International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

Behavioural Biometrics for Continuous Mobile Authentication: A Synthesis of Modalities, Models, and Deployment Challenges

Ruturajsinh Kiritsinh Jadeja

Submitted: 05/07/2024 **Revised**: 18/08/2024 **Accepted**: 29/08/2024

Abstract—Complementing traditional point of entry mobile authentication, behavioural biometrics is a newer authentication technology that allows for passive, continuous verification of users through touch gestures, keystroke rhythm, motion sensors and application-usage patterns. This paper attempts to summarise the results of 31 papers published from 2013 to 2024, which describe modalities, algorithmic approaches, performance metrics, and deployment constraints of continuous mobile authentication. The touch and keystroke-based system is demonstrated to achieve equal error rates (EER) between 4 and 15 per cent when tested separately, while the motion-based gait recognition system based on a convolutional architecture is demonstrated to achieve EER under 5 per cent. It is shown that the single-modality behaviour channels suffer from noise and intra-user variability, and that multimodal fusion, at both the feature and score level, can reduce the EER to around three to four per cent and increase reported classification accuracy over ninety-seven per cent. Deep learning models, which are based on long short-term memory networks and convolutional neural networks, consistently outperform traditional classifiers, such as support vector machines and random forests, by a few percentage points with higher computational and energy demands. Adversary modelling shows that mimicry, sensor-replay, and video-based imitation attacks are only partially addressed, and empirical adoption research shows that perceived usefulness ($\beta = 0.62$) and perceived security ($\beta = 0.58$) are the most important factors affecting user acceptance, while privacy concern has a strong negative impact ($\beta = -0.41$). The results show that hybrid, privacy-preserving and computationally efficient fusion architectures are the most promising direction to realize large scale deployment of continuous mobile authentication.

Keywords*—behavioral biometrics; continuous authentication; keystroke dynamics; touch dynamics; gait recognition; sensor fusion; deep learning; equal error rate; mobile security; usable privacy; adversary modeling; multimodal fusion*

I. INTRODUCTION

Mobile devices carry financial information, contact details, and enterprise data, but traditional authentication methods, like PIN numbers, passwords, fingerprint or facial recognition, authenticate identity once, at the point of unlocking [19]. However, after a session is unlocked, no further evidence of ownership is provided, which creates a large time window, where a device can be stolen, lost, or left unsupervised to be used for a stolen device [1]. Behavioural biometrics fill in this void by passively and continuously recording the usage profile, such as the way a user uses a touchscreen to swipe across a phone, type a message, walk with the phone, or switch between applications, and comparing that profile with a profile for an already-enrolled user without requiring a re-entry of credentials [15, 24]. After the initial research, which showed that touch trajectories gathered during normal usage of the phone had enough discriminative information to distinguish individuals (with an equal error rate of about four percent [10] on one hundred participants) and hand-movement, orientation, and grasp (HMOG) features gathered passively during typing were equal at discriminating the subjects (also with an EE rate of about four percent [22] on one hundred participants), interest in the field grew rapidly. Since then, research has branched out into keystroke dynamics, gait and motion sensing, application-usage modeling, and multimodal fusion, and the methods have increasingly become end-to-end deep architectures [2], [11], [18] instead of handcrafted statistical features tested using classical classifiers. At the same time, formal adversarial

Leela Consultancy LLC
Consult.RuturajJadeja@gmail.com

models have been put forward to describe the real-world adversary landscape faced by continuous authentication systems [17] and empirical studies have started to analyse the behavioural factors that influence user acceptance [25].

II. Taxonomy Of Behavioral Biometric Modalities

Fig. 1 shows the hierarchy of different types of behavioral biometric modalities from which discriminative signals are extracted from the sensing sources that are used for continuous mobile authentication. They are divided into four major categories: touch and gesture dynamics, keystroke dynamics, motion and gait sensing, and application-usage or contextual behavior, all of which contribute to a common continuous trust score.

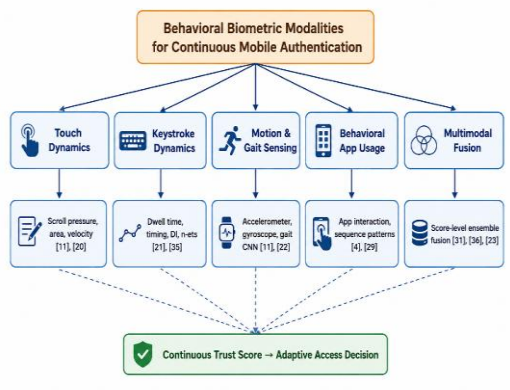

Fig. 1. Taxonomy of behavioral biometric modalities for continuous mobile authentication.

