Journal of Intelligent Computing System (JICS), Vol.1, Issue. 1, pp. 16-36, (Jan 2026)
ISSN: 3139-3616 (online)

EEG-Based Emotion Detection by Bio-Inspired Optimization of LSTM using Artificial Bee Colony and Firefly Algorithms

Gaganjot Kaur, Meenu Gupta, Rakesh Kumar

Department of Computer Science and Engineering, Chandigarh University, Mohali, Punjab, India
sandhu.gaganjot@gmail.com, gupta.meenu5@gmail.com, rakesh77kumar@gmail.com

Abstract: Recognizing human emotions via brain signals is one of the latest advancements in the field of bio-signalling technology. Electroencephalography (EEG) is one of the cost-effective technologies that analyze the biomedical signals generated by a human brain that can be used to interpret human emotions. The primary objective of the paper is to present a multistage emotion classification based on EEG signals. In the initial stage, time and time-frequency-based feature extraction is performed followed by Artificial Bee Colony (ABC) based feature selection in the second stage. Finally, the optimized feature set is used for the training, and classification of the emotions is performed using Long Short-Term Memory (LSTM). Here firefly algorithm is used for the hyperparameters tuning of the classification model. The framework is evaluated against state of art classifiers. Evaluated using the Brainwave dataset, the ABC-F-LSTM model outperforms existing classifiers, achieving a recall of 0.9896. It shows notable improvements of approximately 6.79% over ABC-LSTM, 19.39% over Deep Neural Networks (DNN), and 21.64% over Random Forest, highlighting its superior accuracy and robustness in emotion prediction.

Keywords: *Electroencephalography, Emotion Prediction, Artificial Bee Colony, Machine Learning.*

1. Introduction

In the field of human emotions, the way we think, act and feel has attracted large section of philosophers, psychologists, and scientists all over the globe. Research community have investigated various methods, including brain training, to eliminate the uncertainty in thinking. Recent decades have seen tremendous developments in the field of Neuroscience and the utilization of Neuroimaging Methods for analyzing emotional states of the human brain [1], 2]. Electroencephalography is one such device that allows researchers and clinicians to measure the cognitive processing of different types of emotion through a variety of electrical signals emitted from the brain [3]. EEG (also known as Electroencephalography), is a non-invasive form of Electrical Neurophysiology whereby researchers measure the electrical activity occurring within the human brain. During an EEG procedure, electrodes are placed on the patient's scalp in order to measure the electrical current generated when neurons in different areas of the patient's brain become active.

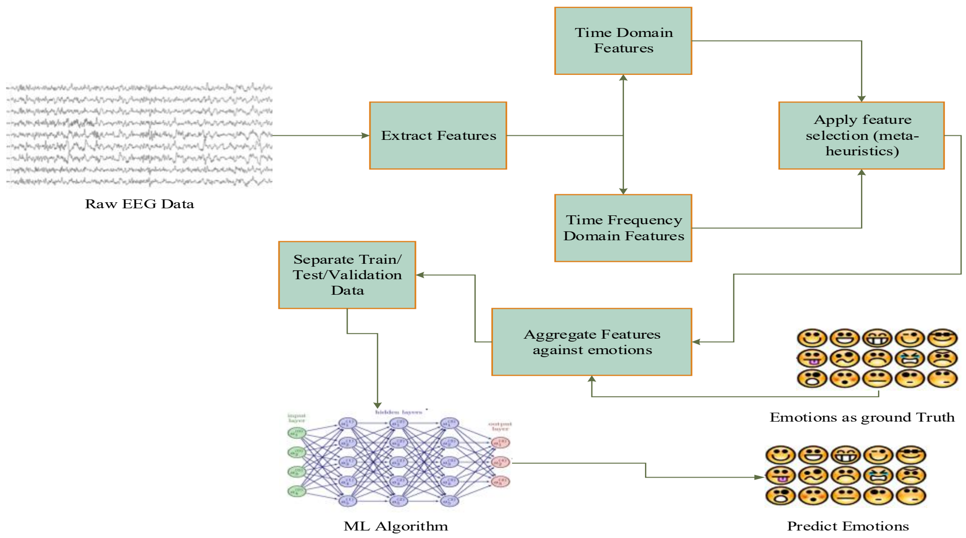

Figure 1. Emotion Recognition Process via. EEG and ML

doi.org/10.67420/109319.1.1.2

These recordings provide a great insight into the quality of brain activity and provide a unique window into the neural processes underlying emotion [4]. EEG utilization has many use cases across several domains, e.g., Psychology, Neuroscience, HCI, Clinical use, etc., due to its great potential for emotion recognition [5]. Nevertheless, many research works have proven that the brain structure differs depending on the varied emotions experienced by an individual [6],[7]. This difference is observed in the EEG records taken of individuals when they are in feelings. Due to this trend in the identification, researchers have been in a position to develop models which categorise the various emotions people undergo basing their classification on EEG records [8]. The utilisation of ML algorithms and the pattern of EEG signals has come up with an effective tool of identifying emotions by evaluation of EEG signals [9]. The integration of the ML algorithms with EEG patterns will enable the researchers to develop strong models to identify emotions based on the data that has already been gathered in the past [10],[11]. Lastly, as depicted in figure 1, the evolution of ML models will enable researchers to create and train their models with data that comprise of emotional cues that will enable them to categorise and anticipate emotional conditions through EEG signals [12].

2. Literature Review

It is suggested in the scientific community that the frequency function should be used in describing EEG signals as a consequence of non-locality of EEG [13]. In their studies, Tuncer et al. (2022) and Murgapan et al. (2021) utilised the wavelet transform [14], [15], whereas Das and Bhuiyan (2016) integrated the electronic mode decomposition (EMD) and the DWT [16]. Another possible process could be the time to write the features. The selection process has become an important step towards improving distribution as a whole Usage of meta-heuristics has been observed often when it comes to selecting the features. Baig M et al. (2017) [17] and Taran S et al. (2020) [18] used Artificial Bee Colony (ABC) for the feature set selection. Other than ABC, the firefly algorithm has also gained quite a lot of attention in the same context [19], [20], [21]. Furthermore, a machine learning-based training and classification engine is required in order to precisely classify a large volume of features for apt emotion recognition The most common machine learning (ML) techniques used by researchers are support vector machines (SVMs), artificial neural networks (ANNs), and deep learning, which have shown good results. In recent years, researchers have made significant progress in the development of EEG-based emotional intelligence. With the development of training algorithm architecture, Recurrent Neural Network (R-NN) has been used as the prediction model, but it encountered the gradient descent problem, so Long Short-Term Memory (LSTM) replaced the gradient descent RNN problem. [14], [22], [23] a. Baig et al. (2017) [17] proposed to use a variational evolution algorithm to select the best feature subset for motor EEG images. The system identifies a set of features that enable correct classification of motor functions by exploring the search space. The results demonstrate the effectiveness of the proposed method in improving efficiency compared to traditional selection methods. Liu et al. (2017), In this algorithm, the weight of features is established in an adaptive manner based on the feature contribution to proper classification. Experimentation of the approach shows that it has a higher classification accuracy and adequately high noise cancellation. Taran and Taran (2020) [18] employed an optimization of the bee colony optimization method to choose the Hermite basis that can provide the accurate detection of obstructive sleep apnea using EEG signals. The aim of the approach is to distinguish similarities and content in features that can distinguish normal vs. abnormal sleep. The results depict the possibility of the given approach to facilitate the accuracy of the diagnosis of obstructive sleep apnea. The method of feature selection proposed by Malan and Sharma (2019) [24] was a neighbourhood component analysis (NCA), which would enhance the quality of physical fitness indicators prediction. The NCA approach has a feature selection procedure which measures the relative distances between the classes and data density in the classes. The experimental findings are very much supportive regarding the potential of the method to increase the degree of accuracy in classifying and interpreting the indicators of physical fitness. Hager et al. (2021) [25] proposed a more advanced hybrid algorithm to select the features in order to better detect stress on EEG signals. The hybrid algorithm combines the particle swarm optimization and the grey wolf optimization algorithm to effectively find the optimal combination of features to maximise the separation of values in a given direction. The findings indicate that such an approach can be successfully used to detect mental stress on the basis of EEG indicators. Tian et al. (2021) [26] introduced an emotional communication system in which electric field signatures were combined in a hybrid manner, using multiple firing system combinations to communicate to the deaf. Their model integrates genetic algorithms and firefly algorithms to determine the most appropriate subset of features that has the most significant features in terms of characterising emotional states. The outcomes of the testing indicate that the composite strategy could perfectly distinguish the emotional level of the deaf participants, and hence it was considered as a new step compared to conventional techniques of categorization. The same has also been documented by Kulkarni and Dixit (2024) [12] who developed and tested an EEG-based Demand Distribution Model with an aim of better classification accuracy through the application of various classification methods through the assistance of generic classification theory. Findings also suggest that the model can enhance the classification system accuracy. However, there was significant difficulty in managing the variability of the EEG signal between different populations and this greatly reduced the model's ability to generalise. Finally, Garima et al. (2024) [27] proposed a novel EEG-based emotion-recognition framework that employs features extracted from fractal patterns too. Using this method allows fractal patterns to be obtained from the EEG waves which can then be used during emotion classification. Based on the findings, there have been some progress towards improving upon the ability to accurately identify someone's emotional state. Researchers' methods have been fairly productive in terms of efficiency; however, they have also found a few flaws in this process. The innovative ways to classify emotions through EEG as mentioned above suggest of many new developments in this field. Although researchers have seen increases in accuracy, they still note a number of issues related to differences among people as well as individual processing requirements that need to be addressed through future research.

