Check for updates

METHOD ARTICLE

REVISED Training neural networks on domain randomized simulations for ultrasonic inspection

[version 2; peer review: 2 approved]

Klaus Schlachter, Kastor Felsner, Sebastian Zambal

Machine Vision, Profactor GmbH, Steyr, 4407, Austria

v2 **First published:** 28 Mar 2022, **2**:43
https://doi.org/10.12688/openreseurope.14358.1

Latest published: 11 May 2022, **2**:43
https://doi.org/10.12688/openreseurope.14358.2

Abstract

To overcome the data scarcity problem of machine learning for nondestructive testing, data augmentation is a commonly used strategy. We propose a method to enable training of neural networks exclusively on simulated data. Simulations not only provide a scalable way to generate and access training data, but also make it possible to cover edge cases which rarely appear in the real world. However, simulating data acquired from complex nondestructive testing methods is still a challenging task. Due to necessary simplifications and a limited accuracy of parameter identification, statistical models trained solely on simulated data often generalize poorly to the real world. Some effort has been made in the field to adapt pre-trained classifiers with a small set of real world data. A different approach for bridging the reality gap is domain randomization which was recently very successfully applied in different fields of autonomous robotics. In this study, we apply this approach for ultrasonic testing of carbon-fiber-reinforced plastics. Phased array captures of virtual specimens are simulated by approximating sound propagation via ray tracing. In addition to a variation of the geometric model of the specimen and its defects, we vary simulation parameters. Results indicate that this approach allows a generalization to the real world without applying any domain adaptation. Further, the trained network distinguishes correctly between ghost artifacts and defects. Although this study is tailored towards evaluation of ultrasound phased array captures, the presented approach generalizes to other nondestructive testing methods.

Keywords

CFRP, Domain Randomization, Nondestructive Testing, Machine Learning, Defect Detection, Ultrasonic Rendering, Raytracing

Open Peer Review

Approval Status ✓ ✓

	1	2
version 2 (revision) 11 May 2022	✓ view	✓ view
version 1 28 Mar 2022	? view	

1. **Carmelo Mineo**, University of Palermo, Palermo, Italy

2. **Iikka Virkkunen**, Aalto University, Espoo, Finland

Any reports and responses or comments on the article can be found at the end of the article.



This article is included in the Horizon 2020 gateway.



This article is included in the Deep Learning collection.

Corresponding author: Klaus Schlachter (klaus.schlachter@profactor.at)

Author roles: Schlachter K: Conceptualization, Investigation, Methodology, Software, Visualization, Writing – Original Draft Preparation; **Felsner K**: Conceptualization, Methodology, Software, Writing – Review & Editing; **Zambal S**: Conceptualization, Methodology, Project Administration, Software, Supervision, Visualization, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This research was financially supported by the European Union's Horizon 2020 research and innovation programme under the grant agreement No 831830, project SonicScan and Clean Sky 2 Joint Undertaking (JU). This work was also supported by the county of Upper Austria and EFRE as part of the FTI strategy, project "DigiManu", and the Growth strategy, project "SmartFactoryLab".

Copyright: © 2022 Schlachter K *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Schlachter K, Felsner K and Zambal S. **Training neural networks on domain randomized simulations for ultrasonic inspection [version 2; peer review: 2 approved]** Open Research Europe 2022, **2**:43 https://doi.org/10.12688/openreseurope.14358.2

First published: 28 Mar 2022, **2**:43 https://doi.org/10.12688/openreseurope.14358.1

REVISED **Amendments from Version 1**

The new version of the article contains the following updates:

- The general overview on simulation of sound propagation was extended: Root finding methods for computing sound paths in multi-layered structures were added to the survey in the introduction chapter.

- Clarifications regarding on how the distortion of the volume masks for training was chosen.

- Clarifications regarding the advantage of the used simulation model.

- Rephrasing a part of Chapter 3, in order to clarify the intended meaning.

Any further responses from the reviewers can be found at the end of the article

1 Introduction

The field of applications of carbon-fiber-reinforced plastics (CFRP) expands with the progress in production of parts with complex geometries. However, the error-prone manufacturing process requires a seamless inspection of safety-critical components[1]. For industrial production, this is typically performed by ultrasonic inspection. While the data acquisition process is often automated, the data still needs to be evaluated by trained experts. Such analyses are arduous and time intensive and lead to results dependent on the judgement of individuals. Thus, automated data evaluation has not only the potential of enormous time and cost savings, but also the potential to enhance the consistency of results.

Automated data evaluation in nondestructive testing (NDT) mostly relies on simple methods such as signal thresholding. For more advanced inspection cases, such as ultrasonic testing, inspectors utilize various signal characteristics. This complex decision does not offer a simple algorithmic description[2]. Machine learning (ML) offers thereby an approach for automated evaluation devoid of manually derived algorithms.

A general problem for applying ML in NDT is the scarcity of flawed training data. For calibration and human training, artificially introduced defects, such as flat-bottom holes and side-drilled holes, are widely used. However, ultrasound signals of artificial flaws do not represent real data distributions[3,4]. While human inspectors can use their physical reasoning to judge real world data based on skills gained from training from simplified examples, training data for ML models need to represent the target domain.

Currently used strategies in NDT to overcome the data scarcity problem can mostly be grouped into data augmentation techniques and domain adaptation techniques. Data augmentation uses scan data from flawed samples. Extracting the flaw signals allows for the introduction of virtual flaws to arbitrary locations of the scan[2]. Domain adaptation is a technique to adapt a model trained on a source domain (e.g. data from specimens with simplified artificial flaws like flat-bottom holes) to work on a target domain (e.g. real world data)[5]. In the simplest case, this can be achieved by fine-tuning a pre-trained model with a smaller set of real world data.

Current approaches in NDT for utilizing simulations try to generate data as close to the real world as possible in order to get results which are useful in the real world or which at least allow the model to be adapted to the real world. The main idea of domain randomization is to introduce variations to the simulator itself. Enough variability might cause the model to see the real world just as another variation of the simulator[6]. Recently, domain randomization was applied very successfully in various fields of autonomous robotics (e.g. for image-based pose estimation for robotic manipulation tasks[7], vision-based drone navigation[8], object category detection[9], and simulation-based robotic policy learning for autonomous quadrotor landing[10]).

This article is structured in the following way. In Section 1.1 we review related work on ML applied in NDT. In Section 1.2 we provide an overview about simulating sound propagation in general. The derived phonon tracing algorithm for simulating total-focusing method (TFM) data is outlined in Section 2.1. The architecture and training of a neural network classifier is described in Section 2.2. An evaluation of the proposed approach and a discussion of results is provided in Section 3. Finally, an outlook is provided in Section 4.

1.1 Related work

In the recent years there has been considerable interest in developing machine learning models to evaluate ultrasonic data. Examples are monitoring of mixing processes[11], assessing carburization of industrial steel tubes[12], and testing of concrete foundation piles[13]. We are particularly interested in applications regarding robot-guided ultrasonic testing of complex compact parts. ML could thereby enable (i) automatic defect detection, and (ii) adaptive path planning to guide the robot to areas of interest which require more detailed scanning[5]. A major challenge towards the use of machine learning models for ultrasonic testing is the lack of representative training data. Data with real flaws tend to be scarce and producing mock-ups is expensive and time-consuming. Further, data is often confidential and therefore rarely shared, as it might reveal insights into the scanned structures[2].

Data augmentation has been suggested as a strategy to overcome the data scarcity problem in ultrasonic testing in the recent literature. The flaw signal is thereby extracted from a real scan and re-introduced at various positions. In 2 a machine learning model is trained with augmented data for flaw detection in phased array ultrasonic data. In order to increase the variety, various transformations, such as rotation and scaling, are applied on the extracted flaw signals.

Transfer learning is another strategy applied in ultrasonic testing. Models pre-trained with data from another domain are thereby adopted to the target domain via transfer learning. In 14 a MobileNet is pre-trained on the ImageNet dataset (general image dataset) and migrated to feature extraction for a welding

defect classificator. In 15 transfer learning is examined to monitor industrial processes via ultrasonic sensing. Two case studies on mixing and cleaning of fouled pipes are investigated and very promising results could be achieved[16].

