking the geometry of an oil and gas pipeline. The pipe presents a defect in the upper region that simulates a damage produced by corrosion. It was simulated the propagation of waves in the z-axis direction. The excitation signal is applied as an external cylindrical source in the central region of the pipe to represent the effect of a transducer. The simulation was performed with a software named Onscale™, which uses the Finite Element Method (FEM).

The simulated structure consists of a 12 mm diameter steel pipe with a thickness of 1 mm and length of 100 mm, see Fig. 7.1. The pipe presents a defect in the upper region, located 40 mm from the center of the cylinder. The length of the defect in the z-axis direction is 4 mm and the width of the defect in the theta direction is 90 degrees. To discretize the domain, a mesh with a square element of size equal to 0.1 mm was used. The excitation signal consists of a ten cycle sine function with a frequency of 500 kHz. The signal is applied as an external cylindrical source in the central region of the pipe, in the direction normal to the pipe circumference. The source thickness in the z-axis direction is 1 mm, which is about the size of the half wavelength of one of the modes being excited by the 500 kHz frequency. The boundary conditions are set as absorber at the ends of the pipe and free in the rest.

Simulations were performed for 100 different defect depth values between

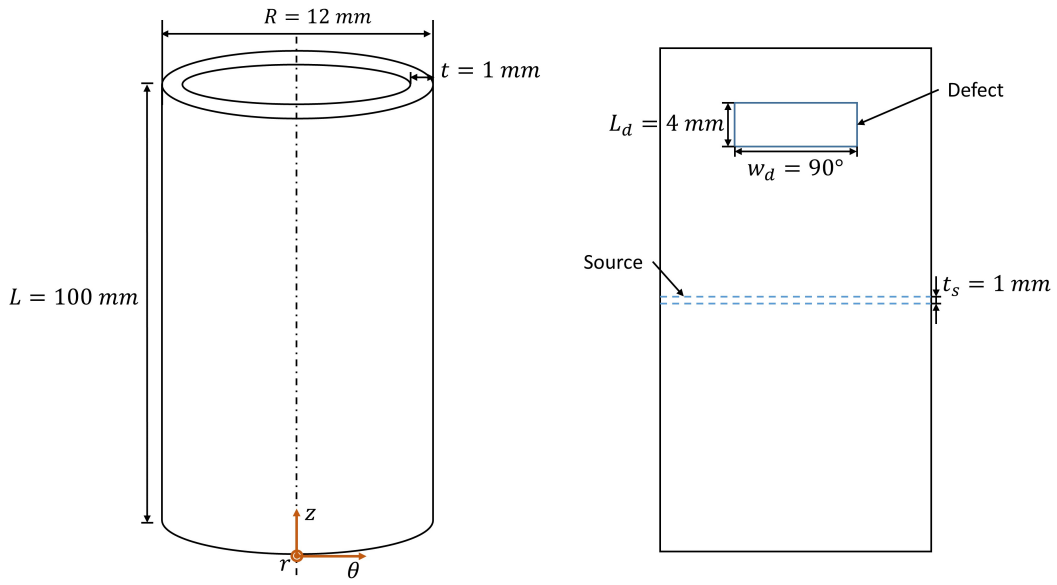
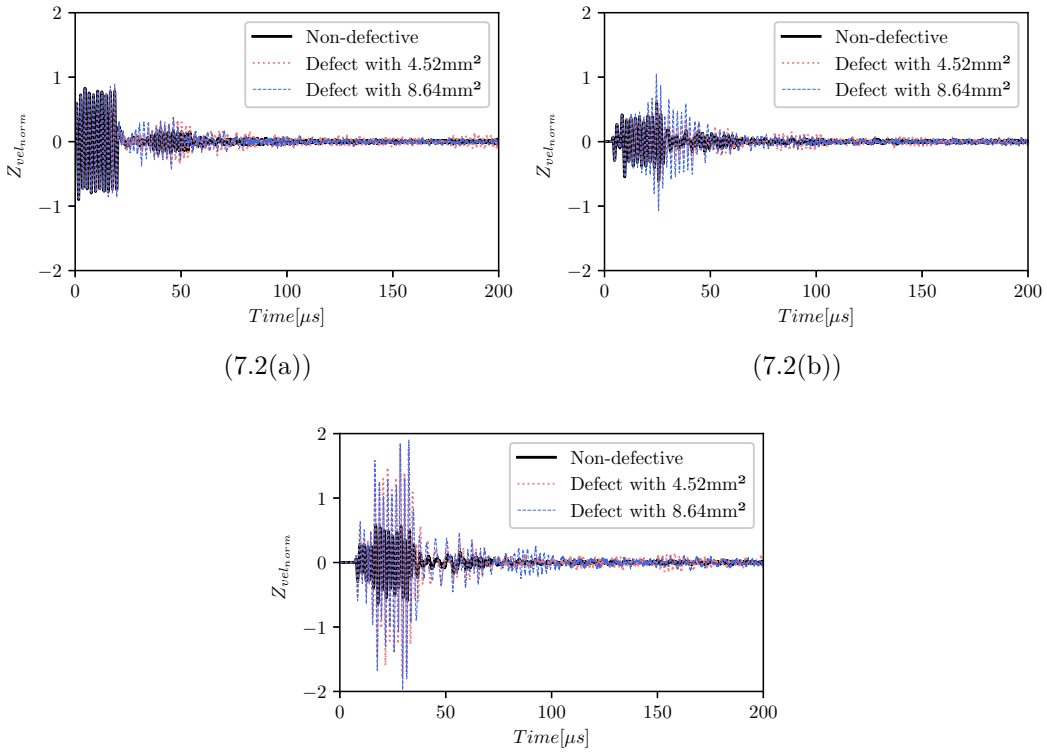


Figure 7.1: Schematic description of the experimental apparatus.

0 mm and 1 mm. The objective of these simulations was to obtain results for different defective conditions representing varied corrosion severity. The numerical simulation computed the z component of the particle velocity wave at the nodes along a line in the cylinder in the z-axis direction passing through the defect. This line starts at the end of the source and ends at the upper end of the pipe. To show the results three points were chosen. The points are located at 1 mm (end of the source), 20.5 mm (about half of the upper part of the pipe), and 40 mm (beginning of the defect). All the distances are measured from the center of the cylinder. The responses acquired at each one of the above mentioned points are depicted in Fig. 7.2. In all figures, three different simulation cases were plotted. Therefore, the black line corresponds to the non-defective case, while the red and the blue lines correspond to defective cases with respective areas equal to 4.52mm^2 and 8.64mm^2. One can notice that the signals representing defective cases are quite distinguishable from the signal representing the non-defective case, specially at 40 mm from the source. On the other hand, signals from different defective cases can be quite similar and thus, estimating the severity of the corrosion is not a trivial task.

7.2 Proposed CNN Architectuture

In this work, a CNN architecture based on LeNet-5 is proposed as means of estimating the area of corrosion-like defects in pipelines. The LeNet-5 is a type of CNN proposed by LeCun et al. [156] in the late 1990s, and it is one of the several widespread architectures developed to date [157, 158, 159, 160,



(7.2(a)) (7.2(b))

(7.2(c))

Figure 7.2: Axial component of the particle velocity wave for different positions along pipe length: at 1mm (a), at 20.5mm (b), and at 40mm (c).

161]. The LeNet-5 was designed to solve the problem of handwritten digits recognition, for which it achieved great success. The structure is comprised of seven layers, including convolutional, pooling, and fully connected ones as depicted in Fig 7.3. In this architecture, the convolutions are done with 5x5 sliding windows, while the subsampling operations are performed with a pool size equal to 2x2. It is also worth noticing that the first set of convolutional and subsampling layers are composed of 6 filters, and this value increases to 16 in the subsequent layers. The dense layers have 120 and 84 nodes, respectively. Finally, the output layer has 10 nodes with Euclidean Radial Basis Function as an activation function.

Despite being firstly employed to solve handwritten digits recognition problems, for which they were very successful, several works have adapted or improved the LeNet-5 architecture aiming at addressing different tasks [162, 163, 164]. Therefore, it has shown to be a flexible algorithm capable of dealing with problems of varied nature after some adaptations.

