A deep learning based methodology for artefact identification and suppression with application to ultrasonic images

Sergio Cantero-Chinchilla*, Paul D. Wilcox and Anthony J. Croxford

Department of Mechanical Engineering, University of Bristol, Bristol, BS8 1TR, UK

ARTICLE INFO

Keywords:
Non-destructive evaluation
Deep learning
Autoencoders
Ultrasound
Phased-arrays
Artefact identification
Full matrix capture
Total focusing method

ABSTRACT

This paper proposes a deep learning framework for artefact identification and suppression in the context of non-destructive evaluation. The model, based on the concept of autoencoders, is developed for enhancing ultrasound inspection and defect identification through images obtained from full matrix capture data and the total focusing method. An experimental case study is used to prove the effectiveness of the method while exploring its practical limitations. A comparison with a state-of-the-art methodology based on image analysis is addressed for the identification and suppression of artefacts. In general, the proposed method efficiently provides accurate suppression of artefacts in complex scenarios, even when the defect is located below the footprint of the ultrasonic probe, and also yields the physical parameters needed for imaging as a by-product.

1. Introduction

Defect characterisation and damage identification are of utmost importance for safety critical industries such as the oil & gas and nuclear industry. To this end, intense structural maintenance and inspection campaigns are typically carried out periodically as a part of scheduled-based maintenance plans. These typically include the use of different non-destructive testing (NDT) techniques, e.g. ultrasound, eddy currents, and radiography [1–3]. However, the identification and characterisation of defects are limited in most non-destructive evaluation (NDE) modalities by a combination of two factors: (1) *noise*, which may be random if it is aleatory or coherent if it stems from the material microstructure; and (2) *artefacts* that originate from structural features, e.g. ultrasonic echoes from structural features. The presence of artefacts in NDT data may lead to costly false positives or potentially catastrophic consequences if they obscure the presence of a defect. Distinguishing between defects and artefacts is one of the big challenges for operators. Therefore, there is a compelling need for artefacts to be removed from the raw NDT data in an efficient and accurate manner given that artefacts restrict the inspection range, hence raising costs.

Amongst the available NDT techniques, ultrasonic testing and imaging is widely employed and may be heavily impacted by artefacts stemming from structural features such as specimen boundaries. Ultrasonic testing can be carried out using a single element or an array probe. The multiple elements of the latter can be simultaneously actuated by using a relative time-delay to steer a beam in the desired direction of inspection (i.e. classical beamforming). Alternatively, the elements can be excited individually while the data is recorded by all the receivers in multiple A-scans, which leads to a larger and more useful amount of information for post-processing (e.g. using synthetic beamforming). This technique is also known as full matrix capture (FMC) [4]. Furthermore, ultrasonic images with defect information can be produced from the FMC data using different algorithms, e.g. the sector B-scan or the total focusing method (TFM) [5–7] using the linear delay-and-sum algorithm. Note that the TFM can focus in both transmission and reception at any point within the image, which makes it flexible and useful for NDE. To obtain even more information out of a single FMC, the multi-view TFM [8] can also be adopted, whereby multiple views from the same region can be obtained by considering multimodal ray paths. However, because structural artefacts are unavoidable in FMC data, these may lead to imaging artefacts at non-physical locations when multi-view images are formed using more complex ray paths. In this context, the structural artefacts recorded in the time-traces of the FMC are unavoidable and reconstruct in the ultrasonic images at locations where defects may be present. Note that the area below the footprint of the array probe

*Corresponding author
Email addresses: sergio.canterochinchilla@bristol.ac.uk (S. Cantero-Chinchilla); p.wilcox@bristol.ac.uk (P.D. Wilcox); a.j.croxford@bristol.ac.uk (A.J. Croxford)
ORCID(s): 0000-0002-6235-5304 (S. Cantero-Chinchilla); 0000-0002-8569-8975 (P.D. Wilcox); 0000-0003-1377-2694 (A.J. Croxford)

is particularly challenging to assess as it is where the largest imaging artefacts are formed from the largest structural feature echoes. A human operator could be trained to identify such artefacts so they are not wrongly identified as defects, but at the cost of lower detection sensitivity in the vicinity of artefacts and greater susceptibility to error.

Removing the artefacts has previously been addressed by masking parts of the image where artefacts are expected to arise based on spatial noise distributions [9]. This suppression is achieved at the cost of masking out areas where the defect may lie within the region of interest (ROI), hence reducing the ROI and potentially hiding defects. Alternatively, the suppression of artefacts can be addressed in the time-domain by identifying and removing the echoes stemming from structural boundaries. This alternative method has the potential to remove the artefacts from the images, while avoiding completely masking parts of the image. Therefore, to effectively increase the ROI and make the images easier to interpret, an artefact removal approach based on the time-domain is desirable. This could be done by adopting the inversion of a physics-based forward model based on parameters inferred from data, whereby the artefacts can be suppressed using their arrival times. However, despite the rigorousness and potential accuracy of inversion approaches, they are computationally expensive, requiring thousands of forward model evaluations [10–12]. Note that a forward model that reproduces the FMC data for a certain set of geometry and material related parameters is an essential part of such an inversion scheme.

Machine and deep learning (DL) approaches [13] can be effective tools to alleviate the computational burden of artefact identification using physics-based models. The applicable methods depend on the amount and complexity of the training data and range from Principal Component Analysis [14], for dimensionality reduction, to neural networks [15] that are able to capture highly non-linear patterns in the data. In the context of ultrasonic NDE, multiple DL methods have been recently adopted to improve the characterisation and identification of defects. For instance, in [16], the authors presented a hybrid temporal and spatial DL model for defect detection using thermography applying previously developed DL networks, such as U-net [17] and Segnet [18]. Convolutional neural networks (CNNs) were used in [19] for defect classification using the eddy current testing technique with high accuracy. Pyle *et al.* [20] provided a DL model based on CNN layers to accurately characterise the length and angle of cracks using data from ultrasonic inspections. These approaches aim at providing direct detection and classification of defects without explicitly dealing with artefacts that may partially obscure them, hence reducing the accuracy of such methods. To identify and suppress artefacts, the candidate DL model should have the ability to learn a reduced artefact-related representation of the input data and reconstruct from the reduced representation an output similar to the original input, but only with artefact-related information. In this context, *autoencoder* models [13] can be adopted due to their inherent structure, which is comprised of two sub-models: (1) an *encoder* that reduces the dimensionality of the input data into a latent representation, which could be forced to include the physical parameters of interest; and (2) a *decoder* that increases the latent dimensions up to the desired output dimensions, which typically are the same as those of the input data. These models have typically been used for denoising applications [21–23], for instance, to further increase the defect classification accuracy of a CNN network using ultrasonic NDT data [24]. Note that artefact removal is particularly attractive for NDE based on ultrasonic images, whereby structural artefacts will always be present and can hide information about defects.

