Joint Time Delay and Frequency Estimation Based on Deep Learning

Mahmoud M. Qasaymeh * and Ahmad Falah Aljaafreh
Tafila Technical University, Tafila, Jordan
Email: qasaymeh@ttu.edu.jo (M.M.Q.); a.aljaafreh@ttu.edu.jo (A.F.A.)
*Corresponding author

***Abstract*—This article introduces to the best of our knowledge a novel approach for simultaneous estimation of time delay and frequencies in noisy complex sinusoidal signals received at two spatially separated sensors. The proposed method comprises two main components. Firstly, a Convolutional Neural Network (CNN) regression model is employed to estimate frequencies using data from the first sensor. The model is trained on a synthetic dataset specifically designed for this task. Secondly, a deep learning model is developed, incorporating densely connected layers and dropout layers for regularization, to effectively estimate the time delay between the received signal copies at the two sensors. Extensive computer simulations demonstrate the effectiveness of the proposed method, showcasing its accuracy in joint time delay and frequency estimation. This deep learning-based technique offers a promising alternative to classical signal processing approaches, enabling advanced signal analysis in diverse engineering domains.**

***Keywords*—delay and frequency estimation, deep learning, convolutional neural networks, temporal patterns**

I. Introduction

The problem of the Time Delay Estimation (TDE) between noisy signals received at two or more remote sensors has various applications such as sonar, acoustics, geophysics, positioning, tracking, speed sensing, Direction of Arrival (DoA) estimation, exploration, tracking the locations of sources estimation of the number of sources, separation of the individual sources and biomedical engineering [1–9]. The recent growth of machine learning methods redevelops the classical signal processing techniques, working around its restrictions [10].

Analogously, accurate frequencies estimation is a necessity in several engineering domains, such as communications [11], 5G, IoT, e-health, radar or sonar [12], and detection of mechanical or structural faults [13]. It is very important also in several applications in the medical field, where the monitoring of the frequency changes of some bio signals acquired from the human body have a decisive role in diagnosis [14]. Standard frequency estimation methods are the Discrete Fourier Transform (DFT) and the Fast Fourier Transform (FFT). These methods work well for numerous applications but can fail if short signals with low frequency are analyzed because of the rough resulting frequency resolution. The large distance between two consecutive spectral lines makes the chance that the position of a spectral line matches the actual frequency to be small. A consequence is the so-called spectral leakage phenomenon. Different methods to increase the accuracy of the frequency estimate are proposed in the literature. One of them is finding the maximum of the curve that crosses two or three points in the DFT spectrum found by interpolation [15–22]. An alternative interpolation method is zero-padding [23]. A comprehensive review is given in [24].

A deep learning network is used to find the frequency of a noisy sinusoidal wave. A three-layer neural network was designed to extract the frequency of sinusoidal waves that had been combined with white noise at a given Signal to Noise Ratio (SNR) of 25dB. One hundred thousand waves were prepared for training and testing the model. The neural network that could achieve a mean squared error of 4×10^{-5} for normalized frequencies. This model was written for the range 1 kHz $\leq f \leq$ 10 kHz. The trained model can find frequency of any previously unseen noisy wave in less than a second [25].

A learning-based approach is proposed for estimating the spectrum of a multi sinusoidal signal from a finite number of samples is considered in [26]. A neural network is trained to estimate the spectra of such signals on simulated data. Numerical experiments show that the approach performs competitively with classical methods designed for additive Gaussian noise at a range of noise levels and is also effective in the presence of impulsive noise. A machine-learning approach to estimate the frequency of each component in a multi-sinusoidal signal from a finite number of noisy samples uses a neural network to output a learned representation with local maxima at the position of the frequency estimates. a neural-network architecture that produces a significantly more accurate representation and combines it with an additional neural-network module trained to detect the number of frequencies [27].

The previous two problems were addressed together as a joint time delay and frequency estimation problem [28–33], several classical methods were applied. Recent advances in machine learning invite us to revisit our problem again.

Manuscript received July 31, 2023; revised August 14, 2023; accepted September 11, 2023; published January 2, 2024.

