CERTAIN INVESTIGATIONS ON FEATURE SELECTION TECHNIQUE USING ARTIFICIAL IMMUNE SYSTEMS FOR EEG COLOR VISUALIZATION CLASSIFICATION

K. Saranya[*], *P.M. Pandiyan*[†], *C.R. Hema*[‡]

Abstract: This paper aims to extract and select the significant features of electroencephalogram (EEG) signals and classify the visual stimulation of distinct colors. In this work, a novel method for selecting distinct colors using EEG signals called affinity artificial immune and Daubechies wavelet time-based learning (AAI-DWTL) is proposed. Initially, the EEG signals were collected in a controlled environment and an in-built band-pass filter was applied to remove the artifacts. The filtered signals were converted into frequency domain signals using least square-based short-term Fourier transform. After that, by utilizing Daubechies wavelet statistical time-based feature extraction model the time domain features were extracted. Followed by, computationally efficient features were selected using an affinity artificial immune-based feature selection model. The selected features were classified using a polynomial kernel multiclass classification-based machine learning algorithm and achieved an accuracy of 97.5% when compared with other methods like linear discriminant analysis (LDA) which obtained only 92%. Furthermore, while utilizing the proposed method classification time was considerably less when compared to LDA. The experimental result shows that the proposed color stimulation of the EEG signals method achieved greater improvement in terms of both classification time and classification accuracy with a minimum false positive rate.

Key words: *electroencephalogram, least square short term Fourier transform, affinity artificial immune, Daubechies wavelet, color visualization classification, polynomial kernel*

Received: August 25, 2023 **DOI:** 10.14311/NNW.2024.34.004
Revised and accepted: February 22, 2024

[*]Kumar Saranya – Corresponding author; Department of Computer Science and Engineering, Sri Ramakrishna Institute of Technology, Pachapalayam, Coimbatore, Tamil Nadu, 641010, India, E-mail: saranyak908798@gmail.com

[†]Paulraj Murugesa Pandiyan; Sri Ramakrishna Institute of Technology, Pachapalayam, Coimbatore, Tamil Nadu, 641010, India, E-mail: paulrajmurugesap98@gmail.com

[‡]C.R. Hema; Department of EEE, Sri Ramakrishna Institute of Technology, CBE-10 E-mail: hemacr345@gmail.com

1. Introduction

From the inception of human civilization, two significant elements intimately involved are color and perception of color. Survival like selecting safe foods, identifying optimal and secured routes, for changing emotional experiences to numerous stimuli, even in the modern world color reshapes the richness of complicated visual information. Nonetheless, the investigation of color perception and its influences on the human brain is interesting as well as significant as it fascinates both practical and theoretical areas of interest.

In order to analyze the brain response while visualizing the colors, non-linear techniques like multi fractal detrended fluctuation analysis and multi fractal detrended cross-correlation analysis (MFDFA and MFDXA) which are time domain features were used. By using these techniques, it has been inferred that an inter and intra-lobe correlation has occurred, but the accuracy was not taken into consideration [1]. [2] have inferred that decoding of colors using the EEG signals can be varied due to brightness and also the features of space [2].

[5] classified the primary colors of EEG signals using multiclass classification techniques along with a novel combination of forward feature selection mechanisms. The dimensionality of the data is reduced using this feature selection technique and the overall result was enhanced effectively resulting in a classification accuracy of 80.6%. [7] performed an Ishihara test to find the response of event-related potentials that emerged from the acquainted EEG signals and inferred that there is variation in response between a partial color-blind person and a normal person.

Motivated by the above facts in this paper, the objective remains to investigate the most effective features of stimulated EEG signals for various color perceptions using a novel feature selection and extraction method. Initially, the preprocessed signals undergo Daubechies wavelet statistical time-based feature extraction algorithm for extracting time domain accurate features followed by a feature selection algorithm namely affinity artificial immune system to select the relevant features that have not been reported ever before.

The contributions of this paper are to investigate Daubechies wavelet statistical time-based feature extraction algorithm for extracting time domain accurate features and to propose affinity artificial immune-based feature selection algorithm for selecting significant features for reducing computational complexity and a polynomial kernel multiclass classification-based machine learning algorithm for classifying the EEG signals for various color stimulation.

2. Related Works

A comprehensive review of stimuli presentation was investigated by [8] which formed a paramount part of any emotion elicitation experiment in effect analysis. Also due to the dearth of excellent recommendations, the researchers utilized their own designed methods that were also not found to be information, therefore not in keeping with ambiguous results. Therefore, a detailed study of the stimuli with an in-depth report was also made. Moreover, the influence of noise on the recognition aspect was also covered by [9].

Certain EEG pieces have been associated with the discernment of temporal magnitudes as a whole. However, to date, there does not remain a clear-cut idea regarding the role of time perception or magnitude perception playing a major role as far as EEG components are concerned. [10] recorded the EEG signals whereas the subject matters were made to make judgments about time or magnitude in a comparative manner, therefore forming mechanics between time and number.

[21] analyzed the effect of various color stimuli to measure signal-to-noise ratio and identified that average frequency provides a good signal to noise ratio than that of low-frequency signals. Yet the ability of color perception for primary colors namely RGB was made utilizing a Type 2 fuzzy classifier and provides interesting application for color identification and better classification accuracy was seen.

[22] designed an efficient method named reflection co-efficient for feature extraction from EEG signals. This also reduces the computational complexities. A combination of reflection coefficient features with support vector machines classifier classifies mental tasks with good accuracy and less computational time.

[22] proposed an ensemble method combining principal component analysis (PCA) with T-statistics for extracting robust features were designed. With this spatial PCA, features were extracted that in turn minimized the signal dimensionality. However, the classification accuracy was not focused. To concentrate on this issue, both time and frequency features were employed by [23] EEG signal classification.

[23] proposed a novel feature extraction method for automatically classifying emotions. First, reliable deep features were obtained by employing five distinct deep convolutional neural networks. Secondly for pre-processing, the obtained signals utilized various techniques like wavelet transform (WT) and continuous wavelet transform (CWT) were used. Finally, for classifying into valence and arousal classes support vector machine method was used, therefore causing a great improvement in accuracy rate. Pritom, [24] employed yet another method called common spatial domain for converting the EEG signals into frequency domains followed by feature extraction and finally classification of EEG signals on time.

