FOREX Publication
Open Access | Rapid and quality publishing

International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 12, Issue 2 | Pages 596-604 | e-ISSN: 2347-470X

Advancing Sleep Stage Classification with EEG Signal Analysis: LSTM Optimization Using Puffer Fish Algorithm and Explainable AI

Srinivasa Rao Vemula[1], Maruthi Vemula[2], Ghamya Kotapati[3], Lokesh Sai Kiran Vatsavai[4], Lakshmi Naga Jayaprada Gavarraju[5] and Ramesh Vatambeti[6*]

[1]Software Test Analyst Senior, FIS Management Services, Durham, North Carolina 27703-8589, USA; srinivas.vemula@fisglobal.com
[2]Student, North Carolina School of Science and Mathematics, Durham, North Carolina 27705, USA; vemula24m@ncssm.edu
[3] Department of AI & ML, School of Computing, Mohan Babu University, Tirupati 517102, India; ghamyakotapati@gmail.com
[4]Department of Information Technology, SRKR Engineering College, Bhimavaram 534204, India; lokesh3069@srkrec.ac.in
[5]Department of Computer Science and Engineering, Malla Reddy College of Engineering and Technology, Hyderabad 500100, India; lakshminagajayaprada.g@mrcet.ac.in
[6]School of Computer Science and Engineering, VIT-AP University, Vijayawada 522237, India; v2ramesh634@gmail.com

***Correspondence:** Ramesh Vatambeti; v2ramesh634@gmail.com

ABSTRACT- In this study, we introduce SleepXAI, a Convolutional Neural Network-Conditional Random Field (CNN-CRF) technique for automatic multi-class sleep stage classification from polysomnography data. SleepXAI enhances classification accuracy while ensuring explainability by highlighting crucial signal segments. Leveraging Long Short-Term Memory (LSTM) networks, it effectively categorizes epileptic EEG signals. Continuous Wavelet Transform (CWT) optimizes signal quality by analyzing eigenvalue characteristics and removing noise. Eigenvalues, which are scalar values indicating the scaling effect on eigenvectors during linear transformations, are used to ensure clean and representative EEG signals. The Puffer Fish Optimization Algorithm fine-tunes LSTM parameters, achieving heightened accuracy by reducing trainable parameters. Evaluation on the Sleep-EDF-20, Sleep-EDF-78, and SHHS datasets shows promising results, with regular accuracy ranging from 85% to 89%. The proposed LSTM-PFOA algorithm demonstrates efficacy for autonomous sleep categorization network development, promising improved sleep stage classification accuracy and facilitating comprehensive health monitoring practices.

Keywords: Sleep Stage Classification; Convolutional neural network; Continuous Wavelet Transform; Puffer Fish Optimization Algorithm.

ARTICLE INFORMATION

Author(s): Srinivasa Rao Vemula, Maruthi Vemula, Ghamya Kotapati, Lokesh Sai Kiran Vatsavai, Lakshmi Naga Jayaprada Gavarraju and Ramesh Vatambeti;

Received: 26/04/2024; **Accepted**: 11/06/2024; **Published**: 25/06/2024;
e-ISSN: 2347-470X;
Paper Id: IJEER 2604-29;
Citation: 10.37391/IJEER.120235
Webpage-link:
https://ijeer.forexjournal.co.in/archive/volume-12/ijeer-120235.html

Publisher's Note: FOREX Publication stays neutral with regard to Jurisdictional claims in Published maps and institutional affiliations.

1. INTRODUCTION

Sufficient sleep is vital for both physical and mental well-being. It supports the body's recovery and rejuvenation [1]. Inadequate sleep quality can lead to various health issues, including stress, obesity, diabetes, and even mortality [2-3]. Research shows that analyzing an individual's sleep habits, particularly EEG signals during sleep, can predict future health concerns. These concerns include dementia, cardiovascular issues, and psychiatric illnesses. For instance, a study involving 32 healthy individuals demonstrated the potential to forecast the onset of Alzheimer's disease based on sleep patterns [4-6].

With advancements in data processing tools and algorithms, more researchers are delving into sleep data analysis, particularly in sleep stage categorization [7]. This categorization is essential for understanding sleep organization, identifying sleep disorders, and monitoring sleep quality. Physiological signals from various devices like EMG machines and pulse oximetry are commonly used for classification [8-9]. Each thirty-second recording or epoch is assigned a distinct sleep stage according to established guidelines.

Automatic sleep staging methods utilize rule-based algorithms, numerical classification approaches, or hybrid systems [10]. Rule-based approaches may struggle to integrate data patterns effectively, while numerical classification methods analyze spectra or time-frequency domains for PSG characteristics. Some studies employ feature selection with numerous characteristics. Hybrid systems combine the strengths of both approaches, but they may be challenging to implement, with rule-based methods often proving most effective [11-12].

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 12, Issue 2 | Pages 596-604 | e-ISSN: 2347-470X

Many aspects of conventional approaches are tailored to sleep, with each wave type and its spectral range being detailed in *tables 1 and 2*, respectively, for each sleep stage [13].

Table 1: Frequency Bands of EEG Beats

Frequency Band (HZ)	Rhythm
0-4	Delta (δ)
30-40	Gamma 1 ($\gamma 1$)
40-49.5	Gamma 2 ($\gamma 1$)
4-8	Theta (θ)
8-12	Alpha (α)
12-15	Sigma (σ)
15-22	Beta 1 ($\beta 1$)
22-30	Beta 2 ($\beta 2$)

Table 2: Relationship Among Characteristic Waves of EEG And Diverse Phases of Sleep

Sleep Stage	Included Characteristics Waves
W	a, β
N1	a, θ
N2	K multifaceted, spindle wave δ, θ
N3	δ spindle wave
REM	a, β, θ

Feature selection algorithms are crucial in classic machine learning to minimize redundancy and select discriminative features. With the rise of deep learning, models can autonomously learn features, offering advantages like end-to-end categorization and handling massive datasets without extensive prior knowledge [14]. Convolutional Neural Networks (CNNs) are widely used in computer vision tasks and biomedical engineering, including sleep EEG categorization [15-16].