This paper is structured as follows. In Section II, the system Adevelopment of the proposed method is presented in Section III. In Section IV, the performance of the proposed method is illustrated through simulations. Finally, some concluding remarks follow in Section V.

II. Problem Formulation

The problem of estimating the power spectral density of a noisy signal has been studied widely [29]. Many methods have been proposed for spectral estimation, which are divided into nonparametric and parametric methods roughly. The nonparametric methods include conventional periodogram, correlogram, temporal windowing, lag windowing, Daniell method, Welch method, Black-Tukey method and so on. The Parametric Methods are Autoregressive (AR), Moving Average (MA), Autoregressive Moving Average (ARMA) and so on. It is possible to estimate the frequencies of complex exponents from the peaks of the spectrum estimated by many methods. Typically, the complex exponentials are the “information bearing” part of the signal. Simply, the estimation of frequencies, that is of interest rather than the spectrum itself. Subspace based methods were applied to estimate the frequencies, using both signal subspace-based methods like Blackman-Tukey, minimum variance, autoregressive and noise subspace-based methods like the minimum norm, Pisarenko, and MUSIC [34].

Consider the discrete-time sinusoidal signals $x(n)$ and $y(n)$ are the measurements of two sensors satisfying:

$$x(n) = s(n) + u(n)$$
$$y(n) = s(n - D) + z(n),\ n = 0,1,\dots,N-1 \tag{1}$$

where

$$s(n) = \sum_{i=1}^{P} a_i e^{j\omega_i n} \tag{2}$$

The source signal $s(n)$ is modeled by a sum of P complex sinusoids where the amplitudes (a_i) are unknown, complex-valued constants, and the normalized radian frequencies (ω_i) are different. Without a loss of generality, we considered $\omega_1 < \omega_2 \dots < \omega_P$. To simplify the problem, we have assumed the number of sources *P* either known or pre estimated. The two terms $u(n)$ and $z(n)$ are representing the two zero mean, additive white complex gaussian noise processes independent of each other. Also, parameters *N* represent the number of samples collected at each channel. The variable D is the delay between the received copies of the signal $s(n)$ at the two separated sensors, which is unknown and is to be estimated. The same system model was used in our previous work in [28, 29].

III. Development of Proposed Method

In this section, we present the development of the proposed method for joint time delay and frequency estimation in noisy complex sinusoidal signals received at two spatially separated sensors. The approach is inspired by recent advancements in deep learning techniques, which have shown promising results in signal processing tasks [35, 36].

A. Frequency Estimation Methodology

The first component of our proposed method focuses on frequency estimation using a Convolutional Neural Network (CNN) regression model. This model is adapted from the work by Bhowmik *et al*. [36], where they successfully applied deep learning for frequency estimation of sinusoidal waves with noise. The CNN model is trained on a synthetic dataset specifically generated to reflect real-world complexities and uncertainties in signal acquisition [2]. We leverage the principles of dropout regularization, as introduced by Srivastava *et al*. [37], to prevent overfitting during training. Additionally, batch normalization, following the approach proposed by Ioffe and Szegedy [38], is incorporated to accelerate the training process and enhance the model's generalization ability. The model architecture comprises multiple layers, including convolutional layers and fully connected layers, designed to capture underlying patterns in the input signals [36]. The synthetic dataset used in this study is generated using a custom function that produces complex-valued signals composed of multiple frequency components. Gaussian noise with a specified noise amplitude is added to the signals to simulate realistic conditions. The visualization of the frequency estimation model architecture is shown in Fig. 1.

The dataset is then divided into training and testing sets. To accommodate the CNN architecture, the input data is reshaped into a three-dimensional array. The CNN regression model is constructed using the Sequential API of Keras. It comprises a series of layers designed to capture the underlying patterns in the input signals. The initial layers consist of a 1D convolutional layer followed by a max pooling layer to extract relevant features. A dropout layer is introduced to prevent overfitting, while batch normalization enhances the model's generalization ability. The subsequent layers flatten the feature maps and pass them through fully connected layers with rectified linear unit (ReLU) activation. The output layer has a number of neurons equal to the desired number of frequency components to predict.

