Received: December 17, 2023. Revised: February 19, 2024.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Classification of Epileptic Seizures Using LSTM Based Zebra Optimization Algorithm with Hyperparameter Tuning

T. Jhansi Rani[1]* **D. Kavitha[2]**

[1]*Department of Computer Science & Engineering, Jawaharlal Nehru Technological University Anantapur, Anantapur, Andhra Pradesh, India*
[2]*Department of Computer Science & Engineering, G. Pulla Reddy Engineering College (Autonomous), Kurnool, Andhra Pradesh, India*
* Corresponding author's Email: satyajhansii@gmail.com

Abstract: Electroencephalogram (EEG) are the neuro-electrophysiology signals, which are commonly used as a diagnostic tool to measure the seizure activity of the brain. The accurate detection and classification of seizures help to provide an optimal solution to diagnose the patient. In this research, a hyperparameter tuning with Zebra Optimization Algorithm (ZOA) is proposed for the fine tuning of features from EEG signals. The EEG signals are taken from the three standard datasets such as Temple University Hospital (TUH) at a rate of sampling signal of 250Hz, Bonn University (BU) at a 173.61 Hz sampling rate, and Bern Barcelona (BB) alongside the sampling frequency of 512 Hz. The EEG signals are pre-processed using Butterworth 8th order filtering method to remove the unwanted noise, and de-noised signals are decomposed by the swarm decomposition method. Features like statistic-based features, frequency-dependent features, multi-scale wavelet transformation, entropy features, and power spectral features are extracted from the decomposition of signals. The extracted features then undergo hyperparameter tuning using ZOA followed by feature selection using Enhanced Spatial bound Whale Optimization Algorithm (WOA) with the combination of Salp Swarm Algorithm (SSA) hybridized with Lens Opposition-based Learning (LOBL) mechanism. The features obtained from the selection algorithm are then fed to hyper parameter optimized Long Short-Term Memory (LSTM) classifier to classify the normality and abnormality of seizures. The attained outcomes of the suggested approach exhibit a better classification rate with 98.43% accuracy on BU dataset, 99.71% accuracy on BB dataset, and 98.43% accuracy on TUH dataset.

Keywords: Electroencephalogram, Enhanced spatial bound whale optimization algorithm, Lens opposition-based learning, Long short-term memory, Salp swarm algorithm, Zebra optimization algorithm.

1. Introduction

Epilepsy is a chronological brain disorder that is caused by the sudden abnormal synchronous discharge behavior in the brain cells. It is the most widespread disease among children and adults after a stroke [1]. Epilepsy causes a lack of consciousness and affects the psychological and neurological condition of a person. Seizures are known to be the temporary interruption of brain functions that can be triggered by abnormal discharge of electrical neurons [2,3]. The brain activities are recorded by the neuro electrophysiologic signals like Electroencephalography (EEG), Magnetoencephalography (MEG), and operational Magnetic Resonance Imaging (MRI) through which the diagnosis and monitoring of seizures can be analyzed [4]. Based on the area of the brain that is stimulated during seizures, epilepsy is often divided into two important classes: partial and generalized. The generalized one began in the entire brain as opposed to the partial one, which originates from a specific area of the brain and lies on one side of the cerebral column [5, 6]. In general, abnormal seizures are complex to identify as it occurs in the partial part of the brain and normal seizures occur in the entire brain. The EEG signals have high frequency waves, where the fluctuations of these waves determine the nature of seizures (normal or abnormal) and the condition of the patient [7,8]. Several Machine Learning (ML) and Deep Learning (DL) methods have been improved to classify as well as determine the seizure types, but are limited to high computational complexity, vanishing gradient, classification errors, and time complexity [9]. Early detection and diagnosis increase the survival rate of patients suffering from epilepsy seizures. The EEG signals from standard datasets consist of various frequency bands. Hence to maintain uniform sensitivity, the signals are decomposed using decomposition methods [10-12]. The decomposed signals are then extracted for generating optimal features to select discriminative features. The selected features are then classified for predicting the normality and abnormality conditions of epilepsy seizures [13]. However certain limitations such as computational and time complexity, overfitting of data, and classification errors of existing methods have left scope for further investigations of seizure detection and classification [14, 15]. The major contributions of this research are described below:

- Bio-electrical artifacts which are not necessary for analysis are to be removed and the external interpretation analysis of EEG signals which are collected has to be done with the help of a Butterworth 8th-order filter. It has the advantage of roll-off around the cut-off frequency and maintains the uniform sensitivity of wanted signals.
- A decay method related to swarm intelligence is used to decompose the denoised signals which reflects the frequency variations in the signals. the set of optimal features is then extracted such as statistic-based features, frequency-dependent features, features derived from multi-scale wavelet transformation, features concerning entropy measure, and power spectral density.
- The obtained features are then fine-tuned with hyper parameter tuning process using ZOA. These optimized features are then selected using WOA with SSA to achieve balance in the process of the exploitation stage and exploration stage of the algorithm and also utilized LOBL strategy to avoid the tendency to fall into local optimum.
- The optimal features are then fed into LSTM to categorize standard and abnormal seizures. The LSTM with a gating mechanism is used in this research, which has the advantage of avoiding the vanishing gradient problem.

The remaining article is coordinated as follows: Section 2 offers the literature survey on existing methods, Section 3 briefs the proposed methodology, Section 4 represents the output results and discussions, and the conclusion of this research work is summarized in section.

2. Literature review

Albaqami [16] implemented a multiclass seizure classification system using dual-tree complex wavelet transform and ML on TUH dataset, where the features of these signals were extracted using Wavelet Transform. The classification of these features was performed using Light Gradient Boosting Machine (LGBM) to classify seven types of seizures and attained better classification results when compared to existing wavelet-based seizure classification methods. However, the tree-leaf-wise split in LGBM causes overfitting of the data.

