Original scientific paper
https://doi.org/10.33180/InfMIDEM2026.101

Informacije MIDEM
Journal of Microelectronics,
Electronic Components and Materials
Vol. 56, No. 1(2026), 3 – 13

Multitude Active Noise cancellation using White Shark Optimized Deep Learning Network

Jabez Daniel Vincent David Michael[1], Mythili Chandra Sekharan[2], Sheela Yovan[3]

[1]Department of Electronics and Communication Engineering, Dr. Sivanthi Aditanar College of Engineering, Thiruchendur, Tamil Nadu, India
[2]Department of Electrical and Electronics Engineering, University College of Engineering, Nagercoil, Tamil Nadu, India
[3]Department of Computer Science and Engineering, Jayaraj Annapackiyam CSI College of Engineering, Nazareth, Tamil Nadu, India

Abstract: Active noise cancellation (ANC) is an essential feature of audio equipment that reduces unwanted background noise. Unwanted signals in information bearing-signal referred to as noise, could degrade the strength of signals in terms of intelligibility and quality. Over the decade, various researchers developed different algorithms to enhance speech signal quality and for noise reduction. To address the issue, a Multitude Active Noise cancellation using White Shark Optimized Convolutional neural network - Long short-term memory Network (MANC Net) has been proposed. Initially, Dual Tree Complex Wavelet Transform (DTCWT) is utilized to enhance the quality of audio signal with a multitude noise and the signal features are extracted using a community detection based Genetic Algorithm. Afterward based on extracted signal, interference and desired signals are classified using Hybrid Convolutional neural network - Long short-term memory (CNN-LSTM). Additionally, the hyperparameters of CNN-LSTM are tuned using White Shark Optimization (WSO) for better accuracy. The efficiency of the proposed method is evaluated using accuracy, specificity, sensitivity, Normalized Mean Squared Error (NMSE), Short-Time Objective Intelligibility (STOI), and Perceptual Evaluation of Speech Quality (PESQ) parameter values in comparison with other conventional methods. The higher accuracy rate and low NMSE in the classification of audio signals evidenced the efficacy of the proposed MANC Net model. The overall accuracy of the proposed is 9.1%, 8.7%, 7.9%, 3.4%, and 1.5% better than Filtered-X Least Mean Square (FxLMS), deep Active Noise Cancellation (deep ANC), Construction Site Noise Network (CsNNet), Multi-Channel Active Noise Cancellation (MCANC), and Generative fixed-Filter Active Noise Control (GFANC), respectively.

Keywords: Active noise Cancellation; Multitude noise; Deep learning; Optimization

Večkratno aktivno odpravljanje hrupa z uporabo optimizirane mreže za globoko učenje White Shark

Izvleček: Aktivno odpravljanje hrupa (ANC) je bistvena lastnost avdio opreme, ki zmanjšuje neželen hrup v ozadju. Neželeni signali v informacijskem signalu, imenovani hrup, lahko poslabšajo moč signalov v smislu razumljivosti in kakovosti. V zadnjem desetletju so različni raziskovalci razvili različne algoritme za izboljšanje kakovosti govornega signala in zmanjšanje hrupa. Za rešitev tega problema je bila predlagano večkratno aktivno odpravljanje hrupa z uporabo optimizirane konvolucionalne nevronske mreže White Shark – mreže z dolgim kratkoročnim spominom (MANC Net). Sprva se za izboljšanje kakovosti avdio signala z večkratnim hrupom uporabi dvojna drevesna kompleksna valovna transformacija (DTCWT), značilnosti signala pa se izločijo z genetskim algoritmom, ki temelji na zaznavanju skupnosti. Nato se na podlagi izvlečenega signala motnje želeni signali razvrstijo z uporabo hibridne konvolucionalne nevronske mreže – Long short-term memory (CNN-LSTM). Poleg tega se hiperparametri CNN-LSTM prilagodijo z uporabo optimizacije White Shark (WSO) za večjo natančnost. Učinkovitost predlagane metode se oceni z uporabo vrednosti parametrov natančnosti, specifičnosti, občutljivosti, normalizirane srednje kvadratne napake (NMSE), kratkotrajne objektivne razumljivosti (STOI) in zaznavne ocene kakovosti govora (PESQ) v primerjavi z drugimi konvencionalnimi metodami. Višja stopnja natančnosti in nizka NMSE pri razvrščanju avdio signalov sta dokazali učinkovitost predlaganega modela MANC Net. Splošna natančnost predlaganega modela je za 9,1 %, 8,7 %, 7,9 %, 3,4 % in 1,5 % boljša od Filtered-X Least Mean Square (FxLMS), deep Active Noise Cancellation (deep ANC), Construction Site Noise Network (CsNNet) Multi-Channel Active Noise Cancellation (MCANC) in Generative fixed-Filter Active Noise Control (GFANC).

Ključne besede: Aktivno odpravljanje hrupa; Večvrstni hrup; Globoko učenje; Optimizacija

** Corresponding Author's e-mail: Jabezdaniel5433@outlook.com*

How to cite:
J. D. V. D. Michael et al., "Multitude Active Noise cancellation using White Shark Optimized Deep Learning Network", Inf. Midem-J. Microelectron. Electron. Compon. Mater., Vol. 56, No. 1(2026), pp. 3–13

1 Introduction

In Active Noise Cancellation (ANC), two waveforms with the same amplitude, but opposing phases, are superimposed on each other to cancel out the noise [1]. This technique is based on the superposition of acoustic waves [2]. The active noise cancellation model utilizes a microphone for sensing noise and then utilizes a speaker for generating (anti-noise) complementary waveforms incorporated with noise for canceling through destructive interference [3]. Because of its small size and ability to effectively reduce low frequencies, ANC has become widely used in consumer audio products, including headphones [4].

Digital Signal Processor Chips (DSP) were used by the Digital ANC devices to suppress noise adaptively despite changes in the surrounding environment [5]. For noise reduction, the devices modify the finite impulse response (FIR) filter coefficient. The Digital ANC was categorized into feed-back and feedforward types [6]. The digital feed-forward ANC utilizes a single speaker and two microphones (error) [7]. The range of ANC operation frequency level is limited to approximately 600 Hertz. In feedback ANC, for maintaining loop stability however it is 1500 Hertz in feed-forwards ANC in the case of earphones [8-10]. Nevertheless, obtaining reference noise becomes challenging, if not impossible, if the noise source is moving within the environment [11].

