AIP Advances

Localizing a target inside an enclosed cylinder with a single chaotic cavity transducer augmented with supervised machine learning

Cite as: AIP Advances **11**, 115104 (2021); https://doi.org/10.1063/5.0068803
Submitted: 27 August 2021 • Accepted: 08 October 2021 • Published Online: 02 November 2021

Tom Sillanpää, Krista Longi, Joni Mäkinen, et al.

Localizing a target inside an enclosed cylinder with a single chaotic cavity transducer augmented with supervised machine learning

Cite as: AIP Advances 11, 115104 (2021); doi: 10.1063/5.0068803
Submitted: 27 August 2021 • Accepted: 8 October 2021 •
Published Online: 2 November 2021

Tom Sillanpää,[1,a)] Krista Longi,[2] Joni Mäkinen,[1] Timo Rauhala,[3] Arto Klami,[2] Ari Salmi,[1] and Edward Hæggström[1]

AFFILIATIONS

[1] Department of Physics, University of Helsinki, P. O. Box 64, Helsinki FI-00014, Finland
[2] Department of Computer Science, University of Helsinki, P. O. Box 68, Helsinki FI-00014, Finland
[3] Altum Technologies, Eteläranta 8, Helsinki FI-00130, Finland

[a)] **Author to whom correspondence should be addressed:** tom.sillanpaa@helsinki.fi

ABSTRACT

Ultrasound is employed in, e.g., non-destructive testing and environmental sensing. Unfortunately, conventional single-element ultrasound probes have a limited acoustic aperture. To overcome this limitation, we employ a modern method to increase the field-of-view of a commercial transducer and to test the approach by localizing a target. In practice, we merge the transducer with a chaotic cavity to increase the effective aperture of the transducer. In conventional pulse-echo ultrasound signal analysis, location estimation is based on determining the time-of-flight with known propagation speed in the medium. In the present case, the dispersing field induces complexity to this inverse problem, also in 2D. To tackle this issue, we use a convolutional neural network-based machine learning approach to study the feasibility of employing one single chaotic cavity transducer to localize an object in 2D. We show that we indeed can localize an inclusion inside a water-filled cylinder. The localization accuracy is one diameter of the inclusion. The area that we can infer increases by 49% in comparison to using the same transducer without applying the proposed chaotic cavity method.

I. INTRODUCTION

Conventional ultrasound probes, used in many applications, such as in non-destructive testing (NDT) and environment sensing applications, have a limited acoustic aperture. With traditional ultrasound probes, one needs to scan across the area of interest or to use a phased array to obtain a wide field-of-view (FOV). Localizing targets inside a volume is important in modern applications such as self-driving vehicles[1] and other location aware systems.[2]

In NDT, it is preferable to use less complex systems. The available computational power allows translating complexity from hardware to software. Methods to decrease the complexity of the ultrasound phased array hardware exist.[3–5] These methods include time-reversal[6] (TR) and compressive sensing[7] (CS). The first approach combines acoustic TR and a transducer or a small array of piezoceramics with a chaotic cavity (CC).[3,4,8–12] This combination can focus ultrasound into a small area/volume in 2D and 3D.[3,10] A CC can be a block of material with a specific geometric shape (e.g., chaotic billiard) or a structure with complex internal structure (resembling a metamaterial), which ensures ray-chaotic dynamics. This kind of dynamic behavior mixes propagation ray angles, which ensures unique reverberations inside the CC. Sending a pressure wave through a CC results in dissimilarities in the signals even with small changes in initial conditions. In this paper, the fixed orientation and static positioning of the CC and transducer combination leaves only one possible cause for the changes in the signals arising from the pressure wave propagation in the media. This cause is, e.g., reflection from an object. The downside of this approach is that reverberating media induce complexity into the ultrasound

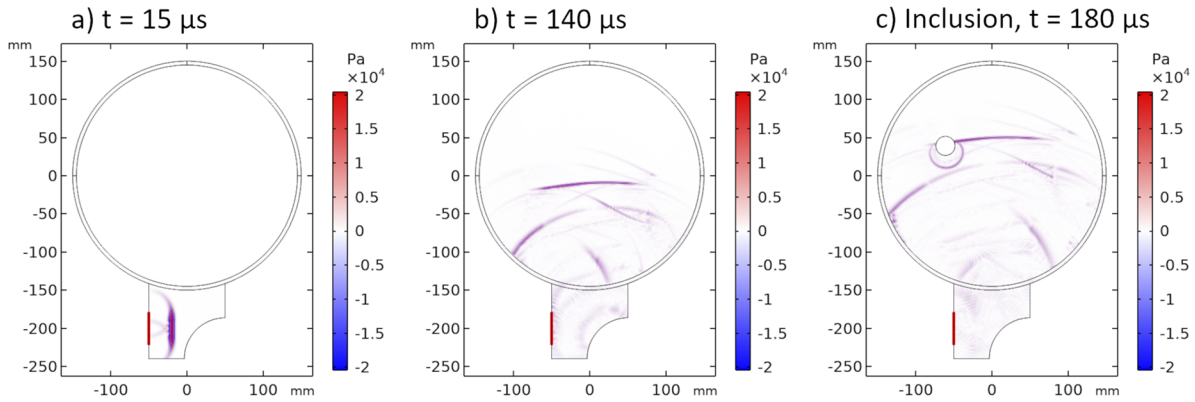

FIG. 1. Frames from the FEM simulations show the pressure wave transmitted through the CC, after which it probes almost the entire water-filled container (not containing an inclusion) and eventually reflects back from the opposite wall back toward the CC. (a) A three-cycle Gaussian-windowed pressure wave was created on the surface of the CC wall, mimicking the one used in the measurements. (b) The wavefront of the pressure wave propagates through the CC and container wall and to the water domain, simultaneously expanding to cover nearly the entire area of the container. (c) If an inclusion is placed in the cylinder, the wavefront reflects back from it toward the CC, and eventually, a signal is captured from the same location as the pressure wave was generated.

signals, which affects signal analysis and makes signal interpretation by conventional means harder.

