engineering proceedings

MDPI

Proceedings

Spatial Damage Prediction in Composite Materials using Multipath Ultrasonic Monitoring, Advanced Signal Feature Selection and Combined Classifier-Regression Artificial Neural Network †

Stefan Bosse [1,*] and Christoph Polle [2]

[1] Dept. Mathemaics & Computer Science, University of Bremen, 28359 Bremen, Germany; sbosse@uni-bremen.de
[2] Faserinstitut Bremen, 28359 Bremen, Germany; polle@Faserinstitut.de
* Correspondence: sbosse@uni-bremen.de
† Presented at the 8th International Electronic Conference on Sensors and Applications, 1–15 November 2021; Available online: https://ecsa-8.sciforum.net.

Abstract: Automated damage detection in Carbon-Fibre and Fibre Metal Laminates is still a challenge. Impact damages are typically not visible from the outside. Different measuring and analysis methods are available to detect hidden damages, e.g., delaminations or cracks. Examples are X-ray computer tomography and methods based on guided ultrasonic waves (GUW). All measuring techniques are characterised by a high-dimensional sensor data, in the case of GUW that is a set of time-resolved signals as a response to a actuated stimulus. We present a simple but powerful two-level method that reduces the input data (time-resolved sensor signals) significantly by a signal feature selection computation finally applied to a damage predictor function. Beside multi-path sensing and analysis, the novelty of this work is a feed-forward ANN posing low complexity and that is used to implement the predictor function that combines a classifier and a spatial regression model.

Keywords: structural health monitoring; multipath monitoring; feature selection; analytical signal; classification; regression; artificial neural network

Citation: Bosse, S.; Polle, C. Spatial Damage Prediction in Composite Materials using Multipath Ultrasonic Monitoring, Advanced Signal Feature Selection and Combined Classifier-Regression Artificial Neural Network. *Eng. Proc.* **2021**, *3*, x. https://doi.org/10.3390/xxxxx

Academic Editor(s):

Published: 1 November 2021

1. Introduction

Structural Health Monitoring (SHM) in Carbon-Fibre and Fibre Metal Laminates (FML) is used to detect and assess mainly hidden damages under the hood. Damage detection, classification, and localisation is part of the lower levels of SHM. SHM is an extremely useful tool for ensuring integrity and safety, detecting the evolution of damage, and estimating performance deterioration of civil infrastructures, but relies heavily on the robustness and accuracy of the underlying damage feature detectors. Early damage detection can avoid situations which can be catastrophic. SHM can allow efficient maintenance works and can avoid unnecessary inspections, furthermore, saving time and money.

One prominent measuring technique for damage detection is the monitoring of guided ultrasonic waves resulting from stimulated ultrasonic emission. Guided Ultrasonic waves (GUW) interact with damages and defects resulting in a modification of the time-resolved ultrasonic sensor signal at a given sensor position. The difference of a signal from a damage interaction with the baseline is typically low and difficult to detect. Although, ML can exploit the relevant damage features from the sensor signals by, e.g., supervised training using a highly non-linear function (function graph implemented by

an Artificial Neural Network ANN) [1], advanced feature selection can improve the damage prediction accuracy and reduces the functional complexity of the predictor function significantly. The wave propagation depends beside material properties and the signal frequency from temperature and moisture (inside the material if it is a composite material). ANNs posing low complexity were already successful applied to damage detections [2,3].

This work addresses a novel two-stage damage detection method that uses supervised Machine Learning (ML) for the training of a damage feature predictor function from experimental data that is able to provide binary damage classification and spatial damage localisation information with high accuracy and reliability even under varying environmental conditions. The output of a non-linear regression function graph model (a traditional Artificial Neural Network with sigmoid transfer functions) is a two-dimensional vector providing an estimated positions of a damage (and the encoded non-damage case). The input of this predictor function is a medium dimensional feature vector that is derived by envelope curve approximation of the measured raw time-resolved ultrasonic wave signal. The ultrasonic waves interact with the damage resulting in a modification of the measured signal finally providing the feature vector [4]. In contrast to other approaches, this approach uses multi-path measurements, i.e., signal recordings of different spatial paths between an actuator and a sensor covering the whole device under test area. The derived features are characteristics of the recorded signals with respect to the desired damage information. Finally, a damage predictor function is trained under varying environmental conditions having impact of the wave interaction and the derived features, here specifically the ambient and device temperature [5].

