

Article

# A Feed-Forward Neural Network Approach for Energy-Based Acoustic Source Localization

Sérgio D. Correia <sup>1,2,\*</sup> , Slavisa Tomic <sup>1</sup>  and Marko Beko <sup>3</sup> 

<sup>1</sup> COPELABS, Universidade Lusófona de Humanidades e Tecnologias, Campo Grande 376, 1749-024 Lisboa, Portugal; slavisa.tomic@ulusofona.pt

<sup>2</sup> VALORIZA—Research Centre for Endogenous Resource Valorization, Instituto Politécnico de Portalegre, Campus Politécnico n. 10, 7300-555 Portalegre, Portugal

<sup>3</sup> Instituto de Telecomunicações, Instituto Superior Técnico, Universidade de Lisboa, 1049-001 Lisbon, Portugal; marko.beko@tecnico.ulisboa.pt

\* Correspondence: scorreia@ipportalegre.pt

**Abstract:** The localization of an acoustic source has attracted much attention in the scientific community, having been applied in several different real-life applications. At the same time, the use of neural networks in the acoustic source localization problem is not common; hence, this work aims to show their potential use for this field of application. As such, the present work proposes a deep feed-forward neural network for solving the acoustic source localization problem based on energy measurements. Several network typologies are trained with ideal noise-free conditions, which simplifies the usual heavy training process where a low mean squared error is obtained. The networks are implemented, simulated, and compared with conventional algorithms, namely, deterministic and metaheuristic methods, and our results indicate improved performance when noise is added to the measurements. Therefore, the current developed scheme opens up a new horizon for energy-based acoustic localization, a field where machine learning algorithms have not been applied in the past.

**Keywords:** acoustic localization; artificial intelligence; artificial neural networks; deep feed-forward networks; deep learning; embedded computing; energy-based localization; wireless sensor networks



**Citation:** Correia, S.D.; Tomic, S.; Beko, M. A Feed-Forward Neural Network Approach for Energy-Based Acoustic Source Localization. *J. Sens. Actuator Netw.* **2021**, *10*, 29. <https://doi.org/10.3390/jsan10020029>

Academic Editors: Diogo Gomes, Mario Antunes and Luis Miguel Contreras-Murillo

Received: 19 February 2021

Accepted: 19 April 2021

Published: 22 April 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The localization of an acoustic source in Wireless Sensors Networks has been commonly employed in several real-life problems. Examples of its application can be found for energy control of buildings [1], ambient assisted living [2], underwater acoustic networks [3], wildlife monitoring [4], smart surveillance [5], shooter detection [6], or as a complementary source of information to other locating platforms [7,8]. The solution to the problem consists of obtaining measurements that represent the distances from an acoustic source to sensors that acquire the measurement. Contrary to range-free methods, physical measurements such as time-of-arrival [9], time-difference-of-arrival [10], or direction-of-arrival [11] have shown promising results for acquiring distance measures; however, they rely either on high-precision hardware for timing purposes or on microphone sensor arrays for angle perception. Contrarily, range-free methods rely on information about connectivity and propagation patterns and are, thus, highly dependent on environmental conditions [12,13]. The acoustic energy decay model as an indirect measure of distance was initially proposed by empirically analyzing the sound emitted from an engine [14]. The localization approach considers averaging the energy of the received acoustic signal data samples, standing out for lower bandwidth since it is sampled at a much lower rate [15]. Additionally, the required hardware becomes very simple, having as the main part a single microphone that converts acoustic pressure into an electrical signal. This model is considered here, that is, the energy measured at each sensor related to the transmitted power and an inverse proportionality to the squared distance between the sensor and the acoustic source.

With regard to artificial intelligence, Artificial Neural Networks (ANNs) have reached a point of maturity that allows for wider use. Regarding its employment in location and positioning, the use of a Multilayer Perceptron (MLP) was firstly evaluated in terms of accuracy, memory, and computational requirements, showing promising results [16]. It was shown that the MLP could potentially achieve higher localization accuracy, requiring less computational effort and lower memory resources when compared with an Extended Kalman filter, for example. More recently, ANNs have reported promising results in applications such as simultaneous localization and mapping technique [17] or Wi-Fi Fingerprint [18]. However, its major drawback lies in their training stage, where the physical presence of training devices is necessary and tedious. Further, Convolutional Neural Network designs were studied with application to localization, by simulating hydrodynamic flow caused by target objects of location, with meritorious accuracy results [19]. Hence, although the application of neural networks can be seen as an option for localization problems, challenges remain to be overcome with regard to their training and structure.

It is worth mentioning that the learning procedure of a neural network can be either supervised or unsupervised. The related works presented above consider that the desired output is already known. As such, during the learning process, the units (or weight values) of such a network are determined given pairs of input/output values. Depending on the difference between the current iterative output and its target (known desired output), an error value is computed with the goal of being minimized. This procedure is called supervised learning, as the current output is being supervised, or monitored, to match the desired one [20]. On the contrary, unsupervised learning does not consider the desired output to be known and has the goal of arranging groups of similar inputs close together. This effect can be used efficiently for pattern classification purposes. Examples of sound classification events can be found in [21], where different sources are separated and classified. Considering the present problem of acoustic localization, a supervised approach is considered, where the network will be trained taking into account observations of an acoustic model, and the corresponding coordinates in a predefined search-space.

Based on the previous discussion, the present work proposes a Deep Feed-Forward Network (DFNN) for solving the energy-based acoustic source localization problem. Several hyperparameters, such as the number of hidden layers, perceptrons, and epochs, are surveyed to build an effective training model. Moreover, the proposed DFNN is trained with noise-free input data and validated against different measurement noise levels. This corresponds to a training stage independent of physical devices, requiring only partial knowledge of the localization layout (dimension of the search space, number of sensors, and an environmental decay factor). The dataset is applied to the new DFNN, where the root-mean-squared error (RMSE) is calculated and compared with several state-of-the-art algorithms. The obtained results indicate that the proposed DFNN is a promising method for determining the location of an acoustic source through energy-based measurements, attaining lower RMSE errors for a wide range of the measurement noise, associated with its low implementation complexity. Besides the simplified methodology for network training, and as far as the authors' knowledge, artificial neural networks have not been applied to the energy-based acoustic source localization problem in the scientific literature. While the existing works on different localization schemes consider hardware-based methodologies for network training, the current work represents the first complete study of the energy-based localization problem via neural networks.

The main insights and contributions of the present work are summarized as follows: (1) a DFNN architecture for solving the energy-based acoustic source localization problem is proposed, where several hyperparameters are tuned; (2) the proposed DFNN is trained with noise-free input data and, later on, validated against different measurement noise levels; (3) the proposed training stage allows for offline network training, without any complementary hardware or previous data acquisition; (4) the proposed architecture supplants traditional methods for a wide range of noise levels; (5) the present work

stamps the appliance of artificial neural networks to the energy-based acoustic source localization problem.

The remaining paper is organized as follows. Section 2 summarizes previous related work and Section 3 provides the theoretical background on energy-based acoustic localization and DFNNs. Section 4 describes the proposed methodology, while Section 5 assesses its performance. Finally, Section 6 concludes the work.

## 2. Related Work

Traditionally, the energy-based acoustic source localization problem has been addressed by the use of deterministic methods [22]. With this proposal, a weighted least-squares method was applied in [23], which was enhanced with a correction technique and presented good results for low values of noise, but its performance is degraded in noisy environments. A closed-form solution was proposed in [24], which exhibits good performance for low noise power, but also suffers considerable degradation for higher levels of noise power. The method proposed in [25] stands for its simplicity, consisting of a bisection approach, where good performance is obtained for low values of the measurement noise. Considering the fact that the problem is highly nonconvex, the use of convex optimization methods was proposed in [26,27] through semidefinite programming relaxations. Considering that the problem is not approached directly, but rather through approximations, its enhancement was proposed in [28–30] by applying second-order cone programming. The methods based on convex optimization performed well, even in noisy environments, with their only drawback being their computational complexity, which increases geometrically with the size of the network. Besides deterministic methods presented, the use of swarm-based optimization—namely, Elephant Herding Optimization (EHO) [31,32]—was initially proposed in [33,34]. The methodology was enhanced in [35,36] by new population initialization strategies, combining computational simplicity with low positioning error. The improved EHO [35,36] demonstrated a high suitability when considering embedded implementations for real-time applications, mainly due to its low latency, although knowledge of the noise statistics is assumed by the estimator.

