

# Impact of Biomimetic Pinna Shape Variation on Clutter Echoes: A Machine Learning Approach

Ibrahim Eshera,\* Sanmeel Lagad, and Rolf Müller

Bats species navigating dense vegetation based on biosonar must obtain sensory information about their environments from “clutter echoes”, i.e., echoes that are superpositions from many unresolved reflecting facets (e.g., leaves) with unpredictable individual waveforms. Prior results suggested that pinna deformations can aid performance in sensing tasks based on deterministic echo patterns, raising the question of whether varying pinna shapes can also have functional significance for biosonar tasks performed on clutter echoes. To test this hypothesis, this work investigates whether different pinna shapes have a consistent effect on clutter echoes despite the random nature of these signals. This is accomplished using a dedicated laboratory setup that produces large amounts of uncorrelated clutter echo data by agitating artificial foliage with fans between echo recordings. Deep learning then identifies the pinna shape that receives a given clutter echo using a data-driven classification approach that learns features directly from echoes without explicit physical modeling. A ResNet-50 achieves 97.8% overall classification accuracy for the pinna shape conformations (true-positive identifications 91.67–100%), whereas a two-dimensional convolutional neural network operating on echo spectrograms still achieves 90% accuracy. These findings demonstrate that even small pinna deformations can impart consistent effects on the clutter echoes.

## 1. Introduction

Bats are a group of animals that have achieved a remarkable evolutionary success in terms of their number of species<sup>[1]</sup> as well as the ecological niches that these species have occupied. A likely key factor behind this evolutionary success is the unique echolocation abilities seen in many bat species.<sup>[2–5]</sup> A considerable

number of bat species live in densely vegetated habitats and therefore routinely navigate in confined spaces between foliage. These environments pose serious challenges to mobility<sup>[6–8]</sup> as well as biosonar sensing.<sup>[9–11]</sup> However, these animals can navigate their challenging habitats by relying on biosonar as their main modality for sensing the environment.<sup>[12–14]</sup>

Two families of bats in particular, horse-shoe bats (Rhinolophidae<sup>[15]</sup>) and Old World round-leaf bats (Hipposideridae<sup>[1]</sup>), each with about 70 species, stand out due to unique adaptations that form a physical substrate for a perception–action loop. The animals have emitter and receiver baffle-like structures, i.e., so-called “noseleaves”<sup>[16]</sup> and the outer ears (pinnae,<sup>[17]</sup>). These structures can deform in very short time intervals through highly intricate musculatures that are unique to these bats, with some species having about twenty muscles on each pinna.<sup>[18]</sup>

Previous work has demonstrated that the noseleaf and pinna dynamics seen in bats have an acoustic impact on the emitted and

received signals, respectively. These changes occur on a similar time scale to the echoes (in one-tenth of a second<sup>[17]</sup>) and are timed so that the respective baffle shapes deform while diffracting the emitted pulse or the received echo.<sup>[17,19,20]</sup> Furthermore, shape changes in the noseleaf or pinna have been shown to encode significant additional sensory information, which could be highly beneficial for tasks such as direction finding. In particular, small deformations have been shown to result in the encoding of additional sensory information between 60% and 80%, significantly enhancing the direction-finding performance.<sup>[21]</sup>

Bats that navigate through dense vegetation face challenges to their biosonar sensing abilities that go far beyond the simple scenarios (especially direction finding for an isolated target) that have been used in the evaluation of the bat pinna rotations and deformations thus far. Echoes that originate from dense clouds of scatterers, e.g., leaves in dense foliage, are known as clutter.<sup>[22]</sup> The defining characteristics of clutter are that the echo waveforms are a superposition of components from many scatterers that have to remain unresolved due to lack of information on the positions, orientations, and shape of the contributing scatterers.<sup>[23]</sup> As a result of this lack of knowledge regarding their components, clutter echoes have to be treated as random and unpredictable by nature which can make interpretation of clutter sonar returns prohibitively difficult.

I. Eshera  
Department of Electrical & Computer Engineering  
Virginia Tech  
Blacksburg, Virginia 24061, USA  
E-mail: ieshera@vt.edu

S. Lagad, R. Müller  
Department of Mechanical Engineering  
Virginia Tech  
Blacksburg, Virginia 24061, USA

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/aisy.202500442>.

© 2025 The Author(s). Advanced Intelligent Systems published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/aisy.202500442

Despite their random nature, foliage echoes could still contain useful information for bats. Prior work has shown, for example, that clutter echoes can be used to detect passageways in foliage without the need to resolve the location of the scatterers.<sup>[24,25]</sup> Similarly, it was possible to identify locations in forest environments on a large scale (i.e., more than 10 kilometers<sup>[26]</sup>) as well as on a small scale (i.e., with a resolution well below 10 meters<sup>[27]</sup>). However, all preliminary studies on the potential use of clutter echoes by bat biosonar or its biomimetic reproductions have used a static sonar periphery. The effects of a dynamic periphery which appears to be a key component of the sonar system of rhinolophid and hipposiderid bats have not been studied in this context so far.

Given that the prominent peripheral dynamics appear to coincide with the ability to navigate and hunt in dense vegetation across different bat groups,<sup>[10,28]</sup> it could be hypothesized that these dynamics are beneficial for the encoding and/or extraction of sensory information from clutter echoes. However, it could also be argued that the random and unpredictable nature of the clutter echoes might obfuscate any informative signatures that may be created by the peripheral dynamics, and would therefore only be of potential use for simpler targets that produce deterministic echo signatures. Finally, it could also be the case that while the signatures introduced by the peripheral dynamics are detectable in the clutter echoes, they serve no functional role in encoding or extracting sensory information.

Determining which of these competing hypotheses applies to operating a biomimetic sonar with peripheral dynamics in a cluttered environment will depend on the objectives of the sensory tasks at hand and the conditions under which they have to be performed. In fact, it could very well be that the peripheral dynamics are important for certain sensory tasks that are performed under certain conditions, but not for other tasks or even the same task performed under different conditions.

To avoid making an arbitrary choice regarding the task and the pertinent conditions at this early stage of investigating the potential role of the peripheral dynamics for encoding/extracting sensory information into/from clutter echoes, the current study has been aimed at establishing whether the peripheral dynamics have a consistent effect on the incoming clutter echoes. If such a consistent effect does not exist, the dynamics cannot have any useful effect on sensory information encoding or extraction. Hence, in the absence of a consistent effect, it would probably not be worth pursuing this line of inquiry any further. If a consistent effect exists, it could still be the case that the effects cannot serve any useful function, so it would still need to be established whether the performance in any specific biosonar task can benefit from these effects. However, the insight that is the goal of the present study would make it worthwhile to pursue this research further and assess the potential impact of the peripheral dynamics on different clutter-based biosonar tasks.

In the present study, we have used a data-driven method based on deep learning classification to establish the existence of a consistent effect of the peripheral dynamics in the clutter echoes. Rather than modeling the echo transformations explicitly, deep neural networks were trained to learn distinguishing features directly from the spectrograms of the recorded echoes. A large number of clutter echoes that are suitable for this purpose were obtained from a laboratory setup that was designed to create data

that matches the structure of clutter echoes from natural environments. If correct classification of the pinna shape conformation that received a given clutter echo is possible, this is experimental evidence for an effect of pinna shape on clutter echoes that is consistent and hence potentially exploitable despite the random nature of the clutter-echo waveforms.