Multitude noise refers to the aggregate or combined noise generated from multiple sources, creating a complex and often unpredictable auditory environment [12]. Despite the advancements in ANC technology, addressing a multitude noise in broader environments remains a complex challenge [13]. Traditional ANC systems are highly effective in controlled settings, like inside headphones, but their effectiveness diminishes in open and variable environments. Future developments in ANC technology aim to enhance its capability to adapt to dynamic noise environments, potentially incorporating advanced algorithms, machine learning, and improved sensor [14] technologies to better analyze and counteract a multitude noise in diverse settings [15]. In this research, a novel Multitude Active Noise cancellation using White Shark Optimized Convolutional neural network - Long short-term memory Network (MANC Net) has been proposed. The major contribution of the proposed method is:

- The primary objective of the research is to develop a Multitude active noise cancellation model-based deep learning techniques to classify the input signal.
- In the proposed method, Dual tree complex Wavelet transform (DTCWT) is utilized to enhance the quality of audio signal with multitude noise, and the signal features are extracted using a community detection based Genetic Algorithm.
- Convolutional neural network - Long short-term memory (CNN-LSTM) is introduced in the proposed model to classify desired signals and interference signals. Additionally, the hyperparameters of CNN-LSTM are tuned using White Shark optimization for better accuracy.
- The efficiency of the MANC Net is evaluated using accuracy, specificity, sensitivity, NMSE, STOI, and PESQ parameter values in comparison with other conventional methods.

The section that follows outlines the paper's organization. Section II provides a summary existing ANC model. Section III provides a detailed list of all the introduced algorithms, which form the basis of the proposed system. The results of the implementation are shown in Section IV. Section V concludes with an overall summary of the system.

2 Literature survey

This section lists reviews of previous studies that examine various active noise-canceling techniques and their various methodologies.

An approach for deep learning-based image processing that uses random search was proposed by Huber, N.R., et al. in 2021 [16]. Compared to the low-dose input picture, the CNN improved SSIM by 76%, reduced RMSE by 34%, and reduced noise by 90% on six clinical examinations that were set aside for testing. An algorithm called Filtered-X Least Mean Square (FxLMS) was proposed in 2022 by Bisht, A. and Patil, H.Y. [17] to address noise produced by motor vehicles, particularly those with combustion engines. To make things simple and easy to use, their source noise signal is narrowband combustion-related white noise that is generated at random while running simulations.

A selective fixed-filter ANC approach based on a simplified two-dimensional Convolutional neural network (2D CNN) was proposed by Shi, D., et al. in 2022 [18] to determine the best control filter for various noise kinds. The effectiveness of the suggested approach in reducing non-stationary noise in the actual world is demonstrated by a numerical simulation that utilizes measurable routes in headphones, as opposed to traditional adaptive algorithms. Zhang, H. and Wang, D. [19] suggested a deep Active Noise Cancellation (deep ANC) technique for active noise reduction in 2021. Methodical assessments employing STOI, NMSE, and PESQ show that the deep ANC model for noise attenuation is reliable and efficient.

To reduce signal noise, Arab, H., et al. (2022) designed a Recurrent Neural Network (RNN) using an LSTM autoencoder architecture. Compared to current methods, the accuracy of 99.93% in classifying the demodulation classes is comparatively greater. Cha, Y.J., et al. (2023) suggested DNoiseNet, a sophisticated DL-based feedback ANC that addresses acoustic delay and other primary and secondary path effects. Through comprehensive parametric and comparative analyses, the suggested approach produces optimal results in terms of root mean square error and noise attenuation metrics.

Deep learning-based active noise cancelation named Construction Site Noise Network (CsNNet) was created in 2023 to reduce various noises found on construction sites, as proposed by Mostafavi, A. and Cha, Y.J. [22]. Modern and conventional ANC algorithms were surpassed by the network that was exhibited. Zhang, H. and Wang, D. (2023) suggested a deep learning-based technique for Multi-Channel Active Noise Cancellation (MCANC) [23]. According to experimental findings, deep MCANC performs well in generalizing to untrained noises and is effective in reducing wideband noise.

A Generative Fixed-filter Active Noise Control based -Kalman technique was proposed by Luo, Z., et al. in 2023 [24]. The suggested method's effectiveness in handling dynamic noises in the real world is confirmed by numerical simulations. An approach for reducing noise, known as Generative Fixed-filter Active Noise Control (GFANC), was proposed by Luo, Z., et al. in 2023 [25]. This approach performs especially effectively when the incoming noise differs noticeably from the pre-training stimuli. The efficacy of the GFANC method is shown by numerical simulations performed on real recorded noises.

Based on the literature review, a number of deep learning techniques were proposed to improve speech understanding in noisy environments. However, issues with ANC enhancement systems still include unreliable voice inputs, high computation costs, and the presence of background noise. To overcome these challenges a novel a novel Multitude Active Noise cancellation using White Shark Optimized CNN-LSTM Network (MANC Net) has been proposed.

3 Proposed MANC net

In this section, a novel Multitude Active Noise cancellation using White Shark Optimized CNN-LSTM Network (MANC Net) has been proposed. The overall block of the proposed MANC Net is depicted in Figure 1.

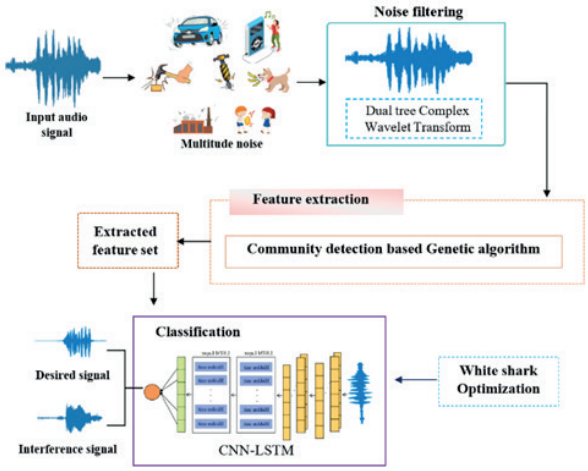


Figure 1: Overall block of proposed MANC Net

3.1 Dataset Description

In this research Audio benchmark dataset is utilized, it comprises 8732 labeled sound excerpts (lesser than or equal to 4s) of various urban sounds belonging to ten classes like car_horn, gun_shot, children-playing, air-conditioner, jackhammer, street music, siren, drilling, engine_idling and dog-bark. The files were pre-sorted to 10 folds like fold-1 to fold-10 aiding in features classification outcomes. Those 8732 audio files consisting of urban sounds are represented in Waveform Audio File Format (WAV format).

