Complex & Intelligent Systems (2026) 12:39
https://doi.org/10.1007/s40747-025-02157-4

ORIGINAL PAPER

Beyond gaze points: augmenting eye movement with brainwave data for multimodal user authentication in extended reality

Matin Fallahi[1,2] · Patricia Arias-Cabarcos[2,3] · Thorsten Strufe[1,2]

Received: 20 May 2025 / Accepted: 25 October 2025

Abstract

Extended Reality (XR) technologies are becoming integral to daily life. However, password-based authentication in XR disrupts immersion due to poor usability, as entering credentials with XR controllers is cumbersome and error-prone. This leads users to choose weaker passwords, compromising security. To improve both usability and security, we introduce a multimodal biometric authentication system that combines eye movements and brainwave patterns using consumer-grade sensors that can be integrated into XR devices. Our prototype, developed and evaluated with 30 participants, achieves an Equal Error Rate (EER) of 0.298%, outperforming eye movement (1.820%) and brainwave (4.920%) modalities alone, as well as state-of-the-art biometric alternatives (EERs between 2.5% and 7%). Furthermore, this system enables seamless authentication through visual stimuli without complex interaction.

Keywords Biometric authentication · EEG authentication · Eye movement authentication · Multimodal authentication

Introduction

Extended reality (XR) is a collective term that encompasses Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR), combining real and virtual environments for interactive user experiences [1–3]. XR is increasingly used across domains such as education [4], healthcare [5], and entertainment [6]. Its immersive 3D environments and real-time interactivity enhance user engagement, but conventional authentication methods like passwords disrupt this experience. This highlights the need for authentication mechanisms that ensure security while integrating seamlessly into XR, with biometric systems offering a promising alternative [7].

Biometric authentication systems employ unique behavioral or physiological traits to identify individuals. However, methods widely used on smartphones and desktops, such as facial recognition, fingerprint scanning, or keystroke logging, are less compatible with XR, as they often require external hardware and disrupt immersion. Meanwhile, XR devices are equipped with multiple sensors, including outward-facing cameras for environment-tracking and inward-facing cameras for eye tracking to enhance user experience. Prior work has demonstrated the potential of eye movements for user authentication [8–12], emphasizing their non-intrusive and hands-free nature. However, performance has been shown to decline at the low frame rates common in consumer-grade devices [8, 11].

In this paper, to improve the reliability of unimodal authentication systems based on eye movements, we adopted a multimodal approach that augments eye movement by measuring brainwave patterns as well. Multimodality has generally been shown to significantly improve authentication accuracy [13, 14], and brainwave patterns, in particular, are a biometric which is unique to each individual and consequently resistant to spoofing, difficult to duplicate, and hands-free [15]. Since brainwave capturing is naturally hands-free, it is as suited to the XR setting as eye tracking; however, the brainwave modality suffers from similar drawbacks as eye movement. Brainwaves are sensitive to noise and

✉ Matin Fallahi
matin.fallahi@kit.edu

Patricia Arias-Cabarcos
pac@mail.upb.de

Thorsten Strufe
strufe@kit.edu

[1] Informatik, Karlsruhe Institute of Technology, Karlsruhe, Germany

[2] KASTEL Security Research Labs, Karlsruhe Institute of Technology, Karlsruhe, Germany

[3] Department of Computer Science, Paderborn University, Paderborn, Germany

Published online: 23 December 2025

artifacts, especially in consumer-grade devices where performance is unreliable [16, 17]. Therefore, we hypothesize that the integration of eye movement and brainwave modalities will yield a secure, robust, and user-friendly authentication mechanism that is fully compatible with XR environments.

To validate our multimodal authentication approach in XR, we conducted a lab study with 30 participants using consumer-grade equipment to record synchronized brainwaves and eye movements. Based on this dataset, we developed a twin neural network system employing an interactive dot stimulus [11] and evaluated multiple feature and score fusion strategies. Since some studies include pupil diameter as a feature [10, 18] while others omit it [8, 11], we also examined its impact on authentication performance. Our results demonstrate the effectiveness of the proposed multimodal system, and we summarize the main contributions as follows:

- *Innovation in Multimodal Authentication* Our authentication system is the first to combine synchronized eye movement and brainwaves. Thus, we offer a novel multimodal authentication solution tailored for XR.
- *Substantial Improvement in Accuracy* We provide experimental confirmation that a multimodal authentication is more effective than a unimodal authentication in the XR context. Our authentication via eye movement shows a notable 81%–83% reduction in EER when augmented with brainwave data.
- *Insight into Pupil Diameter* We investigate the impact of pupil diameter on authentication, offering additional depth to the understanding of eye movement results. The findings serve as a guide for feature selection in the authentication system based on eye movement.

Authentication in XR

Authentication in XR can be categorized into knowledge-based [19, 20] (e.g., passwords), possession-based [21] (e.g., hardware tokens), and inherence-based (biometric) methods [10, 22, 23]. Knowledge-based credentials are hard to recall and input, and possession-based pose availability and security issues [7]. However, Biometrics remove such burdens and benefit from XR's growing sensor integration [7].

While biometrics offer clear usability advantages, there remain challenges related to performance and accuracy. A promising approach is multimodal biometric authentication, particularly combining eye movement and brainwave signals. These modalities are hands-free, difficult to spoof, and have applications beyond authentication [11, 17]. Their combination could mitigate accuracy issues while retaining usability. For a biometric system to be viable, the underlying trait must be universal (present in all individuals), permanent (stable over time), unique, quantifiable, revocable by the user, and resistant to forgery. In addition, sensing must be affordable, and the system must deliver reliable performance.

Studies have demonstrated the temporal stability of brainwave and eye movement-based biometric authentication methods. Maiorana [24] conducted a year-long study, confirming the viability of brainwave authentication. Similarly, Lohr et al. [8] established that eye movement based authentication remains effective over a period of three years. While static biometrics, such as fingerprints or iris scans, cannot be altered if compromised. Lin et al. [25] provided empirical data suggesting that brainwave passwords can be modified by assigning a different task. In terms of usability, several studies have indicated a general willingness among people to adopt eye movement and brainwave authentication methods, although privacy concerns still remain to be addressed [26–28]. Overall, we acknowledge the need for further research to examine other aspects of biometric authentication, but the specific aim of our work is to achieve reliable performance using sensors compatible with XR devices.

The reliability and performance of eye movement and brainwave authentication systems depend heavily on the quality of the recorded data. For instance, Lohr et al. [8] demonstrated that when the sampling rate of an eye-tracker decreases from 1000 to 31 Hz, the EER increases from 3.66% to 23.37%. Moreover, in eye movement authentication, Sluganovic et al. observed that a reduction in sampling rate from 500 to 50 Hz resulted in an 11% increase in the EER. Similarly, in brainwave authentication systems, Arias-Cabarcos et al. [17] compared the performance of two different datasets: a consumer-grade dataset with a 256 Hz sampling rate and a medical-grade dataset with a 1024 Hz sampling rate. Using the same machine learning pipeline for both datasets, they found EERs of 8.5% for the consumer-grade dataset and 1.9% for the medical-grade dataset. Therefore, to effectively utilize brainwaves or eye movement for authentication, it is essential to enhance the performance of authentication systems in consumer-grade devices.

Experimental design and procedures

Our research aims to achieve robust and reliable authentication in XR, specifically focusing on achieving high performance with consumer-grade devices. Since biometric authentication using just single modalities like eye movement [8, 11] and brainwaves [16, 17] have been shown to be unreliable, we investigate the efficacy of combining eye movement and brainwave data. Therefore, we formulate our central research questions as follows: Can synchronized eye movement and brainwave data improve performance compared with a unimodal? Which modality is more reliable? What fusion strategy yields the best results? And, what impact does the feature of pupil diameter have on outcomes? To answer these questions, we have designed a set of experiments that is described in detail in this section.

Technological and methodological blueprint

This section outlines the methodologies, software, and tools employed to implement multimodal authentication through the synchronized integration of eye movement and brainwaves in our experimental design (Fig. 1a):

Authentication task—reflexive saccadic responses

Building on the methodology proposed by Sluganovic et al. [11], we focused on measuring reflexive saccades due to their inherent stability and low susceptibility to temporary changes in mental or emotional conditions, such as attention, mood, or stress—referred to as transient cognitive states. Unlike voluntary saccades, reflexive saccades are driven by automatic mechanisms, making them more reliable for authentication. Our stimulus involved presenting a dot on the screen. When the participant's gaze fixated on the dot, it would disappear, and a new dot would appear in a random position. This sequence was repeated 25 times per round [11], with participants completing 36 rounds. The number of rounds was determined based on the experiment's 25-min time limit. To prevent fatigue and maintain focus, we included a 15-s rest interval between rounds. Previous studies by Sluganovic et al. [11] have demonstrated that this interactive dot task resists replay attacks effectively, as participants must respond to new random dot positions in real-time.

Task implementation—PsychoPy

The authentication task was designed and executed using PsychoPy, a platform commonly used in neuroscience and psychology for creating complex visual and auditory stimuli [29]. This choice was motivated by two key factors: first, its compatibility with our eye tracker enabled the development of an interactive task that dynamically adjusts to the user's gaze position; second, its native Python support facilitated sending event markers to manage synchronized recording of our experiment.

Synchronization—Lab Streaming Layer (LSL)

Given the significance of millisecond-level precision in brainwave [30] and eye movement data [31], synchronization is crucial. We employed the Lab Streaming Layer (LSL),[1] which is already used in literature for synchronized multimodal recording of brainwave and eye movement data [32]. LSL is a system designed for unified time series data collection in research settings. We utilized LSL to achieve synchronized recording of brainwave and eye movement data, along with timestamps streams.