The structure of the CNN consists of two convolutional layers, two average pooling layers, two fully connected layers, and one output layer, as depicted in Fig 7.4. The CNN requires 2D input data. Therefore, the 1D

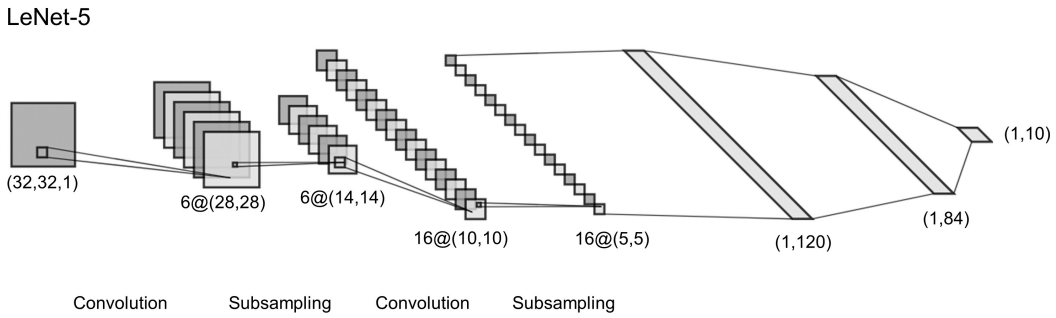


Figure 7.3: Architecture of the LeNet-5 [156].

ultrasonic signals need to be preprocessed before being fed to the CNN. The 1D signals were interpolated, moving from 44,120 to 65,536 data points, and the new signals were converted into 2D grayscale images with a resolution of 256x256 pixels. Fig. 7.5 depicts some examples of images generated from the ultrasonic signals. In the CNN, the automatic feature extraction was performed by two sets of convolutional layers followed by average pooling layers. The structure of the convolutional layers consisted of 30 filters with a 3x3 sliding window (stride of 1) and ReLu activation function, and the shape of the sliding window for the pooling layers was 2x2 (stride of 2). The automatically extracted features were then connected to three dense layers with ReLu activation function. The first dense layer had 56 nodes connected to a second layer with 28 nodes followed by a single node output layer, used to estimate the area of the defect.

7.3 Results

This section is dedicated to evaluating the performance of the proposed CNN architecture. For this purpose, different regression metrics, namely, $RMSE$, R^2, MAE, and $MAXE$, are addressed. A Monte Carlo Cross-Validation approach was adopted to split the data into training and test 100 times. This procedure provides a robust performance estimate, indicating how the model generated from the proposed algorithm is affected by different data partitions.

Table 7.1 summarizes the results obtained throughout the test phase for all data resamples. Both the $RMSE$ and MAE are metrics that measure the average magnitude of the error between the actual and predicted values. The difference is that the $RMSE$ is a quadratic rule that gives high weight to higher errors. The results show that the mean $RMSE$ was higher than the mean MAE indicating variations in the magnitude of the errors. Despite that,

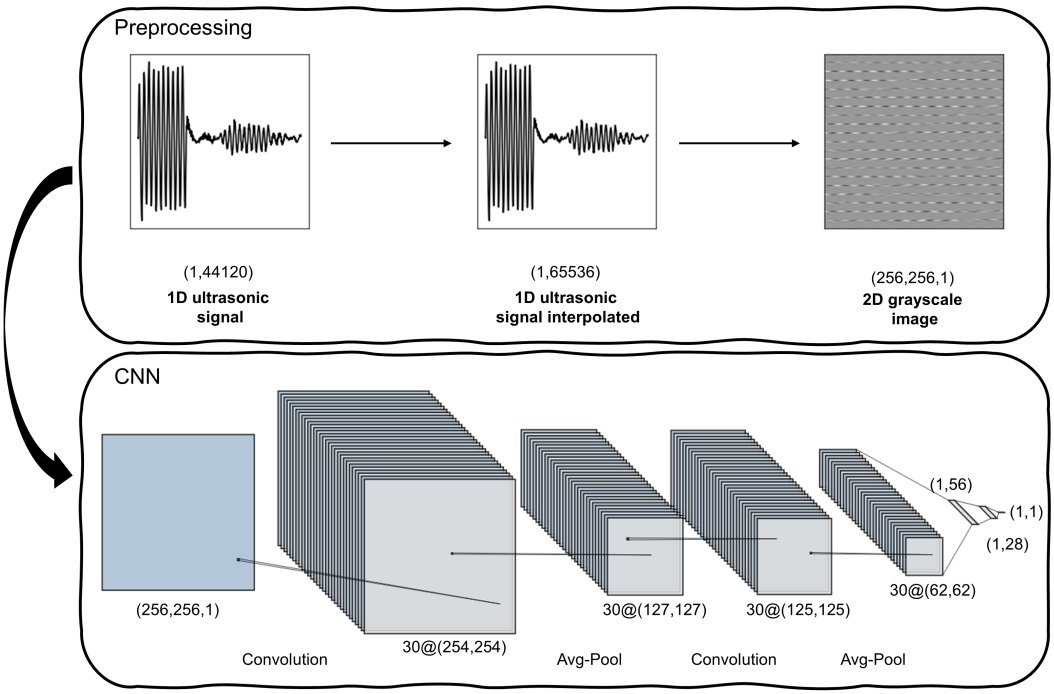


Figure 7.4: Schematic drawing showing the preprocessing procedure and the structure of the proposed CNN architecture for defect estimation.

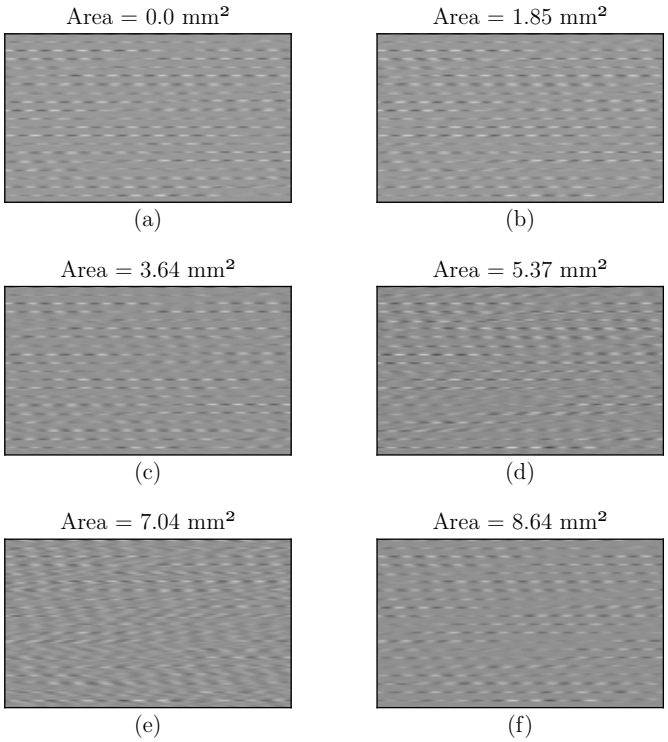


Figure 7.5: Examples of 2D grayscale images generated from the ultrasonic signals. (a) depictsa non-defective sample, and (b) to (f) depicts defective samples with increasing defect area.

Table 7.1: Summary statistics of the regression metrics obtained for all the 100 tests.

	$\boldsymbol{RMSE}$	$\boldsymbol{MAE}$	$\boldsymbol{MAXE}$	$\boldsymbol{R^2}$
Mean	0.4448	0.2719	1.6659	0.9637
Stdev	0.1740	0.0890	0.8160	0.0317

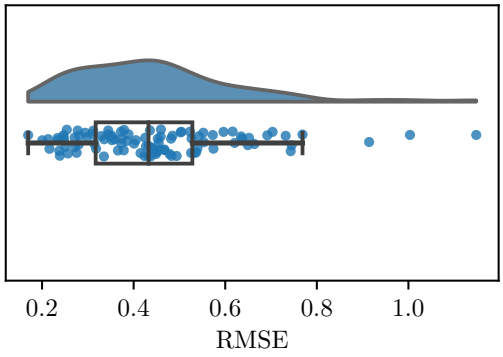
both metrics were close to the ideal value of 0.0. The $MAXE$ represents the maximum error for a set of test samples. The fact that the mean $MAXE$ value was higher than 1.0 indicates that some of the models committed relatively large errors. The R^2 is a measure that represents the goodness of fit of a regression model, and its ideal value of 1.0 is achieved when the predictions are equal to the real values. One can see that the models achieved a mean R^2 value of 0.9637, which is relatively close to the ideal. Therefore, the results presented in Table 7.1 indicate that the proposed architecture is suitable for estimating the area of corrosion-like defects in pipelines.

