Advanced Data Augmentation Strategies for EEG-Based Alzheimer's Disease Classification Using Deep Learning

Achref OTHMANI, Ines BOUZOUITA, Zayneb BRARI and Safya BELGHITH
RISC Laboratory
École Nationale d'Ingénieurs de Tunis
Université de Tunis El Manar
achref.othmani@etudiant-enit.utm.tn, ines.bouzouita@enit.utm.tn,
zayneb.brari@enit.utm.tn, safya.belghith@enit.utm.tn

***Abstract*—This paper discusses the effect of state-of-the-art data augmentation approaches on electroencephalogram (EEG)-based Alzheimer's Disease classification using time-frequency scalogram features computed using the Continuous Wavelet Transform (CWT). To address the problems of limited data and class imbalance, standard image-based data augmentation techniques are compared with artificially generated EEG data using the Variational Autoencoder (VAE) [1]. According to the experimental results, VAE-based data augmentation can increase the robustness and generalize the framework's performance. Specifically, the ResNet-18 model [2] achieves an accuracy of 97.7% at the patient level, a precision of 100%, and a recall of 91.4%. Additionally, the effectiveness of attention-based models for EEG scalogram classification is confirmed by a Vision Transformer (ViT) [3], which achieves an accuracy of 93.5%, a recall of 100%, and an F1-score of 96.8%.**

I. Introduction

Alzheimer's Disease (AD) is a neurodegenerative disorder causing progressive cognitive decline, for which early diagnosis is critical to improving patient outcomes. Among neuroimaging modalities, Electroencephalography (EEG) offers a non-invasive, low-cost alternative with high temporal resolution [4], [5]. Time–frequency representations such as Continuous Wavelet Transform (CWT)-based scalograms have enabled the application of image-based deep learning models to EEG analysis [6], [7]. A key challenge in this domain is the limited size and class imbalance of available AD EEG datasets. Classical image augmentation (rotations, flips, scaling) increases sample count but does not introduce new physiological variability, potentially reducing generalisation [8]. Although their application to CWT scalograms for AD classification is still largely unexplored, generative models such as Variational Autoencoders (VAEs) [1] provide a principled alternative by learning the underlying signal distribution [9]. In order to compare VAE-based generative augmentation with standard augmentation across CNN, ResNet-18 [2], and Vision Transformer (ViT) [3] architectures, this work examines the effect of VAE-based generative augmentation on AD classification using EEG scalograms.

II. Preliminaries and Related Work

Manually extracted spectral, entropy, and nonlinear features combined with conventional classifiers were used in early EEG-based AD diagnosis [4], [10]. Although informative, these techniques do not scale to the variety found in real-world datasets and necessitate considerable domain expertise. The shift toward CWT-based scalograms and spectrograms [5], [7] enabled the use of convolutional architectures [6] and transfer learning. A representative example is the work by Lekkan et al. [11], which combines multiple time–frequency transforms with CNN classifiers. However, such deterministic pipelines are highly sensitive to dataset size and diversity, as they do not explicitly model the distribution of EEG patterns. Most studies use pixel-level augmentation (rotations, flips, intensity shifts) [8] to deal with small datasets, but this approach fails to capture inter-subject physiological variability. Although generative techniques, such as GANs and VAEs [1], have been used for EEG synthesis in motor imagery and emotion recognition tasks [9], [12], [13], their comprehensive evaluation on CWT scalograms for AD classification across diverse architectures is still lacking in the literature. Several studies also report near-perfect accuracy on small cohorts [14], raising concerns about overfitting that further motivate a rigorous comparative evaluation.

III. Proposed Approach

The proposed approach aims to improve EEG-based Alzheimer's disease (AD) classification by addressing data scarcity and class imbalance through augmentation applied to CWT-based EEG scalograms. Operating in the time–frequency domain preserves both the temporal and spectral characteristics of non-stationary EEG signals.

The overall pipeline consists of six main stages, as illustrated in Fig. 1:

1) EEG acquisition from the Florida State University dataset,
2) signal preprocessing,
3) CWT-based time-frequency representation,

4) data augmentation via classical image transformations or VAE-based generation [1],
5) classification using CNN, ResNet-18 [2], or ViT [3],
6) performance evaluation using accuracy, F1-score, and recall.

