Ultrasound Brain Tomography: Comparison of Deep Learning and Deterministic Methods

Manuchehr Soleimani*, Tomasz Rymarczyk, Grzegorz Kłosowski

Abstract*— The general purpose of this document is to develop a lightweight, portable ultrasound computer tomography system that enables non-invasive imaging of the inside of the human head with high resolution. The goal is to analyze the benefits of using a deep neural network containing CNN and LSTM layers compared to deterministic methods. In addition to the CNN+LSTM and LSTM networks, the following methods were used to create tomographic images of the inside of the human head: Truncated Singular Value Decomposition, Linear Back Projection, Gauss-Newton with Regularization Matrix, Tikhonov Regularization, and Levenberg-Marquardt. A physical model of the human head was made. Based on synthetic and real measurements, images of the inside of the brain were reconstructed. On this basis, the CNN+LSTM and LSTM methods were compared with deterministic methods. Based on the comparison of images and quantitative indicators, it was found that the proposed neural network is much more tolerant of noisy and non-ideal synthetic data measurements, which is manifested in the lack of the need to apply filters to the obtained images. Significance: An important finding confirmed by hard evidence is the confirmation of the greater usefulness of neural models in medical ultrasound tomography, which results from the generalization abilities of the deep hybrid neural network. At the same time, research has shown a deficit of these abilities in deterministic methods. Considering the human head's specificity, using hybrid neural networks containing both CNN and LSTM layers in clinical trials is a better choice than deterministic methods.

***Index Terms*—** Brain imaging, speed of sound imaging, ultrasound computed tomography, neural networks, brain phantom

I. Introduction

The quest for non-invasive, early, easily accessible, and effective diagnosis of various diseases has been a long-standing challenge in the medical field. It is particularly critical for conditions that affect the human head, where the early detection of diseases can significantly improve patient outcomes. Among the conditions that can be diagnosed using tomography are hematomas, brain tumors, and cerebral hemorrhages.

A. Motivation

According to the World Health Organization (WHO), neurological disorders, which include various head diseases, are responsible for 12% of total deaths globally [1]. In 2019, stroke was the second leading cause of death, accounting for approximately 11% of all deaths. In absolute numbers, this means 6.1 million people (WHO, Fact Sheets, 2020). This statistic underscores the importance of developing effective diagnostic tools for these conditions.

B. Objective and Novelties

The general aim of the research presented in this paper is to develop a lightweight, portable ultrasound computer tomography (USCT) system that enables non-invasive imaging of the inside of the human head in real time with a resolution that allows reliable identification of lesions.

Hardware and software are the novelties that bring an original scientific contribution to the general state of knowledge. The first innovation is a prototype of a lightweight and portable hardware system that enables ultrasound tomography measurements without the need to transport the patient to the tomograph device. The second innovation is an algorithm based on a unique architecture of neural networks of various types. Thanks to the skillful combination of Convolutional Neural Network (CNN) layers and the Long Short-Term Memory (LSTM) layers, a noise-resistant algorithm with a high tolerance to interference was obtained, enabling the transformation of measurements into high-fidelity images [2]. The superiority of the new CNN+LSTM model was demonstrated through a comparative analysis of its reconstruction capabilities against those of a basic LSTM network [3], [4] and deterministic methods [5]. This evaluation encompassed the utilization of both synthetic and real data.

It is important to acknowledge that the primary purpose of the study was not to construct a predictive model that outperforms all existing methods in terms of accuracy. The emphasis was placed on empirically validating the hypothesis that machine learning-based methods exhibit superior performance in comparison to deterministic approaches within the domain of USCT. Furthermore, the objective of the study was to clarify the precise characteristics that contribute to the aforementioned superiority that was observed.

The results of the study suggest that the quality of raw images obtained using deterministic methods is notably lower in comparison to those acquired through the utilization of a combined approach involving Convolutional Neural Networks and Long Short-Term Memory networks (CNN+LSTM). While it is possible to implement customized filters in order to improve the quality of images in controlled laboratory settings,

M Soleimani (m.soleimani@bath.ac.uk) is with the Engineering Tomography Laboratory (ETL), University of Bath, Bath BA2 7AY, United Kingdom. T Rymarczyk (tomasz@rymarczyk.com) is with Research & Development Centre Netrix S.A., Wojciechowska 31, 20-704 Lublin, Poland. G Kłosowski (g.klosowski@pollub.pl) is with the Faculty of Management, Lublin University of Technology, Nadbystrzycka 38D, 20-618 Lublin, Poland.

achieving the same level of intervention may be impractical in real-world measurements. The aforementioned statement highlights the superiority of machine learning techniques, specifically the heterogeneous CNN+LSTM network, as supported by the empirical findings presented in this research.

In short, this paper describes two new ideas: a prototype of a portable and light USCT system that gets rid of the need to move the patient to a stationary tomograph, and a special algorithm that combines CNN with LSTM layers to make high-fidelity image reconstruction possible.

II. Literature Analysis

A. Foundational Research and Applications in USCT

In recent years, ultrasound tomography has emerged as a promising technique for diagnosing head diseases. A pioneering study by Dines (1981) demonstrated the potential of computed tomography ultrasound to image the human head [6]. This pioneering work laid the groundwork for subsequent research in this field. Further advancements in this field have been made by integrating ultrasound tomography with other imaging techniques. For instance, Escott discussed the protocol for combining positron emission tomography with computed tomography for head and neck cancer imaging [7]. Moreover, in 2017 Tutschek et al. explored the use of computed tomography and ultrasound to determine fetal head station [8]. Although this study focused on obstetric applications, the techniques developed could potentially be adapted to diagnose head diseases. Research on ultrasound tomography for brain diseases has also been conducted. Von Ramm et al. discussed the application of ultrasound tomography for imaging the adult brain [9]. This work represents a significant step towards developing ultrasound-based techniques for diagnosing brain diseases.

B. Technical Advancements and Methodological Refinements in USCT

USCT is a rapidly evolving field with significant potential in medical imaging, particularly for diagnosing conditions affecting the human head. In "Ultrasound Modulated Bioluminescence Tomography," Guillaume Bal and John C Schotland propose a method to reconstruct the density of a luminescent source in a highly-scattering medium from ultrasound-modulated optical measurements [10]. This innovative approach leverages the interaction between ultrasound and light to enhance the imaging capabilities of USCT. Michel Gross, Philippe Goy, and Mohamed Al-Koussa further enhance the detection capabilities of USCT in their work, "Shot-noise detection of ultrasound-tagged photons in ultrasound-modulated optical imaging" [11]. Using a CCD camera, they propose a new method that allows for parallel speckle detection with optimum shot-noise sensitivity. This method could significantly improve the image quality and speed of USCT. In the paper titled "Acousto-electrical Speckle Pattern in Lorentz Force Electrical Impedance Tomography," Pol Grasland-Mongrain et al. explore the phenomenon of acousto-electrical speckle patterns within the context of Lorentz force electrical impedance tomography [12]. This phenomenon, driven by acoustic parameters but due to electrical impedance inhomogeneities, could provide new insights into interpreting USCT images. Tran Quang-Huy and colleagues propose a novel approach to image reconstruction in their work, "Deterministic Compressive Sampling for High-Quality Image Reconstruction of Ultrasound Tomography" [13]. They use chaos measurements in the detection geometry configuration and implement the image reconstruction process using L1 regularization. This approach could enhance the quality of USCT images while reducing the computational complexity of the reconstruction process. In "Ultrasound modulated bioluminescence tomography and controllability of the radiative transport equation," Guillaume Bal, Francis J. Chung, and John C. Schotland propose a method to reconstruct the density of an optical source in a highly scattering medium from ultrasound-modulated optical measurements [14]. This work further demonstrates the potential of combining ultrasound with other modalities to enhance the capabilities of USCT. In 2021 Carlos Cueto and colleagues introduced Stride, an open-source Python library for the solution of large-scale ultrasound tomography problems, in their work "Stride: a flexible platform for high-performance ultrasound computed tomography" [15]. This tool could significantly facilitate the development and application of USCT methods. In "Whole-Body Human Ultrasound Tomography," David C. Garrett and colleagues discuss the development of reflection and transmission modes [16]. This work represents a significant step towards the broad application of USCT in medical diagnostics. Ashkan Javaherian, Felix Lucka, and Ben Cox introduce finite-frequency travel time tomography to medical ultrasound in their work "Refraction-corrected ray-based inversion for three-dimensional ultrasound tomography of the breast" [17]. This method could enhance the accuracy of USCT and expand its applications. In "3D Wave-Equation-Based Finite-Frequency Tomography for Ultrasound Computed Tomography," N. Korta Martiartu, C. Boehm, and A. Fichtner introduce a robust framework for large-scale bent-ray ultrasound tomography in 3D for a hemispherical detector array [18]. This work could significantly improve the spatial resolution and accuracy of USCT. Finally, in "Ultrasound Modulated Bioluminescence Tomography With A Single Optical Measurement," Francis Chung, Tianyu Yang, and Yang Yang propose an alternative solution for this inverse problem which requires only a single optical measurement in order to reconstruct the isotropic source [19]. This approach could simplify the USCT process and enhance its accessibility. These publications highlight the ongoing efforts in the scientific community to refine and improve the methods used in USCT. The paper [20] evaluates the ability of convolutional neural networks (CNNs) to classify and recognize ultrasound imaging of the hip joint in the diagnosis of pediatric dysplasia. The study used transfer learning CNNs such as GoogleNet, SqueezeNet, and AlexNet, with GoogleNet showing the most optimal results. Another example of the use of transfer learning in CNN can be found in [21]. In the publication [22], the LSTM network is used for the remaining useful life (RUL) prediction. The authors propose a health indicator construction method based on Stacked Sparse

Autoencoder (SSAE) and combine SSAE with the LSTM network to predict the remaining RUL. The paper adopts a multi-faceted approach to feature extraction and selection, ultimately integrating it with predictive modeling for enhanced accuracy and robustness.