Training on simulated data. Generating data via simulations provides many advantages to the methods mentioned above. The amount of data is easily scalable and accurate labels are accessible for free. Simulations also enable the integration of edge cases, which by definition rarely occur in the real world. However, statistical models trained in simulations often work poorly in the real world, and there is only very little literature examining this approach in the field of NDT. Impressive results were presented in 17 for defect detection in three-dimensional CT scans of cast aluminum parts. A deep segmentation model is trained solely on utilizing high-fidelity simulations. The segmentation model is able to deal with the huge variety of image artifacts of CT scans and achieves similar results as qualified experts. Further, it is shown that the model can be migrated to specialized tasks via fine-tuning.

Domain randomization. Generating realistic artificial data acquired from complex NDT methods is still a challenging task. Physical models need to make assumptions of the real world to keep the models manageable. More complex models enable reproduction of a broader variety of occurring physical effects, however, typically the number of parameters increases with the complexity of the model. To mimic the target system realistically, parameter identification can thereby be very challenging for more complex models. Domain randomization utilizes low fidelity data for training by introducing variability to the simulator itself. The main assumption is that with just enough variability in the simulator, the real world appears just as another variation[6]. This approach is very successful in the field of autonomous robotics, however, there seems to be very little attention towards this technique in the NDT community.

Ultrasonic image rendering. Simulation of phased array ultrasonic data is used in various fields. In the NDT field simulations are used to evaluate influence factors and to optimize NDT configurations. To accurately account for heterogeneity and anisotropy, wave-based methods are utilized[13,18,19]. However, wave-based simulations are typically very time-consuming, which makes it difficult to generate large data sets as required for deep learning. To reduce simulation time, in 20 a mixture of finite element models in the defect locality and ray-based models for the whole region of interest is used. In 21 a fast deep convolutional generative adversarial network is used to generate synthetic weld TFM images. For training, synthetic data is generated via finite element simulations. In medicine, real-time rendering is of particular interest to enable medical training independent of availability of patients and to include rare diseases in the training[22]. The simulator developed in this work is based on a geometric approach and mostly relies on efforts made in the medical field.

1.2 Simulation of sound propagation

Simulation of sound propagation is a widely studied field, motivated by a variety of use cases, e.g. auralization for audio post processing, auralization of virtual environments, and simulating acoustic properties of classrooms[23–25]. Lately, realistic ultrasound rendering is also studied to enhance medical training[22,26–28]. Two main techniques have evolved for numerically evaluating sound propagation, with one group relying on a wave-based problem formulation, and the other relying on geometric acoustic approaches. The former approximates the solution of the underlying wave equation of linear acoustics for the pressure field $p(\boldsymbol{x}, t)$

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right) p = 0 \tag{1}$$

where c denotes the velocity at which the pressure disturbance propagates in the medium[29]. The solution can be approximated by classical methods, e.g. by finite element method or boundary element method[23,30]. A review of the theoretical basis of these methods and their application for sound simulations can be found in 31. These methods provide very accurate solutions, but the computational costs are high and increase dramatically with the highest considered frequency. Hence, a wave-based approach is especially applicable for simulating low frequencies[32].

Geometric acoustics approximates sound propagation by sound particles moving along directed rays. In general, paths between two points are determined by an underlying variation problem known as Fermat's principle. In homogenous media rays travel along straight lines and change directions at boundaries of media with different wave velocities. Modern root finding methods allow to compute sound paths between a source and target point in multi-layered structures, such as composite parts, very efficiently[33,34]. For this purpose, the problem is modeled based on Snell's law. This approach can be extended for determining sound paths in media with continuously changing material properties by discretizing the media through regions with constant wave velocities. Such an approach can be applied to various problems in the NDT field, e.g. for efficiently determining ultrasonic paths for in-process inspections of welds at the presence of high temperature gradients. While such problems can also be tackled by finite element models, such an approach is of great advantage in terms of the required simulation time[33,34]. In general, a geometric acoustics approach is suitable for waves with high frequencies, i.e. if the wavelengths are short in relation to the objects the wave is interacting with. Splitting the source signal into a low and a high frequency band allows combining geometric approaches with wave-based approaches in order to account for low frequency effects[35].

1.2.1 The rendering equation. Approximating sound propagation via sound particles (phonons) along directed rays is very similar to the approach of geometric optics, and therefore many algorithms from light rendering can be adapted for sound simulations[23,35]. Light rendering is typically formulated as transport problem, where light sources emit energy which is transported in a three dimensional scene by means of reflection and refractions. The rendering equation, which describes this transport problem, was introduced to computer graphics by Kajiya[36]. Thereby, it is assumed that light travels instantaneously

and therefore steady-state is achieved instantaneously[37]. In contrast, sound rendering has to account for the finite speed at which sound waves propagate in the medium. A thorough derivation of the acoustic rendering equation was presented in 38, which represents a time-dependent adaptation of the rendering equation. The rendering equation states that at a point x, the outgoing time-dependent radiance $l(x \rightarrow \Theta)$ from a point x in direction Θ is the sum of the radiance $l_e\ (x \rightarrow \Theta)$ emitted by the surface itself, and the reflected radiance $l_r\ (x \rightarrow \Theta)$. As illustrated in Figure 1 on the left, in general, the reflected radiance results from partial contributions of incoming radiance l(x $\leftarrow\Psi$) from all directions Ψ. Integrating over the hemisphere Ω_x, the (acoustic) rendering equation is stated as

$$l(x \rightarrow \Theta) = l_e(x \rightarrow \Theta) + \underbrace{\int_{\Omega_x} \rho(\Psi,\Theta)\, l(x \leftarrow \Psi) \cos(N_x, \Psi)\, \mathrm{d}\omega_\Psi}_{l_r(x \rightarrow \Theta)} \quad (2)$$

with the acoustic bidirectional reflectance distribution function (BRDF) ρ describing the acoustic reflection properties of the material. The cosine term accounts for geometrical foreshortening with the surface normal N_x.

1.2.2 The importance equation. A rendering algorithm computes the light or sound energy that is visible at every pixel, while each pixel functions as a sensor. The importance equation - the dual of the rendering equation - was introduced to computer graphics by Pattanaik[39]. The importance equation evaluates the contribution w of light or sound energy from point x in direction Θ to the sensor measurement. If the path (x, Θ) reaches the sensor unhindered, the light or sound energy leaving x contributes fully to the measurement and therefore the so-called self-importance w_e equals 1, otherwise it may reach the sensor through reflections and refractions and contributes partly. Given the BRDF ρ the importance equation is stated as

$$w(x \rightarrow \Theta) = w_e(x \rightarrow \Theta) + \int_{\Omega_y} \rho(\Psi,\Theta)\, w(y \rightarrow \Psi) \cos(N_y, \Psi)\, \mathrm{d}\omega_\Psi. \quad (3)$$

The rendering Equation 2 and importance Equation 3 describe the same transport problem; as both equations have the same mathematical structure, the same solution strategies can be applied. Ray tracing for solving the rendering equation sends rays from the sensor into the scene; backwards ray tracing (further referred to as phonon tracing) for solving the importance equation sends rays from the source into the scene[37].

2 Methods

In this section a phonon tracing algorithm which allows to generate artificial TFM data is derived. Subsequently, the architecture and training of a neural network for defect detection is described. The TFM simulator was implemented as Blender (v2.91) add-on and PyTorch (v1.9) was used for training.

2.1 Simulation of TFM images

TFM images result from post-processing of the captures of all transmitter-receiver pairs of a phased array. To emulate this, we simulate impulse responses by stochastic phonon tracing. The main emitting direction of a sender element is orthogonal to its surface, and the distribution of how the energy is spread around the normal is determined by the spatial dimensions of the sensor element with respect to the wave length. Array elements applied in non-destructive testing applications are typically rectangular. Further, the side along which the elements are arranged is typically significantly shorter than the width of the elements. This leads to a fan-shaped distribution of the emitted energy within the sensor plane.