Further analysis of the results is provided by the raincloud plots [165] depicted in Fig. 7.6. These plots comprise information about all the 100 test procedures, providing an overview of the results. Analyzing Fig. 7.6(a), one can see that, despite three outliers, $RMSE$ values ranged between 0.2 and 0.8, which represents good results. Regarding the R^2, the scenario is even better. Fig. 7.6(b) shows that, for most cases, the R^2 value was higher than 0.95 and sometimes even reached values very close to the ideal value of 1.00.

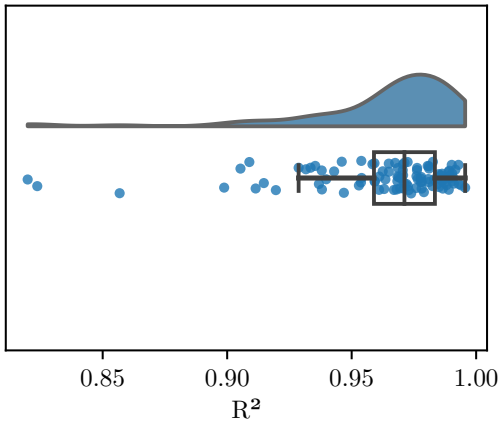
The results of the model with the best performance are depicted in Fig. 7.7. For this case, the $RMSE$, R^2, MAE, and $MAXE$ were equal to 0.1697, 0.9956, 0.1172, and 0.7640, respectively. As seen in the figure, the model provided an excellent approximation between the predicted and actual values. Moreover, the maximum error occurred for a sample where the predicted value was higher than the actual one.

7.4 Discussion

This chapter presented a deep learning-based framework for addressing the problem of corrosion-like defect severity estimation in pipelines. The dataset used was obtained through numerical simulation and it was composed of 100 samples representing varied defect severity. The proposed CNN was tested 100 times in an MCCV fashion, resulting in fair performance estimates. The model achieved low error metrics and mean R^2 of 0.9637. Moreover, for most of the 100 tests performed, the CNN predicted the severity of the defects



(7.6(a))



(7.6(b))

Figure 7.6: Raincloud plots of the $RMSE$ (a), and the R^2 (b).

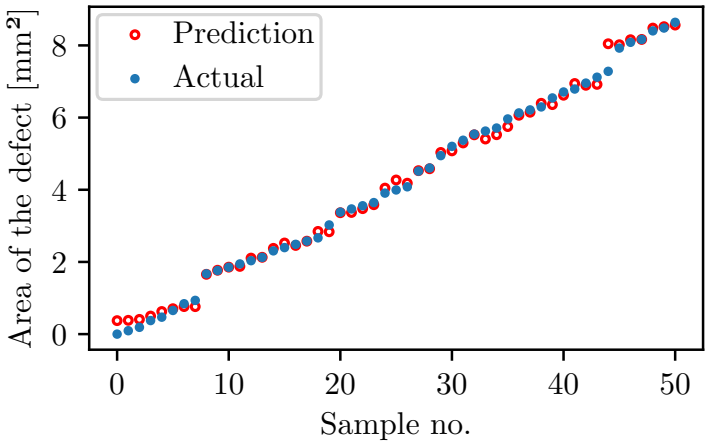


Figure 7.7: Actual and predicted values for the best case amongst all models. The model was capable of predicting the area of the defect with good approximation. It is valuable to notice that, for the case that the error was maximum, the value predicted was higher than the actual area of the defect.

with good approximation as shown by the R^2 values higher than 0.95. Such results indicate the suitability of the proposed CNN to address the task in question.

Part IV

Conclusions

8
Final Considerations

The present work employs ML to solve challenges related to ultrasonic inspection-based evaluation in the OG industry. Several supervised learning paradigms and feature extraction methodologies were adopted to interpret experimental ultrasonic data. Three case studies were developed: (i) a purely data-driven approach to the problem of defect detection in ultrasonically welded thermoplastic composite joints, (ii) an efficient machine learning modeling workflow for an automatic diagnosis system in pipeline welding inspection using ultrasonic imaging, (iii) a deep learning-based framework to estimate the severity of corrosion-like defects in pipelines.

Concerning the first case study, different feature extraction techniques and supervised modeling paradigms were compared for defect detection in ultrasonically welded thermoplastic composite joints. A supervised learning-based approach employing multi-frequency data was developed, enabling the maximization of the data generation process. Moreover, the resampling strategy enabled the use of small data evaluation and provided a fair estimate of the models' performance. Through the proposed approach, we have shown that a proper feature extraction strategy can drastically improve the defect diagnosis as the autoregressive models were able to enhance the performance of all classification-based models tested by at least 72.50%. The best overall results were obtained for the SVM, with the accuracy being improved by 91.50% when using AR features. Moreover, despite the considerable intra-batch differences of the signals, the model was able to achieve satisfactory results, as shown by the mean accuracy of 0.9033.

In the second case study, an extensive comparison amongst shallow and deep learning models employed to detect cracks in a butt-weld of a stainless steel pipe from ultrasonic phased array data was provided. Several metrics regarding the computational performance were proposed to evaluate the suitability of the candidates to serve as an online diagnostic tool. The adoption of a pre-processing procedure can drastically increase the model performance through the appropriate dimensionality reduction. The results show an overall better performance in terms of model prediction and also memory usage when compared to previous works [81]. Specifically, we have shown that with the

proposed approach, it is possible to obtain an efficient model with a smaller size and lower computational requirements. The LGR provided the best balance between classification performance and computational efficiency. This is important as in such application scenarios the computation is usually made distributed geographically; hence the optimization of model deployment is relevant, requiring much lesser computational resources. On the other hand, it also enables measurements at high frequencies to be processed remotely, as sending data would require unfeasible data transfer for such application.

Finally, the third case study addressed a deep learning-based framework for corrosion-like defect estimation in oil pipelines using data obtained through finite element simulations. In total, 100 cases representing different levels of corrosion severity were simulated to enable the evaluation of the deep learning framework. The CNN was trained and tested 100 times, providing rich information about the performance of the proposed architecture. Results indicate the suitability of the developed CNN to the problem at hand.

8.1 Future Works

In light of the results and discussion developed throughout this dissertation, future work can be suggested.

In the context defect detection in ultrasonically welded thermoplastic composite joints, more robust models such as artificial neural networks [72, 88, 68, 110, 166, 167] might be investigated aiming at improving the classification results. Since the dataset is small, another approach to be considered is the use a data augmentation strategy [168, 169, 170] to increase the number of samples, enabling the evaluation of convolutional neural networks [148, 171, 172, 118].

Regarding the second problem, damage severity estimation might be addressed to provide richer information for the decision-making phase. The idea is to develop new models that can predict the depth of the crack present in the welded joint. Another possibility is to test the implementation of the proposed model in dedicated hardware platform like FPGA [173, 174, 175]. Additionally, hardware-in-the-loop tests should also be considered to validate the proposed framework as an online diagnosing tool.

Concerning the third case study, adding new simulation cases varying the position of the defect must be considered as an indication of future work. Another improvement to be considered is the development of an experimental procedure to validate the results obtained. Regarding the machine learning knowledge field, neural architecture search (NAS) [176, 177, 178] algorithms can be used to enhance the results.

Bibliography

[1] DICK, S.. **Artificial intelligence**. Harvard Data Science Review, 1(1), 2019.

[2] MCCARTHY, JOHN. **What is artificial intelligence?** Techical report, Stanford University, Computer Science Department, 2007.

[3] BRAVO, C.; SAPUTELLI, L.; RIVAS, F.; PÉREZ, A. G.; NIKOLAOU, M.; ZANGL, G.; DE GUZMÁN, N.; MOHAGHEGH, S. ; NUNEZ, G.. **State of the art of artificial intelligence and predictive analytics in the e&p industry: a technology survey**. Spe Journal, 19(04):547–563, 2014.

[4] KOROTEEV, D.; TEKIC, Z.. **Artificial intelligence in oil and gas upstream: Trends, challenges, and scenarios for the future**. Energy and AI, 3:100041, 2021.

[5] MITCHELL, T. M.. **Machine Learning**. McGraw-Hill, New York, 1997.

[6] EROFEEV, A.; ORLOV, D.; PERETS, D. ; KOROTEEV, D.. **Ai-based estimation of hydraulic fracturing effect**. SPE Journal, p. 1–12, 2021.

