ACTAUTH: Interaction-Triggered, Continuous User Authentication for Controller-Free Virtual Reality

Haoran Li*, Omar Achkar*, George Zouridakis†‡, Xin Fu‡, and Kyu In Lee*
*Department of Information Science Technology, †Department of Engineering Technology,
‡Department of Electrical and Computer Engineering,
Cullen College of Engineering, University of Houston, Houston, Texas 77204
Email: {hli80, oachkar, gzourida, xfu8, klee48}@cougarnet.uh.edu

2026 IEEE International Conference on Cyber Security and Resilience (CSR) | 979-8-3195-0774-7/26/$31.00 ©2026 IEEE | DOI: 10.1109/CSR69374.2026.11759823

***Abstract*—Virtual Reality (VR) headsets are frequently shared yet conventional authentication only protects initial login, leaving sessions vulnerable to takeover. To address this challenge, we present ACTAUTH, a controller-free continuous user authentication framework that authenticates users from natural hand–object interactions in VR. ACTAUTH uses grab-and-release events to define semantically coherent decision units, operating entirely on headset-native hand tracking without controllers or external sensors. In a proof-of-concept evaluation on a Meta Quest 3S with six users across four manipulation tasks, ACTAUTH achieves a mean true acceptance rate of 94.32% at a false acceptance rate of 2%, with an equal error rate of 0.09%. For takeover detection, ACTAUTH detects 100% of simulated takeovers within a single imposter interaction across all tasks. These results show that ACTAUTH's interaction-centric design achieves strong authentication performance at low false accept rates, providing a practical and deployment-oriented continuous authentication process for shared VR headsets in household-scale settings. Our data and source code are available through public online repository.**

***Index Terms*—VR security, Continuous user authentication, Behavioral authentication**

I. INTRODUCTION

Immersive computing continues to expand rapidly as Virtual and Extended Reality (VR/XR) headsets are deployed in a growing range of environments, including households, research laboratories, classrooms, and clinical settings. Unlike conventional personal devices designed for single-user ownership, VR headsets are frequently shared between users and often left unattended between sessions. In these shared-use contexts, the most probable and practical security risk is not an external adversary defeating the login process but *session takeover*, where another individual wears the headset and resumes activity within an already authenticated session. Conventional authentication methods (i.e., passwords, PINs, or companion device unlocks) verify identity only at initial login and provide no protection when device control changes mid-session. This gap poses fundamental security and privacy risks for VR applications that involve personal data, user accounts, financial transactions, or safety-critical simulations. These limitations motivate the need for *continuous authentication*, an approach that continuously verifies the user's identity during active use, allowing timely detection of takeovers and minimizing data exposure or unauthorized actions.

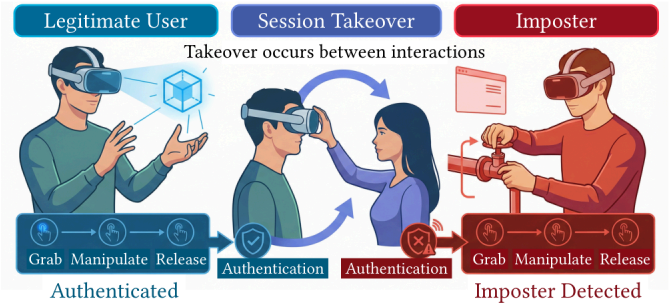


Fig. 1: Continuous authentication on a shared VR headset using hand tracking to detect session takeover after initial login.

Continuous authentication in VR environments has begun to explore modalities such as eye-gaze behavior [1], [2] and subtle head tremors captured by inertial sensors [3]. More broadly, VR research has shown that body movement and interaction behavior can also serve as strong user identifiers, including through controller telemetry, task-driven motion patterns, and longer interactive activity streams [4], [5], [6], [7], [8]. While these results collectively highlight the biometric richness of immersive behavioral signals, they require users to explicitly perform specific motions or depend on sensing channels or hardware (i.e., controllers) that aren't consistently available across devices. This dependency is increasingly misaligned with the current trajectory of the VR/XR ecosystem, which favors lighter, standalone systems that emphasize natural, controller-free hand interaction.

