Journal of Information Systems Engineering and Management
2025, 10(20s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Multimodal Biometric Systems: Combining Electroencephalogram and Facial Recognition for Robust Individual Identification and Verification

ShaluVerma[1], Sanjeevindora[2], RohtashDhiman[3]

[1]ResearchScholar,Department of computer science and engineering ,Deenbandhu, ChhotuRam University of Science and Technology,Murthal,Sonepat,Haryana,(India)

[2]Associate Professor, Department of computer science and engineering, Deenbandhu, ChhotRam University ofScienceandTechnology,Murthal,Sonepat,Haryana,(India)

[3]Assistant Professor, Department of electrical engineering, Deenbandhu,ChhotuRam University of Science and Technology,Murthal,Sonepat,Haryana,(India)

CorrespondingEmail:19001901010shalu@dcrustm.org

ARTICLE INFO

Received: 22 Dec 2024

Revised: 30 Jan 2025

Accepted: 18 Feb 2025

ABSTRACT

The primary factor driving the growing interest in novel biometric features is the vulnerability of traditional methods like fingerprint and facial recognition to forgery. This study focuses on a multimodal biometric identification system that integrates data from electroencephalograms (EEG) and facial features. To derive valuable insights from EEG data, we apply signal processing techniques such as filtering, segmentation, and feature extraction, alongside Daubechies-4 (DB4) wavelet analysis with five decomposition levels. The enhanced facial video features include entropy calculations and tracking of facial measurements. Six classifiers Gaussian Naïve Bayes, K-Nearest Neighbour, Random Forest, AdaBoost, Support Vector Machine, and Multilayer Perceptron were trained utilizing the combined EEG and facial data. Findings reveal that AdaBoost and Random Forest emerged as the most effective classifiers for this application, achieving accuracies of 99.87±0.13% and requiring EEG recording times of 2s and 1.5s, respectively, showcasing excellent precision.

Keywords: Identification, Electroencephalogram (EEG), Biometric systems, Wavelet decomposition, Machine learning algorithms

1 INTRODUCTION

In recent years, technology has increasingly permeated various sectors, including banking, retail, online payments, law enforcement, healthcare, border control, and business operations. The security and privacy of these applications have become a paramount concern. Consequently, there is a rising demand for resilient and enduring solutions [1]. Although existing security systems are already operational, there is a persistent push for enhancements in numerous aspects related to security and privacy. Authentication is vital for identifying or verifying an individual, ultimately deciding whether to permit or deny access to resources and services. Its objective is to assess whether individuals should be allowed access or not [1]. Typically, authentication techniques are based on "something you possess," such as certificates, tokens, and smart cards; "something you know," like passwords; or "something you are," such as biometrics [2]. Every authentication approach carries its own set of advantages and drawbacks [3]. Traditionally, a person's identity is verified through items they have, such as a token or card, or through information they know, such as passwords or identification numbers. While conventional methods are prone to loss, theft, or exposure, biometrics, in contrast, rely on characteristics inherently belonging to an individual. These characteristics can include both physiological and behavioral traits [4]. Physiological attributes encompass physical traits of the human body, including the face, eyes, fingerprints, and retina scans, among other things. On the other hand, behavioral biometrics concentrate on the manner in which a user carries out specific activities

[5]. Furthermore, the bio-electrical signals from the brain (EEG), heart (ECG), and muscles (EMG), when combined with human attributes, have garnered significant interest in biometric technology. These signals are distinctive to each individual and fulfill critical criteria (i.e., they demonstrate uniqueness, universality, and liveness detection) [6]. Each person has unique traits, making them different from one another concerning facial features, fingerprints, mouse dynamics, voice dynamics, keystroke dynamics, and signature verification [7]. Many studies have been carried out to pinpoint effective and suitable biometric modalities. A reliable biometric should adhere to four main criteria [8]: 1) Universality: every individual should possess biometric traits; 2) Uniqueness: every individual should have a characteristic that is not duplicated by anyone else; 3) Permanence: the biometric trait should remain constant over time; and 4) Collectability: the biometric trait should be easily obtainable by a sensor and processed efficiently. Additionally, an effective biometric system should demonstrate high accuracy, be user-friendly, and maintain low rates of false negatives and false positives. EEG-based biometric systems are becoming increasingly popular due to their resistance to replication in hazardous environments, while alternative systems like iris and fingerprint recognition can be easily compromised in threatening situations [9]. Recognizing the growing challenges in various areas where security is crucial, significant advancements have been achieved with biometric systems, particularly those utilizing EEG signals [10]. The neuronal activity of the brain generates a variety of signals that are difficult to replicate since each individual's brain signals have unique characteristics [11]. Among the four aforementioned characteristics that enhance biometric systems, EEG seamlessly meets the most essential criterion, which is uniqueness, as it has been noted that each person exhibits a distinct EEG pattern [12]. Despite its resistance to spoofing attacks, EEG still encounters issues concerning collectability and permanence. The permanence of EEG signals collected from the skull [13] is affected by their variability across different sessions and their low signal-to-noise ratio (SNR) [14].This paper aims to address these specific challenges by proposing a novel hybrid paradigm using face images and EEG signals.Facial images are among the easiest modalities to collect and are more convenient than other biometric methods. When combined with EEG signals, they offer greater security. Each individual has distinct facial features paired with unique EEG signals. Consequently, each user generates a specific set of features. Compared to other modalities, facial characteristics show increased permanence due to their generally stable nature. However, one limitation of facial recognition is its reduced resistance to spoofing, as impostors can potentially use the face of a legitimate user. To mitigate this issue, the aim is to bolster security through a multimodal system. The use of deep learning and machine learning in various wearable biomedical applications is becoming increasingly popular. Several EEG biometric systems and face unimodal biometric systems utilizing machine learning have been introduced previously. Implementing a system based on a single trait always carries the risk of system decline. To create a resilient biometric system, a hybrid approach that integrates both modalities is necessary for effective operation. In the past, hybrid systems have been developed that combine human characteristics with EEG signals to enhance system strength.

1.1 Motivation

This investigation focuses on the two distinct characteristics that comprise a multimodal system: brain signals and facial features. Facial recognition is one of the most reliable methods for identifying individuals. These characteristics depend on extracting features from images with low entropy. A more robust biometric system is necessary, as facial recognition alone is susceptible to issues related to aging and masking. The challenge of addressing these vulnerabilities while enhancing security is the primary motivation behind the creation of a multimodal system. By integrating EEG signals with facial features, the system can be strengthened, ensuring data confidentiality, reducing the risk of signal replication, and preventing data breaches. This research aims to improve user identification and verification in a comprehensive manner by utilizing these two distinct attributes. By harnessing both of these traits, the study intends to enhance the overall process of user identification and verification.

1.2 Significance of the Paper

The proposed multimodal framework contributes in the following ways:

1. This scheme presents two innovative contributions. Firstly, it integrates two biometric traits, specifically facial data and EEG signals. This integration is crucial to the approach put forward. Individual traits can be vulnerable to spoofing and attacks. Therefore, combining the EEG signal with facial recognition increases the system's resilience against unauthorized access. Capturing facial images while recording EEG signals will diminish the interference that may arise if alternative traits are utilized. In our opinion, this represents the inaugural hybrid methodology that merges these two traits.

2. The second key contribution of this study emphasizes an improvement in the accuracy of user identification and verification through the amalgamation of these biometric characteristics and the application of various classifiers.

The document is organized as follows: Section 2 offers a review of relevant literature. The proposed methodology is detailed in the Materials and Methods section, which is Section 3. Results and their corresponding discussions are presented in Section 4. Finally, the paper concludes in Section 5.

2 LITERATURE REVIEW

The current body of research on biometrics can be divided into two primary categories: facial recognition systems and electroencephalogram (EEG)-based biometric systems. These systems emphasize the use of either facial recognition or EEG as distinct traits for both identification and verification. Furthermore, several studies investigate the integration of these traits with additional biometric features to develop multimodal systems.

