https://doi.org/10.7840/kics.2025.50.4.535

Suppressing the Acoustic Effects of UAV Propellers through Deep Learning-Based Active Noise Cancellation

Faisal Ayub Khan◆, Soo Young Shin°

ABSTRACT

This study presents a deep learning-based Active Noise Cancellation (ANC) system for reducing UAV propeller noise using a Convolutional Neural Network (CNN) model. The proposed system effectively minimizes noise in real-time by extracting key audio features such as amplitude, phase, and frequency components, generating and calculating inverse feature values to construct precise anti-noise signals. This approach enables destructive interference, significantly reducing the propeller noise. The model achieved high-performance metrics, including 94.5% accuracy, 93.2% precision, 96.1% recall, and a loss value of 0.115, demonstrating its efficacy in noise cancellation. Deployed on an Nvidia Jetson NX, the ANC system integrates high-quality microphones and strategically placed speakers on a UAV platform, allowing for real- time noise analysis and anti-noise generation. Indoor and outdoor tests validated a substantial reduction in propeller noise up to 36 dB, highlighting the model' s robustness and potential for quieter UAV operation in noise-sensitive settings.

Key Words : Deep Learning, Unmanned Aerial Vehicles (UAVs), Convolutional Neural Network (CNN), Adaptive ANC, Neural Network (NN), Real-time Noise Mapping

Ⅰ. Introduction

Unmanned Aerial Vehicles (UAVs) are finding applications across diverse sectors, including agriculture, surveillance, emergency response, and delivery services. Their ability to perform tasks autonomously in hard-to-reach areas has transformed industries, improving operational efficiency and reducing human risk[1]. However, UAVs face a significant hurdle in terms of their acceptability in urban and densely populated areas due to the noise generated by their propellers[2,3]. This high-frequency, intrusive noise can be disruptive, especially in settings where quiet is crucial[2], such as residential neighborhoods, wildlife monitoring sites, or medical facilities.

Building upon our prior work on noise reduction for UAVs[4], where we developed an initial framework to mitigate propeller noise, this study advances the approach by enhancing model effectiveness, alongside hardware implementation to improve noise cancellation performance. As mentioned before Active Noise

Cancellation (ANC) techniques can tackle the pervasive issues of noise pollution associated with UAVs[5] by employing sophisticated algorithms and sensor arrays to detect and counteract the noise generated by propulsion systems and aerodynamic forces.

※ This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Min- istry of Education" (2018R1A6A1A03024003, 50%) and by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2024-RS-2024-00437190, 50%) supervised by the IITP(Institute for Information Communications Technology Planning Evaluation).
◆ First Author : Kumoh National Institute of Technology, Department of IT Convergence Engineering, faisalayubkhan10@gmail.com
° Corresponding Author : Kumoh National Institute of Technology, Department of IT Convergence Engineering, wdragon@kumoh.ac.kr, 종신회원
논문번호 : 202411-271-A-RU, Received November 7, 2024; Revised December 3, 2024; Accepted December 8, 2024

ANC operates on the principle of "destructive interference" where the systems capture sound through multiple sensors, which then process the data to create anti-noise signals that are equal in amplitude but precisely 180° out of phase[6]. As the two signals converge, they undergo destructive interference, effectively nullifying each other' s effects. Fig. 1 illustrates this fundamental concept, displaying how the primary noise and antinoise interact to achieve a residual noise level that is substantially lower than the original. This noise reduction requires precise calibration: the anti-noise must match the primary noise in both magnitude and phase for optimal cancellation[7].

In UAV applications, the effectiveness of ANC becomes even more impactful when combined with deep learning (DL) models[8]. By training deep neural net works on extensive datasets of UAV noise profiles, the system learns to recognize specific noise signatures generated by UAV propellers and aerodynamic forces[7]. This training enables the neural networks to extract detailed audio features, such as frequency components, amplitude variations, and noise directionality, which contribute to a more precise anti-noise generation process[9]. Additionally, the adaptive capabilities of deep learning enhance the ANC system' s response to changing conditions, such as varying altitudes, wind patterns, and speed, which all impact noise levels. Using an adaptive approach, the system continuously adjusts anti-noise signals to match the dynamic noise sources, ensuring higher cancellation effectiveness and faster response times compared to traditional ANC systems. This data-driven method allows the system to learn and adapt to new noise patterns encountered during operation, improving its performance over time[10]. The synergy between ANC and deep learning enables UAVs to operate more quietly, facilitating their integration into environments where noise control is critical, such as urban areas, medical facilities, and wildlife monitoring zones.

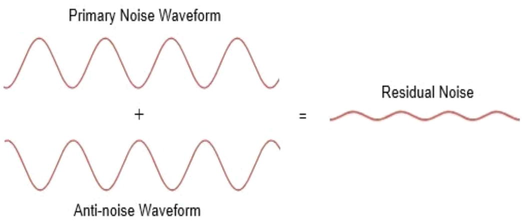


Fig. 1. Illustration of Destructive Interference: Combining primary noise with anti-noise to reduce overall sound through phase cancellation.

Ⅱ. Literature Review

While multiple studies have investigated traditional Active Noise Cancellation (ANC) systems, these techniques generally rely on adaptive filtering methods that process incoming signals to cancel out noise. One of the most widely used approaches is the Least Mean Square (LMS) algorithm, which adapts the filter coefficients based on the error signal. However, LMS has limitations, such as slow convergence rates and its inability to handle non-stationary noise signals effectively[11,12]. Filtered-X LMS (FXLMS) improves upon LMS by incorporating the reference noise signal directly into the update process, which helps with stability in certain applications. Still, FXLMS struggles with highly dynamic noise environments, such as those encountered in UAV applications, and is computationally expensive when dealing with multiple frequencies or complex interference patterns[13,14]. Other traditional ANC techniques, such as the Wiener filter[15] and adaptive notch filtering[16], also exhibit performance limitations in non-stationary environments. While the Wiener filter can be highly effective in stationary noise environments, it suffers from poor adaptability in rapidly changing conditions, like those encountered in UAV applications[17]. Similarly, adaptive notch filtering performs well in filtering out specific frequency bands but cannot adapt to complex and varying noise patterns that are typical in real-time.

