2025 IEEE 22nd International Conference on Mobile Ad-Hoc and Smart Systems (MASS) | 979-8-3315-6599-2/25/$31.00 ©2025 IEEE | DOI: 10.1109/MASS66014.2025.00066

Fingerprint Authentication Using Vibration-boosted Refreshing Touchscreen

Wen Liu
Lyle School Of Engineering
Southern Methodist University
Dallas, USA
wenliu@smu.edu

Zeyu Deng
Lyle School Of Engineering
Southern Methodist University
Dallas, USA
zeyud@smu.edu

Chen Wang
Lyle School Of Engineering
Southern Methodist University
Dallas, USA
cwang6@smu.edu

***Abstract*—The increasing reliance on smartphones for secure digital interactions requires robust, user-friendly authentication methods. Current smartphone authentication depends on dedicated biometric sensors (e.g., fingerprint or facial recognition), which not only raise hardware costs but are also subject to privacy concerns. Meanwhile, most existing methods overlook the potential of passive, natural interactions like tapping the screen with vibration feedback as opportunities for seamless and secure authentication using only built-in smartphone sensors. This paper proposes a versatile behavioral biometric authentication method that leverages active motor vibration during screen taps to verify user identity. To our knowledge, this is the first system to demonstrate that standard touchscreens can capture fingerprint-like biometrics. Specifically, when a user interacts with the device, brief vibrations generated by the smartphone's haptic motor elicit distinctive user-specific behavioral responses. Our method captures and fuses multi-modal signals from the built-in microphone, inertial sensor, and touchscreen without requiring users to touch a specific sensor area or hold the phone at a fixed angle or distance, offering seamless and unobtrusive authentication. Our system integrates a hybrid deep-learning model combining classification models for each modality to classify users based on these subtle patterns. Our system achieved up to 94% accuracy integrating multi-vibration stimuli and a 96% multi-modal authentication accuracy with 16 participants.**

***Index Terms*—User Authentication, Touchscreen, Multisensing, Vibration Signal**

I. Introduction

Smartphones have become ubiquitous in modern life, and the global penetration rate continued its upward trend, reaching an estimated 71% in 2024 [1]. They serve as communication tools and platforms for a wide range of security-critical digital services, including banking, e-commerce, and electronic document signing. As these services increasingly rely on mobile platforms, ensuring secure and seamless user authentication has become a central challenge.

Knowledge-based methods like PINs and passwords remain common due to their simplicity, but they are vulnerable to phishing, guessing, and shoulder-surfing attacks. Additionally, prompting and switching between number and alphabet keyboard windows during sensitive tasks can interrupt the user experience and break the interaction flow [2]. Biometric methods such as facial recognition improve convenience but often require users to align their faces precisely or move the phone to an optimal distance, introducing delay and friction. Moreover, facial recognition systems can be spoofed using processed face images or 3D masks [3], posing serious security concerns. Though widely adopted, fingerprint authentication relies on dedicated hardware sensors that increase manufacturing costs and remain susceptible to spoofing through high-resolution fingerprint replicas or 3D-printed molds. Behavioral biometrics have emerged as an alternative to traditional methods, aiming to authenticate users based on their unique interaction patterns. However, existing approaches often depend on single-modal signals such as keystroke timing or touch pressure and still require explicit user effort or long-term monitoring [4].

All these methods demand explicit user effort, whether aligning or touching specific parts of the device, which can be inconvenient or error-prone. In contrast, tapping the touchscreen is among the most frequent, natural, and low-effort smartphone interactions across all apps and use cases. Touchscreens typically serve to record tap positions, yet their sensing potential creates an opportunity to design authentication systems that are secure, seamless, cost-effective, and naturally integrated into the user's everyday interactions without the need for any dedicated biometric components.

This work is the first to demonstrate that standard smartphone touchscreens when enhanced with vibration stimuli and native multi-modal sensing, can capture fingerprint-related behaviors for authentication. Unlike conventional methods, our system leverages the natural tapping interaction without requiring users to touch a specific sensor area or align the phone at a fixed position or distance. It activates the phone's built-in vibration motor upon a tap and simultaneously records sensor data from the smartphone. These signals are then temporally aligned, segmented into distinct vibration phases, and processed to extract behavioral features. A multi-modal classification model fuses the score vectors from each modality and verifies the user's identity, enabling efficient authentication using only built-in smartphone hardware.