An emitted phonon which hits a sensor element unhindered, contributes fully to the measurement at the time which the phonon took to travel from the emitter to the sensor. Otherwise, depending on the scene geometry, a phonon contributes to the impulse response through several reflections and refractions, and its contribution is determined by Equation 3. At the first interaction with a scene element at y, the energy is partly reflected and refracted. Additionally, the energy is spread out according to a distribution determined by the surface characteristics of the interface. Practically, these partial contributions are integrated numerically via Monte Carlo integration by sampling discrete directions Ψ_i. Assuming that the self-importance equals zero, Monte Carlo integration of Equation 3, i.e.:

$$\hat{w}(x \rightarrow \Theta) = \frac{1}{N} \sum_{i=1}^{N} \frac{\rho \cos(N_y, \Psi)\, w(y \rightarrow \Psi_i)}{p(\Psi_i)}, \quad (4)$$

calculates the importance estimation $\hat{w}$. Thereby, the discrete directions Ψ_i are sampled from the distribution p. The importance

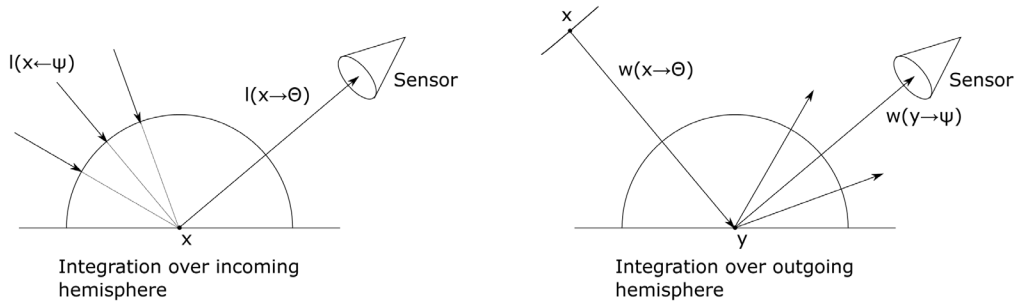


Figure 1. In the rendering equation integration is over the incoming hemisphere (left), whereas in the importance equation integration is over the outgoing hemisphere (right).

values $w(y \rightarrow \Psi_i)$ on the right-hand side of Equation 3 and Equation 4 are, however, unknown. Consequently, another evaluation is necessary. In general, the rays $y \rightarrow \Psi_i$ hit another interface at y′, resulting into nested integrals

$$w(x \rightarrow \Theta) = \int_{\Omega_y} \rho \int_{\Omega_{y'}} \rho w(y' \rightarrow \Psi) \cos(N_{y'}, \Psi)\, d\omega_{y'} \cdot \cos(N_y, \Psi)\, d\omega_y \quad (5)$$

with an increasing recursion depth for an increasing number of interfaces in the scene. In practice, these nested integrals can be solved by nested Monte Carlo integration schemes, leading to trees of paths traced through the scene. The number of nodes thereby increases exponentially with the number of intersections. Another strategy is to sample at every intersection a single direction in which the ray is further traced through the scene. As a result, more rays need to be sent into the scene. In this work each ray that hits a surface generates two follow-up rays – one accounting for reflection and one accounting for refraction. Figure 2 shows several rays emitted from an array element. These rays are traced through the scene, and received by another array element through several reflections and refractions. The simple Monte Carlo integration scheme outlined above leads to an unbiased estimation, i.e. the expected value $E(\hat{w})$ equals the true value w. However, in practice the recursive evaluation requires a stop criterion. If a maximal recursion depth is introduced, paths are stopped via an upper bound for the number of echoes. Therefore, potential important signal paths are cut from the estimation, and a bias is introduced to the estimator. An alternative approach to this problem is known as Russian Roulette. At every interface, paths are stopped with some probability. This keeps the recursion depths manageable, however also a smaller number of long paths are produced. This allows for the exploration of the whole scene, with, in theory, paths with infinite recursion depth. To compensate for the lower probability of longer paths, each sample is weighted with the reciprocal of its continuation probability, which again results in an unbiased simulator[37].

2.1.1 Surface model. At a ray interface intersection, a reflected ray is generated with probability p_R and a refracted ray is generated with probability p_T, representing the fraction of energy which is reflected and transmitted. In order to account for the spreading of the reflected and refracted energy, we use a Phong-like cosine parameterization of the surface similar to 27. A ray is reflected in direction Θ according to the hemispherical distribution

$$\rho = k(R \cdot \Theta)^N \quad (6)$$

with the specular reflection direction R. The exponent N defines the shape of the function with the degenerated case N=0 for a perfect diffuse reflector and $N \rightarrow \infty$ for a perfect specular reflector (see Figure 3). The normalization factor k depends on

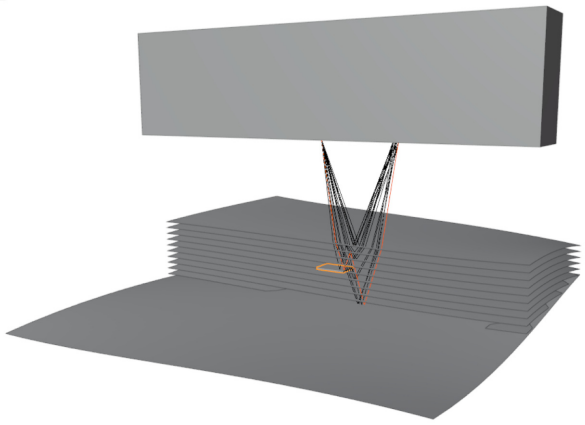
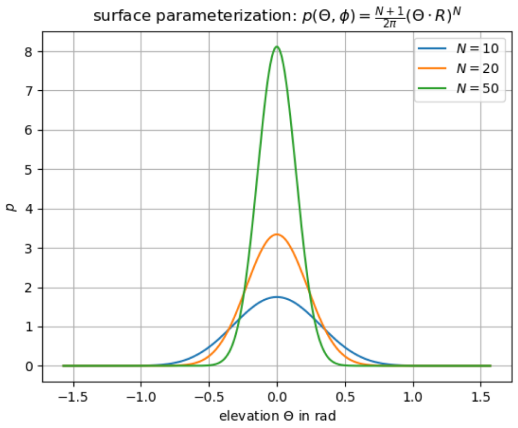


Figure 2. Sample showing a simulation of phased array sensor, artificial specimen, defect, and rays of a single sender-receiver pair.



(a) Example rays at interface.



(b) Cosine-lobe parameterization.

Figure 3. At interfaces, reflected and refracted energy is scattered according to a cosine-lobe surface parameterization (**a**); (**b**) shows the probability distribution used for sampling reflected and refracted rays with the elevation Θ, i.e. the angle between the ray and the specular reflection or refraction direction. The parameter N determines how strongly the energy is spread where $N \rightarrow \infty$ representing the edge case of perfect reflection/refraction.

N and is chosen such that we obtain a valid distribution density function (integrating over the hemisphere must equal 1). The specular reflection direction R is calculated as

$$R = \Theta - 2(\Theta \cdot n)n. \tag{7}$$

The direction of the refracted ray leaving a medium with refraction index η_1 and entering a medium with refraction index η_2 is calculated via Snell's law

$$\eta_1 \sin\Theta_1 = \eta_2 \sin\Theta_2 \tag{8}$$

with the angle of the entering ray Θ_1 and the angle of the refracted ray Θ_1 with respect to the surface normal.

In the context of Monte Carlo integration, in theory, every arbitrary probability function to generate samples produces an unbiased estimator. However, the variance of the estimator can be reduced by a suitable choice of the probability function (this is known as importance sampling). In practice, samples can be generated utilizing the surface model. This is obtained by sampling the azimuth angle uniformly in [0, 2 π]; elevation $R \cdot \Theta$ is sampled by evaluating the inverse cumulative distribution $F^{-1}(u)$ function at the uniform variable $u \in [0, 1]$ where

$$\begin{aligned} F(y) &= 1 - \cos^{N+1} y \\ F^{-1}(u) &= \arccos \sqrt[N+1]{1-u}. \end{aligned} \tag{9}$$

As the cosine lobe is centered at the specular reflection (refraction) direction, there is a probability to sample a direction on the other side of the surface patch. In such a case, we simply resample a new direction.

2.1.2 Implementation and image generation. For image generation, we shoot rays originated at the center of the transducer elements in directions sampled from a normal distribution. The intensity contribution to the final image is accumulated according to Equation 4. In order to sample reflected and refracted rays according to the described surface model, each surface in the scene is parameterized in terms of transmissivity, reflectivity, and the exponent N determining the cosine-surface parameterization in Equation 6. Each emitted phonon carries the following information:

- emitter index,
- receiver index,
- a weighting factor for Russian Roulette compensation and incorporating foreshortening,
- list of interface intersections to reconstruct the travelling time.

