2025 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)

Unsupervised Spectrogram Enhancement Algorithm Based on Bi-LSTM

Hanwen Zhang*‡, Xiruo Su*, Zhijuan Zhu†, Bin Wu*§ and Lingyun Ye *
* College of Biomedical Engineering and Instrument Science, Zhejiang University, Hangzhou, 310058, China
† Detection and Control Technology Research Laboratory, Zhejiang University, Hangzhou, 310058, China
‡ E-mail: hanwen.zhang@zju.edu.cn
§ Corresponding author. E-mail: wubinzju@zju.edu.cn

***Abstract*—Spectrogram enhancement algorithms are essential for the amplification of target tones within complex acoustic environments. Traditional methods, exemplified by the gradient-descent-based Least Mean Square (LMS) algorithm, suffer from a performance that is highly sensitive to the input signal-to-noise ratio (SNR). Conversely, contemporary deep learning techniques, while effective, are predominantly supervised, necessitating extensive labeled data for training, and are often characterized by computationally intensive network structures. This paper introduces an unsupervised spectrogram enhancement algorithm founded on a Bidirectional Long Short-Term Memory (Bi-LSTM) network to surmount these challenges. The proposed algorithm leverages a lightweight Bi-LSTM architecture to achieve exhaustive encoding of tonal features from temporal of the received signals. Experimental validation through simulations indicates that the proposed algorithm yields an SNR gain of $12.58\,$dB when the input SNR is equivalent to the critical SNR of the LMS. Moreover, when evaluated with real-world data under the influence of high-energy narrowband noise, the proposed algorithm demonstrates a robust SNR gain of $8.58\,$dB and exhibits excellent real-time processing capabilities.**

I. Introduction

Analyzing the characteristics of acoustic signals is fundamental to computational auditory applications, including automatic speech recognition, online conferencing, and hearing aid technologies [1], [2]. In complex acoustic environments, where ambient noise often degrades signal quality, spectrogram enhancement algorithms have emerged as critical tools for improving both signal intelligibility and quality. Among these, the gradient-descent-based Least Mean Square (LMS) algorithm is widely employed for adaptive noise reduction, leveraging the distinct correlation properties of narrowband signals and broadband noise to the input signal-to-noise ratio (SNR) [3]–[7]. However, the iterative process in LMS introduces residual noise, leading to a rapid decline in system gain when the input SNR falls below a critical threshold [8]. Consequently, there is an urgent need for spectrogram enhancement algorithms that exhibit greater robustness to low input SNR conditions.

Considering the convergence rate of adaptive methods, there appeared several deep learning methods for noise reduction and line enhancement [9]. Currently, most signal enhancement algorithms based deep learning utilize supervised learning, requiring labeled data, with training datasets typically consisting of strictly matched clean and noisy signal pairs [10], [11]. For example, Xu et al. introduced mapping noisy speech to clean speech signals with a multi-layer deep architecture to perform regression [12], [13]. Maas et al. [14] introduced Recurrent Neural Networks (RNN) for speech enhancement, leveraging their ability to capture long-term temporal dependencies in acoustic signals, which achieves better results than Xu's work. Building on this foundation, Sun et al. [15] employed Long Short-Term Memory (LSTM) networks, which mitigated the gradient explosion and vanishing issues inherent in traditional RNNs when processing extended sequences. Subsequently, Tan et al. [16] proposed a hybrid Convolutional Recurrent Neural Network (CRNN), integrating the spatial feature extraction capabilities of Convolutional Neural Networks (CNNs) with the temporal modeling strengths of RNNs. This architecture achieved more computationally efficient and lightweight speech enhancement, reducing latency while maintaining high-quality signal reconstruction. Shen et al. [17] introduced a sophisticated speech enhancement framework that combines complex graph convolutional networks (GCNs) with a three-channel attention-based RNN. By incorporating decoupled LSTM blocks, their model effectively captures frequency, temporal, and spatial dependencies, enabling robust signal reconstruction across diverse acoustic conditions.

