Speech Enhancement for Noise Reductions Using Hybrid Signal Subspace

D. R. Sandeep[1]; K. Jahnavi[2]; B. L. Gagana[2];
B. S. Devika[2]; A. V. S. S. S. Vyshnavi[2]; A. M. Gowri[2]

[1]Senior Assistant Professor, Department of Electronics and Communication Engineering
Sri Vasavi Engineering College
Tadepalligudem, Andhra Pradesh, India

[2]Department of Electronics and Communication Engineering
Sri Vasavi Engineering College
Tadepalligudem, Andhra Pradesh, India

Publication Date: 2026/05/23

Abstract: Speech signals captured in real time are often messed up by background noise. This noise makes hard to understand and hear clearly. Noise can come from places like traffic, machines or environmental disturbances. It makes speech processing tasks more difficult. Speech enhancement techniques try to remove noise while keeping the important parts of the original speech signal.

In this paper propose a method that combines classical signal processing techniques with subspace-based noise reduction methods. The system examines speech signals frame by frame, breaks down the covariance matrix and projects the signal onto a subspace to reduce noise. Also applying filtering and smoothing operations to improve sound quality.

The system is tested using metrics like Signal-to-Noise Ratio (SNR), Perceptual Evaluation of Speech Quality (PESQ), Short-Time Objective Intelligibility (STOI) and Mean Opinion Score (MOS). The MATLAB tool was used to verify those metrics and the results obtained from the proposed method shows significant improvement in speech quality and intelligibility in noisy conditions.

Keywords: *Signal-to-Noise Ratio, Perceptual Evaluation of Speech Quality, Short-Time Objective Intelligibility and Mean Opinion Score (MOS), Signal Subspace Method, Signal Enhancement.*

How to Cite: D. R. Sandeep; K. Jahnavi; B. L. Gagana; B. S. Devika; A. V. S. S. S. Vyshnavi; A. M. Gowri (2026) Speech Enhancement for Noise Reductions Using Hybrid Signal Subspace. *International Journal of Innovative Science and Research Technology*, 11(5), 1311-1315. https://doi.org/10.38124/ijisrt/26may223

I. INTRODUCTION

Speech communication is a part of many modern technologies, including mobile phones, voice assistants, video conferencing and speech recognition systems. However, in real-life situations speech signals are often mixed with background noise. This noise can come from traffic, machines or other speakers. It degrades speech quality. Reduces intelligibility making it hard for listeners or automated systems to understand spoken information.

Speech Enhancement techniques aim to improve speech quality by reducing noise while keeping the important parts of the speech signal. Over time, researchers have introduced various techniques to solve this problem. Traditional signal processing approaches like subtraction and Wiener filtering are widely used because they are easy to compute and implement.

Recently, the use of machine learning and deep learning techniques has increased significantly to improve speech enhancement performance. These techniques can learn patterns from large datasets and handle various types of noise. However, they often require training datasets and significant computational resources to work.

To address these limitations, this study introduces an approach that integrates signal processing methods with advanced enhancement techniques to achieve improved performance. In this work, covariance matrix analysis and

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may223

eigenvalue decomposition are used to separate speech and noise components

II. LITERATURE REVIEW

Speech Enhancement has gained considerable importance in research area because speech signals are often mixed up by different types of background noise in real-world environments. Various techniques are proposed by many researchers to improve speech quality and intelligibility.

Ioannides and Rallis [1] proposed a real-time speech enhancement method based on spectral subtraction. The approach estimates background noise using minimum statistics and removes it from the speech spectrum. A spectral floor is applied to avoid musical noise and signal distortion. The method is computationally efficient and suitable for real-time systems. Experimental results show improved speech clarity in noisy environments.

Yu, De Ocampo, and Hernandez [7] introduced a speech noise reduction technique using intelligent spectral gain selection. The method adaptively adjusts the gain of frequency components depending on noise levels. This helps suppress noise while preserving important speech information. The approach improves speech intelligibility in different noise conditions. Their results demonstrate effective noise reduction with minimal speech distortion.

