ISSN: 2182-2069 / E-ISSN: 2182-2077

# An Improved EEG Signal Feature Selection Paradigm for Migraine Detection

Asha[1], S.J. Namitha[2], Prameetha Pai[3], Priya Kamath[4*], Rajeshwari Madli[5], and R. Vijaya Arjunan[6*]

[1]Assistant Professor, Department of Artificial Intelligence and Machine Learning, Dayananda Sagar College of Engineering, Bengaluru, India. ashak_aiml@dayanandasagar.edu, https://orcid.org/0009-0000-4458-6806

[2]Assistant Professor, Department of Computer Science and Engineering, JSS Academy of Technical Education Bengaluru, Karnataka, India. namithasj@jssateb.ac.in, https://orcid.org/0009-0006-5488-0009

[3]Assistant Professor, Department of Computer Science and Engineering, B.M.S. College of Engineering, Bengaluru, Karnataka, India. prameethapai.cse@bmsce.ac.in, https://orcid.org/0000-0003-2526-2931

[4*]Assistant Professor - Senior Scale, Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education (MAHE), Manipal, Udupi, Karnataka, India. priya.kamath@manipal.edu, https://orcid.org/0000-0002-5471-8822

[5]Assistant Professor, Department of Computer Science and Engineering, B.M.S. College of Engineering, Bengaluru, Karnataka, India. madlirajeshwari.cse@bmsce.ac.in, https://orcid.org/0009-0001-9809-533X

[6*]Additional Professor, Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education (MAHE), Manipal, Udupi, Karnataka, India. vijay.arjun@manipal.edu, https://orcid.org/0000-0002-1402-6573

Received: March 05, 2024; Revised: May 04, 2024; Accepted: June 10, 2024; Published: August 30, 2024

## Abstract

Recurrent headaches, autonomic nervous system dysfunction, nausea, and vomiting are the hallmarks of migraine, a complicated neurological illness that has a substantial negative influence on suffering individuals. With cutting-edge deep learning techniques and electroencephalography (EEG) data, this study presents a novel approach to automatic migraine detection. The Improved novel Hybrid Feature Extraction for Migraine Detection (IHFEMD) model that has been suggested takes a thorough strategy that includes pre-processing the data, extracting features, selecting features using a hybrid optimization technique, and utilizing a hybrid deep learning framework. To create a highly refined dataset, migraine sufferers' raw EEG data must first undergo a thorough pre-processing step that includes data cleaning and Min-Max normalization. As a result, relevant

*Journal of Internet Services and Information Security (JISIS)*, volume: 14, number: 3 (August), pp. 143-156.
DOI: *10.58346/JISIS.2024.I3.008*

*Corresponding author: Assistant Professor - Senior Scale, Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education (MAHE), Manipal, Udupi, Karnataka, India; Additional Professor, Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education (MAHE), Manipal, Udupi, Karnataka, India.

migraine-related data are obtained using time-frequency analysis methods such as Continuous Wavelet Transform (CWT) and Short-Time Fourier Transform (STFT). The Elephant Herding Optimization (EHO) method is integrated with the Bald Eagle Optimization (BEO) method to optimize the process. Due to this, a much more accurate selection is made. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are used in tandem by the detection framework to combine the advantages of both network types and accomplish migraine detection. Complicated EEG can be identified with the help of this combined approach. The model has been put into practice with the Matlab platform. The result shows an accuracy of 99% and is preferred as a potential diagnostic tool. The novelty of the proposed work focuses on integrating deep learning frameworks with complex pre-processing, sophisticated feature extraction, and hybrid optimization techniques.

**Keywords:** Continuous Wavelet Transform, Short-Time Fourier Transform, Elephant Herding Optimization, Bald Eagle Optimization, CNN.

# 1 Introduction

Migraine is a neurovascular disease, characterized by frequent headaches and generally accompanied by a wide range of symptoms (Youssef & Mack, 2020; Özer & Benlier, 2020). It is a highly prevalent condition mostly affecting females. Migraine usually lasts from 4 to 72 hours if not treated (Fila et al., 2021). Migraine is accompanied by pain that throbs or pulses on one side of the head or sometimes on both sides, vomiting, high sensitivity to light, smell, touch, sound, and vomiting. The migraine pain can be so severe that it even affects our day-to-day activities (Özaltun et al., 2021; Syme & Hagen, 2020). These attacks can last for a few hours or days. For some people, a warning symptom, referred to as aura occurs before the headache (Ashina et al., 2021). An aura may include visual distortions like flashes of blind spots or light, pins or needles sensations in the arms or legs, numbness in one side of the face, loss of vision, and difficulty in speaking (Talbot et al., 2021). Medications help prevent some categories of migraines and help in making them less painful. The lifestyle changes and the right medicines combined with self-help remedies might help in preventing migraines (Taufique et al., 2021). The exact causes of Migraine are not yet identified but genetics, neurotransmitter abnormalities, and environmental factors appear to play a major role (Hougaard et al., 2020). The quality of life of a migraine sufferer can be improved by early diagnosis and effective treatment. With better management of symptoms, early detection, and effective medications, the quality of a person suffering from Migraine generally becomes much easier (Waliszewska-Prosół & Budrewicz, 2021). The creation of standardized migraine diagnostic criteria is challenged by this variety of symptoms (Donnet et al., 2021; Karsan et al., 2021). Determining the presence and intensity of migraine can also be challenging due to individual differences in reporting pain and associated symptoms as well as personal experiences (Kim et al., 2021; Rogers et al., 2020). Electroencephalography (EEG) is a milestone in diagnostic instruments which can record the electrical activities of the brain. Neurological mechanisms behind migraine can be studied by analyzing EEG data, which will help in quick and accurate diagnosis. There are a few obstacles to overcome in the analysis and extraction of relevant data from EEG because of its complexity and high dimensionality. A major portion of data contains raw EEG signals which are not relevant for identifying migraines (Dlugaiczyk et al., 2020; Heiliczer et al., 2022). An efficient feature selection is therefore necessary. A range of machine learning algorithms have been applied to overcome the difficulties associated with the feature selection procedure for the interpretation of EEG. Numerous strategies (Mitrović et al., 2023; Göker, 2023) including filter, wrapper, and embedding methods have been investigated and their advantages and disadvantages are examined. The effectiveness of these strategies is based on the specifics of EEG data (Conti et al., 2024; Yaqoob Yousif & Alsakaa, 2024). Used for Models/Methods shown in Table 1.