Chirasani and Manikandan [17] presented a Deep Neural Network (DNN) for classifying epileptic seizures utilizing a hierarchical attention system to overcome the computational complexity of existing techniques. The classification was performed on raw EEG signals and the features of the signal were weighted utilizing hierarchical attention system. These features were classified using SVM classifier and 1D CNN. The obtained output results achieved better classification accuracy compared to the existing DNN methods when evaluated on BU dataset. However, the proposed framework still has minor computational complexity.

Zhao [18] presented a novel method for classifying epileptic seizure onset zone (SOZ) and non-SOZ based on the partial annotation using intracranial electroencephalogram (iEEG) data to overcome the time-consuming task in existing methods. A Positive Unlabeled (PU) binary classifier was introduced to classify little size of labeled data and big size of unlabeled data that reduces the annotation workload. The obtained classification accuracy was better when compared to existing methods with the use of PU on BB dataset. However, if PU doesn’t know the positive proportion of unlabeled data, this proportion error affects the classification performance.

Mishra [19] presented a Discrete Wavelet Transform (DWT) and Moth Flame Optimization (MFO) based Extreme Learning Machine (ELM) for classifying EEG signals for epileptic seizure detection. The DWT transform was used for decomposing the signal followed by the calculation of the statistical features. Then the classification was performed by calculating the optimal parameters of the classifier using the MFO algorithm. The final classification phase was performed on BU dataset by using achieved better classification accuracy due to efficient learning speed, fast convergence, and ease of implementation. However, the MFO has limitations such as local optima trapping and an imbalance between exploration and exploitation.

Wang and Mengoni [20] presented a Natural Language Processing (NLP) approach to predict seizures using EEG frequency bands and montage selection method. Classifiers namely Random Forest, SVM, and Multi-layer perceptron are used for the categorization of the EEG signals. The result obtained has displayed that the suggested approach achieved better classification accuracy with the use of NLP when evaluated on TUH, CHB-MIT Dataset, the UPenn and Mayo Clinic seizure detection dataset compared to existing methods. However, the inability of NLP to adapt to new domains limits its performance in multiple prediction tasks.

Murariu [21] presented an automatic system that used the Empirical Mode Decomposition (EMD) method for the classification of Epilepsy EEG signals from BB and Cluj-Napoca datasets. The spectral power density of intrinsic mode functions was also applied as features to classify the focal and non-focal types of signals. These signals were classified using KNN and Naïve Bayes classifiers and achieved better classification accuracy but the presented approach is immune to the connectivity between the electrodes and scalp and its subject movement.

Xin [22] developed an Attention Mechanism-related Wavelet CNN (AMWCNN) to classify epilepsy EEG signals from BU and BB datasets. This method was presented to overcome the burden of long-term readings of EEG signals by decomposing them and obtaining their components in various frequency bands using wavelet multi-scale analysis. These decomposed signals were given to CNN for extraction and classification with the attention mechanism and achieved better classification accuracy. However, the time consumption is relatively more for this approach.

Abou-Abbas [23] presented an ML based approach to classifying seizures and seizure-free data from TUH dataset using random forest, KNN, SVM, and gradient-boosting decision tree classifiers. The spectral power with the limited band and the signal complexity were also evaluated to represent the abnormalities in capturing and capturing free EEG signals. The generalization ability to detect seizures and their classification has achieved better results with the use of ML classifiers. However, the medical frequency sub bands are not considered in this work which effects the detection and classification accuracy of seizures.

The common limitations observed from the existing methods on classifying epileptic seizures are overfitting of the data, computational complexity, local optima trapping and an imbalance between exploration and exploitation, and more time consumption. To overcome these limitations, this research has proposed a classification of epileptic seizures using LSTM with ZOA.

3. Methodology

In epileptic seizure detection, the proposed ZOA-LSTM network consists of six steps such as data pre-processing with the help of Butterworth 8th order filter; swarm intelligence-based decomposition method; features extracted to be statistic based features, frequency dependent features, features derived from multi-scale wavelet transformation, features concerning entropy measure and power spectral density are acquired from the decomposition of signals, feature selection; feature selection: enhanced spatial bound WOA with SSA and LOBL;

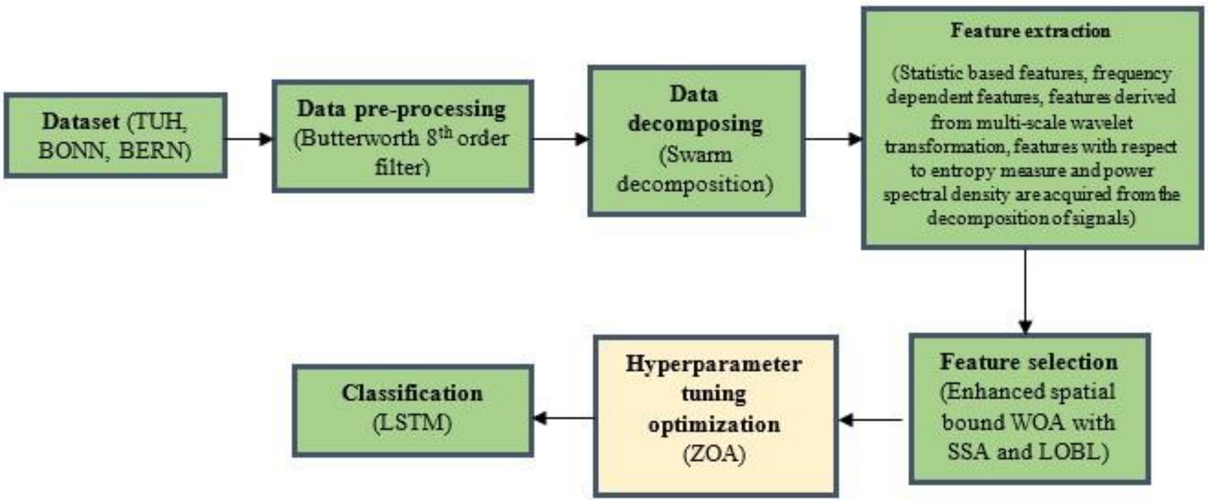

Figure. 1 Block diagram of the proposed EEG classification

hyperparameter tuning optimization: ZOA; and classification: LSTM. The workflow of the suggested approach is represented in Fig. 1.

