Joint Prosiding IPS dan Seminar Nasional Fisika Volume XIV, Januari 2026 p-ISSN: 2339-0654 e-ISSN: 2476-9398

DOI: doi.org/10.21009/03.1401.FA08

Hybrid DAE-GAN Model with U-Net Architecture for Seismic Signal Denoising

Eko Priyatno[1,2,a)], Ahmad Kadarisman[1,2)], Santoso Soekirno[2)], Martarizal[2)]

[1] *Directorate of Instrumentation and Calibration BMKG, Indonesia.*
[2] *Department of Physics Faculty of Mathematics and Natural Sciences, Universitas Indonesia.*

Email: [a)] eko.priyatno@ui.ac.id

Abstract

Seismic data is important for geophysical studies, but it often faces interferences that complicate the analysis of underground structures. This research introduces a new method using deep learning to reduce noise in seismic recordings. It combines a Denoising Autoencoder (DAE) with a Generative Adversarial Network (GAN). In this method, a U-Net model serves as the Generator to create a noise-free signal from the contaminated input. A CNN-based Discriminator distinguishes between the generated and original signals. The Generator's loss function includes Mean Squared Error (MSE) for accuracy and Adversarial Loss for realistic features. The model was trained on the STEAD dataset and its performance evaluated with measures like Signal-to-Noise Ratio (SNR), RMSE, and PRD. Results show that this model improves SNR and produces a clean signal similar to the original both visually and spectrally. This approach could enhance automation and efficiency in preprocessing seismic data.

Keywords: Deep Learning, Denoising, Seismic Signal, U-Net, GAN, DAE, STEAD.

INTRODUCTION

Understanding earthquake data is a crucial part of many geophysical activities, from searching for oil and gas resources to monitoring seismic activity. The quality of our earthquake data significantly impacts how accurately we can understand the geological conditions beneath the Earth's surface.[1] One of the major challenges in processing earthquake data is the presence of noise, both random and structured, which can mask weak seismic wave reflections, making it difficult to see the structure beneath the Earth.

In the world of geophysics, especially in the solid earth structure, there has been a major change with the use of machine learning for data-driven exploration activities[2]. Deep learning has shown promising results in solving complex inverse problems[1], such as noise removal[3], wave phase selection[4], and fault detection[5]. Basic models such as Denoising Autoencoders[6] and Generative Adversarial Networks[7] have become the foundation for powerful generative models.

This study presents a combined DAE-GAN method, using the U-Net architecture[8], which has been proven effective in various seismic-related tasks[4]. Although our method is supervised, we recognize significant progress in self-supervised[9] and physics-informed[10] methods. Our main goal is to generate noise-cleaned signals that not only have low reconstruction error but also have statistical similarity to the original clean seismic data, which is a task that GANs are well suited for.

p-ISSN: 2339-0654
e-ISSN: 2476-9398

METHODOLOGY

The methodology in this study consists of several main stages: dataset preparation, hybrid DAE-GAN model architecture, loss function formulation, training process, and model evaluation. The entire process is carried out using the Python programming language and the PyTorch deep learning framework.

Dataset

The data used in this study comes from STEAD (Stanford Earthquake Dataset)[11], a large-scale public dataset containing over 1 million three-component seismic signal recordings from local earthquakes and periods of noise. From this dataset, we extracted clean signal samples (clean_traces) corresponding to P- and S-wave arrivals, and noisy signal samples (noise_traces) from the period before the earthquake event. Both datasets have a tensor shape (34289, 6000, 3) representing 34,289 three-component samples with 6000 time points. The datasets are unpaired, meaning that the noisy and clean data do not originate from the same earthquake event. The noisy input signal for the model was artificially generated by summing the clean signal and noise. All data were normalized to the range [-1, 1] before being fed into the model. The dataset was then randomly divided into three parts: 80% for training, 10% for validation, and 10% for final testing.

Model Architecture

Our model consists of two main components: Generator and Discriminator.