Beside multi-path sensing and analysis using already recorded sensor data of a CFK plate from the Open Guided Waves data base [6,7], the novelty of this work is a feed-forward ANN used to implement the predictor function that combines a classifier and a spatial regression model, reducing computational and memory complexity, a constraint for the implementation in embedded sensor node systems.

2. Multi-Path Sensor Data

The data source in this work is time-resolved ultrasonic signal data from an active measuring technique, i.e., the signal response is a result of an active stimulus. A piezo-electric actuator that is coupled to the surface of the device under test (or embedded inside a lamnination layer like the CFK plate used in this work) injects a sine-wave like pulse group (about 10 waves). The guided waves propagate through the material and partially on the surface. The guided wave interacts with the material and potential defects and damages. The damage interaction leads to a change of the original damage-free signal only in small part of the signal with respect of the time dimension, shown in Figure 1 (top, right), the Region-of-Interest (ROI). The identification of the ROI is difficult, depending on a-priori knowledge and the strength of damage-wave interaction that can be weak. Beside using the raw time series data, typically transformed in frequency or time-frequency space (like FFT or DWT), characteristic feature parameters should be derived numerically from the signal.

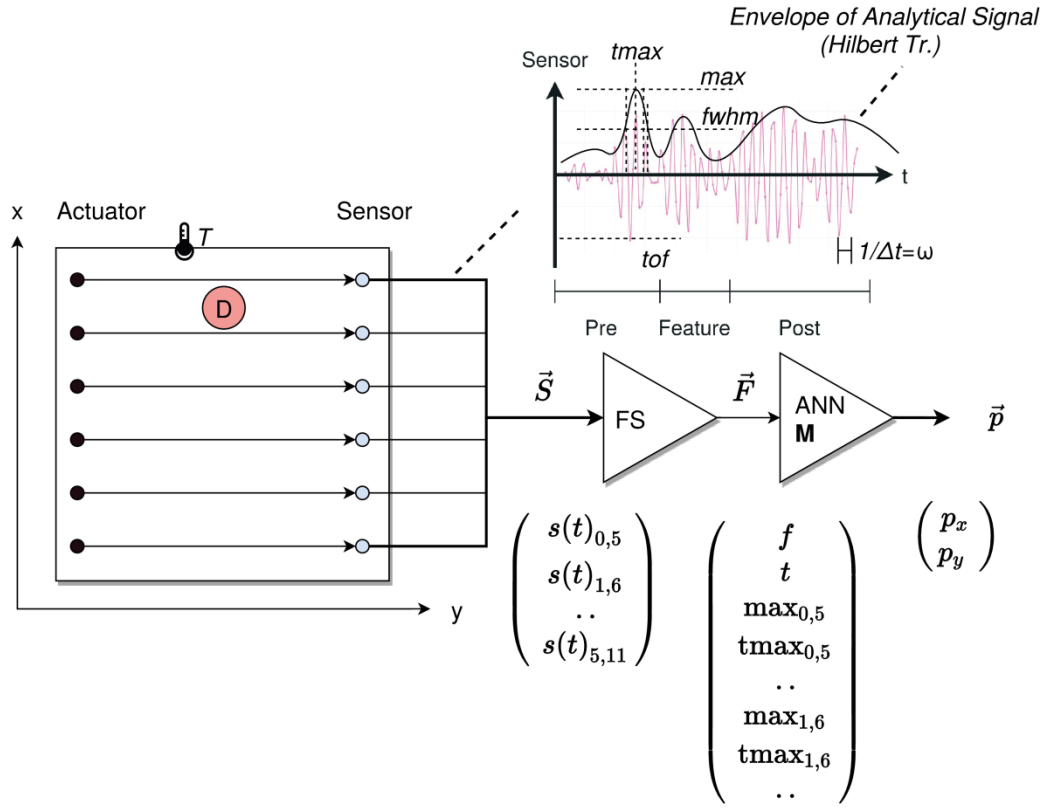

Figure 1. A plate (device under test) is equipped with 12 piezo-electric transducers that can act as an ultrasonic actuator and sensor, too. Six direct paths are measured simultaneously. The time-resolved sensor data is processed by an analytical signal feature selection.