The Sinai billiard is a type of chaotic billiard, in which a circle is removed from the center of a square. This kind of reverberating cavity reduces the number of driven antennas in an ultra-wideband (UWB) imaging system with microwaves.[13] In our current study, we explored, using finite element method (FEM) simulations and experiments, whether a section of this kind of CC coupled to a transducer is a feasible sensor for obstacle detection in the acoustic domain. We simulated a pressure wave propagation in two sensing configurations (SC I and II), and the geometric details are shown in Figs. 3(a) and 3(b). In SC I, a pressure wave is sent through an acrylic block of the shape of a quarter Sinai billiard, from which the wave leaks into a water-filled cylinder. A few frames captured at different times from the simulations are displayed in Fig. 1. They show the dispersing property of this CC with the applied initial conditions. A three-cycle Gaussian enveloped pressure wave is emitted from the wall of the CC (facing the convex surface of the CC), and the propagating wave hits the dispersing surface and creates a wide range of propagating ray angles (creating reverberations in the CC). The beam exiting the CC into the water-filled cylinder is widened and spans almost the entire area of the cylinder. In the simulation, an inclusion was placed inside the water-filled container (see Fig. 1), and by carrying out a pulse-echo measurement, the echoes propagating back to the transmitting location were qualitatively analyzed. This confirmed that the echo from the inclusion propagates back to the transmitting location (the pressure wave transmitted propagates through the CC, leaks into the water container, reflects back from the inclusion, and propagates back to the transmitter location via a reflection from the convex surface of the CC).

The chaotic billiard both at transmission and at reception makes conventional modeling of the reflecting waves impossible. Instead, we are interested in flexible data-driven models, in particular deep neural networks, trained on acoustic measurements for the specific task of localizing the inclusion. By directly learning to perform the task of interest, instead of learning to model the complete ultrasonic field, we can bypass the need for direct modeling of the chaotic element and hence reduce much of the amount of training data required.

Deep learning methods are good at learning internal structures from high dimensionality data and have therefore led to improvements in, e.g., speech recognition and image processing.[14] Especially, convolutional neural networks (CNNs) have been successfully applied to data with temporal or spatial structure, such as images,[15] audio,[16] and wireless signals.[17] CNNs extract progressively higher-level features from the raw measurement signal with convolutional filters and subsequent nonlinear activations, and they are trained in an end-to-end manner based on training measurements with known ground truths. These CNNs are common in image processing where spatial (2D or 3D) convolutions are used but are equally appropriate for temporal measurement signals represented by real vectors when 1D convolutions are employed.[18]

Here, in an NDT manner, we localize an internal structure from the outside of an enclosed container by means of a CC transducer. This study researches the ability of the CC to extend the FOV of a commercial probe and uses a CNN-based machine learning algorithm for localization. We built an experimental setup based on the proposed method and carried out measurements for training the CNN and for evaluating the localization. We conducted FEM simulations to validate the feasibility of the proposed method.

II. MATERIALS AND METHODS

A. Measurements

The experimental setup contains an acrylic CC glued with epoxy (Bison® Epoxy 5 Minutes Blister, 6305448) to the outer wall of an acrylic, cylindrical container (inner diameter Ø = 290 mm, a wall thickness of 5 mm, and a height of 298 mm). The gluing of the CC to the acrylic container was carried out carefully to assure vertical alignment and to avoid air bubbles between the structures. The bottom of the CC was 154 mm above the container floor. An acrylic

pipe (outer diameter Ø = 25 mm and a wall thickness of 2 mm) was used to simulate an internal structure. The setup is displayed in Fig. 2.

The cylinder was filled with tap water, and the measurements were conducted at room temperature (T = 21 °C). Broadband ultrasonic pulses were generated with a pulser (Olympus 5058PR) and a commercial flat focused transducer (Karl Deutsch S24 HB 0.3–1.3 MHz). The mean frequency of the excitation signal was f_{mean} = 800 kHz (a bandwidth of 400 kHz). The echoes were recorded with the same transducer using an integral pre-amplifier (40 dB) of the pulser and a Picoscope 3403D as an oscilloscope. The signal transmission was triggered by a microcontroller (Arduino® Uno). The motorized arm moving the inclusion (acrylic pipe) was built based on OpenBuilds® ACRO system 55 with an additional custom frame to put the XY-stage at a suitable height above the container. The motorized arm was controlled via serial communication, and the measurement was automated together with signal acquisition in MATLAB® (R2018a). The precision of the motorized arm was 0.3 mm.[19] To find the center point of the acrylic container, two points at the inner circumference of the container were determined in the translation stage coordinates (with 0.5 mm precision), and by calculating the intersection point of the two circles (with the same radius as the container), the center of the container was determined. The uncertainty of positioning of the motorized arm was 1 mm.

In both cases, regardless of the CC being applied or not, the same randomly generated (within the limits of a square that fits inside the container) location pattern of the inclusion inside the cylinder was used. From the pulse-echo signals, the feature of a signal that corresponds to the first arriving echo from the inclusion was identified. In the latter case, the transducer was attached to the container wall at the same height as the CC, but on the opposing side of the CC, the location pattern (of the acrylic pipe) was rotated around X, Y with respect to the transmit position, although the location pattern was the same in the translation stage coordinates. In total, 200 signals were captured at each of the 150 locations of the acrylic pipe inside the cylinder. The area of detection plots was produced in both cases, where the envelope of the signals [log 10(envelope)^2] was integrated over chosen time windows and normalized. As a result, the integrational value is proportional to the energy reflected back from the inclusion.

B. Finite element method

FEM simulations were used to display the acoustic wavefront propagation in and out of a water-filled container with and without the CC transducer. The simulations were performed using COMSOL Multiphysics®.[20] COMSOL's Pressure Acoustics and Solid Mechanics modules were used.

The general shape of the simulation geometry is displayed in Fig. 3(a), where the configuration for the chaotic cavity attached to the water-filled cylinder without any inclusion is presented. The inclusion was considered in a set of simulations by removing a circular region from the water domain and setting the boundary condition of these regions to be sonically hard boundaries. In addition, a set of simulations with no CC was run. In these simulations, the CC was substituted by an acrylic block with identical dimensions used in the CC block, although without the CC element (shape of a quarter Sinai billiard), as shown in Fig. 3(b).