In practical scenarios, target tones are typically unknown, and acquiring labeled training data is challenging. Although simulated signals can be used to construct training sets for pre-training networks on tonal features, the complexity of real-world acoustic environments often leads to degraded performance when transitioning from simulation to experimental settings. In contrast, unsupervised learning offers significant advantages by enabling the discovery of underlying data structures from unlabeled datasets and extracting latent features, thereby enhancing adaptability to complex noise conditions [18]. However, current neural network models, particularly those employing deep architectures, often exhibit high computational complexity, posing significant challenges for deployment in resource-constrained real-world applications, such as embedded systems or real-time processing devices.

Therefore, this paper proposes an unsupervised spectrogram enhancement algorithm based on Bidirectional Long Short-Term Memory (Bi-LSTM) network. Through the Bi-LSTM, the algorithm comprehensively considers the influence of both positive and negative directional features of the received signal on the tones, enabling supplementation of signals with dis-

979-8-3315-7206-8/25/$31.00

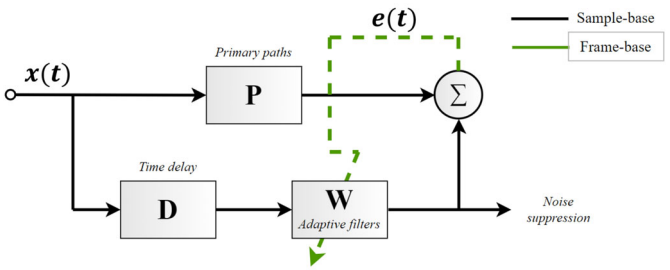


Fig. 1: System block diagram of LMS.

continuous temporal features and ensuring adequate encoding of the tones. Additionally, the algorithm combines adaptive moment estimation (Adam)[19] to update model parameters, effectively handling gradient fluctuations caused by noise, and further enhancing the environmental robustness of the spectrogram enhancement algorithm. Simulations show that the proposed algorithm achieves a gain of $12.58\,\text{dB}$ when the SNR is $-24\,\text{dB}$, which is the critical SNR of LMS. The real-world data indicate that the proposed algorithm effectively addresses the issue of LMS being significantly affected by high-energy narrowband noise. Even when LMS reduces the SNR by $14.99\,\text{dB}$, the proposed algorithm still achieves the SNR gain of $8.58\,\text{dB}$. Furthermore, due to its lightweight network architecture, the proposed algorithm is capable of achieving real-time enhancement performance.

II. Proposed Statement

The LMS is widely considered an effective solution for adaptive noise reduction. It acts as an adaptive narrowband filter taking advantage of the fact that only the spectral components in the input and delayed input are similar to enhance the signal, as shown in Fig. 1.

Assuming that the received signal $\mathbf{x}(t)$ is consist of the tonal signals $\mathbf{s}(t)$ represent and the ambient noise $\mathbf{n}(t)$. Considering the temporal uncorrelation of noise while tonal signals exhibit correlation, the received signal is delayed by a time interval Δ as $\mathbf{x}^D = \mathbf{x}(t+\Delta)$. After the combination with the delayed signal, the output signal of LMS system is defined as

$$y(t) = \sum_{i=1}^{L} w_i(t)x(t+\Delta-i) \tag{1}$$

where L is the the order of the adaptive filter and $w_i(t)$ is the i-th weighting coefficient of LMS.

The loss function is defined as $J(t) = e^2(t)$, where $e(t)$ represents the error signal by

$$e(t) = x(t+\Delta) - w(t)x(t). \tag{2}$$

Thus, the gradient vector and the weight of is updated by

$$\widehat{\nabla}(t) = \frac{\partial[e^2(t)]}{\partial[w(t)]} = -2e(t)x(t) \tag{3}$$

$$w(t+1) = w(t) + 2\mu e(t)x(t), \tag{4}$$

where the gradient vector of the loss function is employed to adjust the weight vector and μ is the learning rate employed to control the convergence speed of the weight vector.