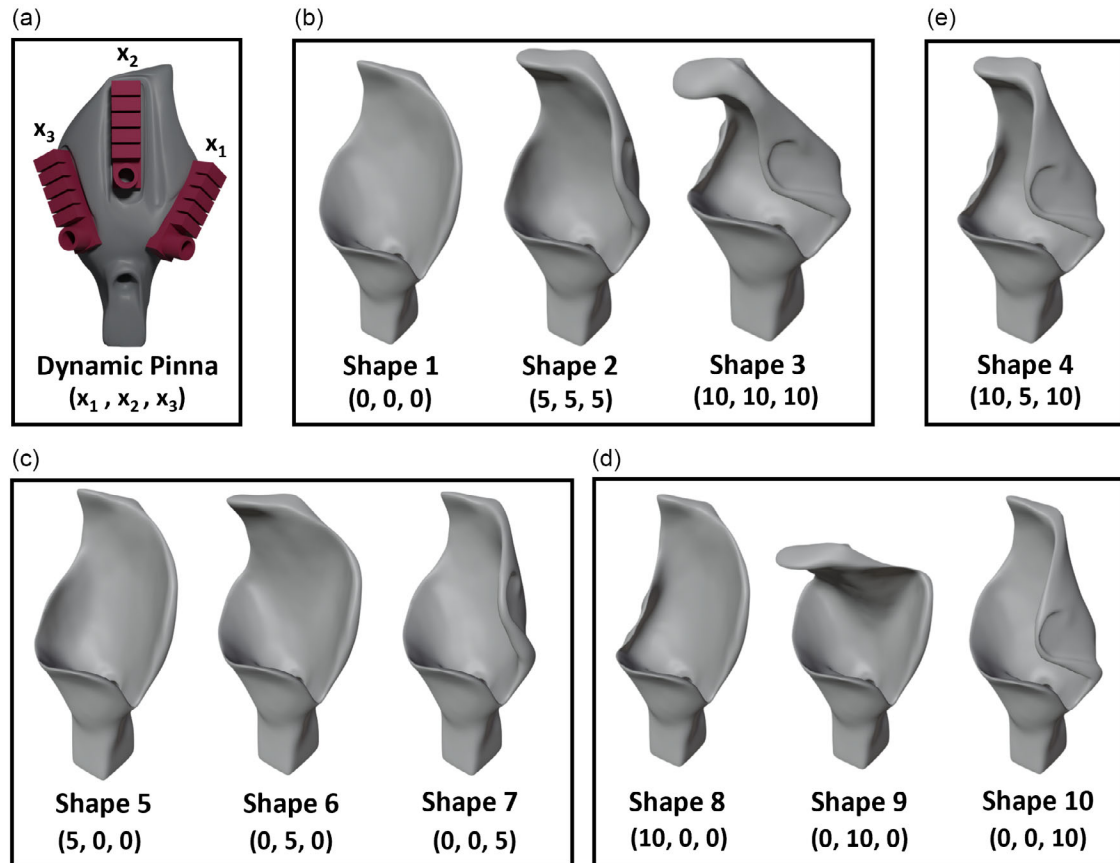
## 2. Experimental Section

### 2.1. Acoustic Setup

The acoustic elements of the biomimetic sonar system used in this research were a single emitter and a single receiver, each with a biomimetic baffle surrounding it that was designed to mimic the respective interfaces in horseshoe bats (family *Rhinolophidae*). The emitter structure was designed to mimic the noseleaf, i.e., a “megaphone-like” emission baffle, of the greater horseshoe bat (*Rhinolophus ferrumequinum*,<sup>[15]</sup>). The noseleaf was scaled up in size by a factor of about 1.7 relative to the biological model. As a result of this scaling, the noseleaf had a total height (tip of lancet to base of anterior leaf) of  $\approx 50$  mm and the “nostrils”, i.e., the outlets for the ultrasound, were  $\approx 3.5$  mm in diameter and spaced 5.5 mm apart. Two electrostatic ultrasonic transducers (Series 600 open-face ultrasonic transducer, diameter 38 mm, SensComp, Livonia, MI, USA) with a  $-6$  dB passband from  $\approx 45$  kHz to 75 kHz and a maximum response at 55 kHz were used to generate the ultrasonic emissions. One transducer each was connected to the nostrils in the biomimetic noseleaf via a conical waveguide  $\approx 10$  cm in length. The receiver structure was a reception baffle designed to mimic the pinna, i.e., the outer ear, of a horseshoe bat. The geometry of the pinna model was adapted from previous simplifications of the greater horseshoe bat (*Rhinolophus ferrumequinum*) pinna geometry.<sup>[29]</sup> The aperture of the pinna, which represents the sound-receiving region relevant for diffraction toward the ear canal, measured  $\approx 50$  mm in height. The ultrasonic echoes were recorded with MEMS capacitive microphones (Monomic, Dodotronic, Rome, Italy) placed at the end of an artificial “ear canal” (length 10 mm) that was attached to each pinna model.

To replicate the nonrigid nature of the pinna motions in horseshoe bats, deformations were generated by virtue of a biomimetic dynamic pinna model with three soft-robotic bending actuators used in prior work.<sup>[30,31]</sup> The bending action of each of the three actuators was then discretized into ten possible states, where state “zero” corresponded to a completely undeformed state of the respective actuator, and state “ten” corresponded to the maximally deformed state of the respective actuator.

Deforming the three actuators independently to any of the ten states would generate a total number of 59 049 possible pinna deformation states. Of this large number of possibilities, ten conformation states with various degrees of shape deformation were chosen (Figure 1) as qualitatively representative examples of the various pinna shape conformations that have been observed in greater horseshoe bats.<sup>[17]</sup> The set of the ten shape conformation states selected for experimentation included symmetric actuator activations (i.e., all actuators activated equally), asymmetric activations (i.e., single-actuator activations), as well as deformations due to partial or full actuator activation. Within the sample set of



**Figure 1.** Shape conformations tested: a) dynamic bat-pinna model with three actuators used to create the set of tested conformations; b) subset containing undeformed, half-deformed, and fully deformed shape conformations created by bending all three actuators on the pinna to the same extent; c) conformation subset created by bending one of the three of the actuators on the pinna at half its maximum and independently of the two others; d) conformation subset created by fully deforming one of the three of the actuators on the pinna independently of the two others; and e) conformation subset created by bending two of the three actuators to their maximum and the remaining actuator to half its maximum.

deformation, the minimum overall displacement applied to any individual actuator from the undeformed state was  $\approx 1$  cm (corresponding to a change in actuator state from 0 to 5), while the maximum overall displacement reached  $\approx 2$  cm (corresponding to a change in actuator state from 0 to 10).

The pulse waveform consisted of a carrier with a linear downward frequency modulation from 100 kHz to 20 kHz over a duration of 3 ms that was gated with a Hamming window that then served as the pulse envelope. While horseshoe bats typically employ constant frequency (CF) components followed by frequency modulated sweeps of a smaller bandwidth (up to 26 kHz,<sup>[8]</sup>), this synthetic pulse was designed to assess the information content and acoustic properties of the pinna models over the entire frequency band of the recording system, independent of species-specific time-frequency biosonar characteristics in the bats.

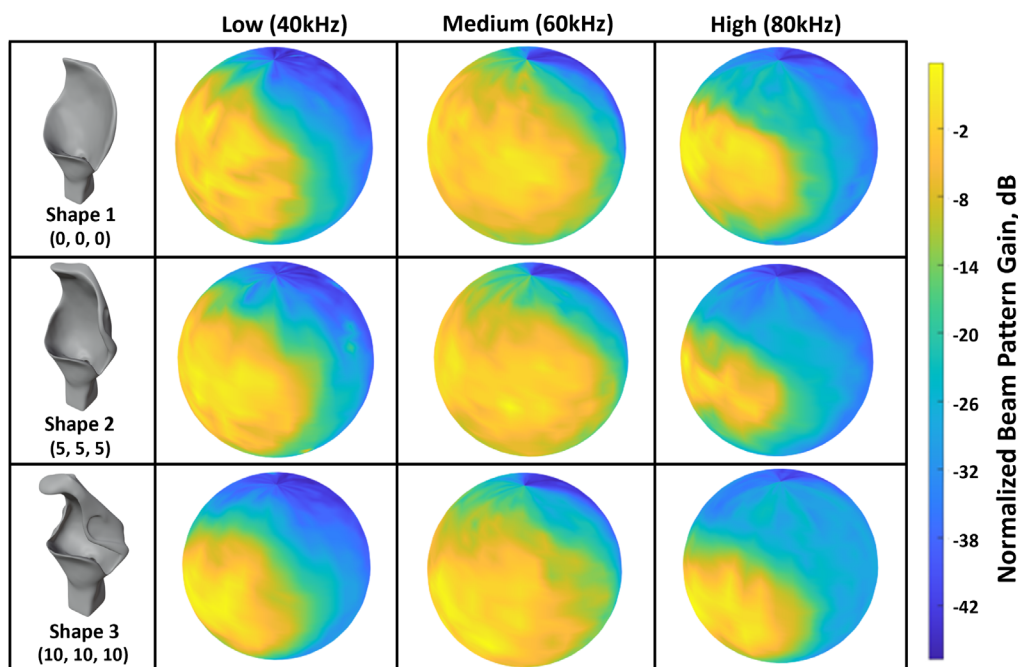
The digital pulse waveform was converted to an analog output with a conversion rate of 500 kHz and a resolution of 16 bits. The echo recordings were digitized with the same sampling rate and resolution that were used for creating the pulse waveforms (PXIe-6356 data acquisition system used for digital-to-analog and analog-to-digital conversion, National Instruments, Austin,

Texas, USA). The returning echoes were recorded over a duration of 15 ms each, synchronized to start with the beginning of the respective pulse. To serve as input for the classification experiments, a signal segment with a duration of 4 ms was selected from each echo recording. The duration of this signal segment was chosen to cover the depth of the foliage (50 cm, i.e., 3 ms time of flight) plus a safety margin to capture any echo components that could arise due to multipath reflections within the foliage.