The six steps of the proposed approach are described and implemented as given in the further sections.

3.1 Dataset

Three open-source and publicly available datasets such as TUH, BERN, BONN are used in this research for classifying raw EEG signals.

Bern Barcelona (BB): The Bern dataset consists of EEG signals of patients suffering from temporal lobe epilepsy that are captured by specialized electrodes. Based on the channel size the dataset contains around 10240 samples, for each file of sample size from 512 or 1024Hz. Dataset link: https://www.upf.edu/web/mdm-dtic/-/1st-testdataset?inheritRedirect=true&redirect=%2Fweb%2Fmdm-dtic%2Fdatasets#.YXK47Z5BzIV

Bonn University (BU): The BONN dataset consists of 4097 samples of EEG signals with a sampling rate of 173.61 Hz. Dataset link: https://www.google.com/url?q=https://epilepsy.uni-freiburg.de/&sa=D&source=hangouts&ust=1634995710030000&usg=AOvVaw0e7ApKfkRdAkSfKOOMAGmz

Temple University Hospital (TUH) dataset: The TUH dataset consists of 30,000 EEG signals of around 10.874 patients from the year 2002 to the current year. Dataset link: mailto:https://isip.piconepress.com/projects/tuh_eeg/

3.2 Data pre-processing

The EEG signals are pre-processed at a frequency of 60Hz using 8th order BWF for eliminating two types of unwanted noises such as electrical and mechanical noises. The maximum order filter is used due to its utility to attenuate the gain beyond the bandwidth. The continuous value of the 8th order filter is measured using Eq. (1).

$$G^2(W) = \frac{G_0^2}{1+\left(\frac{jw}{jw_c}\right)^{2n}} \tag{1}$$

Where, the direct current gain is represented by G_0, the cut-off frequency is given by W_c, and the order of the filter is n.

3.3 Data decomposition

The denoised EEG signals are decomposed using swarm decomposition method which is an impactive decay approach that retrieves the oscillatory elements from the multiple elements EEG signals related to the swarm filter. The denoised EEG signal is taken as an input signal that acts as a prey path to the swarm and this hunting step provides the output of the swarm trajectory [24].

3.4 Feature extraction

After decomposing the signals, various features such as statistic-based features, frequency dependent features, features derived from multi-scale wavelet transformation, features with respect to entropy measure and power spectral density are acquired from the decomposition of signals are extracted from the signals to perform classification.

- **Statistical features:** These features include kurtosis, average value of energy, mean value, variance value and entropy. The statistical features feed the algorithms that are more or less elaborate and that manipulate pixels [25].
- **Frequency features:** These features are extracted with the signal's mean frequency, where signal size is reduced to improve the time complexity and computational complexity of the algorithms [26].
- **Entropy features:** These features can distinguish between different communication signals through state characteristics description of the signals. Entropy measures based on Renyi computation, tsallis measure, Shannon calculation, multiscale permutation, and sure value are the various entropies [27].
- **Spectral features:** To enhance the class separability without any data loss, these features are extracted. Spectral flatness, band power beta, band power gamma, band power delta, band power theta, band power alpha, ratio band power alpha-beta and spectral flux are spectral features [28].
- **Power spectral features:** The density of the power spectrum is one of the spectral features where transforming the EEG signals into power spectra enhances the signals' functionality [29].
- **Multi-scale wavelet transforms:** Energy, variance value, zero crossing rate, waveform length, and standard deviation measure are the type of multi-scale wavelet transforms. These features are extracted when the signal contains not only the frequency components but also the local coordinates information [30].

3.5 Feature selection

After extracting oscillatory features, the features are selected using Enhanced spatial bound Whale optimization with the SSA. The working process of SSA is improved by the WOA, as SSA tends to fall into local optima and also premature convergence. The WOA generates a random population of search agents for a pre-determined number. With the initial iteration, the WOA gradually starts moving alongside the top agent. When the top agent drops into a domestic optimum, the other agents following the leader tend to fall into it automatically. To overcome this, a leader system of SSA is initialized to upgrade the positions of the inheritance and also to improve the exploitation ability of WOA as shown in Eq. (2).

$$X_{1,j}(t+1) = \begin{cases} X_j^*(t) + c_1\left((ub_j - lb_j)c_2 + lb_i\right), c_3 \geq 0.5 \\ X_j^*(t) - \left((ub_j - lb_j)c_2 + lb_i\right), c_3 < 0.5 \end{cases} \tag{2}$$

Where, $X_{1,j}$ and $X_j^*(t)$ depicts the current position of leader and food source in the j^{th} dimension, t represents the number of iterations at the present state, ub_j and lb_j represents the top and bottom bounds in the search space of the j^{th} dimension, c_2 and c_3 are random independent parameters in the range [0,1], c_1 is the balancing factor between exploration and exploitation whose value decreases non-linearly from 2 to 0 as shown in Eq. (3).

$$c_1 = 2e^{-\left(\frac{4t}{T}\right)^2} \tag{3}$$

Where, the current quantity of iterations and high quantity of iterations are denoted by t and T. The updated spot of the leader is provided by Eq. (4).

$$X_{1,j}(t+1) = \frac{1}{2}\left(X_{i.j}(t) + X_{i-1,j}(t)\right) \tag{4}$$

The symbol $i \geq 2$ and $X_{1,j}$ represents the spot of i^{th} follower in the j^{th} dimension.

