International Journal of Computing and Digital Systems
2025, VOL. 17, NO. 1, 1–15

http://dx.doi.org/10.12785/ijcds/1571038488

Migraine types identification based on EEG signals using machine learning techniques

Ayat Yaqoob Yousif[1] and Akeel Alsakaa[1]

[1]*Department of Computer Science, University of Kerbala, Kerbala, Iraq*

Received 13 June 2024, Revised 1 February 2025, Accepted 2 February 2025

Abstract: Migraine (MD) is a neurological disorder that can be accompanied by auditory and visual symptoms called aura and affects the lives of many people around the world. It causes temporary disability and may progress to serious diseases such as epilepsy or stroke, which affect individual health and limit societal productivity. The overlap of migraine symptoms with those of various other diseases makes identifying and diagnosing migraines challenging and time-consuming. We developed two machine learning models to help doctors diagnose and differentiate between types of migraine—both with and without aura—and to enhance patient care beyond traditional methods that rely on interrogation or visual analysis of complex, non-linear Electroencephalography (EEG) signals that require time and expertise. The first model focuses on diagnosing migraine versus healthy controls (HC) using (EEG) obtained from Steady State Auditory Evoked Potentials (SSAEP) and Steady State Visual Evoked Potentials (SSVEP) stimuli from 17 migraine patients and 20 healthy controls. These EEG signals were analyzed using Discrete Wavelet Transform (DWT) and Fast Fourier transform (FFT) to extract the alpha, beta, delta, theta, and gamma frequencies, which were then used to train machine learning algorithms. The second model, on the other hand, expands the visual stimulus data to include 503 participants, enabling the diagnosis of migraine and the differentiation between its two main types—migraine without aura (MwoA) and migraine with aura (MwA)—compared to healthy individuals. Both models achieved a classification accuracy of over 90%, effectively identifying migraine and distinguishing between its main types. This innovative approach enhances the accuracy and efficiency of migraine diagnosis and provides valuable insights into the disorder's neurological underpinnings. By integrating advanced signal processing techniques with machine learning, our model represents a significant advancement in the medical field, offering a more efficient and accurate method for diagnosing migraines and improving patient care.

Keywords: Electroencephalography, Migraine without Aura, Migraine with Aura, Discrete Wavelet Transform, Fast Fourier transform

1. Introduction

Migraine is a debilitating neurovascular condition Characterized by spells of headache with accompanying autonomic and perhaps neurological symptoms, vomiting, and nervous system malfunction [1]. Approximately one-third of migraine patients experience transient neurological disturbances before, during, or after their headaches, known as migraine aura [2].An aura is typically a visual illusion or sensory experience, frequently described as tingling or numbness in the face, arms, or other parts of the body. According to numerous studies, migraine with aura (MwA) is associated with cortical hyper-responsiveness and altered sensory information processing compared to migraine without aura (MwoA) or healthy individuals, even during the interictal phase [3].Migraines are the second most common neurological condition in the world, causing more disability than all other neurological diseases combined [4], [5]. Migraine is frequently misdiagnosed since its symptoms coincide with those of other conditions such as tension headache, epilepsy, and stroke [6].Therefore, studies have shown that relying on and analyzing EEG data is an effective way to detect migraines and other neurological diseases, EEG is a neuroimaging technique that measures electrical impulses produced by electrodes applied to the scalp to record brain activity [7]. The EEG signals are inexpensive, non-radioactive, and non-invasive. As a result, they are now frequently employed to identify brain abnormalities [8], [9]. The most significant benefit of EEG is its exceptionally high temporal resolution, which makes it possible to capture electrical impulses thousands of times per second [7]. Since the nature of EEG signals is non-linear, a trained neurologist is needed to investigate abnormal EEG patterns associated with these disorders. Efficiency varies greatly in the visual evaluation of these signals. Evaluating long

E-mail address: ayat.y@s.uokerbala.edu.iq, akeel.a@uokerbala.edu.iq

EEG recordings can be tedious manually, and results may not always be consistent. With a little human help, the automated system can identify neurological conditions and track brain activity [10] to develop an effective model that supports doctors in making diagnostic decisions. The brain's electrical activity is recorded using electrodes that adhere to international standards, such as the 10-20 system. This record is referred to as electroencephalogram (EEG) activity. EEG data represent the electrical activity of the human brain, which fluctuates in response to neurological conditions. EEG signals are key indicators of neurological diseases and can help with disease diagnosis [11]. Where EEG, magnetic resonance imaging, and computed tomography are employed to supplement expert judgment in the disease diagnosis. However, because EEG is less expensive and requires less equipment, it is favored for disease identification in computer-assisted diagnosis systems [12]. Many researchers employ EEG signals, particularly for the diagnosis of neurological illnesses. Machine learning algorithms classify features extracted from EEG data in various ways. From this perspective, the relevant literature includes research that uses EEG signals and machine learning algorithms to diagnose migraines [13].

EEG recordings are obtained by attaching electrodes to the scalp using a conductive gel. Each active electrode's signal is amplified relative to a reference electrode using a differential amplifier [14]. EEG signals were first discovered in 1875 by Richard Caton, a physician studying electrical brain activity in rabbits and monkeys. In the early 1900s, the first human EEG recordings were made, focusing on absent seizures (when a person loses consciousness for a short time). By the 1930s, the identification of epileptic spikes and seizure patterns spurred significant interest in EEG as a new field [14]. EEG waveforms vary by frequency band. The delta band has the highest amplitude and is frontal in adults and posterior in children. The theta band is common in young children and indicates drowsiness or arousal in adults, often spiking during active inhibition. The alpha band is linked to eye movements and is found in the posterior regions on both sides of the head. The beta band is associated with motor behavior and is located in the frontal regions on both sides of the head [15].

This research aims to develop a binary classification model capable of detecting and diagnosing migraines in comparison to healthy controls. Subsequently, the dataset used in the binary classification model is expanded to create a ternary classification model, which can identify the main types of migraines (MwA and MwoA) versus healthy controls. This is the first study to identify the main types of migraines based on data from Carnegie Mellon University, after expanding and enhancing this data set.

2. Materials And Methods

This section describes the dataset used, preprocessing techniques, and the method for extracting important features for later use as inputs for machine learning classifiers. Section 1 and Figure 1 present the proposed approach scheme, which shows an overview of the proposed method for diagnosing migraine using EEG signals.

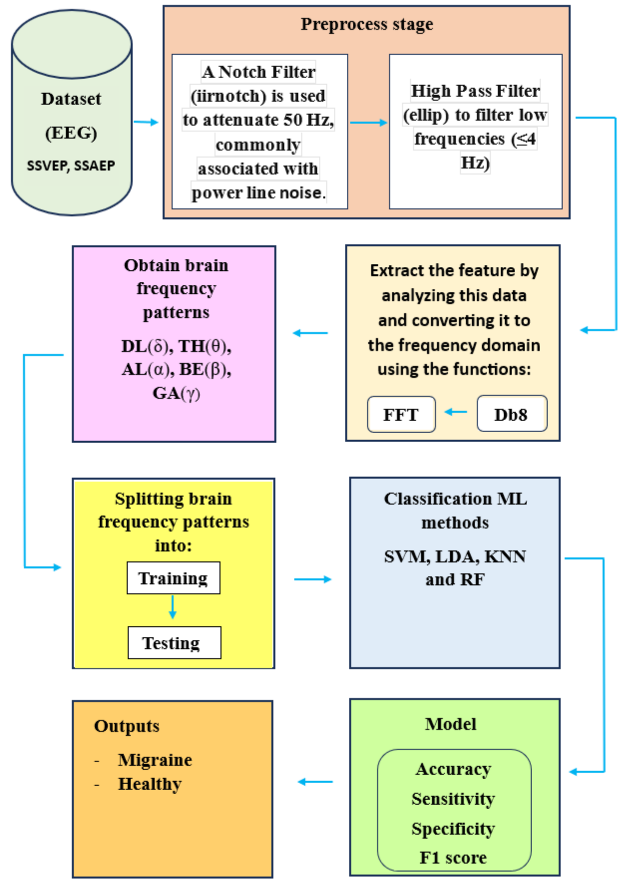

Figure 1. The Migraine Classification System's structural diagram.