C. Integration of Machine Learning and Computational Methods in USCT

As these methods continue to evolve, the potential of USCT as a non-invasive, accessible, and effective diagnostic tool becomes increasingly apparent. In "Neural Operator Learning for Ultrasound Tomography Inversion" (2023) by Dai et al., the authors apply neural operator learning to ultrasound computed tomography, specifically focusing on time-of-flight data [23]. They effectively learn the mapping between this data and the heterogeneous sound speed field, bypassing the need for computationally intensive iterative inverse problem-solving. The work "Interpretable and Intervenable Ultrasonography-based Machine Learning Models for Pediatric Appendicitis" by Marcinkevičs et al. presents interpretable machine learning models for predicting the diagnosis, management, and severity of suspected appendicitis using ultrasound images [24]. This approach provides a valuable tool for clinicians to diagnose and manage this common pediatric condition. Noda et al., in their publication "Blind Signal Separation for Fast Ultrasound Computed Tomography," propose a method named FastUSCT to acquire high-quality images faster than traditional methods for ultrasound tomography [25]. This method concerns transmitting multiple ultrasound waves simultaneously to reduce imaging time and using UNet to separate overlapping waves. Shan et al.'s "Simultaneous reconstruction of the initial pressure and sound speed in photoacoustic tomography using a deep-learning approach" develops a novel data-driven method integrating an advanced deep neural network through model-based iteration [26]. This approach allows for simultaneous photoacoustic tomography reconstruction of initial pressure and sound speed. Lastly, "Amortized Normalizing Flows for Transcranial Ultrasound with Uncertainty Quantification" by Orozco et al. presents a novel approach to transcranial ultrasound computed tomography that utilizes normalizing flows [27]. This method improves imaging speed and quantifies Bayesian uncertainty, offering the potential for real-time predictions.

The publication [28] by Islam et al. serves as evidence that the idea of combining CNN and LSTM layers is present in the literature. The authors propose a deep learning-based approach for the detection of COVID-19 using X-ray images. The architecture combines CNN for feature extraction and LSTM for sequence analysis, achieving high diagnostic metrics including an accuracy of 99.4%. Another example of this idea is the paper [29] by Agga et al., which addresses the issue of short-term photovoltaic power prediction through a hybrid deep learning architecture consisting of CNN and LSTM layers. Utilizing a dataset from Rabat, Morocco, the paper demonstrates superior predictive performance in comparison to traditional machine learning and single deep learning models, as gauged by error metrics such as MAE, MAPE, and RMSE. The study [30] by Liu et al. delves into the problem of video decolorization, focusing on maintaining temporal consistency and local contrast in grayscale video frames. The research integrates CNNs and LSTMs in a novel architecture that not only encodes local semantic content but also refines it using bi-directional LSTMs to mitigate flicker phenomenon and other temporal inconsistencies. The paper [2] by Kłosowski et al. (2023) presents research on the utilization of machine learning algorithms for moisture detection in building walls through Electrical Impedance Tomography (EIT). The study evaluates three separate neural network models—CNN, LSTM, and a hybrid model (CNN+LSTM)—using metrics such as MSE and correlation coefficient, concluding that the hybrid model offers significant advantages in solving the tomographic inverse problem.

This work consists of five sections. Section I, Introduction, includes considerations such as motivation, the objective of the research, and novelties. Section II is a literature review concerning such topics as foundational research in USCT, technical advancements, and the integration of machine learning methods in USCT. Section III describes the technical details of the USCT device, transducers, and human head model. The method of transforming USCT measurements into images is also explained. Creating the training set used for neural network training is discussed. The deterministic methods used are described, and the architecture of the CNN+LSTM and LSTM models is outlined [28]. In Section IV, the results are presented in the form of reconstructions based on both synthetic and real data. Additionally, tabulated quality indicators for various reconstruction cases are presented. The conclusions included in Section V assume the paper.

III. USCT Device and Methods

In medical imaging, USCT has potential to take a key position due to its non-invasive nature and potential for high-resolution imaging. USCT constitutes a system comprising various key elements: transducers, an electronic setup facilitating data acquisition and processing, and the accompanying software designed for solving the inverse problem of transforming measurements into comprehensible images. This synthesis of hardware and software within the USCT system exemplifies an intricate balance, where the quality of the system's output is significantly dependent on the performance of its least efficient component.

Transducers, the primary hardware components, act as the system's sensory organs, converting energy forms to generate and receive ultrasonic waves. The electronic subsystem of the USCT is another pivotal component. It enables the collection of raw data and initiates preliminary processing. Software, the final significant component, has the complex task of solving the inverse problem. It transforms the raw data acquired and initially processed by the electronic system into coherent, visual representations of the body's internal structures.

A. USCT Hardware

In the research and development laboratory of Netrix SA, our team designed and built a prototype USCT tomograph. The

ultrasonic tomograph is a robust, versatile tool configurable to work in two principal modes: full-waveform and transmission. The full-waveform mode provides raw data, while the transmission mode delivers processed time-of-flight (TOF) and pulse amplitude values. Transmission mode, with its reduced data size, offers a more expedient temporal frequency of 4 frames per second (fps), compared to the full-waveform mode's 0.08 fps. This efficiency makes it an ideal choice for real-time monitoring applications.

When the tomograph operates in transmission mode, it measures the signal's travel time and amplitude. The minimum and maximum signal values are auto-detected, and the Analog-Digital Converter (ADC) uses these values to convert the percentage value to a numerical one. A measurement window opens once the signal surpasses a set comparator threshold, within which the largest amplitude value is selected for processing to compute its time-of-flight. Conversely, when in full-waveform mode, the tomograph provides the complete waveform with an adjustable sampling frequency option. The USCT device block diagram is shown in Fig. 1.

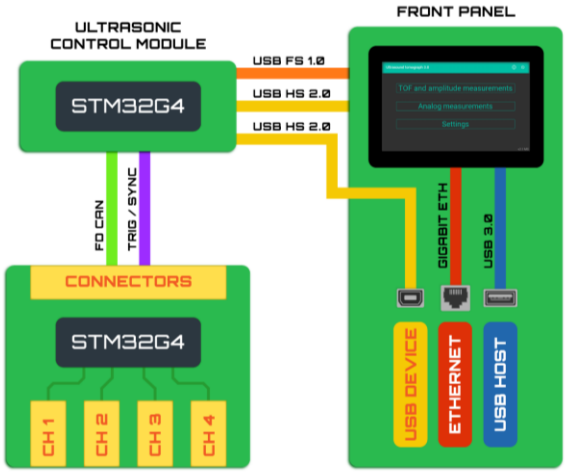

Fig. 1. Ultrasonic tomograph block diagram

The USCT system's hardware design comprises 84 channels distributed over eight four-channel measurement cards. These cards are connected through an FD CAN bus to a measurement module, which is the interface between the microprocessor measuring system and the touch panel or an external control application. In this instance, the touch panel utilizes a Raspberry Pi 4B 2GB RAM board and a 7-inch capacitive touch screen.

Each sensor within the system has its signal conditioning and measurement circuit. This arrangement allows for concurrent measurements for each excitation. The measuring cards can sample at a maximum rate of 4 Mega Bytes per Second (MBPS) per channel. Moreover, each channel has a separate generator for Alternating Current (AC) rectangular waveforms. This generator has an amplitude of up to 144V peak-to-peak (Vp-p) and an instantaneous current capacity of 3 Amperes (A).