The feature selection is performed by considering WOA as the major formation which will be optimized by the SSA system and increased by the LOBL strategy. The LOBL strategy is used in measuring an opposite resolution of the fixed resolution by the candidate during the enhancement. The two solutions are compared to select the best solution to continue the iteration process. According to the LOBL principle, while the distance from the object to the lens is twice the focal length, an inverted and shrinking actual image will be created between one and two times the focal length on the alternate side of the lens. The mathematical expression of this principle is expressed as shown in Eq. (5).

$$\frac{(lb+ub)/2-x}{x^*-(lb+ub)/2} = k \tag{5}$$

To solve optimization issues in multi-dimensional Eq. (5) can be modified as shown in Eq. (6).

$$x_i^* = \frac{lb_i+ub_i}{2} + \frac{lb_i+ub_i}{2} - \frac{x_i}{k} \tag{6}$$

Where, x_i^* represents the opposition resolution of x_i in the i-th dimension, lb_i and ub_i are the bottom and top bounds of the i-th dimensions in the exploration space. The standard Opposition-based Learning (OBL) which special case of LOBL can be expressed as shown in Eq. (7).

$$x_i^* = lb_i + ub_i - x_i \tag{7}$$

LOBL can provide a dynamic opposition solution compared to the OBL. The probability of the approach to overcoming the domestic optimum condition would be improved by adjusting the parameter k. The non-linear argument c_1 of SSA in Eq. (2) is employed with the two strategies of WOA namely bubble-net assaulting phase and enclosing prey phase to stable the exploitation and exploration utility.

- **Fitness Function:** The performance of selected features is measured by the classification error rate as shown in Eq. (8).

$$Fitness\ function = \frac{Number\ of\ wrongly\ classified\ instances}{Total\ number\ of\ instances} \tag{8}$$

The Fitness function used in both WOA and SSA is represented in Eq. (9).

$$Fitness = \alpha\gamma_R(D) + \beta\frac{|C-R|}{|C|} \tag{9}$$

Where, the feature subset is illustrated by R, the total quantity of features is given by C, the categorization accuracy of condition attribute set R based on decision D is given by $\gamma_R(D)$, and α, β are the symmetric parameters to the subset length and classification accuracy.

Related to the classification error and chosen features, Eq. (9) is transformed into minimization problem as displayed in Eq. (10).

$$Fitness = \alpha E_R(D) + \beta \frac{|R|}{|C|} \quad (10)$$

3.6 Feature selection

The objective of hyperparameter optimization is to optimiuze the classification execution of EEG Epileptic seizures by increasing the hyperparameters of LSTM classifier. Hyperparameter tuning is a key aspect before performing classification as it yields satisfactory results by reducing the loss function. Hence for better classification results, the optimization of hyperparameter tuning is necessary. In this research, the ZOA is used for optimizing the parameters such as Dropout [0.1-0.4], Learning Rate [0.003-0.1], L2Regularization [0.003-0.1], and Max-Epoch [5, 10, 15, 20]. Then, the best values of hyperparameter are considered by using accuracy as a fitness function. ZOA is a metaheuristic approach that simulates the foraging attitude and restriction approaches of zebras against predators' attacks. ZOA is an effective optimizer that can resolve enhancement issues by providing an applicable balance between exploitation and exploration. It starts by generating random solutions in the initialization state and tries to increase the classification accuracy till the algorithm meets the stopping criteria. The foraging and the defense strategy against predators are the two important types of behavior of zebras [31]. The natural behavior of zebra presented by the ZOA model is mathematically simulated in three steps:

1) Initialization
2) Foraging behavior and
3) Strategies to defend predators

1) Initialization

The ZOA works on the mechanism of optimizing based on the population in the search space. Here, zebras are considered as population members and each one is estimated as a candidate solution to the problem. The positions of zebras are randomly assigned in the search space and its population matrix is illustrated as shown in Eq. (11).

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_N \end{bmatrix}_{N\times m} = \begin{bmatrix} x_{1,1} \cdots & x_{1,j} \cdots & x_{1,m} \\ \vdots & \vdots & \vdots \\ x_{i,1} \cdots & x_{i,j} \cdots & x_{i,m} \\ \vdots & \vdots & \vdots \\ x_{N,1} \cdots & x_{N,j} \cdots & x_{N,m} \end{bmatrix}_{N\times m} \quad (11)$$

Where, X is the total of zebra inheritance, X_i gives the ith zebra, $x_{i,j}$ is the jth problem variable which is suggested by the ith zebra, N is the quantity of zebras, and the quantity of determination variable is m. As individual zebra is estimated as resolution, the objective function could be measured related to each zebra value. Hence the values attained for the objective function are given as a vector utilizing Eq. (12).

$$F = \begin{bmatrix} F_1 \\ \vdots \\ F_i \\ \vdots \\ F_N \end{bmatrix}_{N\times 1} = \begin{bmatrix} F(X_1) \\ \vdots \\ F(X_i) \\ \vdots \\ F(X_N) \end{bmatrix}_{N\times 1} \quad (12)$$

The symbol F is an objection function vector value and its objective function value for ith zebra is F_i. The least value of zebra is considered the best candidate solution for minimization problems and the highest objective function value is the best solution for maximization problems. The position of the zebras is updated in each iteration and the iteration process continues until the best candidate solution is found.

2) Strategies to defend predators

The positions of the zebras are updated based on the simulation behavior of zebras while searching for forage. The plain zebra which is a pioneer gazer among all the zebras will be considered as the best member of the population which leads other zebras in the direction of its position. Hence upgrading the zebra's spot in the foraging stage could be expressed statistically as shown in Eqs. (13) and (14).

$$x_{i,j}^{new,P1} = x_{i.j} + r.\left(PZ_j - I.x_{i,j}\right) \quad (13)$$

$$X_i = \begin{cases} X_i^{new,P1}, & F_i^{new,P1} < F_i; \\ & X_i, else, \end{cases} \quad (14)$$

Where, $X_i^{new,P1}$ gives the updated status of the ith zebra in the first stage, $x_{i,j}^{new,P1}$ is its jth dimension value, $F_i^{new,P1}$ is the objective function value, PZ_j is the pioneer zebra in its jth dimension, r is an irregular number among [0,1], and I is a parameter that belongs to $\{1,2\}$. If $I = 2$, then there is a high probability of modifications in inheritance motion.