The following steps describe the methodology used in this study:

- Data Collection: Electroencephalographic (EEG) signals were recorded from migraine patients during auditory and visual stimulation using 15 channels which have been reported in the scientific literature as having an association with migraine pain sites. These channels include (Fp1, F7, C3, Pz, Fp2, Fz, F8, Cz, C4, F3, F4, P3, P4, O1, O2), with their locations illustrated in Figure 2. The resulting data for each participant was organized into a matrix of size (15 × 307201), where the first dimension (15) represents the number of EEG channels. and the second dimension (307201) represents the number of time samples extracted within the specified time range from (100 to 700 seconds).
- Signal Cleaning: The signals were processed to remove interference from power lines and artifacts caused by eye movements and muscle activity.

- Signal Analysis: Several feature extraction techniques can be applied to these overlapping EEG signals to facilitate further analysis and processing [13]. One such technique is the use of the Daubechies wavelet transform (db8), which allows for the decomposition of the signal into multiple frequency bands, namely Delta, Theta, Alpha, Beta, and Gamma. After wavelet decomposition, the Fast Fourier Transform (FFT) is applied to convert them from the time domain to the frequency domain. The most frequent frequency in each of the five frequency bands (Delta, Theta, Alpha, Beta, and Gamma) is then extracted for each of the 15 EEG channels. These most common frequencies are stored in the feature matrix for later use in machine learning classifiers [16], [17].

- Data Representation and Dimensionality Reduction: Two methods were tested for data representation to reduce dimensions and organize data in a way that is easier to enter into classifiers and train it.

- Several classification algorithms, such as SVM, LDA, Random Forest, KNN, and Decision Trees, are applied to classify the migraine conditions.

- Data Augmentation: Data augmentation techniques were applied to improve the robustness of the model. This included calculating the minimum and maximum values for the delta, theta, alpha, beta, and gamma bands independently for each band and generating random values within those limits. This was applied independently to all groups (HC, MwoA, and MwA) to ensure that the augmented data represented frequency bands similar to those in the original data while adding sufficient variance to increase the model's ability to generalize.

- Hyperparameter Tuning: Grid search was used to find the best hyperparameters for SVC and RF classifiers.

- Cross-validation: Cross-validation was applied to assess model performance.

- Model Evaluation: Accuracy, sensitivity, specificity, and F1 score were used to evaluate the model's effectiveness in diagnosing migraines.

A. Dataset Collecting Stage

This study categorized migraine sufferers and healthy individuals using an EEG dataset recently released on KiltHub, Carnegie Mellon University's online data repository. The EEG data was recorded using the Bio Semi-Active Two device with a 512 Hz sampling frequency and a 24-bit analog-to-digital (A/D) converter, consisting of 128 channels [18]. The dataset includes data from 18 migraine sufferers and 21 control participants, with ages ranging from 19 to 54 years (13 females, 5 males in the migraine group; 12 females, 9 males in the control group), under visual and auditory stimulation [12]. Participants were recruited from

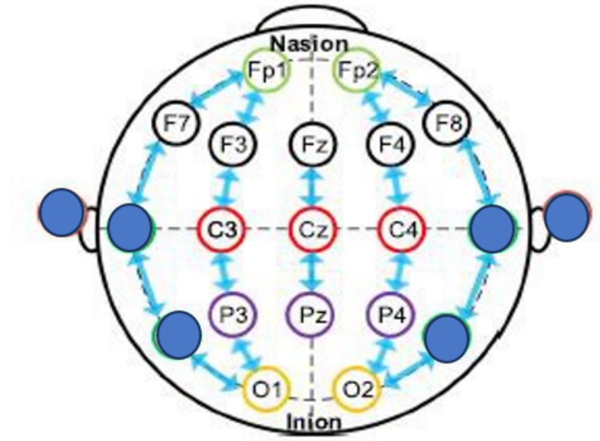

Figure 2. EEG electrode placement.

Carnegie Mellon University and the Pittsburgh area. They did not have any neurological or psychological diagnoses (except for migraine), no history of severe head injury or trauma, normal hearing, and, according to their reports, normal or corrected-to-normal eyesight. All procedures were reviewed and approved by the Carnegie Mellon University Institutional Review Board.

The visual stimulation (SSVEP) involves vertically displaying black-and-white longitudinal wave patterns (gratings) with a spatial frequency of 0.05 cycles per degree of visual angle (cpd), covering an area with a diameter of 5.7 cpd, and displayed at the center of the screen 1 meter away from the participant. Patterns were smoothed using a 2D Gaussian filter to reduce sharp edges. The patterns change in contrast at a time-frequency of 4 Hz or 6 Hz for 2 s each time, followed by an interval of 1 to 1.5 s. Each temporal frequency was presented 100 times in a randomized order. A central fixation cross with a spatial frequency of 0.5 cpd was displayed throughout the experiment, superimposed on the gratings to ensure participants maintained visual focus.

The auditory stimuli (SSAEP) included two-second presentations of 1 kHz tones modulated by a sinusoidal carrier frequency of 4 Hz or 6 Hz. At 44.1 kHz, the stimuli were captured with 16-bit resolution, and 100 repetitions of each carrier frequency were made. The tones were separated by a silent inter-stimulus interval of one to one and a half seconds. Throughout the experiment, participants focused on a black central fixation cross shown on a grey screen while the tones were played using insert earphones (Etymotic Research, Inc.). Reference [19] provides detailed information on the dataset and experimental setting. In our study, participant M13 was excluded due to the lack of electroencephalography (EEG) recordings of the auditory stimulation. In addition, Participant C12 was excluded to reduce the difference between the number of healthy controls and migraine patients, thus reducing potential bias towards the healthy group with a larger number. The other reason for excluding Participant C12 was that the recording duration of his signal was shorter than that of the others.

B. Preprocessing Stage

Two experiments were conducted: the first for visual stimulation and the second for auditory stimulation.
Recordings from 15 channels were chosen because of their association with migraine pain areas, as well as to reduce the complexity of the work. These channels have been mentioned previously. To read these signals and store them in an array, they were converted from BDF to EDF, as it is one of the most commonly used formats for EEG data and is most compatible with our research requirements, using the EDF Browser software. In the preprocessing step, EEG signals are filtered to reduce noise before being separated into specific signal fragments [20]. Filtering techniques have been used to filter out signals from unwanted frequencies. Electroencephalogram (EEG) data is recorded with interference and noise from many sources, such as electrical devices, lighting, and other electronic devices, causing unwanted signals to be introduced to the electrodes of the EEG device. Signal interference can also occur when a patient's eyes move during an EEG recording. This causes changes in the recorded electrical signal.

First, a notch filter removes the specific high frequency of 50 Hz. This means the filter reduces the strength of any signal containing this frequency. The elimination of AC interference caused by Power-line (PLI) from biological signals, such as the ECG and the EEG, to discrete wavelet form is one of the classic uses of digital notch filters [21]. Then, a high pass filter is used to remove low frequencies. In this case, frequencies below 4 Hz are filtered out. This means that the filter removes low-frequency signals and leaves only high-frequency signals. Thus, low-frequency signals (below 4 Hz) and high-frequency signals (above 50 Hz) are eliminated, thus preserving signals in the intermediate frequency range between 4 and 50 Hz. The filtering approach seeks to eliminate all noise and interference, improving the signal-to-noise ratio and thereby improving classification accuracy outcomes [22].