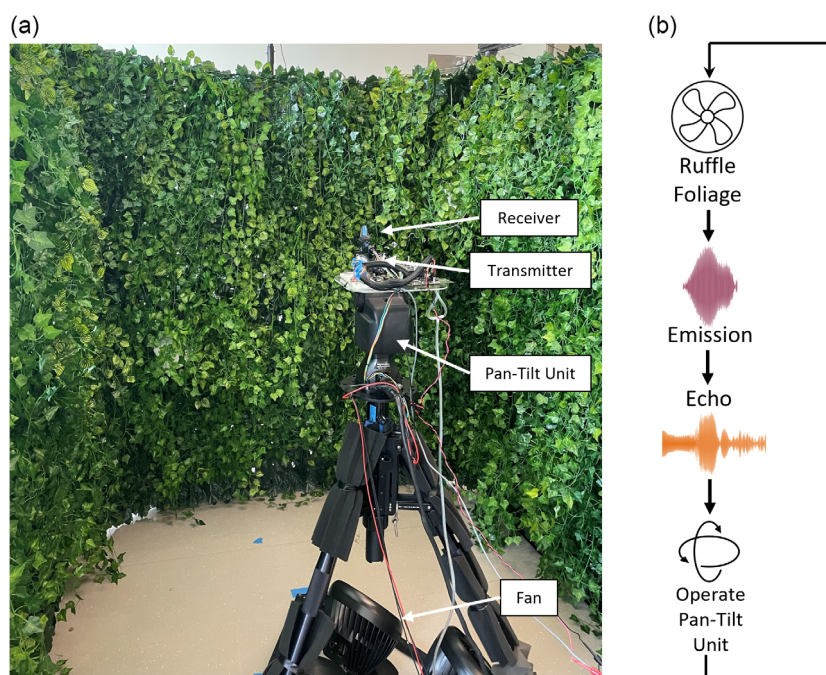
Selected pinna conformation states (numbers 1, 2, and 3, Figure 1) were characterized by virtue of experimental measurements of their beampatterns for a set of different frequencies (40 kHz, 60 kHz, 80 kHz, Figure 2). These beampatterns reflect the shape-dependent acoustic properties of the pinna, including variations in beam directionality and sidelobe structure.

## 2.2. Experimental Design

The sonar setup was placed in an enclosure with approximate dimensions of  $2 \times 2 \times 2$  m (length  $\times$  width  $\times$  height) that was surrounded by artificial foliage on three of the side walls (Figure 3).



**Figure 2.** Example beam patterns measured for three of the ten studied pinna shape conformations (1, 2, 3 – different rows) at a low (40 kHz), medium (60 kHz), and high (80 kHz) frequency (different columns), respectively.



**Figure 3.** Experimental data collection setup: a) physical setup and b) the control scheme for the experiments. Between each echo reception, fans were operated to ensure that the arrangement of the leaves in the artificial foliage differed from echo to echo.

The artificial foliage consisted of plastic leaves that were approximately elliptical in shape, with an average length of 5 cm, a width of 3 cm, and a thickness of 0.2 mm. Since the specific acoustic impedances of plastic materials are similar to those of biological

soft tissues (one to a few MRayl<sup>[32]</sup>) and hence four orders of magnitude larger than air (413 Rayl<sup>[33]</sup>) for any of these materials, the diffraction behavior of the plastic leaves can be expected to be identical to that of biological leaves to a very good approximation.

The total thickness of the foliage layer was  $\approx 0.25$  m. The density of the foliage was estimated by manually counting the number of leaves within representative  $10\text{ cm} \times 10\text{ cm}$  segments of the foliage layer, yielding an estimated density of 16,000 leaves per cubic meter. Hence, with the thickness of the foliage, about 4,000 leaves can be expected within each square-meter segment of the artificial foliage.

The floor and ceiling were not covered with artificial foliage due to the low sonar gains observed in these directions. In this setup, the noise floor in the ultrasonic recording was around  $-66\text{ dB}$  relative to the maximum amplitude of the echoes. The minimum and maximum Fraunhofer distances<sup>[34]</sup> of the biomimetic sonar emitter were estimated as 0.30 m and 1.4 m, respectively, based on a receiver aperture width of 5 cm and the operating frequency range from 20 kHz to 100 kHz. The sonarhead was placed at a minimum distance of 1.4 m away from the artificial foliage in any direction. This placement ensured that all reflections were received under far-field conditions.

The complete biomimetic sonar assembly (transmitter and receiver) was mounted on a pan-tilt unit (D-48 E, FLIR, Wilsonville, Oregon, United States) that was used to sweep the orientation of the sonarhead over a range of different orientations in the azimuth and elevation. These rotations covered a range of  $\pm 15^\circ$  in azimuth and  $\pm 7.5^\circ$  in elevation. These ranges were determined by averaging the beampatterns of all the selected pinna shape conformations and identifying the average  $-6\text{ dB}$  gain region. The limits on the range of sonarhead orientations ensured that all major reflections originated solely from the artificial foliage (Figure 4).

Data acquisition was controlled (Figure 3) so that between each pair of subsequent echo recordings, three large high-velocity drum fans, (Hyper Tough, SFDE-500B3-1, Libertyville, Illinois, United States), with an air throughput of  $3.3\text{ m}^3\text{ S}^{-1}$  each, could be activated to perturb the positions and orientations of the leaves in the foliage. To assess the effectiveness of the control scheme for ensuring that each recorded echo was unique, an experiment was conducted to estimate the correlation structure of echoes that were recorded with and without the use of the fans to alter the artificial foliage between the recordings. For these tests, a single pinna shape was used and four echo datasets were collected under the following conditions: (a) sonarhead in a fixed orientation and without operating the fans, (b) sonarhead in a fixed orientation with operating the fans, (c) varying sonarhead orientations, without operating the fans, and (d) varying orientations with operating the fans. For conditions in which the orientation of the sonarhead was varied, the pan tilt unit was rotated from  $-10^\circ$  to  $+10^\circ$  in azimuth and from  $-3^\circ$  to  $+3^\circ$  in elevation, in steps of  $1^\circ$  for both angular dimensions. For each condition, a dataset consisting of 1000 foliage echo samples was collected. The echo samples were windowed (as described above) to isolate the echo signal from the foliage thus allowing the analysis to focus solely on the effects of foliage agitation. The cross-correlation matrix for each dataset was computed,<sup>[35]</sup> where the diagonal elements represent the autocorrelation of each echo signal while the off-diagonal elements represent the cross-correlation between different echo signals. All entries in the correlation matrix were normalized to an autocorrelation value of one.

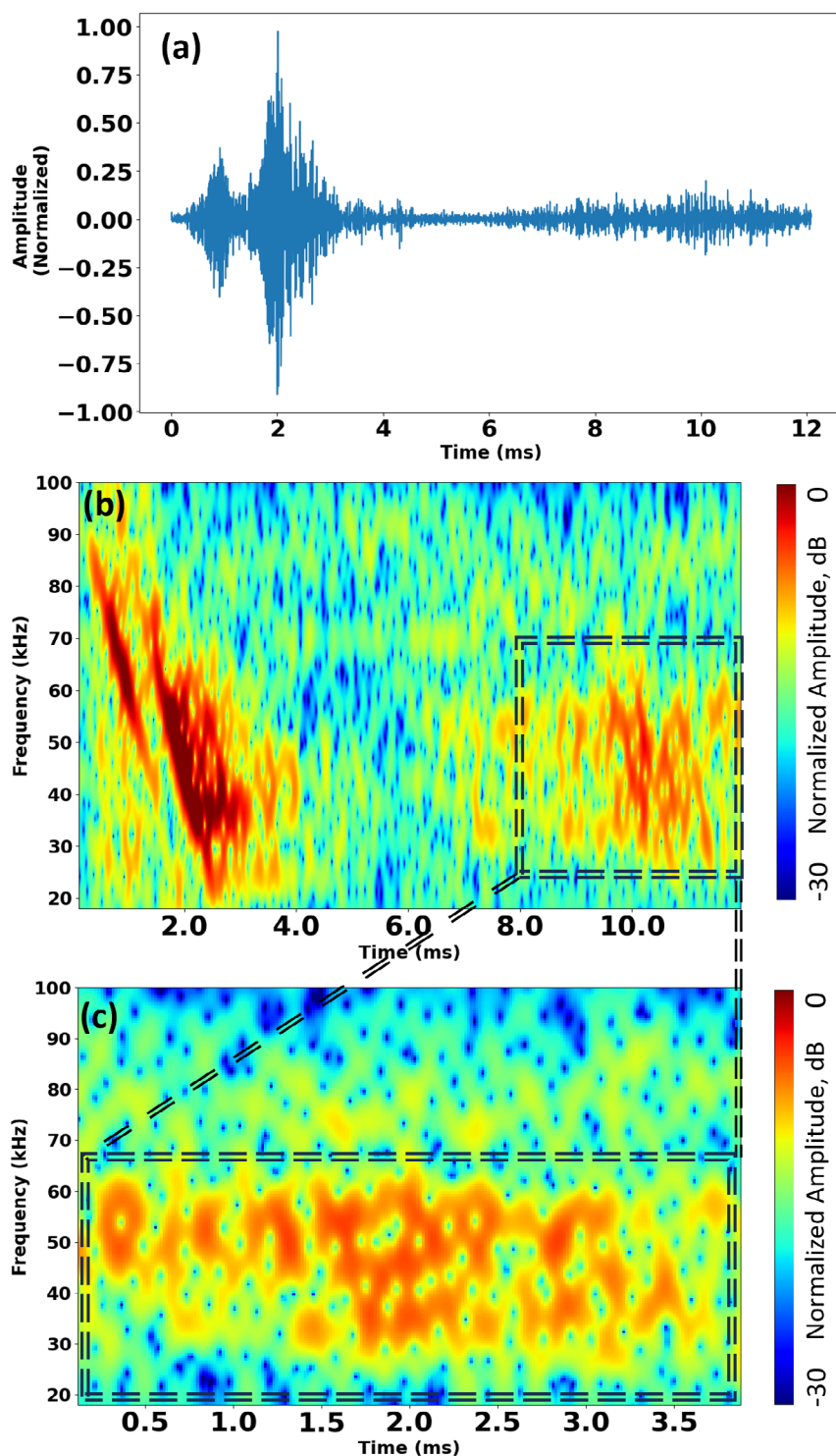
For shape conformation experiments, a total of 14,880 individual echo samples were collected equally across all ten shape

classes, i.e., 1,488 echo samples per pinna shape conformation state. The dataset was split into three subsets as follows: 60% of the dataset was set aside for training, 20% for testing, and 20% for validation. This split remained constant across all experiments conducted to ensure repeatability and consistent evaluation across all tested networks.

