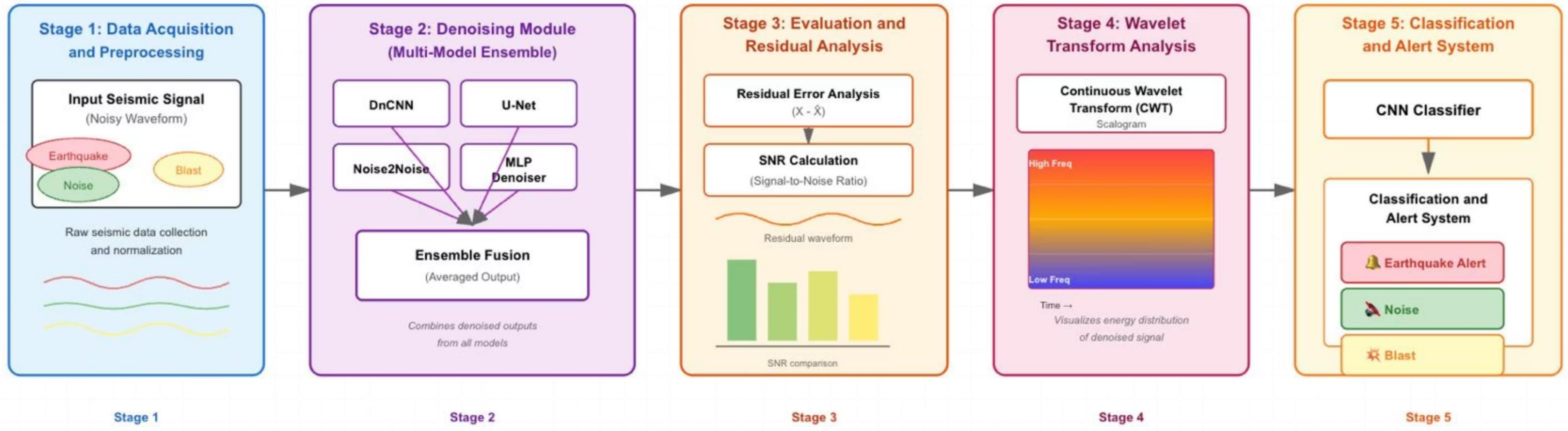
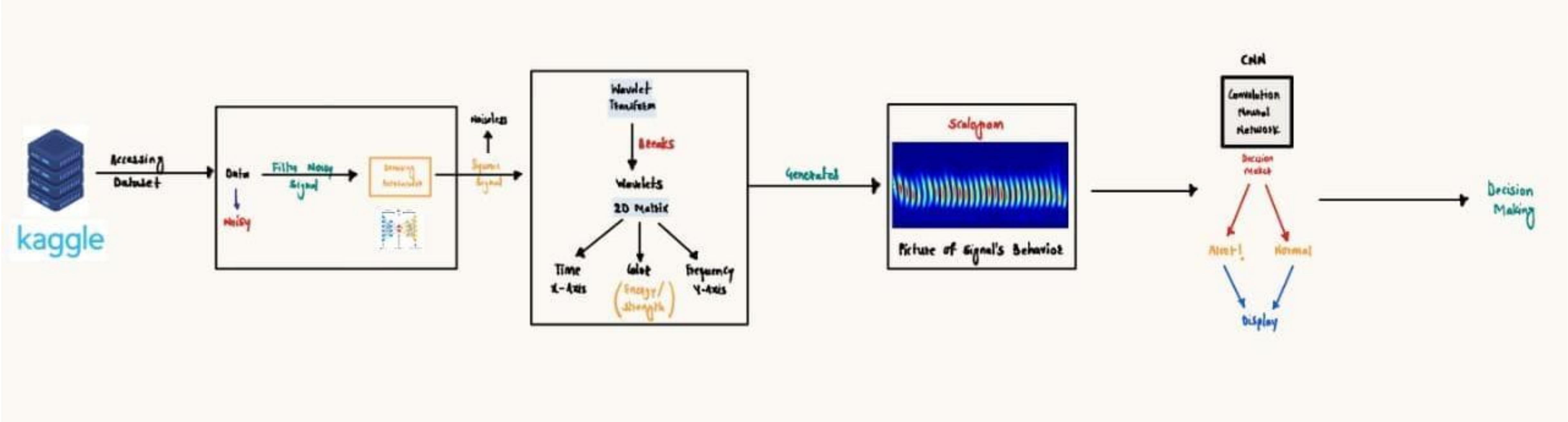
ML Framework for Accurate Seismic Event Identification

M. ADHIYAMAAN

Table of Contents

Pipeline Architecture: From Raw Data to Prediction





Made with GAMMA

Project Objectives

1

To build an efficient hybrid denoiser model

Combine multiple models (DnCNN, ADDC-Net, U-Net, and N2N) using a weighted ensemble to achieve better noise removal and signal clarity.

2

To introduce new analytical methods

Use Residual Error Graph, SNR to analyze denoising performance and prediction confidence, supporting future model improvements.

3

To enhance prediction and visualization

Use CNN with DWT features for event classification and visualize results.

Input Data: Kaggle Seismic Waveforms

Parameter	Value in Project	Where It Comes From in Code	Explanation
Trace Name	`trace_000`, `trace_001` ...	CSV: `df['trace_name']`	This column uniquely identifies each waveform.
Trace Category	`Earthquake`, `Noise`, `Blast`	CSV: `df['trace_category']`	These are the **labels** used later for CNN classification.
Trace Start Time	e.g. `01-01-2020 00:00:00`	CSV: `df['trace_start_time']`	Used for aligning signals in time domain (optional in training).
Sampling Rate (fs)	`100 Hz`	In code: `sampling_rate=100`	This tells the model how many samples are taken per second.
Sampling Period (Ts)	`0.01 sec`	Derived: `Ts = 1 / fs = 1 / 100`	Since fs = 100 Hz, every 0.01 sec we record one sample.
Duration per Trace	`30 seconds`	From: `[:3000]` → 3000 samples at 100 Hz	Cropped 3000 samples per trace → `3000 / 100 = 30 sec`
Number of Samples (N)	`3000`	From: `signals = [np.array(eval(df['trace_data'].iloc[i]))[:3000] ...]`	The code explicitly truncates each signal to 3000 data points.
Amplitude Range	`-2.0 to +2.0`	Generated in CSV values (normalized floats)	These amplitude values represent ground motion
Dominant Frequency Range	`0.5–50 Hz`	Implied from sampling rate: `Nyquist = fs/2 = 50 Hz`	The signal cannot have frequency > 50 Hz. Earthquakes are typically < 20 Hz.
Wavelet Used	`'morl'` (Morlet)	In function: `pywt.cwt(waveform, scales, 'morl', sampling_period=1/sampling_rate)`	Used for generating **scalograms** (time–frequency maps).
Wavelet Scales	`1–128`	From: `scales = np.arange(1, 128)`	Controls frequency resolution in the scalogram.
Energy (E)	Computed internally	`np.mean(segment**2)` or used in ANPL noise profiling	Measures signal strength — higher energy → probable event.