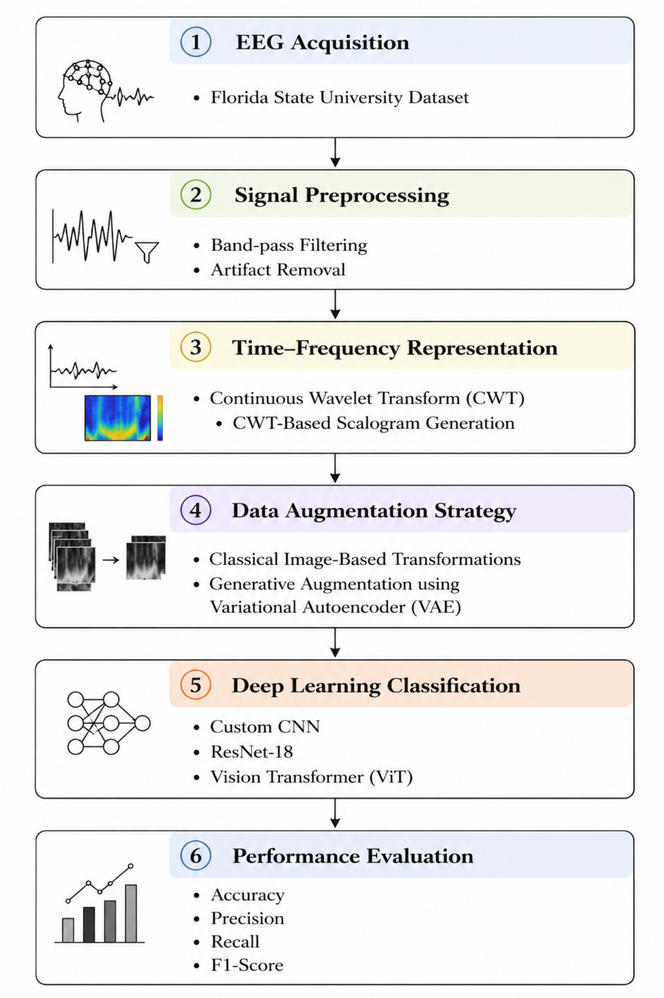


Fig. 1: Block diagram of the proposed EEG-based AD classification framework.

A. Description of the EEG Dataset

The dataset contains EEG recordings collected from elderly participants, including both healthy controls and patients diagnosed with probable Alzheimer's Disease (AD). Recordings are divided into four subgroups based on health condition and eye state (eyes-open / eyes-closed), as shown in Fig. 2.

The dataset comprises 184 subjects in total: 24 healthy elderly individuals and 160 patients with probable AD, equally distributed across the two eye-state conditions.

B. EEG Data Preprocessing

The experiments are conducted on EEG recordings from the Florida dataset. All signals are acquired at a sampling frequency of 128 Hz using 21 scalp electrodes placed according to the international 10–20 system, ensuring standardized spatial coverage of cortical activity. Standard preprocessing steps are applied, including artifact removal, band-pass filtering (0.5–45 Hz), amplitude normalization, and temporal segmentation into fixed-length epochs.

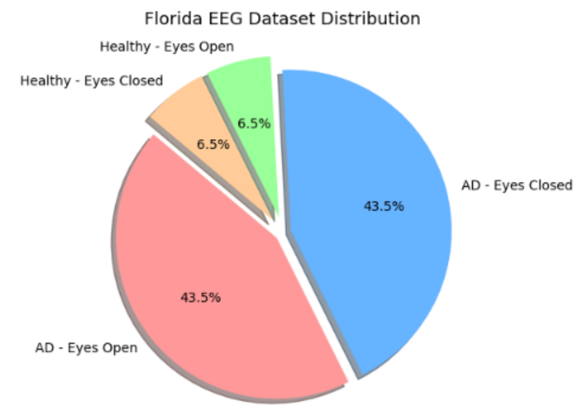


Fig. 2: EEG dataset composition.

C. Continuous Wavelet Transform for EEG Representation

To capture the non-stationary dynamics of EEG signals, a multiresolution time–frequency representation based on the Continuous Wavelet Transform (CWT) is adopted. Unlike the Fourier Transform, which assumes signal stationarity [15], the CWT provides a joint time–frequency decomposition that resolves transient oscillatory patterns and short-lived frequency bursts.