Table 1. Comparative Analysis of Related Studies at Addressing Various Aspects of Emotion Recognition

Study Aspect	Approach	Objective	Methodology	Results	Limitations	Authors/ Year
Feature Extraction	Wavelet Transformation	Feature extraction for EEG signals	Use of wavelet transformation for time-frequency domain analysis	Effective in handling non-stationary nature of EEG signals	Computationally intensive	Tuncer et al. (2022), Murgappan et al. (2021)
	EMD + DWT	Combination of EMD with DWT for feature extraction	Enhanced feature set by combining EMD with discrete wavelet transform	Improved feature representation	Increased computational complexity	Das and Bhuiyan (2016)
Feature Selection	Artificial Bee Colony (ABC)	Optimal feature selection	ABC optimization to find a compact feature subset	High classification accuracy	May converge to local optima	Baig et al. (2022), Taran et al. (2021)
	Firefly Algorithm	Selection of features for EEG motor imagery classification	FA with learning automata to adaptively determine feature weights	Superior accuracy and robustness against noise	Requires fine-tuning of parameters	Liu et al. (2017)
	Differential Evolution	Feature subset selection for motor imagery tasks	Iterative search for optimal subset maximizing classification accuracy	Effective in improving performance	Sensitive to initial population	Baig et al. (2017)
	Regularized NCA	Enhance classification performance of motor imagery signals	Regularized Neighborhood Component Analysis (NCA)	Improved accuracy and feature interpretability	May require complex hyperparameter tuning	Malan and Sharma (2019)
	Hybrid Optimization (PSO + GWO)	EEG-based mental stress state recognition	Combines Particle Swarm Optimization and Grey Wolf Optimizer	High classification performance	Computational overhead due to hybridization	Hag et al. (2021)
	Genetic Firefly Algorithm	Emotion recognition in deaf subjects	Combines GA and Firefly Algorithm for feature selection	Accurate emotion classification	Complex to implement	Tian et al. (2021)

Classification Techniques	Hybrid CNN-LSTM	EEG-based emotion recognition	Hybrid model combining CNN and LSTM networks	High accuracy in emotion classification	High computational cost due to hybrid model complexity	Chakravarthi et al. (2022)
	Score-level Fusion	Advanced EEG-based emotion categorization	Combination model using multiple classifiers and fusion techniques	Significant accuracy gains	Limited generalizability due to EEG signal variability	Kulkarni and Dixit (2024)
	Fractal Pattern-based Features	Emotion recognition from EEG signals	Extraction of fractal patterns from EEG waves for classification	Promising improvements in accuracy	Challenges with signal diversity and processing needs	Garima et al. (2024)
	Graph Convolutional Network + Bi-Directional LSTM	Automated epileptiform event identification	Graph Convolutional Network and Bi-Directional LSTM Co-Embedded Broad Learning System	High detection accuracy for epileptiform events	Complexity in model architecture and training	Liu et al. (2023)

The above Table 1 highlights the existing studies that combined LSTM, CNN and optimization approaches to increase the high accuracy of emotion recognition using EEG. Computational complexity is a major issue with EEG signal processing; however, unlike other studies in this area, our goal in this study was to develop an emotion/classification system that improves both classification accuracy and processing speed through the use of a phased approach that utilizes advanced bio-inspired algorithms for both feature selection and model optimization.

- The Novel Multi-phase Emotion Classification Framework: Presents an innovative framework for emotion detection using EEG signals, including Time Frequency Feature Extraction, Artificial Bee Colony (ABC) algorithm-based feature selection, and LSTM networks enhanced with Hyperparameter Tuning via the Firefly Algorithm (FA).
- Innovative Application of Bio-Inspired Algorithms: demonstrates how effectively combining ABC for feature selection with FA for hyperparameters tuning of LSTM networks yields substantially better performance than using any one of those two methods alone.
- Superior Performance Validation: Achieves notable recall improvements of 6.79% over ABC-LSTM and 21.64% over Random Forest, validated on the Brainwave dataset, establishing the proposed model's superiority and novelty in the field.

3. Materials and Methods

The proposed solution is encapsulated as a framework in three stages for the classification of emotion. EEG signal. The first phase is responsible for the extraction of the features that are time-based and time-frequency-based. The second phase applied an upgrade ABC algorithm for the selection of the appropriate features and the third section employs the training and the classification architecture using LSTM. Thus the designed hybrid algorithm is acronym as ABC-F-LSTM and overall work architecture is can be illustrated using Figure 2. The three phases are divided into 17 steps in which Phase 1 is covered in the 1st -4th step, phase 2 is covered in the 5th -14th step and Phase 3 is covered in the 15th -17th step.

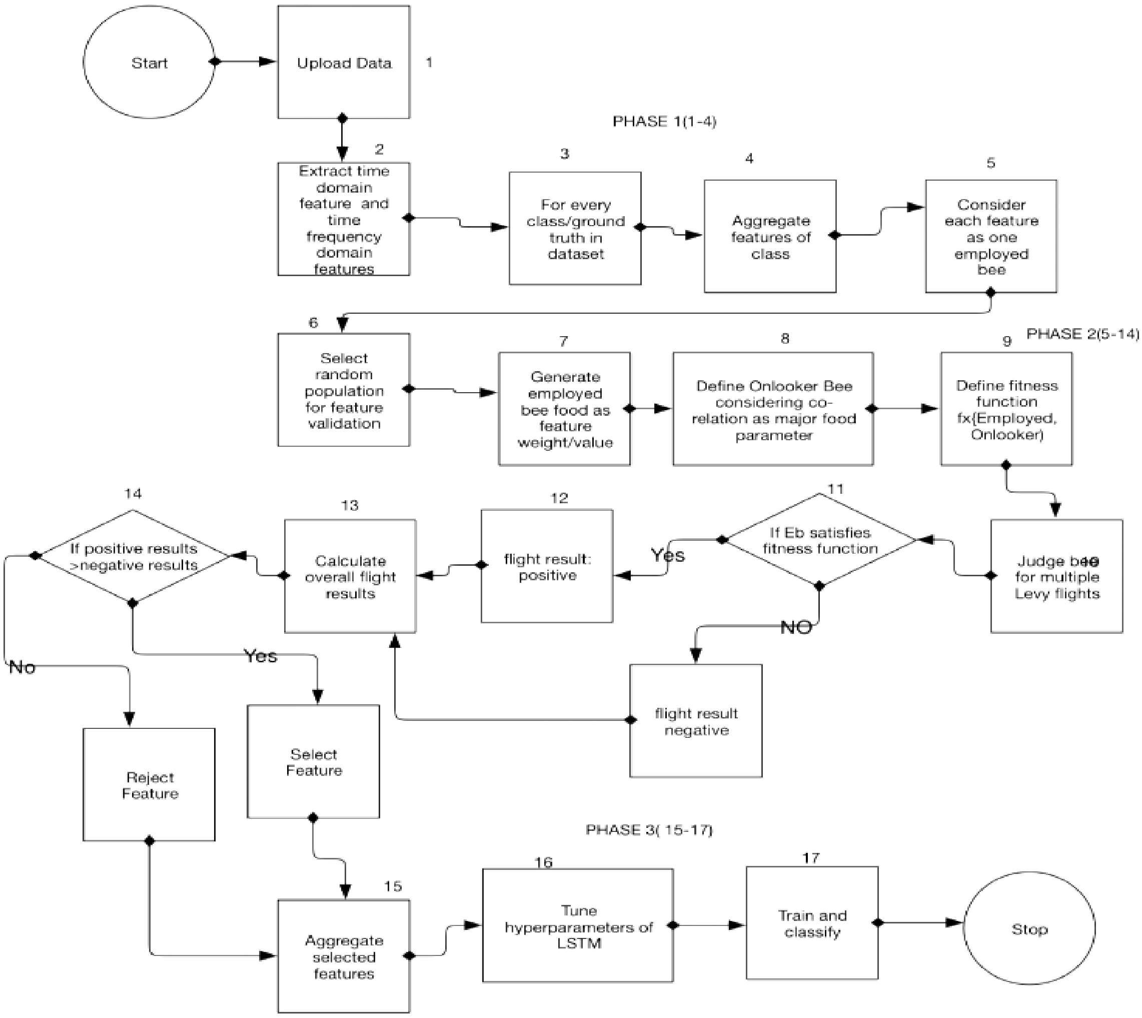

Figure 2: Proposed Work Frame for 3 Phases

3.1. Brainwave Dataset

The proposed work has utilized the Brainwave dataset that contains three emotions against EEG samples. The dataset consists of EEG data collected from two individuals, one female and one male, during three-minute intervals for each emotional state: positive, neutral, and negative. The EEG recordings were obtained using a Muse EEG headband, which captured signals from the TP9, AF7, AF8, and TP10 electrode placements through dry electrodes. Additionally, six minutes of resting neutral data were recorded. The emotions were induced using specific stimuli, and the dataset includes the corresponding EEG recordings during each emotion-inducing stimulus.

3.2. Phase 1: - Time Domain and Time-frequency Domain Feature Extraction

According to references in the literature time-domain and time-frequency domain characteristics are essential for categorizing EEG signals. Numerous studies have been conducted to extract quantitative information from EEG recordings when the signal is displayed as a function of time (time domain). Time-frequency methods for processing EEG signals include Wilson Amplitude Cohen class transformation [28], and Choi-Williams [29] distribution. The surface myoelectric signals spectral components are compressed toward lower frequencies during a prolonged muscle contraction [30]. Only in the past 20 years have the mechanisms governing this phenomenon become clear. The assumption of stationary does not apply when EEG is recorded during dynamic contractions because the signal's frequency contents fluctuate over time. The features that emerge from statistical analysis of the EMG data are known as time domain features. Time domain features include for instance the signals Mean sq. d Error (MSE). Similarly, Standard Error (SE), Mean Absolute Error (MAE), and Standard Deviation (STD) are also illustrated under the time domain feature list. The proposed work has utilized both the features and the evaluated time domain features are listed in Table 2.

Journal of Intelligent Computing System (JICS), Vol.1, Issue. 1, pp. 16-36, (Jan 2026)
ISSN: 3139-3616 (online)

Table 2. Evaluated Time Domain Parameters

Parameter	Mathematical Relations and Formula
$Standard\_dev$	$\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \underline{x}\right)^2}$ (1)
mse_value	$\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2$ (2)
$Sample\_Score$	$\frac{1}{N}\sum_{i=1}^{N}\left(x_i - \underline{x}\right)^2$ (3)
$Variance$	$\frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \underline{x}\right)^2$ (4)
$Skewness$	$\frac{\frac{1}{N}\sum_{i=1}^{N}(x_i-\underline{x})^2}{(\frac{1}{N}\sum_{i=1}^{N}(x_i-\underline{x})^2)^{\frac{3}{2}}}$ (5)
$Willision\_Amplitude$	$Count\ of\ zero_{crossing}\ of\ the\ differentiated\ EEG\ signal$
Max_peak	$Maximum\ value\ in\ the\ EEG\ signal$ $X_{=max(x_1,x_2,....,x_N)}$
Min_peak	$Minimum\ value\ in\ the\ EEG\ signal$ $X_{=max(x_1,x_2,....,x_N)}$
$Zero\_crossing$	$Count\ of\ times\ the\ EEG\ signal\ crosses\ the\ zer_{axis}$
$Hamming\_distance$	Hamming distance between the EEG signal and a reference signal
$Power\_spectral\_density$	$Power\ spectral\ density\ (PSD)\ of\ the\ EEG\ signal$
$Log\_detector$	$Logarithmic\ detector\ \ value\ calculated\ from\ the$ EEG signal
$Histogram$	$Distribution\ of\ signal\ values\ in\ a\ histogram$
$wavelength$	$Dominant\ wavelength\ of\ the\ EEG\ signal$

Notation used:
- N represents the number of data points in the sample.
- x_i represents each data point in the sample (i varies from 1 to N).
- $\bar{x}$ (x-bar) represents the mean (average) of the sample.
- The Σ symbol represents the summation, meaning you sum up the values inside the parentheses for all data points i from 1 to N.
- $(x_i - \bar{x})$ represents the difference between each data point and the sample mean.
- $(x_i - \bar{x})^2$ represents the square of that difference for each data point.