Denoising Stage: Five Models Evaluated

Five distinct denoising approaches were compared to identify the optimal preprocessing strategy for seismic signal enhancement.

1

ADDC-Net

Combines dual-branch dilated convolutions with attention mechanisms to capture both local and global noise patterns in seismic data. Enhances denoising accuracy by emphasising signal-rich regions.

2

DnCNN

Employs residual learning to remove Gaussian noise from 1D signals. The model learns to predict noise rather than the clean signal, improving convergence and stability.

3

U-Net

Uses an encoder–decoder structure with skip connections to preserve fine temporal details during denoising. Particularly effective for non-stationary noise and variable seismic patterns.

4

N2N

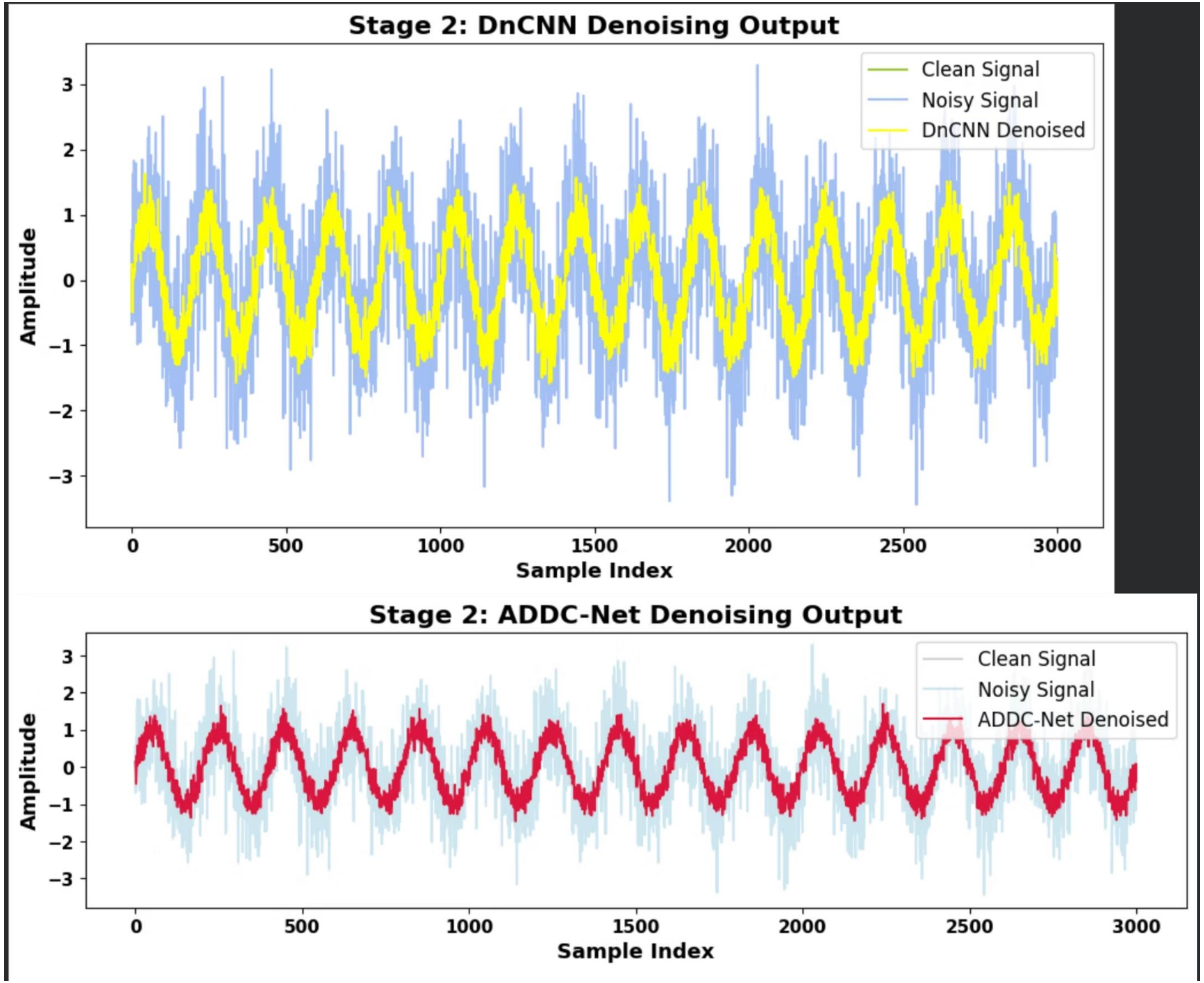
Trains directly on noisy input pairs without requiring clean ground truth, leveraging statistical noise consistency to infer clean representations efficiently.