3) Phase 2: Defense strategies against predators

The defense strategy opposed to predators is simulated in the second phase by updating the spot of zebras in the exploration space. The defense strategy of zebra depends on the predators and the strategies vary from one predator to the other. It is assumed that two scenarios engage through the same probability:

1) Zebra chooses an escape strategy when the lion assaults the zebra
2) Zebra choosing an offensive strategy with alternate predators attacks the zebra

The escape strategy in the first scenario is mathematically modeled using mode S_1 as shown in Eq. (15).

$$x_{i,j}^{new,P2} = \begin{cases} S_1: x_{i,j} + R.(2r-1).\left(1-\frac{t}{T}\right).x_{i,j}, & P_s \le 0.5, \\ S_2: x_{i,j} + r.\left(AZ_j - I.x_{i,j}\right), & else, \end{cases} \quad (15)$$

Where, $x_{i,j}^{new,P2}$ is the updated condition of the ith zebra in jth dimension, R is a constant which is equal to 0.001, repetition contour is t, r is the irregular number between [0,1], T is the high quantity of iterations, P_s is the probability of choosing one strategy among two between the [0,1] interval, AZ_j is the condition of the assaulted zebra in the jth dimension.

The offensive strategy in second scenario is mathematically modeled using the mode S_2. The position of zebra is decided based on the good objective function while updating the zebra's positions. This updating condition is given as shown in Eq. (16).

$$X_i = \begin{cases} X_i^{new,P2}, & F_i^{new,P2} < F_i; \\ X_i, & else, \end{cases} \quad (16)$$

Where, $X_i^{new,P2}$ is the trend condition of the ith zebra related to the second stage, $F_i^{new,P2}$ is the objective function value.

Pseudocode for ZOA

Start ZOA
1. Input: Range of hyperparameters
2. Put the quantity of iteration ($T = 100$) and the quantity of zebras' population ($N = 30$)
3. Preparation a spot of zebras and estimation of the objective function
4. For $t = 1{:}100$
5. Upgrade pioneer zebra (PZ)
6. For $i = 1{:}30$
7. Stage 1: Foraging behavior
8. Estimate trend conditions of the ith zebra utilizing Eq. (12)
9. Upgrade the ith zebra utilizing Eq. (13)
10. Stage 2: Defense approaches opposed to predators
11. If $P_s < 0.5$, $P_s = rand$
12. Approach 1: opposed to lion (exploitation phase)
13. Estimate trend of the ith zebra utilizing mode S_1 in Eq. (14)
14. Else
15. Approach 2: opposed to another predator (exploration phase)
16. Estimate trend condition of ith zebra utilizing mode S_2 in Eq. (14)
17. End if
18. Upgrade the ith zebra utilizing Eq. (16)
19. End for $i = 1{:}30$
20. Save the best candidate resolution achieved up to this point.
21. End for $t = 1{:}100$
22. Output: the better resolution attained by ZOA for a provided range of hyperparameters

Although ZOA can avoid local optima and get a global optimization solution, it is necessary to update the new population to obtain the best optimal solution.

3.7 Classification of EEG signals

The fine-tuned features are classified using the LSTM network where the given and output features are associated with the regularization of the individual layer. Unlike the other classifiers, LSTM has the advantage of overfitting problems, the upper layer is provided to the chosen features that can optimize the features. The LSTM has multiplicative cells, which are collected compared to the temporary and multiplicative units which consist of various characters that handle the data flow in the memory segment. The three gates are forgotten gate f_t, input gate i_t, and output gate o_t as shown in Fig. 2 plays a significant role in storing the memory components and regulates the information flow in and out of the cells.

The forget gate determines which detail to neglect from the cell, the input gate chooses that detail to add into the cell state, and the output gate gives the final output. The processing of nodes in LSTM with three gates is given through Eqs. (17)-(22).

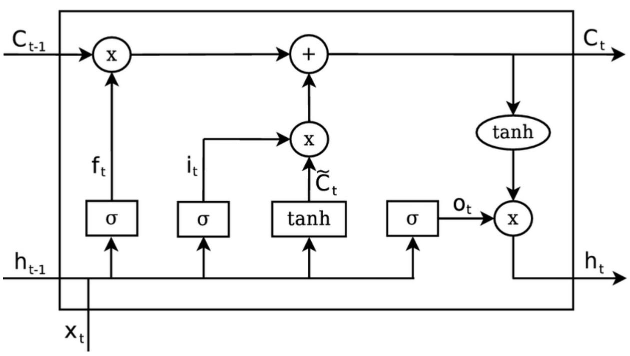

Figure. 2 Architecture of LSTM

$$f_t = \sigma\left(W_f.[h_{t-1}, x_t] + b_f\right) \quad (17)$$

$$i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i) \quad (18)$$

$$\tilde{C}_t = \tanh(W_C.[h_{t-1}, x_t] + b_C) \quad (19)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (20)$$

$$o_t = \sigma(W_0.[h_{t-1}, x_t] + b_0) \quad (21)$$

$$h_t = o_t * \tanh(C_t) \quad (22)$$

Where, f_t denotes forget gate, σ denotes the sigmoid function, h_{t-1} is the hidden state of the prior layer, x_t is the input of the current layer, W and b are the weights and bias state. i_t is the input gate, C_t is the cell state in the next year, $\tilde{C}_t$ is the intermediate temporary state, $C_{t-1}$is cell state present in the prior layer. o_t is the output layer and h_t is the hidden state of the next layer. The parameter setting for LSTM is: 150 is the size of the batch, 0.0015 is the Lambda loss amount, the learning rate is 0.0025, the hidden layer is 32, and the quantity of repetitions is 300.

