


Article

MVDR-LSTM Distance Estimation Model Based on Diagonal Double Rectangular Array

Xiong Zhang ^{1,2,3} , Wenbo Wu ³, Jialu Li ³, Fan Dong ³ and Shuting Wan ^{1,2,3,*}

- ¹ Hebei Key Laboratory of Electric Machinery Health Maintenance & Failure Prevention, North China Electric Power University, Baoding 071003, China; hdjzx@ncepu.edu.cn
- ² Hebei Engineering Research Center for Advanced Manufacturing & Intelligent Operation and Maintenance of Electric Power Machinery, North China Electric Power University, Baoding 071003, China
- ³ Department of Mechanical Engineering, North China Electric Power University, Baoding 071003, China; w15031435169@163.com (W.W.); ljl2481390881@163.com (J.L.); df20000109@163.com (F.D.)
- * Correspondence: 220212224080@ncepu.edu.cn; Tel.: +86-137-0332-3816

Abstract: Deep learning algorithms have the advantages of a powerful time series prediction ability and the real-time processing of massive samples of big data. Herein, a new roller fault distance estimation method is proposed to address the problems of the simple structure and long conveying distance of belt conveyors. In this method, a diagonal double rectangular microphone array is used as the acquisition device, minimum variance distortionless response (MVDR) and long short-term memory network (LSTM) are used as the processing models, and the roller fault distance data are classified to complete the estimation of the idler fault distance. The experimental results showed that this method could achieve high-accuracy fault distance identification in a noisy environment and had better accuracy than the conventional beamforming algorithm (CBF)-LSTM and functional beamforming algorithm (FBF)-LSTM. In addition, this method could also be applied to other industrial testing fields and has a wide range of application prospects.

Keywords: diagonal double rectangular array; MVDR; LSTM; idler failure distance estimation



Citation: Zhang, X.; Wu, W.; Li, J.; Dong, F.; Wan, S. MVDR-LSTM Distance Estimation Model Based on Diagonal Double Rectangular Array. *Sensors* **2023**, *23*, 5094. <https://doi.org/10.3390/s23115094>

Academic Editors: Jinrui Wang, Zongzhen Zhang, Xingkai Yang and Ke Zhao

Received: 19 April 2023

Revised: 20 May 2023

Accepted: 22 May 2023

Published: 26 May 2023



Copyright: © 2023 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

In recent years, due to the increasing emphasis on energy conservation and emission reduction, straw power plants have developed rapidly. As an important piece of equipment for conveying materials, the safety of idlers in belt conveyors has a crucial impact on the smooth operation of the entire conveying system. As the key support and rotating component of the belt conveyor, the chain reaction caused by the failure of the idler poses a safety threat to the entire straw power station. The long-term operation, substantial conveying distances, and huge weight of idlers in harsh working environments necessitate daily inspections and regular maintenance that not only entail significant time and money costs, but also allow for some seriously degraded parts to be overdue for replacement. Therefore, it is of great significance to design non-contact acoustic signal measurement and fault distance estimation methods suitable for belt conveyor idlers in power stations, so as to reasonably arrange maintenance activities; reduce time, manpower, and material costs; and ensure the reliability and safety of mine belt conveyor operation.

In the field of idler bearing fault diagnosis, a method of energy collection was proposed in [1] that was of certain practical value for bearing fault diagnosis. The authors of [2] proposed an electromagnetic energy collector for bearing fault diagnosis that improved the diagnosis efficiency and reduced the difficulty of diagnosis.

With the acceleration of industrial processes, the safety of machinery and equipment has become the top priority, and methods for the fault diagnosis and condition detection of machinery and equipment are also constantly improving and upgrading. In recent years, the use of deep learning to diagnose bearings has also become a popular research

direction. Zhou proposed a novel combination of discrete probabilistic entropy-based health indicators (HIs) and long short-term memory (LSTM)-based methods to predict bearing health. The LSTM model included in the proposed method could accurately predict bearing health [3]. Aljemely et al. designed a combination of long short-term memory and large margin nearest neighbor (LSTM-LMNN), which improved work efficiency and diagnostic accuracy and was superior to traditional diagnostic methods in terms of stability and reliability [4]. In order to increase the effectiveness of convolutional neural networks (CNNs) for bearing fault feature recognition, Xie et al. proposed four different CNN hybrid models, and after an experimental comparison, the CNN-random forest (RF) and CNN-support vector machine (SVM) algorithms were found to make full use of CNN feature extraction capabilities [5]. Yang et al. proposed a method based on the combination of variable mode decomposition (VMD) and a CNN. Firstly, after collecting the motor bearing fault signal, the motor bearing data were denoised through VMD, and the denoising data were extracted through a CNN for diagnosis. Experiments showed that the accuracy of VMD-CNN for motor bearing fault diagnosis was better than that of a CNN [6]. Yu et al. proposed a novel real-time monitoring method for sediment particles, combining a stepoelectric nanogenerator driven by particle droplets (PLDD-TENG) and deep learning technology. By analyzing the output signal of PLDD-TENG, this method could sensitively reflect different particle types and concentrations [7].

Beamforming suppresses signals in non-selected directions and enhances signals in selected directions by merging the signals collected by the microphone array; in addition, it can realize focused pickup in the specified direction, which can effectively improve the signal-to-noise ratio of the received signal. As beamforming algorithms have become more popular, they have also played an increasingly important role in the field of sound source identification [8–11]. MVDR, a commonly used beamforming algorithm, is widely used in the field of fault diagnosis. The main advantages of the MVDR algorithm are summarized below.

Efficient fault feature extraction: The MVDR algorithm can effectively extract fault signal features. Due to its distortion-free nature, the MVDR algorithm can minimize noise and interference while retaining the fault signal, thereby improving the accuracy of fault detection.

Excellent spatial resolution: In terms of fault source location, the high spatial resolution of the MVDR algorithm can provide more accurate fault source location information.

Adaptability: In complex industrial environments, troubleshooting conditions can change. As an adaptive algorithm, MVDR can automatically adjust parameters according to changes in the environment, so as to effectively diagnose faults under various conditions.

Robustness: The MVDR algorithm has good robustness for independent and uniformly distributed noise signals, which allows it to maintain good fault diagnosis performance in various noise environments.

Subramanian et al. used a deep learning neural network to identify and locate multiple sound sources on the basis of direction of arrival estimation, and experimental results showed that the proposed method could classify DOAs with high resolution [12]. Considering the problem of far-field speech processing, Zhang et al. proposed a deep special beamforming technology that combined a microphone array with multi-channel speech enhancement based on deep learning and significantly reduced the probability of a far-field acoustic environment [13]. Yang et al. proposed an MVDR algorithm based on deep learning for extracting target speakers from hybrid speech that first used the target voice signal for estimation as the basis for channel selection and then applied the selected signal to the MVDR algorithm based on deep learning to extract the target speaker from hybrid speech. This method has achieved good results in both simulation and practical applications [14]. Ramezanpour and Mosavi proposed a neural beamformer for estimating the desired signal in noisy and interference environments that consisted of a convolutional neural network (CNN) and bidirectional long short-term memory (BiLSTM) to estimate the interference vector of the signal received by the antenna subarray and the sample of

the desired signal, respectively, achieving a higher signal-to-noise ratio than traditional beamforming algorithms [15].

The combination of deep learning and beamforming algorithms can better solve some tricky problems and has unique advantages in accelerating fault diagnosis efficiency and improving fault diagnosis accuracy, which not only reduces manual operation, but also makes mechanical equipment run more safely and smoothly. In order to solve the problem of the harsh working environment and serious noise interference of idlers, an MVDR-LSTM fault distance estimation model with a diagonal double rectangular array is proposed. This model can greatly suppress the interference of noise and non-coherent sound sources and realize the accurate estimation of the failure distance of idlers. At the same time, compared with the CBF-LSTM and FBF-LSTM models, our model has better robustness.

This article will first explore the relevant theory and historical background of the MVDR-LSTM roller fault distance estimation model. Secondly, the methods and processes adopted in this study will be introduced, including the specific steps of data collection and analysis. Then, we will elaborate on our research results and attempt to provide a deeper understanding of the MVDR-LSTM idler fault distance estimation model by comparing and interpreting these results. Finally, we will summarize our findings and explore their potential impact on the field of fault detection, as well as future research directions.

