Deep Learning based Active Noise Cancellation for reducing UAV propeller sound

1st Faisal Ayub Khan
IT Convergence and Engineering
Kumoh National Institute of Technology
Gumi-si, South Korea
faisal@kumoh.ac.kr

2nd Soo Young Shin
IT Convergence and Engineering
Kumoh National Institute of Technology
Gumi-si, South Korea
wdragon@kumoh.ac.kr

***Abstract*—In this research, we investigate the application of Convolutional Neural Networks (CNN) utilizing deep learning methodologies to actively predict and cancel noise generated by UAV propellers. Our approach involves extracting features from spectrograms and waveform data to train a CNN model capable of predicting and generating counteractive noise signals for effective propeller noise cancellation. A comprehensive assessment of the proposed model was conducted, resulting in an accuracy of 94.5%, a precision of 93.3%, and a recall of 96.1%. Simulated testing in various scenarios showed promising results, demonstrating the model's capability in reducing drone propeller noise effectively. Furthermore, the model's performance remained robust under various simulated environmental conditions, indicating its potential for real-world applications. This research paves the way for more sophisticated noise reduction systems in UAVs, contributing to quieter and more environmentally friendly operations.**

***Index Terms*—Deep Learning, Unmanned Aerial Vehicles (UAVs), Adaptive ANC, Neural Network (NN), Real-time Noise Mapping.**

I. Introduction

Integration of Active Noise Cancellation (ANC) techniques in Unmanned Aerial Vehicles (UAVs) marks a significant advancement in aeronautical engineering, addressing the critical challenge of noise mitigation. As UAVs have become increasingly prevalent in diverse applications—ranging from agriculture and disaster management to environmental monitoring, surveillance, delivery, and reconnaissance—the need to minimize their acoustic footprint has become paramount [1]. ANC techniques can tackle the pervasive issues of noise pollution associated with UAVs by employing sophisticated algorithms and sensor arrays to detect and counteract the noise generated by propulsion systems and aerodynamic forces. By emitting anti-noise signals that are precisely out of phase with ambient noise, ANC effectively reduces the overall acoustic signature of Unmanned Aerial Vehicles. This not only enhances operational stealth but also mitigates environmental impact. The implementation of ANC in UAVs represents a transformative step towards the harmonious integration of unmanned aerial systems in populated or noise-sensitive areas, signifying a major breakthrough in aerospace engineering and facilitating the responsible and widespread adoption of UAV technology in everyday life. ANC operates on the principle of "destructive interference." This technique generates a secondary sound signal, known as anti-noise, which is equal in magnitude but 180° out of phase with the primary noise. The superposition of these primary and secondary sound signals results in residual noise, with an amplitude significantly lower than that of the primary noise [2]. The effectiveness of active noise canceling depends on the precision with which the anti-noise is generated, matching the primary noise in magnitude and phase opposition. Fig. 1 illustrates the basic concept of ANC, showing the interaction between the primary noise and the generated anti-noise to achieve residual noise.

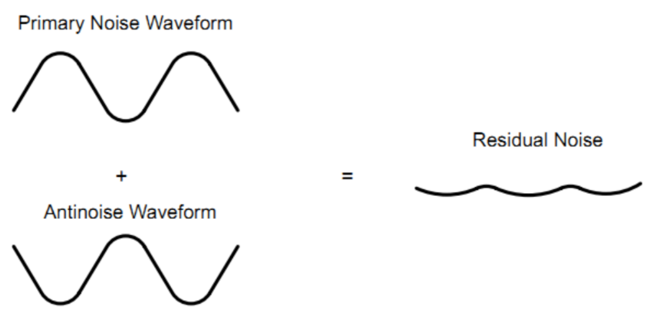

Fig. 1. Active Noise Cancellation Concept.

The application of Active Noise Cancellation (ANC) in Unmanned Aerial Vehicles (UAVs) becomes particularly effective and promising when combined with Deep Learning models approach [2]. Traditional ANC systems analyze incoming noise signals and generate corresponding anti-noise in real-time. Deep neural networks, trained on extensive datasets of UAV noise signatures, can swiftly identify and predict noise patterns during UAV operation. These neural networks extract various features from the UAV propeller noise, which are leveraged to enhance the system's predictive capability. Utilizing these extracted features, the ANC system can proactively generate anti-noise signals tailored to specific noise sources encountered during flight [3]. This dynamic adaptation enables more precise and efficient noise cancellation, even in complex and variable environments. Integrating deep learning into ANC for UAVs not only improves noise prediction accuracy but also allows the system to continuously learn and adapt from new data [4]. This innovative approach to noise mitigation paves the way for quieter, more efficient UAV operations, fostering a future where UAVs can operate seamlessly and unobtrusively in diverse settings.