LMS performs a Δ-step forward prediction on the received signals. This process leverages the predictability of regular signals, which exhibit temporal continuity, while noise, being inherently less predictable, is partially suppressed in the output. The core principle of LMS lies in exploiting the statistical properties of the temporal continuity of target features. However, this continuity is highly dependent on the input SNR. When the input SNR falls below a critical threshold, denoted as SNR_{in}^{min}, the system gain decreases sharply. The value of SNR_{in}^{min} is determined by

$$\text{SNR}_{in}^{min} = 10 \times \log_{10} \frac{2}{2\sqrt{L^2+2M_0}-L}, \tag{5}$$

where $M_0 = \frac{1}{4}\mu\sigma_n^2$, where σ_n^2 represents noise power, and $L \gg 1$.

III. Proposed Method

Considering the effectiveness of LMS structure on real-world data and the adaptability of unsupervised algorithms to diverse datasets, this study enhances the LMS structure depicted in Fig. 1 by modifying its adaptive filtering component. The framework of the proposed algorithm are illustrated in Fig. 2.

To enhance spectrogram features, the proposed algorithm leverages delayed data as training inputs for each received signal, with the original data predicted using a Bi-LSTM network, as illustrated in Fig. 2(a). The Bi-LSTM architecture is adopted due to its superior performance in processing time-series signals [20], [21]. Unlike LSTMs, which rely solely on past and present data to predict future states, the Bi-LSTM incorporates both forward and backward temporal contexts, enabling it to capture the influence of future data on the current time step [22]. This bidirectional learning enhances the extraction of temporal memory features, achieving robust encoding of tonal characteristics critical for spectrogram enhancement.

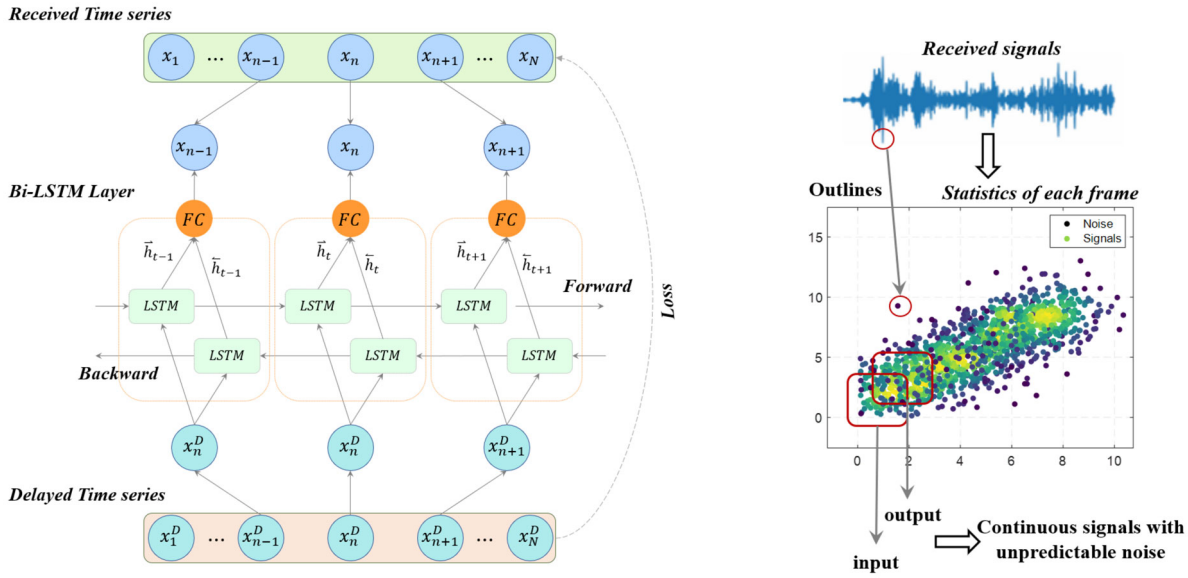
The proposed Bi-LSTM architecture comprises two LSTM layers: LSTM1 for forward sequence encoding, producing the output $\overrightarrow{h_n}$, and LSTM2 for backward sequence encoding, yielding $\overleftarrow{h_n}$. These outputs are concatenated and fed into a fully connected layer (FC), which aligns the combined representation with a predefined number of output response features specified at the start of training. The fully connected layer is parameterized by weights W_r and biases b_r, with the predicted output y_n^* computed as

$$y_n^* = W_r(\overrightarrow{h_n} + \overleftarrow{h_n}) + b_r. \tag{6}$$

To enhance network robustness, a regression layer follows the fully connected layer to constrain the prediction loss, denoted as $\mathcal{L}$. This loss quantifies the deviation between predicted (y_n^*) and actual data (x_n), enabling closed-loop adjustment of learning parameters. The loss is calculated as

$$\mathcal{L} = \frac{1}{2}\sum_{i=1}^{n}(y_i^* - x_i)^2. \tag{7}$$

Parameter optimization is performed using the Adam algorithm, which integrates the benefits of momentum and root