We develop a vibration-boosted touchscreen sensing application and design a hybrid deep-learning framework to address the challenges of weak signals, low refresh rate, and surface dependency. Each sensing modality is independently processed using a dedicated Convolutional Neural Network (CNN) model optimized to capture temporal patterns and discriminative features unique to the sensors' domain. The resulting intermediate score vectors are then combined using soft voting to extract and combine discriminative features across sensing channels. This architecture mitigates the impact of noisy or degraded signals in any modality and strengthens the system's ability to capture user-specific behavioral signatures. As a result, our approach significantly improves authentication accuracy and enhances robustness against zero-effort and impersonation attacks.

We summarize our main contributions as follows:

- We propose the first authentication system leveraging standard touchscreens to extract fingerprint-related behaviors without relying on dedicated biometric hardware.
- We develop a vibration-boosted touchscreen sensing application to address low touchscreen sampling rates, limited signal strength, and environmental variability.
- Our system introduces a hybrid deep learning architecture in which each sensing modality is processed independently using a CNN model, producing intermediate score vectors fused through the decision-level fusion strategy.
- We perform real-user experiments to evaluate the effectiveness of our system and demonstrate high authentication accuracy, along with strong resilience against attacks.

II. Related Work

Mobile authentication has long relied on knowledge-based PINs or passwords. Although easy to deploy, they are vulnerable to phishing, guessing, and peeping attacks, which also interrupt the usage process and induce unsafe behaviors such as cross-site reuse [5], [6]. Static biometrics (fingerprints, faces, and irises) improve usability and security but require dedicated sensors and are easy to forge once leaked [7]. Behavioral biometrics enable continuous, non-touch authentication through various interaction modes, including typing, gait, and touch. Many methods still rely on a single modality or require repeated input, limiting real-world usability and their widespread adoption [8], [9], [10].

Touchscreens are one of the most frequently used components in smartphones. Beyond their primary function of capturing touch coordinates, smartphone touchscreens possess a rich sensing capability that is often utilized in existing authentication systems. Most existing systems use touchscreens solely to detect tap location, timing, or swipe paths [11], [12]. Touchscreen interactions naturally involve subtle mechanical vibrations and response behaviors that can be sensed through other built-in components, such as microphones and IMUs. This presents an opportunity to leverage these inherent device capabilities for more seamless and secure authentication without requiring additional dedicated hardware. However, to our knowledge, no prior work has leveraged the touchscreen with active vibration feedback to capture fingerprint-related behavioral traits during normal tap-based interactions. This opens up a new direction for sensor-free, multi-modal authentication that is seamlessly integrated into common smartphone use.

Existing research has explored the use of vibration signals across various domains, including localization, health monitoring, and user authentication, by leveraging the interaction

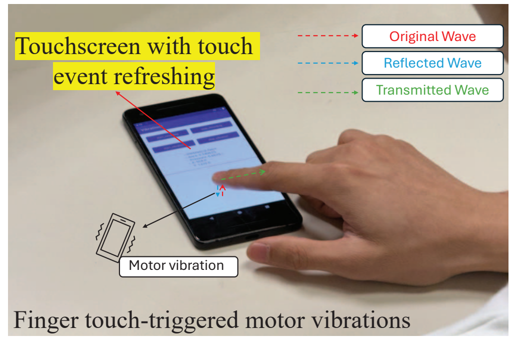


Fig. 1: Motor-induced vibration signal sensing during touchscreen interaction.

between user input and the device's physical properties [13]. Passive methods, such as TapPrint [14], use wearable inertial sensors to collect tapping vibrations for PIN/single-click authentication or use accelerometers to sense heartbeats when pressing the chest. These methods have proven to be feasible but are limited by weak signals, desktop/material differences, and insufficient generalization due to the use of inertial sensors alone. Introducing pressure sensors can enhance sampling but requires additional hardware [15], [16].

