DOI: 10.24996/ijs.2024.65.10.38

ISSN: 0067-2904

Review: Noise Reduction Techniques for Enhancing Speech

Ruqaya Jamal Nasir*, Husam Ali Abdulmohsin
Department of Computer Science, College of Science, University of Baghdad, Baghdad, Iraq

Received: 28/4/2023 Accepted: 20/8/2023 Published: 30/10/2024

Abstract
Speech is a signal produced by humans to interact and communicate. Different information is gained from speech signals, such as the language being spoken, emotion, gender, speaker identification, and other information. Speech signals are exposed to different noises, which can be generated at the beginning of the speech or during the transmission. Due to this problem, noise reduction processes are an interesting field in different communication application systems that cultivate the intelligibility and quality of speech signals. It refers to removing or reducing the background noise in order to obtain an improved quality of the original speech signal without distorting the original (clean) signal. This paper reviews the state-of-the-art research, reviewing different speech enhancement filters and algorithms and comparing their performance to reach a conclusion about which is the best filter or the most effective one based on the kind of noise that was used and the most difficult noise to remove from the signal.

Keywords: Adaptive filtering, Kalman filter, spectral subtraction, wavelet de-noising, wiener filter.

مراجعة: تقنيات تحسين الكلام

رقية جمال ناصر*, حسام علي عبد المحسن
قسم علوم الحاسوب, كلية العلوم, جامعة بغداد, بغداد, العراق

الخلاصة
الكلام هو إشارة تنتج من الإنسان للتفاعل والتواصل. يتم الحصول على معلومات مختلفة من إشارات الكلام ، مثل اللغة التي يتم التحدث بها والعاطفة والجنس وتحديد المتحدث وغير ذلك من المعلومات. تعرضت إشارات الكلام في الماضي إلى ضوضاء مختلفة يمكن أن تتولد من بداية الكلام أو أثناء الإرسال. بسبب هذه المشكلة ، تعد عمليات الحد من الضوضاء مجالًا مثيرًا للاهتمام في أنظمة تطبيقات الاتصالات المختلفة التي تزرع وضوح وجودة إشارة الكلام. يشير إلى إزالة ضوضاء الخلفية أو تقليلها من أجل الحصول على جودة محسنة لإشارة الكلام الأصلية دون تشويه الإشارة الأصلية (النظيفة). يستعرض هذا البحث أحدث الأبحاث ، حيث يقوم بمراجعة مرشحات وخوارزميات تحسين الكلام المختلفة ، ومقارنة أدائها للوصول إلى استنتاج حول أفضل مرشح أو الأكثر فعالية بناء على نوع الضوضاء التي تم استعمالها وأصعب ضوضاء لإزالتها من الإشارة.

*Email: roqia.nasser2101m@sc.uobaghdad.edu.iq

1. Introduction

Speech is a type of communication used to convey ideas. The mechanism for generating the human voice can be subdivided into three parts, such as those presented in Fig. 1. The lungs, the vocal folds within the larynx (voice box), and the articulators. The vocal tract of a human person is used to produce a variety of sounds, including talking, singing, laughing, sobbing, screaming, shouting, humming, and yelling since the voice can be affected by emotion [1]. The vocal folds (vocal cords), which are the main sound source in human sound production, are a special portion of the human voice frequency [2], where the articulators in the mouth and nose, which are responsible for articulation, will influence the airflow in the lungs [3],

The audible frequency range for human beings is from 20 Hz to 20 KHz. Audio signal processing often suffers from noise.

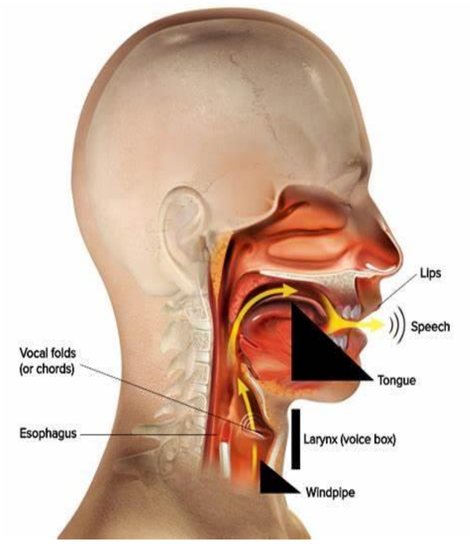

Figure 1: Human speech production [4]

Nowadays, humans are able to interact with computer hardware and others in many aspects of life. Speech processing is widely used in numerous applications, such as teleconferencing systems, speech coding for communications, speech recognition, mobile speech communication, biomedical signal processing, hearing aids, ATM machines, and others. Such applications exist in areas where there is interfering background noise, such as a motor vehicle passing [5]. These interference noises degrade the quality of the original speech in such a way that it does not remain clear anymore.

Noise reduction is a hot research area in signal processing and remains a challenging issue because, in most cases, only noisy speech is available [6].

The most common type of noise that causes the degradation of speech's quality and intelligibility is background noise, which can be stationary or non-stationary and is assumed to be uncorrelated and additive to the speech signal [7].

Due to the varying characteristics of noise over time, it is hard to enhance speech in a noisy environment. Till now, removing noise from noisy speech has been a grueling issue because spectral parcels of non-stationary noise are veritably delicate to estimate and prognosticate. Noise calculation is a serious issue when noise power is greater than speech power because speech content may be removed when treating it as noise.

This survey covers all the work that was implemented. It will be useful to add the statistical and mathematical approaches implemented in noise reduction.

In this paper, we will examine the studies done by other researchers on noise reduction in speech signals and help other researchers make decisions about which filter or method will be used in their work to get the best noise reduction and the fittest filter based on the type of noise.

This paper is structured as follows: The problem is introduced, and a general overview is provided in the introduction in Section 1. Some of the related work of other researchers is explained to the reader, and different researchers' work in different years is compared in the related work and state-of-the-art work in Section 2. A brief definition of noise in our lives is in Section 3. We discuss the most familiar and fundamental strategies for noise reduction in Section 4. Then, in Section 5, the types of filters for noise reduction are classified. Discuss the approaches and filters of other researchers in Section 6, and Section 7 will offer a conclusion.

2. Related Work and State of Art Works

Noise reduction is an attractive field for researchers to explore. Since 1960 until now, many researchers have conducted research and improved noise reduction techniques.

In 2014, the authors proposed [8] (Noise Cancellation in Speech Signal Processing: A Review). They categorize existing noise cancellation schemes and thoroughly investigate various suggestions in each category in order to demonstrate the limitations of existing techniques as well as their effective contributions. Several techniques for filtering noise from a speech waveform have been investigated. They found that the recursive least squares (RLS) algorithm produces the maximum SNR and outperforms the least mean square (LMS) for the lower-order FIR adaptive filter. But for the finite impulse response (FIR) filter Taps, LMS converges more quickly than RLS. By establishing the FIR tap weight, the ideal Mu (LMS) and Lambda (RLS) values have been discovered. Cancellation of acoustic noise: the best method for reducing background noise is adaptive noise cancellation (ANC). The performance of conventional wideband ANC algorithms rapidly declines as the noise's bandwidth and center frequency rise, but they perform better in lower frequency bands.

The authors in 2014 proposed [9] (A Survey on Statistical Based Single Channel Speech Enhancement Techniques). They contrast various estimators (classical and Bayesian estimators). They compare various estimators. The difficulties and possibilities of improving speech are also covered, which facilitates selecting the most effective statistically-based speech enhancement strategy. Techniques based on statistics are described, along with their advantages and disadvantages. It is explained how classical and Bayesian estimators compare. In this work, the fixed window technique's drawbacks are examined. The study of single-channel speech augmentation approaches includes significant and important distinctions between causal and non-causal estimators.

In 2018, the authors proposed [10] (A Review on Various Speech Enhancement Techniques). They reviewed various speech enhancement techniques. They mainly focused on noise removal in speech signals and discussed various single- and multi-sensor speech augmentation techniques. Since it is impossible to totally avoid noise, the authors concentrate on reduction using a variety of criteria. Noise cancellation and echo suppression are also crucial components of speech enhancement. Because the Kalman filter is recursive, it is one of the most effective ways to enhance a signal.

In 2018, authors proposed [11] (A Survey on Techniques for Enhancing Speech). They discuss many such techniques, including the benefits and drawbacks of each. Review the study on alternative machine learning methods, including neural networks, deep networks, convolutional networks, and optimization methods used to improve speech, that was done by other researchers. And neural networks have proven to be the most effective technique. Following simple NN, deep neural networks (DNN) followed, which produced better results but had poor real-world generalization when it came across noise and speech signals that hadn't been visible to it during the training phase. Then came the era of the Convolutional Neural Network (CNN), which has now established itself as a trustworthy instrument for extrapolating real-world noise cancellation issues. During the training phase, it is capable of handling all types of noise signals, whether they are visible to it or not. The effectiveness of neural networks has been demonstrated.

And the authors, in 2022, proposed [12] (Review Paper on Noise Cancellation using Adaptive Filters). The most crucial technique is adaptive noise cancellation (ANC). ANC employs adaptive filters to assess continuously changing real-time data. Numerous algorithms are used in ANC in order to cancel out the noise. The Least Mean Square (LMS) algorithm was the author's main focus. To find the available literature on adaptive filtering in noise reduction using the LMS adaptive algorithm, a thorough review has been conducted. LMS is easy to implement, has a low level of computing complexity, and has a higher rate of convergence. It has been demonstrated that the LMS algorithm performs well when applied to a noise cancellation problem.

