Speech Enhancement using Discrete Wavelet Transform with Long Short-Term Memory Algorithm

Senthamizh Selvi R[1], Gurumita Naidu G[2*], Gandla Tejaswini[2]

[1]Professor and head, Department of Electronics and Communication Engineering, Easwari Engineering College, Chennai-India
[2]Student UG Scholar, Department of Electronics and Communication Engineering, Easwari Engineering College, Chennai-India
Email Id- g.gurumita@gmail.com

This work aims to propose an innovative approach for improving corrupted speech signals through signal processing. Many filtering techniques were used to achieve speech enhancement such as adaptive filtering, spectral subtraction, wiener filtering, and neural network algorithms such as CNN, RNN, LSTM, etc., The Proposed method concentrates in refining the quality of the nature and also intelligibility of the speech signals using Discrete Wavelet Transform and Long Short-Term Memory (DWT – LSTM) Algorithm. The DWT technique is employed as a feature extraction technique to decompose the noisy speech signal into various frequency components and denoising it. The Output of the DWT algorithm is given to the LSTM network as input and it is effectively trained to enhance the denoised signal. From the objective evaluation, it has been observed that t he DWT- LSTM algorithm is more productive in achieving the clean speech signal and also provides the best outcomes for the noises that are stationary. This speech enhancement algorithm demonstrates excellent quality in objective measures, achieving a 30.6dB output SNR at a 15dB input noise level.

Keywords: Speech enhancement; Discrete Wavelet Transform (DWT); Long Short Term Memory (LSTM); Signal to Noise Ratio (SNR).

1. Introduction

Speech enhancement systems are utilized for enhancing the clarity and fidelity of the audio signal that are corrupted by noise [1]. Speech corruption can occur due to real-world noises such as Train Noise, Restaurant Noise, Street Noise, Babble Noise, Car Noise, etc.. The audio signals quality and clarity can be significantly impacted by these noises. Numerous advanced algorithms have been devised, and intensive research has been ongoing for the last three decades [2,3]. Speech enhancement procedures are often categorized into filtering, spectral restoration, and speech model techniques. These algorithms are widely embraced

due to their low complexity in both performance and computation, but they exhibit reduced effectiveness when faced with noise in actual environments [4]. Speech Enhancement involves in many applications where clean communication is important. Some of the main applications of speech enhancement are telecommunications, Entertainment, Healthcare and surveillance. The techniques in speech enhancement reduces the background noise, disturbance and interference by assuring a clean and clear communication in Phone calls, Video conferences and Voice-over-internet-protocol (VoIP) systems. Speech enhancement plays a crucial role in making effective communication, increasing productivity, and enhancing user experience over several domains.Due to the advancements in technology, a hybrid algorithm is proposed by combining a wavelet transform algorithm and a neural network-based algorithm [5].

Maximilian Strake and Bruno Defraene had introduced a speech enhancement using two-step access. The first step is Noise Suppression which was continued by Long Short-Term Memory (LSTM) and the second step is Speech Restoration which was again followed by Convolutional Neural Network (CNN) via Spectral Mapping. The obtained result for this proposed methodology achieved a PESQ score of 0.1 for non- stationary noise types, including intrusive speech. In their methodology, Strake and Defraene had noticed that their model was only able to produce the speech intelligibility only in low SNR conditions [6]. Senthamizh Selvi R, Sathish Kumar P, Sri Krishna R and Surya Rao S had performed analysis on enhancing the speech with the usage of different windows, transformation techniques and overlapping percentage. The different types of windows used in this research was Hanning, Blackmann, Hamming and cosh windows and the various signal processing techniques include Fast Fourier Transform (FFT) and the Discrete Cosine Transform (DCT). The highest SNR value obtained was 44.2409 dB and the overlapping percentage was 50% [7]. Michelle Gut ierrez-Munoz and Marvin cotojimenez utilized a hybrid methodology that combines Deep learning and Wavelet Transformation Techniques. This experiment was taken to test whether the hybrid approach like Deep Learning and Wavelet Tansformation combination was beneficiary for speech enhancement or not under several conditions. This method was mainly to select the parameters of Wavelet Transform and also to check over forty methods of training Deep Neural Networks (DNN). This work is done mainly to check whether the hybrid approach is beneficial in the future implementations [8]. Shruthi, Senthamizh Selvi.R and G.R.Suresh had proposed a work to deal with the issue that predicts the average intelligibility of signals that have been distorted and processed,which was then observed by a group of listeners without hearing impairments. But, This model can give a prediction only for a short-term and not form the long-term. The output of this model was obtained using a fundamental called Mean Square Error (MSE), Which evaluates the amplitude of the noise signal[9]. Vinay N A and Bharathi S H had worked under the design and development of algorithm for the Speech Recognition. Here, the researchers used a speech enhancement model known as Recurrent Neural Network (RNN). Due to issues encountered with the RNN, they proposed a new model called Bidirectional Long Short-Term Memory(BLSTM). In this work, they assessed the word error rate and measured the accuracy with Word based Confusion Matrix (WCM) [10]. Ming Liu had done a work by train ing and conducting comparative analysis on various Speech Enhancement methods. This study involved varying the number of parameters to assess their effectiveness. This work produced an efficient output which tells Long Short-Term Memory

