SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V10I11P112

Volume 10 Issue 11, 128-141, November 2023

Original Article

Advanced EEG Analysis Based Emotion Recognition: A Deep Learning Classifier with Hybrid Feature Selection and Artifact Reduction

T. Manoj Prasath[1], R. Vasuki[2]

[1,2]*Department of Biomedical Engineering, Bharath Institute of Higher Education and Research, Tamilnadu, India.*

[1]*Corresponding Author : manojprasatht@gmail.com*

Received: 23 September 2023 Revised: 26 October 2023 Accepted: 21 November 2023 Published: 02 December 2023

Abstract *- Emotion recognition from Electroencephalogram (EEG) signals has emerged as a crucial area of research with a broad spectrum of applications. To enhance the accuracy and effectiveness of emotion classification, this study presents an innovative approach that combines advanced signal processing and deep learning techniques. The proposed methodology is structured into several distinct phases for optimal performance. Initially, EEG signals are preprocessed using a Fractional Order Butterworth (FOB) filter, which exhibits flexibility, allowing for precise control over the trade-off between preserving relevant emotional information and mitigating unwanted interference. The Short Time Fourier Transform (STFT) is applied to extract time-frequency representations from the preprocessed EEG signals. This transformation captures dynamic changes in spectral content, providing a comprehensive view of the emotional state over time. A hybrid approach is employed to optimize the feature set and enhance the efficiency of emotion classification. This approach combines the Improved Artificial Fish Swarm algorithm with Particle Swarm Optimization (IAFS-PSO). Combining these algorithms efficiently navigates the solution space to select the most informative features, ensuring the subsequent classification process is based on a highly relevant and discriminative set of features. Emotion recognition is accomplished using a hybrid attention-based Convolutional Neural Network combined with Long Short-Term Memory (CNN-LSTM) networks. The CNN component captures spatial features in the EEG data, while the LSTM component is adapted to modelling temporal dependencies. The hybrid architecture is further enhanced with attention mechanisms, allowing the model to focus on critical segments of the EEG data, thereby improving classification accuracy. The evaluation of this approach is conducted rigorously, and the results of this study highlight the proposed methodology's effectiveness and efficiency in accurately classifying emotional states from EEG signals.*

Keywords *- EEG, Fractional Order Butterworth filter, Hybrid attention CNN-LSTM, Hybrid IAFS-PSO, Short Time Fourier Transform.*

1. Introduction

Brain-Computer Interface (BCI) technology makes the brain's capability to communicate through peripheral devices feasible, profoundly affecting everyday human activities [1]. In Human–Computer Interaction (HCI) and computer intelligence, emotional recognition is an essential realm of research [2]. EEG emotion recognition is a significant BCI technique frequently applied to decision-making, interpersonal communication, and the diagnosis of mental illnesses [3, 4]. It simplifies human-machine interaction in real-world initiatives, and robots can communicate and comprehend human emotions [5]. Since emotional awareness helps track and improve students' academic performance, it has become essential for daily living. It aids in the detection and management of a variety of psychological illnesses in medical studies, including autism spectrum disorders and depression [6]. The complete flow of the EEG recognition of emotions framework is shown in Figure 1.

During the recording stage, the signals must be preprocessed, and the resulting brain signals are often distorted and noisy. The artifacts are eye blinks, eye movements, and heartbeat. Brain signals are also unorganized due to muscular action and power line interferences. Artifacts are removed using Common Average Referencing (CAR) [7], Common Spatial Patterns (CSP) [8-10], Independent Component Analysis (ICA) [11, 12], Principal Component Analysis (PCA) [13], Frequency Evaluation [14, 15].

However, these methods have several limitations, such as the necessity for many computations for decomposition. The average and skull exposure analysis is hindered by the limiting density of samples and inadequate head coverage, and electrode placement changes influence classification accuracy. These concerns are overcome by employing Fractional Order Butterworth (FOB) filter, which is effective for signals and artifacts whenever spectrum overlap.

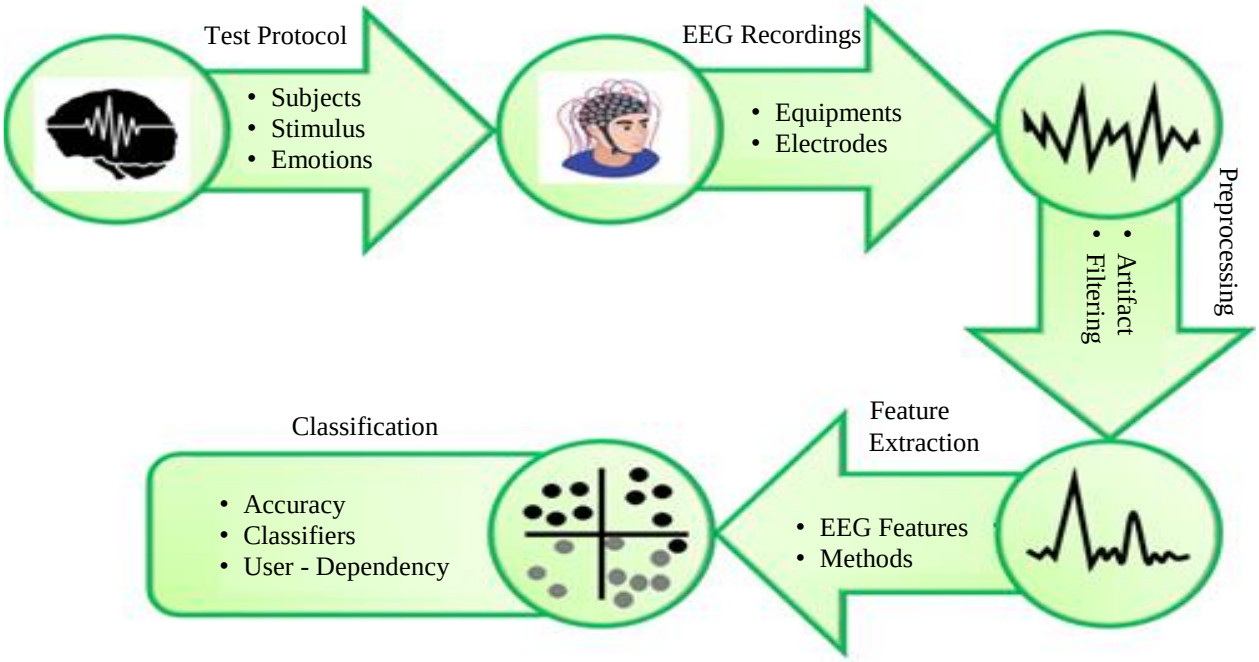

Fig. 1 General flow of human emotion detection

The most significant aspects of the brain signals are collected after obtaining noise-free data from the EEG signal enhancement stage. To extract features from EEG data, techniques which include Adaptive Auto Regressive parameters (AAR) [16], Wavelet Transformations (WT) [17-19], and Wavelet Packet Decomposition (WPD) are implemented [20]. However, these approaches have poor time localization, which leads to failure in every application, and establishing the representation's features for EEG signals is difficult and irrelevant to signals that aren't stationary. As a result, in this study, a Sort Time Fourier Transform (STFT) is proposed, together with a unique IFSA-PSO, to perform efficient feature extraction and feature selection, which efficiently selects the most desirable features.

Following feature extraction and selection, the signals are categorized into several sections using various classifiers. Nearest neighbour classifiers, nonlinear Bayesian classifiers, Support Vector Machine (SVM) [21], Artificial Neural Networks (ANN) [22] and CNN [23] are examples of classifiers for EEG-based emotion detection. However, producing a precise forecast for a suitable category likelihood is very complicated and impossible. The concerns described above are addressed by implementing an attention-based CNN-LSTM classifier. It is also employed to broaden the features of the linear framework that predicts human emotions from brain signals. Henceforth, the significant contribution of the proposed work is summarized as follows:

- Advanced EEG signal analysis based emotion recognition using deep learning algorithm with novel hybrid feature selection approach.
- Adopting the Fractional Order Butterworth (FOB) filter eliminates the EEG artifacts including ocular activity or eye movements, muscle activity, and cardiac activity.
- The time-frequency features are extracted from the preprocessed EEG signal using the Sort Time Fourier Transform (STFT) approach.
- Using the novel hybrid Improved Fish Swarm Algorithm-Particle Swarm Optimization (IFSA-PSO), the optimal features are retrieved for easy classification processing.
- Finally, the proposed Attention-based CNN-LSTM classifier efficiently categorizes human emotions, including valence, arousal and dominance.