Wu and Hung [9] improved speech enhancement models using discrete wavelet transform (DWT) sub-band features. Wavelet analysis provides detailed time–frequency representation of speech signals. These features were integrated with the Adaptive FullSubNet model for better performance. The method improves noise suppression capability in complex environments. Experimental evaluation shows enhanced speech quality compared to traditional models.

Poornimadarshini [12] proposed a hybrid spectral–temporal deep learning model for audio signal enhancement. The model combines spectral and temporal information to effectively remove background noise. It is designed to handle different types of noisy environments. The hybrid approach improves robustness and speech clarity. The results show better performance compared to conventional enhancement methods.

Rosenbaum, Winbrand, Cohen, and Cohen [16] developed a deep learning framework for real-time speech enhancement and dereverberation. The model focuses on reducing both background noise and reverberation. It enhances speech intelligibility in real-world acoustic environments. The framework is optimized for efficient processing and practical applications. Experimental results demonstrate improved speech quality and system performance.

III. PROPOSED METHODOLOGY

The proposed speech enhancement system follows a multi-stage processing approach designed to reduce noise while preserving speech characteristics.

The noisy speech signal is commonly represented as the combination of the original speech signal and an additive noise component:

$$y(n) = s(n) + d(n)$$

where
$y(n)$ represents the noisy speech signal,
$s(n)$ represents the clean speech signal and
$d(n)$ represents the noise signal.
The objective of a speech enhancement system is to recover the original clean speech signal from the observed noisy signal.

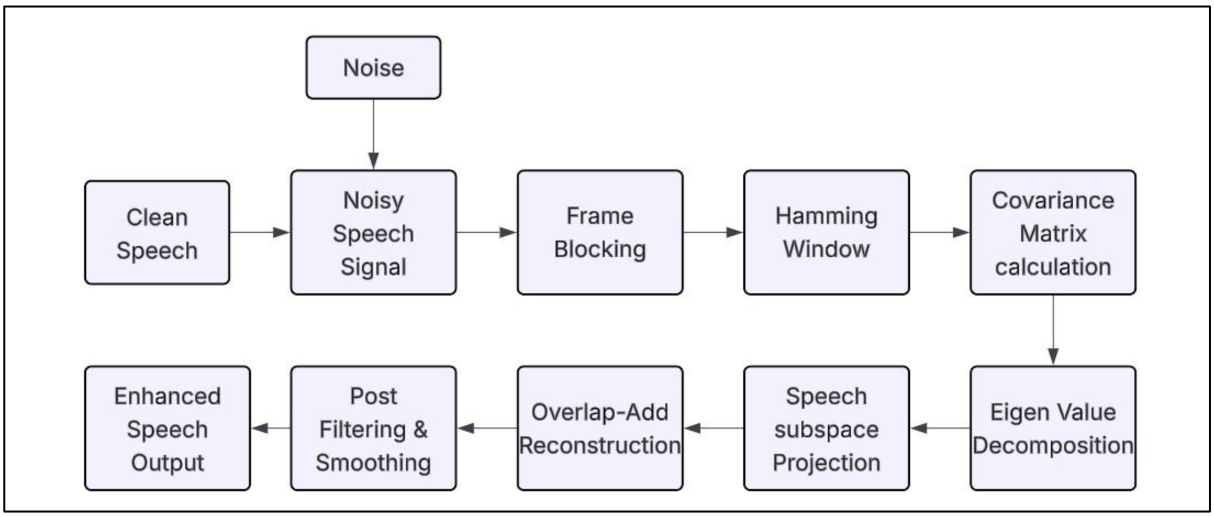

Fig. 1. System Block Diagram

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may223

IV. SOFTWARE IMPLEMENTATION

The implementation of the proposed method was carried out using MATLAB software. The following steps describe the processing procedure.

