DOI 10.5162/sensoren2024/A5.4

Machine Learning Model Based on Signal Difference Features for Damage Localization on Hydrogen Pressure Vessel Using Ultrasonic Guided Waves

Houssam El Moutaouakil[1], Jens Prager[2], Andreas Schütze[1], Tizian Schneider[1]
[1]Lehrstuhl für Messtechnik, Universität des Saarlandes, Saarbrücken, Deutschland
[2]Bundesanstalt für Materialforschung- und prüfung (BAM), Unter den Eichen 87, 12205 Berlin, Deutschland

Abstract

Hydrogen is already shaping the future of energy resources; therefore, its storage must adhere to the highest safety standards due to its nature as a highly explosive gas. Consequently, it is imperative to ensure accurate and reliable detection of damage in pressure vessels at an early stage. Despite the existence of various machine learning methods, particularly those based on deep learning, they often face challenges related to interpretability and overfitting. This paper introduces an explainable machine learning (ML) model capable of localizing structural damage on Composite Overwrapped Pressure Vessels (COPVs) using ultrasonic guided waves, that overcomes the introduced drawbacks.

Keywords: Ultrasonic Guided Waves, Machine Learning, Damage Localization, Composite Overwrapped Pressure Vessels, Hydrogen

Introduction

The utilization of pressure vessels for storage, especially for hydrogen as an environmentally friendly energy source [1], is experiencing significant growth. Therefore, ensuring accurate and reliable detection of damage in pressure vessels at an early stage using ultrasonic guided waves (GW) is crucial. Unlike conventional ultrasonic waves, guided waves can achieve a greater inspection range by leveraging the structure of the vessels and exploiting the reflection and absorption behavior of the ultrasonic waves as they propagate over the structure (Fig. 1) [2]. Measurement using guided waves is conducted by sensors that act as transducers, emitting a pitch signal and capturing the catch signal (pitch-catch procedure) [3].

Various machine learning methods have been utilized to investigate guided waves. However, deep learning-based approaches often face challenges concerning interpretability and overfitting [4]. Qian et al. introduced a neural network based on Wav2Vec2.0 for defect localization in array ultrasonic testing and Structural Health Monitoring (SHM) [5]. Nevertheless, issues regarding the interpretability of neural networks persist.

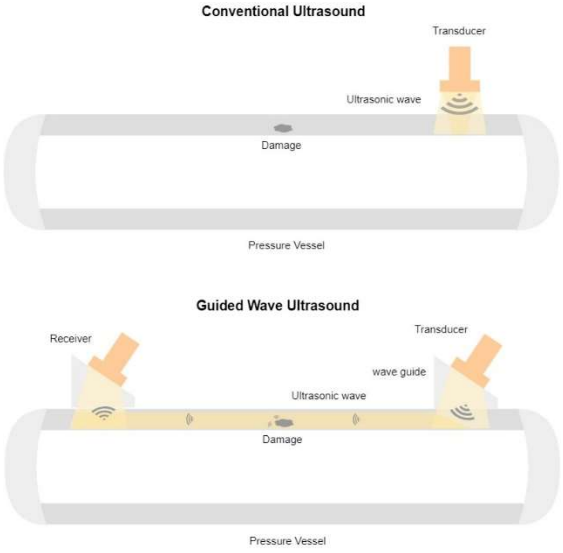


Fig. 1: comparing conventional ultrasound measurement procedure (top) and guided wave measurement (bottom)

In this study, the authors introduce a machine learning algorithm developed using a Matlab-based toolbox created at the Lab for Measurement Technology (LMT-Toolbox) [6]. This model involves benchmarking with several baseline subtraction calculation methods and subsequently applying available feature extraction methods. Then we used the

triangulation method [7] to create targets for the classification. The performed classification predicts those targets, that are used to recover the damage location. Consequently, the authors achieved a model accuracy of nearly 96.5%.

Experimental Setup

To test and validate the developed ML algorithm, we utilized data acquired by the Federal Institute for Materials Research and Testing (BAM) in Berlin (Fig. 2).

