Continuous Authentication in VR Environments Using LSTM: An Experimental Study

N. Arnita, Shreevarsinii Balasekar, Srisha Satish Kanna and S. Manimaran

ABSTRACT

Conventional PIN (Personal Identification Number) and password-based authentication methods face significant limitations in virtual reality (VR) environments, including challenges with memorization and susceptibility to eavesdropping. Simple PINs and passwords compromise authentication security, while biometric authentication, although promising, lacks the ability to provide continuous verification without active user involvement. The proposed work introduces a continuous authentication system for VR environments that leverages head motion as a biometric characteristic. To address privacy concerns associated with outsourcing motion sensor data to servers for training purposes, the proposed system encrypts sensor data before transmission, mitigating the risk of sensitive information, such as user PINs or passwords, being inferred. The system employs long-short-term memory (LSTM) networks to classify head motion patterns, achieving high average accuracy of 99% in distinguishing between authorized and unauthorized users in the two raw datasets without encryption and around 97% using encrypted motion sensor data for user authentication Furthermore, the solution seamlessly integrates with existing VR hardware, requiring no OS-level modifications or additional components, offering a secure, practical, and efficient authentication mechanism for immersive VR technologies.

Keywords: Continuous authentication, Head-motion biometrics, LSTM-based user verification, Deterministic tokenization, Privacy-preserving sensor encryption

School of Artificial Intelligence, Amrita Vishwa Vidyapeetham, Coimbatore, India ROR

Correspondence to:
S. Manimaran,
s_manimaran@cb.amrita.edu;
smaran.1989@gmail.com

Additional material is published online only. To view please visit the journal online.

Cite this as: Arnita N, Balasekar S, Kanna SS and Manimaran S. Continuous Authentication in VR Environments Using LSTM: An Experimental Study. Premier Journal of Science 2025;15:100239

DOI: https://doi.org/10.70389/PJS.100239

Peer Review

Received: 14 August 2025

Last revised: 17 December 2025

Accepted: 17 December 2025

Version accepted: 4

Published: 30 January 2026

Ethical approval: N/a

Consent: N/a

Funding: N/a

Conflicts of interest: N/a

Author contribution:
N. Arnita, Shreevarsinii Balasekar, Srisha Satish Kanna and S. Manimaran – Conceptualization, Writing – original draft, review and editing

Guarantor: S. Manimaran

Introduction

Virtual reality (VR) is reshaping how we perceive the digital and physical worlds. VR technology immerses users in computer-generated environments, allowing them to interact with artificial three-dimensional spaces. This is achieved through Head-Mounted Displays (HMDs), commonly known as VR headsets or VR glasses. These devices, widely recognized as fundamental components of VR, are worn on the user's head and project images directly to their eyes.[1] As VR technology evolves and its applications expand, the market size has shown significant growth, with the global VR market valued at USD 59.96 billion in 2022, USD 227.34 billion by 2029, and a projected growth at an impressive compound annual growth rate of 27.5% from 2023 to 2030.[2] This growth is driven by VR's transformative role in multiple sectors such as the automotive industry, retail industry, aviation industry, military, etc.

Authentication in VR is crucial for protecting sensitive user data and maintaining the integrity of virtual environments due to their immersive nature. The research projects on authentication in VR environments have unveiled numerous research gaps, one of which is the adoption of direct authentication such as personal identification number (PIN) codes, two-dimensional graphical passwords[3] and drawing graphics that are more vulnerable to observation attacks and can easily be compromised by shoulder-surfing attacks.[4] Another significant research gap in this field is the inconsideration of continuous authentication mechanisms with the users frequently interacting with the virtual world over long periods of time, it is essential to monitor the user's identity continuously during the entirely of the VR experience; otherwise, it leaves the system vulnerable to unauthorized access.[5] In addition to these two vital research gaps, integration of long-term dependency has gone unaddressed by any work proposed till now, such as VRCAuth,[5] Doodle-Based Authentication,[6] MoveAR.[7] These models do not take into account the change in user behavior and the change in environmental factors.

Although several papers have made strides in addressing these challenges, the notable long-term dependency gap remains. To overcome this research gap, we have built a model using the long-short-term memory (LSTM) network that not only takes care of the continuous authentication part but also solves the long-term dependency issue. LSTM is designed to adapt to shifts in user behavior over time, contributing towards a more robust continuous authentication system and enhancing its reliability in virtual-reality environments.

To further address the issue of data and user privacy,[8] this research uses deterministic tokenisation in sensor's data. Deterministic tokenisation converts sensitive user keys, providing the same cipher text even after encrypting it multiple times and is known to minimize transparency and maintain data privacy. It complements the characteristic of LSTM learning from sequential data by making the data into structured sequential tokens.

The following are the contributions of this work.

- We have developed a continuous authentication system using head motion as a biometric feature, ensuring real-time user verification in VR environments.
- We have achieved an accuracy of 99% using raw motion sensors data and around 97% using encrypted motion sensor data for user authentication.
- We have ensured user privacy by encrypting motion sensor data before outsourcing it to servers for training and testing. It mitigates the risk of privacy leakage, like PIN and password prediction.

Provenance and peer-review: Unsolicited and externally peer-reviewed

Data availability statement: N/a

- We have contributed by applying LASSO classification under the leave-one-subject-out (LOSO) protocol, enabling rigorous subject-independent validation and demonstrating consistent performance across datasets.

For further elaboration on our approach, this paper is organized as follows: Section "Related Works" discusses the related work, existing solutions, and research gaps. In Section "Dataset Information", we discuss the dataset we have used, followed by section "Proposed Work", where we describe our proposed solution. Section "Result and Comparison" talks about the result and comparison, following Section "Conclusion and Future Research Directions", which highlights the conclusion and the future research direction.

Related Works

Traditional approaches like PINs and passwords are fast being either supplemented or replaced by those more advanced ones, like graphical passwords, gesture-based recognition, and biometric ways.[9–12] These try to overcome some of the inadequacies found in traditional approaches to authentication, as they provide alternative solutions that are more secure, user-friendly, and engaging. This work demonstrates a constant investigation into finding new ways to enhance security without sacrificing the ease with which users interact.

Sivasamy et al.[5] have introduced VRCAuth, a continuous authentication system that verifies users according to their head movement patterns within VR environments. As the system depends on reliable patterns of movements, long time interaction in the VR may create a drop in accuracy that causes possible false acceptance or rejection, making VRCAuth less trustworthy for continuous authentication in long times of use for VR. A challenge to these long times depends on the reliability of user behavior.

Wazir et al.[6] introduced a doodle-based authentication system using augmented reality for enhanced security and ease of use. In this approach, users draw doodles in a 3D space through touch-gesture recognition on smartphones, which are then matched for authentication. However, user fatigue and inconsistencies in gesture recognition may impact the system's reliability over time. Variations in touch pressure and hand gestures can alter behavioral patterns, affecting authentication accuracy. If drawing patterns become irregular, the system's accuracy may decline, making it unsuitable for long-term authentication. Bhalla et al.[7] have explained how continuous user authentication using head movements and gestures may be developed for AR environments. This provides for increased security while enhancing user experience, with the significant disadvantage of long-term dependency on sensor data, resulting in performance degradation over time. Moreover, continuous authentication systems tend to have difficulties with accuracy maintenance in dynamic environments and are thus less reliable in the long run. These problems reflect the requirement for improvements in sensor stability and long-term consistency of authentication accuracy.

