

A DEEP LEARNING APPROACH FOR DENOISING AIR-COUPLED ULTRASONIC RESPONDS DATA

Mikel David Jedrusiak and Frank Weichert

Department of Informatics, Technical University of Dortmund, Germany

ABSTRACT

Ensuring material quality is a central objective in production and manufacturing. Non-contact non-destructive testing methods without the use of coupling media are of particular interest with regard to mechanical or biochemical properties of the material. For this purpose, air-coupled ultrasonic is a useful method for quality control. The challenge is the poor signal-to-noise ratio, which makes it difficult to apply the classical approaches. This makes it impossible to distinguish between defect structures and noise. We are developing a method for denoising air-coupled ultrasonic data by applying deep neural networks by using a geometry-analytical component that detects defect structures. During the evaluation we show that we are able to obtain the data almost free of noise, so that incorrectly classified noisy pixels are mainly located at the edges of the defect structures, which cannot be clearly delimited. It is shown that the quality of the data is significantly improved for detection processes.

KEYWORDS

Air-coupled ultrasonic, Deep neural network, Denoising, Quality assurance, Non-destructive testing

1. INTRODUCTION

In industry, different forms and types of materials are used in production. Materials that can be affected by various influences during the production chain and outside the production cycle. These influences can cause the mechanical or biochemical properties of the material to change, resulting in a high risk potential. Examples include the aircraft and automotive industries, where material defects can have serious consequences [1]. In the event of an accident, vehicles can have a different impact behaviour because the material fails more quickly under the additional force. The same applies to aircraft whose components are exposed to high mechanical loads during take-off and landing as well as during flight [2]. In particular, damage may not be visible on the surface but may be present in the subsurface. Another important domain is the food industry. Due to production errors and defective packaging machines, small metal parts or glass splinters can be found in food products that a customer subsequently consumes. Accordingly, it is important that contaminated products are removed [3]. For this reason, it is necessary to perform quality assurance not only during production, but also at regular intervals in safety-critical areas. A challenge and problem of materials science is the technical detection of damages and impurities in materials, as well as material fatigue. One possibility are destructive inspection methods, which expose the material of a test sample to different loads, so that limit values can be determined. These can be tension or compression, bending or torsion and other forms of loads over variable time intervals. In particular, random samples can also be taken from a running production and tested for the corresponding quality characteristics and predefined loads that the object has to withstand in order to draw conclusions about the complete production. The disadvantage of the destructive methods is that the material or object to be examined is damaged, which means that it can no longer be used. In the case of the above example from aircraft, components would be tested

and damaged to determine that the component could have been used further. With these inspection methods, replacement would be unavoidable, resulting in immense costs. In such situations, non-destructive test methods are recommended for quality assurance [4]. Applied methods are ct, thermographic or ultrasonic techniques using different coupling mediums. With these approaches, the materials remain unaffected, but must be prepared in a pre-process if necessary. However, they cannot always be used, since there are hygienic reasons against thermography in the food industry and the preparation with oils or liquids as coupling medium for ultrasonics is not always possible. For this reason, air-coupled ultrasonic was developed, which avoids a coupling medium. Instead, the air is the transmission medium for the ultrasonic. The resulting advantages include the elimination of a time-consuming pre-process and the contactless inspection possibility, since the ultrasonic sensor does not require direct contact with the material to be tested. The challenge here is the analysis of the data, since the data is very noisy due to a poor signal-to-noise ratio. Our idea is the application of machine learning to optimize the air coupled ultrasonic data. Approaches from the field of deep learning have already shown that good results can be achieved in denoising and analysing data using recurrent networks and LSTMs in combination with an auto encoder, for example in speech recognition [5,6]. The problem that makes the direct apply of these classical approaches impossible is the indistinguishable characteristic of noise with the characteristic of damage in the material. In this paper, we want to present an approach that solves the problem of in distinguish ability and denoises air-coupled ultrasonic data that show defect structures. In this context, our main contribution is the development of a AI-based architecture that provides an additional geometry analyzing component that separates the relevant information from the noise by analyzing the shape of the defects.

Due to the possible and necessary applications of non-destructive testing, as shown in this section, and the importance of noise reduction of the data, we have developed a two-stage architecture from the field of deep learning. In section 2 we first describe the general functionality of ultrasonic sensors and the resulting data. Following this, in section 3 we give an insight into different approaches to denoising that are used in different areas. After that we present in section 4 our developed architecture, that reduces noise of an air-coupled ultrasonic scan. Since this is a two-stage architecture, both parts are presented separately. In section

4.1, we first consider the *Ultrasonic Responds Analyzer*, which is designed to process air-coupled ultrasonic data. The resulting data is further analyzed in the *Shape of Interest Analyzer*, which is presented in section

4.2. Afterwards we evaluate the architecture in section 5 and describe the results numerically and visually. Finally, we discuss the results in section 6 and identify further challenges.

2. AIR-COUPLED ULTRASONIC

Ultrasound is an elastic wave which is above the acoustic threshold of 16 kHz. If such a wave hits a boundary surface of two media with different acoustic impedances, a reflection and transmission occurs. This means that during the transition between the media, depending on their material properties, part of the wave is refracted and travels through the material, while the other part is reflected. Various oils, water or, as in our specific context, air can be used as a transition medium. The basic idea now is to measure the time it takes for the ultrasonic to travel from the transmitter source to the receiver. If there are defects in the material or material fatigue occurs, the mechanical properties also deteriorate, so that the wave velocity changes. Sensor and receiver can either be located on opposite sides of the object to be inspected or on the same side. In the second case, the sensors are angled so that triggering an ultrasonic causes the surface to oscillate. Analogue the time is measured here, which requires the wave to reach the receiver. Since a wave propagates over time, the transit time is measured within a time span [7-9]. The resulting time series is called

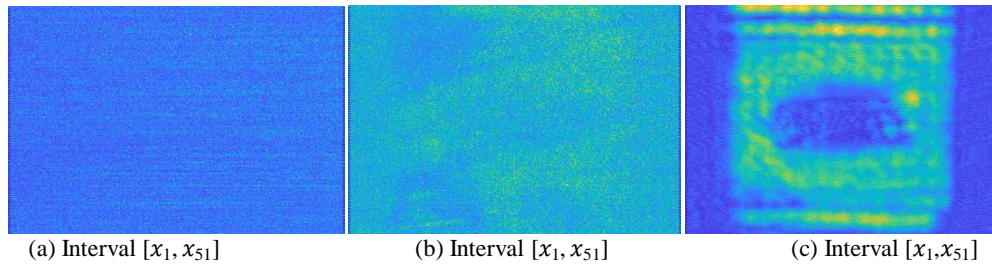


Figure 1. The images (a) – (c) demonstrate the *CScan* visualization of the *Ultrasonic Responds Data* with respect to different time intervals. In figure (c), the impact damage in the centre of the image is shown significantly.