B. Time Delay Estimation Methodology

The second component of our proposed method is dedicated to time delay estimation. For this purpose, we develop a deep learning model that builds upon the concepts used for frequency estimation. The model is inspired by the work of Bhowmik *et al.* [35], where they successfully applied deep learning for time delay estimation in sinusoidal signals received at two spatially separated sensors. In this model, we combine real and imaginary components of the signals to capture their temporal characteristics effectively.

The architecture includes densely connected layers and dropout layers, similar to the frequency estimation model, for regularization and feature extraction [35]. The proposed method employs a deep learning model constructed using the Keras library, a popular framework for developing neural networks. The model architecture comprises multiple layers, including densely connected layers and dropout layers for regularization. To capture the temporal characteristics of signals, we utilize a combination of real and imaginary components. By incorporating both components, our model can extract relevant features and patterns from the input data. The visualization of the time delay estimator model architecture is shown in Fig. 2.

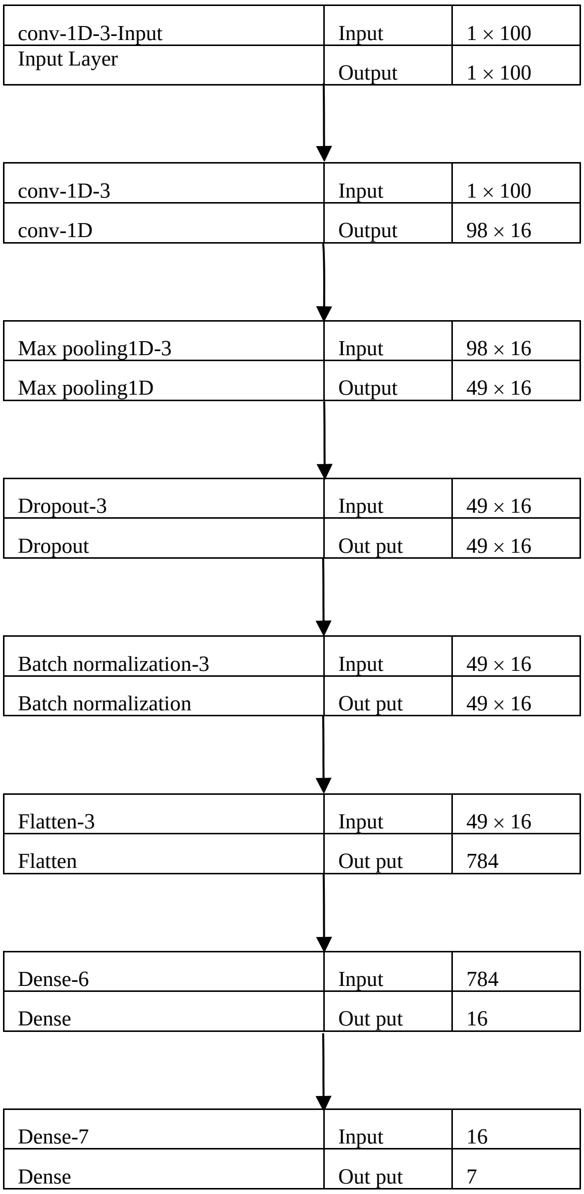

Fig. 1. The visualization of the frequency estimation model architecture.

C. Training and Optimization

Both the frequency estimation and time delay estimation models are trained using Adam optimization, as proposed by Kingma and Ba [1, 2]. The Mean Squared Error (MSE) loss function is employed to measure the discrepancy between predicted and actual frequencies or time delays. The synthetic datasets used for training and evaluation ensure a controlled environment with realistic signal conditions. The training process involves monitoring the loss on both training and validation sets to assess convergence and prevent overfitting [1, 2].