(LSTM) achieved the best result when compared with the other speech enhancement models such as Deep Neural Network (DNN), Convolutional Neural Network (CNN) and also Bidirectional Long Short Term Memory Model (BLSTM). The noise effects in the LSTM model had deducted from 31.23% to 25.89% on the Xiaomi Speaker test set [11]. Xiaofei Li and Radu Horaud had proposed a work relating to Multi-channel Speech Enhancement using Sub-band Long Short-Term Memory (LSTM) Network.The Sub-band Long Short Term memory (LSTM) Network generates the contrast between spatial features of noise and speech. Here the source of the speech taken was non-stationary and clear. Whereas, the source of the noise taken was stationary and less spatially co-related. This work had given a conclusion as that this method performs in the Deep Learning with Full-band method and also it is unsupervised [12].

To overcome the above discussed limitations in the related works, this paper have proposed a novel algorithm using Discrete Wavelet Transform (DWT) with the combination of Long Short-Term Memory (LSTM). This paper is organized as Introduction in Chapter-I, Proposed Methodology in Chapter-II, Results and discussion in Chapter-III and Conclusion in Chapter-IV.

2. MATERIALS AND METHODS

The audio is collected from different databases such as Timit speech corpus, Aurora database and Noisex-92 database and sent as the input. For any Speech Enhancement technique, Pre-Processing plays a main role and must carry on in the first step in the process. So, input is sent to the Pre-Processor block to down-sample by a factor of 2. Pre-Processing must be done before starting any of the Speech Enhancement techniques. Later, the data is sent to the DWT algorithm block, where the denoised signal is obtained. That denoised signal is given as the input for the other algorithm LSTM, where the enhanced speech signal is obtained. The complete process of this algorithm is shown in fig 1.

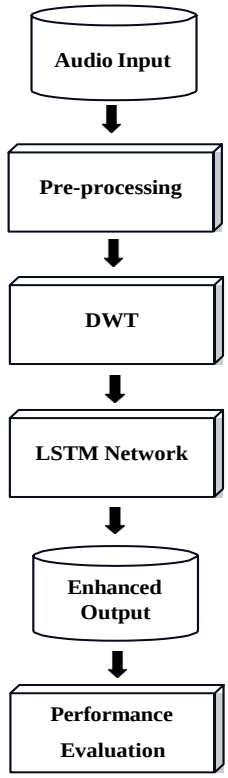


Fig.1. Schematic representation of proposed speech enhancement using DWT-LSTM algorithm

A. AUDIO INPUT

Audio recordings are collected from diverse sources like Timid speech corpus, Aurora database, Noisex-92 database, which includes different levels and types of background noise, reverberation, and speech patterns.

B. PRE-PROCESSING

The first step of any speech enhancement method would be pre-processing. Data cleaning in this process differs from dataset to dataset. It segments the audio data into smaller frames and extract relevant features, and plot it in a graph as spectrograms, to represent the audio signals.

C. DWT

After the pre-processing operation the signal gets transferred to the next block called as DWT. A discrete wavelet transform (DWT) is a technique that breaks down a input signal into multiple sets. These sets consist of time series coefficients that provides the insights of signal evolution in corresponding frequency band. The original signal is a sum of

these wavelets. Coefficients are grouped into packages called sub-bands. Each sub-band collects coefficients resulting from band-pass filtering followed by a sub-sampling. Discrete Wavelet Transform (DWT) is used in many applications such as science, engineering, mathematics and computer science[13]. It is highly notable in the usage of Signal coding, for the representation of a discrete signal by compressing the data. Discrete Wavelet Transform (DWT) block gets the signal and extract the features called as wavelet co-efficients [14]. They carry the information of the weight, which contributes a wavelet basis to the function. DWT decomposes the signal into different scales which capture the both time and frequency information[15,16].

When performing a DWT signal decomposition, the coefficients are obtained by filtering the signal with low- pass (Lo) and high- pass (Hi) filters,and then down-sampling it by a factor 2. The two co- efficients are Approximation (A) and Detail (D) co- efficients.

The basic formulation to find the Approximation and Detail co-efficients are;

$A[n] = (s*Lo)[n]\downarrow 2$

$D[n]=(s*Hi)[n]\downarrow 2$ Where:

s is the noisy speech signal

A[n] = Approximation coefficients, D[n] = Detail coefficients, 2 represents down sampling by a factor of 2

The main goal of selecting DWT algorithm is to extract the main features which are needed for the further process. Here, in our approach the features taken for extraction is wavelet co-efficients [17].

Thus, the Wavelet co-efficients are extracted from this DWT algorithm and obtained as the output of DWT. And this output obtained from DWT will be transferred as the inputs to the next block, LSTM Network.

D. LSTM NETWORK

The main objective in choosing Long Short-Term Memory (LSTM) is because it is an extended version of RNN. LSTM, which is a type of Recurrent Neural Network (RNN), that deals with the vanishing gradient problem in Recurrent Neural Networks (RNN) [18]. Vanishing Gradient problem is an occurrence which occurs when the Deep Neural Network (DNN) gets trained. Because, the gradients which are used to get updated in the network becomes very small or it completely gets vanished. So, LSTM is a solution for not facing this vanishing gradient problem. LSTM process the entire sequence of data instead of individual data points which are done in the traditional Recurrent Neural Networks (RNNs). LSTM is also mainly for the long-term dependencies, which were also not able to process in the traditional Recurrent Neural Networks (RNNs). LSTM can also make predictions based on the datasets that are trained [19].