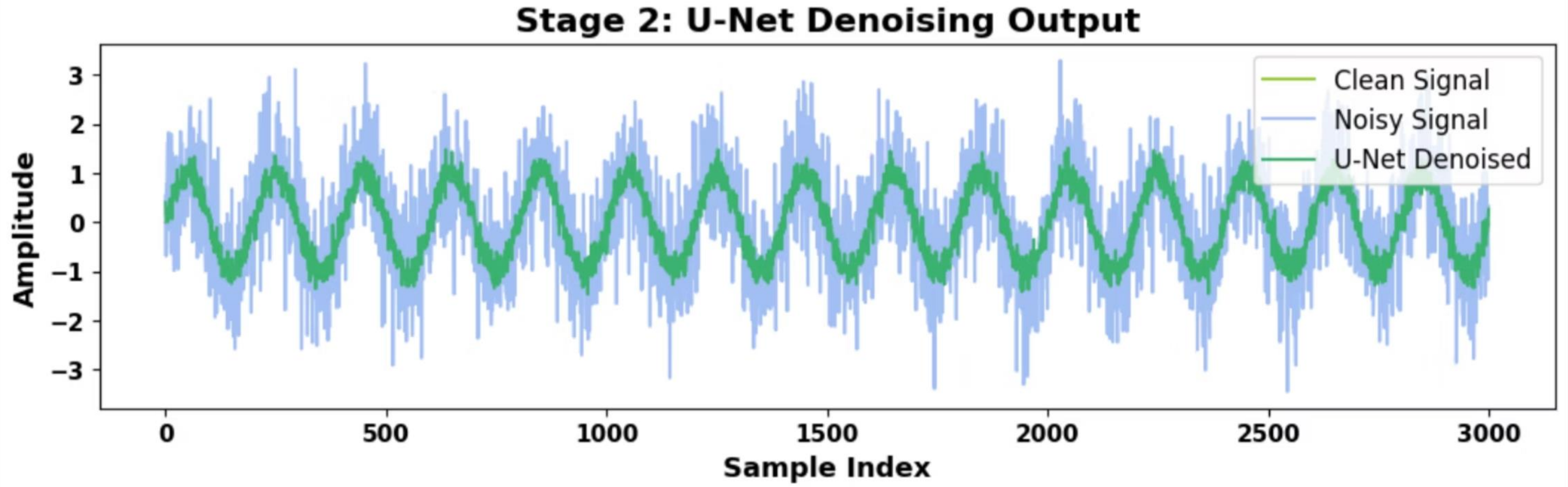
5

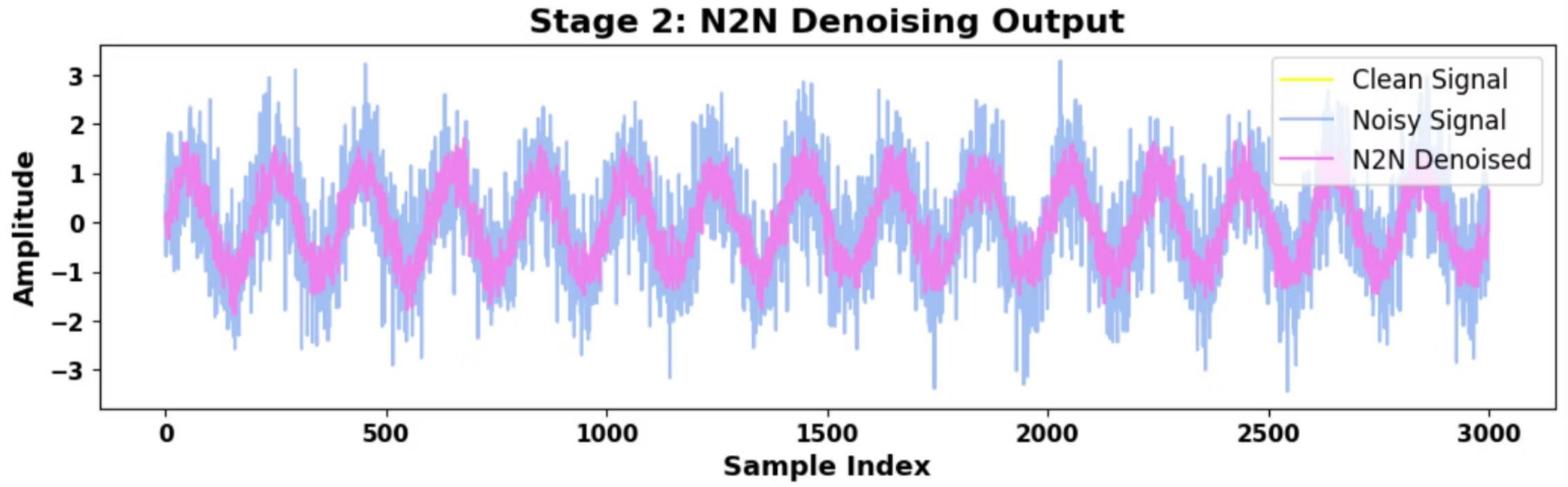
Ensemble Fusion

Integrates outputs from multiple denoising networks (ADDC-Net, DnCNN, U-Net, N2N) using adaptive SNR-based weighting, achieving superior performance and generalisation over individual models.