The limitations of these traditional ANC methods underscore the need for more advanced techniques capable of handling complex, time-varying, and multi-frequency noise signals. To address these challenges, our proposed system introduces a deep learning-based approach, utilizing Convolutional Neural Networks (CNNs) for noise cancellation. CNNs are adaptive at extracting complex features from high-dimensional inputs such as spectrograms, which represent time-frequency information from noisy audio signals. By learning the intricate patterns within the noise data, CNNs can generate highly accurate anti-noise signals tailored to counteract the noise in real-time[18,19]. Unlike traditional ANC methods, which rely on predefined models or manual feature extraction, CNN-based ANC systems adapt to new, unseen noise patterns through training on large datasets. This flexibility allows the system to continuously learn and improve its performance, making it more effective in dynamic environments[18]. Additionally, the reaction time of CNN-based ANC systems is significantly faster compared to traditional methods, as CNNs can process and output noise-canceling signals almost instantly once trained, reducing the delay commonly observed in adaptive filtering methods like LMS and FXLMS[20].

In contrast to traditional ANC systems, CNNs can handle non-stationary and multi-frequency noise efficiently, making them ideal for applications such as UAVs, where the noise environment is highly dynamic and complex[21]. Furthermore, as mentioned in [18], the computational efficiency of deep learning approaches, particularly with advancements in hardware acceleration (e.g., GPUs), allow real-time operation even in complex scenarios. This study focuses on extracting relevant features from the UAV propeller noise, such as amplitude, frequency, and phase information through CNN. Afterward, extracted features are then used to generate inverse values, which are translated into an anti-noise signal. This signal, when played back through speakers directed toward the propeller, results in a reduction of the primary noise due to the phenomenon of destructive interference.

Deep Learning models, particularly Convolutional Neural Networks (CNNs), are highly effective for complex noise cancellation tasks, including reducing propeller noise, which presents unique spectral patterns and time-variant characteristics that challenge traditional signal processing methods. In these models, CNNs analyze time-frequency representations like spectrograms derived from audio recordings[22], where each frame serves as input to the network. Through a series of convolutional and pooling layers, CNN-based ANC systems extract essential audio features-such as amplitude, frequency spectrum, and phase information-by first converting the raw audio into a spectrogram that captures both temporal and frequency data[23]. Convolutional layers apply learnable filters across this spectrogram to detect relevant patterns at different scales, while pooling layers downsample feature maps, preserving important characteristics and optimizing computational efficiency[24]. Deeper layers within the network recognize complex noise patterns, and fully connected layers map these features to antinoise signals, effectively enabling targeted noise cancellation.

Once extracted, these features pass to fully connected layers, which produce an anti-noise signal tailored to counteract the original propeller noise, thus reducing the overall sound level. During training, the CNN minimizes the difference between the generated cancellation signal and the primary propeller noise, learning how to cancel specific noise characteristics effectively[22]. The model' s generalization to various noise scenarios is enhanced by regularization techniques and extensive dataset training, equipping the system to adapt to new audio environments and diverse noise types, making it a robust solution for dynamic noise reduction in UAV applications.

2.1 Proposed Model

The proposed system model, as illustrated in Fig. 2, outlines a Deep Learning-based Active Noise Cancellation (ANC) architecture designed to mitigate UAV propeller noise. This model uses a Convolutional Neural Network (CNN) framework to process the propeller noise signal, $s(t)$, by extracting crucial features that allow for the precise generation of an anti-noise signal, $\hat{s}(t)$, which destructively interferes with the original noise giving residual noise as a result.

The system begins by capturing the propeller noise as an input, represented in the time domain as a continuous audio waveform. This input signal $s(t)$ is then converted into a time-frequency representation, specifically a spectrogram, denoted as $X(f, t)$, to capture both temporal and spectral characteristics. The spectrogram $X(f, t)$, where f denotes frequency and t denotes time, serves as the initial input to the CNN. Within the CNN

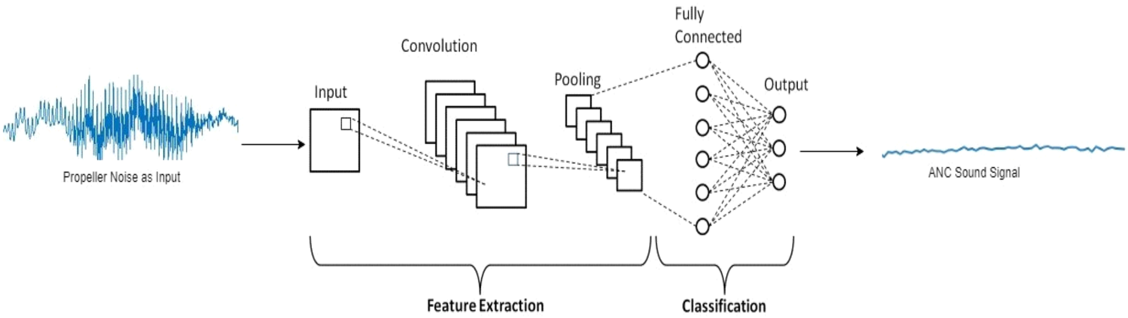


Fig. 2. Proposed Deep Learning model for Active Noise Cancellation of UAV Propeller Noise

, various audio features are extracted to facilitate effective ANC. These features include time-domain characteristics (T_d), which capture the temporal structure of the signal and provide insight into the periodicity and transient properties of the propeller noise. Frequency features (F_s), derived from the spectrogram, reveal dominant frequency components crucial for identifying repetitive noise patterns from the propeller[24]. Amplitude information (A) represents the intensity of the signal at different points, which helps in determining the energy profile of the noise. Phase information (Φ) is also captured, allowing the model to synchronize the anti-noise with the primary noise accurately. Furthermore, the temporal-frequency representation $S(f, t)$ provides a holistic view of how noise characteristics evolve over time, enabling adaptation to dynamic noise patterns.

Each frame of the spectrogram $X(f, t)$ is fed into the CNN, which consists of convolutional layers followed by pooling layers. The convolutional layers apply a set of filters, represented by weights W_c, across the spectrogram, extracting localized patterns and features at various scales[7]. Pooling layers downsample the feature maps, reducing dimensionality while preserving essential characteristics.