In this study, three datasets undergo pre-processing with Continuous Wavelet Transform (CWT), followed by explainable AI for feature extraction. The extracted signals are then classified using an optimizer-based Long Short-Term Memory (LSTM) network. The parameters are tuned by the Puffer Fish Optimization Algorithm (PFOA). This approach integrates advanced techniques to optimize sleep EEG classification, demonstrating the evolution of algorithms in this field. The structure of the paper is organized as follows: *Section 2* discusses related works; *Section 3* elaborates on the proposed approach in detail; *Section 4* presents the results of the study, and finally, *Section 5* provides the conclusion.

2. RELATED WORK

Zhou et al. [17] proposed a novel method for sleep stage classification by combining multivariate phase space reconstruction (MPSR) with a covariance feature matrix architecture. Their approach aimed to capture the geometric and hidden dynamic features of various physiological signals. By treating covariance matrices constructed using MPSR as Symmetric Positive Definite (SPD) matrices, they established a Riemannian manifold space. These matrices were then transformed into tangent space matrices, converting them into feature vectors in Euclidean space. An ensemble learning classifier was employed to achieve the objectives associated with each sleep stage. Ten-fold cross-validation resulted in accuracies of 88.93% and 88.42% for five-stage classification tasks. Leave-one-subject-out cross-validation yielded accuracies of 82.50% for both the Five-class and Six-class sleep stages tasks on the Sleep EDF dataset.

Pei et al. [18] introduced a new model for sleep stage classification that integrates a deep Convolutional Neural Network (CNN) with long-term memory. They utilized Mel-frequency Cepstral Coefficient (MFCC) extracted from EEG and EMG signals as crucial frequency domain characteristics. The model learned features from the frequency domains of various bio-signal channels, extracting MFCC features from multi-channel signals. These features were then fed into an LSTM layer and several convolutional layers. The learned representations were passed to a fully connected classifier for sleep stage classification. The model demonstrated effective sleep stage classification based on the 2D MFCC feature, achieving an accuracy of 82.35% and a Cohen's kappa of 0.75 on the SHHS dataset, and an accuracy of 74.07% and a kappa of 0.63 on the UCDDB dataset.

Liu et al. [19] proposed a method for sleep stage classification based on Dynamic Mode Decomposition (DMD). DMD decomposes polysomnograms into EEG and EOG signals recorded on separate channels, and features are derived from the dynamic mode powers of these signals' epochs. Random forest classification is then employed for categorization. Numerical simulations demonstrated the superiority of this method, achieving a Cohen's Kappa of 0.9980 and a classification accuracy of 99.8748% for the six stages of sleep.

Pan et al. [20] introduced a sleep evaluation technique to assess the sleep organization of Disorders of Consciousness (DOC) patients. Their algorithm consists of an automated model for sleep staging based on CNNs for signal feature extraction from EEGs and EOGs, followed by a bidirectional LSTM (Bi-LSTM) with an attention mechanism to learn sequential information. A classifier evaluates consciousness based on the automated sleep staging results. The CNN-BiLSTM model coupled with an attention sleep network (CBASleepNet) achieved a total accuracy of 81.8% in assessing consciousness levels.

Zhou et al. [21] proposed a sleep stage classification method using Layer-wise Relevance Propagation (LRP) for human-understandable explanations. Raw EEG signals are transformed into time-frequency images using Short-Time Fourier Transform (STFT). The MSSENet model, based on CNNs, processes these images and utilizes the residual squeeze-and-excitation (R-SE) block for classification. LRP evaluates the contribution of each frequency pixel to the model's decision, with experimental results showing superior performance compared to other methods.

Yeh et al. [22] introduced a method for EEG-based sleep stage classification using the Wigner-Ville Distribution (WVD) for

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 12, Issue 2 | Pages 596-604 | e-ISSN: 2347-470X

time-frequency analysis. Particle Swarm Optimization (PSO) is employed to determine thresholds for wakefulness-stage N3. The proposed method demonstrated high accuracy, sensitivity, and a high kappa coefficient in identifying sleep stages in accordance with AASM guidelines.

Heng et al. [23] presented an automatic end-to-end sleep stage classification approach based on EEG signal processing. This approach employs feature learning on critical regions of sleep EEG signals using a CNN and incorporates a bidirectional recurrent unit (Bi-GRU) to capture sleep stage transition rules. An attention mechanism enhances the Bi-GRU's long-term memory capacity and emphasizes crucial features' impact. The proposed method exhibits high accuracy and stability, outperforming competitors with a simpler structure and better representation when tested on the same dataset and model architecture.

3. PROPOSED METHODOLOGY

Three publicly available datasets were utilized in the experiments: Sleep-EDF-20, Sleep-EDF-78, and SHHS [24-26].

Sleep-EDF-20 consists of 39 PSG records, with one night's data for Subject 13 and two nights for others. The dataset includes Sleep Cassette (SC) and Sleep Telemetry (ST) studies on healthy individuals. Each PSG record comprises event markers, EOG, EMG for chin, and two EEG channels (Fpz-Cz and Pz-Oz) with a sampling rate of 100 Hz. Sleep stages (W, N1, N2, N3, N4, REM, Movement (M), and UNKNOWN) were manually scored in 30-second intervals according to the Rechtschaffen & Kales manual.

Sleep-EDF-78 expands on Sleep-EDF-20, containing 153 PSG records from 78 subjects, mostly with two complete nights of data. The structure is similar to Sleep-EDF-20.

SHHS investigates sleep-disordered breathing's effects on cardiovascular health across multiple centers. 329 individuals, meeting regular sleeper criteria (*e.g.,* AHI score < 5), were selected from 6,441 participants. The dataset includes the C4-A1 channel sampled at 125 Hz.

Table 3 illustrates the dataset format.