Equipment

In line with our objective to develop an authentication system for XR environments, we decided against using medical-grade EEG recorders or high-resolution desktop eye trackers commonly employed in other research [10, 11, 33, 34]. Instead, we selected an eyeglass based eye tracker and a neuroheadset tailored for general use. Our aim was to investigate the reliability of these devices for authentication tasks. The following section provides a technical overview of the equipment used (Fig. 1b):

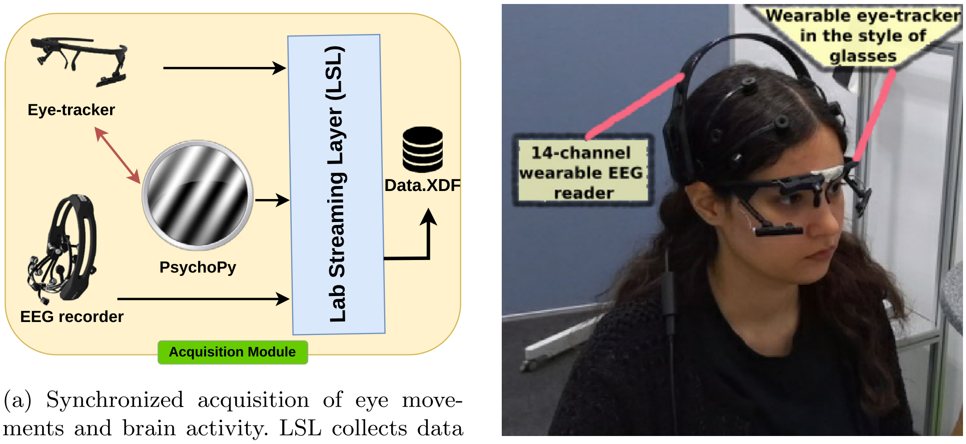
- *Neuroheadset—Emotiv EPOC X* The Emotiv EPOC X neuroheadset,[2] equipped with 14 EEG electrodes, records brainwave data at a sampling rate of 256 Hz. The accompanying Emotiv software provides connectivity and quality metrics. The connectivity refers to how well the device connects to the head of the subject, while signal quality provides a summary measure that considers various factors such as movement, noise, signal amplitude, and other parameters. According to Emotiv's guidelines, optimal electrode contact with the scalp can yield a 100% connectivity score. Nonetheless, achieving high-quality data can be challenging for individuals with long or thick hair.
- *Eye tracker—Pupil Core* For an XR-like experience, we chose the Pupil Core,[3] as an eye tracker. The Pupil Core includes one world camera, which records the surrounding environment and the participant's field of view, and two eye cameras, which simultaneously record the participant's eyes and capture detailed information about gaze direction, pupil diameter, and eye movement at up to 200 Hz. To calibrate the eye tracker, participants were asked to focus on a sequence of five dots on the screen using the Pupil Capture. The procedure was repeated as needed, and calibration was considered successful when the 3D Gaze Mapping alignment visually matched the target points with an estimated accuracy of 1.5°–2.5°, as recommended by Pupil Labs. However, recent advancements in calibration-free eye-tracking systems, such as Pupil Invisible and Tobii Glasses X, suggest a shift toward reducing or eliminating per-user calibration in the near future.

[1] https://github.com/sccn/labstreaminglayer.

[2] https://www.emotiv.com/epoc-x/.

[3] https://pupil-labs.com/products/core/.

Fig. 1 Overview of the synchronization setup and equipment used during data collection


Eye-tracker
PsychoPy
EEG recorder
Lab Streaming Layer (LSL)
Data.XDF
Acquisition Module
Wearable eye-tracker in the style of glasses
14-channel wearable EEG reader

(a) Synchronized acquisition of eye movements and brain activity. LSL collects data from PsychoPy, an eye tracker, and an EEG recorder.

(b) A participant wearing the experimental setup.

Participants and ethical considerations

Our study included 30 participants, 11 women and 19 men. These participants were predominantly young adults (average age 24) affiliated with the university, either as undergraduate, master's, or PhD candidates or as research assistants. The only requirement for participation was that individuals be aged 18 or above.

The university's official social media channels were used to recruit such a cohort. We emphasized that participation was entirely voluntary, and participants had the freedom to withdraw from the study at any stage without any consequences. Participants were reimbursed 15 Euros per hour upon study completion as compensation for their time and contribution.

This study adheres to responsible research practices by maintaining ethical integrity. All procedures, methodologies, consent form, and tools underwent rigorous scrutiny and were approved by our university's Institutional Review Board (IRB). Additionally, for Fig. 1b, we obtained the subject's consent to use her photograph in the paper.

Data collection process

Before participating in an experiment, each subject must read and sign an informed consent form that explains the purposes of the experiment, the types of data to be collected, and how the data will be used. The preparatory step to the experiment itself is the equipment setup. The participating subjects first wear the Pupil Core eye tracker recorder, and next, they wear the Emotiv EPOC X EEG headset. The two devices are then adjusted for each subject, that is, the electrodes of the Emotiv EPOC X are placed to enhance data quality, and the Pupil Core eye tracker is calibrated to achieve precise measurements of gaze points on the screen. To record data, we used the software Emotiv Pro to stream brainwave data and the software Pupil Player to stream eye movement data. Next, we launched the experiment script in PsychoPy to present stimuli and also to manage the event marking stream. Finally, we used the Lab Streaming Layer (LSL) to record data streams from the brainwave recorder, the eye tracker, and PsychoPy (Fig. 1a). Upon completion of the experiment, all devices are carefully removed from the participants, and compensation is provided in accordance with the informed consent agreement.

Authentication architecture

Here, we describe our authentication system architecture and fusion methods employed in the study.

Authentication approach overview

A biometric system consists of four core components: acquisition, preprocessing, recognition, and comparison. These elements are common to all biometric modalities and form the foundation of any biometric authentication system.

1. *Acquisition Module* The initial step involves capturing the user's biometric data through specialized sensors. This phase is essential for collecting the raw information that will be analyzed and compared in subsequent stages. As detailed in the "Experimental Design and Procedures" ("Experimental design and procedures" section), the data acquisition phase involves the collection of user's data through specific sensors, tailored to the biometric trait being analyzed.
2. *Pre-processing Module* The acquired data undergoes a series of preprocessing steps to enhance its quality and make data ready for the next steps. This includes the extraction of relevant time series, filtering, interpolation, and standardization.

3. *Recognition Module* At this stage, the system identifies and extracts specific characteristics or features from the pre-processed data. These features are the unique attributes that distinguish one individual from another. The effectiveness of this step is paramount in ensuring that the system can accurately match the input data with stored templates. We used a twin neural network, inspired by BrainNet[4] [33], to extract biometric features. The network's architecture, guided by a triplet loss function, effectively reduces the dimensionality of the data while preserving individual characteristics.
4. *Comparison Module* The extracted features are then compared to stored biometric templates. We used the Euclidean distance $d = \sqrt{\sum_{i=1}^{n}(e_i - v_i)^2}$ to compute the distance between enrollment and verification samples, and defined the similarity score as $s = -d$ so that higher values indicate greater similarity. Authentication is accepted if $s \geq \tau$, where τ is the decision threshold determined during system calibration.

Preprocessing

The raw data includes multiple time series recorded during the experiment. In order to prepare input for our recognition module, we extract a segment of raw data known as a sample. This sample should consist of recorded biometric data when the subject is exposed to stimuli—in our case, a dot displayed on the screen. We extract these samples based on timestamps where the subject's gaze aligns with the dot's position on the screen.

1. *Data Extraction* We extracted relevant time series data for brain, eye, and event timestamps from raw .xdf files, the output format of the Lab Streaming Layer (LSL). Specifically, for brain activity, we extracted 14 time series corresponding to 14 EEG electrodes. For eye movements, we extracted 12 time series related to the x and y coordinates of both the pupil and gaze point. We excluded time series related to the z-coordinate and gaze confidence, as they could introduce session-specific learning due to the fixed screen-to-user distance, which may vary based on the user's height and dependence of confidence on environmental factors and calibration. Also, we extracted 4 time series related to pupil diameter to investigate authentication with pupil diameter features.
2. *Sample Extraction* Brainwave and eye movement samples were extracted based on the timestamps corresponding to the last hit of the eye-gaze with the dots. We select a duration of 0.4 s for samples, which encompass 0.1 s before the event and extending 0.3 s after it. This duration is chosen to provide data between the dot hit moving to another dot. If a longer duration were selected, it would result in the inclusion of data from multiple dots within each sample, thereby complicating the analysis.
3. *Standardization:* We needed to have a fixed sample size as input for our recognition system. Therefore, through resampling, we achieved a consistent count of 256 data points per second across the dataset for both eye movement and brainwave data.
4. *Data integrity and reliability:* In the eye movement data, we had NaN values caused by blinking. We filtered samples abundant in NaN values to retain only high-quality samples. For the remaining eye movement samples containing NaN values, interpolation techniques were applied solely within the sample to prevent information leakage. Unlike eye data, the brain data didn't require this step, as the brainwave recorder software has a built-in interpolation mechanism.

Feature extraction: Twin Neural Network

To authenticate subjects, we extract unique individual information from eye movement and brainwave signals, which often include noise. We employed a Twin Neural Network (TNN) with a triplet loss function as the core of our feature extraction module. A TNN is a specialized neural network architecture consisting of two or more identical sub-networks connected in parallel. The triplet loss function ensures that embeddings from the same identity are close, while embeddings from different identities are far apart [35]. The triplet loss function L is defined as the Euclidean distance:

$$L(A, P, N) = \max\Big(\|(f(A) - f(P))\|^2 - \|(f(A) - f(N))\|^2 + \alpha, 0\Big). \tag{1}$$

Here, $f(\cdot)$ denotes the embedding function that maps an input sample to its feature representation, A represents an anchor input, P is a positive input (sample from the same subject as A), and N is a negative input from a different subject. The parameter α serves as a margin that enforces a minimum level of dissimilarity between positive and negative pairs, thereby enhancing the differentiation of samples. Let M denote the number of triplets. The objective is to minimize:

$$\sum_{i=1}^{M} \|f(A_i) - f(P_i)\|^2 - \|f(A_i) - f(N_i)\|^2 + \alpha. \tag{2}$$

The indices i correspond to the individual triplet inputs utilized during training. The selection of triplets adheres to a strategy inspired by FaceNet, employing semi-hard negative mining to promote efficient convergence in learning [35].

[4] The hyperparameters: Adam optimizer (learning rate = 0.001), batch size = 128, and 250 epochs.

We employed a CNN for our sub-network architecture due to its demonstrated effectiveness in various brainwave and eye movement authentication studies [8, 24, 33, 36], specifically using the CNN architecture proposed in the BrainNet paper [33].

Comparison in verification mode

The next step is to make an authentication *decision*. The decision-making process can either consider each pair of verification and enrollment samples independently or use a group of them collectively. The number of enrollment samples used can impact data acquisition and performance. Specifically, using fewer samples leads to quicker data acquisition times but compromises the accuracy of the system and vice versa. We explored various scenarios to gain deeper insights into these trade-offs.