Algorithm 1 shows how the described strategies are combined. To explain how phonons contribute to individual pixels, we need to shortly introduce the TFM - the post-processing method of our target system. TFM was introduced in 40 and is one of the most popular methods for post-processing phased array data. The basis for image creation is a full matrix capture (FMC) of a phased array. This refers to a data acquisition that captures the responses of every possible transmitter-receiver combination. Figure 4 illustrates how the FMC captures contribute to pixel intensity values. First, the path length from the sender at (x_i, 0) to the pixel at (x, z) to the receiver at (x_j, 0) is calculated. The Hilbert transform of the captured response function h_{ij} – which is the envelope of the captured time signal – is evaluated at the calculated signal travelling time. Finally, the intensity value I of a pixel at (x, z) is calculated as the sum of the contributions of all sender-receiver pairs

$$I(x,z) = \sum_{i=1}^{N}\sum_{j=1}^{N} h_{ij}\left(\frac{\sqrt{(x-x_i)^2+z^2}+\sqrt{(x-x_j)^2+z^2}}{c}\right). \tag{10}$$

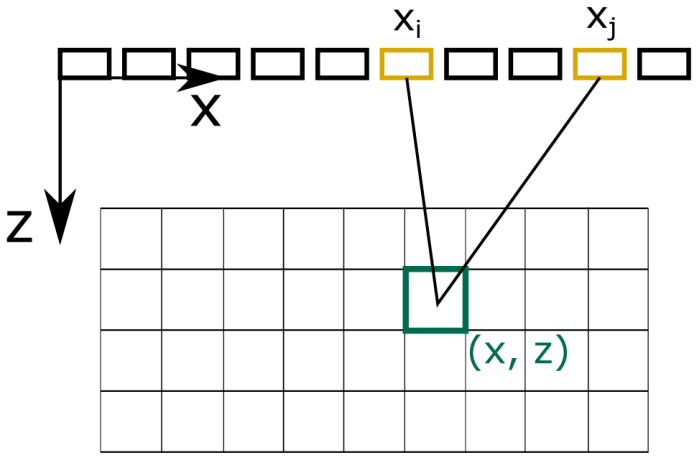


Figure 4. Total focusing method. Full matrix captures contribute to pixel intensity values according to path lengths from a sender *i* to a pixel at (*x*, *z*) to a receiver *j*.

Algorithm 1. Stochastic phonon tracing

```
1 Input: x, Θ, w, max. recursion depth
  if not reached max. recursion depth then
    find next intersection y on x → Θ;
    if intersects with receiver then
      add pixel contributions;
      return;
    Pick u uniformly from [0, 1);
    if u > nonTerminationProbabilty then
      return;
    w *= nonTerminationProbabilty;
    Pick u uniformly from [0, 1)];
    if u < reflectivity then
      sample reflected direction Ψ;
      w *= cos(N_y, Ψ);
      trace reflected ray;
    Pick u uniformly from [0, 1);
    if u < transmissivity then
      sample refracted direction Ψ;
      w *= cos(N_y, Ψ);
      trace refracted vector;
```

To mimic this post-processing, we calculate the travelling time of a traced phonon, and evaluate to which pixels the phonon contributes according to Equation 10. Therefore, every phonon contributes to several pixels lying on an ellipse with foci being the centers of the receiver and transmitter element.

2.2 Domain randomized training

Simulating sound propagation in CFRP parts is a difficult task, as due to its composition its mechanical properties are highly anisotropic. The presented approach is a simple and easy to implement simulation which does not take into account more complicated wave phenomena, which arise due to a directional dependent wave velocity. Further, low frequency effects, such as diffraction, are ignored. This results into less rich, low-fidelity TFM images. However, these low-fidelity images contain some of our physical knowledge about the inspection system, i.e. reverberation, backwall attenuation, and reflections caused by delaminations. Therefore, some of our physical knowledge is exposed to the classifier during training. The hypothesis of domain randomization is that if the variability in simulation during training is sufficiently large, the trained model will generalize to the real world without additional training[6]. Furthermore, by providing a wide range of (unrealistic) data, but which still contains important physical phenomena, that are important to judge the integrity of a specimen, some knowledge can be transferred to and therefore exploited by the ML model.

Parameter identification. For simulating TFM images with our simulator as outlined in the previous chapter, parameter identification is required. Besides the TFM post-processing the considered inspection system also applies a depth-dependent gain to compensate for signal attenuation. To generate training data, we simply adjust the interface parameters by hand in order to achieve a similar outcome compared to our reference system. In particular, we tune the parameters in such a way that we get similar data without applying any post-processing. These values are used as mean values for later parameter randomization. For this purpose, we set the transmissivity of the CFRP sheets to 1 and adjusted the reflectivity in such a way that we get a similar outcome to the reference TFM images. In the same way, as a depth dependent gain corrupts the physical interpretation of received signal energies, the sum of reflected and refracted energy due to this choice is higher than the incident energy. Figure 5 shows examples of the used training data. The shape of the specimen is defined via nurbs surfaces, which are randomized. These randomized shapes are scanned with

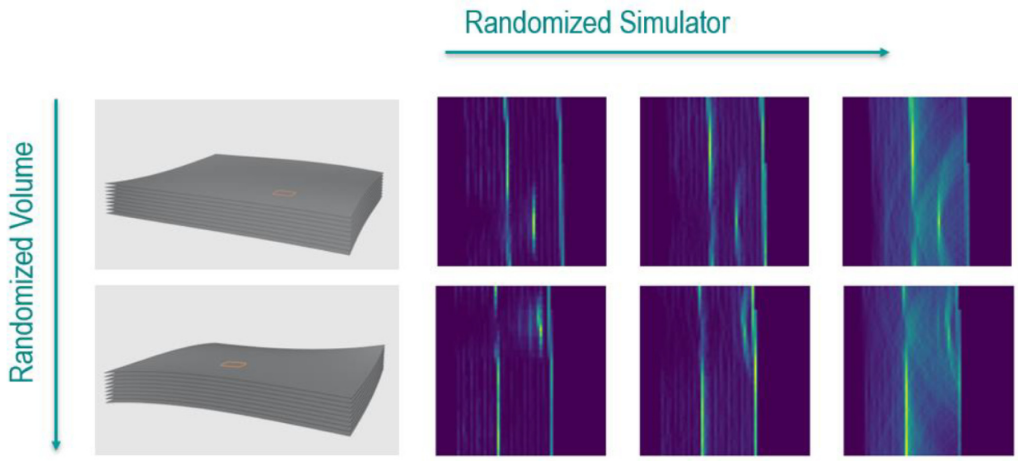


Figure 5. Randomization for artificial ultrasound data generation. The shape of the specimen and the location and size of the defect is randomized (columns). Artificial total-focusing method (TFM) data is simulated with randomized values for transmissivity and reflectivity (row).

the simulator with randomized values for transmissivity and reflectivity.

2.2.1 Model architecture and training. For semantic segmentation we use a convolutional neural network with a U-Net architecture. This architecture - originally proposed for segmentation of medical volume data by Ronneberger, *et al.* in 41 - proved itself to be very successful in a variety of use cases. The network parameters are initialized randomly. Adam optimization is used to minimize the negative log likelihood loss function with mean reduction. As the ideal shape of inspected specimens is typically known, this information is often cheap and easy to integrate for evaluation. Therefore, we decided to additionally provide a distorted version of the volume as input. Distortions are achieved by applying random rotations in the range of [-3°, 3°] and by scaling the object with a factor in the range of [0.9, 1.1]. Those values were chosen in order to sufficiently cover variations as seen in production, and in the data acquisition process assuming a robotic inspection system, such as manufacturing tolerances, calibration inaccuracies, and robot path inaccuracies. Examples of distorted input volume masks are shown in Figure 6. For rendering TFM images according to Algorithm 1 we sent 150 rays per element into the scene. The non-termination probability was set to 0.95 and the maximum recursion depth was set to 30.

3 Results and discussion

The introduced ray-tracing based simulator requires about one minute of computation time on a standard office computer for rendering a single TFM image. Although the implementation was not optimized towards speed efficiency, this allows to generate a large data set in a reasonable period of time. Further, the simulator reproduces important physical effects and image artifacts which are used for data interpretation by human inspectors, i.e. reverberation, backwall attenuation, and reflections caused by delaminations. This makes the proposed method a good candidate for enabling domain randomized training.

To verify the approach, we used scans gathered from a CFRP reference specimen, illustrated in Figure 7. Flat-bottom holes were introduced into the part as artificial reference defects (reflectors). This is a common practice for developing and verifying ultrasonic NDT methods. In total, the sample contains twelve flat-bottom holes, which are placed in the following pattern: four defects with varying radii are introduced within a single plane. Radii from left to right measure: 3 mm, 1.5 mm, 3 mm, 6 mm. This is repeated three times, while in each subsequent plane the defects are introduced in different depths, resulting in four defects with a distance of 1mm from the backwall, in the middle of the part, and with a distance of 1 mm from the frontwall. The reflectors (i.e. flat-bottoms of holes) are parallel to the frontwall of the specimen. The front and backwall of the specimen are slightly tilted resulting in different thicknesses in each plane, ranging from 7 mm to 10 mm.