The simulations were carried out in 2D to approximate the measured system. In the 2D approximation, the wave propagation differs due to the lack of geometric attenuation compared to the experimental measurements. This was most prominent in the case of the chaotic cavity where a part of the wave front was reflecting inside the CC for a much longer time than was expected. To reduce the coda generated by these reflections, parts of the cavity edges were modeled as low-reflecting boundaries [marked with blue lines in Fig. 3(a)]. The transducer was modeled as a boundary displacement in the simulation where a Gaussian modulated sine wave with f_c = 700 kHz central frequency [Fig. 3(a)] was used because the most prominent peak in the frequency of the experimental signal was close to 700 kHz rather than being close to the central frequency (800 kHz) of the transducer.

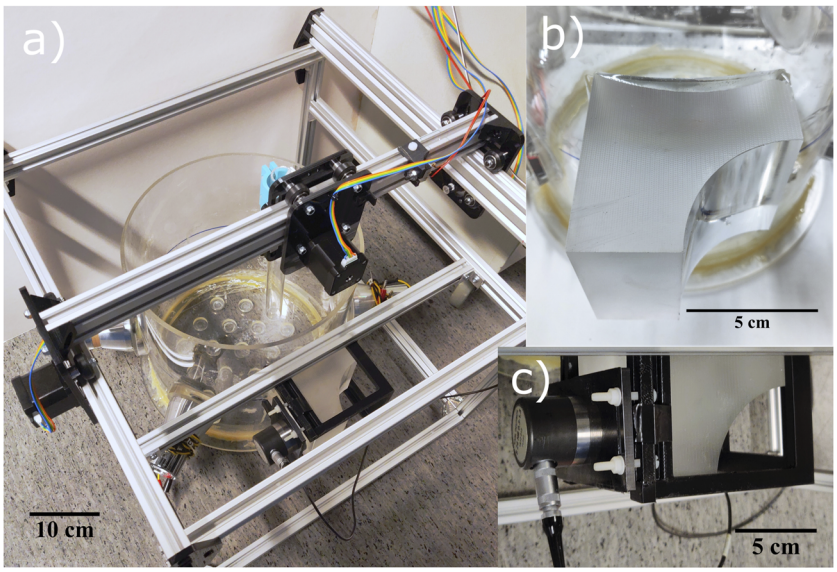

FIG. 2. (a) Setup with a robotic arm to translate the plastic pipe inside the cylinder. (b) CC glued with epoxy to the outer wall of the cylinder. (c) Transducer attachment to the CC.

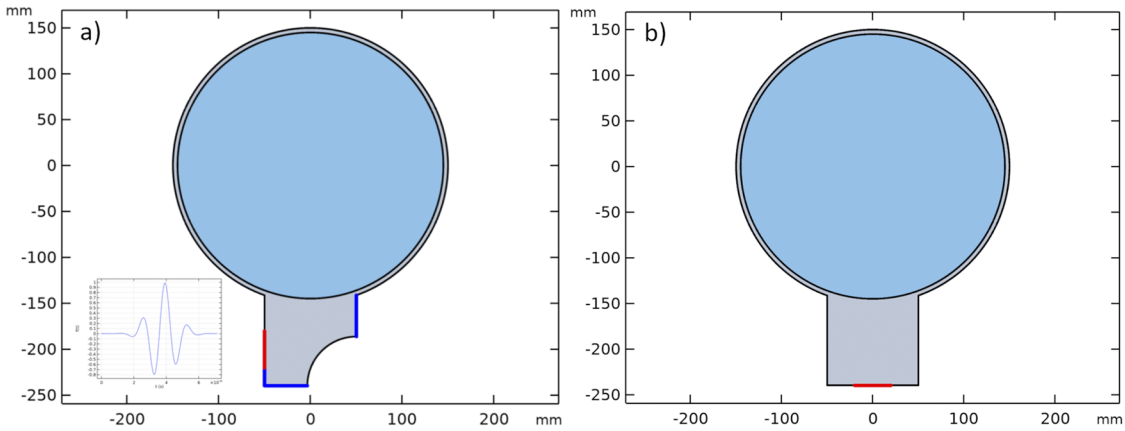

FIG. 3. (a) Simulation geometry with the chaotic cavity (SC I). Inset shows the transmitted signal used in the simulations. The signal is a 700 kHz sine wave modulated by a Gaussian. (b) Acrylic block with identical dimensions to the CC but without the CC element (SC II). The red boundary is set as a boundary displacement to model the transducer. Blue boundaries are set as low-reflecting boundaries. The material of the gray area is acrylic while the blue inside the cylinder is water.

The mesh size in the simulations was calculated from a maximum resolved frequency, which was set to $f_{\max} = 1.5f_c$. The mesh size in the geometry was set to $dx = c/(5f_{\max})$ while using a quadratic Lagrangian discretization for the mesh. The solver was chosen to be generalized alpha in order to set a fixed time stepping scheme. The time step was set to $dt = 0.1dx/c$.[21] The material parameters used in the simulations were obtained from COMSOL's material library for water, c = 1481 m/s and ρ = 998 kg/m^3. For acrylic, Poisson's ratio was set to 0.37,[22] while the density was measured from the volume and mass of a well-defined acrylic piece, ρ = 1010 kg/m^3. Young's modulus, E = 4.1 GPa, was calculated from the speed of sound, density, and Poisson's ratio. The longitudinal speed of sound, c = 2680 m/s, was determined from a pulse-echo measurement of the acrylic piece.

C. Supervised machine learning

1. Convolutional neural network

Our CNN-based machine learning algorithm locates the inclusion based on the captured ultrasound signals. The CNN consists of two 1D convolution layers with max pooling as well as one fully connected layer followed by a dropout and an output layer, as shown in Fig. 4. After each layer, we have an activation function that performs a nonlinear transformation of the outputs, and before each layer, we use batch normalization. The loss function to be minimized is the mean squared error between the predicted location and the true location.