2. Proposed System Description

Among the main challenges in emotional computing, emotion recognition offers several application possibilities and significant research value. However, due to its constraints, emotion detection in the actual situation suffers from low accuracy of emotion recognition classification. In response to this issue, this work proposes a deep learning-based expression-EEG emotion identification approach that integrates (Attention based) CNN with LSTM. Figure 2 illustrates human emotion identification using the Attention based CNN-LSTM classifier. The initial component of BCI signal acquisition is responsible for obtaining and recording the signals generated by brain activity. It also forwards these signals to the preprocessing part for improvement and artifact elimination. Various artifacts , such as power line disturbances, eye movement, and skeletal muscle contractions, constantly contaminate the raw EEG data.

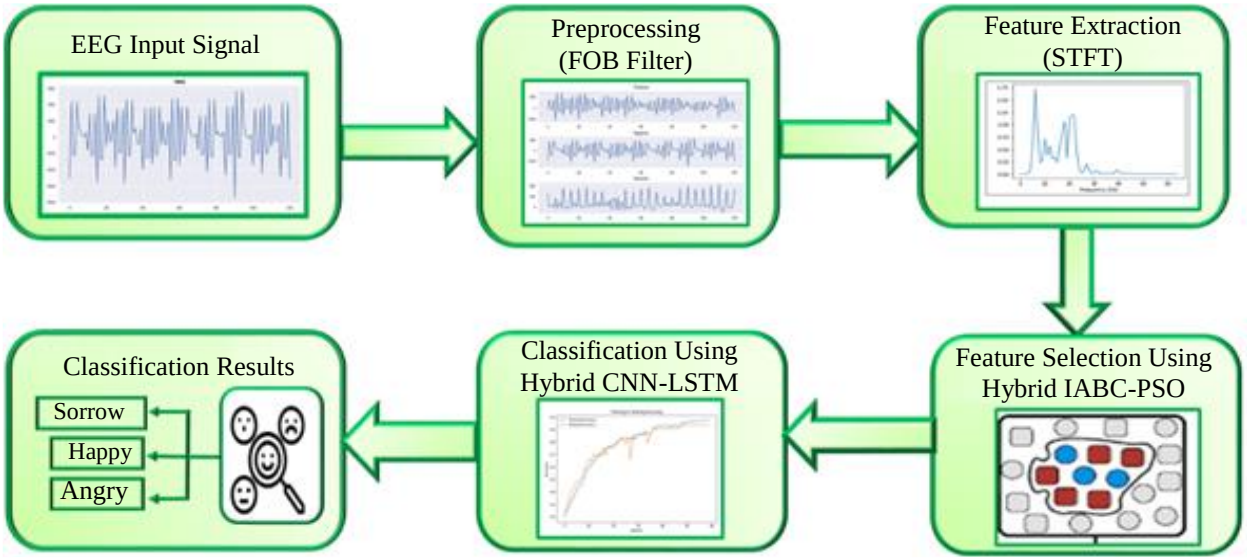

Fig. 2 EEG based emotion detection using attention-based CNN-LSTM

To ensure accurate data analysis, unwanted artifacts are removed, but various noise reduction techniques significantly affect the EEG signal's structure and unique latency, values and amplitude. Consequently, the FOB filter is applied in the present study to minimize noise starting from the EEG data processing stage. EEG signals are divided into gamma, alpha, theta and beta frequency bands employing the FOB filtering.

The STFT approach feature extraction has been used to identify the signals' highest discriminatory properties, including time-frequency. Following that, using the unique hybrid PSO-IFSA algorithm, the foremost desired features from the time-frequency domain are picked. Lastly, the Attention-based CNN-LSTM classifier detects emotions based on the emotion model. This attention-based CNN-LSTM's primary goal is to evaluate the data and identify relevant emotional patterns thoroughly. The following contains detailed descriptions of each part of the proposed system.

2.1. EEG Signal Preprocessing Using Fractional Order Butterworth Filter

The process of transforming the initial EEG signal into clean EEG data by eliminating unwanted noise and artifacts and putting it into a format appropriate for additional analysis is known as EEG data preprocessing. This work proposes a FOB filter to eliminate noises such as muscle, blinking and eye movement, interference with other devices, and power line interference.

Figure 3 depicts the fundamental circuitry for the suggested Fractional Order Butterworth filter structure. A resistor, a fractional inductor & a capacitor with order $\propto, \beta$ and a DVCC as the functional building piece make up the proposed circuit. The subsequent relationships outline the four terminating present mode active analogue components of the DVCC, explaining their terminal current and voltage characteristics.

$$\begin{bmatrix} I_{Y+} \\ I_{Y-} \\ V_X \\ I_Z \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} V_{Y+} \\ V_{Y-} \\ I_X \\ V_Z \end{bmatrix} \quad (1)$$

As a result, the following is the transfer equation of the filter:

$$T(s) = \frac{k_3}{s^{\propto+\beta}+k_1 s^{\propto}+k_2} \quad (2)$$

Where $k_2 = k_3 = 1/LC$ and $k_1 = R/L$. The following is the formula that is characteristic of the system mentioned above:

$$D_{T(j\omega,\sigma,\beta)} = \omega^{(\propto+\beta)} cos\frac{(\propto+\beta)\pi}{2} + k_1\omega^{\propto} cos\frac{\propto\pi}{2} + k_2 + j\left\{\omega^{(\propto+\beta)} sin\frac{(\propto+\beta)\pi}{2} + k_1\omega^{\propto} sin\frac{\propto\pi}{2}\right\} \quad (3)$$

The equivalent cutoff frequency and the power multiplied by the FOB filter reaction are provided by (4) and (5), respectively.

$$|D_{B(j\omega,\sigma,\beta)}|^2 \triangleq K_2^2 + \omega^{2(\propto+\beta)} \quad (4)$$

$$f_c = \frac{1}{2\pi} K_2^{1/(\propto+\beta)} \quad (5)$$

If the squared value of the characteristic formula equals the multiplied value of the proposed filter given in (4), then

the system illustrated in (2) operates like a FOB filter. (6) gives the magnitude squared of the system stated in (3).

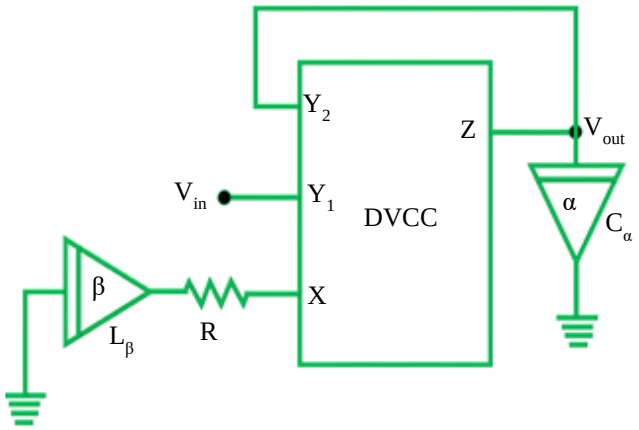

Fig. 3 Layout of FOB filter

$$|D_{T(j\omega,\sigma,\beta)}|^2 = \omega^{2(\propto+\beta)} + K_2^2 + K_1^2\omega^{2\propto} + 2k_1\omega^{2\propto+\beta}cos\frac{\beta\pi}{2} + 2k_2\omega^{\propto}\left\{\omega^{\beta}\,cos\frac{(\propto+\beta)\pi}{2} + k_1 cos\frac{\propto\pi}{2}\right\} \quad (6)$$

Consequently, (7) specifies the conditions under which the system characterized by (4) functions as a FOB filter.

$$K_1^2\omega^{2\propto} + 2k_2\omega^{\propto}\left\{\omega^{\beta}\,cos\frac{(\propto+\beta)\pi}{2} + k_1 cos\frac{\propto\pi}{2}\right\} + 2k_1\omega^{2\propto+\beta}cos\frac{\beta\pi}{2} = 0 \quad (7)$$

As a result, the circumstance of the Butterworth filter relies on both the fractional orders $\propto$ and βas well as on k_1 and k_2 . This increases the amount of independence and adaptability in system layout. At cutoff frequency response, the situation given by (6) is streamlined as indicated by (8).

$$k_1 = \frac{-K_2^{\left(\frac{2\propto+\beta}{\propto+\beta}\right)}\left(2cos\frac{\propto\pi}{2}\right)\pm\left[K_2^{\left(\frac{2\propto+\beta}{\propto+\beta}\right)}\left(2cos\frac{\propto\pi}{2}\right)^2 - 4K_2^{\left(2\frac{2\propto+\beta}{\propto+\beta}\right)}\left(2cos\frac{(\propto+\beta)\pi}{2}\right)\right]^{0.5}}{2K_2^{\left(\frac{2\propto}{\propto+\beta}\right)}} \quad (8)$$

The proposed filter must satisfy both the requirements and be sufficient to exhibit the Butterworth or maximal flat response.