[7] MAO, B.; HAN, L.-G.; FENG, Q. ; YIN, Y.-C.. **Subsurface velocity inversion from deep learning-based data assimilation**. Journal of Applied Geophysics, 167:172–179, 2019.

[8] WOOD, D. A.. **Predicting porosity, permeability and water saturation applying an optimized nearest-neighbour, machine-learning and data-mining network of well-log data**. Journal of Petroleum Science and Engineering, 184:106587, 2020.

[9] BARABOSHKIN, E. E.; ISMAILOVA, L. S.; ORLOV, D. M.; ZHUKOVSKAYA, E. A.; KALMYKOV, G. A.; KHOTYLEV, O. V.; BARABOSHKIN, E. Y. ; KOROTEEV, D. A.. **Deep convolutions for in-depth automated rock typing**. Computers & Geosciences, 135:104330, 2020.

[10] TEMIRCHEV, P.; SIMONOV, M.; KOSTOEV, R.; BURNAEV, E.; OSELEDETS, I.; AKHMETOV, A.; MARGARIT, A.; SITNIKOV, A. ; KOROTEEV, D.. **Deep neural networks predicting oil movement in a development unit**. Journal of Petroleum Science and Engineering, 184:106513, 2020.

[11] HEGDE, C.; GRAY, K.. **Use of machine learning and data analytics to increase drilling efficiency for nearby wells**. Journal of Natural Gas Science and Engineering, 40:327–335, 2017.

[12] TIAN, X.; JIAO, W.; LIU, T.; REN, L. ; SONG, B.. **Leakage detection of low-pressure gas distribution pipeline system based on linear fitting and extreme learning machine**. International Journal of Pressure Vessels and Piping, 194:104553, 2021.

[13] INKPEN, A.; MOFFETT, M. H.. **The global oil & gas industry: Management, strategy and finance**. PennWell Books, LLC, 2011.

[14] DHUTTI, A.; LOWE, S. ; GAN, T.-H.. **Monitoring of critical metallic assets in oil and gas industry using ultrasonic guided waves**. In: ADVANCES IN STRUCTURAL HEALTH MONITORING. IntechOpen, 2019.

[15] WANG, W. D.. **Applications of guided wave techniques in the petrochemical industry**. In: REVIEW OF PROGRESS IN QUANTITATIVE NONDESTRUCTIVE EVALUATION, p. 277–284. Springer, 1999.

[16] VIGGEN, E. M.; MERCIU, I. A.; LØVSTAKKEN, L. ; MÅSØY, S.-E.. **Automatic interpretation of cement evaluation logs from cased boreholes using supervised deep neural networks**. Journal of Petroleum Science and Engineering, 195:107539, 2020.

[17] KAGEYAMA, M.; OKAJIMA, K.; MAESAWA, M.; NONOGUCHI, M.; KOIKE, T.; NOGUCHI, M.; YAMADA, A.; MORITA, E.; KAWASE, S.; KURIBAYASHI, M. ; OTHERS. **X-ray phase-imaging scanner with tiled bent gratings for large-field-of-view nondestructive testing**. NDT & E International, 105:19–24, 2019.

[18] THIYAGARAJAN, J. S.. **Non-destructive testing mechanism for pre-stressed steel wire using acoustic emission monitoring**. Materials, 13(21):5029, 2020.

[19] YIN, W.; TANG, J.; LU, M.; XU, H.; HUANG, R.; ZHAO, Q.; ZHANG, Z. ; PEYTON, A.. **An equivalent-effect phenomenon in eddy current**

non-destructive testing of thin structures. IEEE Access, 7:70296–70307, 2019.

[20] PARK, J.; LEE, J.; MIN, J. ; CHO, Y.. **Defects inspection in wires by nonlinear ultrasonic-guided wave generated by electromagnetic sensors**. Applied Sciences, 10(13):4479, 2020.

[21] BRIZUELA, J.; CAMACHO, J.; COSARINSKY, G.; IRIARTE, J. M. ; CRUZA, J. F.. **Improving elevation resolution in phased-array inspections for ndt**. NDT & E International, 101:1–16, 2019.

[22] CHEN, C. H.. **Pattern recognition in nondestructive evaluation of materials**. In: HANDBOOK OF PATTERN RECOGNITION AND COMPUTER VISION, p. 455–471. World Scientific, 1999.

[23] TAHERI, F.. **18 - advanced fiber-reinforced polymer (frp) composites for the manufacture and rehabilitation of pipes and tanks in the oil and gas industry**. In: Bai, J., editor, ADVANCED FIBRE-REINFORCED POLYMER (FRP) COMPOSITES FOR STRUCTURAL APPLICATIONS, Woodhead Publishing Series in Civil and Structural Engineering, p. 662–704. Woodhead Publishing, 2013.

[24] YAO, S.-S.; JIN, F.-L.; RHEE, K. Y.; HUI, D. ; PARK, S.-J.. **Recent advances in carbon-fiber-reinforced thermoplastic composites: A review**. Composites Part B: Engineering, 142:241–250, 2018.

[25] KARATAŞ, M. A.; GÖKKAYA, H.. **A review on machinability of carbon fiber reinforced polymer (CFRP) and glass fiber reinforced polymer (GFRP) composite materials**. Defence Technology, 14(4):318–326, 2018.

[26] ADAMS, R.; CAWLEY, P.. **A review of defect types and nondestructive testing techniques for composites and bonded joints**. NDT International, 21(4):208–222, 1988.

[27] CAPRIOTTI, M.; KIM, H. E.; SCALEA, F. L. D. ; KIM, H.. **Nondestructive inspection of impact damage in composite aircraft panels by ultrasonic guided waves and statistical processing**. Materials, 10(6):616, 2017.

[28] DE LUCA, A.; CAPUTO, F.; KHODAEI, Z. S. ; ALIABADI, M.. **Damage characterization of composite plates under low velocity impact using ultrasonic guided waves**. Composites Part B: Engineering, 138:168–180, 2018.

[29] TIWARI, K. A.; RAISUTIS, R.; TUMSYS, O.; OSTREIKA, A.; JANKAUSKAS, K. ; JAKUTAVICIUS, J.. **Defect estimation in non-destructive testing of composites by ultrasonic guided waves and image processing**. Electronics, 8(3):315, 2019.

[30] ADEGBOYE, M. A.; FUNG, W.-K. ; KARNIK, A.. **Recent advances in pipeline monitoring and oil leakage detection technologies: Principles and approaches**. Sensors, 19(11):2548, 2019.

[31] BELVEDERESI, C.; THOMPSON, M. S. ; KOMERS, P. E.. **Statistical analysis of environmental consequences of hazardous liquid pipeline accidents**. Heliyon, 4(11):e00901, 2018.

[32] CUNHA, S. B.. **Comparison and analysis of pipeline failure statistics**. Proceedings of the Biennial International Pipeline Conference, IPC, 4:521–530, 2012.

[33] BASTIAN, B. T.; JASPREETH, N.; RANJITH, S. K. ; JIJI, C.. **Visual inspection and characterization of external corrosion in pipelines using deep neural network**. NDT & E International, 107:102134, 2019.

[34] YAACOUBI, S.; MOUNTASSIR, M. E.; FERRARI, M. ; DAHMENE, F.. **Measurement investigations in tubular structures health monitoring via ultrasonic guided waves: A case of study**. Measurement, 147:106800, 2019.

[35] EL MOUNTASSIR, M.; YAACOUBI, S. ; DAHMENE, F.. **Reducing false alarms in guided waves structural health monitoring of pipelines: Review synthesis and debate**. International Journal of Pressure Vessels and Piping, p. 104210, 2020.

[36] ARIYANTO, N. P.; WIDIASTUTI, H. ; PRATAMA, M. Z. S.. **Fabrication of weld defects for phased-array ultrasonic testing**. In: 2018 INTERNATIONAL CONFERENCE ON APPLIED ENGINEERING (ICAE), BATAM, INDONESIA, p. 1–5. IEEE, 2018.

[37] CHEN, Z.; LIU, J. ; QIU, H.. **Solidification crack evolution in high-strength steel welding using the extended finite element method**. Materials, 13(2):1–14, 2020.

[38] JAVADI, Y.; MOHSENI, E.; MACLEOD, C. N.; LINES, D.; VASILEV, M.; MINEO, C.; PIERCE, S. G. ; GACHAGAN, A.. **High-temperature in-process inspection followed by 96-h robotic inspection of intentionally manufactured hydrogen crack in multi-pass robotic**

welding. International Journal of Pressure Vessels and Piping, 189:104288, 2021.

