# Raise2Auth: A Dual-Factor, Adaptive Gesture Authentication System for Enhanced Mobile Security

ANONYMOUS AUTHOR(S)

In the fast-evolving realm of mobile computing, securing authentication methods against an array of vulnerabilities is crucial. Current behavior-biometric systems on smartphones face challenges with robustness and often require ongoing user interaction to extract features, lacking the one-step convenience of fingerprint or facial recognition. To address these issues, we introduce Raise2Auth, an advanced dual-factor, gesture-based mobile authentication system that offers a seamless and more robust solution. Raise2Auth leverages the intrinsic way users interact with their devices, specifically the combined actions of grasping and lifting, to authenticate identity. By analyzing the intricate hand grip patterns and motion trajectories through the smartphone's built-in Inertial Measurement Unit (IMU), Raise2Auth provides a nuanced layer of security less susceptible to spoofing than traditional physiological or behavioral traits. Our comprehensive experimentation with 25 participants revealed the system's resilience, achieving an impressive verification accuracy rate exceeding 97.6% while maintaining a False Acceptance Rate (FAR) of less than 0.1% against both imitation and synthetic attacks. This dual-factor approach, integrating gestural and postural biometrics with IMU data, not only aligns with natural user behavior but also significantly advances the field of authentication technology, bridging the gap between convenience and robust security.

CCS Concepts: • **Human-centered computing → Human computer interaction (HCI)**; *Ubiquitous and mobile computing systems and tools.*

Additional Key Words and Phrases: User authentication, smartphone, behavior biometric

## 1 INTRODUCTION

In this digital age, characterized by the exponential growth of the Internet of Things (IoT), cloud computing, and digital financial services, the significance of user authentication has soared to unprecedented levels. The ubiquity of IoT devices has led to a complex network where secure access to information is paramount. Wearable devices, integrating seamlessly into daily life, have become instrumental in this landscape. This includes a diverse array of gadgets like earphones[7, 18], smartwatches[46], bracelets[16], and rings[42], each offering novel means of user verification. Despite this variety, smartphones continue to dominate as the primary tool for authentication across different platforms and IoT devices. Their widespread adoption and advanced technological capabilities make them the linchpin in the authentication process, linking various devices and platforms in a secure and cohesive manner. This paradigm shift underscores the need for innovative authentication methods that are both secure and user-friendly.

Smartphone-based authentication methods have evolved into three dominant categories: password-based[30], biometric[19, 47], and behavior-based[3, 20]. Password-based systems, while user-friendly and familiar, suffer from significant drawbacks. The tendency of users to select common or easily guessable passwords, often linked to personal information, poses a substantial security risk, especially in public spaces where password entry may be observed. Biometric authentication methods, which include fingerprint scanning, facial recognition, and voiceprint identification, offer enhanced security but are not without their challenges. These methods can be vulnerable to sophisticated attacks and are often influenced by external environmental factors such as lighting conditions and ambient noise [35]. Additionally, the reliability of biometric systems can be compromised under certain conditions, like injuries or voice changes. Behavior-based authentication takes a different approach, focusing on the user's interaction with their device. This includes analyzing patterns in touchscreen gestures[48], typing rhythms[27], and other interaction-based data. However, these methods often require the user to perform specific and sometimes repetitive actions for authentication, potentially impacting the user experience and efficiency.

Confronting the shortcomings of these traditional mechanisms, we propose Raise2Auth, a covert and user-centric authentication system that offers a seamless alternative. Raise2Auth is a novel authentication solution that replaces noticeable verification actions with the natural and automatic movements we use with our smartphones. It leverages the unique way individuals interact with their smartphones, focusing on the natural and often subconscious actions performed while handling the device. By utilizing the smartphone's built-in Inertial Measurement Unit (IMU), Raise2Auth captures the specific nuances of how a user grips, lifts, or moves their phone. This approach not only enhances security by tapping into a user-specific behavioral pattern but also improves user convenience by eliminating the need for deliberate authentication actions. Raise2Auth's methodology is rooted in the understanding that each individual's interaction with their smartphone is distinct and can serve as a reliable identifier, thereby offering a novel and effective solution to the challenges of mobile authentication.

This paper is structured to comprehensively introduce and evaluate Raise2Auth. We begin by discussing the current state and limitations of existing authentication methods. Following this, we delve into the technical details of Raise2Auth, including its design principles and implementation. We then present a thorough analysis of our experimental results, demonstrating the effectiveness of Raise2Auth. Finally, we conclude with a discussion on the broader implications of our findings and potential future directions for research in the field of user authentication.

The contributions of this work mainly are three-folds:

- We introduce the Raise2Auth, a novel authentication framework that combines two distinct biometric descriptors: the motion trajectory of the hand-raising action and the grasp pattern of the hand-holding the device. This dual-factor approach enhances the system's accuracy and security, effectively leveraging the unique aspects of each gesture.
- We utilized 3D trajectory and rotation vectors to characterize the hand-raising gesture, combined with deep learning for feature extraction. This method provides a detailed and accurate representation of the gesture, enabling the system to identify unique user-specific patterns.
- We have carried out comprehensive performance evaluations of the proposed system, verifying its robustness against both imitation and synthetic attack scenario

## 2 RELATED WORK

This section describes research work related to user authentication systems, in the areas of mobile devices and behavior-based biometric solutions.

### 2.1 Mobile User Authentication

Recent advancements in smartphone-based user authentication have witnessed a shift towards exploiting the myriad of built-in sensors beyond traditional biometrics like fingerprints[25], facial recognition[22, 50], and voiceprints[15, 17]. Researchers are innovating to harness unique physiological and behavioral signatures captured through these sensors for more secure and nuanced user verification methods.

The touchscreen, frequently engaged during user interaction with their smartphone, serves as an ideal sensor for capturing user behavior patterns. Abuhamad *et al.* introduced AUToSen [2], a deep-learning-based active authentication approach utilizing consumer-grade smartphone sensors for user identification. Leveraging an LSTM model, AUToSen identifies distinct user behaviors through both active and passive interactions with the device. Similarly, Ali *et al.* developed a behavior-based authentication system that employs swipe movements for continuous user verification post-initial login [5]. Further innovating in this space, Song *et al.* integrated semantic handcrafted features with deep representations through a framework called Feature Regularization Net (FRN), addressing the gap in semantic consistency in previous studies [41].

Acoustic sensing has recently emerged as a rapidly evolving modality capable of detecting subtle object movements. Chen *et al.* proposed a multi-factor authentication system to bolster voice-based authentication, utilizing acoustic signals to monitor chest movements during speech, thereby capturing unique cardiovascular patterns [13]. In a similar vein, Swipepass offers a novel second-factor authentication solution for smartphones by capturing distinct physiological and behavioral characteristics during pattern lock inputs through acoustic signals [12].

Research on IMU sensors for capturing users' biometric traits [45] is also gaining traction. Choi *et al.*'s VibPath, a novel Two-Factor Authentication (2FA) system, enhances smartphone security by analyzing the user's hand neuromuscular system through touchscreen interactions and vibration path responses [14]. HoldPass utilizes hand-based ballistocardiography (Hand-BCG) to record unique heart activity via hand vibrations, presenting an innovative alignment-free authentication scheme that effectively overcomes the challenges of weak signals and motion artifacts, thereby demonstrating its feasibility and accuracy in authentication [24].

As the work most closely related to Raise2Auth, which utilizes hand pick-up movements for mobile behavior-biometric authentication, Lee *et al.* [28] developed a mobile authentication system capturing users' phone pick-up actions. This system employs a weighted multi-dimensional Dynamic Time Warping (DTW) algorithm to effectively measure the similarities between users' pick-up movements. Similarly, Boshoff *et al.* [9] introduced a gesture-based authentication method that leverages the bag-of-movement technique to recognize the motion of retrieving a phone from a pocket, achieving an accuracy of 93%. However, these studies did not delve into the deeper characteristics of the hand-raising gesture but instead directly extracted features from accelerometer and gyroscope data. In everyday scenarios, IMU sensors often encounter substantial noise, including both body motion and environmental factors, as well as noise intrinsic to the sensors themselves.

## 2.2 Behavior-based Authentication

The domain of behavior biometric authentication has undergone significant evolution, transitioning from reliance on static, physical biometric identifiers to dynamic, behavior-based metrics. This shift leverages the intricate and subtle patterns of user interaction with their devices, providing a more personalized and secure means of identity verification. Behavior biometrics delve into the analysis of human activity patterns, such as typing rhythms[4, 6], walking gaits[40, 51], and the unique manner in which a user handles or interacts with their smartphone[11, 38]. Distinct from traditional biometrics, which depend on unchangeable physical traits, behavior biometrics offer a dynamic security layer that adjusts to the user's evolving behaviors.

A pioneering effort in this realm was made by Shen *et al.* [37], who developed a user authentication system based on specific mouse-operation tasks. This system could authenticate users by capturing a detailed characterization of an individual's distinctive mouse behavior. In a similar vein, Li *et al.* [29] investigated continuous authentication by integrating keystroke and mouse dynamics with wrist motion behaviors, effectively bridging the security gap between mouse and keyboard inputs. They introduced a context-aware keystroke latency feature cell generation to address latency fluctuations across different words.

Gait recognition has also been a focal point of behavior biometric studies. Sun *et al.* [43] designed an accelerometer-based gait authentication system, incorporating a speed-adaptive gait cycle segmentation method and a personalized matching threshold generation method to accommodate variable walking speeds. GaitCode [33] represents a continuous authentication system that applies multimodal learning on data from both accelerometers and ground contact force sensors in wearable devices.

The application of deep learning techniques [31] has further refined the analysis and interpretation of complex behavior patterns. Jagmohan *et al.* introduced ContAuth [10], employing online learning models that bolster the robustness of behavior-based authentication systems. ContAuth is notable for its ability to continually adapt to new data (data incremental learning) and to incorporate new users without necessitating retraining (class incremental learning). Wang *et al.* [44] redefined the classification challenge in authentication through deep metric learning, enhancing discrimination capabilities and devising a strategy to deter side-channel attacks by integrating a noise signature into sensing signals, aiming to preserve usability. Additionally, Shi *et al.* [39] proposed an end-to-end deep-learning approach employing domain adaptation techniques to filter out environment- and location-specific information, thereby isolating behavioral biometrics for reliable, environment-independent user authentication.

## 3 THREAT MODEL AND FEASIBILITY STUDY

In this section, we first introduce the attack scenarios of existing behavior-based user verification systems. We further provide the background and the rationale behind the underlying uniqueness from the perspectives of physical behavior properties. We also perform a feasibility study as the proof-of-concept.

### 3.1 Attack Scenarios

Gripping and lifting the smartphone is the most common and natural initial interaction between a user and their device, reflecting a unique and subconscious aspect of how individuals engage with their smartphones. Each user exhibits a distinct manner of holding and manipulating their device, a pattern that is developed and reinforced over time, making it a potent source of behavioral biometrics. Existing mobile biometric authentication methods often suffer from various spoofing and replay attacks [8, 49]. Here we list two primary types of attacks for mobile authentication.

**Imitation Attacks:** In this type of attack, the adversary might attempt to mimic the user's specific way of handling the smartphone. This could involve observing and replicating the user's hand positioning and the typical trajectory they follow when lifting the phone. While this method of attack requires close observation and may be difficult to execute precisely, it's a potential vulnerability, especially if the attacker has prolonged and close access to the user.