A bandpass filter (20 kHz to 100 kHz, 6th-order IIR Butterworth design) was applied to all echo signals to filter out frequencies not covered by the spectrum of the employed pulses. As input for shape-conformation classification experiments, the bandpass-filtered echoes were converted into spectrogram representations (Hanning window with a length of 256 samples, FFT length 256 samples, 50% window overlap). This choice of window length and overlap resulted in a spectrogram representation that contained 14 points along the time dimension. Along the frequency dimension, the spectrogram representations were cropped to the passband (20 kHz to 100 kHz).

Deep learning network architectures with varying complexity and numbers of parameters were evaluated for their ability to determine the pinna shape conformation from the spectrogram representations of the clutter echoes (Figure 5): ResNet-18, ResNet-34, ResNet-50, ResNet-152, and simple 2-D convolutional neural network.<sup>[36,37]</sup> All tested ResNet network architectures (Figure 5a) incorporated the standard ResNet architecture with groups of convolution blocks (Figure 5b) and identity blocks (Figure 5c). Each convolution block employed either a basic block (used for ResNet-18 and ResNet-34) or a bottleneck block (used for ResNet-50 and ResNet-152). The basic block contained two convolutional layers with kernel sizes of  $3 \times 3$  and  $1 \times 1$ , followed by batch normalization and ReLU activation.<sup>[38]</sup> The bottleneck block included three convolution layers: a  $1 \times 1$  convolution for channel reduction, a  $3 \times 3$  convolution for spatial feature extraction, and another  $1 \times 1$  convolution for channel expansion. The identity blocks mirrored the convolution block structure but retained the input dimensions via skip connections, ensuring that the residual path propagated through the network. The overall design (Figure 5a) was complemented by a parallel architecture (Figure 5d) representing a simpler 2-D convolutional neural network. This parallel network consisted of a  $3 \times 3$  convolution, followed by  $2 \times 2$  max pooling, and a series of dense layers leading to a softmax classifier. All networks were implemented in TensorFlow<sup>[39]</sup> (version 2.9.0) via the Keras interface library<sup>[40]</sup> (version 2.4.3) and the Python programming language (version 3.8.18). All training and inference were conducted on a graphics card (GeForce RTX 3090, NVIDIA, Santa Clara, California, USA) with the CUDA application programming interface<sup>[41]</sup> (version 12.6). In training the network, the Adam optimization algorithm<sup>[42]</sup> was used to update the network weights. Categorical cross-entropy was used as a measure for training loss.<sup>[43]</sup> Network performance was found to converge within 100 epochs.

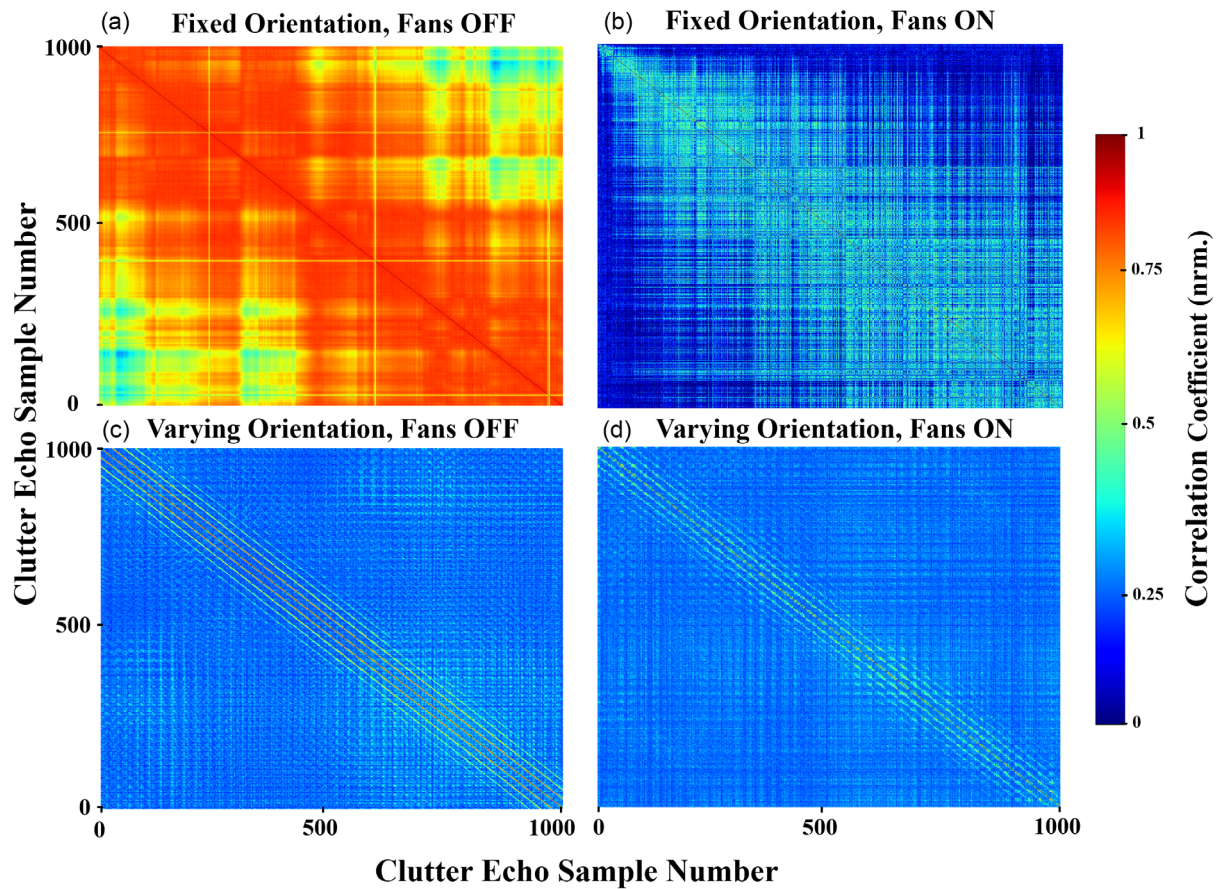
A uniform manifold approximation and projection (UMAP) analysis on the final hidden layer of the network was applied to examine the separation between various pinna shapes in two-dimensional space (Figure 6,<sup>[44]</sup>). The UMAP transformation provided a low-dimensional representation of the high-dimensional features from the network's final hidden layer. This representation preserved the local and global structure of the data, enabling clearer identification of patterns related to distinct pinna shapes.



**Figure 4.** Example of the echo recordings obtained: a) raw echo waveform; b) spectrogram of the full recording including the direct pass-through of the transmit signal 3 ms linear chirp from 100 kHz to 20 kHz, with a Hanning window as envelope) trailed by clutter echoes; and c) clutter echoes example ( $\approx 4$  ms duration) segmented from the recording. The clutter-echo segment was used as input to the deep learning classifier for the pinna shape conformation.

To assess how the information about the shape conformation of the pinna model was distributed over the duration of the echoes, a sliding window was applied to the 4 ms recorded echo

data which divided it into 1 ms segments, with a 0.5 ms overlap between consecutive windows. This created a series of 1 ms snippets, each capturing a slightly different portion of the original