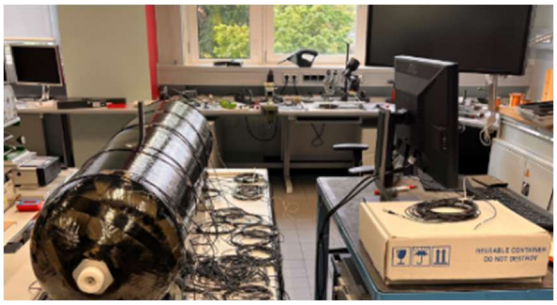


Fig. 2: Measurement setup conducted on pressure vessel by BAM

In the conducted experiment, a Composite Overwrapped Pressure Vessel (COPV) of type AH350-70-4 and manufactured by NPROXX, containing hydrogen, was equipped with a network of 15 sensors (Fig. 3), which served simultaneously as transducers and receivers. Ultrasonic guided waves [2] were emitted and recorded using the pitch-catch procedure [3] with a Verasonics Vantage 64 LF measurement system. As a result, eight experiments and 105 sensor paths per experiment were recorded. This was conducted under a constant temperature of 20°C and a pitch frequency of 130 kHz, as it is suitable for detecting damages down to 4.5 cm, which covers the size of the attached damages.

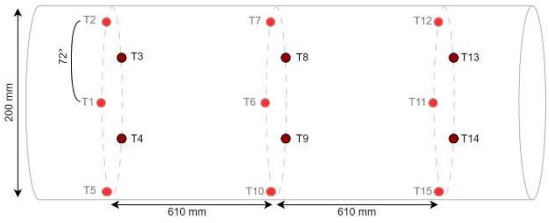


Fig. 3: Arrangement of transducers on pressure vessel where no damage has been induced yet. This setup is utilized to record baseline data

Initially, baseline data were recorded, followed by a set of reversible damages simulated by glued weights of 513g marked as a green dot and 330g marked as blue dot (Fig. 4, Table 1).

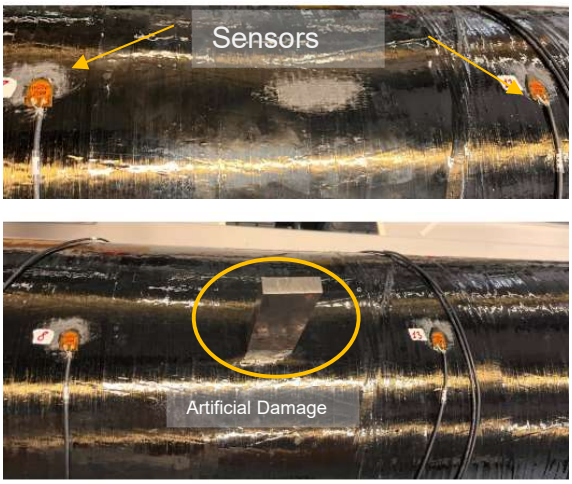
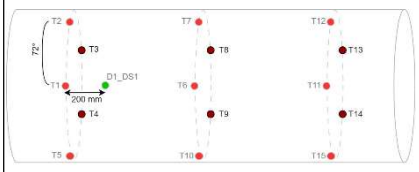
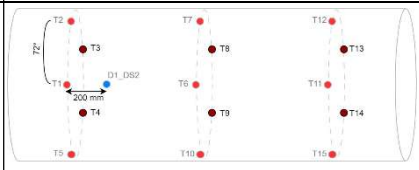
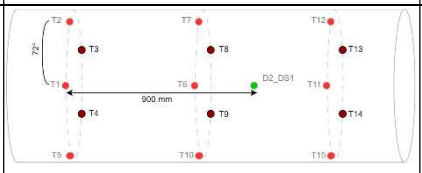
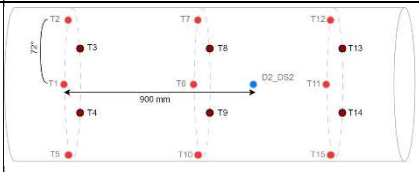
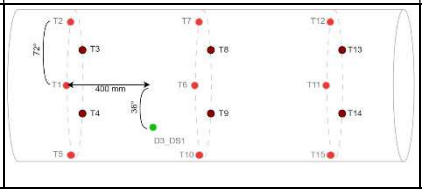
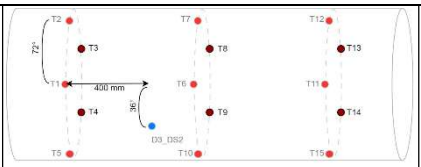
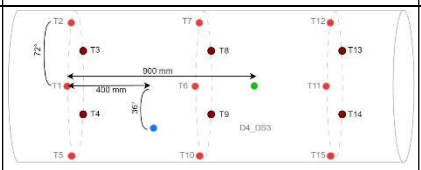
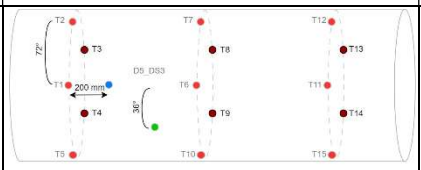


Fig. 4: Scenario involving an undamaged pressure vessel and a scenario featuring a pressure vessel with artificially induced damage

A combination of 8 experiments was then investigated, as shown in Table 1, which details the different damage locations and weights for all performed experiments.