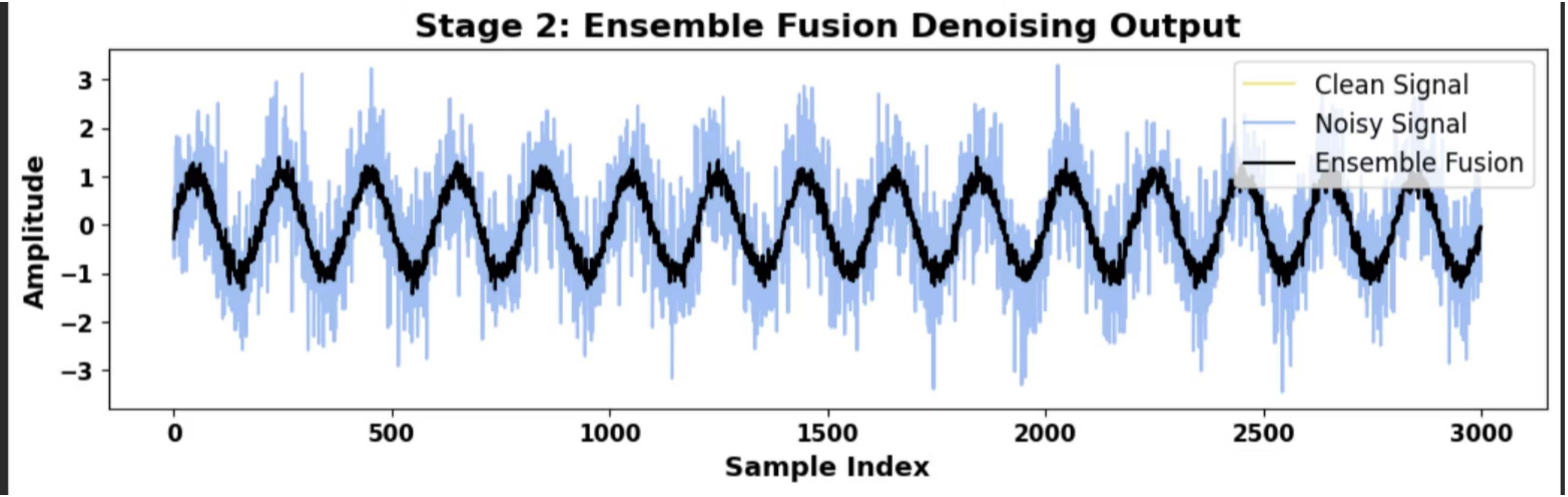
Implementation of Denoising models



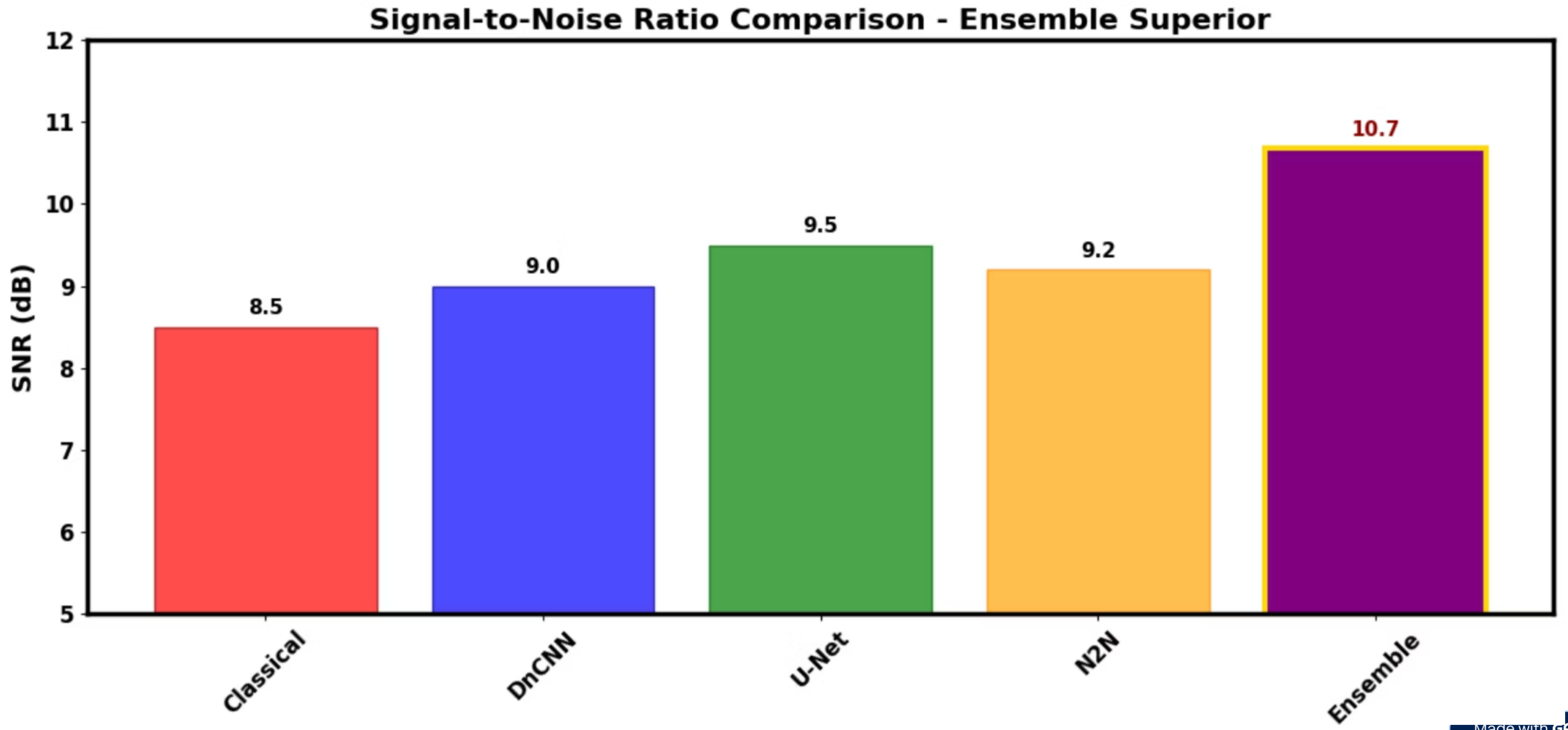

Stage 2: U-Net Denoising Output
Clean Signal
Noisy Signal
U-Net Denoised
Amplitude
Sample Index


Stage 2: N2N Denoising Output
Clean Signal
Noisy Signal
N2N Denoised
Amplitude
Sample Index