- *Fixed Threshold-One Sample (S1):* Fastest query time but potentially least reliable. For this scenario, we use one sample as verification and one sample as enrollment.
- *Fixed Threshold-Best Match (S2):* In most cases, multiple enrollment samples are available for each subject, and the best match between the verification and all enrollment samples is chosen for decision-making. This aims to enhance the system's performance by leveraging multiple enrollment samples.
- *User-specific Threshold (S3):* Alternatively, setting an individualized threshold for each user may yield more accurate results tailored to the unique behavioral characteristics of each individual. This strategy is in line with many existing authentication systems that train a user-specific model [17, 33, 37]. While this method is expected to achieve higher performance, it necessitates an initial calibration phase for each newly enrolled user to determine the optimal threshold in the real-world implementation.

Fusion

In our biometric authentication system, fusion plays a pivotal role in integrating information from the two modalities we employ. We implement fusion at two different levels: at the feature and at the score levels. These fusion levels are designed to enhance the robustness and accuracy of the system by leveraging the complementary information present in both modalities. Below we detail the specific methods and considerations for each fusion level.

Score fusion

To implement score fusion for multimodal biometric authentication, we trained separate twin neural networks for two distinct modalities: eye movement and brainwave data. These networks generate similarity scores that serve as the basis for subsequent fusion techniques. We use several established strategies to combine the similarity scores of eye movement and brainwave data pairs in the evaluation phase. Score fusion occurs in the *Comparison Module* of the authentication system after similarity calculation for each modality. Specifically, the Max method computes the maximum score across each corresponding pair, symbolized as $\max(s_{\text{eye}}, s_{\text{brain}})$, where s_{eye} and s_{brain} represent the scores for eye and brain data, respectively. Conversely, the Min method calculates the minimum score using $\min(s_{\text{eye}}, s_{\text{brain}})$. The Average method takes the mean of both scores, expressed as $\frac{s_{\text{eye}}+s_{\text{brain}}}{2}$. Finally, the product approach multiplies the scores, resulting in $s_{\text{eye}} \cdot s_{\text{brain}}$. We employ these different fusion techniques to generate the final similarity score, thus making our authentication system more robust.

Feature Fusion

As depicted in Fig. 2, we employed two separate convolutional neural networks (CNNs) as subnetworks within a twin neural network for the purpose of feature fusion. In this way, we accommodate the different characteristics of brainwave and eye movement data. In *Architecture A,* each of these CNN networks ends in a 16-dimensional dense layer. Subsequently, these two 16-dimensional layers are concatenated, resulting in a subject-representative layer comprising 32 values. On account of this structure, the feature fusion process is dynamically guided by the loss function, thus ensuring an integrated representation which accentuates the distinct attributes of both modalities. Moreover, in *Architecture B*, to enhance the efficacy of the feature fusion process, we explored an alternative configuration that incorporates an additional 32-dimensional dense layer. This supplementary layer aims to facilitate a more complex integration of features, which, in its own right, will reduce total dependence on the loss function for effective fusion.

Results and testbed

This section delineates the experimental settings of the testbed and presents the outcomes across various metrics for different comparison strategies.

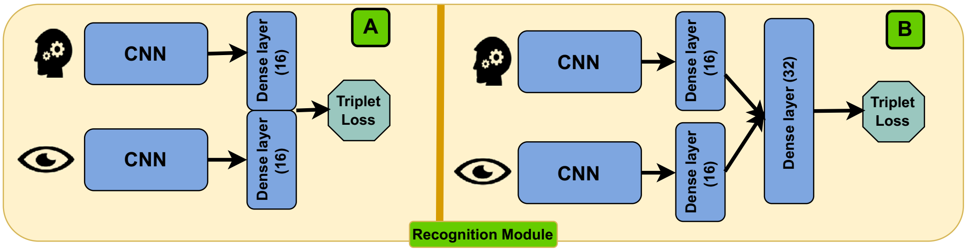


Fig. 2 Twin sub-network architectures (**A**, & **B**) for eye movement and brainwave feature fusion

Testbed and evaluation metrics

For a robust evaluation, we structured the dataset to segregate training and testing subjects. We employed sixfold cross-validation, where each fold contained 25 subjects for training and an additional 5 unseen subjects for testing. Also, during each fold of the cross-validation process, we fit a normalization function[5] on the training data and applied it to both training and test.

The evaluation data were analyzed using the comparison scenarios detailed in "Comparison in verification mode" section. To ensure the integrity of the evaluation, samples originating from the same experimental round as the verification samples were excluded. Consequently, enrollment and verification samples were consistently derived from two separate rounds of the experiment.

Metrics

The Equal Error Rate (EER) served as a summary metric, indicating the point where the False Acceptance Rate (FAR) and the False Rejection Rate (FRR) are equal. Additionally, we report FRR at specific FAR thresholds of 1%, 0.1%, and 0.01%. The FAR represents the success rate of an attacker in a zero-effort attack, and the goal is to achieve a lower FAR while maintaining a reasonable FRR. Moreover, high FRR may lead to additional verification attempts, which can harm the device's usability. Therefore, a balance must be maintained between these two metrics to ensure robust security without compromising the user experience.

The EER serves as a useful comparison metric across studies but is not directly practical for real-world applications. Notably, NIST (2023)/ISO[6] and the European Border Guard Agency Frontex[7] specify that biometric systems must operate at FAR $\leq 0.1\%$. Meanwhile, FIDO[8] and the updated NIST (August 2024)/ISO standards[9] recommend an even stricter FAR $\leq 0.01\%$. Additionally, FIDO and late NIST/ISO standards propose an FRR $\leq 5\%$, ensuring 19 successful logins out of 20 attempts for legitimate users.

Threat model

In alignment with the methodology of Zhang et al. [38], we consider an adversary whose goal is to access sensitive personal information—such as user accounts, photos, or financial data—or to perform unauthorized actions like initiating payments or installing malware on a user's XR device. We assume that the adversary is knowledgeable about the authentication dot task and has physical access to the user's XR headset. Given these assumptions and the adversary's available techniques, we classify the following attacks:

Blind attack

The adversary has no prior knowledge of the legitimate user's eye movement and brainwave patterns. To execute the attack, the attacker wears the user's XR headset and attempts authentication with their own biometric samples. However, since the attacker cannot gain any advantage from observing the subject during authentication—due to brainwave data being completely resistant to observation and eye movement data requiring specialized devices—this attack is effectively equivalent to a mimic attack. In the mimic attack scenario, other threat models consider observers without any additional capabilities (unaided eye) [38].

Random input attack

We consider an adversary capable of circumventing the XR interface to gain access to the API of our biometric system, enabling them to input arbitrary feature vectors. The adversary's goal is to find a feature vector that is close to the genuine user's feature vector. Following Zhao et al. [39], we assume that the feature vectors are normalized and that the number of features is publicly known, with values between 0 and 1. To implement the attack, we generate 1 million

[5] https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html.

[6] https://pages.nist.gov/800-63-3/sp800-63b.html.

[7] https://www.frontex.europa.eu/assets/Publications/Research/Best_Practice_Technical_Guidelines_ABC.pdf.

[8] https://fidoalliance.org/specs/biometric/requirements/Biometrics-Requirements-v4.1-fd-20250106.pdf.

[9] https://pages.nist.gov/800-63-4/sp800-63b.html.

Table 1 This table displays the EER of our biometric authentication system based on three approaches: single modality, score fusion, and feature fusion

		EER (%)		
Approach	**Biometric**	**S1**	**S2**	**S3**
Single Biometric	Brainwaves	14.55	5.560	4.920
	Eye-tracking with pupil	13.42	2.160	1.820
	Eye-tracking without pupil	19.07	3.732	3.639
Score Fusion	Max. with pupil	13.00	4.355	3.763
	Min. with pupil	13.35	1.481	1.231
	Mean with pupil	8.574	0.507	0.385
	Product with pupil	7.098	0.429	0.298
	Max. without pupil	13.18	3.445	2.990
	Min. without pupil	13.85	1.510	1.350
	Mean without pupil	9.601	0.850	0.686
	Product without pupil	9.281	0.906	0.700
Feature Fusion	Architecture A with pupil	8.810	0.917	0.802
	Architecture B with pupil	13.36	1.831	1.674
	Architecture A without pupil	10.47	1.550	1.231
	Architecture B without pupil	16.28	4.373	4.085

The columns represent different authentication strategies ("Comparison in verification mode" section)

samples with 32 values randomly selected from a uniform distribution between 0 and 1. Then, we compare them with the normalized feature vector of the legitimate users.

Overall results

From our data collection, we obtained a total of 22,688 dot samples for analysis. Next, we trained the twin network ("Testbed and evaluation metrics" section). The network was trained separately for brain and eye movement data. For the eye movement, we trained once including pupil diameter and once excluding it. For both brain and eye data, we also applied two different feature fusion architectures ("Fusion" section). The training and evaluation were conducted under the conditions specified in our testbed (5.1). The outcomes, particularly the Equal Error Rates (EER), are summarized in Table 1.

Fixed threshold-best match versus user-specific threshold

Our results clearly show that increasing the number of enrollment samples enhances the system's performance. However, the impact of different threshold strategies becomes evident when we focus on the scenarios which employ multiple enrollment samples, that is, the Fixed Threshold-Best Match scenario (S2) and the User-specific Threshold scenario (S3). For instance, using brainwaves, the EER drops from 5.560% in the S2 to 4.920% in the S3. This comparative analysis reveals that S3 consistently outperforms S2 under the same conditions involving multiple enrollment samples. Moreover, the advantages of using a User-specific Threshold strategy become increasingly apparent when fusion methods are considered; for example, in Score Fusion methods, such as the product with pupil diameter, the EER improves from 0.429% in S2 to 0.298% in S3, indicating a 30% reduction in error.

Single biometrics versus fusion approaches

Table 1 reveals a consistent advantage for Fusion Approaches over Single Biometrics across various scenarios, particularly in the most advanced scenario (S3). For Single Biometrics, Eye tracking with pupil diameter registers the lowest EER, achieving 1.820%. However, this is significantly outperformed by Score Fusion methods, such as the product with pupil diameter, which exhibits an EER of just 0.298%. Likewise, Feature Fusion's Architecture A with pupil demonstrates superior performance with an EER of 0.802%. These results confirm that Fusion Approaches markedly outshine Single Biometrics.

