Universidad Zaragoza
1542

# Master's Thesis

# Biometric identification and authentication based on EEG signals


Author

Patricia García Bautista

Supervisor

Luis Montesano del Campo


Master in Robotics, Graphics and Computer Vision

ESCUELA DE INGENIERÍA Y ARQUITECTURA

2026

# Acknowledgements

I would like to express my gratitude to all the members of the Bitbrain team for their guidance throughout the development process of this project, offering their advice and experience. In particular, I am especially grateful to Luis Montesano and Alejandro Artal, whose support has been fundamental, giving me the opportunity to gain knowledge and make significant progress on this research project.

My gratitude is also extended to my family and friends, who have always placed their trust in me and have provided me with their invaluable support, encouragement and patience throughout this period.

# Abstract

Biometric systems have traditionally relied on external physiological and behavioural traits, such as fingerprints, face, voice, or iris patterns, which can have limitations related to spoofing, permanence, and privacy. In this context, electroencephalography (EEG) has emerged as a promising alternative, as brain signals are inherently individual and difficult to replicate. However, the feasibility, robustness, and privacy implications of EEG-based biometric systems remain open research challenges.

The main objective of this project is to study the feasibility of biometric identification and authentication using EEG signals, a concept commonly referred to as Brainprint, with a focus on inter-session stability, task generalization, scalability, and privacy. To this end, a complete EEG-based biometric pipeline was designed and implemented in Python. Moreover, several machine learning models were explored, including classical approaches, such as support vector machines, as well as convolutional neural networks.

The proposed framework was evaluated using a wide variety of EEG datasets, with publicly available resources and private datasets provided by Bitbrain Technologies. These datasets cover diverse recording conditions and include different numbers of channels, sessions, tasks, and population sizes ranging from tens to several thousand subjects. This experimental design enables a systematic analysis of session variability, task dependency, sensor configuration, and population scale in EEG-based biometric systems.

In addition to biometric performance, this project investigates privacy-related aspects of EEG data. Several signal anonymisation strategies were explored, including Gaussian noise injection, PCA-based obfuscation, and frequency-selective perturbations applied directly to the EEG signals. These methods were evaluated not only in terms of their effect on identification and authentication, but also with respect to their impact on non-biometric EEG-based tasks such as resting-state detection, eyes open/closed classification, and sleep stage classification.

Overall, this work provides a methodological and experimental study of EEG-based biometric systems, offering a deeper understanding of brainprints, their practical limitations and the trade-offs between biometric utility and privacy preservation.

# Index

# Chapter 1

# Introduction

In recent decades, reliable methods for identifying and authenticating individuals have become essential for secure access control, personalisation of services, and protection of sensitive information (Figure 1.1). The conventional authentication methods, including passwords, physical tokens and identification cards, are susceptible to loss, theft and duplication. These limitations have motivated the development of biometric technologies, which aim to exploit stable, individual-specific characteristics for person recognition, including physiological and behavioural traits.

Several biometric modalities have been investigated and deployed in practical applications. Some of them, such as fingerprints and iris patterns, have reached a high level of maturity and are now commonly used in commercial systems. However, as detailed in Section 1.1, these methods present certain limitations, including vulnerability to spoofing, sensitivity to environmental conditions, and potential issues with user consistency [1, 2]. For this reason, the exploration of alternative biometric sources, such as those derived from the human nervous system, is necessary, as these may offer inherent advantages in terms of security and revocability.


*Figure 1.1: Commonly used traditional and biometric recognition methods.*

Among biosignals, electroencephalography (EEG) appears to be a potential candidate for biometric recognition [3], also referred to as Brainprint. EEG is a non-invasive technique that measures electrical brain activity, producing signals that are linked to the neural anatomy and function of each individual. The internal generation of EEG signals makes them difficult to capture or replicate, addressing some key weaknesses of traditional biometrics. However, EEG also presents some challenges, primarily related to signal variability across sessions and its sensitivity to cognitive state and recording conditions [4]. A comparative analysis of the advantages and disadvantages of EEG relative to other modalities is provided in Section 1.1.

In addition to the technical challenges, the utilization of EEG as a biometric modality raises particularly sensitive privacy concerns. As a biomedical signal, EEG may contain additional information about an individual's cognitive state or health [5, 6], which increases the risk associated with its storage and reuse, and calls for enhanced protection measures due to the potential impact of neural data on fundamental rights and personal autonomy [7]. In scenarios where EEG signals are shared for research purposes or for the development of alternative applications, it is essential to ensure that this data does not allow for the direct identification of subjects. This motivates the investigation of signal anonymisation strategies that can selectively reduce identity-related information while preserving the utility of EEG data for other analytical tasks.

Taking these considerations into account, this project has been developed in collaboration with Bitbrain Technologies, a neurotechnology company specialized in the acquisition and analysis of neurophysiological signals for research and applied applications. This collaboration provides a realistic framework in which EEG-based biometric identification, authentication, and anonymisation techniques can be studied using real-world datasets. Bitbrain has previously contributed to EEG-based solutions in areas such as cognitive monitoring, neuroscience research, and sleep analysis, enabling the evaluation of biometric and privacy-preserving approaches under conditions relevant to industrial deployment.

The following sections provide the necessary background and context for this work. Section 1.1 introduces the principles of biometric identification and authentication and reviews existing biometric modalities, including an overview of their main characteristics and limitations, with particular emphasis on the capabilities and challenges of EEG-based systems. Thereafter, Section 1.2 presents the concepts of electroencephalography that are relevant to biometric and machine learning applications.

## 1.1 Biometric identification and authentication

Biometric identification and authentication refer to the set of technologies and methodologies that use measurable physiological or behavioural traits to recognise

or verify the identity of an individual. In an *identification* scenario, the system performs a one-to-many comparison, determining the identity of an unknown individual by matching their biometric data against a group of enrolled subjects. In contrast, *authentication* corresponds to a one-to-one verification process, in which the system validates whether the provided biometric data matches a claimed identity. These approaches depend on distinctive features that vary minimally within individuals but maximally across them.

Biometric modalities are commonly classified into two main categories:

- **Physiological traits**, which are derived from individual physical characteristics, such as fingerprints, facial structure, or iris patterns. These methods have demonstrated a high degree of discriminatory capability and have been extensively adopted within commercial systems.
- **Behavioural traits**, which are based on behavioural patterns, including gait, voice, or keystroke dynamics. These are distinguished by their relative ease of acquisition.

Table 1.1 provides a summary of representative ranges of Equal Error Rate (EER), which is a commonly used evaluation metric that corresponds to the operating point at which the false acceptance rate (FAR) equals the false rejection rate (FRR). Such values provide an indicative comparison of their relative performance under controlled conditions, where lower EER values indicate better overall system performance.

| **Modality** | **Type** | **EER** |
|---|---|---|
| Iris | Physiological | $< 0.01$–$0.1\%$ |
| Fingerprint | Physiological | 0.1–2.0% |
| Face recognition | Physiological | 1.0–5.0% |
| Voice | Behavioural | 2.0–10.0% |

*Table 1.1: Equal Error Rate (EER) of common biometric modalities reported in the literature [8].*

The efficacy of these traditional biometric systems has been demonstrated in many controlled contexts; however, they are also subject to limitations [9]. From a security point of view, physiological biometric data such as fingerprints or facial features are vulnerable to presentation attacks. This is due to the fact that the biometric characteristic is externally observable and can be captured without the user's knowledge. Furthermore, the majority of traditional biometric characteristics are non-revocable: once compromised, they cannot be replaced. This issue is partially addressed by behavioural biometrics; however, these are often characterized by reduced stability over time due to changes in user behaviour or environmental conditions.

In contrast, electroencephalography (EEG) is a technique that captures internally generated neural activity, reflecting both anatomical and functional properties of the brain. This fundamental difference provides EEG-based biometrics with an inherent

resistance to spoofing, as the signal cannot be trivially acquired or replicated without the active participation of the subject. Furthermore, due to the dynamic nature of neural activity, it is possible for EEG patterns to be adaptively adjusted or updated with different recording methods if they are compromised. Moreover, it is permissible to combine EEG with other biometric traits with a view to enhancing system robustness.

Despite these advantages, EEG-based biometric systems also face some challenges. EEG signals are highly non-stationary and sensitive to factors such as electrode placement, cognitive state, fatigue, and recording conditions, which can negatively affect inter-session robustness. Moreover, EEG acquisition typically requires careful sensor placement, artifact mitigation procedures, and the use of devices with high signal-to-noise ratio (SNR), resulting in higher acquisition complexity compared to conventional biometric systems.

## 1.2 Electroencephalography (EEG)

Electroencephalography (EEG) is a non-invasive measurement of the electrical activity of the brain. Electrodes placed on the scalp record voltage potentials resulting from current flow in and around neurons [10]. The recorded output, known as an electroencephalogram, consists of time-varying waveforms that reflect neural dynamics across different temporal scales.

### 1.2.1 Frequency bands

EEG signals span a broad range of frequencies, traditionally categorized into distinct bands [11], as can be observed in Figure 1.2. Each band is associated with different functional and physiological brain states:

- Delta ($0.5 - 4$ Hz): The slowest EEG waves, typically predominant during deep sleep.
- Theta ($4 - 8$ Hz): Often associated with drowsiness, early stages of sleep, and certain cognitive processes.
- Alpha ($8 - 13$ Hz): Prominent during relaxed wakefulness, especially with eyes closed and reduced cognitive demand.
- Beta ($13 - 30$ Hz): Linked to active thinking, alertness, motor behaviour, and concentration.
- Gamma ($> 30$ Hz): Involved in higher cognitive functions, intense focus and information integration.

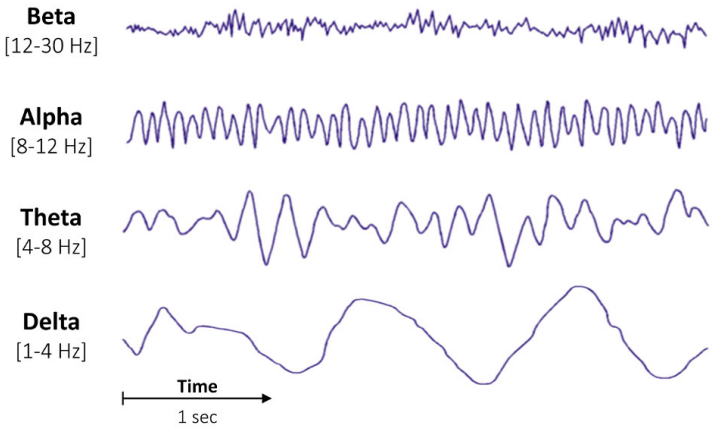



*Figure 1.2: Frequency bands.*

### 1.2.2 Technical features

EEG systems are composed of various technical components that can influence signal quality, usability, and applicability to both research and clinical contexts. These characteristics directly affect the accuracy with which neural activity can be recorded, as well as the efficacy of the system under different experimental conditions. The technical features of an EEG system can be divided into three main areas [12]:

- **The Sensor Layer.** It is the interface between the EEG system and the subject's scalp. Electrode channels on EEG devices typically range from 8 to 128, with an increased number of channels offering enhanced spatial coverage but also a higher setup complexity and greater computational demand. Electrode placement usually follows standardized configurations such as the International 10–20 system, ensuring reproducibility and comparability across studies. Depending on the acquisition setup, electrodes may rely on conductive gel, saline or water-based solutions, or dry-contact technology. Electrodes with higher contact impedance require amplifiers with higher technical specifications, but allow for better usability and easier setup.

- **EEG Amplifier and Signal Acquisition.** Its function is to amplify low-amplitude brain signals, typically in the range of tens of microvolts, and convert them into digital form. The sampling rate is a key parameter, and must be sufficiently high to capture the frequency content of EEG signals (commonly $\geq$ 256 Hz). The resolution of the analog-to-digital converter is also important, as this determines sensitivity to small voltage variations.

- **Connectivity, Power and Portability.** Modern EEG systems offer wired or wireless data transmission, improving subject mobility and comfort but also introducing bandwidth or latency constraints. Furthermore, these devices can operate using external power sources or batteries, enhancing portability.

## 1.3 Objectives and scope of the project

The main objective of this project is to study the feasibility of biometric identification and authentication of individuals using electroencephalographic (EEG) signals. The project will analyse the extent to which subject-specific neural patterns (brainprints) can be reliably extracted and exploited under varying experimental conditions. In particular, the work aims to assess the stability of EEG-based biometric representations across different recording sessions and cognitive tasks, as well as to explore strategies for reducing personal identifiability while preserving signal utility for other applications.

To achieve this, the following specific objectives are defined:

- Review and analyse existing approaches to biometric identification and authentication.
- Design and implement a complete EEG-based biometric system, including signal preprocessing, feature extraction, and the training of machine learning models for both identification and authentication tasks.
- Evaluate the performance of different machine learning models, including traditional and deep learning approaches.
- Analyse the influence of key factors such as the number of subjects, recording sessions, EEG channels and tasks.
- Explore signal anonymisation techniques and assess the impact of anonymisation methods on both biometric identification performance and alternative EEG-based analysis tasks.

In order to address these objectives in a structured manner, this document is organised as follows. The introductory chapter provides background on biometric recognition and EEG signals. The state of the art in EEG-based brainprint identification and signal anonymisation is reviewed in Chapter 2, followed by a description of the datasets used in Chapter 3. The overall system pipeline is presented in Chapter 4, while the machine learning models and evaluation metrics are introduced in Chapter 5. Biometric identification and authentication, with their experimental protocols and results are detailed in Chapter 6. Signal anonymisation methods and their impact on identity suppression and task preservation are analysed in Chapter 7. Finally, the main conclusions and future research directions are discussed in Chapter 8. A summary of the project planning and timeline is provided in the Gantt chart of Appendix A.

# Chapter 2

# Related work

The use of electroencephalography (EEG) signals for biometric purposes has been increasingly studied in recent years, particularly due to advances in machine learning and the proliferation of affordable EEG acquisition devices. This chapter provides a review of the latest research on EEG-based biometric identification and authentication in Section 2.1, with a particular focus on brainprint extraction methods and system performance. Furthermore, previous studies on privacy and anonymisation of EEG signals are reviewed in Section 2.2.

## 2.1 Brainprint: identification based on EEG

Over the last decade, a growing body of research has investigated the feasibility of using electroencephalography (EEG) as a biometric identifier, exploring a wide variety of signal representations, machine learning models, and experimental protocols [13].

Early and classical approaches to EEG-based biometric recognition primarily focused on handcrafted features derived from the spectral properties of the signal. In particular, frequency-domain representations such as power spectral density (PSD) features have consistently been shown to encode discriminative information across individuals. Several studies have demonstrated that EEG spectral patterns remain sufficiently distinctive to enable identification and verification under controlled conditions, even when using relatively simple classifiers [14]. For instance, high authentication accuracies have been reported using PSD-based features with a reduced number of EEG channels. Similarly, it has been shown that PSD features extracted from theta, alpha, and beta frequency bands can achieve identification accuracies above 94% on datasets with limited numbers of subjects, providing early evidence of the discriminative power of spectral EEG representations [15].