The CNN performance depends on multiple hyperparameters that determine the convolutional filter sizes and strides, the nonlinear activation functions, and the dropout parameter. To determine these values that define the network architecture, we used hyperas[23] with the Tree of Parzen Estimator (TPE) algorithm and keras[24] to find the architecture that achieves the lowest error, separately for the two sensing configurations. To determine a single network structure, we used 110 randomly selected locations for training and 40 for evaluating, running the hyperparameter optimization for 100 different networks, each trained for 10 epochs. For both cases, signals obtained with and without the CC, we chose the network with the best performance (achieves the lowest error) on the test set. The experimented hyperparameter values and the chosen structures are

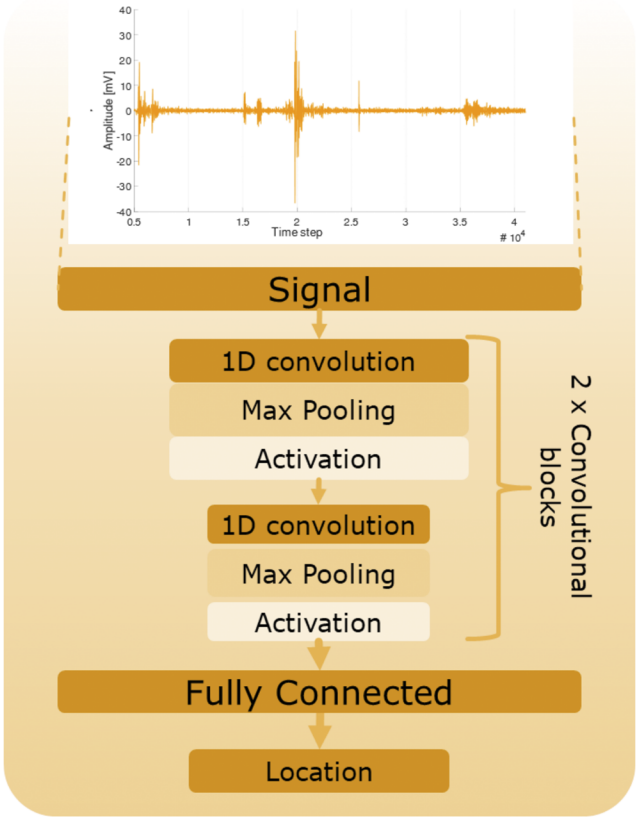

FIG. 4. Structure of convolutional neural network.

TABLE I. List of experimented hyperparameter values and chosen structures for the CNN.

	Chaotic cavity	Without chaotic cavity	Allowed values for hyperparameter optimization
		1D convolution	
Filters	24	24	
Kernel size	1000	600	300, 400, 600, 1000
Strides	20	20	1, 3, 5, 20, 50, 60
Activation	tanh	tanh	tanh, softmax
		Max pooling	
Pooling size	50	200	10, 25, 50, 200, 300
Strides	30	30	1, 3, 5, 10, 30, 50, 60
		1D convolution	
Filters	24	24	
Kernel size	70	70	5, 10, 20, 30, 50, 70
Strides	1	3	1, 2, 3, 5
Activation	Softmax	Softmax	tanh, softmax
		Max pooling	
Pooling size	1	2	1, 2
Strides	2	1	1, 2
Dense layer units	128	128	
Dropout	0.252	0.098	
Activation	relu	relu	relu, tanh

presented in Table I. Note that many different networks performed on a similar level.

After choosing the network structure, we used a cross-validation procedure to evaluate the localization accuracy. We partitioned the 110 data locations that were used as a training set for hyperparameter selection into subsets containing 10 locations each. One subset was kept as evaluation data, while other 10 subsets and the previously used test set of 40 points were used to train the CNN. We repeated this 11 times, using each subset as evaluation data exactly once, and ran the algorithm for 50 training epochs. Importantly, the locations used as the test set for hyperparameter optimization were not used as evaluation data for the final model to avoid re-using the same locations as test instances.

One signal in our dataset is an average of 60 measured signals (out of 200) so that we have 141 slightly different signals for each target location. For each signal, and from each point, we deduct the average value of 100 neighboring points (moving average). We use zero padding for the convolutional layers. Training was carried out using the Adam optimizer[25] with a batch size of 100, and implementations were performed using Tensorflow.[26]

2. Gaussian process

The above-mentioned process provides an estimate of the localization accuracy for each of the 110 training locations. To obtain estimates for the localization accuracy for the whole area of interest, we fitted a Gaussian Process (GP) regressor[27] to visualize the distances between the true location and the predicted location with the trained CNN. GPs are one way of modeling spatial phenomena and providing a way to consider the uncertainty of the predictions. A Gaussian process $f \sim GP(m(x), k(x, x'))$ is specified by a mean function $m(x)$ and a symmetric positive-definite kernel function $k(x, x')$. The kernel function defines the prior covariance between any two function values. With a GP, one can estimate the uncertainty of the predictions of the CNN model outside the measured locations as well based on the assumption that the localization accuracy varies smoothly. Likewise, the operation was also carried out to produce Fig. 7, where the integral of the signal's envelope within the described time window at each location of the inclusion was the input. We fitted the GP using GPy,[28] with a zero mean function and a squared exponential kernel, also known as a radial basis function kernel,

$$k(x, x') = \sigma^2 \exp\left(-\frac{\|x - x'\|^2}{2l^2}\right). \quad (1)$$

Here, σ and l are the variance and lengthscale of the kernel, respectively, and we learn these to maximize the marginal likelihood of the data using the L-BFGS-B optimizer in GPy.

III. RESULTS

Three examples of measured pulse-echo signals (averaged over 200 signals from each location) obtained through the CC with different locations of the inclusion are plotted in Fig. 5 (left). They