2. Theory

2.1. Microphone Array

A microphone array is an advanced technology used to capture sound signals that consists of multiple tiny microphones and can receive sound signals in different directions. Microphone arrays provide a more accurate and clear sound signal than a single microphone, because they suppress ambient noise and echo and can focus on receiving sound in different directions. Microphone arrays are widely used in many fields, such as video conferencing, speech recognition, and audio signal processing. In video conferencing, microphone arrays can provide clear voice signals, improving the efficiency and quality of meetings. In speech recognition, microphone arrays can improve the accuracy of speech recognition, because they reduce the impact of ambient noise. In audio signal processing, microphone arrays can realize functions such as spatial filtering and sound source localization to improve the quality of audio signals. As a result, microphone arrays have become an important part of modern communication and audio technology, and they are constantly improving our daily lives and work experiences [16].

The formation of a microphone array refers to the arrangement of microphones, and different formations can be adapted to different application scenarios. The common forms of microphone arrays are shown in Figure 1.

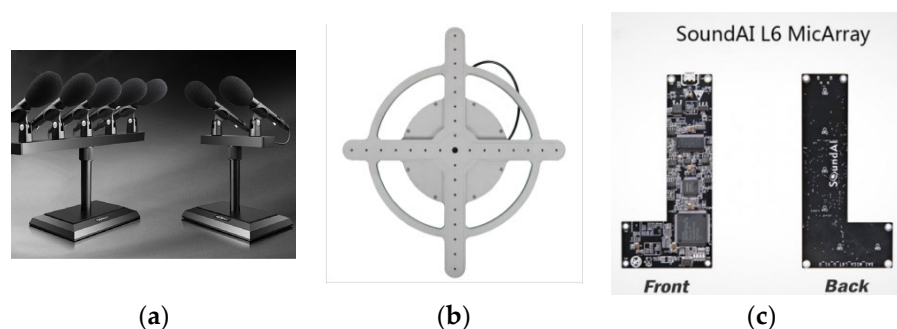


Figure 1. Microphone array formations: (a) line array, (b) cross-shaped array, (c) L-array.

Line array: Multiple microphones arranged in a straight line. This array is suitable for scenarios wherein the sound source is moving horizontally.

Cross-shaped array: An array that positions multiple microphones in a cross-shaped manner, which is suitable for sound source localization and speech recognition scenarios.

Rectangular array: Multiple microphones arranged in a rectangle. This is suitable for scenarios wherein sound sources move horizontally and vertically.

U-shaped array: The microphones are arranged in a special U-shape, suitable for high-quality language recognition. For example, Translator 4.0 developed by iFlytek was paired with a U-shaped microphone array, which helped improve the accuracy and quality of sound capture.

L-Array: Multiple microphones arranged in an L-shape. This array is suitable for scenarios wherein sound needs to be captured within a specific area.

The circular microphone array positions multiple microphones in a circle and has the following benefits: (1) 360-degree omnidirectionality. The circular microphone array can achieve 360-degree omnidirectionality in the horizontal direction and can evenly receive sound from all directions without mechanical rotation or adjusting the orientation of the microphone. This is useful for applications that need to capture the sounds from the whole environment, such as speech recognition, voice conferencing, and voice commands. (2) Eliminating noise and echo. The circular microphone array can use the delay difference and amplitude difference between multiple microphones in the array to achieve spatial sound source localization and sound source separation, so that the influence of noise and echo can be effectively eliminated. This is very helpful for collecting clear speech signals in noisy environments, thus improving the performance of speech recognition and speech processing [17,18].

The diagonal double rectangular array proposed in this article consists of two rectangles. The distance between the adjacent microphones is equal, and the centers of the two rectangles are staggered from each other, forming a diagonal line. The advantages of diagonal double rectangular arrays include: (1) The improved accuracy of sound source positioning. Since the microphones in a diagonal double rectangular array are staggered, more sound direction information can be captured, thereby improving the accuracy of sound source positioning. (2) Reduction in echo and noise interference. The diagonal double rectangular array can use the signals of multiple microphones for noise reduction processing, thereby reducing the interference of echo and noise. This is useful for speech recognition and sound source localization in some noisy environments. (3) Supporting spatial filtering. The diagonal double rectangular array can use the signals of multiple microphones for spatial filtering processing, thereby improving the quality of voice signals. (4) Supporting multi-sound source separation. The diagonal double rectangular array can realize the separation of multiple sound sources by analyzing the time domain and frequency domain characteristics of multiple microphone signals. This is of great benefit for idlers working in environments with complex sounds. In general, the diagonal double rectangular array has unique advantages in improving the positioning accuracy of sound sources, reducing echo and noise interference, supporting spatial filtering, and supporting multi-source separation. Hence, the diagonal double rectangular microphone array was selected as the acquisition array for idler fault distance signals.

2.2. MVDR

Beamforming technology and signal space wavenumber spectrum estimation are the two main research directions of free space signal array processing. In order to maximize the array gain, Capon proposed the MVDR beamformer, also known as the Capon beamformer, in 1969. MVDR is an adaptive beamforming algorithm based on the maximum signal-to-noise ratio (SNR) criterion. The MVDR algorithm can adaptively minimize the power in the desired direction of the array output and maximize the signal-to-drying ratio. It can greatly improve the resolution and suppress noise in spatial spectral estimation [19–21]. This study implemented traditional MVDR beamforming in the frequency domain based on the Acoular library. In the Acoular library, time-domain signals are first processed through the PowerSpectra class, which performs framing and windowing on the time-domain signals. Then, they are converted into frequency-domain signals through fast Fourier transform. Next, the cross spectra between signals at each frequency are calculated, and these cross

spectral values are stored in a matrix. Subsequently, the SteeringVector class is used to calculate the steering vector. The steering vector contains the phase information of the microphone array relative to the signal source. This phase information is used to weight and add the signals of the microphone array to form a sound beam pointing in a specific direction. Afterwards, the cross spectral matrix and steering vector are passed to the BeamformerCapon class. Then, the signal frequency band is divided, and the envelope spectral slope value is used as an indicator to search for the optimal sub-band and select the center frequency. By setting the center frequency f_0 according to the frequency band division method of 1/3 octave bandwidth, the upper frequency of the 1/3 octave band $f_h = f_0 \times \sqrt[3]{2}$, and the lower limit frequency $f_l = f_0 \div \sqrt[3]{2}$. According to the specified center frequency and frequency band index, a synthetic beam is synthesized from the processed microphone channel data to form a sound signal, the feature matrix of the target sound signal is obtained, and the feature matrix is processed as the input layer of the LSTM model.

The specific operation steps of MVDR are as follows. Suppose the desired target location P' in space is (r_0, θ_0) , the jamming signals are $i_j (j = 1, 2, 3, 4 \dots J)$, the interference location is (r_j, θ_j) , the array element noise is n_t , and the n th element signal at the receiving end is

$$x(t) = a(\theta_0)_s(t) + \sum_{j=1}^J a(\theta_j) i_j(t) + n(t) \quad (1)$$

In this formula, the direction of the receiving guide vector $a(\theta)$ is based on (r, θ) , which can be expressed as

$$a(\theta) = \left[1, e^{-j2\pi \frac{f_0 d \sin(\theta)}{c}}, \dots, e^{-j(N-1)2\pi \frac{f_0 d \sin(\theta)}{c}} \right] \quad (2)$$

When the constraints are also satisfied, in order to minimize noise, the result of the objective function optimization is

$$\min_{\omega} \left[|y(t)|^2 \right] = \min_{\omega} (w^H R w) \quad (3)$$

Then, the MVDR weight optimization problem can be expressed as

$$\min_{\omega} (w^H R w) \quad (4)$$

$$\text{s.t.} \begin{cases} w^H a(\theta_d) = 1 \\ w^H a(\theta_{ij}) = 0 \end{cases} \quad (5)$$