II. Methodology

Deep Learning models are increasingly recognized as an effective method for analyzing and addressing complex noise cancellation tasks, particularly in scenarios involving propeller noise reduction. Propeller noise, known for its distinctive spectral patterns and time-varying characteristics, presents challenges that traditional signal processing methods struggle to effectively address. To combat propeller noise, CNNs utilize time-frequency representations, such as spectrograms derived from audio recordings. Each spectrogram frame serves as input to the CNN model. Typically, the CNN architecture consists of multiple convolutional and pooling layers designed for feature extraction and dimensionality reduction. Deeper layers within the network are responsible for learning intricate noise patterns and relationships, ultimately leading to fully connected layers that generate noise cancellation signals based on acquired features.

A CNN model designed for active noise cancellation (ANC) leverages its inherent ability to learn hierarchical representations of audio signals. This allows the model to effectively extract important features such as amplitude, frequency spectrum, and phase information. Initially, the raw audio signal undergoes pre-processing to prepare it for input into the CNN. This typically involves converting the audio waveform into a spectrogram representation, which captures both temporal and frequency domain information. The spectrogram is a 2D matrix where one axis represents time, the other axis represents frequency, and the magnitude of each point in the matrix corresponds to the signal strength at that time-frequency point [5]. Within the CNN architecture, the convolutional layers serve as feature extractors. These layers apply a set of learnable filters (kernels) across the input spectrogram, performing convolutions to extract localized patterns and features. By convolving these filters over the spectrogram, the CNN learns to identify meaningful features at different scales, including variations in amplitude, frequency content, and phase characteristics. The subsequent pooling layers downsample the feature maps, reducing spatial dimensions while preserving important features [6].

The learned features from the convolutional layers are then passed to fully connected layers, which integrate the extracted features into a higher-level representation suitable for decision-making. In the case of Active Noise Cancellation (ANC), the CNN learns to map the extracted audio features to an appropriate anti-noise signal. The CNN figures out how to create an anti-noise signal that can be combined with the original propeller noise to reduce the overall sound. By combining the original propeller noise with the anti-noise signal, the total noise level is reduced. This learned mapping effectively captures the relationship between the input audio features and the desired output, which is a significantly quieter audio signal. During the deep learning model training process, the CNN is optimized to minimize the difference between the predicted output (cancellation signal) and the primary audio signal (propeller sound). This optimization allows the CNN to learn how to cancel out specific characteristics encoded in the input features [7], leading to effective noise cancellation. Regularization techniques and extensive training of the datasets contribute to the model's improved capacity to generalize to unfamiliar audio signals and various types of noise. This ensures that the Deep Learning model can successfully reduce noise in new and different audio scenarios.

A. Proposed Model

The proposed Deep Learning-based approach for Active Noise Cancellation incorporates a convolutional neural network (CNN) to extract features from the UAV propeller noise Fig. 2. The CNN is integral to the model, allowing for precise noise pattern recognition and prediction. The model begins by extracting features from the received propeller audio signals X, including time-domain features (X_{time}), frequency-domain features (X_{freq}), time-frequency domain features (X_{tf}), and temporal characteristics (X_{temp}). After feature extraction, inverse values for all the features are generated which are later used to generate an anti-noise signal for the primary sound. These features are processed through multiple convolutional and pooling layers within the CNN to learn intricate noise patterns [8]. The fully connected layers then integrate these features into a high-level representation suitable for generating an anti-noise signal.

During the training phase, the model parameters W and β are optimized to minimize the difference between the combined audio (original propeller noise X plus the anti-noise signal H) for the target quieter sound, using backpropagation and optimization techniques. The optimization process can be summarized as minimizing the cost function $L(Y, \hat{Y})$, where Y is the target output and $\hat{Y}$ is the predicted output. Once trained, the model can extract features from new real-time audios X' and then generate its anti-noise signal H which is then combined with the primary audio signal to produce the final output Output $=$ SIGN$(X' + H)$. This dynamic and adaptive approach ensures precise and efficient noise cancellation, enhancing operational stealth and reducing the environmental impact of UAVs [9]. The detailed algorithm mentioned is the sections below, captures the step-by-step process of training and predicting with the proposed model, highlighting the systematic feature extraction and integration that underpin the deep learning-driven ANC.

