# NEUROSECURE ID: AN EEG BASED BIOMETRIC SYSTEM

**Ms. Ranaware Purva[1], Ms. Lavale Abhilasha[2], Ms. Jagtap Vaishnavi[3], Mr. S. D. Biradar[4]**

Department of Electronics and Telecommunication VPKBIET Baramati
Savitribai Phule Pune University, Maharashtra, India.

**Abstract:**
**Traditional biometric systems like fingerprint, facial recognition, iris, retina, and voice recognition have long been utilized for identification and authentication due to their unique, personal characteristics. These technologies provide several security benefits, ranging from controlling entry to device unlocking. However, they are not without some disadvantages, especially concerning misuse and ethical concerns. The possibility that these biometrics can be compelled to be used for identification without the subject's permission is one of the most urgent problems. In fact, traditional biometrics like fingerprints, irises, and retinas remain usable even after death, which raises serious concerns about privacy and security, as these identifiers can be exploited. This is where our proposed system comes it EEG based biometric system which evaluates the brain's electrical activity to determine identity. EEG uses electrodes applied to the scalp to record electrical impulses produced by brain neurons. And it cannot be forcefully used as the signals vary on different state and conditions, it cannot be forcefully verified.**

**Keywords: EEG(Electroencephalography), Authentication, Machine learning.**

## 1. INTRODUCTION:

Data protection and individual privacy maintenance are more important than ever in today's digital environment, especially as technology develops further. Robust security protocols are needed to protect confidential data and individual possessions. Conventional techniques for identification include passwords, tokens, and biometrics. These identification techniques are not without limitations as passwords can be easily hacked or manipulated, tokens like keys, cards can be stolen or duplicated and biometrics, which use distinctive biological traits like voice, face, and iris identification to provide more secure alternatives, have their own limitations as they can be used in coercive ways, enabling authentication without the explicit consent of the user, which poses significant risks for privacy and misuse of personal data.

To address these concerns and enhance security, advanced biometrics, like EEG (Electroencephalography), has been introduced. EEG-based biometric systems use brainwave patterns for user identification, which are unique to each individual and difficult to replicate.
Brainwave patterns are a more secure alternative than traditional biometrics since they are practically hard to duplicate and can be affected by various situations. This development provides a more advanced layer of privacy protection, which not only increases data protection but also reduces various ethical difficulties. EEG biometrics safeguards personal data which is only used when users provide authorisation to use it.

## I. WHAT ARE BIOMETRICS:

Measurement and assessment techniques based on living beings' statistical techniques are often the focus of the biometrics area. Biometrics is a concept that is closely connected to the process of identifying people using their biological, physical, or behavioral traits [1]. In order to achieve this, certain biometric intrinsic traits are measured about a person: i) Medical measures such as electromyography (EMG), photoplethysmography (PPG), electrocardiography (ECG), or EEG as it is presented in this work are frequently linked to biological biometrics; ii) The physiological Biometrics are naturally occurring, such as the sound of voice, iris appearance, hand geometry, or fingerprint structure; iii) Behavioral biometrics refer to certain patterns or behaviors of an individual, such as writing style, keystrokes, or gait appearance [1]. To use biometric assessments to determine a known individual, a template of the test results must be known prior subject, this may be understood as "having seen each other before". The necessary authentication criteria are those quantifiable aspects of a person. This procedure is applicable to all forms of identification or communication, including machine-to-machine (M2M), human-to-human (H2H), and machine-to-services (M2S). Because biometric information immediately identifies a person, privacy issues may arise. As a result, extremely careful handling and processing of such data are required.