In the final stage, the detrend function is used with the 'constant' option to eliminate the signal's mean value, essentially eliminating any constant bias or offset. After applying the notch filter to remove power-line interference (50 Hz) and the high-pass filter to reduce low frequencies (below 4 Hz), the detrend function calculates the mean of the filtered signal and subtracts it from each data point. This method enhances signal quality by removing any constant offset, allowing for easier analysis of the signal's true physiological components. For example, if an EEG signal has a baseline drift due to sensor movement, the detrend function will rectify this by centering the signal around zero. But it will retain all the original changes and patterns in the data. This means that the signal will still contain the important information but without any consistent bias or skew, allowing for a more accurate interpretation of brain activity.

C. Feature Extraction

After the waves are read and filtered, a high-dimensional array of overlapping signals is obtained. To reduce their complexity and focus on the most important data for our work, these attributes must be significant for model learning tools. Therefore, they must be discriminative and non-redundant, allowing the data to be fully utilized. This is achieved by selecting optimal features while minimizing the number of features (dimensionality reduction) [23]. Feature extraction involves converting raw data into meaningful features,and capturing relevant information and patterns to facilitate analysis and modeling.

By extracting informative features using wavelet analysis, the data's dimensionality can be reduced, and noise can be filtered out, leading to improved performance of machine learning algorithms for learning and predictions. Different types of wavelets are available, wavelet families can be categorized into two main types: orthogonal and biorthogonal wavelets [24].

Orthogonal, translated, and dilated wavelets are orthogonal to each other. This reduces redundancy in conversion transactions, making data more efficient and less redundant. Among its types: Daubechies wavelets are a popular type of orthogonal wavelet, offering smoother scaling functions than Haar wavelets. They are defined by the number of vanishing moments, with orders ranging from db1 to db45. Higher orders improve the ability to analyze complex signals and filter out noise. In this study, we relied on the db8 wavelet, which uses eight coefficients to process adjacent signal parts. resulting in clearer signal reconstruction and reduced rapid fluctuations, making it smoother and more effective for signal analysis [24].

First, the frequency bands of each wave are determined based on the Nyquist frequency ratio, where the bands are defined as follows: Gamma (30-100 Hz), Beta (12-30 Hz), Alpha (8-12 Hz), Theta (4-8 Hz), and Delta (0.5-4 Hz). Next, bandpass filters are designed using a third-order Butterworth filter for each frequency band, and these filters are applied to the original signal using filtfilt, producing filtered signals representing each band. The Butterworth filter is an Infinite Impulse Response (IIR) filter that aims to generate a frequency response that is as flat as possible within its passband. Stephen Butterworth proposed it in 1930, and it is frequently used in digital signal processing applications such as biomedical, audio, and seismic. The filter refines signal spectra within a specified frequency range, known as the cut-off frequency, and the order influences the sharpness of attenuation in the transition band [25]. After the filtering process, wavelet analysis is applied. Discrete wavelet transform (DWT) allows the use of discrete wavelet coefficients (db8) to characterize EEG signals and transform them into discrete wavelet representations [21] to become increasingly important.

In the Discrete Wavelet Transform (DWT), the initial signal is broken down into two key components: approximation, which captures the low-frequency information, and detail, which highlights the high-frequency details. Following the first stage of decomposition, only the approximation component proceeds to undergo further decomposition, while the detail component remains unchanged. This iterative process continues until a predetermined level of decomposition is achieved [26]. The approximation component in wavelet analysis contains low-frequency information and is typically used to extract low-frequency activities such as theta and delta waves. On the other hand, the detail component often contains high-frequency information, which can be utilized to extract high-frequency activities such as alpha, beta, and gamma waves.

Therefore, the db8 function analyzes the EEG brain signal into several levels, to obtain five main frequency bands: GM (Gamma), BT (Beta), AP (Alpha), TH (Theta), and DT (Delta). Then fast Fourier transform (FFT) transforms these components from the time domain to the frequency domain. The output of a (FFT) is an array of frequency bins, where each bin represents a specific frequency. Each bin contains a composite value indicating the amplitude and phase of the frequency component. Thus, the output representation of the signal in the frequency domain [27]. The dominant frequencies of the alpha, beta, delta, theta, and gamma bands are extracted and stored in an array for later use. This is done through an iterative loop, which calculates the coefficients for each frequency band for each of the 15 EEG channels.

Table I summarizes the main characteristics of different brain wave rhythms and their functional correlates. It provides a useful reference for understanding the neurophysiological basis of various cognitive and behavioral processes.

Thus, we obtain a matrix (15 x5), where the first dimension (15) represents the number of channels that are relied upon to record brain electrical activity, and the second dimension (5) represents the sum of the brain frequencies known as delta, theta, alpha, beta, and gamma.

D. Migraine Classification

At this stage, we will begin by dividing the data into a training and test set to create a model capable of detecting the disease and supporting the opinion of doctors. We focused on machine learning classifiers because the size of the data is small and because most of these classifiers are characterized by a variety of parameters, which allows us to choose the optimal parameters to achieve better accuracy. In addition to their strength with digital data, and also because of their popularity in the field of neurology, and most importantly, they are subject to supervision because Our practice requires data classifications to train migraine-related features and distinguish them from those of healthy people, where after experimentation, it became clear that these classifiers have fast and good classification performance, which makes them better than others. Among these classifiers are:

Linear discriminant analysis (LDA) is frequently used to reduce dimensionality and identify a feature subspace in which the data samples are separable [2], Support vector machine (SVM) is used for regression analysis and classification. It aims to find the optimal hyperplane with the largest margin between classes in an n-dimensional classification space [36]. In this study, the Support Vector Classification (SVC) method was employed for the classification tasks. K-nearest neighbors (KNN) determine a similarity between training and testing instances using Euclidean distance. Nearest neighbors are determined based on these similarities, and the testing sample's class label is determined by majority voting. The choice of distance metric, K value, and majority voting method affect categorization performance [37], Random Forest (RF) classifiers are ensembles of randomly grown trees. Leaf nodes are labeled based on posterior distributions for different classes. Internal nodes have tests for data partitioning [23]. Randomness is introduced through subsampling the data and selecting node tests during training [36].

Classification involves aggregating predictions from individual trees to make the final prediction, Tree Decisions (DTs) are tree-like models used in supervised data mining. They consist of internal nodes representing attribute tests, branches reflecting test results, and leaf nodes indicating class names. The root node stores all tuples, and classification is achieved by branching and splitting based on data properties [38].

At this stage, two models will be created. The first model will focus on diagnosing migraines compared to healthy persons. In contrast, after efforts to expand the dataset, the second model will attempt to diagnose migraines and distinguish between its two main forms (MwoA and MwA) in addition to healthy individuals.

1) The First Model

This model used binary classification for migraine diagnoses versus healthy controls.

As a first step, the sum of the frequencies of each band (delta, theta, alpha, beta, and gamma) is calculated separately for all EEG channels for each participant, thus reducing each participant's data from 15 rows (channel) to one row representing the sum of those Frequencies for each frequency band. This statistical method helps reduce the dimensionality of the data, then we use the sum of frequencies as features, which are statically partitioned into a 75% percent training set and 25% test set using the Holdout method. The SVC algorithm was applied to the visual stimulus dataset with the following parameters: C = 1000, Gamma = 100, Kernel = rbf. The LDA was applied with the parameter discrimType = Linear, and the KNN

TABLE I. The main characteristics of brain wave rhythms EEG

Rhythms	Brain Wave Rhythms	Related Functions
Delta (δ)	0.5–4 Hz	show up in babies and during profound sleep [28], [29].
Theta (θ)	4-8 Hz	Children's brains process tasks in the temporal and parietal areas. The temporal manages hearing, memory, and recognizing faces/emotions, while the parietal handles sensory and motor information [30], [31].
Alpha (α)	8-12 Hz	It appears in an awake adult, where it can be found in the parietal and frontal areas and the scalp. It also appears in the occipital region [32].
Beta (β)	12-30 Hz	Low beta rhythm is associated with movement occurring whether through actual activity, planning, or visualization. This decrease is visible in the corresponding motor cortex, which controls the movement of the body on the other side of it. It can be detected during movement in the frontal and medial lobes of the head [33], [34].
Gamma (γ)	30-100 Hz	The rhythms with frequencies higher than 30 Hz are the higher ones. It has to do with how ideas are formed, how language is processed, and different kinds of learning [35].

algorithm was applied with K = 3. Similarly, the SVC algorithm was applied to the auditory stimulus dataset with the parameters C = 1, Gamma = 100, and Kernel = rbf. The LDA was used with discrimType = Linear, and KNN was applied with K = 3

This step was followed by dividing the data in a balanced manner and training it using the five-fold (K-fold) cross-validation technique, repeated five times. As for the classifier, the random forest algorithm was used, where the best results were searched using two loops, the first for the number of trees (100, 300, 500) and the second for the depth of those trees (5, 10, 15). This is to avoid bias towards a fixed division and to obtain a more comprehensive result than the previous one. The highest accuracy was achieved with 100 trees and a depth of 15 for the visual stimulus dataset, while the best results for the auditory stimulus dataset were also obtained with 100 trees and a depth of 15.

