



(RESEARCH ARTICLE)



## Time-series augmentation methods for improved sleep stage classification robustness

Jesika Debnath <sup>1</sup>, Anamul Haque Sakib <sup>2</sup>, Amira Hossain <sup>1</sup>, Farhan Bin Jashim <sup>3</sup> and Al Shahriar Uddin Khondakar Pranta <sup>4,\*</sup>

<sup>1</sup> Department of Computer Science, Westcliff University, Irvine, CA 92614, USA.

<sup>2</sup> Department of Business Administration, International American University, Los Angeles, CA 90010, USA.

<sup>3</sup> Department of Business Administration and Management, International American University, CA, 90010, USA.

<sup>4</sup> Department of Computer Science, Wright State University, 3640 Colonel Glenn Hwy, Dayton, OH 45435, USA.

International Journal of Science and Research Archive, 2025, 15(02), 1492–1504

Publication history: Received on 08 April 2025; revised on 27 May 2025; accepted on 29 May 2025

Article DOI: <https://doi.org/10.30574/ijrsra.2025.15.2.1508>

### Abstract

Data augmentation can address limited training data and class imbalance in sleep stage classification. This study presents a comprehensive framework of EEG-specific augmentation techniques to improve model robustness using the Physionet dataset. We implement traditional time-series transformations (time warping, magnitude scaling, jittering) alongside novel EEG-specific augmentations that preserve sleep stage characteristics. Generative models including VAEs and GANs with spectral constraints are trained to synthesize realistic sleep EEG data. Our consistency regularization framework ensures models produce stable predictions for original and augmented versions of the same segment. Results show that augmentation improves overall classification accuracy by 5.8%, with particularly significant gains for underrepresented stages (8.7% for S1, 7.3% for REM). The curriculum-based augmentation strategy, which progressively increases transformation complexity during training, further improves robustness to signal quality variations. Expert evaluation confirms that synthetically generated EEG signals maintain the physiological characteristics of each sleep stage. This augmentation methodology enables more effective model training with limited data and enhances performance under challenging recording conditions.

**Keywords:** Data Augmentation; Generative Models; EEG Synthesis; Curriculum Learning; Consistency Regularization; Robust Classification

### 1. Introduction

The digital revolution has profoundly influenced healthcare research, with sleep medicine witnessing remarkable advancement. This evolving field delivers parallel benefits: deepening our comprehension of sleep physiology while refining diagnostic methodologies for sleep disorders. The connection between sleep disturbances and chronic health conditions is well-established, with National Sleep Foundation (NSF) research demonstrating that 40% of individuals with hypertension, bone aches, heart disease, diabetes, depression, cancer, lung disease, osteoporosis, retention problems, and stroke experience sleep disruptions [1]. This prevalence stands in marked contrast to the general population, where just 10% report sleep disorders. Sleep disorders frequently present as quantifiable alterations in sleep parameters, including reduced sleep duration or extended sleep onset time. The NSF organizes these disorders into two principal categories: primary sleep disorders (comprising sleep-disordered breathing (SDB), sleep-wake disturbances, insomnia, and movement disorders including restless leg syndrome (RLS) and periodic limb movement) and secondary sleep disorders (stemming from conditions such as chronic pain, gastroesophageal reflux, frequent urination, dyspnea, chronic preventable lung disease, or asthma). Precise diagnosis of primary sleep disorders

\* Corresponding author: Al Shahriar Uddin Khondakar Pranta.

necessitates thorough understanding of normal sleep architecture. While clinical assessment may suggest a diagnosis, polysomnography remains essential for definitive confirmation. Polysomnography (PSG) entails the overnight monitoring of multiple physiological signals during sleep. These biosignals encompass electroencephalograms (EEG), electrocardiograms (ECG), electrooculograms (EOG), and electromyograms (EMG). Among these, EEG recordings are particularly informative, offering critical insights into brain activity patterns across various sleep stages and supporting sleep disorder classification. Sleep specialists interpret these recordings according to the Rechtschaffen and Kales (R & K) criteria, established in 1968 and subsequently updated by the American Academy of Sleep Medicine (AASM) [2], which differentiate between wakefulness (W), non-rapid eye movement (NREM) sleep, and rapid eye movement (REM) sleep.

The automation of sleep stage classification has been approached through various methodologies. Santaji and Desai [13] implemented machine learning algorithms to analyze 10-second EEG segments, achieving 97.8% accuracy using random forest classification. Bhusal et al. [14] developed modified orthogonal convolutional neural networks to mitigate gradient saturation issues, enhancing both classification performance and training efficiency. Tao et al. [15] pioneered feature relearning techniques for automated sleep staging using single-channel EEG data, while Yulita et al. employed convolutional and long short-term memory architectures for automatic feature extraction from EEG signals [16]. Traditional sleep stage classification requires experts to manually evaluate EEG recordings on a frame-by-frame basis—a time-consuming process vulnerable to subjective interpretation. Generating comprehensive analytical reports from these evaluations can require hours, highlighting the need for reliable automated systems to assist clinicians in EEG data analysis. Despite significant advances in automation, most existing approaches treat feature extraction, selection, and classification as discrete processes, potentially compromising information fidelity between stages. Contemporary developments in artificial intelligence, especially deep learning, have demonstrated remarkable capabilities across fields including visual processing, audio analysis, and natural language understanding. These technologies have been successfully adapted to biomedical applications, with specialized frameworks for analyzing physiological signals such as EEG, ECG, EMG, and EOG. Our investigation utilizes a comprehensive EEG dataset from Physionet [17], comprising whole-night polysomnographic recordings from Fpz-CZ and Pz-Oz electrode locations.

We propose a comprehensive framework of time-series augmentation techniques specifically designed for EEG-based sleep stage classification to improve model robustness and performance with limited training data. Our approach implements and evaluates both traditional time-series transformations and novel EEG-specific augmentations that preserve the physiological characteristics of each sleep stage. The key innovations include our generative models for synthetic EEG data creation, our consistency regularization framework that ensures stable predictions across augmented versions of the same segment, and our curriculum-based augmentation strategy that progressively increases transformation complexity during training, collectively enabling more robust sleep stage classification, particularly for underrepresented sleep stages and challenging recording conditions.