The features extracted in this way are then passed to fully connected layers, which act as a classification and decision-making mechanism[25]. These layers map the features to an anti-noise signal $\hat{s}(t)$, carefully calibrated to mirror the amplitude and phase of the primary noise but precisely 180° out of phase. This anti-noise signal achieves destructive interference with the primary noise $s(t)$, significantly reducing the overall sound level. The anti-noise $\hat{s}(t)$ is produced by computing the inverse values of critical features, such as amplitude A^{-1}, frequency F_s^{-1}, and phase Φ^{-1}. These values are combined through the fully connected layers to form a continuous anti-noise signal that effectively neutralizes the propeller noise when superimposed. This dynamic and adaptive approach enables precise and efficient noise cancellation, enhancing operational stealth and minimizing the environmental impact of UAVs[8]. The algorithm detailed in the following sections outlines the step-by-step process for training the model, emphasizing the systematic feature extraction and integration that form the foundation of deep learning-based ANC.

In summary, the proposed model begins with the UAV propeller noise $s(t)$, transforms it into a spectrogram $X(f, t)$ for feature extraction, and processes this data through convolutional and pooling layers to identify relevant features, including time-domain characteristics, frequency, amplitude, phase, and temporal-frequency representation $S(f, t)$. Fully connected layers utilize these features to generate an anti-noise signal $\hat{s}(t)$, which aligns in magnitude but opposes the primary noise in phase. The output anti-noise signal, when combined with the original propeller noise, reduces the perceived sound through destructive interference. This model effectively integrates deep learning with traditional ANC principles, ensuring precision and adaptability in complex and dynamic UAV environments. The design ensures that the model can adjust to varying noise characteristics, making it a robust solution for UAV noise reduction.

2.2 Algorithm

The proposed algorithm 1 for Active Noise Cancellation (ANC) leverages a Deep Learning approach to effectively reduce UAV propeller noise. Given an input audio signal $s(t)$, the system extracts essential features for noise cancellation, including time-domain characteristics T_d, frequency components F_s, amplitude A, phase Φ, and a temporal-frequency representation $S(f, t)$. These features are then combined into a feature vector X = [T_d, F_s, $S(f, t)$, A, Φ], which serves as input to a Convolutional Neural Network (CNN) model. To produce the anti-noise signal, the model computes the inverse of these features, such as T_d^{-1}, F_s^{-1}, A^{-1}, and Φ^{-1}, creating a feature set that counteracts the primary noise. During training, the model optimizes its parameters W and β which are crucial for the model's learning process. W represents the weights in the convolutional layers, which are learned during training to detect patterns and features in the input audio spectrograms. These weights enable the model to capture the noise characteristics of the UAV propeller sound. β represents the bias terms added to the weighted sum of inputs, allowing the model to adjust its output independently of the input. Together, W and β help the model optimize the noise-canceling signal by minimizing the loss function $\mathscr{L}$ = MSE($s(t)$, $\hat{s}(t)$) + $\lambda\|W\|^2$, where MSE measures the difference between the generated anti-noise signal $\hat{s}(t)$, while $\lambda\|W\|^2$ acts as a regularization term to prevent overfitting. In the real-time noise cancellation phase, the trained model processes new input signals $r(t)$ to generate the anti-noise output $\hat{s}(t)$, which is then combined to achieve destructive interference and reduce the overall noise level.

2.3 Spectrogram Visualization

To facilitate feature extraction for noise cancellation, the input audio signal $s(t)$ is transformed into a timefrequency representation using the Short-Time Fourier Transform (STFT) which is expressed as:

$$X(f,t) = \mathrm{STFT}\{s(t)\}(f,t) = \int s(\tau)w(\tau - t)e^{-j2\pi f\tau}\,d\tau \tag{1}$$

Algorithm 1 : ANC with Deep Learning - Feature extraction, inverse signal generation, and real-time anti-noise synthesis for noise reduction

Require: Audio signals $s(t)$ (training set), Y (labels for training), $\hat{s}(t)$ anti-noise signal, $r(t)$ (new input signals), learning rate η, batch size B, regularization parameter λ

Ensure: Trained model parameters W, β and noise-cancelled output

Stage 1: TRAIN($s(t), Y$)
  $W \leftarrow$ initialize weights randomly
  $\beta \leftarrow$ initialize model parameters
for each batch $B_i \subset s(t)$ **do**
  Stage 2: Extract features from input batch B_i
    $T_d \leftarrow$ time-domain characteristics from B_i
    $F_s \leftarrow$ frequency features from B_i
    $S(f,t) \leftarrow$ temporal-frequency from B_i
    $A \leftarrow$ amplitude information from B_i
    $\Phi \leftarrow$ phase information from B_i
  Stage 3: Form feature vector $\mathbf{X} = [T_d, F_s, S(f,t), A, \Phi]$
  Generate inverse feature values (e.g., $T_d^{-1}, F_s^{-1}, A^{-1}, \Phi^{-1}$)
  Stage 4: Calculate the anti-noise signal $\hat{s}(t) = \text{model}(\mathbf{X}, W, \beta)$
  Stage 5: Compute $\mathscr{L} = \text{MSE}(s(t)), \hat{s}(t)) + \lambda\|W\|^2$
  Update W and β using backpropagation and optimization with learning rate η (SGD or Adam)
end for
return W, β
Stage 6: ADAPTIVE NOISE RESPONSE ($r(t)$)
Stage 7: Extract features from new input audio $r(t)$
  $T_d' \leftarrow$ time-domain characteristics from $r(t)$
  $F_s' \leftarrow$ frequency features from $r(t)$
  $S'(f,t) \leftarrow$ temporal-frequency from $r(t)$
  $A' \leftarrow$ amplitude information from $r(t)$
  $\Phi' \leftarrow$ phase information from $r(t)$
Stage 8: Form feature vector $\mathbf{X}' = [T_d', F_s', S'(f,t), A', \Phi']$
Generate anti-noise signal using trained model
  $\hat{s}(t) \leftarrow \text{model}(\mathbf{X}', W, \beta)$
Stage 9: Combine primary audio with anti-noise signal Output $\leftarrow \text{SIGN}(r(t) + \hat{s}(t))$
return Output

where $X(f, t)$ represents the spectrogram, f is the frequency variable, t denotes time, $s(\tau)$ is the original audio signal, and $w(\tau-t)$ is a window function that localizes the Fourier transform within a specific time frame. The STFT enables the decomposition of $s(t)$ into a sequence of time-dependent frequency components, providing a comprehensive view of how the signal's frequency characteristics evolve over time. Magnitude spectrogram $|X(f, t)|$ is computed as:

$$|X(f,t)| = \sqrt{\mathrm{Re}(X(f,t))^2 + \mathrm{Im}(X(f,t))^2} \tag{2}$$

where $\mathrm{Re}(X(f, t))$ and $\mathrm{Im}(X(f, t))$ represent the real and imaginary parts of $X(f, t)$, respectively. This magnitude spectrogram provides the amplitude of each frequency component, allowing the model to focus on prominent noise frequencies and their variations over time, which is crucial for generating effective anti-noise signals.