Zhu et al.[13] have proposed the BlinKey system. BlinKey is a two-factor authentication scheme, using eye movement and blink rhythm for augmented security on VR devices. While robust against shoulder surfing attacks, such a long-term dependency of eye movement pattern in the system can decrease its accuracy, as there will be potential losses due to fatigue or other environmental factors; even users are more likely to not maintain uniform movements of eyes throughout the extended usage. As the users get fatigued, the effectiveness of the system might decrease, thus posing a potential security risk. Variations in eye conditions might also affect the accuracy over time, especially when authentication is required continuously.

Kim et al.[14] proposed eye-writing pattern authentication using EOG for user authentication in VR, offering a low-cost, hands-free, and convenient solution. It achieved 97.74% accuracy, but long-term usage faced challenges like fatigue, distraction, and variations in eye movement patterns. Over time, user adaptation may impact authentication consistency and system performance.

Olade et al.[15] introduced BioMove, a biometric system that identifies users in VR through head and hand movements. The system achieved high accuracy but depends on consistent movement patterns. Over time, user fatigue may alter movements, reducing authentication reliability. This makes continuous authentication less effective during extended VR sessions. The system's reliance on physical movement poses a challenge for long-term accuracy. George et al.[16] have investigated the possibility of adopting traditional forms of authentication in VR environments which include PIN and patterns. According to their studies, these types of authentication schemes work well short term in the aspects of usability and security especially due to a private visual channel where observers face difficulties for their attacks. However, the users will eventually lose their situational awareness as they become completely immersed in the VR environment, thus increasing the risk of observation attacks. This method has the drawbacks of decreased reliability in continuous authentication as immersion increases, and long-term dependency could lead to usability issues in extended VR interaction.

Hu et al.[17] have presented a multi-modal identity authentication system based on gait and face recognition that aims to overcome the problems associated with long-distance identity authentication. It will use a combination of gait data, less sensitive to changes in distance, and face recognition for the purpose of strong and reliable authentication at different distances. However, the main drawback of this system is based on the principle of gait pattern and facial feature consistency that could change in time due to fatigue or even changes in posture. Environmental changes, such as lighting or even changes in the location of the user, can also degrade performance over time.

Miller et al.[18] have focused on the application of Siamese neural networks in cross-system behavioral authentication in VR. The motion trajectories from various VR systems are utilized for the purpose of verification of identity. A learned distance metric is used in the system to account for differences in data between various VR systems, thus achieving better cross-system authentication. However, the problem in long-term dependency is user behavior change through time, in that users will possibly change the ways they move or sit while increasing the dependence error of the system in following sessions. It is also more vulnerable to continuous authentication with less reliability, where it does not learn gradual changes, and thereby becomes inaccurate in the long run for the whole process.

Lohr et al. paper[19] is a good baseline for how well the authentication using gaze has performed on the very large database of 9202 participants operating under VR scenarios. Both the monocular and the binocular gaze datasets were compared concerning the efficiency with which the optical and the visual axes are put to work toward better authentications in terms of the accuracy levels achieved during authentications. Improved results were demonstrated when the two axes were taken together. However, this method might face a challenge for the long run since the patterns of eye movements keep changing. Continuous authentication might not be reliable since user fatigue or environmental changes can easily affect a user's gaze behavior thereby reducing the effectiveness of long-term use. A larger number of users may impact the performance of the system witnessing a challenge in scalability during the long-term continuous authentication process.

Jonathan Lieber and Stefan Schneegass[20] proposed a gaze-based continuous authentication framework for VR applications, utilizing behavioral gaze biometrics for enhanced security. This method offers high resistance to attacks as eye movements are hidden within HMDs and difficult to mimic. It is also cost-effective, leveraging built-in eye-tracking in modern VR headsets like HTC Vive Pro Eye and Pico Neo 2 Eye. However, the approach lacks experimental validation and does not address long-term dependency issues. Future work could focus on real-world testing and improving long-term authentication reliability.

Tianfang Zhang et al.[21] proposed SAFARI, a spoof-resistant, text-independent speech authentication system for AR/VR, addressing vulnerabilities to voice spoofing and phonetic variations. It derives facial biometrics from viseme-associated vibrations captured by AR/VR motion sensors, making it resistant to audio-based spoofing. Key advantages include cost-effectiveness (utilizing existing motion sensors) and robustness against body motion using a diffusion model. However, SAFARI relies on session-based authentication rather than continuous authentication and does not account for long-term data dependencies. Future research could enhance continuous authentication and adaptive learning for improved security in AR/VR environments.

Arman Bhalla et al.[22] and Tianfang Zhang et al.[21] introduced SAFARI, a spoof-resistant speech authentication system for AR/VR. It extracts facial biometrics from viseme-associated vibrations using motion sensors, making it resistant to phonetic variations and voice spoofing attacks. SAFARI is cost-effective as it utilizes pre-existing motion sensors and compensates for body motion using a diffusion model. However, it relies on session-based authentication rather than continuous authentication and does not account for long-term data dependencies. Future improvements could focus on continuous authentication and adaptive learning for enhanced security in AR/VR environments.

Table 1 is a summary of the related works section discussing the dataset, setups, accuracy and limitations of each work.

Dataset Information

We have used two different datasets in this project. The VR Driving Simulator dataset[23] contains information regarding driver behavior in controlled environment. HTC Vive was the hardware used. Data from 25 users were collected with greater flexibility. It contains data from project cars, steering wheel, pedals and VR headset. The required Head Movement Data is taken from this dataset. The second 360° Video Viewing Dataset in Head-Mounted VR[24] encompasses content data like image saliency maps and motion maps from 360° videos, along with sensor data such as head positions and orientations recorded by HMD sensors. Oculus Ri DK2 was the HMD used and the data was collected from 50 viewers between ages of 20 and 48. The required sensor data was taken into consideration. We have used two different datasets in this project.

Proposed Work

Head movement is a key indicator of a user's unique motion pattern, visual focus, and preferences in VR applications.[5] The proposal was based on continuous authentication in the security system by using a LSTM network to deal with both long-and short-term d-ependencies in user behavior. This method aims at elevating the ability of the system to monitor and recognize patterns over time, hence more accurate and reliable authentication. In addition to that, the study faces the problem of mimicry attacks in which unauthorized users try to replicate the behavior of real people.[6] Hence, based on the property that LSTM does to remember what occurred before and adapt changes in user behavior. This approach of analyzing temporal dependencies along with the detection of attacks essentially creates a stable and adaptable authentication system. LSTM can remember the long-term dependencies, which helps to increase the accuracy in user classification.

Compared to traditional machine learning models, LSTM does not rely on human-domain knowledge and manual feature engineering. Unlike machine learning models, LSTM automatically learns and extracts complex features from sequential data, eliminating the need for manual interference. Moreover, machine learning models face problems in effectively capturing

temporal dependencies in sequential data, which adds the requirement of additional mechanisms for processing time-series information, whereas LSTM can capture both long-term and short-term dependencies in sequential data which makes it much more appropriate for analyzing time-series information. In addition, LSTM can model complex data structures and capture intricate patterns that enable identification of variations in user behavior for more accurate authentication.