B. Algorithm

As seen from Algorithm: 1, the process initiates with a training phase where a diverse set of features is extracted from the input audio signals. These features include characteristics from the time domain (such as amplitude and zero-crossing rate), attributes from the frequency domain (like spectral centroid and Mel-Frequency Cepstral Coefficients (MFCCs)), and properties from the time-frequency domain (including pitch

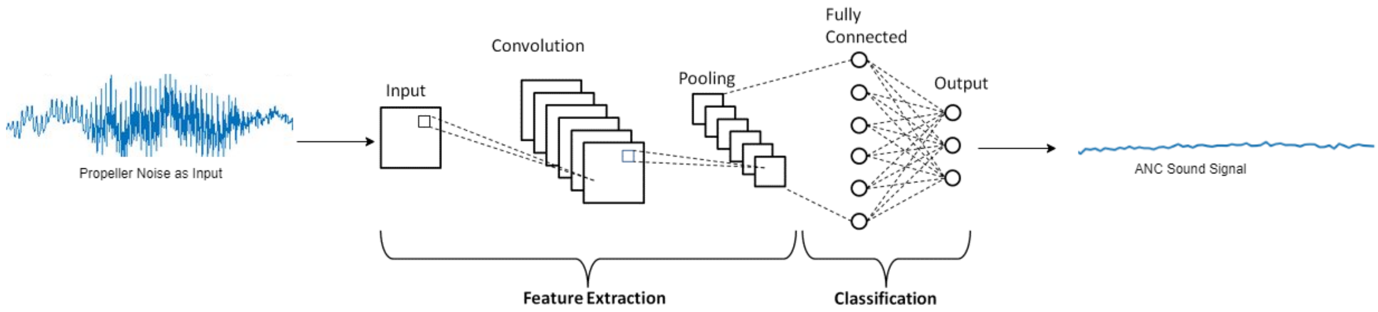

Fig. 2. Proposed Deep Learning model for Active Noise Cancellation of UAV propeller Noise

and phase). The deep learning model is defined and initialized with random weights and parameters, and it undergoes iterative training. During this training, the model learns to generate inverse feature values in response to primary sound features. The model also learns the complex patterns of propeller noise through backpropagation and optimization techniques. In the prediction phase, the trained model extracts features from new input audio signals. It then generates an anti-noise signal specifically designed to counteract the identified propeller noise. By merging this anti-noise signal with the original audio signal, the algorithm achieves effective noise cancellation, substantially reducing the propeller noise. This method exemplifies a sophisticated approach to ANC, leveraging the power of deep learning to adaptively and efficiently suppress propeller sound in real-time scenarios.

C. Spectrogram Visualization

The input audio signal $x(t)$ is transformed into a time-frequency representation using the Short-Time Fourier Transform (STFT):

$$X(t,\omega) = \mathrm{STFT}\{x(t)\}(t,\omega) = \int x(\tau)w(\tau - t)e^{-j\omega\tau}d\tau \tag{1}$$

where *X(t,ω)* represents the spectrogram, ω is the frequency, *t* is time, *x(τ)* is the audio signal, and *w(τ-t)* is a window function.

D. Convolutional Layer Operation and Feature Extraction

The convolution operation is applied to the input spectrogram given by *X(t,ω)* and using the set of learnable filters *h(t,ω)*. The output feature map $F(t',\omega')$ after convolution and activation is calculated as,

$$F\left(t',\omega'\right) = \sigma\left(\sum_{t,\omega} X(t,\omega)h\left(t',\omega'\right)\right) \tag{2}$$

whereas, σ is the activation function such as ReLU, t' and ω' are the indices of the output feature map. The convolutional layers are trained to identify particular characteristics such as edges, frequency patterns, and temporal correlations within the spectrogram. Every filter $h(t',\omega')$ is designed to detect and eliminate specific features that are important for noise detection and cancellation.