AScan. For a visualization an *AScan* is used for each pixel, which was recorded by the ultrasonic sensor. In this case the maximum and minimum amplitude is determined for each pixel within a fixed time interval. The pixel value, which can be transformed into a color space, is specified by the difference of the discrete fourier transformed extrema of the regarded *AScan*. Ultrasonic images constructed in this way are called *CScans*, represent in our special case an impact damage to a plate made of a special synthetic composite and vary depending on the selected time interval, as shown in the figure 1. Based on these images, classification using a *Convolutional Neural Network* [10] or detection using a *Mask-R-CNN* [11] can be performed.

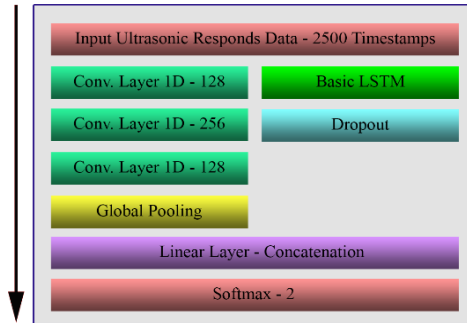
3. RELATED WORKS

With the problem of noise cancelling and reduction have been prepared several publications dealing with different sensor data with different approaches. These can be classical filters or mathematical models that are applied directly to the recorded sensor data. Such data can range from images and video sequences to time series for speech recognition.

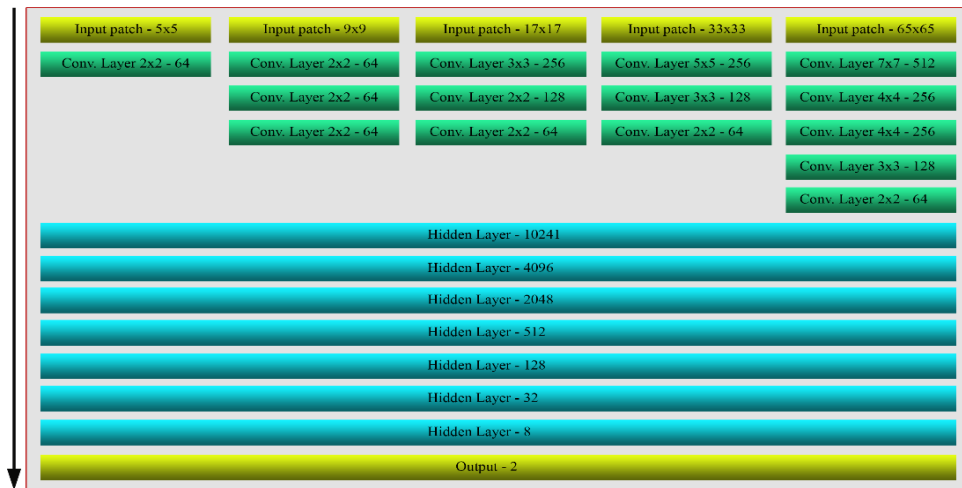
One method of noise reduction from a video sequence is to use additional information from the image sensor [12]. From this information, two different types of noise (*Fixed Pattern Noise* and *Temporal Random Noise*) can be estimated. For the analysis of satellite-derivate time series a double logistic and asymmetric *Gaussian function-fitting* method can be used besides the *Savitzky–Golay filter*, *Mean-value iteration filter* and *ARMD3-ARMA5 filter* [13 - 17]. For the removal of unwanted components of biological origin from neurophysiological recordings, a spatial filter was designed, which uses the separation method *Denoising Source Separation* [18]. For example, to process and remove noise from sensor array data of an *electronic nose system* a *Singular Value Decomposition* can be performed [19]. In this case, the resulting matrix is truncated so that only a few larger singular values remain, from which the noise-free data is finally reconstructed. Another application is the noise reduction of vibration sensor data from a mechanical system. This data is used for machinery fault diagnosis. Due to high noise levels the identification of the fault signature is difficult. For this purpose a method for noise reduction using the time-frequency-multiple concept was developed [20].

Besides the previously mentioned classical techniques, methods in the field of deep learning and machine learning are also used. Neural networks are used in image processing in addition to other methods and achieve high accuracy in classification and detection tasks [21 - 23]. These architectures are used, for example, in medicine for the classification of medical images or the general recognition of objects such as license plates [24, 25]. By using neural networks segmentation or noise reduction can be performed as in the case of our developed approach. developed that denoises the input features [6]. To improve the speech intelligibility for Cochlear

Implant Recipients a two-tier architecture is used, consisting of a noise classifier and a deep denoising autoencoder [26]. Another approach for denoising images is for example the use of Deep CNN [27]. To reduce the noise in a low-dose CT, a method with a Generative Adversial Network was developed [28]. A Convolutional Neural Network was trained with an adversial



(a) Architecture of the *US-Analyzer*



(b) Architecture of the *Shape Analyzer*

Figure 2. The images (a) and (b) show the components of the analyzer parts. The *US-Analyzer* in figure (a) evaluates the ultrasonic data in a *LSTM* and in a concatenation of three 1D-convolutional layers. The result is obtained by concatenation within a linear layer with a following softmax. The *Shape Analyzer* in figure (b) evaluates the patches to be analysed by using a number of convolutional layers, depending on the number and size of the patches. The results of the analysis are then concatenated and the final outcome is determined by using additional linear layers and a softmax.