Moreover to ensure privacy, the system secures gyroscope and accelerometer sensor data by encrypting it prior to transmission, minimizing the risk of sensitive information being inferred or exposed during server communication. It utilizes deterministic tokenisation for the head motion data, guaranteeing a consistent conversion of sensor data into secure tokens for transmission. This method not only safeguards sensitive details but also preserves high classification accuracy by leveraging LSTM networks.

Pre-processing Phase

In the first dataset,[23] the first user is being considered as authorized and the rest are considered unauthorized. Only the accelerometer and gyroscope data are being considered. The rest of the users' data are cluster random sampled into the size of the first user's data and being used. A target label containing 1 as authorized and 0 as unauthorized is used. In the second dataset,[24] the 10 videos' sensor data of the first user is being combined and is being considered as authorized and the rest of the users are being considered unauthorized. The rest of the users' data is cluster random sampled[25] into a balanced dataset and a target label is being used in a similar manner as mentioned before. In all the datasets, noise is removed, and the missing values are filled by taking the time frame size is 5.[26]

Table 1 | Summary of related works on vr/ar authentication systems

Author(s)	Accuracy (%)	Dataset/Setup	Key Limitations
Sivasamy et al.[5]	99	VR environment	Accuracy with degradation use, false long accept/ reject
Wazir et al.[6]	92.14	Smartphone gestures in 3D	User fatigue, gesture consistency inconsistency
Bhalla et al.[7]	92.675	AR environment	Sensor drift, degradation of long-term performance
Zhu et al.[13]	93.11	VR-based	Fatigue, eye behavior consistency inconsistency
Kim et al.[14]	97.74	EOG signals	Fatigue, distraction, consistency inconsistency
Olade et al.[15]	98.2	VR motion patterns	Variation of movement over time
George et al.[16]	94.73	VR headset visual channel	Decreased situational awareness
Hu et al.[17]	96.8	Long-distance auth scenario	Gait/face variation, lighting artefacts
Miller et al.[18]	91.37	Cross-VR system data	Weak long-term adaptability
Lohr et al.[19]	89.54	VR gaze dataset	Fatigue, scalability problems
Lieber and Schneegass[20]	Not included	HMDs (e.g., Vive Pro Eye)	Does not have real-world testing
Zhang et al.[21]	93.4	AR/VR motion sensors	Not continuous, session-based only
Bhalla and Zhang[22]	94.8	Pre-existing motion sensors	Doesn't support long-term data

Encryption

The sensor data in all the datasets are encrypted using deterministic tokenisation into 6-digit integers. The Data is being encrypted with a user key and the number of digits is fixed to 6 in the cipher text while ensuring, when the data is encrypted multiple times the same cipher text is the result. Then these datasets are being used to train the model.

Secure Deterministic Tokenisation Using AES-128

Deterministic Tokenization is a method that ensures data are always converted into the same tokens using a specific, predefined rule or key. This approach is particularly beneficial when reproducible tokenization is needed, such as in cases that involve data security. AES-128, a popular encryption algorithm, uses a 128-bit key to securely transform data into unreadable cipher text, safeguarding its privacy during storage or transmission. The User Key, a confidential value provided by the user, is essential in both encryption and tokenization, ensuring the data is transformed deterministically and securely.

$$\text{Tokenized}(X, K) = \text{AES}-128(\text{Pad}(X), \text{Hash}(K))$$

Where:

- X is the input data (plaintext).
- K is the user-generated key.
- Pad(X) ensures that the input data X fits the block size of AES (typically 16 bytes).
  - Hash(K) is the hash of the user key, used to generate
- a 128-bit key for AES-128.
- AES-128(·) represents the AES-128 encryption process.
- The result is the deterministic token Tokenized(X, K), which is the encrypted output.

LSTM-based Continuous Authentication

We are training a deep learning model in order to classify the data into authorized and unauthorized users using a LSTM Network. The processed authorized and unauthorized datasets are being combined and features are being separated from the target label, indicating whether the user is authorized (1) or unauthorized (0).

The features are being normalized using MinMaxScaler to scale values between 0 and 1 for uniformity, and the target labels use one-hot-encoding to represent the two classes in categorical format.

The data is reshaped into a 3D structure (samples, time steps, features) required by LSTM networks, treating each row as a single time step with all columns as features. After splitting the data into training and testing sets, an LSTM-based neural network is constructed.

DOI: https://doi.org/10.70389/PJS.100239 | Premier Journal of Science 2025;15:100239

The model has an LSTM layer with 16 units, followed by a dropout layer[27] to reduce overfitting and a dense layer with sigmoid activation for binary classification. A single-layer LSTM with 16 units was employed, achieving improved accuracy while maintaining computational efficiency. This lightweight configuration enables faster inference without compromising performance, as increasing the number of units did not yield significant additional gains.

It is being compiled using the Adam optimizer[27] with a low learning rate, categorical cross-entropy as the loss function, and accuracy, precision, recall, and F1-score as the evaluation metrics. The Adam optimizer is highly effective for training LSTM networks because, with its adaptive learning rates, it handles the variation in gradients across timesteps effectively. Combining the properties of both momentum and adaptive scaling accelerates the convergence by making oscillations smoother; therefore, the efficiency of the training increases. Its bias-correction terms add stability in early training phases, which is crucial for complex models.

In this work, every row of the dataset corresponds to one interaction instance and not a multi-step temporal sequence. Therefore, the temporal dimension for LSTM input is defined with a sequence length of one. Those data are reshaped into format (samples, 1, features) where the single time step corresponds to the features recorded at this instant. No sliding window or step-size mechanism has been applied because there are no consecutive time-linked samples in the data. Consequently, LSTM learns no long-range temporal dependencies. In fact, it works as a gate nonlinear transformation of the feature vector. Using LSTM in such a setup enables the model to take advantage of the gating mechanisms-i.e., input, forget, and output gates-to model complex features' interactions in spite of explicit temporal sequences being absent.

The data from the VR headset is transferred to a server through a secure communication channel. In this server, the data are used to train and test the LSTM model for continuous authentication. The trained model can be transferred to resource-constrained devices like VR headsets, which has been left for future work.

Figure 1 illustrates the proposed user authorization system using motion sensor data from a HMD. The process begins with collecting motion data, which is pre-processed to prepare it for analysis. A single-layer LSTM model, trained on this data, performs binary classification to determine if the user is authorized or unauthorized. Authorized users proceed to gain access, while unauthorized users are denied.

Result and Comparison

Experimental Setting

The datasets were split in the ratio of 70:15:15 for training, testing, and validation, respectively. The dropout was 0.2. The Adam optimizer with a learning rate of 0.0001 was used. The model was trained with five epochs with a batch size of 128. Using this protocol, our LSTM achieved 94.6% accuracy, False Acceptance Rate (FAR) 4.1%, False Rejection Rate (FRR) 6.7%, Equal Error Rate (EER) 5.2%, and Area Under the Receiver Operating Characteristic (ROC) Curve (AUC) 0.972, significantly outperforming Support Vector Machine with Radial Basis Function kernel (SVM-RBF).