Score fusion versus feature fusion

Score fusion generally outperforms feature fusion across multiple comparison scenarios. Within score fusion, the mean and product strategies demonstrate a marked advantage over the min and max approaches in reducing the EER. On the feature fusion front, Architecture A consistently yields

Table 2 FRR (%) at FAR levels of 1%, 0.1%, and 0.01% for single-biometric, score-fusion, and feature-fusion methods across three evaluation scenarios (S1–S3)

Approach	Biometric	FRR at FAR = 1%			FRR at FAR = 0.1%			FRR at FAR = 0.01%		
		S1	S2	S3	S1	S2	S3	S1	S2	S3
Single biometric	Brainwaves	59.47	28.86	21.45	83.64	61.22	49.09	94.43	86.92	67.99
	Eye tracking with pupil	52.85	8.15	6.85	74.89	23.25	16.44	86.17	42.08	25.00
	Eye tracking without pupil	68.73	17.10	17.97	88.06	44.86	38.81	96.08	66.33	52.80
Score fusion	Mean with pupil	25.29	0.346	0.356	45.15	2.10	1.61	62.71	6.21	3.79
	Product with pupil	21.88	0.236	0.239	42.96	1.79	1.23	60.99	7.00	3.79
	Mean without pupil	35.39	0.990	0.873	59.03	5.86	5.44	76.60	19.04	11.61
	Product without pupil	35.27	1.123	0.965	59.45	6.85	5.96	77.20	18.72	12.36
Feature fusion	Architecture A with pupil	26.29	1.066	1.101	48.99	6.06	5.26	68.25	20.69	14.69
	Architecture A without pupil	41.16	3.549	3.008	65.98	15.58	10.80	81.94	27.10	20.74

better results than Architecture B, although neither matches the high performance of score fusion techniques employing mean or product strategies.

Influence of pupil diameter

Table 1 clearly illustrates the role of pupil diameter in enhancing authentication performance. When comparing eye tracking methods with and without pupil diameter, there's a consistent improvement in EER across all examined scenarios. For instance, in the S3 scenario, eye tracking with pupil diameter data yields an EER of 1.820%, whereas the approach without pupil diameter data results in a higher EER of 3.639%.

False acceptance rate (FAR) insights

In practical applications, EER is often not the main metric of focus. Instead, the practical usage needs to minimize FAR to bolster system security against zero-effort attacks while maintaining a reasonable FRR To elucidate the trade-offs between these metrics, we present data in Table 2. A comparison between FAR at 0.01%, 0.1%, and 1% reveals that single-biometric approaches suffer from high FRRs, particularly in stringent security settings with low FARs (0.1% and 0.01%). For instance, in the S3 scenario with a FAR of 0.01%, FRRs for single-biometric approaches like Brainwaves and eye tracking with pupil are 67% and 25%, respectively. These high FRRs indicate that single-biometric methods could be impractical for high-security applications. In contrast, multimodal methods, especially those utilizing score fusion, significantly alleviate this issue. For example, the Mean Score Fusion with and without pupil diameter method result in a much lower FRR of 3.79% and 11.61% at a 0.01% FAR, showcasing its effectiveness in balancing security and usability.

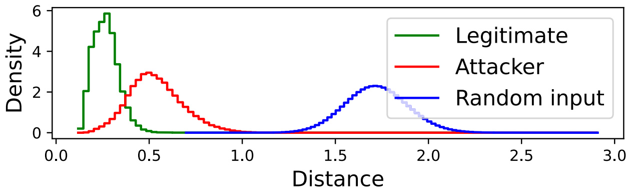


Fig. 3 Density distribution of brainwave similarity scores for legitimate users, human attackers, and random features (similar for eye movement and fusion)

Insight into subject-level EER

Figure 4 reveals distinct patterns in user-level EER, with values ranging from 0 to 2.77% On average, the EER stands at 0.686, but a relatively high variance of 0.397 suggests notable differences in authentication performance across subjects. Specifically, about eight subjects exhibit EERs that are close to zero, underlining the system's effectiveness for these individuals. Conversely, five subjects manifest EERs nearly twice the average, accounting for the high variance and indicating that the system may require optimization for these cases. In terms of cross-validation rounds, some models appear to perform better than others, or certain subjects may have noise in their data samples. Interestingly, even within rounds that have a higher average EER, some subjects still achieve low EER values. This suggests that the higher error rates are likely not a result of model inefficiency but rather may stem from noise in the data samples for specific subjects.

Random input attack

The density distance plot (Fig. 3) demonstrates that the distances between randomly generated features and legitimate users are significantly greater than those between human attackers and legitimate users. This observation confirms that distance-based recognition systems are robust against

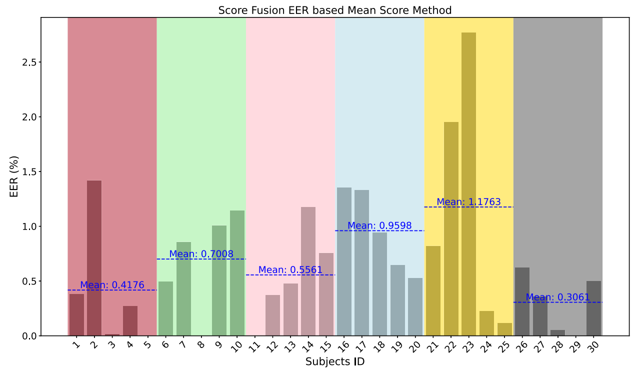


Fig. 4 Bar plot of test subjects' EER using Mean Score Fusion without pupil. Background shades denote cross-validation rounds, and dashed blue lines mark mean EER for each group of five subjects

random feature input attacks [39, 40], suggesting that such attacks are less effective than blind attacker scenarios.

Correlation between modalities

To investigate the correlation between synchronized eye movements and brainwaves, we applied the Pearson correlation coefficient [41] for time domain analysis and magnitude-squared coherence, calculated using Welch's method [42], for frequency domain analysis. We explored correlations between: (1) different channels of eye movement and brainwave data separately; (2) synchronized brainwave and eye movement data; (3) brainwave and eye movement data across different times of the experiment; and (4) brainwave and eye movement data from different subjects. The analysis revealed a consistent trend in both the time and frequency domains, demonstrating a strong[10] correlation within the eye movement data and within the brainwave data (1). However, cross-modality correlations between eye movements and brainwaves were generally weak.[11] Notably, the correlation of synchronized data from the same subject (2) was approximately 1.5 times stronger than that observed in cross-subject and time comparisons (3–4). The synchronized correlations demonstrated a more significant association between x-axis eye movement features and the frontal lobe region of the brain, particularly in the right hemisphere.

In summary, the weak correlation in synchronized data and strong correlation within modalities explain why the fusion of brainwave and eye movement data significantly improves the performance of our authentication system. Thus, we see our hypothesis confirmed that multimodal authentication based on eye movement and brainwaves substantially enhances the reliability of authentication using consumer-grade equipment with low sample recorder rates.

Technical feasibility of XR integration and usability aspects

To effectively integrate multimodal authentication into real-world applications, it is essential to consider both technical feasibility, usability, and privacy. These three aspects are crucial for ensuring successful implementation and user acceptance. In the following, we discuss feasibility, based on the current and projected technological landscape; usability, and privacy, grounded on previous empirical studies evaluating similar interfaces. We also discuss revocability and resistance to coercion.

Technical feasibility of XR integration

While our multimodal authentication system is not immediately deployable in current XR platforms due to the limited availability of integrated EEG sensors, we believe it can be seamlessly adopted in the near future as consumer-grade XR headsets begin to incorporate such capabilities. This belief is supported by four factors: (a) our system is developed using separate but technically integrable consumer-grade devices (b) we use a simple interactive dot task for implicit, hands-free authentication, (c) we have selected two biometric modalities which are well suited for use in XR, and (d) the model is lightweight and compatible with mobile devices, including XR headsets.

Consumer-grade devices

We aim to enhance the potential of consumer-grade devices for use in XR. We pursue this aim by employing only devices that are designed to operate at consumer-grade sampling rates.

For *eye-tracking*, we used the Pupil Core device with a 200Hz sample rate in our experiments. Prior research by Pastel et al. [43] demonstrates the viability of integrating such technology in XR, citing 38 papers that used eye-trackers integrated with head-mounted displays (HMD). Moreover, companies like Pupil Labs and Tobii offer eye-tracking solutions designed for AR and VR devices .[12] Additionally, Varjo provides a VR headset with an integrated eye-tracker operating at a 200Hz sample rate, aligning with our hardware specifications.[13]

Similarly, for *brainwave recording*, we employed the Emotiv EPOC X, a consumer-grade device with a sampling rate of 256 Hz. Recently, Li et al. [44] demonstrated that EEG sponge electrodes can be seamlessly integrated into VR headsets. Additionally, dry EEG electrodes have been

[10] Higher than 0.6.

[11] Lower than 0.1.

[12] https://www.tobii.com/products/integration/xr-headsets.

[13] https://varjo.com/products/vr-3/.

developed and are commercially available[14, 15] and some of these electrodes are already incorporated into VR devices.[16] Soon there will be VR headsets available that incorporate both EEG and eye-tracking capabilities.[17]

Dot task

To facilitate implicit authentication, we implemented an interactive dot task as the stimulus in our system. In a dot task, subjects are instructed to follow a dot displayed on the screen. This task offers high technical feasibility because it is simple and adaptable, thereby making it easy to integrate into daily routines or workflows. Consequently, our interactive dot task is well-suited for implicit authentication in the XR environment.

Broad applications

We selected eye movement and brainwave data as our biometric modalities due to their extensive applications beyond mere authentication. These two modalities are suitable for enhancing human-computer interactions in XR, as supported by existing studies [45, 46]. The tracking of eye movements and brainwave patterns is also applicable in specialized entertainment contexts [47, 48]. Consequently, our chosen biometric modalities are particularly well-suited for integration into XR headset equipment.

Memory, latency and energy

Since our feature extractor is trained and evaluated in an inter-subject setting, and feature comparison is based on Euclidean distance, no training is required during enrollment or verification—only model inference is needed, which constitutes the main computational overhead of our approach. Also, as our best results are based on score fusion, each verification attempt requires two inferences: one for brainwave data and one for eye movement data.

Due to software restrictions on XR devices, we evaluated runtime performance on a Redmi Note 11 S smartphone. This mid-range Android device has hardware comparable to the Meta Quest 3, though it is significantly less powerful than high-end XR devices such as the Apple Vision Pro.[18, 19, 20] Using the official TensorFlow Lite Benchmark Tool, we ran 50 inference iterations on the device. The average inference latency was 14.06 ms, with a peak memory usage of 11.28 MB.

