Reduced Training Data for Laser Ultrasound Signal Interpretation by Neural Networks

Janez Rus[1] · Romain Fleury[1]

Abstract

The performance of machine learning algorithms is conditioned by the availability of training datasets, which is especially true for the field of nondestructive evaluation. Here we propose one reconfigurable specimen instead of numerous reference specimens with known, unchangeable defect properties, which are usually complicated to fabricate. It consist of a shape memory polymer foil with temperature-dependent Young's modulus and ultrasound attenuation. This open a possibility to generate a reconfigurable defect by projecting a heating laser in the form of a short line on the specimen surface. Ultrasound is generated by a laser pulse at one fixed position and detected by a laser vibrometer at another fixed position for 64 different defect positions and 3 different configurations of the specimen. The obtained diversified datasets are used to optimize the neural network architecture for the interpretation of ultrasound signals. We study the performance of the model in cases of reduced and dissimilar training datasets. In our first study, we classify the specimen configurations with the defect position being the disturbing parameter. The model shows high performance on a dataset of signals obtained at all the defect positions, even if trained on a completely different dataset containing signals obtained at only few defect positions. In our second study, we perform precise defect localization. The model becomes robust to the changes in the specimen configuration when a reduced dataset, containing signals obtained at two different specimen configurations, is used for the training process. This work highlights the potential of the demonstrated machine learning algorithm for industrial quality control. High-volume products (simulated by a reconfigurable specimen in our work) can be rapidly tested on the production line using this single-point and contact-free laser ultrasonic method.

Keywords Active wave control · Smart structures · Machine learning · Laser ultrasound · Quality control

1 Introduction

Emerging technologies based on machine learning algorithms or artificial neural networks are transforming the way information is processed [1–3]. Audio signal interpretation by machine learning algorithms is a broadly addressed topic with typical applications for already highly reliable speech recognition and sound classification [4–6]. Can ultrasonic signals be processed in a similar way and provide advantages of automation and improved performance, as it is the case for implementing image recognition algorithms for visual detection and classification of surface damages [7–9] and analysis of radiography images [10, 11]?

Insufficient amount of appropriate training data is the main reason why machine learning algorithms have not successfully penetrated the sector of industrial ultrasonic inspection yet. It is easier to generate a labeled dataset for a problem, which can easily be solved by a human, than a dataset of ultrasonic signals labeled with the target information. In the latter case, it is difficult to obtain diversified and large datasets, as we require a reference inspection method, which is in most cases destructive (e.g. microscopy of a section cut).

Most of the current ultrasonic inspection methods are relying on relatively basic principles based on, for example, observing reflections from internal features or increase of ultrasound attenuation levels caused by delaminations or porosity in through transmission setup [12–14]. More advanced methods use spectral analysis of the signal to

✉ Janez Rus
janez.rus@tum.de

1   Institute of Electrical and Micro Engineering, Laboratory of
    Wave Engineering, Ecole Polytechnique Fédérale de
    Lausanne (EPFL), Station 11, 1015 Lausanne, Switzerland

extract additional information [15–19]. Even more detailed reconstruction of the inspected area could be provided by the synthetic aperture focusing technique, total focusing technique, or certain phased array ultrasonic testing methods [20–26]. Fully new possibilities open if machine learning algorithms are used for the ultrasonic signal interpretation.

1.1 State of the Art: Machine Learning for Ultrasonic Inspection

In the following state of the art review, we address all types of industrial ultrasonic inspection methods where the training datasets were obtained experimentally. In the case of analytically and numerically obtained datasets, we limit us on the methods using lasers to generate or detect ultrasound.

1.1.1 Analytically and Numerically Supported Training Process

The numerical and analytical methods for generating large datasets are facing a strong assumption that the precise physical model of the inspection system is known. Geometric and material properties of a specimen and target features, as well as characteristics of ultrasound generation and detection must be numerically or analytically describable as precisely as possible. There is a question if the machine learning algorithm is really required, if a good physical model is already known. Machine learning algorithms provide a special advantage for the industrial inspection applications scenarios of products with complex shapes, which deliver signals not describable by analytical or numerical models. After the training process, the ultrasound generation, propagation, and detection properties will be encoded in the parameters of the neural network (weights and biases) and the ultrasonic signals can be interpreted without actually knowing the physical model of the measurement setup.

In one of the early realizations, numerically simulated surface wave dispersion curves of layered structures were used as training and testing datasets [27]. The model was experimentally tested on two specimens. Numerical simulations including signal preprocessing based on wavelet decomposition are used as an input to train the neural network [28]. The method is tested on a single composite plate specimen to determine the elastic constants of the material. Theoretical dispersion curves were used for the network training [29]. The method was tested experimentally on one plate by measuring its thickness, Young's modulus, and Poisson's ratio. In an another previous work, dynamic responses calculated by a finite-element method were used for the network training [30]. Location and size of the cuts on the specimen size opposite to the ultrasound excitation and detection were experimentally determined on two specimens.

Numerical simulations of wave propagation in the specimen were used to train the neural network [31]. Maximum, minimum, and peak-to-peak values of Rayleigh waves together with the signal peak frequency and its bandwidth were used as input parameters to determine the location and size of subsurface cracks. Part of the simulated signals and three experimental signals were used to validate the model. A numerical model was utilized to generate ultrasonic signals, which were converted to images and used for training an already pretrained convolutional neural network, which was derived from visual recognition tasks [32]. Four experimentally obtained signals were included in the training and the validation dataset in order to determine location and size of subsurface defects.

An interesting method for training data augmentation was based on a technique used for human inspector trainings, which allowed generation of virtual flaws with variable parameters (e.g. depth or size) by numerical modification of experimentally obtained signals [33, 34]. As it was the case for other methods based on numerically or analytically trained networks, prior knowledge (relevant physical model) on the measurement system and the specimen properties were required for this method as well.

1.1.2 Training Process Using Experimentally Obtained Data

It is technically challenging to produce a high number of reference specimens with known target parameters in order to train a neural network on purely experimental dataset. In previous studies, a larger training data size was typically obtained by repeating the measurement (eventually at different positions) on a small number of specimens. Consequently, the low diversity of the training datasets limited the robustness of the model, which typically provided good performance only for a specific configuration of the experimental system and for one specific inspection task.

Defect classification of porosity, lack of fusion, tungsten inclusions, and intact specimens were performed applying wavelet preprocessing of ultrasonic signals obtained by a piezoelectric transducer [35]. The training and testing datasets were altogether consisting of 240 ultrasonic signals obtained on 9 specimens. Presence of a notch in a metallic plate was detected by a neural network trained on a dataset of 216 numerically and 24 experimentally generated signals, obtained by varying the location of the transducers [36].

An elegant way to obtain larger dataset of ultrasonic signals is to use phased array ultrasonic transducers [37–40]. Signals of individual piezoelectric elements at different excitation parameters have only limited diversification if applied on low number of test objects. Crack orientation and depth were evaluated on a single specimen using wavelet packet decomposition [37]. 12 different defect shapes provided 240 signals of different steering beams angles, which were

divided to the training and testing datasets. The presence of 68 different defects (various holes and crack types) in 6 steel blocks was estimated [38]. The training and testing datasets were composed of more than 4000 linear scan images obtained by phased-array probe and augmented by flipping, random cropping, translation, and visual color adjustments. The performance of neural networks can be improved by a hybrid training process, as shown by two studies addressing detection of holes [39] and pipeline cracks [40]. Both used convolutional neural networks, which were consecutively trained on simulated data and experimental signals obtained by a phased-array probe.

Machine learning algorithms have been used to support ultrasonic quality control of spot welding by classifying them into four quality levels [41] and to predict their static and fatigue behavior [42]. X-ray computed tomography scans were used as a reference to train a neural network to estimate porosity level in carbon fiber reinforced polymers [43]. Similarly, electron backscattered scattering diffraction served as a reference measurement to measure grain size of polycrystalline metals using laser ultrasound [44].

Ultrasonic signals were captured at 10,000 different locations on four samples before and after applying a damage (in form of a mass) [45]. The training dataset was obtained by randomly picking different snapshots of a two-dimensional scan [46]. These two training data augmentations had only limited contribution to the robustness of the machine learning algorithms.

