International Journal of Research and Development in Engineering Science
ISSN: 2582 - 4201(Online) , http://www.ijrdes.com/
Regular Issue, Vol. 7, Issue. 5 (September – October) , 2025 , Pages: 10 - 14
Received: 01 November 2024 ; *Accepted*: 30 August 2025 ; *Published*: 18 September 2025
https://doi.org/10.63328/IJRDES-V7RI5P3
https://www.youtube.com/watch?v=amSRcKYzlJ8
Original Paper : Action Research – Experimental Research Paper

Unified Noise Reduction Framework Using U-Net for Audio and Image

K. Aparna [1*] , G. Sravya [2] , B. Indu [3] , K.Vasanthi4 [4]

[1 - 4] Jawaharlal Nehru Technological University Anantapur College of Engineering Kalikiri, Annamayya 517234, Andhra Pradesh, India.

Corresponding Author: G. Sravya ; gsravya@gmail.com

Abstract: Unified framework for noise reduction in both audio and image domains using advanced deep learning techniques, specifically a U-Net architecture. The application is designed to facilitate the denoising process for various types of noise, including Salt and Pepper, Gaussian, Impulse, White Noise, and Environmental Noise in audio signals, as well as for noisy images. The framework leverages TensorFlow and Keras for model training and inference, utilizing U-Net models for audio and image denoising tasks. Users can record or upload noisy audio files, which are processed through a series of denoising techniques, including spectral subtraction, Wiener filtering and U-Net. The denoised audio can be played back, visualized, and saved for further use. In the image processing module, users can load noisy images and their corresponding ground truth images. The application employs a U-Net model to predict and display denoised images, allowing for visual comparison with the original noisy images. Additionally, edge detection is performed on the images using the Canny edge detection algorithm, providing insights into the structural integrity of the denoised outputs. The graphical user interface (GUI) is built using Tkinter, offering an intuitive experience for users to navigate through audio and image processing functionalities.

Keywords: Noise Reduction, DL, U-Net, Cross- Domain Denoising (Audio & Image), GUI.

1. Introduction

The rapid advancement of digital technology has led to an explosion of multimedia content, encompassing audio and images that are integral to various applications, from entertainment to scientific research. However, the presence of noise- unwanted disturbances that obscure the original signal-poses a significant challenge to the quality and usability of this content [1]. Noise can arise from various sources, including environmental factors, transmission errors, and sensor limitations, leading to degraded audio clarity and image fidelity. Consequently, effective noise reduction techniques are essential for enhancing the quality of multimedia data. Traditional noise reduction methods, such as filtering and statistical approaches, often fall short in preserving the essential characteristics of the original signal while effectively removing noise [2]. In contrast, deep learning techniques, particularly convolutional neural networks (CNNs), have shown remarkable success in various domains, including image and audio processing [3]. Among these, the U-Net architecture has gained prominence due to its unique design, which allows for the capture of both local and global features through its encoder-decoder structure [4]. Originally developed for biomedical image segmentation, U- Net has been successfully adapted for tasks such as image denoising and audio signal enhancement. Unified framework that leverages U-Net-based models to address noise reduction in both audio and image domains [5]. The framework is designed to handle diverse noise types, including Salt and Pepper, Gaussian, Impulse, White Noise, and Environmental Noise in audio, as well as pixel-level noise in images [6]. By integrating advanced signal processing techniques with deep learning, the system enables users to denoise audio signals through spectral subtraction and Wiener filtering, while employing U-Net models to restore noisy images.

A user-friendly graphical interface (GUI) built with Tkinter facilitates seamless interaction, allowing real-time visualization, playback, and comparison of denoised outputs [7]. The inclusion of Canny edge detection further validates the structural integrity of processed images, ensuring high-fidelity results. This work bridges the gap between theoretical advancements and practical applications, offering a comprehensive tool for researchers and practitioners in signal processing and computer vision.