**Physical Synthetic Attacks:** Advanced sensors and recording devices could be used to capture the user's smartphone interaction in detail. For example, a nearby camera could record the speed, angle, and precise movement pattern of the user while lifting and holding the phone. With such data, a programed robot arm could potentially recreate the motion, spoofing the system into granting access.

### 3.2 Attack-resistant User Authentication through Natural Interactions

*3.2.1 Behavior of Grasping and Lifting.* To create a simplified physical model of the behavior of gripping and lifting a smartphone, we can represent it as a series of connected rotational movements involving three rigid bodies. These rigid bodies correspond to the upper arm, the forearm, and the hand (including the smartphone). To simplify the model, we assume each rigid body is homogeneous and they are connected by idealized joints, namely the shoulder joint, the elbow joint, and the wrist joint. This approach allows us to focus on the rotational movements at each joint, simulating the action of lifting a smartphone as a mechanical system.

Model Components:

- Upper Arm Rigid Body: Connected to the shoulder, capable of rotating around the shoulder joint.
- Forearm Rigid Body: Connected at the elbow joint and linked to the upper arm rigid body, capable of rotating around the elbow joint.
- Hand (including Phone) Rigid Body: Connected at the wrist joint and linked to the forearm rigid body, capable of rotating around the wrist joint.

Rotational Joints are defined as:

- Shoulder Joint: Allows the upper arm to rotate around a point in the shoulder.
- Elbow Joint: Connects the upper arm and forearm, allowing the forearm to rotate relative to the upper arm.

- Wrist Joint: Connects the forearm and hand, allowing the hand (including the phone) to rotate relative to the forearm.

#### 3.2.2 *Characterizing Raise2Auth for User authentication.*

The motion of lifting the arm is a complex biomechanical process involving several joints (shoulder, elbow, and wrist) and muscle groups. This motion can be characterized by the unique way in which each individual executes it, influenced by their physiological attributes such as arm length, muscle strength, and joint flexibility.

*Defining the System's Kinetic and Potential Energy:* The kinetic energy $T$ consists of both linear and angular kinetic energies of the arm's segments. For each rigid body (upper arm, forearm, hand), the kinetic energy is expressed as:

$$\mathrm{T}_i = \frac{1}{2} m_i v_i^2 + \frac{1}{2} I_i \omega_i^2 \tag{1}$$

where $m_i$ is the mass, $v_i$ is the linear velocity, $I_i$ is the moment of inertia, and $\omega_i$ is the angular velocity.

The potential energy $V$ is primarily due to gravity and can be expressed as:

$$\mathrm{V}_i = m_i g h_i \tag{2}$$

where $g$ is the acceleration due to gravity, and $h_i$ is the height relative to a reference point.

*Considering Muscle-Generated Torque:* The torque $\tau$ at each joint depends on muscle forces and is influenced by individual muscle strength:

$$\tau_i = f(F_{\text{muscle}}, l_{\text{arm}}, \theta_i) \tag{3}$$

where $\tau_i$ is the torque at joint $i$, $F_{muscle}$ represents muscle forces, $l_{arm}$ is the length of the arm segment, and $\theta_i$ is the joint angle.

*Motion Profile using Lagrangian Equations:* To describe the dynamics of the entire arm, especially when considering the torque generated by muscles, we utilize the Lagrangian equations for complex dynamic systems, particularly when accounting for rotational movements and various types of forces. The Lagrangian $L$ is defined as the system's kinetic energy $T$ minus its potential energy $V$:

$$L = T - V \tag{4}$$

The Lagrangian equation is then formulated as:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}_i}\right) - \frac{\partial L}{\partial \theta_i} = f(F_{\text{muscle}}, l_{\text{arm}}, \theta_i) \tag{5}$$

where $\theta_i$ is the angle of the $i$th joint, and $\dot{\theta}_i$ is the angular velocity.

### 3.3 Pilot Feasibility Study

#### 3.3.1 *Proof-of-concept Setup.*

To investigate the feasibility of characterizing and grasping and lifting behavior for user authentication, we conducted a preliminary study (based on 5 subjects) to perform the picking up smartphone task, ten times per subject. We recorded IMU data using the self-designed app. All IMU data were collected with a 60 Hz sampling rate. As the preliminary study, we performed the experiments under a controlled environment while the participants remained standing still.

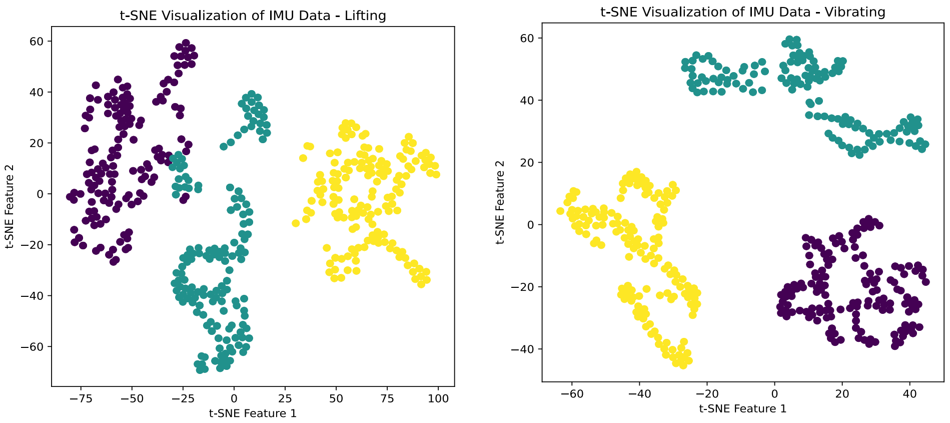


Fig. 1. A pilot study of the distinguishability of the lifting and vibrating modality

.

*3.3.2 Feasibility Analysis.* In our feasibility analysis, we focused on identifying the unique aspects of grasping and lifting behaviors for user authentication by leveraging variations in IMU readings caused by different gripping postures and the motion trajectories during the lifting process. Utilizing the smartphone's IMU, we explored how distinct vibration patterns associated with different grips and the trajectory data from lifting actions can be used for authentication purposes.

We experimented with various combinations of the two authentication modalities: grip and lift. These combinations included scenarios where vibration occurred during the lifting process and others where the device was first lifted and then vibrated. For each of these modalities, we analyzed IMU data collected from five individuals under two distinct modes: 'Vibration + Lift' and 'Lift then Vibrate'.

The IMU data from these modalities were subjected to an initial analysis using the T-SNE (t-Distributed Stochastic Neighbor Embedding) technique for visualization.

*3.3.3 Insights and Summary.* Based on our pilot study, our preliminary findings are as follows: (1) Both vibration and lifting motions exhibit distinct patterns that can differentiate between users, as evidenced in the unique IMU data signatures. However, (2) when vibration and lifting actions occur simultaneously, while they reduce the authentication time, they tend to interfere with each other. This interference, in turn, adversely affects the distinctiveness between different users' data. These findings suggest that while each modality individually contributes to user differentiation, their concurrent execution may not be optimal for maximizing user distinction in authentication processes.

To enhance the extraction of unique characteristics inherent in each authentication modality, we propose the integration of deep learning models. These models are designed to meticulously extract and analyze features from both the vibration and lifting data.

## 4 RAISE2AUTH DESIGN

### 4.1 System Overview

In this study, we propose the Raise2Auth, a behavior-based user authentication system, which utilizes a smartphone's IMU to characterize the grasping and lifting action. The end-to-end framework and processing flow are shown in Fig.2, which consists of four major components: raise gesture detection, dual factor feature representation, authentication model, and decision fusion.

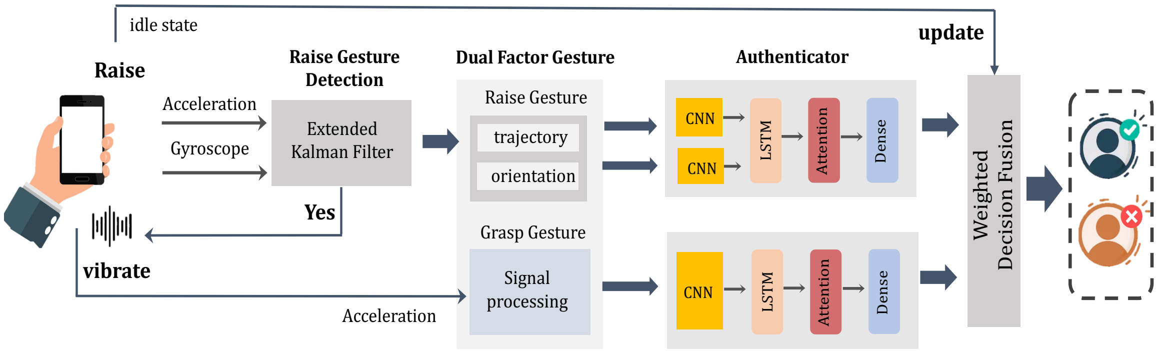


Fig. 2. System overview

***Raise Gesture Detection.*** This module is responsible for identifying the specific action of the user lifting their smartphone. Utilizing the smartphone's IMU, this component detects the initial gesture of grasping and raising the phone. It discerns this action from other similar movements by analyzing the data for specific patterns that indicate a genuine raise gesture, ensuring that the authentication process is triggered by the correct user action.

***Dual Factor Feature Extraction.*** Upon successful detection of the raise gesture, this module triggers a vibration signal, followed by the collection of IMU data to capture the characteristics of the grasping gesture. Concurrently, it generates 3D trajectory and orientation data, providing a comprehensive feature set that represents the nuances of the raising gesture. This dual approach in feature extraction enhances the accuracy of the system in capturing the intricacies of the user's interaction with the device.

***Authentication Modeling.*** The extracted features from both the raising and grasping gestures are then processed by two separate deep-learning-based authenticators. Each authenticator is specialized in analyzing its respective set of features, ensuring a detailed and focused evaluation of the user's unique gestural patterns for authentication.

***Decision Fusion.*** The decision fusion is responsible for integrating the outputs from the dual authenticators to arrive at a conclusive authentication decision. This module synthesizes the information gathered from both the raise and grasping gesture authenticators. Additionally, it employs an idle state to dynamically update the weights of each authenticator, allowing for continuous improvement and adaptation of the system based on ongoing user interactions.

## 5 RAISE GESTURE DETECTION

The event detection module is a critical component of our authentication system, serving as the foundation for accurately identifying and authenticating user actions. Its primary role is to discern specific movements – in our case, the action of lifting a smartphone. The main challenges addressed in this module are:

### 5.1 Defining the Raise Gesture

The lift gesture encompasses a series of movements starting from retrieving a smartphone from one's pocket to holding it up, rotating it for better viewing or interaction angle, and culminating in the unlocking of the device. This gesture sequence involves a combination of arm, wrist, and hand movements, with each phase having its unique characteristics.

*Initiation Phase - Retrieving from the Pocket.* The initiation phase of the gesture begins as the hand reaches into the pocket. This movement is characterized by the fingers grasping the phone, primarily involving the arm and wrist. The main motions in this phase include bending at the elbow and the flexion/extension of the wrist, providing the necessary degrees of freedom to retrieve the phone smoothly from the pocket.