[25] utilized a non-linear feature extractor named Hamsi-Pat using substitution box (S-Box) for feature generation. By using this dimension reduction was achieved and in addition accuracy involved in the classification process was also found to be improved using k nearest neighborhood (KNN) classifier. [26] investigated on deep learning methods for the classification of alcoholic EEG signals based on principal components and ANN as classifier. [16] utilized a multi-scale CNN on EEG signals for the classification purpose, thus improving the classification accuracy to a greater extent.

[15] examines the use of EEG signals in emotion recognition, highlighting the advantages of EEG over other modalities due to its sensitivity and real-time response. (Huang et al) comprehensively reviews various EEG-based feature extraction, machine learning, and deep learning methods for emotion recognition, discussing relevant studies and suggesting future research directions in this field.

[17] propose a hybrid neural network and time-frequency analysis for cardiac abnormality detection in ECG signals, employing visualization techniques and feature reduction via PCA for efficient neural network-based classification. [18] addressing fetal diagnosis, utilize time- frequency features and an ensemble cost-sensitive sup-

port vector machine classifier based on cardiotocography signal transformations and integrating novel image descriptors to predict fetal outcomes amidst data complexity. [19] introduce a smart meter-based approach, employing time-frequency feature combinations and machine learning, specifically random forest and SVM, to infer household characteristics with improved accuracy from real Irish data. Concurrently, [20] focus on brain-computer interface applications, presenting a multi-scale fusion convolutional neural network based on the attention mechanism network for motor imagery EEG decoding. This innovative network integrates convolutional neural networks and attention mechanisms to extract multi-scale features from EEG signals, enhancing sensitivity and performance in classifying brain activities. Their approach, validated using the BCI Competition IV-2a dataset, demonstrates superior classification efficacy compared to conventional methods, thereby showcasing its potential for decoding EEG information flow.

Based on the above observation from various papers, a novel method called affinity artificial immune and Daubechies wavelet time-based learning (AAI-DWTL) for color classification is proposed. An elaborate description of the AAI-DWTL method is given in the following sections.

3. Experiment

3.1 Subject

The dataset comprises nine healthy male participants meticulously selected to ensure they are free from both color blindness and other health conditions that could potentially interfere with the experimental process. Detailed instructions were provided to each participant beforehand, emphasizing the importance of maintaining stillness during the EEG recording sessions to prevent any interference caused by muscle movements or other overt actions that might introduce artifacts to the data. Additionally, participants were advised to wash their hair to minimize high impedance, ensuring optimal signal quality during EEG signal acquisition. These measures were taken to guarantee the reliability and accuracy of the dataset collected for the experimental procedures.

3.2 Experimental Set-Up

EEG signals were captured using a G-Nautilus device with 7 channels. Three channels were selected namely Cz, O1, and O2 for color visualization tasks according to the 10-20 electrode positioning system. The EEG signals were captured using a G-Tec High-Speed base station and recorded in the MATLAB Simulink for further processing. The signals were captured in a controlled environment. Initially, the subjects were asked to be calm, and the base signal was captured for 5 seconds. As a total with a sampling rate of 250 Hz, the band-pass filter is less than 30 Hz, with electrode 1 being CZ, electrode 2 being O1, electrode 3 being O2, and ground set to C3 and left earlobe as reference. The impedance was kept below 5kohm while capturing the signals.

3.3 Experimental Protocol

The subjects were initially made relaxed and a blank screen is shown on an LCD screen of 15 inches. The distance between the subject and the screen was about 50 centimeter. The subjects were asked to visualizing a blank screen for about 10 sec and the interpretation was kept for baseline. Followed by seven colors namely, blue, brown, green, magenta, red, violet, and yellow were shown to each subject for about 5 seconds and the signals were recorded. The sequence was continued five times in five trials. Fig. 1 shows one sequence of the experimental protocol.

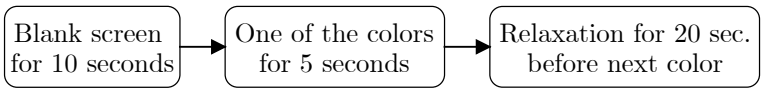

Fig. 1 *Experimental protocol for one sequence.*

4. EEG Signal Processing Methodology

Fig. 2 demonstrates the system design of brain machine interface. Initially, the signals are obtained using color visualization tasks as per the protocol. The signals are then preprocessed using short-term Fourier transform to remove noises and the essential features are extracted using Daubechies wavelet statistical time-based feature extraction model. The extracted features are then selected using the affinity artificial immune system technique which is finally classified using the polynomial kernel multiclass classification technique.

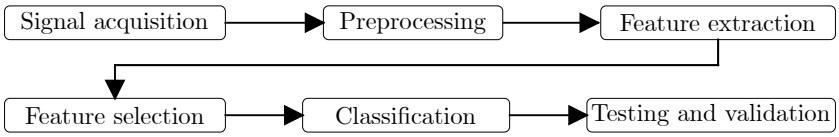

Fig. 2 *Block diagram for EEG signal processing.*

4.1 Preprocessing Technique

Preprocessing is crucial in the removal of artifacts like muscle and overt movements eyeblinks during signal recording and producing clean data. An in-built bandpass filter of < 30 Hz has been applied. The filtered signal undergoes a frequency domain shift using least squares-based short-term Fourier transform so that the local sections of the obtained filtered EEG signal can be obtained over time. Fig. 3 shows the structure of the least squares-based short-term Fourier transform for preprocessing.