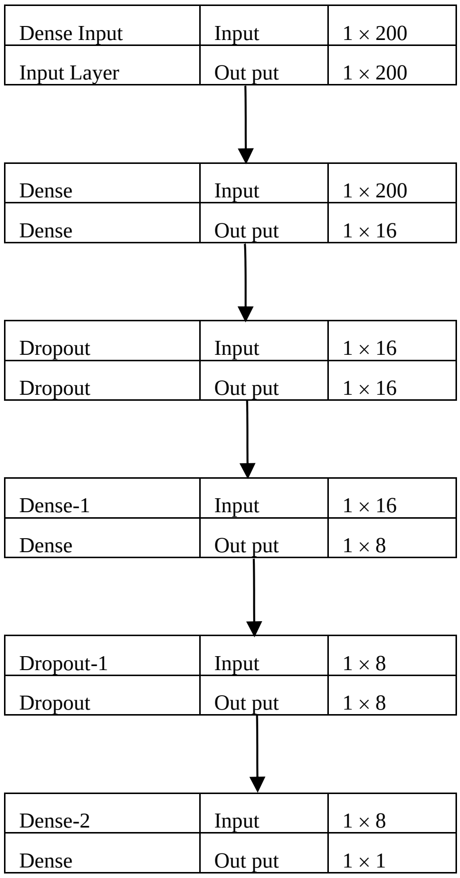

Fig. 2. The visualization of the time delay estimator model architecture.

D. Evaluation and Performance

To evaluate the performance of our proposed method, extensive computer simulations are conducted. The results demonstrate the effectiveness of the deep learning-based approach in accurately estimating joint time delay and frequencies. The models exhibit high accuracy in capturing underlying patterns in complex-valued signals, enabling precise estimation even in the presence of noise.

IV. Simulation Results

Extensive experiments are conducted to evaluate the performance of the proposed CNN regression model. In this section, we present the simulation results of the proposed frequency and time delay estimators.

A. Frequency Estimation

The synthetic dataset, consisting of 2000 samples with three frequency components per sample, is used for training and testing. Normalized frequency range is considered. The model is trained for 30 epochs with a batch size of 32, and a validation split of 0.33 is employed for model selection. The results demonstrate that the proposed CNN regression model achieves high accuracy in frequency prediction tasks. The training and validation loss curves exhibit a decreasing trend, indicating the model's ability to capture the underlying patterns in the complex-valued signals. The evaluation on the testing set confirms the model's effectiveness, with the test loss providing a quantitative measure of its performance. The utilization of deep learning techniques, specifically CNNs, in signal processing tasks has shown promising results. The proposed model leverages the inherent capabilities of CNNs, such as feature extraction through convolutional layers and non-linear mapping through fully connected layers, to capture the underlying patterns in complex-valued signals. The addition of dropout and batch normalization layers further enhances the model's performance and generalization ability. The synthetic *dataset* generation process ensures a controlled environment for training and evaluation. By incorporating multiple frequency components and Gaussian noise, the dataset reflects real-world complexities and uncertainties. The training process utilizes the Adam optimizer and mean squared error loss function which effectively optimizes the model parameters to minimize the discrepancy between predicted and actual frequencies. The experimental results in Fig. 3 demonstrate the model's ability to accurately predict frequencies in complex-valued signals. The decreasing trend in training and validation loss curves indicates the model's capability to learn and generalize from the dataset. The evaluation on the testing set provides quantitative validation of the model's performance.

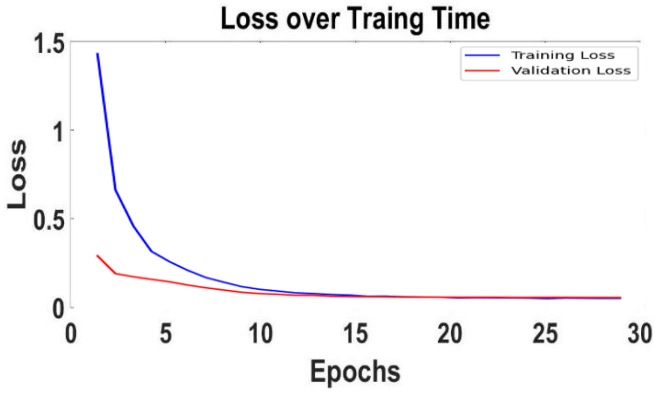

Fig. 3. Training and validation loss for frequencies estimator.