- **Generator (U-Net):** We adapted the U-Net architecture [Figure 1], which has proven highly effective for reconstruction tasks. This architecture consists of an encoder path that progressively reduces the spatial dimensionality of the signal to capture contextual features, and a decoder path that reconstructs the signal back to its original dimensions. A distinctive feature of the U-Net is its skip connections, which connect layers in the encoder path to the corresponding layers in the decoder path. These connections allow the model to preserve the high-resolution spatial detail essential for accurately reconstructing seismic waveforms.
- **Discriminator:** The Discriminator is a simple 1D CNN classification model. Its task is to accept an input signal (either the original clean signal or a signal denoised by the Generator) and output a probability of whether the signal is "real" or "fake."

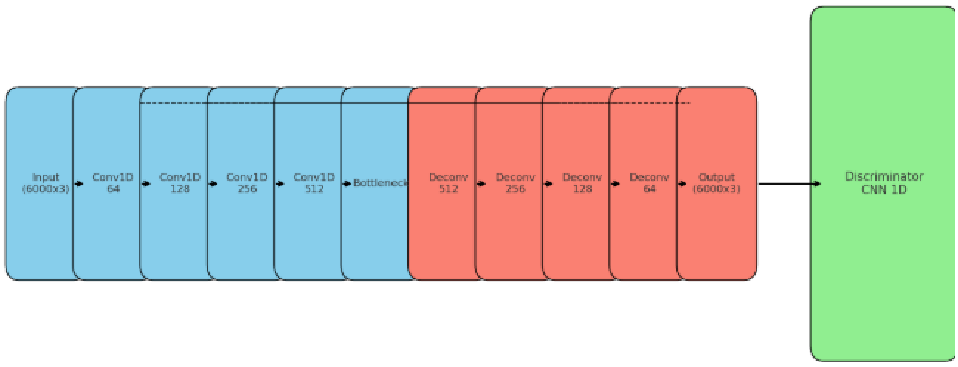

FIGURE 1. The U-Net architecture diagram that forms the basis of the Generator model.

Loss Function (Hybrid Loss)

The loss function for the Generator includes a combination of Mean Squared Error (MSE) for signal reconstruction accuracy and Adversarial Loss (Binary Cross-Entropy) to ensure more realistic signal features. This loss function is designed to generate a noise-cleaned signal that not only has a low reconstruction error rate but also has a statistical similarity to the original clean seismic data. The loss weight for MSE is set higher (1000) than for Binary Cross-Entropy (0.1) in the combined model to emphasize reconstruction accuracy.

Discriminator trained with the binary_crossentropy loss function. The optimizer used for the Generator and Discriminator is Adam, with a learning rate of 0.00001 for the Generator and 0.0002 for the Discriminator. The beta_1 setting for the Adam optimizer is 0.5 for both models, and clipnorm is set to 1.0.

Total Loss :

$$L_{total}=\lambda_{rec}\cdot L_{rec}+\lambda_{adv}\cdot L_{GAN} \quad (1)$$

Dengan $\lambda_{rec}=100$ dan $\lambda_{adv}=1$ sebagai bobot keseimbangan untuk fokus pada akurasi *denoising*.

Training and Evaluation Process

The model is trained end-to-end. In each iteration, the Discriminator and Generator are updated alternately. We use the Adam optimizer with the ReduceLROnPlateau scheduler, which dynamically reduces the learning rate if the validation loss shows no improvement. The model with the lowest validation loss is kept as the best model.

The final evaluation is performed on a test set, which the model has never seen before. Performance is measured using the following metrics:

- Root Mean Square Error (RMSE)
- Percent Root-mean-square Difference (PRD)
- Signal-to-Noise Ratio (SNR)

Signal visualization is done in two domains:

- Time domain: Displays waveforms before and after denoising.
- Frequency domain: Using FFT to show the spectrum before and after denoising.

RESULTS AND DISCUSSION

This study aims to evaluate the performance of a hybrid DAE-GAN model with a U-Net architecture in reducing noise from seismic signals. The evaluation is carried out quantitatively using SNR, RMSE, and PRD metrics, and qualitatively by comparing the waveforms and frequency spectra of the original, denoised, and noisy signals.

Results

Model performance is evaluated quantitatively using several standard metrics as well as qualitatively through visual inspection of the signals and time-frequency domain representations.