1.1.3 Subwavelength Information Extraction by Machine Learning Algorithms

Machine learning algorithms can be used to extract specific information with a resolution below the diffraction limit. The typical approach is to record wave responses scattered from the object in the far-field and perform learning process for different arrangements of the target information (e.g. position or shape of the imaged object). The inconveniences are that an arrangement variation is not always achievable and the method is limited to datasets closely related to the training conditions.

A study was performed using waves on a plate with differently shaped holes as defects using six numerical simulations and five physical specimens [47]. Size and variability of the training datasets were augmented by a random cropping, zooming, flipping, and rotating. The method was able to classify between different shapes, which were used during the training process, however, it was not able to distinguish arbitrary and unknown subwavelength shapes of defects.

It was shown that coupling of subwavelength information to the far field can be improved by placing randomly distributed resonators in the near field of an object [48, 49]. Subwavelength images of digits and numbers drawn by a two-dimensional array of speakers (emitting in the audible frequency range) were reconstructed by a neural network from the signals of four microphones placed in the far field.

In the microwave domain, an object was localized with a subwavelength resolution in a chaotic cavity. Coded-aperture imaging effect was achieved by a metasurface consisting of an array of individually tunable boundary conditions. The signal was captured at a single frequency and at a single location but for a fixed series of random configurations of the metasurface for each of the object locations [50].

1.1.4 Ultrasonic Localization Based on Machine Learning

Two studies were addressing the localization problem by ultrasound analysis based on machine learning. The first was demonstrating a system able to localize a finger touch on a metal plate [51]. The sensing area was surrounded by 4 transmitters and 12 receivers and all the 144 (12 × 12) touch positions were used in the training process. The second addressed the source localization of ultrasound emitted by pencil lead-breaks on the surface of a composite plate [52]. 8 sensors were used to classify between five zones. In both cases, the target localization resolution was not significantly below the training resolution.

1.2 Knowledge Gap

The main challenge is that typically the machine learning algorithms only work reliably under the conditions closely related to those of the training process. They get ineffective if ultrasound acquisition parameters, ultrasound coupling properties, specimen characteristics, probe location, or other conditions change.

Majority of the previous machine learning experimental studies in the domain of ultrasonic inspection suffered of low data diversity, as it is practically challenging to obtain a suitable experimental database containing multiple target feature (defect) types and system configurations.

The motivation of our work is to test how the efficiency of the machine learning algorithms depends on the size and diversity of the training datasets. More precisely, we aim to measure the precision of damage localization in cases when signals used for training are obtained at a low number of damage locations. We address the question if an appropriately trained algorithm can make two dimensional defect localization possible by a single-point ultrasound excitation and detection without performing a scan.

Our study is the first using a reconfigurable defect to experimentally obtain large volumes of diverse datasets of laser ultrasonic signal. We are able to vary both the defect position and the specimen configuration, which both affect the ultrasound propagation. This opens the possibilities to study the robustness of the model and what size of the training

dataset is required for certain performance and how related it should be to the testing dataset.

Our method is able to achieve subwavelength localization with a resolution up to 5 times below the one of the training data.

2 Methods

2.1 Specimen Description

The main element of the specimen was a foil with 0.2-mm thickness made of a shape memory polymer (SMP), (manufacturer: SMP Technologies Inc, Tokyo), with a glass transition temperature in the range between 25 and 90 °C. As its Young's modulus continuously falls by at least a factor 20 by increasing its temperature a few 10 °C above the room temperature [53], it is an appropriate material to simulate a reconfigurable defect by local heating. A localized decrease of Young's modulus and increase of ultrasound absorption has in a certain degree of approximation similar effect on the ultrasound as a crack, local porosity, or a local change of material or geometric properties. However, reconfigurable heat-induced defect is a distinct type of damage, which cannot be directly compared with the real damage, which might feature a complex interaction with the sensing wave: scattering, nonlinear effects, clapping of a delamination, and others.

Its advantage exploited in this study is that the size shape and location of such a defect are easily reconfigurable, if we have control over the temperature field of the SMP foil.

The SMP foil was placed in an arbitrary shaped frame (Fig. 1) consisting of two aluminum plates (with the thickness

of 1 mm each) pressed together. The both plates had identical holes with the disordered shape. The Lamb waves, propagating in the SMP foil, were reflecting from the hole boundaries, where the foil was clamped in the frame. One part of the specimen was covered by carbon powder in order to achieve high light absorption (for efficient laser-based ultrasonic excitation) and another part was covered by a retroreflective foil (for efficient laser-based ultrasonic detection). The mechanic properties of the specimen were changed by sticking a first (100 mg) and a second mass (20 mg) on the SMP foil at the positions marked in Fig. 1a. This provided us three specimen configurations: the first configuration is without mass, the second configuration is with both masses, and the third configuration is with the mass number 1 alone. The masses were attached at the positions shown in Fig. 1 using a temporary adhesive. In order to make the wave propagation even more chaotic, a notch with the length of 2.5 mm and maximal the width of 0.5 mm was induced in the middle of the positions of the reconfigurable defects, which were simulated by heated spots. It was created by cutting the SMP foil by a knife.

2.2 Experimental Setup

Lamb waves were excited at a single arbitrary chosen location by a laser pulse with wavelength of 500 nm, energy of 9 mJ, duration of 5 ns (full width at half-maximum), and repetition rate of 20 Hz using a Surelite SL I-20 pump laser together with a Surelite OPO Plus optical parametric oscillator (manufacturer of both: Continuum). The precise wavelength of the ultrasound excitation laser is not relevant and was chosen arbitrary. An advantage of using an optical parametric
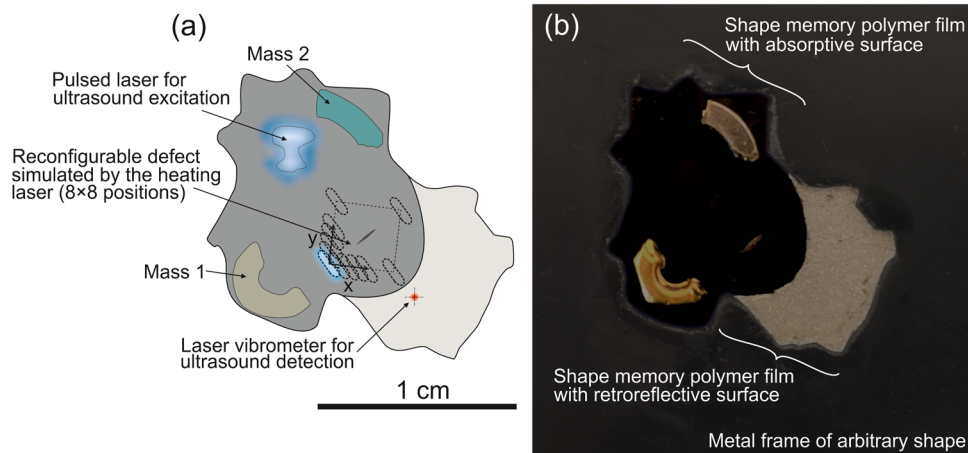