E. Pooling layer Downsampling

The max-pooling operation downsamples the feature maps [10] to reduce spatial dimensions while preserving important features. $P\left(t'',\omega''\right)$ denotes the pooled feature map which is given as:

$$P\left(t'',\omega''\right) = \max_{(t',\omega')\in \text{ pooling region}} F\left(t',\omega'\right) \tag{3}$$

F. Fully Connected Layers

The flattened and pooled feature maps are then passed through fully connected layers (dense layers) to learn high-

---

Algorithm 1 Active Noise Cancellation with Deep Learning

Require: Audio signals X (training set), Y (labels for training), X' (new input signals for prediction)
Ensure: Trained model parameters W, β and noise-cancelled output

```
TRAIN(X, Y)
Extract features from input audio signals X
    X_time ← time-domain features from X
    X_freq ← frequency-domain features from X
    X_tf ← time-frequency domain features from X
    X_temp ← temporal characteristics from X
Define and initialize the deep learning model
    W ← initialize weights randomly
    β ← initialize model parameters
Generate inverse feature values
Train the model using X and Y
    Update β using backpropagation and optimization
return W, β
PREDICT(X')
Extract features from new input audio signals X'
    X'_time ← time-domain features from X'
    X'_freq ← frequency-domain features from X'
    X'_tf ← time-frequency domain features from X'
    X'_temp ← temporal characteristics from X'
Generate anti-noise signal using trained model
    H ← model(X', W, β)
Combine primary audio signal with anti-noise signal
    Output ← SIGN(X' + H)
return Output
```

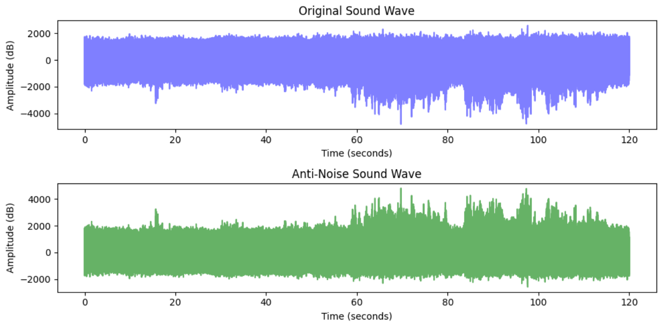

Fig. 3. Primary audio waveform of propeller sound and its Anti-Noise waveform.

level representations:

$$y = \text{softmax}\left(W^{(L)} \ldots \text{ReLU}\left(W^{(1)}x + b^{(1)}\right) \cdots + b^{(L)}\right) \tag{4}$$

where $W^{(l)}$ and $b^{(l)}$ are the weights and biases of layer l, and ReLU denotes the rectified linear unit activation.

G. Generating Anti-Noise Sound Signal

The output, y, represents the predicted cancellation signal. During training, the CNN model for ANC learns to minimize the loss, L, between y and the primary audio signal of the propeller sound:

$$L = \text{MSE}\left(y, x_{\text{clean}}\right) = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - x_{\text{clean},i}\right)^2 \tag{5}$$

To achieve controlled residual sound, a secondary sound waveform must be generated to counteract the primary sound waveform, as illustrated in Fig. 3. This secondary waveform should be out of phase with the primary sound, typically achieving a 180° phase inversion. This requires generating an inverse version of the primary sound signal so that when the primary and anti-noise signals are combined, phase elimination occurs. As mentioned earlier, anti-noise operates using the principle of destructive interference, which means that when sound waves with opposite phases combine, a phase shift occurs, resulting in the mitigation of the sound signal. This phenomenon is further explained below, using a digital signal as an example.

$$s[n] = \left(A \times \sin\left(2\pi f_n \Delta t + \varphi\right)\right) \ldots \tag{6}$$

whereas, s[n] audio signal value at n^{th} sample, Δt shows the total time taken between consecutive audio samples and Fn is the audio signal frequency in Hz. Since an inverse phase signal is needed it is given by:

$$s[n] = -\left(A \times \sin\left(2\pi f_n \Delta t + \varphi\right)\right) \ldots \tag{7}$$

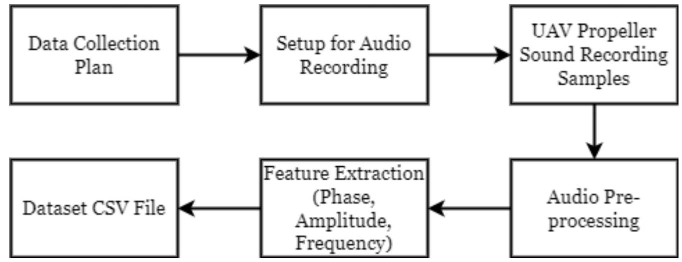

Fig. 4. Flowchart for UAV Propeller Audio Data Collection Process.