As shown in Fig. 3, the procedure for computing least squares-based STFTs is to split the longer time signal into shorter segments possessing definite lengths and then evaluate the Fourier transform independently on each shorter segment. The obtained filtered EEG signal is multiplied by a window function $f(T)$ and the

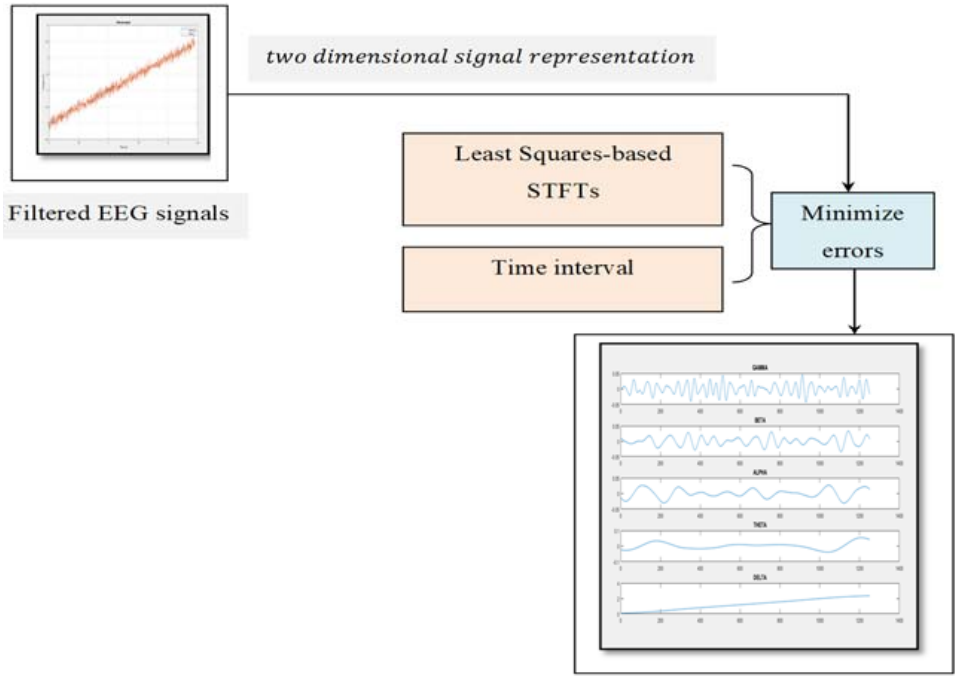

Fig. 3 *Structure of least squares-based short-term Fourier transform for preprocessing.*

Fourier transform of the resulting signal is slid along the time axis α, producing a two-dimensional signal representation and is mathematically stated as given below.

$$STFT\left\{a(T)\right\}(\alpha, f) \equiv A(\alpha, f) = \int_{-\infty}^{\infty} a(T) f(T-\alpha) e^{-\mathrm{i}fT} \mathrm{d}T. \tag{1}$$

From Eq. (1), $f(T)$ forms the Gaussian window function, $a(T)$ represents the obtained filtered EEG signal to be converted to the frequency domain, $A(\alpha, f)$ forms the Fourier transform for time axis T, frequency axis f respectively. Instead of just obtaining the products with sine and cosine waveforms, the least squares-based short term Fourier transform identified a time interval factor ϵ so that the pair would be fine-tuned for unequal powers at samples time T_i, to obtain a better estimate of power at a frequency. This is mathematically stated as given below.

$$\epsilon = \tan 2f\alpha = \frac{\sum_i \sin 2fT_i}{\sum_i \cos 2fT_i}. \tag{2}$$

In addition, the STFT is split in such a manner as to minimize the error φ where the matrix $\mathbf{M}$ of filtered EEG signals minimizes the sum of squared errors W as given below.

$$\varphi = \mathbf{M}W. \tag{3}$$

5. Feature Extraction Method

The preprocessed signals are split into five distinct frequency bands (i.e., alpha, beta, delta, theta, and gamma). To reduce the computational complexity and time complexity Daubechies wavelet statistical time-based feature extraction of EEG signals is utilized. It used to extract relevant information and features from signals, particularly in the context of time-domain analysis. This approach combines the power of wavelet transforms and statistical analysis to capture essential characteristics of a signal's behavior over time.

As shown in Fig. 4 Daubechies wavelet statistical time-based feature extraction model, with the selected features and five distinct bands as input, statistical time-based functions are applied to obtain accurate features (Wicaksono A, 2020) and resulting sub-signals within each frequency band. These statistical measures (FS) capture important characteristics of the signal's behavior in the time domain. Common statistical measures include mean, average energy, and standard deviation.

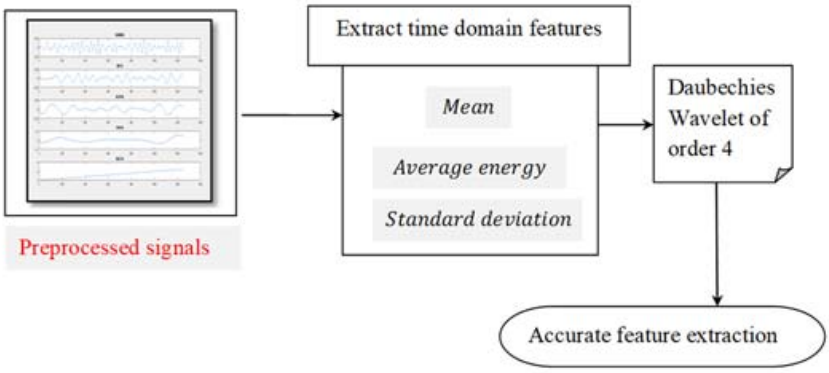

Fig. 4 *Daubechies wavelet statistical time-based feature extraction.*

First, the mean value of each sub-band signal for the corresponding selected features F_i is measured as given below.

$$\mu_i = \frac{1}{n}\sum_{j=1}^{n} FS(\varphi_{ij}) \tag{4}$$

Second, the average energy value of each sub-band signal (φ_{ij}) for the corresponding features selected is mathematically formulated as given below.

$$E_i = \frac{1}{n}\sum_{j=1}^{n} FS(\varphi_{ij})^2 \tag{5}$$

Finally, the standard deviation value of each sub-band signal for the corresponding features selected is mathematically stated as given below.

$$\sigma = \sqrt{\frac{1}{n}\sum_{j=1}^{n} FS[\varphi_{ij}^2]} \tag{6}$$

Each raw EEG signal is decomposed into four details utilizing Daubechies wavelet of order 4 (i.e., with level (n) 4) mathematically formulated as given below.

$$\begin{aligned} \mathrm{DW4}[F_i] \quad = \quad & \frac{1+\sqrt{3}}{4\sqrt{2}}\,[FS(\varphi_{00})] + \frac{3+\sqrt{3}}{4\sqrt{2}}\,[FS(\varphi_{01})] \\ & + \frac{3-\sqrt{3}}{4\sqrt{2}}\,[FS(\varphi_{10})] + \frac{1-\sqrt{3}}{4\sqrt{2}}\,[FS(\varphi_{11})] \end{aligned} \tag{7}$$

From the above resultant features extracted, the definition of each element is shown in Tab. I.