This paper presents a newly developed DL approach based on the fundamentals of autoencoder models for artefact identification and suppression in NDE. The proposed methodology is generic and could be used for any NDT technique; here it is applied to ultrasonic array immersion inspection. A schematic of the generic artefact suppression framework is depicted in Figure 1. The proposed method is comprised of two parts: (1) an encoder that acts as an efficient surrogate inverse problem solver whereby physical parameters (e.g. specimen thickness or probe position) are obtained from the FMC data; and (2) a decoder that receives as input the encoder derived parameters and provides the times of arrival (ToAs) of the structural artefacts, hence acting as a surrogate forward model. Note that this method uses physics-based data and models to train data-based models, which confer the latter with a certain degree of transparency and understanding as opposed to unsupervised learning used for traditional autoencoders. Therefore, we refer to our DL model also as a *grey-box* model in contrast to fully data-driven, black-box, models. In the implementation presented in the current paper, the ToA information is used to suppress artefacts by applying masking windows in the original FMC data. Then, the masked FMC data, along with the inferred encoder generated parameters, are used for ultrasonic multi-view imaging obtaining (1) an effectively larger ROI due to the suppression of large and highly intense artefacts and (2) images that are easier to interpret for the NDE end-user. It is worth highlighting that the proposed DL-based framework is aimed at enhancing existing NDE procedures (through more accurate and easier data interpretation) rather than replacing them. This means that there is less disruption in the actual NDE procedures and regulations,

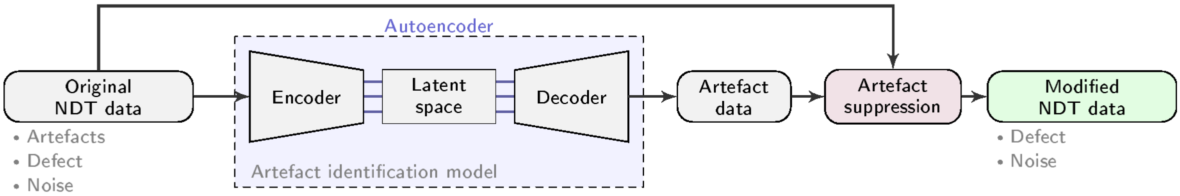


Figure 1: Generic DL-based artefact suppression process.

which could make the proposed approach attractive for near-term industrial implementation. Note also that the DL models are trained using defect-free data alone from simulations using a physics-based forward model.

This paper is organised as follows: Section 2 describes the physics- and data-based models along with the data sampling strategy and metrics used for quantifying the effectiveness of the proposed method; Section 3 illustrates the methodology through an experimental case study; a discussion is provided in Section 4 to assess the methodology against alternative techniques and to analyse the importance of the inference of the physical parameters by the encoder; finally, Section 5 provides concluding remarks.

2. Methodology

The proposed framework relies on model-based defect-free data for the training and validation of the encoder-decoder DL models. Therefore, the applied ultrasonic model and the multi-view TFM imaging are introduced below before describing the proposed models and their associated performance metrics.

2.1. Array imaging and ultrasonic modelling

The generation of one or more ultrasonic images is needed to visually interpret ultrasonic FMC data and to identify whether or not there is a defect in a specimen. To this end, the FMC data is firstly filtered in the frequency domain using a Hilbert transform and a Gaussian function centred at the excitation frequency. TFM images are then obtained as the summation of the time-delayed time-traces, as follows [8]:

$$I_i(\mathbf{r}) = \left| \sum_{T}^{N} \sum_{R}^{N} a_{TR}^{(i)}(\mathbf{r}) \tilde{x}_{TR}^{(i)}\left(\tau_{TR}(\mathbf{r})\right) \right| \tag{1}$$

where $I_i(\mathbf{r})$ is the image intensity of the i-th view at the position $\mathbf{r}$; $a_{TR}^{(i)}$ is an apodisation term, which is not considered in this work, and hence $a_{TR}^{(i)} = 1$; $\tilde{x}_{TR}^{(i)} = \mathcal{H}\left(x_{TR}^{(i)}\right)$ is the Hilbert-transformed FMC data [25]; and N denotes the number of elements in the array. The subscripts T and R refers to the transmitter and receiver elements, respectively. The time delay τ_{TR} is obtained as a function of the travel time between the transmitter T, the point in the image $\mathbf{r}$, and the receiver R. A Lanczos interpolation is used to interpolate the discrete FMC values to the delays given in Equation (1). Different views are generated by considering different transmit and receive ray paths.

The amount of information available in the multi-views for a single FMC dataset is considerable, potentially enabling better defect detection and characterisation. However, the multiple internal reflections also lead to structural echoes that are reconstructed as image artefacts in the TFM views. In this context, removing these artefacts from the FMC data could potentially enable improved performance. Note that the different ray paths are also differentiated between longitudinal (L) and transverse (T) wave modes [9, 26]. The combination of the transmitting and receiving paths associated with these modes will provide the terminology used hereinafter for the different TFM views. For example, LT-T denotes a half skip view with the transmitting path containing two segments (LT) inside the specimen, the first is a L mode from the frontwall to the backwall and the second is the T mode from the backwall to the focus point. The returning path contains only one segment of T mode. Note that the chosen notation does not include any water paths, given that only longitudinal mode is possible within water and it is assumed to be implicit.

The identification of artefacts from the FMC data is addressed using defect-free modelled data. To this end, a multi-frequency ray-based model [26] for ultrasonic array immersion, named here as ARIM[1], is used to obtain the

[1]The ARIM model is publicly available to download from `https://github.com/ndtatbristol/arim`.

FMC time-trace data along with its associated ToAs for each echo (or ray-path) under the assumption of different input parameters (e.g. probe locations and the specimen's material properties). The model considers directivity, transmission and reflection coefficients, and beamspread along with attenuation.

2.2. Deep learning models for artefact identification and suppression

DL models are powerful and computationally efficient tools [13] that offer a high degree of flexibility when addressing complex data. Out of the many available DL models, autoencoders hold a unique structure whereby some information is firstly encoded in an intermediate latent representation of the input data, hence reducing its dimensionality. This information is then decoded into a similar structure to the input data. Autoencoders are typically used for segmentation or denoising purposes given that the latent variables are trained to extract only adequate information from the input data for the ultimate objective of segmenting or denoising. The proposed model architecture was arrived at after comparing multiple model architectures, including overcomplete and undercomplete autoencoders [27].

2.2.1. Grey-box model: encoder-decoder

A grey-box model consists of a blended approach that uses data-driven models which incorporate physics-based information, which provides a certain degree of transparency to the DL model. The ultimate goal of this method is to mask out the regions of the time-traces that correspond to the structural echoes of the specimen under inspection. Based on the autoencoders' structure, the proposed grey-box model for artefact identification and suppression for ultrasonic inspections is comprised of two hierarchical parts: (1) an encoder that provides physical parameters from the FMC data; and (2) a decoder that uses the physical parameters from the encoder to provide the ToAs of the frontwall and backwall echoes. Note that the masking windows are created around the ToA of each echo.

A subset of time-traces $\mathcal{D} \in \mathbb{R}^{N' \times N' \times N_\tau}$ is chosen in the form of a smaller sub-array with $N' < N$ elements of the larger ultrasonic probe instead of the complete FMC dataset in order to train the encoder model. This simplification is done: (1) because the artefact ToAs are smoothly-varying functions across the FMC data; and (2) to avoid memory issues with the available GPUs given that each time-trace contains thousands of individual points, N_τ, and the memory consumption scales up relatively quickly with larger databases. The smaller array is set to 8 transmit-receive elements, which are selected so as to be evenly spaced within the larger array of the ultrasonic probe. This is equivalent to having a probe with a longer pitch between these 8 elements. The training and validation data, which contain the envelopes of the time-traces, the physical parameters, and the ToA of the boundary echoes, are linearly normalised so that the values of the envelopes stay between [0, 1], while the physical parameters and ToA stay within the interval [0, 1].

