<u>Supplementary Materials</u>

1) Significance of Feature Extraction Methods

i. Time-Based Features:

Time-based attributes are desirable for preserving the time-locked amplitude and phase information required for event-related potentials (ERP) EEG signals processing. Besides, the time-domain hallmarks are computationally efficient (without requiring any prior transformation) and precisely retain the temporal resolution of EEG signals, which is critical for accelerated BCI and epilepsy diagnosis systems [1], [2]. The time-based characteristics are selected owing to their utilities in classifying motor imagery [3], mental imagery [3], epilepsy [4], and schizophrenia [5] EEG signals.

ii. Entropy-Based Features:

The EEG is a class of non-stationary signals that fluctuate erratically in response to a prevailing brain function. It is generally known that the diseased brain operations cause a different degree of randomness in EEG signals than healthy brain processes [6]. Entropy-based features are vital to harness the magnitude of randomness in a time-varying signal. Such a property could be helpful in the classification of epilepsy and schizophrenia EEG signals since the presence of a diseased mental state tends to a unique degree of disorder that could be robustly encapsulated with an entropy measure.

iii. Geometrical-Based Features:

The geometric properties of time-varying EEG signals are essential to assess their complexity and variability. For example, epileptic EEG signals are more complex and show a significant variance than the non-epileptic ones. This is because the synapses are more synchronized during the non-ictal activity than during the ictal durations, where synapses are primarily irregular, and the accompanying EEG signal complexity is enhanced. This study added fourteen graphical attributes to the collection of multidomain features based on their importance in classifying epilepsy and schizophrenia EEG signals [7].

iv. Frequency-Based Features:

Spectral characteristics are conventional and effective method for studying motor and mental imagery EEG signals. The transition in mental simulation of a particular task involves the shift in spectral power in specific frequency rhythms (ERD/ERS in μ and β bands). Such a considerable variation in these rhythms presents a useful indicator to anticipate the motor imagery and mental imagery task labels. This work availed Welch power spectral density (PSD) method [8] for frequency-domain feature extraction employing a Hamming window of size 50 with 50% intersection among successive segments.

2) Significance of Feature Selectors:

A brief description and significance of the five optimization techniques is given as follows:

i. ***Marine Predator Algorithm (MPA)*** is a nature-inspired metaheuristic search algorithm that utilizes Levy and Brownian motion foraging strategies adopted by marine predators [9]. The MPA replicates the prey hunting practice embraced by marine predators, where the hunter selects the ideal foraging plan of action depending on the probability of encountering prey. The MPA is favorable for a multidomain feature selection problem due to its minimal parametric dependency, quick computation speed, and high computational accuracy [10].

ii. ***Generalized Normal Distribution Algorithm (GNDA)*** is a metaheuristic search approach based on the Gaussian distribution model. The GNDA is unique in such a way that it does not need initial parameter adjustment and correctly extracts unknown parameters from multi-objective problems [11]. Its primary benefits in multidomain classification include computational speed, adaptability to multivariate distributions, and scaling to noisy data [12].

iii. ***Slime Mould Algorithm (SMA)*** simulates the behavior and morphological changes of slime mould. In the SMA model, variables are weighted according to their contribution to the global optimal solution. It is adaptive to food concentration (interclass separability), where the global solution constantly travels towards the variables with the largest nutritional concentration [13]. The primary advantages of using SMA for a multidomain feature selection model include a bigger search area for distinct variables, parallel processing, numerous optimum solutions, and less information for generalization [14].

iv. ***Equilibrium Optimizer (EO)*** is a physics-based algorithm inspired by the control volume mass balance model to search for an equilibrium state of particles [15]. The com position of EO involves high exploitative and exploratory foraging mechanisms to generate random solutions. The dominant traits of EO are less computational complexity, less prone to outliers, less dependence over the initial points, and faster convergence within a shorter calculation time [16].

v. ***Manta Ray Foraging Optimization (MRFO)*** is a biologically inspired metaheuristic optimization strategy that imitates the foraging behavior of manta rays, such as chain foraging, cyclone foraging, and somersault foraging, to find an optimal solution for a given problem [17]. The MRFO has a close resemblance to the particle swarm optimization (PSO), shark smell optimization (SSO), and Greywolf optimization (GWO), except that the MRFA has fewer initial parameters dependency, supports parallelism, fast convergence speed, and multi-objective optimization [17], which makes it a suitable choice for multidomain feature selection.

The choice of the aforementioned feature selectors is based on their computational efficiency, high computational accuracy, and resilience to adapt to multivariate distributions. As per the referenced studies, each feature selector showcased significant improvement over the conventional optimization-based methods like genetic algorithm (GA), particles swarm optimization (PSO), ant colony optimization (ACO), binary grey wolf (BGW), etc.

3) T-Distributed Stochastic Neighbor Embedding (t-SNE) Method:

The t-SNE is a statistical dimensionality reduction (DR) method for data visualization [18]. The t-SNE minimizes the disparity between high and low-dimensional joint data distributions while maintaining Euclidean linkages. The t-SNE approach has several benefits over other DR methods. *First*, it performs well with linear and non-linear data, while the traditional principal component analysis (PCA) approach creates principal components as a linear mixture of characteristics. As a result, it can't comprehend complex polynomial relationships in data.