## II. EEG (Electroencephelography):

EEG a n electrophysiological monitoring technique is used to record the dynamics of the human brain's electrical activity. EEG recording is entirely passive, non-invasive, and not very costly in comparison to other methods. EEG readings are often explained in terms of the brainwaves' rhythmic activity. Signal frequencies and amplitudes vary from one condition to another. There are five main frequency bands known to exist which are Delta (δ), Theta (θ), Alpha (α), Beta (β), and Gamma (γ) in increasing order of frequency. The frequency ranges and associated brain states are described in Table I [2]. Experimentally, frequency ranges or bands are selected based on the brain condition under investigation.

| Brainwave Type | Frequency Range (Hz) | Brain State |
|---|---|---|
| Delta | 0.5-4 | Calm Deep Sleep |
| Theta | 4-8 | Relaxation and Light sleep |
| Alpha | 8-12 | at ease, calm, and alert state |
| Beta | 12-30 | Thinking proactively and being vigilant |
| Gamma | 30 and above | Increased perception and cognitive functioning |

Table: Different Frequency Bands

## 2. LITERATURE REVIEW:

Biometric authentication using EEG signals has gained considerable interest due to the inherent uniqueness, complexity, and non-replicability of brainwave patterns. Numerous studies have demonstrated the potential of EEG-based systems to serve as robust biometric identifiers under various cognitive and sensory conditions. Ruiz-Blondet et al. [3] developed an EEG-based authentication system using Rapid Serial Visual Presentation (RSVP) that initiates Event-Related Potentials (ERPs)of 400 distinct images. The study involved 50 participants and used a Normalized Cross-Correlation (NCC) method to compare ERP patterns across trials. The authentication system achieved 100% accuracy, emphasizing the high discriminatory power of ERP features, particularly when paired with personalized visual stimuli.

La Rocca et al. [4] (2014) conducted an experiment on 108 individuals where EEG signals were recorded during both open-eye and closed-eye resting states using a high-density 64-channel EEG system. Features such as Power Spectral Density (PSD) and Coherence were extracted to capture the frequency-domain dynamics and association among all the different areas of the brain. Deploying a distance-based classifier

based on Mahalanobis with match-score fusion techniques, the study achieved 100% classification accuracy, showcasing the potential of resting-state EEG for high-reliability biometric identification.
Palaniappan [5] focused on cognitive task-based EEG biometrics involving six participants performing five distinct mental tasks such as mental arithmetic, geometric figure rotation, and word association. EEG data from six channels were used, and features including Auto-Regressive (AR) coefficients, Spectral Power (SP), and Inter-Hemispheric Linear Complexity (IHLC) were extracted.

Palaniappan et al. [5] implemented an authentication system based on visual stimulation through recognition of hand-drawn shapes. EEG signals were collected from 20 subjects, and features were extracted using the Multiple Signal Classification (MUSIC) algorithm. Two classifiers—k-NN and Elman Neural Network (ENN)—were employed for pattern recognition, resulting in a high identification accuracy of 98%, indicating that visually evoked EEG responses can be reliably used for biometric identification.

Ashby et al. [6] performed EEG-based person identification using five subjects engaged in four different cognitive tasks. The system relied on features obtained in time-frequency domain such as AR coefficients, Spectral Power, Inter-Hemispheric Power Difference (IHPD), and Inter-Hemispheric Linear Complexity (IHLC). A Support Vector Machine (SVM) classifier was implemented and achieved a perfect classification rate of 100%, confirming the effectiveness of cognitive engagement in enhancing EEG signal separability.

Chuang et al. [7] carried out a study using a minimal EEG setup—just one EEG channel—to collect data from 15 individuals performing seven different mental tasks. Cosine similarity was computed between EEG segments to form features, which were then classified using a k-Nearest Neighbor (k-NN) algorithm. Despite the use of a single channel and simple feature extraction method, the system achieved 99% accuracy, demonstrating the potential of low-cost EEG devices for user authentication.

Riera et al. [8] designed a multimodal biometric system combining EEG and ECG signals recorded during a relaxation state. Fast Fourier Transform (FFT) and mutual information-based techniques were used to extract time-frequency and cross-modality features. Although the exact number of subjects was not specified, the study reported an identification accuracy of 98% using Fisher Discriminant Analysis (FDA), indicating the value of integrating physiological signals for enhanced biometric performance.

In a subsequent work, Riera et al. [8] investigated the feasibility of minimal-channel EEG-based authentication during relaxation. Data from 51 participants were recorded using only two EEG channels. Features such as Higuchi Fractal Dimension, entropy, skewness, and standard deviation were extracted to capture the non-linear and statistical characteristics of the signals. With the LDA classifier, a 97% accuracy was achieved, supporting the viability of compact, low-channel systems for scalable EEG biometric authentication.