The essence of the MVDR beamformer is to solve the weight coefficients of each array. Then, using the Lagrange multiplier yields

$$L(w) = w^H R w + \lambda [w^H a(\theta_d) - 1] \quad (6)$$

After deriving Equation (6), let it be 0, which can be obtained as

$$\frac{\partial L(w)}{\partial w} = 2Rw + \lambda a(\theta_d) = 0 \quad (7)$$

Then, it is solved as follows:

$$w = \mu R^{-1} a(\theta_d) \quad (8)$$

According to the MVDR criterion, the optimal value of the array weight is obtained:

$$w_{\text{MVDR}} = \frac{R^{-1}a(\theta_d)}{a^H(\theta_d)R^{-1}a(\theta_d)} \quad (9)$$

2.3. LSTM

Recurrent neural networks (RNNs) are robust deep neural networks with high performance that process serial data of varying lengths for end-to-end distribution [22]. Long short-term memory (LSTM) is an advanced RNN architecture that includes the memory unit proposed by Hocklet and Schmiduber to solve the gradient vanishing problem in RNNs. As a time-loop network, it uses the “three-gate structure” to obtain the correlation of large-scale time series data and extract the optimal features [23–25]. The forgetting gate of the LSTM architecture avoids long-term dependencies and trains valuable information from historical network units to obtain more meaningful and autonomous information in sequential data. The data for the input sequence are represented as $X = [x_1, x_2, x_3 \cdots x_m]$. Through each time step t , $x_t \in R_s$ is the input vector, and the hidden layer status is $h_t \in R^l$, providing the following formula:

$$h_t = f(h_{t-1}, x_t) = f(Uh_{t-1} + Wx_{t-1} + b) \quad (10)$$

In this formula, $U \in R^{l \times l}$, $W \in R^{l \times s}$, and $b \in R^l$ are blended learning parameters; f is a nonlinear activation function; m is the sequence length; s is the size of the input; and l is the hidden size.

A diagram of the LSTM unit is shown in Figure 2a with hidden state variables h_t . LSTM retains the encoding time phase t of the storage unit c_t . The efficiency of the storage unit is determined by three gates: the input gate i_t , output gate o_t , and forgetting gate f_t . The upgraded equation can be expressed as:

$$i_t = \text{sigm}(U_i h_{t-1} + W_i x_t + b_i) \quad (11)$$

$$f_t = \text{sigm}(U_f h_{t-1} + W_f x_t + b_f) \quad (12)$$

$$o_t = \text{sigm}(U_o h_{t-1} + W_o x_t + b_o) \quad (13)$$

$$\tilde{c}_t = \tanh(U_c h_{t-1} + W_c x_t + b_c) \quad (14)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (15)$$

$$h_t = o_t \odot \tanh(c_t) \quad (16)$$

where $U \in R^{l \times l}$, $W \in R^{l \times s}$, and $b \in R^l$ are the learning optimization parameters; sigm is an S-type function; \tanh is the hyperbolic cut function; and operator \odot represents the element-level product.

The main operating points of the LSTM unit are defined according to each phase of the time t and comprise the following three steps: (1) From the new input functions x_t and early hidden state h_{t-1} , one can obtain the forgetting gate f_t , whose values range from 0 to 1. When the value of f_t reaches 1, the last memory cell c_{t-1} of incoming data is strongly preserved. On the other hand, when the value of f_t reaches 0, the incoming data are distributed and cannot be retained. (2) Input gates can be removed from the new input functions x_t and early hidden state h_t and added to the memory cell, specifying c_t . (3) Therefore, the output gate must select an item from the memory cell that can be processed to form a new hidden state h_t [26–29].

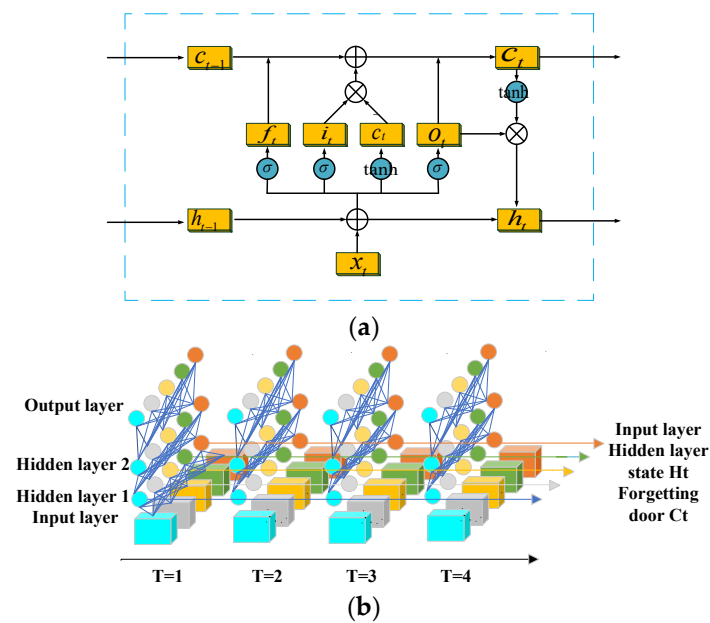


Figure 2. LSTM model: (a) LSTM cell model, (b) schematic diagram of the overall LSTM model.

The overall LSTM model is shown in Figure 2b. The LSTM model has two hidden layers, an input layer, and an output layer. When viewed separately, it is a simple BP neural network, but as time passes, the hidden layer information h , c also begins to be processed, forming a complete LSTM model.

3. Simulation

3.1. MVDR Algorithm Simulation

The MVDR algorithm was simulated by MATLAB, and the parameters set are shown in Table 1. The desired signal angle was set to 10, and two interfering signals were added at angles of -30 and 30 . The simulation analysis showed that the waveform signal with an angle of 10 was convex, and the angle in the -30 and 30 directions was effectively suppressed. It can be seen that the effect of MVDR in the direction of noise suppression was obvious. See Figure 3.

Table 1. Simulation parameters.

Parameter Name	Parameter Value
Number of arrays	18
Desired signal angle	10
Interference signal angle	$-30, 30$
SNR	10
INR	10
Number of stories	100

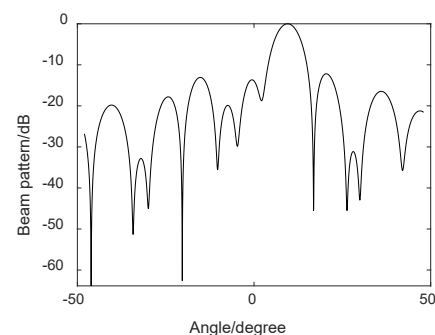


Figure 3. MVDR simulation.

3.2. Microphone Array Formation Selection

The simulation results of the microphone array are shown in Figure 4, and the sound source position of the diagonal double rectangular array was closest to the true value and the suppression effect on the side lobes. The main lobe of the uniform straight array was narrower but it covered a greater area, and the results were less accurate. Although the position of the main sound source could be accurately located, the interference from other sound sources appeared, and the suppression effect on the side lobe was not ideal, producing an error. The cross-shaped array was relatively uniform, and the main lobe of the linear array was more prominent, but there was still a problem of inaccurate positioning. Due to the existence of too many main lobes, the rectangular array had the most obvious inhibition effect on the side lobes, but the localized main lobe area was too large to play the role of precise positioning. Therefore, the proposed double rectangular diagonal array had the best robustness and was most suitable as a microphone array for an idler failure distance estimation model.