Table 1 illustrates the state-of-the-art work related to noise reduction techniques published from 1976 until now.

Table 1 State-of-the-art noise reduction publications

Ref.	Year	Filter Approach	Type of noise	Conclusion	Performance measurement units
[13]	1976	Adaptive Filter Time—Variant digital comb filter	Speakers, Background noise	The authors show how the adaptive system can respond to the test input signal. To demonstrate the degree of intended speaker distortion produced by the systems in this scenario, the filtering system is set up to function as an identity system.	-
[14]	1979	Spectral subtraction, modified spectral subtraction	Broadband noise	To sum up, he subtracts an overestimate of the noise spectrum and keeps the resulting spectral components from falling below a spectral floor, which is the key distinction between their implementation and the classic spectral subtraction approach. Their application of the spectral noise removal method allows for a significant reduction in background noise with little impact on speech comprehension.	According to tests, the intelligibility of the improved speech is equal to that of the unprocessed signal at SNR = +5 dB. The enhancement procedure did not result in any loss of understanding.
[15]	1982	Nonlinear multiband envelope filtering, Logarithmic filtering log—lin filtering	Stationary white noise	When power signals are not linear It was discovered that preprocessing before adding noise improved significantly when filters with a logarithmic or combination log/lin characteristic and high slow level variation compression were used. Better preprocessing for hearing loss due to sensorineural impairment may result from this as well.	The unprocessed and reprocessed signals (including noise) had their intelligibility assessed and compared, the latter using two different FIR filters: one without expansion (0.05 at 0 Hz, +5 db from 2 to 16 Hz, 1 at > 18 Hz) and one with it (0.05 at 0 Hz, +5 db from 2 to 16 Hz, 1 at > 18 Hz). The results were intelligence scores of 64.5, 87.2, and 95.5, respectively.

<table>
<tr>
<td>[16]</td>
<td>1999</td>
<td>Approach is based on the introduction of an auditory model in a subtractive-type enhancement process, Power spectral subtraction, modified spectral subtraction</td>
<td>White Gaussian noise, Car noise, Speechlike noise (long term average Speech spectrum), Aircraft cockpit, Helicopter cockpit, Factory</td>
<td>Background noise is minimized, and residual noise is less organized than with traditional approaches, yet speech distortion remains acceptable. The following are the proposed algorithm's primary benefits:
1. It is efficient in terms of computing.
2. A criterion that is much more correlated with speech perception than the SNR is used to alter the subtraction parameters.
3. It provides the option to alter the trade-off between speech distortion, residual noise, and noise reduction.</td>
<td>proposed algorithm
<table>
<tr><th>Noisy type</th><th colspan="3">Noisy signal</th><th colspan="3">Enhanced signal</th></tr>
<tr><td></td><td>SNR</td><td>IS</td><td>AI</td><td>SNR</td><td>IS</td><td>AI</td></tr>
<tr><td>6 db</td><td>0 db</td><td>3.6</td><td>0.17</td><td>6.5 db</td><td>2.2</td><td>0.38</td></tr>
<tr><td>9 db</td><td>0 db</td><td>2.6</td><td>0.21</td><td>7.7 db</td><td>1.6</td><td>0.41</td></tr>
</table>
</td>
</tr>
<tr>
<td>[17]</td>
<td>2003</td>
<td>Adaptive and Non-adaptive filtering techniques: Beamforming, Adaptive noise cancellation (ANC), Spectral modification</td>
<td>White Gaussian noise, Directional noise condition, Stationary noise</td>
<td>The Wiener filter offers the largest reduction since the noise is stationary. While the NLMS algorithm converges more quickly, the LMS approach takes longer. Both the LMS and NLMS algorithms are outperformed by the RLS algorithm. Spectral modification techniques like spectral subtraction and Wiener filtering are more frequently used to accomplish noise reduction. For comparison, the performance of the ideal Wiener filter is also provided. As can be seen, for most frequencies, the parametric Wiener filter eliminates noise more effectively than the ideal Wiener filter. Compared to a beamformer or an ANC technique, spectral modification can decrease noise more effectively overall.</td>
<td>-</td>
</tr>
<tr>
<td>[18]</td>
<td>2007</td>
<td>Minimum-Mean-Square-Error (MMSE)</td>
<td>White Gaussian Noise (WGN), M109 tank noise, F16 cockpit noise</td>
<td>Results indicate that the proposed algorithm yields better performance than several recently proposed methods.</td>
<td>
<table>
<tr><th rowspan="2">Methods</th><th colspan="2">Gaussian white noise (dB)</th></tr>
<tr><td>5 SNR(db)</td><td>10 SNR(db)</td></tr>
<tr><td>LapMMSE</td><td>15.1721</td><td>13.1893</td></tr>
<tr><td>GamMMSE</td><td>14.8932</td><td>12.2724</td></tr>
</table>
</td>
</tr>
<tr>
<td>[19]</td>
<td>2010</td>
<td>Adaptive Filter Algorithms: (LMS), (NLMS), (RLS), APA, FEDS, FAPA</td>
<td>Office noise</td>
<td>Compared to the LMS algorithm, the NLMS and AP algorithms converge faster. Both the learning curve and the filtered output make this reality clear. The results show that the convergence rates of the FEDS and FAP algorithms are equivalent to those of the RLS method and are quicker than those of the LMS, NLMS, and AP algorithms. They also show that, compared to the LMS, NLMS, and AP algorithms, the RLS technique converges more quickly. The results show that the FEDS and FAP are on par with the RLS algorithm in terms of performance and exceed the LMS, NLMS, and AP algorithms.</td>
<td>Comparison of adaptive filter algorithms
<table>
<tr><th>Algorithm</th><th>SNRI (db)</th></tr>
<tr><td>LMS</td><td>13.5905</td></tr>
<tr><td>NLMS</td><td>16.8679</td></tr>
<tr><td>APA</td><td>20.0307</td></tr>
<tr><td>FEDS</td><td>22.2623</td></tr>
<tr><td>FAPA</td><td>24.9078</td></tr>
<tr><td>RLS</td><td>29.7355</td></tr>
</table>
</td>
</tr>
</table>

