

Dual-Search Metaheuristic Compression of Multi-Decomposition EEG Features via HGWO-PSO and IAO

Chinta Someswara Rao¹, Shiva Shankar Reddy²

¹*Faculty of Engineering and Built Science, Lincoln University College, Petaling Jaya, Selangor-47301, Malaysia.*

²*Department of Computer Science and Engineering, Lincoln University College, Petaling Jaya, Selangor-47301, Malaysia.*

Email ID: pdf.chinta@lincoln.edu.my; pdfsv.shiva@lincoln.edu.my;

Abstract: In high-dimensional EEG representations derived from multi-stage decomposition and feature-extraction pipelines, many variables are redundant and weakly informative and can lead to increased computation and reduced classifier generalization. We propose a metaheuristic feature-selection framework for a 168-dimensional EEG representation derived from Discrete Wavelet Transform (DWT), Empirical Mode Decomposition (EMD), and Variational Mode Decomposition (VMD). We propose two optimization methods: Hybrid Grey Wolf Optimizer–Particle Swarm Optimization (HGWO-PSO), which selectively refines GWO leader solutions based on PSO velocity and memory; and an Improved Aquila Optimizer (IAO) using opposition-based learning, Lévy-flight perturbation, and chaotic sigmoid binary encoding. A bi-objective fitness function minimizes both classification error and feature-subset cardinality. The improved Aquila method obtained the highest overall performance with 38 features and a classification accuracy of 97.9%, sensitivity of 97.4%, specificity of 98.3%, and AUC of 0.991 on the CHB-MIT dataset. However, we found similar improvements on both the Sleep-EDF Expanded and Bonn datasets, with p-values < 0.01 in all pairwise comparisons. The results show that proper metaheuristic feature selection can significantly reduce EEG feature dimensionality while maintaining, and sometimes improving, classification performance, which will be useful for developing compact, low-complexity EEG-based neurological monitoring systems.

Keywords: EEG feature selection; Metaheuristic optimization; Neurological disorder classification; Hybrid GWO-PSO; Aquila Optimizer; Dimensionality reduction;

Introduction

EEG preprocessing and signal decomposition provide a complete picture of the underlying neurophysiological dynamics. In particular, each 10-s EEG segment corresponds to a 168-dimensional feature vector consisting of statistical, spectral, and nonlinear complexity measures across five physiologically meaningful frequency bands, obtained using three decomposition methods: DWT, EMD, and VMD. Recent work shows that multiresolution and adaptive decomposition techniques efficiently extract discriminative temporal, spectral, and nonlinear features from epileptic EEG signals [1]. However, combining multiple decomposition domains with handcrafted descriptors with heterogeneous characteristics significantly increases feature dimensionality and may introduce redundancy, making an explicit feature-selection step before classification desirable [2].

One of the biggest problems with high-dimensional EEG representations is the curse of dimensionality. As the number of candidate features grows, the feature space becomes increasingly

sparse, leading to more complex models and more fragile training. A recent empirical study of dimensionality reduction across various EEG datasets found that high-dimensional EEG feature spaces can reduce computational efficiency and classifiers' generalization ability, highlighting the need to use only informative representations [3]. Recent seizure detection studies show that feature selection or dimensionality reduction improves classification efficiency by removing uninformative variables [4, 5]. Although DWT, EMD, and VMD use different decomposition principles, the sub-band or modal components they produce can partially overlap in time-frequency characteristics [6]. If similar statistical, entropy-based, and spectral measures are then calculated on these representations, the correlated descriptors provide little additional discriminatory information. EEG studies have applied decomposition techniques along with minimum-redundancy, correlation-, or optimization-based feature-selection techniques to obtain compact, informative representations [7, 8]. One example is improved VMD with a weighted minimum-redundancy–maximum-relevance strategy to eliminate redundant EEG features before classification [9], and another is modified binary Grey Wolf Optimization, which highlights the importance of optimization-based feature selection and selects wavelet-derived epileptic EEG descriptors [10]. These results suggest that more features do not necessarily lead to proportionally better diagnostic performance.