4. Results and performance analysis

The effectiveness of the proposed approach is evaluated on MATLAB software on a system setup alongside 8-GB RAM and an Intel core i9 processor. In this section, the classification results are represented in tabular forms and graphical representations in terms of accuracy, sensitivity, specificity, F1-score, and MCC on sets of data such as BB, BU, and TUH. Accuracy is a quantity of percentage that is rightly categorized EEG signals that are compared to the overall examination. Sensitivity and specificity are the amounts of the rightly measured percentage of categorized seizures and non-seizure signals. F1- score is termed by the harmonic mean values of precision and recall, which measures values of the features that are empty values. The value of MCC ranges from 0 to 1, where 1 is the best agreement and 0 is a no agreement. The mathematical expression of accuracy, sensitivity, specificity, F1-score, and MCC is given from Eqs. (23)-(27).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (23)$$

$$Specificity = \frac{TN}{TN+FP} \times 100 \quad (24)$$

$$Sensitivity = \frac{TP}{TP+FN} \times 100 \quad (25)$$

$$F1 - score = \frac{2TP}{2TP+FP+FN} \times 100 \quad (26)$$

Whilst, the True Positive and True Negative are given by TP and TN, and False Positive and False Negative are given by FP and FN.

$$MCC = \frac{TP\times TN - FP\times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \times 100 \quad (27)$$

4.1 Performance analysis

The performance of the LSTM classifier with optimized features is analyzed with different classifiers like Recurrent Neural Network (RNN),

Table 1. Optimized features on different classifiers for hyperparameter tuning

BB dataset					
Classifiers	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**	**F1-score (%)**	**MCC**
RNN	92.15	92.68	92.68	90.98	90.32
MSVM	72.55	70.23	70.23	71.23	72.54
Sparse auto encoder	97.15	93.33	93.83	95.00	92.12
Stacked auto encoder	97.32	97.07	97.75	96.46	97.15
LSTM	99.71	98.54	98.99	97.63	99.12
BU Dataset					
RNN	92.45	92.02	93.86	90.64	93.72
MSVM	58.00	56.91	57.69	56.60	55.42
Sparse auto encoder	94.54	93.08	91.21	94.82	90.54
Stacked auto encoder	92.18	93.95	94.54	90.63	91.42
LSTM	98.43	98.04	97.88	97.09	96.98
TUH dataset					
RNN	92.45	92.02	93.86	90.64	93.72
MSVM	58.00	56.91	57.69	56.60	55.42
Sparse auto encoder	94.54	93.08	91.21	94.82	90.54
Stacked autoencoder	92.18	93.95	94.54	90.63	91.42
LSTM	98.43	98.04	97.88	97.09	96.98

Table 2. Optimized features on different optimizers for hyperparameter tuning

Classifiers	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-score (%)	MCC
BB dataset					
PSO	93.40	93.24	94.51	92.06	92.65
GWO	94.05	93.63	92.93	96.16	95.10
WOA	90.40	90.53	89.88	90.72	91.25
SSA	89.39	86.92	88.75	88.14	90.72
Proposed	99.71	98.54	98.99	97.63	99.12
BU Dataset					
PSO	93.35	92.82	90.57	91.39	93.65
GWO	92.87	90.62	91.40	92.48	92.22
WOA	90.96	89.72	91.41	88.85	90.61
SSA	89.28	87.34	86.52	86.12	89.26
Proposed	98.43	98.04	97.88	97.09	96.98
TUH dataset					
PSO	85.54	85.09	85.05	84.47	86.45
GWO	90.40	90.29	89.13	88.92	90.84
WOA	94.50	94.31	94.23	94.72	93.30
SSA	92.50	93.44	93.25	91.25	92.99
Proposed	99.53	98.99	99.01	97.54	98.43

Table 3. Retrieved features and the selected features for BB, BU, and TUH

Datasets	Extracted features	Selected features
BB	3209	2839
BU	3829	3028
TUH	7620	4729

Multi-class Support Vector Machine (MSVM), Sparse autoencoder, and Stacked auto encoder in Table 1. The proposed LSTM classification achieved highest accuracy with 99.71% on BB dataset and 98.43% on BU and TUH datasets when compared to the classifiers such as RNN, MVSM, Sparse auto encoder, and Stacked auto encoder.

Table 2 represents the performance of different optimizers on optimized features with hyperparameter tuning. The proposed technique has achieved the highest accuracy with 99.71% on BB dataset, 98.43% on BU and 99.53% TUH datasets when compared to the conventional optimizers like Particle Swarm Optimization (PSO), Grey Wolf Optimization (GWO), Whale Optimization Algorithm (WOA), Salp Swarm Optimization (SSO) algorithm.

The retrieved features and the selected features for BB, BU, and TUH are illustrated below as shown in table 3.

4.2 Comparative analysis

The portion provides the execution of the suggested ZOA alongside hyperparameter tuning for the classification of EEG is evaluated on BU, BB, and TUH datasets with the existing research such as Hierarchical Attention based CNN [17], Entropy, FCNN, & PU [18], Discrete Wavelet Transform and Moth Flame Optimization-based Extreme Learning Machine (DM-ELM) [19], Complexity and power for seizure detection [23] in table 4.

The limitations of the existing research are demonstrated by comparing the results of the proposed and the conventional methods in table 4. As observed, the existing methods has achieved superior accuracy values in classifying epileptic seizures from

Table 4. Comparable analysis of the former approach with the suggested approach

Methods	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-score (%)
BU dataset				
Hierarchical Attention based CNN [17]	97.03	97.47	96.04	96.94
DM-ELM [19]	92.00	91.00	93.00	94.00
ZOA-LSTM (proposed)	98.43	98.04	97.88	97.09
BB dataset				
Entropy, FCNN, & PU [18]	76.91	-	-	-
ZOA-LSTM (proposed)	99.71	98.54	98.99	97.63
TUH dataset				
Oscillatory power and complexity [23]	91.07	-	-	91.41
ZOA-LSTM (proposed)	98.43	98.04	97.88	97.09

EEG signals compared to proposed method on BU, BB, and TUH dataset. The existing Hierarchical Attention based CNN [17] has achieved less accuracy compared to proposed method due to computational complexity. Similarly, the DM-ELM [19] has drawback of local optima trapping and an imbalance between exploration and exploitation. Entropy, FCNN, & PU [18] resulted in poor classification due to positive proportion of unlabelled data. Oscillatory power and complexity [23] have resulted in the average classification due to inconsideration of frequency subbands. These limitations have been overcome in the proposed method and achieved better classification accuracy on the considered datasets.