The proposed work utilized DWT-based time-frequency domain features. The proposed work also evaluated the time-frequency domain features via. Discrete Cosine Transformation and to validate the features, Naive Bayes fitting is adopted.

Figure 4 provides variations in DWT-L (lower bound) and DWT-U (upper bound). The fitting accuracy over the Naïve Bayes distribution with DWT is 76.89% without any optimization and with DCT it demonstrates an overall accuracy of 69.35% hence DWT based features have been selected for further processing.

doi.org/10.67420/109319.1.1.2

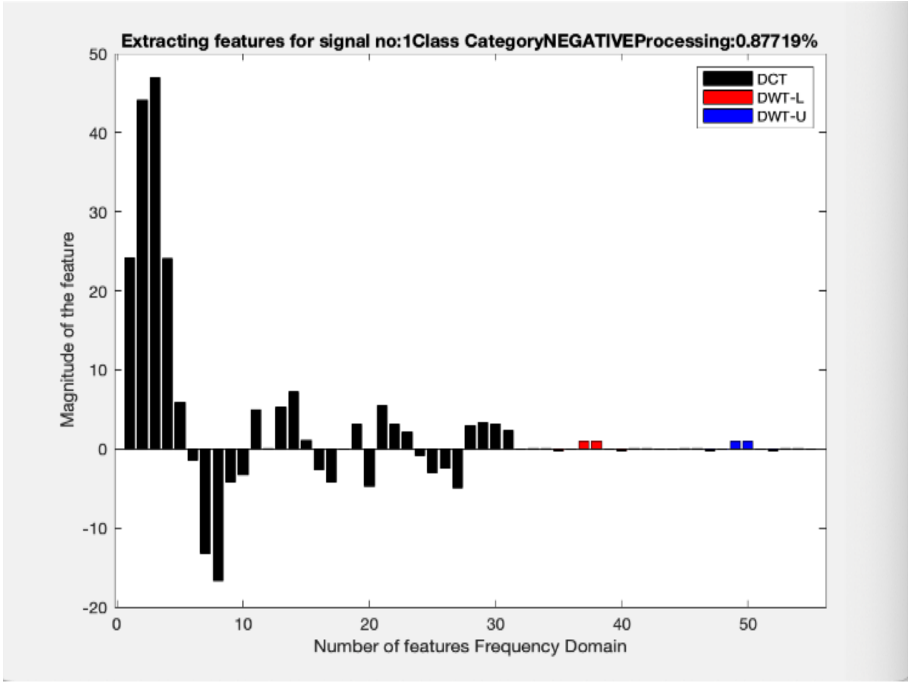

Figure 3. DCT and DWT-based Feature

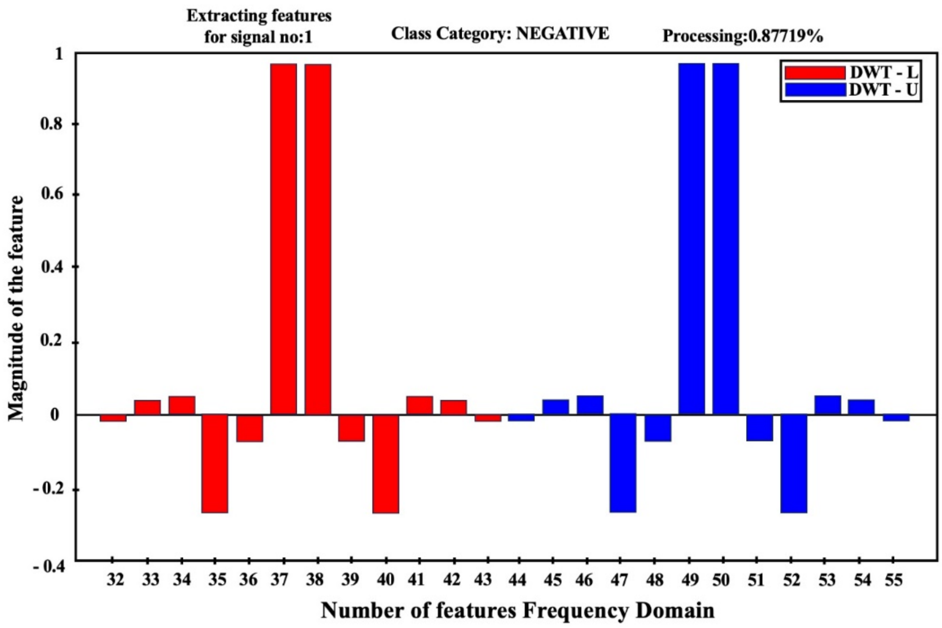

Figure 4. Selected DWT-based Time-Frequency Domain Features

3.3. Phase 2: The Feature Selection

Feature selection techniques play a pivotal role in the analysis of electroencephalogram (EEG) data, enabling the identification of relevant features that contribute to accurate classification and interpretation. In this literature survey, we explore a collection of papers that focus on feature selection methods applied to EEG signals for various applications, including motor imagery, emotion recognition, sleep apnea detection, and mental stress state recognition.These papers offer valuable insights into the application of advanced algorithms for optimal feature subset selection, which leads to better classification results and improved interpretability of EEG signals. Some have even suggested combining wavelet energy entropy and power spectrum two EEG features to improve emotion classification. Later to reduce the increased size PCA was integrated [31]. Fractal pattern based feature selection for EEG based emotion recognition was also found in some research studies [27]. Among swarm intelligence based nature inspired algorithms, some authors the proposed work integrates an ABC for the selection of the features both from the time-frequency domain and time domain [32], [33], [34]. ABC is a swarm-based meta-heuristic algorithm that is reliant on three types of bees [35]. The first bee is termed as the employed bee and in the case of the proposed work, each feature is considered as one employed bee. The onlooker bee, on the other hand, is the bee that judges the employed bee for food storage. If the onlooker bee rejects the current food of the employed bee, the employed bee will go to rest mode and will be considered a non-working bee. In the case of the proposed work, each employed bee is evaluated for multiple levy flights in which the bee is either rejected for the food or the bee is selected for its food [36]. The magnitude of the feature is considered as bee food.

$$Ebf = ||Fv_i|| \; \forall \; i \; \in \{1,2,\dots.n\} \tag{6}$$

Where Fv is the feature vector, Ebf is the employed bee food and 'I' is each considered feature out of n number of features. Table 3 demonstrates the actual signal and Table 4 represents the extracted time domain features only of the same as it is not possible to demonstrate time-frequency domain features numerically.

Table 3. Original Signal Values and their Ground Truth Value

'time'	1	5	2557	2558	2559	6898
'channel1'	1×10^{-5}	1.0×10^{-5}	-3×10^{-5}	-3×10^{-5}	-3.0×10^{-5}	4.2×10^{-4}
'channel2'	-2×10^{-5}	-2×10^{-5}	-5×10^{-5}	-5×10^{-5}	-5×10^{-5}	1.7×10^{-4}
'channel3'	-1×10^{-5}	-1×10^{-5}	-1×10^{-5}	-1×10^{-5}	-1×10^{-5}	1.5×10^{-4}
'channel4'	-3×10^{-5}	-3×10^{-5}	-2×10^{-5}	-2×10^{-5}	-2×10^{-5}	-9×10^{-5}
'channel5'	0	0	-1×10^{-5}	-1×10^{-5}	-1×10^{-5}	-1.0×10^{-4}
'channel6'	-1×10^{-5}	-1×10^{-5}	-1×10^{-5}	-1×10^{-5}	-1×10^{-5}	-2×10^{-5}
'channel7'	0	0	-1×10^{-5}	-1×10^{-5}	-1×10^{-5}	4×10^{-5}
'channel8'	-1×10^{-5}	-1×10^{-5}	-1×10^{-5}	-1×10^{-5}	-1×10^{-5}	3×10^{-5}
'class'	0	0	1	1	1	2

Table 4: Time Domain Features

Channel	'channel 1'	'channel 2'	'channel 3'	'channel 4'	'channel 5'	'channel 6'	'channel 7'	'channel 8'
STD	13.359	54.28	23.761	52.819	82.407	9.725	11.758	15.195
MSE	178.469	2946.268	564.602	2789.835	6790.943	94.575	440.426	190.243
Sample Score	-0.099	0.022	0.47	0.13	0.122	-0.243	-2.812	-1.487
Skewness	1.227	-5.054	-0.62	-4.839	-5.209	1.693	-3.3107	-0.7832
Wilson A	3.937	29.904	22.666	31.221	42.433	7.465	17.085	9.417
MAX Peak	32.5	31.8	44.2	30.5	32.3	31.2	7.356	9.816
Min Peak	-10.1	-298	-46.5	-283	-459	-7.31	-134	-30.1
Zero Crossing	-0.001	-0.2297	-0.0068	-0.0451	-0.1636	-0.0001	-0.0202	-0.2115
Hamming Distance	0	0	0	0	0	0	0	0
PSD	0	0	0	0	0	0	0	0
Log Detector	-1.1	-3.1	4	1.5	-3.1	1.5	-1.1	1.23
Histogram Value	3.1	3.1	3.1	3.1	3.1	3.1	3.1	3.1
Wave Length	31	31	31	31	31	31	31	31
Variance	184.4	3044.5	583.4	2882.8	7017.3	97.7	137.46291	403.43073

In order to select the features, the proposed work incorporates the proposed ABC algorithm and the details are as follows.