To address this need, we present ACTAUTH, *an action-based*, continuous user authentication framework triggered by hand-object interaction in VR. Object manipulation is a fundamental, ubiquitous form of interaction in immersive environments, where users routinely grab, adjust, and release virtual objects as part of normal use. This natural behavior provides an ideal foundation for opportunistic authentication, where identity is inferred implicitly from how users manipulate virtual content without interrupting their task.

The ACTAUTH approach operates entirely on hand-tracking data, leveraging an input modality that is now natively supported across major commercial VR platforms, including Apple Vision Pro, Meta Quest, and Varjo systems with Ultraleap

Authorized licensed use limited to: University of Houston. Downloaded on November 10,2026 at 19:31:14 UTC from IEEE Xplore. Restrictions apply.

tracking. Beyond widespread availability, hand–object interactions provide structured behavioral signals where actions typically follow sequential phases such as approach, grasp, adjustment, and release. This structured decomposition constrains motion variability while preserving user-specific patterns in timing, coordination, and micro-corrections, allowing the system to extract discriminative cues over short, meaningful segments. This property supports frequent re-authentication and efficient detection during normal usage, without introducing artificial prompts or idle intervals.

A central design question in continuous authentication concerns *when* and *how* to issue authentication decisions. Most prior work segments motion continuously using fixed-length sliding windows [2], [7], [9], evaluating each window independently. This approach introduces two issues: (1) long sessions accumulate numerous overlapping windows, so even a low false acceptance rate can compound over time, weakening long-term security; and (2) arbitrary windowing can yield incoherent feature sets that mix distinct behaviors, such as half of a grasp with idle motion, reducing reliability. To overcome these limitations, ACTAUTH ties authentication decisions directly to interaction semantics. Instead of evaluating fixed-interval windows, we treat each complete object manipulation, from grab to release, as a single decision unit. This event-aligned approach issues authentication only during meaningful activity, so the decision frequency scales naturally with user interactions rather than arbitrary time steps, and each evaluation corresponds to a coherent, contextually bounded action that improves interpretability and robustness.

In summary, this work makes the following three contributions:

- **End-to-end user authentication pipeline**: We design and present ACTAUTH, a controller-free continuous user-authentication framework for VR, covering enrollment, interaction segmentation, data conditioning, and online authentication during hand-object interactions.
- **Interaction-aligned segmentation method**: We introduce a new segmentation strategy that uses grab and release events to define interaction units and analyze its performance using multiple machine learning techniques.
- **Session takeover-focused evaluation**: We propose a takeover detection metric, detection within k interactions, and use it alongside standard authentication metrics to evaluate both security and detection speed.

II. Related work

Prior work has shown that behavioral biometrics can identify users in immersive environments from motion and interaction dynamics. Liebers *et al.* [6], [10] and Pfeuffer *et al.* [5] studied VR behavioral identification using task-driven motion, hand-tracking behavior, and body normalization, showing that user-specific movement patterns are distinctive and can be strengthened through appropriate normalization. These studies provide the foundation for using embodied interaction behavior as an authentication signal in VR.

Work on VR biometrics and interaction behavior has emphasized task context and generalization. Goal-oriented interaction studies show that behavior is shaped by task context [11], while *Who is Alyx?* highlights evaluation beyond controlled tasks using a multi-session XR dataset [7]. Cross-system studies further show that device and sensing differences can affect recognition [12], [13]. Together, these works motivate evaluation across task and device variation.

Other work explored alternative sensing modalities. Eye-tracking methods use gaze behavior for authentication [1], [2] and interaction [14], [15], while *VibHead* [16] and *HT-Auth* [3] rely on headset-centric vibration or head-tremor signals. These approaches demonstrate the range of XR sensing options, but they are less directly tied to object manipulation. In contrast, ACTAUTH uses headset-native hand-tracking segments triggered by object-interaction events, aligning authentication with the timing and structure of object-centric VR interaction.

III. System Models

This section describes the assumed usage setting for ACTAUTH and defines the corresponding system and threat models.