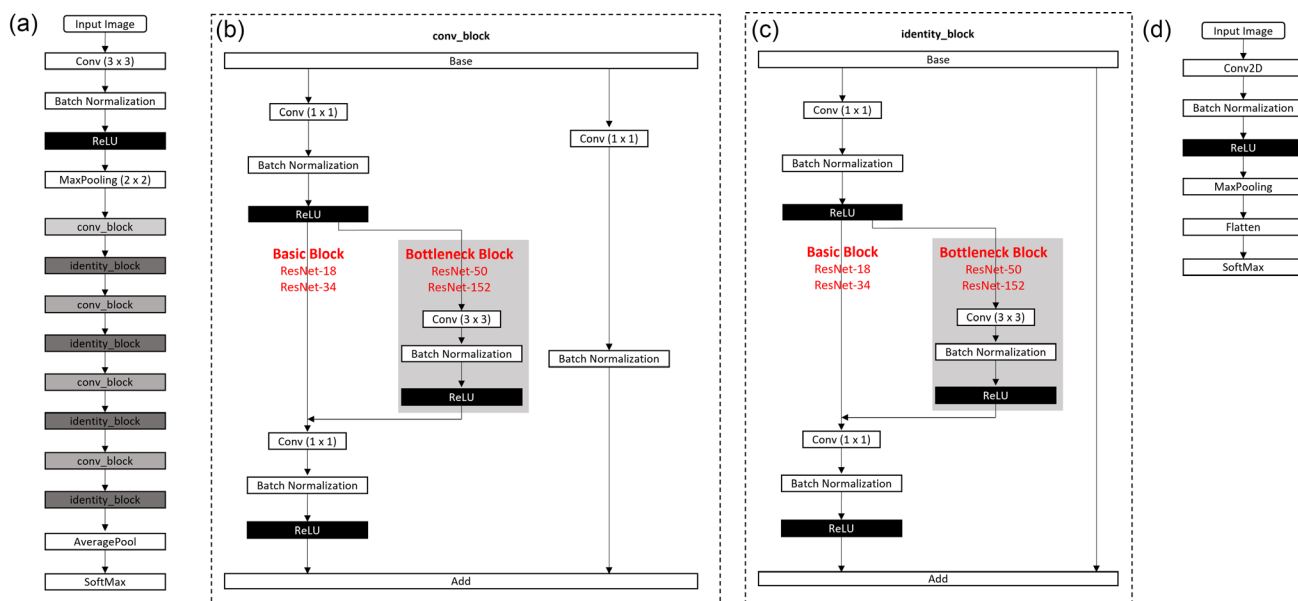


**Figure 5.** Correlation matrix of the collected echo dataset for: a) clutter echo samples collected in a fixed orientation without operation of the fans used to agitate the foliage, leading to a strong correlation in the dataset; b) clutter echo samples collected in a fixed orientation with fans activated to agitate the foliage between each pair of echo recordings, leading to a significant decrease in correlation; c) clutter echo samples collected with a varying orientation without operation of the fans used to agitate the foliage; and d) clutter echo samples collected with a varying orientation with fans activated to agitate the foliage between each pair of echo recordings.

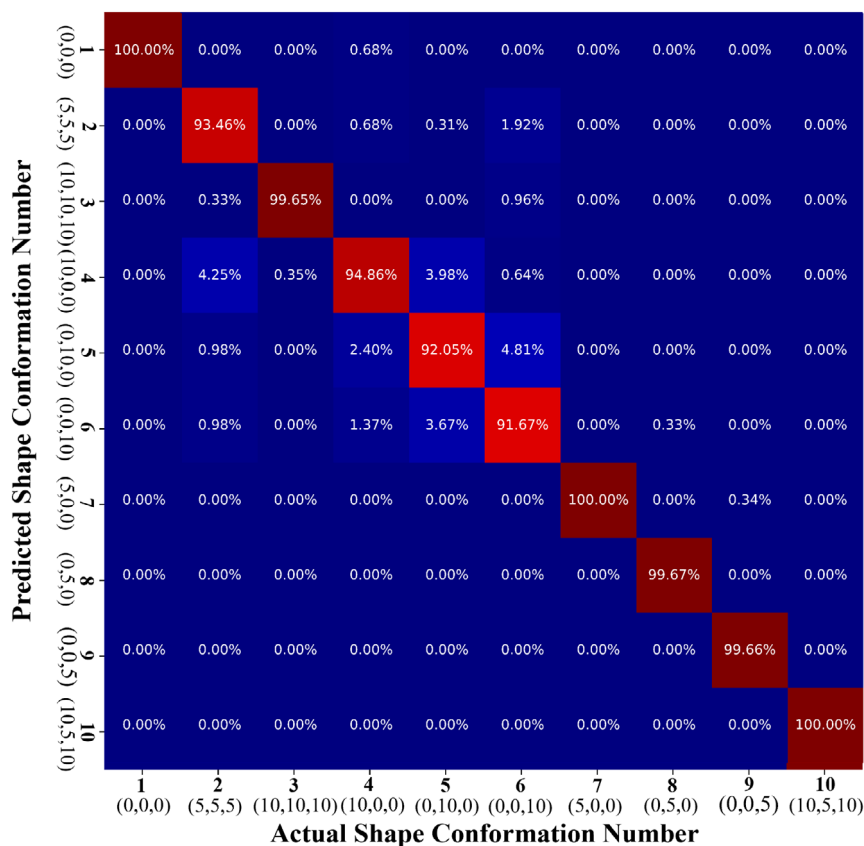
4 ms data. The classifier network was then retrained on each of these 1 ms segments (Figure 7).

The relative importance of information encoding in the time and frequency domain was assessed using the time-frequency trade off of the Fourier transform: Echo spectrograms with window lengths ranging from 0.05% of the input signal length to 100% of the input signal length were tested with ResNet-18, ResNet-34, ResNet-50, ResNet-152, and a simple 2-D convolutional neural network, with  $\approx 58\text{M}$ , 24M, 21M, 12M, and 40k parameters, respectively (Figure 7). Spectrograms were computed by applying short-time Fourier transforms (STFT) with variable window lengths ranging from 0.05% to 100% of the input signal duration. For short windows, the transform emphasized temporal precision with coarse spectral resolution, approximating a time-domain representation. For long windows approaching the full signal length, the transform captured detailed spectral content with no temporal resolution, effectively creating a pure frequency-domain representation. At the other extreme, a pure time-domain representation was used as input to the tested networks.

To investigate the impact of signal bandwidth on classification accuracy, a series of experiments were conducted using a ResNet-50 architecture with fixed STFT parameters. The STFT window length was set to  $\approx 40\%$  of the signal duration. The signal spectrograms were then truncated along the frequency dimension to bandwidths ranging from 10 kHz to 70 kHz, in 10 kHz increments. A bandwidth of 15 kHz was evaluated in addition since it corresponds to the scaled bandwidth of the dominant harmonic in greater horseshoe bats.<sup>[8]</sup> The center frequency for all tested bandwidth values was fixed at 50 kHz, which corresponds to a frequency of 85 kHz on the size scale of a greater horseshoe bat since the noseleaf and pinna models were scaled by a factor of 1.7 relative to the bat. This frequency is close to the CF components in the dominant harmonic of the greater horseshoe bat's echolocation calls (e.g.,  $84 \text{ kHz} \pm 0.5 \text{ kHz}$  in the UK<sup>[45]</sup>). All frequency-truncated spectrograms were passed through the same training pipeline to ensure consistency across conditions. The 1.7 scaling factor was selected to accommodate the physical size of the ultrasonic transducers and simplify fabrication and handling of the structures. The



**Figure 6.** Deep learning network architectures (ResNet-18, ResNet-34, ResNet-50, ResNet-152, and a simple 2-D convolutional neural network) used to identify the pinna shape conformation given the spectrogram of a single clutter echo: a) overall ResNet architecture, b) architecture of an individual convolution block showing the basic block used for ResNet-18 and ResNet-30 and the bottleneck block used for ResNet-50 and ResNet-152, c) the identity block architecture with three convolutional layers and the original input propagated in parallel, and d) the 2-D convolutional neural network architecture used.



**Figure 7.** Performance of the shape classifier: confusion matrix for the ResNet-50 classifier architecture operating on the 4 ms clutter segments (Figure 4c) from the previously unseen test dataset. The average accuracy over all shape conformations was 97.8%.

scaling preserves geometric similarity, allowing key acoustic phenomena such as diffraction and the resulting beam formation to scale proportionally under the assumption of constant wavelength-to-size ratios. As a result, key acoustic properties—in particular beampatterns—were shifted downward in frequency but remained qualitatively and quantitatively identical to the biological model.