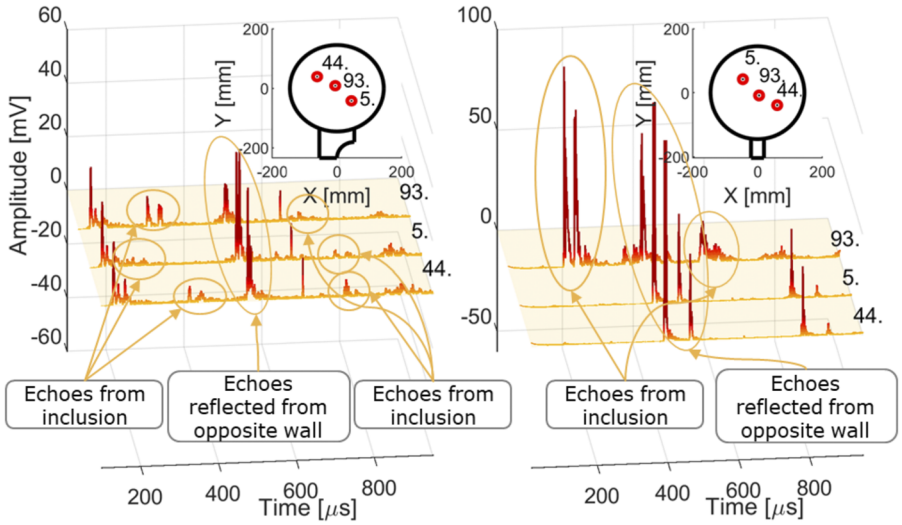

FIG. 5. Envelope of pulse-echo signal obtained through the chaotic cavity (left) for three different positions of the inclusion shows distinguishable features between the different positions. The figure (right) shows measured signals at similar locations without the CC and displays the most prominent echo peak when the inclusion is at the center. No detectable features are present when the inclusion is off the acoustic axis of the transducer. Differences in echo amplitudes are seen in the secondary wall reflection (waves propagated two times forth and back through the container) if the inclusion is located on the acoustic axis of the CC transducer or the normal transducer.

display the ultrasonic signal, which differs as a function of the location of the internal structure. Since the ultrasound beam diverges when it exits the CC, the location information from the inclusion is no more valid when obtained by the conventional time-of-flight analysis. At the beginning ($100 < t < 200$ μs), complex echoes from the chaotic cavity-container wall–water interfaces arrive, comprising several reverberations inside the CC. Then, in the time range $200 < t < 450$ μs, detectable echoes from the inclusion may arrive. An echo of the reflection from the opposite wall (opposite to the CC) arrives at $450 < t < 550$ μs, featuring a relatively high amplitude. Secondary reflections from the inclusion are visible at $550 < t < 850$ μs, and secondary reflections from the opposite wall are seen at the end of the signal ($850 < t < 950$ μs). For comparison, an identical transducer, without the CC, was placed on the container wall opposite to the CC, and an identical set of measurement signals were obtained [see Fig. 5 (right)]. It showed a high amplitude echo from the inclusion, when it was placed in front of the beam as expected (position 93). No interpretable echoes are noticed when the inclusion is moved away from the acoustic axis (positions 5 and 44).

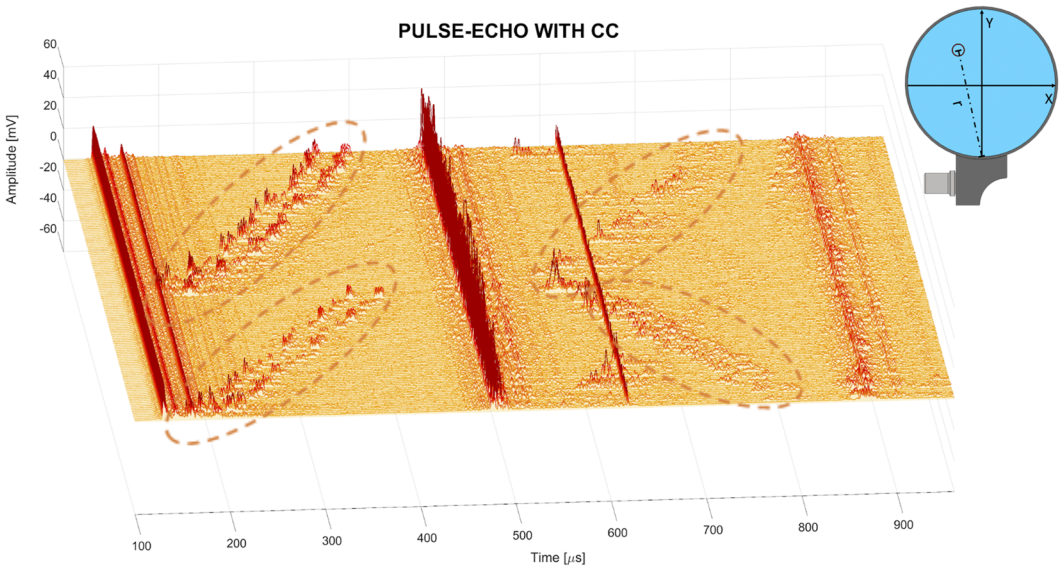

FIG. 6. Envelope of signals received from 150 locations of the inclusion inside the water container. Signals were obtained with the pulse-echo method through the CC. Echoes from the inclusion are emphasized by ellipses that guide the eye. The signals in this figure are sorted in such a way that the distance r (see the top right corner) between the location of the inclusion and the CC (middle) decreases both in the left side (signals top to center) and right side (signals center to bottom).

A. Experiments

To compare an area of detection, with and without the CC, ultrasound signals were acquired at the generated location pattern of the inclusion. An envelope of the signals with the CC is plotted in a waterfall plot in Fig. 6. These signals are sorted for the left (−x-axis) and right (+x-axis) half of the container, and according to the increased distance r depicted in the inset of the figure. The time window ($150 < t < 460\ \mu s$) corresponds to the time range when the first scattered echo from the inclusion arrives back to the transducer. The envelope of the averaged signal [log10(envelope)^2] was first integrated across the corresponding time window and then min–max normalized. The integration result is proportional to the energy of the echo captured with the pulse-echo method. Then, an interpolated surface was fitted to these integral values at the locations with the GP method as prescribed in the section titled "Materials and methods." The result is shown in Fig. 7(a), which depicts a higher signal (red area) in front of the CC than the signals obtained closer to the edges of measured area (blue). For comparison, the same process was conducted without the CC [see Fig. 7(b)]. In that case, the transducer was attached to the outer container wall. A time window ($70 < t < 380\ \mu s$) of identical length to the latter one was chosen, except that it lacks a time delay needed for sound to propagate through the CC. This demonstrates the narrow beam produced by the transducer, which is expected for the straight beam probe.