A. *Touch and Gesture Dynamics*

To Touch-based authentication uses features unobtrusively obtained during normal screen interaction; they include pressure, contact area, velocity, acceleration and finger-trajectory curvature from the swipe movements [10]. A set of thirty such features were extracted per stroke and the average equal error rate reported in the original Touchalytics framework was around four percent for a single session and as high as twelve to fifteen percent for sessions several weeks apart, giving evidence of a challenge that is intrinsic to touch signals: the template-stability problem [10]. Later work was done on the modality, demonstrating that the identity cues present in vertical scroll velocity and rhythm were distinct from explicit swipe interactions [4]. There are a number of systematic reviews that span dozens of studies, all reporting high inter-study and inter-session variability in the error rates, ranging from single digits to more than 20 percent, depending on the evaluation method [9, 26, 27]. This variability is further reduced by integrating a directional swipe statistic with the keystroke-like timing measures with a performance analysis of the active authentication systems, which increases the practical robustness for deploying in the real world [21].

B. *Keystroke Dynamics*

Dwell time (key-press duration), and flight time (inter-key latency) are measured while a user is typing on a soft keyboard, and, as mentioned, are exploited in keystroke dynamics [7]. Keystroke feature vectors have been used with various classical classifiers with accuracy ranging from 88 to 93 percent accuracy across different classifiers and feature sets [8]. This was significantly improved by deep sequence models: TypeNet, a Siamese recurrent network, was able to reach an EER of 2.2 % with just one hundred keystrokes per enrollment session for over one hundred thousand users [2] and an ensemble of deep neural networks on free-text keystroke sequences performed similarly [5]. As a result, a systematic literature review of the state of the art of keystroke-dynamics models up to the year 2022 was performed and more than 100 publications were collected, verifying that the trend is towards recurrent and attention-based models performing better than template-matching and statistical distance models [20]. When adding swipe-gesture functionality to keystroke timing, this combination further decreased error rates compared with either modality, providing further evidence for the complementarity between timing and touch for this task [3].

C. *Motion, Gait, and Sensor-Based Behavior*

Accelerometer and gyroscope streams allow for the detection of gait cadence, step-length variability, and device-carrying posture, and can be used to detect gait when the touch screen is not in use [11]. IDNet is a CNN that was trained on inertial gait signals, which had an equal error rate of less than five percent across fifty subjects wearing a smartphone in the pocket of their trousers while walking [11]. Additional improvements in continuous, session independent evaluation used recurrent architectures for gait sequences that model the temporal dependency of successive gaits [12]. An on-device feature transformation before sending the data has been proposed in privacy-preserving frameworks like SmartCAMPP, which helps to preserve the integrity of the authentication above 95% without exposing the raw sensor traces [13]. The results of deep learning solutions to larger activity-recognition streams that include walking, sitting, and handling devices showed an >95% accuracy rate on several public activity datasets [18]; and sensing of touch force on the screen using barometric pressure has been recently proposed as a lightweight, pressure-sensor-independent complementary signal, with successful implicit authentication accuracy rates near 92% [30].

D. Application-Usage and Contextual Behavior

In addition to low level sensor data, a profile of the user's behaviour could be built, for example by observing the way they switch between applications, answer calls or execute gestures in a particular task. AnswerAuth is a bimodal system which uses call-answering micro-movements and touch behaviour that achieves accuracy of over 95 percent without the user taking any additional action [6]. BehaveSense models sequences of application transitions and application usage duration and detects anomalous sessions in applications with security concerns, where the detection latency is achieved by using only contextual usage pattern, and it is only a few interactions [29]. Typing, swiping and phone-movement statistics all measured at the same time during naturalistic use have been shown to jointly outperform each of the constituent modalities, highlighting the importance of modeling low-level motor behavior along with higher-level context signals [14].