<table>
<tr><td>[20]</td><td>2011</td><td>Novel algorithm</td><td>White, Pink, Buccaneer2, Destroyer engine, Factory noise, Destroyers noise</td><td>The reverberation effect characteristic of the DD approach is diminished by the provided technique since it enhances the estimate of the a priori SNR using the DD approach and has the capacity to follow changes in the speech signal. The simulation findings support the analysis presented above.</td><td>
<table>
<tr><th rowspan="2">Noise type</th><th rowspan="2">Input SNR</th><th>Segmental SNR (dB)</th></tr>
<tr><td>proposed algorithm</td></tr>
<tr><td rowspan="3">White</td><td>0 dB</td><td>4.9109</td></tr>
<tr><td>5 dB</td><td>6.7544</td></tr>
<tr><td>10 dB</td><td>9.1751</td></tr>
</table>
</td></tr>
<tr><td>[21]</td><td>2012</td><td>Perceptual distortion measure. Speech Pre-Processing Algorithm</td><td>White, Babble, Factory, F16</td><td>For all types of noise, the suggested technique performs better than the reference methods. The SNR may be decreased by 3-5 dBs using the suggested technique, according to objective intelligibility prediction findings, without sacrificing intelligibility.</td><td>
<table>
<tr><th></th><th>STOI</th></tr>
<tr><td>SNR</td><td>White noise</td></tr>
<tr><td>0</td><td>>90</td></tr>
<tr><td>5</td><td>100</td></tr>
</table>
</td></tr>
<tr><td>[22]</td><td>2014</td><td>Adaptive Wiener filtering, Spectral subtraction, Traditional frequency-domain Wiener Filtering method, Wavelet.</td><td>Additive White Gaussian Noise (AWGN) as well as colored noise.</td><td>The simulation results show that both colored and additive white Gaussian noise may be filtered effectively using the recommended Wiener filtering approach (AWGN). The outcomes show that the suggested adaptive Wiener filtering approach beats all other speech enhancement methods currently in use for both low and high SNR levels. The proposed filter successfully handles AWGN and colored noise situations. The adaptive characteristics of the filter impulse response account for this. The suggested adaptive Wiener filter also has the advantage of only requiring the noisy signal as a single input.</td><td>
Results for SNR = 5 dB for the AWGN
<table>
<tr><th>Signals</th><th>Metric</th><th>Adaptive Wiener filtering</th></tr>
<tr><td rowspan="2">Male</td><td>SNR</td><td>6.8726</td></tr>
<tr><td>SNRseg</td><td>6.8423</td></tr>
<tr><td rowspan="2">Female</td><td>SNR</td><td>6.8373</td></tr>
<tr><td>SNRseg</td><td>6.8021</td></tr>
</table>
</td></tr>
<tr><td>[23]</td><td>2014</td><td>Wiener Filter, Adaptive Filter (LMS)</td><td>Real time noise, Mixed noise</td><td>The oldest technique for noise cancellation is the Wiener filter; however, it is rather sophisticated. In order to decrease complexity and computational speed, adaptive filters are introduced. The complexity and stability of the systems will rise as the authors attempt to lower the mean square error. The simulation results show that the LMS algorithm is the most appropriate due to its simplicity and lower cost than the Wiener filter. Although LMS is the best algorithm, it has a slow rate of convergence.</td><td>
<table>
<tr><th>Algorithm</th><th>Speed</th><th>Convergence rate</th><th>Complexity</th><th>MSE</th></tr>
<tr><td>Wiener filter</td><td>Low</td><td>-</td><td>High</td><td>Low</td></tr>
<tr><td>LMS</td><td>High</td><td>Low</td><td>Low</td><td>High</td></tr>
</table>
</td></tr>
<tr><td>[24]</td><td>2015</td><td>LMS, RLS, ANC, FIR - LMS, lattice gradient algorithms</td><td>White noise, pink noise, blue noise</td><td>According to studies, the RLS algorithm produces the highest SNR and outperforms the LMS algorithm for lower-order FIR adaptive filters. However, for the Finite Impulse Response (FIR) filter Taps, LMS converges faster than RLS. Fixing the FIR tap weight yielded the best Mu (LMS) and Lambda (RLS) values. Acoustic noise cancellation (ANC) is best for removing ambient noise. Traditional wideband ANC algorithms perform best in lower frequency bands and degrade rapidly as the bandwidth and center frequency of the noise increase. The least mean squares (LMS) and lattice gradient (LG) techniques have been demonstrated to lower background noise strength by at least 20 dB with little to no speech stuttering, making them potentially effective noise suppression pre-processors for voice communication in loud environments.</td><td>
<table>
<tr><td>Original unsmoothed noise</td><td>1</td></tr>
<tr><td>Smoothed white noise</td><td>0.1</td></tr>
<tr><td>Smoothed pink noise</td><td>0.55</td></tr>
<tr><td>Smoothed blue noise</td><td>0.01</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><td>[7]</td><td>2016</td><td>Spectral subtraction & Wiener filter technique</td><td>White noise, Helicopter, noise, Babble noise, Car noise</td><td>In spectral subtraction. The method is not efficient for corrupting speech with non-stationary noise such as car noise, babble noise, and helicopter noise.</td><td>
<table>
<tr><th>Noise Type</th><th>Input SNR (dB)</th><th>Spectral Subtraction</th><th>DD Approach</th><th>TSNR Method</th><th>TSNR and HRNR</th></tr>
<tr><td rowspan="3">White noise</td><td>0</td><td>11.82</td><td>11.7</td><td>12.83</td><td>13.40</td></tr>
<tr><td>2</td><td>13.09</td><td>13.68</td><td>14.35</td><td>45.27</td></tr>
<tr><td>5</td><td>14.76</td><td>16.57</td><td>16.26</td><td>17.61</td></tr>
</table>
</td></tr>
<tr><td>[25]</td><td>2016</td><td>Adaptive Wiener filter
.</td><td>Fan noise</td><td>The concept of an adaptive Wiener filter with TSNR and HRNR methods is used to enhance the noisy speech signal. Two-step noise reduction methods are used to enhance the speech signal. Then harmonic regenerated noise reduction is used to regenerate the harmonics that are lost from the original signal. Experimental results show that the SNR of the input signal is improved by using the TSNR and HRNR methods. The HRNR method gave the best SNR compared to TSNR.</td><td>
<table>
<tr><th>methods</th><th>Input SNR (dB)</th><th>Output SNR (dB)</th><th>Improvement in SNR (dB)</th></tr>
<tr><td>TSNR</td><td>0.649</td><td>1.109</td><td>0.46</td></tr>
<tr><td>HRNR</td><td>0.649</td><td>1.3597</td><td>0.71</td></tr>
</table>
</td></tr>
<tr><td>[26]</td><td>2016</td><td>Deep Neural Networks (DNN)</td><td>Stationary, Non-stationary, Multiple noises</td><td>Results demonstrate the effectiveness of DNN-based voice augmentation techniques in these challenging settings that closely resemble real-world settings. Over all test SNRs, the best model provides an average PESQ increase of 23:97%. This value is about 30% at lower SNRs. In comparison to traditional approaches like log-MMSE, this is substantially better.</td><td>
Average PESQ and STOI
<table>
<tr><th rowspan="2">SNR (db)</th><th colspan="2">Noisy</th><th colspan="2">BD</th><th colspan="2">BSD</th><th colspan="2">BED</th></tr>
<tr><td>PESQ</td><td>STOI</td><td>PESQ</td><td>STOI</td><td>PESQ</td><td>STOI</td><td>PESQ</td><td>STOI</td></tr>
<tr><td>5</td><td>2.11</td><td>0.81</td><td>2.69</td><td>0.87</td><td>2.75</td><td>0.88</td><td>2.74</td><td>0.87</td></tr>
<tr><td>10</td><td>2.53</td><td>0.89</td><td>3.09</td><td>0.92</td><td>3.14</td><td>0.92</td><td>3.14</td><td>0.92</td></tr>
</table>
</td></tr>
<tr><td>[27]</td><td>2017</td><td>Windowing Technique, windows used: Kaiser, Rectangle, Hamming, Hanning, Welch</td><td>power line noise muscle noise EMG noise</td><td>Among all the selected performance factors, Kaiser Windows had the best performance. Closely following this came the rectangle window, which similarly performed well across the board with the exception of SNR. Therefore, it can be inferred that Kaiser window has the best denoising performance, followed by rectangle window and, in some instances, Hamming window as well. Both Kaiser and rectangle windows performed adequately in eliminating power line noise from the ECG data.</td><td>1. MSE: The Kaiser window has the lowest MSE of 0.047. Rectangle and Hamming windows have a close resemblance in performance, with MSE higher than that of Kaiser windows. Hanning and Welch windows have high MSE values of 0.069 and 0.071, which are not suitable for good performance in denoising.
2. SNR: Kaiser, rectangle, and hamming windows have the highest values of denoised SNR, i.e., about 7.5 dB. Hanning and Welch windows, too, have good values of SNR independently.</td></tr>
</table>

<table>
<tr><td>[4]</td><td>2018</td><td>Different linear filtering techniques: LMS, NLMS, Kalman, Wiener</td><td>AWGN noise</td><td>It was discovered that the Kalman filter provides better results than other adaptive filters like LMS and NLMS. We have also discovered from the results that the performance of the Kalman filter is superior to that of the LMS, NLMS, and Wiener filters for AWGN noise, but performance varies for practical noise. In light of the investigation, it can be said that the Kalman filter performs the best among all of the filters for AWGN noise but not for real-world noise.</td><td>
<table>
<tr><th>Sl No.</th><th>Parameter</th><th>LMS</th><th>NLMS</th><th>Kalman</th></tr>
<tr><td colspan="5">Comparisons of the filter with awgn noise</td></tr>
<tr><td>1</td><td>PSNR</td><td>56.2380</td><td>57.8854</td><td>69.6564</td></tr>
<tr><td>2</td><td>MSE</td><td>0.1546</td><td>0.1058</td><td>0.0017</td></tr>
<tr><td>3</td><td>Max error</td><td>2.1650</td><td>1.4057</td><td>1.3563</td></tr>
</table>
</td></tr>
<tr><td>[28]</td><td>2018</td><td>Wavelet filtering</td><td>Additive White Gaussian Noise (AWGN)</td><td>Based on wavelet analysis in speech signals, the use of wavelets from different families is compared, including nine- and tenth-order Daubechies wavelets, fourth-order Symlets, and fifth-order Coiflets. The value of the cross-correlation at point 0 for all types is of the order of 0.8, and the outcomes of noise suppression are sufficiently high.</td><td>The wavelet types employed produced the following average mutual correlation function values at point 0: The tenth order of Daubechies wavelets is 0.79382. The ninth order is 0.78865. The fourth order of symlets is 0.78883. The fifth order of coiflets is 0.78776.</td></tr>
<tr><td>[29]</td><td>2018</td><td>LPC based FEM in Wiener filter</td><td>Engine noise</td><td>The experimental findings demonstrate that our suggested approach yields superior results while preserving all of the original voice signal's information. The observed values and results demonstrate that a formant enhancement approach based on linear predictive coding and Wiener filtering yields greater values of PESQ, SNR, NC, and PSNR. Higher PESQ, SNR, NC, and PSNR values indicate higher and enhanced output signal quality when compared to previous speech enhancement strategies.</td><td>Comparison of different speech enhancement methods based on PESQ, SNR, NC, and PSNR.
<table>
<tr><th rowspan="2">Methods</th><th colspan="3">Objective evaluation</th></tr>
<tr><td>PESQ</td><td>SNR (db)</td><td>PSNR (db)</td></tr>
<tr><td>Noise</td><td>1.80</td><td>5</td><td>55</td></tr>
<tr><td>Wiener + LPC based FEM</td><td>3.54</td><td>12.25</td><td>99.18</td></tr>
</table>
</td></tr>
<tr><td>[30]</td><td>2018</td><td>Kalman filtering</td><td>Non-stationary noise, White noise, Babble noise, F16 noise, Factory noise</td><td>According to experimental findings, it is advantageous to track speech in the log bark power spectral domain while also accounting for the temporal dynamics of each bark subband envelope.</td><td>LSA and pMMSE only yield a PESQ improvement of around 0.4; however, the BSNT method achieves a PESQ improvement of 0.62. PESQ improvements for ST, SNT, and OMLSA are all around 0.5, while improvements for BST are about 0.55. In comparison to the LSA and pMMSE baselines, the PESQ improvement for the BSNT method is greater at about 0.22, 0.12 for OMLSA, 0.1 for ST and SNT, and 0.05 for BST.</td></tr>
<tr><td>[31]</td><td>2019</td><td>Neural network methods ‘ADALINE’
.</td><td>Stationary noise</td><td>The research concluded that the most effective denoising strategy for the majority of speech signals is the adaptive filter employing the LMS algorithm. ADALINE produces the greatest outcomes out of all the neural network techniques. In order to reduce error between network output and the targets, ADALINE employs the adaptive LMS algorithm. The fundamental performance difference between filtering and neural network denoising techniques, however, is that neural networks take longer to execute than filtering techniques. The most complicated technique is the deep learning denoising technique. Deep learning has a very long</td><td>
<table>
<tr><th></th><th>Adaptive Filter</th><th>ADALINE</th><th>Weiner filter</th><th>Kalman Filter</th></tr>
<tr><td>Avg. SNR</td><td>13.2977</td><td>6.7229</td><td>1.0089</td><td>-0.466</td></tr>
<tr><td>Avg. PSNR</td><td>38.4618</td><td>31.8517</td><td>26.1309</td><td>24.6559</td></tr>
</table>
</td></tr>
</table>