# 2 Related Work

Table 1: Used for Models/Methods

| Cited Works | Models/Methods Used | Limitations | Results |
|---|---|---|---|
| Aslan (2021) | Presents a combined method using Q-factor Wavelet transform and ensemble-based learning for detecting acute migraine using EEG signals. | Standardization is needed for TQWT parameters that are found experimentally | An accuracy of 89% was achieved. |
| Subasi et al., (2019) | Investigates the use of different ML models for migraine fdetection using EEG signals. Discrete Wavelet Transform is used for feature extraction. | High potential cost | An accuracy of 85% has been achieved using the Random Forest algorithm. |
| Orhanbulucu et al., (2023) | The recorded EEG signals are processed using "Continuous Wavelet Transform (CWT) and Short-Time Fourier Transform (SFT)" to create spectrogram and scalogram images. | The proposed study is limited to a single dataset. As a result, it may not generalize well across varying populations. | Accuracy of 99.6% is achieved. |
| Cao et al., (2019) | Extraction of SSVEP-based Inherent Fuzzy Entropy in Migraine Patients Using a Wearable Headband EEG signal. | No differences in EEG entropy among groups were observed. | An accuracy of 81% was achieved using the AdaBoost ensemble learning classification model. |
| Li et al., (2022) | Resting-state EEG was recorded in 61 patients with migraine without aura (MwoA), out of which 50 were female patients, and in 66 healthy controls (HC), also with 50 females. Microstate parameters were then compared among these two groups. | It cannot be applied to other forms of migraines. | Among the MwoA patients, microstate B and D classes observed higher time coverage and occurrence compared to class C which had lower time coverage, and occurrence compared to HC. |
| Abbas Abdulhussein et al., (2022) | Discussed the lack of habituation in migraine patients as determined by high-density EEG analysis using the steady state of visual evoked potential | Limited size of the sample and it provides less information on the effects of medication. | Cortical Spreading Depression (CSD) is the primary cause of two types of migraine- one with aura and another without aura. |

# 3 Methodology

Migraine is considered a neurological disorder that is characterized by recurrent headaches accompanied by nausea and vomiting, leading to disturbances in the autonomic nervous system. These episodes of severe pain are prolonged and can significantly impact the quality of life. In this study, a deep learning model named IHFEMD has been proposed to assist the expert judgment in the automatic detection of migraines using EEG signals. The block diagram of the proposed migraine classification model is given in Figure 1.