Table 3: Data Format

Dataset	Array	Array Shape	Array Content
Sleep-EDF	Data	n × 1 × 3000	number × channel × data
	Lables	n × 1	number × channel
SHHS	Data	n × 1 × 3750	number × channel × data
	Lables	n × 1	number × channel

The data sample point for SleepEDF is 3000, with a time of 100 Hz. The expert-selected channel is 1 for the Fpz-Cz channel, and the total is n. When it comes to SHHS, the number of sleep EEG markers for each participant as determined by the expert is denoted by n, the selected channel is C4-A1, and the data sampling point is 3750. The data is sampled at 125 Hz and for 30 seconds.

3.1. Pre-processing

The Wigner-Ville distribution (WVD), a type of time-frequency used for pre-processing the datasets mentioned before. The WVD displays the signal's energy density in the time-frequency domain using a time-frequency analysis. Since the Fourier transform (FT) theory relies on the idea of a stationary signal—which the EEG is not—and because the EEG's frequency content changes with time, the TFA was created to address the shortcomings of the FT. To find the WVD of a time signal x(t), one uses the formula below:

$$WVD_x(t,f) = \int_{-\infty}^{\infty} x\left(t+\frac{\tau}{2}\right) x^*\left(t-\frac{\tau}{2}\right) e^{-j2\pi f\tau} d\tau \quad (1)$$

where * characterizes the. The $WVD_x(t,f)$ is a the creation of x(t + τ2)x∗(t − τ2) owns Hermitian symmetry in τ.

The wavelet transforms further examples of Fourier-based TFA. Essentially, these techniques apply the Fourier transform to signals within a dynamic window. In most cases, the uncertainty principle restricts the usability of Fourier-based TFA.

There are no limitations imposed by the uncertainty principle on WVD as there are on the Fourier-based TFA. Therefore, it is possible to concurrently attain the best time and frequency resolutions. When the signal length is infinite, WVD's marginal spectrum is equal to the power spectrum of the signal, preserving the frequency marginal condition.

$$\int_{-\infty}^{\infty} WVD_x(t,f)dt = |X(f)|^2 \quad (2)$$

The Wigner-Ville Distribution (WVD) offers precise control over brainwave frequency distribution, making it ideal for Time-Frequency Analysis (TFA). However, its major drawback is cross-terms, hindering practical use. Existing approaches primarily filter cross-terms with low-pass window functions, sacrificing time-frequency resolution. Alternatively, integrating the original signal yields a cross-term-free marginal spectrum.

3.2. Continuous Wavelet Transform (CWT)

The Wigner-Ville Distribution (WVD) offers precise brainwave frequency determination but suffers from cross-terms, hindering practical use. Existing methods primarily filter cross-terms with low-pass window functions, sacrificing time-frequency resolution. Alternatively, integrating the original signal yields a cross-term-free marginal spectrum. Therefore, the original signal can be directly integrated to obtain a marginal spectrum that is uncontaminated by the cross-terms.:

$$C_a(b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t)\varphi\left(\frac{t-b}{a}\right) dt \quad (3)$$

In this case, φ(t) is the mother parameter, and φ(t) is the scale parameter. One way to transform the scale into frequency is by:

$$F = \frac{F_c \times F_s}{a} \quad (4)$$

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 12, Issue 2 | Pages 596-604 | e-ISSN: 2347-470X

Where F_c where a scale parameter, wavelet. The time-frequency analysis is greatly impacted by the choice of mother wavelet, which is among all the accessible wavelets. Due to its smaller frequency variance compared to the mother wavelet, the bump wavelet has found application in EEG signal analysis.

3.3. Feature Extraction using Proposed SleepXAI architecture

The model consists of two main components: the sleep encoder, which extracts features from pre-processed EEG input using a time-distributed layer, and the sleep labeler, which employs CNNs and CRFs to correlate features with contextual information and generate probability scores for sleep labels.

3.3.1. Sleep encoder

The layers that make up the sleep encoder are global maxpool, spatial dropout, maxpool, and convolution. To get certain features out of the EEG signal input, the convolution layer is applied in pairs. The network is able to from the second layer after the first layer has extracted low-level ones. In order to get one-dimensional features, 1D-CNN uses EEG signals; different kernels then extract distinct EEG characteristics. This is the way 1D-CNN expresses forward propagation:

$$x_k^l = b_k^l + \sum_{i=1}^{N_{l-1}} conv1D\left(w_{ik}^{l-1}, s_i^{l-1}\right) \quad (5)$$

b_k^l is kernel bias, s_i^{l-1} is output at layer l − 1, w_{ik}^{l-1} is the kernel ith layer l – 1 and the kth layer l.

$$y_k^l = f\left(x_k^l\right) \quad (6)$$

where y_k^l as the intermediate output, and as the activation function, we have f (·).

To reduce dimensionality and maintain prior feature maps, Max pooling downsamples filter-produced feature maps by selecting maximum values. Spatial Dropout enhances feature map independence by replacing strongly correlated maps with single sections. Global maxpooling downsamples input temporally. Dropout addresses overfitting by randomly deactivating neurons. A dense layer modifies encoded sequence dimensions by receiving output from the previous layer's neurons. The sleep encoder's main function is to encode signals, enabling the sleep labeler to better learn long-range dependencies.

3.3.2 Sleep labeler

There are distinct changes from one stage of sleep to another during the night, and the sleep labelers try to learn these changes in order to capture them. The sleep encoder feeds the encoded sequence into the sleep labeler, which uses a unified implementation of CNN and CRF to identify temporal and outputs the likelihood scoring of each sleep label. In this section, the CNN-CRF model and how it is integrated with the problem scenario that has been proposed are explained in depth. In addition, we offer an algorithm that details the entire procedure of the proposed SleepXAI method.