(a) (b)

Fig. 2: (a) System block diagram of unsupervised spectrogram enhancement algorithm based on Bi-LSTM. The received signals is delayed and divided into frames x_n^D, which is regarded as the input to the Bi-LSTM layer. The output signals x_n is the original signals after division. (b) The statistics information of each frame, with deep blue circles as noise and yellow circles as signals.

mean square propagation. Adam accelerates gradient descent by computing a weighted average of gradients (m_n) with a decay rate β_1^n and adaptively adjusts the learning rate based on the weighted average of squared gradients (v_n) with a decay rate β_2^n. The parameter update rule is given by

$$\theta_{n+1} = \theta_n + \alpha \frac{\frac{m_n}{1-\beta_1^n}}{\sqrt{\frac{v_n}{1-\beta_2^n}} + \epsilon}, \tag{8}$$

where θ represents the updated parameters, α is the learning rate, and ϵ is a small constant to ensure numerical stability.

The core of the proposed algorithm lies in using a lightweight network to predict the data itself, where the input data and predicted data share similar or continuous signal components, whereas noise and additional received interference lack such similarity or continuity. The corresponding statistical analysis is presented in Fig. 2(b), which illustrates the statistical characteristics of each frame. In Fig. 2(b), blue dots represent noise, with some outlier noise values deviating from the original statistical distribution, while the signal features appear as continuous yellow bright spots in the central region. This indicates that the prediction process only needs to preserve the continuous statistical components of the received signal. Therefore, unlike traditional deep learning approaches, the proposed algorithm requires no large-scale pre-training dataset and, akin to the LMS, supports online tuning for near real-time tone enhancement.

Finally, the model's parameters are only determined by the Bi-LSTM layer and the fully connected layer. For an input dimension d_x, hidden dimension d_h, and sequence length T, the network has approximately $8 \cdot (d_x d_h + d_h^2 + d_h)$ parameters. The fully connected layer adds $2d_h \cdot d_o + d_o$ parameters, where d_o is the output dimension. In our experiments, d_x and d_o is the frame length, $d_h = 4$, thus the total parameter count is on the order of 10^6 to 10^7, making the model computationally efficient for real-time applications while maintaining robust feature extraction capabilities.

TABLE I: Parameters of LMS and our method

Parameters	LMS	Our method
Filter length	512	-
Frame length	-	2s
Delay time	2s	2s
Learning rate	10^{-7}	10^{-5}

IV. Experiments Analysis

A. Spectrogram enhancement ability

Firstly, the ability of our method to detect signals at low SNR cases compared to the traditional LMS algorithm under white noise is validated. To demonstrate the algorithm's enhancement performance on multi-line spectra, a sampling frequency of $f_s = 16384\,\text{Hz}$ and a duration of $T = 30\,\text{s}$ are employed, with tone signals of equal amplitude added at a frequency of $f_0 = [150, 350, 650, 900\text{Hz}]$. The parameters of LMS and the proposed algorithm are shown in Table I. According to (5), SNR_{in}^{min} under the aforementioned parameter settings is around -24 dB. Therefore, Gaussian white noise is superimposed, varying the input SNR $-36\,\text{dB}$ to $-15\,\text{dB}$.