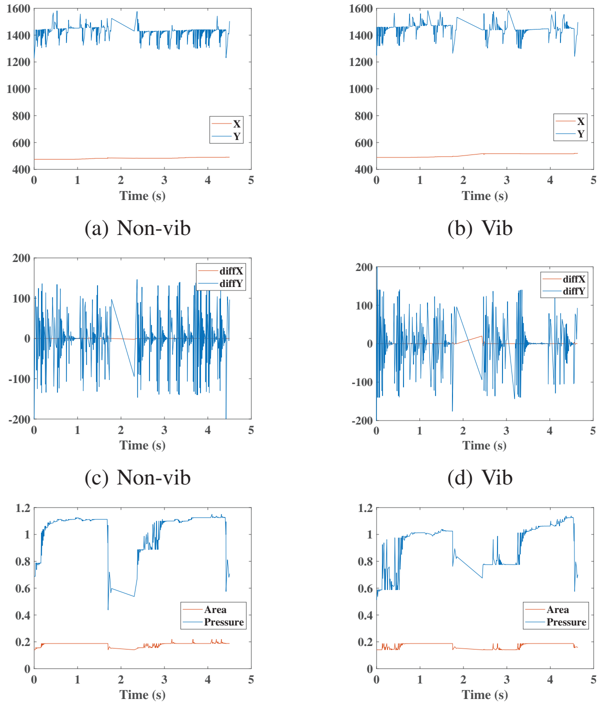
Active vibration methods generate controllable signals through devices or external motors to improve consistency. VibWrite [17], Velody [18], etc. require external installation, and VibID [19] relies on wearables, making it difficult to deploy universally on mobile phones. The most relevant TouchPass [20] drives the mobile phone vibration motor when tapping, analyzes the transient and steady-state segments of the IMU, and proposes a behavior-independent classifier and an anti-spoofing strategy based on a twin network and knowledge distillation, but does not integrate the rich behavior and acoustic response information contained in the touchscreen and microphone. Acoustic solutions (such as PCR [21]) use ultrasound to actively detect identity, but are still sensitive to hand gestures and environmental noise.

Our work addresses challenges in limited touchscreen responsiveness, restricted surfaces, and weak signal. To the best of our knowledge, this is the first work to utilize smartphone touchscreen vibration interactions as a behavioral fingerprinting signal in a multi-modal authentication framework.

III. Background and System Architecture

A. Vibration as Stimuli Signals on Smartphone

Modern smartphones employ miniature motors to generate haptic feedback during user interactions. When a user taps the screen, as shown in Figure 1, the smartphone activates its internal motor, propagating vibrations through the chassis, the user's finger, and back to the touchscreen surface. These vibrations reflect and travel across the finger–smartphone interaction and any external contact surface (e.g., a tabletop or a user's hand), creating distinctive signals detectable by multiple onboard sensors. Our system leverages a vibration-boosted, software-enhanced touchscreen that records touch events at



(a) Non-vib (b) Vib (c) Non-vib (d) Vib (e) Non-vib (f) Vib

Fig. 2: Touchscreen response to static tap with and without vibration stimulus.

elevated sampling rates. Motivated by the underexplored potential of active vibration as a stimulus signal, we conducted signal characterization studies using smartphone built-in sensors. Our method implements a tap-triggered vibration scheme similar to typical typing feedback. This approach establishes a clear sequence in which finger contact initiates motor vibration with subsequent interaction, enabling a reliable and consistent analysis of the characteristics of the response to vibration.

Figure 2 compares touchscreen data under vibrating (vib) and non-vibrating (non-vib) motor conditions. Figure 2a and 2b show that absolute touch positions remain largely stable between the two conditions. However, the differential position traces in 2c and 2d exhibit notable increases during vibration, indicating enhanced positional jitter from vibration effects. The static measurements of touch area and pressure do not seem to be affected by vibration, but there is a clear initial area and pressure confirmation stage, and then stabilizes under vibration, as shown in Figure 2e. These findings indicate that vibrations impact touch stability through subtle positional movements and variations in contact area. Such observations provide valuable features for analyzing user behavior.

Notably, touchscreen responses showed less distinct vibration patterns compared to inertial sensors, and microphones captured signals. That explains why many existing works use inertial sensors and audio methods to authenticate touch-based or vibration. However, inertial sensors experience drift and require calibration, while audios suffer from ambient noise interference. This observation highlights a significant challenge: despite their ubiquity, touchscreen sensors remain underutilized beyond basic positional tracking and require specialized data processing to reveal their full biometric potential.

B. Challenges in Capturing Vibration from Touchscreen

One of the core challenges in using a standard touchscreen for sensing vibrations is the limited responsiveness of touchscreen sensors to high-frequency vibration changes. When a user places a finger on the screen and maintains contact, the touchscreen driver and Android's event pipeline typically suppress motion events. By default, Android only issues fresh motion events (e.g., ACTION MOVE) when there is a significant change in touch location. As a result, the system treats the contact as static and stops delivering new events. This prevents the system from capturing subtle fluctuations caused by vibration, making it challenging to extract meaningful data during continuous contact. To overcome this, we modified the event pipeline to forcibly refresh touch inputs, enabling the touchscreen to sense vibration-induced changes continuously at a higher effective sampling rate.