Fig. 1 Scheme (**a**) and photo (**b**) of the specimen – a shape memory polymer (SMP) foil in an aluminum frame of arbitrary shape. Ultrasound was generated by a laser pulse and detected by a laser vibrometer at fixed positions. An artificial reconfigurable defect was created by locally reducing the Young's modulus and increasing the ultrasound

attenuation level of the SMP by a heating laser. Specimen configuration was changed by applying two masses on the SMP foil. The light and dark grey areas correspond to the SMP polymer regions with high and low laser light reflectivity
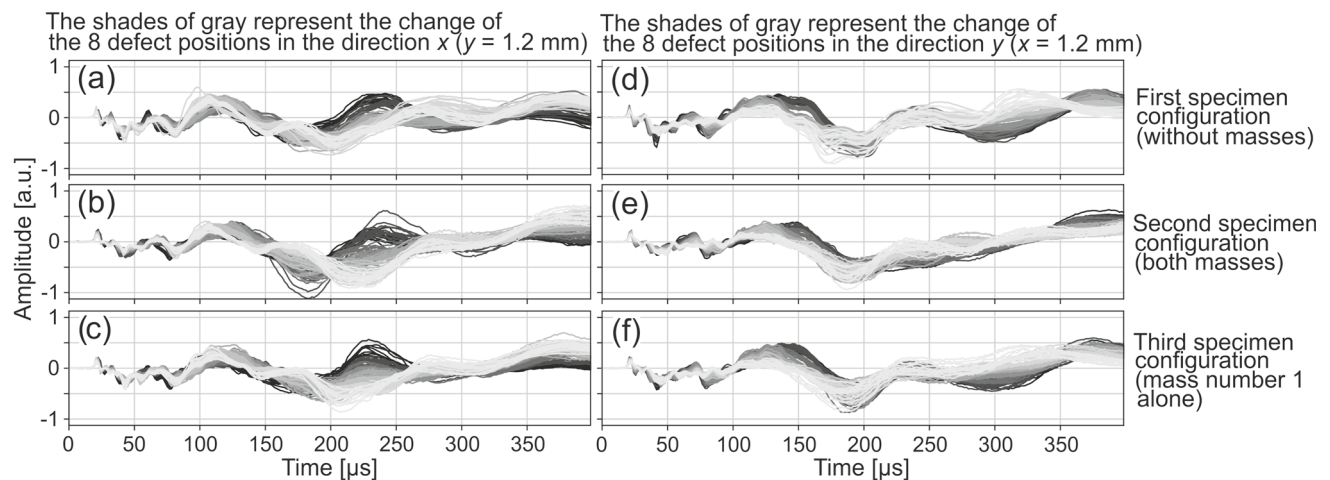
Fig. 2 Raw experimentally obtained signals used for the model training. In the shades of gray, change of the defect position in the direction x (**a–c**) and direction y (**d–f**) is shown for first (**a, d**), second (**b, e**), and third (**c, f**) specimen configuration. Without machine learning algorithms, it is not possible to localize the defect and distinguish between the different specimen configurations. Please note that in order to increase figure clarity, only 624 (13 repetitions $\times$ 8 defect positions $\times$ 2 defect position dimensions $\times$ 3 specimen configurations) signals are shown instead of altogether 3840 signals (20 repetitions $\times$ 8 defect positions x $\times$ 8 defect positions y $\times$ 3 specimen configurations) used in our study

oscillator was that we were able to detune its optimal optical configuration and cause the excitation laser to become unstable, simulating a chaotic excitation scenario. For each of the ultrasound excitation, pulse energy, shape, and diameter of the laser beam were randomly varying for up to 50%. This decreased the repeatability of the ultrasound generation, brought more noise in the measurement and provided more stochastic variations in the signals, which are more challenging but more interesting to be evaluated by machine learning algorithms. This is observable in Fig. 2 where the signals of the same gray shade were subject to stochastic variations despite that they were captured at the same conditions. The motivation for this was to make our experiment more similar to industrial quality control processes, which are typically affected by variations in laser light absorption, environmental noise, or similar disturbing parameters. Ultrasound was detected at a single location using PSV-F-500-HV laser vibrometer (manufacturer: Polytec) with 15 signal averaging.

A third laser used in our experiment was a SuperK supercontinuum white light laser with a SuperK Varia tunable wavelength filter (manufacturer of both: NKT Photonics). Its power was approximately 0.2 W and the wavelength range was arbitrary chosen to range from 400 and 500 nm. A single wavelength laser could also be used instead. XG210 2-axis galvanometer scan head (manufacturer: Mecco) was used to project a 2 mm long and 0.5 mm broad line and locally heat the SMP foil. The orientation of the line was chosen to be perpendicular to the notch length in order to decrease the variation of the heating laser light absorption. It might have decrease if the line was fully overlapping with the notch. By the local drop of the Young's modulus and local increase of the ultrasound absorption, a reconfigurable defect was created. Due to the thermal conductivity of the SMP, the 4 mm wide and 6 mm long (full width at half maximum) area around the projected line was featuring a decrease of the Young's modulus.

2.3 Acquisition of Datasets

Altogether 3840 signals were captured with the length of 250 points and the sampling frequency of 625 kHz. 20 measurements were repeated for each of the 64 reconfigurable defect positions (8×8 positions in two dimensions defined as x and y) and for each of the three specimen configurations (without mass, with both masses, and with a mass number 1 alone). The step distances in the x and y directions between the two defect positions were 0.4 mm. After each change of the defect position, we waited for 20 s in order to reach a new thermal equilibrium.

The 20 repeated signal measurements for each of the defect positions and each of the specimen configurations were firstly divided to 13 signals exclusively used for the model training and the remaining 7 signals exclusively used for the model testing. This remains true for all the studies of our work. As described in the following section, the number of defect positions and specimen configurations was additionally reduced for the model training. This included the cases where the network was trained on only a few defect

positions and tested on all the 64 positions as described in Sect. 3.2.

A part of raw measured signals is shown in Fig. 2. The signals consisted of a superposition of multiple unidentified symmetric and antisymmetric Lamb wave modes. The change in the direction x of the defect positions at $y = 1.2$ mm is coded by the shades of gray in the left-hand column ($x = 0$ mm for the black, $x = 3.2$ mm for the bright gray). The change in the direction y of the defect positions at $x = 1.2$ mm is coded by the shades of gray in the right-hand column ($y = 0$ mm for the black, $y = 3.2$ mm for the bright gray). The signals of the same shade of gray were consecutively obtained by repeating the measurement at the same defect position. Each of the three rows represents from the top to the bottom: the first specimen configuration (without additional mass), the second specimen configuration (with two masses), and the third specimen configuration (with only mass number 1 attached on the SMP foil). Please note that for the purpose of easier visualization, a single x and a single y position was chosen to demonstrate the signal variation in the y and x direction, respectively. For the same reason, only 13 repetitions, which were used for the model training, are shown instead of all the 20 measurements at each of the defect position and at each of the specimen configuration.

The shape of the signal has complex and unknown dependency on the target parameters—the specimen configuration (presence of mass) and the defect position. From the raw signals shown in Fig. 2, it visually appears that the latter has stronger influence on the signal. It is not possible to visually distinguish between the different defect positions also in the case when the signals are converted to the frequency domain. The goal is to interpret the signals by extracting the target parameters without actually knowing the wave propagation and the measurement system properties.

3 Results and Discussion

Two studies are performed on the datasets obtained as described above. The first is a classification of the specimen configuration and the second is a localization of the reconfigurable defect. While the classification problem has a discrete output (label number) the localization has two analog outputs (defect positions x and y).

3.1 Classification of the Specimen Configurations

The aim of the first study is to classify the laser ultrasonic signals obtained at three different specimen configurations (three rows in Fig. 2). In this case, a disturbing factor is the defect position, which significantly affects the signal shape and should be eliminated (marked by the shades of gray in Fig. 2).

3.1.1 Model Architecture and Training Process

At first, the raw signals were converted to the frequency domain by the fast Fourier transform. The first half of the real and imaginary component values (125 values each) were merged together and used as a parallel input for the neural network. Consequently, the full frequency range up to the half of the sampling frequency was considered at the input. Performance dropped by around 10% in the case if raw time-domain signals were used as input for the neural network.

The model consisted of a fully connected neural network with an input of 250 elements, two hidden layers—the first with 200 elements and the second with 20 elements. The last (output) layer had 3 elements for 3 classes: first, second, and third specimen configuration. All three layers were linear and have a rectified linear unit activation function. An Adam optimization algorithm was used as it dynamically adjusts its step size. It is more efficient than stochastic gradient descent. The loss parameter was calculated by a cross entropy loss function. The neural network was trained by 50 epochs—passes of the entire training dataset. Alternative conditions to terminate the training process, which cannot guaranty the constant number of epochs, are not suitable for our study since we vary the size of the training process. The termination conditions might not be reached at lower sizes of the training dataset. For all the training datasets, the convergence rate drastically decreased before 50 epochs. The batch size was always equal to the whole training dataset size, which was a changing variable in our research as described below. Alternative less efficient network architectures are described in the following section.