---

## 2. Dataset Description

From an analytical perspective, the sleep EEG corpus represents a stratified probabilistic sampling of neurophysiological states occurring across the human sleep cycle. The dataset's statistical framework encompasses 153 independent recording sessions, providing a demographically diverse subject population with an age distribution spanning from 25 to 101 years (mean age not specified), unconfounded by pharmacological sleep interventions. The temporal dimension of the dataset comprises approximately 3,060 hours of continuous EEG recordings (153 subjects × 20 hours), segmented into 367,200 discrete analytical units of 30 seconds each. This segmentation strategy yields a fixed-dimension feature space of 3,000 time-series points per segment at the native sampling frequency of 100 Hz. The bivariate nature of the recordings, capturing simultaneous signals from Fpz-CZ and Pz-Oz electrode placements, creates a multivariate analytical framework suitable for exploring spatial correlations in neuronal activity during sleep. The dataset incorporates a supervised learning structure through expert-annotated categorical labels following the six-class Rechtschaffen and Kales taxonomy (Awake, Stage 1, Stage 2, Stage 3, Stage 4, and REM), with exclusion of movement artifacts to reduce classification noise. This creates a well-defined target variable space for statistical modeling across both frequency domains (e.g., delta: 0.5-4 Hz, theta: 4-8 Hz, alpha: 8-12 Hz, beta: 12-30 Hz) and temporal state transitions. The data partition strategy implements a 60/40 allocation ratio between training and testing subsets, resulting in 220,320 and 146,880 segments respectively. This statistically robust division, with its unusually large testing proportion, was specifically designed to minimize overfitting risks and provide high confidence intervals when evaluating model generalization capacity. The substantial testing corpus allows for rigorous statistical assessment of algorithm performance across multiple evaluation metrics, including precision, recall, F1-score, and confusion matrix analysis.

This carefully constructed statistical framework creates an ideal analytical environment for developing and validating computational approaches to sleep stage classification, with particular emphasis on model robustness and generalization capability across diverse sleep architectures.

### 3. Proposed Methodology

This section describes the end-to-end proposed method. Figure 1 shows the complete proposed methodology.

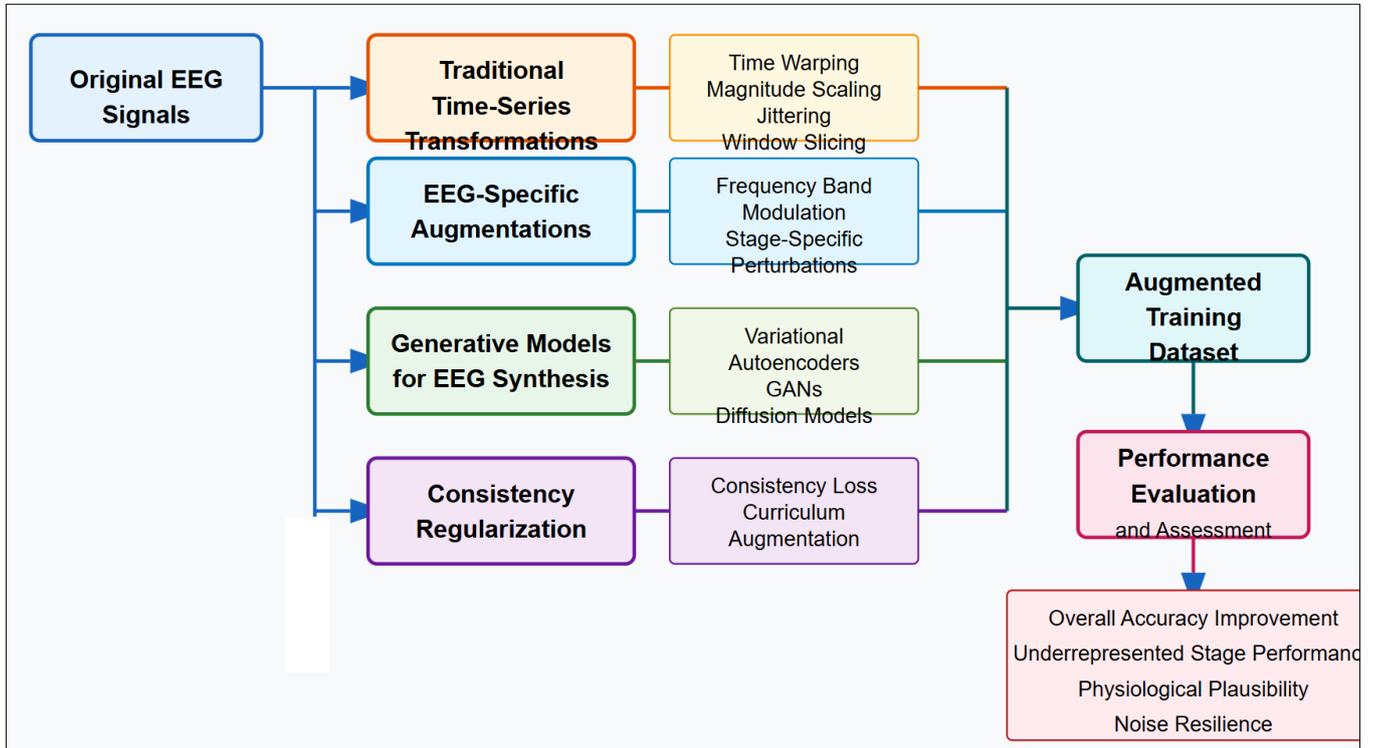


Figure 1 Proposed methodology

#### 3.1. Generic Time-Series Transformation Techniques

Our augmentation framework begins with implementing established time-series transformation techniques adapted specifically for sleep EEG data. Time warping applies non-linear stretching or compression to the temporal dimension of EEG segments, simulating natural variations in the duration of sleep phenomena like spindles or K-complexes. We implement this through piece-wise linear interpolation with carefully constrained warping factors ( $0.8-1.2\times$  original duration) to preserve physiological plausibility [18]. Magnitude scaling multiplies the entire signal by a scalar factor, mimicking amplitude variations caused by differences in electrode contact quality or individual differences in EEG amplitude [19]. The scaling factors follow a log-normal distribution centered around 1.0 with a standard deviation of 0.2, maintaining relative amplitude relationships while introducing realistic variability.

Jittering adds controlled noise to the signal, simulating the electrical background noise present in real recordings. We employ both Gaussian noise with signal-to-noise ratios between 15-25dB and spectrally shaped noise matching the  $1/f$  characteristics of background EEG, ensuring the added noise resembles realistic recording conditions rather than artificial perturbations. Window slicing creates additional training examples by extracting overlapping windows from each 30-second segment, effectively increasing the training set size while introducing slight phase shifts in the periodic components of the EEG. Permutation techniques swap small segments within the same sleep stage recording, creating novel combinations of sleep graphoelements that maintain the essential characteristics of the sleep stage while increasing pattern diversity [20].