Simulation results are presented in Fig. 3. The SNR gain of two algorithms under the input SNR from $-36\,\text{dB}$ to $-15\,\text{dB}$ is shown in Fig. 3(a). The sudden drop in SNR gain of the LMS is visualized in black triangles, and the blue circles illustrate the proposed algorithm exhibiting strong robustness

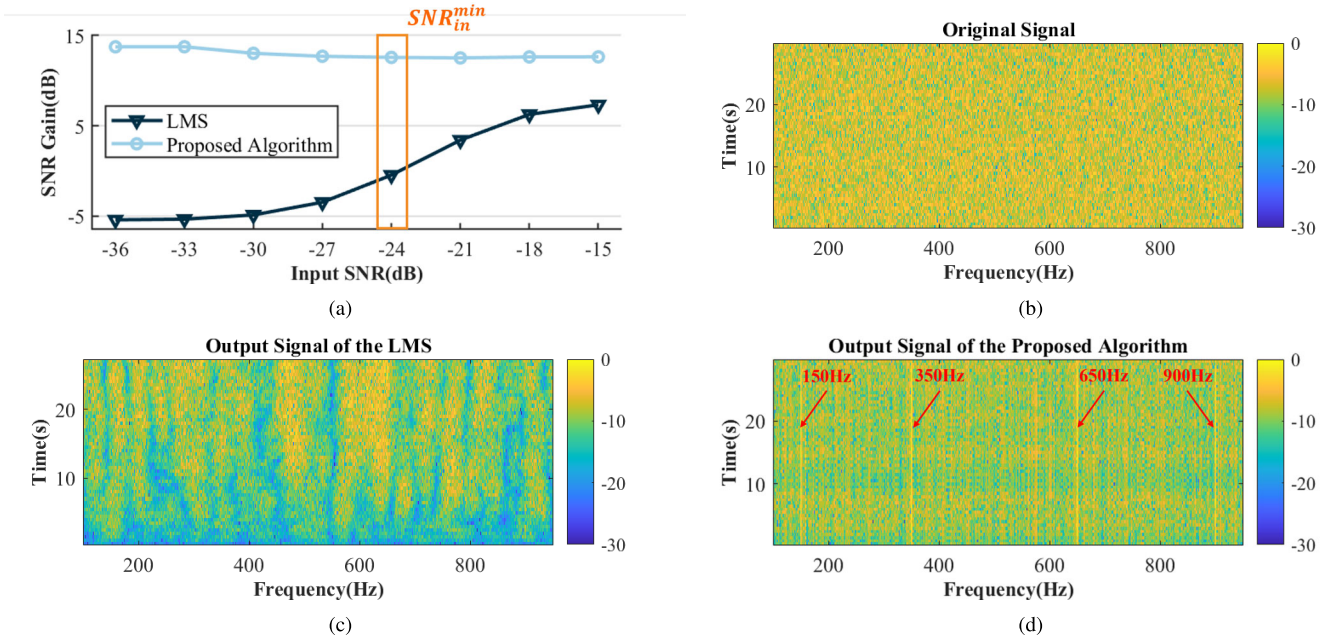


Fig. 3: (a) The SNR gain of the LMS and the proposed algorithm varies with the input SNR. Results of the two algorithms under SNR_{in}^{min}: lofargrams of (b) original signal, (c) output signal of the LMS, (d) output signal of the proposed algorithm.

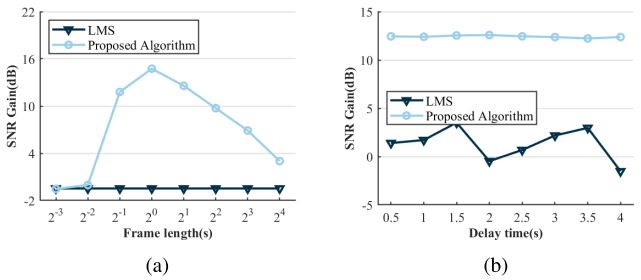


Fig. 4: SNR enhancement performance of LMS and the proposed algorithm influenced by (a) frame length when delay time is 2 s, (b) delay time when frame length is 2 s.

against varying input SNR. In Fig. 3(b), when the input SNR is -24 dB, the target tone is not clearly discernible in the low-frequency analysis and recording spectrum (lofargram). The output signal of the LMS under SNR_{in}^{min} is shown in Fig. 3(c), without discernible tone observed. However, for the proposed algorithm, the tonal signals are visible in Fig. 3(d), with a 12.58 dB SNR gain.