Harmonic filtering is managed effectively within each channel through three eight-order filters. In addition, a two-stage gain control is present in each channel: the first stage varies from +7.5dB to +55.5dB (as managed by the AD8331 component), and the second stage offers a range of +6dB to +36dB, facilitated by the STM32's built-in Programmable Gain Amplifier (PGA).

The device also houses a four-channel high-rectangular-voltage generator capable of generating a three-stage square wave signal (Vpp–GND–Vnn) ranging from +100 Volts (V) to −100 V. This generator circuit comprises four–channel Metal–Oxide–Semiconductor Field–Effect Transistor (MOSFET) drivers connected to double H bridges (component TC8220). The STM32G474 microcontroller within the device manages the synchronization of the converters and controls the generation of appropriate waveforms.

The analog module is a subsystem designed to amplify the ultrasonic transmitting signal. It houses an integrated AD8331 amplifier with gain control via an external Digital-Analog Converter (DAC), a system converting the AC signal to ADL5511 envelopes, and two THS4521 differential amplifiers. The output from this module is a symmetrical differential signal, which reduces noise due to the high-voltage generator being on the same Printed Circuit Board (PCB). The ADC converter speed sets the sampling frequency at 0.25 samples per micro-second. Finally, an integrated envelope converter converts an analogue acoustic signal into the envelope.

The tomograph has small dimensions and is housed in a suitcase, which allows it to be transported in a car (Fig. 2).

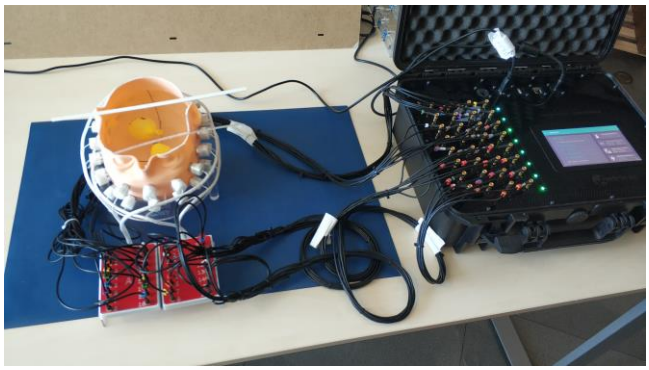

Fig. 2. Prototype of the portable ultrasonic tomograph device connected to the model of the human head

An adult mannequin was used to design a human head model. Fig. 3 shows the head model with marked contact points for the USCT transducers.

Fig. 3. Phantom head with markers

Fig. 4 shows the phantom head with a transmitter-receiver ultrasonic transducer generating a wave with a frequency of 40 kHz.

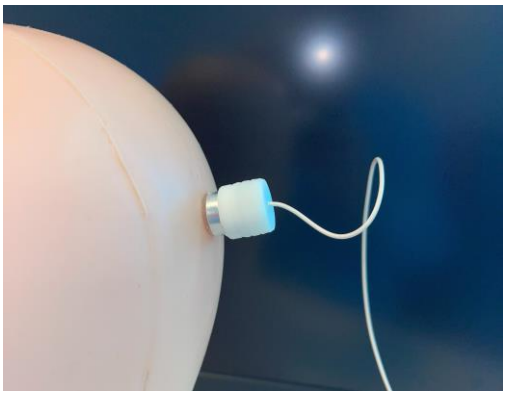

Fig. 4. Phantom head with 40kHz ultrasonic transducer

Fig. 5 shows the transducer with a plastic housing.

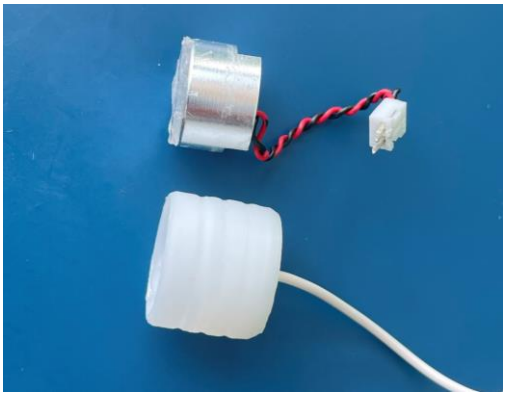

Fig. 5. Ultrasonic transducer 40kHz with housings

B. Method of Image Reconstruction

The software in the USCT system needs to be robust and advanced enough to manage and interpret vast volumes of complex data efficiently. The limitations of the software could potentially lead to inaccurate or low-quality imaging, despite the high functionality of the other components. Thus, each of the components above—electrodes, electronics, and software—plays integral roles in the overall efficacy of the USCT system. The quality of the system is considerably dependent on the performance of its weakest component. Therefore, to ensure optimal functionality, each component within the USCT system must accomplish its respective tasks with high efficiency.

In this study, transmission tomography was used, and the measurement data included time-of-flight (TOF) data. The differential measurement data TOF is derived from the subtraction of the measurement values of the object containing phantoms TOF_{inc} from the background measurements TOF_{back} (object without inclusions). Moreover, the time data is expressed in microseconds (μs) [31].

$$TOF = TOF_{inc} - TOF_{back} \tag{1}$$

In situations involving a travel-time delays model, it is possible to utilize a practical parameterization for the rays that relies on the ray-approximation methodology.

$$\Delta T = \int_{ray} \Delta t(r)\, dl \tag{2}$$

The integral defined above is fundamentally reliant on a solitary ray trajectory, whereby ΔT corresponds to the measured delays in travel time (for instance, the discrepancies derived from subtracting reference data from actual experimental data), while $\Delta t(r)$ signifies the distribution of travel-time delays. Detailed explanations for Eq. (2) can be found in Section 2.2, 'Simulation Forward Model,' in the paper [32].

The Eq. (2) can be written as $\Delta T = \int(\delta c(l)/c)\, dl$ where $\delta c(l)$ is the slowness perturbation at arc length l along the ray trajectory and c is the average sound speed in the medium. This formula states that the TOF perturbation is equal to the integral of the slowness perturbation along the ray trajectory. The slowness perturbation is a measure of how much the sound speed is changed at a given point in the medium.

The equation provided below succinctly explains the linearized representation of the quandary above.

$$\Delta T_m = \sum_{n=1}^{N} J_{m,n}\ \Delta t_n \tag{3}$$

Within Eq. (3), ΔT_m represents the temporal lags engendered by each n^{th} transducer, where $1 < n < N$. $J_{m,n}$ indicates a weighted quantity that elucidates the influence of each distinct wave-ray on every pixel, a phenomenon termed as ray-pixel conjunction. $J_{m,n}$ undergoes normalization between 0 and 1, driven by the summative value across each pixel within the domain. Ultimately, t_n demarcates the territory of the temporal lag distribution. This archetype exhibits simplicity when juxtaposed with the sound-speed model, as it eschews considerations related to the length of the rays and consequently bypasses transformation into the sound-speed domain. To decipher the inverse conundrum embedded in ultrasound tomography, we resort to the sound-speed schema, employing as our measurement the time-of-flight data ΔT. Amplitude data can be uniformly employed to yield an acoustic attenuation map. The linear inverse dilemma in travel-time ultrasound tomography could be construed as the extraction of the alteration in pulse travel time $y = \Delta T$, precipitated by a deviation in the measured boundary acoustic velocity x = Δs.

Eq. (4) presents a linear equation that embodies the principal notion of the forward problem. The system matrix J is calculated on the tenets of ray propagation in space, as expounded in the preceding discourse [32].

$$y = J\, x + e \tag{4}$$

In Eq. (4) J constitutes the forward function and e encapsulates the noise infiltrating the measurements, x is delineated as the changes in acoustic velocity schematic of the region under scrutiny across distinct, discrete instances. Here y represents the changes in measured TOF values. The inverse predicament could be structured as follows:

$$x \approx J^{-1}\, y \tag{5}$$

In the conventional realm, acoustic tomography yields a markedly reduced tally of assessments to the quantity of image pixels destined for reconstruction. Hence, the sensitivity matrix denoted as J diverges from a square matrix and lacks invertibility. Accordingly, the issue transmutes into a non-linear and ill-conditioned quandary, with a routine resolution strategy encompassing J's transpose matrix. This approach earns the appellation of 'linear back-projection reconstruction [33].

$$x \approx J^T y \tag{6}$$

For image reconstruction, each transposed sensitivity matrix is subjected to multiplication by the corresponding sensor's attenuation value, individually projecting back onto the image plane. Subsequently, the aggregation of these matrices yields the back-projected acoustic speed or travel-time or dissipation distributions [34].

C. Forward Problem and Dataset Preparation

An in-depth examination of the ultrasound measurement process is depicted in Fig. 6. The apparatus employed in this study could both emit and capture sound waves using individual transducers.