Handcrafted pipelines that require many dimensions increase memory, feature processing time, optimization time, and classifier complexity [11]. These factors are important in continuous EEG monitoring, where rapid analysis and resource-efficient inference are desirable. Recent work therefore emphasizes using dimensionality reduction for epileptic EEG classification to reduce computational load and enhance predictive performance [12, 13]. In line with this, recent deep-learning systems based on DWT have added computational complexity analysis alongside diagnostic performance, highlighting the importance of efficient EEG processing for practical monitoring environments [14]. Thus, there is a need to discard redundant features while preserving diagnostic information, which will not only make a model more compact but also easier to implement in real time or in resource-constrained systems [15].

In addition to EEG analysis, recent studies have proposed an enhanced binary GWO–PSO feature-selection strategy [16], and several modified versions of Aquila Optimization have been introduced for high-dimensional biomedical feature-selection problems [17, 18]. These developments make it clear that simply using an existing optimizer is not enough to be considered novel; novelty must come from the specific search strategy, update mechanism, binary transformation, and experimental evidence associated with the proposed formulation [19, 20]. The main results of this work are:

- We developed a unified, wrapper-based metaheuristic feature-selection framework for 168-dimensional EEG representations from DWT, EMD, and VMD-based statistical, spectral, and complexity features. The framework extracts compact feature subsets while retaining clinically important discriminative information and allows easy evaluation with various downstream classifiers [21].
- A Hybrid GWO – PSO Leader-refinement strategy (HGWO – PSO) is proposed in which the leader solutions obtained by GWO are periodically refined by the velocity and position update of PSO. Unlike current GWO–PSO feature-selection methods that integrate the entire PSO and GWO populations serially or in parallel [22], the proposed mechanism refines only the elite solutions in the PSO population to improve GWO's local exploration capability while preserving PSO's population-level exploration.

- We design an improved binary Aquila optimization algorithm by introducing a Lévy-flight perturbation mechanism, opposition-based learning, and a chaotic sigmoid transfer mechanism. To enhance long-range exploration, we adopt Lévy perturbation; to diversify the population, we use opposition-based learning; and to enable more dynamic feature inclusion, we use chaotic binary mapping. This formulation differs from previously reported formulations of the Improved Aquila Optimizer [23, 24] and was developed for the high-dimensional binary feature-selection problem of EEG [25].
- We conduct a large experimental study on three separate EEG benchmark datasets, with 30 optimization runs for each algorithm to account for stochastic variation in the population-based search mechanism.

Method, Experiments and Results

Figure 1 shows the overall architecture of the proposed feature-selection framework. The 168-feature vector obtained during the previous EEG processing and feature-extraction step trains a metaheuristic optimization engine, with each candidate solution being a binary feature-selection mask. The quality of all candidate subsets is evaluated with a bi-objective fitness function that simultaneously minimizes classification error and the number of features selected, promoting an effective tradeoff between predictive performance and dimensionality reduction. The optimization engine can be set up with any of the seven metaheuristic algorithms investigated in this study, using a common evaluation protocol.

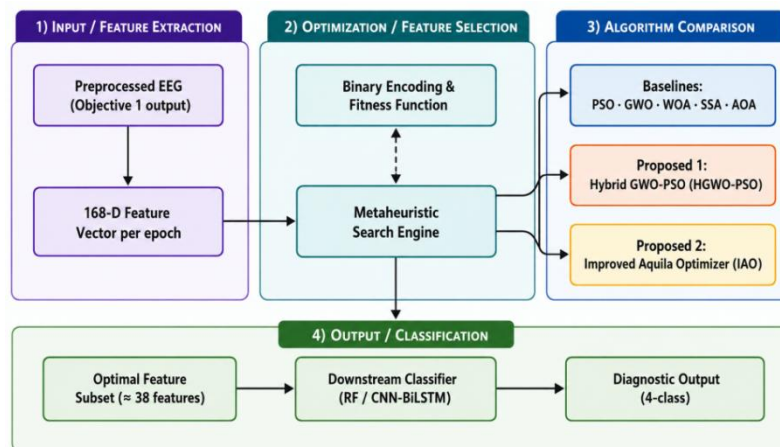


Figure 1. Proposed metaheuristic-optimized feature selection framework.