To explore another way to represent data as input features for machine learning classifiers, we considered all channels for each participant and transformed them into a vector. This resulted in 37 vectors for the visual and 37 vectors for the auditory stimulus conditions. These vectors were statically divided into 75% for training and 25% for testing, and the SVC algorithm was applied to the visual stimulus dataset with the following parameters: C= 0.001, Kernel= polynomial, and Gamma= 10. The LDA was applied with the parameter discrimType = Linear, and the KNN algorithm was used with K = 3. Similarly, the SVC algorithm was applied to the auditory stimulus dataset with the parameters C= 0.001, Kernel= polynomial, and Gamma= 10. The LDA was used with discrimType = Linear, and KNN was applied with K = 3.

To further validate the results, the vectors were partitioned in a balanced manner, the fixed partitioning was replaced with five-fold cross-validation repeated five times. The SVC classifier was evaluated using grid search to find the best parameters for the SVC classifier by testing different values of C [0.001, 0.01, 0.1, 1, 5, 10, 15, 25, 100] and Gamma [0.001, 0.01, 0.1, 1, 5, 10, 15, 25, 100], as well as kernel options (Linear, rbf, polynomial) where Grid Search is used to determine the parameters that achieve the highest average accuracy of the classifier across all folds. It ensures that the classifier trained with the best parameters is tested correctly and calculates various performance metrics. The best parameters for the visual stimulus group were C = 15, Gamma = 5, and Kernel = polynomial, while the best parameters for the auditory stimulus group were C = 25, and Kernel = linear.

Since the linear kernel does not need a gamma parameter to compute, KernelScale does not need to be specified when using KernelFunction = 'linear' in support vector machines. KernelScale is not required because the linear kernel computes the dot product of input features directly.

2) The Second Model

In this model, we relied on the tripartite classification for diagnosing migraine, identifying its two main types (MwoA and MwA) versus healthy controls. Since obtaining real data is challenging due to patient privacy,data augmentation techniques were applied to the visual stimulation data in the frequency domain, that is, after cleaning the signals and extracting the main features represented by alpha, beta, gamma, theta and delta frequencies.

Visual stimulation data was used to augment the data, as it is more commonly utilized in the health sector compared to auditory stimulation. The data for each category (HC, MwoA, and MwA) was expanded by generating random numbers within the minimum and maximum ranges of brain frequencies (alpha, beta, gamma, theta, and delta). This approach ensured that the augmented data represented frequency bands comparable to those in the original data while providing sufficient variability to enhance the model's generalizability. As a result, the dataset size increased from 37 to 503 individuals, distributed as follows: 160 healthy controls (HC), 167 MwA, and 170 MwoA.

Since the Random Forest classifier and the frequency summation approach performed well in the first model, they were also utilized in this model. The data is divided into five folds using K-fold cross-validation, with one-fold assigned for testing and the remaining folds for training. The Random Forest classifier is trained on balanced training data using a "bagging" technique, consisting of three learning cycles with individual decision trees. The test fold is then used for cross-validation to generate predictions.

3. Result

A. *Result Of The First Model*

Here, a binary classification was performed for migraine diagnosis versus healthy controls, using two methods to represent the data and reduce dimensions, given that each participant has 15 rows (channels).

In the first method, the sum of frequencies for each band across all channels for each participant was calculated to form a new dataset, which was then divided into 75% training and 25% test sets. SVC, KNN, and LDA classifiers were applied. The results for the visual and auditory conditions are shown in Figures 3 and Figures 4, respectively:

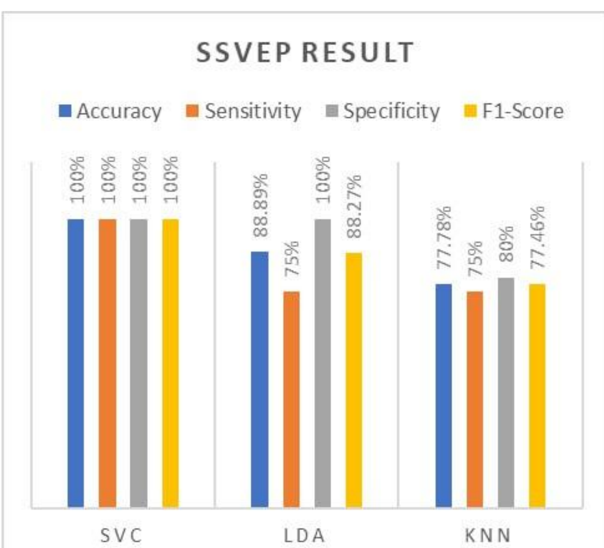

Figure 3. Classification Results for SSVEP Represented as Sum of Frequencies

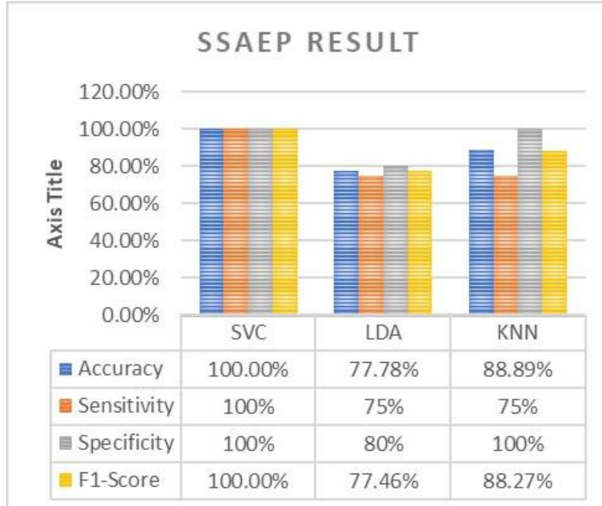

Figure 4. Classification Results for SSAEP Represented as Sum of Frequencies

Thus, we note the superiority of the SVC classifier compared to other classifiers.

The relationship between Gamma and C values significantly impacts the accuracy of the classifier (SVC) for visual stimulation data, as shown in Figure 5. The red dot shows the maximum accuracy achieved using the RBF kernel, emphasizing the importance of choosing appropriate parameters to improve classifier performance.