Table 1: various artificial damage experiments to investigate ultrasonic guided wave propagation in the pressure vessel

Damage ID	Damage Location	Weight
1		513g
2		330g
3		513g
4		330g
5		513g

6		330 g
7		513 g + 330 g
8		513 g + 330 g

Approach Description

The developed ML algorithm is based on two key points.

Firstly (Section A.), it involves extracting features that accurately describe the ultrasonic waves in the presence of damage. This process includes calculating element-wise signal subtraction, optionally applying a signal envelope, and then using conventional feature extractors to reduce dimensionality.

Secondly (Section B.), it entails labeling the extracted features with targets that provide a more accurate description of the damage location based on ellipsoid triangulation. Training the ML algorithm with these targets aids in predicting the position of any existing damage.

The entire pipeline of this ML approach is illustrated in Figure 5.

A. Extracting the features

To find a suitable model for damage detection on pressure vessels, the approach involves benchmarking different methods for calculating baseline subtraction (Table 2). Subsequently, various feature extraction methods (Table 3) are applied to the baseline subtraction.

Baseline subtraction is a common procedure in the analysis of ultrasonic waves, enabling the identification of reflection and absorption caused by damages in the setup under analysis [8].

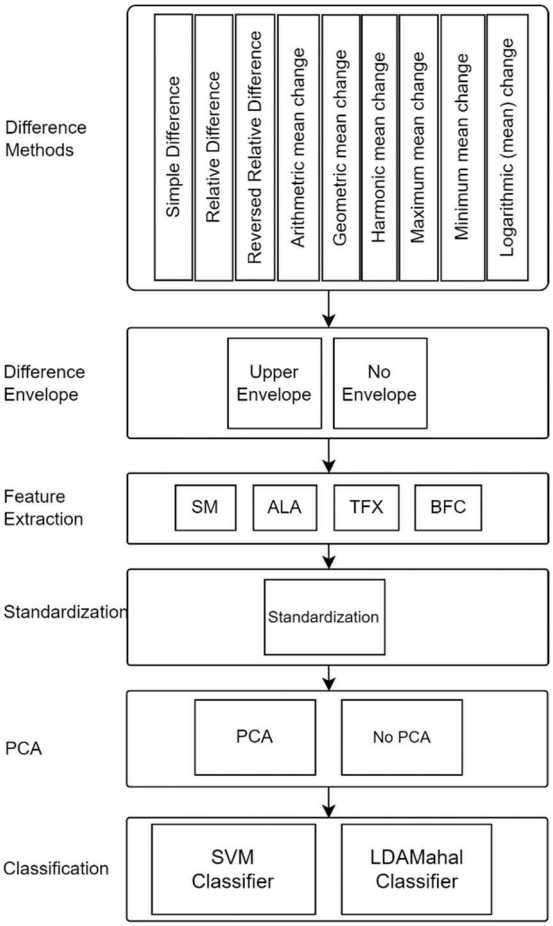


Fig. 5. Schematic describing the signal methods flow, where always one method is picked for each step. As an example, a flow can as follows: Relative Difference → Upper envelope → ALA → Standardization → SVM Classifier

Table 2 Various relative difference method used to calculate baseline subtraction has being proposed in the study [9]. Elementwise damaged signal is noted as $d(t)$ while $b(t)$ represents the elementwise baseline signal