Time-frequency representation is crucial for identifying key noise features associated with the UAV's propeller sounds. By visualizing the spectrogram, it becomes easier to analyze the dominant frequency bands and transient patterns that characterize the noise[26], especially for propeller-driven UAVs, which exhibit repetitive spectral patterns. The CNN model subsequently processes these spectrograms, leveraging the temporal and spectral data to enhance the accuracy of anti-noise generation.

Furthermore, parameters such as the window function $w(\tau-t)$ and the window length influence the resolution of the spectrogram. A shorter window length provides higher temporal resolution, allowing for precise tracking of rapid changes in noise, while a longer window length improves frequency resolution[27], which is beneficial for capturing consistent noise patterns over time. These parameters are selected based on the specific characteristics of the UAV noise, ensuring that the spectrogram accurately represents both the time and frequency domains of the input signal. This allows the model to capture crucial features in both stable and dynamic noise environments.

2.4 Layered Architecture for Feature Extraction

The proposed system leverages a series of Convolutional, Pooling, and Fully Connected layers to process the spectrogram $X(f, t)$ of the UAV propeller noise signal, extracting and refining features essential for noise cancellation. The Convolutional Layer applies learnable filters W_c across the spectrogram, performing a convolution operation at each position (i, j) to produce output $h_c(i, j)$ based on local frequency and time patterns. This operation is defined as:

$$h_c(i,j) = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} W_c(m,n)\cdot X(i+m, j+n) + b_c \tag{3}$$

where M and N denote the filter dimensions. These filters capture crucial noise characteristics, such as prominent frequency peaks and temporal variations tied to propeller noise. Following this, the Pooling Layer reduces the spatial dimensions[28] of the feature maps, capturing only the most prominent features from each region by taking the maximum value within a pooling window, represented by

$$h_p(i,j) = \max_{m,n} h_c(i+m, j+n). \tag{4}$$

This layer helps the model prioritize significant noise components and reduces the impact of minor fluctuations, thus enhancing the ANC system's efficiency. Finally, the Fully Connected Layer flattens the output into a single feature vector $\mathbf{z}$ and computes a linear combination[28] of these features, transforming them into the anti-noise signal $\hat{s}(t)$. This transformation is represented by:

$$h_{fc} = \sigma\left(\sum_{k=1}^{K} w_k z_k + b_{fc}\right), \tag{5}$$

where w_k and b_{fc} are the layer's weights and biases, and σ is an activation function. This layer integrates the extracted features, mapping them to an anti-noise signal that mirrors the UAV noise in amplitude but is precisely out of phase, thus enabling effective noise cancellation through destructive interference.

2.5 Generating Anti-Noise Sound Signal

Once the features have been extracted from the UAV propeller noise spectrogram and their inverse values calculated, the system generates an anti-noise signal $\hat{s}(t)$ designed to mirror the characteristics of the primary noise $s(t)$ but with an opposite phase. The extracted features-including time-domain characteristics, frequency components, amplitude, and phasealong with their respective inverse values, are combined into a final feature vector that serves as input for the anti-noise generation function. Using this vector of inverse features, anti-noise signal $\hat{s}(t)$ is formulated as:

$$\hat{s}(t) = A^{-1}\cos(\omega t + \Phi^{-1}) \tag{6}$$

where A^{-1} and Φ^{-1} represent the inverse amplitude and phase components of the primary noise, respectively, while ω is the angular frequency of the noise, which determines the rate of oscillation of both the primary noise and anti-noise signals. Equation 6 ensures that $\hat{s}(t)$ is aligned in both magnitude and phase opposition to $s(t)$. To achieve perfect cancelation, both the inverse amplitude A^{-1} and inverse phase Φ^{-1} must be precisely synchronized and transmitted to overlap with the primary sound $s(t)$ at the exact same time slot, as illustrated in Fig. 3. The anti-noise signal is further refined by minimizing the Mean Squared Error (MSE) loss between $\hat{s}(t)$ and the target anti-noise, ensuring precise phase alignment for effective noise cancellation.

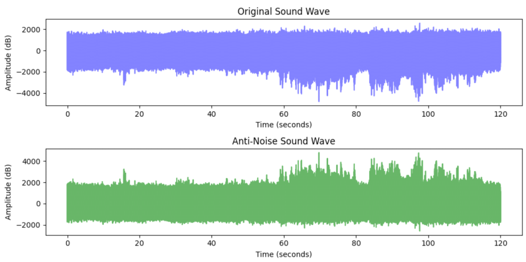


Fig. 3. Primary audio waveform of propeller sound and its Anti-Noise waveform.

Ⅲ. Experiment

The experimentation phase included preparing and pre-processing the dataset, followed by model training and testing. Simulations were first conducted to assess initial model performance in controlled environments, and the trained model was then integrated into the UAV setup for real-time testing. These evaluations, detailed in the following subsections, confirm the model's noise cancellation capabilities in both simulated and physical environments.

3.1 Dataset Acquisition and Preprocessing

An extensive dataset, comprising over 2200+ audio recordings, were carefully curated from two primary sources: publicly available datasets obtained from reliable internet repositories and custom recordings captured in controlled environments. The publicly sourced data provided a diverse range of propeller noise profiles under various operational conditions, while the custom recordings were specifically tailored to mimic real-world scenarios, including noise variations introduced by different propeller types and speeds. Each recording underwent preprocessing to extract critical audio features such as frequency, pitch, phase, amplitude envelope, and temporal characteristics. These features were systematically stored in a structured CSV format to streamline further analysis and model training. The dataset was divided into 70% for training, 15% for validation, and 15% for testing, ensuring an even distribution across various noise profiles and conditions. This comprehensive approach ensured that the dataset effectively represented diverse propeller noise characteristics, enabling the model to generalize well to new and unseen noise scenarios.