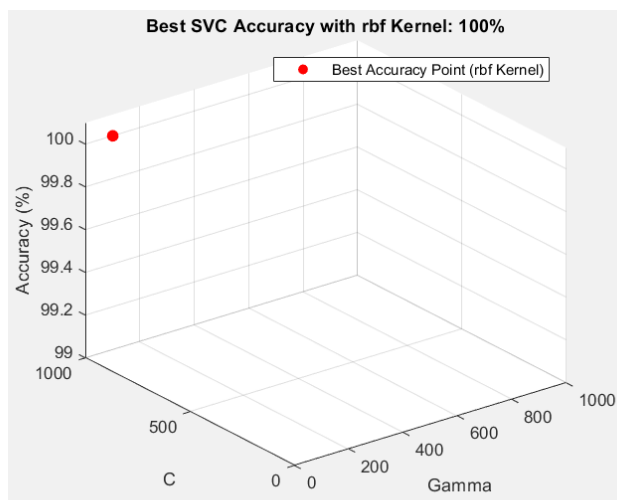

Figure 5. The SVC classifier's accuracy for the SSVEP

As shown in Figure 6, the ROC curve of the SVC classifier shows an optimal classification of the auditory stimulus data in this specific experiment when the data was split into a fixed training and test set, with an area under the curve (AUC) of one. The x-axis shows the false positive rate, and the y-axis represents the true positive rate. The ROC curve depicts how the true positive rate varies with the false positive rate as the decision threshold shifts. The curve approaches the top-left corner, suggesting high model performance. The AUC of one indicates that the model properly identifies the positive (migraine) and negative (healthy) classes when each participant's data is represented by calculating the sum of the frequencies for each band.

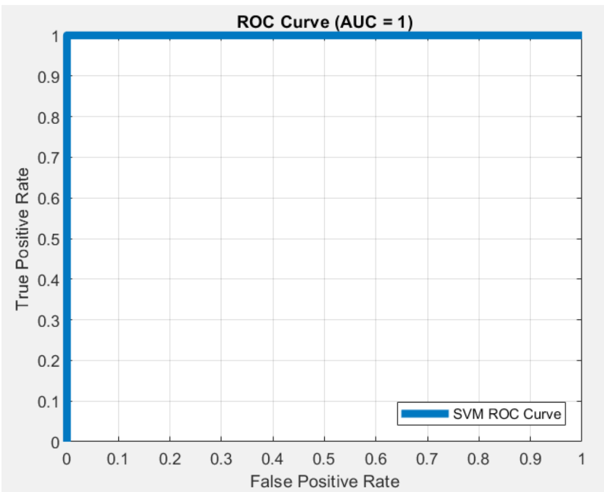

Figure 6. The SVC classifier's accuracy for the SSAEP

To evaluate the data comprehensively and monitor the results on unseen data, a five-fold cross-validation technique was applied, which was repeated five times using an RF classifier. The results are shown in Table II.

Figures 7 and 8 show the accuracy of all iterations, with the parameters used in each iteration. Grid search was relied upon to choose the appropriate parameters each time. Each iteration represents a five-fold cross-validation. Figure 7 shows the results for the visual stimulation data, while Figure 8 shows the results for the auditory stimulation data, both Figures are screenshots of the execution results obtained from MATLAB.

Results Table:

RepeatNumber	BestAccuracy	BestNumTree	BestMaxDepth
1	95	100	15
2	92.5	100	10
3	92.143	100	10
4	91.786	100	10
5	94.643	100	15

Best Repeat Index: 1
Best Overall Accuracy: 95%
Best NumTree: 100
Best MaxDepth: 15

Figure 7. RF Classifier Results for SSVEP with Repeated cross-validation 5 times.

The second method for data representation, the 15 rows

Results Table:

RepeatNumber	BestAccuracy	BestNumTree	BestMaxDepth
1	97.5	100	15
2	97.143	100	10
3	97.5	100	5
4	97.5	300	15
5	97.5	100	15

Best Repeat Index: 1
Best Overall Accuracy: 97.5%
Best NumTree: 100
Best MaxDepth: 15

Figure 8. RF Classifier Results for SSAEP with Repeated cross-validation 5 times.

for each participant were converted into a single vector to form a new dataset and then divided into 75% training and 25% testing. SVC, KNN, and LDA classifiers were applied, and the results for the visual and auditory conditions are shown in Figures 9 and 10, respectively:

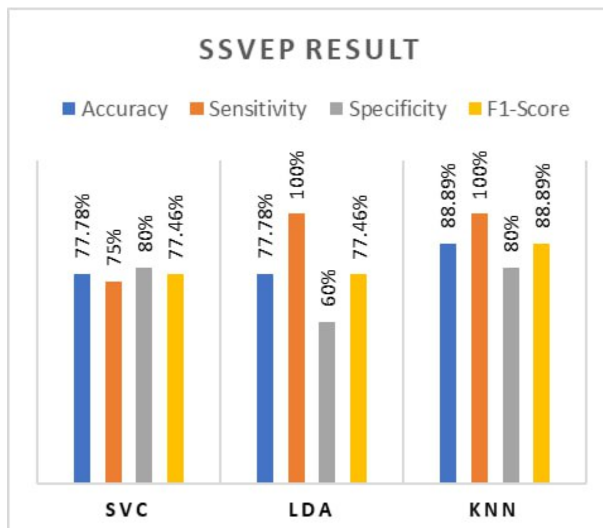

Figure 9. Classification Results for SSVEP Represented as Vector

The KNN classifier showed good performance for both visual and auditory stimuli.

A five-fold cross-validation technique was then applied to comprehensively evaluate the data and monitor the results on the unseen data, which was repeated five times using the SVC classifier. The results are shown in Table III.

The accuracy of all iterations is shown in Figures 11 and 12, where each iteration represents a five-fold cross-validation. Figure 11 shows the accuracy histogram of each iteration applied using SVC on the visual stimulus data, while Figure 12 shows the accuracy histogram of each iteration applied using SVC on the auditory stimulus data.

To evaluate the results of this study, the proposed model

TABLE II. Metrics for RF using 5-fold cross-validation repeated 5 times

The cases	The metrics	Average metrics across repetitions	Best results in the repetitions
SSVEP	Accuracy	93.21	95
	Sensitivity	93.66	95
	Specificity	90	90
	F1-Score	91.54	92.77
SSAEP	Accuracy	97.42	97.5
	Sensitivity	92.33	95
	Specificity	100	100
	F1-Score	95.54	97.14

TABLE III. Metrics for SVC using 5-fold cross-validation repeated 5 times

The cases	The metrics	Average metrics across repetitions	Best results in the repetitions
SSVEP	Accuracy	71.85	76.07
	Sensitivity	69.33	80
	Specificity	74	75
	F1-Score	70.74	74.04
SSAEP	Accuracy	73.07	81.07
	Sensitivity	70	81.66
	Specificity	76	80
	F1-Score	70.20	80.57

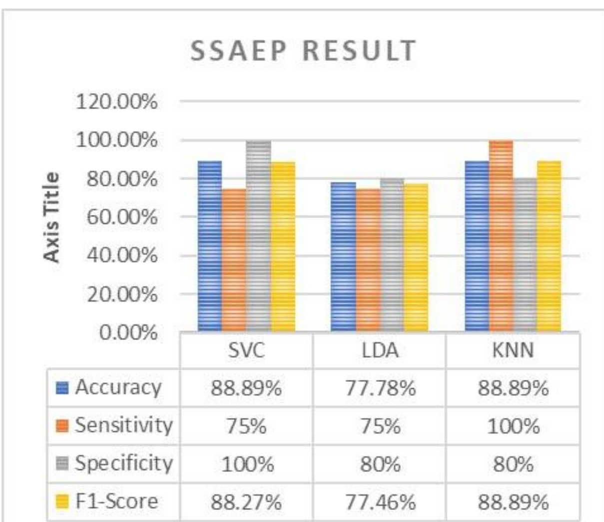

Figure 10. Classification Results for SSAEP Represented as Vector

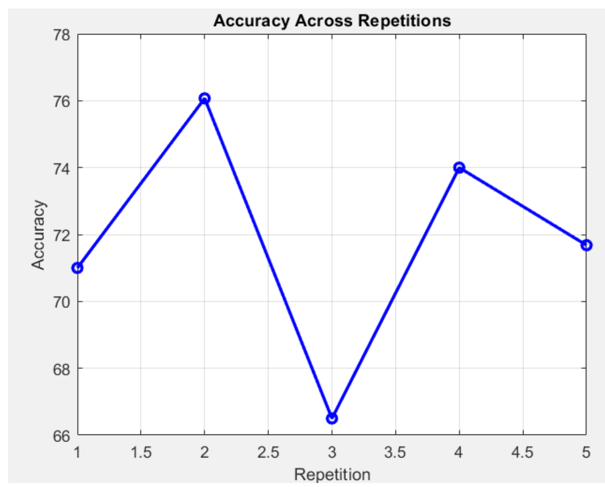

Figure 11. SVC Classifier Accuracy for SSVEP with Repeated cross-validation 5 times.

will be compared with related studies. The comparison results will be presented in Table IV, which shows the number of channels used in each study, the method of extracting important features from the EEG signals, and the conditions under which these signals were recorded. In addition to presenting the strengths, weaknesses, and the degree of accuracy achieved by each study.

