

# DEEP GENERATIVE FIXED-FILTER ACTIVE NOISE CONTROL

Zhengding Luo, Dongyuan Shi, Xiaoyi Shen, Junwei Ji, Woon-Seng Gan

School of Electrical & Electronic Engineering, Nanyang Technological University, Singapore.  
Emails: LUOZ0021@e.ntu.edu.sg; dongyuan.shi@ntu.edu.sg; ewsgan@ntu.edu.sg

## ABSTRACT

Due to the slow convergence and poor tracking ability, conventional LMS-based adaptive algorithms are less capable of handling dynamic noises. Selective fixed-filter active noise control (SFANC) can significantly reduce response time by selecting appropriate pre-trained control filters for different noises. Nonetheless, the limited number of pre-trained control filters may affect noise reduction performance, especially when the incoming noise differs much from the initial noises during pre-training. Therefore, a generative fixed-filter active noise control (GFANC) method is proposed in this paper to overcome the limitation. Based on deep learning and a perfect-reconstruction filter bank, the GFANC method only requires a few prior data (one pre-trained broadband control filter) to automatically generate suitable control filters for various noises. The efficacy of the GFANC method is demonstrated by numerical simulations on real-recorded noises.

**Index Terms**— Active noise control, generative fixed-filter ANC, deep learning, convolutional neural network

## 1. INTRODUCTION

Active noise control (ANC) is widely regarded as an effective noise cancellation technique that produces an "anti-noise" from a secondary source to suppress the unwanted noise [1, 2, 3, 4]. Based on the principle of superposition of acoustic signals, the anti-noise is of equal amplitude and opposite phase to the disturbance [5]. ANC is particularly effective at cancelling low-frequency noises, where passive methods are typically ineffective, prohibitively expensive or cumbersome [6]. Therefore, ANC technologies have been integrated into a wide range of commercial products including headphones, mobile phones, automobiles, etc.

In adaptive ANC systems, LMS-based algorithms such as filtered-X least mean square (FxLMS) are typically employed [7, 8]. Nevertheless, they are less capable of handling rapidly varying or non-stationary noises due to the inherent slow convergence and poor tracking ability. Moreover, their slow response speed and risk of divergence also impact customers' perceptions of the ANC effect [9]. To improve the response speed and robustness, many audio products adopt the fixed-filter ANC methods [10]. However, the fixed control filter is

only suitable for a specific noise type, resulting in degraded noise reduction performance for other types of noises [11].

To select different pre-trained control filters given different noise types, a selective fixed-filter active noise control (SFANC) method based on frequency band matching was proposed in [12]. However, the critical parameters of the method can only be determined by trial and error, limiting its practical applications. Thus, convolutional neural networks (CNNs) are integrated into the SFANC approach [13], where all the parameters are automatically learned from the noise dataset. Nevertheless, the number of pre-trained control filters in the SFANC approach is limited, resulting in poor noise control performance for some noises, particularly when they are quite distinct from the filter-training noises.

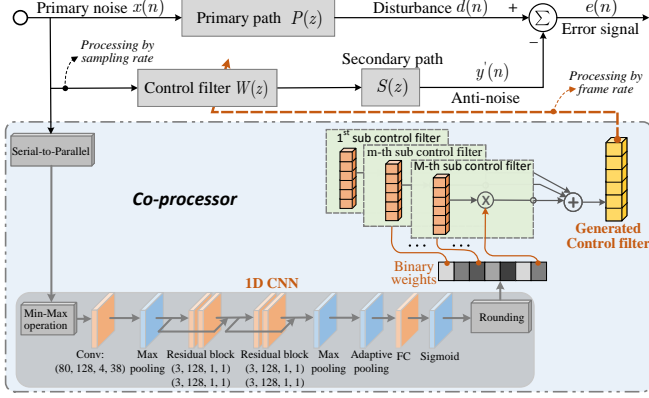
To overcome the limitation of SFANC and generate suitable control filters for various noises, a deep generative fixed-filter active noise control (GFANC) method is proposed in this paper. Firstly, we apply the theory of filter perfect-reconstruction [14] to decompose a pre-trained broadband control filter into multiple sub control filters. On a co-processor, a lightweight one-dimensional CNN (1D CNN) automatically provides the combination weights of sub control filters given the primary noise. After that, a new control filter is generated by the weighted sum of sub control filters and then used for noise cancellation in the real-time controller. For training the 1D CNN, a novel adaptive labelling mechanism is developed to label the training noises without extra human efforts.