2.2. EEG Signal Feature Extraction Using Short Time Fourier Transform

Analyzing the frequency characteristics of time series is frequently done using the Fourier Transform (FT). It cannot determine when every frequency component emerges; instead, it gives the frequency characteristics aggregated throughout the entire signal time duration. As a result, the frequency domain of two signals that differ significantly in the temporal domain may have the same spectrum. Assuming implies that FT appears to make the incorrect assumption that the time sequences are stationary when it encounters EEG signals. The time series ought to be divided into small sections for these non-stationary signal analyses, and the signal waves inside every segment are essentially conceived of as static signals utilized in FT. The idea, or STFT, involves a series of windowed signal Fourier transforms employed to examine variations in a non-stationary signal's frequency spectrum. Gives time-localized frequency data in scenarios where the frequency elements of a signal change over time.

The definition of the STFT computation is:

$$X(\tau,w) = \int_{-\infty}^{\infty} x(t)\omega(t-\tau)e^{-jwt}dt \quad (9)$$

Where $\omega(t)$ denotes the window function, including the hanning window as illustrated in (10) and $x(t)$ denotes the initial signal. It is a linear arrangement of modulating rectangular windows and is commonly employed when low aliasing and decreased spectrum leakage are necessary.

$$w(n) = \frac{1}{2}\left(1 - cos\left(\frac{2\pi n}{N-1}\right)\right) \quad (10)$$

Where the window length is indicated by n and the sampling number is defined by N. In the case of discrete time series, the data could be separated into subsections. In matrices that store frequency and magnitude at every point in time, the complex result of the Fourier transformation of each segment is added. The following formula can be used to calculate STFT for discrete time series:

$$X(m,w) = \sum_{n=-\infty}^{\infty} x[n]\omega[n-m]e^{-jwn} \quad (11)$$

Where the window function is $\omega[n]$and the time series is $x[n]$.

2.3. EEG Signal Feature Selection Using Hybrid IAFS-PSO Approach

A hybrid technique is deployed to maximize the feature set and enlarge the effectiveness of emotion categorization. This method combines Particle Swarm Optimization (PSO) and the Improved Artificial Fish Swarm Algorithm (IAFS-PSO). By successfully navigating the solution space and choosing the most informative features, the combination of these algorithms ensures that the ensuing classification process is founded on a significantly discriminative and applicable set of features, explained in the fourth section.

2.3.1. Particle Swarm Optimization (PSO)

PSOs are classified as swarm-based metaheuristic optimization techniques. The technique employed in PSO to determine the global optimum is population-based exploring. Moving the particles across the exploration field yields the optimum population that best solves the challenge, inspired by

the birds' actions. Particles are generated in a multidimensional exploration region, and each particle adjusts its location according to its own experiences and those of its adjacent particles. The optimal site that a particle and its neighbouring particles have attained also serves as its guide. The PSO's strengths include its simplicity of use and lack of requirement for numerous variable adjustments. The steps that follow are utilized when establishing a change in the velocity item:

If $x > pbestx, v_x - randxa, otherwise, v_x = v_x + randxa.$

If $y > pbesty, v_y - randxa, otherwise, v_y = v_y + randxa.$

The best position ever attained is denoted by $pbest$. A is the constant that adjusts velocity, and $rand$ is a random number in the interval [0, 1]. The following criteria need to be followed when updating the velocity.

If $x > gbestx, v_x - randxb, otherwise, v_x = v_x + randxa.$

If $y > gbesty, v_y - randxb, otherwise, v_y = v_y + randxa.$

Where $gbest$ is the current best location of the entire swarm. The velocity-adjusting constant is denoted by b.On the other hand, slow convergence and local optima trapping are PSO's two critical problems that adversely affect its performance.

Moreover, this method drastically underperforms on high-dimensional issues. In this work, IFSOA is combined with PSO to solve the drawbacks of both algorithms, including low optimization accuracy and the incapacity to consider both local and global information.

2.3.2. Improved Fish Swarm Optimization Algorithm

The global search capabilities of the Artificial Fish Swarm algorithm are limited by its later stages' delayed convergence, resulting in many fake fish doing pointless searches and wasting time. Considering this, a better algorithm is proposed based on swallowing behaviour. The population's range describes the level of dispersal of the fisher swarm. A representation of diversity is going to be $\alpha = [\alpha_1, \alpha_2 \ldots . \alpha_n)$. In the area where$\alpha_i$ is fake fish exhibit adaptive variety.

$$\alpha_i = {min(f, f_{avg})}/{max(f, f_{avg})}, \quad \alpha_i \in (0.1), i = 1,2, \ldots . , n \tag{12}$$

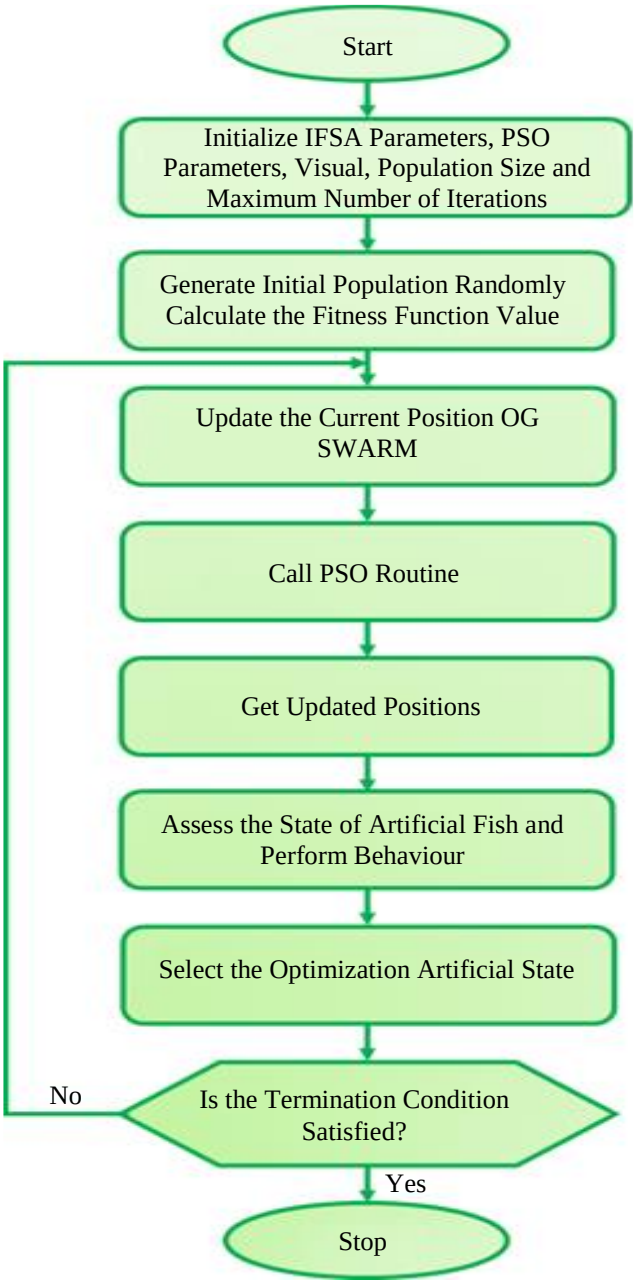

Fig. 4 Flowchart of IFSA-PSO

Where f is the fitness value of the present fish and f_{avg} denotes the average fitness value of the ith iteration.