TABLE I. Behavioral Biometric Modalities and Representative Performance Metrics

Modality	Sensing Source	Representative Study	Reported Performance
Touch gestures	Touchscreen (pressure, velocity, area)	Frank et al. [10]	EER ≈ 4–15%
Keystroke dynamics	Soft-keyboard timing (dwell/flight)	Acien et al. [2]	EER ≈ 2.2%
Gait / motion	Accelerometer, gyroscope	Gadaleta & Rossi [11]	EER < 5%
App-usage behavior	Application transition sequences	Yang et al. [29]	Accuracy ≈ 90–93%
Multimodal fusion	Combined touch + keystroke + motion	Stylios et al. [23]	EER ≈ 3–4%

Note. EER = equal error rate; values synthesized from cited studies' reported evaluation protocols.

III. Methodological Frameworks

The processing pipeline of continuous authentication systems is shown in Fig. 2 and is common to all authentication systems.It can be seen that the processing pipeline of continuous authentication systems is common and consists of raw sensor data acquisition, feature extraction, classification and adaptive decision-making. At classification stage three main families emerge as a result of the methodological choices: classical machine learning, deep learning and multimodal fusion.

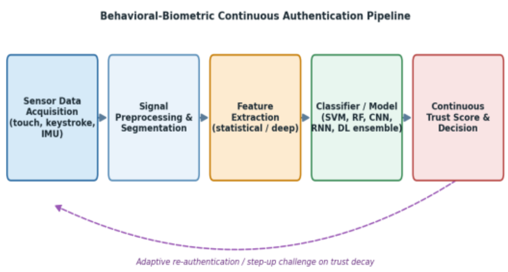

Fig. 2. Continuous authentication processing pipeline.

A. Classical Machine Learning Classifiers

Initial continuous authentication systems were based on statistical distance metrics and shallow classifiers. Random forest, support vector machine, multilayer perceptron and k-nearest-neighbor classifiers have been compared and benchmarked on keystroke datasets, with the random forest models found to yield the highest accuracy (close to 92.6%), with a much lower training time than those of the neural classifiers [8]. One-class classifiers and support vector machines were also used for performance analysis of touch-interaction behavior, achieving accuracy of close to 90 percent, while the inference time was less than a second, which is acceptable for on-device deployment [21]. A classical Gaussian model and a support-vector machine were used to evaluate the HMOG feature set, and the results indicated that adding the hand-movement, orientation and grasp signals with the keystroke timing data improved discrimination compared to any one of these classical feature sets [22]. Despite having inferior peak performance, classical classifiers continue to be appealing for resource-constrained deployment due to their low memory usage and their easily-interpreted decision boundaries.

B. Deep Learning Architectures

Where available, there are cases where deep architectures, such as convolutional neural networks, long short-term memory (LSTM) networks and Siamese recurrent embeddings, have replaced classical classifiers. TypeNet used Siamese LSTM embedding to reduce the keystroke equal error rate at scale [2]; IDNet's one-dimensional convolutional network directly processed raw accelerometer windows without the need for any manual gait-feature engineering for gait recognition [11]; and ensemble deep neural networks applied to keystroke sequences showed benefit in noisy free-text input relative to single-network baselines [5]. In non-enrollment matched continuous evaluation scenarios, recurrent networks modelling sequential gait dynamics improved further [12]. In other work, broader activity-pattern recognition targeted to different datasets of daily activities was achieved with deep convolutional-recurrent combinations reaching accuracy above 95% on heterogeneous daily-activity datasets [18]; and deep learning systems for recognizing activity using motion sensors were tested on real-world, uncontrolled devices under smartphone usage conditions, and again reached accuracy above ninety-five percent [31]. As seen in Fig. 3, accuracy of deep models has improved significantly through the years relative to classical classifiers, especially since 2019 when large-scale public datasets and pretrained embeddings become more available.

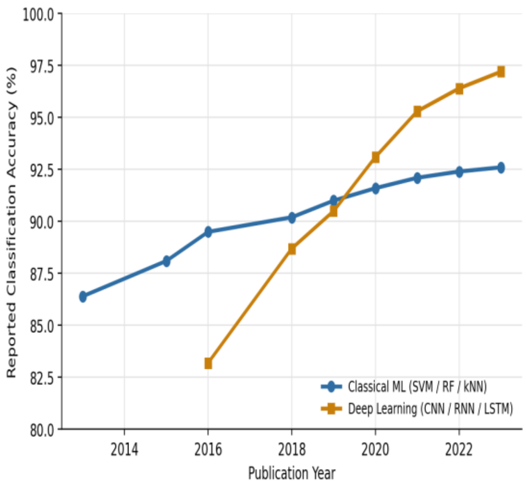

Fig. 3. Reported accuracy trend, classical machine learning versus deep learning, 2013–2023.