The complex Morlet wavelet is selected for its favorable balance between temporal and frequency resolution, making it well suited for tracking rapid neural dynamics [16]. The relationship between wavelet scale a and pseudo-frequency f is given by:

$$f \approx \frac{f_c}{a\,\Delta t} \tag{1}$$

where f_c is the central frequency of the Morlet wavelet and $\Delta t = 1/f_s$ is the sampling period. Scales are chosen to cover the principal EEG bands: delta, theta, alpha, beta, and low gamma.

D. Data Augmentation Strategies

1) Classical Image-Based Augmentation: To reduce overfitting caused by limited dataset size, conventional image-based augmentation is applied to the EEG scalograms, including random rotations, horizontal and vertical flips, and intensity variations [8]. Although these transformations increase the number of training samples and improve model stability, they operate at the pixel level and introduce limited physiological variability, potentially failing to capture realistic inter-subject EEG differences.

2) VAE-Based Generative Augmentation: A Variational Autoencoder (VAE) [1] is employed to generate synthetic scalograms that reflect the underlying distribution of EEG time–frequency patterns. The VAE is trained on CWT-based scalograms and, once converged, produces new samples primarily for the under-represented healthy class. This approach introduces realistic inter-subject variability while preserving the global spectro-temporal structure of EEG signals, unlike pixel-level transformations [9]. A total of 23 synthetic subjects were generated, each represented by 21 electrode-wise CWT scalograms.

The VAE training stability was verified by monitoring the loss evolution across electrodes. Fig. 3 shows the loss curve for electrode 21, which decreases monotonically without abrupt oscillations or divergence, confirming stable convergence and the absence of mode collapse.

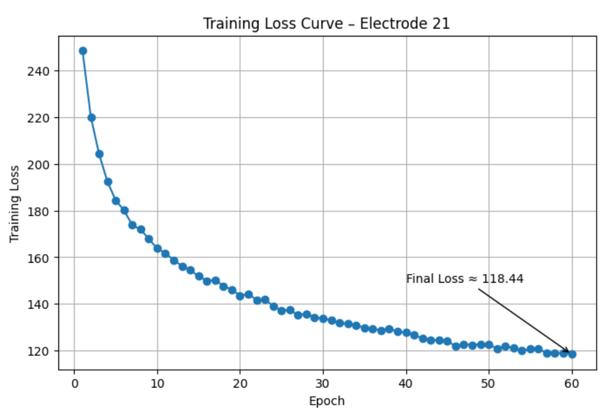


Fig. 3: VAE training loss for electrode 21, showing stable convergence for EEG scalogram reconstruction.

The total VAE loss combines a reconstruction term and a regularization term:

$$\mathcal{L} = \mathcal{L}_{\text{rec}} + \beta\,\mathcal{L}_{\text{KL}} \tag{2}$$

where $\mathcal{L}_{\text{rec}}$ measures pixel-wise reconstruction error and $\mathcal{L}_{\text{KL}}$ is the Kullback–Leibler divergence that regularizes the latent space [1]. For scalograms of size 128×128 pixels ($N_p =$ 16,384):

$$\mathcal{L}_{\text{rec}} = \sum_{p=1}^{N_p} (x_p - \hat{x}_p)^2 \tag{3}$$

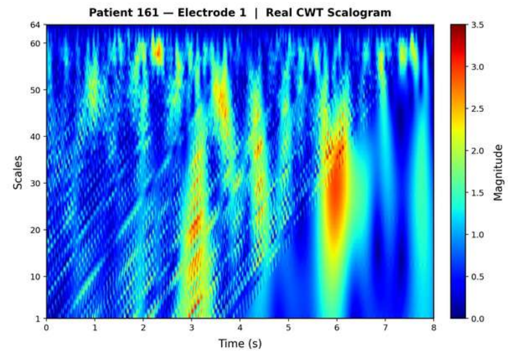
A converged total loss of 118 corresponds to a mean per-pixel error of:

$$\bar{\epsilon} = \frac{\mathcal{L}_{\text{rec}}}{N_p} \approx \frac{118}{16{,}384} \approx 0.007 \tag{4}$$