Another challenge arises from variability in how vibration signals propagate through the phone and its environment. The physical properties of the phone's screen, chassis, and the surface it rests on (e.g., plywood, wood, or paper) significantly influence vibration attenuation, propagation path, coupling, and frequency response. For example, the acoustic impedance and structural stiffness of a wooden desktop and a book cover are significantly different, resulting in distinct attenuation speeds and amplitude changes. These differences alter how vibrations are transmitted through the phone-finger interface and can introduce differences in sensor readings across usage scenarios. Our system is designed to be robust to such material-induced variability to ensure reliable performance.

In addition, adjustable settings such as vibration strength or duration may affect the balance between user comfort and signal detectability. While long and loud vibrations tend to be avoided for user experience, they contribute to the signal fidelity available to sensors. Our system accounts for this trade-off by tuning vibration parameters to minimize user burden without compromising sensing accuracy.

C. Vibration-boosted Refreshing Touchscreen Sensing Design

1) Motor-Induced Vibration Triggered by Tap Events: Each time the user taps the touchscreen, we trigger a short-duration haptic vibration through the phone's built-in motor. This haptic feedback is used in virtual keyboards and other UI interactions. This stimulus vibration acts as a consistent, repeatable, active signal. During the tap and vibration phase, recorded data will capture minute user-specific differences in finger position, pressure, and area. Meanwhile, the inertial sensor and audio responses are recorded as an assistant for further data processing and authentication fusion.

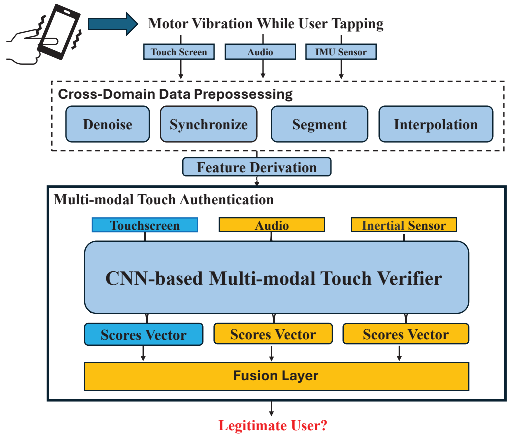


Fig. 3: The flow of vibration-boosted multi-modal authentication system.

2) Three-Stage Vibration Sensing Window: Based on signal analysis, each stimulus vibration interval exhibits a rising edge, an active vibration stage, and a descending edge. We anticipate that interactions between the user's finger and the screen during the active vibration will yield increasingly meaningful patterns. Therefore, we divide the vibration interval into three distinct stages with a time ratio, e.g., 1:3:1. Initially, the **pre-vibration stage** captures the finger's initial contact and stabilization. Subsequently, the **active vibration stage** records primary behavioral signals as vibrations propagate through the user's finger. Lastly, the **post-vibration stage** captures residual movements following motor deactivation and finger-lift dynamics, enriching the authentication data. This structured approach ensures valuable data collection despite touchscreen sensor sampling limitations. It can be further optimized by adjusting the vibration duration or combining multiple vibration intervals to obtain richer information.

D. Multi-modal Authentication System Design

Our work aims to enable fingerprint-like user authentication using only built-in smartphone sensors, eliminating the need for dedicated biometric hardware. To achieve this goal and address the above-mentioned challenges, we develop a multi-modal authentication system that captures user-specific responses to actively generated vibration signals via the touchscreen, microphone, and inertial sensor. While prior work has leveraged audio and inertial sensor signals, our design specifically maximizes the touchscreen sensor's underutilized potential and fuses multi-sensor data to enhance robustness against impersonation and environmental variation.

1) System Flow: The overall system architecture is illustrated in Figure 3. It operates on three different built-in smartphone sensors. Our system consists of three core parts: cross-domain data preprocessing, feature derivation, and multi-modal classification. The Cross-Domain Data Preprocessing first removes out-of-band noise from each sensing channel and temporally aligns them based on the inertial sensor and audio. Next, the Feature Derivation process extracts informative features from each sensor to characterize the vibration response's static and dynamic properties. We compute the difference-based and statistical features of the touchscreen data on position, pressure, and contact area. We extract a range of statistical and frequency-domain features from inertial and audio. All features are computed over sliding windows to preserve temporal resolution, enabling the time-aligned, modality-specific feature sequences for future tasks. Our final stage is Multi-Modal touch authentication. It combines three sensing modalities through a hybrid learning framework. Each modality is first processed by a dedicated branch of a deep learning model, producing intermediate score vectors. These vectors are fused via a fusion layer that learns cross-modal correlations while preserving modality-specific discriminative features. The system outputs a final classification score indicating whether the input matches the enrolled fingerprint profile. This design captures the time-series dynamics and cross-channel correlations necessary for authentication. It enables robustness to noise or signal degradation in any single modality and enhances resilience to impersonation attacks.