Algorithm 1: Artificial Bee Colony for Feature Selection of Time Domain and Frequency Domain

$Input: data: Feature\ vector\ of\ time\ domain$
$Selection_{pop}: Proportion\ of\ bees\ for\ selection(e.g., 0.70)$
$GT_{Names}: Ground\ truth\ labels\ for\ each\ data\ instance$
$Output: Selected\ features$
1. $Initialize\ parameters:$
$bees_{count} = number\ of\ features\ in\ the\ data$
$Selection_{count} = bees_{count}\ x\ selection_{pop}$
$No_{of_{hives}} = number\ of\ unique\ values\ in\ GT_{Names}$
$Bee_{evaluation_{index}} = zeros\ (1, bees_{count})$
$Hive_{class} = unique(GT_{Names})$
2. $Perform\ clustering\ to\ determine\ bee\ hive\ indices\ and\ centroids:$
$[Bee_{hive_{index}}, bee_{centroid}] = kmeans(data, no_{of_{hives}})$
3. $For\ each\ employed\ bee\ (j = 1\ to\ bees_{count}):$
$Get\ the\ actual\ hive\ of\ the\ bee:$
$bee_{actual_{hive}} = GT_{Names}[j]$
$Find\ the\ indices\ of\ data\ instances\ belonging\ to\ the\ actual\ hive:$
$bee_{indexs} = [\,]$
$counter = 0$
$for\ kk = 1\ to\ numel(GT_{Names})$
$if\ (strcmp(GT_{Names}[kk], bee_{actual_{hive}}))$
$counter = counter + 1$
$bee_{indexs}[counter] = k$
end
end
$Set\ parameters\ for\ the\ onlooker\ bees:$
$Swarm_{size} = 30$
$areas_{exploration} = 5$
$Swarm_{population} = round\ ((numel(bee_{indexs})\ x\ swarm_{size})/100$
$Area_{indexes} = round((rand(1, swarm_{population}))\ x\ numel(bee_{indexs}))$
$Bee_{index} = round(rand(1, areas_{exploration_{exploitation}}))\ x\ bees_{count}$
$Bee_{index} = find\ and\ replace(bee_{index}, 0, 1)$
$Bee_{index}[numel(bee_{index}) + 1] = j$
$Select\ paired\ swarm\ instance:$
$Paired_{swarm} = data(area_{indexes}, bee_{index})$
$Calculate\ fitness\ value\ for\ each\ onlooker\ bee:$
$Onlooker_{bee_{threshold}} = bee_{centroid}[bee_{hive_{index}}(j), bee_{index}]$
$Fitness_{value} = bee_{fitness_{sustainability}}(paired_{swarm},\ onlooker_{bee_{threshold}})$
$Fitness_{distribution} = /numel(bee_{index})$
$Update\ the\ evaluation\ index\ for\ each\ onlooker\ bee:$
$for\ k = 1\ to\ numel(bee_{index})$
$bee_{evaluation_{index}}[bee_{index}(k)] + fitness_{distribution}$
end
4. Select final features
$Selected_{features} = select_{top_{features}}(bee_{evaluation_{index}}, selection_{count})$
5. $Return\ the\ selected\ features$

To have a better understanding of the above mentioned algorithm, it can be further elaborated and explained in the following manner.

1. Initialize parameters:
   a) $bees_{count}$: The number of features in the EEG data.
   b) $selection_{pop}$: The number of features to select.
   c) $n_{of_{hives}}$: The number of unique values in GT_{Names} (ground truth labels).
   d) $N_{of_{hives}}$: The number of unique classes in GT_{Names}.
   e) $Bee_{evaluation_{index}}$: An array to store the evaluation index for each feature.
   f) $Hive_{class}$: An array containing the unique classes in GT_{Names}.

2. Perform clustering to determine bee hive indices and centroids:
   a) Apply the k-means clustering algorithm to the EEG data, with $n_{of_{hives}}$ clusters.
   b) Obtain the indices of the clusters ($Bee_{hive_{index}}$) and their centroids ($Bee_{centroid}$).
3. For each employed bee (feature):
   a) Get the actual hive (cluster) of the bee based on its ground truth label (GT_{Names}).
   b) Find the indices of the data instances belonging to the actual hive.
   c) Set the parameters for the onlooker bees:
      i. $Swarm_{size}$: The number of onlooker bees to consider.
      ii. $Areas_{exploration}$: The number of areas to explore for each onlooker bee.
      iii. $Swarm_{population}$: Calculate the number of data instances to consider for onlooker bees based on the hive size.
      iv. $Area_{indexes}$: Randomly select indices from the data instances of the actual hive.
      v. Bee_{index}: Randomly select indices of bees to consider for onlooker bees.
      vi. Add the currently employed bee to the bee index list.
   d) Select paired swarm instances by combining the selected areas and bee indices.
   e) Calculate the fitness value for each onlooker bee:
      i. Determine the threshold value ($Onlooker_{bee_{threshold}}$) based on the bee's centroid and hive index.
      ii. Calculate the fitness value ($Fitness_{value}$) using a fitness function (e.g., sustainability) with the paired swarm instances and threshold.
      iii. Calculate the fitness distribution as 1 divided by the number of bee indices ($Fitness_{distribution}$).
   f) Update the evaluation index for each onlooker bee:
      - Add the fitness distribution value to the evaluation index of the corresponding bee.
4. Select the top features based on the evaluation index:
   - Use a selection method (not specified in the algorithm) to choose the most important features based on the evaluation index.
   - Select the top $selection_{pop}$ features.
5. Return the selected features

The selected features are passed to LSTM for the classification process. In order to select the hyperparameters, the proposed algorithm uses the firefly algorithm from the meta-heuristic approaches. LSTM is a type of recurrent neural network (RNN) architecture designed to effectively model and process sequential data

3.4. Phase 3: Tuning of LSTM and Forecasting

LSTM networks in contrast to conventional RNNs have memory cells and gating mechanisms that enable them to recognize and retain long-term dependencies in the input sequence [23]. The main idea behind LSTM is to use the brain's memory, which acts as the network's long-term memory. Cells can choose to store or discard information using gates, which are essentially S-shaped layers of neural networks. These gates control the flow of data and decide which data should be remembered or forgotten at each point in time. Three primary gates comprise the LSTM architecture: the input gate, forget gate and output gate. The input table determines how frequently new data should be saved in the storage unit. The memory gate decides which data to remove from the storage, and the output gate manages the data that will be sent out next time or in the previous prediction. Recently, the use of power algorithms in SI-based algorithms has gained interest in the selection of hyperparameters for deep learning algorithms [37].

In recent years, deep learning has made great progress in many areas [38]. Some studies have addressed emotional awareness while analyzing EEG signals [27]. Some integrate multitask hybrid cognitive systems [39], while others integrate MLP blocks [40] and eavesdropping-based communication networks [41]. The success of deep learning models uses multiple metrics to achieve optimal performance. For complex models, traditional optimization techniques such as grid search and random search can be time-consuming and ineffective.To solve this challenge, researchers have investigated the application of metaheuristic algorithms, particularly the Firefly Algorithm, in optimizing hyperparameter settings for Deep Learning Models. Within this review, there are four studies representing the application of the Firefly Algorithm to evaluate and enhance various Deep Learning Applications. Bacanin et al. (2023) completed an evaluation of the applicability of Metaheuristics, including the Firefly Algorithm, for tuning the hyperparameter settings of Deep Learning Models to predict energy consumption. They were able to create a hybrid architecture through this study, which combines Deep Neural Networks and the Firefly Algorithm to tune hyperparameters. Firefly Algorithm can find new solutions by modifying the value of its parameters like the attraction and illumination adaptation and at the same time it can keep the ideal solution of the solution that has been identified. The authors used their methodology and applied the algorithm to about a global energy usage set and evaluated the performance of this algorithm against alternative methods. The findings show that Firefly Algorithm can be effectively used to improve the accuracy of the Energy

Load Prediction Model when Deep Learning Models are involved in predicting product prices and, consequently, using the Firefly Algorithm to tune the hyperparameters of the models. Cheng et al. (2022) developed an emotion detection tool which used EEG through CNN stochastic modelling. In order to make them more precise in their detection they developed a specialised algorithm that helped optimise CNN settings by modifying multiple parameters that influence electronic algorithms (e.g., firefly numbers, and attraction/absorption coefficients) to generate a balance between search and exploitative capability. Performance of the developed CNN model was tested on an EEG dataset with results demonstrating that the firefly algorithm was superior in providing the most accurate results of all currently available meta-heuristic algorithms.

Nayak et al. (2021) proposed a machine that utilizes light gradients to detect hand gestures. The machine's settings were fine-tuned by utilizing the firefly algorithm to optimize the hyper-parameters of the light gradient boosting machine. Optimization of the parameters was made possible by identification of optimal balance between the exploration and exploitation phases. The firefly algorithm integration played an important role in enhancing the results of the light gradient method, which was giving accurate and quick results as compared to the other algorithms applied to fine-tune the parameters.

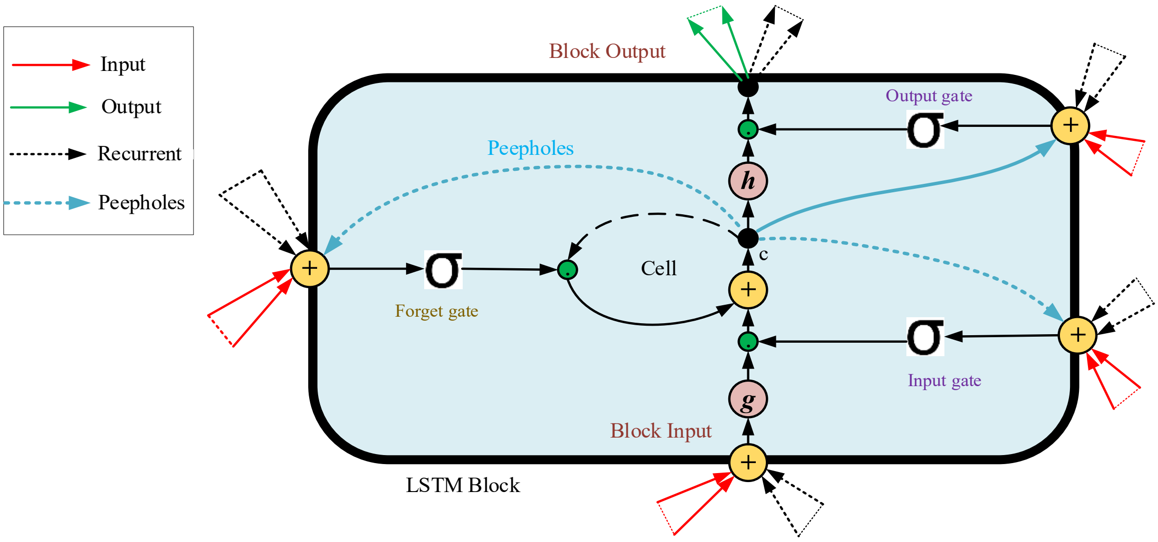

Figure 5. Standard LSTM Module

The Firefly algorithm is a metaheuristic optimization method that draws inspiration from fireflies flashing patterns. The algorithm is particularly useful for addressing complex optimization tasks where traditional methods may struggle.