Jayaratne et al. [9] deployed a mental-task based authentication system where 12 participants were instructed to imagine different four-digit numbers. Common Spatial Patterns (CSP) were used to extract spatial features that distinguish mental representations across users. Classification was carried out using LDA, achieving 97% accuracy. The study illustrates the potential of EEG in recognizing individuals based on internally generated thought patterns without requiring external stimuli.

## 3. METHODOLOGY:

**3.1 Data Acquisition:** The data collection was formed by taking EEG recordings conducted in 12 different experiments from the 20 healthy people. The data includes both auditory stimuli and resting state studies. Both eyes-closed and eyes-opened resting-state EEG signals are included in the data. Six trials make up the section on auditory stimuli: three use in-ear stimuli, while the other three use bone-conducting stimuli. Each example has three stimuli: neutral music, a non-native song, and a native song. The EEG recordings were collected from four electrode channels which are T7, F8, Cz and P4.

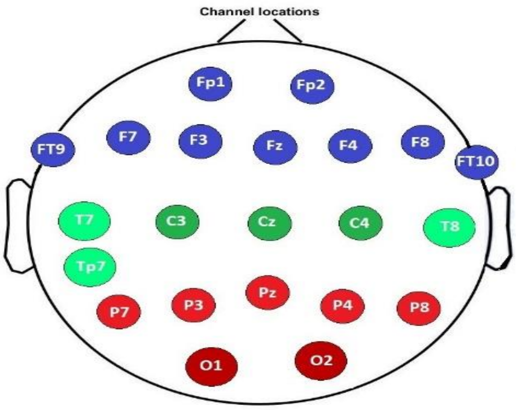


Fig: Channel Electrode Allocation

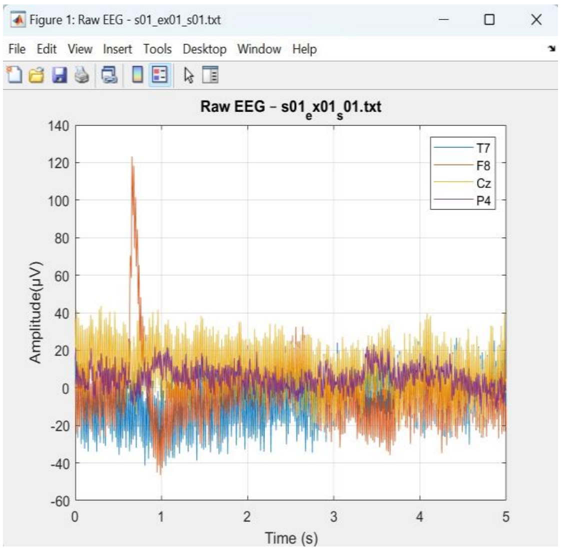


Fig: Raw EEG Signal Obtained From Electrodes- T7, P4, F8 and Cz of The Subject.

**3.2 Data Preprocessing:** In this project, EEG data collected from 20 subjects under six different auditory experimental conditions was pre-processed before classification. The following steps were applied for Pre-Processing:

i. **Filtering:** The EEG dataset was checked for any missing or irregular values, though the dataset was largely clean due to its structured collection from PhysioNet. However, we filtered out any abnormal spikes or noise using bandpass filters to retain only the frequency range relevant to brain activity (e.g., 0.5–50 Hz).