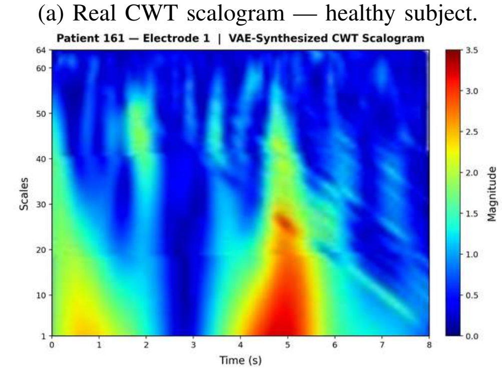
This low per-pixel error confirms that the VAE accurately reconstructs the EEG time–frequency patterns and is suitable for generating realistic synthetic samples.

E. Qualitative Evaluation of VAE-Generated Scalograms

The physiological significance of the data generated from the VAE is demonstrated in Fig. 4, where a real CWT scalogram and a VAE scalogram of electrode Fp1 from a healthy individual are compared.



(a) Real CWT scalogram — healthy subject.



(b) VAE-synthesized CWT scalogram — healthy subject.

Fig. 4: Qualitative comparison of CWT scalograms for electrode Fp1 (healthy subject).

The real Scalogram (a) shows the typical time-frequency energy distribution pattern with localized high magnitude zones that signify the occurrence of natural neural oscillations. Scalogram shows nonstationary spectral-temporal behavior as expected from resting state EEG data collected under controlled circumstances.

The VAE-generated scalogram (b) closely reproduces the global morphology of the real signal. The spatial distribution of energy across scales, the temporal extent of activation patterns, and the relative magnitude range are all qualitatively consistent with the real counterpart. Importantly, the reconstructed scalogram does not replicate the actual pixels from any one training sample; instead, it is an interpolated value in the latent space. [1].

This result confirms that the VAE has successfully captured the underlying generative structure of EEG time–frequency representations, rather than overfitting to individual training instances. The ability to produce such realistic scalograms justifies the use of VAE-based augmentation as a principled strategy for increasing the diversity of the training set, particularly for the under-represented healthy class in the FLORIDA dataset.

F. Deep Learning Architectures

The augmented dataset is evaluated using three complementary deep learning architectures.

1) Residual Network (ResNet-18): ResNet-18 [2] employs residual connections to stabilize the training of deep architectures. Each residual block learns an incremental correction to its input rather than a full mapping, improving gradient flow and mitigating overfitting. A pre-trained ResNet-18 is fine-tuned on EEG scalograms to leverage transferable low-level visual features:

$$y = F(x) + x \tag{5}$$

2) Variational Autoencoder (VAE): The VAE [1] serves as a generative module rather than a classifier. Its latent space is parameterized by a mean μ and variance σ^2, with samples drawn via the reparameterization trick:

$$z = \mu + \sigma \odot \epsilon, \quad \epsilon \sim \mathcal{N}(0, I) \tag{6}$$

3) Vision Transformer (ViT): The Vision Transformer [3] partitions each scalogram into fixed-size patches, which are linearly embedded and enriched with positional encodings before being processed by stacked self-attention layers. Unlike CNNs, ViT captures global time–frequency dependencies across the full input, which is advantageous for modeling complex EEG dynamics:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^\top}{\sqrt{d}}\right)V \tag{7}$$

IV. Achieved Performances and Model Comparison

All models were evaluated at the patient level in order to prevent data leakage between training and testing sets. Performance was assessed using standard metrics, including accuracy, precision, recall, and F1-score. In addition, patient-level confusion matrices were analyzed to better understand classification behavior.

Impact of Model Architecture and Data Augmentation: Figure 6 presents a global comparison of the different architectures and data augmentation strategies.

Models trained on raw scalograms, including CNN and ResNet-18 [2], achieve very high performance across all metrics. Although these results appear optimal, the continuously high scores strongly indicate overfitting, which is probably caused by the limited size and similarity of the original dataset [14].