Therefore, the proposed solution remains lightweight, even when compared with other mobile-compatible neural architectures for biometric authentication. For example, MobileFaceNets [49], which are optimized for on-device face recognition, report inference latencies between 18 and 27 ms.[21]

To estimate model inference energy on the Redmi Note 11 S (MediaTek Helio G96: 2×Cortex-A76 + 6×Cortex-A55), we assume the active high-performance core (A76) operates used which oprate at 0.75 W.[22] With our measured latency of 14.06,ms per modality (two modalities → 28.12 ms per verification), the inference energy is $E = P \times t = 0.75\,\text{W} \times 0.02812\,\text{s} = 21.09\,\text{mJ}$ per verification. This is about 0.00003% of the device's battery (∼69,300 J, from 5000 mAh at 3.85 V), which is a negligible portion of the battery per inference.

Usability aspects of XR integration

Even with a fully implemented prototype, usability remains a pivotal factor for our proposed multimodal authentication system. To ensure the success of a biometric authentication system, it is crucial to understand user perspectives on the modality, its usability, and user concerns. Both brainwave and eye movement are emerging biometric modalities, and fully implemented solutions are not yet available in the market. Therefore, prioritizing usability is crucial as it greatly affects initial impressions and the subsequent adoption of these technologies.

In brainwave authentication usability studies, Chuang et al. [50] and Arias-Cabarcos et al. [17] found visual tasks more appealing than reading or auditory ones. Similarly, Röse et al. [27] and Fallahi et al. [26] confirmed a preference for visual tasks in their usability research with mockup prototypes. For eye movement, Brooks et al. [28] showed that users found PIN entry simpler but considered the eye movement dot task more secure, generally preferring it over reading tasks and PIN entry. Fallahi et al. [26] also reported high usability scores of 78.8 in the SUS scale for the dot task, rated as "good" (A^-) based on Bangor et al. [51] and Sauro et al. [52]. Their results on eye movement and brainwave-based authentication indicate that users value usability, security, and passwordlessness as major advantages, while performance limitations and device overhead are seen as major disadvantages [26].

[14] https://www.bitbrain.com/neurotechnology-products/dry-eeg.

[15] https://www.neurospec.com/Products/Details/1078/dsi-7.

[16] https://www.neurospec.com/Products/Details/1077/dsi-vr300.

[17] https://galea.co/.

[18] https://browser.geekbench.com/v6/cpu/11030128.

[19] https://browser.geekbench.com/v6/cpu/3523396

[20] https://browser.geekbench.com/v6/cpu/4760218.

[21] Note that model size differs from runtime memory usage. MobileFaceNets report model size but do not provide runtime memory usage.

[22] https://en.wikichip.org/wiki/arm_holdings/microarchitectures/cortex-a76.

We selected the dot task as a visually appealing and usable task, and chose extended reality (XR) as the use case, where there is already an assumption of wearing a headset. The potential integration of brainwave and eye movement sensors into these headsets could eliminate the need for additional physical hardware for users. Thus, with a usable dot task and the elimination of extra hardware, our paper focuses on how the fusion of these two modalities could address performance concerns, which are a major consideration for users.

Privacy and ethical risks

Advances toward practical applications are increasing attention to the ethical implications of collecting and using both brainwave and eye movement data for authentication. Prior work has identified significant privacy concerns among researchers [53, 54] and users [26, 27], highlighting the need for robust protection mechanisms. EEG can reveal highly sensitive personal information, such as emotional states [55], medical conditions [56], attention levels [57], and gender [58]. Similarly, eye movement data can reveal cognitive load, reading patterns, personal interests, and medical conditions such as neurological disorders [59], even when collected for authentication purposes. The collection of such data raises the risk of harmful breaches and introduces opportunities for misuse. For example, an honest-but-curious authentication provider could conduct unauthorized behavioral profiling or exploit involuntary physiological responses for commercial gain. These risks are compounded by a general lack of user awareness regarding the sensitivity of brain and eye movement data when using consumer-grade XR devices. Addressing these concerns will require privacy-preserving biometric techniques such as cancelable biometrics, which transform biometric data into an intentionally distorted representation that maintains recognition accuracy but can be regenerated with a new transformation if the stored template is compromised [60]. Homomorphic encryption enables authentication computations to be performed directly on encrypted brainwave or eye movement data, ensuring that raw signals remain inaccessible to the authentication provider [61].

Revocability and coercion attacks

Revocability and coercion resistance are critical for biometric authentication.

Coercion

Coercion affects all authentication methods. We distinguish (1) intentional compliance (voluntary or under duress), and (2) attempts without the user's cooperation. For the latter, our method requires deliberate fixation on randomly positioned targets; authentication succeeds only when measured gaze aligns with the expected target sequence, a behavior expected to be reliably performed only by the legitimate user. In addition, brainwave signals can indicate affective states such as fear [62], and eye movement patterns also reflect emotion [63]. Thus, duress detection is feasible in principle and may trigger lockout or additional verification during suspected coercion. Knowledge-based methods (e.g., passwords and PINs) likewise cannot resist intentional compliance.

Revocability and replay

As described in "Technological and methodological blueprint" section, each trial uses fresh, randomly placed stimuli and requires intentional fixation; comparing recorded fixations with the expected target sequence renders prerecorded data ineffective (replay resistance). Revocability can be provided via cancelable biometrics [60], which apply non-reversible transforms so compromised templates can be replaced. However, for static traits (e.g., fingerprints), if raw data leaks from another source, the underlying trait cannot be changed, leaving systems vulnerable to presentation attacks. In contrast, our modality—especially EEG—allows issuing a new template by slightly changing the task or visual context, which measurably alters the biometric signal and yields a distinguishable template [25].

Related work

In the context of biometric authentication, our study distinguishes itself by focusing on consumer-grade devices that could be used in XR settings. The following discussion elucidates how our contributions relate to existing work in three pivotal domains: brainwave authentication, eye movement authentication, and multimodal authentication approaches.

Brainwave authentication

Considerable effort has been made to optimize performance in the domain of brainwave authentication. Nakanishi et al. [37] conducted an experiment with a sample size of 10 subjects and achieved a 4.4% EER. Arias-Cabarcos et al. [16] expanded the sample size to 50 subjects and achieved a 14.5% EER. In a subsequent study, they improved their results to 8.5% EER through enhanced machine learning techniques [17]. They used the same machine learning pipeline and a sample size of 40 subjects to achieve a 1.9% EER on the medical dataset ERP CORE [65]. These advances in brainwave authentication demonstrate the significant impact of data quality on performance outcomes. Most recently, Fallahi et al. [33] used again the ERP CORE dataset but with a triplet loss twin neural network to further improve performance to

Table 3 Comparative analysis of eye movement authentication studies

Publication	Eye movement authentication				
	S.C.	Device	P.D.	S.R. (Hz)	EER (%)
Zhang et al. [18]	30	Glasses	✓	50	6.9
Sluganovic et al. [11]	30	Desktop	×	500	6.3
Eberz et al. [10]	22	Desktop	✓	500	1.88
Lohr et al. [8]	322	Desktop	×	1000	3.66
Zhao et al. [64]	48	Desktop	×	100	4.3
Our work, Eye	30	Glasses	×	200	3.64
Our work, Eye	30	Glasses	✓	200	1.82
Our work, Eye + EEG	30	Glasses	✓	200	0.298

S.C., Subjects count; Device, Tracking hardware; P.D., Pupil diameter in data; S.R., Sampling rate; EER, Equal error rate

1.37% EER. Similarly, promising results are available on other medical datasets; for example, 0.14% EER [33], 0.19% EER [34], 0.04–11% EER [66], and 1.96% EER [36].

Our study utilizes the same network architecture as BrainNet [33] to achieve a 4.92% EER. Our performance surpasses that of Arias-Cabarcos et al. [17], and while it is essential to acknowledge that our EER is marginally higher by 0.5% compared to the study by Nakanishi et al. [37], we should note that, unlike our approach, Nakanishi et al. did not consider the unknown attacker [67] scenario in their methodology. It is true that our performance did not reach the accuracy levels of medical-grade devices. However, our application is not in the medical sector but instead in XR, where integration of medical-grade devices, which are often bulky and complex to set up, is impractical.

Eye movement authentication

In Table 3, we present key parameters relevant to eye movement authentication in our related works, including device type (desktop or glasses), pupil diameter, number of subjects, EER, and sample rate.

Pupil diameter as a feature

Both Zhang et al. [18] and Eberz et al. [10] incorporated pupil diameter as a feature into their models. In contrast, Sluganovic et al. [11], Lohr et al. [8], and Zhao et al. [64] specifically excluded pupil diameter from their feature sets. Our research indicates that incorporating pupil diameter can enhance model performance by more than 50%, thus making it an important factor to consider in comparisons and feature selection.

EER comparison

Table 3 indicates that our EER outcomes closely align with the state-of-the-art results in unimodal approaches. Our study yielded an EER of 1.820% when pupil diameter data were incorporated and an EER of 3.639% when not. Despite the lower sampling rate of our device, our results compare well to those presented by Eberz et al. [10] and Lohr et al. [8]. We attribute this comparable performance to our refined comparison strategy. Specifically, our S3 strategy uses the remaining rounds for the enrollment set and adheres, as well, to a best-match scenario. This strategy effectively mitigates the effect of noise samples in the enrollment set. By contrast, when we use our simpler S1 strategy, the EER rates rise to 13.420% and 19.070%, which underscores the significant role played by sample rate in outcomes.

Further, our results also underscore the comparative advantage gained by incorporating brainwave data. Our EERs of 0.298% and 0.686% showcase the effectiveness of augmenting multimodal authentication, specifically by brainwave data.

Multimodal authentication

Multiple studies have been conducted on multimodal authentication systems; however, many focus on modalities that are not well-suited for XR environments. For example, Chakladar et al. [68] used EEG and signature-based methods, Zhang et al. [69] combined EEG and gait, Zheng et al. [70] utilized fingerprints with photoplethysmography, and Ammour et al. [71] relied on ECG and fingerprints. In contrast, we found two studies more closely related to our work. First, Wu et al. [72] explored the use of voice and lip movements as biometrics, achieving a 95% True Positive Rate (TPR) and detecting 93.47% of attacks (TNR) with 104 subjects. It is plausible to assume that XR devices could be equipped with a camera to capture lip movements. Second, Peng et al. [73] proposed a system based on voice and hand motion, reporting a 99% TPR and a 0.5% FRR with 32 subjects.