*Transition Phase - Lifting and Orientation.* Following a secure grasp of the phone, the transition phase involves lifting the device from the pocket. This phase is characterized by the extension of the elbow and abduction of the shoulder, allowing the arm to bring the phone up to a more accessible position. The movements in this phase are critical for smoothly transitioning from retrieval to the point where the phone is ready for use.

*Adjustment Phase - Rotating for View/Interaction.* Once the phone is lifted, the adjustment phase begins, where the phone is often rotated to a horizontal position for optimal viewing or interaction. This phase is primarily characterized by wrist pronation/supination along with minor adjustments in the grip. These movements provide the necessary degrees of freedom for the user to comfortably orient the phone for better visibility or to initiate interaction.

In our observations involving 25 volunteers, we noticed that while the initiation phases were performed consistently, variations were observed in the transition phases and adjustment phases due to physiological uniqueness and personal preferences. Interestingly, we observed that the subtleties in wrist rotation, the angle at which the smartphone is lifted, and the speed of movement varied notably among individuals. These variations, coupled with the personalized interaction with the smartphone, such as hand grasping position, contributed to a unique gesture signature for each participant.

## 5.2 Basics and Challenges

*5.2.1 Can Raise Gesture to be differentiated from other gestures.* In the course of everyday smartphone usage, users engage in various actions, many of which entail alterations in the phone's position and orientation. Besides lifting the phone for authentication, these actions include rotating or flipping the phone, shaking it, arm movements while walking or running, and moving the phone to different locations. Each of these actions results in distinct changes in the phone's IMU (Inertial Measurement Unit) readings, reflecting the varying nature of the physical movements.

*5.2.2 Will raise gestures pattern be affected by body motions.* The intricacy of "Raise Gestures" in the context of smartphone interaction raises a pivotal question: are these gesture patterns influenced by broader body motions? This concern is particularly relevant in scenarios where the user is in motion, such as walking, running, or engaging in other activities that involve significant body movement.

Body motions, especially those involving the arms, torso, and head, can indeed impart variations to the IMU readings of a smartphone, thereby potentially influencing the distinctiveness of the "Raise Gesture." As shown in Fig. 3, it is evident that body motion significantly impacts the accelerometer and gyroscope readings during walking. When static, the accelerometer readings are relatively stable, with minimal fluctuations primarily due to minor postural adjustments. In contrast, the walking state introduces pronounced spikes and oscillations across all axes, reflecting the dynamic nature of human gait and its associated accelerations. Similarly, the gyroscope data, which is typically near-baseline in the static state, exhibits substantial variability during walking. This variation can be attributed to the rotational movements and changes in orientation that occur as part of the natural walking motion.

*5.2.3 Different users may raise their phones in various ways.* The variability of the raise gesture extends to the individual differences among users. The way users naturally lift their phones encompasses a spectrum of motion patterns. These patterns are influenced by a multitude of factors including individual ergonomic preferences and habitual movements. One user might lift their phone swiftly and directly but the other user exhibits a more leisurely and varied motion. For example, from Fig. 4, User 1's data might show a steady, controlled motion, User

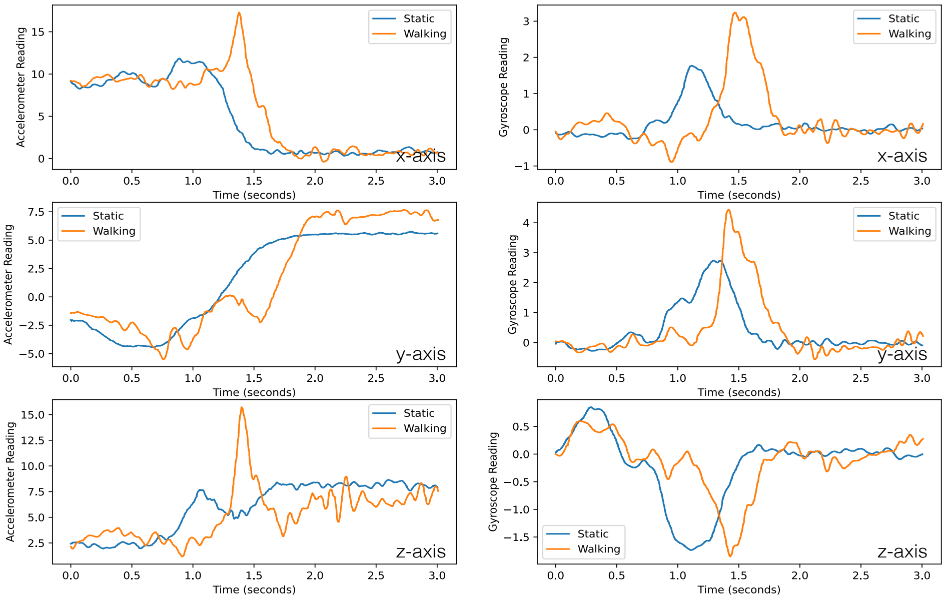


Fig. 3. A pilot study of the distinguishability of the lifting and vibrating modality

.

2's a quick, abrupt movement, and User 3's a variable or even hesitant motion. This variability poses a significant challenge for developing a universal gesture recognition system To accurately segment a user's raise gesture for subsequent authentication modules, algorithms must be tailored to understand and learn the nuances of each individual's movements. The variability demonstrated in the IMU data from different users is a testament to the complexity of this challenge.

## 5.3 Gesture Segmentation Algorithm

*5.3.1 Noise Removal.* The collected raw IMU data might contain high-frequency fluctuations that are not part of the intentional raise gesture. These fluctuations can be due to a variety of sources like micro-movements of the hand, environmental vibrations, or body motions irrelevant to the lifting behavior. To ensure real-time processing capabilities and minimize computational overhead, we applied an Exponential Moving Average (EMA) filter [32] for noise reduction. This filtering technique smooths out the data, effectively reducing the impact of transient movements and vibrations that do not contribute to the raise gesture pattern.

$$\mathrm{EMA}_t = \alpha \cdot \mathrm{IMU}_t + (1 - \alpha) \cdot \mathrm{EMA}_{t-1} \tag{6}$$

where $\mathrm{EMA}_t$ is the filtered value at time $t$, $\mathrm{IMU}_t$ is the raw IMU reading, and $\alpha$ represents the smoothing factor.

*5.3.2 Kinematic Analysis.* When analyzing the kinematic features of raising gestures, the IMU (Inertial Measurement Unit) data of the phone involves rotations in three-dimensional space. These rotations are inherently nonlinear, as they encompass complex movements and changes in orientation. Traditional Kalman filtering techniques are not well-suited for this type of nonlinear system [26]. Consequently, the Extended Kalman Filter (EKF) [34] emerges as a more appropriate choice for such applications. The EKF is adept at handling the nonlinear characteristics introduced by the device's motion and changes in direction.

We use the EKF to estimate an orientation represented as a quaternion. First, we predict the new state (newest orientation) using the immediate measurements of the gyroscopes, then we correct this state using

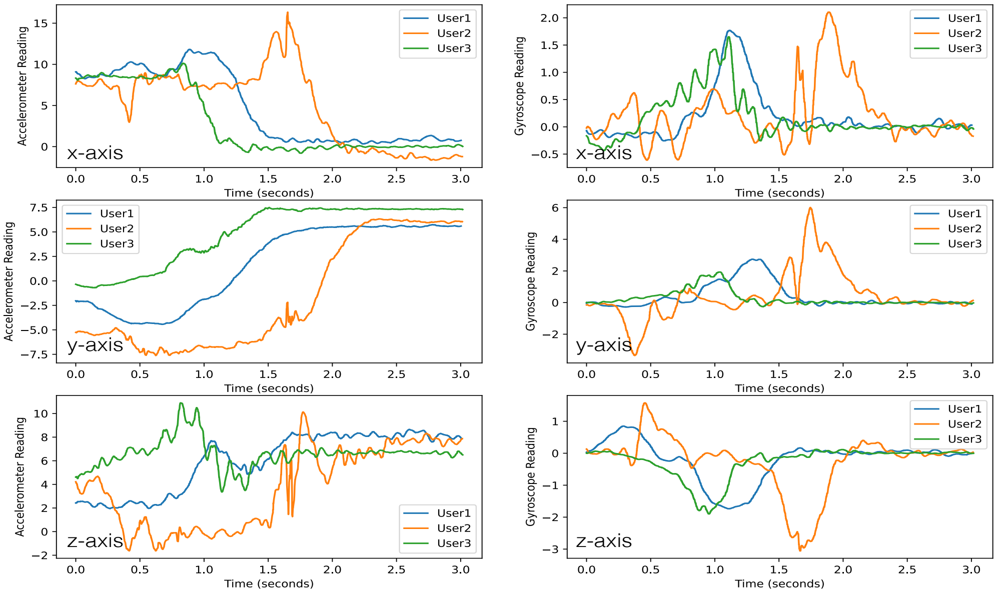


Fig. 4. A pilot study of the distinguishability of the lifting and vibrating modality

.

the measurements of the accelerometers. The process of estimating velocity from IMU data using an Extended Kalman Filter (EKF) [21, 36] involves several steps, which are outlined below:

***Data Acquisition and Preprocessing.*** IMU data, consisting of accelerometer and gyroscope measurements, are acquired and preprocessed:

$$\text{Accelerometer Data: } \mathbf{a} = [a_x, a_y, a_z]$$
$$\text{Gyroscope Data: } \boldsymbol{\omega} = [\omega_x, \omega_y, \omega_z]$$

where $\mathbf{a}$ and $\boldsymbol{\omega}$ are the linear acceleration and angular velocity vectors, respectively.

***State Estimation Using EKF.*** The EKF iteratively estimates the state vector, which includes the orientation of the device. The state update equations in the EKF are as follows:

- **State Vector:** In EKF, the state vector $x$ often includes quaternion components to represent the system's orientation. For instance, the state vector might be:

$$x = \left[q_w,\ q_x,\ q_y,\ q_z\right] \tag{7}$$

  where $q_w, q_x, q_y, q_z$ are the components of the quaternion.
- **State Prediction:**

$$\hat{\mathbf{x}}_{k|k-1} = f(\hat{\mathbf{x}}_{k-1|k-1}, \mathbf{u}_k)$$

  where $\hat{\mathbf{x}}_{k|k-1}$ is the predicted state, $f(\cdot)$ is the state transition function, $\hat{\mathbf{x}}_{k-1|k-1}$ is the previous state estimate, and $\mathbf{u}_k$ is the control input (gyroscope data).