A) CNN-CRF model integration with SleepXAI

The prediction model takes into account the independence of neighboring feature vectors st−1 and st+1, and then, for each label yt, it uses the related input St to assign a probability. The conditional probability delivery for this is given by

$$p(y|s) = \prod_t p(y_t|s_t) = \prod_t exp\left(\Phi(s_t)_{yt}\right)/Z(S_t) \quad (7)$$

Modeling and identifying interdependencies across sleep stages is possible, nevertheless, by making use of CRF qualities and learning the interdependencies of the chains that make up the input sequence.

One way to describe it mathematically is like this:

$$(y|S) = \frac{1}{Z(S)} exp\left\{\sum_{t=1}^{T} \Phi(S_{t-1}, s_t, s_{t+1})_{y_t} + \sum_{t=1}^{T-1} V_{y_t} y_{t+1}\right\} \quad (8)$$

The CRF model can be implemented with the help of neural networks by learning or estimating a set of weights λ and then assigning them to feature vectors. In order to simplify the calculation of the partial derivative, which is provided by, the parameters (lambda) are estimated using maximum likelihood estimation applied to distribution.

$$L(y, S,) = -log\{\prod_{k=1}^{m} P(y^k|s^k, \lambda)\} \quad (9)$$

$$= -\sum_{k=1}^{m} log\left[\frac{1}{Z(s_m)} exp\left\{\sum_{t=1}^{T}\sum_j \lambda_j \Phi_j\left(S^m, t, y_{t-1}^k, y_t^k\right)\right\}\right] \quad (10)$$

Since we take the argmin to apply extreme likelihood to function. Taking the partial derivative with regard to lambda allows us to determine the minimum, and we obtain:

$$\frac{\partial L(y,s,\lambda)}{\partial \lambda} = \frac{-1}{m}\sum_{k=1}^{m} \phi_j(y^k, s^k) + \sum_{k=1}^{m} p(y|s^k, \lambda)\phi_j(y, s^k) \quad (11)$$

λ, an inform is given by

$$\lambda = \lambda + \alpha\left[\sum_{k=1}^{m} \phi_j(y^k, s^k) + \sum_{k=1}^{m} p(y|s^k, \lambda)\phi_j(y, s^k)\right] \quad (12)$$

The limit λ takes a little step each iteration until the values converges, since it is updated repeatedly using gradient descent. Therefore, the SleepXAI architecture concludes with a CRF layer. Predicting sequences using contextual information from a nearby sample is made possible by CRFs, which function as a probabilistic graph model. Thus, in the context of classifying sleep phases, specific stages of sleep are regarded as following or occurring after specific stages of sleep. Classification accuracy for minor examples, like N_1 and N_3, is significantly improved by CRFs' excellent learning of these transitions.

3.4. Feature normalization

To mitigate individual variability, feature normalization was applied independently to each feature. The process involves averaging the 10% lowest and highest values to determine the minimum and maximum. Subsequently, values are scaled between 0 and 1, with values exceeding 1 set to 1 and values below 0 set to 0. Certain features, like "0-30 E" and "22-30 E," are uniquely associated with specific sleep stages, enabling differentiation between REM and N3 and between Wake and N3, respectively.

3.5. Classification using Long Short-Term Memory (LSTM)

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 12, Issue 2 | Pages 596-604 | e-ISSN: 2347-470X

Deep learning, a subset of machine learning, involves training neural networks with multiple hidden layers to perform complex nonlinear operations. Recurrent Neural Networks (RNNs) are particularly adept at processing sequential data, leveraging temporal correlations between data points to make predictions. Each node in an RNN represents a network layer at a given time, connected through weighted connections among input, hidden, and output layers. The weighted connections are denoted by matrices UUU, WWW, and VVV for hidden-hidden, hidden-output, and input-hidden layers respectively. The softmax algorithm transforms the last weight matrix into a binary variable, Ŷ, representing predicted classes, compared to actual classes (Y) using a loss function.

However, RNNs face the challenge of vanishing gradients, where gradients either decrease or grow exponentially over time, affecting the network's ability to learn long-term dependencies. Gated Recurrent Units (GRUs) and Long Short-Term Memory (LSTM) networks are popular solutions to this problem. LSTM, the focus of the present study, utilizes memory blocks composed of interconnected subnetworks capable of retaining information for extended periods. Each block contains at least one unit with input/output gates and forget gates, allowing it to store and retrieve information.

The study employs Python (version 3.6) with TensorFlow v1.0 at its core and the Keras API for modeling and validation. The LSTM model's hyperparameters are detailed in *table 4*. With its ability to analyze the temporal connections in EEG signals, LSTM proves to be a suitable tool for EEG data analysis.

Table 4. Enterprise of Hyper-parameters of LSTM classical

Proposed Model	Optimizer	Batch Size	Loss Function	Drop out rate	Learning rate	Activation Function
LSTM	PFOA	256	Binary Cross entropy	0.2	0.0001	ReLu, Sigmoid

Table 4 shows the LSTM design hyperparameters; using the PFOA optimizer, the ReLu and Sigmoid activation functions get a learning rate of 0.0001 and work quite well.

A generative architectures distribution p (s, y) from feature extraction, where $y = [y1, y2, \ldots, yT]$ is the output label vector and $s = [s1, s2, \ldots, sT]$ are input feature vectors. In sleep stages, namely, $yt \in \{W, N1, N2, N3, REM\}$.

3.5.1. Hyper-parameter tuning using PFOA

The Puffer Fish Optimization Algorithm (PFOA) enhances the LSTM model by optimizing its hyperparameters through a population-based, iteration-based search procedure. Each member of the PFOA population represents a potential solution, with decision variables mapped to a vector. The population explores the search space to identify optimal values for hyperparameters, including learning rate, batch size, dropout rate, and activation functions. This process is mathematically modeled, with initial positions of PFOA members set using *equation (13)* and updated iteratively using *equation (14).*

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_N \end{bmatrix}_{N\times m} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,d} & \cdots & x_{1,m} \\ \vdots & \ddots & \vdots & \iddots & \vdots \\ x_{i,1} & \cdots & x_{i,d} & \cdots & x_{i,m} \\ \vdots & \iddots & \vdots & \ddots & \vdots \\ x_{N,1} & \cdots & x_{N,d} & \cdots & x_{N,m} \end{bmatrix}_{N\times m} \quad (13)$$

$$x_{i,d} = lb_d + r.(ub_d - lb_d) \quad (14)$$

Here, X is the PFOA populace matrix, X_i is the i[th] PFOA member, $x_{i,d}$ is its dth dimension in the N is the sum of population members, m is the sum of r is a random sum in the intermission [0, 1], besides lb_d and ub_d singly.