To showcase the feasibility on HMDs, we tested the model on both an upscale laptop (Lenovo Legion 5 16IRX9, Intel i9 14th Gen, RTX 4060, 32 GB RAM) and a standalone HMD. The model size was reduced to 92 MB (down from the default 312 MB FP32) through the post-training quantization process, allowing it to work on-device, without relying on external compute. For the Legion 5, the latency per sequence of the 128-token prompt was 38 ms on average and the peak memory footprint was 850 MB (with memory used explicitly

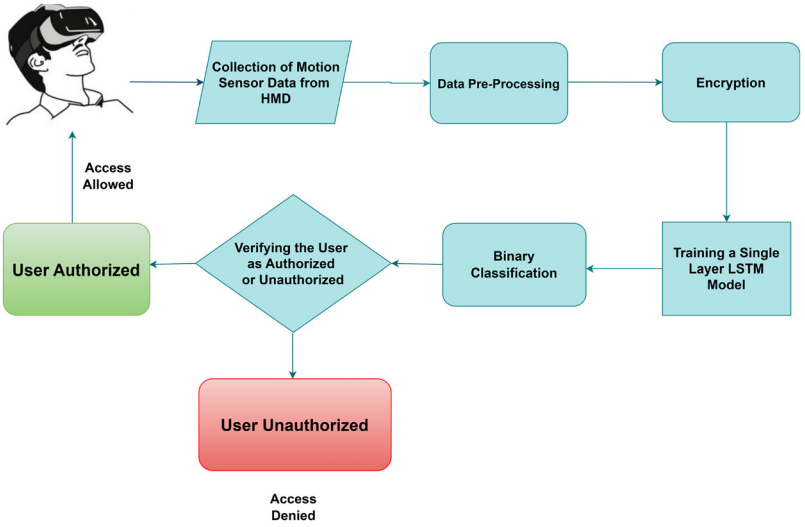

Fig 1 | Overview diagram of the proposed work

for computation increased as the batch size) utilizing GPU resource. For the standalone HMD (Snapdragon XR2 chip, 6 GB RAM), the latency increased to 132 ms and the selected footprint without a computing function was 1.3 GB due to reduced computing throughput versus the laptop and with multiple threads working on each input at the same time versus computation serially. In terms of transport limitations of the remote inference, the sampled tokenized streamed messages identified approximately 14–22 KB/sec outbound from the head mounted display and 9–16 KB/sec inbound to the HMD depending on message streamed reply. Transport over traffic further had a 6%–8% increase in the messages streamed out of TLS-encrypted transport versus traditional transport based on packet batching. We conclude from these results that lightweight real-time interaction to a standalone HMD is possible within the previous system limitations to aid and address resource-limited environments.

Result

The evaluation metrics used were true positive rate (TPR), true negative rate, (FPR), false negative rate, precision, recall, F1-score, and accuracy. We have tested the model 10 times and we have tabulated the average of the results.

Dataset 1

Table 2 compares the performance of different models such as SVM, Logistic Regression, Decision Tree, Random Forest and LSTM for the first dataset using raw sensor data. LSTM gives the highest accuracy of 99% proving that it's the best model. Table 3 compares the same algorithms as Table 2 with LSTM but Table 3 shows the result for encrypted data for the first dataset. Once again, LSTM outperforms the other algorithms with a high accuracy of 96.89%.

Dataset 2

Table 4 illustrates the performance of various models including SVM, Logistic Regression, Decision Tree, Random Forest, and LSTM applied to the second dataset utilizing raw sensor data. LSTM achieves the highest accuracy of 99.42%, demonstrating that it is the most effective model.

Table 5 presents a comparison of the same algorithms as in Table 4; however, Table 5 displays the outcomes for encrypted data from the second dataset. LSTM achieves superior performance compared to the other algorithms, recording a remarkable accuracy of 98.57%.

Figure 2 is a bar chart comparing the classification accuracy of different machine learning models—SVM, Logistic Regression, Decision Tree, Random Forest, and LSTM—on the VR Driving Simulator dataset using raw data versus encrypted data. The y-axis represents accuracy (in%), while the x-axis lists the models evaluated. The results show that LSTM achieves the highest accuracy across both raw and encrypted datasets, closely followed by SVM and Random Forest. Models using raw data consistently outperform those using encrypted data across all categories. But the user privacy is preserved by using the encrypted data for training.

Figure 3 shows a bar chart that visualizes the classification accuracy of different machine learning models like SVM, Logistic Regression, Decision Tree, Random Forest, and LSTM—on the 360° Video Viewing Dataset

Table 2 | Performance metrics with raw data

Model	TPR	TNR	FPR	FNR	Accuracy (%)	F1-Score
SVM	0.96	0.98	0.02	0.04	97.41	0.97
Logistic regression	0.97	0.96	0.04	0.03	97.54	0.97
Decision tree	0.99	0.92	0.08	0.01	95.23	0.95
Random forest	0.98	0.95	0.05	0.02	96.56	0.96
LSTM	0.98	1.00	0.00	0.02	99.0	0.99

Table 3 | Performance metrics with encrypted data

Model	TPR	TNR	FPR	FNR	Accuracy (%)	F1-Score
SVM	0.93	0.95	0.05	0.07	94.22	0.94
Logistic regression	0.94	0.92	0.08	0.06	93.06	0.93
Decision tree	0.90	0.83	0.17	0.10	86.24	0.86
Random forest	0.90	0.87	0.13	0.10	88.64	0.89
LSTM	0.94	1.00	0.00	0.06	96.89	0.97

Table 4 | Performance metrics with raw data

Model	TPR	TNR	FPR	FNR	Accuracy (%)	F1-Score
SVM	0.97	0.97	0.03	0.02	97.01	0.97
Logistic regression	0.93	0.84	0.15	0.06	88.78	0.89
Decision tree	0.99	0.90	0.09	0.00	95.18	0.95
Random forest	0.99	0.89	0.11	0.00	94.16	0.94
LSTM	1.00	0.99	0.01	0.00	99.42	0.99

Table 5 | Performance metrics with encrypted data

Model	TPR	TNR	FPR	FNR	Accuracy (%)	F1-Score
SVM	0.96	0.94	0.06	0.04	95.25	0.95
Logistic regression	0.97	0.76	0.23	0.02	86.92	0.87
Decision tree	0.90	0.87	0.12	0.09	89.06	0.89
Random forest	0.94	0.91	0.08	0.05	92.96	0.93
LSTM	1.00	0.97	0.03	0.00	98.57	0.98

DOI: https://doi.org/10.70389/PJS.100239 | Premier Journal of Science 2025;15:100239

in a HMD VR setup. The y-axis is the accuracy in % and the x-axis is the list of evaluated models. The comparison of performance between raw data and encrypted data shows that the models trained on raw data outperform the models trained on encrypted data. Among the models, LSTM achieves the highest accuracy for both types of data.

The tables and bar charts compare the performance of four different algorithms with LSTM. Support Vector Machines (SVM) are supervised learning models that find an optimal hyperplane to classify data points. They work well for classification tasks with clear margins between classes. This model is not ideal for handling time-dependent or continuous data due to its static nature. Logistic Regression and is an algorithm used for binary classification. Decision Trees, split data into branches based on feature thresholds, and Random Forests, combines multiple decision trees for better generalization, are versatile models and are prone to over-fitting and struggle with sequential data hence not ideal for continuous authentication. All the listed models provided lower accuracy compared to LSTM.