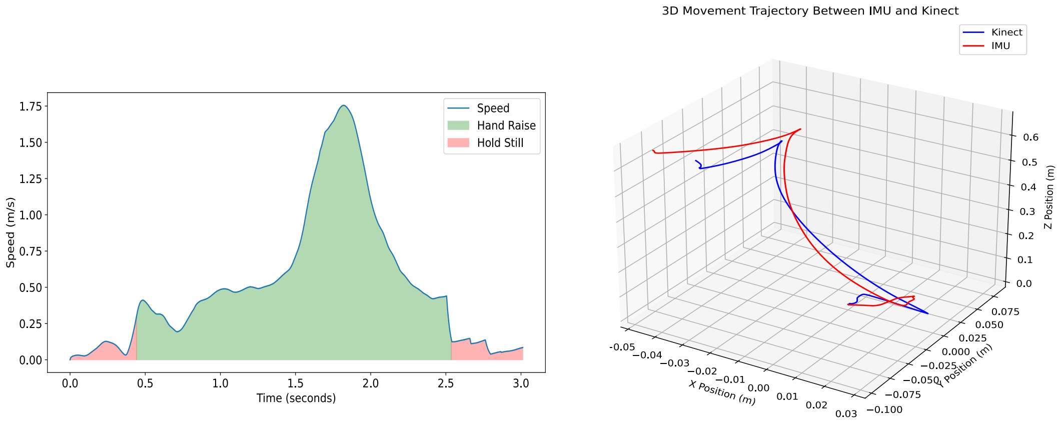


Fig. 5. Motion profile of raise gesture. (a) estimated speed profile by EKF with the segmentation of hand raise; (b) estimated 3D trajectory and groundtruth trajectory using Kinect

- **Update Equations:**

$$\begin{aligned}\mathbf{K}_k &= \mathbf{P}_{k|k-1}\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^T + \mathbf{R}_k)^{-1}\\ \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{z}_k - h(\hat{\mathbf{x}}_{k|k-1}))\\ \mathbf{P}_{k|k} &= (I - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1}\end{aligned}$$

  where $\mathbf{K}_k$ is the Kalman gain, $\mathbf{P}$ is the covariance matrix, $\mathbf{H}_k$ is the measurement matrix, $\mathbf{R}_k$ is the measurement noise covariance, $\mathbf{z}_k$ is the measurement vector (accelerometer data), and $h(\cdot)$ is the measurement function.

***Conversion to Global Coordinate System.*** The orientation estimate is used to transform the accelerometer data to the global coordinate system, and the gravity component is removed:

$$R(\cdot) = \begin{bmatrix} 1-2q_y^2-2q_z^2 & 2q_xq_y-2q_zq_w & 2q_xq_z+2q_yq_w\\ 2q_xq_y+2q_zq_w & 1-2q_x^2-2q_z^2 & 2q_yq_z-2q_xq_w\\ 2q_xq_z-2q_yq_w & 2q_yq_z+2q_xq_w & 1-2q_x^2-2q_y^2 \end{bmatrix} \tag{8}$$

$$\mathbf{a}_{\text{global}} = \mathbf{R}(\hat{\mathbf{x}}_{k|k})\mathbf{a} - \begin{bmatrix}0\\0\\g\end{bmatrix} \tag{9}$$

where $\mathbf{R}(\cdot)$ is the rotation matrix derived from the orientation estimate and $g$ is the acceleration due to gravity.

***Velocity Estimation.*** Finally, the velocity is estimated by integrating the linear acceleration over time:

$$\mathbf{v}_k = \mathbf{v}_{k-1} + \mathbf{a}_{\text{global},k}\Delta t$$

where $\mathbf{v}_k$ is the velocity at time step $k$ and $\Delta t$ is the time interval between measurements.

The norm of the velocity vector gives the speed at each time step:

$$\text{Speed}_k = \|\mathbf{v}_k\|$$

The position trajectory of the smartphone can be later calculated by an integration of the velocity vector. Fig. 5 shows a comparison between the estimated 3D trajectory and the ground truth trajectory using a Kinect camera and joint estimation algorithm [1].

*5.3.3 Rule-based Gesture Detection.* The 'Raise Gesture Detection' component of Raise2Auth employs a rule-based approach to accurately and swiftly identify the gesture of lifting the smartphone. This process hinges on the real-time estimation of the phone's 3D trajectory and orientation, which is accomplished through the application of an Extended Kalman Filter (EKF). Given the typically brief duration of the raising action, our system implements a set of specific rules designed to promptly capture the intent of a raise gesture. The rules for detecting the start and end of the raise gesture, as well as criteria for distinguishing it from other movements, are as follows:

$$V_{\text{start}} > v_1 \quad (\text{e.g., } V_{\text{start}} > 0.2 \text{ m/s}) \tag{10}$$

**Start of Raise Gesture:** where $V_{\text{start}}$ represents the velocity of the phone at the start of the gesture. The motion should be directed upwards, perpendicular to the Earth's surface, to qualify as the initiation of a raise gesture.

$$V_{\text{end}} < v_2 \quad (\text{e.g., } V_{\text{end}} < 0.15 \text{ m/s}) \tag{11}$$

**End of Raise Gesture:** The gesture is considered to end when the phone's velocity decreases to below $v_2$ and the phone remains stationary with its orientation vertically aligned upwards.

$$V_{\text{max during gesture}} > 0.8 \text{ m/s} \tag{12}$$

**Differentiation Criteria:** The raise gesture is confirmed only if the maximum velocity $V_{\text{max}}$ during the gesture exceeds 0.8 m/s. This criterion helps to distinguish the intended raise gesture from other similar movements.

By employing these rule-based criteria, Raise2Auth effectively distinguishes intended raise gestures from incidental movements, enhancing the accuracy of the system. The aforementioned velocity thresholds were obtained through empirical adjustment. This approach ensures that the transition into the grasping feature extraction phase occurs precisely and only in response to a genuine raise gesture, thereby maintaining the reliability and efficiency of the authentication process.

# 6 DUAL-FACTOR AUTHENTICATION

## 6.1 Dual-Factor Gesture Feature

In developing a biometric authentication system based on the concept of hand-raising to unlock a device, we dissect the action into two distinct descriptors: the grasp pattern of the hand holding the device, and the motion trajectory of the hand-raising action. This design allows for a comprehensive analysis of the unlocking gesture, tapping into deep-rooted, individual-specific characteristics that enhance the security and accuracy of the authentication process. Thus, we propose a novel multi-level descriptor of intrinsic behavior biometric that combines the uniqueness of hand raising pattern and the hand-holding pattern. The detailed pipeline is discussed below.

*6.1.1* ***Raise Gesture.*** As discussed in Section 5.3.2, we employ an Extended Kalman Filter (EKF) for precise motion profile estimation from the smartphone's IMU data. The Raise Gesture analysis utilizes the estimated 3D trajectory and orientation shifts during the hand-raising action, as depicted in Fig. **??**. This motion, while outwardly straightforward, involves a complex interplay of rotational and linear movements unique to each individual.

*6.1.2* ***Grasp Gesture.*** The Grasp Gesture Descriptor focuses on the static state when the user grips the device. This state initiates a unique vibration pattern, triggered by the device's interaction with the user's hand. We analyze the intricate IMU patterns that arise during this interaction. To isolate these from the biometrically irrelevant signals, we apply a high-pass filter to eliminate baseline wandering, which are the components with

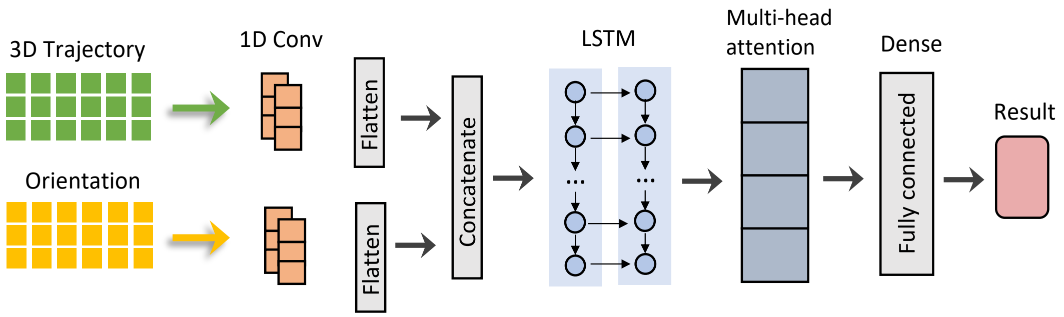


Fig. 6. Architecture of the Multi-view CNN-LSTM with attention.

low-frequency, non-characteristic in the vibration signals. The refined IMU data, free from these irrelevant oscillations, is then processed to the authentication model to discern the nuanced grip patterns unique to the user.

This dual-factor authentication model addresses certain limitations that may arise in situations involving significant body motion or changes in user posture. While the raise gesture may be challenging to authenticate in dynamic scenarios, the grasp gesture is less affected by the user's body posture, making it a reliable factor for authentication. However, it is also more susceptible to imitation attacks. By combining both raise and grasp gestures, our two-factor authentication model achieves a balance, ensuring robustness against motion and posture variability while maintaining high security against potential imitation threats. This integrated approach guarantees both the resilience and the security of the authentication process, making Raise2Auth a pioneering solution in mobile user authentication.

### 6.2 Authentication Modeling

We hypothesize that there exists intrinsic features in the smartphone raising behaviors, including a unique 3D trajectory and vibration pattern. Personal 3D trajectory data may contain complex and subtle individual characteristics. To extract spatial features and understand the temporal characteristics from 3D trajectories, we propose a CNN-LSTM based model with a multi-head attention mechanism [? ], which allows the model to focus on multiple aspects simultaneously in sequence data, such as different points in time or spatial locations within the trajectory.

Figure. 6 shows the general architecture of the CNN-LSTM-based model. The model features two distinct input pathways, each dedicated to a type of input data. The upper stream processes the spatial trajectory information, while the lower stream handles the orientation angle data. These inputs pass through their respective CNN layers, which are adept at extracting spatial features from the sequential data. The convolutional layers are followed by flattening operations that prepare the spatial features for subsequent temporal processing. The outputs of both streams are merged using a concatenation layer. This early fusion strategy ensures that spatial features from both the trajectory and orientation data are combined, allowing the model to capture correlations between the physical movement and the orientation of the object. The combined feature map is fed into an LSTM layer to capture long-term dependencies and temporal dynamics inherent in the sequence data. To further refine the feature representation, the model employs a multi-head attention mechanism after the LSTM layer, enabling the model to attend to different parts of the sequence simultaneously, a vital feature for capturing the nuances of complex temporal patterns.

### 6.3 Weighted Decision Fusion

The Weighted Decision Fusion module [23] is designed to intelligently balance the inputs from the raise and grasp gesture authenticators. Given the distinct advantages each gesture feature offers under varying circumstances, our approach dynamically adjusts their respective weights in the decision-making process to optimize authentication accuracy. The fundamental principle behind this module is the context-aware weighting of the raise and grasp gesture features. Initially, both authenticators are assigned equal weights (1:1), ensuring a balanced contribution from each in standard scenarios. However, the system is designed to detect the current idle state of the smartphone and adjust the weights accordingly:

- **Increased Weight for Raise Gesture:** In situations where the smartphone is detected to be in a static idle state, the weight assigned to the raise gesture authenticator is increased. This adjustment is based on the understanding that when the device is at rest, the raise gesture is likely to be more distinct and less prone to interference from external movements, thereby providing a more reliable authentication factor.
- **Increased Weight for Grasp Gesture:** Conversely, if the idle state involves motion or noticeable body movements, the system shifts the emphasis towards the grasp gesture authenticator by increasing its weight. This change accounts for the fact that in dynamic scenarios, the grasp gesture may offer a more stable and consistent authentication factor, as it is less affected by the broader range of movements typical of an active state.