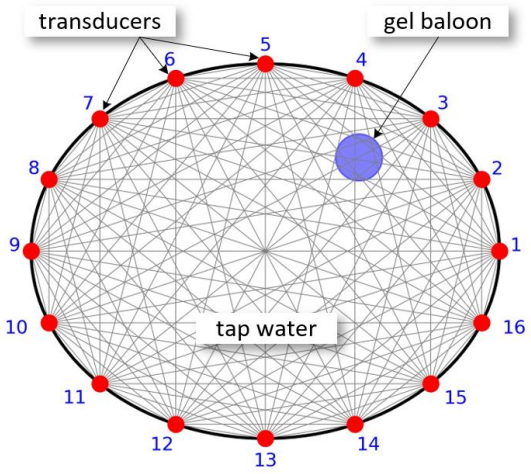

Fig. 6. The way of creating the UST measurements

Consequently, a comprehensive series of measurements was conducted to ascertain the flight time of sound waves between all transducers. It is important to note that each transducer was excluded from receiving its own signal during these measurements. Following this criterion, the total number of measurements required for a system comprising 16 transducers amounted to $16^2 - 16 = 240$. However, by assuming that the TOF of sound waves remains unaffected by the direction of propagation (i.e., $t_{i,j} = t_{j,i}$ where i,j are the numbers of transducers), the overall number of measurements can be halved, resulting in $(256 - 16)/2 = 120$. This calculation is determined by formula $M = (n^2 - n)/2$, where n represents the number of transducers. USCT measurements are transformed into a finite element mesh with a resolution of 2811 pixels. Each pixel gets a specific value in the form of a real number (regression problem). The value of a pixel is correlated with the speed of sound characteristic of the material medium to which a given pixel belongs. Matlab software with the EIDORS (eidors3d.sourceforge.net) toolbox was used to transform USCT measurements into finite element tomographic images.

The visualization presented in Fig. 6 portrays the measurement grid density juxtaposed against the backdrop of the cross-section of the tested model resembling a human head. Within the realm of ultrasonic tomography, a remarkable phenomenon is leveraged: the dynamic variability of sound wave propagation speed contingent upon the surrounding environment. It is worth emphasizing that this propagation speed is intricately linked to time, thereby obviating the need for the USCT system to engage in the arduous task of directly calculating wave propagation speed. Instead, the focus lies on estimating the precise time required for the sound wave to traverse the space between individual transducers. An aspect illustrated in Fig. 6 is the presence of a phantom situated within the sculp, acting as a conduit through which the sound wave traverses, with profound implications for the measurement process.

Fig. 7 shows a single simulation case in which, solving a forward problem, a 120-value measurement vector was assigned to a random phantom plotted on a finite element mesh [35]. A total of 35,000 synthetic training cases were created, with 4,000 of these specifically set aside for validation purposes.

D. Deterministic Methods of Transforming Measurements Into Images

In order to assess the effectiveness of the proposed neural network architecture containing both CNN and LSTM layers, images were reconstructed using a homogenous LSTM model and popular deterministic methods: Truncated Singular Value Decomposition (TSVD), Linear Back Projection (LBP), Gauss-Newton (GN) with Regularization Matrix, Tikhonov Regularization (TR), and Levenberg-Marquardt (LM).

Singular Value Decomposition (SVD) is a mathematical technique that can decompose a matrix into three other matrices. In the context of tomography, TSVD is used to solve the inverse problem. The main idea is to truncate or ignore the

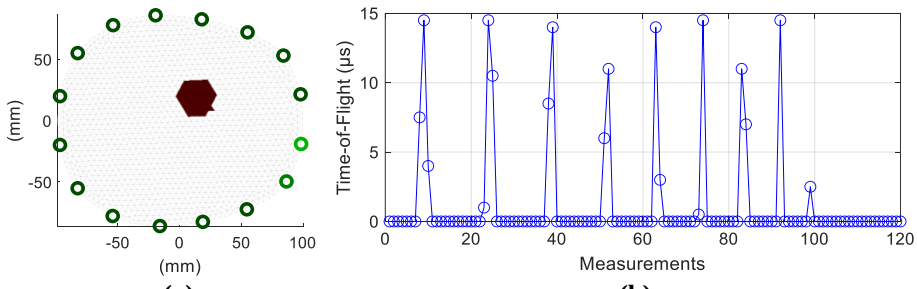

Fig. 7. An observation created with the simulation method of UST: (a) – a random phantom plotted on a finite element mesh, (b) – 120 TOF measurements between different pairs of transducers

smaller singular values, which often correspond to noise. The general SVD decomposition of a matrix J can be represented as follows [36]:

$$J = U\Sigma V^T \tag{7}$$

where U and V are orthogonal matrices, and Σ is a diagonal matrix containing the singular values of J. The truncated SVD solution to the inverse problem $Jx = y$ can be represented as:

$$x_{TSVD} = V\Sigma_{trunc}^{\dagger} U^T y \tag{8}$$

where $\Sigma_{trunc}^{\dagger}$ is the inverse of the truncated singular value matrix.

Linear Back Projection (LBP) is a simple and intuitive method used in tomography. The basic idea is evenly distributing the measured projection value along the projection line. This method can be computationally efficient but often results in blurred images, the same as equation (6), normalization of the Jacobian matrix helps with more stable LBP reconstruction. The formula for the implemented back-projection is [37]:

$$x = (J_n)^T y \tag{9}$$

where J_n is the column-normalized sensitivity matrix, and y is the residual vector (differences between two TOF values).

The Gauss-Newton (GN) algorithm is used for non-linear least squares problems. In the context of tomography, a regularization term is often added to stabilize the solution.

The Gauss-Newton update rule with a regularization matrix R can be written as [38]:

$$x = (J^T J + \lambda R)^{-1} J^T y \tag{10}$$

where J is the Jacobian matrix, λ is the regularization parameter, y is the residual vector, and x is the update to the parameter vector.

Tikhonov Regularization (TR) is a method used to stabilize inverse problems. The idea is to balance fitting the data and keeping the solution smooth. The solution to the TR problem can be written as [39]:

$$x_{reg} = arg\ \min_x \{ \|Jx - y\|_2^2 + \lambda^2 \|Lx\|_2^2 \} \tag{11}$$

where J is the system matrix, y is the measured data, λ is the regularization parameter, and L is the regularization matrix, in this case a Laplacian operator.

The Levenberg-Marquardt (LM) method is an algorithm used for solving non-linear least squares problems. It can be seen as a combination of the Gauss-Newton method and the method of gradient descent. The Levenberg-Marquardt update rule can be written as [40]:

$$x = (J^T J + \lambda I)^{-1} J^T y \tag{12}$$

where J is the Jacobian matrix, I is the identity matrix, λ is the damping factor, y is the residual vector, and x is the update to the parameter vector.

E. Neural network CNN+LSTM approach

The structure of the considered CNN+LSTM neural network is listed in Table I. The network has as many as 57.2 million total learnables.

The deep and varied network architecture is a comprehensive combination of multiple layers, each playing a vital role in processing and learning from the input data. The network starts with a sequence input layer, which accommodates sequence inputs with 120 measurements. This approach enables the network to handle sorted data or any sequence-based data effectively, potentially enhancing its performance on such data types.

TABLE I
CNN+LSTM NETWORK ARCHITECTURE

No	Layer	Description
1	Sequence Input	Sequence input with 120 dimensions
2	1-D Convolution	64 3x120 convolutions with stride 1 and padding 'same'
3	Batch Normalization	Batch normalization with 64 channels
4	ReLU	ReLU
5	Dropout	20% dropout
6	1-D Convolution	128 3x64 convolutions with stride 1 and padding 'same'
7	Batch Normalization	Batch normalization with 128 channels
8	ReLU	ReLU
9	Dropout	20% dropout
10	LSTM	LSTM with 2048 hidden units
11	Dropout	20% dropout
12	LSTM	LSTM with 2048 hidden units
13	Dropout	20% dropout
14	Fully Connected	2811 fully connected layer
15	Regression Output	mean-squared-error as the Loss function

The first transformation applied to the input is a 1-D Convolution, involving 64 filters of size 3×120, with a stride of 1, and a 'same' padding, which ensures the output has the same width as the original input. This layer extracts local features through convolution operations, with two instances escalating the complexity of learned features. Their kernel size and padding ensure a receptive field adept at capturing local patterns without altering the spatial dimensions. This layer's output is normalized by the batch normalization layer with 64 channels. This normalization process improves the network's speed, performance, and stability. After normalization, the network employs a ReLU (rectified linear unit) activation function, which is a popular choice for introducing nonlinearity without impacting the receptive fields of the convolution layer. The first dropout layer drops out 20% of the input data to prevent overfitting. Dropout is a well-known regularization technique that randomly nullifying some of the layer's outputs to make the model more robust and less prone to overfitting. Following the same pattern, the next layers consist of another 1-D Convolution with 128 filters of size 3×64, a stride of 1, and 'same' padding, followed by batch normalization with 128 channels, another ReLU activation, and another dropout at 20%. The model then transitions to the LSTM (Long Short-Term Memory) layer with 2048 hidden units. LSTM layers are excellent for handling sequence data over long periods of time, as they can remember important information and forget irrelevant data, making them especially useful for sequence prediction tasks. A subsequent dropout layer with a 20%

dropout rate follows the first LSTM layer.