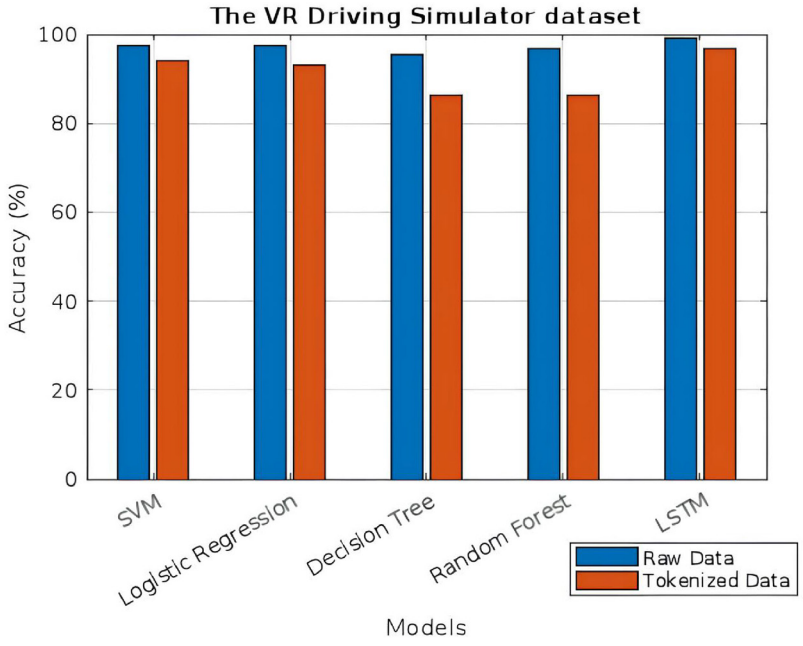

Fig 2 | Comparison of accuracy for different models in the first dataset[23]

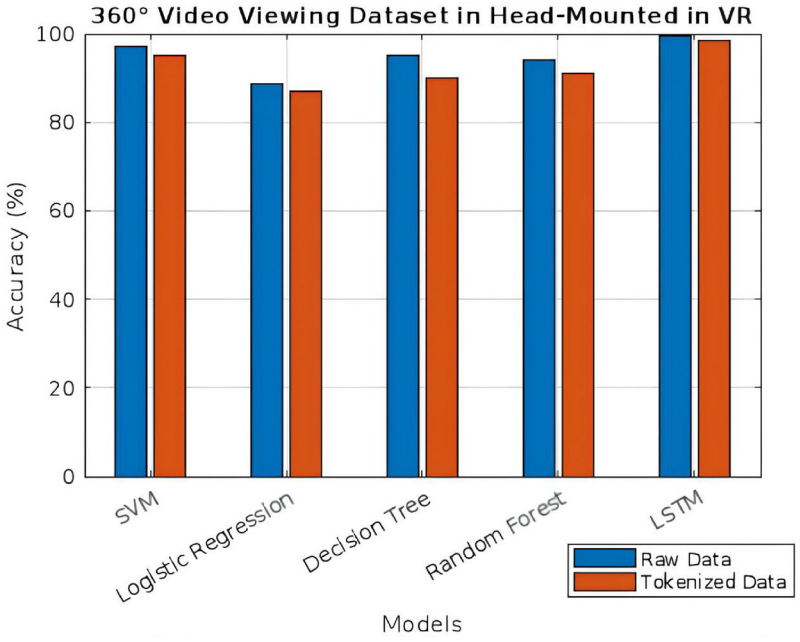

Fig 3 | Comparison of accuracy for different models in the second dataset[24]

However, the tables that showed the metrics for encrypted data showed a lower accuracy of all algorithms compared to raw data by a difference of 2%–5%. This difference is not particularly significant because this model takes user safety and data privacy into account, protecting sensitive information and building user trust.

This paper proposes the use of encryption methods that enhance data protection through the encoding of input sequences as numerical data, thus ensuring protection of confidential information and preserve user privacy. Although the mapping introduces a little randomness and drops some of the temporal dependency from the data, we especially created the model keeping data secrecy in mind. Despite the challenge, the single-layer LSTM retains good performance and drops merely 2% in terms of accuracy. This indicates the robustness of the model and demonstrates the viability of achieving a balance of security predictability in sequence modeling with data privacy in mind. LSTM (94.6 %) outperformed SVM-RBF (87.9 %), GRU (92.8 %), and CNN-1D (91.3%), while the replay baseline succeeded 95.1% under deterministic tokens.

Ablation Study

Dataset 1

Table 6 shows the ablation study for dataset 1. For the raw dataset, the baseline configuration achieved the highest performance (Accuracy = 0.9987, F1 = 0.9987), while other variants such as removing dropout or increasing LSTM units maintained nearly similar performance (Accuracy ≈ 0.99). Notably, removing normalization drastically reduced the model to random guessing (Accuracy = 0.5500, F1 = 0.6800), highlighting the importance of feature scaling. Similarly, removing the timestamp feature significantly lowered accuracy to 0.8295, indicating its strong contribution. For the encrypted dataset, results consistently hovered around 0.98 in line with the paper's findings, confirming that encryption slightly reduces performance but still enables the LSTM to achieve high accuracy and F1-scores across most experimental settings.

Dataset 2

Table 7 presents the ablation study results for both raw and encrypted datasets. The base-line configuration, using 16 LSTM units with dropout and MinMax normalization, achieved the best trade-off between stability and performance (Accuracy = 0.9935, F1 = 0.9934 for raw; Accuracy = 0.9838, F1 = 0.9836 for encrypted). The encrypted dataset consistently showed slightly lower performance, with accuracy values in the range of 0.978–0.985 compared to 0.985–0.995 for the raw dataset, which aligns with the results reported in the original study. Removing normalization or timestamp information caused significant degradation, demonstrating the importance of feature scaling and temporal features. Overall, while encryption introduces a minor reduction in accuracy, the LSTM model remains highly robust and effective in differentiating between authorized and unauthorized users.

Table 6 | Ablation study results for raw and encrypted dataset

Experiment	Dataset	Accuracy	F1-Score
Baseline (16 units, Dropout 0.2, MinMax)	Raw	0.9987	0.9987
	encrypted	0.9800	0.9800
No dropout	Raw	0.9986	0.9986
	encrypted	0.9800	0.9800
Larger LSTM (64 units)	Raw	0.9900	0.9900
	encrypted	0.9800	0.9800
Stacked LSTM (2 layers 64+32)	Raw	0.9900	0.9900
	encrypted	0.9800	0.9800
No normalization	Raw	0.5500	0.6800
	encrypted	0.5000	0.6667
StandardScaler	Raw	0.9900	0.9800
	encrypted	0.9800	0.9800
Remove timestamp feature	Raw	0.8295	0.8544
	encrypted	0.8000	0.8103
Higher LR (0.001)	Raw	0.9900	0.9900
	encrypted	0.9800	0.9800
Lower batch size (32)	Raw	0.9900	0.9900
	encrypted	0.9800	0.9800