CNN to estimate routine dose CT. The task of the generator was to transform the low-dose CT images into the routine-dose CT images. Despite the different types of data and the different approaches to noise reduction, whether classic filtering or machine learning, all the methods listed here have one thing in common. They all work directly on the values received from the sensors. In the first step, our architecture uses *AScans*, which we also receive directly from ultrasonic sensors. The difference to the other methods, however, is that we include another additional source of information in the noise reduction by visualizing the data. Based on the visualization, we do not only rely on the evaluation of the sensor data, but also use a geometric analysis, which makes it possible to include information about the shapes of occurring damages in the denoising process.

4. DEEP ULTRASONIC RESPONDS DENOISING (DUSRD)

Our *Deep Ultrasonic Response Denoising* architecture (DUSRD) is designed to process additional geometric information in combination with the sensor data of air-coupled ultrasonic. The idea of our approach is to use the geometric information of the impact damage to separate damage from noise. However, we focus on the information per pixel and make decisions pixel by pixel as well.

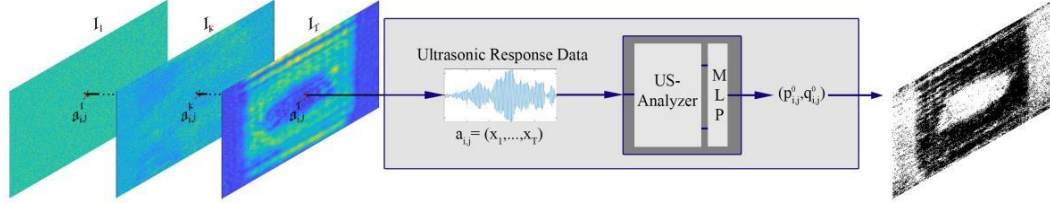


Figure 3. The *Ultrasonic Responds Analyzer* processes the pixel information of an amplitude series. Using the architecture in Figure 2(a), the *Ultrasonic Responds Data* is classified pixel by pixel with respect to the *Impact* and *Undamaged* classes. The result is a first rough noise reduction that acts without considering the pixel environment.

When a *CScan* is constructed, only two values from the selected interval are considered. The remaining values, however, are lost completely. For this reason, we have developed a two-stage architecture to use more information from the *Ultrasonic Response Data* per pixel, so that the result can be nearly noise-free. To do this, we use an *Ultrasonic Responds Analyzer*, which is based on the *Ultrasonic Response Data* per pixel, and a *Shape of Interest Analyzer*, which takes additional information from surrounding pixels. The exact structure regarding the layers used is shown in figure 2.

4.1. Ultrasonic Responds Analyzer (URA)

The *Ultrasonic Responds Analyzer* operates on the data coming directly from the ultrasonic sensor as shown in figure 3. In this process, the sensor records a series of amplitudes $\mathbf{a}_{i,j}^k \in \mathbb{R}^T$ with $i \in \{1, \dots, m\}$, $j \in \{1, \dots, n\}$, $k \in \{1, \dots, T\}$ and $m, n \in \mathbb{N}$ over a predefined time $T \in \mathbb{N}$ for each pixel, corresponding to the reflection of the ultrasonic signal at the measured position. Based on this, this part of the architecture was developed to decide if the object is damaged at a position (respectively at a pixel). In principle, we can distinguish between three classes for a pixel, because in addition to the damage, noise also occurs at the edge of the object. So we consider the classes *undamaged*, *damaged* and *noise*. After some experiments of the *Ultrasonic Response Data* we decided to distinguish only the classes *damaged* and *undamaged*, because the noise has an almost identical signal behaviour as the impact damage. Filtering out the noise is performed with our *Shape of Interest Analyzer* in section 4.2. Since the *Ultrasonic Respond Data* is a time series whose temporal component provides further information to be included in the decision making process, we use a combination of a *Long Short-Term Memory* (LSTM), a special case of a *recurrent network*, and a sequence of one-dimensional *convolutions* [5], as illustrated in figure 2(a). The idea is that we consider the temporal component as a spatial dimension, so that we can use convolutional operations to extract local features like from images. Therefore we define the function

$$\begin{aligned}
 FC_{1D}: \mathbb{R}^T &\rightarrow \mathbb{R}^d \\
 \mathbf{a}_{i,j} &\mapsto \mathbf{x}_f
 \end{aligned} \tag{1}$$

which gets the *Ultrasonic Response Data* $\mathbf{a}_{i,j}$ as input and internally uses a series of convolutional layers. The result is a d -dimensional vector $\mathbf{x}_f \in \mathbb{R}^d$. In parallel we evaluate the vector $\mathbf{a}_{i,j}$ in a

network architecture designed for the analysis of time series. For time-dependent processes, *recurrent networks* and *LSTMs* can be used [29], whereby we use the second in our architecture.

We define the function

$$\begin{aligned} LSTM: \mathbb{R}^T &\rightarrow \mathbb{R}^d \\ \mathbf{a}_{i,j} &\mapsto \mathbf{x}_{lstm} \end{aligned} \quad (2)$$

that calculates a LSTM vector \mathbf{x}_{lstm} after the input of the Ultrasonic Reponse Data $\mathbf{a}_{i,j}$. So we can formalize the result of our Ultrasonic Responds Analyzer using the function

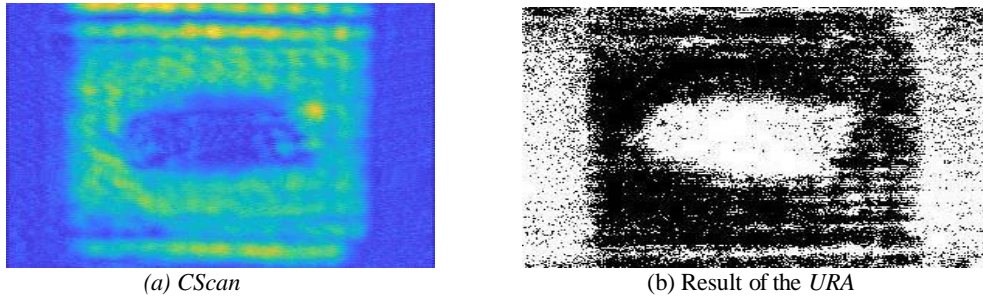


Figure 4. The Image shows the reconstructed noise image after applying the *Ultrasonic Response Data* for each pixel on the *Ultrasonic Responds Analyzer*.