Standard image-based augmentation [8] enhances diversity and somewhat increases data variability, but precision is frequently sacrificed in the process. Such augmentations rely on pixel-level transformations and do not explicitly preserve the physiological structure of EEG time–frequency representations. In contrast, the ResNet-18 model [2] trained with VAE-based augmentation [1], [9] achieves high performance while avoiding perfect scores, indicating improved generalization. This configuration benefits from synthetic scalograms that reflect realistic EEG patterns and increase inter-subject variability. Misclassifications remain limited, as illustrated by the patient-level confusion matrix in Fig. 5.

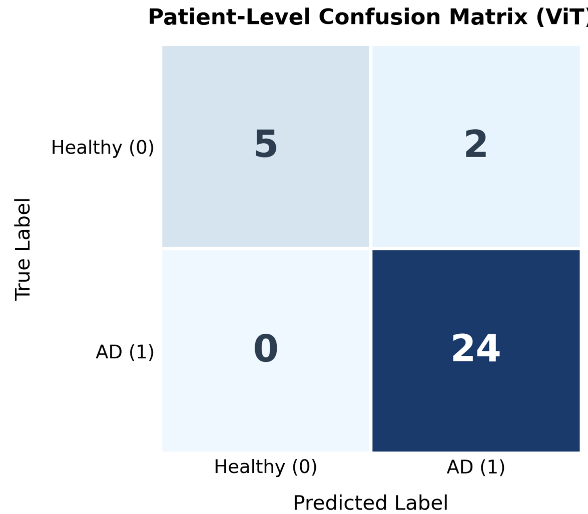


Fig. 5: Patient-level confusion matrix obtained with the Vision Transformer (ViT) model trained on VAE-augmented EEG scalograms.

The Vision Transformer (ViT) model [3] also demonstrates competitive performance. Thanks to its self-attention mechanism, ViT effectively captures long-range dependencies across EEG scalograms from multiple electrodes. However, its performance is more sensitive to data availability, which highlights the importance of VAE-based augmentation for stabilizing training and improving robustness.

1) Quantitative Results with ResNet-18 and VAE Augmentation: Table I summarizes the average patient-level performance of the pre-trained ResNet-18 model [2] trained on VAE-augmented EEG scalograms. The model achieves high accuracy and F1-score, confirming the effectiveness of generative augmentation in improving generalization while maintaining strong discriminative power.

TABLE I: Average performance metrics of ResNet-18 with VAE-augmented EEG scalograms

Metric	Accuracy	Precision	Recall	F1-score
Average	0.977	1.000	0.914	0.952

2) Patient-Level Evaluation with Vision Transformer: Figure 5 presents the patient-level confusion matrix obtained using the Vision Transformer (ViT) [3] model trained on VAE-augmented EEG scalograms. The evaluation confirms strong diagnostic performance, with an accuracy of 93.6%, a recall of 1.000, and an F1-score of 0.960, achieved with an evaluation time of 48.29 seconds.

TABLE II: Patient-level performance metrics of the Vision Transformer (ViT)