➢ *Algorithm*:

Step 1: Load clean speech and noise audio files.
Step 2: Convert audio signals to a common format and sampling rate.
Step 3: Generate a speech signal by adding noise to the clean speech.
Step 4: Divide the signal into overlapping frames.
Step 5: Apply a Hamming window to each frame.
Step 6: Compute the covariance matrix of speech frames.
Step 7: Perform eigenvalue decomposition.
Step 8: Select dominant eigenvectors representing the speech subspace.
Step 9: Project noisy frames onto the speech subspace.
Step 10: Reconstruct enhanced speech using overlap-add technique.
Step 11: Apply smoothing filter to improve quality.
Step 12: Normalize signal energy.
Step 13: Evaluate performance using SNR, PESQ, STOI, and MOS metrics.

V. RESULTS AND DISCUSSION

The performance of the proposed hybrid speech enhancement system was evaluated using speech signals contaminated with background noise. The experiments were conducted in the MATLAB environment using audio signals sampled at a rate of 16 kHz. In the experiment, the clean speech signal was combined with a noise signal to generate a noisy speech input. The proposed algorithm was then applied to reduce the noise components and reconstruct the enhanced speech signal.

To assess the performance of the proposed method, evaluation metrics such as Signal-to-Noise Ratio (SNR), Perceptual Evaluation of Speech Quality (PESQ), Short-Time Objective Intelligibility (STOI), and Mean Opinion Score (MOS) are used. These metrics provide quantitative measurements of speech quality and intelligibility before and after noise reduction.

The results confirm that the proposed technique is capable of reducing background noise effectively while preserving the key features of the speech signal. The SNR improvement of approximately 2.8 dB indicates effective noise suppression.