[39] KOŇÁR, R.; MIČIAN, M.; BOHÁČIK, M. ; GUCWA, M.. **Identification of lack of fusion and incomplete penetration in butt weld joint by ultrasonic phased array method and x-ray method**. Archives of Metallurgy and Materials, 64, 2019.

[40] MOON, J.; LEE, J. J. ; LEE, C. H.. **Reheating cracking susceptibility in the weld heat-affected zone of reduced activation ferritic-martensitic steels and the effect of Cr content**. Journal of Nuclear Materials, 542, 2020.

[41] LI, J.; ZHAN, X.; CHEN, S. L. ; JIN, S. J.. **Research on an ultrasonic phased array defect detection system**. Journal of Applied Sciences, 13(8):1376–1381, 2013.

[42] ZHAN, X.; JIN, S.. **Signal analysis method for automatic flaw classification on pipeline girth weld inspection by ultrasonic phased array system**. In: 2009 2ND INTERNATIONAL CONGRESS ON IMAGE AND SIGNAL PROCESSING, TIANJIN, CHINA, p. 1–5. IEEE, 2009.

[43] ZHOU, Y. L.; QIAN, X.; BIRNIE, A. ; ZHAO, X. L.. **A reference free ultrasonic phased array to identify surface cracks in welded steel pipes based on transmissibility**. International Journal of Pressure Vessels and Piping, 168:66–78, 2018.

[44] BERTOVIC, M.. **Human Factors in Non-Destructive Testing (NDT): Risks and Challenges of Mechanised NDT**. PhD thesis, Technischen Universität Berlin, September 2015.

[45] YE, J.; ITO, S. ; TOYAMA, N.. **Computerized ultrasonic imaging inspection: From shallow to deep learning**. Sensors (Switzerland), 18(11), 2018.

[46] MENG, M.; CHUA, Y. J.; WOUTERSON, E. ; ONG, C. P. K.. **Ultrasonic signal classification and imaging system for composite materials via deep convolutional neural networks**. Neurocomputing, 257:128–135, September 2017.

[47] LATÊTE, T.; GAUTHIER, B. ; BELANGER, P.. **Towards using convolutional neural network to locate, identify and size defects**

in phased array ultrasonic testing. Ultrasonics, 115:106436, August 2021.

[48] MIORELLI, R.; KULAKOVSKYI, A.; CHAPUIS, B.; D'ALMEIDA, O. ; MESNIL, O.. **Supervised learning strategy for classification and regression tasks applied to aeronautical structural health monitoring problems**. Ultrasonics, 113:106372, May 2021.

[49] PARK, S. H.; HONG, J. Y.; HA, T.; CHOI, S. ; JHANG, K. Y.. **Deep learning-based ultrasonic testing to evaluate the porosity of additively manufactured parts with rough surfaces**. Metals, 11(2):1–19, February 2021.

[50] HOLGUIN, C. D. V.; AYALA, H. V. H. ; KUBRUSLY, A. C.. **Improved stress estimation with machine learning and ultrasonic guided waves**. Experimental Mechanics, in press, 2021.

[51] SU, Z.; YE, L.. **Identification of Damage Using Lamb Waves - From Fundamentals to Applications**. Springer, 2 edition, 2009.

[52] MITRA, M.; GOPALAKRISHNAN, S.. **Guided wave based structural health monitoring: A review**. Smart Materials and Structures, 25(5):053001, 2016.

[53] ZHANG, Z.; PAN, H.; WANG, X. ; LIN, Z.. **Machine learning-enriched lamb wave approaches for automated damage detection**. Sensors, 20(6), 2020.

[54] C.KUBRUSLY, A.; VON DER WEID, J. P. ; DIXON, S.. **Experimental and numerical investigation of the interaction of the first four SH guided wave modes with symmetric and non-symmetric discontinuities in plates**. NDT & E International, 108:102175, 2019.

[55] C.KUBRUSLY, A.; FREITAS, M. A.; VON DER WEID, J. P. ; DIXON, S.. **Interaction of sh guided waves with wall thinning**. NDT & E International, 101:94–103, 2019.

[56] SHKERDIN, G.; GLORIEUX, C.. **Lamb mode conversion in a plate with a delamination**. The Journal of the Acoustical Society of America, 116:2089–2100, 2004.

[57] PETCHER, P.; DIXON, S.. **Mode mixing in shear horizontal ultrasonic guided waves**. Nondestructive Testing and Evaluation, 32:113–132, 2017.

[58] DIB, G.; AGHAZADEH, A.; MOUSAVI, A.; NAGARAJAIAH, S.; BARANIUK, R. ; DABAK, A.. **Ensembles of novelty detection classifiers for structural health monitoring using guided waves**. Smart Materials and Structures, 27:015003, 2018.

[59] WANG, Q.; MA, S. ; YUE, D.. **Identification of damage in composite structures using gaussian mixture model-processed lamb waves**. Smart Materials and Structures, 27:045007 (11pp), 2018.

[60] FAKIH, M. A.; MUSTAPHA, S. ; ABDUL-AZIZ, A.. **Robust localization and classification of barely visible indentations in composite structures by fusion of ultrasonic damage indices**. Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems, 2:031004–1 to 12, 2019.

[61] SENO, A. H.; KHODAEI, Z. S. ; ALIABADI, M. F.. **Passive sensing method for impact localisation in composite plates under simulated environmental and operational conditions**. Mechanical Systems and Signal Processing, 129:20–36, 2019.

[62] DE CASTRO RIBEIRO, M. G.; KUBRUSLY, A. C. ; AYALA, H. V. H.. **Damage detection in composite plates with ultrasonic guided-waves and nonlinear system identification**. In: 2020 IEEE SYMPOSIUM SERIES ON COMPUTATIONAL INTELLIGENCE (SSCI), p. 2039–2046. IEEE, 2020.

[63] SU, C.; JIANG, M.; LIANG, J.; TIAN, A.; SUN, L.; ZHANG, L.; ZHANG, F. ; SUI, Q.. **Damage assessments of composite under the environment with strong noise based on synchrosqueezing wavelet transform and stack autoencoder algorithm**. Measurement, 156:107587, 2020.

[64] BHUDOLIA, S. K.; GOHEL, G.; LEONG, K. F. ; ISLAM, A.. **Advances in ultrasonic welding of thermoplastic composites: A review**. Materials, 13(6), 2020.

[65] LI, Y.; LIU, Z.; SHEN, J.; LEE, T. H.; BANU, M. ; HU, S. J.. **Weld quality prediction in ultrasonic welding of carbon fiber composite based on an ultrasonic wave transmission model**. Journal of Manufacturing Science and Engineering, 141(8):081010, 2019.

[66] LI, Y.; LEE, T. H.; BANU, M. ; HU, S. J.. **An integrated process-performance model of ultrasonic composite welding based on**

finite element and artificial neural network. Journal of Manufacturing Processes, 56:1374–1380, 2020.

[67] WANG, B.; LI, Y.; LUO, Y.; LI, X. ; FREIHEIT, T.. **Early event detection in a deep-learning driven quality prediction model for ultrasonic welding**. Journal of Manufacturing Systems, 60:325–336, 2021.

[68] LI, Y.; YU, B.; WANG, B.; LEE, T. H. ; BANU, M.. **Online quality inspection of ultrasonic composite welding by combining artificial intelligence technologies with welding process signatures**. Materials and Design, 194:108912, 2020.

[69] SUN, L.; HU, S. J. ; FREIHEIT, T.. **Feature-based quality classification for ultrasonic welding of carbon fiber reinforced polymer through bayesian regularized neural network**. Journal of Manufacturing Systems, 58:335–347, 2021.

[70] OCHÔA, P.; VILLEGAS, I. F.; GROVES, R. M. ; BENEDICTUS, R.. **Diagnostic of manufacturing defects in ultrasonically welded thermoplastic composite joints using ultrasonic guided waves**. NDT & E International, 107:102126, 2019.

[71] LIU, J.; XU, G.; REN, L.; QIAN, Z. ; REN, L.. **Defect intelligent identification in resistance spot welding ultrasonic detection based on wavelet packet and neural network**. The International Journal of Advanced Manufacturing Technology, 90:2581–2588, 2017.