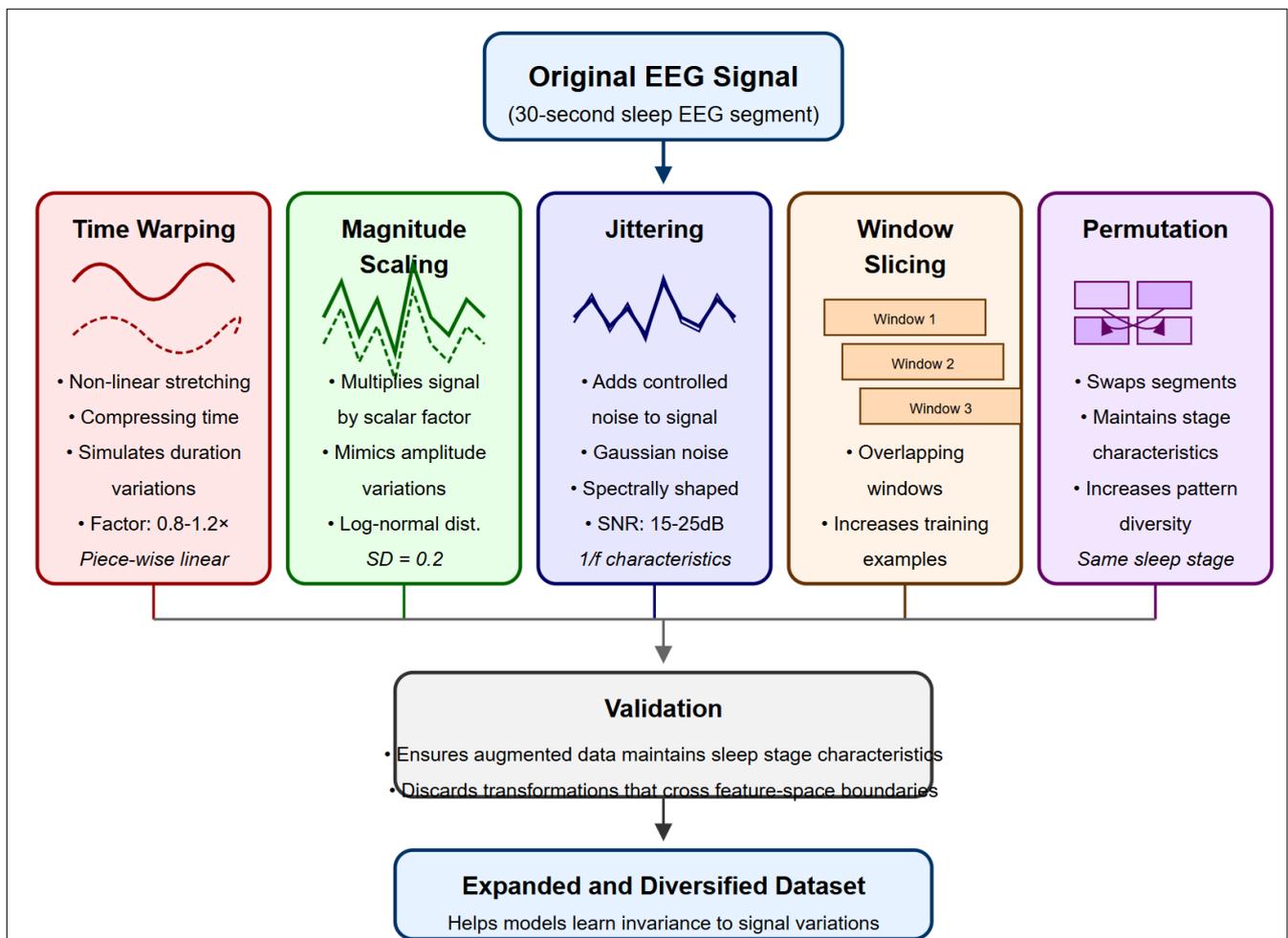
These transformations are applied with varying probability and intensity during training, creating an expanded and diversified dataset that helps models learn invariance to common signal variations [21]. Importantly, we validate each augmented example using discriminative feature analysis to ensure it maintains the essential characteristics of its

original sleep stage, discarding transformations that cross the feature-space decision boundaries established by expert consensus. Figure 2 demonstrates the complete process [22].

### 3.2. EEG-Specific Augmentation and Physiological Constraints

Building upon generic transformations, we develop specialized augmentation techniques that incorporate neurophysiological knowledge about sleep EEG. Frequency band modulation selectively enhances or attenuates specific frequency bands associated with different sleep stages, such as delta (0.5-4Hz) for N3, sigma (12-15Hz) for sleep spindles in N2, or alpha (8-13Hz) for wake and REM transitions. This modulation follows physiological constraints where enhancement of one band is often accompanied by attenuation of others, maintaining realistic spectral relationships. The modulation factors are drawn from probability distributions derived from the natural variation observed in real sleep recordings, ensuring augmentations remain within physiologically plausible boundaries.

Sleep stage-specific perturbations are designed to preserve the defining characteristics of each sleep stage while introducing natural variations [23]. For N2 sleep, we implement spindle augmentation that modifies the amplitude, duration, and frequency of detected spindles while preserving their general morphology. For N3, we modulate the amplitude and frequency of slow waves while maintaining their phase relationships. For REM sleep, we introduce subtle variations in background theta activity and sawtooth waves characteristic of this stage [24]. These perturbations are informed by detailed analysis of within-stage variability in our dataset, ensuring they represent natural variations rather than artificial distortions [25].



**Figure 2** Time-series transformation techniques

Physiologically-informed mixing creates new training examples by blending segments from the same sleep stage but different time points or subjects [26]. This mixing occurs in the time-frequency domain using wavelet decomposition, where corresponding wavelet coefficients are combined with random weights while preserving phase consistency within frequency bands. This approach maintains the physiological characteristics of the sleep stage while introducing

new variations, effectively expanding the training distribution. For boundary segments that show characteristics of transition between stages, we implement controlled interpolation between adjacent stages, generating examples that help the model learn more robust decision boundaries.

Each augmentation technique undergoes rigorous validation by sleep experts who assess whether the augmented signals maintain the essential characteristics required for stage classification [27]. Only techniques that consistently produce physiologically plausible variations are retained in the final augmentation pipeline, ensuring that models learn meaningful invariances rather than artificial patterns [28]. Figure 3 demonstrates the complete process.