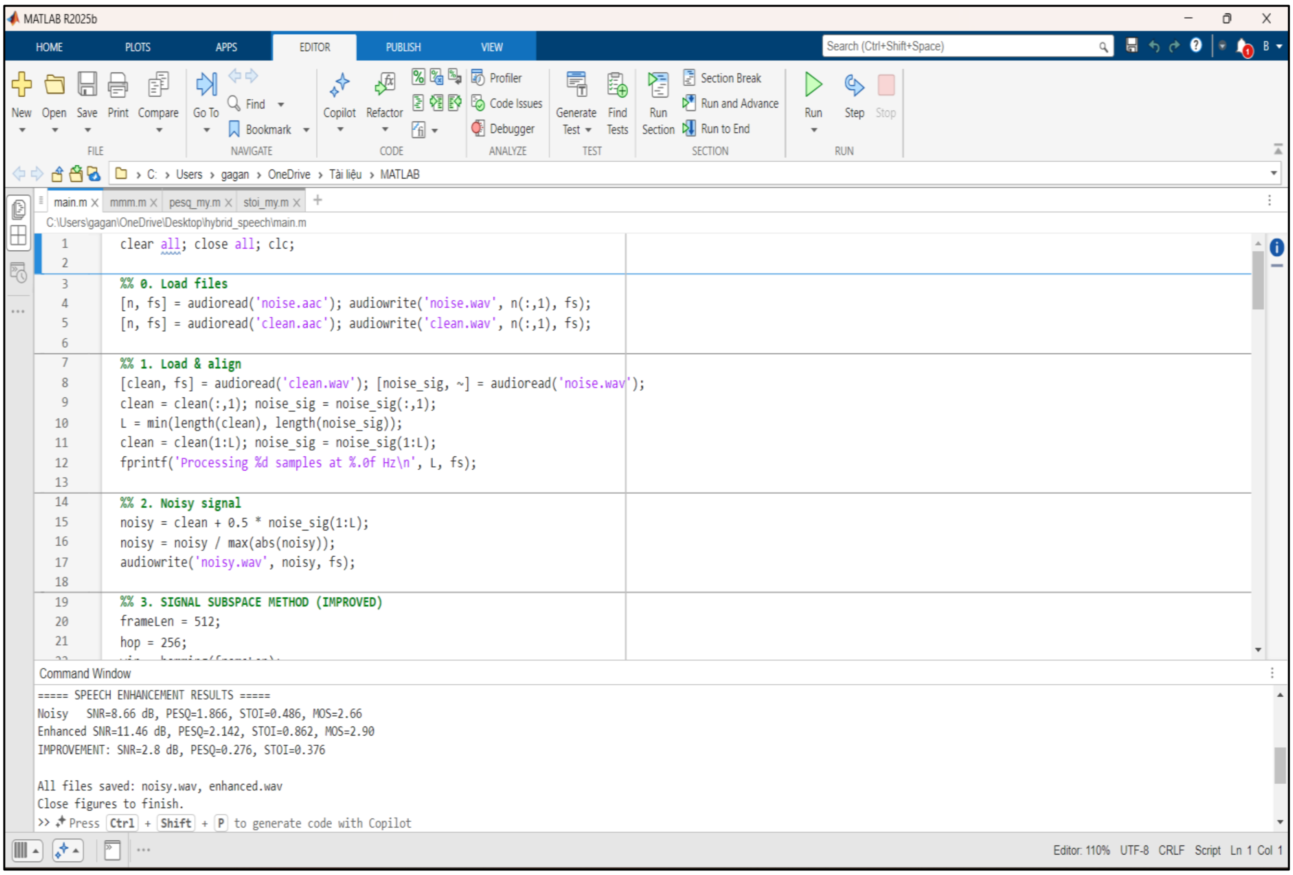

Fig. 2. Performance Evaluation Matrics Obtained From MATLAB

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may223

Noisy
SNR=8.66 dB, PESQ=1.866, STOI=0.486, MOS=2.66

Enhanced
SNR=11.46 dB, PESQ=2.142, STOI=0.862, MOS=2.90

Improvement
SNR=2.8 dB, PESQ=0.276, STOI=0.376

The increase in the PESQ score reflects improved perceptual speech quality after enhancement. A major improvement can be observed in the STOI value, which increased from 0.486 to 0.862. This indicates that the proposed method greatly improves speech intelligibility, making the enhanced speech signal easier to understand.

The MOS value also increased from 2.66 to 2.90, indicating an overall improvement in perceived speech quality. These results confirm that the proposed method effectively reduces background noise while preserving important speech components.

In addition to numerical evaluation, waveform analysis was also performed. The waveform of the noisy speech signal showed strong distortion caused by background noise. After applying the proposed algorithm, the enhanced speech waveform exhibited reduced noise components and clearer speech patterns. Spectrogram analysis further confirmed that noise energy was significantly reduced across multiple frequency bands.

Overall, the experimental results demonstrate that the proposed hybrid speech enhancement approach provides effective noise reduction and improves speech clarity in noisy environments.