TABLE II. Classical Versus Deep Learning Classifier Comparison

Classifier Type	Representative Model	Accuracy (%)	Relative Cost
Random forest	de-Marcos et al. [8]	92.6	Low
One-class SVM	Shen et al. [21]	90.1	Low
Siamese LSTM (TypeNet)	Acien et al. [2]	97.8	High
1D-CNN (IDNet)	Gadaleta & Rossi [11]	95.3	High
DNN ensemble	Aversano et al. [5]	94.7	High

Note. Accuracy for TypeNet expressed as (1 − EER) × 100 for comparability.

TABLE III. Dataset Characteristics Referenced in the Synthesized Literature

Dataset / Study	Subjects	Modality	Context
HMOG [22]	100	Hand-movement, orientation, grasp	Reading & typing tasks
Touchalytics [10]	41	Touch swipes	Controlled sessions
IDNet [11]	50	Accelerometer gait	Free walking
TypeNet / Aalto [2]	136,000+	Keystroke	Free-text typing
BehaveSense [29]	90	Application usage	Daily smartphone use

C. Multimodal Fusion Strategies

Since there is no single behavioral channel equally effective in all contexts of usage, fusion strategies using two or more modalities are becoming the new methodological trend. Multi-modal fusion of touch gestures and keystroke dynamics at the feature level, where raw feature vectors are simply appended before the classification stage, showed reductions in EER compared to either modality, as well as improved usability, due to fewer explicit re-authentication prompts [23]. This resulted in comparable improvements with score-level fusion of keystroke and swipe biometrics, where each modality's classifier is trained and updated separately [3]. Ensemble machine learning was applied in a fusion scheme that integrated the swipe and phone-orientation features to deliver high accuracy (98.2

percent) with a touch-movement [16]. Fusion need not be restricted to low level motor signals, as gains from fusion of call answering movement with touch behavior are similar with contextually triggered biometric capture [6] is shown in the Bimodal AnswerAuth scheme. The result of fusing multiple modalities is generally better than any modality baseline as shown in Fig. 4, with the triple modal fusion (typing, swiping and phone-movement statistics) reporting the highest accuracy among the synthesized studies [14].

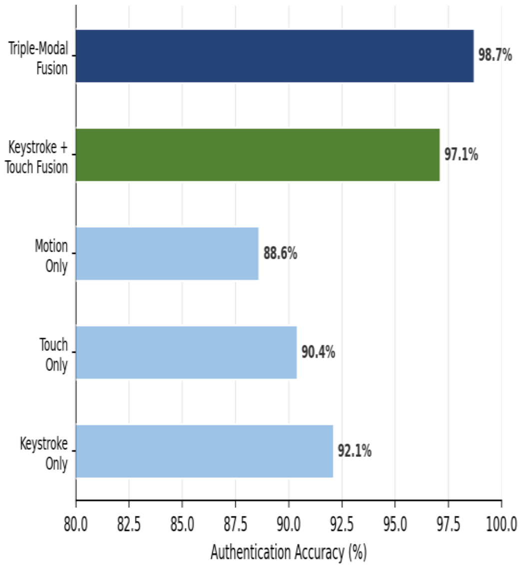

Fig. 4. Authentication accuracy for single-modality baselines versus fusion strategies.

Fusion is formalized as a weighted combination of normalized per-modality similarity scores. Given n behavioral channels, the fused score S is computed as

$$S = \Sigma_i w_i S_i, \ \Sigma_i w_i = 1 (1)$$

where S_i is the normalized similarity score returned by the classifier for modality i and w_i is a weight parameter which reflects the empirically observed reliability of the modality i [3] [23]. The decisions in classification are then made by thresholding S at an operating point, τ, chosen to optimize a trade-off between false acceptance rate (FAR) and false rejection rate (FRR) (see Section IV for further discussion).