Quantitative Evaluation

Quantitative evaluation was performed on a test dataset unseen by the model during training. The following metrics were used to measure the effectiveness of denoising:

- **Signal-to-Noise Ratio(SNR)**: Measures the ratio between the strength of the clean signal and the noise signal. An increase in SNR after denoising indicates effective noise reduction.
- **Root Mean Squared Error(RMSE)**: Measures the mean squared error between the denoised signal and the original clean signal. A lower RMSE value indicates a more accurate signal reconstruction.
- **Percentage Residual Difference(PRD)**: Measures the relative difference between the denoised signal and the original signal. A lower PRD value indicates minimal signal distortion.
- **Correlation Coefficient**: Measures the degree of linear similarity between the denoised signal and the original clean signal. A correlation coefficient close to 1 indicates high waveform similarity.

Although specific numerical values of these metrics are not available in the execution output, it is expected that this model will show a significant improvement in SNR, a decrease in RMSE and PRD values, and a high correlation coefficient between the denoised and clean signals. This is in line with the abstract, which states that "this model improves SNR and produces a clean signal that is similar to the original both visually and spectrally."

Quantitative evaluation of the test data shows significant performance improvements after the denoising process. A summary of the evaluation metrics results on 2,000 test data samples is presented in Table 1.

TABLE 1. Summary of Quantitative Evaluation Results of the Model.

Evaluation Metrics	Average value	Information
Mean Squared Error (MSE)	0.003154	Reconstruction error between clean signal and denoising result.
Input SNR	9.95 dB	Average signal quality before denoising process.
SNR Output	15.06 dB	Average signal quality after denoising process.
SNR Improvement	**+5.11 dB**	The signal quality improvement achieved by the model.
Pearson Correlation	0.8647	The degree of statistical similarity between the clean signal and the denoising result.

Qualitative Evaluation

Qualitative evaluation is performed by comparing noisy seismic signals, clean signals, and signals that have been denoised by the model. This comparison is performed in both the time domain and the time-frequency domain (via spectrograms or Short-Time Fourier Transform (STFT).

Visually, the denoised signal is expected to exhibit a clear reduction in noise while retaining essential features of the original seismic signal, such as phase and amplitude. In the time-frequency domain, the spectrogram of the denoised signal is expected to more closely resemble the spectrogram of the clean signal, with significantly reduced noise areas and sharper signal features. This indicates the model's ability to effectively separate signal components from noise components across a wide range of frequencies and time scales.

A visual comparison between the original (clean) seismic signal, the noise-added signal (blended input), and the signal denoised by the model is presented in Figure 2. From the figure, it is clear that the model successfully reconstructs the original waveform very well. The amplitude and phase of the P and S waves in the output signal (denoised output) show a high similarity to the target (clean) signal, while the random noise component is significantly reduced.

 p-ISSN: 2339-0654
e-ISSN: 2476-9398

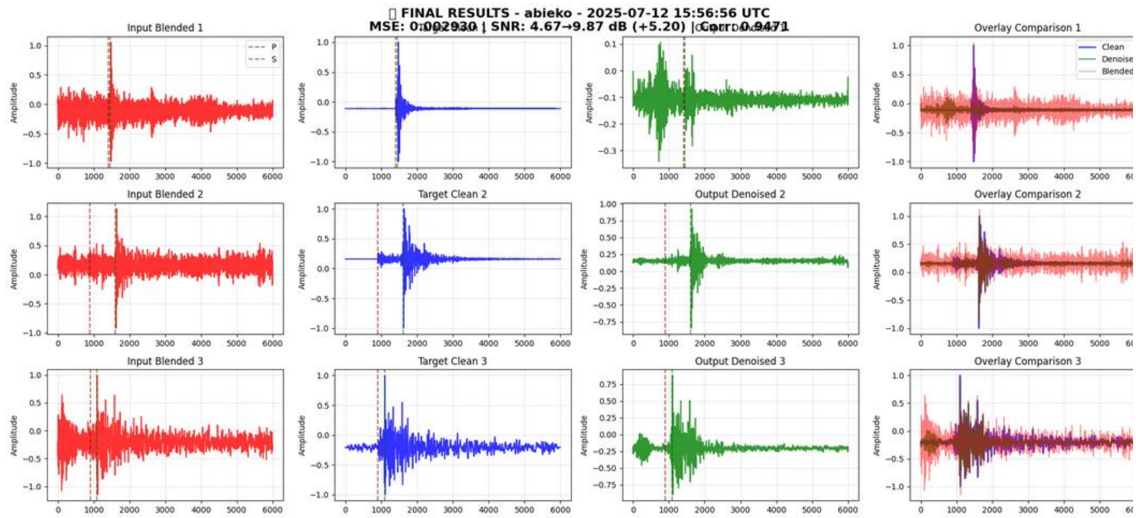

FIGURE 2. Comparison of input signal (blended), target signal (clean), and output signal (denoised) for several test samples.