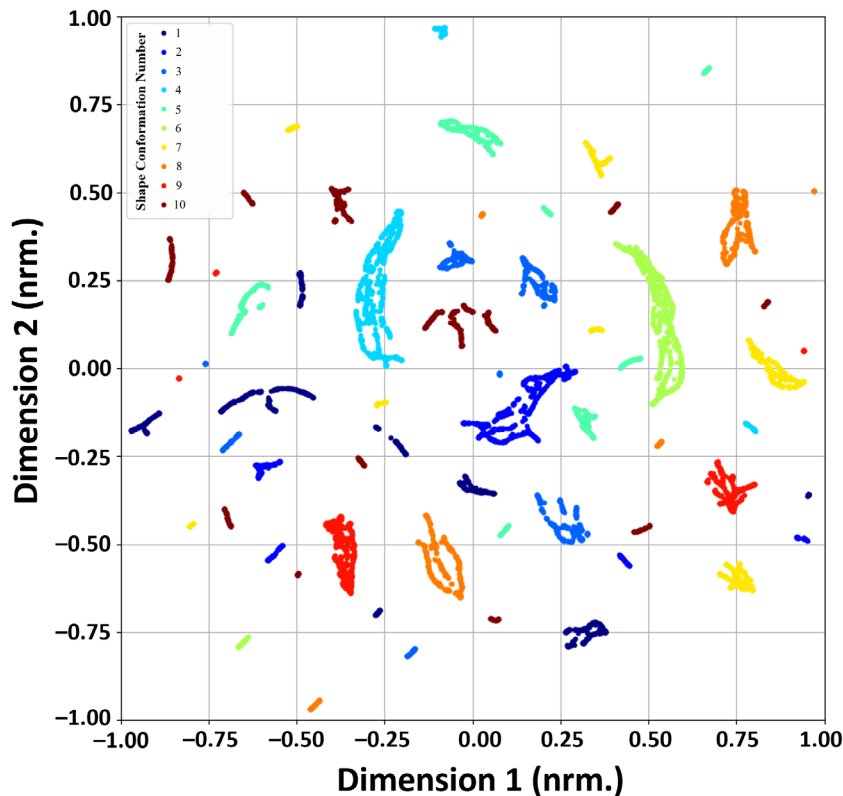
### 3. Results

The reception beamwidth at  $-6$  dB for the different pinna shapes tested was  $\approx 40^\circ (\pm 20^\circ)$  total angular range in both azimuth and elevation. On the surface of the artificial foliage, this corresponded to an illuminated circular footprint with radius  $\approx 1$  m. At a level of  $-10$  dB, the beamwidth had increased to  $\approx 60^\circ (\pm 30^\circ)$  total angular range in both azimuth and elevation, corresponding to an illuminated circle of  $\approx 1.6$  m radius. Based on the sonar footprint estimated from these beamwidths, an estimated number of 16 000 leaves could have contributed to a single echo at  $-6$  dB. For all analyzed frequencies, the different shape conformations (1, 2, 3) resulted in beampatterns that differed in their width as well as in their shape (Figure 2). The average change in beamgain due to changes in pinna shape conformation was  $\approx 13.47$  dB across all pairwise shape comparisons and the three frequencies tested. For the individual frequencies, the average differences were 13.40 dB at 40 kHz, 13.75 dB at 60 kHz, and 13.30 dB at 80 kHz.

The correlation structure of the echo datasets depended on the respective experimental conditions (Figure 8, Table 1). For a static sonarhead and foliage (e.g., no fan agitation, condition a), the correlation matrix showed a broad diagonal ridge of high correlation values (Figure 8a) with average off-diagonal values of 0.76 (Table 1a). Turning the fans on, but keeping the sonar orientation fixed resulted in a similarly broad diagonal correlation structure (Figure 8b) but with greatly reduced correlation values (average 0.22, Table 1b). Varying the orientation of the sonar resulted in a much narrower diagonal (Figure 8c) with average off-diagonal correlation values similar to those achieved by activation of the fans (0.25, Table 1c). Finally, activating the fans in

**Table 1.** Summary statistics for the off-diagonal elements in the correlation matrices for the different experimental setup conditions tested. Each column corresponds to a different experimental condition: conditions: (a) fixed sonar orientation, fans off; (b) fixed sonar orientation, fans on; (c) varying sonar orientation, fans off; and (d) varying orientation, fans on.

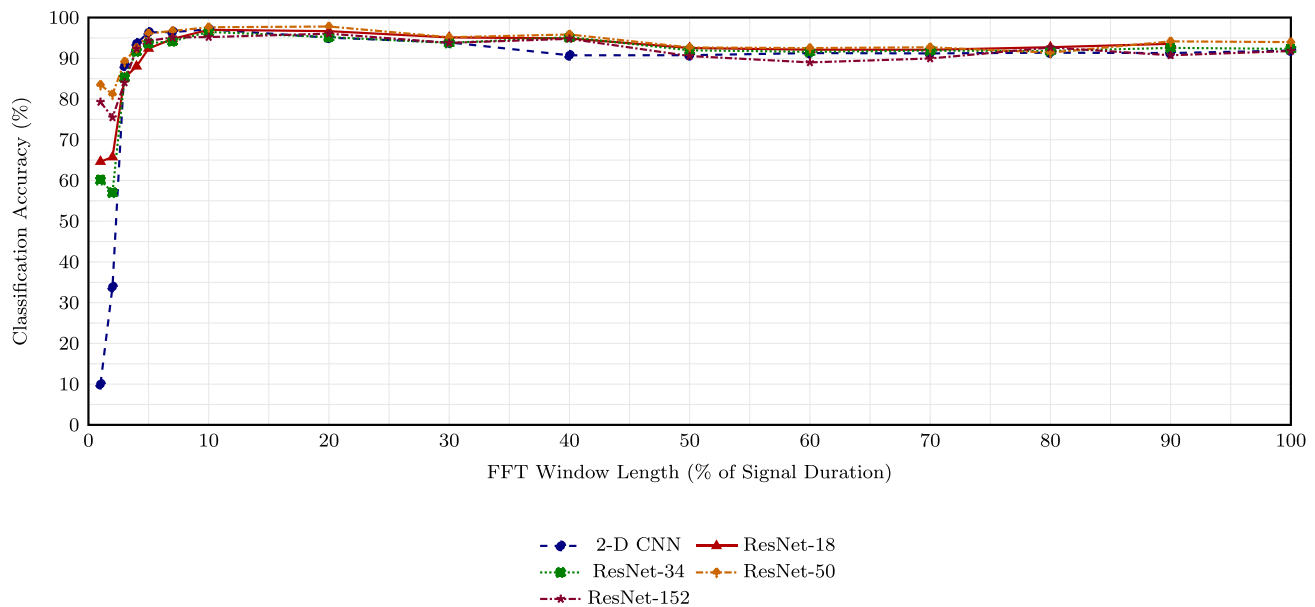
|           | (a)  | (b)  | (c)  | (d)  |
|-----------|------|------|------|------|
| Mean      | 0.76 | 0.22 | 0.25 | 0.24 |
| Median    | 0.80 | 0.22 | 0.24 | 0.23 |
| Std. Dev. | 0.10 | 0.16 | 0.07 | 0.05 |
| Minimum   | 0.24 | 0.31 | 0.14 | 0.12 |
| Maximum   | 0.90 | 0.69 | 0.93 | 0.76 |



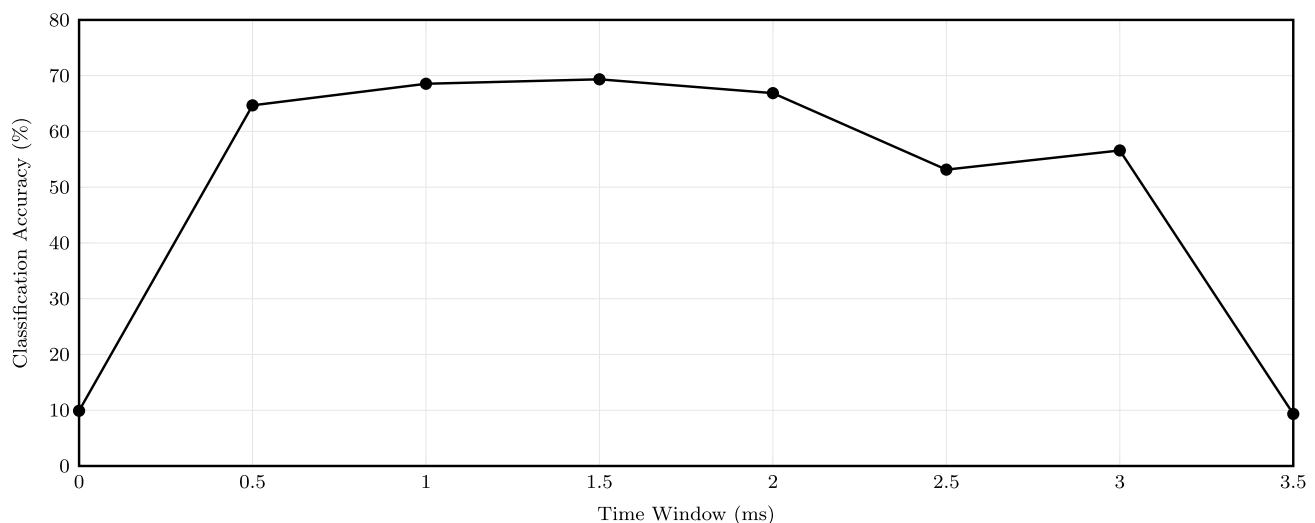
**Figure 8.** Separation of the clutter echo dataset for the different shape conformations: UMAP analysis on the final layer of the ResNet-50 network architecture showing distinct cluster separation for the ten pinna shape conformations in two dimensions.

addition to changing the orientation of the sonarhead, resulted only in a very small decrease of the average correlation values (to 0.24, Table 1d). Regardless of whether the sonarhead was static or rotated, activating the fans did result in a large decrease in the maximum off-diagonal correlation values (from 0.9 to 0.69 for the static sonarhead, and from 0.93 to 0.76 for varying sonar orientation, Table 1). Similarly, the diagonal ridge that was present in the structure of the correlation matrices whenever the sonarhead remained static was reflected in a higher standard deviation values (0.1 and 0.16 for static vs 0.07 and 0.05 for rotations, Table 1).