Swallowed Behaviour

Suppose a fish's diversity value falls below a threshold after a predetermined number of repetitions (say, half of the maximum number of iterations). In that case, the fish remains still, and its space is freed. $if(\alpha_i > Threshold\ 1)$ perform swallowed behaviour; here, the $Threshold\ 1$ represents the diversity threshold.

Breeding Behaviour

The quantity of fish can be decreased by swallowed behaviour, which can significantly shorten the time. However, swallowed behaviour might consume quality fish, making it impossible to identify globally optimal. Thus, an enhanced algorithm that considers ingesting behaviour is put forward.

If($\alpha_i <$ Threshold 2) if($\alpha_i >$ Threshold 1) perform breeding behaviour; here, the Threshold 1 represents the diversity threshold. Locate the fake fish max. X_{max} The most significant goal function value by traversing the created artificial fish. if $Y_i < Y_{max}$, then advance one step to the largest fish of subclass max X; if not, advance one step to the centre of subclass X_c. The hybrid IFSA-PSO algorithm flowchart is represented in the Figure 4. The following section provides a detailed description of the proposed Attention-based CNN-LSTM classifier for efficient human emotion detection.

2.4. EEG Signal Classification Using Hybrid Attention Based CNN-LSTM Method

The proposed classifier comprises the prolonged attention processes, the CNN & LSTM, and the channel-wise attention approach. Figure 5 depicts the architecture of the suggested attention-based CNN-LSTM. The spatial extraction of the features module is displayed on the opposite side of the illustration. Subsequently, the attention method is channel-by-channel into EEG signals to investigate the relative relevance of the various channels of wideband EEG signals (Figure 5). Various EEG in multichannel systems frequently include redundant information during the actual EEG signal collecting process. Specific techniques leverage choosing the channel to pick appropriate channels to increase emotion recognition accuracy.

Despite conventional methods that require artificially selecting significant channels, the intended dynamic channel-wise system can assess each channel's value by considering its information. The EEG samples following processing are represented by $S = \{S_1, S_2, \ldots\ldots, S_n\}$ in the proposed framework, and the $i-th$ EEG sample is represented by $S_i = \{S_1, S_2, \ldots\ldots, S_m\}$ (i = 1,2, … , m) where $S_j (j = 1,2, \ldots, m)$ indicates the $j-th$ channel of EEG signal m and S_i is the overall amount of every sample. To acquire channel-wise characteristics for this framework, initially, the mean pooling is applied for every channel of the EEG sample, as shown below:

$$s^- = [s_1^-, s_2^-, \ldots\ldots\ldots, s_m^-] \quad (13)$$

Where the average of the $j-th$ channel is represented by $s_j^- (j = 1, 2, \ldots, m)$. The channel-based attention system uses Fully Connected (FC) layers centred on non-linearity to minimize complexity while boosting generalization. These layers are a dimension reduction variable. W_1 regards b_1 with elimination proportion r, tanh is activation function and dimension enhancing layer variable W_2 and bias regards b_2. As a result, the channel-based attention-regulating process is stated as follows:

$$v = \text{softmax}(W_2 . (\tanh(W_1 . s^- + b_1) + b_2)) \quad (14)$$

When the distribution likelihood $v = [v1, v2, \ldots, vm]$, which reflects the significance of various channels, is transformed into the importance of channels via the softmax function. In the end, we use likelihood as the weight to recode every channel's data from the EEG sample $Si = [s1, s2, \ldots, sm]$. Consequently, the following can be used to indicate the $j-th (j = 1, 2, \ldots, m)$ attentive channel feature that has been retrieved using channel-wise Attention:

$$c_j = v_j . s_j \quad (15)$$

As a result, $C = \{C_1, C_2, \ldots\ldots, C_n\}$ indicates the acquired channel-wise responsive features. By channels-wise multiplying every component of $v = \{v_1, v_2, \ldots\ldots, v_m\}$ and every channel of $S_i = \{S_1, S_2, \ldots\ldots, S_m\}$One can derive the $i-th$ retrieved feature,$C_i = [C_1, C_2, \ldots\ldots, C_m]$.The spatial data of the EEG signals is subsequently extensively extracted using CNN, whereby the amount of convolution kernels is k, and the amount of electrodes is identical for the kernel height. In this case, the kernel width is similarly intended to investigate the EEG signals' temporal data. Furthermore, the activation function in the convolution processes is the Exponential Linear Unit (ELU) function rather than the Rectified Linear Unit (ReLU) function, which is more commonly used. Therefore, following convolution and activation procedures, the $i-th$ channel responsive feature C_i can be used to produce the feature with the value C_i' $(i = 1, 2, \ldots., n)$. After that, to minimize the number of variables and retrieve more features, a pooling layer is implemented.

The temporal feature extraction unit (Figure 3), which consists of an enhanced self-attention process and a two-layer LSTM, is displayed on the opposite side of the configuration diagram. Since the pattern is built on a recurrent construction, the LSTM network can learn its surroundings. Since the LSTM network can obtain features from EEG data according to temporal dependence, it has been effectively employed for EEG emotion identification. The output at every interval can be seen as the temporal data retrieved from every sample. The number of LSTM layers was selected to two because the LSTM network uses two successive layers to recall and transmit all detected temporal and spatial domains. The hidden state variables of the subsequent layer, $\{ h_i' | h_i' = \text{LSTM}(Q_i), i = 1 \ldots n \}$, are thus the $i-th$ resultant of the LSTM network. By calculating similarities across every sample from various locations may more accurately represent the precise meaning. The resultant z_i' ought to be thought of as a rating vector from $i-th$ sample h_i'.

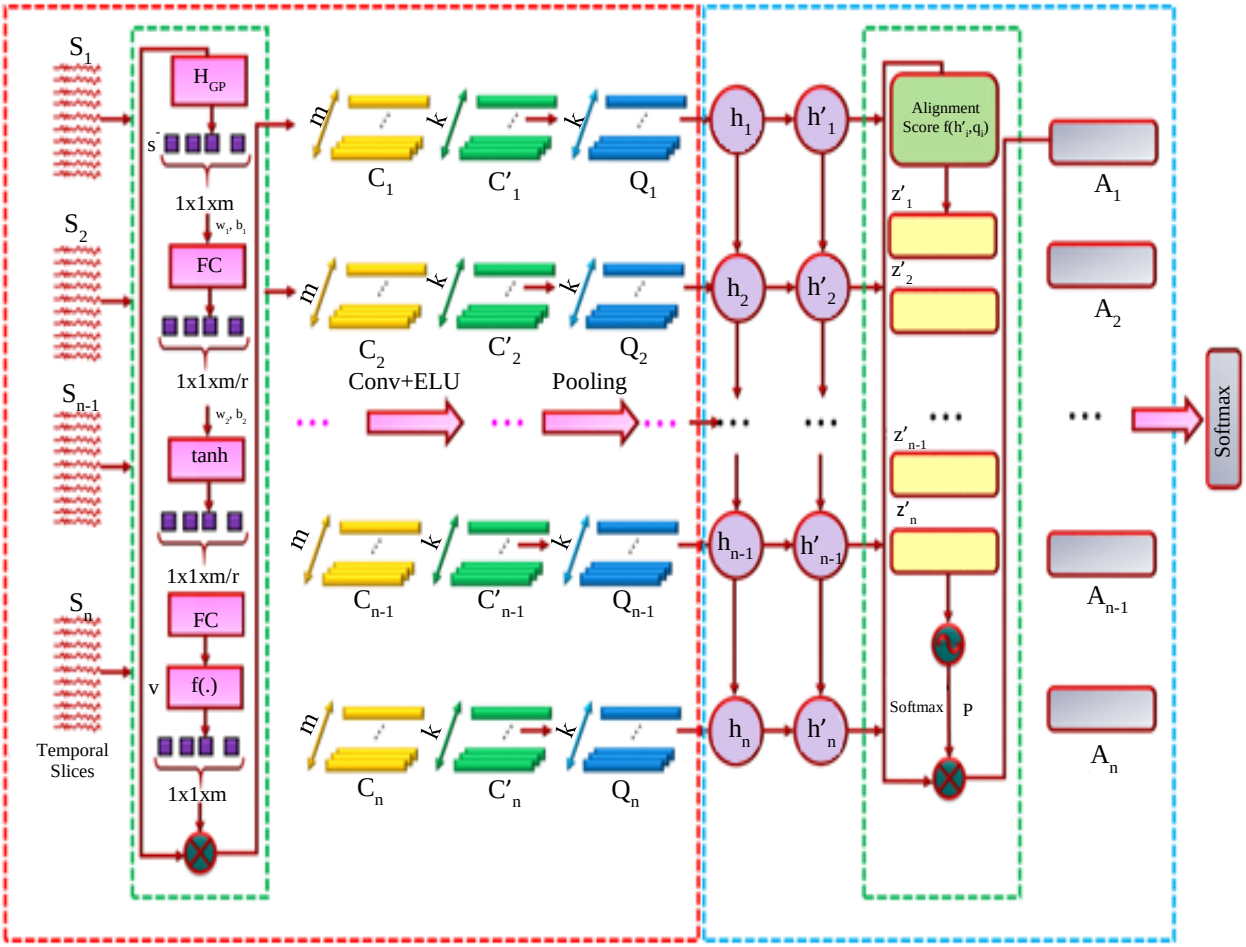

Fig. 5 Attention-based hybrid CNN-LSTM

Furthermore, the prolonged self-attention process adds extra bias terms to each side of the activation function. As a result, $i-th$ feature vector, z_i' is represented as follows:

$$z_i' = f(h_i', q_i) = W^T \sigma(W_1 h_i^1 + W_2 q_i + b_1) + b, \quad (16)$$

Where q_i is the oriented structure vector produced using the feature vector h_i^1 via linear alteration, whereby the dimension is identical to the vector of features, and $f(h_i', q_i)$ reflects the inherent similarities of the $i-th$ encoded EEG sample. In this case,σ, the weighted and bias components of the function are denoted by $W$and b; accordingly, activation function $\sigma(.)$ is an ELU, the weight parameters are W_1, $W_1$and bias terms are b_1. The probabilities of each sample are then indicated by $p = \{p_1, p_2, \dots p_n\}$, and the likelihood of $i-th$ sample is given by the following expression:

$$p_i = \frac{exp(z'^1_i . h_i^1)}{\sum_{i=1}^{n} exp(z'^1_i . h_i^1)} \quad (17)$$

Finally, the enhanced self-attention mechanism's retrieved characteristics are represented by the symbol $A = \{A_1, A_2 \dots, A_n\}$. The following formula is used to determine the $i-th$ attentive feature that the mechanism retrieved:

$$A_i = p_i . h_i^1 \quad (18)$$

The softmax layer classifier is employed in the final section of the attention-based CNN-LSTM. $A = \{A_1, A_2 \dots, A_n\}$ are the retrieved spatiotemporal attention features. The softmax function uses the takeout characteristics as input to identify emotional states in the following ways:

$$P = softmax(WA + b) \quad (19)$$

Where the estimated likelihood of $i-th$ EEG sample is represented by P = {P_1, P_2,,…P_n}, P_i (i=1, 2,…n), b and w are bias and weight elements of the softmax function, correspondingly. Subsequently, the entropy error is assessed for every labelled sample:

$$L = -\sum_{i=1}^{n} \widehat{Y}_i \, log(p_i), \quad (20)$$

Therefore, a channel-wise attention system has been developed to continually allocate the weights of various channels to obtain the inherent data between channels. The spatial data of encrypted EEG data is subsequently retrieved using CNN. Furthermore, the deployed LSTM is used to investigate distinct EEG data's sequential characteristics. Furthermore, a more robust attention system was integrated to allocate weight to EEG data according to their relative significance. Lastly, spatiotemporal attention characteristics for EEG emotion recognition are classified efficiently.

3. Results and Discussion

The current study recognizes and classifies various human emotional states by analyzing EEG signals and the DL approach. The first step in removing the artifacts from the EGG signals is to apply the FOB filter technique. With the use of hybrid IFSA-PSO and STFT, the best features are retrieved and selected appropriately. Consequently, accurate emotion classification is achieved with the use of attention-based CNN-LSTM. The proposed approach has been put into practice on the Python platform to test its functionality, and the findings are described in the section below. It is a brainwave dataset based on EEG, which collects the 32-channel EEG, skin temperature, respiration, galvanic skin response, blood volume pulse, and electromyography from 27 student participants. From 1600 video clips, including cinema, TV programs, and TV news, 28 random video clips containing three categories of continuous emotion (arousal, valorance, and dominance) were chosen randomly. The experiment is split into two sessions, separated by a minimum of 24 hours. Each participant watched 14 video segments in each session, resulting in 14 trials. According to Table 1, the proposed work correctly predicts a variety of emotions, such as arousal (intensity of emotion), valence (pleasantness), and dominance (degree of control). Figure 6 shows the waveform of the input EEG signal, and similarly to what is seen in Figure 7, positive values in the EEG signal create descending waves, and negative values cause upward spirals.

Table 1. Standardized emotional EEG datasets

Data Set	Participants	Stimulus	Obtained Data Modalities	Method	Accuracy	Qualification of Emotion
DREAMER [24]	25 (11 female, 14 male)	18 video clips	14- Channel EEG, Electrocardiogram are peripheral physiological signals.	SVM	71.12%	Continuous type (Arousal, Valance, Dominance
SEED [25]	15(8 female, 7 male)	72 video clips	Data from an EEG with 62 channels and eye tracking	DNN	72.39%	Discrete type (happy, fear, sorrow and neutral)
MPED [26]	23 (13 female, 10 male)	28 video clips	Peripheral physiological signals (Electrocardiogram, breathing, Galvanic Skin Response) and 62-channel EEG	SVM,KNN	80.45%	Discrete type (anger, joy, fear, funny, sad, disgust and neutral)
EEG based on the brainwave data set	27 (12 male, 15 female)	29 video clips	Peripheral physiological data (blood volume pulse, galvanic skin response, respiration, skin temperature, electromyography), 32-channel EEG.	Attention-based CNN-LSTM	96.8%	Continuous type (Arousal, Valance, Dominance)

Artifacts are signals recorded by EEG that are not generated by the brain. Certain artifacts might mimic natural epileptiform aberrations or seizures. The FOB filter is then used to extract the movement artifact, which may be removed from the ECG signal by removing the obtained one, as shown in Figures 8 and 9, respectively. To begin, spectrogram data is obtained from EEG signals of depressive and physically fit people utilizing Short-Time Fourier Transform (STFT) to identify significant time-frequency features as seen in Figure 10. This alteration records fluctuation in spectral content, offering a whole perspective of the emotional state throughout time. The spectrogram data acquired from STFT is subsequently employed as input to the classification model.