The operation on the co-processor (e.g., a mobile phone) can alleviate the computational load of the real-time noise controller. Also, the GFANC approach only requires a single pre-trained control filter, which significantly reduces the ANC's preparation efforts and improves its practicality. Additionally, different from conventional adaptive ANC algorithms, the absence of a feedback error signal [15] in the GFANC method can minimize the risk of divergence. The numerical simulations on real-recorded noises show that the GFANC method can achieve an excellent balance between noise reduction performance and response time.

## 2. THE PROPOSED GFANC APPROACH

Fig. 1 illustrates the GFANC approach's overall architecture, where the co-processor uses the 1D CNN to provide a binary

The code is available at <https://github.com/Luo-Zhengding/GFANC>



**Fig. 1.** Block diagram of the proposed GFANC approach. Convolution layer: (kernel size, channels, stride, padding).

weight vector given each frame's noise data. Through the inner product of this binary weight vector and fixed sub control filters, a new control filter is generated and then utilized for real-time noise control. The weights used are binary weights to reduce the load of computation and memory space. Especially, the co-processor operates at the frame rate, while the real-time controller operates at the sampling rate.

### 2.1. Construction of Sub Control Filters

The construction of sub control filters is an essential step of the proposed method. We used a practical filter decomposition technique based on the theory of filter perfect-reconstruction [16] to obtain sub control filters. Firstly, we used the target ANC system to cancel a broadband primary noise  $\mathbf{x}(n)$  containing the frequency components of interest. Its optimal control filter determined by the FxLMS algorithm [17] is our pre-trained broadband control filter.

It is assumed that the pre-trained broadband control filter has  $N$  taps, which is represented as  $\mathbf{c} = [c(0), \dots, c(N-1)]^T$ .  $\mathbf{c}$  will be decomposed into a perfect-reconstruction filter bank as outlined below. Through the discrete Fourier transform (DFT), its frequency spectrum is derived as

$$\mathbf{C} = \mathbf{F}_N \mathbf{c} = [C(0), \dots, C(k), \dots, C(N-1)]^T, \quad (1)$$

where  $\mathbf{F}_N$  denotes the DFT matrix. According to the conjugate symmetry of the real signal in frequency domain, the control filter  $\mathbf{C}$  is divided into  $M$  sub control filters as

$$\mathbf{C} = \sum_{m=1}^M \mathbf{C}_m, \quad (2)$$

in which the frequency spectrum of  $m$ -th sub control filter can be represented as

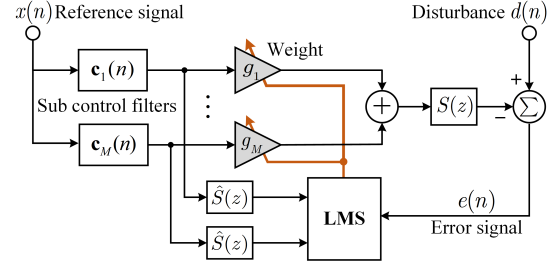
$$\mathbf{C}_m = [C_m(0), \dots, C_m(k), \dots, C_m(N-1)]^T \quad m \in [1, M]. \quad (3)$$

When  $m \neq M$  the elements in (3) are obtained from

$$C_m(k) = \begin{cases} C(k) & k \in [(m-1)I + 1, mI] \\ & \cup [N - mI, N - 1 - (m-1)I] \\ 0 & \text{others;} \end{cases} \quad (4)$$

When  $m = M$ , the sub control filter's frequency spectrum is

$$C_M(k) = \begin{cases} C(k) & k \in [(M-1)I + 1, N - 1 - (M-1)I] \\ 0 & \text{others,} \end{cases} \quad (5)$$



**Fig. 2.** Adaptive labelling mechanism to get the optimal combination weights of sub control filters given the training noise.

where the bandwidth of sub control filter is  $I = \lfloor \frac{N}{2M} \rfloor$ .