Difference method	Formula
Simple Difference	$d(t) - b(t)$
Relative Difference	$(d(t) - b(t))/d(t)$
Reversed Relative Difference	$\big(d(t) - b(t)\big)/b(t)$
Arithmetic mean change	$(d(t) - b(t))/((d(t) + b(t)/2))$
Geometric mean change	$(d(t) - b(t))/\sqrt{d(t) * b(t)})$
Harmonic mean change	$\frac{\big(d(t) - b(t)\big) * \big(d(t) + b(t)\big)}{2d(t) * b(t)}$
Maximum mean change	$(\frac{d(t)}{b(t)} - 1)/\max(1, d(t)/b(t))$
Minimum mean change	$(\frac{d(t)}{b(t)} - 1)/\min(1, d(t)/b(t))$
Logarithmic (mean) change	$\big(ln(d(t)) - ln(b(t))\big)$

Table 3: Methods used to build the presented algorithm. Each pre-processing method is used separately with the rest of signal flow

Methods [7]	
	Feature Extraction
SM	Statistical Moments
ALA	Adaptive Linear Approximation
TFX	Time Frequency Extractor
BFC	Best Fourier Coefficients
	Classification
LDAMahal	Linear Discriminant Analysis transformation and classification based on Mahalanobis Distance
SVM	Support Vector Machines

Typically, the baseline subtraction involves computing the simple difference between the baseline and the measured signal. In this study, the authors also explored relative difference methods, as listed in [9]. While the simple difference entails elementwise subtraction of time series, a relative difference considers the scale of the values being compared in an elementwise manner. It involves dividing the simple difference by a reference value. Previous studies have solely examined the simple difference [10]. This standardization approach helps avoiding unwanted effects when amplitudes are too large, enhancing the robustness of the process of subtracting the baseline from the damaged signal in terms of isolating reflections or absorption caused by a present damage.

Some adaptations were made to these methods to address undefined states, such as division by zero. When the denominator is zero, it is set to one to obtain an equivalent value to the simple difference. Additionally, to prevent result distortion, any absolute value of the denominator that is less than one was set to one. This adjustment prevents large signal difference amplitudes resulting from division by very small numbers, ensuring that the results are not worse than those obtained with a simple difference.

Subsequently, a further processing step involves calculating the upper envelope, as it provides a concise description of the isolated signal shape, simplifying the signal and aiding in accurate calculations, particularly in Time of Flight procedures for damage localization [11].

Following this, we explored various feature extraction methods (Fig. 5 and Table 3). The used parameters for segmenting the signals are listed in Table 4.

Table 4: Feature extraction paramaters used for segmentation

Feature Extraction	Segmentation Paramters
SM	`[10,15]`
ALA	`[40,70,100]`
TFX	`[(1,20),(1,60),(1,100)]`
BFC	`[40,70,100]`

Subsequently, standardization is performed as the further classification method (SVM) is sensitive to it. Then, applying principal component analysis (PCA) is tested [12], as it helps in class separation by changing the weighting of the scattering. The final step involves training the LDA-Mahal or SVM classifier (Table 3). Therefore, the approach investigates a total of 576 algorithms and compares their performance with each other.

B. Targets preparation

The training step is based on the ellipsoid method [7]. Each transducer/receiver pair is labeled based on its ellipsoidal distance from the damage. The label is calculated as the ratio β of the direct path (c) to the indirect path (a+b). The direct path represents the shortest distance between a transducer and a receiver, while the indirect path corresponds to the ultrasonic wave reflected by the damage (see Fig. 6). Once the classifier retrieves the ratio, the damage can be localized as the intersection of multiple ellipsoids.

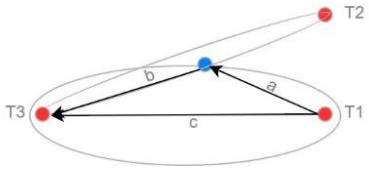


Fig. 6. The ellipsoid distance is described by the ratio of the direct path distance (c) to the sum of the indirect path distances (a+b).The blue dot shows the damage position

Essentially, a β-ratio between 1 and 0.90 indicates damage positions relevant to transducer-receiver pairs as used by the RAPID algorithm introduced in [13]. Therefore, we utilized the ratio interval from 1 to 0.90 to include only neighboring sensor pairs in proximity to the damage.

To construct classes suitable for training classification, we divided the interval [1, 0.90] into 10 classes. The class closer to 1 intentionally spans a smaller range (from 1 down to 0.995) to accurately represent damages lying in the direct path between the transducer and receiver pair. On the other hand, sensor pairs are labeled with a value of zero and are not

considered when their β-ratio to the damage is less than 0.9 (Table 5).