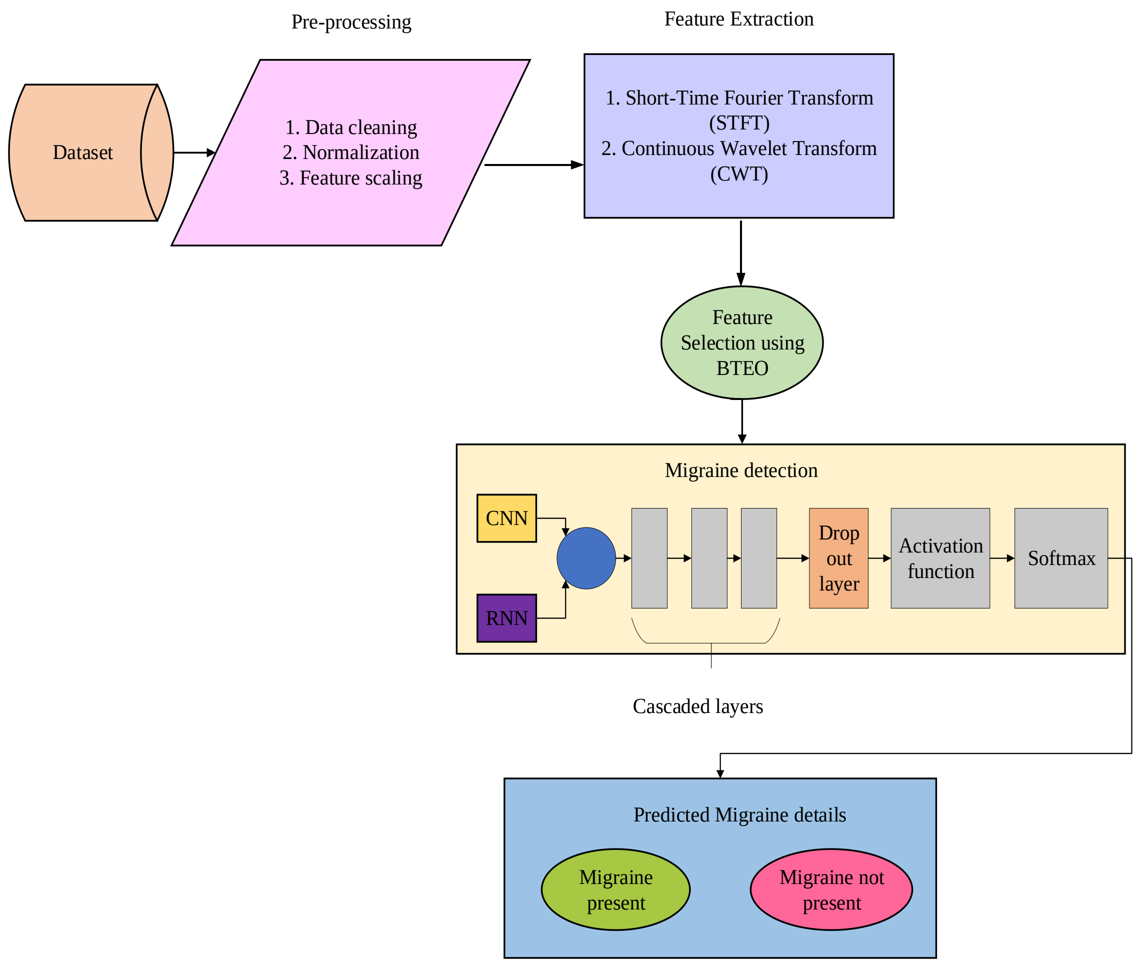


Figure 1: Proposed IHFEMD Model for Migraine Disease Detection

The various steps involved in the proposed methodology are given below.

**a. Pre-processing:** The raw data obtained for migraine patients needs to be cleaned to eliminate the presence of noise or outliers. This process will involve data cleaning, data normalization, and feature scaling.

*Data Cleaning*: Artifacts are inherent in all EEG readings and must be removed to ensure the proper functioning of the classifier. Data cleaning primarily involves the elimination of irregularities from the original data, such as duplicates, missing values, and outliers. Duplicate data is removed when experimental data is repeated. Missing values are imputed using the average of data from preceding and succeeding hours. The formula is as follows Eq. (1):

$$x_i = \frac{x_{i-1}+x_{i+1}}{2} \tag{1}$$

Where $x_i$ refers to the data to be filled, $x_{i-1}$ is the data from the earlier hour of padding data, and $x_{i+1}$ is the data from the subsequent hour of padding data.

*Normalization:* Scaling the feature to a standard scale is done by normalization. All returned characteristics are normalized using Min-Max normalization to improve the performance and accuracy of the classifier. When $X$ is the sample data, $x_{min}$ and $x_{max}$ are the lowest and highest values for each channel, and X is the sample data. The major aim here is to allot all additional data to this 0–1 range while normalizing the lowermost value to 0 and the highest value to 1. The formula is as follows Eq. (2):

$$Norm_{min.max} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (2)$$

*Feature Scaling:* The MCEEG data are then normalized using Feature Scaling (FS), which limits the values between two arbitrary points $a\ and\ b$, and the normalized value is expressed as $h(x)$ using Eq. (3):

$$h(x) = a + \frac{(x - x_{min}) * (b - a)}{x_{max} - x_{min}} \quad (3)$$

Where $x$ denotes the sample data and $x_{min}$ and $x_{max}$ denote the minimum and maximum of each channel. In the suggested technique, the sizes of $x_{min}$ and $x_{max}$ will be $M\ x1.\ a\ and\ b$ take on values 0 and 1.

**b. Feature Extraction:** After pre-processing, relevant features will be extracted from the migraine patient data. A time-frequency analysis technique such as STFT and CWT is used to extract relevant features from the EEG signals.

**Signal's Fourier Transform (STFT)** sequence is employed for feature extraction. It has been chosen as the finest diagnosis tool for detecting abnormalities in a variety of fields, and it works with both stationary and non-stationary data. The STFT is utilized as a spectral analysis in this article to detect the appearance of an abnormality. Notably, spectral analysis is an effective diagnostic technique. As a result, the STFT coefficients are derived by splitting the signal into small portions with a sliding window in the time domain. The conventional mathematical definition of the STFT is illustrated. The formula is as follows Eq. (4):

$$S(n, \omega) = \sum_{m=-\infty}^{\infty} x[m] w[n - m] e^{-j\omega n} \quad (4)$$

Where $x[m]w[n - m]$ is a short-time portion of the input signal $x\ (m)$ at time $n$ and frequency $k$. Additionally, a discrete STFT is defined as follows Eq. (5):

$$S(n, k) = |S(n, \omega|_{\omega = \frac{2\pi k}{N}} \quad (5)$$

Where $N$ is the number of discrete frequencies. As a result, the spectrogram is defined as follows:

$$Spec(n, k) = |S(n, k)|^2 \quad (6)$$

As a result, the spectrogram is a visual solution for representing signal intensity and development in the time-frequency domain; to construct it, we employ the magnitude or squared magnitude of STFT coefficients, as shown in Eq. (6).

In this work, the CWT is used to extract important characteristics from one-dimensional (1D) EEG recordings. Wavelets, which are minuscule oscillations that are highly localized in time, are used by CWT to break down the signals. As a result, CWT is used to build a time-frequency representation of a signal. This provides superior frequency and temporal localization. CWT is used to convert every signal into a coefficient matrix. Eq. (7) displays the CWT transformation equation. The formula used is:

$$CWT_x^{(\psi)}(a, b) = W(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \,.\, \psi^* \left(\frac{t - b}{a}\right) dt \quad (7)$$

Input signal $x(t)$, scale $a$ translation $b$, normalization factor $\psi$, mother wavelet $\left(\frac{t-b}{a}\right)$, and scale factor tab. Mother wavelets are used by CWT to scale and shift the input into parts. To extend or contract the signal in time, the scaling factor, which is inversely related to frequency, is utilized. Consequently, both rapid and gradual changes in the signal may be captured. The shifting factor is utilized to translate the wavelets with various scales over the input signal $x(t)$.