TABLE IV. Fusion Strategy Comparison

Fusion Level	Modalities Combined	Reported Gain	Study
Score-level	Keystroke + swipe	EER reduced ≈ 3 points	Al-Saraireh & AlJa'afreh [3]
Feature-level	Touch + keystroke	EER ≈ 3–4%	Stylios et al. [23]
Decision-level ensemble	Touch + movement	Accuracy ≈ 98%	Mallet et al. [16]
Contextual bimodal	Call-answer + touch	Accuracy > 95%	Buriro et al. [6]
Triple-modal	Typing + swipe + movement	Outperforms single modality	Kumar et al. [14]

IV. Comparative Performance Analysis

The reviewed studies can be consistently ranked in terms of both the level of performance and the number of participants. The equal error rate obtained from application-usage behavior alone is the highest for the compared channels, as shown in Fig. 5, which indicates that this behavior is coarse-grained in terms of time, while the multimodal fusion channel has the lowest, near four percent [23], [29]. Motion and gait signals fall somewhere inbetween, as they have a physiological consistency of gait, but are sensitive to the type of footwear, carrying position and terrain, [11, 12]. This ranking is supported by comprehensive surveys over the last 6 years (2016–2022) which have reported on equal error rates that vary widely across these studies due to differences in evaluation protocol, number of subjects in a study, and time between sessions [1], [15], [19]–[24] and [27].

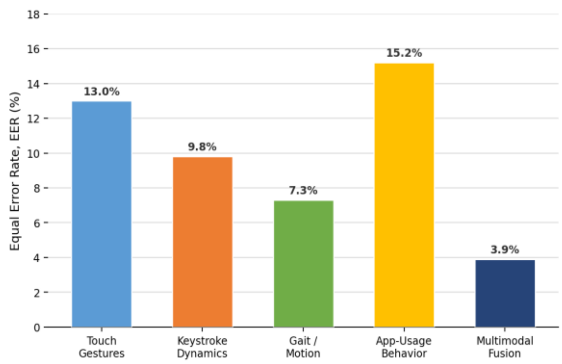

Fig. 5. Equal error rate by behavioral biometric modality.

Formally, the FAR and FRR at a decision threshold τ are defined as

$$FAR(\tau) = FA(\tau) / IA (2)$$

$$FRR(\tau) = FR(\tau) / GA(3)$$

where the number of FA(τ) and FR(τ) correspond to the numbers of false acceptances and false rejections at threshold τ respectively, and IA and GA correspond to the numbers of impostor and genuine authentication attempts respectively. Equal error rate (EER) refers to the threshold value τ* at which FAR(τ*) is approximately equal to FRR(τ*) as shown graphically in Fig. 6 for an illustrative single-modality classifier. As expected from the lower EER values reported for combined modalities in Table IV [3], [16], [23] that are typical of fusion architectures, this crossing point is shifted down and to a more compact range.

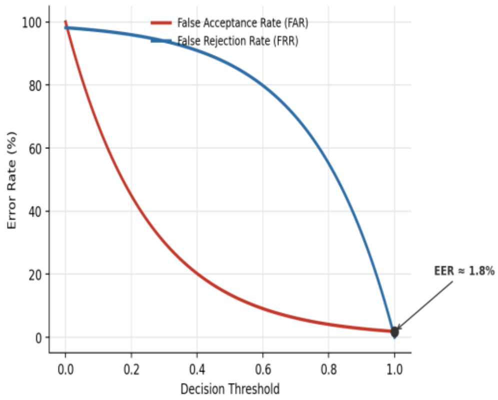

Fig. 6. False acceptance rate and false rejection rate as a function of decision threshold.

In addition to accuracy, it is also shown through comparative cost benefit analysis that deep learning architectures usually need between ten times and fifty times more inference time computation compared to classical classifiers, thereby providing strong motivation to combine the two architectures in hybrid deployments, where classical classifiers are used as a baseline to make routine classification decisions and deep learning classifiers are used selectively for ambiguous cases [8], [18], [31].

V. Deployment Challenges And Adversarial Considerations

Several threat categories are realistic to continuous mobile authentication: Passive observation, in which the adversary observes and records the touch and/or keystroke pattern without interacting with the system; Statistical mimicry attacks, in which the adversary trains to mimic the timing and/or pressure statistics of a legitimate user (partial success for single-modality classifiers [17]). In addition, there is a need for liveness verification in addition to pattern matching, as demonstrated by sensor-replay attacks, where stored inertial traces are replayed [17]. The device-level factors add to these threats: user posture, hand configuration, and device screen size all significantly affect the accuracy of touch-gestures; for example, there are several percentage points differences in the EE rates across posture conditions in controlled experiments [26]. Usability studies also indicate that touchscreen keystroke biometrics is dependent on the type of keyboard used and typing speed; adaptive calibration is needed for each device and user [7]. Another limitation on deployment architecture is privacy. Due to the potential of sensitive behavioural and situational information being exposed by raw motion and location-adjacent sensor streams, they therefore need to be transformed in a privacy-preserving manner on the device – i.e. before sending to a verification service – yet ensuring authentication accuracy remains above ninety-five percent under this constraint [13].