Table 5 quantization used for labeling sensor paths based on ratio of direct path to indirect path between transducers and receivers

Ellipsoid ratio intervals	Quantization value / Labeling	Percentage of observations per class
(0.995 , 1]	10	2.38
(0.98 , 0. 995]	9	2.50
(0.97 , 0.98]	8	1.79
(0.96 , 0.97]	7	0.48
(0.95 , 0.96]	6	0.48
(0.94 , 0.95]	5	0.00
(0.93 , 0.94]	4	0.47
(0.92 , 0.93]	3	0.36
(0.91 , 0.92]	2	0.00
(0.90 , 0.91]	1	0.71
[0 , 0.90]	0	90.83

Model Validation

The ML model validation was conducted using Leave-One-Group-Out Cross-Validation (LOGOCV), where each left-out group represents sensor pair signal data from all eight experiments. This involves splitting the dataset into 105 groups. While 104 groups are used for training the model, the remaining one group is left out for testing and validation. Each group comprises a set of 8 pitch-catch signal collected during the eight experiments (Table 1). This procedure enables testing the model's robustness against changes in sensor position, focusing on features that describe the presence of damage rather than features that describe sensor positions.

Best Model

To identify a suitable model for damage detection, we compared the performance of 576 algorithms, covering all possible combinations illustrated in Fig. 5. The best model achieved an accuracy of 96.55% (Table 6). Notably, applying the following algorithm leads to the best result: raw signal → simple difference → signal envelope → ALA feature extractor with 70 segments → standardization → SVM classifier

Table 6: Listing of the top 15 algorithms with the highest accuracy based on the difference method. The classification method used consists of SVM, as LDA did not give significant results

Difference method	Use of Signal envelope	Feature extraction methods	Segmentation width	Use of PCA	Accuracy %
Simple Difference	Yes	ALA	70	No	96,55
Simple Difference	Yes	ALA	100	No	95,71
Simple Difference	Yes	ALA	40	Yes	95,48
Simple Difference	Yes	ALA	40	Yes	95,12
Simple Difference	Yes	ALA	70	Yes	95,00
Simple Difference	No	ALA	100	No	94,52
Geometric mean change	No	ALA	40	No	92,02
Geometric mean change	No	ALA	70	No	92,02
Geometric mean change	No	ALA	100	No	91,90
Geometric mean change	Yes	ALA	40	No	91,55
Arithmetic mean change	Yes	ALA	100	No	91,55
Geometric mean change	No	ALA	70	No	91,55
Geometric mean change	No	ALA	40	Yes	91,55
Geometric mean change	No	ALA	70	Yes	91,43
Arithmetic mean change	No	ALA	40	Yes	91,31

Result Visualization

As explained in the Model Design chapter, the classifier predicts the quantized β-ratios as listed in Table 5. Therefore, visualizing the predicted damage involves mapping the labels back to their β-ratios in reverse. This process is performed as shown in Table 7.

Table 7: Mapping used to transform Model labels to the corresponding β-ratios

Predicted Labels	Recovered Ellipsoid Ratio
10	0.999
9	0.99
8	0.98
7	0.97
6	0.95
5	0.94
4	0.93
3	0.92
2	0.91
1	0.90
0	no ellipsoid

Following that, the predicted damage location is determined by the intersection of the ellipsoids plotted using their β-ratios. The centroid of the intersection area with the highest amplitude is then marked with a pink dot, labeling the recovered damage localization (see Fig. 7, Fig. 8, Fig. 9 showing the best model selection). In the other hand the green plus symbol represents the position of the actual glued damage. The deviation between the real damage position and predicted average damage position was less than 23 mm for the best model selection. Furthermore, detecting multiple damages is also possible, as shown in Fig. 10.

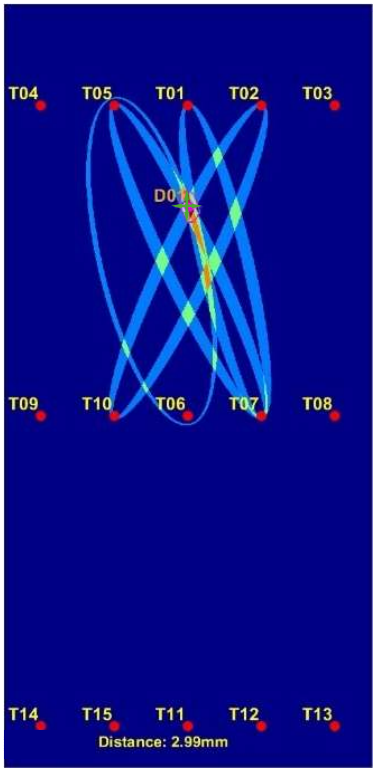


Fig. 7: Visualization of the damage predicted for experiment 1. Distance is the deviation of detection from real damage position