<table>
<tr><td></td><td></td><td></td><td></td><td>execution time, but the outcomes are unsatisfactory. Therefore, deep learning denoising techniques cannot be used for speech applications.</td><td></td></tr>
<tr><td>[32]</td><td>2020</td><td>Electroencephalography (EEG) signals using a Generative adversarial network (GAN) based model, Compare with traditional log (MMSE)</td><td>Gaussian noise</td><td>Utilize cutting-edge deep learning models similar to the ideas of GAN, GRU, TCN regression, and EEG signal processing. This is the first instance where deep learning models have been used to show how to improve spoken pronunciation using EEG characteristics.</td><td>
<table>
<tr><th>Model</th><th>Test Set Avg PESQ</th><th>Enhanced Output Avg PESQ</th></tr>
<tr><td>Log MMSE</td><td>2.4</td><td>2.48</td></tr>
<tr><td>GRU Regression</td><td>2.4</td><td>2.59</td></tr>
<tr><td>TCN Regression</td><td>2.4</td><td>2.48</td></tr>
<tr><td>GRU-TCN Reg.</td><td>2.4</td><td>2.52</td></tr>
<tr><td>GAN</td><td>2.4</td><td>2.50</td></tr>
</table>
</td></tr>
<tr><td>[33]</td><td>2020</td><td>Deep neural network (DNN)-augmented colored noise Kalman filter</td><td>Pink, Buccaneer2, Destroyer engine, Hf channel, Babble, White, Street, Factory</td><td>The suggested method is better equipped to handle color noises seen in real-world settings. Experiments have demonstrated the suggested system’s advantage in two areas. First, the enhanced performance of colored-noise Kalman filtering is significantly enhanced by the use of DNN for parameter estimation and post-subtraction for residual noise suppression. Second, because our suggested approach benefits from both supervised and unsupervised techniques, it is more generalizable. In fact, it gives much superior results on unseen noise, even though it achieves equivalent performance to contemporary DNN-based techniques on visible noise.</td><td>
Seen noise
<table>
<tr><th rowspan="2">Method</th><th colspan="2">PESQ</th><th colspan="2">STOI</th></tr>
<tr><td>3dB</td><td>6dB</td><td>3dB</td><td>6dB</td></tr>
<tr><td>Noisy</td><td>1.68</td><td>1.86</td><td>0.78</td><td>0.83</td></tr>
<tr><td>FNN-MAG</td><td>2.34</td><td>2.55</td><td>0.86</td><td>0.88</td></tr>
<tr><td>FNN-CKFS</td><td>2.32</td><td>2.51</td><td>0.85</td><td>0.88</td></tr>
<tr><td>LSTM-CKFS</td><td>2.38</td><td>2.58</td><td>0.85</td><td>0.88</td></tr>
</table>
</td></tr>
<tr><td>[34]</td><td>2020</td><td>Long short-term memory (LSTM) And (CNN)</td><td>Factory2, Buccaneer1, Destroyer engine, HF channel</td><td>According to the experimental findings, the proposed time-frequency smoothing neural network greatly enhances the voice quality and understandability of improved speech and performs speech enhancement more effectively than previous networks.</td><td>
<table>
<tr><th>Filter size</th><th>5</th><th>7</th><th>9</th><th>13</th></tr>
<tr><td>PESQ</td><td>2.42</td><td>2.45</td><td>2.51</td><td>2.51</td></tr>
<tr><td>STOI (%)</td><td>74.83</td><td>75.57</td><td>77.05</td><td>77.32</td></tr>
<tr><td>Parameters (M)</td><td>0.64</td><td>0.68</td><td>0.72</td><td>0.8</td></tr>
</table>
</td></tr>
</table>

<table>
<tr>
<td>[35]</td>
<td>2020</td>
<td>mask-based long short-term memory (LSTM), convolutional encoder-decoder (CED)</td>
<td>Office, Pub, Traffic noise.</td>
<td>They proposed a brand-new, two-stage strategy for improving speaking. An LSTM network estimates T-F masks for both actual and fictitious portions of the noisy speech spectrum in the first stage, and a convolutional encoder-decoder (CED) network conducts spectral mapping in the second stage. They notice a tremendous improvement of more than 5 dB in SNR over the baselines, but little to no improvement in overall quality (PESQ). When using both stages, it is possible to achieve average PESQ gains of roughly 0.1 MOS points in undetected, highly non-stationary disturbances, such as interfering speech. In comparison to the baselines, our method also enhances STOI in low-SNR situations.</td>
<td>
<table>
<tr><th>SNR</th><th>Method</th><th>PESQ</th><th>STOI</th><th>SNRI</th></tr>
<tr><td rowspan="4">**10**</td><td>**CED-cSA-du**</td><td>2.96</td><td>**0.95**</td><td>23.40</td></tr>
<tr><td>**LSTM-cMSA + DNN-cSA**</td><td>**3.03**</td><td>**0.95**</td><td>**25.88**</td></tr>
<tr><td>**LSTM-cMSA + CED-cSA-du**</td><td>3.12</td><td>**0.96**</td><td>25.11</td></tr>
<tr><td>**LSTM-cMSA + CED-cSA-tr**</td><td>**3.15**</td><td>**0.96**</td><td>**26.06**</td></tr>
</table>
</td>
</tr>
<tr>
<td>[36]</td>
<td>2021</td>
<td>Modified LMS algorithm</td>
<td>Airport, Babble, Restaurant, AWGN</td>
<td>This experiment showed that the proposed method did a better job than the LMS algorithm of improving the quality of the speech signal. This is shown by the three objective measures of dbsnr, LLR, and ISD.</td>
<td>
<table>
<tr><th rowspan="2">S. No.</th><th rowspan="2">Noise Type</th><th rowspan="2">Parameters</th><th colspan="2">Noise Level</th></tr>
<tr><td>5 dB</td><td>10 dB</td></tr>
<tr><td rowspan="5">1</td><td rowspan="5">AWGN</td><td>Noisy SNR</td><td>1.17</td><td>0.44</td></tr>
<tr><td>De-noised SNR</td><td>0.11</td><td>1.77</td></tr>
<tr><td>dB SNR</td><td>4.99</td><td>8.10</td></tr>
<tr><td>LLR</td><td>2.21</td><td>1.85</td></tr>
<tr><td>ISD</td><td>76.6</td><td>117.6</td></tr>
</table>
</td>
</tr>
<tr>
<td>[37]</td>
<td>2021</td>
<td>Filter bank</td>
<td>Acoustic noise, Impulsive noise, Transient noise</td>
<td>While MFCC and other feature extraction algorithms aren’t as good at ignoring noise when getting information, band-pass filtering gets the characteristics of the transmitted signal more accurately and with more reliability. In noisy situations, the voice signal may be accurately extracted using the filter bank analysis approach. The suggested method has demonstrated the ability to extract features from speech signal properties at a substantial pace. When the signal's detailed coefficients are close to the pitch, the performance of the suggested system hasn't degraded much.</td>
<td>
Noise Suppression with Filter Bank.
<table>
<tr><th>Features</th><th>Speech Activity</th><th>Other</th></tr>
<tr><td>Zero Crossing</td><td>0.54</td><td>0.75</td></tr>
<tr><td>Pitch</td><td>0.41</td><td>0.69</td></tr>
<tr><td>Loudness</td><td>0.71</td><td>0.31</td></tr>
<tr><td>Cepstral Peak</td><td>0.84</td><td>0.79</td></tr>
</table>
</td>
</tr>
<tr>
<td>[38]</td>
<td>2021</td>
<td>Wiener filter</td>
<td>Babbling, Cars, Street, Trains, Restaurants, Airport noise.</td>
<td>To estimate clean speech from noisy speech, the implicit Wiener filter with a recursive noise estimation algorithm is suggested and contrasted with the traditional spectral subtraction approach. Compared to spectral subtraction, the results reveal that the envelop of the predicted noise using the implicit Wiener</td>
<td>
<table>
<tr><th rowspan="2">Noise Type</th><th rowspan="2">Input SNR (db)</th><th>LLR</th><th>CD</th><th>WSS</th></tr>
<tr><td colspan="3">W F</td></tr>
<tr><td>Babble</td><td>5</td><td>0.933</td><td>5.584</td><td>75.049</td></tr>
</table>
</td>
</tr>
</table>