The objective function of the problem can be measured with each PFOA member as a potential solution. A vector can be used to represent the set of assessed morals for the problem's objective function in accordance with *equation (15)*.

$$F = \begin{bmatrix} F_1 \\ \vdots \\ F_i \\ \vdots \\ F_N \end{bmatrix}_{N\times 1} = \begin{bmatrix} F(X_1) \\ \vdots \\ F(X_i) \\ \vdots \\ F(X_N) \end{bmatrix}_{N\times 1} \quad (15)$$

In this case, the assessed objective function based on the i[th] PFOA member is denoted as Fi, and the evaluated objective function vector is F.

In order to evaluate the merit of each PFOA member's offered solutions, we can use the evaluated standards of the function. The optimal member, or best candidate key, is represented by the objective function's best evaluated value, while the worst member, or worst candidate solution, is represented by the worst evaluated value. With each iteration, the PFOA members' positions in the space are updated. It is only fair that the best member, determined by comparing newly evaluated function, be updated as well.

A) Mathematical Modelling of PFOA

The suggested PFOA method updates the population's problem-solving space location according to simulated natural interactions between pufferfish and their predators. The pufferfish is the initial target of the predator in this cycle. As a last line of defense, the pufferfish coils up into a ball of sharp spines, causing the predator to flinch and flee.

Phase 1: Marauder Attack towards Pufferfish (Exploration Stage)

The initial step of PFOA involves updating the population's position according to the simulated predator attack technique on pufferfish. A pufferfish's lack of speed makes it an easy target for predators. The PFOA members' sites in the space are updated by simulating the predator's position shift during the attack on the pufferfish. approach to the pufferfish causes the PFOA members to move about a lot, which boosts the algorithm's ability to search the entire ocean.

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 12, Issue 2 | Pages 596-604 | e-ISSN: 2347-470X

Each member of the population acts as a predator in PFOA design; the pufferfish that is most likely to attack is determined by where other members of the population have a higher value function. *Equation (16)* is used to identify the set of members of the population.

$$CP_i = \{X_k : F_k < F_i \ \ and\ k \neq i\}, where\ i = 1,2,\dots,N\ and\ k \in \{1,2,\dots,N\} \quad (16)$$

Here, CP_i is the set of predators, X_k is the populace objective function ith marauder, besides F_k is function worth.

The PFOA project is based on the premise that the predator randomly chooses one pufferfish from the candidate set in the CP set; this pufferfish is called the selected pufferfish (SP). By simulating the predator's approach to the pufferfish, we can use *equation (17)* to choose a new location in space for every member of the PFOA. Following this, the relevant member's old location is replaced by the new one if the enhanced in the novel position, as per *equation (18)*.

$$x_{i,j}^{P1} = x_{i,j} + r_{i,j}.\left(SP_{i,j} - I_{i,j}.x_{i,j}\right) \quad (17)$$

$$X_i = \begin{cases} X_i^{P1}, & F_i^{P1} \leq F_i \\ X_i, & else \end{cases} \quad (18)$$

Here, SP_i is the designated chosen arbitrarily from the CP_i set (i.e., SP_i is an component of the CP_i set), $SP_{i,j}$ is its jth measurement, X_i^{P1} is the novel position calculated for projected PFOA, X_i^{P1} is its jth dimension, F_i^{P1} is its charge, $r_{i,j}$ are accidental statistics from the intermission [0, 1], and $I_{i,j}$ are statistics which are arbitrarily designated as 1 or 2.

Phase 2: Defence Instrument of Pufferfish against Marauders

In phase two of PFOA, inspired by a pufferfish's defence mechanism, the algorithm updates member positions to mimic a predator's retreat. This simulation boosts local search exploitation. Each member's new location is determined by *equation (19)*, based on the predator's position shift. Then, *equation (20)* evaluates this new site's suitability, enhancing the objective function value. *Equation (20)* determines if the new position is embraced or rejected based on its impact on the objective function. This optimization ensures that each member's position aligns with improving the objective function, enhancing the algorithm's efficiency.

$$x_{i,j}^{P2} = x_{i,j} + \left(1 - 2\,r_{i,j}\right).\frac{ub_j - lb_j}{t} \quad (19)$$

$$X_i = \begin{cases} X_i^{P2}, & F_i^{P2} \leq F_i \\ X_i, & else \end{cases} \quad (20)$$

Here, X_i^{P2} is the newfangled position intended for the i[th] phase of the proposed PFOA, $x_{i,j}^{P2}$ jth dimension, F_i^{P2} is its impartial purpose value, $r_{i,j}$ are accidental statistics from the intermission [0, 1], besides t is the repetition counter.

B) Computational Complexity of PFOA

The computational complexity of the PFOA method is O(NmT), where N is the population size and m are the number of decision variables. Each iteration involves two steps, leading to a complexity of O(2NmT). Considering the overall complexity, accounting for fixed numbers, it simplifies to O(NmT).

4. RESULTS AND DISCUSSION

Here you can find a detailed description of the experimental setup. In the end, the paper shows the outcomes of the experiments and compares our method's performance to others using the Sleep-EDF-20, Sleep-EDF-78, and SHHS databases. The accuracy besides loss of the suggested perfect on testing and training data are exposed in *figures 1 and 2*, separately.