III. Experiment

To validate the effectiveness of our model, we conducted comprehensive simulations and tests across various conditions to ensure it met the desired performance metrics. The following subsections will provide detailed discussions on the preparation and preprocessing of the dataset, the training and testing of the Deep Learning model, and the simulation results.

A. Dataset Collection

The first crucial step in this research was establishing an extensive dataset by recording various sounds produced by UAV propellers, as well as acquiring propeller sounds from other sources. This dataset of over 1800 audio recordings served as the foundation for training and evaluating the deep learning models. The primary objective was to ensure the models' effectiveness in accurately identifying and eliminating UAV propeller noise from audio recordings. The process of collecting audio features from the recordings in the dataset is detailed in Fig. 4. The dataset included a range of features such as frequency, pitch, phase, amplitude envelope, and temporal characteristics. This was accomplished by writing a script to extract feature data from audio files, which were then used to generate a CSV file, as illustrated in Fig. 4.

B. Model Training

The deep learning model utilized in this study for active noise cancellation (ANC) was meticulously trained on an extensive dataset of UAV propeller sounds collected from diverse environments, including various altitudes and UAV speeds. Each audio sample was transformed into spectrogram representations using a frame size of 2048 and a hop length of 512, capturing detailed frequency and time-domain information essential for effective noise cancellation.

The CNN architecture comprised multiple convolutional layers, each followed by max-pooling layers for efficient feature extraction. The convolutional layers applied filters to capture spatial patterns within the spectrograms, leveraging cross-correlation for robust feature extraction. Rectified Linear Unit (ReLU) activation functions were incorporated after each convolutional and fully connected layer to introduce non-linearity, enhancing the model's ability to learn complex patterns. The deeper layers of the CNN were tasked with learning intricate wave patterns, while the fully connected

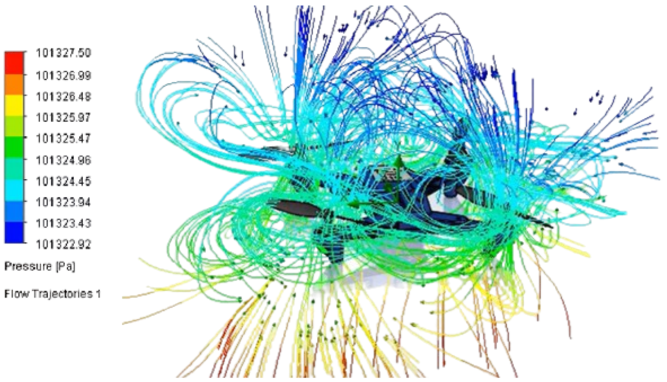

Fig. 5. Visualization of airflow and pressure distribution around a UAV, ranging variations from 101322.92 Pa (blue) to 101327.50 Pa (red)

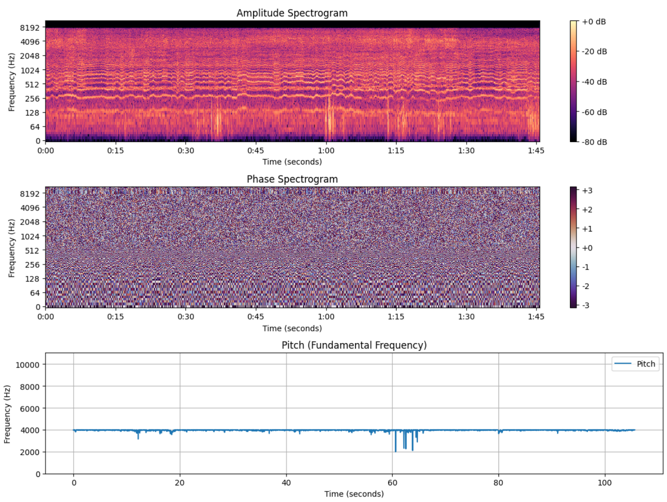

Fig. 6. Spectrogram Analysis for UAV Propeller Noise.

layers generated noise cancellation signals based on the extracted features. The final layer of the CNN employed a linear activation function to produce noise cancellation signal.