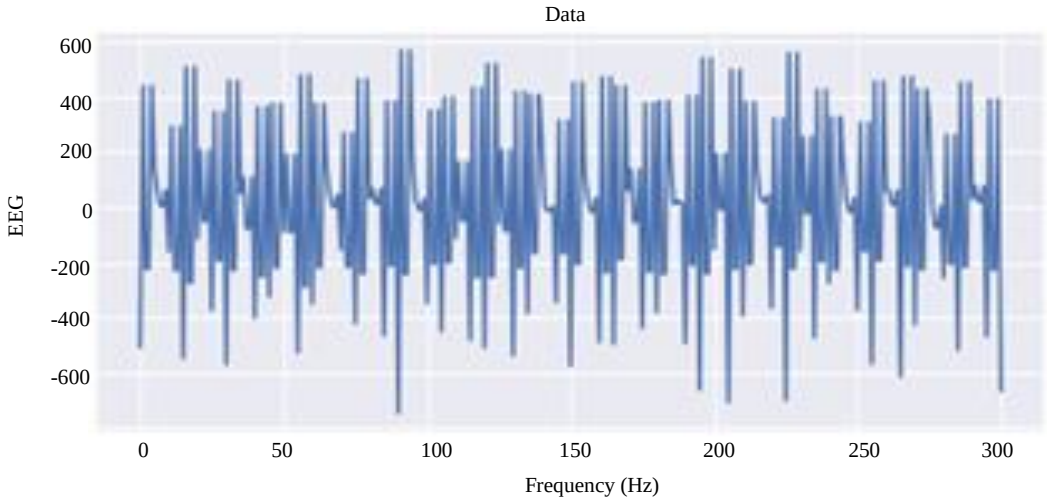

Fig. 6 Input EEG signal

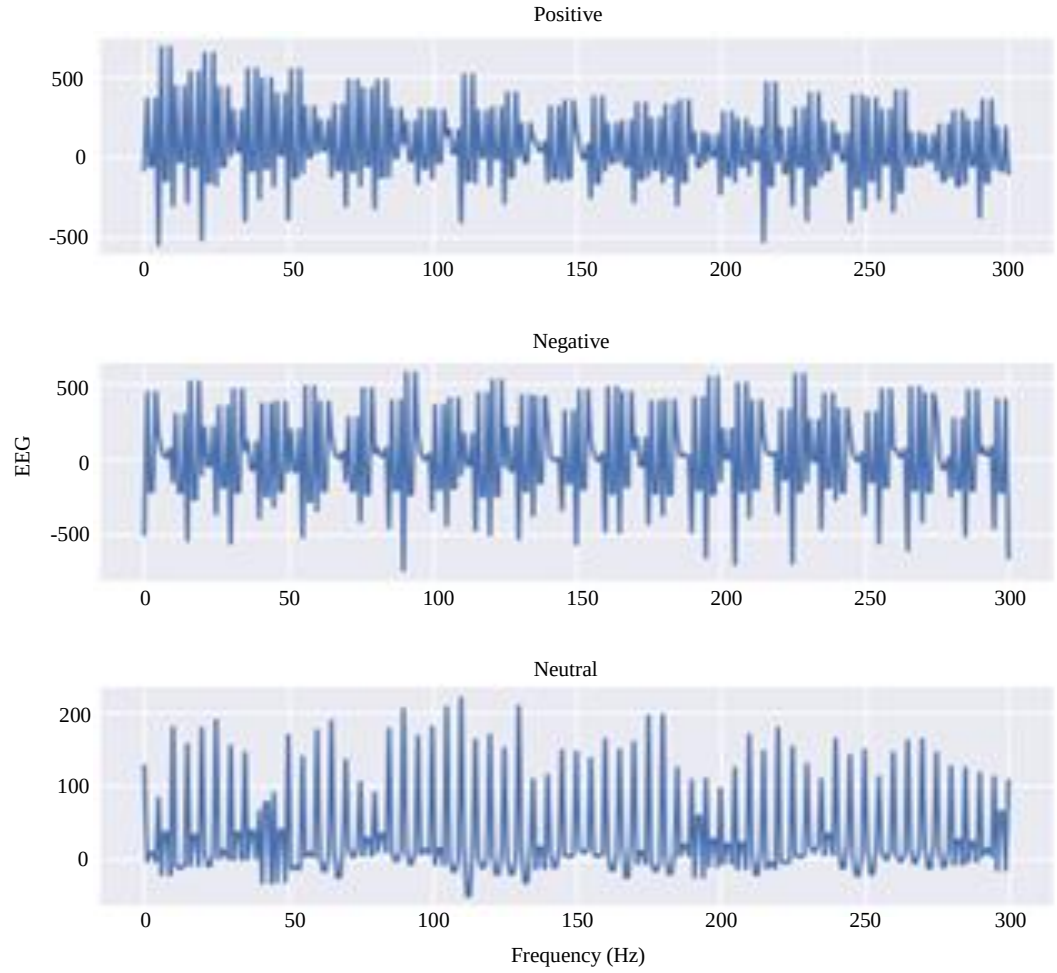

Fig. 7 Input EEG signal's positive, negative and neutral waveform

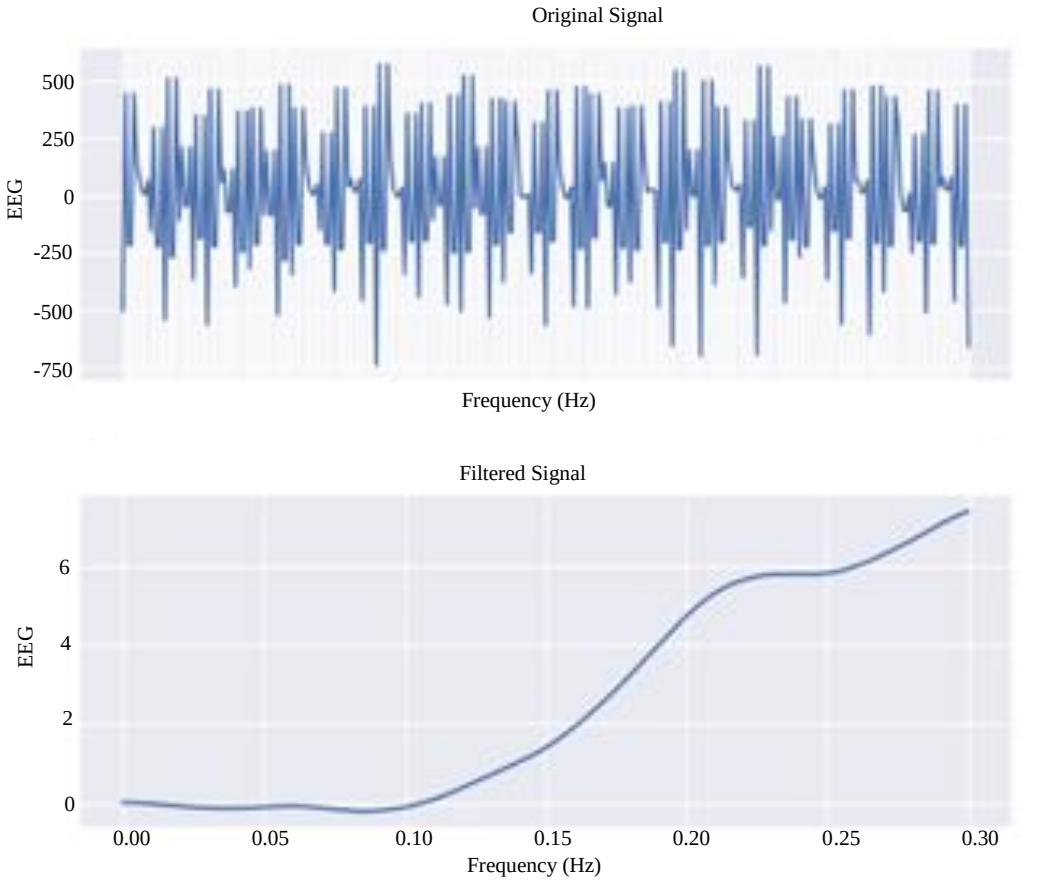

Fig. 8 Filtered EEG signal

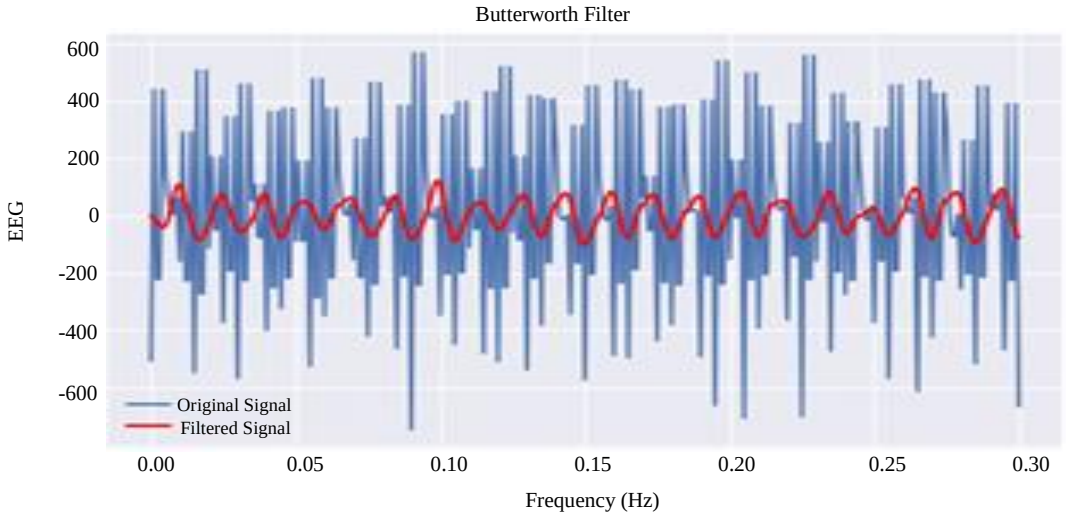

Fig. 9 Artifact removing processing FOB filter

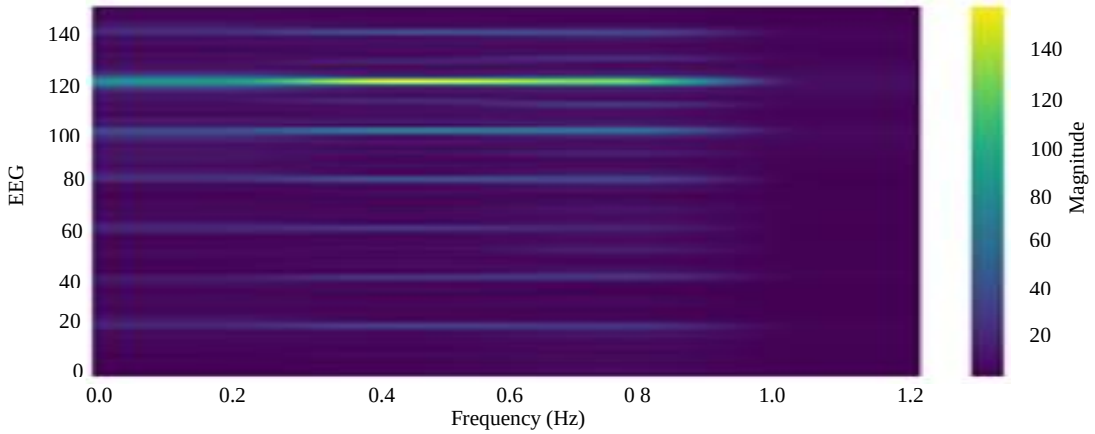

Fig. 10 Feature extraction using the STFT approach

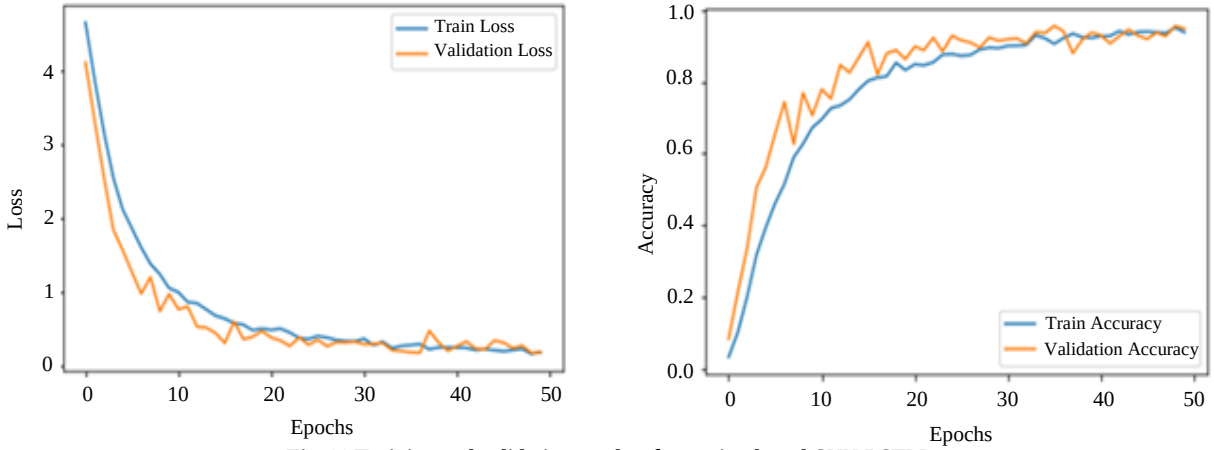

Fig. 11 Training and validation results of attention-based CNN-LSTM

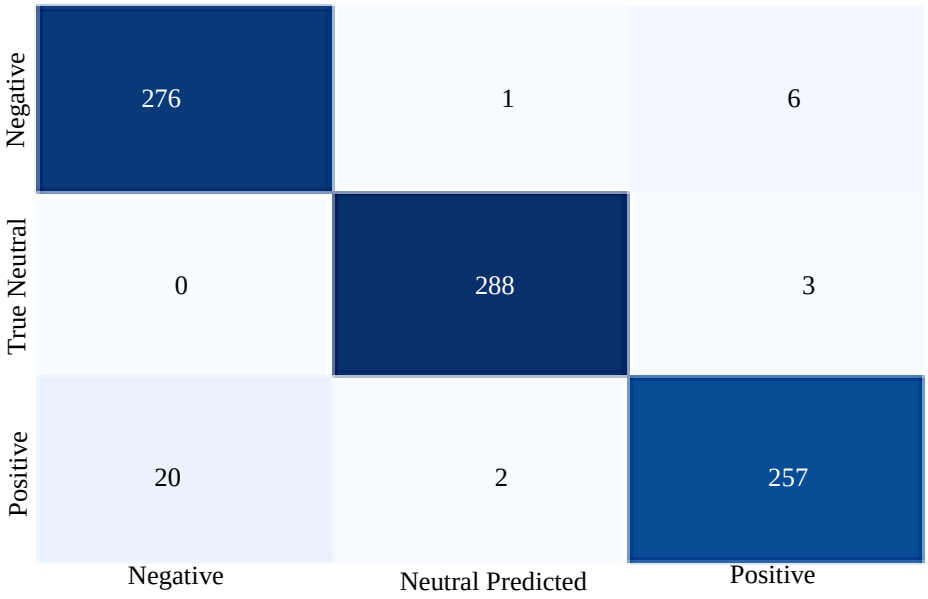

Fig. 12 Confusion matrix

Figure 11 indicates the results of training and validation of Attention based CNN-LSTM; from that, the proposed classifier obtains the highest accuracy value of 96.8 % with decreased loss. The confusion matrix related to the tests employing the brain wave dataset is displayed in Figure 12, based on the outcomes of recognizing every state of mind.

In the confusion matrix, each row represents the target class, while each column represents the predicted class. According to outcomes, neutral emotion can be recognized with excellent precision overall; however, negative emotion is more challenging to identify and often confused with good feeling.