The encoder model $f(\mathcal{D})$ is then defined as a relationship between envelopes of the time-traces in the reduced FMC dataset $\mathcal{D}$ and a set of unknown physical parameters $\boldsymbol{\theta} \in \mathbb{R}^{N_\theta}$, which is assumed to be comprised of $N_\theta = 3$ parameters, i.e. the specimen thickness ($\delta \in \mathbb{R}$), the probe angle ($\alpha \in \mathbb{R}$), and the probe stand-off ($Z_s \in \mathbb{R}$), so $f(\mathcal{D}) : \mathbb{R}^{N' \times N' \times N_\tau} \to \mathbb{R}^{N_\theta}$. The model structure, which is summarised in Figure 2, is loosely based on volumetric segmentation architectures such as the 3D U-net [28] and is comprised of 3D convolutional layers that extract information from the time-traces. Note that 2D convolutional layers could have been used after the third layer instead of 3D ones, but for the sake of generality and simplicity, 3D layers have been used throughout with negligible impact on the computational efficiency. Then, dense layers are used to predict the physical parameters $\boldsymbol{\theta}$. Note also that for training purposes of this encoder model, and to make it more robust and realistic, the input time-traces not only contain structural artefacts, but also Gaussian noise. This is to account for further variability in experimental scenarios. The activation function used throughout the internal layers is the rectified linear unit (ReLu) [13], while the output layers have linear activation functions. As the encoder model is effectively a regression model, the mean squared error (MSE) is used as loss function [13] for each of the three parameters obtained as output.

The decoder model $g(\boldsymbol{\theta})$ is defined as a relationship between the physical parameters $\boldsymbol{\theta}$ and the set of ToAs for each of the frontwall and backwall echoes $\mathcal{T}$, so $g(\boldsymbol{\theta}) : \mathbb{R}^{N_\theta} \to \mathbb{R}^{N' \times N' \times N_\mathcal{T}}$. Note that this DL model is a metamodel (or surrogate model) of the physics-based one described in Section 2.1. In this case, the model is conceived to scale up the dimensionality of the input. Three layers are used: a dense layer to create a larger number of internal data, a 3D convolutional layer, and a dense layer in the output layer to give the prediction of the ToAs. The model is summarised in Figure 3. Note also that the training data is not altered by any type of further variability such as noise, given that the input data are the physical parameters $\boldsymbol{\theta}$. Again, the loss function is calculated using the MSE.

Once both models are independently trained using a scheduled-based learning rate adopted from [29] and the ADAM optimiser [30], they are assembled to work sequentially: firstly encoding the time-traces to the physical

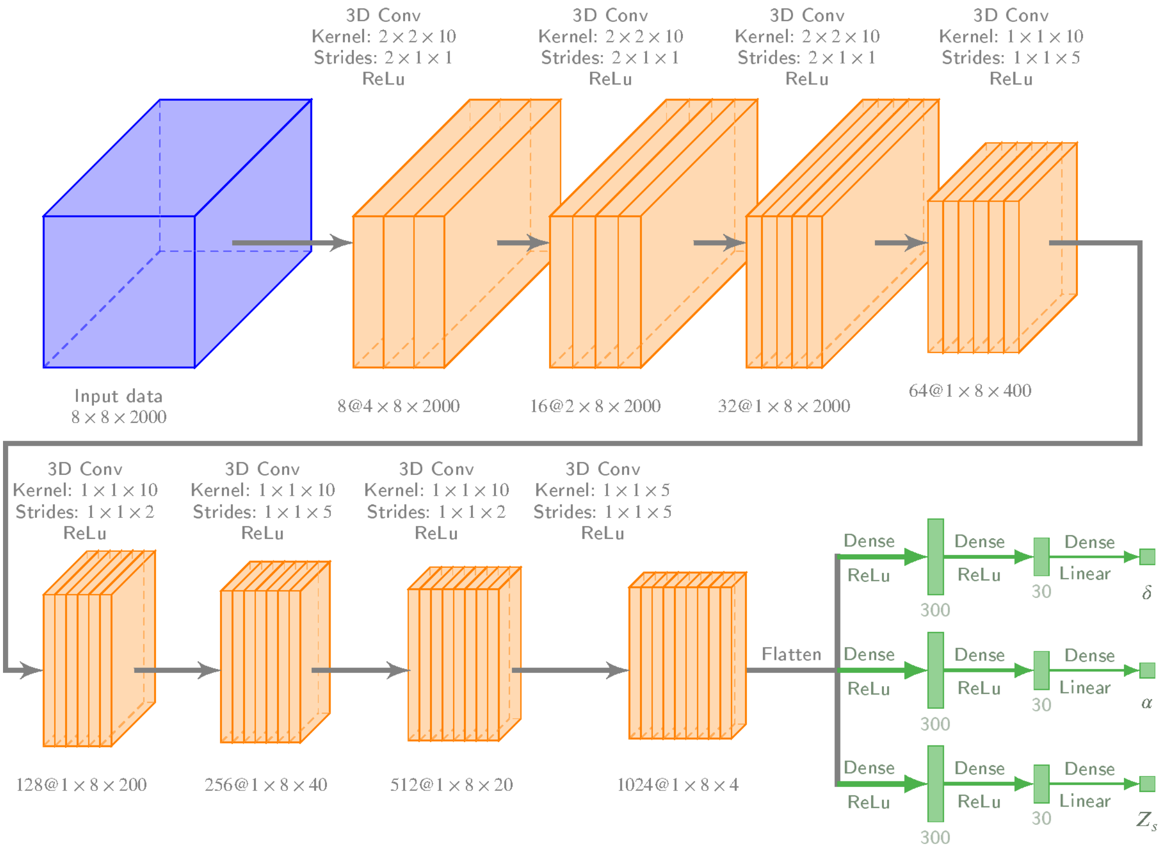


Figure 2: Schematic diagram of the encoder model structure. The number of filters used in each convolutional layer is written before the "@" symbol below each data structure. Orange structures are associated with convolutional layers, while green structures are associated with dense layers.

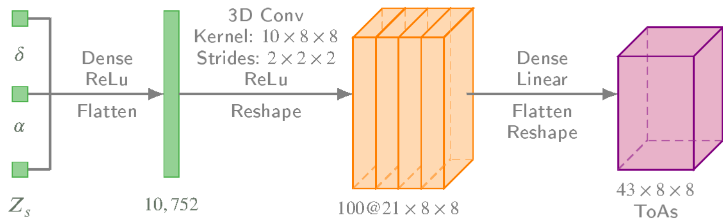


Figure 3: Schematic diagram of the decoder model structure.

parameters and secondly decoding the physical parameters to obtain the ToAs corresponding to the artefacts. Note, however, that the output data is for a reduced set of the time-traces of the larger FMC dataset, so they need to be up-scaled. To this end, the remaining ToAs related to each of the transmit-receive elements other than the 8 used for training are linearly interpolated (for the elements between the 8 chosen for training) and linearly extrapolated if needed (for elements that may be lying outside the reduced set). It is worth highlighting, that this interpolation will introduce a small error between the true and predicted ToA for the intermediate transmit-receive elements of the probe, but it is assumed to be negligible compared to other sources of variability when processing experimental data. Most sources of uncertainty or variability in experimental data are accommodated in the masking stage described below.

The ToA information output by the autoencoder for the complete FMC dataset is then used to produce masking windows in the time-domain. Around each predicted ToA, a masking window with a fixed length L_w, which is related to the number of cycles N_{cyc} and the period of the excitation signal $T_s = 1/f$, is defined as follows:

$L_w = B \cdot N_{\mathrm{cyc}} \cdot T_s$, where B is a constant that is used to accommodate experimental variability and potential model errors. It is worth highlighting that the mask around the frontwall echo needs an additional width ΔL_w to consider the echo's tail stemming from near-field interactions that create additional artefacts in experimental scenarios. A figure of an experimental frontwall echo is provided in the supplemental material to support this assumption. The complete workflow of the proposed artefact suppression method from training to application stages is depicted in Figure 4.