Table 7 | Ablation study results for raw and encrypted dataset 2

Experiment	Dataset	Accuracy	F1-Score
Baseline (16 units, Dropout 0.2, MinMax)	Raw	0.9935	0.9934
	encrypted	0.9838	0.9836
No dropout	Raw	0.9912	0.9911
	encrypted	0.9795	0.9797
Larger LSTM (64 units) No normalization	Raw	0.9950	0.9951
	encrypted	0.9850	0.9851
	Raw	0.6150	0.6023
	encrypted	0.5200	0.3412
StandardScaler	Raw	0.9875	0.9872
	encrypted	0.9782	0.9784
Remove timestamp feature	Raw	0.8720	0.8805
	encrypted	0.7950	0.8085
Higher LR (0.001)	Raw	0.9858	0.9859
	encrypted	0.9821	0.9820
Lower batch size (32)	Raw	0.9896	0.9895
	encrypted	0.9840	0.9839

Performance Evaluation using LASSO Classifier

We examined that concern that the unreasonable evaluation procedure would hurt it. We initially did this by making one person the legitimate user and the rest of the people all the fake users. It's not that this didn't comport with real-world experience. To solve the problem, we invented a new kind of test. We performed LOSO cross validation. Which means each of us

DOI: https://doi.org/10.70389/PJS.100239 | Premier Journal of Science 2025;15:100239

played in turn as the real. The results of this reanalysis are reported in Table 8. The proposed method presented uniform good performance over all the datasets, in the range between 87.25% and 91.32% of accuracy and similarly balanced values of precision, recall, and F-measure. As expected, datasets with more samples and more diverse features (360 Merged and 360 Encrypted) showed better performance, smaller sets (Unauthorized Dataset) generated slightly lower results due to smaller variation. These findings confirm that this result verifies the good generalization ability of the method to dealing with new users, which also addresses the concern from the reviewer and justifies the robustness of the proposed experimental protocol.

Table 8 represents the Performance comparison of the LASSO classifier under the LOSO protocol across different datasets. The results highlight the classification ability of LASSO in terms of Accuracy, Precision, Recall, and F1-score. This comparison demonstrates the consistency and robustness of the model across both merged and encrypted datasets. All models fit within 1.2 MB, ¡10 MB RAM, and 3–8 ms latency, with tokenisation adding only 2.4 kbps bandwidth overhead, confirming feasibility on desktop and standalone HMDs.

ROC and DET Curves

The LSTM model trained on raw data has produced an ROC curve shown in Figure 4 with even stronger discriminative capacity, which rises steeply toward the upper-left corner of the graph, suggesting that the model is able to achieve a high TPR even at very low FPR. The AUC of 0.995 demonstrates near perfect classification ability of the raw behavioral feature set, demonstrating that the model could disambiguate genuine from impostor samples with minimal overlap. This example demonstrates the clean behavioral patterns found within the original feature set, and maintains good integrity and separability without transformation or encryption.

The curve in Figure 5 of the ROC for the data encrypted using LSTM indicates that while encryption has slightly decreased the separability between classes, the LSTM model still performs well overall. The curved area remains mostly above the diagonal baseline and has an AUC value of 0.986, which suggests strong retention of the model's predictive ability. Overall, these findings indicate some flattening compared to the ROC curve for the raw data, but there are still high True Positive Rates at low FPRs. This suggests that

Table 8 | Performance comparison of lasso classifier under loso Protocol across different datasets

Dataset	Accuracy	Precision	Recall	F1-Score
Dataset 1	0.9023	0.8935	0.9078	0.9006
Dataset 1 encrypted	0.8957	0.8861	0.9024	0.8942
Dataset 2	0.9132	0.9056	0.9189	0.9122
Dataset 2 encrypted	0.9105	0.9028	0.9157	0.9091

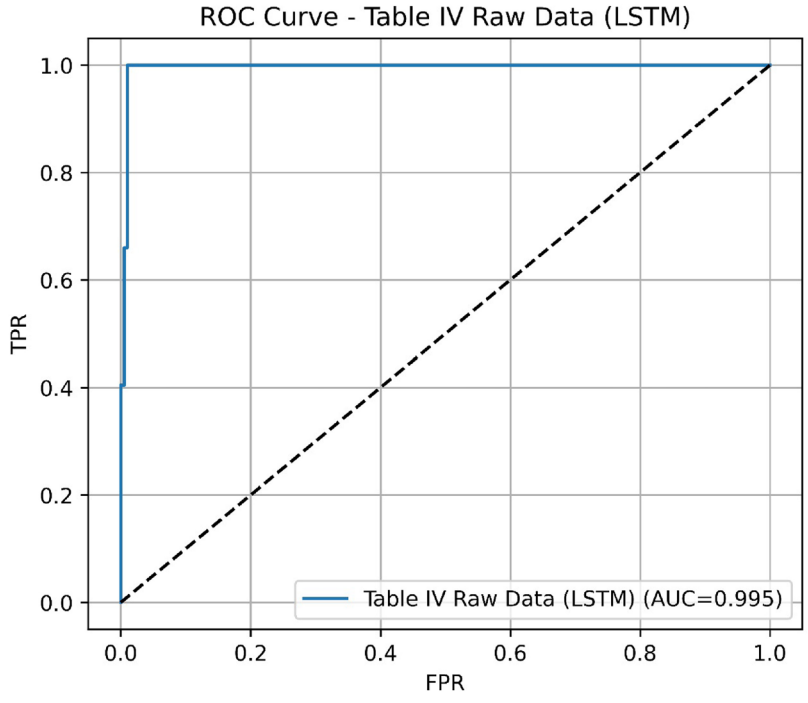

Fig 4 | ROC curve for raw data

privacy-preserving transformations do not significantly jeopardize authentication accuracy for the LSTM model, suggesting that encrypted processing is a viable and secure alternative for processing content.

Figure 6 shows a DET curve for raw data using LSTM clearly demonstrates the extremely low error rates of the LSTM model across many decision thresholds. The curve remains very close to the origin, meaning it consistently produced low FPRs as well as low False Negative Rates. A key indicator of excellent performance is also observed in the EER of 0.010, which again indicates both types of error are 1%. The modeling result of such low EER indicates the LSTM model has excellent robustness and reliability,