For performance comparison, the referenced studies did not report EER, but we can make approximations based on FAR and FRR. In Wu et al.'s work [72], they reported a 5% FAR and 6.53% FRR, leading us to conclude that their system has at least a 5% EER. Similarly, Peng et al. [73] noted a 1% FAR and 0.5% FRR, suggesting a minimum EER of 0.5% (probably in the middle of 0.5 and 1). Comparing these to our best results—0.686% EER without pupil diameter data and 0.298% EER with pupil diameter data—our system outperforms Wu et al. and is more effective than Peng et al. when considering the pupil diameter-based scenario. Moreover, unlike lip cameras, which serve no additional function, EEG and eye-tracking can be applied to various other applications, including human-computer interfaces [46] and entertainment [48]. When compared to voice and hand motion-based systems, we contend that our approach offers the advantage of being hands-free and potentially more effective in a crowded setting.

In a somewhat related study

Krishna et al. [74] explored the feasibility of using EEG and eye movement data for multimodal biometric authentication. Their approach involved combining two independent datasets related to eye movement and brainwave patterns to create a hypothetical multimodal dataset. While the idea is interesting, its practical applicability is doubtful. For the purposes of authentication, the point is not to amalgamate data from multiple subjects but to identify distinct characteristics which are unique to an individual. Furthermore, the results of Krishna et al. [74] showed poor performance in the eye movement modality, with a FAR of 7.4% at an FRR of 36.7%. These results led to no noticeable improvements over unimodal authentication. In fact, the only enhancement observed by the authors occurred in scenarios with 'low-confidence predictions of EEG.' Consequently, despite the innovative aspects of their approach, we do not categorize theirs as a multimodal authentication system based on eye movement and brainwave data.

Comparison with VR authentication works

Reliable authentication in XR environment remains an open challenge. Non-biometric mechanisms, typically classified as knowledge-based or token-based approaches. Knowledge-based methods (e.g., alphanumeric passwords, graphical passwords, gesture patterns) often require memorization and manual input, which is particularly cumbersome in XR and disrupts immersion [75]. For instance, users tend to choose simpler alphanumeric passwords in XR to ease entry using virtual keyboards, which poses significant security risks [76]. Gesture-based inputs may be vulnerable to shoulder-surfing attacks, especially in public or shared spaces and long-term memorability [77–79]. Graphical passwords, while promising in theory, suffer from poor long-term memorability and need for manual input as well. Token-based methods, such as device proximity or the use of external smartphones, face usability challenges due to the need to carry additional hardware and are susceptible to issues such as loss, theft, or unintentional activation [80]. Compared to these alternatives, biometric methods offer seamless, better alignment with XR's immersive paradigm, and eliminate reliance on external devices or memorized secrets.

Biometrics have emerged as a viable solution; however, certain biometric methods require physical activities that may reduce usability and practicality, ranging from discrete hand gestures [81, 82] to more active ones like walking [83, 84] or throwing a virtual ball [85, 86]. Additionally, some solutions utilize sensors that may not be well-suited for XR environments [87]; for instance, Chen et al. [87] employed electrical muscle stimulation, which relies on sensors attached to the hands.

While several studies propose biometrics suitable for XR setups, such as free head and body movement [88], eye-related biometrics [18, 89], skull conductance [90], and brainwaves [91]; their reported EERs range from 2.5% to 7%, comparable to our single-modality results (1.9%–4.9%). However, as shown in Table 2, higher security configurations result in increased FRR. Therefore, the multimodal approach can be a promising alternative to improve performance further.

Multi-factor approaches [92, 93], such as combining biometrics with knowledge-based methods, aim to address performance challenges but inherit the limitations of both factors. An alternative is multimodal biometrics [73, 84], which uses multi biometrics. We adopted this approach by fusing eye movement and brainwaves, reducing the error significantly.

Limitations

Our research faces two primary limitations: sample size and single-session data collection. First, although our sample size is comparable to that used in similar studies in this field [10, 11, 18, 33, 72], it is relatively small when considering broader biometric research such as face and fingerprint recognition. Increasing the sample size could enhance our learning model and facilitate more realistic evaluations. Second, while relying on data from a single session is common in our research domain [10, 17, 37, 72, 73], this approach risks overfitting, which may degrade the performance of our methods in real-world applications over time. To address this, our experiment design avoids using samples from the same round for both enrollment and verification, and incorporates a 15-s rest period between rounds. While we anticipate a higher EER in multi-session scenarios [8, 94], it is important to highlight that our primary contribution lies in demonstrating that the fusion of brainwave and eye movement can improve performance significantly, rather than achieving a specific error rate. Now that we showed feasibility, future work could explore robustness in bigger and varied datasets. Furthermore, prior studies have shown that both EEG and eye movement modalities can exhibit temporal stability across days, months, or even years [8, 24, 95], supporting the long-term viability of these modalities with performance comparable to our single-modality results.

Conclusion and future work

In this study, our investigation shows that this combination of brainwaves and eye movement yields highly promising results. Through our research, we substantially improved authentication accuracy and enhanced resistance against zero-effort attacks. Specifically, our multimodal authentication system achieved an EER of 0.298% and 0.686%, along with FRR of 3.8% and 11.6% at FAR of 0.01% compared with 25% and 52.8% FRR in unimodal eye movement authentication. These results offer a higher level of security with a reasonable FRR, ensuring a smooth user experience without unnecessary disruptions. We provide a straightforward, hands-free authentication method that is both suitable for XR settings and appropriate for consumer-grade devices. Our multimodal authentication system improves authentication accuracy while also holding promise for broader adoption in real-world applications.

In the future, investigating multi-session scenarios and increasing the sample size are essential steps for improving model learning and conducting more comprehensive evaluations. Furthermore, exploring additional tasks alongside the interactive dot task will contribute to a more extensive understanding of multimodal biometric authentication. Moreover, while the dot task can effectively resist against replay attacks, it would be beneficial to explore whether and how it is possible to use correlation in synchronized data to ensure that brainwave and eye movement data are recorded simultaneously, adding an extra layer of security.

Appendix

To examine the generalizability of the findings across different neural network architectures and to further explore the performance of individual modalities, we evaluate our approach using both ShallowNet [96] and ResNet1D [97]. Tables 4 and 5 present the results. The findings confirm similar trends observed with the BrainNet architecture: pupil data has a positive effect, score fusion outperforms feature fusion, mean and product methods perform better than other score fusion strategies, method A performs better than method B in feature fusion, and the user-specific threshold yields better results than other scenarios.

Table 4 This table shows the EER of the biometric authentication system using three approaches on the ShallowNet model: single modality, score fusion, and feature fusion

		EER (%)		
Approach	**Biometric**	**S1**	**S2**	**S3**
Single Biometric	Brainwaves	18.10	8.435	7.375
	Eye-tracking with pupil	21.41	2.319	2.151
	Eye-tracking without pupil	26.65	3.724	3.471
Score Fusion	Max. with pupil	14.95	4.211	3.492
	Min. with pupil	17.10	1.380	1.237
	Mean with pupil	11.76	0.811	0.696
	Product with pupil	11.38	0.925	0.736
	Max. without pupil	18.49	3.269	3.034
	Min. without pupil	17.09	3.981	3.339
	Mean without pupil	13.59	1.672	1.346
	Product without pupil	13.54	1.231	1.007
Feature Fusion	Architecture A with pupil	12.94	2.030	1.575
	Architecture B with pupil	15.26	2.555	2.149
	Architecture A without pupil	15.09	3.113	2.504
	Architecture B without pupil	15.93	4.161	3.368

The columns represent different authentication strategies, specifically: S1—Fixed Threshold with One Sample as the enrollment set; S2—Fixed Threshold with the remainder of the samples used for enrollment; S3—User-specific Threshold per subject with the remainder of the samples used for enrollment

Table 5 This table shows the EER of the biometric authentication system using three approaches on the ResNet1D model: single modality, score fusion, and feature fusion

		EER (%)		
Approach	**Biometric**	**S1**	**S2**	**S3**
Single Biometric	Brainwaves	17.78	8.112	6.720
	Eye-tracking with pupil	15.41	3.480	3.131
	Eye-tracking without pupil	24.49	7.341	7.067
Score Fusion	Max. with pupil	13.13	3.797	3.049
	Min. with pupil	13.04	2.649	2.260
	Mean with pupil	10.51	1.268	1.100
	Product with pupil	9.84	1.253	1.010
	Max. without pupil	21.91	6.972	6.649
	Min. without pupil	17.11	6.935	5.704
	Mean without pupil	13.13	3.118	2.319
	Product without pupil	13.19	2.582	2.028
Feature Fusion	Architecture A with pupil	9.90	1.414	1.054
	Architecture B with pupil	10.90	1.735	1.443
	Architecture A without pupil	12.06	2.198	1.687
	Architecture B without pupil	12.53	2.868	2.389

The columns represent different authentication strategies, specifically: S1—Fixed Threshold with One Sample as the enrollment set; S2—Fixed Threshold with the remainder of the samples used for enrollment; S3—User-specific Threshold per subject with the remainder of the samples used for enrollment

Acknowledgements We thank the technicians of KASTEL Security Research Labs for their assistance with research communication and technical support.

Author contributions Matin Fallahi collected the data, developed the methodology, implemented and conducted the experiments, analyzed the data, and drafted the manuscript. Patricia Arias-Cabarcos conceived the initial idea, developed the methods, supervised study, and contributed to writing–review and editing. Thorsten Strufe conceived the initial idea, developed the methods, supervised the study, and contributed writing–review and editing. All authors read and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. This work was supported by the Helmholtz Association (HGF) within Topic 46.23 "Engineering Secure Systems" (KASTEL Security Research Labs) and by Germany's Excellence Strategy (EXC 2050/1 "CeTI"; ID 390 696 704).

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethics approval This study adheres to responsible research practices by maintaining ethical integrity. All procedures, methodologies, consent form, and tools underwent rigorous scrutiny and were approved by Karlsruhe Institute of Technology university's Institutional Review Board (IRB).

Consent to participate Each participant provided written informed consent by signing a form approved by the Institutional Review Board (IRB), granting permission for their data to be analyzed.