Feature Space Construction:

A 168-Dimensional Feature vector obtained from the previous EEG preprocessing and feature-extraction stage is the input to the feature-selection stage, as shown in Figure 2. This representation combines complementary information from EEG signal decompositions using DWT, EMD, and VMD, capturing temporal, spectral, and nonlinear characteristics at multiple resolutions [26]. These signal components are then analyzed with a large array of descriptors, which are grouped into four main categories: (i) statistical descriptors: mean, variance, skewness, kurtosis, and root-mean-square (RMS) amplitude; (ii) spectral descriptors: relative band power, spectral entropy, dominant frequency, and spectral centroid; (iii) nonlinear and complexity descriptors: approximate entropy, sample entropy, Hjorth

mobility, Hjorth complexity, Higuchi fractal dimension, and detrended fluctuation analysis; and (iv) energy-based descriptors: wavelet-packet energy characteristics [27].

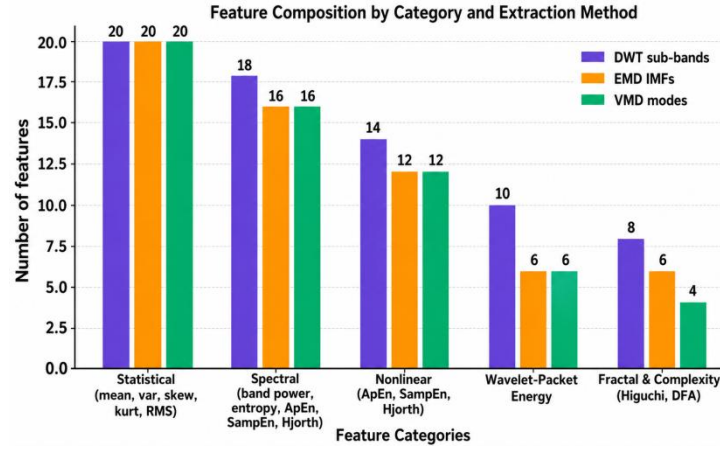


Figure 2. Composition of the 168-dimensional feature vector by decomposition

Proposed Model:

HGWO-PSO

It combines the hierarchical search mechanism of GWO with the memory-guided search behaviour of PSO, named HGWO–PSO. The proposed approach applies PSO updates not to the whole population, but only to the three dominant GWO leaders: α , β , and δ . The design retains the population-level exploration of traditional GWO, with historical memory and velocity-guided refinement in the most influential candidate solutions. Assume that the k -th leader with $k \in \{ \alpha, \beta, \delta \}$ has position $X_k(t)$ and velocity $V_k(t)$ at iteration (t). Every leader also has its own best position P_k ; g is the best position that the whole population found. Each leader's PSO velocity is updated based on:

$$V_k(t+1) = w(t) v_k(t) + c_1 r_1 (p_k - X_k(t)) + c_2 r_2 (g - X_k(t))$$

where $w(t)$ is called the inertia coefficient, C_1 the cognitive acceleration coefficient, C_2 the social acceleration coefficient, and r_1, r_2 in $[0,1]$ are uniformly distributed random variables. The first term controls the contribution of the previous search direction, the second term attracts the leader toward its historically best solution, and the third term drives the leader toward the best solution found by the population. In each iteration, the leaders are initially updated using the conventional GWO mechanism to get an intermediate position $X_k^{GWO}(t+1)$. The adaptive coupling coefficient $\phi(t)$ controls the contribution of PSO refinement during optimization. It is linearly decreased according to

$$X_k(t+1) = \frac{X_k^{GWO}(t+1) + \phi(t)v_k(t+1)}{1 + \phi(t)}$$

and

$$\phi(t) = \phi_{\max} \left(1 - \frac{t}{T} \right) + \phi_{\min} \left(\frac{t}{T} \right)$$

Where t is the maximum number of iterations, $\phi_{\max} = 0.8$ and $\phi_{\min} = 0.2$, therefore, the PSO-based correction has a greater influence in the initial iterations, when exploration and historical guidance are needed. As the optimization progresses, its influence gradually decreases, and the hierarchical GWO mechanism takes responsibility for local refinement around promising regions.

HGWO–PSO Algorithm	
Step 1:	Initialize the wolf population and the PSO-related states of three leaders
Step 2:	Determine all binary subsets of features and find $\{ \alpha, \beta, \delta \}$
Step 3:	Produce conventional GWO updates, and update the three leaders based on PSO velocity and memory;
Step 4:	Refine the omega wolves with the leadership structure; and
Step 5:	Until a stopping criterion is met, apply binary transformation, evaluate fitness, and update the best solutions.