B. SNR enhancement performance

Beyond the inherent learning capabilities of the Bi-LSTM, the performance of the proposed algorithm is also influenced by the frame length and delay time. To further evaluate the algorithm's sensitivity to these parameters, we systematically varied the frame length from 2^{-3} s to 2^{4} s, and the delay module from 0.5 s to 4 s. The corresponding results are presented in Fig. 4.

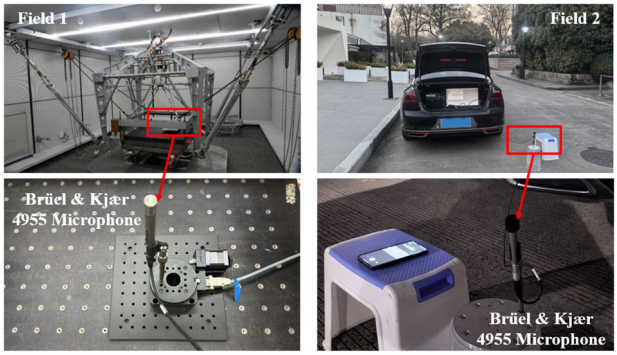
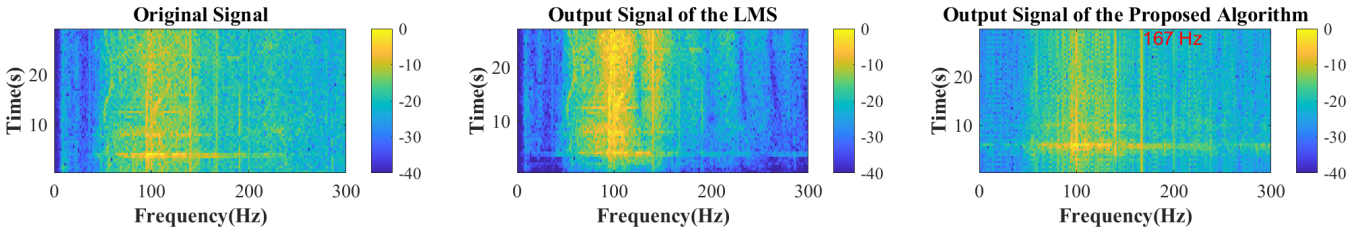


Fig. 5: The signals were sampled using a Brüel & Kjær 4955 microphone at two experimental fields: (a) indoor workspace, (b) outdoor roadway.

The result, shown in Fig. 4(a), demonstrates that the proposed algorithm achieves substantially superior enhancement performance compared to the LMS algorithm under most frame lengths. However, when the frame length is too short or too long, underfitting or overfitting may occur, thereby degrading enhancement performance. By comparing the gain fluctuations with delay between the two algorithms shown in Fig. 4(b), it is evident that the proposed algorithm exhibits lower sensitivity to delay time and demonstrates good robustness.

C. Experiments

In real-world environments, we employed a microphone to capture and process audio signals, with the experimental setup



(a) (b) (c)

Fig. 6: Results of experiments for a sound source level of 40 dB in field 2: lofargrams of (a) original signal, (b) output signal of the LMS, (c) output signal of the proposed algorithm.

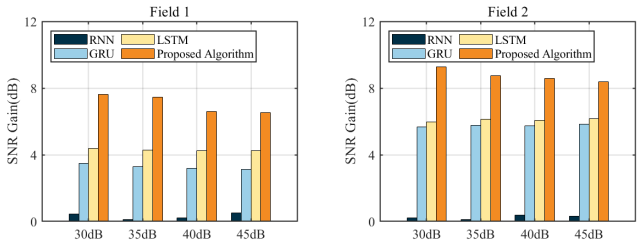
TABLE II: The SNR gain and the execution time of the LMS and the proposed algorithm

Evolve by	Field 1		Field 2	
Source level	LMS	Our method	LMS	Our method
30dB	-13.96	7.61	-14.16	9.26
35dB	-14.04	7.47	-14.55	8.76
40dB	-12.28	6.59	-14.99	8.58
45dB	-6.01	6.54	-6.91	8.39
Execution time(s)	0.36	0.02	0.36	0.02

depicted in Fig. 5. A Brüel & Kjær 4955 microphone was used, with the sound source being of type R1200T II, continuously emitting a 167 Hz sinusoidal signal at a predefined sound source level. The data were sampled at a frequency of $f_s = 16384$ Hz over a duration of $T = 30$ s. Each group of sampled data were processed directly, and the comparative processing results for a sound source level of 40 dB in field 2 are presented in Fig. 6. The SNR gain and the execution time of LMS and the proposed algorithm across four different sound source levels are summarized in Table II.