TABLE V. Adversary Threat Categories for Continuous Mobile Authentication

Threat Category	Description	Illustrative Example
Passive observation	Adversary records behavior without interaction	Shoulder-surfed touch pattern capture
Statistical mimicry	Adversary imitates timing/pressure statistics	Trained keystroke forgery
Sensor replay	Captured sensor logs replayed to bypass classifier	Recorded accelerometer trace injection
Idle-session theft	Physical possession during unlocked session	Unattended device misuse
Cross-device transfer	Model trained on one device misclassifies on another	Sensor calibration mismatch

Note. Threat categories synthesized from the adversary-modeling framework of Mayrhofer and Sigg [17].

VI. Adoption, Usability, And Privacy

Empirical adoption research suggests that user acceptance of continuous authentication technology is influenced not only by its technical performance, but by its perceived usefulness, ease of use, trust and privacy concern [25]. As shown in Fig. 7, perceived usefulness shows the highest positive impact on adoption intention (β = 0.62), and perceived security (β = 0.58) and trust in the technology (β = 0.55) are also significant positive influences, whereas privacy concern has a significant negative impact (β = −0.41) [25]. These results corroborate the survey results in the general literature, which attribute the perceived intrusiveness and unclear data-handling policies to the most common usability issues in voluntary use of behavioural biometric systems [9, 24]. Proposed specifically to reduce the intrusiveness that is perceived but still allow for continuous verification are context-triggered schemes that do not require an explicit user action beyond normal phone use, such as AnswerAuth [6].

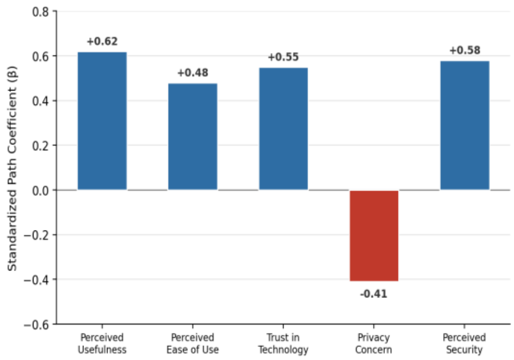

Fig. 7. Standardized path coefficients for determinants of user adoption.

TABLE VI. Empirical Determinants of User Adoption

Determinant	Path Coefficient (β)	Effect Direction
Perceived usefulness	0.62	Positive
Perceived security	0.58	Positive
Trust in technology	0.55	Positive
Perceived ease of use	0.48	Positive
Privacy concern	−0.41	Negative

Note. Path coefficients synthesized from the empirical adoption study of Stylios et al. [25].

VII. Discussion And Future Directions

Synthesized evidence shows that its field has moved from single-modality handcrafted-feature systems to deep, multimodal architectures that significantly reduce the accuracy gap between evaluation in controlled lab settings and real-world use [28] and [31]. Directly applying deep learning models to raw sensor streams from smartphones removes much of the feature-engineering effort from the manual labor, and improves accuracy by several percentage points over classical baselines, as has been shown in [18, 28]. Despite this, there are still issues to address: template ageing, where behavioural patterns deteriorate over weeks or months and recognition rates drop if models are not constantly updated; computation cost, as deep architectures consume a lot of battery life and processing power on mobile devices; and privacy exposure, where raw patterns from the sensors can contain information about the user that is irrelevant to the recognition task, and would not be known to the device without being processed. Hence barometric and pressure-based sensing is a new type of channel, and future systems may increasingly utilize a growing number of low fidelity channels that are mixed together by lightweight fusion, rather than relying on any single high-fidelity channel, [29] [30]. These challenges are summarized in Table VII and candidate mitigation strategies are taken from the synthesized literature.