System model: We consider a consumer VR headset shared among multiple users, such as in households or classrooms. In this setting, several registered users access the same device for everyday VR activities, including interacting with virtual objects such as picking up, moving, placing, or using tools. To be authenticated, each user must first complete a one-time enrollment session on the device, establishing their behavioral baseline for subsequent authentication. The enrolled population is small and household-scale, typically ranging from a few users to at most a dozen. This small closed-group assumption is a deliberate design constraint as ACTAUTH targets shared-device scenarios where the set of enrolled users is known and stable. Scalability to larger populations is left for future work.

Threat model: We assume an adversary who has physical access to the shared headset. The adversary's goal is to use the device while it is already logged in as a legitimate user, including taking over an ongoing session after the legitimate user has started interacting, as illustrated in Fig. 1. The attacker can observe the tasks and objects presented in the application and may attempt to mimic the legitimate user's behavior through normal hand interactions. However, the attacker cannot inject, replay, or modify the underlying sensor streams, and we do not consider OS- or firmware-level compromise or remote attackers without physical device access. Under this model, the objective of ACTAUTH is to maintain a low false acceptance rate, correctly authenticate the legitimate user, and quickly detect session takeovers.

IV. Proposed Authentication Protocol

Fig. 2 illustrates our controller-free authentication protocol, which uses complete hand–object interaction segments (grab→release) as the basic decision units in VR. This section describes the full protocol, spanning user enrollment, interaction segmentation, data conditioning, and authentication.

Authorized licensed use limited to: University of Houston. Downloaded on November 10,2026 at 19:31:14 UTC from IEEE Xplore. Restrictions apply.

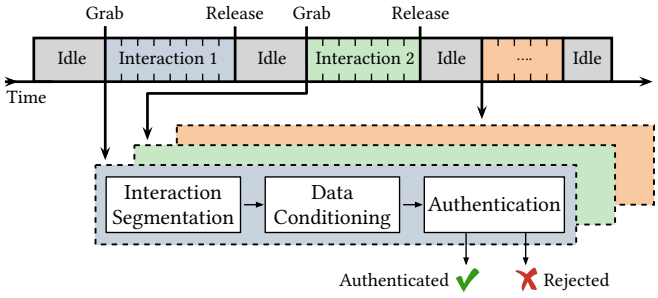


Fig. 2: ACTAUTH protocol pipeline, including interaction segmentation from grab and release events through data conditioning to per-interaction authentication decisions.

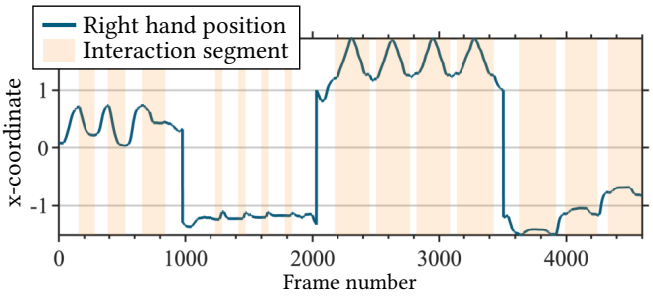


Fig. 3: Example session timeline of right-hand position with shaded regions indicating detected interaction segments used as inputs to authentication.

A. User Enrollment

Enrollment is a one-time session in which each user performs multiple trials of everyday object-manipulation tasks, such as picking up an object, moving it to a target, placing it on a surface, or using it as a tool, to capture natural variability in interaction style. During this process, the system records headset-native hand-tracking signals, including hand and finger kinematics and runtime interaction events (detailed in Section IV-B). Using each grab–release event pair, we extract the corresponding interaction segment and label it with the user identity to form the enrollment dataset. These segmented samples are then used to train authentication models, including traditional machine-learning classifiers and sequence-based neural network models, for authentication during device use. During training, the operating threshold τ is determined from a session-disjoint train/validation/test split with target ratios of 60%/20%/20%, where all interaction segments from the same session are assigned to the same partition.