The ultrasound sensor used for scanning the reference specimen is a 5MHz linear phased array that consists of 114 transducers

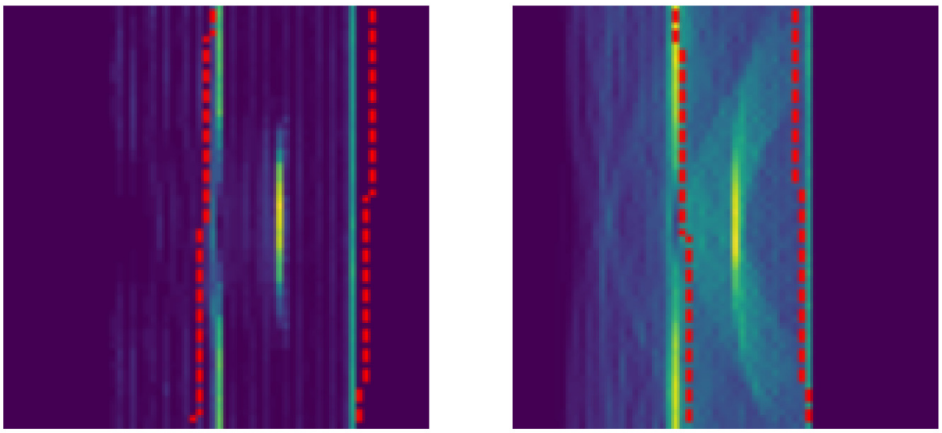


Figure 6. Training was conducted with volume masks as additional input. To disturb the masks, random rotation from -3 to 3 degrees and a random scale from 0.9 to 1.1 were applied. The resultant volume masks are indicated by the red-dotted line.



Figure 7. Flat-bottom holes introduced into reference specimen.

and has a width of 114 mm. The TFM images have a resolution of 0.1 mm and 16 bit depth. When feeding real data to the machine learning model, it outputs reasonable results. An example is shown in Figure 8. Thereby we provided volume masks obtained from a digital model of the specimen.

Due to limited availability of scan data, the evaluation of defect detection on a more quantitative level was difficult. Nevertheless, the developed defect detection method seems very promising. An interesting detail is that the method correctly classifies defects near the frontwall. Furthermore, the ghost artifact, which actually is just an echo of a real defect close to the top surface, is correctly classified as background, i.e. as "no defect" (yellow arrow in Figure 8). The top row in Figure 9 shows the result of defect detection as 3d point cloud, the backwall of the specimen is indicated by a gray surface. Yellow points show the resulting output of the defect detection, blue points on the right show the locations of flat-bottom holes (voxelized ground truth).

In Figure 9, defects in the left two rows are those with 3mm and 1.5mm diameter. The two rows on the right represent defects

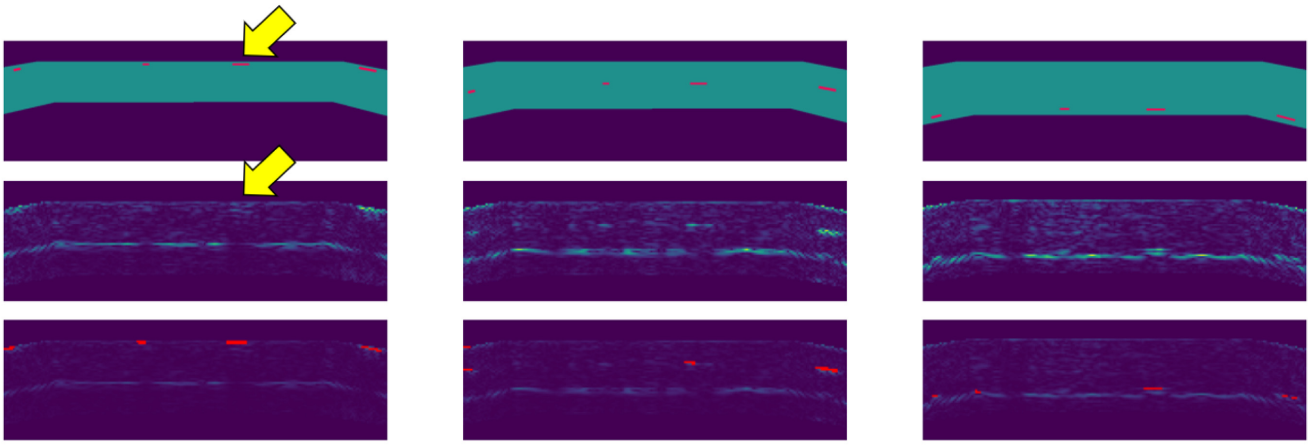


Figure 8. Defect detection on the reference specimen. The individual columns show section images at the location of flat-bottom holes close to the top surface (left), in the middle of the part (center), and close to the backwall (right). The top images show the part geometry and locations of flat-bottom holes. The middle images show original images with full brightness. The bottom images show the original data with reduced brightness and detections overlaid in red. The echo of a defect (yellow arrow[1]) is correctly classified as background.

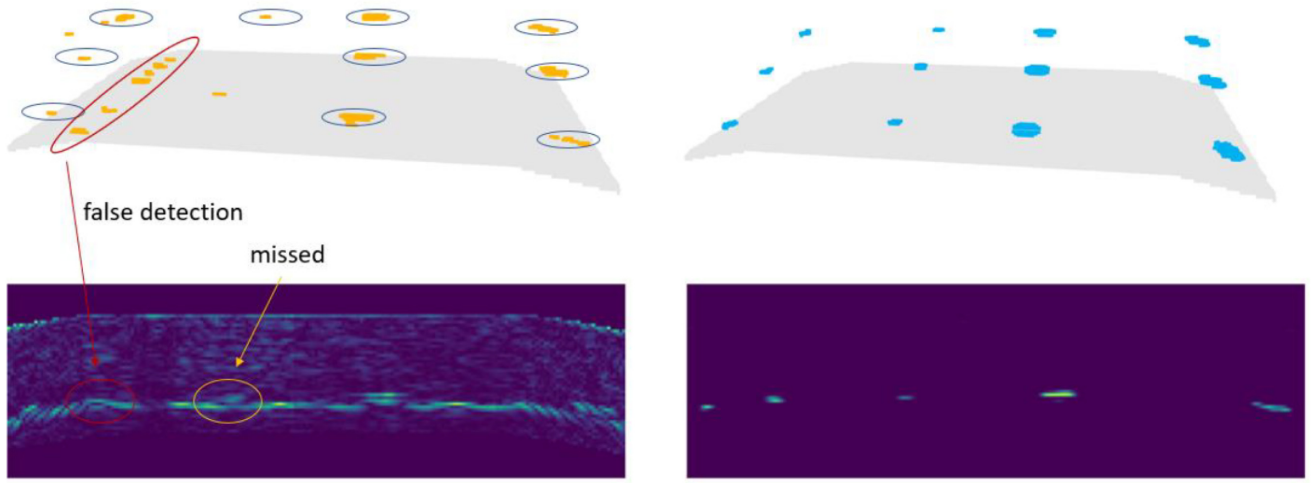


Figure 9. The top row shows the defect detection result (left) and the ground truth data (right) as 3d point clouds. The backwall of the specimen is indicated by a gray surface. The bottom row shows the input image (left) and defect probability output of the U-Net (right) of a single slice.[2]

[1]Annotations were added in Powerpoint.

[2]3d point clouds were visualised with open3d in python and annotations were added via Powerpoint.

with 3mm and 6mm diameter. The larger defects were detected by the trained neural network. Concerning the 1.5 mm diameter defects, the model failed in finding one of them in the middle and one closer to the backwall. Further, false-positives occurred at the transition between the inclined and the horizontal section of the backwall.

4 Conclusion

A simulator with low computational costs for rendering artificial TFM images is introduced. This simulator generates low-fidelity images, which yet reproduce important physical effects and artifacts. Domain randomized training of a semantic segmentation model was conducted without any domain adoption techniques. Thereby, in addition to the TFM images, distorted volume masks are fed into the network. The network, which was trained solely on artificial data, achieves reasonable results when applied to real world data. Further, the network correctly differentiates between real artifacts and ghost artifacts, which is known as a hard problem.