Beyond purely spectral representations, some studies have explored functional connectivity and spatial relationships between brain regions. Brainprint frameworks based on connectivity measures and asymmetry indices have demonstrated that interaction patterns between EEG channels can yield highly discriminative identity

representations [16]. Such approaches have achieved near-perfect identification performance in small-scale experimental settings, showing that competitive results can be obtained using both high-density laboratory systems and low-cost consumer-grade EEG devices.

Conventional machine learning models were widely adopted in early EEG-based biometric systems due to their interpretability and relatively low computational complexity. Support Vector Machines, multilayer perceptrons, and other shallow neural networks have been commonly employed to classify EEG-derived features, achieving competitive authentication performance with a limited number of electrodes [14].

More recent works have emphasized the role of deep learning models for EEG-based identification and authentication, leveraging their ability to automatically learn discriminative features directly from EEG signals. Convolutional and recurrent neural networks have been successfully applied in this context, often achieving very high identification accuracies and low equal error rates [17, 18]. Lightweight convolutional neural networks have also been explored, demonstrating that high identification performance can be achieved using a very small number of EEG channels, which is particularly relevant for practical biometric systems [19].

Another important line of research has focused on the stability of brainprints across time and experimental conditions. Several studies have investigated the effect of incorporating multiple recording sessions into training procedures, showing that leveraging EEG data from distinct sessions significantly improves verification performance and robustness [20]. Furthermore, longitudinal analyses have provided evidence that person-identifying EEG patterns remain embedded in neural activity over time, supporting the existence of stable brainprints despite session-to-session variability [15, 21].

Across the literature, a wide diversity of experimental setups can be observed. Public datasets, such as EEG Motor Movement/Imagery Dataset provided by the PhysioNet BCI [22], are commonly used due to their relatively large number of subjects and standardized acquisition protocols, while many studies also rely on private datasets with smaller cohorts. The number of EEG channels varies substantially, ranging from single-channel or reduced-channel configurations aimed at simplifying acquisition, to high-density systems. While high accuracies are frequently reported in controlled experimental settings and for limited subject populations, the scalability of EEG-based biometric systems and their generalization to larger, more heterogeneous cohorts and real-world conditions remain open challenges.

## 2.2 Privacy and anonymisation of EEG signals

Privacy concerns related to EEG data have been extensively discussed in the context of brain–computer interfaces (BCIs) and neural signal analysis [5]. Unlike many traditional biometric modalities, EEG signals may encode not only identity-related

information but also other sensitive personal attributes, such as gender, cognitive state, emotional responses, or user experience level. Multiple studies have demonstrated that such attributes can be inferred with high accuracy using machine learning models [23], raising significant concerns regarding the storage, sharing, and reuse of raw EEG recordings.

To address these issues, a growing body of research has proposed anonymisation and privacy-preserving techniques for EEG data, aiming to reduce or suppress identity-related information while maintaining task-relevant features. One widely explored strategy is signal perturbation, where controlled noise or transformations are applied to the EEG signals to reduce their discriminative power for identity recognition. Recent work has shown that user-wise perturbations can substantially degrade identity recognition performance across both deep learning and classical machine learning models, while preserving utility for non-biometric tasks [24]. Different perturbation strategies have been investigated, including random noise, synthetic noise, and optimization-based perturbations designed to minimize or maximize classification error.

More structured anonymisation approaches rely on learned transformations, such as autoencoders and generative models. In particular, adversarial frameworks based on generative adversarial networks have been proposed to map EEG signals into alternative representations that mask personal identity while preserving task-related information [25]. In a similar direction, recent studies have explored autoencoder-based architectures and federated learning schemes to generate privacy-preserving EEG representations or synthetic data, often incorporating differential privacy mechanisms to control information leakage [26].

Additionally, cancellable template approaches have been proposed in EEG-based authentication systems to generate non-invertible biometric representations that support revocability and reduce long-term privacy risks [27]. These methods aim to balance biometric utility and privacy preservation, which remains a key challenge for the practical deployment of EEG-based biometric systems.

# Chapter 3

# Datasets

This chapter describes the EEG datasets employed in this work for the evaluation of biometric identification, authentication, and anonymisation methods. A diverse set of datasets was deliberately selected in order to analyse the proposed framework under a wide range of experimental conditions, including variations in population size, number of recording sessions, task paradigms, sensor configurations, and recording environments. Both publicly available and proprietary datasets were considered.

## 3.1 Public datasets

Public EEG datasets were initially selected for experiments due to their availability and widespread use in the literature.

- **PhysioNet EEG Motor Movement/Imagery dataset** [22]. This is one of the most widely used public EEG datasets and has been extensively adopted in brain-computer interface and EEG-based biometric research. It contains EEG recordings from 109 subjects acquired with the BCI2000 system [28], using 64 electrodes placed according to the international 10-10 system. The signals were recorded at a sampling frequency of 160 Hz. Each subject performed a series of motor imagery and motor execution tasks, including imagined or executed movements of the left or right hand and both hands or feet, as well as resting-state periods.

  In this project, the PhysioNet dataset was the first public resource explored during the initial development and validation of the identification pipeline. However, a key limitation of this dataset in the context of biometric evaluation is that recordings are provided within a single session per subject, without explicit session separation over time. As a result, it does not allow a rigorous assessment of inter-session variability, which is a critical factor for evaluating the stability and robustness of brainprint representations.

- **BCI Competition IV dataset 2b** [29]. It is a publicly available EEG dataset

designed for motor imagery classification tasks. It includes recordings from 9 subjects, each participating in multiple recording sessions. EEG data were acquired using 3 channels, with a sampling frequency of 250 Hz. Subjects performed left-hand and right-hand motor imagery tasks across five sessions, recorded on different days.

Despite its relatively small number of subjects and channels, this dataset was selected for this project due to its clear session separation and well-defined experimental protocol. These characteristics make it a valuable dataset for analysing identification and authentication performance under constrained sensing conditions, as well as for studying the impact of session variability on biometric stability.

- **Healthy Brain Network EEG dataset (HBN-EEG)** [30]. This dataset is part of the Healthy Brain Network initiative and provides a large-scale collection of EEG recordings from a pediatric and adolescent population. The dataset includes EEG data from over 3000 participants, acquired using 128 channels. Recordings include a diverse set of task paradigms performed within a single recording session. These tasks are grouped into *active* and *passive* categories. Active tasks correspond to those that require explicit cognitive engagement and goal-directed processing, and include *Contrast Change Detection*, *Sequence Learning*, and *Symbol Search*. Passive tasks, in contrast, involve minimal or no explicit task demands and include *Resting State*, *Surround Suppression*, and movie-watching conditions (*Despicable Me*, *Diary of a Wimpy Kid*, *Fun with Fractals*, and *The Present*).

  The HBN-EEG dataset was incorporated into this project to assess the scalability of EEG-based biometric methods to larger and more heterogeneous populations. Its diversity in terms of subject characteristics and task paradigms provides a realistic and challenging testing environment for evaluating cross-task generalization and identification robustness at scale.

## 3.2 Bitbrain datasets

In addition to common public datasets, several EEG datasets provided by Bitbrain Technologies were used in this project. These datasets were collected in controlled research and applied contexts and include multi-session and longitudinal recordings, allowing the evaluation of EEG-based biometric systems under more realistic conditions.

- **Elevvo Depression dataset**. It was collected as part of a controlled clinical study investigating the cognitive effects of alpha neurofeedback training in patients diagnosed with Major Depressive Disorder [31]. EEG recordings were acquired using 16 electrodes with a sampling rate of 256 Hz and

placed according to the international 10-10 system. They were recorded under controlled laboratory conditions, including resting-state and task-related segments associated with neurofeedback training sessions. The dataset includes 49 subjects with aproximately 8 sessions and 24 subjects with 2 sessions.

The presence of repeated recordings across sessions makes this dataset particularly suitable for evaluating session-wise generalisation and authentication performance.

- **Multi-session Cognitive Tasks datasets**. Two additional datasets were collected as part of a study involving older adults performing a set of cognitive tasks across multiple recording sessions. Each dataset contains recordings from approximately 30 subjects, with around eight sessions per subject, enabling robust analysis of temporal stability in EEG-based biometric features. Both datasets contain signals recorded at a sampling frequency of 256 Hz using two EEG channels. Each recording session comprises three types of tasks: *calibration*, *trial*, and *evaluation*. *Calibration* and *evaluation* tasks consist of two short recordings per session each, while the *trial* tasks are five recordings per session, with a larger duration each. The primary difference between both datasets lies in the recording environment. One dataset was acquired in a controlled laboratory setting, while the other was collected in participants' homes, introducing greater variability in environmental and contextual conditions.

  This dataset was used to analyse how task dependency and session variability jointly affect EEG-based biometric systems, as well as to assess whether reliable identification and authentication can be achieved using a highly reduced number of electrodes.

- **Sleep EEG dataset**. An additional proprietary dataset consists of overnight EEG recordings collected from approximately 30 subjects, each with three sleep sessions recorded on different days. EEG data were acquired at a sampling frequency of 128 Hz, using two EEG channels. The recordings were annotated according to standard sleep stages.

  This dataset was primarily used in this project to evaluate the interaction between biometric anonymisation techniques and downstream EEG-based tasks unrelated to identity, such as sleep stage classification. The multi-session nature of the dataset allows assessment of how subject-specific information persists across nights and how anonymisation strategies affect identification performance under constrained sensing conditions.

- **Bitbrain Open Access Sleep dataset (BOAS)** [32]. This is a publicly available EEG sleep dataset designed to support research bridging the gap between clinical polysomnography (PSG) and emerging wearable EEG technologies. The dataset comprises simultaneous overnight recordings collected from a clinical-grade Brain Quick Plus Evolution PSG system by Micromed and a Bitbrain wearable EEG headband across 128 nights of sleep from 108 healthy

participants. Each recording includes both PSG-derived EEG and wearable EEG data, which are labelled according to consensus human sleep stage annotations following American Academy of Sleep Medicine (AASM) criteria, as well as AI-derived sleep stage labels obtained through automated models trained on the expert annotations.

This dataset was specifically utilised to evaluate anonymization in this work, training and evaluating the performance of a sleep stage classification model on both original and anonymised EEG data.

# Chapter 4

# System Pipeline

This chapter describes the complete processing pipeline implemented in this work for EEG-based biometric analysis. The proposed pipeline provides a modular framework that supports subject identification, authentication, and the evaluation of privacy-preserving signal anonymisation strategies. An overview of the system pipeline is illustrated in Figure 4.1. The pipeline processes raw multichannel EEG recordings and transforms them into biometric decisions through a sequence of well-defined stages, including signal segmentation, feature extraction, machine learning-based modeling, and decision aggregation. An optional anonymisation block can be integrated into the pipeline prior to feature extraction. Each stage is described in detail in the following sections.

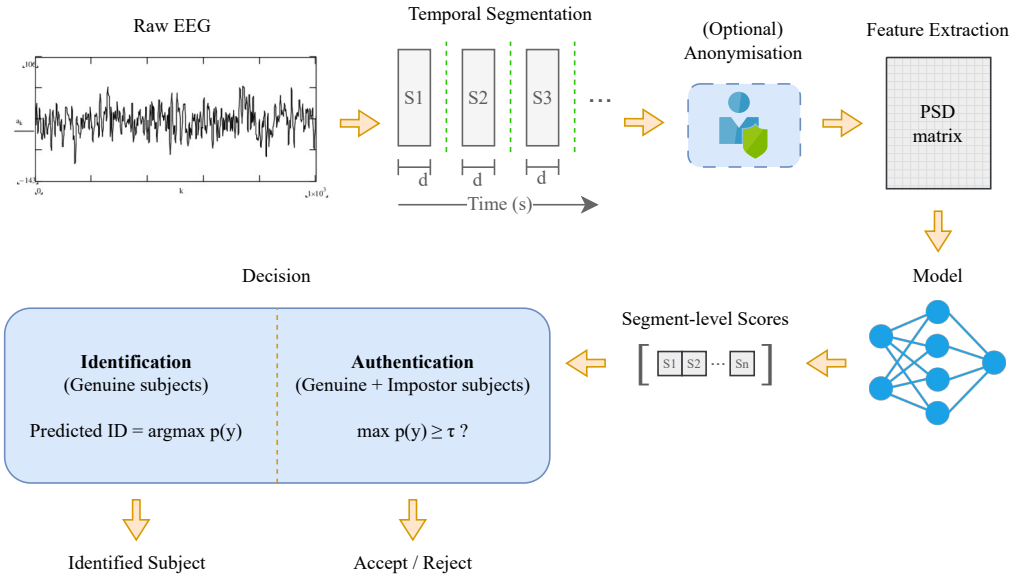



*Figure 4.1: Pipeline of the implemented system.*

## 4.1 Raw EEG acquisition

The input to the system consists of raw EEG recordings acquired using different EEG devices and experimental setups. Depending on the dataset, recordings may include a variable number of channels, ranging from highly constrained configurations with only a few electrodes to high-density systems with more than one hundred channels. Sampling frequencies also vary across datasets.

At this stage, EEG signals are treated as multivariate time series of the form:

$$\mathbf{X} \in \mathbb{R}^{C \times T},$$

where $C$ denotes the number of EEG channels and $T$ the number of temporal samples.

## 4.2 Temporal segmentation

EEG signals are inherently non-stationary and exhibit significant variability over time. To mitigate these effects and increase the number of training samples, EEG recordings are segmented into fixed-length temporal windows. This segmentation strategy ensures that each segment contains sufficient temporal information to capture relevant spectral characteristics, while improving the statistical robustness of the learning process.

Given an EEG recording $\mathbf{X} \in \mathbb{R}^{C \times T}$, the signal is divided into $S$ segments of $D$ seconds, resulting in a segmented representation:

$$\mathbf{X} \rightarrow \{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_S\}, \quad \mathbf{X}_s \in \mathbb{R}^{C \times D}.$$

The duration of each segment depends on the total signal length and the acquisition protocol of each dataset, but the segmentation strategy remains consistent throughout the pipeline.

## 4.3 Optional signal anonymisation

To enable the evaluation of privacy-preserving strategies, the pipeline incorporates an optional anonymisation block that operates at the segment level. This block is placed before feature extraction, ensuring that any privacy-preserving transformation affects all downstream representations and models consistently.

Anonymisation is implemented as a signal-level transformation that modifies each EEG segment while preserving its temporal structure. The specific anonymisation methods explored in this work are described in detail in Chapter 7. From a pipeline perspective, this stage acts as a transparent preprocessing step that outputs a modified

version of each EEG segment:

$$\tilde{\mathbf{X}}_s = \mathcal{A}(\mathbf{X}_s),$$

where $\mathcal{A}(\cdot)$ denotes the anonymisation operator and $\tilde{\mathbf{X}}_s$ represents the anonymised EEG segment. When anonymisation is not applied, $\tilde{\mathbf{X}}_s = \mathbf{X}_s$.

## 4.4 Feature extraction

From each EEG segment, frequency-domain features are extracted in the form of Power Spectral Density (PSD) representations. PSD features are widely used in EEG-based biometric systems due to their robustness and their ability to characterize subject-specific frequency-domain patterns.

For each segment $\tilde{\mathbf{X}}_s$, the PSD is computed independently for each EEG channel and restricted to a frequency range between 1 and 45 Hz, covering the conventional EEG bands while excluding high-frequency noise and power-line artefacts. In order to stabilize the dynamic range and enhance numerical performance during model training, PSD values are expressed on a logarithmic scale.