References

1. Ratcliffe J, Soave F, Bryan-Kinns N, Tokarchuk L, Farkhatdinov I (2021) Extended reality (XR) remote research: a survey of drawbacks and opportunities. In: Proceedings of the 2021 CHI conference on human factors in computing systems, pp 1–13
2. Rauschnabel PA, Felix R, Hinsch C, Shahab H, Alt F (2022) What is XR? Towards a framework for augmented and virtual reality. Comput Hum Behav 133:107289
3. Fast-Berglund Å, Gong L, Li D (2018) Testing and validating extended reality (XR) technologies in manufacturing. Proc Manuf 25:31–38
4. Alnagrat A, Ismail RC, Idrus SZS, Alfaqi RMA (2022) A review of extended reality (XR) technologies in the future of human education: current trend and future opportunity. J Hum Centered Technol 1(2):81–96

5. Mäkinen H, Haavisto E, Havola S, Koivisto J-M (2022) User experiences of virtual reality technologies for healthcare in learning: an integrative review. Behav Inf Technol 41(1):1–17
6. Ansari SZA, Shukla VK, Saxena K, Filomeno B (2022) Implementing virtual reality in entertainment industry. In: Cyber intelligence and information retrieval: proceedings of CIIR 2021. Springer, Berlin, pp 561–570
7. Stephenson S, Pal B, Fan S, Fernandes E, Zhao Y, Chatterjee R (2022) Sok: authentication in augmented and virtual reality. In: 2022 IEEE symposium on security and privacy (SP). IEEE, pp 267–284
8. Lohr D, Komogortsev OV (2022) Eye know you too: toward viable end-to-end eye movement biometrics for user authentication. IEEE Trans Inf Forensics Secur 17:3151–3164
9. Liebers J, Schneegass S (2020) Gaze-based authentication in virtual reality. In: ACM symposium on eye tracking research and applications, pp 1–2
10. Eberz S, Lovisotto G, Rasmussen KB, Lenders V, Martinovic I (2019) 28 blinks later: tackling practical challenges of eye movement biometrics. In: Proceedings of the 2019 ACM SIGSAC conference on computer and communications security, pp 1187–1199
11. Sluganovic I, Roeschlin M, Rasmussen KB, Martinovic I (2018) Analysis of reflexive eye movements for fast replay-resistant biometric authentication. ACM Trans Priv Secur (TOPS) 22(1):1–30
12. Lohr D, Berndt S-H, Komogortsev O (2018) An implementation of eye movement-driven biometrics in virtual reality. In: Proceedings of the 2018 ACM symposium on eye tracking research & applications, pp 1–3
13. Abinaya R, Indira D, Swarup Kumar J (2022) Multimodal biometric person identification system based on speech and keystroke dynamics. In: International conference on computing, communication, electrical and biomedical systems. Springer, Berlin, pp 285–299
14. Bugdol MD, Mitas AW (2014) Multimodal biometric system combining ECG and sound signals. Pattern Recogn Lett 38:107–112
15. Gui Q, Ruiz-Blondet MV, Laszlo S, Jin Z (2019) A survey on brain biometrics. ACM Comput Sur (CSUR) 51(6):1–38
16. Arias-Cabarcos P, Habrich T, Becker K, Becker C, Strufe T (2021) Inexpensive brainwave authentication: new techniques and insights on user acceptance. In: 30th USENIX security symposium (USENIX Security 21), pp 55–72
17. Arias-Cabarcos P, Fallahi M, Habrich T, Schulze K, Becker C, Strufe T (2023) Performance and usability evaluation of brainwave authentication techniques with consumer devices. ACM Trans Priv Secur 26(3):1–36
18. Zhang Y, Hu W, Xu W, Chou CT, Hu J (2018) Continuous authentication using eye movement response of implicit visual stimuli. Proc ACM Interact Mob Wearable Ubiquitous Technol 1(4):1–22
19. Mathis F, Williamson J, Vaniea K, Khamis M (2020) Rubikauth: fast and secure authentication in virtual reality. In: Extended abstracts of the 2020 CHI conference on human factors in computing systems, pp 1–9
20. Düzgün R, Noah N, Mayer P, Das S, Volkamer M (2022) Sok: a systematic literature review of knowledge-based authentication on augmented reality head-mounted displays. In: Proceedings of the 17th international conference on availability, reliability and security, pp 1–12
21. Chan P, Halevi T, Memon N (2015) Glass OTP: secure and convenient user authentication on Google glass. In: Financial cryptography and data security: FC 2015 international Workshops, BITCOIN, WAHC, and Wearable, San Juan, Puerto Rico, January 30, 2015, Revised Selected Papers. Springer, Berlin, pp 298–308
22. Li S, Ashok A, Zhang Y, Xu C, Lindqvist J, Gruteser M (2016) Whose move is it anyway? Authenticating smart wearable devices using unique head movement patterns. In: 2016 IEEE international conference on pervasive computing and communications (PerCom). IEEE, pp 1–9
23. Boutros F, Damer N, Raja K, Ramachandra R, Kirchbuchner F, Kuijper A (2020) Iris and periocular biometrics for head mounted displays: segmentation, recognition, and synthetic data generation. Image Vis Comput 104:104007
24. Maiorana E (2021) Learning deep features for task-independent EEG-based biometric verification. Pattern Recogn Lett 143:122–129
25. Lin F, Cho KW, Song C, Xu W, Jin Z (2018) Brain password: a secure and truly cancelable brain biometrics for smart headwear. In: Proceedings of the 16th annual international conference on mobile systems, applications, and services, pp 296–309
26. Fallahi M, Arias-Cabarcos P, Strufe T (2025) On the usability of next-generation authentication: a study on eye movement and brainwave-based mechanisms. In: Proceedings of the extended abstracts of the CHI conference on human factors in computing systems, pp 1–14
27. Röse M, Kablo E, Arias-Cabarcos P (2023) Overcoming theory: designing brainwave authentication for the real world. In: Proceedings of the 2023 European symposium on usable security, pp 175–191
28. Brooks M, Aragon CR, Komogortsev OV (2013) Perceptions of interfaces for eye movement biometrics. In: 2013 international conference on biometrics (ICB). IEEE, pp 1–8
29. Peirce JW (2009) Generating stimuli for neuroscience using psychopy. Front Neuroinform 2:343
30. Fernández LGC (2017) Brain-motion interaction analysis through the study of brainwaves. In: 2017 international conference on mechatronics, electronics and automotive engineering (ICMEAE). IEEE, pp 26–30
31. Bélanger NN, Rayner K (2015) What eye movements reveal about deaf readers. Curr Dir Psychol Sci 24(3):220–226
32. Meier M, Mason C, Porzel R, Putze F, Schultz T (2018) Synchronized multimodal recording of a table setting dataset. In: Proceedings of the IROS 2018 workshop on latest advances in big activity data sources for robotics & new challenges (Madrid)
33. Fallahi M, Strufe T, Arias-Cabarcos P (2023) Brainnet: improving brainwave-based biometric recognition with Siamese networks. In: 2023 IEEE international conference on pervasive computing and communications (PerCom). IEEE, pp 53–60
34. Schons T, Moreira GJ, Silva PH, Coelho VN, Luz EJ (2018) Convolutional network for EEG-based biometric. In: Progress in pattern recognition, image analysis, computer vision, and applications: 22nd Iberoamerican Congress, CIARP 2017, Valparaíso, Chile, November 7–10, 2017, Proceedings 22. Springer, Berlin, pp 601–608
35. Schroff F, Kalenichenko D, Philbin J (2015) Facenet: a unified embedding for face recognition and clustering. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 815–823
36. Bidgoly AJ, Bidgoly HJ, Arezoumand Z (2022) Towards a universal and privacy preserving EEG-based authentication system. Sci Rep 12(1):2531
37. Nakanishi I, Maruoka T (2019) Biometric authentication using evoked potentials stimulated by personal ultrasound. In: 2019 42nd international conference on telecommunications and signal processing (TSP). IEEE, pp 365–368
38. Zhang T, Ji Q, Ye Z, Rahman MM, Akanda R, Mahdad AT, Shi C, Wang Y, Saxena N, Chen Y (2024) SAFARI: Speech-associated facial authentication for AR/VR settings via robust vibration signatures. In: Proceedings of the 2024 ACM SIGSAC conference on computer and communications security. CCS ’24. Association for Computing Machinery, New York