	Frequency				
Feature	Freq_1	Freq_2	Freq_3	Freq_4	Freq_5
Mean	F_1	F_4	F_7	F_{10}	F_{13}
Average energy	F_1	F_5	F_8	F_{11}	F_{14}
Standard deviation	F_3	F_6	F_9	F_{12}	F_{15}

Tab. I *Feature vector element definition.*

This extraction process where particularly useful when dealing with signals that exhibit dynamic behavior over time, where changes in amplitude and frequency components are crucial to characterizing the different patterns in the signal. As given in Tab. I, three features (i.e., mean, average energy and standard deviation) are extracted from each sub-band signal. With this, a resultant feature vector (i.e., features extracted 'FE') with 15 elements is generated.

6. Feature Selection

The feature selection is considered a significant step in the EEG signal analysis [11]. To reduce the computational complexity with high dimensional data, feature extracted data undergoes this process. Feature selection is considered a binary factor and therefore, a value of 1 for each feature denotes that the feature is included and hence is selected. On the other hand, a value of 0 for each feature denotes that the feature is excluded and hence is eliminated. Fig. 5 shows the structure of the affinity artificial immune-based feature selection model [9, 10].

As shown in Fig. 5, in the affinity-based artificial immune model, two essential elements are antigens and antibodies. Here, antigens denote the problem to be solved (i.e., features to be selected) and constraints (various frequency bands). On the other hand, the antibodies represent the candidate solutions. Let us consider an arbitrarily selected antibody population (PS) to be $AP = PS = \{PS_i\}, PS_i \in$

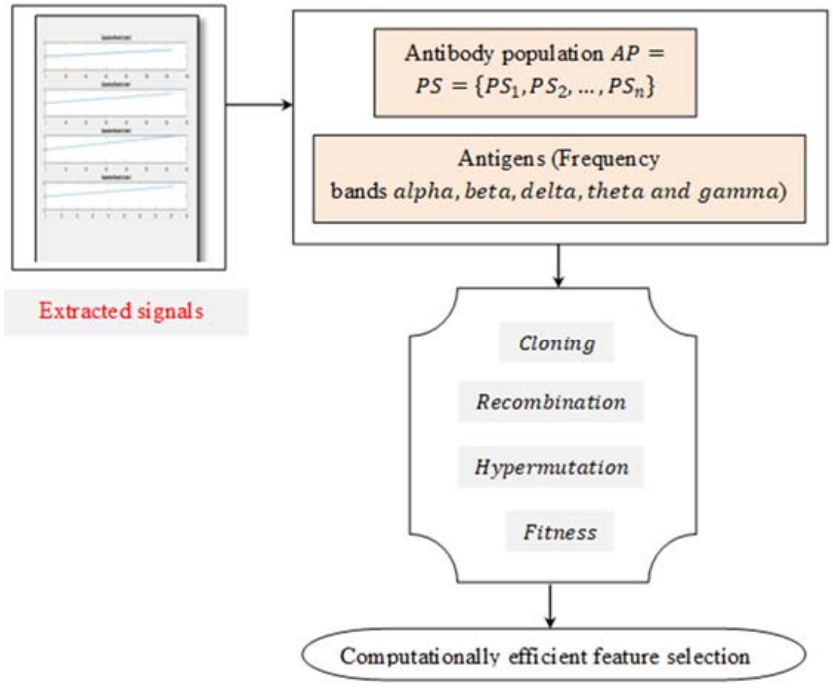

Fig. 5 *Structure of affinity artificial immune-based feature selection model.*

$\Omega, i = 1, 2, \ldots, n$ with n being an m-dimensional group of antibodies with Ω representing the feasible solution and n is the antibody population size. The affinity-based artificial immune model in our work comprises population definition (i.e., population size = 100), cloning, recombination (0.8), hyper-mutation (0.2), and fitness assignment.

Let us consider cloning operator T^c to be modeled on the processed signals $ps_1, ps_2, \ldots, ps_n$ mathematically stated as given below.

$$PS_1(k) = T^c(ps_1, ps_2, \ldots, ps_n) = T^c(ps_1) + T^c(ps_2) + \cdots + T^c(ps_n) \tag{8}$$

From the above Eq. (8), the processed signals output for the corresponding kth generation is modeled based on the cloning operator T^c. Next, let us consider $(c_1, c_2, \ldots, c_n)$ being the after-clone operator, then the recombination value T^R on the updated population is mathematically stated as given below.

$$PS_2(k) = T^R(c_1, c_2, \ldots, c_n) = T^R(c_1) + T^R(c_2) + \cdots + T^R(c_n) \tag{9}$$

From the above Eq. (9), the recombination resultant value $PS_2(k)$ is obtained based on the recombination value T^R on the updated population $T^R(c_1)+T^R(c_2)+\cdots+T^R(c_n)$. Next, the hyper-mutation resultant value after recombination is mathematically formulated as given below.

$$PS_3(k) = T^M(r_1, r_2, \ldots, r_n) = T^M(r_1) + T^M(r_2) + \cdots + T^M(r_n) \tag{10}$$

From the above Eq. (10), hyper-mutation resultant value $PS_3(k)$ is obtained based on the updated recombination processed signals $T^M(r_1) + T^c\ (r_2) + \cdots + T^c(r_n)$. Finally, with this hyper-mutation resultant value, the fitness function result is obtained using affinity measure as given below.

$$\begin{aligned} f &= Dis(Tr[PS_3(k)] - Ts[PS_3(k)]) \\ &= \sqrt{\left(Tr\left[PS_3(k)\right]_j - Ts\left[PS_3(k)\right]_j\right)^2} \end{aligned} \tag{11}$$

In the proposed work the features with an f value that lies between 0.5 and 1 were taken into account for potential feature. Generally, a greater f value indicates a better ability to extract features for differentiation between various colors. $Tr[PS_3(k)]$ and $Ts[PS_3(k)]$ are the transformed and target hyper-mutation resultant value.