[72] CRUZ, F.; FILHO, E. S.; ALBUQUERQUE, M.; SILVA, I.; FARIAS, C. ; GOUVÊA, L.. **Efficient feature selection for neural network based detection of flaws in steel welded joints using ultrasound testing**. Ultrasonics, 73:1–8, 2017.

[73] XIAOKAI, W.; SHANYUE, G.; LIN, H.; BIN, W. ; XIMING, H.. **Classification of spot-welded joint strength using ultrasonic signal timefrequency features and pso-svm method**. Ultrasonics, 91:161–169, 2019.

[74] WANG, X.; GUAN, S.; HUA, L.; WANG, B. ; HE, X.. **Classification of spot-welded joint strength using ultrasonic signal time-frequency features and pso-svm method**. Ultrasonics, 91:161–169, January 2019.

[75] MUNIR, N.; KIM, H.-J.; PARK, J.; SONG, S.-J. ; KANG, S.-S.. **Convolutional neural network for ultrasonic weldment flaw classification in noisy conditions**. Ultrasonics, 94:74–81, 2019.

[76] MUNIR, N.; PARK, J.; KIM, H. J.; SONG, S. J. ; KANG, S. S.. **Performance enhancement of convolutional neural network for ultrasonic flaw classification by adopting autoencoder**. NDT and E International, 111(4):102218, January 2020.

[77] AMIRI, N.; FARRAHI, G. H.; KASHYZADEH, K. R. ; CHIZARI, M.. **Applications of ultrasonic testing and machine learning methods to predict the static & fatigue behavior of spot-welded joints**. Journal of Manufacturing Processes, 52(October 2019):26–34, 2020.

[78] YAN, Y.; LIU, D.; GAO, B.; TIAN, G. Y. ; CAI, Z. C.. **A Deep Learning-Based Ultrasonic Pattern Recognition Method for Inspecting Girth Weld Cracking of Gas Pipeline**. IEEE Sensors Journal, 20(14):7997–8006, 2020.

[79] GANTALA, T.; BALASUBRAMANIAM, K.. **Automated defect recognition for welds using simulation assisted tfm imaging with artificial intelligence**. Journal of Nondestructive Evaluation, 40(1):1–24, February 2021.

[80] KOSKINEN, T.; VIRKKUNEN, I.; SILJAMA, O. ; JESSEN-JUHLER, O.. **The Effect of Different Flaw Data to Machine Learning Powered Ultrasonic Inspection**. Journal of Nondestructive Evaluation, 40(1):1–13, 2021.

[81] VIRKKUNEN, I.; KOSKINEN, T.; JESSEN-JUHLER, O. ; RINTA-AHO, J.. **Augmented Ultrasonic Data for Machine Learning**. Journal of Nondestructive Evaluation, 40(1):1–11, 2021.

[82] ABDELGAYED, T. S.; MORSI, W. G. ; SIDHU, T. S.. **Fault detection and classification based on co-training of semisupervised machine learning**. IEEE Transactions on Industrial Electronics, 65(2):1595–1605, 2017.

[83] CHEN, T.; HILL, D. J. ; WANG, C.. **Distributed Fast Fault Diagnosis for Multimachine Power Systems via Deterministic Learning**. IEEE Transactions on Industrial Electronics, 67(5):4152–4162, 2020.

[84] GOU, B.; XU, Y.; XIA, Y.; WILSON, G. ; LIU, S.. **An Intelligent Time-Adaptive Data-Driven Method for Sensor Fault Diagnosis in**

Induction Motor Drive System. IEEE Transactions on Industrial Electronics, 66(12):9817–9827, 2019.

[85] YIN, Z.; WANG, L.; ZHANG, B.; MENG, L. ; ZHANG, Y.. **An Integrated DC Series Arc Fault Detection Method for Different Operating Conditions**. IEEE Transactions on Industrial Electronics, 0046(c), 2020.

[86] OJO, O.; LANG, H.; KIM, Y.; HU, X.; MU, B. ; LIN, X.. **A Neural Network Based Method for Thermal Fault Detection in Lithium-Ion Batteries**. IEEE Transactions on Industrial Electronics, 68(5):4068–4078, 2021.

[87] GONG, Y.; LIN, Z.; WANG, J. ; GONG, N.. **Bringing machine intelligence to welding visual inspection: Development of Low-Cost portable embedded device for welding quality control**. IS and T International Symposium on Electronic Imaging Science and Technology, p. 2041–2045, 2018.

[88] SILVA, L. C.; FILHO, E. F. S.; ALBUQUERQUE, M. C. ; SILVA, I. C.. **Segmented analysis of time-of-flight diffraction ultrasound for flaw detection in welded steel plates using extreme learning machines**. Ultrasonics, 102:106057, 2020.

[89] SILVA, L. C.; SIMAS FILHO, E. F.; ALBUQUERQUE, M. C.; SILVA, I. C. ; FARIAS, C. T.. **Embedded decision support system for ultrasound nondestructive evaluation based on extreme learning machines**. Computers and Electrical Engineering, 90:106891, March 2021.

[90] ALOQAILY, A.. **Cross Country Pipeline Risk Assessments and Mitigation Strategies**. Gulf Professional Publishing, 2018.

[91] SINGH, M.; POKHREL, M.. **A fuzzy logic-possibilistic methodology for risk-based inspection (RBI) planning of oil and gas piping subjected to microbiologically influenced corrosion (MIC)**. International Journal of Pressure Vessels and Piping, 159:45–54, 2018.

[92] HEIDARY, R.; GABRIEL, S. A.; MODARRES, M.; GROTH, K. M. ; VAHDATI, N.. **A review of data-driven oil and gas pipeline pitting corrosion growth models applicable for prognostic and health management**. International Journal of Prognostics and Health Management, 9(1), 2018.

[93] ZHANG, S.; LIU, B. ; HE, J.. **Pipeline deformation monitoring using distributed fiber optical sensor**. Measurement, 133:208–213, 2019.

[94] SAMPATH, S.; BHATTACHARYA, B.; ARYAN, P. ; SOHN, H.. **A real-time, non-contact method for in-line inspection of oil and gas pipelines using optical sensor array**. Sensors, 19(16):3615, 2019.

[95] BASTIAN, B. T.; JASPREETH, N.; RANJITH, S. K. ; JIJI, C.. **Visual inspection and characterization of external corrosion in pipelines using deep neural network**. NDT & E International, 107:102134, 2019.

[96] LIU, C.; LI, Y.; FANG, L. ; XU, M.. **New leak-localization approaches for gas pipelines using acoustic waves**. Measurement, 134:54–65, 2019.

[97] XIAO, R.; HU, Q. ; LI, J.. **Leak detection of gas pipelines using acoustic signals based on wavelet transform and support vector machine**. Measurement, 146:479–489, 2019.

[98] YAACOUBI, S.; EL MOUNTASSIR, M.; FERRARI, M. ; DAHMENE, F.. **Measurement investigations in tubular structures health monitoring via ultrasonic guided waves: A case of study**. Measurement, 147:106800, 2019.

[99] REN, L.; JIANG, T.; JIA, Z.-G.; LI, D.-S.; YUAN, C.-L. ; LI, H.-N.. **Pipeline corrosion and leakage monitoring based on the distributed optical fiber sensing technology**. Measurement, 122:57–65, 2018.

[100] LIU, J.; ZANG, D.; LIU, C.; MA, Y. ; FU, M.. **A leak detection method for oil pipeline based on markov feature and two-stage decision scheme**. Measurement, 138:433–445, 2019.

[101] BRUNTON, S. L.; KUTZ, J. N.. **Data-driven Science and Engineering - Machine Learning, Dynamical Systems, and Control.** Cambridge University Press, 1 edition, 2019.

[102] KINGSFORD, C.; SALZBERG, S. L.. **What are decision trees?** Biotechnology, 26:1011–1012, 2008.

[103] GERON, A.. **Hands-on machine learning with Scikit-Learn and TensorFlow : concepts, tools, and techniques to build intelligent systems**. O'Reilly Media, Sebastopol, CA, 2017.

[104] BREIMAN, L.. **Random forests**. Machine Learning, 45:5–32, 2001.

[105] BREIMAN, L.. **Bagging predictors**. Machine Learning, 24:123–140, 1996.

[106] BREIMAN, L.. **Pasting small votes for classification in large databases and on-line.** Machine Learning, 36:85–103, 1999.

[107] GÉRON, A.. **Hands-on machine learning with Scikit-Learn and TensorFlow : concepts, tools, and techniques to build intelligent systems**. O'Reilly Media, 2017.