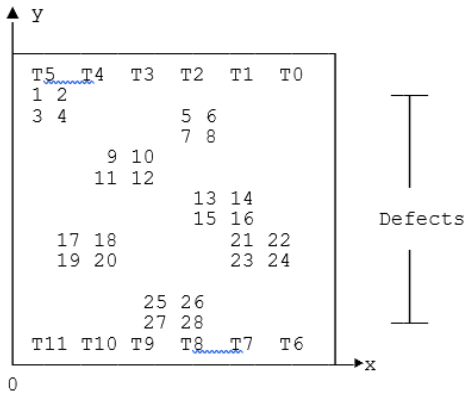

Figure 2. Transducer positions T0..T11 and defect positions D1..D28.

There are two data set groups that are available from the OGW data base feature different measurements with pseudo defects (positions shown in Figure 2):

1. Five data tables with recorded GUW signals with a dynamic temperature profile (20–60 °C) and a sub-set of four defect positions D4, D10, D14, D24, and a base line measurement without a defect (named here dynamic data set);
2. 33 data tables with recorded GUW signals at a static temperature (24 °C) and all 28 defect positions and the base line measurement (named here static data set).

3. Feature Selection

In general, the aim of feature selection is the mapping of the raw time-resolved signal data $s(t)$ on a damage relevant and representative small set of feature parameters $\vec{f}$ by a feature selection function ψ:

$$\begin{aligned} &\psi_{i,j} : \vec{s}_{i,j} \rightarrow \vec{f}_{i,J} \\ &\vec{s} = \langle s(t=0), s(t=1), .. \rangle \\ &\vec{F} = \begin{pmatrix} T \\ \omega \\ \vec{f}_{i,j} \end{pmatrix} \end{aligned} \tag{1}$$

But the GUW depends on the temperature of the medium in which the waves propagate and as a result the damage features are dependent on the environmental state, mostly the temperature [4]. Relevant damage features are contained in the envelope of the signal burst, commonly related to the envelope of the dominant wave group, mainly the hight (*max*), the time point of the maximum (*tmax*), and the full width at half maximum *fwhm*. These parameters strongly dependent on the material temperature as a result of the wav propagation. Details can be found in [4]. To derive the envelope of the signal burst, two numerical approaches can be used:

1. Computing the magnitude of the complex analytical signal by a Hilbert transformation of the time-resolved signal $s(t$ (non-iterative approach);
2. By finding the maximum signal peak and performing a constrained Gaussian peak fitting of the wave group around the maximum, i.e., fitting a Gaussian function to the envelope of the signal group (iterative approach).

For a given time dependent signal $s(t)$ the analytical signal $s_a(t)$ is given as:

$$\begin{aligned} &s_a(t) = s(t) + \mathrm{i}\left[s(t) * \frac{1}{\pi t} \right] = s(t) + \mathrm{i}\mathcal{H}\{s(t)\} \\ &H(s(t)) = |s_a(t)| \end{aligned} \tag{2}$$

The analytical signal bases basically on a convolution operation ($*$) but can be derived by using the discrete Fourier transforms (DFT, and fast version FFT):

$$\begin{aligned} &X[\omega] = \tau \sum_{n=0}^{N-1} s[n] e^{-\mathrm{i}2\pi\omega n\tau} \\ &Z[m] = \begin{cases} X[0] & \text{if } m = 0 \\ 2X[m] & \text{if } 1 \le m \le \frac{N}{2} - 1 \\ X\left[\frac{N}{2}\right] & \text{if } m = \frac{N}{2} \\ 0 & \text{if } \frac{N}{2} + 1 \le m \le N - 1 \end{cases} \\ &H[n] = \frac{1}{NT} \sum_{m=0}^{N-1} Z[m] e^{\mathrm{i}2\pi m \frac{n}{N}} \end{aligned} \tag{3}$$

with τ: sampling interval, ω: frequency, X: forward DFT, Z; Hilbert transform in frequency domain, H: final Hilbert transform in time domain by using the inverse DFT.