The time-domain representation of the  $m$ -th sub control filter is computed by

$$\mathbf{c}_m = \mathbf{F}_N^{-1} \mathbf{C}_m, \quad (6)$$

where  $\mathbf{F}_N^{-1}$  stands for the inverse DFT matrix.

Furthermore, by combing (1), (2), and (6), the control signal  $y(n)$  can be written as

$$\begin{aligned} y(n) &= \mathbf{x}^T(n) \mathbf{c} = \mathbf{x}^T(n) \mathbf{F}_N^{-1} \mathbf{F}_N \mathbf{c} = \mathbf{x}^T(n) \mathbf{F}_N^{-1} \mathbf{C} \\ &= \mathbf{x}^T(n) \mathbf{F}_N^{-1} \sum_{m=1}^M \mathbf{C}_m = \mathbf{x}^T(n) \sum_{m=1}^M \mathbf{c}_m. \end{aligned} \quad (7)$$

It indicates that the control filter  $\mathbf{c}$  can be perfectly reconstructed in time domain by its all sub-band filters  $\mathbf{c}_m$ . Therefore, by decomposing the pre-trained broadband control filter based on the theory of filter perfect-reconstruction, multiple fixed sub control filters can be obtained.

### 2.2. Adaptive Labelling Mechanism

The input and output of the 1D CNN are a frame of 1-second noise and its corresponding binary weight vector, respectively. For automatically labelling the noise dataset, we proposed an adaptive labelling mechanism as shown in Fig. 2 to reduce the required time and effort. Formally, the training noise is input as the reference signal  $\mathbf{x}(n)$ .

Hence, the output signal of the  $m$ -th fixed sub control filter is computed to be

$$y_m(n) = \mathbf{x}^T(n) \mathbf{c}_m \quad m \in [1, M]. \quad (8)$$

The control signal  $y_g(n)$  is subsequently derived from the weighted sum of all sub control filter outputs:

$$y_g(n) = \mathbf{y}^T(n) \mathbf{g}(n), \quad (9)$$

where  $\mathbf{y}(n) = [y_1(n), \dots, y_m(n), \dots, y_M(n)]^T$ , and the combination weight vector is  $\mathbf{g}(n) = [g_1(n), \dots, g_m(n), \dots, g_M(n)]^T$ .

The residual error signal is calculated as

$$e(n) = d(n) - y_g(n) * s(n), \quad (10)$$

where  $d(n)$  and  $s(n)$  represent the disturbance collected by the error microphone and the impulse response of secondary path  $S(z)$ , respectively.  $*$  represents the linear convolution.

To minimize the square of (10) through LMS algorithm [18], the updating formula of the weights is derived as

$$\begin{aligned} \mathbf{g}(n+1) &= \mathbf{g}(n) + \mu \mathbf{y}'(n) e(n), \\ \mathbf{y}'(n) &= \mathbf{y}(n) * s(n), \end{aligned} \quad (11)$$

where  $\mu$  stands for the updating step size.

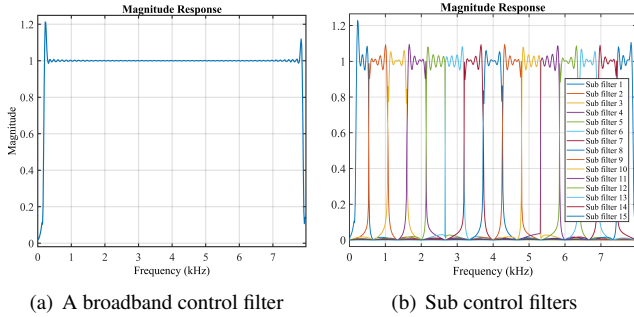


Fig. 3. The sub control filters used in the GFANC approach.