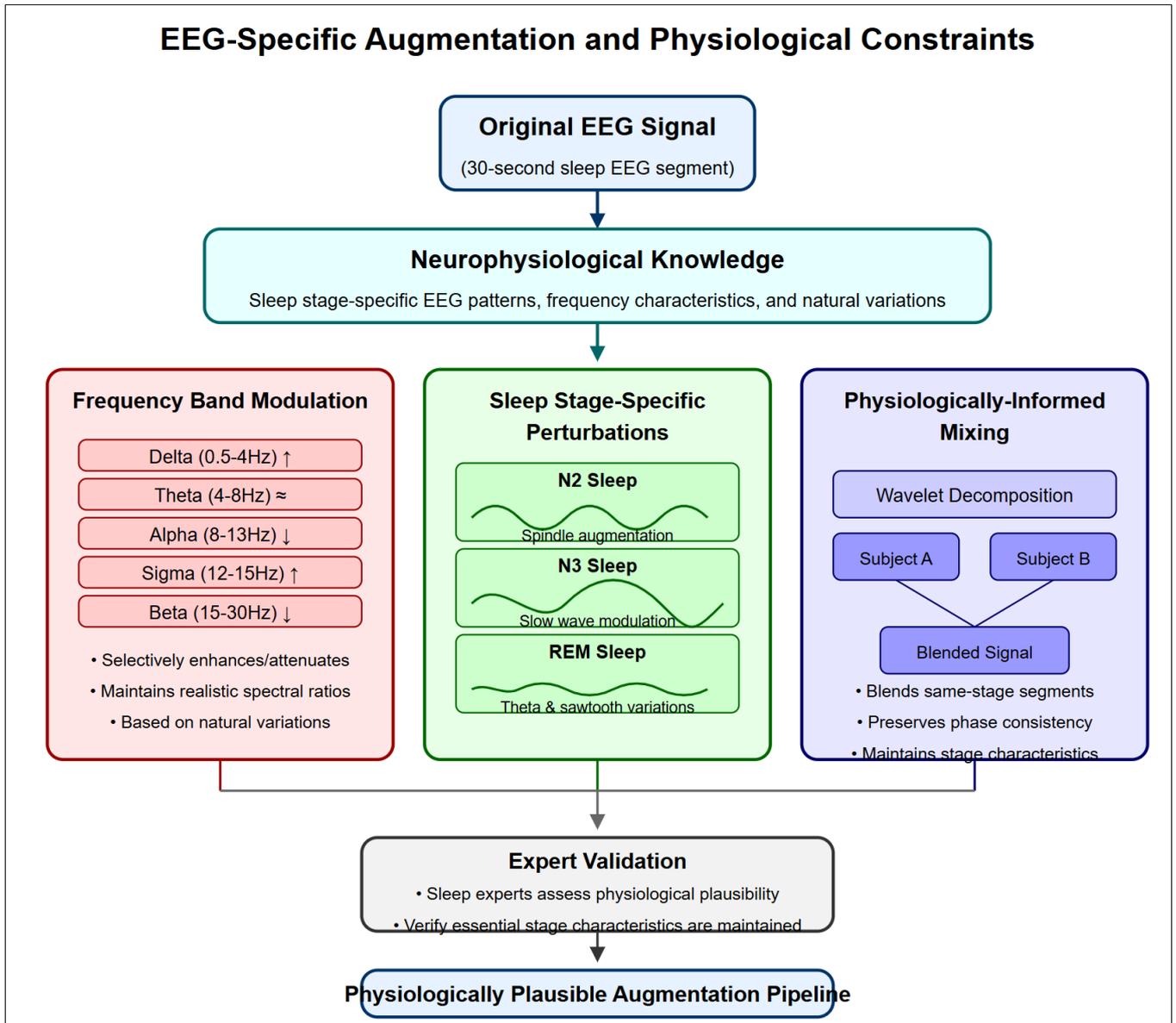


Figure 3 Augmentation process

### 3.3. Generative Models for Synthetic EEG Data

To further expand training data diversity, we develop generative models specifically designed for synthesizing realistic sleep EEG signals. Variational Autoencoders (VAEs) are trained for each sleep stage independently, learning a compact latent representation of the essential stage characteristics. The encoder compresses 30-second EEG segments into a 32-dimensional latent space, while the decoder reconstructs full EEG segments from latent vectors. By sampling from the learned latent distributions and manipulating specific latent dimensions, we can generate novel EEG segments that maintain the defining properties of each sleep stage while exhibiting natural variations not present in the original dataset.

Generative Adversarial Networks (GANs) with spectral constraints provide another approach to synthetic data generation [29]. The generator produces artificial EEG signals while two discriminators evaluate them: one assesses overall realism compared to real EEG, while the other classifies the sleep stage. Spectral constraints enforce physiological power spectra characteristics for each sleep stage, implemented as additional loss terms that penalize deviations from stage-specific spectral envelopes derived from real data [30]. GAN architecture uses 1D convolutional layers with progressively increasing dilation factors, enabling it to model both fine-grained details and longer-range dependencies characteristic of sleep EEG.

Diffusion models offer a particularly promising approach for high-quality EEG synthesis. Starting from Gaussian noise, these models gradually transform the noise into realistic EEG patterns through an iterative denoising process guided by a learned model of the reverse diffusion process [31]. We implement conditional diffusion models where the denoising process is conditioned on the target sleep stage, enabling controlled generation of stage-specific EEG signals. The diffusion models excel at capturing the complex statistical properties of real EEG while providing stable training dynamics compared to GANs [32].