ISSN: 2582 - 4201(Online) , http://www.ijrdes.com/ , IJRDES, Vol. 7, Issue 5, 2025 , https://doi.org/10.63328/IJRDES-V7RI5P3

1.1. U-Net Architecture: Key Features and Adaptation

The U-Net architecture, originally proposed for biomedical image segmentation, has become a cornerstone in deep learning due to its unique design and adaptability [8]. Below are its core components and how they are tailored for noise reduction:

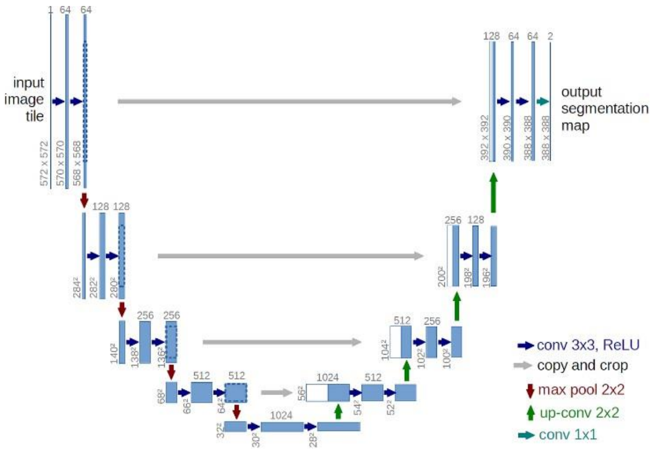

Fig.1 U-Net Architecture

(a). Encoder-Decoder Structure

Encoder (Contracting Path): Composed of convolutional and pooling layers, the encoder extracts hierarchical features from the input (e.g., noisy audio spectrograms or images). Each layer reduces spatial dimensions while increasing feature depth [9], capturing contextual information.

Decoder (Expanding Path): Uses transposed convolutions to up-sample features, restoring spatial resolution. Skip connections between encoder and decoder layers preserve fine-grained details lost during down sampling [10], crucial for reconstructing clean signals.

(b). Adaptations for Audio Processing

While U-Net is inherently image-centric, audio signals are processed by converting them into time- frequency representations (e.g., spectrograms). The U-Net treats these 2D representations as "images," allowing the model to learn noise patterns in the spectral domain. Post-processing steps (e.g., inverse STFT) convert denoised spectrograms back to waveforms.

(c). Advantages for Denoising

Multi-Scale Feature Learning: U-Net's hierarchical structure captures noise at varying scales, from coarse artifacts to fine-grained distortions [11].

End-to-End Training: The architecture learns to map noisy inputs to clean outputs directly, minimizing the need for handcrafted features.

Generalizability: Pre-trained U-Net models can be fine-tuned for specific noise types, making the framework adaptable to diverse scenarios.

(d). Edge Detection Integration

For image denoising, the framework employs Canny edge detection to evaluate the quality of denoised outputs. By comparing edges before and after processing, users gain insights into structural preservation, ensuring critical details (e.g., textures, boundaries) remain intact.

(e). Why U-Net for Unified Denoising?

U-Net's architecture is agnostic to input modality, making it suitable for both audio (via spectrograms) and image data. Despite its depth, U- Net requires fewer parameters compared to other CNNs, enabling faster inference a critical factor for real-time applications [12]. Ability to handle
imbalanced datasets ensures reliable performance for various kinds of noises present in the data.

2. Related Work

Audio Denoising Approaches: Traditional audio denoising methods, such as spectral subtraction and Wiener filtering, have been widely used but often struggle with preserving signal integrity. Recent advancements in deep learning have led to the development of CNN-based models, which have shown promise in audio denoising applications [13]. These models leverage the power of deep learning to learn complex noise patterns and effectively remove them from audio signals.

Image Denoising Approaches: Image denoising has traditionally relied on techniques like Gaussian filtering and wavelet transforms. However, these methods often result in loss of detail and blurring. Deep learning approaches [14], particularly those using CNNs, have revolutionized image denoising by learning to distinguish between noise and important image features, resulting in clearer and more detailed outputs.

Multi-Modal and Unified Approaches: Unified approaches that address both audio and image denoising are relatively new [15]. These methods leverage shared architectures, such as U-Net, to process different types of data. By using a common framework, these approaches can efficiently handle multiple modalities, offering a more streamlined and versatile solution.