4.2.1 Discussions

It is noticed that the suggested ZOA-LSTM has succeeded with good results as shown in Table 4 when compared to existing methods [17] and [19] with a 98.43% accuracy, 98.04% of sensitivity, 97.88% specificity, and 97.09% F1-score on BU dataset. On BB dataset, the performance of the proposed method has achieved 99.71% accuracy, 98.54% sensitivity, 98.99% specificity, and 97.63% f1- score when compared to the existing method [18]. On the TUH dataset, the performance of the proposed method has achieved 98.43% accuracy, 98.04% sensitivity, 97.88% specificity, and 91.41% f1-score. The range of hyperparameters in ZOA is enhanced with the hyperparameter tuning and has overcome the limitations of existing methods such as computational complexity, classification error, the imbalance between exploration and exploitation, and data overfitting by achieving an exceptional set of optimal features for efficient classification of EEG signals to predict epileptic seizures.

5. Conclusion

To detect epilepsy seizures from EEG signals, an effective classification with hyperparameter tuning is proposed in this research. The hyperparameter tuning is performed with the ZOA algorithm to obtain the exceptional set of optimal features for classifying normal and abnormal seizures. Three publicly available datasets such as BB, BU, and TUH are considered for EEG signals of different frequencies. The proposed ZOA-LSTM classification achieved better output results on three datasets when compared to the existing methods such as Hierarchical Attention based CNN, DM-ELM, Entropy, FCNN, & PU classification method, Oscillatory power and complexity detection method 98.43% accuracy, 98.04% sensitivity, 97.88% specificity, and F1-score of 97.09%. Although ZOA can avoid local optima and get a global optimization solution, the future work focuses on the population updating mechanism as it is necessary to update the new population for the best optimal solution.

Notations:

Variables	Notation
G_0	Direct current gain
W_c	Cut-off frequency
$X_{1,j}$	Current position of the leader
$X_j^*(t)$	Current position of food in j^{th} dimension
ub_j	Upper bound of j^{th} dimension in search space
lb_j	Lower bound of j^{th} dimension in search space
c_1	balancing factor between exploration and exploitation
$X_{1,j}$	Position of i^{th} follower in the j^{th} dimension
R	Feature subset
C	Total quantity of features
$\gamma_R(D)$	categorization accuracy of condition attribute set R based on decision D
α, β	Symmetric parameters
X_i	i^{th} zebra
x_i^*	Opposite resolution of x_i in i^{th} dimension
F	Objection function value
$X_i^{new,P1}$	Updated status of i^{th} zebra
$F_i^{new,P1}$	Objective function value
PZ_j	pioneer zebra in jth dimension

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

The paper conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization, have been done by 1st author. The supervision and project administration have been done by 2nd author.

References

[1] J. Escorcia-Gutierrez, K. Beleño, J. Jimenez-Cabas, M. Elhoseny, M.D. Alshehri, and M.M. Selim, "Automated deep learning enabled brain signal classification for epileptic seizure detection on complex measurement systems", *Measurement*, Vol. 196, p. 111226, 2022.

[2] A.M. Hilal, A.A. Albraikan, S. Dhahbi, M.K. Nour, A. Mohamed, A. Motwakel, A.S. Zamani,

and M. Rizwanullah, “Intelligent Epileptic Seizure Detection and Classification Model Using Optimal Deep Canonical Sparse Autoencoder”, *Biology*, Vol. 11, No. 8, p. 1220, 2022.

[3] S.E. Sánchez-Hernández, R.A. Salido-Ruiz, S. Torres-Ramos, and I. Román-Godínez, “Evaluation of Feature Selection Methods for Classification of Epileptic Seizure EEG Signals”, *Sensors*, Vol. 22, No. 8, p. 3066, 2022.

[4] N. Jiwani, K. Gupta, M.H.U. Sharif, N. Adhikari, and N. Afreen, “A LSTM-CNN Model for Epileptic Seizures Detection using EEG Signal”, In: *Proc. of 2022 2nd International Conference on Emerging Smart Technologies and Applications (eSmarTA)*, Ibb, Yemen, pp. 1-5, 2022.

[5] H.K. Fatlawi, and A. Kiss, “An adaptive classification model for predicting epileptic seizures using cloud computing service architecture”, *Applied Sciences*, Vol. 12, No. 7, p. 3408, 2022.

[6] D. Sagga, A. Echtioui, R. Khemakhem, F. Kallel, and A.B. Hamida, “Epileptic Seizures Detection on EEG Signal Using Deep Learning Techniques”, In: *Proc. of 2022 6th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, Sfax, Tunisia, pp. 1-6, 2022.

[7] M.S. Islam, K. Thapa, and S.H. Yang, “Epileptic-Net: An Improved Epileptic Seizure Detection System Using Dense Convolutional Block with Attention Network from EEG”, *Sensors*, Vol. 22, No. 3, p. 728, 2022.

[8] X. Liu, J. Wang, J. Shang, J. Liu, L. Dai, and S. Yuan, “Epileptic Seizure Detection Based on Variational Mode Decomposition and Deep Forest Using EEG Signals”, *Brain Sciences*, Vol. 12, No. 10, p. 1275, 2022.

[9] V. Bhandari, and M.D. Huchaiah, “A new design of epileptic seizure detection using hybrid heuristic-based weighted feature selection and ensemble learning”, *International Journal of Intelligent Robotics and Applications*, Vol. 6, No. 4, pp. 668-693, 2022.

[10] S.Y. Shah, H. Larijani, R.M. Gibson, and D. Liarokapis, “Random neural network based epileptic seizure episode detection exploiting electroencephalogram signals”, *Sensors*, Vol. 22, No. 7, p. 2466, 2022.