B. Supervised machine learning

The CNN localization accuracy was analyzed using a cross-validation procedure. The location estimates of the model are shown

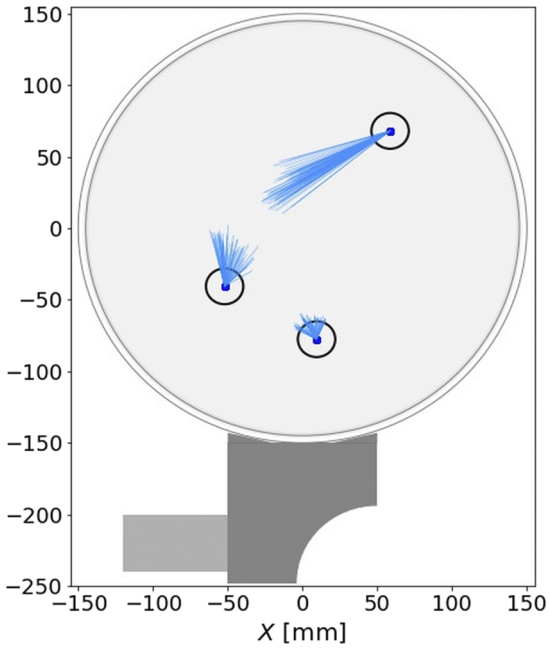

FIG. 8. Each blue line corresponds to a validation test of the CNN with a test signal (141 test signals per location of the inclusion). The blue lines point toward the CNN location estimate at three example locations of the inclusion (marked with a circle and center point). The bottom location agrees closely with the ground truth. The top right location indicates the difficulty of our trained CNN to predict the location for areas where low SNR signals are obtained [see Fig. 7(a)].

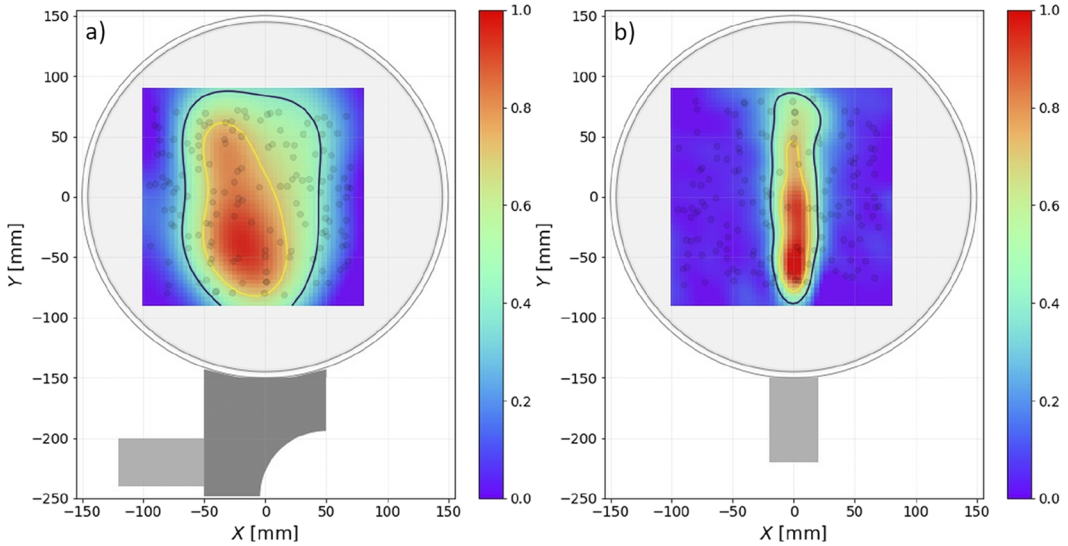

FIG. 7. (a) Dispersed beam through the chaotic cavity increases the FOV inside the water container. (b) The FOV of the ultrasound transducer attached straight to the outside wall of the cylindrical water container. The blue contour line represents 0.4 and the yellow contour line represents 0.7 of the maximum energy reflected back in both cases. The blue dots represent the 150 locations of the inclusion at the pulse-echo measurements.

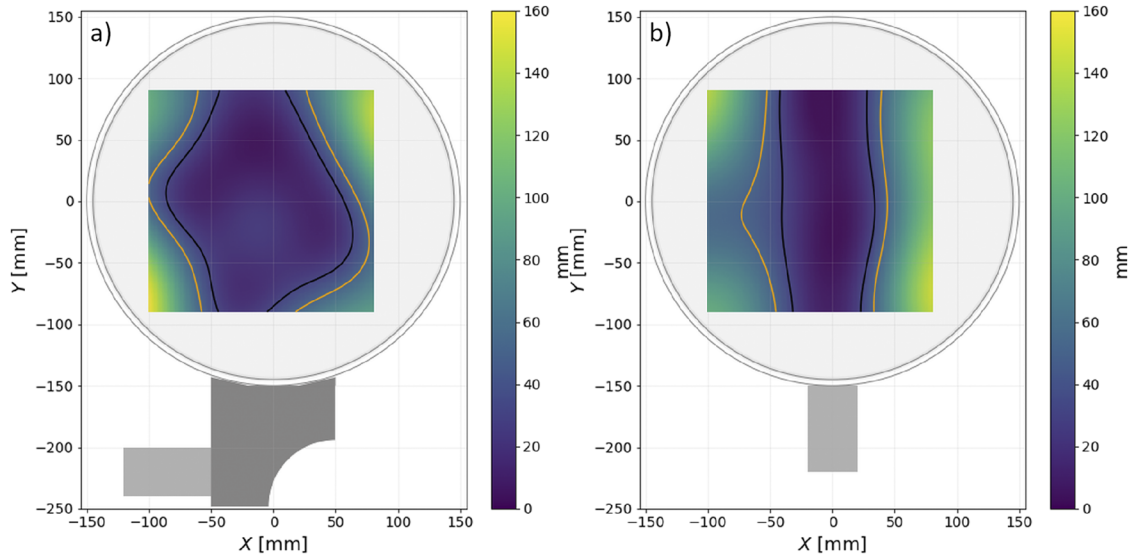

FIG. 9. Mean of the GP estimate as a function of distance between the true location and predicted location (a) with CC and (b) without CC. The black contour line represents 30 mm distance, and orange contour line, 50 mm. These images show the advantage of using the CC transducer to locate the internal structure with only one transducer. As the diameter of the inclusion was 25 mm, the black contour lines confine an area where the error in location detection is similar to the diameter of the inclusion.

(blue lines pointing from the ground truth to the estimated location) for three example locations that were not used to train the model in Fig. 8. The localization precision is affected by the location of the inclusion [see Fig. 7(a)].