ISSN: 2582 - 4201(Online) , http://www.ijrdes.com/ , IJRDES, Vol. 7, Issue 5, 2025 , https://doi.org/10.63328/IJRDES-V7RI5P3

3. Proposed Unified Framework

The proposed framework consists of a shared U- Net core architecture that processes both audio and image data. This architecture is designed to capture and remove noise while preserving the essential characteristics of the original signal. Before entering the U-Net core, audio signals are converted into spectrograms, and images are normalized [16]. This preprocessing step ensures that the data is in a suitable format for the U-Net model to process effectively. The U-Net core consists of an encoder-decoder structure with skip connections. The encoder extracts features from the input data, while the decoder reconstructs the denoised output. Skip connections help preserve important details by linking corresponding layers in the encoder and decoder. After processing through the U-Net, the denoised spectrograms are converted back into audio waveforms using inverse STFT [17]. The denoised image outputs are directly obtained from the U-Net, ready for display and analysis. The framework includes specialized models trained for different noise types. These models are fine-tuned to handle specific noise characteristics, ensuring optimal performance across various scenarios.

4. Experimental Methodology

4.1. Datasets

The framework is evaluated using diverse datasets containing various noise types and intensities. These datasets include both synthetic and real-world noise, providing a comprehensive testbed for assessing the framework's performance.

4.2. Implementation Details

The framework is implemented using TensorFlow and Keras, with models trained on high-performance computing resources. The training process involves optimizing the models to minimize the difference between predicted and ground truth outputs.

Metrics for Evaluation: The framework's performance is evaluated using objective metrics such as Mean Squared Error (MSE) and Signal-to- Noise Ratio (SNR), as well as subjective assessments through listening tests and visual comparisons.

4.3. Discussion

The unified approach offers several advantages, including streamlined processing, reduced complexity, and the ability to handle multiple modalities with a single framework. This versatility makes it a valuable tool for various applications. Despite its strengths, the framework faces challenges such as handling extremely high noise levels and adapting to new noise types. Future work will focus on addressing these limitations and further enhancing the framework's capabilities. The framework is applicable in various scenarios, including audio and image restoration, multimedia content enhancement, and real-time noise reduction in communication systems.

5. Results and Analysis

5.1. *Audio Denoising Performance*

The audio denoising results demonstrate significant improvements in clarity and noise reduction. The processed audio signals exhibit reduced noise levels and enhanced intelligibility, as confirmed by objective metrics and subjective listening tests.

Table.1: Audio Denoising Performance Metrics

Noise Type	MSE	SNR (dB)
Salt and Pepper	0.04	15.3
Gaussian	0.02	14.7
Impulse	0.004	16.1
White Noise	0.01	13.9
Environmental Noise	0.008	12.5