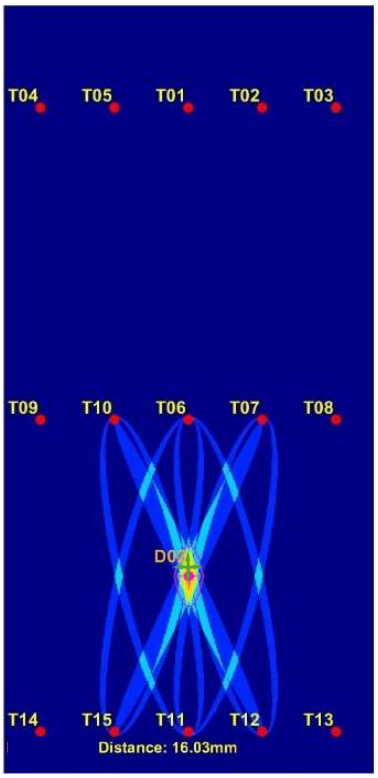


Fig. 9: Visualization of a damage located in the lower vessel region – Experiment 8

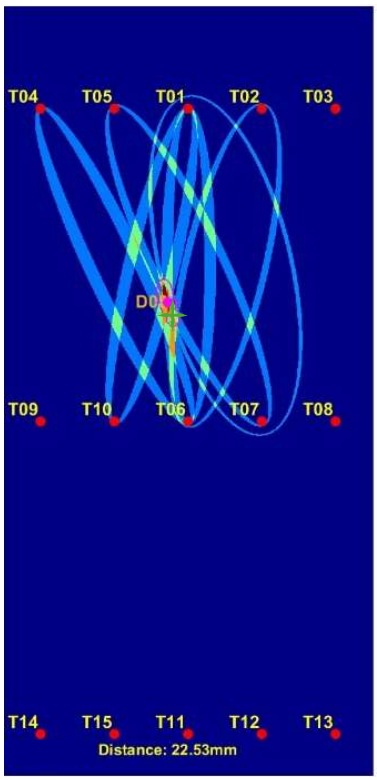


Fig. 8: Visualization of the damage predicted for experiment 4

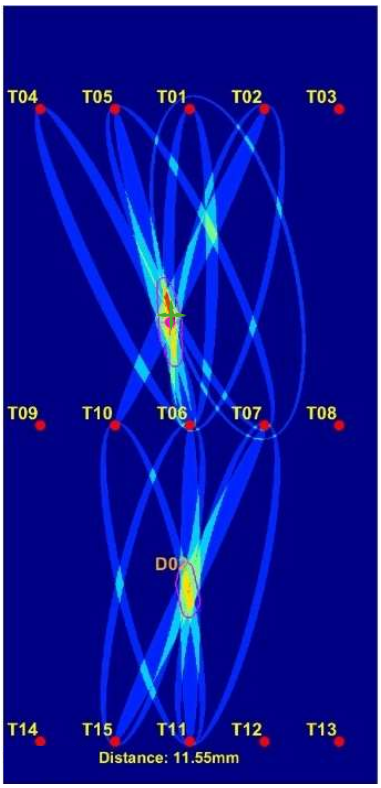


Fig. 10: Visualization of two damages located in the upper and lower vessel regions – Experiment 6

Conclusion

The authors presented an ML algorithm that embodies the advantages of explainable machine learning [14], crucial for applications where traceability is essential. Therefore, weak points of the method can be easily spotted and resolved, starting by investigating the extracted features and finishing with the classification method used to determine its suitability for our case.

The developed method combination underwent testing of various baseline subtraction methods and subsequent application of feature extraction techniques. This was followed by a classification process to localize damages using the pre-trained model, resulting in an accuracy of nearly 96.5%. This result could suffer from overfitting, so it needs to be validated under different vessels and/or with different sensor numbers and sensor positions.

However, there are areas for further investigation. The quantization used to label the sensor pairs requires further exploration. Additionally, the range of sensor pairs involved in damage detection needs to be explored, as pairs with a β-ratio less than 0.90 were not considered. The best way to investigate this consists of using regression instead of classification.