<table>
<tr><td></td><td></td><td></td><td></td><td>filter is pretty similar to the envelop of the noisy speech spectrum. The suggested approach produces an improved speech signal that is spectrally comparable to the clean speech signal and perceptually similar to the clean speech signal. While the voice distortion is still tolerable, the musical noise is less organized than the spectral subtraction.</td><td><table><tr><td>Car</td><td>5</td><td>0.806</td><td>5.122</td><td>65.674</td></tr><tr><td>Airport</td><td>5</td><td>0.822</td><td>5.195</td><td>70.26</td></tr></table></td></tr>
<tr><td>[39]</td><td>2021</td><td>Recursive Least Squares (RLS) algorithms and Least Mean Square (LMS).</td><td>white, colored noises</td><td>This technique reduces the noise levels in a specific signal and makes it simple to analyze the signal's properties. Without lowering the signal's quality, the voice signal's comprehensibility and quality are improved. In comparison to all other algorithms, the system now performs better. The MSE and SNR were computed. We concluded from the findings that the suggested strategy is successful in reducing noise. The suggested technique has successfully been applied in the real-time speech de-noising test, which shows a considerable improvement in speech intelligibility.</td><td>-</td></tr>
<tr><td>[40]</td><td>2021</td><td>Recursive EM (REM) algorithm and Kalman filter.</td><td>air-conditioner (AC) noise, pseudo-diffused babble noise,</td><td>The three issues are all being addressed at once by the suggested statistical approach. Each STFT timeframe employs the E-step and M-step. The Kalman filter is used to implement the E-step. In the M-step, the model parameters are estimated. The echo signal at each channel is assessed using the estimated acoustic route of the far-end signal. The method has convergence capabilities even in time-varying double-talk circumstances. We demonstrate that our approach outperforms rival NLMS-based approaches in terms of speech quality, comprehension, and echo cancellation performance.</td><td><table><tr><td></td><td>LSD</td><td>SRMR</td><td>STOI</td><td>_SER</td></tr><tr><td>Proposed</td><td>3.54</td><td>7.40</td><td>95.57</td><td>23.16</td></tr></table></td></tr>
<tr><td>[41]</td><td>2021</td><td>LBLG and NBLG</td><td>White noise and babble noise.</td><td>LBLE performed well in terms of speech component regeneration. On the other hand, the suggested SEA is a hybrid estimator made up of two stages in a cascade. It is based on a high-characteristic orthogonal transform that improves noisy signals. The suggested SEA's comparative assessment reveals its efficacy and ability to reduce noise in terms of LLR, SIG, OVL, BAK, PESG, and segSNR measures. Simulations of several noisy situations demonstrate that the proposed SEA suppresses unwanted noise more effectively than existing techniques.</td><td>The findings were presented in terms of SegSNR, PESQ, LLR, SIG, BAK, and OVL. The comparison values presented are the averages of ten input signals. The suggested SEA (DKTT-Two stage) provided superior output signals compared to the DCT-NBLE, DCT-LBLE, and DCT-Two stage, yielding the best values. In the instance of babble noise, the suggested SEA produces the greatest results across most testing parameters and settings. It can be shown that for LLR, PESQ, and OVL, the suggested SEA delivers a comparable result to the other estimators for specific levels of SNR only, but for all other levels of SNR, it provides the best results in most circumstances.</td></tr>
<tr><td>[42]</td><td>2021</td><td>(MMSE), linear bilateral Laplacian gain estimator (LBLG), and nonlinear,</td><td>White noise, F-16 noise, speech shaped noise, pink noise.</td><td>In terms of several objective metrics, the suggested estimators outperformed existing approaches such as the standard MMSE approach, the perceptually driven Bayesian estimator, the dual-gain Wiener estimator, and the dual MMSE estimator.
The suggested estimators' results demonstrate their usefulness and capability in decreasing unwanted noise in terms of</td><td>SNR results:<table><tr><td>Noisy type</td><td>SNR (db)</td><td>Proposed LBLG</td><td>Proposed NLBLG</td></tr><tr><td>White</td><td>5</td><td>4.86</td><td>5.17</td></tr><tr><td>F-16</td><td>5</td><td>4.75</td><td>4.71</td></tr><tr><td>speec</td><td>5</td><td>4.39</td><td>4.41</td></tr></table></td></tr>
</table>

<table>
<tr><td></td><td></td><td>bilateral Laplacian gain estimator (NBLG),</td><td></td><td>segSNR and PESQ measures. Furthermore, the suggested estimators outperform the competition since residual noise and speech signal distortion are lower.</td><td><table><tr><td>h shaped</td><td></td><td></td><td></td></tr><tr><td>pink</td><td>5</td><td>4.46</td><td>4.49</td></tr></table></td></tr>
<tr><td>[43]</td><td>2022</td><td>Modified wiener filter</td><td>Exhibition, Airport, Car, Restaurant, Street, Babble noise</td><td>The performance of the Wiener filter is contrasted with that of the modified noise reduction method (MNRM). Based on a comparison of these two filters, our suggested model of the MNST filter boosted the SNR ratio for those various sounds. By utilizing the suggested model, the SNR ratio rose by about 14 to 15%. Modern subband domain DHA can use the MNST for noise reduction. This will result in a better and more relaxing listening experience for the hearing aid user.</td><td><table><tr><th>Various Ambient noises</th><th>SNR by proposed MNRM (db)</th></tr><tr><td>Airport noise</td><td>23.456</td></tr><tr><td>Car noise</td><td>25.345</td></tr><tr><td>Street Noise</td><td>24.456</td></tr><tr><td>Babble Noise</td><td>23.456</td></tr></table></td></tr>
<tr><td>[44]</td><td>2022</td><td>Adaptive LMS filtering</td><td>stationary white Gaussian, engine noise.</td><td>The suggested strategy can enhance the performance and quality of noisy audio signals, according to simulation findings. They have shown through computer simulations that the suggested strategy is highly efficient in reducing noise, particularly in the case of stationary white Gaussian noise.</td><td>For cases where the filter's input power is very unpredictable, the NLM approach is appropriate. It becomes more challenging to statistically examine the NLM algorithm after updating the fully qualified balance. The results show that using the adaptive LMS filter can produce clear voice signals that are of higher quality than signals produced by pure noise.</td></tr>
<tr><td>[45]</td><td>2022</td><td>implicit Wiener filter</td><td>Exhibition, Station, Drone, Helicopter, Airplane, White Gaussian stationary noise</td><td>The implementation findings demonstrate that the suggested speech enhancement algorithm outperforms the spectral subtraction technique for the various types of noise degradations examined. Additionally, the estimated or enhanced speech signal's envelope is quite similar to the clean speech spectrum's envelope. It is demonstrated that the perceived quality of the augmented speech signal and the clean speech signal are comparable. The clean speech spectrogram and the enhanced speech spectrogram share similarities. Additionally, when compared to the spectral subtraction algorithm, the enhanced speech signal produced by the proposed speech enhancement algorithm has a clearer sound.</td><td><table><tr><td rowspan="2">Type of noise</td><td rowspan="2">Input SNR (dB)</td><th colspan="2">PESQ</th></tr><tr><th>SS</th><th>IWF</th></tr><tr><td rowspan="3">AWGN</td><td>0</td><td>1.563</td><td>1.657</td></tr><tr><td>2.5</td><td>1.908</td><td>1.932</td></tr><tr><td>5</td><td>1.656</td><td>2.008</td></tr></table></td></tr>
</table>

3. Noise Definition

In a communication system, noise is essentially undesired or undesirable signals that are added at random to the signal that is really carrying the information or combined with a voice signal at the time of speech signal production or transmission. As a result, the original signal that is being transmitted from one end to another is disturbed. Even when they are not interfering with other signals or may have been purposely created as comfort noise, the term can also be used to describe signals that are random (unpredictable) and provide no meaningful information [46], such as noises in the sonar images [47], or seismic data [48].

To put it another way, noise is a signal that transmits information about its sources and the environment in which it spreads. For instance, background voice dialogues in a busy place might create interference with the hearing of a desired conversation or speech, and the noise from a car engine conveys information about the condition of the engine and how efficiently it is working. There are two different types of noises here:

1- According to life

In our life there is different face of noised in general the noise classified into 4 known types [49]:

- Continuous noise: a noise that is continually created, for example, by machinery that runs continuously.
- Intermittent noise: a noise volume that rapidly rises and falls. This might be caused by a passing train.
- Impulsive noise: most often linked to the building and demolition industries. These loud blasts of sound. Explosions and construction equipment frequently produce impulsive sounds.
- Low-frequency noise: Low-frequency noise is woven into the fabric of our everyday soundscape. We are frequently exposed to low-frequency noise, whether it is the low background hum of a neighboring power plant or the roaring of massive diesel engines.

2- According to signal processing (colored noise):

There are many types of noises that signals carry when generated or transmitted, such as additive noise (white noise, additive white Gaussian noise, black noise, Gaussian noise, pink noise or flicker noise, brownian noise, contaminated Gaussian noise, power-law noise, Cauchy noise, and multiplicative noise), quantization noise, poisson noise, shot noise, transient noise, burst noise, phase noise, background noise, comfort noise, and electromagnetically induced noise [46].