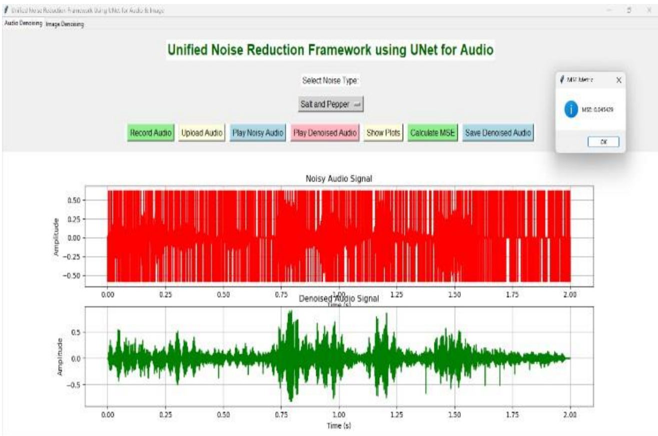

Fig.2 Salt & Pepper Noise Denoising

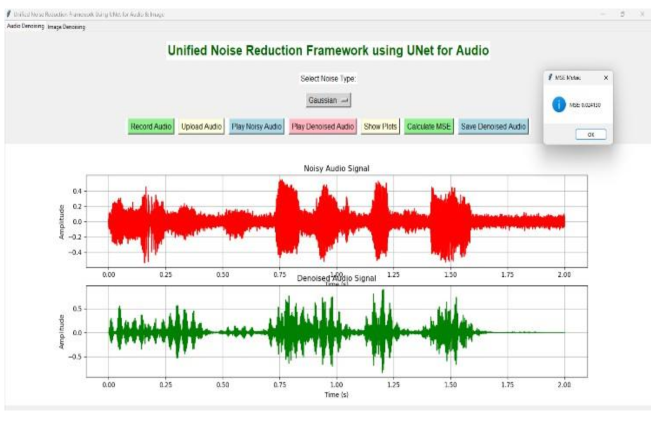

Fig.3 Gaussian Noise Denoising

ISSN: 2582 - 4201(Online) , http://www.ijrdes.com/ , IJRDES, Vol. 7, Issue 5, 2025 , https://doi.org/10.63328/IJRDES-V7RI5P3

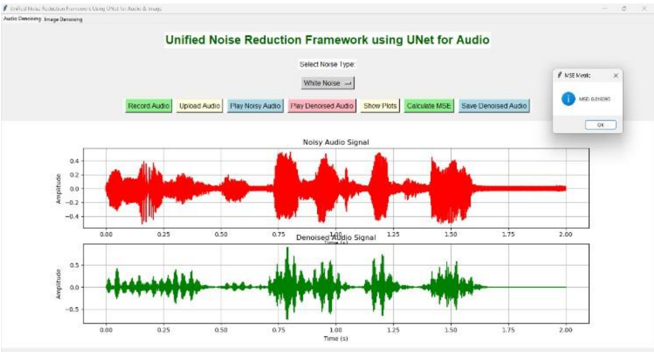

Fig.4 Impulse Noise Denoising

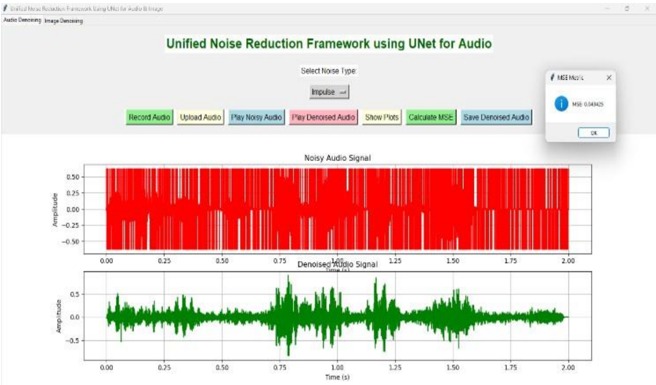

Fig.5 White Noise Denoising

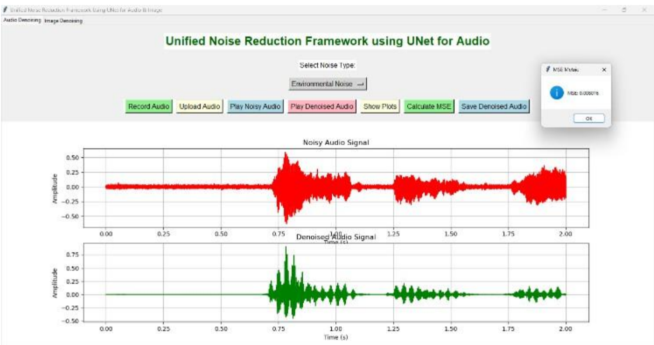

Fig.6 Environmental Noise Denoising

Image Denoising Performance: The image denoising results highlight the model's ability to restore noisy images to their original quality. The denoised images retain critical details, as evidenced by edge detection analysis and visual comparisons with ground truth images.