Once (11) converges, the binary weight vector  $\mathbf{T}$  is gained by rounding the optimal weight vector:

$$t_m = \begin{cases} 1 & g_m^o \geq 0.5, \\ 0 & g_m^o < 0.5, \end{cases} \quad (12)$$

$$\mathbf{T} = [t_1, \dots, t_m, \dots, t_M]^T,$$

where  $g_m^o$  means the optimal combination weight of  $m$ -th sub control filter. Finally, the automatically obtained binary weight vector  $\mathbf{T}$  is seen as the label of the training noise.

### 2.3. The Proposed 1D CNN

The architecture of the proposed 1D CNN used for getting the combination weights is shown in Fig. 1. In the pre-processing stage, there is a min-max operation defined as

$$\hat{x}(n) = \frac{x(n)}{\max[\mathbf{x}(n)] - \min[\mathbf{x}(n)]}, \quad (13)$$

where  $\max[\cdot]$  and  $\min[\cdot]$  mean obtaining the maximum and minimum value of the input vector  $\mathbf{x}(n)$ , respectively. The operation rescales the input range into  $(-1, 1)$  and retains its negative part containing phase information, which is important for ANC applications [19].

Motivated by the work [20], the 1D CNN is designed as a lightweight network including two residual blocks. Every residual block comprises two convolutional layers, batch normalization and ReLU non-linearity. Also, a shortcut connection is utilized in each residual block, as the residual architecture has been shown to be easily optimized [21]. Additionally, the network uses a broad receptive field (RF) in the first convolutional layer and narrow RFs in the rest convolutional layers to fully exploit both global and local information.

In this task, obtaining the binary weight vector belongs to a multi-label classification problem [22]. Thus, Sigmoid function is used as the output layer of the 1D CNN. Also, Binary Cross Entropy (BCE) loss is applied to train the model.

### 2.4. Generation of Control Filter

After training, the 1D CNN can output the binary weight vector depending on the primary noise. Then, a new control filter is formed by the weighted sum of sub control filters. Formally, the generated control filter  $\hat{\mathbf{w}}$  is obtained from

$$\hat{\mathbf{w}} = \sum_{m=1}^M t_m \cdot \mathbf{c}_m \quad m \in [1, M]. \quad (14)$$

Table 1. Performance comparison of different networks.

Network	Testing Accuracy	#Parameters
1D Convolutional Neural Networks		
Proposed Network	<b>97.20%</b>	<b>0.21M</b>
M3 Network	96.64%	0.22M
M5 Network	96.70%	0.56M
M11 Network	96.65%	1.79M
M18 Network	94.35%	3.69M
M34-res Network	96.74%	3.99M
2D Convolutional Neural Networks		
ShuffleNet v2	96.40%	0.25M
Mobilenet v2	96.05%	2.89M
Attention Network	95.80%	4.95M

The generated control filter would be used by the real-time processor for noise cancellation:

$$e(n) = d(n) - \mathbf{x}^T(n) \hat{\mathbf{w}} * s(n). \quad (15)$$

Overall, given the incoming noise, a new control filter can be generated by adaptively combining fixed sub control filters. Instead of selecting from a limited set of control filters, GFANC can generate a more suitable control filter for each type of noise compared to SFANC. Moreover, unlike adaptive ANC algorithms, GFANC does not use the error signal to update the control filter, reducing the risk of divergence.