4. Classification of Noise Reduction Strategies

A broad classification of noise reduction methods can be given as spectral processing and temporal processing methods. The degraded speech goes through processing in the frequency domain in the spectral processing methods, whereas processing will be in the time domain for the temporal processing methods [46].

Several methods were proposed for noise reduction, such as:

The noise in the surroundings corrupts the information signal as it travels in a free environment. Eliminating this noise turns out to be one of everyone's top concerns. There are numerous traditional methods for reducing the noise in the information signal.

4.1. Spectral subtraction

The simplest and most familiar method to remove stationary background noise is spectral subtraction. In this technique, the average magnitude of the noise spectrum is subtracted from the noisy speech spectrum. The average magnitude of the noise spectrum is estimated from the frames of speech absence. The main disadvantage of the spectral subtraction method is that it produces residual noise with irritating and noticeable tonal characteristics known as “musical noise” [7]. Additionally, spectral subtraction does not sufficiently reduce noise during the silent period [22].

4.2. Wiener filtering

The second most familiar method is the Wiener filter, which is a substitute method of spectral subtraction for increasing the quality of the speech signal. A wiener filter is a kind of optimum filter that uses statistical assumptions and previous information to estimate the desired signal from a noisy observation. The main aim is to develop a filter that minimizes the squared difference between the output and the real signal. The drawback of this filter is that it requires previous knowledge of the power spectra of the input, noise, and real signals. In many circumstances, this can be difficult or impracticable to acquire, especially if the signal and noise are non-stationary or non-Gaussian. [50]

4.3. Kalman filter

It is broadly used in speech improvement. The Kalman filter is a model-based system that models a speech signal as an autoregressive (AR) process and also recovers the speech signal. The Kalman filtering system for speech improvement has no supposition of stationary speech signals; it is designed to work with finite data sets; it makes use of models; and it can be made to work with non-stationary signals. The Kalman filter has two major limitations: It assumes that the equations for the system and observation models are both linear, which is unrealistic in many real-world situations. It is assumed that the state belief distribution is Gaussian [4].

4.4. Subspace

Another technique for improving speech is used when speech estimation is seen as a constrained optimization issue. A signal subspace and a noise subspace are formed from the noisy speech signal vector universe. A signal-subspace speech improvement approach was put forward by Surendran et al. [51] employing a perceptual feature and the human auditory system's frequency masking or frequency disguising properties. [52] In comparison to various benchmark speech enhancement techniques, the results of their studies demonstrated the effectiveness of their algorithm. It was demonstrated that their method performed better with white noise compared to colored noise. SNR greater than 10 dB had poor performance.

4.5. Adaptive Filters

Several techniques for filtering noise from a speech waveform have been investigated. The majority of these techniques are based on the concept of adaptive filtering [8]. A system having a linear filter and a transfer function controlled by adjustable parameters and a way to change those parameters in accordance with an optimization technique is called an adaptive filter [53]. Modern digital signal processing (DSP) uses adaptive filters extensively in applications such as active noise control (ANC), adaptive control systems, telephone echo cancellation, noise cancellation, communications channel equalization, and biomedical signal amplification. One of the most frequently suggested ways to reduce the signal corruption brought on by predictable and unpredictable noise is the use of adaptive filters. For almost 50 years, adaptive filters have been employed in a variety of fields. Adaptive filtering configurations include inverse modeling, adaptive equalization, adaptive noise cancellation, adaptive linear prediction, and more [12].

4.6. Least-Mean-Square (LMS)

There are a number of noise reduction algorithms that may be used and implemented using MATLAB. An adaptive algorithm is one that modifies its features during execution according to the available data and prior techniques. The LMS algorithm, or least mean squares algorithm, is a well-known algorithm for adaptive systems that functions as a self-adjusting algorithm. In 1959, Widrow and Hoff produced the LMS (Least Mean Square) algorithm [54], a fairly simple method for noise cancellation. Due to its durability and dependability, the LMS algorithm's simplicity, cheap computing complexity, and quick convergence rate have led many academics to embrace it for hardware implementation. In a noise cancellation problem, the LMS algorithm has demonstrated good performance [12]. Furthermore, with colored interference signals, the LMS suffers from significant performance degradation [24].

5. Types of Filters

In signal processing, a filter is a device or system that removes some undesirable additives or features from a signal. Filters are commonly utilized in electronic and telecommunications applications such as radio, television, audio recording, radar, control systems, music synthesis, image processing, and computer graphics.

There are numerous classifications of filter bases that overlap in a variety of ways; there is no simple hierarchical classification. Filters include [55]:

- Non-linear or linear.
- Time-variant or time-invariant.
- Analog or digital.
- Discrete-time (sampled) or continuous-time
- Passive or active type of continuous-time filter.
- Infinite impulse response (IIR) or finite impulse response (FIR) types of discrete-time or digital filters.

Table 2 illustrates the different speech enhancement techniques and their sub-methods.

Table 1:Speech enhancement techniques

Optimal Filters in the Time Domain	**Optimal Filters in the Frequency Domain**	**Statistical Based Approaches**	**Adaptive Filters**	**Filter Bank**
1. Wiener Filter **2.** Tradeoff Filter **3.** Subspace Approach **4.** Mean. **5.** Median. **6.** Gaussian. **7.** Bilateral. **8.** Comb Filtering **9.** Linear predictive coding (LPC)based Filtering **10.** Adaptive Filtering • Kalman Filtering, • H algorithm **11.** Hidden Markov Model HMM Filtering **12.** Neural Networks **13.** Tradeoff Filters **14.** Subspace Approach	**1.** Wiener Filter **2.** Parametric Wiener Filter **3.** Tradeoff Filter **4.** Fourier-transform **1.** Low-Pass Filter (LPF). **2.** High-Pass Filter (HPF) **5.** Wavelet transforms. • Low-Pass Filter (LPF). • High-Pass Filter (HPF).	**1.** Wiener Filtering **2.** Maximum Likelihood (ML) Estimators **3.** Bayesian Estimators **4.** MMSE Estimators • MMSE Magnitude Estimator • MMSE Complex Exponential Estimator **5.** LogMMSE Estimators **6.** Maximum A Posteriori (MAP) Estimators **7.** Perceptually Motivated Bayesian Estimators.	1. Finite Impulse Response (FIR) Adaptive filters 2. Infinite Impulse Response (IIR) Adaptive filters 3. Kalman Filtering, 4. H algorithm 5. Adaptive wiener filter 6. Adaptive Kalman filter 7. Least Mean Square (LMS) algorithm	1. Discrete Cosine Transform (DCT) Filter Banks 2. Polyphase Filter Banks 3. Gabor Filter Banks 4. Mel Filter Banks 5. Filter Bank Multicarrier (FBMC) 6. Discrete Fourier Transform (DFT) Filter Banks 7. Uniform DFT Filter Bank

6. Discussion

People have become more dependent on communication technology in recent years as it has grown on a huge scale [56], and speech communication is increasingly significant in everyday applications, especially with the invention of mobile phones and Internet services, which enabled the transmission of voice through networks in digital format [57]. For that, we need filters to eliminate the noise of transmitted speech. A noise-reduction filter is used to generate the clean speech estimate during the noise reduction process, which is conceptualized as a filtering issue. With such a formulation, the fundamental challenge of noise reduction is how to create an ideal filter that can fully use the speech and noise statistics to achieve maximal noise suppression without adding perceptibly detectable speech distortion. Although effective filters may be created in the time domain, the majority of techniques now in use operate in the frequency domain. Working in the frequency domain has a variety of benefits, including but not limited to[58] :

1- The quick Fourier transform makes it possible to execute the filtering procedure extremely effectively.
2- There is a great degree of versatility in dealing with colored noise since the filters at different frequencies may be created and managed independently of one another.
3- Since the majority of our knowledge and understanding of speech production and perception is based on frequencies, we can easily apply this knowledge to the frequency domain to improve noise reduction performance.

According to our review, there are currently no perfect techniques or filters to remove noise from speech signals, but in order to achieve the best noise reduction, experts use filters with a variety of parameters or combine several different methods. Based on our review, we discussed various techniques for various authors in Table 1, and we will now discuss various techniques such as spectral subtraction and wiener filtering.