In Fig. 3, we show results of the classification between the three different configurations of the specimen—between the three rows of Fig. 2. We test the performance of the machine learning algorithm on the diversified datasets if a reduced amount of training data is used. Signals obtained at a limited number of defect positions (which is here the disturbing parameter) was used for the model training (starting with a single defect position), while it was tested for all the defect positions. The number of signals used for the training was reaching from 13 (13 repetitions at a single defect position) to 832 (13 repetitions at 64 defect position). The model was always tested on the dataset obtained at all the 64 defect positions with 7 repetitions each—altogether 448 signals, which were always all different from the training dataset.

The final performance of the model tested for all the defect position depends on the defect positions we choose for the model training. The defect positions lying in the middle of the defect position area deliver better performance comparing to those lying at the edge of the scan. In our statistical approach, we repeated the training process 20 times for each number of the defect positions, which were chosen randomly. We determined the labeling accuracy (red full line) and the loss
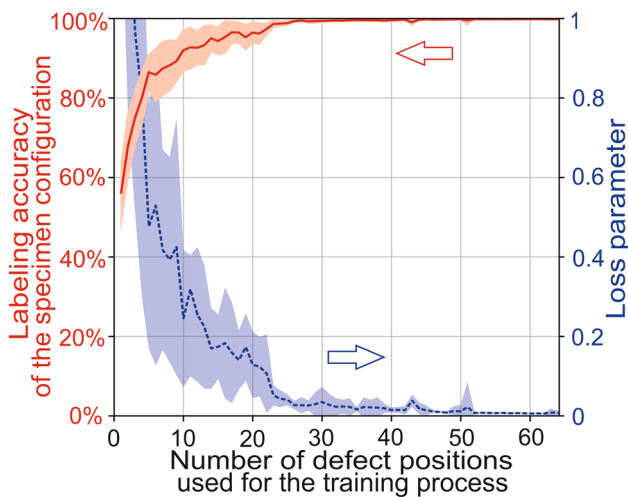
Fig. 3 Labeling accuracy (red full line) and loss parameter (blue dashed line) for the classification between the three specimen configurations in dependency on the number of defect positions used for training. The defect position is in this case a disturbing parameter, which significantly affects the signal shape. The model was tested on the previously unseen dataset containing signals of all defect positions. Good performance and robustness of the model is achieved for reduced training datasets as well

parameter (blue dashed line) at the end of the learning process (Fig. 3). They are represented by the mean value and the shadowed standard deviation range.

The model has a moderate performance on the testing dataset containing signals of all defect positions if the training is performed on a dataset only containing the signals of one defect position (56% labeling accuracy). If additional defect positions are included in the training dataset, the accuracy rapidly increases and reaches 90% if only 8 defect positions are used for the training process. This is remarkable as the defect position (shades of gray in Fig. 2) visually appears to have higher influence on the signal shape than the specimen configuration (three rows in Fig. 2). Influence of the disturbing factor (changing defect position) is eliminated also if training is performed on datasets with limited diversity (limited number of detect locations) and the model robustness is extended to related, but previously unseen data.

3.2 Defect Localization

The second study of our work is using the same experimental data as the first study, but we reverse the problem. Its goal is to localize the defect (which affects the signal as marked by the shades of gray in Fig. 2) and a disturbing factor is the drastic change of the specimen configuration (three rows in Fig. 2) by applying a mass, on which the model should be made robust.

3.2.1 Model Architecture and Training Process

Signal preprocessing for the second study of this work was the same as for the first study: signal conversion to the frequency domain and parallel merging of the real and imaginary components.

The simplest model with sufficient efficiency, which was chosen to be included in the study, consisted of a fully connected neural network with an input of 250 elements, two hidden layers—the first with 500 elements and the second with 100 elements. The last layer had only two elements which were quasi-continuous with high numeric precision and directly represented the defect positions x and y. The loss function was defined as the mean absolute error of these two values. Similarly as for our first study, linear layers and a rectified linear unit activation function were chosen, together with the Adam optimization algorithm and a scheduler reducing the initial learning rate of 0.01 for 10% every 100 epochs, which provided sufficient convergence before the end of the training process. Because of a higher range of possible outputs (defect positions in two dimensions), and because it is about localization and not classification, the learning process for the second study of this work was approximately a hundred times longer than for the first study. The training process comprised 1000 epochs. The batch size was always equal to the whole training dataset size, which was a changing variable (described below), as it was the case for our first study as well.

3.2.2 Less Efficient Alternatives of Model Architecture and Training Procedure

Typically, a higher number of neural network layers and a higher number of elements in each layer improved the performance of the model. However, higher model complexity made the training process slower. The performance of the model was significantly reduced if only one hidden layer was used and a neural network with a single layer (similar to multi-variable regression) was inefficient. If the raw signals are directly used as the input for the neural network (without conversion to the frequency domain), its performance decreases by around 10%. Stochastic gradient descent was less efficient than the Adam optimizer.

The optimization of the network architecture was performed on a dataset consisting of 13 repetitions at 64 defect positions obtained at the third specimen configuration. The test dataset was obtained at the same specimen configuration. As described in Sect. 2.3, the test dataset and the train dataset were not overlapping. In Table 1, we show absolute positioning error and labeling accuracy of the output provided by the trained neural networks having six different sizes, with and without the second hidden layer.

Table 1 Absolute positioning error and labeling accuracy at six different sizes of the neural network

Number of elements—first hidden layer	500	250	125	500	250	125
Number of elements—second hidden layer	100	50	25	–	–	–
Absolute positioning error [μm]	78	80	95	83	85	91
Labeling accuracy	91%	88%	83%	88%	87%	83%

3.2.3 Reduced Training Datasets and Model Robustness

In order to estimate the robustness of the model for the defect localization, we tested its performance with reduced datasets and datasets significantly different to the testing dataset were used for the training process. Please note that during the second study of our work, the 7 signal repetitions measured at each of the defect positions obtained at the third specimen configuration were always used for the model testing. These signals were never included in the training datasets.

For the first training process, we used a dataset consisting of 13 repetitions at a single randomly chosen defect position obtained at the first specimen configuration. The second training process used dataset consisting of 13 repetitions at two randomly chosen (different from each other) defect positions at the first specimen configuration; and so on until the 64th training process, where we used a dataset consisting of 13 repetitions at all of the 64 defect positions at the first specimen configuration. In order to reduce the uncertainty of randomly reducing the training dataset, we repeated this procedure for 20 times. In order to obtain the black dotted curve in Fig. 4a and c, we performed altogether 1280 training processes—at 20 randomly chosen defect positions with the number of different defect positions used for the training process reaching from 1 to 64.

The same procedure was repeated for the second specimen configuration and for the combination of the first and the second specimen configuration. For the latter, we used the same randomly chosen defect positions for datasets obtained at both specimen configurations. This delivered us the blue dashed line and the red dash-dot line in Fig. 4a and c, respectively.

Subsequently, we gradually added ultrasonic signals obtained at the third specimen configuration. Please note that we never used the same ultrasonic signals for the model training and the model testing.

The first training process used a dataset consisting of 13 repetitions at a single randomly chosen defect position at the third specimen configuration, together with 13 repetitions at all defect positions at the first specimen configuration—altogether $13 \times (64 + 1)$ signals. The second training process used a dataset consisting of 13 repetitions at two always randomly chosen (different from each other) defect positions at the third specimen configuration, together with 13 repetitions

at all defect positions at the first specimen configuration—altogether $13 \times (64 + 2)$ signals; and so on until the 64th training process, where we used a dataset consisting of 13 repetitions at 64 defect positions at third specimen configuration and 13 repetitions at all of the 64 defect positions at the first specimen configuration—altogether $13 \times (64 + 64)$ signals. In order to obtain the black dotted curve in Fig. 4b and d, we performed altogether 1280 training processes—at 20 randomly chosen defect positions with the number of different defect positions at the third specimen configuration used for the training process reaching from 1 to 64.