Another LSTM layer with 2048 hidden units is present, followed by a corresponding dropout layer, maintaining the pattern of alternation between computational and regularization layers. Two LSTM layers, with differing output modes, ensure a robust comprehension of temporal dynamics. The penultimate layer is a fully connected layer with 2811 units, integrating the learning from previous layers to perform more complex decisions. This layer consolidates learned features into a format amenable to the regression task. The architecture concludes with a regression output layer, which uses mean squared error as the loss function to optimize the model's performance. This final layer minimizes the average squared difference between the actual and predicted output.

This architecture has been crafted to effectively process sequence-based inputs, with convolution layers catering to localized features and LSTM layers to sequential dependencies. Dropout layers interspersed throughout the network help mitigate overfitting, making this model both robust and capable of handling complex patterns in data. The tradeoff between depth and breadth, local feature learning, sequence learning, and overfitting prevention all potentially contribute to the network's effectiveness. However, the model's performance is ultimately data-dependent, so it has been validated through actual reconstructions obtained from measurements from a physical object.

The training of the described neural network follows a meticulous and well-defined process using the Adam optimization algorithm, which combines the advantages of two other extensions of stochastic gradient descent: AdaGrad and RMSProp [41]. Adam is known for handling sparse gradients on noisy problems, making it a suitable choice. The number of epochs was large, enabling the network to learn complex patterns from the data, given sufficient diversity in the dataset.

The data is divided into mini-batches of size 64 for each iteration of training. This approach, known as mini-batch gradient descent, compromises computational efficiency and the reliable convergence characteristics of stochastic gradient descent. Validation data is provided and used to evaluate the model's performance on unseen data at every 100 epochs. The training algorithm shows validation patience of 20 epochs, i.e., the training will stop if the validation loss does not decrease for 20 consecutive validation checks. This approach is a type of early stopping and effectively prevents overfitting and unnecessary continuation of training.

The initial learn rate is set to 0.001, which determines the step size at each iteration while moving towards a minimum loss function. The learning rate is modulated according to a 'piecewise' schedule, where it drops by a factor of 0.2 every 5 epochs. This approach allows the model to make large changes to the weights at the beginning and gradually smaller changes as the training progresses. The data is shuffled after every epoch, which contributes to the model's robustness by ensuring that the gradient descent does not settle into a local minimum based on the order of the input data. The described training process is optimized to ensure effective training of the model.

The Adam optimizer, mini-batch gradient descent, many epochs, early stopping, learning rate scheduling, and data shuffling contribute to robust and potentially successful training. However, the effectiveness depends on the nature and complexity of the data and the task at hand.

When designing the architecture of the neural network, previous experience was taken into account, which showed that both one-dimensional convolutional layers and recurrent LSTM networks work well for tomographic problems. In order to combine the advantages of both of these structures, the CNN+LSTM network was created. The order and number of layers were selected experimentally. In order to optimize hyperparameters such as "probabilities" in dropout layers, the number of filters in convolutional layers, and the number of hidden units in LSTM layers, the Experiment Manager application, which is part of the Matlab package, was used. An optimization strategy called exhaustive sweep was used. The exhaustive sweep strategy in MATLAB is a technique used to search the hyperparameter space during machine learning or deep learning. When using this strategy, MATLAB executes the experiment for each possible combination of hyperparameter values specified in the hyperparameter table.

F. LSTM approach

The provided information in Table II illustrates the structural composition of a homogeneous neural network, wherein LSTM layers are exclusively employed, and convolutional layers are absent. These are the same LSTM layers used in the CNN+LSTM model, but the network has been simplified by removing the 1-D Convolution, Batch Normalization and ReLU layers. Both the training data and the learning process parameters are the same as those used in the previous CNN+LSTM model.

TABLE II
LSTM NETWORK ARCHITECTURE

No	Layer	Description
1	Sequence Input	Sequence input with 120 dimensions
2	LSTM	LSTM with 2048 hidden units
3	Dropout	20% dropout
4	LSTM	LSTM with 2048 hidden units
5	Dropout	20% dropout
6	Fully Connected	2811 fully connected layer
7	Regression Output	mean-squared-error as the Loss function

IV. RESULTS AND DISCUSSION

In order to assess the quality of the CNN+LSTM and LSTM neural models using quantitative criteria, synthetic measurements are necessary. Comparison of the obtained reconstruction images with the model (reference) images enables the use of indicators considered standard in tomographic examinations.

A. Reconstructions from synthetic measurements

Fig. 8 shows reconstructions based on synthetic test measurements. The first column, labelled "Patterns", displays randomly generated pattern images, from which measurements were computed to tackle a forward problem. The second column shows the results of the CNN+LSTM network, while

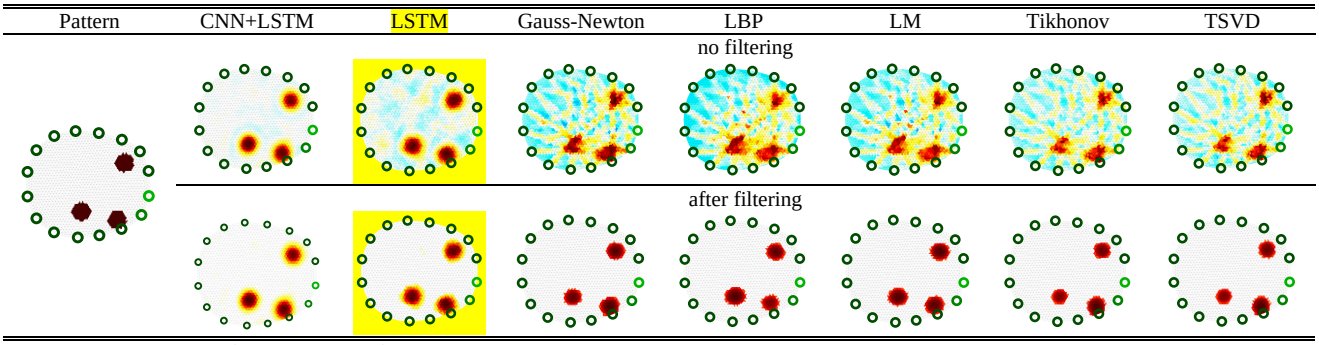

Fig. 8. Comparison of reconstructions from synthetic measurements; images made with deterministic methods (GN, LBP, LM, Tikhonov, and TSVD) were filtered with the Mexican Hat method; CNN+LSTM and LSTM images have been enhanced with a simple cutoff filter.

the next six columns show the USCT reconstructions made using homogenous LSTM and deterministic methods (Gauss-Newton, LBP, LM, Tikhonov, and TSVD). Only one measurement case is shown in Fig. 8. Other reconstructions containing different numbers of phantoms look very similar, so the case shown in Fig. 8 can be treated as representative. The top row shows the raw images obtained directly by converting the measurements to tomograms. In the bottom row of Fig. 8, the same images are improved by data filtering. The reconstructions obtained by deterministic methods were tried to be improved using Gaussian filtering and Mexican Hat filtering, with better results obtained using the latter method [42]. It is clear that, in the case of unfiltered images, the CNN+LSTM method definitely and indisputably outperforms the homogenous LSTM and deterministic methods. In the case of filtered reconstructions, the differences are less noticeable.

Since the human eye is not a reliable and unambiguous measure of image fidelity, four quantitative measures were used. The inaugural metric used for the assessment of the tomographic reconstructions was the mean square error (MSE) computed by Eq. (13)

$$\mathrm{MSE} = \sum_{i=1}^{k} \frac{(x_i - x_i^*)^2}{k} \tag{13}$$

where x_i symbolizes the pattern image composed of n pixels, x represents the reconstructed image. The image resolution $k = 2811$.

The subsequent metric was the peak signal-to-noise ratio (PSNR), delineated by Eq. (14).

$$\mathrm{PSNR} = 10 \cdot \log_{10}(R^2/MSE) \tag{14}$$

For this investigation, we selected $R = 1$, where R signifies the maximum permissible fluctuation in the pixel value utilized to generate the input image. PSNR gauges the signal emanating from the noise in the ultimate reconstructed image. A superior PSNR value indicates a higher-quality image.