Metric	Accuracy	Recall	F1-score
ViT	0.9355	1.000	0.960

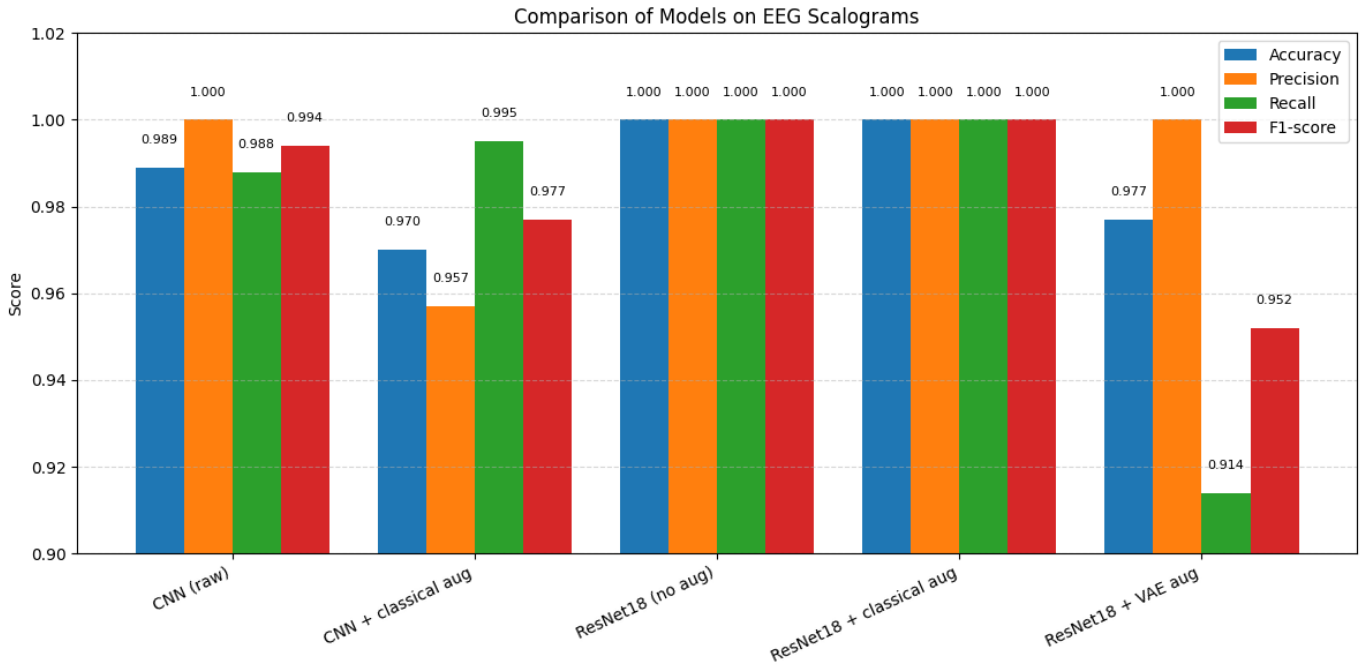


Fig. 6: Global comparison of model performances on EEG scalograms. VAE-based augmentation improves generalization while maintaining high classification accuracy.

The ViT model shows high sensitivity to abnormal EEG patterns, as demonstrated by the perfect recall, which shows that all patients with Alzheimer's disease are accurately identified. Only the healthy class, which remains smaller despite data augmentation, indicates misclassifications. A conservative diagnostic approach that reduces false negatives is reflected in this behaviour, which is expected in medical classification tasks.

The ViT [3] works well because it processes all 21 electrode scalograms together using self-attention, capturing high-level disease patterns across electrodes. VAE-generated synthetic subjects [1], [9] further balance and diversify the training data, improving generalization while keeping predictions clinically meaningful.

V. Discussion and Contributions

This study places itself at the intersection of deep learning techniques using time-frequency representations and conventional EEG data analysis. There are still significant methodological issues in the literature, despite the fact that a number of earlier studies show excellent classification performance on the FLORIDA dataset or similar cohorts. Handmade feature techniques, such as that of Vicchietti *et al.* [17], extract descriptors grounded in information theory, nonlinear dynamics, and graph-based signal models. While these approaches can yield high classification metrics, they require substantial signal-processing expertise and provide no mechanism for addressing dataset imbalance through augmentation. Critically, near-perfect accuracy and specificity values reported in such studies were predominantly obtained on late-stage AD cohorts, which raises legitimate concerns regarding overfitting [14] and restricted generalisation to early-stage or previously unseen individuals.

More recent work has shifted toward deep learning models trained on time–frequency EEG representations. CNNs applied to spectrograms or CWT-based scalograms circumvent the need for manual feature engineering [18]; however, these architectures are typically trained on small datasets and rely on pixel-level augmentation strategies (rotations, flips, intensity shifts) [8] that fail to capture the underlying physiological variability of EEG signals. Among CNN variants, deeper architectures such as ResNet [2] consistently outperform shallower models like LeNet-5, particularly on imbalanced data, owing to their superior capacity for hierarchical feature extraction.

A representative example of this class of methods is the work by Lekkan *et al.* [11], which constructs a deterministic pipeline combining multiple time–frequency transforms with CNN classifiers. Although competitive, this approach does not model the generative distribution of EEG data and therefore remains sensitive to dataset size and inter-subject diversity.

The proposed approach addresses such limitations via the introduction of a Variational Autoencoder (VAE) [1] to facilitate the generation of new EEG scalograms based on the CWT method. By learning a structured latent representation of multi-channel time-frequency patterns, the VAE generates biologically realistic EEG scalograms [9] to overcome class imbalance within the FLORIDA dataset.
This strategy reduces the risk of overfitting that affects both handcrafted-feature pipelines and standard CNN-based approaches.