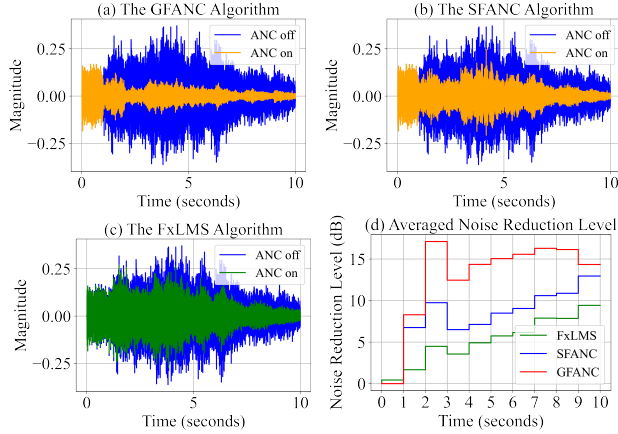
## 3. SIMULATION RESULTS

In the simulations, the pre-trained broadband control filter is divided into 15 sub control filters as depicted in Fig. 3. The length of the control filter and the sampling rate are set to 1,024 taps and 16kHz, respectively. The primary and secondary paths used are bandpass filters with a frequency range of 20Hz-7,980Hz. For training the network, a synthetic noise dataset divided into 3 subsets: 80,000 noise tracks for training, 2,000 noise tracks for validation, and 2,000 noise tracks for testing is used. Each noise track has a 1-second duration.

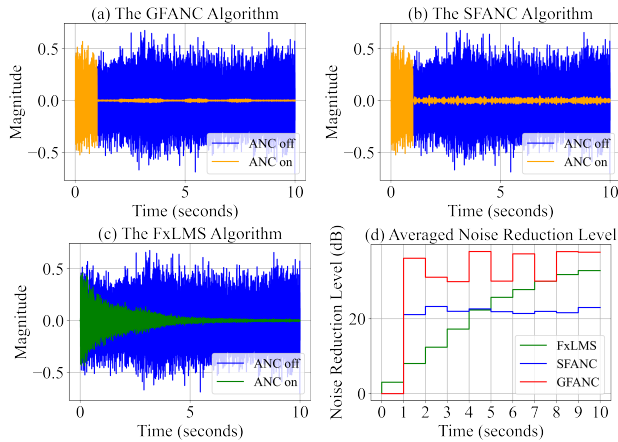
### 3.1. Comparison of Different Networks

The proposed 1D CNN is compared to several 1D networks [20]: M3, M5, M11, M18, and M34-res. Also, some light 2D networks including ShuffleNet v2 [23], Mobilenet v2 [24] and Attention Network [25] are used as baselines. The performance of these networks in the GFANC approach are summarized in Table 1, where the testing accuracy means the accuracy of predicting binary weights on the test dataset.

On the test dataset, the proposed 1D network achieves the highest prediction accuracy of 97.20%, indicating that it can effectively extract discriminative features from raw waveforms to provide suitable binary weights given different noises. Meanwhile, the proposed 1D network is lightweight, with only 0.21M parameters, making it suitable to be implemented on co-processors such as mobile phones. Furthermore, compared to using frequency-domain data in 2D networks, using time-domain waveform directly in 1D networks is more convenient and practical [26].



**Fig. 4.** (a)-(c): Attenuated noise signals by different ANC algorithms, (d): Averaged noise reduction level on the aircraft noise (50Hz-8,000Hz) during every 1 second.

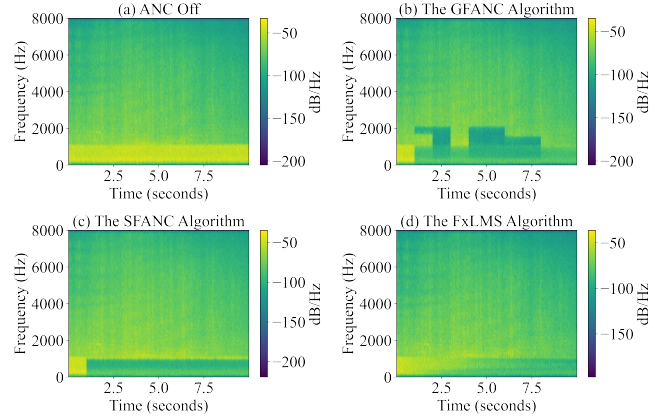


**Fig. 5.** (a)-(c): Attenuated noise signals by different ANC algorithms, (d): Averaged noise reduction level on the traffic noise (40Hz-1,200Hz) during every 1 second.