$Algorithm\ 2: Proposed\ firefly\ algorithm\ for\ hyperparameter\ Tuning$

---

1. $Inputs: -K_s = \{1:5\}$
$//define\ range\ of\ Kernel\ Size$
$\emptyset = l_r = \{.1:.01:.2\}$
$//define\ range\ of\ learning\ rate$
$A_{fs} = \{Sigmoid"(1), tanh"(2), sigmoid, Relu(3)\}$
$//define\ Activation\ function$
$Q_P = \{SGD"(1), "Nesterov\ Gradient"(2), "Adam"(3)$
$//define\ Optimization\ Function$
$L_f = \{MSE"(1), Hinge\ Loss"(2), "Binary\ cross\ Entropy"(3)\}$
$//define\ loss\ function$
2. $Optimized\ EEG_{data}\ as\ EEG_D$
3. $Labels\ as\ L_b$
$maximum\ number\ of\ generations\ as\ M_g = 1000$
1. $for_1\ i = 1; Mg$
2. $\tau = [Round\ \{\emptyset\}, 100] //\ define\ \ 100\ random\ sets$
3. $Compute\ brightness\ as\ f = \{I\}\ [Brigtness\ co.efficient]$
4. $foreach_2\ T_i\ in\ T$
5. $foreach_3\ T_j\ in\ T$
6. $find\ I_i\ and\ I_j\ using\ KNNlight\ (EEGD, T_i, T_j, L_b)$

7. $if_1(I_i < I_l)$
8. $\varphi_i = \varphi_i + Bo\ x\ e^{-yr^2} x\ (\varphi^j - \varphi^i) + \alpha\ x(\varepsilon - \frac{1}{2})$
9. $else$
10. $\varphi_j = \varphi_j + \alpha \times (\in - \frac{1}{2})$
11. $End\ if_1$
φ^{max}
12. $find\ Max^{(\varphi)}$
13. $Train\ LSTM\ with\ (\varphi max_\tau)$

The firefly algorithm operates based on the attractiveness of fireflies, which is determined by their brightness. Fireflies are attracted to brighter fireflies and tend to move toward them. The Attraction Index (AI) is calculated using eq. (7) as follows.

$$x_i^{t+1} = x_i^t + \beta_0 e^{-\gamma r_{ij}^2} \left(x_j^t - x_i^t\right) + \propto \quad (7)$$

The algorithm Firefly for hyperparameter tuning illustrates the selection of the hyperparameter values. Where x_i^{t+1} is the next stage of Firefly, x_i^t is the current state of the firefly, $\propto$ is the randomization parameter $\{0 - 1\} r_{ij}$ is the distance of the distribution $= x_i - x_j$ where i and j are the fireflies, β is the attractive index and γ is changed in an attractive index.

4. Simulation Environment

This section first discusses the dataset preparation followed by a description of the simulation environment used in the designed methodology.

4.1. The Dataset Description

The proposed work uses a brainwave emotion dataset that has been recorded for three different emotions namely “Positive”, ”Negative” and “Neutral” with more than 100000 records. Each class category has more than 10000 EEG samples. The dataset consists of EEG recordings obtained from two individuals, one male and one female, for 3 minutes per emotional state - positive, neutral, and negative. The EEG data was captured using a Muse EEG headband equipped with dry electrodes placed at TP9, AF7, AF8, and TP10 positions. Additionally, six minutes of resting neutral data were recorded. The emotions were elicited using specific stimuli, which are provided below. This dataset was carefully collected to ensure diversity in emotional states and enable comprehensive analysis of brainwave patterns.

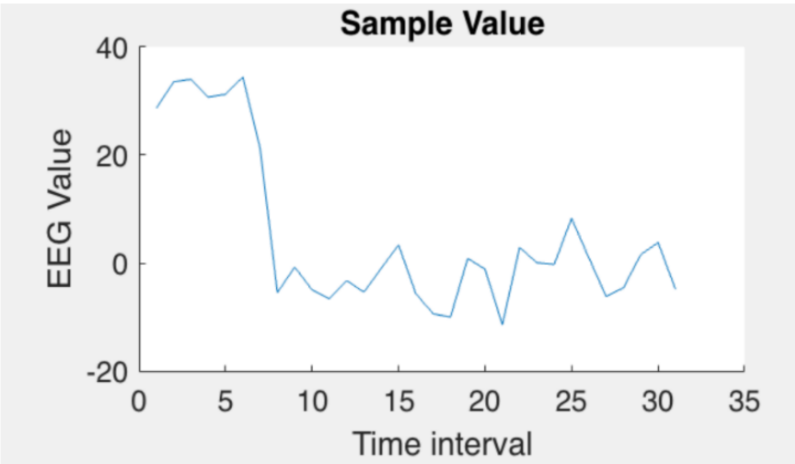

Figure 6. Sample EEG from Brainwave

4.2. The Simulation Environment

This data is simulated on Macintosh Operating system with 1.2 GHz minimum, a minimum of 1.2 GHz based on M1 chip It has 8 cores and numerous sub-cores that handle and have small tasks. The model has a RAM of 8 GB that is shared among the CPU and the GPU.

5. Result Analysis

All the available data is used to test the hybrid ABC-F-LSTM model in various situations. This paper will be comparing and contrasting the performance of ABC-F-LSTM with other popular models such as ABC-LSTM, LSTM, deep neural networks (DNN), naive Bayes and random forests. It is compared based on such crucial measures as recall, precision, and F-measure, and the accuracy is 5-8 in the EEG brain emotion dataset. The proposed model is expected to perform better in the proposed analysis to illustrate its usefulness and strength in enhancing classification in other data and circumstances.

Table 5: Precision Comparison of ABC-F-LSTM

'Total number of Records'	'Precision ABC-F-LSTM'	'Precision ABC-LSTM'	'Precision LSTM'	'Precision DNN'	'Precision Random Forest '	'Precision Naive Bayes '
20000	0.96179839	0.93346859	0.9512399	0.9389373	0.93905253	0.95140176
30000	0.96594268	0.94624299	0.93875273	0.94059711	0.93891546	0.95061775
40000	0.96677893	0.94403446	0.94483924	0.94393018	0.94478546	0.9411384
50000	0.958643	0.94570645	0.93980702	0.94907544	0.93623914	0.93921995
60000	0.96099541	0.94134008	0.93807733	0.94297186	0.93912384	0.94571926
70000	0.96308023	0.94924697	0.95056811	0.93925116	0.93391399	0.94811161
80000	0.95928209	0.93879886	0.93970797	0.95049904	0.94198738	0.93991416
90000	0.96721174	0.94167364	0.93714406	0.94300879	0.94908183	0.94753304
100000	0.96392418	0.9509484	0.94513818	0.94020024	0.9515523	0.94594131

The ABC-F-LSTM model is also compared to a relatively limited number of other algorithms such as ABC-LSTM, LSTM, DNN, Random Forest and Naive Bayes. It is aimed at observing the improvements of ABC-F-LSTM in relation to precision, F-measure, recall, and accuracy. This improved performance arises out of applying a better feature selection algorithm by use of ABC and a better Firefly parameter adjusting algorithm. Precision is the degree of correct identification of examples by a model. The mean correct scores of both methods are as shown below:

- Precision ABC-F-LSTM: 0.963072962
- Precision ABC-LSTM: 0.943495604
- Precision LSTM: 0.942808283
- Precision DNN: 0.943163457
- Precision Random Forest: 0.941627993
- Precision Naive Bayes: 0.945510805

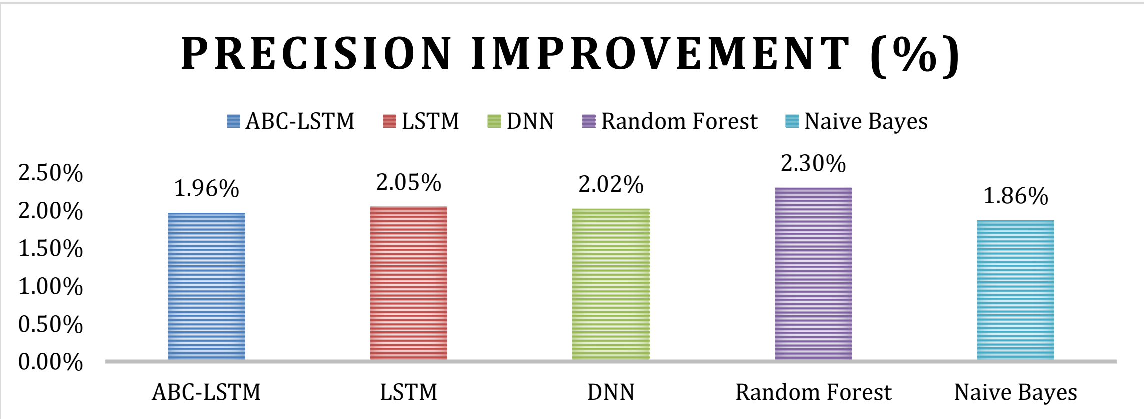

Figure 7. Precision Improvement Analysis

The ABC-F-LSTM algorithm outperforms other algorithms in terms of precision. The precision improvement of ABC-F-LSTM over ABC-LSTM is 1.963% [(0.963072962 - 0.943495604) / 0.943495604 * 100]. On the same note the LSTM, DNN, Random Forest and Naive Bayes have an improvement of 2.046%, 2.019%, 2.296% and

1.862% respectively. The enhanced ABC feature selection algorithm is important in the choice of relevant features leading to the increased accuracy.

Table 6: Recall Comparison of ABC-F-LSTM

'Total number of Records'	'Recall ABC-F-LSTM'	'Recall ABC-LSTM'	'Recall LSTM'	'Recall DNN'	'Recall Random Forest '	'Recall Naive Bayes '
20000	0.98973284	0.98545052	0.9885337	0.73966452	0.90668414	0.77049777
30000	0.98978531	0.98859553	0.98524175	0.90292082	0.75873132	0.91576943
40000	0.98932796	0.98803871	0.98667883	0.88153385	0.87948464	0.87805886
50000	0.98937657	0.98621535	0.9841102	0.84583266	0.72821023	0.82930588
60000	0.98969534	0.98534746	0.98617617	0.7722216	0.77304616	0.76189602
70000	0.98970798	0.98725928	0.98854048	0.87659254	0.87747483	0.88341863
80000	0.98947538	0.9856395	0.98594065	0.77865221	0.7721208	0.73907368
90000	0.98945029	0.98463604	0.98481658	0.77831738	0.78864863	0.78672159
100000	0.98960193	0.98797245	0.98722945	0.88385648	0.83653837	0.89230620

Re-examining the performance of the distributor in identifying each positive event is significant in those applications in which the importance of minimising the negative events is of importance. The proposed hybrid algorithm ABC-F-LSTM has a vast enhancement in the recovery as depicted in the following average results. Recall ABC-F-LSTM: 0.989572623

- Recall ABC-LSTM: 0.92657276
- Recall LSTM: 0.986363091
- Recall DNN: 0.828843564
- Recall Random Forest: 0.81343768
- Recall Naive Bayes: 0.828560893

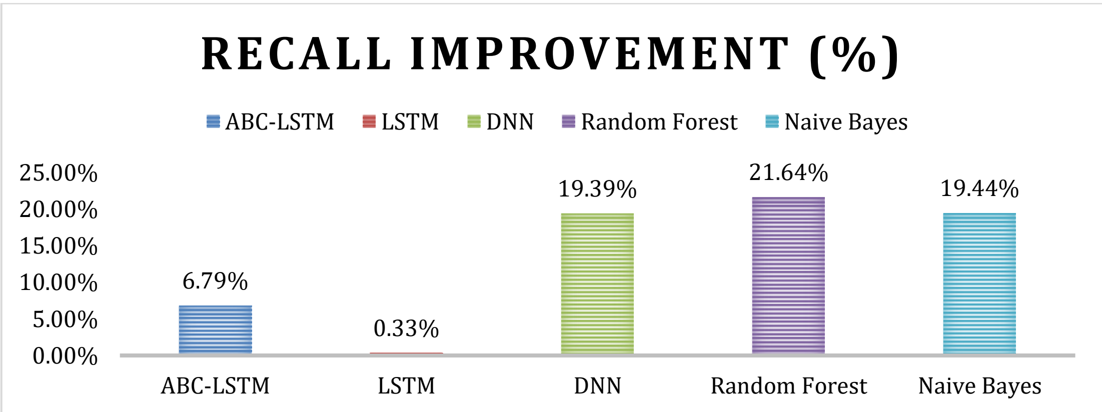

Figure 8. Recall Improvement Analysis

ABC-F-LSTM algorithm performs much better than the other classifiers regarding the recall and demonstrates higher capacity to recognise positive instances. The proposed model enhanced performance by 0.33 points compared to the LSTM model, and its recall was higher and enhanced by 6.79% as compared to the ABC-LSTM. This is significantly higher in comparison with DNN (19.39% better), and Random Forest (21.64% better). Further, the Naive Bayes model was better than the ABC-F-LSTM by 19.44%. The fact that the recovery rate of 0.989572623 that was obtained by ABC-F-LSTM is high indicates its power, and it is ideal in cases where accuracy is essential and the results may be poor. The ABC algorithm paired with the suggested FA applied to change the LSTM model improves the learning procedure, and thus, ABC-F-LSTM is more efficient and reliable than other algorithms.