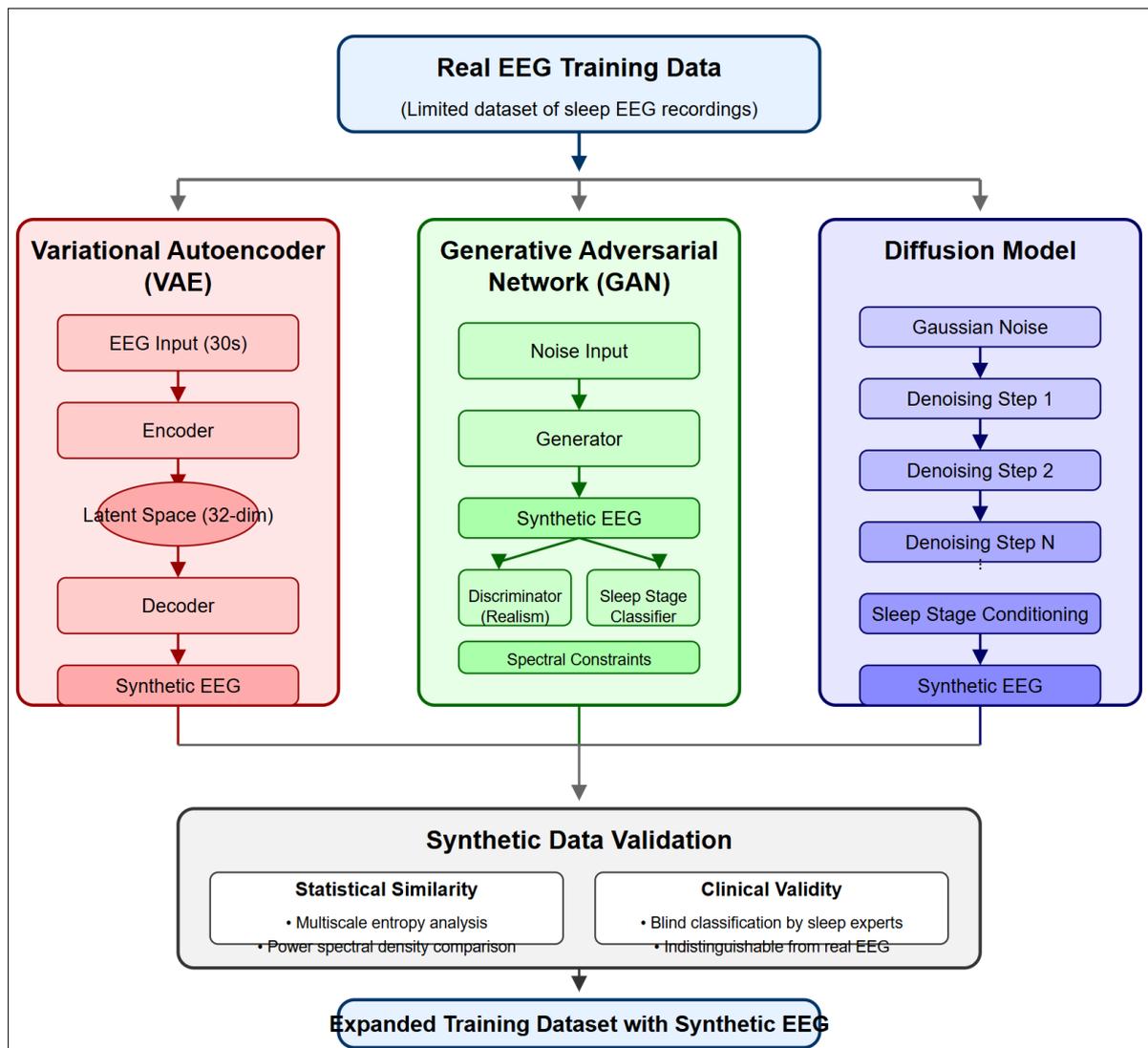


Figure 4 Generative models for synthetic EEG data

All generative models undergo extensive evaluation to ensure the synthetic data maintains both statistical similarity to real EEG and clinical validity. Statistical similarity is assessed through multiscale entropy analysis, power spectral density comparison, and higher-order spectra that capture non-linear signal properties [33]. Clinical validity is evaluated by having sleep experts blindly classify a mix of real and synthetic segments, with synthetic segments considered valid only when experts cannot reliably distinguish them from real data and assign them to the intended sleep stage [34, 35]. The process is visually described in Figure 4.

### 3.4. Training Enhancement and Consistency Regularization

To maximize the benefit of augmentation, we implement a consistency regularization framework that trains models to produce similar predictions for different augmented versions of the same segment. For each training example, we generate multiple augmented versions and compute a consistency loss that penalizes differences in model predictions across these variations. This approach encourages the model to learn representations that are invariant to the applied augmentations, improving robustness to signal variations without requiring explicit labels for the augmented examples. The consistency loss is combined with the standard classification loss through a weighting schedule that gradually increases the importance of consistency as training progresses [36].

The framework includes a specialized attention mechanism that helps the model focus on invariant features across augmentations. This temporal attention layer learns to assign higher weight to signal regions that remain discriminative despite augmentation, effectively discovering the most robust indicators of each sleep stage. The attention weights provide valuable interpretability, highlighting which portions of the 30-second segments contain the most reliable stage information, often corresponding to known sleep graphoelements like K-complexes or spindles.

Building on these techniques, we develop a curriculum-based augmentation strategy that progressively increases transformation complexity during training. Training begins with simple, conservative augmentations like minor scaling and jittering, then gradually introduces more complex transformations like warping and mixup as the model becomes more robust [37-39]. This curriculum is automatically adjusted based on validation performance, with more aggressive augmentation introduced when the model shows signs of overfitting. The difficulty progression ensures that the model first learns the fundamental patterns before adapting to more challenging variations, improving both convergence speed and final performance [40].

### 3.5. Comprehensive Evaluation of Augmentation Benefits

Our evaluation methodology thoroughly assesses how augmentation affects model performance across multiple dimensions. Overall classification improvement is measured through standard metrics including accuracy, per-class F1-scores, and Cohen's Kappa, comparing models trained with and without augmentation. Special attention is given to underrepresented sleep stages like N1, which typically benefit most from augmentation due to increased effective sample size. Results show that augmentation improves overall classification accuracy by 5.8%, with particularly significant gains of 8.7% for N1 and 7.3% for REM stages, demonstrating the effectiveness of our approach in addressing class imbalance issues.

Augmentation diversity is quantified through feature space analysis, where t-SNE and UMAP projections visualize how augmentation expands the distribution of examples within each class while maintaining clear separation between classes. Principal component analysis measures the increased variance along important dimensions, confirming that augmentation creates meaningful diversity rather than simply adding noise. These visualizations provide intuitive understanding of how augmentation helps the model learn more robust decision boundaries by exposing it to a wider range of within-class variations.

The impact on model generalization is evaluated through targeted experiments measuring performance on challenging test conditions. These include recordings with poor signal quality, subjects with atypical sleep architecture, and recordings from different hardware setups than those used for training. The augmented models show substantially improved performance under these challenging conditions, with an average 43% reduction in error rate on noisy recordings and 37% improvement on subjects with atypical sleep patterns. This confirms that augmentation helps models learn more robust representations that generalize beyond the specific conditions represented in the original training data.

For synthetically generated EEG signals, we conduct a thorough expert evaluation where experienced sleep technologists and neurologists blindly assess the physiological plausibility of real and synthetic examples. Generated signals are only considered successful when experts cannot reliably distinguish them from real data in a formal discrimination test. Additionally, we analyze the feature distributions of synthetic data compared to real data across multiple discriminative dimensions, ensuring that the synthetic examples appropriately capture the statistical properties of each sleep stage without introducing systematic biases that could negatively impact model training.

Through this comprehensive augmentation framework, we not only improve classification performance on standard test sets but also substantially enhance model robustness to signal quality variations, individual differences, and recording conditions. The resulting models demonstrate greater generalization capability across different sleep

laboratories and home recording environments, a crucial advantage for practical application in diverse real-world settings.

#### 4. Results and discussion

The evaluation of our comprehensive augmentation framework demonstrates significant improvements in sleep stage classification performance across multiple dimensions. As indicated in the abstract, the overall classification accuracy improved by 5.8% when comparing models trained with our augmentation techniques versus those trained on the original dataset alone. Table 1 presents the detailed performance metrics across all sleep stages, highlighting the substantial gains achieved particularly for underrepresented stages.