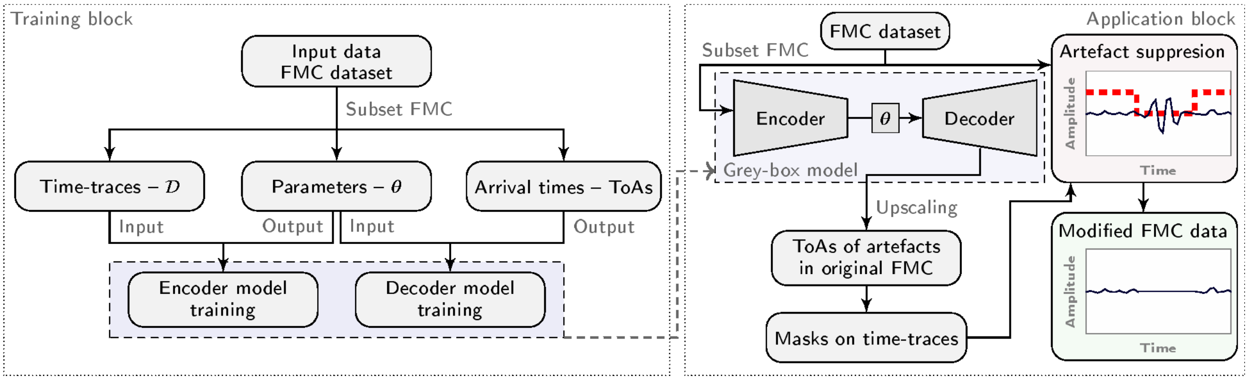


Figure 4: Flowchart explaining the workflow of the encoder-decoder approach.

2.2.2. *Training and validation database*

The training and validation databases used to create the DL models are generated using the ARIM model described in Section 2.1. To this end, the upper and lower bounds of the training space of the set of unknown physical parameters $\boldsymbol{\theta} = \{\delta, \alpha, Z_s\}$ are defined for the desired inspection ranges. Then, a structured 3D grid is created with a step size of $\Delta\boldsymbol{\theta} = \{\Delta\delta, \Delta\alpha, \Delta Z_s\}$. A random perturbation is added to the grid in order to create a more generic training space that eliminates any bias that the structured grid may produce. The uniformly distributed random perturbation is added to each grid point as follows:

$$\boldsymbol{\theta}^{(k)}_{\mathrm{rand}} = \boldsymbol{\theta}^{(k)} + \Xi^{(k)} \quad \text{with} \quad \Xi^{(k)} \sim \mathcal{U}\left(\boldsymbol{\theta}^{(k)} - \Delta\boldsymbol{\theta}/2,\ \boldsymbol{\theta}^{(k)} + \Delta\boldsymbol{\theta}/2\right) \tag{2}$$

where $\boldsymbol{\theta}^{(k)}_{\mathrm{rand}}$ and $\boldsymbol{\theta}^{(k)}$ denote the k-th randomly perturbed and structured grid points, respectively; and $\Xi^{(k)}$ is a perturbation sample extracted from the uniform distribution denoted as $\mathcal{U}(\cdot)$. Two entirely separate subsets are created from the randomly perturbed grid, a training set and a validation one with a relation 70:30. The ARIM model is then evaluated at each grid point and two entities are obtained as output: the FMC time-traces and the ToAs for each of the artefacts.

2.3. Performance metrics

The effectiveness of the proposed masking method is quantified through a series of metrics both in the time and image domains. In the case of the time domain, the relative amplitudes of the signals that pass through the masks (with respect to the total amplitudes) are calculated by using an integral over the time-traces in all three dimensions of the FMC dataset, i.e. the time axis and the rows and columns from the number of emit-receive elements. Mathematically, this time-domain based metric is defined in a decibel scale, as follows:

$$E_t = 20\log_{10}\left(\frac{E_t^M}{E_t^o}\right) \tag{3}$$

where E_t^o and E_t^M are the integrals of the amplitudes of the original and masked time-traces, respectively. These integrals are defined as:

$$E_t^o = \sum_T^N \sum_R^N \int_{t_i}^{t_e} \left|\mathcal{H}\left(x_{TR,o}(t,e,r)\right)\right| dt\, dr\, de \tag{4}$$

$$E_t^M = \sum_T^N \sum_R^N \int_{t_i}^{t_e} \left| \mathcal{H}\left(x_{TR,M}(t,e,r)\right) \right| dt\, dr\, de \tag{5}$$

where $x_{TR,o}(\cdot)$ and $x_{TR,M}(\cdot)$ are the original and masked time-traces, which are dependent on the time point t, the receiver r, and the emitter e. The envelopes of the signals are calculated by using the Hilbert transform, as follows: $|x(t)+i\cdot\mathcal{H}(x(t))|$. The minimisation of the time-domain amplitude metric E_t (Eq. (3)) is adopted herein for establishing the optimal width factor B of the masking windows described in Section 2.2.1.

3. Case study

The proposed framework for artefact identification and suppression is tested and demonstrated in an experimental case study on ultrasonic array imaging.

3.1. Experimental results

The experimental results are obtained using a 5 MHz, linear, 1D array with $N = 64$ elements and pitch of 0.63 mm. The density and longitudinal velocity of the immersed medium (water) are assumed to be known and with values $\rho_w = 1000$ kg/m^3 and $v_w = 1480$ m/s. The material under inspection is plain aluminium with a 0.5mm sawcut made perpendicular to the backwall to represent a surface-breaking crack, density $\rho = 2700$ kg/m^3 and longitudinal and transverse velocities are $v_L = 6300$ m/s and $v_T = 3130$ m/s, respectively. The measured thickness is $\delta = 30$ mm, while the ultrasonic probe is setup at a standoff of $Z_s = 35$ mm (measured from the centre of the array) and an angle of $\alpha = 15°$. A generic immersion test configuration is illustrated in Figure 5.

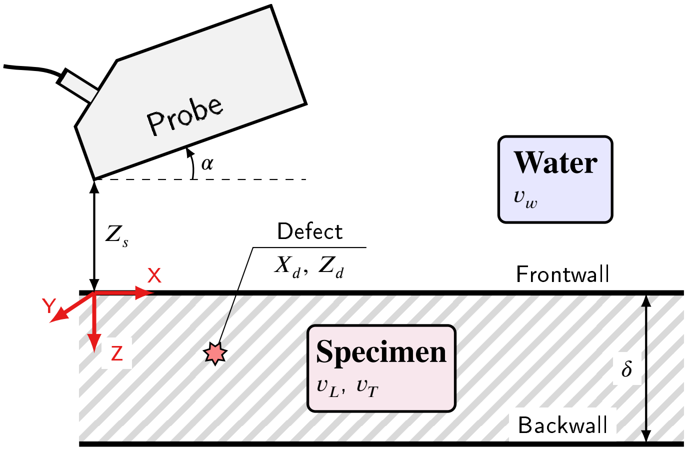


Figure 5: Set-up configuration for immersed ultrasonic inspection.

The proposed DL models have been trained over the following parameter space: $\delta \in [27, 32]$mm $\subset \mathbb{R}$, $\alpha \in [12, 19]° \subset \mathbb{R}$, and $Z_s \in [32, 37]$mm $\subset \mathbb{R}$. Note that the parameters intervals are chosen so that the expected variability in the specimen thickness and probe configuration is fully covered. The models could be trained over a larger parameter space, but in this case the limits are selected to be optimised for a specific inspection. The number of samples used for the training space is 4913 (corresponding to a 3D grid of $17 \times 17 \times 17$). The ARIM forward model is evaluated using the set of 4913 input parameters to estimate both the FMC data along with the 43 ToAs of the chosen ray paths. The training and validation losses for each of the models are given in the supplemental material. The encoder model shows overfitting after around 20 epochs, whereby the training losses start diverging from the validation ones and only training data fit gets better while the network accuracy does not improve on unseen data. The network parameters at the point of minimum validation losses are taken as the optimum ones for subsequent application. On the other hand, the decoder appears to still be learning from the data at 1000 epochs, but at a relatively low rate. Additionally, the level of MSE loss is very small (10^{-8}), which makes it sufficiently accurate for the purpose of the artefacts' ToAs ($\mathcal{T}$) prediction.