Thus, we can observe the convergence of the results between the proposed model, which relied on the DWT method for feature extraction, and the studies of Aslan and Firat Orhanbuluku, despite using a smaller number of channels. Moreover, the proposed model outperformed Subasi's research, which also used the DWT method but applied the Daubechies 4 (db4) wavelet to extract features and relied on only two channels. This highlights the significance of the channels used in this study and the importance of using (db8) along with FFT in selecting the essential features for each channel.

Thus, this study stands out from other studies by relying

TABLE IV. A comparative study of literature research on migraine

Previous studies	Channel count	Feature extraction	number of participants	Status type	Highest accuracy	Strength point	Weaknesses
Kazemi and Katibeh (2018) [39]	19	Welch's method and the Yule-Walker AR method.	43(19 Controls,24 Patients)	SSVEP	93%	Application of the genetic algorithms to find the optimal combination of (features and electrodes) to maximize performance.	This study focuses only on pediatric EEG migraine classification.
Subasi et al. (2019) [6]	2	DWT (db4)	30 (15 Controls,15 patients)	SSVEP	85%	The study's conclusions are strengthened by the application of 10-fold cross-validation (CV) techniques, which guarantee the accuracy of the classification results.	Dependence only on flash stimulation and the use of specific EEG channels.
Aslan (2021) [13]	128	TQWT and statistical features	39 (21 controls,18 patients)	(Resting state)	89%	Using ensemble learning techniques, non-linear feature extraction, and Q-factor wave transformation technique.	The complexity of the calculations and the need to adjust parameters, which takes a long time.
Hanife Göker (2022) [12]	(Experiment1)14	Welch	39 (21 controls,18 patients)	(Resting state)	93.25%	Two experiments were performed with 14 channels and 128 channels, which allowed analysis of the effect of the number of channels on model performance.	The limited number of participants may affect the generalizability of the results.
	(Experiment2)128	Welch	39 (21 controls,18 patients)	(Resting state)	95%		
Aslan (2023) [40]	128	CWT	39 (21 controls,18 patients)	(Resting state)	100%	CNN has a high ability to extract features from large-dimensional data such as EEG signals, which helps improve diagnostic accuracy.	CNN is complex and requires high computational power and a large amount of data.

TABLE IV. (Continued) A comparative study of literature research on migraine

Previous studies	Channel count	Feature extraction	number of participants	Status type	Highest accuracy	Strength point	Weaknesses
Fırat Orhanbu-lucu et al. (2023) [41]	64	CWT	39 (21 controls,18 patients)	(Resting state) SSAEP SSVEP	99.74% 99.44% 98.96%	This study is very detailed as EEG signals recorded in states of rest, visual stimulation, and auditory stimulation were used. The study's DCNN model was created and its performance compared with other widely used CNN architectures.	It is highly uncommon to find research on migraine diagnosis and the use of EEG signals in deep-learning models. Consequently, there wasn't much research to compare the results of this study with.
This study	15	DWT (db8) and FFT	37 (20 controls,17 patients)	SSAEP SSVEP	100% 100%	Combining features extracted from db8 (which provides temporal and frequency information) and features extracted from FFT (which provides precise frequency information) results in a richer representation of the signal, which helps train the model better.	The small number of participants may restrict the results' generalizability.

on DWT (db8) combined with FFT to analyze EEG signals and extract important features for each channel.

B. *Result Of The Second Model*

The results of the RF classifier with five-fold cross-validation are shown in Figure 13. This classifier was used to diagnose and classify migraine types versus healthy subjects by combining original data with data generated under visual stimulation.

The results of the five folds of Random Forest classification, including accuracy, sensitivity, specificity, and F1-Score, are displayed in a line graph as shown in Figure 14. Only three lines are shown, instead of four, due to the overlap between sensitivity and F1-Score, which shows little variation in model performance across different folds and classification efficiency.

We observe the stability of results across folds after data augmentation. This indicates the power of the model itself and the method used to represent the data. This stability means that the model's performance remains similar across different datasets, which is a positive indicator of its reliability and effectiveness.

We found one paper by Alex Frid et al. (2019) [3] in distinguishing between migraine types (MwoA and MwA), which achieved an accuracy of 84.62%. However, we cannot compare results because patients and sample sizes vary widely.

4. Discussion

Migraine symptoms often overlap with other headache symptoms, and even among different types of migraine, many migraine patients suffer from sensitivity to light and loud sounds, which complicates the diagnosis of migraine

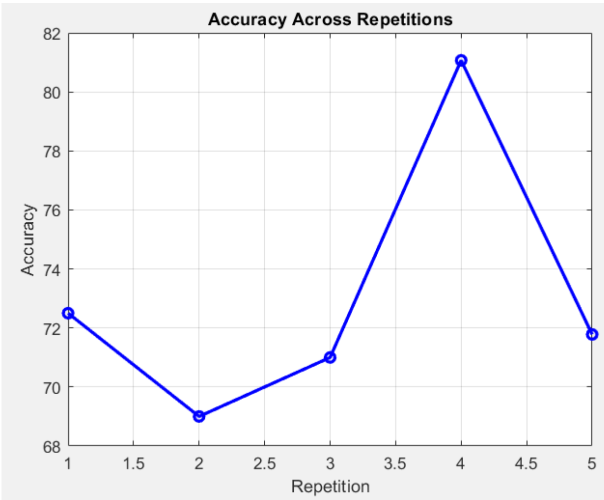

Figure 12. SVC Classifier Accuracy for SSAEP with Repeated cross-validation 5 times.

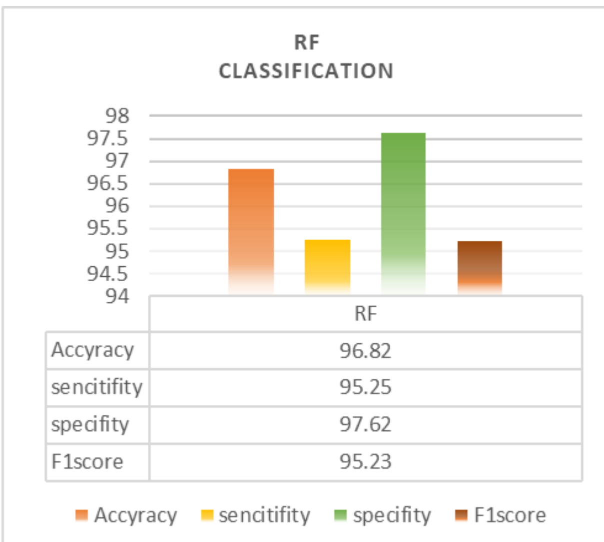

Figure 13. The result of the RF classifier.