**c. Feature Selection:** The extracted features undergo a feature selection process to identify the most relevant ones. This involves employing a hybrid optimization algorithm named BTEO, which integrates elements from both the EHO and BEO algorithms. This combined strategy enables a thorough exploration of the feature space, achieving a balance between cooperative and individual optimization strategies. The detailed procedural steps for the BTEO algorithm are outlined below,

### i) Initialization

In the original EHO, the beginning population for the search is chosen at random. As a result, it is not necessary to know the objective function or restrictions beforehand. Chaotic maps offer superior statistical and dynamic qualities than randomization, which has been shown through experimentation. The typical random number generator is replaced in this article with a pre-programmed Gaussian sequence, which is shown in Eq. (8) as follows.

$$\eta(t+1) = \begin{cases} 0, \eta(t) = 0 \\ \frac{1}{\eta(t)} - \frac{1}{\eta(t)}, Otherwise \end{cases} \quad (8)$$

The number of chaotic maps produced during the current and next generations is shown by the symbols $\eta(t)$ and $\eta(t+1)$. The Gaussian chaos mapping function, which may explore the space more thoroughly and produce superior exploration results, creates the initialized population.

### ii) Fitness Calculation

The fitness can be calculated using the Eq. (9).

$$Fitness = Max(A) \quad (9)$$

Where $A$ represents accuracy. A certain number of clans are created from the population of all agents based on the fitness computation.

### iii) Enhanced Random Wandering Operator

The populations that the clan operator creates have a propensity to stray. As a result, the algorithm has a limited capacity to leave the local optimum. In our opinion, the patriarch should have a ground-breaking and creative updating strategy because he or she is each clan's best-positioned agent. In the exploitation phase, agents concentrate on finding better solutions in the previously investigated areas. As a result, it is important to maximize people's ability to discover the best answer so that the algorithm can quickly converge to a workable local or global optimal solution. Taking this into account, the patriarch's updating approach is modified as follows Eq. (10):

$$x_{best,c_i}^{t+1} = \begin{cases} x_{best}^{t} + C(\sigma) * \left(\frac{Iter}{max_{iter}}\right), If \ \frac{Iter}{max_{iter}} < 0 \\ x_{gbest}^{t} + 2(rand - 0.5)(x_b - x_c), Otherwise \end{cases} \quad (10)$$

Where $x_{best}^{t}$ and $x_{best,c_i}^{t+1}$ are the patriarch's current and most recent locations in clan $c_i$, rand is a random integer between [0,1] created by the tent chaotic map, $x_{gbest}^{t}$ is the best solution globally, and $C(\sigma)$ stands for the Cauchy distributed random number. A random walk based on a Cauchy distribution is capable of aiding in global exploration. Following is a definition of the Cauchy distribution function. The formula is as follows Eq. (11):

$$F(\sigma; a, b) = \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{\sigma - a}{b}\right) \tag{11}$$

Where $a$ is the location parameter and $b$ is the scale parameter. In the standard Cauchy distribution, $a = 0, b = 1$.

### iv) Searching Phase

The bald eagle in this phase searches for prey by moving in different directions within the selected spiral space from the previous phase. Additionally, the best position for the hunting of prey and swooping is determined. This behavior is mathematically defined as follows Eq. (12-16).

$$x_i^{t+1} = x_i^t + z(i) \times \left(x_i^t - x_i^{t+1}\right) + p(i) \times (x_i^t - x_{mean}^t) \tag{12}$$

$$p(i) = \frac{\mathrm{pr}\,(i)}{\max\,(|pr|)}, z(i) = \frac{zr(i)}{\max\,(|zr|)} \tag{13}$$

$$\mathrm{pr}(i) = r(i) \times \cos\big(\theta(i)\big), zr(i) = r(i) \times \sin(\theta(i)) \tag{14}$$

$$\theta(i) = \alpha \times \pi \times r_1 \tag{15}$$

$$r(i) = \theta(i) + R \times r_2 \tag{16}$$

Where $R$ is a constant parameter with a value between [0.5, 2], $\alpha$ is a constant parameter with a value between [0.5, 2], and $r_1\ and r_2$ are two random parameters.

### v) Enhanced Mutation

A variance probability (PSR) is chosen such that most agents will have the chance to mutate. To not go above the population size limit, this option should accept a value between (0,1). It is challenging for the experiment to have a significant impact if PSR is less than 0.2, which indicates that fewer individuals experience mutation. In contrast to the setting's original goal, the algorithm will decide that most people will engage in the mutation if the PSR is higher than 0.8. The mutation operator in BTEO is configured as shown below. The formula is as follows (Eq). (17-19):

$$x_{worst,c_i}^{t+1} = x_{worst,c_i}^t + \delta_m r_1 + K \tag{17}$$

$$K = u_1 e^{-2t/\max iter} \tag{18}$$

$$x_i^{t+1} = \begin{cases} x_i^t + r_2\left(\frac{x_{gbest}^t + x_{clan}}{\alpha . x_{pbest}^t} - x_i^t\right) + r_3\left(\frac{x_{gbest}^t - x_{clan}}{\alpha . x_{pbest}^t} - x_i^t\right), rand < PSR \\ x_i^t + \alpha * r_4(x_{best}^t - x_t), rand \geq PSR \text{ and } rand < 2 * PSR \\ x_i^t + \beta * r_5(x_{worst}^t - x_t), and \geq 2 * PSR \end{cases} \tag{19}$$

