Received 20 January 2025, accepted 4 February 2025, date of publication 5 February 2025, date of current version 13 February 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3539502

RESEARCH ARTICLE

Fine-Tuning EEG Channel Utilization for Emotionally Stimulated Biometric Authentication

CHETAN RAKSHE[1], CHRISTY BOBBY THOMAS[2], MOHANAVELU KALATHE[3], VANTEEMAR S. SREERAJ[4], GANESAN VENKATASUBRAMANIAN[5], DEEPESH KUMAR[1], (Member, IEEE), A. AMALIN PRINCE[6], (Senior Member, IEEE), AND JAC FREDO AGASTINOSE RONICKOM[1]

[1]Computational Neuroscience and Biology Laboratory, School of Biomedical Engineering, Indian Institute of Technology (BHU), Varanasi 221005, India
[2]Department of Electronic and Communication Engineering, M. S. Ramaiah University of Applied Sciences, Bengaluru 560054, India
[3]Defence Bio-Engineering and Electro Medical Laboratory, DRDO, Ministry of Defence, Bangalore 560093, India
[4]National Institute of Mental Health and Neurosciences (NIMHANS), Bengaluru 560029, India
[5]Translational Psychiatry Laboratory, National Institute of Mental Health and Neurosciences (NIMHANS), Bengaluru 560029, India
[6]Department of Electrical and Electronics Engineering, BITS Pilani, K K Birla Goa Campus, Zuarinagar, Goa 403726, India

Corresponding author: A. Amalin Prince (amalinprince@goa.bits-pilani.ac.in)

This work was supported by the Life Sciences Research Board (LSRB), Defence Research and Development Organisation under Grant LSRB/01/15001/LSRB-401/LS&BD/2023.

ABSTRACT Biometric authentication relies on distinct biological traits of individuals to validate their identity, enhancing security measures. However, variations in an individual's emotional state can impact the reliability of the biometric system. In this study, we propose a novel pipeline to evaluate electroencephalography (EEG)-based biometric system across different emotional states and optimize critical brain regions using machine learning algorithms. EEG signals from the DEAP dataset were classified into four emotional states: HAHV, HALV, LALV, and LAHV. We extracted a comprehensive set of statistical, time, frequency, entropy, fractal, spectral, and shape features from each channel. Machine learning classifiers, including Random Forest, Gradient Boosting, Extreme Gradient Boosting, LightGBM, CatBoost, and Bagging, were used for participant authentication. Our results revealed that the CatBoost classifier performed well across all stimuli with average accuracies of 84%, 85%, 86%, and 83% for HAHV, HALV, LALV, and LAHV, respectively. We found that features from channels FC1, Fz, C4 & Pz, and FC1 significantly contributed to EEG authentication on stimuli such as HAHV, HALV, LALV, and LAHV, respectively. Features such as skewness and the theta-to-alpha frequency band ratio consistently performed well across stimuli, demonstrating EEG signals' potential for robust biometric authentication by addressing emotional variations.

INDEX TERMS Biometric, electroencephalography, emotions, brain location, machine learning.

I. INTRODUCTION

Biometric authentication encompasses various modalities that utilize distinctive biological traits for identity verification. There are primarily two categories of biometric systems: conventional and cognitive biometrics. Conventional biometrics use only physiological characteristics, which include fingerprints [1], iris patterns [2], facial characteristics [3], and voice recognition [4] for authentication. These methods, however, have their drawbacks, such as susceptibility to replication or forgery and potential issues with accuracy under various physical conditions or with changes over time [5].

The associate editor coordinating the review of this manuscript and approving it for publication was Vincenzo Conti.

On the other hand, cognitive biometrics uses physiological signals for user recognition to overcome the flaws in conventional biometrics, enhance operating efficiency, and secure critical resources [6], [7].

Cognitive biometrics can be developed using physiological signals such as blood volume pulse (BVP), electrodermal activity (EDA) [7], electrocardiography (ECG) [8], electromyography (EMG) [9], and electroencephalography (EEG) [10]. However, BVP is highly sensitive to physical conditions such as blood pressure and heart rate [11]. Environmental factors, such as temperature and humidity, influence EDA, potentially affecting the reliability of biometric authentication [12]. Additionally, ECG has a lower resolution for distinguishing between different cognitive tasks or states. EMG may depend on physical factors, including muscle fatigue, making them less reliable over time or under different physical conditions [13], [14]. In contrast, EEG-based biometric systems utilize complex brain signals that provide a high level of security, as these are extremely difficult to duplicate or falsify. In this study, we utilized an EEG-based biometric system owing to its distinctive advantages, capitalizing on its sophisticated security and user-aligned authentication.

The selection of elicitation protocols (EPs) or tasks plays a pivotal role in generating distinctive and repeatable EEG patterns for individual identification in EEG-based biometric authentication. It includes motor imagery [15], speech imagery [16], mental computation [17], working memory [17], auditory and visually evoked potential [18] and resting state [19]. It is imperative to acknowledge that the impact of different stimuli on individual emotional states varies significantly. Consequently, individuals with diverse personal emotional experiences may exhibit distinct responses when exposed to the same stimulus. It can be well reflected in the characteristics of EEG signals [20], [21]. It is important to understand the response of the brain to different stimuli, leading to distinct physiological states among individuals based on their personal experiences. Further, incorporating emotional stimuli into EEG-based biometric systems introduces a layer of complexity and uniqueness to the dataset, enhancing the system's resilience against fraudulent attempts at imitation or disguise [22], especially valuable in scenarios requiring the detection of stress or duress during authentication processes [23]. In this study, we analyzed the influence of different emotions in EEG-based biometrics authentication.

EEG signals are captured from various channels, but some of these may not contribute significantly, adding to computational complexity and diminishing authentication accuracy, particularly in datasets with high dimensionality. Thus, it is essential to identify the most relevant subset of channels for effective authentication. Research indicates that reducing the number of EEG channels can notably enhance performance in biometric applications [24]. However, contrary findings suggest that performance may decrease with channel reduction [25]. Additionally, some studies advocate using a single channel for biometric authentication [26], [27]. Notably, channel optimization appears to be indirectly influenced by stimuli [28]. In this investigation, we have fine-tuned the selection of EEG channels for biometric authentication, focusing on wearable applications with a specific emphasis on real-world contexts and emotional stimuli.

The feature extraction helps to identify a set of metrics from the EEG signals of subjects that well represent their identity [29]. Despite the significant advancements in EEG-based biometric authentication, several challenges persist in extracting subject-specific discriminative features due to the non-stationary nature of EEG signals [30], [31]. While various studies have employed statistical [32], time-frequency [33], spectral [34], fractal [35], shape [36], and entropy [37] features for EEG-based biometric authentication, there is a lack of comprehensive approaches that encapsulate the full breadth of information inherent in EEG data. In this study, we have used a wide range of features from multiple domains to improve the performance of an EEG-based authentication system.

The optimization of features extracted from EEG signals stands as a cornerstone in the development of highly accurate EEG-based biometric systems. Studies have used Fisher discriminant ratio, recursive feature elimination [24], mutual information [38], principal component analysis, linear discriminant analysis [39], genetic algorithm, correlation map, ANOVA F-test, logistic regression weights [29], maximum information coefficient, and quantum particle swarm [40] for ideal feature identification. However, studies have minimally explored machine learning algorithms to identify optimal features for EEG-based biometric authentication. In this study, we identified significant features that can reduce computational complexity and enhance the performance of EEG-based authentication systems using machine learning algorithms.

Various machine learning algorithms are used in EEG-based biometric authentication systems to analyze vast amounts of biometric data in real-time, concentrate on the essential aspects of the data, learn and adapt to new patterns and trends, and improve their accuracy and reliability over time. Linear classifiers such as linear support vector machine (SVM) [41], logistic regression [29], and Naive Bayes [42] have been used in EEG-based biometrics applications due to their simplicity and effectiveness. Non-linear classifiers, including artificial neural networks [43], k-nearest neighbors [44], and other SVM variants, have also been utilized, offering flexibility in handling complex, non-linear data patterns. Deep learning techniques, such as convolutional neural networks [10] and recurrent neural networks [45], have been explored for their ability to automatically learn hierarchical representations of data. However, deep learning methods are computationally expensive, require a large amount of data, and are prone to overfitting [46]. On the other hand, ensemble machine learning algorithms like random forest (RF) [47] and extreme gradient boost (XGBoost)

[48] have been underutilized; they offer the advantage of leveraging multiple classifiers to potentially achieve higher accuracy and improved generalization performance compared to individual classifiers. In this study, we used different ensemble machine-learning algorithm variants for EEG-based algorithm variants for EEG-based biometrics authentication systems.

The main contributions of this study are

- We analyzed the influence of different emotions in EEG-based biometrics authentication.
- We have fine-tuned the selection of EEG channels for biometric authentication for wearable applications.
- We have used a wide range of features from multiple domains and optimized them to improve the performance of EEG-based authentication systems.
- We proposed a novel process pipeline for EEG-based biometrics systems.

We have discussed the methodology adopted in this study in section II. Then, we have presented the findings of the study in section III. Section IV is dedicated to discussing the proposed method in comparison to existing approaches. In Section V, we discuss the limitations and future scope of the studies. Finally, Section VI concludes the study by summarizing the key findings and their implications.

II. METHODOLOGY

Figure 1 shows the process pipeline adopted in the study, which includes segregating the EEG signals based on annotations, preprocessing, feature extraction, classifier training, and performance evaluation.