2.1 EEG Based biometric systems

Numerous applications have made use of EEG brain signals, such as sentiment analysis and prediction, across various categories of brain signals. Dustin et al. explored the factors influencing the efficacy of EEG biometrics. The convolutional neural dense connection network (CADCNN) technique, created by Zhang et al., employs a channel attention mechanism to enhance the performance of affective EEG-based individual recognition. In their assessment, they compared this method with other cutting-edge EEG classifiers and models utilized in EEG biometrics studies. Wang et al. introduced a connective graph for EEG biometric identification, which deepens the understanding of human uniqueness by analyzing six aspects, including inter-state stability, computational expense, connectivity metrics, frequency bands, global feature integration, and node centrality. Maiorana's deep learning approach utilizes Siamese convolutional neural networks to extract features for EEG-based biometric verification that remains task-independent. On the other hand, Chen et al. developed a highly secure EEG-based authentication system using RSVP stimuli and dry electrodes. This system employed a knowledge-based method for user validation rather than depending on the extraction of particular intrinsic factors as features from EEG, a common practice in many previous methods. Damasevicius et al. created a biometric cryptosystem based on EEG signals, which they tested using data gathered from 42 participants. Wu et al. devised an innovative identity recognition system that uses EEG with RSVP, incorporating both non-self and self-face images, demonstrating improved accuracy alongside anti-deception features.

2.2 Face Biometric Systems

The identification of human faces in videos has led to the creation of a framework focused on convolutional neural networks (CNN). Ding et al. presented a CNN-based framework designed to address issues related to image blurriness and occlusions. Their proposed model, called the Trunk branch ensemble convolutional neural network, is dedicated to capturing relevant details from facial images. Kharchevnikova et al. aimed to enhance frame quality and developed a lightweight CNN that offers a more efficient frame selection method compared to traditional approaches. Baert et al. formulated a new unsupervised method for assessing face quality to choose images of high quality. This approach prioritizes obtaining high-quality facial crops from video material without revealing the identities of individuals. Kudithalert et al. introduced an innovative extreme learning matching classification technique for verification tasks, named the Siamese extreme learning machine, which effectively handles parallel inputs. Facial images exhibit numerous variations, including changes in pose, expression, and lighting, which can negatively impact the effectiveness of biometric recognition systems. To tackle these issues, Maafiri et al. proposed a novel feature extraction method for face recognition that integrates local binary pattern with wavelet kernel PCA. This strategy aims to extract strong and distinctive features, thereby enhancing performance in face recognition tasks through a non-linear projection method known as RKPCA.

2.3 Multimodal systems

A multimodal biometric system is seen as a more efficient option than an unimodal biometric system since it combines various human attributes or more than two biometric modalities to improve its reliability. Chakladar et al. [32] employed a multimodal Siamese neural network to integrate two different biometric traits—offline signature and EEG signal—for the purpose of user validation. The network captures both spatial and temporal features from the EEG encoder and the image encoder. After the input data is combined into a single feature space, a distance metric is used to evaluate the similarities and differences of the features to generate the verification results. [33] introduces a multimodal deep key-based biometric authentication system that employs EEG signals along with gait signals to tackle the increasing threat of biometric spoofing through devices such as contact lenses, fingerprint films, anti-surveillance masks, and vocoders. Saini et al. [7] presented an innovative method for multimodal user identification by integrating two related traits: signatures and brain signals. Each person's brain shows a unique response while they are signing. Features from both traits are extracted and input into a model that uses a Hidden Markov Model (HMM)-based sequential classifier. In [27], the combination of EEG and ECG is investigated using several classifiers including KNN, ESAVM, and LDA. The two traits are fused at the feature level, followed by a decision-level fusion of the classifiers. Several studies have concentrated on employing EEG and facial recognition systems as unimodal biometric systems. To our knowledge, there has been no research that combines EEG data with video footage of a person’s face. This study suggests a multimodal biometric system that synergizes EEG signal features with facial image features derived from a video of a subject.

3.MATERIALS AND METHODS

3.1 Dataset Description

In this paper, a method is introduced that utilizes a comprehensive dataset to develop a system. This dataset [1][2]was specifically created for recognizing emotions and has been employed in various research efforts. It is a multimodal dataset that comprises recorded electroencephalogram (EEG) and physiological signals from 32 individuals while they were viewing forty-one-minute music videos. Additionally, the dataset includes rating values for valence, dominance, arousal, familiarity, and both liking and disliking of different emotions showcased by participants during the video viewing. A thorough summary of the data will be provided in the subsequent sections.

1. The sampling frequency of the data was reduced to 128 Hz.

2. Artefacts from the electrooculogram (EOG) were removed.

3. A band-pass filter was applied with a frequency range from 4.0 Hz to 45.0 Hz.

4. The data was normalized by averaging it to a common reference point.

5. After removing the baseline 3-s trial, the data was divided into 60-s trials. Most of the researchers have been using this dataset for human emotion classification. However, this dataset is utilized in this work to study EEG-based person identification.

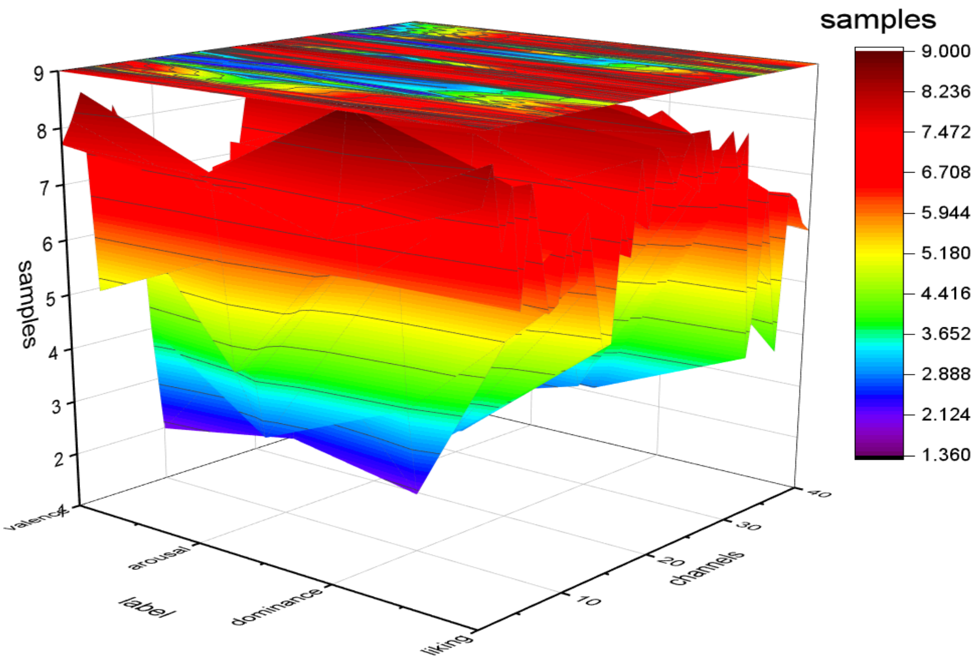


Fig.1: 3D view of EEG data

3.2 Proposed Methodology

This approach identifies prominent characteristics and accurately categorizes them by utilizing brain activity and facial expression information obtained from EEG signal processing and facial video analysis. For time-frequency assessment, EEG data is divided into manageable time segments and converted using wavelet transform into various frequency bands. Analyzing the signals yields statistical measures (mean, variance, skewness, kurtosis), wavelet coefficients, and wavelet energy, which indicate the signal strength across different frequency ranges. The analysis of facial video begins with selecting frames and detecting faces to focus on the area of interest, followed by the extraction of facial features such as basic emotional metrics (e.g., emotional intensity, neutral expression, macro-expressions) and wavelet-derived characteristics for detailed evaluation. Integrating and normalizing the features from both modalities ensures that their scale contributions are balanced. For categorizing emotions, states, and other targets specific to each subject, a machine learning or deep learning approach employs this combined feature vector. The reliability of this classification is ensured by accuracy, precision, and memory. By merging complementary EEG and facial data with wavelet processing and feature integration, this method effectively captures the neurological and expressive elements essential for recognizing emotions, monitoring cognitive status, and analyzing behavior.