## 7 SYSTEM IMPLEMENTATION

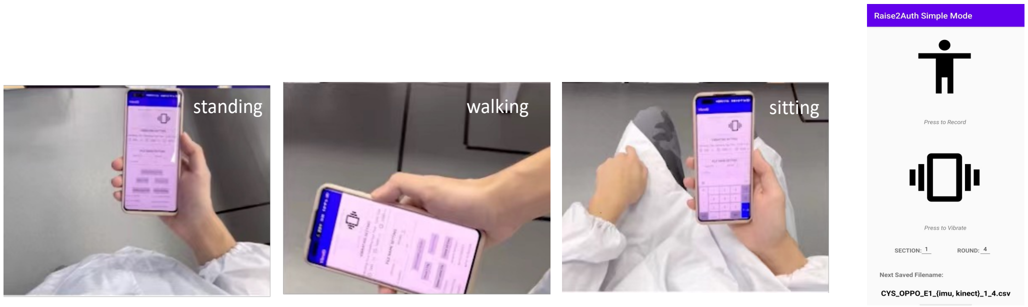
### 7.1 Data Collection

*7.1.1 Participants.* This study encompassed a diverse cohort of 45 participants, including 8 females and 37 males, with an average age of 20 years. The participants were divided into two groups: 25 registered users, who were the legitimate subjects of the experiment, and 19 attackers, who acted as unauthorized users. The attackers were involved in two sets of experiments: one where they did not observe the registered users and collected data 10 times based on their own habits as intruders, and another where they observed the registered users, mimicked their unlocking posture, and then collected 10 iterations of data as imitation attackers. The registered users, on the other hand, were instructed to replicate their unlocking movements 30-50 times for the model evaluation.

*7.1.2 Experimental Protocol and Data Collection.* Our experimental procedure was segmented into two distinct stages. The first stage involved participants extracting a mobile phone from their pant pockets without any motor-induced vibration and lasted for 3 seconds at maximum. The second stage was a simulation of motor vibrations with the phone remaining stationary, lasting for 0.5 seconds. We designed three scenarios for unlocking the phone corresponding to different user postures: standing (E1), sitting (E2), and walking (E3). An additional robustness experiment was also conducted where 5 registered users were selected at random to record their IMU data across three different environments: subway, bus, and indoors. Here, participants were required to remove their phones from their pockets and repeat this action for 10 iterations. The experiments were approved by the Internal Review Board (IRB) of the [University name is hidden for double-blinded review] for human subjects.

*7.1.3 Hardware and Software.* To assess the versatility of Raise2Auth, we conducted evaluations using four distinct smartphone models: OPPO, XiaoMi, HONOR, and Samsung. Each device was selected for its proprietary IMU sensor, which operates at a unique sampling rate: 880Hz, 600Hz, 600Hz, and 450Hz, respectively. The authentication model was developed on TensorFlow (version 2.10) and trained on the NVIDIA GeForce RTX 3090 GPU.

To collect the IMU data while the user is raising the smartphone, we developed an Android app using Android Studio, as shown in Fig. 7. It was designed to meticulously log the data streams from the 9-axis IMU sensor suite, which includes the accelerometer, gyroscope, and magnetometer, along with corresponding timestamps. In



(a)

(b)

Fig. 7. (a) three different data collection scenarios: standing, walking, sitting; (b) User Interface (UI) of the customized app

addition, we have fortified our on-device system evaluation by translating the authentication model to TensorFlow Lite format, integrating it within the mobile app.

*7.1.4 Ground-Truth Labeling.* To accurately measure the hand motions of our participants, we utilized a Microsoft Kinect depth camera to capture the three-dimensional trajectories of three skeletal markers: the hand, wrist, and thumb. This camera operated at a frequency of 30 Hz, providing us with high-fidelity motion data. After collecting the video data, we leveraged Mediapipe [1], an open-source body key points detection tool with stable and sufficient accuracy. The extracted key points were used to estimate the ground truth of the hand motion trajectory. Despite Mediapipe's state-of-the-art (SOTA) capabilities, occasional misannotations were unavoidable. To ensure the highest level of accuracy, we meticulously reviewed and manually corrected any instances where the video data yielded low-confidence scores.

## 7.2 Evaluation Metrics

In addressing this authentication challenge, we employ precision, recall, and the F-1 score as our key evaluation metrics. These metrics are derived from the counts of true positives (TP), false negatives (FN), and false positives (FP), defined as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{13}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{14}$$

Here, a high precision indicates that the system effectively limits access to only authorized users, whereas a high recall ensures that legitimate users are seldom incorrectly denied access. In scenarios where the testing set is imbalanced, precision and recall alone may not provide a complete picture of the system's performance. To address this, we also utilize the F-1 score, which harmonizes the balance between precision and recall, thereby mitigating the limitations posed by imbalanced datasets. The F-1 score is defined as:

$$\text{F-1} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \tag{15}$$

This formula ensures that both the concerns of excluding unauthorized access (precision) and permitting legitimate access (recall) are adequately weighted, offering a more comprehensive measure of the system's effectiveness.

## 7.3 Model Implementation

*7.3.1 Feature representation.* In implementing the Extended Kalman Filter (EKF), we address the scenario where the smartphone is initially positioned in a pocket. We establish the Initial State Estimate using quaternion notation, set as $q_0 = [0.73, -0.5, -0.4, -0.3]$. This quaternion is normalized, ensuring its norm is equal to one, which is essential for representing orientation accurately. The reference frame for gravitational acceleration is set in the North-East-Down (NED) coordinate system, defined as $g_{NED} = [0, 0, 9.81]$ m/s², aligning with Earth's gravitational pull. Measurement noise variances for the accelerometer and gyroscope are carefully determined, set as $\sigma^2\omega = 0.3^2$ and $\sigma^2\alpha = 0.5^2$ respectively, to account for sensor inaccuracies. The state of the EKF is updated at a frequency of 100 Hz, corresponding to every 10 ms, to ensure timely and accurate state estimations.

*7.3.2 Authentication model.* Our authentication framework employs a multi-view CNN-LSTM model, developed and tested using TensorFlow. The model's architecture comprises a two-layer Convolutional Neural Network (CNN) with a kernel size of 3 and 32 filters, designed to extract spatial features from the input data effectively. This is followed by a two-layer Long Short-Term Memory (LSTM) network, each layer containing 50 hidden neurons, to capture temporal dependencies in the data. A multi-head attention mechanism with N=2 heads is incorporated to enhance the model's ability to focus on relevant features across different positions in the input sequence. For optimization, we utilize the Adam optimizer with an initial learning rate set at 0.0004. To mitigate overfitting, a dropout strategy is employed with a rate of 0.5, alongside an L2 regularization with a coefficient $\lambda = 0.01$, ensuring the model's generalization across various datasets.

*7.3.3 Baseline models.* To quantify the efficacy of the dual-factor authentication solution, we create two baseline methods from recent work [9, 28] that utilize pick-up hand motion as the behavior biometric trait. For convenience, we name the work [9] by their feature extraction technique as "bag-of-movement-word" (BOMW), and the other work [28] is named as "SPU" in the remaining of this paper.

***BOMW.*** This method designs some common movement patterns in the dataset and calculates their occurrence frequencies in a certain sample. In the design of the “MOVEMENT-WORDS”, the Minibatch-K-Means algorithm is applied to cluster the data segments. A three-layer MLP model is applied as the final classifier to verify the user. The number of hidden layer units is 256, with an initial learning rate of 0.002.

***SPU.*** This method calculate a weighted multi-dimensional DTW with 3-axis accelerometer data and 3-axis gyroscope data. The weight of the accelerometer and gyroscope is set as 0.6 and 0.4.

## 8 PERFORMANCE EVALUATION

In this section, we aim to evaluate the *Raise2Auth* from the following three perspectives: 1) recognition performance, 2) resistance to various attacks, and 3) robustness quantification over different factors.

### 8.1 Recognition Performance

*8.1.1 Performance over Raise Gesture Detection.* To conserve smartphone energy, a critical aspect of our system's design is the preliminary detection of the raise gesture before initiating the authentication process. The system is programmed to recognize the user's intent to unlock the phone through their hand-raising motion, only activating the more energy-intensive authentication protocols upon confirmation of this intent. To validate the efficacy of this raise gesture detection, we conducted an analysis focusing on two key parameters: the trajectory error of the hand-raising motion and the temporal accuracy concerning the start and end points of this gesture.

The analysis of trajectory error yields an average deviation of 0.11 meters with a standard deviation of 0.043 meters. This precision underscores the system's capability to accurately track hand movements, thereby ensuring that the authentication process is triggered solely in response to a legitimate unlock gesture. Additionally, the temporal analysis regarding the initiation and completion of the gesture compared to the ground truth reveals average discrepancies of 170 milliseconds and 105 milliseconds, with standard deviations of 91.1 and 98.5, respectively. This result shows that the system can reliably discern the correct time frame of the gesture, minimizing the chances of false activation or unnecessary energy expenditure. The precision in detecting the raise gesture not only conserves battery life but also enhances the overall user experience by seamlessly integrating security measures into the natural interaction flow.

### 8.1.2 *Identification and Verification.*

***Identification.*** In this study, we quantify the performance of *Raise2Auth* in regard to classification accuracy for 25 legitimate users. From the collected dataset, we split the data into 20% of testing and 80% of training data. We run the model for 100 epochs to generate the learning representations from the user's vibration responses. We employ the Adam optimizer with a batch size of 32, a learning rate of 0.0004. Under this setting, during training, each subject is a genuine user, while the rest are imposters. Overall, the classification accuracy is 95.85%.

***Verification.*** We then evaluate user verification performance, assuming each user has their own customized pre-trained *Raise2Auth* system on their smartphone. This is regarded as a binary classification problem that has only two labels: genuine and imposter labels. To mitigate the data imbalance issue, we set twice as many imposters as legitimate users and apply the 5-fold cross-validation to enhance the generalization of the model. The result shows an averaged accuracy of 97.61%, precision=96.69%, recall=98.54%, and F1=97.94%.

*8.1.3 Ablation Study on Different Network Structures.* Building upon the foundational CNN-LSTM architecture, our study delved into an exploration of various structural combinations to optimize performance. We experimented with single-view and multi-view approaches to understand how the system benefits from different perspectives of the same data. Additionally, we incorporated attention mechanisms to assess their impact on the model's ability to focus on salient features for authentication.

The experimental results are shown in Table.1, demonstrating that attention mechanisms and multi-view network structures markedly improve the authentication accuracy of the *Raise2Auth* system. Single-view models achieved decent performance, which was notably enhanced by adding attention, signifying its effectiveness in feature refinement. The multi-view approach outperformed single-view models and reached peak accuracy when combined with attention mechanisms, underscoring their combined potential in creating a highly accurate biometric authentication system.

The multi-view model, which inherently captures a more holistic representation of the gesture by considering multiple aspects simultaneously, outperforms the single-view models with an accuracy of 0.92. This suggests that integrating multiple perspectives of the gesture leads to a more robust authentication process.

| **Network Structures** | **Accuracy** |
|---|---|
| Single-view (trajectory) | 0.82 |
| Single-view (orientation) | 0.86 |
| Single-view (trajectory) + attention | 0.85 |
| Single-view (orientation) + attention | 0.89 |
| Multi-view | 0.92 |
| Multi-view + attention | 0.97 |

Table 1. Accuracy of different network structures in the Raise2Auth system.