This process yields a feature representation of the form:

$$\mathbf{F}_s \in \mathbb{R}^{C \times B},$$

where $B$ denotes the number of frequency bins.

## 4.5 Model inference

The extracted feature representations constitute the input to machine learning models trained to perform biometric identification. Both traditional machine learning models and deep learning architectures are supported by the pipeline. All models operate on the same feature representations to ensure comparability across approaches. They are described in detail in Chapter 5.

At inference time, each EEG segment is processed independently by the trained model, producing a vector of class-level scores or posterior probabilities associated with the enrolled subjects. These outputs provide a segment-level estimate of subject identity and form the basis for both identification and authentication decisions.

## 4.6 Decision-making and aggregation

The final stage of the pipeline converts segment-level model outputs into biometric decisions. In identification scenarios, the predicted subject label corresponds to the class with the highest confidence score. In authentication scenarios, confidence scores associated with a claimed identity are compared against a decision threshold to accept or reject the authentication attempt.

To improve robustness and reduce the impact of intra-recording variability, predictions can be aggregated across multiple segments belonging to the same recording or session. Aggregation strategies include majority voting and score averaging, both of which integrate evidence over time to produce a single decision per EEG recording.

# Chapter 5

# Methods

This chapter describes the methodological framework adopted in this work, including the machine learning models employed for EEG-based biometric analysis and the evaluation strategy used throughout the experiments.

## 5.1 Models

To address the EEG-based identification and authentication, both traditional machine learning classifiers and deep learning approaches were implemented. All models operate on the same PSD-based feature representations extracted from segmented EEG signals, as described in the system pipeline (Chapter 4).

### 5.1.1 Traditional Machine Learning

Traditional machine learning classifiers serve as strong baselines and allow for a direct comparison with deep learning approaches in terms of performance, scalability, and interpretability.

**Input representation.** Let $N_c$ denote the number of EEG channels and $N_f$ the number of frequency bins used to compute the PSD. For each EEG segment, the PSD representation is initially computed as a matrix $\mathbf{P} \in \mathbb{R}^{N_c \times N_f}$. This matrix is then reshaped into a one-dimensional feature vector

$$\mathbf{x} \in \mathbb{R}^{N_c \cdot N_f}$$

by concatenating all channel-frequency components.

For multi-subject classification, feature vectors from all subjects are stacked into a global feature matrix $\mathbf{X} \in \mathbb{R}^{N \times (N_c \cdot N_f)}$, where $N$ denotes the total number of EEG segments across all subjects. Each segment is assigned a categorical label corresponding

to the subject identity, resulting in a label vector $\mathbf{y} \in \{1, \ldots, S\}^N$, where $S$ is the number of enrolled subjects.

**Support Vector Machine (SVM).** Support Vector Machines are supervised learning models that aim to find an optimal separating hyperplane by maximising the margin between classes in a high-dimensional feature space. In this work, a multi-class SVM formulation is employed using a one-vs-rest strategy, allowing the model to distinguish among multiple enrolled subjects simultaneously.

The parameters were selected manually. A linear kernel was selected as the primary configuration due to its computational efficiency and its widespread use in EEG biometric literature, where spectral features often exhibit near-linear separability. The regularization parameter $C$ controls the trade-off between maximizing the margin and minimizing classification errors, thereby influencing the model's generalization capability. Potential class imbalance, arising from unequal numbers of EEG segments per subject, was addressed by incorporating balanced class weighting during training.

In order to facilitate authentication analysis based on confidence scores, probabilistic outputs were estimated for each class. This allows the computation of subject-specific confidence values that can later be thresholded to produce binary authentication decisions.

Due to its simplicity, robustness and interpretability, the SVM model serves as a baseline against which more complex approaches can be compared.

**Additional baseline models.** In addition to SVMs, other classical machine learning approaches were implemented as complementary baselines, including Logistic Regression and Multilayer Perceptron (MLP) networks. A multinomial logistic regression model can learn linear decision boundaries in the feature space while producing class-level probability estimations. Despite its limited representational capacity, it achieved competitive performance in scenarios where EEG spectral features exhibited a high degree of separability.

Furthermore, the use of a multilayer perceptron (MLP) classifier was considered as a baseline for a shallow neural network. The MLP consists of multiple fully connected hidden layers with non-linear activation functions and is trained using gradient-based optimization. Although the depth and expressiveness of the MLP are more limited than those of convolutional architectures, it achieved identification results that were, in some configurations, comparable to those obtained with the SVM model. As both models yielded results that were largely consistent with those obtained using SVMs, they were not selected for extensive experimentation.

### 5.1.2 Convolutional Neural Network (CNN)

In addition to traditional classifiers, a deep learning model based on a Convolutional Neural Network (CNN) architecture was implemented to capture more complex patterns in EEG spectral representations. CNNs have demonstrated strong performance in EEG analysis due to their ability to exploit local spatial and frequency-wise correlations present in structured feature representations.

**Input representation.** For the CNN-based approach, each EEG segment is represented as a two-dimensional PSD matrix

$$\mathbf{P} \in \mathbb{R}^{N_c \times N_f},$$

where $N_c$ denotes the number of EEG channels and $N_f$ the number of frequency bins. To comply with the convolutional input format, the PSD matrix is augmented with a singleton dimension, resulting in an input tensor of shape

$$\mathbf{X} \in \mathbb{R}^{1 \times N_c \times N_f}.$$

This representation preserves the spatial structure across channels while enabling the network to learn frequency-dependent patterns through convolutional filters.

**Network architecture.** The proposed CNN architecture is illustrated in Figure 5.1.

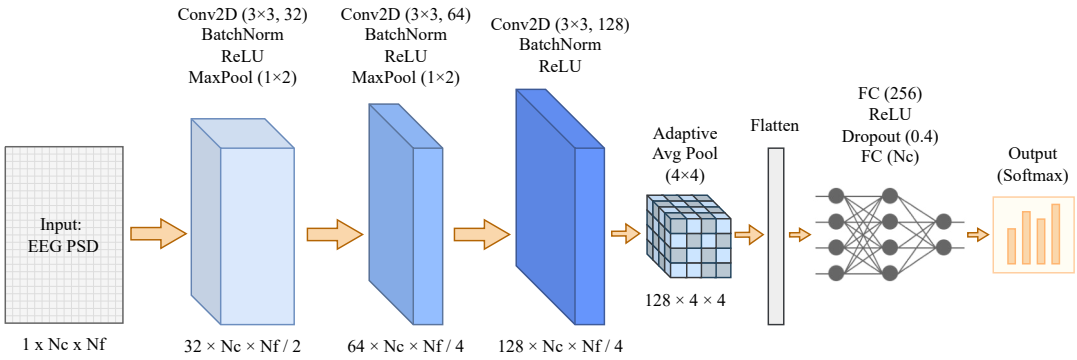



*Figure 5.1: The CNN architecture.*

It consists of a sequence of convolutional blocks, each comprising:

- Two-dimensional convolutional layers with small kernel sizes to capture local spectral and spatial patterns across channels and frequency bins.
- Batch normalization layers to stabilize training and improve convergence.
- Rectified Linear Unit (ReLU) activation functions to introduce non-linearity.

Spatial pooling is applied exclusively along the frequency dimension, progressively reducing spectral resolution while preserving channel-wise structure. Moreover, an adaptive average pooling layer is employed to compress the learned feature maps into a fixed-size representation, enabling robustness to variations in input dimensionality. The resulting features are then passed to a set of fully connected layers, including dropout-based regularization, which integrate the extracted representations and produce class-level output scores. A softmax activation is applied at the output to obtain normalized confidence scores for each enrolled subject.

**Training and inference.** The CNN is trained using the Adam optimiser and a categorical cross-entropy loss function. Mini-batch training is employed, and a portion of the training data is reserved for validation to monitor generalisation performance during training. As with the SVM model, the CNN outputs class-level confidence scores, which are subsequently reused for authentication evaluation and anonymisation analyses.

## 5.2 Evaluation

The performance of the proposed system was evaluated using a set of standard metrics commonly employed in biometric identification and authentication. These metrics were selected to provide both an overall measure of performance and an analysis of classification behaviour under different experimental conditions.

**Accuracy.** Accuracy was used as the primary evaluation metric in both identification and authentication experiments. It measures the proportion of correctly classified samples with respect to the total number of evaluated samples and is defined as:

$$\text{Accuracy} = \frac{N_{\text{correct}}}{N_{\text{total}}}, \tag{5.1}$$

where $N_{\text{correct}}$ is the number of correctly predicted samples and $N_{\text{total}}$ is the total number of samples.

In identification experiments, accuracy reflects the ability of the system to correctly assign an EEG sample to the corresponding enrolled subject in a multi-class setting. In authentication experiments, accuracy indicates the proportion of correct acceptances and rejections when distinguishing between genuine users and impostors.

Accuracy was computed at two different levels: *segment-level accuracy*, where each EEG segment is evaluated independently, and *signal-level accuracy*, obtained by aggregating multiple segment-level predictions belonging to the same recording using a decision fusion strategy based on majority voting. This allows the evaluation of how temporal aggregation affects biometric performance.

**Equal Error Rate.** For authentication scenarios, the Equal Error Rate (EER) was used as a key metric. EER corresponds to the operating point at which the *False Acceptance Rate (FAR)* equals the *False Rejection Rate (FRR)*:

$$\text{EER} = \text{FAR}(\tau) = \text{FRR}(\tau), \tag{5.2}$$

where $\tau$ denotes the decision threshold applied to the authentication scores.

FAR represents the probability of incorrectly accepting an impostor, while FRR represents the probability of incorrectly rejecting a genuine user. EER provides a threshold-independent measure of system performance and is widely adopted in biometric evaluation, particularly for comparing authentication systems under different conditions.

**Precision, recall, and F1-score.** In selected experiments, additional classification metrics were reported to better characterize model performance. Precision, recall, and F1-score are defined as:

$$\text{Precision} = \frac{TP}{TP + FP}, \quad \text{Recall} = \frac{TP}{TP + FN}, \quad \text{F1-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}, \tag{5.3}$$

where $TP$, $FP$, and $FN$ denote true positives, false positives, and false negatives, respectively.

**Confusion matrix and probability analysis.** Confusion matrices were also used in selected experiments to visualize subject-wise classification errors and identify systematic confusions between specific individuals. Additionally, the predicted class probability distributions were analyzed to assess model confidence and uncertainty, providing insight into how decisively the system distinguishes between subjects.

**ROC curve and AUC.** For some authentication experiments, Receiver Operating Characteristic (ROC) curves were computed by varying the decision threshold and plotting the True Positive Rate (TPR) against the False Positive Rate (FPR). The Area Under the ROC Curve (AUC) was used as a scalar summary of the ROC performance, with values closer to one indicating better discriminative capability across thresholds.

**Relative performance loss.** Relative performance loss (RPL) was used in anonymisation experiments to observe the accuracy degradation induced by signal perturbations with respect to a clean, non-anonymised reference condition. It is defined as:

$$\text{Relative performance loss} = \frac{\text{Accuracy}_{\text{ref}} - \text{Accuracy}_{\text{anon}}}{\text{Accuracy}_{\text{ref}}} \times 100, \tag{5.4}$$

where $\text{Accuracy}_{\text{ref}}$ denotes the accuracy obtained using clean EEG signals, and $\text{Accuracy}_{\text{anon}}$ corresponds to the accuracy achieved after anonymisation.

This metric enables direct comparison across anonymisation configurations and noise levels. A relative performance loss of 0% indicates no degradation with respect to the reference condition, while higher values reflect increasing loss of task-relevant information.

Together, these metrics provide a comprehensive evaluation framework that captures both overall biometric performance and finer-grained aspects of model behaviour across identification, authentication, and anonymisation experiments.

# Chapter 6

# Identification and Authentication

This chapter addresses the core biometric problems studied in this work: EEG-based identification and authentication. Building upon the datasets described in Chapter 3, the system pipeline introduced in Chapter 4, and the methodological framework detailed in Chapter 5, this chapter focuses on the formulation, experimental design, and analysis of identity recognition using electroencephalographic signals.

## 6.1 Problem definition

Biometric systems aim to recognize individuals based on distinctive physiological or behavioral characteristics. In the context of EEG-based biometrics, identity recognition relies on the hypothesis that neural activity contains subject-specific patterns, commonly referred to as *brainprints*, which remain sufficiently stable across time and experimental conditions.

This work considers two closely related biometric problems: identification and authentication. Although both rely on the same underlying EEG representations and learning models, they differ in their objectives, decision mechanisms, and evaluation criteria.

### 6.1.1 Identification

EEG-based identification is formulated as a closed-set multi-class classification problem. Given a set of $N$ enrolled subjects, the objective is to assign an EEG recording to the correct individual among this set.

Let $\mathbf{x} \in \mathbb{R}^d$ denote the feature representation extracted from a single EEG segment, where $d$ corresponds to the dimensionality of the PSD-based feature space described in Chapter 4, and let $y \in \{1, \ldots, N\}$ be the corresponding subject label, identification

consists in learning a mapping

$$f : \mathbb{R}^d \to \{1, \ldots, N\},$$

such that $f(\mathbf{x})$ predicts the identity of the subject who generated the EEG segment.

During training, the model is exposed to EEG segments from all enrolled subjects, trying to maximize inter-subject separability while minimizing intra-subject variability.

### 6.1.2 Authentication

EEG-based authentication addresses the problem of determining whether a given EEG recording originates from one of the previously enrolled subjects or from an unknown individual. In this scenario, authentication is formulated as a binary decision problem, distinguishing between *genuine* attempts, corresponding to enrolled subjects, and *impostor* attempts, corresponding to subjects not observed during training.

In this work, authentication is not implemented as an independent binary classifier. Instead, it is derived from the outputs of a closed-set multi-class identification model trained exclusively on enrolled subjects. This design choice reflects practical biometric systems, where a single trained model can support both identification and authentication by exploiting class-level confidence scores.

Given an EEG segment $\mathbf{x}$, the identification model produces a set of confidence scores $\{s(\mathbf{x}, k)\}_{k=1}^{N}$, one for each enrolled subject. Authentication is performed by comparing the maximum confidence score against a predefined threshold $\tau$:

$$\text{accept if } \max_k s(\mathbf{x}, k) \geq \tau, \quad \text{reject otherwise.}$$

## 6.2 Experimental protocol

The experimental protocol for identification and authentication is designed to systematically evaluate the robustness of EEG-based biometric systems under diverse conditions. Rather than defining a single fixed experimental setup, this work adopts a progressive evaluation strategy, starting with exploratory analyses on public EEG datasets and evolving towards more complex and realistic scenarios.

All experiments follow the common system pipeline described in Chapter 4 and rely on the feature extraction and modeling methodology detailed in Chapter 5. Differences between experiments arise from the choice of dataset, data partitioning strategy, sensor configuration, and evaluation objective.

### 6.2.1 Data partitioning and generalization scenarios

To assess the generalization capabilities of EEG-based biometric systems, multiple data partitioning strategies are employed:

- **Session-based generalization**: EEG recordings are split according to recording sessions. Models are trained on a subset of sessions for each subject and evaluated on held-out sessions recorded at different times. This protocol explicitly assesses inter-session variability, which is a critical factor in real-world biometric applications.