Second, the t-SNE method is capable of conserving the structural integrity of the data. This indicates that points adjacent in the high-dimensional data will likely be adjacent in the low dimension. Contrarily, PCA discovers new dimensions that account for most of the variance in the data. Therefore, unlike t-SNE, PCA is unconcerned with its immediate surroundings. *Third*, the t-SNE projects the high-dimensional data into two-dimensional space with maximum variance. Conversely, the other DR methods transform the features into a multidimensional subspace that maximizes the interclass variance. Computing the Euclidean distance among interclass features in a two-dimensional plane is less complex than in a multidimensional space. Owing to the aforesaid virtues, the t-SNE method has been exploited and moulded into a sample reduction technique.

4) References

[1]. Y. Li et al., “Clustering technique-based least square support vector machine for EEG signal classification,” *Comput. Methods Prog. Biomed*., vol. 104, no. 3, pp. 358–372, 2011.

[2]. S. Siuly and Y. Li, “Improving the separability of motor imagery EEG signals using a cross correlation-based least square support vector machine for brain–computer interface,” *IEEE Trans. Neural Syst. Rehabil. Eng*., vol. 20, no. 4, pp. 526–538, Jul. 2012.

[3]. X.Yu, M. Z. Aziz, M. T. Sadiq, Z. Fan, and G. Xiao, “A new framework for automatic detection of motor and mental imagery EEG signals for robust BCI systems,” *IEEE Trans. Instrum. Meas*., vol. 70, pp. 1–12, 2021.

[4]. M. M. or Rashid and M. Ahmad, “Epileptic seizure classification using statistical features of EEG signal,” in *Proc. IEEE Int. Conf. Elect., Comput. Commun. Eng*., 2017, pp. 308–312.

[5]. S. Siuly, S. K. Khare, V. Bajaj, H. Wang, and Y. Zhang, “A computerized method for automatic detection of schizophrenia using EEG signals,” *IEEE Trans. Neural Syst. Rehabil. Eng*., vol. 28, no. 11, pp. 2390–2400, Nov. 2020.

[6]. N.Kannathal,M. L. Choo,U. R.Acharya, and P. Sadasivan, “Entropies for detection of epilepsy in EEG,” *Comput. Methods Prog. Biomed*., vol. 80, no. 3, pp. 187–194, 2005.

[7]. H. Akbari, S. Ghofrani, P. Zakalvand, and M. T. Sadiq, “Schizophrenia recognition based on the phase space dynamic of EEG signals and graphical features,” *Biomed. Signal Process. Control*, vol. 69, 2021, Art. no. 102917.

[8]. Gupta et al., “On the utility of power spectral techniques with feature selection techniques for effective mental task classification in noninvasive BCI,” *IEEE Trans. Syst., Man, Cybern.: Syst*., vol. 51, no. 5, pp. 3080–3092, May 2021.

[9]. Faramarzi, M. Heidarinejad, S. Mirjalili, and A. H. Gandomi, “Marine predators algorithm: A nature-inspired metaheuristic,” *Expert Syst. Appl*., vol. 152, 2020, Art. no. 113377.

[10]. K. Zhong, G. Zhou, W. Deng, Y. Zhou, and Q. Luo, “MOMPA: Multi-objective marine predator algorithm,” *Comput. Methods Appl. Mech. Eng*., vol. 385, 2021, Art. no. 114029.

[11]. M. Chankaya, I. Hussain, A. Ahmad, H. Malik, and F. P. García Márquez, “Generalized normal distribution algorithm-based control of 3-phase 4-wire grid-tied PV-hybrid energy storage system,” *Energies*, vol. 14, no. 14, 2021, Art. no. 4355

[12]. Y. Zhang, Z. Jin, and S. Mirjalili, “Generalized normal distribution optimization and its applications in parameter extraction of photovoltaic models,” *Energy Convers. Manage*., vol. 224, 2020, Art. no. 113301.

[13]. T. Latty and M. Beekman, “Food quality and the risk of light exposure affect patch-choice decisions in the slime mold Physarum polycephalum,” *Ecology*, vol. 91, no. 1, pp. 22–27, 2010.

[14]. T. Latty and M. Beekman, “Speed–accuracy trade-offs during foraging decisions in the acellular slime mould Physarum polycephalum,” *Proc. Roy. Soc. B: Biol. Sci*., vol. 278, no. 1705, pp. 539–545, 2011.

[15]. Faramarzi, M. Heidarinejad, B. Stephens, and S.Mirjalili, “Equilibrium optimizer: A novel optimization algorithm,” *Knowl.-Based Syst*., vol. 191, 2020, Art. no. 105190.

[16]. Y. Gao, Y. Zhou, and Q. Luo, “An efficient binary equilibrium optimizer algorithm for feature selection,” *IEEE Access*, vol. 8, pp. 140936–140963, 2020

[17]. W. Zhao, Z. Zhang, and L. Wang, “Manta ray foraging optimization: An effective bio-inspired optimizer for engineering applications,” *Eng. Appl. Artif. Intell*., vol. 87, 2020, Art. no. 103300.

[18]. M.Wattenberg, F. Viégas, and I. Johnson, "How to use t-SNE effectively," *Distill*, vol. 1, no. 10, 2016, Art. no. e2.