*8.1.4 Performance over Different Feature Representation.* To evaluate the performance of the feature descriptors with multi-view CNN-LSTM, we compared various feature representation strategies, including multi-dimensional DTW [28], movement-words [9], 3D trajectory, orientation and vibration-based handling feature. Fig. 8 shows the detailed performance metrics for different feature representation strategies. The performance analysis of

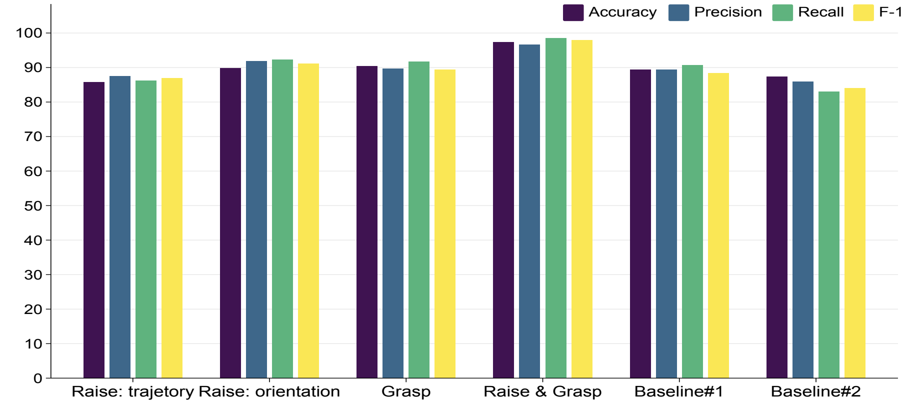


Fig. 8. Performance over different feature representation. Baseline1: BOMW, baseline2: SPU

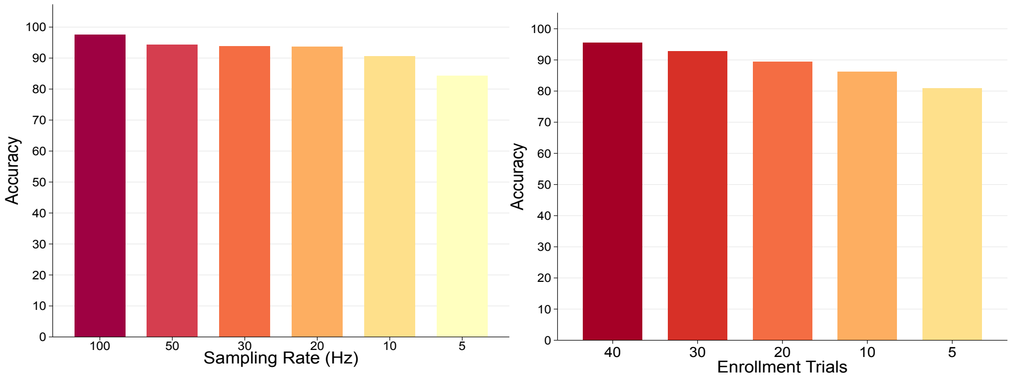


Fig. 9. Performance over different IMU sampling rate (left) and enrollment trails (right).

the authentication features reveals that while individual raise features like trajectory and orientation perform commendably, with accuracies above 93%, the combined Raise and Grasp feature achieves superior performance, reaching an accuracy of 97.31% and outperfored two baseline models.

*8.1.5 Performance over Different Data Sampling Rate.* The sampling rate is a critical factor that influences the precision of an IMU sensor. However, a higher sampling rate also entails increased data volume and computational costs. In pursuit of an optimal balance between system performance and power consumption, we investigated the accuracy of our system's recognition capabilities at various IMU sampling rates. As illustrated in Fig. 9, a reduced sampling rate leads to diminished IMU data quality, thereby impacting the system's performance. Nonetheless, the majority of built-in smartphone IMU sensors, which typically operate at sampling rates above 30 Hz, are capable of sustaining an accuracy level exceeding 90%. This suggests that while there is a performance trade-off at lower sampling rates, standard smartphone IMUs provide sufficient data quality for reliable authentication.

*8.1.6 Performance over Different Enrollment Data.* The quantity of enrollment data, or the data used to register a user with the system, is pivotal to the efficacy of authentication mechanisms. We investigated the effect of varying enrollment dataset sizes on the accuracy of our system's recognition capabilities. By adjusting the amount of enrollment data, we were able to identify the optimal quantity necessary for maximum recognition efficacy. We also pinpointed the threshold beyond which additional data does not yield substantial accuracy gains. Fig.

9 illustrates a decline in performance as the number of enrollment instances decreases. Notably, the system maintains an accuracy rate above 80% even with a minimal enrollment of just 5 instances. Given that raising the smartphone is a frequent and natural user behavior, collecting sufficient enrollment samples for robust training is a feasible task for most users.

#### 8.1.7 *Performance over Intruders.*

To rigorously assess the vulnerability of *Raise2Auth* to unauthorized access attempts, we conducted an experiment involving both legitimate users and potential intruders. In this setup, we designated 25 enrolled subjects as legitimate users, whose data was incorporated into the authentication system for training purposes. Additionally, we introduced a new group of 19 subjects, none of whom had their data previously recorded or used by the system, serving as potential intruders. This approach was intended to simulate a realistic scenario where an intruder attempts to gain unauthorized access, with the system having no prior knowledge of the intruder's biometric data.

In evaluating the system's resistance to these intrusion attempts, we focused on the False Acceptance Rate (FAR), a critical metric that measures the likelihood of the system incorrectly granting access to an unauthorized user. The results of our testing revealed an averaged FAR of 2.1%, with a standard deviation of 2.16. This indicates that, on average, *Raise2Auth* falsely accepted unseen intruders at a rate of just over 2%, demonstrating a commendable level of security against intrusion attempts.

### 8.2 Attack Resistance Performance

We assume that the malicious attacker knows the underlying mechanism of *Raise2Auth* and has recorded the legitimate user's hand-raising behavior through a high-resolution camera at a distance of 30 cm. In our experiment, a 3-second legitimate user's hand motion was recorded.

**Attack via Imitation Attacks:** An imitation attack represents a scenario where a malicious individual, having observed and recorded the legitimate user's smartphone holding position and hand-raising behavior, attempts to manually replicate the motion in an effort to deceive the *Raise2Auth* system. This type of attack is predicated on the attacker's ability to mimic the hand movements closely enough to generate sensor readings that fall within the system's acceptance threshold. In our testing, attackers were provided with footage of the legitimate user's hand position and motions and given time to practice the replication. Their subsequent attempts to authenticate via *Raise2Auth* were meticulously logged to evaluate the system's resilience to human-led spoofing attempts.

**Attack via Synthetic Attacks:** Conversely, a synthetic attack involves the use of a robotic apparatus designed to emulate the user's hand-raising behavior with high precision. After capturing the legitimate user's movements, the exact motion patterns were programmed into a mechanical arm (UFACTORY xArm 6), which then sought to breach the *Raise2Auth* system with a level of accuracy unattainable by human imitation. This form of attack tests the system's robustness against highly sophisticated spoofing methods that leverage technology to replicate biometric behaviors.

For our experimental procedure, we captured the unique smartphone-raising trajectories of each legitimate user (N=25) utilizing a Kinect depth camera. Subsequently, we mounted the smartphone onto the robotic arm's end effector and programmed it to replicate these recorded trajectories, thus enabling the robot to mimic the legitimate user's motion with precision. As illustrated in Fig. 10, the findings indicate that attackers attempting to manually mimic the legitimate user's hand movements achieve minimal success, with a spoofing rate of less than 1% for the Raise gesture authenticator and 4.5% for the Grasp gesture authenticator independently. However, the simultaneous replication of both gestures presents a considerable challenge, substantially enhancing the dual-phase authenticator's effectiveness in mitigating most imitation attacks. Synthetic attacks exhibit even lower success rates for both authenticators, likely attributed to the required precision in movement and rotation, as well as a consistent holding force.

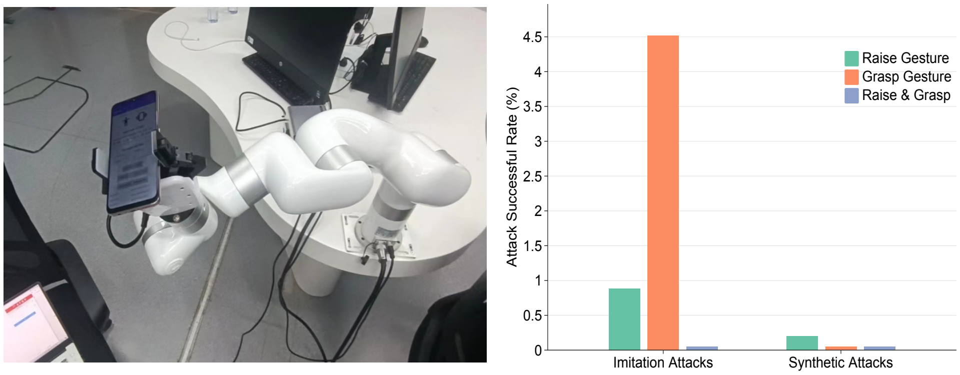


Fig. 10. Synthetic attack setup using robot arm(left) and attack resistance result (right).

Our experiments gauged the effectiveness of the authentication system in differentiating between real human movements and their synthetic counterparts, thereby underscoring the potential vulnerabilities and the required thresholds for secure authentication.

## 8.3 Robustness Quantification

To investigate the robustness of Raise2Auth in real-life application scenarios, we conducted a set of evaluations of performance with respect to different IMU sensors, environmental settings, and body postures.

*8.3.1 Impact of IMU Sensors.* The efficacy of motion-based authentication systems is intricately linked to the quality of data procured from Inertial Measurement Unit (IMU) sensors. These sensors, varying across models and manufacturers, can exhibit differences in noise levels, sensitivity, and data sampling rates—all of which are crucial factors that could potentially impact the performance of such systems. To rigorously evaluate the influence of IMU sensor variability on our system, we conducted tests using four different Android smartphones. Each phone was equipped with a distinct IMU sensor model and operated at a unique sampling frequency. This diverse selection was intentional, aimed at encompassing a broad spectrum of potential user devices.
**Insights:** Our results, depicted in Fig. 11, illustrate that despite variations in IMU sensors and sampling rates across different smartphone models, system accuracy remained consistently high. This suggests that the sampling rates employed were sufficient to capture the defining features of behavior biometrics effectively. However, the slight underperformance observed with Samsung smartphones may be attributed to the unique positioning of the IMU sensor within these devices, which could affect the data quality and, consequently, the system's recognition capabilities.