- **Task-based generalization**: EEG data are partitioned according to task type. Models are trained on a subset of cognitive or experimental tasks and evaluated on unseen tasks, testing the extent to which subject-specific information transfers across different mental states.

In all cases, data from all enrolled subjects are included in both training and testing phases, with the exception of impostor subjects used exclusively for authentication evaluation.

### 6.2.2 Identification and authentication evaluation

In identification, at inference time, the trained model produces a predicted class label and an associated confidence score for each input segment. Identification performance is primarily evaluated using multi-class classification metrics, such as accuracy and confusion matrices, described in Section 5.2, at both the segment level and the signal (session) level. Signal-level decisions are obtained by aggregating predictions across all segments belonging to the same EEG recording, as described in Chapter 5.

For authentication experiments, as described in Section 6.1.2, impostor detection is incorporated by introducing EEG data from subjects not included during model training, and decisions are derived by thresholding class-level confidence scores produced by the identification model. The decision threshold is selected based on validation data by maximizing authentication accuracy, and its effect on the trade-off between false acceptance and false rejection rates is further analyzed using Receiver Operating Characteristic (ROC) curves, Area Under the Curve (AUC), and Equal Error Rate (EER), explained in Section 5.2.

Required deviations from the base protocol, such as changes in channel selection or number of subjects, are explicitly described in the corresponding experimental analysis.

A summary of all experimental configurations considered for identification and authentication, including datasets, number of subjects, session structure, channel configuration, and evaluation objective, is provided in Table 6.1.

| Id | Dataset | Subjects | | Type | Sessions | Tasks | Channels | Motivation |
|---|---|---|---|---|---|---|---|---|
| | | Genuine | Impostors | | | | | |
| **Baseline identification** | | | | | | | | |
| E1 | PhysioNet | 109 | – | Session-based | 1 | 6 | 64 | Feasibility with single-session data. |
| E2 | BCI IV 2b | 9 | – | Session-based | 5 | 2 | 3 | Multi-session identification on a public dataset. |
| E3 | Elevvo Depression | 49 | – | Session-based | ∼8 | 3 | 16 | Identification on clinical multi-session EEG. |
| **Session-wise identification and authentication performance** | | | | | | | | |
| E4 | Elevvo Depression | 49 | 25 | Session-based | ∼8 (2) | 3 (3) | 16 | Session-wise generalization and impostor detection. |
| E5 | Elevvo Depression | 49 | 25 | Session-based | ∼8 (2) | 3 (3) | 16 | Signal-level decision aggregation. Segment vs signal-level prediction. |
| **Impact of channel configuration** | | | | | | | | |
| E6 | Elevvo Depression | 49 | 25 | Session-based | ∼8 (2) | 3 (3) | 1–16 | Impact of channel selection in identification and authentication. |
| **Task-based generalization** | | | | | | | | |
| E7 | Multi-session Cognitive Tasks | 47 | 10 | Session-based | ∼8 (∼8) | 3 (3) | 2 | Robustness across recording environments. |
| E8 | Multi-session Cognitive Tasks | 47 | 10 | Task-based | ∼8 (∼8) | 3 (3) | 2 | Generalization across cognitive tasks. |
| E9 | HBN-EEG | 50 | 10 | Task-based | 1 (1) | ∼8 (∼8) | 129 | Cross-task and task-holdout evaluation. Active vs passive task generalization. |
| **Model comparison and scalability** | | | | | | | | |
| E10 | Multi-session Cognitive Tasks | 47 | 10 | Task-based | ∼8 (∼8) | 3 (3) | 2 | Comparison of classical and deep models. |
| E11 | HBN-EEG | 100–3059 | 30 | Task-based | 1 (1) | ∼8 (∼8) | 129 | Effect of population size on identification. |

*Table 6.1: Summary of configurations for identification and authentication experiments evaluated in this project. When impostor subjects are included, values in parentheses indicate the number of sessions or tasks available for them.*

## 6.3 Results and discussion

This section presents and discusses the experimental results obtained for EEG-based identification and authentication under the protocols described in Section 6.2.2. The analysis is structured to progressively evaluate the robustness of the proposed approach under increasingly challenging conditions, including inter-session variability, task generalization, channel reduction, and scalability with respect to the number of enrolled subjects.

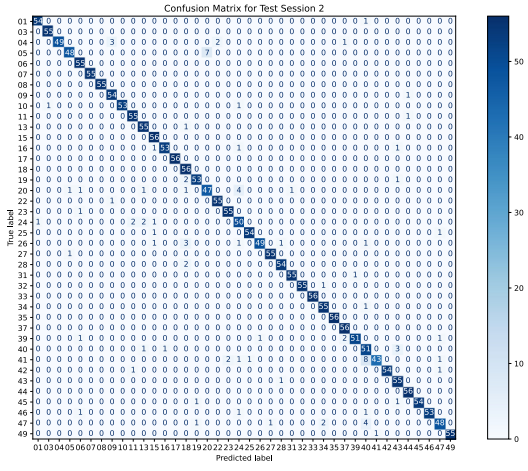
Preliminary experiments assessing the basic feasibility of EEG-based identification under simplified conditions were conducted on multiple public datasets. As these experiments primarily serve as proof-of-concept analyses and do not include authentication or advanced generalization protocols, they are reported separately in Appendix B.1. The remainder of this section concentrates on experiments that explicitly address biometric robustness and real-world applicability.

## Session-wise identification and authentication performance

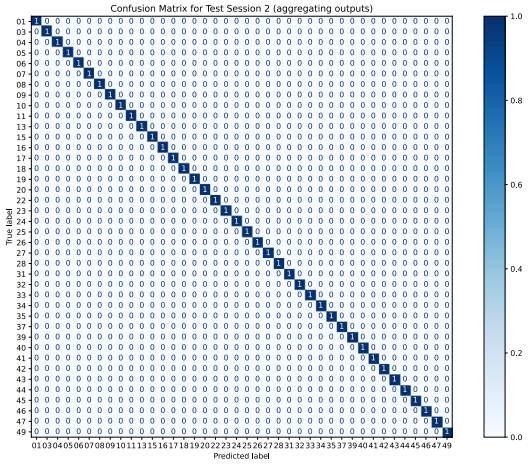
The first set of core experiments evaluates identification and authentication performance under session-wise generalization, using the multi-class SVM classifier and the Elevvo Depression dataset, which provides multiple recording sessions per subject and a larger number of subjects. Initially, decisions are obtained at the segment level, where each EEG window is independently classified. To better reflect realistic biometric usage scenarios, performance is subsequently evaluated at the signal (session) level by aggregating segment-level predictions into a single decision per recording. A session-wise cross-validation protocol was adopted, in which one session per subject was reserved for testing while the remaining sessions were used for training, and performance metrics were averaged across varying test sessions. This protocol explicitly assesses the ability of the system to generalize across recording sessions, which is a critical requirement for biometric applications.

With this configuration, the identification system achieved an average accuracy of 94.05%, with precision and recall values of 94.44% and 94.05%, respectively. Authentication performance was evaluated by incorporating EEG data from 24 non-enrolled subjects as impostor attempts. Segment-level authentication yielded an accuracy of 85%, classifying individuals into genuine and impostor subjects, and an Equal Error Rate (EER) of 14.56%. These results reflect the difficulty of impostor detection when decisions are made at the segment level, where individual EEG windows may be noisy or ambiguous.

By aggregating segment-level predictions into a single decision per recording, performance improves substantially. Identification accuracy increased to 99.06%, with precision of 98% and recall of 99%. The confusion matrices before and after aggregation, shown in Figures 6.1a and 6.1b, clearly illustrate the reduction in spurious misclassifications, as aggregation suppresses segment-level noise and emphasizes consistent subject-specific patterns.



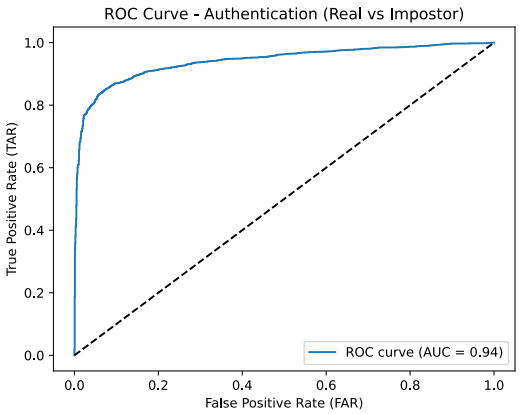

*(a) Evaluation by segments, before being grouped.*



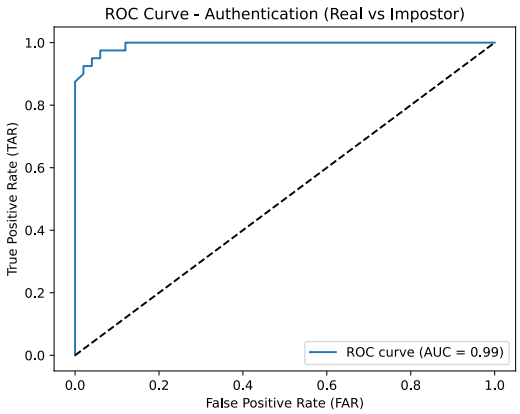

*(b) Evaluation by sessions, after grouping segments.*

*Figure 6.1: Confusion matrices for both segment-level and session-level evaluations.*

Authentication performance also improved markedly. The accuracy increased to 96.67%, while the EER dropped from 14.56% to 4.00%. This reduction indicates a much clearer separation between genuine users and impostors when decisions are made using aggregated evidence rather than individual segments. The ROC curves, presented in Figures 6.2a and 6.2a further confirms the improved discriminative capability of the system under this configuration.




*(a) Evaluation by segments, before being grouped.*




*(b) Evaluation by sessions, after grouping segments.*

*Figure 6.2: ROC curves for both segment-level and session-level evaluations.*

These results highlight the importance of decision-level aggregation in EEG-based biometric systems. While segment-level analysis is useful for feature extraction and model training, reliable biometric decisions benefit significantly from integrating information across longer temporal windows. Consequently, this aggregation strategy has been selected for subsequent experiments.

## Impact of channel configuration

After establishing the importance of session generalization and decision-level aggregation, Experiment E6 focused on analyzing the impact of EEG channel selection on both identification and authentication performance. The objective of this experiment was twofold: first, to assess whether reliable biometric performance can be achieved using a reduced number of electrodes, and second, to identify the most informative channels and channel combinations for EEG-based identification systems. This experiment was conducted using the Elevvo Depression dataset under the same session-based protocol described in the previous section, including both identification and authentication evaluations.

First, identification and authentication were evaluated using a single EEG channel at a time, selected from the 16 available electrodes in the dataset. This analysis allows the contribution of individual channels to be isolated and provides insights into the spatial distribution of subject-specific information.

Table 6.2 summarizes the results obtained for the most representative channels.

The central electrodes C3, C4, and Cz consistently achieved the highest performance, with Cz outperforming all other channels across all metrics. In contrast, O1, Fp2 and Oz showed substantially lower performance, indicating a weaker contribution to subject discrimination.

| **Channel** | **Identification** | | | | **Authentication** | | |
|---|---|---|---|---|---|---|---|
| | ACC | F1 | PREC | REC | ACC | EER | AUC |
| Cz | 0.99 | 0.99 | 0.99 | 0.99 | 0.93 | 0.06 | 0.97 |
| C4 | 0.99 | 0.99 | 0.99 | 0.99 | 0.91 | 0.10 | 0.96 |
| C3 | 0.99 | 0.98 | 0.98 | 0.99 | 0.90 | 0.11 | 0.96 |
| P4 | 0.99 | 0.98 | 0.98 | 0.99 | 0.89 | 0.11 | 0.93 |
| Fz | 0.98 | 0.97 | 0.97 | 0.98 | 0.90 | 0.11 | 0.94 |
| P3 | 0.99 | 0.98 | 0.98 | 0.99 | 0.88 | 0.13 | 0.94 |
| F3 | 0.98 | 0.97 | 0.97 | 0.98 | 0.90 | 0.13 | 0.93 |
| Pz | 0.98 | 0.98 | 0.98 | 0.98 | 0.89 | 0.13 | 0.93 |
| F4 | 0.97 | 0.96 | 0.95 | 0.97 | 0.89 | 0.12 | 0.94 |
| O2 | 0.95 | 0.93 | 0.92 | 0.95 | 0.89 | 0.12 | 0.92 |
| Fp1 | 0.94 | 0.93 | 0.92 | 0.94 | 0.89 | 0.13 | 0.94 |
| P7 | 0.96 | 0.95 | 0.95 | 0.96 | 0.87 | 0.16 | 0.91 |
| P8 | 0.96 | 0.95 | 0.95 | 0.96 | 0.86 | 0.14 | 0.90 |
| Oz | 0.95 | 0.93 | 0.92 | 0.95 | 0.87 | 0.15 | 0.92 |
| Fp2 | 0.96 | 0.95 | 0.94 | 0.96 | 0.84 | 0.17 | 0.89 |
| O1 | 0.94 | 0.92 | 0.92 | 0.94 | 0.84 | 0.17 | 0.89 |

*Table 6.2: Identification and authentication performance per EEG channel, ordered from the highest to the lowest performance.*

These results indicate that subject-discriminative information is strongly concentrated in central regions, although this effect may also be influenced by differences in signal quality across electrodes, as central channels are typically less affected by noise and artefacts than peripheral ones. Importantly, the strong performance obtained using a single central channel, particularly Cz, suggests that practical biometric systems could operate under highly constrained sensor configurations without a severe degradation in performance.

Building upon the single-channel analysis, a second set of experiments evaluated combinations of channels, progressively increasing the number of electrodes while prioritizing those that showed strong individual performance. The goal was to identify compact channel subsets that maximize performance while minimizing system complexity.

Table 6.3 presents the results for the most relevant channel combinations. Among all tested configurations, the combination Fp1–Cz–C4 achieved the best overall performance, followed closely by Fp1–Fz–Cz–C4. Increasing the number of channels beyond these configurations led to diminishing returns, suggesting that a small, well-chosen subset of electrodes is sufficient to capture most of the subject-specific

information.

| Nº Ch. | Channels | Identification | | | | Authentication | | |
|---|---|---|---|---|---|---|---|---|
| | | ACC | F1 | PREC | REC | ACC | EER | AUC |
| 2 | Cz, C4 | 0.993 | 0.991 | 0.990 | 0.993 | 0.951 | 0.055 | 0.985 |
| | Fp1, Fp2 | 0.953 | 0.937 | 0.929 | 0.953 | 0.884 | 0.122 | 0.928 |
| 3 | C3, Cz, C4 | 0.990 | 0.987 | 0.985 | 0.990 | 0.937 | 0.077 | 0.975 |
| | Fz, Cz, C4 | 0.996 | 0.995 | 0.995 | 0.996 | 0.955 | 0.055 | 0.972 |
| | Fp1, Fp2, Cz | 0.987 | 0.983 | 0.981 | 0.987 | 0.956 | 0.062 | 0.978 |
| | **Fp1, Cz, C4** | **0.996** | **0.995** | **0.995** | **0.996** | **0.983** | **0.030** | **0.992** |
| 4 | C3, Cz, C4, P4 | 0.993 | 0.991 | 0.990 | 0.993 | 0.941 | 0.067 | 0.978 |
| | Fz, C3, Cz, C4 | 0.993 | 0.991 | 0.990 | 0.993 | 0.945 | 0.070 | 0.967 |
| | **Fp1, Fz, Cz, C4** | **0.996** | **0.995** | **0.995** | **0.996** | **0.970** | **0.030** | **0.987** |
| 8 | F3, Fz, C3, Cz, C4, P3, Pz, P4 | 0.990 | 0.987 | 0.985 | 0.990 | 0.958 | 0.040 | 0.980 |
| 12 | Fp1, F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, P8, O2 | 0.990 | 0.987 | 0.985 | 0.990 | 0.969 | 0.040 | 0.986 |
| 16 | Fp1, Fp2, F3, Fz, F4, C3, Cz, C4, P7, P3, Pz, P4, P8, O1, Oz, O2 | 0.990 | 0.987 | 0.985 | 0.990 | 0.966 | 0.040 | 0.982 |

*Table 6.3: Identification and authentication performance for different channel configurations.*

## Task-based generalization

While previous experiments focused on session-wise variability, task-based generalization represents an additional challenge. In this setting, the model is trained on EEG data from a subset of tasks and evaluated on unseen tasks, requiring it to generalize across different cognitive tasks rather than across repeated recordings of the same tasks.