3.1.1. Artefact suppression results

The encoder is applied to a subset of 8 time-traces (evenly spaced each 9 elements) of the experimental FMC dataset. Note that the time traces are truncated at the time required for the furthest image point, which is 9.9e-05s for the point located at X = 80mm and Z = 30mm. Additionally, a tuning parameter affecting the scale of the input time-traces is applied to help the encoder infer the physical parameters $\boldsymbol{\theta}$. This tuning parameter, which partially accounts for normalised amplitude discrepancies between modelled and experimental time-traces, is chosen by solving a minimisation problem on the metric of the masked signals E_t with respect to the scaling parameter. The optimal tuning parameter value results to be 2.00 for the experimental dataset evaluated. The inferred physical parameters from the encoder are: $\delta = 29.642$mm, $\alpha = 14.899°$, and $Z_s = 35.048$mm. These values have also relatively small deviation with respect to the measured values of the simulation of $\Delta\delta = 1.19\%$, $\Delta\alpha = 0.68\%$, $\Delta Z_s = 0.14\%$. Note that although the thickness shows a larger variation, this may stem partially from measurement errors of the specimen or a biased value of the longitudinal and transverse velocities of the aluminium material.

The width factor of the masking windows (B) is chosen after performing a parametric study on the impact of the amplitude suppression of the FMC dataset with respect to B. The results, depicted in Figure 6 for both time and image domain metrics, show a strong suppression of amplitudes from $B \geq 1.2$. This is also evident from the numerical derivatives of the metric, i.e. $\partial E_t/\partial B$ and $\partial E_i/\partial B$, represented by an orange line. The width factor chosen for the suppression of artefacts and generation of TFM views is $B = 1.6$. The additional width (ΔL_w) given to the frontwall echo's masking window (L_w^{FW}) is $\Delta L_w = 1.8\ \mu$s.

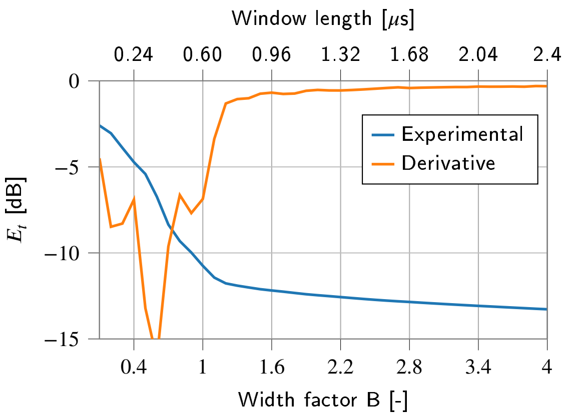


Figure 6: Time-domain metric with regards to the width of the masking windows along with its numerical derivative.

The time-traces are then masked with the ToAs ($\mathcal{T}$) predicted by the decoder using, in turn, the parameters provided by the encoder and a width factor of $B = 1.6$. These are represented in Figure 7 for the signals acquired by all receivers when the first emitter is actuating. The windows are masking the largest amplitudes of the time-traces (coloured regions), which represent both frontwall and backwall echoes. The windows are remarkably well centred around the larger amplitudes corresponding to the structural artefacts, highlighting the effectiveness of the ToA interpolation between the subset and complete FMC data. Observe that there are four dead elements in the ultrasonic probe out of the 64. Nevertheless, these have had a minimal impact on the performance of the grey-box approach given that they are not present in the subset of 8 array elements used for training and evaluation of the DL models. Although for a specific transmitter, these results are typical.

The effectiveness of the approach is further illustrated after imaging using the masked FMC dataset, as shown in Figure 8. Both L-L and LL-L views show no clear indication of any defect in the original unmasked dataset. However, after applying the encoder-decoder models along with the masking windows, the structural artefacts almost disappear, making the defect emerge. The suppression levels in the L-L view are 29 dB and 35 dB for the frontwall and backwall artefacts, respectively, and in the LL-L view 57 dB on average. This makes the defect easily identifiable, even when it overlaps with the structural artefacts. Note also that the physical parameters predicted by the encoder are used to produce the images, showing a remarkably good alignment of the artefacts (in unmasked image) to the axes of the image considering that the aluminium specimen is flat.

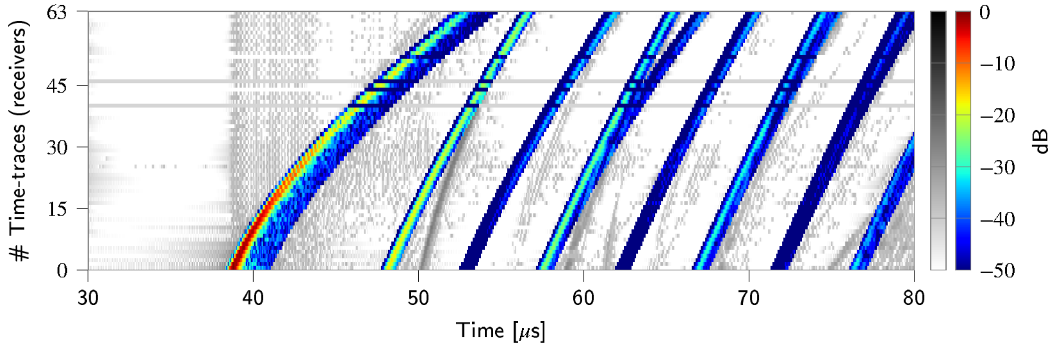
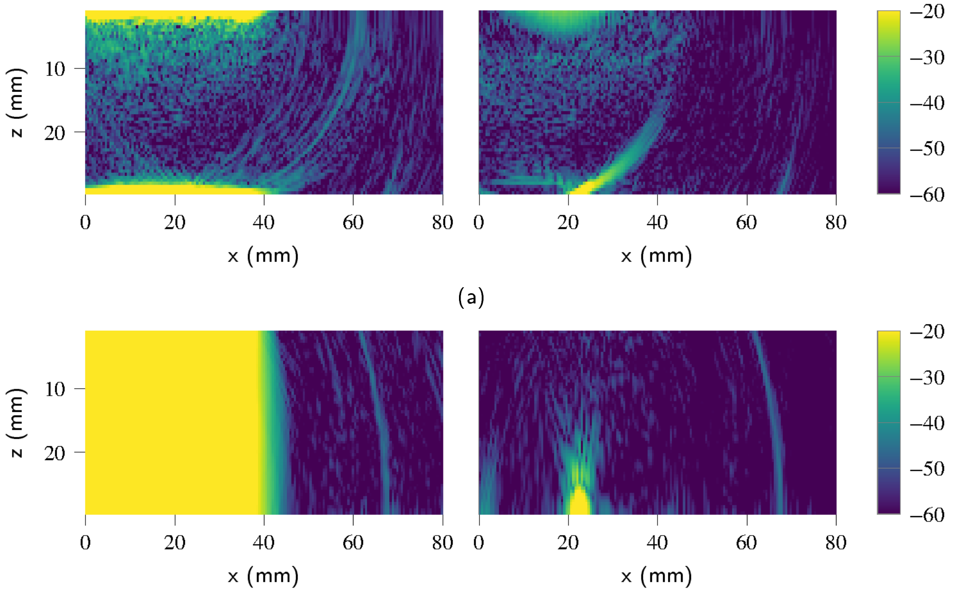


Figure 7: Envelopes of the time-traces and masks (coloured regions) predicted by the decoder using a length factor of 1.6 and the input parameters provided by the encoder for the emitter #1. Reference of the dB is the maximum value present in the unmasked data.