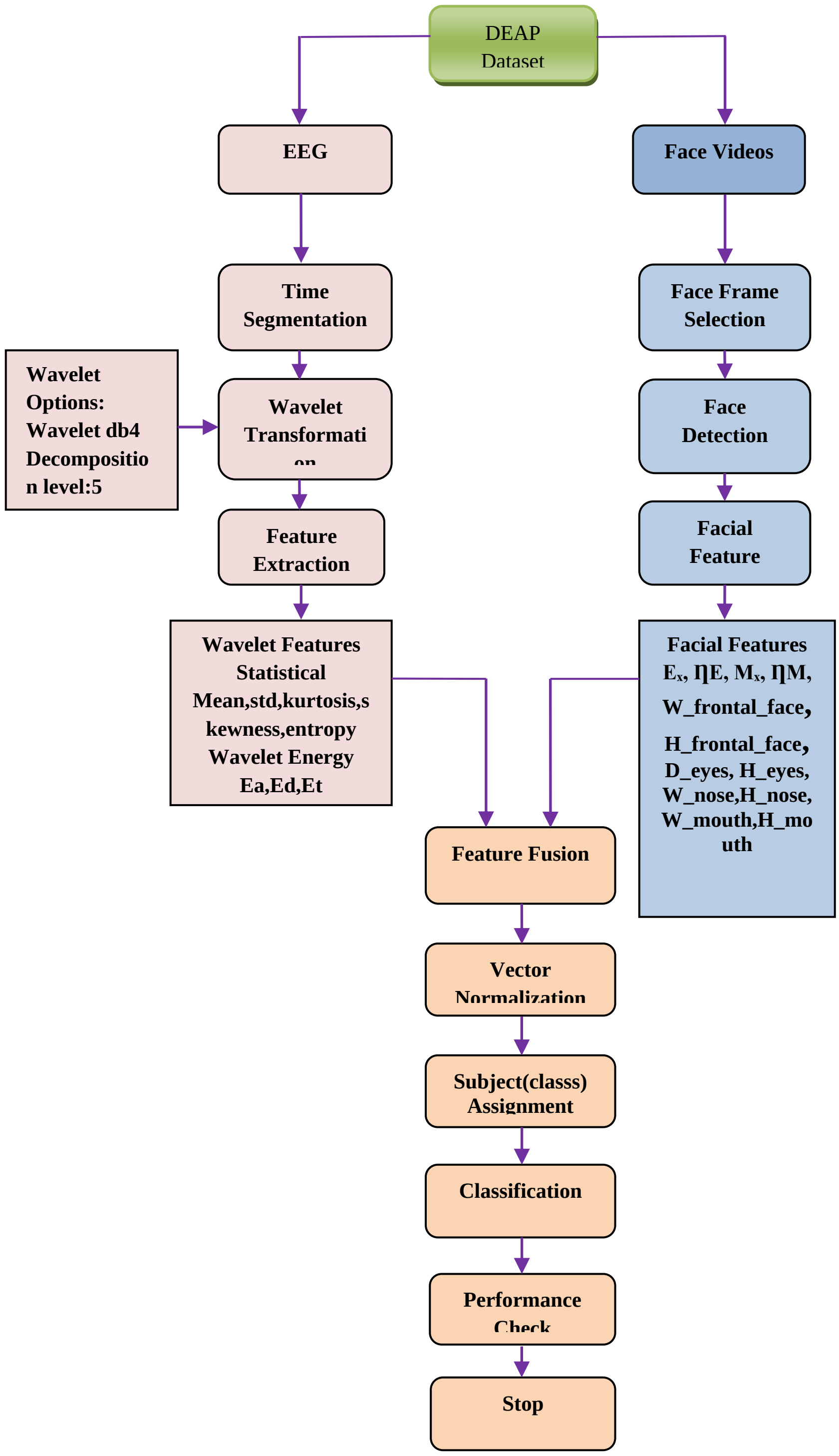


Fig.2: EEG and Face biometric system

3.3 EEG Data Pre-processing

The DEAP dataset comprises records from 32 participants [27]. Each participant's dataset contains EEG information that has been down-sampled to a frequency of 128Hz and has gone through pre-processing. Artifacts, including EOG signals, have been eliminated. A bandpass filter was utilized with a frequency range from 4.0 to 45.0. Before averaging the data and segmenting it into 60-second segments, a 3-second pre-trial baseline was excluded. The original sequence of the trials was rearranged to correspond with the video sequence. Because EEG signals are non-stationary, discrete wavelet transform (DWT) techniques are applied to harness the insightful features of this transformation. Discrete wavelet transforms consist of a series of discrete scales defined by integer multiples of two. Essentially, the initial time-domain signal is divided into two components using a low-pass and a high-pass filter, followed by down-sampling. The coefficients obtained from convolving the original signal with the low-pass filter are called approximation coefficients, while those from the high-pass filter are referred to as detail coefficients. Given that the EEG signal is sampled at 128Hz, the maximum frequency in the signal is 64Hz. Consequently, when the signal is down-sampled by a factor of 2, each node provides a representation that is smaller than the original signal. The EEG signal contains various significant frequency bands, including delta, theta, alpha, beta, and gamma, with specific ranges that may differ among various references. Wavelet analysis is used to illustrate the EEG signal, which entails decomposing the signal into a linear combination of particular wavelets. Various families of wavelets exist, such as the mother wavelet, Morlet, Paul2, Haar, Symlet, and biorthogonal. For this research, the Daubechies-4 wavelet was selected due to its smoothing characteristics, which aid in identifying variations in the signal. The DWT of the function f(t) can be mathematically expressed through the following equation [17].

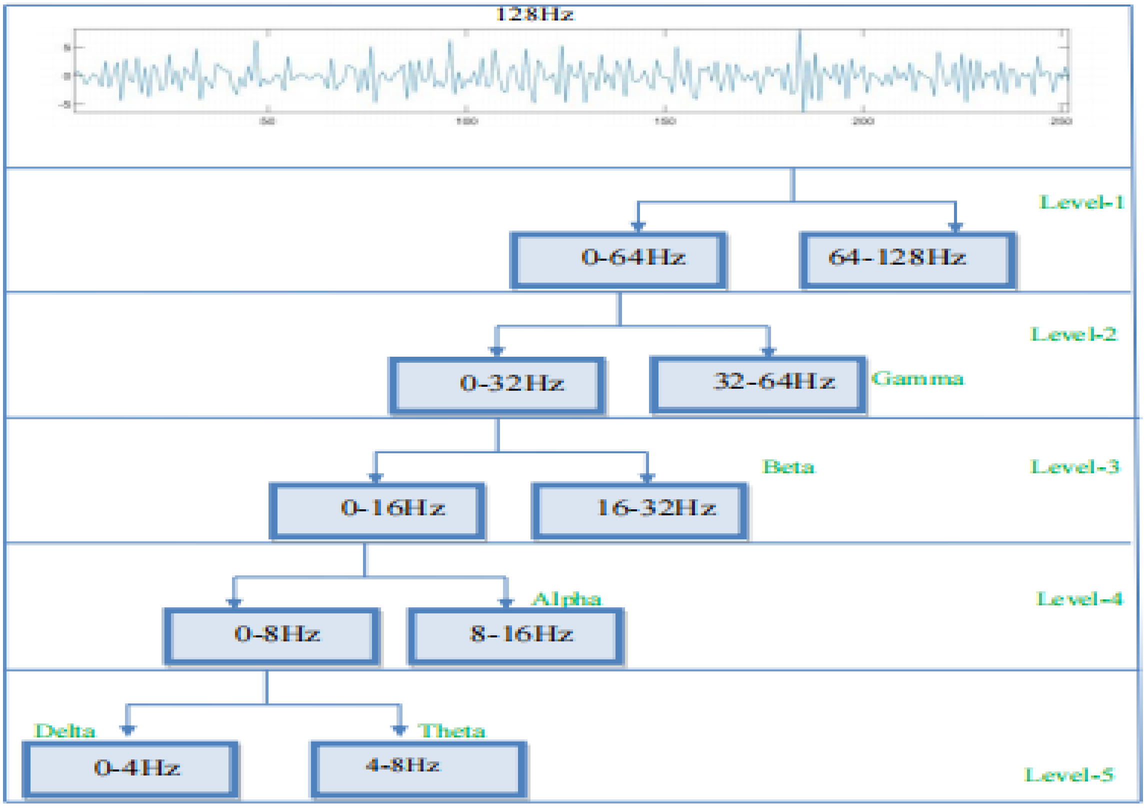


Fig.3: Five levels of decomposition using Discrete wavelet transform[3]

$$DWT_{\Psi}(j,k)=\int_{-\infty}^{\infty} f(t)\Psi_{j,k}(t)dt \quad (1)$$

Five levels of signal decomposition were used in this work to obtain the various frequency bands.