3.2 Model Training

For model training, each audio sample was converted into spectrogram representations using a frame size of 2048 samples and a hop length of 512 samples, with a sampling rate of 22050 Hz. These settings provided a balanced time-frequency resolution to capture UAV noise characteristics effectively. The CNN architecture employed in this study consisted of multiple convolutional layers, each followed by max-pooling layers for efficient feature extraction. The convolutional layers used filters to detect spatial patterns within the spectrogram, leveraging cross-correlation to capture variations in UAV noise characteristics. Activation functions such as the Rectified Linear Unit (ReLU) were applied after each convolutional and fully connected layer to enhance the model' s capacity for learning complex noise patterns. In the deeper layers, the model learned intricate wave patterns characteristic of UAV propeller noise, while the fully connected layers generated noise cancellation signals from these extracted features. The model' s final layer used a linear activation function to produce the ANC signal output.

To optimize model performance, the mean squared error (MSE) loss function with a regularization term λ = 0.0001 was employed to minimize the difference between the generated noise-canceling signal and the primary UAV noise. The loss function is defined as:

$$\mathscr{L} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 + \lambda\|W\|^2 \tag{7}$$

where N is the number of training samples, y_i denotes the primary UAV propeller audio at ith sample which serves as the input for the model. The corresponding target $\hat{y}_i$, or the model's ANC output, is generated to be the inverse of the primary audio, such that when combined with y_i, it results in a reduction of the original primary noise. To create $\hat{y}_i$, the model learns to extract temporal and spectral patterns from y_i and generate the anti-noise signal through a series of convolutional layers. The labeling of the target $\hat{y}_i$ is done by taking the original UAV propeller audio, processing it to obtain the exact inverse signal that will effectively cancel out the noise. This process ensures that during training, the model learns to minimize the error between the generated ANC output and primary noise signal, thereby effectively learning to cancel out the noise. In equation 7, $\lambda\|W\|^2$ represents the regularization term to control overfitting. The training utilized the Adam optimizer with a learning rate of 0.001, proceeding for 75 epochs with a batch size of 32. Early stopping was implemented with patience of 5 epochs based on validation loss, ensuring that the model would generalize effectively across various noise scenarios.

3.3 Evaluation

To assess the performance of our deep learning model, rigorous evaluation was conducted, focusing on both quantitative metrics and visual performance indicators such as spectrograms. Once the model began receiving real-time propeller noise as input, it generated corresponding spectrograms as shown in Fig. 4, suitable for visualizing and identifying UAV noise characteristics. The first plot in Fig. 4 represents the Amplitude Spectrogram of the incoming UAV propeller noise. In this plot, the x-axis shows time, while the y-axis represents frequency. The color intensity corresponds to the amplitude of the frequency components at each time point. Brighter regions indicate higher amplitude, while darker regions indicate lower amplitude.

This spectrogram provides insight into the energy distribution of the noise signal across time and frequency. For UAV propeller noise, prominent frequency bands are observed, particularly in the low-frequency range (below 100 Hz) and high-frequency range (between 1-10 kHz). For example, at around 60 seconds, a peak in the amplitude is observed at approximately 2000 Hz, indicating a strong high-frequency component associated with the UAV' s propeller dynamics. Similarly, at around 37 seconds, the amplitude peaks around 500 Hz, where lower-frequency vibrations are most dominant. These fre-

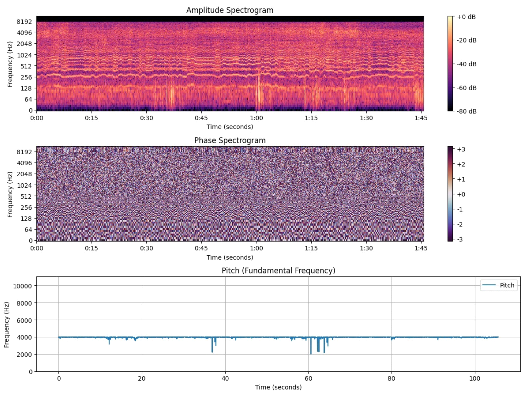


Fig. 4. Spectrogram Analysis for UAV Propeller Noise.

quency bands are significant as they provide insights into the primary sources of noise, which the model uses to focus its cancellation efforts. The second plot shows the Phase Spectrogram of the incoming propeller noise. The x and y-axis represent time and frequency respectively. However, in this plot, the color intensity is related to the phase of the frequency components rather than their amplitude. Phase represents the timing relationship between the different frequency components, which is essential for understanding the signal' s waveform structure.

In relation to UAV propeller noise, phase spectrogram reveals the phase shifts that occur across the frequency components. This information is crucial for the deep learning model in generating accurate anti-noise signals. While phase alone may not provide information about the energy or loudness of the noise, it is vital for synchronizing the cancellation signal to ensure effective destructive interference. A clear alignment of phase information between the original noise and the generated anti-noise signal can lead to more efficient noise cancellation. The third plot in Fig. 4 illustrates the Pitch Spectrogram or Fundamental Frequency of the incoming propeller noise. The pitch corresponds to the perceived frequency of the sound, primarily related to the fundamental frequency of the noise signal.

At around 60 seconds, the pitch spectrogram shows a strong peak around 2000 Hz, which is indicative of the fundamental frequency associated with the UAV' s propeller noise. This periodic pattern is characteristic of a constant rotational speed or a slight fluctuation in the propeller's rotation, which is periodic in nature. Such observations help the deep learning model identify and cancel repetitive elements of the noise. The model extracted critical features from those spectrograms, such as temporal frequency variations, amplitude, and phase shifts, enabling it to generate precise anti-noise signals that aligned with the primary noise. Through this alignment, the destructive interference phenomenon helped significantly in reducing the UAV noise.