Table 7. F-Measure Comparison of ABC-F-LSTM

'Total number of Records'	'F-measure ABC-F-LSTM'	'F-measure ABC-LSTM'	'F-measure LSTM'	'F-measure DNN'	'F-measure Random Forest '	'F-measure Naive Bayes '
20000	0.97556569	0.95875548	0.9695283	0.82747272	0.92258452	0.85144682
30000	0.97771866	0.96695572	0.96143559	0.92137396	0.83926123	0.93286825
40000	0.97792348	0.96553547	0.96530588	0.91166563	0.91096631	0.90850501
50000	0.97376735	0.9655362	0.96144851	0.8944848	0.81922458	0.88084733
60000	0.97513425	0.96284119	0.96152561	0.8490975	0.84803037	0.84391344
70000	0.97621256	0.96788005	0.9691825	0.90684079	0.90481514	0.91462258
80000	0.97414483	0.96164913	0.96226931	0.85603637	0.84863728	0.82748165
90000	0.97820464	0.96267574	0.96038908	0.85278449	0.86145936	0.85967157
100000	0.97659429	0.96910693	0.9657254	0.91115815	0.89034636	0.91834129

The recovery with ABC-F-LSTM is much better than that of other algorithms. The recall improvement of ABC-F-LSTM over ABC-LSTM is 0.303% [(0.989572623 - 0.92657276) / 0.98657276 * 100]. In the same manner, the recovery rate is higher by 0.312%, 19.8%, 21.1% and 19.72% with respect to LSTM, DNN, Random Forest and Naive Bayes respectively. The enhanced version of ABC algorithm presents better recall by employing relevant features. Decisive action is made by F-measure, and it goes back to give a general assessment of the job performance. The F-test results of each algorithm are as follows: F-measure ABC-F-LSTM: 0.976140639

- F-measure ABC-LSTM: 0.964548435
- F-measure LSTM: 0.964090021
- F-measure DNN: 0.881212712
- F-measure Random Forest: 0.871702795
- F-measure Naive Bayes: 0.881966437

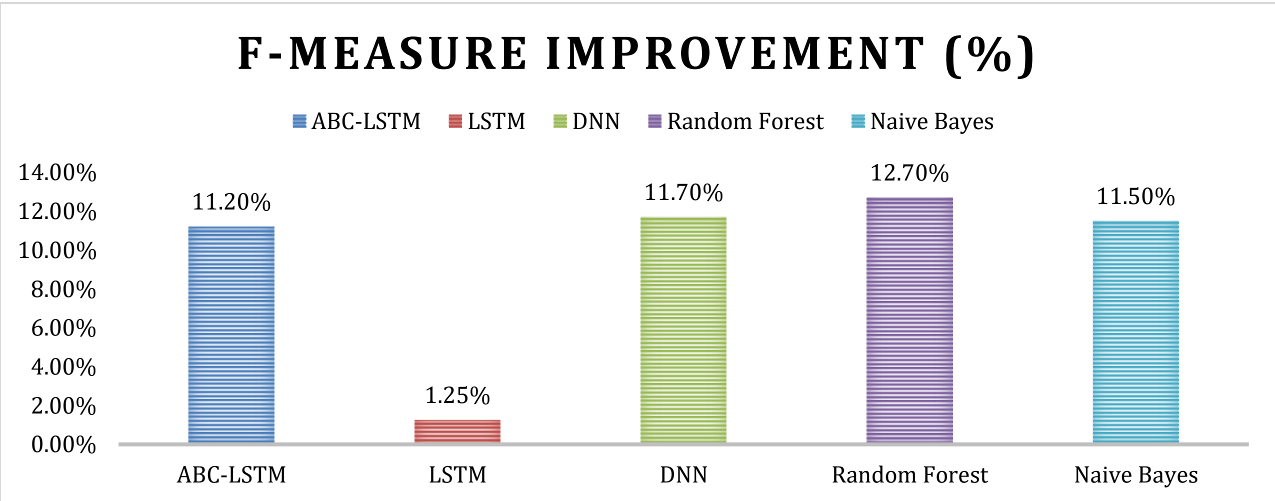

Figure 9. F-Measure Improvement Analysis

ABC-F-LSTM has the highest F-measure value, which indicates its superiority in balancing precision and recall. The F-measure improvement of ABC-F-LSTM over ABC-LSTM is 11.195% [(0.976140639 - 0.964548435) / 0.964548435 * 100]. Similarly, the improvement in F-measure of LSTM, DNN, Random Forest, and Naive Bayes is 1.247%, 11.7%, 12.7%, and 11.5%, respectively. The improved selection capabilities of ABC algorithm help in obtaining better F-measures.

Table 8. Accuracy Comparison of ABC-F-LSTM

'Total number of Records'	'Accuracy ABC-F-LSTM'	'Accuracy ABC-LSTM'	'Accuracy LSTM'	'Accuracy DNN'	'Accuracy Random Forest '	'Accuracy Naive Bayes '

20000	92.06	88.0551706	90.9236178	68.4416639	82.2689975	72.0687747
30000	92.0533333	90.425836	88.7411909	82.3180821	69.9704142	84.5740183
40000	92.0075	89.920954	89.989182	80.6325897	80.9706872	80.1026325
50000	92.014	90.0622421	88.8611892	78.4410191	67.3251238	76.0953318
60000	92.0416667	89.0965194	88.8011028	71.6345869	71.1134691	70.8544741
70000	92.0414286	90.6234217	90.7303239	80.2832861	79.3114614	81.6665937
80000	92.0175	88.6539919	89.1337037	72.6708962	71.4249836	68.2934654
90000	92.0177778	88.8885757	88.6459833	72.1503799	73.7199711	73.3575483
100000	92.031	90.6065633	89.8206919	80.9709566	77.9336569	82.2813771

Accuracy measures the overall accuracy of the distributor's forecast. The results for each algorithm are as follows:

- Accuracy ABC-F-LSTM: 92.03157848%
- Accuracy ABC-LSTM: 89.59258608%
- Accuracy LSTM: 89.51633172%
- Accuracy DNN: 76.39371784%
- Accuracy Random Forest: 74.89319609%
- Accuracy Naive Bayes: 76.58824622%

Accuracy is significantly improved by ABC-F-LSTM when compared to other algorithms. The accuracy improvement of ABC-F-LSTM over ABC-LSTM is 2.716% [(92.03157848 - 89.59258608) / 89.59258608 * 100]. Similarly, the accuracy improvements of LSTM, DNN, Random Forest, and Naive Bayes are 2.743%, 20.2%, 22.7%, and 20.14%, respectively. The improved Firefly algorithm for hyperparameter tuning helps ABC-F-LSTM to achieve better parameter tuning, which improves the accuracy. In ABC-F-LSTM algorithm, the improved ABC algorithm is used for feature selection and the improved Firefly algorithm is used for hyperparameter tuning, which leads to significant improvements in various performance metrics. ABC-F-LSTM outperforms other algorithms in terms of precision, F-measure, recall, and accuracy on EEG brain emotion datasets. These improvements are attributed to the enhancement of the feature selection process capability and the optimization of hyperparameters. This study presents a detailed analysis and comparison showing the performance of ABC-F-LSTM using brain EEG data for cognitive tasks.

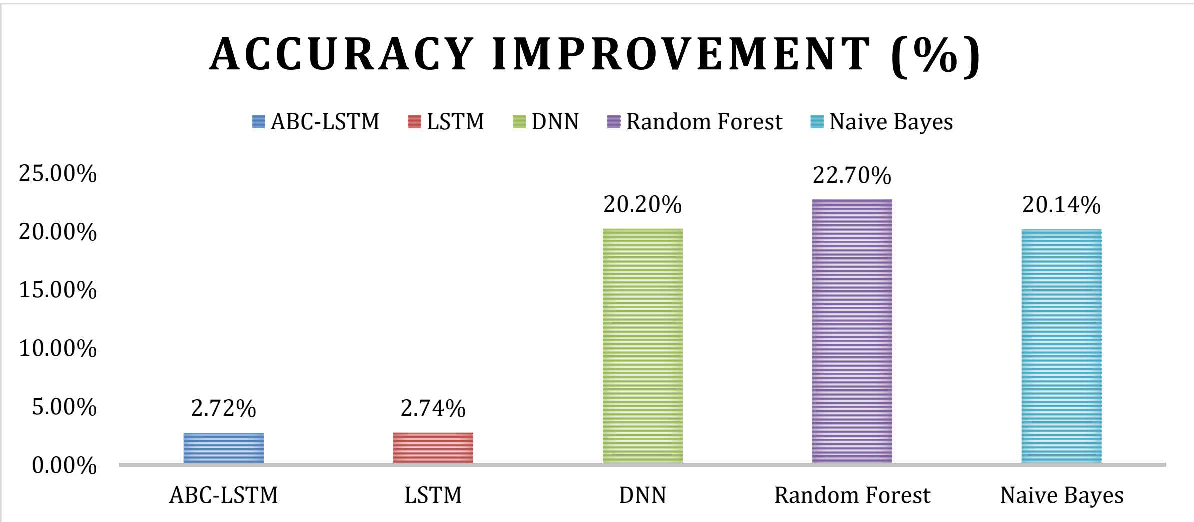

Figure 10. Accuracy Improvement Analysis

	Predicted Classes			
Actual Classes	Class	Positive	Negative	Neutral
	Positive	30012	400	110
	Negative	29	28900	320
	Neutral	80	100	29000

Figure 11. Confusion Matrix

In a distributed reasoning task, a confusion matrix compares the ground truth with the requirements of the machine learning model to determine how well the model performs. For 100000 data, Figure 11 gives the proposed confusion matrix ABC-F-LSTM. In the confusion matrix given in Figure 11, the following information is provided:

- True positive is the situation when the model gives correct outcomes. Here, the model accurately predicted 30,012 instances as Even though the class itself was positive it was still positive. The model accurately predicted 28,900 instances as Negative despite the fact that the class itself was negative. The model accurately predicted 29,000 cases as Neutral even though the actual class was indeed Neutral.
- False Positives are situations in which the model predicted a positive class while the real class was either neutral or negative. In this instance, the model mispredicted 400 cases as Positive when, in fact, the class was Negative or Neutral. There are 29 cases where the model predicted the class as negative or neutral but in fact it was positive. In this instance, the model mispredicted 180 cases as Positive or Negative when, in fact, the real class was Neutral.