7. Classification

Finally, in this section, the polynomial kernel multiclass classification-based machine learning model is applied to the features selected as it has been proven to be good in classification. Fig. 6 shows the structure of polynomial kernel multiclass classification-based machine learning.

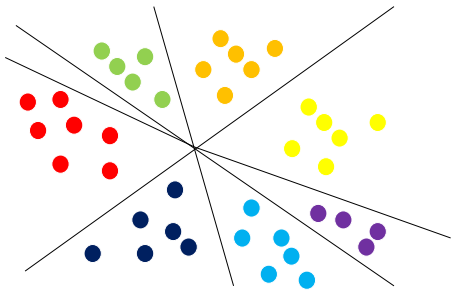

Fig. 6 *Structure of polynomial kernel multiclass classifier model.*

As illustrated in Fig. 6 the polynomial kernel multiclass classifier constructs a hyperplane in such a manner that separates data points belonging to two different classes and then further splits down the multiclass classification for different color stimulations of EEG signals into multiple binary classifications. The objective here remains in mapping the data points to gain mutual linear separation between every two classes. Also, the hyperplane is selected in such a manner to maximize the distance to the closest training data points of any class. To be more specific, the larger the margin is, the higher the classification accuracy and vice versa. To classify N, n-dimensional training data points (i.e., training processed feature selected signals) PS_i, $i = 1, 2, \ldots, N$, that either belong to class-1 or class-2 and the associated labels be $Y_i = (0, \ldots, 14)$ as red for class-1, $Y_i = (15, \ldots, 29)$ as orange for class-2, $Y_i = (30, \ldots, 44)$ as yellow for class-3, $Y_i = (45, \ldots, 59)$ as green for

class-4, $Y_i = (60, \ldots, 74)$ as blue for class-5, $Y_i = (75, \ldots, 89)$ as indigo for class-6 and $Y_i = (90, \ldots, 100)$ as violet for class-7 etc. A decision function is defined as $F(AV_i) = w^n\ AV_i + b$, where w denotes the m-dimensional vector and b represents the bias term respectively. This is mathematically stated as given below.

$$AV_i = [(PS)] \cdot [(FE)]\,. \tag{12}$$

From the above Eq. (12), the average value AV_i is arrived at by utilizing the processed EEG signals PS and the features extracted FE and these values are substituted to model the hyperplane as given below.

$$w^n AV_i + b = 0. \tag{13}$$

With the above hyperplane, the d-degree kernel polynomial is formulated as given below to obtain multiclass classified results.

$$(w^n AV_i + b)^d = 0. \tag{14}$$

With the significant features selected the maximum frequency (obtained from the frequency band) and d-degree kernel polynomial results were employed for detecting different colors. Due to this, the false positive rate involved in analyzing distinct colors were said to be reduced. With the significant features selected the maximum frequency (obtained from the frequency band) and d-degree kernel polynomial results were employed for detecting different colors. Due to this, the false positive rate involved in analyzing distinct colors were said to be reduced.

8. Results and Discussions

The validations are performed to color stimuli using EEG signals obtained from 9 different subjects. Most samples (70%) were used for training, and (30%) were taken for testing. The performance of the AAI-DWT, MFDFA and MFDXA, linear discriminant analysis was evaluated and compared using three parameters namely classification accuracy, classification time and false-positive rate respectively.

8.1 Classification Accuracy

The first parameter involved in the color stimulation of EEG signals is the accuracy involved in the classification process [12].In other words, classification accuracy (CA) refers to the percentage ratio of the number of correct color predictions made. This is mathematically stated as given below.

$$CA = \frac{CP}{TP} \cdot 100\% \tag{15}$$

Tab. II given below lists the classification accuracy obtained from Eq. (15) for three different methods namely AAI-DWT, MFDFA and MFDXA, and multifractal detrended fluctuation analysis (MFD) [1], linear discriminant analysis [2]. The simulations were conducted for subject 1 at different time instances (T_1 to T_5) and analyzed for green color.

Samples	Classification accuracy (%)		
	AAI-DWT	MFDFA and MFDXA	Linear discriminant analysis
500	97.50	95.10	92.50
1000	95.35	92.25	90.35
1500	93.15	90.00	87.15
2000	91.15	88.45	83.00
2500	90.00	86.00	81.15
3000	88.45	84.35	78.00
3500	85.35	81.00	75.25
4000	82.00	78.15	71.00
4500	81.15	75.00	68.00
5000	78.00	72.15	65.25

Tab. II *Tabulation for classification accuracy using AAI-DWT, MFDFA and MFDXA [1], linear discriminant analysis [2].*

Fig. 7 illustrates the classification results obtained using the three methods [2]. It is inferred that increasing the samples causes a significant amount of noise (i.e. while obtaining filtered EEG signals and this in turn results in a decrease in the classification accuracy. However, simulations performed with 5000 samples found an improvement of 97.5% classification accuracy using AAI-DWT, 95.1% classification accuracy using [1] and 92.5% classification accuracy using [2] respectively. The improvement in the classification accuracy using AAI-DWT was owing to the incorporation of Daubechies wavelet statistical time-based feature extraction algorithm. By applying this algorithm first, statistical time-based function, mean, average energy, and standard deviation were applied to the processed distinctly

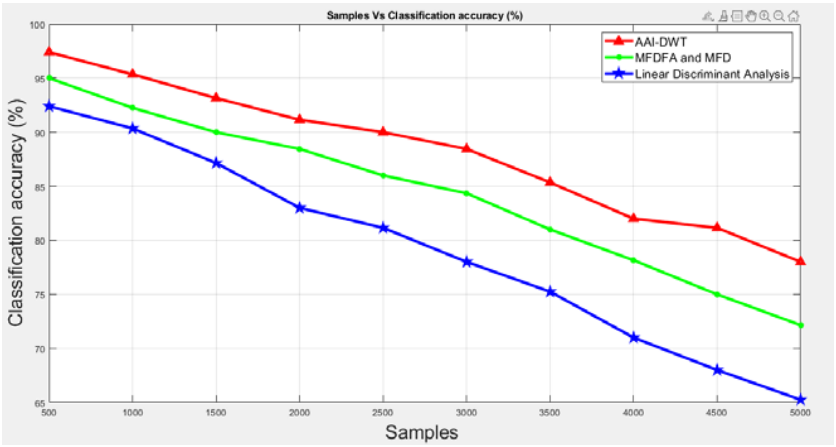

Fig. 7 *Analysis of classification accuracy.*

five band signals (i.e., alpha, beta, theta, delta, and gamma), and the proposed feature extraction technique was used. With this, the classification of EEG signals into distinct colors was made accurately therefore improvement was observed using AAI-DWT by 5% when compared to MFDFA and MFDXA and 12% when compared to LDA.