3.4 Face Image Pre-processing

The handling of images includes identifying faces, applying modifications, and implementing smoothing effects. Various subjects are selected for analysis in the images. The first step in the image pre-processing procedure is face detection. This step pinpoints the area of interest in the input images, which will then be passed on to the next stage shown in Figure 4. Numerous studies have employed the Viola-Jones face detection algorithm [29] to recognize human faces. This technique utilizes AdaBoost as a learning method, choosing a small number of key features from a large pool of potential features.

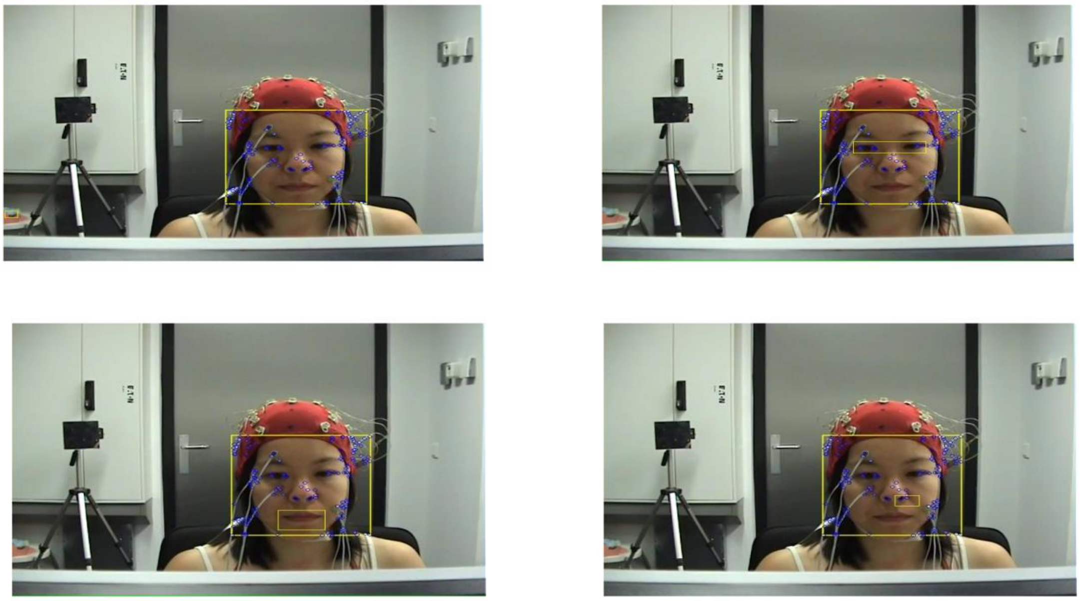


Fig.4: Region of interest includes facial features[2]

3.5 EEG Feature Extraction

An array of research studies has focused on extracting features and classifying EEG signals. A recent study by Carrion-Ojeda has proposed a novel feature called Relative Wavelet Energy, which has demonstrated promise across various EEG applications. This feature has been advantageous in enhancing the analysis and interpretation of EEG signals, offering potential benefits in multiple domains of EEG research. In this investigation, several EEG features are extracted from the signal, such as wavelet energy, relative wavelet energy, mean, standard deviation, skewness, and entropy. Each of these features is computed for every frequency band using distinct equations up to the fifth level of decomposition.

3.6 Face Feature Extraction

Once the pre-processing stage is complete, relevant features can be extracted. As the quality of features significantly influences the accuracy of the system, computer vision has developed numerous techniques. By capturing facial images from videos featuring 22 individuals, both frame entropy and MSE were computed. MSE, or "Mean Squared Error," quantifies the average difference between two signals specifically, between the original facial landmarks and the smoothed landmarks created using a moving average filter. This moving average filter is effective in reducing high-frequency noise within the facial landmark signals, which can lead to a decrease in the MSE between the original and smoothed signals. Theoretically, a larger moving average window will produce a smoother signal, subsequently resulting in a lower MSE. Entropy measures the randomness or unpredictability present within a signal. Regarding facial landmarks, the frame's entropy indicates the extent of variation in the positions of facial features. The moving average filter contributes to the smoothing of high-frequency fluctuations in the facial landmarks, thereby reducing the signal's entropy. Theoretically, a larger moving average window would lead to a smoother signal, which would produce lower entropy. The frame with the lowest entropy is chosen. This study presents a method for identifying facial regions, which are divided into several parts, such as the pair of eyes, mouth, and nose, to extract features from each segment.

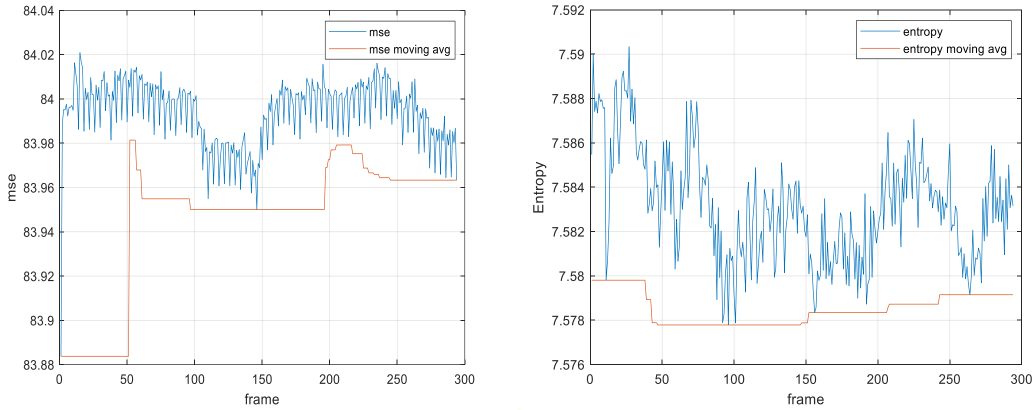


Fig.5: (a) Mean Squared error of signal(b) Frame with minimum entropy

Algorithm 1: Face Frame Selection and Facial Feature Extraction

Require: V (Video of Subject), M(Moving Average Length)

Ensure: FF(Facial Features)

cedure Facial Feature Extraction(V , M)

each frame F ∈ V do

ect face in F

has a face then

5: Calculate and store the entropy of F as Ex

culate and store the MSE (Mean Squared Error) of F as Mx

7: end If

For

9: Calculate moving averages of the entropy values and store them as E¯

10: Calculate moving averages of the MSE values and store them as M¯

11: Select the frame ID with the lowest moving average values of entropy (E¯) and MSE (M¯) and store them as F ID

12: Calculate and return Ex, E¯, Mx, M¯ ,Wfrontal face, Hf rontal face, Deyes, Heyes, Wnose, Hnose, Wmouth, Hmouth