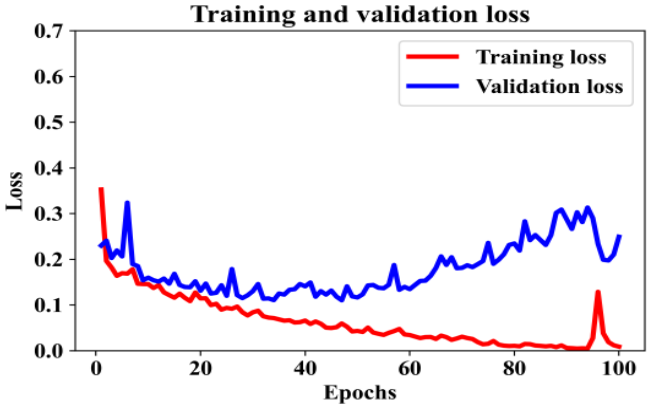

Figure 1: Loss of the projected model for training besides testing data

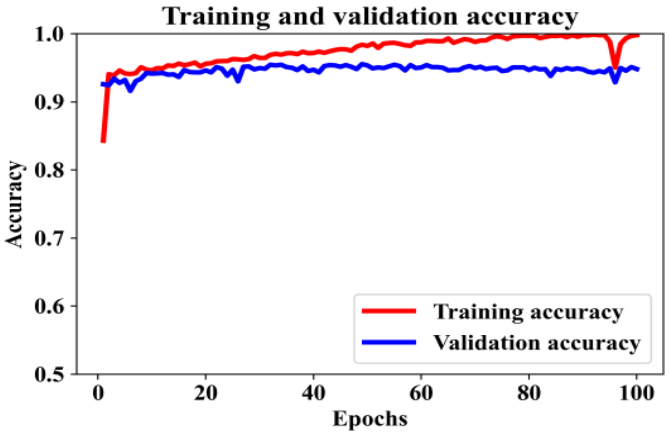

Figure 2: Accuracy of the projected classical for training besides testing data

4.1. Validation analysis of Proposed Model

Figures 3 to 5 provides the untried investigation of projected model with present procedures in terms of different -EDF-20 dataset, Sleep-EDF-78 besides SHHS databases. The existing techniques from related works are tested on different datasets, hence this research work considered three dataset and average results are mentioned in the tables.

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 12, Issue 2 | Pages 596-604 | e-ISSN: 2347-470X

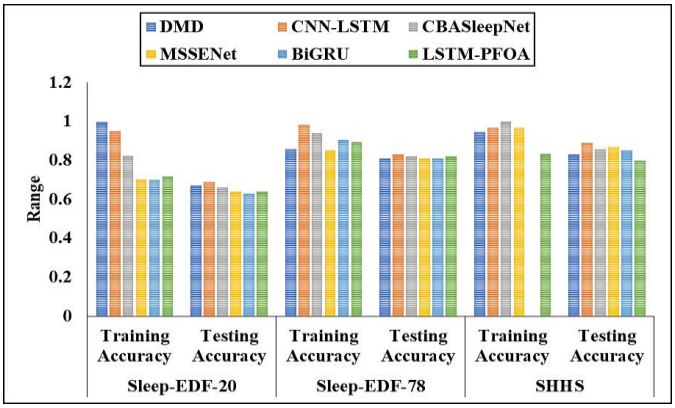

Figure 3: Graphical Analysis of Projected perfect in terms of Accuracy

In *figure 3*, the performance analysis of different metrics on the Sleep-EDF-20 dataset is presented. For the DMD technique [19], the training accuracy is 0.9951, while the precision and F1-score are 0.69 and 0.67, respectively. In the case of the CNN-LSTM technique [18], the training accuracy is 0.9504, with precision and recall both at 0.67, resulting in an F1-score of 0.65. The CBASleepNet technique [20] achieves a training accuracy of 0.8228, and the precision and F1-score are 0.66 and 0.64, respectively. For the MSSENet technique [21], the training accuracy is 0.7045, with precision and recall at 0.59 and 0.64, respectively, resulting in an F1-score of 0.58. In the case of the BiGRU technique [22], the training accuracy is 0.6992, with precision and recall both at 0.63, resulting in an F1-score of 0.60. Finally, the LSTM-PFOA technique achieves a training accuracy of 0.7186, with precision at 0.64, recall at 0.65, and an F1-score of 0.59. These metrics provide insight into the performance of each technique on the Sleep-EDF-20 dataset.

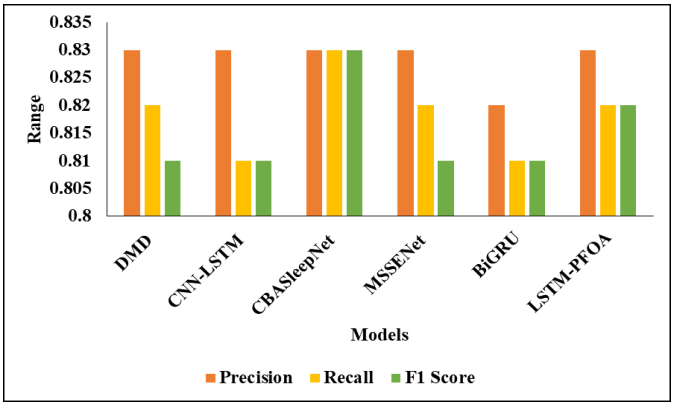

Figure 4: Visual Representation of Projected Perfect for Sleep-EDF-78

Figure 4 illustrates the Validation Analysis of various techniques on the Sleep-EDF-78 dataset. The DMD technique shows a training accuracy of 0.8594 and testing accuracy of 0.81, with precision, recall, and F1-score at 0.83, 0.82, and 0.81 respectively. CNN-LSTM achieves a training accuracy of 0.9825 and testing accuracy of 0.83, with precision, recall, and F1-score at 0.83, 0.81, and 0.81 respectively. CBASleepNet records a training accuracy of 0.9396, with precision, recall, and F1-score at 0.82, 0.83, and 0.83 respectively. MSSENet and BiGRU both demonstrate training accuracies of 0.8515 and 0.9043 respectively, with testing accuracies at 0.81. LSTM-PFOA attains a training accuracy of 0.8949 and testing accuracy of 0.82, with precision, recall, and F1-score at 0.83, 0.82, and 0.82 respectively, offering a comprehensive comparison of techniques.