B. Interaction Segmentation

Rather than evaluating identity over arbitrary fixed-length windows drawn from a continuous motion stream, ACTAUTH segments sessions at application-level interaction boundaries. Interactive VR applications naturally expose runtime events, especially object grab and release actions that provide semantically coherent, deployment-ready opportunities for re-authentication. This formulation aligns with realistic takeover behavior in object-centric VR, where control is more likely to change between interactions than within a single coherent action. We therefore define the authentication decision unit as an *interaction segment*, S_i, bounded by runtime interaction events. Let

$$S_i = \{x_n\}_{n=n_i^{\text{start}}}^{n_i^{\text{end}}},$$

where x_n denotes the n-th hand-tracking observation in the sequential runtime log, and n_i^{start} and n_i^{end} denote the boundaries of the i-th interaction segment (e.g., grab and release events). A user session is represented as an ordered sequence of interaction segments,

$$\{S_1, S_2, \ldots\}.$$

For each segment, the system constructs the corresponding segment buffer by collecting the hand-tracking observations between the start and end markers. The hand-tracking observations include x/y/z-coordinates of hand pose, fingertip and knuckle positions, and finger curl in degrees. Fig. 3 illustrates the session timeline of the right-hand position, where shaded regions indicate the flagged interaction segments used for authentication. These interaction boundaries restrict authentication to brief, task-relevant periods rather than the full continuous stream, leading to more efficient data usage and fewer unnecessary decisions.

C. Data Conditioning

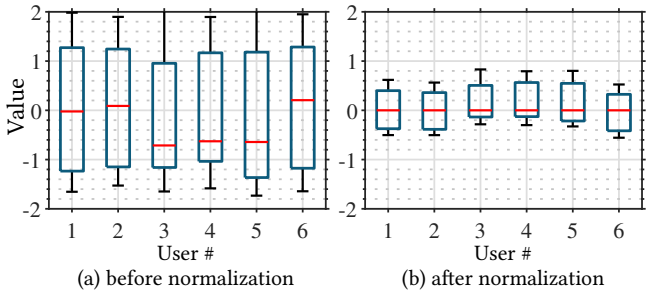
Raw signals within each interaction segment exhibit user/session-dependent offsets and scales (e.g., pose and tracking calibration differences) and may include occasional spikes from hand tracking. We condition each interaction segment through normalization, light denoising, and length standardization to suppress user- and session-dependent offsets and isolate behaviorally relevant cues. For normalization, we apply a percentile-based scaling per feature channel: for channel c, we compute $p_5^{(c)}$ and $p_{95}^{(c)}$ on the training data and transform each sample $x^{(c)}$ as

$$\tilde{x}^{(c)} = \text{clip}\left(\frac{x^{(c)} - p_5^{(c)}}{p_{95}^{(c)} - p_5^{(c)} + \epsilon}, 0, 1\right),$$

which maps the central 90% mass to $[0, 1]$ while saturating outliers. This reduces sensitivity to spikes and suppresses offset/scale cues correlated with anthropometrics. Fig. 4 shows the effect of normalization on a position channel: before normalization, users exhibit clear offset/scale differences (Fig. 4(a)), while after normalization these cross-user magnitude cues are reduced (Fig. 4(b)).

After normalization, we apply lightweight temporal smoothing and convert variable-duration segments into fixed-size inputs for downstream models. We apply a short moving-average filter per channel to suppress frame-level jitter. For sequence models, each segment is converted to a fixed length L via uniform sampling along the segment. For classical machine learning baselines, we aggregate each channel over the segment into a fixed-dimensional feature vector using

Authorized licensed use limited to: University of Houston. Downloaded on November 10,2026 at 19:31:14 UTC from IEEE Xplore. Restrictions apply.



(a) before normalization (b) after normalization

Fig. 4: Effect of normalization on a position channel, illustrating how cross-user offset and scale differences are reduced after normalization.

summary statistics (e.g., mean/std/min/max) and simple first-difference dynamics.