39. Zhao BZH, Asghar HJ, Kaafar MA (2020) On the resilience of biometric authentication systems against random inputs. arXiv preprint arXiv:2001.04056
40. Pagnin E, Dimitrakakis C, Abidin A, Mitrokotsa A (2014) On the leakage of information in biometric authentication. In: International conference on cryptology in India. Springer, Berlin, pp 265–280
41. Cohen I, Huang Y, Chen J, Benesty J, Benesty J, Chen J, Huang Y, Cohen I (2009) Pearson correlation coefficient. In: Noise reduction in speech processing, pp 1–4
42. Welch P (1967) The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. IEEE Trans Audio Electroacoust 15(2):70–73
43. Pastel S, Marlok J, Bandow N, Witte K (2023) Application of eye-tracking systems integrated into immersive virtual reality and possible transfer to the sports sector—a systematic review. Multimed Tools Appl 82(3):4181–4208
44. Li H, Shin H, Zhang M, Yu A, Huh H, Kwon G, Riveira N, Kim S, Gangopadahyay S, Peng J et al (2023) Hair-compatible sponge electrodes integrated on VR headset for electroencephalography. Soft Sci. https://www.oaepublish.com/articles/ss.2023.11%C2%A0
45. Gardony AL, Lindeman RW, Brunyé TT (2020) Eye-tracking for human-centered mixed reality: promises and challenges. In: Optical architectures for displays and sensing in augmented, virtual, and mixed reality (AR, VR, MR), vol 11310. SPIE, pp 230–247
46. Aggarwal S, Chugh N (2022) Review of machine learning techniques for EEG based brain computer interface. Arch Comput Methods Eng 29:3001–3020
47. Sundstedt V (2022) Gazing at games: an introduction to eye tracking control. Springer, Berlin
48. Queiroz Cavalcanti D, Melo F, Silva T, Falcão M, Cavalcanti M, Becker V (2023) Research on brain–computer interfaces in the entertainment field. In: International conference on human–computer interaction. Springer, Berlin, pp 404–415
49. Chen S, Liu Y, Gao X, Han Z (2018) Mobilefacenets: efficient CNNS for accurate real-time face verification on mobile devices. In: Chinese conference on biometric recognition. Springer, Berlin, pp 428–438
50. Chuang J, Nguyen H, Wang C, Johnson B (2013) I think, therefore i am: usability and security of authentication using brainwaves. In: Financial cryptography and data security: FC 2013 workshops, USEC and WAHC 2013, Okinawa, Japan, April 1, 2013, Revised Selected Papers 17. Springer, Berlin, pp 1–16
51. Bangor A, Kortum P, Miller J (2009) Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 4(3):114–123
52. Sauro J (2011) Are both positive and negative items necessary in questionnaires? online publication. Url: https://measuringu.com/positive-negative/, Accessed 07 Aug 2018
53. Höller Y, Uhl A (2018) Do EEG-biometric templates threaten user privacy? In: Proceedings of the 6th ACM workshop on information hiding and multimedia security, pp 31–42
54. Wang M, Wang S, Hu J (2022) Polycosgraph: a privacy-preserving cancelable EEG biometric system. IEEE Trans Depend Secure Comput. https://ieeexplore.ieee.org/abstract/document/9935312/
55. Wang X-W, Nie D, Lu B-L (2014) Emotional state classification from EEG data using machine learning approach. Neurocomputing 129:94–106
56. Sánchez-Reyes L-M, Rodríguez-Reséndiz J, Avecilla-Ramírez GN, García-Gomar M-L, Robles-Ocampo J-B (2021) Impact of EEG parameters detecting dementia diseases: a systematic review. IEEE Access 9:78060–78074
57. Hassan R, Hasan S, Hasan MJ, Jamader MR, Eisenberg D, Pias T (2020) Human attention recognition with machine learning from brain-EEG signals. In: 2020 IEEE 2nd Eurasia conference on biomedical engineering, healthcare and sustainability (ECBIOS). IEEE, pp 16–19
58. Niu Y, Chen X, Chen Y, Yao Z, Chen X, Liu Z, Meng X, Liu Y, Zhao Z, Fan H (2024) A gender recognition method based on EEG microstates. Comput Biol Med 173:108366
59. Wang Y, Li X (2024) Eye movement tracking in ocular, neurological, and mental diseases. Front Neurosci 18:1364078
60. Ragendhu S, Thomas T, Emmanuel S (2024) Cancelable biometric template generation using random feature vector transformations. IEEE Access 12:32064–32079
61. Acar A, Aksu H, Uluagac AS, Conti M (2018) A survey on homomorphic encryption schemes: theory and implementation. ACM Comput Surv 51(4):1–35
62. Serna B, Salazar R, Alonso-Silverio GA, Baltazar R, Ventura-Molina E, Alarcón-Paredes A (2025) Fear detection using electroencephalogram and artificial intelligence: a systematic review. Brain Sci 15(8):815
63. Skaramagkas V, Giannakakis G, Ktistakis E, Manousos D, Karatzanis I, Tachos NS, Tripoliti E, Marias K, Fotiadis DI, Tsiknakis M (2021) Review of eye tracking metrics involved in emotional and cognitive processes. IEEE Rev Biomed Eng 16:260–277
64. Zhao Y, Song X, Huang X, Tian S, Zhou Y (2025) EV-GazeLock: a user authentication system based on micro eye movement with event cameras. ACM Trans Sens Netw. https://dl.acm.org/doi/abs/10.1145/3719010
65. Kappenman ES, Farrens JL, Zhang W, Stewart AX, Luck SJ (2021) ERP CORE: an open resource for human event-related potential research. Neuroimage 225:117465
66. Chaurasia AK, Fallahi M, Strufe T, Terhörst P, Cabarcos PA (2024) Neuroidbench: an open-source benchmark framework for the standardization of methodology in brainwave-based authentication research. J Inf Secur Appl 85:103832
67. Mansfield AJ, Wayman JL (2002) Best practices in testing and reporting performance of biometric devices. https://face-rec.org/databases/Mansfield02.pdf
68. Chakladar DD, Kumar P, Roy PP, Dogra DP, Scheme E, Chang V (2021) A multimodal-Siamese neural network (MSNN) for person verification using signatures and EEG. Inf Fusion 71:17–27
69. Zhang X, Yao L, Huang C, Gu T, Yang Z, Liu Y (2020) Deepkey: a multimodal biometric authentication system via deep decoding gaits and brainwaves. ACM Trans Intell Syst Technol 11(4):1–24
70. Zheng XX, Taha B, Rahman MMU, Masood M, Hatzinakos D, Al-Naffouri T (2025) Multimodal biometric authentication using camera-based PPG and fingerprint fusion. Pattern Recogn Lett . https://www.sciencedirect.com/science/article/pii/S0167865525002454
71. Ammour N, Bazi Y, Alajlan N (2023) Multimodal approach for enhancing biometric authentication. J Imaging 9(9):168. https://www.sciencedirect.com/science/article/pii/S0167865525002454
72. Wu L, Yang J, Zhou M, Chen Y, Wang Q (2019) LVID: a multimodal biometrics authentication system on smartphones. IEEE Trans Inf Forensics Secur 15:1572–1585
73. Peng G, Zhou G, Nguyen DT, Qi X, Yang Q, Wang S (2016) Continuous authentication with touch behavioral biometrics and voice on wearable glasses. IEEE Trans Hum-Mach Syst 47(3):404–416
74. Krishna V, Ding Y, Xu A, Höllerer T (2019) Multimodal biometric authentication for VR/AR using EEG and eye tracking. In: Adjunct of the 2019 international conference on multimodal interaction, pp 1–5
75. Riyadh H, Bhardwaj D, Dabrowski A, Krombholz K (2024) Usable authentication in virtual reality: exploring the usability of pins and gestures. In: International conference on applied cryptography and network security. Springer, Berlin, pp 412–431

76. Kablo E, Last Y, Cabarcos PA, Volkamer M (2025) The (un) suitability of passwords and password managers in virtual reality. arXiv preprint arXiv:2503.18550
77. Länge T, Matheis P, Düzgün R, Volkamer M, Mayer P (2024) Vision: towards fully shoulder-surfing resistant and usable authentication for virtual reality. In: Proceedings of the 2024 symposium on usable security and privacy (USEC)
78. Mathis F, O'Hagan J, Khamis M, Vaniea K (2022) Virtual reality observations: using virtual reality to augment lab-based shoulder surfing research. In: 2022 IEEE conference on virtual reality and 3d user interfaces (VR). IEEE, pp 291–300
79. Xu W, Li X, Tian J, Xiao Y, Qu X, Wang S, Ji X (2018) Which one to go: security and usability evaluation of mid-air gestures. arXiv preprint arXiv:1811.10168
80. Al-Ameen MN, Fatema K, Wright M, Scielzo S (2015) The impact of cues and user interaction on the memorability of System-Assigned Recognition-Based graphical passwords. In: Eleventh symposium on usable privacy and security (SOUPS 2015), pp 185–196
81. Peng G, Zhou G, Nguyen DT, Qi X, Yang Q, Wang S (2016) Continuous authentication with touch behavioral biometrics and voice on wearable glasses. IEEE Trans Hum-Mach Syst 47(3):404–416
82. Chauhan J, Asghar HJ, Kaafar MA, Mahanti A (2014) Gesture-based continuous authentication for wearable devices: the Google glass case. arXiv preprint arXiv:1412.2855
83. Shen Y, Wen H, Luo C, Xu W, Zhang T, Hu W, Rus D (2018) Gaitlock: protect virtual and augmented reality headsets using gait. IEEE Trans Dependable Secure Comput 16(3):484–497
84. Pfeuffer K, Geiger MJ, Prange S, Mecke L, Buschek D, Alt F (2019) Behavioural biometrics in vr: Identifying people from body motion and relations in virtual reality. In: Proceedings of the 2019 CHI conference on human factors in computing systems, pp 1–12
85. Ajit A, Banerjee NK, Banerjee S (2019) Combining pairwise feature matches from device trajectories for biometric authentication in virtual reality environments. In: 2019 IEEE international conference on artificial intelligence and virtual reality (AIVR). IEEE Computer Society, pp 9–97
86. Miller R, Ajit A, Banerjee NK, Banerjee S (2019) Realtime behavior-based continual authentication of users in virtual reality environments. In: 2019 IEEE international conference on Artificial Intelligence and Virtual Reality (AIVR). IEEE, pp 253–2531
87. Chen Y, Yang Z, Abbou R, Lopes P, Zhao BY, Zheng H (2021) User authentication via electrical muscle stimulation. In: Proceedings of the 2021 CHI conference on human actors in computing systems, pp 1–15
88. Mustafa T, Matovu R, Serwadda A, Muirhead N (2018) Unsure how to authenticate on your VR headset? Come on, use your head! In: Proceedings of the fourth ACM international workshop on security and privacy analytics, pp 23–30
89. Luo S, Nguyen A, Song C, Lin F, Xu W, Yan Z (2020) OcuLock: exploring human visual system for authentication in virtual reality head-mounted display. In: Proceedings 2020 network and distributed system security symposium. Internet Society, Reston
90. Schneegass S, Oualil Y, Bulling A (2016) Skullconduct: biometric user identification on eyewear computers using bone conduction through the skull. In: Proceedings of the 2016 CHI conference on human factors in computing systems, pp 1379–1384
91. Lin F, Cho KW, Song C, Xu W, Jin Z (2018) Brain password: a secure and truly cancelable brain biometrics for smart headwear. In: Proceedings of the 16th annual international conference on mobile systems, applications, and services, pp 296–309
92. Zhu H, Jin W, Xiao M, Murali S, Li M (2020) Blinkey: a two-factor user authentication method for virtual reality devices. Proc ACM Interact Mob Wearable Ubiquitous Technol 4(4):1–29
93. Lu D, Huang D, Deng Y, Alshamrani A (2018) Multifactor user authentication with in-air-handwriting and hand geometry. In: 2018 International conference on biometrics (ICB). IEEE, pp 255–262
94. Seha SNA, Hatzinakos D (2019) A new approach for EEG-based biometric authentication using auditory stimulation. In: 2019 International Conference on Biometrics (ICB). IEEE, pp 1–6
95. Fallahi M, Arias-Cabarcos P, Strufe T (2025) Advancing brainwave-based biometrics: a large-scale, multi-session evaluation. arXiv preprint arXiv:2501.17866
96. Schirrmeister RT, Springenberg JT, Fiederer LDJ, Glasstetter M, Eggensperger K, Tangermann M, Hutter F, Burgard W, Ball T (2017) Deep learning with convolutional neural networks for EEG decoding and visualization. Hum Brain Mapp 38(11):5391–5420
97. Zheng Y, Liu Z, Mo R, Chen Z, Zheng W, Wang R (2022) Task-oriented self-supervised learning for anomaly detection in electroencephalography. In: International conference on medical image computing and computer-assisted intervention. Springer, Berlin, pp 193–203