The same procedure of gradually adding more of the defect positions obtained at the third specimen configuration was repeated while exchanging all the defect positions at the first specimen configuration with all the defect positions at the second specimen configuration (blue dashed line in Fig. 4b and d). The procedure was followed for the third time using all the defect positions at the first and the second specimen configuration together (red dash-dot line in Fig. 4b and d). As a reference, we repeated the same procedure for the fourth time without any training data of the defect positions obtained at the first and the second specimen configuration (green full line in Fig. 4b and d).

The positioning error in Fig. 4a and b represents the mean value of absolute errors between the predicted and the real defect positions in the dimensions x and y. If this absolute error is smaller than 0.2 mm for both dimensions, which is the half of the minimal distance between the two defects, its position is labeled correctly. The percentages of correct labeling are shown in Fig. 4c and d.

In Fig. 4a and c, we can observe that the performance of the model on the testing dataset, obtained at the specimen configuration different from the training dataset (third specimen configuration), is increased if the model is trained on a mixed dataset containing signals obtained at two different (first and second) specimen configurations (red dash-dot line). In this case, the model is robust to the change of the specimen configuration and the mean positioning error reaches the values below the half of the minimal distance between two defect positions used for the training process in this study (0.4 mm). Consequently, the labeling accuracy increases to 50%, while it remains low if the undiversified training dataset is used (black dotted and blue dashed line).

If the training process is performed on the dataset obtained at the specimen configurations different to the testing dataset,

Fig. 4 Mean positioning error and labeling accuracy of the defect positions if the model is trained on reduced datasets and different specimen configurations. 20 repetitions of the training process are represented by the mean value (solid line) and the standard deviation range. The thick orange vertical line represents the case shown in Fig. 5. The diversified datasets make the model robust. The labels of the horizontal axes continue in the legend of (**c**)
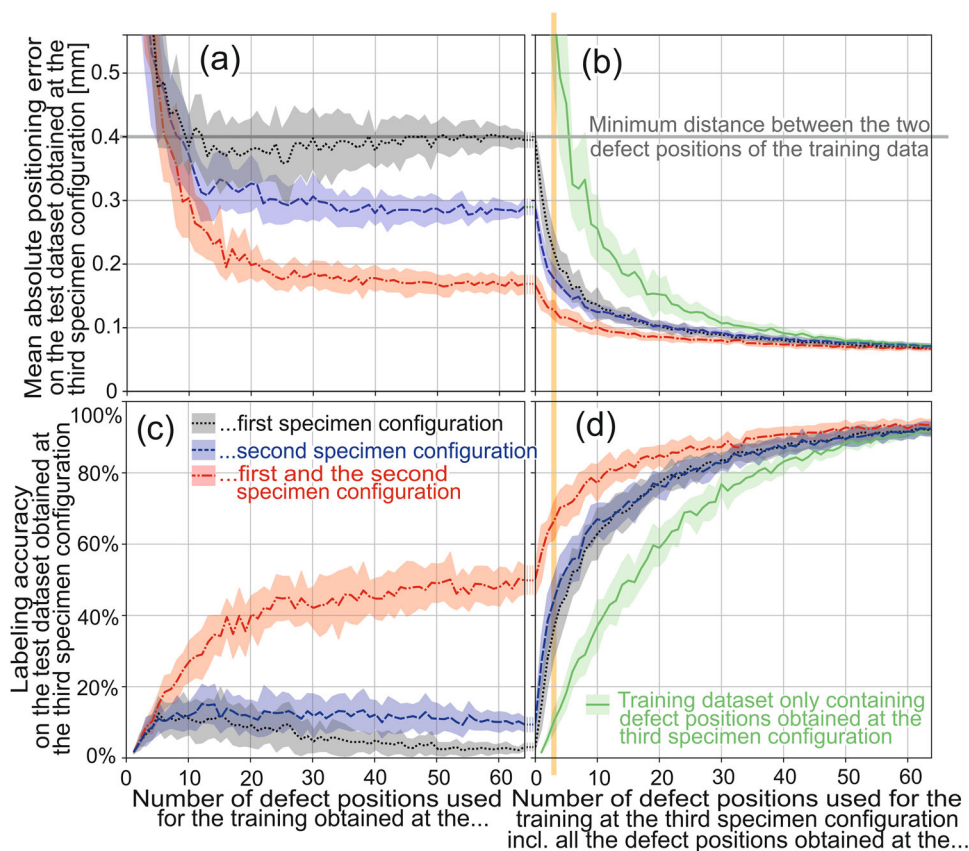
around 21 randomly chosen defect positions already deliver performance close to the one when all the defect positions are used for training.

The end values of Fig. 4a and c are equal to the start values of Fig. 4b and d. If the model is trained by a dataset containing all defect positions obtained at two different specimen configurations (red dash-dot line), both different to the testing datasets, the positioning error is decreased almost for the factor of 2 and the labeling accuracy is improved for the factor of 5, in comparison to the case when a dataset only containing signals obtained at the first (black dotted line) or the second (blue dashed line) specimen configurations.

In this case the model performance is high already when including only a few additional defect positions from the dataset obtained at the third specimen configuration equivalent to the testing. After 11 additional defect positions, positioning error below 0.1 mm and labeling accuracy of 80% is achieved (red dash-dot line in Fig. 4d). Neural network without the training on previous (different) datasets achieves the same performance when 35 different defect positions are used for the training.

Please note that the datasets of the ultrasonic signals obtained at the three specimen configurations are significantly different from each other, as proven by the classification problem described in the first study of this work.

In the first row of Fig. 5, we show the performance of the model during the training process at the situation marked with the thick orange vertical line in Fig. 4. The final performance (after 1000 epochs) for the defect localization on the training dataset (second row) and the testing dataset (third row) is visually represented. In order to provide better comparison, the training dataset always comprised the same 3 defect positions at the third specimen configuration (red dots with full outline). Similar as for Fig. 4, where the defect positions were always chosen randomly, 13 ultrasonic signals repetitions at each of these three positions were used for the first column of Fig. 5. We additionally included all the defect positions obtained at the first specimen configuration (green dots with dashed outline) for the second column in Fig. 5, all the defect positions obtained at second specimen configuration (orange dots without outline) for the third column in Fig. 5, and both datasets (obtained at the first and second specimen configurations) for the fourth column in Fig. 5.