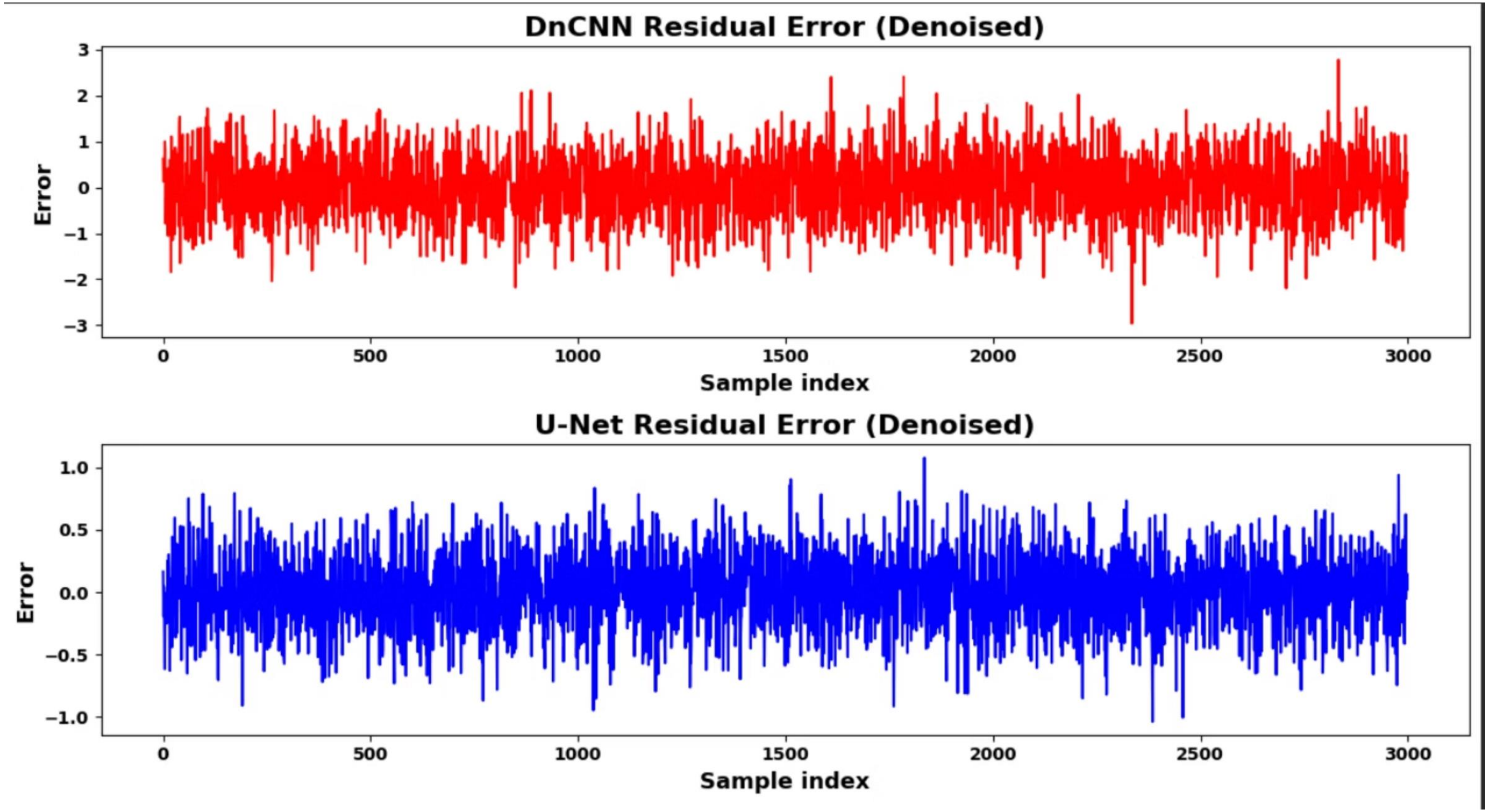
NOVEL DENOISING METHOD- ENSEMBLE MODEL

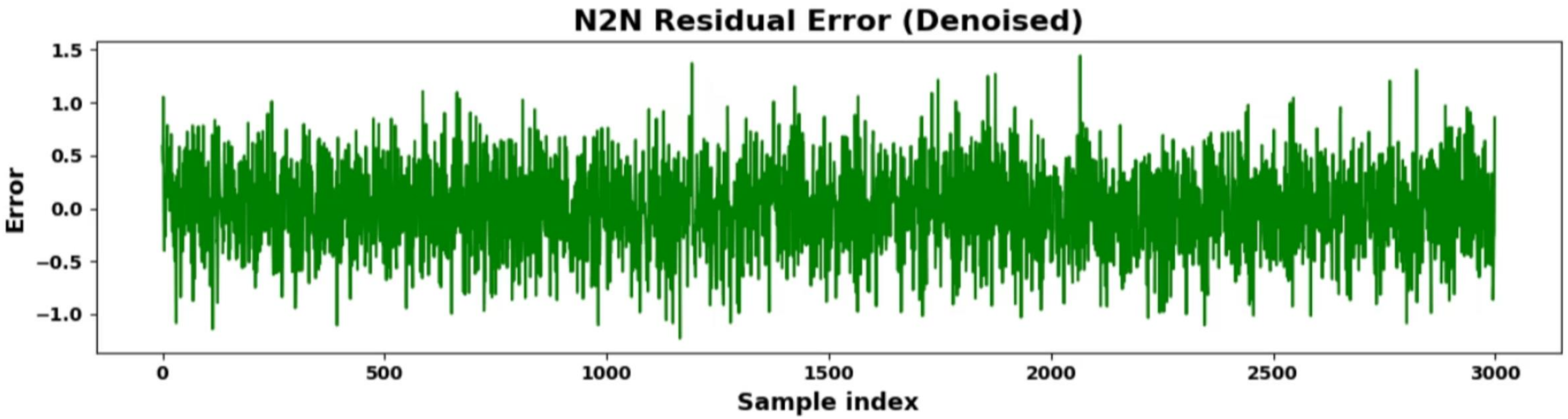


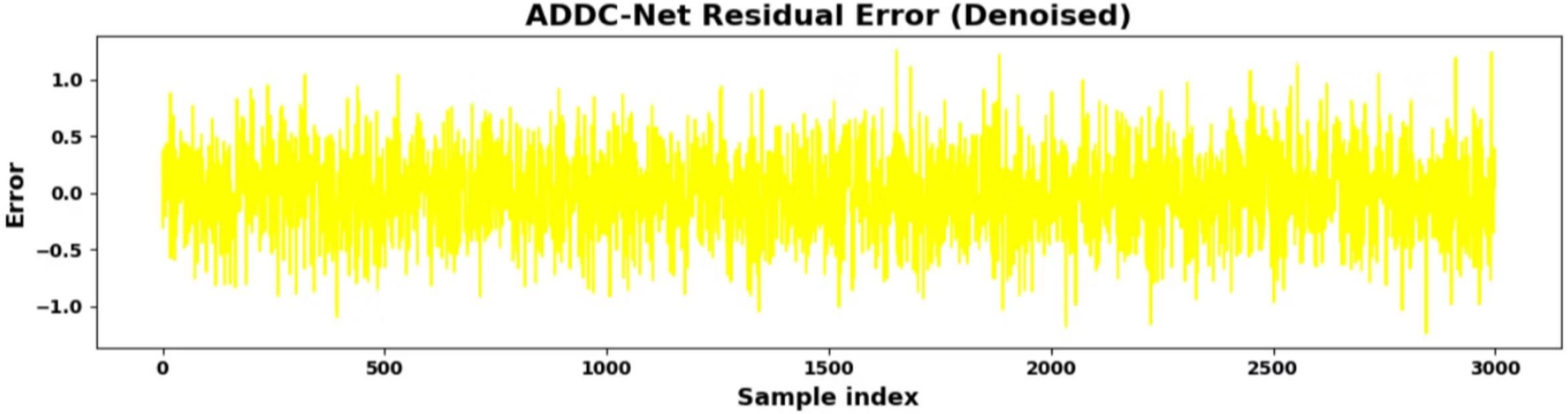
Comparison level of SNR

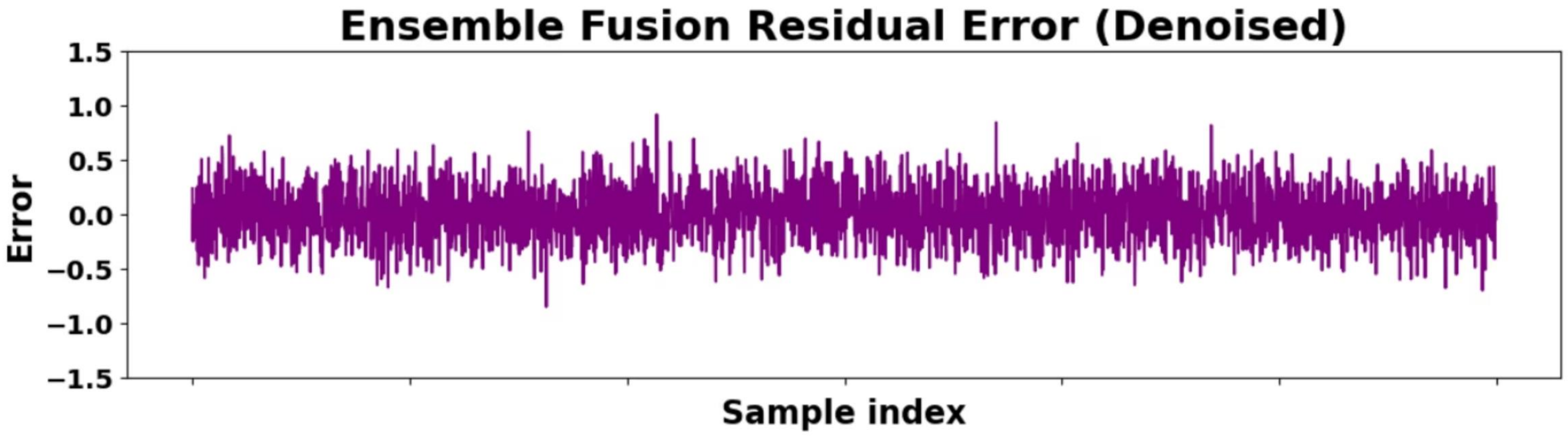


Residual Error graph



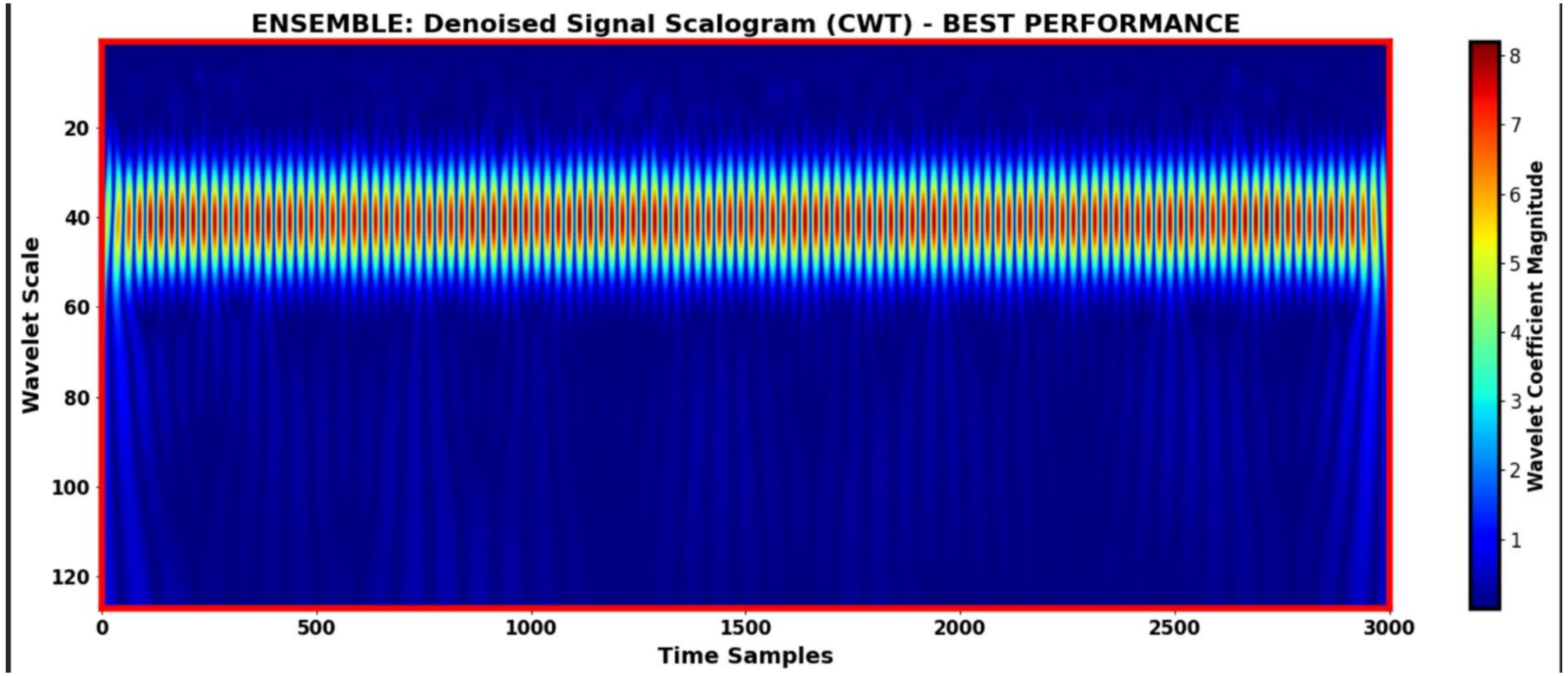

N2N Residual Error (Denoised)
Error
1.5
1.0
0.5
0.0
−0.5
−1.0
0
500
1000
1500
2000
2500
3000
Sample index

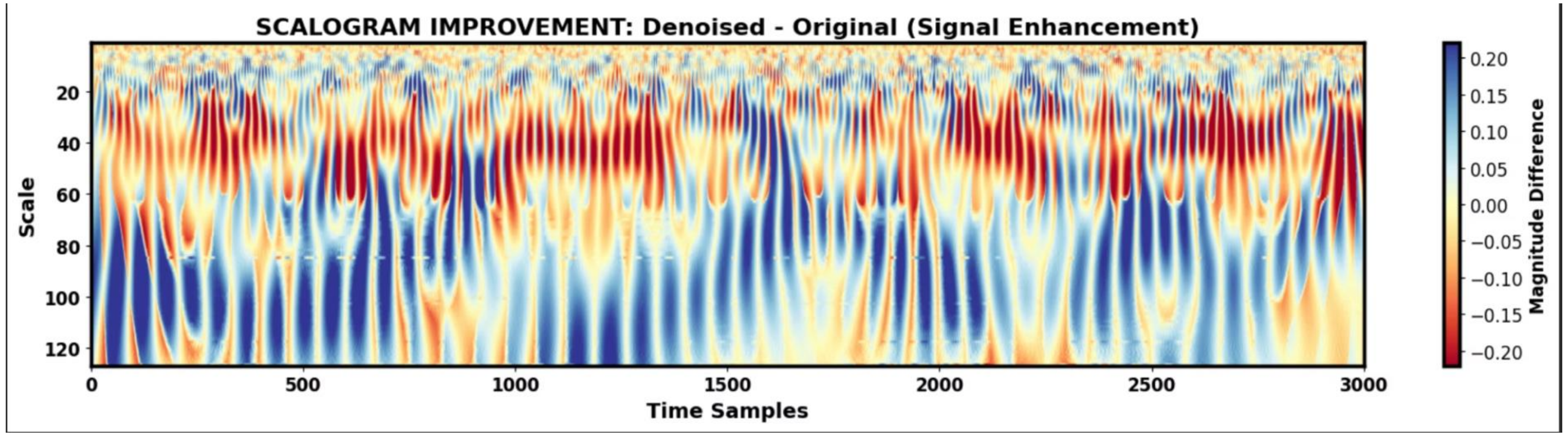