The current state-of-the-art about the acoustic source localization problem is mostly based on triangulation methods that rely on Euclidean geometry. To this end, some measure of distance is acquired by a sensors network with the purpose of inferring geometric properties on the source location, namely, energy, angle, or time [37]. The correlation between the measured quantity and the coordinates of the physical location in space is generally non-linear, non-convex, and subject to different sources of noise. Under near-ideal conditions, these models can very accurately estimate the acoustic physical location without error bounds and among different environments. Nevertheless, sources of measurement noise, simplifications in the ranging models, or complex environment conditions, have great impact on the accuracy and reliability on the system performance. Basically, model-based methods become less reliable when more effects are present that were not considered in its physical mode foundations, i.e., when one has less confidence in the model itself. Theoretically, the Cramer–Rao lower bound is usually applied to demonstrate and state the estimation bounds; its dependencies with the physical model; and, of major importance, the geometry of the problem [15,24,38–41]. Unlike model-based methodologies, data-driven (or learning-based) methods are an alternative to solve constrains that construct a mapping function by acquired knowledge. More specifically, ANNs can behave as universal approximators, and almost automatically discover features relevant to the localization problem by exploiting the increased amount of sensor data and computational power on an offline stage. The localization problem is treated as a regression one, where distances are learned, mapping coordinates in a search-space, having modest or nonexistent physical knowledge intrinsic to the problem under scrutiny [42–45].

### 3. Theoretical Background

The current section aims to provide the theoretical foundations of both the energy-based acoustic localization problem and the deployment of DFF Networks.

#### 3.1. Energy-Based Acoustic Localization

Location of an acoustic source, by exploiting energy measurements acquired by sensors, was firstly addressed in [14]. The proposed model considers  $M$  noisy measurements within a time window  $T = M/f_s$ , where  $f_s$  is the acoustic sampling frequency and averages energy signatures over the time window  $[t - T/2, t + T/2]$ . The obtained measure at the  $i$ th sensor is then modeled as follows [14,15]:

$$y_i = \frac{g_i P}{\|\mathbf{x} - \mathbf{s}_i\|^\beta} + v_i, \quad i = 1, \dots, N, \tag{1}$$

where  $g_i$  is the gain of sensor  $i$ ,  $P$  is the transmitted power,  $\mathbf{x} = (x_x, x_y)^T$  is the unknown location,  $\mathbf{s}_i = (s_{ix}, s_{iy})$  denotes the known location of sensor  $i$ ,  $v_i$  represents the measurement noise modeled as Gaussian,  $N$  is the total number of sensors, and  $\beta$  is a decay factor dependent on environmental conditions. For the sake of simplicity, the decay factor is considered  $\beta = 2$  [14,15], which corresponds to an outdoor setting. In Expression (1), the observed measurements represent the power received per unit surface area where that emitted energy falls, and thus, its unit is  $W/m^2$ . With the purpose of generalizing the physical conditions of the problem, the observations will be numerically treated as dimensionless in the present work. When performing several measurements (minimum of 3 for a bidimensional space), the unknown position would be determined as the intersection point of circumferences, centered on the sensor coordinates, with radii obtained from Expression (1). Due to noise corruption of the measurements, the intersection of the circumferences will likely form an area, rather than a single point (Figure 1). Hence, this work will train the proposed DFNN with noise-free data only, avoiding in this way the need for characterizing the statistical behavior of  $v_i$ . This will imply a simple procedure, independent on real terrain acquisition.

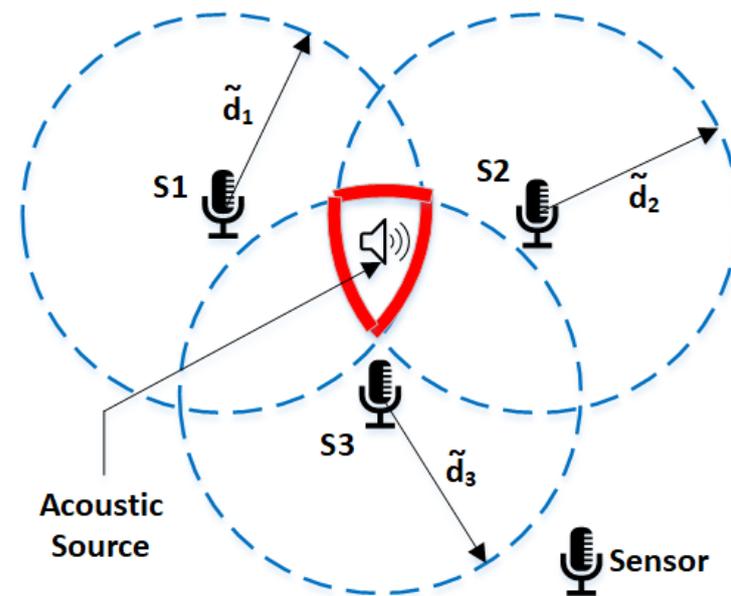


Figure 1. Sensors' distance estimate representation when measurement noise is considered.

Regarding both deterministic and metaheuristic algorithms, all observations from the multiple sensors are aggregated as an estimator of  $\mathbf{x}$ , where the solution of the localization problem is the argument (pair of coordinates) that minimizes the expression [14,15]:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \sum_{i=1}^N \frac{1}{\sigma_{v_i}^2} \left( y_i - \frac{g_i P}{\|\mathbf{x} - \mathbf{s}_i\|^2} \right)^2. \quad (2)$$

The estimator in Expression (2) is highly nonconvex, with singularities in each sensor's coordinates, several suboptimal solutions, and saddle regions. All the enumerated features makes the problem very challenging in the field of numerical optimization, making it a good candidate in the context of regression and ANNs.

The considered energy model (Expression (1)) relies on the fact that the acoustic source is stationary. Targeting moving sound sources is mostly considered with direction or angle measurements [46,47], combined with microphone arrays [48], or mostly relying on the measurement of propagation time [49], given the achievement of very satisfactory accuracy, despite the complexity of the physical devices employed. While in the space-state domain Kalman filters and maximum a posteriori estimation are commonly used [50], ANNs have also been considered [51,52]. In this case, time series prediction is performed with the resource of recurrent neural networks composed of long short-term memory cells [53]. Although network structures are currently well defined in the literature, problems with the need to find the appropriate sample rate, the need to identify an appropriately sized input window, or to archive stability are still challenging.

### 3.2. Deep Feed-Forward Neural Network

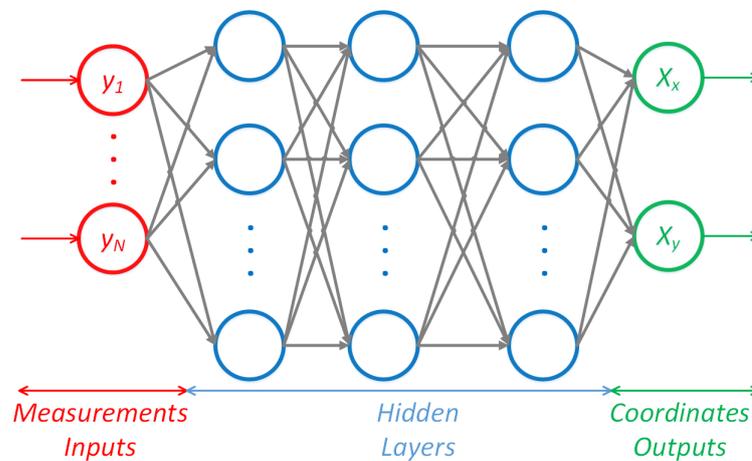
Theoretical results on ANNs, known as the *Universal Approximation Theorem*, assert that a single hidden layer on a sigmoidal feed-forward ANN with a sufficient numbers of nodes is capable of approximating any continuous function with admissible accuracy [54]. The theorem was generalized to feed-forward multilayer architectures in 1991 by Kurt Hornik [55] and, more recently, it was shown that universal approximation also holds for unbounded activation functions such as the rectified linear unit (ReLU) [56]. The mentioned theorem allows the following hypothesis: *if energy values are measured at at least three sensor positions, at a specified point in 2-dimensional space, then a sigmoidal feed-forward artificial neural network may be established, which takes as input the energy values measured at the sensors and predicts the coordinates of the source propagating the acoustic signal*. According to this hypothesis, the key task of this work is to correctly set the network topology and its training. Fundamentals of ANNs rely on modeling the biological neuron and its intercommunication cells called synapses, where an artificial neuron (or perceptron) is obtained. A deep neural network will have an input layer, two or more hidden layers, and one output layer. Each layer is connected to the next one through some synaptic weights, forming a flow of information in one direction, from the inputs to the outputs.