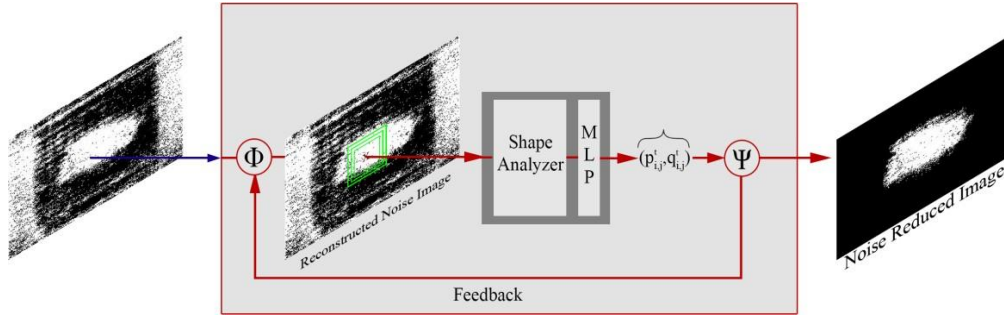


Figure 5. The *Shape of Interest Analyzer* performs a visual analysis of the previous output of the *Ultrasonic Responds Analyzer*. Using the architecture in Figure 2(b), the patches around the viewed pixel are now included in the analysis, changing the classification decision from the previous step if necessary.

This process can be repeated several times using the feedback mechanism. The result is the noise-free ultrasonic image.

$$MLP_{URA}: \mathbb{R}_1^d \times \mathbb{R}_2^d \rightarrow [0,1]^2 \quad (3)$$

in the form $MLP_{URA}(LSTM(\mathbf{a}_{i,j}), FC_{1D}(\mathbf{a}_{i,j}))$. The function 3 concatenates the two calculated parts ($LSTM$ and FC_{1D}) and then determines the probabilities $p_{i,j}^0 \in [0,1]$ and $q_{i,j}^0 \in [0,1]$ to belong to one of the two classes *damaged* or *undamaged*, where $p_{i,j}^0$ is the probability of being a pixel in a damaged area.

4.2. Shape of Interest Analyzer (SoI)

The proposed architecture from subsection 4.1 provides probabilities $p_{i,j}^0$ and $q_{i,j}^0$ of being a part of the impact damage or not for each pixel position $i \in \{1, \dots, n\}$, $j \in \{1, \dots, m\}$. But as indicated in the figure 4(b) in the reconstructed image with the new value $\mathbf{v}_{i,j}^0 = p_{i,j}^0 \cdot \mathbf{c}_2 + q_{i,j}^0 \cdot \mathbf{c}_1$ for

pixel at the position (i, j) , where $\mathbf{c}_1, \mathbf{c}_2 \in [0,1]^3$ are color vectors from the *RGB-Model*, missclassification occur due to noise. In particular, the noise recorded by the ultrasonic sensor at the edge of the object is usually classified as an impact damage, which is represented by the white border bars in figure 4(b). Up to this point, we have considered the pixels independently of each other and excluded their neighbourhoods. But in this case we can not only assume that pixels with an *Ultrasonic Response Data*, that lie within the impact damage, are close together, but also that a certain type of defect, like the impact damage, produces a finite number of shapes visually in the image. The current used dataset consists of impact damages, which basically have an elliptical shape, as shown in figure 4. Based on this property, we want to correct the misclassifications of

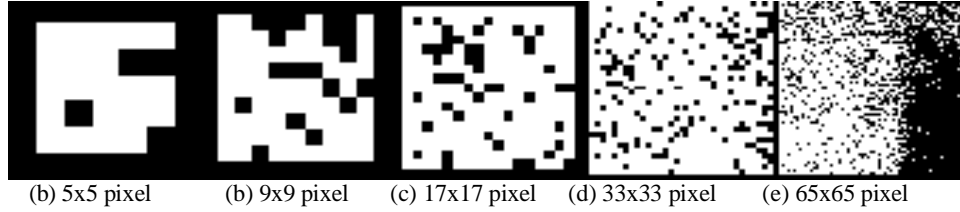


Figure 6. The figures (a) - (e) represent different patches with different sizes at the same position (i, j) of the result of the *Ultrasonic Responds Analyzer*.

single pixels by building an architecture, which is shown in figure 5, that learns the shape of the damage. For this purpose we apply each *Ultrasonic Respond Data* to the *Ultrasonic Responds Analyzer* from subsection 4.1 and arrange the results as an image. Then, for each resulting pixel, we extract a fixed number $l \in \mathbb{N}$ of different patches $\mathbf{A}^k \in [0,1]^{w_k \times h_k}$ with $k \in \{1, \dots, l\}$ and sizes $w_k, h_k \in \mathbb{N}$ around it, as shown in the figure 6. In particular, the selected patches are reduced in area, so that the condition

$$\frac{nm}{2} > r_w^1 \cdot w_1 \cdot r_h^1 \cdot h_1 > r_w^2 \cdot w_2 \cdot r_h^2 \cdot h_2 > \dots > r_w^l \cdot w_l \cdot r_h^l \cdot h_l > 1 \quad (4)$$

with $r_w^k, r_h^k \in [1.5, 2]$ applies. The condition 4 is to guarantee that the patch size does not equal the dimension of the image and that the patches contain sections that differ significantly from each other. If the size would correspond to the given image, features like the position and size of the impact damage could be trained. However, since the damage may only be partially visible during a sensor scan, it is important to filter out local rather than global information. The dimension of the largest patch becomes dependent on the largest extent of the biggest impact damage, so that not all patches, used for a pixel to be analyzed, lie completely within the impact damage. Each patch $\mathbf{A}_{i,j}^k$ is applied to a function

$$FC_{2D}^k : [0,1]^{w_k \times h_k} \rightarrow \mathbb{R}^d \quad (5)$$

$$\mathbf{A}_{i,j}^k \mapsto \mathbf{x}_{i,j}^k$$

with $d \in \mathbb{N}_{\geq 1}$ corresponding to their area, which represents a series of convolutional layers that are used to analyze the individual patches. Based on this, for the classification of the pixel at position (i, j) the results of the function 5 on the corresponding patches are used as input for the

$$MLP: \mathbb{R}^{d_1 + \dots + d_l} \rightarrow [0,1]^2 \quad (6)$$

$$(\mathbf{x}_{i,j}^1, \dots, \mathbf{x}_{i,j}^l) \mapsto (p_{i,j}^t, q_{i,j}^t)$$