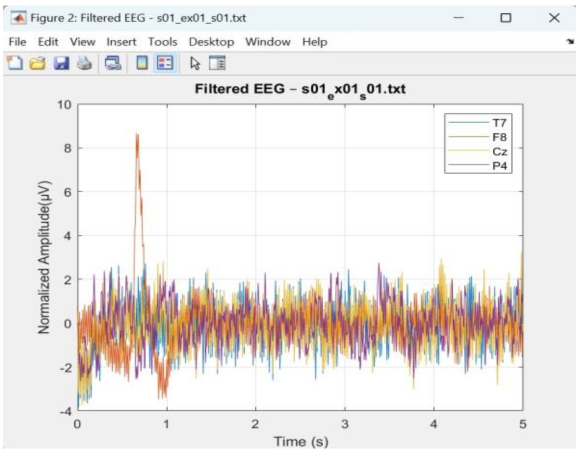


Fig: Filtered EEG Signal of The Subject

ii. **Feature Extraction:** From the four EEG channels, we extracted a variety of time-domain and frequency-domain features, including:

- Mean, Standard Deviation, Skewness, Kurtosis and Variance.

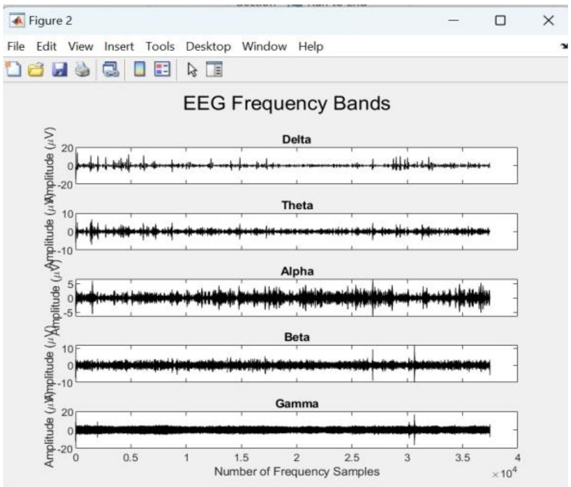


Fig: Extracted Frequency Bands of The Subject

- Power Spectral Density, Bandpower, and Wavelet Transform Coefficients

These features were essential in capturing both temporal and spectral characteristics of the EEG signal relevant to individual biometric patterns.

iii. **Normalization:** The features were normalized using z-score method. Data is converted to a conventional utilizing Z-score normalization to create a normal distribution with a mean of zero ‘0’ and a standard deviation of one ‘1’, also known as standardization.

**3.3 Model Development:**

The goal of the model was to classify EEG signals based on subject ID, effectively serving as a biometric authentication mechanism. We used traditional machine learning models as follows:

**Data Splitting:** The dataset was further split into 20% for testing and 80% for training.

**Classifying Models:** Multiple models were developed and evaluated:

- **Support Vector Machine (SVM)** - with optimized kernel and parameters
- **Random Forest** – evaluated for performance comparison

Models were trained using the extracted feature matrix and corresponding subject labels.

- **k-NN: k-Nearest Neighbors (k-NN)** is a easy, machine learning procedure that is instance-based and utilized for regression and classification. It use the majority label of the 'k' closest neighbors of a data point in the feature space to categorize it. k-NN is non-parametric and relies heavily on distance metrics like Euclidean distance for decision-making.
- **Linear Discriminant Analysis (LDA)** - selected due to to its capacity to maximize class separability and minimize dimensionality, which makes it appropriate for biometric data with well defined boundaries.

i. **Feature Selection:** To enhance accuracy, Wavelet Transform was utilized to identify the most informative features for each classifier, reducing dimensionality and improving model performance.

**3.4 Model Training and Evaluation:**

After extracting and selecting the unique key features of EEG signal, the dataset was used to train different models of machine learning for biometric classification. The aim was to accurately identify individuals based on their unique EEG patterns. To check model performance, the following metrics were computed on the test set:

- Accuracy: In comparison to all occurrences, accuracy is the proportion of correctly anticipated cases (including true positives and true negatives).
- Precision: The amount of true positive predictions among all of the model's positive predictions is known as precision.

- Recall: The proportion of actual positive cases that were accurately anticipated is known as recall.
- F1-score: The F1-score is the average harmonic of recall and accuracy. It is particularly useful when classes are unbalanced since it balances the two metrics.
- Confusion Matrix – to visualize correct vs. misclassified predictions per subject.
- ROC Curve: The Receiver Operating Characteristic Curve is the plotting the True Positive Rate (Recall) versus the False Positive Rate (FPR) at different threshold settings is done via the Receiver Operating Characteristic Curve, or ROC curve.
- Cross-Validation Accuracy: To guarantee the model's stability across several data splits, k-fold cross-validation was employed.