Characteristic features f_i derived from the envelope of the signal $s(t)$ are [4] (see also Figure 1, top right):

- The absolute (normalized) maximum value of the dominant envelope peak *max*;
- The time position at the maximum *tmax*;
- The full width at half maximum of the envelope peak *fwhm*;
- And a time-of-flight parameter *tof*.

All these features are dependent on the signal frequency ω, the temperature T, and for the normalization on the stimulus amplitude, i.e., $f_i = f_i(\omega,T)$, .

4. Predictor Model Function

A classical feed-forward fully connected neuronal network with one or two hidden layers is used to predict the damage position $p = (x,y)$ in normalized coordinates $x = [0,1]$, $y = [0,1]$. An output $|p| < \varepsilon$ indicates the absence of a damage, i.e., x and $y \approx 0$. Therefore, the predictor function combines a classifier and a spatial regression model, shown in Figure 3, reducing computational and memory complexity, a constraint for the implementation in embedded sensor node systems.

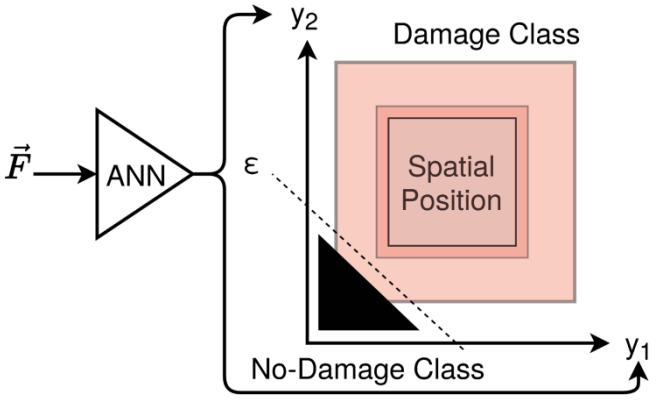

Figure 3. One two-dimensional output predictor function M combines a damage classifier and a damage position ($p_x = y_1$, $p_y = y_2$) regression function

The input of the network is a vector containing the material temperature T, the signal frequency ω, and selected features from all six straight path ($\varphi = 0$) ultrasonic signal measurements. Typical features derived from the feature selction process are *max* and *tmax*. Additional computed feature parameters are time-of-flight *tof* and the full width at half maximum *fwmh*.

$$M(F): \vec{F} \rightarrow \vec{p}\vec{F} = \begin{pmatrix} T \\ \omega \\ \max_{i,j} \\ \mathrm{tmax}_{i,j} \\ .. \end{pmatrix}, \forall (i,j) \in \mathrm{conf}(i,j) \tag{4}$$

$$\mathrm{conf}(i,j) = \{\langle 0,6\rangle, \langle 1,7\rangle, \langle 2,8\rangle, \langle 3,9\rangle, \langle 4,10\rangle, \langle 5,11\rangle\}$$

$$\vec{p} = (x,y)$$

The input layer consists of $||F||$ neurons depending on the selected sub-set of features. The output layer consists of two neurons providing an estimation of the damage x- and y-positions, respectively. The output is normalized to a spatial range of [0.2,0.7] that corresponds to a geometric range of [0,0.5m]. The non-damage case is predicted if $|p| < \varepsilon$ (e.g., $\varepsilon = 0.02$); Any p_x or $p_y \in (0.7,1.0]$ or $(\varepsilon,0.2)$ indicates a prediction error!

5. Functional Scaling

The target host environment for the deployment of the damage predictor function is an embedded sensor node equipped with a low-resource and low-power microcontroller. This sensor node acquires the multi-path raw sensor signals, performs the feature selection pre-processing and the application of the predictor function. Even if the training of the predictor function takes place off-line, the application of the predictor function should be performed on-line. The predictor function consists of the signal pre-processing with the previously introduced feature selection algorithm, and the forward computation of the ANN. Tables 1 and 2 show typical computational times for the feature selection represented by the major part of the Hilbert transformation and the ANN application. Different host computer architectures and processing platforms (i.e., native machine code and VMs node.js and quickjs). For both algorithms there is a JavaScript and a C implementation. The Raspberry Pi Zero is a small low-power embedded computer, although, it is still oversized compared with material-integrated nano computers (less than 100MHz CPU clock and about 100 kB RAM), both algorithms can be implemented and processed on such low-resource systems.