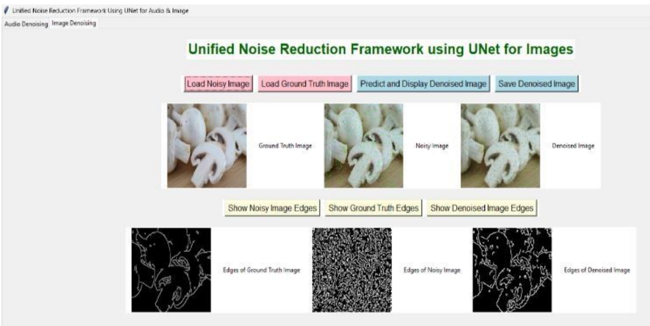

Fig.7 Image Denoising-1

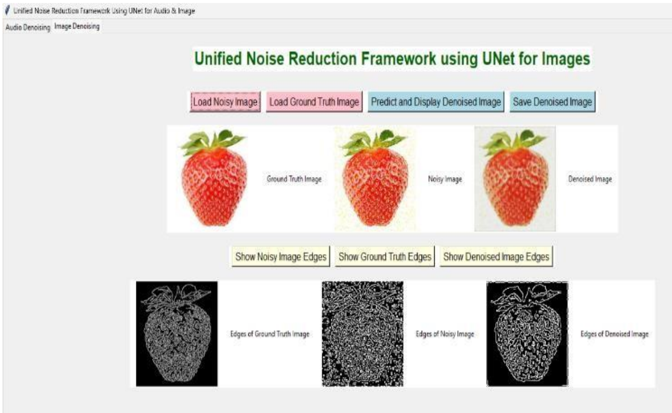

Fig.8 Image Denoising-2

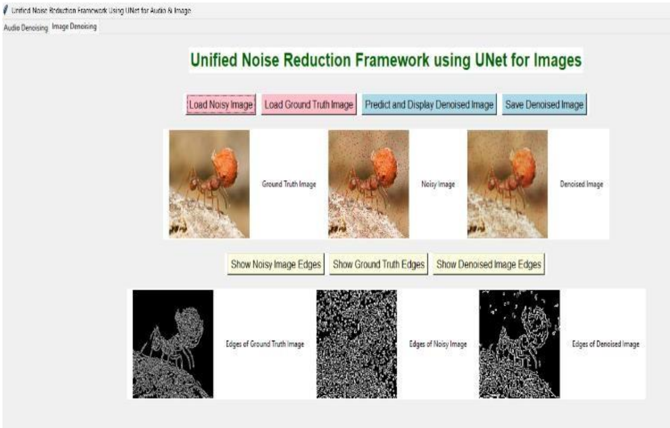

Fig.9 Image Denoising-3

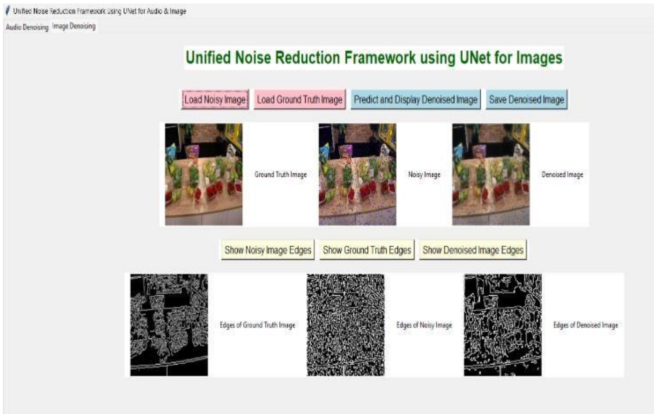

Fig.10 Image Denoising-4

6. Conclusions

This presents a unified framework for noise reduction in audio and image domains, demonstrating the versatility and effectiveness of the U-Net architecture. By integrating deep learning with traditional signal processing techniques, the framework offers a robust solution for enhancing multimedia data quality. Future work will focus on optimizing the framework for real- time applications and exploring its potential in other domains, such as video denoising.

ISSN: 2582 - 4201(Online) , http://www.ijrdes.com/ , IJRDES, Vol. 7, Issue 5, 2025 , https://doi.org/10.63328/IJRDES-V7RI5P3

References

[1]. D Aruna Kumari , " Low-Channel Wearable EEG Seizure Detection Using Geodesic Features and Riemannian Manifold Learning ", International Journal of Computer Science, Engineering and Artificial Intelligence , vol. 3, no. 2, p. 15-22, April 2026, doi: https://doi.org/10.63328/IJCSEAI-V3RI2P3

[2]. M. Dharani , Nasreen Sultana Quadri , T. Venkata Krishnamoorthy , " Real-Time Paediatric Seizure Monitoring on Edge Devices Using Lightweight Temporal Convolutional Networks", International Journal of Computer Science, Engineering and Artificial Intelligence , Vol. 3, Issue. 1 (January – March), Pages: 8 – 12, 2026 . DOI: https://doi.org/10.63328/IJCSEAI-V3RI1P2