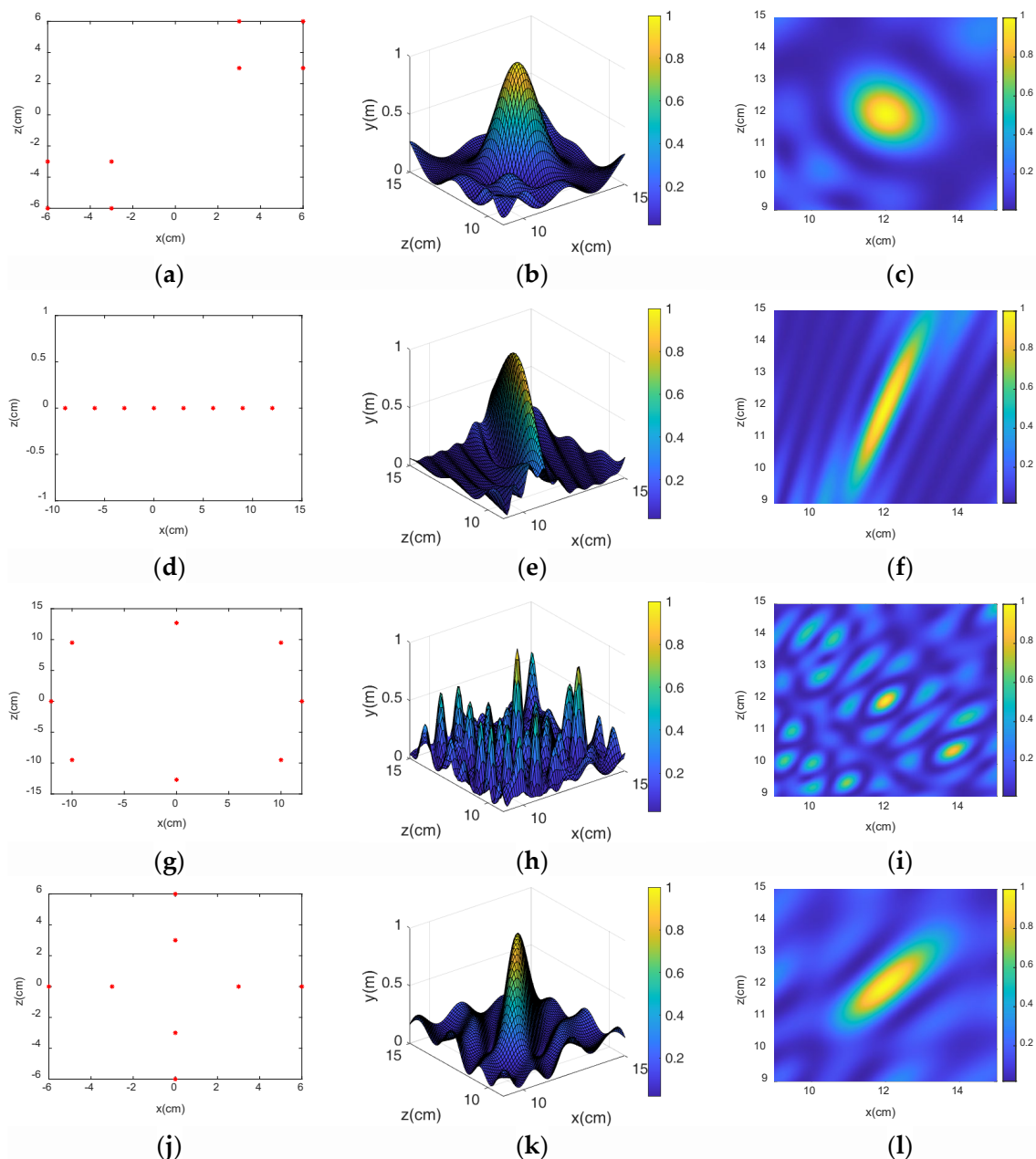


Figure 4. Cont.

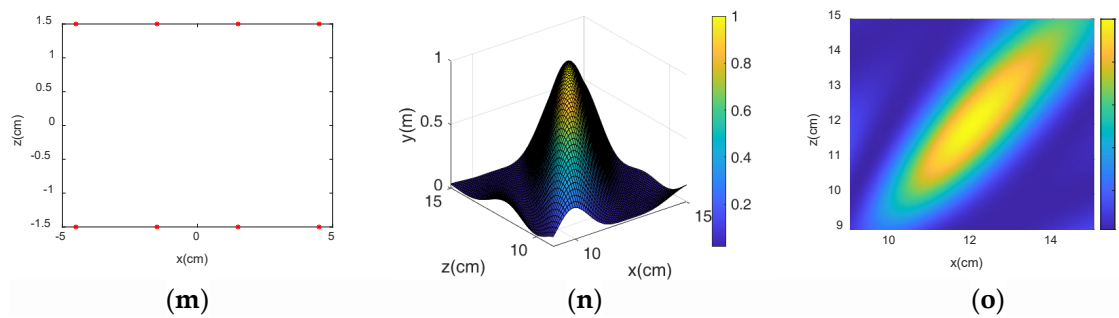


Figure 4. Simulation results of each array: (a–c) schematic diagram of diagonal double rectangular array and two- and three-dimensional positioning diagrams, (d–f) linear array diagram and two- and three-dimensional positioning diagrams, (g–i) circular array schematic diagram and two- and three-dimensional positioning diagrams, (j–l) cross array schematic diagram and two- and three-dimensional positioning diagrams, (m–o) rectangular array schematic diagram and two- and three-dimensional positioning diagrams.

4. Experiments

4.1. LSTM Model Parameter Settings

The LSTM model consisted of one input layer, two hidden layers, one fully connected layer, and one output layer. When the hidden layer was set to one layer, the effect achieved was not ideal. After increasing it to two layers, the demand was already met. When increasing it to three layers, the calculation time of the model increased. Therefore, two hidden layers were used. The accuracy of the model with different numbers of hidden layers is shown in Figure 5. We used the cross-entropy loss function as the objective function to guide the learning of network parameters and selected Softmax as the activation function. Softmax is an activation function specifically designed for multi-class problems. It can convert a set of real numbers into an output and a probability distribution of 1, which is very suitable for the output layer probability prediction of multiple-classification problems. As the optimization method, we selected Adam. Adam optimizers are some of the most popular optimizers in deep learning. They are suitable for a wide variety of problems, including models with sparse or noisy gradients. Their ability to be easily fine-tuned makes it possible to obtain good results quickly; in fact, the default parameter configuration usually achieves good results. Adam optimizers combine the best of AdaGrad and RMSProp. Adam uses the same learning rate for each parameter and adapts independently as the learning progresses.

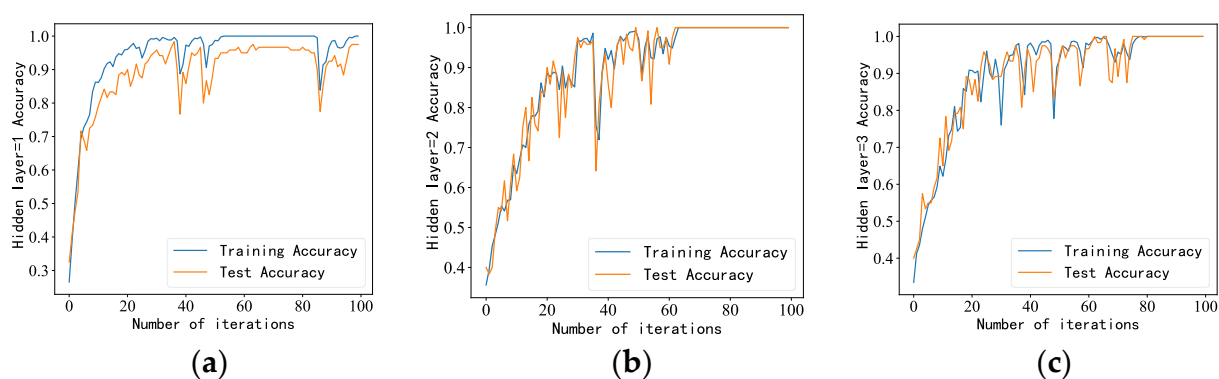


Figure 5. Model accuracy under different numbers of hidden layers: (a) Accuracy when the number of hidden layers is 1, (b) Accuracy when the number of hidden layers is 2, (c) Accuracy when the number of hidden layers is 3.

4.2. Experimental Setup

The diagonal double rectangular microphone array was used as a sound signal acquisition device. The sampling frequency was 16,000 Hz, the number of sampling points was 1024, and the device consisted of 8 MEMS microphones and microphone holders. In the experiment, five positions were selected as sound source points. The distance between each sound source points was 0.5 m, and the microphone array was 1 m away from the ground and 1 m away from the central sound source, as shown in the schematic diagram in Figure 6.