The Sigmoid (or Fermi) activation function was one of the first to be applied in ANNs. The function maps the input to a value between 0 and 1, having a simple derivative that efficiently performs the network training through gradient-based algorithms [54]. Since a continuous signal is faked as output (coordinate values with regard to the search space), the hyperbolic tangent function is considered. The function is similar to the previous sigmoid and shares much of its properties. However, this function allows us to map the input to any value between  $-1$  and  $1$  [57]. Putting a network to work consists of determining the weight that best fits the training data, that is, the set of pairs of inputs and outputs. Essentially, it consists of a regression problem, where the extent of the network can reach a large number of dimensions and, therefore, ANNs will imply good performance [58]. The learning problem is formulated in terms of minimization of an objective function, which measures the performance of a neural network on a predefined data set. Concerning the network training, the backpropagation algorithm is one of the most popular methods to train neural networks [59]. Nonetheless, original backpropagation suffers from slow convergence,

and thus, several variants have been proposed in the literature. The Levenberg–Marquardt (LM) algorithm is an alternative for training ANNs based on a nonlinear optimization [60]. The method employs an approximation of second-order derivatives of the objective function so that better convergence behavior can be obtained [61]. Based on the above discussion, our artificial neuron model will rely on the hyperbolic tangent activation function and the network will be trained by the LM algorithm.

#### 4. Proposed Method

The proposed approach for solving the energy-based acoustic source localization problem relies on a DFNN (Figure 2), where the inputs consist of the measures taken from microphones in the sensors' network. The structure of the network is therefore dependent on the problem layout. For a layout composed of  $N$  acoustic sensors,  $N$  inputs will be present on the DFNN. Regarding the outputs, independent of the number of network entries, these will always be two, corresponding to the estimates of the source coordinates  $(x_x, x_y)$ .



**Figure 2.** Proposed Deep Feed-Forward Neural Network.

Identifying network topology—namely, the number of hidden layers and the number of perceptrons in each layer—has been traditionally based on trial and error [62]. Although it was proven that one single hidden layer can approximate any continuous function [54], the number of required perceptrons in each layer can be as high as the number of training samples [63]. In fact, when considering  $m$  output neurons, the number of perceptrons to train  $N$  samples with some reduced error is given by  $2\sqrt{(m+2)N}$  [63]. In addition, it can be seen that a three-layer feed-forward neural network with  $k$  hidden perceptrons can assign arbitrary analogue values to  $j$  arbitrary inputs [64]. To overcome the physical dimension of the search space and the properties of the model (Expression (1)), a high number of samples must be used to train the proposed network. This situation would lead to a high complexity of the network structure, according to [63]. As such, a two-phase method is considered to determine the optimal number of perceptrons in the hidden layers of a 5-layer network [65].

The overall proposed network consists of one input layer, one output layer, and three hidden layers. The number of layers and the number of perceptrons in each layer are hyperparameters that must be initially specified [66]. Three hidden layers are empirically stated, considering one first layer, a nonlinear separation, one processing layer, and one aggregation layer to provide the outputs. The number of perceptrons in each layer will be under scrutiny and further analyzed and discussed. As mentioned earlier, one of the outcomes of the methodology is its training stage, where random samples are generated over the search space and the observations (or inputs) are calculated under ideal conditions, i.e., when no noise is present. This is because we propose generating the training data

numerically, which will help us by removing the need to make data acquisition in the environment where the network will be implemented. The training stage is summarized in Algorithm 1.

---

**Algorithm 1** Training Procedure

---

```

1: for  $i = 1:1:N_S$  do ▷  $N_S$ —Number of samples
2:    $(x_{x_i}, x_{y_i}) \leftarrow \text{Random} \in [l_b, u_b]$  ▷  $l_b$ —Lower Bound;  $u_b$ —Upper bound
3:   for  $j = 1:1:N$  do ▷  $N$ —Number of sensors
4:      $y_{ij} \leftarrow g_j^P / \|\mathbf{x}_i - \mathbf{s}_j\|^2$  ▷ Apply Expression (1) [noise-free ( $\sigma_{v_i}^2 = 0, \forall i$ )]
5:   end for
6: end for
7:  $\mathbf{y}_{\text{Train}} \leftarrow 70\% \cdot \mathbf{y}$  ▷ Training set (inputs)
8:  $\mathbf{y}_{\text{Valid}} \leftarrow 30\% \cdot \mathbf{y}$  ▷ Validation set (inputs)
9:  $\text{net} \leftarrow \text{train}(\mathbf{y}_{\text{Train}}, \mathbf{y}_{\text{Valid}})$  ▷ Levenberg–Marquardt Algorithm

```

---

Since it is the topology of the network and its training strategy that is under scrutiny, the remaining parameters and tools are considered as well-established methods in the literature. The LM optimization method is applied, using the mean squared error (MSE) as a performance metric.

Looking at the inference stage of a feed-forward neural network, the computational complexity is related to the number of inputs, the number of layers, and the number of perceptrons in each layer. Let  $\mathbf{W}_{ij}$  be the matrix of weights connecting layer  $i$  to layer  $j$ . Obtaining the outputs of layer  $i$  implies computing some  $\mathbf{S}_j = \mathbf{W}_{ij} \cdot \mathbf{Z}_i$ . From this point, the activation function  $f(x)$  is applied as  $\mathbf{Z}_j = f(\mathbf{S}_j)$ , where  $Z_j$  is the output of layer  $j$  obtained from the previous outputs of layer  $i$ ,  $Z_i$ . Thus, if the network has  $L$  layers (including an input and an output layer), the expression is evaluated  $(L - 1)$  times, where, for each layer, a matrix multiplication and an activation function is computed. Considering the MLP case, the weight matrix  $\mathbf{W}_{12} \in \mathbb{R}^{N \times p}$ , where  $p$  is the number of perceptrons in the hidden layer, while  $\mathbf{W}_{23} \in \mathbb{R}^{p \times 2}$ , since there are only two outputs. When considering the DFNN case, the matrices linking hidden layers belong to vector space  $\mathbb{R}^{p \times p}$ . Taking the complexity with regard to the network dimension, one can see that only  $\mathbf{W}_{12}$  will increase accordingly. This situation is valid for both the MLP and DFNN cases. As such, the complexity will then be of order  $O(n)$ , regardless of the network structure.

In summary, Figure 3 shows the two stages that establish the proposed method. Initially, a set of data is generated, where pairs of observations with the corresponding coordinates (lines 2 and 4 of Algorithm 1) are obtained. This data construction depends only on the model (gain, transmitting power, and sensors' positions) and the layout (upper and lower bound) features, and thus, do not need an environment-dependent data acquisition. In this offline stage, the weights of the network are obtained. The second stage matches the online processing phase. Here, noisy measurements are acquired through the sensors and inference in the network, obtaining the coordinates of the acoustic source. It should be noted that, firstly, no assumption regarding statistical features of the noise was made, contrarily to the methods presented in Section 2. Secondly, the inference of the network was performed with noisy measurements, while it was trained with a noise-free collection of pairs of simulated coordinates and observations.

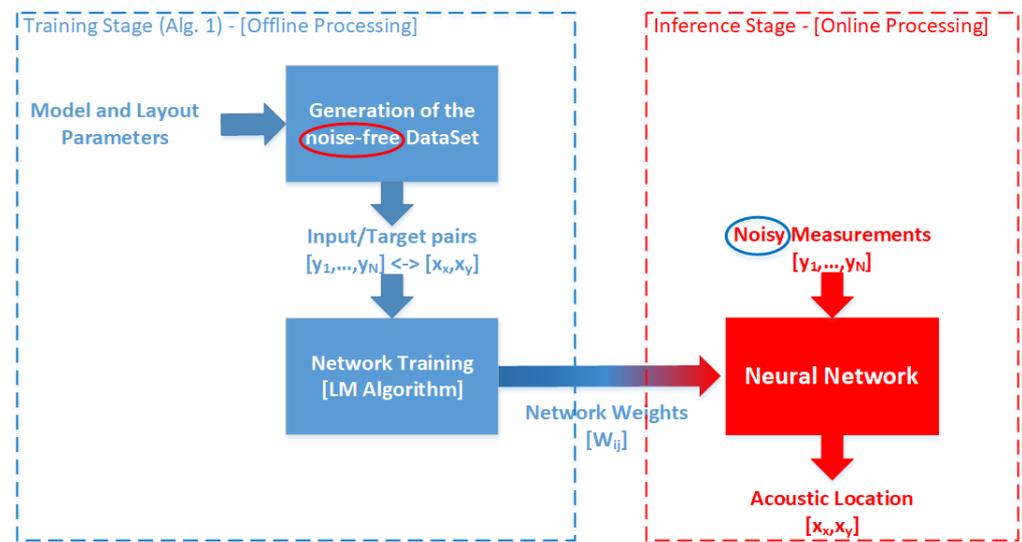


Figure 3. Overall strategy of the network.