D. Per-Interaction Authentication Output

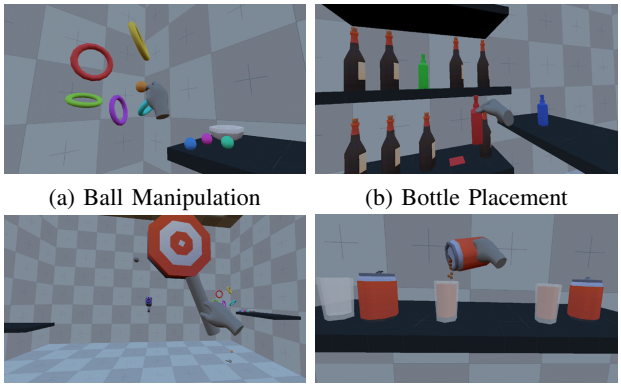
For each interaction segment S_i, the authentication module produces an authentication score s_i indicating consistency with the enrolled user. The interaction score s_i is compared with the user-specific threshold τ determined during enrollment to produce an authenticate/reject decision. The protocol therefore operates as a repeated cycle: (1) detect an interaction boundary, (2) collect and process the corresponding hand-tracking segment, and (3) issue one authentication output. The protocol is model-agnostic at the interface level. Different feature representations and classifiers may be used to compute s_i, but the protocol remains unchanged: one interaction segment enters the authentication module, and one interaction-level score with a corresponding binary decision is produced. This event-aligned design makes the timing of authentication decisions explicit at runtime and avoids dependence on arbitrary fixed-window scheduling, while aligning with the takeover threat where control changes are typically reflected between interactions.

V. Implementation and Evaluation

A. System Implementation and Enrollment

We implemented ActAuth on a Meta Quest 3S headset with native hand tracking. The experimental VR application was built in the Godot game engine and executed on a Windows desktop host, which streamed headset tracking data at a fixed 90 Hz. No handheld controllers, external motion-capture systems, or additional wearable sensors were used; all sensing relied solely on headset-native hand and finger tracking.

We recruited six right-handed participants under institutional IRB approval (STUDY00005734) to enroll in ActAuth for evaluation. Each participant completed 20 sessions involving four object-manipulation tasks: ball manipulation, bottle placement, projectile hitting, and soda pouring, as illustrated in Fig. 5. The dataset is recorded as frame-wise multivariate time-series data, with rows as time steps and columns as hand-tracking features. These objects were selected to capture representative handheld interactions, ensuring diversity in grasp types and motion dynamics. To maintain task consistency without promoting dependence on fixed trajectories, each task incorporated several location variants in which the same object appeared at different positions. Finally, we implemented the authentication portion of the pipeline using both classical and learning-based authentication models, including Random Forest (RF), One-Class SVM (OCSVM), Convolutional Neural Network (CNN), and Transformer, allowing direct comparison across model families under a unified sensing and experimental protocol. To support reproducibility, the full implementation details and experiment settings are available in our public GitHub [1] repository.



(a) Ball Manipulation (b) Bottle Placement

(c) Projectile Hitting (d) Soda Pouring

Fig. 5: Four object-manipulation tasks (ball manipulation, bottle placement, projectile hitting, and soda pouring) implemented in a Godot-based VR application.

B. Evaluation Metrics

We evaluate ActAuth both as an interaction-level authenticator and as a post-login takeover detector.

Interaction-Level Authentication: At the interaction level, each object manipulation constitutes one authentication attempt. For each enrolled user, we label that user's interaction segments as *legitimate* and all other users' segments as *adversary* (imposter). This closed-set evaluation represents a conservative lower bound on discriminability within the enrolled population and is the standard protocol for shared-device authentication studies of this scale. It does not cover open-set scenarios in which the adversary is an individual who has never enrolled. However, we discuss this distinction in Section VI. We report three standard metrics: *True Acceptance Rate (TAR)*, the fraction of legitimate interactions correctly accepted; *False Acceptance Rate (FAR)*, the fraction of adversary interactions incorrectly accepted; and *Equal Error Rate (EER)*, the error rate where false acceptance equals false rejection. For deployment-relevant operating points, we report $\mathrm{TAR@FAR} = 2\%$. Thresholds are selected on validation data to meet FAR targets, then applied unchanged to held-out test data. We also measure rejected genuine interactions per 100

[1] https://github.com/UHSSS/VR-Auth-In-Action

Authorized licensed use limited to: University of Houston. Downloaded on November 10,2026 at 19:31:14 UTC from IEEE Xplore. Restrictions apply.