3.2 Noise Filtering via DTCWT

Audio signal exhibiting nonlinear characteristics that pose challenges for conventional filter denoise removal. Wavelet transform [26] is widely recognized for its ability to analyze signals at multiple resolutions and extract time-frequency features effectively, it has certain disadvantages, such as sensitivity to noise in some applications, difficulty in selecting the optimal wavelet basis for specific tasks, and higher computational complexity [27]. Despite this limitation, the Dual tree complex Wavelet transform (DTCWT) [28] excellent performance and qualitative qualities have made it known to denoise these non-stationary signals.

Two genuine filter banks make up the DTCWT technology (low pass and high pass filter pair). The real and even component of a complex wavelet is provided by the first Discrete Wavelet Transform (DWT) in this approach, while the real and odd or imaginary part is provided by the second DWT. Perfect Reconstruction (PR) requirements are satisfied by each of the two sets of filters used in these two real DWTs. Equation (1) expresses DTCWT as an approximate analytic wavelet.

$$\psi(t) = \psi_q(t) + x\psi_v(t) \quad (1)$$

The values of $\psi_q(t)$ are real and even, but the values of $\psi_v(t)$ are real and odd. The present study utilizes the DTCWT technique to decompose the input audio signal into four levels of detail and approximation coefficients, as depicted in Figure 2.

In contrast to approximate coefficients, which are low-frequency coefficients unaffected by noise, detail coefficients are high-frequency coefficients that are typically influenced by noise. Finally, the audio signal is Enhanced by reducing noise and preserving quality.

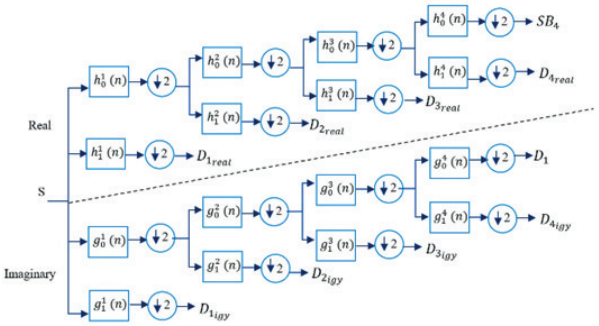


Figure 2: Four Levels of DTCWT technique

3.3 Feature extraction via community based genetic algorithm

The filtered signals are fed into the feature section phase to extract the most relevant features using a community based genetic algorithm [29]. The Community Detection-based Genetic Algorithm for Feature Extraction (CDGAFE) consists of the following primary steps:

Step 1: Measure the relevance of features
The relevance of features was measured. To measure the feature's discriminatory power, the discrimination capability of the feature F_{ii} are evaluated by employing the Fisher score as below in the equation

$$\text{score}_{ii} = \frac{\sum_{k=1}^{C\_i} n_{ii}\left(x_{ii}^{k} - x_i\right)^2}{\sum_{k=1}^{c_i} n_{ii}\left(\sigma_i^k\right)^2} \quad (2)$$

Wherein this C_l denotes the number of dataset classes; N_{1i} denotes the sample count within the classification. the variable x_i indicates the mean of whole patterns in accordance to feature F_{ii}, X_{kl} and σ_k denotes mean of class and variance of class K_{ii} in correspondence with the feature F_{ii}

Larger Score_{ii} value reveals that the feature F_{ii} consists of high discriminative ability. In many instances, features of the fisher score were near each other. In conquering this circumstance, the non-linear normalization method referred as Soft_{max} scaling employed in scaling edge weight into the range [0 1] represented by the equation (3)

$$\text{Score}_{ii} = \frac{1}{1 + \exp\left(-\frac{\text{Score}_{ii} - \text{Score}_1}{\sigma_i^k}\right)} \quad (3)$$

In this equation Score_{ii} implies features F_{ii} fisher score and σ and Score_1 denotes the variance and mean of the entire fisher-score and Score_{ii} reveals normalized fisher-score of features F_{ii}.

Step 2: Feature clustering
For applying a feature-clustering algorithm, the feature's similarity ought to be determined. As a graph-based clustering method, feature space denotes to be a graph. For this mapping of features to their equivalent graph G = (F_i, E_i, W_f) are performed, wherein this F_i = {F_{i1}, F_{i2}, … … ...} represents the group of original-features features E = {(F_{ii}, F_{jj}):$F_{ii}F_{jj}\varepsilon F_{ii}$} were the graph edges and similarity between feature F_ii and F_jj is denoted by W_1i1j and they are connected by an edge (F_i, F_j). The relationship among those features F_{ii} and F_{jj} are defined below equation

$$W_{ij} = \frac{\sum_p \left(X_{ii} - X_{ii1}\right)\left(X_{jj} - X_{jj}\right)}{\sqrt{\sum_p \left(X_{ii} - X_{i1}\right)^2}\sqrt{\sum_p \left(X_{jj} - X_{ijl}\right)^2}} \quad (4)$$

Wherein this variable X_{ii} and X_{jj} denotes the vectors of those features F_{ii} and F_{jj} and X_{ii1} and X_{ii1} implies mean vectors X_{ii} and X_{jj} values. Likewise, towards fisher scores, entire similarity-values were normalized through softmax-scaling.