(a)

(b)

Figure 8: TFM views, i.e. L-L in panel (a) and LL-L in panel (b), for the experimental data with defect before and after applying masks in the left and right columns, respectively. Reference of dB at backwall in the L-L view of the original data containing artefacts and defect.

To further illustrate the performance of the proposed grey-box approach, the fusion of the TFM views by averaging using the masked and complete FMC is addressed, as it includes all 21 views up to TT-TT with many artefacts [31]. Note that the image fusion by averaging is in general a simplistic and suboptimal way of addressing the data fusion in ultrasonic imaging [31]; however, for the purpose of illustrating the effectiveness of the DL-based masking methodology it is appropriate. Figure 9 demonstrates that the artefact suppression is capable of almost removing the frontwall and backwall artefacts while letting the defect through. As a result, the defect can be very well identified even when it is directly below the footprint of the ultrasonic array. Observe also that a significant part of noise present in the original fusion image has also been suppressed after applying the masks.

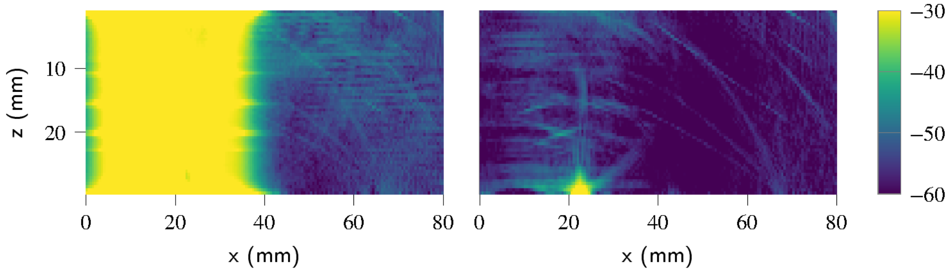


Figure 9: Fusion of TFM views by algebraic averaging for experimental data with defect before and after masks in the left and right panels, respectively. Reference of dB at backwall in the L-L view of the original data containing artefacts and defect.

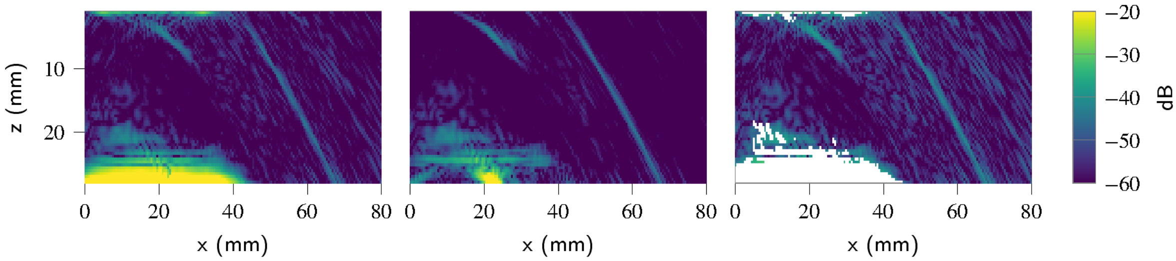
4. Discussion

The proposed encoder-decoder approach for artefact identification and suppression in ultrasonic NDE has been demonstrated in an experimental case study. The combination of a ray tracing physics-based model and DL models has resulted in a highly efficient grey-box approach, whereby data-driven models are trained in such a manner so the physics of the problem are better understood and controlled than in a purely black-box model. The first part of the proposed approach, i.e. the encoder model, is able to provide the physical parameters from the FMC data that drive the imaging and identification process: thickness, probe angle, and probe standoff. These parameters are also used as input to the second part of the framework, i.e. the decoder model. This model provides a set of ToAs of the artefacts (frontwall and backwall echoes) that are used for creating the masking windows. The masked FMC dataset can then be used for defect identification purposes providing more useful information below the probe's footprint, which effectively produces an increase of the ROI in an ultrasonic inspection. The larger ROI also allows a reduction in the number of sweeps needed to inspect a predefined area. As a key point, the proposed framework enables established techniques to be used for defect detection (e.g. thresholding [32]) and characterisation (e.g. -6dB sizing [33]) with enhanced performance over a larger ROI, rather than substituting these techniques completely. Additionally, the methodology has successfully dealt with experimental data after little manipulation (scaling and choosing the first half). This highlights the accuracy and robustness against noise of the proposed grey-box approach in real-world engineering scenarios.

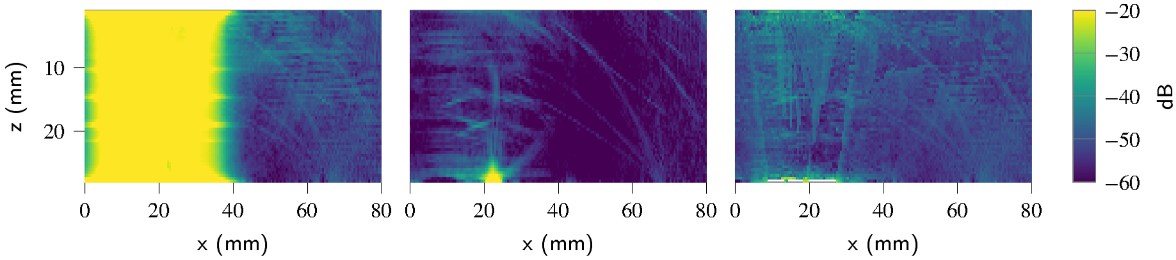
4.1. On the comparison with other techniques

The proposed approach is aimed at removing the artefacts in the time-domain, however there are alternative masking methods that could work on either time or image domains. To comparatively assess our framework with respect to others, the proposed grey-box approach is compared against a recently published method that removes the artefacts in the TFM views [9]. Figure 10 shows an example of both techniques applied to the experimental results obtained in Section 3.1. It is noticeable that there are no areas with zero contribution from any time-trace in the proposed grey-box approach, which makes it possible to identify defects that are coincident with artefacts. This can be appreciated in the individual views (Fig. 10a), where the masks in the image-domain significantly reduce the amount of information while obscuring the defect. The fusion of the views (Fig. 10b) also shows that a more uniform portion of amplitude is let through when removing the artefacts in the time-domain as opposed to the image masking approach, which suppresses a significant amount of defect information. Additionally, the amplitudes that the masks let through are more homogeneous in the whole region using the proposed grey-box approach than using the image-based one, as observed in the amplitude ratio shown in Figure 10c. Overall, the proposed grey-box method provide more usable and consistent results throughout the entire view, making it possible to enlarge the ROI from the right side of the probe to both directly below and on the right of the ultrasonic array.

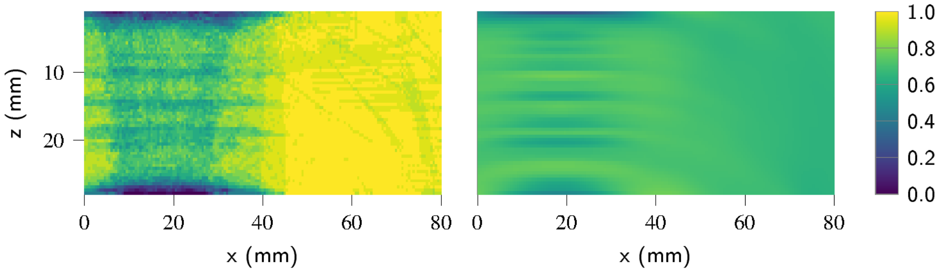
An additional advantage of using the DL-based approach is computational efficiency. The image-based artefact removal approach [9] runs in the order of magnitude of minutes as the experimental parameters are obtained step by step from the raw data and then ARIM is run. Alternatively, the proposed grey-box approach only takes fractions of a second to compute both the encoder and decoder models using the enhanced computational performance provided by GPUs.