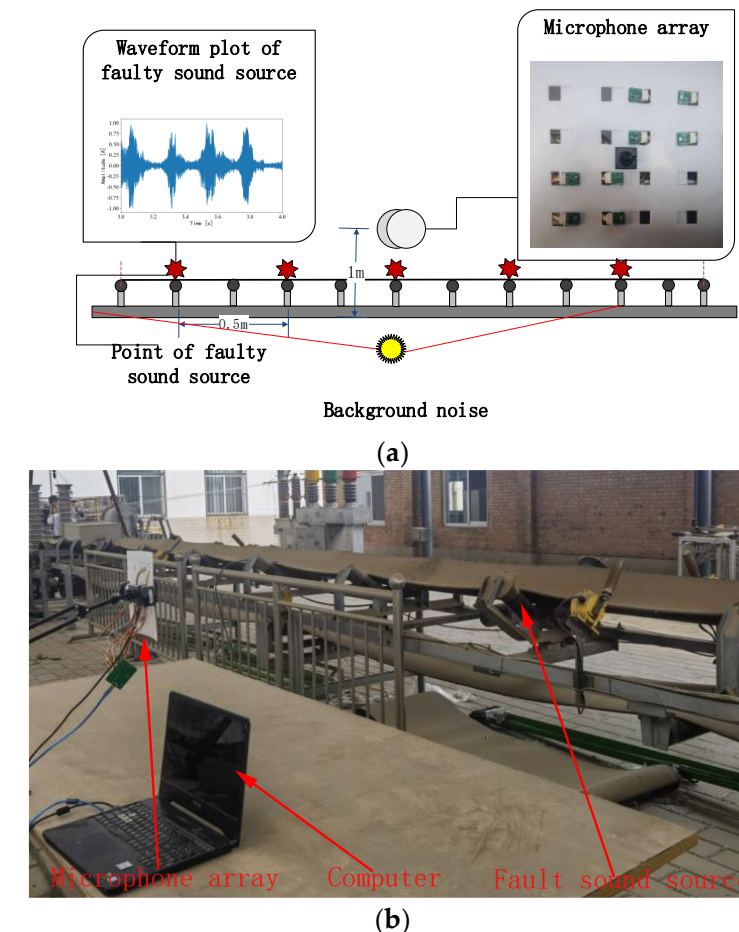


Figure 6. Schematic diagram of the experimental layout: (a) experimental schematic diagram, (b) field experiment diagram.

4.3. Experimental Process

The flow chart of the experiment is shown in Figure 7, and the specific experimental flow is as follows:

- Step 1—The microphone array was used to collect the idler fault audio signal.
- Step 2—MVDR processing was performed on the collected signal to generate idler fault distance data.
- Step 3—We processed the data again to build the dataset.
- Step 4—We divided the data into the training set (90%) and testing set (10%).
- Step 5—We trained and tested the model to draw conclusions.

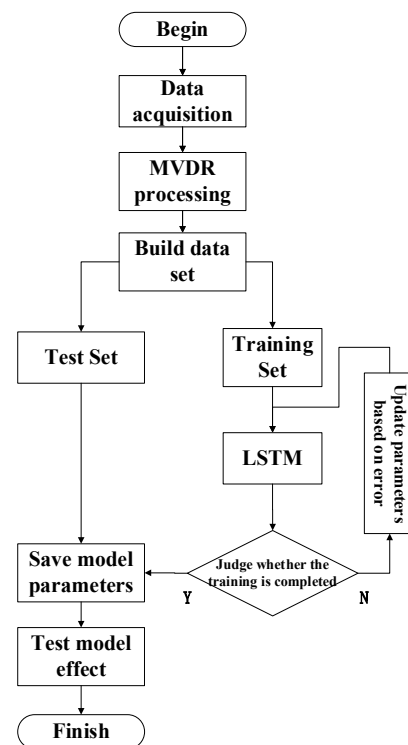


Figure 7. Experimental flowchart.

4.4. Background Noise Interference Experiment

White noise was used as background noise to collect the sound signal of the idler, and the waveform diagram obtained is shown in Figure 8. The fault characteristic information of the idler was masked by the white noise, so as to imitate the environment during on-site inspection.

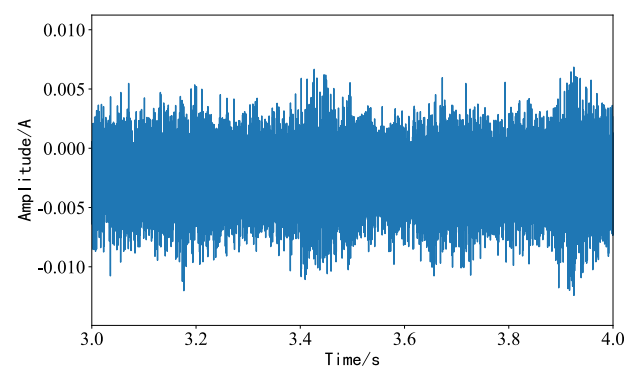


Figure 8. Roller acquisition signal waveform under white noise interference.

After beamforming the acquired sound signal, the processed data were passed to the LSTM model. The data samples were processed 100 times, and the number of data (samples) passed to the program for training at a time (batch_size) was 16. The accuracy and loss rates of the training set and the test set were observed as the model performance, and the accuracy of the training and test sets of the MVDR-LSTM model gradually converged after 65 iterations, with the accuracy remaining at 100% for both. The loss rate remained steady after 65 iterations as the number of iterations increased, and the results are shown in Figure 9.

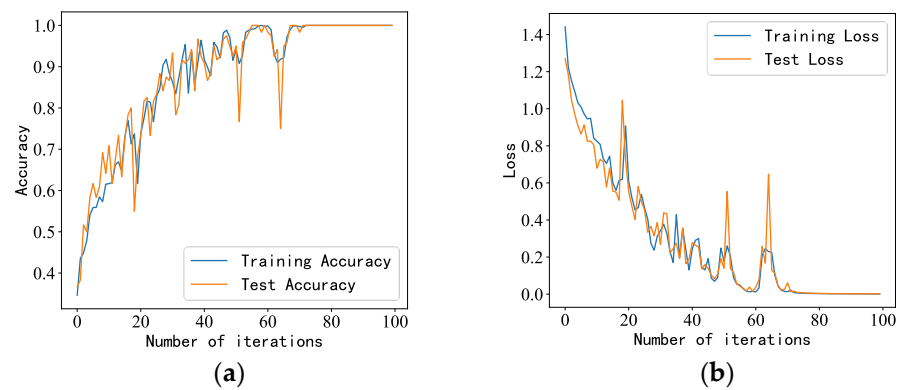


Figure 9. LSTM model performance under background noise interference: (a,b) accuracy and loss rate of training and test sets, respectively.

4.5. Impact Noise Interference Experiment

Under the same experimental conditions, the background noise was replaced with irregular impact interference noise. The obtained waveform is shown in Figure 10, and the impact noise covered part of the fault characteristics, showing a waveform with higher amplitude, which increased the difficulty of estimating the fault distance of the idler.

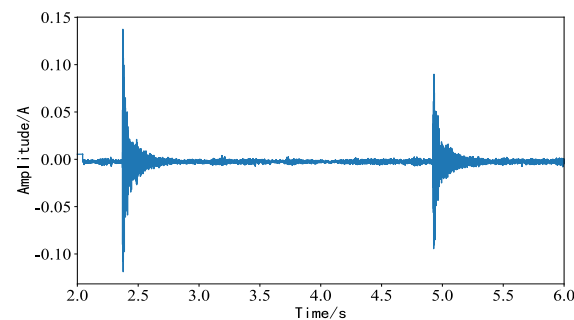


Figure 10. Waveform diagram of idler acquisition signal under impact noise interference.

Similarly, 100 iterations were performed on the newly generated data samples, and the number of data (samples) passed to the program for training (batch_size) was 16. The accuracy and loss rates of the training and test sets are shown in Figure 11. The accuracy of the training and test sets of the MVDR-LSTM model converged after 80 iterations, and the loss rate converged after 85 iterations.