A. DATASET

EEG signals used in the study were obtained from the publicly available Database for Emotion Analysis using the Physiological Signals (DEAP) dataset [49]. It consists of a comprehensive collection of EEG recordings from 32 participants. Each participant was exposed to a series of stimuli designed to elicit various emotional responses. These stimuli were captured as 40 videos, with the intent to assess emotional states corresponding to valence, liking, arousal, and dominance. Each participant's data comprised a robust set of 40 videos x 32 channels x 8064 data matrix that encapsulates numeric data suitable for advanced analytical pursuits. The dataset was meticulously labeled with rating values ranging from 1 to 9 for each emotional label, ensuring a rich framework for examining the nuances of emotional response. The ratings falling between 1 and 5 were categorized as 'Low,' while those ranging from 5 to 9 were designated as 'High.' The resulting four groups are as follows: high valence high arousal (HVHA: valence > 5, arousal > 5), high valence low arousal (HVLA: valence > 5, arousal ≤ 5), low valence high arousal (LVHA: valence ≤ 5, arousal > 5), and low valence low arousal (LVLA: valence ≤ 5, arousal ≤ 5). Through this categorization process, an imbalanced sample size was obtained, comprising 352 HVHA, 288 HVLA, 288 LVHA, and 352 LVLA signals. HVLA and LVLA have fewer samples compared to the HVHA and HVLA classes.

TABLE 1. Dataset description.

Description	Value
Number of participants	32 (Male-16, Female-16)
Age	19-37 (mean age 26.9)
Number of EEG channels	32
Number of videos	40 per participant
Sampling rate	512 Hz
Number of dataset labels	4 labels
Names of labels	Valence, liking, arousal, dominance
Rating values for each label	1 to 9
Data structure for each participant	40 videos x 32 channels x 8064 data

B. PREPROCESSING

The raw EEG signals were recorded at a high frequency of 512 Hz to ensure data integrity. Initially, the data underwent common average referencing. To reduce data volume and computational load, the signals were down-sampled to 128 Hz. Electro-oculographic artifacts caused by eye movements were eliminated using an independent component analysis technique to reduce noise [50]. A bandpass filter ranging from 0.4 to 45 Hz was applied to retain frequencies relevant to emotional processing. Finally, the signals were segmented into 60-second trials, with a 3-second pre-trial baseline removed.

C. FEATURE EXTRACTION

We extracted a total of 76 features for each channel, which consist of statistical, time-domain, frequency-domain, entropy, fractal and chaos theory, spectral, and shape features as shown in Table 2. This comprehensive set of features, derived from each channel's data for every clip and subject, was methodically organized into a cohesive dataset. We ensured a structured and informative dataset by maintaining detailed records of channel, clip number, and subject ID alongside the extracted features.

D. CLASSIFIER TRAINING AND PERFORMANCE EVALUATION

We implemented a range of tree-based machine learning classifiers chosen for the analysis, including bagging classifier (BC), RF, gradient boosting machine (GBM), XGBoost, light gradient boosting machine (LightGBM), and categorical boosting (CatBoost).

1) BAGGING CLASSIFIER

A bagging classifier operates as an ensemble meta-estimator by training base classifiers on random subsets of the original dataset and then combining their individual predictions, typically through voting or averaging, to yield a final prediction [51]. Let $L = \{(y_i, x_i), i = 1, \ldots, N\}$ represent a learning sample comprising N independent observations. Each observation consists of p-dimensional predictor vectors $x_i = (x_{i1}, \ldots, x_{ip}) \in R^P$ and corresponding class labels.

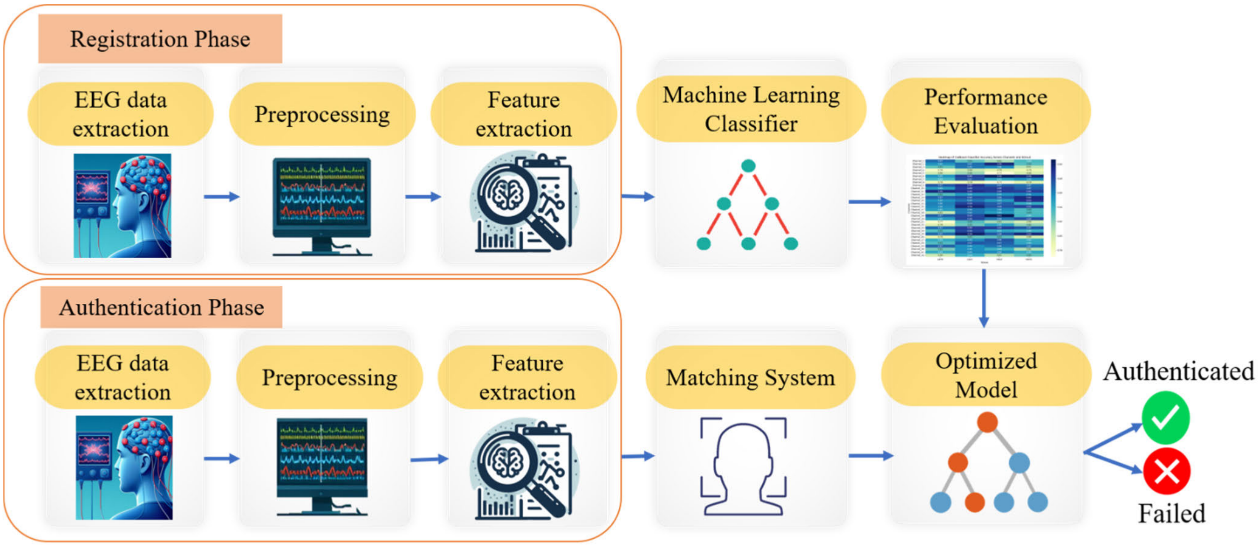


FIGURE 1. Process pipeline of the study.

TABLE 2. Feature description.

Feature Group	Features	Description
Statistical Features	Max, min, mean, median, variance, standard deviation, skewness, kurtosis	Basic statistical measures along with distribution characteristics of the EEG signal.
Time-Domain Features	Root mean square (RMS), Hjorth complexity (HC), Hjorth mobility (HM), Hjorth activity (HA), mean line length (MLL), Teager energy, zero crossing rate (ZCR)	Measures of signal magnitude, complexity, and rate of zero crossings in the time domain.
Frequency-Domain Features	Band powers (delta, theta, alpha, sigma, beta, gamma), total power, high-frequency power, band ratios	Power in specific EEG frequency bands, total power, and ratios of band powers.
Entropy Measures	Entropy, sample entropy (SE)	Quantifies the unpredictability or complexity of signal patterns.
Fractal Features	Hurst exponent(h), Katz fractal dimension (KFD), Higuchi fractal dimension (HFD)	Indicators of the fractal dimension and long-term memory of the EEG signals.
Spectral Features	Mel-frequency cepstral coefficients (MFCC) 1 to MFCC 13	Mel-frequency cepstral coefficients capture the power spectrum of signal frequencies.
Shape Features	Formfactor	Describes the shape of the signal

A new observation $\tilde{x}$ is classified by majority voting:

$$C_A{}^B(\tilde{x}) = argmaxj \in \{1, 2\} \sum_{b=1}^{B} \chi_{\{j\}} \left(C \left(\tilde{x}, L^{\star(b)} \right) \right) \quad (1)$$

where B is a random sample, C_A is an aggregated classifier, L is a learning sample, χ is the indicator function, $C_A{}^B$ is the bagged classifier.

2) RANDOM FOREST

RF is an extended variant of bagging. Initially, a bootstrap sample Z^* was randomly selected from the training set. Considering these samples as the training data, decision trees T_b were established. Then, a subset of M features is randomly selected from the feature set of each node of the decision tree. The RF tree is deployed to improve the efficacy of binding data by iteratively applying the outlined procedures to each terminal node of the decision tree until the decision tree can effectively classify the training dataset while meeting the minimum node size criteria. During model training, the CART algorithm is utilized for node splitting, with the Gini value from the Gini index serving as the criterion for selecting the splitting node. The sample training set Z^* contains different characteristics, and the Gini index of this training set (up to the k-th percentile) is:

$$GINI\ (k) = 1 - \sum_{i=1}^{k} p_i \quad (2)$$

where, p_i is the probability of a category i feature. The number of features corresponding to the sample training set were $\{n_1, n_2, \ldots, n_k\}$, $n = n_1 + n_2 + n_3$, the split Gini index is:

$$GINI\left(M^*\right) = \frac{n_1}{n} GINI\ (M_1) + \frac{n_2}{n} GINI\ (M_2) + \frac{n_3}{n} GINI\ (M_3) \quad (3)$$

Finally, all decision trees $\{T_b\}_1^m$ are aggregated. M1, M2, and M3 represent different subgroups within the entire population. For an input sample, the decision trees of m have recognition results of m, and the RF model inherits all the recognition voting results. Forecasting is performed on the new node, and the class with the most votes is the output $\left\{\hat{C}_b(x)\right\}_1^m$ [52].

3) GRADIENT BOOSTING MACHINE

Gradient boosting is a type of ensemble learning that combines a set of weak learners to construct a strong learner. In contrast to the bagging technique, where the models are made independently, the models in the ensemble boosting technique are made sequentially by iteratively minimizing the error of earlier learned models [53]. It learns a predictive model by combining the M additive tree models (f0, f1, f2, ..., fM) to predict the results by:

$$f(x) = \sum_{m=0}^{M} f_m(x) \tag{4}$$

The tree ensemble model is optimized by reducing the expected generalization error L by:

$$L = \sum_{i}^{n} \left(y_i - \hat{y}_i\right)^2 \tag{5}$$

L is a loss function that measures the delta loss between the target y_i and the prediction $\hat{y}_i$ of a data point.