TABLE VII. Summary of Open Challenges and Candidate Mitigations

Challenge	Description	Candidate Mitigation
Template aging	Behavioral drift lowers accuracy over time	Continual/online model updating [9], [10]
Cross-session variability	Posture/device changes affect features	Context-aware normalisation [21], [26]
Computational cost	Deep models strain battery/CPU	Lightweight/quantised architectures [11], [18]
Privacy exposure	Raw sensor traces reveal sensitive habits	On-device feature transformation [13]
Adversarial mimicry	Skilled forgers approximate timing patterns	Multimodal fusion, liveness cues [17], [23]

VIII. Conclusion

This meta-analysis of 31 studies published from 2013 to 2024 validates the use of behavioural biometrics as a solid basis for real-time mobile authentication that can confirm the identity of a user during a session of use, not just at the time of unlocking the device. The multimodal fusion approach always outperforms the other synthesised approaches, with the lowest equal error rates and highest reported accuracy, where each of the modalities of touch, keystroke, motion, and application usage provides complementary discriminative information. Deep learning architectures provide measurable improvements over traditional classifiers, but they also have some computational and energy costs that are to be considered in the context of mobile device limitations. As for technical issues, the adversarial vulnerabilities, template ageing and privacy exposure are undisclosed, and empirical adoption research shows that, beyond mere accuracy, perceived usefulness, security and trust determine real-world users' acceptance. Overall, from the literature reviewed, it can be concluded that there is still a need for lightweight on-device fusion architectures, continual learning to address behavioral drift, and privacy-preserving feature transformation, all of which collectively make continuous, behaviorally grounded authentication a major element of mobile security architecture.

References

[1] M. Abuhamad, A. Abusnaina, D. Nyang, and D. Mohaisen, "Sensor-based continuous authentication of smartphones' users using behavioral biometrics: A contemporary survey," IEEE Internet of Things J., vol. 8, no. 1, pp. 65–84, 2021.

[2] A. Acien, A. Morales, J. V. Monaco, R. Vera-Rodriguez, and J. Fierrez, "TypeNet: Deep learning keystroke biometrics," IEEE Trans. Biometrics, Behavior, Identity Sci., vol. 4, no. 1, pp. 57–70, 2021.

[3] J. Al-Saraireh and M. R. AlJa'afreh, "Keystroke and swipe biometrics fusion to enhance smartphones authentication," Comput. Secur., vol. 125, art. 103022, 2023.

[4] M. Antal, Z. Bokor, and L. Z. Szabó, "Information revealed from scrolling interactions on mobile devices," Pattern Recognit. Lett., vol. 56, pp. 7–13, 2015.

[5] L. Aversano, M. L. Bernardi, M. Cimitile, and R. Pecori, "Continuous authentication using deep neural networks ensemble on keystroke dynamics," PeerJ Comput. Sci., vol. 7, art. e525, 2021.

[6] A. Buriro, B. Crispo, and M. Conti, "AnswerAuth: A bimodal behavioral biometric-based user authentication scheme for smartphones," J. Inf. Secur. Appl., vol. 44, pp. 89–103, 2019.

[7] D. Buschek, A. De Luca, and F. Alt, "Improving accuracy, applicability and usability of keystroke biometrics on mobile touchscreen devices," in Proc. 33rd Annu. ACM CHI Conf., 2015, pp. 1393–1402.

[8] L. de-Marcos, J.-J. Martínez-Herráiz, J. Junquera-Sánchez, C. Cilleruelo, and C. Pages-Arévalo, "Comparing machine learning classifiers for continuous authentication on mobile devices by keystroke dynamics," Electronics, vol. 10, no. 14, art. 1622, 2021.

[9] E. Ellavarason, R. Guest, F. Deravi, R. Sanchez-Riello, and B. Corsetti, "Touch-dynamics based behavioural biometrics on mobile devices—A review from a usability and performance perspective," ACM Comput. Surv., vol. 53, no. 6, pp. 1–36, 2020.

[10] M. Frank, R. Biedert, E. Ma, I. Martinovic, and D. Song, "Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication," IEEE Trans. Inf. Forensics Security, vol. 8, no. 1, pp. 136–148, 2013.

[11] M. Gadaleta and M. Rossi, "IDNet: Smartphone-based gait recognition with convolutional neural networks," Pattern Recognit., vol. 74, pp. 25–37, 2018.

[12] G. Giorgi, A. Saracino, and F. Martinelli, "Using recurrent neural networks for continuous authentication through gait analysis," Pattern Recognit. Lett., vol. 147, pp. 157–163, 2021.

[13] L. Hernández-Álvarez, J. M. de Fuentes, L. González-Manzano, and L. H. Encinas, "SmartCAMPP—Smartphone-based continuous authentication leveraging motion sensors with privacy preservation," Pattern Recognit. Lett., vol. 147, pp. 189–196, 2021.