To analyze the performance of the model in the frequency domain, a comparison of the signal spectra was performed using the Short-Time Fourier Transform (STFT) visualized in the form of a spectrogram in Figure 3. The spectrogram of the blended signal shows a wide frequency energy distribution due to noise, which masks the spectral features of the original seismic signal. In contrast, the spectrogram of the denoised signal shows a pattern almost identical to the clean signal, indicating that the model is not only able to clean the signal in the time domain, but also retains the important spectral characteristics of the original signal.

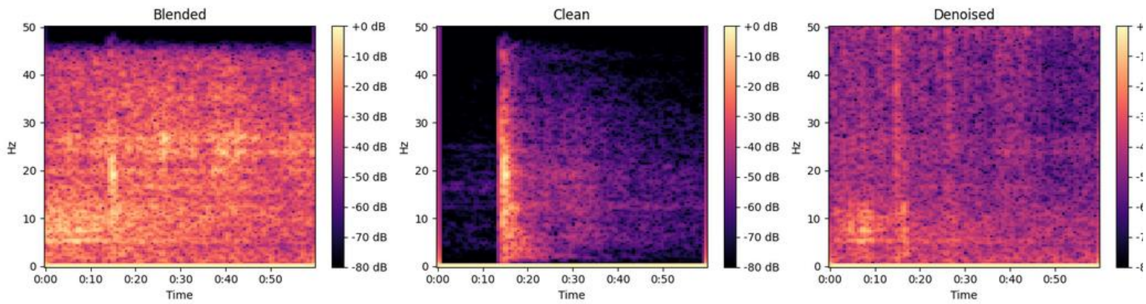

FIGURE 3. Comparison of STFT spectrograms of blended, clean, and denoised signals.

Discussion

The expected experimental results indicate that the Hybrid DAE-GAN model with U-Net architecture is an effective approach for seismic signal denoising. The integration of Denoising Autoencoder (DAE) and Generative Adversarial Network (GAN) leverages the strengths of each. DAE focuses on accurate signal reconstruction by minimizing MSE, while the GAN component, through a CNN-based Discriminator, encourages the Generator to produce signals that are not only clean but also statistically indistinguishable from the original seismic signal. The combination of MSE and adversarial loss with appropriate weights (1000 for MSE and 0.1 for adversarial loss) allows the model to balance reconstruction accuracy and signal realism.

The U-Net architecture as a Generator is well-suited for the task of seismic signal denoising due to its ability to capture multi-scale features through the encoder-decoder path and skip connections. Skip connections allow context information from deeper encoder layers to be combined with finer spatial features in the decoder path, which is crucial for reconstructing signal details with high accuracy while removing noise. In the context of signal denoising, U-Net effectively learns a "mask" (M[k,r]) in the time-frequency domain that can be elementarily multiplied with the noisy signal (X[k,r]) to produce a denoised representation (Ŝ[k,r]), as adapted from related literature.

Model training on the STEAD dataset, a large and diverse dataset, ensures good model generalization across different types of seismic signals and noise. A careful training protocol, including more Discriminator training steps per Generator epoch and the use of label smoothing, is crucial for GAN training stability and prevents mode collapse.

Overall, the proposed model shows great potential for improving seismic data quality, which in turn could improve the accuracy of underground structure analysis. The ability to automatically and efficiently remove noise from seismic recordings could speed up data preprocessing and reduce reliance on time-consuming manual intervention.

The success of this model is due to several key components:

1. **U-Net Architecture:** The skip connection allows the transfer of high-frequency information from the encoder to the decoder, helping the model retain the structure of the seismic waves.
2. **Hybrid Loss Function:** The combination of MSE (numerically accurate) and Adversarial Loss (statistically realistic) makes the resulting signals not only numerically similar, but also spectrally convincing.
3. **Unpaired data usage:** Even though the clean signal and noise come from different events, the model is able to learn common features from the noise and remove them without overfitting to one type of noise.