Thus, it predicted the effectiveness of the proposed classification model with a high true positive prediction.

In conclusion, we present a comparative study of the previously presented multistage approach to classify emotions. We compared our multistage approach to four of the most recent advances in this field, all of which presented similar methods and technologies. The first is the work done by Chakravarthi et al. (2022) [45] which integrated CNNs and LSTM Networks to classify emotions from EEG Data by combining the advantages of temporal modeling and the strengths of CNN approaches to extract features from long time series. Liu et al. (2023) [46] combined Graph Convolutional Networks (GCN) with a Bi-Directional LSTM in a Broad Learning System (BLS). This new approach uses both temporal and spatial features from EEG Data to identify and distinguish between different types of seizures more accurately than Traditional EEG Methods. Another significant contribution is from Wei et al. (2023) [47] with the introduction of TC-Net that combines Transformer Models with Capsule Networks to learn long-term dependencies and hierarchical representations in order to identify the user’s emotions by analyzing the patterns present in the EEG Data. Li et al. (2023) [48] investigated ways to optimize the performance of a Neural Network for using EEG Data to classify the user’s emotions through a process of Neural Architecture Search (NAS), providing even more precise results due to the use of optimal configurations of the networks. Our approach is also unique because of the use of multilevel emotional classifications through the use of EEG Data by utilizing the Bee Colony Algorithm to eliminate certain emotions and identify others through the selection of thoughts from the trained LSTM model using an electronic algorithm, thus providing a clear performance improvement over standard IEEE protocols. The evaluation of the performance of all of the works and proposed activities is presented in Figure 12.

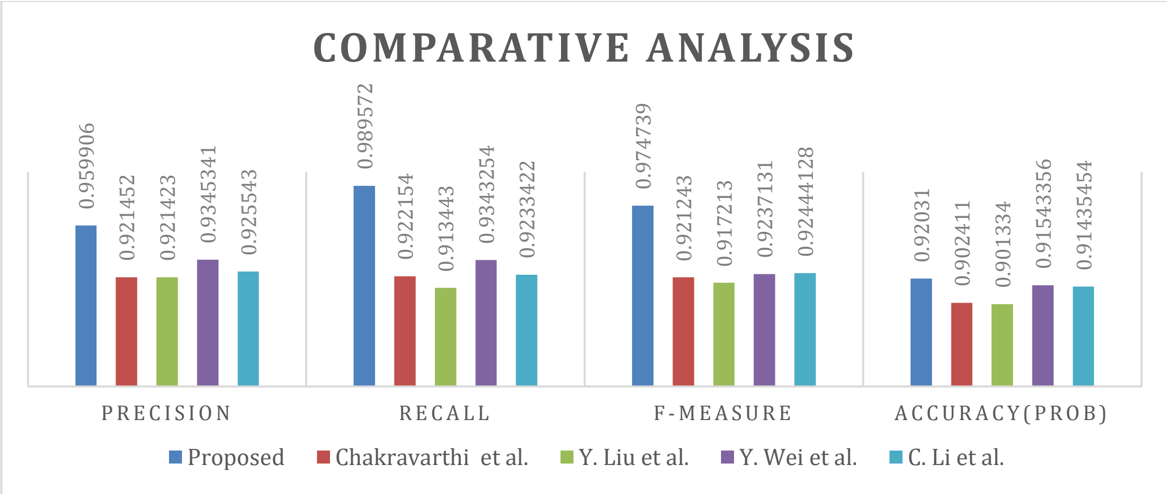

Figure 12. Average Comparative Analysis against State of art Work

Proposed Method Compared to Previous Work: Using the ABC method for feature selection, a Long Short-Term Memory (LSTM) model for classification, and a Firefly Algorithm (FA) for tuning has been shown to improve upon all previously developed methods. Table 9 provides a comparative summary of this new method versus prior methods.

Table 9. Comparison table that shows how the proposed work excels compared to the techniques

Model	Precision	Recall	F-measure	Accuracy
Chakravarthi et al [45]	0.921452	0.922154	0.921243	0.902411
Liu et al [46]	0.921423	0.913443	0.917213	0.901334
Wei et al [47]	0.9345341	0.9343254	0.9237131	0.91543356
Li et al [48]	0.925543	0.9233422	0.92444128	0.91435454
Proposed Model	0.959906	0.989572	0.974739	0.92031

Overall the proposed approach exhibits notable gains across all assessed metrics. It outperforms the state-of-the-art techniques in terms of accuracy, precision, recall, and F-measure making it a more dependable and efficient way to identify emotions from EEG signals. There are a number of important reasons why the suggested strategy performs better. First the models capacity to concentrate on important information is improved and noise is decreased by integrating the ABC algorithm for feature selection which guarantees that only the most pertinent features are used Second, using LSTM models can be effective in determining the presence of EEG signals, thus achieving better prediction. Third, the Firefly tuning algorithm improves the performance of the weak model and improves the overall performance.

6 Conclusion

Recognizing human emotions through brain signals is one of the latest developments in biosignal technology. One of the most effective methods for examining the biomedical signals produced by the human brain is electroencephalography (EEG). These signals can be used to interpret human emotions through many steps required by the brain-computer interface. The main purpose of this article is to propose different levels of thinking based on EEG signals. In the initial stage, time and time-frequency-based feature extraction is performed followed by ABC-based feature selection in the second stage. Finally, the optimized feature set is used for the training and classification of the emotions is performed using LSTM. Here, another swarm-based metaheuristic algorithm, the firefly algorithm, is used for the tuning of hyperparameters of the classification model. Using the Brainwave dataset the suggested framework is assessed against the most advanced classifiers. The simulation analysis shows the outperformance of the tuned LSTM for emotion prediction.

Declaration Statements

Data Availability

The corresponding author will receive the data upon request.

Statement of Conflict of Interest

No conflicts of interest have been declared by any of the writers.

Reference

[1] M. M. Rahman *et al.*, “Recognition of human emotions using EEG signals: A review,” *Comput. Biol. Med.*, vol. 136, p. 104696, Sep. 2021, doi: 10.1016/J.COMPBIOMED.2021.104696.

[2] H. Yadav and S. Maini, “Electroencephalogram based brain-computer interface: Applications, challenges, and opportunities,” *Multimed. Tools Appl.*, vol. 82, no. 30, p. 1, Dec. 2023, doi: 10.1007/S11042-023-15653-X.

[3] A. Phukan and D. Gupta, “Deep feature extraction from EEG signals using xception model for emotion classification,” *Multimedia Tools and Applications 2023 83:11*, vol. 83, no. 11, pp. 33445–33463, Sep. 2023, doi: 10.1007/S11042-023-16941-2.

[4] Garima, N. Goel, and N. Rathee, “Modified multidimensional scaling on EEG signals for emotion classification,” *Multimedia Tools and Applications 2023 82:18*, vol. 82, no. 18, pp. 28547–28568, Feb. 2023, doi: 10.1007/S11042-023-14671-Z.

[5] N. S. Suhaimi, J. Mountstephens, and J. Teo, “EEG-Based Emotion Recognition: A State-of-the-Art Review of Current Trends and Opportunities,” *Comput. Intell. Neurosci.*, vol. 2020, no. 1, p. 8875426, Jan. 2020, doi: 10.1155/2020/8875426.

[6] Y. Zhang, G. Yan, W. Chang, W. Huang, and Y. Yuan, “EEG-based multi-frequency band functional connectivity analysis and the application of spatio-temporal features in emotion recognition,” *Biomed. Signal Process. Control*, vol. 79, p. 104157, Jan. 2023, doi: 10.1016/J.BSPC.2022.104157.

[7] M. N. Dar, M. U. Akram, S. G. Khawaja, and A. N. Pujari, “CNN and LSTM-Based Emotion Charting Using Physiological Signals,” *Sensors 2020, Vol. 20,* vol. 20, no. 16, pp. 1–26, Aug. 2020, doi: 10.3390/S20164551.

[8] L. Gong, W. Chen, M. Li, and T. Zhang, “Emotion recognition from multiple physiological signals using intra- and inter-modality attention fusion network,” *Digit. Signal Process.*, vol. 144, p. 104278, Jan. 2024, doi: 10.1016/J.DSP.2023.104278.

[9] J. Zhang, Z. Yin, P. Chen, and S. Nichele, “Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review,” *Information Fusion*, vol. 59, pp. 103–126, Jul. 2020, doi: 10.1016/J.INFFUS.2020.01.011.

[10] P. Zhong, D. Wang, and C. Miao, “EEG-Based Emotion Recognition Using Regularized Graph Neural Networks,” *IEEE Trans. Affect. Comput.*, vol. 13, no. 3, pp. 1290–1301, 2022, doi: 10.1109/TAFFC.2020.2994159.

[11] M. Jafari *et al.*, “Emotion recognition in EEG signals using deep learning methods: A review,” *Comput. Biol. Med.*, vol. 165, Oct. 2023, doi: 10.1016/j.compbiomed.2023.107450.

[12] D. Kulkarni and V. Vitthalrao Dixit, “EEG-based emotion classification Model: Combined model with improved score level fusion,” *Biomed. Signal Process. Control*, vol. 95, p. 106352, Sep. 2024, doi: 10.1016/J.BSPC.2024.106352.

[13] A. Mert and A. Akan, “Emotion recognition based on time–frequency distribution of EEG signals using multivariate synchrosqueezing transform,” *Digit. Signal Process.*, vol. 81, pp. 106–115, Oct. 2018, doi: 10.1016/J.DSP.2018.07.003.

[14] M. Murugappan, B. S. Zheng, and W. Khairunizam, “Recurrent Quantification Analysis-Based Emotion Classification in Stroke Using Electroencephalogram Signals,” *Arabian Journal for Science and Engineering 2021 46:10*, vol. 46, no. 10, pp. 9573–9588, Feb. 2021, doi: 10.1007/S13369-021-05369-1.