[11] S.K. Pandey, R.R. Janghel, P.K. Mishra, and M.K. Ahirwal, “Automated epilepsy seizure detection from EEG signal based on hybrid CNN and LSTM model”, *Signal, Image and Video Processing*, Vol. 17, No. 4, pp. 1113-1122, 2023.

[12] C. Hinchliffe, M. Yogarajah, S. Elkommos, H. Tang, and D. Abasolo, “Entropy measures of electroencephalograms towards the diagnosis of psychogenic Non-epileptic seizures”, *Entropy*, Vol. 24, No. 10, p. 1348, 2022.

[13] M.L. Giudice, E. Ferlazzo, N. Mammone, S. Gasparini, V. Cianci, A. Pascarella, A. Mammì, D. Mandic, F.C. Morabito, and U. Aguglia, “Convolutional Neural Network Classification of Rest EEG Signals among People with Epilepsy, Psychogenic Non Epileptic Seizures and Control Subjects”, *International Journal of Environmental Research and Public Health*, Vol. 19, No. 23, p. 15733, 2022.

[14] V. Alimisis, G. Gennis, K. Touloupas, C. Dimas, N. Uzunoglu, and P.P. Sotiriadis, “Nanopower Integrated Gaussian Mixture Model Classifier for Epileptic Seizure Prediction”, *Bioengineering*, Vol. 9, No. 4, p. 160, 2022.

[15] M.K. Alharthi, K.M. Moria, D.M. Alghazzawi, and H.O. Tayeb, “Epileptic Disorder Detection of Seizures Using EEG Signals”, *Sensors*, Vol. 22, No. 17, p. 6592, 2022.

[16] H. Albaqami, G.M. Hassan, and A. Datta, “Wavelet-Based Multi-Class Seizure Type Classification System”, *Applied Sciences*, Vol. 12, No. 11, p. 5702, 2022.

[17] S.K.R. Chirasani, and S. Manikandan, “A deep neural network for the classification of epileptic seizures using hierarchical attention mechanism”, *Soft Computing*, Vol. 26, No. 11, pp. 5389-5397, 2022.

[18] X. Zhao, Q. Zhao, T. Tanaka, J. Solé-Casals, G. Zhou, T. Mitsuhashi, H. Sugano, N. Yoshida, and J. Cao, “Classification of the Epileptic Seizure Onset Zone Based on Partial Annotation”, *Cognitive Neurodynamics*, Vol. 17, No. 3, pp. 703-713, 2023.

[19] S. Mishra, S.K. Satapathy, S.N. Mohanty, and C.R. Pattnaik, “A DM-ELM based classifier for EEG brain signal classification for epileptic seizure detection”, *Communicative & Integrative Biology*, Vol. 16, No. 1, p. 2153648, 2023.

[20] Z. Wang, and P. Mengoni, “Seizure classification with selected frequency bands and EEG montages: a Natural Language Processing approach”, *Brain Informatics*, Vol. 9, p. 11, 2022.

[21] M.G. Murariu, F.R. Dorobanțu, and D. Tărniceriu, “A Novel Automated Empirical Mode Decomposition (EMD) Based Method and Spectral Feature Extraction for Epilepsy

EEG Signals Classification", *Electronics*, Vol. 12, No. 9, p. 1958, 2023.
[22] Q. Xin, S. Hu, S. Liu, L. Zhao, and Y.D. Zhang, "An Attention-Based Wavelet Convolution Neural Network for Epilepsy EEG Classification", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 30, pp. 957-966, 2022.
[23] L. Abou-Abbas, I. Jemal, K. Henni, Y. Ouakrim, A. Mitiche, and N. Mezghani, "EEG Oscillatory Power and Complexity for Epileptic Seizure Detection", *Applied Sciences*, Vol. 12, No. 9, p. 4181, 2022.
[24] T.J. Rani, and D. Kavitha, "Effective Epileptic Seizure Detection Using Enhanced Salp Swarm Algorithm-based Long Short-Term Memory Network", *IETE Journal of Research*, 2022.
[25] M. Altaf, T. Akram, M.A. Khan, M. Iqbal, M.M.I. Ch, and C.H. Hsu, "A new statistical features based approach for bearing fault diagnosis using vibration signals", *Sensors*, Vol. 22, No. 5, p. 2012, 2022.
[26] X. Wang, T. Xu, D. An, L. Sun, Q. Wang, Z. Pan, and Y. Yue, "Face Mask Identification Using Spatial and Frequency Features in Depth Image from Time-of-Flight Camera", *Sensors*, Vol. 23, No. 3, p. 1596, 2023.
[27] A.A. Rey, A.C. Frery, M. Lucini, J. Gambini, E.T.C. Chagas, and H.S. Ramos, "Asymptotic Distribution of Certain Types of Entropy under the Multinomial Law", *Entropy*, Vol. 25, No. 5, p. 734, 2023.
[28] K. Zawiślak-Fornagiel, D. Ledwoń, M. Bugdol, P. Romaniszyn-Kania, A. Małecki, A. Gorzkowska, and A.W. Mitas, "The Increase of Theta Power and Decrease of Alpha/Theta Ratio as a Manifestation of Cognitive Impairment in Parkinson's Disease", *Journal of Clinical Medicine*, Vol. 12, No. 4, p. 1569, 2023.
[29] D.M. Mateos, G. Krumm, V.A. Filippetti, and M. Gutierrez, "Power spectrum and connectivity analysis in EEG recording during attention and creativity performance in children", *NeuroSci*, Vol. 3, No. 2, pp. 347-365, 2022.
[30] Y. Tang, X. Xie, and Y. Yu, "Hyperspectral Classification of Two-Branch Joint Networks Based on Gaussian Pyramid Multiscale and Wavelet Transform", *IEEE Access*, Vol. 10, pp. 56876-56887, 2022.
[31] E. Trojovská, M. Dehghani, and P. Trojovský, "Zebra Optimization Algorithm: A New Bio-Inspired Optimization Algorithm for Solving Optimization Algorithm", *IEEE Access*, Vol. 10, pp. 49445-49473, 2022.