The proposed method achieves a classification accuracy of approximately 97%, a result that compares favourably with prior work [11], [18] while remaining within a plausible performance range that does not suggest memorisation of the training set. By integrating generative modelling with state-of-the-art architectures—ResNet-18 [2] and Vision Transformer (ViT) [3]—the framework improves generalisation while ensuring a reproducible and computationally tractable analysis pipeline.

VI. Conclusion and Future Work

This research study combines time-frequency scalograms with several augmentation strategies for EEG-based Alzheimer's disease classification. Although traditional augmentation methods [8] have improved outcomes, they present limitations that result in overfitting issues [14]. A novel generative data augmentation technique based on Variational Autoencoders (VAEs) [1] was applied to EEG scalograms to address these limitations.

The findings demonstrate that this approach maintains strong diagnostic performance at the patient level while enhancing model resilience and generalisation. The proposed framework, which combines ResNet-18 [2] and ViT [3] architectures, reduces reliance on complex traditional signal-processing features and offers a more scalable and clinically realistic solution. Future research will investigate more sophisticated generative models and expand the methodology to larger clinical datasets, including early-stage Alzheimer's disease and multi-modal data.

References

[1] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," 2014.

[2] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[3] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An image is worth 16x16 words: Transformers for image recognition at scale," in *International Conference on Learning Representations (ICLR)*, 2021.

[4] J. Jeong, "EEG dynamics in patients with Alzheimer's disease," *Clinical Neurophysiology*, vol. 115, no. 7, pp. 1490–1505, 2004.

[5] R. Cassani, M. Estarellas, R. San-Martin, F. J. Fraga, and T. H. Falk, "Systematic review on resting-state EEG for Alzheimer's disease diagnosis and progression assessment," *Disease Markers*, vol. 2018, p. 5174815, 2018.

[6] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces," *Journal of Neural Engineering*, vol. 15, no. 5, p. 056013, 2018.

[7] I. A. Fouad, A. Labib, and M. Kholief, "EEG time–frequency analysis for Alzheimer's disease detection," *Biomedical Signal Processing and Control*, vol. 68, p. 102548, 2021.

[8] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, no. 1, p. 60, 2019.

[9] Y. Luo, L.-Z. Zhu, Z.-Y. Wan, and B.-L. Lu, "Data augmentation for enhancing EEG-based emotion recognition with deep generative models," *Journal of Neural Engineering*, vol. 17, no. 5, p. 056021, 2020.

[10] C. J. Stam, "Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field," *Clinical Neurophysiology*, vol. 116, no. 10, pp. 2266–2301, 2005.

[11] A. Lekkan, "Time–frequency transformations for enhanced biomedical signal classification with convolutional neural networks," *Biomedical Signal Processing and Control*, 2025.

[12] K. G. Hartmann, R. T. Schirrmeister, and T. Ball, "EEG-GAN: Generative adversarial networks for electroencephalographic (EEG) data augmentation," *arXiv preprint arXiv:1806.01875*, 2018.

[13] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: A systematic review," *Journal of Neural Engineering*, vol. 16, no. 5, p. 051001, 2019.

[14] F. Pellegrini *et al.*, "Overfitting in machine learning applied to EEG classification," *IEEE Access*, 2021.

[15] P. D. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Transactions on Audio and Electroacoustics*, vol. 15, no. 2, pp. 70–73, 1967.

[16] P. S. Addison, *The Illustrated Wavelet Transform Handbook: Introductory Theory and Applications in Science, Engineering, Medicine and Finance*, 2nd ed. CRC Press, 2017.

[17] M. L. Vicchietti, F. M. Ramos, L. E. Betting, and A. S. L. O. Campanharo, "Computational methods of EEG signals analysis for Alzheimer's disease classification," *Scientific Reports*, vol. 9, p. 13190, 2019.

[18] K. Stefanou, K. D. Tzimourta, C. Bellos, G. Stergios, M. G. Tsipouras, and A. T. Tzallas, "A novel CNN-based framework for Alzheimer's disease detection using EEG spectrogram representations," *Journal of Personalized Medicine*, vol. 15, no. 1, p. 27, 2025.