The output is the probability p^t that the analyzed pixel is within a damage area. This process is performed for each position, so that each pixel can be reclassified. After reconstructing the image

with the new pixel values, individual miscassifications can still occur. For this reason, the previous steps can be repeated under the condition

$$\psi: [0,1]^{m \times n} \times [0,1]^{m \times n} \rightarrow \{0,1\} \quad (7)$$

$$(P_{t-1}, P_t) \mapsto \begin{cases} 1, & \frac{\sum_{i=1}^n |\delta(p_{i,j}^{t-1}) - \delta(p_{i,j}^t)|}{mn} > \varepsilon \\ 0, & \text{else} \end{cases}$$

on the basis of the new values. The matrices $P_{t-1}, P_t \in [0,1]^{m \times n}$ contain in their entries at position (i, j) the corresponding pixel values for the iteration t and $t - 1$. In this case P_0 is the initial matrix and consists of the results of the *Ultrasonic Responds Analyzer*. The function developed in the condition 7 measures the degree of changes compared to the previous iteration.

The more values $p_{i,j}^{t-1} \in P_{t-1}$ and $p_{i,j}^t \in P_t$ differ the closer the function value gets to one. If it is higher than a threshold $\varepsilon \in (0,1)$, the analysis procedure is repeated on the basis of P_t and the



(a) Used sensor and the test setup

(b) The plates used are made of synthetic material

Figure 7. Figure (a) shows the experimental setup used, which is provided by IKT-Stuttgart (Institute in the field of plastics engineering University of Stuttgart) for the experiments. The plates shown in figure (b) are also prepared by IKT-Stuttgart and provided with the corresponding damage.

next iteration $t + 1$ begins. Otherwise, P_t is the output of the *Shape of Interest Analyzer*. The used function

$$\delta: [0,1] \rightarrow \{0,1\} \quad (8)$$

$$x \mapsto \begin{cases} 1, & x \geq 0.5 \\ 0, & \text{else} \end{cases}$$

maps $p_{i,j}^t$ to zero, which represents a pixel on undamaged plate position, or one, which indicates a pixel on a damaged plate position, depending on whether the pixel value is above or below 0.5. In a further iteration, the function

$$\phi: [0,1]^2 \times [0,1]^2 \rightarrow \{0,1\} \quad (9)$$

$$(p_{i,j}^{t-1}, q_{i,j}^{t-1}, p_{i,j}^t, q_{i,j}^t) \mapsto \{0,1\}$$

can be used to decide whether new pixel values are applied or kept from the previous iteration $t - 1$. This reduces the number of pixels to be evaluated per repetition and prevents permanently changing values.

5. RESULTS

In this section we show the quality of our developed architecture by presenting our evaluation results. For this purpose, we describe in section 5.1 and 5.2 the test setup, the system we use for training and evaluation and the recorded data. Finally, in section 5.3 we consider and interpret the obtained results.

5.1. Test Setup

For our experiments and the development of our architecture we have chosen an air-coupled ultrasonic system as shown in figure 7(a). In this case, the system consists of two sensors mounted side by side at a given angle. The transmitter uses ultrasonic to oscillate the surface, which is stronger or weaker depending on the surface structure. The resulting change in reflection is perceived by the receiver, so that changes in the structure due to impact damage, have an effect on the recorded data. To prevent the receiver from already receiving the ultrasonic signal from the transmitter, a barrier is inserted between the two sensor units. Since the vibration propagates