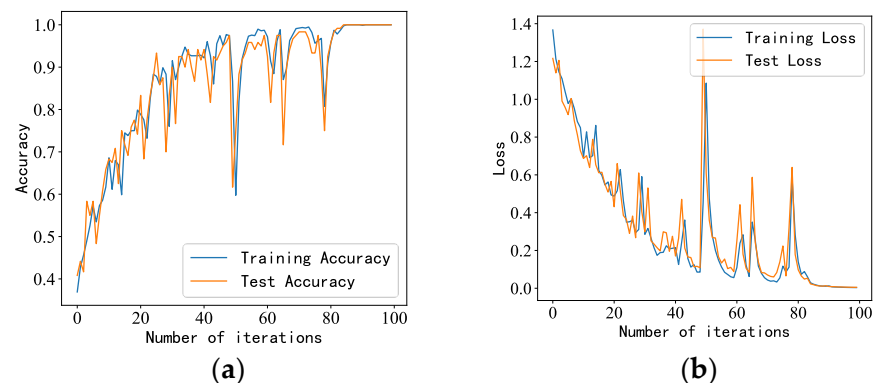


Figure 11. LSTM model performance under impact noise interference: (a,b) accuracy and loss rate of training and test sets, respectively.

4.6. Evaluation Criteria and Comparison of Models

4.6.1. Data Comparison for Diagonal Double Rectangular Microphone Array

In order to verify the superiority of the proposed model, the diagonal double rectangular array was used as the acquisition device, and the MVDR-LSTM, CBF-LSTM, and FBF-LSTM models were used as the comparison models. We set the number of data (samples) passed to the program for training at a time (batch_size) as 16, the activation function was Softmax, and the optimizer was Adam. Under background noise interference, the accuracy of the three models, the confusion matrix classification results, and KPCA dimensionality reduction analysis were used as evaluation criteria. The accuracy results obtained over 10 experiments are shown in Figure 12. The accuracy of the MVDR-LSTM model was higher and more stable than that of the CBF-LSTM and FBF-LSTM models, with average accuracy values of 99.74%, 94.60%, and 74.69%, respectively. The highest values were 100%, 100%, and 86.78%, and the lowest values were 98.32%, 89.52%, and 60.33%, respectively. Based on the data of the 10 experiments, we found that the average accuracy of the MVDR-LSTM model was higher than that of the other two models.

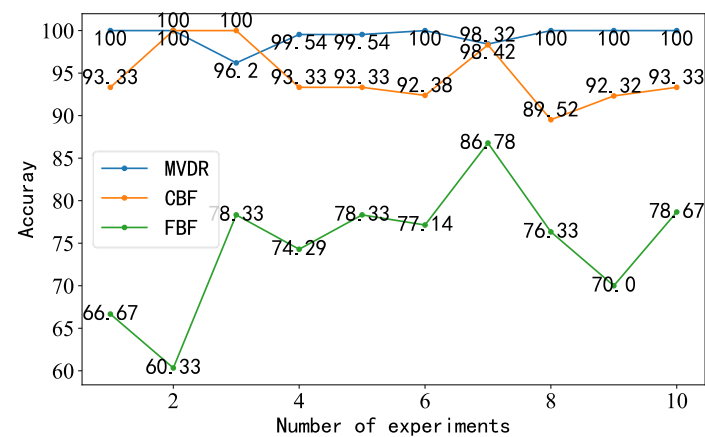


Figure 12. Line chart illustrating the accuracy of the three models for the diagonal double rectangular array.

We used a confusion matrix to compare the classification results with the actual measured values and displayed the accuracy of the classification results in the same confusion matrix. As shown in Figure 13, we found that the MVDR-LSTM model achieved perfect prediction results in all categories without misclassification. The CBF-LSTM model had some classification errors, while FBF-LSTM demonstrated more misclassification, so the MVDR-LSTM model was more effective.

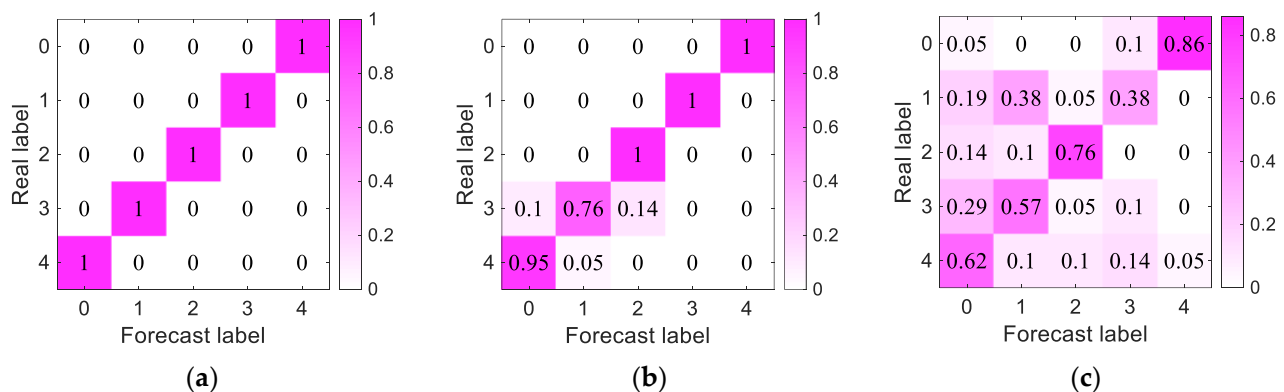


Figure 13. Diagonal double rectangular array per-model confusion matrix results: (a–c) MVDR-LSTM, CBF-MVDR, and FBF-LSTM confusion matrixes, respectively.

The KPCA dimensionality reduction diagram of the idler fault distance data sample is shown in Figure 14. The feature expression distribution learned by the MVDR-LSTM model showed the clearest boundary, so the extracted features were more easily segmented, which meant that it was easier to classify each fault distance category. There was much evidence of feature misclassification by the CBF-LSTM and FBF-LSTM models, and the effect was not good.

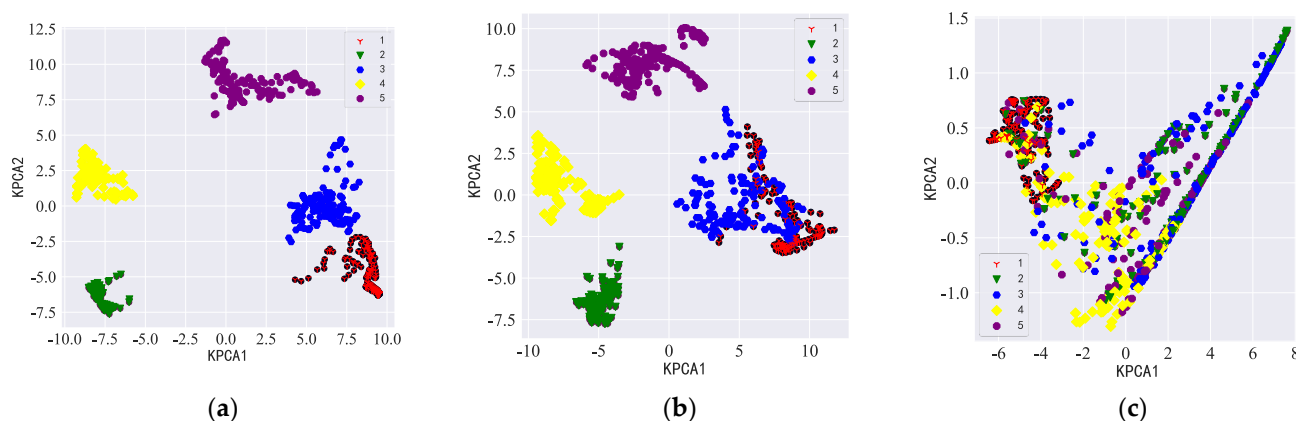


Figure 14. KPCA dimensionality reduction diagram for each model with the diagonal double rectangular array: (a–c) MVDR-LSTM, CBF-MVDR, and FBF-LSTM KPCA dimensionality reduction analyses, respectively.

4.6.2. Data Comparison for Circular Microphone Array

The MVDR-LSTM, CBF-LSTM, and FBF-LSTM models in the circular array were selected as comparison models. We set the number of data (samples) passed to the program for training at a time (batch_size) as 16, the activation function was Softmax, and the optimizer was Adam. Under background noise interference, the accuracy of the three models, the confusion matrix classification results, and KPCA dimensionality reduction analysis were used as evaluation criteria. The accuracy results obtained over 10 experiments are shown in Figure 15. The accuracy of the MVDR-LSTM model was higher and more stable than that of the CBF-LSTM and FBF-LSTM models, with average accuracy values of 89.72%, 85.33%, and 74.53%, respectively. The highest values were 93.33%, 93.33%, and 82.32%, and the lowest values were 86.78%, 54.28%, and 60.00%, respectively. Based on the data of 10 experiments, we found that the average accuracy of the MVDR-LSTM model was higher than that of the other two models.