The best-performing model was later deployed into a user-friendly MATLAB GUI, allowing real-time biometric verification using EEG data.

**3.5 Results:**

Predicted authentication outcomes were compared with actual subject identities, and the results were visualized using MATLAB plots and confusion matrices to assess model performance. Graphs demonstrated how well the model distinguished between genuine users and imposters based on EEG features. Numerous the respondents classifier performance was examined utilising metrics which involves accuracy, precision, recall, and F1-score.

**3.6 EEG Channel-Wise Feature Analysis:**

This section presents the visual and performance-based analysis of EEG signals from selected subjects using wavelet-based feature extraction and classification using SVM, Random Forest (RF), k-NN and LDA.

1. EEG Signal with Wavelet Decomposition

The wavelet transform successfully extracted time-frequency features from Subject’s 4-channel EEG signals. These features captured both low- and high-frequency information, providing distinctive biometric signatures. Among classifiers, SVM achieved the highest classification accuracy for Subjects, indicating clear and separable patterns in the wavelet domain.

2. Random Forest consistently demonstrated strong classification performance across subjects, particularly in cases with higher inter-trial variability. Its ensemble nature allows it to record intricate, non-linear interactions inside the wavelet-extracted EEG features. The model's robustness to noise and overfitting makes it well-suited for biometric systems based on EEG data, where subtle patterns must be identified amid inherent signal variability.

3. k-NN showed competitive accuracy, especially when the feature space retained local structural patterns from wavelet decomposition. Its performance indicates that subject-specific EEG traits are well-clustered in the transformed feature domain. However, its sensitivity to feature scaling and noise means its success depends on effective preprocessing and selection of the optimal 'k' value.

4. LDA provided stable performance on smaller, cleaner EEG feature subsets. While it assumes linear separability, its computational efficiency and interpretability made it suitable for baseline comparisons. LDA was particularly effective when intra-subject consistency was high and EEG noise was minimal, highlighting its potential in well-controlled EEG acquisition scenarios.

**3.7 Classifier Performance Comparison**

The following table provides an overview of the performance indicators of the three classifiers—SVM, Random Forest, k-NN and LDA on the wavelet-extracted EEG features.

| **SrNo.** | **Classifier** | **Accuracy (%)** | **Recommendation** |
|---|---|---|---|
| 1 | SVM | 100 | Highly Recommended |
| 2 | Random Forest | 97.92 | Recommended |

| | | | |
|---|---|---|---|
| 3 | k-NN | 97.92 | Recommended |
| 4 | LDA | 93.75 | Recommended |

Table: Classifier Performance Comparison on Wavelet Features

The table and corresponding figure shows that SVM outperformed the others slightly in overall metrics, but Random Forest and k-NN also showed consistently strong results, especially in noisy or variable cases. LDA performed well in subjects with stable EEG patterns.

**3.8: Training and Testing Evaluation Results using Confusion Matrix:**

**1. Support Vector Machine:**

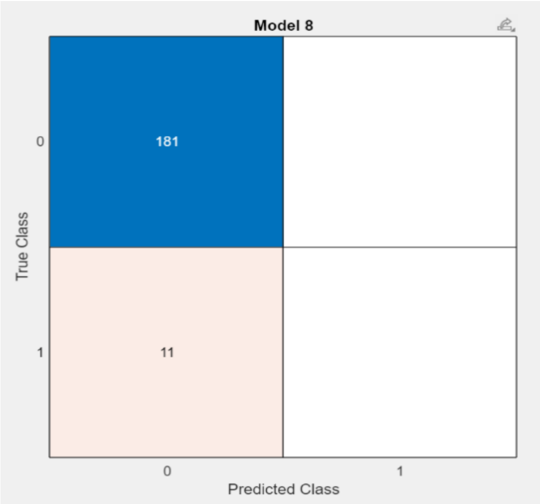


Fig: SVM Confusion Matrix

The execution of SVM is competitive, especially considering the classification's complexity. Compared to simple Random Forest or even k-NN, the decision boundaries are improved by non-linear separation made possible by the employment of a well-tuned kernel function. The model is stable and well-balanced.