TABLE I: Interaction-level authentication performance for each object task, reporting TAR at FAR $= 2\%$ with EER in parentheses for all evaluated models.

Model	Ball	Bottle	Projectile	Soda	Mean
RF	**97.05 (0.00)**	93.06 (0.27)	92.71 (0.10)	**94.44 (0.00)**	**94.32 (0.09)**
OCSVM	65.09 (7.08)	72.22 (5.20)	67.71 (8.85)	79.17 (7.35)	71.05 (7.12)
CNN	96.01 (0.10)	91.67 (0.00)	89.58 (0.10)	84.72 (0.00)	90.50 (0.05)
Transformer	94.11 (0.93)	**94.44 (0.00)**	**93.75 (1.15)**	90.28 (0.41)	93.15 (0.62)

interactions to quantify legitimate user interruptions during continuous use.

Takeover Detection: For post-login takeover evaluation, we simulate realistic attack scenarios by concatenating legitimate interaction sequences with adversarial sequences (imposter), where takeover occurs between interaction boundaries. We report *Time-to-Detect (TTD)*, the number of imposter interactions after takeover until the first reject decision, and detection within k interactions, the fraction of takeovers detected within the first k adversarial interactions ($k = 1, 3, 5$). These metrics quantify how quickly the interaction-triggered framework responds once control changes hands.

C. Authentication Performance

We first analyze interaction-level authentication performance when each interaction segment produces one decision. Table I summarizes TAR at FAR = 2% and EER for four model families, evaluated separately on the Ball, Bottle, Projectile, and Soda tasks and their mean. Across object families, supervised models (RF, CNN, Transformer) achieve high TAR, indicating that hand-object interaction segments carry strong user-specific cues. RF achieves the highest mean TAR (94.32%) with very low EER (0.09%), demonstrating that relatively simple non-sequential models can exploit summary statistics of the normalized interaction features. The Transformer model yields comparable mean TAR (93.15%) with slightly higher EER (0.62%), suggesting that sequence modeling of frame-wise dynamics can match or approach tree-based performance while operating directly on the temporal structure. CNN performance (mean TAR 90.50%, EER 0.05%) is somewhat lower than RF and Transformer but still maintains high acceptance of genuine interactions at the 2% FAR operating point. In contrast, the one-class SVM baseline shows substantially lower TAR (mean 71.05%) and higher EER (7.12%) across tasks. This gap reinforces the benefit of closed-set supervised training in small shared-headset populations, where labeled imposter data from other enrolled users is available and can be leveraged to shape tight decision boundaries.

Fig. 6 illustrates receiver operating characteristic (ROC) curves for each object family, plotting TAR versus FAR for all four model classes. For Ball and Bottle tasks, RF and Transformer curves remain near the upper-left corner over a wide FAR range, indicating robust discrimination between enrolled and imposter interactions. Projectile and Soda tasks show slightly more overlap between models, but supervised methods still dominate the OCSVM baseline, highlighting that

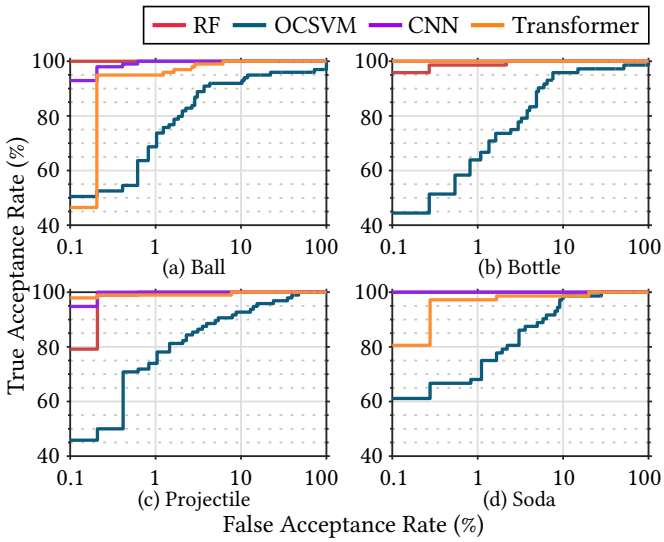


Fig. 6: ROC curves for all models on the (a) Ball, (b) Bottle, (c) Projectile, and (d) Soda tasks, plotting TAR vs. FAR.