The performance metrics further underscore the model's capability, achieving an impressive accuracy of 94.5%, alongside precision of 93.2%, recall of 96.1%, and an F1-score of 94.6%, as illustrated in Fig. 5. The achieved accuracy indicates a high level of reliability in correctly identifying noise events, suggesting that the model performs consistently across various noise scenarios. With 93.2% precision, the model successfully minimizes false positives, ensuring that noise events are rarely misclassified. High recall demonstrates the model's sensitivity to noise, ensuring that it detects nearly all noise events. This is particularly critical for real-world noise cancellation applications, where even subtle noise occurrences need to be captured. Furthermore, the F1-score reflects a well-balanced performance, combining high precision and recall. This indicates that model strikes an optimal balance between minimizing false positives and ensuring comprehensive noise detection. Finally, the loss value of 0.115 shows that the model generates anti-noise signals that closely match the expected output, reinforcing the effectiveness of the noise cancellation process. These high scores highlight the model' s accuracy and responsiveness, emphasizing its potential

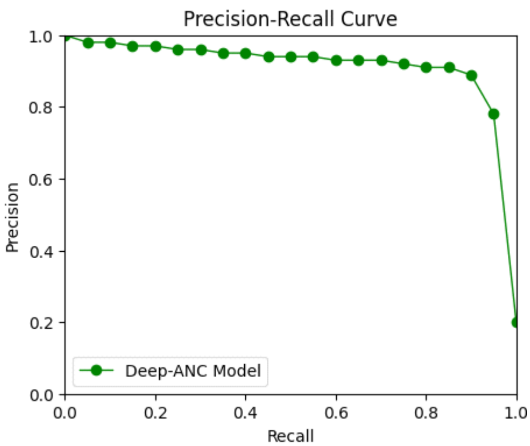


Fig. 5. Precision-Recall Curve for Deep-ANC Model.

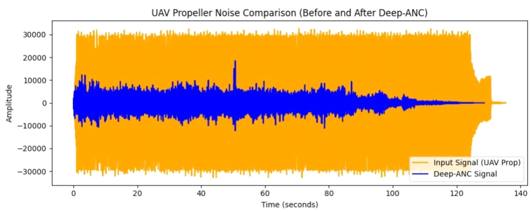


Fig. 6. Waveform comparisons for Input signal (UAV Propeller Noise) with Audio Signal Provided by Deep-ANC model.

for real-world applications. Additionally, Fig. 6 illustrates the amplitude reduction in UAV propeller noise before and after processing, demonstrating the tangible effect of the model's noise cancellation. The model achieved a 36 dB reduction in UAV propeller noise, decreasing the noise level from 85 dB (pre-processed) to 49 dB (post-processed). This significant reduction corresponds to a perceptually noticeable decrease, reinforcing the model's effectiveness in noise cancellation.

3.4 Hardware Implementation

To further validate our approach, we implemented real-time UAV noise cancellation using carefully selected hardware and configurations, allowing us to analyze the model's effectiveness in realistic operational conditions. This practical implementation showcased the model' s capability to mitigate propeller noise effectively during flights. Our setup included a quadcopter UAV equipped with an Nvidia Jetson NX for onboard processing, high-quality microphones (ReSpeaker Mic Array v2.0) for precise noise capturing from all directions, audio interface cards for sound signal transmission, and strategically placed speakers. The hardware setup of the UAV used for experimental validation is illustrated in Fig. 7. This hardware selection was meticulously curated to enable efficient real-time noise reduction, ensuring that components were not only compatible with each other but also with our proposed model for active noise cancellation as well.

Fig. 7. UAV hardware setup for testing the Active Noise Cancellation (ANC) system: The platform includes an Nvidia Jetson NX for real-time processing, ReSpeaker Mic Array for noise capture, and speakers for anti-noise emission, enabling effective reduction of propeller noise.

The UAV setup process involved precisely integrating the microphones and speakers onto the drone platform, ensuring strategic placement to maximize propeller noise capture and efficient anti-noise transmission. The CNN-based ANC model was deployed on the Jetson NX, alongside essential audio processing libraries and dependencies, enabling seamless real-time audio analysis and anti-noise generation.

To assess the ANC model' s impact, we conducted both indoor and outdoor UAV flight tests to evaluate the system' s effectiveness in different environments. Initially, we performed standard UAV flights without activating the ANC system, resulting in high noise levels due to the propeller operation. In the subsequent tests, we activated the ANC model and assessed its noisecanceling performance. During these flights, the propeller noise was captured by microphones and transmitted to the Jetson NX board via the audio interface cards. The model then processed this input to generate anti-noise signals, which were emitted through the speakers positioned towards each propeller. The out door flight test results, shown in Fig. 8, demonstrate the practical efficacy of the proposed system in real-world conditions. Fig. 9 illustrates the waveform plots of the received UAV pro-

Fig. 8. Outdoor flight test of the UAV equipped with the Active Noise Cancellation (ANC) system.

Table 1. Comparison of Traditional ANC Models and the Proposed Deep-ANC Model

Model	Accuracy	Loss	Amplitude Reduction	Sound Reduction	Reaction Time
LMS	83.63%	0.15	26%	14-18 dB	260-300 ms
FXLMS	88.07%	0.12	48%	20-24 dB	140-150 ms
DeepFilterNet2	92.48%	0.13	45%	18-20 dB	140-160 ms
Proposed Model	94.5%	0.115	56%	32-36 dB	50-70 ms

peller noise signal, which is then split across different channels, along with the extracted features. Fig. 10 shows the inverse noise signal generated by the model, and the combined effect of the primary and inverse signals, highlighting the significant noise reduction achieved through destructive interference in outdoor flight scenarios. This setup confirmed that the deep learning-based ANC system effectively reduces noise, facilitating quieter flights and proving its feasibility for real-world UAV applications. Table 1 presents a comparative analysis between traditional ANC approaches and our proposed Deep-ANC model. As shown, the proposed Deep-ANC model outperforms the other approaches in terms of accuracy, noise reduction, reaction time and overall performance.