Experiment E8 evaluated task-based generalization using the Multi-session Cognitive Tasks datasets. For these datasets, EEG recordings were acquired using only two electrodes, representing a highly constrained and realistic acquisition scenario. A session-based split was also evaluated in a separate experiment (E7), reported in Appendix B.2; here, instead of splitting the data by sessions, the dataset was divided by task type according to the task definitions in Chapter 3. Several configurations were evaluated, and all EEG recordings of each task are included, except when evaluating with trial tasks. In this instance, only five recordings were used, from different sessions in one case and from the same session in the other. It should be noted that the results in this setup may also be influenced by the different amount of data available for each task, as the calibration and evaluation recordings are shorter in duration compared to

the trial tasks. Table 6.4 reports task-based identification and authentication at the signal level.

| Train tasks | Test tasks | ACC (Id) | ACC (Auth) | EER |
|---|---|---|---|---|
| Eval + Trial | Calib | 1.00 | 0.90 | 0.10 |
| Trial | Calib + Eval | 1.00 | 0.89 | 0.15 |
| Calib + Eval | Trial (different sessions) | 0.95 | 0.62 | 0.36 |
| Calib + Eval | Trial (same session) | 0.93 | 0.74 | 0.28 |

*Table 6.4: Task-based identification and authentication results on the Multi-session Cognitive Tasks datasets.*

Perfect identification accuracy is achieved in the first two configurations. In contrast, when the model is trained on calibration and evaluation tasks and tested on trial data, performance decreases notably, particularly in authentication. This degradation is likely due to the higher complexity and non-stationarity of trial tasks, as well as the smaller amount of training data available for calibration and evaluation tasks.

The results demonstrate that EEG-based biometric systems can be generalized across cognitive tasks when task characteristics are sufficiently aligned, even under constrained acquisition conditions. However, task complexity, recording conditions, and the amount of available training data can significantly impact on identification and authentication reliability.

Experiment E9 extended task-based evaluation to the HBN-EEG dataset described in Chapter 3, which includes a larger number of tasks across different sessions. As an initial analysis, a reduced subset of 60 subjects (including 10 impostors) was selected. First, a task-holdout protocol was applied, where the model was trained on all tasks except one and evaluated on the held-out task. Compared to previous datasets, overall performance was noticeably higher, as observed in Table 6.5. In particular, evaluation on Task 4 yielded the best results in general, with most subjects being identified with confidence scores above 80%, as observed in the probability distributions of Figure 6.3.

| Test task | Accuracy (Id) | Accuracy (Auth) | EER |
|---|---|---|---|
| 1. Contrast Change Detection | 1.00 | 1.00 | 0.00 |
| 2. Despicable Me | 1.00 | 0.94 | 0.10 |
| 3. Diary of a Wimpy Kid | 1.00 | 0.92 | 0.20 |
| **4. Fun with Fractals** | **1.00** | **1.00** | **0.00** |
| 5. Resting State | 1.00 | 0.97 | 0.10 |
| 6. Surround Suppression | 1.00 | 1.00 | 0.00 |
| 7. Symbol Search | 1.00 | 0.95 | 0.10 |
| 8. The Present | 1.00 | 0.94 | 0.10 |

*Table 6.5: Identification and authentication performance varying test task on HBN dataset. Identification accuracy is reported as segment-level/signal-level.*

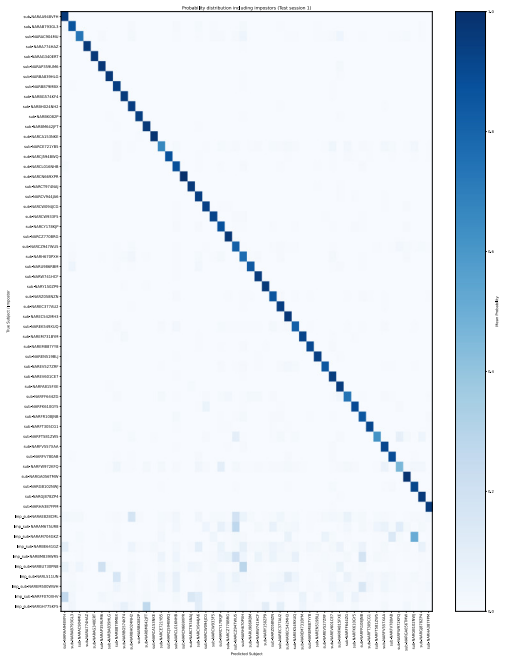


*Figure 6.3: Probability distributions when evaluating on Task 4: Fun with Fractals.*

Second, tasks were grouped into *active* and *passive* categories, following the task descriptions provided in Chapter 3, in order to evaluate cross-task generalization at a higher semantic level.

Based on this categorization, several configurations were explored. These included training the model using all tasks from one category and evaluating it on a task from the opposite category, as well as intra-category generalization, by training on all tasks within a given category except one, which was held out for testing. Finally, a mixed configuration combining tasks from both categories was also considered. For each configuration, the task selected for evaluation corresponds to the task that achieved the highest identification and authentication performance in Table 6.5, ensuring that comparisons focus on the most representative and stable task within each category.

The results, presented in Table 6.6, showed that training on active tasks and evaluating on passive tasks consistently outperformed the other configurations. Probability distributions for this setting revealed sharper subject-wise separation and higher confidence in correct predictions, as can be observed in Figure 6.4.

| **Train / Test** | **Accuracy (Id)** | **Accuracy (Auth)** | **EER** |
|---|---|---|---|
| Passive / Passive | 1.00 | 0.96 | 0.05 |
| Passive / Active | 1.00 | 0.99 | 0.01 |
| Active / Active | 1.00 | 0.97 | 0.05 |
| **Active / Passive** | **1.00** | **0.99** | **0.01** |
| Mix | 1.00 | 0.97 | 0.02 |

*Table 6.6: Identification and authentication performance for different train/test task configurations.*

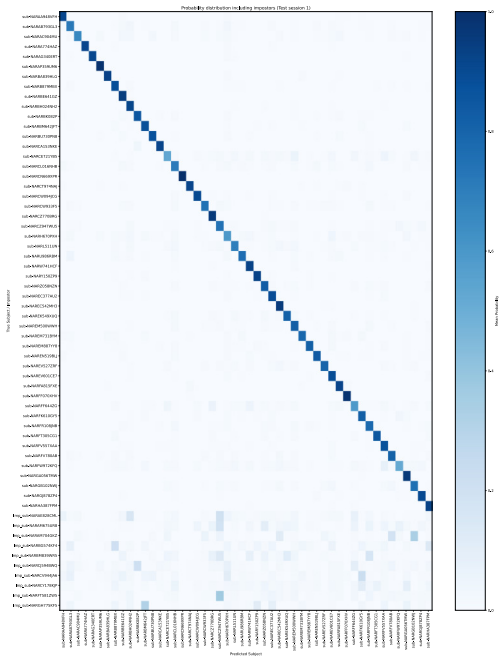


*Figure 6.4: Probability distributions when training on active tasks and evaluating on passive tasks.*

Overall, Experiments E8 and E9 demonstrate that EEG-based biometric systems can generalize across tasks, but performance may depend on task similarity. Moreover, large multi-task datasets such as HBN-EEG appear particularly suitable for learning task-invariant identity representations.

## Model comparison and scalability

After analyzing the impact of sessions, tasks, and channel configurations, the next set of experiments focused on two complementary aspects: model selection and scalability with respect to the number of enrolled subjects. These experiments aim to evaluate whether more expressive models can improve performance under challenging conditions and how the proposed system behaves as the population size increases. In Experiment E10, alternative classification models were evaluated and compared against the baseline Support Vector Machine (SVM) used in all previous experiments. In addition to SVM, a Multilayer Perceptron (MLP) and a Convolutional Neural Network (CNN) were implemented and tested using the Multi-session Cognitive Tasks datasets under the same task-based evaluation protocol described previously for E8, with the three types of tasks: calibration, trial and evaluation tasks.

Table 6.7 summarizes the identification and authentication performance obtained with the three models. While all models achieved comparable identification accuracy, more pronounced differences emerged in authentication performance, particularly in the most challenging task-based split presented in Table 6.4 (training on calibration and evaluation tasks and testing on five trial tasks of diferent sessions).

| Model | Train Data | Test Data | ACC (id) | ACC (auth) | EER |
|---|---|---|---|---|---|
| SVM | Eval + Trial | Cal | 1.00 | 0.90 | 0.10 |
| | Trial | Eval + Cal | 1.00 | 0.89 | 0.15 |
| | Eval + Cal | Trial | 0.95 | 0.62 | 0.36 |
| MLP | Eval + Trial | Cal | 1.00 | 0.93 | 0.10 |
| | Trial | Eval + Cal | 1.00 | 0.89 | 0.10 |
| | Eval + Cal | Trial | 0.98 | 0.75 | 0.30 |
| CNN | Eval + Trial | Cal | 1.00 | 0.92 | 0.10 |
| | Trial | Eval + Cal | 1.00 | 0.90 | 0.11 |
| | Eval + Cal | Trial | 0.97 | 0.84 | 0.20 |

*Table 6.7: Identification and authentication performance across models and train/test data splits.*

The results indicate that neural network-based approaches provide a clear advantage for authentication. The CNN, in particular, achieved the highest authentication accuracy, suggesting that its ability to learn more discriminative and task-invariant representations from PSD inputs is beneficial in complex generalization scenarios. These findings motivated the adoption of CNN-based models in subsequent large-scale experiments.

Experiment E11 evaluated the scalability of the identification and authentication system by progressively increasing the number of enrolled subjects using the HBN-EEG dataset. This experiment represents a key requirement for real-world biometric applications, where systems must operate reliably with hundreds or thousands of users. The CNN model was selected for this experiment due to its superior performance, and results for different population sizes are reported in Table 6.8.

| Nº subjects | Accuracy (Id) | Accuracy (Auth) | EER |
|---|---|---|---|
| 300 | 0.967 | 0.942 | 0.060 |
| 592 | 0.988 | 0.964 | 0.040 |
| 1106 | 0.982 | 0.975 | 0.020 |
| 3059 | 0.958 | 0.943 | 0.060 |

*Table 6.8: Scalability analysis on the HBN-EEG dataset using a CNN model.*

The results demonstrate that the proposed system maintains high identification and authentication performance even as the number of subjects increases to several thousand. Identification accuracy remains above 0.95 in all configurations, while authentication accuracy stays above 0.94 for most population sizes. Notably, classification performance does not degrade monotonically as the number of training subjects increases. Instead, the best results were obtained for intermediate population sizes (approximately 1100 subjects), with an Equal Error Rate as low as 0.02. This outcome is likely attributable to a balance between model capacity and inter-subject similarity. In scenarios where the number of subjects is too limited, the model is

unable to acquire the necessary data to learn robust, generalisable features for subject discrimination. However, beyond an optimal point, the addition of further subjects can increase the probability of encountering individuals sharing highly similar patterns of neural activity in the training set. This can make it more difficult for the model to establish and maintain precise, separable decision boundaries for every individual, thus explaining the slight performance degradation observed when training with 3059 subjects.

# Chapter 7

# Signal anonymisation

While EEG signals have demonstrated strong potential as a biometric modality, their use raises significant concerns regarding privacy, as discussed in Section 2.2. This chapter explores signal anonymisation strategies that are designed to minimize the amount of identity-related information present in EEG recordings, while maintaining their utility for non-biometric tasks. These strategies are built upon the datasets, pipeline, and methodology outlined in previous chapters.

## 7.1 Problem definition

Signal anonymisation aims to limit identity leakage directly at the signal or feature level. Formally, let $\mathbf{x}$ denote an EEG signal segment and let $f_{\text{id}}(\mathbf{x})$ represent an identity recognition model. An anonymisation function $g(\cdot)$ seeks to produce a transformed signal $\tilde{\mathbf{x}} = g(\mathbf{x})$ such that the performance of $f_{\text{id}}(\tilde{\mathbf{x}})$ is significantly degraded, while the performance of task-specific models $f_{\text{task}}(\tilde{\mathbf{x}})$ remains as close as possible to that obtained on the original signal.

Identity suppression is quantified using identification and authentication metrics, while the preservation of signal utility is evaluated using classification accuracy in auxiliary tasks unrelated to identity.

## 7.2 Experimental protocol

Anonymisation experiments follow the same general system pipeline described in Chapter 4, with the addition of a signal transformation stage applied prior to feature extraction.

### 7.2.1 Exploration of signal-level perturbations

Several signal-level anonymisation strategies were initially considered and explored, including stochastic perturbations and dimensionality reduction techniques aimed at suppressing identity-related variance. Of these, two approaches were found to be particularly effective in reducing identification performance:

- Additive noise perturbations, where controlled noise is introduced into the EEG signal to obscure discriminative patterns.
- Subspace-based obfuscation, where the signal is projected onto a reduced-dimensional space and subsequently reconstructed, discarding part of the original information.

However, such generic perturbations provide limited interpretability and may affect a broad range of signal components, potentially degrading performance in other EEG-based tasks. This can be observed in the exploratory analyses reported in Appendix B.3. For this reason, the anonymisation strategy was subsequently refined to incorporate domain knowledge and task-specific frequency characteristics, enabling more targeted and controlled modifications.

### 7.2.2 Frequency-selective noise injection

The primary anonymisation strategy adopted in this work is based on the controlled injection of noise within selected frequency ranges of the EEG signal. This approach is motivated by two main considerations. It should be noted that EEG-based biometric systems rely heavily on spectral features, meaning that frequency-domain perturbations are particularly effective for suppressing identity-related information. Secondly, different EEG-related tasks are known to depend on different frequency bands. This suggests that selective perturbations may affect identity recognition more strongly than other neural decoding tasks.

The anonymisation procedure is applied independently to each EEG segment and each channel. For a given segment and channel, the time-domain EEG signal is transformed into its frequency representation using a Fourier transform. This representation separates the signal into two components: a magnitude spectrum, which captures the energy distribution across frequencies, and a phase spectrum, which encodes temporal information.