Domain randomization could achieve impressive results in various fields, however its potential is yet not much investigated in the NDT community. Results are promising, and we aim at a wider evaluation on real data from industry in the future. Training on simulated data could ease training for specialized cases where data is scarce and data of edge cases are not available. Besides a training solely on artificial data, the proposed method could serve as pre-training step and enable fine-tuning with small real world data sets.

False-positives occur almost exclusively on sharp corners of the specimen. This could be caused by the used specimen model, which is very smooth. In future work, a geometrically richer virtual specimen will be used, which better represents the considered use case.

Data availability

Underlying data

The reference specimen used in this work is owned by Fokker Landing Gear and the National Aerospace Laboratory (NLR). The authors were given access to the specimen as part of the Clean Sky 2 project under grand agreement No 831830. The raw ultrasonic dataset is held under restriction to internal use only, as a third party with access to the raw ultrasonic data may be able to reverse engineer very specific details about material type and laminate layup, which may compromise the GKN Fokker commercial position. In general, data with real flaws tend to be scarce, and data is often confidential and rarely shared. Due to the lack of existence, representative open source data cannot be provided.

Ethics and consent

Ethics approval and consent were not required for this study.

Acknowledgements

The authors would like to thank Peet Vergouwen (GKN Fokker) for providing the sample part and manufacturing data. The authors also wish to thank Andrey Bulavinov, Gerbasi Alexander and Roman Pinchuk (ACS-Solutions GmbH) who conducted ultrasonic measurements and provided insights into data processing of phased array data.

References

1. Okulla M, Düfert U, Bulavinov A, *et al.*: **Fortschrittliche Prüfmethoden zur Prüfung von CFK-Großkomponenten mit komplexer Geometrie**. In *DGZfP Jahrestagung 2017, 22-24 May, Koblenz, Germany (DGZfP 2017)*. 2017.
Reference Source
2. Virkkunen I, Koskinen T, Jessen-Juhler O, *et al.*: **Augmented ultrasonic data for machine learning.** *J Nondestr Eval.* 2021; **40**(1): 1–11.
Publisher Full Text
3. Koskinen T, Virkkunen I, Siljama O, *et al.*: **The effect of different flaw data to machine learning powered ultrasonic inspection.** *J Nondestr Eval.* 2021; **40**(1): 1–13.
Publisher Full Text
4. Kemppainen M, Virkkunen I: **Crack characteristics and their importance to nde.** *J Nondestr Eval.* 2011; **30**(3): 143–157.
Publisher Full Text
5. Gardner P, Fuentes R, Dervilis N, *et al.*: **Machine learning at the interface of structural health monitoring and non-destructive evaluation.** *Philos Trans A Math Phys Eng Sci.* 2020; **378**(2182): 20190581.
PubMed Abstract | Publisher Full Text | Free Full Text
6. Tobin J, Fong R, Ray A, *et al.*: **Domain randomization for transferring deep neural networks from simulation to the real world**. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2017; 23–30.
Publisher Full Text
7. Mehta B, Diaz M, Golemo F, *et al.*: **Active domain randomization**. In Leslie Pack Kaelbling, Danica Kragic, and Komei Sugiura, editors, *Proceedings of the Conference on Robot Learning*. volume 100 of *Proceedings of Machine Learning Research*. PMLR, 30 Oct-01 Nov 2020; 1162–1176.
8. Loquercio A, Kaufmann E, Ranftl R, *et al.*: **Deep drone racing: From simulation to reality with domain randomization.** *IEEE Trans Robot.* 2019; **36**(1): 1–14.
Publisher Full Text
9. Borrego J, Dehban A, Figueiredo R, *et al.*: **Applying domain randomization to synthetic data for object category detection.** *arXiv preprint arXiv: 1807.09834.* 2018.
Reference Source
10. Polvara R, Patacchiola M, Hanheide M, *et al.*: **Sim-to-Real Quadrotor Landing via Sequential Deep Q-Networks and Domain Randomization.** *Robotics.* 2020; **9**(1): 8.
Publisher Full Text
11. Bowler AL, Bakalis S, Watson NJ: **Monitoring mixing processes using ultrasonic sensors and machine learning.** *Sensors (Basel).* 2020; **20**(7): 1813.
PubMed Abstract | Publisher Full Text | Free Full Text
12. Rodrigues LFM, Cruz FC, Oliveira MA, *et al.*: **Carburization level identification in industrial hp pipes using ultrasonic evaluation and machine learning.** *Ultrasonics.* 2019; **94**: 145–151.
PubMed Abstract | Publisher Full Text
13. Zhang M, Li M, Zhang J, *et al.*: **Onset detection of ultrasonic signals for the testing of concrete foundation piles by coupled continuous wavelet**

transform and machine learning algorithms. *Advanced Engineering Informatics.* 2020; **43**: 101034.
Publisher Full Text

14. Pan H, Pang Z, Wang Y, *et al.*: **A new image recognition and classification method combining transfer learning algorithm and mobilenet model for welding defects.** *IEEE Access.* 2020; **8**: 119951–119960.
Publisher Full Text

15. Bowler AL, Watson NJ: **Transfer learning for process monitoring using reflection-mode ultrasonic sensing.** *Ultrasonics.* 2021; **115**: 106468.
PubMed Abstract | Publisher Full Text

16. Escrig J, Simeone A, Woolley E, *et al.*: **Ultrasonic measurements and machine learning for monitoring the removal of surface fouling during clean-in-place processes.** *Food Bioprod Process.* 2020; **123**: 1–13.
Publisher Full Text

17. Fuchs P, Kröger T, Garbe CS: **Defect detection in ct scans of cast aluminum parts: A machine vision perspective.** *Neurocomputing.* 2021; **453**: 85–96.
Publisher Full Text

18. Dominguez N, Reverdy F: **Simulation of ultrasonic testing of composite structures**. In *11th European Conference on Non- Destructive Testing/ECNDT, Oct 6-11, Prague*. 2014; 6–10.

19. Ma J, Xue J, Yang S, *et al.*: **A study of the construction and application of a daubechies wavelet-based beam element.** *Finite Elem Anal Des.* 2003; **39**(10): 965–975.
Publisher Full Text

20. Pyle RJ, Bevan RLT, Hughes RR, *et al.*: **Deep learning for ultrasonic crack characterization in NDE.** *IEEE Trans Ultrason Ferroelectr Freq Control.* 2020; **68**(5): 1854–1865.
PubMed Abstract | Publisher Full Text

21. Gantala T, Balasubramaniam K: **Automated defect recognition for welds using simulation assisted TFM imaging with artificial intelligence.** *Journal of Nondestructive Evaluation.* 2021; **40**(28): 1–24.
Publisher Full Text

22. Østergaard ML, Nielsen KR, Albrecht-Beste E, *et al.*: **Simulator training improves ultrasound scanning performance on patients: a randomized controlled trial.** *Eur Radiol.* 2019; **29**(6): 3210–3218.
PubMed Abstract | Publisher Full Text

23. Lakka E, Malamos AG, Pavlakis KG, *et al.*: **Spatial sound rendering-a survey.** *International Journal of Interactive Multimedia and Artificial Intelligence.* 2018; **5**(3): 33–45.
Publisher Full Text

24. Hulusic V, Harvey C, Debattista K, *et al.*: **Acoustic rendering and auditory-visual cross-modal perception and interaction**. In *Computer Graphics Forum*. 2012; **31**(1): 102–131.
Publisher Full Text

25. Röber N, Kaminski U, Masuch M: **Ray acoustics using computer graphics technology**. In *10th International Conference on Digital Audio Effects, 11-14 September, Bordeaux, France*. 2007; 117–124.
Reference Source

26. Mattausch O, Goksel O: **Image-based reconstruction of tissue scatterers using beam steering for ultrasound simulation.** *IEEE Trans Med Imaging.* 2017; **37**(3): 767–780.
PubMed Abstract | Publisher Full Text

27. Mattausch O, Makhinya M, Goksel O: **Realistic ultrasound simulation of complex surface models using interactive monte-carlo path tracing**. In *Computer Graphics Forum.* 2018; **37**(1): 202–213.
Publisher Full Text

28. Yeo L, Romero R: **Optical ultrasound simulation-based training in obstetric sonography.** *J Matern Fetal Neonatal Med.* 2020; 1–16.
Publisher Full Text

29. Thanedar BD: **Monte Carlo investigation of transient acoustic fields in partially or completely bounded medium**. Ph. D. Thesis. dissertation, University of Virginia, 1972.
Reference Source