IV. Method Design

A. Motor-induced Vibration Stimuli Sensing

Our system uses active stimulus vibration as a consistent sensing signal and captures user-specified responses. This vibration propagates through smartphone-finger interaction, imprinting a unique response pattern on all three sensing channels, including the touchscreen, microphone, and inertial sensor. We configure a vibration stimulus interval with a specific duration (e.g., 300 ms) active vibration period to capture sufficient user-specific responses during a touch event. Specifically, when the user taps the screen, the vibration starts after a short delay (e.g., 100 ms), and the user's finger remains on the screen until the vibration completes, at which point the finger is naturally lifted. The entire process takes less than 500 ms. This design enables us to segment each interaction into three distinct sensing stages. The lengths of delay and vibration are selected to accommodate the low sampling rate of touchscreen sensors, ensuring that sufficient data points can be collected.

B. Cross-Domain Data Processing

Collecting data from the touchscreen, inertial sensors, and microphone, we preprocess the three-channel data as follows:

Denoising. A 200 Hz low-pass Butterworth filter suppresses high-frequency noise in the inertial and audio signals, while a moving-mean filter removes outliers in the touchscreen.

Synchronizing & Segmentation. We align audio and inertial data by cross-correlating the filtered audio with the accelerometer's z-axis; the peak lag reveals their offset, and we zero-pad the earlier stream for alignment. Short-time energy analysis then identifies vibration onset and offset in the inertial signal. Because each vibration originates from a screen tap, we refine these boundaries using the isVibrating flag to indicate when vibration starts and ends. We segment them to include the touchscreen's pre-vibration and post-vibration edges. This ensures that all three modalities cover the complete vibration cycle: rising, active vibration, and decay.

C. Feature Derivation

Our feature set is designed to capture subtle variations in touchscreen, inertial sensor, and audio signals during a user's finger interaction with the phone screen under active vibration. We derive 108 features across these three domains, forming separate feature matrices as input to the learning-based model. Specifically, we extract 15 features from touchscreen data, 21 from the inertial sensor, and 42 from audio. We apply a sliding-window approach to each channel, creating a multi-dimensional feature time series by concatenating over time.

Touchscreen. We extract 15 statistical features for each *x*, *y*, *diff_x*, *diff_y*, *area* and *pressure*. These features are computed within short overlapping windows to capture temporal dynamics, ensuring that transient variations in touchscreen interactions are preserved.

Inertial Sensor. Accelerometer and gyroscope readings capture subtle user touch motions and smartphone vibration movements. We exploit these variations to differentiate among interaction patterns, extracting 59 representative statistical features and frequency domain features across the inertial sensor.

Audio. To reveal acoustic patterns associated with tapping and vibration events, we compute Mel-Frequency Cepstral Coefficients, along with their first and second-order derivatives. These 42 features represent the spectral shape and its temporal evolution, enabling robust audio-based vibration analysis.

D. Multimodal Touch Fusion Verifier

We develop a CNN-based authentication system that performs multi-modal classification to identify different users based on their interaction patterns. Also, our system already involves user's behavior inconsistency (e.g., touch pressure variation), and each interaction requires the user to press the touchscreen with naturally varying gestures. We initially train three separate models (TouchCNN, InertialCNN, and AudioCNN) using the same architectural pattern but with different input channels to process each modality independently. Specifically, each modality consists of a stack of 1D convolutions (with batch norm + dropout and ReLU), followed by a pooling and a dropout layer, then a couple of dense layers and an output of the probabilities core. We use the Adam optimizer with gradient clipping for training to prevent gradient explosion and cross-entropy as a loss function. Additionally, we implement early stopping with a patience parameter to halt training when the validation accuracy stops improving, which helps prevent overfitting while ensuring optimal performance.

After that, we implement a late fusion strategy to get better performance. Each model independently processes its respective input data and generates probability scores. These scores are then integrated using a soft voting approach, where weights are determined based on the reliability and accuracy of each modality in different environmental conditions. Each model will learn modality-specific cues. Separate training for each modality model allows each to specialize without interference. Each model is an independent classifier as long as it is reasonably accurate and independent. Combining them can significantly boost overall accuracy. Even if one modality may perform weakly, the combined performance is still much better. In addition, we can update the weak-mode model without affecting the others' performance.