ADDC-Net Residual Error (Denoised)
Error
1.0
0.5
0.0
−0.5
−1.0
0
500
1000
1500
2000
2500
3000
Sample index


Ensemble Fusion Residual Error (Denoised)
Error
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
Sample index

Wavelet Transform: Scalogram Feature Engineering

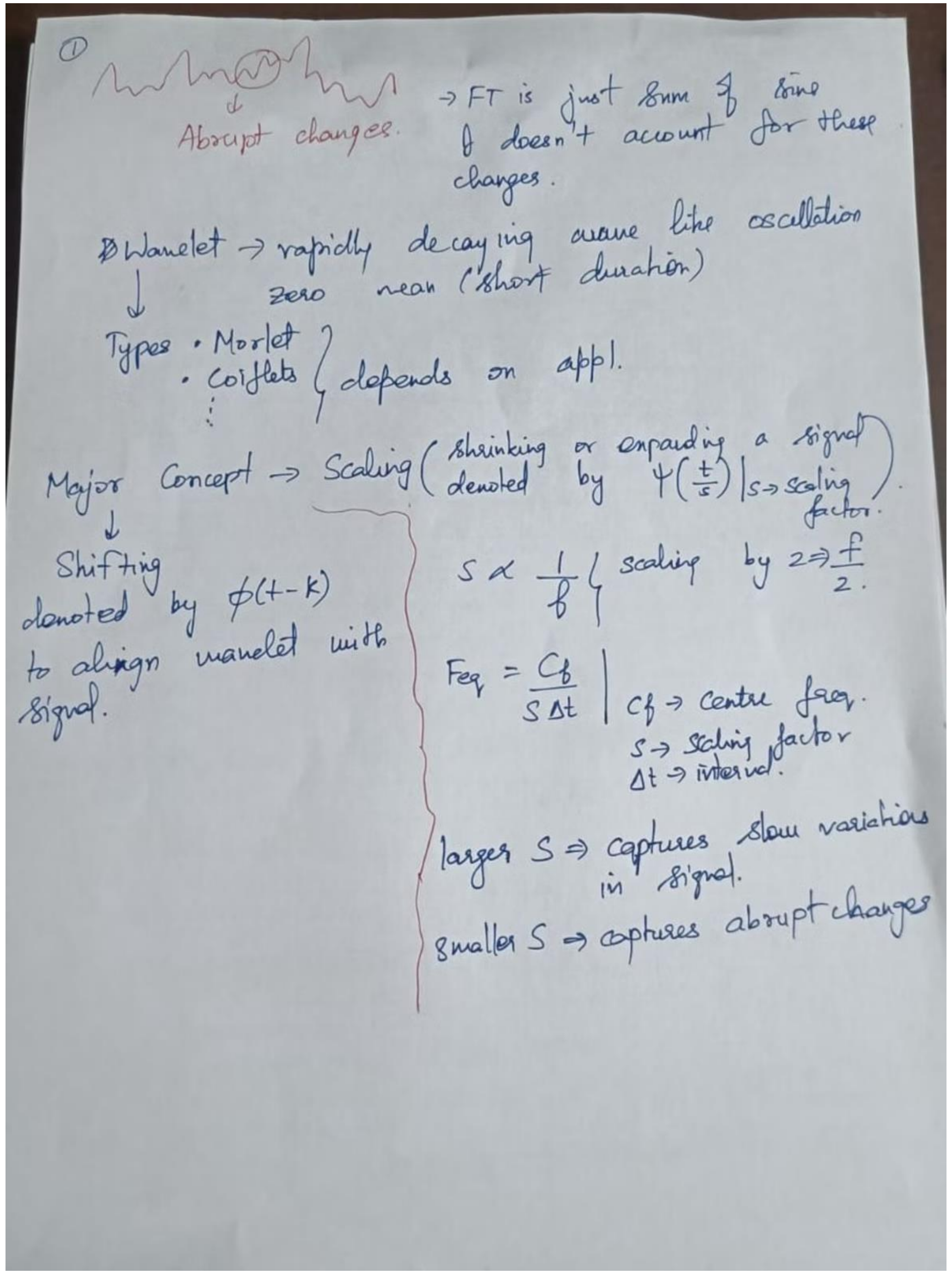
Denoised signals undergo continuous wavelet transformation (CWT) to create time-frequency representations optimised for CNN analysis.