Fig. 6(b) illustrates that, for the 167 Hz tone at the critical SNR, the enhancement provided by the LMS is limited. Although it effectively suppresses some noise, the tone remains difficult to observe in the lofargram, with a decrease of 14.99 dB in the SNR. This occurs because, in an environment with broadband noise, the LMS updates the weighting coefficients to suppress irrelevant noise components and detect narrowband signals that remain stable over time. However, the lofargram of the original data, shown in Fig. 6(a), reveals a significant amount of high-energy narrowband noise near the 167 Hz tone, which is caused by the operation of air conditioning units and other rotary mechanical equipment in the experimental environment. This portion of the noise cannot be effectively suppressed by the LMS algorithm, especially when the energy of the narrowband noise exceeds that of the target signal, the LMS exacerbates the energy difference between strong narrowband noise and weak tones, thereby suppressing the characteristic components of the tones and hindering the enhancement of weak tones.

Compared to LMS, the proposed algorithm is less affected by narrowband noise and achieves a significant enhancement, as shown in Fig. 6(c), with a 8.58 dB improvement in SNR. The results in Table II show that under different levels of sound source and environmental conditions, the proposed algorithm



(a) (b)

Fig. 7: SNR enhancement performance of the RNN, GRU, LSTM and Bi-LSTM for different sound source level from (a) field 1, (b) field 2.

always remains with significant enhancement effects even when the LMS cannot enhance the tone. At the same time, with a lightweight network structure, the proposed algorithm has a fast response speed and is capable of achieving real-time enhancement.

In order to more comprehensively demonstrate the advantages of the Bi-LSTM architectures in the proposed algorithm, three common networks—RNN, GRU, and LSTM—are also selected as reference methods for signal enhancement. Fig. 7 illustrates the processing results of the proposed algorithm and the three aforementioned networks for signals with varying power levels in two experiment fields. The proposed method demonstrates superior signal enhancement performance, thereby fully confirming the specific effectiveness of Bi-LSTM architectures for time-series signals.

V. Conclusion

This paper proposes an unsupervised spectrogram enhancement algorithm based on Bi-LSTM. This algorithm utilizes the strong nonlinear capabilities of neural network and the ability of Bi-LSTM to capture forward and backward dependencies in time series to address the limitation of the LMS algorithm, which experiences a decline in gain under critical SNR conditions. Simulations show that the proposed algorithm performs good robustness to spectrogram enhancement. When operating at the critical SNR of the LMS, the proposed algorithm still achieves good enhancement effects on target tone, with an SNR improvement of 12.58 dB. Experiments demonstrate that the proposed algorithm also addresses the challenge of weak tone

enhancement and enhances tonal signals in real-time. Under the influence of high-energy narrowband noise, the proposed algorithm achieves a gain of $8.58\,\mathrm{dB}$ dB even when the LMS experiences a SNR decline of $14.99\,\mathrm{dB}$ dB. Compared to other commonly used lightweight neural networks, it also demonstrates significant advantages in terms of signal enhancement.

References

[1] D. Michelsanti, Z.-H. Tan, S.-X. Zhang, *et al.*, "An overview of deep-learning-based audio-visual speech enhancement and separation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 1368–1396, 2021.

[2] M. J. Bianco, P. Gerstoft, J. Traer, *et al.*, "Machine learning in acoustics: Theory and applications," *The Journal of the Acoustical Society of America*, vol. 146, no. 5, pp. 3590–3628, Nov. 2019.