### 3.2. Real-world Noise Cancellation

The GFANC algorithm is compared to SFANC [27] and FxLMS in controlling real-world noises, which do not belong to the training dataset. The step size of FxLMS algorithm is 0.0001. In the first simulation in Fig. 4, the ANC system utilized different algorithms to cancel the aircraft noise. It is observed that the proposed GFANC algorithm outperforms SFANC and FxLMS algorithms in terms of response speed and noise reduction level. Specifically, during 2s-3s, the average noise reduction level achieved by GFANC is approximately 7dB and 12dB greater than that of SFANC and FxLMS, respectively. This suggests that the GFANC algorithm is able to track and respond rapidly to varying noise. However, GFANC and SFANC cannot handle the first-second noise because they update the control filter coefficients for the next second based on the first-second noise.

Also, these ANC methods are used to attenuate the low-frequency traffic noise, as presented in Fig. 5. Though the SFANC method responds to the noise faster than the FxLMS



**Fig. 6.** The spectrogram of noise signal obtained (a) without ANC algorithm, (b)-(d) with different ANC algorithms, on the mixed noise of aircraft and traffic noise.

method in the first 4s, the FxLMS algorithm performs better after convergence [28]. In comparison, the proposed GFANC method always performs best during the noise reduction process. The results indicate that the control filter generated by the GFANC method is more suitable for the noise than that provided by the SFANC method. Furthermore, our previous research [27] demonstrated that combining fixed-filter and adaptive ANC algorithms can achieve rapid convergence and satisfactory noise reduction level.

Furthermore, we added the aircraft and traffic noises together in time domain to get a mixed noise. The spectrograms of the mixed noise attenuated by different ANC algorithms are depicted in Fig. 6. The spectrogram represents the power spectral density of the noise. Noticeably, most components of the mixed noise can be effectively attenuated by the GFANC algorithm. However, SFANC is less effective at removing the noise components above 1,000Hz. The reason is that SFANC can only choose control filters from a limited number of candidates, whereas GFANC is capable of generating more appropriate control filters. Also, it is found that the FxLMS algorithm responds much slower on the mixed noise than GFANC and SFANC.

## 4. CONCLUSION

This paper proposed a novel GFANC method that generates different control filters given different primary noises. A lightweight 1D CNN operated on a co-processor extracts the noise features to form a binary weight vector. Using the weight vector, a new control filter is generated by combining sub control filters obtained from the decomposition of a pre-trained broadband control filter. Owing to the flexibility of generating control filters and avoiding optimization progress, the GFANC achieves better response speed and noise reduction level than SFANC and FxLMS. Additionally, the training dataset is automatically labelled via the adaptive mechanism instead of manual labelling. Also, only one pre-trained control filter is required, significantly reducing the effort for pre-training control filters in real-world applications.