An additional indicator employed to evaluate the proposed method is the structural similarity index (SSIM) [43]. SSIM is distinctive in that it amalgamates three features into a singular indicator. Quantities such as contrast, local image structure, and luminance are compiled into one synthetic measure. The feature referred to as "structure" is the intensity patterns of the finite elements of the FEM mesh, which consider the adjacent pixels. Structures are normalized in terms of luminance and contrast.

The human visual system is adapted to observe structural differences, which is why the SSIM index aligns with human assessment. In the context of tomography, this is a specific advantage. The indicator is defined by Eq. (15)

$$\mathrm{SSIM} = \frac{(2\mu_{x^*}\mu_x + \mathrm{C}_1)(2\sigma_{x^*x} + \mathrm{C}_2)}{(\mu_{x^*}^2 + \mu_x^2 + \mathrm{C}_1)(\sigma_{x^*}^2 + \sigma_x^2 + \mathrm{C}_2)} \tag{15}$$

where $\mu_{x^*}, \mu_x, \sigma_{x^*}, \sigma_x \; and \; x$ are the local means, standard deviations, and cross-covariances for images x^*, x; $C_1 = (0.01 \cdot L)^2$ and $C_2 = (0.03 \cdot L)^2$. If the pixel values are within the range (0,1), as is the case here, then $L = 1$.

The final measure utilized in this research is the image correlation coefficient (ICC), which is calculated according to Eq. (16)

$$\mathrm{ICC} = \frac{\sum_{i=1}^{k}(x_i - \bar{x})(x^*{}_i - \overline{x^*})}{\sqrt{\sum_{i=1}^{k}(x_i - \bar{x})^2 \sum_{i=1}^{k}(x^*{}_i - \overline{x^*})^2}} \tag{16}$$

where $\overline{x^*}$ and $\bar{x}$ are the distributions of the average values of the reconstructed and reference images, respectively.

Table III displays the mean values of the metrics (MSE, PSNR, SIMM, and ICC) for a dataset consisting of 4,000 validation cases. These cases were not utilized in the process of training the neural network.

The index values for the variant of unfiltered reconstructions are placed in the upper rows of Table III. The bottom four rows show indicators for enhanced images. Bold values indicate how the given indicator achieved the best value. For MSE, the smaller the value, the better the reconstruction. The higher the

TABLE III
COMPARISON OF THE RECONSTRUCTIONS THROUGH INDICATORS

Variant	Metrics	CNN+ LSTM	LSTM	Gauss-Newton	LBP	LM	Tikho-nov	TSVD
no filtering	MSE	**0.002**	0.003	0.006	0.008	0.007	0.006	0.006
	PSNR	**26.98**	25.86	21.95	20.87	21.64	22.09	22.24
	SIMM	**0.707**	0.516	0.300	0.212	0.283	0.300	0.292
	ICC	**0.901**	0.874	0.665	0.608	0.649	0.662	0.671
after filtering	MSE	**0.002**	0.003	0.008	0.010	0.008	0.007	0.007
	PSNR	**27.10**	26.96	21.32	20.28	21.01	21.51	21.84
	SIMM	**0.772**	0.723	0.721	0.676	0.699	0.735	0.755
	ICC	0.904	0.898	**0.935**	0.889	0.929	0.917	0.913

PSNR, SIMM, and ICC indicators, the better. For ICC, a special case of the Pearson correlation coefficient for images, high values (close to +1) indicates a strong agreement between two images, while values close to −1 indicate inverse solid relationships. Values near 0 suggest a lack of correlation.

In evaluating image processing techniques using the provided dataset, we observed distinct performance characteristics under two scenarios: images without filtering and images post-filtering. For unfiltered images, two methods, namely CNN+LSTM and LSTM, demonstrated superior performance concerning Mean Squared Error (MSE), all yielding the lowest values. It suggests that these methods exhibit the least amount of error. Regarding the Peak Signal-to-Noise Ratio (PSNR), the CNN+LSTM method outperformed all other methods, indicating the highest signal-to-noise ratio. For the SIMM metric, the CNN+LSTM method also ranked highest. This is significant because SIMM is the metric most akin to human perception. In terms of ICC, the CNN+LSTM method showed the second most optimal performance, which does not change the fact that in the overall comparison, the heterogeneous CNN+LSTM model turned out to be the clear winner.

In the scenario involving filtered images, both the CNN+LSTM and LSTM methods maintained the lowest MSE values, thereby continuing to exhibit the least amount of error. In the case of PSNR, the CNN+LSTM method retained its superior performance, providing the highest signal-to-noise ratio. When evaluating the SIMM metric, the CNN+LSTM method again ranked highest. However, we observed a slight deviation from the previous trend with the ICC metric. The Gauss-Newton method demonstrated the highest performance, although the differences with the Levenberg-Marquardt, Tikhonov, TSVD, and CNN+LSTM methods are insignificant. The CNN+LSTM method showed the second most optimal performance, which does not change the fact that in the overall comparison, the heterogeneous CNN+LSTM model turned out to be the clear winner. All deterministic methods achieved $ICC > 0.9$, so the results can be considered correct. LSTM also showed high quality, obtaining an ICC close to 0.9.

In conclusion, an overarching analysis of the performance metrics across filtered and unfiltered images suggests that the CNN+LSTM method consistently demonstrates robust performance across all metrics. Nevertheless, it is noteworthy that the Levenberg-Marquardt method exhibits significant competence with the ICC metric in the context of filtered images. These findings underscore the importance of method selection in image processing, which is contingent upon the specific requirements of the task, the metrics prioritized, and the nature of the images, whether filtered or unfiltered.

In light of the obtained results, the CNN+LSTM method demonstrates substantial resilience to noise, coupled with a significant capacity for generalization. This robustness enables it to process and interpret data effectively, even when it considerably deviates from the training dataset. Such an attribute is particularly beneficial in image processing, where the ability to adapt to various data is paramount. Contrarily, deterministic methods do not appear to exhibit the same adaptability or effectiveness without image filtering. Their performance seems to be reliant on the precondition of image filtering, without which their utility is limited. These findings highlight the inherent distinction between machine learning methods such as CNN+LSTM, which can learn and adapt from data, and deterministic methods, whose performance is more rigid and constrained by their predefined rules and assumptions.

Therefore, in scenarios that involve diverse and noisy datasets, the use of adaptable machine learning methods like CNN+LSTM may be more favorable than deterministic methods, emphasizing the importance of selecting the appropriate method based on the specific characteristics of the data and the task at hand.

B. Reconstructions from actual measurements

Figs. 9 and 10 show three identical measurement cases, respectively. However, Fig. 9 shows unfiltered reconstructions, while Fig. 10 shows images subjected to methods improving their readability. As in the case of synthetic measurements, the high efficiency of the CNN+LSTM network for imaging raw, unfiltered tomograms is noteworthy. Since the reconstructions in question were made based on actual measurements obtained using a physical model of the human head, it is not possible to compare the obtained results with reference measurements. In the case of a real object, such as a human head, performing a reference (background) measurement is impossible. In the case of a physical model, you can perform a reference measurement by removing all phantoms from the head model. This results in a differential measurement ($\Delta x = x_{raw} - x_{ref}$), where x_{raw} is a raw measurement and x_{ref} is the reference measurement vector, so theoretically, the phantom-free reconstruction pixels have zero values while the inclusion pixels are non-zero. In this case $\Delta x = TOF$ (see Eq. 1), so if the phantoms are used to slow down the sound waves relative to the surrounding medium, then the phantom pixels in the tomograms will have positive values ($\mathrm{TOF}_{\mathrm{phantom}} > 0$), and the inclusions will be red in the images. If ($\mathrm{TOF}_{\mathrm{phantom}} < 0$), the phantoms would be blue. The phantoms in the images are red, which means that the gel contained in the balloons slows down the sound waves relative to the tap water. The Matlab Parula colormap was used in this study but can be changed based on subjective choice.

It should be considered that it is impossible to make a reference measurement of a living human head. A synthetic reference measurement should be used to obtain a tomographic reconstruction performed in clinical conditions, which will certainly not perfectly reflect the physical conditions of the human head. For this reason, the CNN+LSTM neural network with a high generalization ability, tolerating deviations from the data on which it was trained, has a much greater potential for generating faithful reconstructions than deterministic methods.