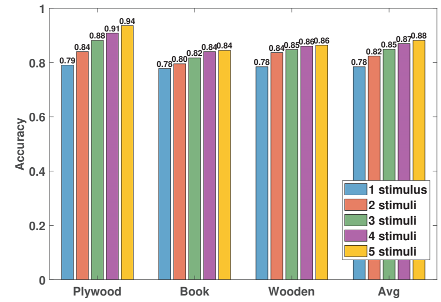


Fig. 4: Multiple vibration stimuli result on different surfaces.

V. Performance Evaluation

A. Experiment Setup

To assess the effectiveness of our proposed authentication system, we implemented an Android prototype that concurrently captures touchscreen, inertial, and audio data for every vibration event triggered by a single tap—mirroring the one-tap experience of fingerprint unlock. Sixteen participants used a Google Pixel 2 (Android 10) across multiple sessions. To incorporate natural behavioral variability (e.g., changes in touch pressure), each participant performed multiple taps with different pressures. We also examined surface effects by placing the phone on plywood, solid wood, and a cushioned book, collecting 40 trials per surface. The dataset was evenly divided into training and testing sets, and for each participant, we trained a binary classifier to distinguish their taps from those of all other users. The study was approved by the university's IRB, and all participants provided informed consent.

B. Authentication Performance

1) Touchscreen-only Performance: Based on the three-stage touch-vibration stimulation sensing design discussed in Section III-C, we study two interaction schemes: single-vibration and multi-vibration authentication.

Single vibration stimulus. For a single stimulus interval, the system starts the phone's linear motor 100 ms after the user touches the screen with one finger and maintains a constant amplitude vibration for 300 ms; the single interaction process is controlled within 500 ms, similar to commodity smartphones' fingerprint sensing time. This minimal stimulus yields virtually identical accuracies on all three surfaces, as presented in the blue bars in Figure 4.

Multi-vibration stimuli. We explore CNN score-based integration of multiple vibration stimuli to improve touch-

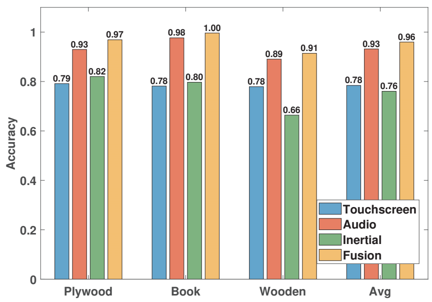


Fig. 5: Multi-modal fusion result on different surfaces.

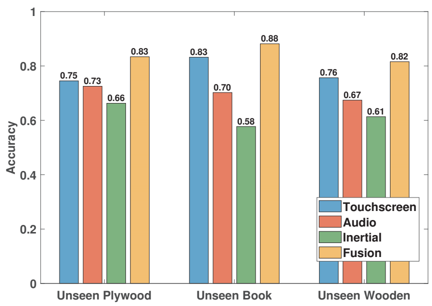


Fig. 6: Placing the phone on different unseen surfaces.

screen sensing robustness. The total duration after integrating 5-vibration stimuli is still within 2.5 s. As shown in Figure 4, two vibration stimuli already deliver a significant performance boost. Notably, there is a 15% performance gain on the plywood surface after integrating five stimuli, while the average accuracy across all surfaces climbs from 78% to 88%. Considering the trade-off between speed and robustness, three stimuli (around 1.3 s) strikes an attractive balance for daily authentication, while the 4- and 5-stimuli mode can be reserved for high-security purposes where the extra authentication time is acceptable.

2) Multi-modal Authentication: This study uses a single vibration stimulus as the baseline to focus on multimodal factors that improve performance. As shown in Figure 5, the touchscreen is a stable channel and performs well. The audio has high individual accuracy compared with the inertial, but also has varying performance on different surfaces. Across all surfaces, the multi-modal model outperforms every single-modal model, even on the wooden surface that decreases acoustic and inertial performance, with up to 31% increase. Additionally, we present that the multi-modal model which integrates the touchscreen channel has robust performance on unseen surfaces.

3) Unseen-Surfaces Comparison: We evaluate how well the proposed models incorporate with previously unseen surfaces by training on data collected from two surfaces and testing on the third surface (e.g., the model trained with wood and book data is tested on an unseen plywood surface). The three

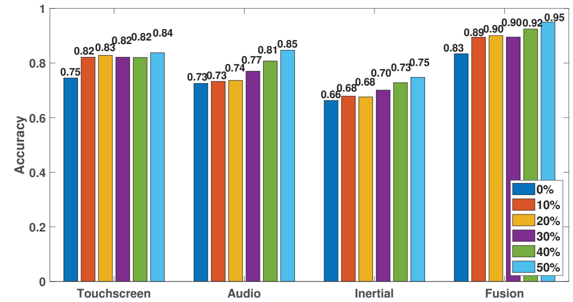


Fig. 7: Few-shot learning to incorporate on unseen surfaces.

combinations of train-test are reported in Figure 6. It reveals a sharp drop for audio and inertial (down to 58 %), whereas the touchscreen remains stable. Decision-level fusion still leads (82 –88 %), but its edge shrinks when other channels falter.