### 5. Results and Discussion

The first hyperparameter to be addressed concerns the number of layers and the number of perceptrons in each layer. In order to study this hyperparameter, several scenarios were created that include arranging  $N$  ( $N = 3, 6, 9, 12,$  and  $15$ ) acoustic sensors on a circle centered at the middle of the search space, with radius equal to 50 m. The training data are randomly generated over the search space and made up of 10,000 coordinates  $(X_{x_i}, X_{y_i})$ , where 70% of the samples form the training set, and the remaining 30% form the validation set. Firstly, all of the samples are generated using MATLAB® R2020b on an AMD Ryzen™ R7-4700U Octa-Core processor, featuring 2 GHz and 16 GB of RAM, running on a Windows™ 10 operating system. The generated coordinates and the sensors' distribution over the search space are represented in Figure 4 when considering  $N = 6$ , where a distance of 5 m is considered to avoid singularities that would arise when the source matches the coordinates of the sensors (Expression (1)).

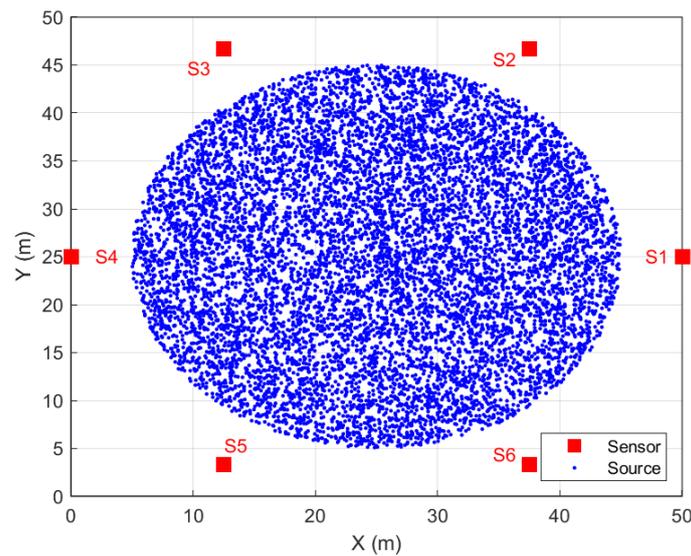


Figure 4. Random source distribution over the search space for  $N = 6$ .

Secondly, the corresponding observations are calculated under ideal noise-free conditions, by applying Expression (1) with  $v_i = 0$ . This strategy implies that the network is trained offline, with “measurements” obtained numerically, and as such is not dependent

on field acquisitions. Thirdly, the input data are scaled and normalized athwart Batch Normalization [67]. With this procedure over the input values, the fastest and more stable network is obtained, by recentering and rescaling the observation values [68]. Basically, the Batch process performs a rescaling of the input data over the domain of the activation function. As it is a linear mathematical operation, the neighboring information is preserved, and a centralizing and stretching effect is observed. The histogram of Figure 5 shows the distribution of the data obtained with the model (Expression (1)) in red and the result of applying the Batch Normalization algorithm in blue, corresponding to the total dataset when  $N = 3$ . It should be noticed that without this rescaling, the hyperbolic tangent activation function would behave as a linear one, since all values would tend to zero from positive values. The rescaling also highlights the fact that the input data approaches a Rayleigh distribution. This behavior is expected since the three sensor readings are independent and identically distributed Gaussian with equal variance and zero mean [69].

Regarding the training parameters, the LM optimization method is applied, using the mean squared error (MSE) as a performance metric, over 1000 epochs, a minimum performance gradient of  $10^{-7}$ , and a learning rate of 0.001. We consider that the training is finished when the total number of epochs is reached, the performance gradient falls below  $10^{-7}$ , or the performance is minimized to its goal, which that would be hypothetically null ( $MSE \rightarrow 0$ ). The transmitted power is set to  $P = 5$  and the sensors gain to  $g_i = 1$  for  $i = 1, \dots, N$ .

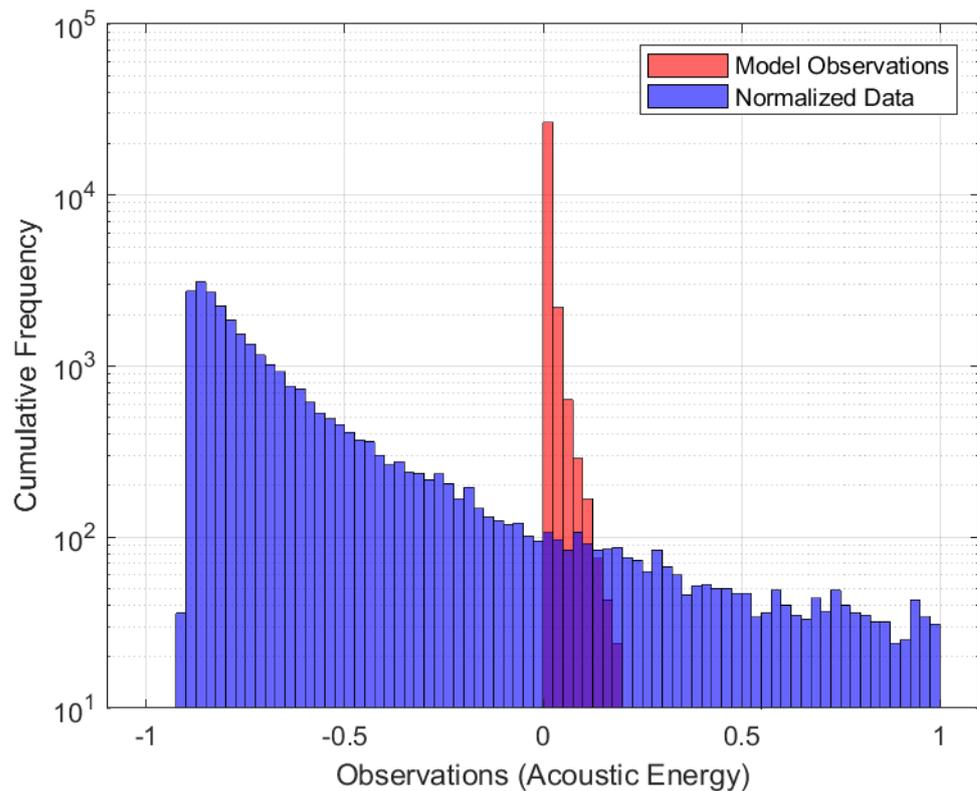
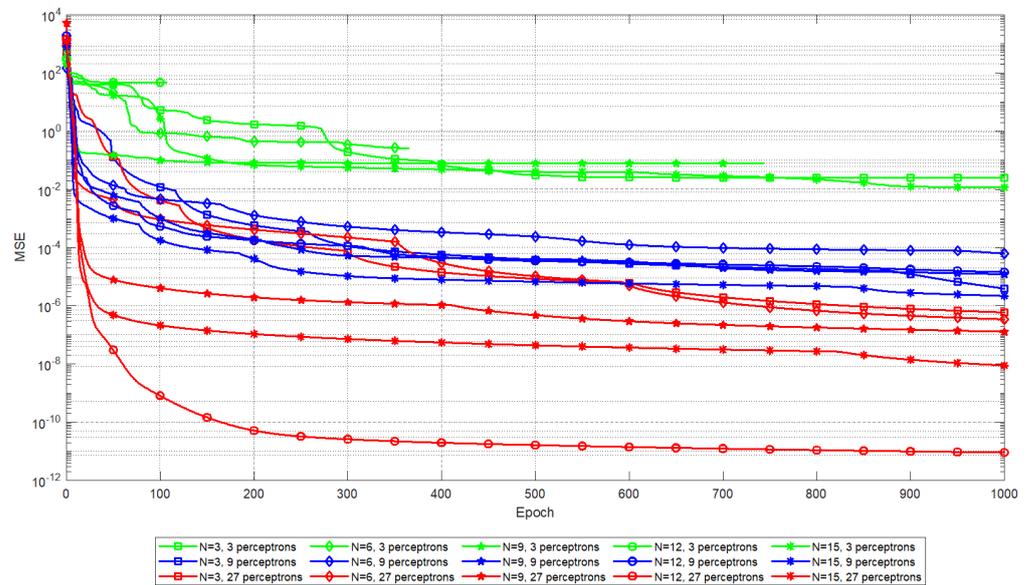