The output attained from DWT process is given as the input for this LSTM Network block. Firstly, this LSTM Network should be trained by the datasets taken. LSTM Architecture mainly have three gates. Which are Forget Gate, Input or Store Gate and the Output Gate.

In the LSTM Network, the first step is the Forget gate. The function of the forget gate, which

is used to check if the information achieved from the previous time stamp is to remember or to forget. The activation function in this gate is Sigmoid function, which has the output values as either “0” or “1”.If the value obtained from the forget gate equation is “0”,then it tells the network to forget the previous information and if the value obtained is “1”, then it tells the network to remember the information. The formula to evaluate the forget gate is;

$$f_t = \sigma(Y_t * W_i + H_{t-1} * W_f) \quad (1)$$

Where, Yt represents the input given to current timestamp,

Wi denotes the weight which is linked with input, Ht-1 signifies the hidden state from the previous timestamp, and Wf stands for the matrix of weight that is connected with the state that is hidden.

The second gate is known as Store gate. Its purpose is modifing the data based on new information received. The activation function used in this gate is the hyberbolic tangent (tanh) function, which constraints the new information value to be within the range of -1 to 1. If Nt has a negative value, the input gate will perform subtraction on the cell state using the data.If Nt has a positive value , then the input gate adds or updates the data in the cell state. The formula for the input or the store gate is;

$$i_t = \sigma(Y_t * W_i + H_{t-1} * W_{ih}) \quad (2)$$

Yt represents the input at time t, Wi signifies the weight matrix corresponding to the input, Ht-1 denotes the hidden state from the previous timestamp, and Wih represents the weight matrix incorporating both the input and the hidden state. The formula for obtaining new information is

$$N_t = \tanh(Y_t * W_c + H_{t-1} * W_s) \quad (3)$$

After attaining the new information, the formula for updating or adding the cell state is;

$$C_t = f_t * C_{t-1} + i_t * N_t \quad (4)$$

Ct-1 represents the state of the cell at the present timestamp and the values that are left are which we have computed before.

The third and the last gate in the LSTM Network in a LSTM unit is said to be an Output gate. The purpose of the output gate is to transfer the updated data from the input gate to the output.The function that provides activation for the Output gate is the Sigmoid function, where the values of the output are either “0” or “1”.

$$O_t = \sigma(X_t * U_o + H_{t-1} * W_o) \quad (5)$$

To compute the present hidden state;

$$H_t = O_t * \tanh(C_t) \quad (6)$$

The state which is hidden is determined by a combination of the long-term memory (Ct) and the current output to obtain the result of present timestamp, apply SoftMax activation function to the hidden layer, Ht,

$$Output = Soft\max(H_t) \quad (7)$$

The value with the topmost result in the output, is the prediction done by the LSTM Network.

E. AUDIO OUTPUT

The output is obtained from the LSTM Network. LSTM Network gets trained and takes the input to produce the enhanced speech signal.

F. PERFORMANCE EVALUATION

After the enhanced output is attained from the LSTM block, subjective and objective analysis is been carried out for the evaluation of the enhanced signal of the speech.Subjective analysis is performed by making the listener listen it it if the obtained speech signal is clear or noise is reduced when compared to the original noise speech signal.

Objective analysis is to compute by the values obtained. The objective analysis carried out here is Signal to Noise Ratio (SNR). Signal to Noise Ratio (SNR) is expressed as the proportion of the power of a signal obtained to the power of the background noise present in the speech signal. The unit of SNR is decibels (dB).

The formula to compute SNR is;

$$SNR(dB) = 10.\log_{10}(P_s / P_n) \quad (8)$$

Where;

P_s is the power of the signal

P_n is the noise of the signal

The signal's power is computed by taking the averaging the square of its amplitude values of the signal over a specified time period or frequency range. The formula of power of the signal (Ps) is;

$$P_s = N_1 \sum n = 1N \,|\, (n) \,|\, \text{^}2 \quad (9)$$

The amplitude of the signal at time index n is represented by x(n),and N denotes the total number of samples in the signal.

The power of the noise is similarly computed by taking the average of the squared amplitude values of the noise over the same time period or frequency range. The formula for the power of the noise (Pnoise) is;

$$P_n = N_1 \sum n = 1N \,|\, \text{noise}(n) \,|\, \text{^}2 \quad (10)$$

Where noise(n) is the amplitude of the noise at time index n, and the total number of samples in the noise signal is represented by N. Higher SNR values indicate a higher quality signal with less interference from noise [20].

3. RESULTS AND DISCUSSION

The TIMIT corpus and Aurora database of read speech is anticipated in supplying the speech

data for acoustic-phonetic studies [21]. It is also intended for the evolution and estimation of the systems which are automatic in speech recognition. TIMIT involves the recordings of broadband with 630 number of speakers in eight major dialects of American English, each reading ten phonetically rich sentences. The TIMIT corpus transcriptions have been hand verified. The degraded signals are collected from many types of noises at different SNRs that ranges from 5dB to 15dB [22]. The Full process is done in MATLAB R2023a. It is a simulation software in which the Audio toolbox and Deep learning toolbox is utilized to achieve Enhanced Speech. The waveform for Noisy audio and Denoised audio which is visualized in Fig.1. and Fig.7.