[108] RUSSELL, S.; NORVIG, P.. **Artificial Intelligence: A Modern Approach**. Prentice Hall, USA, 3 edition, 2010.

[109] HAYKIN, S.. **Neural Networks: A Comprehensive Foundation**. Prentice Hall, 2 edition, 1999.

[110] CHEN, Y.; CHEN, B.; YAO, Y.; TAN, C. ; FENG, J.. **A spectroscopic method based on support vector machine and artificial neural network for fiber laser welding defects detection and classification**. NDT & E International, 108:102176, 2019.

[111] BOSWELL, D.. **Introduction to support vector machines**. 2002.

[112] CHANG, C.-C.; LIN, C.-J.. **LIBSVM: A library for support vector machines**. ACM Transactions on Intelligent Systems and Technology, 2, 2011.

[113] LECUN, Y.; BENGIO, Y. ; HINTON, G.. **Deep learning**. Nature, 521(7553):436–444, 2015.

[114] SHARMA, N.; JAIN, V. ; MISHRA, A.. **An analysis of convolutional neural networks for image classification**. Procedia Computer Science, 132:377–384, 2018. International Conference on Computational Intelligence and Data Science.

[115] LIAO, S.; WANG, J.; YU, R.; SATO, K. ; CHENG, Z.. **Cnn for situations understanding based on sentiment analysis of twitter data**. Procedia Computer Science, 111:376–381, 2017. The 8th International Conference on Advances in Information Technology.

[116] WANG, Z.; YAN, W. ; OATES, T.. **Time series classification from scratch with deep neural networks: A strong baseline**. Proceedings of the International Joint Conference on Neural Networks, p. 1578–1585, 2017.

[117] LIVIERIS, I. E.; PINTELAS, E. ; PINTELAS, P.. **A cnn–lstm model for gold price time-series forecasting**. Neural Computing and Applications, 32(23):17351–17360, 2020.

[118] EWALD, V.; GROVES, R. M. ; BENEDICTUS, R.. **Deepshm a deep learning approach for structural health monitoring based on guided lamb wave technique**. In: SENSORS AND SMART STRUCTURES TECHNOLOGIES FOR CIVIL, MECHANICAL, AND AEROSPACE SYSTEMS 2019, p. 1–16, Denver, Colorado, USA, 2019.

[119] NAIR, V.; HINTON, G. E.. **Rectified linear units improve restricted boltzmann machines**. In: PROCEEDINGS OF THE 27TH INTERNATIONAL CONFERENCE ON MACHINE LEARNING, p. 807–814, 2010.

[120] WOLPERT, D.; MACREADY, W.. **No free lunch theorems for optimization**. IEEE Transactions on Evolutionary Computation, 1(1):67–82, 1997.

[121] KHALID, S.; KHALIL, T. ; NASREEN, S.. **A survey of feature selection and feature extraction techniques in machine learning**. In: 2014 SCIENCE AND INFORMATION CONFERENCE, p. 372–378, 2014.

[122] LIN, Y.-Z.; NIE, Z.-H. ; MA, H.-W.. **Structural damage detection with automatic feature-extraction through deep learning**. Computer-Aided Civil and Infrastructure Engineering, 32(12):1025–1046, 2017.

[123] AMEZQUITA-SANCHEZ, J. P.; ADELI, H.. **Signal processing techniques for vibration-based health monitoring of smart structures**. Archives of Computational Methods in Engineering, 23(1):1–15, 2016.

[124] NIGAM, S.; SINGH, R. ; MISRA, A.. **Efficient facial expression recognition using histogram of oriented gradients in wavelet domain**. Multimedia tools and applications, 77(21):28725–28747, 2018.

[125] SUBASI, A.; GURSOY, M. I.. **Eeg signal classification using pca, ica, lda and support vector machines**. Expert systems with applications, 37(12):8659–8666, 2010.

[126] LAW, M.; JAIN, A.. **Incremental nonlinear dimensionality reduction by manifold learning**. IEEE Transactions on Pattern Analysis and Machine Intelligence, 28(3):377–391, 2006.

[127] JOLLIFFE, I.. **Principal Component Analysis**. Springer, 2 edition, 2002.

[128] JIMÉNEZ, A. A.; MUÑOZ, C. Q. G. ; MÁRQUEZ, F. P. G.. **Dirt and mud detection and diagnosis on a wind turbine blade employing guided waves and supervised learning classifiers**. Reliability Engineering and System Safety, 184:2–12, 2019.

[129] JIMÉNEZ, A. A.; MÁRQUEZ, F. P. G.; MORALEDA, V. B. ; MUÑOZ, C. Q. G.. **Linear and nonlinear features and machine learning for wind turbine blade ice detection and diagnosis**. Renewable Energy, 132:1034–1048, 2019.

[130] GUI, G.; PAN, H.; LIN, Z.; LI, Y. ; YUAN, Z.. **Data-driven support vector machine with optimization techniques for structural health monitoring and damage detection**. KSCE Journal of Civil Engineering, 21(2):523–534, 2017.

[131] BILLINGS, S. A.. **Nonlinear system identification : NARMAX methods in the time, frequency, and spatio-temporal domains**. Wiley, 1 edition, 2013.

[132] LJUNG, L.; SÖDERSTRÖM, T.. **Theory and Practice of Recursive Identification**. The MIT Press, 1983.

[133] ROSE, J. L.. **Ultrasonic Guided Waves in Solid Media**. Cambridge University Press, 1 edition, 2014.

[134] GUYON, I.; GUNN, S.; NIKRAVESH, M. ; ZADEH, L. A.. **Feature extraction: foundations and applications**, volumen 207. Springer, 2008.

[135] HASTIE, T.; TIBSHIRANI, R. ; FRIEDMAN, J.. **The Elements of Statistical Learning - Data Mining, Inference, and Prediction**. Springer, 2 edition, 2008.

[136] KUHN, M.; JOHNSON, K.. **Applied predictive modeling**. Springer, New York, USA, 2013.

[137] GHAWI, R.; PFEFFER, J.. **Efficient hyperparameter tuning with grid search for text categorization using knn approach with bm25 similarity**. Open Computer Science, 9(1):160–180, 2019.

[138] BERGSTRA, J.; BENGIO, Y.. **Random search for hyper-parameter optimization**. Journal of Machine Learning Research, 13:281–305, 2012.

[139] FRIEDRICHS, F.; IGEL, C.. **Evolutionary tuning of multiple svm parameters**. Neurocomputing, 64:107–117, 2005. Trends in Neurocomputing: 12th European Symposium on Artificial Neural Networks 2004.

[140] SPRINGENBERG, J. T.; KLEIN, A.; FALKNER, S. ; HUTTER, F.. **Bayesian optimization with robust bayesian neural networks**. Advances in neural information processing systems, 29:4134–4142, 2016.

[141] DUMONT, V.; GARNER, C.; TRIVEDI, A.; JONES, C.; GANAPATI, V.; MUELLER, J.; PERCIANO, T.; KIRAN, M. ; DAY, M.. **Hyppo: A surrogate-based multi-level parallelism tool for hyperparameter optimization**. arXiv preprint arXiv:2110.01698, 2021.

[142] DAGGUMATI, S.; PAEPEGEM, W. V.; DEGRIECK, J.; XUB, J.; LOMOV, S. ; VERPOEST, I.. **Local damage in a 5-harness satin weave composite under static tension: Part ii – meso-fe modelling**. Composites Science and Technology, 70:1934–1941, 2010.

[143] OCHÔA, P.; FERNANDEZ VILLEGAS, I.; GROVES, R. M. ; BENEDICTUS, R.. **Experimental assessment of the influence of welding process parameters on lamb wave transmission across ultrasonically welded thermoplastic composite joints**. Mechanical Systems and Signal Processing, 99:197 – 218, 2018.

[144] PICARD, R. R.; COOK, R. D.. **Cross-validation of regression models**. Journal of the American Statistical Association, 79(387):575–583, 1984.

[145] XU, Q.-S.; LIANG, Y.-Z.. **Monte carlo cross validation**. Chemometrics and Intelligent Laboratory Systems, 56(1):1–11, 2001.