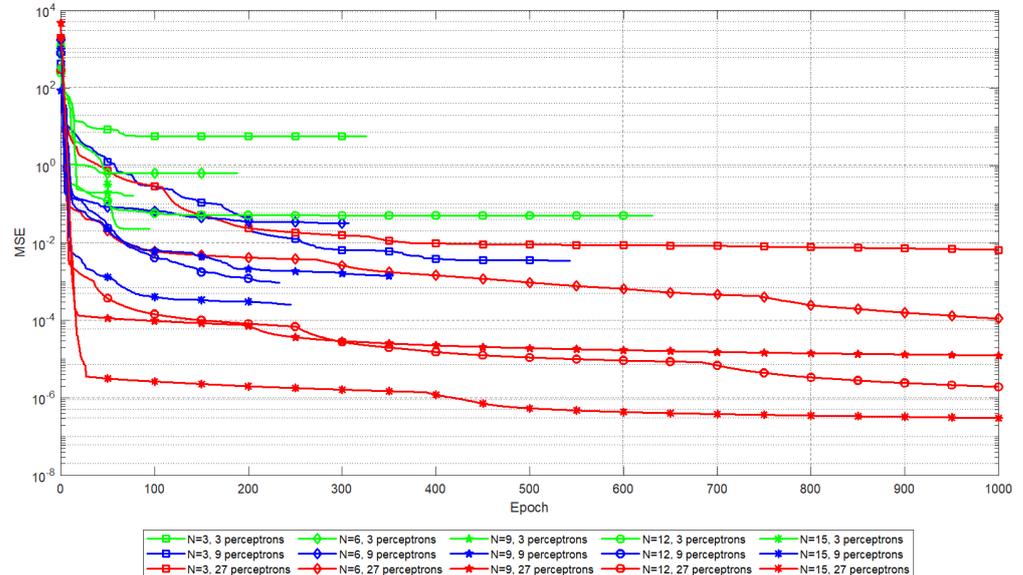
Figure 5. Histogram of the input data before and after normalization and rescaling.

To assert the performance of the proposed DFNN, all scenarios ( $N = 3$  to  $N = 15$  with increments of 3) are trained with network structures of 3, 9, and 27 perceptrons per layer (Figure 6). To demonstrate the effectiveness of the *Deep Learning* methodology, the same procedure is done on an MLP, consisting of only one input layer, one hidden layer, and one output layer (Figure 7). This comparison, the DFNN with the MLP, indirectly asserts the number of layers in the network, since networks with one and three hidden layers are being balanced out. The case of using only one hidden layer is identified as  $MLP_1$  throughout the

results discussion, and the DFNN with 3 hidden layers is identified as  $DFNN_3$ . It should be noted that the acronym DFNN is used here to distinguish the two networks in terms of the number of layers and that a DFNN is also an MLP network, only with a higher number of hidden layers.



**Figure 6.** Training Performance considering 3, 9, and 27 perceptrons per layer, for  $N = 3, 6, 9, 12,$  and 15 Sensors (DFNN Network).



**Figure 7.** Training Performance considering 3, 9, and 27 perceptrons per layer, for  $N = 3, 6, 9, 12,$  and 15 Sensors (MLP Network).

When analyzing the obtained MSE over the 1000 epochs of the  $MLP_1$  (Figure 7), one can see that the training stopped prematurely (due to gradient performance) for all scenarios considered, with different values for the number of sensors  $N$  for 3 and 9 perceptrons. The networks consisting of 27 perceptrons per layer reached 1000 epochs with an MSE of approximately  $6.5 \times 10^{-3}$  and  $3 \times 10^{-7}$  for  $N = 3$  and  $N = 15$ , respectively. Concerning the  $DFNN_3$ , only the case of  $N = 3, N = 6,$  and  $N = 9$  with 3 perceptrons per layer stopped prematurely, but the  $DFNN_3$  with 27 perceptrons per layer reached much lower MSE values ( $6 \times 10^{-7}$  for  $N = 3$  and  $9 \times 10^{-12}$  for  $N = 15$ ). This situation

demonstrates the superiority of the proposed DFNN over a standard MLP, justifying its implementation, although with higher complexity in training or network structure. Since the training is performed in an offline stage, and online processing involves mostly matrix summation and multiplication, this complexity increase does not directly affect the network behavior, increasing linearly with the number of sensors, as demonstrated. The performance of the newly presented localization method will be compared in terms of root-mean-squared error (RMSE) defined as

$$RMSE = \sqrt{\frac{\sum_{i=1}^M \|\mathbf{x}_i - \hat{\mathbf{x}}_i\|^2}{M}}, \tag{3}$$

where  $\mathbf{x}_i$  is the true (unknown) source location of the  $i$ th sample,  $\hat{\mathbf{x}}_i$  is the estimated coordinate through the DFNN, and  $M$  is the size of the data set. Since the RMSE measures the standard deviation of the residuals, the error maintains the dimension of the variables, allowing cross-checking of the error with the search space in meters. For that purpose, the  $MLP_1$  is considered for comparison as a simpler neural network structure (composed of one hidden layer). Based on the previous analysis (Figure 7), 27 perceptrons are considered for the hidden layer. Additionally, two more methods are selected from state-of-the-art deterministic and metaheuristic algorithms. With regard to the deterministic one, although second-order cone programming can be considered as the current state-of-the-art [28,29], providing good results even for noisy environments, the bisection method proposed in [25] is considered due to its simplicity and reliability, denoted here by “EXACT”. With regard to metaheuristic algorithms, the most recent version of EHO for acoustic localization is considered [32], and named “EEHO”. Both methods that are compared with the present work have complexity in the order of  $O(n)$  [33]. The employed strategy to validate the four algorithms consists of generating 10,000 samples randomly distributed over the search space (Figure 4). Secondly, seven new sample sets are created by adding different noise levels, generated from a Normal distribution. More specifically, the noise term  $v_i$  in Expression (1) follows a normal distribution with zero mean and variance  $\sigma_{v_i}^2$ , that is,  $v_i \sim \mathcal{N}(0, \sigma_{v_i}^2)$ . The variance is considered in a logarithmic scale ( $\sigma_{v_i}^2(\text{dB}) = 10 \times \log_{10}(\sigma_{v_i}^2)$ ) on the interval  $[-80, -50]$  dB, with increments of 5 dB. The performance comparison methods is evaluated in terms of RMSE. The noise variance interval intends to cover a wide range of the signal-noise-ratio (SNR). When considering a source at 1 m of the sensor, the SNR varies from 74 dB to 104 dB (dB ref. 1 W/m<sup>2</sup>), but when considering a distance of 50 m—which would correspond to a source located at the edge of the search space—the SNR varies from approximately  $-5$  dB to 30 dB (dB ref. 1 W/m<sup>2</sup>).

The obtained results for both layouts ( $N = 3$  and  $N = 12$  sensors) are represented in Figure 8, with regard to the newly proposed  $DFNN_3$  and the three algorithms chosen for comparison ( $MLP_1$ ,  $EEHO$ , and  $EXACT$ ). Concerning lower values of the measurement noise, the performance of the considered methods is quite diverse. In this situation, the  $EXACT$  method has the highest error (8.45 m for  $\sigma_{v_i}^2 = -80$  dB and  $N = 3$ ); followed by  $MLP_1$  (3.41 m for  $\sigma_{v_i}^2 = -80$  dB and  $N = 3$ );  $EEHO$  (1.57 for  $\sigma_{v_i}^2 = -80$  dB and  $N = 3$ ); and finally, with the lowest error, the proposed  $DFNN_3$  (0.53 for  $\sigma_{v_i}^2 = -80$  dB and  $N = 3$ ). When considering  $N = 12$ , the proposed  $DFNN_3$  reaches values of error as low as 0.096 m for  $\sigma_{v_i}^2 = -80$  dB.

When analyzing the methods’ behavior for higher values of measurement noise, the results obtained tend to overlap and the differences of the errors are not as pronounced. Considering  $N = 12$  and  $\sigma_{v_i}^2 = -50$  dB, while the  $EXACT$  method shows an error of 17.53 m,  $MLP_1$ ,  $DFNN_3$ , and  $EEHO$  show 9.80 m, 8.77 m, and 8.50 m, respectively. This means that, although the  $EXACT$  method stands out in a negative sense, the remaining methods show an equivalent behavior. Even so, the proposed  $DFNN_3$  assumes greater simplicity in terms of its implementation when compared to  $EEHO$ . As a last remark, it is worth mentioning that  $DFNN_3$  for  $N = 3$  performs concurrently with its counterparts when considering  $N = 12$ . This situation implies that the proposed method needs less

sensors to achieve the performance of others. This  $DFNN_3$  network behavior is not surprising since the network has been trained with synthetic or ideal observations. Even so, it works relatively well in scenarios where the noise is not small (or at least not worse than the other methods).