TABLE II: Number of rejected genuine interactions per 100 interactions at FAR $= 2\%$ for each task and model.

Model	Ball	Bottle	Projectile	Soda
RF	**2.95**	6.94	7.29	**5.56**
OCSVM	34.91	27.78	32.29	20.83
CNN	3.99	8.33	10.42	15.28
Transformer	5.89	**5.56**	**6.25**	9.72

interaction-triggered segmentation yields consistent verification performance across diverse manipulation styles.

To assess the stability of continuous use, Table II reports the mean number of rejected genuine interactions per 100 interactions at the FAR = 2% threshold. For RF, rejection rates remain between 2.95 and 7.29 per 100 interactions across tasks, while CNN and Transformer exhibit similar low rejection frequencies, especially for Ball and Bottle. These values correspond to infrequent interruptions in a sustained session, suggesting that an application could tolerate such occasional rejections by prompting lightweight re-authentication only when necessary. In contrast, OCSVM rejects roughly 20–35 genuine interactions per 100, which would cause frequent disruptions and is therefore less suitable for continuous deployment at the same security operating point.

Overall, these results indicate that interaction-aligned segmentation, combined with supervised models, can provide high interaction-level accuracy at low FAR while keeping legitimate user interruptions limited during normal behavior. The strong performance across four object families also supports the generality of the interaction-triggered formulation beyond a single task.

D. Takeover Detection

We next evaluate how quickly ACTAUTH can detect session takeover when a different user takes control between interactions. We focus on the best-performing supervised model, RF,

Authorized licensed use limited to: University of Houston. Downloaded on November 10,2026 at 19:31:14 UTC from IEEE Xplore. Restrictions apply.

TABLE III: Session takeover detection latency for the best-performing RF model, including median time-to-detect and detection rates within 1, 3, and 5 imposter interactions.

Task	Med. TTD	$\mathbf{TTD \leq 1}$	$\mathbf{TTD \leq 3}$	$\mathbf{TTD \leq 5}$
Ball	1	100.0%	100.0%	100.0%
Bottle	1	100.0%	100.0%	100.0%
Projectile	1	100.0%	100.0%	100.0%
Soda	1	100.0%	100.0%	100.0%

under the FAR = 2% operating point and evaluate *interaction-count latency* on takeover sequences that concatenate a genuine prefix with an imposter suffix at an interaction boundary. Table III summarizes takeover detection latency, measured in number of imposter interactions across four tasks. For all four object families, the median TTD is exactly one interaction, meaning that half of the simulated takeovers are detected after the very first imposter manipulation following takeover. Moreover, the fraction of takeovers detected within one interaction ($k = 1$) is 100% for every task, and detection within three and five interactions ($k = 3, 5$) are also 100%, indicating that RF-based authentication almost always rejects imposter control at or immediately after the first adversarial interaction segment. These results highlight a key advantage of the interaction-triggered protocol: by tying decisions directly to grab and release events, the system obtains dense, semantically meaningful opportunities to re-evaluate identity and does not require majority voting over long windows or multiple segments. In a shared-headset setting, this translates into rapid containment of session takeover with minimal delay in interaction count between control change, limiting the adversary's opportunity to access sensitive content or perform unauthorized actions.

VI. Discussion

We discuss several challenges and future directions for ActAuth.

Security–usability trade-off: Threshold selection creates a direct trade-off between reducing false accepts and avoiding unnecessary rejections of genuine users. Stricter thresholds improve security but may increase friction during sustained use, while looser thresholds improve usability at higher risk. Future work should study application-level responses to rejections, such as gating sensitive actions, requesting re-authentication, or flagging suspicious behavior.

Generalization across tasks and environments: Although our evaluation covers multiple task families, broader deployment may introduce variation from lighting, occlusion, camera viewpoint, headset tracking pipelines, physical space, and behavioral drift over time. Future evaluations should consider longer time gaps, different environments and headsets, unseen interaction styles, and the amount of enrollment data needed to maintain stable thresholds.