Step 3: Initialization of Population
The set of chromosome populations has been initialized randomly. The count of original features n, will be equal to every length of chromosome. The total selected features count in every chromosome should be K * ω such that *k* denotes cluster count and the user-specified parameters are denoted by ω that controls the final feature subset size

Step 4: Calculation of Fitness values
Once the initial population is created, the fitness function score for entire features should be computed. For this objective, the multi-fitness function was been introduced. The fit of feature subset FS_i^k within an iteration (t) represented through $J\left(F_s^k(t)\right)$ and it is measured through equation (5)

$$J\left(F_s^k\left(t\right)\right)=\frac{CA(FS^k\left(t_i\right)}{\frac{2}{F_{sk}\left(t\right)*F_s^k\left(t\right)-1\mid}} \quad (5)$$

In this equation, the extraction accuracy of chosen features-subset $FS_i^k\left(t\right)$ indicated by CA ¬$FS_i^k\left(t\right)$ and the selected features FS_i^k(t) denotes the size of the subset. The similarity among the features F_{ii} and F_{jj} is stated by variable FS_i^k(t).

Step 5: Perform phase Crossover operation and Mutation operation

New chromosomes or features are generated through cross-over operators and mutation operators. This chromosome gene proceeds predefined mutation probability, if or not to get selected for mutation.

Step 6: Performing Repair Operation

The repair process is used to recalculate the number of features selected from each group among all features developed. If the number of randomly chosen features in a cluster is less than this ω_i, then a single feature is chosen at random and the associated feature is changed to a value of 1.

Step 7: Termination Criterion

In suppose a case if it has the count of iteration, that is higher than the maximum allowable iteration counts, it will continue or else another step is proceeded in the calculation of fitness score values.

Step 8: Finalisation of Feature Subset Selection

Under the fitness values of the feature, the strongest chromosome of the last hyperparameters, which had higher stronger fisher score values, implies the optimized selection of features subset, that paves the higher accuracy and performance in denoise the noise signals from features extracted signals.

3.4 CNN-LSTM based classification

The CNN-LSTM receives the extracted features and uses them to distinguish between the interference signal and the intended signal. To classify and denoise audio signals, CNN and LSTM are integrated into the proposed CNN-LSTM algorithm.

3.4.1 Convolutional neural network (CNN)

In addition to the input and output layers, the CNN [30] architecture also includes several hidden layers. Several convolution layers make up these hidden layers. Convolution Layers are composed of many kernels. The activation function described by act(.), similarly pooling(.) denoted by the pooling function, the local neighborhood is represented by $R_{i,j}$. The down-sampling process in the training phase in each layer's feature maps was computed as below.

$$\mathrm{yf}_{\mathrm{i,j,k}}=\mathrm{pooling}\left(\mathrm{act}\left(\mathrm{Zf}_{\mathrm{i,j,k}}\right)\right)\forall\left(\mathrm{m,n}\right)\ \epsilon\ \mathrm{R}_{\mathrm{i,j}} \quad (6)$$

The fully-connected layer might get placed after this convolutional layer and then the pooling layer. In any classification issues of features to categorize as desired and interference signal, this can be handled through the softmax function in the CNN model and is generally employed in CNN output later (last layer). This loss could also be retrieved by computation below,

$$L=\sum_{\mathrm{n=1}}^{\mathrm{N}}\mathrm{l}\left(\theta;\ \mathrm{y}^{(\mathrm{n})},\mathrm{o}^{(\mathrm{n})}\right)\ \mathrm{n}\ \epsilon\ \left[1....\mathrm{N}\right]\frac{1}{\mathrm{N}} \quad (7)$$

Wherein the n^{th} input-data denoted by $x^{(n)}$ and the real target label of n^{th} input-data. Similarly, the nth output of classification by CNN is denoted by $o^{(n)}$ and θ points out all parameters. The pooling layer further generates a lower dimension matrix as output after receiving the value of the organized layer as input. Lastly, the signal is classified by the fully connected layer using the pooling layer's output.

3.4.2 LSTM

With the use of feedback connections, the LSTM [31] layer can learn long-term dependencies. Three primary gates input, output, and forget combined with a memory cell to form an LSTM layer. By using this design, the long-term dependency maintenance mechanism (LSTM) may determine which data to "remember" and which to "forget."

$$fg_t=\sigma(W_{fg}[hs_{t-1}/\ p_t\ ]+u_{fg} \quad (8)$$

$$ip_t=\sigma(W_{ip}[hs_{t-1}/\ p_t\ ]+u_{ip} \quad (9)$$

$$\widetilde{cd_t}=tanh(W_{ch}[hs_{t-1}/\ p_t\ ]+u_{ch} \quad (10)$$

$$cd_t=ip_t\times\widetilde{cd_t}+fg_t\times cd_{t-1} \quad (11)$$

$$op_t=\sigma(W_{op}[hs_{t-1}/\ p_t\ ]+u_{op} \quad (12)$$

$$hs_t=op_t\times\tanh\left(cd_t\right) \quad (13)$$

Hence, at time t, the states of the output gate, input gate, and forget gate respectively, fg_t, ip_t, op_t. Each component has weight matrices denoted by W_{fg}, W_{ip}, W_{ch}, W_{op}, and bias vectors denoted by υ_{fg}, υ_{ip}, υ_{ch}, υ_{op}. $\widetilde{cd_t}$ is the memory cell's candidate state value at time t, as determined by the tanh function. s is the sigmoid function, and c_{dt} is the memory cell state at time *t*.

3.4.3 Hybrid CNN-LSTM

The combination of CNN and LSTM for signal denoise reduction in audio classification. Figure 3 shows the proposed hybrid model for active noise cancellation in categorization. The system consists of twenty layers: one fully connected layer, one LSTM layer, five pooling layers, twelve convolutional layers, and one output layer that uses the softmax function. Beyond that, there is a dropout layer with a 25% dropout rate. Features are extracted using a convolutional layer with a 3×3 kernel size, which is initiated by the ReLU function. The convolution section (none, 7, 7, 512) is followed in determining the final shape. The input size of the LSTM layer has been lowered by the reshaping method (49, 251). The structure classifies the spectrogram images into desired and interference signal categories by first analyzing the temporal properties and then sorting the images through a completely linked layer. In Figure 3, the CNN-LSTM architecture.