The computational complexity of the ANN is neglectible compared with the feature computation process (about 1:1000). For each prediction, m Hilbert transformations must be performed for m paths. But even using the slowest but embeddable JavaScript quickjs platform, the entire prediction requires less than two seconds on a RP Zero. Assuming a computational power ratio of 1:100 comparing the RP Zero with a material-integrated nano computer (e.g., the ancient Micro Mote M3), a native code implementation of the full predictor program requires only three second computation time, which can be still considered as sufficient. Probably FIR/IIR filter-bank approach approximating the Hilbert transform can provide an additional reduction of the computational complexity.

Table 1. Computation times for the central feature selection (Computation of analytical signal by Hilbert transform) and the application of the ANN predictor model for various processing platforms (gcc: C and native machine code, node.js/quickjs: JavaScript code).

Host	Platform	Function	Performance
Intel i5-3427U 1.80GHz	node.js 8.12	dhrystone benchmark	5400 k dhry/sec
Intel i5-3427U 1.80GHz	node.js 8.12	ann-model-14-8-2	10 µsec/pred
Intel i5-3427U 1.80GHz	gcc 6.3 -O2	ann-model-14-8-2	0.8 µs/pred
Intel i5-3427U 1.80GHz	node.js 8.12	hilbert-fft (4096 points)	1 ms/trans
Intel i5-3427U 1.80GHz	gcc 6.3 -O2	hilbert-fft (4096 points)	0.5 ms/trans

Table 2. Computation times for the central feature selection (Computation of analytical signal by Hilbert transform) and the application of the ANN predictor model for various processing platforms (gcc: C and native machine code, node.js/quickjs: JavaScript code).

Host	Platform	Function	Performance
Raspberry Pi Zero W 1.0GHz	node.js 8.12	dhrystone benchmark	200 k dhry/sec
Raspberry Pi Zero W 1.0GHz	quickjs 2021.3	dhrystone benchmark	8 k dhry/sec
Raspberry Pi Zero W 1.0GHz	node.js 8.12	ann-model-14-8-2	30 µsec/pred
Raspberry Pi Zero W 1.0GHz	quickjs 2021.3	ann-model-14-8-2	300 µsec/pred
Raspberry Pi Zero W 1.0GHz	gcc 6.3 -O2	ann-model-14-8-2	5 µs/pred
Raspberry Pi Zero W 1.0GHz	node.js 8.12	hilbert-fft (4096 points)	25 ms/trans

Host	Platform	Function	Performance
Raspberry Pi Zero W 1.0GHz	quickjs 2021.3	hilbert-fft (4096 points)	300 ms/trans
Raspberry Pi Zero W 1.0GHz	gcc 6.3 -O2	hilbert-fft (4096 points)	5 ms/trans

6. Evaluation

The existing GUW signal data [6] was taken from the OGW server and a broad set records signal data sets were stored in SQL tables for further processing. The feature selection computation using the Hilbert transform and conventional peak analysis algorithms was originally performed with Python code [4]. The predictor function was implemented with a modified version of the Neataptic JavaScript ANN framework [8]. The experimental matrix consists of:

- Two data sets: (D) Dynamic temperature profile (T = 20–50 °C, 4 defect positions), (S) Static temperature (T = 24 °C, 28 defect positions), recorded GUW sensor date from a 500 × 500 mm CFK plate with attached pseudo defects;
- Different feature parameter sets: {*max*, *tmax*, *T*}, {*max*, *tmax*, *fwhm*, *T*}, {*max*, *tmax*, *fwhm*, *tof*, *T*}
- Different network architectures: [*I*,*H*₁,*H*₂,...,*O*]
- Two different input variable scaling methods: Static, i.e., a feature is scaled equally for all paths with a fixed scale; Auto: a feature is scaled automatically and independently for all paths.
- Different Training- and Testset combinations: {D/D, D/S, SD/SD, SD/Dm SD/S}