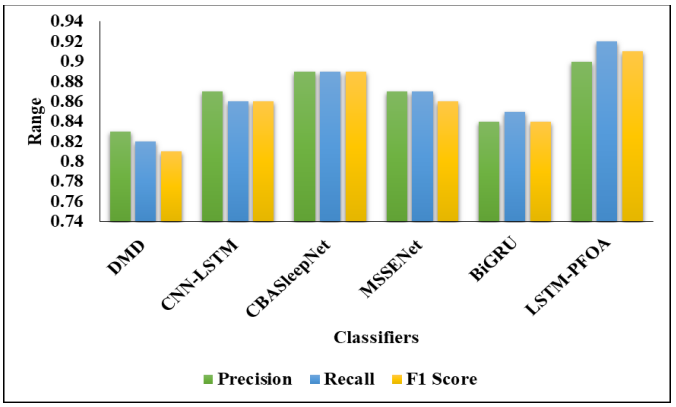

Figure 5: Visual Representation of projected model with existing practices for SHHS dataset

Figure 5 displays the Validation Investigation of various techniques on the SHHS dataset. DMD shows a training accuracy of 0.9476 and testing accuracy of 0.83, with precision, recall, and F1-score at 0.83, 0.82, and 0.81 respectively. CNN-LSTM achieves training accuracy of 0.9670 and testing accuracy of 0.89, with precision, recall, and F1-score at 0.87, 0.86, and 0.86 respectively. CBASleepNet records a training accuracy of 1.0000 and testing accuracy of 0.86, with precision, recall, and F1-score at 0.89, 0.89, and 0.89 respectively. MSSENet and BiGRU both demonstrate training accuracies of 0.9696 and 0.9997 respectively, with testing accuracies at 0.87 and 0.85. LSTM-PFOA attains a training accuracy of 0.8325 and testing accuracy of 0.80, with precision, recall, and F1-score at 0.90, 0.92, and 0.91 respectively, offering insights into each technique's performance on the SHHS dataset.

We can see the optimised batch formation option in *table 5*. As the batch size is decreased, the test recall performance is seen to decline. Because it achieves the highest possible test recall, the suggested work's batch size of 128 is ideal.

Table 5: Optimization in assortment of batch creation

Batch Size	Test Recall
512	96.17
256	98.83
128	99.08
64	93.37
32	91.99

Table 5 depicts the optimization in the assortment of batch creation. For a batch size of 512, the recall is 96.17%. Decreasing the batch size to 256 increases the recall to 98.83%. Further reducing the batch size to 128 results in a recall of 99.08%. However, when the batch size is decreased to 64, the

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 12, Issue 2 | Pages 596-604 | e-ISSN: 2347-470X

recall decreases to 93.37%. Lastly, with a batch size of 32, the recall is 91.99%. These findings highlight the impact of batch size on the model's performance, indicating higher recall values with smaller batch sizes up to a certain threshold, beyond which there is a decline in performance.

5. CONCLUSION

In this study, we introduced SleepXAI, a Convolutional Neural Network-Conditional Random Field (CNN-CRF) technique, for automatic multi-class sleep stage classification. By leveraging Long Short-Term Memory (LSTM) networks and optimizing signal quality through Continuous Wavelet Transform (CWT), we enhanced classification accuracy and ensured explainability. The Puffer Fish Optimization Algorithm (PFOA) significantly contributed to hyperparameter tuning, balancing exploration and exploitation to achieve optimal model performance. Evaluation on the Sleep-EDF-20, Sleep-EDF-78, and SHHS datasets demonstrated promising results, with accuracies ranging from 85% to 89%. Our approach successfully categorized sleep stages, providing a reliable method for autonomous sleep monitoring and facilitating comprehensive health assessments. Future research will focus on addressing classification accuracy for the N1 sleep stage and exploring methods to enhance model efficiency and conciseness. Overall, SleepXAI represents a significant advancement in sleep stage classification, combining state-of-the-art techniques to deliver high accuracy and explainability in sleep health monitoring.

REFERENCES

[1] Dutt, M., Redhu, S., Goodwin, M., & Omlin, C. W. (2023). SleepXAI: An explainable deep learning approach for multi-class sleep stage identification. Applied Intelligence, 53(13), 16830-16843.

[2] Kwon, K., Kwon, S., & Yeo, W. H. (2022). Automatic and accurate sleep stage classification via a convolutional deep neural network and nanomembrane electrodes. Biosensors, 12(3), 155.

[3] Zhou, D., Wang, J., Hu, G., Zhang, J., Li, F., Yan, R., ... & Cong, F. (2022). SingleChannelNet: A model for automatic sleep stage classification with raw single-channel EEG. Biomedical Signal Processing and Control, 75, 103592.

[4] Zhao, D., Jiang, R., Feng, M., Yang, J., Wang, Y., Hou, X., & Wang, X. (2022). A deep learning algorithm based on 1D CNN-LSTM for automatic sleep staging. Technology and Health Care, 30(2), 323-336.

[5] Prabhakar, S. K., Rajaguru, H., Ryu, S., Jeong, I. C., & Won, D. O. (2022). A holistic strategy for classification of sleep stages with EEG. Sensors, 22(9), 3557.

[6] Li, C., Qi, Y., Ding, X., Zhao, J., Sang, T., & Lee, M. (2022). A deep learning method approach for sleep stage classification with EEG spectrogram. International Journal of Environmental Research and Public Health, 19(10), 6322.

[7] Pei, W., Li, Y., Siuly, S., & Wen, P. (2022). A hybrid deep learning scheme for multi-channel sleep stage classification. Computers, Materials and Continua, 71(1), 889-905.