Where $r_1, r_2, r_3$, and $r_4$ represent random numbers that are distributed uniformly between 0 and 1. $\alpha\ and\ \beta$ refers to the scaling factor. $x_{worst,c_i}^{t+1}$ indicates the agent's position to be modified and δ refers to the variation factor. In this paper, δ = 0.1 $*\ (x_{max} - x_{min})$.. u 1 is a random variable of [− 1, 1]. $x_{worst,c_i}^{t+1}$ indicates the solution that is optimal at the t$^{th}$ iteration.

### vi) Greedy Selection Strategy

A greedy selection approach then assesses the search agents' solutions. If the fitness value of the new agent is greater than the fitness value of the existing agent, it is replaced and added to a round of iteration procedures. Guarantee BTEO convergence is the goal. The formula is as follows (Eq). (20):

$$x_{gbest}^t = \begin{cases} x_i^t, f\left(x_i^t\right) < f(x_{gbest}^t) \\ x_{gbest}^t, else \end{cases} \tag{20}$$

**vii) Save and Return**

$$x_{pbest}^{t} \text{ and } x_{best}^{t}.$$

**viii) Terminate the Process**

After selecting the features, the selected features are given to the classifiers called CNN and RNN for further classification of migraine disease.

**d. Classification:** A hybrid deep learning approach will be used for migraine detection. Specifically, a CNN and a RNN are combined to form a hybrid deep-learning architecture. The spatial features from the EEG signals are extracted using CNN and the temporal dependencies are captured using RNN. In the context of migraine detection, the CNN can identify spatial patterns in EEG signals that are indicative of migraine events. Combining the strengths of CNNs and RNNs in a hybrid architecture allows for a more comprehensive analysis of EEG data. This holistic approach may lead to improved accuracy in migraine detection. The performance metrics for both the proposed model and the existing techniques like Bi-LSTM, RNN, CNN (Conti et al., 2004), ANN (Orhanbulucu et al., 2023), and RF (Göker, 2023) along with the comparison are shown in Table 2.

# 4 Results and Discussion

The performance and effectiveness of the BTEO algorithm in optimizing the feature selection process for the detection of Migraine are evaluated. By integrating Elephant Herding Optimization (EHO) with Bald Eagle Optimization (BEO) algorithms, BTEO aims to identify the most relevant features from extracted data, thereby enhancing the accuracy and reliability of migraine diagnosis models. The effectiveness of BTEO with traditional approaches in choosing the best feature subsets has been compared and the results are tabulated.

Table 2. presents the performance metrics of various classification models, evaluated using different measures such as sensitivity and specificity, and the same is plotted in Figure 2.

Table 2: Overall Comparison of the Proposed IHFEMD Model with Existing Models

| Technique | Sensitivity | Specificity | Accuracy | Precision | Recall | F-Measure | NPV | FPR | FNR | MCC |
|---|---|---|---|---|---|---|---|---|---|---|
| **Proposed** | 0.9640 | 0.9995 | 0.9900 | 0.9878 | 0.9895 | 0.9842 | 0.9964 | 0.0188 | 0.0548 | 0.9726 |
| **Bi-LSTM** | 0.9377 | 0.9940 | 0.9744 | 0.9825 | 0.9822 | 0.9856 | 0.9934 | 0.0283 | 0.1301 | 0.9608 |
| **RNN** | 0.9280 | 1.0002 | 0.9600 | 0.9767 | 0.9788 | 0.9798 | 0.9901 | 0.0749 | 0.1550 | 0.8839 |
| **CNN [18]** | 0.9054 | 0.9861 | 0.9484 | 0.9695 | 0.9759 | 0.9755 | 0.9800 | 0.0777 | 0.1432 | 0.8760 |
| **ANN [ 22]** | 0.8299 | 0.9789 | 0.9219 | 0.8980 | 0.8945 | 0.8976 | 0.9769 | 0.0637 | 0.2172 | 0.8092 |
| **RF [17]** | 0.8038 | 0.9344 | 0.9035 | 0.8923 | 0.8912 | 0.8907 | 0.9306 | 0.1132 | 0.2957 | 0.7474 |

Sensitivity, also referred to as recall, helps in measuring the model's ability to correctly detect true positive cases. In this context, the IHFMD model shows a high sensitivity value of 0.9640, which means 96.40% of the positive cases are correctly identified. Specificity, in contrast, assesses the capability of the model to correctly detect the negative cases. The IHFEMD technique showcases an exceptional specificity of 0.9995, signifying that it accurately classifies 99.95% of actual negative cases. The efficiency of the IHFEMD model in accurately detecting positive and negative instances is suggested by its high sensitivity and specificity values. It's important to consider the application's requirements when interpreting these values—favoring high sensitivity when missing positive cases is critical, and prioritizing high specificity when minimizing false positives is paramount. Figure 3 shows the comparison of accuracy and precision. The proposed IHFEMD model properly predicts 99.00% of all cases in the context of the table, earning a high accuracy score of 0.9900. Contrarily, precision evaluates the model's ability to make accurate positive predictions. The proposed IHFEMD method specifically exhibits an accuracy of 0.9878, which means that 98.78% of occurrences projected as positive are indeed true positives. These impressive accuracy and precision ratings for the proposed approach highlight its prowess in creating correct forecasts generally and in making predictions that are favorable for classes, respectively.