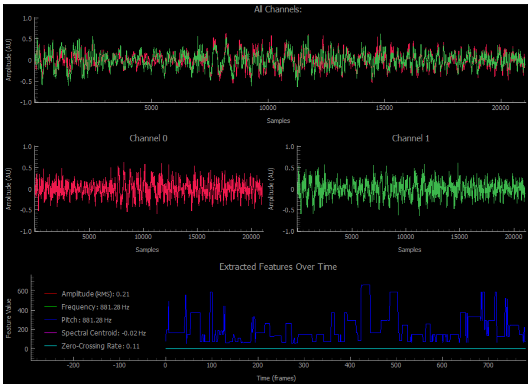


Fig. 9. Outdoor flight test results demonstrating various waveform-plots of received UAV Propeller Noise Signal, splitting across different channels and extracted features.

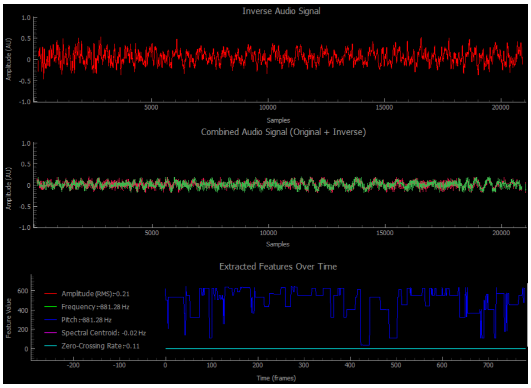


Fig. 10. Experimental Results illustrates the inverse noise signal generated by the model, and combined effect of the primary and inverse signals, showing significant noise reduction through destructive interference in outdoor flight scenarios.

Ⅳ. Conclusion

This study proposed a deep learning-based CNN model for active noise cancellation (ANC) of UAV propeller noise, demonstrating high effectiveness in noise reduction by leveraging CNN capabilities. The model underwent comprehensive training, testing, and validation, achieving impressive performance metrics: an accuracy of 94.5%, a precision of 93.2%, and a recall of 96.1%. The noise reduction results were validated across various flight scenarios, with the successful hardware implementation highlighting the model' s capacity to perform in dynamic operational settings. Overall the proposed system was able to achieve a remarkable 36 dB reduction in UAV propeller noise. This study underscores the potential of deep learning approaches for tackling complex noise challenges in UAV operations, offering a promising foundation for future noise mitigation strategies.

Future work will focus on enhancing the system' s efficiency and exploring much broader range of UAV mobility scenarios. Practical implementation in diverse conditions will serve as key performance indicators (KPIs) for evaluating the system' s robustness and adaptability. By expanding the range of scenarios and refining the ANC model, we aim to further improve noise cancellation effectiveness and operational efficiency in real-world applications.

References

[1] S. Sharma, P. Kulkarni, and P. Pathak, "Applications of unmanned aerial vehicles (uavs) for improved business management," in *2022 IIHC*, pp. 53-57, 2022. (https://doi.org/10.1109/IIHC55949.2022.10060638)

[2] S. Tyagi, A. R. Várkonyi-Kóczy, and S. Szénási, "A comprehensive investigation into the noise reduction techniques for speech," in *2023 IEEE 21st World Symp. Applied Mach. Intell. and Inf.* (*SAMI*), pp. 207-212, 2023. (https://doi.org/10.1109/SAMI58000.2023.10044486)

[3] B. Kang, H. Ahn, and H. Choo, "A software platform for noise reduction in sound sensor equipped drones," *IEEE Sensors J.*, vol. 19, no. 21, pp. 10121-10130, 2019. (https://doi.org/10.1109/JSEN.2019.2927370)

[4] F. A. Khan and S. Y. Shin, "Deep learning based active noise cancellation for reducing uav propeller sound," in *2024 The 15th Int. Conf. ICT Convergence* (*ICTC*), 2024. (https://doi.org/10.1109/3ICT56508.2022.9990841)

[5] S. M. Akarsh, R. Biradar, and P. V. Joshi, "Implementation of an active noise cancellation technique using deep learning," in *2022 ICICT*, pp. 249-253, 2022. (https://doi.org/10.1109/ICICT54344.2022.9850807)

[6] K. Srijomkwan, C. Weerasakul, S. Kittiweerawong, and J. Puntree, "The cancellation of human sounds using synthesized soundwaves," *J. Physics Conf. Series*, vol. 1380, no. 1, p. 12123, 2019. (https://doi.org/10.1088/1742-6596/1380/1/012123)

[7] B. S. K. Reddy, J. SaiGnaneswar, K. Balaji, K. S. Vivek Reddy, and K. L. Reddy, "Deep learning based image enhancing environment with noise suppression," in *2023 Second Int. Conf. Electr. and Renewable Syst.* (*ICEARS*), pp. 1339-1344, 2023. (https://doi.org/10.1109/ICEARS56392.2023.10084977)

[8] L. Wang and A. Cavallaro, "Deep-learning-assisted sound source localization from a flying drone," *IEEE Sensors J.*, vol. 22, no. 21, pp. 20828-20838, 2022. (https://doi.org/10.1109/JSEN.2022.3207660)

[9] M. Libina and K. Mahan, "Hybrid noise reduction and enhancement of audio quality using deep learning," in *2023 Second ICEARS*, pp. 1131-1135, 2023. (https://doi.org/10.1109/ICEARS56392.2023.10085247)

[10] D. Utebayeva and A. Yembergenova, "Study a deep learning-based audio classification for detecting the distance of UAV," in *2024 IEEE Int. Conf. Evolving and Adaptive Intell. Syst.* (*EAIS*), pp. 1-7, 2024. (https://doi.org/10.1109/EAIS58494.2024.10569107)

[11] S. Haykin and B. Widrow, "Deep learning based image enhancing environment with noise suppression," *Springer Sci.*, pp. 19-34, 2003. (https://doi.org/08.119/ICEARS56392.2003.10010077)

[12] T. S. Alexander, *Adaptive signal processing: Theory and applications*, Springer Science & Business Media, 2012. (https://doi.org/5499/ARRFS56392.2012.1846357)

[13] X. Li, W. Chen, Z. Liu, C. Lu, and M. Sun, "A survey on filtered-x least mean square-based active noise control systems with emphasis on reducing computational complexity," *EURASIP J. Advances in Signal Process.*, vol. 2023, no. 1, p. 128, 2023. (https://doi.org/11.7846/ICEARS56397.2023.35849675)