*8.3.2 Impact of Environment Scenarios.* In our pursuit to authenticate the performance of the system in real-world settings, we extended our experimentation to encompass various environmental scenarios where phone unlocking is routinely performed. Specifically, we collected data in three distinct environments: indoor settings, aboard public buses, and within subway trains. These locations were chosen to represent a range of ambient conditions that users commonly encounter, thus providing a realistic assessment of the system's adaptability and reliability.
**Insights:** Our analysis, detailed in Fig. 11, indicates that environmental context plays a significant role in the authentication process. Indoor environments, with their relatively controlled conditions, demonstrated the lowest false reject rate. In contrast, the fluctuating conditions aboard buses and the unique movement dynamics found in subway scenarios introduced more variables, affecting the system's performance to varying degrees. Particularly notable were the scenarios within subway environments, where the FRR was impacted not only by the inherent movement of the train but also by the varying degrees of crowding at different times and locations. This crowding,

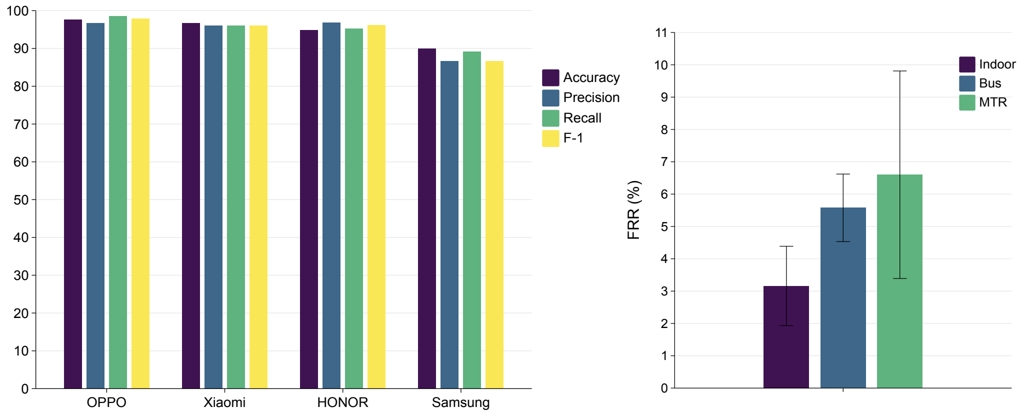


Fig. 11. Performance over different smartphones (left) and different environment settings (right).

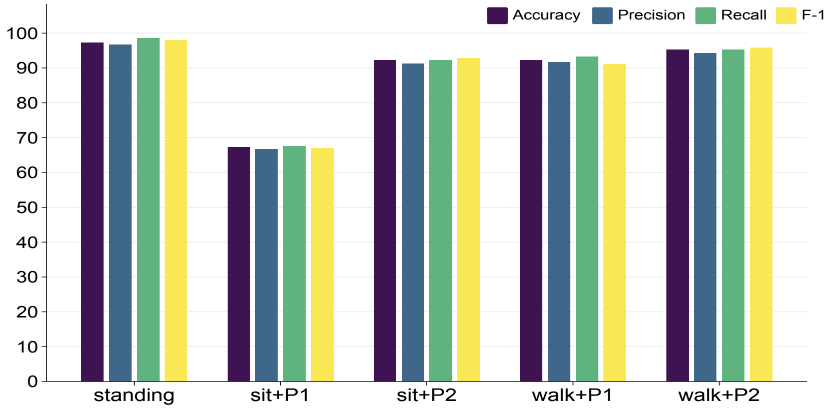


Fig. 12. Performance among different body postures. P1: raise gesture, P2: raise gesture and grasp gesture

coupled with the vehicle's motion, contributed to a higher standard deviation in FRR. Despite these challenges, the system maintained a commendable level of accuracy, underscoring its robust design.

*8.3.3 Reliability on Body Postures.* Our experiment meticulously recorded the unlocking process across three common body postures: standing, sitting, and walking, which are representative of the typical scenarios in which users interact with their devices. To ensure a comprehensive understanding of how body posture affects authentication, we captured the nuances of each posture's impact on the system's reliability.

**Insights:** As depicted in Fig. 12, our findings present a detailed examination of how authentication performance correlates with user posture. While the postures of standing and sitting showed distinct differences, with sitting posture posing certain limitations and potential inaccuracies when relying solely on the raise gesture for authentication, walking posture demonstrated minimal impact on performance. This suggests that the dynamic nature of walking does not significantly affect the raise gesture's effectiveness. Conversely, the grasp feature's performance remains uninfluenced by changes in body posture, thereby underscoring its value in enhancing system robustness. The integration of both raise and grasp gestures significantly improves the system's reliability across various postures. These insights reinforce the efficacy of our dual-factor authentication approach, particularly in its robustness against the variances introduced by different user states,

Table 2. System overhead on different smartphones

| | OPPO | HONOR | Samsung |
|---|---|---|---|
| **Execution Time (ms)** | 110.1 | 131.2 | 126.3 |
| **Memory Usage (MB)** | 8.69 | 9.28 | 10.64 |

[1] Results are tested given the same 3.5-second IMU sample and averaged for 20 runs.

### 8.4 On-Device Evaluation

To assess the practical applicability and real-time performance of Raise2Auth as a mobile authentication solution, we conducted an on-device evaluation. This involved collecting and analyzing data directly from smartphones to understand the system's operational latency when deployed in real-world scenarios. As detailed in Table. 2, our evaluation focused on a typical authentication session involving the detection of hand-raising and grip posture actions, segmented into two primary computational phases: feature extraction and deep learning inference. Our findings reveal that, on average, the entire authentication process from initial gesture detection to final user verification registers a latency between 100 to 200 milliseconds across a range of smartphones. This latency encompasses both the feature extraction and model inference stages, and remains within a tolerable range for daily use authentication, indicating that Raise2Auth can seamlessly integrate into users' routine interactions with their devices without causing noticeable delays.

## 9 DISCUSSION AND FUTURE WORK

### 9.1 Comparison with Existing Gesture Based Mobile Authentication Methods

Raise2Auth stands as a significant advancement in the realm of gesture-based mobile authentication, differentiating itself through a unique focus on the natural interaction of users with their smartphones, particularly the hand-raising gesture. While Lee *et al.* [28] and Boshoff *et al.* [9] have laid the groundwork in utilizing hand pick-up movements for behavior-biometric authentication, their approaches primarily hinge on the direct extraction of features from accelerometer and gyroscope data. These systems, though innovative, do not explore the deeper, more nuanced characteristics of the hand-raising gesture and often struggle with the significant noise inherent in IMU sensor data, caused by both environmental factors and body movements. In addition, their methodologies primarily rely on simplistic segmentation of the hand-raising gesture based on changes in acceleration data. This approach does not effectively distinguish between the specific gesture of raising the phone and other routine activities involving the device,

In contrast, Raise2Auth capitalizes on a sophisticated analytical framework that combines motion trajectory and hand grip dynamics during phone pick-up, making full use of the smartphone's built-in Inertial Measurement Unit (IMU). This is achieved through the implementation of an Extended Kalman Filter (EKF), which integrates accelerometer and gyroscope data to extract detailed information on movement trajectories and phone rotation. Further, it employs deep learning models to unearth deeper, more intricate features from this data. This methodological innovation enables Raise2Auth to offer a nuanced and secure authentication mechanism that effectively counteracts the impact of sensor noise and environmental disturbances. Moreover, Raise2Auth introduces a dual-factor authentication model that addresses potential limitations encountered during significant body motions or changes in user posture. While the gesture of raising the phone forms the primary basis of authentication, it may be susceptible to constraints arising from extensive body movements or postural adjustments. To mitigate these issues, Raise2Auth incorporates the phone grip posture as a complementary authentication factor. The grip posture, though potentially vulnerable to imitation attacks, is less likely to be influenced by the user's posture, offering a balanced approach that ensures both robustness and security. This two-factor authentication model synergizes the distinct advantages of both gesture and grip analyses, ensuring a high degree of accuracy and resilience against sophisticated attempts at unauthorized access.

### 9.2 Exploring Gesture-independent Authentication

The current iteration of Raise2Auth has been designed around the specific gesture of extracting and lifting the smartphone as a means of authentication. This process, while effective, represents a single instance of interaction between the user and their device. There is a compelling opportunity to expand our system's capabilities beyond fixed gestures to include continuous authentication. This would involve analyzing a broader spectrum of natural interactions that users have with their phones throughout the day. By capturing more spontaneous and varied grip and hand movements, we can potentially extract a richer set of physiological characteristics inherent to the user's limb movements.

### 9.3 Extending Smartphone-Based Authentication to Wearable Devices

Building upon our existing research with Raise2Auth, extending our authentication approach to include wearable devices such as smartwatches and smart rings opens up unprecedented opportunities. The persistent wear of these devices not only facilitates the convenient collection of a broader array of data but also enables a deeper exploration into users' behavioral and physiological characteristics. The continuous wearability of such devices stands to significantly augment our authentication system. Unlike smartphones, which are intermittently handled, wearables like smartwatches remain in constant contact with the user, providing continuous streams of data. This constant data flow allows for the ongoing authentication that is both seamless and more secure, leveraging real-time physiological and behavioral biometrics. Moreover, the addition of smart rings into the authentication framework introduces the capability to capture finer-grained actions, such as specific finger movements, that are beyond the reach of smartphone sensors. This level of detail could unveil new dimensions of user interaction, offering insights into unique gestural inputs that are highly individualized and difficult to replicate.

## 10 CONCLUSION

In conclusion, by leveraging the natural interactions users have with their smartphones, such as the distinctive ways they grasp and lift their devices, Raise2Auth introduces a novel approach to authentication that is both intuitive and robust. Utilizing the built-in Inertial Measurement Unit (IMU) to analyze the intricate dynamics of hand grip and motion trajectory during phone pick-up, this system sets itself apart from traditional authentication methods that rely on physiological or behavioral traits vulnerable to spoofing. Our comprehensive evaluation of Raise2Auth, involving 25 participants, has not only demonstrated the system's ability to achieve an impressive verification accuracy rate of over 97.6% but also its capability to maintain a False Acceptance Rate (FAR) of less than 0.1% against both imitation and synthetic attacks. These results underscore the system's exceptional reliability and security, showcasing the potential of combining gestural and postural biometrics with advanced IMU data analysis for authentication purposes.

## REFERENCES

[1] 2023. https://developers.google.com/mediapipe. [Online; accessed 19-Oct-2023].

[2] Mohammed Abuhamad, Tamer Abuhmed, David Mohaisen, and DaeHun Nyang. 2020. AUToSen: Deep-learning-based implicit continuous authentication using smartphone sensors. *IEEE Internet of Things Journal* 7, 6 (2020), 5008–5020.

[3] Mohammed Abuhamad, Ahmed Abusnaina, DaeHun Nyang, and David Mohaisen. 2020. Sensor-based continuous authentication of smartphones' users using behavioral biometrics: A contemporary survey. *IEEE Internet of Things Journal* 8, 1 (2020), 65–84.

[4] Alejandro Acien, Aythami Morales, John V Monaco, Ruben Vera-Rodriguez, and Julian Fierrez. 2021. TypeNet: Deep learning keystroke biometrics. *IEEE Transactions on Biometrics, Behavior, and Identity Science* 4, 1 (2021), 57–70.

[5] Adnan Bin Amanat Ali, Vasaki Ponnusamy, Anbuselvan Sangodiah, Roobaea Alroobaea, NZ Jhanjhi, Uttam Ghosh, and Mehedi Masud. 2021. Smartphone security using swipe behavior-based authentication. *Intelligent Automation & Soft Computing* 29, 2 (2021), 571–585.