[146] SIROJAN, T.; PHUNG, T. ; AMBIKAIRAJAH, E.. **Enabling deep learning on embedded systems for IoT sensor data analytics: Opportunities and challenges**. In: 2018 IEEE INTERNATIONAL CONFERENCE ON INFORMATION AND AUTOMATION FOR SUSTAINABILITY (ICIAFS), COLOMBO, SRI LANKA, p. 1–5. IEEE, 2018.

[147] ZHANG, H.; HOU, Y.; ZHAO, J.; WANGB, L.; XI, T. ; LI, Y.. **Automatic welding quality classification for the spot welding based on the hopfield associative memory neural network and chernoff face description of the electrode displacement signal features**. Mechanical Systems and Signal Processing, 85:1035–1043, 2017.

[148] MELVILLE, J.; ALGURI, K. S.; DEEMER, C. ; HARLEY, J. B.. **Structural damage detection using deep learning of ultrasonic guided**

waves. In: PROCEEDINGS OF 44TH ANNUAL REVIEW OF PROGRESS IN QUANTITATIVE NONDESTRUCTIVE EVALUATION, p. 230004–1 to 7, Provo, Utah, 2018.

[149] SECHIDIS, K.; TSOUMAKAS, G. ; VLAHAVAS, I.. **On the stratification of multi-label data**. In: JOINT EUROPEAN CONFERENCE ON MACHINE LEARNING AND KNOWLEDGE DISCOVERY IN DATABASES, p. 145–158. Springer, 2011.

[150] VIRKKUNEN, I.; HAAPALAINEN, J.; PAPULA, S. ; SARIKKA, T.. **Effect of feedback and variation on inspection reliability**. In: PROCEEDINGS OF THE 7TH EUROPEAN-AMERICAN WORKSHOP ON RELIABILITY OF NDE, p. 1–9, September 2017.

[151] KOSKINEN, T.; VIRKKUNEN, I.; PAPULA, S.; SARIKKA, T. ; HAAPALAINEN, J.. **Producing a POD curve with emulated signal response data**. Insight: Non-Destructive Testing and Condition Monitoring, 60(1):42–48, 2018.

[152] CHASSIGNOLE, B.; EL GUERJOUMA, R.; PLOIX, M.-A. ; FOUQUET, T.. **Ultrasonic and structural characterization of anisotropic austenitic stainless steel welds: Towards a higher reliability in ultrasonic non-destructive testing**. NDT & E International, 43(4):273–282, 2010.

[153] ZÖLLER, M.-A.; HUBER, M. F.. **Benchmark and survey of automated machine learning frameworks**. Journal of Artificial Intelligence Research, 70:409–472, 2021.

[154] HE, X.; ZHAO, K. ; CHU, X.. **Automl: A survey of the state-of-the-art**. Knowledge-Based Systems, 212:106622, 2021.

[155] PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M. ; DUCHESNAY, E.. **Scikit-learn: Machine learning in Python**. Journal of Machine Learning Research, 12:2825–2830, 2011.

[156] LECUN, Y.; BOTTOU, L.; BENGIO, Y. ; HAFFNER, P.. **Gradient-based learning applied to document recognition**. Proceedings of the IEEE, 86(11):2278–2323, 1998.

[157] KRIZHEVSKY, A.; SUTSKEVER, I. ; HINTON, G. E.. **Imagenet classification with deep convolutional neural networks**. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, volumen 25, 2012.

[158] SZEGEDY, C.; LIU, W.; JIA, Y.; SERMANET, P.; REED, S.; ANGUELOV, D.; ERHAN, D.; VANHOUCKE, V. ; RABINOVICH, A.. **Going deeper with convolutions**. In: 2015 IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR), p. 1–9, 2015.

[159] HE, K.; ZHANG, X.; REN, S. ; SUN, J.. **Deep residual learning for image recognition**. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, p. 770–778, 2016.

[160] SIMONYAN, K.; ZISSERMAN, A.. **Very deep convolutional networks for large-scale image recognition**. 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings, p. 1–14, 2015.

[161] ZEILER, M. D.; FERGUS, R.. **Visualizing and understanding convolutional networks**. In: COMPUTER VISION – ECCV 2014, p. 818–833. Springer International Publishing, 2014.

[162] WEI, G.; LI, G.; ZHAO, J. ; HE, A.. **Development of a lenet-5 gas identification cnn structure for electronic noses**. Sensors, 19(1), 2019.

[163] ZHANG, C. W.; YANG, M. Y.; ZENG, H. J. ; WEN, J. P.. **Pedestrian detection based on improved LeNet-5 convolutional neural network**. Journal of Algorithms and Computational Technology, 13, 2019.

[164] WAN, L.; CHEN, Y.; LI, H. ; LI, C.. **Rolling-element bearing fault diagnosis using improved lenet-5 network**. Sensors, 20(6), 2020.

[165] ALLEN, M.; POGGIALI, D.; WHITAKER, K.; MARSHALL, T. R.; VAN LANGEN, J. ; KIEVIT, R. A.. **Raincloud plots: a multi-platform tool for robust data visualization**. Wellcome Open Research, 4:63, 2021.

[166] ZHAO, D.; ZHAO, K.; REN, D. ; GUO, X.. **Ultrasonic welding of magnesium–titanium dissimilar metals: A study on influences of welding parameters on mechanical property by experimentation and artificial neural network**. Journal of Manufacturing Science and Engineering, 139:031019, 2017.

[167] SATPATHY, M. P.; MISHRA, S. B. ; SAHOO, S. K.. **Ultrasonic spot welding of aluminum-copper dissimilar metals: A study on joint strength by experimentation and machine learning techniques**. Journal of Manufacturing Processes, 33:96–110, 2018.

[168] RASHID, K. M.; LOUIS, J.. **Times-series data augmentation and deeplearning for construction equipment activity recognition**. Advanced Engineering Informatics, 42:100944, 2019.

[169] LE GUENNEC, A.; MALINOWSKI, S. ; TAVENARD, R.. **Data augmentation for time series classification using convolutional neural networks**. In: ECML/PKDD WORKSHOP ON ADVANCED ANALYTICS AND LEARNING ON TEMPORAL DATA, 2016.

[170] LEMLEY, J.; BAZRAFKAN, S. ; CORCORAN, P.. **Smart augmentation learning anoptimal data augmentation strategy**. IEEE Access, 5:5858–5869, 2017.

[171] LIU, H.; ZHANG, Y.. **Deep learning based crack damage detection technique for thin plate structures using guided lamb wave signals**. Smart Materials and Structures, 29, 2019.

[172] ZHANG, S.; LI, C. M. ; YE, W.. **Damage localization in plate-like structures using time-varying feature and one-dimensional convolutional neural network**. Mechanical Systems and Signal Processing, 147:107107, 2021.

[173] KARA, K.; ALISTARH, D.; ALONSO, G.; MUTLU, O. ; ZHANG, C.. **FPGA-accelerated dense linear machine learning: A precision-convergence trade-off**. In: 2017 IEEE 25TH ANNUAL INTERNATIONAL SYMPOSIUM ON FIELD-PROGRAMMABLE CUSTOM COMPUTING MACHINES (FCCM), NAPA, USA, p. 160–167. IEEE, 2017.

[174] AYALA, H. V. H.; MUÑOZ, D. M.; LLANOS, C. H. ; DOS SANTOS COELHO, L.. **Efficient hardware implementation of radial basis function neural network with customized-precision floating-point operations**. Control Engineering Practice, 60:124–132, 2017.

[175] NGADIUBA, J.; LONCAR, V.; PIERINI, M.; SUMMERS, S.; GUGLIELMO, G. D.; DUARTE, J.; HARRIS, P.; RANKIN, D.; JINDARIANI, S.; LIU, M.; PEDRO, K.; TRAN, N.; KREINAR, E.; SAGEAR, S.; WU, Z. ; HOANG, D.. **Compressing deep neural networks on FPGAs to binary and ternary precision with hls4ml**. 2(1):015001, 2020.

[176] BAKER, B.; GUPTA, O.; NAIK, N. ; RASKAR, R.. **Designing neural network architectures using reinforcement learning**. In: INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS - ICLR, 2017.

[177] ELSKEN, T.; METZEN, J. H. ; HUTTER, F. H.. **Efficient multi-objective neural architecture search via lamarckian evolution**. In: INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS - ICLR, 2019.

[178] REAL, E.; AGGARWAL, A.; HUANG, Y. ; LE, Q. V.. **Regularized evolution for image classifier architecture search**. In: AAAI CONFERENCE ON ARTIFICIAL INTELLIGENCE, 2019.