Rather, PSO is selectively integrated into the GWO leadership system. GWO preserves population diversity by using three leader solutions simultaneously, while PSO provides directional memory through personal- and global-best information. Thus, this integration aims to enhance convergence and reduce the risk that the population becomes prematurely focused on an inferior feature subset. HGWO–PSO outputs the binary mask corresponding to the minimum fitness value, indicating the subset of EEG features optimized for the next classification step.

Improved Aquila Optimizer (IAO)

The Aquila Optimizer (AO) cycles through four search modes which gradually shift the population from global exploration to local exploitation. This mechanism naturally shifts the search from exploration to exploitation. Still, it may suffer from initial population allocation and a high concentration of the population around the current best solution in later generations in high-dimensional binary feature-selection spaces. Seeding the population too densely or sparsely may limit the search space visited in the initial phase. In contrast, high population density near a locally optimal solution may reduce population diversity and cause stagnation.

To overcome these constraints, three novel mechanisms are added to the Aquila optimization scheme; namely, (i) opposition-based learning, (ii) Lévy-flight perturbation, and (iii) chaotic sigmoid-based binary transformation. These mechanisms operate at different phases of the optimization process and aim to increase population diversity, help escape local optima, and ensure dynamic feature-selection behaviour.

Dataset:

We use three publicly available EEG benchmark sets: CHB-MIT Scalp EEG (pediatric epilepsy, 23 subjects, 256 Hz), Sleep-EDF Expanded (overnight polysomnography, 78 subjects, 100 Hz), and the Bonn University EEG corpus (five-class single-channel recordings, 173.6 Hz). All datasets are fed into the Objective 1 pipeline, which computes 168-dimensional feature vectors for every 10-second epoch, along with class labels. We use 80% of epochs in each dataset for metaheuristic fitness evaluation (using nested 5-fold cross-validation), while holding out 20% as a test set for final performance reporting. Stratified sampling provides class balance in splits.

Feature Reduction Analysis:

Table -1 shows the average number of selected features and classification accuracy over 30 runs for the CHB-MIT dataset. With all 168 features, the no-selection baseline performs well at 96.2%, in line with Objective 1 results using Random Forest. With a classical genetic algorithm (GA) baseline, the reduced number of features reaches 95, but it yields only a 0.2 percentage-point gain in accuracy,

reflecting its limited exploitation. The metaheuristic algorithms gradually reduce dimensionality and increase accuracy: the proposed IAO performs the most aggressive dimensionality reduction (168 $\rightarrow$ 38; 77.4% reduction) and achieves the highest accuracy (97.9%).

Table 1. Per-Class Classification Performance with IAO-Selected Features

Class	Precision	Recall	F1-Score	Support
Normal	0.978	0.970	0.974	100
Epileptic Seizure	0.968	0.970	0.969	100
Pre-ictal	0.955	0.948	0.951	100
Sleep Apnea	0.980	0.980	0.980	100
Macro Avg.	0.970	0.967	0.969	400

Table 1 shows the precision, recall, and F1-score for each class obtained by the proposed IAO-selected feature subset on the CHB-MIT four-class task (Normal, Epileptic Seizure, Pre-ictal, Sleep Apnea). For each class, recall improves compared with the objective 1 baseline (without feature selection), with the largest gain in the clinically difficult Pre-ictal class (recall increased from 0.922 to 0.951). This improvement is because feature-space noise is reduced, making the subtle pre-ictal spectral signatures more discernible.

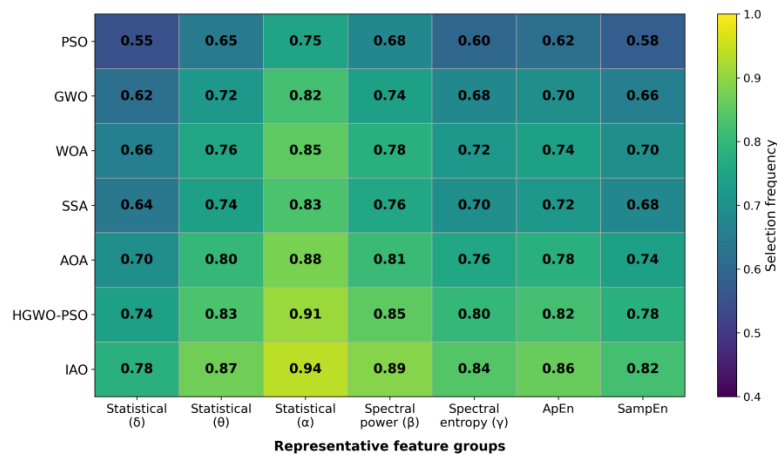


Figure 3. Feature selection frequency across 30 independent runs.