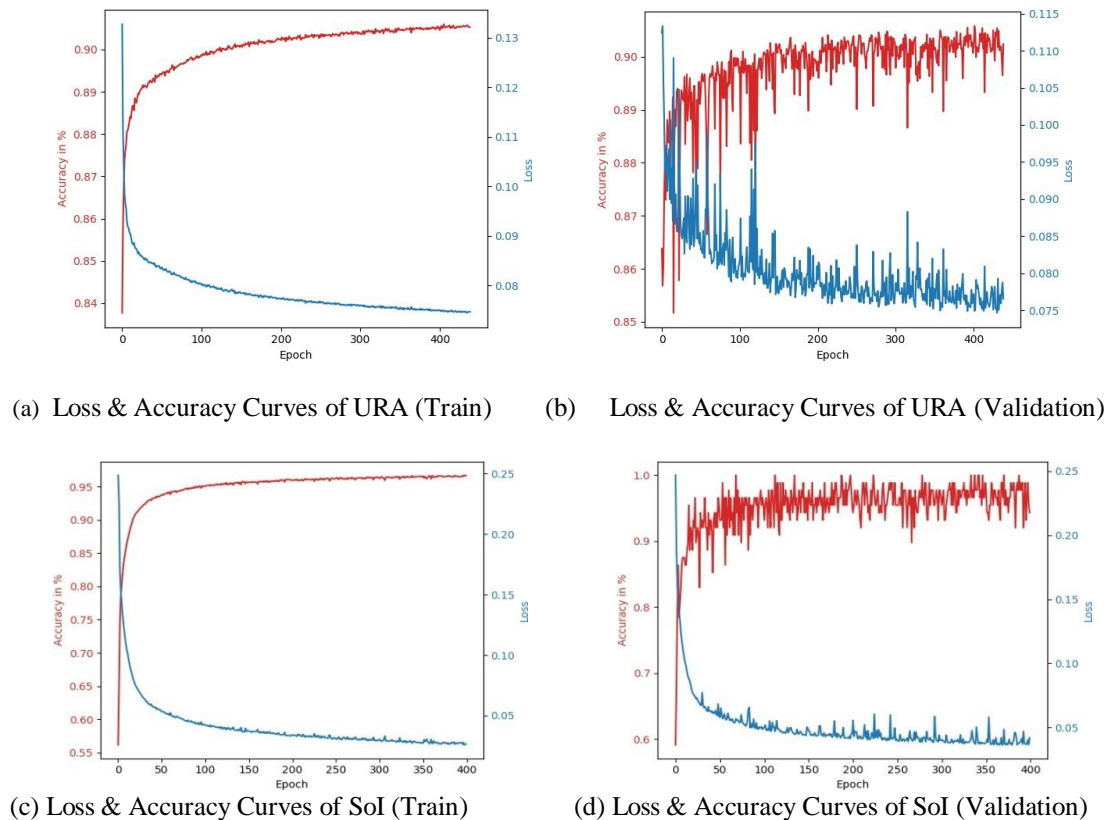


Figure 8. The diagrams (a) – (b) show the development of the loss value and the accuracy of the training and validation data for the *Ultrasonic Responds Analyzer* and the *Shape of Interest Analyzer*.

in the surface over a time t , the received reflection is measured for each position x and y , resulting in an amplitude response diagram. Our test objects are plates made of special synthetic material, which are shown in the photograph 7(b), with a width of 10 cm and a length of 20 cm. Each plate contains an impact damage in the center, which was caused by impacts of different intensities. The ultrasonic images were taken for 61.103 positions (301 along the x -axis and 203 along the y -

axis) with 2.500 time stamp each. The angle between the sensor units is selected individually for each plate according to the maximum signal-to-noise ratio. For a larger dataset, additional variations in the angle were allowed, which did not provide the maximum of the signal-to-noise ratio.

5.2. Data & Hardware

For the evaluation task, the *Ultrasonic Responds Analyzer* and the *Shape of Interest Analyzer* are evaluated separately. This is possible because both analyzers work independently in their processing. Both components of the architecture were trained, validated and evaluated on a system with an Nvidia RTX2080TI with 11GB memory, an Intel Core i9-9900K with 3.6 Ghz and 32GB of Ram. A total of 369.730 data points, where a data point is an *Ultrasonic Response Data*, are available for the *Ultrasonic Responds Analyzer* and 109.370 data points for the *Shape of Interest Analyzer*, split into 60% training, 20% validation and 20% evaluation data. Accordingly, for the first part 221.838 data points are available for the training processes and 73.946 data points for the validation and evaluation process. For the second part the split is 65.622 data points for training and 21.874 data points for validation and evaluation. The data is balanced so that each class appears nearly equal. For the *Ultrasonic Responds Analyzer*, *Ultrasonic Response Data* are used that were recorded with a total of 2.500 time steps. The *Shape of Interest Analyzer* is

	Label: Impact	Label: Undamaged		Label: Impact	Label: Undamaged
Classified: Impact	32546	4427	Classified: Impact	9700	1237
Classified: Undamaged	2604	34369	Classified: Undamaged	179	10758
(a)	Confusion matrice (URA)		(b)	Confusion matrice (SoI)	

Table 1. Confusion matrices with respect to the *Ultrasonic Responds Analyzer* and the *Shape of Interest Analyzer*. They show the classification results based on the evaluation data in the form of absolute numbers and indicate which class causes most misclassifications.

evaluated based on previously calculated reconstructed noise images and the corresponding patches. Important in this context is that for the training, validation and evaluation processes, data points of different plates are used which are not 100% homogeneous, since, amongst other things, minimal differences in the thickness of the plates can occur.

5.3. Evaluation

Our architecture, which is focused on improving air-coupled ultrasonic recordings by using deep learning for noise cancelling, performs its application purpose, as the following evaluation results confirm. Consider first the training process of the two parts of the architecture regarding the evaluation of training and validation data, which is shown in the figure 8. The two diagrams show the development of the accuracy of the classification and the error with respect to the loss-function on the different data sets. A convergence can be seen in both architectures. Accuracy reaches values above 90% while the loss value falls below 0.05 and 0.08 respectively. However, as shown by the diagrams 8(b) and 8(d), there are strong variances in the validation data in both training processes, which become smaller as the number of epochs increases. This behaviour results from the ambiguous separation between the damaged and undamaged areas. The *Ultrasonic Responds Data* from these transitions was also annotated for the training and validation process. As a result, these data points in the training data set are learned by the architecture, while they cause increased

variance in the validation data set. Above all, the variances are more significant in the *Ultrasonic Responds Analyzer* than in the *Shape of Interest Analyzer*. However, since the two architectures expect different types of data as input, this explains the different intensity of variance in the validation data sets. For a deeper analysis, we look at the corresponding confusion matrices, which can be found in the table 1, based on the evaluation data. The values in the matrices correspond to the numbers for the correct classification and misclassification with regard to the individual classes. It can already be seen that the absolute number of misclassifications for the impact class is significantly lower than for the undamaged class in both parts of the architecture. For a better representation we consider the relative values including the total accuracy in the diagrams in figure 9. assumption that the undamaged class is more difficult to classify and therefore more problematic. Even though the total accuracy is above 90%, these results benefit from balancing the data. The Impact class is detected about 4% better than the Undamaged class in the *Ultrasonic Responds Analyzer* and almost 10% better in the *Shape of Interest Analyzer*, resulting in the high total accuracy. However, the distribution of the classes on a given synthetic plate with an impact damage is not uniform, but usually consists of a larger proportion of the undamaged class. This means that the total accuracy evaluated on such test objects may vary depending on the class distribution. The image in figure 10(b) show the results of the *Ultrasonic Responds Analyzer* applied to the air-coupled ultrasonic scans of two synthetic plates with impact damage. The dark dots represent the pixels classified as impact, while the bright dots represent the undamaged pixels. Color representation of pixels that cannot be clearly classified by the architecture is not available, because the results of *softmax* are always 0 for undamaged class or 1 for damaged class.