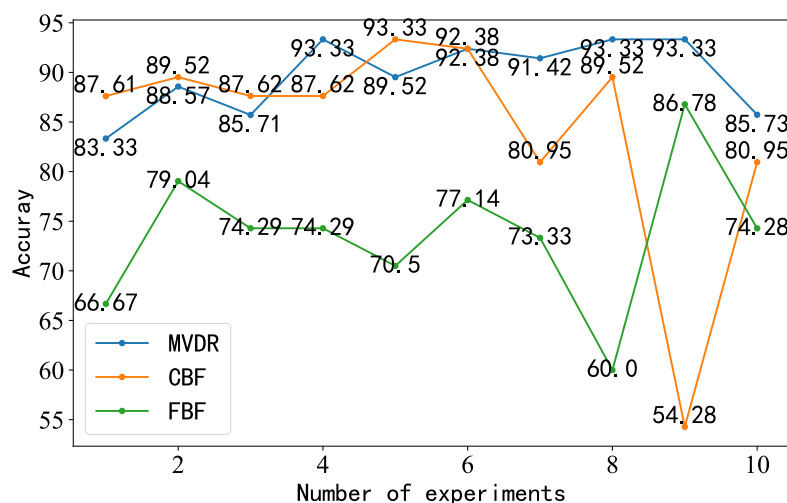


Figure 15. Line chart illustrating the accuracy of the three models for the circular array.

The confusion matrix of the roller failure distance data sample of the circular array is shown in Figure 16. We found that the sample classification results of the MVDR-LSTM model included less misclassification, and there were further advantages over the other two models.

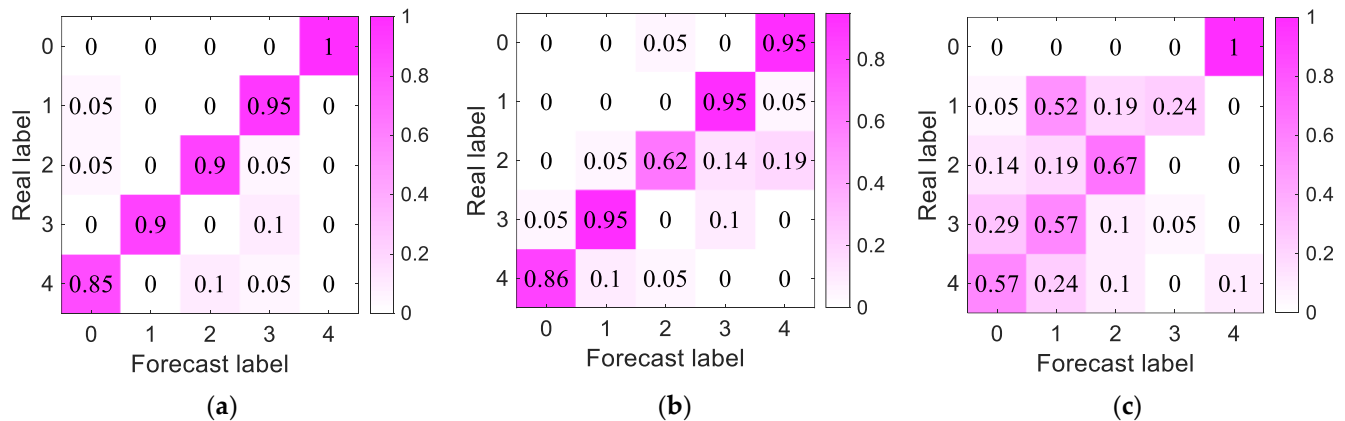


Figure 16. Circular array per-model confusion matrix results: (a–c) MVDR-LSTM, CBF-MVDR, and FBF-LSTM confusion matrixes, respectively.

The KPCA dimensionality reduction diagram of the idler fault distance data sample obtained by the circular array acquisition signal is shown in Figure 17. Although the boundary shown by the feature distribution expression of the three types of models was not obvious, the classification effect of MVDR-LSTM was the best, and it demonstrated less data misclassified compared with the other two models.

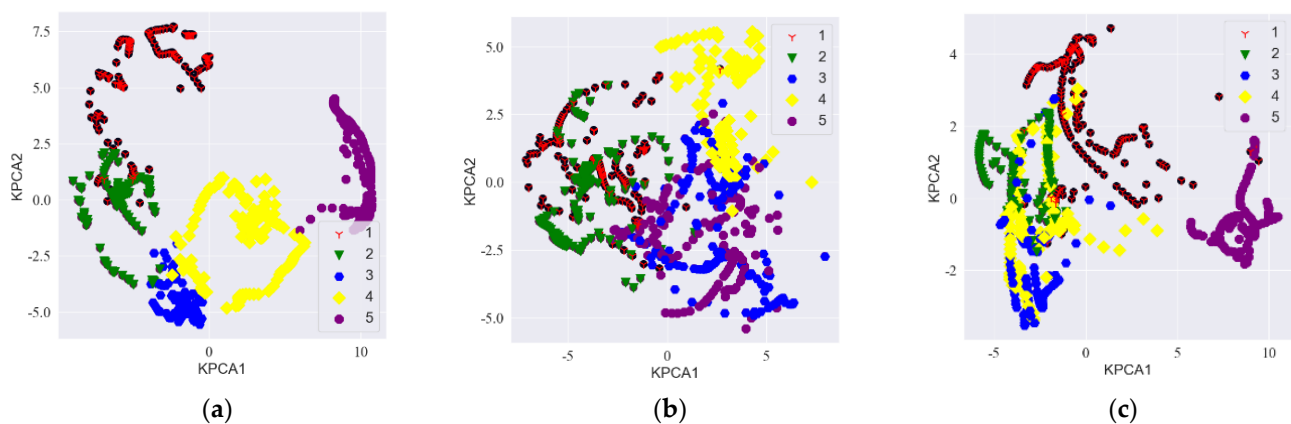


Figure 17. KPCA dimensionality reduction diagram for each model with the circular array: (a–c) MVDR-LSTM, CBF-MVDR, and FBF-LSTM KPCA dimensionality reduction analyses, respectively.

Through the comparative analysis of the above methods, we observed that the effect of acquiring signals using a diagonal double rectangular array in a noisy interference environment was better than when using a circular array, and the MVDR-LSTM model presented substantial advantages in all aspects compared with the other five models, so the MVDR-LSTM model was selected as the fault distance estimation model.

5. Conclusions

In this paper, an MVDR-LSTM model for idler fault distance estimation was proposed. The model was tested in two different interference environments, and the results were compared with five other models to reach the following conclusions:

1. Due to the special structure of an idler, the human diagnosis steps are cumbersome and time is wasted. Based on deep learning, our model removed the tedious step of

manually extracting idler fault features, which could greatly shorten the time required and improve the efficiency of idler fault diagnosis.

2. Through simulation experiments, we found that the diagonal dual rectangular microphone array had higher resolution and noise resistance compared to the other microphone arrays and could be beneficial to the estimation of roller fault distance.
3. After adjusting the model structure and parameters, we trained the generated idler fault distance samples. The accuracy of the proposed model was 100%, with better performance than the other models and results closer to the true values.
4. Five fault locations during idler operation were analyzed, and the experimental results showed that the proposed model had the ability to estimate the fault distance and provide better robustness. It could be used as a standard for judging the fault distance of idlers and provide ideas for combining beamforming algorithms and deep learning with idler fault diagnosis.

The main advantage of LSTM is its ability to learn and remember long-term sequence dependencies. However, compared to CNNs, LSTM models have higher computational complexity because they contain gating mechanisms and time-step dependencies, resulting in greater computational costs in the training and inference process, which may require a longer training time and higher computing resources. Thus, it is necessary to continue to optimize the model in future research to reduce the computational costs and shorten the computing time.