The testing dataset was always the same—all defect positions at the third specimen configuration. In order to better distinguish between the different defect positions, red and blue colors are alternating in x dimension and full and empty filling are alternating in y dimension. In the background, the gray shading shows the mean absolute error of each of the defect with the black representing the largest error.
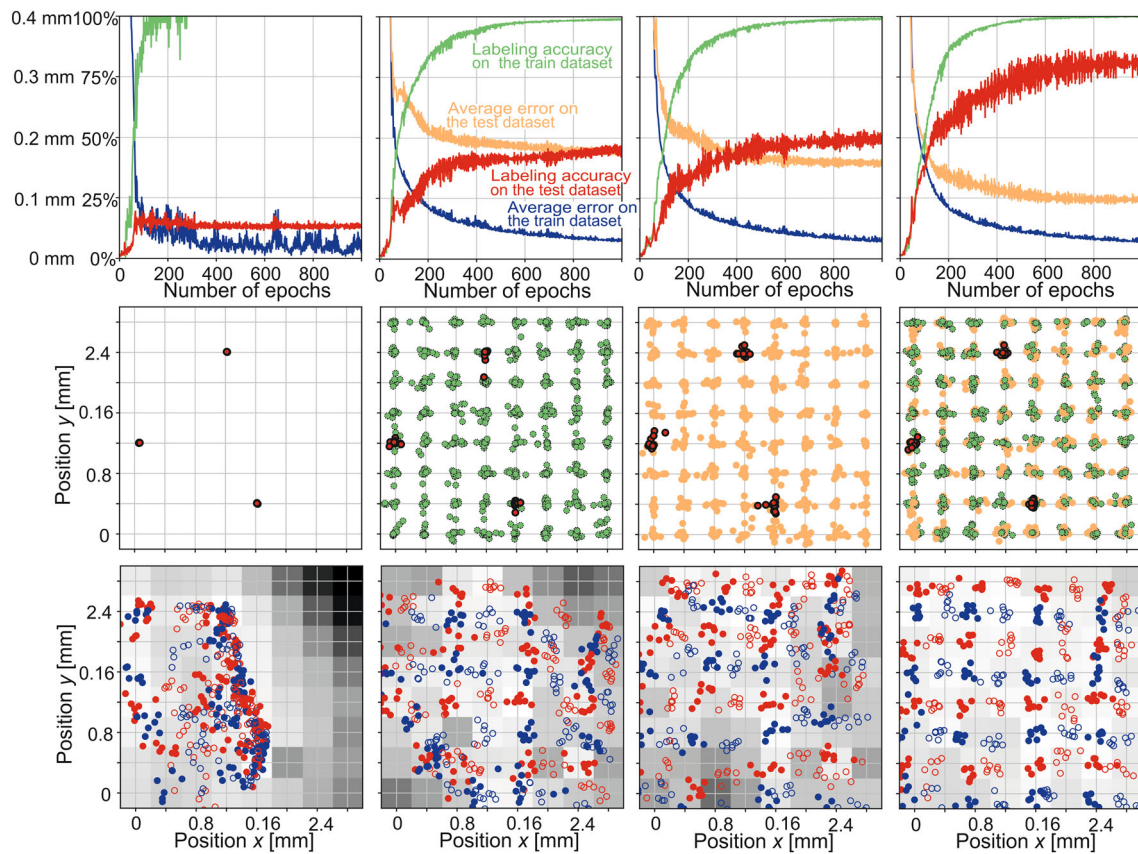
Fig. 5 Performance of the model during a single training process (first row) and its final performance on the training (second row) and testing (third row) datasets. No localization is possible if only three defect positions from the dataset obtained at the specimen configuration equivalent to the testing dataset are used for the training process (first column). The performance of the model is partly improved if an additional dataset containing all defect positions is used for the training process even if they were obtained at the specimen configuration different to the one of testing (second and third column). The localization error decreases to 0.25 mm (which delivers the labeling accuracy of 80%) if the robustness of the model is improved by a diversified dataset containing all defect positions obtained at two different specimen configurations and only three defect positions obtained at the specimen configuration equivalent to the testing dataset (fourth column)

The localization is not possible if only three defect position are used to train the model (first column in Fig. 5), because of the lack of the information about the rest of the defect positions. The performance of the model significantly improves, if other defect positions are included in the dataset, even though they were not obtained at the same specimen configuration. Comparing the second and third column of Fig. 5, we can observe that the second specimen configuration is more similar to the third specimen configuration, because of its slightly better training performance. However, labeling accuracy does not lift above 50% and the mean absolute error is not much below 0.17 mm.

The model become robust to the change of the specimen configuration if all the defect positions obtained at the first and second specimen configuration are used for training. Labeling accuracy of 80% and mean absolute error of 0.25 mm is achieved already if only three defect positions at the specimen configuration equivalent to the testing dataset are used for training.

Predicted defect positions are scattered also when the method is tested on the training dataset. This is showing the high stochastic variability of all the datasets because of the low repeatability of the ultrasound excitation, which was induced artificially.

Table 2 show mean absolute positioning error and labeling accuracy at the end of the training process for eleven different training datasets (13 repetitions for each of the defect position as specified in Sect. 2.3). The test dataset was the same as for Figs. 4 and Fig. 5—separately obtained at the third specimen configuration.

Table 2 Mean absolute positioning error and labeling accuracy at different sizes of the train dataset

	Number of defect positions used for training										
First specimen configuration	64	0	64	0	64	0	64	0	64	0	64
Second specimen configuration	0	64	64	0	0	64	64	0	0	64	64
Third specimen configuration	0	0	0	3	3	3	3	32	32	32	32
Absolute positioning error [μm]	397	291	168	603	220	178	129	104	84	87	80
Labeling accuracy	3%	9%	49%	10%	36%	44%	66%	77%	85%	85%	87%

4 Conclusions

We introduce a novel experimental method to generate large labeled datasets of laser ultrasonic inspection signals. This open a possibility to validate and optimize the architecture of machine learning algorithms for the application of ultrasound signal interpretation. We study their performance and robustness if limited and/or only partly related training datasets are used for the model training. We prove that by the use of machine learning algorithms, it is possible to extract information, difficult to access by conventional ultrasonic signal processing methods. Due to the complex effect of the inspected property on the signal shape and subwavelength size of target features (defects), it is impossible to directly interpret the raw signal, for example, by observing the time of flight or the amplitude variations. The algorithm can be trained to identify relevant signal parameters with their interdependences and to correlate them to the desired output. The mechanism behind cannot be fully interpreted by humans, as well as it is in the most cases impossible to visually differentiate between differently labeled ultrasonic signals. The relevant information can be extracted from the ultrasonic signals without knowing the physical model of the system i.e. without actually understanding the wave propagation (differentiating between the wave types, reflections etc.). In the case of our study, it would not be possible because of the disordered specimen shape and undefined and unstable laser excitation properties. The additional advantage of machine learning ultrasound interpretation is that the ultrasound excitation and detection devices do not require to be each of them separately calibrated. This is proven by our setup since the lasers for ultrasound excitation and detection are not calibrated and the ultrasound propagation in the complex shape of the sample are not modeled. In a certain way, the training process itself is the calibration process of the whole system including wave propagation in the specimen, ultrasound excitation, detection, and coupling properties, together with the signal processing elements, which can all be undefined. Our results show that only a small dataset obtained at new conditions is required in order to restitute the algorithm's performance after the change of environmental conditions (specimen configuration). The training process is more flexible method of calibration, which allows the system being only partly calibrated by less-related or small training datasets. The increased number of trained parameters can make the system robust to variations of system properties without manually switching between different calibration settings related to the specific system properties.

In the first study of our work, the model trained on a reduced dataset (reduced number of different defect positions) is able to discern with high reliability the specimen configurations and is robust to the change of the defect position, which is otherwise having a stronger effect on the signal shape. We reverse the problem for the second study of our work. If a mixed dataset of two specimen configurations is used for the training process, the model for defect localization becomes robust to the different related change of the specimen configuration. The localization error is smaller than 0.1 mm (four times less than the minimum resolution of the training dataset) if only 10 additional training defect positions obtained at the specimen configuration equivalent to the one of the testing dataset are included. This is giving some evidence that if the training dataset is sufficiently diversified, it is possible to make the neural network for ultrasound interpretation robust to the change of the specimen properties. Our study suggests the potential for this being also true for the change in ultrasound excitation (e.g. energy, position, form) detection (e.g. position, gain) and other system properties. Likewise, the neural network can be made robust to some disturbances (in our case variation of ultrasound excitation properties) and to the certain degree of noise level. Our results show also that model performance, sufficient for multiple applications, can also be achieved if the model is trained on a small-sized dataset partly different from the testing dataset.

These outcomes of our research show a high potential of the machine learning algorithms for the laser ultrasound signal interpretation for the specific industrial inspection applications as, for example, two-dimensional defect localization on a flat sample without performing a scan. Reduced size of training datasets would suffice for numerous cases to achieve sufficient performance and the training is not even necessary to be performed on datasets closely related

to the testing dataset. If the testing object or the inspection system is moderately changed, only a small size of the training data is required to achieve the same correctness. After this change and additional training, the model becomes increasingly more robust to this type of change. This has practical advantages for modern lean industries demanding high flexibility, autonomous production, and interconnectivity of machines, products, and personnel. Only several test objects are necessary to be inspected destructively for the purpose of the model training, while the quality state of the rest can be with high fidelity estimated by non-destructive inspection method and ultrasound signal interpretation based on machine learning method described in this work, which is an essential part of the cyber-physical, non-destructive evaluation concept known under the term NDE 4.0 [54–61]. We expect that the results of our work are potentially transferable to other broadband ultrasonic inspection methods (e.g. using a single-point contact piezoelectric transducer) and possibly to additional defect characteristics (e.g. size or orientation).

Declarations

Conflict of interest The authors have no competing interests, or other interests that might be perceived to influence the results and/or discussion reported in this paper. A PCT patent application has been filed related to the topic of this publication.