We fit a GP regressor[29] to visualize the average distance between true locations and the predicted locations with the trained CNN. We plotted the mean of the GP fit in both cases with the CC in Fig. 9(a) and without the CC in Fig. 9(b). These plots demonstrate an improvement in the target location close to the edges of the 324 cm^2 area (area of a square determined between minimum and maximum measurement points in x and y directions) when we used the CC transducer: the area inside the 30 mm contour line (black) in the plotted area ($x \in [-100, 80]$ and $y \in [-90, 90]$) is 49% larger than that with solely the standard transducer (186 vs 125 cm^2).

The mean error of the location output was 29 ± 26 mm when using the CC and 44 ± 35 mm without it. The cross-validation results are shown in Table II. As a reference, the true location is measured with 2 mm accuracy, and the internal structure has a diameter of 25 mm. Consequently, the uncertainty in location estimation is on the order of the size of the inclusion.

TABLE II. Summary of cross-validation error mean and std of the localization results.

	Error mean (mm)	Std (mm)
With CC	29	26
Without CC	44	35

IV. DISCUSSION

Compared to employing a standard transducer, by using a CC transducer, we predicted locations of inclusions accurately on a larger area (186 vs 125 cm^2). The increase in the inferenced area is mainly due to lateral widening of the ultrasound field. In both cases, with and without the CC, the machine learning based location estimates suffered larger uncertainty at locations further away from the acoustic axis. This is due to the lower signal amplitude in these locations (see Fig. 5). We trained and validated the CNN by averaging 60 signals and ended up having 141 slightly different training signals at each location of the inclusion (the total number of raw signals per location was 200). We assume that averaging each raw signal more and other possible improvements to the SNR could help the CNN extract higher-level features that would improve the location information. Getting noiseless data from experiments is unlikely; thus, we also considered teaching the CNN by signals obtained with the FEM. This approach was unfortunately infeasible for a real-life scenario due to an extensive computation time (one day per signal with the available computation resources). The possible influence arising from the changing effective area of the sensor was initially addressed by FEM simulations in the two sensing configurations (SC I and II). The results with the CC showed agreement in the pulse-echo signals with the three locations of the inclusion in comparison to the experimentally obtained signals. In addition, in the simulations with SC II, the signal amplitude from the inclusion diminished below the detection limit (signal amplitudes from the inclusions aside from the central axis of the generated beam were roughly −40 dB lower than those in the central location). Due to this, we assumed that even without the delay block (not a CC type), the signals would not be detectable and simplified

the experimental design by not having a separate acrylic container nor a delay line without the CC element for SC II (see Sec. II B). The simulation results also suggest that the increased area of target localization is not related to the effective area of the transducer. On the contrary, they indicate that the improved results arise from the dispersing properties of the proposed CC method as expected [see Fig. 1(c)].

The current study concerned a simple setup where only one inclusion is inserted into the sensor's detection range. The case of several inclusions is outside the scope of this study. In more realistic scenarios, several reflecting boundaries, e.g., many different size inclusions, need to be studied. The studied inclusion is larger than the resolution of the ultrasound pulse (wavelength in water 1480 m/s/700 kHz = 2.1 mm); hence, its size could be decreased and tested to determine the minimum resolvable features that could be attained with this type of CC transducer. This information could then be brought into the CNN's training to extract size information about the inclusion. In conjunction with our previous study,[18] other kinds of feature extraction with the CNNs could be plausible, e.g., fouling detection from the obtained signals.

By definition, the action of CCs depends on the initial conditions, where a small perturbation (e.g., a different placement on the cylinder wall) may change the characteristics of the CC sensor. In this study, we reduced this effect by gluing the CC block onto the surface of the container. This approach prevented us from studying how the CC works if the attachment of the CC sensor is changed and, moreover, how such uncertainties affect the performance of the AI. The ultrasound beam exiting the current CC was axially more sensitive on one side. This feature can be altered by modifying the shape of the CC and allows us to design asymmetric sensors that are feasible in applications where attention needs to be focused to specific areas.

Our machine learning model was a relatively shallow CNN since the more flexible network structures were ruled out by the small amount of training data, resulting in merely 150 distinct locations. With a richer training collection, deeper CNNs as well as recurrent neural networks would be applicable. These could be developed with a protocol that combines network structure specification with automatic hyperparameter optimization for fine-tuning. Having access to a larger data collection would most likely extend the area of sufficient localization accuracy further into the peripheral areas of low signal amplitude, but extremely low signal amplitudes would still remain challenging.

An important notion for machine learning approaches is the generalizability of the model. In this work, we carefully evaluated the accuracy of the localizing performance of the CNN with a dataset that was left out from the model training, providing a within-data estimate for the generalization ability. Since the experiment was conducted on one specific structure (CC attached to a water-filled cylinder), it provides no information on how the results generalize to, e.g., different structures or chaotic elements. Questions such as this can only be addressed with experimental setups that cover multiple structures. Our expectation is that the proposed model architecture and training procedure would generalize for new structures and other minor modifications of the measurement setup with similar qualitative performance, but the neural network would need to be re-trained (including hyperparameter search) for each one specifically. Training a network to perform localization inside arbitrary future structures with no re-training is unrealistic, but solutions that require less data for new structures could be developed based on transfer learning and domain adaptation.[28]

The FEM simulations were performed with a geometry identical to that used in the measurements although limited to 2D. By simplifying the model to 2D, the computation time is reduced significantly at the cost of losing wave propagation properties of the actual 3D case, e.g., geometric attenuation. One other difference between the FEM simulations and the real case is that the inclusion is modeled as a perfect reflector instead of as a water-filled structure (acrylic pipe) where the ultrasound can propagate through. This transmission property is realized as a feature in the differences between the signals obtained with measurement or FEM simulations. In measured signals, the echos arising from the inclusion (features emphasized in Fig. 5) have two high peaks adjacent in time, whereas in the FEM simulations, only the first peak is observed (as the inclusion was modeled as a circle with an acoustically hard boundary). Despite that difference, we observed agreement between the signals obtained with both measurements and FEM simulations. To the authors' knowledge, this is the first paper that introduces this kind of CC transducer coupled with AI-based location detection for potential use in NDT and environment sensing applications. This method nearly doubles the FOV of a commercial transducer in the lateral dimension. The method reduces the number of required probes, thereby reducing costs and complexity of operation.