4) EXTREME GRADIENT BOOSTING

The XGBoost model, a variant of GBM, effectively amalgamates the predictions of numerous weak sub-learners, typically regression trees, each of which performs slightly better than random guessing. Employing a gradient learning approach, this model iteratively builds a strong learner. Initially, a weak sub-learner is trained, followed by sub-learners tailored to the residuals of the preceding learners. This sequential training process continues until a predefined stopping criterion is met, resulting in an ensemble model weighted by the contributions of these individual weak sub-learner's predictions. Let x_i be a feature vector; then the XGBoost model can be seen as the ensemble of K additive functions.

$$\hat{y}_i = \emptyset(x_i) = \sum_{k=1}^{K} f_k(x_i), f_k \in F_j \tag{6}$$

where F_j is a group of regression trees. The function f_k makes the k prediction based on a certain output. XGBoost simplifies its training process to prevent overfitting by using a loss function combined with penalty terms. The XGBoost model is built on the loss + penalty objective function.

$$Obj^{(t)} = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i\right) + \sum_{i=1}^{k} \Omega(f_i) \tag{7}$$

where l is a loss function that is used to measure the difference between the target y_i and the prediction $\hat{y}_i$. Ω is used to penalize the complexity of the model, and it is calculated based on the scores of each leaf and the number of leaves [54].

5) LIGHT GRADIENT BOOSTING MACHINE

LightGBM [54] is a gradient-boosting framework that uses decision tree-based learning algorithms. It builds a multi-classifier system by combining weak models into a strong model using an iterative way. Specifically, when building initial classifiers, a decision tree $F_{k,0}(x)$ is created for each label category by:

$$F_{k,0}(x) = 0.5 \times log\left(\frac{\sum_{i=1}^{N} y_i}{\sum_{i=1}^{N}(1 - y_i)}\right) \tag{8}$$

where N is sample features, k is the label category, and y is the real label of the sample. We use multi-class log loss as the loss function, $L(y, F)$, and the probability $p_k(x)$ is calculated based on the Softmax function [55].

$$L(y, F) = -\sum_{k=1}^{n} y_k log p_k(x) \tag{9}$$

6) CATEGORICAL BOOSTING

CatBoost is a type of gradient boosting on the decision trees that can handle categorical, ordered features and the overfitting of the model by Bayesian estimators. It is designed to improve an objective function by creating a sequence of decision trees. The main function of CatBoosting can be expressed as [56]:

$$\text{Obj}(\theta) = \sum_{i=1}^{N} L(y_i, F(x_i)) + \sum_{t=1}^{T} \Omega(f_t) \tag{10}$$

where θ represents the model parameters, $F(x_i)$ is the ensemble prediction for input x_i, L is the loss function measuring the discrepancy between predicted and true labels, and $\Omega(f_t)$ is a regularization term for the t-th decision tree. Each decision tree is constructed by recursively partitioning the feature space based on binary splits. The splitting process minimizes the loss function by determining optimal threshold values for each feature.

E. PERFORMANCE METRICS CALCULATION AND HYPERPARAMETER OPTIMIZATION

For each emotional stimulus category (HAHV, HALV, LALV, and LAHV), the classifiers were trained channel-wise to evaluate the performance of each EEG channel independently. Cross-validation was conducted with a 5-fold strategy to ensure the reliability and generalizability of the model. Each test fold in HVHA and LVLA contained approximately 70 samples, except for two folds, which included 71 samples. Correspondingly, the training sets consisted of either 282 or 281 samples, depending on the test fold composition. Similarly, each test fold in HVLA and LVHA contained approximately 58 samples, except for two folds, which included 57 samples. Correspondingly, the training sets consisted of either 231 or 230 samples, depending on the test fold composition. Performance metrics, including accuracy, F1 score, recall, and precision, were calculated for each fold and then averaged to provide a comprehensive evaluation of each classifier's effectiveness. From the evaluation, we have selected the best-performing classifier, CatBoost, for further analysis. In the case of the CatBoost, hyperparameter tuning was conducted through grid search coupled with cross-validation to determine the optimal model parameters, balancing model complexity and learning efficiency, as tabulated in Table 3.

TABLE 3. Catboost hyperparameter optimization.

Parameters	Values
Depth	4, 6, 8
Learning rate	0.05, 0.1, 0.2
L2_leaf_reg	3, 5, 10
RSM	0.7, 0.8, 0.9
Iterations	100, 200, 500
Random seed	42

III. RESULTS

Figure 2 visualizes the performance of six machine learning algorithms (BC, CatBoost, GBM, Light GBM, RF, and XGBoost) through the range of performance metrics (accuracy, F1 score, precision, and recall) for four different stimuli (HAHV, HALV, LAHV, and LALV) and generalized model (GM). The performance metrics have been aggregated across multiple channels to show the distribution of each metric for each model. We can see that the performance of each model varies significantly across different channels, suggesting that the characteristics of each channel uniquely influence the model's effectiveness. It can be observed from Figure 2 (a) that CatBoost performed exceptionally well for the HAHV stimulus, achieving an average accuracy of 84%, an F1 score of 83%, a precision of 86%, and a recall of 84%. This was followed by strong performances from Random Forest (RF) and LightGBM. However, GBM showed poor performance compared to other classifiers. We can observe similar results in Figure 2 (b) for HALV, with CatBoost producing high average accuracy, F1 score, precision, and recall at 85%, 83%, 84%, and 85%, respectively. In Figure 2 (c), it is evident that CatBoost is performing well for LAHV with average accuracy, F1 score, precision, and recall at 83%, 81%, 83%, and 84%, respectively. For LALV, Catboost again emerged as the best classifier with average accuracy, F1 score, precision, and recall at 86%, 85%, 88%, and 86%, respectively, as seen in Figure 2 (d). It shows a smaller interquartile range for accuracy and precision, suggesting more consistent performance. A similar trend has been observed for four different stimuli, i.e., HAHV, HALV, LAHV, and LALV. CatBoost shows relatively high average accuracy values across the channels and is further finetuned to identify optimal channels. We combined all the data and built classifier models, with the results displayed in Figure 2 (e). From the figure, CatBoost emerged as the best performer, achieving a median accuracy of 91%, precision of 92%, F1 score of 91%, and recall of 91%, with minimal variability. LightGBM and Random Forest followed closely, while GBM and BC showed lower and less consistent performance. The aggregated analysis highlights CatBoost's robustness and ability to generalize effectively across multi-stimuli datasets, making it a reliable choice for classification tasks.

Figure 3 shows the accuracy of a CatBoost classifier across various channels for four different stimuli: HAHV, HALV, LAHV, LALV, and GM. Most channels have a high accuracy rate, primarily between 80% and 90%. This indicates that CatBoost consistently produced good performance across different channels. Considering the stimulus HAHV, channel FC1 produced the highest accuracy of 90%, followed by channel C4 (89%) and channel PO3 (89%). In the case of the HALV stimulus, channel Fz showed the highest accuracy of 91%, followed by channel AF3 and FC1, showing an accuracy of 90%. For the LAHV stimulus, the highest accuracy was achieved by channel P3 (90%), followed by channels CP5 and PO3 (87%). We received a peak accuracy of 90% for LALV using channels C4 and Pz, followed by channel P3 (89%). It can be noted that the classifier's performance varies across different stimuli. LAHV and LALV stimuli tend to have slightly higher accuracy rates in several channels compared to HALV and HAHV. This suggests that the classifier is more effective at distinguishing between low arousal states compared to high arousal states, irrespective of the valence. When comparing the high valence conditions (LAHV and HAHV), the accuracy fluctuates minimally, suggesting that the classifier's performance is relatively stable across different levels of arousal for high valence stimuli. We found that channels P3, Pz, FC2, and C4 were common in the top 10 performing channels for all four stimuli, which can be utilized for wearable applications. When considering the GM analysis, the highest accuracy of 94% was observed in channels CP1 and CP5, followed closely by channels C4, Fz, and Pz, all achieving an accuracy of 93%. This performance highlights the classifier's robustness in aggregating data across all stimuli, with some channels consistently performing well across conditions. Channels P3, Pz, FC2, and C4 were common among the top-performing channels for all stimuli and the overall measure, further emphasizing their reliability for wearable applications.

Figure 4 (a-e) presents an important feature analysis of the top ten performing channels for HAHV, HALV, LAHV, LALV emotional stimuli, and GM obtained after employing the CatBoost classifier. The graph shows several prominent peaks for features such as skewness, theta-to-alpha ratio, beta-to-alpha ratio, MFCC 4, MFCC 5, HFD, and SE features standing out as especially important. These peaks are crucial as they represent features that the classifier deems most informative for distinguishing individual biometric patterns.