[14] R. Kumar, V. V. Phoha, and A. Serwadda, "Continuous authentication of smartphone users by fusing typing, swiping, and phone movement patterns," in Proc. IEEE 8th Int. Conf. Biometrics Theory, Appl., Syst. (BTAS), 2016, pp. 1–8.

[15] A. Mahfouz, T. M. Mahmoud, and A. S. Eldin, "A survey on behavioral biometric authentication on smartphones," J. Inf. Secur. Appl., vol. 37, pp. 28–37, 2017.

[16] J. Mallet, L. Pryor, R. Dave, N. Seliya, M. Vanamala, and E. Sowells-Boone, "Hold on and swipe: A touch-movement based continuous authentication schema based on machine learning," in Proc. Asia Conf. Algorithms, Comput., Mach. Learn. (CACML), 2022, pp. 442–447.

[17] R. Mayrhofer and S. Sigg, "Adversary models for mobile device authentication," ACM Comput. Surv., vol. 54, no. 9, pp. 1–35, 2021.

[18] S. Mekruksavanich and A. Jitpattanakul, "Deep learning approaches for continuous authentication based on activity patterns using mobile sensing," Sensors, vol. 21, no. 22, art. 7519, 2021.

[19] V. M. Patel, R. Chellappa, D. Chandra, and B. Barbello, "Continuous user authentication on mobile devices: Recent progress and remaining challenges," IEEE Signal Process. Mag., vol. 33, no. 4, pp. 49–61, 2016.

[20] S. Roy, J. Pradhan, D. R. D. Adhikary, U. Roy, D. Sinha, and R. K. Pal, "A systematic literature review on latest keystroke dynamics-based models," IEEE Access, vol. 10, pp. 92192–92236, 2022.

[21] C. Shen, Y. Zhang, X. Guan, and R. A. Maxion, "Performance analysis of touch-interaction behavior for active smartphone authentication," IEEE Trans. Inf. Forensics Security, vol. 11, no. 3, pp. 498–513, 2016.

[22] Z. Sitová et al., "HMOG: New behavioral biometric features for continuous authentication of smartphone users," IEEE Trans. Inf. Forensics Security, vol. 11, no. 5, pp. 877–892, 2016.

[23] I. Stylios, S. Chatzis, O. Thanou, and S. Kokolakis, "Continuous authentication with feature-level fusion of touch gestures and keystroke dynamics to solve security and usability issues," Comput. Secur., vol. 132, art. 103363, 2023.

[24] I. Stylios, S. Kokolakis, O. Thanou, and S. Chatzis, "Behavioral biometrics & continuous user authentication on mobile devices: A survey," Inf. Fusion, vol. 66, pp. 76–99, 2021.

[25] I. Stylios, S. Kokolakis, O. Thanou, and S. Chatzis, "Key factors driving the adoption of behavioral biometrics and continuous authentication technology: Empirical research," Inf. Comput. Secur., vol. 30, no. 4, pp. 562–582, 2022.

[26] Z. Syed, J. Helmick, S. Banerjee, and B. Cukic, "Touch gesture-based authentication on mobile devices: The effects of user posture, device size, configuration, and inter-session variability," J. Syst. Softw., vol. 149, pp. 158–173, 2019.

[27] P. S. Teh, N. Zhang, A. B. J. Teoh, and K. Chen, "A survey on touch dynamics authentication in mobile devices," Comput. Secur., vol. 59, pp. 210–235, 2016.

[28] H. C. Volaka, G. Alptekin, O. E. Basar, M. Isbilen, and O. D. Incel, "Towards continuous authentication on mobile phones using deep learning models," Procedia Comput. Sci., vol. 155, pp. 177–184, 2019.

[29] Y. Yang, B. Guo, Z. Wang, M. Li, Z. Yu, and X. Zhou, "BehaveSense: Continuous authentication for security-sensitive mobile apps using behavioral biometrics," Ad Hoc Netw., vol. 84, pp. 9–18, 2019.

[30] M. Yao, D. Tao, R. Gao, J. Wang, S. Helal, and S. Mao, "PressAFE: Barometer-based on-screen pressure assisted implicit authentication for smartphones," IEEE Internet of Things J., vol. 10, no. 1, pp. 1–17, 2023.

[31] T. Zhu, Z. Weng, G. Chen, and L. Fu, "A hybrid deep learning system for real-world mobile user authentication using motion sensors," Sensors, vol. 20, no. 14, art. 3876, 2020.