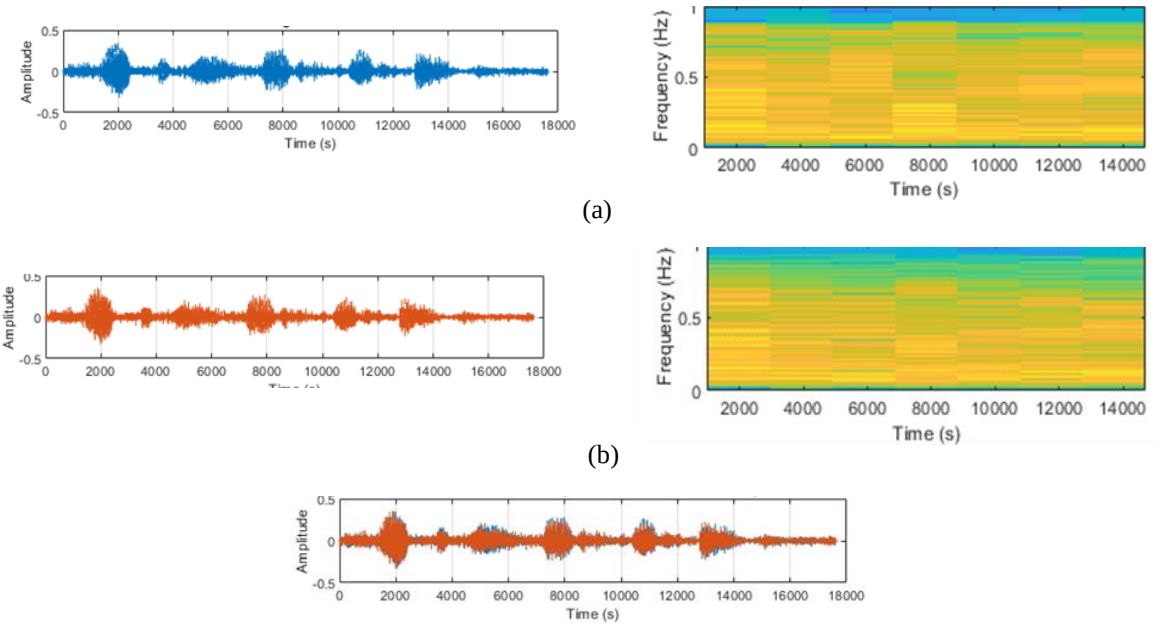
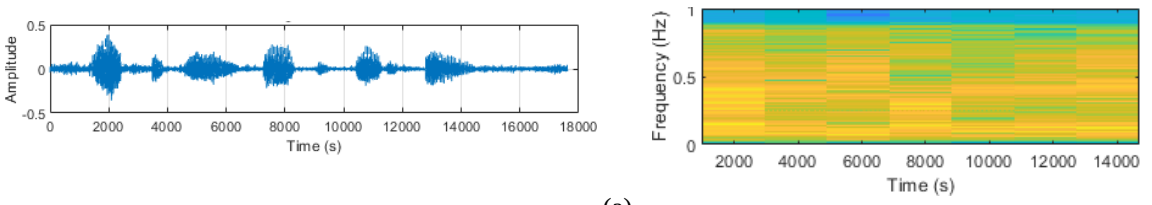
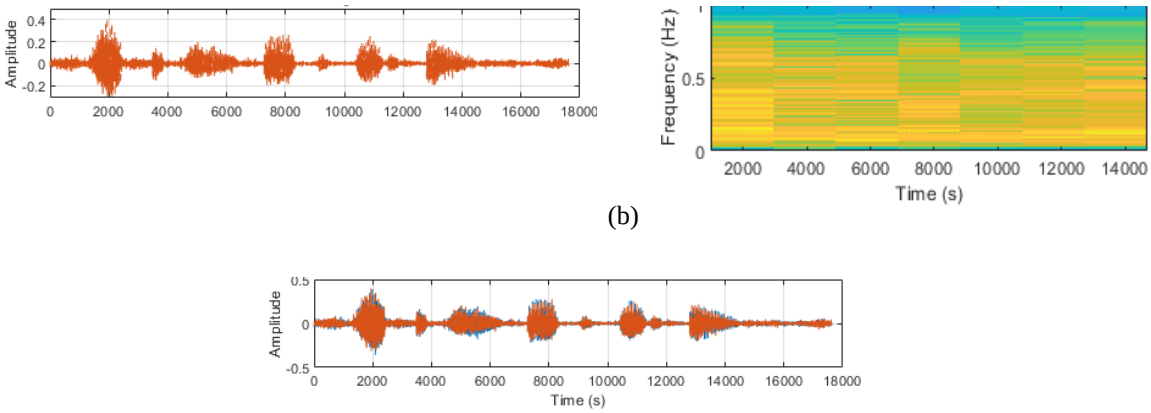


Fig.2. (a) Waveform and Spectrogram of Noisy speech (Babble Noise with SNR 5dB) (b) Waveform and Spectrogram of Denoised speech (Babble Noise with SNR 5dB) (c) Waveform of Difference Between Noisy and Denoised speech.