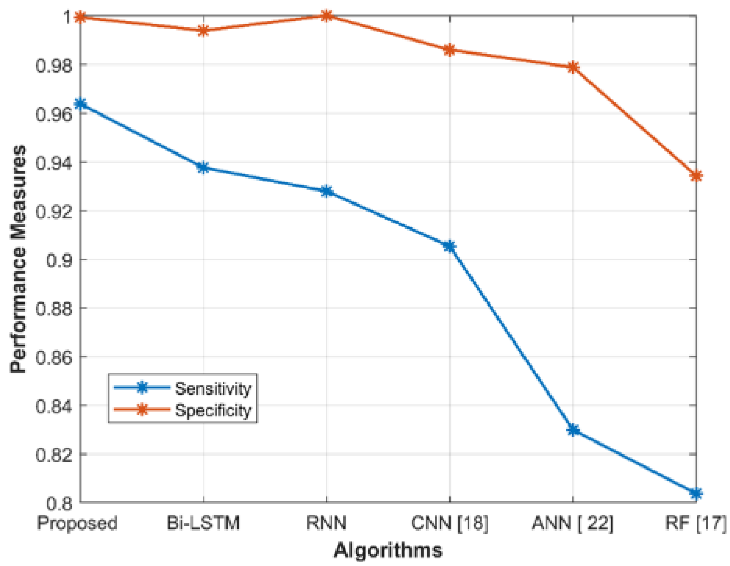


Figure 2: Comparison of the Sensitivity and Specificity

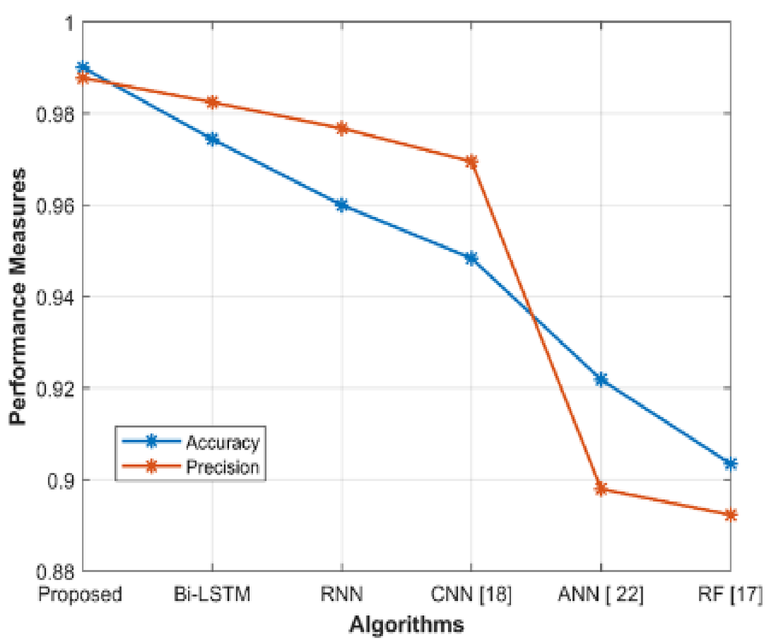


Figure 3: Comparison of the Accuracy and Precision

In Figure 4, the FNR and FPR are displayed. The statistic known as the False Positive Rate (FPR) quantifies the model's tendency to predict favorably when the actual outcome is negative. In the table, the proposed approach's FPR is 0.0188. It can be seen from this that the model predicts false positives at a rate of 1.88%, meaning that very few negative events are inadvertently classified as positive. The False Negative Rate (FNR) assesses the model's propensity to forecast occurrences as negative when the actual result is positive. When the cost of falsely predicting something negatively, FNR is important. By dividing the total of true positives and false negatives by the number of erroneous negative forecasts, it is determined. The FNR for the IHFMD method in the table is 0.0548. This indicates that only a tiny percentage of real positive cases are missed by the model, which has a false negative prediction rate of 5.48%.