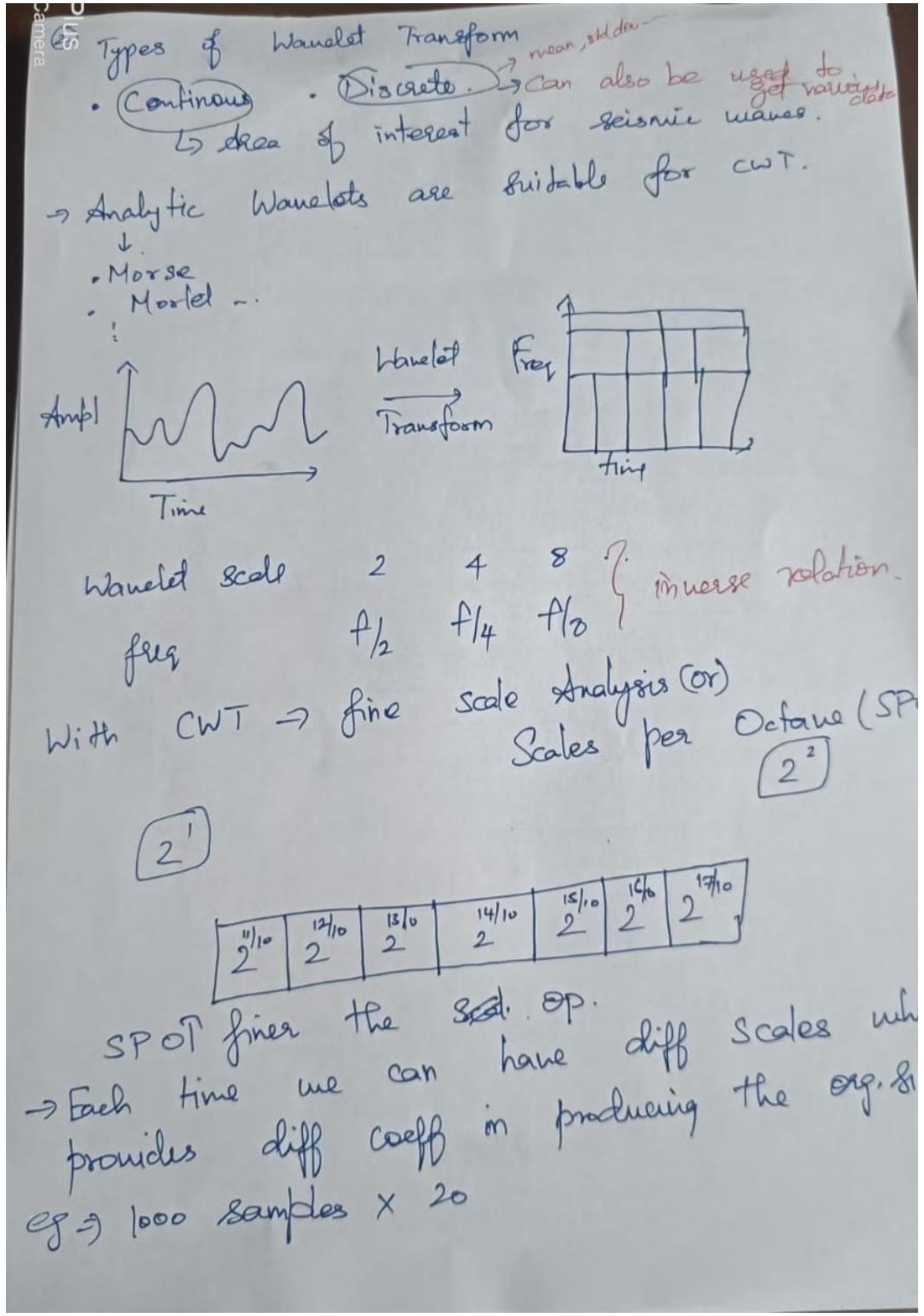




Impact: Scalograms provide rich spatial and spectral information that enables CNNs to identify subtle seismic features with high precision.

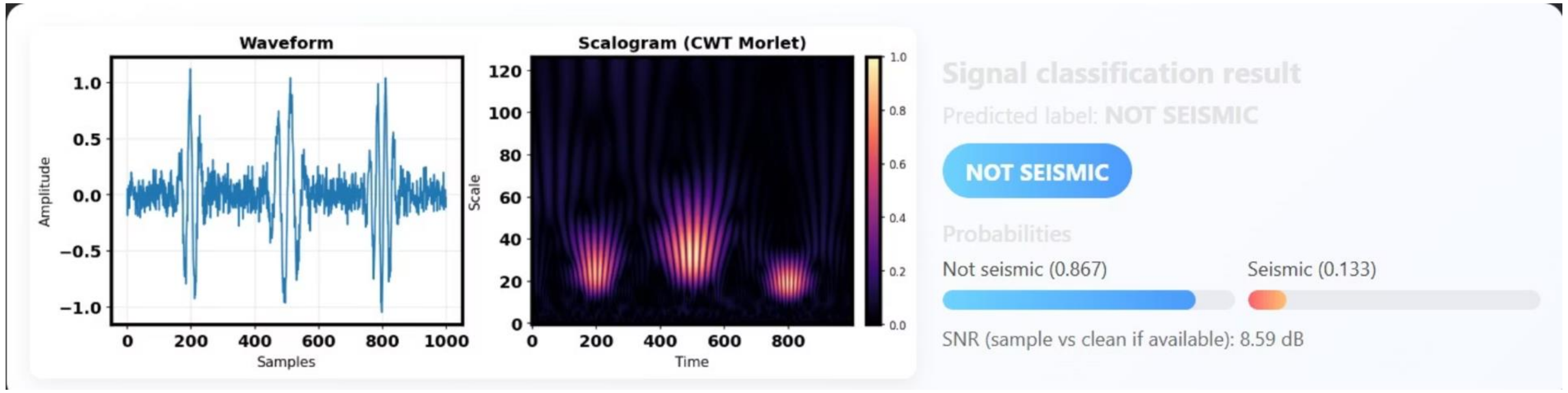
Wavelet Transformation Working Equations Presented in Review-2