During the training process, mean squared error (MSE) was utilized as the loss function to minimize the disparity between the predicted noise-canceled spectrogram and the primary audio spectrogram. The Adam optimizer was employed with a learning rate of 0.001, facilitating efficient convergence. The training was conducted over 50 epochs with a batch size of 32 and a sampling rate of 22050 Hz. To prevent overfitting and ensure generalization across various noise scenarios, early stopping was implemented based on validation loss given by:

$$\text{MSE (Loss Function)} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \quad (8)$$

whereas, N denotes number of training samples used, y_i represents the actual target (primary audio) for the i^{th} sample and $\hat{y}_i$ is the predicted output (ANC signal) for the i^{th} sample.

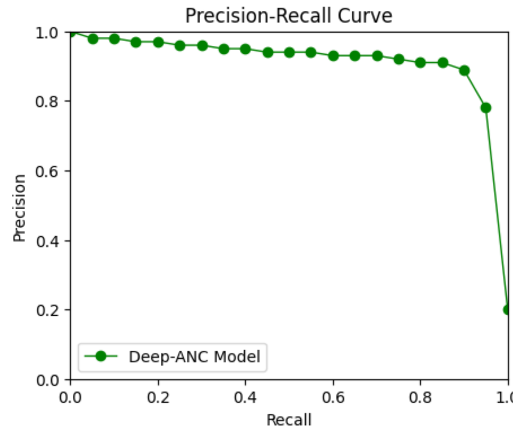

Fig. 7. Precision-Recall Curve for Deep-ANC Model.

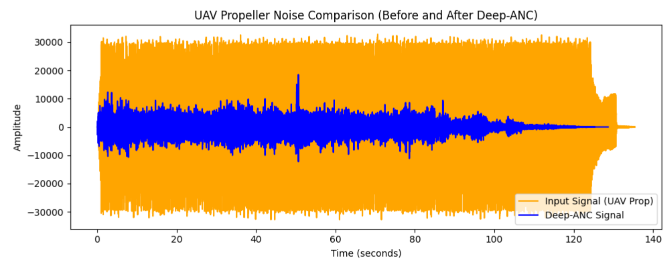

Fig. 8. Waveform comparisons for Input signal (UAV Propeller Noise) with Audio Signal Provided by Deep-ANC model.

C. Evaluation

In this section, we will assess the performance of our proposed deep learning model designed for active noise cancellation of UAV propeller noise. Noise is generated around UAV propellers due to the turbulent airflow and rapid pressure changes that occur during UAV operations Fig. 5. During the evaluation, we comprehensively analyzed several parameters, including amplitude and phase spectrograms, accuracy, precision, recall, and F1-score. Upon receiving the audio input, the Deep-ANC model analyzes various features in real-time. Utilizing the data from these extracted features, an inverse audio signal is generated to counteract the actual UAV propeller noise. After a few iterations, the model also predicts upcoming variations in the propeller sound. The model incorporates three essential wave plots that depict significant aspects of UAV propeller noise analysis Fig. 6. The phase spectrogram wave plot showcases the temporal progression of phase characteristics across different frequency bands, offering insights into the fluctuation of the propeller noise phase over time. This information is utilized to generate an inverse phase, which is 180° out of phase with the original sound. The amplitude spectrogram wave plot emphasizes the distribution of noise amplitudes, crucial for understanding the frequency components and variations in the intensity of the propeller noise. Finally, the pitch (fundamental frequency) plot visually captures changes in the dominant frequency of the noise signal, aiding in the identification of fundamental pitch patterns. Together, these wave plots provide a comprehensive overview of the acoustic characteristics of UAV propeller noise, enabling informed analysis and subsequent noise cancellation strategies using our CNN-based deep learning model. The system extracted various features from primary audio with their respective value ranges. The proposed model effectively generates precise anti-noise signals that counter the UAV propeller noise as a result of the destructive interference phenomenon, resulting in significantly quieter operation. The proposed Deep Learning-ANC model achieved impressive performance metrics, with an accuracy of 94.5%, precision of 93.2%, recall of 96.1%, and an F1 score of 94.6%. Fig. 7 illustrates the performance of the Deep-ANC model in predicting UAV noise. The model maintains high precision across a wide range of recall values, demonstrating effective noise prediction and cancellation capabilities with a trade-off between precision and recall as the recall increases. Additionally, the model demonstrated a low loss value of 0.115, indicating its effectiveness in accurately identifying and canceling UAV propeller noise. Fig. 8 shows the comparison between the original UAV sound (before) and the audio after applying our CNN-based deep learning model. The results demonstrate that our model significantly reduces the propeller noise to a minimal level.