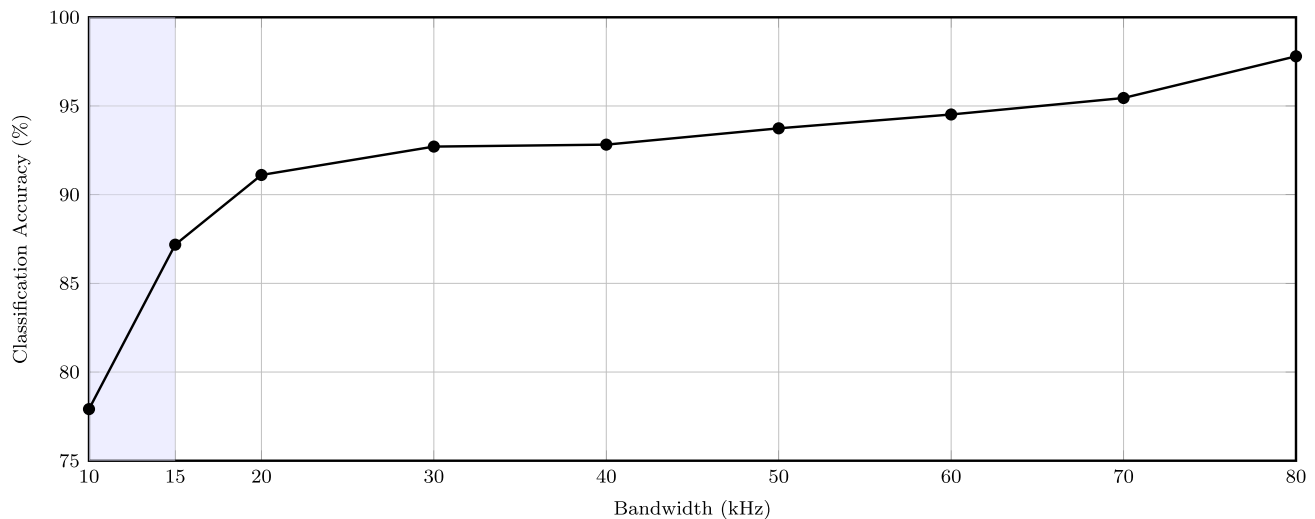
For the classification of the different pinna shape conformations, the overall highest performing network was a ResNet-50 architecture operating on spectrogram inputs computed with a window length of  $\approx 40\%$  of the signal duration and showing an overall classification accuracy of 97.8% (Figure 9). The overall lowest performing network was a 2-D convolutional neural network architecture operating on time-domain inputs (10%). For the ResNet-50 architecture with the best classifier performance, the highest confusion (4.25%) was found between shape conformations 2 and 4 (Figure 10). The true positive



**Figure 9.** Effect of the time–frequency resolution trade-off on shape classification accuracy using different network-architectures: 2-D convolutional neural network (blue), ResNet-18 (red), ResNet-34 (green), ResNet-50 (yellow), and ResNet-152 (purple). The highest classification accuracy was achieved with a ResNet-50 architecture with an overall classification accuracy of 97.8%. The lowest classification accuracy achieved was 10% with the 2-D convolutional neural network architecture.



**Figure 10.** Classification of the different pinna shape conformations based on different time segments of the clutter echoes: test accuracy based on different time segments of the echoes extracted by virtue of 1 ms-duration sliding windows that were applied across the 4 ms clutter-echo segment with 0.5 ms overlap.



**Figure 11.** Classification accuracy as a function of bandwidth centered around a frequency of 50 kHz (corresponding to 85 kHz on the size scale of the greater horseshoe bats). The shaded region indicates the bandwidths of the second harmonic in the biosonar pulses of greater horseshoe bats (26 kHz,<sup>[8]</sup> scaled according to the size of the model used here (15 kHz).

rate for the different pinna shape conformations ranged from 91.67% to 100%. The three shape conformations with a perfect recall were the upright conformation (1) and two conformations with the largest deformations (7 and 10, Figure 10).

Classification performance for all studied network architectures was found to be susceptible to the time-frequency trade-off in the input spectrograms (Figure 9): The lowest classification performances for all networks were found on pure time-domain signal representations increasing rapidly with fast Fourier transform (FFT) window length and hence frequency resolution. Classification accuracy reached a broad optimum for FFT window lengths around 20%–30% of the signal duration. For longer FFT windows, a slight decrease in classification performance occurred (down to 91.78%).

The identification of the different pinna shape conformations was found to be possible on 1 ms-long sliding windows that were positioned across the 4 ms clutter-echo recording in steps of 0.5 ms overlap (Figure 7). The best performance achieved for these windows was a classification accuracy of 69.35% with classification performance decreasing to about 10% at both ends of the echo recording.

Classification accuracy increased monotonically as the available bandwidth increased. With a narrow 10 kHz passband, the model achieved only 77.9% accuracy. Accuracy improved steadily with bandwidth, exceeding 95% for bandwidths of 60 kHz and greater. The highest performance of 97.8% accuracy was achieved when the entire passband of the transmitted pulses (20 kHz to 100 kHz) was used (Figure 11).

The results of the UMAP consisted of distinct clusters for the different shape conformations (Figure 6). However, each shape conformation was mapped onto multiple clusters in the UMAP domain, some containing very few data points.

## 4. Conclusion

The leaf density estimated for the artificial foliage falls within values that can be estimated for natural foliage: for example, a leaf area index (LAI) of  $6 \text{ m}^2 \text{ m}^{-3}$ <sup>[46]</sup> with an average leaf area of  $20 \text{ cm}^2$  and a foliage thickness of 20 cm would correspond to 16 000 leaves per cubic meter. Similarly, the correlation structure of the laboratory data obtained here (fast decorrelation away from the diagonal down to a baseline) matches what has been reported for echo data collected from natural vegetation in the field.<sup>[26,47]</sup> The average off-diagonal correlation values obtained here (0.22 for fans and motion) fell between the values that have been reported for field data (mean values of 0.13<sup>[26]</sup> and 0.53<sup>[47]</sup>). This demonstrates that while the artificial foliage used in this study may not replicate the structure of any natural foliage exactly, it was able to emulate the random nature of the natural clutter echoes.

The current results have hence established that a compact laboratory setup can be used to generate large amounts of physical clutter echoes that are not correlated beyond a baseline that is due to the common input pulse, emulating the randomness of echoes from natural cluttered environments. The results also further underscore the sensitivity of clutter echoes to even minimal perturbations, such as small rotations of the sonar or fan agitation of the foliage. The susceptibility to the outdoor equivalents of these factors, such as minimal positioning sonar positioning errors or environmental factors that change leaf orientation (e.g., wind, rain, change in light conditions) suggests that clutter echoes are not a reliable substrate for deterministic template matching approaches.

Despite the random nature of the clutter echoes, a deep learning classifier was able to successfully distinguish between the different pinna shape conformations with a high level of accuracy. Hence, the results provide a clear answer to the main research question of the present study in that clutter echoes contain

consistent effects that are due to the pinna shape they are received with. Furthermore, even the simplest architecture (2-D convolutional neural network) with a small number of parameters (40 k) was capable of achieving high classification performance (96.44%), highlighting that the ready accessibility of the pinna shape information in the clutter echoes. While not a necessary condition – as is the consistency – ready accessibility of echo features that are due to different pinna shapes would offer substantial advantages to exploiting any effects of variable pinna shapes that could enhance encoding of sensory information.

The influence that the time-frequency resolution of the echo spectrogram exerted on classification performance indicates that the classification-relevant information was most readily accessible to the classifier networks in the joint time-frequency domain. The poor performance of all classifier networks on time-domain only-representations (down to 10%) shows that a lack of frequency-domain information made it hard for the classifiers to accomplish their task, whereas the poor temporal resolution associated with the long FFT window lengths tested only produced a small dent in classifier performance (not more than 5.05%). Since the observed spread between the true positive rates associated with the different shape conformations was non-negligible (about 8%), it may indicate that some shape conformations result in clearer or hence more separable signatures in the clutter echoes. This could mean that some shape conformations may be better suited for the encoding of certain sensory information than others.

The observed improvements in classification accuracy with increasing bandwidth suggest that the strength of the signal features that are reflections of the respective pinna shape conformation increases with increasing bandwidth, especially around the low end of the bandwidth values that were studied here. For the dominant (second) harmonic of the biosonar pulses of greater horseshoe bats, bandwidth values of up to 26 kHz have been reported.<sup>[8]</sup> Scaled to the size of the model used here, this would correspond to a bandwidth of  $\approx 15$  kHz. While this bandwidth still resulted in a classification performance of 87.18% overall correct classification, well below the observed maximum, this value still suggests that greater horseshoe bats should have ready access to consistent effects of their pinna shape conformation on clutter echoes. It should also be noted that the performance in the present classification task is not necessary an upper bound on the performance in any hypothetical sensory task that would be performed based on signal features introduced into the clutter echoes by virtue of the changing ear shapes since the time-variant signatures could be used without telling different pinna shapes apart.

Due to the random nature of the clutter echoes, it is highly unlikely that the observed classification performance can be linked to simple deterministic features of the echo spectrograms. Instead, it is more likely that the classifier networks have learned statistical invariants of the echo signals, but the exact nature of these features will require additional in-depth research, e.g., using suitable transparent AI techniques. The UMAP results amplify the finding that the different shape conformations can be readily classified based on the clutter echoes recorded through them. However, while the UMAP analysis provides evidence of separability and distinct clustering, it does not reveal the exact

nature of the features being used for classification. It is also noteworthy in this context that the echoes associated with each of the different pinna shape conformations did not end up in single compact clusters suggesting that mapping the echo features into a two-dimensional representation posed a challenge.