Ethics approval and consent to participate Not applicable.

Consent for publication The Authors gives the Publisher the permission to publish the Work.

References

1. Chen, X.W., Lin, X.: Big data deep learning: challenges and perspectives. IEEE Access **2**, 514–525 (2014). https://doi.org/10.1109/ACCESS.2014.2325029
2. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**(7553), 436–444 (2015). https://doi.org/10.1038/nature14539
3. Uijlings, J.R.R., van de Sande, K.E.A., Gevers, T., Smeulders, A.W.M.: Selective search for object recognition. Int. J. Comput. Vis. **104**(2), 154–171 (2013). https://doi.org/10.1007/s11263-013-0620-5
4. Guodong, G., Li, S.Z.: Content-based audio classification and retrieval by support vector machines. IEEE Trans. Neural Netw. **14**(1), 209–215 (2003). https://doi.org/10.1109/TNN.2002.806626
5. McLoughlin, I., Zhang, H., Xie, Z., Song, Y., Xiao, W.: Robust sound event classification using deep neural networks. IEEE/ACM Trans. Audio Speech Lang. Process. **23**(3), 540–552 (2015). https://doi.org/10.1109/TASLP.2015.2389618
6. Noda, K., Yamaguchi, Y., Nakadai, K., Okuno, H.G., Ogata, T.: Audio-visual speech recognition using deep learning. Appl. Intell. **42**(4), 722–737 (2015). https://doi.org/10.1007/s10489-014-0629-7
7. Caggiano, A., Zhang, J., Alfieri, V., Caiazzo, F., Gao, R., Teti, R.: Machine learning-based image processing for on-line defect recognition in additive manufacturing. CIRP Ann. **68**(1), 451–454 (2019). https://doi.org/10.1016/j.cirp.2019.03.021
8. Ngan, H.Y.T., Pang, G.K.H., Yung, N.H.C.: Automated fabric defect detection—a review. Image Vis. Comput. **29**(7), 442–458 (2011). https://doi.org/10.1016/j.imavis.2011.02.002
9. Tao, X., Zhang, D., Ma, W., Liu, X., Xu, D.: Automatic metallic surface defect detection and recognition with convolutional neural networks. Appl. Sci. (2018). https://doi.org/10.3390/app8091575
10. Fuchs, P., Kröger, T., Garbe, C.S.: Defect detection in CT scans of cast aluminum parts: a machine vision perspective. Neurocomputing **453**, 85–96 (2021). https://doi.org/10.1016/j.neucom.2021.04.094
11. Schlotterbeck, M., Schulte, L., Alkhaldi, W., Krenkel, M., Toeppe, E., Tschechne, S., Wojek, C.: Automated defect detection for fast evaluation of real inline CT scans. Nondestruct. Test. Eval. **35**(3), 266–275 (2020). https://doi.org/10.1080/10589759.2020.1785446
12. Rus, J., Gustschin, A., Mooshofer, H., Grager, J.-C., Bente, K., Gaal, M., Pfeiffer, F., Grosse, C.U.: Qualitative comparison of non-destructive methods for inspection of carbon fiber-reinforced polymer laminates. J. Compos. Mater. **57**(27), 4325–4337 (2020). https://doi.org/10.1177/0021998320931162
13. Rus, J., Kulla, D., Grager, J.C., Grosse, C.U.: Air-coupled ultrasonic inspection of fiber-reinforced plates using an optical microphone. In: Proceedings of German Acoustical Society, DAGA Rostock, pp. 763–766 (2019)
14. Blomme, E., Bulcaen, D., Declercq, F.: Air-coupled ultrasonic NDE: experiments in the frequency range 750 kHz–2 MHz. NDT E Int. **35**, 417–426 (2002). https://doi.org/10.1016/S0963-8695(02)00012-9
15. Rus, J., Grosse, C.U.: Thickness measurement via local ultrasonic resonance spectroscopy. Ultrasonics **109**, 106261 (2021). https://doi.org/10.1016/j.ultras.2020.106261
16. Rus, J., Grosse, C.U.: Local ultrasonic resonance spectroscopy: a demonstration on plate inspection. J. Nondestruct. Eval. (2020). https://doi.org/10.1007/s10921-020-00674-5
17. Migliori, A., Sarrao, J.L.: Resonant Ultrasound Spectroscopy: Applications to Physics, Materials Measurements, and Nondestructive Evaluation. A Wiley-Interscience publication, New York (1997)
18. Jüngert, A., Grosse, C., Krüger, M.: Local acoustic resonance spectroscopy (LARS) for glass fiber-reinforced polymer applications.

J. Nondestr. Eval. **33**(1), 23–33 (2014). https://doi.org/10.1007/s10921-013-0199-3

19. Solodov, I., Bai, J., Busse, G.: Resonant ultrasound spectroscopy of defects: case study of flat-bottomed holes. J. Appl. Phys. **113**(22), 223512 (2013). https://doi.org/10.1063/1.4810926

20. Blouin, A., Lévesque, D., Néron, C., Drolet, D., Monchalin, J.P.: Improved resolution and signal-to-noise ratio in laser-ultrasonics by SAFT processing. Opt. Express **2**(13), 531–539 (1998). https://doi.org/10.1364/OE.2.000531

21. Fendt, K.T., Mooshofer, H., Rupitsch, S.J., Ermert, H.: Ultrasonic defect characterization in heavy rotor forgings by means of the synthetic aperture focusing technique and optimization methods. IEEE Trans. Ultrason. Ferroelectr. Freq. Control **63**(6), 874–885 (2016). https://doi.org/10.1109/TUFFC.2016.2557281

22. Fendt, K.T., Mooshofer, H., Rupitsch, S.J., Lerch, R., Ermert, H.: Investigation of the synthetic aperture focusing technique resolution for heavy rotor forging ultrasonic inspection. IEEE Int. Ultrason. Symp. (IUS) (2013). https://doi.org/10.1109/ULTSYM.2013.0476

23. Ni, C.-Y., Chen, C., Ying, K.-N., Dai, L.-N., Yuan, L., Kan, W.-W., Shen, Z.-H.: Non-destructive laser-ultrasonic synthetic aperture focusing technique (SAFT) for 3D visualization of defects. Photoacoustics **22**, 100248 (2021). https://doi.org/10.1016/j.pacs.2021.100248

24. Camacho, J., Atehortua, D., Cruza, J.F., Brizuela, J., Ealo, J.: Ultrasonic crack evaluation by phase coherence processing and TFM and its application to online monitoring in fatigue tests. NDT E Int. **93**, 164–174 (2018). https://doi.org/10.1016/j.ndteint.2017.10.007

25. Piedade, L.P., Painchaud-April, G., Duff, A.L., Bélanger, P.: Compressive sensing strategy on sparse array to accelerate ultrasonic TFM imaging. IEEE Trans. Ultrason. Ferroelectr. Freq. Control **70**(6), 538–550 (2023). https://doi.org/10.1109/TUFFC.2023.3266719

26. Schmid, S., Wei, H., Grosse, C.U.: On the uncertainty in the segmentation of ultrasound images reconstructed with the total focusing method. SN Appl. Sci. **5**(4), 108 (2023). https://doi.org/10.1007/s42452-023-05342-7

27. Tsai, C.D., Wu, T.T., Liu, Y.H.: Application of neural networks to laser ultrasonic NDE of bonded structures. NDT E Int. **34**(8), 537–546 (2001). https://doi.org/10.1016/S0963-8695(01)00015-9

28. Yang, J., Cheng, J., Berthelot, Y.H.: Determination of the elastic constants of a composite plate using wavelet transforms and neural networks. J. Acoust. Soc. Am. **111**(3), 1245–1250 (2002). https://doi.org/10.1121/1.1451071

29. Lefevre, F., Jenot, F., Ouaftouh, M., Duquennoy, M., Poussot, P., Ourak, M.: Laser ultrasonics and neural networks for the characterization of thin isotropic plates. Rev. Sci. Instrum. **80**(1), 014901 (2009). https://doi.org/10.1063/1.3070518