IV. DISCUSSIONS

A. CLASSIFIER PERFORMANCE VARIABILITY

In this study, we employed six machine learning algorithms, such as BC, RF, GBM, XGBoost, LightGBM, and

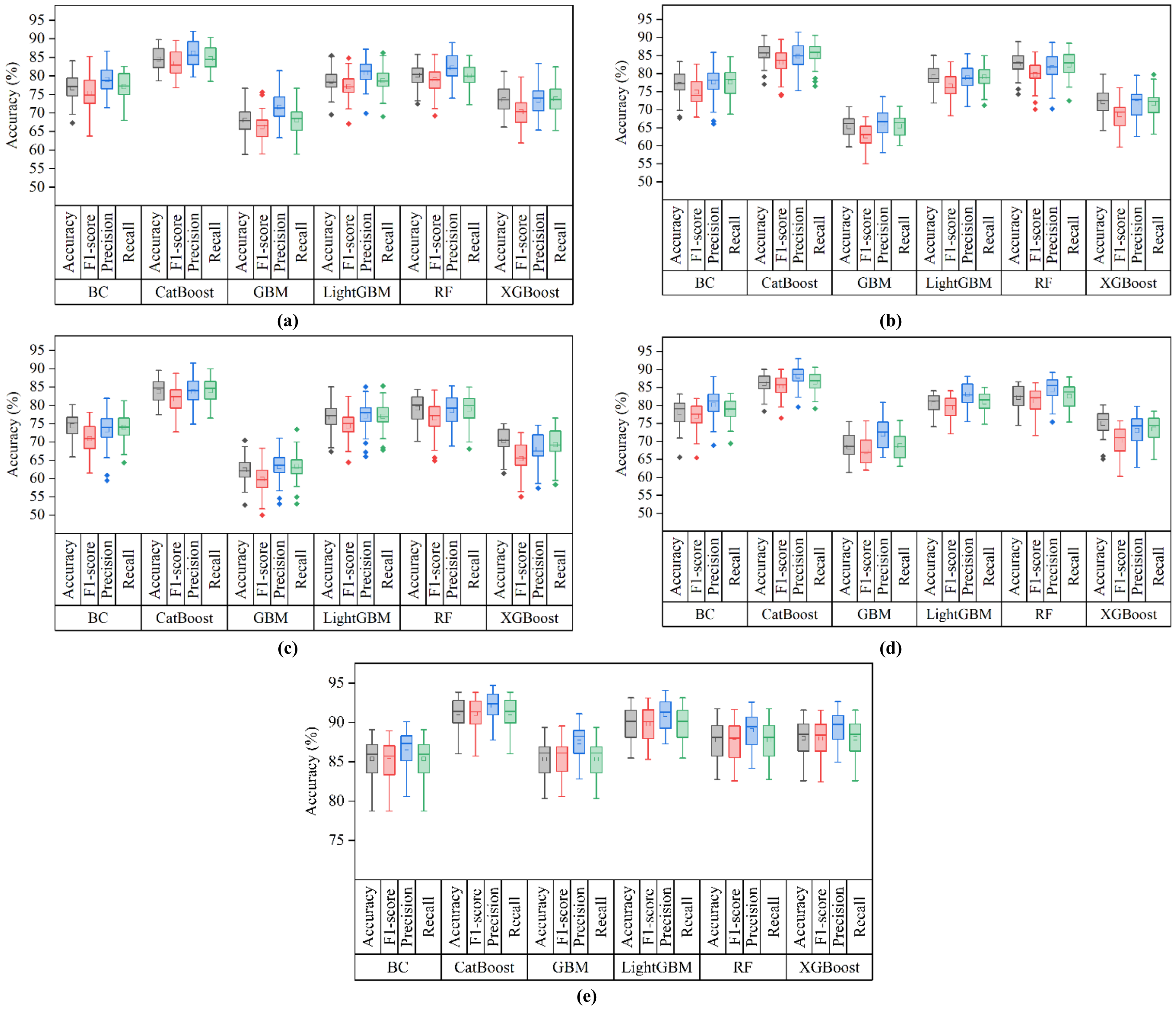


FIGURE 2. Average performance of classifiers for (a) HAHV, (b) HALV, (c) LAHV, (d) LALV stimulus and (e) GM.

CatBoost, to build an EEG-based biometric authentication system. Our results suggest that CatBoost is performing well irrespective of the stimuli and GM compared to other classifiers. We achieved an accuracy of 84%, 89%, 76%, 81%, 85%, and 91% by BC, RF, GBM, XGBoost, LightGBM, and CatBoost, respectively. Most of the time, maximum accuracy is achieved for HALV and HAHV stimuli. It can be noted that CatBoost performed well compared to other classification models. The superior performance shown by CatBoost may be due to its advantage in dealing with multicollinearity in features and the use of regularization terms, which avoids overfitting and improves generalization on unseen data [57], [58]. CatBoost has never been used for EEG-based authentication systems; however, it was implemented in emotion recognition, which outperformed other ensemble classifiers [59]. It can be noted that a few studies have reported higher accuracy than our results in EEG-based authentication systems. However, these studies used a large number of channels [60] and computationally expensive deep-learning models [61], [62]. Several studies have also reported lower accuracy than the accuracy obtained in our study, which shows the advantage of our proposed model [63], [64].

B. CHANNEL PERFORMANCE

We evaluated the contribution of 32 channels of EEG signals used in the study using machine learning algorithms. Our results suggest that channels such as P3, Pz, FC2, and C4 were common in the top 10 performing channels for all four stimuli. P3 and Pz electrodes are positioned atop the parietal lobe, while the FC2 electrode is situated over the

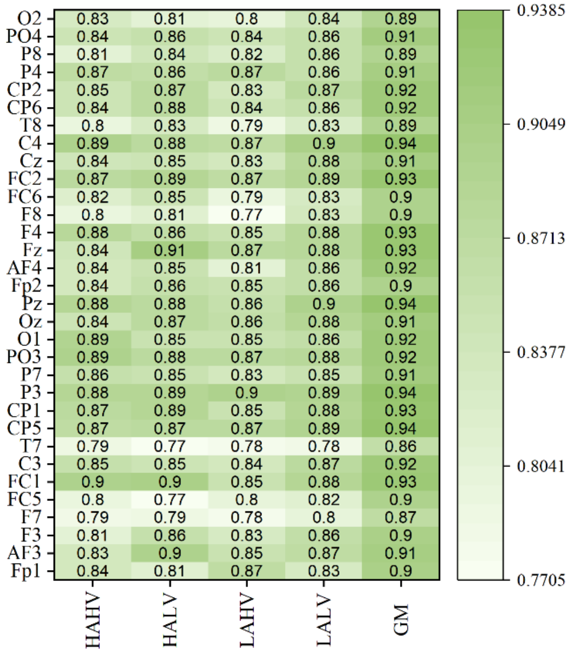


FIGURE 3. Performance of CatBoost classifier across the channels for different stimuli HAHV, HALV, LAHV, LALV, and GM.

supplementary motor area of the frontal lobe. C4 is over the central region of the brain and is commonly used to locate the motor cortex [59]. It reveals the importance of utilizing channels from the frontal and parietal lobes in EEG-based biometrics. Further, the channels P3, Pz, FC2, and C4 are located in areas of the brain that are involved in various cognitive and emotional processes [65]. Studies have reported the influence of P3 [66], Pz [67], FC2 [35], and C4 [28] channels in EEG-based authentication systems. Research indicates that reducing the number of EEG channels can notably enhance performance in biometric applications [68]. However, contrary findings suggest that performance may decrease with channel reduction [25]. Additionally, a recent study advocates for using a single channel for biometric authentication [66]. Notably, channel optimization appears to be indirectly influenced by stimuli [28].

C. PERFORMANCE VARIABILITY ACROSS STIMULI

We analyzed the performance of the EEG-based authentication system on four different emotional stimuli: HAHV, HALV, LAHV, and LALV. Our results revealed that the proposed authentication system performs well in LALV stimulus followed by HALV, HAHV, and LAHV. It can be noted that the developed authentication system performs better on low arousal (LAHV and LALV) compared to high arousal (HALV and HAHV). The high performance of the system during the low arousal state reveals that individuals may be more relaxed and able to maintain steady attention and focus, leading to more consistent and discernible EEG patterns. In contrast, high arousal stimuli can trigger heightened physiological and cognitive responses, causing fluctuations in attention and arousal levels, which may affect the stability and reliability of EEG signals and, thereby, the authentication system. Alternatively, high arousal states may lead to early saturation of the signals, and interindividual variability may be less. Our findings align with prior research indicating that EEG-based biometrics exhibit improved performance during calm emotional states compared to stressed or excited states, although this may vary depending on signal duration [69]. A recent study demonstrated that LVLA outperformed HVLA for EEG-based biometrics [20], [21]. Conversely, another study found that resting-state stimuli yielded superior results to affective stimuli [70]. A study employed emotional stimuli for EEG-based authentication, but it never examined the impact of various emotional stimuli [21]. Additionally, research suggests that training EEG systems with recordings from similar emotional states enhances identification accuracy compared to training with recordings from disparate emotional states [71]. Another notable outcome is the improved performance of the GM compared to stimulus-specific models. By aggregating data across all stimuli, we likely captured more comprehensive and consistent EEG signatures that generalize well across individuals.

A study demonstrated that combining multiple visual stimuli in visual evoked potential-based EEG biometrics resulted in more distinctive and stable EEG patterns, thereby improving system accuracy [72].

D. FEATURE SIGNIFICANCE

We analyzed the effect of 76 features from different domains in an EEG-based biometric authentication system. Our results revealed that features such as skewness, theta-to-alpha ratio, beta-to-alpha ratio, MFCC 4, MFCC 5, HFD, and SE significantly contributed to the authentication system, irrespective of the stimuli. These features are from multiple domains such as statistical (skewness), frequency (theta to alpha ratio, beta to alpha ratio), spectral (MFCC 4, MFCC 5), fractal dimension (HFD), and entropy measures (SE). Skewness measures the asymmetry of a distribution in EEG signals, which can indicate underlying brain dynamics. As per the study by [73], skewness has influential contributions to EEG-based person authentication. It has been reported that the beta/alpha ratio is highly related to the arousal dimension of emotions [74], [75]. A study highlights the significance of alpha waves in EEG-based biometrics for triggering emotions. It demonstrates that changes in these brainwave patterns during alertness and anticipation could influence the accuracy of the authentication system [76]. A study by [71] revealed that MFCC features performed well for emotion-elicited EEG-based biometrics compared to other features for emotion-elicited EEG-based biometrics.