- First, speech enhancement is the technique of increasing the quality of a speech transmission by reducing background noise and other unpleasant noises. The clarity, consistency, and comprehension of a voice signal are typically used to determine its quality [59].
- The spectral subtraction method has been found to be a good method but not the best since it produces musical noise. And in [14], the spectral noise removal approach reduces background noise significantly while having minimal influence on speech intelligibility. Formal testing has revealed that at SNR = +5 dB, the improved speech has the same intelligibility as the untreated signal.
- Spectral subtraction in other author comparisons failed; in [22], when compared to Wiener filter and adaptive Wiener filter, SNR results were (spectral subtraction method SNR = 5.0439 dB, Wiener filtering method SNR = 4.9880 dB, adaptive Wiener filtering method SNR = 6.8726 dB) in the time domain. Here, adaptive Wiener filtering showed that it outperformed spectral subtraction and Wiener at both low and high SNR values, and it works in both the additive white Gaussian noise (AWGN) and colored noise scenarios.
- Another researcher uses the Wiener filter [38] when their results show that the suggested method consistently and successfully enhances all forms of noise examined. And the suggested approach produces an improved speech signal that is spectrally comparable to the clean speech signal and perceptually similar to the clean speech signal, where the result of the highest log-likelihood ratio (LLR) was 1.230 in 5 SNR (db).
- Another researcher used the Wiener filter but made some modifications to it, like [43], where they reached the result that approximately 14 to 15% of the SNR ratio increased by using the MNRM compared to the Wiener filter, and the highest SNR was in car noise, about 25.345 SNR (db).
- For more noise reduction, other authors used an adaptive Wiener filter but with TSNR and HRNR [25], and their result was more efficient. To improve the voice signal, two-step noise reduction approaches are applied. Then, harmonic regenerated noise reduction is applied to recreate the harmonics lost in the original signal. The experimental results suggest that employing the TSNR and HRNR methods improves the SNR of the input signal. When compared to TSNR, the HRNR approach produced the highest SNR, where the improvement in SNR (dB) in TSNR was 0.46 and in HRNR was 0.71.
- When FEDS and FAP were compared to classical adaptive filters like LMS, NLMS, AP, and RLS, the authors [19] found that RLS had the highest SNRI (db) at 29.7355. Compared to the LMS, NLMS, and AP algorithms, the RLS method offers a quicker

convergence speed. The FEDS and FAP algorithms outperform the LMS, NLMS, and AP algorithms and are comparable to the RLS algorithm. In another study [24], the authors found the RLS algorithm produces the highest SNR and outperforms LMS in terms of performance. However, with the Finite Impulse Response (FIR) filter taps, LMS converges quicker than RLS.

- When CNN is compared to filters, neural networks have been shown to be an effective technique. as well as CNN, which has been shown to be a viable technique for generalizing real-world noise suppression, and it is noticed that CNN in the training step takes a lot of time to process.
- Finally, from our review of different researchers' work, we notice that the most popular data bases that are used by the researchers are the NOIZUS data bases. This database comprises 30 sentences from the IEEE sentence database spoken by three male and three female speakers, and it is easily available in clean and noisy voices and does not need preprocessing. The Noisex-92 database is also available, but just the noisy file may need preprocessing, as is the TIMIT database. This dataset is not easily available for noisy voices, but in the field of recognition, it is available on many websites.

7. Conclusion

Noise reduction is an interesting and complex field to solve due to the fact that speech enhancement is affected by several types of noise, and there are many algorithms and techniques for noise reduction. In this paper, an overview of several noise reduction methods is discussed and compared with the performance that the other researcher reached with respect to various parameters.

In this review, we concluded that the most efficient filter is the adaptive Kalman filter for both stationary and non-stationary noise. The results show that it is better in white Gaussian (WGN) noise, but performance differs as the noise becomes usable in some cases. The Wiener filter is next, which works best when the noise is stationary. After that comes the adaptive filter algorithm LMS, which works well for low cost, complexity, and increasing the SNR in different noises and in color noise, but the RLS, FEDS, and FAP methods converge faster than the LMS. And then adaptive wiener, which performs well in both colored and additive white Gaussian noise (AWGN) and low and high SNR levels. Then spectral subtraction comes in on the list of good methods, but it produces residual noise, and its shortcoming is the use of noisy phases that produce a roughness in the quality of speech.

And in terms of neural network methods, they're rather good because, when compared to filtering methods, they found that neural networks execute more slowly. The deep learning de-noising method is the most complex and takes a very long time to complete, yet the results are subpar. ADALINE is the best of the neural network methods. The hardest noise to remove from a speech signal is non-stationary noise, and real-world unknown natural noise (mixed noise) is the hardest noise, followed by white Gaussian noise (WGN) and then colored noise.

Acknowledgements

Our sincere gratitude to all researchers in the field of noise reduction for the huge efforts they have given in this field and for helping us to finish this review paper.

References

[1] H. A. Abdulmohsin, "A new proposed statistical feature extraction method in speech emotion recognition," Computers & Electrical Engineering, vol. 93, p. 107172, July. 2021, doi: https://doi.org/10.1016/j.compeleceng.2021.107172.

[2] Haque, M. and Bhattacharyya, K. "Speech Background Noise Removal Using Different Linear Filtering Techniques," *In: Bhattacharyya, S., Gandhi, T., Sharma, K., Dutta, P. (eds) Advanced Computational and Communication Paradigms. Lecture Notes in Electrical Engineering*, Springer, Singapore, vol. 475, pp. 297-307, 2018. https://doi.org/10.1007/978-981-10-8240-5_33.

[3] H. A. Abdulmohsin, B. Al-Khateeb, S. S. Hasan, and R. Dwivedi, "Automatic illness prediction system through speech," Computers and Electrical Engineering, vol. 102, p. 108224, July. 2022, doi: https://doi.org/10.1016/j.compeleceng.2022.108224.

[4] S. Tom, "Why speech is a human innovation," *Knowable Magazine*, February 21, 2019. Available: https://knowablemagazine.org/article/mind/2019/why-speech-human-innovation.

[5] D. Labala, and D. Rajitha, "Implementation of Speech Enhancement Using Combination of Digital Audio Effects with Kalman Filter," *Journal of Engineering Sciences*, Vol. 13, pp. 737-743, July 2022,

[6] D. Soumasunderaswari, and K. Prashanthini, "A survey on various multichannel speech enhancement algorithms," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 4, pp. 254-255, January 2015.

[7] S. Vihari, A. S. Murthy, P. Soni and D.C. Naik, "Comparison of Speech Enhancement Algorithms," *Procedia computer science*, vol. 8, pp. 666-676, June 2016.

[8] S. Lakshmikanth, KR. Natraj, and KR. Rekha, "Noise Cancellation in Speech Signal Processing-A Review," *International Journal of Advanced Research in Computer Communication Engineering*, vol. 3, pp. 5175-5186, January 2014.

[9] V. Sunnydayal, N. Sivaprasad and T.K.K. Kumar, "A Survey on Statistical Based Single Channel Speech Enhancement Techniques," *International Journal of Intelligent Systems and Applications*, vol. 6, pp. 69- 85, November 2014.

[10] A. Rajani and S.V.S. Soundarya, "A Review on Various Speech Enhancement Techniques," *International Journal of Advanced Research in Computer Communication Engineering*, vol. 5, pp. 296-301, August 2016.

[11] T.M.F. Taha, A. Adeel and A. Hussain, "A Survey on Techniques for Enhancing Speech," *International Journal of Computer Applications*, vol. 179, no.17, pp. 1-14, February 2018.

[12] D. JAIN and P. BENIWAL. "Review Paper on Noise Cancellation Using Adaptive Filters," *International Journal of Engineering Research & Technology (IJERT)*, vol. 11, pp. 241-244, January, 2022.

[13] R.H. Frazier, S. Samsam, L.D. Braida and A.V. Oppenheim, "Enhancement of Speech by Adaptive Filtering," in 1976 *ICASSP'76. IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1976, pp. 251- 253.

[14] M. Berouti, R. Schwartz and J. MAKHOUL, "Enhancement of Speech Corrupted by Acoustic Noise," in 1979 *ICASSP'79. IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1979, pp. 208- 211.

[15] T. Langhans and H.W. Strube, "Speech Enhancement by Nonlinear Multiband Envelope Filtering," in 1982 *ICASSP'82. IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1982, pp. 156- 159.

[16] N. Virag, "Single Channel Speech Enhancement Based on Masking Properties of the Human Auditory System," *IEEE Transactions on speech audio processing*, vol. 7, no. 2, pp. 126-137, March 1999.

[17] J. Chen, Y. Huang and J. "Benesty, *Adaptive Signal Processing: Filtering Techniques for Noise Reduction and Speech Enhancement,"* J. Benesty and Y. Huang, Eds. Verlag Berlin Heidelberg: Springer, 2003, pp. 129–154.

[18] X. Zou and X. Zhang, "Speech Enhancement Using an MMSE Short Time DCT Coefficients Estimator with Supergaussian Speech Modeling," *Journal of Electronics*, vol. 24, no.3, pp. 332-337, May 2007. Available: https://doi.org/10.1007/s11767-005-0174-y.

[19] S.A. Hadei and M. Lotfizad, "A Family of Adaptive Filter Algorithms in Noise Cancellation for Speech Enhancement," arXiv, 2011. Available: https://doi.org/10.48550/arXiv.1106.0846.

[20] C. Deng, X. Liu, H. Liu and Z. Wang, "Noisy Speech Enhancement Using a Novel a Priori SNR Estimation," *In: Jin, D., Lin, S. (eds) Advances in Computer Science, Intelligent System and Environment. Advances in Intelligent and Soft Computing*, Springer, Berlin, Heidelberg, vol 105, pp. 139–145, 2011. https://doi.org/10.1007/978-3-642-23756-0_23

[21] C.H. Taal, R.C. Hendriks and R. Heusdens, "A Speech Preprocessing Strategy for Intelligibility Improvement in Noise Based on A Perceptual Distortion Measure," in 2012 *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* 2012. pp. 4061- 4064.

[22] M.A. Abd El-Fattah, M I. Dessouky, A M. Abbas, S M. Diab, E M. El-Rabaie, W. Al-Nuaimy, S.A. Alshebeili and F.E. Abd El-Samie, "Speech Enhancement with an Adaptive Wiener Filter," *International Journal of Speech Technology*, vol. 17, pp. 53-64, August 2013, Available: https://doi.org/10.1007/s10772-013-9205-5.