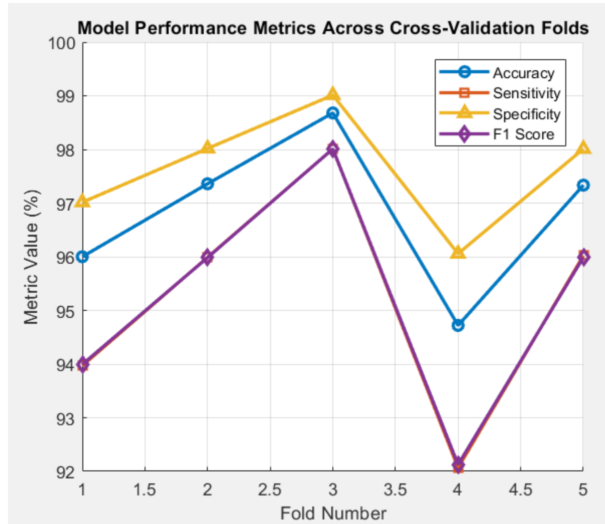

Figure 14. Five-Fold Cross-Validation Results for RF with Sum of Frequencies.

with or without aura, making accurate diagnosis difficult and time-consuming. This challenge is exacerbated when communicating with children or the elderly, as they may forget symptoms during medical inquiries. Therefore, there is no point in relying on traditional methods that are question-based, nor is it possible to rely on visual examination of brain signals, as they are complex and non-linear and require effort, experience, and time. To address these problems, we suggest relying on the models proposed in this study, as they automatically analyze these signals, simplifying the diagnosis process. It ensures appropriate and timely treatment.

Where EEG signals from visual and auditory stimulation were analyzed, focusing on specific channels (Fp1, F7, C3, Pz, Fp2, Fz, F8, Cz, C4, F3, F4, P3, P4, O1, O2). These signals underwent preprocessing to remove noise and environmental interference. They were then decomposed into frequency bands using the db8 wavelet transform. Afterward, FFT was applied to convert these extracted bands from the time domain to the frequency domain and identify the dominant frequency within each band Each participant's data includes multiple rows corresponding to different channels, necessitating dimensionality reduction for efficient classification. For the first model, two methods were used to represent data and reduce dimensions:

1) Frequency values for each band (delta, theta, alpha, beta, and gamma) were summed across all 15 channels corresponding to each participant. And

2) All 15 rows corresponding to each participant were converted into a single vector.

The method that summed the frequency values for each band (delta, theta, alpha, beta, and gamma) independently across all rows (channels) corresponding to each participant had great success, especially with the RF and SVC classifiers in diagnosing of migraine versus the healthy controls. The auditory stimulus data had the highest accuracy, followed by the visual.

On the other hand, representing the data as vectors resulted in good accuracy with the SVC classifier, especially for audio data when cross-validation was used. In contrast, the KNN classifier performed better with fixed partitioning for both datasets (SSVEP and SSAEP). However, the frequency sum method outperformed the vector representation method in auditory and visual stimulation conditions. It has proven to be the most effective, achieving much better results than the vector-based method.

In the second model, data for visual stimuli, previously cleaned and analyzed in the first model, were expanded due to their broader applicability compared to auditory stimuli. The method of summing frequencies for each band (delta, theta, alpha, beta, and gamma) across all channels corresponding to each participant was used, because it showed superior performance in the first model when applied with RF and SVC classifiers. RF was chosen as a classifier for this model because it relies on a triple classification to distinguish between types of migraines.

This makes using the SVC classifier impractical, as it is designed for binary classification and its performance deteriorates in triple classification scenarios. We employed five-fold cross-validation to test as many datasets as possible, and excellent results were obtained, confirming the success of this approach in identifying migraine types versus healthy controls.

The proposed study has some limitations. First, there are no publicly available EEG datasets on migraine other than those from Carnegie Mellon University, which prevented testing the study with different datasets. Another limitation is the scarcity of studies aiming to identify types of migraine based on EEG, making it difficult to test the proposed methodology in such a context.

5. Conclusions And Future Work

This study developed two electroencephalogram(EEG) –based models to diagnose migraines and classify their types using machine learning. The first model distinguishes migraine patients from healthy individuals, while the second classifies migraine with aura (MwA), migraine without aura (MwoA), and healthy controls. By utilizing 15 EEG channels, the study effectively reduced computational complexity while achieving reliable accuracy in diagnosing migraines and their types. The Butterworth filter efficiently removed irrelevant frequencies, and the db8 wavelet, combined with FFT, facilitated a detailed analysis of brainwave patterns across key bands (delta, theta, alpha, beta, gamma). Among the data representation techniques used, the frequency sum method showed superior accuracy compared to the vector method. Among the classifiers used, the RF classifier achieved the highest accuracy for both models while the SVC classifier particularly stood out in the first model for diagnosing migraine versus healthy controls because it was designed for binary classification. Moreover, comparing patients and healthy controls, an increase in beta and theta band values was observed, with a clear decrease in alpha band values in migraine patients, especially in the case of migraine accompanied by an aura.

In the future, the data set can be expanded to include other types of migraines and other types of neurological diseases can be included, such as tension and sinus headaches, as well as epilepsy, because their symptoms overlap with those of migraines, thus developing EEG-based care for this type of neurological disorder.

References

[1] A. Y. Yousif and A. Alsakaa, "Majority vote for electroencephalography (eeg)-based migraine classification," 2024.

[2] K. Mitrović, I. Petrušić, A. Radojičić, M. Daković, and A. Savić, "Migraine with aura detection and subtype classification using machine learning algorithms and morphometric magnetic resonance imaging data," *Frontiers in neurology*, vol. 14, p. 1106612, 2023.

[3] A. Frid, M. Shor, A. Shifrin, D. Yarnitsky, and Y. Granovsky, "A biomarker for discriminating between migraine with and without aura: Machine learning on functional connectivity on resting-state eegs," *Annals of Biomedical Engineering*, vol. 48, pp. 403–412, 2020.

[4] V. L. Feigin, E. Nichols, T. Alam, M. S. Bannick, E. Beghi, N. Blake, W. J. Culpepper, E. R. Dorsey, A. Elbaz, R. G. Ellenbogen *et al.*, "Global, regional, and national burden of neurological disorders, 1990–2016: a systematic analysis for the global burden of disease study 2016," *The Lancet Neurology*, vol. 18, no. 5, pp. 459–480, 2019.

[5] A. K. Eigenbrodt, H. Ashina, S. Khan, H.-C. Diener, D. D. Mitsikostas, A. J. Sinclair, P. Pozo-Rosich, P. Martelletti, A. Ducros, M. Lantéri-Minet *et al.*, "Diagnosis and management of migraine in ten steps," *Nature Reviews Neurology*, vol. 17, no. 8, pp. 501–514, 2021.

[6] A. Subasi, A. Ahmed, E. Aličković, and A. R. Hassan, "Effect of photic stimulation for migraine detection using random forest and discrete wavelet transform," *Biomedical signal processing and control*, vol. 49, pp. 231–239, 2019.

[7] L. Tong, "Evaluation of different brain imaging technologies," in *2021 International Conference on Public Art and Human Development (ICPAHD 2021)*. Atlantis Press, 2022, pp. 692–696.

[8] S. K. Khare, V. Bajaj, and G. R. Sinha, "Adaptive tunable q wavelet transform-based emotion identification," *IEEE transactions on instrumentation and measurement*, vol. 69, no. 12, pp. 9609–9617, 2020.

[9] L. Angrisani, P. Arpaia, D. Casinelli, and N. Moccaldi, "A single-channel ssvep-based instrument with off-the-shelf components for trainingless brain-computer interfaces," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 10, pp. 3616–3625, 2018.

[10] J. Vandana and N. Nirali, "A review of eeg signal analysis for diagnosis of neurological disorders using machine learning," *Journal of Biomedical Photonics & Engineering*, vol. 7, no. 4, p. 40201, 2021.

[11] K. Jindal, R. Upadhyay, H. Singh, M. Vijay, A. Sharma, K. Gupta, J. Gupta, and A. Dube, "Migraine disease diagnosis from eeg signals using non-linear feature extraction technique," in *2018 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*. IEEE, 2018, pp. 1–4.