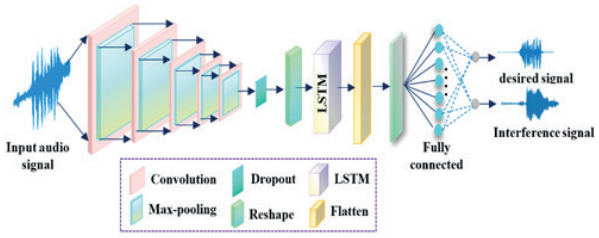


Figure 3: Architecture of CNN-LSTM

3.4.3.1 Hyper parameter turing via white shark optimization (WSO)

An intelligent metaheuristic model with practical applications, the White Shark Optimizer (WSO) [32] can solve a wide range of optimization problems in a continuous search domain. Importantly, the collective hunting strategies, swimming capability, and highly developed auditory and olfactory capabilities utilized in prey identification all have an impact on the WSO algorithm. Equation (16) can be used to pinpoint the position of a white shark:

$$s=\begin{bmatrix} s_1^1 & s_2^1 & \cdots & s_l^1 \\ s_1^2 & s_2^2 & \cdots & s_l^2 \\ \vdots & \vdots & \vdots & \vdots \\ s_1^x & s_2^x & \cdots & s_l^x \end{bmatrix} \tag{14}$$

The p^{th} spot of the white shark to the p^{th} dimension is indicated by s_q^p. The algorithm searches regions in the q^{th} dimension can be computed as follows by utilizing the upper u_{bq} and lower lb_q Limits:

$$s_q^p = lb_p + rand \times \left(ub_q - lb_q\right) \tag{15}$$

where a random number within the range of [0, 1] is indicated by *rand*. Here, the starting fitness values for the first solutions given by Equation (16) can be calculated.

$$r_{k+1}^p = \mu\left(r_k^p + m_1\left[s_{gbest_k} - r_k^p\right]\times b_1 + m_1\left[s_{best}^{r_k^p} - s_k^p\right]\times b_2\right) \tag{16}$$

In iterations $k + 1$ and k, the improved speed of the p^{th} white sharks are represented by r_{k+1}^p and r_k^p, accordingly. The white shark influences on s_{gbest_k} and $s_{best}^{r_k^p}$ are symbolized by m_1 and m_2. While s_{gbest_k} and $s_{best}^{r_k^p}$ provide the optimal global position at the K^{th} iteration, s_k^p indicates the location of the p^{th} white sharks in repetition k. White shark convergence behavior is analyzed using the WSO constriction factor, μ, and the parameter $s_{best}^{r_k^p}$, which represents the i th best-defined position of the swarm throughout the frequency *k* iteration process.

$$s = \left[x \times rand\left(1, x\right)\right] + 1 \tag{17}$$

where the random integers in the vector *rand*(1, *x*) are between 0 and 1. Utilizing the following relationships, the variables b_1 and b_2 provided in Equation (18) can be computed:

$$b_1 = b_{max} + \left(b_{max} + b_{min}\right) \times e^{-\left(\frac{4k}{K}\right)^2} \tag{18}$$

$$b_2 = b_{min} + \left(b_{max} + b_{min}\right) \times e^{-\left(\frac{4k}{K}\right)^2} \tag{19}$$

where the highest and lowest velocities of movement for great white sharks are denoted by b_{max} and b_{min}. A 1.5 and a 0.5 are assumed for b_{max} and b_{min} in this situation. Based on the following formula (22) the white shark position is updated.

$$s_{k+1}^p = \begin{cases} s_k^p \cdot \neg \oplus \tau_0 + ub \cdot a + lb \cdot a & if\ rand < mv \\ s_k^p + \dfrac{s_k^p}{g} & if\ rand \geq mv \end{cases} \tag{20}$$

where τ_0 is a logical vector, *ub* and *lb* are the upper and lower limits of the search space, *a* and *b* represent for binary vectors, *g* represents the white shark's wavy motion frequency, and $\neg$ is the negation operator. One approaches the best shark in the following manner:

$$s_{k+1}^{,p} = s_{gbest_k} + v_1 \overrightarrow{H_s} \times sgn\left(v_2 - 0.5\right) if\ v_3 < W_s \tag{23}$$

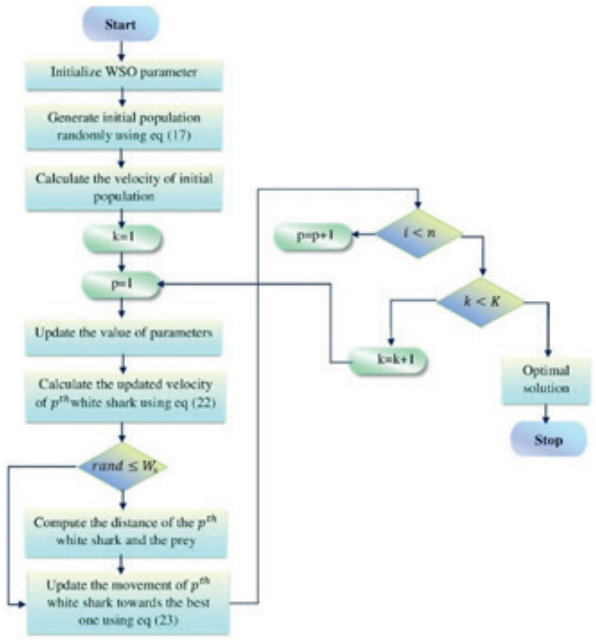


Figure 4: Flow chart of White Shark Optimization

Where v_1, v_2, v_3, are random numbers between 0 and 1, and S_{k+1}^{p} represents the new location of the i th white shark to the prey. Given that it returns either 1 or −1, $sgn(v_2 - 0.5)$ can be used to reverse the search's direction. H_s indicates how far away the prey is from the white shark, while W_s describes how strong the shark's sight and smell senses are. The flowchart provided in Figure 4 illustrates the shark optimizer's operating principle.

4 Result and discussion

The efficiency of the proposed MANC Net method is assessed in its performance through comparative deep learning-based ANC methods. To develop the proposed framework's performance evaluation by comparing its parameter values for accuracy, specificity, sensitivity, NMSE, STOI, and PESQ to those of other traditional techniques.

Figure 5 represents the input signals and the outcomes of the resultant denoised signals after the active noise cancellation process. The input signals comprising different intensities of waveforms are subjected to the proposed active noise cancellation process using hybrid CNN-LSTM, the outcomes of classified outcomes yield precise text voice signals with no interference of noise signals in the outcome end.