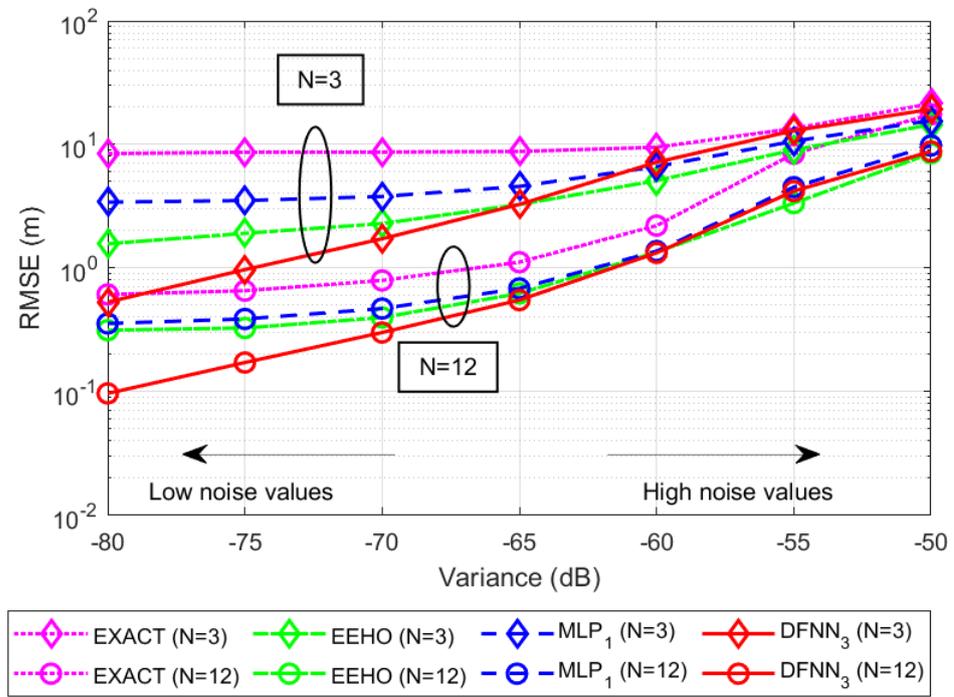


Figure 8. RMSE Comparison of  $EXACT$ ,  $EEHO$ ,  $MLP_1$ , and  $DFNN_3$  for  $N = 3$  and  $N = 12$ .

As can be seen from Figure 8, when noise power is low to medium, there is quite a performance margin between the proposed method and the remaining ones (e.g., for  $N = 3$ , noise variance =  $-80$  dB, the proposed method outperforms  $EEHO$ ,  $MLP_1$ , and  $EXACT$  for around 1, 3, and 8 m, respectively). This behavior can be explained by the fact that the proposed method is trained with noise-free data, which allows for better accuracy for low-to-medium noise power. As the noise power gets higher, one can see that the performance of all methods deteriorates significantly, and that all of them have roughly the same location accuracy.

With the goal of getting an even better insight on the performance of the proposed  $DFNN_3$ , we employ another performance metric in Figures 9 and 10. The figures illustrate the Cumulative Distribution Function (CDF) of the localization error (LE),  $LE = \|x_i - \hat{x}_i\|_2$ , over the  $M$  samples, when  $\sigma_{v_i}^2 = -80$  dB for  $N = 3$  and  $N = 12$ . From Figure 9, for  $N = 3$ , one can see that the proposed solution achieves  $LE < 1$  m in 90% of the cases, while  $LE < 2.5$  m and  $LE < 5$  m achieved for  $EEHO$  and  $MLP_1$  in the same percentage, respectively, whereas  $EXACT$  achieves  $LE \leq 10$  m in around 75% of the cases. When  $N = 12$  (Figure 10), one can observe  $LE < 0.15$  m for  $DFNN_3$ , while the  $EEHO$ ,  $MLP_1$ , and  $EXACT$  methods achieve only  $LE < 0.32$  m,  $LE < 0.45$  m, and  $LE < 0.5$  m in the same percentage of cases, respectively. The obtained results are in line with the previous analysis regarding RMSE performance, confirming the effectiveness of the proposed solution.

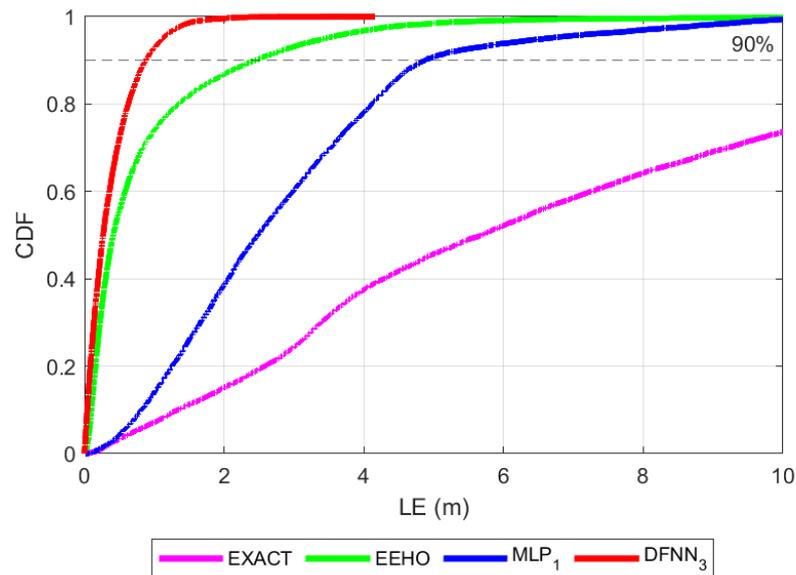


Figure 9. Cumulative Distribution Function of LE for  $\sigma_{v_i}^2 = -80$  dB and  $N = 3$ .

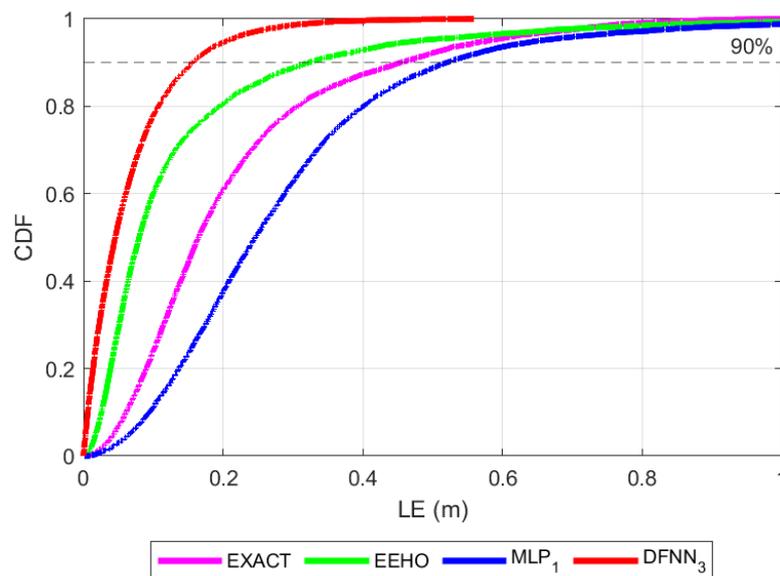


Figure 10. Cumulative Distribution Function of LE for  $\sigma_{v_i}^2 = -80$  dB and  $N = 12$ .

### 6. Conclusions

A new method based on DFNNs for solving the energy-based acoustic localization problem is proposed in the present work. Its particularity consists of a reliable and straightforward training stage, where the dataset is generated under ideal environmental conditions and noise-free measurements. The simulation results demonstrate that the methodology exceeds the performance of the state-of-the-art for lower values of measurement noise, while it matches its performance in noisy environments. Concerning environments with a higher noise level, the proposed method matches the performance of its counterparts. The new methodology paves the way to machine learning, more specifically, the use of Neural Networks, for the acoustic localization problem.

This work considered the localization of a single stationary source. Generalizing the presented algorithm for the localization of nonstationary sources through extending the feed-forward network to process time serialized data, is to be considered as future work.