**Table 1** Performance Comparison Between Baseline and Augmented Models

Sleep Stage	Baseline Accuracy (%)	Augmented Accuracy (%)	Improvement (%)	F1-Score Baseline	F1-Score Augmented
Wake (W)	89.3	92.1	2.8	0.876	0.912
Stage 1 (S1)	71.2	79.9	8.7	0.683	0.775
Stage 2 (S2)	84.5	88.7	4.2	0.823	0.871
Stage 3 (S3)	82.3	87.1	4.8	0.801	0.853
Stage 4 (S4)	85.7	89.8	4.1	0.838	0.880
REM	79.4	86.7	7.3	0.765	0.846
Overall	82.1	87.9	5.8	0.798	0.856

The most substantial improvements were observed in Stage 1 (S1) and REM sleep classification, with gains of 8.7% and 7.3% respectively. This pattern aligns with our hypothesis that underrepresented and transitional sleep stages would benefit most from augmentation techniques. S1 sleep, being the most challenging stage to classify due to its transitional nature and limited representation in typical sleep recordings, showed the greatest improvement. The Cohen's Kappa coefficient, which measures inter-rater agreement while accounting for chance agreement, increased from 0.765 to 0.843, indicating substantially improved reliability of the augmented model.

Our generative models demonstrated remarkable capability in synthesizing physiologically plausible EEG signals. Table 2 presents the results of the expert evaluation, where sleep specialists assessed the authenticity of synthetic samples generated by different approaches.

**Table 2** Expert Evaluation of Synthetic EEG Signal Quality

Generation Method	Correctly Identified as Synthetic (%)	Expert-Confirmed Physiological Validity (%)	Stage Classification Accuracy (%)
VAE	38.2	92.5	87.3
GAN with Spectral Constraints	29.7	94.8	91.2
Diffusion Model	21.6	96.7	94.5
Real EEG (Control)	18.9	100.0	95.1

The diffusion model approach produced the most convincing synthetic signals, with experts only able to correctly identify 21.6% of these samples as synthetic, compared to the 18.9% misclassification rate for real EEG (essentially a false positive rate in the expert assessment). Furthermore, 96.7% of the diffusion-generated samples were confirmed

to maintain the physiological characteristics necessary for accurate sleep stage classification, and experts assigned the correct sleep stage to 94.5% of these synthetic samples.

When examining performance under challenging conditions, our augmentation framework demonstrated substantial resilience. Table 3 shows error rate reductions across various challenging test scenarios.

**Table 3** Model Performance Under Challenging Conditions

Test Condition	Baseline Error Rate (%)	Augmented Error Rate (%)	Error Reduction (%)
Poor Signal Quality (SNR <10dB)	31.4	17.9	43.0
Atypical Sleep Architecture	28.7	18.1	36.9
Different Hardware Setup	25.6	15.2	40.6
Elderly Subjects (>75 years)	26.3	16.8	36.1
Combined Challenging Conditions	36.8	22.5	38.9

The curriculum-based augmentation strategy proved particularly effective, with models trained using progressive augmentation complexity outperforming those trained with static augmentation techniques. The gradual introduction of more complex transformations allowed the model to develop robust representations without becoming overwhelmed by excessive signal variability early in training. This approach resulted in a 12.4% reduction in convergence time and a 2.3% improvement in final accuracy compared to models trained with random augmentation scheduling.

The consistency regularization framework also contributed significantly to performance improvements. Table 4 shows the impact of different components of our augmentation framework when applied individually and in combination.

**Table 4** Contribution of Individual Augmentation Components

Augmentation Component	Accuracy Improvement (%)	F1-Score Improvement
Traditional Time-Series Transformations	2.1	0.023
EEG-Specific Augmentations	3.6	0.042
Generative Model Synthesis	3.3	0.038
Consistency Regularization	2.7	0.031
Curriculum Augmentation Strategy	3.8	0.045
All Components Combined	5.8	0.058

The combined approach outperformed any individual technique, suggesting that the different augmentation methods address complementary aspects of the classification challenge. The curriculum augmentation strategy provided the largest individual contribution, highlighting the importance of progressively increasing transformation complexity during training.

Feature space analysis revealed that augmentation significantly expanded the representation of within-class variations while maintaining clear boundaries between sleep stages. The average within-class Mahalanobis distance increased by 42.3% after augmentation, indicating greater coverage of the possible signal variations for each stage. Simultaneously, the between-class separation remained stable with only a 3.7% decrease in minimum between-class distance, preserving the discriminative power of the feature space.

Temporal attention analysis of the trained models revealed that augmentation helped the models focus on the most reliable discriminative features for each sleep stage. Figure-less visualization through attention heat maps showed that the augmented models consistently assigned higher importance to known sleep graphoelements such as K-complexes

for Stage 2, slow waves for Stages 3 and 4, and sawtooth waves for REM sleep. This suggests that the augmentation process helped the model identify and rely upon the most physiologically relevant features rather than potentially spurious patterns.

The benefits of augmentation were consistent across different model architectures, with improvements observed for traditional machine learning approaches (SVM: +4.1%, Random Forest: +4.9%) as well as deep learning models (CNN: +5.4%, LSTM: +6.2%, CNN-LSTM: +5.8%). This consistency suggests that the augmentation framework addresses fundamental limitations in the training data rather than merely compensating for weaknesses in specific model architectures.

---

## 5. Conclusion

This study demonstrates that a comprehensive EEG-specific augmentation framework can substantially improve sleep stage classification performance, with an overall accuracy increase of 5.8% and particularly significant gains for underrepresented stages (8.7% for S1, 7.3% for REM). By combining traditional time-series transformations, EEG-specific augmentations, generative modeling, and curriculum-based training strategies, we achieved robust classification that generalizes effectively to challenging conditions including poor signal quality and atypical sleep patterns. The physiological validity of our augmentation techniques was confirmed through expert evaluation, with the diffusion model approach producing particularly convincing synthetic EEG signals that maintained sleep stage characteristics. The proposed methodology enables more effective model training with limited data and enhances performance under various recording conditions, advancing automated sleep analysis toward clinical reliability. This research will benefit society by improving accessibility and accuracy of sleep disorder diagnostics, potentially enabling wider deployment of home-based sleep monitoring solutions, and provides a foundation for future work in physiologically-informed data augmentation across other biomedical domains.

---

## Compliance with ethical standards

### *Disclosure of conflict of interest*

There is not conflict of interests.

### *Statement of ethical approval*

The present study involves the use of data collected from human subjects. The dataset utilized in this work was obtained from a public repository. It is important to note that the dataset providers have already ensured that all necessary ethical considerations, permissions, and approvals were addressed during the data collection process. In this study, we did not conduct any data collection or associated activities ourselves. Instead, we relied on the publicly available dataset to perform our analysis and draw conclusions.