Due to the limited data set variance, the test of the model accuracy was tested with training data with different combinations of the dynamic and static temperature data sets. Although, Monte Carlo simulation was used to augment training data by adding Gaussian noise, no further data augmentation was performed. Therefore, the results shown in Tables 3 and 4 cannot conclude any generalization quality of the trained model. The position error threshold was set to 100 mm (error above is classified as an incorrect position prediction). The non-damage detection threshold was set to 0.05/1.0. The results in Tables 3 and 4 show the average defect position estimation accuracy delivered by the predictor function, the fraction of incorrectly located defects (position error too large), and the binary defect classification rates true-positive (TP, damage), true-negative (TN, no damage) with their negative counterparts false-positive (FP) and false-negative (FN).

Table 3. Prediction results for ω = 40 kHz (D: Dynamic temperature data set, S: Static temperature).

Features $F^{\rightarrow}$(40 kHz)	Model [Layer]	Scaling	Training Data [%]	Test Data [%]	Mean. Pos. Error [%]	Position Incorrect [%]	TP/FP, TN/FN [%]
max, *tmax*, *T*	[14,8,2]	Fixed	D = 100, S = 100	D = 100, S = 100	3	4	100/0, 100/0
max, *tmax*, *T*	[14,8,3,2]	Fixed	D = 100, S = 100	D = 100, S = 100	5	9	100/0, 100/0
max, *tmax*, *T*	[14,8,3,2]	Auto	D = 100, S = 0	D = 100, S = 0	2	0	100/0, 100/0
max, *tmax*, *fwhm*, *T*	[20,8,2]	Fixed	D = 100, S = 100	D = 100, S = 100	9	16	100/0, 91/9
max, *tmax*, *fwhm*, *T*, *tof*	[26,8,2]	Fixed	D = 100, S = 100	D100, S = 100	6	21	100/0, 100/0
max, *tmax*, *T*	[14,8,3,2]	Fixed	D = 100, S = 0	[1]D = 100, S = 0 [2]D = 0, S = 100	[1]3 [2]48	[1]0 [2]70	[1]100/0, 100/0 [2]20/80, 100/0

Table 4. Prediction results for ω = 80 kHz (D: Dynamic temperature data set, S: Static temperature).

Features $F^{\rightarrow}$(80kHz)	Model [Layer]	Scaling	Training Data [%]	Test Data [%]	Mean. Pos. Error [%]	Position Incorrect [%]	TP/FP, TN/FN [%]
max, *tmax*, *T*	[14,8,2]	Fixed	D = 100, S = 100	D = 100, S = 100	20	50	100/0, 33/67
max, *tmax*, *T*	[14,8,4,2]	Fixed	D = 0, S = 100	D = 0, S = 100	16	36	100/0, 0/100
max, *tmax*, *T*	[14,8,3,2]	Auto	D = 100, S = 0	D = 100, S = 0	4	3	100/0, 100/0
max, *tmax*, *T*	[14,8,3,2]	Auto	D = 100, S = 100	D = 100, S = 100	12	27	100/0, 80/20
max, *tmax*, *T*	[14,8,3,2]	Auto	D = 100, S = 0	[1]D = 100, S = 0, [2]D = 0, S = 100	[1]5, [2]60	[1]3, [2]85	[1]100/0, 100/0 [2]0/100, 100/0

We observed that there is:

- A high accuracy of defect classification (100% TP, 100% TN) even under temperature variations in the range 20–50 °C can be achieved by a network with only one hidden layer (8 neurons) and by using the major features *max* and *tmax*;
- The other minor features *fwhm* and *tof* can be discarded, they show no benefit (in contrast, including them degrades model accuracy until the training damps them);
- A reasonable defect localisation with an average position error below 20 mm is possible;
- A high sensitivity of prediction results to feature parameter noise (even if low as 5% Gaussian noise) and feature variable scaling (static and fixed versa dynamic and automatic);
- Training process and prediction accuracy shows high sensitivity on data normalization (scaling);
- Probably only a specialized model was trained (due to low variances in defect positions and variations of environmental parameters);
- Suitable learning rate were chosen between 0.05 and 0.2 depending
- A model trained by the dynamic data set (only four defects) show low accuracy for the prediction of the static data set (even concerning the four damages contained in the static data set, too)
- The accuracy of the prediction model depends on the signal frequency (40kHz outperforms 80kHz)
- One hidden layer is typically suitable for achieving a high accuracy showing a low non-linearity degree of this problem using;
- The training time for one model is about some minutes on a generic desktop computer (JavaScript processed by node.js or in the WEB browser).