V. Conclusion

This research conducted a comparative analysis of machine learning and deterministic methods in medical applications related to ultrasonic brain tomography. A deep neural network model, comprising both Convolutional Neural Network (CNN)

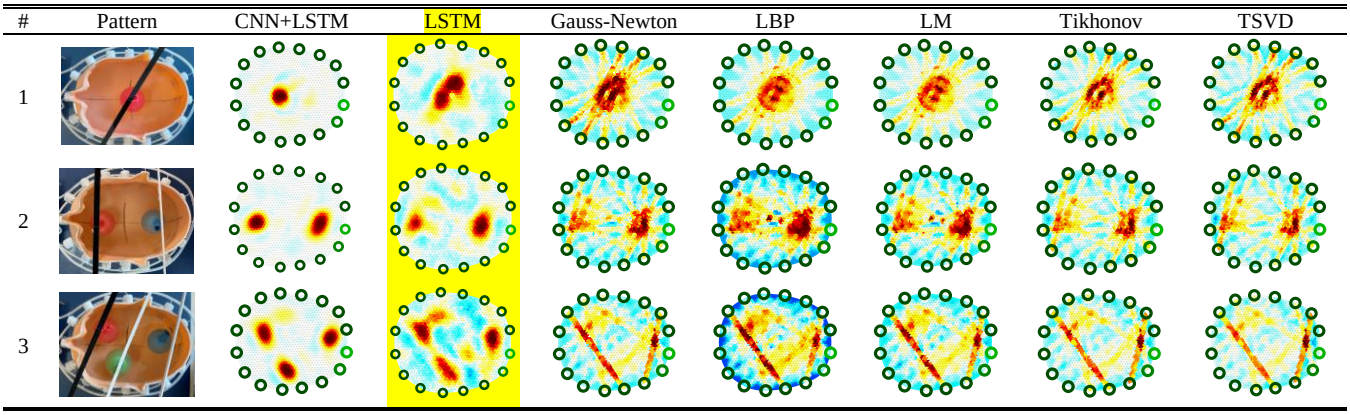

Fig. 9. Comparison of reconstructions from real measurements – images were not subjected to any improvement methods

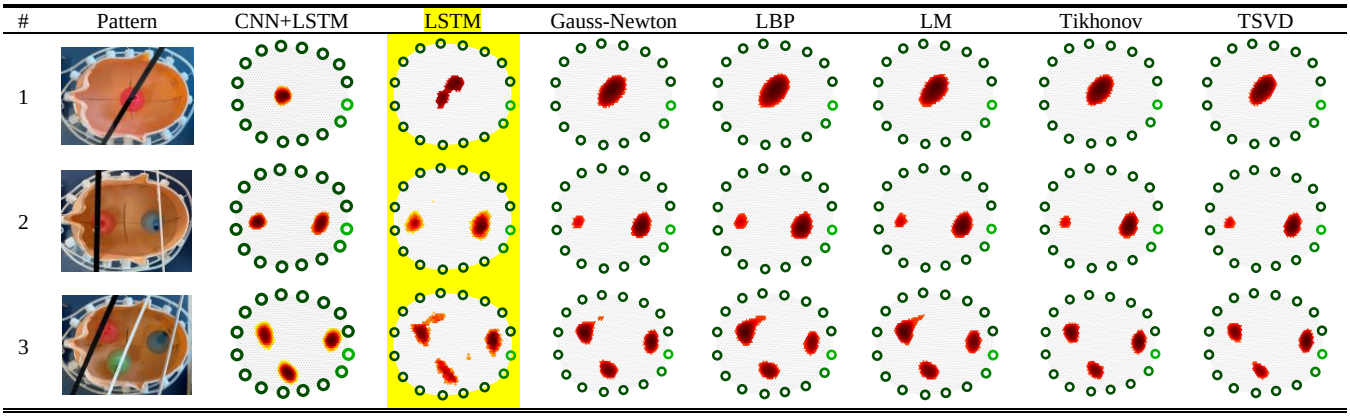

Fig. 10. Comparison of improved reconstructions from actual measurements; images made with deterministic methods (GN, LBP, LM, Tikhonov, and TSVD) were filtered with the Mexican Hat method; CNN+LSTM and LSTM images have been enhanced with a simple cutoff filter.

and Long Short-Term Memory (LSTM) layers, was compared with five popular deterministic methods (Gauss-Newton, LBP, Levenberg-Marquardt, Tikhonov, and TSVD) and a homogenous LSTM network. The development of a properly functioning neural model requires the tuning of numerous parameters. Learnables are variables determined by the network architecture and optimized during training. However, more critical are the preparation of an appropriate training set, the development of network architecture, and the establishment of the training process workflow.

The current study demonstrated that it is possible to train a heterogenous neural model containing CNN and LSTM layers that outperforms both homogenous LSTM and deterministic methods. Particularly noteworthy is that reconstructions obtained by the five deterministic methods differed little from each other. The CNN+LSTM model's greatest advantage was its resilience to measurement disturbances and values deviating from the ideal training data derived mathematically. It is particularly relevant because in the subsequent phase, which will involve clinical trials, it will not be possible to precisely determine the background measurement vector. In such a case, the generalization ability of the CNN+LSTM model will gain importance.

Ultrasound tomography holds great promise for the non-invasive, early, easily accessible, and effective diagnosis of various head diseases. Continued research in this field is necessary to refine these techniques further and expand their clinical applications. Future research will focus on using multi-network models capable of selecting (categorizing) the optimal neural network for use in a specific measurement case.

REFERENCES

[1] G. Bal and J. C. Schotland, “Ultrasound Modulated Bioluminescence Tomography,” *Phys Rev E Stat Nonlin Soft Matter Phys*, vol. 89, no. 3, Oct. 2013, doi: 10.1103/PhysRevE.89.031201.

[2] G. Kłosowski, A. Hoła, T. Rymarczyk, M. Mazurek, K. Niderla, and M. Rzemieniak, “Using Machine Learning in Electrical Tomography for Building Energy Efficiency through Moisture Detection,” *Energies (Basel)*, vol. 16, no. 4, 2023, doi: 10.3390/en16041818.

[3] G. Kłosowski, A. Hoła, T. Rymarczyk, M. Mazurek, K. Niderla, and M. Rzemieniak, “Use of the double-stage LSTM network in electrical tomography for 3D wall moisture imaging,” *Measurement (Lond)*, vol. 213, 2023, doi: 10.1016/j.measurement.2023.112741.

[4] P. E. Jebarani and N. Umadevi, “Grey Wolf Optimization Based Breast Cancer Detection using 1D Convolution LSTM Classifier,” *Przeglad Elektrotechniczny*, vol. 99, no. 1, 2023, doi: 10.15199/48.2023.01.01.

[5] J. Gocławski, J. Sekulska-Nalewajko, and E. Korzeniewska, “Prediction of textile pilling resistance using optical coherence tomography,” *Sci Rep*, vol. 12, no. 1, 2022, doi: 10.1038/s41598-022-23230-9.

[6] K. DINES, “Computerized ultrasound tomography of the human head: Experimental results,” *Ultrason Imaging*, vol. 3, no. 4, pp. 342–351, Oct. 1981, doi: 10.1016/0161-7346(81)90175-9.

[7] P. Grasland-Mongrain *et al.*, “Acousto-electrical speckle pattern in Lorentz force electrical impedance tomography,” *Phys Med Biol*, vol. 60, no. 9, pp. 3747–3757, Jun. 2015, doi: 10.1088/0031-9155/60/9/3747.

[8] B. Tutschek, T. Braun, F. Chantraine, and W. Henrich, “Computed tomography and ultrasound to determine fetal head station,”

Ultrasound in Obstetrics & Gynecology, vol. 49, no. 2, pp. 279–280, Feb. 2017, doi: 10.1002/UOG.17291.

[9] O. T. von Ramm, S. W. Smith, and J. A. Kisslo, "Ultrasound Tomography of the Adult Brain," *Ultrasound in Medicine*, pp. 261–267, 1978, doi: 10.1007/978-1-4613-4021-8_73.

[10] G. Bal and J. C. Schotland, "Ultrasound Modulated Bioluminescence Tomography," *Phys Rev E Stat Nonlin Soft Matter Phys*, vol. 89, no. 3, Oct. 2013, doi: 10.1103/PhysRevE.89.031201.

[11] M. Gross, P. Goy, and M. Al-Koussa, "Shot-noise detection of ultrasound-tagged photons in ultrasound-modulated optical imaging," Apr. 2008, Accessed: Jul. 15, 2023. [Online]. Available: https://arxiv.org/abs/0804.2126v1

[12] P. Grasland-Mongrain *et al.*, "Acousto-electrical speckle pattern in Lorentz force electrical impedance tomography," *Phys Med Biol*, vol. 60, no. 9, pp. 3747–3757, Jun. 2015, doi: 10.1088/0031-9155/60/9/3747.