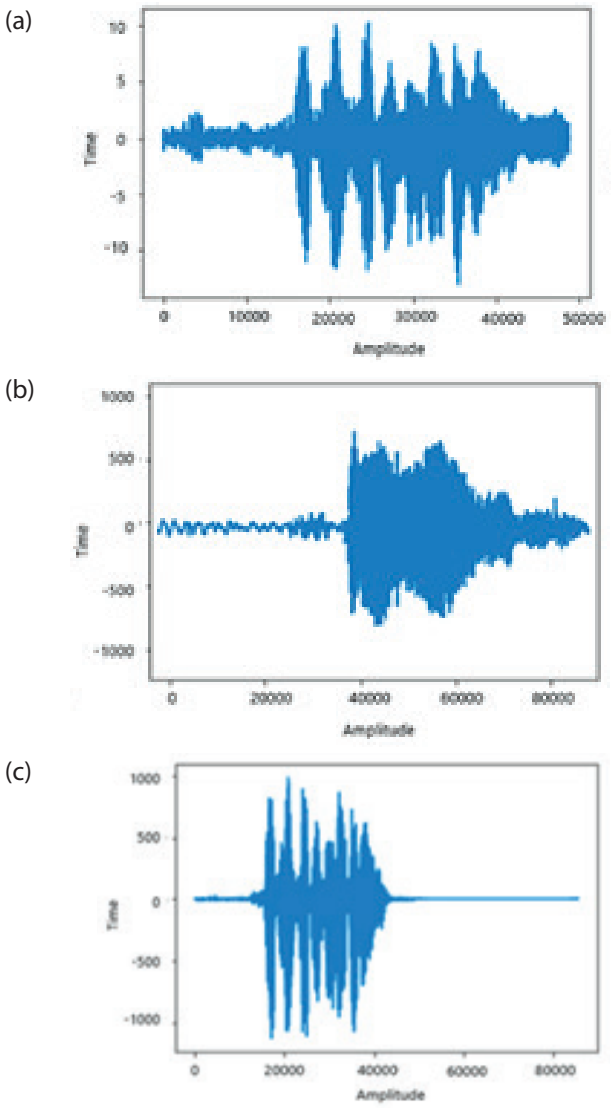


Figure 5: (a) Input signal (b) Interference signal (c) Desired Signal

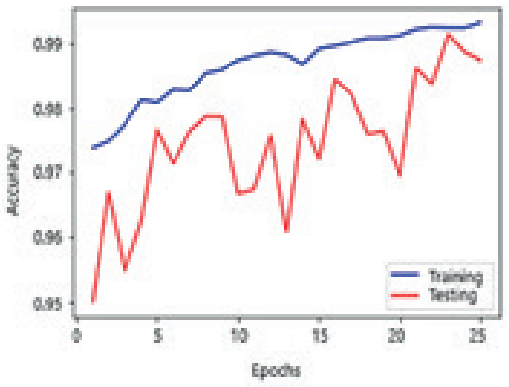
4.1 Performance metric

The results enumerated the performance assessment of the proposed method in terms of accuracy, specificity, sensitivity, NMSE, STOI, and PESQ with other noise cancelation systems.

Short-Time Objective Intelligibility (STOI): The correlation coefficient between the temporal envelopes of time-aligned reference signals and processed signals inside shorter time overlapping segments served as the basis for the STOI metric.

Perceptual Evaluation of Speech Quality (PESQ): The PESQ metric refers to a set of standards that include test methods for automatically evaluating speech quality based on connections made through a calling system's user experience.

(a)



(b)

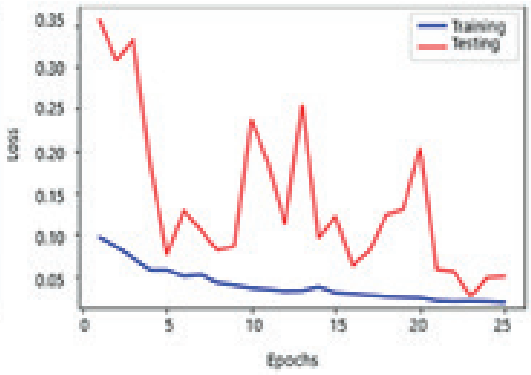


Figure 6: Training and testing of proposed method (a) Accuracy (b) Loss

Normalized Mean Squared Error (NMSE): The MSE measure that has been normalized by signal power is called the NMSE. In the ANC model, the error-signal power was typically employed as a quality parameter for noise attenuation.

Figure 6 (a) explicates the accuracy of the training set and the validation set if the classified outcomes are brought out perfectly. The accuracy of the training results in denoise of audio signals produced the classification outcomes at a higher rate of accuracy. Figure 6 (b) presents the loss of signals while performing the ANC process. The loss attained in the training phase seems to be lower in comparison to the validation loss.

Table 1: PESQ estimation for different noise

NOISE	PESQ (0 dB)	PESQ (5 dB)	PESQ (10 dB)	PESQ (15 dB)
Car	4.1578	4.1001	4.1481	4.1785
Children playing	4.1485	4.1675	4.0458	4.2475
Drilling	4.1078	4.1369	4.1785	4.2785
Street music	4.2745	4.1284	4.2758	4.2775
Dog bark	4.1201	4.1234	4.2785	4.2845

Table 1 shows the PESQ scores for different noise types at various input PESQ levels: 0 dB, 5 dB, 10 dB, and 15 dB. At 10 dB, dog barking sound (4.2785) and Street noise (4.2758) lead to performance. At 15 dB, dog barking noise achieves the highest PESQ score (4. 2845), followed by drilling noise (4. 2785). Overall, the proposed method performs best under Street noise at 0 dB, and dog barking noise at 15 dB.

4.2 Comparative analysis

The effectiveness of the proposed MANC Net approach was assessed in comparison to that of other approaches to demonstrate its accuracy and efficiency. The proposed MANC Net approach is compared with the current techniques, including Custom FxLMS [17], deep ANC [19], CsNNet [22], MCANC [23], and GFANC [25].