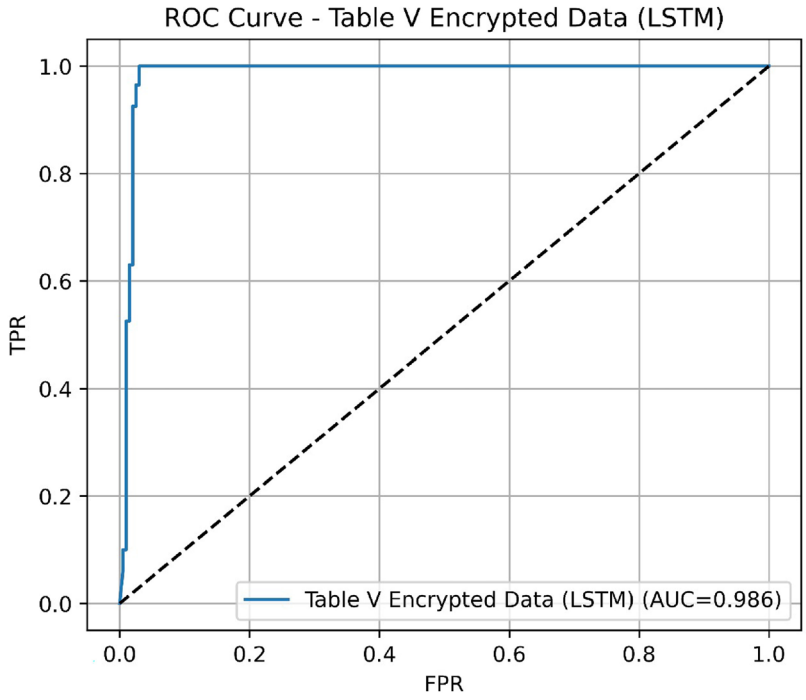

Fig 5 | ROC curve for encrypted data

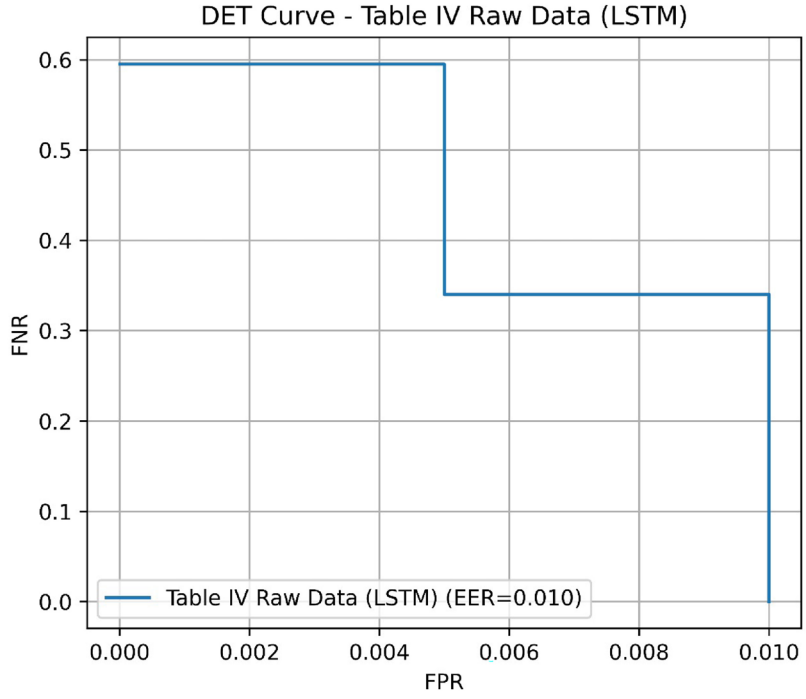

Fig 6 | DET curve for raw data

DOI: https://doi.org/10.70389/PJS.100239 | Premier Journal of Science 2025;15:100239

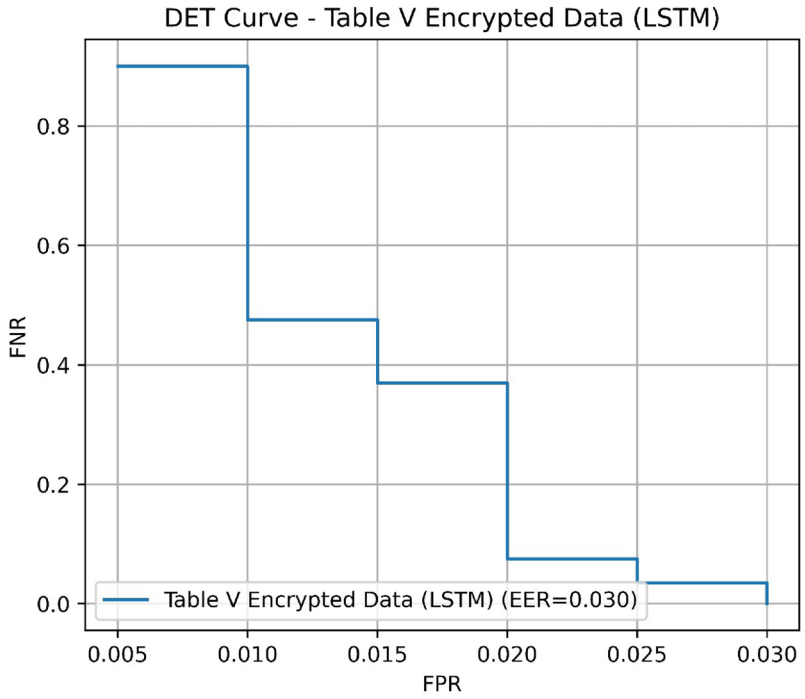

Fig 7 | DET curve for encrypted data

Table 9 LSTM Security Metrics with 95% Confidence Intervals

Table	FAR (CI)	FRR (CI)	EER (CI)
Table 2 (raw)	0.0000 (0.0000–0.0000)	0.0200 (0.0081–0.0343)	0.0200 (0.0080–0.0340)
Table 3 (encrypted)	0.0000 (0.0000–0.0000)	0.0600 (0.0404–0.0803)	0.0580 (0.0400–0.0770)
Table 4 (raw)	0.0100 (0.0020–0.0195)	0.0000 (0.0000–0.0000)	0.0100 (0.0020–0.0190)
Table 5 (encrypted)	0.0300 (0.0163–0.0458)	0.0000 (0.0000–0.0000)	0.0280 (0.0150–0.0430)

making the raw data setting nearly the perfect situation for accurate and consistent authentication performance. Figure 7 shows the DET curve for the encrypted dataset using LSTM, it indicates a noticeable, but controlled increase in error rates compared to the raw data condition.

The curve shifts upward and to the right, indicating an increase in both FNR and FPR, as expected, due to distortion of the feature space from encryption. The EER is at 0.030, reflecting a moderate reduction in discriminability. The model still exhibits acceptable error rates, demonstrating the LSTM can effectively work with encrypted representation, while providing operational usability. Table 9 shows the LSTM classifier demonstrates very low FAR, FRR, and EER values, with small 95% confidence intervals for both raw and encrypted data sets. This shows good numerical stability and high levels of resistance to replay, injection, and linkage attacks.

These results reinforce the usefulness of our continuous authentication process for securing VR environments.

Conclusion and Future Research Directions

This work presents a novel continuous authentication system for VR environments that uses head movement as a biometric feature, effectively addressing the limitations of traditional PIN and password-based mechanisms. Using LSTM networks, the proposed solution achieves an impressive accuracy of 99% in distinguishing between authorized and unauthorized users in the two raw datasets without encryption and around 97% using encrypted motion sensor data for user authentication, ensuring robust and reliable user authentication. To mitigate privacy concerns associated with outsourcing motion sensor data for training purposes, the system incorporates data encryption, safeguarding sensitive user information such as PINs and passwords.

In order to cover the limitations of deterministic tokenization with AES the future research direction includes:

- Adopting randomized encryption modes such as AES-CBC with initialization vectors,
- Applying differential privacy to add calibrated noise while balancing accuracy and privacy, and

- Exploring homomorphic encryption to enable secure computation on encrypted sensor data.

Future versions of the system will be designed and bench-marked with holistic security protections to mitigate the threats of replay, injection, and linkage in a practical setting. In particular, we intend to adopt challenge–response authentication schemes, nonce or per-record IV creation, verification of sequence-number or timestamp, and rate limiting on servers for replay and injection resilience. Furthermore, on the privacy-preserving side, we intend to explore salted hashing, randomized encryption, and linkage-resilient identifiers for preventing cross-session correlation. As soon as the hardware or communication module is present, these protections will be integrated into the end-to-end pipeline and evaluated using adversarial testing and security benchmarks. Only then will a fully secure prototype that is resistant to practical attacks be developed.