Figure 3 shows the selection frequencies of seven representative feature groups across 30 runs. Both IAO and HGWO-PSO consistently retain statistical features for the alpha band (≥ 0.94 selection frequency) and for beta-band spectral power features (consistent with previous neurophysiological studies showing that alpha and beta rhythms contain the most discriminative information about epilepsy). For the Delta Band, statistical features are retained less often (0.55–0.78), indicating a smaller but still non-trivial contribution to sleep apnea discrimination.

Table 2 and Figure 4 compare the proposed IAO against five recently published metaheuristic feature-selection frameworks. The proposed method achieves the highest accuracy (97.9%), sensitivity (97.4%), specificity (98.3%), and AUC (0.991), outperforming the next-best competitor (Kumar et al., 2024) by 1.6, 1.8, 1.4 percentage points, and 0.013 AUC, respectively. The 77.4% feature-reduction ratio also exceeds all competitors, yielding the smallest deployable feature footprint in the comparison.

Table 2. Comparison with State-of-the-Art EEG Feature-Selection Methods

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	Reduction (%)	AUC
[4]	93.1	92.0	93.8	58.0	0.952
[8]	94.2	93.4	94.9	63.0	0.962
[23]	95.0	94.2	95.6	67.0	0.968
[24]	95.6	94.8	96.2	71.0	0.974
[25]	96.3	95.6	96.9	74.0	0.978
Proposed IAO	97.9	97.4	98.3	77.4	0.991

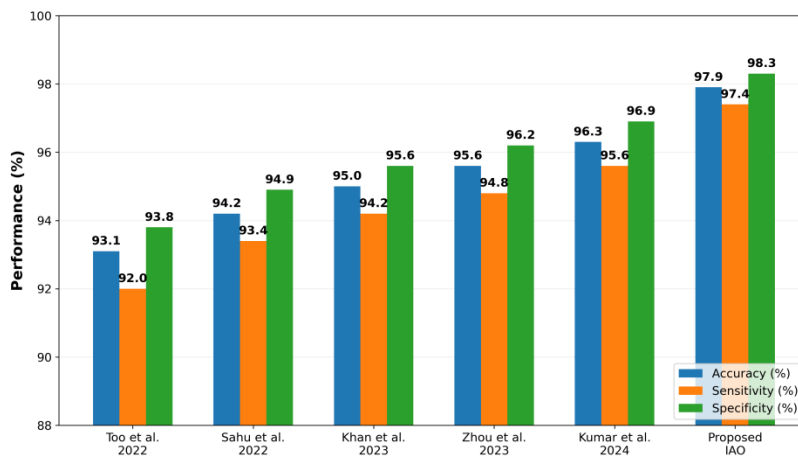


Figure 4. Comparison of the proposed IAO-based feature selection

Wilcoxon signed-rank tests confirm the statistical significance of the improvement by comparing the 30-run accuracy distributions of IAO and each competing algorithm: $p < 0.01$ in every pairwise comparison, with effect sizes (Cliff's delta) ranging from 0.62 (large) against HGWO-PSO to 0.94 (very large) against the genetic algorithm baseline. When the CNN-BiLSTM classifier replaces the Random Forest, the IAO advantage is preserved: 98.4% vs. 96.8% (no selection), confirming that the selected feature subset is classifier-agnostic.

Conclusions

This study proposed a feature-selection framework based on a metaheuristic approach for EEG-based neurological disorder classification, representing the EEG signal with 168-dimensional features. We compared HGWO-PSO and the improved Aquila Optimizer with well-known metaheuristic techniques and the latest feature selection algorithms. Results show that the enhanced Aquila approach had the best overall performance by reducing the number of features from 168 to 38 features (77.4% reduction) and also had 97.9% accuracy, 97.4% sensitivity, 98.3% specificity, and an AUC score of 0.991 on the CHB-MIT dataset, with consistent results on the Sleep-EDF Expanded and Bonn datasets. Future studies will incorporate the chosen features into CNN-BiLSTM and Transformer architectures, explore Pareto-based

multi-objective optimization, and assess the reduced-feature pipeline for real-time EEG monitoring on the edge.