[15] T. Tuncer, S. Dogan, M. Baygin, and U. Rajendra Acharya, “Tetromino pattern based accurate EEG emotion classification model,” *Artif. Intell. Med.*, vol. 123, p. 102210, Jan. 2022, doi: 10.1016/J.ARTMED.2021.102210.

[16] A. B. Das and M. I. H. Bhuiyan, “Discrimination and classification of focal and non-focal EEG signals using entropy-based features in the EMD-DWT domain,” *Biomed. Signal Process. Control*, vol. 29, pp. 11–21, Aug. 2016, doi: 10.1016/J.BSPC.2016.05.004.

[17] M. Z. Baig, N. Aslam, H. P. H. Shum, and L. Zhang, “Differential evolution algorithm as a tool for optimal feature subset selection in motor imagery EEG,” *Expert Syst. Appl.*, vol. 90, pp. 184–195, Dec. 2017, doi: 10.1016/J.ESWA.2017.07.033.

[18] S. Taran and V. Bajaj, "Sleep Apnea Detection Using Artificial Bee Colony Optimize Hermite Basis Functions for EEG Signals," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 2, pp. 608–616, Feb. 2020, doi: 10.1109/TIM.2019.2902809.

[19] A. Liu, K. Chen, Q. Liu, Q. Ai, Y. Xie, and A. Chen, "Feature Selection for Motor Imagery EEG Classification Based on Firefly Algorithm and Learning Automata," *Sensors 2017, Vol. 17,* vol. 17, no. 11, Nov. 2017, doi: 10.3390/S17112576.

[20] H. He, Y. Tan, J. Ying, and W. Zhang, "Strengthen EEG-based emotion recognition using firefly integrated optimization algorithm," *Appl. Soft Comput.*, vol. 94, p. 106426, Sep. 2020, doi: 10.1016/J.ASOC.2020.106426.

[21] V. Gupta, M. Mittal, V. Mittal, S. Diwania, R. Singh, and V. Gupta, "A Firefly based Deep Belief Signal Specification based Novel Hybrid Technique for EEG Signal Analysis," *IETE J. Res.*, vol. 70, no. 5, pp. 5263–5269, May 2024, doi: 10.1080/03772063.2023.2220698.

[22] F. Shen, G. Dai, G. Lin, J. Zhang, W. Kong, and H. Zeng, "EEG-based emotion recognition using 4D convolutional recurrent neural network," *Cognitive Neurodynamics 2020 14:6*, vol. 14, no. 6, pp. 815–828, Sep. 2020, doi: 10.1007/S11571-020-09634-1.

[23] M. K. Chowdary, J. Anitha, and D. J. Hemanth, "Emotion Recognition from EEG Signals Using Recurrent Neural Networks," *Electronics 2022, Vol. 11,* vol. 11, no. 15, Jul. 2022, doi: 10.3390/ELECTRONICS11152387.

[24] N. S. Malan and S. Sharma, "Feature selection using regularized neighbourhood component analysis to enhance the classification performance of motor imagery signals," *Comput. Biol. Med.*, vol. 107, pp. 118–126, Apr. 2019, doi: 10.1016/J.COMPBIOMED.2019.02.009.

[25] A. Hag *et al.*, "Enhancing EEG-Based Mental Stress State Recognition Using an Improved Hybrid Feature Selection Algorithm," *Sensors 2021, Vol. 21,* vol. 21, no. 24, Dec. 2021, doi: 10.3390/S21248370.

[26] Z. Tian, D. Li, Y. Song, Q. Gao, Q. Kang, and Y. Yang, "EEG-Based Emotion Recognition of Deaf Subjects by Integrated Genetic Firefly Algorithm," *IEEE Trans. Instrum. Meas.*, vol. 70, 2021, doi: 10.1109/TIM.2021.3121473.

[27] Garima, N. Goel, and N. Rathee, "Novel fractal pattern based features for EEG-based emotion identification," *Biomed. Signal Process. Control*, vol. 96, p. 106505, Oct. 2024, doi: 10.1016/J.BSPC.2024.106505.

[28] L. Cohen, "Generalized Phase-Space Distribution Functions," *J. Math. Phys.*, vol. 7, no. 5, pp. 781–786, May 1966, doi: 10.1063/1.1931206.

[29] H. I. Choi and W. J. Williams, "Improved Time-Frequency Representation of Multicomponent Signals Using Exponential Kernels," *IEEE Trans. Acoust.*, vol. 37, no. 6, pp. 862–871, 1989, doi: 10.1109/ASSP.1989.28057.

[30] P. Bonato, S. H. Roy, M. Knaflitz, and C. J. De Luca, "Time frequency parameters of the surface myoelectric signal for assessing muscle fatigue during cyclic dynamic contractions," *IEEE Trans. Biomed. Eng.*, vol. 48, no. 7, pp. 745–753, 2001, doi: 10.1109/10.930899.

[31] Q. Gao, C. han Wang, Z. Wang, X. lin Song, E. zeng Dong, and Y. Song, "EEG based emotion recognition using fusion feature extraction method," *Multimedia Tools and Applications 2020 79:37*, vol. 79, no. 37, pp. 27057–27074, Jul. 2020, doi: 10.1007/S11042-020-09354-Y.

[32] R. S. Parpinelli, G. F. Plichoski, R. S. Da Silva, and P. H. Narloch, "A review of techniques for online control of parameters in swarm intelligence and evolutionary computation algorithms," *International Journal of Bio-Inspired Computation*, vol. 13, no. 1, pp. 1–20, 2019, doi: 10.1504/IJBIC.2019.097731.

[33] N. Nedjah, L. de Macedo Mourelle, and R. G. Morais, "Inspiration-wise swarm intelligence meta-heuristics for continuous optimisation: A survey - part II," *International Journal of Bio-Inspired Computation*, vol. 16, no. 4, pp. 195–212, 2020, doi: 10.1504/IJBIC.2020.112340.

[34] R. S. Parpinelli and H. S. Lopes, "New inspirations in swarm intelligence: A survey," *International Journal of Bio-Inspired Computation*, vol. 3, no. 1, pp. 1–16, 2011, doi: 10.1504/IJBIC.2011.038700.

[35] I. Boussaïd, J. Lepagnot, and P. Siarry, "A survey on optimization metaheuristics," *Inf. Sci. (N Y).*, vol. 237, pp. 82–117, Jul. 2013, doi: 10.1016/J.INS.2013.02.041.

[36] C. Dong, Z. Xiong, X. Liu, Y. Ye, Y. Yang, and W. Guo, "Dual-Search Artificial Bee Colony Algorithm for Engineering Optimization," *IEEE Access*, vol. 7, pp. 24571–24584, 2019, doi: 10.1109/ACCESS.2019.2899743.

[37] L. Jovanovic *et al.*, "Machine learning tuning by diversity oriented firefly metaheuristics for Industry 4.0," *Expert Syst.*, vol. 41, No. 2, P. E13293, Feb. 2024, doi: 10.1111/Exsy.13293;Journal:Journal:14680394;Csubtype:String:Special;Page:String:Article/Chapter.

[38] D. Maheshwari, S. K. Ghosh, R. K. Tripathy, M. Sharma, and U. R. Acharya, "Automated accurate emotion recognition system using rhythm-specific deep convolutional neural network technique with multi-channel EEG signals," *Comput. Biol. Med.*, vol. 134, p. 104428, Jul. 2021, doi: 10.1016/J.COMPBIOMED.2021.104428.

[39] Q. Zhou, C. Shi, Q. Du, and L. Ke, "A multi-task hybrid emotion recognition network based on EEG signals," *Biomed. Signal Process. Control*, vol. 86, p. 105136, Sep. 2023, doi: 10.1016/J.BSPC.2023.105136.

[40] W. Li, Y. Tian, B. Hou, J. Dong, S. Shao, and A. Song, "A Bi-Stream hybrid model with MLPBlocks and self-attention mechanism for EEG-based emotion recognition," *Biomed. Signal Process. Control*, vol. 86, p. 105223, Sep. 2023, doi: 10.1016/J.BSPC.2023.105223.

[41] L. Gong, M. Li, T. Zhang, and W. Chen, "EEG emotion recognition using attention-based convolutional transformer neural network," *Biomed. Signal Process. Control*, vol. 84, p. 104835, Jul. 2023, doi: 10.1016/J.BSPC.2023.104835.

[42] N. Bacanin, C. Stoean, M. Zivkovic, M. Rakic, R. Strulak-Wójcikiewicz, and R. Stoean, "On the Benefits of Using Metaheuristics in the Hyperparameter Tuning of Deep Learning Models for Energy Load Forecasting," *Energies 2023, Vol. 16,* vol. 16, no. 3, Feb. 2023, doi: 10.3390/EN16031434.

[43] W. X. Cheng, R. Gao, P. N. Suganthan, and K. F. Yuen, "EEG-based emotion recognition using random Convolutional Neural Networks," *Eng. Appl. Artif. Intell.*, vol. 116, p. 105349, Nov. 2022, doi: 10.1016/J.ENGAPPAI.2022.105349.

[44] J. Nayak, B. Naik, P. B. Dash, A. Souri, and V. Shanmuganathan, "Hyper-parameter tuned light gradient boosting machine using memetic firefly algorithm for hand gesture recognition," *Appl. Soft Comput.*, vol. 107, p. 107478, Aug. 2021, doi: 10.1016/J.ASOC.2021.107478.

[45] B. Chakravarthi, S. C. Ng, M. R. Ezilarasan, and M. F. Leung, "EEG-based emotion recognition using hybrid CNN and LSTM classification," *Front. Comput. Neurosci.*, vol. 16, p. 1019776, Oct. 2022, doi: 10.3389/FNCOM.2022.1019776/BIBTEX.

[46] Y. Liu, H. Zhou, M. Guan, F. Feng, and J. Duan, "Scalp EEG-Based Automatic Detection of Epileptiform Events via Graph Convolutional Network and Bi-Directional LSTM Co-Embedded Broad Learning System," *IEEE Signal Process. Lett.*, vol. 30, pp. 448–452, 2023, doi: 10.1109/LSP.2023.3263433.

[47] Y. Wei, Y. Liu, C. Li, J. Cheng, R. Song, and X. Chen, "TC-Net: A Transformer Capsule Network for EEG-based emotion recognition," *Comput. Biol. Med.*, vol. 152, p. 106463, Jan. 2023, doi: 10.1016/J.COMPBIOMED.2022.106463.

[48] C. Li, Z. Zhang, R. Song, J. Cheng, Y. Liu, and X. Chen, "EEG-Based Emotion Recognition via Neural Architecture Search," *IEEE Trans. Affect. Comput.*, vol. 14, no. 2, pp. 957–968, Apr. 2023, doi: 10.1109/TAFFC.2021.3130387.