**2. Random Forest:**


Fig: Random Forest Confusion Matrix

Random Forest's ensemble structure offers resilience against overfitting when managing complex boundaries. Extremely dependable and applicable in this situation.

**3. k-Nearest Neighbor:**

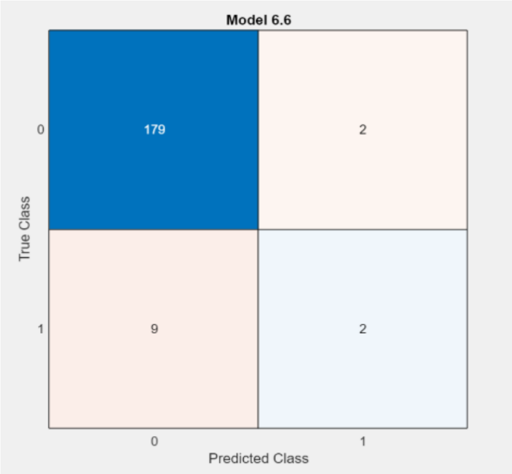


Fig: k-NN Confusion Matrix

The k-NN model's confusion matrix shows a decent but less consistent performance compared to the SVM model. The k-NN classifier's accuracy is moderate, and its effectiveness may be enhanced with better feature scaling or optimized choice of 'k'. However, it lacks the robustness and adaptability seen in the Random Forest model, limiting its effectiveness for high-precision biometric applications.

**4. Linear Discriminant Analysis:**

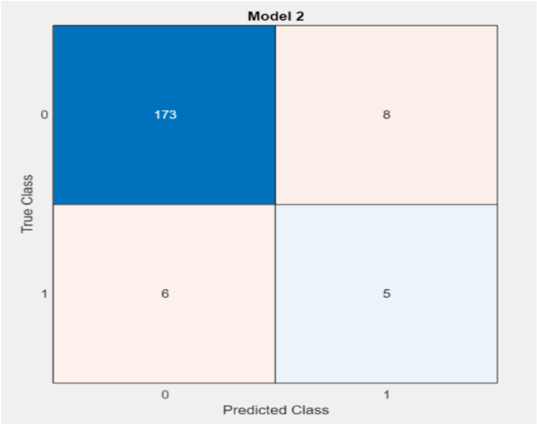


Fig: LDA Confusion Matrix

The confusion matrix for LDA reveals a relatively lower performance in contrast to alternative approaches. This suggests that the model struggles to find a linear boundary that separates the genuine and imposter users effectively. Consequently, the model fails to record the data's underlying structure, making it the least suitable among the classifiers evaluated for the purpose of biometric system.

**RESULTS AND DISCUSSION:**

The EEG-based biometric system demonstrated high accuracy and robustness using wavelet-based feature extraction. The selected classifiers SVM, Random Forest, k-NN and LDA showed reliable performance, with SVM achieving the highest overall accuracy of 100 %.

Compared to traditional methods that rely on statistical features alone, wavelet transforms provided a richer feature set by Obtaining both time and frequency domain characteristics. The classification model performance metrics (precision, recall, and F1-score) confirmed the system's ability to distinguish subjects effectively using their EEG signatures.

The system's flexibility in identifying genuine users versus imposters makes it suitable for real-time access control and secure biometric applications.