30. Leckey CAC, Parker FR: **Nde and shm simulation for cfrp composites**. In *Proc. of American Society for Composites 29th Technical Conference, 8 - 10 September, La Jolla, California, USA*. 2014; 454.
Reference Source

31. Atalla N, Bernhard RJ: **Review of numerical solutions for low-frequency structural-acoustic problems.** *Applied Acoustics.* 1994; **43**(3): 271–294.
Publisher Full Text

32. Bertram M, Deines E, Mohring J, *et al.*: **Phonon tracing for auralization and visualization of sound**. In *Proc. of the IEEE Visualization, Minneapolis, USA*. 2005; 151–158.
Publisher Full Text

33. Mineo C, Lines D, Cerniglia D: **Generalised bisection method for optimum ultrasonic ray tracing and focusing in multi-layered structures.** *Ultrasonics.* 2021; **111**: 106330.
PubMed Abstract | Publisher Full Text

34. Mineo C, Cerniglia D, Mohseni E: **Solving ultrasonic ray tracing in parts with multiple material layers through Root-Finding methods.** *Ultrasonics.* 2022; **124**: 106747.
PubMed Abstract | Publisher Full Text

35. Rabenstein R, Schips O, Stenger A: **Acoustic rendering of buildings**. In *5th Int. IBPSA Conference Building Simulation, September 8-10, Prague, Czech Republic*. 1997; **2**: 181–188.
Reference Source

36. Kajiya JT: **The rendering equation**. In *Proceedings of the 13th annual conference on Computer graphics and interactive techniques (SIGGRAPH ' 86)*. 1986; **20**(3): 143–150.
Publisher Full Text

37. Dutré P, Bekaert P, Bala K: **Advanced global illumination**. 2nd edition. CRC Press, 2016.
Reference Source

38. Siltanen S, Lokki T, Kiminki S, *et al.*: **The room acoustic rendering equation.** *J Acoust Soc Am.* 2007; **122**(3): 1624–1635.
PubMed Abstract | Publisher Full Text

39. Pattanaik SN, Mudur SP: **The potential equation and importance in illumination computations**. In *Computer Graphics Forum.* 1993; **12**(2): 131–136.
Publisher Full Text

40. Holmes C, Drinkwater BW, Wilcox PD: **Post-processing of the full matrix of ultrasonic transmit-receive array data for non-destructive evaluation.** *Ndt & E International.* 2005; **38**(8): 701–711.
Publisher Full Text

41. Ronneberger O, Fischer P, Brox T: **U-net: Convolutional networks for biomedical image segmentation.** *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015.* 2015; **9351**: 234–241.
Publisher Full Text

Open Peer Review

Current Peer Review Status: ✔ ✔

Version 2

Reviewer Report 23 May 2022

https://doi.org/10.21956/openreseurope.15998.r29193

© **2022 Virkkunen I.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Iikka Virkkunen**
Aalto University, Espoo, Finland

The authors address an important topic, i.e. the various ways to handle data scarcity in using machine learning techniques for NDE data evaluation. In particular, the authors target defect detection in TFM data frames acquired from carbon fibre composite components using U-net style segmentation network. The main contribution is to train the network using low fidelity simulated responses with the hypothesis that even substantially limited simulation based training data will provide sufficient generalization for finding defects, provided that the training data is sufficiently varied.

As a previous reviewer has already pointed out, there's wide variety of available methodologies for simulating ultrasonic responses, and the selection here is by no means obvious. Thus, a better justification of the chosen modelling approach would be warranted. Perhaps the choice was made in favor of a fairly simplistic modeling approach in order to study the value of even limited fidelity training material.

The results from the ML segmentation are compared against flat-bottom holes of various sizes. The authors state, that these kind of simplified flaws are commonly used develop and verify ultrasonic NDT methods. This is correct as far as providing reference signals for human inspectors go. However, using such simplified flaws to assess quality of ML training requires more justification. In particular here, where the object of study is the value extracted from simplified training data, comparing with simplified reference flaws may yield overly optimistic results. To balance, the authors do not claim high performance, instead, the performance is described as "promising" or "reasonable". Nevertheless, some discussion on the limitations of the current approach would be warranted. In addition, it would be good to give some sort of reference to the reader on the difficulty of these flaws: would a human inspector be expected to find all of them easily? Would the missed defects be considered difficult for the human inspector?

The above potential improvements notwithstanding, the paper presents a clear contribution to the field and suggests an additional tool to aid training state of the art machine learning networks with limited NDE data.

Is the rationale for developing the new method (or application) clearly explained?
Yes

Is the description of the method technically sound?
Yes

Are sufficient details provided to allow replication of the method development and its use by others?
Yes

If any results are presented, are all the source data underlying the results available to ensure full reproducibility?
Partly

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: NDE reliability, Machine learning for NDE

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 12 May 2022

https://doi.org/10.21956/openreseurope.15998.r29253

© 2022 Mineo C. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Carmelo Mineo
University of Palermo, Palermo, Italy

The authors have addressed all my comments appropriately.

Is the rationale for developing the new method (or application) clearly explained?
Yes

Is the description of the method technically sound?
Yes

Are sufficient details provided to allow replication of the method development and its use

by others?
Yes

If any results are presented, are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Non-destructive testing, Ultrasonics, Ray Tracing, Phased Array, Composite materials, Robotic and autonomous inspections

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 11 April 2022

https://doi.org/10.21956/openreseurope.15495.r28936

© 2022 Mineo C. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Carmelo Mineo
University of Palermo, Palermo, Italy

The authors present an interesting work to overcome the data scarcity problem of machine learning for non-destructive testing. They propose a method to enable training of neural networks on simulated data. This is aligned with my research interests and I appreciated the invitation to review this manuscript. The article is well written. The background motivations are clear. I think this work contributes to the state-of-the-art related to automated flaw detection in ultrasonic non-destructive testing. The authors should address the following questions/issues:

About ultrasonic ray-tracing, the authors state: “The solution can be approximated by classical methods, e.g. by finite element method or boundary element method[21,28].” The authors seem to not be aware of recent solutions that have been proposed to solve the ultrasonic ray-tracing problem using iterative root-finding methods. These methods are described in the following publications:

1. Mineo, C., Lines, D., & Cerniglia, D. (2021). Generalised bisection method for optimum

ultrasonic ray tracing and focusing in multi-layered structures. Ultrasonics, 111, 106330.

2. Mineo, C., Cerniglia, D., & Mohseni, E. (2022). Solving ultrasonic ray tracing in parts with multiple material layers through Root-Finding methods. Ultrasonics, 124, 106747.

The authors should explain the advantages offered by their ray-tracing solution. The description of the approach used for the simulation of sound propagation occupies most part of the manuscript. In my opinion, the authors should explain why existing approaches were not suitable to support their work on neural network training. In section 2.2, the authors state: “The presented approach is a simple and easy to implement simulation which does not take into account more complicated wave phenomena, which arise due to a directional dependent wave velocity.”. It is not clear why the authors have not tried to use existing and more accurate solutions. Was that deemed unnecessary? If that is the case, why?

The ray tracing approach proposed by the authors seems to be capable of dealing with complex geometries. The authors say they have randomized the part geometries by applying random rotations in the range of [-3°, 3°] and by scaling the object with a factor from [0.9 to 1.1]. It is not clear if these ranges were limited by the capability of the ray-tracing approach to work with excessive object curvature of if they were randomly chosen by the authors. If the latter is the case, what is the maximum object curvature the approach can tolerate?

In the text following Equation 8, replace “refracted ray Θ_1” with “refracted ray Θ_2”.

In section 3, the following statement seems to contain a contradiction: “An interesting detail is that the method correctly classifies defects near the frontwall. Furthermore, the echo from a defect close to the top surface is correctly classified as background, i.e. as “no defect” (yellow arrow in Figure 8).”. I think the two sentences in the text above refer to the same subject, but have the opposite meaning.

References

1. Mineo C, Lines D, Cerniglia D: Generalised bisection method for optimum ultrasonic ray tracing and focusing in multi-layered structures.*Ultrasonics*. 2021; **111**: 106330 PubMed Abstract | Publisher Full Text
2. Mineo C, Cerniglia D, Mohseni E: Solving ultrasonic ray tracing in parts with multiple material layers through Root-Finding methods. *Ultrasonics*. 2022; **124**. Publisher Full Text

Is the rationale for developing the new method (or application) clearly explained?

Yes

Is the description of the method technically sound?