Table 2: Performance of Proposed ANC model

Models	NMSE	STOI	PESQ
Unprocessed	0	0.79	1.95
Custom FxLMS [17]	-4.54	0.71	1.84
deep ANC [19]	-6.55	0.69	1.73
CsNNet [22]	-10.6	0.72	1.71
MCANC [23]	-10	0.84	2.26
GFANC [25]	-11	0.88	2.30
Proposed	-12	0.92	2.35

Table 2 depicts the numerical values of STOI, PESQ, and NMSE metrics values of existing Custom FxLMS [17], deep ANC [19], CsNNet [22], MCANC [23], and GFANC [25] methods and it compares with proposed MANC Net method. The proposed method possesses a higher PESQ value, (2.35) and a higher STOI value (0.92) revealing the higher performance in denoise behavior. The least values of NMSE (-12) values are gained for proposed active noise cancellation techniques.

Table 3: Comparative analysis of proposed model classifier for different datasets

Methods	ANN	RNN	CNN	LSTM	Proposed Optimized CNN-LSTM
ESC-10	70.15	72.15	78.31	82.15	91.32
ESC-50	72.36	61.54	60.1	85.45	92.35
UrbS8K	79.08	83.7	86.15	89.45	98.56

Table 3 represents the noise cancellation outcomes from input signals of the proposed Hybrid CNN-LSTM classification approach in terms of accuracy. The outcomes in denoising of input signals, from different datasets such as ESC-10, ESC-50, and UrbanSound8K dataset. For the UrbanSound8K dataset, different clas-

sifiers yield higher accuracy in classification for the Optimized CNN-LSTM classifier (98.56%).

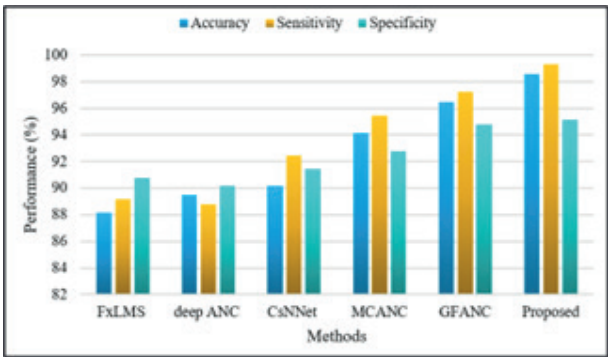


Figure 7: Performance analysis of proposed with existing methods

Figure 7 shows the comparative analysis performed in assessing the classification accuracy of the proposed MANC Net model with other FxLMS [17], deep ANC [19], CsNNet [22], MCANC [23], and GFANC [25]. From the figure, it explicates that the proposed noise cancellation method, shows higher training accuracy (98.56%), higher specificity (95.12%), and high sensitivity (99.32%) than other existing ANC methods. The overall accuracy of the proposed is 9.1%, 8.7%, 7.9%, 3.4%, and 1.5% better than FxLMS, deep ANC, CsNNet, MCANC, and GFANC respectively.

5 Conclusion

In this research, Multitude Active Noise cancellation using White Shark Optimized CNN-LSTM Network (MANC Net) has been proposed. Initially, Dual tree complex Wavelet transform is utilized to enhance the quality of audio signal with multitude noise and the signal features are extracted using community detection based Genetic Algorithm. Afterwards based on extracted signal, interference and desired signals are classified using hybridized Convolutional neural network - Long short-term memory (CNN-LSTM). Additionally, the hyper parameters of CNN-LSTM are tuned using White Shark optimization for better accuracy. The efficiency of the proposed method is evaluated using accuracy, specificity, sensitivity, NMSE, STOI and PESQ. The proposed method possesses higher PESQ value, (2.35) and higher STOI value (0.92) reveals the higher performance in denoise behaviour. The overall accuracy of the proposed is 9.1%, 8.7%, 7.9%, 3.4%, and 1.5% better than FxLMS, deep ANC, CsNNet, MCANC, and GFANC respectively.

6 Acknowledgments

The author would like to express his heartfelt gratitude to the supervisor for his guidance and unwavering support during this research for his guidance and support.

7 Conflict of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

8 References

1. D. Shi, W.S. Gan, B. Lam and S. Wen, "Feedforward selective fixed-filter active noise control: Algorithm and implementation", IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 28, pp. 1479-1492, 2020. https://doi.org/10.1109/TASLP.2020.2989582
2. B. Lam, W.S. Gan, D. Shi, M. Nishimura and S. Elliott, "Ten questions concerning active noise control in the built environment", Build. Environ., vol.200, pp. 107928, 2021. https://doi.org/10.1016/j.buildenv.2021.107928
3. T. Li, M. Wang, Y. He, N. Wang, J. Yang, R. Ding, and K. Zhao. "Vehicle engine noise cancellation based on a multi-channel fractional-order active noise control algorithm", Mach., vol. 10, no. 8, pp. 670, 2022. https://doi.org/10.3390/machines10080670
4. H. Zhang, and D. Wang, "A Deep Learning Approach to Active Noise Control", In Interspeech, vol. 16, no. 3, pp. 141-1145, 2020. http://dx.doi.org/10.21437/Interspeech.2020-1768
5. C.Y. Ho, K.K. Shyu, C.Y. Chang and S.M. Kuo, "Efficient narrowband noise cancellation system using adaptive line enhancer", IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 28, pp. 1094-1103, 2020. https://doi.org/10.1109/TASLP.2020.2982578
6. B. Lam, W.S. Gan, D. Shi, M. Nishimura and S. Elliott, "Ten questions concerning active noise control in the built environment", Build. Environ., vol. 200, pp. 107928, 2021. https://doi.org/10.1016/j.buildenv.2021.107928
7. C.Y. Chang, A. Siswanto, C.Y. Ho, T.K. Yeh, Y.R. Chen and S.M. Kuo, "Listening in a noisy environment: Integration of active noise control in audio products", IEEE Consum. Electron. Mag., vol. 5, no. 4, pp. 34-43, 2016. https://doi.org/10.1109/MCE.2016.2590159