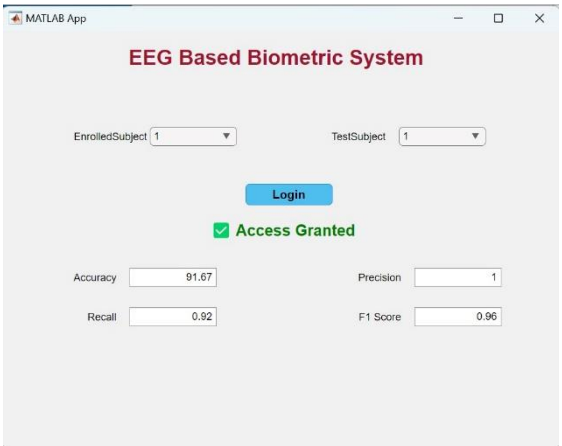


Fig: EEG Authentication Result for Same Subjects

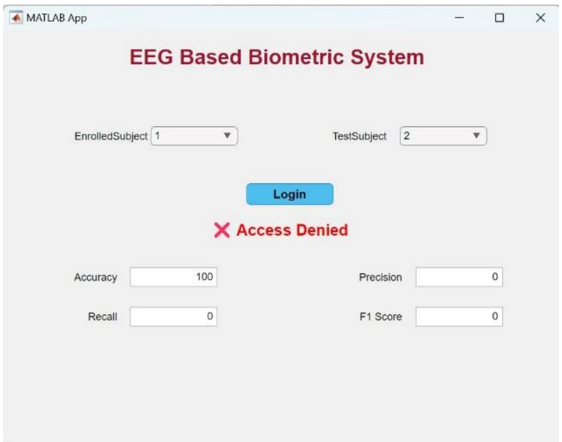


Fig: EEG Authentication Result for Different Subjects

## 5. CONCLUSION:

This project demonstrates the effectiveness of combining wavelet-based extraction of features with machine learning methods like, SVM, Random Forest, k-NN and LDA for authentication using EEG signals. The system successfully distinguishes between enrolled subjects and imposters with high accuracy, proving its potential as a secure and intelligent access control method.

Future work will focus on:

- Real-time signal acquisition and classification,
- Incorporating advanced deep learning models,
- Expanding the dataset for cross-subject and cross-session variability handling,
- Testing under cognitive tasks and varying emotional states to ensure robustness.

## REFERENCES:

[1] Jan Herbst, Jan Petershans, Matthias R¨ ub, Christoph Lipps, Ann-Kathrin Beck, Joana C. Carmo, "Inception based Deep Learning: Biometric Identification using Electroencephalography (EEG)", in IEEE International Symposium on Networks, Computers and Communications (ISNCC), 2023.

[2] Isuru Jayarathne,Michael Cohen and Senaka Amarakeerthi, "Survey of EEG-Based Biometric Authentication",2017.

[3] Hanbyul Jang, Sangin Park, Jincheol Woo, Jihyeon Ha, Laehyun Kim, "Authentication System Based on Event-related Potentials Using AR Glasses", 11th International Winter Conference on Brain-Computer Interface (BCI), 2023.

[4] F. La Rocca, D. Campisi, J. C. La Foresta, G. Scarano, and A. Tistarelli, "Human Brain Distinctiveness Based on EEG Spectral Coherence Connectivity," in IEEE Transactions on Information Forensics and Security, vol. 9, no. 5, pp. 782–793, 2014.
[5] R. Palaniappan and D. P. Mandic, "Biometrics from Brain Electrical Activity: A Machine Learning Approach," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 4, pp. 738–742, 2007.
[6] Amit Bhatia, Jacob Vogelstein, Francesco Tenore and Corey Ashby. "Low-Cost Electroencephalogram (EEG)- Based Authentication." 1999.
[7] Bru. Quintela, J. P. da Silva Cunha and A. Zquete, "Biometric authentication using brain responses to visual stimuli.", 2010.
[8] A. Riera, P. García, and M. Dunne, "Feasibility of EEG-Based Biometric Authentication with Two-Channel Systems During Resting State," in Sensors, vol. 21, no. 14, pp. 4750, 2021.
[9] D. Jayaratne, C. H. Chuang, and C. T. Lin, "Cognitive EEG Biometric Modeling for User Recognition During Imagined Activities," in IEEE Transactions on Human-Machine Systems, vol. 47, no. 6, pp. 701–712, 2017.
[10] H. Mou, Z. Pei, W. Yang, F. Li, S. Zhang and X. Wu, "An Overview of Brain Fingerprint Identification Based on Various Neuroimaging Technologies," in IEEE Transactions on Cognitive and Developmental Systems, 2024.