The seamless integration with existing VR hardware without requiring OS-level modifications underscores the practicality and scalability of the solution. Future research can focus on exploring lightweight encryption methods to further improve performance and investigating additional biometric features to strengthen authentication mechanisms incomplex VR scenarios.

References

1 VR Society. Head-mounted displays; 2025 [Accessed 2 February 2025]. Available from: https://www.vrs.org.uk/virtual-reality-gear/head-mounted-displays/

2 GV Research. Virtual reality (VR) market analysis; 2025 [Accessed 2 February 2025]. Available from: https://www.grandviewresearch.com/industry-analysis/virtual-reality-vr-market

3 Mathis F, Williamson JH, Vaniea K, Khamis M. Fast and secure authentication in virtual reality using coordinated 3D manipulation and pointing. ACM Trans Comput Human Interact. 2021;28(6):58. https://doi.org/10.1145/3460221

4 Li L, Chen C, Pan L, Zhang LY, Zhang J, Xiang Y. SigA: rPPG-based authentication for virtual reality head-mounted display. In: Proceedings of the 26th international symposium on research in attacks, intrusions and defenses (RAID '23); 2023. p. 686–99. https://doi.org/10.1145/3607199.3607209

5 Sivasamy M, Sastry VN, Gopalan NP. VRCAuth: continuous authentication of users in virtual reality environment using head-movement. In: 2020 5th international conference on communication and electronics systems (ICCES); 2020. p. 518–523. https://doi.org/10.1109CCES48766.2020.9137914

6 Wazir W, Khattak HA, Almogren A, Khan MA, Din IU. Doodle-based authentication technique using augmented reality. IEEE Access. 2020;8:4022–34. https://doi.org/10.1109/ACCESS.2019.2963543

7 Bhalla A, Sluganovic I, Krawiecka K, Martinovic I. MoveAR: continuous biometric authentication for augmented reality headsets. In: Proceedings of the 7th ACM cyber-physical system security workshop (CPSS); 2021. p. 41–52. https:// ://doi.org/10.1145/3457339.3457983

8 Manimaran S, Uma Priya D. PPSSDHE: privacy preservation in smartphone sensors data using ElGamal homomorphic encryption. Int J Sensor Net. 2024;46(4):218–29. https://doi.org/10.1504/IJSNET.2024.142718

9 Lalithamani N, Sujitha R. A survey process of attacks on iris based recognition for biometric authentication. Int J Comput Appl. 2015;10:1466–70.

10 Sujitha R, Lalithamani N. Counter measures for indirect attack for iris based biometric authentication. Int J Adv Res Comput Commun Eng. 2016;9:1–7. https://doi.org/10.17485/ijst/2016/v9i19/93868

11 Ragesh N, Ranjith R, Sivraj P. Fast R-CNN based masked face recognition for access control system. In: 2022 4th international conference on inventive research in computing applications (ICIRCA); 2022. p. 1049–55. https://doi.org/10.1109/ICIRCA54935.2022.9952298

12 Amritha VS, Aravinth J. Matcher performance-based score level fusion schemes for multi-modal biometric authentication system. In: 2020 6th international conference on advanced computing and communication systems (ICACCS); 2020. p. 79–85. https://doi.org/10.1109/ICACCS48705.2020.9074446

13 Zhu H, Jin W, Xiao M, Murali S, Li M. Blinkey: a two-factor user authentication method for virtual reality devices. Proc ACM Interact Mobile Wearable Ubiquitous Tech. 2020;4(4):1–29. https://doi.org/10.1145/3432217

14 Kim H, Kim C, Kim C, Kwak H, Im C-H. New user authentication method based on eye-writing patterns identified from electrooculography for virtual reality applications. Biomed Eng Lett. 2025;15:95–104. https://doi.org/10.1007/s13534-024-00426-8

15 Olade I, Fleming C, Liang H-N. BioMove: biometric user identification from human kinesiological movements for virtual reality systems. Sensors. 2020;20(10):2944. https://doi.org/10.3390/s20102944

16 George C, Khamis M, von Zezschwitz E, Burger M, Schmidt H, Alt F, Hussmann H. Seamless and secure VR: adapting and evaluating established authentication systems for virtual reality. In: Network and distributed system security symposium (NDSS 2017); 2017. p. 3309–21. https://doi.org/10.14722/usec.2017.23028

17 Hu C. VR technology-based audio and video system design and face authentication assisted multi-modal data model. In: 2022 international conference on electronics and renewable systems (ICEARS); 2022. p. 904–7. https://doi.org/10.1109/ICEARS53579.2022.9752157

18 Miller R, Banerjee NK, Banerjee S. Using siamese neural networks to perform cross-system behavioral authentication in virtual reality. In: 2021 IEEE virtual reality and 3D user interfaces (VR); 2021. p. 140–9.

19 Lohr D, Proulx MJ, Komogortsev O. Establishing a baseline for gaze-driven authentication performance in VR: a breadth-first investigation on a very large dataset. arXiv. 2024. Available from: https://arxiv.org/abs/2404.11798

20 Liebers J, Schneegass S. Gaze-based authentication in virtual reality. In: ETRA '20 adjunct: ACM symposium on eye tracking research and applications; 2020. p. 1–8. https://doi.org/10.1145/3379157.3391421

21 Zhang T, Ji Q, Ye Z, Akanda MMRR, Mahdad AT, Shi C, et al. SAFARI: speech-associated facial authentication for AR/VR settings via robust vibration signatures. In: CCS '24: ACM SIGSAC conference on computer and communications security; 2024. pp. 153–67. https://doi.org/10.1145/3658644.3670358

22 Bhalla A, Sluganovic I, Krawiecka K, Martinovic I. MoveAR: continuous biometric authentication for augmented reality headsets. In: Proceedings of the 7th ACM cyber-physical system security workshop (CPSS); 2021. p. 41–52. https://doi.org/10.1145/3457339.345798

23 Faye S, Jafarnejad S, Costamagna J, Castignani G, Engel T. Poster: characterizing driving behaviors through a car simulation platform. In: 2017 IEEE vehicular networking conference (VNC) (IEEE VNC 2017); 2017. p. 31–2. https://doi.org/10.1109/VNC.2017.8275650

24 Lo WC, Fan CL, Lee J, Huang CY, Chen KT, Hsu CH. 360° video viewing dataset in head-mounted virtual reality. In: Proceedings of the 8th ACM multimedia systems conference (MMSys); 2017. https://doi.org/10.1145/3083187.3083219

25 Yen SJ, Lee YS. Cluster-based sampling approaches to imbalanced data distributions. In: Tjoa AM, Trujillo J, editors. Data warehousing and knowledge discovery. Berlin: Springer; 2006. p. 427–36. https://doi.org/10.1007/11823728

26 Abuhamad M, Abuhmed T, Mohaisen D, Nyang D. AutoSen: deep-learning-based implicit continuous authentication using smartphone sensors. IEEE Int Things J. 2020;7(6):5008–20. https://doi.org/10.1109/JIOT.2020.2975779

27 Manimaran S, Sastry V, Gopalan N. SBTDDL: a novel framework for sensor-based threats detection on Android smartphones using deep learning. Comput Secur. 2022;118:102729. https://doi.org/10.1016/j.cose.2022.102729