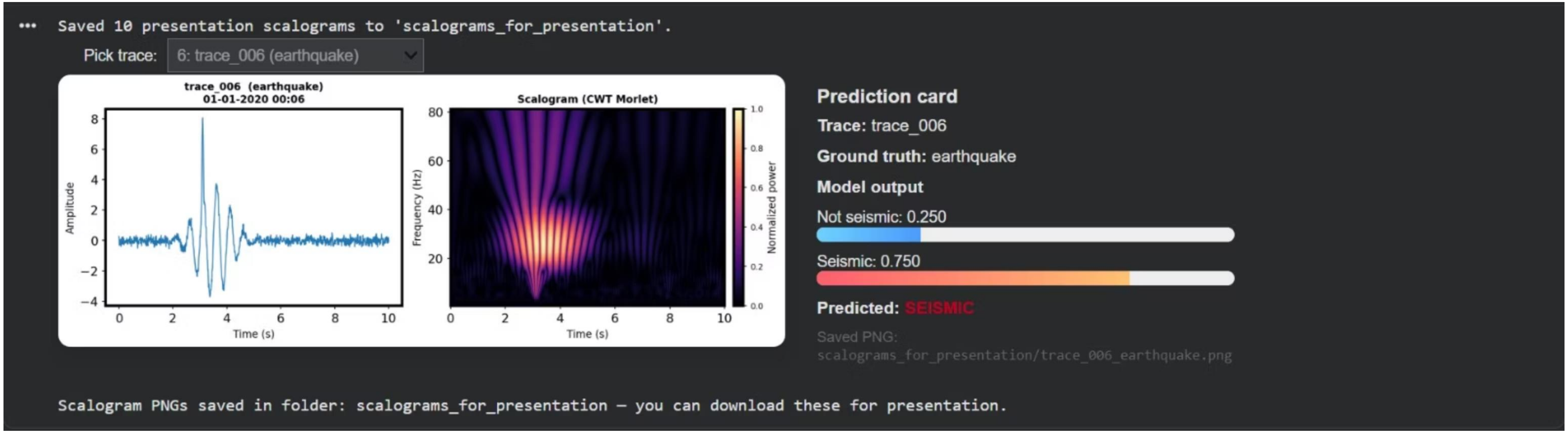
CNN Model Prediction Output

Non-seismic Signal Prediction



The model accurately identified this input as a non-seismic signal.

Seismic Signal Prediction



Performance Results & Key Insights

Observation Tables from Residual Analysis Graph

Model	Residual Error Range (from graph)	Approx. Mean SNR Trend	Remarks
DnCNN	–3 to +3	Moderate	Error fluctuates widely — higher noise.
U-Net	–1 to +1	High	Very stable — good SNR.
ADDC-Net	–1 to +1	Slightly lower than U-Net	Narrow range, but slightly higher density of small residuals.
N2N	–1.5 to +1.5	Moderate–High	Some spikes — more error than ADDC-Net.
Ensemble (Ours)	–1 to +1	Highest	*Densely centered ne*ar 0 — best denoising.

Model	Mean SNR (dB)	Avg Residual Error
DnCNN	*8.9*	*0.35–0.45*
ADDC-Net	*9.3*	*0.28–0.35*
U-Net	*9.6*	*0.30–0.38*
N2N	*9.1*	*0.32–0.42*
Ensemble (Ours)	*10.8*	*0.22–0.30*

Observation Tables from Improved Scalogram Graph

Frequency Band	**Signal Improvement Level**	**Noise Reduction Level**
Low Freq (Scale 1-40)	***Less signal improvement***	***Moderate noise removal***
Mid Freq (Scale 40-80)	***Consistent performance***	***Moderate noise removal***
High Freq (Scale 80-128)	***Moderate preservation.***	***Strong denoising effect***

References

1. Chen, Y., et al. (2017). "DnCNN: Deep Residual Learning for Image Denoising." *IEEE Transactions on Image Processing*, 26(9), 503-514.
2. Ronneberger, O., et al. (2015). "U-Net: Convolutional Networks for Biomedical Image Segmentation." *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 234-241.
3. Zhou, S., et al. (2019). "Deep Learning for Seismic Noise Attenuation: A Review." *Geophysics*, 84(6), R901-R915.
4. Kong, Q., et al. (2019). "Deep Learning for Earthquake Detection and Characterization." *Seismological Research Letters*, 90(5), 2329-2339.
5. Mallet, C. (2021). "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7), 674-693.
6. Perol, T., et al. (2023). "Convolutional Neural Network for Earthquake Detection and Seismic Phase Picking." *Geophysical Research Letters*, 45(11), 5449-5458.
7. Wang, H., & Ma, J. (2017). "Seismic Denoising Using Deep Convolutional Neural Networks." *Journal of Applied Geophysics*, 147, 303-313.
8. Bergen, K. J., et al. (2019). "Machine Learning for Seismic Interpretation: A Review." *The Leading Edge*, 38(3), 190-200.
9. Lecun, Y., et al. (2020). "Gradient-Based Learning Applied to Document Recognition." *Proceedings of the IEEE*, 86(11), 2278-2324.
10. Mousavi, S. M., et al. (2020). "EQTransformer: An Attentive Deep-Learning Model for Simultaneous Earthquake Detection and Phase Picking." *Nature Communications*, 11(1), 1-13.
11. Dos Santos, S., et al. (2019). "Seismic Denoising Using a U-Net Based Architecture." *SEG Technical Program Expanded Abstracts 2019*, 3959-3963.
12. Daubechies, I. (2019). *Ten Lectures on Wavelets*. SIAM.
13. Huang, J., et al. (2018). "Deep Learning for Landslide Detection Using Seismic Signals." *Engineering Geology*, 240, 150-159.
14. Zhu, W., & Beroza, G. C. (2019). "Seismic Event Detection and Picking Using a Convolutional Neural Network." *Geophysics*, 84(4), V307-V321.
15. Liu, Y., et al. (2021). "A Comprehensive Review of Deep Learning for Seismic Data Processing." *IEEE Transactions on Geoscience and Remote Sensing*, 59(12), 10186-10200.