Closed-set supervision and open-set deployment: Our current evaluation draws adversaries from the enrolled population, which reflects household-scale use cases where users are known. Less controlled deployments require rejecting unseen individuals, where closed-set classifiers may lack calibrated thresholds. The OCSVM baseline provides an initial one-class comparison, but its higher EER suggests that robust open-set authentication remains an important future direction.

VII. Conclusion and Future Work

We presented ActAuth, a continuous authentication framework that aligns verification decisions with hand–object interaction boundaries in controller-free VR. Evaluated across four tasks with six users, the system achieves a mean EER as low as 0.09% (RF), with 94.32% true acceptance at false acceptance rate of 2%. Session takeovers are detected within a single imposter interaction at a 100% rate, while legitimate user rejections remain between 3–7 per 100 interactions for the best models. These results confirm that ActAuth provides a reliable and unobtrusive foundation for continuous authentication in shared VR environments. Future work will expand the participant cohort to better assess robustness and generalizability across a broader user population. We will also explore generalization across unseen tasks and interaction styles to better understand how well the framework transfers beyond the specific activities studied here.

Acknowledgments

All study procedures were reviewed and approved by the University of Houston Institutional Review Board under IRB ID STUDY00005734. This work was supported by National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2024-00463802).

References

[1] Y. Zhang *et al.*, "Continuous Authentication Using Eye Movement Response of Implicit Visual Stimuli," *ACM IMWUT*, 2018.
[2] J. Liebers *et al.*, "Using Gaze Behavior and Head Orientation for Implicit Identification in Virtual Reality," in *ACM VRST '21*.
[3] Z. He *et al.*, "HT-Auth: Secure VR Headset Authentication via Subtle Head Tremors," *ACM IMWUT*, 2025.
[4] J. Liebers *et al.*, "Kinetic Signatures: A Systematic Investigation of Movement-Based User Identification in Virtual Reality," in *ACM CHI '24*.
[5] K. Pfeuffer *et al.*, "Behavioural Biometrics in VR: Identifying People from Body Motion and Relations in Virtual Reality," in *ACM CHI '19*.
[6] J. Liebers *et al.*, "Understanding User Identification in Virtual Reality Through Behavioral Biometrics and the Effect of Body Normalization," in *ACM CHI '21*.
[7] C. Rack *et al.*, "Who is Alyx? A new behavioral biometric dataset for user identification in XR," *Frontiers in Virtual Reality*, 2023.
[8] R. Miller *et al.*, "Realtime Behavior-Based Continual Authentication of Users in Virtual Reality Environments," in *IEEE AIVR*, 2019.
[9] M. Li *et al.*, "Motion Forecasting Attacks on Behavioral Biometric Authentication Systems in Virtual Reality," in *ACM VRST '25*.
[10] J. Liebers *et al.*, "Identifying Users by Their Hand Tracking Data in Augmented and Virtual Reality," *Int. J. Hum.-Comput. Interact.*, 2024.
[11] S. Prange *et al.*, "Go for GOLD: Investigating User Behaviour in Goal-Oriented Tasks," ser. CHI EA '19.
[12] R. Miller *et al.*, "Within-System and Cross-System Behavior-Based Biometric Authentication in Virtual Reality," in *IEEE VRW '20*.
[13] R. Miller *et al.*, "Using Siamese Neural Networks to Perform Cross-System Behavioral Authentication in Virtual Reality," in *IEEE VR '21*.
[14] K. Pfeuffer *et al.*, "Empirical Evaluation of Gaze-enhanced Menus in Virtual Reality," in *ACM VRST '20*.
[15] K. Pfeuffer *et al.*, "Gaze + pinch interaction in virtual reality," ser. SUI '17.
[16] F. Li *et al.*, "VibHead: An Authentication Scheme for Smart Headsets through Vibration," *ACM Trans. Sensor Networks*, 2024.

Authorized licensed use limited to: University of Houston. Downloaded on November 10,2026 at 19:31:14 UTC from IEEE Xplore. Restrictions apply.