Author Contributions: Conceptualization, F.D. and J.L.; methodology, S.W.; software, W.W.; validation, S.W., W.W. and X.Z.; formal analysis, S.W.; investigation, F.D. and J.L.; resources, X.Z.; data curation, W.W.; writing—original draft preparation, S.W.; writing—review and editing, X.Z. and W.W.; visualization, S.W.; supervision, F.D. and W.W.; project administration, X.Z.; funding acquisition, X.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (52105098), the Natural Science Foundation of Hebei Province (E2021502038), and the Fundamental Research Funds for the Central Universities (2023MS130).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Zhang, L.; Zhang, F.; Qin, Z.; Han, Q.; Wang, T.; Chu, F. Piezoelectric energy harvester for rolling bearings with capability of self-powered condition monitoring. *Energy* **2022**, *238*, 121770. [\[CrossRef\]](#)
2. Zhang, Y.; Cao, J.; Zhu, H.; Lei, Y. Design, modeling and experimental verification of circular Halbach electromagnetic energy harvesting from bearing motion. *Energy Convers. Manag.* **2019**, *180*, 811–821. [\[CrossRef\]](#)
3. Zhou, Y.; Kumar, A.; Gandhi, C.P.; Vashishtha, G.; Tang, H.; Kundu, P.; Singh, M.; Xiang, J.W. Discrete entropy-based health indicator and LSTM for the forecasting of bearing health. *J. Braz. Soc. Mech. Sci. Eng.* **2023**, *45*, 12. [\[CrossRef\]](#)
4. Aljemely, A.H.; Xuan, J.; Al-Azzawi, O.; Jawad, F.K. Intelligent fault diagnosis of rolling bearings based on LSTM with large margin nearest neighbor algorithm. *Neural Comput. Appl.* **2022**, *34*, 19401–19421. [\[CrossRef\]](#)
5. Xie, W.; Li, Z.; Xu, Y.; Gardoni, P.; Li, W. Evaluation of Different Bearing Fault Classifiers in Utilizing CNN Feature Extraction Ability. *Sensors* **2022**, *22*, 3314. [\[CrossRef\]](#)
6. Yang, Q.; Zhang, J.; Chen, L.; Wu, D.S. Fault diagnosis of motor bearing based on improved convolution neural network based on VMD. In Proceedings of the 31st Chinese Control and Decision Conference (CCDC), Nanchang, China, 3–5 June 2019; pp. 405–409.
7. Yu, J.; Wen, Y.; Yang, L.; Zhao, Z.; Guo, Y.; Guo, X. Monitoring on triboelectric nanogenerator and deep learning method. *Nano Energy* **2021**, *92*, 106698. [\[CrossRef\]](#)
8. Chen, L.; Choy, Y.S.; Wang, T.G.; Chiang, Y.K. Fault detection of wheel in wheel/rail system using kurtosis beamforming method. *Struct. Health Monit.* **2020**, *19*, 495–509. [\[CrossRef\]](#)
9. Cabada, E.C.; Leclerc, Q.; Antoni, J.; Hamzaoui, N. Fault detection in rotating machines with beamforming: Spatial visualization of diagnosis features. *Mech. Syst. Signal Process.* **2017**, *97*, 33–43. [\[CrossRef\]](#)

10. Sun, S.L.; Wang, T.Y.; Yang, H.X.; Chu, F.L. Damage identification of wind turbine blades using an adaptive method for compressive beamforming based on the generalized minimax-concave penalty function. *Renew. Energy* **2022**, *181*, 59–70. [\[CrossRef\]](#)
11. He, T.; Xiao, D.H.; Pan, Q.; Liu, X.D.; Shan, Y.C. Analysis on accuracy improvement of rotor–stator rubbing localization based on acoustic emission beamforming method. *Ultrasonics* **2014**, *54*, 318–329. [\[CrossRef\]](#)
12. Subramanian, A.S.; Weng, C.; Watanabe, S.; Yu, M.; Yu, D. Deep learning based multi-source localization with source splitting and its effectiveness in multi-talker speech recognition. *Comput. Speech Lang.* **2022**, *75*, 101360. [\[CrossRef\]](#)
13. Zhang, X.L. Deep ad-hoc beamforming. *Comput. Speech Lang.* **2021**, *68*, 101201. [\[CrossRef\]](#)
14. Yang, Z.; Guan, S.; Zhang, X.L. Deep ad-hoc beamforming based on speaker extraction for target-dependent speech separation. *Speech Commun.* **2022**, *140*, 87–97. [\[CrossRef\]](#)
15. Ramezanpour, P.; Mosavi, M.-R. Two-Stage Beamforming for Rejecting Interferences Using Deep Neural Networks. *IEEE Syst. J.* **2021**, *15*, 4439–4447. [\[CrossRef\]](#)
16. Tao, T.; Zheng, H.; Yang, J.; Guo, Z.; Zhang, Y.; Ao, J.; Chen, Y.; Lin, W.; Tan, X. Sound Localization and Speech Enhancement Algorithm Based on Dual-Microphone. *Sensors* **2022**, *22*, 715. [\[CrossRef\]](#)
17. Ahamed, P.S.S.; Duraiswamy, P. Virtual Sensing Active Noise Control System with 2D Microphone Array for Automotive Applications. In Proceedings of the International Conference on Signal Processing and Integrated Networks, Noida, India, 7–8 March 2019; pp. 151–155.
18. Wakabayashi, Y.; Yamaoka, K.; Ono, N. Rotation-Robust Beamforming Based on Sound Field Interpolation with Regularly Circular Microphone Array. In Proceedings of the ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 6–11 June 2021; pp. 771–775.
19. Kidav, J.U.; Sivamangai, N.M.; Pillai, M.P.; Sreejeesh, S.G. A broadband MVDR beamforming core for ultrasound imaging. *Integration* **2021**, *81*, 221–233. [\[CrossRef\]](#)
20. Huang, Q.H.; Hu, R.; Fang, Y. Real-valued MVDR beamforming using spherical arrays with frequency invariant characteristic. *Digit. Signal Process.* **2016**, *48*, 239–245. [\[CrossRef\]](#)
21. Li, J.; White, P.R.; Bull, J.M.; Leighton, T.G.; Roche, B.; Davis, J.W. Passive acoustic localisation of undersea gas seeps using beamforming. *Int. J. Greenh. Gas Control* **2021**, *108*, 103316. [\[CrossRef\]](#)
22. Ngoc, H.V.; Mayer JR, R.; Bitar-Nehme, E. Deep learning LSTM for predicting thermally induced geometric errors using rotary axes’powers as input parameters. *CIRP J. Manuf. Sci. Technol.* **2022**, *37*, 70–80. [\[CrossRef\]](#)
23. Nemani, V.P.; Lu, H.; Thelen, A.; Hu, C.; Zimmerman, A.T. Ensembles of probabilistic LSTM predictors and correctors for bearing prognostics using industrial standards. *Neurocomputing* **2022**, *491*, 575–596. [\[CrossRef\]](#)
24. Liu, J.; Pan, C.; Lei, F.; Hu, D.; Zuo, H. Fault prediction of bearings based on LSTM and statistical process analysis. *Reliab. Eng. Syst. Saf.* **2021**, *214*, 107646. [\[CrossRef\]](#)
25. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.* **2014**, *61*, 85–117. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Wang, Q.; Yu, Y.; Ahmed, H.O.A.; Darwish, M.; Nandi, A.K. Open-Circuit Fault Detection and Classification of Modular Multilevel Converters in High Voltage Direct Current Systems (MMC-HVDC) with Long Short-Term Memory (LSTM) Method. *Sensors* **2021**, *21*, 4159. [\[CrossRef\]](#)
27. Yin, A.; Yan, Y.; Zhang, Z.; Li, C.; Sánchez, R.-V. Fault Diagnosis of Wind Turbine Gearbox Based on the Optimized LSTM Neural Network with Cosine Loss. *Sensors* **2020**, *20*, 2339. [\[CrossRef\]](#)
28. Zheng, J.; Liao, J.; Chen, Z. End-to-End Continuous/Discontinuous Feature Fusion Method with Attention for Rolling Bearing Fault Diagnosis. *Sensors* **2022**, *22*, 6489. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.