We introduce a few-shot learning step to minimize this gap: each new user contributes a small fraction of samples from the new surface while keeping the enrollment burden low. The effect of adding 0% to 50% of the targeted unseen surface data is summarized in Figure 7. A few-shot calibration alleviates this: adding only 10 % samples from the new surface lifts fusion accuracy to 89 %, and 50 % pushes it to 95 %. Thus, a generic model plus a handful of enrollment taps can quickly regain near-surface performance, making the system practical across diverse contact materials.

C. Impact Factors

We explore how various system parameters, including vibration length, amplitude, and phone placement surfaces, affect authentication performance. These experiments help assess trade-offs between usability, robustness, and sensing fidelity.

1) Vibration Length: We evaluated the impact of vibration duration on authentication accuracy by testing five different vibration lengths through 100, 300, 500, 750, and 1000 ms. Figure 8 presents the results for each modality and their fused decision outcome. Accuracy peaks at 300 ms, where each modality performs best and fusion reaches 100 %. Even at the shortest 100 ms, fusion still achieves 94 %, showing strong cross-modal robustness. For longer taps, audio remains reliable, touchscreen accuracy improves, and inertial performance drifts slightly, yet the fused model stays above 91 % across all lengths. Hence, we set 300 ms as the default duration.

2) Vibration Amplitude: We further studied the effect of motor amplitude on authentication performance by evaluating five levels: 50, 100, 150, 200, and 255(maximum). Figure 9 presents that the performance of each modality improves with increasing amplitude. Low amplitude (50) yields weak inertial signals and pulls fused accuracy down to 88 %, even though touchscreen and audio remain 90 %. As amplitude rises, all channels' performance improve and the fused model reaches 100 % at the maximum setting. So we adopt the maximum amplitude for reliable cross-modal authentication.

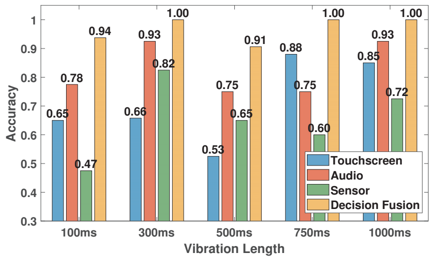


Fig. 8: Different vibration lengths.

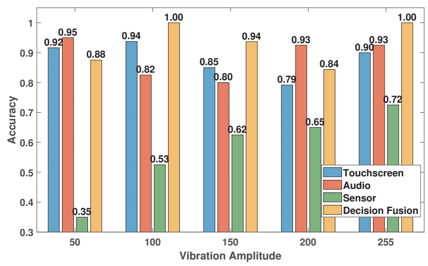


Fig. 9: Different vibration amplitudes.

VI. Conclusion

In this work, we introduce a vibration-based, multi-modal smartphone authenticator that repurposes the touchscreen, microphone, and inertial sensor already present in commodity phones. A brief motor pulse during a normal tap excites distinctive mechanical and acoustic responses; three lightweight CNNs independently extract features from each modality, and a decision fusion module produces the result. This architecture eliminates the need for dedicated sensors like fingerprint or face hardware, overcomes the limited resolution of the touchscreen, and is a hedge against single-sensor noise, surface variability, and spoofing attempts. Through real-world experiments with 16 participants, our system achieved up to 96% authentication accuracy. We further demonstrated its resilience and practical robustness for different surface comparisons in adjusting multi-vibration stimuli patterns and incorporating with multi-channel signals, and analyzed its performance across different vibration settings and usage scenarios. Our results suggest active vibration-stimulus behavioural biometrics can enable practical, secure, cost-effective authentication solutions using only built-in smartphone sensors. Our work opens new directions for sensor-free mobile security systems that work passively in the background of natural user interactions.

Acknowledgment

This work is partially supported by NSF CNS-2450046 and CNS-2440238.