13: end Procedure

4 RESULTS AND DISCUSSION

4.1 User identification result

The EEG signal was segmented into multiple time intervals, specifically 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2, 2.25, and 2.5 seconds. To mimic potential variations that could occur between recordings in a real-world scenario, the segmentation started at random points. The results produced by the classifiers were assessed based on sensitivity, specificity, and accuracy for each of the time segments. The outcomes reveal that when five levels of decomposition are used, starting from the 2-second mark, there are no notable changes in the classifiers' performance. For identifying subjects, a 5-fold cross-validation method was employed, utilizing 80% of the dataset for training while reserving 20% for testing. With a recording duration of 1.5 seconds and five levels of wavelet decomposition, AdaBoost excelled compared to the other classifiers evaluated, achieving sensitivity, specificity, and estimated accuracy of 100%, as illustrated in Table 1. Each set of rows corresponds to a specific time segment with an assigned time value. Within each set, rows represent different algorithms (KNN, SVM, etc.). The values listed in each row indicate the performance metrics (Precision, Recall, etc.) of that specific algorithm on the particular segment at the designated time value. This data facilitates a comparison of how each classification algorithm performs across various datasets and time intervals. By analyzing the recorded performance metrics, one can determine the appropriateness of different algorithms for various time values. The analysis uncovers intriguing insights regarding the performance of these algorithms across datasets with differing time intervals. In the dataset with a time value of 0.50, both KNN and SVM exhibited exceptional performance. They achieved high Precision, Recall, and Accuracy scores, highlighting their effectiveness with this specific dataset. RF consistently produced excellent results across a range of time values, attaining nearly perfect Precision, Recall, and Accuracy. Its Specificity remained consistently high, demonstrating its capability to accurately identify negative cases. KNB and AdaBoost also showed impressive results, especially in cases with time values of 1.25, 1.50, and 1.75, consistently achieving high scores in Precision, Recall, and Accuracy, making them strong candidates for these datasets. MLP maintained a commendable performance, but its scores were slightly lower compared to the other algorithms in this evaluation. It delivered competitive results but displayed minor fluctuations in performance across different datasets and time intervals. It is important to acknowledge that these findings offer useful insights for selecting algorithms but should be viewed in relation to the specific problem and dataset at hand.

4.2.1 UserIdentification Algorithm

Algorithm 2: EEG and Face-Based Subject Identification

Require: DEAP Dataset (EEG and Videos of Subjects), S (SegmentSize), Ψ (Wavelet Function), L (Decomposition Level), M (Moving Average Length)

Ensure: ACC(Accuracy)

1: Procedure SubjectIdentification(DEAP Dataset, S, Ψ, L, M)

2: For each subject n in DEAP Dataset do

3: For each trial t of the subject in the dataset do

4: FF (Facial Features) ← Extract Facial Features using Algorithm 1

5: EEG (Segmented EEG Data) ← Load and Segment EEG data with SegmentSize S

6: Decoded EEG ← Decode the EEG using the specified wavelet function Ψ up to Decomposition Level L

7: AW (Alpha Waves), BW (Beta Waves), and DW (Delta Waves) ← Extract Alpha, Beta, and Delta Waves from the decoded EEG

8: Mean (μ), Standard Deviation (σ), Kurtosis (κ), Skewness (γ), and Entropy (ϵ) ← Extract Mean, Standard Deviation, Kurtosis, Skewness, and Entropy from the EEG

9: RWE (Relative Wavelet Energy's) ← Extract Relative Wavelet Energy's from the EEG

10: EEG Features ← Fuse AW, BW, DW, μ, σ, κ, γ, ϵ, and RWE of the EEG

11: F used Features ← Fuse the Facial Features alongside EEG Features

12: Assign Subject as class y to each feature instance

13: end For

14: end For

15: Use F used Features and class y to train the classifier: Y ← Classifier(Features, y)

16: To obtain predicted subject class Y

17: Evaluate Accuracy (ACC) using true class y and predicted class Y

18: end Procedure

In this algorithm, the EEG signals undergo pre-processing and are integrated with facial features through a series of steps:

1. Retrieve the EEG dataset and choose a specific channel: The code retrieves the EEG data for a particular subject from a designated dataset directory and picks a single channel from the available options.

2. Segment the EEG signals: The code divides the EEG signals into smaller segments of a specified size using a function named "EEG_Segment". This facilitates the extraction of pertinent features from designated time intervals of the EEG signals.

3. Implement wavelet transform: The code conducts a wavelet transform on each EEG segment to extract relevant frequency-domain features. The wavelet transform is executed through the "wavedec" function, which decomposes the EEG signal into various frequency sub-bands.

4. Calculate relative wavelet energy: The code utilizes the coefficients gained from the wavelet transform to determine the relative wavelet energy (RWE) for each frequency sub-band. RWE quantifies the energy present in each sub-band relative to the overall energy of the signal.

5. Derive statistical features: The code also computes statistical features from the EEG signals, such as mean, standard deviation, kurtosis, skewness, and entropy.

6. Standardize the feature vector: The code standardizes the extracted feature vector by scaling all features to a range between 0 and 1. This process helps eliminate any bias resulting from variations in feature scales.

7. Merge EEG and facial features: The normalized feature vector obtained from the EEG signals is combined with the facial features extracted from video frames.

By executing these pre-processing steps, the EEG signals are converted into a collection of relevant features suitable for classifying the subject.

4.2.2 Classification

This research addresses the "closed set" recognition challenge, which pertains to a pattern recognition issue where the categories of the input data are established and predetermined. This implies that the potential classifications of the input data are confined to a specific, limited set of classes, and the objective is to accurately categorize new input into one of these known classes. A frequently employed approach to solve the closed-set recognition problem involves the use of supervised learning algorithms. In these methods, a model is trained on a labeled dataset, where each instance is associated with a class label. During training, the model acquires the correlations between the characteristics of the input data and the respective class labels. Once training is complete, the model can classify new, unseen data by identifying the class label that most closely corresponds to the input data based on the relationships acquired during training. Additionally, Fig. 6 illustrates the confusion matrix for each classifier at the decomposition level 5 using a 1.5-second recording.

Various classifiers apply different criteria for performing classification: The fundamental concept underlying closed-set recognition challenges, which is particularly suited for SVM, is to find a hyperplane that maintains the maximum distance from the nearest data points of each class, known as support vectors. Support vectors are the crucial data points that significantly impact the positioning of the hyperplane that is nearest to them. Moreover, despite having minimal training data, our classifier demonstrates impressive performance and can also manage noisy data. The K-Nearest Neighbour (KNN) technique is simple and effective, functioning by identifying the closest neighbor to a data point and assigning class labels based on the majority voting among those neighbors. In addition to its benefits, such as simplicity and ease of use, KNN does not necessitate any assumptions regarding the data distribution and can accommodate non-linear decision boundaries.

Random Forest (RF) is particularly renowned for its robustness and exceptional accuracy in the presence of outliers. The fundamental concept is that each decision tree tends to overfit the training data; however, aggregating multiple decision trees can mitigate overfitting, resulting in a more resilient model with enhanced generalization performance. The improvement in performance can stem from decreased correlation among the trees due to random feature selection and bootstrapped training data. A popular variation of the naive Bayes algorithm is Gaussian Naive Bayes (GNB). In this approach, it is assumed that the conditional probability distribution of each feature given the class label follows a Gaussian distribution, implying that each feature is expected to conform to a normal distribution. The Gaussian Naive Bayes classification algorithm utilizes Bayes' theorem to compute the posterior probabilities of each class label based on the input features. The algorithm selects the class label with the highest probability as the predicted outcome by evaluating these probabilities.