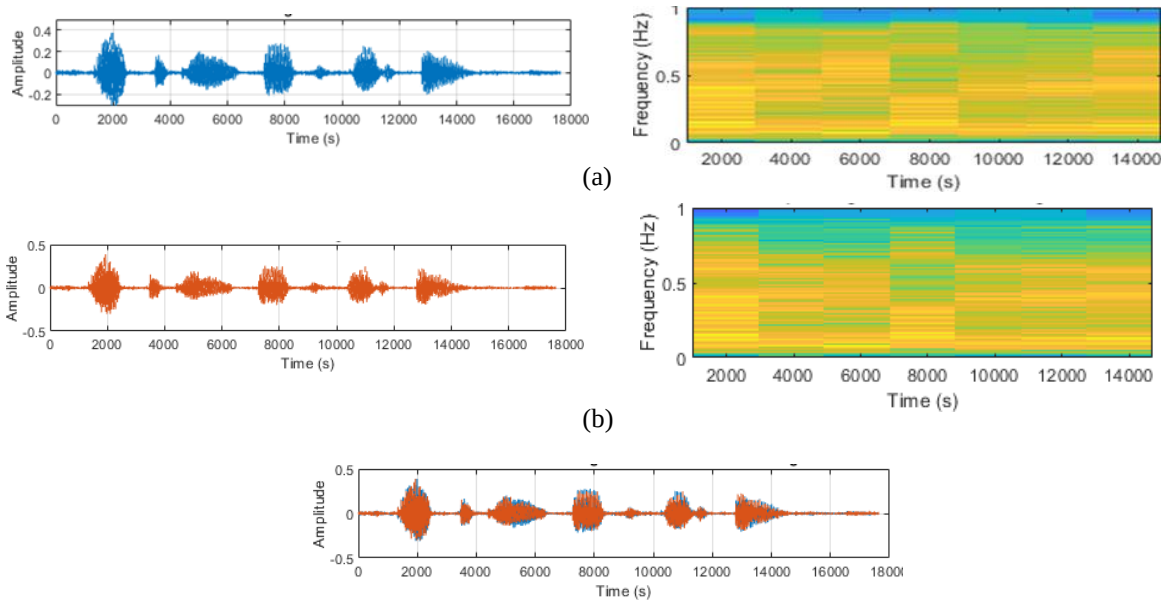




(b)

(c)

Fig.3. (a) Waveform and Spectrogram of Noisy speech (Babble Noise with SNR 10dB) (b) Waveform and Spectrogram of Denoised speech (Babble Noise with SNR 10dB) (c) Waveform of Difference Between Noisy and Denoised speech.



(a)

(b)

(c)

Fig.4. (a) Waveform and Spectrogram of Noisy speech (Babble Noise with SNR 15dB) (b) Waveform and Spectrogram of Denoised speech (Babble Noise with SNR 15dB) (c) Waveform of Difference Between Noisy and Denoised speech.

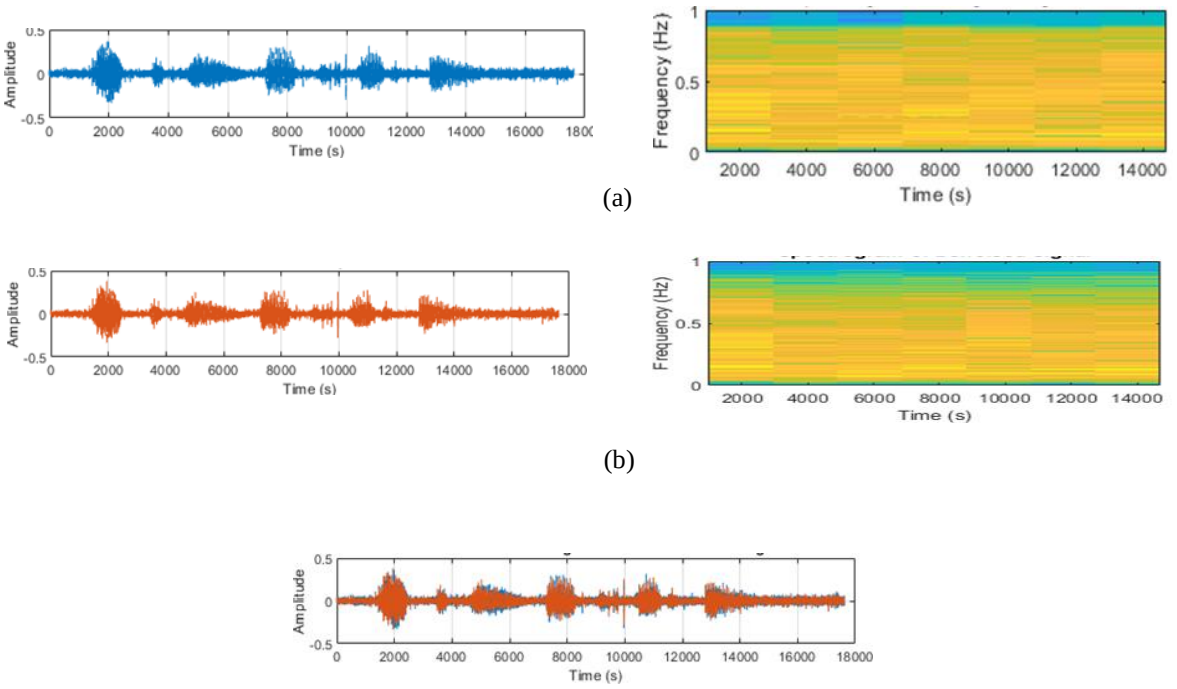


Fig.5. (a) Waveform and Spectrogram of Noisy speech (Street Noise with SNR 5 dB) (b) Waveform and Spectrogram of Denoised speech (Street Noise with SNR 5 dB) (c) Waveform of Difference Between Noisy and Denoised speech.