Noise is selectively introduced into the magnitude spectrum at a predefined set of frequency bins corresponding to the frequency bands under analysis. These frequency bins are selected according to the target anonymisation configuration and remain fixed across all segments within each experiment. For each selected frequency component, the magnitude is perturbed by adding a noise term whose amplitude is proportional

to the original spectral magnitude:

$$\tilde{M}_i = M_i + \epsilon_i, \tag{7.1}$$

where

$$\epsilon_i = M_i \cdot \alpha \cdot \mathcal{N}(0, 1). \tag{7.2}$$

In this formulation, $\alpha$ defines the *noise level* and controls the relative strength of the perturbation, while $\mathcal{N}(0, 1)$ denotes a standard Gaussian random variable. Defining the noise in relative terms ensures that frequency components with higher energy are perturbed more strongly, leading to an energy-adaptive anonymisation process.

The phase spectrum is preserved unchanged throughout the process. After perturbation, the modified magnitude spectrum is recombined with the original phase information, and the anonymised EEG segment is reconstructed by applying the inverse Fourier transform. This design preserves the overall temporal structure of the signal while selectively altering its spectral content in the targeted frequency ranges.

By varying both the selected frequency ranges and the noise level parameter $\alpha$, this method enables a systematic exploration of the trade-off between identity suppression and task preservation. Anonymisation is applied directly to each raw EEG segment prior to feature extraction, ensuring that its effects propagate consistently through the entire processing pipeline and can be reliably evaluated in downstream biometric and non-biometric tasks.

### 7.2.3 Use of auxiliary EEG classification tasks

To assess whether anonymisation strategies selectively degrade identity-related information while preserving task-relevant EEG features, several auxiliary EEG classification models were employed. These models address non-biometric tasks and serve as reference points to evaluate signal utility after anonymisation.

Initial anonymisation strategies based on additive Gaussian noise and PCA projection were evaluated using two simple classification tasks: discrimination between active and passive cognitive tasks, and resting-state prediction. In the frequency-selective anonymisation experiments, an additional eyes-open versus eyes-closed classification task was implemented. This task is strongly associated with changes in specific EEG frequency bands, particularly within the alpha range, and therefore provides a meaningful benchmark for evaluating frequency-dependent perturbations.

Finally, the effects of anonymisation were evaluated using a sleep stage classification model developed by Bitbrain, representing a real-world industrial application of EEG analysis.

### 7.2.4 Anonymisation evaluation

For each anonymisation configuration, identification and authentication experiments are conducted following the same protocols described in Chapter 6. To evaluate task preservation, anonymised signals are also used to train and test task-specific classifiers. Task performance is measured using classification accuracy.

A summary of the experimental configurations considered for anonymisation evaluation is provided in Table 7.1.

| Id | Dataset | Subjects | | Type | Sessions | Tasks | Channels | Motivation |
|---|---|---|---|---|---|---|---|---|
| | | Genuine | Impostors | | | | | |
| **Generic signal-level anonymisation baselines** | | | | | | | | |
| E12 | HBN-EEG | 300 | 30 | Task-based | 1 (1) | ∼8 (∼8) | 129 | Anonymisation with Gaussian noise and PCA. |
| **Frequency-selective anonymisation guided by task preservation** | | | | | | | | |
| E13 | HBN-EEG | 300 | 30 | Task-based | 1 (1) | ∼8 (∼8) | 129 | Frequency-band noise injection and comparison with open/closed eyes model. |
| **Frequency-selective anonymisation under realistic deployment scenarios** | | | | | | | | |
| E14 | Sleep | ∼30 | 5 | Session-based | 3 (3) | 1 (1) | 2 | Noise in test data and comparison with sleep stage classification model. |
| E15 | Sleep | ∼30 | 5 | Session-based | 3 (3) | 1 (1) | 2 | Effect of noise in training and evaluation, comparing to sleep stage classification model. |

*Table 7.1: Summary of configurations for anonymisation experiments evaluated in this project. Values in parentheses indicate the number of sessions or tasks available for impostors.*

## 7.3 Results and discussion

The objective of the anonymisation experiments is to determine whether EEG signals can be transformed so that biometric identifiability is substantially weakened while their utility for non-biometric applications is preserved. This scenario reflects a realistic use case in which EEG recordings should be stored or shared without revealing the identity of the subject, yet remaining suitable for downstream analyses such as behavioural monitoring or clinical assessment.

The experiments were organized progressively. First, generic and task-agnostic transformations were explored, as explained in Section 7.2.1. The results showed that generic perturbations offer limited control over the privacy–utility trade-off, motivating the development of a more structured anonymisation strategy based on frequency-selective perturbations, which constitutes the focus of the following experiments discussed in this section.

## Frequency-selective anonymisation guided by task preservation

Experiment E13 aims to investigate whether identity-related information could be selectively attenuated by injecting noise into specific EEG frequency bands, while minimally affecting a non-biometric classification task. To evaluate this idea, biometric identification and authentication were assessed jointly with an auxiliary eyes open / eyes closed (EO/EC) classifier, a well-established paradigm with clear spectral correlates. The anonymisation was applied consistently during both training and evaluation of all models, with noise injected over the frequencies of PSD features extracted from signals of HBN-EEG dataset.

The baseline EO/EC classifier trained with clean signals achieved an accuracy of 0.748. After introducing frequency-specific noise at varying levels and across different frequency bands, a differentiated impact was observed between biometric and non-biometric tasks, as summarized in Table 7.2.

| Noise level | Frequencies | ACC (Id) | ACC (Auth) | EER | ACC (EO/EC) |
|---|---|---|---|---|---|
| 50 | Delta | 0.93 | 0.89 | 0.11 | 0.70 |
| | Theta | 0.87 | 0.86 | 0.14 | 0.70 |
| | Alpha | 0.63 | 0.76 | 0.26 | 0.68 |
| | **Beta** | **0.21** | **0.65** | **0.37** | **0.68** |
| | Gamma | 0.34 | 0.65 | 0.35 | 0.71 |
| 20 | Beta | 0.65 | 0.73 | 0.25 | 0.73 |
| | Beta, Delta | 0.51 | 0.72 | 0.32 | 0.70 |
| | Beta, Gamma | 0.31 | 0.68 | 0.32 | 0.70 |
| | Theta, Gamma | 0.52 | 0.68 | 0.34 | 0.70 |
| 10 | **Beta, Gamma** | **0.50** | **0.71** | **0.29** | **0.73** |
| | Alpha, Beta, Gamma | 0.33 | 0.67 | 0.34 | 0.69 |
| 5 | Alpha, Beta, Gamma | 0.72 | 0.77 | 0.25 | 0.72 |
| | Delta, Theta, Beta, Gamma | 0.64 | 0.75 | 0.26 | 0.73 |
| | All | 0.45 | 0.72 | 0.30 | 0.70 |

*Table 7.2: Impact of frequency-selective noise injection on biometric performance and EO/EC classification.*

Perturbations concentrated in the beta band produced the largest decrease in identification accuracy, down to 0.21 at the highest noise level, which is a more substantial decrease in accuracy than in the case of EO/EC, where it remained at 0.68. Conversely, the impact of noise injected into the delta or theta bands was found to have a limited effect on identification, in addition to a reduction in EO/EC accuracy by a comparable amount. This suggests that these bands are more critical for the auxiliary

task.

A comparable pattern was observed for combinations of bands at lower noise magnitudes. It should be noted that configurations which preserved the alpha band while perturbing beta and gamma frequencies achieved a significant reduction in biometric performance, with a much smaller impact on EO/EC classification. These results initially suggest that identity-discriminating information is not uniformly distributed across the spectrum and that anonymisation can be guided by prior neurophysiological knowledge about the spectral relevance of the target task.

## Frequency-selective anonymisation under realistic deployment scenarios

As demonstrated in previous experiments, biometric information can be attenuated through frequency-selective perturbations in a controlled setting. However, those analyses relied on the HBN-EEG dataset and a relatively simple auxiliary task. Experiments E14 and E15 extend this investigation to a more realistic scenario based on sleep EEG recordings, where the non-biometric objective, sleep stage classification, is directly relevant to clinical applications. The goal is to evaluate whether EEG signals can be stored in an anonymised form that limits identity leakage while still supporting reliable hypnogram estimation.

To explore this trade-off under practical deployment conditions, the anonymisation strategy was refined further by modifying the stage at which noise was applied. Instead of injecting perturbations directly into the power spectral density (PSD), noise was added to the raw EEG signals prior to feature extraction. This change was motivated by the need to ensure methodological consistency with the proprietary sleep stage classification pipeline, enabling a more meaningful comparison across tasks. Subsequently, two complementary configurations were defined. Experiment E14 analyses the case in which noise is applied only at evaluation time. In this setting, biometric identification models are trained on clean signals from the Sleep dataset and evaluated on perturbed Sleep data, while the sleep stage classifier is trained on clean BOAS recordings and evaluated on perturbed signals from the Sleep dataset. This configuration assesses the immediate impact of anonymisation on pre-existing systems and highlights the vulnerability of current pipelines to posterior signal transformations. Alternatively, Experiment E15 addresses a different scenario, in which anonymised signals are available from the outset. In this case, noise is present during both training and evaluation. Identification models are trained and evaluated on perturbed Sleep recordings, while sleep stage classification is trained and evaluated on perturbed BOAS data, allowing the models to adapt to the altered signal distribution.

Using the Sleep dataset for E14, frequency-selective noise was applied exclusively at evaluation time. For the sleep staging task, relative performance loss is computed with respect to the accuracy obtained when the classifier trained on clean BOAS data is evaluated on clean Sleep recordings, which is taken as the 100% reference

performance. Table 7.3 summarizes the results for multiple noise magnitudes and spectral configurations.

| **Noise level** | **Frequencies** | **Identification** | | **Authentication** | | **Sleep** | |
|---|---|---|---|---|---|---|---|
| | | ACC | RPL (%) | ACC | RPL (%) | ACC | RPL (%) |
| 5 | Delta | 72.48 | 25.27 | 66,96 | 16.29 | 12.67 | 87.33 |
| | Theta | 11.19 | 88.46 | 69.92 | 12.59 | 65.92 | 34.08 |
| | Alpha | 25.05 | 74.17 | 62.04 | 22.44 | 73.54 | 26.46 |
| | Beta | 4.80 | 95.05 | 55.68 | 30.39 | 56.10 | 43.90 |
| | Gamma | 30.38 | 68.68 | 62.10 | 22.37 | 82.12 | 17.88 |
| 3 | Delta | 87.40 | 9.89 | 68.79 | 14.01 | 45.5 | 54.50 |
| | Theta | 15.98 | 83.52 | 69.40 | 13.25 | 74.37 | 25.63 |
| | Alpha | 29.31 | 69.78 | 64.47 | 19.41 | 77.76 | 22.24 |
| | Beta | 9.59 | 90.11 | 56.56 | 29.30 | 76.17 | 23.83 |
| | Gamma | 41.04 | 57.69 | 64.34 | 19.57 | 88.81 | 11.19 |
| 1 | Theta | 63.96 | 34.06 | 77.92 | 2.60 | 93.69 | 6.31 |
| | Alpha | 60.76 | 37.36 | 67.06 | 16.17 | 94.47 | 5.53 |
| | **Beta** | **39.96** | **58.80** | **59.64** | **25.44** | **93.77** | **6.23** |
| | Gamma | 80.47 | 17.04 | 73.32 | 8.35 | 97.45 | 2.55 |
| | Beta, Gamma | 51.16 | 47.25 | 61.80 | 22.74 | 92.72 | 7.28 |
| | Alpha, Beta, Gamma | 35.70 | 63.19 | 59.49 | 25.63 | 88.63 | 11.37 |
| | Theta, Alpha, Gamma | 51.91 | 46.48 | 64.94 | 18.82 | 95.19 | 4.81 |
| | All | 39.44 | 59.34 | 61.60 | 22.99 | 83.84 | 16.16 |
| 0.5 | Theta | 88.99 | 8.25 | 78.41 | 1.98 | 97.61 | 2.39 |
| | Alpha | 85.27 | 12.09 | 73.96 | 7.55 | 97.86 | 2.14 |
| | Beta | 70.34 | 27.48 | 67.99 | 15.01 | 97.68 | 2.32 |
| | Gamma | 95.93 | 1.10 | 79.23 | 0.96 | 98.86 | 1.14 |
| | Beta, Gamma | 68.74 | 29.13 | 69.28 | 13.40 | 97.38 | 2.62 |
| | **Alpha, Beta, Gamma** | **52.76** | **45.60** | **64.57** | **19.28** | **96.1** | **3.90** |
| | Theta, Alpha, Gamma | 85.99 | 11.35 | 73.64 | 7.92 | 97.21 | 2.79 |
| | All | 53.29 | 45.06 | 65.45 | 18.18 | 93.2 | 6.80 |
| 0.3 | Beta | 91.14 | 6.04 | 74.59 | 6.76 | 98.88 | 1.12 |
| | Beta, Gamma | 89.54 | 7.69 | 74.26 | 7.17 | 98.71 | 1.29 |
| | Alpha, Beta | 85.27 | 12.09 | 71.76 | 10.29 | 98.19 | 1.81 |
| | Alpha, Beta, Gamma | 77.78 | 19.78 | 69.91 | 12.61 | 98.02 | 1.98 |
| | All | 76.21 | 21.43 | 70.93 | 11.33 | 95.67 | 4.33 |

*Table 7.3: Performance of identification, authentication, and sleep stage classification when frequency-selective noise is applied only in evaluation data.*

The impact of perturbations was highly dependent on the affected frequency band. Noise injected into the delta band produced a dramatic reduction in sleep stage performance, with relative loss above 80% for the highest noise level. This behaviour is consistent with the central role of slow-wave activity in sleep staging and confirms that delta components cannot be heavily distorted without compromising clinical utility. In contrast, perturbations concentrated in the beta band led to a severe degradation of biometric identification, with a higher difference in relative performance loss in comparison with sleep classification. Similar trends were observed for combinations of bands that excluded delta, where identification accuracy dropped sharply whereas the sleep task exhibited moderate or limited deterioration.

At lower noise magnitudes, such as 0.5, the dissociation between tasks became more evident. Several configurations, particularly those that include beta band, reduced identification performance by 20–30% while maintaining sleep accuracy above 96%, corresponding to less than 4% relative loss.

In Experiment E15, frequency-selective noise was incorporated consistently during the training and evaluation of all models. Identification and authentication models were trained and evaluated on perturbed Sleep recordings, while the sleep stage classifier was trained and tested on perturbed BOAS data. In this case, relative performance loss for sleep staging is computed with respect to the baseline accuracy obtained when training and evaluating the classifier on clean BOAS signals (85.7%).

Table 7.4 provides a summary of the obtained results. In comparison with E14, where noise was introduced exclusively at evaluation time, the degradation of biometric performance is noticeably attenuated. This behaviour is expected: when exposed to perturbed signals during training, the models can partially adapt and learn representations that compensate for the added noise. Consequently, significantly higher noise magnitudes were necessary to achieve reductions in identification accuracy analogous to those observed in E14.

Despite this adaptation effect, a clear dissociation between tasks remains observable. Configurations affecting theta, alpha, beta, and gamma bands while preserving delta activity consistently produced larger relative losses in identification than in sleep staging. For instance, at noise level 30, combinations such as *Theta, Alpha, Beta* reduced identification accuracy by more than 17%, while sleep performance decreased by only 5%. Conversely, elevated noise levels (e.g., $\geq 100$) resulted in a substantial degradation of both tasks, indicating that excessive perturbation can eventually result in the destruction of task-relevant information.