IV. Conclusion

Unmanned Aerial Vehicles (UAVs) are increasingly being utilized across various fields, including agriculture, surveillance, delivery services, and environmental monitoring. However, to make them more acceptable and noise-less in urban and sensitive environments, it is crucial to address the issue of noise generated by their propellers. This study proposed a deep learning-based CNN model for active noise cancellation of UAV propeller noise. The proposed model is able to predict the upcoming UAV propeller noise, allowing it to effectively anticipate and counteract noise disturbances. It has proven to be highly effective in reducing UAV propeller noise by leveraging the power of CNN. The model underwent rigorous training, testing, and validation processes, yielding impressive performance metrics: an accuracy of 94.5%, precision of 93.2%, and recall of 96.1%. The proposed model achieved remarkable noise reduction results, validated across various flight scenarios. This study underscores the potential of deep learning approaches in addressing complex noise challenges in UAV operations, providing a promising avenue for future noise mitigation strategies. Future work will focus on enhancing the system's efficiency, hardware implementation, and analyzing numerous noise scenarios during UAV flights. Practical implementation in diverse environments and conditions will be key performance indicators (KPIs) for evaluating the system's robustness and adaptability. By expanding the scope of scenarios and refining the ANC model, we aim to further improve noise cancellation effectiveness and operational efficiency in real-world applications to achieve quieter operations.

Acknowledgment

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2024-RS-2024-00437190) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation).

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ICAN (ICT Challenge and Advanced Network of HRD) program (IITP-2024-RS-2022-00156394) supervised by the IITP (Institute of Information & Communications Technology Planning & Evaluation).

References

[1] S. Tyagi, A. R. Várkonyi-Kóczy, and S. Szénási, "A comprehensive investigation into the noise reduction techniques for speech," in *2023 IEEE 21st World Symposium on Applied Machine Intelligence and Informatics (SAMI)*, 2023, pp. 000 207–000 212.

[2] Sivanantham, 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 International Conference on Electronics and Renewable Systems (ICEARS)*, 2023, pp. 1339–1344.

[3] L. M and M. K, "Hybrid noise reduction and enhancement of audio quality using deep learning," in *2023 Second International Conference on Electronics and Renewable Systems (ICEARS)*, 2023, pp. 1131–1135.

[4] D. Utebayeva and A. Yembergenova, "Study a deep learning-based audio classification for detecting the distance of uav," in *2024 IEEE International Conference on Evolving and Adaptive Intelligent Systems (EAIS)*, 2024, pp. 1–7.

[5] E. Kazemi, F. Taherkhani, and L. Wang, "Semisupervised learning for noise suppression using deep reinforcement learning of contrastive features," *IEEE Sensors Letters*, vol. 7, no. 4, pp. 1–4, 2023.

[6] S. Park and D. Park, "Integrated 3d active noise cancellation simulation and synthesis platform using tcl," in *2023 IEEE 16th International Symposium on Embedded Multicore/Many-core Systems-on-Chip (MCSoC)*, 2023, pp. 111–116.

[7] R. Sule, A. Kolekar, K. Patel, and A. Tandle, "Selective noise cancellation using machine learning," in *2023 International Conference in Advances in Power, Signal, and Information Technology (APSIT)*, 2023, pp. 645–650.

[8] Y. Tsa, H.-C. Chu, S.-H. Fang, J. Lee, and C.-M. Lin, "Adaptive noise cancellation using deep cerebellar model articulation controller," *IEEE Access*, vol. 6, pp. 37 395–37 402, 2018.

[9] L. Wang and A. Cavallaro, "Deep-learning-assisted sound source localization from a flying drone," *IEEE Sensors Journal*, vol. 22, no. 21, pp. 20 828–20 838, 2022.

[10] A. S. M, R. Biradar, and P. V. Joshi, "Implementation of an active noise cancellation technique using deep learning," in *2022 International Conference on Inventive Computation Technologies (ICICT)*, 2022, pp. 249–253.