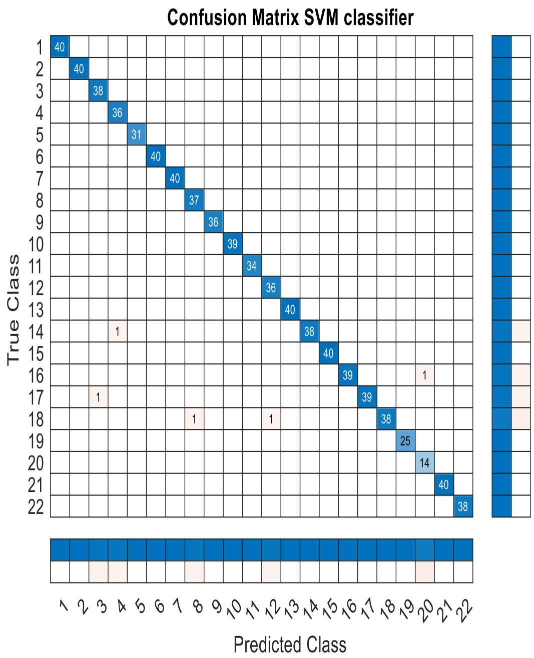
AdaBoost (AB) is a well-known ensemble learning technique. It operates by focusing on the misclassified instances from prior iterations, adjusting the weights at each step of the AdaBoost procedure to iteratively train a weak classifier on a weighted version of the training dataset. The ultimate classifier is formed by aggregating the weak classifiers, each receiving a weight based on its performance during training [34].

The Multilayer Perceptron (MLP) is a type of neural network that is proficient at capturing complex non-linear relationships between input variables and output results. This capability makes MLPs particularly effective for addressing closed-set recognition challenges. MLPs can be trained utilizing labeled training data to uncover underlying patterns and generate accurate predictions for new data points. During the training phase, the weights and biases of the network are adjusted to reduce a cost function like mean squared error or cross-entropy.

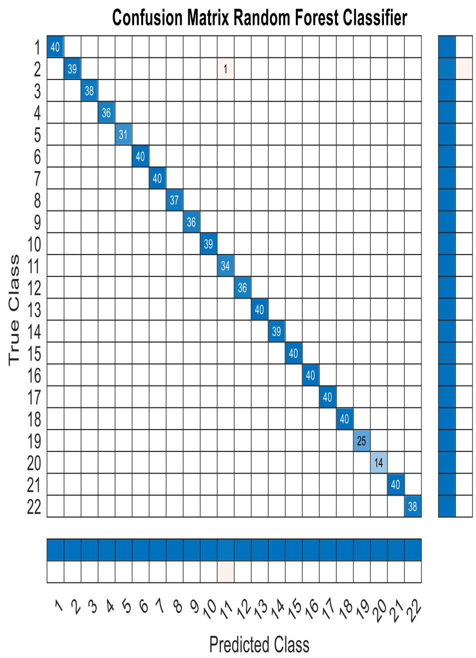
Once the network has been trained, it can be used to classify unseen data points by inputting the feature values into the network and receiving the output from the final output layer [35].

4.3 User verification

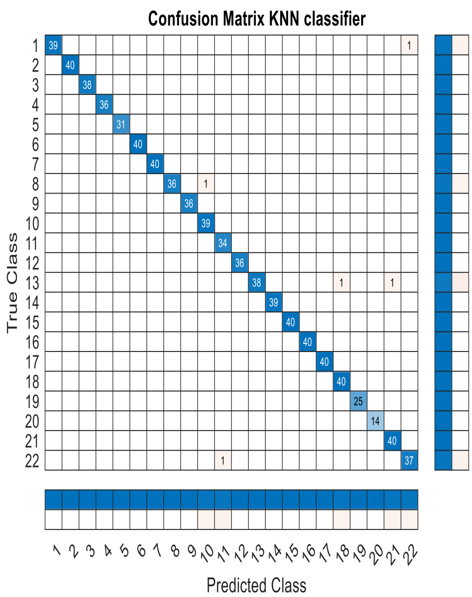
Determine a score d(X,F') that indicates how closely the input sample resembles a group of identified individuals. This can be achieved using a distance metric, such as the Euclidean distance, where d represents the metric and Fi denotes the feature vector for the i-th identified individual. Assess d(X,F') against a predetermined threshold value T. If S(X,F') is less than or equal to T, classify X as not associated with any known individual. Conversely, if S(X,F') exceeds T, classify X as associated with the known person whose feature vector is closest. Specifically, if min(d(F', Fi)) is greater than T, then X does not correspond to any known individual; otherwise, X is associated with the recognized person whose feature vector Fk minimizes the distance, such that d(F', Fk) = min(d(F', Fi)). In this context, the if-statement evaluates whether the shortest distance between the feature vector of the input sample and those of recognized individuals in S surpasses the threshold T. If it does, the input sample is regarded as not linked to any known individual; if it doesn’t, the input sample is classified as belonging to the known person with the feature vector that is closest, determined by the minimum distance to the feature vector of the input sample.