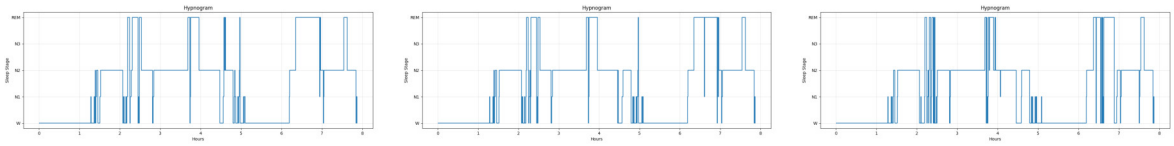
Empirically, configurations resulting in more than approximately 6% relative performance loss in the sleep task were associated with visibly distorted hypnograms and unstable stage transitions (Figure 7.1). Therefore, the objective cannot be maximal biometric suppression, but rather the determination of operating points that ensure sleep degradation remains below this limit. Within this constraint, noise levels of 30, 10, and 5 emerge as favourable trade-off regions. In these configurations, identification accuracy is reduced by approximately 10–20%, while sleep performance remains within

| Noise level | Frequencies | Identification | | Authentication | | Sleep | |
|---|---|---|---|---|---|---|---|
| | | ACC | RPL (%) | ACC | RPL (%) | ACC | RPL (%) |
| 300 | Alpha | 66.00 | 31.95 | 58.18 | 27.27 | 68.98 | 19.50 |
| | Theta, Alpha, Gamma | 65.77 | 33.19 | 61.30 | 23.37 | 76.53 | 10.70 |
| 200 | Alpha, Gamma | 76.00 | 21.64 | 64.41 | 19.48 | 74.98 | 12.50 |
| | Theta, Alpha, Gamma | 69.66 | 28.18 | 61.10 | 23.62 | 76.57 | 10.65 |
| 100 | Theta, Alpha, Gamma | 74.49 | 23.20 | 62.66 | 21.67 | 77.33 | 9.76 |
| 50 | Theta, Alpha, Gamma | 77.17 | 20.44 | 66.33 | 17.08 | 78.34 | 8.58 |
| | Theta, Alpha, Beta | 81.58 | 15.42 | 71.24 | 10.95 | 79.00 | 7.05 |
| 30 | Theta, Alpha, Gamma | 81.00 | 16.49 | 64.72 | 19.10 | 80.10 | 6.53 |
| | **Theta, Alpha, Beta** | **79.47** | **17.53** | **70.68** | **11.65** | **81.40** | **5.01** |
| | Alpha, Beta, Gamma | 77.37 | 20.23 | 70.43 | 11.96 | 81.80 | 4.55 |
| 10 | Theta, Alpha, Gamma | 85.20 | 12.16 | 65.45 | 18.18 | 81.80 | 4.55 |
| | Theta, Alpha, Beta | 83.68 | 13.32 | 70.60 | 11.75 | 82.80 | 3.38 |
| | Alpha, Beta, Gamma | 86.32 | 11.01 | 68.88 | 13.90 | 83.10 | 3.03 |
| 5 | **Theta, Alpha, Gamma** | **94.21** | **2.87** | **68.48** | **14.40** | **84.10** | **1.86** |
| | Theta, Alpha, Beta | 87.37 | 9.63 | 71.24 | 10.95 | 84.40 | 1.51 |
| | Alpha, Beta, Gamma | 90.53 | 6.67 | 71.51 | 10.61 | 83.30 | 2.80 |

*Table 7.4: Performance of identification, authentication, and sleep stage classification when frequency-selective noise is applied during both training and evaluation.*

the acceptable range. It is important to note that a reduction of this magnitude is sufficient to bring biometric performance below the thresholds typically required for a reliable identification system, effectively compromising its practical usability.

Overall, this demonstrates that frequency-selective noise injection, when applied consistently during both training and evaluation, enables controllable and interpretable trade-offs between privacy and utility. Rather than seeking to achieve complete identity removal, the proposed approach aims to compromise biometric reliability to



*(a) No noise* *(b) Relative loss = 3.90* *(c) Relative loss = 6.23*

*Figure 7.1: Effect of increasing noise levels on hypnogram structure.*

a point where identification becomes impractical, while ensuring that EEG retains its functionality for clinical analysis. This balance may provide a realistic path towards the privacy-preserving storage of neurophysiological data.

# Chapter 8

# Conclusions

This study has presented a detailed investigation into the use of electroencephalography (EEG) signals as a biometric modality, exploring three key concepts: feasibility, robustness, and privacy. The primary objective of this project was to develop a comprehensive framework for EEG-based subject identification and authentication, while addressing the significant privacy concerns that are associated with the handling of neural data. The work successfully integrates theoretical research with practical implementation, resulting in a reliable and scalable biometric process, along with techniques to address the issue of data privacy.

The main achievement is the design and implementation of a complete and modular EEG biometric system, where the architecture is inherently flexible and supports both identification and authentication.

Experimental validation is a fundamental contribution of this thesis. The system has been evaluated under a range of conditions to determine its practical feasibility. Using a diverse collection of public and private datasets, the research has provided key insights:

- **Inter-session robustness:** A primary challenge for EEG biometrics is signal variability between different recording sessions. The system has been explicitly evaluated under session generalisation protocols, in which models were trained in some sessions and evaluated in completely different ones. The result obtained was that the system was capable of identifying and authenticating subjects with a very high degree of accuracy, despite the fact that the recordings were distributed over a period of time.
- **Sensor efficiency and channel optimization:** Practical implementation necessitates minimal, user-friendly hardware. A systematic analysis of channel contribution was conducted, revealing that discriminatory subject information is not uniformly distributed across the scalp. The analysis indicated that central electrodes appeared to be the most informative. It is important to note that the system demonstrated high performance even with very limited configurations: reliable identification and authentication were possible using only a single channel

or a compact subset. This results in a substantial reduction in system complexity, cost, and setup time, thereby enhancing its practical applicability.

- **Generalization across cognitive tasks:** Beyond temporal stability, the system's capacity to recognise individuals when performing different cognitive tasks has also been evaluated. In cross-task experiments, the models, particularly the CNN, exhibited remarkable generalisation ability. The findings demonstrated that performance remained high, indicating that learned identity representations are not strictly tied to specific task-related brain activations. Rather, they appear to capture more fundamental, subject-specific neural features.

- **Scalability to Large Populations:** A critical criterion for evaluating the efficacy of any biometric system is its performance as the registered population increases. Experiments conducted with the large-scale HBN-EEG dataset confirmed the scalability of the system. The performance did not collapse; instead, identification accuracy remained above 95% and authentication EER below 6%.

A further significant contribution of this project is the exploration of EEG signal anonymisation. Recognising that EEG is a particularly sensitive biometric data, the present work explores solutions to preserve the usefulness of the data while protecting the privacy of the subjects. In order to achieve this objective, an interpretable methodology of selective frequency noise injection was developed. This approach was founded on the neurophysiological knowledge that different EEG bands are associated with different functions. The primary conclusion drawn from this analysis is that information related to identity and usefulness for other tasks can be dissociated into distinct spectral patterns. For instance, the injection of noise in the beta band resulted in a substantial deterioration of biometric performance, while showing a comparatively minor effect on the open-eye/closed-eye classification task. In contrast, perturbation of the delta band resulted in compromised sleep stage classification, a process dependent on slow wave activity, while biometric identification maintained a reasonable degree of accuracy.

By carefully selecting the frequency bands and magnitude of the perturbations, it has been possible to obtain EEG signals that significantly compromise re-identification while maintaining high accuracy in sleep stage classification. This result is of great importance to the neurotechnology industry and biomedical research. It provides a concrete and robust pathway for companies such as Bitbrain to store, share, or use EEG datasets without preserving their biometric utility. In the event of a data breach, these anonymised signals would be much less likely to compromise individual identity, thus aligning technological innovation with ethical data management.

In summary, this project makes a contribution to the fields of biometrics and neurotechnology. The findings demonstrate that EEG-based identification and authentication are not only feasible, but can also be robust, scalable and efficient. Moreover, it addresses the significant ethical and practical challenges posed by the

privacy of EEG data. The work thus establishes a solid foundation for the development of secure and privacy-preserving brain-computer interfaces.

# Future Work

This research opens up new directions for future work that could also be explored:

- **Advanced model architecture and hyperparameter optimisation:** Although CNN demonstrated high performance, a systematic and exhaustive search for hyperparameters in the tested models could yield further enhancements. Exploration of more advanced architectures, such as transformer-based models or hybrid CNN-LSTM networks, could capture long-range temporal dependencies more effectively and potentially improve robustness against noise and variability.
- **Adversarial anonymization frameworks:** The idea behind adversarial anonymization is to generate more complex, data-driven noise that improves privacy while preserving utility. The current frequency selection method, although effective, is adjusted manually. A more powerful and automated approach would involve the training of an adversarial neural network or an autoencoder specialised in privacy preservation. The objective would be to develop a transformation function that directly optimises the balance between privacy and utility. This framework would automatically discover the optimal perturbation, which could achieve greater anonymisation with less loss of utility than manually designed spectral filters.
- **Longitudinal stability and template updating mechanisms:** A long-term study is essential to establish the stability of brainprints over several years. Furthermore, the development of adaptive algorithms that can gradually update the user's biometric template over time to account for slow neural changes due to ageing, learning or lifestyle would be crucial for lifelong biometric systems.
- **Large-scale population studies:** While this work demonstrates the feasibility and robustness of EEG-based biometric identification on datasets ranging from tens to several thousand subjects, future research should extend the evaluation to substantially larger populations, well beyond the 3000-subject scale, enabling a more meaningful contextual comparison with established biometric modalities in terms of scalability, error rates, and operational constraints.

# Chapter 9

# Bibliography


[1] A.K. Jain, A. Ross, and S. Prabhakar. An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):4–20, 2004.

[2] Emanuela Marasco and Arun Ross. A survey on antispoofing schemes for fingerprint recognition systems. *ACM Comput. Surv.*, 47(2), November 2014.

[3] Ramaswamy Palaniappan. Two-stage biometric authentication method using thought activity brain waves. *International journal of neural systems*, 18:59–66, 03 2008.

[4] Emanuele Maiorana, Daria La Rocca, and Patrizio Campisi. On the permanence of eeg signals for biometric recognition. *IEEE Transactions on Information Forensics and Security*, 11(1):163–175, 2016.

[5] Kun Xia, Wlodzislaw Duch, Yu Sun, Kedi Xu, Weili Fang, Hanbin Luo, Yi Zhang, Dong Sang, Xiaodong Xu, Fei-Yue Wang, and Dongrui Wu. Privacy-preserving brain–computer interfaces: A systematic review. *IEEE Transactions on Computational Social Systems*, 10(5):2312–2324, 2023.

[6] Agencia Española de Protección de Datos (AEPD) & European Data Protection Supervisor (EDPS). Techdispatch sobre neurodatos. Techdispatch, AEPD / EDPS, 2024.

[7] Daniel Hain, Roman Jurowetzki, Mariagrazia Squicciarini, and Lihui Xu. *Unveiling the Neurotechnology Landscape: Scientific Advancements, Innovations and Major Trends*. 07 2023.

[8] Julien Mahier, Marc Pasquet, Christophe Rosenberger, and Félix Cuozzo. Biometric authentication. In *Encyclopedia of Information Science and Technology*, pages 1–12. IGI Global, August 2008.

[9] Wencheng Yang, Song Wang, Jiankun Hu, Guanglou Zheng, and Craig Valli. Security and accuracy of fingerprint-based biometrics: A review. *Symmetry*, 11(2), 2019.

[10] Andrea Biasiucci, Benedetta Franceschiello, and Micah M. Murray. Electroencephalography. *Current Biology*, 29(3):R80–R85, 2019.

[11] Mohammed Abo-Zahhad, Sabah Ahmed, and Sherif Nagib Seha. A new eeg acquisition protocol for biometric identification using eye blinking signals. *International Journal of Intelligent Systems and Applications (IJISA)*, 07:48–54, 05 2015.

[12] The Bitbrain team. The most important features of eeg systems explained, Aug 2025.

[13] Marcos Del Pozo-Banos, Jesus B. Alonso, Jaime R. Ticay-Rivas, and Carlos M. Travieso. Electroencephalogram subject identification: A review. *Expert Systems With Applications*, 41(15):6537–6554, May 2014.

[14] Mahsa Zeynali and Hadi Seyedarabi. Eeg-based single-channel authentication systems with optimum electrode placement for different mental activities. *Biomedical Journal*, 42(4):261–267, 2019.

[15] Yao-Yuan Yang, Angel Hsing-Chi Hwang, Chien-Te Wu, and Tsung-Ren Huang. Person-identifying brainprints are stably embedded in eeg mindprints. *Scientific Reports*, 12(1):17031, 2022.

[16] Jordan Ortega-Rodriguez, Kevin Martin-Chinea, Jose Francisco Gomez-Gonzalez, and Ernesto Pereda. Brainprint based on functional connectivity and asymmetry indices of brain regions. *IET Biometrics*, 12(3):129–145, 2023.

[17] Yingnan Sun, Frank P.-W. Lo, and Benny Lo. Eeg-based user identification system using 1d-convolutional long short-term memory neural networks. *Expert Systems With Applications*, 125:259–267, 2019.

[18] Ali Seyfizadeh, Robert L. Peach, Philip Tovote, Ioannis U. Isaias, Jens Volkmann, and Muthuraman Muthuraman. Enhancing security in brain–computer interface applications with deep learning. *Expert Systems With Applications*, 253:124218, 2024.

[19] Walaa Alsumari, Muhammad Hussain, Laila Alshehri, and Hatim A. Aboalsamh. Eeg-based person identification and authentication using deep convolutional neural network. *Axioms*, 12(1):74, 2023.

[20] Renata Plucinska, Konrad Jedrzejewski, Urszula Malinowska, and Jacek Rogala. Leveraging multiple distinct eeg training sessions for improvement of spectral-based biometric verification results. *Sensors*, 23(4):2057, 2023.

[21] Emanuele Maiorana and Patrizio Campisi. Longitudinal evaluation of eeg-based biometric recognition. *IEEE Transactions on Information Forensics and Security*, 13(5):1123–1138, 2018.

[22] Ary L. Goldberger, Luis A. N. Amaral, Leon Glass, Jeffrey M. Hausdorff, Plamen Ch. Ivanov, Roger G. Mark, Joseph E. Mietus, George B. Moody, Chung-Kang Peng, and H. Eugene Stanley. Physiobank, physiotoolkit, and physionet. *Circulation*, 101(23):e215–e220, 2000.

[23] Lubin Meng, Xue Jiang, Tianwang Jia, and Dongrui Wu. Protecting multiple types of privacy simultaneously in eeg-based brain-computer interfaces, 2024.

[24] Xiaoqing Chen, Siyang Li, Yunlu Tu, Ziwei Wang, and Dongrui Wu. User-wise perturbations for user identity protection in eeg-based bcis, November 2024.

[25] Shiya Liu, Yue Yao, Chaoyue Xing, and Tom Gedeon. Disguising personal identity information in eeg signals, October 2020.