B. Time Delay Estimation

We generated a synthetic dataset consisting of 3000 samples, each with a varying delay and two signal components. The delays were randomly generated within a specified range, while the frequencies of the signal components were uniformly distributed. We added AWG noise to both signal components. to simulate realistic conditions. The resulting signals were concatenated, forming the input data for the deep learning model. The dataset was divided into training and testing sets, with 2400:600 split. The architecture of the deep learning model employed for signal delay estimation is shown in Fig. 4. The figure provides a visual representation of the sequential arrangement of the model's layers. The model consists of three types of layers: dense (fully connected) layers, dropout layers, and the final output layer. The input layer receives signals that are concatenated representations of the real and imaginary ccomponents. The model's architecture includes two dense layers with 16 and 8 neurons, respectively, which are responsible for learning and extracting relevant features from the input data. Dropout layers are incorporated after each dense layer to prevent overfitting by randomly setting a fraction of the input units to zero during training. The final layer is a dense layer with a single neuron, serving as the output layer responsible for estimating the delay. The scatter plot in Fig. 4 illustrates the comparison between the predicted delays and the true delays in the test set. Each data point represents a sample, where the x-coordinate represents the true delay, and the y-coordinate represents the corresponding predicted delay. The points are depicted in blue with an alpha value of 0.5 to indicate the density of the data. Additionally, a red line is plotted, representing the ideal case where the predicted delays perfectly align with the true delays. By visually assessing the plot, we can observe the proximity of the data points to the red line, which serves as a measure of the accuracy of the model's delay estimation.

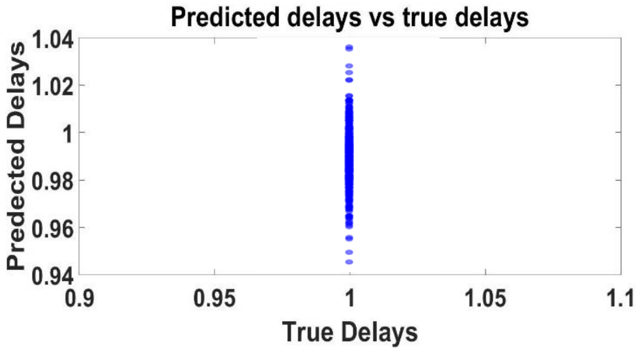

Fig. 4. Predicted delays vs true delays.

Fig. 5 presents the training and validation loss curves during the model's training process. The plot displays the loss values on the y-axis and the number of epochs on the x-axis. The blue line represents the training loss, while the red line represents the validation loss. By examining the plot, we can analyze the convergence and generalization capability of the model. A decreasing trend in both training and validation loss indicates that the model effectively learns the underlying patterns in the data without overfitting. Conversely, a significant divergence between the two lines could suggest potential overfitting or underfitting issues. Thus, Fig. 4 serves as a valuable tool for assessing the model's performance and determining the optimal number of training epochs. These two figures provide a comprehensive visual representation of the model's performance and training progress. They offer insights into the accuracy of the delay estimation and the model's convergence during training, further supporting the effectiveness and reliability of our proposed deep learning-based approach.

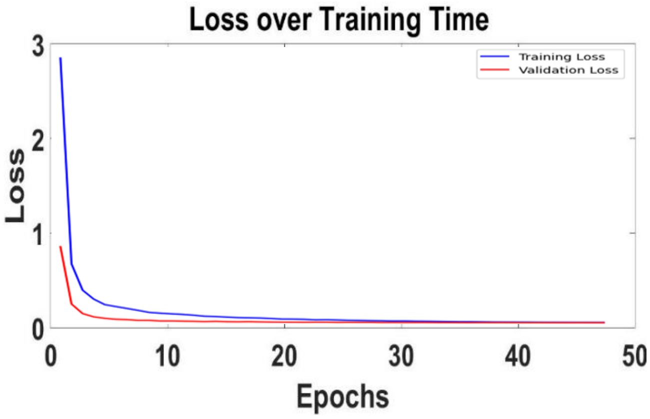

Fig. 5. Training and validation Loss for time delay estimator.