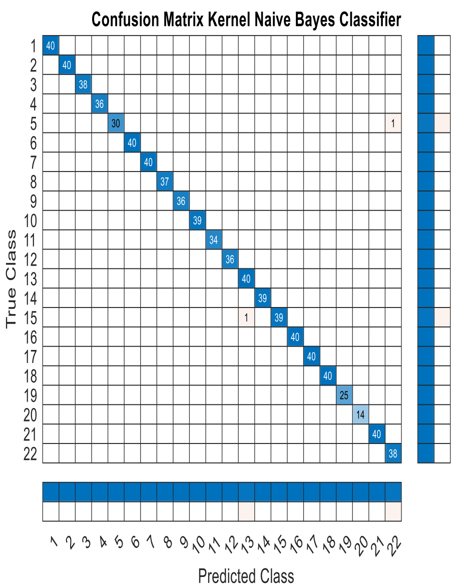
(a) SVM



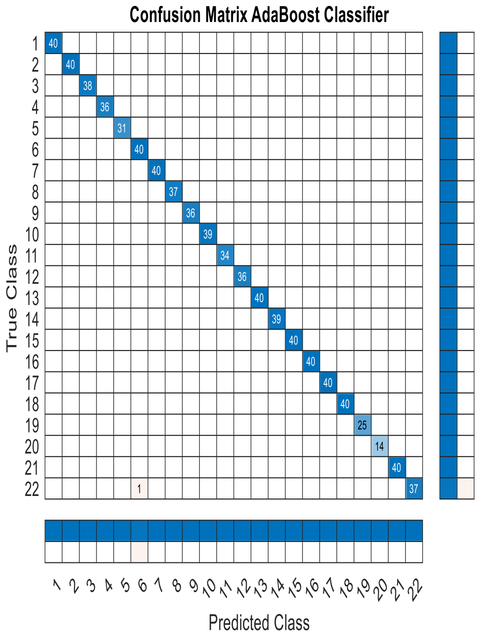
(b) RF



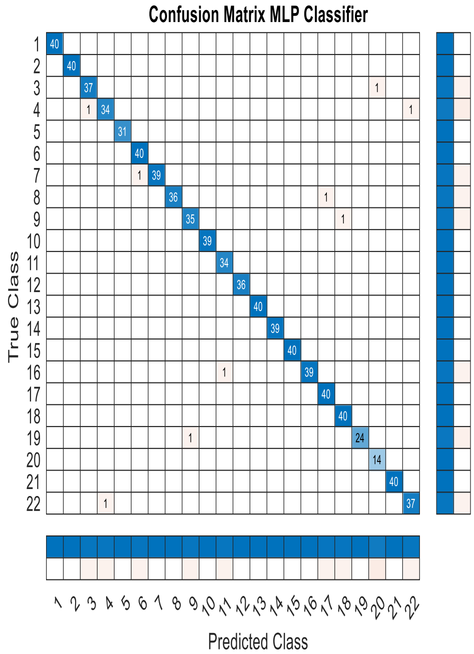
(c)KNN



(d)GNB



(e)AB



(f)MLP

Fig.6: Confusion matrices of the classifiers using 1.5 s of recording

4.4 Results

Each recording consists of 32 EEG channels, 12 peripheral channels, three channels that are not used, and one status channel. The signals in this research were down-sampled from 512 Hz to 128 Hz prior to performing the tests. A band-pass frequency filter was applied to restrict the signal range to between 4 Hz and 45 Hz, after eliminating electrooculogram (EOG) artefacts. Subsequently, the EEG signals were divided into Segment Size windows and underwent wavelet decomposition using the "wavedec" function at a defined number of levels. The wavelet coefficients were preserved, and the detailed coefficients from levels 1 to 5 were extracted for the delta, alpha, beta, and gamma frequency bands. A single-row vector containing the coefficients for the alpha, beta, and delta bands was normalized to ensure a zero mean and unit variance for each component. The energy characteristics of the wavelet coefficients, along with statistical variables such as mean (mu), standard deviation (sigma), kurtosis (Ku), skewness (Sk), and entropy (Ent), were incorporated into this normalized vector. The concatenation also included the topic number. This comprehensive method transformed EEG signals into a feature vector that encompassed wavelet-derived characteristics, statistical features, and facial attributes for further analysis. K-Nearest Neighbours (KNN), Support Vector Machine (SVM), Random Forest, and AdaBoost were trained on the extracted features to establish classification or prediction patterns. In this scenario of multi-class classification, a confusion matrix was utilized to assess the accuracy (Acc) of the models, as well as Macro-averaging Specificity (Sp), precision (Pe), recall (Re), and F1-score (Fs). These metrics indicate the ability of the model to distinguish between classes and its classification performance. The formulas used to compute these metrics are as follows:

$$Acc= \left(\sum_{i=1}^{l} \frac{Tp_i+Tn_i}{Tp_i+Fn_i+Fp_i+Tn_i}\right) / l \quad (2)$$

$$Sp= \left(\sum_{i=1}^{l} \frac{Tn_i}{Tn_i+Fp_i}\right) / l \quad (3)$$

Pe= $\left(\sum_{i=1}^{l} Tp_i\right)/\left(\sum_{i=1}^{l} Tp_i + Fp_i\right)$ (4)

Re= $\frac{\left(\sum_{i=1}^{l} Tp_i\right)}{\left(\sum_{i=1}^{l} Tp_i + Fn_i\right)}$ (5)

Fs=2*(Pe*Re)/(Pe+Re) (6)

Table1: Classification results using combined features with five levels of DWT

Time (sec)	Algorithm	Precision	Recall	Accuracy	Specificity	F_score
	KNN	0.995411	0.992296	0.994987	0.999761	0.993702
	SVM	0.993515	0.994156	0.993734	0.999702	0.993732
	RF	0.998771	0.998864	0.998747	0.99994	0.998802
	KNB	0.998891	0.998864	0.998747	0.99994	0.998863
	AdaBoost	0.998891	0.998864	0.998747	0.99994	0.998863
0.5	MLP	0.993053	0.993122	0.992481	0.999641	0.992982
	KNN	0.994099	0.990907	0.993734	0.999701	0.992313
	SVM	0.994246	0.994199	0.993734	0.999701	0.994146
	RF	0.998737	0.998864	0.998747	0.99994	0.998784
	KNB	0.998891	0.998834	0.998747	0.99994	0.998849
	AdaBoost	0.998864	0.998864	0.998747	0.99994	0.998849
0.75	MLP	0.993593	0.99397	0.993734	0.999702	0.993693
	KNN	0.992831	0.994535	0.994987	0.999762	0.99355
	SVM	0.992645	0.992551	0.992481	0.999642	0.992496
	RF	0.998534	0.998701	0.998747	0.999941	0.998596
	KNB	0.998891	0.998701	0.998747	0.99994	0.99878
	AdaBoost	0.998891	0.995868	0.998747	0.99994	0.997274
1	MLP	0.992057	0.992617	0.992481	0.999643	0.99225
	KNN	0.991649	0.99403	0.993734	0.999703	0.992674
	SVM	0.994344	0.994318	0.993734	0.9997	0.99423
	RF	0.998891	0.998864	0.998747	0.99994	0.998863
	KNB	0.998182	0.998864	0.998747	0.999941	0.998497
	AdaBoost	0.998891	0.998864	0.998747	0.99994	0.998863
1.25	MLP	0.993859	0.994192	0.993734	0.999702	0.993928
	KNN	0.991683	0.994192	0.993734	0.999703	0.992775
	SVM	0.994247	0.994192	0.993734	0.999701	0.994145
1.5	RF	0.998771	0.995868	0.998747	0.99994	0.997213

	KNB	0.995076	0.997727	0.997494	0.999882	0.996298
	AdaBoost	1	1	1	1	1
	MLP	0.992781	0.992997	0.992481	0.999641	0.992793
	KNN	0.99323	0.994127	0.993734	0.999703	0.993537
	SVM	0.994218	0.993308	0.993734	0.999701	0.993691
	RF	0.998864	0.998864	0.998747	0.99994	0.998849
	KNB	0.998891	0.998864	0.998747	0.99994	0.998863
	AdaBoost	0.998891	0.998804	0.998747	0.99994	0.998833
1.7	MLP	0.994708	0.995425	0.994987	0.999762	0.995011
	KNN	0.994127	0.994096	0.993734	0.999701	0.994063
	SVM	0.99397	0.994066	0.993734	0.999701	0.993906
	RF	1	1	1	1	1
	KNB	0.998891	0.998864	0.998747	0.99994	0.998863
	AdaBoost	0.998771	0.998804	0.998747	0.99994	0.998771
2	MLP	0.992436	0.992797	0.992481	0.999642	0.992511
	KNN	0.994246	0.994156	0.993734	0.999701	0.994126
	SVM	0.994071	0.993971	0.993734	0.999701	0.993955
	RF	0.998534	0.998804	0.998747	0.999941	0.998649
	KNB	0.998891	0.998864	0.998747	0.99994	0.998863
2.25	AdaBoost	0.998864	0.998804	0.998747	0.99994	0.998819
	MLP	0.99183	0.993153	0.993734	0.999702	0.992328
	KNN	0.99315	0.994	0.993734	0.999703	0.993458
	SVM	0.993664	0.994192	0.993734	0.999701	0.993887
	RF	0.998891	0.998834	0.998747	0.99994	0.998849
	KNB	0.994949	0.997727	0.997494	0.999883	0.996218
	AdaBoost	0.998891	0.998864	0.998747	0.99994	0.998863
2.5	MLP	0.994035	0.994289	0.993734	0.999701	0.994083

The True Positives (Tp) is the actual positive classes that were categorized as genuine subject True Negatives (Tn) is actual negative class categorized as negative,false positive (Fp)is the actual negative classes that were categorized as positive are imposter subjects, (Fn) are positive classes but categorized as negative. Each class were added, and the True Accuracy (Acc) was computed by dividing the sum by the total number of instances (*l*) in the dataset. In this study, a 5-fold cross-validation approach is employed, wherein the dataset is divided into five subsets or folds. Each fold was sequentially used as a testing set, while the remaining four folds were combined to form the training set. This approach ensures that each fold is used as the testing set once, providing a comprehensive evaluation of the model's performance. Specifically, approximately 20% of the data was allocated for testing in each fold, while the remaining 80% was utilized for training the model. By employing cross-validation, the study aims to assess the generalization ability and performance of the model across multiple independent subsets of the data. Additionally, the data present in each fold is the same across all categorization algorithms.

5.CONCLUSION AND FUTURE SCOPE

This paper introduced a multimodal biometric framework that combines EEG signals with face video data, aiming to enhance user identification and verification. Traditional unimodal EEG-based methods, though harder to forge than facial images, still confront collectability and session variability. By merging EEG features with facial characteristics, recording time requirements (1.5–2.0s) are reduced while system robustness is increased. Six classifiers were tested, showing that AdaBoost and Random Forest offered the most reliable performance, exceeding 99.8% accuracy. Future work may explore deep neural network architectures to determine an optimal balance of EEG recording duration against accuracy. Employing larger and more diverse datasets could further validate generalization. The experiments here were executed in MATLAB on an Intel(R) Core(TM) i7-10870H system with 16 GB RAM. As technology advances, integrating even more biometric traits such as speech and iris could yield further improvements in system security and user convenience.