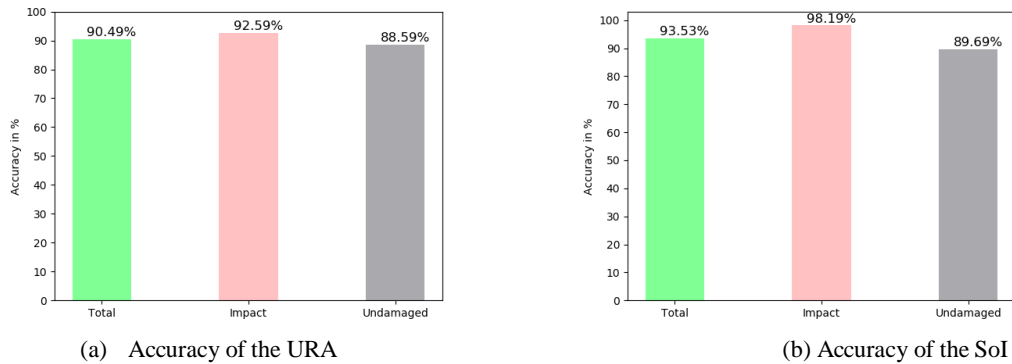


Figure 9. The diagrams (a) and (b) show the total accuracy as well as the accuracy of the individual classes for the *Ultrasonic Responds Analyzer* and the *Shape of Interest Analyzer*.

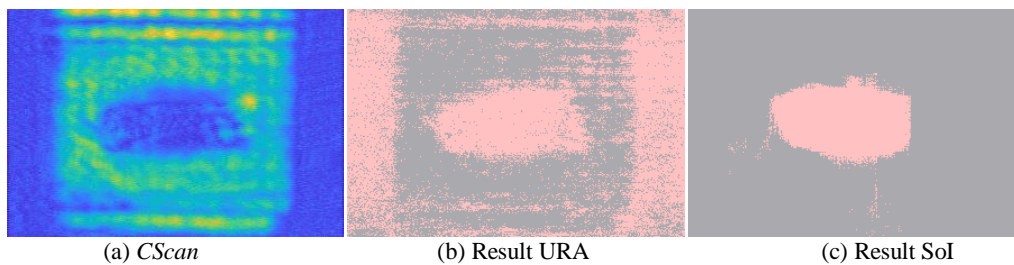


Figure 10. The resulting image (c) after applying figure (a) to our architecture from figure 2(a) and 2(b) after the intermediate result from figure (b).

Compared to the *CScan* in figure 10(a), the *Ultrasonic Responds Analyzer* initially delivers a slightly noise-reduced image with respect to the undamaged plate. The edge noise is classified as impact damage as described in section 4. Visually, the resulting images correspond to the accuracies from diagram 9. In order to achieve the goal of the representation as free of noise as

possible, we apply the *Shape of Interest Analyzer* to the image 10(b). As a result we obtain the classification visualized in the images 10(c). It can be clearly seen that the background and edge noise have been almost completely eliminated, which is initially inconsistent with the accuracy analysis. However, if we compare figures 10(b) and 10(c), we can see that the misclassifications are primarily located around the elliptical impact damage, apart from a few outliers. This behaviour corresponds in particular to our expectations, as the limits of the impact damage cannot be clearly defined, so that misclassifications can occur at precisely these points. Our results show that the combination of a time series analysis with the analysis of geometric shapes using deep learning can reduce noise to a minimal area. In particular, the evaluations show that the addition of information on the geometric structure of errors leads to a significant improvement in noise cancelling. This means that this approach can be adapted to problems and recordings whose error objects can be reduced to a finite number of geometric shapes. In addition, some of the misclassifications can be simply not taken into account because they lie in an area that cannot be accurately classified, which means that the resulting images have a better accuracy than the exact evaluation results reflect. Here it is important that this statement is only relevant if the accuracy of the classification of an error pixel is as high as possible, because only under this condition it can be guaranteed that the existing error can be reconstructed correctly. Based on this, visibly good results can be achieved.

6. DISCUSSION & CONCLUSIONS

The results of our evaluation show that neural networks can be used to denoise noisy data. In the following sections we describe the benefits of our approach for industrial applications and discuss future work.

6.1. Contribution

With our two-stage network architecture we have increased the practical relevance of air-coupled ultrasonic by using deep learning to significantly improve the quality of the sensor data. Our evaluation results have shown that we have achieved our goal of making the data noise-free, so we have created the basis for further analysis steps. Our denoising approach, which builds on the strengths of neural networks by additionally exploiting visual information, enables a more interference-free evaluation of the data. In particular, the denoising developed here is not limited to the use of air-coupled ultrasonic data, but can be applied to all data based on geometric shapes. Thus, especially in our context, we create the potential to apply air-coupled ultrasonic in industrial applications in safety-critical areas such as aircraft.

6.2. Conclusion

With our pixel-wise classification and accuracies of over 90%, we can provide better defect classification and detection of *Ultrasonic Responds Data*. The results of our evaluation show that by adding geometric information, noise was almost completely removed. In particular, it can be concluded from the results that the majority of the misclassifications are located at the edge of the damage, so that they are not significant due to the ambiguous delimitation of the damage. These conclusions make it possible to use the architecture for other types of damage with different characteristics. In particular, the occurrence of overlapping material damage in the ultrasonic data can be recorded and detected by the geometric component, so that quality assurance can be performed.

6.3. Future Work

With our architecture we have set the basis for further development. Apart from optimizations in the implementation by parallelization of pixel-wise classification, which leads to a significant reduction of the time required, our goal is to extend the approach presented here to a processing pipeline for air-coupled ultrasonic data. The existing architecture is to be improved with respect to a more generalized detection of object edges, so that the noise that occurs there can be removed more easily. In addition, the pipeline will be used to locate and classify defects based on the noise-free data. Such a processing pipeline can be used in many areas of industry, including as part of a semi-automatic defect classification software.

7. ACKNOWLEDGEMENT

This work was encouraged and supported by the projects *Ki-LiveS* within the BMBF funding announcement and *ADJutaNT* within a ZIM cooperation project.