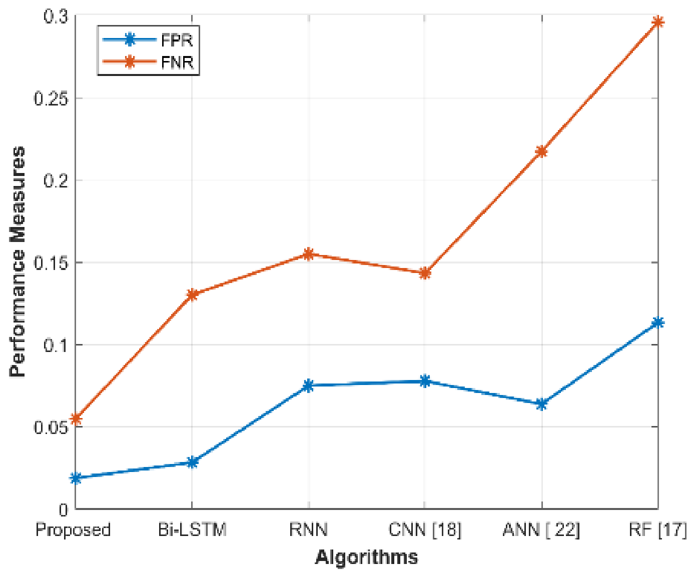


Figure 4: Comparison of the FNR and FPR

## 5 Conclusion

The most challenging problem of identifying Migraine automatically using deep learning methods on EEG data has been carried out in this work. The proposed model IHFMD provides effective pre-processing of data followed by enhanced feature extraction. The feature selection phase is performed using a novel hybrid optimization procedure along with a hybrid architecture of a deep learning model. The dataset is pre-processed by cleaning the raw EEG data collected from persons suffering from Migraine. The cleaned data is then normalized using the min-max normalization. Using sophisticated time-frequency analysis techniques, Migraine-related relevant features are retrieved from the processed EEG signals. The combination of CWT along with STFT enables the extraction of necessary features essential in the Migraine patterns. As a result, the dataset becomes more relevant for further analysis when the ideal subset of features is identified by the BTEO algorithm. Therefore, there is a major need for developing an innovative hybrid deep learning framework. The most unique qualities of both RNN and CNN are identified and combined synergistically. While RNN detects the temporal relationships of typical episodes, the CNN component efficiently extracts the spatial characteristics from the recordings of EEG signals.

The final accuracy of the forecast is greatly increased by combining the two findings. Therefore, the proposed IHFMD model outperforms other state-of-the-art methods in w.r.t. accuracy, recall, precision, and f1-score.

## References

[1] Abbas Abdulhussein, M., Alyasseri, Z. A. A., Mohammed, H. J., & An, X. (2022). Lack of habituation in migraine patients based on high-density eeg analysis using the steady state of visual evoked potential. *Entropy*, *24*(11), 1688. https://doi.org/10.3390/e24111688.

[2] Ashina, M., Goadsby, P. J., Reuter, U., Silberstein, S., Dodick, D. W., Xue, F., & Mikol, D. D. (2021). Long-term efficacy and safety of erenumab in migraine prevention: results from a 5-year, open-label treatment phase of a randomized clinical trial. *European Journal of Neurology*, *28*(5), 1716-1725.

[3] Aslan, Z. (2021). Migraine detection from EEG signals using tunable Q-factor wavelet transform and ensemble learning techniques. *Physical and Engineering Sciences in Medicine*, *44*(4), 1201-1212.

[4] Cao, Z., Lin, C. T., Lai, K. L., Ko, L. W., King, J. T., Liao, K. K., & Wang, S. J. (2019). Extraction of SSVEPs-based inherent fuzzy entropy using a wearable headband EEG in migraine patients. *IEEE Transactions on Fuzzy Systems*, *28*(1), 14-27.

[5] Conti, M., Bovenzi, R., Palmieri, M. G., Placidi, F., Stefani, A., Mercuri, N. B., & Albanese, M. (2024). Early effect of onabotulinumtoxin A on EEG-based functional connectivity in patients with chronic migraine: A pilot study. *Headache: The Journal of Head and Face Pain*. https://doi.org/10.1111/head.14750.

[6] Dlugaiczyk, J., Habs, M., & Dieterich, M. (2020). Vestibular evoked myogenic potentials in vestibular migraine and Menière’s disease: cVEMPs make the difference. *Journal of Neurology*, *267*, 169-180.

[7] Donnet, A., Ducros, A., Radat, F., Allaf, B., Chouette, I., & Lanteri-Minet, M. (2021). Severe migraine and its control: A proposal for definitions and consequences for care. *Revue Neurologique*, *177*(8), 924-934.

[8] Fila, M., Chojnacki, C., Chojnacki, J., & Blasiak, J. (2021). Nutrients to improve mitochondrial function to reduce brain energy deficit and oxidative stress in migraine. *Nutrients*, *13*(12), 4433. https://doi.org/10.3390/nu13124433.

[9] Göker, H. (2023). Automatic detection of migraine disease from EEG signals using bidirectional long-short term memory deep learning model. *Signal, Image and Video Processing*, *17*(4), 1255-1263.

[10] Heiliczer, S., Wilensky, A., Gaver, T., Georgiev, O., Hamad, S., Nemirovski, A., & Haviv, Y. (2022). Salivary endocannabinoid profiles in chronic orofacial pain and headache disorders: an observational study using a novel tool for diagnosis and management. *International Journal of Molecular Sciences*, *23*(21), 13017. https://doi.org/10.3390/ijms232113017.

[11] Hougaard, A., Younis, S., Iljazi, A., Sugimoto, K., Ayata, C., & Ashina, M. (2020). Intravenous endothelin-1 infusion does not induce aura or headache in migraine patients with aura. *Headache: The Journal of Head and Face Pain*, *60*(4), 724-734.

[12] Karsan, N., Bose, P., Newman, J., & Goadsby, P. J. (2021). Are some patient-perceived migraine triggers simply early manifestations of the attack?. *Journal of Neurology*, *268*, 1885-1893.

[13] Kim, S., Bae, D. W., Park, S. G., & Park, J. W. (2021). The impact of pain-related emotions on migraine. *Scientific Reports*, *11*(1), 577. https://doi.org/10.1038/s41598-020-80094-7.

[14] Li, Y., Chen, G., Lv, J., Hou, L., Dong, Z., Wang, R., & Yu, S. (2022). Abnormalities in resting-state EEG microstates are a vulnerability marker of migraine. *The journal of headache and pain*, *23*(1), 45. https://doi.org/10.1186/s10194-022-01414-y.