[12] H. Göker, "Automatic detection of migraine disease from eeg signals using bidirectional long-short term memory deep learning model," *Signal, Image and Video Processing*, vol. 17, no. 4, pp. 1255–1263, 2023.

[13] Z. Aslan, "Migraine detection from eeg signals using tunable q-factor wavelet transform and ensemble learning techniques," *Physi-*

cal and Engineering Sciences in Medicine, vol. 44, no. 4, pp. 1201–1212, 2021.

[14] M.-P. Hosseini, A. Hosseini, and K. Ahi, “A review on machine learning for eeg signal processing in bioengineering,” *IEEE reviews in biomedical engineering*, vol. 14, pp. 204–218, 2020.

[15] M.-P. Hosseini, D. Pompili, K. Elisevich, and H. Soltanian-Zadeh, “Random ensemble learning for eeg classification,” *Artificial intelligence in medicine*, vol. 84, pp. 146–158, 2018.

[16] A. A. Al-Sakaa, Z. H. Nasralla, M. H. Hussein, S. A. Abd, H. Alsaqaa, K. Nermend, and A. Borawska, “The distinction of logical decision according to the model of the analysis of brain signals (eeg),” *Karbala International Journal of Modern Science*, vol. 8, no. 3, pp. 438–445, 2022.

[17] A. Borawska and A. Alsakaa, “Prediction of decision outcome via observation of brain activity signals during decision-making process,” in *Neuroeconomic and Behavioral Aspects of Decision Making: Proceedings of the 2016 Computational Methods in Experimental Economics (CMEE) Conference*. Springer, 2017, pp. 255–269.

[18] F. Orhanbulucu, F. Latifoğlu, and R. Baydemir, “A new hybrid approach based on time frequency images and deep learning methods for diagnosis of migraine disease and investigation of stimulus effect,” *Diagnostics*, vol. 13, no. 11, p. 1887, 2023.

[19] A. Chamanzar, S. M. Haigh, P. Grover, and M. Behrmann, “Abnormalities in cortical pattern of coherence in migraine detected using ultra high-density eeg,” *Brain Communications*, vol. 3, no. 2, p. fcab061, 2021.

[20] A. Caliskan and S. Rencuzogullari, “Transfer learning to detect neonatal seizure from electroencephalography signals,” *Neural Computing and Applications*, vol. 33, no. 18, pp. 12 087–12 101, 2021.

[21] M. Arumugam and A. K. Sangaiah, “Arrhythmia identification and classification using wavelet centered methodology in ecg signals,” *Concurrency and computation: practice and experience*, vol. 32, no. 17, p. e5553, 2020.

[22] F. A. Alturki, K. AlSharabi, A. M. Abdurraqeeb, and M. Aljalal, “Eeg signal analysis for diagnosing neurological disorders using discrete wavelet transform and intelligent techniques,” *Sensors*, vol. 20, no. 9, p. 2505, 2020.

[23] M. S. Nafea and Z. H. Ismail, “Supervised machine learning and deep learning techniques for epileptic seizure recognition using eeg signals—a systematic literature review,” *Bioengineering*, vol. 9, no. 12, p. 781, 2022.

[24] M. Dalal and M. Juneja, “Evaluation of orthogonal and biorthogonal wavelets for video steganography,” *Information Security Journal: A Global Perspective*, vol. 29, no. 1, pp. 40–50, 2020.

[25] M. Shouran and E. Elgamli, “Design and implementation of butterworth filter,” *International Journal of Innovative Research in Science Engineering and Technology*, vol. 9, no. 9, pp. 7975–7983, 2020.

[26] M. M. Naqiuddin, M. S. Leong, L. Hee, and M. Azrieasrie, “Ultrasonic signal processing techniques for pipeline: A review,” in *MATEC Web of Conferences*, vol. 255. EDP Sciences, 2019, p. 06006.

[27] N. R. Hatwar, U. G. Gawande, C. B. Thaokar, and R. F. Hatwar, “Identification of appropriate channels and feature types that differentiate the normal and stress data of eeg signals,” *Int. J. Intell. Syst. Appl. Eng*, vol. 11, no. 11, pp. 102–120, 2023.

[28] B. Lechat, K. L. Hansen, Y. A. Melaku, A. Vakulin, G. Micic, R. J. Adams, S. Appleton, D. J. Eckert, P. Catcheside, and B. Zajamsek, “A novel electroencephalogram-derived measure of disrupted delta wave activity during sleep predicts all-cause mortality risk,” *Annals of the American Thoracic Society*, vol. 19, no. 4, pp. 649–658, 2022.

[29] I. Hussain, M. A. Hossain, R. Jany, M. A. Bari, M. Uddin, A. R. M. Kamal, Y. Ku, and J.-S. Kim, “Quantitative evaluation of eeg-biomarkers for prediction of sleep stages,” *Sensors*, vol. 22, no. 8, p. 3079, 2022.

[30] A. J. Anderson, S. Perone, and M. A. Gartstein, “Context matters: Cortical rhythms in infants across baseline and play,” *Infant Behavior and Development*, vol. 66, p. 101665, 2022.

[31] D. M. Mateos, G. Krumm, V. Arán Filippetti, and M. Gutierrez, “Power spectrum and connectivity analysis in eeg recording during attention and creativity performance in children,” *NeuroSci*, vol. 3, no. 2, pp. 347–365, 2022.

[32] V. Rakhshan, P. Hassani-Abharian, M. Joghataei, M. Nasehi, and R. Khosrowabadi, “Effects of the alpha, beta, and gamma binaural beat brain stimulation and short-term training on simultaneously assessed visuospatial and verbal working memories, signal detection measures, response times, and intrasubject response time variabilities: A within-subject randomized placebo-controlled clinical trial,” *BioMed Research International*, vol. 2022, no. 1, p. 8588272, 2022.

[33] H. T. Darch, N. L. Cerminara, I. D. Gilchrist, and R. Apps, “Pre-movement changes in sensorimotor beta oscillations predict motor adaptation drive,” *Scientific Reports*, vol. 10, no. 1, p. 17946, 2020.

[34] M. S. Al-Quraishi, I. Elamvazuthi, S. A. Daud, S. Parasuraman, and A. Borboni, “Eeg-based control for upper and lower limb exoskeletons and prostheses: A systematic review,” *Sensors*, vol. 18, no. 10, p. 3342, 2018.

[35] J. L. Ulloa, “The control of movements via motor gamma oscillations,” *Frontiers in human neuroscience*, vol. 15, p. 787157, 2022.

[36] E. Y. Boateng, J. Otoo, and D. A. Abaye, “Basic tenets of classification algorithms k-nearest-neighbor, support vector machine, random forest and neural network: a review,,” *Journal of Data Analysis and Information Processing*, vol. 8, no. 4, pp. 341–357, 2020.

[37] J. Gou, L. Sun, L. Du, H. Ma, T. Xiong, W. Ou, and Y. Zhan, “A representation coefficient-based k-nearest centroid neighbor classifier,,” *Expert Systems with Applications*, vol. 194, 2022.

[38] Priyanka and D. Kumar, “Decision tree classifier: a detailed survey,,” *International Journal of Information and Decision Sciences*, vol. 12, no. 3, pp. 246–269, 2020.

[39] S. Kazemi and P. Katibeh, “Comparison of parametric and non-parametric eeg feature extraction methods in detection of pediatric migraine without aura,,” *Journal of Biomedical Physics and Engineering*, 2018.

[40] Z. Aslan, “Deep convolutional neural network-based framework in the automatic diagnosis of migraine,,” *Circuits, Systems, and Signal Processing*, vol. 42, no. 5, pp. 3054–3071, 2023.

[41] F. Orhanbulucu, F. Latifoğlu, and R. Baydemir, "A new hybrid approach based on time-frequency images and deep learning methods for diagnosis of migraine disease and investigation of stimulus effect,," *Diagnostics*, vol. 13, no. 11, May 2023.