REFERENCES

[1] O. L. Finnegan et al., "The utility of behavioral biometrics in user authentication and demographic characteristic detection: a scoping review," Syst. Rev., vol. 13, no. 1, pp. 1–17, 2024, doi: 10.1186/s13643-024-02451-1.

[2] W. Alsumari, M. Hussain, L. Alshehri, and H. A. Aboalsamh, "EEG-Based Person Identification and Authentication Using Deep Convolutional Neural Network," Axioms, vol. 12, no. 1, 2023, doi: 10.3390/axioms12010074.

[3] S. Guennouni, A. Mansouri, and A. Ahaitouf, "Biometric Systems and Their Applications," in Vision Impairment and Blindness – What We Know and What We Have to Know, InTech, 2020, pp. 1–12.

[4] A. Riera, A. Soria-Frisch, M. Caparrini, I. Cester, and G. Ruffini, "Multimodal Physiological Biometrics Authentication," [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.108.3421

[5] I. Traore, M. Alshahrani, and M. S. Obaidat, "State of the art and perspectives on traditional and emerging biometrics: A survey," Security and Privacy, vol. e44, 2018.

[6] M. Benouis, L. Mostefai, N. Costen, and M. Regouid, "ECG based biometric identification using one-dimensional local difference pattern," Biomed. Signal Process. Control, p. 102226, 2021.

[7] R. Saini, B. Kaur, P. Singh, and P. Kumar, "A bio-signal based framework to secure mobile devices," J. Netw. Comput. Appl., vol. 89, pp. 62–71, 2017, doi: 10.1016/j.jnca.2017.02.011.

[8] T. Nakamura, V. Goverdovsky, and D. P. Mandic, "In-ear EEG biometrics for feasible and readily collectable real-world person authentication," IEEE Trans. Inf. Forensics Secur., vol. 13, no. 3, pp. 648–661, 2018.

[9] K. Gorur, E. Olmez, Z. Ozer, and O. Cetin, "EEG-Driven Biometric Authentication for Investigation of Fourier Synchrosqueezed Transform-ICA Robust Framework," Arabian J. Sci. Eng., 2023, doi: 10.1007/s13369-023-07798-6.

[10] B. Kaur, D. Singh, and P. P. Roy, "A Novel framework of EEG-based user identification by analyzing music-listening behavior," Multimed. Tools Appl., vol. 76, no. 24, pp. 25581–25602, 2017.

[11] A. K. Jaiswal and H. Banka, "Brain signal based biometric identification using one-dimensional local gradient pattern and artificial neural network," in Proceedings of the International Conference on Smart Computing and Informatics, 2017, pp. 525–536.

[12] Q. Wu, Y. Zeng, C. Zhang, L. Tong, and B. Yan, "An EEG-based person authentication system with open-set capability combining eye blinking signals," Sensors, vol. 18, no. 2, p. 335, 2018.

[13] J. W. Lee, W. K. Lee, and S. Y. Sohn, "Patenting trends in biometric technology of the Big Five patent offices," World Patent Information, vol. 64, p. 102040, 2021.

[14] A. Rahman et al., "Robust biometric system using session invariant multimodal EEG and keystroke dynamics by the ensemble of self-ONNs," Comput. Biol. Med., vol. 150, p. 105238, 2022.

[15] H. Gauba et al., "Prediction of advertisement preference by fusing EEG response and sentiment analysis," Neural Netw., vol. 92, pp. 77–88, 2017.

[16] P. Kumar, R. Saini, P. P. Roy, and D. P. Dogra, "A bio-signal based framework to secure mobile devices," J. Netw. Comput. Appl., vol. 89, pp. 62–71, 2017.

[17] D. Carrión-Ojeda, R. Fonseca-Delgado, and I. Pineda, "Analysis of factors that influence the performance of biometric systems based on EEG signals," Expert Syst. Appl., vol. 165, p. 113967, 2021.

[18] M. Wang, J. Hu, and H. A. Abbass, "BrainPrint: EEG biometric identification based on analyzing brain connectivity graphs," Pattern Recognit., vol. 105, p. 107381, 2020.

[19] Y. Chen et al., "A High-Security EEG-Based Login System with RSVP Stimuli and Dry Electrodes," IEEE Trans. Inf. Forensics Secur., vol. 11, no. 12, pp. 2635–2647, 2016.

[20] R. Damaševičius, R. Maskeliunas, E. Kazanavičius, and M. Woźniak, "Combining Cryptography with EEG Biometrics," Comput. Intell. Neurosci., vol. 2018, 2018.

[21] 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, pp. 1–16, 2018.

[22] C. Ding and D. Tao, "Trunk-Branch Ensemble Convolutional Neural Networks for Video-Based Face Recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 4, pp. 1002–1014, 2018.

[23] A. Kharchevnikova and A. V. Savchenko, "Efficient video face recognition based on frame selection and quality assessment," PeerJ Comput. Sci., vol. 7, p. e391, 2021.

[24] M. Baert, S. Leroux, and P. Simoens, "Intelligent Frame Selection as a Privacy-Friendlier Alternative to Face Recognition," arXiv:2101.07529, 2021.

[25] W. Kudisthalert, K. Pasupa, A. Morales, and J. Fierrez, "SELM: Siamese extreme learning machine with application to face biometrics," Neural Comput. Appl., 2022, doi: 10.1007/s00521-022-07100-z.

[26] F. Khelifi, A. Bir-jmel, and O. Elharrouss, "LWKPCA: A New Robust Method for Face Recognition Under Adverse Conditions," IEEE Access, vol. 10, pp. 64819–64831, 2022.

[27] K. Bashar, "ECG and EEG Based Multimodal Biometrics for Human Identification," in 2018 IEEE Int. Conf. Syst. Man, Cybern., 2018, pp. 4345–4350.

[28] R. Saini et al., "Don't just sign use brain too: A novel multimodal approach for user identification and verification," Inf. Sci., vol. 430–431, pp. 163–178, 2018.

[29] J. Howie, "Cohomology of one-relator products of locally indicable groups," J. London Math. Soc., vol. s2-30, no. 3, pp. 419–430, 1984.

[30] I. S. Koelstra et al., "DEAP: A Database for Emotion Analysis using Physiological Signals," IEEE Trans. Affect. Comput., 2011.

[31] A. Kurani, P. Doshi, A. Vakharia, and M. Shah, "A Comprehensive Comparative Study of Artificial Neural Network (ANN) and Support Vector Machines (SVM) on Stock Forecasting," Ann. Data Sci., vol. 10, no. 1, pp. 183–208, 2023.

[32] X. Xu, X. Geng, Z. Gao, H. Yang, Z. Dai, and H. Zhang, "Optimal Heart Sound Segmentation Algorithm Based on K-Mean Clustering and Wavelet Transform," Appl. Sci., vol. 13, no. 2, 2023.

[33] Q. Abbas, A. Hussain, and A. R. Baig, “CAD-ALZ: A Blockwise Fine-Tuning Strategy on Convolutional Model and Random Forest Classifier for Recognition of Multistage Alzheimer’s Disease,” Diagnostics, vol. 13, no. 1, 2023.

[34] M. T. Ramakrishna, V. K. Venkatesan, I. Izonin, M. Havryliuk, and C. R. Bhat, “Homogeneous Adaboost Ensemble Machine Learning Algorithms with Reduced Entropy on Balanced Data,” Entropy, vol. 25, no. 2, 2023.

[35] M. Hemmat Esfe, F. Amoozadkhalili, and D. Toghraie, “Determining the optimal structure for accurate estimation of the dynamic viscosity of oil-based hybrid nanofluid containing MgO and MWCNTs nanoparticles using multilayer perceptron neural networks with Levenberg-Marquardt Algorithm,” Powder Technol., vol. 415, p. 118085, 2023.