V. LIMITATIONS AND FUTURE SCOPE

In the realm of EEG-based biometrics, our study has identified several areas for improvement and future exploration. Firstly, the performance of channel combinations warrants

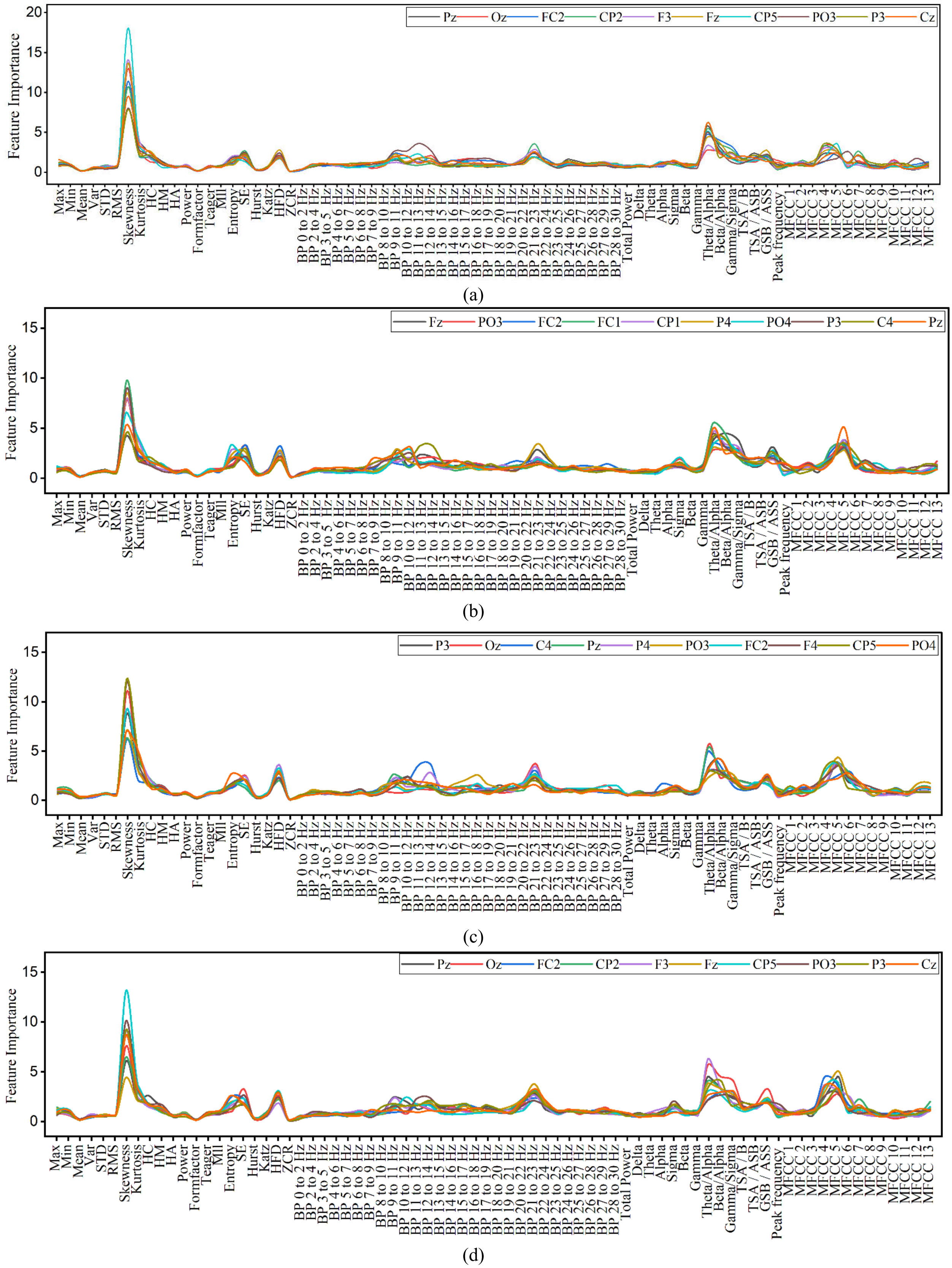
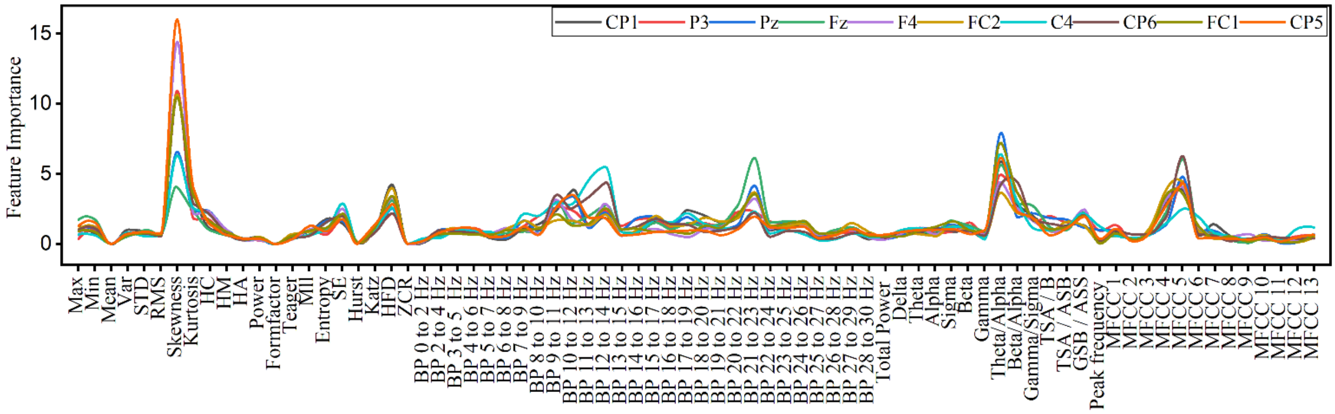


FIGURE 4. Feature importance of CatBoost classifier across the top channels for different stimuli (a) HAHV, (b) HALV, (c) LAHV, (d) LALV. (e) GM. (TSA/B (ratio of theta sum alpha to beta), TSA/ ASB (ratio of theta sum alpha to alpha sum beta, GSB/ASS (ratio of gamma sum beta to alpha sum sigma).



(e)

FIGURE 4. Feature importance of CatBoost classifier across the top channels for different stimuli (a) HAHV, (b) HALV, (c) LAHV, (d) LALV. (e) GM. (TSA/B (ratio of theta sum alpha to beta), TSA/ ASB (ratio of theta sum alpha to alpha sum beta, GSB/ASS (ratio of gamma sum beta to alpha sum sigma).

further investigation to enhance authentication accuracy, which may require high computational power. Secondly, optimizing the length of video stimuli for real-world applications is a critical need. To enhance the model's generalizability, it is imperative to evaluate its performance across diverse datasets, incorporating an expanded test set and real-time data representing individuals from varied demographic backgrounds. This approach ensures a larger and more representative sample size, thereby improving the robustness and applicability of the model. We performed 5-fold cross-validation with limited iterations, which may impact the stability and robustness of the results, though future work will address this through repeated cross-validation.

While our study utilized time and frequency domain features for authentication, the incorporation of advanced features could reveal intricate details that may boost the efficiency of the authentication system. Looking ahead, there are several promising avenues for enhancing the robustness and accuracy of emotion-based authentication systems. Future work will focus on analyzing inter-session repeatability using in-house recorded EEG data with multiple sessions to enhance the applicability of EEG-based biometric authentication for wearable applications. One such avenue is the integration of multimodal data to complement EEG signals. Another is the exploration of lightweight deep learning approaches, which can automatically extract hierarchical features from raw EEG data. It potentially improves performance without compromising computational complexity in handling the inherent variability and complexity of EEG signals. Further, we can improve the robustness of an EEG-based authentication system by training the model on one emotion category and testing it on other emotion categories. Finally, customizing EEG-based emotion recognition systems to leverage high-performing channels and accommodate individual variability in emotional processing could enhance their applicability in real-world scenarios, such as mental health monitoring or adaptive human-computer interfaces.

VI. CONCLUSION

In this study, a comprehensive pipeline was proposed to examine the performance of an EEG-based authentication system across different emotional stimuli, channels, and features, utilizing machine learning algorithms. Our results demonstrated that the CatBoost classifier consistently outperformed other classifiers, achieving notable accuracies across all stimuli. Specifically, accuracies of 84%, 89%, 76%, 81%, 85%, and 91% were attained by BC, RF, GBM, XGBoost, LightGBM, and CatBoost, respectively. We found that the proposed authentication system performs well in LALV stimulus followed by HALV, HAHV, and LAHV. Channels such as P3, Pz, FC2, and C4 emerged as consistently strong performers across all stimuli, suggesting their relevance in EEG-based authentication. Moreover, features including skewness, theta-to-alpha ratio, beta-to-alpha ratio, MFCC 4, MFCC 5, HFD, and SE were found to be particularly influential in enhancing authentication accuracy, irrespective of the emotional stimulus. Overall, our findings highlight the efficacy of EEG signals in biometric authentication, offering a robust framework that accounts for emotional variations and identifies key brain locations and features for optimal authentication performance. By providing insights into the reliability of EEG-based authentication systems across different emotional states, this research contributes to advancing the development of secure and adaptable biometric authentication technologies.

ACKNOWLEDGMENT

The authors would like to thank the support of the PARAM Shivay Supercomputer Facility of the Indian Institute of Technology (BHU), Varanasi, India, used for this research.