7. Conclusions

Using multi-path sensing of guided ultrasonic waves, advanced feature selection, and a simple artificial neural network we were able to detect pseudo defects applied to a CFK plate with a high probability (typically nearly 100%) and position accuracy (typically below 20 mm or better) in a wide range of material temperature (20–50 °C). The advanced feature selection bases on a Hilbert transform of the time-resolved signal data and maximum peak analysis. The computed features are the input vector for the ANN predictor functions that combines a binary damage/defect classifier with a two-dimensional position regression of the damage. Typically only one hidden layer with a few neurons is suitable to achieve high accuracy and TP/TN rates.

It could be shown that the proposed analysis method is suitable to be implemented in embedded systems including material-integrated nano computers providing damage detection within 10 s after signal measurement, which is sufficient for a broad range of

applications in SHM. Using signal down-sampling and optimized implementations of the FFT and Hilbert transform should provide prediction times below 1 s (on an embedded nano computer) The feature selection reduces the input data vector dimension from 4096 samples × 6 paths (24576) to lowest dimension of 14 (maximal 26 depending on the selected feature sub-set)!

This work bases on already existing data lacking variance with respect to defect positions, material properties, and environmental conditions. Further investigations using GUW measurements for a Fibre-Metal Laminate plate are under work and should create a suitable training and test data set that allows assessment of the robustness and generalization degree of the trained model.

Institutional Review Board Statement:

Informed Consent Statement:

Data Availability Statement:

References

1. Sarkar, S.; Reddy, K.K.; Giering, M.; Gurvich, M.R. *Deep Learning for Structural Health Monitoring: A Damage Characterization Application*; In Annual conference of the prognostics and health management society, 2016; pp. 176–182.
2. Bosse, S.; Lehmhus, D. Robust Detection of Hidden Material Damages Using Low-Cost External Sensors and Machine Learning. In Proceedings of the 6th International Electronic Conference on Sensors and Applications (ECSA), 15–30 November 2019.
3. Bosse, S. Learning Damage Event Discriminator Functions with Distributed Multi-instance RNN/LSTM Machine Learning—Mastering the Challenge. In Proceedings of the 5th International Conference on System-Integrated Intelligence Conference, Bremen, Germany, 11–13 November 2020
4. Polle, C.; Koerdt, M.; Maack, B.; Focke, O.; Herrmann, A.S. Introduction of the Temperature Scaling Method for Structural Health Monitoring with Guided Ultrasonic Waves. In Proceedings of the IWSHM 2021: 13th International Workshop on Strictural Health Monitoring, Stanford University, Stanford, CA, USA, 15–17 March 2022.
5. Wang, Y.-S.; Gao, L.; Yuan, S.; Qiu, L.; Qing, X. An Adaptive Filter-based Temperature Compensation Technique for Structural Health Monitoring. *J. Intell. Mater. Syst. Struct.* **2014**, *25*, 2187–2198, doi:10.1177/1045389.
6. Open Guided Waves Data Base. Available online: http://openguidedwaves.de/downloads (accessed on 1 January 2020).
7. Moll, J.; Kexel, C.; Kathol, J.; Fritzen, C.-P.; Moix-Bonet, M.; Willberg, C.; Rennoch, M.; Koerdt, M.; Herrmann, A. Guided Waves for Damage Detection in Complex Composite Structures: The Influence of Omega Stringer and Different Reference Damage Size. *Appl. Sci.* **2020**, *10*, 3068, doi:10.3390/app10093068
8. Wagenaar, T. Available online: https://wagenaartje.github.io/neataptic (accessed on 1 January 2020).