V. Conclusion

We proposed a Deep Learning-Based technique for joint time delay and frequencies estimation of sinusoidal signals received at two separated sensors. We trained the deep learning model on the training set and evaluated its performance on the testing set. The model was optimized using mean squared error loss and the Adam optimizer. During training, we monitored the loss on both the training and validation sets to assess the model's convergence and generalization capability. Overall, the proposed method showcases its potential as an alternative to classical signal processing techniques, offering advanced signal analysis capabilities across various engineering domains.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Mahmoud M. Qasaymeh presented the proposed approach in this paper, conducted the research, analyzed the data, wrote the paper, and reviewed the code of the ML algorithm. Ahmad Falah Aljaafreh wrote the code and reviewed the paper itself. Both authors worked on the concept, literature survey, provided the corresponding interpretations; all authors had approved the final version.

References

[1] S. Qiyan and X. C. Ma. "High-resolution time delay estimation algorithms Through cross-correlation post-processing," *IEEE Signal Processing Letters*, vol 28, pp. 479–483, 2021.

[2] F. L. Ding *et al.*, "Variational bayesian inference time delay estimation for passive sonars," *Journal of Marine Science and Engineering*, p. 194, 2023.

[3] Q. Song and X. Ma, "Block sparse bayesian learning using weighted laplace prior for super-resolution estimation of multi-path parameters," *Global Oceans: Singapore*, U.S. Gulf Coast, 2020.

[4] R. M. R. B. Boopathi and A. R. Mohanty, "Time delay estimation using wavelet denoising maximum likelihood method for underwater reverberant environment," *IET Radar, Sonar and Navigation*, vol. 6, 2020.

[5] R.W u, W. Wang, and Q. Jia, "FFT-based efficient algorithms for time delay estimation," vol. 56, 2012.

[6] J. Ma *et al.*, "Super-resolution time delay estimation using exponential kernel correlation in impulsive noise and multipath environments," *Digit. Signal Process*, vol. 133, 2022.

[7] S. Bagchi and R. D. Fréin, "Evaluating large delay estimation techniques for assisted living environments," *Electronics Letters*, vol. 58, 2022.

[8] R. Wu *et al.*, "Application of RELAX in time delay estimation: Principles and applications of RELAX: A robust and universal estimator," pp 101–157, 2019.

[9] S. Y. Yao, Q. Meng, C. Y. Chen, and I. Tariq, "A DPTF algorithm for the time-delay estimation in the reflected environment," *Digital Signal Processing*, vol. 127, 2022.

[10] L. Houegnigan, P. Safari, C. Nadeu *et al.*, "Neural networks for high performance time-delay estimation and acoustic source localization," *Journal of Computer Science and Information Technology*, pp. 137–146, 2017.

[11] C. Johnson, W. Sethares, and A. Klein, "Software receiver design: build your own digital communications system in five easy steps," 2011.

[12] W. Knight, R. Pridham, and S. Kay, "Digital signal processing for sonar," in *Proc. the IEEE*, vol. 69, no. 11, pp. 1451–1506, 1981.

[13] G. R. Gillich, I. C. Mituletu, Z. I. Praisach, I. Negru, and M. Tufoi, "Method to enhance the frequency readability for detecting incipient structural damage," *IJST-T Mech. Eng.*, vol, 41, pp. 233–242, 2017.

[14] C. Park and D. Lee, "Classification of respiratory states using spectrogram with convolutional neural network," *Appl. Sci.*, vol. 12, pp. 1895, 2022.

[15] B. G. Quinn, "Estimating frequency by interpolation using fourier coefficients," *IEEE Signal Proces*, vol. 42, pp. 1264–1268, 1994.

[16] E. Jacobsen and P. Kootsookos, "Fast, accurate frequency estimators," *IEEE Signal Proc. Mag.*, vol. 24, no. 3, pp. 123–125, 2007.

[17] C. Candan, "A method for fine resolution frequency estimation from three DFT samples," *IEEE Signal Proc. Let.*, vol. 18, no. 6, pp. 351354, June 2011.