REFERENCES

[1] C. Lin and A. Kumar, "Matching contactless and contact-based conventional fingerprint images for biometrics identification," *IEEE Trans. Image Process.*, vol. 27, no. 4, pp. 2008–2021, Apr. 2018, doi: 10.1109/TIP.2017.2788866.

[2] A. S. Al-Waisy, R. Qahwaji, S. Ipson, S. Al-Fahdawi, and T. A. M. Nagem, "A multi-biometric iris recognition system based on a deep learning approach," *Pattern Anal. Appl.*, vol. 21, no. 3, pp. 783–802, Oct. 2017, doi: 10.1007/s10044-017-0656-1.

[3] T. Phillips, X. Zou, F. Li, and N. Li, "Enhancing biometric-capsule-based authentication and facial recognition via deep learning," in *Proc. 24th ACM Symp. Access Control Models Technol.*, New York, NY, USA, May 2019, pp. 141–146, doi: 10.1145/3322431.3325417.

[4] S. Shakil, D. Arora, and T. Zaidi, "Feature based classification of voice based biometric data through machine learning algorithm," *Mater. Today, Proc.*, vol. 51, pp. 240–247, Jan. 2022, doi: 10.1016/j.matpr.2021.05.261.

[5] W. Yang, S. Wang, J. Hu, G. Zheng, and C. Valli, "Security and accuracy of fingerprint-based biometrics: A review," *Symmetry*, vol. 11, no. 2, p. 141, Jan. 2019, doi: 10.3390/sym11020141.

[6] K. Revett and S. T. de Magalhães, "Cognitive biometrics: Challenges for the future," in *Global Security, Safety, and Sustainability*, vol. 92, S. T. de Magalhães, H. Jahankhani, and A. G. Hessami, Eds., Berlin, Germany: Springer, 2010, pp. 79–86.

[7] E. Piciucco, E. Di Lascio, E. Maiorana, S. Santini, and P. Campisi, "Biometric recognition using wearable devices in real-life settings," *Pattern Recognit. Lett.*, vol. 146, pp. 260–266, Jun. 2021, doi: 10.1016/j.patrec.2021.03.020.

[8] S. Aziz, M. U. Khan, Z. A. Choudhry, A. Aymin, and A. Usman, "ECG-based biometric authentication using empirical mode decomposition and support vector machines," in *Proc. IEEE 10th Annu. Inf. Technol., Electron. Mobile Commun. Conf. (IEMCON)*, Oct. 2019, pp. 0906–0912, doi: 10.1109/IEMCON.2019.8936174.

[9] M. U. Khan, Z. A. Choudry, S. Aziz, S. Z. H. Naqvi, A. Aymin, and M. A. Imtiaz, "Biometric authentication based on EMG signals of speech," in *Proc. Int. Conf. Electr., Commun., Comput. Eng. (ICECCE)*, Jun. 2020, pp. 1–5, doi: 10.1109/ICECCE49384.2020.9179354.

[10] H. Vadher, P. Patel, A. Nair, T. Vyas, S. Desai, L. Gohil, S. Tanwar, D. Garg, and A. Singh, "EEG-based biometric authentication system using convolutional neural network for military applications," *Secur. Privacy*, vol. 7, no. 2, p. e345, Mar. 2024, doi: 10.1002/spy2.345.

[11] F. Sabry, T. Eltaras, W. Labda, K. Alzoubi, and Q. Malluhi, "Machine learning for healthcare wearable devices: The big picture," *J. Healthcare Eng.*, vol. 2022, pp. 1–25, Apr. 2022, doi: 10.1155/2022/4653923.

[12] B. Fernandez-Saavedra, J. Liu-Jimenez, and R. Sanchez-Reillo, "Environmental testing methodology of biometric system performance, standardization," in *Encyclopedia of Biometrics*, S. Z. Li and A. K. Jain, Eds., Boston, MA, USA: Springer, 2015, pp. 416–420.

[13] L.-I. Hsu, K.-W. Lim, Y.-H. Lai, C.-S. Chen, and L.-W. Chou, "Effects of muscle fatigue and recovery on the neuromuscular network after an intermittent handgrip fatigue task: Spectral analysis of electroencephalography and electromyography signals," *Sensors*, vol. 23, no. 5, p. 2440, Feb. 2023, doi: 10.3390/s23052440.

[14] M. S. Qasim, D. S. Bari, and Ø. G. Martinsen, "Influence of ambient temperature on tonic and phasic electrodermal activity components," *Physiological Meas.*, vol. 43, no. 6, Jun. 2022, Art. no. 065001, doi: 10.1088/1361-6579/ac72f4.

[15] D. Lee, H.-J. Lee, and S.-G. Lee, "Motor imagery EEG classification method using EMD and FFT," *J. KIISE*, vol. 41, no. 12, pp. 1050–1057, Dec. 2014, doi: 10.5626/jok.2014.41.12.1050.

[16] D. Lopez-Bernal, D. Balderas, P. Ponce, and A. Molina, "A state-of-the-art review of EEG-based imagined speech decoding," *Frontiers Human Neurosci.*, vol. 16, Apr. 2022, Art. no. 867281, doi: 10.3389/fnhum.2022.867281.

[17] H.-L. Chan, P.-C. Kuo, C.-Y. Cheng, and Y.-S. Chen, "Challenges and future perspectives on electroencephalogram-based biometrics in person recognition," *Frontiers Neuroinform.*, vol. 12, p. 66, Oct. 2018, doi: 10.3389/fninf.2018.00066.

[18] H. Zhao, Y. Chen, W. Pei, H. Chen, and Y. Wang, "Towards online applications of EEG biometrics using visual evoked potentials," *Exp. Syst. Appl.*, vol. 177, Sep. 2021, Art. no. 114961, doi: 10.1016/j.eswa.2021.114961.

[19] D. Kim and K. Kim, "Resting state EEG-based biometric system using concatenation of quadrantal functional networks," *IEEE Access*, vol. 7, pp. 65745–65756, 2019, doi: 10.1109/ACCESS.2019.2917918.

[20] J. Jeswani, P. K. Govarthan, A. Selvaraj, A. A. Prince, J. Thomas, M. Kalathe, V. S. Sreeraj, V. Subramaniam, and J. F. A. Ronickom, "Low valence low arousal stimuli: An effective candidate for EEG-based biometrics authentication system," *Stud. Health Technol. Inform.*, vol. 302, pp. 257–261, May 2023, doi: 10.3233/shti230114.

[21] R. Khalil, A. Arasteh, and A. K. Sarkar, "EEG based biometrics using emotional stimulation data," in *Proc. IEEE Region 10 Humanitarian Technol. Conf. (R10-HTC)*, Dec. 2017, pp. 246–249, doi: 10.1109/R10-HTC.2017.8288949.

[22] T.-M. Li, H.-C. Chao, and J. Zhang, "Emotion classification based on brain wave: A survey," *Human-centric Comput. Inf. Sci.*, vol. 9, no. 1, p. 42, Dec. 2019, doi: 10.1186/s13673-019-0201-x.

[23] K. Erat, E. B. Şahin, F. Doğan, N. Merdanoğlu, A. Akcakaya, and P. O. Durdu, "Emotion recognition with EEG-based brain-computer interfaces: A systematic literature review," *Multimedia Tools Appl.*, vol. 83, no. 33, pp. 79647–79694, Mar. 2024, doi: 10.1007/s11042-024-18259-z.

[24] Y. Bai, Z. Zhang, and D. Ming, "Feature selection and channel optimization for biometric identification based on visual evoked potentials," in *Proc. 19th Int. Conf. Digit. Signal Process.*, Aug. 2014, pp. 772–776, doi: 10.1109/ICDSP.2014.6900769.

[25] E. C. Ketola, M. Barankovich, S. Schuckers, A. Ray-Dowling, D. Hou, and M. H. Imtiaz, "Channel reduction for an EEG-based authentication system while performing motor movements," *Sensors*, vol. 22, no. 23, p. 9156, Nov. 2022, doi: 10.3390/s22239156.

[26] B. Singh, S. Mishra, and U. S. Tiwary, "EEG based biometric identification with reduced number of channels," in *Proc. 17th Int. Conf. Adv. Commun. Technol. (ICACT)*, Jul. 2015, pp. 687–691, doi: 10.1109/ICACT.2015.7224883.

[27] B. C. Armstrong, M. V. Ruiz-Blondet, N. Khalifian, K. J. Kurtz, Z. Jin, and S. Laszlo, "Brainprint: Assessing the uniqueness, collectability, and permanence of a novel method for ERP biometrics," *Neurocomputing*, vol. 166, pp. 59–67, Oct. 2015, doi: 10.1016/j.neucom.2015.04.025.

[28] L. A. Moctezuma and M. Molinas, "Multi-objective optimization for EEG channel selection and accurate intruder detection in an EEG-based subject identification system," *Sci. Rep.*, vol. 10, no. 1, p. 5850, Apr. 2020, doi: 10.1038/s41598-020-62712-6.

[29] N. A. Alzahab, M. Baldi, and L. Scalise, "Efficient feature selection for electroencephalogram-based authentication," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, Jun. 2021, pp. 1–6, doi: 10.1109/MeMeA52024.2021.9478700.

[30] M. Saeidi, W. Karwowski, F. V. Farahani, K. Fiok, R. Taiar, P. A. Hancock, and A. Al-Juaid, "Neural decoding of EEG signals with machine learning: A systematic review," *Brain Sci.*, vol. 11, no. 11, p. 1525, Nov. 2021, doi: 10.3390/brainsci11111525.