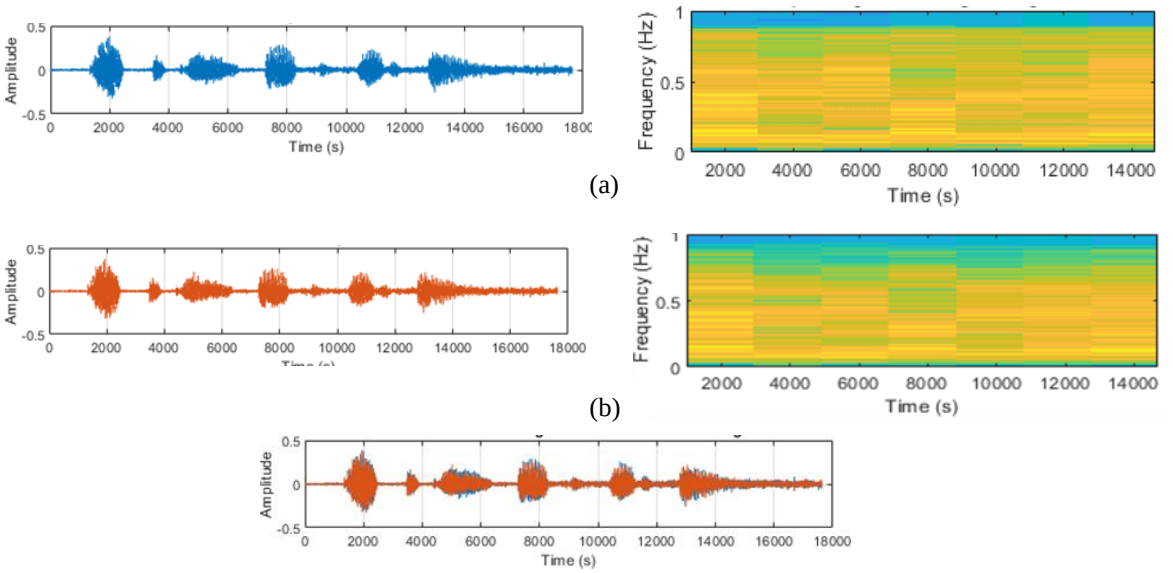
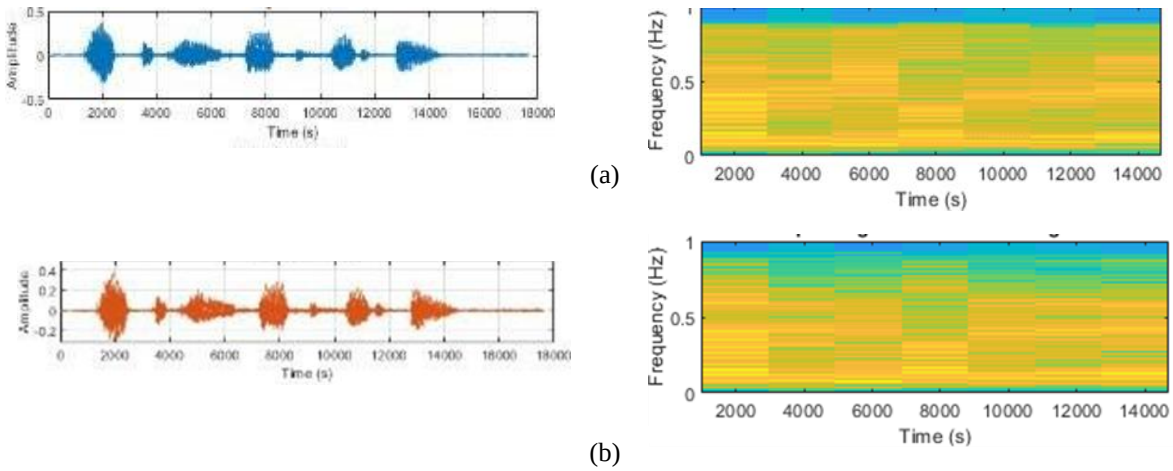
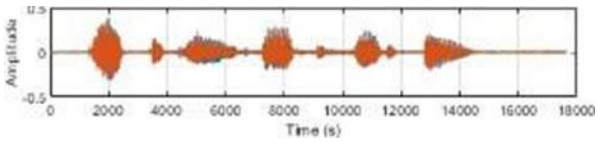


Fig.6. (a) Waveform and Spectrogram of Noisy speech (Street Noise with SNR 10dB) (b) Waveform and Spectrogram of Denoised speech (Street Noise with SNR 10dB) (c) Waveform of Difference Between Noisy and Denoised speech.