➢ *Speech Enhancement Waveforms :*

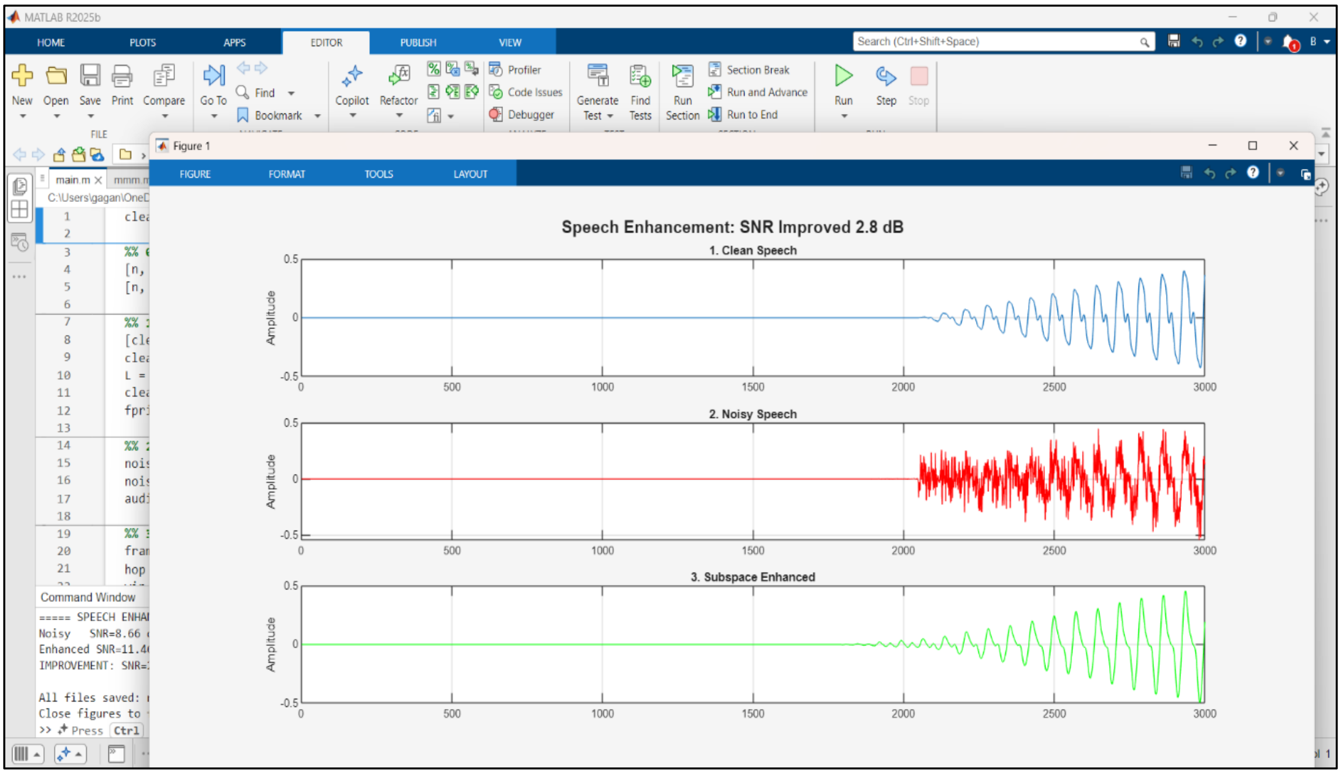

Fig. 3. Speech Enhancement waveforms

VI. CONCLUSION

In this work, a speech enhancement technique was implemented to reduce noise from corrupted speech signals. The algorithm processes the noisy speech signal and generates an enhanced output with improved speech quality.

The results obtained from waveform analysis and performance metrics show that the proposed metrics show that the proposed system effectively reduces background noise while preserving important speech information. Therefore, the developed speech enhancement system can be useful in applications that require clear speech communication.

FUTURE SCOPE

Although the proposed hybrid speech enhancement method shows promising results in reducing background noise and improving speech quality, there are several directions in which the work can be further extended.

In this work advanced deep learning models such, as Convolutional Neural Networks (CNNs) can be explored. Recurrent Neural Networks can be combined with the signal processing framework to reduce noise in changing environments. These models can learn relationships between clean and noisy speech signals and work better in non-stationary noise conditions. The proposed algorithm can also

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may223

be implemented in real-time systems. The current implementation is done using MATLAB for analysis and testing.

Future research can focus on implementing the algorithm on embedded platforms like DSP processors, FPGA boards or microcontrollers to enable real-time speech enhancement in applications. Additionally, the system can be tested with types of noise environments, such as traffic noise, crowd noise and industrial noise to evaluate its robustness under various real-world conditions. Using more diverse speech datasets may further improve the reliability of the enhancement system.