**Author Contributions:** All authors (S.D.C., S.T., M.B.) have contributed equally to the work reported. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Fundação para a Ciência e a Tecnologia under Projects UIDB/04111/2020 and foRESTER PCIF/SSI/0102/2017; and Instituto Lusófono de Investigação e Desenvolvimento (ILIND) under Project COFAC/ILIND/COPELABS/1/2020.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Uziel, S.; Elste, T.; Kattanek, W.; Hollosi, D.; Gerlach, S.; Goetze, S. Networked embedded acoustic processing system for smart building applications. In Proceedings of the 2013 Conference on Design and Architectures for Signal and Image Processing, Cagliari, Italy, 8–10 October 2013; pp. 349–350.
2. Quintana-Suárez, M.; Sánchez-Rodríguez, D.; Alonso-González, I.; Alonso-Hernández, J. A Low Cost Wireless Acoustic Sensor for Ambient Assisted Living Systems. *Appl. Sci.* **2017**, *7*, 877. [[CrossRef](#)]
3. Toky, A.; Singh, R.P.; Das, S. Localization schemes for Underwater Acoustic Sensor Networks—A Review. *Comput. Sci. Rev.* **2020**, *37*, 100241. [[CrossRef](#)]
4. Taylor, C.E.; Huang, Y.; Yao, K. Distributed sensor swarms for monitoring bird behavior: An integrated system using wildlife acoustics recorders. *Artif. Life Robot.* **2016**, *21*, 268–273. [[CrossRef](#)]
5. Lopatka, K.; Kotus, J.; Czyzewski, A. Detection, classification and localization of acoustic events in the presence of background noise for acoustic surveillance of hazardous situations. *Multimed. Tools Appl.* **2015**, *75*, 10407–10439. [[CrossRef](#)]
6. Maroti, M.; Simon, G.; Ledeczi, A.; Sztipanovits, J. Shooter localization in urban terrain. *Computer* **2004**, *37*, 60–61. [[CrossRef](#)]
7. Lammert-Siepmann, N.; Bestgen, A.K.; Edler, D.; Kuchinke, L.; Dickmann, F. Audiovisual communication of object-names improves the spatial accuracy of recalled object-locations in topographic maps. *PLoS ONE* **2017**, *12*, e0186065. [[CrossRef](#)] [[PubMed](#)]
8. Reagan, I.; Baldwin, C.L. Facilitating route memory with auditory route guidance systems. *J. Environ. Psychol.* **2006**, *26*, 146–155. [[CrossRef](#)]
9. Xu, E.; Ding, Z.; Dasgupta, S. Source Localization in Wireless Sensor Networks From Signal Time-of-Arrival Measurements. *IEEE Trans. Signal Process.* **2011**, *59*, 2887–2897. [[CrossRef](#)]
10. Yang, L.; Ho, K.C. An Approximately Efficient TDOA Localization Algorithm in Closed-Form for Locating Multiple Disjoint Sources With Erroneous Sensor Positions. *IEEE Trans. Signal Process.* **2009**, *57*, 4598–4615. [[CrossRef](#)]
11. Ali, A.M.; Yao, K.; Collier, T.C.; Taylor, C.E.; Blumstein, D.T.; Girod, L. An Empirical Study of Collaborative Acoustic Source Localization. In Proceedings of the 2007 6th International Symposium on Information Processing in Sensor Networks, Cambridge, MA, USA, 25–27 April 2007. [[CrossRef](#)]
12. He, T.; Huang, C.; Blum, B.M.; Stankovic, J.A.; Abdelzaher, T.F. Range-Free Localization and Its Impact on Large Scale Sensor Networks. *ACM Trans. Embed. Comput. Syst.* **2005**, *4*, 877–906. [[CrossRef](#)]
13. Singh, S.P.; Sharma, S. Range Free Localization Techniques in Wireless Sensor Networks: A Review. *Procedia Comput. Sci.* **2015**, *57*, 7–16. [[CrossRef](#)]
14. Hu, Y.H.; Li, D. Energy based collaborative source localization using acoustic micro-sensor array. In Proceedings of the 2002 IEEE Workshop on Multimedia Signal Processing, St. Thomas, VI, USA, 9–11 December 2002; pp. 371–375. [[CrossRef](#)]
15. Sheng, X.; Hu, Y. Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks. *IEEE Trans. Signal Process.* **2005**, *53*, 44–53. [[CrossRef](#)]
16. Shareef, A.; Zhu, Y. Localization Using Extended Kalman Filters in Wireless Sensor Networks. In *Kalman Filter Recent Advances and Applications*; I-Tech: Vienna, Austria, 2009.
17. Wang, X.; Mao, S. Deep learning for indoor localization based on bimodal CSI data. *Appl. Mach. Learn. Wirel. Commun.* **2019**, *81*, 343–369. [[CrossRef](#)]
18. Wang, Y.; Gao, J.; Li, Z.; Zhao, L. Robust and Accurate Wi-Fi Fingerprint Location Recognition Method Based on Deep Neural Network. *Appl. Sci.* **2020**, *10*, 321. [[CrossRef](#)]
19. Wolf, B.J.; Wolfshaar, J.V.D.; Netten, S.M.V. Three-dimensional multi-source localization of underwater objects using convolutional neural networks for artificial lateral lines. *J. R. Soc. Interface* **2020**, *17*, 20190616. [[CrossRef](#)] [[PubMed](#)]
20. Livingstone, D.J. *Artificial Neural Networks: Methods and Applications*; Humana Press: Totowa, NJ, USA, 2009; ISBN 978-1-60327-101-1. [[CrossRef](#)]
21. Parathai, P.; Tengtrairat, N.; Woo, W.L.; Abdullah, M.A.M.; Rafiee, G.; Alshabrawy, O. Efficient Noisy Sound-Event Mixture Classification Using Adaptive-Sparse Complex-Valued Matrix Factorization and OvsO SVM. *Sensors* **2020**, *20*, 4368. [[CrossRef](#)] [[PubMed](#)]
22. Meng, W.; Xiao, W. Energy-Based Acoustic Source Localization Methods: A Survey. *Sensors* **2017**, *17*, 376. [[CrossRef](#)]
23. Meesookho, C.; Mitra, U.; Narayanan, S. On Energy-Based Acoustic Source Localization for Sensor Networks. *IEEE Trans. Signal Process.* **2008**, *56*, 365–377. [[CrossRef](#)]
24. Ho, K.C.; Sun, M. An Accurate Algebraic Closed-Form Solution for Energy-Based Source Localization. *IEEE Trans. Audio Speech Lang. Process.* **2007**, *15*, 2542–2550. [[CrossRef](#)]