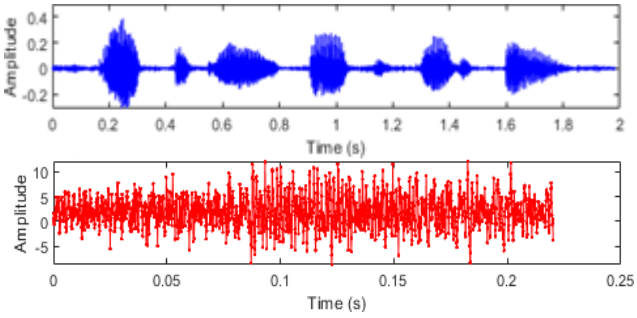




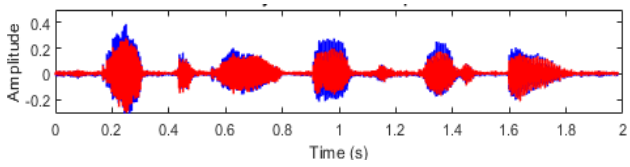
(c)

Fig.7. (a) Waveform and Spectrogram of Noisy speech (Street Noise with SNR 15dB) (b) Waveform and Spectrogram of Denoised speech (Street Noise with SNR 15dB) (c) Waveform of Difference Between Noisy and Denoised speech.

The waveform obtained from the denoised speech serve as the progressed signal of the discrete wavelet transform algorithm. The denoised signal signifies the depletion of the noisy part existed in the signal. The denoised signal is further given as input to LSTM Network. The LSTM Network is trained and then the enhanced speech signal is obtained. The waveform for Enhanced speech obtained from LSTM network which is visualized in Fig.8. and Fig.9.

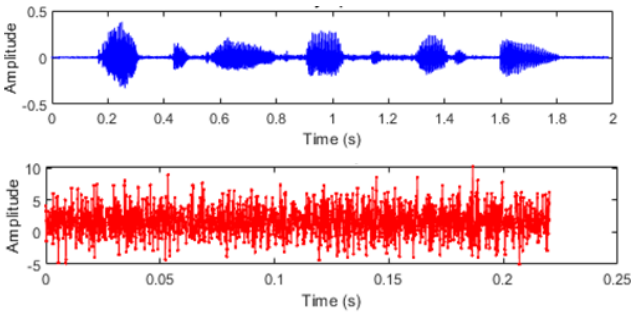


(a)

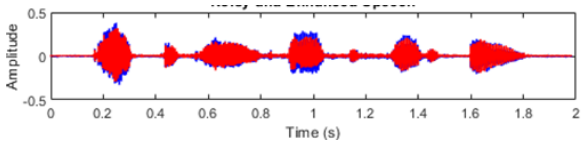


(b)

Fig.8. (a) Waveform of Denoised Speech(DWT) and Enhanced Speech (LSTM)(Babble Noise with SNR 15dB) (b) Waveform of Denoised speech (Babble Noise with SNR 15dB) (c) Waveform for Difference Between Denoised and Enhanced speech.



(a)



(b)

Fig.9. (a) Waveform of Denoised Speech(DWT) and Enhanced Speech (LSTM)(Street Noise with SNR 15dB) (b) Waveform of Denoised speech (Street Noise with SNR 15dB) (c) Waveform for Difference Between Denoised and Enhanced speech.

Table 2: Objective evaluation of DWT-LSTM Model at different noise levels

NOISE TYPE	INPUT SNR (dB)	OUTPUT SNR (dB)		
		DWT	LSTM	DWT-LSTM
Train Noise	5	15.56	14.34	18.50
	10	17.87	16.78	22.33
	15	19.90	19.43	27.86
Restaurant Noise	5	16.67	15.67	17.56
	10	18.34	18.32	24.69
	15	19.87	19.30	28.78
Babble Noise	5	17.65	13.43	16.43
	10	18.54	14.98	23.39
	15	20.00	16.40	26.50
Street Noise	5	17.98	15.77	20.21
	10	18.02	18.43	24.65
	15	19.43	20.21	**30.46**

Thus, the system which is developed achieves a good performance of different SNR levels at different noises. Using DWT- LSTM, the output SNR obtained is 30.46 dB for street noise with 15 dB noise.

4. CONCLUSION

This work proposed an innovative approach for improving corrupted speech signals through signal processing. Many filtering techniques were used to achieve Speech Enhancement such as adaptive filtering, spectral subtraction, wiener filtering, and neural network algorithms such as CNN, RNN, LSTM, etc., Our experimental results shows that the proposed Speech Enhancement system using DWT-LSTM outperforms the existing algorithms in terms of output, 30.6 dB for the street noise of SNR 15 dB, also with waveform and spectrogram respresentation respectively.