Table 2. Outcome of the proposed FOB filter

Order		Max. Pass-Band Error of Fractional Order Filter	
α	β	In ref [27]	FOB
1.1	0.6	9 dB	7 dB
0.9	0.9	12 dB	10 dB
0.6	0.6	1.25 dB	1 dB
0.5	1	5 dB	4 dB

Furthermore, the FOB filter's effectiveness contrasts with the conventional filter concerning the maximum gain error. Table 2 illustrates the outcome of this ratio. It is discovered that the highest pass band magnitude error for the FOB filter is lower in every instance than for the previous one, so the proposed FOB filter is better than the current design.

Table 3. Accuracy comparison

Classifier	Accuracy
KNN [28]	86.75 %
DBN [29]	87.62 %
MLP [30]	78.11%
Proposed Attention Based CNN-LSTM	96.8 %

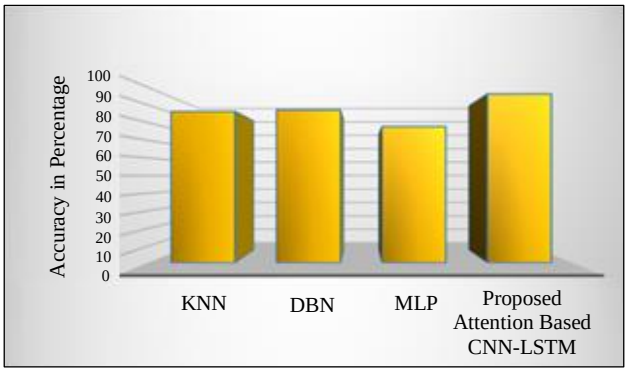

Fig. 13 Accuracy comparison

Table 3 compares proposed Attention based CNN-LSTM to different classifiers, such as KNN [28], DBN [29] and MLP [30], and Figure 13 depicts the related graph. According to the data, the proposed classification model has a fantastic accuracy of 96.8% for forecasting human emotions. Similarly, according to Table 4, the attention-based CNN-LSTM surpasses previous algorithms regarding specificity, FI, precision, sensitivity, and accuracy.