[8] Sekkal, R. N., Bereksi-Reguig, F., Ruiz-Fernandez, D., Dib, N., & Sekkal, S. (2022). Automatic sleep stage classification: From classical machine learning methods to deep learning. Biomedical Signal Processing and Control, 77, 103751.

[9] Goshtasbi, N., Boostani, R., & Sanei, S. (2022). SleepFCN: A fully convolutional deep learning framework for sleep stage classification using single-channel electroencephalograms. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 30, 2088-2096.

[10] Kim, D., Lee, J., Woo, Y., Jeong, J., Kim, C., & Kim, D. K. (2022). Deep learning application to clinical decision support system in sleep stage classification. Journal of Personalized Medicine, 12(2), 136.

[11] Satapathy, S. K., & Loganathan, D. (2022). Automated classification of sleep stages using single-channel EEG: A machine learning-based method. International Journal of Information Retrieval Research (IJIRR), 12(2), 1-19.

[12] ElMoaqet, H., Eid, M., Ryalat, M., & Penzel, T. (2022). A deep transfer learning framework for sleep stage classification with single-channel EEG signals. Sensors, 22(22), 8826.

[13] Leino, A., Korkalainen, H., Kalevo, L., Nikkonen, S., Kainulainen, S., Ryan, A., ... & Myllymaa, K. (2022). Deep learning enables accurate automatic sleep staging based on ambulatory forehead EEG. IEEE Access, 10, 26554-26566.

[14] Wang, H., Guo, H., Zhang, K., Gao, L., & Zheng, J. (2022). Automatic sleep staging method of EEG signal based on transfer learning and fusion network. Neurocomputing, 488, 183-193.

[15] Sri, T. R., Madala, J., Duddukuru, S. L., Reddipalli, R., & Polasi, P. K. (2022, April). A systematic review on deep learning models for sleep stage classification. In 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI) (pp. 1505-1511). IEEE.

[16] You, Y., Zhong, X., Liu, G., & Yang, Z. (2022). Automatic sleep stage classification: A light and efficient deep neural network model based on time, frequency and fractional Fourier transform domain features. Artificial Intelligence in Medicine, 127, 102279.

[17] Zhou, X., Ling, B. W. K., Ahmed, W., Zhou, Y., Lin, Y., & Zhang, H. (2024). Multivariate phase space reconstruction and Riemannian manifold for sleep stage classification. Biomedical Signal Processing and Control, 88, 105572.

[18] Pei, W., Li, Y., Wen, P., Yang, F., & Ji, X. (2024). An automatic method using MFCC features for sleep stage classification. Brain Informatics, 11(1), 1-13.

[19] Liu, J., Ling, B. W. K., Li, R., Shao, J., Lin, S., Che, J., & Liu, Q. (2024). Sleep stage classification via dynamic mode decomposition approach. Signal, Image and Video Processing, 18(1), 535-544.

[20] Pan, J., Yu, Y., Wu, J., Zhou, X., He, Y., & Li, Y. (2024). Deep Neural Networks for Automatic Sleep Stage Classification and Consciousness Assessment in Patients with Disorder of Consciousness. IEEE Transactions on Cognitive and Developmental Systems.

[21] Zhou, D., Xu, Q., Zhang, J., Wu, L., Xu, H., Kettunen, L., ... & Cong, F. (2024). Interpretable Sleep Stage Classification Based on Layer-wise Relevance Propagation. IEEE Transactions on Instrumentation and Measurement.

[22] Yeh, P. L., Ozgoren, M., Chen, H. L., Chiang, Y. H., Lee, J. L., Chiang, Y. C., & Chiang, R. P. Y. (2024). Automatic Wake and Deep-Sleep Stage Classification Based on Wigner–Ville Distribution Using a Single Electroencephalogram Signal. Diagnostics, 14(6), 580.

[23] Heng, X., Wang, M., Wang, Z., Zhang, J., He, L., & Fan, L. (2024). Leveraging discriminative features for automatic sleep stage classification based on raw single-channel EEG. Biomedical Signal Processing and Control, 88, 105631.

[24] B. Kemp, A. H. Zwinderman, B. Tuk, H. A. C. Kamphuisen, and J. J. L. Oberye, “Analysis of a sleep-dependent neuronal feedback loop: The slow-wave microcontinuity of the EEG,” IEEE Trans. Biomed. Eng., vol. 47, no. 9, pp. 1185–1194, Sep. 2000.

[25] A. L. Goldberger et al., “Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals,” Circulation, vol. 101, no. 23, pp. e215–e220, 2000.

International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 12, Issue 2 | Pages 596-604 | e-ISSN: 2347-470X

[26] G.-Q. Zhang et al., "The national sleep research resource: Towards a sleep data common," J. Amer. Med. Inform. Assoc., vol. 25, no. 10, pp. 1351–1358, 2018.

[27] P. Fonseca, N. den Teuling, X. Long, and R. M. Aarts, "Cardiorespiratory sleep stage detection using conditional random fields," IEEE J. Biomed. Health Informat., vol. 21, no. 4, pp. 956–966, Jul. 2017.

[28] Khare S.K. and Bajaj V., 2020. Constrained based tunable Q wavelet transform for efficient decomposition of EEG signals. Applied Acoustics, 163, p.107234.

[29] Thirumalraj, A., & Balasubramanian, P. K. Designing a Modified Grey Wolf Optimizer Based Cyclegan Model for Eeg Mi Classification in Bci.

[30] Thirumalraj, A., Anusuya, V. S., & Manjunatha, B. (2024). Detection of Ephemeral Sand River Flow Using Hybrid Sandpiper Optimization-Based CNN Model. In Innovations in Machine Learning and IoT for Water Management (pp. 195-214). IGI Global.

[31] Al-Baik, O., Alomari, S., Alssayed, O., Gochhait, S., Leonova, I., Dutta, U., ... & Dehghani, M. (2024). Pufferfish Optimization Algorithm: A New Bio-Inspired Metaheuristic Algorithm for Solving Optimization Problems. Biomimetics, 9(2), 65.