[26] Shouvik Paul and Garima Bajwa. Privacy-preserving eeg data generation: A federated split learning approach using privacy-adaptive autoencoders and secure aggregation with gflownet, 2025.

[27] Amir Jalaly Bidgoly, Hamed Jalaly Bidgoly, and Zeynab Arezoumand. Towards a universal and privacy preserving eeg-based authentication system. *Scientific Reports*, 12(1):2531, February 2022.

[28] Gerwin Schalk, Dennis Mcfarland, Thilo Hinterberger, NR Birbaumer, and Jonathan Wolpaw. Bci2000: a general-purpose brain-computer interface (bci) system. *IEEE Trans. Biomed. Eng.*, 51:1034–, 07 2004.

[29] Michael Tangermann, Klaus-Robert Müller, Ad Aertsen, Niels Birbaumer, Christoph Braun, Clemens Brunner, Robert Leeb, Carsten Mehring, Kai J. Miller, Gernot Mueller-Putz, Guido Nolte, Gert Pfurtscheller, Hubert Preissl, Gerwin Schalk, Alois Schlögl, Carmen Vidaurre, Stephan Waldert, and Benjamin Blankertz. Review of the bci competition iv. *Frontiers in Neuroscience*, Volume 6 - 2012, 2012.

[30] Seyed Yahya Shirazi, Alexandre Franco, Mauricio Scopel Hoffmann, Nathalia Esper, Dung Truong, Arnaud Delorme, Michael Milham, and Scott Makeig. HBN-EEG: The FAIR implementation of the healthy brain network (HBN) electroencephalography dataset. *bioRxiv*, page 2024.10.03.615261, 3 October 2024.

[31] Carlos Escolano, Mayte Navarro-Gil, Javier Garcia-Campayo, Marco Congedo, Dirk De Ridder, and Javier Minguez. A controlled study on the cognitive effect of alpha neurofeedback training in patients with major depressive disorder. *Frontiers in Behavioral Neuroscience*, Volume 8 - 2014, 2014.

[32] Eduardo López-Larraz, María Sierra-Torralba, Sergio Clemente, Galit Fierro, David Oriol, Javier Minguez, Luis Montesano, and Jens G. Klinzing. "bitbrain open access sleep dataset", 2025.

# List of Figures

# List of Tables

# Appendices

# Appendix A

# Project planning and timeline

This appendix presents the planning and timeline followed during the development of this project. Figure A.1 shows the Gantt chart summarizing the main tasks and their temporal distribution.

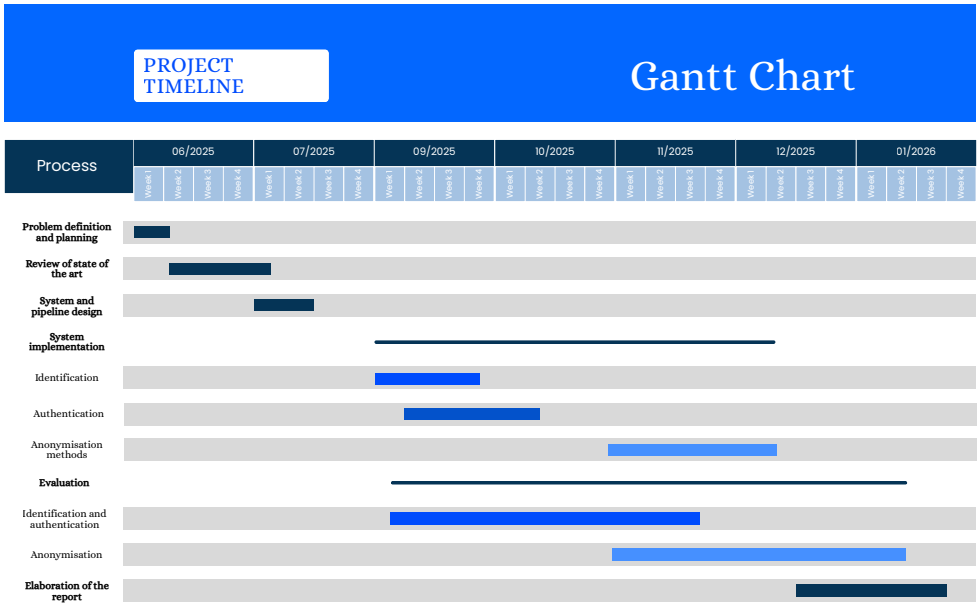



*Figure A.1: Gantt chart of the project.*

# Appendix B

# Additional Experiments

## B.1 Experiments E1, E2 and E3: Baseline EEG-based identification

This appendix reports a set of preliminary experiments aimed at evaluating the basic feasibility of EEG-based person identification under simplified conditions. These experiments focus exclusively on multi-class identification and do not include authentication protocols or advanced generalization strategies. Their primary purpose is to establish that subject-specific information can be reliably extracted from EEG signals across datasets with different characteristics.

Three publicly available datasets with progressively more realistic conditions were considered: PhysioNet EEG Motor Imagery, BCI Competition IV 2b, and the Elevvo Depression dataset. Identification accuracy was evaluated at the segment level using a multi-class SVM classifier. The main characteristics and results of these experiments are summarized in Table B.1.

| **Dataset** | **Subjects** | **Sessions** | **Channels** | **Accuracy (%)** |
|---|---|---|---|---|
| PhysioNet | 109 | 1 | 64 | 99 |
| BCI Competition IV 2b | 9 | 5 | 3 | 96 |
| Elevvo Depression | 49 | 8 | 16 | 92 |

*Table B.1: Baseline EEG-based identification results obtained on three different datasets.*

Initial experiments were conducted on the PhysioNet EEG Motor Imagery dataset, which includes recordings from only a single session per individual. Despite this limitation, very high identification performance was achieved, with a classification accuracy of 99%. This result confirms that EEG signals contain strong subject-specific patterns that can be exploited for identification under controlled conditions. However, the absence of multiple sessions prevents the evaluation of inter-session variability, which is a critical factor in real biometric applications. For this reason, although the PhysioNet dataset was useful as a proof of concept, it was not considered suitable for

further analyses focused on generalization and robustness.

To incorporate temporal variability, experiments were then conducted using the BCI Competition IV 2b dataset, which provides EEG recordings from multiple sessions per subject. Despite the limited number of subjects and the use of only three EEG channels, the system achieved an identification accuracy of 96%. This result suggests that reliable identification may be possible even under constrained hardware conditions, although scalability cannot be assessed in this setting.

Finally, a more realistic and challenging scenario was evaluated using the Elevvo Depression dataset, with a larger number of subjects and sessions. Identification experiments conducted on this dataset achieved an accuracy of 92%, which highlights the impact of inter-session variability and reinforces the importance of multi-session evaluation when assessing EEG-based biometric systems.

Overall, these baseline experiments demonstrate the feasibility of EEG-based identification across datasets with diverse characteristics. Nevertheless, they do not address authentication performance or robust generalization, which are essential for practical biometric applications. For this reason, the main body of the thesis focuses on more advanced experimental protocols incorporating session-wise and task-based generalization.

## B.2 Experiment E7: Session-based generalization on Multi-session Cognitive Tasks datasets

Experiment E7 evaluates session-based generalization on the Multi-session Cognitive Tasks datasets. These results are referenced in Section 6.3, where task-based generalization with the same dataset is discussed. The datasets provide a particularly suitable testbed, as EEG recordings were collected across multiple sessions under two distinct acquisition conditions: a controlled laboratory environment and an unconstrained home environment. This allows not only the evaluation of session-based generalization, but also a comparison between recording contexts with different levels of noise and variability.

The experiment was conducted using the two datasets collected within the older adults study:

- **Laboratory dataset**: EEG recordings acquired in a controlled experimental environment.
- **Home dataset**: EEG recordings collected in participants' homes, introducing additional variability due to environmental noise and less controlled recording conditions.

A session-based data partitioning strategy was adopted in all configurations. For each subject, EEG data from a subset of sessions were used for training, while

one session was held out for testing. This protocol explicitly evaluates inter-session generalization, which is a critical requirement for biometric systems.

Three experimental configurations were evaluated:

1. Training and evaluation using only the *Home* dataset.

2. Training and evaluation using only the *Laboratory* dataset.

3. Training and evaluation using the combined *Home + Laboratory* dataset.

The same preprocessing, feature extraction, and modeling pipeline described in Chapter 4 was applied consistently across all configurations. Identification and authentication performance was evaluated at the session level by aggregating segment-level predictions.

### B.2.1 Results

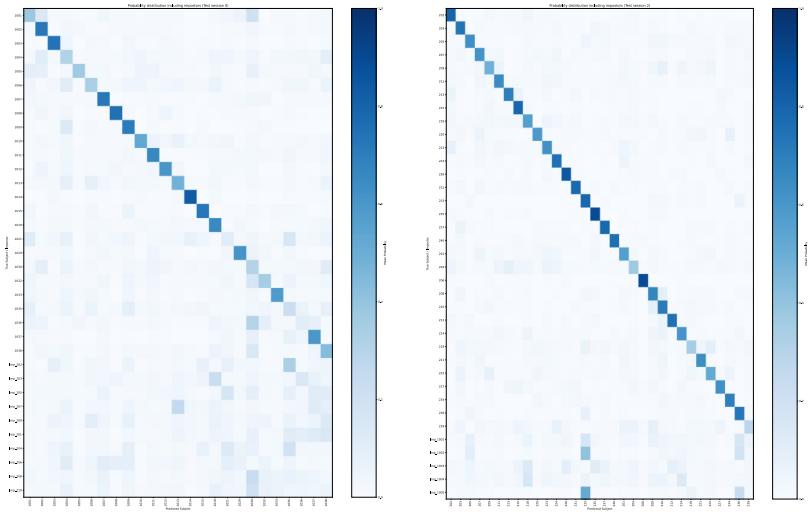
Table B.2 summarizes the session-based identification and authentication results obtained for the three configurations.

| **Dataset** | **ACC (Id)** | **ACC (Auth)** | **EER** |
|---|---|---|---|
| Home | 0.975 | 0.911 | 0.172 |
| Laboratory | 1.00 | 0.903 | 0.117 |
| Home + Laboratory | 0.986 | 0.883 | 0.151 |

*Table B.2: Session-based identification and authentication performance on the Multi-session Cognitive Tasks datasets.*

The laboratory dataset achieved perfect identification accuracy and the lowest Equal Error Rate (EER), indicating a higher degree of biometric stability under controlled acquisition conditions. In contrast, the home dataset exhibited slightly reduced performance, particularly in authentication, reflecting the increased variability introduced by less controlled environments. The combined Home + Laboratory configuration yielded intermediate results. While identification accuracy remained high, authentication accuracy decreased and EER increased compared to the results obtained using only the laboratory data.

These results highlight the strong influence of recording conditions on EEG-based biometric performance. The superior results obtained with laboratory recordings can be attributed to more stable electrode placement, reduced environmental noise, and consistent experimental protocols. This trend is also evident in the corresponding probability distributions of Figure B.1, where laboratory recordings exhibit sharper and more confident probabilities compared to home recordings.



*(a) Home dataset.* *(b) Laboratory dataset.*

*Figure B.1: Probability distributions for the Home and Laboratory datasets.*

## B.3 Experiment E12: Generic signal-level anonymisation strategies

This appendix reports the results of **Experiment E12**, which constitutes a preliminary exploration of generic, task-agnostic anonymisation strategies for EEG signals. As introduced in Chapter 7, the objective of this experiment is to evaluate whether simple perturbations can substantially reduce biometric identifiability while preserving the utility of EEG data for non-biometric downstream tasks. Specifically, this experiment investigates two widely used and conceptually simple anonymisation approaches: *Gaussian noise injection* and *PCA-based obfuscation*. These methods are applied globally to the EEG signals. The insights gained from these analyses motivate the development of more structured anonymisation strategies discussed in the Section 7.3.

All anonymisation transformations were applied directly to the raw EEG signals prior to segmentation and feature extraction, following the pipeline described in Chapter 4. Two independent models were trained for the non-biometric tasks: one for resting-state prediction and one for active versus passive task classification.

### B.3.1 Gaussian Noise Injection

In the first set of experiments, Gaussian noise was added to the EEG signals at different Signal-to-Noise Ratio (SNR) levels. This approach aims to reduce subject-specific information by increasing signal variability in a task-agnostic manner.

Table B.3 summarizes the results obtained with Gaussian noise injection.

For the non-biometric tasks, Gaussian noise injection also led to a degradation in performance, although the impact was less pronounced than for identification.

| Task | No noise | SNR 20 | SNR 10 | SNR 5 |
|---|---|---|---|---|
| Identification | 0.98 | 0.76 | 0.68 | 0.61 |
| Resting-state prediction | 0.70 | 0.61 | 0.60 | 0.59 |
| Active vs. passive tasks | 0.61 | 0.57 | 0.56 | 0.57 |

*Table B.3: Performance under Gaussian noise injection at different SNR levels.*

This indicates that while Gaussian noise weakens biometric identifiability, it does so at the cost of reducing the overall signal quality, limiting its effectiveness as a privacy-preserving transformation with fine-grained control over the privacy–utility trade-off.

### B.3.2 PCA-based Obfuscation

In the second set of experiments, anonymisation was performed using Principal Component Analysis (PCA). EEG signals were projected onto a reduced subspace by retaining only a limited number of principal components, thereby suppressing components associated with lower variance, which may encode subject-specific information.

The number of retained components was progressively reduced from the full representation down to a single component. Table B.4 reports the corresponding results.

| Task | All | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 |
|---|---|---|---|---|---|---|---|---|---|---|
| Identification | 0.98 | 0.97 | 0.96 | 0.85 | 0.90 | 0.86 | 0.80 | 0.67 | 0.60 | 0.50 |
| Resting-state prediction | 0.70 | 0.77 | 0.74 | 0.80 | 0.79 | 0.79 | 0.77 | 0.76 | 0.77 | 0.70 |
| Active vs. passive tasks | 0.61 | 0.72 | 0.73 | 0.72 | 0.73 | 0.73 | 0.72 | 0.71 | 0.72 | 0.65 |

*Table B.4: Performance under PCA-based anonymisation for different numbers of components.*

Unlike Gaussian noise, PCA-based obfuscation enabled a more controlled degradation of biometric performance. Identification accuracy progressively decreased as fewer components were retained. Importantly, the performance of the non-biometric tasks remained stable across most configurations and, in some cases, even improved relative to the original signal.

This behaviour suggests that PCA suppresses identity-related variability while preserving, or even enhancing, task-relevant structure. The improvement observed in resting-state and active/passive task classification may be attributed to the removal of noise and subject-specific components that are not informative for these tasks.

The results of Experiment E12 highlight the limitations of generic, task-agnostic anonymisation strategies. While Gaussian noise injection can reduce biometric

identifiability, it lacks precise control over the privacy–utility trade-off and leads to a general degradation of signal quality. Moreover, high signal-level identification accuracy combined with low model confidence indicates that such perturbations do not reliably suppress identity-related information.

In contrast, PCA-based obfuscation provides a more favourable trade-off. It enables a gradual reduction of biometric identifiability while largely preserving, and sometimes improving, performance on non-biometric tasks. Nevertheless, PCA remains a global transformation that does not explicitly target identity-related frequency components, limiting its effectiveness as a standalone anonymisation strategy.