Yes

Are sufficient details provided to allow replication of the method development and its use by others?

Partly

If any results are presented, are all the source data underlying the results available to ensure full reproducibility?

Partly

Are the conclusions about the method and its performance adequately supported by the findings presented in the article?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Non-destructive testing, Ultrasonics, Ray Tracing, Phased Array, Composite materials, Robotic and autonomous inspections

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 09 May 2022

Klaus Schlachter

Dear Carmelo Mineo, Thank you very much for your constructive review of our work. We appreciate your comments and suggestions to improve the manuscript. Please find our answers to the points raised and comments below. Our responses are shown in italics:

- About ultrasonic ray-tracing, the authors state: "The solution can be approximated by classical methods, e.g. by finite element method or boundary element method[21,28]." The authors seem to not be aware of recent solutions that have been proposed to solve the ultrasonic ray-tracing problem using iterative root-finding methods. These methods are described in the following publications:

1. Mineo, C., Lines, D., & Cerniglia, D. (2021). Generalised bisection method for optimum ultrasonic ray tracing and focusing in multi-layered structures. Ultrasonics, 111, 106330.
2. Mineo, C., Cerniglia, D., & Mohseni, E. (2022). Solving ultrasonic ray tracing in parts with multiple material layers through Root-Finding methods. Ultrasonics, 124, 106747.

Thank you very much for pointing this out. Indeed, the introduction section regarding geometric acoustics was rather general, and we think it benefits from including more specific research that applies geometric acoustics for solving problems related to ultrasonic testing. We have added the following paragraph to the manuscript:
"In general, paths between two points are determined by an underlying variation problem known as Fermat's principle. In homogenous media rays travel along straight lines and change directions at boundaries of media with different wave velocities. Modern root finding methods allow to compute sound paths between a source and target point in multi-layered structures, such as composite parts, very efficiently [Mineo2021, Mineo2022]. For this purpose, the problem is modeled based on Snell's law. This approach can be extended for determining sound paths in media with continuously changing material properties by discretizing the media through regions with constant wave velocities. Such an approach can be applied to various problems in the NDT field, e.g. for efficiently determining ultrasonic paths for in-process inspections of welds at the presence of high temperature gradients. While such problems can also be tackled by finite

element models, such an approach is of great advantage in terms of the required simulation time [Mineo2021, Mineo2022].”

- The ray tracing approach proposed by the authors seems to be capable of dealing with complex geometries. The authors say they have randomized the part geometries by applying random rotations in the range of [-3°, 3°] and by scaling the object with a factor from [0.9 to 1.1]. It is not clear if these ranges were limited by the capability of the ray-tracing approach to work with excessive object curvature or if they were randomly chosen by the authors. If the latter is the case, what is the maximum object curvature the approach can tolerate?

As shown in Figure 6 and described in subchapter “Model architecture and training” we provide a volume mask as an additional input channel to the segmentation model because the ideal shape of an inspected specimen is typically know. However, due to manufacturing tolerances, mesh registration inaccuracies (assuming an automated robotic inspection system), and limited robot trajectory accuracy, it is not reasonable to assume that captured TFM images are perfectly aligned with production drawings. Therefore, this additional input channel is meant to provide context information, however, in order to prevent the neural network from simply relying on the volume mask, the volume mask was distorted within a reasonable range.
Indeed, the virtual model used for training is pretty smooth, though, we do not see any obvious reason why the presented approach should be limited by curvature. We think that the training would benefit from a geometrically richer virtual specimen. In future work this would be interesting for further investigations. In order to clarify this, we added text in bold to the manuscript:
"Distortions are achieved by applying random rotations in the range of [-3°, 3°] and by scaling the object with a factor in the range of [0.9, 1.1]. Those values were chosen in order to sufficiently cover variations as seen in production, and in the data acquisition process assuming a robotic inspection system, such as manufacturing tolerances, calibration inaccuracies, and robot path inaccuracies."

- The authors should explain the advantages offered by their ray-tracing solution. The description of the approach used for the simulation of sound propagation occupies most part of the manuscript. In my opinion, the authors should explain why existing approaches were not suitable to support their work on neural network training. In section 2.2, the authors state: “The presented approach is a simple and easy to implement simulation which does not take into account more complicated wave phenomena, which arise due to a directional dependent wave velocity.”. It is not clear why the authors have not tried to use existing and more accurate solutions. Was that deemed unnecessary? If that is the case, why?

Thank you for the point raised. The scarcity of flawed training data is a general problem in the NDT field. This is even a common problem for training human inspectors. Machine learning requires even larger data sets, and the lack of representative flawed training data is a major challenge towards machine learning in the NDT field. Approaches towards training with synthetic data in current literature typically try to use accurate but very time-consuming FE simulations (most often, additionally, to a smaller set of real-world data). We added two recent and very interesting articles to our survey which speed up simulations by (i) mixing FE simulations (used in the defect locality) and ray-based simulations and by (ii) a deep learning approach. To clarify our point, we added the following to the paragraph “Ultrasonic image rendering” in chapter 1.1:
“However, wave-based simulations are typically very time-consuming, which makes it difficult to generate large data sets as required for deep learning. To reduce simulation time, in [Pyle2021] a

mixture of finite element models in the defect locality and ray-based models for the whole region of interest is used. In [Gantala2021] a fast deep convolutional generative adversarial network is used to generate synthetic weld TFM images. For training, synthetic data is generated via finite element simulations."

In the medical field, a lot of effort is put into developing methods for real-time rendering ultrasonic images (for training humans) based on a global illumination model. We think that the NDT field could benefit from these developments. Global illumination models are able to reproduce important image artefacts, as they incorporate indirect sound paths, but they typically do not incorporate low frequency effects. However, signal properties and image artefacts which are important for interpretation and utilized by human inspectors, such as wall-echo shadows and ghost artefacts, are reproduced with the presented simulation method. The hypothesis of this work is that with enough variability in simulation but reproducing import physical effects, the model will generalize to the real world (see chapter 2.2). The huge advantage of these global illumination models is that with a ray tracing approach the solution can be approximated very time efficient (medical sonography can be simulated in real-time). With the implemented simulator it is possible to generate a large data set within a short period of time, although the implementation is not optimized in any means (a single TFM image can be rendered in about a minute). Further, a major aim of our work is to investigate, if it is possible to achieve reasonable results using simple simulation models via domain randomization. Investigations of the influence of the accuracy of the used model on the detection accuracy would be interesting for future work. To clarify the advantage of the proposed ray-tracing solution, we added the following paragraph at the beginning of chapter 3:
"The introduced ray-tracing based simulator requires about one minute of computation time on a standard office computer for rendering a single TFM image. Although the implementation was not optimized towards speed efficiency, this allows to generate a large data set in reasonable period of time. Further, the simulator reproduces important physical effects and image artifacts which are used for data interpretation by human inspectors, i.e. reverberation, backwall attenuation, and reflections caused by delaminations. This makes the proposed method a good candidate for enabling domain randomized training."

References:
[Pyle2021] Pyle, Richard J., et al. "Deep learning for ultrasonic crack characterization in NDE." IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control 68.5 (2020): 1854-1865.
[Gantala2021] Gantala T., Balasubramaniam K. "Automated defect recognition for welds using simulation assisted TFM imaging with artificial intelligence." Journal of Nondestructive Evaluation 40.1 (2021): 1-24.

- In the text following Equation 8, replace "refracted ray Θ_1" with "refracted ray Θ_2".

Thank you very much for pointing out that typographical error. We corrected the index in the manuscript.

- In section 3, the following statement seems to contain a contradiction: "An interesting detail is that the method correctly classifies defects near the frontwall. Furthermore, the echo from a defect close to the top surface is correctly classified as background, i.e. as "no defect" (yellow arrow in Figure 8).". I think the two sentences in the text above refer to the same subject, but have the opposite meaning.

Unfortunately, that was not cleary phrased. The defect close to the top surface causes several

echos, therefore, the defect is visible a second time in the TFM images. However, only the signal close to the top surface corresponds to the real defect. The second echo is a ghost artifact. Such ghost artefacts are typical in TFM images and are also produced by the developed simulator used in this work. We wanted to point out that the trained network correctly classifies this second echo as no defect. The text was changed to: "An interesting detail is that the method correctly classifies defects near the frontwall. Furthermore, the ghost artifact, which actually is just an echo of a real defect close to the top surface, is correctly classified as background, i.e. as "no defect" (yellow arrow in Figure 8).

Competing Interests: No competing interests were disclosed.