30. Oishi, A., Yamada, K., Yoshimura, S., Yagawa, G., Nagai, S., Matsuda, Y.: Neural network-based inverse analysis for defect identification with laser ultrasonics. Res. Nondestr. Eval. **13**(2), 79–95 (2001). https://doi.org/10.1080/09349840109409688

31. Zhang, K., Lv, G., Guo, S., Chen, D., Liu, Y., Feng, W.: Evaluation of subsurface defects in metallic structures using laser ultrasonic technique and genetic algorithm-back propagation neural network. NDT E Int. **116**, 102339 (2020). https://doi.org/10.1016/j.ndteint.2020.102339

32. Guo, S., Feng, H., Feng, W., Lv, G., Chen, D., Liu, Y., Wu, X.: Automatic quantification of subsurface defects by analyzing laser ultrasonic signals using convolutional neural networks and wavelet transform. IEEE Trans. Ultrason. Ferroelectr. Freq. Control **68**(10), 3216–3225 (2021). https://doi.org/10.1109/TUFFC.2021.3087949

33. Virkkunen, I., Koskinen, T., Jessen-Juhler, O., Rinta-aho, J.: Augmented ultrasonic data for machine learning. J. Nondestr. Eval. **40**(1), 4 (2021). https://doi.org/10.1007/s10921-020-00739-5

34. Koskinen, T., Virkkunen, I., Siljama, O., Jessen-Juhler, O.: The effect of different flaw data to machine learning powered ultrasonic inspection. J. Nondestr. Eval. **40**(1), 24 (2021). https://doi.org/10.1007/s10921-021-00757-x

35. Sambath, S., Nagaraj, P., Selvakumar, N.: Automatic defect classification in ultrasonic NDT using artificial intelligence. J. Nondestr. Eval. **30**(1), 20–28 (2011). https://doi.org/10.1007/s10921-010-0086-0

36. Rai, A., Mitra, M.: Lamb wave based damage detection in metallic plates using multi-headed 1-dimensional convolutional neural network. Smart Mater. Struct. **30**(3), 035010 (2021). https://doi.org/10.1088/1361-665x/abdd00

37. Yang, X., Chen, S., Jin, S., Chang, W.: Crack orientation and depth estimation in a low-pressure turbine disc using a phased array ultrasonic transducer and an artificial neural network. Sensors (2013). https://doi.org/10.3390/s130912375

38. Medak, D., Posilović, L., Subašić, M., Budimir, M., Lončarić, S.: Automated defect detection from ultrasonic images using deep learning. IEEE Trans. Ultrason. Ferroelectr. Freq. Control **68**(10), 3126–3134 (2021). https://doi.org/10.1109/TUFFC.2021.3081750

39. 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 (2021). https://doi.org/10.1016/j.ultras.2021.106436

40. Pyle, R.J., Bevan, R.L.T., Hughes, R.R., Rachev, R.K., Ali, A.A.S., Wilcox, P.D.: Deep learning for ultrasonic crack characterization in NDE. IEEE Trans. Ultrason. Ferroelectr. Freq. Control **68**(5), 1854–1865 (2021). https://doi.org/10.1109/TUFFC.2020.3045847

41. Martín, Ó., López, M., Martín, F.: Artificial neural networks for quality control by ultrasonic testing in resistance spot welding. J. Mater. Process. Technol. **183**(2), 226–233 (2007). https://doi.org/10.1016/j.jmatprotec.2006.10.011

42. 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. J. Manuf. Process. **52**, 26–34 (2020). https://doi.org/10.1016/j.jmapro.2020.01.047

43. Chen, D., Zhou, Y., Wang, W., Zhang, Y., Deng, Y.: Ultrasonic signal classification and porosity testing for CFRP materials via artificial neural network. Mater. Today Commun. **30**, 103021 (2022). https://doi.org/10.1016/j.mtcomm.2021.103021

44. Xue, R., Wang, X., Yang, Q., Xu, D., Sun, Y., Zhang, J., Krishnaswamy, S.: Grain size distribution characterization of aluminum with a particle swarm optimization neural network using laser ultrasonics. Appl. Acoust. **180**, 108125 (2021). https://doi.org/10.1016/j.apacoust.2021.108125

45. Melville, J., Alguri, K.S., Deemer, C., Harley, J.B.: Structural damage detection using deep learning of ultrasonic guided waves. AIP Conf. Proc. **1949**(1), 230004 (2018). https://doi.org/10.1063/1.5031651

46. Shukla, K., Di Leoni, P.C., Blackshire, J., Sparkman, D., Karniadakis, G.E.: Physics-informed neural network for ultrasound nondestructive quantification of surface breaking cracks. J. Nondestr. Eval. **39**(3), 61 (2020). https://doi.org/10.1007/s10921-020-00705-1

47. Song, H., Yang, Y.: Super-resolution visualization of subwavelength defects via deep learning-enhanced ultrasonic beamforming: a proof-of-principle study. NDT E Int. **116**, 102344 (2020). https://doi.org/10.1016/j.ndteint.2020.102344

48. Lerosey, G., de Rosny, J., Tourin, A., Fink, M.: Focusing beyond the diffraction limit with far-field time reversal. Science **315**(5815), 1120–1122 (2007). https://doi.org/10.1126/science.1134824

49. Orazbayev, B., Fleury, R.: Far-field subwavelength acoustic imaging by deep learning. Phys. Rev. X **10**(3), 031029 (2020). https://doi.org/10.1103/PhysRevX.10.031029

50. del Hougne, M., Gigan, S., del Hougne, P.: Deeply subwavelength localization with reverberation-coded aperture. Phys. Rev. Lett. **127**(4), 043903 (2021). https://doi.org/10.1103/PhysRevLett.127.043903

51. Chang, C.-S., Lee, Y.-C.: Ultrasonic touch sensing system based on lamb waves and convolutional neural network. Sensors (2020). https://doi.org/10.3390/s20092619

52. Sikdar, S., Liu, D., Kundu, A.: Acoustic emission data based deep learning approach for classification and detection of damage-sources in a composite panel. Compos. B Eng. **228**, 109450 (2022). https://doi.org/10.1016/j.compositesb.2021.109450

53. Firouzeh, A., Salerno, M., Paik, J.: Stiffness control with shape memory polymer in underactuated robotic origamis. IEEE Trans. Rob. **33**(4), 765–777 (2017). https://doi.org/10.1109/TRO.2017.2692266

54. Meyendorf, N., Ida, N., Singh, R., Vrana, J.: NDE 4.0: progress, promise, and its role to industry 4.0. NDT E Int. **140**, 102957 (2023). https://doi.org/10.1016/j.ndteint.2023.102957

55. Singh, R., Fernandez, R.S., Vrana, J.: Principles for successful deployment of NDE 4.0. J. Non Destruct. **19**(4), 28–34 (2022)

56. Valeske, B., Osman, A., Römer, F., Tschuncky, R.: Next generation NDE sensor systems as IIoT elements of Industry 4.0. Res. Nondestr. Eval. **31**(5–6), 340–369 (2020). https://doi.org/10.1080/09349847.2020.1841862

57. Vrana, J.: The core of the fourth revolutions: industrial internet of things, digital twin, and cyber-physical loops. J. Nondestr. Eval. **40**(2), 46 (2021). https://doi.org/10.1007/s10921-021-00777-7

58. Vrana, J., Meyendorf, N., Ida, N., Singh, R.: Introduction to NDE 4.0. Handbook of Nondestructive Evaluation 4.0, pp. 3–30 (2022). https://doi.org/10.1007/978-3-030-73206-6_43

59. Vrana, J., Singh, R.: NDE 4.0—a design thinking perspective. J. Nondestr. Eval. **40**(1), 8 (2021). https://doi.org/10.1007/s10921-020-00735-9

60. Vrana, J., Singh, R.: Value Creation in NDE 4.0: What and How. Handbook of Nondestructive Evaluation 4.0, pp. 1–27 (2021). https://doi.org/10.1007/978-3-030-48200-8_41-1

61. Vrana, J., Singh, R.: Cyber-physical loops as drivers of value creation in NDE 4.0. J. Nondestr. Eval. **40**(3), 61 (2021). https://doi.org/10.1007/s10921-021-00793-7