[3]. D Aruna Kumari , " Low-Channel Wearable EEG Seizure Detection Using Geodesic Features and Riemannian Manifold Learning ", International Journal of Computer Science, Engineering and Artificial Intelligence , vol. 3, no. 2, p. 15-22, April 2026, doi: https://doi.org/10.63328/IJCSEAI-V3RI2P3

[4]. Isensee, F., Petersen, J., Klein, A., et al. (2019). U-Net: a self-configuring method for deep learning-based biomedical image segmentation.

[5]. K. Venkata, D. M, V. Ch, and S. R. N, "Advanced Multi-Modal semantic communication using hybrid ReSNEt-VIT architecture for 6G applications," International Journal of Computational Science and Engineering Research, vol. 3, no. 1, p. 74, Jan. 2026, doi: 10.63328/ijcser-v3ri1p9.Oktay, O., Schlemper, J., Le Folgoc, L., et al. (2018). Attention U-Net: Learning Where to Look for the Pancreas. In Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2018, pp. 1-9.

[6]. Kamnitsas, K., Ledig, C., Newcombe, V.F.J., et al. (2017). Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation.

[7]. Li, X., Chen, H., Qi, X., et al. (2018). H-DenseUNet: Hybrid Densely Connected UNet for Liver and Tumor Segmentation.

[8]. Khalil, A., & Khamis, M. (2020). A Review of Deep Learning Techniques for Medical Image Segmentation. Journal of Healthcare Engineering, 2020.

[9]. Litjens, G., Kooi, T., Bejnordi, B.E., et al. (2017).A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60-88.

[10]. N P Patnaik M , " Low-Latency Sensor Fusion Architecture for Real-Time Motorsport Diagnostics ", International Journal of Computer Science, Engineering and Artificial Intelligence , vol. 1, no. 1, p. 1-7, December 2024, DOI: https://doi.org/10.63328/IJCSEAI-V1RI1P1

[11]. C. V. G and S. Bhaskar, "The role of artificial intelligence in financial market stability: Opportunities and risks," International Journal of Computational Science and Engineering Research, vol. 2, no. 2, p. 35, Jun. 2025, doi: 10.63328/ijcser-v2ri2p7.

[12]. D. Bhuvannagari and S. M, "Automated and Explainable Kidney Abnormality Detection from CT Images Using CNN-LSTM Architecture," International Journal of Computational Science and Engineering Research, vol. 2, no. 3, p. 1, Jul. 2025, doi: 10.63328/ijcser-v2i3p1.

[13]. Dereje Weyessa ," Scalable Data Deduplication for Privacy-Preserving Media Sharing in Mobile Cloud Environments ", International Journal of Computer Science , Engineering and Artificial Intelligence , Vol. 2, Issue. 2 (April – June) , 2025 , Pages: 30 – 34. DOI : https://doi.org/10.63328/IJCSEAI-V2RI2P5

[14]. M. M. T, "Internet Enhanced Smart Energy Network Efficiency using Artificial Intelligence," International Journal of Research and Development in Engineering Sciences, vol. 8, no. 2, p. 1, Mar. 2026, doi: 10.63328/ijrdes-v8ri2p1.

[15]. S. B. Bokka, "Enhancing brain stroke prediction using machine learning for early intervention," International Journal of Research and Development in Engineering Sciences, vol. 7, no. 1, p. 46, Feb. 2025, doi: 10.63328/ijrdes-v7ri1p8.

[16]. S. R. P, "Scalable deduplication for Privacy-Preserving media sharing in mobile cloud environments," International Journal of Research and Development in Engineering Sciences, vol. 8, no. 2, p. 25, Mar. 2026, doi: 10.63328/ijrdes-v8ri2p4.

[17]. Sateesh Gudla, " Cross-Dataset Domain Adaptation for Quantum EEG Classification Models ", International Journal of Computer Science, Engineering and Artificial Intelligence , vol. 3, no. 1, p. 1-7, January 2026, DOI: https://doi.org/10.63328/IJCSEAI-V3RI1P1