REFERENCES

- [1] Wang, X. Y., Li, C. C., Wu, H. D., Zhang, D., Zhang, X. D., & Gong, X., (2019) "NDE Application of Air-Coupled Transducer for Surface Crack Detection", In *2019 Symposium on Piezoelectricity, Acoustic Waves and Device Applications (SPAWDA)*, IEEE, pp1-4.
- [2] Findeis, D., & Gryzagoridis, J. (2014) "Digital shearography and vibration excitation for NDT of aircraft components", In *AIP Conference Proceedings*, American Institute of Physics, Vol. 1600, No. 1, pp33-38.
- [3] Hægström, E., & Luukkala, M. (2001) "Ultrasound detection and identification of foreign bodies in food products", *Food Control*, Vol. 12, No. 1, pp37-45.
- [4] Gholizadeh, S. (2016) "A review of non-destructive testing methods of composite materials", *Procedia Structural Integrity*, Vol. 1, pp50-57.
- [5] Karim, F., Majumdar, S., Darabi, H., & Chen, S. (2017) "LSTM fully convolutional networks for time series classification", *IEEE access*, Vol. 6, pp1662-1669.
- [6] Maas, A., Le, Q. V., O'neil, T. M., Vinyals, O., Nguyen, P., & Ng, A. Y. (2012) "Recurrent neural networks for noise reduction in robust ASR".
- [7] Gräfe, B. (2009) "Luftgekoppeltes Ultraschallecho-Verfahren für Betonbauteile".
- [8] Unger, A. (2019) "Air Coupled Ultrasonic Transducers for Industrial Applications" (Doctoral dissertation, Technische Universität).
- [9] Döring, D. (2011) "Luftgekoppelter Ultraschall und geführte Wellen für die Anwendung in der zerstörungsfreien Werkstoffprüfung".
- [10] Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Chen, T and et al., (2018) "Recent advances in convolutional neural networks", *Pattern Recognition*, Vol. 77, pp354-377.
- [11] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017) "Mask r-cnn", In *Proceedings of the IEEE international conference on computer vision*, pp2961-2969.
- [12] Bosco, A., Findlater, K., Battiato, S., & Castorina, A. (2003) "A noise reduction filter for full-frame data imaging devices", *IEEE Transactions on Consumer Electronics*, Vol. 49, No. 4, pp676-682.
- [13] Hird, J. N., & McDermid, G. J. (2009) "Noise reduction of NDVI time series: An empirical comparison of selected techniques", *Remote Sensing of Environment*, Vol.113, No. 1, pp248- 258.
- [14] Jonsson, P., & Eklundh, L. (2002) "Seasonality extraction by function fitting to time-series of satellite sensor data", *IEEE transactions on Geoscience and Remote Sensing*, Vo. 40, No. 8, pp1824-1832.
- [15] Chen, J., Jönsson, P., Tamura, M., Gu, Z., Matsushita, B., & Eklundh, L. (2004) "A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter", *Remote sensing of Environment*, Vol. 91, No. 3-4, pp332-344.
- [16] Ma, M., & Veroustraete, F. (2006) "Reconstructing pathfinder AVHRR land NDVI time-series data for the Northwest of China", *Advances in Space Research*, Vol. 37, No. 4, pp835-840.

- [17] Filipova-Racheva, D., & Hall-Beyer, M. (2000) "Smoothing of NDVI time series curves for monitoring of vegetation changes in time". In *Ecological monitoring and assessment network national science meeting*, pp17-22.
- [18] de Cheveigné, A., & Simon, J. Z. (2008), "Denoising based on spatial filtering", *Journal of neuroscience methods*, Vol 171, No. 2, pp331-339.
- [19] Jha, S. K., & Yadava, R. D. S. (2010) "Denoising by singular value decomposition and its application to electronic nose data processing", *IEEE Sensors Journal*, Vol. 11, No. 1, pp35-44.
- [20] He, Q., Wang, X., & Zhou, Q. (2014) "Vibration sensor data denoising using a time-frequency manifold for machinery fault diagnosis", *Sensors*, Vol. 14, No. 1, pp382-402.
- [21] Massaro, A., Dipierro, G., Cannella, E., & Galiano, A. M. (2020) "Comparative Analysis among Discrete Fourier Transform, K-Means and Artificial Neural Networks Image Processing Techniques Oriented on Quality Control of Assembled Tires.", *Information*, Vol. 11, No. 5, pp257.
- [22] Giri, S., & Joshi, B. (2019) "Transfer Learning Based Image Visualization Using CNN", *International Journal of Artificial Intelligence and Application*, Vol. 10, No. 4, pp47-55.
- [23] Egmont-Petersen, M., de Ridder, D., & Handels, H. (2002) "Image processing with neural networks—a review", *Pattern recognition*, Vol. 35, No. 10, pp2279-2301.
- [24] Shi, Z., He, L., Suzuki, K., Nakamura, T., & Itoh, H. (2009) "Survey on neural networks used for medical image processing", *International journal of computational science*, Vol. 3, No. 1, pp86.s
- [25] Koval, V., Turchenko, V., Kochan, V., Sachenko, A., & Markowsky, G. (2003) "Smart license plate recognition system based on image processing using neural network", In *Second IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, pp123-127.
- [26] Lai, Y. H., Tsao, Y., Lu, X., Chen, F., Su, Y. T., Chen, K. C., Lee, C. H. and et all (2018) "Deep learning-based noise reduction approach to improve speech intelligibility for cochlear implant recipients", *Ear and hearing*, Vol. 39, No. 4, pp795-809.
- [27] Zhang, K., Zuo, W., Chen, Y., Meng, D., & Zhang, L. (2017) "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising", *IEEE Transactions on Image Processing*, Vol. 26, No. 7, pp3142-3155.
- [28] Wolterink, J. M., Leiner, T., Viergever, M. A., & Išgum, I. (2017) "Generative adversarial networks for noise reduction in low-dose CT", *IEEE transactions on medical imaging*, Vol. 36, No. 12, pp2536-2545.
- [29] Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., & Schmidhuber, J. (2016) "LSTM: A search space odyssey", *IEEE transactions on neural networks and learning systems*, Vol. 28, No. 10, pp2222-2232.