Further optimization techniques can be applied to reduce complexity and improve processing speed making the algorithm more suitable for mobile devices and IoT-based speech communication systems.

REFERENCES

[1]. G. Ioannides and V. Rallis, "Real-Time Speech Enhancement Using Spectral Subtraction with Minimum Statistics and Spectral Floor," arXiv, 2023.

[2]. P. C. Loizou, "Speech Enhancement: Theory and Practice," CRC Press, Boca Raton, FL, USA, 2013.

[3]. J. S. Lim and A. V. Oppenheim, "All-Pole Modeling of Degraded Speech," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 26, no. 3, pp. 197–210, 1978.

[4]. R. Martin, "Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics," IEEE Transactions on Speech and Audio Processing, vol. 9, no. 5, pp. 504–512, 2001.

[5]. A. Pandey and D. Wang, "A New Framework for Supervised Speech Enhancement in the Time Domain," IEEE/ACM Transactions on Audio, Speech and Language Processing, vol. 28, pp. 282–295, 2020.

[6]. Y. Ephraim and D. Malah, "Speech Enhancement Using a Minimum Mean-Square Error Short-Time Spectral Amplitude Estimator," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 32, no. 6, pp. 1109–1121, 1984.

[7]. H. Yu, A. L. De Ocampo, and R. Hernandez, "Speech Noise Reduction via Intelligent Spectral Gain Selection and Modification," International Journal of Intelligent Systems and Applications in Engineering, 2024

[8]. S. Boll, "Suppression of Acoustic Noise in Speech Using Spectral Subtraction," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 27, no. 2, pp. 113–120, 1979.

[9]. Z.-T. Wu and J.-W. Hung, "Improving the Speech Enhancement Model with Discrete Wavelet Transform Sub-Band Features in Adaptive FullSubNet," Electronics, vol. 14, 2025.

[10]. I. Cohen and B. Berdugo, "Speech Enhancement for Non-Stationary Noise Environments," Signal Processing, vol. 81, no. 11, pp. 2403–2418, 2001.

[11]. P. Scalart and J. V. Filho, "Speech Enhancement Based on a Priori Signal-to-Noise Estimation," IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 629–632, 1996.

[12]. S. Poornimadarshini, "Robust Audio Signal Enhancement Using Hybrid Spectral–Temporal Deep Learning Models in Noisy Environments," National Institute of STEM Research, India, 2025.

[13]. Y. Hu and P. C. Loizou, "Evaluation of Objective Quality Measures for Speech Enhancement," IEEE Transactions on Audio, Speech and Language Processing, vol. 16, no. 1, pp. 229–238, 2008.

[14]. S. F. Boll and J. H. Porter, "Spectral Subtraction for Speech Enhancement," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 27, no. 2, pp. 113–120, 1979.

[15]. 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.

[16]. T. Rosenbaum, E. Winbrand, O. Cohen, and I. Cohen, "Deep Learning Framework for Efficient Real-Time Speech Enhancement and Dereverberation," Sensors, vol. 25, 2025.

[17]. K. Tan and D. Wang, "A Convolutional Recurrent Neural Network for Real-Time Speech Enhancement," Interspeech Conference, pp. 3229–3233, 2018.

[18]. Y. Xu, J. Du, L. Dai, and C. Lee, "A Regression Approach to Speech Enhancement Based on Deep Neural Networks," IEEE Transactions on Audio, Speech and Language Processing, vol. 23, no. 1, pp. 7–19, 2015

[19]. C. H. Taal, R. C. Hendriks, R. Heusdens, and J. Jensen, "An Algorithm for Intelligibility Prediction of Time–Frequency Weighted Noisy Speech," IEEE Transactions on Audio, Speech and Language Processing, vol. 19, no. 7, pp. 2125–2136, 2011.

[20]. J. Benesty, S. Makino, and J. Chen, "Speech Enhancement," Springer Handbook of Speech Processing, Springer, Berlin, Germany, 2008.