(a) LT-L view (Original - Grey-box approach - Image-based method)



(b) Fusion (Original - Grey-box approach - Image-based method)



(c) Average amplitude ratio that passes through the masks (Image-based approach - Grey-box approach)

Figure 10: Masking comparison between the image-based approach and the time-domain one using the grey-box model. Panel (a) shows the LT-L view, panel (b) shows the data fusion, and panel (c) shows the amplitude ratio passing through the masks in the fusion.

4.2. On the parameter inference

In addition to the identification and suppression of artefacts, the proposed method is capable of dealing with certain misalignment or biased parameters used for imaging. This occurs, for instance, if using a set of parameter values (thickness, probe angle and/or standoff) for imaging that are incorrectly determined due to experimental errors. Note that, in order to do the same with the image-based masking approach discussed above, the masks may need to be dilated to accommodate variability in all parameters, hence reducing the ROI. Moreover, images would be poorly focused as the parameters cannot be determined on defect data. If the proposed grey-box approach is used, the masking tracks the variability in the specified physical parameters, so the information loss is less (in addition to obtaining the parameters). To further illustrate this case, Figure 11 represents the L-L view in three additional experimental measurements in which the probe angle and standoff slightly vary from the originally measured parameters described in Section 3.1 ($\delta = 30$ mm, $\alpha = 15°$, $Z_s = 35$ mm): (1) for a smaller angle of $\alpha = 14°$; (2) for a longer standoff of $Z_s = 36$ mm; and (3) for both higher angle and longer standoff of $\alpha = 16°$ and $Z_s = 36$ mm respectively. Observe in Figure 11 that misalignment in the setup (e.g. due to measurement errors) produces a strong deviation of the theoretical position of the frontwall and backwall artefacts. The corrected views are generated with the original complete FMC data and the encoder-predicted parameters, and therefore the artefacts are visible in the views. This highlights the flexibility and greater usability of the proposed approach in comparison with other masking approaches, where no physical information would have been obtained. Besides, the inference of the physical parameters potentially enables the use of less precise and probably cheaper physical rigs, further reducing costs.

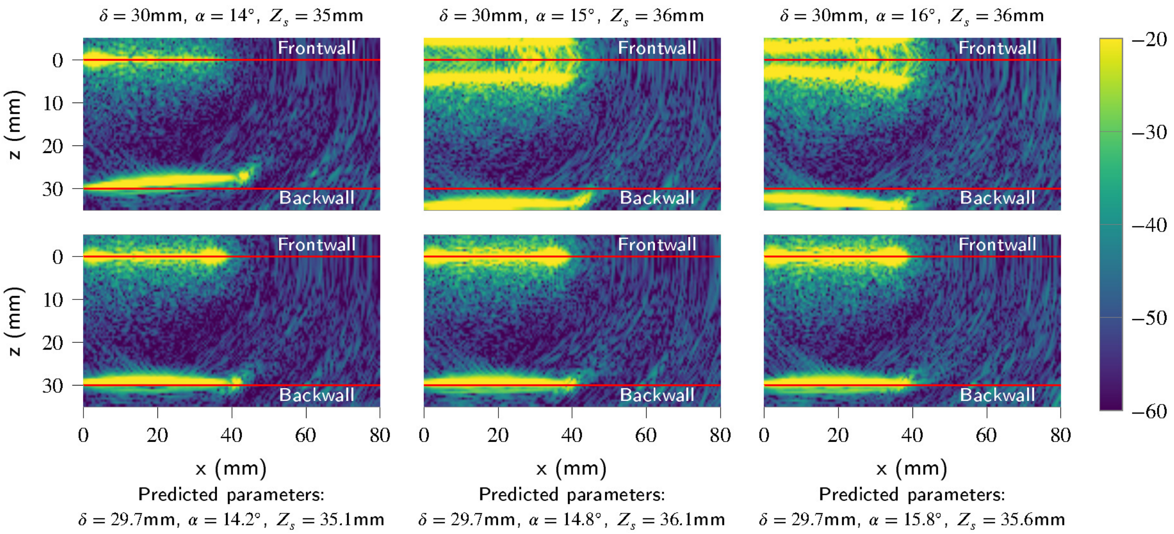


Figure 11: Comparison of the L-L view using the originally measured parameters ($\delta = 30$ mm, $\alpha = 15°$, $Z_s = 35$ mm) in first row and the images using the predicted parameters (shown below the x-labels) by the encoder in the second row.

The robustness of the encoder model in providing consistent and accurate predictions of the model parameters can also be found on experimental measurements. Note that additional measurements were made on the same aluminium specimen described in Section 3.1 to produce Figure 11. The performance of the encoder in inferring the parameters $\boldsymbol{\theta}$ has been remarkably consistent, with a relatively small variation from the whole set of FMC measurements, as can be observed in Table 1. Nonetheless, there are still physical parameters that are assumed, such as the propagation velocities, that would be useful to infer along with their associated uncertainties from the FMC along with the thickness and relative probe location. Note that this scenario would lead to an ill-posed inverse problem where there could be more than one viable solution of distances and velocities. Therefore, a desirable extension of the proposed method is to account for more unknown model parameters, e.g., propagation velocities, while quantifying their uncertainty. To do this without compromising the computational efficiency, several methods could be adopted, such as probabilistic DL models [34] or the adoption of Bayesian inverse problem for inferring the model parameters from the data through an approximation, e.g. using the approximate Bayesian computation algorithm [35].

Table 1
Predicted parameters from multiple experimental FMC measurements.

Measurement No	Thickness [mm]	Angle [°]	Stand-off [mm]
#1	29.642	14.898	35.048
#2	29.526	14.976	35.053
#3	29.532	14.930	35.047
#4	29.580	14.909	35.090
#5	29.598	14.886	35.127
#6	29.586	14.909	35.124
#7	29.518	14.886	35.016
#8	29.544	14.907	35.025
#9	29.606	14.886	35.070
#10	29.585	14.911	35.073
Mean	29.572	14.910	35.067
Standard deviation	0.0381	0.0255	0.0357

The results shown in Table 1 highlight the aforementioned effectiveness and robustness of the proposed method in real-world scenarios. The application of this method in industrial environments is a desirable extension of the proposed method, whereby additional sources of variability and larger noise can be evaluated in the proposed method. Nonetheless, they grey-box approach has proven to be an effective tool to remove artefacts, enlarge the ROI, while make it easier to identify defects in challenging scenarios.

5. Conclusions

A DL approach based on the fundamentals of autoencoder models have proposed in this paper to tackle the artefact identification and suppression in NDT data. The approach, that is general, has been implemented for ultrasonic array immersion inspection. The proposed approach consists of three parts: (1) an encoder model that infers physical parameters from the FMC data; (2) a decoder model that predicts the arrival time of the artefacts; and (3) the application of masking windows to the input FMC data. The resulting data is then used for ultrasonic imaging for defect identification. An experimental case study with FMC data has been used to illustrate the methodology. The following conclusions can be drawn:

- The proposed method identifies and suppresses the artefacts contained in NDT data in a very efficient and accurate manner.
- The grey-box approach accurately also provides useful physical parameters which can be used for ultrasonic imaging.
- The encoder demonstrates a remarkable consistency and robustness when dealing with experimental data.
- The methodology enhances the interpretation of NDT data by making it easier through the suppression of artefacts from the raw data.
- While the grey-box methodology suppresses artefacts and extracts physical parameters from the raw data, the subsequent imaging process is standard, and hence established methods may be used for defect detection and characterisation; in this sense, the process is anticipated to be more straightforward to certify than a black-box approach.