[15] Mitrović, K., Petrušić, I., Radojičić, A., Daković, M., & Savić, A. (2023). Migraine with aura detection and subtype classification using machine learning algorithms and morphometric magnetic resonance imaging data. *Frontiers in Neurology*, *14*, 1106612. https://doi.org/10.3389/fneur.2023.1106612

[16] Orhanbulucu, F., Latifoğlu, F., & Baydemir, R. (2023). A new hybrid approach based on time frequency images and deep learning methods for diagnosis of migraine disease and investigation of stimulus effect. *Diagnostics*, *13*(11), 1887. https://doi.org/10.3390/diagnostics13111887.

[17] Özaltun, M. F., Geyik, S., & Yılmaz, Ş. G. (2021). Screening for copy number variations of the 15q13. 3 hotspot in CHRNA7 gene and expression in patients with migraines. *Current Issues in Molecular Biology*, *43*(2), 1090-1113.

[18] Özer, G., & Benlier, N. (2020). Neck pain: is it part of a migraine attack or a trigger before a migraine attack?. *Acta neurologica Belgica*, *120*, 289-293.

[19] Rogers, D. G., Protti, T. A., & Smitherman, T. A. (2020). Fear, avoidance, and disability in headache disorders. *Current Pain and Headache Reports*, *24*, 1-8.

[20] Subasi, A., Ahmed, A., Aličković, E., & Hassan, A. R. (2019). Effect of photic stimulation for migraine detection using random forest and discrete wavelet transform. *Biomedical Signal Processing and Control*, *49*, 231-239.

[21] Syme, K. L., & Hagen, E. H. (2020). Mental health is biological health: Why tackling "diseases of the mind" is an imperative for biological anthropology in the 21st century. *American Journal of Physical Anthropology*, *171*, 87-117.

[22] Talbot, J., Stuckey, R., Crawford, L., Weatherby, S., & Mullin, S. (2021). Improvements in pain, medication use and quality of life in onabotulinumtoxinA-resistant chronic migraine patients following erenumab treatment–real world outcomes. *The Journal of Headache and Pain*, *22*, 1-10.

[23] Taufique, Z., Zhu, B., Coppola, G., Shoaran, M., & Altaf, M. A. B. (2021). A low power multi-class migraine detection processor based on somatosensory evoked potentials. *IEEE Transactions on Circuits and Systems II: Express Briefs*, *68*(5), 1720-1724.

[24] Waliszewska-Prosół, M., & Budrewicz, S. (2021). The unusual course of a migraine attack during COVID-19 infection—Case studies of three patients. *Journal of Infection and Public Health*, *14*(7), 903-905.

[25] Yaqoob Yousif, A., & Alsakaa, A. (2024). Migraine types identification based on EEG signals using machine learnings techniques. *International Journal of Computing and Digital Systems*, *16*(1), 1-15.

[26] Youssef, P. E., & Mack, K. J. (2020). Episodic and chronic migraine in children. *Developmental Medicine & Child Neurology*, *62*(1), 34-41.

## Author's Biography

**K. Asha,** holds a Bachelor of Engineering (B.E.) in Computer Science and Engineering, Master of Technology in Computer Science and Engineering. She is currently working as an Assistant Professor in Artificial Intelligence and Machine Learning department in Dayananda Sagar College of Engineering, Bangalore, with over 12 years of teaching experience. Her research areas of interest include Image processing, Machine learning, and Computer Vision.

**S.J. Namitha,** holds a Bachelor of Engineering (B.E.) in Computer Science and Engineering, Master of Technology in Computer networking and Engineering, she is currently working as an assistant Professor in Computer Science and engineering department in JSS Academy of Technical Education Bangalore, with over 10 years of teaching experience, her focused researched fields are Computer networks, mobile application develpment, image processing, she is a member of IEI, she has an excellent knowledge in user interface design techniques for a Good Web Design.

**Prameetha Pai,** holds a Bachelor of Engineering (B.Eng.) in Computer Science and Engineering, Master of Technology in Computer networking and Engineering. She is also pursing her Ph. D in deep learing and precision agriculture from Visvesvaraya Technological University, Belgaum. She is currently an assistant professor in department of Computer Science and engg in B.M.S. College of Engineering. Her research areas of interest include Image processing, Deep learning, and Digital signal Processing.

**Priya Kamath,** holds a Bachelor of Engineering (B.E.) and Master of Technology (M.Tech.) in Computer Science and Engineering. She is currently pursuing the Ph.D. degree with the Manipal Academy of Higher Education, Manipal, India. She is an Assistant Professor with the Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education. Her research interests include Natural LAnguage Processing, Machine Learning, Deep Learning, Sentiment Analysis, and Opinion Mining.

**Rajeshwari Madli,** received the B.E. degree in Computer Science and Engineering in 2007, and the M.Tech. degree in Computer Science and Engineering in 2015, from Visvesvaraya Technological University, Belgaum, India. She is also pursuing her Ph.D. in the field of Vehicular Ad-hoc Networks from Visvesvaraya Technological University, Belgaum, India. She is currently working as Assistant Professor in Department of Computer Science and Engg, B.M.S. College of Engineering. Her research interests include vehicular networks and sensor networks, cloud computing and machine learning.

**R. Vijaya Arjunan,** has been serving as an Additional Professor in the Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, since July 2010. During his tenure at Manipal Institute of Technology, he was on deputation to the School of Engineering and IT, Manipal, Dubai, from 2014 to 2017. He obtained his Master of Engineering and Ph.D. in Computer Science and Engineering from Sathyabama Institute of Science, Technology and Sankara University in 2005 and 2013 respectively. He has published around 50+ research articles in various International Conferences and Journals. His research interests include Computer Vision, Image Processing, Machine Learning, Deep Learning, and Data Mining.