[14] S. Gaur and V. Gupta, "A review on filtered-x lms algorithm," *Int. J. Signal Process. Syst.*, vol. 4, no. 2, pp. 172-176, 2016. (https://doi.org/5471/ICDHW6382.2016.548617)

[15] J. Liu, H. Sun, M. Wu, and J. Yang, "Gpu

implementation of a fast multichannel wiener filter algorithm for active noise control," *IEEE Signal Process. Lett.*, 2024. (https://doi.org/21.3356/ IEEES5253.2024.10084977)

[16] W. Wu, Y. Xiao, J. Lin, L. Ma, and K. Khorasani, "An efficient filter bank structure for adaptive notch filtering and applications," *IEEE/ACM Trans. Audio, Speech, and Lang. Process.*, vol. 29, pp. 3226-3241, 2021. (https://doi.org/10.1109/IEEES56392.2021.849562)

[17] A. Mohanaprasad Kothandaraman, "Adaptive noise cancellation using stationary wavelet transform based wiener filter," *Third Int. Conf. Sensors and Related Networks* (*SENNET-12*), vol. 3, pp. 326-341, 2015. (https://doi.org/06.5836/SENNET5189.2015.875483984)

[18] H. Zhang and D. Wang, "Deep anc: A deep learning approach to active noise control," *Neural Netw.*, vol. 141, pp. 1-10, 2021, ISSN: 0893-6080. (https://doi.org//10.1016/j.neunet.2021.03.037) [Online] Available: https://www.sciencedirect.com/science/article/pii/S0893608021001258)

[19] L. Nanni, G. Maguolo, S. Brahnam, and M. Paci, "An ensemble of convolutional neural networks for audio classification," *Applied Sci.*, vol. 11, no. 13, p. 5796, 2021. (https://doi.org/10.5368/mdpi5189.2021.5315860)

[20] S. Im, S. Kim, S. Woo, et al., "Deep learning-assisted active noise control in a time-varying environment," *J. Mechanical Sci. and Technol.*, vol. 37, no. 3, pp. 1189-1196, Mar. 2023, Publisher Copyright: © 2023, The Korean Society of Mechanical Engineers and Springer-Verlag GmbH Germany, part of Springer Nature, ISSN: 1738-494X. (https://doi.org/10.1007/s12206-023-0206-2)

[21] S. Im, S. Kim, S. Woo, et al., "Drone noise reduction using deep convolutional autoencoder for uav acoustic sensor networks," *IEEE 16th Int. Conf. Mobile Ad Hoc and Sensor Syst. Wkshp.* (*MASSW*), vol. 7, no. 3, pp. 9-19, 2019, Publisher Copyright: © 2023, The Korean Society of Mechanical Engineers and Springer-Verlag GmbH Germany, part of Springer Nature. (https://doi.org/05.1118/ gh252-019-816752)

[22] R. Sule, A. Kolekar, K. Patel, and A. Tandle, "Selective noise cancellation using machine learning," in *2023 Int. Conf. Advances in Power, Signal, and Inf. Technol.* (*APSIT*), pp. 645-650, 2023. (https://doi.org/10.1109/APSIT58554.2023.10201773)

[23] E. Kazemi, F. Taherkhani, and L. Wang, "Semisupervised learning for noise suppression using deep reinforcement learning of contrastive features," *IEEE Sensors Lett.*, vol. 7, no. 4, pp. 1-4, 2023. (https://doi.org/10.1109/LSENS.2023.3264998)

[24] S. Park and D. Park, "Integrated 3d active noise cancellation simulation and synthesis platform using tcl," in *2023 IEEE 16th Int. Symp. Embedded Multicore/Many-core Systems-on-Chip* (*MCSoC*), pp. 111-116, 2023. (https://doi.org/10.1109/MCSoC60832.2023.00024)

[25] E. Nerantzis, A. Kazakis, G. Symeonidis, and G. A. Papakostas, "The effects of fully connected layers adjustment for lightweight convolutional neural networks," in *2022 Int. Conf. Innovation and Intell. Inf., Computi., and Technol.* (*3ICT*), pp. 50-57, 2022. (https://doi.org/10.1109/3ICT56508.2022.9990841)

[26] A. Z. bin Sha'ameri and T. J. Lynn, "Spectrogram time-frequency analysis and classification of digital modulation signals," in *2007 IEEE Int. Conf. Telecommun. and Malaysia Int. Conf. Commun.*, pp. 113-118, 2007. (https://doi.org/10.1109/ICTMICC.2007.4448616)

[27] H. Uyanık and M. Köseoğlu, "Performance evaluation of different window functions for audio fingerprint based audio search

algorithm," in *2020 4th Int. Symp. Multidisciplinary Studies and Innovative Technol.* (*ISMSIT*), pp. 1-4, 2020. (https://doi.org/10.1109/ISMSIT50672.2020.9255225)

[28] T. Anjum, T. E. Chowdhury, S. Sakib, and S. Kibria, "Performance analysis of convolutional neural network architectures for the identification of covid-19 from chest x-ray images," in *2022 IEEE 12th Annual CCWC*, pp. 0446-0452, 2022. (https://doi.org/10.1109/CCWC54503.2022.9720862)

Faisal Ayub Khan

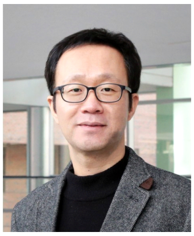
Jul. 2021 : BS(CS). degree, School of Computer Science and Computer Engineering, HITEC University

Mar. 2023 : MS. degree, School of IT Convergence and Engineering, Kumoh National Institute of Technology

<Research Interest> Artificial Intelligence, Machine Learning, Computer Vision, 5G/6G Wireless Communication Network Systems

[ORCID:0009-0002-1558-1453]

Soo Young Shin



Feb. 1999 : B.Eng. degree, School of Electrical and Electronic Engineering, Seoul National University

Feb. 2001 : M.Eng. degree, School of Electrical, Seoul National University

Feb. 2006 : Ph.D. degree, School of Electrical Engineering and Computer Science, Seoul National University

Sept. 2010-Current : Professor, School of IT Convergence and Engineering, Kumoh National Institute of Technology

<Research Interests> 5G/6G Wireless Communications and Networks, Signal Processing, Internet of Things, Mixed Reality, and Drone Applications.

[ORCID:0000-0002-2526-2395]