## 5. REFERENCES

- [1] Colin N Hansen, *Understanding active noise cancellation*, CRC Press, 2002.
- [2] Stefan Liebich, Johannes Fabry, Peter Jax, and Peter Vary, "Signal processing challenges for active noise cancellation headphones," in *Speech Communication; 13th ITG-Symposium*. VDE, 2018, pp. 1–5.
- [3] Feiran Yang, Yin Cao, Ming Wu, Felix Albu, and Jun Yang, "Frequency-domain filtered-x lms algorithms for active noise control: A review and new insights," *Applied Sciences*, vol. 8, no. 11, pp. 2313, 2018.
- [4] Xiaojun Qiu, *An introduction to virtual sound barriers*, CRC Press, 2019.
- [5] Philip A Nelson and Stephen J Elliott, *Active control of sound*, Academic press, 1991.
- [6] Sen M Kuo and Dennis R Morgan, "Active noise control: a tutorial review," *Proceedings of the IEEE*, vol. 87, no. 6, pp. 943–973, 1999.
- [7] Stephen J Elliott and Philip Arthur Nelson, "Active noise control," *IEEE signal processing magazine*, vol. 10, no. 4, pp. 12–35, 1993.
- [8] Dongyuan Shi, Woon-Seng Gan, Bhan Lam, Shulin Wen, and Xiaoyi Shen, "Active noise control based on the momentum multichannel normalized filtered-x least mean square algorithm," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*. Institute of Noise Control Engineering, 2020, vol. 261, pp. 709–719.
- [9] Kazuyuki Arikawa, Shoichi Koyama, and Hiroshi Saruwatari, "Spatial active noise control based on individual kernel interpolation of primary and secondary sound fields," in *2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2022, pp. 1056–1060.
- [10] Jianfeng Guo, Feiran Yang, and Jun Yang, "Convergence analysis of the conventional filtered-x affine projection algorithm for active noise control," *Signal Processing*, vol. 170, pp. 107437, 2020.
- [11] Daocheng Chen, Longbiao Cheng, Dingding Yao, Junfeng Li, and Yonghong Yan, "A secondary path-decoupled active noise control algorithm based on deep learning," *IEEE Signal Processing Letters*, vol. 29, pp. 234–238, 2021.
- [12] Dongyuan Shi, Woon-Seng Gan, Bhan Lam, and Shulin 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.
- [13] Dongyuan Shi, Bhan Lam, Kenneth Ooi, Xiaoyi Shen, and Woon-Seng Gan, "Selective fixed-filter active noise control based on convolutional neural network," *Signal Processing*, vol. 190, pp. 108317, 2022.
- [14] Martin Vetterli and Didier Le Gall, "Perfect reconstruction fir filter banks: Some properties and factorizations," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 1057–1071, 1989.
- [15] Bernard Widrow, John R Glover, John M McCool, John Kautitz, Charles S Williams, Robert H Hearn, James R Zeidler, JR Eugene Dong, and Robert C Goodlin, "Adaptive noise cancelling: Principles and applications," *Proceedings of the IEEE*, vol. 63, no. 12, pp. 1692–1716, 1975.
- [16] Hai Huyen Dam, Sven Nordholm, Antonio Cantoni, and Jan Mark de Haan, "Iterative method for the design of dft filter bank," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 51, no. 11, pp. 581–586, 2004.
- [17] Feiran Yang, Jianfeng Guo, and Jun 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.
- [18] Xiaoyi Shen, Dongyuan Shi, Woon-Seng Gan, and Santi Peksi, "Adaptive-gain algorithm on the fixed filters applied for active noise control headphone," *Mechanical Systems and Signal Processing*, vol. 169, pp. 108641, 2022.
- [19] Hao Zhang and DeLiang Wang, "Deep anc: A deep learning approach to active noise control," *Neural Networks*, vol. 141, pp. 1–10, 2021.
- [20] Wei Dai, Chia Dai, Shuhui Qu, Juncheng Li, and Samarjit Das, "Very deep convolutional neural networks for raw waveforms," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2017, pp. 421–425.
- [21] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [22] Jonatas Wehrmann, Ricardo Cerri, and Rodrigo Barros, "Hierarchical multi-label classification networks," in *International conference on machine learning*. PMLR, 2018, pp. 5075–5084.
- [23] Ningning Ma, Xiangyu Zhang, Hai-Tao Zheng, and Jian Sun, "Shufflenet v2: Practical guidelines for efficient cnn architecture design," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 116–131.
- [24] Sainath Adapa, "Urban sound tagging using convolutional neural networks," in *Proceedings of the Detection and Classification of Acoustic Scenes and Events 2019 Workshop (DCASE2019)*, 2019, pp. 5–9.
- [25] Zhengding Luo, Junting Li, and Yuesheng Zhu, "A deep feature fusion network based on multiple attention mechanisms for joint iris-periocular biometric recognition," *IEEE Signal Processing Letters*, vol. 28, pp. 1060–1064, 2021.
- [26] Yedid Hoshen, Ron J Weiss, and Kevin W Wilson, "Speech acoustic modeling from raw multichannel waveforms," in *2015 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. IEEE, 2015, pp. 4624–4628.
- [27] Zhengding Luo, Dongyuan Shi, and Woon-Seng Gan, "A hybrid sfanc-fxnllms algorithm for active noise control based on deep learning," *IEEE Signal Processing Letters*, vol. 29, pp. 1102–1106, 2022.
- [28] Yoshinobu Kajikawa, Woon-Seng Gan, and Sen M Kuo, "Recent advances on active noise control: open issues and innovative applications," *APSIPA Transactions on Signal and Information Processing*, vol. 1, 2012.