Acknowledgment

This paper acknowledges funding from the German Federal Ministry for Education and Research (BMBF) under grant number 03VP10464 for the project "Künstliche Intelligenz für das Ultraschall-Monitoring von Wasserstoff-Druckbehältern" (KIMono). The authors express their appreciation to the BMBF for their support and to the Federal Institute for Material Research and Testing (BAM), Berlin, Germany, for conducting the experiments.

References

[1] Shuangmiao Z.; Shaoping Z.; Shaojie C.& Bin Y.; Yong, L. Novel Defect Location Method for Pressure Vessel by Using L (0, 2) Mode Guided Wave. Journal of Pressure Vessel Technology. 2018. doi: 141. 10.1115/1.4039502

[2] Zima, B.; Moll, J, Theoretical and experimental analysis of guided wave propagation in plate-like structures with sinusoidal thickness variations. Archiv.Civ.Mech.Eng 23, 34 (2023). https://doi.org/10.1007/s43452-022-00564-9

[3] Zatar, W.A.; Nguyen, H.D.; Nghiem, H.M. Ultrasonic pitch and catch technique for non-destructive testing of reinforced concrete slabs. J Infrastruct Preserv Resil 1, 12 (2020). doi: 10.1186/s43065-020-00012-z

[4] Capineri, L.; Bulletti, A. Ultrasonic Guided-Waves Sensors and Integrated Structural Health Monitoring Systems for Impact Detection and Localization: A Review. Sensors 2021, 21, 2929. doi: 10.3390/s21092929

[5] Qian, L.; Liu, S.; Fan, G.; Liu, X; Zhang, H.; Mei, Y.; Xing, Y; Wang, Z., Damage localization method using ultrasonic lamb waves and Wav2Vec2.0 neural network. 2023. doi:10.3389/fmats.2023.1212909

[6] Dorst, T.; Schneider, T.; Eichstädt, S.; Schütze, A. Uncertainty-aware automated machine learning toolbox - Technisches Messen, vol. 90, no. 3, 2023, pp. 141-153. doi: 10.1515/teme-2022-0042

[7] Guo, J.; Zeng, X.; Liu, Q.; Qing, X. Lamb Wave-Based Damage Localization and Quantification in Composites Using Probabilistic Imaging Algorithm and Statistical Method. Sensors 2022, 22, 4810. doi:10.3390/s22134810

[8] McKeon, P.; Yaacoubi, S.; Declercq, N.F.; Ramadan, S.; Yaacoubi, W.K. Baseline subtraction technique in the frequency-wavenumber domain for high sensitivity damage detection. Ultrasonics. 2014 Feb;54(2):592-603. doi: 10.1016/j.ultras.2013.08.010

[9] Vartia, Yrjö O. (1976). Relative changes and index numbers. ETLA A 4. Helsinki: Research Institute of the Finnish Economy. ISBN 951-9205-24-1

[10] Dawson, A.J.; Michaels, J.E.; Michaels, T.E. Isolation of ultrasonic scattering by wavefield baseline subtraction, Mechanical Systems and Signal Processing, 2016, Volumes 70–71, Pages 891-903, doi: 10.1016/j.ymssp.2015.09.008.

[11] Salmanpour, MS.; Sharif Khodaei, Z.; Aliabadi, MH. Guided wave temperature correction methods in structural health monitoring, Journal of Intelligent Material Systems and Structures, 2016. doi: 10.1177/1045389X16651155

[12] Yang, K.; Kim. S.; Yue, R.; Yue, H.; Harley, J.B. Long-term guided wave structural health monitoring in an uncontrolled environment through long short-term principal component analysis. Structural Health Monitoring. 2022;21(4):1501-1517. doi: 10.1177/14759217211035532

[13] Hua, J.; Lin, J.; Zeng, L., High-resolution damage detection based on local signal difference coefficient model. Structural Health Monitoring. 2015;14(1):20-34. doi:10.1177/1475921714546060

[14] Schnur, C.; Goodarzi, P.; Lugovtsova, Y.; Bulling, J.; Prager, J.; Tschöke, K.; Moll, J.; Schütze, A.; Schneider, T. Towards Interpretable Machine Learning for Automated Damage Detection Based on Ultrasonic Guided Waves. Sensors 2022, 22, 406. doi: 10.3390/s22010406