References

[1] A. Sherif. (2025) Global smartphone penetration rate as share of population from 2016 to 2024. [Online]. Available: https://www.statista.com/statistics/203734/global-smartphone-penetration-per-capita-since-2005/

[2] S. Trewin, C. Swart, L. Koved, J. Martino, K. Singh, and S. Ben-David, "Biometric authentication on a mobile device: a study of user effort, error and task disruption," in *Proceedings of the 28th annual computer security applications conference*, 2012, pp. 159–168.

[3] D. Oberhaus. (2017) iphone x's face id can be fooled with a 3d-printed mask. [Online]. Available: https://www.vice.com/en/article/iphone-x-face-id-mask-spoof/

[4] G. Kambourakis, D. Damopoulos, D. Papamartzivanos, and E. Pavlidakis, "Introducing touchstroke: keystroke-based authentication system for smartphones," *Security and Communication Networks*, vol. 9, no. 6, pp. 542–554, 2016.

[5] M. H. Barkadehi, M. Nilashi, O. Ibrahim, A. Z. Fardi, and S. Samad, "Authentication systems: A literature review and classification," *Telematics and Informatics*, vol. 35, no. 5, pp. 1491–1511, 2018.

[6] A. Das, J. Bonneau, M. Caesar, N. Borisov, and X. Wang, "The tangled web of password reuse." in *NDSS*, vol. 14, no. 2014, 2014, pp. 23–26.

[7] B. Ammour, L. Boubchir, T. Bouden, and M. Ramdani, "Face–iris multimodal biometric identification system," *Electronics*, vol. 9, no. 1, p. 85, 2020.

[8] A. Mahfouz, T. M. Mahmoud, and A. S. Eldin, "A survey on behavioral biometric authentication on smartphones," *Journal of information security and applications*, vol. 37, pp. 28–37, 2017.

[9] Y. Liang, S. Samtani, B. Guo, and Z. Yu, "Behavioral biometrics for continuous authentication in the internet-of-things era: An artificial intelligence perspective," *IEEE Internet of Things Journal*, vol. 7, no. 9, pp. 9128–9143, 2020.

[10] A. Alzubaidi and J. Kalita, "Authentication of smartphone users using behavioral biometrics," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1998–2026, 2016.

[11] C. Bo, L. Zhang, X.-Y. Li, Q. Huang, and Y. Wang, "Silentsense: silent user identification via touch and movement behavioral biometrics," in *Proceedings of the 19th annual international conference on Mobile computing & networking*, 2013, pp. 187–190.

[12] Y. Song, Z. Cai, and Z.-L. Zhang, "Multi-touch authentication using hand geometry and behavioral information," in *2017 IEEE symposium on security and privacy (SP)*. IEEE, 2017, pp. 357–372.

[13] Y. Huang and K. Wu, "Vibration-based pervasive computing and intelligent sensing," *CCF Transactions on Pervasive Computing and Interaction*, vol. 2, no. 4, pp. 219–239, 2020.

[14] W. Chen, L. Chen, Y. Huang, X. Zhang, L. Wang, R. Ruby, and K. Wu, "Taprint: Secure text input for commodity smart wristbands," in *The 25th Annual International Conference on Mobile Computing and Networking*, 2019, pp. 1–16.

[15] L. Wang, K. Huang, K. Sun, W. Wang, C. Tian, L. Xie, and Q. Gu, "Unlock with your heart: Heartbeat-based authentication on commercial mobile phones," *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies*, vol. 2, no. 3, pp. 1–22, 2018.

[16] H. Cao, H. Jiang, K. Yang, S. Chen, W. Wu, J. Liu, and S. Dustdar, "Data-augmentation-enabled continuous user authentication via passive vibration response," *IEEE Internet of Things Journal*, vol. 10, no. 16, pp. 14 137–14 151, 2023.

[17] J. Liu, C. Wang, Y. Chen, and N. Saxena, "Vibwrite: Towards finger-input authentication on ubiquitous surfaces via physical vibration," in *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 2017, pp. 73–87.

[18] J. Li, K. Fawaz, and Y. Kim, "Velody: Nonlinear vibration challenge-response for resilient user authentication," in *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, 2019, pp. 1201–1213.

[19] L. Yang, W. Wang, and Q. Zhang, "Vibid: User identification through bio-vibrometry," in *2016 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. IEEE, 2016, pp. 1–12.

[20] X. Xu, J. Yu, Y. Chen, Q. Hua, Y. Zhu, Y.-C. Chen, and M. Li, "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*, 2020, pp. 1–13.

[21] L. Huang and C. Wang, "Pcr-auth: Solving authentication puzzle challenge with encoded palm contact response," in *2022 IEEE Symposium on Security and Privacy (SP)*. IEEE, 2022, pp. 1034–1048.