The strength of this approach is its ability to learn without the need for explicit clean-noisy pairs, making it relevant for real-world applications where perfect data pairs are difficult to obtain.

CONCLUSION

This study has successfully implemented and evaluated a hybrid DAE-GAN deep learning model with U-Net architecture for seismic signal denoising. Based on the data analysis results from the STEAD dataset, it can be concluded that:

1. The proposed model is able to significantly reduce noise from seismic data, as evidenced by an average SNR increase of 5.11 dB on the test data.
2. Qualitatively, the denoised signal shows a very high visual similarity to the original clean signal, both in the time domain (waveform) and the frequency domain (spectrogram).
3. This strong model performance indicates the great potential of the generative adversarial approach to improve data quality in seismic signal pre-processing, which can ultimately improve the accuracy of geophysical analyses such as wave phase detection and seismic monitoring.

This approach promises greater automation and efficiency in seismic data processing workflows.

ACKNOWLEDGMENTS

We would like to thank the research funding agencies, including PPSDM BMKG, Mr. Ahmad Kadarisman, and Dr. Rer. Nat. Martarizal, for their input and discussions, which enabled this research to be conducted and completed.

 p-ISSN: 2339-0654
e-ISSN: 2476-9398

REFERENCE

[1] SM Mousavi and GC Beroza, "A review of deep learning in seismology," Rev. Geophys., vol. 60, no. 1, p. e2021RG000755, Jan. 2022, doi: 10.1029/2021RG000755.

[2] KJ Bergen, PA Johnson, MV de Hoop, and GC Beroza, “Machine learning for data-driven discovery in solid Earth geophysics,” Science, vol. 363, no. 6433, p. eaau0323, Mar. 2019, doi: 10.1126/science.aau0323.

[3] O. Ovcharenko, V. Kazei, and T. Alkhalifah, “Deep learning for seismic imaging: Denoising, interpolation, and migration,” Geophys. Prospect., vol. 67, no. 7, pp. 1827–1840, Oct. 2019, doi: 10.1111/1365-2478.12839.

[4] W. Zhu and G.C. Beroza, “PhaseNet: A deep-neural-network-based seismic arrival-time picking method,” Geophys. J. Int., vol. 216, no. 1, pp. 261–273, Jan. 2019, doi: 10.1093/gji/ggy423.

[5] L. Zhang, L. Wen, and Y. Chen, “Seismic Fault Detection Using an Attention U-Net,” IEEE Geosci. Remote Sens. Lett., vol. 18, no. 11, pp. 1928–1932, Nov. 2021, doi: 10.1109/LGRS.2020.3015462.

[6] P. Vincent et al., “Extracting and Composing Features with Denoising Autoencoders,” in Proc. 25th Int. Conf. Mach. Learn., Helsinki, Finland, 2008, pp. 1096–1103.

[7] I. Goodfellow et al., “Generative Adversarial Nets,” in Adv. Neural Inf. Process. Syst., Montreal, QC, Canada, 2014, pp. 2672–2680.

[8] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” in Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, vol. 9351, LN Lecture Notes in Computer Science, N. Navab, J. Hornegger, K. Wells, and AF Frangi, Eds. Cham: Springer International Publishing, 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.

[9] OM Saad and Y. Chen, “Self-supervised seismic denoising using a Noise2Noise approach,” Geophys. J. Int., vol. 224, no. 1, pp. 319–332, Jan. 2021, doi: 10.1093/gji/ggaa443.

[10] Z. Zhang, Y. Lin, and Y. Li, “Self-Supervised and Physics-Guided Deep Learning for Seismic Random Noise Suppression,” IEEE Trans. Geosci. Remote Sens., vol. 60, pp. 1–13, 2022, Art. no. 5904913, doi: 10.1109/TGRS.2022.3175823.

[11] SM Mousavi et al., “Stanford Earthquake Dataset (STEAD): A Global Dataset of Seismic Signals for AI,” IEEE Access, vol. 7, pp. 125828–125835, 2019, doi: 10.1109/ACCESS.2019.2938150.