Table 4. Classification outputs of EEG signal validation and training (valence and arousal)

Classifier	Valance			Arousal		
	F1	Precision	Recall	F1	Precision	Recall
Proposed	95.56	95.8	95.8	95.47	95.87	95.99
KNN [3]	93.32	93.09	93.18	93.35	93.60	93.95
DT [3]	92.06	92.65	91.23	91.28	91.22	91.56
NB [3]	91.72	91.93	92.51	92.80	92.32	92.62

4. Conclusion

This work proposes an innovative strategy that employs advanced signal processing and deep learning approaches to enlarge the accuracy and efficacy of emotion classification.

Initially, EEG signals are preprocessed with a FOB filter, which is sufficiently adaptable to regulate the balance between conserving significant emotional information and minimizing undesired interference. The STFT is used to derive time-frequency representations from EEG signals that have been preprocessed. A hybrid IAFS-PSO strategy is employed to optimize the feature set and boost the efficiency of emotion categorization. Hybrid Attention-based CNN-LSTM networks are used to recognize emotions. The CNN component identifies spatial characteristics in EEG data, whereas the LSTM component is skilled at modelling temporal dependencies.

The Python software is used to create an evaluation framework that analyzes the proposed model's specificity, sensitivity, and accuracy. Based on the findings, the proposed strategy is adequate to predict human emotions, and the proposed attention-based CNN-LSTM strategy has a higher accuracy of 96.8%.

References

[1] Theerawit Wilaiprasitporn et al., "Affective EEG-Based Person Identification Using the Deep Learning Approach," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 12, no. 3, pp. 486–496, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[2] Md. Rabiul Islam et al., "Emotion Recognition from EEG Signal Focusing on Deep Learning and Shallow Learning Techniques," *IEEE Access*, vol. 9, pp. 94601–94624, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[3] Rania Alhalaseh, and Suzan Alasasfeh, "Machine-Learning-Based Emotion Recognition System Using EEG Signals," *Computers*, vol. 9, no. 4, pp. 1-15, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[4] Aya Hassouneh, A.M. Mutawa, and M. Murugappan, "Development of a Real-Time Emotion Recognition System Using Facial Expressions and EEG Based on Machine Learning and Deep Neural Network Methods," *Informatics in Medicine Unlocked*, vol. 20, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[5] Alireza Samavat et al., "Deep Learning Model with Adaptive Regularization for EEG-Based Emotion Recognition Using Temporal and Frequency Features," *IEEE Access*, vol. 10, pp. 24520–24527, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[6] Wei Tao et al., "EEG-Based Emotion Recognition via Channel-Wise Attention and Self-Attention," *IEEE Transactions on Affective Computing*, vol. 14, no. 1, pp. 382-393, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[7] Syahrull Hi Fi Syam et al., "Comparing Common Average Referencing to Laplacian Referencing in Detecting Imagination and Intention of Movement for Brain Computer Interface," *MATEC Web of Conferences*, *2017 International Conference on Emerging Electronic Solutions for IoT (ICEESI 2017)*, vol. 140, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[8] Vasilisa Mishuhina, and Xudong Jiang, "Feature Weighting and Regularization of Common Spatial Patterns in EEG-Based Motor Imagery BCI," *IEEE Signal Processing Letters*, vol. 25, no. 6, pp. 783–787, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[9] Simon Geirnaert, Tom Francart, and Alexander Bertrand, "Fast EEG-Based Decoding of the Directional Focus of Auditory Attention Using Common Spatial Patterns," *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 5, pp. 1557–1568, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[10] P. Rithwik, V.K. Benzy, and A.P. Vinod, "High Accuracy Decoding of Motor Imagery Directions from EEG-Based Brain Computer Interface Using Filter Bank Spatially Regularised Common Spatial Pattern Method," *Biomedical Signal Processing and Control*, vol. 72, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[11] Sebastian Michelmann et al., "Data-Driven Re-Referencing of Intracranial EEG Based on Independent Component Analysis (ICA)," *Journal of Neuroscience Methods*, vol. 307, pp. 125–137, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[12] Souvik Phadikar, Nidul Sinha, and Rajdeep Ghosh, "Automatic EEG Eyeblink Artifact Identification and Removal Technique Using Independent Component Analysis in Combination with Support Vector Machines and Denoising Autoencoder," *IET Signal Processing*, vol. 14, no. 6, pp. 396–405, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[13] Fiorenzo Artoni, Arnaud Delorme, and Scott Makeig, "Applying Dimension Reduction to EEG Data by Principal Component Analysis Reduces the Quality of Its Subsequent Independent Component Decomposition," *NeuroImage*, vol. 175, pp. 176–187, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[14] Qiang Gao et al., "EEG Based Emotion Recognition Using Fusion Feature Extraction Method," *Multimedia Tools and Applications*, vol. 79, pp. 27057–27074, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[15] Peiyang Li et al., "EEG Based Emotion Recognition by Combining Functional Connectivity Network and Local Activations," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 10, pp. 2869–2881, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[16] Xin Chai et al., "A Fast, Efficient Domain Adaptation Technique for Cross-Domain Electroencephalography (EEG)-Based Emotion Recognition," *Sensors*, vol. 17, no. 5, pp. 1-21, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[17] Yong Zhang, Suhua Zhang, and Xiaomin Ji, "EEG-Based Classification of Emotions Using Empirical Mode Decomposition and Autoregressive Model," *Multimedia Tools and Applications*, vol. 77, pp. 26697–26710, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[18] Vipin Gupta, Mayur Dahyabhai Chopda, and Ram Bilas Pachori, "Cross-Subject Emotion Recognition Using Flexible Analytic Wavelet Transform from EEG Signals," *IEEE Sensors Journal*, vol. 19, no. 6, pp. 2266–2274, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[19] Arti Anuragi, Dilip Singh Sisodia, and Ram Bilas Pachori, "EEG-Based Cross-Subject Emotion Recognition Using Fourier-Bessel Series Expansion Based Empirical Wavelet Transform and NCA Feature Selection Method," *Information Sciences*, vol. 610, pp. 508–524, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[20] Jingxia Chen, Dongmei Jiang, and Yanning Zhang, "A Common Spatial Pattern and Wavelet Packet Decomposition Combined Method for EEG-Based Emotion Recognition," *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 23, no. 2, pp. 274–281, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[21] Kranti S. Kamble, and Joydeep Sengupta, "Ensemble Machine Learning-Based Affective Computing for Emotion Recognition Using Dual-Decomposed EEG Signals," *IEEE Sensors Journal*, vol. 22, no. 3, pp. 2496–2507, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[22] S. Thejaswini et al., "EEG Based Emotion Recognition Using Wavelets and Neural Networks Classifier," *Cognitive Science and Artificial Intelligence, Springer Briefs in Applied Sciences and Technology*, pp. 101–112, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[23] Heng Cui et al., "EEG-Based Emotion Recognition Using an End-to-End Regional-Asymmetric Convolutional Neural Network," *Knowledge-Based Systems*, vol. 205, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[24] Stamos Katsigiannis, and Naeem Ramzan, "DREAMER: A Database for Emotion Recognition through EEG and ECG Signals from Wireless Low-Cost Off-the-Shelf Devices," *IEEE Journal of Biomedical & Health Informatics*, vol. 22, no. 1, pp. 98–107, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[25] Wei-Long Zheng et al., "Emotionmeter: A Multimodal Framework for Recognizing Human Emotions," *IEEE Transactions on Cybernetics*, vol. 49, no. 3, pp. 1110–1122, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[26] Tengfei Song et al., "MPED: A Multi-Modal Physiological Emotion Database for Discrete Emotion Recognition," *IEEE Access*, vol. 7, pp. 12177–12191, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[27] Shalabh Kumar Mishra, Maneesha Gupta, and Dharmendra Kumar Upadhyay, "Active Realization of Fractional Order Butterworth Lowpass Filter Using DVCC," *Journal of King Saud University-Engineering Sciences*, vol. 32, no. 2, pp. 158–165, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[28] Mi Li et al., "Emotion Recognition from Multichannel EEG Signals Using K-Nearest Neighbor Classification," *Technology and Health Care*, vol. 26, no. S1, pp. 509–519, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[29] Wei-Long Zheng et al., "EEG-Based Emotion Classification Using Deep Belief Networks," *2014 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1-6, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[30] Adnan Mehmood Bhatti et al., "Emotion Recognition and Analysis in Response to Audio Music Using Brain Signals," *Computers in Human Behavior*, vol. 65, pp. 267–275, 2016. [CrossRef] [Google Scholar] [Publisher Link]