[23] G.V.P.C.S. Yadav, BA. Krishna and M. Kamaraju, "Performance of Wiener Filter and Adaptive Filter for Noise Cancellation in Real-Time Environment," *International journal of computer applications*, vol. 97, no. 15, pp. 503-511, July 2014.

[24] M.A.J. Sathya, and SP. Victor, "Noise Reduction Techniques and Algorithms for Speech Signal Processing," *International Journal of Scientific Engineering Research*, vol. 6, pp. 317-322, December 2015. Available: https://www.ijser.org/researchpaper/Noise-Reduction-Techniques-and-Algorithms-For-Speech-Signal-Processing.pdf.

[25] O.R. Shruthi and J.C. Saldanha, "Speech Enhancement Using Filtering Techniques," in 2016 *3rd national conference on emerging trends in electronics and communication (NCETEC-16)*, 2016, pp. 1-6.

[26] A. Kumar and D. Florencio, "Speech Enhancement in Multiple-Noise Conditions Using Deep Neural Networks," arXiv, 2016. Available: https://doi.org/10.48550/arXiv.1605.02427 .

[27] Priyanka and. G. Kaur, "Noise Removal in ECG Signal Using Windowing Technique and Its Optimization," *Advances in Biotechnology & Microbiology,* vol. 6, pp. 001-006, August 29, 2017.

[28] H.A. Lafta, "Reduce the Noise in Speech Signals Using Wavelet Filtering," *Journal of University of Babylon, Engineering Sciences*, vol. 26, no. 5, pp. 157-165, March 12, 2018.

[29] K. Bhatt, CS. Vinitha and R. Gupta, "Secure Speech Enhancement Using LPC based FEM in Wiener filter," in *Advances in Intelligent Systems and Computing*, Springer, Singapore, vol. 542, pp. 657–665, 2018. Available: https://doi.org/10.1007/978-981-10-3223-3_63.

[30] N. Dionelis and M. Brookes, "Speech enhancement using Kalman filtering in the logarithmic Bark power spectral domain," in *2018 26th European Signal Processing Conference (EUSIPCO),* 2018. Available: 10.23919/EUSIPCO.2018.8553291.

[31] G.K. Rajini, V. Harikrishnan, M.J P. Priyadarisini and S. Balaji, "A Research on Different Filtering Techniques and Neural Networks Methods for Denoising Speech Signals," *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, vol. 8, pp. 503-511, July 2019.

[32] G. Krishna, C. Tran, Y. Han, M. Carnahan and A.H. Tewfik, "Spoken Speech Enhancement using EEG," arXiv, 2019. Available: https://doi.org/10.48550/arXiv.1909.09132.

[33] H. Yu, W. Zhu and B. Champagne, “Speech Enhancement Using a DNN-Augmented Colored-Noise Kalman Filter,” *Speech Communication*, vol. 125, pp. 142-151, November 4, 2020.

[34] W. Yuan, “A Time–Frequency Smoothing Neural Network for Speech Enhancement,” *Speech Communication*, vol. 124, pp. 75-84, November 2020. Available: https://doi.org/10.1016/j.specom.2020.09.002.

[35] M. Strake, B. Defraene, K. Fluyt, W. Tirry, and T. Fingscheidt, “Speech Enhancement By LSTM-Based Noise Suppression Followed By CNN-Based Speech Restoration,” *EURASIP Journal on Advances in Signal Processing*, vol. 49, pp. 1-26, December 10, 2020.

[36] J.S. Jakati and S.S. Kuntoji, “A Noise Reduction Method Based on Modified LMS Algorithm of Real Time Speech Signals,” *WSEAS transactions on Systems and Control*, vol. 16, pp. 162-170, March 30, 2021.

[37] J. Padmapriya, T. Sasilatha, R. Karthickmanoj, G. Aagash and V. Bharathi, “Voice Extraction from Background Noise using Filter Bank Analysis for Voice Communication Applications,” in *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, 2021, pp. 269-273.

[38] R. Jaiswal and D. Romero, “Implicit Wiener Filtering for Speech Enhancement in Non-Stationary Noise,” in *2021 11th International Conference on Information Science and Technology (ICIST)*, 2021, pp. 39- 47.

[39] S.C. Venkateswarlu, N.U. Kumar and A. Karthik, “Speech Enhancement Using Recursive Least Square Based on Real-Time Adaptive Filtering Algorithm,” in *2021 6th International Conference for Convergence in Technology (I2CT)*, 2021, pp. 1-4.

[40] N. Cohen, G. Hazan, B. Schwartz, and Sh. Gannot, “An Online Algorithm for Echo Cancellation, Dereverberation and Noise Reduction Based on A Kalman-EM Method,” *EURASIP Journal on Audio, Speech, Music Processing*, vol. 21, pp. 1-17, August 28, 2021.

[41] B.M. Mahmmod, SH. Abdulhussain, MA. Naser, M Alsabah, J. Mustafina, “Speech Enhancement Algorithm Based on a Hybrid Estimator,” in *2021 IOP Conference Series: Materials Science and Engineering*, 2021, pp. 1-11.

[42] B.M. Mahmmod, A.R. Ramli, SH. Abdulhussian, SAR. WA. Jassim, “Low-Distortion MMSE Speech Enhancement Estimator Based on Laplacian Prior,” *IEEE Access*, vol. 5, pp. 9866-9881, April 28, 2017.

[43] C.R. Kumar, M.P. Chitra, “Implementation of Modified Wiener Filtering in Frequency Domain in Speech Enhancement,” *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 2, pp. 434-439, 2022, Available: 10.14569/IJACSA.2022.0130251.

[44] A. Kukker, Y. Nigam, T. Sawle, and R. Kumar, “Audio Quality Enhancement Using Adaptive Filters,” *IRE Journals*, vol. 6, pp. 165-169, Jul. 2022.

[45] R.K. Jaiswal, S.R. Yeduri and L.R. Cenkeramaddi, “Single-Channel Speech Enhancement Using Implicit Wiener Filter For High-Quality Speech Communication,” *International Journal of Speech Technology*, vol. 25, pp. 745-758, August 01, 2022.

[46] T. VYACHESLAV, *Signal Processing Noise: Electrical Engineering and Applied Signal Processing Series*. Boca Raton: CRC Press, 2010.

[47] N. H. Resham, H. K. Abbas, H. J. Mohamad, and A. H. Al-Saleh, "Noise reduction, enhancement and classification for sonar images," Iraqi Journal of Science, vol. 62, no. 11, pp. 4439-4452, 2021, doi: 10.24996/ijs.2021.62.11(SI).25.

[48] A. H. Ali and A. M. Al-Rahim, "Linear Noise Removal Using Tau-P Transformation on 3D Seismic Data of Al-Samawah Area-South West of Iraq," Iraqi Journal of Science, vol. 60, no. 12, pp. 2664-2671, 2019, doi: 10.24996/ijs.2019.60.12.16.

[49] Noise News, “What are the 4 Different Types of Noise?,” April 13, 2020. Available: https://www.cirrusresearch.co.uk/blog/2020/04/4-different-types-noise/.

[50] Linked in, "What are the advantages and disadvantages of using a wiener filter for noise reduction in DSP?," July 1, 2023. Available: https://www.linkedin.com/advice/3/what-advantages-disadvantages-using-wiener.

[51] S. Surendran, and T.k. Kumar, "Perceptual Subspace Speech Enhancement With SSDR Normalization," in *2016 International Conference on Microelectronics, Computing and Communications (Microcom),* 2016. Available: https://ieeexplore.ieee.org/abstract/document/7522568.

[52] F. Jabloun and B. Champagne, "Incorporating the Human Hearing Properties in the Signal Subspace Approach for Speech Enhancement," *IEEE Transactions on Speech Audio Processing*, vol. 11, no. 6, pp. 700-708, November 2003.

[53] N.V. Thakor and Y.S. Zhu, "Applications of Adaptive Filtering to ECG Analysis: Noise Cancellation and Arrhythmia Detection," *IEEE Transactions on Biomedical Engineering*, vol. 38, no. 8, pp. 785-794, August 1991.

[54] L. Yang, X. Mingli and T. Yong, "A Noise Reduction Method Based on LMS Adaptive Filter of Audio Signals," in *2013 3rd International Conference on Multimedia Technology (ICMT-13)*, 2013, pp. 1001-1008.

[55] R. Markell, "Better than Bessel" Linear Phase Filters for Data Communications," *McCarthy Blvd, Milpitas,* pp. 1-16, 1994.

[56] M.K. Ibrahem, H.A. Kassim, "VoIP Speech Encryption System Using Stream Cipher with Chaotic Key Generator," *Iraqi Journal of Science,* Special Issue, pp. 240-248, Jan.14, 2021.

[57] E. Abd Alsalam, S.A. Razoqi, E.F. Ahmed, "Effects of Using Static Methods with Contourlet Transformation on Speech Compression," *Iraqi Journal of Science,* vol. 62, no. 8, pp. 2784-2795, Nov.8, 2020.

[58] J. Benesty, J. Chen and YA. Huang, "On Widely Linear Wiener and Tradeoff Filters for Noise Reduction," *Speech Communication,* vol. 52, pp. 427-439, February 3, 2010.

[59] H.S. Hassan, J.S. Harbi, M.A. Kodher, "Hybrid Filter for Enhancing Input Microphone-Based Discriminative Model," *Iraqi Journal of Science,* vol. 61, no. 9, pp. 2434-2439, Jan.21, 2020.