8.2 Classification Time

The second parameter of significance is the time taken for classification. It refers to the time involved in the classification of multiclass results. The equation is given below.

$$CT = \sum_{i=1}^{n} Samples_i \cdot Time\left[w^n AV_i + b\right] \tag{16}$$

From Eq. (16), the classification time CT refers to the time taken for the classification process in $Samples_i$ and the actual time involved in the classification process $Time[w^n AV_i + b]$. Milliseconds (ms) is the measure used for this parameter. Results of classification time is shown in Tab. III which is arrived from Eq. (16) for three different methods, AAI-DWT, MFDFA, and MFDXA [1], linear discriminant analysis [2]. The simulations were conducted for the subject Gowtham at five different time instances ranging between T_1 and T_5 with the in-depth analysis made for one color.

Samples	Classification time (ms)		
	AAI-DWT	MFDFA and MFDXA	Linear discriminant analysis
500	75	115	140
1000	93	123	151
1500	123	152	183
2000	141	194	224
2500	162	212	253
3000	183	223	294
3500	201	283	318
4000	225	312	353
4500	283	334	412
5000	302	382	443

Tab. III *Tabulation for classification time using AAI-DWT, MFDFA and MFDXA [1], linear discriminant analysis [2].*

Fig. 8 shows classification time using several samples obtained from subject 2. In the above figure, the horizontal axis refers to the number of sample signals involved, and the vertical axis represents the classification time measured in terms of milliseconds (ms). It is inferred that with the increase in the number of sample signals there is an increase in the number of candidate solutions and therefore

increase in estimating the recombination and hypermutation for differentiation between various colors. As a result, there is an increase in the classification time also. However, processing with 500 samples showed a segmentation time of 75 ms using the AAI-DWT method. 115 ms using MFDFA and MFDXA, 140 ms using LDA method respectively. From these results, it can be inferred that the classification time using the AAI-DWT method was found to be lesser than MFDFA and linear discriminant analysis. The affinity artificial immune-based feature selection algorithm contributed to this result. By applying this algorithm, an affinity measure was applied instead of a conventional artificial immune system, wherein the corresponding cloning, recombination, and hyper-mutation were performed for the corresponding signals. Finally, significant features were obtained via affined input measures substitution and fitness function to obtain significant features. After utilizing the proposed feature selection algorithm, the classification time has been considerably reduced by 31% compared to MFDFA and MFDXA 47%compared to LDA respectively.

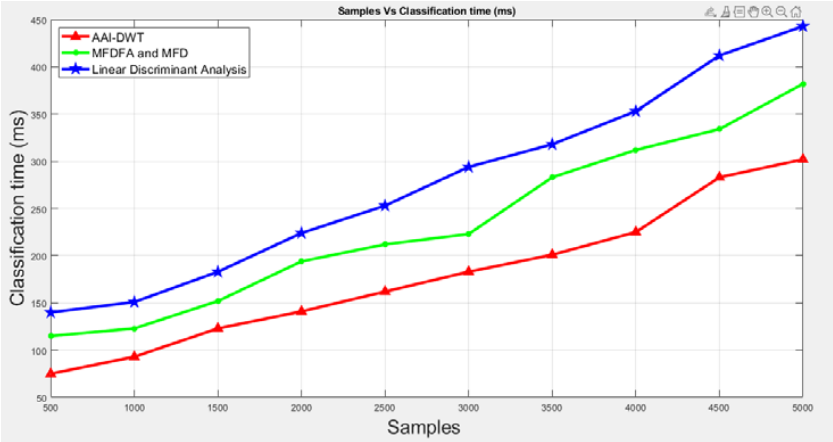

Fig. 8 *Analysis of various method using classification time.*

8.3 False Positive Rate

False positive rate is a measure used for differentiating the number of false positives or red color stimulus wrongly categorized as red color and the number of actual color stimulus that does not belong to red.

$$FPR = \frac{FP}{FP + TN} \tag{17}$$

From the above Eq. (17), the false positive rate *FPR* results are obtained, and true negative rate *TN*. Tab. IV given below gives the false positive rate results arrived based on the results from Eq. (17) for three different methods namely AAI-DWT, MFDFA and MFDXA, linear discriminant analysis.

Finally, Fig. 9 given above represents the false positive rate for differing samples. In the above figure, the x-axis represents the samples, and the y-axis denotes

Samples	False positive rate (%)		
	AAI-DWT	MFDFA and MFDXA	Linear discriminant analysis
500	7.00	10.00	13.00
1000	7.35	10.25	13.15
1500	7.55	10.45	13.35
2000	8.00	10.85	13.85
2500	8.15	11.00	14.00
3000	8.25	11.35	14.15
3500	8.40	11.55	14.25
4000	8.85	11.80	15.00
4500	9.00	12.00	15.35
5000	9.35	12.35	16.00

Tab. IV *Tabulation for false positive rate using AAI-DWT, MFDFA and MFDXA [1], linear discriminant analysis [2].*

its corresponding false positive rate. Increasing the samples causes an increase in the false positive rate also. This is because of the reason that with the increase in the samples provided as input, the color analysis may also pose a great challenge to the user in identifying the color and this in turn results in a significant amount of falsification of color analysis also. However, simulations conducted with 500 samples observed a false positive rate of 7% when applied with AAI-DWT, 10% using MFDFA and MFDXA, and 13% using LDA respectively. With this analysis, the false positive rate of the AAI-DWT was found to be comparatively lesser than the methods discussed in [1] and [2]. polynomial kernel multiclass classification-based machine learning algorithm contributed to this improvement. By applying this