25. Beck, A.; Stoica, P.; Li, J. Exact and Approximate Solutions of Source Localization Problems. *IEEE Trans. Signal Process.* **2008**, *56*, 1770–1778. [[CrossRef](#)]
26. Wang, G. A Semidefinite Relaxation Method for Energy-Based Source Localization in Sensor Networks. *IEEE Trans. Veh. Technol.* **2011**, *60*, 2293–2301. [[CrossRef](#)]
27. Beko, M. Energy-based localization in wireless sensor networks using semidefinite relaxation. In Proceedings of the 2011 IEEE Wireless Communications and Networking Conference, Cancun, Mexico, 28–31 March 2011; pp. 1552–1556. [[CrossRef](#)]
28. Beko, M. Energy-Based Localization in Wireless Sensor Networks Using Second-Order Cone Programming Relaxation. *Wirel. Pers. Commun.* **2014**, *77*, 1847–1857. [[CrossRef](#)]
29. Yan, Y.; Shen, X.; Hua, F.; Zhong, X. On the Semidefinite Programming Algorithm for Energy-Based Acoustic Source Localization in Sensor Networks. *IEEE Sens. J.* **2018**, *18*, 8835–8846. [[CrossRef](#)]
30. Shi, J.; Wang, G.; Jin, L. Robust Semidefinite Relaxation Method for Energy-Based Source Localization: Known and Unknown Decay Factor Cases. *IEEE Access* **2019**, *7*, 163740–163748. [[CrossRef](#)]
31. Deb, S.; Fong, S.; Tian, Z. Elephant Search Algorithm for optimization problems. In Proceedings of the 2015 Tenth International Conference on Digital Information Management (ICDIM), Jeju, Korea, 21–23 October 2015; pp. 249–255. [[CrossRef](#)]
32. Wang, G.G.; Coelho, L.D.S.; Gao, X.Z.; Deb, S. A new metaheuristic optimisation algorithm motivated by elephant herding behaviour. *Int. J. Bio Inspired Comput.* **2016**, *8*, 394. [[CrossRef](#)]
33. Correia, S.D.; Beko, M.; Cruz, L.; Tomic, S. Elephant Herding Optimization for Energy-Based Localization. *Sensors* **2018**, *18*, 2849. [[CrossRef](#)]
34. Correia, S.D.; Beko, M.; Da Silva Cruz, L.A.; Tomic, S. Implementation and Validation of Elephant Herding Optimization Algorithm for Acoustic Localization. In Proceedings of the 2018 26th Telecommunications Forum (TELFOR), Belgrade, Serbia, 20–21 November 2018; pp. 1–4. [[CrossRef](#)]
35. Correia, S.D.; Beko, M.; Tomic, S.; Da Silva Cruz, L.A. Energy-Based Acoustic Localization by Improved Elephant Herding Optimization. *IEEE Access* **2020**, *8*, 28548–28559. [[CrossRef](#)]
36. Correia, S.D.; Fé, J.; Tomic, S.; Beko, M. Development of a Test-Bench for Evaluating the Embedded Implementation of the Improved Elephant Herding Optimization Algorithm Applied to Energy-Based Acoustic Localization. *Computers* **2020**, *9*, 87. [[CrossRef](#)]
37. Cobos, M.; Antonacci, F.; Alexandridis, A.; Mouchtaris, A.; Lee, B. A Survey of Sound Source Localization Methods in Wireless Acoustic Sensor Networks. *Wirel. Commun. Mob. Comput.* **2017**, *2017*, 3956282. [[CrossRef](#)]
38. Sheng, X.; Hu, Y.H. Energy Based Acoustic Source Localization. In *Information Processing in Sensor Networks Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2003; pp. 285–300.
39. Meng, W.; Xiao, W.; Xie, L. An Efficient EM Algorithm for Energy-Based Multisource Localization in Wireless Sensor Networks. *IEEE Trans. Instrum. Meas.* **2011**, *60*, 1017–1027. [[CrossRef](#)]
40. Lu, L.; Wu, H.C. Novel energy-based localization technique for multiple sources. In Proceedings of the 2012 IEEE International Conference on Communications (ICC), Ottawa, ON, Canada, 10–15 June 2012. [[CrossRef](#)]
41. Chang, S.; Li, Y.; He, Y.; Wang, H. Target Localization in Underwater Acoustic Sensor Networks Using RSS Measurements. *Appl. Sci.* **2018**, *8*, 225. [[CrossRef](#)]
42. Morar, A.; Moldoveanu, A.; Mocanu, I.; Moldoveanu, F.; Radoi, I.E.; Asavei, V.; Gradinaru, A.; Butean, A. A Comprehensive Survey of Indoor Localization Methods Based on Computer Vision. *Sensors* **2020**, *20*, 2641. [[CrossRef](#)] [[PubMed](#)]
43. Mao, G. Localization algorithms and strategies for wireless sensor networks monitoring and surveillance techniques for target tracking. In *Information Science Reference*; IGI Global: Hershey, PA, USA, 2009; Chapter 3. [[CrossRef](#)]
44. Bhatti, G. Machine Learning Based Localization in Large-Scale Wireless Sensor Networks. *Sensors* **2018**, *18*, 4179. [[CrossRef](#)] [[PubMed](#)]
45. Ahmadi, H.; Bouallegue, R. Comparative study of learning-based localization algorithms for Wireless Sensor Networks: Support Vector regression, Neural Network and Naive Bayes. In Proceedings of the 2015 International Wireless Communications and Mobile Computing Conference (IWCMC), Dubrovnik, Croatia, 24–28 August 2015. [[CrossRef](#)]
46. Alexandridis, A.; Mouchtaris, A. Multiple Sound Source Location Estimation in Wireless Acoustic Sensor Networks using DOA estimates: The Data-Association Problem. *IEEE/ACM Trans. Audio Speech Lang. Process.* **2017**. [[CrossRef](#)]
47. Guo, Y.; Zhu, H.; Dang, X. Tracking multiple acoustic sources by adaptive fusion of TDOAs across microphone pairs. *Digit. Signal Process.* **2020**, *106*, 102853. [[CrossRef](#)]
48. Evers, C.; Dorfan, Y.; Gannot, S.; Naylor, P.A. Source tracking using moving microphone arrays for robot audition. In Proceedings of the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 5–9 March 2017. [[CrossRef](#)]
49. Sun, S.; Zhang, X.; Zheng, C.; Fu, J.; Zhao, C. Underwater Acoustical Localization of the Black Box Utilizing Single Autonomous Underwater Vehicle Based on the Second-Order Time Difference of Arrival. *IEEE J. Ocean. Eng.* **2020**, *45*, 1268–1279. [[CrossRef](#)]
50. Tomic, S.; Beko, M.; Dinis, R.; Tuba, M.; Bacanin, N. Bayesian methodology for target tracking using combined RSS and AoA measurements. *Phys. Commun.* **2017**, *25*, 158–166. [[CrossRef](#)]
51. Frank, R.J.; Davey, N.; Hunt, S.P. Time Series Prediction and Neural Networks. *J. Intell. Robot. Syst.* **2001**, *31*, 91–103. [[CrossRef](#)]
52. Koh, B.H.D.; Lim, C.L.P.; Rahimi, H.; Woo, W.L.; Gao, B. Deep Temporal Convolution Network for Time Series Classification. *Sensors* **2021**, *21*, 603. [[CrossRef](#)] [[PubMed](#)]

53. Sangiorgio, M.; Dercole, F. Robustness of LSTM neural networks for multi-step forecasting of chaotic time series. *Chaos Solitons Fractals* **2020**, *139*, 110045. [[CrossRef](#)]
54. Cybenko, G. Approximation by superpositions of a sigmoidal function. *Math. Control Signals Syst.* **1989**, *2*, 303–314. [[CrossRef](#)]
55. Hornik, K. Approximation capabilities of multilayer feedforward networks. *Neural Netw.* **1991**, *4*, 251–257. [[CrossRef](#)]
56. Sonoda, S.; Murata, N. Neural network with unbounded activation functions is universal approximator. *Appl. Comput. Harmon. Anal.* **2017**, *43*, 233–268. [[CrossRef](#)]
57. Kalman, B.L.; Kwasny, S.C. Why tanh: Choosing a sigmoidal function. In Proceedings of the IJCNN International Joint Conference on Neural Networks, Baltimore, MD, USA, 7–11 June 1992; Volume 4, pp. 578–581. [[CrossRef](#)]
58. Abdulkarim, M.; Shafie, A.; Ahmad, W.F.W.; Razali, R. Performance Comparison between MLP Neural Network and Exponential Curve Fitting on Airwaves Data. In *Lecture Notes in Electrical Engineering Advances in Computer Science and Its Applications*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 1079–1087. [[CrossRef](#)]
59. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. [[CrossRef](#)]
60. Battiti, R. First- and Second-Order Methods for Learning: Between Steepest Descent and Newton’s Method. *Neural Comput.* **1992**, *4*, 141–166. [[CrossRef](#)]
61. Hagan, M.T.; Menhaj, M.B. Training feedforward networks with the Marquardt algorithm. *IEEE Trans. Neural Netw.* **1994**, *5*, 989–993. [[CrossRef](#)]
62. Stathakis, D. How many hidden layers and nodes? *Int. J. Remote Sens.* **2009**, *30*, 2133–2147. [[CrossRef](#)]
63. Huang, G. Learning capability and storage capacity of two-hidden-layer feedforward networks. *IEEE Trans. Neural Netw.* **2003**, *14*, 274–281. [[CrossRef](#)] [[PubMed](#)]
64. Tamura, S. Capabilities of a three layer feedforward neural network. In Proceedings of the 1991 IEEE International Joint Conference on Neural Networks, Singapore, 18–21 November 1991. [[CrossRef](#)]
65. Shin-ike, K. A two phase method for determining the number of neurons in the hidden layer of a 3-layer neural network. In Proceedings of the SICE Annual Conference 2010, Taipei, Taiwan, 18–21 August 2010; pp. 238–242.
66. Sheela, K.G.; Deepa, S.N. Review on Methods to Fix Number of Hidden Neurons in Neural Networks. *Math. Probl. Eng.* **2013**, *2013*, 425740. [[CrossRef](#)]
67. Sun, J.; Cao, X.; Liang, H.; Huang, W.; Chen, Z.; Li, Z. New Interpretations of Normalization Methods in Deep Learning. *Proc. AAAI Conf. Artif. Intell.* **2020**, *34*, 5875–5882. [[CrossRef](#)]
68. Ioffe, S.; Szegedy, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. In Proceedings of the 32nd International Conference on International Conference on Machine Learning-ICML’15, Lille, France, 6–11 July 2015; Volume 37, pp. 448–456.
69. Beichelt, F. Functions of Random Variables. *Appl. Probab. Stoch. Process.* **2018**, 155–198. [[CrossRef](#)]