[3] A. Ali, M. Moinuddin, and T. Y. Al-Naffouri, "Nlms is more robust to input-correlation than lms: A proof," *IEEE Signal Processing Letters*, vol. 29, pp. 279–283, 2022.

[4] S. Maiti, D. Adusumalli, K. Keshan, S. Sharma, S. H. Pauline, and S. Dhanalakshmi, "Leaky lms algorithm based low complexity adaptive noise cancellation," in *2025 International Conference on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence (RAEEUCCI)*, 2025, pp. 1–6.

[5] S. Sanei, T. K. Lee, and V. Abolghasemi, "A new adaptive line enhancer based on singular spectrum analysis," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 2, pp. 428–434, 2011.

[6] X. Su, D. Shi, B. Wu, L. Ye, and W.-S. Gan, "Co-forecasting of time-varying spatial-frequency map for selective fixed-filter multichannel anc based on dynamic factor graph," *IEEE Transactions on Audio, Speech and Language Processing*, 2025.

[7] C.-Y. Ho, K.-K. Shyu, C.-Y. Chang, and S. M. Kuo, "Efficient narrowband noise cancellation system using adaptive line enhancer," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 1094–1103, 2020.

[8] A. Nehorai and D. Malah, "On the stability and performance of the adaptive line enhancer," in *ICASSP '80. IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, 1980, pp. 478–481.

[9] P. C. Loizou, *Speech Enhancement: Theory and Practice*, 2nd. USA: CRC Press, Inc., 2013, ISBN: 1466504218.

[10] Y. Wang, A. Narayanan, and D. Wang, "On training targets for supervised speech separation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 12, pp. 1849–1858, 2014.

[11] D. Wang and J. Chen, "Supervised speech separation based on deep learning: An overview," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, no. 10, pp. 1702–1726, 2018.

[12] Y. Xu, J. Du, L.-R. Dai, and C.-H. Lee, "An experimental study on speech enhancement based on deep neural networks," *IEEE Signal Processing Letters*, vol. 21, no. 1, pp. 65–68, 2014.

[13] Y. Xu, J. Du, L.-R. Dai, and C.-H. Lee, "A regression approach to speech enhancement based on deep neural networks," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 1, pp. 7–19, 2015.

[14] A. L. Maas, Q. V. Le, T. M. O. 'Neil, O. Vinyals, and A. Y. Ng, "Recurrent neural networks for noise reduction in robust asr," in *Conference of the International Speech Communication Association*, 2012.

[15] L. Sun, J. Du, L.-R. Dai, and C.-H. Lee, "Multiple-target deep learning for lstm-rnn based speech enhancement," in *2017 Hands-free Speech Communications and Microphone Arrays (HSCMA)*, 2017, pp. 136–140.

[16] K. Tan, X. Zhang, and D. Wang, "Real-time speech enhancement using an efficient convolutional recurrent network for dual-microphone mobile phones in close-talk scenarios," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 5751–5755.

[17] X. Shen and W.-P. Zhu, "Multichannel speech enhancement using complex-valued graph convolutional networks and triple-path attentive recurrent networks," in *2024 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 2024, pp. 1–6.

[18] M. Wang, X. Bai, C. Zhang, and Y. Zhong, "Unsupervised weak speech enhancement using periodic mixing invariant training," *Circuits, Systems, and Signal Processing*, pp. 1–22, 2025.

[19] C. Chen, L. Shen, W. Liu, and Z.-Q. Luo, "Efficient-adam: Communication-efficient distributed adam," *IEEE Transactions on Signal Processing*, vol. 71, pp. 3257–3266, 2023.

[20] Z. Zhang, H. Luo, C. Wang, C. Gan, and Y. Xiang, "Automatic modulation classification using cnn-lstm based dual-stream structure," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 11, pp. 13 521–13 531, 2020.

[21] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Transactions on Neural Networks*, vol. 5, no. 2, pp. 157–166, 1994.

[22] Y. Li, Y. Xu, G. Li, S. Liu, W. Liu, and J. Hu, "Deep learning-enhanced anti-jamming decoder for otfs systems: A cnn bi-lstm hybrid approach," *IEEE Transactions on Cognitive Communications and Networking*, pp. 1–1, 2025.