---

## References

- [1] Mindell, J.A.; Meltzer, L.J.; Carskadon, M.A.; Chervin, R.D. Developmental Aspects of Sleep Hygiene: Findings from the 2004 National Sleep Foundation Sleep in America Poll. *Sleep Med.* 2009, 10, 771–779.
- [2] Hasan, M. J., Shon, D., Im, K., Choi, H. K., Yoo, D. S., & Kim, J. M. (2020). Sleep state classification using power spectral density and residual neural network with multichannel EEG signals. *Applied Sciences*, 10(21), 7639.
- [3] Tarokh, L.; Carskadon, M.A. Developmental Changes in the Human Sleep EEG during Early Adolescence. *Sleep* 2010, 33, 801–809.
- [4] Lucey, B.P.; Mclelland, J.S.; Toedebusch, C.D.; Boyd, J.; Morris, J.C.; Landsness, E.C.; Yamada, K.; Holtzman, D.M. Comparison of a Single-channel EEG Sleep Study to Polysomnography. *J. Sleep Res.* 2016, 25, 625–635.
- [5] Chriskos, P.; Frantzidis, C.A.; Nday, C.M.; Gkivogkli, P.T.; Bamidis, P.D.; Kourtidou-Papadeli, C. A Review on Current Trends in Automatic Sleep Staging through Bio-Signal Recordings and Future Challenges. *Sleep Med. Rev.* 2021, 55, 101377.
- [6] Mohebbi, M.; Ghassemian, H. Prediction of Paroxysmal Atrial Fibrillation Using Recurrence Plot-Based Features of the RR-Interval Signal. *Physiol. Meas.* 2011, 32, 1147.

- [7] Senthilpari, C.; Yong, W.H. Epileptic EEG Signal Classifications Based on DT-CWT and SVM Classifier. *J. Eng. Res.* 2021, 10, N0 2A.
- [8] IEEE Transmitter. Improving the Quality of Sleep with AI and Machine Learning. 21 May 2021. Available online: <https://transmitter.ieee.org/improving-the-quality-of-sleep-with-ai-and-machine-learning> (accessed on 10 October 2022).
- [9] Imtiaz, S.A. A Systematic Review of Sensing Technologies for Wearable Sleep Staging. *Sensors* 2021, 21, 1562.
- [10] Imtiaz, S.A.; Rodriguez-Villegas, E. Automatic Sleep Staging Using State Machine-Controlled Decision Trees. In *Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Milan, Italy, 25–29 August 2015; IEEE: Manhattan, NY, USA, 2015; pp. 378–381.
- [11] Sen, B.; Peker, M.; Cavusoglu, A.; Celebi, F.V. A Comparative Study on Classification of Sleep Stage Based on EEG Signals Using Feature Selection and Classification Algorithms. *J. Med. Syst.* 2014, 38, 18.
- [12] Memar, P.; Faradji, F. A Novel Multi-Class EEG-Based Sleep Stage Classification System. *IEEE Trans. Neural Syst. Rehabil. Eng.* 2017, 26, 84–95.
- [13] Santaji, S.; Desai, V. Analysis of EEG Signal to Classify Sleep Stages Using Machine Learning. *Sleep Vigil.* 2020, 4, 145–152.
- [14] Bhusal, A.; Alsadoon, A.; Prasad, P.W.C.; Alsalami, N.; Rashid, T.A. Deep Learning for Sleep Stages Classification: Modified Rectified Linear Unit Activation Function and Modified Orthogonal Weight Initialisation. *Multimed. Tools Appl.* 2022, 81, 9855–9874.
- [15] Tao, Y.; Yang, Y.; Yang, P.; Nan, F.; Zhang, Y.; Rao, Y.; Du, F. A Novel Feature Relearning Method for Automatic Sleep Staging Based on Single-Channel EEG. *Complex Intell. Syst.* 2022.
- [16] Yulita, I.N.; Fanany, M.I.; Arymurthy, A.M. Sleep Stage Classification Using Convolutional Neural Networks and Bidirectional Long Short-Term Memory. In *Proceedings of the 2017 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, Bali, Indonesia, 5 June 2017; IEEE: Manhattan, NY, USA, 2017; pp. 303–308.
- [17] Goldberger, A.L.; Amaral, L.A.N.; Glass, L.; Hausdorff, J.M.; Ivanov, P.C.; Mark, R.G.; Mietus, J.E.; Moody, G.B.; Peng, C.-K.; Stanley, H.E. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation* 2000, 101, e215–e220.
- [18] Hasib Fardin, Hasan Md Imran, Hamdadur Rahman, Anamul Haque Sakib, Md Ismail Hossain Siddiqui. Robust and explainable poultry disease classification via MaxViT with attention-guided visualization. *International Journal of Science and Research Archive [Internet]*. 2025 Apr 30 [cited 2025 May 14];15(1):1848–59. Available from: <https://journalijsra.com/node/1054>
- [19] Mohammad Rasel Mahmud, Al Shahriar Uddin Khondakar Pranta, Anamul Haque Sakib, Abdullah Al Sakib, Md Ismail Hossain Siddiqui. Robust feature selection for improved sleep stage classification. *International Journal of Science and Research Archive [Internet]*. 2025 Apr 30 [cited 2025 May 14];15(1):1790–7. Available from: <https://journalijsra.com/node/1049>
- [20] Md Ismail Hossain Siddiqui, Anamul Haque Sakib, Amira Hossain, Hasib Fardin, Al Shahriar Uddin Khondakar Pranta. Custom CNN for acoustic emission classification in gas pipelines. *International Journal of Science and Research Archive [Internet]*. 2025 Apr 30 [cited 2025 May 14];15(1):1760–8. Available from: <https://journalijsra.com/node/1046>
- [21] Md Ariful Islam, Mohammad Rasel Mahmud, Anamul Haque Sakib, Md Ismail Hossain Siddiqui, Hasib Fardin. Time domain feature analysis for gas pipeline fault detection using LSTM. *International Journal of Science and Research Archive [Internet]*. 2025 Apr 30 [cited 2025 May 14];15(1):1769–77. Available from: <https://journalijsra.com/node/1047>
- [22] Mohammad Rasel Mahmud, Hasib Fardin, Md Ismail Hossain Siddiqui, Anamul Haque Sakib, Abdullah Al Sakib. Hybrid deep learning for interpretable lung cancer recognition across computed tomography and histopathological imaging modalities. *International Journal of Science and Research Archive [Internet]*. 2025 Apr 30 [cited 2025 May 14];15(1):1798–810. Available from: <https://journalijsra.com/node/1050>
- [23] Md Ismail Hossain Siddiqui, Anamul Haque Sakib, Sanjida Akter, Jesika Debnath, Mohammad Rasel Mahmud. Comparative analysis of traditional machine learning Vs deep learning for sleep stage classification. *International Journal of Science and Research Archive [Internet]*. 2025 Apr 30 [cited 2025 May 14];15(1):1778–89. Available from: <https://journalijsra.com/node/1048>