[31] I. Rakhmatulin, M.-S. Dao, A. Nassibi, and D. Mandic, "Exploring convolutional neural network architectures for EEG feature extraction," *Sensors*, vol. 24, no. 3, p. 877, Jan. 2024, doi: 10.3390/s24030877.

[32] M. A. H. Akhand, M. A. Maria, M. A. S. Kamal, and K. Murase, "Improved EEG-based emotion recognition through information enhancement in connectivity feature map," *Sci. Rep.*, vol. 13, no. 1, p. 13804, Aug. 2023, doi: 10.1038/s41598-023-40786-2.

[33] C. J. Ortiz-Echeverri, O. Paredes, S. Salazar-Colores, J. Rodríguez-Reséndiz, and R. Romo-Vázquez, "A comparative study of time and frequency features for EEG classification," in *Proc. 8th Latin American Conf. Biomed. Eng. 42nd Nat. Conf. Biomed. Eng.*, vol. 75, C. A. G. Díaz, C. C. González, E. L. Leber, H. A. Vélez, N. P. Puente, D.-L. Flores, A. O. Andrade, H. A. Galván, F. Martínez, R. García, C. J. Trujillo, and A. R. Mejía, Eds., Cancún, México. Cham, Switzerland: Springer, Oct. 2020, pp. 91–97.

[34] C. M. Jijomon and A. P. Vinod, "EEG-based biometric identification using frequently occurring maximum power spectral features," in *Proc. IEEE Appl. Signal Process. Conf. (ASPCON)*, Dec. 2018, pp. 249–252, doi: 10.1109/ASPCON.2018.8748581.

[35] L. A. Moctezuma and M. Molinas, "Towards a minimal EEG channel array for a biometric system using resting-state and a genetic algorithm for channel selection," *Sci. Rep.*, vol. 10, no. 1, p. 14917, Sep. 2020, doi: 10.1038/s41598-020-72051-1.

[36] Q. Wu, B. Yan, Y. Zeng, C. Zhang, and L. Tong, "Anti-deception: Reliable EEG-based biometrics with real-time capability from the neural response of face rapid serial visual presentation," *Biomed. Eng. OnLine*, vol. 17, no. 1, p. 55, May 2018, doi: 10.1186/s12938-018-0483-7.

[37] S. Li and S.-H. Cha, "Feature extraction based on high order statistics measures and entropy for EEG biometrics," in *Proc. 7th Int. Workshop Biometrics Forensics (IWBF)*, May 2019, pp. 1–6, doi: 10.1109/IWBF.2019.8739183.

[38] C. Lee, J.-H. Kang, and S.-P. Kim, "Feature slection using mutual information for EEG-based biometrics," in *Proc. 39th Int. Conf. Telecommun. Signal Process. (TSP)*, Jun. 2016, pp. 673–676, doi: 10.1109/TSP.2016.7760968.

[39] B. Kaliraman and M. Duhan, "A new hybrid approach for feature extraction and selection of electroencephalogram signals in case of person recognition," *J. Reliable Intell. Environments*, vol. 7, no. 3, pp. 241–251, Sep. 2021, doi: 10.1007/s40860-021-00148-z.

[40] W. Chen, Y. Cai, A. Li, Y. Su, and K. Jiang, "EEG feature selection method based on maximum information coefficient and quantum particle swarm," *Sci. Rep.*, vol. 13, no. 1, p. 14515, Sep. 2023, doi: 10.1038/s41598-023-41682-5.

[41] A. J. Bidgoly, H. J. Bidgoly, and Z. Arezoumand, "Towards a universal and privacy preserving EEG-based authentication system," *Sci. Rep.*, vol. 12, no. 1, p. 2531, Feb. 2022, doi: 10.1038/s41598-022-06527-7.

[42] A. D. Wibawa, N. Fatih, Y. Pamungkas, M. Pratiwi, and P. A. Ramadhani, "Time and frequency domain feature selection using mutual information for EEG-based emotion recognition," in *Proc. 9th Int. Conf. Electr. Eng., Comput. Sci. Informat. (EECSI)*, Oct. 2022, pp. 19–24, doi: 10.23919/EECSI56542.2022.9946522.

[43] *Development of an EEG-Based Biometric Identification and Authentication System Using Artificial Neural Networks (ANN)*. Accessed: Apr. 3, 2024. [Online]. Available: https://www.researchgate.net/publication/314087244Developmentof_an_EEGbased_Biometric_Identification_and_Authentication_System_Using_Artificial_Neural_Networks_ANN

[44] S. P. Kamaraju, K. Das, and R. B. Pachori, "EEG based biometric authentication system using multivariate FBSE entropy," Indian Inst. Technol. Indore, Indore, India, Tech. Rep., May 2023, doi: 10.36227/techrxiv.23244209.v1.

[45] Y. Mohamed, A. M. Anter, and A. B. Zaky, "Recurrent neural networks (RNNs) to improve EEG-based person identification," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jul. 2023, pp. 616–621, doi: 10.1109/imsa58542.2023.10217750.

[46] I. H. Sarker, "Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions," *Social Netw. Comput. Sci.*, vol. 2, no. 6, p. 420, Aug. 2021, doi: 10.1007/s42979-021-00815-1.

[47] Q. Gui, Z. Jin, and W. Xu, "Exploring EEG-based biometrics for user identification and authentication," in *Proc. IEEE Signal Process. Med. Biol. Symp. (SPMB)*, Dec. 2014, pp. 1–6, doi: 10.1109/SPMB.2014.7002950.

[48] A. Rahman, M. E. H. Chowdhury, A. Khandakar, S. Kiranyaz, K. S. Zaman, M. B. I. Reaz, M. T. Islam, M. Ezeddin, and M. A. Kadir, "Multimodal EEG and keystroke dynamics based biometric system using machine learning algorithms," *IEEE Access*, vol. 9, pp. 94625–94643, 2021, doi: 10.1109/ACCESS.2021.3092840.

[49] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis; Using physiological signals," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 18–31, Jan. 2012, doi: 10.1109/T-AFFC.2011.15.

[50] N. Ille, Y. Nakao, S. Yano, T. Taura, A. Ebert, H. Bornfleth, S. Asagi, K. Kozawa, I. Itabashi, T. Sato, R. Sakuraba, R. Tsuda, Y. Kakisaka, K. Jin, and N. Nakasato, "Ongoing EEG artifact correction using blind source separation," *Clin. Neurophysiol.*, vol. 158, pp. 149–158, Feb. 2024, doi: 10.1016/j.clinph.2023.12.133.

[51] T. Hothorn and B. Lausen, "Bagging tree classifiers for laser scanning images: A data- and simulation-based strategy," *Artif. Intell. Med.*, vol. 27, no. 1, pp. 65–79, Jan. 2003, doi: 10.1016/s0933-3657(02)00085-4.

[52] X. Wang, G. Gong, N. Li, and S. Qiu, "Detection analysis of epileptic EEG using a novel random forest model combined with grid search optimization," *Frontiers Hum. Neurosci.*, vol. 13, p. 52, Feb. 2019, doi: 10.3389/fnhum.2019.00052.

[53] D. Sunaryono, R. Sarno, and J. Siswantoro, "Gradient boosting machines fusion for automatic epilepsy detection from EEG signals based on wavelet features," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 34, no. 10, pp. 9591–9607, Nov. 2022, doi: 10.1016/j.jksuci.2021.11.015.

[54] J. Wu, T. Zhou, and T. Li, "Detecting epileptic seizures in EEG signals with complementary ensemble empirical mode decomposition and extreme gradient boosting," *Entropy*, vol. 22, no. 2, p. 140, Jan. 2020, doi: 10.3390/e22020140.

[55] Y. Hua, "An efficient traffic classification scheme using embedded feature selection and LightGBM," in *Proc. Inf. Commun. Technol. Conf. (ICTC)*, May 2020, pp. 125–130, doi: 10.1109/ICTC49638.2020.9123302.

[56] M. A. O. Ahmed, Y. A. Satar, E. M. Darwish, and E. A. Zanaty, "Synergistic integration of multi-view brain networks and advanced machine learning techniques for auditory disorders diagnostics," *Brain Informat.*, vol. 11, no. 1, p. 3, Jan. 2024, doi: 10.1186/s40708-023-00214-7.

[57] A. Badr and A. Abdul-Hassan, "CatBoost machine learning based feature selection for age and gender recognition in short speech utterances," *Int. J. Intell. Eng. Syst.*, vol. 14, no. 3, pp. 150–159, Jun. 2021, doi: 10.22266/ijies2021.0630.14.

[58] T. Wu, X. Kong, Y. Zhong, and L. Chen, "Automatic detection of abnormal EEG signals using multiscale features with ensemble learning," *Frontiers Hum. Neurosci.*, vol. 16, Sep. 2022, Art. no. 943258, doi: 10.3389/fnhum.2022.943258.

[59] P. Mehra and P. Jain, "ERIL: An algorithm for emotion recognition from Indian languages using machine learning," *Wireless Pers. Commun.*, vol. 126, no. 3, pp. 2557–2577, Oct. 2022, doi: 10.1007/s11277-022-09829-1.

[60] U. Barayeu, N. Horlava, A. Libert, and M. Van Hulle, "Robust single-trial EEG-based authentication achieved with a 2-Stage classifier," *Biosensors*, vol. 10, no. 9, p. 124, Sep. 2020, doi: 10.3390/bios10090124.

[61] T. Yu, C.-S. Wei, K.-J. Chiang, M. Nakanishi, and T.-P. Jung, "EEG-based user authentication using a convolutional neural network," in *Proc. 9th Int. IEEE/EMBS Conf. Neural Eng. (NER)*, Mar. 2019, pp. 1011–1014, doi: 10.1109/NER.2019.8716965.