Future works are under consideration on: (1) the inference of additional physical parameters, e.g. wave velocities, using the encoder; (2) the quantification of uncertainty during the inference of physical parameters; and (3) the application of the proposed artefact suppression procedure to alternative NDE modalities (e.g. ultrasonic guided-waves or X-ray imaging) and to data collected in industrial environments.

Acknowledgements

The work reported is part of a pilot project (Grant number 100374) funded by Lloyd's Register Foundation and the Alan Turing Institute Data-Centric Engineering Programme. This work was carried out using the computational facilities of the Advanced Computing Research Centre, University of Bristol – `http://www.bris.ac.uk/acrc/`.

CRediT authorship contribution statement

Sergio Cantero-Chinchilla: Conceptualization, Methodology, Software, Investigation, Validation, Formal analysis, Visualization, Writing - Original Draft. **Paul D. Wilcox:** Conceptualization, Data curation, Investigation, Writing - Review & Editing, Resources, Project administration. **Anthony J. Croxford:** Conceptualization, Data curation, Investigation, Writing - Review & Editing, Supervision.

References

[1] S. Gholizadeh, A review of non-destructive testing methods of composite materials, Procedia Structural Integrity 1 (2016) 50–57.
[2] A. Sophian, G. Tian, D. Taylor, J. Rudlin, Electromagnetic and eddy current NDT: a review, Insight 43 (2001) 302–306.
[3] J. Blitz, G. Simpson, Ultrasonic methods of non-destructive testing, volume 2, Springer Science & Business Media, 1995.

[4] C. Holmes, B. W. Drinkwater, P. D. Wilcox, Post-processing of the full matrix of ultrasonic transmit–receive array data for non-destructive evaluation, NDT & E International 38 (2005) 701–711.
[5] A. J. Hunter, B. W. Drinkwater, P. D. Wilcox, The wavenumber algorithm for full-matrix imaging using an ultrasonic array, IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control 55 (2008) 2450–2462.
[6] J. Zhang, B. W. Drinkwater, P. D. Wilcox, Efficient immersion imaging of components with nonplanar surfaces, IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control 61 (2014) 1284–1295.
[7] J. Zhang, T. Barber, A. Nixon, P. Wilcox, Investigation into distinguishing between small volumetric and crack-like defects using multi-view total focusing method images, in: AIP Conference Proceedings, volume 1806, AIP Publishing LLC, 2017, p. 040003.
[8] J. Zhang, B. W. Drinkwater, P. D. Wilcox, A. J. Hunter, Defect detection using ultrasonic arrays: The multi-mode total focusing method, NDT & E International 43 (2010) 123–133.
[9] R. L. Bevan, J. Zhang, N. Budyn, A. J. Croxford, P. D. Wilcox, Experimental quantification of noise in linear ultrasonic imaging, IEEE transactions on ultrasonics, ferroelectrics, and frequency control 66 (2019) 79–90.
[10] J. Chiachío, M. Chiachío, S. Sankararaman, Bayesian Inverse Problems: Fundamentals and Engineering Applications, CRC Press, Taylor & Francis Group, Boca Raton, FL, USA, 2021.
[11] J. Chiachío, N. Bochud, M. Chiachío, S. Cantero, G. Rus, A multilevel Bayesian method for ultrasound-based damage identification in composite laminates, Mechanical Systems and Signal Processing 88 (2017) 462–477.
[12] S. Cantero-Chinchilla, J. Chiachío, M. Chiachío, D. Chronopoulos, A. Jones, A robust Bayesian methodology for damage localization in plate-like structures using ultrasonic guided-waves, Mechanical Systems and Signal Processing 122 (2019) 192–205.
[13] I. Goodfellow, Y. Bengio, A. Courville, Y. Bengio, Deep learning, volume 1, MIT Press Cambridge, 2016.
[14] H. Abdi, L. J. Williams, Principal Component Analysis, Wiley Interdisciplinary Reviews: Computational Statistics 2 (2010) 433–459.
[15] J. Schmidhuber, Deep learning in neural networks: An overview, Neural Networks 61 (2015) 85–117.
[16] Q. Luo, B. Gao, W. L. Woo, Y. Yang, Temporal and spatial deep learning network for infrared thermal defect detection, NDT & E International 108 (2019) 102164.
[17] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 234–241.
[18] V. Badrinarayanan, A. Kendall, R. Cipolla, Segnet: A deep convolutional encoder-decoder architecture for image segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence 39 (2017) 2481–2495.
[19] P. Zhu, Y. Cheng, P. Banerjee, A. Tamburrino, Y. Deng, A novel machine learning model for eddy current testing with uncertainty, NDT & E International 101 (2019) 104–112.
[20] R. J. Pyle, R. L. Bevan, R. R. Hughes, R. K. Rachev, A. A. S. Ali, P. D. Wilcox, Deep Learning for Ultrasonic Crack Characterization in NDE, IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control (2020).
[21] F. Gao, B. Li, L. Chen, X. Wei, Z. Shang, C. He, Ultrasonic signal denoising based on autoencoder, Review of Scientific Instruments 91 (2020) 045104.
[22] D. Perdios, A. Besson, M. Arditi, J.-P. Thiran, A deep learning approach to ultrasound image recovery, in: 2017 IEEE International Ultrasonics Symposium (IUS), Ieee, 2017, pp. 1–4.
[23] B. Li, K. Xu, D. Feng, H. Mi, H. Wang, J. Zhu, Denoising convolutional autoencoder based B-mode ultrasound tongue image feature extraction, in: ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2019, pp. 7130–7134.
[24] N. Munir, J. Park, H.-J. Kim, S.-J. Song, S.-S. Kang, Performance enhancement of convolutional neural network for ultrasonic flaw classification by adopting autoencoder, NDT & E International 111 (2020) 102218.
[25] A. V. Oppenheim, J. R. Buck, R. W. Schafer, Discrete-time signal processing, volume 2, Upper Saddle River, NJ: Prentice Hall, 2001.
[26] N. Budyn, R. L. Bevan, J. Zhang, A. J. Croxford, P. D. Wilcox, A model for multiview ultrasonic array inspection of small two-dimensional defects, IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control 66 (2019) 1129–1139.
[27] M. A. Ponti, L. S. F. Ribeiro, T. S. Nazare, T. Bui, J. Collomosse, Everything you wanted to know about deep learning for computer vision but were afraid to ask, in: 2017 30th SIBGRAPI conference on graphics, patterns and images tutorials (SIBGRAPI-T), IEEE, 2017, pp. 17–41.
[28] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, O. Ronneberger, 3D U-Net: learning dense volumetric segmentation from sparse annotation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2016, pp. 424–432.
[29] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin, Attention is all you need, arXiv preprint arXiv:1706.03762 (2017).
[30] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980 (2014).
[31] P. D. Wilcox, A. J. Croxford, N. Budyn, R. L. Bevan, J. Zhang, A. Kashubin, P. Cawley, Fusion of multi-view ultrasonic data for increased detection performance in non-destructive evaluation, Proceedings of the Royal Society A 476 (2020) 20200086.
[32] M. Sezgin, B. Sankur, Selection of thresholding methods for nondestructive testing applications, in: Proceedings 2001 International Conference on Image Processing (Cat. No. 01CH37205), volume 3, IEEE, 2001, pp. 764–767.
[33] M. V. Felice, Z. Fan, Sizing of flaws using ultrasonic bulk wave testing: A review, Ultrasonics 88 (2018) 26–42.
[34] S. Earp, A. Curtis, Probabilistic neural network-based 2D travel-time tomography, Neural Computing and Applications 32 (2020) 17077–17095.
[35] M. Chiachio, J. L. Beck, J. Chiachio, G. Rus, Approximate Bayesian computation by subset simulation, SIAM Journal on Scientific Computing 36 (2014) A1339–A1358.