References

1. Wang, Li & Zheng, Weiguang & Ma, Xiaojun & Lin, Shiming. ”Denoising Speech Based on Deep Learning and Wavelet Decomposition. Scientific Programming”,Scientific Programming, vol. 2021. https://doi.org/10.1155/2021/8677043
2. Q. Zhang, X. Qian, Z. Ni, A. Nicolson, E. Ambikairajah and H. Li, "A Time-Frequency Attention Module for Neural Speech Enhancement," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 31, pp. 462-475, 2023 https://doi.org/10.1109/TASLP.2022.3225649
3. Sun, Lei & Du, Jun & Dai, Li-Rong & Lee, Chin-Hui,“Multiple-target deep learning for LSTM-RNN based speech enhancement”. https://doi.org/10.1109/HSCMA.2017.7895577
4. Dharshini, R. R., Selvi, R. S., Suresh, G. R., & Raja, S. K. S.,”Enhancement Of Speech Intelligibility And Quality In Hearing Aid Using Fast Adaptive Kalman Filter Algorithm”,Vol no 9,2017
5. Garg A, "Speech enhancement using long short term memory with trained speech features and adaptive wiener filter,”Multimed Tools Appl. 2023;82(3):3647-3675. https://doi.org/10.1007/s11042-022-13302-3
6. Strake, M., Defraene, B., Fluyt, K. et al. “Speech enhancement by LSTM- based noise suppression followedby CNN-based speech restoration”. EURASIP J. Adv. Signal Process,2020. https://doi.org/10.1186/s13634-020-00707-1
7. Senthamizh Selvi R , Sathish Kumar P , Sri Krishna R , Surya Rao S “Speech Enhancement using Adaptive Filtering with Different Window Functions and Overlapping Sizes” Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(13), 1886–1894,2021 https://doi.org/10.17762/turcomat.v12i13.8841.
8. Gutiérrez-Muñoz, M.; Coto-Jiménez, “M. An Experimental Study on Speech Enhancement Based on a Combination of Wavelets and Deep Learning”. Computation 2022, 10,102. https://doi.org/10.3390/computation10060102.
9. V.Shruthi, R.Senthamizhselvi , G.R.Suresh “Speech intelligibility prediction and near end listening enhancement for mobile appliciation”,IJARSE, vol.no.07,Issue No.02 , 2018
10. Vinay, N A, Bharathi, S H “Design and Development of Algorithm for Recognition of speech Disfluency using Enhanced Correlative method”, 2021
11. M. Liu, Y. Wang, J. Wang, J. Wang and X. Xie, "Speech Enhancement Method Based On LSTM Neural Network for Speech Recognition",14th IEEE International Conference on Signal Processing (ICSP), Beijing, China,pp. 245- 249, 2018 https://doi.org/10.1109/ICSP.2018.8652331
12. X. Li and R. Horaud, "Multichannel Speech Enhancement Based On Time-Frequency Masking Using Subband Long Short-Term Memory", 2019 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), New Paltz, NY, USA, 2019, pp. 298-302. https://doi.org/10.1109/WASPAA.2019.8937218
13. Yuliani, Asri & Amri, Mohd & Suryawati, Endang & Ramdan, Ade & Pardede, Hilman. ,”Speech Enhancement Using Deep Learning Methods: A Review”. Jurnal Elektronika dan Telekomunikasi vol. 21, no. 1, pp. 19- 26,2021 https://doi.org/10.14203/jet.v21.19-26.
14. Haripriya, R. & Mathew, Lani & Gopakumar, K, "Performance evaluation of dwt based speech enhancement". https://doi.org/442-446. 10.1109/NETACT.2017.8076812
15. O. Jane, E. Elbaşi, “Hybrid non-blind watermarking based on DWT and SVD”, Journal of Applied Research and Technology, vol. 12, issue 4, pp. 750- 761, 2014. https://doi.org/10.1016/S1665-6423(14)70091-4
16. N. Al Bassam, V. Ramachandran, S. E. Parameswaran, “ Wavelet Theory and Application in Communication and Signal Processing”,Chapter,IntechOpen,2021. https://doi.org/10.5772/intechopen.95047
17. Elfouly, F. H., Mahmoud, M. I., Dessouky, M. I. M., & Deyab, S, "Comparison Between Haar

And Daubechies Wavelet Transformions On Fpga Technology",International Journal of Computing, 6(3), 23-29. https://doi.org/10.47839/ijc.6.3.447.

18. Tan, K., Wang, D, A "Convolutional Recurrent Neural Network for Real- Time Speech Enhancement",Proc Interspeech 2018, 3229-3233 https://doi.org/10.21437/Interspeech.2018-1405
19. Shanthini, V., Senthamizh, S. R., & Suresh, G. R, " Hybridization of model based and template based techniques for speech enhancement. Advances in Natural and Applied Sciences, 10(9 SE), 336+,2016.
20. Weninger, Felix & Erdogan, Hakan & Watanabe, Shinji & Vincent, Emmanuel & Le Roux, Jonathan & Hershey, John & Schuller, Björn, "Speech Enhancement with LSTM Recurrent Neural Networks and its Application to Noise-Robust ASR", 2015 https://doi.org/10.1007/978-3-319-22482-4_11
21. Y. Wang, A. Narayanan and D. Wang, "On Training Targets for Supervised Speech Separation," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 22, no. 12, pp. 1849-1858, 2014 https://doi.org/10.1109/TASLP.2014.2352935
22. R. Senthamizh Selvi, G.R. Suresh, S. Kanaga Suba Raja, "Speech Enhancement Using Harmonic-Model with Multichannel Wiener Filter "Jour of Adv Research in Dynamical & Control Systems, Vol. 9, No. 2, 2017