Based on results from prior research, pinna motions observed in horseshoe bats could serve the animals' sensing in three different, but nonexclusive ways, i.e., through (i) a rigid component that reorients the beam pattern,<sup>[48,49]</sup> (ii) a nonrigid component that changes the pinna's linear characteristics (beam pattern) by virtue of the pinna geometry,<sup>[17,49]</sup> and (iii) a nonlinear, especially Doppler-based component.<sup>[19,50]</sup> The present work has extended (ii) from deterministic input signals to clutter echoes. It remains to be seen how rigid rotations and nonlinear signal transformations in the periphery of a biomimetic sonar system could affect clutter echoes.

The consistency and ready accessibility of the effects demonstrated here suggest that further exploration of potential uses of variant pinna shapes for the extraction of sensory information from clutter echoes could be worthwhile. The next challenge to continue this line of inquiry would be to identify a suitable task for demonstrating the utility of pinna shape variations. It should be possible to achieve moderate baseline performance for any candidate task without shape deformations to demonstrate principal feasibility. At the same time, the performance with a single static shape should be low enough to leave room for improvement through the incorporation of pinna deformations. Examples for such tasks could be found in all principal categories of estimation tasks, i.e., target detection, localization, and classification – as long as they are carried out in cluttered environments and require the evaluation of clutter echoes.

## Acknowledgements

This research has been supported by the Naval Engineering Education Consortium (NEEC, contract number N001741910001) and the Office of Naval Research (ONR, MURI N00014-17-1-2736). The authors would like to thank Gaurav Duggal, Joseph Sutlive, Max Engel, Colin Fox, and Henry Forsyth for contributions to the research.

## Conflict of Interest

The authors declare no conflict of interest.

## Data Availability Statement

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

## Keywords

biomimetic, biosonar, classification, clutter, machine learning, robotics

Received: April 18, 2025

Revised: August 17, 2025

Published online:

- [1] S. Solari, R. J. Baker, *J. Mammal.* **2007**, *88*, 824.
- [2] G. Neuweiler, *Trends Ecol. Evol.* **1989**, *4*, 160.
- [3] M. B. Fenton, *Can. J. Zool.* **1990**, *68*, 411.
- [4] H.-U. Schnitzler, C. F. Moss, A. Denzinger, *Trends Ecol. Evol.* **2003**, *18*, 386.
- [5] Y. Gustafson, H.-U. Schnitzler, *J. Comp. Physiol. A* **1979**, *131*, 161.
- [6] U. M. Norberg, J. M. V. Rayner, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **1987**, *316*, 335.
- [7] B. Falk, J. Kasnadi, C. F. Moss, *J. Exp. Biol.* **2015**, *218*, 3678.
- [8] B. Tian, H. U. Schnitzler, *J. Acoust. Soc. Am.* **1997**, *101*, 2347.
- [9] J. Habersetzer, G. Schuller, G. Neuweiler, *J. Comp. Physiol. A* **1984**, *155*, 559.
- [10] G. Neuweiler, W. Metzner, U. Heilmann, R. Rübsamen, M. Eckrich, H. Costa, *Behav. Ecol. Sociobiol.* **1987**, *20*, 53.
- [11] T. A. Vaughan, *Afr. J. Ecol.* **1977**, *15*, 237.
- [12] H.-U. Schnitzler, E. K. V. Kalko, *BioScience* **2001**, *51*, 557.
- [13] G. Neuweiler, E. Covey, *The Biology Of Bats*, Oxford University Press, Oxford, UK **2000**.
- [14] G. Jones, J. M. V. Rayner, *Behav. Ecol. Sociobiol.* **1989**, *25*, 183.
- [15] G. Csorba, P. Ujhelyi, N. Thomas, *Horseshoe Bats Of The World: (Chiroptera: Rhinolophidae)*, Alana Books, Bishop's Castle, Shropshire, UK **2003**.
- [16] L. Feng, L. Gao, H. Lu, R. Müller, *PLoS ONE* **2012**, *7*, e34685.
- [17] L. Gao, S. Balakrishnan, W. He, Z. Yan, R. Müller, *Phys. Rev. Lett.* **2011**, *107*, 214301.
- [18] H. Schneider, F. P. Möhres, *Z. Vergl. Physiol.* **1960**, *44*, 1.
- [19] X. Yin, R. Müller, *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 12270.
- [20] L. Zhang, L. Yang, R. Zhang, R. Müller, *J. Acoust. Soc. Am.* **2020**, *148*, 954.
- [21] R. Müller, A. K. Gupta, H. Zhu, M. Pannala, U. S. Gillani, Y. Fu, P. Caspers, J. R. Buck, *Phys. Rev. Lett.* **2017**, *118*, 158102.
- [22] D. S. Jacobs, A. Bastian, *Bat Echolocation: Adaptations For Prey Detection And Capture*, 13–30, Springer International Publishing, Cham, ISBN 978-3-319-32492-0 **2016**.
- [23] R. Müller, R. Kuc, *J. Acoust. Soc. Am.* **2000**, *108*, 836.
- [24] R. Wang, R. Müller, *Bioinspir. Biomim.* **2021**, *16*, 066022.
- [25] R. Wang, Y. Liu, R. Müller, *Bioinspir. Biomim.* **2022**, *17*, 056009.
- [26] L. Zhang, R. Müller, *Bioinspir. Biomim.* **2022**, *17*, 026011.
- [27] L. Zhang, A. Farabow, P. Singhal, R. Müller, *Bioinspir. Biomim.* **2023**, *18*, 026009.
- [28] A. R. C. Britton, G. Jones, *J. Exp. Biol.* **1999**, *202*, 1793.
- [29] P. B. Caspers, Ph.D. Dissertation, Virginia Polytechnic Institute and State University (Virginia Tech), Blacksburg, VA, USA 2017, <https://vtechworks.lib.vt.edu>.
- [30] J. Sutlive, R. Müller, *Bioinspir. Biomim.* **2019**, *14*, 066014.
- [31] J. Sutlive, A. Singh, S. Zhang, R. Müller, *Bioinspir. Biomim.* **2020**, *16*, 016016.
- [32] F. Duck, *Physical Properties Of Tissues: A Comprehensive Reference Book*, Academic Press, London **1990**.
- [33] L. E. Kinsler, A. R. Frey, A. B. Coppens, J. V. Sanders, *Fundamentals Of Acoustics*, Wiley, New York, 4th ed., **2000**.
- [34] E. Hecht, *Optics*, Addison Wesley, 4th intern ed., San Francisco, CA **2002**.
- [35] A. Oppenheim, *Discrete-Time Signal Processing*, Pearson Education, Upper Saddle River, NJ **1999**.
- [36] K. He, X. Zhang, S. Ren, J. Sun, *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.* **2016**, 770.
- [37] A. Krizhevsky, I. Sutskever, G. E. Hinton, *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 1097.
- [38] X. Glorot, A. Bordes, Y. Bengio, in (Eds: G. Gordon, D. Dunson, M. Dudík), *Proceedings of the Fourteenth Inter. Conf. on Artificial Intelligence and Statistics, Proceedings of Machine Learning Research*. PMLR, Fort Lauderdale, FL, USA, **2011**, vol. 15, pp. 315–323, <https://proceedings.mlr.press/v15/glorot11a.html>.
- [39] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. J. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Józefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. G. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, et al., *CoRR* **2016**, abs/1603.04467.
- [40] F. Chollet, *Deep Learning With Python*, Manning Publications, Shelter Island, NY, 2nd ed. **2021**.
- [41] NVIDIA, P. Vingelmann, F. H. Fitzek, Cuda, release: 10.2.89 **2020**, <https://developer.nvidia.com/cuda-toolkit>.
- [42] D. P. Kingma, J. Ba, *Adam: A method for stochastic optimization* **2017**, <https://arxiv.org/abs/1412.6980>.
- [43] K. P. Murphy, *Machine learning: A probabilistic perspective*, MIT Press, Cambridge, Mass. [u.a.] **2013**.
- [44] L. McInnes, J. Healy, N. Saul, L. Großberger, *J. Open Source Softw.* **2018**, *3*, 861.
- [45] M. M. Andrews, P. T. Andrews, D. F. Wills, S. M. Bevis, *Acta Chiropt.* **2006**, *8*, 197.
- [46] D. Clark, P. Olivas, S. Oberbauer, D. Clark, M. Ryan, *Ecol. Lett.* **2008**, *11*, 163.
- [47] A. Bhardwaj, M. O. Khyam, R. Müller, *Bioinspir. Biomim.* **2021**, *16*, 046026.
- [48] V. A. Walker, H. Peremans, J. C. T. Hallam, *J. Acoust. Soc. Am.* **1998**, *104*, 569.
- [49] X. Yin, P. Qiu, L. Yang, R. Müller, *J. Acoust. Soc. Am.* **2017**, *141*, 3011.
- [50] X. Yin, R. Müller, *Nat. Mach. Intell.* **2021**, *3*, 507.