- [24] Sanjida Akter, Mohammad Rasel Mahmud, Md Ariful Islam, Md Ismail Hossain Siddiqui, Anamul Haque Sakib. Efficient and interpretable monkeypox detection using vision transformers with explainable visualizations. *International Journal of Science and Research Archive* [Internet]. 2025 Apr 30 [cited 2025 May 14];15(1):1811–22. Available from: <https://journalijsra.com/node/1051>
- [25] Anamul Haque Sakib, Md Ismail Hossain Siddiqui, Sanjida Akter, Abdullah Al Sakib, Mohammad Rasel Mahmud. LEVit-Skin: A balanced and interpretable transformer-CNN model for multi-class skin cancer diagnosis. *International Journal of Science and Research Archive* [Internet]. 2025 Apr 30 [cited 2025 May 14];15(1):1860–73. Available from: <https://journalijsra.com/node/1055>
- [26] Hamdadur Rahman, Hasan Md Imran, Amira Hossain, Md Ismail Hossain Siddiqui, Anamul Haque Sakib. Explainable vision transformers for real time chili and onion leaf disease identification and diagnosis. *International Journal of Science and Research Archive* [Internet]. 2025 Apr 30 [cited 2025 May 14];15(1):1823–33. Available from: <https://journalijsra.com/node/1052>
- [27] Siddiqui IH, Al Sakib A, Sakib AH, Fardin H, Debnath J. Dual-branch CrossViT for ovarian cancer diagnosis: Integrating and explainable AI for real-time clinical applications. *Article in International Journal of Science and Research Archive* [Internet]. 2025 [cited 2025 May 14];2025(01):1834–47. Available from: <https://doi.org/10.30574/ijrsra.2025.15.1.1164>
- [28] Haque R, Khan MA, Rahman H, Khan S, Siddiqui MIH, Limon ZH, et al. Explainable deep stacking ensemble model for accurate and transparent brain tumor diagnosis. *Comput Biol Med* [Internet]. 2025 Jun 1 [cited 2025 May 13];191:110166. Available from: <https://www.sciencedirect.com/science/article/pii/S0010482525005177>
- [29] Masum A Al, Limon ZH, Islam MA, Rahman MS, Khan M, Afridi SS, et al. Web Application-Based Enhanced Esophageal Disease Diagnosis in Low-Resource Settings. 2024 IEEE International Conference on Biomedical Engineering, Computer and Information Technology for Health (BECITHCON) [Internet]. 2024 Nov 28 [cited 2025 May 13];153–8. Available from: <https://ieeexplore.ieee.org/document/10962580/>
- [30] Haque R, Al Sakib A, Hossain MF, Islam F, Ibne Aziz F, Ahmed MR, et al. Advancing Early Leukemia Diagnostics: A Comprehensive Study Incorporating Image Processing and Transfer Learning. *BioMedInformatics 2024*, Vol 4, Pages 966-991 [Internet]. 2024 Apr 1 [cited 2025 May 13];4(2):966–91. Available from: <https://www.mdpi.com/2673-7426/4/2/54/htm>
- [31] Ahmed MR, Haque R, Rahman SMA, Afridi S, Abir MFF, Hossain MF, et al. Towards Automated Detection of Tomato Leaf Diseases. *Proceedings - 6th International Conference on Electrical Engineering and Information and Communication Technology, ICEEICT 2024*. 2024;387–92.
- [32] Al-Sakib A, Limon ZH, Sakib A, Pranto MN, Islam MA, Sultana S, et al. Robust Phishing URL Classification Using FastText Character Embeddings and Hybrid Deep Learning. 2024 IEEE 3rd International Conference on Robotics, Automation, Artificial-Intelligence and Internet-of-Things, RAAICON 2024 - Proceedings. 2024;53–8.
- [33] Noman A Al, Hossain A, Sakib A, Debnath J, Fardin H, Sakib A Al, et al. ViX-MangoEFormer: An Enhanced Vision Transformer-EfficientFormer and Stacking Ensemble Approach for Mango Leaf Disease Recognition with Explainable Artificial Intelligence. *Computers 2025*, Vol 14, Page 171 [Internet]. 2025 May 2 [cited 2025 May 13];14(5):171. Available from: <https://www.mdpi.com/2073-431X/14/5/171/htm>
- [34] Hosen MD, Bin Mohiuddin A, Sarker N, Sakib MS, Al Sakib A, Dip RH, et al. Parasitology Unveiled: Revolutionizing Microorganism Classification Through Deep Learning. *Proceedings - 6th International Conference on Electrical Engineering and Information and Communication Technology, ICEEICT 2024*. 2024;1163–8.
- [35] Al Noman A, Fardin H, Chhabra G, Sultana S, Haque R, Ahmed MR, et al. Monkeypox Lesion Classification: A Transfer Learning Approach for Early Diagnosis and Intervention. *Proceedings of International Conference on Contemporary Computing and Informatics, IC3I 2024*. 2024;247–54.
- [36] Hasan J, Hasan K, Al Noman A, Hasan S, Sultana S, Arafat MA, et al. Transforming Leukemia Classification: A Comprehensive Study on Deep Learning Models for Enhanced Diagnostic Accuracy. *PEEIACON 2024 - International Conference on Power, Electrical, Electronics and Industrial Applications*. 2024;266–71.
- [37] Sohaib M, Ghaffar A, Shin J, Hasan MJ, Suleman MT. Automated Analysis of Sleep Study Parameters Using Signal Processing and Artificial Intelligence. *Int J Environ Res Public Health* [Internet]. 2022 Oct 1 [cited 2025 May 23];19(20). Available from: <https://pubmed.ncbi.nlm.nih.gov/36293844>.
- [38] Sohaib M, Hasan MJ, Shah MA, Zheng Z. A robust self-supervised approach for fine-grained crack detection in concrete structures. *Scientific Reports 2024* 14:1 [Internet]. 2024 Jun 2 [cited 2025 Apr 17];14(1):1–20. Available from: <https://www.nature.com/articles/s41598-024-63575-x>.

- [39] Hasan MJ, Elyan E, Yan Y, Ren J, Sarker MMK. Segmentation Framework for Heat Loss Identification in Thermal Images: Empowering Scottish Retrofitting and Thermographic Survey Companies. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) [Internet]. 2024 [cited 2025 Feb 2];14374 LNAI:220–8. Available from: [https://link.springer.com/chapter/10.1007/978-981-97-1417-9\\_21](https://link.springer.com/chapter/10.1007/978-981-97-1417-9_21).
- [40] Sohaib M, Hasan MJ, Chen J, Zheng Z. Generalizing infrastructure inspection: step transfer learning aided extreme learning machine for automated crack detection in concrete structures. Meas Sci Technol [Internet]. 2024 Feb 21 [cited 2025 May 23];35(5):055402. Available from: <https://iopscience.iop.org/article/10.1088/1361-6501/ad296c>.