[6] Blaine Ayotte, Mahesh Banavar, Daqing Hou, and Stephanie Schuckers. 2020. Fast free-text authentication via instance-based keystroke dynamics. *IEEE Transactions on Biometrics, Behavior, and Identity Science* 2, 4 (2020), 377–387.

[7] Hongliang Bi, Yuanyuan Sun, Jiajia Liu, and Lihao Cao. 2021. SmartEar: Rhythm-based tap authentication using earphone in information-centric wireless sensor network. *IEEE Internet of Things Journal* 9, 2 (2021), 885–896.

[8] Anusha Bodepudi and Manjunath Reddy. 2020. Spoofing Attacks and Mitigation Strategies in Biometrics-as-a-Service Systems. *Eigenpub Review of Science and Technology* 4, 1 (2020), 1–14.

[9] Dutliff Boshoff, Alexander Scriba, Raphael Nkrow, and Gerhard P Hancke. 2023. Phone Pick-up Authentication: A Gesture-Based Smartphone Authentication Mechanism. In *2023 IEEE International Conference on Industrial Technology (ICIT)*. IEEE, 1–6.

[10] Jagmohan Chauhan, Young D Kwon, Pan Hui, and Cecilia Mascolo. 2020. Contauth: Continual learning framework for behavioral-based user authentication. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 4 (2020), 1–23.

[11] Huijie Chen, Fan Li, Wan Du, Song Yang, Matthew Conn, and Yu Wang. 2020. Listen to your fingers: User authentication based on geometry biometrics of touch gesture. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 3 (2020), 1–23.

[12] Yongliang Chen, Tao Ni, Weitao Xu, and Tao Gu. 2022. SwipePass: Acoustic-based second-factor user authentication for smartphones. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 3 (2022), 1–25.

[13] Yanjiao Chen, Meng Xue, Jian Zhang, Qianyun Guan, Zhiyuan Wang, Qian Zhang, and Wei Wang. 2021. Chestlive: Fortifying voice-based authentication with chest motion biometric on smart devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 4 (2021), 1–25.

[14] Seokmin Choi, Junghwan Yim, Se Jun Kim, Yincheng Jin, Di Wu, and Zhanpeng Jin. 2023. VibPath: Two-Factor Authentication with Your Hand's Vibration Response to Unlock Your Phone. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 3 (2023), 1–26.

[15] Haipeng Dai, Wei Wang, Alex X Liu, Kang Ling, and Jiajun Sun. 2019. Speech based human authentication on smartphones. In *2019 16th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*. IEEE, 1–9.

[16] Deniz Ekiz, Yekta Said Can, Yagmur Ceren Dardagan, and Cem Ersoy. 2020. Can a smartband be used for continuous implicit authentication in real life. *IEEE Access* 8 (2020), 59402–59411.

[17] Yang Gao, Yincheng Jin, Jagmohan Chauhan, Seokmin Choi, Jiyang Li, and Zhanpeng Jin. 2021. Voice in ear: Spoofing-resistant and passphrase-independent body sound authentication. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 1 (2021), 1–25.

[18] Yang Gao, Wei Wang, Vir V Phoha, Wei Sun, and Zhanpeng Jin. 2019. EarEcho: Using ear canal echo for wearable authentication. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–24.

[19] Raul Garcia-Martin and Raul Sanchez-Reillo. 2020. Vein biometric recognition on a smartphone. *IEEE Access* 8 (2020), 104801–104813.

[20] Sandeep Gupta, Rajesh Kumar, Mouna Kacimi, and Bruno Crispo. 2022. IDeAuth: A novel behavioral biometric-based implicit deauthentication scheme for smartphones. *Pattern Recognition Letters* 157 (2022), 8–15.

[21] Jouni Hartikainen, Arno Solin, and Simo Särkkä. 2011. Optimal filtering with Kalman filters and smoothers. *Department of biomedica engineering and computational sciences, Aalto University School of Science, 16th August* (2011).

[22] Jong-Hyuk Im, Seong-Yun Jeon, and Mun-Kyu Lee. 2020. Practical privacy-preserving face authentication for smartphones secure against malicious clients. *IEEE Transactions on Information Forensics and Security* 15 (2020), 2386–2401.

[23] Byeungwoo Jeon and David A Landgrebe. 1999. Decision fusion approach for multitemporal classification. *IEEE transactions on geoscience and remote sensing* 37, 3 (1999), 1227–1233.

[24] Kevin Jiokeng, Gentian Jakllari, and André-Luc Beylot. 2022. I Want to Know Your Hand: Authentication on Commodity Mobile Phones Based on Your Hand's Vibrations. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 2 (2022), 1–27.

[25] Young-Hoo Jo, Seong-Yun Jeon, Jong-Hyuk Im, Mun-Kyu Lee, et al. 2016. Security analysis and improvement of fingerprint authentication for smartphones. *Mobile Information Systems* 2016 (2016).

[26] Simon J Julier and Jeffrey K Uhlmann. 1997. New extension of the Kalman filter to nonlinear systems. In *Signal processing, sensor fusion, and target recognition VI*, Vol. 3068. Spie, 182–193.

[27] Rajesh Kumar, Vir V Phoha, and Abdul Serwadda. 2016. Continuous authentication of smartphone users by fusing typing, swiping, and phone movement patterns. In *2016 IEEE 8th international conference on biometrics theory, applications and systems (BTAS)*. IEEE, 1–8.

[28] Wei-Han Lee, Xiaochen Liu, Yilin Shen, Hongxia Jin, and Ruby B Lee. 2017. Secure pick up: Implicit authentication when you start using the smartphone. In *Proceedings of the 22nd ACM on symposium on access control models and technologies*. 67–78.

[29] Borui Li, Wei Wang, Yang Gao, Vir V Phoha, and Zhanpeng Jin. 2020. Wrist in motion: A seamless context-aware continuous authentication framework using your clickings and typings. *IEEE Transactions on Biometrics, Behavior, and Identity Science* 2, 3 (2020), 294–307.

[30] Zengpeng Li, Ding Wang, and Eduardo Morais. 2020. Quantum-safe round-optimal password authentication for mobile devices. *IEEE transactions on dependable and secure computing* 19, 3 (2020), 1885–1899.

[31] Sakorn Mekruksavanich and Anuchit Jitpattanakul. 2021. Deep learning approaches for continuous authentication based on activity patterns using mobile sensing. *Sensors* 21, 22 (2021), 7519.

[32] Masafumi Nakano, Akihiko Takahashi, and Soichiro Takahashi. 2017. Generalized exponential moving average (EMA) model with particle filtering and anomaly detection. *Expert Systems with Applications* 73 (2017), 187–200.

[33] Ioannis Papavasileiou, Zhi Qiao, Chenyu Zhang, Wenlong Zhang, Jinbo Bi, and Song Han. 2021. GaitCode: Gait-based continuous authentication using multimodal learning and wearable sensors. *Smart Health* 19 (2021), 100162.

[34] Maria Isabel Ribeiro. 2004. Kalman and extended kalman filters: Concept, derivation and properties. *Institute for Systems and Robotics* 43, 46 (2004), 3736–3741.

[35] Zhang Rui and Zheng Yan. 2018. A survey on biometric authentication: Toward secure and privacy-preserving identification. *IEEE access* 7 (2018), 5994–6009.

[36] Angelo Maria Sabatini. 2011. Kalman-filter-based orientation determination using inertial/magnetic sensors: Observability analysis and performance evaluation. *Sensors* 11, 10 (2011), 9182–9206.

[37] Chao Shen, Zhongmin Cai, Xiaohong Guan, Youtian Du, and Roy A Maxion. 2012. User authentication through mouse dynamics. *IEEE Transactions on Information Forensics and Security* 8, 1 (2012), 16–30.

[38] Chao Shen, Yuanxun Li, Yufei Chen, Xiaohong Guan, and Roy A Maxion. 2017. Performance analysis of multi-motion sensor behavior for active smartphone authentication. *IEEE Transactions on Information Forensics and Security* 13, 1 (2017), 48–62.

[39] Cong Shi, Jian Liu, Nick Borodinov, Bruno Leao, and Yingying Chen. 2020. Towards environment-independent behavior-based user authentication using wifi. In *2020 IEEE 17th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*. IEEE, 666–674.

[40] Wei Song, Hong Jia, Min Wang, Yuezhong Wu, Wanli Xue, Chun Tung Chou, Jiankun Hu, and Wen Hu. 2022. Pistis: Replay Attack and Liveness Detection for Gait-Based User Authentication System on Wearable Devices Using Vibration. *IEEE Internet of Things Journal* 10, 9 (2022), 8155–8171.

[41] Yunpeng Song and Zhongmin Cai. 2022. Integrating Handcrafted Features with Deep Representations for Smartphone Authentication. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 1 (2022), 1–27.

[42] Jack Sturgess, Simon Birnbach, Simon Eberz, and Ivan Martinovic. 2022. RingAuth: Wearable Authentication using a Smart Ring. *arXiv preprint arXiv:2301.03594* (2022).

[43] Fangmin Sun, Chenfei Mao, Xiaomao Fan, and Ye Li. 2018. Accelerometer-based speed-adaptive gait authentication method for wearable IoT devices. *IEEE Internet of Things Journal* 6, 1 (2018), 820–830.

[44] Cong Wang, Yanru Xiao, Xing Gao, Li Li, and Jun Wang. 2021. A framework for behavioral biometric authentication using deep metric learning on mobile devices. *IEEE Transactions on Mobile Computing* 22, 1 (2021), 19–36.

[45] Xiangyu Xu, Jiadi Yu, Yingying Chen, Qin Hua, Yanmin Zhu, Yi-Chao Chen, and Minglu Li. 2020. TouchPass: Towards behavior-irrelevant on-touch user authentication on smartphones leveraging vibrations. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*. 1–13.

[46] Xiaojing Yu, Zhijun Zhou, Mingxue Xu, Xuanke You, and Xiang-Yang Li. 2020. Thumbup: Identification and authentication by smartwatch using simple hand gestures. In *2020 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE Computer Society, 1–10.

[47] Xinman Zhang, Dongxu Cheng, Pukun Jia, Yixuan Dai, and Xuebin Xu. 2020. An efficient android-based multimodal biometric authentication system with face and voice. *IEEE Access* 8 (2020), 102757–102772.

[48] Xinchen Zhang, Yafeng Yin, Lei Xie, Hao Zhang, Zefan Ge, and Sanglu Lu. 2020. TouchID: User authentication on mobile devices via inertial-touch gesture analysis. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 4 (2020), 1–29.

[49] Zheng Zheng, Qian Wang, and Cong Wang. 2023. Spoofing Attacks and Anti-Spoofing Methods for Face Authentication over Smartphones. *IEEE Communications Magazine* (2023).

[50] Bing Zhou, Zongxing Xie, Yinuo Zhang, Jay Lohokare, Ruipeng Gao, and Fan Ye. 2021. Robust human face authentication leveraging acoustic sensing on smartphones. *IEEE Transactions on Mobile Computing* 21, 8 (2021), 3009–3023.

[51] Tiantian Zhu, Lei Fu, Qiang Liu, Zi Lin, Yan Chen, and Tieming Chen. 2020. One cycle attack: Fool sensor-based personal gait authentication with clustering. *IEEE Transactions on Information Forensics and Security* 16 (2020), 553–568.