V. CONCLUSIONS

First, the concept of applying a CC to a transducer to increase the FOV of an ultrasound probe was studied in FEM simulations. Based on those findings, we built an experimental setup to test the CC transducer as an acoustic location detector for, e.g., NDT or environment sensing applications. The inverse problem arising from the dispersing ultrasound field is tackled by training a CNN-based machine learning algorithm that predicts the object location. Labeled data for teaching the AI algorithm were acquired, and the models were trained and evaluated using a cross-validation setup. The trained CNN showed a precision that was on the order of the defect's diameter. The combination of the CC transducer and machine learning demonstrates a capability for object localization with a 49% increase in the detection area inside an enclosed container in comparison to using the transducer alone.

ACKNOWLEDGMENTS

The Academy of Finland (Grant No. 324852) is acknowledged for funding.

AUTHOR DECLARATIONS

Conflict of Interest

T.R. works for Altum Technologies, a company that produces industrial ultrasound cleaning systems. A.S. and E.H. have received compensation as scientific advisors from the company. The company provided some of the experimental instruments for this research. T.S., K.L., J.M., and A.K. declare no potential conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1]J. Rhee and J. Seo, "Low-cost curb detection and localization system using multiple ultrasonic sensors," Sensors **19**, 1389 (2019).

[2]G. Deak, K. Curran, and J. Condell, "A survey of active and passive indoor localisation systems," Comput. Commun. **35**, 1939–1954 (2012).

[3]G. Montaldo, D. Palacio, M. Tanter, and M. Fink, "Building three-dimensional images using a time-reversal chaotic cavity," IEEE Trans. Ultrason. Ferroelectrics Freq. Control **52**, 1489–1497 (2005).

[4]T.-D. Luong, T. Hies, and C.-D. Ohl, "A compact time reversal emitter-receiver based on a leaky random cavity," Sci. Rep. **6**, 36096 (2016).

[5]P. Kruizinga, P. van der Meulen, A. Fedjajevs, F. Mastik, G. Springeling, N. de Jong, J. G. Bosch, and G. Leus, "Compressive 3D ultrasound imaging using a single sensor," Sci. Adv. **3**, e1701423 (2017).

[6]M. Fink, G. Montaldo, and M. Tanter, "Time reversal acoustics," in *2004, IEEE Ultrasonics Symposium* (IEEE, 2004), Vol. 2, pp. 850–859.

[7]R. Baraniuk, "Compressive sensing [lecture notes]," IEEE Signal Process. Mag. **24**, 118–121 (2007).

[8]C. Draeger, J.-C. Aime, and M. Fink, "One-channel time-reversal in chaotic cavities: Experimental results," J. Acoust. Soc. Am. **105**, 618–625 (1999).

[9]Y. Choi, H. Lee, H. Hong, and W.-S. Ohm, "Two-dimensional virtual array for ultrasonic nondestructive evaluation using a time-reversal chaotic cavity," J. Acoust. Soc. Am. **130**, 2720–2727 (2011).

[10]N. Etaix, M. Fink, and R. K. Ing, "Acoustic imaging device with one transducer," J. Acoust. Soc. Am. **131**, EL395–EL399 (2012).

[11]B. Arnal, M. Pernot, M. Fink, and M. Tanter, "Tunable time-reversal cavity for high-pressure ultrasonic pulses generation: A tradeoff between transmission and time compression," Appl. Phys. Lett. **101**, 064104 (2012).

[12]O. Bou Matar, Y. F. Li, and K. Van Den Abeele, "On the use of a chaotic cavity transducer in nonlinear elastic imaging," Appl. Phys. Lett. **95**, 141913 (2009).

[13]T. Fromenteze, E. L. Kpré, C. Decroze, and D. Carsenat, "Passive compression technique applied to UWB beamforming and imaging architectures," Int. J. Microw. Wirel. Technol. **8**, 815–823 (2016).

[14]Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature **521**, 436–444 (2015).

[15]A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems, NIPS'12* (Curran Associates Inc., Lake Tahoe, Nevada, 2012), Vol. 1, pp. 1097–1105.

[16]O. Abdel-Hamid, A. Mohamed, H. Jiang, and G. Penn, "Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition," in *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE, 2012), pp. 4277–4280.

[17]X. Li, F. Dong, S. Zhang, and W. Guo, "A survey on deep learning techniques in wireless signal recognition," Wireless Commun. Mobile Comput. **2019**, 179.

[18]C. Rajani, A. Klami, A. Salmi, T. Rauhala, E. Haggstrom, and P. Myllymaki, "Detecting industrial fouling by monotonicity during ultrasonic cleaning," in *2018 IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP)* (IEEE, 2018), pp. 1–6.

[19]R. Stehlik and M. Carew, "OpenBuilds® actuator testing," (2016), http://openbuilds.com/uploadfiles/Part\%20Specs/OpenBuilds\%c2\%ae\%20Actuator\%20Tests_V2.pdf; accessed 29 November 2019.

[20]COMSOL Multiphysics® software version 5.4, COMSOL AB, Stockholm, Sweden, 2019, http://www.comsol.com.

[21]"Resolving time-dependent waves," (Knowledge Base, 2019), http://www.comsol.com/support/knowledgebase/1118/; accessed 22 October 2019.

[22]"Overview of materials for acrylic: Cast," (2019), http://www.matweb.com/search/datasheet.aspx?bassnum=O1303&ckck=1; accessed 13 August 2019.

[23]M. Pumperla, "Keras + hyperopt: A very simple wrapper for convenient hyperparameter optimization: Maxpumperla/hyperas," (2019), http://github.com/maxpumperla/hyperas; accessed 07 August 2019.

[24]F. Chollet *et al.*, "Home: Keras documentation," (2015), http://keras.io/; accessed 07 August 2019.

[25]D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv:1412.6980 (2014).

[26]M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. G. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: A system for large-scale machine learning," in *Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation, OSDI'16* (USENIX Association, Savannah, GA, 2016), pp. 265–283.

[27]"GPy: A Gaussian process framework in python," (2012), http://github.com/SheffieldML/GPy; accessed 07 August 2019.

[28]M. Long, H. Zhu, J. Wang, and M. I. Jordan, "Deep transfer learning with joint adaptation networks," arXiv:1605.06636 (2017).

[29]C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning, Adaptive Computation and Machine Learning* (MIT Press, 2006).