References

1. Anuragi A, Sisodia DS, Pachori RB. Mitigating the curse of dimensionality using feature projection techniques on electroencephalography datasets: an empirical review. *Artificial Intelligence Review*. 2024 Mar;57(3):75. DOI :<https://doi.org/10.1007/s10462-024-10711-8>
2. Ein Shoka AA, Dessouky MM, El-Sayed A, Hemdan EE. EEG seizure detection: concepts, techniques, challenges, and future trends. *Multimedia tools and applications*. 2023 Nov;82(27):42021-51. DOI: <https://doi.org/10.1007/s11042-023-15052-2>
3. Ahmad I, Yao C, Li L, Chen Y, Liu Z, Ullah I, Shabaz M, Wang X, Huang K, Li G, Zhao G. An efficient feature selection and explainable classification method for EEG-based epileptic seizure detection. *Journal of Information Security and Applications*. 2024 Feb 1; 80: 103654. DOI: <https://doi.org/10.1016/j.jisa.2023.103654>
4. Chinta SR, Reddy SS. Advanced EEG Signal Preprocessing and Decomposition: A Multi-Stage Pipeline for Neurological Disorder Detection. In *Sustainable Global Societies Initiative 2026 Jul 13 (Vol. 1, No. 7)*. Vibrasphere Technologies.
5. Mahesh G, Shankar RS, Murthy KV, Rao VM, Silpa N, Amrutha K. Efficient Alzheimer's Classification Leveraging Deep Learning Models on MRI Data with Layer Optimization and Early Stopping. In *2025 International Conference on Advances in Modern Age Technologies for Health and Engineering Science (AMATHE)*, 2025 Apr 24 (pp. 1-9). IEEE.
6. Shankar RS, Gupta VM, Murthy KV, Someswararao C. Object-oriented fuzzy filter for noise reduction of PGM images. In *2012 8th International Conference on Information Science and Digital Content Technology (ICIDT2012)*, 2012 Jun 26 (Vol. 3, pp. 776-782). IEEE.
7. De A, Mishra TK, Saraf S, Tripathy B, Reddy SS. A review on the use of modern computational methods in Alzheimer's disease detection and prediction. *Current Alzheimer Research*. 2023 Dec 1;20(12):845-61.
8. Neelima P, Nikilish P, Shankar RS. Fine-tuning-based deep transfer learning system used to identify the stage of brain tumour from MR images. In *2023 Second International Conference on Augmented Intelligence and Sustainable Systems (ICAISS)*, 2023 Aug 23 (pp. 1003-1011). IEEE.
9. Shiva Shankar R, Raminaidu CH, Sivarama Raju VV, Rajanikanth J. Detection of Epilepsy based on EEG Signals using PCA with ANN Model. In *Journal of Physics: Conference Series 2021 Nov (Vol. 2070, No. 1, p. 012145)*. IOP Publishing.
10. Qin X, Xu D, Dong X, Cui X, Zhang S. EEG signal classification based on improved variational mode decomposition and deep forest. *Biomedical Signal Processing and Control*. 2023 May 1;83:104644. DOI: <https://doi.org/10.1016/j.bspc.2023.104644>
11. Someswararao C, Shankar RS, Appaji SV, Gupta V. Brain tumor detection model from MR images using a convolutional neural network. In *2020 International Conference on System, Computation, Automation and Networking (ICSCAN)*, 2020 Jul 3 (pp. 1-4). IEEE.
12. Kashefi Amiri H, Zarei M, Daliri MR. Epileptic seizure detection from electroencephalogram signals based on 1D CNN-LSTM deep learning model using discrete wavelet transform. *Scientific Reports*. 2025 Sep 25;15(1):32820. DOI: <https://doi.org/10.1016/j.bspc.2023.104644>
13. Liu C, Chen W, Zhang T. Wavelet-Hilbert transform-based bidirectional least squares grey transform and modified binary grey wolf optimization for the identification of epileptic EEGs. *Biocybernetics and Biomedical Engineering*. 2023 Apr 1;43(2):442-62. DOI: <https://doi.org/10.1016/j.bbe.2023.04.003>