8. F. Liu, X. Zhao, Z. Zhu, Z. Zhai and Y. Liu, "Dual-microphone active noise cancellation paved with Doppler assimilation for TADS", Mech. Syst. Signal Process., vol. 184, pp. 109727, 2023. https://doi.org/10.1016/j.ymssp.2022.109727
9. L. Jenifer, X. Cheng, A. Ahilan, and P.J. Shermila, "A Novel Internet Of Things-Based Electrocardiogram Denoising Method Using Median Modified Weiner And Extended Kalman Filters".
10. C. Shi, F. Du and Q. Wu, "A digital twin architecture for wireless networked adaptive active noise control", IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 30, pp. 2768-2777, 2022. https://doi.org/10.1109/TASLP.2022.3199992
11. D. Shi, W.S. Gan, B. Lam and S. Wen, "Feedforward selective fixed-filter active noise control: Algorithm and implementation", IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 28, pp. 1479-1492, 2020. https://doi.org/10.1109/TASLP.2020.2989582
12. Y.X. Li, and L. Wang, "A novel noise reduction technique for underwater acoustic signals based on complete ensemble empirical mode decomposition with adaptive noise, minimum mean square variance criterion and least mean square adaptive filter", Def. Technol., vol. 16, no. 3, pp. 543-554, 2020. https://doi.org/10.1016/j.dt.2019.07.020
13. F. Yang, J. Guo, and J. Yang, "Stochastic analysis of the filtered-x LMS algorithm for active noise control", IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 28, pp.2252-2266, 2020. https://doi.org/10.1109/TASLP.2020.3012056
14. M.A. Stebi, and A. Jeyam, "Estimation of household appliances and monitorization for impact reduction using electro chemical sensor," *International Journal of System Design and Computing*, vol.*1*, no. 01, pp.26-34, 2023.
15. Y. Zhu, H. Zhao, X. Zeng, and B. Chen, "Robust generalized maximum correntropy criterion algorithms for active noise control", IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 28, pp. 1282-1292, 2020. https://doi.org/10.1109/TASLP.2020.298203
16. N.R. Huber, A.D. Missert, H. Gong, S.S. Hsieh, S. Leng, L. Yu, and C.H. McCollough, "Random search as a neural network optimization strategy for Convolutional-Neural-Network (CNN)-based noise reduction in CT", In Medical Imaging 2021: Image Processing, vol. 11596, pp. 509-515, 2021. https://doi.org/10.1117/12.2582143
17. Bisht, and H.Y. Patil, "Active Noise Cancellation System in Automobile Cabins Using an Optimized Adaptive Step-Size FxLMS Algorithm", In Applied Information Processing Systems: Proceedings of ICCET 2021 Springer Singapore, pp. 187-200, 2022. https://doi.org/10.1007/978-981-16-2008-9_18
18. D. Shi, B. Lam, K. Ooi, X. Shen and W.S. Gan, "Selective fixed-filter active noise control based on convolutional neural network", Signal Process., vol. 190, pp. 108317, 2022. https://doi.org/10.1016/j.sigpro.2021.108317
19. H. Zhang and D. Wang, "Deep ANC: A deep learning approach to active noise control", Neural Networks, vol. 141, pp. 1-10, 2021. https://doi.org/10.1016/j.neunet.2021.03.037
20. H. Arab, I. Ghaffari, R.M. Evina, S.O. Tatu, and S. Dufour, "A hybrid LSTM-ResNet deep neural network for noise reduction and classification of V-band receiver signals", Ieee Access, vol. 10, pp. 14797-14806, 2022. https://doi.org/10.1109/ACCESS.2022.3147980
21. Y.J. Cha, A. Mostafavi and S.S. Benipal, "DNoiseNet: Deep learning-based feedback active noise control in various noisy environments", Eng. Appl. Artif. Intell., vol. 121, pp. 105971, 2023. https://doi.org/10.1109/ACCESS.2022.3147980
22. Mostafavi and Y.J. Cha, "Deep learning-based active noise control on construction sites", Autom. Constr., vol. 151, pp. 104885, 2023. https://doi.org/10.1016/j.autcon.2023.104885
23. H. Zhang and D. Wang, "Deep MCANC: A deep learning approach to multi-channel active noise control", Neural Networks, vol. 158, pp. 318-327, 2023. https://doi.org/10.1016/j.neunet.2022.11.029
24. Z. Luo, D. Shi, X. Shen, J. Ji, and W.S. Gan, "Gfanc-kalman: Generative fixed-filter active noise control with cnn-kalman filtering", IEEE Signal Process Lett., 2023. https://doi.org/10.1109/LSP.2023.3334695
25. Z. Luo, D. Shi, X. Shen, J. Ji and W.S. Gan, "Deep generative fixed-filter active noise control", In ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1-5, 2023, June. https://doi.org/10.1109/ICASSP49357.2023.10095205
26. M.S. Hussain, M.B.I. Reaz, M.I. Ibrahimy, A.F. Ismail, and F. Mohd-Yasin, "Wavelet based noise removal from EMG signals", *Informacije MIDEM-Journal of Microelectronics Electronic Components and Materials*, *37*(2), pp.94-97. 2007.
27. A. Amsaveni, and P.T. Vanathi, "Reversible data hiding based on radon and integer lifting wavelet transform," *Informacije MIDEM*, vol. *47*(2), pp.91-100, 2017.
28. N. Prashar, M. Sood, and S. Jain, "Design and implementation of a robust noise removal system in ECG signals using dual-tree complex wavelet

transform," *Biomedical signal processing and control*, vol. *63*, p.102212, 2021.
29. M. Rostami, K. Berahmand, and S. Forouzandeh, "A novel community detection based genetic algorithm for feature selection," Journal of Big Data, 8(1), p.2. 2021.
30. L. Nanni, G. Maguolo, S. Brahnam, and M. Paci, "An ensemble of convolutional neural networks for audio classification," *Applied Sciences*, vol. *11*, no. 13, pp.5796, 2021.
31. K. Banuroopa, and D. Shanmuga Priyaa, "MFCC based hybrid fingerprinting method for audio classification through LSTM," *International Journal of Nonlinear Analysis and Applications*, *12*(Special Issue), pp.2125-2136,2021.
32. M. Braik, A. Hammouri, J. Atwan, M.A. Al-Betar, and M.A. Awadallah, "White Shark Optimizer: A novel bio-inspired meta-heuristic algorithm for global optimization problems," *Knowledge-Based Systems*, vol. *243*, pp.108457, 2022.

Arrived: 24. 09. 2024
Accepted: 21. 02. 2025