[18] E. Aboutanios and B. Mulgrew, "Iterative frequency estimation by interpolation on Fourier coefficients," *IEEE T. Signal Proces.*, vol. 53, no. 4, pp. 1237–1242, May 2005.

[19] T. Grandke, "Interpolation algorithms for discrete fourier transforms of weighted signals," *IEEE T. Instrum. Meas.*, vol. 32, pp. 350–355, June 1983.

[20] V. K. Jain, W. L. Collins, and D. C. Davis, "High-Accuracy Analog Measurements via Interpolated FFT," *IEEE T. Instrum. Meas.*, vol. 28, pp. 113–122, June 1979.

[21] K. Ding, C. Zheng, and Z. Yang, "Frequency estimation accuracy analysis and improvement of energy barycenter correction method for discrete spectrum," *J. Mech. Eng.*, vol. 46, no. 5, pp. 43–48, 2010.

[22] P. Voglewede, "Parabola approximation for peak determination," *Global DSP Magazine*, vol. 3, no. 5, pp. 13–17, May 2004.

[23] J. Z. Xiang, S. Qing, and C. Wei, "A novel single tone frequency estimation by interpolation using DFT samples with zero-padding," in *Proc. IEEE ICSP*, pp. 277–281, March 2017.

[24] A. A. Minda, C. I. Barbinita, and G. R. Gillich, "A review of Interpolation methods used for frequency estimation," *Rom. J. Acoust. Vib.*, vol. 17, no. 1, pp. 21–26, June 2020.

[25] S. Iman and R. Junsuk, "Accurate and instant frequency estimation from noisy sinusoidal waves by deep learning," *Nano Convergence*, vol. 27, 2019.

[26] G. Izacard, B. Bernstein, and A. Fernandez-Granda, "A learning-based framework for line-spectra super-resolution," in *Proc. 44th IEEE International Conference on Acoustics, Speech, and Signal Processing*, ICASSP, pp. 3632–23636, 2019,

[27] I. Gautier and M. Sreyas, and F. G. Carlos, "Advances in neural information processing systems," in *Proc. 29th Annual Conference on Neural Information Processing Systems*, vol. 32, 2019.

[28] H. Shatnawi, H. Gami, and M. Qasaymeh, "High resolution joint time delay and frequency estimation," *SARNOF*, 2009.

[29] M. Qasaymeh *et al.*, "Joint time delay and frequency estimation without eigen-decomposition," *Signal Processing Letters*, pp. 422–425, 2009.

[30] Y. Wu, H. C. So, and P. C. Ching, "Joint time delay and frequency estimation via state-space realization," *IEEE Signal Processing Letters*, vol. 10, pp. 339–342, 2003.

[31] X. Qian and R. Kumaresan, "Joint estimation of time delay and pitch of voiced speech signals," in *Proc. Conf. Rec. 29th Asilomar Conf. Signals, Systems Computers*, vol. 1, pp. 735–739, 1995.

[32] G. Liao, H. C. So, and P. C. Ching, "Joint time delay and frequency estimation of multiple sinusoids," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Processing*, pp. 3121–3124, 2001.

[33] G. Liao, H. C. So, and P. C. Ching, "Joint time delay and frequency estimation of multiple sinusoids" in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing Proceedings*, pp. 3121–3124, vol. 5, 2001.

[34] Hayes, *Statistical Digital Signal Processing*, Wiley, 1996.

[35] U. K. Bhowmik, M. S. Hossain, and A. Ahmad, "Deep learning for time delay estimation in sinusoidal signals received at two spatially separated sensors," *IEEE Transactions on Signal Processing*, vol. 69, pp. 2762–2777, 2021.

[36] U. K. Bhowmik, M. S. Hossain, and A. Ahmad, "Deep learning-based frequency estimation of sinusoidal waves with noise," *Journal of Electrical Systems and Information Technolo*gy, vol. 7, no. 1, pp. 1–14. 2020.

[37] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958. 2014.

[38] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. International Conference on Machine Learning*, pp. 448-456. 2015.