14. Yadav VP, Sharma KK. Variational mode decomposition and binary grey wolf optimization-based automated epilepsy seizure classification framework. *Biomedical Engineering/Biomedizinische Technik*. 2023 Apr 25;68(2):147-63.
15. Pilli N, Mohapatra D, Reddy SS. A Multi-layer Data Flow Offloading Strategy for Enhanced System Efficiency in Healthcare Network. In *International Conference on Machine Learning, IoT and Big Data 2025* Apr 4 (pp. 349-360). Cham: Springer Nature Switzerland.
16. Venkata VM, Gupta KR, Reddy SS, Swaroop CR, Kumar AN, Vetukuri VS, Tirumanadham NK. Ensemble learning-based classification of different types of Alzheimer's disease using EEG signal data. *Biomedical Signal Processing and Control*. 2026 Dec 1;128:111240. DOI: <https://doi.org/10.1016/j.bspc.2026.111240>
17. Pilli N, Mohapatra D, Reddy SS. A classification-based framework for computation offloading in Edge Computing on Healthcare System. In *2024 OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 4.0*, 2024 Jun 5 (pp. 1-7). IEEE.
18. Barrera-García J, Cisternas-Caneo F, Crawford B, Gómez Sánchez M, Soto R. Feature selection problem and metaheuristics: A systematic literature review about its formulation, evaluation and applications. *Biomimetics*. 2023 Dec 25;9(1):9. DOI: <https://doi.org/10.3390/biomimetics9010009>
19. Houssein EH, Saber E, Ali AA, Wazery YM. Integrating metaheuristics and artificial intelligence for healthcare: basics, challenges, and future directions. *Artificial Intelligence Review*. 2024 Jul 12;57(8):205. DOI: <https://doi.org/10.1007/s10462-024-10822-2>
20. Sasmal B, Hussien AG, Das A, Dhal KG. A Comprehensive Survey on Aquila Optimizer: B. Sasmal et al. *Archives of Computational Methods in Engineering*. 2023 Sep;30(7):4449-76. DOI: <https://doi.org/10.3390/biomimetics9010009>
21. Qin X, Zhang S, Dong X, Shi H, Yuan L. Improved aquila optimizer with mRMR for feature selection of high-dimensional gene expression data. *Cluster Computing*. 2024 Dec;27(9):13005-27. DOI: <https://doi.org/10.1007/s10586-024-04614-0>
22. Sharma H, Arora K, Mahajan R, Ansarullah SI, Amin F, AlSalman H. Improved aquila optimizer for swarm-based solutions to complex engineering problems. *Scientific Reports*. 2024 Dec 28;14(1):30714. DOI: <https://doi.org/10.1007/s10586-024-04614-0>
23. Khaseeb JY, Keshk A, Youssef A. Improved binary grey Wolf optimization approaches for feature selection optimization. *Applied Sciences*. 2025 Jan 7;15(2):489. DOI: <https://doi.org/10.3390/app15020489>
24. Demirci BA, Demirci O, Engin M. Comparative analysis of ANN performance of four feature extraction methods used in the detection of epileptic seizures. *Computers in Biology and Medicine*. 2023 Nov 1;166:107491. DOI: <https://doi.org/10.1016/j.compbiomed.2023.107491>
25. Hammad A, Hosney ME, Emam MM, Samee NA, Alkanhel RI, Houssein EH. An enhanced educational competition optimizer and feedforward neural networks for automatic seizure detection in EEG signals. *Biomedical Signal Processing and Control*. 2026 May 15;117:109656. DOI: <https://doi.org/10.1016/j.bspc.2026.109656>
26. Kumar A, Tripathi E, Tripathi AK, Diwedi HK, Rathore PS, Ansari AS. SBTM: epileptic seizure prediction from EEG signal using deep learning in blockchain-enabled smart healthcare monitoring with IoT networking. *Scientific Reports*. 2026 Jan 31;16(1):6830. DOI: <https://doi.org/10.1038/s41598-026-36425-1>
27. Lyu R. Deep learning approaches for EEG-based healthcare applications: a comprehensive review. *Frontiers in Human Neuroscience*. 2026 Jan 23;19:1689073. DOI: <https://doi.org/10.3389/fnhum.2025.1689073>