[62] H. Y. Yap, Y.-H. Choo, Z. I. M. Yusoh, and W. H. Khoh, "An evaluation of transfer learning models in EEG-based authentication," *Brain Informat.*, vol. 10, no. 1, p. 19, Aug. 2023, doi: 10.1186/s40708-023-00198-4.

[63] E. Maiorana, "Deep learning for EEG-based biometric recognition," *Neurocomputing*, vol. 410, pp. 374–386, Oct. 2020, doi: 10.1016/j.neucom.2020.06.009.

[64] B. B. Das, S. K. Ram, K. S. Babu, R. K. Mohapatra, and S. P. Mohanty, "Person identification using autoencoder-CNN approach with multitask-based EEG biometric," *Multimedia Tools Appl.*, vol. 83, no. 35, pp. 83205–83225, Mar. 2024, doi: 10.1007/s11042-024-18693-z.

[65] B. Chen, K. W. Cho, C. Xu, Z. Li, F. Lin, Z. Jin, and W. Xu, "A stimulus-response based EEG biometric using Mallows distance," *CCF Trans. Netw.*, vol. 3, no. 2, pp. 128–139, Oct. 2020, doi: 10.1007/s42045-020-00033-y.

[66] M. Zeynali and H. Seyedarabi, "EEG-based single-channel authentication systems with optimum electrode placement for different mental activities," *Biomed. J.*, vol. 42, no. 4, pp. 261–267, Aug. 2019, doi: 10.1016/j.bj.2019.03.005.

[67] M. Abo-Zahhad, S. M. Ahmed, and S. N. Abbas, "State-of-the-art methods and future perspectives for personal recognition based on electroencephalogram signals," *IET Biometrics*, vol. 4, no. 3, pp. 179–190, Sep. 2015, doi: 10.1049/iet-bmt.2014.0040.

[68] J.-H. Kang, Y. C. Jo, and S.-P. Kim, "Electroencephalographic feature evaluation for improving personal authentication performance," *Neurocomputing*, vol. 287, pp. 93–101, Apr. 2018, doi: 10.1016/j.neucom.2018.01.074.

[69] *EEG-based Person Authentication for Security Systems—University of Canberra Research Portal*. Accessed: Apr. 11, 2024. [Online]. Available: https://researchprofiles.canberra.edu.au/en/studentTheses/eeg-based-person-authentication-for-security-systems

[70] P. Arnau-González, S. Katsigiannis, M. Arevalillo-Herráez, and N. Ramzan, "BED: A new data set for EEG-based biometrics," *IEEE Internet Things J.*, vol. 8, no. 15, pp. 12219–12230, Aug. 2021, doi: 10.1109/JIOT.2021.3061727.

[71] P. Arnau-González, S. Katsigiannis, M. Arevalillo-Herráez, and N. Ramzan, "Image-evoked affect and its impact on eeg-based biometrics," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2019, pp. 2591–2595, doi: 10.1109/ICIP.2019.8803315.

[72] H. Qu, H. Zhao, Q. Wei, W. Pei, X. Gao, and Y. Wang, "Combing multiple visual stimuli to enhance the performance of VEP-based biometrics," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 7982–7993, 2024, doi: 10.1109/TIFS.2024.3452628.

[73] D. Jeswani, P. K. Govarthan, A. Selvaraj, C. B. Thomas, J. Thomas, and J. F. A. Ronickom, "A feasibility study on using EEG for biometric trait authentication system," *Current Directions Biomed. Eng.*, vol. 9, no. 1, pp. 690–693, Sep. 2023, doi: 10.1515/cdbme-2023-1173.

[74] Q. Zhang and M. Lee, "Emotion recognition in natural scene images based on brain activity and gist," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Jun. 2008, pp. 3050–3057, doi: 10.1109/IJCNN.2008.4634229.

[75] D. P. Bos, "EEG-based emotion recognition the influence of visual and auditory stimuli," Indian Dept. Comput. Sci., Univ. Twente, Enschede, The Netherlands, Tech. Rep., 2007.

[76] H. J. Lee, H. S. Kim, and K. S. Park, "A study on the reproducibility of biometric authentication based on electroencephalogram (EEG)," in *Proc. 6th Int. IEEE/EMBS Conf. Neural Eng. (NER)*, Nov. 2013, pp. 13–16, doi: 10.1109/NER.2013.6695859.

CHETAN RAKSHE received the bachelor's degree in mechanical engineering from Pimpri Chinchwad College of Engineering and Research, Ravet, in 2021. He is currently pursuing the Ph.D. degree in EEG-based biometric authentication systems with the School of Biomedical Engineering, IIT (BHU), Varanasi, India. He subsequently worked as a Junior Research Fellow at the Computational Neuroscience and Biology Laboratory, Indian Institute of Technology (BHU). His research interests include bio-signal and image processing, brain-computer interfaces, computation neuroscience, and artificial intelligence in healthcare.

CHRISTY BOBBY THOMAS received the Bachelor of Engineering degree in electronics and communication engineering from Bharathidasan University, Tiruchirappalli, and the M.S. and Ph.D. degrees in biomedical image and signal processing from Anna University, Chennai. She is currently working as a Professor with the M. S. Ramaiah University of Applied Sciences, Bengaluru, Karnataka, India, with more than 20 years of teaching experience. She has got more than 15 years of research experience in the fields of biomedical image processing, signal processing, and machine learning. She is also working as the Head of the Centre for Biomedical Systems and 3D Printing Research, M. S. Ramaiah University of Applied Sciences. She has published her research works in many International Journals and Conferences in the area of medical imaging. Her current research interests include cognitive neuroscience, neuroimaging, bio signal and image processing, and bone biomechanics.

MOHANAVELU KALATHE received the B.E. degree in instrumentation and control from the University of Madras, Chennai, India, and the M.Tech. degree in biomedical engineering from IIT Madras. He is currently working as a Scientist G and the Head of the Biomedical Technology, DEBEL, DRDO, Bengaluru. He has been working as a Scientist with Defence Bioengineering and Electromedical Laboratory (DEBEL) under the Defence Research and Development Organisation (DRDO), since 2002. His research interests include telemedicine, wearable physiological monitoring, augmentative exoskeleton, physical, cognitive performance of military personnel, cognitive exoskeleton, augmented cognition, gait analysis, and brain–computer interface. He is a member of the Biomedical Engineering Society of India (BMESI), the Institute of Smart Structure and Systems (ISSS), and the Instrument Society of India.

VANTEEMAR S. SREERAJ is currently working as an Associate Professor of psychiatry with the Department of Psychiatry, NIMHANS. He did Psychiatry Residency training at three of the esteemed psychiatric institutes of India namely, the Central Institute of Psychiatry (CIP), Ranchi; the Institute of Mental Health and Hospital (IMHH), Agra; and the NIMHANS, Bengaluru including postdoctoral fellowship in clinical neurosciences and therapeutics in schizophrenia. His research interests include neurocognitive and neurobiological markers of psychiatric disorders, translational psychiatry, and brain stimulation.

GANESAN VENKATASUBRAMANIAN is currently working as a Professor of psychiatry and consultant psychiatrist with Schizophrenia Clinic and the metabolic clinic in psychiatry with the National Institute of Mental Health and Neurosciences (NIMHANS), Bengaluru, where he is also the Head of the Translational Psychiatry Laboratory, Neurobiology Research Centre. He coordinates the InSTAR Program for Schizophrenia (www.instar-program.org) and the WISER Neuromodulation Program (www.wiser-program.org) at NIMHANS. His research interests include clinical neurobiology and metabolic abnormalities in schizophrenia, transcranial direct current stimulation (tDCS) in psychiatry, and brain imaging in psychiatry.

DEEPESH KUMAR (Member, IEEE) the Ph.D. degree in technology-assisted screening and balance training systems for stroke patients from Indian Institute of Technology (IIT) Gandhinagar, India, in 2018. After his Ph.D., he joined the National University of Singapore (NUS), as a Research Fellow, in July 2018 and was affiliated with the N.1 Institute for Health (previously, SINAPSE). At N.1, he worked on Neuromorphic tactile and vision sensing for neurorobotics and neuroprosthesis applications before moving to IIT (BHU), Varanasi, in July 2021, where he is currently working as an Assistant Professor with the School of Biomedical Engineering. His research interests include neuromorphic tactile and vision sensing, virtual reality-based neurorehabilitation interventions, biomedical signal processing, eye tracking for health monitoring, and machine learning.

A. AMALIN PRINCE (Senior Member, IEEE) received the Ph.D. degree from BITS Pilani, Pilani, India, in 2011. He is currently working as a Professor and the Head of the Department of Electrical and Electronics Engineering, BITS Pilani, Goa, India. His research interests include FPGA-based system design, hardware accelerated data processing, affective computing, and applications of artificial intelligence.

JAC FREDO AGASTINOSE RONICKOM received the degree in electrical and electronics engineering and the Ph.D. degree in neuroinformatics from Anna University, Chennai, in 2015. He continued his postdoctoral research at Brain Development Imaging Laboratory, San Diego State University, USA, in 2016. He joined as a Research Fellow at Nanyang Technological University, Singapore, in 2018 and got Advanced Research Opportunities (AROP) Scholarship at RWTH Aachen, Germany, in 2020. He is currently working as an Associate Professor with the School of Biomedical Engineering, Indian Institute of Technology (BHU), Varanasi, India. His research interests include bio-signal and image processing, biomedical instrumentation, computational neuroscience, and neuroinformatics.