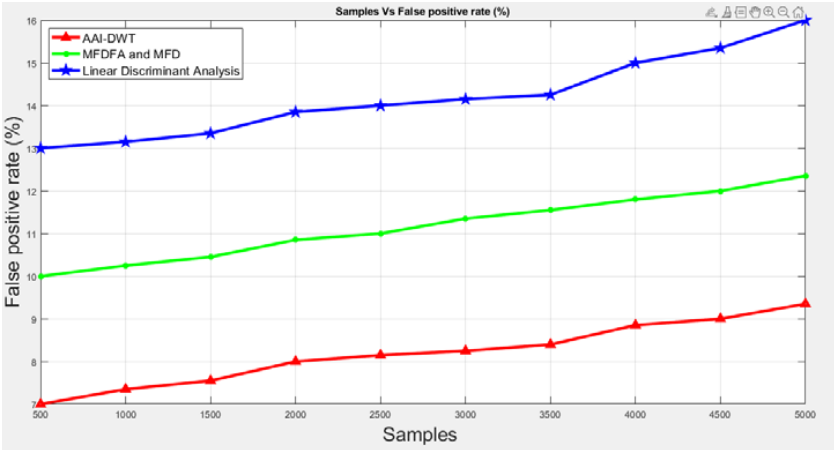

Fig. 9 *False positive rate.*

algorithm, significant features extracted from maximum frequency and d-degree kernel polynomial results were employed for identifying distinct colors. Next, by employing multiclass classification via support vector machine, data points were mapped to gain mutual linear separation between every two classes. Moreover, the hyperplane was derived in such a way that the distance was maximized to the closest training data points of any class. Due to this, falsification of color analysis was found to be reduced and therefore minimizing the false positive rate involved in analyzing distinct colors by 27% compared to [1] and 42% compared to [2].

The confusion matrix (i.e., validating visual colors are predicted) for different color perceptions with 500 samples is shown in Tab. V.

Samples	B	O	G	I	V	R	Y	ACU (%)
Blue	75	2	0	2	0	0	1	98
Orange	0	28	0	0	2	0	0	95
Green	0	0	65	0	0	3	0	97
Indigo	0	0	0	232	0	0	0	100
Violet	1	1	1	0	26	0	1	98
Red	2	2	0	0	2	34	0	97
Yellow	0	0	2	2	0	0	16	95

Tab. V *Confusion matrix with 500 samples.*

A confusion matrix is a trendy measure employed to address classification issues. It is used for binary classification as well as for multiclass classification issues [16, 17]. Where, R – red, O – orange,, Y – yellow, G – green, B – blue, I – indigo, V – violet, and ACU – accuracy.

The results show that the AAI-DWT technique works well in EEG signal analysis, with good classification accuracy seen across a range of sample sizes. When compared to alternative methods such as MFDFA, MFDXA, and linear discriminant analysis, there is a significant improvement. Furthermore, the AAI-DWT method reduces classification time significantly, owing to the effective feature selection algorithm used. Furthermore, the false positive rate study demonstrates the AAI-DWT approach's superiority in color analysis error minimization, highlighting its potential for real-world applications in EEG signal processing. The study focused solely on the feature extraction procedure and subsequent classification processes; therefore optimization or machine learning algorithms were not substantially examined. The purpose was to demonstrate the effectiveness of the suggested method, AAI-DWTL, in feature selection and classification accuracy for EEG signal processing. While optimization and machine learning techniques are important in signal processing, the scope was deliberately limited to provide a thorough analysis of the suggested method's performance. However, the significance of optimization and machine learning techniques in improving classification results is recognized, and their potential incorporation will be examined in future research discussion.

9. Conclusion

This system utilizes a novel affinity artificial immune and Daubechies wavelet time-based learning (AAI-DWTL) method for color stimulation from EEG signals has been used. The AAI-DWT method combines Daubechies wavelet statistical time-based feature extraction using a statistical time-based function with affinity artificial immune-based feature selection model. The reason behind the improvement was that first to remove the noise by converting filtered EEG signals to the frequency domain employing least squares-based short-term Fourier transform algorithm for preprocessing model. Second, Daubechies wavelet statistical time-based feature extraction algorithm was applied to extract accurate features for color analysis. Third, affinity artificial immune-based feature selection algorithm was applied that selected significant features based on frequency in a significant manner using cloning, recombination, and hyper-mutation. Finally, the classification was made by employing the polynomial kernel multiclass classification method for nine different subjects and the result shows that the classification accuracy was 97.5% using the proposed method which outperforms the MFDFA and MFDXA and LDA method. In addition to that the false positive rate and the classification time of AAI-DWT are lesser when compared to the other aforementioned methods.

Declaration

Ethics Approval and Consent to Participate: No participation of humans takes place in this implementation process.

Human and Animal Rights: No violation of Human and Animal Rights is involved.

Funding: No funding is involved in this work.

Data availability statement: Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Conflict of Interest: Conflict of Interest is not applicable in this work.

Authorship contributions: There is no authorship contribution.

Acknowledgement

There is no acknowledgement involved in this work.

References

[1] ROY S., BANERJEE A., ROY C., NAG S., SANYAL S., SENGUPTA R., GHOSH D. Brain response to color stimuli: an EEG study with non-linear approach. Cognitive Neurodynamics, Springer. [Multi Fractal Detrended Fluctuation Analysis and Multi Fractal Detrended cross-correlation Analysis (MFDFA and MFDXA)], 2021, doi: 10.1007/s11571-021-09692-z.

[2] HAJONIDES J.E., NOBRE A.C., VAN EDE F., STOKES M.G. Decoding visual color from scalp electroencephalography measurements. Neuro-Image, Elsevier, [Linear Discriminant Analysis], 2021, doi: 10.1016/j.neuroimage.2021.118030.

[3] GHOSH R., DEB N., SENGUPTA K., PHUKAN A., CHOUDHURY N., KASHYAP S., PHADIKAR S., SAHA R., DAS P., SINHA N., DUTTA P. SAM 40: Dataset of 40 subject EEG recordings to monitor the induced-stress while performing Stroop color-word

test, arithmetic task, and mirror image recognition task. Data in Brief, Elsevier, 2022, doi: 10.1016/j.dib.2021.107772.