[13] T. Quang-Huy, T. Duc-Tan, H. H. Tue, and N. Linh-Trung, "Deterministic Compressive Sampling for High-Quality Image Reconstruction of Ultrasound Tomography," Aug. 2015, Accessed: Jul. 15, 2023. [Online]. Available: https://arxiv.org/abs/1508.00647v1

[14] G. Bal, F. J. Chung, and J. C. Schotland, "Ultrasound modulated bioluminescence tomography and controllability of the radiative transport equation," *SIAM Journal on Mathematical Analysis*, vol. 48, no. 2, pp. 1332–1347, Jun. 2015, doi: 10.1137/15M1026262.

[15] C. Cueto *et al.*, "Stride: a flexible platform for high-performance ultrasound computed tomography," *Comput Methods Programs Biomed*, vol. 221, Oct. 2021, doi: 10.1016/j.cmpb.2022.106855.

[16] D. C. Garrett, J. Xu, G. Ku, and L. V. Wang, "Whole-Body Human Ultrasound Tomography," Jun. 2023, Accessed: Jul. 15, 2023. [Online]. Available: https://arxiv.org/abs/2307.00110v1

[17] A. Javaherian, F. Lucka, and B. T. Cox, "Refraction-corrected ray-based inversion for three-dimensional ultrasound tomography of the breast," *Inverse Probl*, vol. 36, no. 12, p. 125010, Dec. 2020, doi: 10.1088/1361-6420/ABC0FC.

[18] N. K. Martiartu, C. Boehm, and A. Fichtner, "3-D Wave-Equation-Based Finite-Frequency Tomography for Ultrasound Computed Tomography," *IEEE Trans Ultrason Ferroelectr Freq Control*, vol. 67, no. 7, pp. 1332–1343, Jul. 2020, doi: 10.1109/TUFFC.2020.2972327.

[19] F. Chung, T. Yang, and Y. Yang, "Ultrasound Modulated Bioluminescence Tomography With A Single Optical Measurement," *Inverse Probl*, vol. 37, no. 1, Aug. 2020, doi: 10.1088/1361-6420/abc8aa.

[20] Y. Bilynsky *et al.*, "Convolutional neural networks for early computer diagnosis of child dysplasia," *Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Srodowiska*, vol. 13, no. 2, 2023, doi: 10.35784/iapgos.3499.

[21] Y. E. Karabacak, "Deep learning-based CNC milling tool wear stage estimation with multi-signal analysis," *Eksploatacja i Niezawodność – Maintenance and Reliability*, vol. 25, no. 3, p. 2023, Jun. 2023, doi: 10.17531/EIN/168082.

[22] X. Zhan, Z. Liu, H. Yan, Z. Wu, C. Guo, and X. Jia, "A novel method of health indicator construction and remaining useful life prediction based on deep learning," *Eksploatacja i Niezawodność – Maintenance and Reliability*, 2023, doi: 10.17531/ein/171374.

[23] H. Dai, M. Penwarden, R. M. Kirby, and S. Joshi, "Neural Operator Learning for Ultrasound Tomography Inversion," Apr. 2023, Accessed: Jul. 16, 2023. [Online]. Available: https://arxiv.org/abs/2304.03297v2

[24] R. Marcinkevičs *et al.*, "Interpretable and Intervenable Ultrasonography-based Machine Learning Models for Pediatric Appendicitis," Feb. 2023, doi: 10.5281/zenodo.7669442.

[25] T. Noda, Y. Jinnai, N. Tomii, and T. Azuma, "Blind Signal Separation for Fast Ultrasound Computed Tomography," Apr. 2023, Accessed: Jul. 16, 2023. [Online]. Available: https://arxiv.org/abs/2304.14424v1

[26] H. Shan, C. Wiedeman, G. Wang, and Y. Yang, "Simultaneous reconstruction of the initial pressure and sound speed in photoacoustic tomography using a deep-learning approach," p. 4, Jul. 2019, doi: 10.1117/12.2529984.

[27] R. Orozco, M. Louboutin, A. Siahkoohi, G. Rizzuti, T. van Leeuwen, and cwinl Felix Herrmann, "Amortized Normalizing Flows for Transcranial Ultrasound with Uncertainty Quantification," *Proceedings of Machine Learning Research-Under Review*, pp. 1–19, Mar. 2023, Accessed: Jul. 16, 2023. [Online]. Available: https://arxiv.org/abs/2303.03478v1

[28] M. Z. Islam, M. M. Islam, and A. Asraf, "A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images," *Inform Med Unlocked*, vol. 20, 2020, doi: 10.1016/j.imu.2020.100412.

[29] A. Agga, A. Abbou, M. Labbadi, Y. El Houm, and I. H. Ou Ali, "CNN-LSTM: An efficient hybrid deep learning architecture for predicting short-term photovoltaic power production," *Electric Power Systems Research*, vol. 208, 2022, doi: 10.1016/j.epsr.2022.107908.

[30] S. Liu, H. Wang, and X. Zhang, "Video Decolorization Based on the CNN and LSTM Neural Network," *ACM Transactions on Multimedia Computing, Communications and Applications*, vol. 17, no. 3, 2021, doi: 10.1145/3446619.

[31] M. Soleimani and T. Rymarczyk, "Ultrasound Tomography for Lung Imaging: An Experimental Phantom Study," *IEEE Sens J*, vol. 23, no. 8, 2023, doi: 10.1109/JSEN.2023.3252340.

[32] P. Koulountzios, T. Rymarczyk, and M. Soleimani, "Ultrasonic time-of-flight computed tomography for investigation of batch crystallisation processes," *Sensors (Switzerland)*, vol. 21, no. 2, 2021, doi: 10.3390/s21020639.

[33] F. Santos and M. Vogelius, "A backprojection algorithm for electrical impedance imaging," *SIAM J Appl Math*, vol. 50, pp. 216–243, 1990.

[34] R. A. Rahim, M. H. F. Rahiman, and M. N. Mohd Taib, "Non-invasive ultrasonic tomography: Liquid/gas flow visualization," *2005 1st International Conference on Computers, Communications and Signal Processing with Special Track on Biomedical Engineering, CCSP 2005*, pp. 243–247, 2005, doi: 10.1109/CCSP.2005.4977199.

[35] G. Kłosowski, T. Rymarczyk, K. Niderla, M. Kulisz, Ł. Skowron, and M. Soleimani, "Using an LSTM network to monitor industrial reactors using electrical capacitance and impedance tomography – a hybrid approach," *Eksploatacja i Niezawodność – Maintenance and Reliability*, vol. 25, no. 1, p. 2023, Jan. 2023, doi: 10.17531/EIN.2023.1.11.

[36] J. Schwab, S. Antholzer, R. Nuster, G. Paltauf, and M. Haltmeier, "Deep Learning of truncated singular values for limited view photoacoustic tomography," p. 111, Jan. 2019, doi: 10.1117/12.2508418.

[37] P. Müller, M. Schürmann, and J. Guck, "The Theory of Diffraction Tomography," Jul. 2015, Accessed: Jul. 19, 2023. [Online]. Available: https://arxiv.org/abs/1507.00466v3

[38] T. Lukić, J. Lindblad, and N. Sladoje, "Regularized image denoising based on spectral gradient optimization," *Inverse Probl*, vol. 27, no. 8, p. 085010, Jul. 2011, doi: 10.1088/0266-5611/27/8/085010.

[39] M. Gehre, B. Jin, and X. Lu, "An Analysis of Finite Element Approximation in Electrical Impedance Tomography," *Inverse Probl*, vol. 30, no. 4, Dec. 2013, doi: 10.1088/0266-5611/30/4/045013.

[40] C. Li, M. Karamehmedović, E. Sherina, and K. Knudsen, "Levenberg-Marquardt algorithm for acousto-electric tomography based on the complete electrode model," *J Math Imaging Vis*, vol. 63, no. 4, pp. 492–502, Dec. 2019, doi: 10.1007/s10851-020-01006-y.

[41] R. Zaheer and H. Shaziya, "A Study of the Optimization Algorithms in Deep Learning," in *Proceedings of the 3rd International Conference on Inventive Systems and Control, ICISC 2019*, 2019. doi: 10.1109/ICISC44355.2019.9036442.

[42] N. H. Lestriandoko and R. Sadikin, "Circle detection based on hough transform and Mexican Hat filter," in *Proceeding - 2016 International Conference on Computer, Control, Informatics and its Applications: Recent Progress in Computer, Control, and Informatics for Data Science, IC3INA 2016*, 2017. doi: 10.1109/IC3INA.2016.7863041.

[43] W. Fu, Y. Sheng, and D. Xiong, "Optimal reconstruction design for JPEG-coded image using structural similarity index," *Proceedings of the 2017 12th IEEE Conference on Industrial Electronics and Applications, ICIEA 2017*, vol. 2018-February, pp. 180–184, Feb. 2018, doi: 10.1109/ICIEA.2017.8282837.