[4] YOTO A., KATSUURA T., IWANAGA K., SHIMOMURA Y. Effects of Object Color Stimuli on Human Brain Activities in Perception and Attention Referred to EEG Alpha Band Response. Journal of Physiological Anthropology, doi: 10.2114/jpa2.26.373.

[5] CHAUDHARY M., MUKHOPADHYAY S., LITOIU M., SERGIO L.E., ADAMS M.S. Understanding Brain Dynamics for Color Perception using Wearable EEG headband. CASCON'20, 2020, pp. 10–13, doi: 10.48550/arXiv.2008.07092.

[6] BEKHTEREVA V., MÜLLER M.M. Bringing color to emotion: The influence of color on attentional bias to briefly presented emotional images. Cognitive, Affective and Behavioral Neural Science, Springer, 2017, doi: 10.3758/s13415-017-0530-z.

[7] WICAKSONO A., MENGKO1 T., IRAMINA K. Investigation of EEG Signal Response Using Event-Related Potential (ERP) Towards Ishihara Pseudo-Isochromatic Visual Stimulus. IEEE Signal Processing in Medicine and Biology, 2020, doi: 10.1109/SPMB50085.2020.9353640.

[8] SARMA P., BARMA S. Review on Stimuli Presentation for AffectAnalysis Based on EEG. IEEE Access, 2020, doi: 10.1109/ACCESS.2020.2980893.

[9] WANG Q., YIBO L., XUEPING L. The influence of photo elements on EEG signal recognition. EURASIP Journal on Image and Video Processing, 2018, doi: 10.1186/s13640-018-0367-6.

[10] SCHLICHTING N., DE JONG R., VAN RIJN H. Performance-informed EEG analysis reveals mixed evidence for EEGsignatures unique to the processing of time. Psychological Research, Springer, 2018, doi: 10.1007/s00426-018-1039-y.

[11] KALAIVANI K., KSHIRSAGARR P.R., SIRISHA DEVI J., BANDELA S.R., COLAK I., NAGESWARA RAO J., RAJARAM A. Prediction of biomedical signals using deep learning techniques. Journal of Intelligent & Fuzzy Systems, (Preprint), 2023, pp. 1–14, doi: 10.3233/JIFS-230399.

[12] CHIRANJEEVI P., RAJARAM A. A lightweight deep learning model based recommender system by sentiment analysis. Journal of Intelligent & Fuzzy Systems, (Preprint), 2023, pp. 1–14, doi: 10.3233/JIFS-223871.

[13] ASHOK BABU P., MAZHER IQBAL J.L., SIVA PRIYANKA S., JITHENDER REDDY M., SUNIL KUMAR G., RAJARAM A. Power Control and Optimization for Power Loss Reduction Using Deep Learning in Microgrid Systems. Electric Power Components and Systems, pp. 1–14, 2023.

[14] SHEKHAR H., BHUSHANMAHATO C., SUMAN S., SINGH S., BHAGYALAKSHMI L., PRASAD SHARMA L., LAXMI KANTHA B., HELAN VIDHYA T., KUMAR AGRAHARAM S., RAJARAM A. Demand Side Control for Energy Saving in Renewable Energy Resources Using Deep Learning Optimization. Electric Power Components and Systems, pp. 1–17, 2023.

[15] ALMADHOR A., SAMPEDRO G.A., ABISADO M., ABBAS S. Efficient Feature-Selection-Based Stacking Model for Stress Detection Based on Chest Electrodermal Activity. Sensors, 2023, 23(15), 6664.

[16] HUANG M., ZHANG X., CHEN X., MAI Y., WU X., ZHAO J., FENG Q. Joint-channel-connectivity-based feature selection and classification on fNIRS for stress detection in decision-making. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2022, 30, pp. 1858–1869.

[17] SHAFI I., AZIZ A., DIN S., ASHRAF I. Reduced features set neural network approach based on high-resolution time-frequency images for cardiac abnormality detection. Computers in Biology and Medicine, 2022, 145, 105425.

[18] ZENG R., LU Y., LONG S., WANG C., BAI J. Cardiotocography signal abnormality classification using time-frequency features and Ensemble Cost-sensitive SVM classifier. Computers in Biology and Medicine, 2021, 130, 104218.

[19] YAN S., LI K., WANG F., GE X., LU X., MI Z., CHANG S. Time–frequency feature combination based household characteristic identification approach using smart meter data. IEEE Transactions on Industry Applications, 2020, 56(3), pp. 2251–2262.

[20] LI D., XU J., WANG J., FANG X., JI Y. A multi-scale fusion convolutional neural network based on attention mechanism for the visualization analysis of EEG signals decoding. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2020, 28(12), pp. 2615–2626.

[21] DUART X., QUILES E., SUAY F., CHIO N., GARCÍA E., MORANT F. Evaluating the effect of stimuli color and frequency on SSVEP. Sensors, 2020, 21(1), p. 117.

[22] RAHMAN M.M., FATTAH S.A. An efficient feature extraction scheme for classification of mental tasks based on inter-channel correlation in wavelet domain utilizing EEG signal. Biomedical Signal Processing and Control, 2020, 61, p. 102033.

[23] HARPALE V., BAIRAG V. An adaptive method for feature selection and extraction for classification of epileptic EEG signal in significant states. Journal of King Saud University-Computer and Information Sciences, 2021, 33(6), pp. 668–676.

[24] SAHA P.K., RAHMAN M.A., ALAM M.K., FERDOWSI A., MOLLAH M.N. Common spatial pattern in frequency domain for feature extraction and classification of multichannel EEG signals. SN Computer Science, 2021, 2, pp. 1–11.

[25] TUNCER T. A new stable nonlinear textural feature extraction method based EEG signal classification method using substitution Box of the Hamsi hash function: Hamsi pattern. Applied Acoustics, 2021, 172, p. 107607.

[26] FARSI L., SIULY S., KABIR E., WANG H. Classification of alcoholic EEG signals using a deep learning method. IEEE Sensors Journal, 2020, 21(3), pp. 3552–3560.